



Universidade Federal de Santa Catarina
Centro Tecnológico
Programa de Pós-Graduação em Engenharia de Produção

O método de Equação Integral com Quadratura Gaussiana
para otimizar os parâmetros do gráfico de controle
multivariado de Somas Acumuladas

Custodio da Cunha Alves

Tese de Doutorado

Orientador
Robert Wayne Samohyl, Ph.D.

Florianópolis
2009

Universidade Federal de Santa Catarina
Centro Tecnológico
Programa de Pós-Graduação em Engenharia de Produção

**O método de Equação Integral com Quadratura Gaussiana
para otimizar os parâmetros do gráfico de controle
multivariado de Somas Acumuladas**

Custodio da Cunha Alves

Tese apresentada à
Universidade Federal de Santa Catarina
para obtenção do título de
Doutor em Engenharia de Produção

Orientador:
Robert Wayne Samohyl, Ph.D.

Florianópolis
2009

Catálogo na fonte pela Biblioteca Universitária da
Universidade Federal de Santa Catarina

A474 Alves, Custodio da Cunha
O método de Equação Integral com Quadratura Gaussiana
para otimizar os parâmetros do gráfico de controle multivariado
de Somas Acumuladas [tese] / Custodio da Cunha Alves ;
orientador, Robert Wayne Samohyl - Florianópolis, SC, 2009.
196f.: il., tabs., grafs.

Tese (doutorado) - Universidade Federal de Santa Catarina,
Centro Tecnológico. Programa de Pós-Graduação em
Engenharia de Produção.

Inclui bibliografia

1. Engenharia de produção. 2. Matemática para engenharia.
3. Pesquisa operacional. 4. Estatística - Controle de qualidade.
5. Planejamento experimental - Estatística. I. Samohyl,
Robert Wayne. II. Universidade Federal de Santa Catarina.
Programa de Pós-Graduação em Engenharia de Produção.
III. Título.

CDU 658.5

Custodio da Cunha Alves

**O método de Equação Integral com Quadratura Gaussiana
para otimizar os parâmetros do gráfico de controle
multivariado de Somas Acumuladas**

Esta tese foi julgada e aprovada para a obtenção do título de **Doutor em Engenharia de Produção** no Programa de Pós Graduação em Engenharia de Produção da Universidade Federal de Santa Catarina

Florianópolis, 06 de março de 2009

Prof. Antônio Sérgio Coelho, Dr
Coordenador do Programa de Pós-Graduação em Engenharia de Produção

BANCA EXAMINADORA

Prof. Adriano Mendonça Souza, Dr
Universidade Federal de Santa Maria
Examinador Externo

Prof. Milton Procópio de Borba, Dr
Universidade do Estado de SC
Examinador Externo

Prof. Pedro Alberto Barbeta, Dr
Universidade Federal de SC
Examinador/Moderador

Prof. Marcelo Menezes Reis, Dr
Universidade Federal de SC
Examinador

Prof^a Vera do Carmo C. de Vargas, Dr^a
Universidade Federal de SC
Examinadora

Prof. Robert Wayne Samohyl, Ph.D.
Universidade Federal de SC
Orientador

Minhas homenagens:

A mãe de meus filhos, Cecília,
a quem amo, e que se faz
sempre amiga e companheira

Aos meus filhos
Cássio, Cassiano e Cláudio
que me dão alegria e amor

Aos meus pais (em memória) que
sempre me ajudaram e apoiaram ao longo
ao longo de toda minha vida e a quem
sempre estarei em dívida

Ao meu irmão Dario (em memória) por
ter despertado em mim o interesse
por Ciências Exatas

As minhas cinco irmãs que a
cada dia escrevem uma nova página
na história de nossa família

AGRADECIMENTOS

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

Um agradecimento especial ao meu orientador professor Robert Wayne Samohyl pela amizade, atenção, incentivo e, principalmente, paciência e dedicação na elaboração deste trabalho.

Agradeço ainda:

aos meus familiares, o apoio constante;

ao meu filho primogênito, Cássio Alves, pela valiosa contribuição no desenvolvimento dos programas computacionais propostos neste trabalho;

aos professores e funcionários do PPGEP/UFSC com quem tive a oportunidade de conviver e aprender;

aos meus colegas Manoel, Éder, Andréa, Leandro, Gueibe e Rodrigo, pela amizade e companheirismo;

à colega Elisa, pela amizade e palavras de incentivo;

ao colega Ary, pela enorme colaboração no fornecimento dos dados reais do processo de usinagem utilizado em sua dissertação de mestrado;

aos professores Adriano Souza, Vera de Vargas, Pedro Barbeta, Marcelo Reis e Milton de Borba, Robert Samohyl, Rubson Rocha e Evandro da Silva as sugestões propostas para a melhoria deste trabalho por ocasião tanto do Exame de Qualificação quanto da Defesa de Doutorado;

aos professores, funcionários e acadêmicos da UNIVILLE e SENAI onde leciono que me encorajaram ao longo de todo o doutorado. Não me refiro ninguém em especial para não cometer a injustiça de esquecer alguém;

à Universidade Região de Joinville, o incentivo moral e a ajuda financeira e, em especial, à professora Sandra Furlan, Pró-Reitora de Pesquisa e Pós-Graduação, a quem sou muito grato por sua política de valorização do professor, responsável pelo meu crescimento pessoal e profissional;

A todas as pessoas que não foram nominalmente mencionadas, mas que de alguma forma contribuíram para viabilizar este trabalho.

BIOGRAFIA DO AUTOR

Custodio da Cunha Alves é graduado em Matemática e possui especialização em Matemática Aplicada pela Universidade da Região de Joinville (UNIVILLE), em 1990 e 1995, respectivamente. É mestre em Engenharia de Produção pela Universidade Federal de Santa Catarina (UFSC), em 2003. O autor atua como professor de ensino de Matemática e Estatística desde 1980. Atua, também como instrutor de ensino na área de Estatística Industrial. É consultor ad hoc da Fundação de Ciência e Tecnologia do Estado de Santa Catarina (FUNCITEC). Trabalhou na Empresa Brasileira de Compressores S.A. (EMBRACO), em Joinville-SC, de 1989 a 1998, na área de Controle da Qualidade. Atualmente, é professor de ensino do Departamento de Engenharia de Produção Mecânica da UNIVILLE e SENAI onde leciona as disciplinas de Cálculo Diferencial e Integral, Álgebra Linear e Geometria Analítica, Estatística, Pesquisa Operacional e Controle Estatístico de Processos. Entre seus principais interesses por trabalhos acadêmicos incluem a pesquisa e o ensino de matemática e estatística na engenharia, tais como, técnicas de controle estatístico da qualidade, planejamento de experimentos e aplicação da metodologia de pesquisa operacional a problemas de sistemas de manufatura. É, também, autor e co-autor de artigos técnicos nessas áreas publicados em anais de grandes eventos nacionais, tais como, Encontro Nacional de Engenharia de Produção (Enegep), Simpósio Nacional de Probabilidade e Estatística (Sinape) e Simpósio de Engenharia de Produção (Simpep). Atualmente, o autor continua estudando técnicas de controle estatístico com enfoque em aplicações eficientes de projetos de gráficos de controle univariados e multivariados de Somas Acumuladas (CUSUM e MCUSUM) no monitoramento de processos industriais.

RESUMO

ALVES, C.C. **O método de Equação Integral com Quadratura Gaussiana para otimizar os parâmetros do gráfico de controle multivariado de Somas Acumuladas.** Tese de Doutorado. Programa de Pos-Graduação em Engenharia de Produção e Sistemas. Universidade Federal de Santa Catarina, 2009.

O rápido crescimento das tecnologias de aquisição de dados e a utilização de computadores têm incrementado o interesse pelas técnicas de controle estatístico emergentes, tais como os gráficos de controle multivariados aplicados no monitoramento de processos que envolvem o controle simultâneo de características da qualidade ou variáveis de processo correlacionadas. Neste trabalho propõe-se o método de Equação Integral com Quadratura Gaussiana para otimizar os principais parâmetros ARL, k e h do gráfico de controle multivariado de Somas Acumuladas (MCUSUM). Este método envolve a derivação analítica de uma equação integral, cuja resolução numérica via Quadratura Gaussiana possibilita ao usuário obter a solução aproximada desses parâmetros utilizando um programa computacional desenvolvido a partir deste método. Além disso, propõe-se também, a partir desta metodologia, uma aproximação sistemática de projetar estatisticamente um gráfico de controle MCUSUM ótimo baseado nestes parâmetros, que minimize o custo médio de operação. Isso, orienta o usuário a decidir que tamanho de mudança do vetor de médias é realmente importante detectar, e tendo isto como base, selecionar aqueles parâmetros ótimos do gráfico MCUSUM que contemplem aspectos estatísticos como a minimização do número de falsos alarmes e a maximização da capacidade de detecção de mudanças reais. Os resultados obtidos deste trabalho revelam o método de Equação Integral proposto como uma excelente alternativa para otimizar os parâmetros do gráfico MCUSUM por ser um procedimento mais versátil que fornece melhores resultados para o valor de ARL com maior rapidez de cálculo comparado com o método de simulações e a relativa simplicidade de implementação em um programa computacional.

Palavras-chave: Otimização de parâmetros, equação integral, métodos numéricos, gráfico de controle multivariado de somas acumuladas.

ABSTRACT

ALVES, C.C. **The Method of Integral Equation with Gaussian Quadrature to optimize the parameters of the multivariate cumulative sum control chart. Doctorate Thesis.** Programa de Pos-Graduação em Engenharia de Produção e Sistemas. Universidade Federal de Santa Catarina, 2009.

The fast growth of data technologies acquisition and the use of computers have increased the interest in techniques of statistical emergent control, such as applied multivariate control charts for monitoring processes that involve the simultaneous control of the quality characteristics or process variables correlated. In this work the method of Integral Equation with Gaussian Quadrature to optimize the main parameters ARL, k and h of the multivariate cumulative sum control chart (MCUSUM) is proposed. This method involves the derivation of an analytical integral equation, whose numerical resolution way Gaussian Quadrature allows the user to obtain the approximate solution of these parameters using a computer program developed from this method. Furthermore, from this methodology, a systematic approach to design a statistical chart of MCUSUM optimal control based on these parameters, which minimizes the average cost of operation, is also proposed. This, guides the user to decide what size to change the vector of means is important to detect, and taking this as a base, selecting those parameters MCUSUM great chart that address statistical issues such as minimizing the number of false alarms and maximizing the ability of detection of real changes. The results of this study reveal the method of Integral Equation proposed as an excellent alternative to optimize the parameters of the chart MCUSUM to be a more versatile procedure that provides better results for the value of ARL with higher speed of calculation compared with the method of simulations and the relative simplicity of implementation in a computer program.

Keywords: Optimization of parameters, integral equation, numerical methods, multivariate cumulative sum control chart.

LISTA DE FIGURAS

Figura 1.1	Fases da metodologia utilizada no trabalho.....	23
Figura 2.1	Região de controle para duas variáveis independentes analisadas de forma conjunta adaptada de Montgomery 2004).....	28
Figura 2.2	Região de controle para duas variáveis correlacionadas analisadas de forma conjunta adaptada de Montgomery (2004).....	29
Figura 2.3	Forma bidimensional de pirâmide, cone e parabolóide.....	45
Figura 2.4	Valores de ARL dos gráficos MCUSUM e T^2 de Hotelling para $ARL_0=200$	60
Figura 4.1	Curvas de ARL do gráfico MCUSUM para $n=1$ e $ARL_0=200, 500$ e 1000	77
Figura 4.2	Desempenho de ARL dos gráficos MCUSUM, MEWMA e T^2 de Hotelling.....	78
Figura 4.3	Desempenho de ARL do gráfico MCUSUM com aplicação de SIM, MCM e MEI.....	79
Figura 4.4	Resumo das etapas para desenvolvimento do projeto estatístico MCUSUM.....	81
Figura 4.5	Regiões de Máxima e mínima potência de um gráfico de controle.....	82
Figura 4.6	Função perda para as características da qualidade do exemplo de aplicação.....	83
Figura 4.7	Regiões de Máxima e mínima potência do MCUSUM do exemplo de aplicação.....	86
Figura 4.8	Caixa de diálogo com os dados de entrada do exemplo de aplicação.....	87
Figura 4.9	Esboço dos furos 1 e 2 do bloco de motor para veículo de passeio.....	89
Figura 4.10	Fluxo resumido do processo de usinagem do bloco de motor.....	90
Figura 4.11	Histograma das características da qualidade X_1 e Y_1 (Processo XY).....	93
Figura 4.12	Gráfico de probabilidade qui-quadrado (Q-Q plot) para os dados do processo XY.....	94
Figura 4.13	Correlogramas das características da qualidade X_1 e Y_1 (Processo XY).....	95
Figura 4.14	Função Autocorrelação (FAC) das variáveis X_1 com Y_1 (Processo XY).....	96
Figura 4.15	Análise de Variância (ANOVA) e Estatística Descritiva para as variáveis X_1 e Y_1	96
Figura 4.16	Modelo matemático para otimizar os pontos de máxima e mínima potência.....	98
Figura 4.17	Pontos de máxima e mínima potência para otimizar o gráfico MCUSUM.....	99
Figura 4.18	Tela do R para carregar a rotina do gráfico MCUSUM.....	100

Figura 4.19	Tela do R para executar a rotina do gráfico MCUSUM.....	101
Figura 4.20	Gráfico MCUSUM do processo XY (furação de acabamento do furo 1).....	101
Figura 4.21	Histogramas das características da qualidade X_1 , Y_1 e D_{12} do Processo XYD.....	103
Figura 4.22	Gráfico de Probabilidade Normal das variáveis X_1 , Y_1 e D_{12} do Processo XYD.....	104
Figura 4.23	Gráfico de probabilidade qui-quadrado (Q-Q plot) para os dados do Processo XYD.....	105
Figura 4.24	Correlogramas das características da qualidade X_1 , Y_1 e D_{12} do Processo XYD.....	106
Figura 4.25	Função Autocorrelação (FAC) para as variáveis X_1 com D_{12} e Y_1 e D_{12}	107
Figura 4.26	Análise de Variância (ANOVA) e Estatística Descritiva para as variáveis X_1 e D_{12}	108
Figura 4.27	Análise de Variância (ANOVA) e Estatística Descritiva para as variáveis as variáveis Y_1 e D_{12}	108
Figura 4.28	Gráfico de controle MCUSUM para os dados do processo XYD.....	110

LISTA DE TABELAS

Tabela 4.1	Tabela para a implementação do algoritmo de soma acumulada O parâmetro ótimo do gráfico, k para projetar o MCUSUM e os correspondentes limites de controle, h com o mínimo ($ARL_{min.}$) para $p=2$, ARL sob controle (ARL_o) de 200, 500 e 1000, e vários tamanhos de mudança, d.....	73
Tabela 4.2	Tabela para a implementação do algoritmo de soma acumulada O parâmetro ótimo do gráfico, k para projetar o MCUSUM e os correspondentes limites de controle, h com o mínimo ($ARL_{min.}$) para $p=3$, ARL sob controle (ARL_o) de 200, 500 e 1000, e vários tamanhos de mudança, d	74
Tabela 4.3	Tabela para a implementação do algoritmo de soma acumulada O parâmetro ótimo do gráfico, k para projetar o MCUSUM e os correspondentes limites de controle, h com o mínimo ($ARL_{min.}$) para $p=4$, ARL sob controle (ARL_o) de 200, 500 e 1000, e vários tamanhos de mudança, d.....	74
Tabela 4.4	Comparativo das propriedades estatísticas do MCUSUM via MCM x MEI.....	75
Tabela 4.5	Análise de sensibilidade: Valores de k e h para o ARL de 200 quando o processo está sob controle e os valores de ARL quando o processo está fora de controle com $p=2$ e $n=1$	76
Tabela 4.6	Desempenho de ARL entre gráficos de controle multivariado.	110

LISTA DE SIGLAS E SÍMBOLOS

RL	número de amostras coletadas até à emissão de um sinal (<i>Run Length</i>)
ARL	número médio de amostras coletadas até à emissão de um sinal (<i>Average Run Length</i>)
ARL _o	número médio de amostras coletadas até a emissão de um sinal durante o período sob controle (ARL para o desvio nulo)
ARL _d	número médio de amostras coletadas até a emissão de um sinal que indique uma situação de fora de controle (ARL para o desvio d)
MRL	número mediano de amostras coletadas até à emissão de um sinal
SPRT	Teste Sequencial da Razão de Probabilidades (<i>Sequential Probability Ratio Test</i>)
FIR	Resposta Inicial Rápida (<i>Fast Initial Response</i>)
MEI	Método de Equação Integral
MCM	Método de Cadeias de Markov
MCU	Múltiplos gráficos CUSUM univariados
SIM	Método de Simulação
CEP	Controle Estatístico de Processos
CUSUM	Soma acumulada (<i>Cumulative Sum</i>)
FAC	Função Autocorrelação
ACP	Análise de Componentes Principais
f(x)	função distribuição de probabilidade da variável aleatória X
H _o	hipótese nula
H ₁	hipótese alternativa
k	valor de referência do gráfico de controle
h	limite de controle ou intervalo de decisão padronizado
N(0,1)	distribuição normal padronizada com média zero e desvio padrão unitário
α	probabilidade de um erro tipo I
β	probabilidade de um erro tipo II
d	tamanho da mudança em unidades de desvio padrão
μ	vetor de médias
μ_o	vetor de médias para a situação sob controle
μ_1	vetor de médias para uma situação de fora de controle

SUMÁRIO

CAPÍTULO 1 INTRODUÇÃO.....	18
1.1 Motivação para aplicação de Gráficos de Controle Multivariados.....	20
1.2 Tema e Justificativa.....	20
1.3 Objetivos.....	21
1.4 Método de Trabalho.....	22
1.5 Delimitações da Metodologia.....	24
1.5 Estrutura da Tese.....	24
CAPÍTULO 2 REVISÃO DE LITERATURA.....	26
2.1 Gráficos de Controle Estatístico de Processos Multivariados.....	26
2.2 Gráfico de Controle T^2 de Hotelling (Shewhart Multivariado).....	30
2.2.1 Subgrupos racionais	30
2.2.2 Observações individuais multivariadas.....	32
2.3 Gráfico de Controle Multivariado MEWMA.....	35
2.4 Gráfico de Controle Multivariado MCUSUM.....	37
2.4.1 Gráfico de Controle CUSUM.....	37
2.4.2 Gráficos de controle multivariado MCUSUM.....	38
2.5 O ARL - Número Médio de Amostras até a emissão de um sinal.....	45
2.5.1 Método de Equação Integral.....	47
2.5.2 Método das Cadeias de Markov.....	49
2.5.3 Método de Simulação.....	50
2.6 Parâmetros dos Gráficos de Controle CUSUM e MCUSUM.....	50
2.7 Projeto ótimo aplicado para obter os parâmetros do gráfico MCUSUM.....	55
2.8 Método para otimização dos parâmetros do gráfico MCUSUM.....	58
2.8.1 Otimização do MCUSUM baseado em Regiões de Máxima e Mínima Potência.....	58
2.9 Integração Numérica.....	61
2.9.1 Quadratura Gaussiana.....	62
2.10 Síntese do Capítulo.....	64

CAPÍTULO 3 METODOLOGIA	65
3.1 Introdução.....	65
3.2 Método de Equação Integral com Quadratura Gaussiana para otimização dos parâmetros do gráfico MCUSUM.....	65
3.3 Síntese do Capítulo.....	68
CAPÍTULO 4 APLICAÇÃO DA METODOLOGIA.....	70
4.1 Introdução.....	70
4.2 Desenvolvimento Experimental da Metodologia.	71
4.3 Programas Implementados.	71
4.4 Aplicação do MEI para determinar os parâmetros do gráfico de controle MCUSUM.....	72
4.4.1 Resultados da aplicação do MEI para os parâmetros do gráfico MCUSUM.....	73
4.5 Projeto estatístico ótimo para um gráfico MCUSUM baseado em ARL e sob a ótica de regiões de máxima e mínima potência.....	79
4.5.1 Exemplo de aplicação para o projeto estatístico ótimo do gráfico MCUSUM baseado em ARL e sob a ótica de regiões de máxima e mínima potência.....	82
4.6 Aplicação da metodologia proposta utilizando dados reais de um processo de usinagem... ..	88
4.6.1 Apresentação da empresa.....	88
4.6.2 Descrição do produto selecionado para análise estatística multivariada do processo de usinagem.....	89
4.6.3 Descrição do processo selecionado para o monitoramento de características da qualidade via gráfico MCUSUM.....	90
4.6.4 Exploração dos dados do processo para aplicação da metodologia proposta.....	91
4.7 Análise estatística multivariada dos dados: Processo XY.....	92
4.7.1 Normalidade.....	92
4.7.2 Autocorrelação.....	94
4.7.3 Otimização dos pontos que delimitam as regiões de máxima e mínima potência do gráfico MCUSUM para o processo XY.....	97

4.7.4	Desenvolvimento do gráfico MCUSUM projetado estatisticamente para o processo XY.....	99
4.8	Análise estatística multivariada dos dados: Processo XYD.....	102
4.8.1	Normalidade.....	102
4.8.2	Autocorrelação.....	105
4.8.3	Projeto ótimo do gráfico MCUSUM em um ponto.....	109
4.9	Síntese do Capítulo.....	111
CAPÍTULO 5 CONCLUSÕES.....		112
5.1	Desenvolvimento e Aplicação do Método de Equação Integral.....	112
5.2	Projeto Estatístico Ótimo do Gráfico de Controle MCUSUM.....	113
5.3	Projeto Otimização do MCUSUM com Regiões de Máxima e mínima.....	113
5.4	Recursos Computacionais Aplicados e Parâmetros Ótimos obtidos através do MEL.....	115
5.5	Aplicação da Metodologia e Resultados obtidos com a Utilização de Dados Reais.....	116
5.6	Recomendações para Investigações Futuras.....	117
REFERÊNCIAS BIBLIOGRÁFICAS.....		119
APÊNDICE A Integração Gaussiana por Polinômios de Legendre.....		125
APÊNDICE B Teoremas e Demonstrações: Equação Integral para o ARL sob controle.....		131
APÊNDICE C Distribuição Normal Multivariada.....		137
APÊNDICE D Autocorrelação e Estabilidade Estatística.....		144
APÊNDICE E Teste Seqüencial da Razão de Probabilidades e Teoria da Soma Acumulada em Gráficos de controle.....		147
APÊNDICE F A Função Perda de Taguchi para a Tomada de Decisões.....		154
APÊNDICE G Método da Secante.....		159
APÊNDICE H Interpolação Polinomial.....		161
APÊNDICE I Relatórios: resultados obtidos com a resolução de modelos matemáticos.....		163

APÊNDICE J	O R Project e Rotina Desenvolvida para gerar o Gráfico MCUSUM.....	167
APÊNDICE K	Programas Computacionais Desenvolvidos em Ambiente Matlab.....	170
APÊNDICE L	Código Computacional dos Programas em Ambiente Matlab.....	184
ANEXO A	Dados do processo furação de acabamento do furo1 para as características da qualidade X_1 , Y_1 e D_{12} (medidas em mm).....	193
ANEXO B	Parâmetros Ótimos do Gráfico MCUSUM via Método Cadeias de Markov (Tabela 1).....	194
ANEXO C	Parâmetros Ótimos do Gráfico MCUSUM via Método Cadeias de Markov (Tabela 2).....	195
ANEXO D	Parâmetros Ótimos do Gráfico MCUSUM via Método Cadeias de Markov (Tabela 3).....	196

CAPÍTULO 1

INTRODUÇÃO

A indústria atual se encontra diante de uma economia globalizada que estabelece desafios cada vez mais exigentes. Hoje em dia, nosso competidor pode estar em qualquer lugar do planeta e a sobrevivência do setor industrial está baseada na melhoria contínua de seus processos e produtos. Diante deste quadro de extrema competência, a utilização de ferramentas estatísticas cada vez mais específicas direcionadas ao aperfeiçoamento da qualidade tem sido uma prática na maior parte das organizações para elevar seus níveis de qualidade a um baixo custo, permitindo, a longo prazo, a sobrevivência e a otimização de cotas de mercado.

No século XIX, os esforços para eliminar as variações inerentes ao processo eram muitas vezes bem sucedidos devido à simplicidade de seus produtos manufaturados. Atualmente, estamos atravessando uma terceira revolução industrial, a revolução da informação, em que dados estão cada vez mais abundantes e acessíveis. Isso, demanda uma maior complexidade dos sistemas de fabricação e montagem que exige uma maior atenção no monitoramento das características da qualidade para desenvolver um produto cada vez melhor em torno do valor nominal e dentro das normas técnicas exigidas.

O Controle Estatístico de Processos (CEP) e em particular as técnicas de Controle de Qualidade, tais como gráficos de controle, têm sido cada vez mais importantes pelo fato de desempenharem papel primordial na indústria moderna. Nas últimas décadas, mudanças consideráveis no cenário econômico mundial tem sido acompanhada por um crescimento significativo de aplicações de técnicas de controle da qualidade emergentes tais como o controle estatístico de processos multivariados. A aplicação dessas técnicas estatísticas tais como o gráfico de controle multivariado de Somas Acumuladas (MCUSUM), objeto de estudo desse trabalho, tem como alvo proporcionar o monitoramento efetivo para assegurar a qualidade de processos e produtos industriais.

O aumento da utilização de técnicas de controle estatístico multivariado notadamente nos últimos anos em razão dos inúmeros recursos da informatização hoje existentes tem sido fundamental para atender a complexidade dos atuais processos industriais. Estes recursos em procedimentos de controle industriais têm melhorado significativamente a qualidade da produção na indústria moderna. Na indústria de manufatura, por exemplo, procedimentos automatizados de inspeção nos permitem obter medições, em cada peça fabricada, de variáveis

correlacionadas num mesmo processo tais como a dimensão, peso, etc. Na prática, este cenário requer a necessidade do monitoramento simultâneo e on-line de mais de uma característica da qualidade via gráficos de controle multivariados, uma vez que os recursos dos sistemas de informação atuais facilitam o compilamento e o armazenamento de dados utilizados na maioria de nossos processos industriais que infelizmente são de natureza multivariada.

No Brasil, a utilização de gráficos de controle com metodologia para processos de multivariáveis não tem sido comum no controle de qualidade industrial devido às dificuldades inerentes às técnicas multivariadas. Felizmente, com o avanço da informática nas últimas décadas, programas de computadores aplicados aos processos de produção têm sido desenvolvidos para atender as tendências mais modernas de métodos de controle estatísticos multivariados que possam auxiliar tais processos na busca pela qualidade.

No monitoramento de processos industriais multivariados, muitas vezes se faz necessário detectar pequenas e moderadas mudanças das variáveis correlacionadas, situação onde são recomendados os gráficos de controle com memória, tais como o gráfico multivariado de Somas Acumuladas (MCUSUM) e o gráfico de Média Móvel Exponencialmente Ponderada (MEWMA). Estes gráficos acumulam a informação mais recente com informações anteriores e, com isso, detectam pequenas mudanças nos parâmetros de um processo com um número médio de amostras coletadas até à emissão de um sinal (ARL) menor do que fariam os tradicionais gráficos de controle T^2 de Hotelling.

O gráfico de controle MCUSUM, focalizado neste trabalho, é considerado uma ferramenta estatística versátil uma vez que seu processo de decisão baseia-se nas somas acumuladas dos resultados e não em observações isoladas de amostras. Esta característica torna este gráfico mais sensível para detectar mais rapidamente pequenas trocas no vetor de médias do processo que implica em menores valores para o ARL. A detecção rápida destas pequenas alterações nos parâmetros de um processo é uma tarefa realmente importante numa perspectiva de aumentar a consistência da qualidade de processos e produtos. Diante disso, a otimização de alguns parâmetros que tornam esse gráfico mais robusto em termos de desempenho de ARL é fundamental para reduzir de forma significativa a probabilidade de falsos alarmes.

Este trabalho propõe um projeto estatístico ótimo para o gráfico de controle MCUSUM baseado na otimização sistemática de ARL com a melhor combinação dos parâmetros valor de referência k com o limite superior de controle h (intervalo de decisão para assegurar o desempenho deste gráfico. Para determinar o ARL uma equação integral é analiticamente derivada. A solução dessa equação é obtida através do método de Equação

Integral com Quadratura Gaussiana. O procedimento deste projeto ótimo é uma extensão multivariada do gráfico de controle CUSUM univariado (Gan,1991,1993) que utiliza o método de Equação Integral para determinar os parâmetros ótimos desse gráfico. Além disso, realizar um estudo comparativo da otimização do gráfico MCUSUM em termos de desempenho de ARL via Cadeias de Markov (Lee e Khoo, 2006) com o Método de Equação Integral proposto neste trabalho nas mesmas condições e restrito a observações individuais com $p = 2, 3$ e 4 variáveis (características da qualidade) e para o ARL_o de 200, 500 e 1000.

1.1 Motivação para aplicação de Gráficos de Controle Multivariados

Os gráficos de controle amplamente aplicados ao monitoramento de processos são ainda, sem dúvida, os tradicionais gráficos que utilizam técnicas univariadas de controle estatístico de processos. Apesar de eficazes, estas ferramentas não são indicadas em situações onde se faz necessário o monitoramento simultâneo de duas ou mais características da qualidade (ou variáveis) correlacionadas. Para estas situações é fundamental a utilização de gráficos de controle multivariados na análise estatística de processos.

Na implementação de um método para monitorar diversas variáveis, simultaneamente, procura-se uma técnica que preserve essencialmente os seguintes tópicos:

- Produzir uma única resposta para constatar que o processo está sob controle estatístico.
- Manter a probabilidade especificada (α) de considerar o processo fora de controle quando na verdade ele não está.
- Estabelecer o método que leve em consideração a interdependência entre as variáveis.

1.2 Tema e Justificativa

Este trabalho tem como tema principal a aplicação do método de Equação Integral com Quadratura Gaussiana cuja resolução numérica da equação integral derivada analiticamente fornece o ARL. Para selecionar a combinação ótima dos parâmetros k com h que fornece ARL mínimo do gráfico MCUSUM para a mudança do vetor de médias que é importante detectar utiliza-se programas computacionais em ambiente Matlab, algoritmos de métodos numéricos, tais como, a interpolação polinomial não linear para obter-se valores intermediários tabelados dos parâmetros desejados e o método iterativo da secante para aproximar a solução do limite superior de controle h . Isso, é fundamental para a análise de

sensibilidade preliminar desses parâmetros e imprescindível para projetar estatisticamente o gráfico de controle MCUSUM ótimo. Alguns pesquisadores como Crosier (1988), Pignatiello e Runger (1990), Lowry, Woodall e Champ (1992), Ngai e Zang (2001), etc, têm apresentado tabelas e métodos gráficos para selecionar parâmetros ótimos de gráficos multivariados tais como os valores de k , h para algumas magnitudes de mudanças. Em aplicações práticas estes recursos apresentam um número muito reduzido de situações sendo portanto, muito limitada a sua aplicação.

O tema proposto nesta tese tem sua relevância fundamentada na otimização do gráfico MCUSUM cujo estudo referente a este tema nas últimas décadas direcionado aos gráficos de controle multivariados T^2 de Hotelling e MEWMA. Pouca atenção tem sido direcionada ao gráfico de controle multivariado MCUSUM, objeto de estudo deste trabalho. Além disso, propor uma metodologia cuja seleção dos parâmetros ótimos do gráfico MCUSUM se leve em consideração aspectos estatísticos como a minimização do número de falsos alarmes (sinalização da ocorrência de uma causa especial quando na verdade não houve alteração do processo) e por conseqüência a maximização da capacidade de detecção de mudanças reais. Para isso, propõe-se uma metodologia que contemple a otimização sistemática de parâmetros levando em consideração as regiões de máxima e mínima potência do gráfico MCUSUM. Na bibliografia consultada só é encontrado a referência de Woodall (1985), que trata análise estatística de gráficos de controle segundo a ótica de regiões de máxima e mínima potência referidas ao caso univariado. No caso multivariado para o gráfico MCUSUM, a investigação neste sentido é nula.

1.3 Objetivos

O objetivo principal deste trabalho é desenvolver o método de Equação Integral com Quadratura Gaussiana cujos procedimentos metodológicos incluem as etapas relacionadas desde a derivação analítica da equação integral até sua aplicação para obter a solução aproximada de ARL que otimize o gráfico de controle MCUSUM quando o processo está sob controle. Com base nesta informação, são utilizados outros métodos numéricos adequados para estender o estudo deste parâmetro à situações quando o processo está fora de controle. Além disso, a implementação deste método em um ambiente industrial, cuja averiguação para validade prática é realizada a partir de dados reais resultante do monitoramento do vetor de médias de um processo de produção com características da qualidade correlacionadas.

Os objetivos específicos mais relevantes deste trabalho são os seguintes:

- a) revisar a literatura sobre gráficos de controle multivariado de Somas Acumuladas (MCUSUM), suas peculiaridades e aplicações;
- b) derivar analiticamente uma equação integral para otimizar os parâmetros fundamentados no desempenho de ARL do gráfico MCUSUM;
- c) desenvolver o método de Equação Integral com Quadratura Gaussiana para otimizar os parâmetros ARL, k e h essenciais para projetar estatisticamente o gráfico MCUSUM sob a ótica de regiões de máxima e mínima potência;
- d) aplicar um modelo estatístico que possibilite encontrar os valores ótimos dos parâmetros do gráfico MCUSUM segundo a filosofia de máxima e mínima potência deste gráfico que minimize o custo médio de operação;
- e) elaborar um algoritmo numérico computacional utilizando o método de Equação Integral com Quadratura Gaussiana sob a ótica de regiões de máxima e mínima potência para otimizar os parâmetros do projeto estatístico do gráfico MCUSUM em um ponto para qualquer magnitude de mudança do vetor de médias de um processo. Isso, orienta o usuário a detectar mudanças significativas para situações de variáveis correlacionadas e monitoradas simultaneamente, tamanho da amostra e ARL dados;
- f) realizar análise de sensibilidade para mostrar como varia a probabilidade do erro tipo I para diferentes combinações de (k,h) da região ótima do gráfico MCUSUM.

1.4 Método de Trabalho

O método de trabalho proposto para o desenvolvimento deste trabalho constitui-se na aplicação de técnicas úteis elaboradas a partir de um modelo matemático adequado ao controle estatístico de processos multivariados cujas restrições para a otimização do gráfico MCUSUM incluem além do ARL os parâmetros k e h , valor de referência e limite superior de controle, respectivamente. A metodologia utilizada na presente tese consiste na adoção de um referencial teórico inicial do tema proposto para testar a metodologia proposta no escopo do trabalho através de exemplos práticos da literatura e de uma aplicação com dados reais para comprovar sua validade prática. O desenvolvimento desta metodologia compreende duas fases bem diferenciadas. A fase de busca de informação e planejamento e a fase puramente experimental conforme figura 1.1

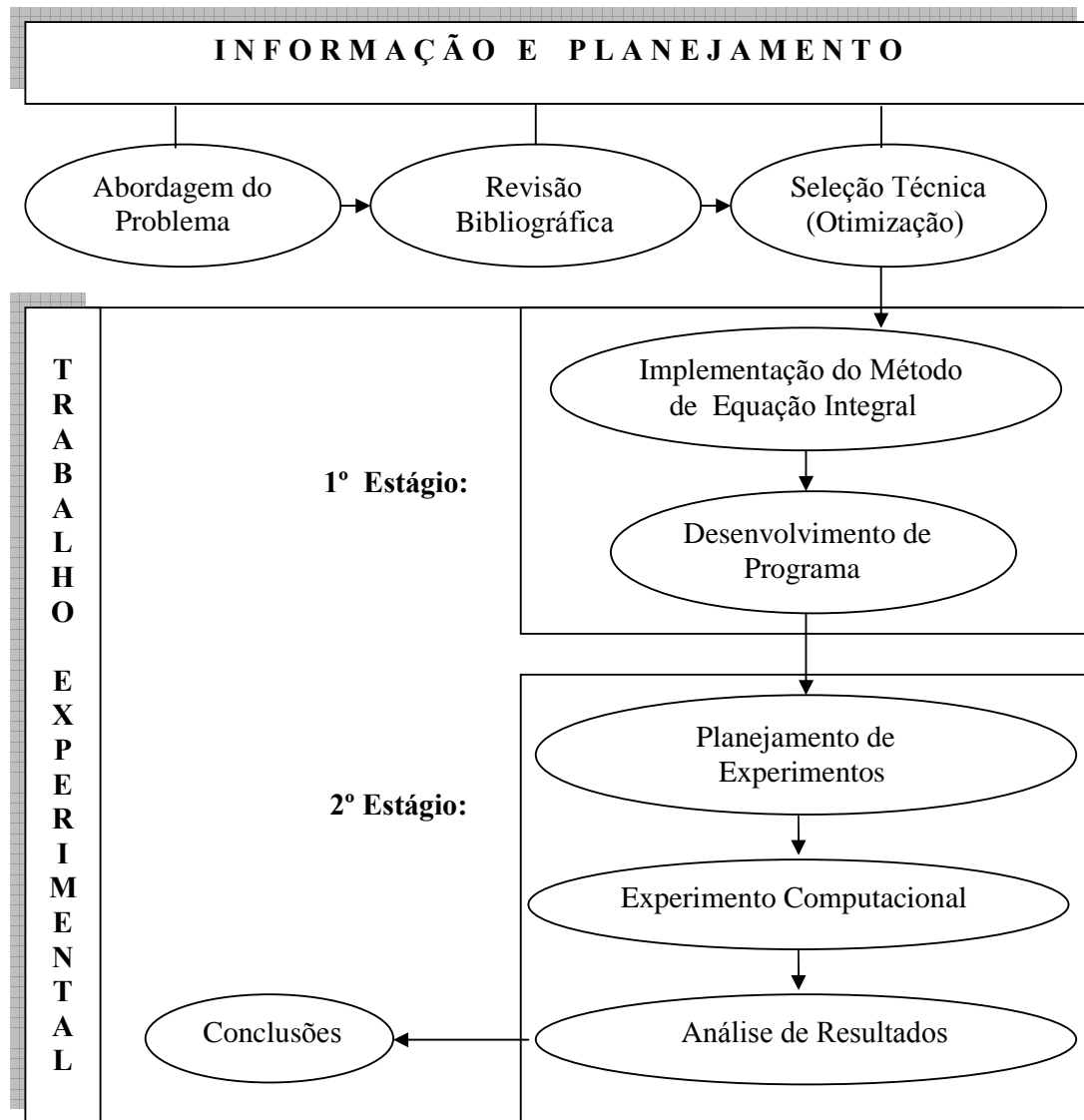


Figura 1.1 Fases da metodologia utilizada neste trabalho

A metodologia utilizada na segunda fase considerada como trabalho experimental inicia com a implementação do método de integração numérica que incluem a derivação analítica e a adequação da equação integral ao método proposto. Uma vez definida a equação integral para o método é desenvolvido um programa computacional (amigável) que nos permita otimizar os parâmetros do gráfico MCUSUM. Como o conjunto de possíveis combinações de parâmetros (k,h) é elevado propõe-se um planejamento de experimentos para poder encontrar em quais condições estes parâmetros otimizam o gráfico MCUSUM. O experimento computacional é realizado em computadores do Laboratório de Informática da Universidade da Região de Joinville. Uma análise estatística dos resultados obtidos do experimento computacional nos

permite analisar a influência das diferentes combinações dos parâmetros k com h que associado ao ARL interferem na otimização do gráfico MCUSUM. Por último, uma análise de sensibilidade da solução ótima desses parâmetros nos leva a extrair as conclusões finais da investigação.

1.5 Delimitações da Metodologia

As delimitações da metodologia proposta neste trabalho referem-se a otimização dos principais parâmetros ARL, k e h para projetar estatisticamente o gráfico de controle multivariado de Somas Acumuladas restrito a observações individuais devido a escassez de material bibliográfico referente a esse tipo gráfico para subgrupos racionais. Esta metodologia inclui um modelo matemático para o projeto do gráfico MCUSUM ótimo com as seguintes restrições:

- a) o limite superior de controle, h do gráfico MCUSUM para o ARL mínimo é determinado apenas para ARL_0 de 200, 500 e de 1000;
- b) o número p de variáveis (características da qualidade) é considerado apenas com $p = 2, 3$ e 4.
- c) a estatística MCUSUM de controle Y_i é demarcada num gráfico unilateral com intervalo de decisão $0 \leq Y_i \leq h$ onde 0 e h são respectivamente, os limites inferior e superior de controle.

1.6 Estrutura da Tese

O presente trabalho está estruturado em cinco capítulos, os quais estão brevemente comentados a seguir :

O capítulo 1 trata das preliminares e objetivos do trabalho que incluem a motivação com sua respectiva justificativa, necessária para o entendimento da situação em que se aplica o controle estatístico de processos multivariados e para estabelecer-se os objetivos do presente trabalho.

O capítulo 2 apresenta o referencial teórico que incluem a revisão de literatura sobre gráfico de controle multivariado de Somas Acumuladas (MCUSUM) e seus principais parâmetros. Além disso, apresenta os métodos numéricos utilizados para aproximar estes parâmetros destacando a aplicação do Método de Equação Integral com Quadratura Gaussiana, objeto de estudo deste trabalho.

O capítulo 3 expõe a metodologia utilizada para alcançar os objetivos identificados no capítulo 1 constante no desenvolvimento de modelos que posteriormente são otimizados.

O capítulo 4 trata do desenvolvimento e da utilização do programa computacional (amigável) para aplicar o método de Equação Integral com Quadratura Gaussiana proposto que otimize os parâmetros essenciais do gráfico MCUSUM em ambiente industrial.

No capítulo 5 encontram-se as conclusões, ou seja, os resultados mais relevantes obtidos nos capítulos anteriores e as recomendações para trabalhos futuros.

CAPÍTULO 2

REVISÃO DE LITERATURA

2.1 Gráficos de Controle Estatístico de Processos Multivariados

Nas últimas décadas diversas áreas do setor produtivo vêm experimentando acelerados avanços tecnológicos e, desta forma, exigindo ferramentas estatísticas cada vez mais específicas para monitorar e avaliar a atual complexidade de seus processos. Este rápido crescimento em tecnologias de aquisição de dados e o uso de computadores para o monitoramento on-line de processos possibilitam, em tempo real, e com menores custos, o monitoramento simultâneo de várias características da qualidade ou variáveis de processo correlacionadas. Diante desta realidade, tem-se verificado nos últimos anos um crescente interesse em desenvolver técnicas de controle de qualidade mais robustas, que incorporem todas estas múltiplas variáveis correlacionadas de uma única vez tais como os procedimentos de gráficos de controle estatístico de processos multivariados (Alves, Henning e Samohyl, 2008a).

O tradicional método de controle estatístico de processos onde apenas uma característica da qualidade (ou variável) é monitorada a partir da utilização de alguns gráficos de controle univariados disponíveis tais como Shewhart, CUSUM, EWMA, etc., é ainda o mais conhecido e amplamente aplicado no setor industrial devido a sua simplicidade e facilidade de operacionalização (Montgomery, 2004). No entanto, são muitas as situações no cenário industrial onde num mesmo processo é necessário o controle simultâneo e on-line de duas ou mais características da qualidade. Embora a aplicação de gráficos de controle univariados a cada variável individual seja uma solução possível pode levar a conclusões errôneas se não levarmos em conta a probabilidade conjunta de erro tipo I e a estrutura de correlação entre as variáveis consideradas (Ferrer, 2005).

Suponhamos por exemplo, que se deseja monitorar duas variáveis normalmente distribuídas aplicando a cada uma delas um gráfico \bar{X} de Shewhart com limites de controle situados segundo o habitual critério 3σ . Se as duas variáveis estão sob controle, a probabilidade de que uma dessas variáveis exceda os limites de controle (probabilidade de erro tipo I) é igual a 0,0027. No entanto, a probabilidade conjunta de ambas variáveis excederem seus limites de controle, simultaneamente quando ambas estão sob controle, é

$(0,0027).(0,0027)=0,00000729$, o que é consideravelmente menor que $0,0027$. Porém, para o controle conjunto das duas variáveis, supondo que são independentes, a probabilidade de que as duas variáveis sejam marcadas fora dos limites de controle quando o processo está realmente sob controle é $1-[(1-0,0027).(1-0,0027)]=0,0054$. Além disso, a probabilidade de que as duas variáveis sejam marcadas dentro dos limites de controle quando o processo está sob controle é $(0,9973).(0,9973)= 0,99460729$. Portanto, a utilização de dois gráficos \bar{X} independentes alteram o monitoramento simultâneo das duas variáveis, considerando que um erro tipo I e a probabilidade de um ponto ser representado corretamente não sejam iguais aos seus níveis anunciados para os gráficos de controle individuais. Essa distorção no procedimento de monitoramento do processo será maior à medida que aumenta o número de variáveis que estamos monitorando simultaneamente (Montgomery, 2004). Em geral, se há p variáveis (características da qualidade) estatisticamente independentes para um produto em particular, e se o gráfico \bar{X} com probabilidade de erro tipo I é mantido para cada uma, então a verdadeira probabilidade α (erro tipo I) para cada procedimento de controle conjunto, α' , é:

$$\alpha' = 1 - (1 - \alpha)^p \quad (2.1)$$

onde $(1 - \alpha)^p$ é a probabilidade de que todas as variáveis sejam representadas no gráfico simultaneamente dentro de seus limites de controle (Montgomery, 2004).

A equação (2.1) pode ser utilizada quando as variáveis são independentes para determinar os valores apropriados de α para cada gráfico de controle quando α' é fixado, e então obter convenientemente os correspondentes limites de controle.

Para um gráfico de controle multivariado com p características de qualidade, a probabilidade para que o gráfico indique uma situação sob controle no vetor de médias do processo μ_0 é $1 - \alpha$. No caso em que α seja escolhido de forma que $\frac{\alpha}{2p} = 0,00135$, então $1 - 0,0027p$ é igual a $1 - \alpha$. Quando as p características são independentes a diferença entre as duas probabilidades é praticamente zero. Portanto, para os gráficos de controle separados, a probabilidade de que cada uma das p médias assume um valor dentro de seus limites de controle é $(1 - 0,0027)^p$, muito próximo de $1 - 0,0027p$ para o gráfico multivariado. Por exemplo, suponhamos que se pretende controlar simultaneamente a média do processo para duas características da qualidade independentes \bar{X}_1 e \bar{X}_2 de uma distribuição normal bivariada. Para monitorar ambas as características da qualidade pode-se aplicar o gráfico \bar{X} de Shewhart cujos limites de controle são fixados segundo o valor de $Z_{\frac{\alpha}{2p}} = Z_{0,00135}$ para cada

característica. O processo, neste caso, estará sob controle somente se as médias amostrais \bar{X}_1 e \bar{X}_2 representadas no gráfico estiverem dentro de seus respectivos limites de controle. Isto equivale a representar o par de médias (\bar{X}_1, \bar{X}_2) dentro da região retangular de controle conforme figura 2.1.

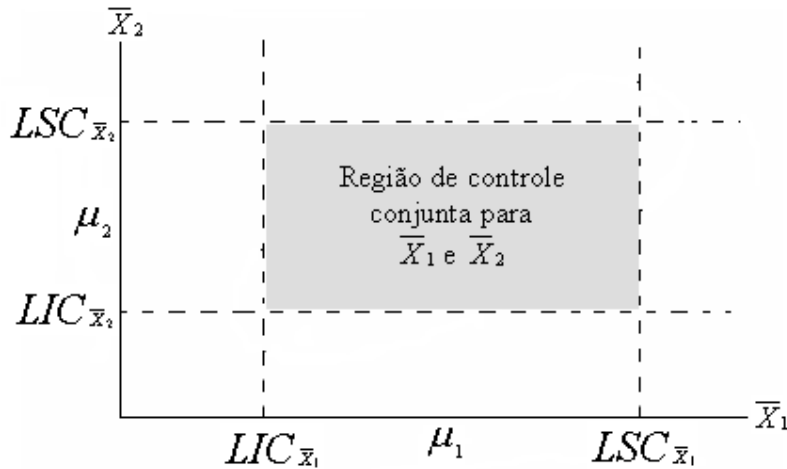


Figura 2.1 Região de controle para duas variáveis independentes analisadas de forma conjunta adaptada de Montgomery (2004)

A probabilidade que as $p = 2$ características da qualidade sob controle, assumem uma medida dentro de seus limites de controle é $(1 - 0,0027)^2 = 0,995$. Porém, quando a probabilidade conjunta de ambas as características da qualidade, não excedam seus limites de controle é $1 - [(0,0027) \cdot (2)] = 0,9946$. Isto comprova que quando as características da qualidade são independentes praticamente não existe diferença entre ambas as probabilidades (Montgomery, 2004).

No entanto, quando as características da qualidade não forem independentes, ou seja, correlacionadas que é o caso mais freqüente em aplicações industriais, isto representa uma diferença considerável. Neste caso, um procedimento mais complexo, deve ser empregado para determinar o valor apropriado α' e com isso obter corretamente os limites de controle. Por simplicidade, continuamos considerando o caso bivariado supondo que o gráfico \bar{X} de Shewhart é usado e que ambas variáveis (médias \bar{X}_1 e \bar{X}_2) seguem uma distribuição normal bivariada de correlação ρ com vetor de médias $\mu_0 = \mu$ e matriz de covariâncias $\Sigma_0 = \Sigma$ respectivamente conhecidos.

$$\mu_0' = (\mu_{0,1}, \mu_{0,2}) \quad \Sigma_0 = \begin{pmatrix} \sigma_{0,1}^2 & \rho\sigma_{0,1}\sigma_{0,2} \\ \rho\sigma_{0,1}\sigma_{0,2} & \sigma_{0,2}^2 \end{pmatrix}$$

Para isso, um valor deve ser calculado a partir da equação 2.1

$$\begin{aligned} \alpha' &= 1 - (1 - \alpha)^p \\ \alpha' &= 1 - P\left[-a < \frac{(\bar{X}_1 - \mu_{0,1})\sqrt{n}}{\sigma_{0,1}} < a \cap -a < \frac{(\bar{X}_2 - \mu_{0,2})\sqrt{n}}{\sigma_{0,2}} < a\right] \\ \alpha' &= 1 - P(-a < Z_1 < a \cap -a < Z_2 < a) \\ \alpha' &= 1 - \int_{-a}^a \int_{-a}^a f(z_1, z_2) dz_1 dz_2 \end{aligned} \tag{2.2}$$

onde n é o tamanho da amostra e $f(z_1, z_2)$ é a função densidade de probabilidade de uma distribuição normal padronizada bivariada com correlação ρ . A região de controle conjunta neste caso, para um determinado valor α' , é uma elipse cuja inclinação depende da correlação ρ existente entre as variáveis consideradas conforme figura 2.2.

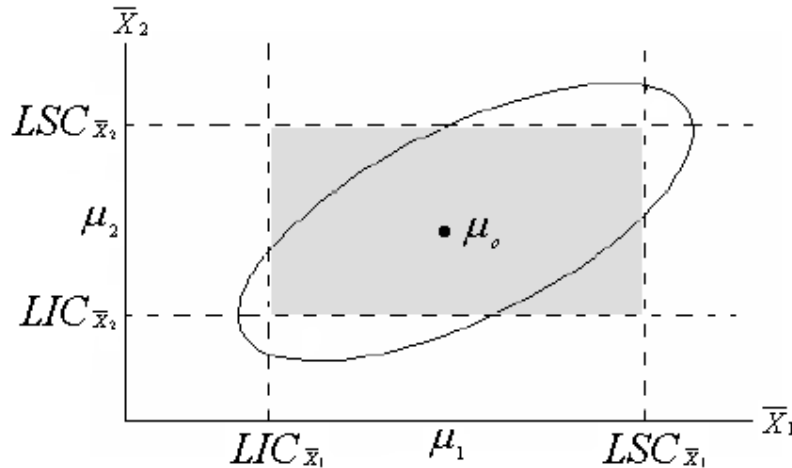


Figura 2.2 Região de controle para duas variáveis correlacionadas analisadas de forma conjunta adaptada de Montgomery (2004)

Os limites de controle fixados segundo o critério 3σ univariado para as variáveis \bar{X}_1 e \bar{X}_2 são sobrepostos na elipse para realizar a comparação. A área que representa o processo bivariado sob controle é limitada pela elipse. Como as duas variáveis \bar{X}_1 e \bar{X}_2 são dependentes qualquer ponto localizado fora da elipse de controle, indica a presença de uma causa atribuível, mesmo que esteja dentro dos limites de controle individuais para \bar{X}_1 e \bar{X}_2 .

Os gráficos de controle estatístico multivariados podem ser separados em dois grupos distintos: os esquemas de monitoramento direcionalmente específicos e os esquemas de monitoramento direcionalmente invariantes. No primeiro deles, a direção específica do

deslocamento do vetor de médias μ_o do processo é conhecida visto que no outro, esta direção é desconhecida (Pham, 2006).

Nesta seção, o objetivo maior foi justificar a pertinência da utilização de gráfico de controle estatístico multivariado para monitorar simultaneamente duas ou mais características da qualidade correlacionadas e medidas sobre o mesmo processo frente à utilização de gráficos de controle individuais para cada uma das variáveis separadamente.

Nas seções seguintes, é realizado um estudo pormenorizado dos três principais gráficos de controle atualmente utilizado no controle estatístico de processos multivariados o gráfico T^2 de Hotelling, o gráfico MEWMA e o gráfico MCUSUM.

2.2 Gráfico de Controle T^2 de Hotelling (Shewhart Multivariado)

O procedimento de controle multivariado mais conhecido e atualmente utilizado para monitorar o vetor de médias de um processo é ainda o gráfico T^2 de Hotelling. É um procedimento análogo ao gráfico de controle de Shewhart para o caso multivariado. Nesta seção são apresentadas as duas versões do gráfico T^2 de Hotelling: uma para dados em subgrupos racionais, e outra para dados individuais.

2.2.1 Subgrupos racionais

Os gráficos de controle multivariados são desenvolvidos a partir de medições dos resultados do processo para múltiplas variáveis. As medições são apresentadas, geralmente, para subgrupos de itens coletados de tal forma que estejam sob atuação de somente causas comuns. São chamados na literatura de subgrupos racionais.

O projeto de desenvolvimento de gráficos de controle T^2 de Hotelling para subgrupos racionais é separado em duas fases distintas (Lowry e Montgomery, 1995). A primeira fase (fase I) consiste da análise de dados passados com o objetivo de verificar se o processo estava sob controle estatístico quando os primeiros subgrupos foram obtidos, bem como retirar os pontos suspeitos de estarem fora dos limites de controle e estimar a distribuição do processo para teste de pontos futuros. A fase seguinte (fase II) consiste em usar o gráfico de controle para detectar qualquer fuga do processo em relação à sua distribuição estimada na fase 1.

Para a fase I considere que obtivemos m subgrupos independentes de tamanho n ($n > p$) provenientes de distribuições $N_p(\mu_i, \Sigma)$ $i = 1, 2, \dots, m$ em que os vetores de médias sejam desconhecidos e cujas matrizes de covariâncias são desconhecidas, mas iguais. Nesta primeira fase, precisamos estimar o vetor de médias μ e a matriz de covariâncias Σ a partir de uma

análise de m amostras de tamanho n , tomadas quando se admite que o processo está sob controle estatístico. Esses estimadores para $\boldsymbol{\mu}$ e $\boldsymbol{\Sigma}$, são dados respectivamente, por

$$\bar{\mathbf{x}} = \frac{1}{m} \sum_{i=1}^m \bar{\mathbf{X}}_i \quad (2.3)$$

$$\bar{\mathbf{S}} = \frac{1}{m} \sum_{i=1}^m \mathbf{S}_i \quad (2.4)$$

$$\mathbf{S}_m = \frac{1}{m-1} \sum_{i=1}^m (\mathbf{X}_k - \bar{\mathbf{X}}_i)(\mathbf{X}_k - \bar{\mathbf{X}}_i)' \quad (2.5)$$

onde m representa o número de amostras preliminares de tamanho n retiradas do processo e X_k é a k -ésima observação p -variável referente a i -ésima amostra.

Na fase seguinte (fase II) utilizaremos as estimativas dos parâmetros para obter a estatística teste cujo procedimento nesta forma é denominado gráfico de controle T^2 de Hotelling, ou seja:

$$T^2 = n(\bar{\mathbf{X}} - \bar{\bar{\mathbf{X}}})' \mathbf{S}^{-1} (\bar{\mathbf{X}} - \bar{\bar{\mathbf{X}}}) \quad (2.6)$$

onde $\bar{\bar{\mathbf{X}}}$ e \mathbf{S} representam, respectivamente, as estimativas para o vetor de médias $\boldsymbol{\mu}$ e matriz de covariâncias $\boldsymbol{\Sigma}$ do processo.

Para um gráfico T^2 de Hotelling os limites de controle são dados pela escolha de um valor α , tal que $P(T^2 \leq LIC) = 1 - \alpha$, conforme limites de controle quando a fase I é estabelecida. Na fase 1, os limites de controle do gráfico T^2 de Hotelling são dados por:

$$LSC = \frac{p(m-1)(n-1)}{mn - m - p + 1} F_{\alpha, p, mn-m-p+1}$$

$$LIC = 0 \quad (2.7)$$

Caso ocorram pontos plotados acima do LSC_{fase1} esses pontos são retirados do conjunto de dados e é recalculado um novo LSC_{fase1} para o teste dos pontos remanescentes. Esse processo é realizado até que todos os pontos restantes estejam abaixo do LSC_{fase1} calculado por último. Assumimos a partir daí que esses pontos restantes são amostras aleatórias de uma distribuição $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ e essas informações serão usadas para a construção do gráfico de controle para futuras observações do processo (Montgomery, 2004).

Na fase II, os novos limites de controle do gráfico T^2 de Hotelling são dados por:

$$LSC = \frac{p(m+1)(n-1)}{mn - m - p + 1} F_{\alpha, p, mn-m-p+1}$$

$$LIC = 0 \quad (2.8)$$

Quando um ponto excede esse limite, dizemos que o processo está fora de controle. Alguma causa especial, também chamada de causa assinalável, que não pertence ao sistema de causas comuns do processo deve estar atuando e tem de ser investigada. Para os limites de controle equações (2.7) e (2.8), $F_{\alpha, p, mn-m-p+1}$ representa o percentil da distribuição F de Snedecor com p e $(mn-m-p+1)$ graus de liberdade, obtido a partir da probabilidade α de alarme falso adotada.

2.2.2 Observações individuais multivariadas

Em alguns processos a técnica de agrupamento de dados via subgrupos racionais é inadequada de se aplicar, como por exemplo, indústrias de processos (químicas, petroquímicas, de mineração e outras) situações onde podemos considerar um item como sendo uma porção de material coletado instantaneamente de determinada corrente ou de um lote homogêneo de algum produto. Nesse caso, temos geralmente subgrupos de itens de tamanho 1 devido às restrições de amostragem. Além disso, a suposição de independência dos itens estaria muito comprometida caso se definisse como amostra um subgrupo de material do lote ou da corrente. Em indústria de peças e componentes, a suposição de independência dos itens pode ser adequada, mas às vezes aumenta o custo para se obter medições múltiplas de subgrupos de itens. Isso, se justifica pelo elevado tempo de ciclo de medição ou pelo seu custo. Observações individuais são muito comuns nesses casos (Pham, 2006).

O projeto de desenvolvimento de gráficos de controle T^2 de Hotelling para observações individuais é separado em duas fases distintas.

A primeira fase consiste em um estudo retrospectivo dos dados, que mostra por meio de testes, se o processo está ou não sob controle estatístico quando os dados representam uma amostra de observações tomadas em todos os pontos do processo, ou seja, uma amostra representativa. Por exemplo, uma amostra retirada durante um específico período de tempo. Esta fase é de primordial importância para estabelecer-se o controle estatístico e se encontrar os limites de controle apropriados para a fase seguinte.

Para a fase I considere o caso em que p características correlacionadas existem, são medidas simultaneamente e estão necessariamente sob controle estatístico. Assume-se que estas características seguem uma distribuição normal multivariada p -dimensional com vetor de médias $\boldsymbol{\mu}' = (\mu_1, \mu_2, \dots, \mu_p)$ e matriz de covariâncias $\boldsymbol{\Sigma}$, onde μ_i é a média da i -ésima característica e $\boldsymbol{\Sigma}$ uma matriz de dimensão $p \times p$ que contém as variâncias-covariâncias das p -características. Assumindo que o processo está sob controle nessa fase, os dados obtidos são usados para estimar os verdadeiros valores da distribuição normal $\boldsymbol{\mu}$ e $\boldsymbol{\Sigma}$. Suponha que

inicialmente se tenha m elementos amostrados, após a retirada dos pontos atípicos restando dessa forma n elementos amostrais com $n \leq m$.

Um problema significativo, no caso das observações individuais é a estimação da matriz de covariâncias do processo. Sullivan e Woodall (1996) apresentam critérios alternativos para estimar a matriz de covariâncias de processos cujo tamanho da amostra é $n=1$. Esses autores propõem alguns procedimentos para obtenção de \mathbf{S} (estimadores) que tornam o deslocamento abrupto (mudança súbita) no processo e deslocamento gradativo (tendência ou mudança gradativa) no vetor de médias do processo. Dois desses estimadores merecem destaque. Um deles, é o estimador “usual” denominado \mathbf{S}_1 , obtido pela simples combinação de todas as m observações, ou seja,

$$\mathbf{S}_1 = \frac{1}{m} \sum_{i=1}^m (\mathbf{X}_i - \bar{\mathbf{X}})' (\mathbf{X}_i - \bar{\mathbf{X}}) \quad (2.9)$$

O outro estimador é uma abordagem multivariada do estimador de amplitudes móveis (Holmes e Mergen, 1993) dado por

$$\mathbf{S}_2 = \frac{1}{2} \frac{\mathbf{V}'\mathbf{V}}{(m-1)} \quad (2.10)$$

onde \mathbf{V} é a matriz contendo os vetores linha das diferenças entre duas observações sucessivas p -variadas dadas por $\mathbf{V}_i = \mathbf{x}_{i+1} - \mathbf{x}_i$, $i = 1, 2, 3, \dots, m-1$ e \mathbf{S}_2 representa a matriz de covariâncias dessas diferenças. Usando \mathbf{S}_2 ao invés de \mathbf{S} na estatística de controle T^2 de Hotelling (2.6), obtém-se um gráfico de controle mais sensível a possíveis mudanças, abruptas ou gradativas, no vetor de médias do processo.

O projeto de desenvolvimento do gráfico de controle estatístico nessas condições para a observação X_i , utiliza a seguinte estatística T^2 de Hotelling para observações individuais

$$T^2 = (\mathbf{X}_i - \bar{\mathbf{X}})' \mathbf{S}^{-1} (\mathbf{X}_i - \bar{\mathbf{X}}) \quad (2.11)$$

onde \mathbf{X}_i e \mathbf{S} representam, respectivamente, as estimativas do vetor de médias e matriz de covariâncias do processo.

Nesta primeira fase selecionam-se m amostras preliminares e representativas para estabelecer-se o controle estatístico (processo padrão) e encontrar os limites de controle apropriados. Os limites de controle para a estatística T^2 , nesta fase, segundo Tracy, Young e Mason (1992), deve basear-se em uma distribuição Beta dados por:

$$LSC = \frac{(m-1)^2}{m} \beta_{[\alpha/2; \frac{p}{2}, (m-p-1)/2]}$$

$$LIC = \frac{(m-1)^2}{m} \beta_{[1-\alpha/2; \frac{p}{2}, (m-p-1)/2]} \quad (2.12)$$

onde $_{[\alpha/2; \frac{p}{2}, (m-p-1)/2]}$ é o $1-\alpha$ percentil da distribuição Beta; $p/2$ e $(n-p-1)/2$ são os parâmetros da distribuição Beta. Como as tabelas para a distribuição Beta não são facilmente avaliadas, pode-se utilizar a seguinte relação $\frac{[p/(n-p-1)]F_{(\alpha; p, n-p-1)}}{1+[p/(n-p-1)]F_{(\alpha; p, n-p-1)}} = \beta_{[\alpha; p/2, (n-p-1)/2]}$ entre as variáveis aleatórias com distribuição Beta e F. Utilizando essa relação obtém-se em termos de percentil da distribuição F, os seguintes limites de controle do gráfico T^2 de Hotelling para a fase I:

$$LSC = \frac{(m-1)^2}{m} \times \frac{[p/(m-p-1)]F_{(\alpha/2; p, m-p-1)}}{1+[p/(m-p-1)]F_{(\alpha/2; p, m-p-1)}}$$

$$LIC = \frac{(m-1)^2}{m} \times \frac{[p/(m-p-1)]F_{(1-\alpha/2; p, m-p-1)}}{1+[p/(m-p-1)]F_{(1-\alpha/2; p, m-p-1)}} \quad (2.13)$$

onde F é a distribuição F, α o nível de confiança, p o número de variáveis observadas e n é o número de observações.

Na fase I, o gráfico de controle multivariado principalmente para o caso bivariado pode ser plotado tanto num gráfico T^2 de Hotelling, em que se tem a idéia do comportamento do processo no tempo, mas também é possível marcar uma variável contra a outra num gráfico de Dispersão, em que os limites de controle são visualizados como uma elipse. A elipse de controle desse gráfico é expressada em termos da amostra X_i de m observações é dada por

$$(\bar{\mathbf{X}}_i - \boldsymbol{\mu}_0)' \mathbf{S}^{-1} (\bar{\mathbf{X}}_i - \boldsymbol{\mu}_0) \leq \chi_{p; \alpha/m}^2 \quad (2.14)$$

onde $\boldsymbol{\mu}_0$ é o vetor de médias do processo sob controle. Este tipo gráfico é normalmente utilizado para visualizar e examinar o comportamento dos dados (Pham, 2006).

Os procedimentos para a fase I na construção de gráficos de controle multivariados principalmente para dados individuais têm recebido menos atenção na literatura que os da fase II.

A segunda fase se caracteriza pelo monitoramento do processo com a utilização dos limites de controle estabelecidos na fase anterior para testar futuras observações. Nesta fase as estimativas obtidas para os parâmetros do processo sob controle estatístico a partir da seleção

das m amostras preliminares fornecem os limites de controle para a estatística T^2 de Hotelling são:

$$LSC = \frac{p(m+1)(m-1)}{m(m-p)} F_{\alpha, p, m-p}$$

$$LIC = 0 \quad (2.15)$$

Quando o número m de amostras preliminares é grande ($m > 100$) é comum utilizar-se um limite de controle aproximado, ou os limites de controle:

$$LSC = \frac{p(m-1)}{m-p} F_{\alpha, p, m-p} \quad (2.16)$$

ou

$$LSC = \chi_{\alpha, p}^2 \quad (2.17)$$

onde $\chi_{\alpha, p}^2$ representa o percentil da distribuição qui-quadrado com p graus de liberdade, obtido a partir da probabilidade α de alarme falso adotada. Para $m > 100$, o limite de controle (equação 2.15) é uma aproximação razoável. No entanto, o limite qui-quadrado (equação 2.16) é apropriado apenas se a matriz de covariância for conhecida, muito embora seja amplamente usado como uma aproximação (Lowry e Montgomery, 1995).

A utilização de gráficos de controle T^2 de Hotelling a partir de observações individuais no monitoramento de processos segundo Mason e Young (2002) exige uma amostra na qual o número de observações m excede o número de variáveis p . Se $m < p$ nem a matriz inversa da covariância Σ^{-1} e nem a estimativa S^{-1} podem ser calculadas. Além disso, um grande número de parâmetro deve ser estimado quando o vetor de médias e a matriz de covariâncias são desconhecidos. Neste caso, n deve ser suficientemente grande para gerar estimativas precisas.

2.3 Gráfico de Controle Multivariado MEWMA

O gráfico de controle T^2 de Hotelling apresentado na seção anterior é uma extensão multivariada do gráfico de Shewhart para a média, onde cada conjunto de dados ou cada valor inserido no gráfico individualmente e a sua relação com outros pontos é determinado apenas pelo gráfico. Para este tipo de gráfico nenhuma estatística que envolva todos os dados anteriores é utilizada, ou seja, apenas a informação do último ponto demarcado no gráfico é considerada. Portanto, insensível a pequenas mudanças no vetor de médias de um processo.

Os modelos de gráficos de controle univariados com memória tais como o gráfico de Média Móvel Exponencialmente Ponderada EWMA (Exponentially Weighted Moving Average) e o de Somas Acumuladas CUSUM (Cumulative Sum) foram desenvolvidos para

oferecer maior sensibilidade a pequenos desvios na média que passam despercebidos pelo gráfico de Shewhart. A extensão do desenvolvimento destes gráficos de controle para o caso multivariado (MEWMA e MCUSUM) se constitui como alternativa ao gráfico T^2 de Hotelling para detectar com maior rapidez mudanças de pequena magnitude no vetor de médias de um processo. Sullivan e Woodall (1998), recomendam a utilização dos gráficos de controle MCUSUM e MEWMA para a análise preliminar de observações multivariadas.

A primeira referência sobre gráfico de controle multivariado MEWMA (Multivariate Exponentially Weighted Moving Average), se deve a Lowry, Woodall e Champ (1992) que definem o MEWMA como uma extensão lógica do controle EWMA univariado ($p = 1$) cuja estatística de controle para monitorar o valor médio do processo é dada por:

$$Z_i = rx_i + (1-r)Z_{i-1} \quad i \geq 1 \quad (2.18)$$

onde $0 < r \leq 1$ é uma constante e o valor inicial (exigido como 1ª amostra em $i=1$) é o valor nominal do processo de modo que $Z_0 = \mu_0$.

Para a versão multivariada ($p > 1$) do gráfico EWMA a estatística da equação (2.18) é expandida conforme equação (2.19)

$$\mathbf{Z}_i = R\mathbf{X}_i + (I - R)\mathbf{Z}_{i-1} \quad (2.19)$$

onde \mathbf{X}_i é o vetor p -dimensional de observações amostrais referente a i -ésima amostra unitária, \mathbf{I} é a matriz identidade de ordem p , R é a matriz diagonal (r_1, r_2, \dots, r_p) que contém p constantes de ponderação $\{r_i \in (0,1]\}$ e \mathbf{Z}_{i-1} é o vetor p -dimensional dos escores referentes à amostra $i - 1$ tomando geralmente como partida $\mathbf{Z}_0 = \boldsymbol{\mu}_0$. Quando $r = 1$ o gráfico de controle MEWMA é equivalente ao gráfico de controle χ^2 de Hotelling.

A estatística de controle utilizada para o gráfico MEWMA é definida como:

$$T_i^2 = \mathbf{Z}_i' \Sigma_{Z_i}^{-1} \mathbf{Z}_i \quad (2.20)$$

onde $\Sigma_{Z_i}^{-1}$ é a inversa da matriz de covariâncias de \mathbf{Z}_i . A matriz de covariâncias de \mathbf{Z}_i é dada segundo a equação

$$\Sigma_{Z_i} = \frac{r[1 - (1-r)^{2i}].\Sigma}{2-r} \quad (2.21)$$

Para o procedimento do gráfico de controle MEWMA pode-se tomar a matriz de covariância assintótica dada por

$$\Sigma_Z = \lim_{i \rightarrow \infty} \Sigma_{Z_i} = \left(\frac{r}{2-r} \right) \Sigma \quad (2.22)$$

de forma análoga ao que ocorre no caso univariado para observações individuais. No caso cujo tamanho da amostra $n > 1$, obtém-se a equação (2.23) corrigida por n :

$$\Sigma_z = \left(\frac{r}{(2-r)n} \right) \Sigma \quad (2.23)$$

O processo é considerado fora de controle se $T_i^2 > h$ na equação (2.20). O valor de h (limite superior de controle) é escolhido a partir do desempenho de ARL para o gráfico de controle MEWMA.

2.4 Gráfico de Controle Multivariado MCUSUM

O gráfico de controle multivariado com memória MEWMA apresentado na seção anterior se constitui como uma das alternativas ao gráfico de controle T^2 de Hotelling em situações onde a detecção de pequenas mudanças nos parâmetros do processo é importante.

Nesta seção apresenta-se o gráfico de controle multivariado com memória MCUSUM como uma outra alternativa ao gráfico T^2 de Hotelling para atender a tais situações. Inicialmente, é apresentado uma breve introdução ao procedimento básico do gráfico de controle univariado CUSUM (Soma Acumulada). Uma apresentação de forma analítica deste tipo de gráfico de controle pode ser encontrada em Hawkins (1998) e Alves (2003, 2004).

2.4.1 Gráfico de controle CUSUM

O gráfico de controle de soma acumulada CUSUM (Cumulative Sum Control Charts) introduzido por Page (1954) é uma boa alternativa ao gráfico de controle de Shewhart para detectar pequenas mudanças na distribuição característica da qualidade, manter um controle apertado sobre o processo e dar uma estimativa do novo nível do processo ou da nova média. Este tipo de gráfico com memória incorpora, diretamente, toda a seqüência de informações demarcando as somas acumuladas dos desvios dos valores da amostra de um valor objetivo. Supondo que amostras de tamanho $n \geq 1$ são coletadas, e que x_j seja a observação da j -ésima amostra. Então, se μ_o é o valor desejado (alvo ou valor nominal) para a média do processo, a soma acumulada CUSUM para observações individuais é obtida demarcando a quantidade resultante da estatística de controle dada por:

$$C_i = \sum_{j=1}^i (x_j - \mu_o) \quad (2.24)$$

onde C_i é a soma acumulada de todos os desvios do valor nominal desde o período 1 até o período i . No entanto, se desejarmos a soma acumulada para um gráfico CUSUM que utiliza as médias de um subgrupo racional substituiremos na equação (2.23) cada observação x_j pela média amostral \bar{x}_j .

Uma maneira mais usual para monitorar a média de um processo é o procedimento CUSUM Tabular ou Algoritmo utilizado para calcular as somas acumuladas unilaterais C_i^+ (Cusum superior) e C_i^- (Cusum inferior) conforme equações (2.25) e (2.26):

$$C_i^+ = \text{máx}[0, X_i - (\mu_o + K) + C_{i-1}^+] \quad (2.25)$$

$$C_i^- = \text{máx}[0, (\mu_o - K) - X_i + C_{i-1}^-] \quad (2.26)$$

onde os valores iniciais $C_i^+ = C_i^- = 0$. Nessas equações, K é o valor de referência e corresponde aproximadamente à metade do desvio entre o valor nominal μ_o e o valor da média fora de controle μ_1 que se tem interesse em detectar rapidamente ou

$$K = \frac{\delta}{2} \sigma = \frac{|\mu_1 - \mu_o|}{2} \quad (2.27)$$

onde δ é o tamanho da mudança em unidades de desvios padrões e σ é o desvio padrão do processo. No algoritmo de soma acumulada para cada amostra são obtidos os sucessivos valores acumulados das estatísticas de controle C_i^+ e C_i^- a serem demarcados no gráfico CUSUM. As somas acumuladas dos desvios C_i^+ e C_i^- são comparadas com o intervalo de decisão h , limite que pode ser atingido pela soma acumulada unilateral.

2.4.2 Gráfico de controle multivariado MCUSUM

O gráfico de controle multivariado mais conhecido e atualmente utilizado com maior frequência para monitorar o vetor de médias de um processo é ainda o tradicional gráfico T^2 de Hotelling onde cada conjunto de dados ou cada valor é inserido no gráfico individualmente e a sua relação com outros pontos é determinado apenas pelo gráfico (Bersimis, Psarakis e Panaretos, 2007). Apesar de extremamente eficaz não é a única ferramenta disponível para a análise estatística multivariada de processos industriais. Em alguns casos, outros tipos de gráficos de controle multivariados podem ser aplicados com a mesma finalidade, e com vantagens. É o caso do gráfico de controle multivariado MCUSUM e o MEWMA apresentado na seção anterior. Esses gráficos detectam pequenas mudanças dos parâmetros de um processo multivariado com um ARL bem menor que o gráfico T^2 de Hotelling. Assim, podem

complementar ou substituir com vantagens o gráfico o T^2 de Hotelling, podendo permitir em função do caso em análise, a obtenção de uma solução mais precisa, a um custo e prazos menores que os requeridos pelas metodologias tradicionais.

O gráfico de controle multivariado de Somas Acumuladas (MCUSUM - Multivariate CUSUM Control Charts) é um procedimento que utiliza a soma acumulada dos desvios de cada vetor aleatório previamente observado em relação ao valor nominal para monitorar o vetor de médias de um processo multivariado.

Os procedimentos de controle estatístico multivariado baseado na filosofia CUSUM são discriminados em duas principais categorias: (i) procedimentos de controle que utilizam múltiplos gráficos de controle CUSUM univariados (abreviados por MCU), desconsiderando assim a correlação entre as variáveis e (ii) procedimentos de controle que utilizam um gráfico de controle CUSUM multivariado (abreviado por MCUSUM), isto é, utilizam a matriz Σ_o de covariâncias das variáveis para obter uma aproximação do gráfico CUSUM em processos multivariados. Portanto, a primeira delas consiste em reduzir as observações multivariadas a um escalar, enquanto a outra consiste em elaborar um gráfico MCUSUM para analisar diretamente estas observações multivariadas.

O primeiro procedimento de controle multivariado MCU (Multiple Univariate CUSUM) foi proposto por Woodall e Ncube (1985), descrevem como um processo multivariado pode ser monitorado utilizando múltiplos gráficos CUSUM univariados para a p variáveis originais ou aplicando gráficos CUSUM univariados para as p componentes principais. A partir de um estudo comparativo entre a utilização do gráfico T^2 de Hotelling e o gráfico de controle bivariado de MCU baseado em componentes principais estes autores concluem que os gráficos MCU são mais sensíveis para detectar mais rapidamente pequenos e moderados deslocamentos no vetor de médias do processo. Neste procedimento, a média de cada variável é monitorada através da estatística de controle e [equações (2.25) e (2.26)], com limites de controle dados por h . Cada um dos p gráficos de controle sinaliza que a i -ésima variável monitorada sofreu um deslocamento quando $C_i^+ > h$ ou $C_i^- < -h$ para valores específicos de K e h . Quando pelo menos um dos p gráficos univariados detectar algum desvio em relação à média da respectiva variável, o gráfico MCU indicará que o processo está fora de controle. Segundo Pignatiello e Runger (1990), os gráficos MCU devem ser aplicados num processo onde existe o interesse em detectar desvios numa direção específica preestabelecida em relação ao vetor μ_o . Estes desvios podem ser caracterizados através de deslocamentos das variáveis do processo sobre seus eixos (desvios específicos em relação as médias das

variáveis) ou na direção dos eixos dos componentes principais (variáveis independentes que são combinações lineares das variáveis originais).

Alwan (1986) propõe um gráfico CUSUM multivariado baseado em testes seqüenciais de proporções de probabilidade. Com essa abordagem multivariada obtém um modelo de decisão linear para este gráfico utilizando uma aproximação para a distribuição χ^2 não central.

Healy (1987) desenvolveu um gráfico CUSUM multivariado para demonstrar que, quando existe interesse em detectar um desvio da média em uma determinada direção preestabelecida, um gráfico CUSUM univariado baseado na combinação linear das p variáveis originais, estruturada naquela direção, fornecerá bons resultados em termos de desempenho ARL. Para desenvolver este gráfico, o autor utiliza como estratégia o fato de que os gráficos CUSUM podem ser vistos como uma série de testes de probabilidade de proporções seqüenciais. Essa estratégia pode ser pouco eficiente se num processo multivariado existir interesse em detectar deslocamentos em várias direções (Pignatiello e Runger, 1990).

Seja a seqüência $X_1, X_2, X_3, \dots, X_i$, variáveis aleatórias independentes distribuídas segundo $N_p(\mu_o, \Sigma_o)$ com um vetor $p \times 1$ de médias μ_o sob controle e uma matriz de covariâncias $p \times p$ conhecida Σ_o .

A estatística CUSUM (Soma acumulada) para detectar uma mudança de μ_o para μ_1 pode ser expressa como:

$$S_i = \max\left[0, S_{i-1} + \mathbf{a}'(\mathbf{X}_i - \mu_o) - \frac{1}{2}d\right] \quad (2.28)$$

onde d é a raiz quadrada do parâmetro de não centralidade e $\mathbf{a}' = \left(\frac{1}{d}\right) \cdot (\mu_1 - \mu_o)' \Sigma_o^{-1}$. A estatística CUSUM S_i indicará se o processo está fora de controle quando $S_i > h$ cujo h é o intervalo de decisão (limite de controle).

Crosier (1988) propõe dois procedimentos para o gráfico de controle CUSUM multivariado. O primeiro procedimento é baseado na raiz quadrada da estatística T^2 de Hotelling enquanto o segundo pode ser obtido substituindo a quantidade de escalares nas expressões de somas acumuladas pelos correspondentes vetores.

O primeiro procedimento, denominado gráfico de controle soma de T, COT ou CUSUM of T, proposto por Crosier (1988), consiste em reduzir as observações multivariadas a escalares. A estatística de soma acumulada para este procedimento é dada por:

$$S_i = \max(0, S_{i-1} + T_i - k) \quad (2.29)$$

Este procedimento baseia-se em realizar uma soma acumulada (CUSUM) da raiz quadrada da estatística T_i^2 para a i -ésima amostra, isto é:

$$T_i = \sqrt{(\mathbf{X}_i - \boldsymbol{\mu}_0)' S^{-1} (\mathbf{X}_i - \boldsymbol{\mu}_0)} \quad (2.30)$$

onde $S_0 \geq 0$, $k > 0$ e X_i é a i -ésima observação p variada do processo e k é o valor de referência. O procedimento indicará um processo fora de controle quando $C_i > h$ onde h é o intervalo de decisão (limite de controle).

O segundo procedimento denominado MCUSUM (CUSUM de vetores) se constitui numa extensão multivariada do gráfico de controle CUSUM univariado de Crosier (1986). Nesse procedimento, as quantidades escalares são substituídas por vetores.

Define-se C_i como:

$$C_i = \sqrt{[(\mathbf{S}_{i-1} + \mathbf{X}_i - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}^{-1} (\mathbf{S}_{i-1} + \mathbf{X}_i - \boldsymbol{\mu}_0)]} \quad (2.31)$$

onde S_i são as somas acumuladas expressadas como:

$$\mathbf{S}_i = \begin{cases} \mathbf{0}, & \text{se } C_i \leq k \\ (\mathbf{S}_{i-1} + \mathbf{X}_i) \left(1 - \frac{k}{C_i}\right), & \text{se } C_i > k \end{cases} \quad (2.32)$$

com $\mathbf{S}_0 = \mathbf{0}$ e valor de referência $k > 0$.

A estatística de controle a ser demarcada no gráfico de controle é dada por:

$$Y_i = \sqrt{\mathbf{S}_i' \boldsymbol{\Sigma}^{-1} \mathbf{S}_i} \quad (2.33)$$

O método sinaliza uma situação fora de controle se $Y_i > h$ onde h é o intervalo de decisão (limite superior de controle). Crosier (1988), demonstra que estes dois procedimentos de gráficos MCUSUM apresentam vantagens em relação ao gráfico T^2 de Hotelling pois podem ser projetados para detectar mudanças específicas no vetor de médias do processo, isto é, características tais como a aplicação do procedimento Resposta Inicial Rápida FIR (*Fast Initial Response*) para melhorar a sensibilidade do gráfico no início do processo ou após o MCUSUM ter dado um sinal de fora de controle. Além disso, a partir de um estudo comparativo entre estes dois procedimentos de gráfico MCUSUM o autor demonstra que o segundo procedimento MCUSUM (CUSUM de Vetores) apresenta um desempenho ARL superior em relação ao procedimento COT (Soma de T).

Dois outros procedimentos para gráficos de controle multivariados CUSUM foram também propostos por Pignatiello e Runger (1990): os gráficos de controle MC1 e MC2. A diferença entre estes dois procedimentos está centrada na forma como a somatória dos vetores

de médias é realizada. No procedimento MC1, realiza-se inicialmente a somatória dos valores de médias e após isso, calcula-se o quadrado dessa somatória. No procedimento MC2, cada vetor de médias é elevado ao quadrado e após isso, determina-se a somatória.

O procedimento, MC1, é baseado no vetor de somas para o tempo t dado por:

$$\mathbf{C}_t = \sum_{i=t-n_t+1}^t (\mathbf{X}_i - \boldsymbol{\mu}_0) \quad (2.34)$$

onde n é o número de subgrupos desde a última renovação, isto é, valor zero do CUSUM.

Então $\frac{1}{n_t} \mathbf{C}_t$ pode ser escrito como:

$$\frac{1}{n_t} \mathbf{C}_t = \left(\frac{1}{n_t} \sum_{i=t-n_t+1}^t \mathbf{X}_i \right) - \boldsymbol{\mu}_0 \quad (2.35)$$

onde o vetor $\frac{1}{n_t} \mathbf{C}_t$ representa a diferença entre a média acumulada da amostra e a média do processo sob controle. Para o tempo t , esta média do processo multivariado pode ser estimada por $\frac{1}{n_t} \mathbf{C}_t + \boldsymbol{\mu}_0$. A extensão de \mathbf{C}_t é

$$\|\mathbf{C}_t\| = \sqrt{\mathbf{C}_t' \boldsymbol{\Sigma}^{-1} \mathbf{C}_t} \quad (2.36)$$

onde \mathbf{C}_t é considerada como uma medida estimada da distância entre a média do processo e o valor alvo $\boldsymbol{\mu}_0$ para a média. O gráfico de controle CUSUM multivariado pode ser construído definindo MC1 como

$$MC1_t = \max(0, C_t - kn_t) \quad (2.37)$$

e

$$n_t = \begin{cases} n_{t-1} + 1, & \text{se } MC1_{t-1} > 0 \\ 1, & \text{caso contrário} \end{cases} \quad (2.38)$$

onde $MC1_t$ é a estatística de controle e k é o valor de referência. O método sinaliza uma situação fora de controle se $MC1_t > h$ onde h é o intervalo de decisão (limite de controle).

O segundo procedimento proposto, considera o quadrado da distância D_t^2 da t -ésima amostra do valor alvo $\boldsymbol{\mu}_0$, onde

$$D_t^2 = (\mathbf{X}_t - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}^{-1} (\mathbf{X}_t - \boldsymbol{\mu}_0) \quad (2.39)$$

tem uma distribuição χ^2 com p graus de liberdade quando o processo está sob controle, e uma distribuição χ^2 não central com p graus de liberdade de parâmetro $\lambda^2(\mu_1)$ quando o processo está fora de controle.

Um gráfico de controle CUSUM univariado pode agora ser formado como

$$MC2_t = \text{máx}(0, MC2_{t-1} + D_t^2 - k) \quad (2.40)$$

com $MC2_0 = 0$, onde $MC2_t$ é a estatística de controle e k o valor de referência. O método sinaliza uma situação fora de controle se $MC2_t > h$ onde h é o intervalo de decisão (limite de controle).

Pignatiello e Runger (1990) demonstram a partir de um estudo comparativo que o gráfico de controle MC1 possui melhor desempenho ARL em relação ao gráfico T^2 de Hotelling na detecção de deslocamentos na média do processo inferiores a 3 (três). Esses autores também mostram que o gráfico MC1 possui, em geral, um melhor desempenho de ARL em relação ao gráfico de controle MCU de Woodall e Ncube (1985) quando o deslocamento no vetor de médias do processo for provocado por desvios simultâneos nas médias de mais de uma variável.

Hawkins (1991) sugeriu para o controle da qualidade multivariado um método de aproximação baseado em variáveis de regressão ajustada aplicado aos gráficos de controle CUSUM e Shewhart com observações individuais. Para esse método, o autor aperfeiçoou a estratégia sugerida por Healy (1987), propondo um MCU, que se mostrou mais eficiente em relação ao MCU apresentado por Woodall e Ncube (1985) para detectar desvios no vetor μ_o em várias direções.

Ngai e Zhang (2001) desenvolveram a partir de simulações dois novos projetos de gráficos de controle CUSUM multivariados com direcionalidade invariante (PPCUSUM e FPCUSUM) baseados em métodos de projeção. Além disso, realizaram um estudo comparativo entre estes dois novos gráficos de controle e os gráficos MC1 (Pignatiello e Runger, 1990), MEWMA₁ e MEWMA₂ (Lowry, Woodall, Champ e Rigdon, 1992) para avaliar tanto o desempenho de ARL destes gráficos quando o desempenho de outros parâmetros tais como SRL (desvio padrão de número de amostras até o sinal), ADRL (média atrasada de número de amostras até o sinal) e SDRL (desvio padrão atrasado de número de amostras até o sinal). Resultados obtidos deste estudo a partir de simulações indicam que estes dois novos projetos de gráficos CUSUM são mais eficazes para evitar problemas de inércia e situações quando deparamos com deslocamentos atrasados da média do processo provenientes dos

gráficos MC_1 , $MEWMA_1$ e $MEWMA_2$. Por esta razão, recomendam os gráficos PPCUSUM e FPCUSUM para detectar trocas atrasadas da média (ADRL) em processos industriais que apresentam alguma anormalidade de característica da qualidade e a utilização dos gráficos $MEWMA_1$, $MEWMA_2$, MC_1 e MCUSUM para detectar trocas iniciais da média desses processos.

Qiu e Hawkins (2001) sugerem um procedimento não paramétrico do gráfico CUSUM multivariado para detectar uma mudança no vetor de médias das medições de um processo baseado nas linhas da seção transversal das medições. Qiu e Hawkins (2003) propõem um outro gráfico de controle multivariado CUSUM também não paramétrico baseado em informações sobre a ordem entre os componentes e de medição da ordem das informações entre a medição e os seus componentes na média sob controle.

Runger e Testik (2004) apresentam uma descrição e uma análise geométrica detalhada de diversas extensões do gráfico de controle CUSUM com intervalo de decisão bilateral simétrico $[-h, h]$ tais como as vantagens e desvantagens de cada uma. As interpretações geométricas apresentadas e os nomes propostos se fundamentam em suas características geométricas distintivas análogas à máscara V. A descrição geométrica conforme característica da extensão multivariada do gráfico CUSUM seguem a forma geométrica de pirâmide, cone, parabolóide e a forma invariante que dão origem aos gráficos denominados MPIRÂMIDE, MCONE, MPARABOLÓIDE E MINVARIANTE.

a) Gráfico MPIRÂMIDE (formato de pirâmide)

O procedimento de controle é baseado em múltiplos gráficos CUSUM univariados com intervalo de decisão bilateral simétrico onde cada valor é operado como uma máscara V. Geometricamente, isto define o formato de uma pirâmide com $p+1$ dimensões que incluem uma linha central para o tempo.

b) MCONE (formato de cone)

É uma extensão direta da máscara V para múltiplas dimensões. Uma vantagem do gráfico MCONE é que é igualmente sensível a deslocamentos (distância de Mahalanobis) de mesmo valor em todos os sentidos do alvo.

c) MPARABOLÓIDE (formato de parabolóide)

Esta derivação é uma etapa fácil previamente fornecida para o gráfico MCONE que pode ser operada com forma de parabolóide situada na direção da distância da observação mais recente.

d) MINVARIANTE (forma invariante)

A derivação deste gráfico é uma contribuição nova obtida de um argumento em que o problema é invariante sob transformações ortogonais quando o valor da média fora de controle for especificado e somente o sentido for desconhecido.

A figura 2.3 é uma ilustração gráfica da geometria de pirâmide, cone e parabolóide para $p=2$ dimensões.

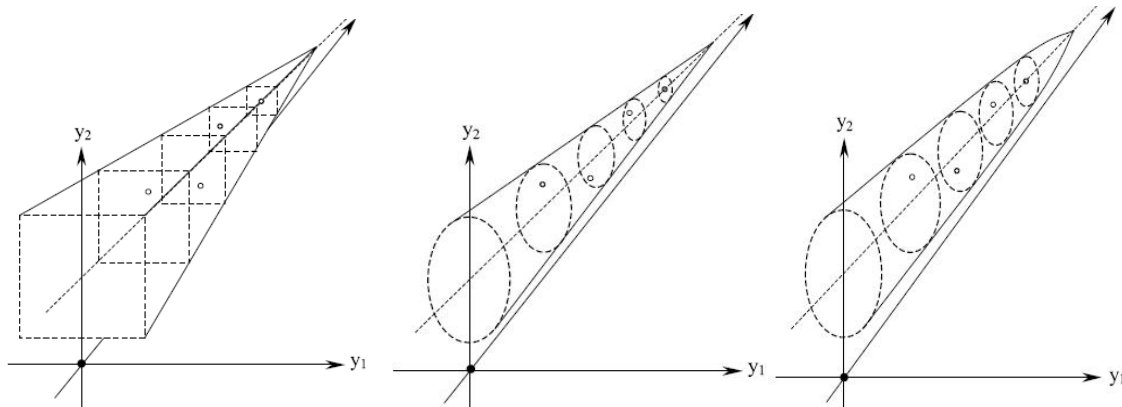


Figura 2.3 Forma bidimensional de pirâmide, cone e parabolóide

Como podemos observar, para todos os gráficos multivariados, o cone, o parabolóide e a forma invariante compartilham de uma mesma característica importante: uma base circular. A diferença está na função que determina os raios destes círculos em pontos precedentes de tempo.

Lee e Khoo (2006) propõem uma extensão multivariada dos procedimentos para o projeto estatístico do gráfico CUSUM univariado baseada tanto em desempenho ARL (número médio de amostras até o sinal) de Gan (1991) quanto em desempenho MRL (número mediano de amostras até o sinal) de Gan (1992). Essa extensão multivariada com base em MRL se constituiu em uma nova contribuição, pois na literatura existente o número mediano de amostras até o sinal só era utilizado em gráficos de controle univariados tais como EWMA (Crowder, 1989) e CUSUM (Gan, 1992).

2.5 O ARL - Número médio de amostras coletadas até à emissão de um sinal

A avaliação e a comparação de diferentes tipos de gráficos de controle multivariados são realizadas mediante a utilização de indicadores de desempenho estatísticos e econômicos. O ARL (Average Run Length) é o número médio de amostras coletadas até à emissão de um sinal. Este sinal, pode ser tanto um falso alarme como um sinal de que o processo realmente está fora de controle após um desvio médio do valor nominal. É o indicador estatístico mais

utilizado para avaliar o desempenho de um gráfico de controle e comparar vários outros procedimentos. Este parâmetro leva em conta os valores dos erros Tipo I e Tipo II associados às tomadas de decisão e as conseqüências econômicas deles resultante, isto é, o custo associado à procura do problema inexistente e o custo associado à fraca qualidade que se obtém no produto final desde quando a mudança ocorre até que seja detectada. Por este motivo, para avaliar os parâmetros de um gráfico de controle, costuma-se estudar o comportamento do ARL face a várias amplitudes de mudança pois pretende-se que o ARL do gráfico seja grande quando o processo está sob controle e bastante pequeno quando o processo está fora de controle.

Quando o processo está sob controle, e os pontos que se marcam no gráfico são independentes, o ARL é dado por

$$ARL_0 = \frac{1}{\alpha} \quad (2.41)$$

onde ARL_0 é o número médio de amostras coletadas até que ocorra um falso alarme, quando o processo está sob controle e α é a probabilidade de cometer um erro Tipo I (probabilidade de falso alarme). No controle estatístico para o caso univariado como, por exemplo, o gráfico \bar{X} de Shewhart com os limites 3σ , toma-se $\alpha = 0,0027$ o que supõe um $ARL_0 \cong 370$, isto é, mesmo que o processo permaneça sob controle, um sinal fora de controle será emitido a cada 370 amostras, em média. Na bibliografia relativa a gráficos de controle multivariados o valor $\alpha = 0,005$, é encontrado com freqüência. Vamos considerar neste trabalho, o valor de $\alpha = 0,005$, $\alpha = 0,002$ e $\alpha = 0,001$, o que supõe um $ARL_0 = 200$, $ARL_0 = 500$ e $ARL_0 = 1000$, respectivamente.

Para determinar uma mudança de valor d , o ARL é dado por

$$ARL_d = \frac{1}{1 - \beta} \quad (2.42)$$

onde ARL_d é o número médio de amostras coletadas até que seja sinalizado uma situação de fora de controle e β é a probabilidade de cometer um erro Tipo II (probabilidade de não detectar que o processo está fora de controle supondo independência das amostras). Portanto, a equação 2.42 só é aplicável aos gráficos de controle do tipo Shewhart como os gráficos \bar{X} e o T^2 de Hotelling e não aos gráficos univariados e multivariados com memória, pois nesses gráficos os pontos plotados não são independentes.

A determinação exata dos parâmetros de um gráfico de controle CUSUM tais como o ARL nem sempre é possível pelo fato das estatísticas de controle serem variáveis aleatórias

dependentes. No entanto existem alguns métodos numéricos que permitem otimizar estes parâmetros, entre as quais podemos destacar o Método de Equação Integral, o Método de Cadeias de Markov e o Método de Simulação.

2.5.1 Método de Equação Integral

O método de equação integral para o caso de um gráfico CUSUM Unilateral Superior é aplicado na obtenção de mudanças positivas (desvios) na média de uma sucessão de variáveis aleatórias independentes e normalmente distribuídas. O estudo para o caso do gráfico CUSUM Unilateral Inferior é semelhante e para o gráfico CUSUM Bilateral recorreremos à relação de Kemp (1971) definida por

$$\frac{1}{ARL^b} = \frac{1}{ARL^u} + \frac{1}{ARL^l} \quad (2.43)$$

que exprime o ARL do projeto Bilateral, ARL^b , à custa dos ARL's dos projetos CUSUM Unilaterais Superior e Inferior, ARL^u e ARL^l , respectivamente.

Consideremos então um gráfico CUSUM Unilateral Superior com valor de referência k e intervalo de decisão h . Este gráfico é equivalente a uma sucessão de testes sequenciais de Wald (1945) com limites em 0(zero) e em h , sendo que o teste termina quando um desses limites é atingido. Kemp (1967, 1971) e Waldmann (1986) apresentam expressões para determinar estes limites, entre o estudo de outras propriedades do gráfico de controle CUSUM. Nas aplicações práticas, em geral, é suficiente trabalhar com limites superiores e inferiores para o ARL de um plano facilitando deste modo o estudo.

Seja X a variável contínua a ser controlada e $f(x)$ a função densidade de probabilidade de X . Representamos por $P(z)$ a probabilidade de um teste que inicia em z vir a terminar porque é absorvido no limite inferior (com $0 \leq z \leq h$), por $N(z)$ o número de amostras consideradas no teste e por $L(z)$ o ARL do gráfico CUSUM, quando este inicia-se no valor z , mas com acumulações subsequentes iniciando em zero. Então, a função densidade de probabilidade conforme definição de soma acumulada, satisfaz as equações integrais descritas por Page (1954) através de um método numérico popularmente conhecido como Método de Equação Integral definido segundo as equações:

$$P(z) = \int_{-\infty}^z f(x)dx + \int_0^h P(x)f(x-z)dx \quad (2.44)$$

$$N(z) = 1 + \int_0^h N(x)f(x-z)dx \quad (2.45)$$

$$L(z) = 1 + L(0)F(-z) + \int_0^h L(x)f(x-z)dx, \quad (0 \leq z \leq h) \quad (2.46)$$

Relacionando estas três equações obtém-se a relação

$$L(z) = N(z) + L(0)P(z)$$

que permite obter o ARL de um projeto que se inicia no valor z do intervalo $[0, h]$.

Fazendo $z = 0$ obtemos $L(0) = N(0) + L(0)P(0)$. Quando $N(0)$ e $P(0)$ são conhecidos, o valor $L(0)$, que corresponde ao valor ARL do gráfico numa situação de controle, é dada pela equação $L(0) = \frac{N(0)}{1 - P(0)}$. Esta equação deve ser aplicada com cuidado uma vez que

numa situação de controle $P(0)$ é um valor muito próximo de 1, para qual o denominador $1 - P(0)$ tende para zero. Assim a estimativa que se aplica para $P(0)$ deve ser bastante precisa de modo que a fórmula anterior forneça valores satisfatórios para o ARL do gráfico.

Apesar do cálculo $P(z)$ e de $N(z)$ recorrendo às formulas anteriores ser mais fácil do que o cálculo de $L(z)$, vários autores tem proposto métodos numéricos aproximados para calcular $P(z)$ e $N(z)$, os quais consistem na substituição das equações integrais mencionadas por sistemas de equações lineares algébricas.

Para o caso quando X é uma variável normalmente distribuída com média θ e desvio padrão 1, temos

$$P(z) = \phi(-z - \theta) + \int_0^h P(x)(2\pi)^{-\frac{1}{2}} e^{-\frac{(x-z-\theta)^2}{2}} dx \quad (2.47)$$

$$e \quad N(z) = 1 + \int_0^h N(x)(2\pi)^{-\frac{1}{2}} e^{-\frac{(x-z-\theta)^2}{2}} dx \quad (2.48)$$

$$\text{onde} \quad \Phi(z) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^z e^{-\frac{t^2}{2}} dt \quad (2.49)$$

Para obter aproximações de $P(z)$ e $N(z)$, as equações integrais de (2.44) e (2.45) são substituídas por sistemas lineares de equações algébricas e resolvidas para as variáveis desconhecidas.

Kemp (1967) apresenta expressões para calcular estimativas para $P(z)$ e $L(z)$ razoáveis no caso da distribuição normal e para os outros tipos de distribuição apresenta expressões que permitem obter estimativas para $L(0)$ mais precisas do que as estimativas que se obtém assumindo a hipótese de normalidade.

Goel e Wu (1971) apresentam expressões simplificadas para estas equações integrais sugeridas por Page (1954), bem como o sistema de equações algébricas lineares que fornece

as estimativas de $P(z)$ e $L(z)$ e do número de amostras consideradas num determinado teste quando a média do processo segue uma distribuição normal reduzida.

2.5.2 Método de Cadeias de Markov

O Método das Cadeias de Markov foi introduzido por Brook e Evans (1972) para determinar as propriedades de um gráfico CUSUM Unilateral. Mais tarde outros autores ampliaram este método de modo a ser aplicável também a gráficos CUSUM Bilaterais. Este método é aplicado tanto para variáveis contínuas quanto para variáveis discretas.

Nesta abordagem, o gráfico é identificado como uma Cadeia de Markov com um número finito de estados. Assim, quando o atributo associado ao processo que tem de ser controlado está associado a uma variável aleatória contínua, é necessário fazer a discretização da estatística de controle, bem como da sua função distribuição de probabilidade, a qual consiste na partição do espaço de estados contínuos da estatística CUSUM, num número finito de intervalos de classe.

Lucas e Crosier (1982) propõem a ampliação dos resultados apresentados no cálculo do ARL de um gráfico Cusum Unilateral para o caso de um gráfico CUSUM Bilateral discretizando os valores de ambas as estatísticas CUSUM, de modo que o gráfico bilateral possa ser representado por uma Cadeia de Markov com um número finito de estados. Woodall (1984) melhorou esta abordagem minimizando o número de estados contidos na Cadeia de Markov para tornar o método tão eficiente quanto possível.

O número de estados que se deve considerar é escolhido de modo a obtermos uma boa aproximação para a matriz das probabilidades de transição e a escolha da amplitude dos intervalos correspondentes a esses estados deve ser realizada de modo conveniente, uma vez que as propriedades de ARL são sensíveis a isso. Quanto menor for a amplitude destes intervalos, melhor será a aproximação. Nesta abordagem, a matriz das probabilidades de

transição, $P(\delta)$, quando ocorre uma mudança de amplitude $\delta = \frac{(\mu_1 - \mu_o)\sqrt{n}}{\sigma}$ no valor médio

de variáveis padronizadas do processo, pode ser aproximada pela matriz em blocos

$$P(\delta) = \begin{pmatrix} Q(\delta) & (I - Q(\delta))1 \\ 0^t & 1 \end{pmatrix}$$
 onde $Q(\delta)$ é a sub-matriz das probabilidades de transição entre

estados e transientes, I representa a matriz identidade e 1 é um vetor coluna de 1 's. Em gráficos de controle o ARL pode ser determinado mediante a aplicação do modelo aproximado

de Cadeias de Markov, partindo da matriz de probabilidades de transição através dos diferentes estados em que se pode modelar o gráfico de controle Brook e Evans (1972).

Neste método, é hábito dividir o procedimento usando matrizes de probabilidades de transição de tamanhos diferentes e o valor ARL do plano é obtido a partir da aproximação

$$ARL = A + \frac{B}{m} + \frac{C}{m^2} \quad (2.50)$$

usando o Método de Mínimos Quadrados, sendo m o número de estados transientes considerados e A o valor de ARL assintótico.

2.5.3 Método de Simulação

Quando se utiliza este método para determinar o ARL, gera-se uma seqüência de observações de uma distribuição pré-especificada e seguidamente calcula-se o valor da estatística de controle até que ele fique fora dos limites de controle.

O número médio de observações necessárias até à emissão de um sinal, isto é, o valor de ARL, varia de seqüência para seqüência. A estimativa para o ARL obtém-se repetindo o procedimento um elevado número de vezes (10.000 ou mais para se obter uma estimativa precisa) e calculando-se os valores de ARL. Para facilitar a implementação deste método, determina-se o ARL assumindo que as mudanças no valor médio em relação ao valor pretendido μ_0 ocorrem no instante em que inicia o procedimento.

2.6 Parâmetros dos Gráficos de Controle CUSUM e MCUSUM

A utilização de métodos numéricos para otimizar os parâmetros do gráfico CUSUM, tais como a determinação da distribuição de ARL e dos parâmetros k e h que melhore o desempenho deste gráfico objeto de estudo de alguns pesquisadores conforme abordado nas três últimas seções; vale lembrar que o aprimoramento e a aplicação destes métodos vêm sendo estudada e proposta por vários outros pesquisadores nas últimas décadas. Lucas (1976) sugere a utilização do método de Equação Integral com o procedimento Quadratura Gaussiana usando 24 pontos para obter aproximações numéricas dos valores de ARL do projeto do gráfico CUSUM a partir da resolução de sistemas de equações lineares algébricas que aproximam as equações integrais (Page, 1954) para as quantidades exigidas. Além disso, o autor sugere a utilização de curvas características de operação como alternativa aos nomogramas desenvolvidos por Goel e Wu (1971) para determinação do parâmetro h do gráfico CUSUM.

Para melhorar a sensibilidade do gráfico CUSUM no momento da partida Lucas e Crosier (1982) propõem a implementação do procedimento Resposta Inicial Rápida (Fast Initial Response - FIR) cuja utilização permite a redução do valor de ARL_0 e uma redução significativa do ARL_δ .

Um método aprimorado para projetar gráficos de controle com base no seu desempenho estatístico, a partir dos valores dos parâmetros especificados para a região sob controle e fora de controle, é proposto por Woodall (1985), que recomenda tomar como critério deste projeto para gráfico de controle a seleção do tamanho da amplitude na média do processo $d = \frac{|\mu_1 - \mu_0| \cdot \sqrt{n}}{\sigma}$ que seja importante detectar.

Vance (1986) desenvolveu um programa computacional para o cálculo dos valores de ARL do gráfico CUSUM, sendo que os resultados obtidos para pequenos valores de ARL's através deste programa, segundo Vance, são aproximadamente iguais quando comparados numericamente com os valores encontrados por Goel e Wu (1971) e Lucas (1976). Este programa apresenta certas vantagens, tais como: (1) facilitar o procedimento do projeto econômico de soma acumulada, pois para o projeto econômico é necessário minimizar a função perda iterativamente, o que requer muitos cálculos de ARL; (2) maior precisão de aproximação da função de distribuição normal cumulativa padronizada.; (3) evitar a utilização de nomogramas em situações onde não são convenientes; (4) maior sensibilidade ao número de pontos utilizado da Gaussiana e (5) modificar facilmente o programa quando da utilização de Quadratura Gaussiana de 24 pontos para até 50 pontos.

O desenvolvimento de um método sofisticado de Quadratura Gaussiana baseado na regra de ponto central de produto (product midpoint) para obter a solução da equação integral (2.50) é proposto por Champ e Rigdon (1991). Estes autores também propõem a aplicação de um formulário de Quadratura Gaussiana para transformar esta equação integral num sistema de equações lineares algébricas. Concluíram através deste estudo que o Método da Equação Integral é preferível quando pode ser encontrada uma equação para a solução aproximada de um problema exigido.

Gan (1991) sugeriu um projeto ótimo para o gráfico de controle CUSUM obtido a partir de uma adaptação do procedimento do gráfico de controle EWMA, proposto Crowder (1989), e de uma otimização seguida de análise de sensibilidade que complementa o projeto de gráfico CUSUM sem o FIR adotado por Lucas (1976), cujos parâmetros k e h para obtenção do perfil de ARL formam um subconjunto muito pequeno em relação ao projeto ótimo do gráfico CUSUM. Este gráfico CUSUM proposto é definido pelo autor como ótimo por possuir

uma especificação do valor do ARL_0 (sob controle) e um mínimo valor de ARL_1 (fora de controle) para detectar rapidamente uma mudança menor da média do processo considerada importante.

Um procedimento relativamente simples, mas bastante preciso para o cálculo de ARL do gráfico de controle CUSUM baseado em uma equação de aproximação para ambas as situações de um processo sob controle e fora de controle é proposto por Hawkins (pág. 37- 43, 1992). A precisão da otimização utilizada neste procedimento exige a aplicação de tabelas cujos valores de h e k pertençam aos intervalos $[0,9)$ e $(-2,3]$, respectivamente.

Jun e Choi (1993) desenvolveram duas aproximações para melhorar a estimação de ARL dos gráficos de controle CUSUM utilizando o método de simulação e técnicas de redução de variância. A primeira delas utiliza o risco total como uma variável de controle conforme proposto por Ross (1990) e a segunda aproximação utiliza estimadores de proporção baseado no comprimento de ciclos, onde um ciclo é completado se o gráfico CUSUM for reajustado à origem ou emitir um sinal.

O desenvolvimento de um programa computacional que envolve o Método da Equação Integral bem como a utilização de Quadratura Gaussiana com 24 pontos para computar a função de probabilidade e os percentuais de ARL do gráfico de controle CUSUM é proposto por Gan (1993). Este programa possui também uma rotina que fornece o ARL e o esboço para a curva da função probabilidade e da função de distribuição cumulativa correspondente.

Hawkins e Olwell (1998) propõem dois programas computacionais integrado ao livro “Cumulative Sum Charts and Charting for Quality Improvement” para calcular facilmente os valores dos parâmetros do gráfico CUSUM k , h e ARL relacionados. A utilização destes dois programas segundo os autores deste livro, facilitam a escolha de parâmetros para o gráfico CUSUM adaptados às necessidades de um processo pois ponderam a taxa de falsos alarmes que se pode tolerar.

Um algoritmo rápido e preciso é proposto por Luceño e Puig-Pey (2000) para determinar a distribuição de probabilidade do número de amostras coletadas até à emissão de um sinal (RL) do gráfico CUSUM a partir da utilização numérica de uma fórmula de recorrência estável baseada no Método de Quadratura Gaussiana capaz de melhorar a eficiência e precisão dos métodos existentes. Este algoritmo pode ser aplicado para descobrir se a aproximação adequada é geométrica ou não e, quando possível, permite trocar para o recurso geométrico.

Rao, Disney e Pignatiello (2001) propõem um método de singularidade e

convergência que utiliza uma única equação integral $L(y) = 1 + L(0)F(k - y) + \int_0^h L(x)dF(x + k - y)$ e um algoritmo numérico baseado em Gauss-Legendre e Quadratura Simpson para obter aproximações precisas de ARL resolvendo o sistema de equações lineares algébricas a partir desta única equação sujeita a certas condições de regularidade que na prática são frequentemente satisfeitas.

Luceño e Puig-Pey (2002) desenvolveram um programa de computação para o algoritmo rápido e preciso proposto por eles em (2000) para determinar a distribuição de probabilidade de RL do gráfico CUSUM com parâmetros conhecidos para monitorar a média de processos com distribuições contínuas tanto normais quanto não simétricas. Este programa possui um diferencial de aproximação com a substituição de recursos numéricos de matrizes usados por Gan (1993) pela utilização de fórmulas de recorrência eficientes que permite estender as avaliações da distribuição de RL para um número de amostras muito maior. Além disso, este programa, possui uma rotina precisa que computa as abscissas de Gauss-Legendre e pesos, dispensando a utilização de tabelas para armazenar estas magnitudes e, mais importante do que isso, é a facilidade para efetuar a troca do número de pontos da Gaussiana.

Um método de aproximação numérica é proposto por Jones, Champ e Rigdon (2004) para a distribuição de RL e momentos do gráfico CUSUM com estimação de parâmetros para avaliar o desempenho deste gráfico cujos parâmetros são calculados a partir de uma variedade de situações práticas. Neste método são estudados os efeitos da estimação de parâmetros para o procedimento do gráfico CUSUM tais como o efeito da derivação da distribuição do comprimento de corrida de um gráfico Cusum unilateral condicionada a valores específicos de estimativas realizadas pela média do processo e desvio padrão.

Como podemos observar, é ampla a quantidade de trabalhos desenvolvidos nas últimas décadas que abordam o aprimoramento e a aplicação de métodos numéricos para obter de forma aproximada os parâmetros para avaliar o desempenho do gráfico CUSUM univariado. No entanto, quando se trata da otimização dos parâmetros do gráfico de controle multivariado MCUSUM poucos trabalhos que abordam o referido tema com expressiva relevância têm sido desenvolvidos neste mesmo período principalmente envolvendo o gráfico de controle MCUSUM (Crosier, 1988), objeto de estudo desta pesquisa.

Woodall e Ncube (1985) aplicaram o método Cadeias de Markov para obter aproximações de ARL para o desenvolvimento do primeiro projeto de gráfico de controle multivariado que envolve múltiplos gráficos CUSUM univariados (MCU). Além disso,

realizam um estudo comparativo dos valores de ARL para avaliar o desempenho entre os gráficos de controle T^2 de Hotelling e CUSUM Bivariado de MCU baseado em variáveis originais ou em componentes principais cujos resultados obtidos, segundo os autores, indicam o gráfico CUSUM Bivariado mais sensível para detectar rapidamente pequenas e moderadas mudanças do processo consideradas importante.

Crosier (1988) propõe a generalização do gráfico de controle CUSUM univariado sugerido por ele em 1986, para situações multivariadas, substituindo as quantidades escalares por correspondentes vetores nas expressões de somas acumuladas do gráfico CUSUM univariado. Para isso, desenvolveu um estudo comparativo para avaliar o desempenho de ARL entre o gráfico T^2 de Hotelling e os dois novos projetos de gráficos multivariados CUSUM (CUSUM COT e MCUSUM) propostos por ele para monitorar a média de processos com multivariáveis normais a partir de um estudo comparativo. A aproximação de ARL para o gráfico CUSUM COT com ou sem o FIR nesse estudo é obtida a partir do método Cadeias de Markov, gráfico MCUSUM do método de simulação e gráfico T^2 de Hotelling do método analítico. Os resultados obtidos, segundo o autor, demonstram que ambos os gráficos CUSUM COT e MCUSUM são mais sensíveis que o gráfico T^2 de Hotelling para detectar mais rapidamente pequenos deslocamentos no vetor de médias do processo. Além disso, justifica sua preferência pelo gráfico MCUSUM em relação ao gráfico CUSUM COT pelo fato do vetor de médias do MCUSUM indicar o sentido em que a média está sendo deslocada.

Pignatiello e Runger (1990) desenvolveram um estudo comparativo para avaliar o desempenho de ARL entre diferentes gráficos de controle tais como MCU (CUSUM Bivariado), T^2 de Hotelling além dos dois novos projetos de gráficos de controle CUSUM multivariados (MC_1 e MC_2) propostos por eles para monitorar a média de processos. Nesse estudo, utilizam o método da Cadeia de Markov de Brook e Evans (1972) para obter a aproximação de ARL do gráfico MCU e o método de Simulações de Monte Carlo para os demais projetos de gráficos de controle multivariados deste estudo. Os resultados obtidos, segundo os autores, demonstram que o gráfico MC_1 proposto possui, em geral, um melhor desempenho de ARL em relação a MCU quando o deslocamento no vetor de médias do processo for provocado por desvios simultâneos nas médias de mais de uma variável.

Lee e Khoo (2006) aplicaram o método de Cadeias de Markov para obter em situações sob controle os parâmetros ótimos do gráfico de controle MCUSUM para observações individuais baseados em ARL e MRL, extensão multivariada dos gráficos de controle CUSUM univariados ótimos de Gan (1991),(1992), respectivamente e o método de Simulação para

situações fora de controle. Apresentam exemplos de projetos estatísticos ótimos para o gráfico MCUSUM baseado em ARL e MRL cujos parâmetros ótimos tais como o valor de referência k e o limite superior de controle h correspondente são determinados a partir de valores tabelados para $p = 2, 3$ e 4 características da qualidade tanto para o ARL quanto para o MRL sob controle de 100, 200, 370, 500 e 1000.

Neste trabalho, propõe-se o método de Equação Integral com Quadratura Gaussiana para otimizar os parâmetros do gráfico MCUSUM (CUSUM de Vetores de Crosier, 1988) para no máximo quatro variáveis uma vez que os procedimentos convencionais de construção de gráficos multivariados são razoavelmente eficazes desde que o número de variáveis p do processo não seja grande. Segundo Montgomery (2004), a construção de gráficos de controle multivariados com elevado número de variáveis é inviável pois à medida em que p aumenta o desempenho de ARL para detectar uma mudança específica dessas variáveis para gráficos multivariados também aumenta. Uma alternativa para processos com elevado número de variáveis (características da qualidade) são os métodos de estrutura latente tais como Análise de Componentes Principais (ACP). Alves, Henning e Samohyl (2008b) propõem o desenvolvimento de gráficos de controle multivariados baseados na projeção de dados via ACP como uma alternativa adequada para o tratamento de dados de processos com elevado número de variáveis. Utilizando uma rotina numérica no pacote GNU R mostram o quanto a ACP é fundamental para a redução da dimensionalidade da matriz original de dados que permite a visualização da variação dos resultados e são úteis, sobretudo para identificar quais variáveis estão bem explicadas no modelo. Souza (2000) desenvolveu uma metodologia para auxiliar o monitoramento e a realimentação de um processo de produção multivariado aplicando ACP aos gráficos de controle cuja aplicação além de contribuir para a redução do número de variáveis a serem investigadas proporcionou também a identificação da variável ou conjunto de variáveis causadoras da instabilidade deste processo.

2.7 Projeto ótimo aplicado para obter os parâmetros do gráfico MCUSUM

No controle estatístico de processo univariado uma única característica da qualidade quantitativa com média μ_o e desvio padrão σ_o é monitorada, o tamanho da mudança d produzido na média quando esta passa de μ_o para μ_1 , e se mantém constante a dispersão, é medido pela expressão:

$$d = \frac{|\mu_1 - \mu_o|}{\sigma_o} \quad (2.51)$$

onde $d=0$ quando o processo está sob controle e $d > 0$ quando está fora de controle.

No entanto, quando se deseja controlar simultaneamente duas ou mais características da qualidade correlacionadas (controle estatístico multivariado) então o tamanho da mudança a ser detectado é medido através da Distância de Mahalanobis (tamanho da mudança de um vetor de médias ou distância entre vetores) utilizada em análise multivariada. Neste caso, a distância entre o vetor de médias original e o novo vetor de médias é

$$d = \sqrt{(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)} \quad (2.52)$$

Em análise de processos, em geral, não é viável que o procedimento de controle escolhido gere muitos falsos alarmes. Portanto, a estratégia utilizada no projeto de um gráfico de controle baseia-se na escolha de um valor de ARL grande quando o processo estiver sob controle. Este valor é representado por ARL_0 e indica em média o número de amostras coletadas até que ocorra um falso alarme. De posse do valor de ARL_0 o gráfico de controle deve ser projetado de modo que o tamanho da mudança d , que se deseja detectar, seja sinalizado no gráfico a partir de um número de amostras coletadas até que seja sinalizada uma situação fora de controle referente a um desvio do valor nominal.

O procedimento de projeto ótimo em um ponto do gráfico de controle consiste em especificar o ARL, e a magnitude da mudança no processo que desejamos detectar, selecionando então a combinação ótima (k,h) que proporcione o mínimo ARL fora de controle (ARL_d) para um certo número de variáveis a controlar e um determinado tamanho de amostras.

O par de parâmetros (k,h) é ótimo, no sentido que para uma probabilidade de erro tipo I fixado (ARL_0) produz a menor probabilidade de erro tipo II (ARL_d) possível para uma mudança especificada. Este par de valores (k,h) ótimo, em geral, dependerá da magnitude da mudança. Uma análise de sensibilidade pode mostrar como varia a probabilidade do erro tipo I para diferentes pares (k,h) próximos da combinação (k,h) ótima (Gan, 1991).

A melhoria de desempenho dos parâmetros do gráfico CUSUM no controle estatístico de processos industriais tem sido objeto de estudo de alguns autores que propõem tabelas e métodos gráficos para selecionar os parâmetros ótimos. Estas tabelas apresentam um número muito reduzido de situações, sendo, portanto, muito limitada a sua aplicação. A otimização do gráfico CUSUM univariado resulta, por exemplo, de referências tais como de Lucas e Crosier (1982) que sugerem para análise do perfil de desempenho deste gráfico valores tabelados dos parâmetros ótimos k e h em função do tamanho da mudança a ser detectada para um determinado valor de ARL_0 . Para estes valores tabelados se desejarmos

escolher k em relação ao tamanho da mudança que desejamos detectar; utilizamos $k = \frac{1}{2}d$, onde d é o tamanho da mudança em unidades de desvio padrão. Essa abordagem tende a minimizar o valor de ARL_d para detectar uma mudança de tamanho d , para um valor fixado. Conforme tabelas propostas por Lucas e Crosier (1982) se escolhermos o valor de referência $k=0,5$ e intervalo de decisão $h=5$ resulta no valor $ARL_o=465$. Aplicando o FIR utilizando-se o valor inicial de vantagem $S_o = \frac{h}{2} = 2,5$, recomendado por Lucas e Crosier (1982), o valor de ARL_o cai para 430. Um procedimento de soma acumulada projetado para detectar, por exemplo, uma mudança $\delta=1$ seria detectada em média em 10 amostras num procedimento padrão ($S_o = 0$). Com o uso do FIR ($S_o = 2,5$) é detectado em 6 amostras. O valor de ARL_δ com o uso do FIR diminui à medida que diminui o valor de ARL_o .

No entanto, o projeto ótimo de um gráfico CUSUM multivariado é baseado no ARL_o desejado, tamanho da mudança que se deseja detectar e no número de variáveis a controlar. Crosier (1988) apresenta para análise do desempenho do gráfico de controle multivariado MCUSUM tabelas obtidas por simulação para $p = 2, 5, 10$ e 20 os valores ótimos do par de parâmetros (k,h) com e sem a implementação do FIR tanto para o $ARL_o=200$ quanto para o $ARL_o=500$. Além disso, apresenta uma tabela de ARL 's obtidos por simulação para $ARL_o=200$, com tamanho de mudança d iguais a $d=0,5, 1, 1,5, 2, 2,5, 3$ ótimos do par de parâmetros (k,h) iguais a $(0,5, 5,5)$, $(1, 2,99)$ e $(1,5, 1,87)$.

Por exemplo, se desejarmos detectar uma mudança $d=1$ para $p=2$ e um $ARL_o=200$, o gráfico MCUSUM ótimo (Crosier,1988) a ser utilizado teria um valor de $k=0,5$ e $h=5$.

Pignatiello e Runger (1990) apresentam tabelas obtidas tanto pelo Método de Simulação de Monte Carlo como pelo Método de Cadeias de Markov para um $ARL_o=200$, $d=0,5, 1, 1,5, 2, 2,5, 3$ e $p=2, 3$ e 10 (variáveis independentes) para os gráficos multivariados MC_1, MC_2, MCU (Woodall e Ncube, 1985) e T^2 de Hotelling. Além disso, apresentam tabelas para $ARL_o=200$, $d=0,5, 1, 1,5, 2, 2,5, 3$ e $p=2$ (variáveis dependentes) para o gráfico MCU Bivariado (Woodall e Ncube, 1985).

No presente trabalho é proposto um método de cálculo baseado em equações integrais para determinar tanto o ARL como a melhor combinação dos parâmetros k e h do gráfico de controle MCUSUM para qualquer mudança do vetor de médias considerada importante que se deseja detectar. Os valores tabelados existentes atualmente na literatura para estes parâmetros são muito limitados para aplicações em situações práticas.

2.8 Método para Otimização dos Parâmetros do Gráfico MCUSUM

Conforme abordado em seções anteriores, o gráfico de controle multivariado T^2 de Hotelling tem como memória apenas o último ponto demarcado no gráfico, ou seja, só leva em consideração a informação atual. Como as informações são analisadas separadamente, é difícil detectar padrões cíclicos e tendenciosos nos dados, o que se traduz na pouca eficiência deste gráfico para detectar pequenas mudanças no vetor de médias do processo. Para melhorar a eficácia na detecção de pequenas amplitudes de mudança, foram desenvolvidos gráficos multivariados de memória tais como o MEWMA e o MCUSUM que ponderam todas as informações do processo acumulando a informação mais recente com informações anteriores, e, com isso, detectam pequenas mudanças no vetor de médias com melhores valores de ARL que o gráfico T^2 de Hotelling.

Na abordagem sobre projeto econômico de gráficos de controle, Woodall (1986,1987) justifica que em alguns projetos econômicos a probabilidade de erro tipo I do gráfico de controle é consideravelmente maior que o desejado no projeto estatístico, e que isto conduzirá a um número maior de falsos alarmes, situação esta indesejável.

Em situações em que o processo é muito capaz ou dificilmente ajustável pode não ser interessante detectar mudanças de pequena magnitude. Segundo Woodall (1985), tentar ajustar um processo quando a mudança da média é muito pequena pode levar ao fenômeno de sobre ajuste e da introdução de variabilidade extra no processo. Portanto, é mais interessante decidir que tamanho de mudança é realmente importante detectar, e tendo isto como base, selecionar aquele gráfico de controle que seja muito eficiente quando realmente é necessário, e que tenha uma probabilidade de falsos alarmes realmente baixa. Este trabalho, propõe o desenvolvimento de um método numérico que permita ao usuário determinar on-line os parâmetros ótimos do gráfico multivariado MCUSUM para monitorar processos em que não resulte apenas a detecção de mudanças de pequena magnitude, mas sobretudo seja ao mesmo tempo um método muito eficaz para detectar mudanças realmente importantes.

2.8.1 Otimização do MCUSUM baseado em Regiões de Máxima e Mínima Potência

O método usual de comparação da potência de diferentes gráficos de controle se fundamenta nos valores de ARL. Para que dois projetos de gráficos de controle sejam comparados, ambos, devem apresentar o mesmo ARL quando o processo se encontra sob controle, isto é, quando não existe nenhuma mudança na média da característica da qualidade

a controlar. O gráfico de controle mais eficiente ou de maior potência para detectar mudanças será aquele que apresente um ARL grande quando o processo se encontra sob controle e um ARL pequeno quando o processo se encontra fora de controle.

Por outro lado, é sabido que sob o ponto de vista econômico se considera que o custo de operação do processo (custo médio por hora de produção quando há uma mudança na média do processo) é proporcional ao número de falsos alarmes, e que sob o ponto de vista estatístico pode ocorrer que o número de falsos alarmes seja uma fonte de variabilidade extra no processo. Diante do que anteriormente foi exposto, Woodall (1985) sugere um projeto estatístico para gráfico de controle e recomenda tomar como critério deste projeto, a seleção do tamanho da mudança que seja importante detectar. Para isso, propõe a definição de três regiões de controle: região sob controle, região indiferente, e região fora de controle. Estas regiões são delimitadas por dois valores (A e B) do tamanho da mudança a ser detectada conforme figura (2.4). No caso univariado este tamanho da mudança é medido em unidades de desvios padrão conforme equação (2.51) enquanto que no caso multivariado é medido a partir da Distância de Mahalanobis conforme equação (2.52).

As regiões de máxima e mínima eficácia podem ser definidas como:

- a) Região Sob Controle, $[0, A]$. Esta região corresponde a um estado equivalente ao de estado sob controle e corresponde ao tamanho de mudança compreendido entre $d=0$ e $d=A$. Nesta região não se deseja detectar alguma mudança. No entanto, deseja-se um ARL máximo. Se o gráfico mostrar um sinal fora de controle considera-se então este sinal como um falso alarme.
- b) Região Fora de Controle, $[B, \infty)$, corresponde ao valor de mudança $d=B$, a partir do qual se requer a máxima eficácia de detecção. Além disso, se deseja um ARL mínimo.
- c) Região Indiferente, (A, B) , compreendida entre $d=A$ e $d=B$. Nesta região é indiferente detectar ou não uma mudança no processo.

Como exemplo, na figura 2.4 pode-se observar a curva de ARL para dois gráficos de controle multivariados do tipo MCUSUM e T^2 de Hotelling.

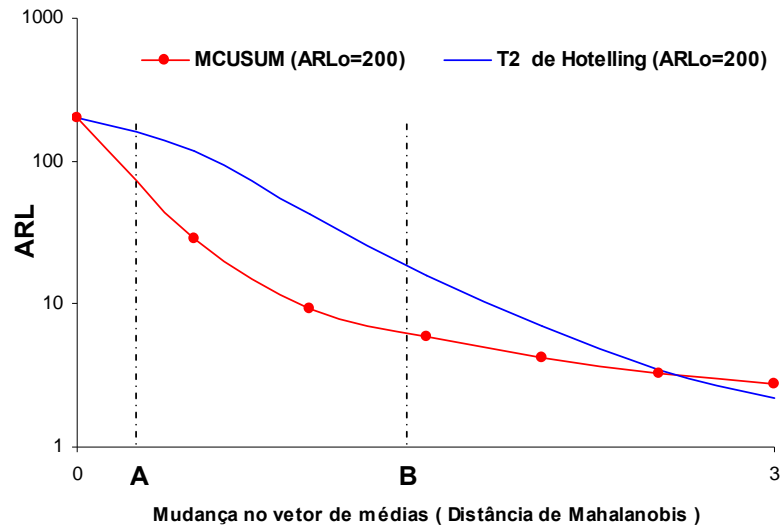


Figura 2.4 Valores de ARL dos gráficos MCUSUM e T^2 Hotelling para $ARL_0=200$

Conforme figura 2.4 podemos observar que ambos os gráficos têm o mesmo ARL para $d = 0$, sendo o gráfico ótimo aquele que apresenta maior ARL na região sob controle e menor ARL na região fora de controle. Isto significa que apresenta menor probabilidade de falso alarme e é mais eficaz para detectar mudanças que devem ser rapidamente detectadas. Observamos realmente que quando, por exemplo, nosso processo for muito capaz, tentar detectar mudanças muito pequenas pode não ser interessante.

O enfoque de Woodall (1985) que aborda a utilização de regiões de máxima e mínima potência em projetos de otimização para gráficos de controle se encontra justificado em situações que aparecem freqüentemente na indústria. Entre outros, e como exemplo, vamos analisar três tipos de processos característicos:

- Processos muito capazes: um tipo de processo onde pode resultar pouco interesse em detectar pequenas mudanças na média do processo referente ao valor nominal são aqueles que apresentam elevada capacidade real ($C_{pk} = 1,5$), típicos de setores como o da automação.
- Processos dificilmente ajustáveis: em alguns processos industriais, independentemente de considerações econômicas, acontece que por razões técnicas ou físicas tenham um comportamento que os transforma em dificilmente ajustáveis, quando no início o processo já está fora do valor nominal, típicos de processos onde uma ferramenta de corte sofre desgastes contínuos e graduais.
- Processos com um elevado custo econômico de ajuste: em muitos processos como o de estampagem de chapas, por exemplo, o ajuste para que se trabalhe no valor nominal centrado nas especificações requer a parada do processo e a regulagem da máquina utilizada neste

trabalho. Segundo, Aparisi e Díaz (2007), esta operação pode requerer um período elevado de tempo com a conseqüente perda de produção associada a esse tempo produtivo.

2.9 Integração Numérica

A integração numérica é uma excelente ferramenta matemática, geralmente empregada por cientistas e engenheiros, como o único recurso para resolver em ambientes computacionais muitos problemas de engenharia que geram integrais para as quais soluções analíticas não são sempre possíveis.

Designamos, de um modo geral, por integração numérica o processo de obter $I(f) = \int_a^b f(x)dx$ em que f é uma função integrável no intervalo $[a,b] \subset \mathfrak{R}$.

Atualmente, os inúmeros recursos computacionais existentes, têm sido responsável pelo desenvolvimento de uma grande variedade de métodos numéricos aplicados para simplificar a integral. Esses métodos consistem em aproximar a função $f(x)$ por um polinômio interpolador e determinar analiticamente a integral desse polinômio no intervalo $[a,b]$ (Campos Filho, 2001). Neste trabalho, utiliza-se aproximação de integração numérica via método de Quadratura Gaussiana para determinar os parâmetros do gráfico de controle multivariado MCUSUM tais como o ARL, k e h .

A idéia básica da integração numérica é de aproximar a função $f(x)$ por um polinômio de grau n , $P_n(x)$ e, então, realizar a integração deste polinômio de aproximação, visto que cada termo no polinômio pode ser integrado analiticamente. A precisão da integração numérica depende da escolha adequada desse polinômio de aproximação (Billo, 2007).

Os métodos de integração numérica podem ser classificados em dois grupos, as Fórmulas de Integração de Newton-Cotes e as Quadraturas Gaussianas.

As Fórmulas de Newton-Cotes usam valores da função em pontos uniformemente espaçados. Essa restrição é conveniente quando as fórmulas são combinadas de modo a fornecer regras compostas, mas essa utilização pode reduzir de maneira significativa a precisão da aproximação (Jeannequin, 2003).

As Quadraturas Gaussianas fornecem uma flexibilidade não somente em escolher os pesos, mas em localizar pontos onde as funções são avaliadas para otimizar a aproximação em vez de considerar apenas pontos regularmente espaçados (Kiusalaas, 2005). Como resultado, as Quadraturas Gaussianas fornecem duas vezes mais pontos de precisão do que as Fórmulas de Newton-Cotes com o mesmo número de funções de avaliação. Se a função é conhecida e

suave, as Quadraturas Gaussianas, em geral, têm vantagens e eficiência decisivas (Jeannequin, 2003).

2.9.1 Quadratura Gaussiana

A Quadratura Gaussiana também denominada Quadratura de Gauss-Legendre é um método de integração numérica robusto, que apresenta uma série de vantagens em várias situações, tais como a escolha de pontos discretos x_i na qual a função $f(x)$ é avaliada. Quando esta escolha for realizada adequadamente pode levar a uma maior precisão da avaliação da função (Canale e Chapra, 2008). A idéia central deste método para n pontos de quadratura é obter a solução da integral de uma função $f(x)$ em termos de uma soma da forma:

$$I(f) = \int_a^b f(x) dx \cong w_1 f(x_1) + w_2 f(x_2) + \dots + w_n f(x_n) \cong \sum_{i=1}^n w_i f(x_i) \quad (2.53)$$

onde os termos w_i são os pesos (coeficientes ponderados) e x_i são os pontos de quadratura.

No cálculo do valor aproximado da integral definida utiliza-se combinação linear dos valores da função $f(x)$ em certos pontos discretos x_i , $a \leq x_i \leq b$. Mais formalmente, utiliza-se o processo de quadratura que consiste em substituir uma integral com a soma de suas partes. Os nós x_1, x_2, \dots, x_n são n pontos distintos de $[a, b]$ para os quais os valores de w_i e x_i são obtidos de maneira que a integral seja exata para polinômios de grau $\leq 2n - 1$. Os valores tanto de x_i quanto de w_i são escolhidos de modo a minimizar o erro esperado no cálculo da aproximação. Para medir esta precisão, admite-se que a melhor escolha desses valores é a que produz resultado exato para a maior classe de polinômios, isto é, aquela escolha que fornece um maior grau de precisão.

A condição da equação (2.53) ser exata para polinômios de grau $\leq 2n - 1$ nos leva a um sistema não linear de equações algébricas, cuja solução nos fornece os valores de x_i e w_i .

Para ilustrar o procedimento de escolha dos parâmetros associados, é mostrado como selecionar os pesos e nós quando $n=2$, cujo intervalo de integração é $[-1, 1]$. Suponha que se deseja determinar w_1, w_2 e x_1, x_2 de maneira que a fórmula

$$I(f) = \int_{-1}^1 f(x) dx \cong w_1 f(x_1) + w_2 f(x_2)$$

de um resultado exato sempre que $f(x)$ seja um

polinômio de grau $2(2)-1 = 3$ ou menor, ou seja, quando $f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3$ para qualquer conjunto de constantes a_0, a_1, a_2, a_3 .

Como $\int (a_0 + a_1x + a_2x^2 + a_3x^3)dx = a_0 \int 1dx + a_1 \int xdx + a_2 \int x^2dx + a_3 \int x^3dx$, este fato equivale a mostrar que a fórmula dá resultados exatos quando $f(x)$ é igual a $1, x, x^2$ e x^3 . Assim sendo, precisamos de w_1, w_2 e x_1, x_2 de maneira que se constitui num sistema com 4 (quatro) equações e 4 (quatro) incógnitas, ou seja,

$$\begin{cases} w_1 \cdot 1 + w_2 \cdot 1 = \int_{-1}^1 1dx = 2 \\ w_1 \cdot x_1 + w_2 \cdot x_2 = \int_{-1}^1 xdx = 0 \\ w_1 \cdot x_1^2 + w_2 \cdot x_2^2 = \int_{-1}^1 x^2dx = \frac{2}{3} \\ w_1 \cdot x_1^3 + w_2 \cdot x_2^3 = \int_{-1}^1 x^3dx = 0 \end{cases} \quad (2.54)$$

Resolvendo o sistema de equações (2.54) obtém-se as soluções únicas $w_1 = 1, w_2 = 1,$

$x_1 = \frac{-1}{\sqrt{3}} = -0,577350269189626, x_2 = \frac{1}{\sqrt{3}} = 0,577350269189626$ que permite escrever a

fórmula de aproximação de Gauss-Legendre de dois pontos

$$\int_{-1}^1 f(x)dx \cong w_1 f(x_1) + w_2 f(x_2) = f\left(\frac{-1}{\sqrt{3}}\right) + f\left(\frac{1}{\sqrt{3}}\right).$$

Essa fórmula possui grau de precisão 3,

isto é, produz resultado exato para todo polinômio de 3º grau ou menor. Através de desenvolvimento semelhante ao apresentado para $n=2$, é possível encontrar os termos de ponderações w_i e x_i para a integração Gaussiana de ordem 3,4, e maior.

O procedimento de integração numérica Quadratura Gaussiana tem por principal vantagem o fato de poder ser facilmente inserido num programa computacional destinado à análise de estruturas pelo método dos elementos finitos para aplicações que envolvem integrais cujas soluções analíticas nem sempre são possíveis.

Um outro método alternativo de integração Gaussiana pode ser encontrado no Apêndice A deste trabalho.

2.10 Síntese do Capítulo

Neste capítulo além de alguns fundamentos básicos preliminares ao estudo de gráficos de controle multivariados são abordados os principais gráficos de controle multivariados: o tradicional T^2 de Hotelling e os gráficos com memória MEWMA e o MCUSUM, objeto de estudo deste trabalho.

Um estudo dos principais parâmetros do gráfico MCUSUM propostos neste trabalho tais como o valor de referência k , o limite superior de controle h , o número médio de amostras coletadas até à emissão de um sinal (ARL) e os métodos numéricos para otimizar estes parâmetros enfatizando o Método de Equação Integral proposto nesta pesquisa.

A apresentação de um procedimento para projetar estatisticamente o gráfico MCUSUM ótimo baseado nestes parâmetros segundo a ótica de regiões de máxima e mínima potência para selecionar o tamanho da mudança que seja importante detectar e tendo isto como referência optar pelo gráfico MCUSUM cujos parâmetros sejam capazes de minimizar o número de falsos alarmes, ou seja, maximizar a capacidade de detectar mudanças reais. Além disso, alguns tópicos referentes a métodos de análise numérica com ênfase a um método de integração numérica robusto denominado Quadratura Gaussiana que neste trabalho é inserido ao Método de Equação Integral (MEI).

Alguns tópicos complementares cujo referencial teórico fundamenta a Aplicação da Metodologia (Capítulo 4) tais como normalidade, autocorrelação, métodos numéricos, etc, estão disponíveis nos Apêndices deste trabalho.

CAPÍTULO 3

METODOLOGIA

3.1 Introdução

Uma equação integral para determinar o ARL de um gráfico de controle multivariado de Somas Acumuladas (MCUSUM) quando o processo está sob controle é analiticamente derivada, e um método numérico para obter uma solução aproximada desta equação é proposto neste capítulo. Esta equação é utilizada para determinar de forma aproximada os parâmetros que otimizam o gráfico MCUSUM, tais como o intervalo de decisão h para vários valores de referência k , a dimensão de p características da qualidade mensurada, bem como o ARL mínimo.

Conforme revisão de literatura, vários autores têm apresentado métodos gráficos e tabelas para selecionar estes parâmetros para o gráfico MCUSUM. Em situações práticas, estes recursos atendem a um número muito reduzido de situações, sendo, portanto, muito limitada a sua aplicação.

3.2 Método de Equação Integral com Quadratura Gaussiana para otimização dos parâmetros do Gráfico de Controle MCUSUM

O método de equação integral, proposto neste trabalho, envolve a solução de uma equação integral, que substituída por um sistema de equações lineares algébricas, cuja resolução numérica via Quadratura Gaussiana fornece os valores aproximados de ARL do gráfico MCUSUM para um processo supostamente com distribuição normal multivariada. Estes valores de ARL são indispensáveis para avaliação de outros parâmetros tais como k e h , que envolvem a mensuração de índices de falsos alarmes e análise da sensibilidade do gráfico MCUSUM para monitorar processos em diferentes situações.

O recurso computacional para obter a solução aproximada de ARL consiste no desenvolvimento de um programa computacional como uma ferramenta sensível, que possa proporcionar ao usuário uma avaliação de desempenho do gráfico MCUSUM. Com este recurso computacional é possível determinar com segurança e em tempo real os valores dos parâmetros ótimos k e h associados ao ARL que devem ser selecionados para se obter níveis de qualidade aceitáveis.

Primeiramente, apresenta-se o método de integração para a solução numérica de ARL do gráfico CUSUM univariado, tomando como referência publicações de alguns pesquisadores, tais como Lucas (1976), Lucas e Crosier (1982), Gan (1991,1993) e Rao, Disney e Pignatiello (2001).

A derivação de uma equação integral tendo como referência a equação original de Page (1954) para otimização de ARL de um gráfico CUSUM Unilateral, via método de Quadratura Gaussiana, objeto de estudo de Rao, Disney e Pignatiello (2001), é apresentada neste trabalho por ser a mais recente.

Se X_i é uma variável contínua e aleatória a ser controlada e se $L(z)$ é o ARL esperado quando a soma acumulada inicia-se em $S_o = z$ com $z \in [0, h]$, pode-se mostrar que o ARL do gráfico CUSUM satisfaz a equação integral

$$L(z) = 1 + L(0)F(k - z) + \int_0^h L(x)f(x + k - z)dx \quad (3.1)$$

onde

$L(x)$ é o ARL quando a soma acumulada inicia-se em x ;

$L(z)$ é o ARL esperado quando a soma acumulada inicia-se em $S_o = z$ com $z \in [0, h]$;

$f(x)$ é a função densidade de probabilidade de uma distribuição normal $(\mu, 1)$;

$F(\cdot)$ é a função densidade de probabilidade cumulativa de uma distribuição normal $(\mu, 1)$;

Segundo estes autores, a função $L(x)$ apresenta uma solução única e contínua em $[0, h]$ para o ARL dessa equação, pois está sujeita à certas condições de regularidade na distribuição da função $F(\cdot)$ de amostra X_i no gráfico CUSUM em (3.1), tais como:

- a) A distribuição da função $F(\cdot)$ é contínua e diferenciável com uma primeira derivada $f(\cdot)$
- b) O projeto de parâmetros k e h , são tais que $F(k+h) < 1$.

Estas condições de regularidade são usualmente satisfeitas na prática. Por exemplo, o monitoramento da média de um processo via gráfico CUSUM cuja amostra estatística tem distribuição normal que satisfaz as condições de regularidade. Como na equação integral (3.1), $L(x)$ é contínua em $[0, h]$ e satisfaz as condições de regularidade a) e b) mencionadas, segue que $L(x)$ tem uma solução única e contínua para o ARL dessa equação.

A solução numérica desta equação integral para otimizar o ARL do gráfico CUSUM envolve a substituição da equação por uma Quadratura Gaussiana que resulta em

$$L(z) = 1 + L(0)F(k - z) + \sum_{i=1}^n w_i L(x_i) f(x_i + k - z) + \rho(z) \quad (3.2)$$

onde $\{x_i\}_{i=1}^n$, $\{w_i\}_{i=1}^n$, são, respectivamente, os nós e os pesos da quadratura gaussiana no intervalo $[0, h]$ e $\rho(z)$ é um erro muito pequeno e desconhecido devido à substituição da integral pela quadratura. Avaliando $L(\cdot)$ para $z = 0, z_1, z_2, \dots, z_n$,

$$\tilde{L}(0) = 1 + \tilde{L}(0)F(k) + \sum_{i=1}^n w_i \tilde{L}(x_i) f(x_i + k) \quad (3.3)$$

$$\tilde{L}(z_i) = \tilde{L}(0)F(k - z_i) + \sum_{i=1}^n w_i \tilde{L}(x_i) f(x_i + k - z_i) \quad (3.4)$$

onde $\tilde{L}(\cdot)$ é uma aproximação para $L(\cdot)$. A aproximação de $\tilde{L}(\cdot)$ é obtida a partir da resolução de um sistema de equações algébricas lineares e usando a equação

$$\tilde{L}(z) = 1 + \tilde{L}(0)F(k - z) + \sum_{i=1}^n w_i \tilde{L}(x_i) f(x_i + k - z), \quad \text{para } \forall z \in [0, h] \quad (3.5)$$

Resolvendo este sistema de equações obtém-se $\tilde{L}_1 = \tilde{L}_1(x_1), \tilde{L}_2 = \tilde{L}_2(x_2), \dots, \tilde{L}_n = \tilde{L}_n(x_n)$ e substituindo $\tilde{L}(x_i)$ em (3.3) e (3.5), determina-se $\tilde{L}(0)$ e $\tilde{L}(z)$, respectivamente.

A derivação analítica de uma equação integral utilizando integração numérica Quadratura Gaussiana proposta neste trabalho para o gráfico MCUSUM é um método alternativo ainda não existente na literatura para determinar o ARL deste tipo de gráfico de controle. Este procedimento segue uma extensão multivariada da equação integral utilizada para determinar o ARL de um gráfico de controle CUSUM univariado com uma adaptação da equação integral para determinar o ARL de um gráfico de controle MEWMA sugerida por Rigdon (1995).

Suponha que desejamos monitorar um processo repetitivo onde há p características de qualidade em cada unidade produzida. Sejam $\mathbf{X}_1, \mathbf{X}_2, \dots$ os vetores aleatórios da seqüência p que representam a saída do processo. Esses vetores aleatórios podem representar tanto as médias dos subgrupos quanto observações individuais. Suponhamos ainda, que $\mathbf{X}_1, \mathbf{X}_2, \dots$ são vetores aleatórios normais multivariados independentes e identicamente distribuídos com vetor de médias $\boldsymbol{\mu}$ e matriz de covariância Σ , onde $\boldsymbol{\mu}$ e Σ são supostamente conhecidos.

A equação integral analiticamente derivada neste trabalho para o gráfico MCUSUM está sujeita à suposição de que o vetor de médias sob controle é $\boldsymbol{\mu}_0 = \mathbf{0}$ e a matriz de covariância é a identidade I . Embora esta suposição possa parecer restritiva, na realidade não é. Para justificar a razão pela qual isto se comporta, é necessário os seguintes teoremas que são demonstrados no Apêndice B deste trabalho.

Teorema 1:

Se $\mathbf{X}_1, \mathbf{X}_2, \dots$ são variáveis aleatórias i.i.d $N_p(\mu, \Sigma)$, onde Σ é positiva definida. Se o gráfico MCUSUM com parâmetros k e h é aplicado para estas variáveis X 's, então o ARL sob controle para gráfico MCUSUM com parâmetros k e h aplicado para um processo com vetor de médias $\boldsymbol{\mu}_0 = \mathbf{0}$ e matriz de covariância \mathbf{I} . Isto é, portanto suficiente por determinar o ARL para o processo com vetor de médias $\mathbf{0}$ e matriz de covariâncias \mathbf{I} .

Demonstração: (conforme Apêndice B)

Teorema 2:

O ARL sob controle do gráfico de controle MCUSUM aplicado para o processo com vetor de médias $\boldsymbol{\mu}_0 = \mathbf{0}$ e matriz de covariância \mathbf{I} dependem tão somente do valor inicial $S_0 = z$ através do quadrado de sua magnitude $\delta = \mathbf{z}'\mathbf{z}$.

Demonstração: (conforme Apêndice B)

Teorema 3:

Supomos que o gráfico MCUSUM com parâmetro k aprimorado seja aplicado para um processo p dimensional com vetor de médias $\boldsymbol{\mu}_0 = \mathbf{0}$ e matriz de covariância \mathbf{I} . Se $L(\delta/h)$ denota o ARL dado para que a estatística MCUSUM inicial $S_0 = z$ satisfaça $\mathbf{S}'_0 \mathbf{S}_0 = \delta$ e dado que o intervalo de decisão é h . Então a função L satisfaz a equação integral

$$L(\delta/h) = 1 + L(0/h)F(x/\delta) + \int_0^h L(x)f(x | \mathbf{S}'_0 \mathbf{S}_0 = \delta)dx$$

onde $f(x | \mathbf{S}'_0 \mathbf{S}_0 = \delta)$ é a função densidade de probabilidade da distribuição Qui-quadrado não central com p graus de liberdade e parâmetro de não centralidade $\sum_{i=1}^p \lambda_i^2$.

Demonstração: (conforme Apêndice B)

3.3 Síntese de Capítulo

Neste capítulo uma equação integral é derivada analiticamente para a metodologia proposta cuja resolução numérica via Quadratura Gaussiana fornece os valores aproximados de ARL do gráfico de controle MCUSUM para uma determinada mudança do vetor de médias

do processo. Além disso, para aplicação da metodologia é sugerido no próximo capítulo deste trabalho, um programa computacional (amigável) em ambiente Matlab desenvolvido a partir desta equação integral que determina os principais parâmetros para otimizar o gráfico MCUSUM.

CAPÍTULO 4

APLICAÇÃO DA METODOLOGIA

4.1 Introdução

O capítulo 2 deste trabalho teve por finalidade fornecer a revisão de literatura necessária ao entendimento e a aplicação dos conceitos e fundamentos que envolvem a otimização dos principais parâmetros associados ao desempenho do gráfico de controle multivariado de Somas Acumuladas (MCUSUM). No capítulo 3 são aplicados os recursos matemáticos e estatísticos tais como os teoremas e as definições para derivar analiticamente uma equação integral cuja resolução numérica via Quadratura Gaussiana fornece o ARL do gráfico MCUSUM.

Agora, torna-se necessário, a aplicação da metodologia proposta neste trabalho a partir da execução de um algoritmo de CEP multivariado que contemple o Método de Equação Integral (MEI) com Quadratura Gaussiana como uma alternativa preliminar para otimizar o ARL do gráfico MCUSUM. Além disso, outros parâmetros tais k e h deste gráfico são essenciais, pois envolvem também a mensuração de índices de falsos alarmes e análise de sensibilidade para monitorar processos industriais em diferentes situações.

Inicialmente, são apresentados os valores dos parâmetros do gráfico MCUSUM determinado com o MEI para as combinações de ARL_0 de 200, 500 e 1000 com $p = 2, 3$ e 4 . Estes valores otimizados para os parâmetros são gerados por um programa computacional desenvolvido em ambiente Matlab.

A partir dos valores de ARL, k e h é desenvolvido um projeto estatístico ótimo do gráfico MCUSUM para observações individuais baseado nos valores desses parâmetros, sob a ótica de análise de regiões de máxima e mínima potência capaz de minimizar a probabilidade de falsos alarmes e maximizar a capacidade de detecção de mudanças reais.

Com o objetivo de aprimorar os conceitos e teorias até aqui apresentados, indispensáveis para auxiliar na interpretação de resultados preliminares é desenvolvido um exemplo de aplicação e, por fim, para comprovar a validade prática da metodologia proposta, é utilizado dados reais de um processo de usinagem que é monitorado com duas e três características de qualidade numa fundição na cidade de Joinville- SC.

4.2 Desenvolvimento Experimental da Metodologia

O desenvolvimento da metodologia aplicada neste trabalho compreende duas fases bem diferenciadas. A fase pela busca de informação e planejamento, e a fase puramente experimental. A descrição e a execução da metodologia nesta segunda fase denominada fase experimental é orientada conforme sugerido na seção 1.4 do capítulo 1 (figura 1.1). Esta fase inicia com a implementação do Método de Equação Integral (MEI). Uma vez derivada analiticamente a equação integral; esta equação é aplicada ao programa computacional desenvolvido em ambiente Matlab para determinar aproximadamente a solução ótima dos principais parâmetros. Os valores obtidos dessa solução ótima são utilizados para projetar estatisticamente o gráfico MCUSUM para observações individuais.

Como na prática o projeto estatístico do gráfico de controle MCUSUM em termos de desempenho de ARL está sujeito a um número muito elevado de combinações dos parâmetros k com h , é imprescindível a aplicação de planejamento de experimentos para poder obter-se qual combinação mais adequada desses parâmetros otimiza este tipo de gráfico. A variável resposta, ou seja, a combinação de k com h escolhida como melhor para ser aplicada será, logicamente aquele par de parâmetros (k,h) que associado ao ARL seja capaz de otimizar o gráfico MCUSUM.

O experimento computacional foi realizado em computadores do Laboratório de Informática da Universidade da Região de Joinville.

Uma análise estatística dos resultados obtidos do experimento computacional nos permite averiguar a influência das diferentes combinações do parâmetro k com h que associados ao ARL interferem significativamente na otimização do gráfico MCUSUM. Por último, uma análise de sensibilidade da solução ótima desses parâmetros para várias magnitudes de mudança do vetor de médias nos leva a extrair as conclusões finais da investigação. Nas próximas seções deste capítulo são abordados alguns aspectos relevantes da metodologia utilizada nesta segunda fase.

4.3 Programas Implementados

Para alcançar os objetivos propostos neste trabalho, definidos na seção 1.3 do capítulo 1, desenvolveu-se três programas computacionais em ambiente Matlab conforme a seguir:

- Programa Computacional 1: gera os fatores de quadratura (raízes e pesos) e aplica o MEI com Quadratura Gaussiana para otimizar os parâmetros ARL, k e h do gráfico MCUSUM.
- Programa Computacional 2: aproxima o limite superior de controle h do gráfico MCUSUM

com a aplicação de um método iterativo de aproximação conhecido em análise numérica como método da Secante.

- Programa Computacional 3: determina a matriz de covariância e sua inversa, a matriz de correlação e os coeficientes de assimetria e curtose utilizados por Mardia (1970, 1974) para verificar a normalidade multivariada. Além disso, determina a função perda multivariada de Taguchi e a função para a distância de Mahalanobis que é utilizada para calcular os pontos ótimos A e B que delimitam as regiões de máxima e mínima potência. Um outro programa computacional, considerado complementar é desenvolvido para realizar a análise de sensibilidade da solução ótima. Este programa aplica a interpolação polinomial de um conjunto de pontos (valores de ARL's) conhecidos quando o processo está sob controle para estimar os valores de ARL entre estes pontos para uma situação fora de controle.

O tutorial com as instruções operacionais destinadas a orientar os usuários de cada um destes três programas propostos encontra-se disponível no Apêndice K deste trabalho.

Uma vez desenvolvido o programa computacional correspondente ao que havia sido proposto para otimizar os parâmetros do gráfico MCUSUM; conclui-se que o planejamento desta 2ª fase experimental é considerado consistente para realizar o planejamento de experimentos e a verificação do experimento computacional correspondente. O tipo de planejamento de experimentos, realizado neste trabalho, foi um planejamento fatorial equilibrado 2^k ou balanceado replicado com k fatores onde a cada tratamento possível se experimenta um mesmo número de vezes.

4.4 Aplicação do MEI para determinar os parâmetros do gráfico de controle MCUSUM

Definidos os parâmetros ótimos de entrada para o MEI utilizou-se um algoritmo com o objetivo de determinar adequadamente os valores ótimos de ARL associados as melhores combinações dos parâmetros k com h capazes de minimizar o número de falsos alarmes. Como o resultado ótimo depende da sensibilidade desejada para o gráfico MCUSUM, diferentes valores de k foram combinados experimentalmente com o parâmetro h para otimizar o ARL. Estes valores obtidos experimentalmente são utilizados para a aproximação sistemática de projetar estatisticamente um gráfico MCUSUM baseado no ARL. O limite superior de controle h do gráfico MCUSUM neste trabalho é determinado para o ARL_o de 200, 500 e 1000 aplicando-se o MEI para o número de variáveis p=2, 3 e 4.

A avaliação de desempenho do gráfico MCUSUM proposta neste trabalho segue a metodologia proposta por Crosier (1988) que recomenda o valor de referência k ótimo deste

gráfico para detectar uma mudança no vetor de médias do processo seja obtido aproximadamente como a metade dessa mudança para todos os ARL's quando o processo está sob controle. Este valor ótimo para k pode minimizar o ARL e um deslocamento particular para um determinado ARL_o . Num intervalo o parâmetro ótimo do gráfico MCUSUM é determinado pelo tamanho (amplitude) de mudança do vetor de médias especificado onde k (ponto médio dessa amplitude de mudança) é considerado o ponto ótimo do deslocamento.

4.4.1 Resultados da aplicação do MEI para os parâmetros do gráfico MCUSUM

Os resultados obtidos com a aplicação do MEI para os principais parâmetros do gráfico MCUSUM conforme tabelas 4.1, 4.2, e 4.3. Estas tabelas mostram para vários tamanhos (amplitudes) de mudança, d as combinações ótimas dos parâmetros k com h associadas ao ARL mínimo ($ARL_{min.}$) para o ARL_o de 200, 500 e 1000 e número de variáveis $p=2, 3$ e 4. Estes valores que representam as soluções ótimas globais encontradas para diferentes tamanhos de mudança são calculados com o Programa Computacional 1 que utiliza o MEI com Quadratura Gaussiana para obter os parâmetros ótimos do gráfico de controle MCUSUM. Além disso, os valores obtidos destes principais parâmetros para projetar estatisticamente o gráfico MCUSUM via MEI são comparados conforme tabela 4.4 com os valores obtidos destes mesmos parâmetros obtidos via método de Cadeias de Markov (MCM).

Tabela 4.1 O parâmetro ótimo do gráfico, k para projetar o MCUSUM e os correspondentes limites de controle, h com o ARL mínimo ($ARL_{min.}$) para $p=2$, ARL sob controle (ARL_o) de 200, 500 e 1000, e vários tamanhos (amplitudes) de mudança, d.

p = 2		ARL _o = 200		ARL _o = 500		ARL _o = 1000	
d	k	h	ARL mín.	h	ARL mín.	h	ARL mín.
0,1	0,05	15,670	136,47	22,589	245,28	28,920	351,76
0,2	0,1	13,245	83,69	18,226	134,95	22,635	182,95
0,3	0,15	11,334	54,07	14,975	79,53	18,059	102,06
0,4	0,2	9,834	36,98	12,568	50,58	14,768	61,83
0,5	0,25	8,658	26,76	10,791	34,72	12,420	40,86
0,6	0,3	7,732	20,40	9,471	25,55	10,747	29,30
0,7	0,35	6,997	16,25	8,477	19,92	9,540	22,52
0,8	0,4	6,402	13,42	7,706	16,23	8,645	18,25
0,9	0,45	5,911	11,38	7,087	13,66	7,947	15,35
1	0,5	5,493	9,84	6,566	11,74	7,370	13,21
1,1	0,55	5,126	8,62	6,110	10,23	6,864	11,51
1,2	0,6	4,796	7,62	5,699	8,99	6,400	10,10
1,3	0,65	4,492	6,77	5,321	7,96	5,966	8,90
1,4	0,7	4,209	6,03	4,973	7,08	5,558	7,87
1,5	0,75	3,946	5,39	4,653	6,34	5,182	7,00
1,6	0,8	3,704	4,83	4,364	5,71	4,843	6,28
1,7	0,85	3,486	4,35	4,108	5,20	4,546	5,70
1,8	0,9	3,295	3,94	3,886	4,77	4,295	5,24
1,9	0,95	3,135	3,61	3,696	4,43	4,091	4,89
2	1,0	3,010	3,36	3,535	4,15	3,928	4,62
2,5	1,25	2,892	3,13	2,942	3,21	3,342	3,76
3	1,5	2,732	2,82	2,792	2,99	2,912	3,21

Tabela 4.2 O parâmetro ótimo do gráfico, k para projetar o MCUSUM e os correspondentes limites de controle, h com o ARL mínimo ($ARL_{\min.}$) para $p=3$, ARL sob controle (ARL_o) de 200, 500 e 1000, e vários tamanhos (amplitudes) de mudança, d.

p = 3		ARL _o = 200		ARL _o = 500		ARL _o = 1000	
d	k	h	ARL mín.	h	ARL mín.	h	ARL mín.
0,1	0,05	19,564	144,71	28,215	268,29	36,113	393,02
0,2	0,1	16,562	91,26	22,757	133,72	28,075	205,33
0,3	0,15	14,197	60,29	18,678	73,07	22,306	115,21
0,4	0,2	12,338	41,85	15,652	44,73	18,207	70,17
0,5	0,25	10,875	30,51	13,413	30,73	15,311	46,53
0,6	0,3	9,718	23,28	11,751	23,22	13,254	33,41
0,7	0,35	8,793	18,49	10,500	18,85	11,766	25,63
0,8	0,4	8,042	15,18	9,535	16,05	10,648	20,66
0,9	0,45	7,418	12,79	8,763	14,10	9,764	17,24
1	0,5	6,885	11,00	8,117	12,61	9,024	14,71
1,1	0,55	6,419	9,59	7,553	11,40	8,375	12,72
1,2	0,6	6,001	8,46	7,045	10,36	7,788	11,10
1,3	0,65	5,620	7,52	6,577	9,42	7,253	9,76
1,4	0,7	5,269	6,73	6,144	8,56	6,767	8,67
1,5	0,75	4,946	6,06	5,745	7,78	6,333	7,78
1,6	0,8	4,650	5,50	5,384	7,06	5,951	7,07
1,7	0,85	4,383	5,03	5,064	6,42	5,618	6,49
1,8	0,9	4,148	4,64	4,787	5,85	5,326	6,03
1,9	0,95	3,945	4,32	4,555	5,36	5,059	5,64
2	1,0	3,777	4,07	4,364	4,96	4,797	5,28
2,5	1,25	3,392	3,54	3,782	3,67	3,952	4,48
3	1,5	2,982	3,04	3,492	2,99	3,710	4,07

Tabela 4.3 O parâmetro ótimo do gráfico, k para projetar o MCUSUM e os correspondentes limites de controle, h com o ARL mínimo ($ARL_{\min.}$) para $p=4$, ARL sob controle (ARL_o) de 200, 500 e 1000, e vários tamanhos (amplitude) de mudança, d.

p = 4		ARL _o = 200		ARL _o = 500		ARL _o = 1000	
d	k	h	ARL mín.	h	ARL mín.	h	ARL mín.
0,1	0,05	22,958	150,15	33,190	285,13	42,551	426,10
0,2	0,1	19,489	96,85	26,842	163,39	33,172	226,26
0,3	0,15	16,756	65,25	22,058	99,02	26,352	128,03
0,4	0,2	14,604	45,98	18,483	63,93	21,455	78,11
0,5	0,25	12,903	33,86	15,824	44,08	17,971	51,66
0,6	0,3	11,548	25,97	13,845	32,36	15,495	36,95
0,7	0,35	10,457	20,65	12,357	25,09	13,717	28,28
0,8	0,4	9,561	16,92	11,213	20,34	12,404	22,83
0,9	0,45	8,812	14,20	10,304	17,04	11,386	19,13
1	0,5	8,171	12,16	9,549	14,61	10,549	16,41
1,1	0,55	7,611	10,57	8,893	12,70	9,817	14,27
1,2	0,6	7,112	9,30	8,300	11,14	9,149	12,50
1,3	0,65	6,662	8,27	7,750	9,83	8,525	10,99
1,4	0,7	6,253	7,42	7,238	8,71	7,941	9,72
1,5	0,75	5,880	6,71	6,761	7,77	7,402	8,64
1,6	0,8	5,541	6,12	6,326	6,98	6,917	7,76
1,7	0,85	5,235	5,64	5,941	6,33	6,494	7,06
1,8	0,9	4,961	5,23	5,612	5,82	6,137	6,51
1,9	0,95	4,717	4,90	5,345	5,43	5,846	6,10
2	1,0	4,501	4,62	5,143	5,16	5,612	5,78
2,5	1,25	3,722	3,77	4,820	4,74	4,712	4,73
3	1,5	3,262	3,37	3,722	3,55	4,032	4,09

Os resultados obtidos via MEI para os principais parâmetros do gráfico MCUSUM (tabelas 4.1, 4.2 e 4.3) indicam que os valores para as melhores combinações dos parâmetros k com h

que otimizam ARL's mínimos para um mesmo ARL_o de 200, 500 e 1000 são bastante similares aos resultados obtidos via Método de Cadeias de Markov (Lee e Khoo, 2006) para os mesmos parâmetros e nas mesmas condições neste trabalho conforme tabelas 1, 2 e 3 dos Anexos B, C e D, respectivamente.

Um comparativo entre os resultados obtidos via MCM x MEI para o parâmetro ótimo do gráfico, k e os correspondentes limites de controle, h associados com o ARL mínimo ($ARL_{mín.}$) obtido para $p = 2, 3$ e 4 e para o mesmo $ARL_o=200$ é mostrado conforme tabela 4.4 para vários tamanhos (amplitudes) de mudança, d .

Tabela 4.4 Comparativo entre valores dos parâmetros do MCUSUM obtidos via MCM x MEI

d	k	ARL _o =200	p=2		p=3		p=4	
			MCM	MEI	MCM	MEI	MCM	MEI
0,2	0,1	h	13,054	13,245	16,386	16,562	19,355	19,489
		ARL(mín.)	79,12	83,69	87,61	91,26	94,44	96,85
0,4	0,2	h	9,755	9,834	12,261	12,338	14,542	14,604
		ARL(mín.)	35,86	36,98	40,73	41,85	45,00	45,98
0,6	0,3	h	7,776	7,732	9,761	9,718	11,583	11,548
		ARL(mín.)	20,82	20,40	23,71	23,28	26,34	25,97
0,8	0,4	h	6,449	6,402	8,087	8,042	9,597	9,561
		ARL(mín.)	13,78	13,42	15,66	15,18	17,41	16,92
1	0,5	h	5,491	5,493	6,883	6,885	8,169	8,171
		ARL(mín.)	9,90	9,84	11,20	11,00	12,45	12,16
1,2	0,6	h	4,763	4,796	5,970	6,001	7,089	7,112
		ARL(mín.)	7,49	7,62	8,46	8,46	9,40	9,30
1,4	0,7	h	4,188	4,209	5,250	5,269	6,239	6,253
		ARL(mín.)	5,90	6,03	6,66	6,73	7,38	7,42
1,6	0,8	h	3,722	3,704	4,667	4,650	5,551	5,541
		ARL(mín.)	4,80	4,83	5,40	5,50	5,98	6,12
1,8	1,9	h	3,334	3,295	4,183	4,148	4,980	4,961
		ARL(mín.)	3,99	3,94	4,48	4,64	4,96	5,23
2	1	h	3,008	3,010	3,775	3,777	4,499	4,501
		ARL(mín.)	3,39	3,36	3,79	4,07	4,19	4,62

Como se pode verificar, a combinação ótima (k,h) associada ao ARL mínimo obtido via MEI para $p=2, 3$ e 4 e com mesmo $ARL_o=200$ cujos valores conforme tabela 4.4 apresentam uma certa similaridade com os valores desses parâmetros obtidos via MCM nas mesmas condições.

Uma análise de sensibilidade do gráfico de controle MCUSUM é efetuada para $p=2$ variáveis e $n=1$ com uma aproximação de ARL_o's selecionando-se um conjunto de valores de

referência $k = 0,2, 0,35, 0,50, 0,65$ e $0,80$. Usando aproximação via Método de Equação Integral obtém-se os limites de controle deste gráfico $9,834, 6,997, 5,493, 4,492$ e $3,704$. Uma vez determinado estes valores para $ARL_0=200$ quando o processo está sob controle, os ARL 's para vários tamanhos (amplitudes) de mudança, d quando o processo está fora de controle são estimados neste trabalho via interpolação polinomial cujos valores conforme tabela 4.5.

Tabela 4.5 Análise de Sensibilidade: Valores de k e h para o ARL de 200 com $p=2$ e $n=1$ quando o processo está sob controle e os valores estimados para o ARL quando o processo está fora de controle.

d	$k = 0,20$ $h = 9,834$	$k = 0,35$ $h = 6,997$	$k = 0,50$ $h = 5,493$	$k = 0,65$ $h = 4,492$	$k = 0,80$ $h = 3,704$
0,10	138,38	138,11	136,47	138,24	138,52
0,20	87,28	86,63	85,36	87,40	89,42
0,30	56,08	55,74	54,96	56,15	57,07
0,40	37,86	37,72	37,25	37,87	38,07
0,50	27,13	27,08	26,76	27,11	27,06
0,60	20,58	20,56	20,32	20,55	20,47
0,70	16,38	16,37	16,18	16,36	16,30
0,80	13,55	13,53	13,37	13,53	13,50
0,90	11,52	11,50	11,36	11,50	11,50
1,00	9,98	9,96	9,84	9,97	9,99
1,10	8,76	8,74	8,64	8,75	8,78
1,20	7,75	7,73	7,64	7,74	7,77
1,30	6,88	6,86	6,79	6,88	6,90
1,40	6,12	6,11	6,04	6,12	6,14
1,50	5,46	5,45	5,39	5,46	5,47
1,60	4,88	4,87	4,82	4,88	4,89
1,70	4,39	4,38	4,33	4,38	4,39
1,80	3,98	3,97	3,93	3,97	3,97
1,90	3,65	3,64	3,60	3,64	3,65
2,00	3,41	3,40	3,36	3,40	3,41

Conforme tabela 4.5, observa-se que o gráfico ótimo é tão sensível quanto os demais gráficos MCUSUM para tamanhos de mudança, d entre $d=0,8$ até $d=2,0$ considerados suficientemente importantes para serem rapidamente detectadas. Os valores obtidos indicam que em $d=1$ a combinação preferida (k,h) capaz de produzir um gráfico MCUSUM ótimo é $k=0,50$ e $h=5,493$.

Uma representação gráfica (Curvas de ARL , em escala logarítmica) ilustra para vários tamanhos de mudança do vetor de médias os resultados obtidos para os principais parâmetros do gráfico MCUSUM via MEI (tabelas 4.1, 4.2 e 4.3) conforme figura 4.1.

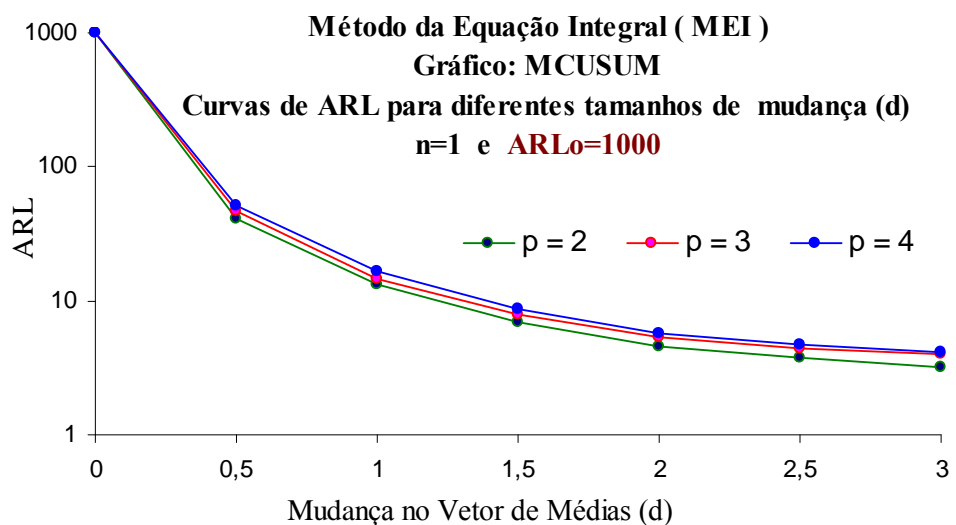
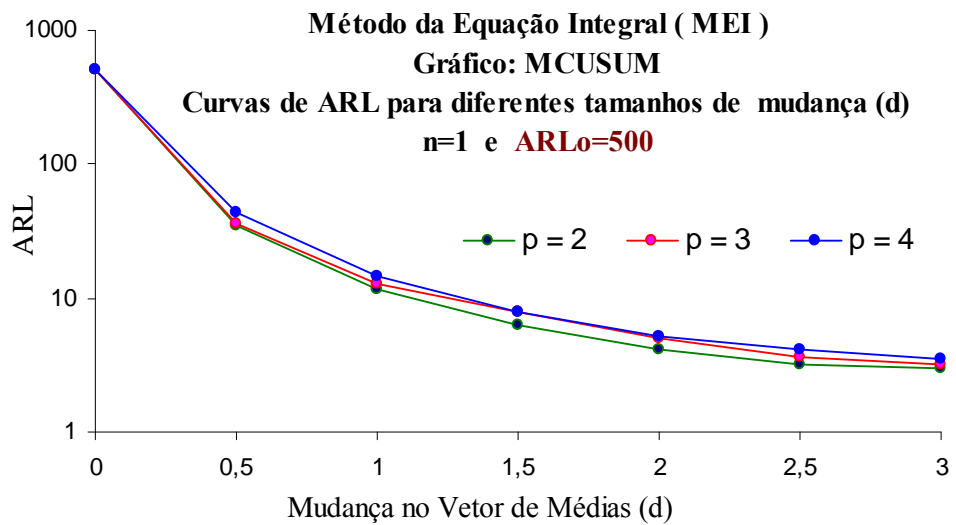
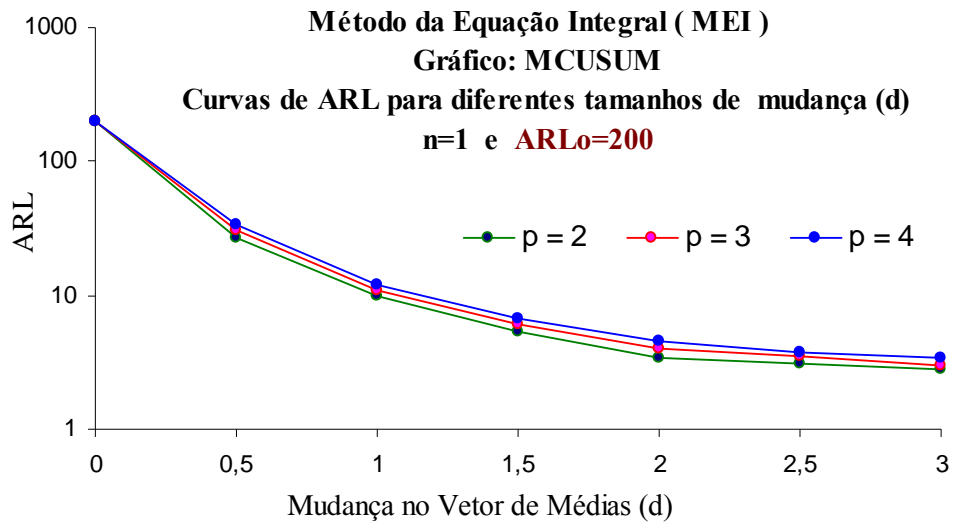


Figura 4.1 Curvas de ARL do gráfico MCUSUM para $n=1$ e ARL_0 de 200, 500 e 1000

Conforme figura 4.1, o valor de ARL diminui conforme a amplitude da mudança a ser detectada aumenta. Além disso, conclui-se que o gráfico de controle MCUSUM bivariado é mais sensível para amplitudes de mudança do vetor de médias, $d \geq 1$ e que a medida que aumenta o número de variáveis esta sensibilidade é reduzida.

O desempenho de ARL (em escala logarítmica) para vários tamanhos (amplitudes) de mudança do vetor de médias dos principais gráficos de controle multivariados T^2 de Hotelling, MCUSUM e MEWMA para $p=2$, $n=1$ e $ARL_0=200$ conforme figura 4.2.

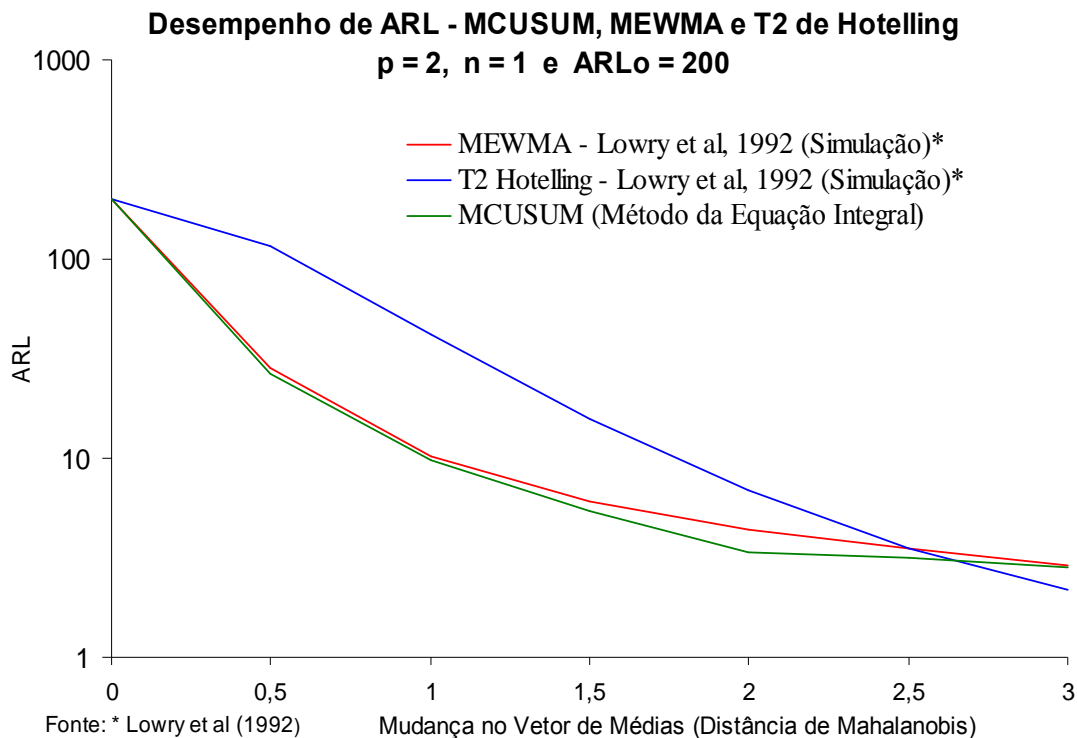


Figura 4.2 Desempenho de ARL dos gráficos MCUSUM , MEWMA e T² de Hotelling

Como se pode observar, o gráfico T² de Hotelling é mais sensível para grandes mudanças no vetor de médias do processo. No entanto, para pequenas mudanças os gráficos MCUSUM e MEWMA são mais sensíveis. A diferença existente entre o desempenho destes dois gráficos para esta situação (quando $p=2$ variáveis) é considerada significativa conforme figura 4.2 para tamanhos de mudança $d \geq 1$, onde o gráfico MCUSUM é mais sensível.

O desempenho de ARL do gráfico MCUSUM (Crosier,1988) para vários tamanhos de mudança do vetor de médias decorrente da aplicação do Método de Simulação (SIM), Método de Cadeias de Markov e o presente Método de Equação Integral (MEI) proposto neste trabalho conforme figura 4.3.

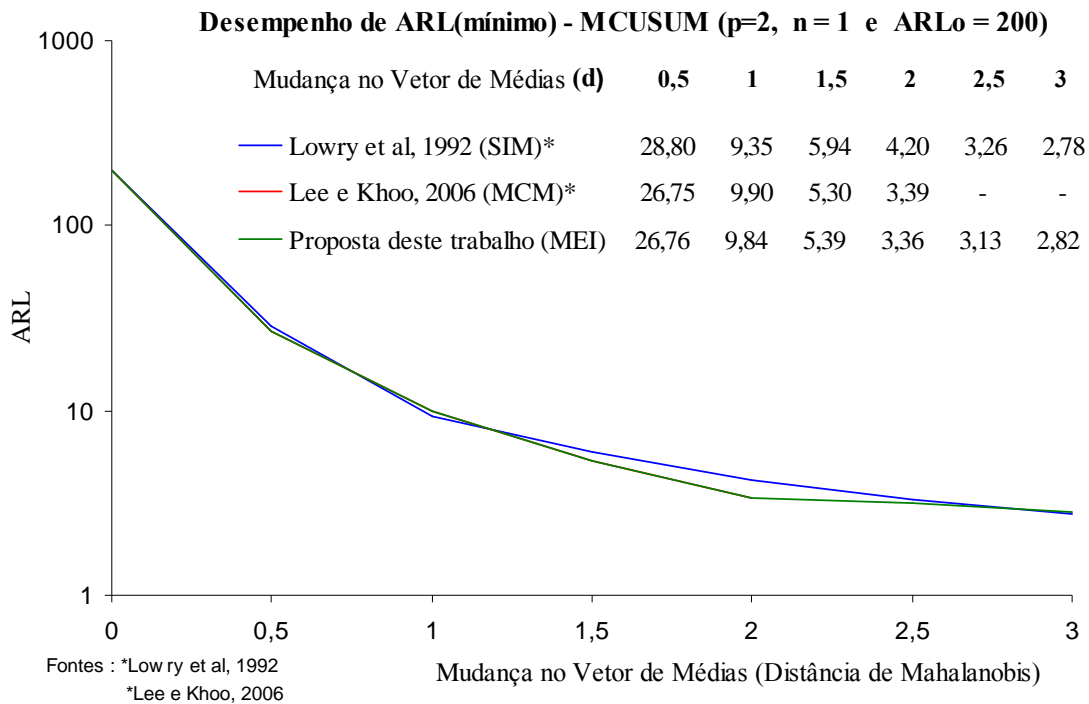


Figura 4.3 Desempenho de ARL do gráfico MCUSUM com a aplicação de SIM, MCM e MEI
Conforme figura 4.3, os métodos numéricos MEI e MCM fornecem melhores resultados para aproximação de ARL que os tradicionais métodos de simulação. Esta diferença considerada significativa ocorre entre os tamanhos de mudança, d=1 e d=2,5.

4.5 Projeto estatístico ótimo para um gráfico MCUSUM baseado em ARL e sob a ótica de regiões de máxima e mínima potência

Um procedimento simples baseado em ARL pode ser desenvolvido para o projeto estatístico ótimo de um gráfico MCUSUM. Gan (1991,1992) adaptou o procedimento do projeto ótimo do gráfico EWMA univariado de Crowder (1989) que contém quatro etapas a um gráfico CUSUM univariado baseado em ARL e MRL. Lee e Khoo (2006) propõem um procedimento também com quatro etapas como uma extensão multivariada do gráfico CUSUM univariado de Gan (1991).

Um procedimento simples baseado em ARL pode ser desenvolvido para o projeto estatístico ótimo de um gráfico MCUSUM. Gan (1991,1992) adaptou o procedimento que contém quatro etapas do projeto ótimo de um gráfico EWMA univariado de Crowder (1989) a um gráfico CUSUM univariado baseado em ARL e MRL. Lee e Khoo (2006) propõem um

procedimento também com quatro etapas como uma extensão multivariada para o projeto ótimo do gráfico CUSUM univariado de Gan (1991).

Neste trabalho aplica-se o procedimento para projetar estatisticamente o gráfico MCUSUM ótimo sugerido por Lee e Khoo (2006), porém com a proposta de adicionar ao algoritmo de procedimentos uma etapa extra que contemple a otimização de parâmetros suficiente para maximizar a capacidade de detecção de mudanças reais deste gráfico. Esta etapa extra se constitui como uma alternativa de avaliação prévia para otimizar os parâmetros do gráfico MCUSUM baseado em ARL e sob a ótica de regiões do gráfico que apresentam uma elevada potência. Isso, é imprescindível para detectar no processo mudanças consideradas importantes. Com esta alternativa, torna-se possível detectar mudanças significativas, e que, simultaneamente, apresente uma probabilidade de falso alarme realmente baixa quando se produzem mudanças cuja detecção não seja de interesse prático. Este procedimento alternativo com a inclusão desta nova etapa envolve a análise de regiões de máxima e mínima potência e se constitui numa extensão multivariada do projeto estatístico de gráficos de controle de qualidade univariado proposto por Woodall (1985). Para isso, propõe-se neste trabalho a função perda multivariada de Taguchi (Teeravaraprug and Cho, 2002) para determinar utilizando a ferramenta do solver do MS-Excel os valores dos pontos ótimos A e B que delimitam as regiões de máxima e mínima potência do gráfico MCUSUM.

Diante destas considerações, propõe-se um algoritmo cujos procedimentos incluem esta etapa complementar que é algo ainda não disponível na literatura para o gráfico de controle MCUSUM. Portanto, as etapas para o desenvolvimento do projeto estatístico para otimizar os parâmetros do gráfico MCUSUM propostas neste trabalho conforme a seguir:

- Etapa 1: Escolher o menor ARL_0 aceitável.
- Etapa 2: Decidir o menor deslocamento (tamanho de mudança) aceitável, d no vetor de médias do processo que é extremamente importante ser detectado rapidamente. Para isso, é fundamental escolher corretamente o parâmetro ótimo do gráfico, k capaz de produzir o ARL mínimo para o deslocamento do vetor de médias, d , baseado no ARL_0 especificado na etapa 1.
- Etapa 3 (**Proposta**): Aplicar o método baseado na função perda multivariada de Taguchi que neste trabalho é utilizada para determinar os pontos ótimos A e B que delimitam as regiões de máxima e mínima potência. Com isso, é possível determinar a região sob controle, $d \leq A$ e a região fora de controle, $d \geq B$ onde devemos detectar o mais rápido possível uma mudança. Além disso, para quantificar a distância (desvio) entre o vetor de médias das características da

qualidade e o vetor de valores nominais dessas características utiliza-se a Distância de Mahalanobis, d onde o interesse neste trabalho é sem dúvida minimizar esta distância, d implementando um modelo matemático para obter o ponto ótimo, ou seja, o ponto que tenha um deslocamento (distância) mínimo em relação ao vetor de valores nominais.

- Etapa 4: Conhecido o valor ótimo de k , determinar o limite de controle h , de modo que o gráfico MCUSUM produza o ARL_0 especificado na etapa 1.
- Etapa 5 Análise de sensibilidade de desempenho para comparar o ARL fora de controle para a combinação ótima de k e de h para outras escolhas de k e de h para produzir o mesmo ARL_0 . A combinação ótima de k e de h capaz de produzir o desempenho total, ou seja, o mais desejável possível em termos de ARL são selecionados para a implementação.

A figura 4.4 ilustra de forma reduzida as etapas para o desenvolvimento deste projeto estatístico aplicado ao gráfico MCUSUM.

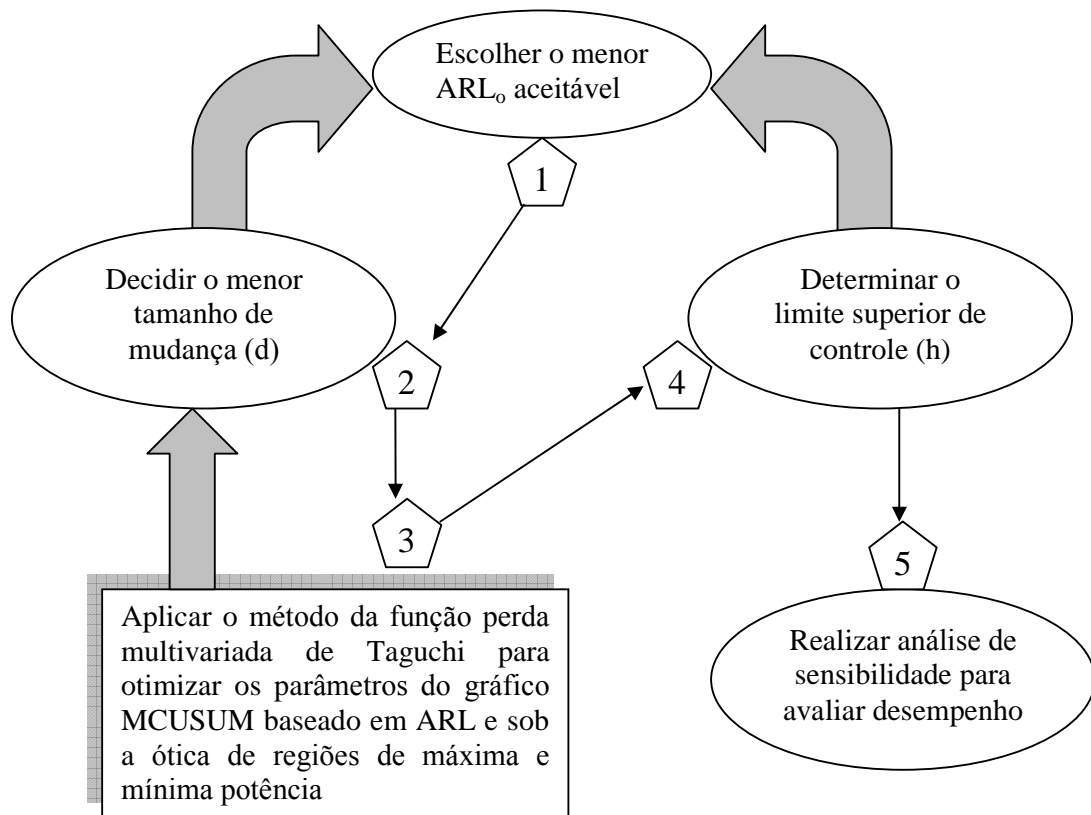


Figura 4.4 Resumo das etapas para o desenvolvimento do projeto estatístico do MCUSUM

4.5.1 Exemplo de aplicação para o projeto estatístico ótimo do gráfico MCUSUM baseado em ARL e sob ótica de regiões de máxima e mínima potência

Num determinado processo de usinagem de eixos, o diâmetro interno e externo são as características da qualidade de interesse monitoradas estatisticamente pelo gráfico de controle MCUSUM. Os valores nominais (T) para estas características da qualidade, com matriz de covariâncias Σ e matriz de coeficientes de perda K deste processo são dados por.

$$T = \begin{pmatrix} 8 \\ 20 \end{pmatrix} \quad \Sigma = \begin{pmatrix} 0,49 & 0 \\ 0 & 0,25 \end{pmatrix} \quad \text{e} \quad K = \begin{pmatrix} 3 & 0,5 \\ 0,5 & 2 \end{pmatrix}$$

Supomos que a política de recursos financeiros para a produtividade desta empresa considera um custo ou perda depreciável de 0,1 u.m e um custo ou perda inadmissível de 2 u.m. (unidades monetárias). Como o exemplo trata da mudança de valor do vetor de médias de um processo segundo a ótica de análise de regiões de máxima e mínima potência, determinam-se os valores dos pontos A e B que delimitam estas regiões conforme figura 4.5.

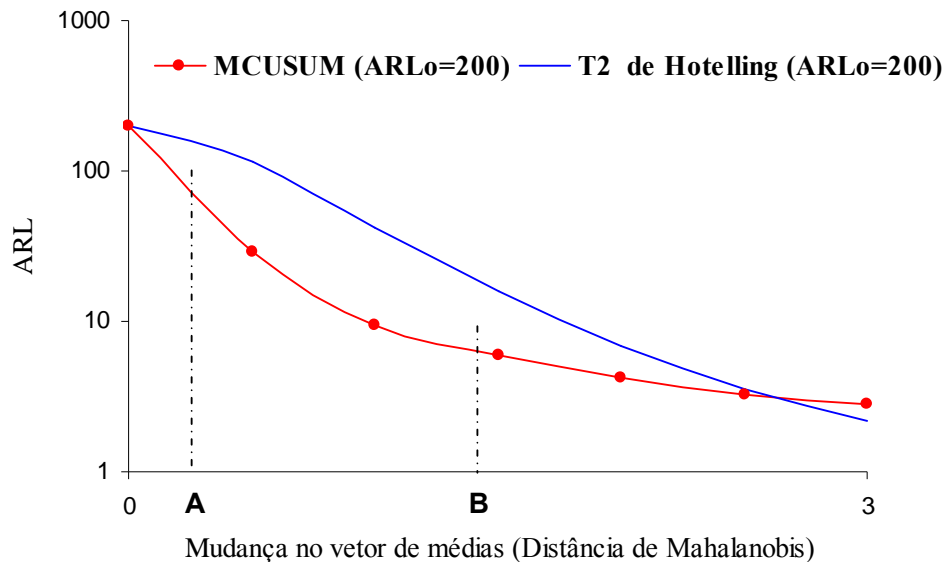


Figura 4.5 Regiões de máxima e mínima potência de um gráfico de controle

A região sob controle, $[0,A]$ é aquela região onde não há interesse em detectar alguma mudança, que se estende de $d=0$ até $d=A$ (perda depreciável). A região, (A,B) compreendida entre $d=A$ até $d=B$ é a região onde é indiferente detectar ou não uma mudança no processo e a região fora de controle, $[B,\infty)$ corresponde a região em que deseja-se detectar alguma mudança em $d=B$, a partir do qual se requer a máxima eficácia de detecção, isto é, um ARL mínimo.

A função perda multivariada de Taguchi segundo Kapur e Cho (1996) pode ser

expressa como

$$L(y_1, y_2, \dots, y_p) = \sum_{i=1}^p \sum_{j=1}^i k_{ij} (y_i - T_i)(y_j - T_j) \quad (4.1)$$

onde $L(y_1, y_2, \dots, y_p)$ é a perda transmitida à sociedade e $k_{ij} = k_{ji}$ é a constante de proporcionalidade que especifica os custos dos desvios entre as características da qualidade e seus valores nominais. Por exemplo, k_{11} é o custo individual por unidade ocasionado pela característica da qualidade y_1 por desviar-se de T_1 e k_{12} o custo por unidade adicional ocorrido quando as características y_1 e y_2 estão simultaneamente fora de seus valores nominais T_1 e T_2 . A equação (4.1) pode ser escrita na forma matricial como $L(y) = (y - T)'K(y - T)$.

A função perda multivariada $L(y_1, y_2)$ para os dados deste exemplo é:

$$L(y_1, y_2) = k_{11}(y_1 - T_1)^2 + k_{22}(y_2 - T_2)^2 + 2k_{12}(y_1 - T_1)(y_2 - T_2)$$

$$L(y_1, y_2) = 3(y_1 - 8)^2 + 2(y_2 - 20)^2 + (y_1 - 8)(y_2 - 20)$$

$$L(y_1, y_2) = 3y_1^2 + 2y_2^2 - 68y_1 - 88y_2 + y_1y_2 + 1152$$

Na figura 4.6, observa-se a superfície gerada pela função perda bivariada com um mínimo nos valores nominais $T_1=8$ e $T_2=20$ e as linhas de contorno ou curvas de nível.

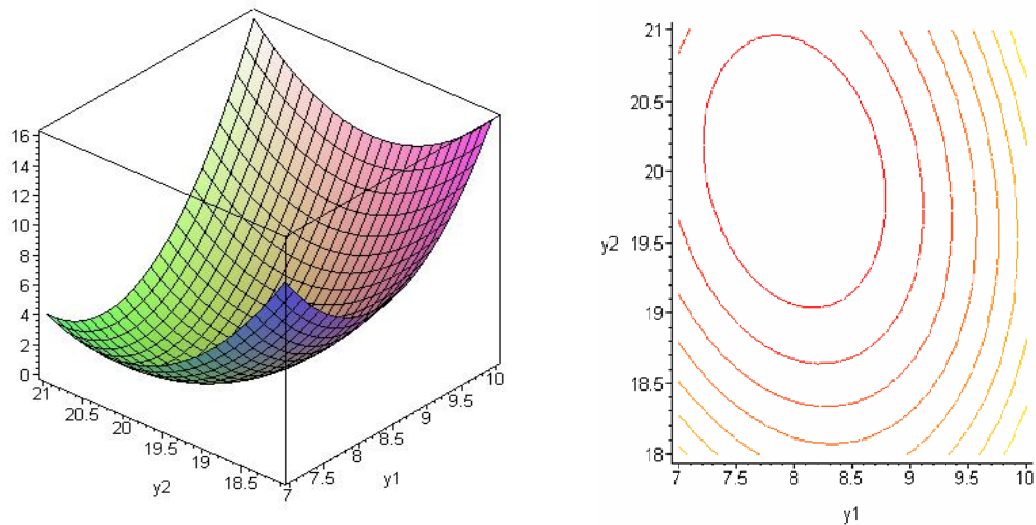


Figura 4.6 Função perda bivariada para as características da qualidade do exemplo de aplicação

Agora, determina-se a distância de Mahalanobis (medida de distância entre dois vetores), d que neste exemplo é a distância entre o vetor de valores obtido $\mathbf{Y}=(y_1, y_2)$ e o vetor de valores nominais \mathbf{T} .

$$d = \sqrt{(\mathbf{Y} - \mathbf{T})' \Sigma^{-1} (\mathbf{Y} - \mathbf{T})} \quad (4.2)$$

Esta é portanto, a forma utilizada para quantificar o desvio ou a descentralização produzida sobre o valor nominal, ou seja, o desvio (distância) entre o vetor de valores das características da qualidade e o vetor de valores nominais. Denomina-se por exemplo, como d_A o desvio entre o vetor de valores das características da qualidade observadas no eixo usinado em A , Y e o vetor de valores nominais (alvo) T . O interesse é sem dúvida obter o ponto que tenha um desvio mínimo em relação a T . Neste caso, trata-se de minimizar a distância de Mahalanobis, d daqueles pontos que se encontram sobre a linha de contorno (curvas de nível).

Um modelo matemático em ambiente MS-Excel (usando a ferramenta Solver) é aplicado para determinar o ponto que minimize a distância d , ou seja, o menor desvio em relação ao valor nominal T .

Para minimizar a Distância de Mahalanobis, d daqueles pontos que se encontram sobre as linhas de contorno, aplica-se a equação 4.2. Assim, para os dados do exemplo temos:

$$d(x, y) = \sqrt{(y_1 - 8) \begin{pmatrix} 2,041 & 0 \\ 0 & 4 \end{pmatrix} (y_2 - 20)}$$

$$d(x, y) = (2,041y_1^2 - 32,656y_1 + 1730,624 + 4y_2^2 - 160y_2)^{\frac{1}{2}}$$

A região fora de controle $[B, \infty)$ em que há interesse em detectar mudanças no vetor de médias do processo em $d=B$ (distância associada a perda considerada como inadequada, neste exemplo, $d = 2,0$). O procedimento utilizado para determinar o valor do ponto B a seguir:

O procedimento de minimização: Ponto B (Resolução do Modelo Matemático):

$$\text{Min } d(x, y) = (2,041y_1^2 - 32,656y_1 + 1730,624 + 4y_2^2 - 160y_2)^{\frac{1}{2}}$$

Dado:

$$3(y_1 - 8)^2 + 2(y_2 - 20)^2 + (y_1 - 8)(y_2 - 20) = 2,0$$

$$y_1 \geq 0, \quad y_2 \geq 0,$$

$$\text{Minimizar } (d, x, y) = \begin{pmatrix} 7,204 \\ 19,901 \end{pmatrix}$$

Resolução de d :

$$x = 7,204$$

$$y = 19,901$$

$$d(x, y) = (2,041y_1^2 - 32,656y_1 + 1730,624 + 4y_2^2 - 160y_2)^{\frac{1}{2}}$$

$$d(x, y) = 1,15 \text{ (ponto B)}$$

Desta forma, encontra-se o desvio mínimo da região fora de controle a partir do qual é necessário detectar qualquer mudança no vetor de médias que é $d=1,15$.

No entanto, a região sob controle $[0,A]$ em que não há interesse em detectar alguma mudança no vetor de médias do processo compreendida entre $d=0$ até $d=A$ (distância associada à perda considerada depreciable, neste exemplo, $d=0,1$) cujo procedimento utilizado para determinar o valor do ponto A conforme a seguir:

O procedimento de minimização: Ponto A (Resolução do Modelo Matemático):

$$\text{Min } d(x, y) = (2,041y_1^2 - 32,656y_1 + 1730,624 + 4y_2^2 - 160y_2)^{\frac{1}{2}}$$

Dado:

$$3(y_1 - 8)^2 + 2(y_2 - 20)^2 + (y_1 - 8)(y_2 - 20) = 0,1$$

$$y_1 \geq 0, \quad y_2 \geq 0,$$

$$\text{Minimizar } (d, x, y) = \begin{pmatrix} 8,178 \\ 20,022 \end{pmatrix}$$

Resolução de d :

$$x = 8,178$$

$$y = 20,022$$

$$d(x, y) = (2,041y_1^2 - 32,656y_1 + 1730,624 + 4y_2^2 - 160y_2)^{\frac{1}{2}}$$

$$d(x, y) = 0,26 \text{ (ponto A)}$$

Desta forma, encontra-se o desvio máximo da região sob controle que é $d=0,26$.

Logo os pontos que delimitam as regiões de máxima e mínima potência do gráfico de controle MCUSUM para este exemplo são $A=0,26$ e $B=1,15$, conforme figura 4.7.

Os resultados obtidos com a resolução deste modelo matemático desenvolvido em ambiente MS-Excel com a aplicação da ferramenta “solver” para otimizar os pontos de máxima e mínima potência para os dados deste exemplo de aplicação tais como os relatórios de respostas, de sensibilidade e de limites conforme Apêndice I deste trabalho.

A figura 4.7 ilustra os valores dos pontos A e B que delimitam as regiões de máxima e mínima potência determinados com o modelo matemático proposto para os dados deste exemplo.

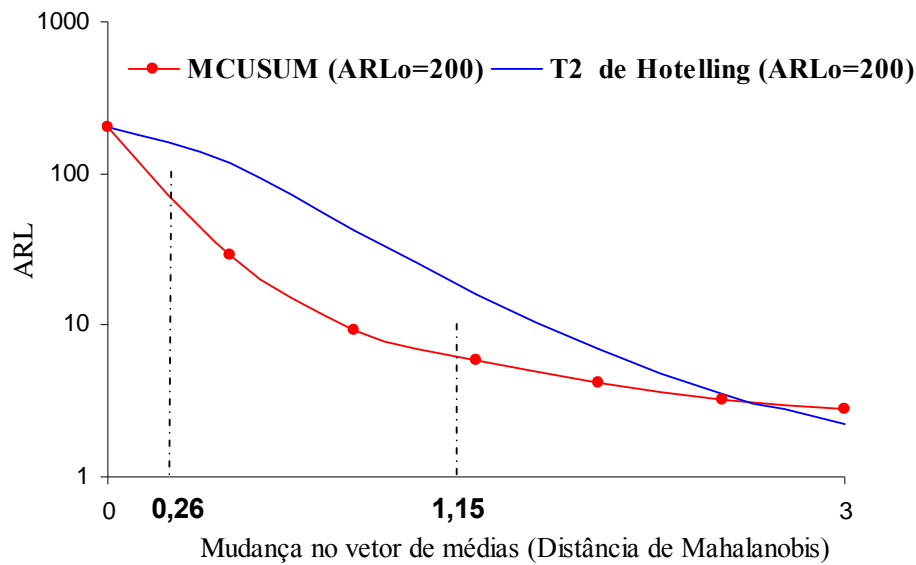


Figura 4.7 Regiões de máxima e mínima potência do MCUSUM do exemplo de aplicação

Conforme figura 4.7, observa-se que o gráfico MCUSUM apresenta uma sensibilidade considerável em relação ao gráfico T^2 de Hotelling para tamanhos de mudança, d do vetor de médias entre $d=0,26$ e $d=1,15$. No entanto, para tamanhos de mudança, $d > 2,5$ observa-se que o gráfico T^2 de Hotelling é mais sensível que o gráfico MCUSUM. Portanto, os pontos A e B que delimitam as regiões de máxima e mínima potência deste gráfico assumem valores maiores que 2,5.

Conhecidos os pontos ótimos A e B do gráfico que delimitam as regiões de máxima e mínima potência do gráfico conforme etapa 3 proposta para o algoritmo de procedimentos do projeto estatístico ótimo do gráfico MCUSUM. Neste trabalho, conclui-se que para este exemplo o tamanho (amplitude) da mudança desejado é $d=1,15$ ($d=B$). Portanto, o interesse neste caso, deve estar focado em detectar mudanças de vetor de médias do processo para $d \geq 1,15$ ($d_B \geq 1,15$). Selecionando os valores ótimos para os parâmetros do gráfico MCUSUM obtidos via Método de Equação Integral com Quadratura Gaussiana proposto neste trabalho, conforme tabelas 4.1, 4.2 e 4.3, nota-se que não existe o valor tabelado para o tamanho (amplitude) de mudança $d=1,15$. No entanto, conforme proposta deste trabalho estes valores para os parâmetros que otimizam o gráfico MCUSUM podem ser facilmente obtidos utilizando-se o Programa Computacional 1 conforme tutorial apresentado no Apêndice K. Então, para os dados deste exemplo de aplicação se desejarmos um $ARL_0 = 200$ para estas duas características da qualidade utilizando uma aproximação com 53 pontos de quadratura para o gráfico MCUSUM com tamanho de amostra $n=1$, obtém-se os seguintes parâmetros

para o $ARLo=200$, $p=2$, $d=1,15$. Os dados de entrada na caixa de diálogo para a execução do Programa Computacional 1 conforme figura 4.8.

Figura 4.8 Caixa de diálogo com os dados de entrada do exemplo de aplicação

Resultados (de saída) gerados após a execução do programa para os dados deste exemplo de aplicação conforme a seguir :

>> calculodearl

QUADRATURA GAUSSIANA

GRAFICO DE CONTROLE MULTIVARIADO DE SOMA ACUMULADA (MCUSUM)

UNIVERSIDADE FEDERAL DE SANTA CATARINA

Custodio da Cunha Alves custodio.alves@univille.br

***** RESULTADOS PARA 2 variaveis e $ARLo = 200$ *****

k	h	integral	ARL
0.575	4.957181329240516	10.539966470625162	8.098000749739537

FIM!! Se deseja parar tecle q:

Conforme resultados obtidos a partir da execução do Programa Computacional 1 para os dados deste exemplo de aplicação, conclui-se que a melhor combinação de parâmetros k com h capaz de produzir o $ARL_{\text{mínimo}} = 8,098$ sob a ótica de regiões de máxima e mínima potência é $k=0,575$ e $h=4,957$.

4.6 Aplicação da metodologia proposta utilizando dados reais de um processo de usinagem

A metodologia proposta neste trabalho é aplicada a um conjunto de dados reais de um processo de usinagem da Fundação Tupy Ltda localizada na cidade de Joinville-SC. Estes mesmos dados foram autorizados para Soares (2006) comprovar a validade prática de um modelo proposto em sua em sua dissertação de mestrado que inclui o cálculo de índice de capacidade multivariado em uma linha de usinagem.

4.6.1 Apresentação da empresa

A Fundação Tupy Ltda com sede própria e principal parque fabril na cidade de Joinville foi fundada em 1938 cuja produção inicial na época se restringia a fabricar tão somente artefatos de ferro utilizando os conhecimentos rudimentares de fundição. No entanto, ao longo dos anos a empresa seguiu os passos da industrialização do Brasil e da cidade de Joinville onde a partir da década de 50 com o desenvolvimento da indústria automotiva brasileira passou a fabricar peças especiais para este segmento e, a partir de 1995 inaugura em Joinville a unidade de blocos e cabeçotes de motores. Parceira das indústrias automotivas, a Tupy hoje concentra 75% de sua produção destinada ao setor automobilístico com grande destaque para as exportações. Atualmente, a empresa desenvolve e fabrica componentes em ferro fundido para os setores automotivo, ferroviário e de máquinas e equipamentos. Além disso, produz vários tipos de conexões de ferro maleável para diversas modalidades de aplicação. Nas últimas décadas a empresa dobrou sua capacidade produtiva e hoje se posiciona entre as cinco maiores fundições do planeta.



4.6.2 Descrição do produto selecionado para análise estatística multivariada do processo de usinagem

O produto selecionado para a aplicação da metodologia proposta via análise estatística multivariada do processo de usinagem é um bloco de motor para veículo de passeio que é produzido em ferro fundido e posteriormente usinado. Para preservar as informações sigilosas referentes ao projeto do produto utilizado no trabalho não é apresentado o desenho deste, mas tão somente um esboço considerado essencial para o entendimento das características da qualidade do processo de usinagem propostas nesta tese para serem monitoradas via gráfico de controle MCUSUM conforme figura 4.9.

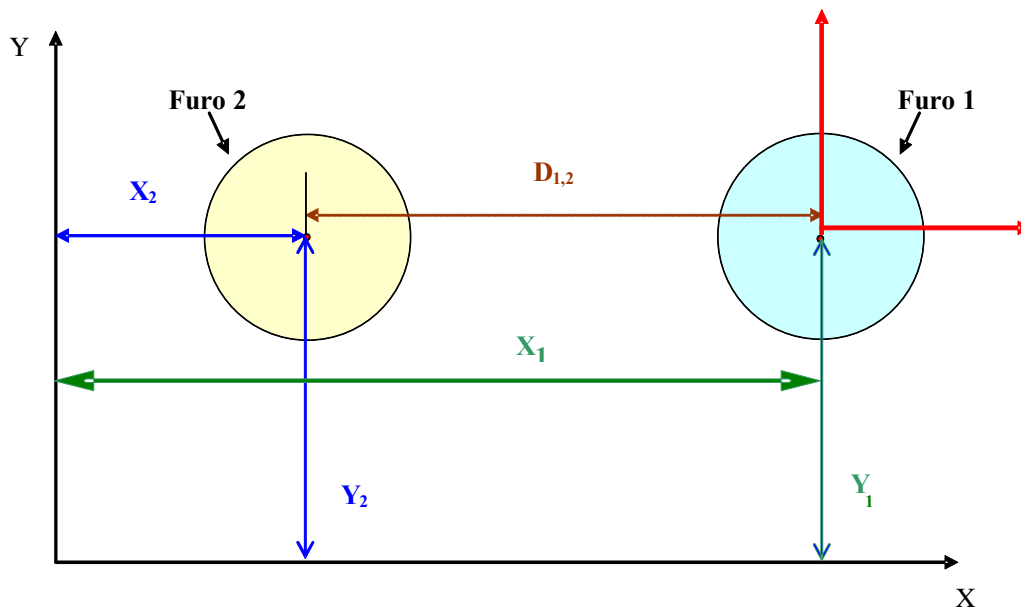


Figura 4.9 Esboço dos furos 1 e 2 do bloco de motor para veículo de passeio

Conforme figura 4.9 a posição do centro de cada furo é definida pela distância aos eixos X:Y com tolerância delimitada por um cilindro de diâmetro de 0,16 mm. O eixo de cada cilindro é perpendicular ao plano formado pelos eixos X e Y. O centro do furo 1 apresenta como especificação a localização de 5 mm em relação a X, 103,25 mm em relação a Y e distância de 194,27 mm em relação ao centro do furo 2. A especificação da região de tolerância do processo nas direções dos eixos x e y, é definida pela equação:

$$\text{tolerância_no_eixo} = (\text{tolerância} - \text{posicional}) / 2\sqrt{2} \quad (4.1)$$

Assim, a $\text{tolerância_no_eixo} = (\text{tolerância} - \text{posicional}) / 2\sqrt{2} = 0,0566$. Desta forma, a especificação no eixo X é 103,25 +/- 0,0566 mm e no eixo Y é 5 +/- 0,0566 mm. Estes valores delimitam a zona de tolerância quadrada das coordenadas X e Y.

4.6.3 Descrição do processo selecionado para o monitoramento de características da qualidade via gráfico MCUSUM

O processo de furação do bloco de motor em questão é realizado em duas etapas de operação: a operação pré-furo e a operação de furação de acabamento. Nesse processo os dois furos 1 e 2 do bloco são referências para o posicionamento da peça no dispositivo de fixação a cada operação de usinagem realizada em diferentes máquinas. Na primeira operação de usinagem é realizada a furação em desbaste das guias do bloco com a utilização de broca de metal duro e a alargador. Nesta etapa inicial do processo o objetivo é o limite inferior de especificação do diâmetro do furo enquanto que na operação furação de acabamento o objetivo é o valor nominal. Após as operações iniciais é realizada a usinagem de acabamento do furo guia do bloco acabado utilizando-se barra de mandrilar com guias de cermet. O objetivo nesta operação é tanto o nominal do diâmetro quanto o nominal do posicional. O posicional é utilizado como guia para em seguida serem realizadas as operações de acabamento na peça. O fluxo resumido para a obtenção do bloco de motor deste processo de usinagem é esquematizado conforme figura 4.10.

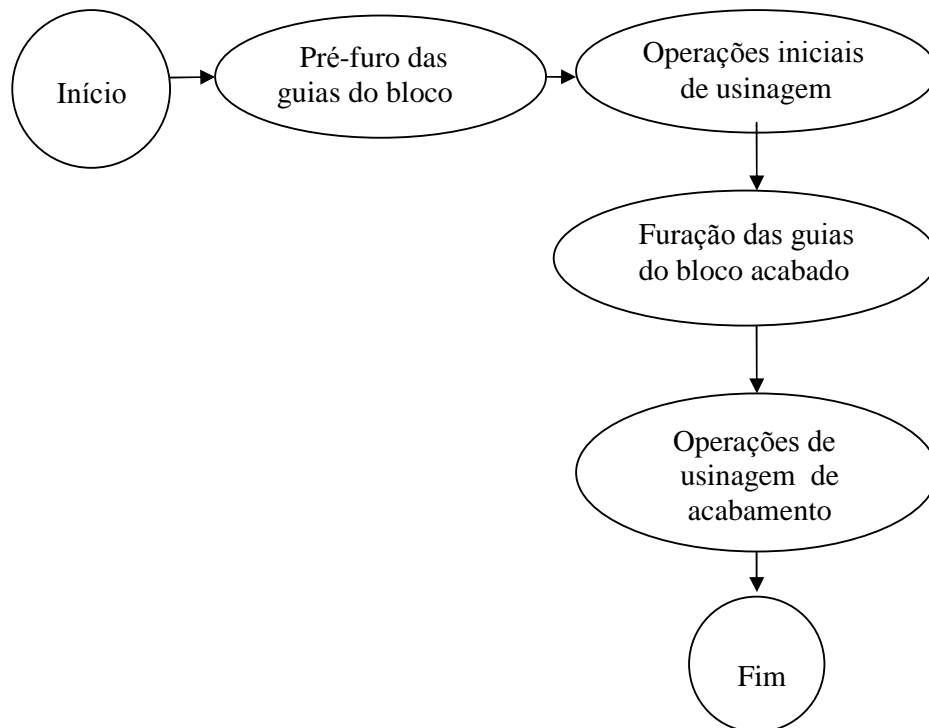


Figura 4.10 Fluxo resumido do processo de usinagem do bloco de motor

A metodologia proposta neste trabalho é aplicada apenas para os dados referentes a operação furação de acabamento do furo 1. As características da qualidade monitoradas

via gráfico MCUSUM no processo furação de acabamento proposto neste trabalho conforme figura 4.9 são:

X_1 : posicional do furo 1 em relação a coordenada X;

Y_1 : posicional do furo 1 em relação a coordenada Y;

$D_{1,2}$: distância entre os centros dos furos 1 e 2 ;

Para monitorar o processo furação de acabamento com duas variáveis ou características da qualidade utilizou-se as variáveis X_1 e Y_1 e, para o estudo de três variáveis incluiu-se uma terceira variável que é a distância $D_{1,2}$ entre os furos.

A escolha deste processo se justifica dentre os demais por se tratar de um processo onde se pode monitorar múltiplas características da qualidade que apresentam distribuição normal multivariada e com pequenas variações para o vetor de médias do processo. Além disso, a empresa possui maior quantidade de dados históricos sobre o controle deste processo, por estar utilizando o CEP de forma consolidada há mais tempo; tais dados são essenciais para o projeto estatístico do gráfico de controle MCUSUM.

4.6.4 Exploração dos dados do processo para aplicação da metodologia proposta

A exploração preliminar dos dados das características da qualidade em estudo neste processo é conduzida por um algoritmo próprio para CEP multivariado cujos procedimentos incluem a verificação de certos pressupostos como a verificação de suposição de multinormalidade dos dados quanto à existência de autocorrelação e estabilidade estatística (inexistência de causas especiais). A validade dessas três suposições para avaliar os dados do processo em questão foram já verificadas por Soares (2006) em sua dissertação de mestrado para cálculo de índice de capacidade multivariado. No entanto, alguns comentários extras que justificam a validade de tais pressupostos são ressaltados ao longo deste capítulo tendo como referencial teórico alguns tópicos descritos no Apêndice C e D deste trabalho.

Para facilitar o entendimento da metodologia proposta cujo estudo envolve a otimização dos parâmetros para o gráfico MCUSUM no processo furação de acabamento do furo1, denominamos o processo de usinagem neste trabalho conforme o número de variáveis monitoradas simultaneamente.

PROCESSO: XY

É a denominação dada a este processo que envolve o monitoramento simultâneo de duas características da qualidade (X_1 e Y_1).

X_1 (posicional do furo1 em relação a coordenada X) cuja especificação é $5 \pm 0,08$ mm (Valor Nominal)

Y_1 (posicional do furo1 em relação a coordenada Y) cuja especificação é $103,25 \pm 0,08$ mm (Valor Nominal)

Cada uma destas características da qualidade é avaliada a partir de amostras de tamanho $n=1$.

PROCESSO: XYD

É a denominação dada a este processo que envolve o monitoramento simultâneo de três características da qualidade (X_1 , Y_1 e D_{12}), ou seja, as duas características da qualidade do processo XY com a inclusão de uma terceira variável ou característica da qualidade (D_{12} - distância entre os centros dos furos 1 e 2) também avaliada a partir de amostras de tamanho $n=1$ com o valor nominal de $194,27 \pm 0,02$ mm.

Os dados para a análise estatística multivariada destes dois processos foram coletados durante a etapa de aprovação do PPAP (Processo de Aprovação de Produção da Peça) e consistiu na produção seqüencial das peças seguido de medição logo após a operação em uma máquina de medição de coordenadas. Nesta etapa iniciou-se a produção de 34 peças, de forma a garantir que no mínimo 30 peças estivessem aprovadas para envio ao cliente. Ao final da produção foram obtidas 31 peças. Os dados das 31 peças (blocos de motor) para as três características da qualidade por ordem de coleta da amostra conforme Anexo A. Neste trabalho, para facilitar a interpretação dos dados no desenvolvimento do projeto estatístico ótimo do gráfico MCUSUM utiliza-se o valor da observação individual em cada uma das três características da qualidade, enquanto Soares (2006) utilizou o desvio de cada observação em relação ao valor nominal das características da qualidade cujo interesse foi o cálculo de índices de capacidade.

4.7 Análise estatística multivariada dos dados: Processo XY

A análise estatística multivariada preliminar dos dados do processo furação de acabamento do furo 1 para as duas características da qualidade X_1 e Y_1 (processo XY) inclui inicialmente a verificação de suposição de multinormalidade dos dados e a existência de autocorrelação.

4.7.1 Normalidade

Para a análise estatística multivariada é fundamental que os dados sigam uma

distribuição normal multivariada, quando não exatamente, pelo menos, aproximadamente. Se tal fato for verificado, ou seja, os dados forem distribuídos segundo uma distribuição normal multivariada, evita-se o problema de procurar procedimentos de transformação de variáveis, que na sua maioria são robustos, de tal forma a normalizar o conjunto de dados para, a partir daí, trabalhar-se com os dados multivariados normalmente distribuídos. A aplicação de métodos gráficos tais como o gráfico de probabilidade qui-quadrado (Q-Q plot) podem ser utilizados para auxiliarem na verificação de normalidade multivariada. Este método gráfico se constitui em determinar se os dados das amostras estão de acordo com a suposta distribuição, baseado em um exame visual subjetivo dos dados. Além disso, um outro procedimento é utilizado neste trabalho para detecção de normalidade multivariada relacionado aos testes de hipóteses proposto por Mardia (1970,1974) baseados nos coeficientes de assimetria e curtose da distribuição normal multivariada. Maiores detalhes sobre os métodos práticos de verificação da hipótese de normalidade aplicados neste trabalho conforme Apêndice C.

Examinando visualmente os histogramas das variáveis X_1 e Y_1 deste processo observa-se conforme figura 4.11 que individualmente ambas as características da qualidade (variáveis) X_1 e Y_1 seguem uma distribuição normal.

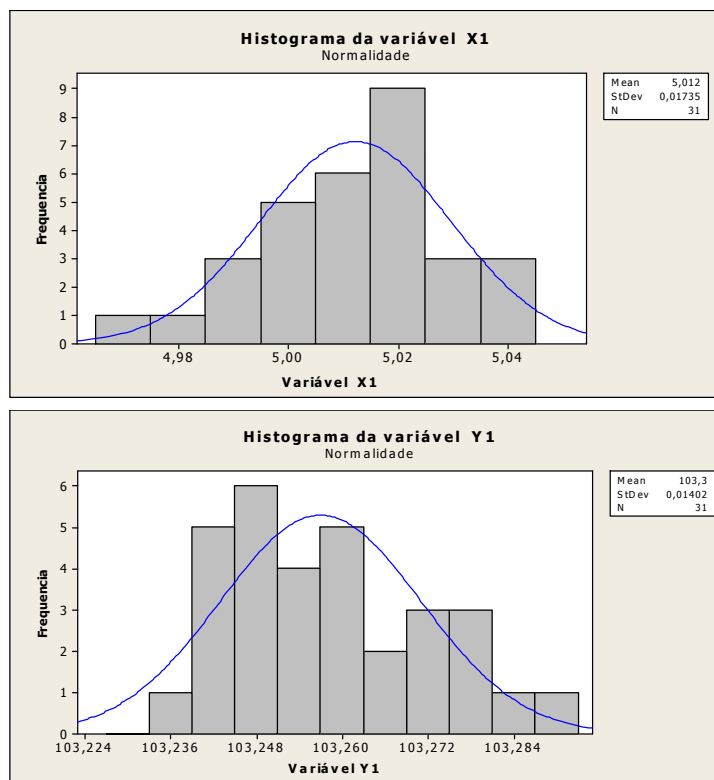


Figura 4.11 Histogramas das características da qualidade X_1 e Y_1 (Processo XY)

Agora, verificando visualmente a suposição de normalidade multivariada através do gráfico de probabilidade qui-quadrado (Q-Q plot) desenvolvido neste trabalho para dos dados deste processo (variáveis X_1 e Y_1) em ambiente MS-Excel cuja representação gráfica conforme figura 4.12.

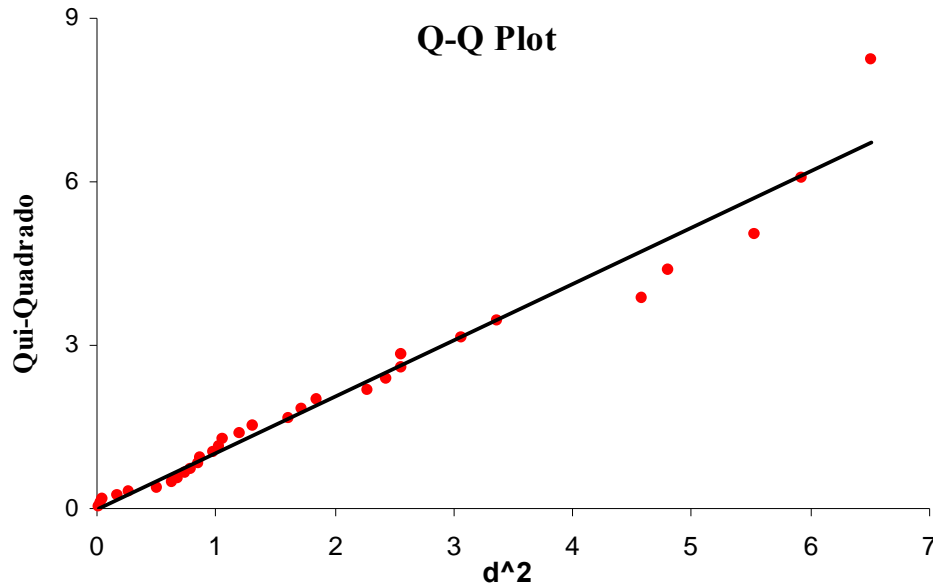


Figura 4.12 Gráfico de probabilidade qui-quadrado (Q-Q plot) para os dados do processo XY

Podemos constatar visualmente que a disposição dos pontos indica a proximidade em relação à reta, o que nos leva a não rejeitar a suposição de normalidade multivariada. Além disso, analisando a multinormalidade dos dados deste processo segundo os testes de hipóteses de Mardia baseados nos coeficientes de assimetria e curtose conforme resultados obtidos via Programa Computacional 3 indicam que o valor do percentil tanto para $\beta_{1,p}$ quanto para $\beta_{2,p}$ são suficientes para aceitar a hipótese de normalidade. Desta forma, pode-se afirmar que as observações multivariadas deste processo se distribuem segundo uma multinormal.

4.7.2 Autocorrelação

Para que o gráfico MCUSUM atinja um desempenho razoável no monitoramento simultâneo das variáveis X_1 e Y_1 é necessário que a suposição de independência estatística entre as observações dos dados gerados pelo processo XY em questão seja satisfatória. Conforme literatura existente, muitos autores aplicam um modelo apropriado de séries temporais às observações para trabalhar com dados autocorrelacionados. Segundo Mason

(1997, 2002) e Noorossana e Vaghefi (2006) o ajuste de um modelo de séries temporais para dados multivariados também é possível.

Para verificar a presença de autocorrelação dos dados deste processo aplicou-se a metodologia desenvolvida tanto por Mason (1997, 2002) quanto por Noorossana e Vaghefi (2006) que sugere aplicações de séries temporais envolvendo a análise de função de autocorrelação (correlogramas) para medir o grau de autocorrelação em cada variável. Além disso, para avaliar o efeito comparativo da análise de correlogramas propõem também a aplicação de técnicas de estatística inferencial tais como Análise de Variância (ANOVA) para o desenvolvimento de um modelo de Análise de Regressão AR(1) da forma:

$y_t = b_0 + b_1 y_{t-1} + erro$. As observações geradas por AR(1) são utilizadas para investigar o efeito da autocorrelação no desempenho do gráfico de controle.

A figura 4.13 ilustra os correlogramas para as características da qualidade X_1 e Y_1 deste processo.

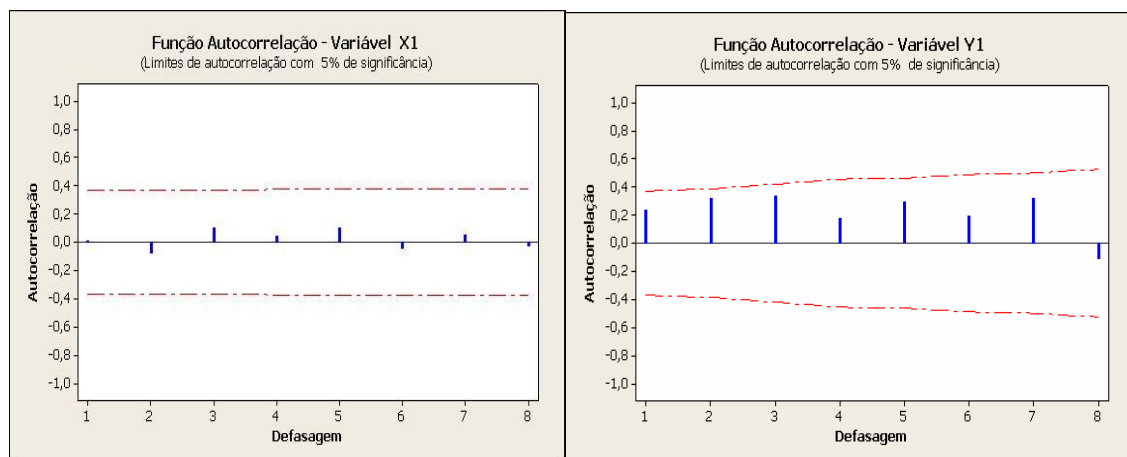


Figura 4.13 Correlogramas das características da qualidade X_1 e Y_1 (Processo XY)

Como se pode observar a função autocorrelação (FAC) tanto da característica da qualidade X_1 quanto da característica da qualidade Y_1 revelam que nenhum coeficiente de correlação é significativo. Além disso, verificada a adequação do modelo AR(1) e, por ser estacionário, vale, então a hipótese de estabilidade estatística deste processo.

A função autocorrelação (FAC) e o comparativo dos correlogramas das variáveis X_1 e Y_1 com a inferência estatística análise de variância (ANOVA) cujos resultados conforme figuras 4.14 e 4.15

Função Autocorrelação (FAC) - Variáveis X₁; Y₁

FAC - correlaciona X₁(t) e Y₁(t+k)

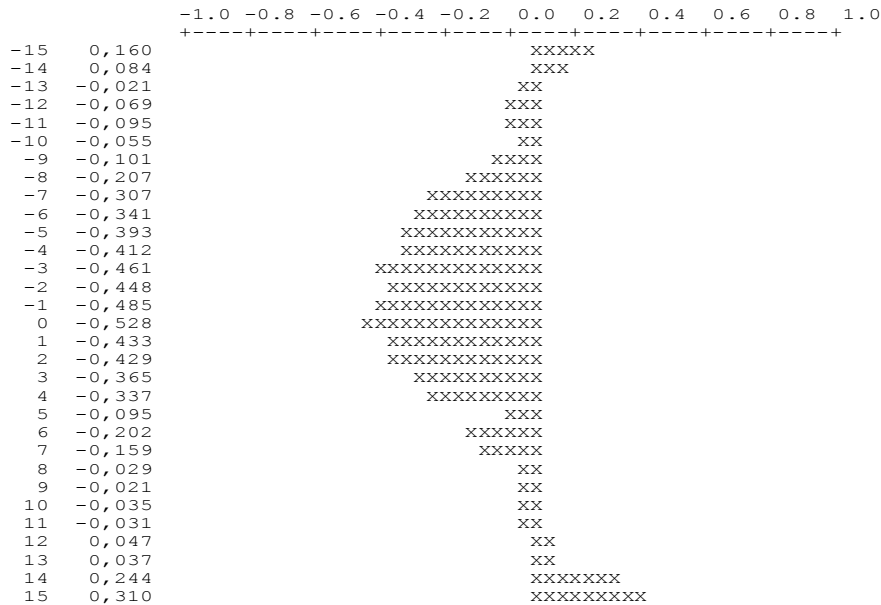


Figura 4.14 Função Autocorrelação (FAC) das variáveis X₁ com Y₁ (Processo XY)

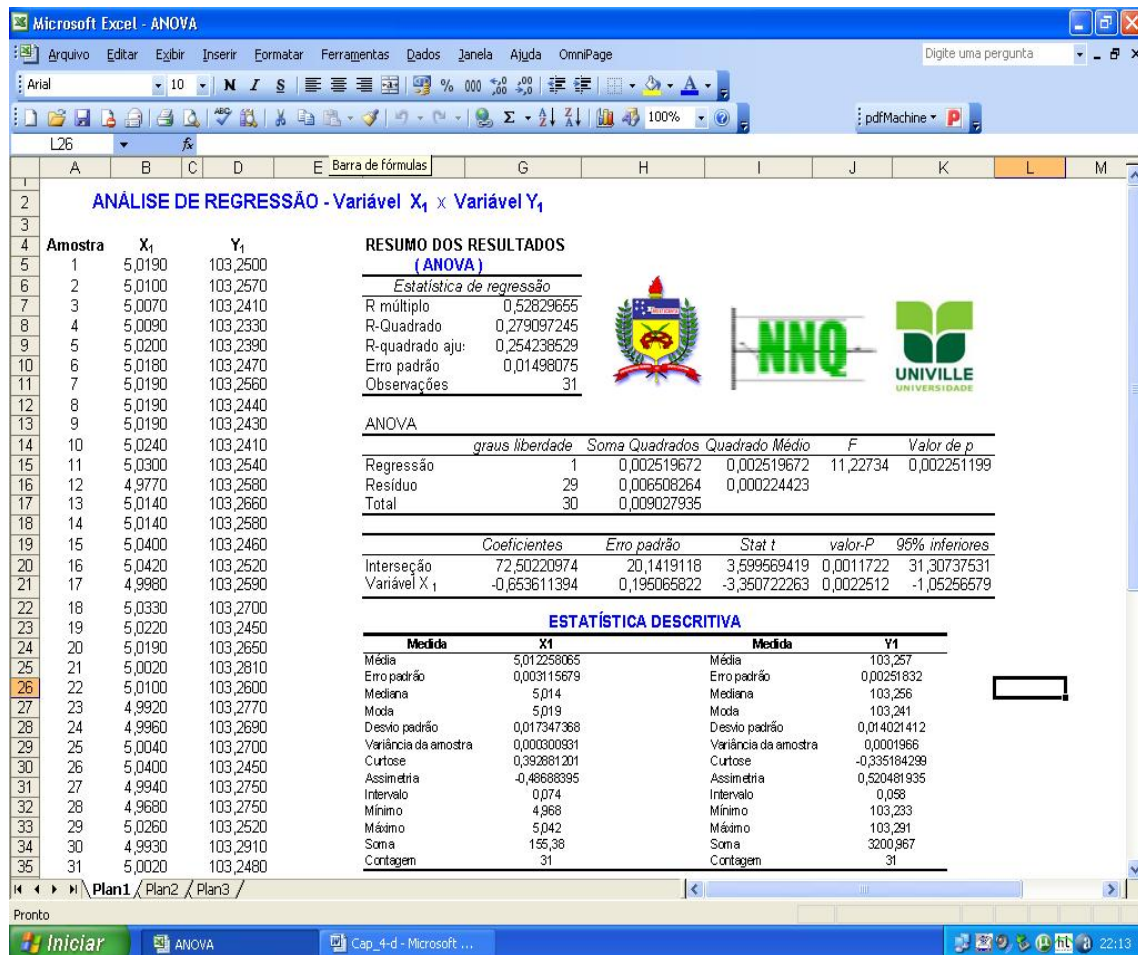


Figura 4.15 Análise de Variância (ANOVA) e Estatística Descritiva para as variáveis X₁ e Y₁

Conforme figuras 4.14 e 4.15 pode-se verificar os resultados obtidos da Função Autocorrelação (FAC) e o efeito comparativo da análise de correlogramas com a Análise de Variância (ANOVA) que indicam a não existência de autocorrelação nos dados deste processo. Além disso, os resultados obtidos a partir da Estatística Descritiva revelam a estabilidade estatística (inexistência de causas especiais de variação conforme medidas de localização e dispersão obtidas a análise estatística dos dados deste processo).

4.7.3 Otimização dos pontos que delimitam as regiões de máxima e mínima potência do gráfico MCUSUM para o processo XY

Um método baseado na função perda multivariada de Taguchi e na distância de Mahalanobis conforme seção 4.5 deste trabalho é proposto para determinar os pontos ótimos A e B que delimitam as regiões de máxima e mínima potência do gráfico MCUSUM deste processo. Para otimizar estes pontos aplica-se um modelo matemático desenvolvido em ambiente MS-Excel onde com o auxílio da ferramenta Solver é possível calcular a menor distância d , ou seja, encontrar o menor desvio em relação ao valor nominal das características da qualidade X_1 e Y_1 . Para complementar este procedimento de otimização bem como a análise estatística multivariada do gráfico MCUSUM para os dados do processo XY, aplica-se em ambiente Matlab o Programa Computacional 3 proposto na seção 4.3 cujo tutorial conforme Apêndice K deste trabalho. Os resultados (de saída) dos dados do Processo XY gerados após a execução deste Programa Computacional 3 encontram-se também no Apêndice K deste trabalho.

Uma vez determinado a função perda multivariada de Taguchi e a distância de Mahalanobis utilizam-se os dados obtidos via Programa Computacional 3 para formular um modelo matemático em ambiente MS-Excel. Na resolução deste modelo matemático utiliza-se a ferramenta Solver para encontrar os pontos ótimos A e B que delimitam as regiões de máxima e mínima potência do gráfico MCUSUM. Para a otimização destes pontos é necessário conhecer a política de recursos financeiros que a empresa direciona a melhoria de produtividade e qualidade. Para este processo supõe-se que a política de recursos financeiros da empresa considere um custo ou perda depreciável de 0,0012 u.m. e um custo ou perda inadmissível de 0,0042 u.m. (unidade monetária).

Assim, o modelo matemático para otimizar os pontos de máxima e mínima potência do gráfico de controle MCUSUM desenvolvido numa planilha de MS-Excel para os dados do processo XY conforme figura 4.16.

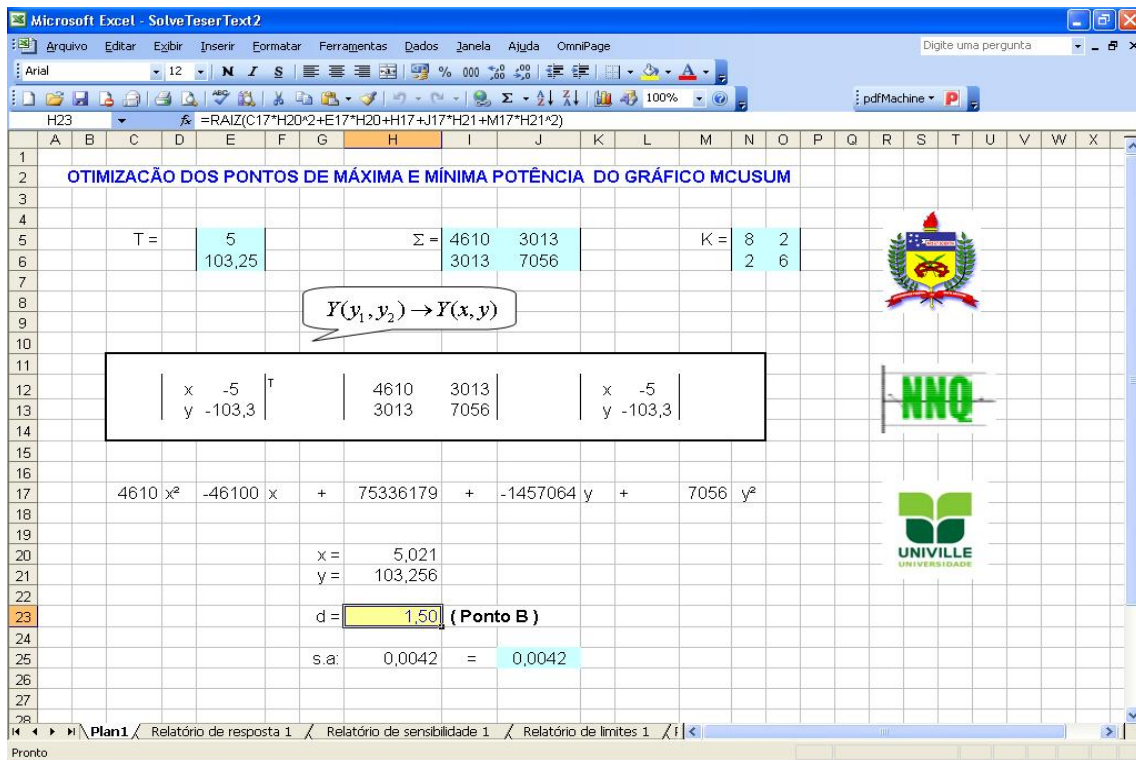
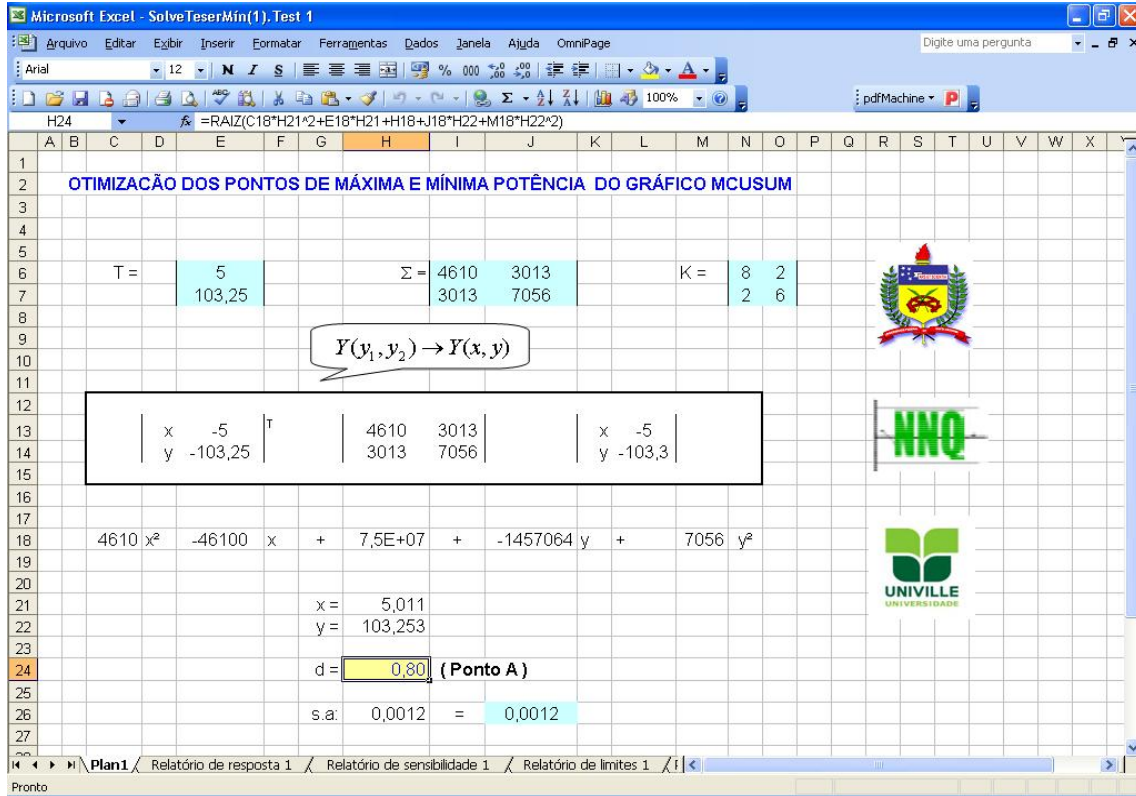


Figura 4.16 Modelo matemático para otimizar os pontos de máxima e mínima potência

Os relatórios de resposta, análise de sensibilidade e de limites deste modelo matemático conforme Apêndice I.

A resolução do modelo matemático (figura 4.16) fornece os pontos de máxima e mínima potência ($A=0,80$ e $B=1,50$) para otimizar o gráfico MCUSUM do processo XY conforme figura 4.17.

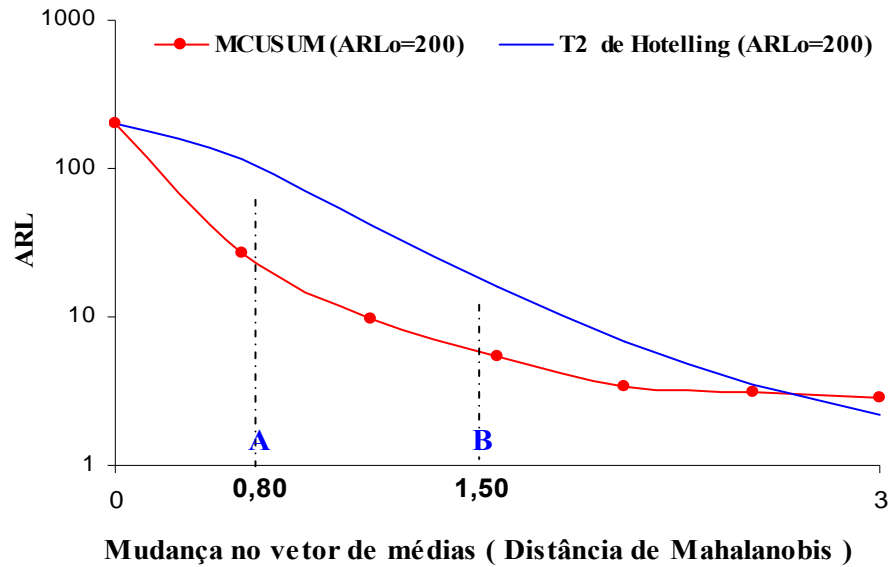


Figura 4.17 Pontos de máxima e mínima potência para otimizar o gráfico MCUSUM

Conhecidos os pontos ótimos A e B do gráfico que delimitam as regiões de máxima e mínima potência conforme etapa 3 sugerida para o algoritmo de procedimentos do projeto estatístico ótimo do gráfico MCUSUM (seção 4.5 deste trabalho), conclui-se que para este processo o tamanho (amplitude) de mudança desejado é $d=1,50$ (ponto B). Portanto, nosso interesse neste caso deve estar focado em detectar mudanças de vetor de médias do processo para $d \geq 1,50$. Consultando a tabela 4.1 para ($p=2$ e $ARL_0=200$) para selecionar os valores dos parâmetros do gráfico MCUSUM obtidos via Método de Equação Integral com Quadratura Gaussiana proposto neste trabalho, ou utilizando o programa computacional 1, obtém-se então os parâmetros para projetar o gráfico MCUSUM ótimo deste processo. Assim, os parâmetros ótimos deste processo determinados para um tamanho (amplitude) de mudança $d=1,50$:

$k=0,75$, $h=3,95$ cujo desempenho é o $ARL_{\text{mínimo}}= 5,39$.

4.7.4 Desenvolvimento do gráfico MCUSUM projetado estatisticamente para o processo XY

Nas últimas décadas com o avanço da tecnologia programas computacionais comerciais têm possibilitado um maior desenvolvimento e implantação de métodos

estatísticos multivariados na indústria como ferramenta de controle da qualidade. No entanto, no que diz respeito ao desenvolvimento específico de gráficos de controle multivariados, pouca atenção tem sido direcionada a atender as especificidades dessas ferramentas. Felizmente, programas computacionais livres, como por exemplo, o r-project tem se propagado principalmente no meio acadêmico na última década como uma ferramenta de computação estatística com uma enorme variedade de funções para análise de dados. Alves, Henning e Samohyl (2008a) propõem o desenvolvimento de gráficos de controle multivariados MCUSUM e MEWMA em ambiente R como um procedimento alternativo e adequado ao usuário para análise estatística de processos multivariados. Destacam, sobretudo a importância deste recurso computacional livre para complementar algumas funções ainda não disponíveis na maioria dos softwares comerciais de análise estatística de dados, como por exemplo, o desenvolvimento do gráfico MCUSUM

Neste trabalho desenvolveu-se uma rotina computacional no pacote GNU R para gerar o gráfico MCUSUM que é aplicado para monitorar as duas características da qualidade X_1 e Y_1 do processo XY (furação de acabamento do furo1) conforme Apêndice J.

As telas do R para carregar e executar a rotina do gráfico de controle MCUSUM conforme figuras 4.18 e 4.19.

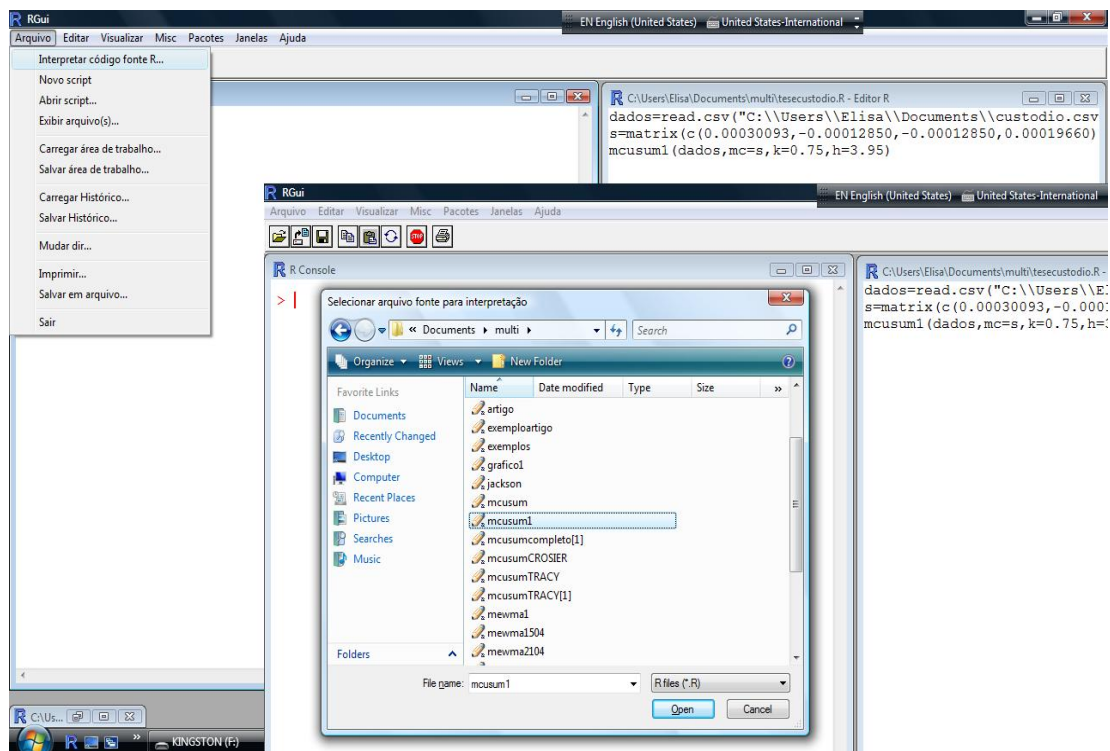


Figura 4.18 Tela do R para carregar a rotina do gráfico MCUSUM

```

RGui - [R Console]
Arquivo  Editar  Visualizar  Misc  Pacotes  Janelas  Ajuda
EN English (United States)  United States-International

> source("C:\\Users\\Elisa\\Documents\\multi\\mcusum1.R")
> dados=read.csv("C:\\Users\\Elisa\\Documents\\custodio.csv",header=T,sep=";",dec=",")
> s=matrix(c(0.00030093,-0.00012850,-0.00012850,0.00019660),ncol=2)
> mcusum1(dados,mc=s,k=0.75,h=3.95)

SESTATISTICA.MCUSUM
[1] 0.0000000 0.0000000 0.8407274 2.5198994 3.0094172
[6] 2.7138907 1.6834068 2.2442800 2.8066369 3.3694582
[11] 3.1527953 2.7835555 1.1032062 0.2729868 0.8229845
[16] 1.7166936 0.2603180 1.3490627 0.7138810 1.2068018
[21] 1.6627228 1.0098517 1.7061497 1.8257864 1.9603177
[26] 1.0480491 1.6087304 2.8718326 1.2565484 2.7636057
[31] 1.8783451

Smatriz.cov
      [,1]      [,2]
[1,] 0.00030093 -0.0001285
[2,] -0.00012850 0.0001966

Smedia
      [,1]
[1,] 5.011935
[2,] 103.258065

SH
[1] 3.95

SK
[1] 0.75

Spontos.fora
[1] 0

Svalores
numeric(0)

> |
    
```

Figura 4.19 Tela do R para executar a rotina do gráfico MCUSUM

O gráfico MCUSUM gerado para os dados do processo XY conforme figura 4.20

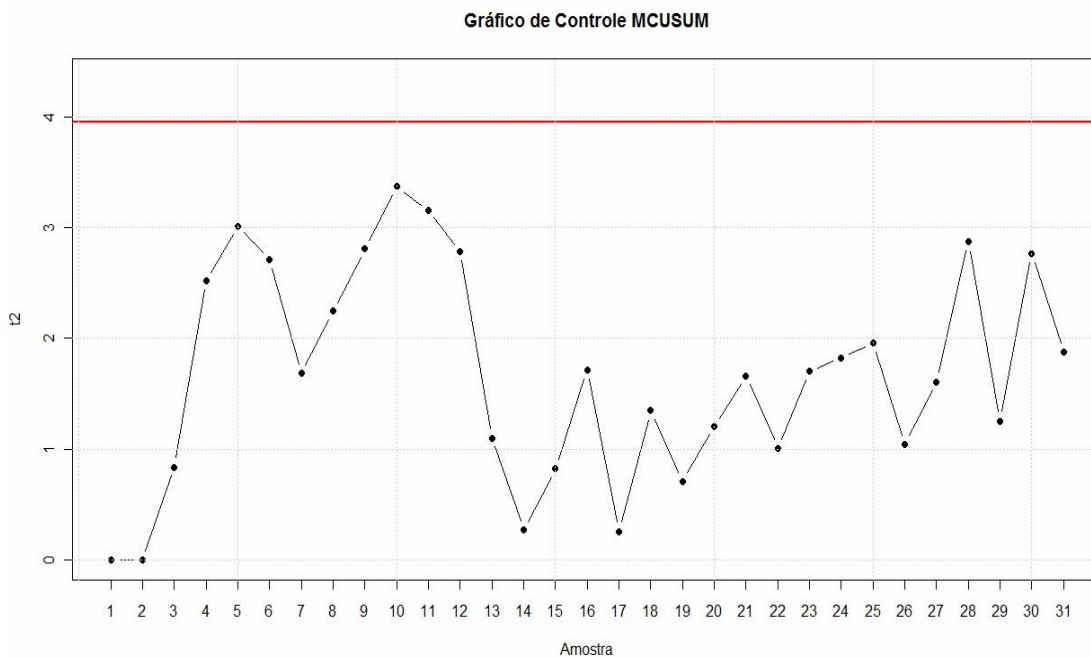


Figura 4.20 Gráfico MCUSUM do processo XY(furação de acabamento do furo 1)

Uma análise preliminar deste processo conforme figura 4.20 revela que, aparentemente, se trata de um processo estável, apesar de indicar uma tendência inicial ascendente (amostras 3 a

5), uma queda abrupta (amostra 7), uma nova tendência ascendente (amostras 8 a 10) e uma queda (amostras 11 a 14). No entanto, a partir da 15ª amostra observa-se que há uma melhoria do processo cujos pontos demarcados mostram um comportamento aleatório dos valores. Além disso, pode-se verificar que não há presença de causas especiais de variação já que todos os pontos se encontram dentro do limite superior de controle, h .

4.8 Análise estatística multivariada dos dados: Processo XYD

A análise estatística multivariada preliminar dos dados do processo furação de acabamento do furo 1 para as três características da qualidade X_1 , Y_1 e D_{12} (Processo XYD) que envolve agora uma terceira característica da qualidade D_{12} (distância entre os centros dos furos 1 e 2) inclui inicialmente a verificação de suposição de multinormalidade dos dados e a existência de autocorrelação.

4.8.1 Normalidade

A verificação de normalidade multivariada agora para as três características da qualidade aplicando tanto o método gráfico de probabilidade qui-quadrado Q-Q Plot quanto testes de hipóteses baseados nos coeficiente de assimetria e curtose de Mardia, procedimentos de verificação também propostos anteriormente para as duas primeiras variáveis X_1 e Y_1 . Maiores detalhes dos métodos práticos de verificação de normalidade multivariada aplicados neste trabalho conforme Apêndice C.

Como recurso computacional em ambiente Matlab para verificação de normalidade multivariada além do Programa Computacional 3 desenvolveu-se um outro programa denominado Teste de Mardia que determina os coeficientes de assimetria e curtose aplicados nos testes de hipóteses de normalidade multivariada de Mardia. Além disso, este programa determina também a matriz de covariância, sua inversa e a matriz de correlação. Os dados de entrada deste programa em ambiente Matlab são transferidos de uma planilha em MS-Excel. Maiores detalhes deste programa conforme Apêndice K deste trabalho.

Examinando visualmente o histograma das características da qualidade deste processo X_1 , Y_1 e D_{12} observa-se que individualmente estas três variáveis seguem uma distribuição normal conforme figura 4.21.

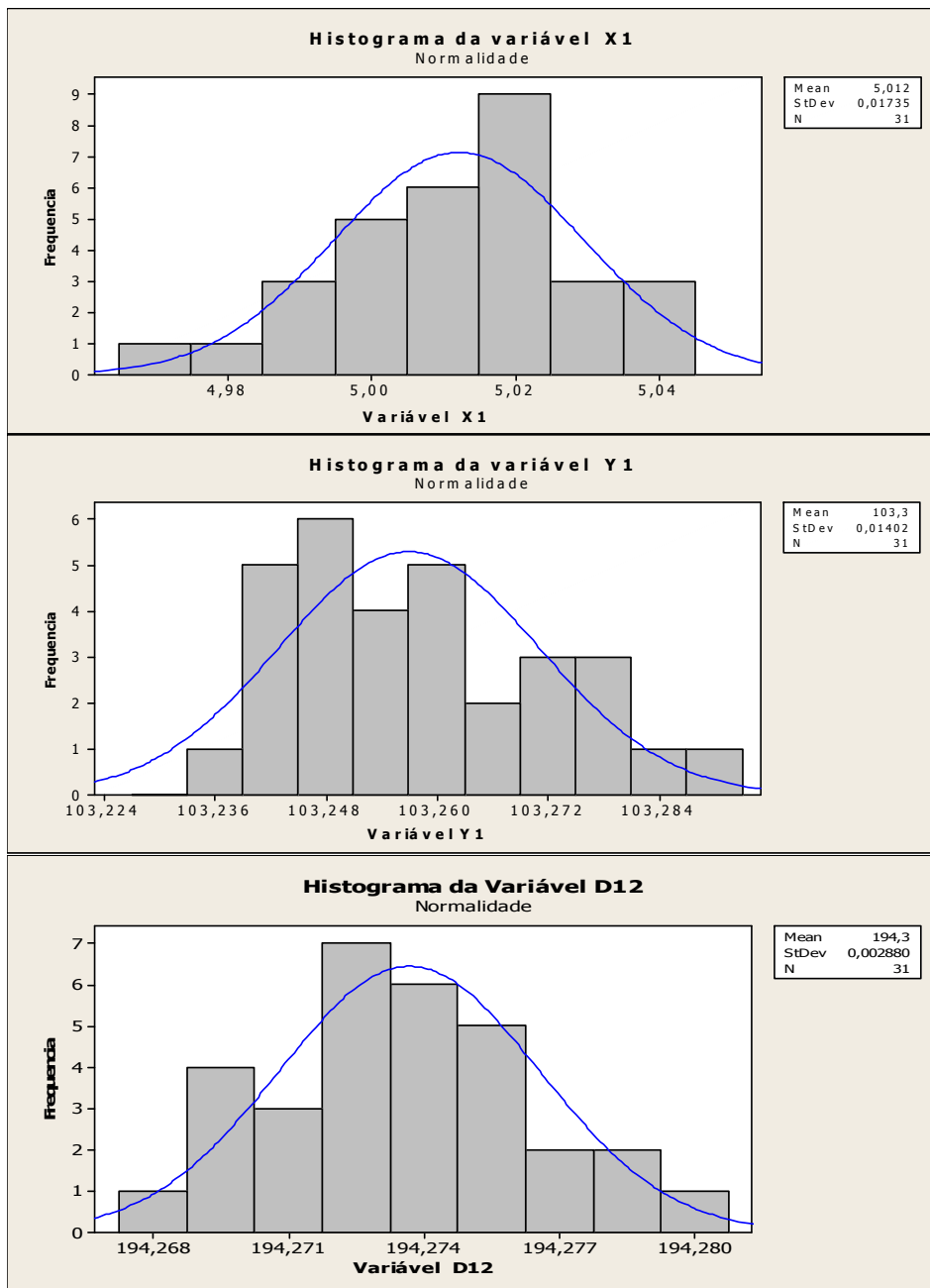


Figura 4.21 Histogramas das características da qualidade X_1 , Y_1 e D_{12} (Processo XYD)

Este fato pode ser constatado conforme gráfico de probabilidade normal (figura 4.22) utilizado para avaliar a hipótese de normalidade univariada de cada uma das três características da qualidade deste processo onde as probabilidades de significância correspondentes estão acima do valor de referência 0,05 indicando que a suposição de normalidade é plausível.

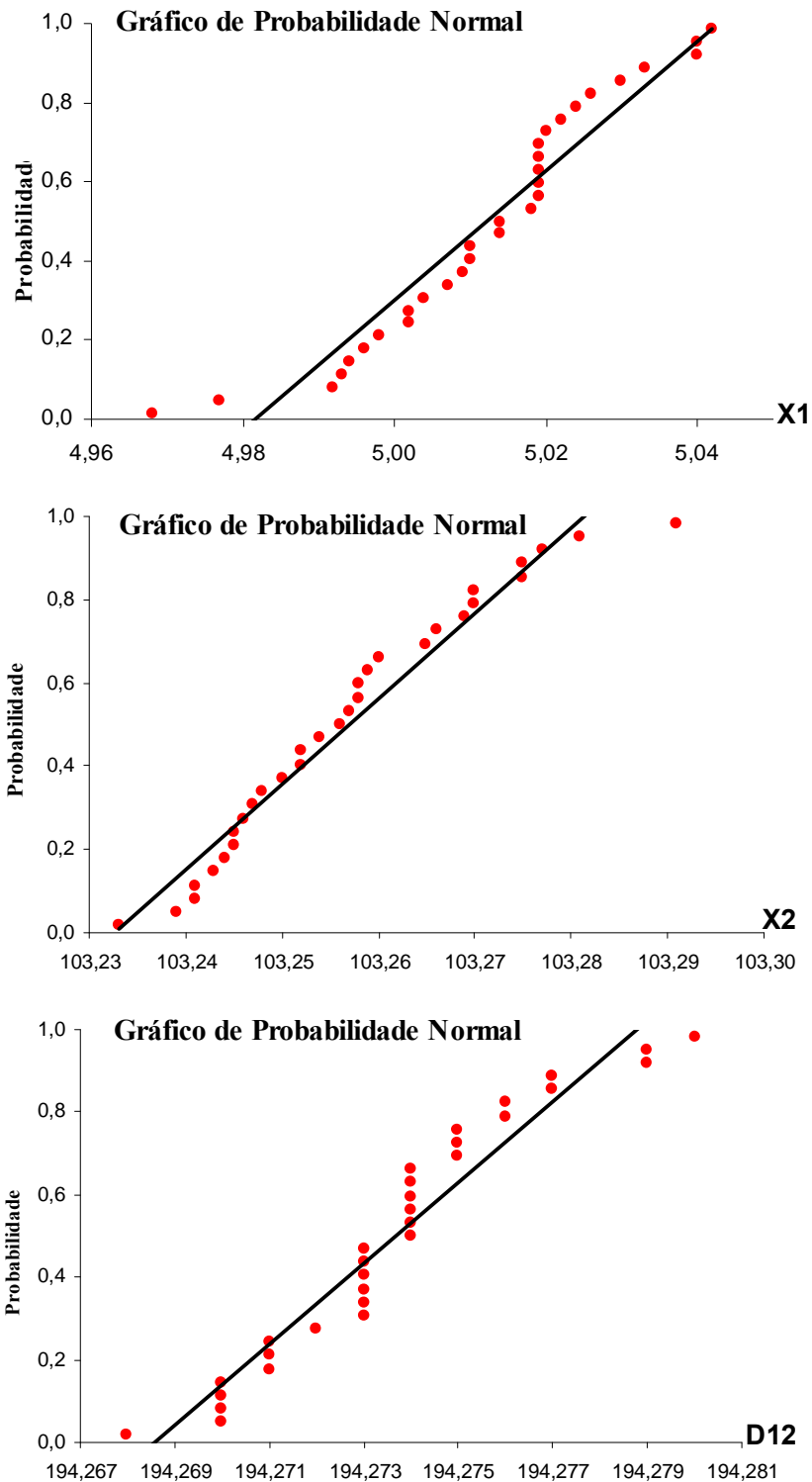


Figura 4.22 Gráfico de Probabilidade Normal das variáveis X_1 , Y_1 e D_{12} (Processo XYD)

Agora, verificando visualmente a suposição de normalidade multivariada através do gráfico de probabilidade qui-quadrado (Q-Q plot) desenvolvido para as características da qualidade deste processo conforme figura 4.23.

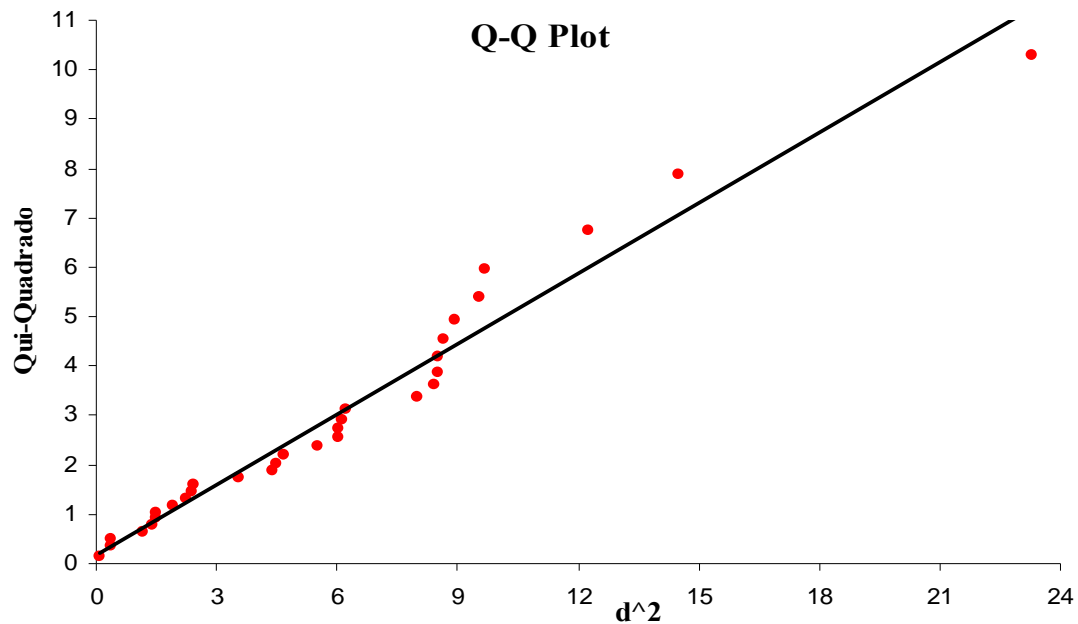


Figura 4.23 Gráfico de probabilidade qui-quadrado (Q-Q plot) para os dados do processo XYD

Podemos constatar visualmente que a disposição dos pontos indica a proximidade em relação a uma reta, o que nos leva a não rejeitar a suposição de normalidade multivariada. Além disso, analisando a multinormalidade dos dados deste processo segundo os testes de hipóteses de Mardia baseados nos coeficientes de assimetria e curtose conforme resultados obtidos via Programa Computacional 3 cujos resultados indicam que o valor do percentil tanto para $\beta_{1,p}$ quanto para $\beta_{2,p}$ são suficientes para aceitar a hipótese de normalidade. Desta forma, pode-se afirmar que as observações multivariadas deste processo se distribuem segundo uma multinormal.

4.8.2 Autocorrelação

Para verificar a presença de autocorrelação dos dados deste processo aplicou-se a metodologia desenvolvida tanto por Mason (1997, 2002) quanto por e Noorossana e Vaghefi (2006) que sugere aplicações de séries temporais envolvendo a análise de função de autocorrelação (correlogramas) para medir o grau de autocorrelação em cada variável. Além disso, para avaliar o efeito comparativo da análise de correlogramas propõem também a aplicação de técnicas de estatística inferencial tais como Análise de Variância (ANOVA) para o desenvolvimento de um modelo de Análise de Regressão AR(1) da forma:

$y_t = b_o + b_1 y_{t-1} + erro$. As observações geradas por AR(1) são utilizadas para investigar o efeito da autocorrelação no desempenho do gráfico de controle. A figura 4.24 ilustra os

os correlogramas das três características da qualidade X_1 , Y_1 e D_{12} deste processo.

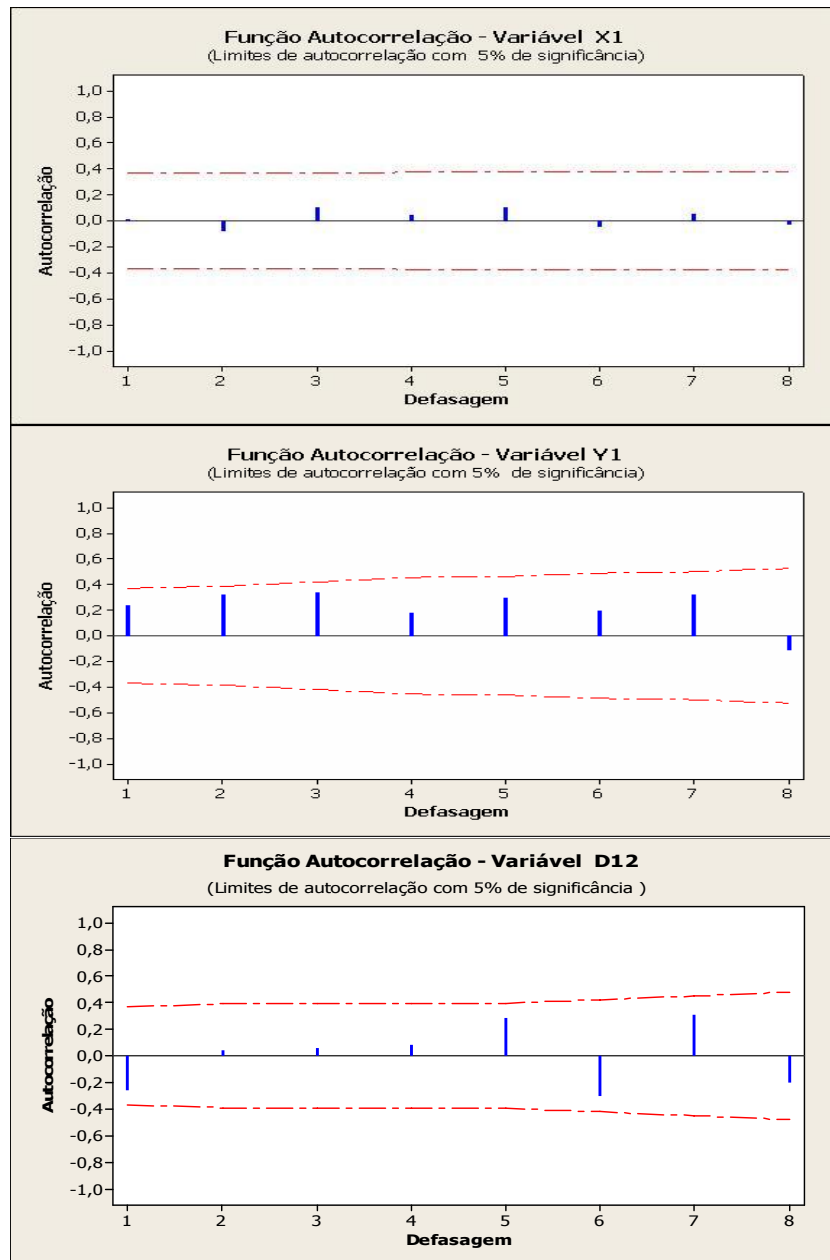


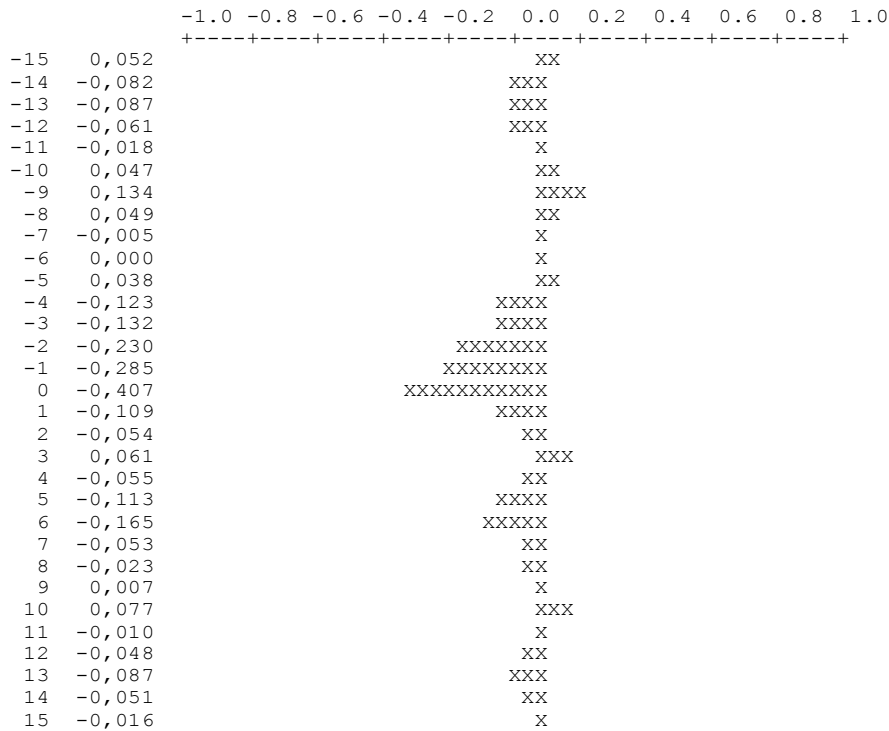
Figura 4.24 Correlogramas das características da qualidade X_1 , Y_1 e D_{12} (Processo XYD)

Como se pode observar a função autocorrelação (FAC) de cada uma das três características da qualidade X_1 , Y_1 e D_{12} revelam que nenhum coeficiente de correlação é significativo. Além disso, verificada a adequação do modelo AR(1) e, por ser estacionário, vale, então a hipótese de estabilidade estatística deste processo.

A função FAC das variáveis X_1 com D_{12} , Y_1 com D_{12} e o comparativo do correlograma das variáveis X_1 , Y_1 e D_{12} com a inferência estatística análise de variância (ANOVA) de X_1 com D_{12} e Y_1 com D_{12} cujos resultados conforme figuras 4.25, 4.26 e 4.27.

Função Autocorrelação (FAC) - Variáveis X1; D12

FAC - correlaciona X1(t) e D12(t+k)



Função Autocorrelação (FAC) - Variáveis Y1; D12

FAC - Correlaciona Y1(t) e D12(t+k)

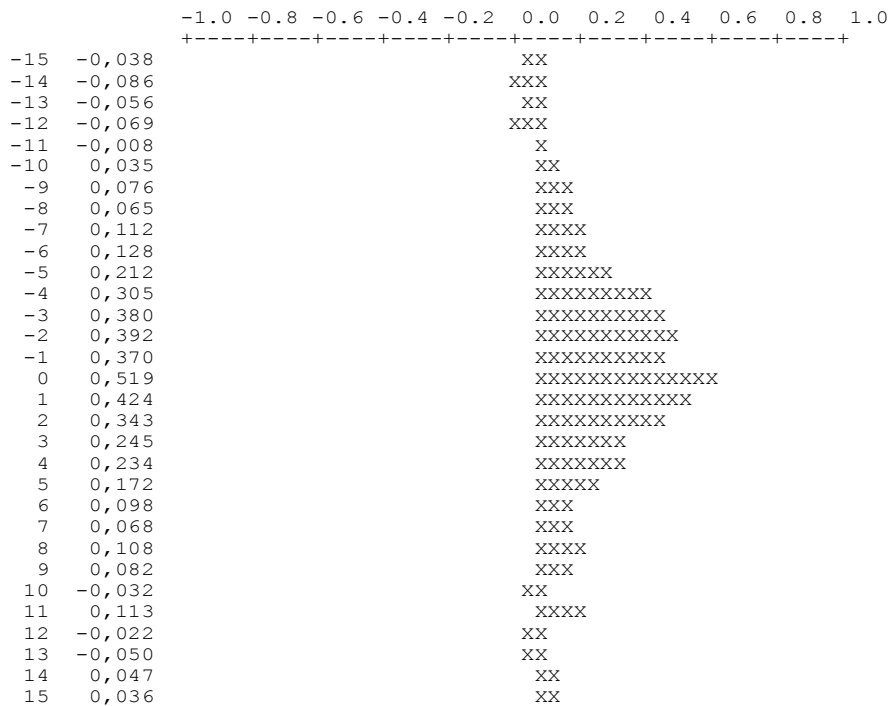


Figura 4.25 Função Autocorrelação (FAC) para as variáveis X1 com D12 e Y1 com D12

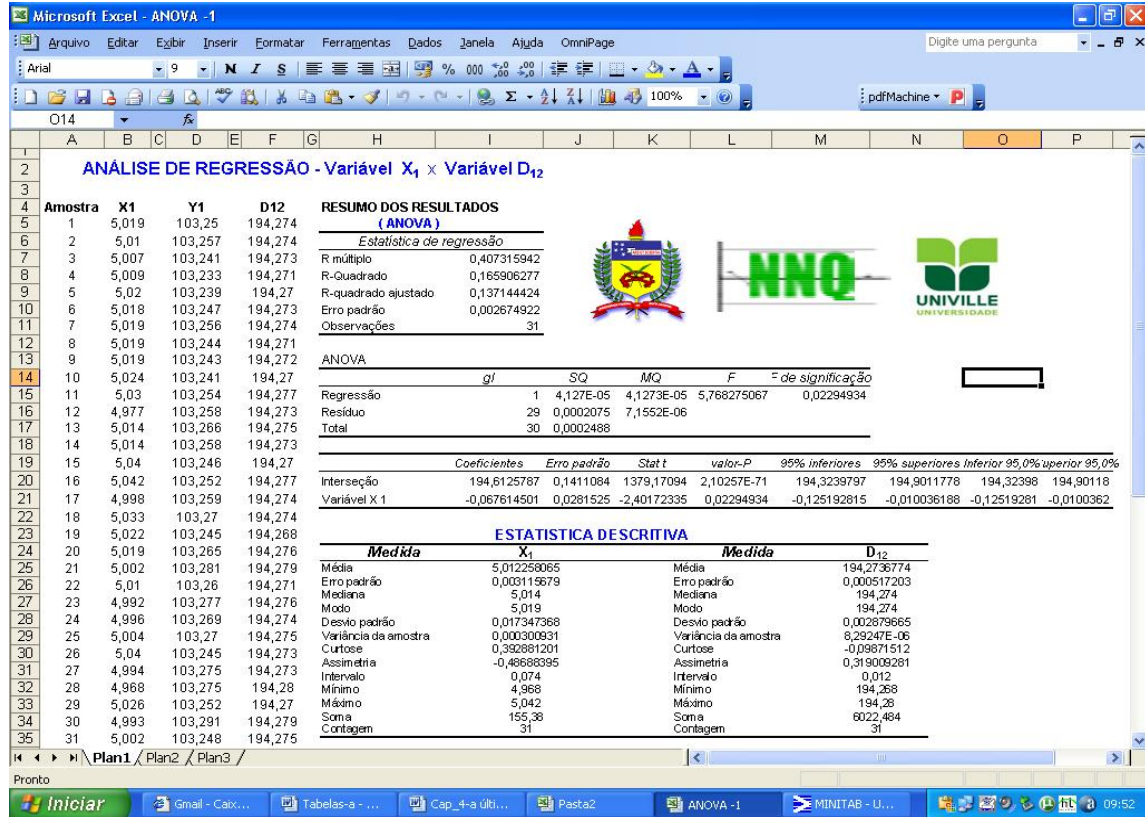


Figura 4.26 Análise de Variância (ANOVA) e Estatística Descritiva para as variáveis X₁ e D₁₂

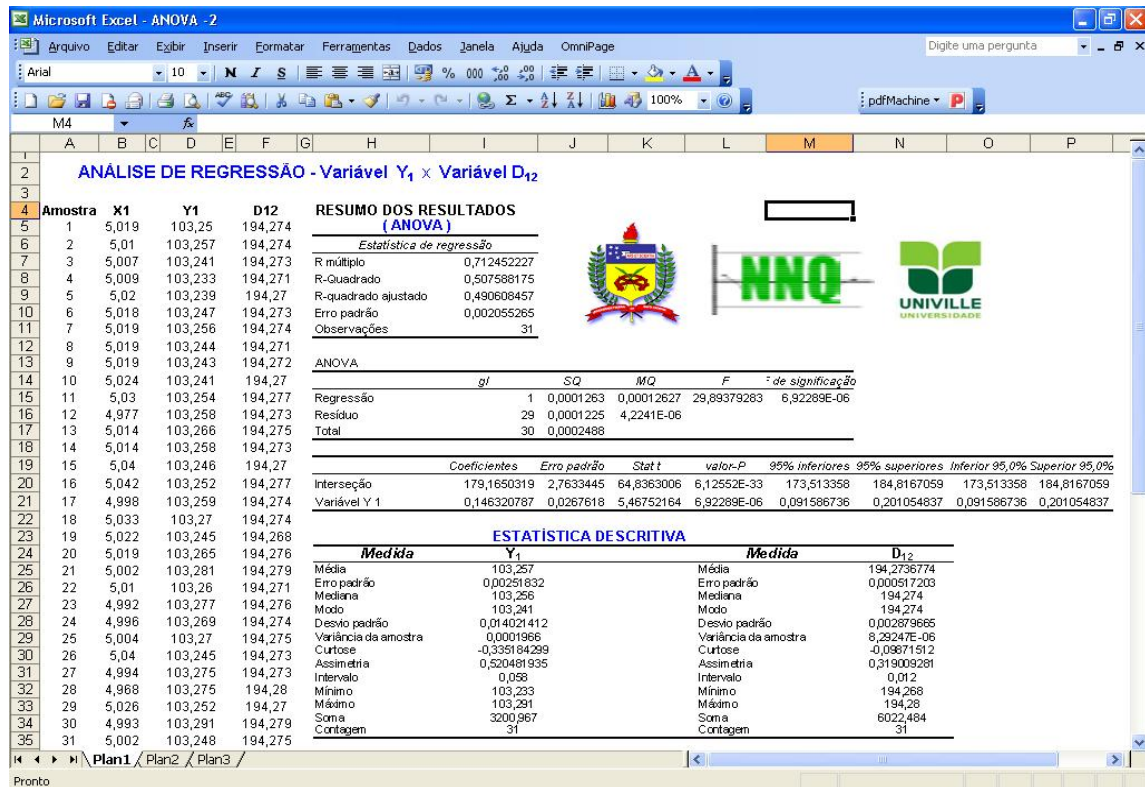


Figura 4.27 Análise de Variância (ANOVA) e Estatística Descritiva para as variáveis Y₁ e D₁₂

Conforme figuras 4.25, 4.26 e 4.27 pode-se verificar os resultados obtidos da Função Autocorrelação (FAC) e o efeito comparativo da análise de correlogramas com a Análise de Variância (ANOVA) que indicam a não existência de autocorrelação nos dados deste processo. Além disso, os resultados obtidos a partir da Estatística Descritiva revelam a estabilidade estatística (inexistência de causas especiais de variação conforme medidas de localização e dispersão obtidas a análise estatística dos dados deste processo).

4.8.3 Projeto ótimo do gráfico MCUSUM em um ponto

O projeto ótimo do gráfico de controle MCUSUM em um ponto para monitorar simultaneamente as características da qualidade X_1 , Y_1 e D_{12} para os dados do processo XYD quando este se encontra sob controle estatístico, o vetor de médias e a matriz de variâncias-covariâncias são

$$\boldsymbol{\mu}_o = \begin{pmatrix} 5 \\ 103,25 \\ 194,27 \end{pmatrix} \quad \Sigma_o = \begin{pmatrix} 0,000300928 & -0,000128501 & -0,000020346 \\ -0,000128501 & 0,000196602 & 0,000028765 \\ -0,000020346 & 0,000028765 & 0,000008291 \end{pmatrix}$$

cujo tamanho da amostra utilizado é $n=1$. Deseja-se, que este processo possa ter a melhor proteção contra uma mudança na distância de Mahalanobis igual a 1, $d=1$. Para este tamanho de mudança, deseja-se um valor de $ARL_o=200$.

$$\boldsymbol{\mu}_1 = \begin{pmatrix} 5,01 \\ 103,26 \\ 194,27 \end{pmatrix}$$

Nestas condições, para $p=3$, $n=1$ e $ARL_o=200$ via programas computacionais desenvolvidos obtém-se a melhor combinação para os parâmetros k e h , ou seja, o par de parâmetros ótimo (k,h) que otimize o gráfico de controle MCUSUM (minimize ARL_d) para os dados deste processo em $d=1$:

$k=0,5$, $h=6,885$ e $ARL_{min.} = 11,0$.

Um comparativo entre o desempenho em termos de ARL_d (ARL mínimo) obtido pelo gráfico de controle MCUSUM para os dados do processo XYD em $d=1$ com os gráficos T^2 de Hotelling e MEWMA para as mesmas condições conforme tabela 4.6.

Tabela 4.6 Desempenho de ARL entre gráficos de controle multivariados

Gráfico de Controle	ARL _d (ARL mínimo)
T ² de Hotelling (*)	52,6
MEWMA (*)	11,3
MCUSUM (*)	11,2
MCUSUM (MCM)	11,2
MCUSUM (MEI)	11,0

Fonte: (*) Simulação (Lowry et al, 1992), MCM (Lee and Khoo, 2006) e MEI (Proposta deste trabalho)

O gráfico MCUSUM conforme tabela 4.6 nas mesmas condições que os demais gráficos multivariados é mais eficiente para detectar este tamanho de mudança, $d=1$. Para essa magnitude de mudança nas condições de $p=3$ características da qualidade para um $ARL_0=200$ conforme revisão de literatura o gráfico MCUSUM é bem sensível do que o gráfico T² de Hotelling, ou seja, é quase cinco vezes mais rápido do que o T² de Hotelling. Este fato justifica a escolha do gráfico MCUSUM para monitorar as características da qualidade deste processo.

Os valores da estatística MCUSUM obtidos com a execução de uma rotina desenvolvida em ambiente R para os dados do processo XYD são utilizados para desenvolver o gráfico MCUSUM deste processo em ambiente MS-Excel conforme figura 4.28.

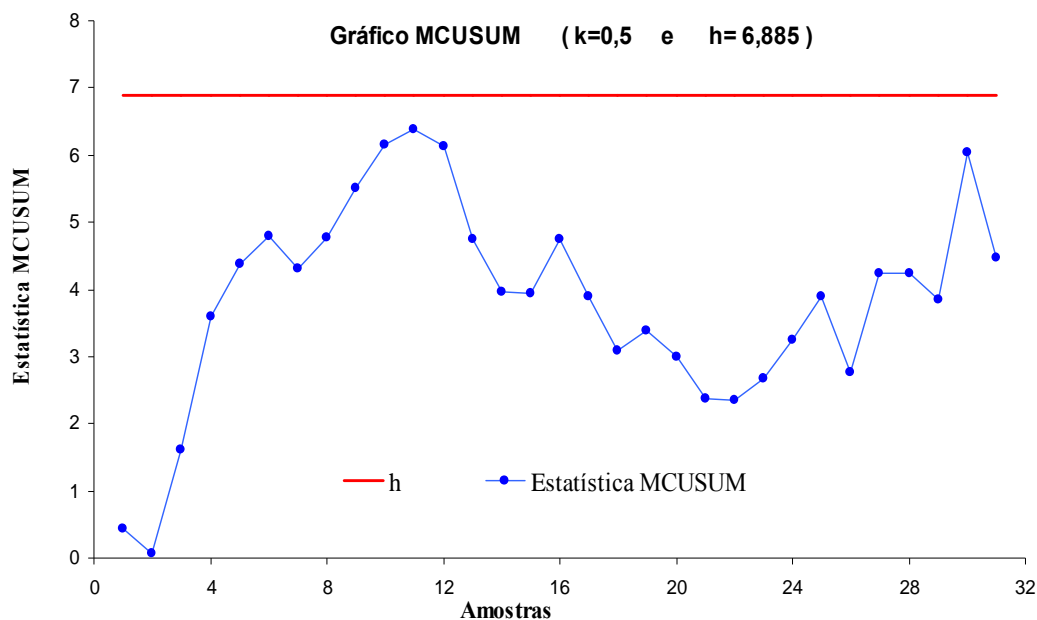


Figura 4.28 Gráfico de controle MCUSUM para os dados do processo XYD

Uma análise preliminar deste processo conforme figura 4.28 revela que, aparentemente, se

trata de um processo estável, apesar de indicar uma tendência inicial ascendente (amostras 3 a 11) e uma queda abrupta (amostras 12 a 14). No entanto, a partir da 15^a amostra observa-se que há uma melhoria do processo cujos pontos demarcados mostram um comportamento aleatório dos valores. Além disso, pode-se verificar que não há presença de causas especiais de variação já que todos os pontos se encontram dentro do limite superior de controle, h .

4.9 Síntese do Capítulo

O presente capítulo nos orienta a extrair as seguintes conclusões:

1. O Método de Equação Integral com Quadratura Gaussiana proposto neste trabalho diante dos resultados obtidos é considerado um excelente método para a otimização de parâmetros do gráfico MCUSUM superando a outros métodos numéricos de otimização existentes.
2. Mediante a execução de um algoritmo específico para o CEP multivariado que inclui a utilização de um programa computacional é possível obter valores dos principais parâmetros (k , h e ARL) que otimizam o gráfico MCUSUM. O programa computacional proposto permite ao usuário obter de forma aproximada os parâmetros ótimos do gráfico MCUSUM em diferentes situações.
3. Este capítulo abre novas possibilidades para otimização de parâmetros com a aplicação do MEI com Quadratura Gaussiana para projetar estatisticamente o gráfico MCUSUM ótimo em situações de difíceis soluções cujo tradicional método de simulações pode não ser uma alternativa viável. O Método de Equação Integral (MEI) é um excelente método de otimização de parâmetros pelo fato de ser um procedimento mais versátil que fornece melhores resultados para o valor de ARL com maior rapidez de cálculo comparado com o tradicional método de simulações. Além disso, a relativa simplicidade de implementação deste método em um programa computacional.

CAPÍTULO 5

CONCLUSÕES

A aplicação de gráficos de controle multivariados para monitorar processos industriais tem aumentado nas últimas décadas motivada pelo crescimento de novas tecnologias de aquisição on-line de dados existente. Hoje, computadores realizam em tempo real, várias mensurações acerca das variáveis de processo. Diante desta realidade, existe um crescente interesse em desenvolver ferramentas mais robustas, que incorporem essa grande massa de dados no monitoramento de processos. Tais ferramentas devem permitir o monitoramento simultâneo de variáveis de interesse considerando sua estrutura de correlação.

No controle estatístico de processos multivariados o principal objetivo é a eliminação de variabilidade e o gráfico de controle multivariado é uma ferramenta eficaz que permite a redução sistemática desta variabilidade nas características da qualidade do produto representadas pelas variáveis monitoradas simultaneamente pelo gráfico. O gráfico de controle multivariado de Somas Acumuladas (MCUSUM) objeto de estudo deste trabalho, por exemplo é uma excelente ferramenta de controle estatístico multivariado utilizado para monitorar a qualidade de um processo. Este gráfico pode ser utilizado com a mesma finalidade, e com vantagens sobre os tradicionais gráficos T^2 de Hotelling quando o objetivo for detectar mudanças de pequena magnitude.

5.1 Desenvolvimento e Aplicação do Método de Equação Integral

O desenvolvimento da metodologia proposta neste trabalho inicia-se com a derivação analítica de uma equação integral cuja resolução numérica através de Quadratura Gaussiana fornece os valores aproximados de ARL do gráfico MCUSUM para um processo supostamente com distribuição normal multivariada. A partir dessa equação derivada analiticamente aplicou-se o Método de Equação Integral com Quadratura Gaussiana no desenvolvimento de um programa computacional em ambiente Matlab para determinar os parâmetros propostos para otimizar o gráfico MCUSUM.

O objetivo principal deste trabalho é a aplicação do Método de Equação Integral com Quadratura Gaussiana para otimizar os parâmetros que envolvem tanto a mensuração do desempenho de ARL quanto a seleção da combinação ótima dos parâmetros k e h que abrange

a mensuração de índices de falsos alarmes e a análise de sensibilidade do gráfico MCUSUM para monitorar processos em diferentes situações.

5.2 Projeto Estatístico Ótimo do Gráfico de Controle MCUSUM

A sugestão de uma metodologia para projetar estatisticamente o gráfico MCUSUM se deve ao fato de optarmos por um gráfico com memória sensível a pequenas mudanças no vetor de médias de um processo monitorado com poucas variáveis cuja vantagem principal é que este gráfico acumula a informação mais recente com informações anteriores e, com isso, detecta pequenas mudanças no vetor de médias do processo multivariado com um ARL bem menor do que faria tradicional gráfico T^2 de Hotelling.

O projeto estatístico ótimo em um ponto do gráfico MCUSUM proposto neste trabalho consiste em determinar os valores ótimos para o par de parâmetros (k, h) capaz de produzir um ARL mínimo para um tamanho de mudança do vetor de médias, d . Os valores tabelados desses parâmetros conforme literatura existente são muito limitados, ou seja, não são suficientes para orientar os usuários a realizarem uma escolha adequada para em situações práticas projetarem estatisticamente o gráfico MCUSUM ótimo. É importante salientar ainda que a implementação de qualquer gráfico de controle que se considere, pressupõe uma escolha adequada dos seus parâmetros, uma vez que a sua eficiência depende em parte desta escolha.

5.3 Projeto Otimização do MCUSUM com Regiões de Máxima e Mínima Potência

Neste trabalho, a análise estatística multivariada dos dados termina com a segunda fase da metodologia, ou seja, a fase do experimento computacional que nos permite analisar a influência das diferentes combinações dos parâmetros k e h que associadas ao ARL interferem significativamente na otimização do gráfico MCUSUM. Por último, uma análise de sensibilidade da solução ótima desses parâmetros nas diferentes situações nos levou a extrair as conclusões finais da investigação.

O próximo passo dessa segunda fase do trabalho é a sugestão do desenvolvimento de um procedimento alternativo com análise de regiões de máxima e mínima potência para o gráfico MCUSUM ainda não existente na literatura e que se constitui numa extensão multivariada do procedimento proposto por Woodall(1985) para gráficos de controle de qualidade univariados. Este procedimento de projetar estatisticamente o gráfico MCUSUM sob a ótica de regiões de máxima e mínima potência é uma etapa alternativa e complementar ao algoritmo de projetar estatisticamente o gráfico MCUSUM proposto por Lee e Khoo

(2006). Um exemplo de aplicação para o projeto estatístico ótimo do gráfico MCUSUM baseado nos valores de ARL, k e h e sob a ótica de regiões de máxima e mínima potência e com a inclusão desta etapa complementar ao algoritmo do projeto estatístico do gráfico MCUSUM ótimo conforme seção 4.5 deste trabalho. Para implementar este novo procedimento foi desenvolvido em ambiente Matlab o Programa Computacional 3 que além de determinar a matriz de covariância e sua inversa, a matriz de correlação, a distância de Mahalanobis, d e os coeficientes de assimetria e curtose que fundamentam os testes hipóteses sugeridos por Mardia (1970, 1974) para verificar a normalidade multivariada. Além disso, determina e gera graficamente a função perda multivariada de Taguchi que é utilizada para determinar os pontos A e B que delimitam as regiões de máxima e mínima potência. De posse desses resultados é formulado um modelo matemático em ambiente MS-Excel onde com a utilização da ferramenta solver determina-se estes pontos A e B, ou seja, a distância de Mahalanobis, d em cada um destes pontos.

A metodologia proposta neste trabalho contempla a aproximação sistemática de projetar estatisticamente o gráfico MCUSUM ótimo baseado nos parâmetros ARL, k , e h segundo filosofia de análise de regiões de máxima e mínima potência. Isso, é relevante por exemplo, quando o processo é muito capaz ou dificilmente ajustável, onde não é interessante detectar mudanças de pequena magnitude mas é fundamental sobretudo decidir que tamanho de mudança é realmente importante detectar. Uma alternativa que aborda tanto a determinação da magnitude de mudança que desejamos detectar e a que não desejamos detectar tem sido o enfoque de Taguchi. Métodos práticos desenvolvidos com a utilização da função perda multivariada de Taguchi por exemplo, determinam os valores dos pontos que delimitam as regiões de máxima e mínima potência do gráfico MCUSUM.

O projeto ótimo do gráfico MCUSUM com regiões de máxima e mínima potência é significativo, pois teoricamente sabe-se que o estudo comparativo da eficiência entre diferentes gráficos de controle é realizado em termos de ARL. Assim, dois gráficos de controle serão comparáveis se ambos apresentarem o mesmo ARL quando o processo se encontra sob controle, isto é, quando não existe nenhuma mudança na média da variável controlada. Diante dessas considerações, o gráfico de controle mais eficiente ou de maior potência para detectar uma determinada mudança, d é aquele que apresenta um menor valor de ARL. Um bom projeto de controle é aquele que apresenta um elevado valor de ARL_o (quando o processo se encontra sob controle) e um mínimo ARL_d (quando o processo se encontra fora de controle). Por outro lado, sabe-se que do ponto de vista econômico considera-se que o custo de operação

do processo é proporcional ao número de falsos alarmes, e do ponto de vista estatístico pode ocorrer que o número de falsos alarmes seja uma fonte de variabilidade extra no processo.

5.4 Recursos Computacionais Aplicados e Parâmetros Ótimos obtidos através do MEI

Os recursos computacionais aplicados neste trabalho que envolvem a otimização de parâmetros, análise estatística multivariada e apresentação de dados destacam-se os três principais programas desenvolvidos em ambiente Matlab. Além disso, planilhas desenvolvidas em ambiente MS-Excel com a ferramenta solver e o desenvolvimento de rotinas computacionais no pacote GNU R.

Os programas computacionais desenvolvidos no ambiente Matlab são implementados para o experimento computacional realizado nos últimos meses no laboratório de informática da Univille. Para esse experimento aplicou-se o Método de Equação Integral com Quadratura Gaussiana para as combinações de ARL_o de 200, 500 ou 1000 com $p=2, 3$ e 4 para determinar aproximadamente a solução ótima dos parâmetros propostos neste trabalho que otimize o gráfico MCUSUM. Os resultados obtidos nesse experimento computacional desses parâmetros para vários tamanhos de mudança, d do gráfico MCUSUM foram tabelados conforme tabelas 4.1, 4.2 e 4.3 (Capítulo 4). Os valores desses parâmetros obtidos através do MEI e mostrados nessas tabelas apresentam uma certa similaridade com os valores obtidos através do MCM. Um estudo comparativo entre os valores desses parâmetros para (ARL_o de 200 com $p=2, 3$ e 4 e $n=1$) para vários tamanhos de mudança, d do gráfico MCUSUM obtidos através do MCM (Lee e Khoo, 2006) e os obtidos através do MEI (Proposta deste trabalho) é mostrado conforme tabela 4.4 (Capítulo 4). Os valores dos parâmetros mostrados nessa tabela revelam que há uma pequena diferença entre os valores obtidos através do MCM e os valores obtidos através do MEI. Além disso, uma análise de sensibilidade para vários tamanhos mudança, d é realizada a partir dos valores dos parâmetros estimados com interpolação polinomial para um ARL_o de 200 quando $p=2$ e $n=1$ para averiguar a partir dessas mesmas condições a situação quando o processo estiver fora de controle. Os valores estimados se mostraram satisfatórios cuja diferença é algo em torno de 3 % ou menos conforme tabela 4.5 (Capítulo 4).

Uma representação gráfica de curvas de ARL do gráfico MCUSUM (em escala logarítmica) dos valores de ARL obtidos através do MEI para um ARL_o de 200, 500 ou 1000 com $p= 2, 3$ e 4 e $n=1$ é ilustrada conforme figura 4.1 (Capítulo 4). Além disso, uma representação gráfica de ARL em escala logarítmica do desempenho de ARL quando $p=2, n=1$ e $ARL_o = 200$ dos gráficos multivariados com memória MCUSUM e MEWMA e o gráfico T^2

de Hotelling conforme figura 4.2 (Capítulo 4). Os resultados obtidos do desempenho de ARL conforme esta figura indicam que os métodos MCM e MEI apresentam melhores valores de ARL que os tradicionais métodos de simulação. Essa representação gráfica revela ainda o intervalo da mudança do vetor de médias d mais apropriado para otimização do gráfico MCUSUM.

O desempenho de ARL especificamente para o gráfico MCUSUM obtido pelos métodos numéricos: Simulação (SIM), Cadeias de Markov (MCM) e Equação Integral (MEI) para um ARL_0 de 200 com $p=2, 3$ e 4 e $n=1$ é ilustrado conforme figura 4.3 (Capítulo 4) indicando que os valores de ARL obtidos através dos métodos MCM e MEI são similares e são melhores que os valores obtidos através método SIM para um tamanho de mudança, $d > 1$.

5.5 Aplicação da Metodologia e Resultados Obtidos com a Utilização de Dados Reais

A aplicação da metodologia proposta utilizando dados reais de um processo de usinagem da Fundação Tupy Ltda localizada em Joinville-SC foram de extrema relevância neste trabalho pois serviu como referencial para comprovar a validade prática da metodologia proposta.

Na aplicação dos dados reais foram avaliados dois processos de usinagem. O primeiro deles, o processo XY (furação de acabamento do furo 1) envolve duas características da qualidade que são monitoradas simultaneamente X_1 (posicional do furo 1 em relação a coordenada x) e Y_1 (posicional do furo 1 em relação a coordenada y). O segundo processo, XYD (distância entre os centros dos furos 1 e 2) é composto de três características da qualidade que são monitoradas simultaneamente, ou seja, além das duas primeira variáveis X_1 e Y_1 é incluída uma 3ª variável, D_{12} . Definido os dois processos a serem monitorados realizou-se a exploração dos dados para aplicar a metodologia proposta. Para tais dados verificou-se a suposição de normalidade, a autocorrelação bem como a estabilidade estatística. Para compilar estes dados reais utilizou-se também os programas computacionais propostos neste trabalho. No projeto estatístico ótimo do gráfico MCUSUM desses dois processos avaliados utilizou-se os principais parâmetros otimizados através do a Método de Equação Integral com Quadratura Gaussiana. O gráfico de controle MCUSUM para a análise estatística multivariada dos dados reais de cada um dos dois processos analisados foi gerado a partir do desenvolvimento de rotina no pacote GNU R conforme seção 4.7.4 (telas do R para carregar e executar a rotina) e Apêndice J (codificação computacional das rotinas desenvolvidas em ambiente R).

O Método de Equação Integral com Quadratura Gaussiana proposto neste trabalho representa uma nova alternativa de otimização de parâmetros para projetar estatisticamente o gráfico MCUSUM ótimo em situações de difíceis soluções cujo tradicional método de simulação pode não ser viável. Os resultados obtidos com a aplicação desta metodologia para o gráfico MCUSUM revelam o Método de Equação Integral (MEI) como um excelente método numérico de otimização de parâmetros pelo fato de ser um procedimento mais versátil que fornece melhores resultados para o valor de ARL com maior rapidez de cálculo comparado com o tradicional método de simulação. Além disso, a relativa simplicidade de implementação deste método em um programa computacional.

5.6 Recomendações para Investigações Futuras

A seleção de parâmetros para o projeto estatístico de um gráfico de controle envolve propriedades estatísticas tais como as probabilidades de erro tipo I e tipo II, o ARL e o ATS, ou seja, para esta seleção se devem levar em consideração aspectos estatísticos como a minimização do número de falsos alarmes e a maximização da capacidade de detecção de mudanças reais. Além disso, aspectos econômicos como o custo de amostragem, perdas de produção de produtos de baixa qualidade e custos de localização de causas especiais quando o diagnóstico das causas assinaladas indicam um desvio do processo.

Conforme revisão de literatura sobre projeto econômico-estatístico é ampla a quantidade de trabalhos direcionados aos gráficos de controle univariados. No entanto, pouca atenção tem sido dedicada ao projeto econômico-estatístico que envolvem gráficos multivariados. Alguns trabalhos recentes abordam o projeto econômico do gráfico de controle multivariado MEWMA, com o objetivo de minimizar o custo de operação deste gráfico. Linderman e Love (2000a. e 2000b) propõem um modelo econômico tendo como base o modelo de custo unificado de Lorenzen e Vance (1986) que incluem restrições tanto estatísticas quanto econômicas.

Como recomendação de investigação futura propõe-se o desenvolvimento de um projeto econômico-estatístico sob a ótica de regiões de máxima e mínima potência para o gráfico MCUSUM ainda não existente na literatura. Este modelo, deve contemplar tanto as restrições de modelos estatísticos que nos permita obter gráficos MCUSUM com boas propriedades estatísticas quanto as restrições econômicas capazes de reduzir custos de operação. Com esse projeto é possível quantificar o projeto estatístico proposto neste trabalho para o gráfico MCUSUM e que de fato ele seja economicamente motivado. Para isso, sugere-

se um modelo econômico para o gráfico MCUSUM segundo a filosofia de análise de regiões de máxima e mínima potência para obter-se planos de amostragem ótimos que proporcionem custos médios de operação por unidade de tempo mínimos para satisfazer restrições estatísticas impostas para o modelo em forma de ARL. Diante dessas considerações, este modelo proposto deve estar sujeito as restrições de um mínimo valor de ARL sob controle (ARL_o) e um máximo valor de ARL quando o processo se encontra fora de controle (ARL_d). Alternativamente, pode-se utilizar restrições em termos de ATS (Average Time to Signal), que expressa o tempo médio esperado até o sinal). Este projeto econômico-estatístico deve ter como objetivo a otimização do gráfico de controle MCUSUM através do Método de Equação Integral com Quadratura Gaussiana que minimizam o custo incluso no processo de produção. Tipicamente, este custo inclui o custo de amostragem, o custo de não qualidade devido à produção de produto não conforme e o custo de investigação de situação fora de controle que realmente são falsos alarmes.

REFERÊNCIAS BIBLIOGRÁFICAS

ALVES, C.C. **Gráficos de controle CUSUM: um enfoque dinâmico para a análise estatística de processos.** Dissertação de Mestrado em Engenharia de Produção, UFSC, Florianópolis, 2003.

ALVES, C.C. e SAMOHYL, R.W. **A utilização dos gráficos de controle CUSUM para o monitoramento de processos industriais.** XXIV Encontro Nacional de Engenharia de Produção - Florianópolis, SC, Brasil, 2004.

ALVES, C.C., HENNING, E. e SAMOHYL, R.W. **O desenvolvimento de gráficos de controle MCUSUM e MEWMA em ambiente R como um procedimento alternativo para análise estatística de processos multivariado.** XXIVIII Encontro Nacional de Engenharia de Produção – Rio de Janeiro, RJ, Brasil, 2008a.

ALVES, C.C., HENNING, E. e SAMOHYL, R.W. **Análise de dados multivariados em ambiente R: uma proposta para o gráfico de controle MEWMA baseado na projeção de dados via ACP.** XV Simpósio de Engenharia de Produção, Bauru, SP, Brasil, 2008b.

ALWAN, L.C. **Cusum Quality Control. – Multivariate Approach.** Communications in Statistics, 15 (12), 3531-3543, 1986.

APARISI, F. and DÍAZ, J.C.G. **Design and optimization of EWMA control charts for in-control, indifference, and out-of-control regions.** Computers & Operations Research 34, pp. 2096–2108, 2007.

BERSIMIS, S., PSARAKIS, S. and PANARETOS, J. **Multivariate Statistical Process Control Charts: An Overview.** Quality and Reliability Engineering International, 23, pp.517-543, 2007.

BILLO, J.B. **Excel for Scientists and Engineers – Numerical Methods.** Wiley-Interscience, 2007.

BROOK, D. and EVANS, D.A. **An Approach to the Probability Distribution of Cusum Run Length.** Biometrika, v.59, p.539-549, 1972.

CAMPOS FILHO, F.F. **Algoritmos Numéricos.** 1ª ed., LTC Editora, 2001.

CANALE R. P., CHAPRA S.C. **Métodos Numéricos para Engenharia.** Editora McGraw-Hill Interamericana, 5ª ed., 2008.

CHAMP, C.W. and RIGDON S.E. **A comparison of the Markov Chain and the integral equation approaches for evaluating the run length distribution of quality control Chart.** Comun. Statist. 20(1) 191-204, 1991.

CHAPMAN, S.J. **Programação em Matlab para Engenheiros.** Editora Pioneira Thomson Learning, 2003.

- CHENG S.W. and THAGA,K. **Max-Cusum chart for autocorrelated processes**. *Statistica Sinica*, v.15, pp.527-546, 2005.
- CHOU, C.Y, LIU, H.R. and HUANG, X.R. **Economic-Statistical Design of Multivariate Control Charts Using Quality Loss Function**. *Advanced Manufacturing Technology*, 20, pp.916-924, 2002.
- COLLIER, N. and KAW,A. **Department of Mechanical Engineering, University of South Florida Tampa**, <http://numericalmethods.eng.usf.edu/>, 2008 (Acessado em 28/12/2008).
- CROSIER, R.B. et al. **A New Two-Sided Cumulative Sum Control Scheme**. *Technometrics*, 28, p.187-194, 1986.
- CROSIER, R.B. et al. **Multivariate Generalizations of Cumulative Sum Quality-control Schemes**. *Technometrics*, 30 (3), p.291-303, 1988.
- CROWDER, S.V. **Design of Exponentially Weighted Moving Average Schemes**. *Journal of Quality Technology*, 21 (3) p.155-162, 1989.
- D'AGOSTINO, R. and STEPHENS, M.A. **Godness-of-Fit Techniques**. New York, Marcel Dekker, 1986.
- FERRER, A. **Técnicas estadísticas multivariantes para el control estadístico de procesos altamente automatizados**. III Verano Estadística Industrial CIMAT, Guanajuato (MX), 2005 http://www.cimat.mx/Eventos/veicep/CIMAT_MSPC_Ferrer.pdf Acessado em 16/12/2006.
- GAN, F. F. **An optimal design of CUSUM quality control charts**. *Journal of Quality Technology*, 23, 279–286, 1991.
- GAN, F.F. **An optimal design of cumulative sum control charts based on median run length**. Research Report, N° 536, Department of Mathematics, National University of Singapore, 1992.
- GAN, F. F. **The Run Length of a Cumulative Sum Control Chart.**, *Journal of Quality Technology*, 25, 205-215, 1993.
- GOEL, A.L. and WU, S.M. **Determination of ARL and a Contour Nomogram for Cusum Charts to Control Normal Mean**, *Technometrics*, Vol.13, p.221-230, 1971.
- HAWKINS, D.M. **Multivariate Quality control Based on Regression-Adjusted**. *Journal of Technometrics*, 33(1), 61-75, 1991.
- HAWKINS, D.M. and OLWELL, D.H **Cumulative Sum Charts and Charting for Quality Improvement**, *Statistics for Engineering and Physical Science*, Springer, 1998.
- HEALY, J.D. **A note on Multivariate CUSUM Procedures**. *Technometrics*, 29(4), p.409-412, 1987.
- HOLMES, D. S. & MERGEN, A. E. **Improving the Performance of the T2 Control Chart**. *Quality Engineering*, v.5, p.619-625, 1993.

JARRET, J.E and PAN, X. **The quality control chart for monitoring multivariate autocorrelated process.** Computational Statistics and Data Analysis, v.51, pp. 3862-3870, 2007

JEANNEQUIN, N. **Quadrature.** Dissertation, Oxford University Department of Comlab, 2003.

JONES, A.L., CHAMP, C.W. and RIGDON, S.E. **The Run Length Distribution of the CUSUM with Estimated Parameters.** Journal of Quality Technology, 36(1), 95-108, 2004.

JUN, Chi-Houck and CHOI, M.S. **Simulating the average run length for CUSUM schemes using variance reduction techniques,** Communications in Statistics B: Simulation, 22(3), p.877-887,1993.

KALGONDA,A.A. and KULKARN, S.R. **Multivariate quality control for autocorrelated Process.** Journal of Applied Statistics, 31(3), 317-327, 2004.

KAPUR, K.C. and CHO, B. **Economic design of the specification region for multiple quality characteristics.** IIE Transactions, 28, 237-248, 1996.

KARRIS, S.T. **Numerical Analysis using Matlab and Excel.** Orchard Publications, 3^a ed., 2007.

KEMP, K.W. **A simple procedure for determining upper and lower limits for the average sample run length of a cumulative sum scheme.** Journal of the Royal Statistical Society, Series B, 23, Methodological, Vol.29, p.263-265,1967.

KEMP, K. W. **Formal Expressions Which can be Applied to Cusum Charts, With Discussion,** Journal of The Royal Statistical Society, SER B, v.33, p.331-360, 1971.

KALGONDA,A.A. and KULKARN, S.R. **Multivariate quality control for autocorrelated Process.** Journal of Applied Statistics, 31(3), 317-327, 2004.

KIUSALAAS, Jann **Numerical Methods in Engineering with MATLAB.** Cambridge University Press, 2005.

LEE, M.H. and KHOO, M.B. **Optimal Statistical Designs of a Multivariate Cusum Chart based on ARL and MRL** International Journal of Reliability, Quality and Safety Engineering, Vol. 13, No.5, pp. 479–497, 2006.

LINDERMAN, K. and LOVE, T.E. **Implementing Economic and Economic Statistical Designs for MEWMA Charts.** Journal of Quality Technology, 32(4), 2000.

LINDERMAN, K. and LOVE, T.E. **Economic and Economic Statistical Designs for MEWMA Control Charts.** Journal of Quality Technology, 32(4), 2000.

LORENZEN, T.J. and VANCE, L.C. **The Economic Design of Control Charts: A Unified Approach..** Technometrics, 28, p.3-10, 1986.

LOWRY, C.A., WOODALL, W.H.; CHAMP, C.W. & RIGDON, S.E. **A Multivariate Exponentially Weighted Moving Average Control Chart**. *Technometrics*, 34(1), p.46-53, 1992.

LOWRY, C. A. & MONTGOMERY, D. C. **A Review of Multivariate Control Charts**. *IIE Transactions*, v. 27, p. 800-810, 1995.

LUCAS, J. M. **The Design and Use of Cumulative Sum Quality Control Schemes**. *Journal of Quality Technology*, v. 8, n° 1, p. 1-12, 1976.

LUCAS, J.M. and CROSIER, R.B. **Fast Initial Response for CUSUM Quality-Control Schemes: Give Your CUSUM a Head Start**. *Technometrics*, 24(3), p.199-205, 1982.

LUCENÑO, A. and PUG-PEY, J. **Evaluation of the Run Length Probability Distribution for CUSUM Charts assessing Chart Performance**. *Technometrics*, 42, pp.411-416, 2000.

LUCENÑO, A. and PUG-PEY, J. **Computing the Run Length Probability Distribution for CUSUM Charts**. *Journal of Quality Technology*, 34, pp.209-215, 2002.

LYSHEUSKI, S.E. **Engineering and Scientific Computation using Matlab**. Willey-Interscience, 2003.

MARDIA, K.V. **Measures of multivariate skewness and kurtosis with applications**. *Biometrika*, 57(3), pp. 519-539, 1970.

MARDIA, K.V. **Applications of some measures of multivariate skewness and kurtosis in testing normality and robustness studies**. *The Indian Journal of Statistics*, V. 36, series B, pp. 115-128, 1974.

MARTINEZ, W.L and MARTINEZ, A.R. **Computational Statistics Handbook with Matlab**. Chapman & Hall/CRC, Washington, 2002.

MASON, R. L., TRACY, N. D. and YOUNG, J. C. **A practical approach for interpreting multivariate T2 control chart signals**. *J. Qual. Technol.* 29:(4) , pp. 396-406, 1997.

MASON, R.L. and YOUNG, J.C. **Multivariate Statistical Process Control With Industrial Applications**. Philadelphia: ASA-SIAM, PA, 2002.

MINGOTI, S.A. **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada**. Belo Horizonte: Editora UFMG, 2005.

MONTGOMERY, D. C. **Introdução ao Controle Estatístico da Qualidade**. 4ª Ed. LTC, 2004.

NEYMAN, J. and PEARSON, E.S. **On the Problem of the Most Efficient Tests of Statistical Hypotheses**. *Philosophical Transactions Royal Society Series A*, pp.289-337, 1933.

NGAI, H. and ZHANG, J. **Multivariate Cumulative Sum Control Charts based on Projection Pursuit** *Statistica Sinica*, 11, 747-766, 2001.

NOOROSSANA, R. and VAGHEFI, S.J.M. **Effect of Autocorrelation on performance of the MCUSUM control chart.** Quality and Reliability Engineering International, V.22, pp.191-197, 2006.

PAGE, E.S. **Continuous Inspection Schemes,** Biometrika, v.41, p.100-115, 1954.

PAGE, E.S. **Cumulative Sum Charts,** Technometrics, 3, p.1-9, 1961.

PAO, Y.C. **Engineering Analysis Interactive Methods and Programs with FORTRAN, QuickBASIC, MATLAB, and Mathematica,** CRC Press, Washington, 2001.

PEÑA, D. **Análisis de Datos Multivariantes.** 1ª Ed. Interamericana de España, McGraw Hill, 2002.

PIGNATIELLO, J.J., Jr., and RUNGER, G.C. **Comparisons of Multivariate CUSUM charts.** Journal of Quality Technology, 22(3), p.173-186, 1990.

PHAM, H. **Springer Handbook of Engineering Statistics.** Rutgers the State University of New Jersey, 2006.

QIU, P. and Hawkins, D. **A nonparametric multivariate CUSUM procedure for detecting shifts in all direction.** JRSS-D (The Statistician) 52(2), p.151-164, 2003.

QIU, P. and Hawkins, D. **A rank based multivariate CUSUM procedures.** Technometrics, v.43, p.120-132, 2001.

RAO, B.V, DISNEY,R.L and PIGNATIELLO, J.J. **Uniqueness and convergence of solutions to average run length integral equations for cumulative sum and other control charts.** IIE Transactions,33, pp.463-469, 2001.

RENCHER, A.C. **Methods of Multivariate Analysis** 2ª ed., Wiley-Interscience, New York, 2002.

RIGDON, S.E. **An integral equation for the in-control average run length of a multivariate exponentially weighted moving average control chart** Journal Statist.Comput.Simul. V.52, pp. 351-365, 1995.

ROSS, S.M. **Variance Reduction in simulation via random hazards.** Probability in the Engineering and Informational Sciences, v. 4, pp. 299-309, 1990.

RUNGER G. C and TESTIK M.C., **Multivariate extensions to cumulative sum control charts,** Quality and Reliability Engineering International 20, pp.587–606, 2004.

SALLEH, S., ZOMAYA, A.Y. and BAKAR, S.A. **Computing for numerical methods using Visual C⁺⁺.** Wiley-Interscience Publication, 2007.

SOARES, A.A. **O índice de capacidade multivariado como instrumento para avaliação em uma operação de usinagem.** Dissertação de Mestrado em Engenharia de Produção,UFSC, Florianópolis, 2006.

SOUZA, A.M. **Monitoração e ajuste de realimentação em processos produtivos multivariados**. Tese de Doutorado em Engenharia de Produção, UFSC, Florianópolis, 2000.

SULLIVAN, J. H. & WOODALL, W. H. **Comparison of Multivariate Control Charts for Individual observations**. *Journal of Quality Technology*, 28(4), p. 398-408, 1996.

TEERAVARAPRUG, J. and CHO, B.R. **Designing the optimal process target levels for multiple quality characteristics**. *International Journal of Production Research*, 40(1), 37-54, 2002.

TRACY, N. D., YOUNG, J. C. & MASON, R. L. **Multivariate Control Charts for individual Observations**. *Journal of Quality Technology*, v. 24, n° 2, p. 88-95, 1992.

VANCE, L.C. **Average Run Lengths of Cumulative Sum Control Charts for Controlling Normal Means**. *Journal of Quality Technology*, 18, 189-193, 1986.

WALD, A. **Sequential tests in industrial statistics**. *Ann Math Stat.*, v.6 pp117-186, 1945.

WALD, A. (1947). **Sequential Analysis**, Wiley, New York., 1947.

WALDMANN, K.H. **Bounds for the Distribution of the Run Length of one-sided and Two-sided CUSUM Quality Control Schemes**. *Technometrics*, v.28, p.61-67, 1986.

WOODALL, W.H. **The Statistical design of quality control charts** *The Statistician*, V. 34 pp.155-160, 1985.

WOODALL, W.H. **On the Markov Chain Approach to the Two-Sided CUSUM Procedure**, *Technometrics*, 26, 41-46, 1984.

WOODALL, W.H. and NCUBE, M.M. **Multivariate CUSUM Quality Control Procedures** *Technometrics*, 27(3), p.285-292, 1985.

WOODALL, W.H.. **The Design of CUSUM Quality Control Charts**, *Journal of Quality Technology*, 18(2) p.99-102, 1986.

WOODALL, W.H.. **Conflicts Between Deming's Philosophy and the Economic Design of control charts**. *Frontiers in Statistical Quality Control*, Physica-Verlag, Viena, 1987.

YANG, K and TREWN, J. **Multivariate Statistical Methods in Quality Engineering**. New York, Mcgraw-Hill, 2004.

APÊNDICE A - Integração Gaussiana por Polinômios de Legendre

Um método alternativo de integração Gaussiana baseado em uma aproximação de polinômios de Legendre (polinômios ortogonais) é aplicado para determinar mais facilmente os termos de ponderação w_i e x_i , principalmente para $n > 2$. Esta alternativa se fundamenta em mostrar o integrando da função cuja integral se deseja encontrar, como valores que representam raízes de polinômios ortogonais.

Um conjunto de polinômios de Legendre de grau n $\{P_0(x), P_1(x), \dots, P_n(x), \dots\}$ é dado sob a forma de uma equação diferencial ordinária de ordem n denominada de fórmula de recorrência de Rodrigues (Salleh, Zomaya e Bakar, 2007).

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n \quad (\text{A.1})$$

Os cinco primeiros polinômios de Legendre conforme equação (A.2) obtidos a partir da equação diferencial ordinária de ordem n (A.1) e sua representação gráfica conforme figura (A.1):

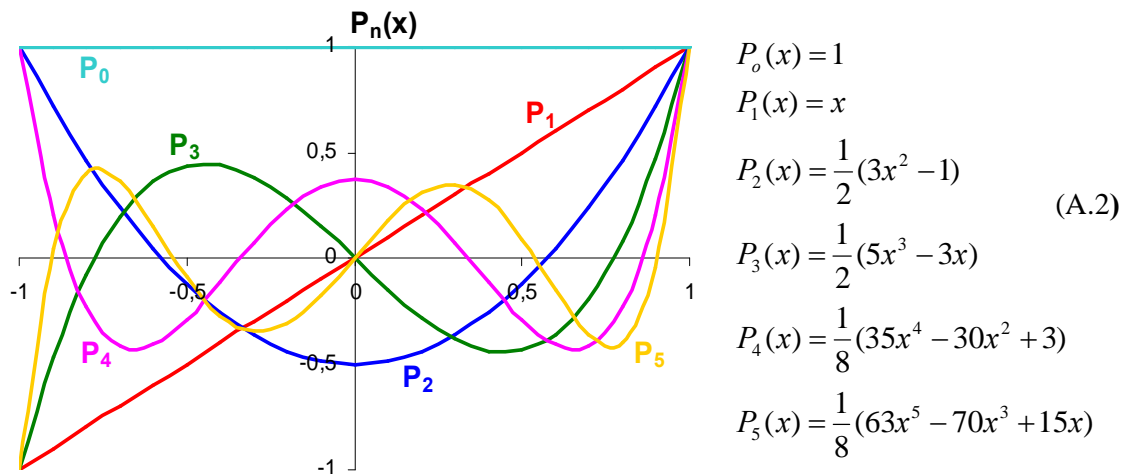


Figura A.1 Representação gráfica dos cinco primeiros polinômios de Legendre

O propósito deste método alternativo via polinômio de Legendre é discutir a fórmula de integração Gaussiana que aproxima $\int_{-1}^1 f(x) dx$ e mostrar que com uma simples troca de

variável é possível estender os limites de integração a valores $[-1,1]$. Assim, a aproximação da integral definida pode ser expressa como

$$\int_{-1}^1 f(x)dx = w_1 f(x_1) + w_2 f(x_2) + w_3 f(x_3) + \dots + w_i f(x_i) = \sum_{i=1}^n w_i f(x_i) \quad (A.3)$$

onde w_1, w_2, \dots, w_n são os pesos (coeficientes de ponderação) e x_1, x_2, \dots, x_n são as raízes (nós) do polinômio de Legendre $P_n(x)$. O problema consiste em encontrar as $(2n)$ constantes (w_i, x_i) . Para determinar estas constantes, partimos da suposição básica de que qualquer polinômio de Legendre resultante da fórmula de recorrência de Rodrigues representa exatamente, um polinômio de ordem $2n - 1$ ou menor.

A aplicação deste método alternativo de integração Gaussiana para determinar os pontos de quadratura (nós) consiste em considerarmos $P_n(x_1) = 0$; $P_n(x_2) = 0$; $P_n(x_3) = 0$;; $P_n(x_n) = 0$, o que implica que os pontos de quadratura $x_1, x_2, x_3, \dots, x_n$ são as raízes do polinômio de Legendre $P_n(x)$. Assim, para $P_n(x) \in [-1,1]$ existem n raízes reais distintas. Por exemplo, para $n = 2$, $P_n(x) = P_2(x) = \frac{1}{2}(3x^2 - 1) = 0$, cujas raízes são

$$x = \pm \frac{1}{\sqrt{3}} = \pm 0,577350269189626, \text{ enquanto que para } n=3, P_n(x) = P_3(x) = \frac{1}{2}(5x^3 - 3x) = 0,$$

cujas raízes são $x = 0, x = \pm \sqrt{\frac{3}{5}} = \pm 0,774596669241483$. Este procedimento pode estender-se para diferentes valores de n , ou seja, $n = 4, 5, 6, \dots$

No entanto, para a determinação de w_i coeficientes de ponderação (pesos) novamente levamos em consideração o requisito estabelecido em (A.3), isto é, se o integrando $f(x)$ é um polinômio de grau $2n - 1$ ou menor, a equação não envolve uma aproximação. Além disso, o polinômio de Lagrange para aproximar qualquer polinômio $h(x)$ de grau $n-1$, que passa por n pontos $x_1, x_2, x_3, \dots, x_n$ pode ser expresso como:

$$h(x) = \sum_{i=1}^n h(x_i) L_i(x) \quad (A.4)$$

Então $\int_{-1}^1 h(x) dx = \int_{-1}^1 \sum_{i=1}^n h(x_i) L_i(x) dx$. Sabendo que $h(x_i)$ é uma constante temos,

$$\int_{-1}^1 h(x) dx = \sum_{i=1}^n h(x_i) \int_{-1}^1 L_i(x) dx \quad (A.5)$$

Comparando (A.3) com (A.5), teremos:

$$w_i = \int_{-1}^1 L_i(x) dx \quad i = 0, 1, 2, 3, \dots \quad (\text{A.6})$$

É comum encontrar a definição L_i e portanto de w_i em termos de polinômio de Legendre, conforme a seguir:

O polinômio $\frac{P_n(x)}{x - x_i} = 0$ para todo $x = x_j$, $j = 1, 2, \dots, n$, porém $j \neq i$

De acordo com a Regra de L'Hôpital

$$\lim_{x \rightarrow x_i} \frac{P_n(x)}{x - x_i} = \left[\frac{\frac{dP_i(x)}{dx}}{\frac{d(x - x_i)}{dx}} \right]_{x=x_i} = \frac{dP_i(x_i)}{dx} = P'_n(x_i) \quad (\text{A.7})$$

Sabendo que a derivada do denominador é igual a 1 e que x_i é uma das raízes do polinômio de Legendre $P_n(x) = 0$. Assim, o polinômio de Lagrange conforme (A.4) pode ser expresso

$$\text{agora como } L_i = \frac{1}{P'_n(x_i)} \int_{-1}^1 \frac{P_n(x)}{x - x_i} dx \quad (\text{A.8})$$

Por exemplo, para $n = 2$,

$P_n(x) = P_2(x) = \frac{1}{2}(3x^2 - 1)$ cujas raízes são $x_1 = \frac{-1}{\sqrt{3}}$, $x_2 = \frac{1}{\sqrt{3}}$ e sua derivada

$$P'_2(x) = \frac{1}{2}(6x) = 3x.$$

Então:

$$w_1 = \frac{1}{3 \left(\frac{-1}{\sqrt{3}} \right)} \int_{-1}^1 \frac{\frac{1}{2}(3x^2 - 1)}{x + \frac{1}{\sqrt{3}}} dx \Rightarrow \frac{-\sqrt{3}}{2} \int_{-1}^1 \left(x - \frac{1}{\sqrt{3}} \right) dx = \frac{-\sqrt{3}}{2} \left[\frac{x^2}{2} - \frac{1}{\sqrt{3}} x \right]_{-1}^1 = 1$$

$$w_2 = \frac{1}{3 \left(\frac{1}{\sqrt{3}} \right)} \int_{-1}^1 \frac{\frac{1}{2}(3x^2 - 1)}{x + \frac{1}{\sqrt{3}}} dx \Rightarrow \frac{\sqrt{3}}{2} \int_{-1}^1 \left(x + \frac{1}{\sqrt{3}} \right) dx = \frac{\sqrt{3}}{2} \left[\frac{x^2}{2} + \frac{1}{\sqrt{3}} x \right]_{-1}^1 = 1$$

Enquanto que para $n = 3$,

$$w_1 = \frac{1}{\frac{3}{2} \left(5 \frac{3}{5} - 1 \right)} \int_{-1}^1 \frac{\frac{1}{2}(5x^3 - 3x)}{x + \sqrt{\frac{3}{5}}} dx \Rightarrow \frac{5}{6} \int_{-1}^1 \left(x^2 - \sqrt{\frac{3}{5}} x \right) dx = \frac{5}{6} \left[\frac{x^3}{3} - \frac{1}{2} \sqrt{\frac{3}{5}} x^2 \right]_{-1}^1 = \frac{5}{9} = (0,55556)$$

$$w_2 = \frac{1}{\frac{3}{2}(5.0-1)^{-1}} \int_{-1}^1 \frac{\frac{1}{2}(5x^3 - 3x)}{x-0} dx \Rightarrow \frac{-1}{3} \int_{-1}^1 (5x^2 - 3) dx = \frac{-1}{3} \left[\frac{5x^3}{3} - 3x \right]_{-1}^1 = \frac{8}{9} = (0,88889)$$

$$w_3 = \frac{1}{\frac{3}{2}\left(5\frac{3}{5}-1\right)^{-1}} \int_{-1}^1 \frac{\frac{1}{2}(5x^3 - 3x)}{x - \sqrt{\frac{3}{5}}} dx \Rightarrow \frac{5}{6} \int_{-1}^1 \left(x^2 + \sqrt{\frac{3}{5}}x \right) dx = \frac{5}{6} \left[\frac{x^3}{3} + \frac{1}{2} \sqrt{\frac{3}{5}}x^2 \right]_{-1}^1 = \frac{5}{9} = (0,55556)$$

Este procedimento pode estender-se para diferentes valores de n, o que implica que devemos tomar n pontos. A tabela A.1 mostra algumas raízes x_i (nós) e coeficientes de ponderação w_i (peso), para n até 5. Neste trabalho foi desenvolvido um programa computacional em ambiente Matlab para determinar estes valores de x_i e w_i para até 53 pontos de quadratura. Este valor n=53 resulta da observação que com n maior, os resultados das integrais não se alteram mais.

Tabela A.1 Fórmulas de Gauss-Legendre $\int_{-1}^1 f(x)dx \cong \sum_{i=0}^n w_i f(x_i)$ com n pontos

Número de pontos (n)	x_i (Raiz)	w_i (Peso)
2	-0,577350269189626	1,000000000000000
	0,577350269189626	1,000000000000000
3	-0,774596669241483	0,5555555555556000
	0,000000000000000	0,8888888888889000
	-0,774596669241483	0,5555555555556000
4	-0,339981043584856	0,652145154862546
	-0,861136311594053	0,347854845137454
	0,861136311594053	0,347854845137454
	0,339981043584856	0,652145154862546
5	-0,906179845938664	0,236926885056189
	-0,538469310105683	0,478628670499366
	0,000000000000000	0,568888888888889
	0,538469310105683	0,478628670499366
	0,906179845938664	0,236926885056189

Sabendo-se que os limites de integração associados com este desenvolvimento são -1 e 1 em um problema de aplicação deverá ser ajustado o procedimento da Quadratura Gaussiana aos limites da aplicação particular mediante uma simples mudança de variável, ou seja, adotamos $[a,b] = [-1,1]$. Para isso, definimos uma relação linear com a nova variável conforme a seguir:

$$t = \frac{2x - a - b}{b - a}$$

$$x = \frac{(b-a)t + a + b}{2}$$

$$dx = \frac{b-a}{2} dt$$

Neste caso, $\int_a^b f(x)dx = \frac{b-a}{2} \int_{-1}^1 f\left(\frac{(b-a)t_i + (b+a)}{2}\right) dt$.

Sabendo que a Quadratura de Gauss-Legendre no intervalo [-1,1] é definida como

$$I(f) = \int_{-1}^1 f(x)dx \cong \sum_{i=1}^n w_i f(x_i) \tag{A.9}$$

Esta integral para um intervalo [a,b]⊂R pode ser expressada com a mudança de variáveis como:

$$I(f) = \int_a^b f(x)dx = \frac{(b-a)}{2} \sum_{i=1}^n w_i f\left(\frac{(b-a)t_i + (b+a)}{2}\right) \tag{A.10}$$

Assim, podemos reescrever a fórmula de integração Gaussiana ou integração de

Gauss-Legendre $I(f) = \int_a^b f(x)dx \cong w_0 f(x_0) + w_1 f(x_1) + \dots + w_n f(x_n) \cong \sum_{i=0}^n w_i f(x_i)$ definida

inicialmente no capítulo (2) para n pontos de quadratura como:

$$I(f) = \int_a^b f(x)dx = \frac{b-a}{2} \int_{-1}^1 f\left(\frac{(b-a)t_i + (b+a)}{2}\right) dt \cong \frac{(b-a)}{2} \sum_{i=1}^n w_i f\left(\frac{(b-a)t_i + (b+a)}{2}\right) \tag{A.11}$$

Como exemplo de aplicação, o cálculo da integral definida $I(f) = \int_0^{5,49} (x^2 - 4x + \sqrt{2})dx$ é

mostrado tanto analiticamente quanto numericamente por meio da fórmula de Quadratura de Gauss-Legendre (para os métodos de dois e cinco pontos).

$$I(f) = \int_0^{5,49} (x^2 - 4x + \sqrt{2})dx = \left[\frac{x^3}{3} - 2x^2 + \sqrt{2}x \right]_0^{5,49} = \mathbf{2,6402154574283}$$

(Valor real com 13 decimais)

Calculando agora numericamente pela fórmula de Quadratura de Gauss-Legendre (A.10):

Para o método n =2 pontos,

$$I(f) = \int_a^b f(x)dx = \frac{(b-a)}{2} \sum_{i=1}^n w_i f\left(\frac{(b-a)t_i + (b+a)}{2}\right)$$

$$I(f) = \frac{(5,49 - 0)}{2} \left[1,0f \left(\frac{(5,49 - 0) \left(\frac{-1}{\sqrt{3}} \right) + (5,49 + 0)}{2} \right) + 1,0f \left(\frac{(5,49 - 0) \left(\frac{1}{\sqrt{3}} \right) + (5,49 + 0)}{2} \right) \right]$$

$$I(f) = (2,745) \cdot [0,9618271247462]$$

$$I(f) = \mathbf{2,6402154574283}$$
 (Valor aproximado com 13 decimais)

O resultado obtido não apresenta erro pois $f(x)$ é um polinômio de grau $\leq 2n - 1 = 3$.

Para o método $n = 5$ pontos, o cálculo da integral é efetuado com o auxílio de uma planilha eletrônica em ambiente MS-Excel, conforme figura A.2:

The screenshot shows an Excel spreadsheet titled "QuadratGaussCustodio". The main content is a table for Gaussian quadrature with $n=5$ points. The table has columns for k , t_k , w_k , the transformed node $[(b-a)t_k + (b+a)]/2$, the function value $f\{[(b-a)t_k + (b+a)]/2\}$, and the weighted function value $f\{[(b-a)t_k + (b+a)]/2\} * w_k$. The function being integrated is $f(x) = x^2 - 4x + \sqrt{2}$. The interval is $[a, b] = [0, 5.49]$. The numerical result of the quadrature is 2,640215457428290, which is compared to the exact integral value of 2,640215457428290.

k	t_k	w_k	$[(b-a)t_k + (b+a)]/2$	$f\{[(b-a)t_k + (b+a)]/2\}$	$f\{[(b-a)t_k + (b+a)]/2\} * w_k$
1	-0,906179845938664	0,23692688505619	0,257536322898368	0,45039322839164	0,106710264653231
2	-0,538469310105683	0,47862867049937	1,266901743759900	-2,04835338432463	-0,980400657052174
3	0,000000000000000	0,568888888888889	2,745000000000000	-2,03076143762690	-1,155277617849970
4	0,538469310105683	0,47862867049937	4,223098256240100	2,35637941927087	1,127830748637680
5	0,906179845938664	0,23692688505619	5,232463677101630	7,86303498615450	1,862964386357420
Soma					0,961827124746188
a	0			Soma * (b-a)/2	2,640215457428290
b	5,49				(Valor Aproximado)
					2,640215457428290
					(Valor Real)

Figura A.2 Exemplo de Aplicação: Método Quadratura Gaussiana ($n = 5$)

Podemos observar que, com $n = 5$ o valor calculado também não apresenta erro.

APÊNDICE B -Teoremas e Demonstrações: Equação Integral para o ARL sob controle

A equação integral analiticamente derivada neste trabalho para o gráfico MCUSUM está sujeita à suposição de que o vetor de médias sob controle é $\boldsymbol{\mu}_0 = \mathbf{0}$ e a matriz de covariância é a identidade I. Embora esta suposição possa parecer restritiva, na realidade não é. Para justificar a razão pela qual isto se comporta, é necessário a demonstração dos seguintes teoremas:

Teorema 1:

Se $\mathbf{X}_1, \mathbf{X}_2, \dots$ são variáveis aleatórias i.i.d $N_p(\boldsymbol{\mu}, \Sigma)$, onde Σ é positiva definida. Se o gráfico MCUSUM com parâmetros k e h é aplicado para estas variáveis X_i , então o ARL sob controle para gráfico MCUSUM com parâmetros k e h aplicado para um processo com vetor de médias $\boldsymbol{\mu}_0 = \mathbf{0}$ e matriz de covariância I. Isto é, portanto suficiente por determinar o ARL para o processo com vetor de médias $\mathbf{0}$ e matriz de covariâncias I.

Demonstração:

Se $S_0 = z \quad 0 \leq z \leq h$

$$S_i = \max[0, S_{i-1} + \mathbf{X}_i - (\boldsymbol{\mu}_0 + k)], \quad i \geq 1$$

e se o processo está sob controle, então o vetor de médias é dado por $E(S_i) = z$

Visto que, Σ e Σ^{-1} são definidos positivos, existe uma matriz triangular superior não singular C tal que $C'C = \Sigma^{-1}$. O produto $C'C$ é a denominada decomposição de Cholesky de Σ^{-1} .

Agora define-se

$$\mathbf{W}_i = C(\mathbf{X}_i - \mathbf{z}), \quad i \geq 1$$

Então $\mathbf{W}_1, \mathbf{W}_2, \dots$ são p variáveis aleatórias normais i.i.d. com vetor de médias

$$E(\mathbf{W}_i) = C[E(\mathbf{X}_i) - \mathbf{z}] = C \cdot \mathbf{0} = \mathbf{0}$$

e a matriz de covariância

$$\text{cov}(\mathbf{W}_i) = C \text{cov}(\mathbf{X}_i - \mathbf{z})C' = C \Sigma C' = C(C'C)^{-1}C' = I$$

Assim $\mathbf{W}_1, \mathbf{W}_2, \dots$ são i.i.d. $N_p(\mathbf{0}, I)$. Agora, define-se:

$$S_0^* = \mathbf{0}$$

$$S_i^* = S_{i-1}^* + \mathbf{W}_i - (\boldsymbol{\mu}_0 - K) \quad i \geq 1$$

Provamos por indução que $\mathbf{S}_i^* = C(\mathbf{S}_i - \mathbf{z})$. Se $i=0$ nós temos $\mathbf{S}_0^* = \mathbf{0}$, e

$$C(\mathbf{S}_0 - \mathbf{z}) = C(\mathbf{z} - \mathbf{z}) = \mathbf{0}$$

Assim, isto é verdadeiro de modo que $\mathbf{S}_0^* = C(\mathbf{S}_0 - \mathbf{z})$. Agora supõe-se que $\mathbf{S}_i^* = C(\mathbf{S}_i - \mathbf{z})$.

Então

$$\begin{aligned} \mathbf{S}_{i+1}^* &= \mathbf{S}_i + \mathbf{W}_{i+1} \\ &= C(\mathbf{S}_i - \mathbf{z}) + C(\mathbf{X}_{i+1} - \mathbf{z}) \\ &= C(\mathbf{S}_i - \mathbf{z} + \mathbf{X}_{i+1}) \\ &= C(\mathbf{S}_{i+1} - \mathbf{z}) \end{aligned}$$

Deste modo se completa a indução, demonstrando que $\mathbf{S}_i^* = C(\mathbf{S}_i - \mathbf{z})$.

As estatísticas Y_i e Y_i^* MCUSUM desses dois desenvolvimentos são portanto

$$Y_i = \left[\left((\mathbf{S}_{i-1} + \mathbf{X}_i - \boldsymbol{\mu}_0) \left(1 - \frac{k}{C_i}\right) \right)' \Sigma^{-1} \left((\mathbf{S}_{i-1} + \mathbf{X}_i - \boldsymbol{\mu}_0) \left(1 - \frac{k}{C_i}\right) \right) \right]^{\frac{1}{2}} \quad e$$

$$Y_i^* = \left[\left((\mathbf{S}_{i-1}^* + \mathbf{X}_i - \mathbf{0}) \left(1 - \frac{k}{C_i}\right) \right)' I^{-1} \left((\mathbf{S}_{i-1}^* + \mathbf{X}_i - \mathbf{0}) \left(1 - \frac{k}{C_i}\right) \right) \right]^{\frac{1}{2}}$$

Esta última estatística pode ser escrita como

$$Y_i^* = \mathbf{S}_i^{*'} \mathbf{S}_i^*$$

$$Y_i^* = \left[\left(C(\mathbf{S}_{i-1} + \mathbf{X}_i) \left(1 - \frac{k}{C_i}\right) \right)' \left(C(\mathbf{S}_{i-1} + \mathbf{X}_i) \left(1 - \frac{k}{C_i}\right) \right) \right]^{\frac{1}{2}}$$

$$Y_i^* = \left[\left(C(\mathbf{S}_{i-1} + \mathbf{X}_i) \left(1 - \frac{k}{C_i}\right) \right)' C' C \left(C(\mathbf{S}_{i-1} + \mathbf{X}_i) \left(1 - \frac{k}{C_i}\right) \right) \right]^{\frac{1}{2}}$$

$$Y_i^* = \left[\left(C(\mathbf{S}_{i-1} + \mathbf{X}_i) \left(1 - \frac{k}{C_i}\right) \right)' \left(C(\mathbf{S}_{i-1} + \mathbf{X}_i) \left(1 - \frac{k}{C_i}\right) \right) \right]^{\frac{1}{2}}$$

$$Y_i^* = Y_i \quad \text{ou ainda}$$

$$\begin{aligned}
 Y_i^* &= \left[\left(C(\mathbf{S}_{i-1} + \mathbf{X}_i) \left(1 - \frac{k}{C_i} \right) \right)' \left(C(\mathbf{S}_{i-1} + \mathbf{X}_i) \left(1 - \frac{k}{C_i} \right) \right) \right]^{\frac{1}{2}} \\
 \{Y_i^*\}^2 &= \left\{ \left[\left(C(\mathbf{S}_{i-1} + \mathbf{X}_i) \left(1 - \frac{k}{C_i} \right) \right)' \left(C(\mathbf{S}_{i-1} + \mathbf{X}_i) \left(1 - \frac{k}{C_i} \right) \right) \right]^{\frac{1}{2}} \right\}^2 \\
 (Y_i^*)^2 &= \left(1 - \frac{k}{C_i} \right)^2 C_i^2 \\
 (Y_i^*)^2 &= \left(1 - \frac{2k}{C_i} + \frac{k^2}{C_i^2} \right) C_i^2 \\
 (Y_i^*)^2 &= (C_i^2 - 2kC_i + k^2) \\
 (Y_i^*)^2 &= (C_i - k)^2 \\
 Y_i^* &= (C_i - k) \\
 Y_i^* &= Y_i
 \end{aligned}$$

Assim, quando o processo permanece sob controle, o gráfico MCUSUM para o processo \mathbf{X} que tem vetor de médias $\boldsymbol{\mu}$ e matriz de covariância Σ , um sinal fora de controle é emitido, ao mesmo tempo em que o gráfico MCUSUM para o processo \mathbf{W} tem vetor de médias $\mathbf{0}$ e matriz de covariância I . Os ARL's para as duas seqüências de MCUSUM são, portanto, iguais.

Teorema 2:

O ARL sob controle do gráfico de controle MCUSUM aplicado para o processo com vetor de médias $\boldsymbol{\mu}_0 = \mathbf{0}$ e matriz de covariância I dependem tão somente do valor inicial $S_0 = z$ através do quadrado de sua magnitude $\delta = \mathbf{z}'\mathbf{z}$.

Demonstração:

Se \mathbf{z} e \mathbf{z}^* indicam dois valores iniciais para o gráfico MCUSUM e que satisfaz $\mathbf{z}'\mathbf{z} = \delta = \mathbf{z}^{*\prime}\mathbf{z}^*$. Para isso, vamos mostrar que o ARL é o mesmo que utilizarmos \mathbf{z} ou \mathbf{z}^* . Visto que \mathbf{z} tem a mesma magnitude que \mathbf{z}^* e como \mathbf{z}^* é uma rotação de \mathbf{z} ; assim existe uma matriz ortogonal A tal que $\mathbf{z}^* = A\mathbf{z}$. Se $\mathbf{X}_1, \mathbf{X}_2, \dots$ é uma seqüência de variáveis aleatórias i.i.d. $N_p(\mathbf{0}, I)$, e se $\mathbf{X}_i^* = A\mathbf{X}_i$. Então $E(\mathbf{X}_i^*) = A\mathbf{0} = \mathbf{0}$ e visto que

$\text{cov}(\mathbf{X}_i^*) = \text{cov}(A\mathbf{X}_i) = A\text{cov}(\mathbf{X}_i)A' = AIA' = I$, a seqüência \mathbf{X}_i^* é também i.i.d. $N_p(\mathbf{0}, I)$. Se

\mathbf{S}_i^* , $i = 1, 2, \dots$ é a seqüência MCUSUM

$$\mathbf{S}_i^* = \mathbf{S}_{i-1}^* + \mathbf{X}_i - (\boldsymbol{\mu}_0 + k), \quad i \geq 1$$

Uma demonstração por indução, similar ao que foi demonstrado no teorema 1, onde:

$$\mathbf{S}_i^* = A\mathbf{S}_i$$

Se Y_i e Y_i^* indicam as estatísticas MCUSUM para estas duas seqüências. Então, devido A ser uma matriz ortogonal nós temos

$$\begin{aligned} Y_i^* &= \left[\left((\mathbf{S}_{i-1}^* + \mathbf{X}_i - \mathbf{0}) \left(1 - \frac{k}{C_i}\right) \right)' I^{-1} \left((\mathbf{S}_{i-1}^* + \mathbf{X}_i - \mathbf{0}) \left(1 - \frac{k}{C_i}\right) \right) \right]^{\frac{1}{2}} \\ &= \left[(A\mathbf{S}_{i-1})' (A\mathbf{S}_{i-1}) \right]^{\frac{1}{2}} \\ &= \left(\mathbf{s}_{i-1}' A' A \mathbf{s}_{i-1} \right)^{\frac{1}{2}} \\ &= \left(\mathbf{s}_{i-1}' \mathbf{s}_{i-1} \right)^{\frac{1}{2}} \end{aligned}$$

Assim as duas seqüências da estatística Y_i são iguais. Isto implica que o gráfico MCUSUM que inicia com \mathbf{z} e \mathbf{z}^* emitirá um sinal na mesma fase de amostragem. Então os ARL's determinados são iguais.

Teorema 3:

Supomos que o gráfico MCUSUM com parâmetro k aprimorado seja aplicado para um processo p dimensional com vetor de médias $\boldsymbol{\mu}_0 = \mathbf{0}$ e matriz de covariância I. Se $L(\delta/h)$ denota o ARL dado para que a estatística MCUSUM inicial $S_0 = z$ satisfaça $\mathbf{S}_0' \mathbf{S}_0 = \delta$ e dado que o intervalo de decisão é h. Então a função L satisfaz a equação integral

$$L(\delta/h) = 1 + L(0/h)F(x/\delta) + \int_0^h L(x)f(x | \mathbf{S}_0' \mathbf{S}_0 = \delta) dx$$

onde $f(x | \mathbf{S}_0' \mathbf{S}_0 = \delta)$ é a função densidade de probabilidade da distribuição Qui-quadrado não central com p graus de liberdade e parâmetro de não centralidade $\sum_{i=1}^p \lambda_i^2$.

Demonstração:

Com a suposição de que o valor do vetor de médias é zero e a matriz de covariâncias é a

identidade, nós temos $Y_i = (\mathbf{S}_i' \mathbf{S}_i)^{\frac{1}{2}}$

Seja um vetor \mathbf{S}_i de p componentes distribuídas segundo uma seqüência de variáveis aleatórias $N(\mathbf{v}, T)$ não singular, então $\mathbf{S}_i' T^{-1} \mathbf{S}_i$ é distribuído de acordo com uma distribuição χ^2 não central com graus de liberdade e parâmetro de não centralidade $\mathbf{v}' T^{-1} \mathbf{v}$. Agora, se admitirmos C uma matriz não singular tal que $CT^{-1}C' = I$ e definindo \mathbf{Z} como $\mathbf{Z} = C\mathbf{S}_i$. Então \mathbf{Z} é normalmente distribuído com média $E(\mathbf{Z}) = CE(\mathbf{S}_i) = C\mathbf{v} = \boldsymbol{\lambda}$ e matriz de covariância

$$E(\mathbf{Z} - \boldsymbol{\lambda})(\mathbf{Z} - \boldsymbol{\lambda})' = E(\mathbf{S}_i - \mathbf{v})(\mathbf{S}_i - \mathbf{v})' C' = CT^{-1}C' = I.$$

Então $\mathbf{S}_i' T^{-1} \mathbf{S}_i = \mathbf{Z}' (C')^{-1} T^{-1} C^{-1} \mathbf{Z} = \mathbf{Z}' (CT^{-1}C')^{-1} \mathbf{Z} = \mathbf{Z}' \mathbf{Z} = Z_1^2 + Z_2^2 + \dots + Z_p^2$. Analogamente

temos $\mathbf{v}' T^{-1} \mathbf{v} = \boldsymbol{\lambda}' \boldsymbol{\lambda}$. Portanto, $\mathbf{S}_i' T^{-1} \mathbf{S}_i$ é distribuído como $\sum_{i=1}^p Z_i^2$ onde Z_1, Z_2, \dots, Z_p são independentes e normalmente distribuídos com médias $\lambda_1, \lambda_2, \dots, \lambda_p$, respectivamente, e variância 1. Por definição esta distribuição, denotada por o $\chi_{(p, \lambda)}^2$, é uma χ^2 não central com

parâmetros de não centralidade $\sum_{i=1}^p \lambda_i^2$.

Se $f(n, x)$ indica a função densidade de probabilidade conjunta de variáveis aleatórias N e X . A fdp de N pode ser escrita através da integral de todos os possíveis valores de X quando $n = 1, 2, \dots$. Isto fornece:

$$\begin{aligned} L(\delta/h) &= 1 + L(0/h)F(x|\delta) + \sum_{n=1}^{\infty} n \int_0^h f(n+1, x | \mathbf{S}_i' \mathbf{S}_i = \delta) dx \\ &= 1 + L(0/h)F(x|\delta) + \sum_{n=1}^{\infty} n \int_0^h [(P(N = n+1) | \mathbf{S}_i' \mathbf{S}_i = \delta, \mathbf{S}_1' \mathbf{S}_1 = x)] \times f(x | \mathbf{S}_i' \mathbf{S}_i = \delta) dx \\ &= 1 + L(0/h)F(x|\delta) + \left(\sum_{n=1}^{\infty} n P(N = n+1) | \mathbf{S}_i' \mathbf{S}_i = \delta, \mathbf{S}_1' \mathbf{S}_1 = x \right) \times f(x | \mathbf{S}_i' \mathbf{S}_i = \delta) \end{aligned} \quad (B.1)$$

onde $f(x | \mathbf{S}_i' \mathbf{S}_i = \delta)$ é a função densidade de probabilidade da distribuição $\chi_{(p, \lambda)}^2$.

Dado que $\mathbf{S}_i' \mathbf{S}_i = x$ a probabilidade entre parêntesis em (B.1) não depende de $\delta = \mathbf{S}_i' \mathbf{S}_i$. Esta somatória pode então ser escrita como

$$\sum_{n=1}^{\infty} n P(N - 1 = n) | \mathbf{S}_1' \mathbf{S}_1 = x$$

Esta soma infinita é justamente o valor esperado do número adicional de amostras coletadas

até o sinal (RL) conforme o primeiro subgrupo dado que $\mathbf{S}_1' \mathbf{S}_1 = x$. Portanto, esta soma $L(x)$ é obtida a partir da solução da equação integral

$$L(\delta/h) = 1 + L(0/h)F(x|\delta) + \int_0^h L(x)f(x|\mathbf{S}_1' \mathbf{S}_1 = \delta)dx \quad (\text{B.2})$$

na função L desconhecida. É justamente nesta integral que se aplica o método numérico de Quadratura Gaussiana.

$$L(\delta) = 1 + L(0h)F(x|\delta) + \sum_{i=1}^n w_i L(x_i) f(x_i|\mathbf{S}_1' \mathbf{S}_1 = \delta) + \rho(\delta) \quad (\text{B.3})$$

onde $\{x_i\}_{i=1}^n$, $\{w_i\}_{i=1}^n$, são, respectivamente os nós e os pesos da quadratura gaussiana no intervalo $[0,h]$ e $\rho(z)$ é um erro muito pequeno desconhecido devido a substituição da integral pela quadratura. Avaliando $L(\cdot)$ para $\delta = 0, \delta_1, \delta_2, \dots, \delta_n$,

$$\tilde{L}(0) = 1 + \tilde{L}(0/h)F(x|\delta) + \sum_{i=1}^n w_i \tilde{L}(x_i) f(x_i) \quad (\text{B.4})$$

$$\tilde{L}(\delta_i) = \tilde{L}(0/h)F(x|\delta) + \sum_{i=1}^n w_i \tilde{L}(x_i) f(x_i|\mathbf{S}_1' \mathbf{S}_1 = \delta) \quad (\text{B.5})$$

onde $\tilde{L}(\cdot)$ é uma aproximação para $L(\cdot)$. A aproximação de $\tilde{L}(\cdot)$ é obtida a partir da resolução do sistema de equações algébricas lineares e usando a equação

$$\tilde{L}(\delta) = 1 + \tilde{L}(0/h)F(x|\delta) + \sum_{i=1}^n w_i \tilde{L}(x_i) f(x_i|\mathbf{S}_1' \mathbf{S}_1 = \delta) \quad (\text{B.6})$$

Resolvendo o sistema de equações cujo procedimento de resolução é análogo ao caso univariado obtém-se $\tilde{L}_1 = \tilde{L}_1(x_1), \tilde{L}_2 = \tilde{L}_2(x_2), \dots, \tilde{L}_n = \tilde{L}_n(x_n)$ e substituindo $\tilde{L}(x_i)$ em (B.4) e (B.6), determina-se $\tilde{L}(0/h)$ e $\tilde{L}(\delta)$, respectivamente. Para aproximar a solução por (B.1) utiliza-se o método Quadratura Gaussiana com $n=53$ pontos de quadratura.

APÊNDICE C - Distribuição Normal Multivariada

C.1 Introdução

A generalização da densidade normal univariada para duas ou mais dimensões desempenha um papel fundamental na análise multivariada. De fato, a maioria das técnicas multivariadas parte do pressuposto de que os dados foram gerados de uma distribuição normal multivariada. Apesar dos dados originais não serem quase nunca “ exatamente” normal multivariados, a densidade normal se constitui muitas vezes numa aproximação adequada e útil da verdadeira distribuição populacional (Pena, 2002).

A distribuição normal, além da sua atratividade pela sua facilidade de tratamento matemático, possui duas razões práticas que justificam a sua utilidade. A primeira, diz que a distribuição normal é a mais adequada para modelos populacionais em várias situações; e a segunda refere-se ao fato da distribuição amostral de muitas estatísticas multivariadas ser aproximadamente normal, independentemente da forma da distribuição da população original, devido ao efeito do teorema do limite central (Yang e Trewn, 2004).

C.2 Densidade Normal Multivariada em Controle Estatístico de Processos

A abordagem da densidade normal multivariada em controle estatístico de processos multivariados é análoga ao caso univariado cujas alterações do modelo matemático estão relacionadas apenas à adição de p variáveis. No controle estatístico univariado se a característica da qualidade X monitorada é contínua, geralmente assumimos que esta variável segue uma distribuição de probabilidade normal $N \sim (\mu, \sigma)$ cuja função de densidade univariada é dada por

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty \leq x \leq \infty \quad (C.1)$$

onde μ, σ representam respectivamente a média e o desvio padrão da variável X e $\left(\frac{x-\mu}{\sigma}\right)^2$ é o termo que mede o quadrado da distância padronizada (em unidades de desvio padrão) do valor x à média μ . Definindo $t = \frac{x-\mu}{\sigma}$ como a variável aleatória normal padrão $N \sim (0,1)$ a equação (C.1) resulta em

$$f(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}t^2} \quad (C.2)$$

Esta mesma abordagem de aproximação pode ser utilizada no caso do controle estatístico de processos multivariados

Suponhamos agora que estamos interessados em monitorar simultaneamente p características da qualidade X_1, X_2, \dots, X_p medidas sobre uma peça e que se encontram correlacionadas. Podemos considerar que o vetor aleatório ou as p-variáveis aleatórias $\mathbf{X}' = [x_1, x_2, \dots, x_p]$ seguem uma distribuição normal multivariada com vetor de médias $\boldsymbol{\mu}' = [\mu_1, \mu_2, \dots, \mu_p]$ e matriz de variâncias-covariâncias Σ de dimensão p x p onde $|\Sigma|$ representa o determinante desta matriz. Neste caso, o quadrado da distância (generalizada) padronizada ou quadrado da distância entre os vetores \mathbf{X}' e $\boldsymbol{\mu}'$

$$d^2 = (\mathbf{X}', \boldsymbol{\mu}') = (\mathbf{x} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) \quad (C.3)$$

conhecida em análise multivariada como a distância de Mahalanobis que se distribui como uma χ^2 com p graus de liberdade. Este conceito é muito importante em controle de qualidade multivariado pois nos permite quantificar a troca produzida em um vetor aleatório amplamente utilizado no presente trabalho.

A função de densidade normal multivariada pode então ser obtida simplesmente pela substituição do termo $\left(\frac{x - \mu}{\sigma}\right)^2$ por $(x - \mu)' \Sigma (x - \mu)$ e pela mudança do termo $\frac{1}{\sqrt{2\pi\sigma^2}}$ para uma forma mais geral que torne a área sob a função densidade de probabilidade unitária, independente do valor de p. Portanto, a função de densidade de probabilidade normal multivariada $N_p \sim (\mu, \Sigma)$ é

$$f(x) = f(x_1, x_2, \dots, x_p) = \frac{1}{(2\pi)^{\frac{p}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})} \quad \text{onde } -\infty \leq x_1, x_2, \dots, x_p \leq \infty \quad (C.4)$$

Esta expressão revela que os contornos de densidade de probabilidade constante têm a forma de um elipsóide centrado em μ e gerados quando o termo da exponencial for constante, isto é,

$$d^2 = (\mathbf{x} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) = \text{cte}$$

A figura C.1. ilustra um exemplo de distribuição normal bivariada desenvolvido a partir de uma planilha de dados em ambiente MS-Excel

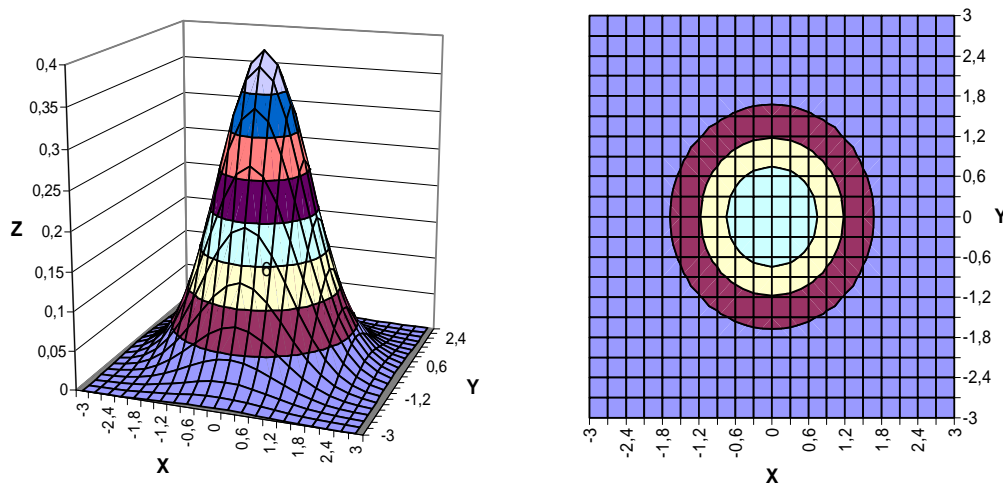


Figura C.1 Exemplo de distribuição normal bivariada

C.3 Normalidade Multivariada

Alguns problemas de modelagem estatística ou de testes de hipóteses estatísticas, entre outros têm como pressuposto básico que os dados sejam distribuídos normalmente, independentes e identicamente distribuídos, sendo crucial verificar-se se os dados com que se está trabalhando satisfazem a condição da distribuição inicial. Para muitas análises estatística, é realmente importante que os dados sigam uma distribuição normal multivariada, quando não exatamente, pelo menos, aproximadamente. Se tal fato for verificado, ou seja, os dados forem distribuídos segundo uma distribuição normal multivariada, evita-se o problema de procurar por procedimentos de transformação de variáveis, que na sua maioria são robustos, de tal forma a normalizar o conjunto de dados para, a partir daí, trabalhar-se com os dados multivariados normalmente distribuídos.

Para verificar a hipótese de multinormalidade dos dados, D'Agostino (1986) recomenda a realização de três testes:

- Verificar a normalidade de cada variável em separado, através de um método gráfico.
- Aplicar um teste estatístico univariado para cada variável.
- Verificar a multinormalidade, através de um teste estatístico multivariado.

Segundo D'Agostino (1986) mesmo que cada variável tenha distribuição normal, não significa que o conjunto seja multinormal. No entanto, se uma variável não é normalmente distribuída, o conjunto não é multinormal, daí a importância dos testes univariados.

A aplicação de métodos gráficos tais como o gráfico de probabilidade qui-quadrado (Q-Q plot) podem ser utilizados para auxiliarem na verificação de normalidade multivariada.

Este método gráfico se constitui em determinar se os dados das amostras estão de acordo com a suposta distribuição, baseado em um exame visual subjetivo dos dados. Mingoti (2005) propõe um algoritmo para o desenvolvimento do gráfico de probabilidade qui-quadrado (Q-Q plot). Com a utilização deste algoritmo desenvolveu-se neste trabalho em ambiente MS-Excel o gráfico Q-Q plot para três variáveis do exemplo numérico proposto pela autora conforme figura C.2 Podemos constatar visualmente que a disposição dos pontos indica a proximidade em relação à reta, o que nos leva a não rejeitar a suposição de normalidade multivariada.

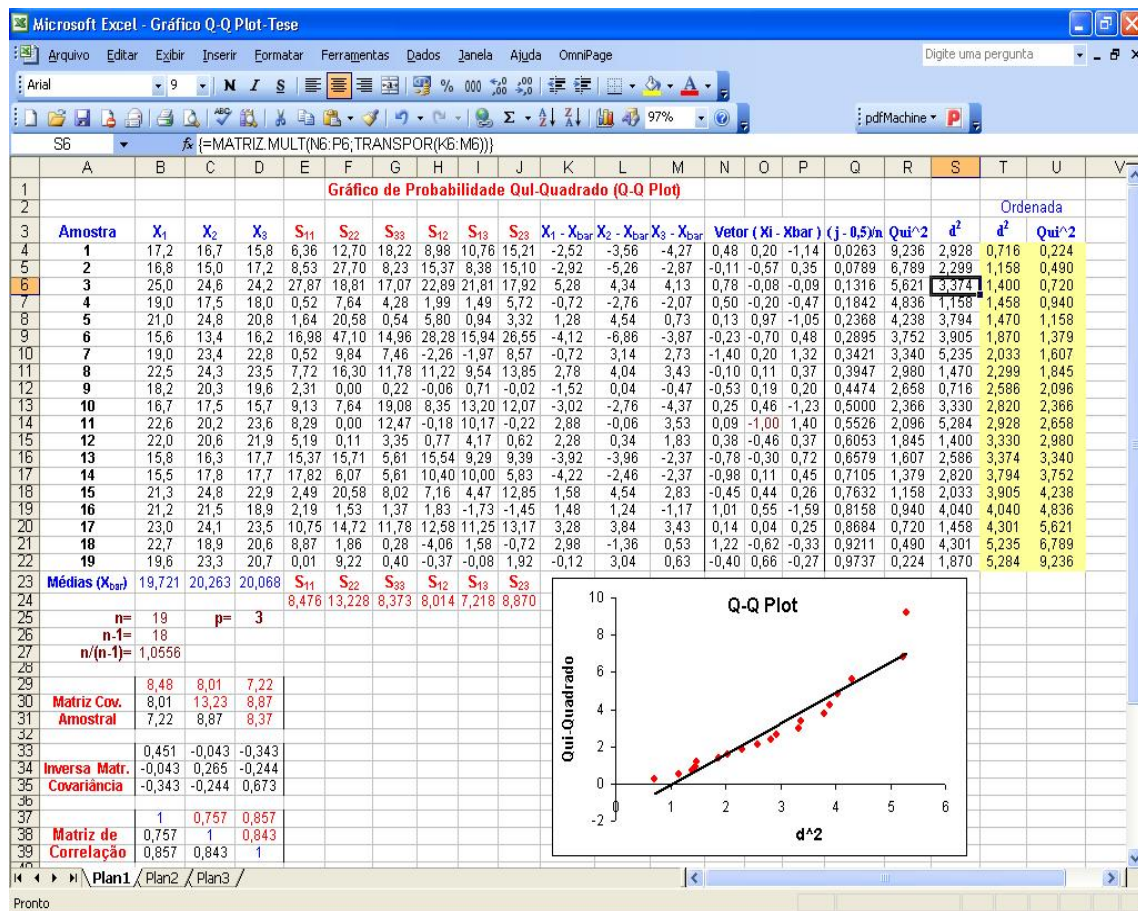


Figura C.2 Gráfico de probabilidade qui-quadrado (Q-Q plot) do exemplo de Mingoti (2005) Podemos constatar visualmente que a disposição dos pontos indica a proximidade em relação à reta, o que nos leva a não rejeitar a suposição de normalidade multivariada.

Além de métodos gráficos, outros procedimentos podem ser utilizados para verificar a suposição de normalidade multivariada. Segundo Rencher (2002) um dos tratamentos matemáticos usuais relativamente simples, para verificar a suposição de normalidade multivariada, é usar testes de hipóteses baseados nas medidas de assimetria e curtose

multivariada de Mardia (1970,1974). Esse na realidade, é um teste descritivo e um excelente procedimento para a verificação de tal suposição. Para utilizar os testes de hipóteses na verificação de suposição de normalidade é necessário determinar antes as medidas de assimetria e curtose por intermédio dos seguintes momentos:

$$\beta_{1,p} = E[(y - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (x - \boldsymbol{\mu})]^3 \quad (\text{C.5})$$

Sendo x independente de y como em todas as distribuições e

$$\beta_{2,p} = E[(y - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (y - \boldsymbol{\mu})]^2 \quad (\text{C.6})$$

de tal forma que as esperanças nas expressões $\beta_{1,p}$ (C.5) e $\beta_{2,p}$ (C.6) existam. Para a distribuição normal multivariada, espera-se que a assimetria seja $\beta_{1,p} = 0$ e a curtose $\beta_{2,p} = p(p + 2)$.

Para a amostra de tamanho n, os estimadores $\beta_{1,p}$ e $\beta_{2,p}$ serão obtidos por:

$$\widehat{\beta}_{1,p} = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n g_{ij}^3 \quad (\text{C.7})$$

$$\widehat{\beta}_{2,p} = \frac{1}{n} \sum_{i=1}^n g_{ij}^2 = \frac{1}{n} \sum_{i=1}^n d_i^4 \quad (\text{C.8})$$

onde $g_{ij} = (y_i - \bar{y})' S_n^{-1} (y_i - \bar{y})$ e $d_i = \sqrt{g_{ij}}$

essa é a versão amostral da distância de Mahalanobis entre y_i e \bar{y} (que é a estimação da média populacional μ) Mardia (1970).

Observa-se que os valores obtidos por $\widehat{\beta}_{1,p}$ (igual ao quadrado do coeficiente amostral de assimetria quando p=1) e $\widehat{\beta}_{2,p}$ (igual ao coeficiente de curtose quando p=1) de tal forma que são não negativos. Para os dados com distribuição normal multivariada, espera-se que $\widehat{\beta}_{1,p}$ seja aproximadamente zero. Partindo de uma simetria esférica (ou seja, correlação zero e variâncias iguais), espera-se que $\widehat{\beta}_{2,p}$ bem maiores. A quantidade $\widehat{\beta}_{2,p}$ é usada para indicar os pontos com comportamento extremo no quadrado de distancia de Mahalanobis das observações em relação a média amostral. Mas, $\widehat{\beta}_{2,p}$ e $\widehat{\beta}_{1,p}$ podem ser utilizados para detectar a normalidade multivariada. Em seu trabalho, Mardia (1970) mostra que, para grandes amostras $k_1 = \frac{n\widehat{\beta}_{1,p}}{6}$

segue uma distribuição Qui-Quadrado com $\frac{p(p+1)(p+2)}{6}$ graus de

liberdade e $k_2 = \frac{[\beta_{2,p} - p(p+2)]}{\sqrt{\frac{8p(p+2)}{n}}}$ segue uma distribuição normal padrão. Para amostras

pequenas, existe a tabela dos valores críticos para este teste estatístico desenvolvido por Mardia(1974). Mardia, recomenda que se ambas as hipóteses são aceitas, a teoria normal para testes do vetor de média ou a matriz de variância-convariância poderá se usada. Portanto, na presença de não normalidade, a teoria dos testes normais na média são sensíveis a $\hat{\beta}_{1,p}$. Visto que testes na matriz de variância-convariância são influenciados por $\hat{\beta}_{2,p}$. Pode-se observar que a curtose multivariada de Mardia será usada como uma medida de detectar os outliers (ou pontos extremos) dos dados que são supostamente distribuídos segundo uma normal multivariada.

É importante ressaltar também que em análise estatística multivariada, além da distribuição normal outras distribuições intimamente relacionadas à distribuição normal são aplicadas. Uma delas é a distribuição Qui-Quadrado (χ^2) definida em termos da distribuição normal e importante por ser base de procedimentos de inferência estatística, procedimentos estes que envolvem a utilização de técnicas de gráficos de controle multivariados em sucessivas amostras para monitorar um processo onde a média é uma normal multivariada que pode ser interpretada com repetidos testes de significância da forma

$$H_o : \mu = \mu_o$$

$$H_1 : \mu \neq \mu_o$$

onde μ representa a média normal multivariada do processo cujo valor é desconhecido e μ_o é o valor nominal (alvo) para o parâmetro.

Suponhamos que p-vetores aleatórios \mathbf{X}_i são independentes e identicamente distribuídos segundo uma multivariável normal com matriz de covariância Σ conhecida e constante, isto é, $N_p \sim (\mu, \Sigma)$ onde \mathbf{X}_i denota a (amostra) do vetor de médias. Para testar a hipótese, deve ser rejeitada a hipótese (H_o) quando $\chi^2 > \chi^2_{(v,\alpha)}$ onde

$$\chi^2 = (\mathbf{x} - \boldsymbol{\mu}_o)' \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}_o) \tag{C.9}$$

e $\chi^2_{(v,\alpha)}$ é superior a 100α pontos percentuais da distribuição χ^2 com v graus de liberdade.

Uma outra distribuição associada a distribuição χ^2 é a distribuição Qui-Quadrado não central (χ'^2) cujo tamanho da mudança entre os vetores $\boldsymbol{\mu}$ e $\boldsymbol{\mu}_o$ é reportando em termos de

uma quantidade

$$\lambda^2(\boldsymbol{\mu}) = (\boldsymbol{\mu} - \boldsymbol{\mu}_0)' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - \boldsymbol{\mu}_0) \quad (\text{C.10})$$

onde $\lambda^2(u)$ é o parâmetro de não centralidade da distribuição $\chi^2_{(v,\lambda)}$ com ν graus de liberdade frequentemente utilizado para representar uma medida de distância em p dimensões entre os vetores de média $\boldsymbol{\mu}$ e $\boldsymbol{\mu}_0$. A maior diferença entre as distribuições $\chi^2_{(v,\alpha)}$ e $\chi^2_{(v,\lambda)}$ é o adicional parâmetro λ referente ao parâmetro de não centralidade. Quando os p -vetores aleatórios \mathbf{X}_i anteriormente mencionados são i.i.d. e se distribuem segundo $\mathbf{X}_i \sim N_p(0, I)$, o parâmetro de não centralidade pode ser escrito como $\lambda^2(\boldsymbol{\mu}) = \boldsymbol{\mu}' \boldsymbol{\mu}$.

APÊNDICE D - Autocorrelação e Estabilidade Estatística

Para que seja possível a utilização de qualquer tipo de gráfico de controle duas suposições devem ser validadas, partindo-se da situação de processo para o controle estatístico. A primeira suposição é a de independência estatística dos dados gerados no processo. A segunda é a de que os dados de processo seguem uma distribuição normal com média μ e desvio padrão σ . Quando ambas as suposições são satisfeitas, para o controle estatístico, os dados de processo (x_t , $t=1,2,\dots$) seguem o modelo

$$x_t = \mu + \varepsilon_t \quad (D.1)$$

Quando a suposição de independência entre as observações de um processo não é satisfeita, o gráfico de controle acaba não oferecendo um desempenho satisfatório. Esta presença de autocorrelação entre as observações tem profundos efeitos nos gráficos de controle desenvolvidos usando a suposição de independência entre as observações. Um destes efeitos é o aumento de frequência com que sinais falsos são gerados. Mesmo correlações fracas produzem distúrbios nos gráficos de controle levando a conclusões erradas sobre o estado de controle do processo. Na presença de um sinal fora do controle, por exemplo, torna-se difícil distinguir se o mesmo é ocasionado pela presença de uma causa especial ou se é um alarme falso induzido pela estrutura de autocorrelação dos dados do processo.

A autocorrelação nada mais é do que um mecanismo existente no processo, faz com que os dados não sejam mais independentes entre si ao longo do tempo. A intensidade e o grau de autocorrelação podem ser medidos analiticamente. A autocorrelação ao longo de um modelo de séries temporais é medida pela função de autocorrelação

$$\rho_k = \frac{Cov(x_t, x_{t-k})}{V(x_t)} \quad k = 0, 1, 2, \dots \quad (D.2)$$

Esta função permite que se entenda melhor o comportamento da dependência estatística entre as observações e, posteriormente, pode ser útil quando da determinação de qual modelo de séries temporais utilizar para representar o processo.

A abordagem de diversos autores para trabalhar com dados autocorrelacionados baseia-se no ajuste de um modelo apropriado de séries temporais.

Mason (1997,2002) sugere séries temporais envolvendo a análise de funções de autocorrelação (correlogramas) e a aplicação de técnicas de estatística inferencial tais como

Análise de Variância (ANOVA) para o desenvolvimento de um modelo de Análise de Regressão AR(1) da forma $y_t = b_0 + b_1 y_{t-1} + erro$.

A aplicação de um modelo de séries temporais para investigar as propriedades ARL do gráfico de controle MCUSUM quando as observações são autocorrelacionadas é proposto por Noorossana e Vaghefi (2006). Aplicando séries temporais utilizam neste modelo além de correlogramas as observações geradas por uma Análise de Regressão AR(1) para investigar o efeito da autocorrelação sobre o desempenho do gráfico MCUSUM. A aplicação deste modelo para dados de um exemplo numérico comprova que a violação de independência que comumente ocorre pode conduzir a um desempenho insatisfatório do gráfico de controle MCUSUM.

Jarret e Pan (2007) sugerem para processos independentes, o de controle de controle gráfico multivariado e para processos autocorrelacionados, os gráficos de controle univariados separadamente. Para isso, propõem um vetor autoregressivo (VAR) para gráficos de controle como uma combinação tabelada para autocorrelação resídua univariada e multivariada do gráfico para processos independentes.

Para monitorar o vetor de médias de um processo, Kalgonda e Kulkarn (2004) propõem um modelo autoregressivo VAR(1) de primeira ordem. Este modelo apresenta como recurso as ilustrações que incluem conhecimentos apropriados tais como CEP multivariado e autocorrelação.

Cheng e Thaga (2005) sugerem um gráfico de controle CUSUM capaz de detectar mudanças na média e desvio padrão de dados autocorrelacionados. Afirmam com base no ARL que o novo gráfico de controle é útil para o acompanhamento de processos de produção modernos onde se produz com alta qualidade e com a mínima parte de unidades de produtos de saída não conformes.

A aplicação de qualquer gráfico de controle estatístico a processos com dados autocorrelacionados merece uma atenção especial de forma a evitar conclusões incorretas sobre o comportamento do processo. O efeito de não ser considerada a autocorrelação quando se aplica o CEP pode se manifestar da seguinte forma:

- Considerar estável (sob controle estatístico) um processo, quando realmente existem causas especiais de variação.
- Identificar causas especiais de variação, quando na realidade o processo se encontra sob controle estatístico.
- Estimar incorretamente os parâmetros do processo.

- Dada a imprecisão na estimação dos parâmetros do processo, considerar incorreto um processo capaz ou não capaz.
- Investigação da raiz de um eventual problema que na realidade não existe no processo, o que conduz desperdício de recursos e descrédito do CEP.

Portanto, na análise estatística de processos quando se considera a autocorrelação em gráficos de controle facilita tanto a verificação da evidência de qualquer causa especial de variação quanto o estudo da capacidade do processo.

APÊNDICE E - Teste Seqüencial da Razão de Probabilidades e Teoria da Soma Acumulada em Gráficos de Controle

Os fundamentos teóricos que fornecem apoio técnico ao estudo de somas acumuladas disponíveis na literatura são abordados a partir de duas linhas de trabalho. Um deles é a proposta de Page (1954) que sugere a soma acumulada de desvios em relação ao valor de referência. Esta soma, segundo Page deve ser diagnosticada para verificar o aumento máximo desta para um ponto imediatamente superior.

Uma outra linha de trabalho é o Teste Seqüencial da Razão de Probabilidades (SPRT-Sequential Probability Ratio Test) desenvolvida por Wald (1945). Análogo ao princípio de verossimilhança da soma acumulada, o SPRT utiliza dados seqüencialmente, adicionando quando possível, a informação de cada nova observação, X_n .

Nos testes SPRT avalia-se uma simples hipótese nula (H_0) contra uma simples hipótese alternativa (H_1). Para cada uma destas hipóteses é associada uma função densidade de probabilidade $f_0(x)$ e $f_1(x)$, respectivamente, para a medida X_n .

Suponha uma determinada seqüência de observações independentes $\{X_i\}$ de tamanho n que se deseja decidir entre H_0 e H_1 . No SPRT, a razão de verossimilhança, Λ_n é dada por

$$\Lambda_n = \prod_{i=1}^n \frac{f_1(X_i)}{f_0(X_i)} \quad (\text{E.1})$$

O teste aceita H_0 se Λ_n for inferior ou igual ao valor do ponto de corte A , e rejeita H_0 em favor de H_1 se Λ_n for maior que o valor do outro ponto de corte, B . No entanto, se $A < \Lambda_n < B$, então o SPRT apela para uma outra observação X_{n+1} e atualiza a razão de verossimilhança para incorporá-lo.

Geralmente, é mais fácil trabalhar com logaritmos da razão de verossimilhança

$$\ln \Lambda_n = \sum_{i=1}^n \ln \left[\frac{f_1(X_i)}{f_0(X_i)} \right] \quad (\text{E.2})$$

A equação (E.2) pode ser escrita também como $\ln \Lambda_n = \sum_{i=1}^n \ln Z_i$ então

$$Z_i = \ln \left[\frac{f_1(X_i)}{f_0(X_i)} \right] \quad (\text{E.3})$$

onde o escore da variável Z_i é apenas uma transformação da quantidade aleatória. Por isso, é também uma variável aleatória cuja distribuição pode ser calculada a partir do X . De acordo com o pressuposto de que os valores originais de X_i são independentes, então os valores de Z_i são também independentes.

Com isso:

- Aceita-se H_0 se $\ln \Lambda_n \leq \ln A$
- Aceita-se H_1 se $\ln \Lambda_n \geq \ln B$
- Apela para uma outra observação se $\ln A < \ln \Lambda_n < \ln B$

A figura E.1 ilustra isto com $\ln A = -3$ e $\ln B = 3$, onde teria sido aceito a observação 4.

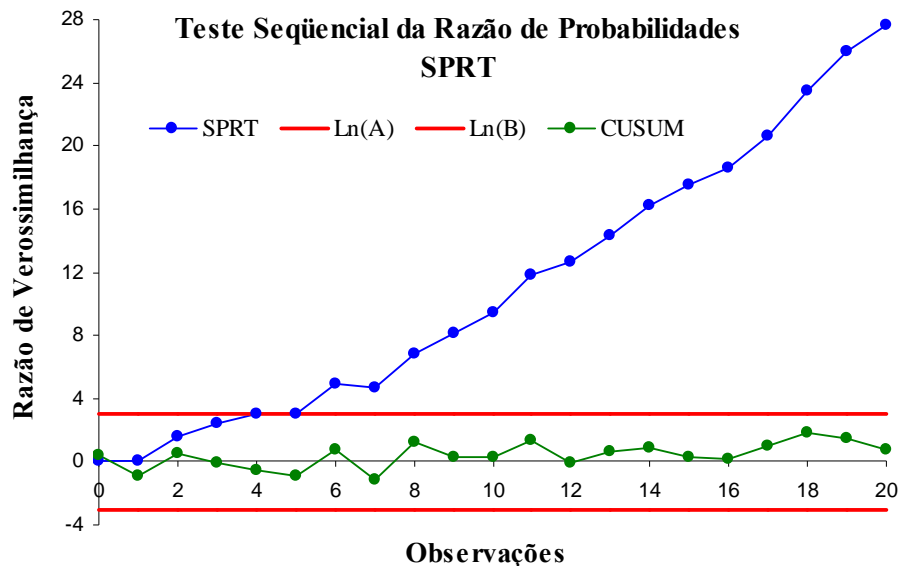


Figura E.1 Representação gráfica do SPRT (Adaptado de Hawkins, 1998, pág. 137)

Como se pode observar há uma notável semelhança entre a representação gráfica de SPRT e a forma original da Máscara V do gráfico MCUSUM.

Wald (1947) supõe o teste seqüencial da razão de probabilidades como sendo um teste ótimo para decidir entre as duas hipóteses. Portanto, aplica o SPRT para minimizar o número de pontos amostrados antes que uma decisão possa ser alcançada.

Na prática, o teste SPRT com estes logaritmos da razão de verossimilhança, ou $\ln(\Lambda_n)$, se transforma num procedimento baseado na soma acumulada dos Z_i . Com base nesta soma é possível decidir aceitar, rejeitar, ou continuar a amostragem. Se a hipótese H_1 for verdadeira, os valores esperados de Z_i serão positivos. Como os pontos são adicionados, a

soma segue uma tendência ascendente. No entanto se a hipótese H_0 for verdadeira, os valores esperados de Z_i serão negativos e a soma segue uma tendência descendente.

As constantes A e B determinam as probabilidades de erros tipo I e tipo II. Os valores necessários para estas constantes A e B são fixados através das probabilidades:

$$\alpha = P(\text{erro tipo I}) = P(\text{rejeitar } H_0 | H_0) \quad \text{e} \quad (\text{E.4})$$

$$\beta = P(\text{erro tipo II}) = P(\text{falhar em rejeitar } H_0 | H_1) \quad (\text{E.5})$$

Conforme lema de Neyman-Pearson (1933), o melhor teste para H_0 versus H_1 é a razão de verossimilhança, enunciada como

$$\text{rejeitar } H_0 \text{ se } \frac{\prod_{i=1}^n f_1(X_i)}{\prod_{i=1}^n f_0(X_i)} > C \quad (\text{E.6})$$

$$\text{Sejam } L_0(x_1, x_2, \dots, x_n) = \prod_{i=1}^n f_0(X_i) \quad \text{e} \quad (\text{E.7})$$

$$L_1(x_1, x_2, \dots, x_n) = \prod_{i=1}^n f_1(X_i) \quad (\text{E.8})$$

as funções de verossimilhança sob as hipóteses H_0 e H_1 . A razão $\Lambda_n = \frac{L_1}{L_0}$ é chamada de razão de verossimilhança sob as hipóteses H_0 e H_1 . O teste da razão de verossimilhança rejeita H_0 quando a razão $\Lambda_n = \frac{L_1}{L_0}$ é maior que o valor C pré-estabelecido.

Segundo Neyman e Pearson (1933), o princípio do teste sequencial é o de coletar amostras até que se tenha forte evidência ou de H_0 ou de H_1 . Continua-se coletando amostras até que a razão de verossimilhança seja ou muito grande, e neste caso H_1 parece ser a hipótese provável, ou muito pequena; neste caso H_0 parece ser a hipótese provável. Coleta-se amostras enquanto a razão de verossimilhança se mantenha entre os limites A e B , como em (E.9)

$$A \leq \frac{L_1(x_1, x_2, \dots, x_n)}{L_0(x_1, x_2, \dots, x_n)} \leq B \quad (\text{E.9})$$

Os limites A e B são tais que indicam as regiões de evidência da hipótese H_0 ou dada hipótese H_1 e a região em que se um ponto for demarcado deve-se continuar amostrando. A figura E.2 ilustra estas regiões.

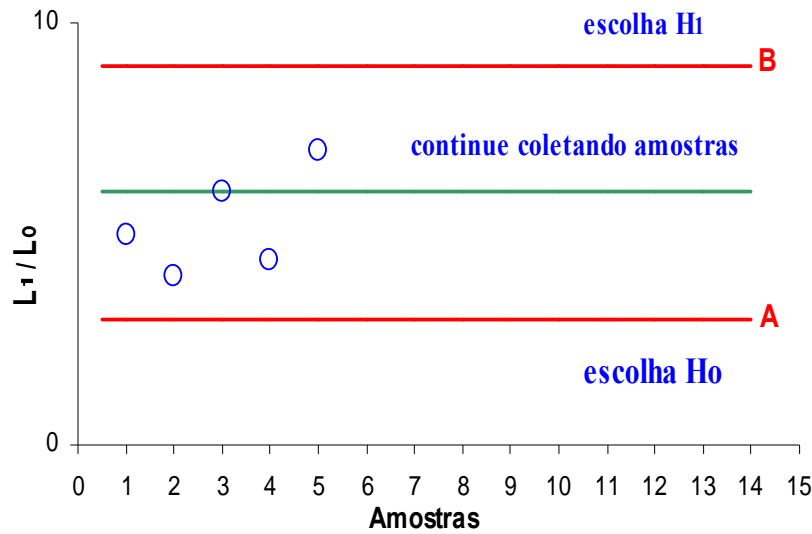


Figura E.2 Regiões para escolha das alternativas H_0 e H_1

Seja o conjunto A_m tal que

$$A_m = \{(x_1, x_2, \dots, x_m) \mid \text{encerra a amostragem após precisamente } m \text{ observações e escolhe } H_1\}$$

Então para todos $(x_1, x_2, \dots, x_m) \in A_m$ a razão de verossimilhança é tal que

$$\frac{L_1(x_1, x_2, \dots, x_m)}{L_0(x_1, x_2, \dots, x_m)} > B$$

Então $\sum_{(x_1, x_2, \dots, x_m) \in A_m} L_1(x_1, x_2, \dots, x_m) > B \times \sum_{(x_1, x_2, \dots, x_m) \in A_m} L_0(x_1, x_2, \dots, x_m)$, o que representa as probabilidades

$$P(\text{escolher } H_1 \mid H_1) > B \times P(\text{escolher } H_1 \mid H_0)$$

$$1 - \beta > B \times \alpha$$

De maneira similar o limite A é obtido

$$A \times (1 - \alpha) > \beta$$

Testando-se a hipótese de ocorrência de um desvio $\delta = \mu_1 - \mu_0$ em um processo normal, tem-se:

$$H_0 = \mu = \mu_0$$

$$H_1 = \mu = \mu_0 + \delta$$

Para $x_1, x_2, \dots, x_n \sim N(\mu, \sigma^2)$ independentes e identicamente distribuídas, em que δ é o desvio conhecido. Sem perda de generalidade, admitindo-se que $\mu_1 > \mu_0$. Tendo observado n observações, a razão de verossimilhança será

$$\Lambda_n = \frac{\prod_{i=1}^n \left\{ \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x_i - \mu_0)^2}{2\sigma^2}\right] \right\}}{\prod_{i=1}^n \left\{ \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x_i - \mu_0 - \delta)^2}{2\sigma^2}\right] \right\}} \quad (\text{E.10})$$

A seguir a demonstração algébrica para a razão de verossimilhança, Λ_n associada as constantes A e B, ou seja aos limites que indicam no gráfico as regiões para escolha das alternativas H_0 e H_1 .

$$\begin{aligned} \Lambda_n &= \frac{L_1(x_1, x_2, \dots, x_n)}{L_o(x_1, x_2, \dots, x_n)} = \frac{\prod_{i=1}^n \left\{ \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x_i - \mu_0)^2}{2\sigma^2}\right] \right\}}{\prod_{i=1}^n \left\{ \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x_i - \mu_0 - \delta)^2}{2\sigma^2}\right] \right\}} \\ \Lambda_n &= \prod_{i=1}^n \exp\left[-\frac{(x_i - \mu_0)^2}{2\sigma^2} + \frac{(x_i - \mu_0 - \delta)^2}{2\sigma^2}\right] \\ \Lambda_n &= \exp\left[-\frac{(x_1 - \mu_0)^2}{2\sigma^2} + \frac{(x_1 - \mu_0 - \delta)^2}{2\sigma^2}\right] \times \exp\left[-\frac{(x_2 - \mu_0)^2}{2\sigma^2} + \frac{(x_2 - \mu_0 - \delta)^2}{2\sigma^2}\right] \times \\ &\dots \times \exp\left[-\frac{(x_n - \mu_0)^2}{2\sigma^2} + \frac{(x_n - \mu_0 - \delta)^2}{2\sigma^2}\right] = \\ \Lambda_n &= \exp\left[\frac{-\delta^2 + 2x_1\delta - 2\mu_0\delta}{2\sigma^2}\right] \times \exp\left[\frac{-\delta^2 + 2x_2\delta - 2\mu_0\delta}{2\sigma^2}\right] \times \exp\left[\frac{-\delta^2 + 2x_n\delta - 2\mu_0\delta}{2\sigma^2}\right] \\ \Lambda_n &= \exp\left[-\frac{(\mu_1 - \mu_0)^2}{2\sigma^2} + \frac{(\mu_1 - \mu_0)}{\sigma^2}(x_1 - \mu_0)\right] \times \exp\left[-\frac{(\mu_1 - \mu_0)^2}{2\sigma^2} + \frac{(\mu_1 - \mu_0)}{\sigma^2}(x_2 - \mu_0)\right] \\ &\dots \times \exp\left[-\frac{(\mu_1 - \mu_0)^2}{2\sigma^2} + \frac{(\mu_1 - \mu_0)}{\sigma^2}(x_n - \mu_0)\right] \\ \Lambda_n &= \exp\left[-\frac{n\delta^2}{2\sigma^2} + \frac{\delta}{\sigma^2} \sum_{i=1}^n (x_i - \mu_0)\right] \end{aligned}$$

Logo a razão de verossimilhança é tal que

$$\begin{aligned} \Lambda_n &= \frac{L_1(x_1, x_2, \dots, x_n)}{L_o(x_1, x_2, \dots, x_n)} = \exp\left[\frac{\delta}{\sigma^2} \sum_{i=1}^n (x_i - \mu_0) - \frac{n\delta^2}{2\sigma^2}\right]. \text{ Como } \delta = \mu_1 - \mu_0, \text{ então} \\ \Lambda_n &= \frac{L_1(x_1, x_2, \dots, x_n)}{L_o(x_1, x_2, \dots, x_n)} = \exp\left[\frac{\mu_1 - \mu_0}{\sigma^2} \sum_{i=1}^n (x_i - \mu_0) - \frac{n(\mu_1 - \mu_0)^2}{2\sigma^2}\right] \quad (\text{E.11}) \end{aligned}$$

O procedimento de teste é o de coletar amostras enquanto a razão de verossimilhança esteja entre as constantes (limites) A e B, então

$$\begin{aligned}
 A &\leq \exp \left[\frac{\mu_1 - \mu_0}{\sigma^2} \sum_{i=1}^n (x_i - \mu_0) - \frac{n(\mu_1 - \mu_0)^2}{2\sigma^2} \right] \leq B \\
 \ln A &\leq \left[\frac{\mu_1 - \mu_0}{\sigma^2} \sum_{i=1}^n (x_i - \mu_0) - \frac{n(\mu_1 - \mu_0)^2}{2\sigma^2} \right] \leq \ln B \\
 \frac{\sigma^2}{\mu_1 - \mu_0} \left\{ \ln A + \frac{n(\mu_1 - \mu_0)^2}{2\sigma^2} \right\} &\leq \sum_{i=1}^n (x_i - \mu_0) \leq \frac{\sigma^2}{\mu_1 - \mu_0} \left\{ \ln B + \frac{n(\mu_1 - \mu_0)^2}{2\sigma^2} \right\} \quad (E.12)
 \end{aligned}$$

em que os limites A e B são dados por $A = \frac{\beta}{1 - \alpha}$ e $B = \frac{1 - \beta}{\alpha}$

O monitoramento de uma seqüência de variáveis aleatórias independentes e normalmente distribuídas como procedimento CUSUM está intimamente relacionado com os testes seqüenciais da razão de probabilidades (SPRT) de Wald.

Os gráficos de controle podem ser interpretados como testes seqüenciais para interpretar a hipótese nula (H_0) indicando sob controle e a hipótese alternativa (H_1) indicando fora de controle. Uma das aplicações de teste seqüencial da razão de probabilidades (SPRT) em gráficos de controle CUSUM é a de continuar acompanhando o processo até a emissão de um sinal fora de controle, ou seja, até $\Lambda_n > A$.

Page desenvolveu em 1954 na Inglaterra o gráfico de controle Somas Acumuladas (CUSUM) como uma alternativa viável ao gráfico \bar{X} de Shewhart para detectar mudanças no processo de pequena magnitude. Este gráfico é baseado nas somas acumuladas das médias ou das observações individuais e é derivado do SPRT. No procedimento de soma acumulada deste gráfico é incorporado diretamente, toda a seqüência de informações para demarcar as somas acumuladas dos desvios de \bar{x}_j em relação ao valor-alvo μ_0 (valor nominal). Supondo que amostras de tamanho $n \geq 1$ são coletadas, \bar{x}_j é a média de j-ésima amostra e μ_0 é o valor desejado para a média do processo, a estatística CUSUM é formada demarcando a quantidade da equação (E.13) com a amostra i

$$\begin{aligned}
 C_1 &= (\bar{x}_1 - \mu_0) \\
 C_2 &= (\bar{x}_1 - \mu_0) + (\bar{x}_2 - \mu_0) = C_1 + (\bar{x}_2 - \mu_0) \\
 C_3 &= (\bar{x}_1 - \mu_0) + (\bar{x}_2 - \mu_0) + (\bar{x}_3 - \mu_0) = C_2 + (\bar{x}_3 - \mu_0) \\
 &\dots \\
 C_i &= \sum_{j=1}^i (\bar{x}_j - \mu_0) = C_{i-1} + (\bar{x}_i - \mu_0) \quad (E.13)
 \end{aligned}$$

onde C_i é a soma acumulada incluindo a i -ésima amostra, uma vez que combinam informações de diversas amostras.

Se o processo permanecer sob controle para o valor desejado μ_o , as somas acumuladas definidas em (E.13) descrevem um percurso aleatório com média zero. Porém, se a média muda para algum valor acima $\mu_1 > \mu_o$, então a tendência ascendente se desenvolverá na soma acumulada C_i . Reciprocamente, se a média muda para algum valor abaixo $\mu_1 < \mu_o$, a soma acumulada C_i terá uma direção negativa. Por esta razão, se nos pontos demarcados aparecer uma tendência para cima ou para baixo, deve-se considerar isto como uma evidência de que a média do processo mudou e uma busca de causas assinaláveis deve ser realizada.

Page (1961) ao aperfeiçoar o gráfico CUSUM reconhece explicitamente que as suas somas acumuladas (CUSUM) resultam em uma sucessão de testes sequenciais de Wald. O gráfico de controle CUSUM unilateral pode ser pensado como repetidos testes sequenciais de razão de probabilidades. A principal diferença entre CUSUM e SPRT de Wald é que na soma acumulada a hipótese de controle, (H_o), nunca é aceita. Por isso, é aconselhável que reinicie o teste cada vez que a evidência favorece a hipótese de que o processo está sob controle. Sempre que a soma acumulada Z_i for negativa, a evidência favorece a hipótese de controle, (H_o). Quando isto acontece, considera-se esta soma igual a zero, ou em termos algébricos,

$$C_i^+ = \max(0, C_{i-1}^+ - k + Z_i) \quad (\text{E.14})$$

A estatística C_i^+ é comparada a uma constante h^+ . Se $C_i^+ > h^+$, o gráfico CUSUM indica que ocorreu um aumento na média do processo. Para um tamanho de mudança, $\delta < 0$, o SPRT similarmente ao procedimento de soma acumulada revela uma redução na média do processo. Neste caso, conforme equação (E.15)

$$C_i^- = \max(0, C_{i-1}^+ - k - Z_i) \quad (\text{E.15})$$

o gráfico CUSUM é comparado a uma constante h^- . Se $C_i^- < h^-$, o gráfico indica que ocorreu uma redução na média do processo.

Na abordagem do gráfico CUSUM, rejeita-se a hipótese nula (H_o) em favor da hipótese alternativa (H_1). O escore da função Z_i passa a ser a mesma função definida no SPRT; é uma função de X_i , e depende de alguns parâmetros sob controle e fora de controle que possa ter para as distribuições de $f_o(x_i)$ e $f_1(x_i)$.

APÊNDICE F - A Função Perda de Taguchi para a Tomada de Decisões

F.1 Função Perda de Taguchi (Univariada e Multivariada)

Um método baseado na função perda de Taguchi é proposto neste trabalho para determinar sob a ótica de regiões de máxima e mínima potência em processos multivariados, a região sob controle de $d \leq A$, e a região onde devemos detectar o mais rápido possível uma mudança $d \geq B$.

O objetivo de se buscar a otimização de um processo produtivo multivariado não é o de apenas manter as características de qualidade estáveis, dentro de limites previamente estabelecidos em projetos: é, também, o de reduzir a variabilidade em torno de um valor médio enfocado por Chou, Liu e Huang (2002).

A qualidade de um produto segundo a filosofia de Taguchi é a perda mínima imposta por este produto a sociedade no decorrer de sua utilização Kapur e Cho, (1996). Do ponto de vista social este conceito de qualidade é único porque inclui fabricantes, clientes e a sociedade como um todo. A filosofia Taguchi enfatiza a importância econômica de alcançar uma baixa variabilidade, uma coerência funcional e sobretudo uma alta qualidade.

No conceito tradicional de avaliação de um sistema de qualidade, um produto é classificado como não conforme se a característica da qualidade do produto sob controle se encontra fora dos limites de especificação incorrendo em uma determinada perda econômica. No caso contrário, o produto é classificado como conforme e sem perda econômica. Neste enfoque caracterizado com filosofia ocidental utiliza-se uma avaliação binário de qualidade com especificações ou tolerâncias onde por exemplo, uma peça é classificada como absolutamente conforme se está dentro de tolerâncias e como não conforme quando está fora de tolerâncias.

O enfoque japonês de Taguchi contempla a avaliação contínua da qualidade de um produto baseado na perda econômica que significa a variação das características da qualidade referente aos valores nominais definidos. Para Taguchi a perda é nula somente quando o valor da característica da qualidade coincide com o valor nominal e aumenta de forma contínua à medida que se afasta do valor nominal ainda que se cumpram as especificações. As idéias anteriores se modelam na conhecida função perda de Taguchi para a tomada de decisões. Esta

função avalia a perda, em termos econômicos, de um produto em relação com o valor de uma de suas características quantitativas da qualidade mediante uma função quadrática do tipo:

$$L(y) = k(y - T)^2 \tag{F.1}$$

onde y é o valor da característica da qualidade de Y considerado, T é o valor nominal (valor alvo), k é o coeficiente de perda (constante de proporcionalidade) e $L(y)$ é a perda econômica produzida pelo fato da característica da qualidade de interesse desviar-se do valor nominal. A utilização desta função perda $L(y)$ é fundamental para atender propósitos gerenciais como a valor monetário da consequência de qualquer aperfeiçoamento em qualidade. Esta expressão matemática estabelece uma medida financeira para o cálculo do desvio de uma característica do produto com relação ao valor nominal.

Consideremos que o processo se encontra sob controle e centrado no valor nominal T . Os limites das especificações são T_1 e T_2 . Seja C_A o custo ou perda depreciável quando o desvio ocorrido em relação ao valor nominal é y_A e C_B o custo ou perda depreciável quando o desvio ocorrido em relação ao valor nominal é y_B . Podemos escrever $C_A=L(y_A)$ e $C_B=L(y_B)$. Com isso, se cumpre que $T \leq y_A \leq y_B \leq T_2$ ou também $T_1 \leq y_B \leq y_A \leq T$, conforme figura F.1

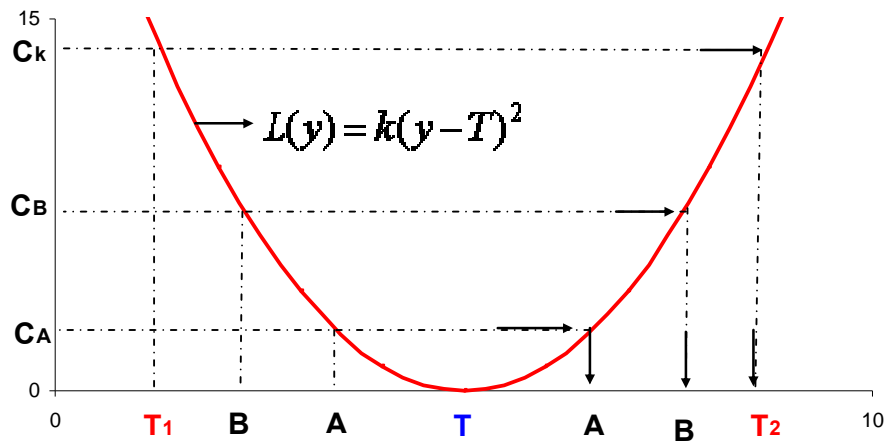


Figura F.1 Função perda univariada de Taguchi

A função perda de Taguchi dada pela equação F.1 neste caso, toma-se na seguinte forma

$$L(y_A) = k(y_A - T)^2$$

$$L(y_B) = k(y_B - T)^2 \tag{F.2}$$

onde para determinar k devemos conhecer o valor de $L(y)$ associada a um valor da variável Y diferente do valor nominal que pode ser tanto o custo como a perda associada a um valor que

se encontra fora dos limites de especificações. A este valor de, $L(y_k)$, chamamos de C_k e aplicando a equação (F.2) se obtém o valor de k :

$$k = \frac{L(y_k)}{(y_k - T)^2} \tag{F.3}$$

Conhecido o valor de descentralização $d = \frac{|y - T|}{\sigma}$ e os valores dos custos C_A , C_B e C_k e os desvios do valor nominal y_A e y_B é possível determinar o desvio relativo ao valor nominal em unidade de σ .

$$\begin{aligned} d_A &= |y_A - T| \\ d_B &= |y_B - T| \end{aligned} \tag{F.4}$$

Estes são os valores que delimitam as regiões de máxima e mínima potência, ou seja, $A = d_A$ e $B = d_B$. A figura F.2 ilustra os pontos que delimitam tais regiões.

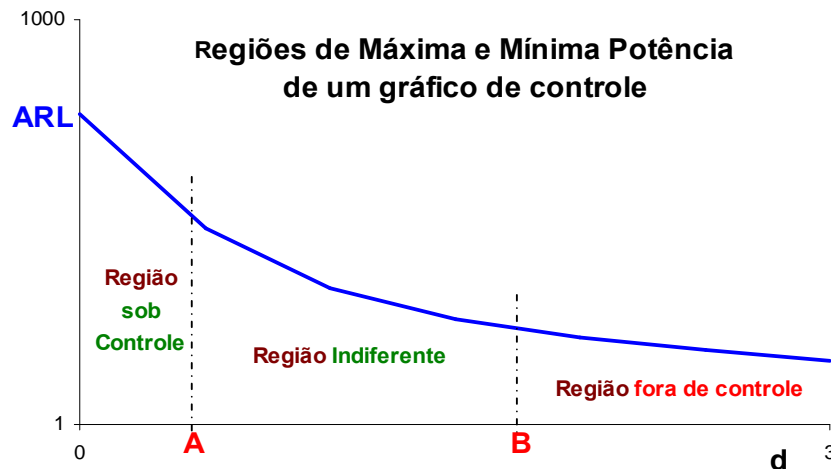


Figura F.2 Pontos A e B que delimitam as regiões de máxima e mínima potência

Quando se fala da fabricação de um produto se faz referência ao conjunto de unidades fabricadas pelo processo considerado. Tais unidades fabricadas apresentarão variabilidade referente a característica da qualidade considerada onde cada uma delas terá qualidade diferente. Esta qualidade se obtém encontrando o valor esperado da função de perdas.

Seja Y uma variável aleatória de média μ e desvio padrão σ , a perda média por peça fabricada é:

$$E[L(y)] = k[\sigma^2 + (\mu - T)^2] \tag{F.5}$$

onde $E[L(y)]$ é o valor esperado da função de perdas e $k = \frac{A}{(LSE - T)^2}$. Os valores de A e

LSE representam respectivamente, o custo de se produzir um produto fora do especificado e o limite superior especificado para a característica da qualidade em questão.

Agora considerando o controle de qualidade de um processo multivariado onde desejamos monitorar simultaneamente p variáveis que estão correlacionadas entre si. De forma análoga ao caso univariado, o produto ou o processo transmitirá uma perda à variabilidade associada ao afastamento de algumas das característica da qualidade consideradas. Como em um processo multivariado as características de qualidade usualmente estão correlacionadas entre si e geralmente a perda global não é a soma das perdas individuais de cada uma das características da qualidade torna-se necessário o desenvolvimento de uma função de perda multivariada.

Kapur e Cho (1996) estende o conceito da função perda univariada de Taguchi no caso multivariado. Suponhamos que y_1, y_2, y_p são os valores de p características da qualidade de uma mesma unidade de um determinado produto e T_1, T_2, \dots, T_p são, respectivamente, os valores nominais das p características de qualidade consideradas. Desenvolvendo em modelo de séries de Taylor a função perda e ignorando os termos de ordem superior, a função de perda multivariada pode ser expressa como:

$$L(y_1, y_2, \dots, y_p) = \sum_{i=1}^p \sum_{j=1}^i k_{ij} (y_i - T_i)(y_j - T_j) \quad (\text{F.6})$$

onde $L(y_1, y_2, \dots, y_p)$ é a perda transmitida à sociedade e onde $k_{ij} = k_{ji}$ é a constante de proporcionalidade que especifica os custos do desvio da característica da qualidade ao afastar-se de seus valores nominais. Por exemplo, k_{11} é o custo individual por unidade pelo fato da característica de qualidade y_1 desviar-se de T_1 e k_{12} é o custo por unidade adicional incluído quando as características y_1 e y_2 estão simultaneamente fora de seus valores nominais T_1 e T_2 .

No caso particular de que as características de qualidade y_1 e y_2 sejam independentes então $k_{ij} = 0$ e a equação (F.6) se reduz a soma de p funções de perda independentes Teeravaraprug e Cho (2002). As constantes k_{ij} podem ser determinadas mediante métodos de regressão Kapur e Cho (1996).

F.2 Pontos A e B que delimitam das Regiões de Máxima e Mínima Potência de um gráfico de controle

Geralmente em um produto existem duas ou mais características quantitativas de qualidade. O imediato é pensar que podemos monitorar o processo de fabricação mediante o

controle de perda de uma das características separadamente e somar as perdas individuais. Isto é incorreto pois em geral, as perdas não são aditivas já que existe um efeito conjunto dos desvios de pares de variáveis sobre a perda global. Com relação, por exemplo no caso bivariado, (p=2) a equação (F.6) pode ser escrita como:

$$L(y_1, y_2) = k_{11}(y_1 - T_1) + k_{22}(y_2 - T_2) + 2k_{12}(y_1 - T_1)(y_2 - T_2) \quad (F.7)$$

No caso em que y_1 e y_2 são independentes ($k_{12}=0$), temos que $L(y_1, y_2) = L(y_1) + L(y_2)$

Por exemplo, seja um processo bivariado cujas características da qualidade tenham como valores nominais dados por T, cuja matriz de variâncias-covariâncias seja Σ com uma matriz de coeficientes de perda K

$$T = \begin{pmatrix} T_1 \\ T_2 \end{pmatrix} \quad \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} \quad K = \begin{pmatrix} k_{11} & k_{12} \\ k_{21} & k_{22} \end{pmatrix}$$

Dada uma perda conjunto de valor m, por exemplo, representada pela linha de contorno $C = m$ e considerando os valores da variável bivariada representada pelos vetores y_A , y_B e y_C então:

$$L(y_{A_1}, y_{A_2}) = L(y_{B_1}, y_{B_2}) = L(y_{C_1}, y_{C_2}) = m \quad (F.8)$$

Agora, o problema a ser resolvido é medir em uma peça produzida o valor do desvio da característica da qualidade em relação ao valor nominal. Em análise estatística multivariada a distância de Mahalanobis mede o tamanho de mudança de um vetor. Em nosso caso é a distância entre o vetor de valores nominais T e o vetor de valores obtidos $Y(y_1, y_2)$:

$$d = \sqrt{(y_1 - T_1)' \Sigma^{-1} (y_1 - T)} \quad (F.9)$$

Esta é portanto a forma que vamos utilizar para medir o afastamento sobre o valor nominal.

APÊNDICE G - Método da Secante

O método da secante é um método iterativo de análise numérica bastante utilizado na prática para determinadas funções cuja derivada é muito complicada ou sujeita a erros e com isso não é viável determinar $f'(x)$. Uma das grandes dificuldades com o método de Newton é determinar a derivada $f'(x)$ para algumas funções. Esta exigência acrescenta uma sobrecarga no computador, e pode ainda acrescentar um custo computacional para a resolução do problema. Na verdade, é uma alternativa ao método de Newton pois há uma substituição da derivada por uma linha secante. A figura G.1 ilustra geometricamente o método que inicia a secante com suposição inicial de dois pontos x_0 e x_1 , onde $f(x_0) \neq f(x_1)$. Nesse método, partimos de duas aproximações x_0 e x_1 e determinamos a reta que passa por $[x_0, f(x_0)]$ e $[x_1, f(x_1)]$. A intersecção dessa reta com o eixo $x = 0$ determina a próxima iteração x_2 . Continuamos o processo a partir de x_2 . Temos assim o seguinte processo iterativo:

$$x_{i+2} = \frac{x_i f(x_{i+1}) - x_{i+1} f(x_i)}{f(x_{i+1}) - f(x_i)} \quad (\text{G.1})$$

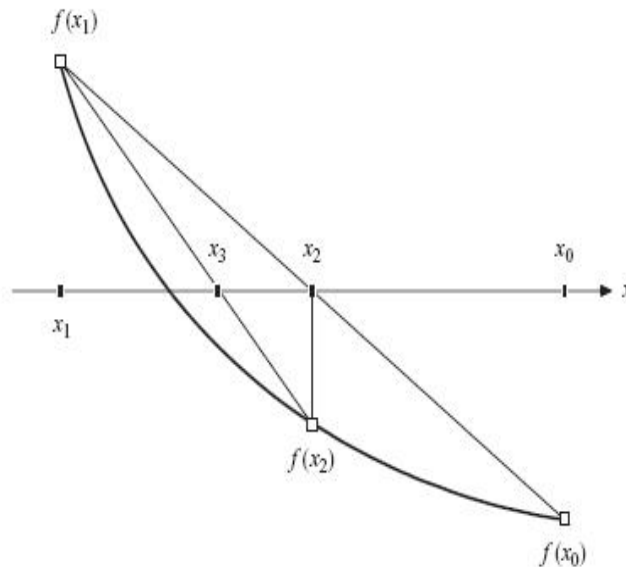


Figura G.1 Método da Secante

Neste trabalho é desenvolvido um programa em ambiente Matlab denominado Programa Computacional 2 para determinar o limite superior de controle h (intervalo de decisão) apropriado do gráfico MCUSUM para o valor desejado de ARL quando o processo

está sob controle. Para tal programa, este método iterativo de análise numérica denominado Método da Secante é aplicado para aproximar a solução de h à equação não linear $ARL(0|h) - ARL_0 = 0$. A saída dos dados deste programa inclui as sucessivas aproximações para o valor de h em um determinado número de iterações.

A figura G.2 ilustra um exemplo de aplicação do método da secante desenvolvido em ambiente MS-Excel para a função $f(x) = x^3 - x^2 - 2$

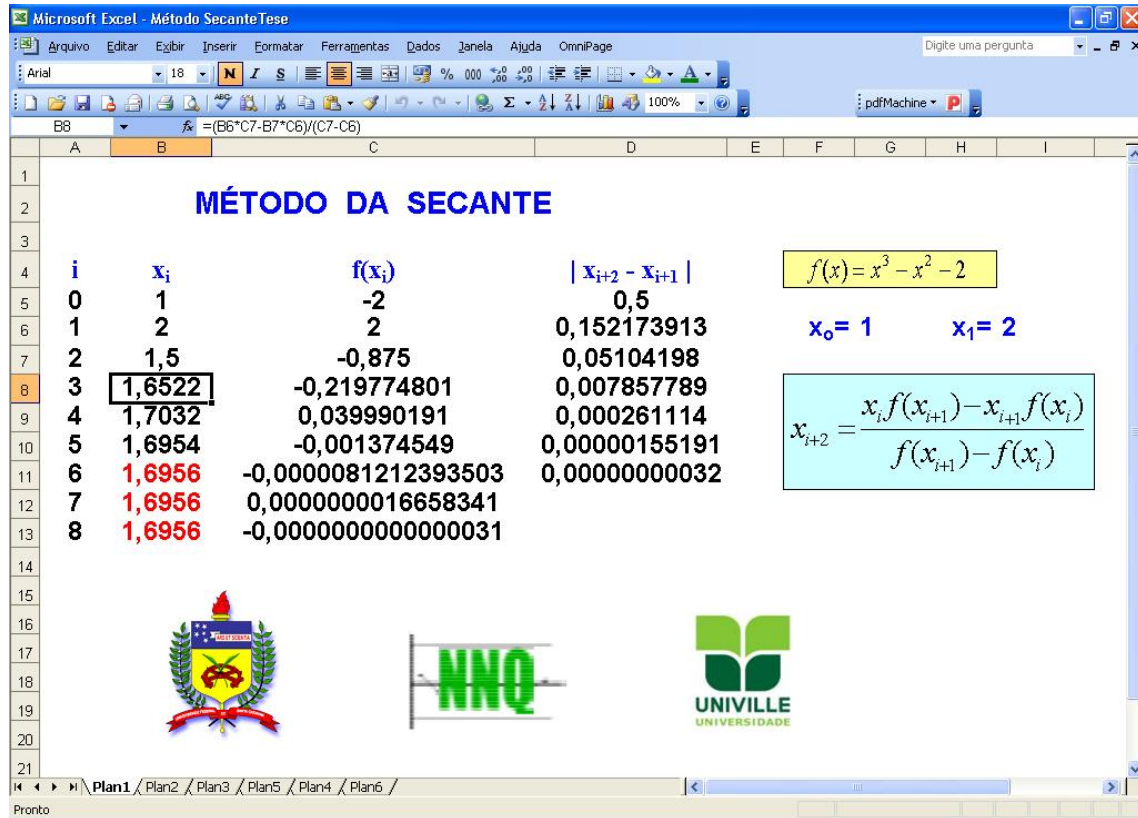


Figura G.2 Exemplo de aplicação do Método da Secante

APÊNDICE H - Interpolação Polinomial

A interpolação é um procedimento no qual uma fórmula matemática é usada para fornecer o valor exato dos pontos pertencentes a um conjunto de dados e um valor estimado entre esses pontos. Abordaremos aqui a interpolação polinomial empregando-se um único polinômio, para qualquer número de pontos envolvidos, ou seja, para n pontos existe um polinômio de ordem $n-1$ que passa por todos esses pontos. Para 5 pontos, por exemplo, o polinômio é de 4ª ordem.

Consideremos um conjunto de pontos (designados nós de interpolação) x_0, x_1, \dots, x_n a que estão associados os valores de uma função: f_0, \dots, f_n , respectivamente. Pretendemos encontrar um polinômio:

$$p(x) : p(x_i) = f_i \quad \text{para } i = 0, \dots, n.$$

O polinômio de 3º grau interpola a função em 4 pontos conforme figura H.1

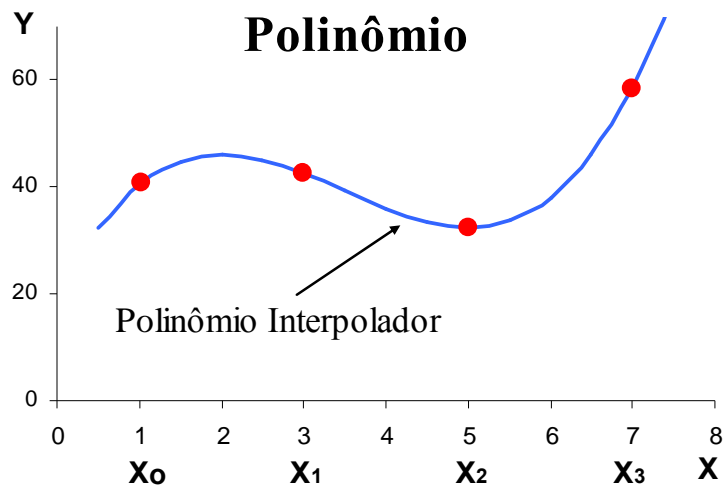


Figura H.1 Polinômio de 3º grau

Escrevendo $p(x) = a_0 + a_1 x + \dots + a_m x^m$, obtemos o sistema:

$$a_0 + a_1 x_0 + \dots + a_m x_0^m = f_0$$

...

$$a_0 + a_1 x_n + \dots + a_m x_n^m = f_n$$

e para que este sistema seja possível e determinado é pelo menos necessário que $m=n$.

Obtemos assim o sistema linear :

$$\begin{bmatrix} 1 & x_0 & \dots & x_0^n \\ 1 & x_1 & \dots & x_1^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \dots & x_n^n \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ f_n \end{bmatrix}$$

em que a matriz do sistema é conhecida como Matriz de Vandermonde.

A existência e unicidade do polinômio interpolador é equivalente a assegurar que o sistema é possível e determinado para quaisquer x_0, \dots, x_n distintos.

Teorema:

Dados $n+1$ nós, x_0, \dots, x_n e os respectivos valores f_0, \dots, f_n , existe um e um só, polinômio interpolador de grau $\leq n$, para esses valores.

Demonstração:

Supondo que existem dois polinômios interpoladores p e q de grau $\leq n$, então o polinômio $p(x) - q(x)$ tem grau $\leq n$ e $n+1$ raízes, já que, sendo polinômios interpoladores, verificam :

$$p(x_i) = f_i = q(x_i) \quad \text{para } i = 0, \dots, n.$$

Conseqüentemente, como tem $n+1$ raízes e grau $\leq n$, o polinômio $p(x)-q(x)$ terá que ser nulo, logo $p=q$

Neste trabalho, aplicou-se interpolação polinomial para obter-se os valores estimados para o ARL quando o processo está fora de controle e com estes valores realizar a análise de sensibilidade.

APÊNDICE I - Relatórios: resultados obtidos com a resolução dos modelos matemáticos

I.1 Relatórios de resposta, análise de sensibilidade e limites do modelo matemático formulado para os dados do exemplo de aplicação em ambiente MS-Excel (Solver)

$$T = \begin{bmatrix} 8 \\ 20 \end{bmatrix} \quad \Sigma = \begin{bmatrix} 0,49 & 0 \\ 0 & 0,25 \end{bmatrix} \quad K = \begin{bmatrix} 3 & 0,5 \\ 0,5 & 2 \end{bmatrix}$$

$$Y(y_1, y_2) \rightarrow Y(x, y)$$

$$\begin{bmatrix} x & -8 \\ y & -20 \end{bmatrix}^T \begin{bmatrix} 2,041 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} x & -8 \\ y & -20 \end{bmatrix}$$

$$2x^2 - 33x + 1730,6 + -160y + 4y^2$$

$$x = 8,178$$

$$y = 20,022$$

$$d = 0,26 \text{ (Ponto A)}$$

$$s.a: 0,100 = 0,1$$



Microsoft Excel 11.0 Relatório de resposta
Planilha: [Otimização do Gráfico MCUSUM.xls]
Pontos de máxima e mínima potência
Relatório criado: 10/1/2009 17:12:21

Célula de destino (Min)

Célula	Nome	Valor original	Valor final
\$H\$21	d = S =	0,26	0,26

Células ajustáveis

Célula	Nome	Valor original	Valor final
\$H\$18	x = S =	8,178	8,178
\$H\$19	y = S =	20,022	20,022

Restrições

Célula	Nome	Valor da célula	Fórmula	Status	Transigência
\$H\$23	s.a: S =	0,100	\$H\$23=\$J\$23	Sem agrupar	0
\$H\$18	x = S =	8,178	\$H\$18>=0	Sem agrupar	8,178
\$H\$19	y = S =	20,022	\$H\$19>=0	Sem agrupar	20,022

Microsoft Excel 11.0 Relatório de sensibilidade
Planilha: Planilha: [Otimização do Gráfico MCUSUM.xls]
Pontos de máxima e mínima potência
Relatório criado: 10/1/2009 17:12:21

Células ajustáveis

Célula	Nome	Final Valor	Reduzido Gradiente
\$H\$18	x = S =	8,178	0,000
\$H\$19	y = S =	20,022	0,000

Restrições

Célula	Nome	Final Valor	Lagrange Multiplicador
\$H\$23	s.a: S =	0,100	1,291

Microsoft Excel 11.0 Relatório de limites
Planilha: Planilha: Planilha: [Otimização do Gráfico MCUSUM.xls]
Pontos de máxima e mínima potência
Relatório criado: 10/1/2009 17:12:21

Destino

Célula	Nome	Valor
\$H\$21	d = S =	0,26

Ajustável

Célula	Nome	Valor	Inferior Limite	Destino Resultado	Superior Limite	Destino Resultado
\$H\$18	x = S =	8,178	8,178	0,258	8,178	0,258
\$H\$19	y = S =	20,022	20,022	0,258	20,022	0,258

NNQ

UNIVILLE
UNIVERSIDADE

$$T = \begin{bmatrix} 8 \\ 20 \end{bmatrix} \quad \Sigma = \begin{bmatrix} 0,5 & 0 \\ 0 & 0,25 \end{bmatrix} \quad K = \begin{bmatrix} 3 & 0,5 \\ 0,5 & 2 \end{bmatrix}$$

$$Y(y_1, y_2) \rightarrow Y(x, y)$$

$$\begin{bmatrix} x & -8 \\ y & -20 \end{bmatrix}^T \quad \begin{bmatrix} 2,041 & 0 \\ 0 & 4 \end{bmatrix} \quad \begin{bmatrix} x & -8 \\ y & -20 \end{bmatrix}$$

$$2x^2 - 33x + 1730,6 - 160y + 4y^2$$

$x = 7,204$
 $y = 19,901$
 $d = 1,15$ (Ponto B)
 s.a: 2,000 = 2



Microsoft Excel 11.0 Relatório de resposta
 Planilha: [Otimização do Gráfico MCUSUM.xls]
 Pontos de Máxima e mínima potência
 Relatório criado: 10/1/2009 19:02:41

Célula de destino (Mín)

Célula	Nome	Valor original	Valor final
\$H\$21	d = S =	1,15	1,15

Células ajustáveis

Célula	Nome	Valor original	Valor final
\$H\$18	x = S =	7,204	7,204
\$H\$19	y = S =	19,901	19,901

Restrições

Célula	Nome	Valor da célula	Fórmula	Status	Transigência
\$H\$23	s.a: S =	2,000	\$H\$23=\$J\$23	Sem agrupar	0
\$H\$18	x = S =	7,204	\$H\$18>=0	Sem agrupar	7,204
\$H\$19	y = S =	19,901	\$H\$19>=0	Sem agrupar	19,901

Microsoft Excel 11.0 Relatório de sensibilidade
 Planilha: [Otimização do Gráfico MCUSUM.xls]
 Pontos de Máxima e mínima potência
 Relatório criado: 10/1/2009 19:02:41

Células ajustáveis

Célula	Nome	Final Valor	Reduzido Gradiente
\$H\$18	x = S =	7,204	0,000
\$H\$19	y = S =	19,901	0,000

Restrições

Célula	Nome	Final Valor	Lagrange Multiplicador
\$H\$23	s.a: S =	2,000	0,289

Microsoft Excel 11.0 Relatório de limites
 Planilha: [Otimização do Gráfico MCUSUM.xls]
 Pontos de Máxima e mínima potência
 Relatório criado: 10/1/2009 19:02:41

Destino		
Célula	Nome	Valor
\$H\$21	d = S =	1,15

Ajustável		Inferior Limite	Destino Resultado	Superior Limite	Destino Resultado
Célula	Nome	Valor			
\$H\$18	x = S =	7,204	7,204	1,154	7,204
\$H\$19	y = S =	19,901	19,901	1,154	19,901

I.2 Relatórios de resposta, análise de sensibilidade e limites do modelo matemático formulado para os dados do Processo XY (furação de acabamento do furo 1) em ambiente MS-Excel (Solver)

Microsoft Excel 11.0 Relatório de resposta
 Planilha: [SolveTesorMin.5.xls]Plan1
 Relatório criado: 25/1/2009 22:27:00

Célula de destino (Mín)

Célula	Nome	Valor original	Valor final
\$H\$21	d = S =	8679,64	0,80

Células ajustáveis

Célula	Nome	Valor original	Valor final
\$H\$18	x = S =	0,000	5,011
\$H\$19	y = S =	0,000	103,253

Restrições

Célula	Nome	Valor da célula	Fórmula	Status	Transigência
\$H\$23	s.a: S =	0,0012	\$H\$23=\$J\$23	Sem agrupar	0
\$H\$18	x = S =	5,011	\$H\$18>=0	Sem agrupar	5,011
\$H\$19	y = S =	103,253	\$H\$19>=0	Sem agrupar	103,253

Microsoft Excel 11.0 Relatório de sensibilidade
 Planilha: [SolveTesorMin.5.xls]Plan1
 Relatório criado: 25/1/2009 22:27:00

Células ajustáveis

Célula	Nome	Final Valor	Reduzido Gradiente
\$H\$18	x = S =	5,011	0,000
\$H\$19	y = S =	103,253	0,000

Restrições

Célula	Nome	Final Valor	Lagrange Multiplicador
\$H\$23	s.a: S =	0,0012	335,3218

Microsoft Excel 11.0 Relatório de limites
 Planilha: [SolveTesorMin.5.xls]Relatório de limites 1
 Relatório criado: 25/1/2009 22:27:00

Destino		
Célula	Nome	Valor
\$H\$21	d = S =	0,80

Ajustável			Inferior Limite	Destino Resultado	Superior Limite	Destino Resultado
Célula	Nome	Valor				
\$H\$18	x = S =	5,011	5,011	0,804	5,011	0,804
\$H\$19	y = S =	103,253	103,253	0,804	103,253	0,804

Microsoft Excel 11.0 Relatório de resposta
Planilha: [SolveTesorMín.3.xls]Plan1
Relatório criado: 25/1/2009 22:29:08

Célula de destino (Mín)

Célula	Nome	Valor original	Valor final
\$H\$21	d = S =	8679,64	1,50

Células ajustáveis

Célula	Nome	Valor original	Valor final
\$H\$18	x = S =	0,000	5,021
\$H\$19	y = S =	0,000	103,256

Restrições

Célula	Nome	Valor da célula	Fórmula	Status	Transigência
\$H\$23	s.a: S =	0,0042	\$H\$23=\$J\$23	Sem agrupar	0
\$H\$18	x = S =	5,021	\$H\$18>=0	Sem agrupar	5,021
\$H\$19	y = S =	103,256	\$H\$19>=0	Sem agrupar	103,256

Microsoft Excel 11.0 Relatório de sensibilidade
Planilha: [SolveTesorMín.3.xls]Plan1
Relatório criado: 25/1/2009 22:29:08

Células ajustáveis

Célula	Nome	Final Valor	Reduzido Gradiente
\$H\$18	x = S =	5,021	0,000
\$H\$19	y = S =	103,256	0,000

Restrições

Célula	Nome	Final Valor	Lagrange Multiplicador
\$H\$23	s.a: S =	0,0042	178,9174

Microsoft Excel 11.0 Relatório de limites
Planilha: [SolveTesorMín.3.xls]Relatório de limites 1
Relatório criado: 25/1/2009 22:29:08

Destino		
Célula	Nome	Valor
\$H\$21	d = S =	1,50

Ajustável		
Célula	Nome	Valor
\$H\$18	x = S =	5,021
\$H\$19	y = S =	103,256

Inferior Limite	Destino Resultado
5,021	1,504
103,256	1,504

Superior Limite	Destino Resultado
5,021	1,504
103,256	1,504

APÊNDICE J - O R Project e Rotina Desenvolvida para gerar o Gráfico MCUSUM

J.1 Introdução

A internet, instrumento de globalização do conhecimento, vem de encontro a grande necessidade atual de se democratizar a informação cujo papel de extrema importância é direcionado a uma sociedade tecnologicamente desenvolvida. Conhecer algumas ferramentas computacionais livres que possam auxiliar usuários a resolver seus problemas em diversas áreas tem sido facilitado cada vez mais pelo acesso a rede mundial. Portanto, é natural a procura pelo desenvolvimento de aplicações que auxiliem usuários a terem acesso rápido a informações claras, precisas e objetivas.

A aplicação de pacotes computacionais a análise estatística multivariada de processos tais como MINITAB[®], STATISTICA[®], etc., vêm se fazendo cada vez mais presentes, cuja demanda tem sua origem tanto no âmbito da pesquisa nas universidades quanto no monitoramento de processos em ambientes industriais. Como característica comum está o fato de todos serem pacotes comerciais, implicando no pagamento de licenças para acompanhar as várias versões lançadas num curto espaço de tempo. Felizmente, programas computacionais livres, como por exemplo, o R Project tem se propagado principalmente no meio acadêmico na última década como uma ferramenta de computação estatística com uma enorme variedade de funções para análise de dados. Este recurso computacional livre segundo Alves, Henning e Samohyl (2008a) é fundamental para complementar algumas funções ainda não disponíveis na maioria dos softwares comerciais de análise estatística de dados, como por exemplo, o desenvolvimento do gráfico de controle MCUSUM. Para gerar este tipo de gráfico de controle desenvolveu-se neste trabalho uma rotina computacional no pacote GNU R.

J.2 O R Project

O R Project é uma linguagem e um ambiente para computação estatística. Como é um projeto GNU, baseado no conceito de software livre, pode ser usado sem custos de licença. Ele permite, de acordo com os autores, uma grande variedade de análises estatísticas como, por exemplo, análise exploratória de dados, testes estatísticos, regressão linear e não linear análise de séries temporais, entre diversas outras. Como ponto forte está a facilidade com que

gráficos bem delineados e de alta qualidade para impressão podem ser produzidos. Além disso, possui inúmeras funções para manipulação, importação e exportação de dados, sendo multi-plataforma, contendo versões para Windows, MacOS, GNU/Linux e Unix. É possível carregar dados externos das mais diversas formas, incluindo de planilhas eletrônicas, como o Microsoft Excel® e Open-Office, banco de dados e até de outros pacotes como MINITAB® e SPSS®, por exemplo. O termo ambiente pretende caracterizar o R como um sistema totalmente planejado e coerente ao invés de uma aglomeração de ferramentas muito específicas e inflexíveis, permitindo aos usuários acrescentar funcionalidade adicional por definição de novas funções.

J.3 Rotina Computacional no Pacote GNU R para gerar o MCUSUM do processo XY

A rotina computacional no pacote GNU R para gerar o gráfico MCUSUM (figura 4.20 - Capítulo 4) que é aplicado para monitorar simultaneamente as duas características da qualidade X_1 e Y_1 no processo XY (furação de acabamento do furo 1) conforme a seguir:

```
mcusum1=function(x,media=NULL,mc=NULL,k=0.5,h=5.5)
{
  x=as.matrix(x)
  if (is.null(mc)) {
    m=nrow(x)
    v=matrix(c(0),nrow=(nrow(x)-1),ncol=ncol(x))
    for (i in 1:(m-1)) {v[i,]=x[i+1,]-x[i,]}
    vt=t(v)
    vv=vt%*%v
    mc=matrix(c(0),ncol=ncol(x),nrow=ncol(x))
    for (i in 1:ncol(mc)) {mc[i,]=(1/(2*(m-1)))*(vv[i,])}
    if (is.null(media)) {
      media=matrix(c(0),ncol=1,nrow=ncol(x))
      for (i in 1:ncol(x)) {media[i,1]=mean(x[,i])}
      s=matrix(c(0),ncol=ncol(x),nrow=nrow(x))
      st=matrix(c(0),ncol=ncol(x),nrow=nrow(x))
      ci=matrix(c(0),ncol=1,nrow=nrow(x))
      p=matrix(c(0),ncol=1,nrow=nrow(x))
      vmc=matrix(c(0),ncol=ncol(x),nrow=nrow(x))
      for (i in 1:ncol(x)) {x[,i]=x[,i]-media[i,1]}
      s[1,]=s[1,]+x[1,]
      st[1,]=s[1,]%*% solve(mc)
      ts=t(s)
      ci[1]=sqrt(st[1,]%*% ts[,1])
      p[1]=1-k/ci[1]
      if (ci[1]>k) {vmc[1,]=(vmc[1,]+x[1,])*p[1]} else
(vmc[1,]=matrix(c(0),ncol=ncol(x)))
      for (i in 2:nrow(x)) {
        s[i,]=vmc[i-1,]+x[i,]
        st[i,]=s[i,]%*%solve(mc)
        ts=t(s)
        ci[i]=sqrt(st[i,]%*%ts[,i])
        p[i]=1-k/ci[i]
        if (ci[i]>k) {vmc[i,]=(vmc[i-1,]+x[i,])*p[i]}
      }
    }
  }
}
```

```
else (vmc[i,]=matrix(c(0),ncol=ncol(x)))}
  vmc1=vmc %*% solve(mc)
  tc=t(vmc)
  t2=0
  for (i in 1:nrow(x)){t2[i]=sqrt(vmc1[i,]%*%tc[,i])}
  pontos.fora=subset(t2,t2>h)
  amostra=seq(1,nrow(x))
  plot(amostra,t2,type='b',col=ifelse(t2>h,2,1),pch=ifelse(t2>h,2
3,19),bg='red',ylim=c(0,max(1.1*h,1.1*max(t2))),main='Gráfico de
Controle MCUSUM',xlab='Amostra',axes=FALSE)
  axis(1,1:nrow(x))
  axis(2)
  box()
  abline(h=h,lwd=2,col=2)
  text(2,h+1,paste("LS =",h))
  grid()
  structure(list(ESTATISTICA.MCUSUM=t2,matriz.cov=mc,Media=media,
H=h,K=k,pontos.fora=length(pontos.fora),valores=pontos.fora))
}
```

APÊNDICE K - Programas Computacionais Desenvolvidos em Ambiente Matlab

K.1 Introdução

Neste trabalho foram desenvolvidos três programas computacionais em ambiente Matlab para determinar conforme metodologia proposta os principais parâmetros do gráfico de controle MCUSUM. Estes programas foram aplicados neste trabalho tanto a realização de experimentos computacionais quanto para a aplicação da metodologia. As principais referências bibliográficas utilizadas para o desenvolvimento destes três programas foram: Martinez (2002), Chapmann (2003), PAO(2003), Lysheuski (2003), Karris (2007) e Collier (2008). A versão do Matlab utilizada para o desenvolvimento destes programas foi a R2008b.

O acesso a cada um desses programas é o obtido mediante a digitação em ambiente Matlab do comando de entrada indicado para cada programa.

- Programa Computacional 1 (comando de entrada : **calculodearl**)
- Programa Computacional 2 (comando de entrada : **metodosecante**)
- Programa Computacional 3 (comando de entrada : **analiseestatistica**)

K.2 Tutorial do Programa Computacional 1

O objetivo deste programa desenvolvido em ambiente Matlab é proporcionar ao usuário uma ferramenta sensível e eficaz para determinar os principais parâmetros que otimizam o gráfico MCUSUM. Neste programa o MEI com Quadratura Gaussiana é aplicado como critério preliminar para a aproximação sistemática dos valores ótimos de ARL utilizados para projetar estatisticamente um gráfico de controle MCUSUM.

A utilização deste programa exige do usuário apenas a entrada correta de dados em cada uma das oito etapas (opções) descritas na caixa de diálogo conforme figura K.1.

1ª) Digite o limite inferior da integral: **0** (o limite inferior do gráfico de controle MCUSUM unilateral é zero).

2ª) Digite o número de pontos de quadratura: **53** (escolha um número de pontos entre 1 e 53. É importante lembrar que quanto maior o número de pontos escolhidos maior será a aproximação).

3ª) Escolha o número de variáveis: **2** (escolha apenas uma das opções para o número de variáveis: 2, 3 ou 4).

4ª) Escolha o valor de ARL_o : **200** (escolha apenas uma das opções para o valor de ARL_o : 200, 500 ou 1000).

5ª) Escolha o tamanho (amplitude) de mudança d ou o valor de referência k : **k** (se a opção desejada for o tamanho de mudança, digite d . Caso contrário, permanece o valor de $k = \frac{d}{2}$, Crosier, 1988).

6ª) Digite o primeiro valor de d ou de k : **0** (escolha um valor para d ou k conforme sua opção na etapa anterior)

7ª) Digite o último valor de d ou de k : **1.0** (escolha um valor para d ou k conforme sua opção na 5ª etapa. Se desejar obter apenas um único valor para o parâmetro desejado neste programa, digite 1. No entanto, se a opção for obter um conjunto de valores deste parâmetro, digite a variação desejada entre os valores de d ou de k na 8ª etapa. Por exemplo, se a opção for determinar os valores do parâmetro entre $d=1$ e $d=2$ com variação de 0,25, ou seja, $d=1$, $d=1,25$, $d=1,5$ e $d=2$, o programa determinará um conjunto de valores para o parâmetro desejado.

8ª) Digite a variação desejada entre os valores de d ou de k : **1** (escolha esta variação desejada)

O comando de entrada para acesso ao programa ao Programa Computacional 1 em ambiente Matlab é: (digite: **calculodearl** e o programa responderá como saída a caixa de diálogo conforme figura K.1)

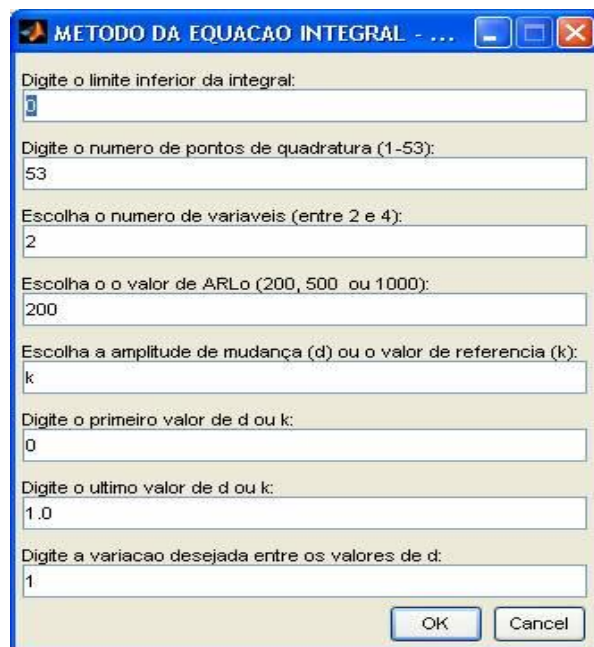


Figura K.1 Caixa de diálogo para acesso ao Programa Computacional 1

Os exemplos 1 e 2 ilustram didaticamente as instruções para a alimentação da caixa de diálogo bem como a saída de dados do Programa Computacional 1 (resultados obtidos após o cálculo).

Exemplo 1:

Cálculo do valor de ARL_{\min} para $d = 1$ com 53 pontos de quadratura escolhendo $ARL_0 = 200$ para $p = 2$.

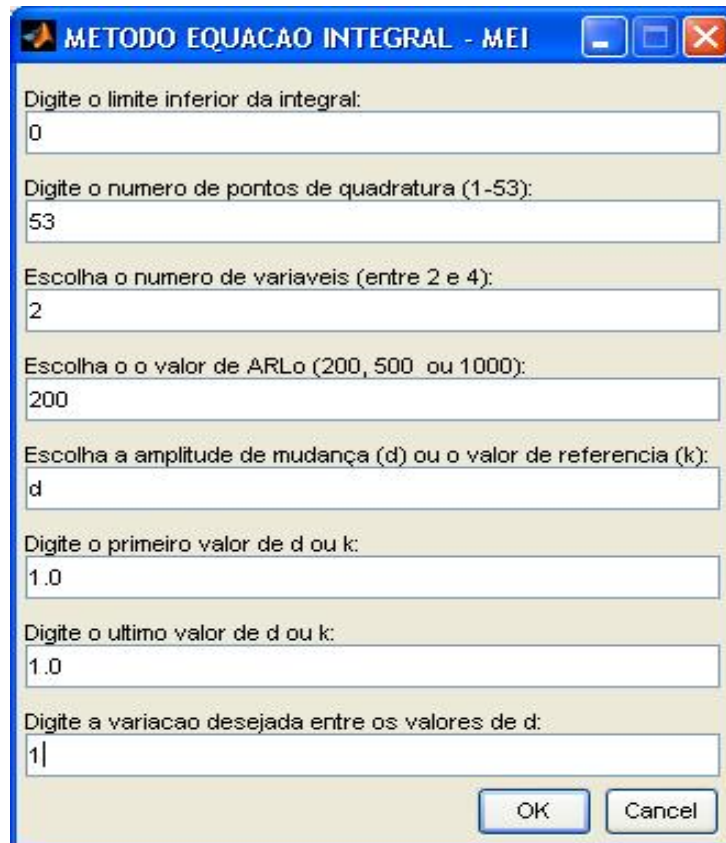


Figura K.2 Caixa de diálogo com os dados de entrada do exemplo 1

Resultados (de saída) gerados após a execução do programa para os dados do exemplo 1:

```
>> calculodearl
```

```
***** METODO DA EQUACAO INTEGRAL PARA O CALCULO DE ARL*****
```

```
QUADRATURA GAUSSIANA
```

```
GRAFICO DE CONTROLE MULTIVARIADO DE SOMA ACUMULADA (MCUSUM)
```

```
UNIVERSIDADE FEDERAL DE SANTA CATARINA
```

```
Custodio da Cunha Alves custodio.alves@univille.br
```

***** RESULTADOS PARA 2 variaveis e $ARL_o = 200$ ***

k	h	integral	ARL
0.50	5.492999999999995	12.282873467360808	9.840907746475184

FIM!! Se deseja parar tecle q:

Exemplo 2:

Cálculo para os valores de ARL_{min} entre $k=0,5$ e $k=1,25$ com variação de 0,25 usando 53 pontos de quadratura e escolhendo $ARL_o=200$ para $p=2$.

Figura K.3 Caixa de diálogo com os dados de entrada do exemplo 2

Resultados (de saída) gerados após a execução do programa para os dados do exemplo 2:

>> calculodearl

***** METODO DA EQUACAO INTEGRAL PARA O CALCULO DE ARL*****

QUADRATURA GAUSSIANA

GRAFICO DE CONTROLE MULTIVARIADO DE SOMA ACUMULADA (MCUSUM)

UNIVERSIDADE FEDERAL DE SANTA CATARINA

Custodio da Cunha Alves custodio.alves@univille.br

***** RESULTADOS PARA 2 variáveis e $ARL_0 = 200$ *****

k	h	integral	ARL
0.50	5.492999999999995	12.282873467360808	9.840907746475184
0.75	3.9459999999999973	7.828760121979740	5.386794401094115
1.00	3.0099999999999963	5.804001481838699	3.362035760953074
1.25	2.8919999999999964	5.571131871892412	3.129166151006787

FIM!! Se deseja parar tecle q:

É importante ressaltar que ao final de cada operação existe neste programa a opção de continuar realizando cálculos de ARL para outras situações. No entanto, se esta não for sua opção, isto é, se desejar parar digite: q e o programa será encerrado. Uma vez encerrado o programa, neste mesmo ambiente poderá ser efetuado o comando de entrada para acesso a um dos outros programas propostos neste trabalho. Para isso, basta apenas digitar o comando de entrada do novo programa.

K.3 Tutorial do Programa Computacional 2

O objetivo deste programa desenvolvido em ambiente Matlab é proporcionar ao usuário uma ferramenta sensível e eficaz para determinar o limite superior de controle h (intervalo de decisão) adequado do gráfico MCUSUM para o valor desejado de ARL quando o processo está sob controle. Neste programa, um método iterativo de análise numérica denominado método da Secante é aplicado para aproximar a solução de h à equação não linear $ARL(0|h) - ARL_0 = 0$. A saída dos dados deste programa inclui as sucessivas aproximações para o valor de h em um determinado número de iterações.

A utilização deste programa exige do usuário apenas a entrada correta de dados em cada uma das seis etapas (opções) descritas na caixa de diálogo conforme figura K.4.

1ª) Escolha o número de variáveis: **2** (escolha apenas uma das opções para o número de variáveis p : 2, 3 ou 4).

2ª) Escolha o valor de ARL_0 : **200** (escolha apenas uma das opções para o ARL_0 : 200, 500 ou 1000).

3ª) Escolha um valor para o ARL: **0** (escolha um número real positivo)

4ª) Digite o limite inferior do intervalo utilizado na determinação de h : **0** (digite: 0).

5ª) Digite o limite superior do intervalo utilizado na determinação de h: **10** (para maior rapidez na determinação de h, é importante otimizar o número de iterações (minimizar o número de iterações desnecessárias). Para isso, é aconselhável escolher um número inteiro positivo ≤ 30 para $p=2$, ≤ 40 para $p=3$ e ≤ 50 para $p=4$).

6ª) Número máximo de iterações: **1000** (digite um número inteiro positivo ≤ 1000 . Evite um número muito elevado de iterações).

O comando de entrada para acesso ao Programa Computacional 2 em ambiente Matlab é: (digite: **metodosecante** e o programa responderá como saída a caixa de diálogo conforme figura K.4).

O exemplo a seguir ilustra didaticamente as instruções para a alimentação da caixa de diálogo bem como a saída de dados do Programa Computacional 2 (resultados obtidos após o cálculo).

Exemplo:

Determine o valor do limite de controle h do gráfico MCUSUM para o valor de $ARL = 9,84$.

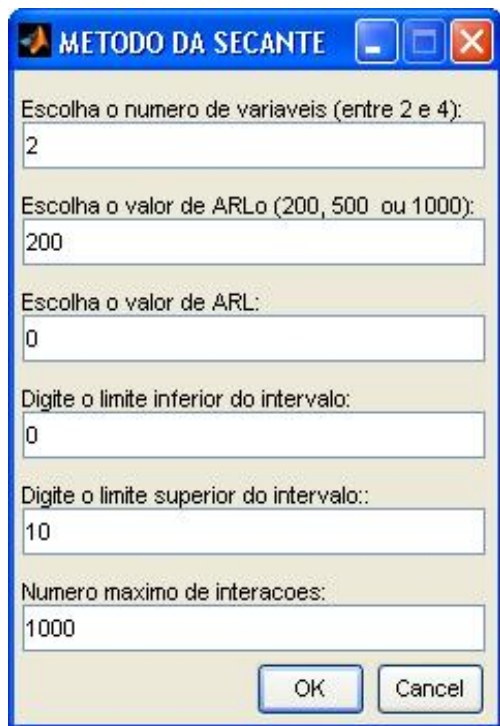


Figura K.4 Caixa de diálogo para acesso ao Programa Computacional 2

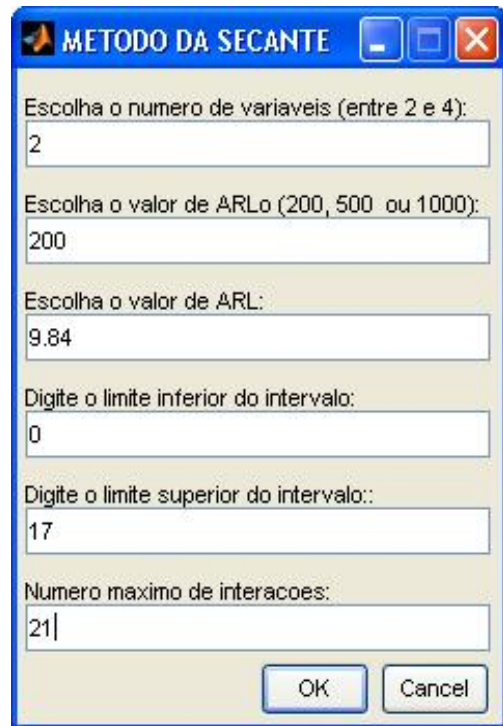


Figura K.5 Caixa de diálogo com os dados de entrada do exemplo

Resultados (de saída) gerados após a execução do programa para os dados de entrada do exemplo digitados na caixa de diálogo (figura K.5):

>> metodossecante

***** METODO DA SECANTE:*****

Intervalo de Decisao h (Limite Superior de Controle)

Grafico de Controle MCUSUM

UNIVERSIDADE FEDERAL DE SANTA CATARINA

Custodio da Cunha Alves custodio.alves@univille.br

Iteracoes	ARL(h)	h
1	12.281966	17.000000000000000
2	163.983848	1.184537221906142
3	9.877918	2.083089286257319
4	8.230409	6.571956081941941
5	4.344818	5.021025010897645
6	1.546811	5.428212491193808
7	0.222992	5.496801485317234
8	0.014164	5.492705149346553
9	0.000120	5.492739521955622

O valor da Raiz (h) e: 5.4927, encontrada com: 9 iteracoes.

FIM!! Se deseja parar tecle q:

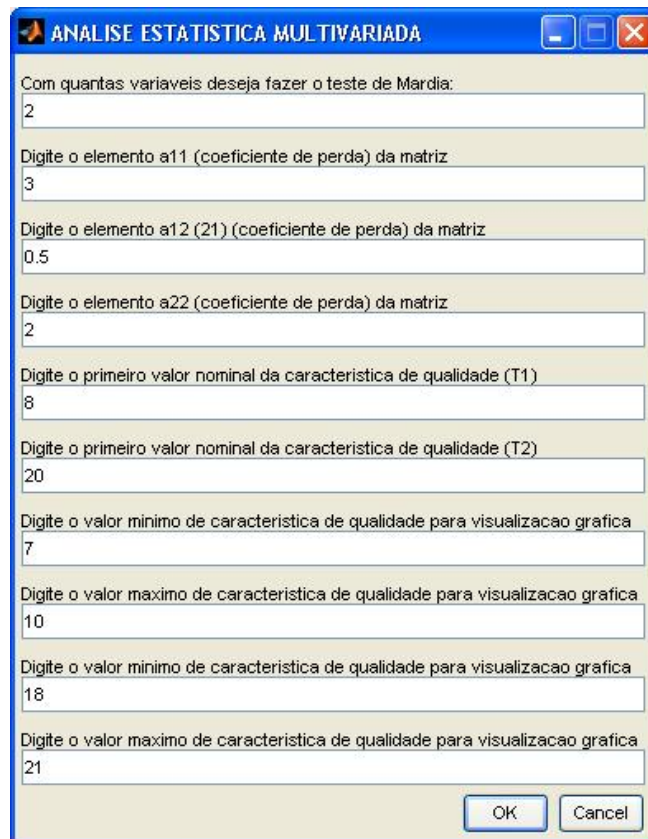
K.4 Tutorial do Programa Computacional 3

O objetivo deste programa desenvolvido em ambiente Matlab é proporcionar ao usuário uma ferramenta sensível e eficaz para realizar a análise estatística multivariada dos dados para projetar o gráfico MCUSUM baseado nos principais parâmetros e sob a ótica de análise de regiões de máxima e mínima potência. Este programa além de determinar a matriz de covariância e sua inversa, a matriz de correlação, a distância de Mahalanobis, d e os coeficientes de assimetria e curtose que fundamentam os testes hipóteses sugeridos por Mardia (1970, 1974) para verificar a normalidade multivariada. Além disso, determina e gera graficamente a função perda multivariada de Taguchi que é utilizada para determinar os pontos A e B que delimitam as regiões de máxima e mínima potência.

A utilização deste programa exige do usuário apenas a entrada correta de dados em cada uma das dez etapas (opções) descritas na caixa de diálogo conforme figura K.6.

- 1ª) Digite o número de variáveis desejadas para o teste de Mardia: (escolha apenas uma das opções para o número de variáveis: 2, 3 ou 4).
- 2ª) Digite o elemento a_{11} (coeficiente de perda) da matriz de coeficientes de perdas: (digite o elemento da matriz desejada)
- 3ª) Digite o elemento a_{12} (coeficiente de perda) da matriz de coeficientes de perdas.
- 4ª) Digite o elemento a_{22} (coeficiente de perda) da matriz de coeficientes de perdas.
- 5ª) Digite o primeiro valor nominal da característica de qualidade: (digite o valor nominal da 1ª característica de qualidade)
- 6ª) Digite o segundo valor nominal da característica de qualidade: (digite o valor nominal da 2ª característica de qualidade)
- 7ª) Digite um valor mínimo para a 1ª característica de qualidade que facilite a visualização gráfica da função perda multivariada de Taguchi.
- 8ª) Digite o valor máximo da 1ª característica de qualidade que facilite a visualização gráfica da função perda multivariada de Taguchi.
- 9ª) Digite o valor mínimo da 2ª característica de qualidade que facilite a visualização gráfica da função perda multivariada de Taguchi.
- 10ª) Digite o valor máximo da 2ª característica de qualidade que facilite a visualização gráfica da função perda multivariada de Taguchi.

O comando de entrada para acesso ao Programa Computacional 3 em ambiente Matlab é: (digite: **analiseestatistica** e o programa responderá como saída uma caixa de diálogo conforme figura K.6).



A caixa de diálogo intitulada "ANALISE ESTADISTICA MULTIVARIADA" apresenta os seguintes campos de entrada:

- Com quantas variaveis deseja fazer o teste de Mardia: 2
- Digite o elemento a11 (coeficiente de perda) da matriz: 3
- Digite o elemento a12 (21) (coeficiente de perda) da matriz: 0.5
- Digite o elemento a22 (coeficiente de perda) da matriz: 2
- Digite o primeiro valor nominal da caracteristica de qualidade (T1): 8
- Digite o primeiro valor nominal da caracteristica de qualidade (T2): 20
- Digite o valor minimo de caracteristica de qualidade para visualizacao grafica: 7
- Digite o valor maximo de caracteristica de qualidade para visualizacao grafica: 10
- Digite o valor minimo de caracteristica de qualidade para visualizacao grafica: 18
- Digite o valor maximo de caracteristica de qualidade para visualizacao grafica: 21

Botões "OK" e "Cancel" estão localizados na base direita da caixa.

Figura K.6 Caixa de diálogo para acesso ao Programa Computacional 3

Para ilustrar didaticamente tanto as instruções de alimentação da caixa de diálogo quanto da saída de dados (resultados obtidos) após a execução do Programa Computacional 3 são utilizados os dados do processo XY. Os resultados obtidos com os dados deste processo após a execução deste programa são utilizados tanto para realizar a análise estatística multivariada quanto para formular um modelo matemático em ambiente MS-Excel onde com a utilização da ferramenta Solver é possível determinar os pontos que delimitam as regiões de máxima e mínima potência do gráfico MCUSUM

Uma vez, transferido os dados (valores das características da qualidade X_1 e Y_1) do processo XY da planilha do MS-Excel para o ambiente Matlab e realizado a operação salvar arquivo; inicia-se a operação de entrada destes dados em ambiente Matlab na figura K.7.

Dados de entrada do Processo: XY (Processo furação de acabamento do furo1)

Características da qualidade: X_1 (posicional do furo1 em relação a coordenada X)

Valor nominal : $5 \pm 0,08$ mm

Características da qualidade : Y_1 (posicional do furo1 em relação a coordenada Y)

Valor nominal : $103,25 \pm 0,08$ mm.

Para este processo supõe-se que a política de recursos financeiros da empresa considere um custo ou perda depreciable de 0,0012 u.m e um custo ou perda inadmissível de 0,0042 u.m.

(unidades monetárias) e que a matriz coeficientes de perda seja $K = \begin{pmatrix} 8 & 2 \\ 2 & 6 \end{pmatrix}$.

Para este processo supõe-se que a política de recursos financeiros da empresa considere um custo ou perda depreciable de 0,0012 u.m e um custo ou perda inadmissível de 0,0042 u.m.

(unidades monetárias) e que a matriz coeficientes de perda seja $K = \begin{pmatrix} 8 & 2 \\ 2 & 6 \end{pmatrix}$.

Os dados de entrada do processo na caixa de diálogo para execução do Programa Computacional 3 conforme figura K.7.

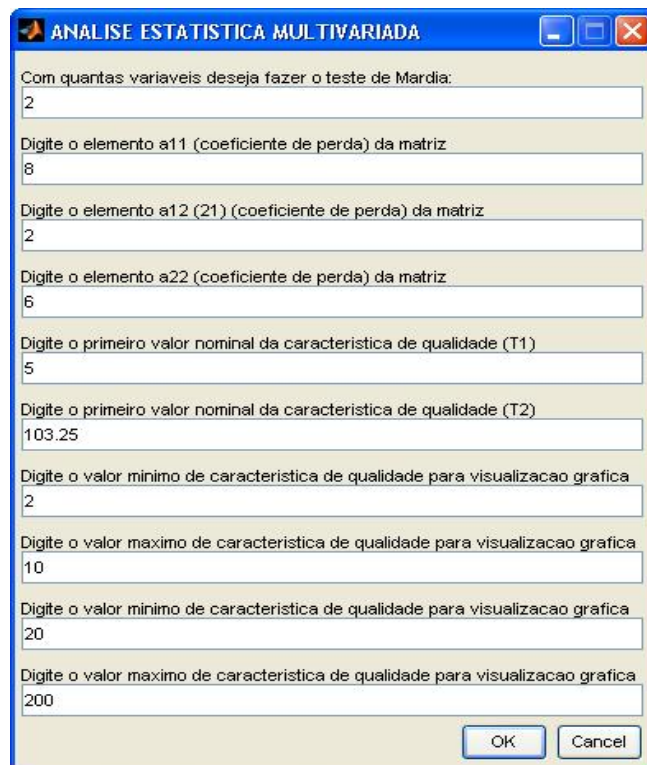


Figura K.7 Caixa de diálogo com os dados de entrada do processo

Resultados (de saída) gerados após a execução do programa para os dados do processo:

```
>> analiseestatistica
```

```
***** ANALISE ESTATISTICA DE DADOS MULTIVARIADOS:*****
```

```
UNIVERSIDADE FEDERAL DE SANTA CATARINA
```

```
Custodio da Cunha Alves custodio.alves@univille.br
```

**** RESULTADOS PARA 2 VARIÁVEIS

A matriz de Covariância mc e:

```
mc = 1.0e-003 *
    0.300931182795699 -0.128500000000017
   -0.128500000000017  0.196600000000008
```

A matriz de Inversa de Covariância imc e:

```
imc = 1.0e+003 *
    4.609524448106236  3.012837698787922
    3.012837698787922  7.055695037101938
```

A matriz de Correlação mcorre e: mcorre =

```
1.000000000000001 -0.528296550359601
-0.528296550359601  1.000000000000000
```

O Valor de beta(1,2) e: 23.7966

O Valor do coeficiente de assimetria, k(1): 122.9492

O Valor de beta(2,2) e: 6.9726

O Valor do coeficiente de curtose, k(2): -3.9813

O valor da Função perda de Taguchi (Multivariada):

$$R = 8*y_1^2 - 493*y_1 + 529827/8 + 6*y_2^2 - 1259*y_2 + 4*y_1*y_2$$

A Distância de (d) Mahalanobis:

$$d_{y_1, y_2} = \frac{1}{2097152} * (20272902916842224*y_1^2 - 2938977997246686396*y_1 + 344998913380151109561 + 26501200659353648*y_1*y_2 - 6540464278683474936*y_2 + 31031274941339984*y_2^2)^{(1/2)}$$

FIM!! Se deseja parar tecla q:

O gráfico da função perda multivariada de Taguchi conforme figura K.8.

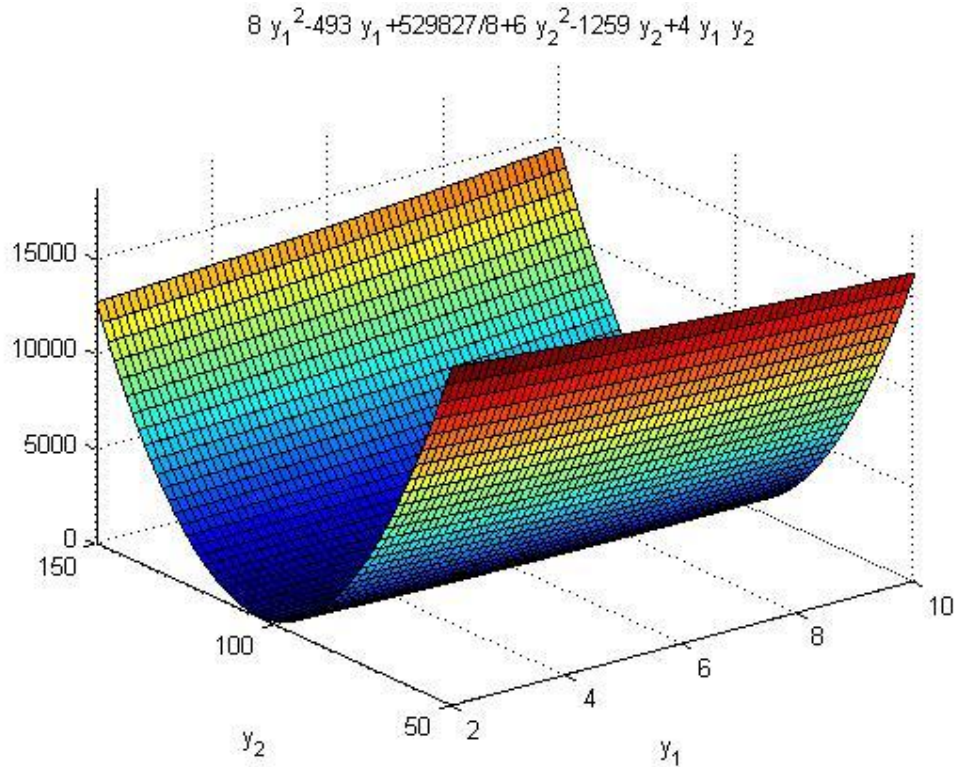


Figura K.8 Função perda para as características da qualidade X_1 e Y_1 do processo

K.5 Teste de Mardia: Programa Computacional

Como recurso computacional em ambiente Matlab para verificação de normalidade multivariada além do Programa Computacional 3 desenvolveu-se um outro programa denominado Teste de Mardia que determina os coeficientes de assimetria e curtose aplicados nos testes de hipóteses de normalidade multivariada de Mardia. Além disso, este programa determina também a matriz de covariância, sua inversa e a matriz de correlação. Os dados de entrada deste programa em ambiente Matlab são transferidos de uma planilha em MS-Excel

O comando de entrada para acesso ao Programa Computacional Teste de Mardia em ambiente Matlab é: (digite: **testemardia**) e o programa responderá como saída uma caixa de diálogo onde informa-se o número desejado de variáveis conforme figura K.9.

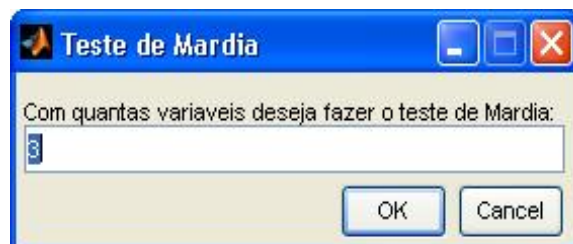


Figura K.9 Caixa de diálogo com a escolha do número de variáveis

A utilização deste programa computacional exige do usuário apenas a entrada correta de dados que consiste em transferir os dados (valores das características da qualidade) de uma planilha do MS-Excel para o ambiente Matlab e após realizado a operação salvar arquivo; informar na caixa de diálogo o número de variáveis (características da qualidade) conforme a figura K.9.

Resultados (de saída) gerados após a execução do programa para os dados do processo XYD:

```
>> testemardia
```

```
***** ANALISE DESCRITIVA DE DADOS MULTIVARIADOS:*****
```

```
UNIVERSIDADE FEDERAL DE SANTA CATARINA
```

```
Custodio da Cunha Alves custodio.alves@univille.br
```

```
**** RESULTADOS PARA 3 VARIAVEIS ****
```

A matriz de Covariância mc e:

```
mc = 1.0e-003 *
```

```
0.300931182795699 -0.128500000000017 -0.020347311827910  
-0.128500000000017 0.196600000000008 0.028766666666629  
-0.020347311827910 0.028766666666629 0.008292473118270
```

A matriz de Inversa de Covariância imc e:

```
imc = 1.0e+005 *
```

```
0.046219804785464 0.027650759844724 0.017489184106398  
0.027650759844724 0.119838996826610 -0.347875694185799  
0.017489184106398 -0.347875694185799 2.455610253813275
```

A matriz de Correlação mcorre e:

```
mcorre =
```

```
1.000000000000001 -0.528296550359601 -0.407315942237409  
-0.528296550359601 1.000000000000000 0.712452226575601  
-0.407315942237409 0.712452226575601 1.000000000000000
```

O Valor de beta (1,3) e: 65.6157

O Valor do coeficiente de assimetria, k(1): 339.0146

O Valor de beta (2,3) e: 13.6953

O Valor do coeficiente de curtose, k(2): -3.6922

FIM!! Se deseja parar tecle q:

APÊNDICE L - Código Computacional dos Programas em Ambiente Matlab

L.1 Programa Computacional 1

```
function calculodeARL
% Este programa calcula os fatores (raizes e pesos) de quadratura
% e aplica o Metodo da Equacao Integral e Quadratura Gaussiana para
% aproximar
% as propriedades estatisticas (ARL, k, h) do grafico de controle MCUSUM

format long ;
% INFORMAÇÕES
disp(sprintf('\n***** METODO DA EQUACAO INTEGRAL PARA O CALCULO DE
ARL:***** '))
disp(sprintf('\n QUADRATURA GAUSSIANA '))
disp(sprintf('\n GRAFICO DE CONTROLE MULTIVARIADO DE SOMA ACUMULADA
(MCUSUM) '))
disp(sprintf('\n UNVERSIDADE FEDERAL DE SANTA CATARINA'))
disp(sprintf('\n Custodio da Cunha Alves   custodio.alves@univille.br\n'))

%***** QUADRATURA GAUSSIANA *****
%Adaptacao dos Autores
%http://numericalmethods.eng.usf.edu/contact.html

    x_values = zeros(n,n) ;
    c_values = zeros(n,n) ;
    i=1 ;
    x_values(i,1) = 0.0 ;
    c_values(i,1) = 2.0 ;

    for i=2:n
        for ii=1:i
            x_values(i,ii) = x(ii) ;
            c_values(i,ii) = A(ii) ;
        end
    end

%Se deseja mostrar os pontos de quadratura gauss e pesos

        %if (((2*i)==nn1) && ((cont == 1) | (cont==2 && valor(1) ==
0)))
        %   disp(sprintf('RAIZES E PESOS DA QUADRATURA '))
        %   disp(sprintf('\n \n RAIZES DA QUADRATURA '))
        %   for ci=1:n
        %       disp(sprintf('
x%i=%1.16f',ci,x_values(n,ci)))
        %   end
        %   disp(sprintf('\n \n PESOS DA QUADRATURA '))
        %   for cii=1:n
        %       disp(sprintf('
c%i=%1.16f',cii,c_values(n,cii)))
        %   end
```



```

%*****
    for i=1:n
        xv(i)=(b-a)/2*x_values(n,i)+(b+a)/2 ;
    end
    approx = 0 ;
    for i=1:n-1
        approx = approx + c_values(n,i)*f(xv(i)) ;
    end
    approx = approx + c_values(n,n)*f(xv(n)) ;

    approx = (b-a)/2 * approx ;

    exact = quad(f,a,b) ;

%***** APRESENTACAO DOS RESULTADOS *****

% Comando necessario para comparar com um valor exato
    exact = quad(f,a,b) ;

    integral = approx;

end %roots

%***** CALCULO DA FUNCAO DO ARL FINAL

    integrall = integral;

    F = 0;
    if (ARL0 == 200 && v == 2 )
        F = pa200_2(4);
        ARL = F + integrall;
        disp(sprintf('\n %1.4f %1.15f %1.15f %1.15f
',valor(cont), h, approx, ARL ) )
    end
    if (ARL0 == 500 && v == 2)
        F = pa500_2(4);
        ARL = F + integrall;
        disp(sprintf('\n %1.4f %1.15f %1.15f %1.15f
',valor(cont), h, approx, ARL ) )
    end
    if (ARL0 == 1000 && v == 2)
        F = pa1000_2(4);
        ARL = F + integrall;
        disp(sprintf('\n %1.4f %1.15f %1.15f %1.15f
',valor(cont), h, approx, ARL ) )
    end
    if (ARL0 == 200 && v == 3)
        F = pa200_3(4);
        ARL = F + integrall;
        disp(sprintf('\n %1.4f %1.15f %1.15f %1.15f
',valor(cont), h, approx, ARL ) )
    end
    if (ARL0 == 500 && v == 3)
        F = pa500_3(4);

```

```

        ARL = F + integrall;
        disp(sprintf('\n %1.4f %1.15f %1.15f %1.15f
',valor(cont), h, approx, ARL ) )
    end
    if (ARL0 == 1000 && v == 3)
        F = pa1000_3(4);
        ARL = F + integrall;
        disp(sprintf('\n %1.4f %1.15f %1.15f %1.15f
',valor(cont), h, approx, ARL ) )
    end
    if (ARL0 == 200 && v == 4)
        F = pa200_4(4);
        ARL = F + integrall;
        disp(sprintf('\n %1.4f %1.15f %1.15f %1.15f
',valor(cont), h, approx, ARL ) )
    end
    if (ARL0 == 500 && v == 4)
        F = pa500_4(4);
        ARL = F + integrall;
        disp(sprintf('\n %1.4f %1.15f %1.15f %1.15f
',valor(cont), h, approx, ARL ) )
    end
    if (ARL0 == 1000 && v == 4)
        F = pa1000_4(4);
        ARL = F + integrall;
        disp(sprintf('\n %1.4f %1.15f %1.15f %1.15f
',valor(cont), h, approx, ARL ) )
    end

end % ARL PARA K E D DIFERENTE DE 0

end %contador %for para os varios d e k
%***** PARAR *****
parar = input('\n \n \n FIM!! Se deseja parar tecle q: ','s');
if parar == 'q'
    break;
end %parar
end %while (1)
end

function [p,dp] = legendre(t,n)
p0=1.0;
p1=t;
for k=1:n-1
    p=((2*k+1)*t*p1-k*p0)/(k+1);
    p0=p1;
    p1=p;
end
dp=n*(p0-t*p1)/(1-t^2);
end

```

L.2 Programa Computacional 2

```

function metodosecante
% Este programa usa um metodo interativo (Metodo Secante)
%para aproximar o intervalo de decisao h (limite superior)
%do grafico de controle MCUSUM.
%Adaptado de:
%KIUSALAAS, Jann. Numerical Methods in Engineering with
%MATLAB. Cambridge University Press, 2005.

disp(sprintf('\n\n***** METODO DA SECANTE:***** '))
disp(sprintf('\n Intervalo de Decisao h (Limite Superior de Controle) '))
disp(sprintf('\n Grafico de Controle MCUSUM '))
disp(sprintf('\n UNVERSIDADE FEDERAL DE SANTA CATARINA'))
disp(sprintf('\n Custodio da Cunha Alves   custodio.alves@univille.br \n'))

format long

while(1)
prompt={'Escolha o numero de variaveis (entre 2 e 4):',...
      'Escolha o valor de ARLo (200, 500 ou 1000):',...
      'Escolha o valor de ARL:',...
      'Digite o limite inferior do intervalo:',...
      'Digite o limite superior do intervalo:',...
      'Numero maximo de interacoes:'};

%NOME DA CAIXA DE DIALOGO
name='METODO DA SECANTE';

%numero de linhas visiveis para entrada
numlines=1;

%resposta padrao
defaultanswer={'2','200','0','0','10','1000'};

% cria caixa de dialogo
answer=inputdlg(prompt,name,numlines,defaultanswer);

v = str2num ( answer{1} );
ARL0 = str2num ( answer{2} );
ARL = str2num ( answer{3} );
x0 = str2num ( answer{4} );
x1 = str2num ( answer{5} );
maxintera = str2num ( answer{6} );
%tolerancia = str2num ( answer{7} );
tolerancia = 1e-8;

%*****

disp(sprintf('\n          interacoes          ARL(h)          h ' ) )

itera = 0;
if ( (v ~= 2) & (v ~= 3) & (v ~= 4) );
    helpdlg('Voce escolheu o numero de variaveis diferente de 2, 3 e 4. O
programa escolheu o numero de variaveis igual a 2.');
```

```

        v=2;
    end
    if ARL == 0;
        helpdlg('Escolha um valor de ARL diferente de zero. Reinicie o programa
e desconsidere o resultado');
        break;
    end

    if ARL ~= 0;
    while (itera<maxintera) & (abs(funcao(x1,ARL0,ARL,v))>tolerancia)

        if ( (ARL0 ~= 200) & (ARL0 ~= 500) & (ARL0 ~= 1000) );
            helpdlg('Voce escolheu um valor de ARLo diferente de 200, 500 ou
1000. O programa escolheu ARLo = 200. ');
            ARL0=200;
            break;
        end

        itera = itera + 1 ;

        f0 = funcao(x0,ARL0,ARL,v);

        f1 = funcao(x1,ARL0,ARL,v);

        x2 = x0-f0*(x1-x0)/(f1-f0);

        x0 = x1;

        x1 = x2;
        disp(sprintf('\n                %1.0f                % .6f                %.15f ',
itera, abs(f0), x0 ) )

    end

    if itera==maxintera

        disp(' Nenhuma raiz encontrada')

    else

        disp(sprintf('\n\n\n'))
        disp(['O valor da Raiz (h) e: ' num2str(x1,10) ', encontrada com: '
num2str(itera) ' iteracoes.'])

    end

    %*****          PARAR          *****
    parar = input('\n \n \n FIM!! Se deseja parar tecle q: ','s');
    if parar == 'q'
        break;
    end %parar
end
end %while

```

L.3 Programa Computacional 3

```
function analiseestatistica
% Este programa calcula a matriz de covariancia e sua inversa,
%a matriz decorrelacao e os coeficientes de assimetria e curtose fornecidos
%por Mardia (1974) para testar a normalidade para tres variaveis

disp(sprintf('\n\n***** ANALISE ESTATISTICA DE DADOS
MULTIVARIADOS:***** '))
disp(sprintf('\n UNVERSIDADE FEDERAL DE SANTA CATARINA'))
disp(sprintf('\n Custodio da Cunha Alves   custodio.alves@univille.br \n'))

format long

%***** CAIXAS *****
while (1)
%notice this is a cell array!
prompt={'Com quantas variaveis deseja fazer o teste de Mardia:',...
'Digite o elemento a11 (coeficiente de perda) da matriz ',...
'Digite o elemento a12 (21) (coeficiente de perda) da matriz ',...
'Digite o elemento a22 (coeficiente de perda) da matriz ',...
'Digite o primeiro valor nominal da caracteristica de qualidade
(T1)',...
'Digite o primeiro valor nominal da caracteristica de qualidade
(T2)',...
'Digite o valor minimo de caracteristica de qualidade para visualizacao
grafica',...
'Digite o valor maximo de caracteristica de qualidade para visualizacao
grafica',...
'Digite o valor minimo de caracteristica de qualidade para visualizacao
grafica',...
'Digite o valor maximo de caracteristica de qualidade para visualizacao
grafica'};

%NOME DA CAIXA DE DIALOGO
name='ANALISE ESTATISTICA MULTIVARIADA';

%numero de linhas visiveis para entrada
numlines=1;

%resposta padrao
defaultanswer={'2','3','0.5','2','8','20','7','10','18','21'};

% cria caixa de dialogo
answer=inputdlg(prompt,name,numlines,defaultanswer);

p = str2num ( answer{1} );
k11 = str2num ( answer{2} );
k12 = str2num ( answer{3} );
k22 = str2num ( answer{4} );
c11 = str2num ( answer{5} );
c12 = str2num ( answer{6} );
Ly1a = str2num ( answer{7} );
Ly1b = str2num ( answer{8} );
Ly2a = str2num ( answer{9} );
Ly2b = str2num ( answer{10} );
```

```

%***** DADOS REAIS *****

dados = xlsread('dados');

m=dados(:,1);
n=dados(:,2);
if p==3
    o=dados(:,3);
    mediao=mean(o);
end

%m=m'; n=n'; o=o';
N=length(n);
mediam=mean(m);
median=mean(n);

if p==3
    media=[mediam, median, mediao];
    mc=[m,n,o];
end
if p==2
    media=[mediam, median];
    mc=[m,n];
end

%***** Matriz de Covariancia

disp(sprintf('\n**** RESULTADOS PARA %i VARIAVEIS',p))

for i=1:N
    for j=1:p
        d(i,j)= mc(i,j) - media(j);
    end
end

dt=d';
disp(sprintf('\n A matriz de Covariancia mc e: '))
mc= (1/(N-1))*(dt * d)
disp(sprintf('\n A matriz de Inversa de Covariancia imc e: '))
imc= inv(mc)

%***** Matriz de Correlacao
for i=1:N
    for j=1:p
        dd(i,j)= d(i,j) / sqrt(mc(j,j));
    end
end

ddt=dd';
disp(sprintf('\n A matriz de Correlacao mcorre e: '))
mcorre= (ddt * dd)/(N-1)
imc= inv(mc);

%*****Teste da Assimetria ****

for i=1:N
    for j=1:N

```

```

        r(i,j)=d(i,:)*imc*dt(:,j);
    end
end
rN=length(r(1,:));

for i=1:rN
    for j=1:rN
        r3(i,j)=r(i,j)^3;
    end
end

kr=sum(r3);
beta1=sum(kr)/16;
k1=N*beta1/6.0;
grau=p*(p+1)*(p+2)/6;
disp(sprintf('\n    O Valor de beta(1,%i) e: %1.4f ', p, beta1 ));
disp(sprintf('\n    O Valor do coeficiente de assimetria, k(1): %1.4f ', k1
));

%*****Teste da Curtose ****

diagr=diag(r);
rN=length(diagr);

for i=1:rN
    r2(i)=diagr(i)^2;
end

beta2=sum(r2)/rN;
k2=(beta2-p*(p+2))/(sqrt(8*p*(p+2))/N);
disp(sprintf('\n    O Valor de beta(2,%i) e: %1.4f ', p, beta2 ));
disp(sprintf('\n    O Valor do coeficiente de curtose, k(2): %1.4f ', k2 ));

%*****

if p==2
    syms y1 y2
    A=[c11, c12];
    At=A';
    B=[y1, y2];
    Bt=[y1; y2];
    C=B-A;
    Ct=Bt-At;

    K=[k11, k12; k12, k22];
    d2_y1_y2= (k11*((y1-A(1,1))^2)+(k22*(y2-A(1,2))^2)+(2*(k12*(y1-
A(1,1))*(y2-A(1,2)))));

    disp(sprintf('\n    O valor da Funcao perda de
Taguchi (Multivariada):'));
    R=expand(d2_y1_y2)

    disp(sprintf('\n    A Distancia de (d) Mahalanobis:'));
    d_y1_y2=sqrt( C*imc*Ct )

```

```
%re= subs(R, y1, 7.204);
%ret =subs(re, y2, 19.9);

ezsurf(R, [Ly1a, Ly1b], [Ly2a, Ly2b])

end

%*****          PARAR          *****

parar = input('\n \n \n FIM!! Se deseja parar tecle q: ','s');

if parar == 'q'
    break;
end %parar

end %while
%ezplot(re)
%quiver3(R)
```


ANEXO A - Dados do processo furação de acabamento do furo 1 para as características da qualidade X_1 , Y_1 e D_{12} (medidas em mm)

Desvio de cada observação em relação ao valor nominal das características de qualidade

Observação	1	2	3	4	5	6	7	8	9	10
X_1	0,019	0,010	0,007	0,009	0,020	0,018	0,019	0,019	0,019	0,024
Y_1	0,000	0,007	-0,009	-0,017	-0,001	-0,003	0,006	-0,006	-0,007	-0,009
D_{12}	0,004	0,004	0,003	0,001	0,000	0,003	0,004	0,001	0,002	0,000

Observação	11	12	13	14	15	16	17	18	19	20
X_1	0,030	-0,023	0,014	0,014	0,040	0,042	-0,002	0,033	0,022	0,019
Y_1	0,004	0,008	0,016	0,008	-0,004	0,002	0,009	0,020	-0,005	0,015
D_{12}	0,007	0,003	0,005	0,003	0,000	0,007	0,004	0,004	-0,002	0,006

Observação	21	22	23	24	25	26	27	28	29	30
X_1	0,002	0,010	-0,008	-0,004	0,004	0,040	-0,006	-0,032	0,026	-0,007
Y_1	0,031	0,010	0,027	0,019	0,020	-0,005	0,025	0,025	0,002	0,041
D_{12}	0,009	0,001	0,006	0,004	0,005	0,003	0,003	0,010	0,000	0,009

Observação	31
X_1	0,002
Y_1	-0,002
D_{12}	0,005

X_1 (Posicional do furo 1 em relação a coordenada X)
 Y_1 (Posicional do furo 1 em relação a coordenada Y)
 D_{12} (Distância entre os centros dos furos 1 e 2)

Valor da observação individual das características da qualidade

	1	2	3	4	5	6	7	8	9	10
X_1	5,019	5,010		5,009	5,020	5,018	5,019	5,019	5,019	5,024
Y_1	103,250	103,257	103,241	103,233	103,239	103,247	103,256	103,244	103,243	103,241
D_{12}	194,274	194,274	194,273	194,271	194,270	194,273	194,274	194,271	194,272	194,270

Observação	11	12	13	14	15	16	17	18	19	20
X_1	5,030	4,977	5,014	5,014	5,040	5,042	4,998	5,033	5,022	5,019
Y_1	103,254	103,258	103,266	103,258	103,246	103,252	103,259	103,270	103,245	103,265
D_{12}	194,277	194,273	194,275	194,273	194,270	194,277	194,274	194,274	194,260	194,276

Observação	21	22	23	24	25	26	27	28	29	30
X_1	5,002	5,010	4,992	4,996	5,004	5,040	4,994	4,968	5,026	4,993
Y_1	103,281	103,260	103,277	103,269	103,270	103,245	103,275	103,275	103,252	
D_{12}	194,279	194,271	194,276	194,274	194,275	194,273	194,273	194,280	194,270	194,279

Observação	31
X_1	5,002
Y_1	103,248
D_{12}	194,275

X_1 (Posicional do furo 1 em relação a coordenada X)
 Y_1 (Posicional do furo 1 em relação a coordenada Y)
 D_{12} (Distância entre os centros dos furos 1 e 2)

ANEXO B - Parâmetros Ótimos do Gráfico MCUSUM via Método Cadeias de Markov

Tabela 1 O parâmetro ótimo do gráfico, k para projetar o MCUSUM e os correspondentes limites de controle, h com o ARL mínimo (ARL_{\min}) para $p=2$, ARL sob controle (ARL_0) de 200, 500 e 1000, e vários tamanhos (amplitude) de mudança, δ .

δ	k	ARL_0	100	200	370	500	1000
0.1	0.05	H	11.566	15.668	20.124	22.587	28.918
		ARL_{\min}	80.32	136.47	204.83	245.28	351.76
0.2	0.1	H	10.035	13.054	16.088	17.673	21.505
		ARL_{\min}	54.00	79.12	105.51	119.67	154.24
0.3	0.15	H	8.836	11.17	13.408	14.544	17.218
		ARL_{\min}	37.38	51.11	64.50	71.44	87.96
0.4	0.2	H	7.878	9.755	11.506	12.380	14.411
		ARL_{\min}	27.43	35.86	43.91	47.98	57.45
0.5	0.25	H	7.098	8.656	10.083	10.789	12.418
		ARL_{\min}	21.08	26.75	32.02	34.68	40.80
0.6	0.3	H	6.451	7.776	8.975	9.565	10.921
		ARL_{\min}	16.77	20.82	24.56	26.41	30.67
0.7	0.35	H	5.905	7.054	8.085	8.591	9.75
		ARL_{\min}	13.70	16.75	19.51	20.86	24.00
0.8	0.4	H	5.437	6.449	7.353	7.794	8.806
		ARL_{\min}	11.41	13.78	15.92	16.97	19.36
0.9	0.45	H	5.032	5.935	6.738	7.129	8.025
		ARL_{\min}	9.71	11.57	13.27	14.08	16.00
1.0	0.5	H	4.678	5.491	6.213	6.564	7.368
		ARL_{\min}	8.35	9.90	11.25	11.93	13.46
1.1	0.55	H	4.364	5.104	5.759	6.078	6.807
		ARL_{\min}	7.28	8.55	9.69	10.24	11.51
1.2	0.6	H	4.085	4.763	5.362	5.654	6.320
		ARL_{\min}	6.41	7.49	8.44	8.91	9.98
1.3	0.65	H	3.835	4.460	5.012	5.280	5.893
		ARL_{\min}	5.70	6.61	7.43	7.83	8.73
1.4	0.7	H	3.608	4.188	4.700	4.948	5.516
		ARL_{\min}	5.11	5.90	6.60	6.94	7.72
1.5	0.75	H	3.403	3.944	4.420	4.651	5.180
		ARL_{\min}	4.61	5.30	5.90	6.20	6.88
1.6	0.8	H	3.216	3.722	4.167	4.384	4.879
		ARL_{\min}	4.20	4.80	5.33	5.58	6.18
1.7	0.85	H	3.044	3.519	3.938	4.141	4.606
		ARL_{\min}	3.83	4.36	4.84	5.07	5.59
1.8	0.9	H	2.887	3.334	3.729	3.921	4.359
		ARL_{\min}	3.52	3.99	4.41	4.62	5.09
1.9	0.95	H	2.741	3.165	3.537	3.719	4.133
		ARL_{\min}	3.24	3.67	4.05	4.24	4.65
2.0	1.0	H	2.607	3.008	3.361	3.533	3.926
		ARL_{\min}	3.01	3.39	3.73	3.90	4.28

ANEXO C - Parâmetros Ótimos do Gráfico MCUSUM via Método Cadeias de Markov

Tabela 2 O parâmetro ótimo do gráfico, k para projetar o MCUSUM e os correspondentes limites de controle, h com o ARL mínimo (ARL_{\min}) para $p=3$, ARL sob controle (ARL_0) de 200, 500 e 1000, e vários tamanhos (amplitude) de mudança, δ .

δ	k	ARL_0	100	200	370	500	1000
0.1	0.05	H	14.444	19.562	25.134	28.213	36.111
		ARL_{\min}	83.34	144.71	222.49	268.30	393.02
0.2	0.1	H	12.614	16.386	20.155	22.111	26.794
		ARL_{\min}	58.46	87.61	118.50	135.06	175.53
0.3	0.15	H	11.146	14.040	16.784	18.161	21.363
		ARL_{\min}	41.58	57.66	73.32	81.34	100.34
0.4	0.2	H	9.956	12.261	14.378	15.422	17.816
		ARL_{\min}	30.84	40.73	50.02	54.72	65.54
0.5	0.25	H	8.979	10.873	12.578	13.411	15.309
		ARL_{\min}	23.83	30.45	36.54	39.53	46.42
0.6	0.3	H	8.165	9.761	11.180	11.869	13.434
		ARL_{\min}	19.00	23.71	27.96	30.03	34.86
0.7	0.35	H	7.477	8.849	10.060	10.646	11.974
		ARL_{\min}	15.53	19.03	22.18	23.72	27.21
0.8	0.4	H	6.887	8.087	9.140	9.649	10.800
		ARL_{\min}	12.94	15.66	18.07	19.22	21.89
0.9	0.45	H	6.377	7.440	8.370	8.818	9.833
		ARL_{\min}	11.00	13.13	15.03	15.97	18.05
1.0	0.5	H	5.930	6.883	7.714	8.115	9.022
		ARL_{\min}	9.46	11.20	12.73	13.47	15.16
1.1	0.55	H	5.534	6.397	7.148	7.510	8.329
		ARL_{\min}	8.24	9.67	10.94	11.55	12.94
1.2	0.6	H	5.183	5.970	6.654	6.984	7.730
		ARL_{\min}	7.26	8.46	9.51	10.03	11.19
1.3	0.65	H	4.867	5.590	6.218	6.521	7.206
		ARL_{\min}	6.45	7.47	8.37	8.80	9.79
1.4	0.7	H	4.582	5.250	5.830	6.110	6.743
		ARL_{\min}	5.78	6.66	7.42	7.80	8.64
1.5	0.75	H	4.324	4.944	5.483	5.743	6.331
		ARL_{\min}	5.21	5.98	6.64	6.96	7.70
1.6	0.8	H	4.088	4.667	5.170	5.412	5.961
		ARL_{\min}	4.73	5.40	5.98	5.27	6.91
1.7	0.85	H	3.871	4.414	4.885	5.113	5.627
		ARL_{\min}	4.31	4.90	5.42	5.67	6.24
1.8	0.9	H	3.672	4.183	4.626	4.840	5.325
		ARL_{\min}	3.95	4.48	4.94	5.16	5.67
1.9	0.95	H	3.489	3.971	4.389	4.591	5.048
		ARL_{\min}	3.82	4.32	4.76	4.97	5.46
2.0	1.0	H	3.319	3.775	4.171	4.362	4.795
		ARL_{\min}	3.37	3.79	4.16	4.34	4.75

Adaptado de: Lee e Khoo (2006) pág. 486

ANEXO D - Parâmetros Ótimos do Gráfico MCUSUM via Método Cadeias de Markov

Tabela 3 O parâmetro ótimo do gráfico, k para projetar o MCUSUM e os correspondentes limites de controle, h com o ARL mínimo (ARL_{\min}) para $p=4$, ARL sob controle (ARL_0) de 200, 500 e 1000, e vários tamanhos (amplitude) de mudança, δ .

δ	k	ARL_0	100	200	370	500	1000
0.1	0.05	H	16.930	22.956	29.541	33.188	42.549
		ARL_{\min}	85.45	150.16	234.32	285.13	426.10
0.2	0.1	H	14.887	19.355	23.818	26.127	31.620
		ARL_{\min}	62.11	94.44	129.47	148.25	194.00
0.3	0.15	H	13.210	16.633	19.858	21.466	25.167
		ARL_{\min}	45.04	63.20	80.99	90.03	111.45
0.4	0.2	H	11.831	14.542	17.008	18.213	20.947
		ARL_{\min}	33.80	45.00	55.51	60.81	72.81
0.5	0.25	H	10.688	12.901	14.870	15.822	17.969
		ARL_{\min}	26.27	33.77	40.62	43.99	51.59
0.6	0.3	H	9.730	11.583	13.208	13.990	15.747
		ARL_{\min}	21.03	26.34	31.09	33.42	38.69
0.7	0.35	H	8.917	10.501	11.878	12.539	14.021
		ARL_{\min}	17.21	21.15	24.65	26.34	30.15
0.8	0.4	H	8.220	9.597	10.788	11.358	12.637
		ARL_{\min}	14.41	17.41	20.06	21.33	24.25
0.9	0.45	H	7.615	8.830	9.876	10.377	11.500
		ARL_{\min}	12.22	14.61	16.69	17.70	19.97
1.0	0.5	H	7.086	8.169	9.101	9.547	10.547
		ARL_{\min}	10.52	12.45	14.14	14.94	16.76
1.1	0.55	H	6.618	7.594	8.434	8.835	9.735
		ARL_{\min}	9.18	10.75	12.13	12.79	14.31
1.2	0.6	H	6.201	7.089	7.851	8.215	9.034
		ARL_{\min}	8.08	9.40	10.54	11.10	12.34
1.3	0.65	H	5.827	6.640	7.338	7.672	8.421
		ARL_{\min}	7.17	8.30	9.27	9.74	10.79
1.4	0.7	H	5.490	6.239	6.882	7.189	7.881
		ARL_{\min}	6.42	7.38	8.22	8.62	9.53
1.5	0.75	H	5.184	5.878	6.474	6.759	7.400
		ARL_{\min}	5.79	6.62	7.34	7.69	8.48
1.6	0.8	H	4.904	5.551	6.106	6.371	6.969
		ARL_{\min}	5.25	5.98	6.61	6.91	7.60
1.7	0.85	H	4.648	5.253	5.772	6.020	6.580
		ARL_{\min}	4.79	5.43	5.99	6.25	6.86
1.8	0.9	H	4.413	4.980	5.468	5.701	6.227
		ARL_{\min}	4.38	4.96	5.45	5.69	6.23
1.9	0.95	H	4.195	4.730	5.189	5.409	5.905
		ARL_{\min}	4.03	4.55	4.99	5.20	5.69
2.0	1.0	H	3.994	4.499	4.933	5.141	5.610
		ARL_{\min}	3.73	4.19	4.59	4.78	5.21

Adaptado de: Lee e Khoo (2006) pág. 487