

Dissertação de Mestrado:

**“Uma Avaliação do Desempenho de Núcleo-Estimadores no Controle de
Processos Multivariados”**

por:

Fernando Augusto Alves Glória

Orientador(a): Prof^a. Dr^a. Sueli Aparecida Mingoti

Co-orientador: Prof. Dr. Gregório Saravia Atuncar

Belo Horizonte, junho de 2006.

Aos meus pais e namorada.

Agradecimentos

Acho que uma página é pouco para que eu possa agradecer a todos que me ajudaram nessa caminhada. Em primeiro lugar agradeço a Deus pela força de vontade, disposição e paciência dada a mim para que eu pudesse chegar até aqui, mesmo quando eu pensei em jogar tudo para o alto. Agradeço a meus pais pela vida que me deram, pelo apoio incondicional nas decisões que eu tomei, pelo carinho, atenção, pelas orações da minha mãe para que Deus me ajudasse e pelo orgulho que eles sentem de mim por tudo que eu faço. Também me orgulho muito de ter vocês como meus pais, vocês são os melhores pais do mundo. Agradeço ao meu irmão pela preocupação que sempre teve comigo principalmente quando se dispôs a me ajudar quando viu que o computador de casa teimava em não funcionar quando eu mais precisava dele, tentando conseguir um computador para que eu pudesse cumprir minhas tarefas e prazos. Não poderia deixar de agradecer e pedir desculpas à pessoa que mais sofreu com todo esse tempo que se transcorreu o meu mestrado, minha namorada Érica. Meu amor, peço desculpas: pelos finais de semana solitários em que você ficava sozinha e eu tinha que virar noites na frente do computador fazendo meus programas funcionar, pelas brigas infundadas que inevitavelmente tivemos por causa do meu stress contínuo e por todos os momentos que poderíamos ter ficado juntos e não ficamos. Agradeço a você por confiar em mim e por continuar comigo até hoje apesar das adversidades. Garanto que conseguirei mostrar que depois da dissertação de mestrado existe vida e que o Fernando de outros tempos ainda vive. Não poderia deixar de agradecer aos professores que fizeram parte da minha formação, desde o jardim de infância até o mestrado, e principalmente gostaria de agradecer a meus orientadores pelo apoio no mestrado e especialmente a Dra. Sueli Aparecida Mingoti, que acompanha meus passos há quase 6 anos me orientando em todos os projetos que desenvolvi em minha vida acadêmica. Sueli, obrigado pela paciência, pela ajuda e pela compreensão. Agradeço à família Werkema que me deu apoio quando precisei, confiou no meu trabalho e torceu a meu favor nessa grande batalha. Agradeço às minhas grandes amigas: Taynãna (Nerd), Juliana (Nerdzinha), Flávia (Pequena 14) e ao meu amigo Fábio (Chupiscão), que sempre tiveram uma palavra de apoio e ajudaram quando eu mais precisava (valeu!!!). Agradeço a minha avó, avô, tios, tias, primas, primos, enfim, obrigado a todos que me apoiaram e que eu não lembrei nessa página. Vocês sempre estarão presentes no meu coração.

Resumo

O objetivo desse trabalho é utilizar as técnicas estatísticas não-paramétricas de núcleo-estimador e função de distribuição empírica no controle de qualidade de processos multivariados com distribuição normal e não-normal, autocorrelacionada e não-autocorrelacionada. Foram abordadas três estatísticas para a avaliação do processo multivariado: a estatística T^2 de Hotelling (1947), a estatística M de Hayter e Tsui (1994) e as componentes principais (Jackson, 1959). Para determinação da janela ótima h na metodologia de núcleo-estimador foram utilizadas as metodologias de Polansky e Baker (2000) e Bessegato (2001).

Foi observado, através de simulações realizadas, que a estatística M de Hayter e Tsui forneceu melhores resultados na análise dos processos simulados mostrando uma baixa taxa de alarmes falsos (altos valores de *ARL in control*) e uma alta taxa de alarmes verdadeiros (baixos valores de *ARL out of control*). Observou-se que a obtenção da constante $C_{R,\alpha}$ proposta por Hayter e Tsui (1994), quando utilizamos as metodologias de núcleo-estimador para sua estimação, foi bastante efetiva mesmo nos casos em que pequenas amostras ($m=50$) eram utilizadas. Além de bastante relevante, essa análise é inédita na literatura de controle de qualidade de processos multivariados uma vez que Hayter e Tsui, em 1994, dizem que é necessário, para processos normais, gerar amostras de tamanho igual a 100.000 para a estimação dessa constante para processos normais e igual a 500 para processos não-normais.

Através das simulações realizadas foi observado que a técnica não-paramétrica que apresentou pior desempenho na estimação dos limites de controle das estatísticas supracitadas foi a função de distribuição empírica. Dentre as metodologias de núcleo-estimador analisadas nessa dissertação, em outra análise inédita na literatura, foi observado que o desempenho do método *plug-in* de multi-estágio estudado por Polansky e Baker (2000) com estágio $b=5$ e o método de *plug-in* proposto por Bessegato (2001) fornecem resultados bem próximos mostrando que essas duas técnicas são efetivas na obtenção dos limites de controle para os processos multivariados.

Palavras chaves: Processo multivariado; estatística T^2 de Hotelling e estatística M de Hayter e Tsui; núcleo-estimador; constante $C_{R,\alpha}$.

Abstract

The purpose of this dissertation is to evaluate the performance of kernel estimators and the empirical distribution function when applied in quality control of multivariate processes with and without autocorrelation with Normal and non-normal distributions. Three estimators were considered: T^2 Hotelling's statistic (1947), Hayter and Tsui's statistic (1994) and principal components analysis (Jackson, 1959). Kernel estimators and the empirical distribution function were used to determine the distribution function of these statistics. Two methods were used to obtain the optimal window (h) for kernel estimators: Polansky e Baker's (2000) and Bessegato's (2001).

These results from Monte Carlo simulation showed that in general the M statistic had a better performance than the other statistics presenting a smaller rates of false alarms (higher ARL in control values) and higher rates of true alarms (lower ARL out of control values) even for small sample sizes when kernel estimators are used for estimate this distribution. The Hayter and Tsui's M statistic had better performance for normal and non-normal process with and without autocorrelation.

The empirical distribution function had bad performance in all cases considered. As far as the method to estimate the window h is considered the plug-in multi-stage from Polansky and Baker's (2000) with $b=5$ and Bessegato's (2001) presented similar results and were very adequate in the estimation of the control limits for multivariate process.

Key words: Multivariate Process; T^2 Hotelling's statistic; Hayter and Tsui M statistic; Kernel estimator; Empirical distribution function; Principal Component Analysis.

Sumário

| | |
|--|----|
| Capítulo 1 – Introdução | 15 |
| 1.1 – Histórico e Revisão Bibliográfica | 16 |
| 1.2 – Objetivos | 20 |
| 1.3 – Resultados relevantes | 21 |
| 1.4 – Organização | 21 |
| Capítulo 2 – Cartas de Controle | 23 |
| 2.1 – Cartas de Controle Univariadas para a Média do Processo | 22 |
| 2.1.1 – Cartas de Controle de Shewhart para Processos Não-Autocorrelacionados | 22 |
| 2.1.2 – O <i>ARL in control</i> e o <i>ARL out of control</i> | 25 |
| 2.2 – Carta de Controle Multivariada para o Vetor de Médias de Processos Não-Autocorrelacionados | 27 |
| 2.2.1 – A Distribuição Normal Multivariada | 28 |
| 2.2.2 – A Estatística T^2 de Hotelling | 31 |
| 2.2.3 – As Componentes Principais | 33 |
| 2.2.4 – A Estatística M de Hayter e Tsui | 35 |
| 2.2.5 – Processos Autocorrelacionados Multivariados | 37 |
| Capítulo 3 – Técnicas Estatísticas Não Paramétricas para a construção de Cartas de Controle | 42 |
| 3.1 –A Função de Distribuição Empírica | 44 |
| 3.2 –A metodologia de Núcleo-Estimador | 45 |
| 3.2.1 –A escolha do Parâmetro de Suavidade | 46 |
| 3.2.1 –A escolha do Parâmetro de Suavidade | 48 |
| 3.2.1.1–Método de Bessegato para Determinação da Janela Ótima | 48 |
| 3.2.1.2–Método de Polansky e Baker para Determinação da Janela Ótima | 49 |
| 3.3 –Exemplo de Aplicação | 51 |

| | |
|---|-----|
| Capítulo 4 – Estratégia de Execução: Processos Multivariados Normais Não-Autocorrelacionados | 55 |
| 4.1 – Modelos Simulados | 56 |
| 4.2 – Determinação dos Limites de Controle e Janela Ótima | 60 |
| 4.3 – Determinação dos <i>ARL</i> | 61 |
| 4.4 – Aspectos Computacionais | 64 |
| Capítulo 5 – Análise de Resultados: Processos Multivariados Normais Não Autocorrelacionados | 66 |
| 5.1 – Comparação dos Métodos para Determinação dos Limites de Controle | 67 |
| 5.1.1 – Processos bivariados ($p=2$) | 67 |
| 5.1.2 – Processos multivariados com $p=5$ variáveis | 72 |
| 5.1.3 – Processos multivariados com $p=10$ variáveis | 76 |
| 5.1.4 – Processos multivariados com $p=20$ variáveis | 80 |
| 5.1.5 – Nota sobre a estatística T^2 de Hotelling | 83 |
| 5.2 – Análise dos <i>ARL in control</i> e <i>out of control</i> | 85 |
| 5.2.1 – Processos Multivariados: $p=2$, $p=5$ e $p=10$ | 85 |
| 5.2.2 – Processos Multivariados com $p=20$ variáveis | 92 |
| 5.2.3 – Conclusão | 97 |
| Capítulo 6 – Estratégia de Execução: Processos Multivariados Não-Normais e Não-Autocorrelacionados | 99 |
| 6.1 – Processo 1: O Processo Bivariado estudado por Hayter e Tsui (1994) | 100 |
| 6.1.1 – Descrição do processo | 100 |
| 6.1.2 – Análise dos resultados | 102 |
| 6.2 – Processo 2: O Processo Lognormal Bivariado | 104 |
| 6.2.1 – Descrição do processo | 104 |
| 6.2.2 – Análise dos resultados | 106 |
| 6.3 – Processo 3: Seno Hiperbólico Multivariado ($p=5$) | 110 |

| | |
|---|------------|
| 6.3.1 – Descrição do processo | 110 |
| 6.3.2 – Análise dos resultados | 111 |
| 6.4 – Conclusão | 117 |
| Capítulo 7 – Estratégia de execução: Processos Normais e Não-Normais Multivariados Autocorrelacionados | 118 |
| 7.1 – Processos Normais: Modelo $VAR(1)$ Multivariado | 119 |
| 7.2 – Análise de Resultados: Processos Normais Multivariados Autocorrelacionados | 121 |
| 7.3 – Processos Não Normais: Modelo $VAR(1)$ Multivariado Não-Normal | 126 |
| 7.4 – Análise de Resultados: Processos Multivariados Não-Normais Autocorrelacionados | 129 |
| 7.5 – Conclusão | 131 |
| Capítulo 8 – Exemplo de Aplicação | 132 |
| Capítulo 9 – Conclusões Finais e Sugestões para Trabalhos Futuros | 138 |
| Referências Bibliográficas | 141 |
| Anexo A | 146 |
| Anexo B | 152 |
| Anexo C | 158 |

Lista de Figuras

Capítulo 2 – Cartas de Controle

| | |
|---|----|
| Figura 2.1: Carta de controle: exemplo de <i>ARL in control</i> e <i>ARL out of control</i> | 26 |
| Figura 2.2: Cartas de controle individuais para um processo bi-variado ($\rho = 0,90$) | 27 |
| Figura 2.3: Distribuição normal multivariada com $p=2$, $\mu_1 = \mu_2 = 0$, $\sigma_{11} = \sigma_{22} = 1$ e $\rho_{12} = 0$ | 29 |
| Figura 2.4: Distribuição normal multivariada com $p=2$, $\mu_1 = \mu_2 = 0$, $\sigma_{11} = \sigma_{22} = 1$ e $\rho_{12} = 0,9$ | 29 |
| Figura 2.5: Comparação da elipse de confiança com os limites de controle univariados com $\alpha = 0,0027$ | 31 |
| Quadro 2.1: Obtenção da constante $C_{R,\alpha}$ no caso da distribuição normal p-variada | 36 |
| Quadro 2.2: Estimativa não-paramétrica de $C_{R,\alpha}$ | 36 |

Capítulo 3 – Técnicas Estatísticas Não Paramétricas para a construção de Cartas de Controle

| | |
|--|----|
| Figura 3.1: Carta de valores individuais para o processo supondo distribuição normal (a) e distribuição exponencial (b) | 43 |
| Figura 3.2: Comparação da função de distribuição empírica (---) para uma amostra de tamanho 50 (a) e 1000 (b) proveniente de uma distribuição teórica $N(10,4)$ (—) | 45 |
| Figura 3.3: Comparação das funções densidade estimadas via núcleo estimador (—) com a função densidade teórica $N(10,4)$ (- - -) | 47 |
| Figura 3.4: Comparação das funções de distribuição estimadas via núcleo estimador (—) com função de distribuição teórica $N(10,4)$ (- - -) | 48 |
| Figura 3.5: Gráfico de dispersão das variáveis X_1 e X_2 do processo e da distribuição Qui-quadrado | 52 |
| Figura 3.6: Comportamento da estatística T^2 de Hotelling e da estatística M | 52 |
| Figura 3.7: Comportamento da primeira componente principal | 54 |

Capítulo 5 – Análise de Resultados: Processos Multivariados Normais Não Autocorrelacionados

| | |
|--|----|
| Figura 5.1: Comparação do Erro e Erro Quadrático dos limites de controle obtidos à 95% de confiança para processos com correlação entre 0,75 e 0,90, $m=50$, 200 e 1000 e $p=2$ | 70 |
| Figura 5.2: Comparação do Erro e Erro Quadrático dos limites de controle obtidos à 95% de confiança para processos com correlação entre 0,25 e 0,50, $m=50$ e 200 e $p=2$ | 71 |
| Figura 5.3: Comparação do Erro e Erro Quadrático dos limites de controle obtidos à 95% de confiança para processos com correlação igual a zero, $m=50$ e 200 e $p=2$ | 72 |
| Figura 5.4: Comparação do Erro e Erro Quadrático dos limites de controle obtidos à 95% de confiança para processos com correlação entre 0,75 e 1, $m=50$ e 200 e $p=5$ | 74 |
| Figura 5.5: Comparação do Erro e Erro Quadrático dos limites de controle obtidos à 95% de confiança para processos com correlação entre 0,50 e 0,75, $m=50$ e 200 e $p=5$ | 75 |
| Figura 5.6: Comparação do Erro e Erro Quadrático dos limites de controle obtidos à 95% de confiança para processos com correlação igual zero, $m=50$ e 200 e $p=5$ | 76 |
| Figura 5.7: Comparação do Erro e Erro Quadrático dos limites de controle obtidos à 95% de confiança para processos com correlação entre 0,75 e 1, $m=50$ e 200 e $p=10$ | 78 |
| Figura 5.8: Comparação do Erro e Erro Quadrático dos limites de controle obtidos à 95% de confiança para processos com correlação entre 0,50 e 0,75, $m=50$ e 200 e $p=10$ | 79 |
| Figura 5.9 Comparação do Erro e Erro Quadrático dos limites de controle obtidos à 95% de confiança para processos com correlação igual a zero, $m=50$ e 200 e $p=10$ | 80 |
| Figura 5.10: Comparação do Erro e Erro Quadrático dos limites de controle obtidos à 95% de confiança para processos com correlação igual a zero, $m=50$ e 200 e $p=10$ | 82 |
| Figura 5.11: Comparação da estatística T^2 de Hotelling com a distribuição F | 84 |

Capítulo 6 – Estratégia de Execução: Processos Multivariados Não-Normais e Não-Autocorrelacionados

| | |
|---|-----|
| Figura 6.1: Curvas de nível amostral da função densidade conjunta de Z_1 e Z_2 | 101 |
| Figura 6.2: Função densidade conjunta amostral de Z_1 e Z_2 | 101 |
| Figura 6.3: Curvas de nível amostral da função de densidade do vetor aleatório Z com distribuição lognormal bivariada | 106 |
| Figura 6.4: Função densidade amostral do vetor aleatório Z com distribuição lognormal bivariada | 106 |
| Figura 6.5: Curvas de nível amostral da função densidade do vetor aleatório Z com distribuição do seno hiperbólico | 111 |

| | |
|---|-----|
| Figura 6.6: Função densidade amostral do vetor aleatório Z com distribuição do seno hiperbólico | 111 |
|---|-----|

Capítulo 7 – Estratégia de execução: Processos Normais e Não-Normais Multivariados Autocorrelacionados

| | |
|---|-----|
| Figura 7.1: Curvas de nível amostral da função densidade do processo normal bivariado autocorrelacionado $VAR(I)$ | 120 |
|---|-----|

| | |
|---|-----|
| Figura 7.2: Função densidade amostral do processo normal bivariado $VAR(I)$ | 121 |
|---|-----|

| | |
|--|-----|
| Figura 7.3: Curvas de nível amostral da função densidade do processo com distribuição não-normal bivariada autocorrelacionada $VAR(I)$ | 128 |
|--|-----|

| | |
|---|-----|
| Figura 7.4: Função densidade amostral do processo com distribuição não-normal bivariada autocorrelacionada $VAR(I)$ | 128 |
|---|-----|

Capítulo 8 – Exemplo de Aplicação

| | |
|---|-----|
| Figura 8.1: Gráfico seqüencial das variáveis consideradas do processo de fundição do gusa | 134 |
|---|-----|

| | |
|--|-----|
| Figura 8.2: Gráfico de normalidade das variáveis estudadas do processo de fundição do gusa | 135 |
|--|-----|

| | |
|--|-----|
| Figura 8.3: Carta de controle para o processo de produção do ferro gusa com limite de controle obtido de acordo o método de Polansky e Baker (2000) com $b=5$ e $m=3600$ | 136 |
|--|-----|

| | |
|--|-----|
| Figura 8.4: Carta de controle para o processo de produção do ferro gusa com limite de controle obtido de acordo o método de Bessegato e $m=3600$ | 136 |
|--|-----|

Lista de Tabelas

Capítulo 3 – Técnicas Estatísticas Não Paramétricas para a construção de Cartas de Controle

Tabela 3.1: Limites superior de controle para a estatística T^2 de Hotelling e estatística M ($\alpha = 0,05$)53

Tabela 3.2: Cálculo dos limites de controle da 1ª componente principal ($\alpha = 0,05$)54

Capítulo 5 – Análise de Resultados: Processos Multivariados Normais Não Autocorrelacionados

Tabela 5.1: Erros médio e Quadrático médio dos limites de controle obtidos à 95% de confiança para processos com $m=50, 200$ e 1000 e $p=2$ variáveis68

Tabela 5.2: Erros médio e Quadrático médio dos limites de controle obtidos à 95% de confiança para processos com $m=50$ e 200 e $p=5$ variáveis73

Tabela 5.3: Erros médio e Quadrático médio dos limites de controle obtidos à 95% de confiança para processos com $m=50$ e 200 e $p=10$ 77

Tabela 5.4: Erros médio e Quadrático médio dos limites de controle obtidos à 95% de confiança para processos com correlação misturada, $m=50, 100, 200$ e 1000 e $p=20$ 81

Tabela 5.5: *ARL in control* e *out of control* de processos com distribuição normal bivariada $-0,75 \leq \rho \leq 0,90$ e $m=50$ 87

Tabela 5.6: *ARL in control* e *out of control* de processos com distribuição normal bivariada $(0,75 \leq \rho \leq 0,90)$ e $m=200$ 88

Tabela 5.7: *ARL in control* e *out of control* de processos com distribuição normal bivariada $(\rho = 0,90)$ e $m=1000$ 89

Tabela 5.8: *ARL in control* e *out of control* de processos com distribuição normal bivariada $(0,25 \leq \rho \leq 0,50)$ e $m=50$ 90

Tabela 5.9: *ARL in control* e *out of control* de processos com distribuição normal bivariada $(0,25 \leq \rho \leq 0,50)$ e $m=200$ 91

Tabela 5.10: *ARL in control* e *out of control* de processos com distribuição normal bivariada $(\rho = 0)$ e $m=50$ e 200 92

Tabela 5.11: *ARL in control* e *out of control* de processos com distribuição normal p-variada ($p=20$) e $m=50$ 94

Tabela 5.12: *ARL in control* e *out of control* de processos com distribuição normal p-variada ($p=20$) e $m=100$ 95

| | |
|--|----|
| Tabela 5.13: <i>ARL in control</i> e <i>out of control</i> de processos com distribuição normal p-variada ($p=20$) e $m=200$ | 96 |
|--|----|

| | |
|---|----|
| Tabela 5.14: <i>ARL in control</i> e <i>out of control</i> de processos com distribuição normal p-variada ($p=20$) e $m=1000$ | 97 |
|---|----|

Capítulo 6 – Estratégia de Execução: Processos Multivariados Não-Normais e Não-Autocorrelacionados

| | |
|--|-----|
| Tabela 6.1: Impacto dos choques implementados nas médias do vetor aleatório X sobre as médias do vetor aleatório Z | 102 |
|--|-----|

| | |
|---|-----|
| Tabela 6.2: Limites de controle, <i>ARL in control</i> e <i>out of control</i> . Processo com distribuição não-normal bivariada estudada por Hayter e Tsui ($m=50$ e 200) | 103 |
|---|-----|

| | |
|---|-----|
| Tabela 6.3: Limites de controle, <i>ARL in control</i> e <i>out of control</i> . Processo com distribuição lognormal bivariada ($m=50$) | 108 |
|---|-----|

| | |
|--|-----|
| Tabela 6.4: Limites de controle, <i>ARL in control</i> e <i>out of control</i> . Processo com distribuição lognormal bivariada ($m=200$) | 109 |
|--|-----|

| | |
|--|-----|
| Tabela 6.5: Limites de controle, <i>ARL in control</i> e <i>out of control</i> . Processo com distribuição do seno hiperbólico de uma distribuição normal multivariada com $p=5$ variáveis, correlação forte ($0,75 < \rho \leq 1$) e $m=50$ | 113 |
|--|-----|

| | |
|---|-----|
| Tabela 6.6: Limites de controle, <i>ARL in control</i> e <i>out of control</i> . Processo com distribuição do seno hiperbólico de uma distribuição normal multivariada com $p=5$ variáveis, correlação forte ($0,75 < \rho \leq 1$) e $m=200$ | 114 |
|---|-----|

| | |
|--|-----|
| Tabela 6.7: Limites de controle, <i>ARL in control</i> e <i>out of control</i> . Processo com distribuição do seno hiperbólico de uma distribuição normal multivariada com $p=5$ variáveis correlação intermediária ($0,50 < \rho \leq 0,75$) e $m=50$ | 115 |
|--|-----|

| | |
|---|-----|
| Tabela 6.8: Limites de controle, <i>ARL in control</i> e <i>out of control</i> . Processo com distribuição do seno hiperbólico de uma distribuição normal multivariada com $p=5$ variáveis correlação intermediária ($0,50 < \rho \leq 0,75$) e $m=200$ | 116 |
|---|-----|

| | |
|--|-----|
| Tabela 6.9: Limites de controle, <i>ARL in control</i> e <i>out of control</i> . Processo com distribuição do seno hiperbólico de uma distribuição normal multivariada com $p=5$ variáveis correlação nula ($\rho = 0$) $m=50$ e $m=200$ | 117 |
|--|-----|

Capítulo 7 – Estratégia de execução: Processos Normais e Não-Normais Multivariados Autocorrelacionados

| | |
|--|-----|
| Tabela 7.1: Limites de controle, <i>ARL in control</i> e <i>out of control</i> . Processo com distribuição normal bivariada autocorrelacionada. Modelo $VAR(1)$ ($m=50$) | 123 |
|--|-----|

Tabela 7.2: Limites de controle, *ARL in control* e *out of control*. Processo com distribuição normal bivariada autocorrelacionada. Modelo *VAR (1)* ($m=100$)124

Tabela 7.3: Limites de controle, *ARL in control* e *out of control*. Processo com distribuição normal bivariada autocorrelacionada. Modelo *VAR (1)* ($m=200$)125

Tabela 7.4: Limites de controle, *ARL in control* e *out of control*. Processo com distribuição não-normal bivariada autocorrelacionada *VAR(1)* ($m=50, 100$ e 200)130

Capítulo 8 – Exemplo de Aplicação

Tabela 8.1: Estatísticas descritivas das variáveis consideradas no processo de fundição do gusa134

Anexo A

Tabela A.1: *ARL in control* e *out of control* de processos com distribuição normal multivariada com forte correlação entre as variáveis ($0,75 < \rho \leq 1$) e $m=50$ 147

Tabela A.2: *ARL in control* e *out of control* de processos com distribuição normal multivariada com forte correlação entre as variáveis ($0,75 < \rho \leq 1$) e $m=200$ 148

Tabela A.3: *ARL in control* e *out of control* de processos com distribuição normal multivariada com correlação intermediária entre as variáveis ($0,50 < \rho \leq 0,75$) e $m=50$ 149

Tabela A.4: *ARL in control* e *out of control* de processos com distribuição normal multivariada com correlação intermediária entre as variáveis ($0,50 < \rho \leq 0,75$) e $m=200$..150

Tabela A.5: *ARL in control* e *out of control* de processos com distribuição normal multivariada com correlação nula entre as variáveis ($\rho = 0$) e $m=50$ e $m=200$ 151

Anexo B

Tabela B.1: *ARL in control* e *out of control* de processos com distribuição normal multivariada com forte correlação entre as variáveis ($0,75 < \rho \leq 1$) e $m=50$ 153

Tabela B.2: *ARL in control* e *out of control* de processos com distribuição normal multivariada com forte correlação entre as variáveis ($0,75 < \rho \leq 1$) e $m=200$ 154

Tabela B.3: *ARL in control* e *out of control* de processos com distribuição normal multivariada com correlação intermediária entre as variáveis ($0,50 < \rho \leq 0,75$) e $m=50$ 155

Tabela B.4: *ARL in control* e *out of control* de processos com distribuição normal multivariada com correlação intermediária entre as variáveis ($0,50 < \rho \leq 0,75$) e $m=200$..156

Tabela B.5: *ARL in control* e *out of control* de processos com distribuição normal multivariada com correlação nula entre as variáveis ($\rho = 0$) e $m=50$ e $m=200$ 157

Capítulo 1 – Introdução

1.1 – Histórico e Revisão Bibliográfica

O estudo de processos de produção teve início na década de 20 quando o Dr. Walter A. Shewhart dos Laboratórios Bell criou as cartas de controle univariadas, conhecidas como cartas de controle de Shewhart, para o monitoramento de processos. Nas décadas seguintes, principalmente após a 2ª Guerra mundial, a utilização desta ferramenta foi bastante intensificada e teve sua popularização evidenciada durante a implantação do Programa da Qualidade Total, no Japão, pelas mãos de Deming e Juran.

O objetivo da carta de controle de Shewhart (1926) era criar limites de controle que facilitassem o monitoramento da qualidade de um processo verificando se este era estável em relação à parâmetros pré-estabelecidos e se existiam causas especiais que atuavam sobre ele. Quando causas especiais eram detectadas nas cartas de controle (pontos fora dos limites de controle) o processo deveria sofrer modificações corretivas para que voltasse à condição de controle estatístico. Na maioria das vezes as cartas de controle de Shewhart tinham como suposição a distribuição normal da variável de estudo no processo.

Um dos problemas dessas cartas e que elas eram apropriadas apenas para a análise de uma variável do processo, e quando neste existia mais de uma variável (processo multivariado), várias cartas de controle deveriam ser utilizadas simultaneamente, e de forma independente, para que fosse feito o monitoramento do processo. Dessa forma, quando as variáveis eram correlacionadas entre si, as cartas de controle não incorporavam essa correlação o que podia gerar conclusões incorretas sobre a estabilidade do processo.

Outro fato que influenciava negativamente a análise de processos multivariados utilizando várias cartas de controle de forma independente era que essa análise promovia um aumento do nível de significância global do processo causando um aumento do número de alarmes falsos detectados pelas cartas de controle. Alguns autores, utilizando a metodologia proposta por Bonferroni (1930), tentaram fazer uma correção nesse nível de significância fazendo com que o processo tivesse um nível de significância global fixo, idéia que logo foi abandonada uma vez que as cartas de controle com os limites corrigidos apresentavam limites de controle maiores do que os obtidos sem a correção do nível de significância, fazendo com que perdessem poder de detecção dos alarmes verdadeiros.

Para resolver o problema da incorporação da correlação na análise do processo multivariado alguns autores desenvolveram técnicas estatísticas que incorporavam a correlação das variáveis do processo às cartas de controle. Um dos primeiros trabalhos desenvolvidos na literatura que aborda essa incorporação foi o trabalho de Harold Hotelling, que em 1947, propôs a criação da estatística T^2 . Essa estatística, que mais tarde ficou

conhecida como estatística T^2 de Hotelling, foi estimada de forma a fornecer valores que medem a distância entre cada observação do processo e o vetor de média amostral (sendo essa distância ponderada pela inversa da matriz de covariâncias do processo). Com base nesses valores, e seguindo a suposição de o que processo fornece informações segundo uma distribuição normal multivariada, Hotelling construiu uma carta de controle para os valores de T^2 tendo como limite de controle os valores teóricos da distribuição Qui-quadrado ou a distribuição F de Fisher-Snedecor, distribuições derivadas por ele e que designavam a distribuição da estatística T^2 .

A estatística proposta por Hotelling, além de incorporar a correlação entre as variáveis, permitia que o nível de significância global da carta de controle permanecesse fixo sem a necessidade de correção do nível de significância o que fazia com que essa carta fornecesse um baixo número de alarmes falsos e detectasse rapidamente as mudanças ocorridas na média do processo. Devido a essas qualidades, essa estatística ganhou destaque na literatura sendo estudada por diversos autores, entre eles: Chou, *et. al* (1999 e 2001), Colengui e Mingoti (2005a e 2005b), Thomsen (2005), Souza e Rigão (2005).

Seguindo a idéia de incorporar a correlação entre as variáveis e buscando novas técnicas para o monitoramento de processos multivariados, Jackson (1959) utilizou a técnica de componentes principais para a análise de processos normais bivariados. A aplicação dessa técnica a um processo multivariado com p variáveis produzia p componentes principais, sendo cada componente não correlacionada entre si, promovendo a construção de p cartas de controle univariadas. A vantagem da construção dessas cartas de controle é que a correlação entre as variáveis do processo multivariado são “*dissolvidas*” dentre a componentes principais sendo essas não correlacionadas entre si. Porém um problema dessa abordagem está relacionado ao nível de significância global da análise. Jackson (1959), preocupado com o nível de significância global da análise, construiu p cartas de controle univariadas utilizando para cada uma delas o nível de significância corrigido através da técnica proposta por Bonferroni (1930), porém, apesar de manter o nível de significância global da análise fixo, os limites de controle das cartas ficaram maiores, fazendo com que as cartas de controle perdessem capacidade de detectar mudanças no processo. Jonhson e Wichern (2002) e Mingoti (2005) mostram em seus livros que, quando aplicada a amostras que seguem a distribuição normal p -variada, as componentes principais seguem individualmente a distribuição normal, ou seja, tem-se p componentes principais normalmente distribuídas o que permite o cálculo dos limites de controle paramétricos para cada carta de controle.

Outros autores que desenvolveram técnicas estatísticas para a análise de processos multivariados foram Hayter e Tsui que em 1994 propuseram o cálculo da chamada estatística M e da constante $C_{R,\alpha}$ com o objetivo de construir uma carta de controle que competissem com a carta da estatística T^2 de Hotelling. A partir dessa carta de controle era possível verificar a estabilidade do processo e, quando este está fora de controle, essa técnica permite a identificação imediata de qual(is) variável(is) do processo é(são) responsável(is) pela falta de controle. Outros autores como Mingoti e Glória (2003 e 2005b), Rocon (2005), Colengui e Mingoti (2005a e 2005b) e Thomsen (2005) estudaram exaustivamente a estatística M e a constante $C_{R,\alpha}$ de Hayter e Tsui (1994) em seus trabalhos.

A maioria dos artigos citados nos parágrafos anteriores abordam a utilização da estatística T^2 de Hotelling, da estatística M e da componente principal para processos que seguem a distribuição normal multivariada, entretanto ao analisar processos reais, não é raro observar processos que seguem uma distribuição não-normal multivariada (Chou *et. al.*, 2001). Para esses processos as técnicas desenvolvidas sob a suposição da normalidade multivariada podem fornecer resultados inválidos uma vez que não se levou em consideração a distribuição de probabilidade das variáveis estudadas.

Com o objetivo de avaliar os processos com distribuição não-normal Chou *et. al* (2001) utilizaram a estatística T^2 de Hotelling (1947) para construir cartas de controle e verificar a estabilidade de processos que tem distribuição exponencial multivariada. Nesse caso, para obter os limites de controle os autores utilizaram a técnica não-paramétrica de núcleo-estimador com o método de *plug-in* de multi-estágio proposto por Polansky (2000) e Polansky e Baker (2000), para a determinação da janela ótima h . Outros métodos para estimar a janela ótima h foram estudados por Silverman (1992), Wand e Jones (1995) e Bessegato (2001) entre outros, sendo esse último utilizado nessa dissertação juntamente com o método estudado por Polansky e Baker (2000) onde será efetuada uma comparação desses métodos.

Uma outra forma de obtenção dos limites de controle para processos não-normais multivariados foi proposta por Hayter e Tsui em 1994 que mostram que a estatística M e a constante $C_{R,\alpha}$ podem ser obtidas apenas através das informações amostrais do processo caso tenha-se disponível uma amostra de no mínimo 500 observações. Neste caso, a distribuição da estatística M é estimada pela função de distribuição empírica e a constante $C_{R,\alpha}$ é determinada a partir dessa estimativa. No entanto, Mingoti e Glória (2005b) mostram que para a distribuição normal multivariada será necessário pelo menos 5000 observações para que a constante $C_{R,\alpha}$ fosse determinada corretamente.

Até agora as técnicas descritas nos parágrafos consideram apenas processos que possuem distribuição normal e não-normal multivariada, porém, em grande parte dos processos industriais, principalmente processos de fluxo contínuo (Mason e Young, 2002), além da correlação entre as variáveis do processo, possuem uma correlação significativa entre os valores de uma variável no tempo t e os valores no tempo $t-1$, nesse caso, dizemos que o processo é autocorrelacionado. A análise de processos autocorrelacionados através de técnicas convencionais pode causar sério impacto na carta de controle causado pelo fato da autocorrelação ter sido desconsiderada (Kalgonda e Kulkarni, 2004).

A partir das estatísticas usadas para o estudo dos processos normais multivariados, e com o objetivo de analisar processos autocorrelacionados, Kalgonda e Kulkarni em 2004 estudaram processos bivariados autocorrelacionados de ordem 1 ($VAR(1)$) utilizando conceitos de series temporais multivariadas (Reinsel, 1997; Wei, 1990; Box e Jenkins, 1994), uma variante da estatística M de Hayter e Tsui (1994) e estudaram uma forma alternativa de determinação da constante $C_{R,\alpha}$. Essa variante leva em consideração a matriz de correlação-cruzada que tem a informação da variação temporal estimada através do sistema linear de Yule-Walker (Reinsel, 1997). Embora tenham feito poucas simulações, os resultados fornecidos por esses autores mostram que a proposta da variante da estatística M é adequada para o controle de processos autocorrelacionados.

Todos os autores mencionados nos parágrafos anteriores têm sempre preocupação muito grande com a qualidade das análises realizadas pelas cartas de controle. A tradução numérica da qualidade de uma carta de controle é feita através da obtenção de duas grandezas que determinam o quanto os limites de controle das cartas construídas detectam problemas no processo, sem que de fato eles ocorram (alarmes falsos) e o quanto os limites de controle são capazes de detectar mudanças ocorridas no processo quando elas de fato ocorrem (alarme verdadeiro). Essas grandezas são aquelas as quais conhecemos como *ARL in control* e *ARL out of control*, respectivamente.

Nessa dissertação utilizaremos de forma intensiva os valores estimados dos *ARL in control* e *out of control* para avaliação da qualidade de uma carta de controle. Mais detalhes sobre a obtenção desses valores, bem como ferramentas para a análise de processos univariados e multivariados podem ser encontradas em Montgomery (2004).

1.2 – Objetivos

Como pode ser observado não há na literatura nenhuma análise sobre a aplicação da técnica de núcleos estimadores para determinação da distribuição da estatística M de Hayter e Tsui (1994) independente da distribuição fornecida pelo processo multivariado.

Outra observação que pode ser feita e que não há nenhuma comparação conclusiva entre a qualidade dos limites de controle obtidos via estatística T^2 de Hotelling, estatística M de Hayter e Tsui e componentes principais, nem no caso em que os limites de controle são obtidos de forma paramétrica nem nos casos em que os limites são obtidos de forma não paramétrica.

Além disso, não existe na literatura nenhuma comparação entre os resultados fornecidos pela técnica de núcleo-estimador quando se usa o método multi-estágio estudado por Polansky e Baker (2000) para a determinação da janela ótima h com resultados fornecidos pelo método estudado por Bessegato (2001). Pode-se salientar que, além disso, não existem estudos que mostrem qual o valor de b deveria ser utilizado da determinação do estágio do método multi-estágio de Polansky e Baker (2000).

Dessa forma, foi feito nessa dissertação um estudo extensivo sobre a construção de cartas de controle tanto para a estatística T^2 , quanto para a estatística M e para as componentes principais para processos que possuem distribuição normal multivariada autocorrelacionada ou não e processos com distribuição não-normal multivariada autocorrelacionada ou não. Para cada uma dessas estatísticas, considerando cada uma dessas distribuições, serão calculados limites de controle paramétricos (quando a distribuição do processo era a normal autocorrelacionada ou não) e limites de controle não-paramétricos sendo esses limites determinados através de quatro métodos diferentes: o método estudado por Polansky e Baker (2000) utilizando o estágio $b=1$ e $b=5$, o método estudado por Bessegato (2001) e a função de distribuição empírica.

Após obter os limites de controle para cada uma dessas técnicas um dos nossos objetivos será comparar, utilizando o erro, erro quadrático e erro absoluto, a diferença entre os limites de controle paramétricos e os limites de controle não-paramétricos de forma a determinar qual dos métodos não-paramétricos forneceram melhores estimativas para os limites de controle teóricos do processo (menores erros) e, além disso, verificar qual valor do estágio do núcleo-estimador estudado por Polansky e Baker (2000) é o mais adequado na determinação da função de distribuição de uma variável.

Com o objetivo de comparar a qualidade de cada uma das cartas de controle construídas, calcularemos empiricamente os valores dos *ARL in control* e *ARL out of control*.

Com o conhecimento desses valores será possível verificar quais das cartas de controle têm melhor desempenho na detecção de um menor número de alarmes falsos e um maior número de alarmes verdadeiros e quais deverão ser utilizadas no monitoramento do processo.

1.3 – Resultados relevantes

Ao final dessas análises concluímos, com base nas simulações realizadas que a estatística M e a constante $C_{R,\alpha}$ de Hayter e Tsui (1994) são as estatísticas mais confiáveis para o estudo de um processo multivariado, seja ele normal ou não, autocorrelacionado ou não um vez que na maioria dos casos, as cartas de controle para essa estatística sempre forneceram bons valores de *ARL in control* e *ARL out of control*.

Em relação ao método de obtenção dos limites de controle de forma não-paramétrica observou-se que os métodos estudados por Polansky e Baker (2000) com $b=5$ e Bessegato (2000) sempre possuíram um bom desempenho (quando comparado com os limites de controle teóricos) sendo a função de distribuição empírica, mesmo em casos em que as amostras analisadas eram grandes, o estimador não-paramétrico que teve o pior desempenho frente aos demais métodos não paramétricos (núcleo-estimador).

O resultado mais relevante, e que representa uma inovação para a área, foi observado quando o cálculo dos limites de controle do processo, utilizando a estatística M de Hayter e Tsui (1994), é feito através do método de núcleo estimador. Foi possível observar que mesmo nos casos em que possuímos amostras pequenas ($m=50$ amostras) os limites calculados via núcleo-estimador forneceram resultados bem precisos frente aos limites calculados de forma teórica (que são calculados utilizando simulações em que se precisa gerar vetores aleatórios de tamanho 100.000 do processo, no caso de se possuir a distribuição normal ou ainda, para o caso não-normal, possuir no mínimo 500 valores amostrais).

1.4 – Organização

O conteúdo dessa dissertação está organizado em 9 capítulos e em 3 anexos. No Capítulo 2 apresentamos a descrição das estatísticas utilizadas nessa dissertação para sumarizar as informações multivariadas para a construção das cartas de controle. No Capítulo 3 apresentamos a descrição das técnicas estatísticas não-paramétricas para a construção das cartas de controle. Nos Capítulos 4 e 5 apresentamos a estratégia de execução das simulações dos processos normais multivariados não-autocorrelacionados, os resultados e as análises dessas simulações. No Capítulo 6 apresentamos a estratégia de execução das simulações dos

processos não-normais multivariados e não-autocorrelacionados, os resultados e as análises dessas simulações. No Capítulo 7 apresentamos os resultados dos processos normais e não-normais multivariado autocorrelacionado, os resultados e as análises dessas simulações. No Capítulo 8 é apresentado um exemplo prático de aplicação das técnicas abordadas nessa dissertação e finalmente no Capítulo 9 as considerações finais dessa dissertação.

Capítulo 2 – Cartas de Controle

Neste capítulo apresentamos as cartas de controle de Shewhart para processos univariados e a definição de ARL (*Average Run Length*). Para processos multivariados nos quais não há correlação entre as observações no que se refere a ordem de coleta (tempo) apresentamos as cartas construídas via estatística T^2 de Hotelling (1947), estatística M de Hayter e Tsui (1994) e a técnica de componentes principais (Jackson, 1959, Souza e Rigão, 2005). Para os processos multivariados em que há correlação das observações na ordem de coleta (tempo), o que é chamado de processo autocorrelacionado, apresentamos a abordagem de Kalgonda e Kulkarni (2004).

2.1 – Cartas de Controle Univariadas para a Média do Processo

2.1.1 – Cartas de Controle de Shewhart para Processos Não-Autocorrelacionados

Considere m amostras aleatórias de tamanho n , da variável de interesse X , proveniente de um processo. Na terminologia de controle de qualidade diz-se que se têm m subgrupos de tamanho n , $n \geq 1$ (Shewhart, 1929). Segundo Montgomery (2004) para verificar a estabilidade e a existência de causas especiais que atuam sobre o processo, é comum utilizar cartas de controle univariadas de Shewhart (Shewhart, 1929). Essas cartas são construídas a partir do valor de uma estatística $T(x)$ que é calculada para cada amostra do processo e logo em seguida os valores dessa estatística são colocados em um gráfico ou carta de controle.

No controle de qualidade tradicional a obtenção dos limites de controle para a média ou variabilidade do processo parte do pressuposto de que a distribuição da característica de qualidade considerada é normal sendo os limites de controle para a média do processo definidos por:

$$LSC = \mu + k\sigma; LM = \mu; LIC = \mu - k\sigma \quad (2.1)$$

sendo μ a média do processo e σ o desvio-padrão. Em geral utiliza-se o valor $k = 3$ para a obtenção dos limites de controle, gerando assim, uma carta de controle com um erro do tipo I igual a $\alpha = 0,0027$. Isso significa que se o processo é estável a probabilidade de que um valor amostral de $T(x)$ esteja fora dos limites de controle por razões puramente aleatórias e não devido a mudanças no processo é igual a 0,0027. Este erro é chamado de “alarme falso”.

Um outro erro que pode ser observado na construção da carta de controle é o erro do tipo II e representa a chance de não detectamos uma mudança significativa no processo quando ela de fato existe sendo a probabilidade desse erro ocorrer denotada por β . Assim a probabilidade de detectar-se na carta de controle uma mudança significativa no processo quando ela de fato existe (“alarme verdadeiro”) é dada por $1 - \beta$. Cartas de controle com

baixos valores de β devem ser preferidas e uma das formas de minimizar essa probabilidade é aumentando o número de amostras observadas do processo (Montgomery , 2004).

Para a construção de cartas de controle univariadas é desejável conhecer os parâmetros teóricos do processo, porém na prática, esses parâmetros são desconhecidos e estimados através de amostras aleatórias do processo de interesse. Para m amostras com subgrupos de tamanho $n=1$ é usual estimar μ e σ^2 por:

$$\hat{\mu} = \bar{x} = \frac{1}{m} \sum_{i=1}^m x_i \quad (2.2)$$

e

$$\hat{\sigma}^2 = s^2 = \frac{1}{m-1} \sum_{i=1}^m (x_i - \bar{x})^2 \quad (2.3)$$

e para m amostras com subgrupos de tamanho maior $n > 1$, é usual estimar μ e σ^2 por:

$$\hat{\mu} = \bar{\bar{x}} = \frac{1}{m} \sum_{i=1}^m \bar{x}_i \quad (2.4)$$

e

$$\hat{\sigma}^2 = s^2 = \frac{\sum_{i=1}^m \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2}{\sum_{i=1}^m (n_i - 1)} \quad (2.5)$$

sendo \bar{x}_i a média amostral do subgrupo i e x_{ij} o valor da observação j do subgrupo i , $j=1,2,\dots,n$, $i=1,2,\dots,m$.

Existem outros estimadores para a variância σ^2 e para o desvio padrão σ porém esses estimadores não serão utilizados nessa dissertação. Para maiores detalhes consultar Montgomery (2004). Nessa dissertação iremos abordar apenas o caso em que o tamanho dos subgrupos é $n=1$.

2.1.2 – O *ARL in control* e o *ARL out of control*

Duas medidas de desempenho que são utilizadas para avaliar a qualidade da carta de controle são *ARL in control* (*Average Run Length in control*) e o *ARL out of control* (*Average Run Length out of control*). O *ARL in control* representa o número de amostras observadas do processo até que um alarme falso seja detectado, já o *ARL out of control* representa o número de amostras observadas do processo até que um alarme verdadeiro seja detectado. Essas medidas são definidas, respectivamente por:

$$ARL \text{ in control} = \frac{1}{\alpha} \quad (2.6)$$

$$ARL \text{ out of control} = \frac{1}{1-\beta} \quad (2.7)$$

Considerando-se a variável aleatória Y como o número de amostras observadas do processo até que se obtenha o primeiro alarme falso, então Y tem uma distribuição geométrica com esperança e variância dadas por:

$$E(Y) = \frac{1}{\alpha} \quad (2.8)$$

$$Var(Y) = \frac{1-\alpha}{\alpha^2} \quad (2.9)$$

Por analogia, se Y é a variável aleatória que representa o número de amostras observadas do processo até que se obtenha o primeiro alarme verdadeiro, então Y tem uma distribuição geométrica com esperança e variância dados por:

$$E(Y) = \frac{1}{1-\beta} \quad (2.10)$$

$$Var(Y) = \frac{\beta}{(1-\beta)^2} \quad (2.11)$$

A Figura 2.1 apresenta uma ilustração das situações onde observamos os valores do $ARL \text{ in control}$ e $ARL \text{ out of control}$. Na situação (a) nenhuma mudança foi realizada na média do processo mas a 10ª amostra apresenta um valor fora dos limites de controle enquanto na situação (b) houve uma alteração após a 10ª amostra e esta mudança somente foi detectada na 7ª amostra coletada a partir do momento da mudança. Nesse exemplo houve um $ARL \text{ in control}$ igual a 10 e um $ARL \text{ out of control}$ igual a 7.

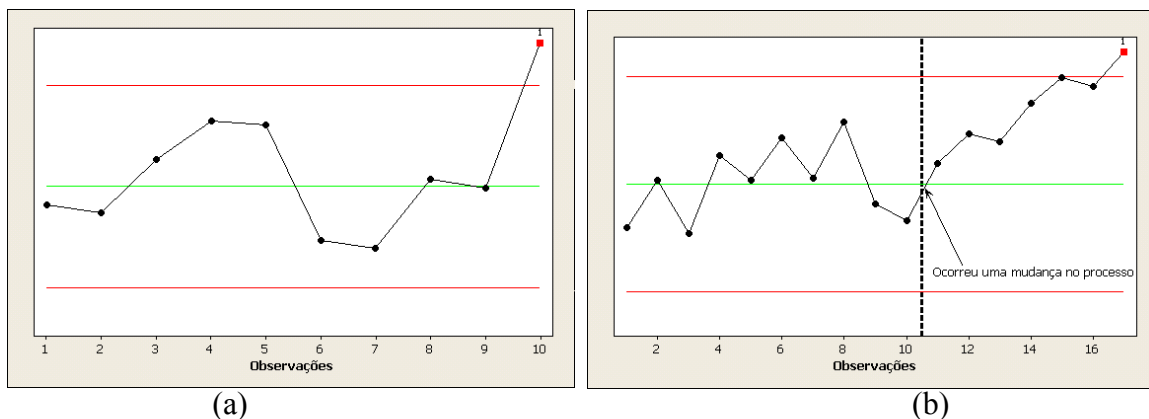


Figura 2.1: Carta de controle: exemplo de $ARL \text{ in control}$ e $ARL \text{ out of control}$.

A análise conjunta do *ARL in control* e *ARL out of control* indica a eficácia da carta de controle. Quanto maior for o *ARL in control* e quanto menor o *ARL out of control* mais eficiente é a carta pois seus limites conseguem detectar rapidamente a falta de controle do processo quando ele sofre alguma modificação e demora a acusar falta de controle quando ele não sofre nenhuma modificação.

2.2 – Carta de Controle Multivariada para o Vetor de Médias de Processos Não-Autocorrelacionados

Alguns usuários das cartas de controle ao monitorar um processo multivariado com p variáveis de interesse utilizam p cartas controle univariadas de Shewhart simultaneamente para verificar a estabilidade do processo. A Figura 2.2 apresenta a ilustração de duas cartas construídas separadamente, com a probabilidade do erro do tipo I igual a $\alpha=0,0027$, para 25 amostras de um processo bi-variado em que as variáveis são provenientes da distribuição normal padrão e possuem uma correlação de 0,90 ($\rho = 0,90$).

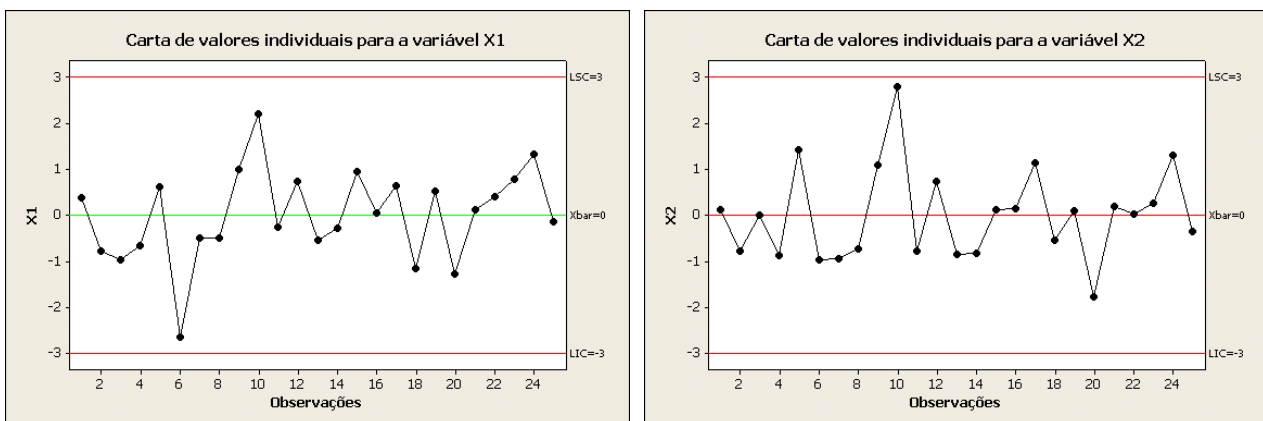


Figura 2.2: Cartas de controle individuais para um processo bi-variado ($\rho = 0,90$).

Considerando-se o nível de significância de α na obtenção dos limites de controle de cada carta, a probabilidade do erro do tipo I global (α_δ) do processo descrito no exemplo é dado por:

$$\alpha_\delta = 1 - (1 - \alpha)^p \quad (2.12)$$

e dessa forma a probabilidade do erro do tipo I global é igual a 0,00539 que é praticamente duas vezes maior que o α pré-estabelecido originalmente.

A metodologia de Bonferroni (Bonferroni, 1930) pode ser usada para manter o nível de significância global pré-estabelecido. Nesse caso, cada carta de controle é feita utilizando-se o nível de significância igual a (α/p) e assim o nível de significância global da análise do

processo multivariado será igual ao valor fixado a priori α . Um problema que ocorre ao utilizarmos esse método é que a diminuição do valor do nível de significância das cartas univariadas aumenta a amplitude dos limites de controle das cartas, tornando mais difícil a detecção da falta de controle do processo. Por exemplo para $p=10$, se o nível de significância global for $\alpha = 0,0027$, usando o método de Bonferroni, cada carta de controle univariada deveria ser construída assumindo um nível de significância de $\alpha = 0,00027$ o que acarretaria num baixo poder de detecção de mudança do processo quando essas de fato existirem. Assim, é necessário construir cartas de controle multivariadas para o monitoramento de processos que mantenham o nível de significância global no valor pré-fixado.

Nas seções a seguir apresentamos as cartas de controle para processos multivariados que serão discutidos nessa dissertação. Como estas dependem em alguns casos da suposição da distribuição normal multivariada ela será introduzida inicialmente.

2.2.1 – A Distribuição Normal Multivariada

Considere $X = [X_1, X_2, \dots, X_p]'$ um vetor aleatório contendo p características de qualidade. Dizemos que X tem distribuição normal multivariada se a função de densidade de X tem a forma:

$$f(x) = (2\pi)^{-\frac{p}{2}} |\Sigma|^{-\frac{1}{2}} \times \exp\left[-\frac{1}{2}(x - \mu)' \Sigma^{-1}(x - \mu)\right] \quad (2.13)$$

onde $x \in \mathfrak{R}^p, \mu \in \mathfrak{R}^p$ sendo $\mu = [\mu_1, \mu_2, \dots, \mu_p]'$ um vetor de dimensão $p \times 1$ que representa a esperança matemática do vetor aleatório X , isto é, $\mu_i = E(X_i), i=1,2,\dots,p$ e Σ uma matriz positiva definida de dimensão $p \times p$ que representa a matriz de variâncias e covariâncias do vetor aleatório X e é denotada por:

$$\Sigma_{p \times p} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1p} & \sigma_{2p} & \cdots & \sigma_{pp} \end{bmatrix} \quad (2.14)$$

em que $\sigma_{ii} = \sigma_i^2 = Var(X_i), i = 1, 2, \dots, p$ e $\sigma_{ij} = Cov(X_i, X_j), i, j = 1, 2, \dots, p, (i \neq j)$, com $\sigma_{ii} \in (0, \infty)$ e $\sigma_{ij} \in (-\infty, \infty)$. A correlação entre as variáveis X_i e X_j é definida por:

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}} \sqrt{\sigma_{jj}}}, \quad i, j = 1, 2, \dots, p \quad \text{com} \quad -1 \leq \rho_{ij} \leq 1.$$

No caso em que $p=2$ a função de densidade de X pode ser escrita em função do coeficiente de correlação como, isto é,

$$f(x) = \frac{1}{2\pi\sqrt{\sigma_{11}\sigma_{22}(1-\rho_{12}^2)}} \times \exp\left\{\frac{1}{2(1-\rho_{12}^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_{11}}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_{22}}\right)^2 - 2\rho_{12}\left(\frac{x_1-\mu_1}{\sqrt{\sigma_{11}}}\right)\left(\frac{x_2-\mu_2}{\sqrt{\sigma_{22}}}\right)\right]\right\} \quad (2.15)$$

onde $x \in \mathcal{R}^2$, $\mu_i \in \mathcal{R}^2$, $-1 \leq \rho_{12} \leq 1$, $\sigma_{ii} \in (0, \infty)$. Como ilustração apresenta-se alguns gráficos dessa distribuição, e suas respectivas curvas de nível, nas Figuras 2.3 e 2.4.

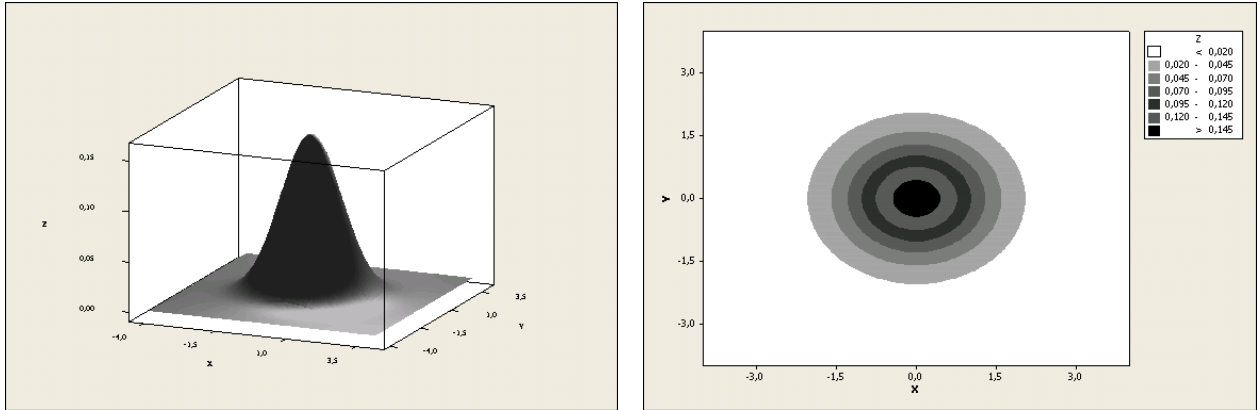


Figura 2.3: Distribuição normal multivariada com $p=2$, $\mu_1 = \mu_2 = 0$, $\sigma_{11} = \sigma_{22} = 1$ e $\rho_{12} = 0$.

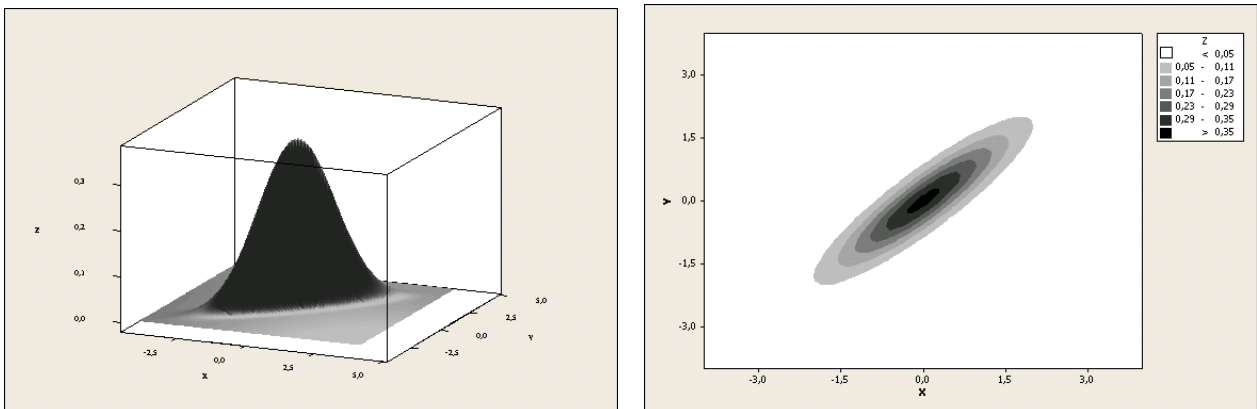


Figura 2.4: Distribuição normal multivariada com $p=2$, $\mu_1 = \mu_2 = 0$, $\sigma_{11} = \sigma_{22} = 1$ e $\rho_{12} = 0,9$.

Em geral os parâmetros μ e Σ são desconhecidos e estimados a partir das amostras aleatórias observadas no processo. Para estimar estes parâmetros considere uma amostra aleatória de tamanho m ($m > 1$), denotada por $X = (X_1, X_2, \dots, X_m)$ sendo $X_i = (X_{i1}, X_{i2}, \dots, X_{ip})'$ onde X_{ij} representa a medida da variável j no elemento amostral i ($j=1, 2, \dots, p$; $i=1, 2, \dots, m$). A média amostral da j -ésima variável é definida por:

$$\bar{X}_j = \frac{1}{m} \sum_{i=1}^m X_{ij} \quad , \forall j = 1, 2, \dots, p \quad (2.16)$$

e o vetor de médias amostral é dado por:

$$\bar{X} = (\bar{X}_1, \bar{X}_2, \dots, \bar{X}_p)' \quad (2.17)$$

A matriz de variâncias e covariâncias amostral é definida por:

$$S = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1p} \\ S_{21} & S_{22} & \cdots & S_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ S_{p1} & S_{p2} & \cdots & S_{pp} \end{bmatrix} \quad (2.18)$$

em que

$$\begin{cases} S_{jj} = \frac{1}{m-1} \sum_{i=1}^m (X_{ij} - \bar{X}_j)^2 & , j = 1, 2, \dots, p \\ S_{jk} = \frac{1}{m-1} \sum_{i=1}^m (X_{ij} - \bar{X}_j)(X_{ik} - \bar{X}_k) & , j, k = 1, 2, \dots, p, (j \neq k) \end{cases} \quad (2.19)$$

Esses estimadores são não viciados para μ e Σ respectivamente (Anderson, 1984) e são obtidos pelo método dos momentos (Berger e Casella, 2002). As estimativas de máxima verossimilhança de μ e Σ são dadas por:

$$\hat{\mu} = \bar{X} \text{ e } \hat{\Sigma} = \frac{(m-1)S}{m}$$

Quando o vetor aleatório X tem distribuição normal multivariada e os parâmetros μ e Σ são conhecidos é possível mostrar que a variável $(X_i - \mu)' \Sigma^{-1} (X_i - \mu)$ têm distribuição Qui-quadrado com p graus de liberdade. No entanto, quando μ e Σ são estimados por \bar{X} e S , respectivamente, a variável $(X_i - \bar{X})' S^{-1} (X_i - \bar{X})$ tem distribuição $c \times F_{p, m-p}$ onde $c = p(m^2 - 1) / [m(m - p)]$ e $F_{p, m-p}$ representa a distribuição F de Fisher-Snedecor com p e $m-p$ graus de liberdade, respectivamente.

Para ilustrar o problema motivado no início da seção 2.2 sobre a utilização de cartas de controle univariadas para o controle do processo multivariado, considere um processo no qual X tem distribuição normal bi-variada em que a média das variáveis desse processo são iguais a zero, as variâncias são iguais a 1 e a correlação entre as variáveis é igual a 0,9. Quando sobrepomos os limites de controle univariados calculados através das expressões definidas em (2.1) com $k=3$ (o que equivale a utilizar um nível de significância individual de $\alpha = 0,0027$ para cada carta) sobre a elipse de confiança do processo (obtida através da distribuição Qui-Quadrado com p graus de liberdade a um nível de significância global de $\alpha = 0,0027$), obtemos os resultados da Figura 2.5 na qual observamos que existe uma diferença entre as regiões de controle fornecidas pela elipse de confiança e o cruzamento dos limites de controle

univariados. Observa-se que existe uma região em que amostras fora dos limites de controle univariados estão dentro da elipse de confiança (áreas mais escuras) e outra região em que amostras dentro dos limites de controle univariados estão fora da elipse de confiança (áreas mais claras). Esse fato mostra a importância de se utilizar a densidade conjunta das variáveis quando se faz o monitoramento de processos multivariados.

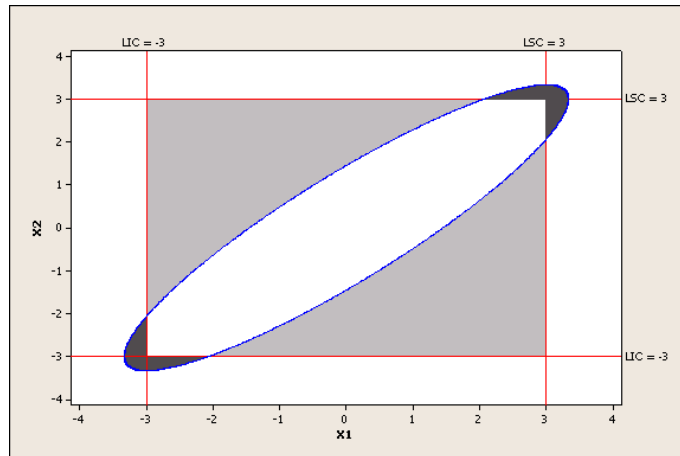


Figura 2.5: Comparação da elipse de confiança com os limites de controle univariados com $\alpha = 0,0027$.

2.2.2 – A Estatística T^2 de Hotelling

Segundo Johnson e Wichern (2002) e Mason e Young (1999), a estatística T^2 proposta por Hotelling (1947) é utilizada para medir a distância Euclidiana ao quadrado, ponderada pela matriz inversa de covariâncias, entre o vetor observado X_i e o vetor de médias populacional μ . O valor desta estatística é equivalente ao valor do teste *t-student* elevado ao quadrado para o caso univariado. O valor da estatística T^2 de Hotelling (1947) para o *i-ésimo* elemento amostral X_i quando μ e Σ são conhecidos é dado por:

$$T_i^2 = (X_i - \mu)' \Sigma^{-1} (X_i - \mu), \quad i = 1, 2, \dots, m \quad (2.20)$$

e por:

$$T_i^2 = (X_i - \bar{X})' S^{-1} (X_i - \bar{X}), \quad i = 1, 2, \dots, m \quad (2.21)$$

quando μ e Σ são desconhecidos e estimados através dos estimadores \bar{X} e S . A estatística descrita em (2.20) tem distribuição Qui-Quadrado com p graus de liberdade enquanto a estatística (2.23) tem distribuição relacionada com F de Fisher-Snedecor, isto é:

$$T^2 \sim \frac{p(m^2 - 1)}{m(m - p)} F_{p, m-p} \quad (2.22)$$

onde $F_{p, m-p}$ é a distribuição F com p e $m-p$ graus de liberdade como é mostrado em Hotelling (1947).

Chou, *et al.* (1999) estudaram outros três estimadores para a matriz de covariâncias amostral Σ e mostraram que apesar do estimador S ser considerado sensível à presença de *outliers*, ele forneceu maior sensibilidade na detecção de pontos fora de controle quando a estatística T^2 de Hotelling (1947) foi utilizada na construção de cartas de controle para processos multivariados. Dessa forma, nesta dissertação, apenas a matriz de covariâncias S , definida em (2.18) será utilizada.

Para construir uma carta de controle para o vetor de médias do processo usando a estatística T^2 de Hotelling (1947) precisamos calcular o valor do limite superior de controle (*LSC*) uma vez que os valores dessa estatística são sempre não negativos. Quando μ e Σ são conhecidos esse limite é dado por:

$$LSC = \chi_{p;1-\alpha}^2 \quad (2.23)$$

e quando μ e Σ são desconhecidos esse limite é dado por:

$$LSC = \frac{p(m^2 - 1)}{m(m - p)} F_{1-\alpha; p; m-p} \quad (2.24)$$

onde $(1 - \alpha)$ é o nível de confiança da carta, $\chi_{p;1-\alpha}^2$ é a ordenada da distribuição Qui-quadrado com p graus de liberdade e $F_{1-\alpha; p; m-p}$ é a ordenada da distribuição F de Fisher-Snedecor com $p, m-p$ graus de liberdade para a probabilidade acumulada igual a $1 - \alpha$.

Um dos problemas ao utilizar a estatística T^2 de Hotelling (1947) para controle de processos multivariados é que, quando observamos valores fora dos limites de controle a identificação de qual variável do processo é responsável pela falta de controle não é imediata, sendo necessário uma análise em separado para essa identificação. Segundo Mahalanobis (1936) essa identificação pode ser feita da seguinte forma: Considere a estatística d_j definida por:

$$d_j = T_i^2 - T_{i(j)}^2$$

em que T_i^2 é o valor da estatística T^2 para o i -ésimo elemento amostral e $T_{i(j)}^2$ é o valor da estatística T^2 para o i -ésimo elemento amostral excluindo-se a j -ésima variável do processo com $i=1,2,\dots,m$ e $j=1,2,\dots,p$. Dessa forma, quando algum valor da estatística T^2 de Hotelling (1947) é verificado fora do limite de controle, o maior valor da estatística d_j ($j=1,2,\dots,p$) indica qual das p variáveis do processo é responsável pela falta de controle. Colenghi e Mingoti (2005a e 2005b) implementaram em S-plus o teste de T^2 de Hotelling (1947) com a identificação dessas variáveis. Nessa dissertação não será utilizada essa identificação da variável responsável pela falta de controle do processo.

A estatística T^2 de Hotelling (1947) pode ser obtida quando possuímos subgrupos de tamanho maior que 1 ($n > 1$), porém, a sua forma de obtenção e o estudo desses tipos de processos não serão mostrados nessa dissertação. Maiores detalhes podem ser obtidos em Montgomery (2004).

Quando a distribuição do vetor aleatório X não é a distribuição normal multivariada, a estatística T^2 de Hotelling (1947) não possui distribuição facilmente determinada (Chou, *et al.*, 2001) dificultando assim a obtenção dos limites teóricos para a construção da carta de controle. Uma das formas de resolvermos esse problema é utilizando-se técnicas estatísticas não-paramétricas para estimar a distribuição de interesse, como a função de distribuição empírica e a técnica de núcleo-estimadores (Chou, *et al.*, 2001) que serão apresentadas no Capítulo 3.

2.2.3 – As Componentes Principais

Uma outra forma de analisar um processo multivariado é utilizando-se a análise de componentes principais, isto é, construindo-se combinações lineares não correlacionadas independentes entre si das p variáveis e monitorando-se o processo a partir dessas combinações.

A análise de componentes principais tem como objetivo explicar a estrutura de variâncias e covariâncias de um conjunto de variáveis de interesse através de combinações lineares das variáveis X_1, X_2, \dots, X_p que são correlacionadas entre si (Mingoti, 2005). Geometricamente, essas combinações lineares representam a seleção de um novo sistema de coordenadas obtidas através da rotação do sistema original, tendo como coordenadas os novos eixos Y_1, Y_2, \dots, Y_p (Johnson e Wichern, 2002 e Mingoti 2005).

A análise de componentes principais não depende da suposição de normalidade multivariada do vetor aleatório X do processo. No entanto, quando o vetor X tem distribuição normal multivariada as componentes principais terão individualmente a distribuição normal univariada e serão independentes (Johnson e Wichern, 2002).

Seja $X = [X_1, X_2, \dots, X_p]'$ um vetor aleatório de interesse com matriz de variâncias e covariâncias Σ conhecida. A partir da decomposição espectral da matriz de covariâncias Σ é possível encontrar os autovalores $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ e os respectivos autovetores normalizados (vetores de comprimento igual a 1) e_1, e_2, \dots, e_p , construindo-se as combinações lineares mostradas a seguir (Anderson, 1984):

$$\begin{aligned}
Y_1 &= e_1' X = e_{11}X_1 + e_{12}X_2 + \dots + e_{1p}X_p \\
Y_2 &= e_2' X = e_{21}X_1 + e_{22}X_2 + \dots + e_{2p}X_p \\
&\vdots \\
Y_p &= e_p' X = e_{p1}X_1 + e_{p2}X_2 + \dots + e_{pp}X_p
\end{aligned} \tag{2.25}$$

Sendo:

$$\begin{aligned}
E(Y_l) &= e_l' \mu \quad , \forall l = 1, 2, \dots, p \\
Var(Y_l) &= e_l' \Sigma e_l = \lambda_{ll} \quad , \forall l = 1, 2, \dots, p \\
Cov(Y_l, Y_k) &= e_l' \Sigma e_k = 0 \quad , \forall l, k = 1, 2, \dots, p \ (l \neq k)
\end{aligned} \tag{2.26}$$

Como as componentes principais são não-correlacionadas, cada componente pode ser analisada de forma independente o que possibilita a construção de cartas de controle univariadas de Shewhart usando-se os valores de cada componente principal.

A soma dos p autovalores obtidos através da decomposição espectral da matriz de covariâncias Σ é igual a soma das variâncias das p variáveis de interesse, sendo essa quantidade chamada de variância total, isto é:

$$\sigma_{11} + \sigma_{22} + \dots + \sigma_{pp} = \lambda_1 + \lambda_2 + \dots + \lambda_p$$

A proporção da variância total explicada pela l -ésima componente principal é dada por:

$$Prop_l = \frac{\lambda_l}{\lambda_1 + \lambda_2 + \dots + \lambda_p} \quad , \forall l = 1, 2, \dots, p \tag{2.27}$$

Apesar da análise de componentes principais fornecer p variáveis não-correlacionadas possibilitando a construção de p cartas de controle univariadas separadas, a análise dessas cartas, em termos práticos pode ser inviável, pela dificuldade de interpretação de todas as componentes e pela não identificação imediata de qual das variáveis originais estariam causando a falta de controle do processo (Hayter e Tsui, 1994). Como a primeira componente principal explica um maior percentual da variância total, uma alternativa é construir apenas uma carta de controle usando-se os valores da primeira componente principal e assim, analisar o processo de acordo com essa carta.

Quando o vetor aleatório X tem distribuição normal p -variada pode-se utilizar duas componentes principais construindo-se uma região de confiança como apresentado na Figura 2.5. Nesta dissertação utilizaremos apenas a primeira componente principal e construiremos a carta de controle a partir dos seus valores. Quando o vetor aleatório X não possui distribuição normal p -variada a distribuição da componente principal Y_i pode ser estimada utilizando-se

as técnicas estatísticas não-paramétricas como a da função de distribuição empírica e núcleo-estimadores. Essas técnicas serão apresentadas no Capítulo 3.

2.2.4 – A Estatística M de Hayter e Tsui

Hayter e Tsui em 1994 propuseram uma metodologia alternativa para o controle de vetores de médias de processos multivariados. O objetivo destes autores era construir uma carta de controle que permitisse detectar automaticamente qual(is) variável(is) era(m) responsável(is) pela falta de controle do processo evitando assim os problemas de identificação que ocorrem na estatística T^2 de Hotelling (1947). Com esse objetivo os autores construíram um intervalo de confiança para a média de cada variável usando um valor crítico comum que leva em consideração a correlação entre as variáveis de interesse. O ponto crítico para este intervalo de confiança foi chamado de $C_{R,\alpha}$. Seja X_1, X_2, \dots, X_m uma amostra aleatória de tamanho m do vetor aleatório $X = [X_1, X_2, \dots, X_p]'$ tendo X uma distribuição normal p -variada com vetor médias μ e matriz de variâncias e covariâncias Σ e $X_i = (X_{i1}, X_{i2}, \dots, X_{ip})'$ o i -ésimo elemento amostral ($i=1, 2, \dots, m$). Seja $C_{R,\alpha}$ uma constante tal que:

$$P\left(\left|\frac{X_j - \mu_j}{\sigma_j}\right| \leq C_{R,\alpha}, \forall j = 1, 2, \dots, p\right) = 1 - \alpha$$

que é equivalente a

$$P([X_j - \sigma_j C_{R,\alpha}; X_j + \sigma_j C_{R,\alpha}] \subset \mu_j, \forall j = 1, 2, \dots, p) = 1 - \alpha \quad (2.28)$$

ou seja, tem-se um intervalo de confiança para cada μ_j do tipo:

$$[X_j - \sigma_j C_{R,\alpha}; X_j + \sigma_j C_{R,\alpha}] \quad , \forall j = 1, 2, \dots, p. \quad (2.29)$$

O processo é então considerado sob controle estatístico se todos os intervalos determinados por (2.29) contiverem os valores teóricos de μ_j , com $j = 1, 2, \dots, p$. A constante $C_{R,\alpha}$ é determinada usando-se a distribuição do máximo do valor absoluto das coordenadas do vetor aleatório X padronizado. Assim, para um dado vetor de observações $X_i = [X_{i1}, X_{i2}, \dots, X_{ip}]'$ o processo será considerado fora de controle quando:

$$M_i = \max_{1 \leq j \leq p} \left| \frac{X_{ij} - \mu_j}{\sigma_j} \right| > C_{R,\alpha}, \quad \forall i = 1, 2, \dots, m. \quad (2.30)$$

O procedimento gráfico para traçar a carta de controle consiste portanto em grafar os pontos relacionados aos valores da estatística M tendo como limite superior de controle o valor

da constante $C_{R,\alpha}$, obtida à um nível de significância de α ($0 < \alpha < 1$). Quando algum ponto é localizado acima do limite superior de controle pode-se identificar automaticamente as variáveis responsáveis pela falta de controle do processo bastando observar quais delas apresentaram valores não pertencentes aos limites $\mu_j \pm \sigma_j C_{R,\alpha}$, $j = 1, 2, \dots, p$.

Quando μ e Σ são conhecidos e $p=2$, a constante $C_{R,\alpha}$ pode ser obtida parametricamente através de tabelas criadas por Bechhofer e Dunnett (1988), porém, quando $p > 2$, não existem tabelas que forneçam os valores da distribuição da estatística M , pois esta se torna muito complexa. Uma solução nesses casos é utilizar técnicas computacionais de simulação para determinar sua distribuição.

Supondo-se que X tem distribuição normal p-variada, a obtenção da constante $C_{R,\alpha}$ é feita de acordo com o procedimento descrito no Quadro 2.1 (Hayter e Tsui, 1994). Neste algoritmo é necessário gerar aleatoriamente N vetores da distribuição normal p-variada padronizada com vetor de médias zero e matriz de variâncias e covariâncias igual a matriz de correlação amostral. Segundo Hayter e Tsui (1994) o número ideal de N é 100000, no entanto, Mingoti e Glória (2003) mostram que para $N=10000$ os valores de $C_{R,\alpha}$ já são obtidos com boa precisão.

Quadro 2.1: Obtenção da constante $C_{R,\alpha}$ no caso da distribuição normal p-variada.

1. Gerar N vetores aleatórios da distribuição normal p-variada com vetor de médias igual a zero e matriz de correlação $R_{p \times p}$, denotados por Z^1, Z^2, \dots, Z^N ;
2. Calcular o valor da estatística M para o i -ésimo vetor aleatório amostral ($i = 1, 2, \dots, N$) da seguinte forma:

$$M_i = \max_{1 \leq j \leq p} |Z_j^i| \quad \forall i = 1, 2, \dots, N$$

em que Z_j^i é a observação da j -ésima variável do i -ésimo vetor aleatório amostral.

3. Encontrar a ordenada correspondente ao $100(1-\alpha)\%$ percentil dos valores de M_1, M_2, \dots, M_N e utilizar este valor como estimativa paramétrica para a constante $C_{R,\alpha}$.

Quando o vetor aleatório X não tem distribuição normal multivariada o valor da constante $C_{R,\alpha}$ poderá ser obtido de forma não-paramétrica através dos valores amostrais observados do processo conforme descrito no Quadro 2.2. Segundo Hayter e Tsui (1994), é necessário observar do processo uma amostra de tamanho mínimo igual a $N=500$ para se obter uma boa distribuição da estatística M . No entanto, Mingoti e Glória (2005b), mostram

que no caso da distribuição normal multivariada para obter-se valores confiáveis da constante $C_{R,\alpha}$ através de procedimentos não-paramétricos de Hayter e Tsui (1994) é necessário uma amostra de tamanho maior ou igual a 5000.

Quadro 2.2: Estimativa não-paramétrica de $C_{R,\alpha}$.

1. Calcular o vetor de médias amostral \bar{X} e matriz de variâncias e covariâncias amostral S a partir da amostra de tamanho N observada do processo.
2. Calcular o valor da estatística M para o i -ésimo vetor amostral, $(i = 1, 2, \dots, N)$ da seguinte forma:

$$M_i = \max_{1 \leq j \leq p} \left| \frac{(X_{ij} - \bar{X}_j)}{\sqrt{s^2_{jj}}} \right| \quad \forall i = 1, 2, \dots, N$$

em que X_{ij} , \bar{X}_j e s^2_{jj} são respectivamente a i -ésima observação da j -ésima variável, a média amostral e a variância amostral da j -ésima variável.

3. Encontrar a ordenada correspondente ao $100(1 - \alpha)\%$ percentil dos valores de M_1, M_2, \dots, M_N e utilizar este valor como estimativa não-paramétrica para a constante $C_{R,\alpha}$.

Assim, seria interessante ter-se um procedimento alternativo para encontrar o valor da constante $C_{R,\alpha}$ nos casos em o tamanho da amostra do processo não fosse muito grande e seus valores não provenientes da distribuição normal p -variada.

A metodologia de Hayter e Tsui (1994) foi proposta considerando o caso em que se têm m amostras (subgrupos) de tamanho $n=1$, no entanto, é possível fazer uma extensão dessa metodologia para o caso que se tem subgrupos de tamanho $n>1$ (Mingoti e Colenghi, 2005a e 2005b). Nessa dissertação apenas o caso de subgrupos de tamanho $n=1$ será considerado.

2.2.5 – Processos Autocorrelacionados Multivariados

As técnicas abordadas até agora levam em consideração apenas processos multivariados em que as observações do processo são independentes ao longo do tempo ou em relação a ordem de coleta das informações. No entanto, existem vários casos nos quais a autocorrelação está presente. Mason e Young (2002) citam que a maioria das operações industriais de fluxo contínuo, principalmente processos químicos, possuem esse tipo de autocorrelação e uma das principais causas para a sua ocorrência é o desgaste gradual de componentes críticos do processo, isso porque o que acontece em um processo no tempo t ,

dependendo da condição do equipamento, pode influenciar no que ocorrerá no tempo $t+h$, $h>0$.

Um dos efeitos ao analisar-se um processo autocorrelacionado através de cartas de controle de Shewhart sem levar em conta a presença da autocorrelação é que a carta construída tem a sua variabilidade alterada em função da autocorrelação das observações do processo. Quando possuímos uma autocorrelação positiva entre as observações têm-se cartas de controle com limites de controle mais estreitos gerando uma grande quantidade de alarmes falsos e quando possuímos uma autocorrelação negativa as cartas de controle têm limites de controle mais amplos dificultando a identificação de causas especiais no processo (Alwan e Roberts, 1995).

Para monitorar processos multivariados autocorrelacionados Kalgonda e Kulkarni (2004) propuseram uma modificação na forma de obtenção da estatística M e da determinação de $C_{R,\alpha}$ de Hayter e Tsui (1994) de modo a fazer com que essa nova carta fosse eficiente para detectar a mudança no vetor de médias de processos autocorrelacionados.

Antes de introduzir o conceito de processo multivariado autocorrelacionado vamos definir a forma de um processo em que não há autocorrelação. Seja uma amostra aleatória de tamanho m de um processo multivariado, com observações independentes entre si, e que pode ser representado da seguinte forma:

$$Y_t = \mu + \varepsilon_t, \quad t = 1, 2, \dots, m \quad (2.31)$$

em que $Y_t = [Y_{t1}, Y_{t2}, \dots, Y_{tp}]'$ representa o vetor aleatório de dimensão $p \times 1$ no tempo t , $\mu = [\mu_1, \mu_2, \dots, \mu_p]'$ é um vetor de médias de dimensão $p \times 1$ que não depende do tempo t e $\varepsilon_t = [\varepsilon_{t1}, \varepsilon_{t2}, \dots, \varepsilon_{tm}]'$ é um vetor aleatório de dimensão $p \times 1$ para o tempo t proveniente de uma distribuição normal multivariada com vetor de médias nulo e matriz de covariâncias igual a Σ_ε . Considera-se nesse caso que os vetores $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_m$ são independentes.

Um processo multivariado autocorrelacionado pode ser representado por diversos modelos de séries temporais estacionários. A classe mais conhecida é a de modelos estacionários autoregressivos e de média móvel ARMA(k, q), cujo modelo é definido genericamente como (Reinsel, 1997):

$$Y_t = \mu + \varepsilon_t + \sum_{j=1}^k \Phi_j (Y_{t-j} - \mu) - \sum_{j=1}^q \Theta_j \varepsilon_{t-j}, \quad t = 1, 2, \dots, m. \quad (2.32)$$

sendo μ, ε_t e Σ_ε como definidos anteriormente, Φ_j e Θ_j as matrizes de dimensões $p \times p$ que contém os parâmetros autoregressivos e de médias móveis do modelo de séries temporais

multivariadas, e Y_{t-j} e ε_{t-j} representam os valores das observações do processo e dos erros no tempo $t-j$ respectivamente. Quando Θ_j é a matriz nula tem-se o modelo autoregressivo multivariado de ordem k ($VAR(k)$) e quando Φ_j é uma matriz nula tem-se o modelo de média móvel multivariado de ordem q ($VMA(q)$). Kalgonda e Kulkarni (2004) propõem a construção de cartas de controle específicas para monitoração de processos bivariados autocorrelacionados de ordem 1 ($VAR(1)$).

A matriz de covariâncias cruzadas, que é utilizada para a construção do modelo proposto por Kalgonda e Kulkarni (2004), é definida da seguinte maneira. Considere a matriz $\Gamma(t, t+h)$ que é a matriz de covariâncias cruzadas entre os vetores Y_t e Y_{t+h} sendo cada elemento (l, k) dessa matriz definido por:

$$\gamma_{lk}(h) = E\{(Y_{l,t} - \mu_{l,t})(Y_{k,t+h} - \mu_{k,t+h})\} \quad (2.33)$$

Como $\mu_t = \mu$ é constante ao longo do tempo, a matriz $\Gamma(t, t+h)$ pode ser re-escrita apenas em função do lag h sendo definida como $\Gamma(h)$. A matriz de correlação cruzada para o modelo $VAR(1)$ no lag h é dado por:

$$\rho(h) = V^{-1/2} \Gamma(h) V^{-1/2} \quad (2.34)$$

em que

$$V = \text{diag}(\gamma_{11}(0), \gamma_{22}(0), \dots, \gamma_{pp}(0))$$

Para que a matriz de correlação cruzada possa ser calculada é necessário conhecer a matriz de covariâncias cruzadas no lag 0 ($\Gamma(0)$). Essa matriz é obtida através da solução das equações de Yule-Walker. Essas equações são mostradas por Reinsel (1997) e são obtidas por:

$$\Gamma(0) = \Phi \Gamma(0) \Phi' + \Sigma_\varepsilon \quad (2.34)$$

em que as matrizes Φ e Σ_ε são conhecidas. Para o processo multivariado gaussiano autocorrelacionado e estacionário de ordem 1 ($VAR(1)$) Y_t terá distribuição normal multivariada com vetor de médias μ , matriz de covariâncias $\Gamma(0)$ e matriz de correlação $\rho(0)$.

Kalgonda e Kulkarni (2004) propõem o uso da carta de Hayter e Tsui (1994) para a análise de processos normais autocorrelacionados calculando-se a constante $C_{R,\alpha}$ a partir da matriz $\rho(0)$, constante chamada por eles de $C_{\rho(0),\alpha}$. Assim tem-se que:

$$P\left(\left|\frac{Y_{j,t} - \mu_j}{\gamma_j(0)}\right| \leq C_{\rho(0),\alpha}, \forall j = 1, 2, \dots, p\right) = 1 - \alpha$$

em que $\gamma_j(0) = \sqrt{\gamma_{jj}(0)}$, $\forall j = 1, 2, \dots, p$. Essa equação é equivalente a:

$$P([Y_{j,t} - \gamma_j(0) \times C_{\rho(0),\alpha}; Y_{j,t} + \gamma_j(0) \times C_{\rho(0),\alpha}] \subset \mu_j, \forall j = 1, 2, \dots, p) = 1 - \alpha \quad (2.35)$$

ou seja, tem-se um intervalo de confiança para cada μ_j do tipo:

$$[Y_{j,t} - \gamma_j(0)C_{\rho(0),\alpha}; Y_{j,t} + \gamma_j(0)C_{\rho(0),\alpha}] \quad , \forall j = 1, 2, \dots, p. \quad (2.36)$$

O processo é então considerado sob controle estatístico se todos os intervalos determinados por (2.36) contiverem os respectivos valores de μ_j com $j = 1, 2, \dots, p$. A constante $C_{\rho(0),\alpha}$ é determinada computacionalmente usando-se a distribuição do máximo do valor absoluto das coordenadas do vetor aleatório Y_t padronizado pelo vetor de médias e pelos desvios padrões obtidos a partir da diagonal principal da matriz de covariâncias $\Gamma(0)$. Assim, para um dado vetor de observações $Y_t = [Y_{t1}, Y_{t2}, \dots, Y_{tp}]'$ o processo será considerado fora de controle quando:

$$M_t^* = \max_{1 \leq j \leq p} \left| \frac{Y_{tj} - \mu_j}{\sqrt{\gamma_{jj}(0)}} \right| > C_{\rho(0),\alpha} \quad , \forall t = 1, 2, \dots, m \quad (2.37)$$

Observe que a estatística M^* apresentada na equação (2.37) é equivalente a estatística M descrita por Hayter e Tsui (1994) exceto pela utilização do desvio padrão $\sqrt{\gamma_{jj}(0)}$ que leva em consideração o modelo de série temporal multivariado do processo. Dessa forma, o procedimento gráfico para traçar a carta de controle consiste em grafar os pontos relacionados ao valores da estatística M^* tendo como limite superior de controle o valor da constante $C_{\rho(0),\alpha}$, obtida à um nível de significância de α ($0 < \alpha < 1$).

Quando algum ponto é localizado acima do limite superior de controle pode-se verificar automaticamente qual a variável foi responsável pela falta de controle do processo bastando observar quais variáveis apresentaram valores que não pertencem aos limites de controle $\mu_j \pm \gamma_j(0)C_{\rho(0),\alpha}$, $j = 1, 2, \dots, p$.

Quando que Y_t tem distribuição normal multivariada com vetor de médias μ e matriz de covariâncias $\Gamma(0)$, a obtenção da constante $C_{\rho(0),\alpha}$ é feita de acordo com o procedimento descrito anteriormente no Quadro 2.1 (página 36) substituindo-se a matriz R pela matriz $\rho(0)$ estimada.

Quando o vetor aleatório Y_t não tem distribuição normal multivariada o valor da constante $C_{\rho(0),\alpha}$ poderá ser obtido de forma não-paramétrica de maneira análoga ao Quadro 2.2 (página 37) trocando nesse caso, a matriz de variâncias e covariâncias $\hat{\Sigma}$ pela matriz de covariâncias $\hat{\Gamma}(0)$.

A partir da modificação da obtenção da estatística M e da constante $C_{\rho(0),\alpha}$ feita por Kalgonda e Kulkarni (2004), vamos utilizar nessa dissertação, uma modificação na estatística T^2 de Hotelling (1947) e nas componentes principais quando analisamos processos autocorrelacionados. Essa modificação consiste em utilizar a matriz de covariâncias $\Gamma(0)$ no lugar da matriz de variâncias e covariâncias Σ e $\rho(0)$ estimada do lugar de R nas equações (2.20) e (2.21) (página 31).

Na prática as matrizes Σ_ε e Φ são estimadas. Nessa dissertação será utilizado o método de mínimos quadrados via *stepwise* (Neumaier e Schneider, 2001). Esse método está computacionalmente implementado no *software* estatístico R podendo ser facilmente acessado através do pacote *mAr.est* (*Estimate of Multivariate AR(p) Model*). Logo após estimar as matrizes Σ e Φ é utilizado o sistema linear de Yule-Walkar, descrito na equação (2.38) para se obter o valor de $\hat{\Gamma}(0)$.

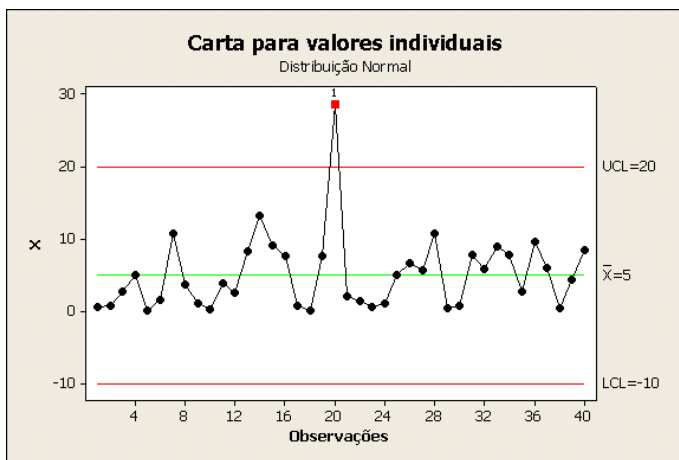
Para processos não-normais, nesta dissertação, serão avaliados os métodos não-paramétricos da função de distribuição empírica e núcleo-estimador para a determinação da constante $C_{\rho(0),\alpha}$.

Capítulo 3 – Técnicas Estatísticas Não Paramétricas para a construção de Cartas de Controle

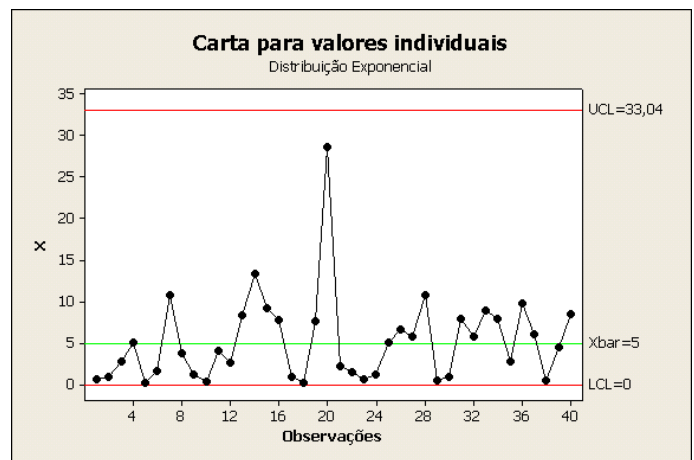
Uma das etapas mais importantes da construção da carta de controle é a identificação da distribuição de probabilidade da estatística de interesse $T(x)$ usada na construção da carta de controle, para que a partir dessa possamos determinar os limites de controle que serão utilizados no monitoramento de processos.

Para ilustrar a importância da identificação da distribuição da estatística de interesse $T(x)$, considere um processo que fornece dados provenientes de uma distribuição exponencial com parâmetro igual a 5. A Figura 3.1 (a) apresenta a carta de controle construída sob a suposição de normalidade para essa amostra de tamanho $m=40$. Observa-se que existe uma amostra localizada fora dos limites de controle da carta indicando que o processo está fora de controle, e além disso, que o limite inferior de controle é negativo, o que não é compatível com o domínio da distribuição exponencial.

O teste de Anderson-Darling resultou numa probabilidade de significância de $p>0,25$ para a hipótese nula que esses dados seguem uma distribuição exponencial. Na Figura 3.1 (b) apresenta-se a carta de controle construída de acordo com a distribuição exponencial a qual indica um processo sob controle estatístico. Assim, o uso incorreto da distribuição normal resultaria em uma carta de controle com um grande número de alarmes falsos.



(a)



(b)

Figura 3.1: Carta de valores individuais para o processo supondo distribuição normal (a) e distribuição exponencial (b)

Neste exemplo, como os dados são provenientes de uma distribuição conhecida, é possível construir uma carta de controle usando métodos paramétricos. No entanto, existem várias situações nas quais a forma matemática da distribuição gerada dos dados não faz parte de uma classe paramétrica conhecida sendo portanto de difícil identificação. Nesses casos, é possível utilizar técnicas não-paramétricas como a função de distribuição empírica e núcleos

estimadores para estimar a função de distribuição acumulada $F(x)$, ou a função de densidade $f(x)$.

Embora em alguns casos a variável original X possa ser transformada utilizando-se a transformação de Box-Cox (Box e Cox, 1964) numa variável Y que tenha distribuição aproximadamente normal, nessa dissertação iremos abordar apenas o caso de estimação não-paramétrica.

3.1 – A Função de Distribuição Empírica

A função de distribuição empírica é um dos estimadores mais utilizados quando se deseja estimar a função de distribuição F de uma variável aleatória (Silverman, 1992). Considere uma amostra de tamanho m denotada por X_1, X_2, \dots, X_m da variável aleatória X que tem função de distribuição contínua F definida por:

$$F(x) = P(X \leq x) \quad (3.1)$$

A função de distribuição empírica no ponto x_i , denotada por \hat{F}_n , é definida por:

$$\hat{F}_n(x_i) = \hat{P}(X \leq x_i) = \frac{\text{número de valores observados na amostra no intervalo } (-\infty; x_i]}{m} \quad (3.2)$$

para cada valor amostral $i=1,2,\dots,m$. É possível mostrar que $\hat{F}_n(x)$ é um estimador não viciado e consistente de $F(x)$, para x fixo (Tacas, 1971 e Serfling, 1980). A ordenada do percentil de ordem $100(\alpha)\%$ é estimada a partir da função inversa de \hat{F}_n e é definida como:

$$x_\alpha = \hat{F}_n^{-1}(\alpha) \quad (3.3)$$

A função de distribuição empírica \hat{F}_n algumas vezes apresenta problemas na estimativa das ordenadas correspondentes aos percentis. Isto é explicado pelo fato de que \hat{F}_n é uma função escada que pondera cada x_i amostral pelo valor m^{-1} , e então, quando o tamanho de amostra é pequeno, os valores estimados muitas vezes não se aproximam dos valores reais da função teórica F . Dessa forma, a aderência de \hat{F}_n à F é melhor quando o tamanho de amostra é grande. Por exemplo, considere uma amostra aleatória de tamanho 50 proveniente de uma distribuição normal com média 10 e variância igual a 4. Estimando-se a função de distribuição teórica F através da função de distribuição empírica pode-se comparar os resultados como pode ser visto na Figura 3.2(a). Observa-se que em alguns pontos a função de distribuição empírica superestima a função de distribuição teórica de X e em outros subestima.

Na Figura 3.2(b) tem-se o resultado para uma amostra de tamanho 1000 proveniente da mesma distribuição normal. Neste caso a função de distribuição empírica apresenta uma boa aderência à função de distribuição teórica de X evidenciando que para amostras maiores, espera-se uma melhoria na estimação da função de distribuição F_x quando se usa a função de distribuição empírica.

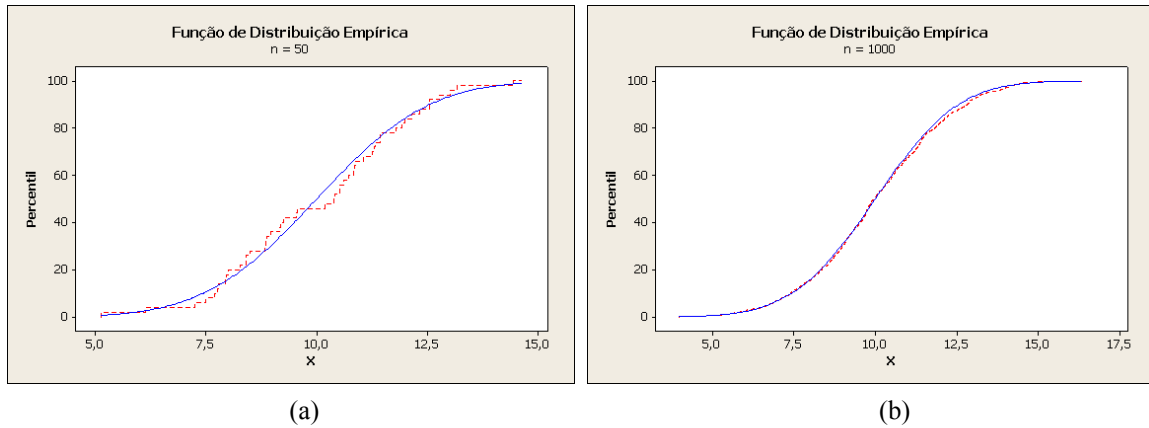


Figura 3.2: Comparação da função de distribuição empírica (---) para uma amostra de tamanho 50 (a) e 1000 (b) proveniente de uma distribuição teórica $N(10,4)$ (—).

3.2 – A metodologia de Núcleo-Estimador

A metodologia de núcleo-estimador é utilizada quando desejamos estimar a função de distribuição ou densidade de uma variável aleatória de interesse. Por suas propriedades assintóticas, esse estimador vem ganhando cada vez mais espaço em artigos nas mais diversas áreas de estudo (Silverman, 1992 e Wand e Jones, 1996, Polansky e Baker, 2000).

Através da escolha de um parâmetro de suavização h , conhecido como janela ótima, determina-se a taxa de suavidade e convergência do estimador. A escolha da janela ótima é a parte fundamental deste método de estimação (Silverman, 1992). Considere uma amostra aleatória de tamanho m definida por X_1, X_2, \dots, X_m de uma variável aleatória X com função de distribuição F . O núcleo estimador de F no ponto x é definido por:

$$\hat{F}_n(x) = \frac{1}{m} \sum_{i=1}^m W\left(\frac{x - X_i}{h}\right) \quad (3.4)$$

onde W é uma função de distribuição que representa a função núcleo e h é a janela definida como parâmetro de suavidade do estimador. Uma outra forma de estimar-se a função de distribuição F no ponto fixo x e através da função de densidade $\hat{f}(x)$ obtida via núcleo estimador. O núcleo estimador da densidade $f(x)$ é definido por:

$$\hat{f}(x) = \frac{1}{mh^*} \sum_{i=1}^m w\left(\frac{x - X_i}{h^*}\right) \quad (3.5)$$

onde w é uma função de densidade que representa o núcleo e h^* é a janela definida como parâmetro de suavidade do estimador. Neste caso, a função de distribuição \hat{F} é estimada por:

$$\hat{F}_{1n}(x) = \int_{-\infty}^x \hat{f}(x) dx \quad (3.6)$$

Nesta dissertação abordamos apenas o estimador definido em (3.4).

3.2.1 – A escolha do Parâmetro de Suavidade

A escolha do parâmetro de suavidade h é parte fundamental do método de estimação das funções de distribuição e densidade via núcleos estimadores pelo fato desse parâmetro ser responsável pelo controle da suavidade da estimativa de $\hat{F}_n(x)$ sendo sua escolha uma questão crítica para a estimação da função de distribuição teórica F . Em geral, h é escolhido de maneira que $\hat{F}_n(x)$ seja um ótimo estimador de F de acordo com uma medida de desempenho e diversos autores já demonstraram que o impacto exercido pela escolha do núcleo w para a estimação da função de densidade ou W para a estimação da função de distribuição, é pequeno (Silverman, 1992).

Para visualizar a influência da escolha da janela h na estimativa de $f(x)$ e $F(x)$, consideramos uma amostra aleatória de tamanho m denotada por: X_1, X_2, \dots, X_m proveniente de uma distribuição normal com média 10 e variância igual a 4. Seja w e W as funções de densidade e distribuição da normal padronizada definidas por:

$$w(u) = \frac{1}{\sqrt{2\pi}} \times \exp\left\{-\frac{u^2}{2}\right\} \quad (3.7)$$

$$W(u) = \int_{-\infty}^u \frac{1}{\sqrt{2\pi}} \times \exp\left\{-\frac{u^2}{2}\right\} du \quad (3.8)$$

Nas Figuras 3.3 e 3.4 apresenta-se os estimadores de $f(x)$ e $F(x)$ considerando os valores $h_1 = 0,1$, $h_2 = 0,3$, $h_3 = 0,5$, $h_4 = 0,7$, $h_5 = 0,9$, $h_6 = 1,0$, $h_7 = 2,0$ e $h_8 = 5,0$ para a janela h . As funções teóricas também são apresentadas para efeito de comparação.

Observa-se que os valores de h^* iguais a 0,7, 0,9 e 1 foram os que resultaram nas funções de densidade estimadas mais próximas da função de densidade teórica. Para valores muito pequenos de h as densidades estimadas apresentaram muitas irregularidades (sub-suavizada) e para valores de h muito grandes a densidade teórica foi super-suavizada.

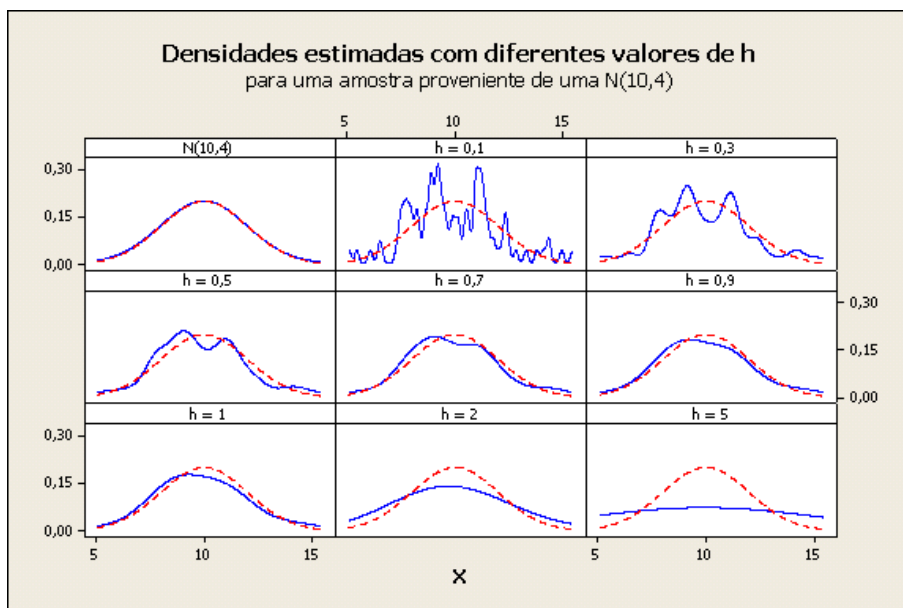


Figura 3.3: Comparação das funções densidade estimadas via núcleo estimador (—) com a função densidade teórica $N(10,4)$ (- - -).

Quanto a função de distribuição as janelas h que resultaram em melhores aproximações foram aquelas em que $h < 1$. Para valores de h maiores que 1, a distribuição teórica foi subestimada, portanto, verifica-se que as estimativas da função de densidade f ou da distribuição F dependem do valor h e além disso, pode-se verificar que para uma mesma amostra, o valor de h que define uma boa estimativa para a função densidade de X não necessariamente é o valor que define uma boa estimativa para a função de distribuição de X . Silverman (1992) e Wand e Jones (1995) salientam que a amplitude da janela ótima h do núcleo estimador da função de densidade é da ordem de $m^{-1/5}$ enquanto para a função de distribuição é na ordem de $m^{-1/3}$.

Eventualmente a metodologia de núcleo estimador pode fornecer resultados não muito adequados quando estimamos a função de densidade ou de distribuição para variáveis que são obtidas de distribuições assimétricas (seja à direita ou à esquerda) ocorrendo o que é chamado de problema de fronteira (Travassos, 2003). Nesses casos é necessário utilizar uma correção para melhorar as estimativas de $f(x)$ e $F(x)$. Esse tipo de correção não será abordada nessa dissertação e para maiores detalhes ver Travassos (2003).

Na literatura existem vários trabalhos que definem métodos de escolha do valor da janela h quando desejamos estimar a distribuição de uma variável de interesse. Nessa dissertação utilizaremos dois métodos: o método *plug-in* estudado por Bessegato (2001) e o de *plug-in* de multiestágio estudado por Polansky e Baker (2000). Esses métodos são apresentados a seguir.

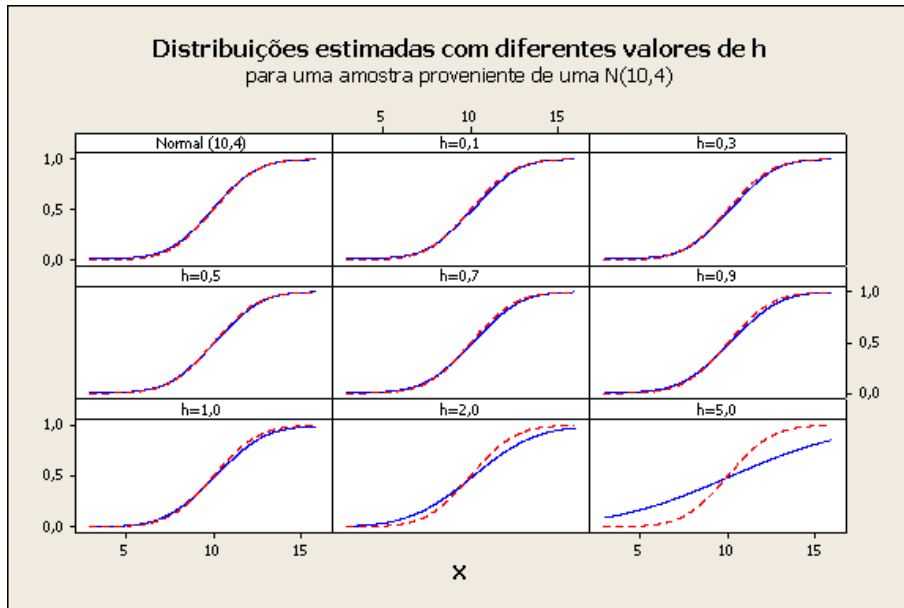


Figura 3.4: Comparação das funções de distribuição estimadas via núcleo estimador (—) com função de distribuição teórica $N(10,4)$ (- - -).

3.2.1 – A escolha do Parâmetro de Suavidade

3.2.1.1–Método de Bessegato para Determinação da Janela Ótima

O critério para a análise da qualidade dos núcleos estimadores consiste em avaliar a proximidade das estimativas obtidas através do núcleo estimador com os verdadeiros valores da função densidade f ou de distribuição F . Como dito anteriormente, a janela ótima h é obtida de forma que a diferença entre o estimador $\hat{F}_n(x)$ seja o mais próximo possível do estimador de $F(x)$, e assim o valor de h é escolhido de forma que o Erro Quadrático Médio Integrado (*Mean Integrated Squared Error*) que é dado por

$$MISE(h) = E \int \{ \hat{F}_n(x) - F(x) \}^2 dx$$

seja minimizado.

Assim, Bessegato (2001) definiu, provou e implementou, de acordo com o método *plug-in*, uma forma para determinação da janela ótima h que é definido por:

$$\hat{h}_{opt} = \left\{ \frac{\int W(x)[1 - W(x)] dx}{\left[\int z^2 dW(z) \right]^2 H} \right\} m^{-\frac{1}{3}} \quad (3.9)$$

em que W é a função de distribuição da normal padronizada definida em (3.8) e $H = \int [F''(x)]^2 dx$ sendo estimado por \hat{H} dado em (3.10) obtido pelo método de *plug-in* através da equação:

$$\hat{H} = \frac{1}{\pi} \int_0^{\Lambda} \lambda^2 \left[|\hat{\phi}(\lambda)|^2 - \frac{1}{m} \right] d\lambda \quad (3.10)$$

em que $|\hat{\phi}(\lambda)|^2$ é a função característica empírica, $\Lambda = \min\{\lambda : |\hat{\phi}(\lambda)|^2 \leq \frac{C}{m}\}$ para algum valor de $C > 1$ e o termo $1/m$ é usado para reduzir a variação de $|\hat{\phi}(\lambda)|^2$ para $\lambda > \Lambda$, identificada a parte que contém a maioria da informação de $f(x)$. Da mesma forma que definido por Bessegato (2001) utilizaremos nessa dissertação o valor de $C=3$. Para maiores detalhes ver capítulo 3 de Bessegato (2001).

3.2.1.2–Método de Polansky e Baker para Determinação da Janela Ótima

Uma outra forma de obter o valor da janela ótima h é descrito por Polansky e Baker (2000) e se baseia no método *plug-in* de multi-estágio. Segundo estes autores a janela ótima \hat{h}_b , que também é definida de forma que a diferença entre o estimador $\hat{F}_n(x)$ seja o mais próximo possível do estimador de $F(x)$, é definida através do algoritmo proposto a seguir: A função de distribuição $F(x)$ pode ser reescrita da seguinte maneira (conhecida agora como função Ψ_r):

$$\Psi_r = \int_{-\infty}^{\infty} f^{(r)}(x) f(x) dx$$

com $r \geq 2$ para qualquer valor inteiro. Essa função pode ser estimada por $\hat{\Psi}_r(g)$ que é definida por:

$$\hat{\Psi}_r(g) = m^{-2} g^{-r-1} \sum_{i=1}^m \sum_{j=1}^m L^{(r)} \left\{ \frac{x_i - x_j}{g} \right\}, g > 0 \quad (3.11)$$

em que L é a função de densidade da distribuição normal padronizada que é utilizada como núcleo, g é um valor positivo e r o estágio do estimador. Para obter o valor ótimo da janela h pelo método *plug-in* de multi-estágio considere o valor da função $\hat{\Psi}_r^{NR}$, que é o valor inicial do algoritmo, definida por:

$$\hat{\Psi}_r^{NR} = \frac{(-1)^{\frac{r}{2}} r!}{(2\hat{\sigma})^{r+1} (r/2)! \sqrt{\pi}} \quad (3.12)$$

em que $\hat{\sigma} = \min\{s, IQR / 1,349\}$ é o desvio-padrão estimado através do amostra aleatória e r é o valor do estágio do estimador, $r \in N$. Assumindo um valor de $b > 0$, inteiro, considere os seguintes passos:

1. Calcule o valor inicial do método usando $\hat{\Psi}_{2b+2}^{NR}$ como descrito em (3.12);
2. Considere um valor de j tal que $j=b$ e obtenha o valor da função $\hat{\Psi}_{2j}(\hat{g}_{2j})$ sendo a constante \hat{g}_{2j} definida por:

$$\hat{g}_{2j} = \left[\frac{2L^{(2j)}(0)}{-m\mu_2(L)\hat{\Psi}_{2j+2}} \right]^{\frac{1}{2j+3}} \quad (3.13)$$

em que o valor de $\hat{\Psi}_{2j+2}$ é definido com base no seguinte critério:

$$\hat{\Psi}_{2j+2} = \begin{cases} \hat{\Psi}_{2b+2}^{NR} & \text{quando } j = b \\ \hat{\Psi}_{2j+2}(\hat{g}_{2j+2}) & \text{quando } j < b \end{cases}$$

ou seja, quando $j < b$ o valor $\hat{\Psi}_{2j+2}$ é obtido de acordo com a equação (3.11) tendo como base o valor da constante \hat{g}_{2j+2} ;

3. Repita o passo 2 diminuindo de uma em uma unidade o valor de j procedendo dessa forma até que $j=1$;
4. O valor de h_b no b -ésimo estágio é definido por:

$$\hat{h}_b = \left[\frac{\rho(k)}{-m\mu_2^2(k)\hat{\Psi}_2(\hat{g}_2)} \right]$$

com $\rho(k) = 2 \int_{-\infty}^{\infty} xk(x)K(x)dx = 0,56418958323$ e $\mu_2(k)=1$, quando k é a função de densidade da normal padrão.

Polansky e Baker (2000) discutem em seu artigo a eficiência do método da escolha da janela ótima h frente à função de distribuição empírica. Esses autores concluem com base em exemplos (um deles aplicados a dados censurados) e através de simulações realizadas (onde foram estimadas as funções de distribuição para dados com distribuição normal, distribuição levemente assimétrica, distribuição altamente assimétricas, distribuição bimodal e distribuição bimodal assimétrica) que as distribuições F estimadas através do núcleo estimador usando o método de *plug-in* de multi-estágio forneceram melhores resultados frente a função de distribuição empírica. Maiores detalhes sobre essa metodologia podem ser obtidos no artigo dos referidos autores.

Serão executadas simulações para verificar o comportamento das funções de distribuição estimadas através do método de núcleo estimador (usando o métodos de Polansky e Baker (2000) com $b=1$ e $b=5$ e Bessegato (2001)) e função de distribuição empírica na construção das cartas de controle para monitoramento de processos multivariados. Os resultados dessas simulações são apresentados nos capítulos 5, 6 e 7.

3.3 – Exemplo de Aplicação

Nessa seção apresentamos uma aplicação das técnicas descritas nos Capítulos 2 e 3 com o intuito de verificar a aplicação em processos com distribuição normal multivariada não autocorrelacionada. Nesse exemplo serão mostrados os valores dos limites de controle teóricos obtido para a estatística T^2 de Hotelling (1947), para a primeira componente principal e para os valores da estatística M de Hayter e Tsui (1994). Serão mostrados também os valores dos limites de controle quando utilizamos as técnicas não-paramétricas de núcleo estimador com a janela ótima h estimada através das metodologias estudadas por Polansky e Baker (2000) e Bessegato (2001), além da função de distribuição empírica. Considere uma amostra aleatória de tamanho $m=100$ proveniente de um processo normal bi-variado com parâmetros desconhecidos e estimados por:

$$\hat{\mu} = \bar{X} = [100,41 \quad 50,24]'$$
$$S_{2 \times 2} = \begin{bmatrix} 4,25 & 1,99 \\ 1,99 & 1,08 \end{bmatrix}$$

O teste de Anderson-Darling (D'Agostinho e Stephens, 1986), usado para verificar a normalidade de cada variável separadamente, indicou que cada variável possui distribuição normal (valor-p > 0,05 em ambos os casos a um nível de 5% de significância).

Na Figura 3.5 (a) e (b) apresenta-se o gráfico de dispersão dos valores das variáveis X_1 e X_2 e o gráfico de probabilidade da distribuição Qui-quadrado conhecido como *QQ plot*, utilizado para avaliar a normalidade multivariada (Mingoti, 2005). Para a construção desse gráfico calcula-se para cada elemento amostral i , a estatística

$$d_i^2 = (X_i - \bar{X})' S^{-1} (X_i - \bar{X}), i = 1, 2, \dots, m$$

na qual a suposição de normalidade p-variada deveria ter distribuição aproximadamente Qui-quadrado com p graus de liberdade. Os valores de d_i^2 são ordenadas e a função de distribuição empírica calculada para cada valor de d_i^2 ordenado. Compara-se então esse valor com a ordenada correspondente da distribuição Qui-quadrado obtida considerando-se a probabilidade acumulada estimada para d_i^2 . A comparação é feita graficamente. Quando a normalidade p-variada é válida este gráfico deve resultar numa reta com coeficiente angular próximo a 1. A Figura 3.5 (b) indica a normalidade bi-variada é plausível para a função de densidade gerada dos dados desse exemplo.

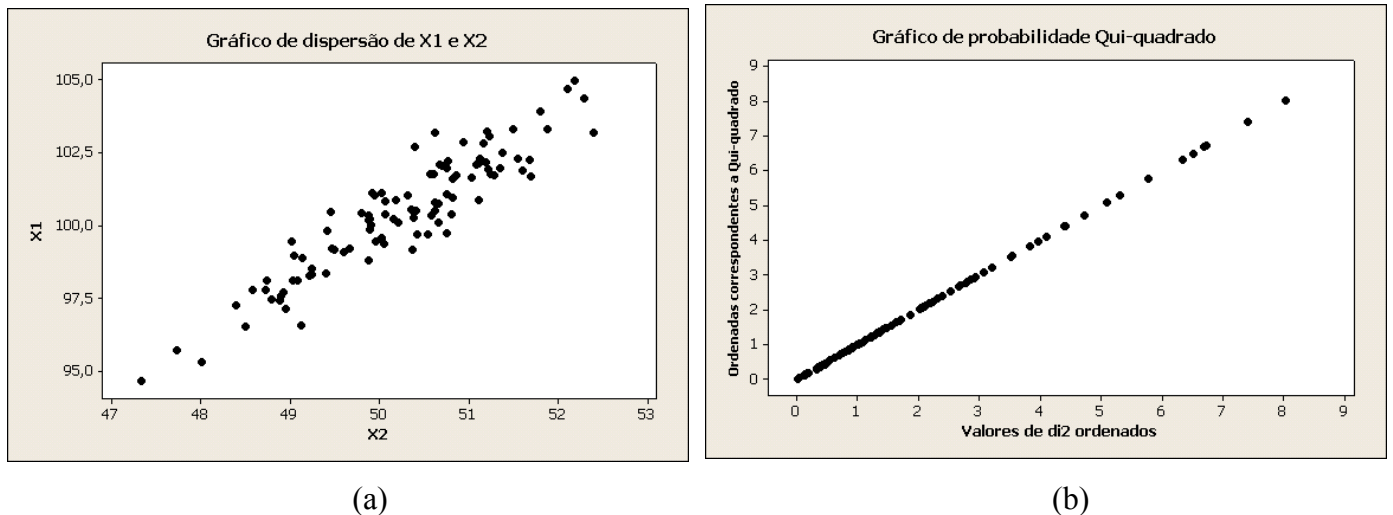


Figura 3.5: Gráfico de dispersão das variáveis X1 e X2 do processo e da distribuição Qui-quadrado.

O histograma dos valores amostrais da estatística T^2 de Hotelling (1947), calculada através da fórmula (2.21) está apresentado na Figura 3.6 (a) e indica uma distribuição assimétrica à direita. Na Figura 3.6 (b) apresenta-se o histograma dos valores amostrais da estatística M de Hayter e Tsui (1994), obtidos através da fórmula (2.30) para os valores das $m=100$ amostras analisadas. Observa-se neste gráfico que a distribuição dessa estatística também é assimétrica à direita.

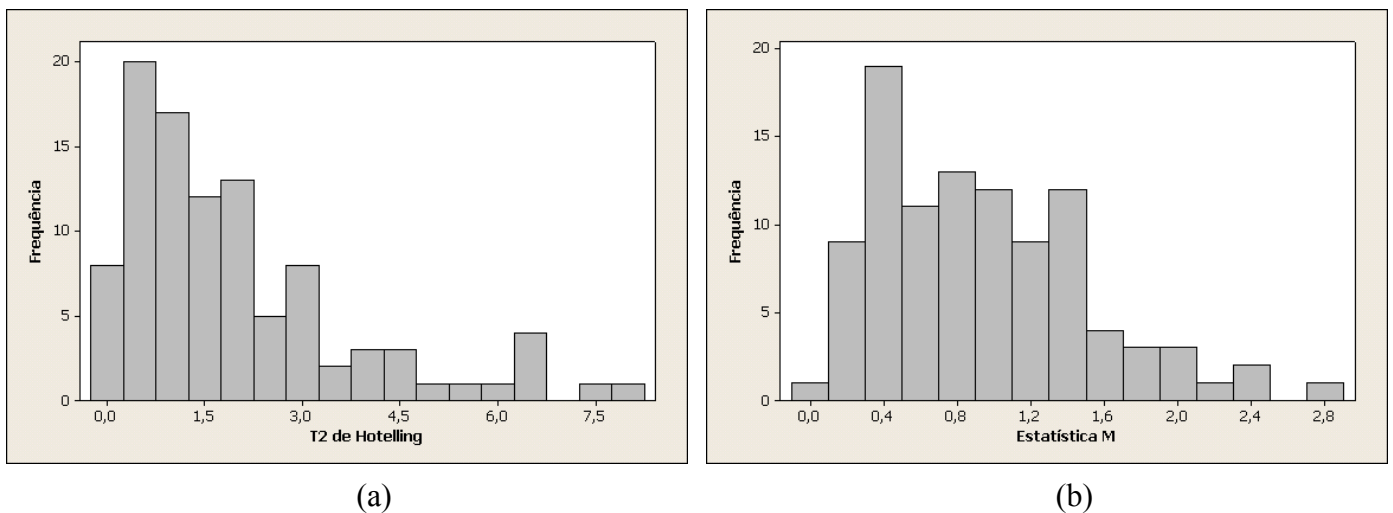


Figura 3.6: Comportamento da estatística T^2 de Hotelling e da estatística M .

A determinação dos limites de controle via estatística T^2 de Hotelling (1947) e via estatística M de Hayter e Tsui (1994), a um nível de 95% de confiança, foi efetuada utilizando-se as técnicas descritas nos capítulos 2 e 3. Para a estatística T^2 de Hotelling (1947) o limite superior de controle paramétrico (Teo) foi estimado de acordo com (2.22) e para estatística M de Hayter e Tsui (1994) foi estimado de acordo com o procedimento descrito no

Quadro 2.1 (página 36) utilizando um total de $N=10000$ amostras como visto por Minogti e Glória (2003). Os limites superior de controle foram obtidos utilizando a técnica de núcleo estimador com a janela ótima h escolhida de acordo com a metodologia estudada por Polansky e Baker (2000) com os valores de $b=1$ (*PB1*) e $b=5$ (*PB5*) e de acordo com Bessegato (2001) (*Bes*) além daquele obtido através da função de distribuição empírica (*Emp*). A função núcleo utilizada foi a da distribuição normal padronizada. Os resultados para a estatística T^2 de Hotelling e estatística M são apresentados na Tabela 3.1.

Tabela 3.1: Limites superior de controle para a estatística T^2 de Hotelling e estatística M ($\alpha = 0,05$).

| Técnica utilizada | LSC | Erro | EQ |
|---|------------|-------------|-----------|
| <i>T² de Hotelling</i> | | | |
| Limite teórico (<i>Teo</i>) | 6,3039 | - | - |
| Polansky e Baker com $b=1$ (<i>PB1</i>) | 6,3052 | -0,0013 | 0,0000 |
| Polansky e Baker com $b=5$ (<i>PB5</i>) | 6,3412 | -0,0373 | 0,0014 |
| Bessegato (<i>Bes</i>) | 6,3177 | -0,0138 | 0,0002 |
| Empírico (<i>Emp</i>) | 6,3319 | -0,0280 | 0,0008 |
| <i>Estatística M</i> | | | |
| Limite teórico (<i>Teo</i>) | 2,0898 | - | - |
| Polansky e Baker com $b=1$ (<i>PB1</i>) | 2,1038 | -0,0140 | 0,0002 |
| Polansky e Baker com $b=5$ (<i>PB5</i>) | 2,0965 | -0,0067 | 0,0000 |
| Bessegato (<i>Bes</i>) | 2,1059 | -0,0161 | 0,0003 |
| Empírico (<i>Emp</i>) | 2,0717 | 0,0181 | 0,0003 |

Legenda: LSC: Limite Superior de Controle; EQ: Erro Quadrático

Para a estatística T^2 de Hotelling (1947) observa-se que o limite superior de controle que mais se aproximou do valor do limite teórico (*Teo*) foi aquele obtido por núcleo estimador com a janela ótima h escolhida de acordo com Polansky e Baker (2000) utilizando o estágio $b=1$ pois o erro (que mede a diferença entre o valor teórico e as estimativas não paramétricas) e o erro quadrático (que é o valor do erro elevado ao quadrado) foram menores para esse caso. Para a estatística M de Hayter e Tsui (1994) os limites superiores de controle, utilizando as técnicas de núcleo estimador estudadas por Polansky e Baker (2000) com $b=5$ forneceu melhor resultado em comparação com o valor teórico dessa estatística.

Para verificar quais os limites de controle desse processo utilizando as componentes principais calculamos os autovalores e autovetores normalizados à partir da decomposição espectral da matriz de covariância $S_{2 \times 2}$ que são:

$$\hat{\lambda} = [5,22 \quad 0,12]' \quad \hat{e}_1 = [0,90 \quad 0,43]' \quad \hat{e}_2 = [-0,43 \quad 0,90]'$$

A primeira componente principal explica 97,7% da variância total do processo (igual a 5,34). Supondo-se que X_1 e X_2 tem conjuntamente uma distribuição normal bi-variada, então a primeira componente principal segue uma distribuição normal com média igual a $\hat{e}_1' \bar{X} = 112,26$ e variância $\hat{\lambda} = 5,22$. O histograma dos valores da primeira componente principal, obtida através da formula (2.25), está apresentado na Figura 3.7. O valor da probabilidade de significância do teste de Anderson-Darling foi igual a 0,27 indicando a normalidade da primeira componente principal. Os limites superior e inferior de controle, à 95% de confiança, são apresentados na Tabela 3.2. Considerando os valores dos erros e da média dos erros quadráticos observa-se que os limites de controle estimados a partir da técnica de núcleo estimador, em que a janela ótima h foi escolhida de acordo com a metodologia de Polansky e Baker (2000) com $b=5$, resultaram em uma melhor aproximação em relação aos limites teóricos, pois utilizando essa metodologia obtemos o menor erro quadrático médio.

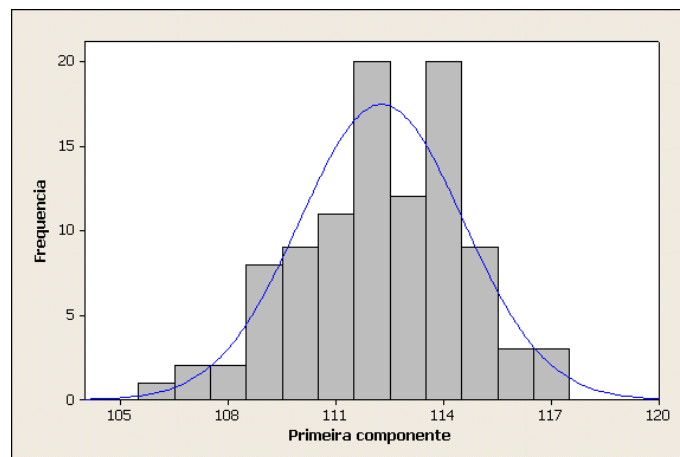


Figura 3.7: Comportamento da primeira componente principal.

Tabela 3.2: Cálculo dos limites de controle da 1ª componente principal ($\alpha = 0,05$).

| Técnica utilizada | LIC | Erro LIC | EQ LIC | LSC | Erro LSC | EQ LSC | Média EQ |
|---|----------|----------|--------|----------|----------|--------|----------|
| Limite teórico (<i>Teo</i>) | 107,7879 | - | - | 116,7417 | - | - | - |
| Polansky e Baker com $b=1$ (<i>PB1</i>) | 107,1198 | 0,6681 | 0,4464 | 116,6717 | 0,0700 | 0,0049 | 0,2256 |
| Polansky e Baker com $b=5$ (<i>PB5</i>) | 107,1243 | 0,6636 | 0,4404 | 116,6661 | 0,0756 | 0,0057 | 0,2230 |
| Bessegato (<i>Bes</i>) | 107,1154 | 0,6725 | 0,4523 | 116,6794 | 0,0623 | 0,0039 | 0,2281 |
| Empírico (<i>Emp</i>) | 106,7006 | 1,0873 | 1,1822 | 116,7174 | 0,0243 | 0,0006 | 0,5914 |

Legenda: **LIC**: Limite Inferior de Controle; **LSC**: Limite Superior de Controle; **EQ LIC**: Erro Quadrático para o Limite Inferior de Controle; **EQ LSC**: Erro Quadrático para o Limite Superior de Controle; **Média EQ**: Soma do Erro Quadrático do Limite Inferior de Controle com o Erro Quadrático do Limite Superior de Controle.

Nos próximos capítulos apresentaremos resultados de simulações extensivas feitas para verificar a adequabilidade das metodologias não-paramétricas para a determinação dos limites de controle.

Capítulo 4 – Estratégia de Execução: Processos Multivariados Normais Não Autocorrelacionados

4.1 – Modelos Simulados

Para verificar a adequação das técnicas estatísticas já apresentadas nessa dissertação vários processos normais e não-normais, autocorrelacionados ou não, foram computacionalmente simulados. Nessa seção serão apresentadas as estratégias de execução apenas para os modelos normais multivariados não autocorrelacionados. Para cada processo foram geradas amostras de tamanho $m=50$, 100 e 200 , e em alguns casos, $m=1000$ sendo que cada amostra tem apenas um único elemento ($n=1$).

Para cada processo simulado e cada m fixo, foram geradas $k=1000$ (réplicas) do processo com o intuito de avaliar os estimadores tratados nessa dissertação. Maiores e menores números de réplicas foram gerados, mas os resultados não diferenciaram significativamente e dessa forma esses resultados não serão apresentados. O número de variáveis consideradas foram $p=2$, 5 , 10 e 20 . Para todos os casos as matrizes de variâncias e covariâncias foram pré-fixadas respeitando faixas de correlação. Para $p=2$ as matrizes de covariâncias foram determinadas de acordo com 3 faixas de correlação: correlação forte ($0,75 \leq \rho \leq 0,9$), correlação média ($0,25 \leq \rho \leq 0,50$) e correlação nula ($\rho=0$). Para $p=5$ e 10 essas matrizes foram determinadas também de acordo com 3 faixas de correlação: correlação forte ($0,75 \leq \rho < 1$), correlação intermediária ($0,50 \leq \rho < 0,75$) e correlação nula ($\rho=0$). Para $p=20$ a matriz de variâncias e covariâncias foi determinada de forma que um conjunto de 5 variáveis tivessem uma correlação forte ($0,75 \leq \rho < 1$), 5 variáveis com correlação intermediária ($0,50 \leq \rho < 0,75$), 5 variáveis com correlação média ($0,25 \leq \rho < 0,50$) e com 5 variáveis com correlação fraca ($0 \leq \rho < 0,25$). As matrizes de variâncias dos processos simulados para cada p foram fixadas e são apresentadas a seguir:

- Processos com correlação forte ($0,75 \leq \rho \leq 0,9$) e $p=2$

$$\Sigma_1 = \begin{bmatrix} 1,00 & 0,90 \\ 0,90 & 1,00 \end{bmatrix} \Sigma_2 = \begin{bmatrix} 1,00 & 1,80 \\ 1,80 & 4,00 \end{bmatrix} \Sigma_3 = \begin{bmatrix} 1,00 & 0,75 \\ 0,75 & 1,00 \end{bmatrix} \Sigma_4 = \begin{bmatrix} 1,00 & 1,50 \\ 1,50 & 4,00 \end{bmatrix}$$

- Processos com correlação intermediária ($0,25 \leq \rho \leq 0,5$) e $p=2$

$$\Sigma_5 = \begin{bmatrix} 1,00 & 0,50 \\ 0,50 & 1,00 \end{bmatrix} \Sigma_6 = \begin{bmatrix} 1,00 & 1,00 \\ 1,00 & 4,00 \end{bmatrix} \Sigma_7 = \begin{bmatrix} 1,00 & 0,25 \\ 0,25 & 1,00 \end{bmatrix} \Sigma_8 = \begin{bmatrix} 1,00 & 0,50 \\ 0,50 & 4,00 \end{bmatrix}$$

- Processos com correlação nula ($\rho = 0$) e $p=2$

$$\Sigma_9 = \begin{bmatrix} 1,00 & 0,00 \\ 0,00 & 1,00 \end{bmatrix} \Sigma_{10} = \begin{bmatrix} 1,00 & 0,00 \\ 0,00 & 4,00 \end{bmatrix}$$

- Processos com correlação forte ($0,75 \leq \rho < 1$) e $p=5$

$$\Sigma_{11} = \begin{bmatrix} 1,000 & 0,805 & 0,826 & 0,754 & 0,797 \\ 0,805 & 1,000 & 0,826 & 0,900 & 0,875 \\ 0,826 & 0,826 & 1,000 & 0,924 & 0,812 \\ 0,754 & 0,900 & 0,924 & 1,000 & 0,853 \\ 0,797 & 0,875 & 0,812 & 0,853 & 1,000 \end{bmatrix}$$

- Processos com correlação intermediária ($0,50 \leq \rho < 0,75$) e $p=5$

$$\Sigma_{12} = \begin{bmatrix} 1,000 & 0,534 & 0,731 & 0,703 & 0,712 \\ 0,534 & 1,000 & 0,730 & 0,585 & 0,512 \\ 0,731 & 0,730 & 1,000 & 0,618 & 0,588 \\ 0,703 & 0,585 & 0,618 & 1,000 & 0,673 \\ 0,712 & 0,512 & 0,588 & 0,673 & 1,000 \end{bmatrix}$$

- Processos com correlação nula ($\rho = 0$) e $p=5$

$$\Sigma_{13} = \begin{bmatrix} 1,000 & 0,000 & 0,000 & 0,000 & 0,000 \\ 0,000 & 1,000 & 0,000 & 0,000 & 0,000 \\ 0,000 & 0,000 & 1,000 & 0,000 & 0,000 \\ 0,000 & 0,000 & 0,000 & 1,000 & 0,000 \\ 0,000 & 0,000 & 0,000 & 0,000 & 1,000 \end{bmatrix}$$

- Processos com correlação forte ($0,75 \leq \rho < 1$) e $p=10$

$$\Sigma_{14} = \begin{bmatrix} 1,000 & 0,853 & 0,759 & 0,771 & 0,869 & 0,785 & 0,829 & 0,752 & 0,893 & 0,809 \\ 0,853 & 1,000 & 0,861 & 0,861 & 0,852 & 0,896 & 0,816 & 0,784 & 0,866 & 0,780 \\ 0,759 & 0,861 & 1,000 & 0,825 & 0,802 & 0,836 & 0,841 & 0,837 & 0,764 & 0,772 \\ 0,771 & 0,861 & 0,825 & 1,000 & 0,800 & 0,887 & 0,796 & 0,772 & 0,860 & 0,845 \\ 0,869 & 0,852 & 0,802 & 0,800 & 1,000 & 0,790 & 0,858 & 0,797 & 0,759 & 0,767 \\ 0,785 & 0,896 & 0,836 & 0,887 & 0,790 & 1,000 & 0,809 & 0,733 & 0,850 & 0,755 \\ 0,829 & 0,816 & 0,841 & 0,796 & 0,858 & 0,809 & 1,000 & 0,849 & 0,854 & 0,862 \\ 0,752 & 0,784 & 0,837 & 0,772 & 0,797 & 0,733 & 0,849 & 1,000 & 0,768 & 0,800 \\ 0,893 & 0,866 & 0,764 & 0,860 & 0,759 & 0,850 & 0,854 & 0,768 & 1,000 & 0,879 \\ 0,809 & 0,780 & 0,772 & 0,745 & 0,767 & 0,755 & 0,862 & 0,800 & 0,879 & 1,000 \end{bmatrix}$$

- Processos com correlação intermediária ($0,50 \leq \rho < 0,75$) e $p=10$

$$\Sigma_{15} = \begin{bmatrix} 1,000 & 0,508 & 0,506 & 0,658 & 0,727 & 0,514 & 0,541 & 0,725 & 0,648 & 0,626 \\ 0,508 & 1,000 & 0,543 & 0,522 & 0,518 & 0,528 & 0,679 & 0,703 & 0,599 & 0,572 \\ 0,506 & 0,543 & 1,000 & 0,646 & 0,558 & 0,717 & 0,585 & 0,676 & 0,692 & 0,585 \\ 0,658 & 0,522 & 0,646 & 1,000 & 0,555 & 0,736 & 0,677 & 0,588 & 0,603 & 0,587 \\ 0,727 & 0,518 & 0,558 & 0,555 & 1,000 & 0,644 & 0,524 & 0,530 & 0,747 & 0,572 \\ 0,514 & 0,528 & 0,717 & 0,736 & 0,644 & 1,000 & 0,576 & 0,584 & 0,572 & 0,607 \\ 0,541 & 0,679 & 0,585 & 0,677 & 0,524 & 0,576 & 1,000 & 0,584 & 0,553 & 0,716 \\ 0,725 & 0,703 & 0,676 & 0,588 & 0,530 & 0,584 & 0,584 & 1,000 & 0,511 & 0,530 \\ 0,648 & 0,599 & 0,692 & 0,603 & 0,747 & 0,572 & 0,553 & 0,511 & 1,000 & 0,522 \\ 0,626 & 0,572 & 0,585 & 0,587 & 0,572 & 0,607 & 0,716 & 0,530 & 0,522 & 1,000 \end{bmatrix}$$

- Processos com correlação nula ($\rho = 0$) e $p=10$

$$\Sigma_{16} = \begin{bmatrix} 1,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 \\ 0,000 & 1,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 \\ 0,000 & 0,000 & 1,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 \\ 0,000 & 0,000 & 0,000 & 1,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 \\ 0,000 & 0,000 & 0,000 & 0,000 & 1,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 \\ 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 1,000 & 0,000 & 0,000 & 0,000 & 0,000 \\ 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 1,000 & 0,000 & 0,000 & 0,000 \\ 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 1,000 & 0,000 & 0,000 \\ 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 1,000 & 0,000 \\ 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 0,000 & 1,000 \end{bmatrix}$$

Processo com correlação misturada e p=20

$$\Sigma_{17} = \begin{bmatrix} 1,00 & 0,20 & 0,07 & 0,16 & 0,02 & 0,01 & 0,04 & 0,03 & 0,03 & 0,02 & 0,00 & -0,01 & -0,01 & 0,01 & -0,01 & 0,04 & 0,03 & 0,04 & 0,05 & 0,06 \\ 0,20 & 1,00 & 0,14 & 0,18 & 0,06 & -0,04 & -0,02 & -0,05 & -0,01 & -0,01 & -0,01 & -0,01 & -0,04 & -0,02 & -0,05 & -0,05 & -0,06 & -0,02 & -0,03 & -0,04 \\ 0,07 & 0,14 & 1,00 & 0,08 & 0,03 & 0,01 & 0,01 & 0,01 & -0,01 & 0,04 & -0,02 & 0,04 & 0,01 & -0,01 & 0,01 & -0,01 & -0,01 & 0,02 & 0,02 & -0,01 \\ 0,16 & 0,18 & 0,08 & 1,00 & 0,03 & 0,01 & -0,01 & -0,04 & -0,01 & -0,01 & -0,04 & -0,04 & 0,00 & -0,02 & -0,03 & -0,03 & -0,02 & -0,01 & -0,03 & -0,04 \\ 0,02 & 0,06 & 0,03 & 0,03 & 1,00 & 0,02 & 0,00 & 0,03 & -0,03 & 0,02 & -0,04 & -0,02 & -0,02 & -0,01 & -0,01 & 0,03 & 0,03 & 0,03 & 0,03 & 0,04 \\ 0,01 & -0,04 & 0,01 & 0,01 & 0,02 & 1,00 & 0,47 & 0,49 & 0,34 & 0,48 & -0,03 & 0,02 & 0,01 & 0,01 & 0,04 & 0,03 & 0,05 & 0,04 & 0,03 & 0,04 \\ 0,04 & -0,02 & 0,01 & -0,01 & 0,00 & 0,47 & 1,00 & 0,42 & 0,45 & 0,41 & 0,02 & 0,05 & 0,03 & 0,01 & 0,04 & -0,01 & 0,01 & 0,01 & 0,00 & 0,03 \\ 0,03 & -0,05 & 0,01 & -0,04 & 0,03 & 0,49 & 0,42 & 1,00 & 0,36 & 0,30 & 0,05 & 0,05 & 0,06 & 0,05 & 0,06 & 0,01 & 0,01 & 0,02 & 0,02 & 0,03 \\ 0,03 & -0,01 & -0,01 & -0,01 & -0,03 & 0,34 & 0,45 & 0,36 & 1,00 & 0,42 & 0,03 & 0,02 & 0,03 & 0,03 & 0,03 & -0,02 & 0,01 & 0,01 & -0,02 & 0,02 \\ 0,02 & -0,01 & 0,04 & -0,01 & 0,02 & 0,48 & 0,41 & 0,30 & 0,42 & 1,00 & 0,03 & 0,06 & 0,05 & 0,06 & 0,08 & 0,06 & 0,09 & 0,07 & 0,06 & 0,09 \\ 0,00 & -0,01 & -0,02 & -0,04 & -0,04 & -0,03 & 0,02 & 0,05 & 0,03 & 0,03 & 1,00 & 0,69 & 0,60 & 0,69 & 0,52 & -0,01 & 0,01 & 0,00 & 0,00 & 0,02 \\ -0,01 & -0,01 & 0,04 & -0,04 & -0,02 & 0,02 & 0,05 & 0,05 & 0,02 & 0,06 & 0,69 & 1,00 & 0,58 & 0,72 & 0,63 & 0,00 & 0,00 & -0,01 & 0,00 & 0,01 \\ -0,01 & -0,04 & 0,01 & 0,00 & -0,02 & 0,01 & 0,03 & 0,06 & 0,03 & 0,05 & 0,60 & 0,58 & 1,00 & 0,71 & 0,67 & -0,02 & 0,01 & -0,01 & 0,00 & 0,01 \\ 0,01 & -0,02 & -0,01 & -0,02 & -0,01 & 0,01 & 0,01 & 0,05 & 0,03 & 0,06 & 0,69 & 0,72 & 0,71 & 1,00 & 0,74 & -0,01 & 0,01 & -0,01 & 0,00 & 0,01 \\ -0,01 & -0,05 & 0,01 & -0,03 & -0,01 & 0,04 & 0,04 & 0,06 & 0,03 & 0,08 & 0,52 & 0,63 & 0,67 & 0,74 & 1,00 & 0,00 & 0,01 & 0,00 & 0,00 & 0,01 \\ 0,04 & -0,05 & -0,01 & -0,03 & 0,03 & 0,03 & -0,01 & 0,01 & -0,02 & 0,06 & -0,01 & 0,00 & -0,02 & -0,01 & 0,00 & 1,00 & 0,92 & 0,83 & 0,90 & 0,85 \\ 0,03 & -0,06 & -0,01 & -0,02 & 0,03 & 0,05 & 0,01 & 0,01 & 0,01 & 0,09 & 0,01 & 0,00 & 0,01 & 0,01 & 0,01 & 0,92 & 1,00 & 0,86 & 0,88 & 0,89 \\ 0,04 & -0,02 & 0,02 & -0,01 & 0,03 & 0,04 & 0,01 & 0,02 & 0,01 & 0,07 & 0,00 & -0,01 & -0,01 & -0,01 & 0,00 & 0,83 & 0,86 & 1,00 & 0,89 & 0,87 \\ 0,05 & -0,03 & 0,02 & -0,03 & 0,03 & 0,03 & 0,00 & 0,02 & -0,02 & 0,06 & 0,00 & 0,00 & 0,00 & 0,00 & 0,00 & 0,90 & 0,88 & 0,89 & 1,00 & 0,83 \\ 0,06 & -0,04 & -0,01 & -0,04 & 0,04 & 0,04 & 0,03 & 0,03 & 0,02 & 0,09 & 0,02 & 0,01 & 0,01 & 0,01 & 0,01 & 0,85 & 0,89 & 0,87 & 0,83 & 1,00 \end{bmatrix}$$

O estudo foi dividido em 2 partes: na primeira abordou-se apenas a análise dos estimadores não-paramétricos em relação aos limites teóricos que são conhecidos para cada processo simulado. Para este estudo o vetor de médias do processo foi sempre o vetor nulo. Diante dos resultados obtidos na primeira fase procedeu-se com a análise dos valores dos valores de *ARL in control* e *out of control* como será explicado na seção 4.3.

4.2 – Determinação dos Limites de Controle e Janela Ótima

Para determinar os limites de controle para os processos com distribuição normal multivariada não autocorrelacionada utilizou-se o valor teórico do limite de controle e os estimadores de núcleo e da função de distribuição empírica para as estatísticas T^2 de Hotelling (1947), primeira componente principal e M de Hayter e Tsui (1994). Os valores amostrais dessas estatísticas foram obtidos sempre com base nos dados de cada amostra simulada utilizando as seguintes fórmulas:

- Estatística T^2 de Hotelling (1947)

$$T_i^2 = (X_i - \bar{X})' S^{-1} (X_i - \bar{X}), \quad i = 1, 2, \dots, m$$

- Primeira componente principal

$$Y_{li} = \hat{e}_l' X_i = \hat{e}_{l1} X_{1i} + \hat{e}_{l2} X_{2i} + \dots + \hat{e}_{lp} X_{pi}, \quad i = 1, 2, \dots, m$$

- Estatística M de Hayter e Tsui (1994)

$$M_i = \max_{1 \leq j \leq p} \left| \frac{X_{ij} - \bar{X}_j}{s_{jj}} \right|, \quad i = 1, 2, \dots, m$$

sendo os limites de controle teóricos obtidos como:

- Estatística T^2 de Hotelling (1947)

$$LSC = \frac{p(m^2 - 1)}{m(m - p)} F_{1-\alpha; p; m-p}$$

em que $F_{1-\alpha; p; m-p}$ é a distribuição de Fisher-Snedecor no percentil $1 - \alpha$ com p e $m-p$ graus de liberdade.

- Primeira componente principal

$$LIC = \hat{e}_l \bar{X} - \Phi_{\alpha/2}^{-1} \sqrt{\hat{\lambda}_l}$$

$$LSC = \hat{e}_l \bar{X} + \Phi_{\alpha/2}^{-1} \sqrt{\hat{\lambda}_l}$$

em que $\Phi_{\alpha/2}^{-1}$ é a função inversa da distribuição normal padronizada no percentil $\alpha / 2$.

- Estatística M de Hayter e Tsui (1994)

$$LSC = C_{R,\alpha}$$

em que a constante $C_{R,\alpha}$ é obtida através do algoritmo descrito no Quadro 2.1 (página 36) utilizando a matriz de variâncias e covariâncias amostral de cada caso simulado.

Na técnica de núcleo estimador utilizou-se os métodos de Polansky e Baker (2000) com $b=1$ e $b=5$ e Bessegato (2001) para determinação da janela ótima h sendo a função núcleo a distribuição normal padronizada. Em todos casos foi considerado um nível de confiança de 95% para determinação das cartas.

Para determinar quais são as melhores técnicas para determinação dos limites de controle utilizou-se três medidas estatísticas: o Erro Médio (EM), o Erro Quadrático Médio (EQM) e o Erro Absoluto Médio (EAM) que são obtidos da seguinte maneira:

$$Erro\ Médio = \frac{\sum_{i=1}^k (LC_{Teórico} - LC_{Estimado})}{k}$$

$$Erro\ Quadrático\ Médio = \frac{\sum_{i=1}^k (LC_{Teórico} - LC_{Estimado})^2}{k}$$

$$Erro\ Absoluto\ Médio = \frac{\sum_{i=1}^k |LC_{Teórico} - LC_{Estimado}|}{k}$$

em que $LC_{Teórico}$ é o limite de controle obtido à partir da função de distribuição teórica da estatística em consideração, o $LC_{Estimado}$ é o limite de controle estimado através das técnicas não-paramétricas e k é a quantidade de réplicas (amostras) dos processos simulados. Quanto mais próximos de zero observamos os valores desses erros, melhor é considerada a técnica de determinação não-paramétrica correspondente.

4.3 – Determinação dos ARL

A determinação dos ARL 's para cada m e p fixo e para cada carta de controle foi feita da seguinte forma:

- Num primeiro momento definiu-se os limites de controle a partir da amostra aleatória de tamanho m gerada do processo sob consideração.

- Posteriormente procedeu-se com o cálculo do *ARL in control* do seguinte modo: Estabelecido os limites de controle novas observações do processo sob controle foram geradas uma a uma até que ocorresse a primeira amostra fora dos limites de controle. O número de observações geradas até a ocorrência deste evento foi a estimativa do *ARL in control*.
- Para determinar o *ARL out of control* procedeu-se do seguinte modo: Estabelecido os limites de controle observações fora de controle foram geradas uma a uma do processo modificado de acordo com um “choque” dado no vetor de médias do processo, até que ocorresse a primeira amostra fora dos limites de controle. O número de observações geradas até a ocorrência desse evento foi a estimativa do *ARL out of control*.

As amostras utilizadas para o cálculo dos *ARL's* são geradas de acordo com a distribuição utilizada na geração das amostras utilizadas no cálculo dos limites de controle. Para a determinação do *ARL out of control* o “choque” dado no vetor de médias definido no parágrafo anterior faz com que a média considerada na geração da amostra passe a ser chamada de μ_{choque} seguindo a mesma matriz de variâncias e covariâncias teóricas do processo. A intensidade dos “choques” aplicados ao vetor de médias do processo e a forma como eles foram implementados são apresentados seguir:

- Processos com $p=2$ variáveis

- As médias das p variáveis mudam ordenadamente na mesma escala

$$\mu_{choque} = \mu_{processo} + choque \quad (choque = 0,5;1;2;3;4)$$

- A média da 1ª variável permanece constante e a média da 2ª variável muda ordenadamente

$$\begin{aligned} \mu_{1,choque} &= \mu_{1,processo} \\ \mu_{2,choque} &= \mu_{2,processo} + choque \quad (choque = 0,5;1;2;3;4) \end{aligned}$$

- A média da 1ª variável muda em 0,5 e a média da 2ª variável muda ordenadamente

$$\begin{aligned} \mu_{1,choque} &= \mu_{1,processo} + 0,5 \\ \mu_{2,choque} &= \mu_{2,processo} + choque \quad (choque = 0,5;1;2;3;4) \end{aligned}$$

- Processos com $p=5$ variáveis

- As médias das p variáveis mudam ordenadamente na mesma escala

$$\mu_{choque} = \mu_{processo} + choque \quad (choque = 0,5;1;2;3;4)$$

- As médias de quatro primeiras variáveis permanecem constantes e a média da 5ª variável muda ordenadamente

$$\begin{aligned}\mu_{i,choque} &= \mu_{i,processo} \quad i = 1,2,3,4 \\ \mu_{5,choque} &= \mu_{5,processo} + choque \quad (choque = 0,5;1;2;3;4)\end{aligned}$$

- A média da 1ª variável muda em 0,5, a média da 2ª variável muda em 3 e a média das outras variáveis mudam em 1

$$\begin{aligned}\mu_{1,choque} &= \mu_{1,processo} + 0,5 \\ \mu_{2,choque} &= \mu_{2,processo} + 3 \\ \mu_{i,choque} &= \mu_{i,processo} + 1 \quad (i = 3,4,5)\end{aligned}$$

- Processos com p=10 variáveis

- As médias das p variáveis mudam ordenadamente na mesma escala

$$\mu_{choque} = \mu_{processo} + choque \quad (choque = 0,5;1;2;3;4)$$

- As médias de nove primeiras variáveis permanecem constantes e a média da 10ª variável muda ordenadamente

$$\begin{aligned}\mu_{i,choque} &= \mu_{i,processo} \quad i = 1,2, \dots, 9 \\ \mu_{10,choque} &= \mu_{10,processo} + choque \quad (choque = 0,5;1;2;3;4)\end{aligned}$$

- A média da 1ª variável muda em 0,5, a média da 2ª variável muda em 3, a média da 3ª, 4ª e 5ª variáveis mudam em 1 e as demais permanecem constantes

$$\begin{aligned}\mu_{1,choque} &= \mu_{1,processo} + 0,5 \\ \mu_{2,choque} &= \mu_{2,processo} + 3 \\ \mu_{i,choque} &= \mu_{i,processo} + 1 \quad (i = 3,4,5) \\ \mu_{i,choque} &= \mu_{i,processo} \quad (i = 6,7, \dots, 10)\end{aligned}$$

- Processos com p=20 variáveis

- As médias das p variáveis mudam ordenadamente na mesma escala

$$\mu_{choque} = \mu_{processo} + choque \quad (choque = 0,5;1;2;3;4)$$

- As médias das dezenove primeiras variáveis permanecem constantes e a média da 20ª variável muda ordenadamente

$$\begin{aligned}\mu_{i,choque} &= \mu_{i,processo} \quad i = 1, 2, \dots, 19 \\ \mu_{20,choque} &= \mu_{20,processo} + choque \quad (choque = 0,5;1;2;3;4)\end{aligned}$$

- A média da 1ª variável muda em 0,5, a média da 2ª variável muda em 3, a média da 3ª, 4ª e 5ª variáveis mudam em 1 e as demais permanecem constantes

$$\begin{aligned}\mu_{1,choque} &= \mu_{1,processo} + 0,5 \\ \mu_{2,choque} &= \mu_{2,processo} + 3 \\ \mu_{i,choque} &= \mu_{i,processo} + 1 \quad (i = 3, 4, 5) \\ \mu_{i,choque} &= \mu_{i,processo} \quad (i = 6, 7, \dots, 19)\end{aligned}$$

A análise conjunta do *ARL in control* e *ARL out of control* indica a qualidade da carta de controle uma vez que maiores valores de *ARL in control* e menores valores de *ARL out of control* indicam que a carta de controle é eficiente para detectar mudanças significativas no processo e, além disso, geram poucos alarmes falsos.

4.4 – Aspectos Computacionais

Como dito anteriormente o método de Polansky e Baker (2000) depende de um parâmetro inicial b ($b \geq 1$) que é utilizado para determinar o estágio do estimador que fornece o valor da janela ótima h . Como não possuíamos valores pré-fixados de b , dois valores diferentes ($b=1$ e $b=5$) foram escolhidos para estimar a janela h . A escolha desses valores foi feita de forma que os cálculos do algoritmo proposto por Polansky e Baker (2000) não fossem computacionalmente muito complexos. Por exemplo, ao trabalhar com um valor de $b=5$ o estimador multi-estágio necessita da décima derivada da função utilizada como núcleo (no caso dessa dissertação a distribuição normal) para que o método seja aplicado.

A metodologia de determinação da janela ótima h estudada por Bessegato (2001) já havia sido computacionalmente implementada em sua dissertação, tanto na linguagem de programação C++ (para simulações mais extensas), quanto no *software S-plus for Windows* sendo que esse último foi transcrito para a linguagem de programação do *software R for Windows* nessa dissertação. Os programas para determinação da janela ótima h são apresentados no Anexo 3 dessa dissertação.

Esses programas são produto dessa dissertação e esperamos disponibilizá-los na internet para facilitar o trabalho de usuários de métodos de núcleo estimador no que se refere à obtenção da janela ótima h . Os programas são gerais e podem ser utilizados em outras situações diferentes das apresentadas nessa dissertação.

A determinação da constante $C_{R,\alpha}$ de Hayter e Tsui (1994), chamada nessa dissertação de limite teórico, foi feita para cada carta de controle gerada do processo de interesse, ou seja, para cada uma das $k=1000$ réplicas de um processo com p e Σ fixos, têm-se

uma estimativa do valor da constante $C_{R,\alpha}$ com base no valor da matriz de correlação amostral \hat{R} o que nos gera ao final da simulação 1000 valores diferentes da constante $C_{R,\alpha}$. Os valores dessa constante foram sumarizados e são apresentados nas tabelas contidas nos capítulos seguintes. Essa constante foi calculada de acordo com o procedimento descrito no Quadro 2.1 (página 36) usando um valor de $N=10000$ seguindo a determinação de Mingoti e Glória (2005b).

Capítulo 5 – Análise de Resultados: Processos Multivariados Normais Não Autocorrelacionados

Neste capítulo apresentamos os principais resultados obtidos nas simulações realizadas considerando-se processos normais p -variados não autocorrelacionados. Inicialmente se fará uma análise da qualidade dos métodos de estimação não-paramétrica para determinação dos limites de controle e posteriormente uma análise das cartas de controle em termos da avaliação do *ARL in control* e *out of control*.

5.1 – Comparação dos Métodos para Determinação dos Limites de Controle

Para comparação da metodologia no que se refere a determinação dos limites de controle e para avaliação da qualidade dos mesmos levou-se em consideração os valores do Erro Médio (*EM*), Erro Absoluto Médio (*EAM*) e Erro Quadrático Médio (*EQM*). Vamos apresentar apenas os resultados para o *EM* e *EQM* pois as soluções com *EAM* são similares. Tabelas e gráficos *boxplot* serão utilizados para apresentar os resultados gerais para cada combinação (p , m e Σ) onde p é o número de variáveis, m o tamanho da amostra e Σ a matriz de variâncias e covariâncias teórica dos processos simulados. Serão apresentados apenas os resultados para $m=50$ e 200 já que os obtidos para $m=100$ são semelhantes aos de $m=200$. Além disso, para o caso em que Σ é uma matriz diagonal, ou seja, as p variáveis são não correlacionadas não serão mostrados as estimativas das componentes principais visto que esta técnica não tem utilidade prática em casos desse tipo pois cada componente principal é uma das variáveis originais (Mingoti, 2005). Em um caso em que $p=2$ e no caso em que temos $p=20$ variáveis foram simuladas amostras de tamanho $m=1000$ para verificar o comportamento do estimador quando possuímos amostras de tamanho elevado.

5.1.1 – Processos bivariados ($p=2$)

Os erros médio e quadrático médio dos limites de controle para os processos bivariados com em que a correlação das variáveis do processo estão entre $0,75 \leq \rho \leq 0,90$, $0,25 \leq \rho \leq 0,50$ e $\rho = 0$ para $m=50$ e 200 são apresentados na Tabela 5.1. Nas Figuras 5.1 a 5.3 apresenta-se gráficos de *boxplot* de cada um desses erros. Para a situação em que $\rho=0,9$ apresenta-se também os valores obtidos para amostras de tamanho $m=1000$ considerando-se $k=1000$ réplicas.

Tabela 5.1: Erros médio e Quadrático médio dos limites de controle obtidos à 95% de confiança para processos com $m=50, 200$ e 1000 e $p=2$ variáveis.

| Tipo de Erro | Técnica | $m=50$ e $0,75 \leq \rho \leq 0,9$ | | | | $m=200$ e $0,75 \leq \rho \leq 0,9$ | | | |
|-----------------------|---------|-------------------------------------|---------|---------|---------|-------------------------------------|---------|--------------------|---------|
| | | T2 | LSC-CP | LIC-CP | M | T2 | LSC-CP | LIC-CP | M |
| Erro Médio | Pb=1 | 0,8908 | -0,0922 | 0,0842 | -0,0126 | 0,2585 | -0,0322 | 0,0346 | -0,0060 |
| | Pb=5 | 0,9299 | -0,0750 | 0,0672 | -0,0029 | 0,2716 | -0,0288 | 0,0312 | -0,0027 |
| | Bes | 0,9068 | -0,1142 | 0,1054 | -0,0178 | 0,2697 | -0,0368 | 0,0392 | -0,0046 |
| | Emp | 0,9773 | 0,2702 | 0,3217 | 0,0402 | 0,3649 | 0,1557 | -0,0160 | 0,0329 |
| Erro Quadrático Médio | Pb=1 | 1,1820 | 0,2900 | 0,2865 | 0,0126 | 0,1976 | 0,0694 | 0,0731 | 0,0045 |
| | Pb=5 | 1,2939 | 0,2901 | 0,2874 | 0,0135 | 0,2135 | 0,0697 | 0,0734 | 0,0047 |
| | Bes | 1,2182 | 0,2993 | 0,2945 | 0,0118 | 0,2106 | 0,0694 | 0,0730 | 0,0045 |
| | Emp | 1,4808 | 0,4012 | 0,5477 | 0,0264 | 0,2818 | 0,1129 | 0,1054 | 0,0079 |
| Tipo de Erro | Técnica | $m=1000$ e $\rho=0,90$ | | | | $m=50$ e $0,25 \leq \rho \leq 0,5$ | | | |
| | | T2 | LSC-CP | LIC-CP | M | T2 | LSC-CP | LIC-CP | M |
| Erro Médio | Pb=1 | 0,1097 | 0,0339 | -0,0302 | -0,0017 | 0,8836 | -0,0899 | 0,0938 | -0,0072 |
| | Pb=5 | 0,1139 | 0,0341 | -0,0304 | -0,0007 | 0,9224 | -0,0743 | 0,0789 | 0,0006 |
| | Bes | 0,1138 | 0,0336 | -0,0299 | -0,0010 | 0,9001 | -0,1094 | 0,1133 | -0,0147 |
| | Emp | 0,1337 | 0,0739 | -0,0524 | 0,0075 | 0,9687 | 0,2425 | 0,3120 | 0,0557 |
| Erro Quadrático Médio | Pb=1 | 0,0423 | 0,0103 | 0,0096 | 0,0013 | 1,1826 | 0,2447 | 0,2539 | 0,0145 |
| | Pb=5 | 0,0444 | 0,0104 | 0,0096 | 0,0013 | 1,2934 | 0,2443 | 0,2539 | 0,0150 |
| | Bes | 0,0443 | 0,0103 | 0,0096 | 0,0013 | 1,2214 | 0,2527 | 0,2615 | 0,0139 |
| | Emp | 0,0502 | 0,0169 | 0,0137 | 0,0017 | 1,4765 | 0,3353 | 0,4913 | 0,0285 |
| Tipo de Erro | Técnica | $m=200$ e $0,25 \leq \rho \leq 0,5$ | | | | $m=50$ e $\rho=0$ | | $m=200$ e $\rho=0$ | |
| | | T2 | LSC-CP | LIC-CP | M | T2 | M | T2 | M |
| Erro Médio | Pb=1 | 0,2545 | -0,0322 | 0,0275 | -0,0014 | 0,8826 | -0,0021 | 0,2726 | 0,0015 |
| | Pb=5 | 0,2670 | -0,0285 | 0,0240 | 0,0006 | 0,9191 | 0,0051 | 0,2876 | 0,0035 |
| | Bes | 0,2650 | -0,0364 | 0,0317 | -0,0016 | 0,8970 | -0,0105 | 0,2853 | 0,0010 |
| | Emp | 0,3616 | 0,1387 | -0,0178 | 0,0390 | 0,9684 | 0,0637 | 0,3831 | 0,0441 |
| Erro Quadrático Médio | Pb=1 | 0,1974 | 0,0614 | 0,0608 | 0,0047 | 1,1696 | 0,0138 | 0,2068 | 0,0047 |
| | Pb=5 | 0,2135 | 0,0618 | 0,0609 | 0,0048 | 1,2784 | 0,0144 | 0,2251 | 0,0048 |
| | Bes | 0,2103 | 0,0612 | 0,0607 | 0,0046 | 1,2033 | 0,0132 | 0,2218 | 0,0046 |
| | Emp | 0,2840 | 0,0989 | 0,0866 | 0,0084 | 1,4738 | 0,0283 | 0,3008 | 0,0088 |

Legenda: T2: Limite superior de controle da estatística T^2 de Hotelling (1947); LSC-CP: Limite superior de controle da 1ª componente principal; LIC-CP: Limite inferior de controle da 1ª componente principal; M: Limite superior de controle da estatística M de Hayter e Tsui (1994); Pb1: Limite de controle obtido pela a técnica de Polansky e Baker (2000) com $b=1$; Pb5: Limite de controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; Bes: Limite de controle obtido pela a técnica de Bessegato (2001); Emp: Limite de controle obtido pela função da função de distribuição empírica.

Observa-se na Tabela 5.1 que para m fixo as mudanças nas faixas de correlação do processo não afetam significativamente os valores do erro médio e do erro quadrático médio independentemente da técnica não-paramétrica utilizada.

Embora não apresentado na Tabela 5.1 um estudo foi feito para avaliar a influência das variâncias das variáveis na determinação dos limites de controle via técnicas não-paramétricas. Esse estudo mostrou que a mudança na variância em uma das variáveis, mantendo a correlação entre as variáveis constante, exerceu pequena influência nos limites de controle obtidos de forma não-paramétrica quando analisamos os valores das estatísticas T^2 de Hotelling (1947) e a estatística M de Hayter e Tsui (1994). Já para a primeira componente principal o aumento das variâncias fez com que em média as estimativas dos limites de controle fossem maiores do que aquelas obtidas para iguais a 1.

Os limites de controle teóricos da estatística T^2 de Hotelling (1947) possuem sempre valores fixos para cada tamanho de amostra m , uma vez que esses limites dependem apenas do valor de m e do valor de p para um α fixo.

É possível observar que a técnica de núcleo estimador quando utilizamos o método de Polansky e Baker (2000) com $b=5$ e Bessegato (2000) para determinar o valor da janela ótima h , fornece menores erros médios e erros quadráticos médios, ou seja, fornecem estimativas mais próximas dos limites de controle teóricos. Em alguns casos os valores dos erros para o método de Polansky e Baker com $b=1$ forneceram valores semelhantes aos do método com $b=5$ e em outros casos semelhantes aos valores do método estudado por Bessegato (2000). Observamos que em todos os casos, inclusive naquele em que $m=1000$ a estimativa via função de distribuição empírica forneceu maiores valores dos erros médio e quadrático médio indicando que essa técnica não foi adequada para obtenção dos limites de controle do processo.

Entre as técnicas analisadas observa-se que independente da faixa de correlação, do valor de m , e da variância das variáveis que fazem parte do processo, os valores dos erros obtidos para a estatística M de Hayter e Tsui (1994) foram menores (e mais próximos de zero) que os valores observados para as outras estatísticas independente da técnica utilizada para estimação dos limites de controle. Isso mostra que o uso da estatística M é o mais recomendado para processos bivariados. É interessante observar que mesmo nos casos em que o tamanho de amostra é pequeno ($m=50$) a técnica de núcleo-estimador forneceu boas estimativas para os limites de controle quando a estatística M foi utilizada fato que pode ser observado também nos gráficos de *boxplot* pois a variabilidade dos erros é bem menor para essa estatística em comparação com as outras. Para todos os valores de m os gráficos *boxplot* dos erros da estatística T^2 de Hotelling (1947) apresentam sempre maiores valores numéricos (e mais longe de zero, que seria o valor ideal) e uma maior variabilidade nos limites de controle obtidos. Em geral, a primeira componente principal fornece valores intermediários dos erros sendo os valores localizados sempre entre os valores dos erros da estatística T^2 de Hotelling (1947) e os da estatística M de Hayter e Tsui (1994).

É possível observar visualmente que os erros obtidos quando a técnica da função de distribuição empírica é utilizada, em geral, fornece uma maior variabilidade. Como esperado quanto maior o tamanho m da amostra menor é o valor do erro médio e quadrático médio de todas as técnicas avaliadas.

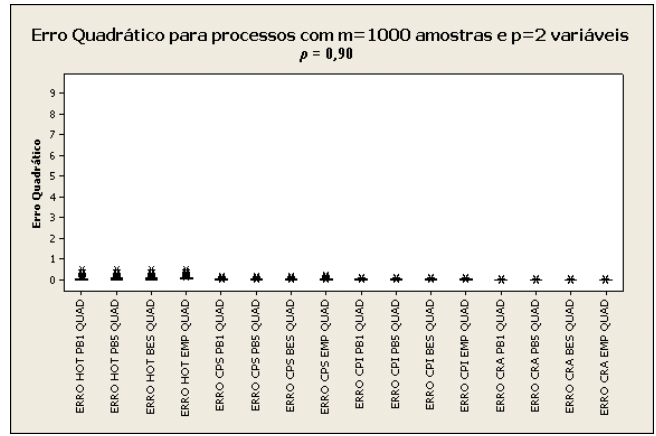
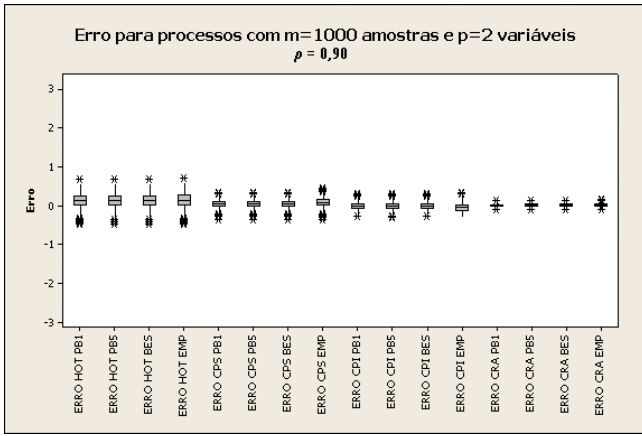
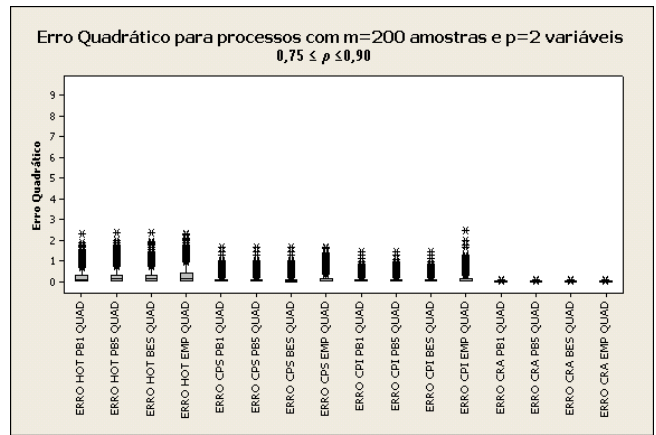
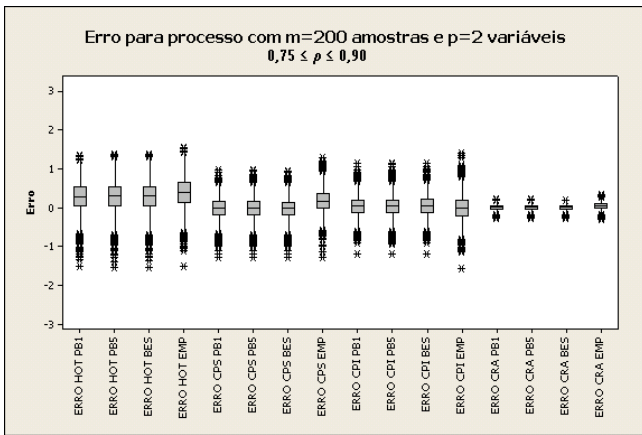
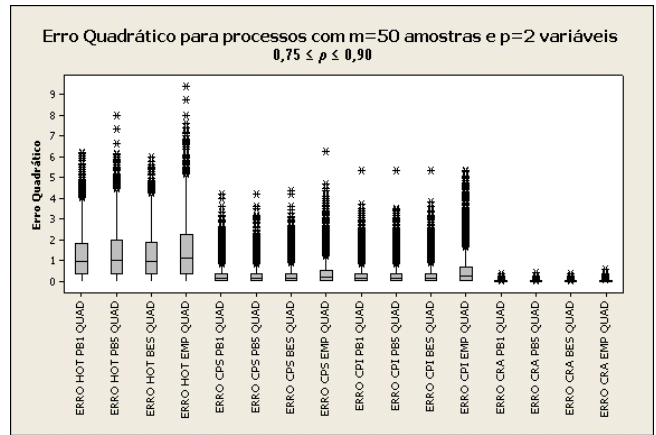
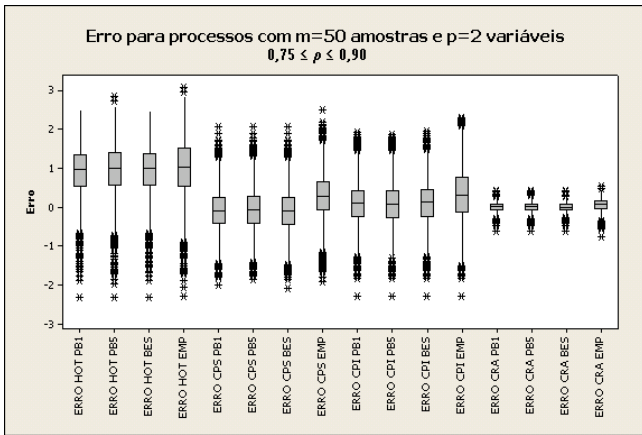


Figura 5.1: Comparação do Erro e Erro Quadrático dos limites de controle obtidos à 95% de confiança para processos com correlação entre 0,75 e 0,90, $m=50, 200$ e 1000 e $p=2$.

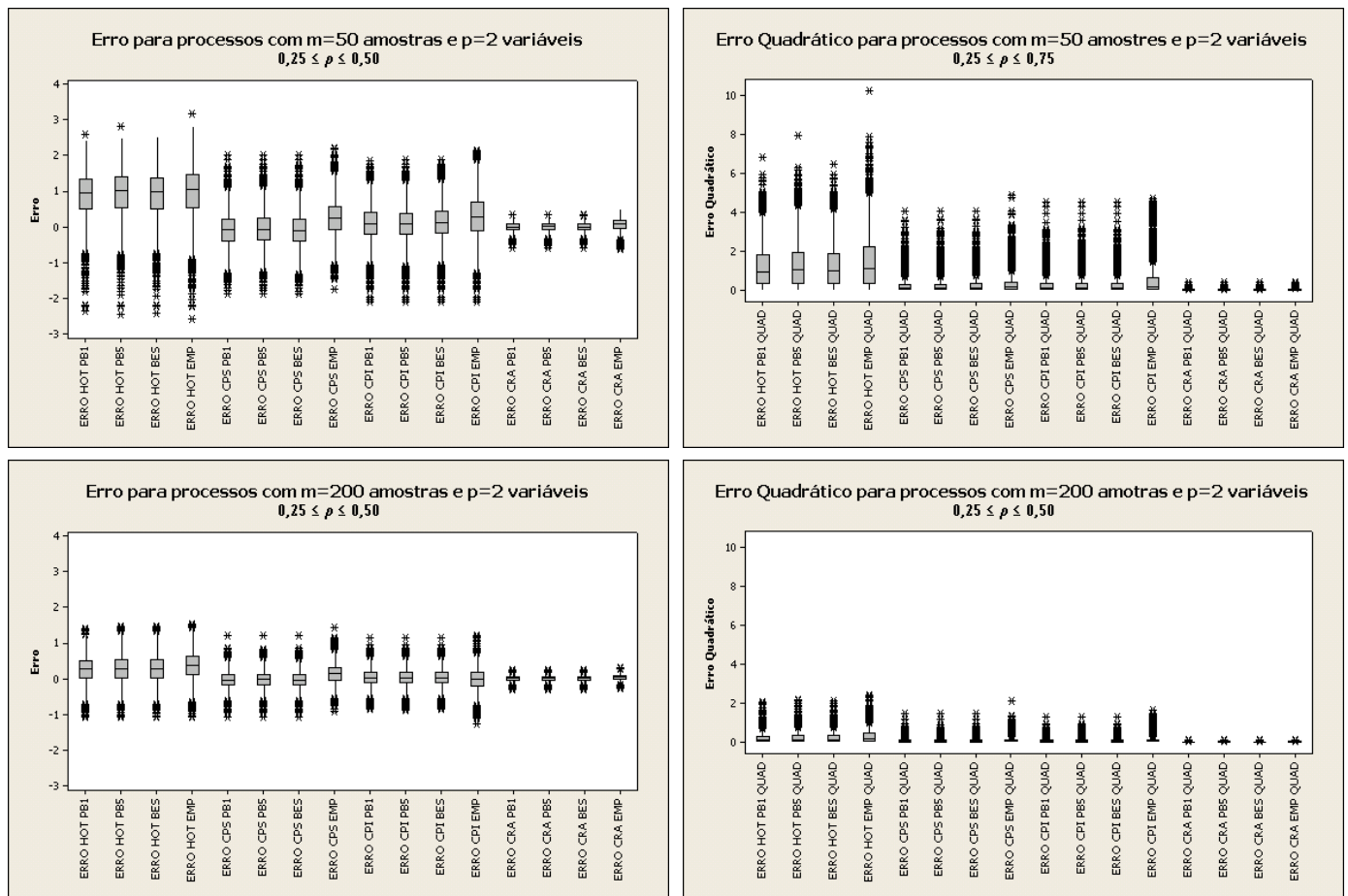


Figura 5.2: Comparação do Erro e Erro Quadrático dos limites de controle obtidos à 95% de confiança para processos com correlação entre 0,25 e 0,50, $m=50$ e 200 e $p=2$.

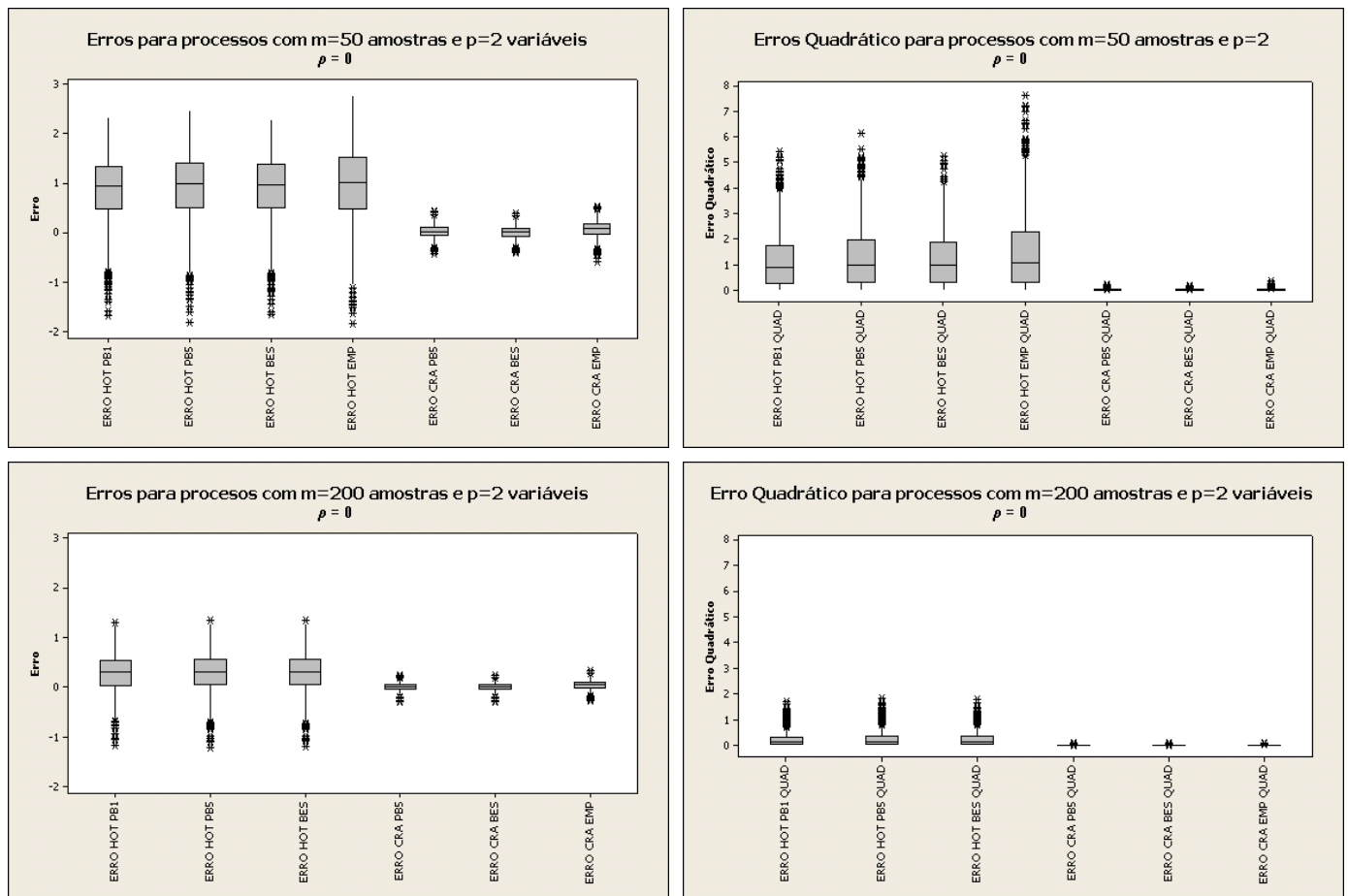


Figura 5.3: Comparação do Erro e Erro Quadrático dos limites de controle obtidos à 95% de confiança para processos com correlação igual a zero, $m=50$ e 200 e $p=2$.

5.1.2 – Processos multivariados com $p=5$ variáveis

Na Tabela 5.2 e Figuras 5.4 a 5.6 são apresentados os valores dos erros médios e dos erros quadráticos médios dos limites de controle para os processos multivariados com $p=5$ variáveis em que essas variáveis possuem faixa de correlação $0,75 \leq \rho < 1$, $0,50 \leq \rho < 0,75$ e $\rho = 0$ para $m=50$ e 200 .

Observa-se na Tabela 5.2 que as técnicas não-paramétricas para obtenção dos limites de controle para m fixo, forneceram resultados próximos tanto para os erros médios quanto para os erros quadráticos médios, independentemente da mudança nas faixas de correlação, mostrando que a correlação das variáveis não exerce, de forma significativa, efeito sobre as estimativas dos limites de controle obtidos.

Em geral observa-se que a técnica de núcleo estimador, quando os métodos de Polansky e Baker (2000) com $b=5$ e de Bessegato (2000) são utilizados, fornece menores erros médios e erros quadráticos médios indicando que esses dois métodos são os mais adequados para a obtenção da janela ótima quando se constrói as cartas de controle para o

processo utilizando-se técnicas não-paramétricas. Os valores dos erros para o método de Polansky e Baker com $b=1$ em alguns casos são até melhores frente às outras técnicas, porém na maioria das vezes, semelhantes com os obtidos usando $b=5$ ou com aqueles obtidos usando o método de Bessegato. Observa-se também que a função de distribuição empírica, sempre forneceu maiores valores dos erros médios e erros quadráticos médios, mostrando ser uma técnica pouco adequada na determinação dos limites de controle. A função de distribuição empírica apresenta, na maioria das vezes, erros médio e quadrático maiores que os fornecidos pelas técnicas de núcleo estimador, fato que ocorria também nos processos bivariados.

Tabela 5.2: Erros médio e Quadrático médio dos limites de controle obtidos à 95% de confiança para processos com $m=50$ e 200 e $p=5$ variáveis.

| Tipo de Erro | Técnica | $m=50$ e $0,75 < \rho \leq 1$ | | | | $m=200$ e $0,75 < \rho \leq 1$ | | | |
|-----------------------|---------|----------------------------------|---------|----------------------|---------|-----------------------------------|---------|--------|---------|
| | | T2 | LSC-CP | LIC-CP | M | T2 | LSC-CP | LIC-CP | M |
| Erro Médio | Pb=1 | 2,9863 | -0,1775 | 0,1663 | -0,0101 | 0,7754 | -0,0907 | 0,0966 | -0,0095 |
| | Pb=5 | 3,0276 | -0,1583 | 0,1465 | -0,0014 | 0,7889 | -0,0865 | 0,0929 | -0,0065 |
| | Bes | 2,9662 | -0,2016 | 0,1913 | -0,0150 | 0,7522 | -0,1152 | 0,1202 | -0,0151 |
| | Emp | 3,2376 | 0,2474 | 0,5099 | 0,0442 | 0,9588 | 0,1371 | 0,0466 | 0,0284 |
| Erro Quadrático Médio | Pb=1 | 9,5020 | 0,4683 | 0,4554 | 0,0137 | 0,8271 | 0,1187 | 0,1147 | 0,0051 |
| | Pb=5 | 9,7650 | 0,4623 | 0,4496 | 0,0144 | 0,8534 | 0,1184 | 0,1145 | 0,0052 |
| | Bes | 9,3570 | 0,4843 | 0,4713 | 0,0129 | 0,7824 | 0,1224 | 0,1182 | 0,0050 |
| | Emp | 11,3700 | 0,5423 | 1,0702 | 0,0267 | 1,2153 | 0,1555 | 0,1508 | 0,0081 |
| Tipo de Erro | Técnica | $m=50$ e $0,50 < \rho \leq 0,75$ | | | | $m=200$ e $0,50 < \rho \leq 0,75$ | | | |
| | | T2 | LSC-CP | LIC-CP | M | T2 | LSC-CP | LIC-CP | M |
| Erro Médio | Pb=1 | 3,0158 | -0,1706 | 0,1631 | 0,0058 | 0,7932 | -0,0785 | 0,0612 | -0,0005 |
| | Pb=5 | 3,0607 | -0,1533 | 0,1454 | 0,0139 | 0,8072 | -0,0747 | 0,0571 | 0,0015 |
| | Bes | 2,7453 | -0,3060 | 0,2901 | -0,0491 | 0,7697 | -0,1005 | 0,0835 | -0,0058 |
| | Emp | 3,2833 | 0,2398 | 0,4645 | 0,0622 | 0,9815 | 0,1236 | 0,0063 | 0,0391 |
| Erro Quadrático Médio | Pb=1 | 9,6110 | 0,3553 | 0,3796 | 0,0144 | 0,8210 | 0,0864 | 0,0886 | 0,0050 |
| | Pb=5 | 9,9000 | 0,3508 | 0,3743 | 0,0153 | 0,8479 | 0,0862 | 0,0888 | 0,0051 |
| | Bes | 8,0280 | 0,4665 | 0,4951 | 0,0153 | 0,7758 | 0,0893 | 0,0905 | 0,0048 |
| | Emp | 11,5630 | 0,4368 | 0,8603 | 0,0302 | 1,2088 | 0,1178 | 0,1209 | 0,0082 |
| Tipo de Erro | Técnica | $m=50$ e $\rho = 0$ | | $m=200$ e $\rho = 0$ | | | | | |
| | | T2 | M | T2 | M | | | | |
| Erro Médio | Pb=1 | 2,9766 | 0,0277 | 0,7886 | 0,0126 | | | | |
| | Pb=5 | 3,0179 | 0,0336 | 0,8031 | 0,0138 | | | | |
| | Bes | 2,7127 | -0,0255 | 0,7648 | 0,0080 | | | | |
| | Emp | 3,2378 | 0,0915 | 0,9841 | 0,0518 | | | | |
| Erro Quadrático Médio | Pb=1 | 9,4130 | 0,0198 | 0,8318 | 0,0060 | | | | |
| | Pb=5 | 9,6760 | 0,0207 | 0,8587 | 0,0061 | | | | |
| | Bes | 7,8840 | 0,0193 | 0,7867 | 0,0058 | | | | |
| | Emp | 11,3580 | 0,0357 | 1,2295 | 0,0102 | | | | |

Legenda: T2: Limite superior de controle da estatística T^2 de Hotelling (1947); LSC-CP: Limite superior de controle da 1ª componente principal; LIC-CP: Limite inferior de controle da 1ª componente principal; M: Limite superior de controle da estatística M de Hayter e Tsui (1994); Pb1: Limite de controle obtido pela a técnica de Polansky e Baker (2000) com $b=1$; Pb5: Limite de controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; Bes: Limite de controle obtido pela a técnica de Bessegato (2001); Emp: Limite de controle obtido pela função da função de distribuição empírica.

Entre as técnicas analisadas, independente da faixa de correlação e do tamanho da amostra simulada, a estatística M de Hayter e Tsui (1994) forneceu sempre menores valores numéricos de erros médio e quadrático médio mostrando ser a mais adequada na análise dos processos com $p=5$ variáveis. Mesmo para amostras de tamanho pequeno a técnica de núcleo

estimador forneceu boa estimativa para os limites de controle quando a estatística M foi utilizada e em alguns casos com precisão de duas casas decimais de diferença entre os valores teóricos e os valores estimados. A estatística T^2 de Hotelling (1947) forneceu os maiores valores numéricos dos erros médio e quadrático médio enquanto a primeira componente principal forneceu valores intermediários em relação às outras duas técnicas.

Os gráficos de *boxplot* dos erros e erros quadráticos para a estatística M de Hayter e Tsui (1994) apresentam erros com mediana mais próxima a zero e menor variabilidade mostrando que as estimativas dos limites de controle foram realizadas de forma satisfatória. Os gráficos de *boxplot* da estatística T^2 de Hotelling (1947) mostram erros com mediana mais distantes de zero e maior variabilidade. Já a primeira componente principal forneceu erros sempre com valores intermediários em relação às outras técnicas.

Dentre as técnicas não-paramétricas utilizadas para a estimação dos limites de controle, a técnica da função de distribuição empírica apresenta sempre maior mediana e maior amplitude (quando comparado às estimativas realizadas pelas técnicas de núcleo-estimador) mostrando que mesmo para amostras grandes ($m=200$) devemos preferir a técnica de núcleos estimadores para a obtenção dos limites de controle.

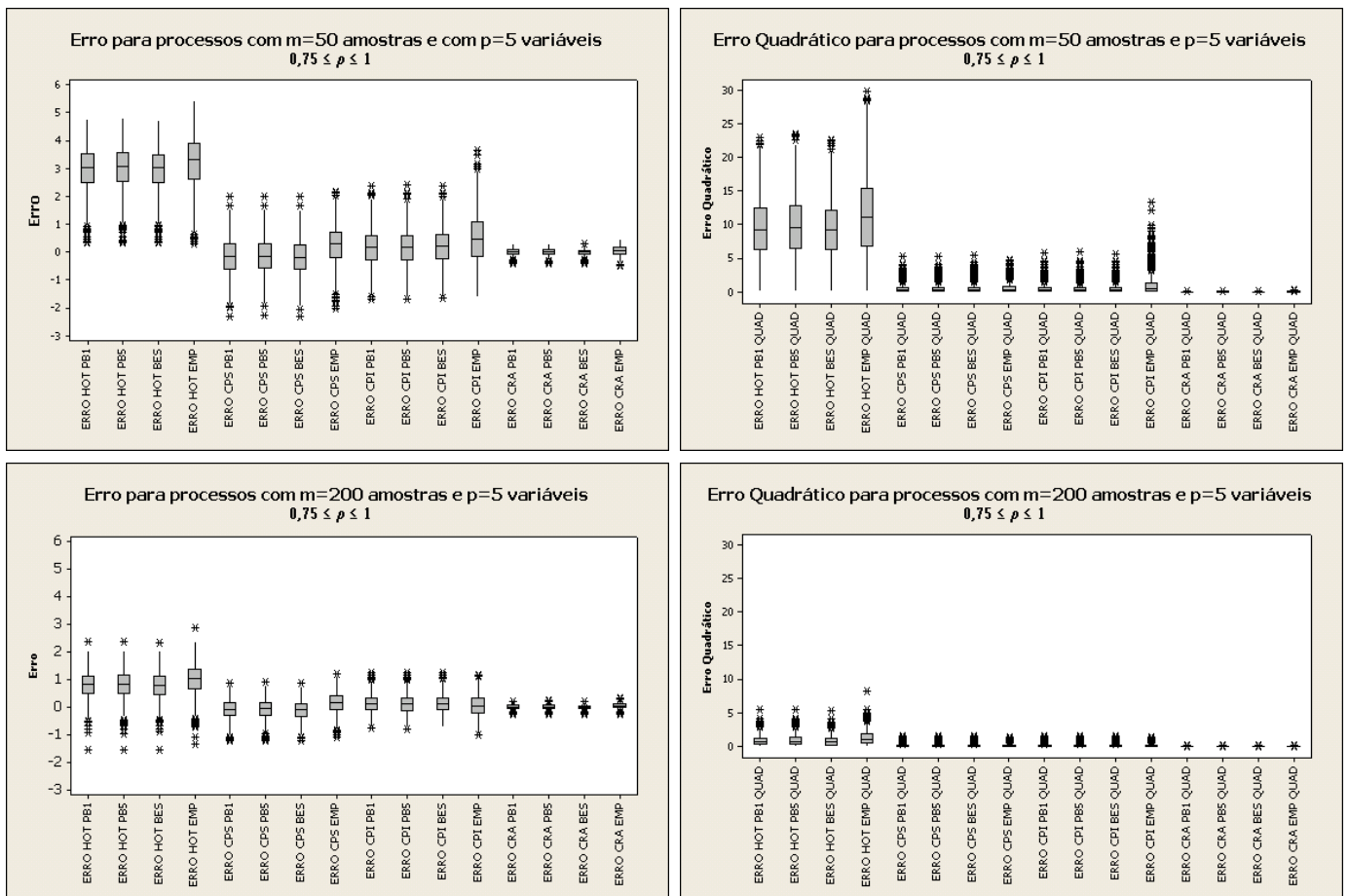


Figura 5.4: Comparação do Erro e Erro Quadrático dos limites de controle obtidos à 95% de confiança para processos com correlação entre 0,75 e 1, $m=50$ e 200 e $p=5$.

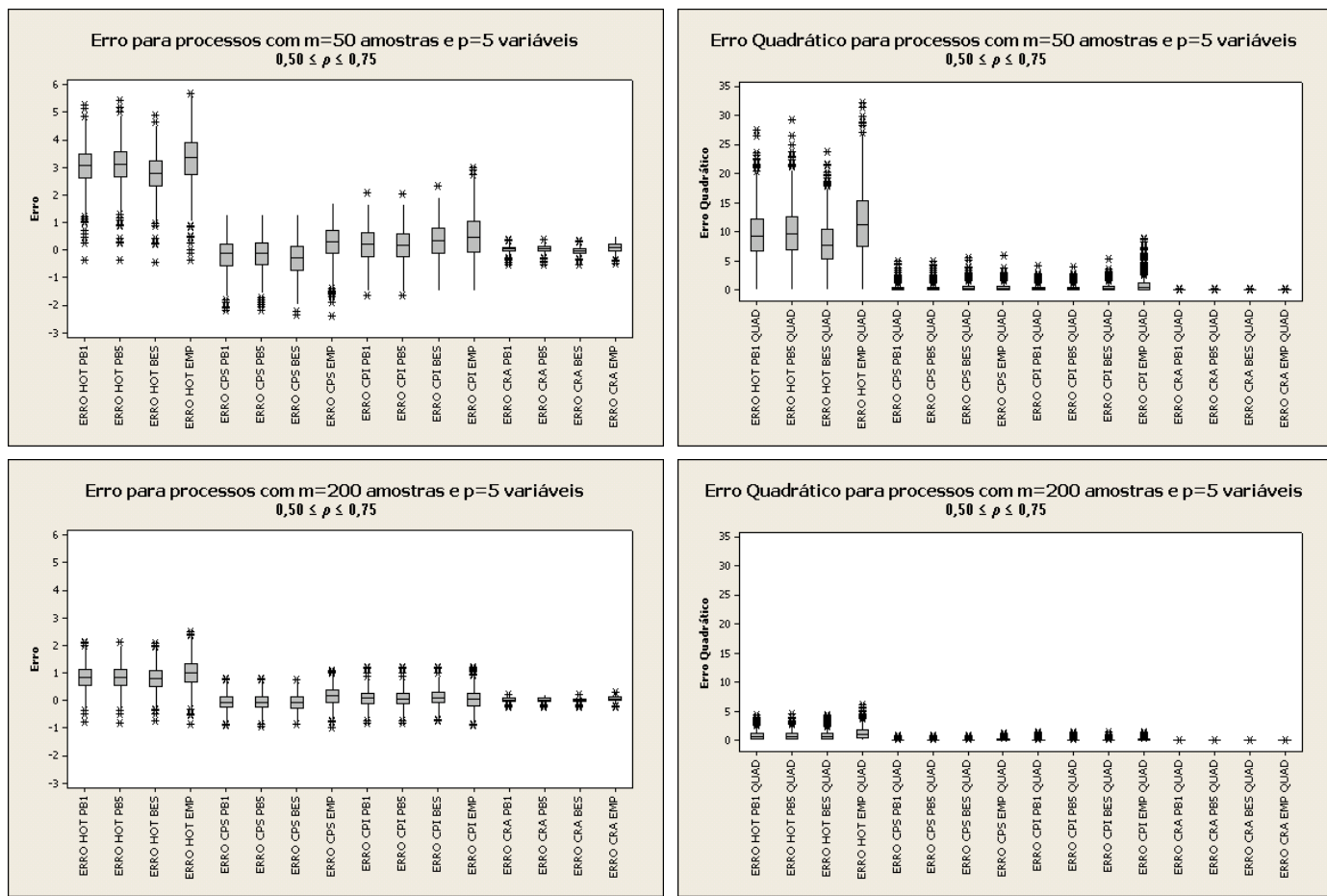


Figura 5.5: Comparação do Erro e Erro Quadrático dos limites de controle obtidos à 95% de confiança para processos com correlação entre 0,50 e 0,75, $m=50$ e 200 e $p=5$.

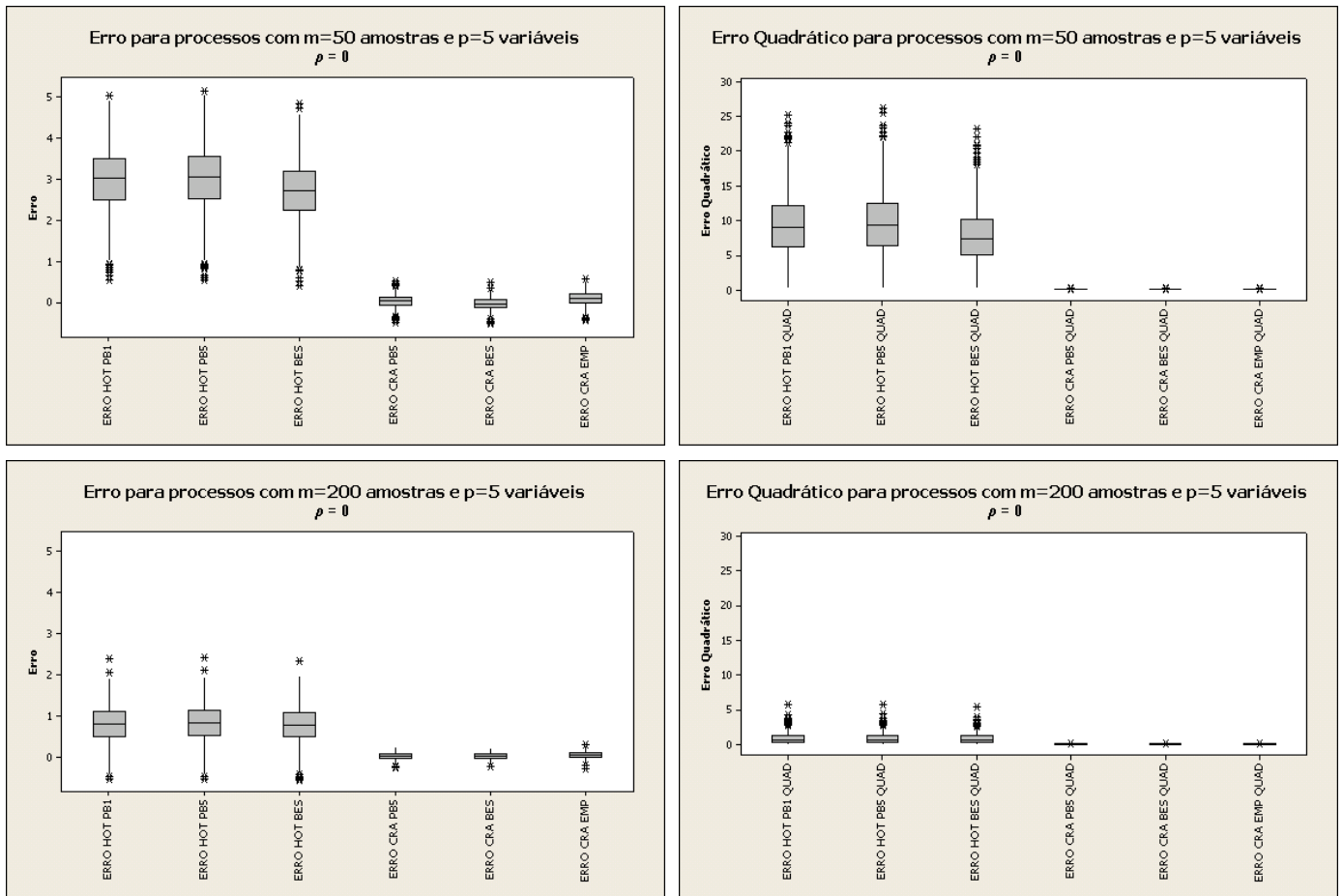


Figura 5.6: Comparação do Erro e Erro Quadrático dos limites de controle obtidos à 95% de confiança para processos com correlação igual zero, $m=50$ e 200 e $p=5$.

5.1.3 – Processos multivariados com $p=10$ variáveis

Apresentamos na Tabela 5.3 e Figuras 5.7 a 5.9 os valores dos erros médios e os erros quadráticos médios dos limites de controle para os processos multivariados com $p=10$ variáveis em que essas variáveis possuem correlação forte ($0,75 \leq \rho < 1$), correlação intermediária ($0,50 \leq \rho < 0,75$) e correlação nula ($\rho = 0$) para amostras de tamanho $m=50$ e $m=200$.

Tabela 5.3: Erros médio e Quadrático médio dos limites de controle obtidos à 95% de confiança para processos com $m=50$ e 200 e $p=10$.

| Tipo de Erro | Técnica | $m=50$ e $0,75 < \rho \leq 1$ | | | | $m=200$ e $0,75 < \rho \leq 1$ | | | |
|-----------------------|---------|----------------------------------|---------|----------------------|---------|-----------------------------------|---------|--------|---------|
| | | T2 | LSC-CP | LIC-CP | M | T2 | LSC-CP | LIC-CP | M |
| Erro Médio | Pb=1 | 8,9846 | -0,2653 | 0,2343 | 0,0000 | 1,9837 | -0,1270 | 0,1235 | -0,0040 |
| | Pb=5 | 9,0276 | -0,2394 | 0,2068 | 0,0081 | 1,9970 | -0,1205 | 0,1163 | -0,0013 |
| | Bes | 8,6164 | -0,4547 | 0,4319 | -0,0545 | 1,9478 | -0,1604 | 0,1584 | -0,0094 |
| | Emp | 9,4190 | 0,3124 | 0,6984 | 0,0550 | 2,2694 | 0,2029 | 0,0296 | 0,0347 |
| Erro Quadrático Médio | Pb=1 | 81,5900 | 0,9548 | 0,9352 | 0,0135 | 4,2660 | 0,2298 | 0,2074 | 0,0049 |
| | Pb=5 | 82,3860 | 0,9484 | 0,9328 | 0,0144 | 4,3210 | 0,2290 | 0,2079 | 0,0050 |
| | Bes | 75,0990 | 1,2126 | 1,1961 | 0,0149 | 4,1183 | 0,2363 | 0,2138 | 0,0047 |
| | Emp | 90,0260 | 1,0540 | 2,1000 | 0,0280 | 5,5870 | 0,3032 | 0,2843 | 0,0083 |
| Tipo de Erro | Técnica | $m=50$ e $0,50 < \rho \leq 0,75$ | | | | $m=200$ e $0,50 < \rho \leq 0,75$ | | | |
| | | T2 | LSC-CP | LIC-CP | M | T2 | LSC-CP | LIC-CP | M |
| Erro Médio | Pb=1 | 8,9825 | -0,2762 | 0,2417 | 0,0140 | 1,9994 | -0,1325 | 0,0925 | 0,0049 |
| | Pb=5 | 9,0256 | -0,2532 | 0,2181 | 0,0201 | 2,0116 | -0,1270 | 0,0872 | 0,0068 |
| | Bes | 8,6095 | -0,4483 | 0,4152 | -0,0382 | 1,9641 | -0,1620 | 0,1218 | 0,0000 |
| | Emp | 9,4620 | 0,2326 | 0,6548 | 0,0743 | 2,3013 | 0,1491 | 0,0347 | 0,0456 |
| Erro Quadrático Médio | Pb=1 | 81,5880 | 0,6849 | 0,7166 | 0,0159 | 4,3358 | 0,1744 | 0,1685 | 0,0055 |
| | Pb=5 | 82,3840 | 0,6779 | 0,7099 | 0,0166 | 4,3881 | 0,1737 | 0,1685 | 0,0056 |
| | Bes | 75,0200 | 0,9019 | 0,9252 | 0,0161 | 4,1878 | 0,1806 | 0,1724 | 0,0054 |
| | Emp | 90,8500 | 0,7304 | 1,6307 | 0,0315 | 5,7292 | 0,2189 | 0,2375 | 0,0096 |
| Tipo de Erro | Técnica | $m=50$ e $\rho = 0$ | | $m=200$ e $\rho = 0$ | | | | | |
| | | T2 | M | T2 | M | | | | |
| Erro Médio | Pb=1 | 8,9864 | 0,0551 | 2,0087 | 0,0144 | | | | |
| | Pb=5 | 9,0339 | 0,0601 | 2,0211 | 0,0157 | | | | |
| | Bes | 8,9279 | 0,0482 | 2,0090 | 0,0140 | | | | |
| | Emp | 9,4502 | 0,1142 | 2,3033 | 0,0524 | | | | |
| Erro Quadrático Médio | Pb=1 | 81,5620 | 0,0212 | 4,3383 | 0,0056 | | | | |
| | Pb=5 | 82,4400 | 0,0222 | 4,3918 | 0,0057 | | | | |
| | Bes | 80,4960 | 0,0201 | 4,3360 | 0,0055 | | | | |
| | Emp | 90,4960 | 0,0380 | 5,7075 | 0,0095 | | | | |

Legenda: T2: Limite superior de controle da estatística T^2 de Hotelling (1947); LSC-CP: Limite superior de controle da 1ª componente principal; LIC-CP: Limite inferior de controle da 1ª componente principal; M: Limite superior de controle da estatística M de Hayter e Tsui (1994); Pb1: Limite de controle obtido pela a técnica de Polansky e Baker (2000) com $b=1$; Pb5: Limite de controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; Bes: Limite de controle obtido pela a técnica de Bessegato (2001); Emp: Limite de controle obtido pela função da função de distribuição empírica.

Assim como no caso em que $p=2$ e 5 , observa-se na Tabela 5.3 para cada técnica fixada, e m tanto para os erros médios quanto para os erros quadráticos médios são semelhantes, independentemente da mudança nas faixas de correlação.

Em geral, na maioria das vezes a técnica de núcleo estimador usando as metodologias de Polansky e Baker (2000) com $b=5$ e Bessegato (2000) forneceu menor valor de erro médio e erro quadrático médio indicando que esses dois métodos são os mais apropriados para determinação do valor da janela ótima h na construção das cartas de controle para o processo quando o núcleo-estimador é utilizado.

A função de distribuição empírica forneceu sempre maiores valores dos erros médios e quadráticos médios, mostrando novamente ser pouco adequada na determinação dos limites de controle.

Independente da faixa de correlação e do valor de m os limites de controle obtidos via estatística M foram melhores estimados pois os erros médio e quadrático médio são sempre menores e próximos a zero. Os valores dos erros médios e quadráticos médios obtidos para a estatística T^2 de Hotelling (1947) quando consideramos $m=50$ são muito elevados e, quando comparados aos resultados obtidos para $p=5$ chegaram quase quadruplicar de valor, o que mostra que o uso dessa estatística pode não ser eficiente quando deseja-se estimar os limites de controle de forma não-paramétrica e temos uma amostra pequena. A primeira componente principal forneceu erros médios e quadráticos médios entre os valores obtidos para a estatística M e para estatística T^2 . Em linhas gerais, os gráficos *boxplot* indicam as mesmas conclusões que aquelas obtidas para os processos com $p=2$ e 5 variáveis.

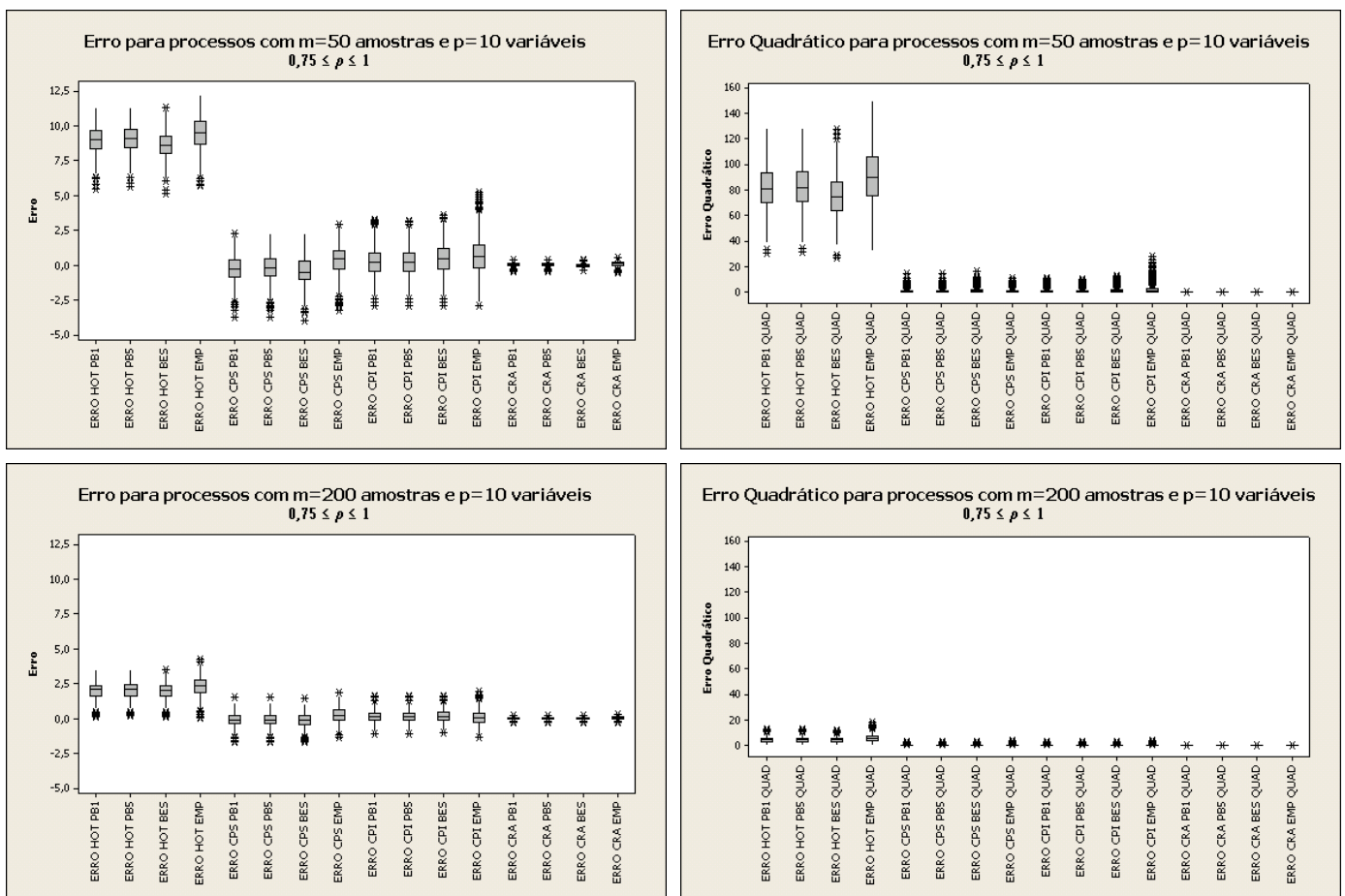


Figura 5.7: Comparação do Erro e Erro Quadrático dos limites de controle obtidos à 95% de confiança para processos com correlação entre 0,75 e 1, $m=50$ e 200 e $p=10$.

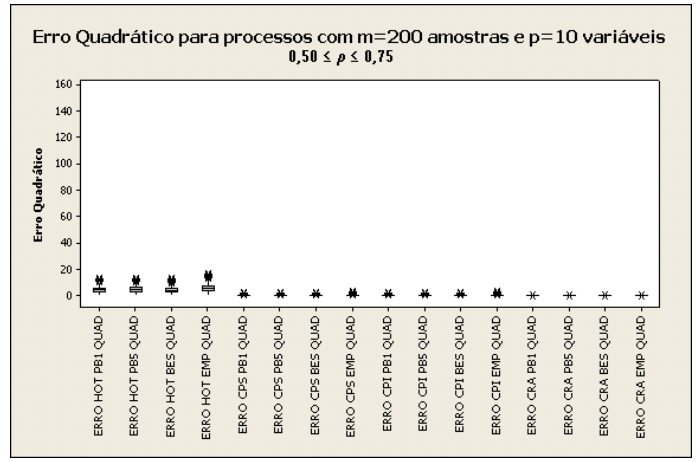
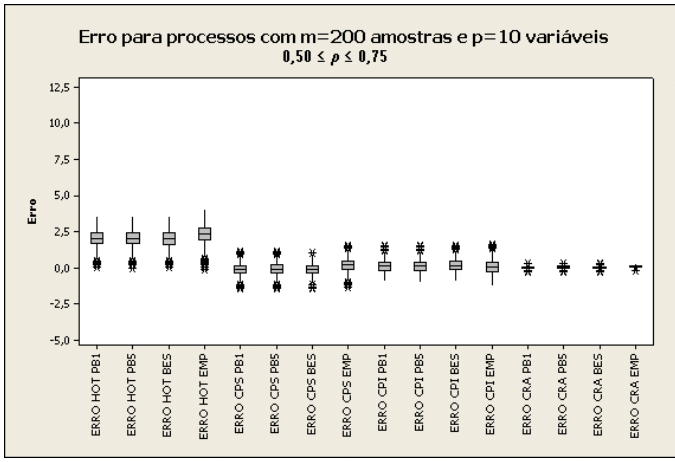
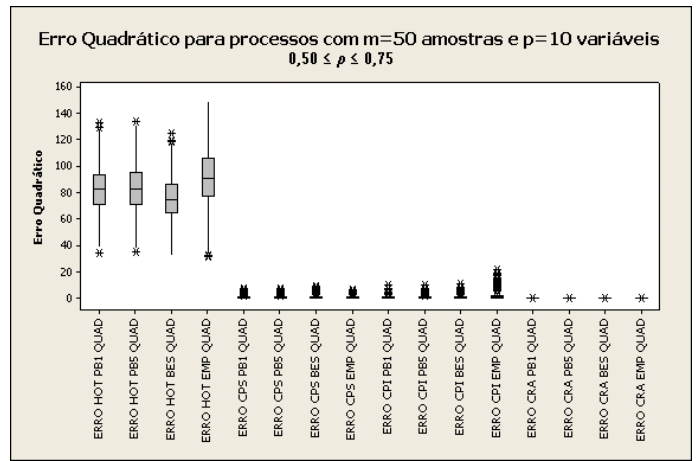
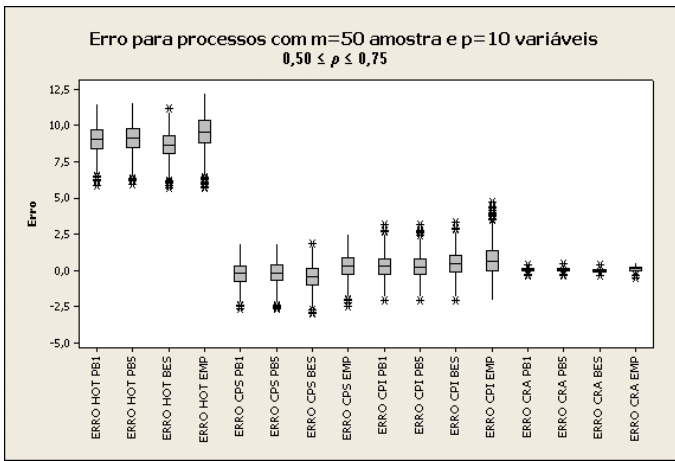


Figura 5.8: Comparação do Erro e Erro Quadrático dos limites de controle obtidos à 95% de confiança para processos com correlação entre 0,50 e 0,75, $m=50$ e 200 e $p=10$.

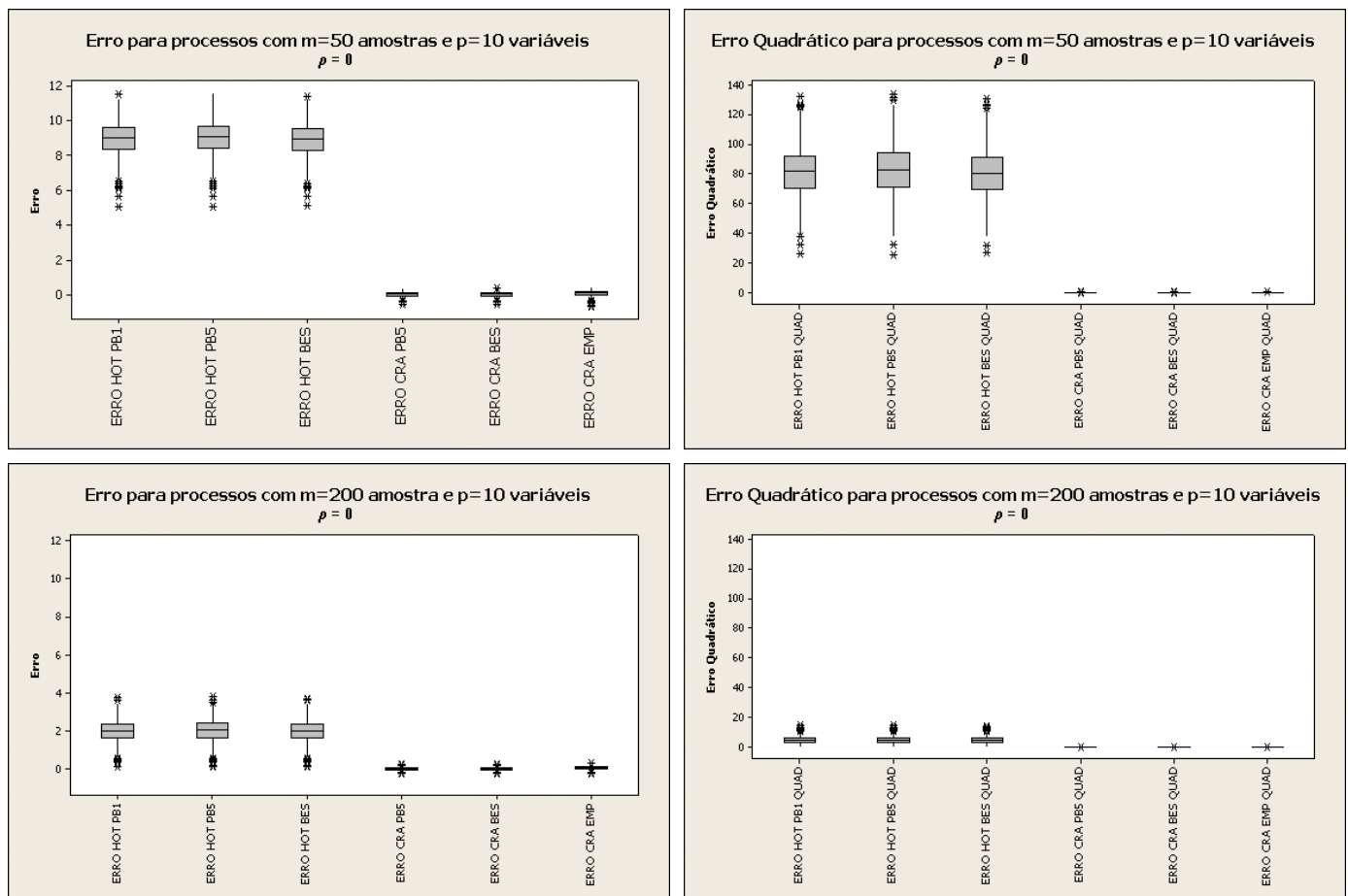


Figura 5.9 Comparação do Erro e Erro Quadrático dos limites de controle obtidos à 95% de confiança para processos com correlação igual a zero, $m=50$ e 200 e $p=10$.

5.1.4 – Processos multivariados com $p=20$ variáveis

Apresentamos na Tabela 5.4 e Figura 5.10 os valores dos erros médios e os erros quadráticos médios dos limites de controle para os processos multivariados com $p=20$ variáveis em que essas variáveis possuem correlações misturadas (ver página 59). Foram simulados também amostras com $m=1000$ com a finalidade de verificar se a qualidade das técnicas de estimação melhorava com o aumento da amostra, já que $m=50$, 100 e 200 podem ser considerados valores pequenos para estimação dos limites de controle quando $p=20$. Isto se deve ao fato de que o número de parâmetros da matriz de variâncias e covariâncias teórica do processo (Σ) é igual a $p(l+p)/2=210$.

Tabela 5.4: Erros médio e Quadrático médio dos limites de controle obtidos à 95% de confiança para processos com correlação misturada, $m=50, 100, 200$ e 1000 e $p=20$.

| Tipo de Erro | Técnica | m=50 e ρ misturado | | | | m=100 e ρ misturado | | | |
|-----------------------|---------|--------------------------|---------|--------|--------|---------------------------|---------|--------|--------|
| | | T2 | LSC-CP | LIC-CP | M | T2 | LSC-CP | LIC-CP | M |
| Erro Médio | Pb=1 | 36,373 | -0,568 | 0,570 | 0,057 | 12,9090 | -0,3547 | 0,3749 | 0,0278 |
| | Pb=5 | 36,424 | -0,544 | 0,548 | 0,062 | 12,9390 | -0,3420 | 0,3609 | 0,0302 |
| | Bes | 36,291 | -0,597 | 0,598 | 0,051 | 12,8770 | -0,3699 | 0,3903 | 0,0260 |
| | Emp | 37,059 | -0,071 | 0,976 | 0,113 | 13,6390 | -0,1128 | 0,5730 | 0,0944 |
| Erro Quadrático Médio | Pb=1 | 1324,000 | 0,762 | 0,768 | 0,022 | 167,4600 | 0,3151 | 0,3377 | 0,0112 |
| | Pb=5 | 1327,700 | 0,740 | 0,747 | 0,023 | 168,2500 | 0,3094 | 0,3310 | 0,0114 |
| | Bes | 1318,000 | 0,802 | 0,810 | 0,021 | 166,6200 | 0,3247 | 0,3482 | 0,0110 |
| | Emp | 1374,700 | 0,509 | 1,889 | 0,039 | 187,0600 | 0,2919 | 0,7228 | 0,0222 |
| Tipo de Erro | Técnica | m=200 e ρ misturado | | | | m=1000 e ρ misturado | | | |
| | | T2 | LSC-CP | LIC-CP | M | T2 | LSC-CP | LIC-CP | M |
| Erro Médio | Pb=1 | 5,8021 | -0,2000 | 0,1953 | 0,0191 | 1,3458 | -0,0494 | 0,0390 | 0,0109 |
| | Pb=5 | 5,8122 | -0,1949 | 0,1908 | 0,0203 | 1,3473 | -0,0489 | 0,0384 | 0,0110 |
| | Bes | 5,7955 | -0,2052 | 0,2002 | 0,0186 | 1,3462 | -0,0498 | 0,0394 | 0,0109 |
| | Emp | 6,2457 | 0,0437 | 0,1354 | 0,0566 | 1,4710 | 0,0187 | 0,0069 | 0,0215 |
| Erro Quadrático Médio | Pb=1 | 34,1780 | 0,1480 | 0,1443 | 0,0061 | 1,9503 | 0,0295 | 0,0253 | 0,0016 |
| | Pb=5 | 34,2980 | 0,1467 | 0,1429 | 0,0061 | 1,9542 | 0,0295 | 0,0253 | 0,0016 |
| | Bes | 34,0970 | 0,1494 | 0,1456 | 0,0060 | 1,9511 | 0,0296 | 0,0253 | 0,0016 |
| | Emp | 39,6980 | 0,1507 | 0,1737 | 0,0104 | 2,3357 | 0,0326 | 0,0312 | 0,0021 |

Legenda: T2: Limite superior de controle da estatística T^2 de Hotelling (1947); LSC-CP: Limite superior de controle da 1ª componente principal; LIC-CP: Limite inferior de controle da 1ª componente principal; M: Limite superior de controle da estatística M de Hayter e Tsui (1994); Pb1: Limite de controle obtido pela a técnica de Polansky e Baker (2000) com $b=1$; Pb5: Limite de controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; Bes: Limite de controle obtido pela a técnica de Bessegato (2001); Emp: Limite de controle obtido pela função da função de distribuição empírica.

Na Tabela 5.4 podemos observar os erros diminuem à medida que o tamanho da amostra aumenta, no entanto, para $m=50$ e 100 os erros são muito grandes para a estatística T^2 de Hotelling (1947) tendendo a diminuir para $m=200$ e $m=1000$.

Em geral a técnica de núcleo estimador, com os métodos de Polansky e Baker com $b=5$ e Bessegato (2000), forneceu menores erros médios e quadráticos médios em todas as simulações. Observa-se que a função de distribuição empírica, mesmo no caso em que a amostra tem tamanho $m=1000$, sempre forneceu maiores valores dos erros médios e quadráticos médios, mostrando ser uma técnica pouco adequada na determinação dos limites de controle no caso de $p=20$.

É possível observar que em todas as situações, mesmo quando o tamanho da amostra gerada é pequeno, os valores de erros da estatística M são próximos a zero e além disso são sempre menores que os observados para as outras estatísticas, fato que independe do tamanho da amostra gerada. Isso é um fato importante pois, independente do número de variáveis, a estatística M tende a produzir bons resultados tanto para amostras pequenas quanto para amostras grandes.

Os gráficos *boxplot* dos erros, para o processo com $p=20$ variáveis, em linhas gerais, resultaram nas mesmas conclusões obtidas para $p=2, 5$ e 10 .

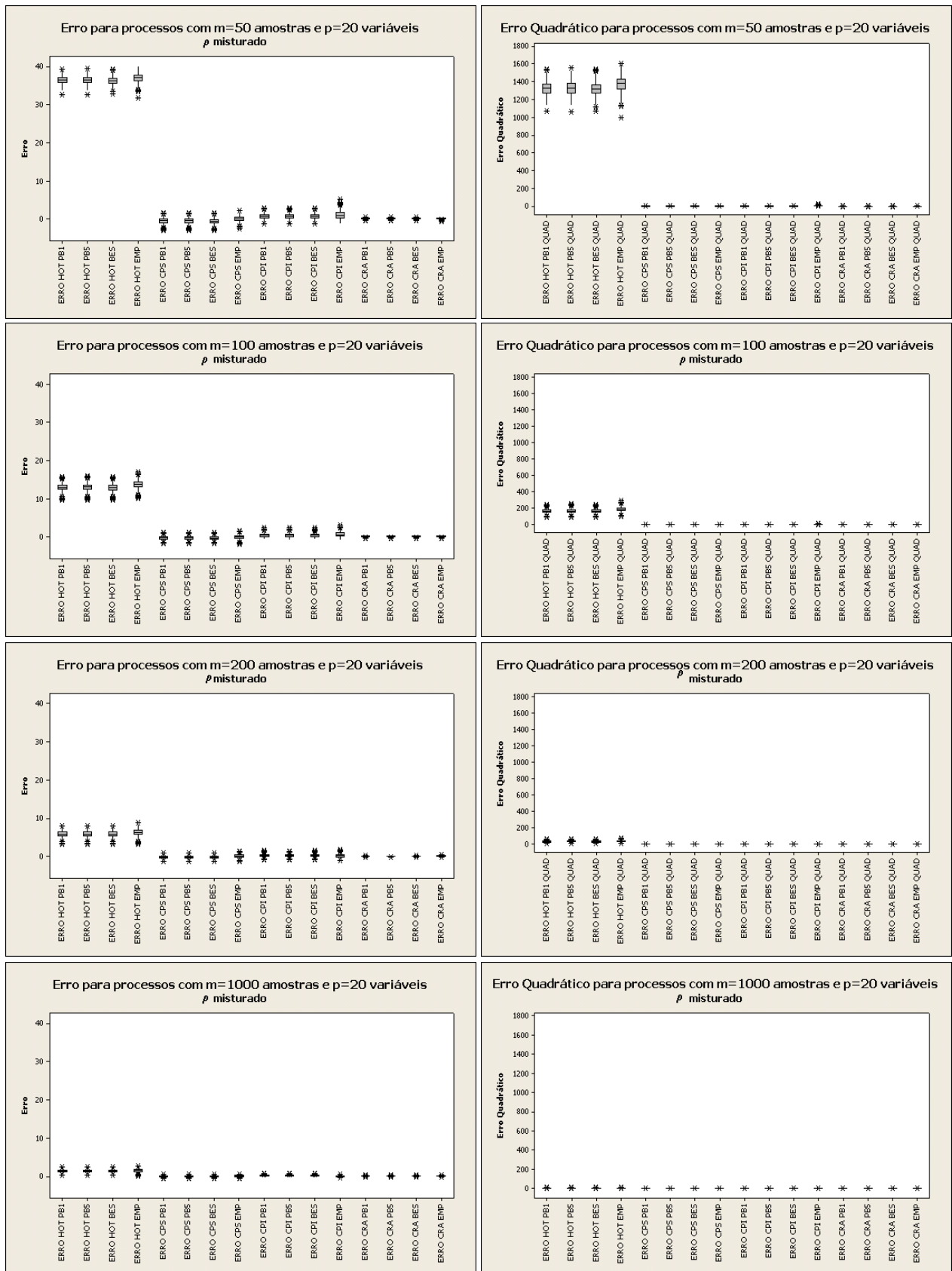


Figura 5.10 Comparação do Erro e Erro Quadrático dos limites de controle obtidos à 95% de confiança para processos com correlação igual a zero, $m=50$ e 200 e $p=10$.

5.1.5 – Nota sobre a estatística T^2 de Hotelling

Em todos os processos simulados, a estatística T^2 de Hotelling (1947) forneceu sempre maiores erros, principalmente para amostras menores, em todas as técnicas não-paramétricas avaliadas nessa dissertação. A técnica de núcleo estimador capta a informação que a amostra tem sobre a função de distribuição da variável gerada dos dados amostrais. Desse modo, é importante verificar até que ponto os dados amostrais indicavam de fato que se tinha uma distribuição F de Fisher-Snedecor no caso do ajuste da estatística T^2 de Hotelling. Assim, pode-se verificar se a alta magnitude dos erros foi devido ao uso do núcleo estimador como método de estimação da distribuição T^2 ou devido ao fato de que a distribuição amostral de T^2 não era próxima da distribuição F .

Na Figura 5.11 apresentamos a comparação de alguns histogramas, construídos para valores amostrais da estatística T^2 em comparação com os valores da distribuição F de Fisher-Snedecor. Para cada situação mostrada na Figura 5.11 foi gerada uma amostra que tinha tamanho $m=50, 200$ e 10000 com $p=2, 5, 10$ e 20 variáveis sendo que as amostras geradas tinham como base a distribuição normal multivariada com parâmetro teórico de média descrito por um vetor nulo e com parâmetro teórico de variabilidade descrito por matrizes de covariâncias que, no caso em que $p=2$ era igual a matriz descrita no processo 1 (página 56), no caso em que $p=5$ era igual a matriz descrita no processo 11 (página 57), no caso em que $p=10$ era igual a matriz descrita no processo 14 (página 56) e no caso em que $p=20$ era igual a matriz descrita no processo 17 (página 59). Para cada m e p fixo a estatística T^2 foi calculada da seguinte forma:

$$T_i^2 = (X_i - \bar{X})' S^{-1} (X_i - \bar{X}) , i = 1, 2, \dots, m$$

Para verificar a adequação da estatística F ao dados amostrais foi calculada a função de densidade teórica da distribuição F , e essa densidade plotada sobre o histograma da estatística T^2 . Os resultados são apresentados a seguir:

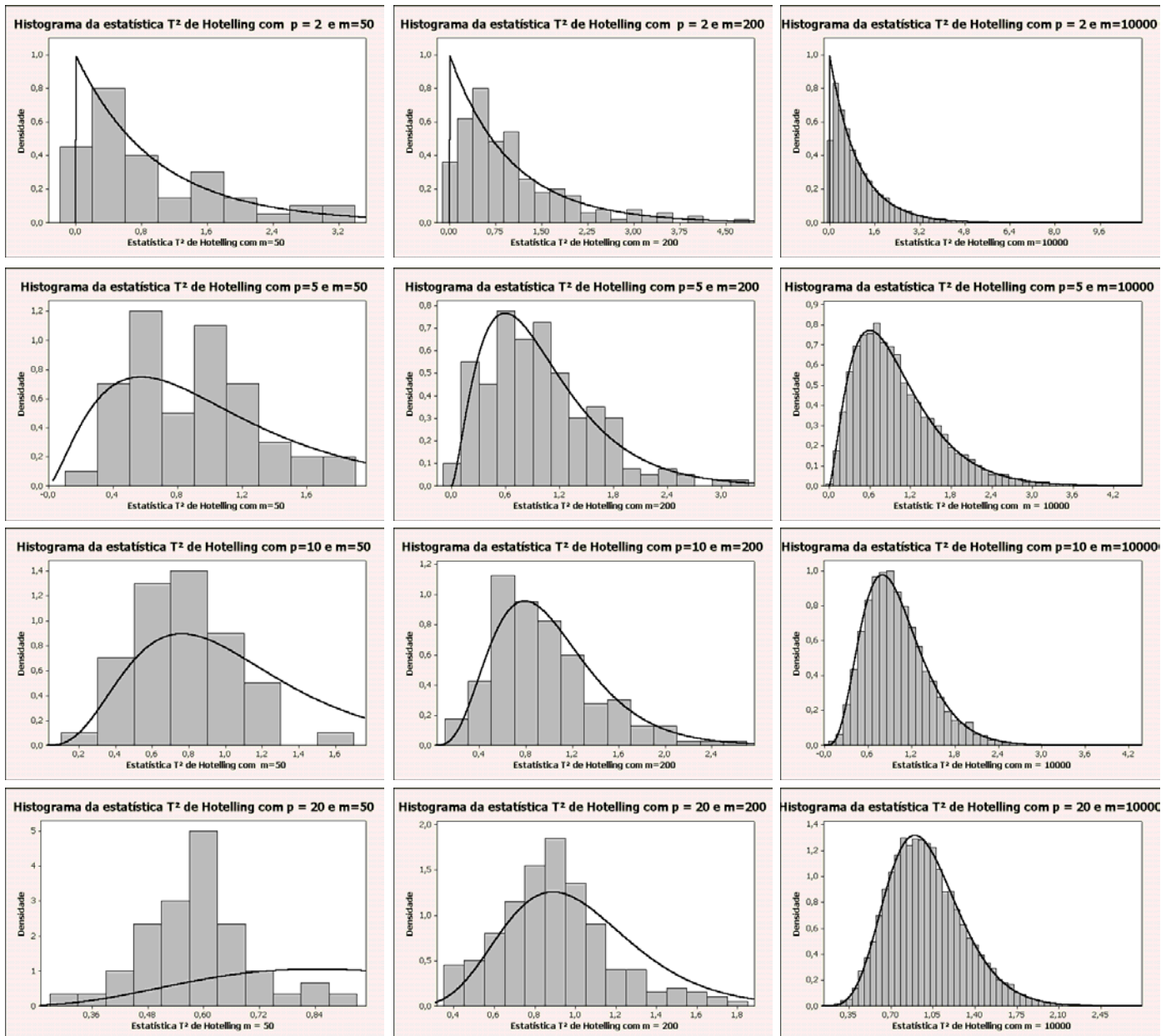


Figura 5.11 – Comparação da estatística T^2 de Hotelling com a distribuição F .

Observa-se que para $m=50$ a distribuição da estatística T^2 de Hotelling (1947) não se aproxima da distribuição F teórica. Para $m=200$ observa-se que a aproximação das distribuições amostral e teóricas são próximas sendo que para $m=10000$ as distribuições amostral e teóricas de T^2 de Hotelling (1947) se aproximam muito mais. Mesmo gerando apenas uma amostra para comparar os valores da distribuição amostral da estatística T^2 com a sua distribuição teórica (distribuição F de Fisher-Snedecor) pode-se inferir que os altos valores dos erros, que foram gerados à partir da comparação dos limites de controle estimados através das técnicas não-paramétricas com os limites de controle teóricos, deve-se à

problemas relacionados à aproximação desses limites e não à baixa eficácia das técnicas não-paramétricas na estimação dos limites de controle.

5.2 – Análise dos *ARL in control* e *out of control*

5.2.1 – Processos Multivariados: $p=2$, $p=5$ e $p=10$

Nessa seção apresentamos análises dos valores observados dos *ARL in control* e *ARL out of control* dos processos simulados em que a distribuição das variáveis é normal p -variada não-autocorrelacionada. São apresentados os resultados para os tamanhos amostrais $m=50$, 200 e em alguns casos $m=1000$, para $p=2$, 5 e 10 . Os *ARL in control* foram determinados considerando-se amostras do processo com vetor de médias nulo e matrizes de covariâncias e correlação como descritos na seção 4.1. Os *ARL out of control* foram determinados com a mudança no vetor de médias do processo (“choques”) como descrito na seção 4.3.

Os *ARL in control* e *out of control* foram calculados considerando os limites teóricos e os limites de controle obtidos através da técnica de núcleo estimador com os métodos de determinação da janela ótima h estudados por Polansky e Baker (2000) com $b=5$ e por Bessegato (2001). Os resultados referentes ao método de Polansky e Baker (2000) com $b=1$ não serão apresentados pois na seção 5.1 observa-se que os melhores resultados da estimação dos limites de controle ocorreram para $b=5$. Por apresentar piores resultados na estimação dos limites de controle frente à técnica de núcleo-estimador, os resultados da função de distribuição empírica também não serão apresentados aqui. O nível de confiança de todas as cartas de controle é de 95%. Os resultados para $m=100$ não serão apresentados pois são semelhantes aos valores obtidos para $m=200$. As tabelas com os resultados do *ARL in control* e *out of control* para $p=5$ e $p=10$ são apresentadas nos Anexos A e B. Em todos os casos foram simuladas $k=1000$ réplicas.

Nas Tabelas 5.5 a 5.7 apresenta-se os valores do *ARL in control* e *ARL out of control* para processos com $p=2$ em que as variáveis têm correlação forte ($0,75 \leq \rho \leq 0,90$) com exceção do caso em que $m=1000$ onde temos apenas correlação igual a 0,90; nas Tabelas 5.8 e 5.9 os valores do *ARL in control* e *ARL out of control* para processos com $p=2$ em que as variáveis têm correlação média ($0,25 \leq \rho \leq 0,50$) e na Tabela 5.10 os valores do *ARL in control* e *ARL out of control* para processos com $p=2$ em que as variáveis têm correlação nula ($\rho = 0$).

É possível observar que a estatística T^2 de Hotelling (1947) forneceu, em todos os processos simulados, menores valores dos *ARL in control* e *ARL out of control* o que indica

que mesmo com o processo sob controle muitas observações foram geradas fora dos limites de controle gerando uma grande quantidade de alarmes falsos. A primeira componente principal, forneceu maiores valores de *ARL in control*, mas em compensação gerou maiores valores de *ARL out of control* o que indica que mesmo quando o processo está fora de controle estatístico muitas observações são consideradas sob controle gerando uma alta taxa de não identificação de “alarmes verdadeiros”. A estatística *M*, por sua vez apresenta valores intermediários de *ARL in control* e *ARL out of control* em comparação com os *ARL*'s da estatística *T² de Hotelling* e da primeira componente principal. Esses resultados em geral indicam que a estatística *M* de Hayter e Tsui (1994) deve ser preferida na análise dos processos com distribuição normal bivariada.

Mantendo as faixas de correlação fixas, verificamos que na maioria das vezes o valor do *ARL in control* para as estatísticas *T²* e *M* aumentam e da primeira componente principal diminui, quando aumentamos os tamanhos das amostras. O valor do *ARL out of control* para as estatísticas *T²* e *M* mantém-se constantes e da primeira componente principal diminui quando aumentamos os tamanhos das amostras. Os valores dos *ARL out of control* são menores para aqueles “choques” que possuem maior intensidade (a média do processo se desloca em um valor igual a 4), sendo que o choque mais facilmente percebido é aquele em que as médias de todas as variáveis do processo mudam simultaneamente.

Apesar de não mostrado nas tabelas apresentadas, foi realizado um estudo que mostra que os valores do *ARL in control* e *ARL out of control* foram melhores para aqueles processos em que as variâncias das variáveis eram menores em comparação com aqueles em que variâncias eram maiores.

As cartas de controle construídas utilizando-se a técnica de núcleo-estimador possuíram resultados semelhantes, na maioria das vezes, tanto para o método de Polansky e Baker (2000) com $b=5$ quanto para o de Bessegato (2001). As Tabelas A1 a A5 para $p=5$ e B1 a B5 para $p=10$ (ver anexos A e B) indicam conclusões similares aos observados para $p=2$.

Tabela 5.5: *ARL in control e out of control* de processos com distribuição normal bivariada - $0,75 \leq \rho \leq 0,90$ e $m=50$.

| Distribuição bivariada com $0,75 \leq \rho \leq 0,90$ e $m=50$ | | Medidas Descritivas | Estatística T ² de Hotelling | | | Estatística M | | | Componentes Principais | | | | |
|--|--|---------------------|---|---------|---------|---------------|---------|---------|------------------------|---------|---------|---------|---------|
| | | | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | | |
| ARL in control | | Média | 21,6950 | 14,3080 | 14,3860 | 17,5230 | 18,2210 | 20,2890 | 23,6010 | 24,7980 | 20,2890 | | |
| | | Desvio | 24,9630 | 17,3310 | 17,3530 | 19,6690 | 21,5500 | 19,4520 | 29,4650 | 31,1530 | 19,4520 | | |
| | | Mediana | 14,0 | 9,0 | 9,0 | 11,0 | 11,0 | 14,0 | 14,0 | 15,0 | 14,0 | | |
| Todas as médias mudam simultaneamente na mesma escala | | Choque = 0,5 | | Média | 14,7360 | 10,1240 | 10,2510 | 12,4380 | 12,3980 | 12,7370 | 14,3160 | 16,9850 | 18,2030 |
| | | | | Desvio | 16,7410 | 11,8650 | 12,0130 | 14,5780 | 14,3050 | 14,7640 | 13,8730 | 20,7110 | 22,4150 |
| | | | | Mediana | 9,0 | 6,0 | 6,5 | 8,0 | 8,0 | 8,0 | 10,0 | 10,0 | 11,0 |
| | | Choque = 1 | | Média | 7,3220 | 5,3855 | 5,4248 | 6,1020 | 6,1980 | 6,3180 | 7,9780 | 10,0780 | 10,5090 |
| | | | | Desvio | 8,2370 | 5,8959 | 5,9004 | 6,3520 | 6,6040 | 6,6810 | 8,0540 | 14,4180 | 14,9330 |
| | | | | Mediana | 5,0 | 3,0 | 3,0 | 4,0 | 4,0 | 4,0 | 5,0 | 6,0 | 6,0 |
| | | Choque = 2 | | Média | 2,1450 | 1,8343 | 1,8362 | 1,9760 | 1,9988 | 2,0195 | 3,0673 | 3,6242 | 3,7228 |
| | | | | Desvio | 1,7685 | 1,3956 | 1,3889 | 1,5369 | 1,5782 | 1,6043 | 3,0306 | 4,5537 | 4,8465 |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 |
| | | Choque = 3 | | Média | 1,2095 | 1,1438 | 1,1443 | 1,1787 | 1,1882 | 1,1912 | 1,6868 | 1,8282 | 1,8683 |
| | | | | Desvio | 0,5650 | 0,4457 | 0,4455 | 0,4862 | 0,4984 | 0,5047 | 1,3591 | 1,7029 | 1,8198 |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| Choque = 4 | | Média | 1,0287 | 1,0195 | 1,0197 | 1,0245 | 1,0260 | 1,0260 | 1,2375 | 1,2867 | 1,3005 | | |
| | | Desvio | 0,1716 | 0,1419 | 0,1427 | 0,1546 | 0,1592 | 0,1592 | 0,6031 | 0,7377 | 0,7634 | | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | | |
| Uma média permanece constante e a outra muda arbitrariamente | | Choque = 0,5 | | Média | 13,3170 | 9,0860 | 9,1170 | 14,7960 | 15,1840 | 15,5730 | 17,8580 | 21,0270 | 22,1330 |
| | | | | Desvio | 17,8130 | 11,4460 | 11,3540 | 18,2040 | 19,1480 | 19,4200 | 17,5030 | 26,7850 | 28,3680 |
| | | | | Mediana | 8,0 | 5,0 | 5,0 | 9,0 | 9,0 | 10,0 | 13,0 | 13,0 | 13,0 |
| | | Choque = 1 | | Média | 6,2730 | 4,7450 | 4,7650 | 9,8500 | 10,5390 | 10,7460 | 13,2190 | 16,4740 | 17,2570 |
| | | | | Desvio | 8,6360 | 6,7680 | 6,7580 | 12,2440 | 15,4600 | 15,4310 | 12,8850 | 22,1580 | 22,9280 |
| | | | | Mediana | 3,0 | 3,0 | 3,0 | 6,0 | 6,0 | 6,0 | 9,0 | 9,0 | 10,0 |
| | | Choque = 2 | | Média | 2,2473 | 1,9355 | 1,9382 | 4,3207 | 4,4105 | 4,4930 | 6,3768 | 7,9890 | 8,3150 |
| | | | | Desvio | 2,8866 | 2,2634 | 2,2537 | 5,3160 | 5,5706 | 5,6271 | 5,9879 | 9,4520 | 9,8030 |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 | 4,0 | 5,0 | 5,0 |
| | | Choque = 3 | | Média | 1,5345 | 1,3972 | 1,3982 | 2,8515 | 2,8675 | 2,9215 | 4,0423 | 4,7662 | 4,9028 |
| | | | | Desvio | 1,1609 | 0,9395 | 0,9415 | 2,7550 | 2,8483 | 2,9312 | 3,5251 | 5,1627 | 5,3577 |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 | 3,0 | 3,0 | 3,0 |
| Choque = 4 | | Média | 1,0873 | 1,0647 | 1,0650 | 1,5653 | 1,5928 | 1,6078 | 2,1825 | 2,4213 | 2,4868 | | |
| | | Desvio | 0,3907 | 0,3444 | 0,3403 | 1,2606 | 1,3967 | 1,4131 | 1,6935 | 2,2041 | 2,3039 | | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 | | |
| Uma média muda pouco e outra muda arbitrariamente | | Choque = 0,5 | | Média | 14,8440 | 10,4040 | 10,4550 | 12,5240 | 12,6220 | 13,0150 | 14,8510 | 17,6650 | 18,5940 |
| | | | | Desvio | 17,1140 | 11,9480 | 11,9720 | 14,8670 | 14,6140 | 14,9870 | 14,8640 | 22,6870 | 23,5670 |
| | | | | Mediana | 9,0 | 7,0 | 7,0 | 8,0 | 8,0 | 8,0 | 10,0 | 10,0 | 11,0 |
| | | Choque = 1 | | Média | 11,0890 | 7,9650 | 8,0230 | 9,3570 | 9,4920 | 9,7330 | 10,2560 | 12,7650 | 13,4690 |
| | | | | Desvio | 14,1220 | 10,0890 | 10,0480 | 11,1040 | 11,5280 | 11,7300 | 10,0080 | 16,3920 | 17,1310 |
| | | | | Mediana | 6,0 | 4,0 | 5,0 | 6,0 | 6,0 | 6,0 | 7,0 | 7,0 | 8,0 |
| | | Choque = 2 | | Média | 3,8658 | 3,0128 | 3,0203 | 4,4872 | 4,5713 | 4,6693 | 5,1668 | 6,1070 | 6,3250 |
| | | | | Desvio | 5,7118 | 4,0194 | 4,0104 | 5,5791 | 6,0641 | 6,1537 | 5,1206 | 7,3230 | 7,6430 |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 3,0 | 3,0 | 3,0 | 3,0 | 4,0 | 4,0 |
| | | Choque = 3 | | Média | 1,7790 | 1,6063 | 1,6115 | 2,4720 | 2,4868 | 2,5195 | 2,9795 | 3,4207 | 3,5175 |
| | | | | Desvio | 1,8061 | 1,5272 | 1,5387 | 2,9578 | 2,9452 | 2,9794 | 2,5917 | 3,5790 | 3,6680 |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 |
| Choque = 4 | | Média | 1,2092 | 1,1572 | 1,1572 | 1,5685 | 1,5817 | 1,5932 | 1,8947 | 2,0363 | 2,0675 | | |
| | | Desvio | 0,6675 | 0,5649 | 0,5600 | 1,2682 | 1,2920 | 1,3133 | 1,3965 | 1,7878 | 1,8222 | | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | | |

Legenda: **Teo:** ARL para o Limite de Controle teórico; **Pb5:** ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com b=5; **Bes:** ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

Tabela 5.6: *ARL in control e out of control* de processos com distribuição normal bivariada - $(0,75 \leq \rho \leq 0,90)$ e $m=200$.

| Distribuição bivariada com $0,75 \leq \rho \leq 0,90$ e $m=200$ | | Medidas Descritivas | Estatística T ² de Hotelling | | | Estatística M | | | Componentes Principais | | |
|---|--------------|---------------------|---|---------|---------|---------------|---------|---------|------------------------|---------|---------|
| | | | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes |
| ARL in control | | Média | 18,3500 | 16,1570 | 16,1780 | 17,9450 | 18,3690 | 18,4410 | 20,4410 | 21,4020 | 21,6340 |
| | | Desvio | 18,5490 | 16,6050 | 16,6170 | 17,8300 | 18,6850 | 18,7500 | 19,8940 | 22,1260 | 22,3020 |
| | | Mediana | 12,0 | 11,0 | 11,0 | 12,0 | 12,0 | 12,0 | 14,0 | 14,0 | 15,0 |
| Todas as médias mudam simultaneamente na mesma escala | Choque = 0,5 | Média | 12,2610 | 11,0930 | 11,1170 | 11,9980 | 12,2590 | 12,2980 | 14,4050 | 15,3890 | 15,5460 |
| | | Desvio | 11,9660 | 11,2240 | 11,2630 | 12,3870 | 12,9650 | 12,9840 | 14,4330 | 16,8810 | 17,1890 |
| | | Mediana | 8,0 | 7,0 | 7,5 | 8,0 | 8,0 | 8,0 | 10,0 | 10,0 | 10,0 |
| | Choque = 1 | Média | 5,9655 | 5,5043 | 5,5010 | 5,9643 | 5,9948 | 6,0073 | 8,0450 | 8,5080 | 8,5910 |
| | | Desvio | 5,8197 | 5,2617 | 5,2469 | 5,6835 | 5,7765 | 5,7819 | 8,4260 | 9,5230 | 9,6730 |
| | | Mediana | 4,0 | 4,0 | 4,0 | 4,0 | 4,0 | 4,0 | 5,0 | 6,0 | 6,0 |
| | Choque = 2 | Média | 1,9863 | 1,9275 | 1,9275 | 1,9918 | 1,9928 | 1,9978 | 3,0297 | 3,1788 | 3,1945 |
| | | Desvio | 1,4701 | 1,4326 | 1,4326 | 1,4043 | 1,4084 | 1,4136 | 2,9291 | 3,2458 | 3,2699 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 2,0 | 1,0 | 2,0 | 2,0 | 2,0 | 2,0 |
| | Choque = 3 | Média | 1,1748 | 1,1575 | 1,1573 | 1,1895 | 1,1903 | 1,1918 | 1,7267 | 1,7572 | 1,7610 |
| | | Desvio | 0,4683 | 0,4379 | 0,4377 | 0,4808 | 0,4839 | 0,4858 | 1,3570 | 1,4555 | 1,4668 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| Choque = 4 | Média | 1,0205 | 1,0175 | 1,0175 | 1,0210 | 1,0212 | 1,0212 | 1,2342 | 1,2510 | 1,2520 | |
| | Desvio | 0,1417 | 0,1311 | 0,1311 | 0,1434 | 0,1442 | 0,1442 | 0,6281 | 0,6709 | 0,6716 | |
| | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| Uma média permanece constante e a outra muda ordenadamente | Choque = 0,5 | Média | 10,8320 | 9,7970 | 9,8020 | 14,5210 | 14,7640 | 14,8330 | 17,6510 | 18,9340 | 19,1510 |
| | | Desvio | 12,2040 | 10,9670 | 10,9640 | 14,8390 | 15,3500 | 15,3810 | 16,9140 | 19,6680 | 19,8660 |
| | | Mediana | 7,0 | 6,0 | 6,0 | 10,0 | 10,0 | 10,0 | 13,0 | 13,0 | 13,0 |
| | Choque = 1 | Média | 5,4990 | 5,0975 | 5,1047 | 9,7490 | 9,7850 | 9,8450 | 13,3660 | 14,5320 | 14,6530 |
| | | Desvio | 6,5510 | 6,2565 | 6,2626 | 11,2140 | 11,1120 | 11,2140 | 12,9100 | 15,4400 | 15,6810 |
| | | Mediana | 3,0 | 3,0 | 3,0 | 6,0 | 6,0 | 6,0 | 9,0 | 9,0 | 9,0 |
| | Choque = 2 | Média | 1,9815 | 1,9195 | 1,9190 | 4,3705 | 4,4045 | 4,4173 | 6,5680 | 6,9650 | 7,0070 |
| | | Desvio | 2,0640 | 1,9900 | 1,9897 | 5,2678 | 5,3729 | 5,3830 | 6,2058 | 7,0590 | 7,0890 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 | 5,0 | 5,0 | 5,0 |
| | Choque = 3 | Média | 1,4300 | 1,3880 | 1,3890 | 2,7268 | 2,7518 | 2,7595 | 3,9475 | 4,1900 | 4,2215 |
| | | Desvio | 0,9580 | 0,8940 | 0,8952 | 2,4572 | 2,4895 | 2,4929 | 3,4438 | 3,8942 | 3,9316 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 | 3,0 | 3,0 | 3,0 |
| Choque = 4 | Média | 1,0705 | 1,0670 | 1,0677 | 1,5168 | 1,5292 | 1,5310 | 2,1858 | 2,2650 | 2,2770 | |
| | Desvio | 0,3171 | 0,3099 | 0,3117 | 1,1318 | 1,1688 | 1,1700 | 1,7119 | 1,8771 | 1,8961 | |
| | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 | |
| Uma média muda em 0,5 e a outra muda ordenadamente | Choque = 0,5 | Média | 12,5950 | 11,2870 | 11,3010 | 12,1220 | 12,3200 | 12,3730 | 14,5120 | 15,3450 | 15,4880 |
| | | Desvio | 12,6040 | 11,4070 | 11,4150 | 12,4010 | 12,9000 | 12,9590 | 14,2690 | 15,9490 | 16,0630 |
| | | Mediana | 9,0 | 8,0 | 8,0 | 8,0 | 8,0 | 8,0 | 10,0 | 10,0 | 10,0 |
| | Choque = 1 | Média | 9,3520 | 8,5360 | 8,5380 | 9,1240 | 9,2030 | 9,2550 | 10,3840 | 11,1940 | 11,3030 |
| | | Desvio | 10,5360 | 9,6980 | 9,6980 | 9,5440 | 9,7350 | 9,7740 | 10,1820 | 11,8470 | 11,9010 |
| | | Mediana | 6,0 | 5,0 | 5,0 | 6,0 | 6,0 | 6,0 | 7,0 | 7,0 | 7,5 |
| | Choque = 2 | Média | 3,3328 | 3,1132 | 3,1140 | 4,5365 | 4,5570 | 4,5758 | 5,2035 | 5,4838 | 5,5250 |
| | | Desvio | 4,3101 | 3,9353 | 3,9365 | 5,2496 | 5,4849 | 5,4985 | 5,0639 | 5,9257 | 5,9512 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 3,0 | 3,0 | 3,0 | 4,0 | 4,0 | 4,0 |
| | Choque = 3 | Média | 1,6465 | 1,5972 | 1,5960 | 2,3458 | 2,3772 | 2,3813 | 2,9325 | 3,0305 | 3,0400 |
| | | Desvio | 1,4883 | 1,3973 | 1,3958 | 2,4081 | 2,5221 | 2,5250 | 2,6012 | 2,7877 | 2,7859 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 |
| Choque = 4 | Média | 1,1650 | 1,1530 | 1,1530 | 1,5687 | 1,5833 | 1,5845 | 1,8933 | 1,9638 | 1,9703 | |
| | Desvio | 0,5476 | 0,5298 | 0,5298 | 1,2428 | 1,2736 | 1,2749 | 1,4028 | 1,5460 | 1,5596 | |
| | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |

Legenda: **Teo:** ARL para o Limite de Controle teórico; **Pb5:** ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com b=5; **Bes:** ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

Tabela 5.7: *ARL in control e out of control* de processos com distribuição normal bivariada- ($\rho = 0,90$) e $m=1000$.

| Distribuição normal bivariada com $\rho=0.9$ e $m=1000$ | | Medidas Descritivas | Estatística T ² de Hotelling | | | Estatística M | | | Componentes Principais | | |
|--|--------------|---------------------|---|---------|---------|---------------|---------|---------|------------------------|---------|---------|
| | | | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes |
| <i>ARL in control</i> | | Média | 17,1390 | 16,3140 | 16,3140 | 17,3320 | 17,4560 | 17,4880 | 18,9460 | 17,9830 | 17,9790 |
| | | Desvio | 16,7113 | 16,1347 | 16,1347 | 16,5041 | 16,9176 | 16,9470 | 17,9007 | 17,0833 | 17,0867 |
| | | Mediana | 12,0 | 11,0 | 11,0 | 12,0 | 12,0 | 12,0 | 13,0 | 13,0 | 13,0 |
| Todas as médias mudam simultaneamente na mesma escala | Choque = 0,5 | Média | 12,0250 | 11,5510 | 11,5510 | 11,3990 | 11,4910 | 11,4910 | 12,3070 | 11,9630 | 11,9530 |
| | | Desvio | 11,5047 | 11,0702 | 11,0702 | 11,0269 | 11,0068 | 11,0068 | 11,4260 | 11,1555 | 11,1428 |
| | | Mediana | 8,0 | 8,0 | 8,0 | 8,0 | 8,0 | 8,0 | 9,0 | 9,0 | 9,0 |
| | Choque = 1 | Média | 6,9140 | 6,5110 | 6,5170 | 5,3460 | 5,3450 | 5,3450 | 5,7460 | 5,4770 | 5,4770 |
| | | Desvio | 6,4710 | 6,0449 | 6,0463 | 4,9615 | 4,9370 | 4,9370 | 5,3834 | 5,2603 | 5,2603 |
| | | Mediana | 5,0 | 4,5 | 4,5 | 4,0 | 4,0 | 4,0 | 4,0 | 4,0 | 4,0 |
| | Choque = 2 | Média | 2,2070 | 2,1920 | 2,1920 | 1,8160 | 1,8210 | 1,8210 | 1,8360 | 1,8240 | 1,8250 |
| | | Desvio | 1,5597 | 1,5593 | 1,5593 | 1,1496 | 1,1766 | 1,1766 | 1,1447 | 1,1464 | 1,1470 |
| | | Mediana | 2,0 | 2,0 | 2,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| | Choque = 3 | Média | 1,2370 | 1,2280 | 1,2280 | 1,1490 | 1,1530 | 1,1530 | 1,1590 | 1,1520 | 1,1520 |
| | | Desvio | 0,5283 | 0,5121 | 0,5121 | 0,3885 | 0,3997 | 0,3997 | 0,4073 | 0,4038 | 0,4038 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| Choque = 4 | Média | 1,0290 | 1,0280 | 1,0280 | 1,0090 | 1,0100 | 1,0100 | 1,0110 | 1,0120 | 1,0120 | |
| | Desvio | 0,1738 | 0,1710 | 0,1710 | 0,0945 | 0,0995 | 0,0995 | 0,1044 | 0,1089 | 0,1089 | |
| | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| Uma média permanece constante e a outra muda ordenadamente | Choque = 0,5 | Média | 5,4090 | 5,2830 | 5,2740 | 12,3620 | 12,4230 | 12,4360 | 17,3740 | 16,3860 | 16,3900 |
| | | Desvio | 5,1622 | 5,0516 | 5,0494 | 12,0481 | 12,1952 | 12,1940 | 16,3242 | 15,2926 | 15,2891 |
| | | Mediana | 4,0 | 4,0 | 4,0 | 8,0 | 8,5 | 9,0 | 13,0 | 13,0 | 13,0 |
| | Choque = 1 | Média | 1,8630 | 1,8260 | 1,8260 | 5,9340 | 5,9750 | 5,9750 | 12,7460 | 12,2530 | 12,2570 |
| | | Desvio | 1,2024 | 1,1761 | 1,1761 | 5,2687 | 5,2817 | 5,2817 | 11,6882 | 11,4320 | 11,4369 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 4,0 | 4,0 | 4,0 | 9,0 | 9,0 | 9,0 |
| | Choque = 2 | Média | 1,0100 | 1,0100 | 1,0100 | 2,0400 | 2,0270 | 2,0290 | 5,5140 | 5,3160 | 5,3160 |
| | | Desvio | 0,0995 | 0,0995 | 0,0995 | 1,5148 | 1,4820 | 1,4820 | 4,9560 | 4,8074 | 4,8074 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 4,0 | 4,0 | 4,0 |
| | Choque = 3 | Média | 1,0000 | 1,0000 | 1,0000 | 1,1940 | 1,1970 | 1,1970 | 3,0390 | 2,9730 | 2,9750 |
| | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,5066 | 0,5143 | 0,5143 | 2,5892 | 2,5398 | 2,5391 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 |
| Choque = 4 | Média | 1,0000 | 1,0000 | 1,0000 | 1,0240 | 1,0240 | 1,0240 | 1,8500 | 1,8160 | 1,8130 | |
| | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,1531 | 0,1531 | 0,1531 | 1,2062 | 1,2081 | 1,1989 | |
| | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| Uma média muda em 0,5 e a outra muda ordenadamente | Choque = 0,5 | Média | 12,5380 | 12,0130 | 12,0130 | 11,0260 | 10,9540 | 10,9760 | 12,6320 | 11,8130 | 11,8140 |
| | | Desvio | 12,8659 | 12,4724 | 12,4724 | 10,9654 | 10,6527 | 10,6775 | 13,0782 | 11,9272 | 11,9078 |
| | | Mediana | 9,0 | 8,0 | 8,0 | 8,0 | 8,0 | 8,0 | 9,0 | 8,0 | 8,0 |
| | Choque = 1 | Média | 4,3340 | 4,2290 | 4,2290 | 6,5200 | 6,5900 | 6,5900 | 8,6710 | 8,3810 | 8,3890 |
| | | Desvio | 3,7436 | 3,6196 | 3,6196 | 5,8788 | 6,1032 | 6,1032 | 8,5293 | 8,2186 | 8,2513 |
| | | Mediana | 3,0 | 3,0 | 3,0 | 5,0 | 5,0 | 5,0 | 6,0 | 6,0 | 6,0 |
| | Choque = 2 | Média | 1,0870 | 1,0830 | 1,0830 | 2,0270 | 2,0440 | 2,0450 | 4,0470 | 3,9170 | 3,9170 |
| | | Desvio | 0,3091 | 0,3003 | 0,3003 | 1,5319 | 1,5409 | 1,5432 | 3,5667 | 3,4243 | 3,4243 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 | 3,0 | 3,0 | 3,0 |
| | Choque = 3 | Média | 1,0020 | 1,0020 | 1,0020 | 1,2000 | 1,2040 | 1,2040 | 2,1900 | 2,1560 | 2,1460 |
| | | Desvio | 0,0447 | 0,0447 | 0,0447 | 0,5003 | 0,5144 | 0,5144 | 1,6615 | 1,6347 | 1,6264 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 |
| Choque = 4 | Média | 1,0000 | 1,0000 | 1,0000 | 1,0240 | 1,0240 | 1,0240 | 1,5920 | 1,5620 | 1,5620 | |
| | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,1531 | 0,1531 | 0,1531 | 1,0211 | 0,9865 | 0,9865 | |
| | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |

Legenda: **Teo:** ARL para o Limite de Controle teórico; **Pb5:** ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com b=5; **Bes:** ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

Tabela 5.8: *ARL in control e out of control* de processos com distribuição normal bivariada - $(0,25 \leq \rho \leq 0,50)$ e $m=50$.

| Distribuição bivariada com $0,25 \leq \rho \leq 0,50$ e $m=50$ | | Medidas Descritivas | Estatística T ² de Hotelling | | | Estatística M | | | Componentes Principais | | | | |
|--|--|---------------------|---|---------|---------|---------------|---------|---------|------------------------|---------|---------|---------|---------|
| | | | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | | |
| ARL in control | | Média | 20,9640 | 13,9720 | 14,0680 | 16,3060 | 16,8590 | 17,4860 | 20,6420 | 24,5890 | 25,9780 | | |
| | | Desvio | 24,8470 | 16,2790 | 16,2810 | 18,7740 | 20,3310 | 21,4520 | 21,2000 | 32,3150 | 33,9600 | | |
| | | Mediana | 13,0 | 9,0 | 9,0 | 10,0 | 10,0 | 11,0 | 14,0 | 14,0 | 15,0 | | |
| Todas as médias mudam simultaneamente na mesma escala | | Choque = 0,5 | | Média | 14,6810 | 10,1700 | 10,2070 | 11,2440 | 11,5380 | 11,9400 | 14,5210 | 17,8360 | 18,7470 |
| | | | | Desvio | 16,6350 | 11,7680 | 11,7840 | 12,2650 | 13,3420 | 13,8230 | 15,8530 | 22,8530 | 23,9990 |
| | | | | Mediana | 9,0 | 6,0 | 6,0 | 7,0 | 7,0 | 7,0 | 10,0 | 10,0 | 11,0 |
| | | Choque = 1 | | Média | 7,1230 | 5,3458 | 5,3572 | 5,5410 | 5,6130 | 5,7597 | 7,9190 | 9,5980 | 10,1190 |
| | | | | Desvio | 7,5640 | 5,6950 | 5,6879 | 5,8042 | 6,0140 | 6,1548 | 9,1190 | 13,6670 | 14,7150 |
| | | | | Mediana | 5,0 | 3,0 | 3,0 | 4,0 | 4,0 | 4,0 | 5,0 | 5,0 | 5,0 |
| | | Choque = 2 | | Média | 2,1527 | 1,8670 | 1,8750 | 1,8752 | 1,8733 | 1,9003 | 3,1293 | 3,6553 | 3,7835 |
| | | | | Desvio | 1,7851 | 1,4378 | 1,4493 | 1,4686 | 1,4374 | 1,4904 | 4,0158 | 5,3221 | 5,5639 |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 |
| | | Choque = 3 | | Média | 1,2148 | 1,1610 | 1,1608 | 1,1562 | 1,1632 | 1,1660 | 1,8430 | 1,9502 | 1,9813 |
| | | | | Desvio | 0,5312 | 0,4490 | 0,4477 | 0,4426 | 0,4540 | 0,4577 | 2,6250 | 2,3346 | 2,3854 |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| Choque = 4 | | Média | 1,0280 | 1,0170 | 1,0170 | 1,0168 | 1,0153 | 1,0155 | 1,3193 | 1,3505 | 1,3633 | | |
| | | Desvio | 0,1795 | 0,1368 | 0,1368 | 0,1322 | 0,1246 | 0,1256 | 1,6508 | 1,6644 | 1,6875 | | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | | |
| Uma média permanece constante e a outra muda ordenadamente | | Choque = 0,5 | | Média | 17,2780 | 12,1610 | 12,2160 | 14,3760 | 14,7870 | 15,2020 | 17,7350 | 21,4600 | 22,4110 |
| | | | | Desvio | 19,4360 | 14,2470 | 14,2430 | 15,7340 | 17,2290 | 17,5830 | 17,3140 | 27,6370 | 28,6580 |
| | | | | Mediana | 11,0 | 8,0 | 8,0 | 9,0 | 9,0 | 10,0 | 12,0 | 13,0 | 13,0 |
| | | Choque = 1 | | Média | 11,4900 | 8,2550 | 8,2830 | 9,8990 | 10,1610 | 10,4080 | 12,2690 | 15,4120 | 16,1910 |
| | | | | Desvio | 14,2570 | 10,0470 | 10,0060 | 11,2420 | 12,3300 | 12,5930 | 12,1150 | 19,3730 | 20,4170 |
| | | | | Mediana | 7,0 | 5,0 | 5,0 | 6,0 | 6,0 | 6,0 | 8,0 | 9,0 | 9,0 |
| | | Choque = 2 | | Média | 4,8070 | 3,7012 | 3,7110 | 4,5420 | 4,6610 | 4,7450 | 5,6585 | 7,0460 | 7,3240 |
| | | | | Desvio | 6,2830 | 4,4590 | 4,4498 | 5,8563 | 6,1537 | 6,2315 | 6,2885 | 9,4340 | 9,8460 |
| | | | | Mediana | 3,0 | 2,0 | 2,0 | 3,0 | 3,0 | 3,0 | 4,0 | 4,0 | 4,0 |
| | | Choque = 3 | | Média | 2,5090 | 2,1172 | 2,1243 | 2,5058 | 2,5120 | 2,5565 | 3,1490 | 3,4980 | 3,6060 |
| | | | | Desvio | 3,0270 | 2,2432 | 2,2545 | 2,9823 | 3,0196 | 3,1230 | 4,3300 | 3,9469 | 4,1052 |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 |
| Choque = 4 | | Média | 1,6525 | 1,4713 | 1,4738 | 1,6767 | 1,6968 | 1,7128 | 2,0290 | 2,1610 | 2,2033 | | |
| | | Desvio | 1,4798 | 1,1345 | 1,1385 | 1,5117 | 1,5313 | 1,5562 | 2,3921 | 2,3387 | 2,3939 | | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | | |
| Uma média muda em 0,5 e outra muda ordenadamente | | Choque = 0,5 | | Média | 14,6700 | 10,1960 | 10,2940 | 11,3130 | 11,4400 | 11,7700 | 14,2770 | 17,2910 | 18,2750 |
| | | | | Desvio | 16,3290 | 11,2380 | 11,3680 | 12,3820 | 12,9370 | 13,2670 | 15,4930 | 21,8390 | 23,3760 |
| | | | | Mediana | 9,0 | 7,0 | 7,0 | 7,0 | 7,0 | 8,0 | 10,0 | 10,0 | 11,0 |
| | | Choque = 1 | | Média | 10,7320 | 7,8420 | 7,8760 | 8,3520 | 8,5000 | 8,7070 | 9,7980 | 12,0300 | 12,5850 |
| | | | | Desvio | 12,4470 | 8,8810 | 8,8770 | 9,1100 | 9,4980 | 9,7380 | 10,7060 | 15,7050 | 16,3290 |
| | | | | Mediana | 7,0 | 5,0 | 5,0 | 5,0 | 5,0 | 6,0 | 7,0 | 7,0 | 7,0 |
| | | Choque = 2 | | Média | 5,2880 | 4,1093 | 4,1443 | 4,4655 | 4,5068 | 4,5655 | 4,6840 | 5,6020 | 5,8260 |
| | | | | Desvio | 7,2260 | 5,4233 | 5,4669 | 5,6723 | 5,7448 | 5,8025 | 5,0040 | 6,9960 | 7,3370 |
| | | | | Mediana | 3,0 | 2,0 | 2,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 |
| | | Choque = 3 | | Média | 2,7995 | 2,3208 | 2,3305 | 2,4715 | 2,4985 | 2,5258 | 2,6402 | 2,9782 | 3,0540 |
| | | | | Desvio | 3,2210 | 2,5130 | 2,5225 | 2,7903 | 2,8843 | 2,9144 | 2,7672 | 3,2162 | 3,3156 |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 |
| Choque = 4 | | Média | 1,7830 | 1,5685 | 1,5718 | 1,6688 | 1,6705 | 1,6970 | 1,7758 | 1,9000 | 1,9310 | | |
| | | Desvio | 1,7250 | 1,2999 | 1,3005 | 1,5077 | 1,4765 | 1,5337 | 1,8842 | 2,0668 | 2,0946 | | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | | |

Legenda: **Teo:** ARL para o Limite de Controle teórico; **Pb5:** ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes:** ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

Tabela 5.9: *ARL in control e out of control* de processos com distribuição normal bivariada- $(0,25 \leq \rho \leq 0,50)$ e $m=200$.

| Distribuição bivariada com $0,25 \leq \rho \leq 0,50$ e $m=200$ | | Medidas Descritivas | Estatística T ² de Hotelling | | | Estatística M | | | Componentes Principais | | |
|---|--------------|---------------------|---|---------|---------|---------------|---------|---------|------------------------|---------|---------|
| | | | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes |
| ARL in control | | Média | 18,5310 | 16,2540 | 16,2510 | 17,3110 | 17,4540 | 17,5470 | 20,3560 | 21,9990 | 22,2670 |
| | | Desvio | 19,2680 | 17,0600 | 17,0820 | 17,4900 | 18,3440 | 18,4010 | 19,6220 | 23,0300 | 23,4290 |
| | | Mediana | 13,0 | 11,0 | 11,0 | 12,0 | 12,0 | 12,0 | 14,0 | 15,0 | 15,0 |
| Todas as médias mudam simultaneamente na mesma escala | Choque = 0,5 | Média | 12,5080 | 11,3730 | 11,3810 | 11,7510 | 11,8550 | 11,9140 | 13,6990 | 14,6110 | 14,8060 |
| | | Desvio | 12,4970 | 11,6660 | 11,6670 | 11,6940 | 12,1970 | 12,2440 | 13,4980 | 15,7260 | 15,9460 |
| | | Mediana | 9,0 | 8,0 | 8,0 | 8,0 | 8,0 | 8,0 | 10,0 | 10,0 | 10,0 |
| | Choque = 1 | Média | 6,1505 | 5,7510 | 5,7600 | 5,6238 | 5,7027 | 5,7240 | 7,7640 | 8,1660 | 8,2460 |
| | | Desvio | 6,1177 | 5,6943 | 5,6957 | 5,4752 | 5,6624 | 5,6868 | 8,7330 | 9,5970 | 9,7300 |
| | | Mediana | 4,0 | 4,0 | 4,0 | 4,0 | 4,0 | 4,0 | 5,0 | 5,0 | 5,0 |
| | Choque = 2 | Média | 1,9682 | 1,8872 | 1,8878 | 1,8452 | 1,8518 | 1,8547 | 3,0265 | 3,1820 | 3,1952 |
| | | Desvio | 1,4516 | 1,3880 | 1,3891 | 1,3102 | 1,3387 | 1,3437 | 3,3541 | 3,7085 | 3,7175 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 |
| | Choque = 3 | Média | 1,1748 | 1,1555 | 1,1555 | 1,1520 | 1,1537 | 1,1537 | 1,7433 | 1,7790 | 1,7868 |
| | | Desvio | 0,4683 | 0,4386 | 0,4386 | 0,4312 | 0,4372 | 0,4372 | 1,5395 | 1,5868 | 1,5956 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| | Choque = 4 | Média | 1,0218 | 1,0170 | 1,0172 | 1,0175 | 1,0172 | 1,0172 | 1,2857 | 1,2960 | 1,2973 |
| | | Desvio | 0,1542 | 0,1312 | 0,1321 | 0,1349 | 0,1340 | 0,1340 | 0,7268 | 0,7463 | 0,7484 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| Uma média permanece constante e a outra muda ordenadamente | Choque = 0,5 | Média | 14,8150 | 13,1470 | 13,1660 | 14,1710 | 14,1170 | 14,1660 | 17,2170 | 18,1270 | 18,2960 |
| | | Desvio | 15,3700 | 13,6590 | 13,6760 | 14,6010 | 14,8310 | 14,8420 | 16,6220 | 18,9510 | 19,1350 |
| | | Mediana | 10,0 | 9,0 | 9,0 | 10,0 | 9,0 | 10,0 | 12,0 | 12,0 | 12,0 |
| | Choque = 1 | Média | 9,4640 | 8,6150 | 8,6230 | 9,8110 | 9,8640 | 9,9090 | 11,5870 | 12,6590 | 12,7540 |
| | | Desvio | 10,6370 | 9,5030 | 9,5150 | 10,6120 | 10,8730 | 10,9110 | 11,1690 | 13,3800 | 13,4240 |
| | | Mediana | 6,0 | 6,0 | 6,0 | 6,0 | 6,0 | 6,0 | 8,0 | 8,0 | 8,0 |
| | Choque = 2 | Média | 4,2023 | 3,8720 | 3,8738 | 4,4088 | 4,4398 | 4,4542 | 5,3705 | 5,7142 | 5,7840 |
| | | Desvio | 4,9311 | 4,4119 | 4,4058 | 4,9300 | 5,0601 | 5,0711 | 4,8990 | 5,5452 | 5,6578 |
| | | Mediana | 2,0 | 2,0 | 2,0 | 3,0 | 3,0 | 3,0 | 4,0 | 4,0 | 4,0 |
| | Choque = 3 | Média | 2,2007 | 2,1035 | 2,1045 | 2,4290 | 2,4163 | 2,4235 | 2,9010 | 2,9807 | 2,9945 |
| | | Desvio | 2,2382 | 2,1207 | 2,1231 | 2,5845 | 2,5628 | 2,5745 | 2,5148 | 2,6782 | 2,6848 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 |
| | Choque = 4 | Média | 1,5390 | 1,4978 | 1,4970 | 1,6540 | 1,6530 | 1,6532 | 1,8568 | 1,8968 | 1,9033 |
| | | Desvio | 1,2351 | 1,1730 | 1,1727 | 1,4008 | 1,4075 | 1,4070 | 1,3855 | 1,4564 | 1,4622 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| Uma média muda em 0,5 e outra muda ordenadamente | Choque = 0,5 | Média | 12,5280 | 11,3000 | 11,3080 | 11,5110 | 11,6110 | 11,6410 | 13,5530 | 14,6760 | 14,7990 |
| | | Desvio | 11,9380 | 11,1710 | 11,2240 | 11,4150 | 11,8920 | 11,9250 | 13,3690 | 15,4690 | 15,5710 |
| | | Mediana | 9,0 | 8,0 | 8,0 | 8,0 | 8,0 | 8,0 | 9,0 | 10,0 | 10,0 |
| | Choque = 1 | Média | 9,4800 | 8,5580 | 8,5600 | 8,4180 | 8,5790 | 8,5840 | 9,4900 | 10,0930 | 10,2070 |
| | | Desvio | 10,2360 | 9,4460 | 9,4480 | 8,8440 | 9,3840 | 9,3080 | 9,6130 | 10,9370 | 11,1180 |
| | | Mediana | 6,0 | 6,0 | 6,0 | 6,0 | 6,0 | 6,0 | 6,0 | 7,0 | 7,0 |
| | Choque = 2 | Média | 4,7080 | 4,3133 | 4,3175 | 4,3248 | 4,3800 | 4,3805 | 4,4548 | 4,7563 | 4,7803 |
| | | Desvio | 5,6302 | 5,1334 | 5,1515 | 4,8871 | 5,1127 | 5,1089 | 4,2348 | 4,9560 | 5,0039 |
| | | Mediana | 3,0 | 2,0 | 2,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 |
| | Choque = 3 | Média | 2,5182 | 2,3763 | 2,3812 | 2,4412 | 2,4565 | 2,4630 | 2,5135 | 2,6017 | 2,6125 |
| | | Desvio | 2,7362 | 2,4935 | 2,5186 | 2,5581 | 2,6732 | 2,6817 | 2,1355 | 2,2743 | 2,2797 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 |
| | Choque = 4 | Média | 1,6350 | 1,5763 | 1,5787 | 1,6002 | 1,6022 | 1,6028 | 1,6512 | 1,6640 | 1,6663 |
| | | Desvio | 1,4106 | 1,3167 | 1,3199 | 1,3332 | 1,3572 | 1,3572 | 1,1790 | 1,2037 | 1,2051 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |

Legenda: **Teo:** ARL para o Limite de Controle teórico; **Pb5:** ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com b=5; **Bes:** ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

Tabela 5.10: *ARL in control e out of control* de processos com distribuição normal bivariada - ($\rho = 0$) e $m=50$ e 200 .

| Distribuição normal bivariada com $\rho = 0$ e $m=50$ e 200 | Medidas Descritivas | m=50 | | | | | | m=200 | | | | | | |
|---|---------------------|---|---------|---------|---------------|---------|---------|---|---------|---------|---------------|---------|---------|---------|
| | | Estatística T ² de Hotelling | | | Estatística M | | | Estatística T ² de Hotelling | | | Estatística M | | | |
| | | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | |
| ARL in control | Média | 21,1400 | 14,1270 | 14,3070 | 15,8210 | 16,2870 | 16,7220 | 17,9220 | 15,9820 | 15,9780 | 17,0560 | 17,0020 | 17,0230 | |
| | Desvio | 24,9590 | 17,2940 | 17,4200 | 17,1270 | 18,8640 | 19,1420 | 18,7410 | 17,2850 | 17,2640 | 17,8350 | 17,4680 | 17,4150 | |
| | Mediana | 13,0 | 9,0 | 9,0 | 10,0 | 10,0 | 10,0 | 12,0 | 11,0 | 11,0 | 11,0 | 11,0 | 11,0 | |
| Todas as médias mudam simultaneamente na mesma escala | Choque = 0,5 | Média | 13,4350 | 9,4350 | 9,5320 | 10,7180 | 11,1250 | 11,3000 | 11,7380 | 10,4420 | 10,4450 | 11,3830 | 11,4240 | 11,4920 |
| | | Desvio | 16,5380 | 10,6480 | 10,7720 | 11,9360 | 12,5920 | 12,6900 | 11,1380 | 10,1750 | 10,1730 | 11,3480 | 11,3780 | 11,4160 |
| | | Mediana | 8,0 | 6,0 | 6,0 | 7,0 | 7,0 | 7,0 | 8,0 | 7,0 | 7,0 | 8,0 | 8,0 | 8,0 |
| | Choque = 1 | Média | 6,1980 | 4,6910 | 4,7320 | 5,2740 | 5,4020 | 5,4950 | 5,1480 | 4,7780 | 4,7780 | 5,2800 | 5,3180 | 5,3440 |
| | | Desvio | 6,6760 | 4,8760 | 4,9420 | 5,4110 | 5,8090 | 5,8550 | 4,8580 | 4,5410 | 4,5410 | 4,9500 | 5,0530 | 5,0670 |
| | | Mediana | 4,0 | 3,0 | 3,0 | 3,0 | 3,0 | 4,0 | 4,0 | 3,0 | 3,0 | 4,0 | 4,0 | 4,0 |
| | Choque = 2 | Média | 1,7955 | 1,6005 | 1,6045 | 1,7880 | 1,8025 | 1,8305 | 1,7010 | 1,6130 | 1,6125 | 1,7740 | 1,7625 | 1,7655 |
| | | Desvio | 1,3321 | 1,0661 | 1,0685 | 1,2597 | 1,3260 | 1,3578 | 1,1934 | 1,0563 | 1,0573 | 1,2388 | 1,2078 | 1,2088 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| | Choque = 3 | Média | 1,0960 | 1,0725 | 1,0725 | 1,1085 | 1,1140 | 1,1180 | 1,0875 | 1,0775 | 1,0780 | 1,1150 | 1,1085 | 1,1090 |
| | | Desvio | 0,3344 | 0,2780 | 0,2780 | 0,3766 | 0,3835 | 0,3926 | 0,3129 | 0,2925 | 0,2932 | 0,3603 | 0,3447 | 0,3452 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| Choque = 4 | Média | 1,0120 | 1,0080 | 1,0085 | 1,0150 | 1,0135 | 1,0140 | 1,0080 | 1,0085 | 1,0085 | 1,0140 | 1,0135 | 1,0135 | |
| | Desvio | 0,1089 | 0,0891 | 0,0918 | 0,1256 | 0,1197 | 0,1217 | 0,0891 | 0,0918 | 0,0918 | 0,1175 | 0,1154 | 0,1154 | |
| | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| Uma média permanece constante e a outra muda ordenadamente | Choque = 0,5 | Média | 17,6160 | 12,2890 | 12,4530 | 13,5380 | 14,0350 | 14,4290 | 14,9490 | 13,0620 | 13,0620 | 14,3280 | 14,4520 | 14,4740 |
| | | Desvio | 20,8180 | 14,6120 | 14,5400 | 14,3480 | 16,2720 | 16,5820 | 14,7210 | 12,9400 | 12,9390 | 14,3980 | 14,6700 | 14,7250 |
| | | Mediana | 11,0 | 7,0 | 7,0 | 9,0 | 8,0 | 9,0 | 10,0 | 9,0 | 9,0 | 9,0 | 9,0 | 9,0 |
| | Choque = 1 | Média | 11,6590 | 8,6240 | 8,6700 | 9,3710 | 9,6940 | 9,9770 | 10,7420 | 9,6160 | 9,6340 | 10,3540 | 10,3080 | 10,3340 |
| | | Desvio | 14,2730 | 11,0790 | 11,0650 | 10,7600 | 11,5520 | 11,8560 | 11,5920 | 10,3750 | 10,3950 | 10,5560 | 10,4930 | 10,5670 |
| | | Mediana | 7,0 | 5,0 | 5,0 | 6,0 | 6,0 | 6,0 | 7,0 | 6,0 | 6,0 | 7,0 | 7,0 | 7,0 |
| | Choque = 2 | Média | 5,5590 | 4,3700 | 4,3950 | 4,6990 | 4,8620 | 4,9950 | 4,9960 | 4,5620 | 4,5620 | 4,8000 | 4,7580 | 4,7650 |
| | | Desvio | 7,6590 | 6,3810 | 6,4160 | 6,5600 | 7,0660 | 7,3470 | 6,1110 | 5,4780 | 5,4780 | 5,5610 | 5,5790 | 5,5760 |
| | | Mediana | 3,0 | 2,0 | 2,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 |
| | Choque = 3 | Média | 2,8085 | 2,3185 | 2,3350 | 2,4820 | 2,5005 | 2,5470 | 2,5985 | 2,3915 | 2,3995 | 2,5670 | 2,5940 | 2,6060 |
| | | Desvio | 3,2384 | 2,4114 | 2,4396 | 2,7207 | 2,7629 | 2,8409 | 2,8673 | 2,4788 | 2,4898 | 2,7367 | 2,8620 | 2,9072 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| Choque = 4 | Média | 1,8000 | 1,6215 | 1,6280 | 1,6815 | 1,6885 | 1,7130 | 1,6995 | 1,6315 | 1,6305 | 1,6350 | 1,6350 | 1,6360 | |
| | Desvio | 1,6187 | 1,3598 | 1,3782 | 1,5599 | 1,5260 | 1,6081 | 1,5084 | 1,3604 | 1,3572 | 1,3721 | 1,3863 | 1,3862 | |
| | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| Uma média muda em 0,5 e outra muda ordenadamente | Choque = 0,5 | Média | 13,5250 | 9,4460 | 9,5310 | 11,1300 | 11,4770 | 11,7520 | 11,5970 | 10,2990 | 10,2900 | 11,4440 | 11,4630 | 11,4920 |
| | | Desvio | 15,3050 | 10,7560 | 10,9340 | 11,9800 | 13,6720 | 13,8100 | 11,6100 | 10,4540 | 10,4420 | 11,2860 | 11,5020 | 11,5380 |
| | | Mediana | 9,0 | 6,0 | 6,0 | 7,0 | 7,0 | 7,0 | 8,0 | 7,0 | 7,0 | 8,0 | 8,0 | 8,0 |
| | Choque = 1 | Média | 9,7200 | 7,2050 | 7,2330 | 8,3070 | 8,4900 | 8,6850 | 8,6100 | 7,6840 | 7,6900 | 8,3850 | 8,3660 | 8,4100 |
| | | Desvio | 11,0800 | 8,1150 | 8,1150 | 9,8170 | 10,1840 | 10,4130 | 8,6220 | 7,6550 | 7,6640 | 8,1670 | 8,2370 | 8,2940 |
| | | Mediana | 6,0 | 5,0 | 5,0 | 5,0 | 5,0 | 5,0 | 6,0 | 5,0 | 5,0 | 6,0 | 6,0 | 6,0 |
| | Choque = 2 | Média | 4,7240 | 3,7140 | 3,7250 | 4,3130 | 4,3530 | 4,4610 | 4,2260 | 3,9205 | 3,9240 | 4,2090 | 4,1950 | 4,2080 |
| | | Desvio | 6,2470 | 4,5380 | 4,5340 | 5,3760 | 5,5020 | 5,6810 | 4,6760 | 4,3633 | 4,3637 | 4,6990 | 4,6540 | 4,6720 |
| | | Mediana | 3,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 |
| | Choque = 3 | Média | 2,7565 | 2,3135 | 2,3215 | 2,5305 | 2,6135 | 2,6645 | 2,3860 | 2,2320 | 2,2345 | 2,3795 | 2,3775 | 2,3810 |
| | | Desvio | 3,4659 | 2,5388 | 2,5527 | 2,8187 | 3,1677 | 3,2144 | 2,4427 | 2,2079 | 2,2084 | 2,5147 | 2,4962 | 2,5034 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| Choque = 4 | Média | 1,7840 | 1,5860 | 1,5870 | 1,6285 | 1,6420 | 1,6605 | 1,6945 | 1,6300 | 1,6315 | 1,6735 | 1,6880 | 1,6885 | |
| | Desvio | 1,6991 | 1,3162 | 1,3154 | 1,3555 | 1,3892 | 1,4287 | 1,4840 | 1,3878 | 1,3884 | 1,5053 | 1,5076 | 1,5088 | |
| | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |

Legenda: Teo: ARL para o Limite de Controle teórico; Pb5: ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com b=5; Bes: ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

5.2.2 – Processos Multivariados com $p=20$ Variáveis

Apresentamos nas Tabela 5.11 a 5.14 os valores do *ARL in control* e *ARL out of control* para processos com $p=20$ variáveis e amostras de tamanho $m=50, 100, 200$ e 1000 .

É possível observar que para $m=50$ os valores dos *ARL in control* são muito baixos para a estatística *T² de Hotelling* o que pode ser devido ao fato m ser igual a 50. Já a primeira

componente principal tem um valor muito elevado tanto do *ARL in control* quanto do *ARL out of control*, indicando que mesmo com o processo sofrendo modificações a carta com esta estatística não consegue identificar com rapidez a falta de controle do processo. A estatística *M* apresenta valores de *ARL in control* e *ARL out of control* não tão baixos como o da estatística *T² de Hotelling* e não tão altos como a primeira componente principal mostrando que deveria ser a preferida para monitoramento do processo multivariado com $p=20$ variáveis. As cartas de controle construídas usando os métodos de Polansky e Baker (2000) com $b=5$ e de Bessegato (2001) para determinação da janela ótima forneceram valores bem semelhantes.

Em geral, é possível observar que na maioria das vezes o valor do *ARL in control* para as estatísticas *T²* e *M* aumenta e da primeira componente principal diminui à medida que aumentamos os tamanhos das amostras processo que também ocorre para os valores do *ARL out of control*. Os valores dos *ARL out of control* são menores para aqueles “choques” que possuem maior intensidade (a média do processo se desloca em um valor igual a 4), sendo que o choque mais facilmente percebido é aquele em que as médias de todas as variáveis do processo mudam simultaneamente.

Tabela 5.11: ARL in control e out of control de processos com distribuição normal p-variada ($p=20$) e $m=50$.

| Distribuição normal multivariada $p=20$ com ρ misturado e $m=50$ | | Medidas Descritivas | Estatística T^2 de Hotelling | | | Estatística M | | | Componentes Principais | | | |
|---|--|--------------------------|--------------------------------|---------|--------|-----------------|---------|---------|------------------------|---------|---------|---------|
| | | | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | |
| <i>ARL in control</i> | | Média | 24,1840 | 1,5660 | 1,5700 | 12,6480 | 11,8110 | 12,0830 | 29,2050 | 58,3070 | 61,8780 | |
| | | Desvio | 30,2250 | 0,9222 | 0,9252 | 12,6246 | 13,4732 | 13,6760 | 29,9677 | 79,9639 | 86,4213 | |
| | | Mediana | 14,0 | 1,0 | 1,0 | 9,0 | 8,0 | 8,0 | 20,0 | 34,0 | 36,0 | |
| Todas as médias mudam simultaneamente na mesma escala | Choque = 0,5 | Média | 15,5910 | 1,3850 | 1,3870 | 6,4760 | 6,0360 | 6,2030 | 17,0970 | 35,6580 | 38,3560 | |
| | | Desvio | 20,7353 | 0,7371 | 0,7374 | 5,7643 | 5,7390 | 5,8929 | 20,8112 | 51,0337 | 57,8211 | |
| | | Mediana | 9,0 | 1,0 | 1,0 | 5,0 | 4,0 | 4,0 | 10,0 | 17,0 | 18,0 | |
| | Choque = 1 | Média | 5,8110 | 1,1420 | 1,1430 | 2,7290 | 2,5760 | 2,5980 | 9,1800 | 17,4670 | 18,7740 | |
| | | Desvio | 6,4983 | 0,4243 | 0,4251 | 2,2910 | 2,2807 | 2,2973 | 16,3612 | 37,5992 | 42,3682 | |
| | | Mediana | 3,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 | 4,0 | 6,0 | 6,0 | |
| | Choque = 2 | Média | 1,3210 | 1,0020 | 1,0020 | 1,0820 | 1,0740 | 1,0750 | 4,8040 | 8,5940 | 8,8750 | |
| | | Desvio | 0,7256 | 0,0447 | 0,0447 | 0,2956 | 0,2909 | 0,2923 | 11,7533 | 27,2806 | 27,6556 | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | |
| | Choque = 3 | Média | 1,0030 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 3,5450 | 6,0150 | 6,4050 | |
| | | Desvio | 0,0547 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 12,1486 | 22,6606 | 24,4909 | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| | Choque = 4 | Média | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 2,6810 | 4,7340 | 5,0400 | |
| | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 7,5313 | 18,6267 | 20,1367 | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| | Uma média permanece constante e a outra muda ordenadamente | Choque = 0,5 | Média | 22,8750 | 1,4850 | 1,5000 | 12,2470 | 11,4380 | 11,6540 | 27,2160 | 51,1440 | 54,2040 |
| | | | Desvio | 27,4730 | 0,8380 | 0,8590 | 13,2020 | 14,2290 | 14,3060 | 27,8830 | 62,8170 | 66,5790 |
| | | | Mediana | 13,0 | 1,0 | 1,0 | 8,5 | 7,0 | 7,0 | 18,5 | 31,0 | 32,0 |
| Choque = 1 | | Média | 20,0240 | 1,4350 | 1,4410 | 10,2790 | 9,2110 | 9,4100 | 27,6820 | 58,1270 | 61,0080 | |
| | | Desvio | 25,7350 | 0,8050 | 0,8110 | 10,7100 | 11,3320 | 11,5180 | 30,3830 | 83,8880 | 85,7550 | |
| | | Mediana | 12,0 | 1,0 | 1,0 | 7,0 | 6,0 | 6,0 | 19,0 | 31,0 | 34,0 | |
| Choque = 2 | | Média | 10,2370 | 1,2390 | 1,2430 | 4,3200 | 4,0890 | 4,1600 | 28,0970 | 55,8140 | 59,2450 | |
| | | Desvio | 13,5860 | 0,5880 | 0,5920 | 4,2210 | 4,4010 | 4,4490 | 30,8940 | 75,2090 | 80,0810 | |
| | | Mediana | 6,0 | 1,0 | 1,0 | 3,0 | 2,0 | 3,0 | 19,0 | 30,0 | 32,0 | |
| Choque = 3 | | Média | 4,9110 | 1,1070 | 1,1040 | 1,8040 | 1,7710 | 1,7800 | 26,7650 | 55,0650 | 57,7700 | |
| | | Desvio | 5,2110 | 0,3680 | 0,3620 | 1,2230 | 1,2390 | 1,2650 | 28,0090 | 85,6680 | 87,3350 | |
| | | Mediana | 3,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 18,0 | 30,5 | 32,0 | |
| Choque = 4 | | Média | 2,5650 | 1,0310 | 1,0310 | 1,1560 | 1,1320 | 1,1400 | 26,8070 | 55,2900 | 59,3930 | |
| | | Desvio | 2,6910 | 0,1850 | 0,1850 | 0,4470 | 0,4010 | 0,4320 | 27,4900 | 79,6860 | 83,2610 | |
| | | Mediana | 2,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 18,0 | 29,0 | 31,0 | |
| 5 médias mudam 15 ficam constantes | | Choque = 0,5, 3, 1, 1, 1 | Média | 2,9340 | 1,0260 | 1,0290 | 1,4870 | 1,4290 | 1,4500 | 9,3380 | 16,8900 | 17,6730 |
| | | | Desvio | 3,0970 | 0,1713 | 0,1794 | 0,9427 | 0,8842 | 0,9211 | 13,8749 | 32,8995 | 33,8047 |
| | | | Mediana | 2,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 5,0 | 7,0 | 7,0 |

Legenda: **Teo:** ARL para o Limite de Controle teórico; **Pb5:** ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes:** ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

Tabela 5.12: *ARL in control* e *out of control* de processos com distribuição normal p-variada ($p=20$) e $m=100$.

| Distribuição normal multivariada $p=20$ com ρ misturado e $m=100$ | | Medidas Descritivas | Estatística T^2 de Hotelling | | | Estatística M | | | Componentes Principais | | |
|--|--------------------------|---------------------|--------------------------------|--------|--------|---------------|---------|---------|------------------------|---------|---------|
| | | | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes |
| ARL in control | | Média | 19,3630 | 3,1880 | 3,1880 | 14,3250 | 13,8620 | 14,0340 | 24,4140 | 36,6050 | 37,5520 |
| | | Desvio | 20,6248 | 2,6739 | 2,6642 | 13,9447 | 15,0266 | 15,4700 | 23,8541 | 40,0085 | 40,1369 |
| | | Mediana | 12,0 | 2,0 | 2,0 | 9,0 | 9,0 | 9,0 | 17,0 | 24,0 | 25,0 |
| Todas as médias mudam simultaneamente na mesma escala | Choque = 0,5 | Média | 12,1220 | 2,5430 | 2,5850 | 7,6270 | 7,2660 | 7,3310 | 14,5390 | 21,2540 | 21,8910 |
| | | Desvio | 12,3611 | 2,0563 | 2,1296 | 7,7774 | 7,3865 | 7,4124 | 15,6560 | 27,0016 | 27,4973 |
| | | Mediana | 8,0 | 2,0 | 2,0 | 5,0 | 5,0 | 5,0 | 9,0 | 13,0 | 13,0 |
| | Choque = 1 | Média | 4,2190 | 1,5220 | 1,5290 | 2,9460 | 2,8200 | 2,8420 | 7,8420 | 10,6510 | 10,7870 |
| | | Desvio | 4,0997 | 0,8901 | 0,8944 | 2,4362 | 2,3944 | 2,4118 | 12,3700 | 16,1987 | 16,2873 |
| | | Mediana | 3,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 | 4,0 | 5,0 | 5,0 |
| | Choque = 2 | Média | 1,1000 | 1,0140 | 1,0150 | 1,0870 | 1,0860 | 1,0860 | 4,3850 | 5,2440 | 5,3840 |
| | | Desvio | 0,3466 | 0,1258 | 0,1296 | 0,2992 | 0,2944 | 0,2944 | 14,0459 | 14,2483 | 14,4803 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 |
| | Choque = 3 | Média | 1,0010 | 1,0000 | 1,0000 | 1,0010 | 1,0000 | 1,0000 | 2,4960 | 3,0310 | 3,1180 |
| | | Desvio | 0,0316 | 0,0000 | 0,0000 | 0,0316 | 0,0000 | 0,0000 | 5,4227 | 7,0091 | 7,4658 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| | Choque = 4 | Média | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 2,1510 | 2,8200 | 2,8380 |
| | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 5,6907 | 9,1162 | 9,1854 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| Uma média permanece constante e a outra muda ordenadamente | Choque = 0,5 | Média | 19,0760 | 3,3180 | 3,3230 | 14,1540 | 13,6760 | 13,8360 | 24,8720 | 38,5540 | 39,3100 |
| | | Desvio | 20,1229 | 2,7325 | 2,7335 | 14,0618 | 15,7476 | 15,8416 | 25,3274 | 43,0629 | 43,5137 |
| | | Mediana | 13,0 | 2,0 | 2,0 | 10,0 | 9,0 | 9,0 | 17,5 | 23,5 | 24,0 |
| | Choque = 1 | Média | 14,4850 | 2,8490 | 2,8630 | 11,2500 | 10,5020 | 10,5950 | 24,6270 | 36,6900 | 37,9180 |
| | | Desvio | 16,1597 | 2,3417 | 2,3537 | 10,8901 | 11,0531 | 11,1196 | 24,6502 | 38,5815 | 39,5582 |
| | | Mediana | 9,0 | 2,0 | 2,0 | 8,0 | 7,0 | 7,0 | 17,0 | 25,0 | 25,0 |
| | Choque = 2 | Média | 6,9290 | 2,0850 | 2,0930 | 4,4970 | 4,2880 | 4,3470 | 23,8440 | 34,9550 | 36,6250 |
| | | Desvio | 6,7404 | 1,5499 | 1,5585 | 4,2594 | 4,0422 | 4,1016 | 24,0924 | 38,4824 | 40,5850 |
| | | Mediana | 5,0 | 2,0 | 2,0 | 3,0 | 3,0 | 3,0 | 16,0 | 23,0 | 24,0 |
| | Choque = 3 | Média | 3,2050 | 1,3560 | 1,3610 | 1,7650 | 1,7670 | 1,7670 | 23,6460 | 34,7580 | 36,0180 |
| | | Desvio | 3,1754 | 0,6984 | 0,7079 | 1,1667 | 1,2346 | 1,2346 | 25,2346 | 40,1648 | 42,1027 |
| | | Mediana | 2,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 16,0 | 21,0 | 22,0 |
| | Choque = 4 | Média | 1,7510 | 1,1250 | 1,1260 | 1,1690 | 1,1700 | 1,1680 | 23,1370 | 33,5270 | 35,0330 |
| | | Desvio | 1,2444 | 0,3654 | 0,3664 | 0,4320 | 0,4374 | 0,4335 | 23,3510 | 39,4432 | 41,8717 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 15,0 | 21,0 | 22,0 |
| 5 médias mudam 5 ficam constantes | Choque = 0,5, 3, 1, 1, 1 | Média | 1,9970 | 1,1580 | 1,1580 | 1,5060 | 1,4730 | 1,4780 | 8,1020 | 11,0070 | 11,2470 |
| | | Desvio | 1,5854 | 0,4280 | 0,4280 | 0,9267 | 0,8866 | 0,8980 | 11,5550 | 16,4180 | 16,8067 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 4,0 | 5,0 | 5,0 |

Legenda: **Teo:** ARL para o Limite de Controle teórico; **Pb5:** ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes:** ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

Tabela 5.13: *ARL in control* e *out of control* de processos com distribuição normal p-variada ($p=20$) e $m=200$.

| Distribuição normal multivariada $p=20$ com ρ misturado e $m=200$ | | Medidas Descritivas | Estatística T ² de Hotelling | | | Estatística M | | | Componentes Principais | | | |
|--|--|--------------------------|---|---------|--------|---------------|---------|---------|------------------------|---------|---------|---------|
| | | | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | |
| ARL in control | | Média | 19,2610 | 6,7950 | 6,8120 | 16,9980 | 16,7640 | 16,8750 | 21,6540 | 27,8760 | 28,0150 | |
| | | Desvio | 19,9597 | 6,1468 | 6,1515 | 16,1961 | 17,5336 | 17,6404 | 21,5030 | 30,8440 | 30,8018 | |
| | | Mediana | 13,0 | 5,0 | 5,0 | 12,0 | 11,0 | 11,0 | 14,5 | 17,5 | 18,0 | |
| Todas as médias mudam simultaneamente na mesma escala | Choque = 0,5 | Média | 10,7930 | 4,6070 | 4,6280 | 8,3660 | 8,0190 | 8,0800 | 13,6840 | 16,5660 | 16,6750 | |
| | | Desvio | 10,1614 | 4,2653 | 4,3061 | 7,7264 | 7,5591 | 7,6109 | 14,8772 | 18,6682 | 18,7331 | |
| | | Mediana | 8,0 | 3,0 | 3,0 | 6,0 | 6,0 | 6,0 | 9,0 | 10,0 | 10,0 | |
| | Choque = 1 | Média | 3,2990 | 1,9790 | 1,9810 | 3,0080 | 2,8890 | 2,9050 | 7,4000 | 8,7470 | 8,7710 | |
| | | Desvio | 2,7823 | 1,3551 | 1,3581 | 2,4190 | 2,2978 | 2,3318 | 9,4977 | 11,8770 | 11,8829 | |
| | | Mediana | 2,0 | 1,5 | 1,5 | 2,0 | 2,0 | 2,0 | 4,0 | 5,0 | 5,0 | |
| | Choque = 2 | Média | 1,0710 | 1,0200 | 1,0200 | 1,0870 | 1,0870 | 1,0870 | 3,0630 | 3,8700 | 3,8780 | |
| | | Desvio | 0,2684 | 0,1401 | 0,1401 | 0,2924 | 0,2958 | 0,2958 | 6,7022 | 10,8193 | 10,8525 | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| | Choque = 3 | Média | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 2,1550 | 2,4730 | 2,4790 | |
| | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 5,3976 | 6,9976 | 7,0006 | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| | Choque = 4 | Média | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,7540 | 1,9810 | 1,9810 | |
| | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 3,9057 | 6,5110 | 6,5110 | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| | Uma média permanece constante e a outra muda ordenadamente | Choque = 0,5 | Média | 17,9370 | 6,5970 | 6,6040 | 16,0350 | 15,5180 | 15,5660 | 21,5340 | 27,4440 | 27,5970 |
| | | | Desvio | 17,7587 | 6,1808 | 6,1853 | 16,6625 | 17,6106 | 17,6809 | 19,3413 | 28,5516 | 28,5469 |
| | | | Mediana | 12,0 | 5,0 | 5,0 | 11,0 | 10,0 | 10,0 | 16,0 | 20,0 | 20,0 |
| Choque = 1 | | Média | 14,1490 | 5,5820 | 5,5900 | 12,1450 | 11,6030 | 11,6110 | 22,8820 | 28,6260 | 29,0120 | |
| | | Desvio | 14,0954 | 5,1441 | 5,1469 | 12,0237 | 11,6977 | 11,6988 | 22,4886 | 28,4745 | 28,8277 | |
| | | Mediana | 10,0 | 4,0 | 4,0 | 8,0 | 8,0 | 8,0 | 16,0 | 19,0 | 20,0 | |
| Choque = 2 | | Média | 6,5280 | 3,0670 | 3,0850 | 4,6170 | 4,5330 | 4,5480 | 22,9250 | 27,9060 | 28,1120 | |
| | | Desvio | 6,0501 | 2,3783 | 2,4066 | 4,3117 | 4,1550 | 4,1842 | 22,4412 | 28,5800 | 28,7888 | |
| | | Mediana | 5,0 | 2,0 | 2,0 | 3,0 | 3,0 | 3,0 | 16,5 | 19,0 | 19,0 | |
| Choque = 3 | | Média | 2,6920 | 1,7330 | 1,7350 | 1,8260 | 1,7930 | 1,7980 | 21,9090 | 27,3940 | 27,6890 | |
| | | Desvio | 2,1535 | 1,2415 | 1,2427 | 1,2990 | 1,2907 | 1,2942 | 21,6299 | 32,1408 | 32,6652 | |
| | | Mediana | 2,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 15,0 | 18,0 | 18,0 | |
| Choque = 4 | | Média | 1,5270 | 1,2130 | 1,2140 | 1,1650 | 1,1650 | 1,1650 | 21,2460 | 27,9920 | 28,5860 | |
| | | Desvio | 0,8900 | 0,4815 | 0,4821 | 0,4218 | 0,4335 | 0,4335 | 21,1014 | 31,2170 | 32,0458 | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 14,0 | 18,0 | 18,0 | |
| 5 médias mudam 15 ficam constantes | | Choque = 0,5, 3, 1, 1, 1 | Média | 1,7330 | 1,3160 | 1,3160 | 1,5340 | 1,5140 | 1,5140 | 7,1250 | 8,2600 | 8,4190 |
| | | | Desvio | 1,1684 | 0,6531 | 0,6531 | 0,9304 | 0,9136 | 0,9136 | 9,3850 | 11,5315 | 12,2353 |
| | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 4,0 | 5,0 | 5,0 |

Legenda: **Teo:** ARL para o Limite de Controle teórico; **Pb5:** ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes:** ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

Tabela 5.14: *ARL in control* e *out of control* de processos com distribuição normal *p*-variada ($p=20$) e $m=1000$.

| Distribuição normal multivariada com ρ misturado e $p=20$ ($n=1000$) | | Medidas Descritivas | Estatística T^2 de Hotelling | | | Estatística M | | | Componentes Principais | | | |
|---|---------|--------------------------|--------------------------------|---------|---------|-----------------|---------|---------|------------------------|---------|---------|---------|
| | | | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | |
| ARL in control | | Média | 20,5110 | 14,7960 | 14,7960 | 18,6100 | 17,9190 | 17,9190 | 22,0270 | 23,3380 | 23,4290 | |
| | | Desvio | 20,7929 | 15,2664 | 15,2664 | 17,7711 | 17,5541 | 17,5541 | 21,2179 | 23,5546 | 23,6064 | |
| | | Mediana | 14,0 | 10,0 | 10,0 | 13,0 | 12,0 | 12,0 | 15,0 | 16,0 | 16,0 | |
| Todas as médias mudam simultaneamente na mesma escala | | Choque = 0,5 | Média | 9,5980 | 7,3410 | 7,3410 | 8,8790 | 8,7370 | 8,7520 | 11,8080 | 12,2720 | 12,3140 |
| | | | Desvio | 9,1581 | 7,0823 | 7,0823 | 8,6302 | 8,6490 | 8,6745 | 12,0389 | 12,2151 | 12,2999 |
| | | | Mediana | 7,0 | 5,0 | 5,0 | 6,0 | 6,0 | 6,0 | 8,0 | 8,0 | 8,0 |
| | | Choque = 1 | Média | 3,2050 | 2,7410 | 2,7410 | 3,0370 | 2,9810 | 2,9810 | 5,0940 | 5,4080 | 5,4120 |
| | | | Desvio | 2,5969 | 2,0804 | 2,0804 | 2,5668 | 2,5654 | 2,5654 | 5,1546 | 5,8931 | 5,8955 |
| | | | Mediana | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 4,0 | 4,0 | 4,0 |
| | | Choque = 2 | Média | 1,0410 | 1,0330 | 1,0330 | 1,0990 | 1,0920 | 1,0920 | 1,7090 | 1,7310 | 1,7310 |
| | | | Desvio | 0,2034 | 0,1842 | 0,1842 | 0,3276 | 0,3125 | 0,3125 | 1,2985 | 1,3403 | 1,3403 |
| | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| | | Choque = 3 | Média | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,1720 | 1,1420 | 1,1420 |
| | | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 1,3814 | 0,5178 | 0,5178 |
| | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| Choque = 4 | Média | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0290 | 1,0270 | 1,0270 | | |
| | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,2150 | 0,2008 | 0,2008 | | |
| | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | | |
| Uma média permanece constante e a outra muda ordenadamente | | Choque = 0,5 | Média | 16,8870 | 12,7320 | 12,7440 | 16,9400 | 16,4500 | 16,4500 | 19,9110 | 20,9080 | 20,9320 |
| | | | Desvio | 16,5400 | 12,8867 | 12,9047 | 16,1656 | 16,0408 | 16,0408 | 20,1730 | 21,2669 | 21,3200 |
| | | | Mediana | 12,0 | 8,5 | 9,0 | 12,0 | 11,0 | 11,0 | 14,0 | 14,0 | 14,0 |
| | | Choque = 1 | Média | 11,8460 | 9,4120 | 9,4120 | 13,4950 | 13,0340 | 13,0340 | 20,6260 | 22,0630 | 22,0830 |
| | | | Desvio | 11,1399 | 8,7509 | 8,7509 | 13,4607 | 13,1748 | 13,1748 | 19,6463 | 22,4478 | 22,4486 |
| | | | Mediana | 8,0 | 7,0 | 7,0 | 9,0 | 9,0 | 9,0 | 15,0 | 15,0 | 15,0 |
| | | Choque = 2 | Média | 5,8790 | 4,8890 | 4,8890 | 4,6610 | 4,5580 | 4,5580 | 19,5360 | 21,1190 | 21,1530 |
| | | | Desvio | 5,4677 | 4,4064 | 4,4064 | 4,0105 | 4,0575 | 4,0575 | 18,6685 | 21,2533 | 21,2643 |
| | | | Mediana | 4,0 | 4,0 | 4,0 | 3,0 | 3,0 | 3,0 | 14,0 | 14,0 | 14,0 |
| | | Choque = 3 | Média | 2,5600 | 2,2360 | 2,2360 | 1,8700 | 1,8540 | 1,8540 | 20,8860 | 21,4190 | 21,5140 |
| | | | Desvio | 2,0136 | 1,6053 | 1,6053 | 1,3337 | 1,3094 | 1,3094 | 20,0452 | 20,7235 | 20,8105 |
| | | | Mediana | 2,0 | 2,0 | 2,0 | 1,0 | 1,0 | 1,0 | 15,0 | 15,0 | 15,0 |
| Choque = 4 | Média | 1,4800 | 1,3840 | 1,3840 | 1,1570 | 1,1600 | 1,1600 | 20,6840 | 21,7770 | 21,9370 | | |
| | Desvio | 0,8150 | 0,7036 | 0,7036 | 0,4056 | 0,4130 | 0,4130 | 19,8084 | 20,8600 | 21,2943 | | |
| | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 15,0 | 16,0 | 16,0 | | |
| 5 médias mudam 15 ficam constantes | | Choque = 0,5, 3, 1, 1, 1 | Média | 1,6250 | 1,5220 | 1,5220 | 1,5550 | 1,5430 | 1,5430 | 5,7570 | 6,0210 | 6,0210 |
| | | | Desvio | 0,9702 | 0,8685 | 0,8685 | 0,9153 | 0,9050 | 0,9050 | 6,2271 | 6,5959 | 6,5959 |
| | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 4,0 | 4,0 | 4,0 |

Legenda: **Teo**: ARL para o Limite de Controle teórico; **Pb5**: ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes**: ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

5.2.3 – Conclusão

De modo geral, observa-se que a carta de controle com a estatística M de Hayter e Tsui (1994) se mostrou mais sensível na detecção da falta de controle do processo e produziu, em geral, um baixo número de alarmes falsos quando utiliza-se o núcleo estimador para estimação dos limites de controle.

Observa-se que a mudança na variância das variáveis do processo reflete em limites de controle mais amplos. Observa-se que a medida que o tamanho de amostra aumenta os valores dos *ARL in control* aumentam quando utilizamos a estatística T^2 de Hotelling (1947) e a estatística M de Hayter e Tsui (1994) para construção das cartas e diminuem quando utilizamos a primeira componente principal para construção da carta de controle. É possível observar que os valores do *ARL out of control*, para processos com $p=2$, se mantêm à medida que o tamanho da amostra aumente (quando utilizamos a estatística T^2 de Hotelling e a

estatística M de Hayter e Tsui) e tendem a diminuir quando utilizamos a primeira componente principal para construir a carta de controle para o processo. Nos casos em que $p=5, 10$ e 20 , os valores dos ARL *out of control* em geral tendem a aumentar (quando utilizamos a estatística T^2 de Hotelling e a estatística M de Hayter e Tsui) e diminuir quando utilizamos a primeira componente principal para construir a carta de controle para a análise do processo. Em geral os valores dos ARL *out of control* são menores para os “choques” mais intensos e quando esses são aplicados em todas as variáveis do processo, mostrando que essas mudanças são sempre mais facilmente detectados pelos limites de controle do processo nestes casos.

As cartas de controle construídas via núcleos estimadores, usando o método de Polansky e Baker (2000) com $b=5$ e Bessegato (2001), forneceram resultados similares, indicando que qualquer um dos dois métodos seria adequado na determinação dos limites de controle para um processo normal multivariado sem autocorrelação.

Observa-se que a técnica de núcleo estimador mesmo para amostras pequenas ($m=50$) fornece boas estimativas dos limites de controle quando a estatística M de Hayter e Tsui (1994) é usada. Este é um fato importante já que Hayter e Tsui (1994) sugerem o uso de no mínimo $N=100.000$ observações para a obtenção da constante $C_{R,\alpha}$ no caso da distribuição normal enquanto Mingoti e Gloria (2003) sugerem $N=10.000$, de acordo com o procedimento de simulação descrito no Quadro 2.1 (página 36). No entanto, usando núcleo-estimador, obtém-se boas estimativas da constante $C_{R,\alpha}$ e com um tamanho de amostra menor que a sugerida por esses autores.

**Capítulo 6 – Estratégia de Execução: Processos
Multivariados Não-Normais e Não-
Autocorrelacionados**

Neste capítulo faremos a avaliação das técnicas não-paramétricas estudadas nessa dissertação aplicadas a processos multivariados que não possuem distribuição normal e são não-autocorrelacionados. Com essa finalidade alguns processos não-normais não-autocorrelacionados foram simulados. Para cada processo foram geradas $k=1000$ amostras de tamanho $m=50, 100$ e 200 . Como feito para o processo normal multivariado será amostrado apenas o resultado das amostras de tamanho $m=50$ e 200 .

A estimação não-paramétrica dos limites de controle foi feita utilizando a metodologia de núcleo estimadores utilizando-se os métodos de Polansky e Baker (2000) com $b=1$ e $b=5$ e Bessegato (2001) e a função de distribuição empírica. A função núcleo utilizada em todos os casos foi a da distribuição normal padrão. Todos os limites de controle são obtidos à 95% de confiança. Como no capítulo 5 observamos que o método estudado por Polansky e Baker (2000) com $b=1$ fornecia valores dos limites de controle muito próximos aos valores fornecidos pelas demais técnicas de núcleo-estimador e como a função de distribuição empírica tinha sempre um pior desempenho em relação as técnicas de núcleo-estimador, não exibiremos nesse capítulo as análises do processo com essa metodologia. A seguir descrevemos os modelos não-normais simulados e apresentamos os resultados dessas simulações.

6.1 – Processo 1: O Processo Bivariado estudado por Hayter e Tsui (1994)

6.1.1 – Descrição do processo

Considere uma amostra de tamanho m do vetor $X = [X_1, X_2]'$ com distribuição normal bivariada com os parâmetros:

$$\mu = [0 \quad 0]' \text{ e } \Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Sejam as variáveis Z_1 e Z_2 definidas por:

$$Z_1 = \max(|X_1|, |X_2|)$$

$$Z_2 = X_1^2 + X_2^2$$

Dessa forma, teremos uma amostra de tamanho m do vetor aleatório $Z = [Z_1, Z_2]'$. A partir da qual definiremos os limites de controle e encontraremos os valores dos *ARL in control* e *out of control*. Esse processo foi um dos simulados por Hayter e Tsui (1994) para aplicação de controle de processos via estatística M para distribuições não-normais multivariadas não-autocorrelacionadas.

Para ilustrar o comportamento probabilístico do vetor aleatório Z foi gerada uma amostra aleatória de tamanho $m=100000$ com base na distribuição normal bivariada descrita em (6.1) e os valores do vetor aleatório Z foram calculados. Os gráficos da função densidade conjunta de Z_1 e Z_2 e das curvas de nível amostrais são apresentados nas Figuras 6.1 e 6.2.

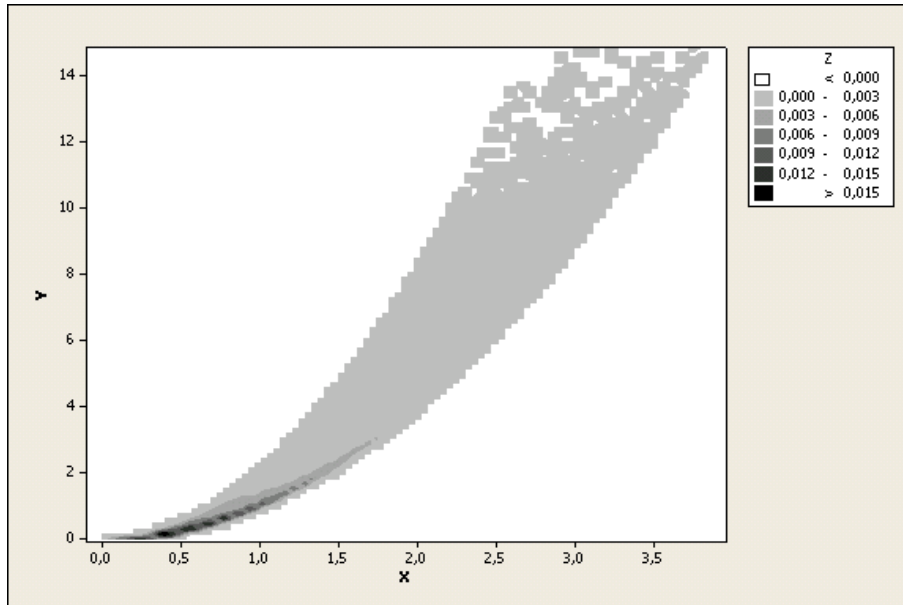


Figura 6.1: Curvas de nível amostral da função densidade conjunta de Z_1 e Z_2 .

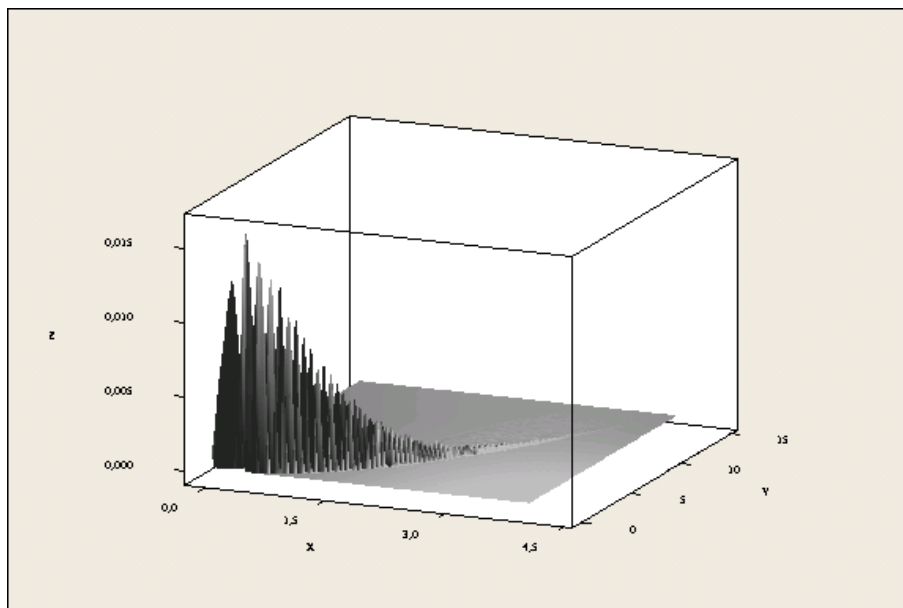


Figura 6.2: Função densidade conjunta amostral de Z_1 e Z_2 .

A forma de obtenção dos *ARL in control* e *out of control* é bem similar àquela mostrada no capítulo 4, porém, pela forma em que o vetor Z é obtido, apenas algumas mudanças no vetor de médias foram implementadas. Os “*choques*” que foram implementados para o cálculo do *ARL out of control* (implementados nas médias do vetor aleatório X) apresentam os seguintes impactos no vetor Z :

Tabela 6.1: Impacto dos choques implementados nas médias do vetor aleatório X sobre as médias do vetor aleatório Z.

| Média X_1 | Média X_2 | Média Z_1 | Média Z_2 |
|-------------|-------------|-------------|-------------|
| 0,0 | 0,0 | 1,12876 | 2,0007 |
| 0,5 | 0,5 | 1,26391 | 2,5001 |
| 1,0 | 1,0 | 1,61367 | 3,9971 |
| 2,0 | 2,0 | 2,56582 | 10,0108 |
| 0,0 | 0,5 | 1,19866 | 2,2519 |
| 0,0 | 1,0 | 1,39874 | 3,0001 |
| 0,0 | 2,0 | 2,09888 | 5,9921 |

Observa-se na Tabela 6.1 que ao aplicar-se “choques” de valor igual a 2 na variável X_2 a variável Z_2 já sofre alto impacto na sua média, o que explica o fato de termos utilizados “choques” maiores que 2 na mudança das variáveis uma vez que tais choques provavelmente seriam detectados quase imediatamente após a realização da mudança. Por apresentar resultados similares ao caso em que temos “choques” na qual apenas a variável X_2 e Z_2 mudam, não foi implementado o “choque” em que as variáveis X_1 e Z_1 mudam um pouco e as variáveis X_2 e Z_2 mudam ordenadamente.

6.1.2 – Análise dos resultados

Apresenta-se na Tabela 6.2 as estimativas dos limites de controle, *ARL in control* e *ARL out of control* para o processo não-normal simulado por Hayter e Tsui (1994). Nessa tabela são exibidos apenas os *ARL out of control* para as mudanças das médias apresentadas na Tabela 6.1.

Observa-se que tanto para processos simulados com $m=50$ quanto para $m=200$ os valores do *ARL in control* para a estatística M de Hayter e Tsui (1994) são intermediários quando comparados aos fornecidos pelas estatísticas T^2 de Hotelling (1947) e de componentes principais. Observa-se também que os valores do *ARL out of control* para essa estatística foram, na maioria das vezes, inferiores aos fornecidos pelas outras indicando que os limites de controle obtidos via estatística M de Hayter e Tsui (1994) são mais adequados pois se observa uma baixa quantidade de alarmes falsos e uma alta taxa de alarmes verdadeiros.

Pode-se observar que os valores do *ARL in control* para a estatística T^2 de Hotelling e componentes principais aumentaram e os valores da estatística M diminuíram quando a amostra aumentou de $m=50$ para $m=200$. Os valores do *ARL out of control* para a estatística T^2 de Hotelling e a primeira componente principal aumentaram quando e a estatística M de Hayter e Tsui diminuiu quando o tamanho das amostras simuladas aumentou de $m=50$ para

$m=200$. Em geral, a falta de controle do processo é mais rapidamente detectada quando os valores dos “choques” implementados para o cálculo dos *ARL out of control* mudam em uma maior escala, sendo que o “choque” em que as duas variáveis do processo mudam simultaneamente em um valor igual a 2 foram mais facilmente detectados.

Tabela 6.2: Limites de controle, *ARL in control* e *out of control*. Processo com distribuição não-normal bivariada estudada por Hayter e Tsui ($m=50$ e 200).

| Processo não-normal multivariado simulado com em Hayter e Tsui (1994) | | Medidas Descritivas | Estatística T ² de Hotelling | | Estatística M | | Componentes principais | | | | |
|---|--|---------------------|---|---------|---------------|---------|------------------------|---------|---------|---------|--|
| | | | Pb=5 | Bes | Pb=5 | Bes | Pb=5 | | Bes | | |
| m=50 | Limites de controle | Média | 6,2679 | 6,2766 | 2,1373 | 2,1488 | 0,0868 | 7,4732 | 0,0868 | 7,5210 | |
| | | Desvio | 1,4222 | 1,4125 | 0,2293 | 0,2183 | 0,0606 | 1,5067 | 0,0606 | 1,5019 | |
| | | Mediana | 6,1311 | 6,1353 | 2,1052 | 2,1149 | 0,0735 | 7,3304 | 0,0735 | 7,3635 | |
| | ARL in control | Média | 18,9450 | 19,0840 | 23,3060 | 23,5390 | 22,9480 | | 23,5440 | | |
| | | Desvio | 26,9068 | 27,0007 | 35,4742 | 35,4485 | 32,7380 | | 33,0954 | | |
| | | Mediana | 10,0 | 11,0 | 12,0 | 12,0 | 12,0 | | 13,0 | | |
| | Todas as médias mudam simultaneamente na mesma escala | Choque = 0,5 | Média | 11,4290 | 11,4640 | 12,5630 | 12,5360 | 16,2270 | | 16,4180 | |
| | | | Desvio | 13,5431 | 13,5597 | 21,6557 | 21,4545 | 27,5418 | | 27,6128 | |
| | | | Mediana | 7,0 | 7,0 | 7,0 | 7,0 | 9,0 | | 9,0 | |
| | | Choque = 1 | Média | 4,7030 | 4,7030 | 4,6720 | 4,6790 | 6,6120 | | 6,6970 | |
| | | | Desvio | 4,7133 | 4,7133 | 4,6459 | 4,6452 | 8,0623 | | 8,1617 | |
| | | | Mediana | 3,0 | 3,0 | 3,0 | 3,0 | 4,0 | | 4,0 | |
| | | Choque = 2 | Média | 1,3700 | 1,3700 | 1,3650 | 1,3680 | 1,5740 | | 1,5780 | |
| | | | Desvio | 0,7319 | 0,7319 | 0,6887 | 0,6892 | 0,9957 | | 1,0015 | |
| | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | | 1,0 | |
| | Uma média permanece constante e a outra muda ordenadamente | Choque = 0,5 | Média | 16,4770 | 16,5070 | 18,6740 | 18,6250 | 19,0960 | | 19,3930 | |
| | | | Desvio | 22,3444 | 22,3451 | 27,9616 | 27,5927 | 22,9068 | | 23,2376 | |
| | | | Mediana | 9,0 | 9,0 | 10,0 | 10,0 | 11,0 | | 11,0 | |
| | | Choque = 1 | Média | 8,2210 | 8,2250 | 8,8900 | 8,9620 | 11,2310 | | 11,3680 | |
| | | | Desvio | 9,4045 | 9,4048 | 10,8606 | 10,8913 | 14,8353 | | 14,9721 | |
| Mediana | | | 5,0 | 5,0 | 5,0 | 5,0 | 6,0 | | 7,0 | | |
| Choque = 2 | | Média | 2,5690 | 2,5700 | 2,4200 | 2,4170 | 3,1370 | | 3,1890 | | |
| | | Desvio | 2,2218 | 2,2213 | 1,9441 | 1,9112 | 2,9902 | | 3,0982 | | |
| | | Mediana | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | | 2,0 | | |
| m=200 | Limites de controle | Média | 5,8720 | 5,8720 | 2,0712 | 2,0726 | 0,0329 | 7,7771 | 0,0328 | 7,7799 | |
| | | Desvio | 0,8314 | 0,8311 | 0,1426 | 0,1413 | 0,0205 | 0,8423 | 0,0206 | 0,8403 | |
| | | Mediana | 5,7618 | 5,7618 | 2,0605 | 2,0616 | 0,0291 | 7,7279 | 0,0292 | 7,7261 | |
| | ARL in control | Média | 20,8500 | 20,8500 | 22,0300 | 22,0300 | 33,1800 | | 33,5000 | | |
| | | Desvio | 22,7000 | 22,7000 | 24,0700 | 24,0700 | 33,7700 | | 34,0200 | | |
| | | Mediana | 13,0 | 13,0 | 14,0 | 14,0 | 23,0 | | 23,0 | | |
| | Todas as médias mudam simultaneamente na mesma escala | Choque = 0,5 | Média | 12,1300 | 12,1300 | 10,9910 | 10,9910 | 21,5600 | | 21,6200 | |
| | | | Desvio | 11,4290 | 11,4290 | 10,5920 | 10,5920 | 27,2500 | | 27,2700 | |
| | | | Mediana | 8,0 | 8,0 | 8,0 | 8,0 | 13,0 | | 13,0 | |
| | | Choque = 1 | Média | 4,7950 | 4,7950 | 4,4190 | 4,4190 | 6,7020 | | 6,7580 | |
| | | | Desvio | 4,4050 | 4,4050 | 4,1120 | 4,1120 | 6,7090 | | 6,7370 | |
| | | | Mediana | 3,0 | 3,0 | 3,0 | 3,0 | 4,0 | | 4,0 | |
| | | Choque = 2 | Média | 1,3395 | 1,3395 | 1,3302 | 1,3302 | 1,5256 | | 1,5256 | |
| | | | Desvio | 0,6198 | 0,6198 | 0,6323 | 0,6323 | 0,8417 | | 0,8417 | |
| | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | | 1,0 | |
| | Uma média muda em 0,5 e a outra muda ordenadamente | Choque = 0,5 | Média | 16,1600 | 16,1600 | 15,7300 | 15,6000 | 30,7800 | | 30,8500 | |
| | | | Desvio | 15,4600 | 15,4600 | 15,1000 | 14,9700 | 36,7600 | | 36,7200 | |
| | | | Mediana | 11,0 | 11,0 | 11,0 | 11,0 | 19,0 | | 19,0 | |
| | | Choque = 1 | Média | 8,5490 | 8,5720 | 8,1210 | 8,0700 | 12,2470 | | 12,2470 | |
| | | | Desvio | 8,9430 | 8,9380 | 8,1280 | 8,1380 | 13,3150 | | 13,3150 | |
| Mediana | | | 6,0 | 6,0 | 6,0 | 6,0 | 8,0 | | 8,0 | | |
| Choque = 2 | | Média | 2,7490 | 2,7490 | 2,4470 | 2,4600 | 3,6470 | | 3,6470 | | |
| | | Desvio | 2,0470 | 2,0470 | 1,7580 | 1,7580 | 3,2830 | | 3,2830 | | |
| | | Mediana | 2,0 | 2,0 | 2,0 | 2,0 | 3,0 | | 3,0 | | |

Legenda: **Pb5**: ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes**: ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

Hayter e Tsui (1994) afirmam que para a determinação da constante $C_{R,\alpha}$ pelo método não-paramétrico por eles proposto para processos não-normais, é necessário ter-se no mínimo $N=500$ observações. Para o respectivo processo simulado nesta seção esses autores mostram em seu artigo que o valor da constante $C_{R,\alpha}$ é igual a 2,08 obtido via o método não-paramétrico (ver Quadro 2.2). Nos resultados das simulações apresentadas na Tabela 6.2 observamos que para amostras de tamanho $m=200$ a técnica de núcleo-estimador é eficaz na obtenção da constante $C_{R,\alpha}$ pois fornece valores próximos a 2,07, tanto pelo método de Polansky e Baker (2000) com $b=5$ quanto pelo método de Bessegato (2001). Já para $m=50$ a estimativa de $C_{R,\alpha}$ não é próxima de 2,08.

6.2 – Processo 2: O Processo Lognormal Bivariado

6.2.1 – Descrição do processo

Considere o vetor aleatório $X = [X_1, X_2]'$ com distribuição normal bivariada com parâmetros:

$$\mu = [0 \quad 0]' \text{ e } \Sigma = \begin{bmatrix} 0,25 & 0,20 \\ 0,20 & 0,25 \end{bmatrix}$$

Considere agora dois parâmetros definidos por:

$$\lambda_i = \frac{1}{\sqrt{\exp(2 \times \sigma_i^2) - \exp(\sigma_i^2)}}, \quad i = 1,2$$

$$\delta_i = \frac{-\exp(\sigma_i^2 / 2)}{\lambda_i^{-1}}, \quad i = 1,2$$

em que $\sigma_i^2 = \text{var}(X_i), i = 1,2$. Seja $Z = [Z_1, Z_2]'$ o vetor aleatório definido por:

$$Z_i = \lambda_i \times \exp(X_i) + \delta_i, \quad i = 1,2 \quad (6.1)$$

O vetor aleatório $Z = [Z_1, Z_2]'$ têm distribuição lognormal bivariada (Johnson, 1984) com esperança e variância definidas por:

$$E(Z_i) = \lambda_i \times \exp(\mu_i + \sigma_i^2 / 2) + \delta_i, \quad i = 1,2$$

$$\text{Var}(Z_i) = \lambda_i^2 \times (\exp(2\mu_i + 2\sigma_i^2) - \exp(2\mu_i + \sigma_i^2)) \quad i = 1,2$$

Dessa forma, como os valores do vetor aleatório Z são obtidos de acordo com os valores do vetor aleatório X (que tem distribuição normal bivariada), é necessário criar uma forma de implementar os valores dos “choques” no vetor aleatório Z para que os valores do *ARL out of control* possa ser corretamente calculado. Os valores desses “choques” devem

primeiramente ser implementados no vetor aleatório X , e dessa forma, a partir da transformação realizada em (6.1) esses valores sejam refletidos no vetor aleatório Z .

Para realizar esses cálculos, podemos determinar os valores teóricos dos parâmetros λ_i, δ_i que apresentam, de acordo com os valores teóricos dos parâmetros da distribuição normal bivariada, os seguintes resultados:

$$\lambda_{Teo} = \frac{1}{\sqrt{\exp(2 \times \sigma_{Teo}^2) - \exp(\sigma_{Teo}^2)}} \Rightarrow \frac{1}{\sqrt{\exp(2 \times 0,25) - \exp(0,25)}} = 2,74$$

$$\delta_{Teo} = \frac{-\exp(\sigma_{Teo}^2 / 2)}{\lambda_{Teo}^{-1}} \Rightarrow \frac{-\exp(0,25 / 2)}{(2,742011)^{-1}} = -3,11$$

E dessa forma, calculando a função inversa da $E(Z_i)$, temos os seguinte resultado:

$$\mu_{X;choque} = \log\left(\frac{choque - \delta_{Teo}}{\lambda_{Teo}}\right) - \frac{\sigma_{Teo}^2}{2} \Rightarrow \log\left(\frac{choque + 3,11}{2,74}\right) - \frac{0,25}{2}, \text{ choque} = 0,5; 1; 2; 3; 4$$

Que é a fórmula que será utilizada para determinar qual deverá ser a mudança empregada no vetor aleatório X para que o vetor aleatório Z tenha em média os valores 0,5; 1; 2; 3 e 4. Para exemplificar o raciocínio utilizado, considere que seja de interesse obter um vetor aleatório Z com vetor de média igual a:

$$\mu_Z = [0,5 \ 0,5]'$$

Assim, as médias do vetor de média do vetor aleatório X deverão ser iguais a:

$$\mu_{X;0,5} = \log\left(\frac{0,5 + 3,11}{2,74}\right) - \frac{0,25}{2} = 0,15$$

logo,

$$\mu_{X;0,5} = [0,15 \ 0,15]'$$

Para ilustrar o comportamento probabilístico do vetor aleatório Z foi gerada uma amostra de tamanho $m=100000$ com base na distribuição normal bivariada com parâmetros descritos em (6.2) e os valores do vetor aleatório Z . A função densidade conjunta de Z_1 e Z_2 e as curvas de nível amostrais são apresentados nas Figuras 6.3 e 6.4.

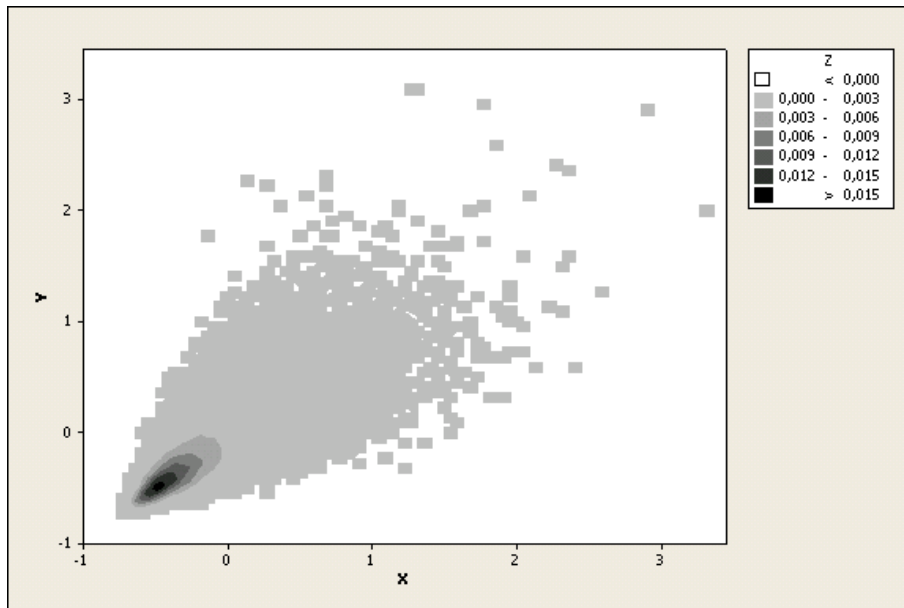


Figura 6.3: Curvas de nível amostral da função de densidade do vetor aleatório Z com distribuição lognormal bivariada.

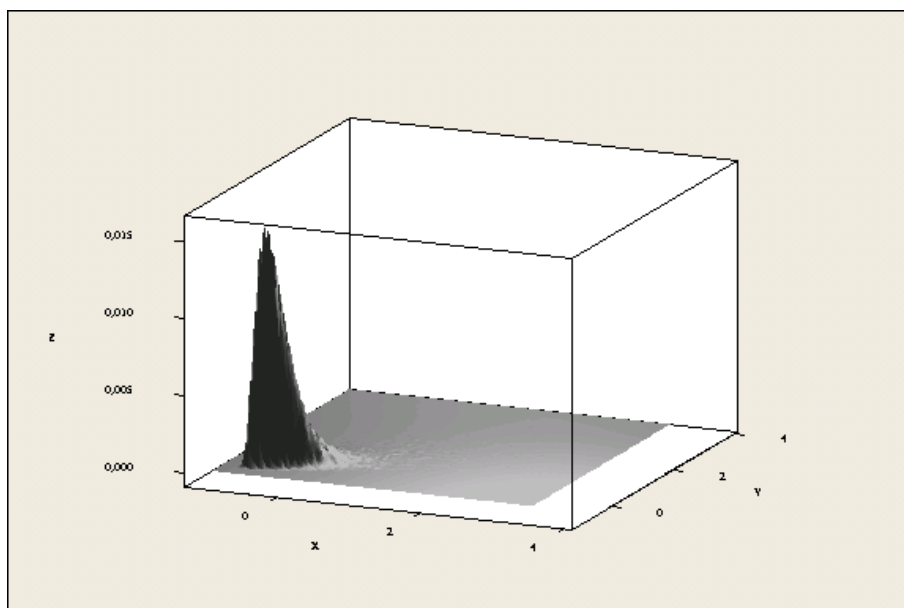


Figura 6.4: Função densidade amostral do vetor aleatório Z com distribuição lognormal bivariada.

6.2.2 – Análise dos resultados

Apresentamos nas Tabelas 6.3 e 6.4 as estimativas dos limites de controle, *ARL in control* e *out of control* para processos com a distribuição lognormal bivariada para amostras de tamanho $m=50$ e $m=200$. Os valores dos choques implementados para avaliar os *ARL out of control* foram definidos conforme descrito na seção 6.2.1.

Observa-se que os valores do *ARL in control* e *ARL out of control* da estatística *M* de Hayter e Tsui (1994) são intermediários em relação aos valores das estatísticas T^2 de Hotelling (1947) e da primeira componente principal. Em alguns casos é possível observar que o valor do *ARL out of control* da estatística *M* chega a ser menor que os valores dos *ARL out of control* fornecidos pelas outras estatísticas mostrando que essa é uma técnica bem eficaz na detecção da falta de controle estatístico do processo. As cartas de controle via núcleo estimador com a janela ótima *h* estimada pelos métodos de Polansky e Baker (2000) com $b=5$ e por Bessegato (2001) forneceram valores semelhantes, o que mostra que ambos poderiam ser utilizados na construção dos limites de controle para o processo lognormal bivariado.

Os valores do *ARL in control* calculados para a estatística T^2 de Hotelling, estatística *M* de Hayter e Tsui e pela primeira componente principal diminuíram quando se aumentou o tamanho da amostra de $m=50$ para 200 , fato que também ocorreu na maioria dos valores do *ARL out of control*. Observa-se que os “choques” que mais rapidamente detectados pelos limites de controle são aqueles em que as médias mudam simultaneamente para um valor igual a 4.

Tabela 6.3: Limites de controle, *ARL in control* e *out of control*. Processo com distribuição lognormal bivariada ($m=50$).

| Distribuição lognormal bivariada com $\rho=0.8$ e $m=50$ | | Medidas Descritivas | Estatística T ² de Hotelling | | Estatística M | | Componentes Principais | | | |
|--|--------------|---------------------|---|----------|---------------|----------|------------------------|---------|----------|--------|
| | | | Pb=5 | Bes | Pb=5 | Bes | Pb=5 | | Bes | |
| | | | | | | | | | | |
| Limites de controle | | Média | 7,7129 | 7,7182 | 2,3458 | 2,3504 | -1,7945 | 3,4383 | -1,8017 | 3,4641 |
| | | Desvio | 1,8326 | 1,8245 | 0,3140 | 0,3085 | 0,1570 | 0,7519 | 0,1627 | 0,7420 |
| | | Mediana | 7,5504 | 7,5425 | 2,3383 | 2,3418 | -1,7894 | 3,3514 | -1,7939 | 3,3783 |
| ARL in control | | Média | 23,31900 | 23,42600 | 26,42900 | 26,43800 | 27,89500 | | 28,60900 | |
| | | Desvio | 39,22680 | 39,28800 | 45,62990 | 45,56000 | 39,07910 | | 39,99260 | |
| | | Mediana | 12,0 | 12,0 | 14,0 | 14,0 | 15,0 | | 16,0 | |
| Todas as médias mudam simultaneamente na mesma escala | Choque = 0,5 | Média | 11,44700 | 11,42000 | 12,18500 | 12,20300 | 19,08500 | | 19,31600 | |
| | | Desvio | 19,90110 | 19,88750 | 23,66230 | 23,67080 | 35,03030 | | 35,12890 | |
| | | Mediana | 5,0 | 5,0 | 6,0 | 6,0 | 8,0 | | 8,5 | |
| | Choque = 1 | Média | 5,66600 | 5,68300 | 5,68800 | 5,72200 | 9,48100 | | 9,52200 | |
| | | Desvio | 8,35780 | 8,37730 | 8,67940 | 8,69500 | 13,25300 | | 13,22800 | |
| | | Mediana | 3,0 | 3,0 | 3,0 | 3,0 | 5,0 | | 5,0 | |
| | Choque = 2 | Média | 2,65500 | 2,65400 | 2,50700 | 2,53100 | 3,88300 | | 3,90400 | |
| | | Desvio | 2,55230 | 2,55280 | 2,53070 | 2,55060 | 6,03290 | | 6,04870 | |
| | | Mediana | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | | 2,0 | |
| | Choque = 3 | Média | 1,76200 | 1,76200 | 1,69000 | 1,69300 | 2,10300 | | 2,11300 | |
| | | Desvio | 1,25890 | 1,25890 | 1,18460 | 1,18670 | 1,69570 | | 1,70210 | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | | 1,0 | |
| Choque = 4 | Média | 1,42700 | 1,42900 | 1,36300 | 1,37100 | 1,60900 | | 1,61000 | | |
| | Desvio | 0,87830 | 0,87960 | 0,80860 | 0,81360 | 1,15130 | | 1,15120 | | |
| | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | | 1,0 | | |
| Uma média permanece constante e a outra muda ordenadamente | Choque = 0,5 | Média | 11,54100 | 11,54900 | 14,10700 | 14,22800 | 21,91600 | | 22,20600 | |
| | | Desvio | 17,39740 | 17,40260 | 23,40610 | 23,52360 | 34,67660 | | 34,81760 | |
| | | Mediana | 6,0 | 6,0 | 7,0 | 7,0 | 11,0 | | 11,0 | |
| | Choque = 1 | Média | 5,44500 | 5,44400 | 7,74900 | 7,70200 | 15,98000 | | 16,37000 | |
| | | Desvio | 7,22350 | 7,22340 | 10,09140 | 10,05200 | 23,33490 | | 24,15680 | |
| | | Mediana | 3,0 | 3,0 | 4,0 | 4,0 | 8,0 | | 9,0 | |
| | Choque = 2 | Média | 2,13500 | 2,13500 | 3,23500 | 3,23400 | 8,87300 | | 8,98500 | |
| | | Desvio | 1,81660 | 1,81660 | 3,31860 | 3,31760 | 12,89750 | | 13,09720 | |
| | | Mediana | 1,0 | 1,0 | 2,0 | 2,0 | 5,0 | | 5,0 | |
| | Choque = 3 | Média | 1,39700 | 1,39900 | 1,98000 | 1,98600 | 5,18100 | | 5,19000 | |
| | | Desvio | 0,84030 | 0,84050 | 1,47710 | 1,48050 | 6,94900 | | 6,91490 | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 3,0 | | 3,0 | |
| Choque = 4 | Média | 1,16400 | 1,16400 | 1,54300 | 1,55000 | 3,45200 | | 3,49300 | | |
| | Desvio | 0,51900 | 0,51900 | 0,99660 | 1,01120 | 3,79530 | | 3,85810 | | |
| | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | | 2,0 | | |
| Uma média muda em 0,5 e outra muda ordenadamente | Choque = 0,5 | Média | 11,14300 | 11,14800 | 11,71300 | 11,73000 | 18,59400 | | 18,89500 | |
| | | Desvio | 15,77360 | 15,77370 | 17,97150 | 17,96510 | 28,35660 | | 28,57990 | |
| | | Mediana | 6,0 | 6,0 | 6,0 | 6,0 | 9,0 | | 9,0 | |
| | Choque = 1 | Média | 6,52300 | 6,53100 | 7,11300 | 7,13300 | 13,97700 | | 14,12300 | |
| | | Desvio | 9,75360 | 9,75400 | 9,77150 | 9,77020 | 35,50840 | | 35,56270 | |
| | | Mediana | 4,0 | 4,0 | 4,0 | 4,0 | 6,0 | | 6,5 | |
| | Choque = 2 | Média | 2,52700 | 2,53000 | 3,33500 | 3,34100 | 7,26000 | | 7,32400 | |
| | | Desvio | 2,47500 | 2,47490 | 3,75720 | 3,77030 | 11,44490 | | 11,46630 | |
| | | Mediana | 2,0 | 2,0 | 2,0 | 2,0 | 4,0 | | 4,0 | |
| | Choque = 3 | Média | 1,52400 | 1,52500 | 1,99800 | 1,99800 | 3,99300 | | 4,03400 | |
| | | Desvio | 0,96140 | 0,96140 | 1,56670 | 1,56670 | 5,05260 | | 5,13260 | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | | 2,0 | |
| Choque = 4 | Média | 1,22900 | 1,22800 | 1,52000 | 1,52200 | 3,07700 | | 3,11300 | | |
| | Desvio | 0,53740 | 0,53690 | 0,94370 | 0,94470 | 3,52240 | | 3,53960 | | |
| | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | | 2,0 | | |

Legenda: **Pb5:** ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes:** ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

Tabela 6.4: Limites de controle, *ARL in control* e *out of control*. Processo com distribuição lognormal bivariada ($m=200$).

| Distribuição lognormal bivariada com $\rho=0.8$ e $m=200$ | | Medidas Descritivas | Estatística T ² de Hotelling | | Estatística M | | Componentes Principais | | | |
|--|--------------|---------------------|---|---------|---------------|---------|------------------------|--------|---------|--------|
| | | | Pb=5 | Bes | Pb=5 | Bes | Pb=5 | | Bes | |
| Limites de controle | | Média | 7,5178 | 7,5159 | 2,3095 | 2,3101 | -1,8179 | 3,4095 | -1,8243 | 3,4112 |
| | | Desvio | 1,0071 | 1,0021 | 0,1804 | 0,1800 | 0,0683 | 0,4453 | 0,0681 | 0,4441 |
| | | Mediana | 7,4768 | 7,4765 | 2,3103 | 2,3109 | -1,8220 | 3,3583 | -1,8314 | 3,3569 |
| ARL in control | | Média | 21,5430 | 21,5430 | 21,2020 | 21,2020 | 26,6580 | | 26,7880 | |
| | | Desvio | 24,1918 | 24,1918 | 21,5102 | 21,5102 | 29,0856 | | 29,1036 | |
| | | Mediana | 14,0 | 14,0 | 14,0 | 14,0 | 18,0 | | 18,0 | |
| Todas as médias mudam simultaneamente na mesma escala | Choque = 0,5 | Média | 9,1460 | 9,1410 | 8,7800 | 8,7600 | 15,7130 | | 15,7770 | |
| | | Desvio | 10,2771 | 10,2798 | 10,2417 | 10,2432 | 18,5753 | | 18,9160 | |
| | | Mediana | 6,0 | 6,0 | 5,0 | 5,0 | 10,0 | | 10,0 | |
| | Choque = 1 | Média | 5,1930 | 5,1970 | 5,0400 | 5,0440 | 8,3570 | | 8,3560 | |
| | | Desvio | 4,9665 | 4,9648 | 5,1599 | 5,1702 | 9,9421 | | 9,9519 | |
| | | Mediana | 4,0 | 4,0 | 4,0 | 4,0 | 5,0 | | 5,0 | |
| | Choque = 2 | Média | 2,4090 | 2,4060 | 2,2080 | 2,2090 | 3,1440 | | 3,1440 | |
| | | Desvio | 1,8877 | 1,8875 | 1,6792 | 1,6787 | 3,0546 | | 3,0546 | |
| | | Mediana | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | | 2,0 | |
| | Choque = 3 | Média | 1,7020 | 1,7020 | 1,6130 | 1,6110 | 1,9300 | | 1,9300 | |
| | | Desvio | 1,1200 | 1,1200 | 1,0528 | 1,0473 | 1,4364 | | 1,4364 | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | | 1,0 | |
| Choque = 4 | Média | 1,3780 | 1,3760 | 1,3430 | 1,3430 | 1,5440 | | 1,5420 | | |
| | Desvio | 0,7250 | 0,7247 | 0,6662 | 0,6662 | 0,9268 | | 0,9258 | | |
| | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | | 1,0 | | |
| Uma média permanece constante e a outra muda ordenadamente | Choque = 0,5 | Média | 9,9940 | 10,0120 | 12,0630 | 12,0960 | 20,4700 | | 20,5340 | |
| | | Desvio | 11,2590 | 11,2606 | 15,8590 | 15,8589 | 23,4428 | | 23,3281 | |
| | | Mediana | 6,0 | 6,0 | 8,0 | 8,0 | 13,0 | | 13,0 | |
| | Choque = 1 | Média | 4,7830 | 4,7780 | 6,2620 | 6,2630 | 14,6620 | | 14,7570 | |
| | | Desvio | 4,8288 | 4,8274 | 6,1644 | 6,1641 | 16,9438 | | 16,9902 | |
| | | Mediana | 3,0 | 3,0 | 4,0 | 4,0 | 9,0 | | 9,0 | |
| | Choque = 2 | Média | 2,0260 | 2,0250 | 2,9660 | 2,9680 | 7,1840 | | 7,1720 | |
| | | Desvio | 1,4613 | 1,4610 | 2,4114 | 2,4106 | 7,5949 | | 7,4855 | |
| | | Mediana | 2,0 | 2,0 | 2,0 | 2,0 | 5,0 | | 5,0 | |
| | Choque = 3 | Média | 1,3700 | 1,3700 | 1,8640 | 1,8640 | 4,7550 | | 4,7550 | |
| | | Desvio | 0,7468 | 0,7468 | 1,3451 | 1,3451 | 5,0162 | | 5,0162 | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 3,0 | | 3,0 | |
| Choque = 4 | Média | 1,1450 | 1,1440 | 1,5020 | 1,5020 | 3,1610 | | 3,1610 | | |
| | Desvio | 0,4197 | 0,4189 | 0,8756 | 0,8756 | 2,8750 | | 2,8750 | | |
| | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | | 2,0 | | |
| Uma média muda em 0,5 e outra muda ordenadamente | Choque = 0,5 | Média | 9,1370 | 9,1370 | 9,0210 | 9,0250 | 14,9560 | | 14,9570 | |
| | | Desvio | 10,5722 | 10,5563 | 10,7104 | 10,7100 | 19,1968 | | 19,7522 | |
| | | Mediana | 6,0 | 6,0 | 5,0 | 5,0 | 9,0 | | 9,0 | |
| | Choque = 1 | Média | 5,5740 | 5,5740 | 5,8990 | 5,8990 | 10,5590 | | 10,3190 | |
| | | Desvio | 5,1567 | 5,1567 | 5,8154 | 5,8154 | 14,2656 | | 13,2264 | |
| | | Mediana | 4,0 | 4,0 | 4,0 | 4,0 | 6,0 | | 6,0 | |
| | Choque = 2 | Média | 2,3790 | 2,3790 | 2,8200 | 2,8200 | 5,5120 | | 5,5220 | |
| | | Desvio | 1,9129 | 1,9129 | 2,4750 | 2,4754 | 5,5556 | | 5,5525 | |
| | | Mediana | 2,0 | 2,0 | 2,0 | 2,0 | 4,0 | | 4,0 | |
| | Choque = 3 | Média | 1,5200 | 1,5200 | 1,8360 | 1,8360 | 3,7230 | | 3,7230 | |
| | | Desvio | 0,8958 | 0,8958 | 1,2151 | 1,2151 | 3,6949 | | 3,6949 | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | | 2,0 | |
| Choque = 4 | Média | 1,2210 | 1,2210 | 1,4910 | 1,4910 | 2,8380 | | 2,8380 | | |
| | Desvio | 0,5920 | 0,5920 | 0,8959 | 0,8959 | 2,5140 | | 2,5140 | | |
| | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | | 2,0 | | |

Legenda: **Pb5:** ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes:** ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

6.3 – Processo 3: Seno Hiperbólico Multivariado (p=5)

6.3.1 – Descrição do processo

Considere o vetor aleatório $X = [X_1, X_2, X_3, X_4, X_5]'$ com distribuição normal multivariada com $p=5$. Essa distribuição possui os mesmos parâmetros utilizados para simular os processos normais multivariados com $p=5$ não-autocorrelacionados descritos no capítulo 4 (ver páginas 57 e 58). Considere agora o vetor aleatório $Z = [Z_1, Z_2, Z_3, Z_4, Z_5]'$ sendo:

$$Z_i = \sinh(X_i) \quad , i = 1, 2, 3, 4, 5$$

Dessa forma, a variável aleatória Z_i tem esperança e variância definidas por:

$$E(Z_i) = \exp\left(\frac{\text{var}(X_i)}{2}\right) \sinh(\mu_i) \quad , i = 1, 2, 3, 4, 5$$

$$\text{Var}(Z_i) = [\exp(\text{var}(X_i)) - 1] \times \{1 + [\cosh(2 \times \mu_i) \times \exp(\text{var}(X_i))] / 2\} \quad , i = 1, 2, 3, 4, 5$$

em que $\sinh(\bullet)$ e $\cosh(\bullet)$ representam o seno e o cosseno hiperbólicos respectivamente (Johnson, 1984). Como serão geradas amostras do vetor aleatório X é necessário descobrir qual a intensidade dos “choques” que deverão ser aplicados no vetor de médias de X de forma a observar um choque de 0,5; 1; 2; 3 e 4 em uma ou mais médias do vetor aleatório Z para o cálculo dos *ARL out of control*. As intensidades desses choques são determinadas de acordo com os valores teóricos da distribuição do vetor aleatório X utilizando a função inversa da esperança da distribuição do seno hiperbólico. As intensidades dos choques são obtidas da seguinte maneira:

$$\mu_{X;choque} = \text{asinh}\left(\frac{\text{choque}_{Teo}}{\exp(\text{var}(X)/2)}\right) \Rightarrow \text{asinh}\left(\frac{\text{choque}_{Teo}}{\exp(1/2)}\right) \quad , \text{choque} = 0,5; 1; 2; 3; 4$$

em que $\text{asinh}(\bullet)$ é o arco seno hiperbólico. Se obtivermos amostras do vetor aleatório X segundo a distribuição normal bivariada com vetor de médias

$$\mu_{X;0,5} = [0,2987993 \ 0,2987993 \ 0,2987993 \ 0,2987993 \ 0,2987993]'$$

e matriz de variâncias e covariâncias como as definidas no capítulo 4 (em que as variâncias das $p=5$ varáveis são iguais a 1), teremos amostras do vetor aleatório Z com distribuição do seno hiperbólico da distribuição normal multivariada com vetor de médias igual a

$$\mu_Z = [0,5 \ 0,5 \ 0,5 \ 0,5 \ 0,5]'$$

Apesar do vetor aleatório Z possuir $p=5$ variáveis, para ilustrar o comportamento probabilístico de Z estimamos a função de densidade conjunta de Z_1 e Z_2 considerando uma amostra de tamanho $m=100000$ de uma distribuição normal bivariada com vetor de média

nulo, variâncias igual a 1 e correlação igual a 0,9. Os gráficos das curvas de nível da função de densidade conjunta amostral de Z_1 e Z_2 são apresentados nas Figuras 6.5 e 6.6.

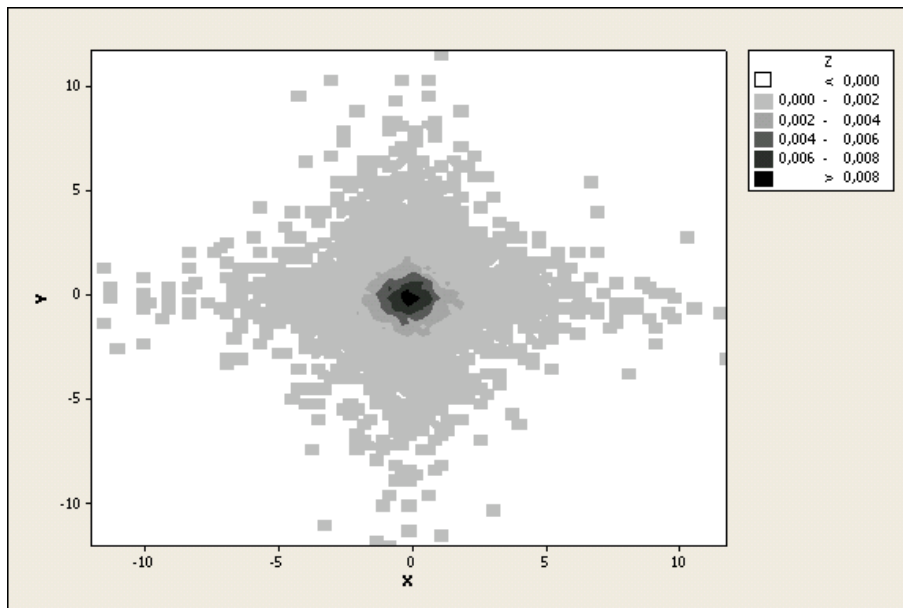


Figura 6.5: Curvas de nível amostral da função densidade do vetor aleatório Z com distribuição do seno hiperbólico.

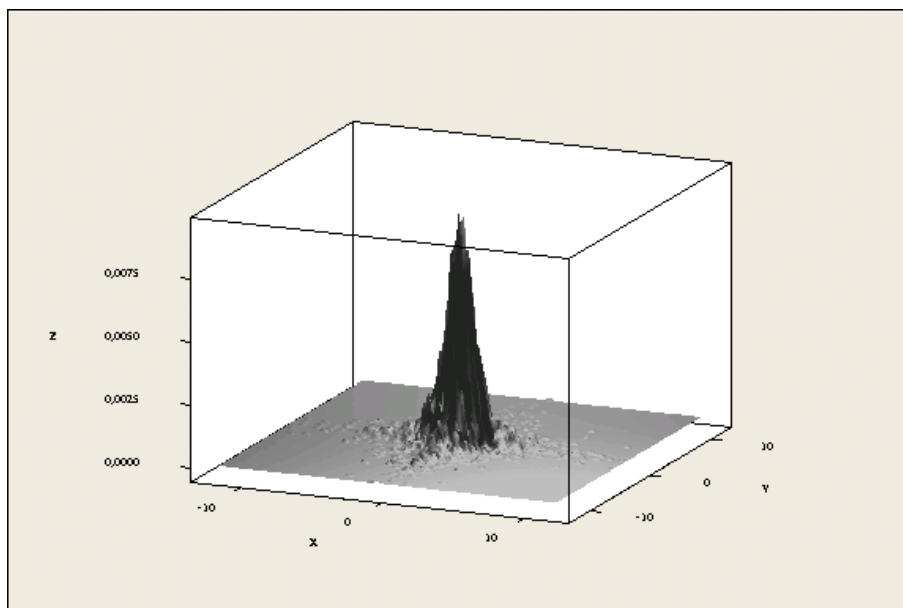


Figura 6.6: Função densidade amostral do vetor aleatório Z com distribuição do seno hiperbólico.

6.3.2 – Análise dos resultados

Apresentamos nas Tabelas 6.5 até 6.9 as estimativas dos limites de controle, *ARL in control* e *out of control* para o processo com a distribuição do seno hiperbólico de uma distribuição normal multivariada não autocorrelacionada com $p=5$ variáveis para os tamanhos

de amostra igual a $m=50$ e $m=200$. Nas Tabelas 6.5 e 6.6 apresentamos os resultados para o processo em que as variáveis são fortemente correlacionadas ($0,75 < \rho \leq 1$), nas Tabelas 6.7 e 6.8 apresentamos os resultados para correlação intermediária ($0,50 < \rho \leq 0,75$) e na Tabela 6.9 os resultados para processos com correlação nula ($\rho = 0$), sendo que nesse caso os valores da primeira componente principal não foram apresentados.

Observa-se que a média dos valores do *ARL in control* e *out of control* gerados via estatística M de Hayter e Tsui (1994) forneceram valores intermediários quando comparados com os valores da estatística T^2 de Hotelling (1947) e da primeira componente principal. Os valores dos *ARL in control* e *ARL out of control* da primeira componente principal forneceram na maioria das vezes os valores altos das três técnicas enquanto os valores da estatística T^2 de Hotelling (1947) sempre forneceram menores valores. Em alguns casos, principalmente nas amostras não correlacionadas, é possível observar que os valores das medianas dos *ARL out of control* das estatísticas T^2 de Hotelling e M .

Em geral, não foram observadas mudanças significativas nos valores dos *ARL in control* e *out of control* quando mudamos a faixa de correlação das variáveis. Os valores do *ARL in control* são praticamente os mesmos quando consideramos mudanças do tamanho das amostras simuladas de $m=50$ para $m=200$ com exceção do caso em que a correlação entre as variáveis é nula (nesse caso, os valores do *ARL in control* para amostras de tamanho $m=200$ são maiores que aqueles em que $m=50$). Os valores do *ARL in control* para a estatística T^2 de Hotelling (1947) aumentam e os das componentes principais diminuem à medida que o tamanho de amostra aumenta. Em alguns casos, o comportamento do *ARL out of control* é bastante parecido com os descritos para o *ARL in control* com exceção de alguns casos em que a estatística M apresenta maiores valores do *ARL out of control* à medida que o tamanho de amostra aumenta de $m=50$ para $m=200$.

Em geral, detecta-se com mais facilidade aqueles “choques” que possuem maior amplitude sendo que aqueles “choques” em que todas as variáveis do processo mudam simultaneamente possuem sempre menor valor do *ARL out of control*.

Tabela 6.5: Limites de controle, *ARL in control* e *out of control*. Processo com distribuição do seno hiperbólico de uma distribuição normal multivariada com $p=5$ variáveis, correlação forte ($0,75 < \rho \leq 1$) e $m=50$.

| Distribuição do seno hiperbólico $p=5$ $0,75 < \rho \leq 1$ e $m=50$ | | Medidas Descritivas | Estatística T ² de Hotelling | | Estatística M | | Componentes Principais | | | |
|---|--------------|---------------------|---|---------|---------------|---------|------------------------|---------|---------|--------|
| | | | Pb=5 | Bes | Pb=5 | Bes | Pb=5 | | Bes | |
| | | | Média | Desvio | Mediana | Média | Desvio | Mediana | Média | Desvio |
| Limites de Controle | | Média | 18,4058 | 18,4189 | 2,9960 | 3,0011 | -7,4857 | 7,4718 | -7,5646 | 7,5520 |
| | | Desvio | 4,1607 | 4,1403 | 0,3866 | 0,3808 | 2,5293 | 2,5318 | 2,5267 | 2,5256 |
| | | Mediana | 17,5914 | 17,5631 | 2,9645 | 2,9632 | -6,9680 | 6,9931 | -7,0706 | 7,0489 |
| ARL in control | | Média | 13,4890 | 13,4930 | 20,0560 | 20,1510 | 20,5140 | | 21,0520 | |
| | | Desvio | 16,6318 | 16,6293 | 32,0360 | 32,0401 | 27,5460 | | 27,8878 | |
| | | Mediana | 8,0 | 8,0 | 11,0 | 11,0 | 12,0 | | 12,0 | |
| Todas as médias mudam simultaneamente na mesma escala | Choque = 0,5 | Média | 11,1040 | 11,1500 | 16,1300 | 16,0620 | 17,0130 | | 17,6240 | |
| | | Desvio | 12,1246 | 12,1345 | 22,3991 | 21,9487 | 19,6657 | | 20,9070 | |
| | | Mediana | 7,0 | 7,0 | 9,0 | 9,0 | 10,0 | | 11,0 | |
| | Choque = 1 | Média | 8,4010 | 8,4050 | 10,6050 | 10,6280 | 12,0130 | | 12,2880 | |
| | | Desvio | 9,8745 | 9,8780 | 13,3523 | 13,3707 | 14,9705 | | 15,1052 | |
| | | Mediana | 5,0 | 5,0 | 6,0 | 6,0 | 7,0 | | 7,0 | |
| | Choque = 2 | Média | 4,5010 | 4,5060 | 5,3010 | 5,2870 | 6,1200 | | 6,1640 | |
| | | Desvio | 4,4773 | 4,4795 | 5,4635 | 5,4430 | 7,6192 | | 7,5682 | |
| | | Mediana | 3,0 | 3,0 | 4,0 | 4,0 | 4,0 | | 4,0 | |
| | Choque = 3 | Média | 3,1110 | 3,1130 | 3,4670 | 3,4670 | 3,9460 | | 3,9940 | |
| | | Desvio | 2,7648 | 2,7651 | 3,2366 | 3,2356 | 4,1944 | | 4,2280 | |
| | | Mediana | 2,0 | 2,0 | 2,0 | 2,0 | 3,0 | | 3,0 | |
| Choque = 4 | Média | 2,3360 | 2,3360 | 2,5110 | 2,5080 | 2,6780 | | 2,6970 | | |
| | Desvio | 1,9336 | 1,9336 | 2,0939 | 2,0939 | 2,4937 | | 2,4963 | | |
| | Mediana | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | | 2,0 | | |
| Uma média permanece constante e a outra muda ordenadamente | Choque = 0,5 | Média | 11,9380 | 11,9440 | 18,0530 | 18,1170 | 20,8460 | | 21,2200 | |
| | | Desvio | 15,2990 | 15,2989 | 28,6495 | 28,6881 | 30,2548 | | 30,4799 | |
| | | Mediana | 7,0 | 7,0 | 10,0 | 10,0 | 12,0 | | 12,0 | |
| | Choque = 1 | Média | 8,7460 | 8,7610 | 14,1020 | 14,0880 | 18,2110 | | 19,0400 | |
| | | Desvio | 10,4098 | 10,4044 | 20,1767 | 20,1734 | 25,2032 | | 26,8792 | |
| | | Mediana | 5,0 | 5,0 | 8,0 | 8,0 | 11,0 | | 11,0 | |
| | Choque = 2 | Média | 4,9280 | 4,9280 | 8,8170 | 8,6900 | 16,3760 | | 16,6650 | |
| | | Desvio | 5,4833 | 5,4837 | 14,2637 | 13,2623 | 21,3693 | | 21,4809 | |
| | | Mediana | 3,0 | 3,0 | 5,0 | 5,0 | 9,0 | | 9,0 | |
| | Choque = 3 | Média | 3,0020 | 2,9960 | 5,2080 | 5,2500 | 12,9100 | | 13,3700 | |
| | | Desvio | 3,3724 | 3,3492 | 5,6475 | 5,6887 | 15,4953 | | 16,3453 | |
| | | Mediana | 2,0 | 2,0 | 3,0 | 3,0 | 8,0 | | 8,0 | |
| Choque = 4 | Média | 2,1990 | 2,1990 | 4,1230 | 4,1280 | 10,7600 | | 10,9950 | | |
| | Desvio | 3,3572 | 3,3572 | 4,8028 | 4,7978 | 12,9758 | | 13,1573 | | |
| | Mediana | 1,0 | 1,0 | 3,0 | 3,0 | 6,0 | | 7,0 | | |

Legenda: **Pb5:** ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes:** ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

Tabela 6.6: Limites de controle, *ARL in control* e *out of control*. Processo com distribuição do seno hiperbólico de uma distribuição normal multivariada com $p=5$ variáveis, correlação forte ($0,75 < \rho \leq 1$) e $m=200$.

| Distribuição do seno hiperbólico $p=5$ $0,75 < \rho \leq 1$ e $m=200$ | | Medidas Descritivas | Estatística T ² de Hotelling | | Estatística M | | Componentes Principais | | | | |
|--|---------|---------------------|---|---------|---------------|---------|------------------------|---------|---------|---------|--|
| | | | Pb=5 | Bes | Pb=5 | Bes | Pb=5 | | Bes | | |
| | | | | | | | | | | | |
| Limites de controle | | Média | 18,6949 | 18,6897 | 2,9372 | 2,9375 | -7,3702 | 7,3936 | -7,3759 | 7,3989 | |
| | | Desvio | 2,8482 | 2,8407 | 0,2733 | 0,2729 | 1,2770 | 1,3794 | 1,2734 | 1,3767 | |
| | | Mediana | 18,5140 | 18,5135 | 2,9670 | 2,9670 | -7,2579 | 7,2653 | -7,2627 | 7,2730 | |
| ARL in control | | Média | 18,3060 | 18,2980 | 19,7000 | 19,6790 | 20,0220 | | 20,1030 | | |
| | | Desvio | 17,9224 | 17,9194 | 19,3502 | 19,3179 | 21,0074 | | 21,0034 | | |
| | | Mediana | 13,0 | 13,0 | 14,0 | 14,0 | 13,5 | | 14,0 | | |
| Todas as médias mudam simultaneamente na mesma escala | | Choque = 0,5 | Média | 15,9790 | 15,9790 | 16,0990 | 16,0990 | 17,4090 | | 17,2390 | |
| | | | Desvio | 17,5009 | 17,5009 | 16,2792 | 16,2792 | 19,1121 | | 18,6610 | |
| | | | Mediana | 11,0 | 11,0 | 11,0 | 11,0 | 11,0 | | 11,0 | |
| | | Choque = 1 | Média | 11,0380 | 11,0350 | 11,2800 | 11,2800 | 12,0940 | | 12,0940 | |
| | | | Desvio | 11,3455 | 11,3448 | 11,1331 | 11,1331 | 12,0185 | | 12,0185 | |
| | | | Mediana | 7,0 | 7,0 | 8,0 | 8,0 | 8,0 | | 8,0 | |
| | | Choque = 2 | Média | 5,5440 | 5,5440 | 5,2850 | 5,2850 | 5,6890 | | 5,6920 | |
| | | | Desvio | 5,1453 | 5,1453 | 4,7970 | 4,7970 | 5,5677 | | 5,5660 | |
| | | | Mediana | 4,0 | 4,0 | 4,0 | 4,0 | 4,0 | | 4,0 | |
| | | Choque = 3 | Média | 3,3810 | 3,3810 | 3,3450 | 3,3450 | 3,3080 | | 3,3080 | |
| | | | Desvio | 2,7939 | 2,7939 | 2,8952 | 2,8952 | 2,8833 | | 2,8833 | |
| | | | Mediana | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | | 2,0 | |
| Choque = 4 | Média | 2,4470 | 2,4460 | 2,3010 | 2,3010 | 2,3460 | | 2,3460 | | | |
| | Desvio | 1,9948 | 1,9938 | 1,7474 | 1,7474 | 1,8264 | | 1,8264 | | | |
| | Mediana | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | | 2,0 | | | |
| Uma média permanece constante e a outra muda ordenadamente | | Choque = 0,5 | Média | 17,7610 | 17,7630 | 19,1830 | 19,2280 | 21,5670 | | 21,5960 | |
| | | | Desvio | 18,4660 | 18,4662 | 19,1214 | 19,2582 | 21,7986 | | 21,7931 | |
| | | | Mediana | 12,0 | 12,0 | 14,0 | 14,0 | 15,0 | | 15,0 | |
| | | Choque = 1 | Média | 11,1630 | 11,1670 | 13,6970 | 13,6970 | 18,8130 | | 18,8640 | |
| | | | Desvio | 11,5314 | 11,5296 | 14,1941 | 14,1941 | 19,9703 | | 19,9674 | |
| | | | Mediana | 8,0 | 8,0 | 9,0 | 9,0 | 12,0 | | 12,0 | |
| | | Choque = 2 | Média | 6,1690 | 6,1690 | 8,6260 | 8,6290 | 16,1070 | | 16,1150 | |
| | | | Desvio | 6,1973 | 6,1973 | 8,6676 | 8,6669 | 16,9267 | | 16,9357 | |
| | | | Mediana | 4,0 | 4,0 | 6,0 | 6,0 | 11,0 | | 12,0 | |
| | | Choque = 3 | Média | 3,5840 | 3,5790 | 5,2850 | 5,3030 | 12,2950 | | 12,2970 | |
| | | | Desvio | 3,5780 | 3,5753 | 5,5193 | 5,5199 | 12,3941 | | 12,3922 | |
| | | | Mediana | 2,0 | 2,0 | 4,0 | 4,0 | 9,0 | | 9,0 | |
| Choque = 4 | Média | 2,5850 | 2,5850 | 3,8270 | 3,8270 | 9,4990 | | 9,5230 | | | |
| | Desvio | 2,0917 | 2,0917 | 3,4048 | 3,4048 | 9,4802 | | 9,4932 | | | |
| | Mediana | 2,0 | 2,0 | 3,0 | 3,0 | 6,0 | | 6,0 | | | |

Legenda: **Pb5:** ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes:** ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

Tabela 6.7: Limites de controle, *ARL in control* e *out of control*. Processo com distribuição do seno hiperbólico de uma distribuição normal multivariada com $p=5$ variáveis correlação intermediária ($0,50 < \rho \leq 0,75$) e $m=50$.

| Distribuição do seno hiperbólico $p=5$ $0,50 < \rho \leq 0,75$ e $m=50$ | | Medidas Descritivas | Estatística T ² de Hotelling | | Estatística M | | Componentes Principais | | | | |
|--|---------|---------------------|---|---------|---------------|---------|------------------------|---------|---------|---------|--|
| | | | Pb=5 | Bes | Pb=5 | Bes | Pb=5 | | Bes | | |
| | | | | | | | | | | | |
| Limites de controle | | Média | 17,4465 | 17,4621 | 3,3178 | 3,3250 | -6,8676 | 6,8164 | -6,9421 | 6,8858 | |
| | | Desvio | 3,7534 | 3,7287 | 0,4043 | 0,3971 | 2,1946 | 2,1630 | 2,1874 | 2,1532 | |
| | | Mediana | 16,7609 | 16,7562 | 3,2876 | 3,2909 | -6,4789 | 6,4212 | -6,5525 | 6,4905 | |
| ARL in control | | Média | 13,8510 | 13,8870 | 18,9150 | 18,9780 | 23,0640 | | 23,5170 | | |
| | | Desvio | 16,6415 | 16,6614 | 24,6922 | 24,8045 | 32,2287 | | 32,2224 | | |
| | | Mediana | 8,5 | 8,5 | 11,0 | 10,0 | 13,0 | | 13,0 | | |
| Todas as médias mudam simultaneamente na mesma escala | | Choque = 0,5 | Média | 12,1590 | 12,1790 | 16,1190 | 16,2590 | 19,4430 | | 20,1850 | |
| | | | Desvio | 16,5667 | 16,5674 | 22,5274 | 22,8308 | 25,0150 | | 26,0383 | |
| | | | Mediana | 7,0 | 7,0 | 10,0 | 10,0 | 11,0 | | 11,0 | |
| | | Choque = 1 | Média | 7,9310 | 7,9520 | 10,0070 | 10,0500 | 12,1230 | | 12,4130 | |
| | | | Desvio | 8,8609 | 8,8681 | 13,0204 | 13,0536 | 16,9627 | | 17,0942 | |
| | | | Mediana | 5,0 | 5,0 | 6,0 | 6,0 | 7,0 | | 7,0 | |
| | | Choque = 2 | Média | 4,1020 | 4,1020 | 4,6450 | 4,6500 | 5,9190 | | 6,0470 | |
| | | | Desvio | 4,3564 | 4,3592 | 4,8092 | 4,8123 | 8,7078 | | 8,8794 | |
| | | | Mediana | 3,0 | 3,0 | 3,0 | 3,0 | 4,0 | | 4,0 | |
| | | Choque = 3 | Média | 2,8220 | 2,8180 | 3,0870 | 3,0870 | 3,4780 | | 3,5310 | |
| | | | Desvio | 2,5410 | 2,5407 | 3,0602 | 3,0573 | 4,0465 | | 4,1243 | |
| | | | Mediana | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | | 2,0 | |
| Choque = 4 | Média | 2,0510 | 2,0480 | 2,1840 | 2,1920 | 2,3850 | | 2,3880 | | | |
| | Desvio | 1,5495 | 1,5421 | 1,7491 | 1,7556 | 2,2013 | | 2,1996 | | | |
| | Mediana | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 | | 2,0 | | | |
| Uma média permanece constante e a outra muda ordenadamente | | Choque = 0,5 | Média | 12,8910 | 12,9250 | 17,0740 | 17,1680 | 22,9940 | | 23,8240 | |
| | | | Desvio | 15,7236 | 15,7359 | 22,1856 | 22,2030 | 32,6166 | | 32,9887 | |
| | | | Mediana | 8,0 | 8,0 | 10,0 | 10,0 | 13,0 | | 14,0 | |
| | | Choque = 1 | Média | 9,8910 | 9,8980 | 15,6740 | 15,7590 | 20,3960 | | 20,7990 | |
| | | | Desvio | 12,0507 | 12,0698 | 19,4352 | 19,4901 | 25,3838 | | 25,6307 | |
| | | | Mediana | 6,0 | 6,0 | 9,0 | 9,0 | 12,0 | | 12,0 | |
| | | Choque = 2 | Média | 6,2070 | 6,2170 | 10,0160 | 10,0310 | 16,6200 | | 17,1640 | |
| | | | Desvio | 9,1654 | 9,1674 | 13,0266 | 13,0175 | 22,5584 | | 23,4406 | |
| | | | Mediana | 4,0 | 4,0 | 6,0 | 6,0 | 9,5 | | 10,0 | |
| | | Choque = 3 | Média | 3,7240 | 3,7390 | 6,4090 | 6,4450 | 13,6920 | | 13,9040 | |
| | | | Desvio | 4,3154 | 4,3603 | 8,1260 | 8,1402 | 20,4970 | | 20,5844 | |
| | | | Mediana | 2,0 | 2,0 | 4,0 | 4,0 | 8,0 | | 8,0 | |
| Choque = 4 | Média | 2,6380 | 2,6480 | 4,4230 | 4,4350 | 9,5030 | | 9,7230 | | | |
| | Desvio | 2,7482 | 2,7669 | 4,9104 | 4,9210 | 10,5487 | | 10,6224 | | | |
| | Mediana | 2,0 | 2,0 | 3,0 | 3,0 | 6,0 | | 6,0 | | | |

Legenda: **Pb5:** ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes:** ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

Tabela 6.8: Limites de controle, *ARL in control* e *out of control*. Processo com distribuição do seno hiperbólico de uma distribuição normal multivariada com $p=5$ variáveis correlação intermediária ($0,50 < \rho \leq 0,75$) e $m=200$.

| Distribuição do seno hiperbólico $p=5$ $0,50 < \rho \leq 0,75$ e $m=200$ | | Medidas Descritivas | Estatística T ² de Hotelling | | Estatística M | | Componentes Principais | | | | |
|---|--|---------------------|---|---------|---------------|---------|------------------------|---------|---------|---------|--|
| | | | Pb=5 | Bes | Pb=5 | Bes | Pb=5 | | Bes | | |
| Limites de controle | | Média | 17,9900 | 17,9900 | 3,3295 | 3,3301 | -6,5630 | 6,6616 | -6,5693 | 6,6663 | |
| | | Desvio | 2,4710 | 2,4691 | 0,2495 | 0,2491 | 1,0252 | 1,1185 | 1,0224 | 1,1158 | |
| | | Mediana | 17,7225 | 17,7208 | 3,3209 | 3,3216 | -6,4665 | 6,5236 | -6,4771 | 6,5252 | |
| ARL in control | | Média | 17,8220 | 17,8220 | 18,5900 | 18,6080 | 20,1820 | | 20,2400 | | |
| | | Desvio | 17,8752 | 17,8752 | 18,2998 | 18,3140 | 19,6058 | | 19,6278 | | |
| | | Mediana | 12,0 | 12,0 | 13,0 | 13,0 | 15,0 | | 15,0 | | |
| Todas as médias mudam simultaneamente na mesma escala | Choque = 0,5 | Média | 15,9860 | 15,9860 | 15,8290 | 15,8290 | 16,2700 | | 16,2940 | | |
| | | Desvio | 17,1526 | 17,1526 | 16,3631 | 16,3631 | 18,2305 | | 18,1305 | | |
| | | Mediana | 10,0 | 10,0 | 10,0 | 10,0 | 10,0 | | 10,0 | | |
| | Choque = 1 | Média | 9,9810 | 9,9800 | 10,2070 | 10,2770 | 10,5750 | | 10,5830 | | |
| | | Desvio | 9,3985 | 9,3990 | 9,6530 | 9,8489 | 10,5935 | | 10,5961 | | |
| | | Mediana | 7,0 | 7,0 | 7,0 | 7,0 | 7,0 | | 7,0 | | |
| | Choque = 2 | Média | 4,9280 | 4,9370 | 4,9210 | 4,9210 | 4,4300 | | 4,4320 | | |
| | | Desvio | 4,5600 | 4,5789 | 4,4107 | 4,4107 | 3,6574 | | 3,6566 | | |
| | | Mediana | 3,0 | 3,0 | 4,0 | 4,0 | 3,0 | | 3,0 | | |
| | Choque = 3 | Média | 3,3640 | 3,3640 | 3,2460 | 3,2480 | 3,0050 | | 3,0070 | | |
| | | Desvio | 2,9894 | 2,9894 | 2,8548 | 2,8542 | 2,6889 | | 2,6878 | | |
| | | Mediana | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | | 2,0 | | |
| | Choque = 4 | Média | 2,2500 | 2,2480 | 2,2080 | 2,2080 | 2,0360 | | 2,0360 | | |
| | | Desvio | 1,7189 | 1,7168 | 1,6412 | 1,6412 | 1,5030 | | 1,5030 | | |
| | | Mediana | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | | 2,0 | | |
| | Uma média permanece constante e a outra muda ordenadamente | Choque = 0,5 | Média | 17,2740 | 17,2830 | 18,4130 | 18,3900 | 19,3630 | | 19,3960 | |
| | | | Desvio | 18,0809 | 18,0775 | 18,8022 | 18,7317 | 18,8112 | | 18,8074 | |
| | | | Mediana | 11,0 | 11,0 | 12,0 | 12,0 | 14,0 | | 14,0 | |
| | | Choque = 1 | Média | 12,6720 | 12,6720 | 14,6620 | 14,6760 | 18,4930 | | 18,4930 | |
| | | | Desvio | 13,9705 | 13,9705 | 16,2428 | 16,2392 | 20,1666 | | 20,1666 | |
| | | | Mediana | 8,0 | 8,0 | 9,0 | 9,0 | 12,0 | | 12,0 | |
| | | Choque = 2 | Média | 7,3150 | 7,3050 | 10,1140 | 10,1060 | 13,9870 | | 13,9960 | |
| | | | Desvio | 7,5084 | 7,5095 | 10,9701 | 10,9636 | 14,3889 | | 14,3925 | |
| | | | Mediana | 5,0 | 5,0 | 6,0 | 6,0 | 10,0 | | 10,0 | |
| Choque = 3 | | Média | 4,4410 | 4,4410 | 6,3790 | 6,3790 | 10,5940 | | 10,5940 | | |
| | | Desvio | 4,2992 | 4,2992 | 6,5663 | 6,5663 | 10,1973 | | 10,1973 | | |
| | | Mediana | 3,0 | 3,0 | 4,0 | 4,0 | 8,0 | | 8,0 | | |
| Choque = 4 | | Média | 3,2100 | 3,2100 | 4,6790 | 4,6790 | 9,2960 | | 9,2830 | | |
| | | Desvio | 2,8527 | 2,8527 | 4,5269 | 4,5269 | 9,4315 | | 9,4229 | | |
| | | Mediana | 2,0 | 2,0 | 3,0 | 3,0 | 6,0 | | 6,0 | | |

Legenda: **Pb5:** ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes:** ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

Tabela 6.9: Limites de controle, *ARL in control* e *out of control*. Processo com distribuição do seno hiperbólico de uma distribuição normal multivariada com $p=5$ variáveis correlação nula ($\rho = 0$) $m=50$ e $m=200$.

| Distribuição do seno hiperbólico $p=5$ $\rho = 0$ $m=50$ e $m=200$ | | Medidas Descritivas | m=50 | | | | m=200 | | | |
|---|--------------|---------------------|---|---------|---------------|---------|---|---------|---------------|---------|
| | | | Estatística T ² de Hotelling | | Estatística M | | Estatística T ² de Hotelling | | Estatística M | |
| | | | Pb=5 | Bes | Pb=5 | Bes | Pb=5 | Bes | Pb=5 | Bes |
| Limites de controle | | Média | 15,9348 | 15,9602 | 3,6203 | 3,6301 | 16,8805 | 16,8803 | 3,7041 | 3,7048 |
| | | Desvio | 2,9552 | 2,9271 | 0,4028 | 0,3931 | 1,9176 | 1,9162 | 0,2382 | 0,2374 |
| | | Mediana | 15,5188 | 15,5468 | 3,5819 | 3,5842 | 16,7999 | 16,7999 | 3,7034 | 3,7052 |
| <i>ARL in control</i> | | Média | 13,3410 | 13,3550 | 15,9300 | 15,9960 | 18,9880 | 19,0150 | 18,9880 | 19,0150 |
| | | Desvio | 16,9281 | 16,8044 | 19,6421 | 19,6392 | 22,1672 | 22,1653 | 22,1672 | 22,1653 |
| | | Mediana | 8,0 | 8,0 | 9,0 | 9,0 | 12,0 | 12,0 | 12,0 | 12,0 |
| Todas as médias mudam simultaneamente na mesma escala | Choque = 0,5 | Média | 10,1510 | 10,1670 | 12,0920 | 12,1810 | 13,6140 | 13,6140 | 14,6330 | 14,6330 |
| | | Desvio | 12,3750 | 12,3679 | 13,9577 | 13,9760 | 13,7648 | 13,7648 | 15,8842 | 15,8842 |
| | | Mediana | 6,0 | 6,0 | 7,0 | 7,0 | 9,0 | 9,0 | 9,0 | 9,0 |
| | Choque = 1 | Média | 5,7440 | 5,7500 | 7,1420 | 7,1790 | 7,7110 | 7,7110 | 8,5330 | 8,5240 |
| | | Desvio | 6,7663 | 6,7684 | 8,1594 | 8,1838 | 7,8439 | 7,8439 | 8,7263 | 8,7233 |
| | | Mediana | 4,0 | 4,0 | 5,0 | 5,0 | 5,0 | 5,0 | 6,0 | 6,0 |
| | Choque = 2 | Média | 2,5880 | 2,5910 | 3,3160 | 3,3350 | 3,2270 | 3,2270 | 3,7540 | 3,7540 |
| | | Desvio | 2,9569 | 2,9578 | 3,7598 | 3,7727 | 3,1192 | 3,1192 | 3,6385 | 3,6385 |
| | | Mediana | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 3,0 | 3,0 |
| | Choque = 3 | Média | 1,6310 | 1,6310 | 2,0740 | 2,0760 | 1,9170 | 1,9170 | 2,2660 | 2,2660 |
| | | Desvio | 1,0545 | 1,0545 | 1,7394 | 1,7456 | 1,3190 | 1,3190 | 1,6709 | 1,6709 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 |
| Choque = 4 | Média | 1,3150 | 1,3160 | 1,5510 | 1,5540 | 1,3990 | 1,3990 | 1,6150 | 1,6160 | |
| | Desvio | 0,6799 | 0,6802 | 0,9635 | 0,9613 | 0,7684 | 0,7684 | 0,9538 | 0,9547 | |
| | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| Uma média permanece constante e a outra muda ordenadamente | Choque = 0,5 | Média | 12,6670 | 12,7070 | 15,3900 | 15,3750 | 17,7690 | 17,7690 | 18,3920 | 18,3920 |
| | | Desvio | 16,0380 | 16,1920 | 19,7850 | 19,7860 | 19,6147 | 19,6147 | 19,8168 | 19,8168 |
| | | Mediana | 8,0 | 8,0 | 9,0 | 9,0 | 12,0 | 12,0 | 12,0 | 12,0 |
| | Choque = 1 | Média | 10,4700 | 10,4830 | 12,6300 | 12,8120 | 14,3680 | 14,3680 | 14,7030 | 14,7170 |
| | | Desvio | 13,7710 | 13,7690 | 15,6050 | 16,0960 | 15,4720 | 15,4720 | 15,6337 | 15,7027 |
| | | Mediana | 6,0 | 6,0 | 7,0 | 7,0 | 9,0 | 9,0 | 10,0 | 10,0 |
| | Choque = 2 | Média | 7,5510 | 7,5310 | 8,9170 | 8,9850 | 9,4530 | 9,4870 | 10,0850 | 10,0300 |
| | | Desvio | 10,0870 | 9,8970 | 11,5490 | 11,8730 | 9,1814 | 9,1815 | 9,8964 | 9,8389 |
| | | Mediana | 5,0 | 5,0 | 6,0 | 6,0 | 7,0 | 7,0 | 7,0 | 7,0 |
| | Choque = 3 | Média | 4,8680 | 4,8950 | 5,8700 | 5,9210 | 6,3040 | 6,3040 | 6,4580 | 6,4480 |
| | | Desvio | 5,1630 | 5,2570 | 6,2910 | 6,3380 | 6,3237 | 6,3237 | 6,3905 | 6,3903 |
| | | Mediana | 3,0 | 3,0 | 4,0 | 4,0 | 4,0 | 4,0 | 4,0 | 4,0 |
| Choque = 4 | Média | 3,9770 | 3,9720 | 4,4960 | 4,5210 | 4,7550 | 4,7550 | 4,8720 | 4,8700 | |
| | Desvio | 4,4070 | 4,4060 | 4,8150 | 4,8370 | 4,6954 | 4,6954 | 4,7263 | 4,7275 | |
| | Mediana | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | |

Legenda: **Pb5**: ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes**: ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

6.4 – Conclusão

De modo geral, para os processos multivariados com distribuição não-normal e não-autocorrelacionados simulados nessa dissertação, observa-se que estatística M de Hayter e Tsui (1994) foi a que forneceu melhores valores de *ARL in control* e *out of control*, mesmo para pequenas amostras com $m=50$. As cartas de controle com o método de Polansky e Baker (2000) com $b=5$ e com o de Bessegato (2001) forneceram resultados bem próximos, indicando que qualquer um dos métodos é apropriado na determinação dos limites de controle dos processos estudados.

**Capítulo 7 – Estratégia de execução: Processos
Normais e Não-Normais Multivariados
Autocorrelacionados**

7.1 – Processos Normais: Modelo VAR(1) Multivariado

Neste capítulo serão estudados processos que seguem uma distribuição normal bivariada autocorrelacionada. O processo estudado neste capítulo é o mesmo estudado por Kalgonda e Kulkarni (2004) e é descrito a seguir: Considere o vetor aleatório $Y_t = [Y_{t1} \ Y_{t2}]'$ tal que:

$$Y_t = \mu + \Phi(Y_{t-1} - \mu) + \varepsilon_t, t = 1, 2, \dots, m \quad (7.1)$$

em que

$$\mu = [0 \ 0]'$$

$$\Phi = \begin{bmatrix} 0,5 & 0 \\ 0 & 0,7 \end{bmatrix}$$

e $\varepsilon_t = [\varepsilon_{t1} \ \varepsilon_{t2}]'$ é o vetor de erros aleatório no tempo t , com distribuição normal bivariada com vetor médias nulo e matriz de covariâncias Σ_ε definida por:

$$\Sigma_\varepsilon = \begin{bmatrix} 1 & 0,5 \\ 0,5 & 1 \end{bmatrix}$$

A esperança de Y_t é dada por:

$$E(Y_t) = \mu + \Phi E(Y_{t-1} - \mu) + E(\varepsilon_t) = \mu + \mu_\varepsilon = 0$$

As matrizes de covariâncias e correlação cruzadas teóricas deste processo são dadas por:

$$\Gamma(0) = \begin{bmatrix} 1,33 & 0,77 \\ 0,77 & 1,96 \end{bmatrix} \quad \rho(0) = \begin{bmatrix} 1,00 & 0,48 \\ 0,48 & 1,00 \end{bmatrix}$$

Este processo é chamado de autoregressivo bivariado de ordem 1 e foi simulado nessa dissertação para amostras de tamanho $m=50$, 100 e 200 . Como o processo é autocorrelacionado de ordem 1, a observação amostral simulada para o tempo t está correlacionada com a observação amostral simulada do tempo $t-1$. Para se assegurar que as amostras geradas estivessem de fato representando este processo a quantidade de elementos amostrais (m) sempre foi simulada em dobro do valor desejado, ou seja, se o interesse era simular uma amostra de tamanho $m=200$ então eram gerados $m'=400$ elementos amostrais e logo em seguida os primeiros 200 elementos eram excluídos da análise. Este procedimento foi utilizado pois, através de testes computacionais, verificou-se resultados aceitáveis quando ele era utilizado. Na literatura, esse procedimento é chamado de “*burn-in*”. A matriz de autocorrelação estimada $\hat{\Phi}$, necessária para a obtenção da matriz de covariâncias-cruzada $\hat{\Gamma}(0)$ foi estimada de acordo o algoritmo *mAr.est* implementado no *software* estatístico *R for*

Windows, que calcula essa matriz pelo método de mínimos quadrados via *stepwise* (Neumaier e Schneider, 2001). A matriz de covariâncias $\hat{\Sigma}_\varepsilon$ também é obtida através desse algoritmo.

Para verificar a qualidade das cartas de controle para processos dessa natureza, calculamos os valores do *ARL in control* e *out of control*. Para calcular os valores do *ARL out of control* foram implementadas mudanças no vetor de médias do erro ε_t do processo. Dessa forma, os “choques” implementados seguiram o seguinte critério:

- As médias das p variáveis do vetor de erro mudam ordenadamente na mesma escala

$$\mu_{\varepsilon;choque} = \mu_\varepsilon + choque \quad (choque = 0,5;1;2;3;4)$$

- A média da 1ª variável permanece constante e a média da 2ª variável muda ordenadamente

$$\mu_{1,\varepsilon;choque} = \mu_{1,\varepsilon}$$

$$\mu_{2,\varepsilon;choque} = \mu_{2,\varepsilon} + choque \quad (choque = 0,5;1;2;3;4)$$

- A média da 1ª variável muda em 0,5 e a média da 2ª variável muda ordenadamente

$$\mu_{1,\varepsilon;choque} = \mu_{1,\varepsilon} + 0,5$$

$$\mu_{2,\varepsilon;choque} = \mu_{2,\varepsilon} + choque \quad (choque = 0,5;1;2;3;4)$$

sendo que a matriz de variâncias e covariâncias dos erros permanece a mesma.

Para ilustrar o comportamento da função densidade conjunta de Y_{t1} e Y_{t2} as curvas dessa função foi gerada uma amostra de tamanho $m=100000$ de acordo com o processo mostrado em (7.1) e os resultados observados são apresentados nas Figuras 7.1 e 7.2.

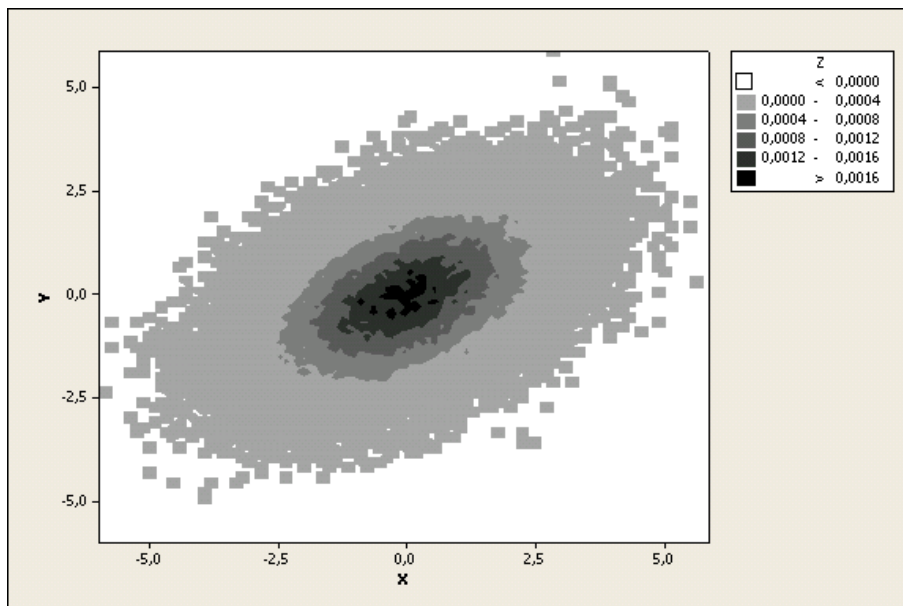


Figura 7.1: Curvas de nível amostral da função densidade do processo normal bivariado autocorrelacionado VAR (1).

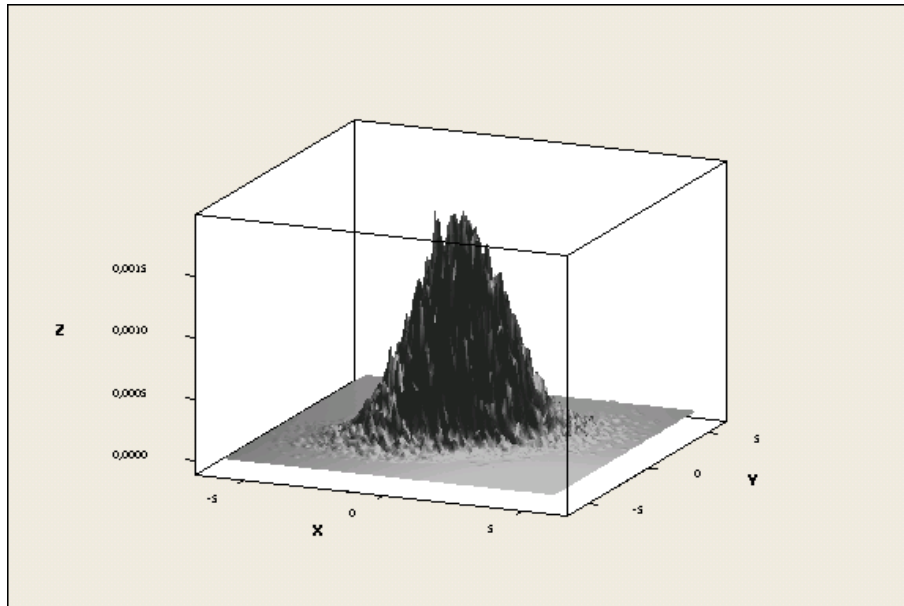


Figura 7.2: Função densidade amostral do processo normal bivariado $VAR(1)$.

Para esse processo foram obtidos limites de controle à 95% de confiança através da técnica de núcleo-estimador, tendo como função núcleo a distribuição normal padrão sendo as janelas ótimas h determinadas pelos métodos de Polansky e Baker (2000) com $b=1$ e $b=5$ e Bessegato (2001), e através da função de distribuição empírica. Foram simuladas num total $k=1000$ réplicas desse processo.

7.2 – Análise de Resultados: Processos Normais Multivariados Autocorrelacionados

As cartas de controle foram construídas usando as mesmas estatísticas usadas no capítulo 5 para processos normais multivariados não-autocorrelacionados. Nas Tabelas 7.1 a 7.3 apresenta-se os limites de controle estimados a 95% de confiança e as estimativas dos ARL *in control* e *out of control* para amostras de tamanho $m=50, 100$ e 200 .

Para obter os valores paramétricos das estatísticas T^2 de Hotelling (1947), estatística M de Hayter e Tsui(1994) e componentes principais foram utilizadas as matrizes de covariâncias e correlação cruzadas $\Gamma(0)$ e $\rho(0)$ e as equações descritas no capítulo 2 dessa dissertação. Os limites de controle não-paramétricos foram obtidos com base nos valores amostrais das estatísticas T^2 de Hotelling (1947), estatística M de Hayter e Tsui(1994) e componentes principais utilizando as matrizes de covariância e correlação cruzadas $\hat{\Gamma}(0)$ e $\hat{\rho}(0)$ (que são obtidas através da utilização do algoritmo *mAr.est* e da solução do sistema linear de Yule-Walker descrito na seção 2.2.5 dessa dissertação).

Como nos processos que tinham distribuição normal multivariada não-autocorrelacionada (capítulo 5) o método Polansky e Baker (2000) com $b=1$ forneceu estimativas dos limites de controle muito próximos aos valores do método com $b=5$. Como a função de distribuição empírica não apresentou um bom resultado, foi realizada uma análise preliminar dos resultados do processo simulado nesse capítulo, quando onde foi verificado um comportamento bastante similar ao que ocorria nos processos simulados no capítulo 5. Dessa forma, os resultados que serão apresentados referem-se apenas àqueles em que os limites de controle amostrais foram obtidos pelos métodos de Polansky e Baker (2000) com $b=5$ e Bessegato (2001).

Observa-se nas Tabelas 7.1 a 7.3 que a estatística T^2 de Hotelling (1947) forneceu na maioria dos casos menores valores dos *ARL in control* quando se utiliza o método de núcleo-estimador para estimar os limites de controle gerando muitos alarmes falsos. Já primeira componente principal forneceu maiores valores dos *ARL out of control* indicando que mesmo quando o processo está fora de controle os limites traçados levam muito tempo para detectar mudanças nos vetores de média do processo. Os valores do *ARL in control* fornecidos pela estatística M estão localizados entre os valores do *ARL in control* da estatística T^2 de Hotelling e da primeira componente principal.

Os valores do *ARL out of control* da estatística M , para aqueles choques em que as duas médias do processo são modificadas, são sempre menores que os *ARL out of control* da estatística T^2 de Hotelling e da primeira componente principal. Para os choques em que apenas uma das médias muda e nos choques em que uma média muda um pouco e a outra muda ordenadamente, observa-se que na maioria das vezes o menor *ARL out of control* é fornecido pela da estatística T^2 de Hotelling sendo que os valores do *ARL out of control* para a estatística M são sempre muito próximos aos valores da estatística T^2 de Hotelling, porém sempre maiores. Os valores do *ARL out of control* da primeira componente principal são sempre maiores que os valores das outras estatísticas indicando a dificuldade dessa estatística em detectar a falta de controle do processo.

As cartas construídas usando os métodos de Polansky e Baker (2000) com $b=5$ e Bessegato (2001) forneceram valores de *ARL in control* e *out of control* muito semelhantes mostrando que qualquer um dos métodos pode ser utilizado para a estimar-se os limites das cartas de controle.

Como era esperado, a medida que os tamanhos das amostras aumentam, aumentam os valores do *ARL in control* da estatística T^2 de Hotelling e M e diminuem os valores do *ARL in control* da primeira componente principal. Os valores do *ARL out of control* seguiram o mesmo comportamento dos menores do *ARL in control*.

Em geral, a presença da autocorrelação não alterou significativamente os valores dos *ARL in control* e *out of control* quando comparamos seus resultados com os resultados obtidos para os processos normais não-autocorrelacionados. Os limites de controle conseguiram detectar com maior rapidez aqueles choques em que todas as médias do processo mudam simultaneamente na mesma escala conclusão que também vai de encontro às conclusões obtidas para os processos normais não-autocorrelacionados com $p=2$ variáveis simulados no capítulo 5.

Tabela 7.1: Limites de controle, *ARL in control* e *out of control*. Processo com distribuição normal bivariada autocorrelacionada. Modelo VAR (1) ($m=50$).

| Distribuição normal bivariada autocorrelacionada ($m=50$) | Medidas Descritivas | Estatística T ² de Hotelling | | | Estatística M | | | Componentes Principais | | | | | |
|---|---------------------|---|---------|---------|---------------|---------|---------|------------------------|---------|---------|---------|---------|---------|
| | | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | | Pb=5 | | Bes | |
| Limites de Controle | Média | 6,6447 | 5,5383 | 5,5619 | 2,2101 | 2,1621 | 2,1774 | -3,0851 | 3,0851 | -3,1165 | 3,0954 | -3,1518 | 3,1276 |
| | Desvio | 0,0000 | 0,7868 | 0,7609 | 0,0236 | 0,1347 | 0,1293 | 0,0000 | 0,0000 | 0,7017 | 0,7147 | 0,7011 | 0,7174 |
| | Mediana | 6,6447 | 5,4360 | 5,4536 | 2,2116 | 2,1509 | 2,1675 | -3,0851 | 3,0851 | -3,0638 | 3,0528 | -3,1162 | 3,0905 |
| ARL in control | Média | 23,6080 | 16,6060 | 16,7920 | 21,4380 | 19,9340 | 20,4320 | 28,1590 | 32,3680 | 34,2080 | 34,2080 | 34,2080 | 34,2080 |
| | Desvio | 29,8291 | 31,0708 | 31,2186 | 28,9162 | 25,2607 | 25,4702 | 30,0998 | 61,8097 | 64,2631 | 64,2631 | 64,2631 | 64,2631 |
| | Mediana | 14,0 | 9,0 | 9,0 | 13,0 | 11,0 | 12,0 | 19,0 | 17,0 | 17,0 | 17,0 | 17,0 | 17,0 |
| Todas as médias mudam simultaneamente na mesma escala | Choque = 0.5 | Média | 7,9140 | 6,1360 | 6,1700 | 6,3470 | 6,0440 | 6,1090 | 6,8330 | 7,8700 | 8,1430 | 8,1430 | 8,1430 |
| | | Desvio | 8,0122 | 6,0208 | 6,1190 | 6,2631 | 5,8953 | 5,9959 | 5,5546 | 9,4335 | 10,1030 | 10,1030 | 10,1030 |
| | | Mediana | 5,0 | 4,0 | 4,0 | 5,0 | 4,0 | 4,0 | 5,0 | 5,0 | 5,0 | 5,0 | 5,0 |
| | Choque = 1 | Média | 2,9570 | 2,5710 | 2,5790 | 2,6270 | 2,4810 | 2,5180 | 2,7250 | 2,8640 | 2,9040 | 2,9040 | 2,9040 |
| | | Desvio | 2,0502 | 1,8117 | 1,8235 | 1,7762 | 1,7078 | 1,7282 | 1,6734 | 2,0351 | 2,0773 | 2,0773 | 2,0773 |
| | | Mediana | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 |
| | Choque = 2 | Média | 1,3770 | 1,2980 | 1,2990 | 1,2950 | 1,2740 | 1,2810 | 1,3450 | 1,3610 | 1,3730 | 1,3730 | 1,3730 |
| | | Desvio | 0,5650 | 0,5152 | 0,5156 | 0,5121 | 0,4992 | 0,5043 | 0,5406 | 0,5718 | 0,5798 | 0,5798 | 0,5798 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| | Choque = 3 | Média | 1,0530 | 1,0360 | 1,0360 | 1,0360 | 1,0340 | 1,0350 | 1,0580 | 1,0700 | 1,0740 | 1,0740 | 1,0740 |
| | | Desvio | 0,2241 | 0,1864 | 0,1864 | 0,1864 | 0,1813 | 0,1839 | 0,2339 | 0,2592 | 0,2657 | 0,2657 | 0,2657 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| Choque = 4 | Média | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0010 | 1,0010 | 1,0010 | 1,0010 | 1,0010 | |
| | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0316 | 0,0316 | 0,0316 | 0,0316 | 0,0316 | |
| | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| Uma média permanece constante e a outra muda ordenadamente | Choque = 0.5 | Média | 8,3650 | 6,4110 | 6,4140 | 8,7390 | 8,4380 | 8,6020 | 11,3480 | 14,0120 | 14,2860 | 14,2860 | 14,2860 |
| | | Desvio | 8,3843 | 6,4952 | 6,4330 | 8,7364 | 9,2522 | 9,3816 | 12,1974 | 21,7922 | 21,8431 | 21,8431 | 21,8431 |
| | | Mediana | 6,0 | 5,0 | 5,0 | 6,0 | 6,0 | 6,0 | 8,0 | 8,0 | 8,0 | 8,0 | 8,0 |
| | Choque = 1 | Média | 3,3800 | 2,9520 | 2,9700 | 3,6780 | 3,5530 | 3,5850 | 5,1670 | 5,3770 | 5,5500 | 5,5500 | 5,5500 |
| | | Desvio | 2,2722 | 1,9615 | 1,9715 | 2,4549 | 2,4391 | 2,4456 | 4,1755 | 4,7583 | 5,0041 | 5,0041 | 5,0041 |
| | | Mediana | 3,0 | 2,0 | 2,5 | 3,0 | 3,0 | 3,0 | 4,0 | 4,0 | 4,0 | 4,0 | 4,0 |
| | Choque = 2 | Média | 1,6450 | 1,5090 | 1,5130 | 1,7460 | 1,7040 | 1,7160 | 2,4060 | 2,4520 | 2,4760 | 2,4760 | 2,4760 |
| | | Desvio | 0,7180 | 0,6452 | 0,6452 | 0,7925 | 0,7739 | 0,7797 | 1,5111 | 1,8355 | 1,8365 | 1,8365 | 1,8365 |
| | | Mediana | 2,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 |
| | Choque = 3 | Média | 1,2180 | 1,1630 | 1,1660 | 1,2890 | 1,2740 | 1,2780 | 1,6530 | 1,6390 | 1,6520 | 1,6520 | 1,6520 |
| | | Desvio | 0,4155 | 0,3696 | 0,3723 | 0,4730 | 0,4638 | 0,4658 | 1,4624 | 0,9019 | 0,9098 | 0,9098 | 0,9098 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| Choque = 4 | Média | 1,0380 | 1,0250 | 1,0250 | 1,0840 | 1,0730 | 1,0760 | 1,3210 | 1,3330 | 1,3390 | 1,3390 | 1,3390 | |
| | Desvio | 0,1965 | 0,1562 | 0,1562 | 0,2775 | 0,2603 | 0,2651 | 0,6873 | 0,5696 | 0,5713 | 0,5713 | 0,5713 | |
| | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| Uma média muda pouco e outra muda ordenadamente | Choque = 0.5 | Média | 7,9800 | 6,0660 | 6,0970 | 6,5390 | 6,3260 | 6,4250 | 6,5050 | 7,7910 | 8,0440 | 8,0440 | 8,0440 |
| | | Desvio | 8,8044 | 6,1025 | 6,1058 | 7,0932 | 7,0337 | 7,0762 | 5,4328 | 9,6200 | 10,1717 | 10,1717 | 10,1717 |
| | | Mediana | 5,0 | 4,0 | 4,0 | 5,0 | 4,0 | 4,0 | 5,0 | 5,0 | 5,0 | 5,0 | 5,0 |
| | Choque = 1 | Média | 3,7890 | 3,2910 | 3,2990 | 3,3990 | 3,3470 | 3,3700 | 3,6080 | 3,7070 | 3,7620 | 3,7620 | 3,7620 |
| | | Desvio | 2,5564 | 2,2359 | 2,2411 | 2,4827 | 2,4691 | 2,4967 | 2,4332 | 2,8230 | 2,8424 | 2,8424 | 2,8424 |
| | | Mediana | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 |
| | Choque = 2 | Média | 1,8120 | 1,6530 | 1,6580 | 1,7790 | 1,7430 | 1,7530 | 1,9490 | 1,9960 | 2,0180 | 2,0180 | 2,0180 |
| | | Desvio | 0,8083 | 0,7437 | 0,7467 | 0,8092 | 0,8072 | 0,8079 | 0,9356 | 0,9833 | 0,9933 | 0,9933 | 0,9933 |
| | | Mediana | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 |
| | Choque = 3 | Média | 1,2650 | 1,2020 | 1,2020 | 1,2590 | 1,2480 | 1,2510 | 1,4130 | 1,4470 | 1,4500 | 1,4500 | 1,4500 |
| | | Desvio | 0,4550 | 0,4066 | 0,4066 | 0,4473 | 0,4390 | 0,4407 | 0,5734 | 0,6063 | 0,6065 | 0,6065 | 0,6065 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| Choque = 4 | Média | 1,0710 | 1,0490 | 1,0490 | 1,0760 | 1,0680 | 1,0700 | 1,2160 | 1,2310 | 1,2390 | 1,2390 | 1,2390 | |
| | Desvio | 0,2570 | 0,2160 | 0,2160 | 0,2689 | 0,2519 | 0,2553 | 0,4895 | 0,4794 | 0,5100 | 0,5100 | 0,5100 | |
| | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |

Legenda: Teo: Limites de controle e ARL para o Limite de Controle obtido de forma teórica; Pb5: Limites de controle e ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; Bes: Limites de Controle e ARL para o Limite de Controle obtido pela a técnica de Besseगत (2001).

Tabela 7.2: Limites de controle, *ARL in control* e *out of control*. Processo com distribuição normal bivariada autocorrelacionada. Modelo VAR (1) ($m=100$).

| Distribuição normal bivariada autocorrelacionada ($m=100$) | | Medidas Descritivas | Estatística T^2 de Hotelling | | | Estatística M | | | Componentes Principais | | | | | |
|--|--|---------------------|--------------------------------|---------|---------|---------------|---------|---------|------------------------|---------|---------|---------|---------|--------|
| | | | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | | Pb=5 | | Bes | |
| Limites de Controle | | Média | 6,3039 | 5,7416 | 5,7477 | 2,2128 | 2,1901 | 2,1962 | -3,0851 | 3,0851 | -3,1848 | 3,1964 | -3,2070 | 3,2188 |
| | | Desvio | 0,0000 | 0,5541 | 0,5475 | 0,0218 | 0,1059 | 0,1035 | 0,0000 | 0,0000 | 0,5366 | 0,5737 | 0,5338 | 0,5718 |
| | | Mediana | 6,3039 | 5,6730 | 5,6837 | 2,2126 | 2,1824 | 2,1878 | -3,0851 | 3,0851 | -3,1552 | 3,1700 | -3,1724 | 3,1922 |
| ARL in control | | Média | 25,2140 | 19,2230 | 19,2640 | 22,6400 | 22,1970 | 22,5950 | 27,8000 | | 34,8120 | | 36,4350 | |
| | | Desvio | 30,5373 | 23,3064 | 23,3467 | 26,1421 | 27,5639 | 28,0965 | 27,8751 | | 61,7996 | | 63,2763 | |
| | | Mediana | 15,5 | 12,0 | 12,0 | 14,0 | 14,0 | 14,0 | 14,0 | | 20,0 | | 21,0 | |
| Todas as médias mudam simultaneamente na mesma escala | | Choque = 0.5 | | Média | 7,7130 | 6,6380 | 6,6640 | 6,5980 | 6,4140 | 6,4530 | 6,3900 | 7,4540 | 7,6120 | |
| | | | | Desvio | 7,3261 | 6,2037 | 6,2130 | 6,0303 | 5,6596 | 5,6883 | 4,8451 | 7,8261 | 7,9248 | |
| | | | | Mediana | 6,0 | 5,0 | 5,0 | 5,0 | 5,0 | 5,0 | 5,0 | 5,0 | 5,0 | |
| | | Choque = 1 | | Média | 2,9670 | 2,7400 | 2,7380 | 2,6720 | 2,6570 | 2,6610 | 2,6960 | 2,8630 | 2,8950 | |
| | | | | Desvio | 1,9161 | 1,7296 | 1,7298 | 1,6337 | 1,6407 | 1,6404 | 1,5780 | 1,7667 | 1,7897 | |
| | | | | Mediana | 3,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 3,0 | 3,0 | |
| | | Choque = 2 | | Média | 1,3540 | 1,3170 | 1,3170 | 1,2740 | 1,2780 | 1,2790 | 1,3160 | 1,3600 | 1,3700 | |
| | | | | Desvio | 0,5541 | 0,5261 | 0,5261 | 0,4890 | 0,4909 | 0,4913 | 0,5083 | 0,5646 | 0,5687 | |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| | | Choque = 3 | | Média | 1,0550 | 1,0400 | 1,0400 | 1,0350 | 1,0340 | 1,0350 | 1,0450 | 1,0610 | 1,0610 | |
| | | | | Desvio | 0,2281 | 0,1961 | 0,1961 | 0,1839 | 0,1813 | 0,1839 | 0,2074 | 0,2395 | 0,2395 | |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| Choque = 4 | | Média | 1,0020 | 1,0010 | 1,0010 | 1,0000 | 1,0010 | 1,0010 | 1,0000 | 1,0030 | 1,0030 | | | |
| | | Desvio | 0,0447 | 0,0316 | 0,0316 | 0,0000 | 0,0316 | 0,0316 | 0,0000 | 0,0547 | 0,0547 | | | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | | | |
| Uma média permanece constante e a outra muda ordenadamente | | Choque = 0.5 | | Média | 8,7420 | 7,5620 | 7,5760 | 9,6920 | 9,6210 | 9,6980 | 11,1510 | 13,8910 | 14,0950 | |
| | | | | Desvio | 7,9115 | 6,9204 | 6,9287 | 10,6389 | 10,8304 | 10,8764 | 9,4034 | 16,3284 | 16,6112 | |
| | | | | Mediana | 6,0 | 6,0 | 6,0 | 7,0 | 7,0 | 7,0 | 8,0 | 9,0 | 9,0 | |
| | | Choque = 1 | | Média | 3,3230 | 3,1330 | 3,1380 | 3,7400 | 3,6780 | 3,6920 | 4,8980 | 5,4350 | 5,5090 | |
| | | | | Desvio | 1,9325 | 1,8635 | 1,8672 | 2,4422 | 2,3668 | 2,4288 | 3,4802 | 4,4269 | 4,5167 | |
| | | | | Mediana | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 4,0 | 4,0 | 4,0 | |
| | | Choque = 2 | | Média | 1,6370 | 1,5690 | 1,5680 | 1,7850 | 1,7740 | 1,7810 | 2,2530 | 2,3580 | 2,3670 | |
| | | | | Desvio | 0,7012 | 0,6676 | 0,6662 | 0,8182 | 0,8060 | 0,8134 | 1,2436 | 1,2859 | 1,2861 | |
| | | | | Mediana | 2,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | |
| | | Choque = 3 | | Média | 1,1930 | 1,1620 | 1,1620 | 1,2730 | 1,2570 | 1,2600 | 1,5390 | 1,5790 | 1,5870 | |
| | | | | Desvio | 0,4049 | 0,3740 | 0,3740 | 0,4524 | 0,4463 | 0,4479 | 0,6472 | 0,6725 | 0,6730 | |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| Choque = 4 | | Média | 1,0280 | 1,0200 | 1,0200 | 1,0600 | 1,0570 | 1,0570 | 1,2270 | 1,2500 | 1,2540 | | | |
| | | Desvio | 0,1651 | 0,1401 | 0,1401 | 0,2376 | 0,2320 | 0,2320 | 0,4401 | 0,4513 | 0,4535 | | | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | | | |
| Uma média muda pouco e outra muda ordenadamente | | Choque = 0.5 | | Média | 7,5960 | 6,6660 | 6,6550 | 6,4430 | 6,3790 | 6,4090 | 6,4330 | 7,6300 | 7,7800 | |
| | | | | Desvio | 6,7414 | 6,0363 | 6,0202 | 5,4616 | 5,5983 | 5,6122 | 4,9844 | 7,0305 | 7,1489 | |
| | | | | Mediana | 6,0 | 5,0 | 5,0 | 5,0 | 5,0 | 5,0 | 5,0 | 6,0 | 6,0 | |
| | | Choque = 1 | | Média | 3,8490 | 3,5110 | 3,5190 | 3,5130 | 3,4360 | 3,4650 | 3,4660 | 3,7210 | 3,7710 | |
| | | | | Desvio | 2,6765 | 2,5104 | 2,5152 | 2,4409 | 2,3583 | 2,3922 | 2,3367 | 2,6699 | 2,7214 | |
| | | | | Mediana | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | |
| | | Choque = 2 | | Média | 1,7940 | 1,7180 | 1,7180 | 1,7700 | 1,7510 | 1,7600 | 1,8780 | 1,9660 | 1,9720 | |
| | | | | Desvio | 0,7989 | 0,7610 | 0,7610 | 0,7923 | 0,7910 | 0,7919 | 0,9066 | 0,9777 | 0,9799 | |
| | | | | Mediana | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | |
| | | Choque = 3 | | Média | 1,2620 | 1,2310 | 1,2310 | 1,2760 | 1,2670 | 1,2680 | 1,3930 | 1,4350 | 1,4440 | |
| | | | | Desvio | 0,4445 | 0,4288 | 0,4288 | 0,4583 | 0,4538 | 0,4543 | 0,5629 | 0,5711 | 0,5755 | |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| Choque = 4 | | Média | 1,0600 | 1,0510 | 1,0510 | 1,0620 | 1,0630 | 1,0630 | 1,1630 | 1,1790 | 1,1850 | | | |
| | | Desvio | 0,2376 | 0,2201 | 0,2201 | 0,2413 | 0,2431 | 0,2431 | 0,3723 | 0,3887 | 0,3936 | | | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | | | |

Legenda: **Teo:** Limites de controle e ARL para o Limite de Controle obtido de forma teórica; **Pb5:** Limites de controle e ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes:** Limites de Controle e ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

Tabela 7.3: Limites de controle, *ARL in control* e *out of control*. Processo com distribuição normal bivariada autocorrelacionada. Modelo VAR (1) ($m=200$).

| Distribuição normal bivariada autocorrelacionada ($m=200$) | | Medidas Descritivas | Estatística T ² de Hotelling | | | Estatística M | | | Componentes Principais | | | | | | | |
|--|--|---------------------|---|---------|---------|---------------|---------|---------|------------------------|--------|---------|--------|---------|--------|---------|--|
| | | | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | | Pb=5 | | Bes | | | |
| Limites de Controle | | Média | 6,1443 | 5,8610 | 5,8629 | 2,2130 | 2,2086 | 2,2109 | -3,0851 | 3,0851 | -3,1465 | 3,1559 | -3,1544 | 3,1635 | | |
| | | Desvio | 0,0000 | 0,4290 | 0,4267 | 0,0195 | 0,0808 | 0,0797 | 0,0000 | 0,0000 | 0,3954 | 0,3909 | 0,3944 | 0,3902 | | |
| | | Mediana | 6,1443 | 5,8258 | 5,8246 | 2,2126 | 2,2052 | 2,2071 | -3,0851 | 3,0851 | -3,1471 | 3,1489 | -3,1577 | 3,1535 | | |
| <i>ARL in control</i> | | Média | 25,6920 | 23,5920 | 23,5950 | 24,4040 | 24,5180 | 24,5950 | 28,3170 | | 31,0780 | | 31,4750 | | | |
| | | Desvio | 28,0919 | 24,4897 | 24,4880 | 25,7354 | 26,8083 | 26,8479 | 26,6182 | | 34,6719 | | 34,9177 | | | |
| | | Mediana | 18,0 | 16,0 | 16,0 | 15,0 | 15,0 | 15,5 | 20,0 | | 20,0 | | 20,0 | | | |
| Todas as médias mudam simultaneamente na mesma escala | | Choque = 0.5 | | Média | 7,7980 | 7,1760 | 7,1830 | 6,7280 | 6,7320 | 6,7430 | 6,6280 | | 7,1410 | | 7,1620 | |
| | | | | Desvio | 6,5913 | 6,0476 | 6,0442 | 5,6037 | 5,6617 | 5,6596 | 5,1771 | | 6,3433 | | 6,3536 | |
| | | | | Mediana | 6,0 | 6,0 | 6,0 | 5,0 | 5,0 | 5,0 | 5,0 | | 5,0 | | 5,0 | |
| | | Choque = 1 | | Média | 2,8630 | 2,7380 | 2,7420 | 2,6780 | 2,6570 | 2,6610 | 2,5910 | | 2,6880 | | 2,6980 | |
| | | | | Desvio | 1,8143 | 1,7316 | 1,7304 | 1,6562 | 1,6686 | 1,6730 | 1,5219 | | 1,6332 | | 1,6448 | |
| | | | | Mediana | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | | 2,0 | | 2,0 | |
| | | Choque = 2 | | Média | 1,3370 | 1,3220 | 1,3220 | 1,2820 | 1,2880 | 1,2900 | 1,3240 | | 1,3590 | | 1,3610 | |
| | | | | Desvio | 0,5213 | 0,5163 | 0,5163 | 0,4740 | 0,4830 | 0,4839 | 0,5170 | | 0,5481 | | 0,5486 | |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | | 1,0 | | 1,0 | |
| | | Choque = 3 | | Média | 1,0400 | 1,0350 | 1,0350 | 1,0310 | 1,0280 | 1,0280 | 1,0260 | | 1,0370 | | 1,0380 | |
| | | | | Desvio | 0,1961 | 0,1839 | 0,1839 | 0,1734 | 0,1651 | 0,1651 | 0,1592 | | 0,1889 | | 0,1913 | |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | | 1,0 | | 1,0 | |
| Choque = 4 | | Média | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0010 | | 1,0010 | | 1,0010 | | | |
| | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0316 | | 0,0316 | | 0,0316 | | | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | | 1,0 | | 1,0 | | | |
| Uma média permanece constante e a outra muda ordenadamente | | Choque = 0.5 | | Média | 8,2440 | 7,6730 | 7,6810 | 9,0540 | 9,0400 | 9,1090 | 11,0460 | | 12,0030 | | 12,0420 | |
| | | | | Desvio | 6,9694 | 6,5513 | 6,5587 | 7,9232 | 8,1956 | 8,2639 | 9,7271 | | 11,1881 | | 11,1891 | |
| | | | | Mediana | 6,0 | 6,0 | 6,0 | 7,0 | 7,0 | 7,0 | 8,0 | | 9,0 | | 9,0 | |
| | | Choque = 1 | | Média | 3,3560 | 3,2120 | 3,2130 | 3,8560 | 3,8190 | 3,8230 | 5,0100 | | 5,3240 | | 5,3390 | |
| | | | | Desvio | 2,0311 | 1,9852 | 1,9858 | 2,4489 | 2,4495 | 2,4506 | 3,2858 | | 4,0056 | | 4,0058 | |
| | | | | Mediana | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 4,0 | | 4,0 | | 4,0 | |
| | | Choque = 2 | | Média | 1,6190 | 1,5800 | 1,5810 | 1,7550 | 1,7460 | 1,7500 | 2,2110 | | 2,2860 | | 2,2900 | |
| | | | | Desvio | 0,6755 | 0,6618 | 0,6617 | 0,7386 | 0,7362 | 0,7362 | 1,1008 | | 1,1788 | | 1,1812 | |
| | | | | Mediana | 2,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 | 2,0 | | 2,0 | | 2,0 | |
| | | Choque = 3 | | Média | 1,1860 | 1,1760 | 1,1760 | 1,2840 | 1,2810 | 1,2820 | 1,5470 | | 1,5830 | | 1,5850 | |
| | | | | Desvio | 0,3893 | 0,3810 | 0,3810 | 0,4578 | 0,4607 | 0,4612 | 0,6198 | | 0,6539 | | 0,6536 | |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | | 2,0 | | 2,0 | |
| Choque = 4 | | Média | 1,0250 | 1,0210 | 1,0210 | 1,0530 | 1,0530 | 1,0530 | 1,2150 | | 1,2340 | | 1,2350 | | | |
| | | Desvio | 0,1562 | 0,1435 | 0,1435 | 0,2241 | 0,2241 | 0,2241 | 0,4183 | | 0,4352 | | 0,4381 | | | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | | 1,0 | | 1,0 | | | |
| Uma média muda pouco e outra muda ordenadamente | | Choque = 0.5 | | Média | 7,2300 | 6,6780 | 6,6790 | 6,4570 | 6,3960 | 6,3910 | 6,2010 | | 6,8480 | | 6,8680 | |
| | | | | Desvio | 5,9102 | 5,6361 | 5,6358 | 5,2532 | 5,3486 | 5,3343 | 4,8380 | | 5,8881 | | 5,9028 | |
| | | | | Mediana | 6,0 | 5,0 | 5,0 | 5,0 | 5,0 | 5,0 | 5,0 | | 5,0 | | 5,0 | |
| | | Choque = 1 | | Média | 3,8400 | 3,7030 | 3,7050 | 3,6120 | 3,5880 | 3,5960 | 3,4470 | | 3,6140 | | 3,6350 | |
| | | | | Desvio | 2,4038 | 2,3401 | 2,3400 | 2,3244 | 2,2842 | 2,2821 | 2,1725 | | 2,3674 | | 2,3739 | |
| | | | | Mediana | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | | 3,0 | | 3,0 | |
| | | Choque = 2 | | Média | 1,7980 | 1,7580 | 1,7580 | 1,7640 | 1,7610 | 1,7640 | 1,8490 | | 1,8940 | | 1,8990 | |
| | | | | Desvio | 0,7548 | 0,7416 | 0,7416 | 0,7422 | 0,7513 | 0,7542 | 0,8191 | | 0,8681 | | 0,8681 | |
| | | | | Mediana | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | | 2,0 | | 2,0 | |
| | | Choque = 3 | | Média | 1,2680 | 1,2570 | 1,2570 | 1,2600 | 1,2640 | 1,2650 | 1,3950 | | 1,4120 | | 1,4130 | |
| | | | | Desvio | 0,4499 | 0,4440 | 0,4440 | 0,4479 | 0,4500 | 0,4505 | 0,5686 | | 0,5767 | | 0,5769 | |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | | 1,0 | | 1,0 | |
| Choque = 4 | | Média | 1,0580 | 1,0500 | 1,0500 | 1,0640 | 1,0620 | 1,0620 | 1,1400 | | 1,1570 | | 1,1590 | | | |
| | | Desvio | 0,2339 | 0,2181 | 0,2181 | 0,2449 | 0,2413 | 0,2413 | 0,3472 | | 0,3640 | | 0,3659 | | | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | | 1,0 | | 1,0 | | | |

Legenda: **Teo:** Limites de controle e ARL para o Limite de Controle obtido de forma teórica; **Pb5:** Limites de controle e ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes:** Limites de Controle e ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

7.3 – Processos não normais: Modelo VAR(1) multivariado não-normal

O processo que será simulado nesse capítulo também é um processo bivariado autoregressivo de ordem 1 ($VAR(1)$) como descrito em (7.1), porém, os erros desse modelo não seguem a distribuição normal bivariada, e sim a distribuição uniforme. Considere o vetor aleatório $Y_t = [Y_{t1} \ Y_{t2}]'$ tal que:

$$Y_t = \mu + \Phi(Y_{t-1} - \mu) + \varepsilon_t, t = 1, 2, \dots, m \quad (7.2)$$

em que

$$\mu = [0 \ 0]'$$

$$\Phi = \begin{bmatrix} 0,5 & 0 \\ 0 & 0,7 \end{bmatrix}$$

e $\varepsilon_t = [\varepsilon_{t1} \ \varepsilon_{t2}]'$ é o vetor de erros aleatório no tempo t , com distribuição uniforme definida por:

$$\varepsilon_{t,1} \sim U(-2,2)$$

$$\varepsilon_{t,2} \sim U(-2,2)$$

sendo que ε_{t1} e ε_{t2} são independentes. Dessa forma, a esperança e a variância dos erros são iguais a:

$$E(\varepsilon_{t,i}) = 0, i = 1, 2$$

$$Var(\varepsilon_{t,i}) = 1,3333, i = 1, 2$$

Como a distribuição de cada variável do vetor do erro ε_t é uniforme e ε_{t1} e ε_{t2} são independentes, de forma indireta, pode-se considerar a matriz de covariâncias dos erros como:

$$\Sigma_\varepsilon = \begin{bmatrix} 1,333 & 0 \\ 0 & 1,333 \end{bmatrix}$$

Sendo a esperança de Y_t dada por:

$$E(Y_t) = \mu + \Phi E(Y_{t-1} - \mu) + E(\varepsilon_t) = 0 \quad (7.3)$$

As matrizes $\hat{\Gamma}(0)$ e $\hat{\Sigma}_\varepsilon$ são obtidas através do método de mínimos quadrados via *stepwise* (Neumaier e Schneider, 2001) utilizando o pacote *mAr.est* já implementado no *software R for Windows*. Para verificar a qualidade dos limites de controle para esse processo estimamos os valores do *ARL in control* e *out of control*. Para o cálculo do *ARL out of control* foram implementados “choques” no vetor de médias do erro do processo.

Para avaliar a sensibilidade dos limites de controle em detectar mudanças na média do processo foram implementados “choques” na média do processo. Como a média do vetor

aleatório Y_t depende da média do processo (μ) e da média dos erros (μ_ε), como pode ser verificado em (7.3) optou-se nessa dissertação efetuar a mudança na média do vetor de erros mantendo a média do processo fixa no valor zero ($\mu=0$).

Como o vetor dos erros tem uma distribuição uniforme, para efetuar a mudança na média do vetor de erro deve se preocupar também em manter o valor da variância do modelo constante. Dessa forma foi criado um sistema linear que leva em consideração os parâmetros da distribuição uniforme e que, ao efetuar a mudança na média do vetor de erros a variância permaneça constante. A partir desse sistema é possível determinar qual deve ser a intensidade dos “choques” de forma que se observe a mudança no vetor de média do processo. Assim, considerando a distribuição uniforme com parâmetros a e b ($U(a,b)$) e considerado a variância constante em 1,333, temos o seguinte sistema linear:

$$\begin{cases} \mu_{choque} = \frac{b_{choque} + a_{choque}}{2} \Rightarrow b_{choque} = 2\mu_{choque} - a_{choque} \\ 1,3333 = \frac{(b_{choque} - a_{choque})^2}{12} \Rightarrow a_{choque} = b_{choque} - \sqrt{16} \Rightarrow a_{choque} = b_{choque} - 4 \end{cases}$$

Assim, considerando esse sistema linear, suponha que desejamos gerar um vetor Y_t com média 0,5 e a variância igual a 1,333. Então, os valores dos parâmetros da distribuição uniforme, responsável pelo choque, devem ser iguais a:

$$\begin{cases} b_{choque} = 2 \times 0,5 - a_{choque} \Rightarrow b_{choque} = 1 - a_{choque} \Rightarrow b_{choque} = 2,5 \\ a_{choque} = b_{choque} - 4 \Rightarrow a_{choque} = 1 - a_{choque} - 4 \Rightarrow a_{choque} = -1,5 \end{cases}$$

ou seja, gerando um vetor de erros com distribuição uniforme com parâmetros $a_{choque} = -1,5$ e $b_{choque} = 2,5$ teremos vetor o Y_t com médias iguais a 0,5 e com variância igual a 1,333.

Para ilustrar o comportamento probabilístico da função densidade conjunta de Y_{11} e Y_{12} e das curvas de nível dessa distribuição, foi gerada uma amostra de tamanho igual a $m=100000$ de acordo com o processo descrito em (7.2) com os erros seguindo a distribuição $U(-2,2)$. Os gráficos da função densidade amostral e das curvas de nível desse processo são apresentados nas Figuras 7.3 e 7.4.

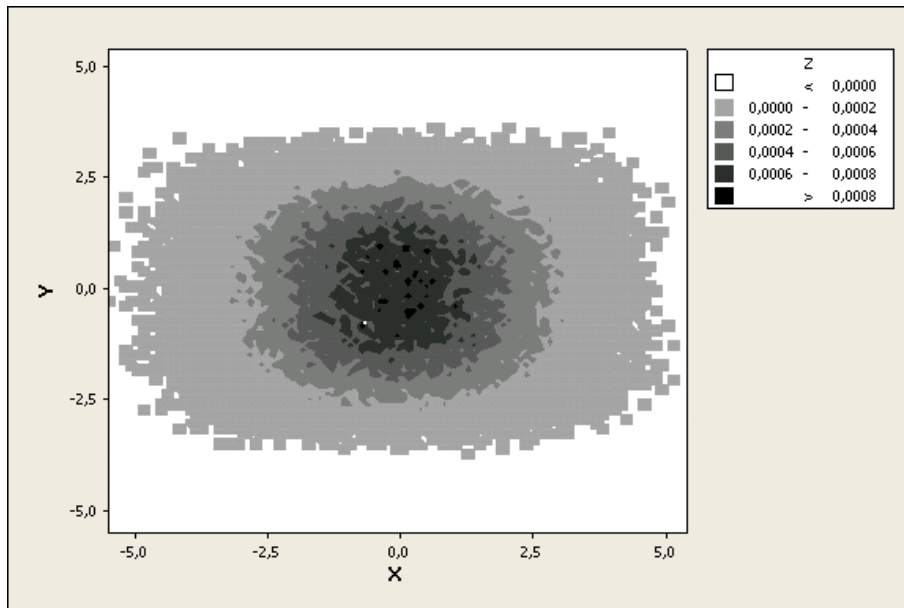


Figura 7.3: Curvas de nível amostral da função densidade do processo com distribuição não-normal bivariada autocorrelacionada $VAR(1)$.

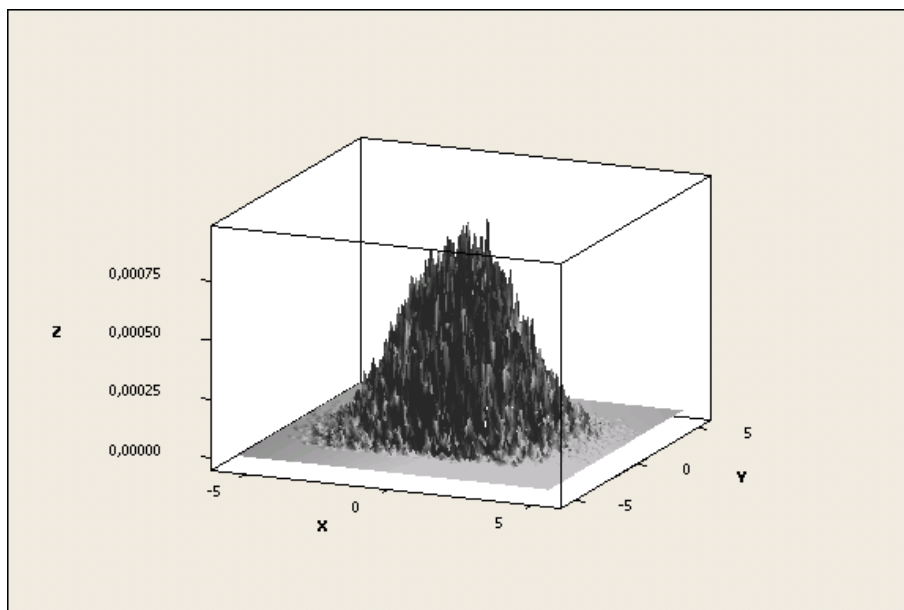


Figura 7.4: Função densidade amostral do processo com distribuição não-normal bivariada autocorrelacionada $VAR(1)$.

Para o processo não-normal bivariado autocorrelacionado foram obtidos limites de controle à 95% de confiança através da técnica de núcleo estimador, utilizando os métodos de Polansky e Baker (2000) com $b=1$ e $b=5$ e Bessegato (2001) para determinação da janela ótima h tendo como função núcleo a normal padrão, e através da função de distribuição empírica. Foram consideradas amostras de tamanho $m=50, 100$ e 200 e foram feitas $k=1000$ réplicas do processo simulado.

7.4 – Análise de Resultados: Processos Multivariados Não-Normais Autocorrelacionados

Na Tabela 7.4 são apresentadas as estimativas dos limites de controle a 95% de confiança considerando-se o núcleo estimador com os métodos de Polansky e Baker (2000) com $b=5$ e por Bessegato (2001) e os valores do *ARL in control* e *ARL out of control* do processo simulado considerando-se amostras de tamanhos $m=50, 100$ e 200 .

Serão apresentados apenas os resultados da estatística T^2 de Hotelling (1947) e estatística M de Hayter e Tsui (1994). Para esse processo em específico, como a matriz Σ_ε é uma matriz diagonal, ou seja, o vetor $\varepsilon_t = [\varepsilon_{t1} \ \varepsilon_{t2}]'$ foi gerado de modo que ε_{t1} e ε_{t2} tenham correlação igual a zero, e como a matriz Φ também é diagonal, logo, a matriz $\Gamma(0)$ também será diagonal e dessa forma, os resultados da técnica de componentes principais não são úteis pois eles representam as variáveis originais isoladamente. No entanto, se Σ_ε ou Φ não fossem diagonais a técnica de componentes principais poderia ser usada normalmente.

Serão exibidos nessa seção apenas os cálculos dos limites de controle obtidos utilizando o método estudado por Polansky e Baker (2000) com $b=5$ e o estudado por Bessegato (2001), uma vez que o método estudado por Polansky e Baker (2000) com $b=1$ apresentou desempenho similar ao observado para o caso em que $b=5$ e a função de distribuição empírica sempre apresentou pior desempenho. Como o vetor aleatório Y_t não possui a distribuição normal, os valores dos limites de controle teóricos também não serão apresentados.

Analisando a Tabela 7.4 observa-se que a estatística T^2 de Hotelling forneceu em todos os casos um menor valor do *ARL in control* e na maioria das vezes um menor valor do *ARL out of control* indicando que esses processos fornecem muitos alarmes falsos. Já a estatística M , quando utilizamos a técnica de núcleo-estimador, apresenta valores maiores do *ARL in control* e *ARL out of control* em comparação com os *ARL's* da estatística T^2 de Hotelling (1947).

Observa-se ainda nessa tabela que ao efetuar a mudança na média do processo, quando simulamos amostras de tamanho $m=50$ e 100 , os valores do *ARL out of control* são mais altos, que os valores fornecidos para a simulação em que o tamanho da amostra é igual a 200. Foi observado que em geral todos os “choques” iguais a 0,5 demoram a ser detectados pelos limites de controle indicando que essas cartas precisam de um grande número de valores ser traçado na carta para que pequenas mudanças no vetor de média sejam detectado. Dessa forma, para esse tipo de choque, eliminamos os resultados apresentados na Tabela 7.4.

A medida que aumentamos os tamanho da amostra gerada, observa-se que se aumenta também os valores do *ARL in control* tanto da estatística *M* quanto da estatística T^2 de Hotelling. Em geral, comportamento parecido com esse ocorre para os valores do *ARL out of control*. Entretanto, em alguns casos, o *ARL ou of control* também diminui quando aumentamos o tamanho da amostra. Em geral as mudanças do processo em que os “choques” de valor 4 são aplicados em todas as médias do processo são mais rapidamente detectadas pelas cartas de controle.

Conclui-se que a estatística *M* deve ser preferida para o cálculo dos limites de controle do processo, uma vez que possuímos para esses casos um maior *ARL in control*. Os métodos de Polansky e Baker (2000) com $b=5$ e Bessegato (2001) forneceram resultados parecidos o que aponta para as mesmas conclusões dos demais modelos simulados.

Tabela 7.4: Limites de controle, *ARL in control* e *out of control*. Processo com distribuição não-normal bivariada autocorrelacionada VAR(1) ($m=50, 100$ e 200).

| Distribuição não-normal multivariada autocorrelacionada $m=50, 100, 200$ | Médidas Descritivas | m= 50 | | | | m=100 | | | | M=200 | | | | |
|--|---------------------|-----------------------------|---------|---------------|---------|-----------------------------|---------|---------------|---------|-----------------------------|---------|---------------|---------|--------|
| | | T ² de Hotelling | | Estatística M | | T ² de Hotelling | | Estatística M | | T ² de Hotelling | | Estatística M | | |
| | | Pb=5 | Bes | Pb=5 | Bes | Pb=5 | Bes | Pb=5 | Bes | Pb=5 | Bes | Pb=5 | Bes | |
| Limites de Controle | Média | 4,9512 | 4,9872 | 2,0414 | 2,0573 | 5,1678 | 5,1791 | 2,0694 | 2,0761 | 5,2566 | 5,2615 | 2,0770 | 2,0799 | |
| | Desvio | 0,5150 | 0,4928 | 0,1043 | 0,1017 | 0,4110 | 0,4038 | 0,0786 | 0,0766 | 0,3144 | 0,3112 | 0,0583 | 0,0576 | |
| | Mediana | 4,9036 | 4,9393 | 2,0350 | 2,0518 | 5,1346 | 5,1476 | 2,0676 | 2,0737 | 5,2324 | 5,2386 | 2,0726 | 2,0766 | |
| ARL in control | Média | 17,2410 | 17,5420 | 23,6510 | 24,7650 | 19,7740 | 19,9070 | 25,9900 | 26,5210 | 25,5630 | 25,5450 | 26,3570 | 26,7110 | |
| | Desvio | 22,0959 | 21,5952 | 32,1441 | 32,7322 | 19,5596 | 19,6114 | 33,7372 | 34,0229 | 27,2320 | 27,2591 | 27,6841 | 28,3725 | |
| | Mediana | 11,0 | 11,0 | 13,0 | 14,0 | 14,0 | 14,0 | 15,0 | 16,0 | 17,0 | 17,0 | 17,0 | 17,0 | |
| Todas as médias mudam simultaneamente na mesma escala | Choque = 1 | Média | 3,8000 | 3,9000 | 4,1000 | 4,2000 | 3,9000 | 3,9000 | 4,1000 | 4,1000 | 3,8000 | 3,8000 | 4,1000 | 4,1000 |
| | | Desvio | 4,0000 | 4,0000 | 4,0000 | 4,0000 | 3,0000 | 3,0000 | 3,0000 | 3,0000 | 3,0000 | 3,0000 | 3,0000 | 3,0000 |
| | | Mediana | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 |
| | Choque = 2 | Média | 1,6000 | 1,6000 | 1,7000 | 1,8000 | 1,7000 | 1,7000 | 1,8000 | 1,8000 | 1,6000 | 1,6000 | 1,7000 | 1,7000 |
| | | Desvio | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 |
| | | Mediana | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 1,0 | 1,0 | 2,0 | 2,0 |
| | Choque = 3 | Média | 1,2000 | 1,2000 | 1,2000 | 1,2000 | 1,2000 | 1,2000 | 1,2000 | 1,2000 | 1,2000 | 1,2000 | 1,2000 | 1,2000 |
| | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| | Choque = 4 | Média | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,1000 | 1,1000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 |
| | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| Uma média permanece constante e a outra muda ordenadamente | Choque = 1 | Média | 5,7920 | 5,8780 | 6,5900 | 6,8270 | 5,9090 | 5,9200 | 6,3110 | 6,3670 | 5,6200 | 5,6160 | 5,7470 | 5,7850 |
| | | Desvio | 6,0788 | 6,1500 | 6,7475 | 7,0915 | 4,8195 | 4,8364 | 5,0147 | 5,0515 | 4,4373 | 4,4362 | 4,5144 | 4,6102 |
| | | Mediana | 4,0 | 4,0 | 5,0 | 5,0 | 5,0 | 5,0 | 5,0 | 5,0 | 4,0 | 4,0 | 5,0 | 5,0 |
| | Choque = 2 | Média | 2,3560 | 2,3710 | 2,4500 | 2,4830 | 2,4290 | 2,4350 | 2,5390 | 2,5440 | 2,4790 | 2,4810 | 2,5120 | 2,5180 |
| | | Desvio | 1,3539 | 1,3679 | 1,3702 | 1,3826 | 1,3125 | 1,3189 | 1,3742 | 1,3733 | 1,2819 | 1,2827 | 1,2867 | 1,2874 |
| | | Mediana | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 |
| | Choque = 3 | Média | 1,6080 | 1,6160 | 1,6470 | 1,6600 | 1,6260 | 1,6310 | 1,6580 | 1,6630 | 1,6280 | 1,6290 | 1,6390 | 1,6410 |
| | | Desvio | 0,6992 | 0,7022 | 0,7396 | 0,7463 | 0,7160 | 0,7179 | 0,7152 | 0,7183 | 0,6929 | 0,6941 | 0,6820 | 0,6831 |
| | | Mediana | 1,0 | 2,0 | 2,0 | 2,0 | 1,5 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 |
| | Choque = 4 | Média | 1,2780 | 1,2800 | 1,2950 | 1,3020 | 1,3150 | 1,3150 | 1,3260 | 1,3310 | 1,3020 | 1,3030 | 1,3070 | 1,3070 |
| | | Desvio | 0,4700 | 0,4710 | 0,4819 | 0,4848 | 0,4817 | 0,4817 | 0,4816 | 0,4834 | 0,4744 | 0,4748 | 0,4764 | 0,4764 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| Uma média muda pouco e outra muda ordenadamente | Choque = 1 | Média | 5,4900 | 5,6030 | 6,1240 | 6,3760 | 5,4060 | 5,4260 | 5,7900 | 5,8280 | 5,0670 | 5,0720 | 5,2520 | 5,2780 |
| | | Desvio | 5,6670 | 5,9360 | 7,5840 | 8,5300 | 4,1939 | 4,2086 | 4,9667 | 4,9861 | 3,6551 | 3,6666 | 3,8524 | 3,9018 |
| | | Mediana | 4,0 | 4,0 | 4,0 | 4,0 | 4,0 | 4,0 | 4,0 | 4,0 | 4,0 | 4,0 | 4,0 | 4,0 |
| | Choque = 2 | Média | 2,4260 | 2,4380 | 2,5340 | 2,5560 | 2,4790 | 2,4800 | 2,5420 | 2,5580 | 2,4360 | 2,4390 | 2,4580 | 2,4640 |
| | | Desvio | 1,4350 | 1,4420 | 1,4510 | 1,4780 | 1,3658 | 1,3658 | 1,3864 | 1,3894 | 1,2921 | 1,2938 | 1,2853 | 1,2886 |
| | | Mediana | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 |
| | Choque = 3 | Média | 1,6120 | 1,6160 | 1,6580 | 1,6700 | 1,6330 | 1,6350 | 1,6650 | 1,6700 | 1,6170 | 1,6170 | 1,6130 | 1,6160 |
| | | Desvio | 0,7360 | 0,7380 | 0,7410 | 0,7510 | 0,6977 | 0,6973 | 0,7009 | 0,7040 | 0,6729 | 0,6744 | 0,6780 | 0,6804 |
| | | Mediana | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 |
| | Choque = 4 | Média | 1,3640 | 1,3670 | 1,4070 | 1,4140 | 1,4090 | 1,4120 | 1,4030 | 1,4100 | 1,4080 | 1,4090 | 1,3920 | 1,3930 |
| | | Desvio | 0,5310 | 0,5320 | 0,5400 | 0,5450 | 0,5367 | 0,5372 | 0,5262 | 0,5274 | 0,5136 | 0,5157 | 0,5006 | 0,5008 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |

Legenda: **Pb5**: Limites de controle e ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes**: Limites de Controle e ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

7.5 – Conclusão

Podemos observar que, da mesma forma que para os processos normais e não-normais autocorrelacionados, a estatística M de Hayter e Tsui (1994) mostra ser uma boa alternativa na construção das cartas de controle para processos autocorrelacionados, fornecendo valores intermediários do *ARL in control* e *out of control*. Consideramos que o estudo realizado para processos autocorrelacionados é apenas um estudo inicial para construção das cartas de controle, sendo sugerido que futuramente, outros modelos de séries temporais multivariadas sejam estudados.

Capítulo 8 – Exemplo de Aplicação

Apresentamos nesse capítulo um exemplo de aplicação das técnicas abordadas nessa dissertação. Os dados analisados foram gentilmente cedidos pela indústria siderúrgica Siderpa – Siderúrgica Paulino Ltda e fazem parte das informações referentes à produção de ferro gusa do alto forno 2 da empresa. Nesse alto forno foi implantado um sistema computacional que monitora as variáveis e coleta as informações do processo a cada segundo, ou seja, durante um dia são produzidas em torno de $m=86400$ observações amostrais.

O alto forno é a unidade na qual o minério de ferro é transformado em gusa. Um alto forno é de uma torre de material refratário (tijolos ou pedras que não se alteram sob a ação do calor) revestida externamente com chapas de aço. O alto forno pode ter até 40 m de altura e 10 m de diâmetro na largura máxima e pode funcionar ininterruptamente durante anos.

O minério de ferro, o coque e os fundentes (que constituem a carga do alto forno) são introduzidos na parte superior do alto forno (goela) alternadamente, em camadas sucessivas. Na parte inferior é injetado o ar aquecido (a uma temperatura próxima de 1.000°C). O coque entra então em combustão e produz monóxido de carbono (CO_2), que é o principal agente redutor, que reage com o minério e permite a liberação do ferro, separando-o dos elementos com os quais estava combinado.

O ferro fundido misturado ao carbono (ou seja, a gusa) é recolhido no cadinho que se localiza na parte inferior do alto forno, de onde sai através de um conduto denominado “furo de corrida”. Nesse mesmo cadinho é recolhida também a escória, que fica suspensa sobre o metal em estado líquido.

Nesse alto forno específico a cada segundo são coletadas 15 variáveis do processo de produção de ferro gusa nodular especial (teores de fósforo e manganês abaixo de 0,05% e de enxofre abaixo de 0,012%). Exemplos de aplicação de controle multivariado do processo de produção do ferro gusa podem ser encontrados em Thomsen (2005) que mostra inclusive o *software* desenvolvido pelo autor para o monitoramento de um alto forno.

Com base nas informações contidas em Thomsen (2005) das 15 variáveis coletados do processo, 4 delas são mais críticas para análise do processo são elas: temperatura de coroa, pressão do topo, vazão do ar e pressão de leito.

Em uma análise inicial calculamos as estatísticas descritivas e construímos gráficos sequenciais para essas variáveis (Figura 8.1) considerando uma amostra de tamanho 3600 que é equivalente a aproximadamente uma hora de produção contínua do alto forno. Foram consideradas nessa análise dados do dia 13/06/2004 uma vez que esse dia foi considerado pela empresa um dia de produção normal do processo.

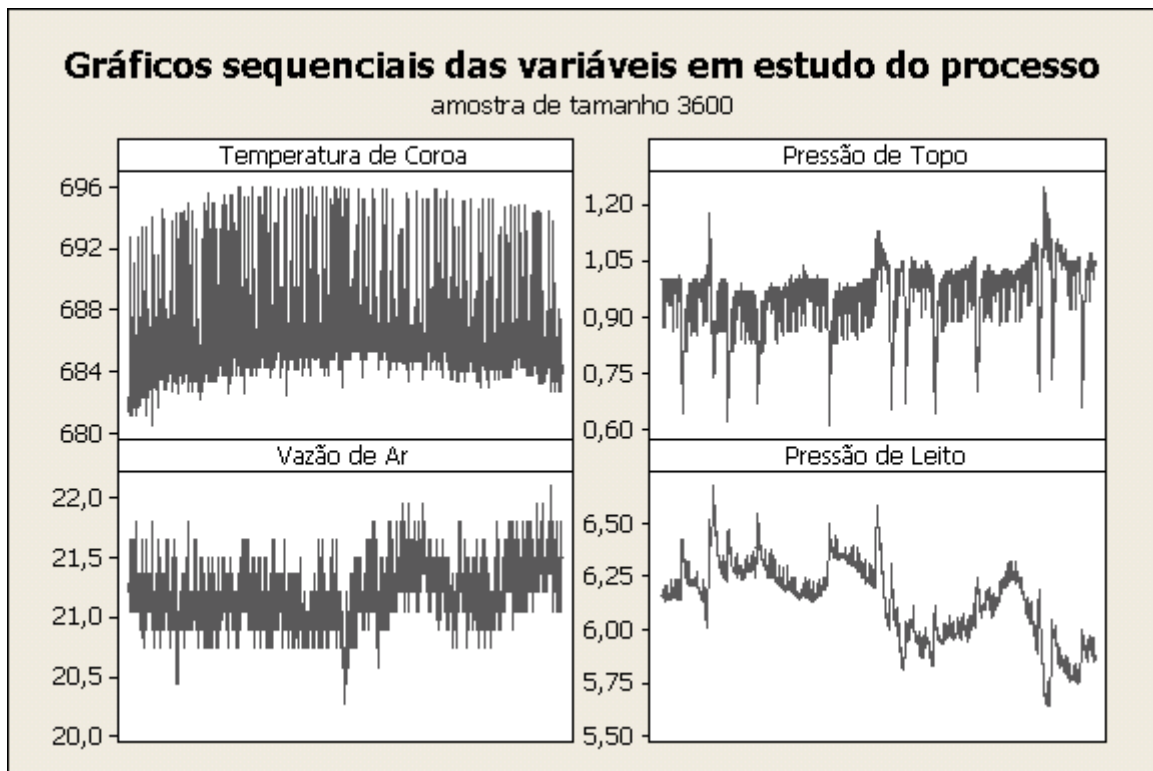


Figura 8.1: Gráfico sequencial das variáveis consideradas do processo de fundição do gusa.

Tabela 8.1: Estatísticas descritivas das variáveis consideradas no processo de fundição gusa.

| Variáveis | m | Média | Desvio-padrão | Mínimo | Q1 | Mediana | Q3 | Máximo |
|----------------------|------|---------|---------------|---------|---------|---------|---------|---------|
| Temperatura de Coroa | 3600 | 685,98 | 2,48 | 680,47 | 684,65 | 685,44 | 686,51 | 696,09 |
| Pressão de Topo | 3600 | 0,97235 | 0,0801 | 0,61000 | 0,95000 | 0,98000 | 1,02000 | 1,25000 |
| Vazão de Ar | 3600 | 21,164 | 0,228 | 20,260 | 21,050 | 21,210 | 21,360 | 22,110 |
| Pressão de Leito | 3600 | 6,1331 | 0,1817 | 5,6400 | 5,9900 | 6,1600 | 6,2600 | 6,6800 |

As matrizes de covariâncias e de correlação amostrais são dadas por:

$$S = \hat{\Sigma} = \begin{bmatrix} 6,1259448 & -0,0097273 & -0,0302624 & 0,0063185 \\ -0,0097273 & 0,0064268 & -0,0005203 & -0,0063459 \\ -0,0302624 & -0,0005203 & 0,0517903 & -0,0219854 \\ 0,0063185 & -0,0063459 & -0,0219854 & 0,0330038 \end{bmatrix}$$

$$R = \hat{\rho} = \begin{bmatrix} 1,000 & -0,049 & -0,054 & 0,014 \\ -0,049 & 1,000 & -0,029 & -0,436 \\ -0,054 & -0,029 & 1,000 & -0,532 \\ 0,014 & -0,436 & -0,532 & 1,000 \end{bmatrix}$$

É possível verificar pela Figura 8.1 que não existe nenhum comportamento atípico, tampouco nenhum valor discrepante dos dados analisados. A maior correlação encontra-se entre as variáveis: pressão de leito e pressão de topo (-0,436) e vazão de ar e pressão de leito

(-0,532). A temperatura de coroa não se relaciona fortemente com nenhuma outra variável. Na Figura 8.2 apresenta-se os gráficos de probabilidade normal para essas variáveis e como pode ser visto, nenhuma delas segue a distribuição normal univariada (valor-p do teste de normalidade de Anderson-Darling $<0,05$) e desse modo, a normalidade multivariada é rejeitada.

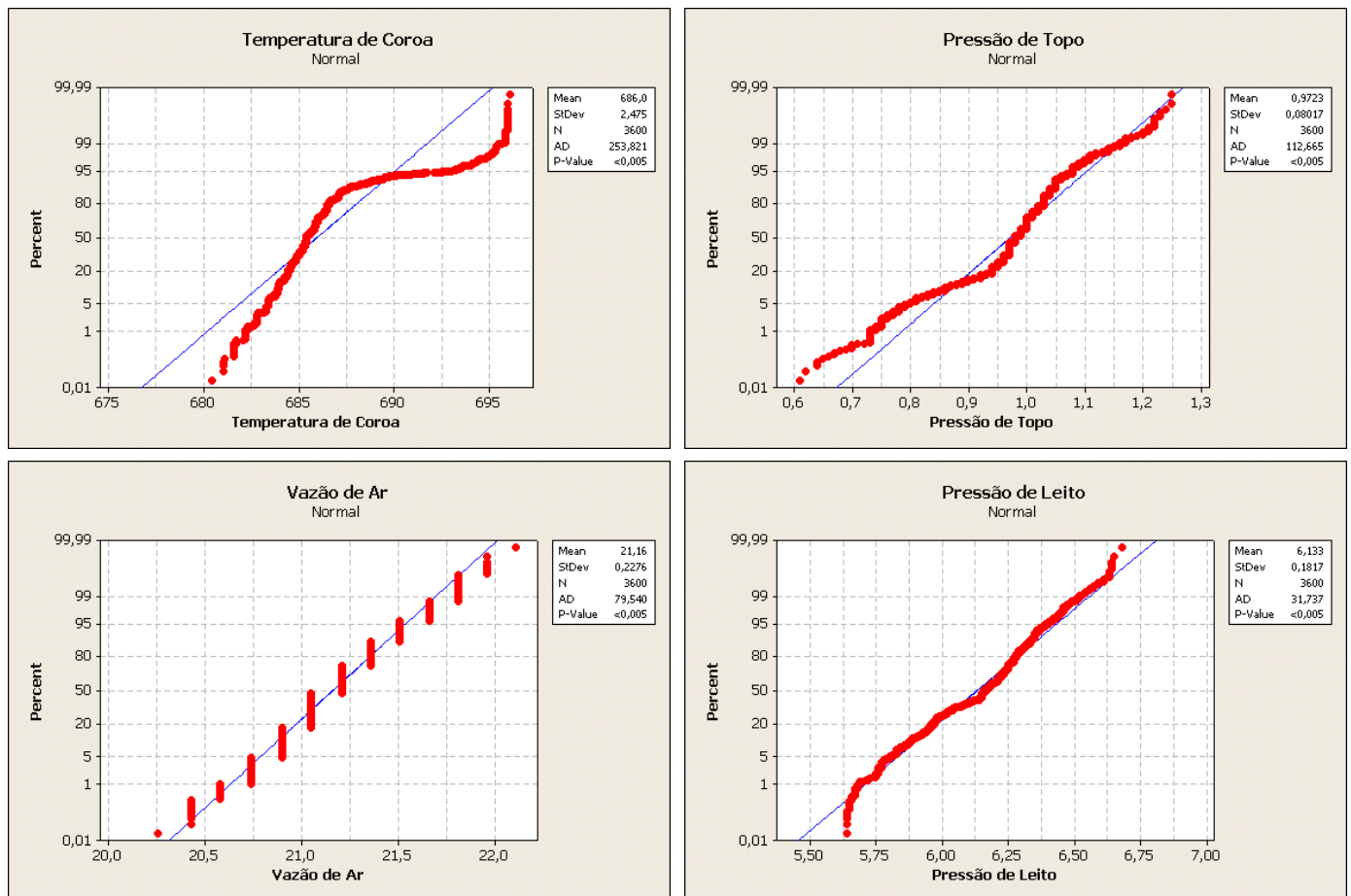


Figura 8.2: Gráfico de normalidade das variáveis estudadas do processo de fundição do gusa.

Para analisar esses dados, foram calculados os valores da estatística M de Hayter e Tsui (1994) e para monitoramento do processo foram construídos limites de controle a 95% de confiança para os valores dessa estatística, para isso, foi utilizada a técnica de núcleo-estimador utilizando os métodos de estimação da janela ótima h de Polansky e Baker (2000) com $b=5$ e de Bessegato (2000). A estatística M foi escolhida por apresentar, nas simulações analisadas nessa dissertação, melhores resultados na construção dos limites de controle. Os resultados dessas análises são apresentados na Figuras 8.3 e 8.4. A estimativa do limite superior de controle é igual a 3,097 para o método de Polansky e Baker (2000) com $b=5$ e igual a 3,116 para o método de Bessegato (2001).

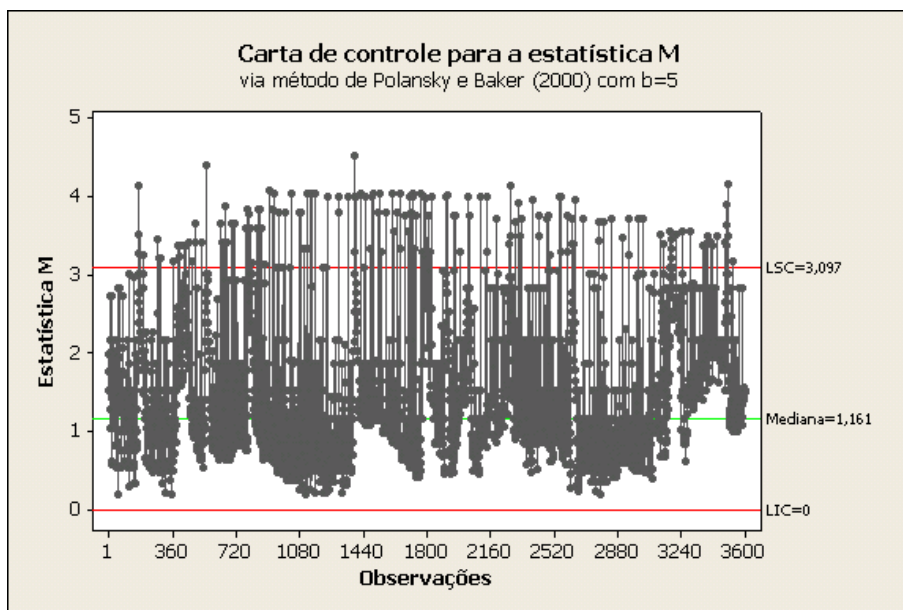


Figura 8.3: Carta de controle para o processo de produção do ferro gusa com limite de controle obtido de acordo o método de Polansky e Baker (2000) com $b=5$ e $m=3600$.

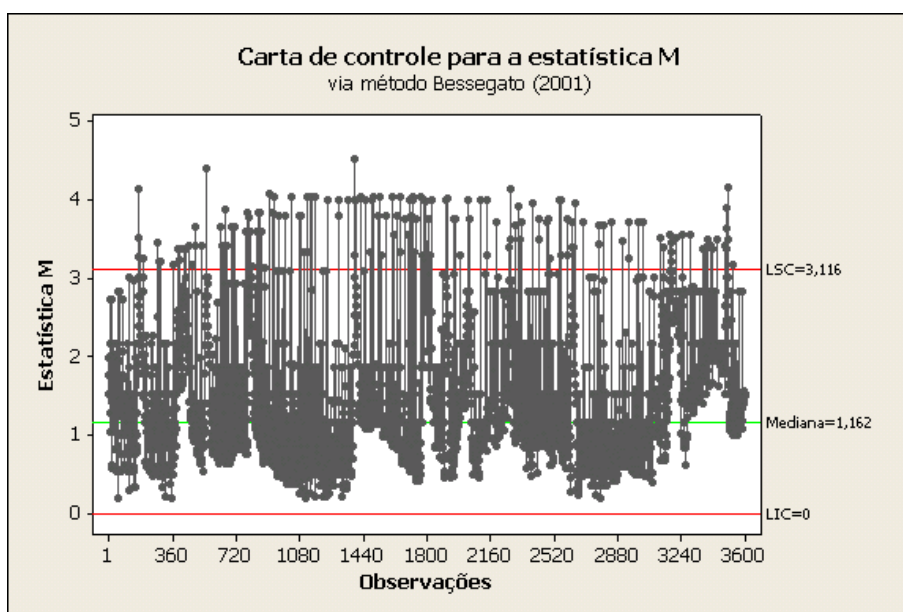


Figura 8.4: Carta de controle para o processo de produção do ferro gusa com limite de controle obtido de acordo o método de Bessegato e $m=3600$.

Pode-se observar nas cartas de controle construídas que existem vários pontos da amostra localizados fora do limite superior de controle para os dois métodos utilizados (Polansky e Baker (2000) com $b=5$ e Bessegato (2001)). Apesar de cada método fornecer limites de controle diferentes, observamos um total de 176 pontos localizados fora dos limites de controle o que equivale a 4,89% do total de pontos da amostra que tamanho igual a 3600.

Quando se aplica as técnicas não-paramétricas de estimação dos limites de controle tratados nessa dissertação sempre haverá uma quantidade de pontos que estará fora dos limites de controle devido à própria forma de construção das cartas. Isso significa que para uma carta

de controle com 95% de confiança, por exemplo, sempre teremos 5% dos pontos localizados fora dos limites de controle. Como citado anteriormente, foi considerado pela empresa que o dia no qual ocorreu a coleta dos dados houve um comportamento normal do processo. Dessa forma, como nenhum valor muito atípico ocorreu nas cartas de controle construídas, conclui-se que os limites de controle que devem ser usados para o monitoramento do processo são iguais a 3,097 para o método de Polansky e Baker (2000) com $b=5$ e 3,116 para o método de Bessegato (2001), utilizando a estatística M de Hayter e Tsui (1994).

Capítulo 9 – Conclusões Finais e Sugestões para Trabalhos Futuros

Durante todas as simulações realizadas, foi possível observar que a estatística M de Hayter e Tsui (1994), foi mais efetiva na detecção da falta de controle do processo multivariado, mostrando-se mais adequada para o monitoramento do processo, pois fornece rápida detecção de alarmes verdadeiros e baixa taxa de alarme falsos.

Foi possível observar também os métodos de obtenção da janela ótima h da técnica de núcleo-estimador forneceram valores bem semelhantes entre si, mostrando que os métodos estudados por Polansky e Baker (2000) e Bessegato (2001) poderiam ser utilizados para o cálculo dos limites das cartas de controle do processo multivariado, independente do número de variáveis e da distribuição conjunta das mesmas.

Foi observado que, em geral, a mudança nas faixas de correlação dos processos, não alteram significativamente as estimativas dos limites de controle do processo. Quando são efetuadas mudanças nas variâncias das variáveis do processo, mantendo-se fixa a correlação entre as variáveis, os limites de controle em geral tendem a ficar maiores, uma vez que o grau de incerteza sobre o conjunto de variáveis aumenta.

Em geral a mudança no vetor de médias mais facilmente captada pelas cartas de controle é aquele em que o “choque” no vetor de médias atua sobre todas as variáveis, fato que ocorreu para todas as técnicas de estimação dos limites de controle.

Observa-se que a técnica de núcleo estimador para amostras com distribuição normal multivariada não-autocorrelacionada, mesmo quando trabalhamos com amostras pequenas ($m=50$) fornece boa estimativa do limite de controle ($C_{R,\alpha}$) quando utiliza-se a estatística M de Hayter e Tsui (1994). Quando consideramos o processo simulado por Hayter e Tsui (1994) com a amostra seguindo a distribuição não-normal multivariada não-autocorrelacionada simulada por esses autores, observa-se a estatística M , com amostra de tamanho $m=200$, e utilizando as técnicas de núcleo-estimador usando o método de Polansky e Baker (2000) com $b=5$ e de Bessegato (2001), fornecem boas estimativas do limite de controle ($C_{R,\alpha}$) quando comparado à forma de obtenção descrita no Quadro 2.2 e usando $N=500$ como suposto por Hayter e Tsui (1994), fato questionado por Mingoti e Glória (2005b) mostrando que para a distribuição normal multivariada é necessário um $N=5000$.

Esses dois fatos são muito importantes pois fazem com que o núcleo estimador seja uma boa alternativa frente ao algoritmo discutido no Quadro 2.1 para a obtenção dos limites de controle ($C_{R,\alpha}$) via estatística M no caso em que o processo segue a distribuição normal multivariada não-autocorrelacionada. A mesma observação pode ser feita para obtenção da constante $C_{R,\alpha}$ quando o processo segue uma distribuição não-normal multivariada não-

autocorrelacionada. Em termos computacionais, quando o processo é normal multivariado, a obtenção da constante $C_{R,\alpha}$ via simulação descrita no Quadro 2.1 é muito rápida, no entanto, para os processos não-normais, há um ganho considerável em se usar o núcleo-estimador pois é possível estimar $C_{R,\alpha}$ com boa precisão sem a necessidade de possuir uma amostra muito grande do processo.

Para os processos normais autocorrelacionados, observa-se que os resultados fornecidos pela técnica de núcleo-estimador são bem próximas àquelas fornecidas para o processo com distribuição normal multivariada não-autocorrelacionada. No caso em que a distribuição é não normal autocorrelacionada observou-se que apenas para amostras maiores que 200 ($m > 200$) a técnica de núcleo-estimador forneceu limites de controle confiáveis (baixa detecção de alarmes falsos e alta detecção de alarmes verdadeiros).

Pela natureza assimétrica da distribuição das estatísticas T^2 de Hotelling (1947) e M de Hayter e Tsui (1994) fica como sugestão para trabalhos futuros a utilização da metodologia de núcleo-estimador mediante à correção do problema de fronteira, mencionado no capítulo 3 e não implementado nessa dissertação, a fim de melhorar ainda mais os resultados fornecidos por essa metodologia na obtenção dos limites de controle do processo.

Além disso, sugere-se a implementação das estatísticas para sumarização de informações multivariadas aplicadas a processos nos quais se trabalha com subgrupos de tamanho maior ou igual a 1, utilizando para estimar a distribuição dessas estatísticas a metodologia de núcleo-estimador, fato que não foi abordado nessa dissertação.

Sugere-se também que os processos autocorrelacionados sejam mais bem estudados nos casos em que se tem autocorrelação de ordem maior que 2 ($VAR(p)$, $p \geq 2$) tanto para processos normais quanto para processos não-normais, já que consideramos o estudo feito para os processos $VAR(1)$, abordados no capítulo 7, apenas preliminar. Sugere-se ainda que processos do tipo $ARMA(p,q)$ sejam abordados, fato que também não foi considerado nessa dissertação.

Referências Bibliográficas

- ANDERSON, T. W. *An introduction to multivariate statistical analysis*. New York, Wiley e Sons, 2nd. ed , 1984.
- ALWAN, L. C., ROBERTS, H.V., The problem of misplaced control limits, *Applied Statistics*, vol. 44, nº 3, 269-278.
- BESSEGATO, L. F., *Escolha do parâmetro de suavidade na estimativa da função de distribuição*. Belo Horizonte, Departamento de Estatística, UFMG, 2001 (Dissertação de mestrado).
- BECHHOFFER, R. E., DUNNET, C. W., Percentage points of multivariate Student t Distributions. *American Mathematical Society*, 1988.
- BONFERRONI, C. *Elementi di statistica generale*, Libreria Seber, Firenze, 1930.
- BOX, G.E.P., COX, D.R. An analysis of transformations, *Journal of Royal Statistical Society*, Serie B (Methodology), vol. 26, nº 2, p. 211-252 (1964).
- BOX, G.E.P., JENKINS, G.M. *Time Series Analysis: Forecasting and Control*, 3 rd ed, New York, Prentice Hall, 1976.
- CASELLA, G., BERGER, R. L. *Statistical inference. 2nd ed*, Thomson Learning, 2002.
- CHEN, H. A multivariate process capability index over a rectangular solid tolerance Zone. *Statistica Sinica*, vol. 4, p. 749-758, 1994.
- CHOU, Y-M., MASON, R.L., YOUNG, J. C., Power comparisons for Hotelling T^2 Statistic. *Communications in Statistical Simulations* 28, nº 4, p.1031-1050, 1999.
- CHOU, Y-M., MASON, R. L., YOUNG, J. C. The control chart for individual observations from a multivariate non-normal distribution. *Communications in Statistical Simulations* vol. 30, nº 8 e 9, p.1937-1949, 2001.

- COLENGUI, F.K.R, MINGOTI, S. A. *Implementação de testes estatísticos multivariados em S-plus*, Belo Horizonte, Departamento de Estatística, 2005 (Monografia).
- COLENGUI, F.K.R, MINGOTI, S. A. Implementação computacional em S-Plus dos testes estatísticos de Hotelling e Hayter e Tsui e suas aplicações em controle de qualidade. *XII SIMPEP*, Bauru – SP. Anais do evento, 2005 (em cd-rom).
- D'AGOSTINHO, R.B., STEPHENS, M.A. *Goodness-of-Fit Techniques*, Marcel Deker, 1986.
- HAYTER, A. J., TSUI, K-L. Identification and quantification in multivariate quality control problems. *Journal of Quality Technology* 26, nº 3, p.197-208, 1994.
- HOTELLING, H., Multivariate quality control. *Techniques of statistical analysis*. New York, Mc Graw Hill, p.111-184, 1947.
- JACKSON, J.E., Quality control methods for several related variables. *Technometrics*, vol.1, nº 4, 1959.
- JOHNSON, M. E., WANG, C., RAMBERG, J. S. Generation of continuous multivariate distributions for statistical applications. *American Journal of Mathematical and Management Sciences*, 1984.
- JONHSON, R.A., WICHERN, D.W., *Applied Multivariate Statistical analysis*. Prentice Hall, 6th ed, 2002.
- KALGONDA, A. A., KULKARNI, S. R., Multivariate quality control chart for autocorrelated process. *Journal of Applied Statistics* 31, nº 3, p. 317-327, 2004.
- MAHALANOBIS, P. C. On the generalized distance in statistics. *Proceedings of the National Institute of Sciences of India*, vol. 12, p. 49-55, 1936.
- MASON, R., YOUNG, J. C. Improving the sensitivity of the T^2 statistic in multivariate process control. *Journal of Quality Technology* 31, nº 2, p. 155-165, 1999.

- MASON, R., YOUNG, J.C., *Multivariate statistical process control with industrial applications*. Alexandria, Society for Industrial and Applied Mathematics, 2002.
- MINGOTI, S.A., *Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada*. Belo Horizonte, Editora da UFMG, 2005.
- MINGOTI, S.A., GLÓRIA, F.A.A., Uma proposta de modificação do Índice de Capacidade Multivariado de Chen. *XXIII ENEGEP*, Ouro Preto - MG. Anais do Evento, 2003 (em cd-rom).
- MINGOTI, S. A., GLÓRIA, F. A. A. A modification of Chen's multivariate capability index and a comparison with Niverthi and Dey capability indices. (Artigo submetido a revista internacional), 2005.
- MINGOTI, S.A., GLÓRIA, F.A.A., Comparando os métodos paramétrico e não-paramétrico na determinação do valor crítico do teste estatístico de médias proposto por Hayter e Tsui. *Revista Produto e Produção*, v15, nº.2, 251-262, Maio-Agosto, 2005.
- MONTGOMERY, D.C. *Introduction to Statistical Quality Control*. New York, John Wiley & Sons, 5 th ed, 2004.
- NEUMAINER, A., SCHNEIDER, T., Estimation of parameters and eigenmodes of multivariate autoregressive models. *ACM Transactions on Mathematical Software*, vol 27, nº 1, pp 27-57, 2001.
- POLANSKY, A.M., An Algorithm for computing a smooth non-parametric process capability estimate. *Journal of Quality Technology* 32, nº 3, p. 284-289, 2000.
- POLANSKY, A. M., BAKER, E. R., Multistage plug-in bandwidth selection for Kernel distribution function estimates. *Journal of Statistical Computation and Simulation* 65, p. 63-80, 2000.
- REINSEL, G. C., *Elements of multivariate time series analysis*, Madison, Springer, 1997.

- ROCON, G. M., *Comparação de cartas de controle multivariadas usadas no monitoramento de processos univariados autocorrelacionados*, Belo Horizonte, Departamento de Estatística, 2005 (Dissertação de Mestrado).
- SERFLING, R. J. *Approximation theorems of mathematical statistics*, New York, John Wiley and Sons, 1980.
- SHEWHART, W. A. Quality control charts. *Bell System Technical Journal*, 1926.
- SILVERMAN, B.W., *Density Estimation for Statistics and Data Analysis*. Londres, Chapman & Hall, 1992.
- SOUZA, A. M., RIGÃO, M.A., Identificação de variáveis fora de controle em processos produtivos multivariados, *Revista produto e produção*, vol 14, nº 1, p.74-86, Janeiro-Abril, 2005.
- TACAKS, L., On the comparison of a theoretical and an empirical distribution function, *Journal of Applied Probability*, vol. 8, nº. 2, p. 321-330, 1971.
- THOMSEN, E. P. R. *Controle de processo de produção do alto forno por meio de técnicas estatísticas multivariadas em tempo real*. Belo Horizonte, Departamento de Estatística, 2005 (Dissertação de Mestrado).
- TRAVASSOS, A. P. A., *Problemas de fronteiras dos núcleo-estimadores e suas abordagens*. Belo Horizonte, Departamento de Estatística, 2003 (Dissertação de mestrado).
- WAND, M. P., JONES, M. C. *Kernel Smoothing*. Londres, Chapman & Hall, 1995.
- WEI, W.W.S, *Times Series Analysis*, New York, Addison-Wesley Publishing Company, 1990.

Anexo A – Tabelas do *ARL in control* e *ARL out of control*. Processo com distribuição normal multivariada não-autocorrelacionada com $p=5$ variáveis

Tabela A.1: *ARL in control* e *out of control* de processos com distribuição normal multivariada com forte correlação entre as variáveis ($0,75 < \rho \leq 1$) e $m=50$.

| Distribuição multivariada com $p=5$, $0,75 < \rho \leq 1$ e $m=50$ | | Medidas Descritivas | Estatística T ² de Hotelling | | | Estatística M | | | Componentes Principais | | | |
|---|--|------------------------|---|---------|---------|---------------|---------|---------|------------------------|---------|---------|---------|
| | | | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | |
| ARL in control | | Média | 21,8660 | 9,1390 | 9,2570 | 18,0220 | 18,7880 | 19,2840 | 20,5790 | 24,0240 | 25,7170 | |
| | | Desvio | 24,4211 | 10,8814 | 10,9151 | 20,2389 | 23,0685 | 23,5679 | 21,4584 | 29,8830 | 31,8566 | |
| | | Mediana | 14,0 | 6,0 | 6,0 | 11,0 | 11,0 | 11,0 | 14,0 | 15,0 | 15,0 | |
| Todas as médias mudam simultaneamente na mesma escala | Choque = 0,5 | Média | 19,0720 | 8,1050 | 8,1640 | 12,0200 | 12,2600 | 12,5690 | 12,2920 | 17,0040 | 17,7600 | |
| | | Desvio | 22,8030 | 9,1935 | 9,2567 | 13,1844 | 14,0294 | 14,4405 | 11,5573 | 24,9707 | 26,3844 | |
| | | Mediana | 12,0 | 5,0 | 5,0 | 8,0 | 8,0 | 8,0 | 9,0 | 10,0 | 10,0 | |
| | Choque = 1 | Média | 10,8940 | 5,2380 | 5,3250 | 5,5630 | 5,6350 | 5,7300 | 5,5620 | 7,2620 | 7,4310 | |
| | | Desvio | 11,7606 | 5,5750 | 5,7012 | 5,5356 | 6,1594 | 6,1887 | 5,1108 | 9,5281 | 9,6021 | |
| | | Mediana | 7,0 | 3,0 | 3,0 | 4,0 | 4,0 | 4,0 | 4,0 | 4,0 | 4,0 | |
| | Choque = 2 | Média | 3,6840 | 2,2490 | 2,2600 | 1,8000 | 1,7970 | 1,8070 | 1,7360 | 1,9630 | 1,9970 | |
| | | Desvio | 3,5798 | 1,8348 | 1,8439 | 1,2221 | 1,1913 | 1,1930 | 1,0533 | 1,4511 | 1,4605 | |
| | | Mediana | 3,0 | 2,0 | 2,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| | Choque = 3 | Média | 1,6130 | 1,2910 | 1,2960 | 1,1440 | 1,1430 | 1,1460 | 1,1470 | 1,1750 | 1,1800 | |
| | | Desvio | 1,0641 | 0,6362 | 0,6394 | 0,4018 | 0,3984 | 0,4035 | 0,4093 | 0,4611 | 0,4646 | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| | Choque = 4 | Média | 1,1190 | 1,0400 | 1,0430 | 1,0120 | 1,0130 | 1,0130 | 1,0090 | 1,0120 | 1,0120 | |
| | | Desvio | 0,3886 | 0,2108 | 0,2173 | 0,1089 | 0,1133 | 0,1133 | 0,0945 | 0,1089 | 0,1089 | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| | Uma média permanece constante e a outra muda ordenadamente | Choque = 0,5 | Média | 10,9200 | 5,2000 | 5,2660 | 14,3120 | 14,5850 | 15,0820 | 19,4090 | 26,3080 | 28,0590 |
| | | | Desvio | 11,8480 | 4,9883 | 5,0793 | 16,8861 | 17,7305 | 18,2991 | 18,8222 | 33,0894 | 37,6266 |
| | | | Mediana | 7,0 | 4,0 | 4,0 | 9,0 | 9,0 | 9,0 | 13,0 | 15,0 | 15,0 |
| Choque = 1 | | Média | 3,0560 | 1,9800 | 1,9940 | 8,3770 | 8,7280 | 8,9220 | 18,3210 | 24,7840 | 26,0320 | |
| | | Desvio | 2,6400 | 1,4886 | 1,5087 | 9,7631 | 10,2868 | 10,4916 | 17,4249 | 31,8057 | 33,9886 | |
| | | Mediana | 2,0 | 1,0 | 1,0 | 5,0 | 5,0 | 5,0 | 13,0 | 14,0 | 15,0 | |
| Choque = 2 | | Média | 1,0560 | 1,0200 | 1,0200 | 2,4590 | 2,5520 | 2,5750 | 14,2380 | 18,4510 | 19,5760 | |
| | | Desvio | 0,2468 | 0,1401 | 0,1401 | 2,1953 | 2,3687 | 2,3770 | 15,3803 | 26,6454 | 29,0199 | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 | 9,0 | 10,0 | 11,0 | |
| Choque = 3 | | Média | 1,0010 | 1,0010 | 1,0010 | 1,2750 | 1,2790 | 1,2900 | 10,5490 | 14,7990 | 15,5320 | |
| | | Desvio | 0,0316 | 0,0316 | 0,0316 | 0,6097 | 0,6062 | 0,6311 | 10,7490 | 21,0954 | 22,4478 | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 7,0 | 8,0 | 8,0 | |
| Choque = 4 | | Média | 1,0000 | 1,0000 | 1,0000 | 1,0400 | 1,0430 | 1,0440 | 7,5050 | 10,0760 | 10,6390 | |
| | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,2201 | 0,2263 | 0,2283 | 7,0684 | 13,1168 | 14,9977 | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 5,0 | 6,0 | 6,0 | |
| Todas as médias mudam | | Choque=0,5, 3, 1, 1, 1 | Média | 1,0000 | 1,0000 | 1,0000 | 1,3150 | 1,3480 | 1,3460 | 3,4890 | 4,1930 | 4,3820 |
| | | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,7002 | 0,7625 | 0,7462 | 2,8625 | 4,7777 | 5,1018 |
| | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 3,0 | 3,0 | 3,0 |

Legenda: **Teo:** Limites de controle e ARL para o Limite de Controle obtido de forma teórica; **Pb5:** Limites de controle e ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes:** Limites de Controle e ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

Tabela A.2: *ARL in control e out of control* de processos com distribuição normal multivariada com forte correlação entre as variáveis ($0,75 < \rho \leq 1$) e $m=200$.

| Distribuição multivariada com $p=5$, $0,75 < \rho \leq 1$ e $m=200$ | | Medidas Descritivas | Estatística T ² de Hotelling | | | Estatística M | | | Componentes Principais | | | | |
|--|--|--------------------------|---|---------|---------|---------------|---------|---------|------------------------|---------|---------|---------|---------|
| | | | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | | |
| ARL in control | | Média | 18,9590 | 14,4850 | 14,7260 | 18,5860 | 18,7620 | 18,9900 | 20,7240 | 22,5760 | 23,3730 | | |
| | | Desvio | 19,6913 | 15,1765 | 15,4445 | 18,8841 | 19,8995 | 20,0893 | 21,3420 | 23,7478 | 24,4776 | | |
| | | Mediana | 12,0 | 10,0 | 10,0 | 13,0 | 12,0 | 13,0 | 14,0 | 15,0 | 15,0 | | |
| Todas as médias mudam simultaneamente na mesma escala | | Choque = 0,5 | | Média | 14,2940 | 10,7780 | 10,9090 | 10,4240 | 10,6290 | 10,8500 | 11,0190 | 12,4910 | 12,8340 |
| | | | | Desvio | 14,4994 | 10,8313 | 10,6318 | 10,0975 | 10,7745 | 11,0458 | 11,1297 | 13,7168 | 14,0317 |
| | | | | Mediana | 10,0 | 7,0 | 7,0 | 7,0 | 7,0 | 7,0 | 8,0 | 8,0 | 9,0 |
| | | Choque = 1 | | Média | 8,8610 | 7,2230 | 7,2210 | 5,3210 | 5,5380 | 5,6270 | 5,4460 | 6,1870 | 6,2650 |
| | | | | Desvio | 8,7595 | 7,0644 | 7,0470 | 4,9742 | 5,3630 | 5,5300 | 5,2683 | 6,4673 | 6,5042 |
| | | | | Mediana | 6,0 | 5,0 | 5,0 | 4,0 | 4,0 | 4,0 | 4,0 | 4,0 | 4,0 |
| | | Choque = 2 | | Média | 2,9560 | 2,5850 | 2,6070 | 1,7920 | 1,8020 | 1,8130 | 1,7300 | 1,7960 | 1,8170 |
| | | | | Desvio | 2,5645 | 2,1301 | 2,1557 | 1,1367 | 1,1341 | 1,1547 | 1,0818 | 1,1540 | 1,1802 |
| | | | | Mediana | 2,0 | 2,0 | 2,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| | | Choque = 3 | | Média | 1,4560 | 1,3730 | 1,3740 | 1,1250 | 1,1180 | 1,1210 | 1,1120 | 1,1280 | 1,1300 |
| | | | | Desvio | 0,8239 | 0,7406 | 0,7421 | 0,3841 | 0,3608 | 0,3640 | 0,3572 | 0,3921 | 0,3940 |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| Choque = 4 | | Média | 1,0650 | 1,0530 | 1,0530 | 1,0150 | 1,0170 | 1,0180 | 1,0110 | 1,0140 | 1,0140 | | |
| | | Desvio | 0,2662 | 0,2372 | 0,2372 | 0,1216 | 0,1293 | 0,1330 | 0,1044 | 0,1175 | 0,1175 | | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | | |
| Uma média permanece constante e a outra muda ordenadamente | | Choque = 0,5 | | Média | 8,4110 | 6,8660 | 6,9180 | 14,3320 | 14,1950 | 14,5880 | 19,3410 | 21,2840 | 21,9950 |
| | | | | Desvio | 7,8373 | 6,3940 | 6,4486 | 14,6071 | 14,1705 | 14,5315 | 18,7357 | 21,4801 | 22,1465 |
| | | | | Mediana | 6,0 | 5,0 | 5,0 | 9,0 | 9,0 | 9,0 | 13,0 | 15,0 | 15,0 |
| | | Choque = 1 | | Média | 2,5510 | 2,2430 | 2,2530 | 8,2590 | 8,2440 | 8,3220 | 18,6650 | 20,8310 | 21,4560 |
| | | | | Desvio | 2,0649 | 1,7294 | 1,7338 | 7,7302 | 7,7292 | 7,7802 | 18,0204 | 22,1498 | 22,6484 |
| | | | | Mediana | 2,0 | 2,0 | 2,0 | 6,0 | 6,0 | 6,0 | 14,0 | 14,0 | 15,0 |
| | | Choque = 2 | | Média | 1,0310 | 1,0210 | 1,0210 | 2,3700 | 2,4270 | 2,4420 | 15,1400 | 17,0820 | 17,9140 |
| | | | | Desvio | 0,1734 | 0,1435 | 0,1435 | 1,7630 | 1,8314 | 1,8401 | 14,9705 | 18,0545 | 19,2506 |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 | 11,0 | 11,0 | 11,0 |
| | | Choque = 3 | | Média | 1,0000 | 1,0000 | 1,0000 | 1,3000 | 1,3020 | 1,3050 | 11,0340 | 11,9590 | 12,2130 |
| | | | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,6515 | 0,6612 | 0,6621 | 10,5294 | 12,1000 | 12,1939 |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 8,0 | 9,0 | 9,0 |
| Choque = 4 | | Média | 1,0000 | 1,0000 | 1,0000 | 1,0300 | 1,0330 | 1,0340 | 7,7940 | 8,6240 | 8,7660 | | |
| | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,1707 | 0,1787 | 0,1813 | 7,0375 | 8,4439 | 8,5771 | | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 6,0 | 6,0 | 6,0 | | |
| Todas as médias mudam | | Choque = 0,5, 3, 1, 1, 1 | | Média | 1,0000 | 1,0000 | 1,0000 | 1,2930 | 1,2960 | 1,3050 | 3,3440 | 3,6890 | 3,7350 |
| | | | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,5861 | 0,6283 | 0,6359 | 2,6097 | 3,1503 | 3,1709 |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 3,0 | 3,0 | 3,0 |

Legenda: **Teo:** Limites de controle e ARL para o Limite de Controle obtido de forma teórica; **Pb5:** Limites de controle e ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes:** Limites de Controle e ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

Tabela A.3: *ARL in control e out of control* de processos com distribuição normal multivariada com correlação intermediária entre as variáveis ($0,50 < \rho \leq 0,75$) e $m=50$.

| Distribuição multivariada com $p=5$, $0,50 < \rho \leq 0,75$ e $m=50$ | | Medidas Descritivas | Estatística T ² de Hotelling | | | Estatística M | | | Componentes Principais | | |
|--|--|--------------------------|---|----------|----------|---------------|----------|----------|------------------------|----------|----------|
| | | | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes |
| ARL in control | | Média | 21,91700 | 8,73000 | 9,66000 | 17,25700 | 16,76800 | 20,02700 | 20,51200 | 26,17100 | 32,63600 |
| | | Desvio | 25,05310 | 10,11940 | 11,54000 | 22,45390 | 23,04780 | 26,36880 | 20,30460 | 34,62800 | 46,40130 |
| | | Mediana | 14,0 | 6,0 | 6,0 | 11,0 | 10,0 | 11,0 | 14,0 | 15,0 | 18,0 |
| Todas as médias mudam simultaneamente na mesma escala | Choque = 0,5 | Média | 16,716 | 7,311 | 7,974 | 10,575 | 10,663 | 11,926 | 10,905 | 14,644 | 19,202 |
| | | Desvio | 19,2368 | 7,6622 | 8,4527 | 12,0704 | 12,4340 | 14,1213 | 10,1839 | 18,2177 | 27,0967 |
| | | Mediana | 10,0 | 5,0 | 5,0 | 7,0 | 6,0 | 7,0 | 8,0 | 9,0 | 10,0 |
| | Choque = 1 | Média | 9,437 | 4,442 | 4,780 | 4,734 | 4,720 | 5,160 | 4,395 | 5,841 | 6,738 |
| | | Desvio | 10,0994 | 4,2857 | 4,6563 | 4,9195 | 5,0372 | 5,4085 | 3,9454 | 6,3756 | 8,1125 |
| | | Mediana | 6,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 4,0 | 4,0 |
| | Choque = 2 | Média | 2,931 | 1,944 | 2,019 | 1,612 | 1,622 | 1,698 | 1,527 | 1,665 | 1,800 |
| | | Desvio | 2,6910 | 1,5242 | 1,6097 | 1,1219 | 1,1707 | 1,2940 | 0,9634 | 1,2475 | 1,4588 |
| | | Mediana | 2,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| | Choque = 3 | Média | 1,396 | 1,171 | 1,189 | 1,081 | 1,086 | 1,090 | 1,065 | 1,078 | 1,090 |
| | | Desvio | 0,7936 | 0,4690 | 0,4955 | 0,3042 | 0,3142 | 0,3194 | 0,2662 | 0,3033 | 0,3194 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| | Choque = 4 | Média | 1,054 | 1,022 | 1,023 | 1,002 | 1,004 | 1,007 | 1,001 | 1,005 | 1,006 |
| | | Desvio | 0,2513 | 0,1534 | 0,1565 | 0,0447 | 0,0632 | 0,0834 | 0,0316 | 0,0706 | 0,0773 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| | Uma média permanece constante e a outra muda ordenadamente | Choque = 0,5 | Média | 12,811 | 5,869 | 6,456 | 13,955 | 14,099 | 15,900 | 19,319 | 24,028 |
| Desvio | | | 14,8158 | 6,1071 | 6,7078 | 16,1647 | 16,9658 | 19,0936 | 18,7904 | 31,2473 | 42,9957 |
| Mediana | | | 8,0 | 4,0 | 4,0 | 9,0 | 9,0 | 10,0 | 14,0 | 14,0 | 17,0 |
| Choque = 1 | | Média | 5,037 | 2,860 | 2,981 | 9,300 | 9,485 | 10,780 | 18,034 | 24,571 | 30,939 |
| | | Desvio | 5,3387 | 2,5434 | 2,6525 | 9,9289 | 11,0030 | 12,6951 | 17,4877 | 31,1043 | 40,5512 |
| | | Mediana | 3,0 | 2,0 | 2,0 | 6,0 | 6,0 | 7,0 | 13,0 | 14,0 | 16,0 |
| Choque = 2 | | Média | 1,324 | 1,149 | 1,174 | 2,920 | 2,964 | 3,152 | 13,422 | 18,073 | 21,820 |
| | | Desvio | 0,6764 | 0,4231 | 0,4816 | 2,8969 | 3,1424 | 3,3003 | 12,4858 | 21,1652 | 29,3291 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 | 9,0 | 10,0 | 12,0 |
| Choque = 3 | | Média | 1,015 | 1,002 | 1,003 | 1,373 | 1,378 | 1,452 | 8,654 | 12,528 | 15,365 |
| | | Desvio | 0,1216 | 0,0447 | 0,0547 | 0,7606 | 0,7627 | 0,9893 | 8,1109 | 16,8171 | 22,7235 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 6,0 | 7,0 | 8,0 |
| Choque = 4 | | Média | 1,000 | 1,000 | 1,000 | 1,063 | 1,061 | 1,081 | 6,109 | 7,846 | 9,315 |
| | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,2741 | 0,2745 | 0,3202 | 5,6892 | 10,1042 | 11,9436 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 4,0 | 5,0 | 5,0 |
| Todas as médias mudam | | Choque = 0,5, 3, 1, 1, 1 | Média | 1,392 | 1,181 | 1,200 | 1,403 | 1,392 | 1,440 | 3,086 | 3,704 |
| | Desvio | | 0,7422 | 0,4842 | 0,5102 | 0,7856 | 0,7700 | 0,8574 | 2,5233 | 3,5498 | 4,5719 |
| | Mediana | | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | 2,5 | 3,0 |

Legenda: **Teo:** Limites de controle e ARL para o Limite de Controle obtido de forma teórica; **Pb5:** Limites de controle e ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes:** Limites de Controle e ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

Tabela A.4: *ARL in control e out of control* de processos com distribuição normal multivariada com correlação intermediária entre as variáveis ($0,50 < \rho \leq 0,75$) e $m=200$.

| Distribuição multivariada com $p=5$, $0,50 < \rho \leq 0,75$ e $m=50$ | | Medidas Descritivas | Estatística T ² de Hotelling | | | Estatística M | | | Componentes Principais | | | | |
|--|--|--------------------------|---|---------|---------|---------------|---------|---------|------------------------|---------|---------|---------|---------|
| | | | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | | |
| ARL in control | | Média | 19,7200 | 15,1080 | 15,3270 | 17,9250 | 18,4520 | 18,7870 | 19,4440 | 21,2440 | 22,1770 | | |
| | | Desvio | 18,7045 | 15,1870 | 15,4120 | 18,2764 | 20,1756 | 20,5888 | 18,2687 | 22,1056 | 22,8815 | | |
| | | Mediana | 14,0 | 10,0 | 11,0 | 12,0 | 12,0 | 12,0 | 14,0 | 14,0 | 15,0 | | |
| Todas as médias mudam simultaneamente na mesma escala | | Choque = 0,5 | | Média | 14,6070 | 11,5900 | 11,6500 | 11,1000 | 11,2450 | 11,3920 | 11,1740 | 11,6970 | 11,8930 |
| | | | | Desvio | 14,8770 | 11,6898 | 11,7191 | 10,7168 | 11,4232 | 11,6189 | 10,3855 | 11,4008 | 11,5330 |
| | | | | Mediana | 9,0 | 8,0 | 8,0 | 8,0 | 8,0 | 8,0 | 8,0 | 8,0 | 8,0 |
| | | Choque = 1 | | Média | 8,1570 | 6,5320 | 6,6060 | 4,8960 | 4,8570 | 4,9000 | 4,7050 | 5,0410 | 5,1210 |
| | | | | Desvio | 7,8556 | 6,4220 | 6,4840 | 4,8422 | 4,7526 | 4,7616 | 4,1436 | 4,8362 | 4,8516 |
| | | | | Mediana | 6,0 | 4,0 | 4,0 | 3,0 | 3,0 | 3,0 | 3,0 | 4,0 | 4,0 |
| | | Choque = 2 | | Média | 2,4480 | 2,1690 | 2,1820 | 1,6450 | 1,6560 | 1,6650 | 1,5340 | 1,5720 | 1,5820 |
| | | | | Desvio | 1,9362 | 1,6103 | 1,6197 | 1,0631 | 1,0977 | 1,1045 | 0,8830 | 0,9347 | 0,9519 |
| | | | | Mediana | 2,0 | 2,0 | 2,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| | | Choque = 3 | | Média | 1,2790 | 1,2270 | 1,2360 | 1,0830 | 1,0840 | 1,0880 | 1,0610 | 1,0740 | 1,0750 |
| | | | | Desvio | 0,6337 | 0,5690 | 0,5767 | 0,2970 | 0,3017 | 0,3072 | 0,2477 | 0,2768 | 0,2783 |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| Choque = 4 | | Média | 1,0290 | 1,0260 | 1,0260 | 1,0080 | 1,0080 | 1,0080 | 1,0070 | 1,0070 | 1,0070 | | |
| | | Desvio | 0,1679 | 0,1592 | 0,1592 | 0,0891 | 0,0891 | 0,0891 | 0,0834 | 0,0834 | 0,0834 | | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | | |
| Uma média permanece constante e a outra muda ordenadamente | | Choque = 0,5 | | Média | 10,5790 | 8,5800 | 8,6460 | 14,5970 | 14,7340 | 14,9160 | 18,9450 | 21,6410 | 22,6220 |
| | | | | Desvio | 10,5097 | 8,9347 | 8,8988 | 15,3785 | 15,7976 | 15,8832 | 19,8329 | 24,5197 | 25,6573 |
| | | | | Mediana | 7,0 | 6,0 | 6,0 | 10,0 | 10,0 | 10,0 | 13,0 | 14,0 | 15,0 |
| | | Choque = 1 | | Média | 4,0360 | 3,3560 | 3,4210 | 8,8580 | 8,7990 | 8,8550 | 17,7270 | 20,1610 | 20,6930 |
| | | | | Desvio | 3,6038 | 2,8519 | 2,9107 | 9,0805 | 9,1340 | 9,1879 | 18,3542 | 22,3412 | 22,9520 |
| | | | | Mediana | 3,0 | 3,0 | 3,0 | 6,0 | 6,0 | 6,0 | 12,0 | 13,0 | 13,0 |
| | | Choque = 2 | | Média | 1,2300 | 1,1850 | 1,1860 | 2,8290 | 2,8270 | 2,8480 | 13,2930 | 14,6720 | 15,0820 |
| | | | | Desvio | 0,5229 | 0,4658 | 0,4665 | 2,3063 | 2,2186 | 2,2534 | 12,6991 | 14,8168 | 15,1569 |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 | 9,0 | 10,0 | 10,0 |
| | | Choque = 3 | | Média | 1,0060 | 1,0030 | 1,0030 | 1,3460 | 1,3450 | 1,3530 | 8,3390 | 9,5330 | 9,7970 |
| | | | | Desvio | 0,0773 | 0,0547 | 0,0547 | 0,6788 | 0,6873 | 0,6992 | 7,9679 | 10,3635 | 10,6182 |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 6,0 | 6,0 | 6,0 |
| Choque = 4 | | Média | 1,0000 | 1,0000 | 1,0000 | 1,0520 | 1,0520 | 1,0520 | 5,9990 | 6,5490 | 6,6750 | | |
| | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,2310 | 0,2310 | 0,2310 | 5,5411 | 6,5453 | 6,6154 | | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 4,0 | 5,0 | 5,0 | | |
| Todas as médias mudam | | Choque = 0,5, 3, 1, 1, 1 | | Média | 1,2840 | 1,2280 | 1,2280 | 1,3750 | 1,3810 | 1,3830 | 2,9850 | 3,1440 | 3,2250 |
| | | | | Desvio | 0,6047 | 0,5141 | 0,5141 | 0,7049 | 0,7116 | 0,7119 | 2,3619 | 2,7758 | 2,8908 |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 |

Legenda: **Teo:** Limites de controle e ARL para o Limite de Controle obtido de forma teórica; **Pb5:** Limites de controle e ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes:** Limites de Controle e ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

Tabela A.5: *ARL in control* e *out of control* de processos com distribuição normal multivariada com correlação nula entre as variáveis ($\rho = 0$) e $m=50$ e $m=200$.

| Distribuição normal multivariada com $p=5$, $\rho=0$ $m=50$ e $m=200$ | | Medidas Descritivas | m=50 | | | | | | m=200 | | | | | | |
|--|--|--------------------------|---|---------|---------|---------------|---------|---------|---|---------|---------|---------------|---------|---------|---------|
| | | | Estatística T ² de Hotelling | | | Estatística M | | | Estatística T ² de Hotelling | | | Estatística M | | | |
| | | | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | |
| ARL in control | | Média | 23,7540 | 8,9520 | 9,8330 | 14,3150 | 14,2380 | 16,6890 | 18,9310 | 14,3010 | 14,4910 | 17,0280 | 17,4390 | 17,5790 | |
| | | Desvio | 28,4530 | 9,9140 | 11,5760 | 14,5480 | 16,7270 | 22,0060 | 19,5330 | 14,4370 | 14,5860 | 16,3830 | 18,2750 | 18,3570 | |
| | | Mediana | 14,0 | 6,0 | 6,0 | 10,0 | 9,0 | 10,0 | 13,0 | 10,0 | 10,0 | 12,0 | 12,0 | 12,0 | |
| Todas as médias mudam simultaneamente na mesma escala | Choque = 0,5 | Média | 10,5180 | 5,1890 | 5,5080 | 8,3820 | 8,4590 | 9,3330 | 7,8850 | 6,5160 | 6,5620 | 9,0550 | 8,8530 | 8,9430 | |
| | | Desvio | 11,2730 | 5,4950 | 5,7900 | 8,8210 | 9,7360 | 10,1800 | 7,7630 | 6,6170 | 6,6310 | 8,8830 | 9,2430 | 9,2700 | |
| | | Mediana | 7,0 | 4,0 | 4,0 | 5,0 | 5,0 | 6,0 | 5,0 | 4,0 | 5,0 | 6,0 | 6,0 | 6,0 | |
| | Choque = 1 | Média | 3,2590 | 2,1030 | 2,2110 | 3,4320 | 3,2870 | 3,6340 | 2,9480 | 2,5590 | 2,5650 | 3,7570 | 3,5190 | 3,6070 | |
| | | Desvio | 3,0680 | 1,6290 | 1,8130 | 2,9850 | 3,0390 | 3,4030 | 2,5604 | 2,1064 | 2,1072 | 3,3140 | 2,9770 | 3,0825 | |
| | | Mediana | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 3,0 | 2,0 | 2,0 | 2,0 | 3,0 | 3,0 | 3,0 | |
| | Choque = 2 | Média | 1,1040 | 1,0400 | 1,0430 | 1,1960 | 1,2010 | 1,2290 | 1,0580 | 1,0380 | 1,0380 | 1,2370 | 1,2350 | 1,2400 | |
| | | Desvio | 0,3370 | 0,1960 | 0,2030 | 0,4750 | 0,5280 | 0,5520 | 0,2464 | 0,1965 | 0,1965 | 0,5632 | 0,5693 | 0,5733 | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| | Choque = 3 | Média | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0010 | 1,0030 | 1,0000 | 1,0000 | 1,0000 | 1,0030 | 1,0040 | 1,0040 | |
| | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0320 | 0,0550 | 0,0000 | 0,0000 | 0,0000 | 0,0547 | 0,0632 | 0,0632 | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| | Choque = 4 | Média | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | |
| | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| | Uma média permanece constante e a outra muda ordenadamente | Choque = 0,5 | Média | 18,0060 | 7,7020 | 8,3330 | 12,6600 | 12,4360 | 13,9960 | 16,0130 | 12,2670 | 12,3300 | 15,2840 | 14,8300 | 15,0190 |
| | | | Desvio | 19,7990 | 8,2990 | 8,8250 | 12,7520 | 13,7150 | 15,0090 | 16,1420 | 11,8000 | 11,8470 | 15,0760 | 15,4490 | 15,4810 |
| | | | Mediana | 12,0 | 5,0 | 5,0 | 8,0 | 8,0 | 9,0 | 11,0 | 9,0 | 9,0 | 11,0 | 10,0 | 10,0 |
| Choque = 1 | | Média | 12,1050 | 5,3560 | 5,6480 | 8,1510 | 8,1890 | 9,1780 | 9,9930 | 8,3160 | 8,3760 | 9,4850 | 9,2350 | 9,3450 | |
| | | Desvio | 15,3130 | 5,5280 | 5,6930 | 8,0370 | 8,7860 | 10,4270 | 9,7280 | 8,2540 | 8,2730 | 9,9470 | 9,7460 | 9,8260 | |
| | | Mediana | 8,0 | 4,0 | 4,0 | 6,0 | 5,0 | 6,0 | 7,0 | 6,0 | 6,0 | 6,0 | 6,0 | 6,0 | |
| Choque = 2 | | Média | 4,3390 | 2,5650 | 2,7360 | 3,1220 | 3,1200 | 3,3820 | 3,3930 | 2,9710 | 2,9900 | 2,9840 | 2,9790 | 3,0090 | |
| | | Desvio | 4,7430 | 2,3830 | 2,8260 | 2,8850 | 3,0830 | 3,3820 | 3,0423 | 2,6186 | 2,6527 | 2,4976 | 2,6320 | 2,6472 | |
| | | Mediana | 3,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | 2,0 | |
| Choque = 3 | | Média | 1,8050 | 1,3400 | 1,3510 | 1,4400 | 1,4010 | 1,4640 | 1,5830 | 1,4900 | 1,4920 | 1,4160 | 1,4070 | 1,4100 | |
| | | Desvio | 1,4340 | 0,7380 | 0,7490 | 0,9160 | 0,8640 | 0,9650 | 1,0159 | 0,8767 | 0,8790 | 0,7560 | 0,7642 | 0,7658 | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| Choque = 4 | | Média | 1,1890 | 1,0740 | 1,0820 | 1,0610 | 1,0540 | 1,0660 | 1,1300 | 1,0990 | 1,1010 | 1,0750 | 1,0720 | 1,0730 | |
| | | Desvio | 0,4890 | 0,2940 | 0,3090 | 0,2390 | 0,2310 | 0,2640 | 0,3915 | 0,3366 | 0,3419 | 0,2854 | 0,2809 | 0,2824 | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | |
| Todas as médias mudam | | Choque = 0,5, 3, 1, 1, 1 | Média | 1,4210 | 1,1870 | 1,2090 | 1,3250 | 1,3130 | 1,3520 | 1,2840 | 1,2260 | 1,2300 | 1,3380 | 1,3390 | 1,3460 |
| | | | Desvio | 0,8381 | 0,4922 | 0,5231 | 0,7196 | 0,6823 | 0,7298 | 0,5913 | 0,5208 | 0,5248 | 0,6650 | 0,6712 | 0,6743 |
| | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |

Legenda: **Teo:** Limites de controle e ARL para o Limite de Controle obtido de forma teórica; **Pb5:** Limites de controle e ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes:** Limites de Controle e ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

Anexo B – Tabelas do *ARL in control* e *ARL out of control*. Processo com distribuição normal multivariada não autocorrelacionada com $p=10$ variáveis

Tabela B.1: *ARL in control e out of control* de processos com distribuição normal multivariada com forte correlação entre as variáveis ($0,75 < \rho \leq 1$) e $m=50$.

| Distribuição multivariada com $p=10$, $0,75 < \rho \leq 1$ e $m=50$ | | Medidas Descritivas | Estatística T ² de Hotelling | | | Estatística M | | | Componentes Principais | | |
|--|--|--------------------------|---|--------|--------|---------------|---------|---------|------------------------|---------|---------|
| | | | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes |
| ARL in control | | Média | 22,4700 | 4,4860 | 4,7880 | 17,9780 | 19,1980 | 21,3840 | 20,6680 | 25,7400 | 32,2330 |
| | | Desvio | 25,4224 | 4,5772 | 4,8519 | 23,9266 | 26,1010 | 27,5254 | 20,1617 | 34,7125 | 44,4016 |
| | | Mediana | 14,0 | 3,0 | 3,0 | 10,0 | 10,5 | 12,0 | 14,0 | 15,0 | 18,0 |
| Todas as médias mudam simultaneamente na mesma escala | Choque = 0,5 | Média | 19,0520 | 3,9050 | 4,2050 | 10,9310 | 10,7900 | 12,2820 | 12,1610 | 15,8830 | 20,1620 |
| | | Desvio | 21,6999 | 3,9163 | 4,1243 | 11,2095 | 11,7152 | 13,9016 | 11,4853 | 22,7059 | 31,2978 |
| | | Mediana | 12,0 | 3,0 | 3,0 | 7,0 | 7,0 | 8,0 | 9,0 | 9,0 | 11,0 |
| | Choque = 1 | Média | 13,3220 | 3,0950 | 3,2370 | 4,7160 | 4,9490 | 5,3490 | 5,0220 | 6,4990 | 7,5800 |
| | | Desvio | 14,9415 | 2,7128 | 2,8211 | 4,6833 | 5,7243 | 6,1977 | 4,5204 | 8,2825 | 10,8292 |
| | | Mediana | 8,0 | 2,0 | 2,0 | 3,0 | 3,0 | 3,0 | 4,0 | 4,0 | 4,0 |
| | Choque = 2 | Média | 5,4660 | 1,9640 | 2,0370 | 1,7150 | 1,7250 | 1,8080 | 1,7080 | 1,9400 | 2,0310 |
| | | Desvio | 5,8357 | 1,4501 | 1,5317 | 1,1905 | 1,2316 | 1,3262 | 1,0972 | 1,5469 | 1,6849 |
| | | Mediana | 4,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| | Choque = 3 | Média | 2,1350 | 1,2520 | 1,2680 | 1,1120 | 1,1180 | 1,1290 | 1,1030 | 1,1420 | 1,1710 |
| | | Desvio | 1,7157 | 0,5699 | 0,5818 | 0,3655 | 0,3718 | 0,3880 | 0,3354 | 0,4336 | 0,5020 |
| | | Mediana | 2,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| | Choque = 4 | Média | 1,2610 | 1,0400 | 1,0450 | 1,0090 | 1,0090 | 1,0100 | 1,0090 | 1,0150 | 1,0150 |
| | | Desvio | 0,5772 | 0,2060 | 0,2168 | 0,0945 | 0,0945 | 0,0995 | 0,0945 | 0,1216 | 0,1216 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| | Uma média permanece constante e a outra muda ordenadamente | Choque = 0,5 | Média | 3,5580 | 1,5440 | 1,6010 | 13,7570 | 14,7700 | 16,8180 | 20,2220 | 23,9690 |
| Desvio | | | 3,5150 | 0,9576 | 1,0241 | 15,8802 | 18,8743 | 21,1531 | 20,1340 | 32,1523 | 40,2064 |
| Mediana | | | 2,0 | 1,0 | 1,0 | 9,0 | 9,0 | 10,0 | 14,0 | 14,0 | 16,0 |
| Choque = 1 | | Média | 1,0950 | 1,0090 | 1,0110 | 9,7010 | 9,7260 | 11,1930 | 18,8390 | 26,7210 | 33,7850 |
| | | Desvio | 0,3195 | 0,0945 | 0,1044 | 12,3048 | 12,4143 | 14,2337 | 18,5316 | 39,4050 | 62,1330 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 6,0 | 6,0 | 7,0 | 13,0 | 14,0 | 16,5 |
| Choque = 2 | | Média | 1,0000 | 1,0000 | 1,0000 | 2,8670 | 2,8840 | 3,1150 | 17,7870 | 23,8530 | 29,9750 |
| | | Desvio | 0,0000 | 0,0000 | 0,0000 | 2,6549 | 2,8229 | 3,0485 | 17,3933 | 37,1640 | 46,1583 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 | 13,0 | 14,0 | 16,0 |
| Choque = 3 | | Média | 1,0000 | 1,0000 | 1,0000 | 1,3760 | 1,3720 | 1,4250 | 16,1360 | 22,3840 | 28,1840 |
| | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,7611 | 0,8511 | 0,9345 | 14,7287 | 30,6518 | 41,5274 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 11,0 | 12,0 | 14,0 |
| Choque = 4 | | Média | 1,0000 | 1,0000 | 1,0000 | 1,0430 | 1,0500 | 1,0550 | 14,6390 | 19,7250 | 23,7620 |
| | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,2218 | 0,2440 | 0,2531 | 13,8895 | 26,6382 | 31,0441 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 10,0 | 11,0 | 14,0 |
| 5 médias mudam 5 ficam constantes | | Choque = 0,5, 3, 1, 1, 1 | Média | 1,0000 | 1,0000 | 1,0000 | 1,4390 | 1,4360 | 1,4720 | 4,2440 | 5,4120 |
| | Desvio | | 0,0000 | 0,0000 | 0,0000 | 0,8312 | 0,8140 | 0,8520 | 3,5444 | 6,4781 | 7,6813 |
| | Mediana | | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 3,0 | 3,0 | 4,0 |

Legenda: **Teo:** Limites de controle e ARL para o Limite de Controle obtido de forma teórica; **Pb5:** Limites de controle e ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes:** Limites de Controle e ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

Tabela B.2: *ARL in control* e *out of control* de processos com distribuição normal multivariada com forte correlação entre as variáveis ($0,75 < \rho \leq 1$) e $m=200$.

| Distribuição multivariada com $p=10$, $0,75 \leq \rho \leq 1$ e $m=200$ | | Medidas Descritivas | Estatística T ² de Hotelling | | | Estatística M | | | Componentes Principais | | |
|--|--|--------------------------|---|---------|---------|---------------|---------|---------|------------------------|---------|---------|
| | | | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes |
| ARL in control | | Média | 19,0940 | 11,0440 | 11,3030 | 19,9900 | 20,7990 | 21,0630 | 20,8870 | 23,2210 | 23,8690 |
| | | Desvio | 18,8140 | 10,7690 | 11,0200 | 21,0520 | 22,7000 | 22,8840 | 21,0200 | 24,5290 | 25,1800 |
| | | Mediana | 13,0 | 8,0 | 8,0 | 13,0 | 13,0 | 13,0 | 14,0 | 16,0 | 16,0 |
| Todas as médias mudam simultaneamente na mesma escala | Choque = 0,5 | Média | 16,4990 | 9,8800 | 9,9610 | 11,4140 | 11,6310 | 11,7210 | 12,2010 | 14,0330 | 14,2570 |
| | | Desvio | 15,6170 | 9,7220 | 9,7240 | 11,3120 | 11,6760 | 11,7030 | 12,3500 | 15,6310 | 15,7630 |
| | | Mediana | 12,0 | 7,0 | 7,0 | 8,0 | 8,0 | 8,0 | 9,0 | 9,0 | 9,0 |
| | Choque = 1 | Média | 10,5640 | 6,7440 | 6,8930 | 5,1060 | 5,1710 | 5,2460 | 5,2660 | 5,8460 | 5,9540 |
| | | Desvio | 11,0460 | 6,1640 | 6,2580 | 4,5630 | 4,6810 | 4,7370 | 4,8050 | 5,6960 | 5,7960 |
| | | Mediana | 7,0 | 5,0 | 5,0 | 4,0 | 4,0 | 4,0 | 4,0 | 4,0 | 4,0 |
| | Choque = 2 | Média | 3,6790 | 2,8329 | 2,8471 | 1,7071 | 1,7271 | 1,7414 | 1,6971 | 1,7586 | 1,7743 |
| | | Desvio | 3,1130 | 2,2214 | 2,2225 | 1,1810 | 1,1870 | 1,2034 | 1,1155 | 1,1830 | 1,1855 |
| | | Mediana | 3,0 | 2,0 | 2,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| | Choque = 3 | Média | 1,6500 | 1,4414 | 1,4471 | 1,0914 | 1,0871 | 1,0886 | 1,0843 | 1,0929 | 1,0986 |
| | | Desvio | 1,0613 | 0,8223 | 0,8278 | 0,3344 | 0,3112 | 0,3131 | 0,3120 | 0,3141 | 0,3258 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| | Choque = 4 | Média | 1,1543 | 1,1029 | 1,1057 | 1,0057 | 1,0057 | 1,0057 | 1,0100 | 1,0114 | 1,0114 |
| | | Desvio | 0,4062 | 0,3267 | 0,3301 | 0,0754 | 0,0754 | 0,0754 | 0,0996 | 0,1064 | 0,1064 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| | Uma média permanece constante e a outra muda ordenadamente | Choque = 0,5 | Média | 2,8300 | 2,0871 | 2,1014 | 15,1060 | 16,1130 | 16,2540 | 21,6630 | 24,0000 |
| Desvio | | | 2,3471 | 1,5893 | 1,6019 | 15,4640 | 17,0830 | 17,1700 | 22,0180 | 26,6500 | 27,5900 |
| Mediana | | | 2,0 | 2,0 | 2,0 | 10,0 | 11,0 | 11,0 | 14,0 | 15,0 | 15,0 |
| Choque = 1 | | Média | 1,0357 | 1,0200 | 1,0200 | 9,5690 | 10,1390 | 10,3340 | 20,0640 | 21,8840 | 22,6430 |
| | | Desvio | 0,1933 | 0,1401 | 0,1401 | 9,8860 | 10,9790 | 11,1960 | 20,8080 | 24,1780 | 25,2240 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 7,0 | 7,0 | 7,0 | 13,0 | 14,0 | 15,0 |
| Choque = 2 | | Média | 1,0000 | 1,0000 | 1,0000 | 2,8629 | 2,8614 | 2,8829 | 18,0500 | 20,2610 | 20,7010 |
| | | Desvio | 0,0000 | 0,0000 | 0,0000 | 2,2684 | 2,3419 | 2,3622 | 17,7720 | 21,5220 | 21,8640 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 | 12,0 | 13,0 | 13,0 |
| Choque = 3 | | Média | 1,0000 | 1,0000 | 1,0000 | 1,3757 | 1,4000 | 1,4029 | 16,3770 | 18,5190 | 18,9800 |
| | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,7118 | 0,7412 | 0,7396 | 15,8970 | 19,7170 | 20,0010 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 11,0 | 12,0 | 13,0 |
| Choque = 4 | | Média | 1,0000 | 1,0000 | 1,0000 | 1,0557 | 1,0557 | 1,0600 | 13,3100 | 15,8100 | 16,3140 |
| | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,2417 | 0,2417 | 0,2494 | 11,9300 | 17,7170 | 18,0510 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 10,0 | 11,0 | 11,0 |
| 5 médias mudam 5 ficam constantes | | Choque = 0,5, 3, 1, 1, 1 | Média | 1,0000 | 1,0000 | 1,0000 | 1,4500 | 1,4429 | 1,4529 | 4,3430 | 4,5330 |
| | Desvio | | 0,0000 | 0,0000 | 0,0000 | 0,7708 | 0,7665 | 0,7691 | 3,6710 | 3,8240 | 3,9470 |
| | Mediana | | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 3,0 | 3,0 | 3,0 |

Legenda: **Teo:** Limites de controle e ARL para o Limite de Controle obtido de forma teórica; **Pb5:** Limites de controle e ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes:** Limites de Controle e ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

Tabela B.3: *ARL in control* e *out of control* de processos com distribuição normal multivariada com correlação intermediária entre as variáveis ($0,50 < \rho \leq 0,75$) e $m=50$.

| Distribuição multivariada com $p=10$, $0,50 < \rho \leq 0,75$ e $m=50$ | | Medidas Descritivas | Estatística T ² de Hotelling | | | Estatística M | | | Componentes Principais | | | | |
|---|--|--------------------------|---|---------|---------|---------------|---------|---------|------------------------|---------|---------|---------|---------|
| | | | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | | |
| ARL in control | | Média | 21,7010 | 4,1940 | 4,4830 | 15,6630 | 16,2190 | 18,3610 | 21,3090 | 28,7230 | 35,7650 | | |
| | | Desvio | 25,2594 | 3,9542 | 4,3188 | 18,0162 | 21,1802 | 23,1805 | 20,0964 | 49,8088 | 62,2087 | | |
| | | Mediana | 13,0 | 3,0 | 3,0 | 10,0 | 10,0 | 11,0 | 15,0 | 15,0 | 19,0 | | |
| Todas as médias mudam simultaneamente na mesma escala | | Choque = 0,5 | | Média | 18,9590 | 3,9710 | 4,2480 | 9,2620 | 9,2400 | 10,1850 | 10,9150 | 15,0090 | 18,2580 |
| | | | | Desvio | 22,0251 | 3,6007 | 3,9971 | 9,8113 | 10,3779 | 11,3567 | 10,5467 | 19,9000 | 26,5646 |
| | | | | Mediana | 12,0 | 3,0 | 3,0 | 6,0 | 6,0 | 7,0 | 7,0 | 8,0 | 10,0 |
| | | Choque = 1 | | Média | 11,4740 | 3,0350 | 3,1820 | 4,2150 | 4,1340 | 4,5450 | 4,2440 | 5,4130 | 6,2870 |
| | | | | Desvio | 14,0071 | 2,9689 | 3,0891 | 4,0449 | 4,1176 | 4,7166 | 3,8211 | 6,3749 | 8,7117 |
| | | | | Mediana | 7,0 | 2,0 | 2,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 4,0 |
| | | Choque = 2 | | Média | 3,9750 | 1,6470 | 1,6950 | 1,5040 | 1,4970 | 1,5470 | 1,4280 | 1,5740 | 1,6590 |
| | | | | Desvio | 4,0108 | 1,1161 | 1,1582 | 0,9439 | 0,9343 | 0,9699 | 0,8388 | 1,0685 | 1,2117 |
| | | | | Mediana | 3,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| | | Choque = 3 | | Média | 1,7290 | 1,1240 | 1,1340 | 1,0540 | 1,0540 | 1,0660 | 1,0370 | 1,0580 | 1,0680 |
| | | | | Desvio | 1,3294 | 0,3725 | 0,3927 | 0,2390 | 0,2390 | 0,2602 | 0,1992 | 0,2621 | 0,3090 |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| Choque = 4 | | Média | 1,1220 | 1,0140 | 1,0170 | 1,0060 | 1,0050 | 1,0060 | 1,0040 | 1,0040 | 1,0050 | | |
| | | Desvio | 0,3678 | 0,1175 | 0,1293 | 0,0773 | 0,0706 | 0,0773 | 0,0632 | 0,0632 | 0,0706 | | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | | |
| Uma média permanece constante e a outra muda ordenadamente | | Choque = 0,5 | | Média | 12,3590 | 2,9270 | 3,0670 | 12,7820 | 12,7810 | 14,6480 | 20,2840 | 25,7220 | 30,9810 |
| | | | | Desvio | 15,4222 | 2,5648 | 2,6810 | 14,3028 | 15,7064 | 18,1091 | 21,4126 | 31,1263 | 41,7865 |
| | | | | Mediana | 7,0 | 2,0 | 2,0 | 8,0 | 8,0 | 9,0 | 13,5 | 16,0 | 18,0 |
| | | Choque = 1 | | Média | 4,0690 | 1,6410 | 1,7150 | 10,0870 | 10,0200 | 11,7030 | 19,0140 | 26,3640 | 33,1710 |
| | | | | Desvio | 4,2730 | 1,1033 | 1,2277 | 10,9044 | 11,2102 | 13,5661 | 17,9202 | 38,3571 | 47,2795 |
| | | | | Mediana | 3,0 | 1,0 | 1,0 | 7,0 | 7,0 | 7,0 | 13,0 | 14,0 | 16,5 |
| | | Choque = 2 | | Média | 1,1180 | 1,0170 | 1,0190 | 3,6260 | 3,5310 | 3,8230 | 18,1230 | 25,2110 | 30,2790 |
| | | | | Desvio | 0,3552 | 0,1293 | 0,1437 | 3,7121 | 3,4807 | 3,7728 | 17,9714 | 51,3985 | 57,3218 |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 3,0 | 13,0 | 14,0 | 17,0 |
| | | Choque = 3 | | Média | 1,0000 | 1,0000 | 1,0000 | 1,5380 | 1,5420 | 1,5920 | 16,3770 | 22,6030 | 28,6840 |
| | | | | Desvio | 0,0000 | 0,0000 | 0,0000 | 1,0778 | 1,1169 | 1,1674 | 15,6477 | 30,3816 | 40,1618 |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 12,0 | 13,0 | 16,0 |
| Choque = 4 | | Média | 1,0000 | 1,0000 | 1,0000 | 1,0970 | 1,0920 | 1,1040 | 12,7980 | 16,9330 | 20,3280 | | |
| | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,3251 | 0,3188 | 0,3336 | 12,2660 | 22,0791 | 28,8806 | | |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 9,0 | 10,0 | 11,0 | | |
| 5 médias mudam 5 ficam constantes | | Choque = 0,5, 3, 1, 1, 1 | | Média | 1,7460 | 1,1560 | 1,1670 | 1,5290 | 1,5340 | 1,5850 | 3,4920 | 4,3250 | 4,8030 |
| | | | | Desvio | 1,2079 | 0,4493 | 0,4640 | 0,9727 | 1,0039 | 1,0835 | 2,9938 | 5,1203 | 5,6609 |
| | | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | 3,0 | 3,0 |

Legenda: **Teo:** Limites de controle e ARL para o Limite de Controle obtido de forma teórica; **Pb5:** Limites de controle e ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes:** Limites de Controle e ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

Tabela B.4: *ARL in control e out of control* de processos com distribuição normal multivariada com correlação intermediária entre as variáveis ($0,50 < \rho \leq 0,75$) e $m=200$.

| Distribuição multivariada com $p=10$, $0,50 < \rho \leq 0,75$ e $m=200$ | | Medidas Descritivas | Estatística T ² de Hotelling | | | Estatística M | | | Componentes Principais | | |
|--|--------------------------|---------------------|---|---------|---------|---------------|---------|---------|------------------------|---------|---------|
| | | | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes |
| ARL in control | | Média | 18,5480 | 11,1230 | 11,2990 | 20,1710 | 22,2080 | 22,8610 | 18,3560 | 18,4020 | 18,8090 |
| | | Desvio | 18,1302 | 10,9118 | 11,0069 | 18,9949 | 22,4063 | 23,0855 | 18,0069 | 18,6503 | 19,0943 |
| | | Mediana | 13,0 | 8,0 | 8,0 | 14,0 | 16,0 | 16,0 | 12,0 | 12,0 | 12,0 |
| Todas as médias mudam simultaneamente na mesma escala | Choque = 0,5 | Média | 15,2940 | 9,0210 | 9,1760 | 9,9610 | 10,1580 | 10,3060 | 10,2280 | 11,4780 | 11,8190 |
| | | Desvio | 15,3707 | 8,9064 | 9,1407 | 9,2932 | 10,1083 | 10,3060 | 9,7127 | 12,0111 | 12,3648 |
| | | Mediana | 10,0 | 6,0 | 6,0 | 7,0 | 7,0 | 7,0 | 7,0 | 7,0 | 8,0 |
| | Choque = 1 | Média | 8,9720 | 6,1310 | 6,1930 | 4,3800 | 4,4330 | 4,4400 | 4,1000 | 4,4970 | 4,5510 |
| | | Desvio | 8,4256 | 5,6958 | 5,7389 | 3,9994 | 4,2208 | 4,1663 | 3,4418 | 4,1016 | 4,1748 |
| | | Mediana | 6,0 | 4,0 | 4,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 | 3,0 |
| | Choque = 2 | Média | 3,0730 | 2,4100 | 2,4500 | 1,5510 | 1,5430 | 1,5530 | 1,4420 | 1,4590 | 1,4670 |
| | | Desvio | 2,6135 | 1,8459 | 1,8940 | 0,8719 | 0,8700 | 0,8798 | 0,7676 | 0,8018 | 0,8060 |
| | | Mediana | 2,0 | 2,0 | 2,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| | Choque = 3 | Média | 1,3830 | 1,2550 | 1,2570 | 1,0720 | 1,0770 | 1,0770 | 1,0520 | 1,0550 | 1,0550 |
| | | Desvio | 0,7409 | 0,5885 | 0,5944 | 0,2737 | 0,2813 | 0,2813 | 0,2310 | 0,2367 | 0,2367 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| | Choque = 4 | Média | 1,0480 | 1,0290 | 1,0290 | 1,0020 | 1,0030 | 1,0030 | 1,0020 | 1,0020 | 1,0020 |
| | | Desvio | 0,2139 | 0,1679 | 0,1679 | 0,0447 | 0,0547 | 0,0547 | 0,0447 | 0,0447 | 0,0447 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| Uma média permanece constante e a outra muda ordenadamente | Choque = 0,5 | Média | 9,2400 | 6,0820 | 6,1270 | 16,0560 | 16,0970 | 16,2880 | 20,0780 | 22,4630 | 23,3820 |
| | | Desvio | 9,3631 | 5,8223 | 5,8408 | 15,6757 | 15,8350 | 16,0491 | 18,8750 | 22,3187 | 23,1646 |
| | | Mediana | 6,0 | 4,0 | 4,0 | 11,0 | 11,0 | 11,0 | 14,0 | 15,0 | 16,0 |
| | Choque = 1 | Média | 3,0400 | 2,3520 | 2,3630 | 11,0360 | 11,0820 | 11,2400 | 19,9870 | 20,7720 | 21,4150 |
| | | Desvio | 2,6977 | 1,9047 | 1,9149 | 11,8008 | 11,8008 | 11,9764 | 18,7663 | 20,5944 | 20,9181 |
| | | Mediana | 2,0 | 2,0 | 2,0 | 7,0 | 7,0 | 7,0 | 14,0 | 15,0 | 15,0 |
| | Choque = 2 | Média | 1,0490 | 1,0280 | 1,0280 | 3,5510 | 3,4680 | 3,5110 | 17,9920 | 20,1350 | 20,9830 |
| | | Desvio | 0,2251 | 0,1651 | 0,1651 | 3,3446 | 3,1842 | 3,2141 | 17,3038 | 21,6814 | 22,6383 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 3,0 | 13,0 | 14,0 | 14,0 |
| | Choque = 3 | Média | 1,0000 | 1,0000 | 1,0000 | 1,4970 | 1,4830 | 1,4880 | 15,0380 | 17,2570 | 17,7220 |
| | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,8665 | 0,8651 | 0,8710 | 14,1823 | 18,1270 | 18,5155 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 11,0 | 11,0 | 11,0 |
| | Choque = 4 | Média | 1,0000 | 1,0000 | 1,0000 | 1,0740 | 1,0750 | 1,0770 | 12,6300 | 14,0870 | 14,5160 |
| | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,2874 | 0,2854 | 0,2918 | 12,3526 | 14,5673 | 14,8347 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 9,0 | 9,0 | 10,0 |
| 5 médias mudam 5 ficam constantes | Choque = 0,5, 3, 1, 1, 1 | Média | 1,4000 | 1,2460 | 1,2480 | 1,5290 | 1,5130 | 1,5220 | 3,3590 | 3,6520 | 3,7140 |
| | | Desvio | 0,7916 | 0,5847 | 0,5872 | 0,9675 | 0,9438 | 0,9521 | 2,9149 | 3,3589 | 3,4430 |
| | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,5 |

Legenda: **Teo:** Limites de controle e ARL para o Limite de Controle obtido de forma teórica; **Pb5:** Limites de controle e ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes:** Limites de Controle e ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

Tabela B.5: *ARL in control* e *out of control* de processos com distribuição normal multivariada com correlação nula entre as variáveis ($\rho = 0$) e $m=50$ e $m=200$.

| Distribuição normal multivariada com $p=10$, $\rho=0$ $m=50$ e $m=200$ | | Medidas Descritivas | m=50 | | | | | | m=200 | | | | | | |
|---|---------|--|---|--------------------------|--------|---------------|---------|---------|---|---------|---------|---------------|---------|---------|---------|
| | | | Estatística T ² de Hotelling | | | Estatística M | | | Estatística T ² de Hotelling | | | Estatística M | | | |
| | | | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | Teo | Pb=5 | Bes | |
| <i>ARL in control</i> | | Média | 23,6370 | 4,1300 | 4,2280 | 12,5330 | 10,9630 | 11,2580 | 19,4760 | 12,0270 | 12,1200 | 16,8690 | 17,0960 | 17,1620 | |
| | | Desvio | 29,6320 | 3,7500 | 3,8970 | 13,3060 | 12,4440 | 12,7750 | 19,6372 | 11,6225 | 12,1740 | 16,7391 | 18,0633 | 18,1503 | |
| | | Mediana | 14,0 | 3,0 | 3,0 | 8,0 | 7,0 | 7,0 | 14,0 | 8,0 | 9,0 | 11,5 | 11,0 | 11,0 | |
| Todas as médias mudam simultaneamente na mesma escala | | Choque = 0,5 | Média | 9,0200 | 2,5790 | 2,6100 | 7,0470 | 6,5360 | 6,7570 | 6,8510 | 4,5980 | 4,6220 | 8,1780 | 8,0440 | 8,0490 |
| | | | Desvio | 9,6120 | 2,1240 | 2,1590 | 7,1810 | 8,2860 | 8,6720 | 6,5079 | 4,1998 | 4,2555 | 8,1407 | 8,1669 | 8,1613 |
| | | | Mediana | 6,0 | 2,0 | 2,0 | 5,0 | 4,0 | 4,0 | 5,0 | 3,0 | 3,0 | 6,0 | 5,0 | 5,0 |
| | | Choque = 1 | Média | 2,2980 | 1,3020 | 1,3110 | 2,6490 | 2,4570 | 2,5090 | 1,9040 | 1,5650 | 1,5680 | 2,9620 | 2,9100 | 2,9120 |
| | | | Desvio | 1,9440 | 0,6920 | 0,7050 | 2,1660 | 2,0360 | 2,0860 | 1,3808 | 0,9668 | 0,9697 | 2,5010 | 2,6678 | 2,6663 |
| | | | Mediana | 2,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 | 1,0 | 1,0 | 1,0 | 2,0 | 2,0 | 2,0 |
| | | Choque = 2 | Média | 1,0080 | 1,0000 | 1,0000 | 1,0760 | 1,0640 | 1,0660 | 1,0030 | 1,0000 | 1,0000 | 1,0840 | 1,0810 | 1,0810 |
| | | | Desvio | 0,0890 | 0,0000 | 0,0000 | 0,3010 | 0,2680 | 0,2720 | 0,0547 | 0,0000 | 0,0000 | 0,3115 | 0,3075 | 0,3075 |
| | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| | | Choque = 3 | Média | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0010 | 1,0010 | 1,0010 |
| | | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0316 | 0,0316 | 0,0316 |
| | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| | | Choque = 4 | Média | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 | 1,0000 |
| | | | Desvio | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 |
| | | | Mediana | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| | | Uma média permanece constante e a outra muda ordenadamente | | Choque = 0,5 | Média | 20,3000 | 4,0810 | 4,1310 | 11,7060 | 11,1090 | 11,3060 | 16,6470 | 10,1950 | 10,2350 | 15,1590 |
| Desvio | 22,5880 | | | | 3,5480 | 3,6380 | 11,4020 | 11,8140 | 12,0590 | 16,3819 | 10,0936 | 10,1133 | 14,8111 | 15,6518 | 15,6401 |
| Mediana | 12,0 | | | | 3,0 | 3,0 | 8,0 | 7,0 | 7,5 | 11,0 | 7,0 | 7,0 | 11,0 | 10,0 | 10,0 |
| Choque = 1 | Média | | | 14,0310 | 3,4270 | 3,4910 | 9,1730 | 8,3730 | 8,5800 | 12,2990 | 7,4040 | 7,4040 | 10,6180 | 10,7030 | 10,7230 |
| | Desvio | | | 14,9980 | 2,9610 | 3,0970 | 9,7050 | 9,2830 | 9,4270 | 11,9111 | 7,2364 | 7,2364 | 10,3482 | 11,5013 | 11,5035 |
| | Mediana | | | 10,0 | 2,0 | 2,0 | 6,0 | 5,0 | 6,0 | 9,0 | 5,0 | 5,0 | 7,0 | 7,0 | 7,0 |
| Choque = 2 | Média | | | 6,4760 | 2,1210 | 2,1410 | 3,7480 | 3,5750 | 3,6080 | 4,6970 | 3,3460 | 3,3630 | 3,8640 | 3,8000 | 3,8150 |
| | Desvio | | | 7,5020 | 1,6340 | 1,6620 | 3,6770 | 3,6830 | 3,7060 | 4,3373 | 2,9660 | 2,9732 | 3,2970 | 3,3267 | 3,3298 |
| | Mediana | | | 4,0 | 2,0 | 2,0 | 3,0 | 2,0 | 2,0 | 3,0 | 2,0 | 2,0 | 3,0 | 3,0 | 3,0 |
| Choque = 3 | Média | | | 2,5910 | 1,3790 | 1,3780 | 1,6640 | 1,6310 | 1,6460 | 2,0030 | 1,6700 | 1,6760 | 1,5730 | 1,5740 | 1,5760 |
| | Desvio | | | 2,2460 | 0,8280 | 0,8120 | 1,1320 | 1,1590 | 1,1770 | 1,4342 | 1,0575 | 1,0621 | 1,0686 | 1,0861 | 1,0860 |
| | Mediana | | | 2,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 2,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| Choque = 4 | Média | | | 1,5020 | 1,0990 | 1,1030 | 1,1260 | 1,1110 | 1,1140 | 1,3180 | 1,1930 | 1,1910 | 1,1240 | 1,1200 | 1,1200 |
| | Desvio | | | 0,8970 | 0,3340 | 0,3410 | 0,3950 | 0,3500 | 0,3620 | 0,6255 | 0,4711 | 0,4677 | 0,3778 | 0,3711 | 0,3711 |
| | Mediana | | | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |
| 5 médias mudam 5 ficam constantes | | | | Choque = 0,5, 3, 1, 1, 1 | Média | 1,4130 | 1,0850 | 1,0840 | 1,4140 | 1,3550 | 1,3770 | 1,2240 | 1,1420 | 1,1420 | 1,4080 |
| | | Desvio | 0,8060 | | 0,3190 | 0,3180 | 0,7770 | 0,7150 | 0,7400 | 0,5119 | 0,4025 | 0,4025 | 0,7656 | 0,7697 | 0,7697 |
| | | Mediana | 1,0 | | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 | 1,0 |

Legenda: **Teo:** Limites de controle e ARL para o Limite de Controle obtido de forma teórica; **Pb5:** Limites de controle e ARL para o Limite de Controle obtido pela a técnica de Polansky e Baker (2000) com $b=5$; **Bes:** Limites de Controle e ARL para o Limite de Controle obtido pela a técnica de Bessegato (2001).

Anexo C – Programas utilizados na obtenção da janela ótima h segundo as metodologias de Polansky e Baker (2000) e Bessegato (2001)

```

##### Método estudado por Bessegato ###

func.phi2=function(vetor.data, lambda)
{
  n = length(vetor.data)
  soma.real = 0
  soma.imag = 0
  for(j in 1:n) {
    arg = lambda * vetor.data[j]
    soma.real = soma.real + cos(arg)
    soma.imag = soma.imag + sin(arg)
  }
  phi2 = ((soma.real)^2 + (soma.imag)^2)/(n^2)
  phi2
}

alg.lambda=function(vetor.data, cota.sup = 8, cota.inf = 0){

n = length(vetor.data)
corte = 3/n
prec = 0.001      # Precisaõ ate'a 3a. casa
vetor.phi2 = numeric(0)
vetor.lambda = seq(cota.inf, cota.sup, prec)
vetor.phi2 = func.phi2(vetor.data, vetor.lambda)
cota.sup = min(vetor.lambda[vetor.phi2 <= corte])
cota.sup

}

alg.H=function(vetor.dados, lambda, int = 50){

n = length(vetor.dados)
vet = vetor.dados
h = lambda/(2 * int)
soma.par = 0
for(k in 1:(int - 1)) {
  x = 2 * h * k
  func = (func.phi2(vet, x) - 1/n) * (x^2)
  soma.par = soma.par + func
}
soma.impar = 0
for(k in 1:int) {
  x = h * (2 * k - 1)
  func = (func.phi2(vet, x) - 1/n) * (x^2)
  soma.impar = soma.impar + func
}

func.lambda = (func.phi2(vet, lambda) - 1/n) * (lambda^2)
soma = (h * (func.lambda + 2 * soma.par + 4 * soma.impar))/3
aga = soma/pi
aga

}

alg.hop.F=function(vetor.dados, H.hat, flag = 1){
pos = 3
if(flag == 0) {
  #Nucleo Normal
  int.bow = 0.56418958323
  var = 1
}
}

```

```

else {
# Nucleo Epanechnikov
  int.bow = 0.257142857
  var = 0.2
}
n = length(vetor.dados)
cte = exp(log(int.bow/(var * var * n))/3)
H.hat
h.opt = cte/exp(log(H.hat)/3)
h.opt
}

func.F.norm=function(x, hop, VET = vet){
#
# OBS.: Verificar insercao do vetor de dados na atribuicao da funcao
# *** Algoritmo para Calculo de Fn^ ***
# Nucleo Normal
# VET = vetor de dados
# lin = linha da matriz de dados
# MAT.hop = matriz com h=opt
# hop = largura da janela ótima
# n = numero de elementos da amostra
# *** Declaracao de Variaveis ***
  n = length(VET)
  soma = 0
  for(i in 1:n) {
    argh = ((x - VET[i])/hop)
    soma = soma + pnorm(argh, 0, 1)
  }
  Fn = soma/n
  Fn
}

#FUNCAO QUE ESTIMA O PERCENTIL

func.percentil=function(lamb, vet, hop.f)
{
#xx= valor de x e funcaoFnorm=F(x)

maximo=max(vet)
minimo=min(vet)
amplitude=maximo-minimo

E=(amplitude/10)
while (E > 0.001)
{
  if (lamb >0.50)
  {
    xx=max(vet)
    funcaoFnorm=func.F.norm(xx, hop.f, vet)
    while(funcaoFnorm >lamb)
    {
      xx=xx-E
      funcaoFnorm=func.F.norm(xx, hop.f, vet)
    }
  }
  else
  {
    xx=min(vet)

```



```

        funcaoFnorm=func.F.norm(xx, hop.f, vet)
        while(funcaoFnorm <lamb){
            xx=xx+E
            funcaoFnorm=func.F.norm(xx, hop.f, vet)
        }
    }

    #fim do while
E=E/2
}

    vetorpercentil=c(xx, funcaoFnorm,E*2)
    vetorpercentil
#FIM DA FUNCAO
}

bessegato=function(dados, alfa){

lambda=alg.lambda(dados, cota.sup = 12, cota.inf = 0)
H=alg.H(dados, lambda, int = 50)
hotimo=alg.hop.F(dados, H, flag = 0)
percentil=func.percentil(alfa,dados,hotimo)
percentil[1]

}

```

```

#### Método estudado por Polansky e Baker #####

psi.norm.r=function(dados,r){
desvio=min(stdev(dados),IQR(dados)/1.349)
psi=(((-
1)**(r/2))*factorial(r))/(((2*desvio)**(r+1))*factorial(r/2)*sqrt(pi))
psi
}

#### Função L #####

L=function(x,der){

if(der==2){
derivada=sqrt(1/(2*pi))*exp(-0.5*x^2)*(-1+x^2)
}
if(der==4){
derivada = sqrt(1/(2*pi))*exp(-0.5*x^2)*(3 - 6*x^2 + x^4)
}
if(der==6){
derivada = sqrt(1/(2*pi))*exp(-0.5*x^2)*(-15 + 45*x^2 - 15*x^4 + x^6)
}
if(der==8){
derivada = sqrt(1/(2*pi))*exp(-0.5*x^2)*(105 -420*x^2 + 210*x^4 -28*x^6
+x^8)
}

if(der==10){
derivada = sqrt(1/(2*pi))*exp(-0.5*x^2)*(-945 + 4725*x^2 - 3150*x^4 +
630*x^6 - 45*x^8 + x^10)
}

derivada
}

##### função gdoisjota #####

gdoisjota=function(dados,j,psi){

n=length(dados)
a=((2*L(0,2*j))/(-1*n*psi))
g2j=abs(a)^(1/(2*j+3))
g2j
}

##### Função Psi.r(g) #####
psi.j.g=function(dados,r,g){

n=length(dados)
soma=0
for(i in 1:n){
soma=soma+sum(L(((dados-dados[i])/g),r))
}

psichap=((1/n)^2)*((1/g)^(r+1))*soma
psichap

}

```

```
##### h otimo #####

h.otimo=function(dados,psichap){

int.bow=0.56418958323
n=length(dados)
h=(int.bow/(-n*1*psichap))^(1/3)
h
}

polanskybaker=function(dados,j,alfa){

psi=psi.norm.r(dados,2*j+2)
for(k in j:1){
g2j=gdoisjota(dados,k,psi)
psig=psi.j.g(dados,2*k,g2j)
psi=psig
}

h1=h.otimo(dados,psi)
hotimo=h1
percentil=func.percentil(alfa,dados,hotimo)
percentil[1]

}

```

