

UNIVERSIDADE FEDERAL DO AMAZONAS
INSTITUTO DE CIÊNCIAS EXATAS
PÓS-GRADUAÇÃO EM MATEMÁTICA

**ANÁLISE DE DESEMPENHO DO GRÁFICO DE CONTROLE POR GRUPOS
COM BASE EM ESTATÍSTICAS DE ORDEM**

LUIZ FELIPE MARTINS LIMA

MANAUS-AM

2024



UFAM

ANÁLISE DE DESEMPENHO DO GRÁFICO DE CONTROLE POR GRUPOS
COM BASE EM ESTATÍSTICA DE ORDEM

Luiz Felipe Martins Lima

Dissertação de Mestrado apresentada ao
Programa de Pós-graduação em Matemática,
da Universidade Federal do Amazonas, como
parte dos requisitos necessários à obtenção do
título de Mestre em Matemática

Orientadores: Jeremias da Silva Leão

Max Brandão de Oliveira

Manaus

Abril de 2024



Ministério da Educação
Universidade Federal do Amazonas
Coordenação do Programa de Pós-Graduação em Matemática

FOLHA DE APROVAÇÃO

"ANÁLISE DE DESEMPENHO DO GRÁFICO DE CONTROLE POR GRUPOS COM BASE EM ESTATÍSTICA DE ORDEM"

LUIZ FELIPE MARTINS LIMA

Dissertação de Mestrado defendida e aprovada pela banca examinadora constituída pelos Professores:

Prof. Dr. Max Brandão de Oliveira - UFPI - (Presidente)

Prof. Dr. João Welliandre Carneiro Alexandre - UFC - (Membro Externo)

Prof. Dr. Jhonnata Bezerra de Carvalho - UFPE - (Membro Externo)

Manaus, 06 de Maio de 2024



Documento assinado eletronicamente por **Max Brandão de Oliveira, Usuário Externo**, em 06/05/2024, às 12:41, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **João Welliandre Carneiro Alexandre, Usuário Externo**, em 06/05/2024, às 15:16, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Jhonnata Bezerra de Carvalho, Usuário Externo**, em 06/05/2024, às 17:50, conforme horário oficial de Manaus, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufam.edu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2036135** e o código CRC **7322F3DE**.

Av. General Rodrigo Octávio, 6200 - Bairro Coroado 1 Campus Universitário Senador Arthur Virgílio Filho,
Setor Norte - Telefone: (92) 3305-1181 / Ramal 2405
CEP 69080-900, Manaus/AM, pos-matematica@ufam.edu.br

Referência: Processo nº 23105.007291/2024-70

SEI nº 2036135

Ficha Catalográfica

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

L732a Lima, Luiz Felipe Martins
Análise de desempenho do gráfico de controle por grupos com base em estatística de ordem. / Luiz Felipe Martins Lima . 2024
58 f.: il. color; 31 cm.

Orientador: Jeremias da Silva Leão
Coorientador: Max Brandão de Oliveira
Dissertação (Mestrado em Matemática - Estatística) -
Universidade Federal do Amazonas.

1. Análise de desempenho. 2. Fluxos Correlacionados. 3. Densidade do máximo. 4. Controle Estatístico de Processo. I. Leão, Jeremias da Silva. II. Universidade Federal do Amazonas III. Título

Dedico este trabalho aos meus pais, cujo amor incondicional e apoio incansável têm sido a luz que ilumina meu caminho em todos os momentos da vida. A vocês, que sempre acreditaram em mim e me encorajaram a perseguir meus sonhos, dedico este trabalho com profunda gratidão e admiração.

Aos meus queridos irmãos, companheiros de jornada e fonte constante de alegria, aprendizado e inspiração. Obrigado por estarem sempre ao meu lado, compartilhando momentos inesquecíveis. Esta conquista também é de vocês, e é com imenso carinho que dedico esta dissertação.

Agradecimentos

Aos professores Jeremias da Silva Leão e Máx Brandão de Oliveira, por toda atenção, incentivo e orientações importantes e necessárias que recebi de ambos.

À minha família, cujo apoio incondicional foi fundamental em cada etapa do meu percurso acadêmico. Em especial, dedico à minha mãe, cujo amor, incentivo e compreensão foram luzes orientadoras durante todo o período do mestrado.

Aos meus colegas de mestrado, cuja jornada compartilhada foi marcada por colaboração, apoio mútuo e aprendizado conjunto. Suas contribuições e amizade tornaram esta jornada acadêmica uma experiência enriquecedora e memorável.

Aos professores do Departamento de Estatística da UFAM, pelo esforço coletivo para que as atividades profissionais de estatística ocupem mais espaço no mercado de trabalho, devido à sua importância incontestável.

À todos que me ajudaram de forma direta e indireta para a conclusão desta etapa na minha vida.

À Banca Examinadora, pela valorosa revisão.

À Fundação de Amparo a Pesquisa do Estado do Amazonas (FAPEAM), pelo suporte financeiro.

Resumo da Dissertação apresentada ao Programa de Pós-Graduação em Matemática, da Universidade Federal do Amazonas, como parte dos requisitos necessários para a obtenção do grau de Mestre em Matemática. (M.Sc.)

ANÁLISE DE DESEMPENHO DO GRÁFICO DE CONTROLE POR GRUPOS
COM BASE EM ESTATÍSTICA DE ORDEM

Luiz Felipe Martins Lima

Abril/2024

Orientadores: Jeremias da Silva Leão

Max Brandão de Oliveira

Área de Concentração: Estatística

O uso de gráficos de Shewhart não é recomendado para monitorar processos com múltiplos fluxos de produção, que podem gerar amostras com itens de diferentes populações. Essa situação pode afetar a eficácia dos gráficos de controle na identificação de causas especiais de variação. Uma abordagem comum é usar um par de gráficos de Shewhart para cada fluxo, mas isso pode tornar o monitoramento mais complexo e demorado. Outra opção é usar o Gráfico de Controle por Grupos, que permite controlar vários fluxos com um único par de gráficos. No entanto, ambos os tipos de gráficos podem ser influenciados pela presença de correlação entre os fluxos, violando o pressuposto de independência das amostras. Diante disso, este trabalho propõe um método de estimação da distribuição do máximo e do mínimo de uma Normal Multivariada Dependente, que pode lidar com a dependência entre as observações e permitir a simulação de diferentes cenários, alterando a média e a variância do processo. O método é usado para comparar o desempenho do Controle por Grupos em relação ao gráfico tradicional de Shewhart, mostrando-se competitivo em cenários com baixa correlação entre os dados e se saindo melhor quando a correlação é alta. Este método é aplicado a um conjunto de dados práticos, demonstrando

sua aplicabilidade e ilustrando como o gráfico por grupos é sensível à alterações na média ou na variância do processo, com detecções de pontos fora dos limites de controle tornando-se cada vez mais rápidas. Além disso, constatou-se que o método para obtenção do NMA com base no método de proposto neste estudo foi mais eficiente do que o de Monte Carlo, evidenciando mais robustez teórica.

Palavras-chave: Análise de desempenho; Fluxos Correlacionados; Densidade do máximo; Controle Estatístico de Processo.

Abstract of Dissertation presented to Postgraduate in Mathematics, of the Federal University of Amazonas, as a partial fulfillment of the requirements for the degree of Master of Mathematics. (M.Sc.)

PERFORMANCE ANALYSIS OF GROUP CONTROL CHART
BASED ON ORDER STATISTICS

Luiz Felipe Martins Lima

April/2024

Advisors: Jeremias da Silva Leão

Max Brandão de Oliveira

Research lines: Statistics

The use of Shewhart charts is not recommended for monitoring processes with multiple production flows, which can generate samples with items from different populations. This situation can affect the effectiveness of control charts in identifying special causes of variation. A common approach is to use a pair of Shewhart charts for each flow, but this can make monitoring more complex and time-consuming. Another option is to use the Group Control Chart, which allows controlling multiple flows with a single pair of charts. However, both types of charts can be influenced by the presence of correlation between flows, violating the assumption of independence of samples. Therefore, this work proposes a method for estimating the distribution of the maximum and minimum of a Dependent Multivariate Normal, which can handle the dependence between observations and allow the simulation of different scenarios by changing the mean and variance of the process. The method is used to compare the performance of Group Control with the traditional Shewhart chart, showing competitiveness in scenarios with low correlation between data and performing better when the correlation is high. This method is applied to a practical dataset, demonstrating its applicability and illustrating how the group chart is sensitive to changes in the mean or variance of the process, with detections of points

outside the control limits becoming faster. Additionally, it was found that the method for obtaining the NMA based on the proposed method in this study was more efficient than the Monte Carlo method, demonstrating greater theoretical robustness.

Keywords: Performance Analysis; Correlated Flows; Maximum Density; Statistical Process Control.

Sumário

| | |
|--|-------------|
| Lista de Figuras | xii |
| Lista de Tabelas | xiii |
| 1 Introdução | 1 |
| 1.1 Contextualização | 1 |
| 1.2 Formulação do problema e definição dos objetivos | 2 |
| 1.3 Estrutura da dissertação | 3 |
| 1.4 Suporte computacional | 4 |
| 2 Material e métodos | 5 |
| 2.1 Gráficos de Controle tradicionais | 5 |
| 2.1.1 GC para $\bar{X} - R$ | 7 |
| 2.1.2 GC para $\bar{X} - S$ | 7 |
| 2.2 Processos paralelos | 8 |
| 2.3 Gráficos por Grupos (GG) | 10 |
| 2.3.1 GG para \bar{X} | 12 |
| 2.3.2 GG para R | 12 |
| 2.3.3 NMA para a avaliação do GG | 13 |
| 2.3.4 Análise de Correlação | 18 |
| 3 Métodos e simulação | 20 |
| 3.1 Metodologia de Pesquisa | 20 |
| 3.2 Complexidade Teórica | 21 |
| 3.3 Método Arellano para a estimação da distribuição do máximo e do mínimo | 21 |
| 3.3.1 Estatística de ordem | 21 |

| | |
|---|-----------|
| <i>SUMÁRIO</i> | xi |
| 3.3.2 Método Arellano para a estimação da distribuição do máximo . . . | 23 |
| 3.3.3 Método Arellano para a estimação da distribuição do mínimo . . . | 25 |
| 3.4 Simulação | 27 |
| 3.5 O algoritmo desenvolvido | 28 |
| 4 Um estudo de desempenho entre GG, GG A e o gráfico tradicional de Shewhart | 31 |
| 4.1 Desempenho do GG A | 31 |
| 4.1.1 Eficiência entre o GG e o GG A para o NMA_1 | 37 |
| 4.2 Eficiência entre o GG A, GG e o GCS na presença de correlação e alterações na média e variância do processo | 39 |
| 4.2.1 Alteração na média do processo (δ) | 39 |
| 4.2.2 Alteração na variância do processo (λ) | 41 |
| 4.2.3 Alteração na média e na variância do processo | 44 |
| 4.2.4 Eficiência com alteração conjunta na média e na variância do processo | 45 |
| 5 Aplicação | 47 |
| 5.1 Elaboração do GG | 49 |
| 5.2 NMA_1 para a Aplicação | 50 |
| 5.3 Alteração na média do processo (δ) | 51 |
| 5.4 Alteração na variância do processo (λ) | 52 |
| 5.5 Alteração conjunta na média e na variância do processo | 52 |
| 6 Considerações Finais | 54 |
| 6.1 Conclusões | 54 |
| 6.2 Trabalhos Futuros | 55 |
| Referências Bibliográficas | 57 |

Lista de Figuras

| | | |
|-----|---|----|
| 3.1 | Densidade do máximo e do mínimo para uma amostra de tamanho 1.000. | 26 |
| 3.2 | Densidade do máximo e do mínimo para uma amostra de tamanho 1.000.000. | 27 |
| 3.3 | Fluxograma da simulação | 30 |
| 4.1 | NMA_1 para fluxos independentes (exato e simulado) e para fluxos dependentes com $\mu = \sigma^2 = 100$ | 33 |
| 4.2 | NMA_1 para fluxos dependentes com $\mu = 1 < \sigma^2 = 100$ e $r = 0,1;0,2;0,3;0,5;0,7$ e $0,9$ | 35 |
| 4.3 | NMA_1 para fluxos dependentes com $\mu = 100 > \sigma^2 = 9^2$ e $r = 0,1;0,2;0,3;0,5; e 0,7$ | 36 |
| 5.1 | Gráfico GG para o Processo de pintura a pó de fogões | 49 |
| 5.2 | Gráfico da correlação entre os fluxos (posições) do processo de pintura a pó de fogões. | 50 |

Lista de Tabelas

| | | |
|-----|--|----|
| 2.1 | Esquema de montagem geral do Gráfico de Controle (GC) tradicional para a média (\bar{X}), amplitude (R) e desvio-padrão (S). | 6 |
| 2.2 | Esquema de montagem para o gráfico por grupos para a média (\bar{X}), amplitude (R) e desvio-padrão (S) | 11 |
| 4.1 | NMA_1 para fluxos independentes (exato e simulado) e para fluxos dependentes com $\mu = \sigma^2 = 100$, obtido pelo algoritmo descrito no Capítulo 3 variando a correlação. | 32 |
| 4.2 | NMA_1 para fluxos dependentes com $\mu = 1 < \sigma^2 = 100$ e $r = 0,1; 0,2; 0,3; 0,5; 0,7$ e $0,9$ | 34 |
| 4.3 | NMA_1 para fluxos dependentes com $\mu = 100 > \sigma^2 = 9^2$, obtido pelo algoritmo proposto variando a correlação $r = 0,1; 0,2; 0,3; 0,5; 0,7$ | 36 |
| 4.4 | Eficiência relativa para NMA_1 para fluxos dependentes com $\mu = \sigma^2 = 100$, $NMA_{1.GG}$ obtido pelo algoritmo proposto em Oliveira (2013) e o $NMA_{1.GGA}$ pelo algoritmo variando a correlação | 37 |
| 4.5 | Eficiência relativa para NMA_1 para fluxos dependentes com $\mu = 1 < \sigma^2 = 100$, $NMA_{1.GG}$ obtido pelo algoritmo proposto em Oliveira (2013) e o $NMA_{1.GGA}$ pelo algoritmo variando a correlação. | 38 |
| 4.6 | Eficiência relativa para NMA_1 para fluxos dependentes com $\mu = 100 > \sigma^2 = 9^2$, $NMA_{1.GG}$ obtido pelo algoritmo proposto em Oliveira (2013) e o $NMA_{1.GGA}$ pelo algoritmo variando a correlação. | 38 |
| 4.7 | NMA_{21} para o GG A com correlação entre os fluxos e um deslocamento na média. | 40 |
| 4.8 | Eficiência entre o GCS e o GG A para um deslocamento na média do processo e com fluxos correlacionados. | 40 |

| | | |
|------|---|----|
| 4.9 | Eficiência entre o GG e o GG A para um deslocamento na média do processo e com fluxos correlacionados. | 41 |
| 4.10 | NMA_{22} para o GG A com correlação entre os fluxos e um deslocamento na variância. | 42 |
| 4.11 | Eficiência entre o GC e o GG A para um deslocamento na variância do processo e com fluxos correlacionados. | 42 |
| 4.12 | Eficiência entre o GG e o GG A para um deslocamento na variância do processo e com fluxos correlacionados. | 43 |
| 4.13 | NMA_{23} para o Gráfico por Grupos A com correlação $r = 0, 1; 0, 5; 0, 9$ e um deslocamento na média (δ) e na variância (λ) para $k = 3, 5, 10$ | 44 |
| 4.14 | Eficiência Relativa do NMA_{23} entre GCS e o GG A com correlação $r = 0, 1; 0, 5; 0, 9$ e um deslocamento na média (δ) e na variância (λ) para $k = 3, 5, 10$ | 45 |
| 4.15 | Eficiência Relativa do NMA_{23} entre GG e o GG A com correlação $r = 0, 1; 0, 5; 0, 9$ e um deslocamento na média (δ) e na variância (λ) para $k = 3, 5, 10$ | 46 |
| 5.1 | Tabela com o conjunto de dados "Processo de pintura a pó de fogões" . . . | 48 |
| 5.2 | NMA_{21} para o Processo de pintura a pó de fogões. | 51 |
| 5.3 | NMA_{22} para o Processo de pintura a pó de fogões. | 52 |
| 5.4 | NMA_{23} para o Processo de pintura a pó de fogões. | 53 |

Capítulo 1

Introdução

1.1 Contextualização

Em qualquer processo de produção, seja na indústria alimentícia, automotiva, farmacêutica ou em serviços, é necessário assegurar a qualidade do produto final. Para isso, é fundamental monitorar de maneira contínua o processo de produção, verificando se o produto final está dentro dos padrões exigidos. Dentre as diversas aplicações desse monitoramento, temos que as ferramentas do controle estatístico de processo (CEP) desempenham um papel crucial em análise de processos, podemos, por exemplo, que ressaltar a importância no setor alimentício, em que as empresas que fabricam alimentos precisam garantir que seus produtos sejam seguros e estejam dentro dos parâmetros de qualidade estabelecidos pelos órgãos reguladores. Para isso, é enfatizado por Montgomery (2000) que o monitoramento estatístico da qualidade das matérias-primas, do processo de fabricação e do produto final é essencial.

Dessa forma, temos que o CEP é uma das principais ferramentas utilizadas na indústria para monitorar e controlar a qualidade dos produtos e serviços. Desde que foi proposto o CEP tem modelos adaptativos e alternativos, com o objetivo de torná-lo mais eficiente e adequado às necessidades dos processos industriais. Nesse sentido, duas técnicas se destacam: o gráfico de Shewhart e o gráfico de controle por grupos. O gráfico de Shewhart, proposto por Shewhart (1931), é a técnica de CEP mais utilizada na indústria. Baseia-se na coleta de uma única amostra em cada ponto de coleta, com o objetivo de monitorar a média e a variação do processo ao longo do tempo. No entanto, conforme destacado por Costa *et al.* (2005), em ambientes de processos paralelos, nos quais várias

amostras são coletadas simultaneamente de diferentes fluxos, o gráfico de Shewhart pode não ser eficiente.

Nesse sentido, o gráfico de controle por grupos surge como uma alternativa mais adequada. Essa técnica, descrita em Montgomery (2000), utiliza uma amostra de cada grupo coletadas em intervalos regulares de tempo e as agrupa em conjuntos (grupos), em que os valores do máximo e mínimo da média de cada grupo são calculados e plotados no gráfico, permitindo uma análise da estabilidade do processo ao longo do tempo. Adicionalmente, temos que o gráfico de controle por grupos permite detectar variações locais que podem ocorrer em processos paralelos, identificando quais fluxos apresentam variações significativas e atuando de forma pontual para corrigir problemas específicos, como discutido por Costa *et al.* (2005). Essa técnica também permite uma redução do tempo de coleta de dados e uma melhor compreensão das variações no processo e das causas de problemas, como mencionado por Montgomery (2000).

Diante dessas considerações, é importante destacar a relevância do gráfico de controle por grupos como uma alternativa mais eficiente e adequada ao gráfico de Shewhart em ambientes de processos paralelos. Essa técnica permite uma análise mais precisa e rápida do processo, identificando variações locais e auxiliando na determinação das causas de problemas, conforme abordado em Costa *et al.* (2005).

1.2 Formulação do problema e definição dos objetivos

O desenvolvimento dos recursos tecnológicos, aliado à produção em larga escala, possibilita a implementação de processos produtivos com vários fluxos trabalhando simultaneamente. Cada fluxo fabrica o mesmo produto e pode interagir de forma isolada ou não com os demais, ou seja, pode existir uma relação entre eles. Do ponto de vista prático, isso pode ocorrer com o ajuste individual ou geral dos fluxos. Se o ajuste é executado em cada um separadamente, espera-se que não seja constatada relação entre eles, caso contrário, existem indícios dessa relação ser descrita como uma forma de dependência entre os fluxos, proporcionando a ocorrência de sucessivos alarmes falsos no processo de monitoramento, caso essa dependência seja incorporada no modelo.

Estudos mostram que essa dependência, quantificada pelo coeficiente de autocorrelação, afeta diretamente o desempenho do gráfico de controle caso não seja identificada,

como visto em Guimarães & Epprecht (2000). Em processos com múltiplos fluxos, a dependência também pode estar presente entre os fluxos na forma de correlação, o que leva à pergunta desta pesquisa: como o Gráfico de Controle por Grupos \bar{X} e R se comporta, em termos de desempenho e eficiência, com o Gráfico de Shewhart, na presença de correlação entre os fluxos?

É fundamental verificar a forma como a dependência entre os fluxos afeta o monitoramento de processos paralelos, para que seja possível propor um método capaz de corrigir essa distorção e impedir uma análise equivocada, já que existe a possibilidade de aplicação errônea dos métodos tradicionais em cenários cujos fluxos apresentam uma estrutura de correlação inerente ao processo. O desempenho dos gráficos de controle por grupos na presença de correlação entre os fluxos foi analisado em Oliveira (2013) através do método de Monte Carlo. Porém, devido a dificuldade da convergência da função de distribuição acumulada da normal, não se sabe como o número médio de amostras (NMA) até o alarme falso se comporta através da distribuição exata do máximo e do mínimo quando há uma estrutura de correlação entre os fluxos.

Portanto, o objetivo deste estudo é realizar uma investigação sobre o desempenho dos gráficos de controle por grupos quando aplicados a variáveis aleatórias com distribuição normais com dependência entre os fluxos produtivos. Serão utilizadas as abordagens tradicionais, como o gráfico de Shewhart para efeito de comparação, e será explorada a eficácia desses métodos em diferentes contextos, considerando a normalidade dos dados, o grau de relação entre os fluxos e os deslocamentos na média e na variância do processo.

1.3 Estrutura da dissertação

A presente dissertação é constituída por cinco capítulos. No primeiro capítulo, introduzimos ao leitor o contexto do trabalho, destacando a problemática abordada, justificativa para a pesquisa, os objetivos propostos e a aplicação prática dos resultados obtidos. Em seguida, o segundo capítulo, foi dedicado à exposição dos gráficos de controle tradicionais, análise de processos paralelos e a discussão detalhada sobre os gráficos de controle por grupos. Dessa forma, este capítulo fornecerá uma base teórica, apresentando as principais ferramentas e conceitos que serão explorados ao longo do trabalho. No terceiro capítulo abordamos a metodologia empregada no desenvolvimento da pesquisa. Detalha-

mos os procedimentos e recursos utilizados, os métodos empregados na simulação e os conceitos essenciais para compreender a estatística de ordem. Este capítulo servirá como guia para a compreensão da abordagem metodológica, estabelecendo as bases necessárias para a análise e interpretação dos resultados apresentados nos capítulos subsequentes. No quarto capítulo, apresentamos o desempenho e a eficiência do gráfico de médias em resposta a uma perturbação que afeta tanto a média quanto a variabilidade do processo (de maneira individual e conjunta). O quinto capítulo é dedicado a uma aplicação prática, comparando a eficiência com outros dois estudos. Por fim, no sexto capítulo, discutimos as principais características observadas no estudo de desempenho do gráfico de controle por grupos com base em estatística de ordem e perspectivas de trabalhos futuros.

1.4 Suporte computacional

A realização de um estudo de Monte Carlo para simulação dos cenários e a construção de gráficos por grupos representam um desafio complexo que demanda um robusto suporte computacional. Neste contexto, a escolha do software R como ferramenta central para conduzir a análise é crucial. O R Core Team (2023) desenvolveu uma linguagem de programação e ambiente estatístico, em que oferece uma vasta gama de pacotes e funcionalidades especializadas em simulação estatística, permitindo a implementação eficiente de modelos complexos. Sua natureza de código aberto e a comunidade ativa de usuários contribuem para uma flexibilidade significativa e em constante evolução. O uso do R proporciona não apenas a capacidade de conduzir estudos de Monte Carlo detalhados, mas também a flexibilidade necessária para explorar e visualizar os resultados de maneira intuitiva, facilitando a interpretação dos cenários simulados para a construção de gráficos por grupos.

Capítulo 2

Material e métodos

2.1 Gráficos de Controle tradicionais

Os gráficos de controle são ferramentas que permitem monitorar a média, variabilidade de um processo e detectar desvios que possam comprometer a qualidade do produto final. Os gráficos de controle tradicionais são aqueles que utilizam dados individuais, médias e desvios-padrão para monitorar a qualidade do processo. Segundo Garvin (1988), a gestão da qualidade é uma das estratégias competitivas mais importantes para as empresas, e a utilização de ferramentas estatísticas é fundamental para o sucesso desse processo. Shewhart (1931), por sua vez, foi o primeiro a desenvolver os gráficos de controle e propor sua utilização na indústria como forma de monitorar a qualidade dos processos. Os gráficos de controle tradicionais visam inspecionar a variabilidade natural de um processo ao longo do tempo, permitindo a identificação de desvios significativos que podem indicar a presença de causas especiais. Em Montgomery & Mastrangelo (2015), a ênfase é colocada na importância de controlar variáveis-chave de maneira eficiente para assegurar a qualidade do produto.

A construção de gráficos de controle tradicionais envolve a coleta sistemática de dados ao longo do tempo, frequentemente em amostras sequenciais. Em Amiri & Solis (2018), destaca-se que a seleção adequada de estatísticas descritivas e limites de controle é crucial para garantir a eficácia desses gráficos. Baseando-se na suposição de normalidade da característica da qualidade X , com média μ e desvio-padrão σ , temos que \bar{X} também é normalmente distribuída com média μ e desvio $\sigma_{\bar{X}} = \sigma/\sqrt{n}$, bem como a probabilidade do intervalo $[\mu_{\bar{X}} - z_{\alpha/2}\sigma_{\bar{X}}, \mu_{\bar{X}} + z_{\alpha/2}\sigma_{\bar{X}}]$ conter a verdadeira média μ é $(1 - \alpha)$.

Diante dessas informações vindas da suposição de normalidade, uma forma geral para a construção de gráficos de controle foi elaborada a partir da Tabela 2.1.

Tabela 2.1: Esquema de montagem geral do Gráfico de Controle (GC) tradicional para a média (\bar{X}), amplitude (R) e desvio-padrão (S).

| Subgrupo | Descrição das amostras | | | | \bar{X} | R | S |
|----------|------------------------|----------|----------|----------|-------------|----------|----------|
| 1 | X_{11} | X_{12} | \cdots | X_{1n} | \bar{X}_1 | R_1 | S_1 |
| 2 | X_{21} | X_{22} | \cdots | X_{2n} | \bar{X}_2 | R_2 | S_2 |
| \vdots | \vdots | \vdots | \ddots | \vdots | \vdots | \vdots | \vdots |
| k | X_{k1} | X_{k2} | \cdots | X_{kn} | \bar{X}_k | R_k | S_k |

A construção da Tabela 2.1 é função dos chamados parâmetros dos GC, n , h e L , em que n representa o tamanho da amostra, ou seja, o número de medições a serem realizadas dentro de cada amostragem; h é o intervalo entre as coletas; e L é o fator de largura do intervalo. A variação dessas quantidades interfere diretamente na análise e, tendo conhecimento da manipulação dos mesmos, pode-se achar uma combinação ótima entre eles que promova uma análise mais eficiente, no sentido de identificar rapidamente onde ocorreu um defeito na produção. Destaca-se ainda que a elaboração da referida tabela é baseada nos GC tradicionais, que são balanceados e não possuem dados faltantes.

Estes gráficos de controle tradicionais consistem na plotagem dos dados individuais ou das médias das amostras coletadas em um gráfico de controle, juntamente com linhas que marcam os limites de controle superior, central e inferior. O intuito é monitorar o processo e verificar se ele está dentro dos limites de controle, que são definidos a partir da média e do desvio-padrão dos dados coletados. Ramos (2000) destaca que os gráficos de controle tradicionais são eficazes para detectar desvios acentuados, como a presença de pontos fora dos limites de controle. No entanto, eles podem não ser tão eficazes para detectar desvios menores ou tendências no processo. Li & Tsung (2015) destacam a limitação desses gráficos quando aplicados a processos multivariados, ressaltando a necessidade de métodos mais avançados nesses cenários específicos.

Os gráficos de controle tradicionais são amplamente utilizados na indústria para monitorar a qualidade dos processos e controlar a variabilidade. Eles permitem uma visualização clara do processo e a identificação de desvios que possam comprometer a qualidade do produto final. É importante ressaltar que eles devem ser utilizados em conjunto com outras ferramentas estatísticas para uma gestão efetiva da qualidade. Os principais gráficos de controle são os gráficos \bar{X} , σ e R , estes gráficos são usados para monitorar a média

e a variabilidade do processo, respectivamente. O gráfico \bar{X} monitora a média do processo. Os gráficos σ e R podem ser usados para monitorar a variabilidade do processo. Se a dispersão aumenta, pode indicar um problema no processo de produção. O gráfico para σ é indicado para um tamanho amostral $n > 10$ (Ramos, 2000).

2.1.1 GC para $\bar{X} - R$

Os limites de controle para o gráfico \bar{X} são definidos por

$$LIC = \bar{\bar{X}} - A_2 \bar{R},$$

$$LC = \bar{\bar{X}},$$

$$LSC = \bar{\bar{X}} + A_2 \bar{R},$$

em que $A_2 = \frac{3}{\sqrt{nd_2}}$. Enquanto para limites de controle para o gráfico \bar{R} são

$$LIC_R = D_3 \cdot \bar{R},$$

$$LC = \bar{R},$$

$$LSC_R = D_4 \cdot \bar{R},$$

em que d_2 , D_3 e D_4 são constantes tabeladas, com $D_3 = 1 - 3\left(\frac{d_3}{d_2}\right)$ e $D_4 = 1 + 3\left(\frac{d_3}{d_2}\right)$, encontradas em Montgomery (2000), $\bar{\bar{X}} = \frac{1}{nk} \sum_{i=1}^n \sum_{j=1}^k \bar{X}_{ij}$, $R_i = \max(X_{ij}) - \min(X_{ij})$ e $\bar{R} = \frac{1}{k} \sum_{i=1}^k R_i$.

2.1.2 GC para $\bar{X} - S$

Os limites de controle para o gráfico para a média \bar{X} são definidos por

$$LIC_{\bar{X}} = \bar{\bar{X}} - \frac{3\bar{S}}{\sqrt{nc_4}},$$

$$LC = \bar{\bar{X}},$$

$$LSC = \bar{\bar{X}} + \frac{3\bar{S}}{\sqrt{nc_4}},$$

e os limites de controle para S são

$$LIC_S = B_3 \cdot \bar{S},$$

$$LC = \bar{S},$$

$$LSC_S = B_4 \cdot \bar{S}.$$

Em que $\bar{S} = \frac{1}{k} \sum_{i=1}^k S_i$ e S_i é o desvio padrão, as constantes c_4, B_3, B_4 são constantes tabeladas e podem ser encontradas em Montgomery (2000).

2.2 Processos paralelos

Os processos paralelos referem-se a sistemas de produção nos quais múltiplos fluxos operam simultaneamente, uma prática comum nas indústrias. Conforme destacado por Vasconcelos (2009), a adoção de processos produtivos nos quais um produto pode ser fabricado em diferentes fluxos ou até mesmo simultaneamente em conjuntos distintos de equipamentos.

Os benefícios inerentes aos processos paralelos são vastos e impactam diretamente a resiliência e eficiência operacional das operações industriais. A redundância é um dos principais trunfos dessa abordagem, em caso de falha em uma linha de produção, as outras linhas podem continuar operando, garantindo uma continuidade crucial para a entrega de produtos. A flexibilidade é outra vantagem significativa. Os processos paralelos permitem ajustes rápidos e mudanças na produção conforme necessário, respondendo ágil e eficazmente às demandas dinâmicas do mercado. A eficiência é maximizada com a capacidade de produção total aumentada, proporcionando não apenas uma resposta mais rápida às necessidades do mercado, mas também assegurando uma produção consistente e eficiente em termos de recursos.

O exame individual de cada fluxo em processos produtivos em paralelo, além de ser inviável devido à necessidade de múltiplos GC, suscita questionamentos sobre a aplicabilidade desses GC em processos paralelos, especialmente em relação à independência. A teoria de Shewhart, além de pressupor a normalidade, exige a independência entre as observações, condição nem sempre assegurada em contextos com múltiplos fluxos. A independência, fundamental na teoria de Shewhart, implica correlação zero, quando se

trata de uma distribuição normal. A validade da teoria dos GC depende da satisfação de suas suposições. Guimarães & Epprecht (2000) enfatizam que, na presença de dados correlacionados, os GC podem fornecer resultados distorcidos, gerando alarmes falsos (que ocorre quando um ponto está fora dos limites de controle, mas o processo está livre de causas especiais). A autocorrelação, medida bidimensional da correlação em uma única variável, é uma consideração importante, que será aprofundada nas seções subsequentes.

Em casos de múltiplos fluxos de produção, além da presença de autocorrelação dentro de cada fluxo, pode haver correlação entre eles. Nesse cenário, é crucial abordar a análise com cautela. Para um único fluxo, avaliamos a autocorrelação e a normalidade dos dados. Se essas condições não forem violadas, podemos aplicar o GC sem modificações. Caso contrário, é necessário realizar transformações nos dados para atingir a normalidade ou empregar métodos para contornar a autocorrelação, temas que serão discutidos nas seções seguintes.

Quando a indústria opera com vários fluxos de produção, a alternativa é elaborar dois GCs para cada fluxo, mesmo que em grande quantidade. Isso se torna essencial, pois a identificação da origem de uma falha (fluxo e grupo) torna-se impossível se as amostras contiverem informações misturadas de todos os fluxos. A independência facilita a análise, pois a presença de um fator que interfira na estrutura compromete diretamente a análise. A independência refere-se, por exemplo, ao ajuste das máquinas. Se os fluxos forem ajustados individualmente, isso indica independência; caso contrário, se for um ajuste para todos, pode indicar dependência. Portanto, é crucial realizar testes que avaliem a independência entre os fluxos, indo além da situação prática, uma vez que o ajuste não garante uma estrutura de independência.

A complexidade de aplicar o GC tradicional a processos paralelos e a consideração da dependência entre os fluxos tornam o monitoramento ainda mais desafiador. A relação entre os fluxos gera uma interferência generalizada na análise. Se uma linha produtiva estiver fortemente relacionada com as demais, seu comportamento influenciará diretamente as outras. Assim, a análise isolada de cada fluxo, além de inviável, não é capaz de detectar a presença de causas especiais, especialmente em relação à estrutura de dependência. Portanto, do ponto de vista da viabilidade, seria necessária um GC mais adequado para os processos paralelos.

2.3 Gráficos por Grupos (GG)

Os gráficos de controle por grupos são ferramentas estatísticas utilizadas para monitorar e controlar a qualidade de processos industriais em paralelo. No que diz respeito à eficiência na inspeção, os gráficos de controle tradicionais demandam a construção de um gráfico individual para cada fluxo de processo, resultando em uma abordagem burocrática e potencialmente onerosa em termos de custos de inspeção. Em contrapartida, os gráficos por grupos (GG) podem ser empregados para monitorar os processos paralelos através de um único par de gráficos, simplificando o monitoramento e reduzindo a complexidade operacional.

A redução de custos é uma vantagem adicional proporcionada pelos GG's ao inspecionar processos paralelos em relação aos GC tradicionais de Shewhart. Enquanto a construção de pares de gráficos (média e dispersão) para cada fluxo pode aumentar consideravelmente o tempo e o custo associado ao CEP, a utilização de um único par de gráficos para monitorar vários fluxos resulta em otimização de recursos de inspeção. O uso dos gráficos de grupos foi proposto por Boyd (1950). Boyd apresentou os gráficos de controle como uma forma de monitorar a variabilidade dos processos e detectar desvios que possam comprometer a qualidade dos produtos.

Os gráficos por grupos consistem na plotagem dos dados de uma amostra em um gráfico, em que é possível visualizar a média e a amplitude dos valores coletados. A partir disso, é possível definir limites de controle e identificar quando os valores se encontram fora desses limites. Isso indica que o processo está fora de controle e deve ser investigado para identificar a causa do problema e tomar as medidas necessárias para corrigi-lo. Ramos (2000) destaca que os gráficos de grupos podem ser utilizados em processos contínuos ou em bateladas, permitindo a monitoração da variabilidade do processo ao longo do tempo ou em lotes de produção. Essa ferramenta permite uma visão ampla do processo, possibilitando identificar problemas antes que eles afetem a qualidade do produto final. Em Oliveira (2013) é mostrado que o esquema do controle para os GG tem dois objetivos preliminares:

- Detectar um deslocamento na média de todos os fluxos;
- Detectar um deslocamento na média de um fluxo individual do produto.

Destaca-se que o esquema geral apresentado na Tabela 2.2 se baseia na característica

Tabela 2.2: Esquema de montagem para o gráfico por grupos para a média (\bar{X}), amplitude (R) e desvio-padrão (S)

| Grupo | Amostra | Descrição das amostras | | | \bar{X} | R | S |
|----------|----------|------------------------|----------|-----------|----------------|----------|----------|
| | | Prod.1 | ... | Prod.n | | | |
| 1 | 1 | X_{111} | ... | X_{11n} | \bar{X}_{11} | R_{11} | S_{11} |
| | \vdots | \vdots | ... | \vdots | \vdots | \vdots | \vdots |
| | k | X_{1k1} | ... | X_{1kn} | \bar{X}_{1k} | R_{1k} | S_{1k} |
| \vdots | \vdots | \vdots | \ddots | \vdots | \vdots | \vdots | |
| m | 1 | X_{m11} | ... | X_{m1n} | \bar{X}_{m1} | R_{m1} | S_{m1} |
| | \vdots | \vdots | ... | \vdots | \vdots | \vdots | \vdots |
| | k | X_{mk1} | ... | X_{mkn} | \bar{X}_{mk} | R_{mk} | S_{mk} |

da qualidade $X \sim N(\mu, \sigma^2)$, em que X_{ijl} representa a l -ésima observação, do j -ésimo subgrupo do i -ésimo grupo, em que $l = 1, 2, \dots, n$; $j = 1, 2, \dots, k$ e $i = 1, 2, \dots, m$, no qual m é o número de grupos, k o número de subgrupos e n o tamanho da amostra, bem como $\bar{X}_{ij} = \frac{\sum_{l=1}^n X_{ijl}}{n}$, $R_{ij} = \max(X_{ijl}) - \min(X_{ijl})$ e $S_{ij} = \sqrt{\frac{\sum_{l=1}^n (X_{ijl} - \bar{X}_{ij})^2}{n-1}}$.

O gráfico GG é uma alternativa mais prática em relação ao gráfico de controle tradicional quando existe estratificação (distinção entre os fluxos), pois fornece apenas um par de gráficos para todos os fluxos, sendo de grande importância decidir a forma na qual a estratificação será elaborada, evidenciando a necessidade de critérios rigorosos na definição e amostragem desses grupos. Determinada a formação dos subgrupos, em Ramos (2000) é dada a seguinte sequência para a construção do GG para a média e para variabilidade:

- Inicia-se a elaboração do gráfico através da coleta sistemática de amostras de cada fluxo, como se fossem construir gráficos para a média e o gráfico R para a variabilidade em cada fluxo, seguindo o gráfico de Shewhart;
- Calcula-se, em seguida, a média e a amplitude para cada amostra selecionada e calcula-se a média geral e a amplitude média;
- Determinam-se limites de controle para o gráfico por grupos para a média e o desvio-padrão da mesma maneira como é feito para o gráfico tradicional de Shewhart;
- Plotam-se a maior e menor média de cada grupo, no gráfico para a média, e a maior amplitude, no gráfico para o desvio-padrão.

Na seção seguinte são apresentados os limites de controle para o gráfico por grupos para a média e para a amplitude do processo.

2.3.1 GG para \bar{X}

Considerando X_{ijl} apresentado na Tabela 2.2, os limites para o gráfico por grupos de \bar{X} são dados por

$$LIC_{\bar{X}} = \bar{\bar{X}} - A_2\bar{R} \quad (2.1)$$

$$LC_{\bar{X}} = \bar{\bar{X}} \quad (2.2)$$

$$LSC_{\bar{X}} = \bar{\bar{X}} + A_2\bar{R}. \quad (2.3)$$

em que,

$$\bar{\bar{X}} = \sum_{i=1}^m \sum_{j=1}^k \frac{\bar{X}_{ij}}{mk}$$
$$\bar{R} = \sum_{i=1}^m \sum_{j=1}^k \frac{R_{ij}}{mk}$$

e os índices i e j e t das variáveis \bar{X} e R se referem , respectivamente, ao grupo e à amostra, seguindo os índices dados na Tabela 2.2. A constante A_2 é tabelada, pode ser encontrada em Montgomery (2000).

2.3.2 GG para R

Assumindo uma unidade amostral X_{ijl} , os limites de controle para o gráfico por grupos para amplitude R são dados por

$$LIC_R = \bar{R}D_3$$

$$LC_R = \bar{R}$$

$$LSC_R = \bar{R}D_4.$$

O cálculo dos limites para Gráfico por Grupos é análogo ao tradicional, diferindo-se essencialmente nas estimativas e nas constantes tabeladas (D_3 e D_4) que podem ser encontrada em Montgomery (2000), estas constantes se diferem pelo número de fluxos e grupos, no gráfico de Shewhart o parâmetro de interesse para o cálculo das contantes era o tamanho da amostra n . Como é possível notar na Tabela 2.2 temos que o conjunto de dados são divididos em Grupos, por isso o cálculo destas estatísticas são diferentes.

2.3.3 NMA para a avaliação do GG

O "Average Run Length" (ARL), também conhecido como "Número Médio de Amostras" (NMA), é uma métrica utilizada no controle estatístico de processos. Ela quantifica o número médio de observações ou amostras esperadas até a ocorrência de um evento. No contexto do CEP, o NMA é associado ao número média de amostras esperadas até a ocorrência de um alarme falso ou até que o GC detecte uma alteração na média ou na variância do processo (ou em ambas simultaneamente). Essa estatística é adotada na avaliação da eficiência do sistema de monitoramento. Desta forma, é associa-se o GC a um teste com as hipóteses:

$$\begin{cases} H_0 : \mu = \mu_0 \rightarrow \text{O processo está sob controle,} \\ H_1 : \mu \neq \mu_0 \rightarrow \text{O processo está fora de controle.} \end{cases}$$

O teste de hipóteses é um método estatístico que nos permite tomar decisões sobre uma afirmação (hipótese) com base em evidências amostrais. Com base no paradigma de teste de hipóteses aplicado ao CEP, a construção do indicador de desempenho do processo considera os erros do tipo I e II, referenciados como ETI e ETII, respectivamente. No contexto do CEP, o erro do tipo I ocorre quando se declara erroneamente que o processo está fora de controle, enquanto, na verdade, ele está sob controle, ou seja, uma medida se posiciona fora dos limites de controle dado que o processo está sob controle. Por outro lado, em Montgomery (2000) é dito que o erro do Tipo II ocorre quando erroneamente se considera que o processo está sob controle, quando, na verdade, ele não está. Isto é, os pontos flutuam entre os limites de controle, mas há uma causa especial no processo.

A compreensão desses erros é fundamental para avaliar a eficácia de um teste de hipóteses e do sistema de controle estatístico, proporcionando uma análise das consequências associadas a decisões equivocadas sobre a estabilidade do processo. Portanto, de forma probabilística, as probabilidades dos erros do tipo I e II para o gráfico \bar{X} , podem ser denotadas por,

$$P(ETI) = P(\bar{X} \notin [LIC, LSC] \mid \mu = \mu_0) \quad (2.4)$$

$$P(ETII) = P(\bar{X} \in [LIC, LSC] \mid \mu = \mu^*), \quad \forall \mu^* \neq \mu_0. \quad (2.5)$$

Devido à característica do NMA, que se refere ao número médio de amostras até a ocorrência de um evento de interesse, a distribuição geométrica é empregada no cálculo dessa medida. De acordo com Casella & Berger (2021), a distribuição geométrica é aplicada quando cada tentativa possui apenas dois resultados possíveis, como sucesso ou fracasso, e a probabilidade de sucesso é constante em cada ensaio. Isto é, seja Y uma variável aleatória com distribuição geométrica,

$$Y \sim G(p)$$

em que Y é o Número de ensaios até a ocorrência do primeiro sucesso ou fracasso, tem a seguinte distribuição de probabilidade

$$P(Y = y) = p(1 - p)^{y-1} I_{\{1,2,\dots\}}^{(y)}$$

com média e variância respectivamente dadas por

$$\begin{aligned} E(Y) &= \frac{1}{p}, \\ V(Y) &= \frac{1-p}{p^2}. \end{aligned} \tag{2.6}$$

Essa distribuição é particularmente útil em contextos nos quais estamos interessados em contar o número de tentativas necessárias até que um evento específico ocorra pela primeira vez. Nesse contexto, o NMA_1 se refere ao número de inspeções até a ocorrência do primeiro alarme falso. Supondo que o parâmetro p é a probabilidade de alarme falso ($P(ETI)$) em cada (ensaio) inspeção, logo

$$E(Y) = \frac{1}{P(ETI)} = NMA_1. \tag{2.7}$$

em que o NMA_1 representa o número médio de amostras até a ocorrência do primeiro alarme falso, que será um indicador utilizado neste estudo e fornece uma indicação do desempenho do GC. No caso, quanto maior o valor do NMA_1 , mais proteção o processo produtivo oferece contra alarmes falsos. Tal cenário é desejado, já que uma alta probabilidade de alarme falso indica que as medidas têm tendência de sair dos limites de controle, mas o processo permanece estável. Isto é, do ponto de vista gerencial, seria necessário intervir a produção para detectar uma causa especial que de fato não existe. Assim, o NMA_1

é uma métrica crucial para avaliar a performance de um sistema de controle estatístico de processos.

No contexto do GG, foco do estudo, o indicador é determinado levando em conta a quantidade de fluxos. Isso significa que as equações (2.4) e (2.5) são ajustadas para refletir a presença de k fluxos,

$$P(ETI_k) = P\left(\bigcup_{i=1}^k \bar{X}_i \notin [LIC, LSC] \mid \mu = \mu_0\right) \quad (2.8)$$

$$P(ETII_k) = P\left(\bigcap_{i=1}^k \bar{X}_i \in [LIC, LSC] \mid \mu = \mu^*\right). \quad (2.9)$$

Utilizando a propriedade de complementaridade das probabilidades, podemos expressar de forma geral, a partir da equação (2.4), que

$$P(ETI_k) = 1 - P(\bar{X} \in LC \mid \mu = \mu_0)^k \quad (2.10)$$

$$P(ETII_k) = 1 - P(\bar{X} \notin LC \mid \mu = \mu^*)^k. \quad (2.11)$$

Considerando a existência de k fluxos e utilizando a propriedade da probabilidade complementar, conforme apresentado em (2.10), a equação (2.8) utilizada no cálculo do *NMA* pode ser reescrita por

$$P(ETI_k) = 1 - P(LIC \leq \bar{X}_1 \leq LSC \cap \dots \cap LIC \leq \bar{X}_k \leq LSC) \quad (2.12)$$

Assumindo a satisfação da condição de independência entre os fluxos, as médias são independentes e identicamente distribuídas $\left(\bar{X}_i \sim N\left(\mu, \frac{\sigma^2}{n}\right)\right)$, tem-se ainda, segundo Mood *et al.* (1974), que

$$P(ETI_k) = 1 - P(LIC \leq \bar{X} \leq LSC)^k. \quad (2.13)$$

Perceba que para $k = 1$, o processo torna-se o GC Shewhart tradicional,

$$P(ETI_1) = 1 - P(LIC \leq \bar{X} \leq LSC).$$

Determinada a probabilidade do Erro do Tipo I, é necessária a aplicação de $P(ETI_k)$

na esperança da distribuição geométrica em 2.6,

$$NMA_1 = E(Y) = \frac{1}{P(ETI_k)} = \frac{1}{1 - P(LIC \leq \bar{X}_1 \leq LSC)^k} \quad (2.14)$$

fornecendo a interpretação de que, em média, serão necessárias $\frac{1}{P(ETI_k)}$ amostras até a ocorrência do primeiro alarme falso. É de se ressaltar o interesse em obter pequenos valores para $P(ETI_k)$, visto que aumentará o NMA_1 , o que é favorável à análise, já que quanto maior o valor do indicador, mais amostras serão necessárias para se ter o primeiro alarme falso. No GC tradicional de Shewhart, considerando $L = 3$, esse valor é de aproximadamente um alarme falso a cada 370 inspeções. Nesse estudo, o Número Médio de Amostras até a ocorrência do primeiro alarme falso será representado por NMA_1 .

Outra forma de analisar o desempenho do GC está associado à detecção rápida de um deslocamento na média ou na variância do processo, que utiliza a probabilidade do erro do tipo II. Isto é, assume-se um deslocamento nos parâmetros $\mu^* = \mu_0 + \delta\sigma_{\bar{X}}$, $\mu^* \neq \mu_0$ e $\sigma^{*2} = \lambda^2\sigma_0^2$, respectivamente, de forma individual (deslocamento apenas na média ou na variância) e conjunta (na média e na variância ao mesmo tempo). Em outras palavras, pode se calcular o Número Médio de Amostras até a detecção de uma perturbação no processo, denominado NMA_2 , que se baseia nas probabilidades

$$P(ETII_{21}) = P(\bar{X} \notin (LIC; LSC) \mid \mu = \mu^*) = 1 - P(LIC < \bar{X} < LSC \mid \mu = \mu^*) \quad (2.15)$$

$$P(ETII_{22}) = P(\bar{X} \notin (LIC; LSC) \mid \sigma = \sigma^{*2}) = 1 - P(LIC < \bar{X} < LSC \mid \sigma = \sigma^{*2}). \quad (2.16)$$

Desenvolvendo (2.15), tem-se

$$P(ETII_{21}) = 1 - \left[P\left(Z < \frac{\mu_0 + \frac{L\sigma_X}{\sqrt{n}} - \mu_0 - \delta\sigma_X}{\frac{\sigma_X}{\sqrt{n}}}\right) - P\left(Z < \frac{\mu_0 - \frac{L\sigma_X}{\sqrt{n}} - \mu_0 - \delta\sigma_X}{\frac{\sigma_X}{\sqrt{n}}}\right) \right],$$

logo,

$$P(ETII_{21}) = 1 - [\Phi(-\delta\sqrt{n} + L) - \Phi(-\delta\sqrt{n} - L)]^k. \quad (2.17)$$

Ocorrendo a perturbação na variância, e a partir de (2.16),

$$P(ETII_{22}) = 1 - \left[P\left(Z < \frac{\mu + \frac{L\sigma_X}{\sqrt{n}} - \mu}{\frac{\sigma^*}{\sqrt{n}}}\right) - P\left(Z < \frac{\mu - \frac{L\sigma_X}{\sqrt{n}} - \mu}{\frac{\sigma^*}{\sqrt{n}}}\right) \right]^k,$$

assim,

$$P(ETII_{22}) = 1 - \left[\Phi\left(\frac{L}{\lambda}\right) - \Phi\left(-\frac{L}{\lambda}\right) \right]^k. \quad (2.18)$$

Por fim, o último caso consiste em uma variação conjunta na média e na variância do processo. Agora, tem-se que,

$$P(ETII_{23}) = 1 - (P(LIC_{\bar{X}} < \bar{X} < LSC_{\bar{X}} | \mu^*, \sigma^*))^k \quad (2.19)$$

$$= 1 - P\left(\frac{\mu - L\sigma_{\bar{X}} - \mu - \delta\sigma}{\lambda\sigma_{\bar{X}}} < Z < \frac{\mu + L\sigma_{\bar{X}} - \mu - \delta\sigma}{\lambda\sigma_{\bar{X}}}\right)^k \quad (2.20)$$

$$= 1 - P\left(-\frac{L + \delta\sqrt{n}}{\lambda} < Z < \frac{L - \delta\sqrt{n}}{\lambda}\right)^k \quad (2.21)$$

assim, a probabilidade do Erro do Tipo II que contempla as duas variações simultâneas $P(ETII_{23})$ é dada por,

$$P(ETII_{23}) = 1 - \left(\Phi\left(\frac{L - \delta\sqrt{n}}{\lambda}\right) - \Phi\left(-\frac{L + \delta\sqrt{n}}{\lambda}\right) \right)^k. \quad (2.22)$$

Assim, de forma análoga ao $P(ETI_k)$, $NMA_2 = \frac{1}{P(ETII_2)^k}$ é função da largura do intervalo L , do tamanho da amostra, da variação da média em termos do desvio padrão δ e do número de fluxos k . Para cada tipo de deslocamento será atribuído um tipo de indicador, seguindo a mesma nomenclatura utilizada em Oliveira (2013). NMA_{21} representa o Número Médio de Amostras até a detecção de um deslocamento na média do processo, NMA_{22} está relacionado ao deslocamento em σ^2 e NMA_{23} está associado a uma perturbação conjunta em μ e σ^2 . É importante lembrar que, inicialmente, o processo é considerado sob controle estatístico do processo.

Tanto NMA_1 quanto NMA_2 , são apresentados para gráficos de controle independentes, descartando qualquer forma de relação entre as observações. O cálculo teórico do indicador exige a satisfação das duas condições de aplicação do GC, caso contrário, as quantidades baseadas em (2.12) serão alteradas, pois a presença de uma estrutura de dependência impede o uso da propriedade da probabilidade de eventos independentes, que possibilitou a modificação de (2.12) e de (2.13). No caso deste estudo, assegura-se a condição de normalidade, no entanto, a de independência é violada, já que a investigação se dá em processos com fluxos correlacionados. Esse é um dos motivos pela utilização

de técnicas de simulação para alcançar os objetivos propostos, os demais motivos serão abordados nas sessões futuras.

2.3.4 Análise de Correlação

De acordo com Bussab & Morettin (2017) a análise de correlação é uma técnica estatística amplamente empregada para mensurar a relação entre duas ou mais variáveis, sendo de grande relevância no cenário industrial para a identificação de padrões e tendências nos dados. O coeficiente de correlação, que varia de -1 a 1 , é utilizado como métrica, no qual valores próximos de -1 indicam uma correlação negativa forte, enquanto valores próximos de 1 apontam para uma correlação positiva forte. Um valor próximo de 0 sugere uma correlação fraca ou inexistente.

Em James *et al.* (2011) é dito que essa abordagem analítica é importante para compreender variáveis complexas que não podem ser manipuladas em pesquisas experimentais. A análise de correlação proporciona percepções estatísticas sobre como as variáveis se relacionam, sendo aplicável em diversas áreas. Em Levine *et al.* (2014) destaca-se a utilidade dessa análise na previsão do comportamento de variáveis com base em dados históricos, essa capacidade preditiva possibilita antecipar problemas e implementar medidas preventivas, contribuindo significativamente para evitar falhas ou acidentes.

Os gráficos por grupos representam uma variação dos gráficos de controle tradicionais propostos por Shewhart, sendo adaptados para monitorar fluxos específicos de produção. Em situações em que ocorre autocorrelação entre as médias dos subgrupos, violando a condição de independência dos dados, uma abordagem sugerida por Costa *et al.* (2005) envolve a utilização do coeficiente de autocorrelação amostral de grau $L(r_L)$:

$$r_L = \frac{\sum_{i=1}^{n-L} (\bar{X}_i - \bar{\bar{X}})(\bar{X}_{i+L} - \bar{\bar{X}})}{\sum_{i=1}^n (\bar{X}_i - \bar{\bar{X}})^2}.$$

Esse coeficiente é empregado para avaliar o grau de interferência no monitoramento do processo. No entanto, torna-se essencial determinar a significância desse valor para influenciar a análise. Ramos (2000) propõe um teste de hipóteses, com H_0 indicando a ausência de autocorrelação no processo e H_1 indicando a presença de autocorrelação. A estatística do teste é dada por:

$$e = \frac{2}{\sqrt{L}},$$

em que L é o "lag", representa o intervalo de tempo entre observações no cálculo de r_L . Se o coeficiente de correlação exceder $\pm e$, a autocorrelação é considerada significativa com 95% de confiança. Vale ressaltar que Oliveira (2013) destaca a importância desse coeficiente ao abordar a questão do desvio elevado (presença de muitos "outliers").

A consideração conjunta do coeficiente de autocorrelação e do teste proposto proporciona uma abordagem abrangente para avaliar e lidar com a autocorrelação no processo de monitoramento. Essas técnicas são cruciais para garantir a validade e a confiabilidade dos resultados obtidos nos Gráficos por Grupos, especialmente quando há indícios de autocorrelação entre as médias dos subgrupos, o que pode afetar a interpretação correta dos padrões e tendências identificados nos dados de controle.

Capítulo 3

Métodos e simulação

Este capítulo oferece uma justificativa clara para a escolha da metodologia de simulação, descrevendo de forma detalhada o método aplicado, a execução do algoritmo e a plataforma utilizada no desenvolvimento.

3.1 Metodologia de Pesquisa

Esta seção tem como objetivo estabelecer os principais conceitos, colocando o pesquisador em contato com as contribuições já existentes sobre o tema. A análise dos cenários que envolvem a correlação entre os fluxos de processos paralelos, monitorados pelos gráficos por grupos, constitui a essência deste estudo. A construção do referencial teórico baseou-se em fontes diversas, como livros, material de cursos e artigos de periódicos, oferecendo uma base sólida para a compreensão dos conceitos fundamentais necessários à elaboração deste trabalho.

Os métodos propostos serão calculados e comparados através de simulações e informações disponíveis na literatura consultada. A obtenção dos resultados envolverá o cálculo de indicadores de desempenho para a ferramenta, variando parâmetros como o número de fluxos, tamanho da amostra, média e variância do processo. Além disso, este estudo visa investigar o comportamento da densidade do máximo e do mínimo em cada cenário analisado, proporcionando uma compreensão mais profunda das nuances do comportamento dos processos paralelos.

Essa abordagem expandida contribuirá para uma compreensão abrangente não apenas do desempenho global da ferramenta, mas também de como as características específicas

de cada cenário influenciam a distribuição dos valores máximos e mínimos. A inserção de perturbações nos conjuntos de dados e a comparação entre o gráfico por grupos e o método tradicional de Shewhart finalizarão a análise, proporcionando a compreensão e otimização de processos paralelos na presença de correlação.

3.2 Complexidade Teórica

Vasconcelos (2009) conduziu uma avaliação probabilística no NMA do GG e o comparou com o método tradicional de Shewhart. Posteriormente, analisou a eficiência entre esses métodos diante de alterações na média e variabilidade do processo. O autor assegurou o cumprimento das condições necessárias para a aplicação do instrumento, considerando a independência e normalidade dos dados. Ele baseou seu estudo em cenários que obedecem às condições de independência, aplicando as leis clássicas da probabilidade.

Entretanto, é relevante observar que Oliveira (2013) introduziu uma estrutura de correlação entre os fluxos, violando assim a condição de independência dos dados. Além disso, ele desconsiderou a condição de normalidade ao desenvolver o estudo por meio das distribuições do mínimo e do máximo, as quais foram tratadas como desconhecidas e inexatas.

Diferentemente, em nossa pesquisa, destacamos o conhecimento das distribuições do mínimo e do máximo, utilizando o método de Arellano. Essa abordagem robusta proporciona uma base sólida para lidar com as distribuições, garantindo uma maior precisão e confiabilidade na análise estatística do NMA em relação ao GG.

3.3 Método Arellano para a estimação da distribuição do máximo e do mínimo

3.3.1 Estatística de ordem

A estatística de ordem é uma estatística que permite estudar a distribuição de um conjunto de dados ordenados. Segundo Casella & Berger (2021), a estatística de ordem é definida como uma função dos dados que leva em conta sua ordem, a posição que cada um dos dados ocupa quando eles são ordenados. Matematicamente, a estatística de

ordem é definida como uma função que mapeia o conjunto de dados ordenados em um número real. Essa função pode ser representada por $X_{(n)}$, em que \mathbf{X} é o conjunto de dados ordenados e n é o tamanho da amostra. O valor de $X_{(n)}$ depende da posição do valor na ordem dos dados e, portanto, é uma função não paramétrica que não requer nenhuma suposição sobre a distribuição dos dados.

As estatísticas de ordem contribuem significativamente para a compreensão da variabilidade entre os grupos e dentro de cada grupo. Ao ordenar as observações, é possível discernir tendências, identificar "outliers" e avaliar a dispersão dos dados em cada conjunto. Essa abordagem proporciona uma representação mais abrangente da distribuição, permitindo uma interpretação mais refinada das mudanças no desempenho do processo.

Além disso, as estatísticas de ordem são fundamentais para avaliar a consistência das medições entre grupos, possibilitando uma comparação mais precisa da performance ao longo do tempo. Isso aprimora a sensibilidade do gráfico por grupos, tornando-o uma ferramenta mais eficaz na detecção de desvios significativos e na monitorização contínua da estabilidade do processo.

Existem várias estatísticas de ordem que podem ser utilizadas para analisar um conjunto de dados. Algumas das mais comuns incluem a estatística de ordem mínima, que é o menor valor do conjunto de dados, e a estatística de ordem máxima, que é o maior valor do conjunto de dados. Além disso, existem outras estatísticas de ordem como a mediana, os quartis e os percentis, que são frequentemente utilizados em análises exploratórias e inferenciais.

Por exemplo, a estatística de ordem do mínimo é definida como:

$$\text{Min}(\mathbf{X}) = X_{(1)}, \quad (3.1)$$

em que $X_{(1)}$ é o menor valor no conjunto de dados ordenados \mathbf{X} . De forma semelhante, a estatística de ordem do máximo é definida como

$$\text{Max}(\mathbf{X}) = X_{(n)} \quad (3.2)$$

em que $X_{(n)}$ é o maior valor no conjunto de dados ordenados \mathbf{X} . A mediana é uma estatística de ordem comum e pode ser definida como

$$\text{mediana}(\mathbf{X}) = \begin{cases} X_{(\frac{n}{2})} & \text{se } n \text{ é ímpar} \\ \frac{1}{2}(X_{(\frac{n}{2})} + X_{(\frac{n}{2}+1)}) & \text{se } n \text{ é par} \end{cases} \quad (3.3)$$

em que n é o tamanho amostral em cada fluxo. Os quartis e percentis também são estatísticas de ordem comuns e podem ser definidos de forma semelhante.

A estatística de ordem é uma ferramenta muito útil em diversas áreas da estatística, como na análise de dados de sobrevivência, na análise de experimentos, análise de séries temporais, entre outras. Ela permite estudar a distribuição dos dados de forma não paramétrica e é particularmente útil quando não há informações sobre a distribuição dos dados ou quando as suposições de normalidade não são válidas. É definida como uma função que mapeia o conjunto de dados ordenados em um número real. Através dela, é possível estudar a distribuição dos dados de forma não paramétrica e obter informações valiosas sobre o conjunto de dados em análise, como a distribuição do máximo e do mínimo. A distribuição do máximo de uma estatística de ordem de acordo com Casella & Berger (2021), com função densidade de probabilidade f e distribuição acumulada F é dada por

$$f(X_{(n)}) = nf(x)F(x)^{n-1}.$$

e a distribuição do mínimo é dada por

$$f(X_{(1)}) = nf(x)(1 - F(x))^{n-1}.$$

com n sendo o tamanho da amostra.

3.3.2 Método Arellano para a estimação da distribuição do máximo

O método proposto em Arellano-Valle & Genton (2008) visa calcular a função de densidade de probabilidade exata do valor máximo de um conjunto de variáveis aleatórias dependentes, todas seguindo uma distribuição normal. A abordagem baseia-se na exploração das relações intrínsecas entre a distribuição do valor máximo e a **família de distribuições FUS** (*Fundamental Skew Symmetric*). Seja $\mathbf{X} = (\mathbf{Y}|\mathbf{Z})$, em que $\mathbf{Y} \in \mathbb{R}^k$ é um vetor aleatório com função de densidade $f_{\mathbf{Y}}$ e $\mathbf{Z} \in \mathbb{R}^m$ é um vetor aleatório, com $\mathbf{Z} \geq 0$.

Para poder aplicar este método é importante notar que a densidade do máximo:

$$f_{x_{(n)}}(x) = \sum_{i=1}^n f_{X_i}(x) F_{\mathbf{X}_{-i}|X_i=x}(x \mathbf{1}_{n-1}), x \in \mathbb{R}^n, \quad (3.4)$$

em que f_{X_i} é a densidade marginal de X_i e $F_{\mathbf{X}_{-i}|X_i=x}$ é a distribuição condicional acumulada de $\mathbf{X}_{-i} = (X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n)^T$ dado $X_i = x$, pode ser reescrita no formato de uma distribuição FUS,

$$f_x(x) = K_m^{-1} f_Y(x) Q_m(X), \quad x \in \mathbb{R}^n, \quad (3.5)$$

em que $Q_m(X) = P(\mathbf{Z} \geq 0 | \mathbf{Y} = \mathbf{x})$, $K_m = E(Q_m(\mathbf{Y})) = P(\mathbf{Z} \geq 0)$ e f_Y é a densidade de uma distribuição simétrica ($f_Y(-y) = f_Y(y)$) para todo $y \in \mathbb{R}^k$. Considerando uma família de distribuições com contorno elíptico. Seguindo Fang (2018), sendo $\mathbf{X} \in \mathbb{R}^n$ pertencente a família de distribuições com contorno elíptico, com vetor de locação $\boldsymbol{\mu}$, e matriz de escala $\boldsymbol{\Sigma}$ e função de geradora de densidade $h^{(n)}$, a probabilidade dessa função de densidade é dada pela seguinte forma,

$$f_n(x; \boldsymbol{\mu}; \boldsymbol{\Sigma}; h^{(n)}) = |\boldsymbol{\Sigma}|^{-1/2} h^{(n)}[(x - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1} (x - \boldsymbol{\mu})], x \in \mathbb{R}^n. \quad (3.6)$$

Considerando todas as condições dadas nessa seção e aplicando para o caso particular da normal multivariada, é descrito em Arellano-Valle & Genton (2008) a seguinte expressão para determinar o valor máximo em uma distribuição multivariada dependente, a qual segue uma distribuição normal.

$$f_{X_{(n)}}(x) = n \phi_1(x; \boldsymbol{\mu}_i, \sigma^2) \Phi_{n-1}(\sqrt{1-rz} \mathbf{1}_{n-1}; \mathbf{0}, \mathbf{I}_{n-1} + r \mathbf{1}_{n-1} \mathbf{1}_{n-1}^\top), x \in \mathbb{R}^n.$$

em que, n é o tamanho da amostra, $\boldsymbol{\mu}$ é vetor de médias, $\boldsymbol{\Sigma}$ matriz de variância-covariância, $\Phi(\cdot)$ função de distribuição acumulada da distribuição normal padrão e $\phi(\cdot)$ é função de densidade de probabilidade da distribuição normal padrão.

3.3.3 Método Arellano para a estimação da distribuição do mínimo

Em contextos nos quais as variáveis de interesse seguem uma distribuição normal multivariada, a consideração da simetria dessas variáveis é de grande importância. A simetria implica que as funções de probabilidade das variáveis e de suas contrapartes transformadas devem ser semelhantes.

Neste contexto, uma transformação comum que preserva a estrutura da normal multivariada é a transformação simétrica, em que uma variável \mathbf{X}_n tem mesma distribuição que \mathbf{X}_1 . Esta transformação é fundamentada na propriedade de simetria da distribuição normal multivariada, que assume que os valores à esquerda e à direita da média são igualmente prováveis. A transformação $f(\mathbf{X}) = f(-\mathbf{X})$ é uma maneira direta e intuitiva de garantir que a distribuição original e a transformada permaneçam simétricas em relação à média.

Essa abordagem é discutida detalhadamente em Fang (2018). Nesta obra, são exploradas as propriedades das distribuições multivariadas simétricas, fornecendo uma base teórica sólida para a aplicação da transformação simétrica em variáveis de uma normal multivariada.

Na Figura 3.1 temos uma simulação com uma amostra de tamanho 1.000 e na Figura 3.2 uma de tamanho 1.000.000, a simulação é feita a partir de uma normal Multivariada, com $\boldsymbol{\mu} = 0$, $\boldsymbol{\Sigma} = (1 - r)\mathbf{I}_5 + r\mathbf{1}_5\mathbf{1}_5^\top$ com a densidade do máximo sendo representada pela cor preta e a densidade do mínimo representado pela cor vermelha, com r é variando a cada curva, para demonstrar esta simetria:

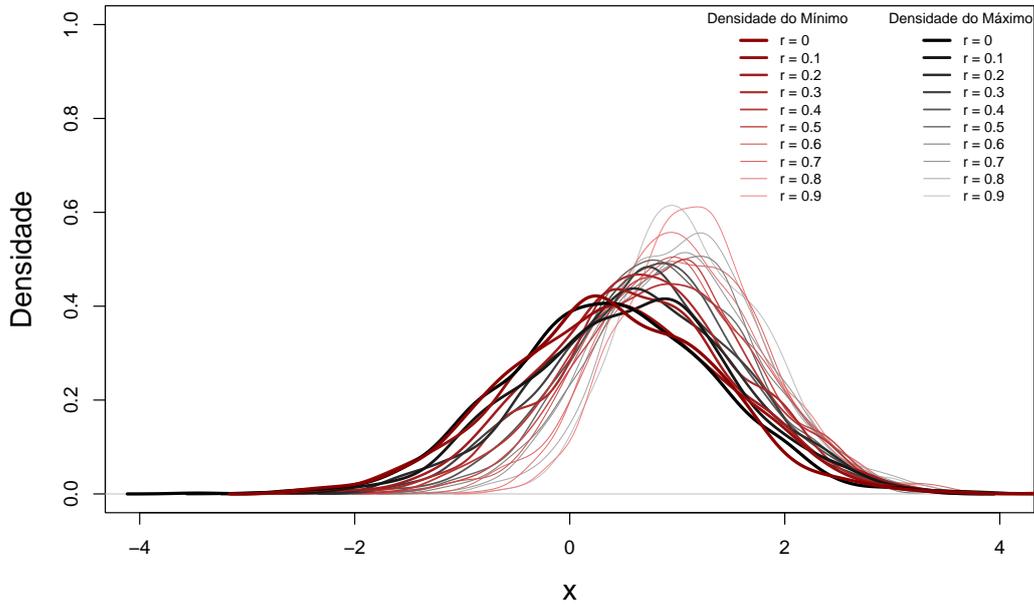


Figura 3.1: Densidade do máximo e do mínimo para uma amostra de tamanho 1.000.

Observando as figuras 3.1 e 3.2 é possível notar uma diferença significativa na proximidade entre as curvas vermelhas e pretas em relação ao tamanho da amostra, a densidade do máximo e do mínimo se sobrepõem em cada curva, cada curva representa uma mudança no valor de r , que variava no intervalo de 0,1 e 0,9. Para o conjunto de dados com 1.000 observações, as curvas vermelhas e pretas estavam próximas, mas ainda apresentavam alguma separação, indicando uma certa variabilidade nos valores máximos e mínimos. No entanto, para o conjunto de dados com 1.000.000 observações, as curvas estavam praticamente sobrepostas, sugerindo uma convergência quase perfeita entre os valores máximos e mínimos.

A convergência das curvas vermelhas e pretas em amostras maiores destaca a importância do tamanho da amostra na precisão dos resultados estatísticos. Amostras maiores tendem a oferecer uma representação mais fiel da população, permitindo uma estimativa mais precisa dos parâmetros estatísticos, incluindo os valores máximos e mínimos. Isso implica que, à medida que o tamanho da amostra aumenta, a densidade dos valores máximos e mínimos se aproxima, indicando uma maior confiabilidade nas conclusões tiradas com base nesses extremos.

Os resultados obtidos nas figuras 3.1 e 3.2 mostraram que, à medida que a covariância entre as variáveis aumentava, a densidade dos valores máximos e mínimos se deslocava

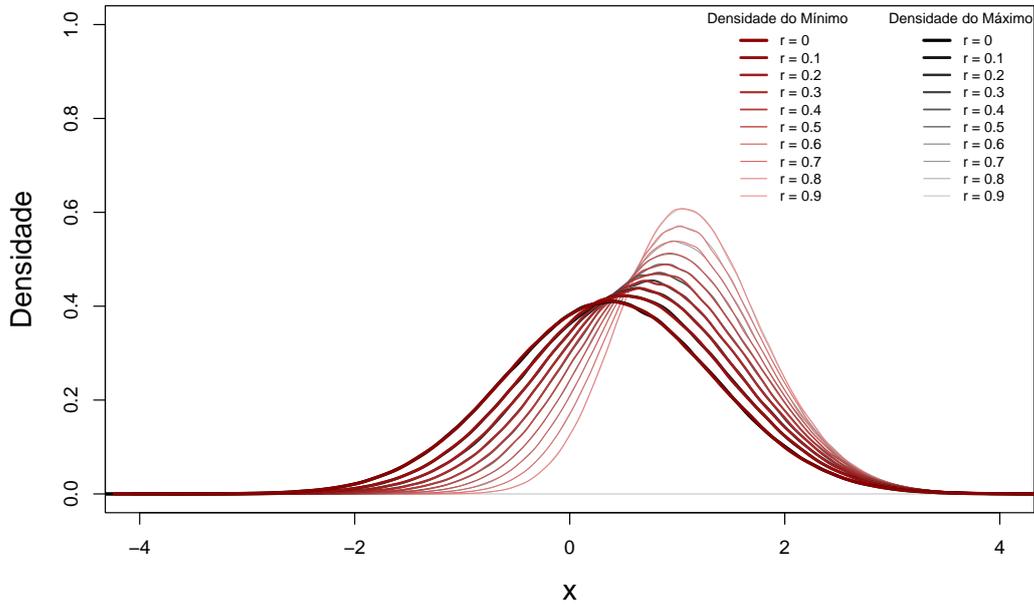


Figura 3.2: Densidade do máximo e do mínimo para uma amostra de tamanho 1.000.000.

para a direita. Isso significa que, quando as variáveis estão mais fortemente correlacionadas, os valores extremos têm sua média deslocada para direita.

3.4 Simulação

Neste estudo, buscamos criar um ambiente que simule o contexto de aplicação do CEP, levando em consideração a presença de correlação entre os fluxos. Em particular, estamos replicando cenários previamente examinados por Oliveira (2013) com o intuito de realizar uma comparação direta.

Dentro de um contexto de simulação, é crucial compreender a diferenciação entre dois tipos principais: o determinístico e o probabilístico. De acordo com Law & Kelton (2000), a abordagem determinística considera a obtenção de dados "reais" sem associar probabilidades diretamente às observações selecionadas. Essa metodologia busca replicar o sistema sem incorporar explicitamente elementos de aleatoriedade nas simulações. Por outro lado, Banks & Carson (1984) destacam o enfoque probabilístico na simulação, que, conforme descrito pelos autores, integra o comportamento probabilístico no relacionamento interno do sistema. Essa abordagem visa refletir a natureza inerentemente probabilística das variáveis do sistema, introduzindo elementos estocásticos nas simulações

para representar adequadamente a incerteza associada aos processos simulados.

De acordo com Banks & Carson (1984); Law & Kelton (2000) essa dualidade entre determinístico e probabilístico proporciona uma base conceitual essencial para os estudos de simulação, permitindo que os modeladores escolham a abordagem mais adequada com base na natureza do sistema em análise e nos objetivos específicos da simulação. Considerando os cenários previamente simulados por Oliveira (2013), desenvolvemos um algoritmo no software R capaz de reproduzir um ambiente que reflita a aplicação da ferramenta em um contexto específico. Esse algoritmo oferece a possibilidade de configurar uma situação controlada por meio de parâmetros incorporados à função desenvolvida, tais como o número de fluxos, o tamanho das amostras e a correlação entre eles. Além disso, o algoritmo realiza a virtualização de deslocamentos na média e na variabilidade do processo. O objetivo é realizar uma comparação detalhada entre o modelo do gráfico por grupos e o gráfico de controle tradicional de Shewhart, proporcionando uma análise abrangente dos diferentes cenários simulados.

3.5 O algoritmo desenvolvido

Em Fishman (2013) é dito que o processo de Monte Carlo envolve a repetição de experimentos usando números aleatórios para representar diferentes cenários possíveis. Esses resultados são então combinados para fornecer uma estimativa estatística das propriedades do sistema em análise.

Conforme discutido em Rubinstein & Kroese (2016) os benefícios de Monte Carlo incluem a capacidade de lidar com incertezas, modelar sistemas complexos e avaliar riscos. No entanto, é importante notar que a eficácia do método depende da precisão e representatividade dos modelos utilizados, além de uma cuidadosa consideração dos parâmetros e das distribuições de probabilidade associadas.

O método apresentado na Seção 3.3.2 surge como uma alternativa viável devido à sua capacidade de obter a densidade exata do máximo. Isso permite o cálculo direto do *NMA* utilizando essa densidade, eliminando a necessidade de realizar uma simulação de Monte Carlo. O GG utilizando este método será chamado de GG A (Gráfico por Grupos - Arellano).

O procedimento de geração das quantidades citadas, no caso de fluxos dependentes,

consiste na execução de uma função a partir de uma sequência de passos:

- I. A função calcula o *NMA* do processo e tem como argumentos, o número de fluxos (k), o tamanho da amostra (n), a correlação entre os fluxos (r), o fator de largura utilizado (L), o deslocamento na média do processo (δ) e o deslocamento na variância (λ);
- II. Uma estimação de μ e σ^2 é realizada a partir de distribuições normais com os parâmetros fornecidos no passo anterior.
- III. Com os valores de μ e σ^2 , obtém-se a densidade do máximo de uma variável normal utilizando o método de Arellano, como explicado na Seção 3.3.2.
- IV. Os limites de controle são determinados seguindo a Equação 2.1.
- V. O valor de $P(ETI)$, que representa a probabilidade de o máximo observado $X_{(n)}$ estar fora dos limites de controle (LIC, LSC), é calculado conforme a Equação 2.4. De maneira semelhante, $P(ETII)$ pode ser calculado seguindo a Equação 2.5, que representa a probabilidade de $X_{(n)}$ estar dentro dos limites de controle (LIC, LSC), quando não deveria.
- VI. O *NMA* é calculado como o inverso de $P(ETI)$ ou $P(ETII)$, conforme a Equação 2.7.

Na Figura 3.3 temos um fluxograma para melhor compreensão dos passos explicados nesta seção.

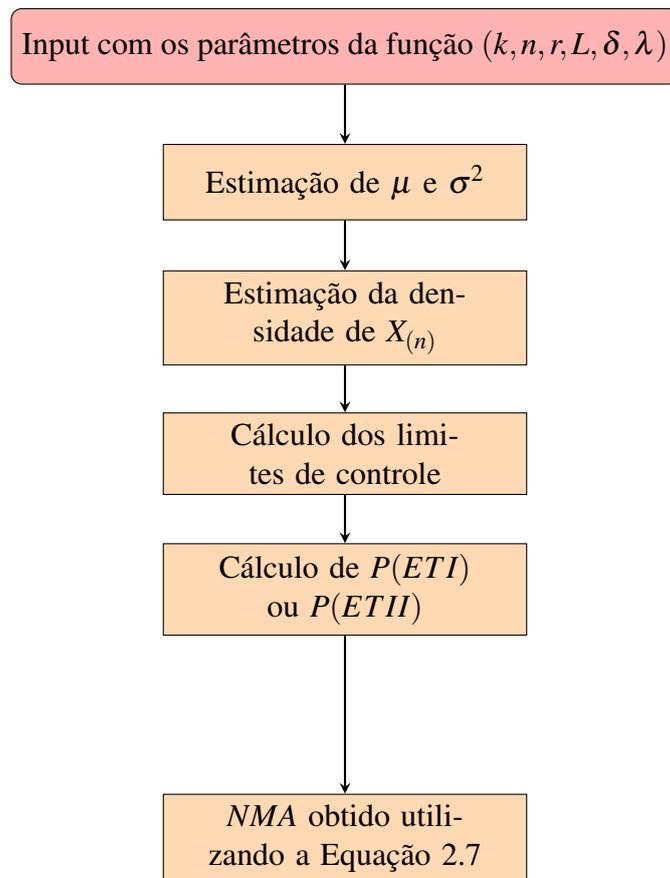


Figura 3.3: Fluxograma da simulação

Capítulo 4

Um estudo de desempenho entre GG, GG A e o gráfico tradicional de Shewhart

Neste capítulo, discutiremos o desempenho e a eficiência do gráfico de médias em resposta a uma perturbação que afeta tanto a média quanto a variabilidade do processo, tanto de forma individual quanto conjunta. Isso será explorado especialmente em situações em que existe correlação entre os fluxos de dados.

4.1 Desempenho do GG A

O algoritmo elaborado tem o intuito de analisar o comportamento do NMA_1 em processos paralelos com fluxos correlacionados, que será comparado com três métodos distintos: o exato, obtido de forma probabilística, e o simulado através das distribuições do mínimo e do máximo, ambos adquiridos respeitando a condição de independência e um simulado que não respeita esta condição de independência. Essa comparação é realizada visando averiguar o comportamento do NMA_1 para a taxa de alarmes falsos com fluxos dependentes em relação NMA_1 com fluxos independentes.

Este trabalho desenvolve um procedimento análogo ao executado na obtenção dos resultados com fluxos independentes, no entanto, agrega um diferencial, que é inserção da correlação como argumento da função e não obrigatoriedade da condição de independência. Vasconcelos (2009) e Oliveira (2013) calcularam o NMA_1 considerando

$L = 3$ para processos com $X \sim N(100, 10^2)$ e com o número de fluxos variando em $k = 1, 2, 3, \dots, 10, 15, 20$, sendo, portanto, estes valores usados na construção do algoritmo para viabilizar a comparação entre os métodos.

Para efeito de comparação, utilizou-se o algoritmo para elaborar a Tabela 4.1, que contém os resultados desenvolvidos de forma probabilística apresentados por Vasconcelos (2009), e os obtidos pelo código desenvolvido no capítulo 3, na presença de correlação entre os fluxos, bem como a média e a variância são iguais a 100. Essa tabela mostra a confiabilidade do modelo proposto calculando o indicador em cenários distintos, sendo no caso, a correlação alterada para visualizar o comportamento do NMA_1 com um nível de dependência acompanhado gradualmente ao longo do crescimento da correlação r .

Tabela 4.1: NMA_1 para fluxos independentes (exato e simulado) e para fluxos dependentes com $\mu = \sigma^2 = 100$, obtido pelo algoritmo descrito no Capítulo 3 variando a correlação.

| k | Fluxos independentes | | Fluxos dependentes (r) | | | | | |
|----|----------------------|----------|----------------------------|--------|--------|--------|--------|--------|
| | Exato | Simulado | 0,1 | 0,2 | 0,3 | 0,5 | 0,7 | 0,9 |
| 1 | 370,40 | 384,57 | - | - | - | - | - | - |
| 2 | 185,40 | 202,32 | 125,27 | 194,44 | 188,57 | 181,43 | 148,15 | 181,08 |
| 3 | 123,80 | 120,94 | 82,98 | 79,14 | 110,12 | 109,18 | 126,11 | 121,05 |
| 4 | 93,00 | 87,00 | 53,35 | 41,17 | 41,86 | 32,15 | 34,32 | 81,45 |
| 5 | 74,50 | 74,03 | 38,29 | 37,75 | 45,56 | 29,08 | 56,50 | 68,04 |
| 6 | 62,10 | 62,32 | 28,08 | 13,64 | 41,74 | 31,77 | 60,41 | 53,56 |
| 7 | 53,30 | 52,72 | 29,70 | 13,24 | 25,07 | 34,70 | 33,51 | 42,71 |
| 8 | 46,70 | 45,98 | 17,65 | 22,42 | 11,07 | 16,49 | 32,98 | 44,34 |
| 9 | 41,60 | 34,95 | 10,23 | 7,65 | 7,34 | 10,33 | 16,53 | 30,51 |
| 10 | 37,50 | 37,92 | 3,54 | 6,98 | 15,64 | 9,39 | 18,61 | 33,41 |
| 15 | 25,20 | 27,15 | 2,52 | 1,67 | 4,52 | 10,03 | 1,72 | 11,81 |
| 20 | 19,00 | 19,02 | 0,70 | 0,68 | 1,78 | 1,52 | 3,40 | 3,79 |

Considerando o contexto do Gráfico de controle tradicional de Shewhart (GCS) tradicional, a presença de correlação geralmente resulta em um aumento de alarmes falsos, devido à violação da condição de independência. No entanto, no estudo em questão, a correlação está presente entre os fluxos, eliminando a existência de autocorrelação dentro de cada fluxo. Isso levanta questionamentos sobre a similaridade no comportamento entre os dois tipos de correlação (entre fluxos e dentro dos fluxos).

Ao analisar os valores contidos na Tabela 4.1, observa-se inicialmente a ausência do primeiro NMA_1 simulado para fluxos dependentes. Isso se justifica pelo fato de que a correlação é uma medida de natureza bidimensional, tornando inviável a análise com apenas um fluxo. Além disso, ao fixar r e variar k , destaca-se a redução do NMA_1 nas situações com dependência em comparação com as situações de independência.

A Tabela 4.1 também fornece os valores correspondentes ao NMA_1 , variando a correlação r nos seguintes níveis: 0,1;0,2;0,3;0,5;0,7 e 0,9. Esses valores originaram o gráfico representado na Figura 4.1, com cada curva representando uma mudança na correção, a curva com maior densidade tem $r = 0,9$, conforme a correlação diminui a densidade também diminui e a calda da densidade alonga. O modelo desenvolvido apresentou um comportamento divergente para correlações acima de 0,50. Observa-se que, para um elevado grau de dependência positiva, os valores gerados são significativamente superiores aos demais. Essa disparidade pode ser explicada pela geração dos valores iniciais, se esses valores estiverem posicionados próximos à linha central e possuírem uma alta correlação, pode ser necessário um número considerável de amostras antes de alcançar a estabilidade de controle.

Ao compararmos com as simulações de Oliveira (2013) a partir de uma correlação superior a 0,50 ou envolvendo pelo menos 10 fluxos, o método que não incorpora a condição de independência demonstra um desempenho superior.

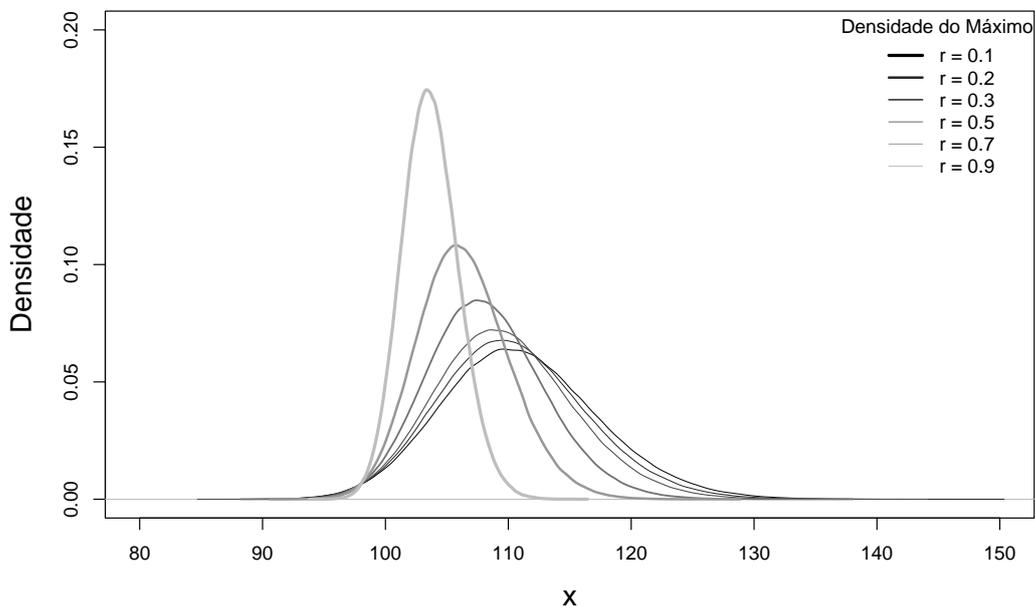


Figura 4.1: NMA_1 para fluxos independentes (exato e simulado) e para fluxos dependentes com $\mu = \sigma^2 = 100$.

É fundamental ressaltar a importância da geração inicial dos valores para a entrada no laço interno do algoritmo, onde ocorre o procedimento responsável pelo cálculo do

NMA_1 . Se o valor inicial gerado for significativamente alto, próximo ao Limite Superior de Controle (LSC), e a correlação for mantida em 0,90, poucas amostras serão necessárias para desencadear um alarme falso. Em contrapartida, mesmo com uma correlação de 0,90, se o valor estiver próximo ao Limite Inferior de Controle (LIC), serão necessárias várias amostras até que ocorra um alarme falso.

A Figura 4.1 foi feita utilizando a mesma simulação utilizada no Capítulo 3 para mostrar a distribuição do máximo, com outros parâmetros, $\mu = 100$ e $\sigma = 10$. A figura ilustra claramente que, com uma alta correlação, o comportamento assemelha-se ao modelo com independência. À medida que o valor de r aumenta, observa-se uma convergência lenta, similar aos demais modelos com correlação inferior. Esse fenômeno pode ser explicado pelo seguinte raciocínio: ao considerar uma pequena quantidade de fluxos, se a correlação for elevada, o comportamento entre eles será essencialmente o mesmo, aproximando-se do modelo tradicional.

Realizou-se uma modificação na média e na variância do processo ($\mu = 1 < \sigma^2 = 100$) para aprofundar na análise. Os resultados, apresentados na Tabela 4.2 e na Figura 4.2 temos o segundo cenário com $\mu < \sigma^2$, a distribuição dos dados é igual ao da Figura 4.1 mudando a média para $\mu = 1$, com cada curva representando uma mudança na correção, a curva com maior densidade tem $r = 0,9$, conforme a correlação diminui a densidade também diminui e a calda da densidade alonga, em relação a Figura 4.1 a calda desta densidade é mais pesada. Independente do valor de r o NMA A apresenta resultados muito próximos. Neste cenário, o método utilizado em Oliveira (2013) tem um desempenho superior, visto que os valores do NMA são menores.

Tabela 4.2: NMA_1 para fluxos dependentes com $\mu = 1 < \sigma^2 = 100$ e $r = 0,1; 0,2; 0,3; 0,5; 0,7$ e $0,9$.

| k | Fluxos dependentes (r) | | | | | |
|-----|----------------------------|--------|--------|--------|--------|--------|
| | 0,1 | 0,2 | 0,3 | 0,5 | 0,7 | 0,9 |
| 2 | 168,53 | 115,52 | 138,40 | 203,04 | 200,77 | 142,88 |
| 3 | 91,40 | 54,01 | 87,15 | 133,36 | 89,37 | 139,25 |
| 4 | 49,93 | 39,65 | 40,89 | 39,62 | 99,43 | 107,27 |
| 5 | 39,77 | 55,02 | 41,91 | 41,95 | 31,99 | 73,10 |
| 6 | 18,32 | 23,02 | 28,54 | 31,15 | 30,45 | 48,44 |
| 7 | 20,06 | 17,46 | 10,91 | 24,25 | 26,06 | 45,80 |
| 8 | 21,52 | 9,91 | 17,60 | 19,95 | 24,15 | 48,78 |
| 9 | 12,39 | 11,27 | 18,01 | 14,62 | 15,75 | 31,98 |
| 10 | 13,32 | 16,66 | 9,39 | 10,93 | 17,85 | 28,06 |
| 15 | 2,68 | 4,04 | 4,20 | 3,32 | 9,65 | 5,33 |
| 20 | 0,57 | 0,89 | 1,31 | 1,61 | 3,55 | 3,61 |

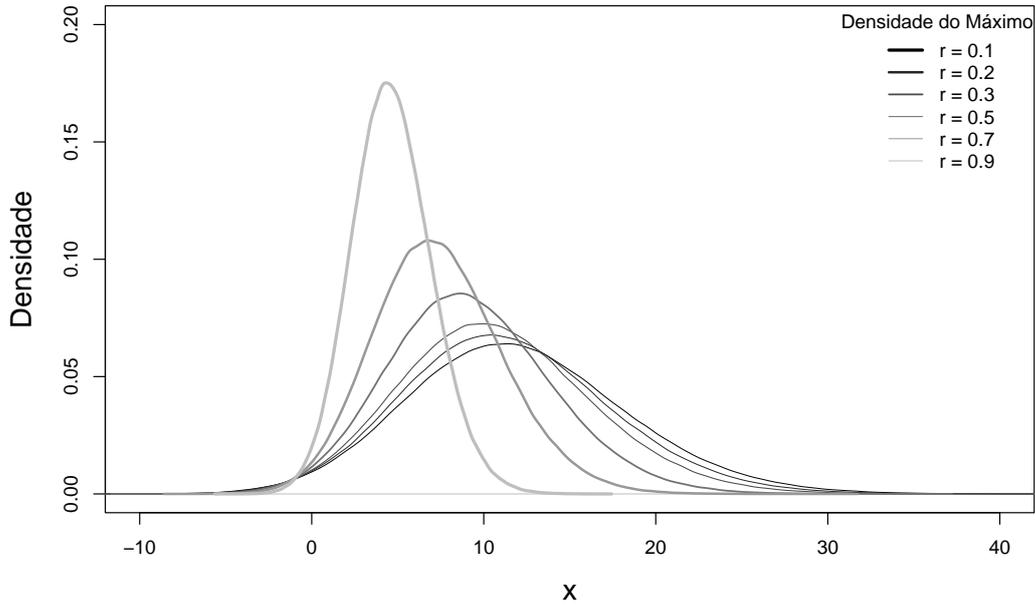


Figura 4.2: NMA_1 para fluxos dependentes com $\mu = 1 < \sigma^2 = 100$ e $r = 0, 1; 0, 2; 0, 3; 0, 5; 0, 7$ e $0, 9$.

De maneira análoga, replicamos a Tabela 4.3 que apresenta os valores de NMA_1 para uma média superior à variância ($\mu = 100 > \sigma^2 = 9^2$), utilizando os resultados apresentados em Oliveira (2013) e os obtidos a partir do algoritmo apresentado no Capítulo 3. Os dados correspondentes estão visualizados na Figura 4.3 para diferentes níveis de correlação, neste cenário com $\mu > \sigma^2$ é possível notar um comportamento parecido com o da Figura 4.1, mas com uma cauda curta, por conta de um σ^2 menor. No entanto, vale notar a ausência da coluna referente à correlação de 0,90. Através de uma análise conduzida no software R, Oliveira (2013) identificou que a matriz de covariância não é definida positiva, levando à exclusão da sexta coluna dessa análise.

Tabela 4.3: NMA_1 para fluxos dependentes com $\mu = 100 > \sigma^2 = 9^2$, obtido pelo algoritmo proposto variando a correlação $r = 0, 1; 0, 2; 0, 3; 0, 5; \text{ e } 0, 7$.

| k | Fluxos dependentes (r) | | | | |
|-----|----------------------------|--------|--------|--------|--------|
| | 0,1 | 0,2 | 0,3 | 0,5 | 0,7 |
| 2 | 131,43 | 133,29 | 161,41 | 211,87 | 124,95 |
| 3 | 62,98 | 54,84 | 70,49 | 68,03 | 86,86 |
| 4 | 61,15 | 29,75 | 61,97 | 51,15 | 81,15 |
| 5 | 31,59 | 30,84 | 45,31 | 64,74 | 33,86 |
| 6 | 28,66 | 25,32 | 27,82 | 20,85 | 37,37 |
| 7 | 16,55 | 18,89 | 20,52 | 20,08 | 38,59 |
| 8 | 13,58 | 9,75 | 16,76 | 22,13 | 17,88 |
| 9 | 18,71 | 14,22 | 13,01 | 14,52 | 25,57 |
| 10 | 5,81 | 7,57 | 10,89 | 21,16 | 12,20 |
| 15 | 1,92 | 2,09 | 3,95 | 3,71 | 4,68 |
| 20 | 0,68 | 0,44 | 1,91 | 1,60 | 2,75 |

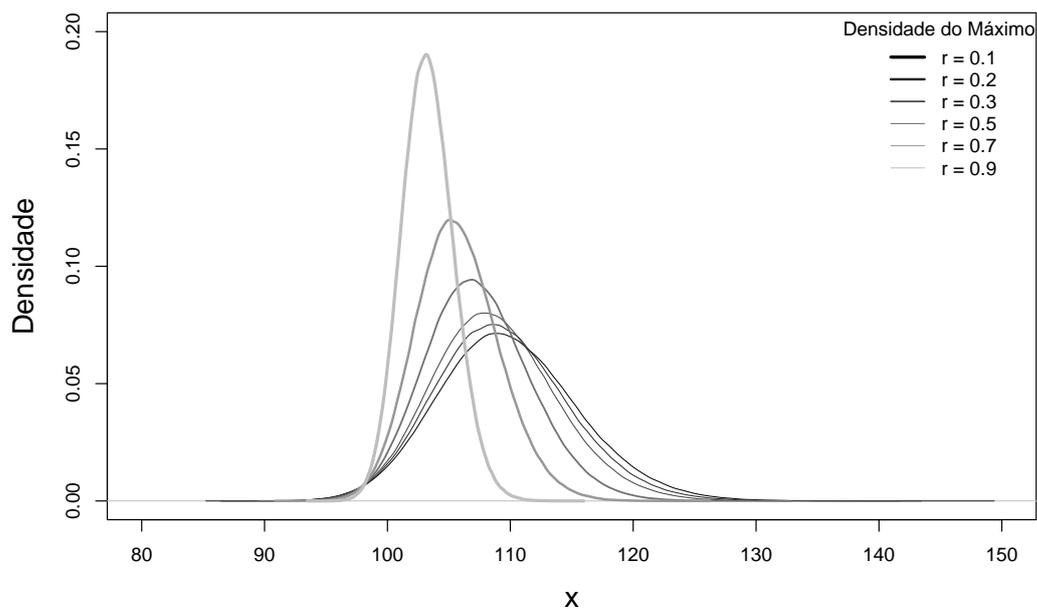


Figura 4.3: NMA_1 para fluxos dependentes com $\mu = 100 > \sigma^2 = 9^2$ e $r = 0, 1; 0, 2; 0, 3; 0, 5; \text{ e } 0, 7$.

No cenário apresentado na Tabela 4.3 é possível notar que o GG utilizado nesta dissertação teve uma performance superior em todos os cenários. Isto pode indicar que este método tem uma eficácia maior quando sua média é maior que a variância, além de ser competitivo no cenário em que a média é igual a variância.

4.1.1 Eficiência entre o GG e o GG A para o NMA_1

O cálculo da eficiência consiste na razão entre os valores encontrados para o NMA_1 do GG apresentados em Oliveira (2013) e do GG A apresentados nesta Seção 4. Em resumo, divide-se o valores do NMA_1 do GG pelos da Tabela 4.1, conforme a Equação 4.1. Os resultados são apresentados na Tabela 4.4.

$$ER = \frac{NMA_{1.GG}}{NMA_{1.GGA}}, \text{ se } \begin{cases} = 1 \rightarrow \text{Desempenho equivalente} \\ < 1 \rightarrow \text{GC A melhor do que o GG} \\ > 1 \rightarrow \text{GG melhor do que o GG A.} \end{cases} \quad (4.1)$$

Tabela 4.4: Eficiência relativa para NMA_1 para fluxos dependentes com $\mu = \sigma^2 = 100$, $NMA_{1.GG}$ obtido pelo algoritmo proposto em Oliveira (2013) e o $NMA_{1.GGA}$ pelo algoritmo variando a correlação

| k | Fluxos dependentes (r) | | | | | |
|-----|----------------------------|------|------|------|------|------|
| | 0,1 | 0,2 | 0,3 | 0,5 | 0,7 | 0,9 |
| 1 | - | - | - | - | - | - |
| 2 | 1,15 | 0,67 | 0,78 | 0,88 | 1,04 | 1,03 |
| 3 | 0,99 | 0,99 | 0,78 | 0,90 | 0,76 | 1,00 |
| 4 | 0,92 | 1,31 | 1,27 | 1,53 | 1,68 | 1,18 |
| 5 | 0,91 | 0,90 | 0,94 | 1,30 | 0,86 | 1,07 |
| 6 | 0,93 | 2,16 | 0,69 | 0,88 | 0,56 | 0,97 |
| 7 | 0,77 | 1,34 | 0,79 | 0,63 | 0,99 | 1,10 |
| 8 | 0,93 | 0,58 | 1,22 | 1,00 | 0,64 | 0,87 |
| 9 | 0,95 | 1,40 | 1,36 | 1,25 | 0,96 | 1,08 |
| 10 | 2,47 | 1,20 | 0,66 | 1,24 | 0,83 | 0,69 |
| 15 | 0,97 | 2,47 | 0,58 | 0,44 | 2,84 | 0,90 |
| 20 | 1,22 | 1,09 | 1,11 | 1,34 | 0,89 | 1,63 |

No primeiro cenário, caracterizado por $\mu = \sigma^2$, a análise da Tabela 4.4 revela que, para fluxos inferiores a 10, os resultados apresentam proximidade entre os métodos considerados. Contudo, a partir de $k = 10$, observa-se um aumento nos resultados, com valores superiores a 1, indicando uma eficiência superior para o método $NMA_{1.GG}$ em cenários com maior quantidade de fluxos. De maneira análoga, a Tabela 4.5 foi elaborada para comparar a eficiência entre os dois métodos em um cenário em que $\mu < \sigma^2$, com $\mu = 1$ e $\sigma^2 = 100$.

No segundo cenário em análise, observamos que a eficiência relativa é aproximadamente homogênea para todas as combinações de parâmetros k e r . Não temos uma indicação sobre a preferência por um método, dado que $\mu < \sigma^2$. Além dos cenários previamente discutidos, é crucial abordar a análise do terceiro cenário, caracterizado por

Tabela 4.5: Eficiência relativa para NMA_1 para fluxos dependentes com $\mu = 1 < \sigma^2 = 100$, $NMA_{1.GG}$ obtido pelo algoritmo proposto em Oliveira (2013) e o $NMA_{1.GGA}$ pelo algoritmo variando a correlação.

| k | Fluxos dependentes (r) | | | | | |
|-----|----------------------------|------|------|------|------|------|
| | 0,1 | 0,2 | 0,3 | 0,5 | 0,7 | 0,9 |
| 1 | - | - | - | - | - | - |
| 2 | 0,90 | 1,18 | 0,98 | 0,63 | 0,81 | 1,32 |
| 3 | 0,85 | 1,52 | 0,89 | 0,64 | 1,13 | 0,92 |
| 4 | 1,19 | 1,30 | 1,20 | 1,29 | 0,60 | 0,85 |
| 5 | 0,88 | 0,70 | 0,82 | 1,01 | 1,60 | 1,02 |
| 6 | 1,63 | 1,16 | 0,89 | 0,90 | 1,13 | 1,21 |
| 7 | 0,90 | 1,14 | 1,70 | 0,92 | 0,90 | 0,91 |
| 8 | 0,74 | 1,47 | 0,87 | 0,89 | 0,82 | 0,75 |
| 9 | 0,82 | 1,08 | 0,69 | 0,93 | 1,01 | 0,98 |
| 10 | 0,59 | 0,53 | 0,99 | 1,20 | 0,76 | 0,83 |
| 15 | 0,93 | 0,80 | 0,65 | 1,70 | 0,63 | 1,81 |
| 20 | 1,38 | 1,00 | 0,94 | 0,80 | 0,61 | 1,45 |

$\mu = 100$ e $\sigma^2 = 9^2$. Esta configuração particular estabelece um contexto distinto, em que μ é maior que σ^2 .

Tabela 4.6: Eficiência relativa para NMA_1 para fluxos dependentes com $\mu = 100 > \sigma^2 = 9^2$, $NMA_{1.GG}$ obtido pelo algoritmo proposto em Oliveira (2013) e o $NMA_{1.GGA}$ pelo algoritmo variando a correlação.

| k | Fluxos dependentes (r) | | | | | |
|-----|----------------------------|------|------|------|------|-----|
| | 0,1 | 0,2 | 0,3 | 0,5 | 0,7 | 0,9 |
| 1 | - | - | - | - | - | - |
| 2 | 1,04 | 1,05 | 0,89 | 0,67 | 1,47 | - |
| 3 | 1,28 | 1,46 | 1,03 | 1,29 | 1,30 | - |
| 4 | 1,03 | 1,85 | 0,84 | 1,38 | 0,94 | - |
| 5 | 1,18 | 1,24 | 0,88 | 0,68 | 1,82 | - |
| 6 | 0,85 | 1,07 | 0,97 | 1,52 | 1,37 | - |
| 7 | 1,20 | 1,05 | 1,15 | 1,20 | 1,08 | - |
| 8 | 1,07 | 1,84 | 0,90 | 1,01 | 1,98 | - |
| 9 | 0,80 | 0,89 | 0,82 | 1,13 | 0,98 | - |
| 10 | 0,57 | 1,11 | 0,92 | 0,56 | 2,30 | - |
| 15 | 1,52 | 1,62 | 0,96 | 0,77 | 1,67 | - |
| 20 | 1,41 | 1,70 | 0,43 | 1,58 | 1,51 | - |

Neste terceiro cenário, com $\mu = 100$ e $\sigma^2 = 9^2$, observando a Tabela 4.6, não é possível uma formulação de uma conclusão precisa devido à proximidade dos resultados obtidos pelos métodos considerados. Observa-se uma notável similaridade nos resultados em relação ao cenário anterior (cenário 2, onde $\mu < \sigma^2$), o que dificulta a identificação de vantagens ou desvantagens distintas entre os métodos em questão neste contexto específico. A proximidade dos resultados sugere uma relativa insensibilidade dos métodos às

variações nos parâmetros μ e σ^2 nesse intervalo, tornando a interpretação mais desafiadora.

4.2 Eficiência entre o GG A, GG e o GCS na presença de correlação e alterações na média e variância do processo

Na Seção 4.1, realizou-se um estudo para investigar o comportamento do NMA_1 em relação ao GG A, considerando a média na presença de correlação entre os fluxos. Nesta seção, examinamos o desempenho dos GG A em termos de eficiência em comparação com o modelo de Shewhart e o Monte Carlo apresentado em Oliveira (2013), quando há alterações na média e na variância do processo, tanto isoladamente como conjuntamente. Essas alterações são representadas, respectivamente, pelas constantes δ e λ .

Para permitir comparações apropriadas, utilizamos o mesmo algoritmo para calcular o NMA_2 para o GG A com fluxos correlacionados. A comparação entre GG A, GCS e GG requer a manutenção da mesma taxa de alarme falso, estabelecida por meio de uma modificação no fator de largura L . Os valores para este estudo são obtidos através do mesmo algoritmo, com $\mu = \sigma^2 = 100$.

4.2.1 Alteração na média do processo (δ)

Considerando a característica da qualidade X com média μ e variância σ^2 , supomos que, no instante t , ocorreu uma perturbação δ com valores de 0,1;0,3;0,5;1,0;2,0;3,0. Essa perturbação desloca a média para $\mu_1 = \mu + \delta\sigma$. Em outras palavras, estamos investigando os NMA_{21} para o GG com correlação e para o GCS com coeficientes de correlação r iguais a 0,10;0,50 e 0,90. Os valores foram devidamente calculados e estão apresentados na Tabela 4.5. Importante ressaltar que estamos seguindo o mesmo cenário de NMA_{21} utilizado em Oliveira (2013) para permitir uma comparação consistente dos resultados.

Fixando k e r e variando δ , percebe-se que o NMA_{21} no GG A tem um desempenho superior em detectar o deslocamento na média quando a perturbação cresce. Mantendo k e δ fixos, percebe-se que quando r aumenta, o NMA_{21} apresenta um comportamento de crescimento também, os valores do NMA_{21} aumentam junto com r . Quando r e δ

Tabela 4.7: NMA_{21} para o GG A com correlação entre os fluxos e um deslocamento na média.

| δ | $r = 0,1$ | | | $r = 0,5$ | | | $r = 0,9$ | | |
|----------|-----------|--------|-------|-----------|--------|-------|-----------|--------|--------|
| | 3 | 5 | 10 | 3 | 5 | 10 | 3 | 5 | 10 |
| 0,1 | 69,164 | 32,090 | 8,928 | 120,480 | 42,752 | 8,118 | 144,820 | 42,364 | 18,084 |
| 0,3 | 31,242 | 13,240 | 4,826 | 38,364 | 11,826 | 5,870 | 59,976 | 23,236 | 10,266 |
| 0,5 | 7,546 | 4,218 | 1,140 | 8,908 | 7,818 | 1,620 | 16,390 | 9,624 | 5,074 |
| 1,0 | 1,750 | 1,000 | 1,000 | 2,260 | 1,336 | 1,000 | 3,032 | 1,630 | 1,000 |
| 2,0 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |
| 3,0 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |

permanecem constantes, o NMA_{21} decresce quando k aumenta. Outro fator de destaque está presente na rápida convergência de NMA_{21} para 1 quando a perturbação δ passa de 1,0 para 2,0, cuja detecção do deslocamento na média ocorreria na próxima inspeção.

O cálculo da eficiência consiste na razão entre os valores encontrados para o NMA_{21} do GG A, GG e do GCS. Em resumo, divide-se o valores da Tabela 4.7 pelos valores do GCS e GG apresentados em Oliveira (2013), conforme a Equação 4.2. Os resultados são apresentados na Tabela 4.8.

$$ER = \frac{NMA_{21.GCS}}{NMA_{21.GGA}}, \text{ se } \begin{cases} = 1 \rightarrow \text{Desempenho equivalente} \\ < 1 \rightarrow \text{GCS melhor do que o GG A} \\ > 1 \rightarrow \text{GG A melhor do que o GCS} \end{cases} \quad (4.2)$$

Tabela 4.8: Eficiência entre o GCS e o GG A para um deslocamento na média do processo e com fluxos correlacionados.

| δ | $r = 0,1$ | | | $r = 0,5$ | | | $r = 0,9$ | | |
|----------|-----------|------|------|-----------|------|------|-----------|------|------|
| | 3 | 5 | 10 | 3 | 5 | 10 | 3 | 5 | 10 |
| 0,1 | 1,23 | 1,60 | 2,93 | 0,76 | 1,36 | 4,05 | 0,97 | 2,53 | 4,30 |
| 0,3 | 0,61 | 0,90 | 1,32 | 0,57 | 1,24 | 1,56 | 0,54 | 1,15 | 2,03 |
| 0,5 | 0,73 | 0,85 | 1,92 | 0,73 | 0,60 | 2,02 | 0,59 | 0,85 | 1,34 |
| 1,0 | 0,67 | 1,06 | 1,01 | 0,59 | 0,90 | 1,10 | 0,54 | 0,95 | 1,44 |
| 2,0 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 |
| 3,0 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 |

Os resultados apresentados na Tabela 4.8 destacam que, para um mesmo valor de δ , observa-se um desempenho superior do método de Shewhart em relação ao gráfico por grupos à medida que o número de fluxos k aumenta. Isso é evidenciado pela visível diminuição na eficiência do gráfico por grupos com o aumento de k . Além disso, para valores fixos de δ e k , a eficiência do gráfico por grupos melhora conforme a correlação entre os fluxos aumenta, indicando uma superioridade do GG A em relação ao GCS.

Nota-se que o GCS é capaz de detectar um deslocamento na média do processo mais rapidamente do que o GG, especialmente quando $\delta \leq 1$, em que $ER \leq 1$ na maioria dos casos. É importante ressaltar que, quando a perturbação atinge valores superiores a 1, a convergência de ER para 1 é rápida, sugerindo uma equivalência entre os dois modelos.

O GCS tradicional demonstra maior eficiência do que o GG A para pequenas perturbações (entre 0,3 e 1,0), pois requer menos amostras para detectar um deslocamento na média do processo. À medida que δ aumenta, o uso do GG A torna-se mais vantajoso devido à sua velocidade e facilidade de aplicação. De maneira análoga a Equação 4.2 foi feita a Tabela 4.9 com a eficiência relativa entre os métodos GG e GG A, os valores do GG são da simulação feita em Oliveira (2013).

Tabela 4.9: Eficiência entre o GG e o GG A para um deslocamento na média do processo e com fluxos correlacionados.

| δ | $r = 0,1$ | | | $r = 0,5$ | | | $r = 0,9$ | | |
|----------|-----------|------|-------|-----------|------|-------|-----------|------|-------|
| | 3 | 5 | 10 | 3 | 5 | 10 | 3 | 5 | 10 |
| 0,1 | 4.42 | 9.28 | 27.51 | 2.56 | 6.70 | 33.82 | 2.12 | 6.75 | 15.18 |
| 0,3 | 3.36 | 6.24 | 14.87 | 2.69 | 7.48 | 13.41 | 2.08 | 4.75 | 9.71 |
| 0,5 | 4.32 | 5.86 | 16.16 | 3.50 | 3.71 | 14.75 | 2.50 | 4.02 | 7.38 |
| 1,0 | 2.18 | 2.89 | 1.79 | 1.68 | 2.70 | 2.41 | 1.74 | 2.90 | 3.44 |
| 2,0 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 |
| 3,0 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 |

Ao analisar a Tabela 4.9, é evidente que a detecção do GG A demonstra uma velocidade superior à do GG quando os valores de δ são inferiores a 2. Além disso, para valores de δ de 2,0 e 3,0, ambos os métodos exibem desempenho equivalente. Nota-se também que o número de fluxos (k) exerce uma influência significativa no desempenho, sendo que um aumento neste número resulta em uma detecção mais rápida do GG A. Adicionalmente, observa-se uma relação inversa com a correlação, onde um aumento na correlação está associado a uma diminuição na discrepância de desempenho entre os dois métodos.

4.2.2 Alteração na variância do processo (λ)

O NMA_{22} representa um deslocamento na variância do processo, e assim como no caso da alteração na média, a introdução de uma estrutura de dependência impacta a aplicação da propriedade de independência, tornando-a impraticável. Nesse contexto, a técnica de simulação foi novamente empregada para determinar os NMA_{22} em res-

posta a uma mudança na variância, onde $\sigma^{*2} = \sigma^2\lambda^2$, com λ assumindo os valores 1,5; 2,0; 2,5; 3,0; 5,0. Os resultados dessas simulações estão apresentados na Tabela 4.10.

No caso do GG A, foram calculados os NMA_{22} para diferentes valores de k ($k = 3, 5, 10$) considerando uma perturbação λ variando entre 1,5 e 5,0. Esses valores específicos estão detalhados na Tabela 4.10, destacando uma sensibilidade em relação ao deslocamento na variância. Em outras palavras, observa-se que, no contexto do gráfico para a média tratado neste estudo, um deslocamento na variância influencia significativamente o desempenho do processo.

Tabela 4.10: NMA_{22} para o GG A com correlação entre os fluxos e um deslocamento na variância.

| λ | $r = 0,1$ | | | $r = 0,5$ | | | $r = 0,9$ | | |
|-----------|-----------|-------|-------|-----------|-------|-------|-----------|-------|-------|
| | 3 | 5 | 10 | 3 | 5 | 10 | 3 | 5 | 10 |
| 1,5 | 3,990 | 1,948 | 1,000 | 5,828 | 2,738 | 1,000 | 7,102 | 5,452 | 1,942 |
| 2,0 | 1,722 | 1,000 | 1,000 | 2,088 | 1,000 | 1,000 | 2,960 | 1,574 | 1,000 |
| 2,5 | 1,056 | 1,000 | 1,000 | 1,300 | 1,000 | 1,000 | 1,892 | 1,280 | 1,000 |
| 3,0 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,306 | 1,000 | 1,000 |
| 5,0 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |

Conforme evidenciado na Tabela 4.10, observa-se que a detecção de perturbações da variância do processo é rapidamente alcançada, indicada pela rápida convergência para o valor 1 quando $\lambda \geq 2,5$. Adicionalmente, nota-se que o incremento no número de fluxos contribui para aprimorar o desempenho do NMA_{22} , e o mesmo padrão é observado com o aumento de λ . A detecção eficiente de deslocamentos é caracterizada por uma convergência rápida para o valor unitário, sendo essa eficiência intensificada tanto pelo aumento do número de fluxos quanto pelo crescimento de λ . De maneira similar a Equação 4.2 foi calculado a eficiência relativa entre os 3 métodos, com o deslocamento na variância do processo e com fluxos correlacionados.

Tabela 4.11: Eficiência entre o GC e o GG A para um deslocamento na variância do processo e com fluxos correlacionados.

| δ | $r = 0,1$ | | | $r = 0,5$ | | | $r = 0,9$ | | |
|----------|-----------|------|------|-----------|------|------|-----------|------|------|
| | 3 | 5 | 10 | 3 | 5 | 10 | 3 | 5 | 10 |
| 1,5 | 0,55 | 0,81 | 1,16 | 0,39 | 0,60 | 1,20 | 0,34 | 0,32 | 0,66 |
| 2,0 | 0,69 | 1,05 | 1,00 | 0,57 | 1,05 | 1,00 | 0,41 | 0,67 | 1,00 |
| 2,5 | 1,00 | 1,01 | 1,00 | 0,81 | 1,01 | 1,00 | 0,56 | 0,79 | 1,00 |
| 3,0 | 1,02 | 1,00 | 1,00 | 1,02 | 1,00 | 1,00 | 0,78 | 1,00 | 1,00 |
| 5,0 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 |

Observando a Tabela 4.11 o método GG A demonstrou ser competitivo em relação

ao GC para todos os valores de δ no intervalo de 2,5 a 5. Essa eficiência relativa apresenta um aumento conjunto com o parâmetro k , indicando que quanto maior o número de fluxos (k), maior é a eficiência relativa do método GG A. Em essência, o método GG A revelou-se menos eficaz na detecção de deslocamentos na média do processo com fluxos correlacionados em que o δ é menor que 2,5.

Tabela 4.12: Eficiência entre o GG e o GG A para um deslocamento na variância do processo e com fluxos correlacionados.

| δ | $r = 0,1$ | | | $r = 0,5$ | | | $r = 0,9$ | | |
|----------|-----------|------|------|-----------|------|------|-----------|------|------|
| | 3 | 5 | 10 | 3 | 5 | 10 | 3 | 5 | 10 |
| 1,5 | 2,94 | 4,91 | 5,41 | 1,88 | 3,08 | 5,28 | 1,47 | 1,43 | 2,34 |
| 2,0 | 1,97 | 2,30 | 1,35 | 1,55 | 2,37 | 1,17 | 0,96 | 1,21 | 1,13 |
| 2,5 | 1,64 | 1,16 | 1,00 | 1,31 | 1,26 | 1,00 | 0,87 | 0,86 | 1,00 |
| 3,0 | 0,67 | 1,06 | 1,01 | 0,59 | 0,90 | 1,10 | 0,54 | 0,95 | 1,44 |
| 5,0 | 1,38 | 1,00 | 1,00 | 1,24 | 1,00 | 1,00 | 0,97 | 1,00 | 1,00 |

Conforme os resultados apresentados na Tabela 4.12, destaca-se que o desempenho do GG A demonstrou superioridade em situações onde $\lambda < 2,5$. Em intervalos de λ diferentes, ambos os métodos revelaram-se altamente competitivos, deixando a escolha entre eles ao critério do analista, dependendo dos critérios específicos utilizados na análise.

4.2.3 Alteração na média e na variância do processo

O cálculo probabilístico do NMA_{23} , considerando a introdução de uma estrutura de correlação entre os fluxos, assemelha-se aos métodos empregados nas seções anteriores, demandando a utilização de abordagens computacionais. Nesse contexto, assume-se que $\mu_1 = \mu + \delta\sigma$ e $\sigma^{*2} = \sigma^2\lambda^2$.

Tabela 4.13: NMA_{23} para o Gráfico por Grupos A com correlação $r = 0,1; 0,5; 0,9$ e um deslocamento na média (δ) e na variância (λ) para $k = 3, 5, 10$.

| δ | λ | $r = 0,1$ | | | $r = 0,5$ | | | $r = 0,9$ | | |
|----------|-----------|-----------|-------|-------|-----------|-------|-------|-----------|-------|-------|
| | | 3 | 5 | 10 | 3 | 5 | 10 | 3 | 5 | 10 |
| 0,1 | 1,5 | 4,922 | 2,27 | 0,724 | 4,404 | 3,03 | 0,662 | 7,386 | 4,80 | 3,040 |
| | 2,0 | 1,926 | 1,038 | 0,398 | 1,800 | 1,226 | 0,454 | 3,374 | 1,796 | 0,868 |
| | 3,0 | 0,848 | 0,564 | 0,226 | 0,998 | 0,524 | 0,284 | 1,428 | 0,938 | 0,344 |
| | 5,0 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |
| 0,3 | 1,5 | 3,928 | 1,794 | 0,436 | 4,876 | 2,952 | 0,826 | 5,612 | 3,954 | 1,320 |
| | 2,0 | 1,646 | 0,928 | 0,360 | 1,994 | 0,932 | 0,334 | 2,688 | 2,392 | 0,848 |
| | 3,0 | 1,000 | 1,000 | 1,000 | 1,206 | 1,000 | 1,000 | 1,398 | 1,000 | 1,000 |
| | 5,0 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |
| 0,5 | 1,5 | 2,178 | 1,172 | 0,474 | 3,288 | 1,728 | 0,548 | 5,090 | 2,818 | 1,422 |
| | 2,0 | 1,254 | 0,656 | 0,236 | 1,712 | 1,168 | 0,290 | 2,342 | 1,652 | 0,700 |
| | 3,0 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,492 | 1,000 | 1,000 |
| | 5,0 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |
| 1,0 | 1,5 | 1,236 | 0,632 | 0,272 | 1,376 | 1,066 | 0,380 | 1,968 | 1,238 | 0,488 |
| | 2,0 | 1,010 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |
| | 3,0 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |
| | 5,0 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |
| 2,0 | 1,5 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |
| | 2,0 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |
| | 3,0 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |
| | 5,0 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 | 1,000 |

Com base nos dados apresentados na Tabela 4.13, surgem evidências para analisar o perfil do NMA_{23} para o GG A na presença de correlação entre os fluxos. Assim como observado nos resultados das Tabelas 4.6 e 4.7, o NMA_{23} aumenta em função da correlação r quando δ , λ e k permanecem fixos. Entretanto, o NMA_{23} diminui em relação a k quando δ , λ e r são mantidos constantes. Em outras palavras, o NMA_{23} aumenta com o crescimento de r e diminui com o aumento de k . Este fenômeno é constatado para ambos os tipos de perturbação.

O desempenho para o GG A com correlação entre os fluxos e perturbação na média e na variância do processo apresentou um comportamento, em parte, análogo aos demais. Fixando as perturbações δ e λ e a correlação r , o NMA_{23} decresce à medida que o número de fluxos k aumenta. Fixando δ , k e r , o NMA_{23} também diminui quando λ aumenta. O comportamento é semelhante quando fixam-se λ , k e r e variando δ .

4.2.4 Eficiência com alteração conjunta na média e na variância do processo

Foi realizado o cálculo da eficiência relativa, comparando GG A com GCS e GG A com o GG, o cálculo é feito de maneira análoga ao da Equação 4.2.

Tabela 4.14: Eficiência Relativa do NMA_{23} entre GCS e o GG A com correlação $r = 0, 1; 0, 5; 0, 9$ e um deslocamento na média (δ) e na variância (λ) para $k = 3, 5, 10$.

| δ | λ | $r = 0, 1$ | | | $r = 0, 5$ | | | $r = 0, 9$ | | |
|----------|-----------|------------|------|------|------------|------|------|------------|------|------|
| | | 3 | 5 | 10 | 3 | 5 | 10 | 3 | 5 | 10 |
| 0,1 | 1,5 | 1,08 | 1,51 | 2,79 | 1,29 | 1,25 | 3,58 | 0,88 | 0,98 | 1,05 |
| | 2,0 | 0,95 | 1,32 | 1,08 | 1,05 | 1,16 | 1,11 | 0,60 | 0,86 | 1,19 |
| | 3,0 | 1,13 | 1,03 | 1,00 | 1,14 | 1,03 | 1,00 | 0,80 | 1,04 | 1,00 |
| | 5,0 | 1,01 | 1,00 | 1,00 | 1,01 | 1,00 | 1,00 | 1,01 | 1,00 | 1,00 |
| 0,3 | 1,5 | 0,91 | 1,34 | 1,55 | 0,80 | 0,94 | 1,88 | 0,81 | 0,87 | 1,90 |
| | 2,0 | 1,01 | 1,28 | 1,05 | 0,87 | 1,34 | 1,09 | 0,69 | 0,61 | 1,16 |
| | 3,0 | 1,12 | 1,02 | 1,00 | 0,93 | 1,03 | 1,00 | 0,81 | 1,03 | 1,00 |
| | 5,0 | 1,01 | 1,00 | 1,00 | 1,01 | 1,00 | 1,00 | 1,01 | 1,00 | 1,00 |
| 0,5 | 1,5 | 1,03 | 1,39 | 1,20 | 0,76 | 1,10 | 1,44 | 0,57 | 0,83 | 1,29 |
| | 2,0 | 1,14 | 1,16 | 1,02 | 0,88 | 1,06 | 1,06 | 0,69 | 0,81 | 1,12 |
| | 3,0 | 1,10 | 1,02 | 1,00 | 1,11 | 1,02 | 1,00 | 0,75 | 1,03 | 1,00 |
| | 5,0 | 1,01 | 1,00 | 1,00 | 1,01 | 1,00 | 1,00 | 1,01 | 1,00 | 1,00 |
| 1,0 | 1,5 | 0,93 | 1,04 | 1,00 | 0,90 | 1,05 | 1,04 | 0,69 | 1,00 | 1,13 |
| | 2,0 | 1,08 | 1,02 | 1,00 | 1,13 | 1,05 | 1,01 | 1,19 | 1,09 | 1,03 |
| | 3,0 | 1,04 | 1,00 | 1,00 | 1,05 | 1,01 | 1,00 | 1,06 | 1,01 | 1,00 |
| | 5,0 | 1,01 | 1,00 | 1,00 | 1,01 | 1,00 | 1,00 | 1,01 | 1,00 | 1,00 |
| 2,0 | 1,5 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 |
| | 2,0 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 |
| | 3,0 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 |
| | 5,0 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 |

Ao examinar a Tabela 4.14, observa-se que o GCS demonstra um desempenho superior quando $k = 3$, enquanto o GG A exibe um desempenho superior quando $k = 10$. No caso de $k = 5$, ambos os métodos apresentam desempenhos semelhantes, particularmente quando $\delta \neq 1$. Notavelmente, quando $\delta = 2$, os desempenhos dos dois métodos são iguais. Essas conclusões evidenciam a influência da escolha de k no desempenho relativo entre o GCS e o GG A, destacando variações significativas dependendo do valor específico de δ .

Tabela 4.15: Eficiência Relativa do NMA_{23} entre GG e o GG A com correlação $r = 0, 1; 0, 5; 0, 9$ e um deslocamento na média (δ) e na variância (λ) para $k = 3, 5, 10$.

| δ | λ | $r = 0, 1$ | | | $r = 0, 5$ | | | $r = 0, 9$ | | |
|----------|-----------|------------|------|------|------------|------|------|------------|------|------|
| | | 3 | 5 | 10 | 3 | 5 | 10 | 3 | 5 | 10 |
| 0,1 | 1,5 | 2,35 | 2,48 | 2,50 | 2,23 | 1,74 | 2,62 | 1,29 | 1,09 | 0,54 |
| | 2,0 | 1,62 | 1,57 | 1,00 | 1,62 | 1,34 | 1,00 | 0,81 | 0,85 | 1,00 |
| | 3,0 | 1,27 | 1,00 | 1,00 | 1,28 | 1,00 | 1,00 | 0,91 | 1,00 | 1,00 |
| | 5,0 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 |
| 0,3 | 1,5 | 1,99 | 2,45 | 1,42 | 1,48 | 1,39 | 1,19 | 1,17 | 1,02 | 1,20 |
| | 2,0 | 1,72 | 1,58 | 1,00 | 1,38 | 1,54 | 1,00 | 1,03 | 0,58 | 1,00 |
| | 3,0 | 1,25 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 0,84 | 1,00 | 1,00 |
| | 5,0 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 |
| 0,5 | 1,5 | 2,33 | 2,34 | 1,00 | 1,46 | 1,44 | 1,00 | 0,98 | 0,93 | 0,70 |
| | 2,0 | 1,74 | 1,27 | 1,00 | 1,36 | 1,07 | 1,00 | 0,91 | 0,77 | 1,00 |
| | 3,0 | 1,18 | 1,00 | 1,00 | 1,23 | 1,00 | 1,00 | 0,79 | 1,00 | 1,00 |
| | 5,0 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 |
| 1,0 | 1,5 | 1,58 | 1,02 | 1,00 | 1,41 | 1,08 | 1,00 | 0,88 | 0,99 | 1,00 |
| | 2,0 | 1,37 | 1,00 | 1,00 | 1,41 | 1,00 | 1,00 | 1,29 | 1,00 | 1,00 |
| | 3,0 | 1,13 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,01 | 1,00 | 1,00 |
| | 5,0 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 |
| 2,0 | 1,5 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 |
| | 2,0 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 |
| | 3,0 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 |
| | 5,0 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 | 1,00 |

Ao analisar a Tabela 4.15, observa-se que o GG A apresenta um desempenho superior em quase todas as variações, com exceção quando $r = 0, 9$, momento em que o GG demonstra maior eficácia na detecção de alarmes verdadeiros. Quando $\delta = 2$, os desempenhos dos dois métodos são equivalentes, e à medida que λ aumenta, os dois métodos se aproximam. Essas conclusões destacam a influência significativa da escolha de r no desempenho relativo entre o GG e o GG A, evidenciando pequenas variações conforme os valores de δ e λ são variados.

Capítulo 5

Aplicação

Neste capítulo, aplicaremos o método desenvolvido e discutido no Capítulo 3 do desempenho do GG a um conjunto de dados reais com o objetivo de analisar e extrair informações valiosas. A aplicação prática do método proposto é essencial para avaliar sua eficácia e utilidade em cenários aplicados.

A base de dados selecionada para esta aplicação é o “Processo de pintura a pó de fogões”, pode ser encontrada em Alexandre *et al.* (2006). Esta base de dados oferece uma oportunidade significativa para explorar os conceitos e técnicas discutidos neste estudo em um contexto aplicado e relevante. Antes de prosseguir com a análise dos dados, é crucial compreender a estrutura e as características da base de dados selecionada. A seguir, apresentaremos uma descrição detalhada da base de dados, incluindo suas fontes, tamanho, formato e as variáveis disponíveis.

A produção de fogões inclui uma etapa de pintura a pó eletrostática, composta por três linhas de produção. Cada linha é composta por pistolas móveis, verticalmente distribuídas em pontos estratégicos, que aplicam o pó sobre as peças, e por um transportador suspenso que direciona horizontalmente as peças a serem pintadas. O pó é lançado através da pressão da linha de ar e direcionado para a peça por meio de um campo magnético criado entre a peça e a pistola. Dessa forma, o pó adere à peça devido à tensão eletrostática, permanecendo nela até que seja aquecido e adquira as características de uma pintura convencional. Devido à natureza dinâmica do processo, podem ocorrer falhas na distribuição do pó, como camadas de espessuras desiguais, devido a entupimentos, paradas no transportador, entre outros problemas.

O indicador que caracteriza a distribuição do pó na peça é a espessura da camada de

tinta. Essa espessura é monitorada por um aparelho de medição chamado eucômetro, e é expressa em micrômetros (μm). Os limites de espessura especificados em projeto variam de 30 a 50 micrômetros. Camadas de tinta com espessura abaixo do limite mínimo estão sujeitas à oxidação precoce da peça, enquanto camadas com espessura acima do limite máximo estão sujeitas à quebra da tinta.

Para monitorar a distribuição do pó, foram coletadas amostras consecutivas a cada duas horas, consistindo em duas laterais de fogão. O procedimento de medição adotado foi manter o mesmo utilizado pela empresa para cada lateral, sendo seis medições realizadas para cada posição: superior, média e inferior. Para a construção do GG, foram utilizadas as médias das medições realizadas em cada posição de cada lateral.

Tabela 5.1: Tabela com o conjunto de dados "Processo de pintura a pó de fogões"

| Grupo | Amostra | Posição | A1 | A2 | B1 | B2 | \bar{X}_1 | \bar{X}_2 | \bar{X} | Amplitude |
|-------|---------|----------|----|----|----|----|-------------|-------------|-----------|-----------|
| 1 | 1 | Superior | 32 | 30 | 30 | 35 | 31,00 | 32,50 | 31,75 | 1,50 |
| | 2 | Meio | 34 | 30 | 33 | 32 | 32,00 | 32,50 | 32,25 | 0,50 |
| | 3 | Inferior | 32 | 30 | 30 | 32 | 31,00 | 31,00 | 31,00 | 0,00 |
| 2 | 1 | Superior | 33 | 33 | 36 | 35 | 33,00 | 35,50 | 34,25 | 2,50 |
| | 2 | Meio | 37 | 41 | 33 | 35 | 39,00 | 34,00 | 36,50 | 5,00 |
| | 3 | Inferior | 35 | 40 | 35 | 31 | 37,50 | 33,00 | 35,25 | 4,50 |
| 3 | 1 | Superior | 37 | 31 | 36 | 34 | 34,00 | 35,00 | 34,50 | 1,00 |
| | 2 | Meio | 30 | 32 | 30 | 30 | 31,00 | 30,00 | 30,50 | 1,00 |
| | 3 | Inferior | 34 | 33 | 31 | 35 | 33,50 | 33,00 | 33,25 | 0,50 |
| 4 | 1 | Superior | 30 | 30 | 32 | 34 | 30,00 | 33,00 | 31,50 | 3,00 |
| | 2 | Meio | 32 | 30 | 32 | 30 | 31,00 | 31,00 | 31,00 | 0,00 |
| | 3 | Inferior | 31 | 30 | 30 | 31 | 30,50 | 30,50 | 30,50 | 0,00 |
| 5 | 1 | Superior | 38 | 34 | 39 | 36 | 36,00 | 37,50 | 36,75 | 1,50 |
| | 2 | Meio | 30 | 35 | 34 | 34 | 32,50 | 34,00 | 33,25 | 1,50 |
| | 3 | Inferior | 32 | 35 | 35 | 37 | 33,50 | 36,00 | 34,75 | 2,50 |
| 6 | 1 | Superior | 43 | 35 | 39 | 40 | 39,00 | 39,50 | 39,25 | 0,50 |
| | 2 | Meio | 36 | 33 | 37 | 36 | 34,50 | 36,50 | 35,50 | 2,00 |
| | 3 | Inferior | 33 | 39 | 33 | 35 | 36,00 | 34,00 | 35,00 | 2,00 |
| 7 | 1 | Superior | 31 | 30 | 37 | 32 | 30,50 | 34,50 | 32,50 | 4,00 |
| | 2 | Meio | 32 | 31 | 32 | 30 | 31,50 | 31,00 | 31,25 | 0,50 |
| | 3 | Inferior | 36 | 33 | 32 | 30 | 34,50 | 31,00 | 32,75 | 3,50 |
| 8 | 1 | Superior | 34 | 38 | 36 | 38 | 36,00 | 37,00 | 36,50 | 1,00 |
| | 2 | Meio | 32 | 40 | 37 | 38 | 36,00 | 37,50 | 36,75 | 1,50 |
| | 3 | Inferior | 40 | 37 | 37 | 40 | 38,50 | 38,50 | 38,50 | 0,00 |
| 9 | 1 | Superior | 31 | 35 | 38 | 37 | 33,00 | 37,50 | 35,25 | 4,50 |
| | 2 | Meio | 31 | 34 | 34 | 38 | 32,50 | 36,00 | 34,25 | 3,50 |
| | 3 | Inferior | 34 | 38 | 31 | 31 | 36,00 | 31,00 | 33,50 | 5,00 |
| 10 | 1 | Superior | 39 | 39 | 39 | 34 | 39,00 | 36,50 | 37,75 | 2,50 |
| | 2 | Meio | 39 | 34 | 33 | 34 | 36,50 | 33,50 | 35,00 | 3,00 |
| | 3 | Inferior | 36 | 34 | 34 | 31 | 35,00 | 32,50 | 33,75 | 2,50 |
| 11 | 1 | Superior | 38 | 34 | 30 | 30 | 36,00 | 30,00 | 33,00 | 6,00 |
| | 2 | Meio | 33 | 31 | 31 | 30 | 32,00 | 30,50 | 31,25 | 1,50 |
| | 3 | Inferior | 34 | 35 | 38 | 37 | 34,50 | 37,50 | 36,00 | 3,00 |

Fonte: Alexandre *et al.* (2006)

5.1 Elaboração do GG

Foi empregado o método mostrado no Capítulo 3 para a elaboração do gráfico utilizando a base de dados exposta na Tabela 5.1. Este procedimento permitiu a representação visual dos dados analisados, visando destacar padrões, tendências ou relações relevantes identificadas durante a análise.

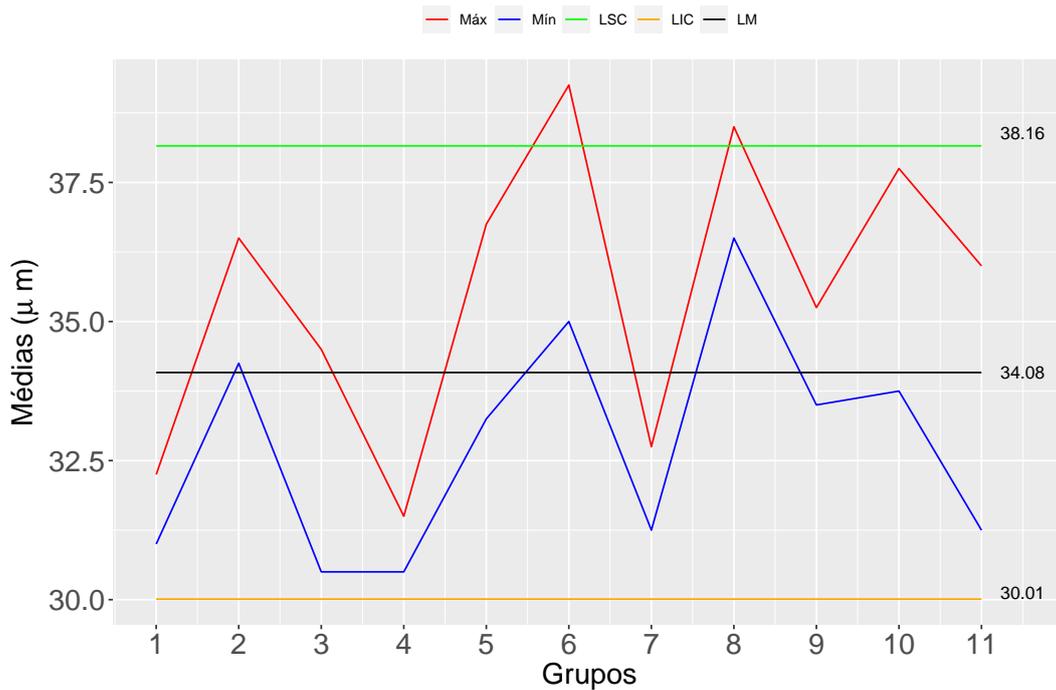


Figura 5.1: Gráfico GG para o Processo de pintura a pó de fogões

Observando a Figura 5.1, destaca-se que apenas os Grupos 6 e 8 apresentam pontos que ultrapassam os limites de controle, especificamente nos pontos máximos. Tal observação sugere que o processo em questão não está completamente fora de controle, sugere-se a possibilidade de que problemas relacionados à medição ou falhas nos equipamentos possam estar ocorrendo. Assim, recomenda-se uma análise mais detalhada e a realização de eventuais reparos nos equipamentos de medição para garantir a precisão e confiabilidade das medições realizadas.

5.2 NMA_1 para a Aplicação

Para avaliar o Gráfico por Grupos (GG) utilizado em 5.1, foi realizado o cálculo do NMA_1 utilizando a expressão fornecida na Equação 2.7:

$$E(Y) = \frac{1}{P(ETI)} = NMA,$$

onde $P(ETI) = P(\bar{X} \notin [LIC, LSC] | \mu = \mu_0)$.

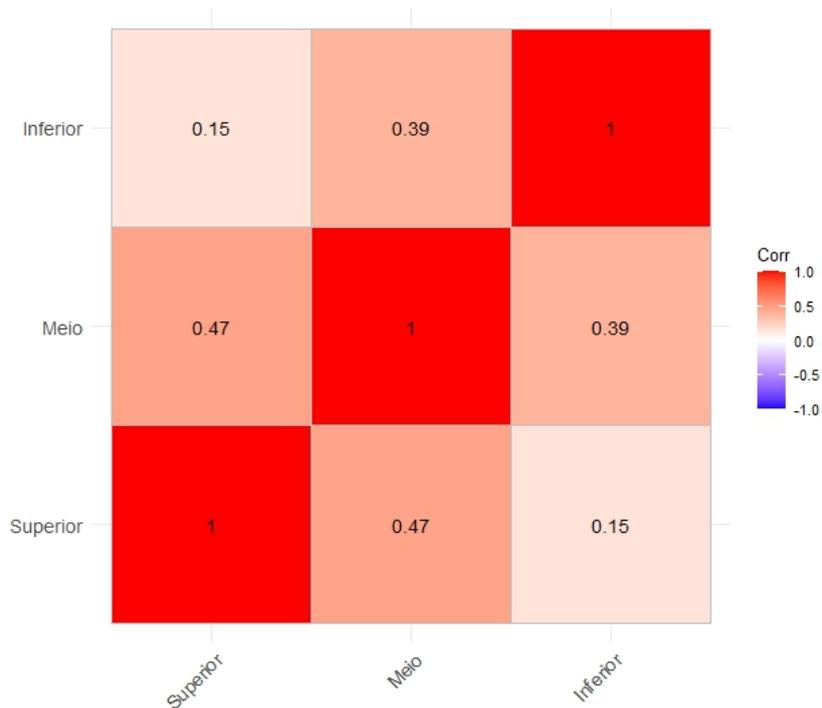


Figura 5.2: Gráfico da correlação entre os fluxos (posições) do processo de pintura a pó de fogões.

Para calcular o NMA_1 utilizando o método apresentado no Capítulo 3, é necessário estimar a matriz de correlação, conforme ilustrado na Figura 5.2. Analisando a Figura 5.2, observa-se uma correlação considerável entre as variáveis, sugerindo que uma análise presumindo independência pode não ser a abordagem mais adequada. Utilizando o método descrito no Capítulo 3, foi realizada a estimação da densidade do máximo e do mínimo para calcular o NMA_1 .

O NMA_1 foi calculado em dois cenários distintos: o primeiro considerando a matriz de correlação fornecida na Figura 5.2 e o segundo pressupondo a ausência de correlação entre os fluxos. Isto é, supondo covariância nula entre as posições. No primeiro caso, o

NMA_1 resultou em 3,57, enquanto no segundo cenário foi obtido um valor de 4,22. Esses resultados evidenciam que a suposição de independência não era a mais apropriada, já que o primeiro cenário sinalizou um alarme falso mais rápido, com um NMA_1 menor em relação ao cenário com independência. Em um cenário real, como o da aplicação apresentada no Capítulo 5, o cenário com dependência é o mais comum, por isso a sinalização mais rápida neste cenário é de extrema importância, podendo reduzir perdas com produtos fora do padrão exigido na indústria.

5.3 Alteração na média do processo (δ)

Nesta seção, faremos uso de um resultado obtido no Capítulo 3, em que foi estabelecido que a partir do vetor de médias μ e da matriz de covariância Σ de uma distribuição normal multivariada, é possível obter a distribuição do máximo de variáveis aleatórias dependentes. Esta distribuição é fundamental para calcular o NMA da variável de interesse. O NMA é calculado utilizando a Equação 2.4, em que os limites são determinados a partir do gráfico por grupos apresentado na Seção 5.1.

A Tabela 5.2 foi feita com alterações na média, $\mu_1 = \mu + \delta\sigma$, com δ variando entre, 0,1;0,3;0,5;1;2;3 e o tamanho da amostra n está variando entre 2,3 e 4.

Tabela 5.2: NMA_{21} para o Processo de pintura a pó de fogões.

| n | δ | | | | | |
|-----|----------|------|------|------|------|------|
| | 0,1 | 0,3 | 0,5 | 1 | 2 | 3 |
| 2 | 3,42 | 3,14 | 2,90 | 2,41 | 1,77 | 1,42 |
| 3 | 1,70 | 1,62 | 1,55 | 1,39 | 1,19 | 1,08 |
| 4 | 1,42 | 1,36 | 1,31 | 1,21 | 1,08 | 1,01 |

Analisando os dados apresentados na Tabela 5.2, observa-se uma relação entre a magnitude da alteração e a eficácia do NMA_{21} . Constatou-se que à medida que a alteração no vetor de médias μ aumenta, a capacidade de detecção de alarmes verdadeiros pelo NMA_{21} também aumenta, indicando uma detecção mais rápida de alarmes verdadeiros.

Um comportamento similar acontece com o tamanho da amostra, conforme o tamanho da amostra aumenta, a capacidade de detecção do NMA_{21} também aumenta, resultando em uma detecção mais rápida de alarmes verdadeiros. Este fenômeno sugere uma relação positiva entre o tamanho da amostra e a eficácia do método de detecção baseado no NMA_{21} .

5.4 Alteração na variância do processo (λ)

Semelhante à Seção 5.3, neste tópico a alteração será na variância do processo (considerando a média sem deslocamento), $\sigma^{*2} = \sigma^2 \lambda^2$. A Tabela 5.3 foi feita com λ variando entre 1,5; 2; 3; 5 e n variando entre 2, 3 e 4.

Tabela 5.3: NMA_{22} para o Processo de pintura a pó de fogões.

| n | λ | | | |
|-----|-----------|------|------|------|
| | 1,5 | 2,0 | 3,0 | 5,0 |
| 2 | 2,64 | 2,26 | 1,92 | 1,66 |
| 3 | 1,58 | 1,50 | 1,41 | 1,34 |
| 4 | 1,37 | 1,32 | 1,28 | 1,24 |

Comparativamente ao NMA_{21} , observou-se uma detecção mais ágil ao utilizar valores maiores de λ e n . Notadamente, a detecção torna-se mais rápida à medida que esses parâmetros aumentam, sendo evidente que o tamanho da amostra n possui uma influência relevante na capacidade de detecção.

5.5 Alteração conjunta na média e na variância do processo

Nesta seção a alteração será na média $\mu_1 = \mu\sigma + \delta$ e variância do processo $\sigma^{*2} = \sigma^2 \lambda^2$ conjuntamente. A Tabela 5.4 foi elaborada com λ variando entre 1,5; 2; 3; 5 e δ variando entre 0, 1; 0,3; 0,5; 1; 2; 3.

Tabela 5.4: NMA_{23} para o Processo de pintura a pó de fogões.

| n | δ | λ | | | |
|-----|----------|-----------|------|------|------|
| | | 1,5 | 2 | 3 | 5 |
| 2 | 0,1 | 2,56 | 2,20 | 1,88 | 1,65 |
| | 0,3 | 2,41 | 2,11 | 1,82 | 1,61 |
| | 0,5 | 2,28 | 2,02 | 1,77 | 1,58 |
| | 1 | 2,00 | 1,82 | 1,65 | 1,50 |
| | 2 | 1,61 | 1,53 | 1,45 | 1,38 |
| | 3 | 1,37 | 1,34 | 1,31 | 1,28 |
| 3 | 0,1 | 1,55 | 1,47 | 1,40 | 1,33 |
| | 0,3 | 1,50 | 1,43 | 1,37 | 1,31 |
| | 0,5 | 1,45 | 1,39 | 1,34 | 1,29 |
| | 1 | 1,34 | 1,31 | 1,28 | 1,25 |
| | 2 | 1,19 | 1,18 | 1,18 | 1,17 |
| | 3 | 1,09 | 1,10 | 1,11 | 1,12 |
| 4 | 0,1 | 1,35 | 1,31 | 1,27 | 1,23 |
| | 0,3 | 1,31 | 1,28 | 1,24 | 1,21 |
| | 0,5 | 1,27 | 1,25 | 1,22 | 1,20 |
| | 1 | 1,20 | 1,19 | 1,17 | 1,16 |
| | 2 | 1,09 | 1,09 | 1,10 | 1,11 |
| | 3 | 1,02 | 1,03 | 1,05 | 1,06 |

Observa-se, a partir da Tabela 5.4, um padrão semelhante entre os comportamentos do NMA_{21} e do NMA_{22} , evidenciando a influência significativa de todos os parâmetros na análise. É notável que à medida que os valores de n , δ e λ aumentam, ocorre uma diminuição do NMA_{23} , indicando uma detecção mais sensível de causas especiais no processo. Isto é altamente benéfico, dado que tais amostras estão fora dos limites do gráfico por grupos.

Um aspecto interessante a respeito dos NMA_{21} e NMA_{22} é a sua capacidade de detecção mais rápida, o que era previsto, considerando que esses parâmetros exercem uma influência direta nos resultados dos NMA. Esta detecção precoce sugere uma resposta mais ágil do método diante de alterações nas condições do processo, ressaltando sua utilidade em monitoramentos industriais.

Capítulo 6

Considerações Finais

Após uma análise dos métodos de monitoramento de processos, o método proposto neste estudo, o gráfico de Shewhart e o método Monte Carlo apresentado em Oliveira (2013) temos que cada abordagem tem vantagens e desvantagens. Por exemplo, o gráfico de Shewhart é amplamente difundido e compreendido, demonstrando eficácia em processos com independência entre as observações, o GG oferece uma alternativa para o monitoramento de múltiplos fluxos de produção, permitindo uma análise mais abrangente e eficiente, especialmente em processos com dependência entre as variáveis. Dito isto, temos que o método apresentado neste estudo revelou-se competitivo em cenários com independência entre os fluxos, fornecendo uma ferramenta robusta para estimar a densidade do máximo em processos com correlações entre as variáveis, além de possibilitar a realização de simulações de cenários diversos. A comparação entre esses métodos destaca a importância de selecionar a abordagem mais adequada para cada contexto específico de monitoramento de processos, levando em consideração a natureza dos dados e os objetivos da análise. Uma aplicação prática utilizando o método proposto neste estudo feita no Capítulo 5, no qual foram alcançados resultados superiores aos obtidos ao considerar a suposição de independência entre os fluxos, corroborando com a suposição que existe uma alta probabilidade de correlação entre os fluxos em cenários reais.

6.1 Conclusões

Diante dos objetivos estabelecidos no Capítulo 1, as etapas realizadas forneceram um método para determinar a densidade do máximo de variáveis com distribuição normal

considerando a dependência entre os fluxos produtivos, conforme abordado no Capítulo 3, representa um avanço importante na análise estatística desses processos. Este método, baseado na estimação da densidade do máximo, permite uma abordagem mais robusta para o monitoramento de processos, levando em consideração a correlação entre as variáveis envolvidas.

A comparação dos resultados obtidos com o método desenvolvido neste estudo com os já existentes, apresentados no Capítulo 4, proporcionou uma compreensão da eficácia e aplicabilidade dessa nova abordagem. A análise comparativa entre o método proposto, o gráfico de Shewhart e o GG destacou as vantagens e limitações de cada técnica, contribuindo para uma compreensão mais abrangente do monitoramento de processos em ambientes industriais e produtivos.

O Capítulo 5 apresentou uma aplicação prática do método desenvolvido, utilizando dados reais sobre o processo de pintura a pó de fogões. Esta aplicação demonstrou a eficácia e relevância do método em um contexto industrial real, fornecendo uma percepção para o monitoramento e controle de processos em diferentes setores produtivos. A análise detalhada dos resultados obtidos nesta aplicação destaca a capacidade do método proposto em lidar com a complexidade dos processos industriais, considerando não apenas a média e a variância, mas também a correlação entre as variáveis envolvidas.

Em síntese, este estudo oferece uma abordagem inovadora e eficaz para lidar com a dependência entre os fluxos produtivos.

6.2 Trabalhos Futuros

Para futuras pesquisas, sugere-se a realização de estudos mais aprofundados em cenários multivariados, considerando variações na correlação entre as variáveis. Esta área oferece diversas oportunidades para avançar no entendimento do comportamento de processos complexos e para desenvolver técnicas de monitoramento mais eficazes. Algumas possíveis direções para trabalhos futuros incluem:

Modelagem da Correlação: explorar métodos de modelagem mais avançados para capturar a estrutura de correlação entre as variáveis em cenários multivariados;

Desenvolvimento de Métodos de Monitoramento: propor novas abordagens para o monitoramento de processos em cenários multivariados, levando em consideração a vari-

ação da correlação.

Essas são apenas algumas das possíveis direções para futuras pesquisas na área de monitoramento de processos em cenários multivariados. Ao explorar essas direções, espera-se contribuir para o desenvolvimento de metodologias mais robustas e eficazes para o monitoramento e controle de processos industriais complexos.

Referências Bibliográficas

- Alexandre, J. W. C., Rodrigues, L. F. M., Rodrigues, F. P. & Freitas, S. M. d. (2006). Aplicação do gráfico de controle por grupos em uma indústria manufatureira do estado do ceará.
- Amiri, A. & Solis, A. A. (2018). An overview of statistical process control applications in the automotive industry. *Quality Engineering*, **30**(3), 317–332.
- Arellano-Valle, R. B. & Genton, M. G. (2008). On the exact distribution of the maximum of absolutely continuous dependent random variables. *Statistics & Probability Letters*, **78**(1), 27–35.
- Banks, J. & Carson, J. S. (1984). *Discrete-Event System Simulation*. Prentice-Hall.
- Boyd, D. (1950). Applying the group chart for \bar{X} and r . *Industrial Quality Control*, **7**, 22–25.
- Bussab, W. O. & Morettin, P. A. (2017). *Estatística Básica*. Saraiva.
- Casella, G. & Berger, R. L. (2021). *Statistical inference*. Cengage Learning.
- Costa, A., Epprecht, E. & Carpinetti, L. (2005). Controle estatístico de qualidade. 2ª edição. São Paulo: Editora Atlas.
- Fang, K. W. (2018). *Symmetric multivariate and related distributions*. CRC Press.
- Fishman, G. S. (2013). *Monte Carlo: Concepts, Algorithms, and Applications*. Springer Science & Business Media.
- Garvin, D. A. (1988). *Managing quality: The strategic and competitive edge*. Simon and Schuster.
- Guimarães, N. & Epprecht, E. (2000). Uma proposta prática para o controle estatístico de processos com correlação. *Anais do ENEGEP*.
- James, G., Witten, D., Hastie, T. & Tibshirani, R. (2011). Understanding correlation - bivariate data analysis for statistics. *Journal of Machine Learning Research*, **12**, 2319–2349.

- Law, A. M. & Kelton, W. D. (2000). *Simulation Modeling and Analysis*. McGraw-Hill, third edition.
- Levine, D. M., Szabat, K. A. & Stephan, D. F. (2014). *Estatística Aplicada à Administração e Economia*. Bookman.
- Li, R. & Tsung, F. (2015). Statistical process control of multivariate processes. *Quality and Reliability Engineering International*, **31**(7), 1087–1100.
- Montgomery, D. C. (2000). *Introdução Ao Controle Estatístico Da Qualidade .. Grupo Gen-LTC*.
- Montgomery, D. C. & Mastrangelo, C. M. (2015). Statistical process control for univariate and multivariate applications. *Quality and Reliability Engineering International*, **31**(8), 1421–1435.
- Mood, A. M., Graybill, F. A. & Boes, D. C. (1974). *Introduction to the Theory of Statistics*. McGraw-Hill International Editions, hird revised edition.
- Oliveira, M. B. d. (2013). Eficiência entre os gráficos de controle por grupos para a média e o tradicional de shewhart em processos com fluxos correlacionados.
- R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Ramos, A. W. (2000). *CEP para processos contínuos e em bateladas*. Editora Blucher.
- Rubinstein, R. Y. & Kroese, D. P. (2016). *Simulation and the Monte Carlo Method*. John Wiley & Sons.
- Shewhart, W. A. (1931). Statistical method from an engineering viewpoint. *Journal of the American Statistical Association*, **26**(175), 262–269.
- Vasconcelos, A. P. d. (2009). Estudo do gráfico de controle por grupos como uma alternativa ao modelo de shewhart em processos paralelos.