

UNIVERSIDADE FEDERAL DE ITAJUBÁ

RENAN RYUJI MURASHITA TAKENAKA

ESTUDO COMPARATIVO ENTRE O GRÁFICO DE CONTROLE  
SIMULTÂNEO (X-BARRA,S) E OS GRÁFICOS DE SHEWHART PARA  
MÉDIA E DESVIO-PADRÃO

ITAJUBÁ

2023

RENAN RYUJI MURASHITA TAKENAKA

ESTUDO COMPARATIVO ENTRE O GRÁFICO DE CONTROLE  
SIMULTÂNEO (X-BARRA,S) E OS GRÁFICOS DE SHEWHART PARA  
MÉDIA E DESVIO-PADRÃO

Dissertação submetida ao Programa de Pós-Graduação  
em Engenharia de Produção da Universidade Federal  
de Itajubá como parte dos requisitos para obtenção do  
título de Mestre em Engenharia de Produção.

Linha de Pesquisa: Qualidade e Produtos

Orientador: Prof. Dr. Paulo Henrique da Silva Campos

Coorientador: Prof. Dr. Anderson Paulo de Paiva

ITAJUBÁ

2023

Dedico este trabalho aos meus pais, irmão, tia Eloisa e minha namorada Milena pelo carinho, apoio, compreensão e inspiração durante todo o período do programa de mestrado.

## **AGRADECIMENTOS**

Em primeiro lugar, agradeço a Deus por permitir toda minha jornada até o presente momento, concedendo saúde e força necessárias para atingir meus objetivos.

Agradeço meus pais, Edilene e Valter, meu irmão Fernando e minha tia Eloisa, pelo apoio, carinho, compreensão e inspiração durante todo o meu período dedicado ao programa de mestrado. Sem essa força, tudo seria muito mais difícil.

Agradeço a minha namorada Milena por estar ao meu lado, sempre me incentivando e apoiando minhas decisões para alcançar meus objetivos.

Agradeço ao meu orientador, Prof. Dr. Paulo Henrique da Silva Campos, por toda a orientação e paciência, ao Prof. Dr. Anderson Paulo de Paiva, por aceitar a coorientação deste trabalho, enriquecendo com o seu conhecimento.

Agradeço à banca examinadora pelas valiosas contribuições na qualificação, a partir de reflexões e direcionamentos importantes para minha formação.

Agradeço a UNIFEI que permitiu a expansão do meu conhecimento e crescimento da minha carreira. Sou grato a todos os professores por todo o conhecimento compartilhado.

## RESUMO

Os gráficos de controle são ferramentas do Controle Estatístico de Processos amplamente utilizados para monitorar as características de qualidade chaves de um produto ou serviço. Por serem extremamente eficientes no monitoramento e resultarem em melhorias expressivas de qualidade, a pesquisa se faz necessária na investigação de diferentes abordagens. Entretanto, a diversidade de gráficos e metodologias de aplicação dificultam a escolha da melhor combinação para implementação em um processo. Além disso, alguns trabalhos apresentam pouca ou nenhuma comparação com outros trabalhos, dificultando o processo de escolha. O gráfico simultâneo (X-barra,S) é um caso em que as comparações apresentadas deixam dúvidas quanto à eficácia em relação até mesmo aos gráficos de Shewhart para média (X-barra) e desvio-padrão (S). Dessa forma, este trabalho visa realizar a comparação entre os gráficos de controle de Shewhart para média (X-barra) e desvio-padrão (S) e o gráfico simultâneo (X-barra,S) quanto a facilidade em identificar o estado do processo, a performance com o número médio de amostras até alarme (NMA) e a capacidade em determinar o momento em que o desvio ocorreu na série amostral. Para isso, foram apresentados os procedimentos de construção dos gráficos de controle, simulados os dados para cálculo do NMA e realizado a avaliação visual das informações presentes nos gráficos. Como resultado, concluiu-se que as inovações apresentadas no gráfico simultâneo (X-barra,S) geram vantagens inferiores aos tradicionais gráficos de Shewhart.

**Palavras-chave:** Gráfico de controle por variáveis Gráfico de controle simultâneo. Monitoramento de processos.

## ABSTRACT

Control charts are tools from Statistical Process Control widely used to monitor key quality characteristics of a product or service. For being extremely efficient in monitoring and improving quality, research is needed to investigate different approaches. However, the diversity of control charts and application methodologies makes the selection of the best combination to implement in a process harder. Furthermore, some works shows a few or no comparison between related subjects, making the selection even harder. The simultaneous ( $\bar{X}$ ,S) control chart is a case where the presented comparison raise doubts about its effectiveness even compared to Shewhart charts for mean ( $\bar{X}$ ) and standard deviation (S). Therefore, this study aims to compare Shewhart control charts for mean ( $\bar{X}$ ) and standard deviation (S) with the simultaneous ( $\bar{X}$ ,S) chart in terms of ease of identifying the process state, performance with the average run length (ARL), and ability to determine when the deviation occurred in the sample series. To do so, the procedures for constructing control charts were presented, data were simulated to calculate the NMA, and a visual evaluation of the information presented in the charts was carried out. As a result, the innovations presented in the simultaneous ( $\bar{X}$ ,S) chart shows inferior advantages compared to traditional Shewhart charts.

**Keywords:** Control charts for variables. Simultaneous control chart. Process monitoring.

## LISTA DE ILUSTRAÇÕES

Figura 1 - Quantidade de artigos publicados por ano contendo a palavra Control Chart.....	13
Figura 2 - Quantidade de artigos publicados por ano contendo a palavra Control Chart.....	15
Figura 3 - Gráfico de controle por variáveis da média amostral (X-barra) .....	16
Figura 4 - Representação do gráfico de controle (X-barra,S).....	23
Figura 5 - Fluxograma de execução da simulação deste trabalho .....	28
Figura 6 - Gráficos de controle X-barra e S para cada valor de n .....	30
Figura 7 - Projeção dos limites das Equações 26 a 31 .....	32
Figura 8 - Gráficos de controle simultâneo (X-barra,S) para cada valor de n.....	33
Quadro 1 – Resumo dos resultados obtidos pelas comparações .....	36

## LISTA DE TABELAS

Tabela 1 - Valores prefixados para $c_4$ .....	29
Tabela 2 - Valores médios e limites de controle dos gráficos de Shewhart .....	29
Tabela 3 - Valores de $\lambda(\alpha)$ para $n$ e $\alpha=0,0027$ .....	31
Tabela 4 - Valores de NMA dos gráficos estudados de acordo com o $n$ e desvios da média e desvio-padrão.....	34

## SUMÁRIO

<b>1 INTRODUÇÃO .....</b>	<b>10</b>
<b>1.1 Contextualização.....</b>	<b>10</b>
<b>1.2 Objetivos.....</b>	<b>12</b>
1.2.1 Objetivo Geral .....	12
1.2.2 Objetivos Específicos .....	12
<b>1.3 Justificativa .....</b>	<b>13</b>
<b>1.4 Limitações e delimitações.....</b>	<b>14</b>
<b>2 REVISÃO BIBLIOGRÁFICA .....</b>	<b>14</b>
<b>2.1 Controle Estatístico de Processo (CEP).....</b>	<b>15</b>
<b>2.2 Gráfico de Controle .....</b>	<b>16</b>
2.2.1 Variabilidade do processo .....	17
2.2.2 Gráfico de Controle por Variáveis.....	18
2.2.2.1 Gráfico de Controle da Média Amostral (X-barra) .....	19
2.2.3 Gráfico de Controle Simultâneo (X-barra,S).....	20
2.2.3.1 Método de Projeção do Limite de Controle do Gráfico de Controle (X-barra,S) .....	21
2.2.3.2 Método de Identificação de Condição Fora de Controle por AIC.....	22
<b>2.3 Teste de hipóteses .....</b>	<b>24</b>
<b>2.4 Fases de aplicação dos gráficos de controle .....</b>	<b>24</b>
<b>2.5 Cálculo do NMA por simulação de Monte Carlo .....</b>	<b>25</b>
<b>3 METODOLOGIA.....</b>	<b>26</b>
<b>3.1 Classificação do trabalho .....</b>	<b>26</b>
<b>3.2 Instrumentos de pesquisa e tratamento dos dados .....</b>	<b>27</b>
<b>4 APRESENTAÇÃO E ANÁLISE DE RESULTADOS .....</b>	<b>28</b>
<b>4.1 Construção dos gráficos de controle de Shewhart.....</b>	<b>28</b>
<b>4.2 Construção do gráfico de controle (X-barra,S) .....</b>	<b>31</b>
<b>4.3 Cálculo do NMA .....</b>	<b>33</b>

<b>5 CONCLUSÃO.....</b>	<b>35</b>
<b>REFERÊNCIAS .....</b>	<b>38</b>
<b>APÊNDICE A – CÓDIGO PARA CONSTRUÇÃO E CÁLCULO DO NMA DOS GRÁFICOS DE SHEWHART.....</b>	<b>40</b>
<b>APÊNDICE B – CÓDIGO PARA CONSTRUÇÃO E CÁLCULO DO NMA DO GRÁFICO SIMULTÂNEO (<math>\bar{X}</math>,S) .....</b>	<b>42</b>

# 1 INTRODUÇÃO

## 1.1 Contextualização

Gráficos de controle são ferramentas do Controle Estatístico de Processo (CEP) para monitoramento de processos amplamente utilizados com o objetivo de detectar rapidamente causas especiais de variabilidade. Aplicados mundialmente nas indústrias, são objetos muito atrativos de estudos em CEP e muitos pesquisadores têm trabalhado para desenvolver e investigar diferentes tipos de gráficos de controle.

Os gráficos de controle são representações visuais que visam monitorar processos de acordo com os requisitos do usuário, auxiliando na garantia e melhoria da qualidade dos produtos. Para isso, é necessário definir estados de um processo sob ou fora de controle, visto que o objetivo principal é auxiliar na redução da variabilidade das saídas do processo.

Shewhart (1926) foi o pioneiro no desenvolvimento das cartas de controle, afirmando a existência de causas aleatórias e especiais que impossibilitam a obtenção de saídas idênticas. Conforme o autor, a primeira refere-se àqueles fatores que influenciam positiva ou negativamente um determinado processo, de forma que não é possível separá-las (como vibrações, ruídos, perdas de energia, entre outros) porque há pouco ou nada a ser feito para removê-las do processo em questão. As causas especiais, por outro lado, são os fatores que também alteram as saídas, entretanto são conhecidas e existem formas de controlar ou removê-las do processo, além de possuírem um impacto indesejável superior às causas aleatórias.

As causas aleatórias de um processo geram saídas aceitáveis da característica de qualidade controlada e o investimento necessário para conter sua influência é muito alto pela pequena redução resultante, com a possibilidade de ser praticamente imperceptível. Visto que as causas especiais possuem um grande impacto no processo e que possuem formas de controlá-las, são alvos dos esforços de melhoria nos projetos de redução da variabilidade da característica monitorada.

Os gráficos de controle ganharam uma grande popularidade na pesquisa e na indústria por permitirem diversas abordagens e aplicações teóricas para os diversos tipos de monitoramento de processos. Além disso, a evolução dos gráficos de controle nas pesquisas tem gerado melhorias de qualidade significativas quando aplicadas corretamente. Montgomery (2013) cita cinco motivos à essa notoriedade: ser comprovada para melhoria da qualidade; serem efetivas na prevenção de defeitos; evitam ajustes desnecessários; fornecem informações de diagnóstico; e transmite a capacidade do processo.

Desde o gráfico de controle da média amostral proposto por Shewhart (1926), diversos outros foram desenvolvidos e aprimorados com o intuito de expandir a aplicação do monitoramento de processos. Entre os principais gráficos de controle, Costa, Epprecht e Carpinetti (2012) apresentam quatro principais classificações: por variáveis que é aplicado para monitorar características de qualidade cuja variável é contínua, como dimensões; processos autocorrelacionado, em que as características de qualidade possuem interdependência, como a temperatura de caldeiras; somas acumuladas (CUSUM) e média móvel ponderada exponencialmente (EWMA) que são indicados para monitoramento de processos sujeitos a pequenas perturbações, por exemplo calibrações de equipamentos, de forma que a decisão sobre o estado do processo se baseia na informação acumulada das amostras; e por atributos que monitora características de qualidade não numéricas, baseando-se em classificações como defeituoso e não defeituoso, como percentual de parafusos defeituosos.

Montgomery (2013) afirma que, em vários casos, as características de qualidade podem ser definidas por uma distribuição normal e, por isso, profissionais aplicam os gráficos da média ( $\bar{X}$ -barra) e variabilidade (R, S ou  $s^2$ ) em conjunto para monitorar as saídas do processo. Gan (1995) percebeu que a maioria dos gráficos de controle são projetados para monitorar um único parâmetro de processo, mas que é necessário utilizar os gráficos de média e variabilidade de forma conjunta para facilitar a percepção de alterações nos parâmetros.

McCracken e Chakraborti (2013) estudaram diversas estruturas de gráficos de controle para média e variabilidade, considerando esquemas de um ou dois gráficos. Entretanto, aqueles que possuíam uma modelagem para visualização simultânea em um único gráfico se mostraram mais promissores considerando aplicações reais.

Dentre os diversos trabalhos disponíveis para monitoramento de processos com esquemas de gráfico único, Takemoto e Arizono (2018) realizaram uma melhoria do gráfico ( $\bar{X}$ -barra,S) de Kanagawa, Arizono, Ohta (1997) que tem o objetivo de monitorar a média e o desvio-padrão do processo. Takemoto e Arizono (2018) implementaram um método para identificar um sistema fora de controle por meio do critério informativo de Kullback-Leibler, uma ferramenta utilizada para selecionar o melhor modelo matemático entre outros modelos estatísticos, definir áreas no gráfico que representem momentos de atenção e alerta com o critério informativo de Akaike (AIC), e acrescentar uma visualização informativa da série temporal desenhada no gráfico.

O trabalho de Takemoto e Arizono (2018) apresenta uma representação gráfica de fácil compreensão, pelo fato de sua projeção ser semelhante a uma folha de tiro ao alvo, porém possui um problema sobre a ordem temporal dos eventos porque o plano bidimensional impede uma

visualização intuitiva com o aumento de pontos inseridos. Além disso, os autores não apresentam uma comparação de performance com os tradicionais gráficos de Shewhart para média e desvio-padrão, tornando a eficácia questionável. Entre os indicadores de performance, o principal critério é o número médio de amostras até um alarme (NMA) porque representa quantas amostras precisam ser coletadas até que o próximo indique um estado fora de controle (YEGANEH e SHADMAN, 2021).

O presente trabalho realizará uma comparação entre o gráfico simultâneo (X-barra,S) de Takemoto e Arizono (2018) e os gráficos de controle tradicionais de média e desvio-padrão de Shewhart. Para isso, a obtenção das amostras para o cálculo do NMA foi feita por simulação de Monte Carlo, visto que é uma forma prática de se obter os valores teóricos necessários à construção dos gráficos e obtenção do NMA (SIMÕES, COSTA e MACHADO, 2020). Outros fatores considerados foram a complexidade de implementação dos gráficos e a facilidade em identificar o estado do sistema durante toda a série amostral.

## **1.2 Objetivos**

A realização do estudo possui objetivos gerais e específicos que visam, respectivamente, explicitar a ideia central do trabalho e delimitar os assuntos abordados do tema.

### **1.2.1 Objetivo Geral**

O trabalho tem como objetivo geral a comparação do gráfico de controle simultâneo (X-barra,S) com os gráficos de Shewhart para média e desvio-padrão para avaliar a performance e determinar aquele que melhor identifica alterações nos parâmetros monitorados.

### **1.2.2 Objetivos Específicos**

São objetivos específicos deste trabalho:

- Realizar uma revisão bibliográfica dos gráficos de controle (X-barra,S) e de Shewhart para média e desvio-padrão;
- Aplicar os gráficos de controle simultâneo (X-barra,S) e de Shewhart para média e desvio-padrão;

- Determinar o gráfico de controle que apresenta a mais eficácia no monitoramento de processos conforme o valor do NMA, complexidade de implementação e facilidade de análise.

### 1.3 Justificativa

Desde a sua origem no ano de 1926, diversos trabalhos sobre gráficos de controle têm sido publicados com novos modelos e métodos para monitorar diferentes características de qualidade, sempre buscando melhorar a eficiência do monitoramento do processo e, conseqüentemente, a qualidade das saídas.

Conforme a Figura 1, houve um crescimento na quantidade de artigos publicados contendo as palavras *control chart* na base da Scopus (2022), sendo que o período levantado foi de 2011 a 2021. A ascensão dos gráficos de controle se deve à popularização de seus conceitos e os resultados promissores na busca por melhorarias na eficiência produtiva.

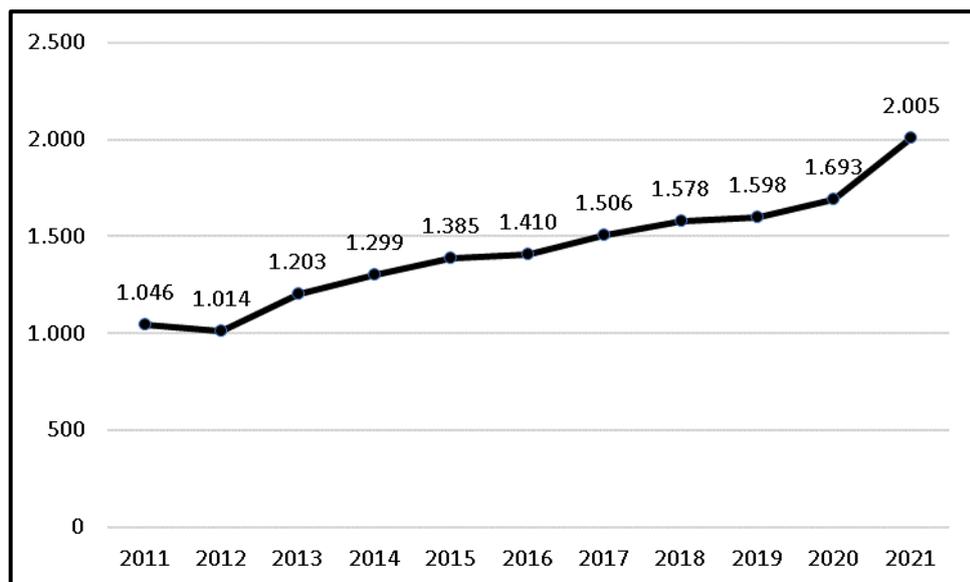


Figura 1 - Quantidade de artigos publicados por ano contendo a palavra Control Chart

Fonte: Scopus (2022), elaborado pelo autor

O aumento da variedade de gráficos de controle com diversos métodos e modelos dificultam a escolha da combinação mais adequada à característica de qualidade a ser monitorada, visto que diversas teorias podem ser aplicadas na resolução de um mesmo problema.

Muitos dos trabalhos publicados apresentam uma perspectiva inovadora e bastante interessante, mas existem aqueles que realizam poucas ou nenhuma comparação com modelos já aceitos pela comunidade, aumentando as dúvidas quanto à eficácia de sua proposta e, portanto, diminuem sua relevância na academia e nas indústrias.

Visto a importância de se realizar estudos comparativos de novas teorias com trabalhos efetivados pela comunidade científica e mercado, este trabalho se justifica por ampliar o conhecimento do gráfico de controle simultâneo (X-barra,S), apresentando padrões de comparação que auxiliem futuros estudos na escolha dos gráficos de controle a serem utilizados.

#### **1.4 Limitações e delimitações**

Este trabalho não possui restrições físicas e geográficas porque os cálculos e comparações serão realizados por simulação, utilizando a programação na linguagem Python.

O escopo do estudo é a comparação dos gráficos de controle simultâneo (X-barra,S) e os de Shewhart de acordo com o NMA obtido por simulação, o procedimento de construção e avaliação das informações visuais dos gráficos, como percepção temporal das amostras e a determinação do estado da característica de qualidade monitorada.

Visto que o gráfico de controle simultâneo (X-barra,S) de Takemoto e Arizono (2018) utiliza estimadores de máxima verossimilhança assumindo dados normalmente distribuídos, este trabalho realizou todas as simulações por meio do uso de dados que seguem uma distribuição normal.

Além disso, os gráficos serão avaliados como se estivessem na fase II, em que os parâmetros de construção dos gráficos de controle foram determinados quando o processo já se encontrava estável, reduzindo outras variáveis na realização das comparações. Gráficos de controle da fase I são geralmente utilizados na primeira implementação do monitoramento de processos, pois ainda está em etapa de ajustes e permite o uso de limites de controles testes.

Entretanto, o escopo desta pesquisa não inclui o cálculo teórico do NMA do gráfico simultâneo (X-barra,S) nem a determinação da probabilidade de alarmes falsos.

## **2 REVISÃO BIBLIOGRÁFICA**

Nesta seção, serão apresentados os principais conceitos necessários para atingir o objetivo proposto do trabalho.

## 2.1 Controle Estatístico de Processo (CEP)

O CEP é uma abordagem analítica utilizada tradicionalmente no monitoramento de processos de manufatura para diferenciar causas de variação aleatórias de especiais em qualquer sistema ao longo do tempo (KNAUER *et. al.*, 2018). Para isso, o CEP acompanha o comportamento das estatísticas das saídas da característica de qualidade desejada do sistema, auxiliando a tomada de decisão acerca das ações corretivas conforme causas especiais forem detectadas (CORRÊA e CORRÊA, 2011). Com isso, o CEP tem por finalidade monitorar a presença de não-conformidades que, por consequência, melhora a qualidade de produtos e serviços e reduz os custos de operação.

Montgomery (2013) define o CEP como uma coleção poderosa de ferramentas para solução de problemas, atingir estabilidade de processo e melhorar sua capacidade por meio da redução da variabilidade das saídas. Apesar de ser constituída basicamente de ferramentas, tem o papel de criar um ambiente em que todos os indivíduos busquem a melhoria contínua da qualidade e produtividade.

Entre as ferramentas do CEP, tem-se: histograma, folha de controle, gráfico de Pareto, diagrama de causa e efeito (ou diagrama de Ishikawa), diagrama de concentração de defeito, diagrama de dispersão e gráfico de controle – sendo que este último é o mais sofisticado tecnicamente em função da aplicação de conceitos estatísticos (MONTGOMERY, 2013).

A Figura 2 é uma representação adaptada de Montgomery (2013) que mostra as entradas e saídas de um processo produtivo, ilustrando o papel que as ferramentas podem exercer.

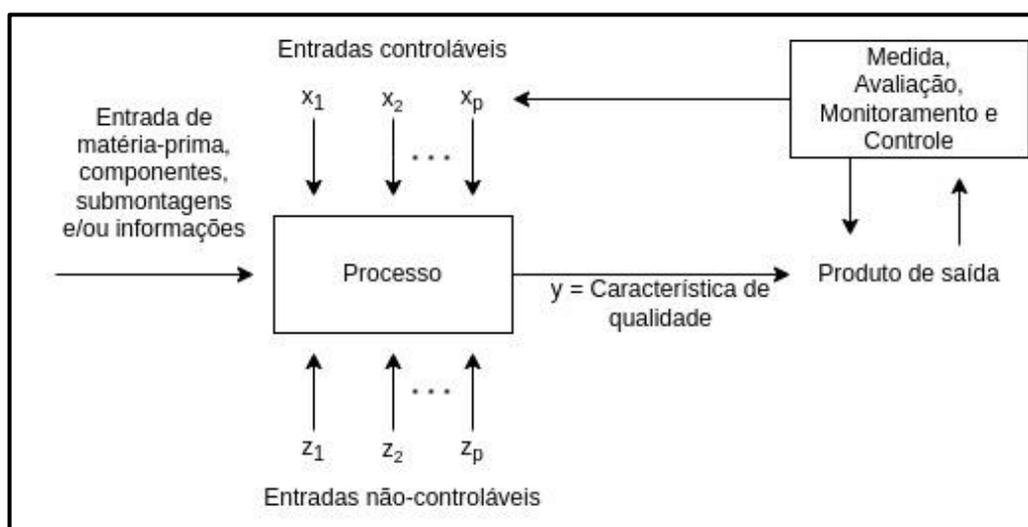


Figura 2 - Quantidade de artigos publicados por ano contendo a palavra Control Chart

Fonte: Adaptado de Montgomery (2013)

O processo se inicia com a entrada de matéria-prima, componentes, submontagens e/ou informações. Durante a execução, o processo sofre influência de entradas controláveis (x) e por entradas não-controláveis (z). As primeiras são variáveis que possuímos controle, como pressão e taxa de alimentação, enquanto a segunda é aquela que não possuímos controle, como fatores ambientais e propriedades de materiais fornecidos por terceiros. O processo realiza a transformação das entradas, resultando em um produto acabado constituído de diversas características de qualidade (y), cujas principais serão medidas, avaliadas, monitoradas e controladas. Durante essa etapa, as ferramentas do CEP são aplicadas provendo informações sobre quais são as entradas controláveis e não-controláveis e auxiliando na tomada de decisão quanto a necessidade de intervenção e em quais variáveis atuar a fim de manter o processo sob controle.

## 2.2 Gráfico de Controle

Os gráficos de controle são representações gráficas do CEP, cujo principal objetivo é monitorar a estabilidade de uma variável de processo, auxiliando a garantir e melhorar a qualidade de produtos e serviços.

Uma vez definida a média de uma característica de qualidade específica, a sua produção deve ser projetada de forma que as saídas associadas ao produto sejam iguais ou muito próximas à média. Se os produtos apresentarem variações na qualidade desejada, essa variação precisa estar dentro de um intervalo pré-definido (PRIYA e KANTAM, 2017). A Figura 3 é uma representação do gráfico proposto por Shewhart que visa monitorar a média amostral ao longo do tempo.

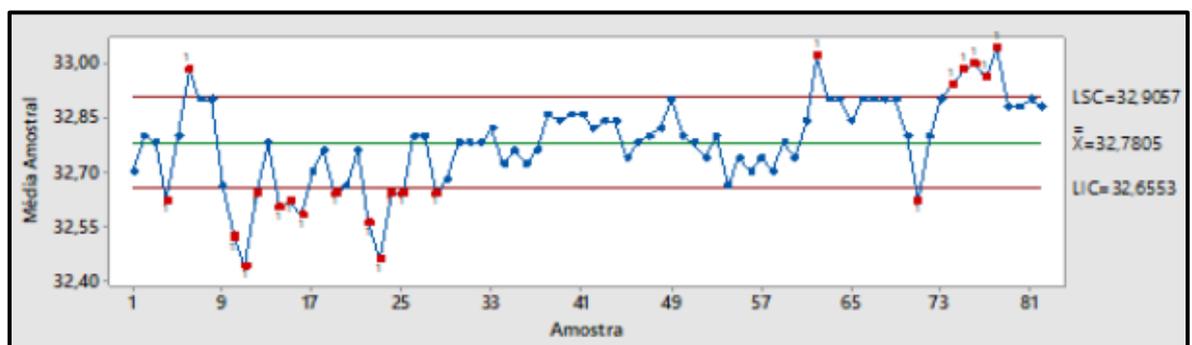


Figura 3 - Gráfico de controle por variáveis da média amostral (X-barra)

Fonte: Adaptado de Shewhart (1926)

Os pontos projetados no gráfico são médias amostrais de mesmo tamanho, denominados subgrupos racionais, e eles devem variar ao longo da linha verde, denominada Linha Média (LM), que representa o valor médio da característica de qualidade a ser monitorada e pode ser o valor definido em projeto ou então a média amostral do conjunto utilizado na elaboração do gráfico de controle. As linhas vermelhas representam os limites superior e inferior de controle (LSC e LIC, respectivamente), que são os valores cujos pontos não devem sair desse intervalo, porque configuram um processo fora de controle.

Shewhart (1924) publicou a definição de variabilidade do processo, conceito que justifica a necessidade de realizar o monitoramento dos processos por meio da estatística de suas saídas. Em 1926, Shewhart publicou os primeiros gráficos de controle por variáveis: de média (X-barra) e desvio-padrão (s) amostral.

Kanagawa, Arizono, Ohta (1997) identificaram a necessidade de unificar a visualização dos gráficos de média e desvio-padrão amostral em um único gráfico, sendo o principal objetivo agilizar a identificação do estado do sistema produtivo de forma imediata. Takemoto e Arizono (2018) aperfeiçoaram o gráfico de Kanagawa, Arizono, Ohta (1997) com a divisão de áreas que representam tendências de alterações nos parâmetros da média e desvio-padrão amostral, além de aplicarem conceitos de visualização informacional para obter a informação temporal da amostra projetada no gráfico.

### **2.2.1 Variabilidade do processo**

Em 1924, Shewhart questionou as variações existentes nas medições, pois era comum adotar um valor verdadeiro e valores diferentes serem considerados erros dos processos de medição. No seu trabalho, apontou que as teorias molecular e de quanta foram determinantes na alteração de alguns conceitos sobre medição, porque tornou necessário considerar a média e a variação das observações individuais, uma vez que a precisão é limitada por incertezas estatísticas.

Todo e qualquer processo, independentemente da adequação ao projeto ou do grau de controle, apresenta ao menos um componente responsável por gerar variabilidade nas saídas e que é impossível de ser removido (SHEWHART, 1924). Costa, Epprecht e Carpinetti (2012) defendem que esses componentes consistem na variabilidade natural do processo, sendo séries de pequenas perturbações contra as quais pouco ou nada se pode fazer, denominadas de causas aleatórias.

Por outro lado, os processos estão sujeitos a perturbações maiores, capazes de deslocar os valores esperados da distribuição de probabilidades esperada. Denominadas causas especiais, essas perturbações resultam em um estado anormal do processo e que, portanto, devem ser corrigidos ou eliminados (COSTA, EPPRECHT E CARPINETTI, 2012).

Uma vez conhecidas a distribuição de probabilidades esperada da característica de qualidade a ser monitorada, é possível determinar a média do processo e a variabilidade aceitável durante a inspeção do processo. Com essas informações, pode-se construir os gráficos de controle adequados para realizar o monitoramento do processo e, assim, aumentar a eficiência da operação.

### **2.2.2 Gráfico de Controle por Variáveis**

O gráfico de controle por variáveis é destinado ao monitoramento de processos que considera toda característica de qualidade que é refletida em por uma ou mais variáveis contínuas, ou seja, observam-se os valores de um conjunto de medidas numéricas, como pressão, volume, *lead time*, entre outros (LOUZADA, DINIZ, FERREIRA e FERREIRA, 2013). Os autores ainda defendem que os gráficos de controle por variáveis são os de maior destaque por causa da sua ampla aplicação em relação aos outros tipos.

Conforme Siqueira (1997), são três principais gráficos de controle por variáveis: o gráfico de X-barra, destinado ao estudo da média; o gráfico de R para monitoramento da amplitude dos dados; e o gráfico de S que acompanha o desvio-padrão para avaliação da variabilidade do processo.

Quando a característica de qualidade é uma variável contínua, geralmente são monitorados o seu valor médio e a sua variação de forma simultânea, sendo que a dispersão pode ser acompanhada pelo gráfico de R ou pelo gráfico de S. A escolha do gráfico para monitorar a variabilidade do processo é baseada no tamanho da amostra: se menor ou igual a 10, utiliza-se o gráfico de R, quando a amostra possui tamanho maior que 10, opta-se pelo gráfico de S.

Diversos trabalhos ainda são publicados com esse assunto, entretanto o artigo de Takemoto e Arizono (2018) se destacou por apresentar uma visualização simultânea aos gráficos de X-barra e S e que propõe facilitar a detecção de amostras que apresentem uma tendência de ter o parâmetro da média e/ou do desvio-padrão alterado ao longo do processo. Portanto, foi um dos gráficos escolhidos para este trabalho.

### 2.2.2.1 Gráfico de Controle da Média Amostral (X-barra)

Quando a característica de qualidade a ser monitorada é uma variável contínua, é essencial realizar o acompanhamento de sua centralidade ao longo do tempo. Para isso, o gráfico de controle da média amostral (X-barra) é o mais comum de ser utilizado (COSTA, EPPRECHT E CARPINETTI, 2012).

A linha média do gráfico de X-barra representa o valor esperado para a média das amostras do processo, enquanto os limites de controle são estabelecidos a três desvios padrão da média ( $\sigma$ ). A relação entre a variância ( $\sigma^2$ ) das observações individuais, o valor da média populacional ( $\mu$ ) e o tamanho da amostra ( $n$ ), é representada pela Equação 1:

$$\sigma^2 = \frac{\sigma^2}{n} \quad (1)$$

Sabendo-se que o desvio-padrão ( $\sigma$ ) é a raiz quadrada da variância, é possível obter a linha média ( $LM_{X\text{-barra}}$ ) e os limites superior e inferior de controle ( $LSC_{X\text{-barra}}$  e  $LIC_{X\text{-barra}}$ , respectivamente) por meio das Equações 2 a 4, nesta ordem:

$$LM_{X\text{-barra}} = \mu \quad (2)$$

$$LSC_{X\text{-barra}} = \mu + 3 \frac{\sigma}{\sqrt{n}} \quad (3)$$

$$LIC_{X\text{-barra}} = \mu - 3 \frac{\sigma}{\sqrt{n}} \quad (4)$$

De acordo com Costa, Epprecht e Carpinetti (2012), o processo está sob controle se os subgrupos racionais flutuarem em torno da linha média sem ultrapassar os limites de controle. Caso contrário, é necessário realizar uma análise de causa raiz para verificar se o subgrupo fora dos limites configurou um erro pontual ou se o processo realmente apresenta alterações na sua média.

### 2.2.2.2 Gráfico de Controle do Desvio-padrão Amostral (S)

Quando aplicado o gráfico de X-barra, é importante o acompanhamento da variabilidade do processo de forma conjunta. Isso se deve ao fato de que a centralidade dos subgrupos seja mantida apesar de apresentar uma dispersão muito grande (COSTA, EPPRECHT E CARPINETTI, 2012).

Um dos gráficos de controle por variáveis publicados por Shewhart (1926) foi o gráfico de desvio-padrão (S) que foi base para desenvolvimento do gráfico de Kanagawa, Arizono e Ohta (1997).

Os cálculos da linha média (LMS) e limites superior e inferior de controle (LSCS e LICs, respectivamente) podem ser realizados conforme as Equações 5 a 7, nesta ordem:

$$LM_S = c_4\sigma \quad (5)$$

$$LSC_S = c_4\sigma + 3\sigma\sqrt{1 - c_4^2} \quad (6)$$

$$LIC_S = \max \{0, c_4\sigma - 3\sigma\sqrt{1 - c_4^2}\} \quad (7)$$

Sendo  $\sigma$  o desvio-padrão da variável X-barra e  $c_4$  uma constante tabelada baseada no tamanho amostral  $n$ .

Assim como no gráfico de X-barra, Costa, Epprecht e Carpinetti (2012) afirmam que o processo se encontra sob controle se os subgrupos racionais flutuarem em torno da linha média e dentro dos limites de controle, pois o contrário resulta em alterações na dispersão do processo.

### 2.2.3 Gráfico de Controle Simultâneo (X-barra,S)

O gráfico de controle simultâneo (X-barra,S), desenvolvido por Kanagawa, Arizono e Ohta (1997), é um dos gráficos simultâneos para monitorar a média e o desvio-padrão, com o principal objetivo de julgar a condição do processo no momento da amostragem individual. Para isso, os autores representaram o gráfico em coordenadas X-barra-S e determinaram que os limites de controle seguiriam os critérios de informação de Kullback-Leibler, uma ferramenta para comparar a distância entre duas distribuições.

Com o intuito de determinar alguns parâmetros de processo responsáveis por um sinal de fora de controle no gráfico de controle (X-barra,S), Takemoto, Satoh e Arizono (2013) decidiram acrescentar critérios de informação de Akaike (AIC). Assim como o critério de informação de Kullback-Leibler, o AIC também realiza a comparação entre dois modelos estatísticos e, no caso, foi utilizado para determinar os limites dos modelos que caracterizam alterações na média e/ou desvio-padrão do processo.

O último trabalho publicado do gráfico de controle (X-barra,S) por Takemoto e Arizono (2018) identificou que, assim como outros gráficos de controle simultâneos, sua principal desvantagem é que a natureza de ordenação de tempo é perdida visualmente. Para contornar esse problema, foi implementado uma informação visual sobre o progresso temporal. Ainda nesse trabalho, os autores alocaram áreas de atenção, para mudanças na média e/ou desvio-padrão do processo, e alerta, para uma condição fora de controle por meio do AIC.

Nesse trabalho, não foi considerada a informação visual sobre o progresso temporal porque o objetivo principal é realizar uma comparação quanto à identificação de um sistema

fora de controle, além de que a solução apontada pelos autores não foi satisfatória para contornar problema da natureza de ordenação de tempo.

### 2.2.3.1 Método de Projeção do Limite de Controle do Gráfico de Controle (X-barra,S)

Adote que a respectiva característica de qualidade dos itens  $x_i$ ,  $i = 1, 2, \dots, n$  é normalmente distribuída, em que  $n$  é o tamanho da amostra. Então, defina que a condição sob controle do processo de qualidade seja  $N(\mu_0, \sigma_0^2)$  como condição alvo, em que  $N(\mu, \sigma^2)$  represente uma distribuição normal de média  $\mu$  e variância  $\sigma^2$ .

Empregando a estatística base do critério informativo de Kullback-Leibler, é possível discriminar entre as condições de processo sob e fora de controle pela Equação 8:

$$\lambda = \frac{n}{2} \left( \log \frac{\sigma_0^2}{s^2} - 1 + \frac{s^2}{\sigma_0^2} + \frac{(\bar{x} - \mu_0)^2}{\sigma_0^2} \right) \quad (8)$$

sendo  $\bar{X}$ -barra e  $S^2$  são estimadores de máxima verossimilhança como segue:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (9)$$

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (10)$$

Dessa forma, a regra de discriminação entre as condições sob e fora de controle de um processo é formulada a seguir:

$$\begin{cases} \text{se, } \lambda \leq \lambda(\alpha), \text{ então condição sob controle} \\ \text{senão, condição fora de controle} \end{cases} \quad (11)$$

onde  $\alpha$  é a probabilidade de erro do tipo I e  $\lambda(\alpha)$  é o percentil superior  $100\alpha$  da distribuição de  $\lambda$ . Assim,  $\lambda(\alpha)$  é dada pela seguinte equação:

$$\lambda(\alpha) = \frac{K_2}{2K_1} \chi_{\phi}^2(\alpha) \quad (12)$$

em que

$$\phi = \frac{2K_1^2}{K_2} \quad (13)$$

$$K_1 = \frac{-n}{2} \varphi^{(0)} \left( \frac{n-1}{2} \right) + \frac{n}{2} \log \frac{n}{2} \quad (14)$$

$$K_2 = \frac{n^2}{4} \varphi^{(1)} \left( \frac{n-1}{2} \right) - \frac{n}{2} \quad (15)$$

$\chi_{\phi}^2(\alpha)$  é o percentil superior  $100\alpha$  da distribuição de qui-quadrado com  $\phi$  graus de liberdade.  $K_1$  e  $K_2$  implicam a média e variância da estatística  $\lambda$  na condição sob controle. Além disso,  $\varphi^{(j)}(z)$  são funções polígama definidas utilizando a função gama  $\Gamma(z)$  como:

$$\varphi^{(j)}(z) = \frac{d^{j+1}}{dz^{j+1}} \log \Gamma(z) \quad (16)$$

Usualmente,  $\lambda(\alpha)$  é denominado limite de controle e obtido com o conjunto de pontos que satisfaçam a relação da Equação 17:

$$\lambda(\alpha) = \frac{n}{2} \left( \log \frac{\sigma_0^2}{s^2} - 1 + \frac{s^2}{\sigma_0^2} + \frac{(\bar{x} - \mu_0)^2}{\sigma_0^2} \right) \quad (17)$$

### 2.2.3.2 Método de Identificação de Condição Fora de Controle por AIC

Takemoto e Arizono (2018) expandiram o trabalho de Takemoto, Satoh e Arizono (2013) com o desenvolvimento de modelos estatístico para diferenciação das condições do processo utilizando as seguintes funções de densidade de probabilidade:

$$f(x; \mu; \sigma_0^2) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp \left\{ -\frac{(x-\mu)^2}{2\sigma_0^2} \right\} \quad (18)$$

$$f(x; \mu_0; \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(x-\mu_0)^2}{2\sigma^2} \right\} \quad (19)$$

$$f(x; \mu; \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(x-\mu)^2}{2\sigma^2} \right\} \quad (20)$$

$$f(x; \mu_0; \sigma_0^2) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp \left\{ -\frac{(x-\mu_0)^2}{2\sigma_0^2} \right\} \quad (21)$$

Os respectivos AICs dos modelos estatísticos definidos pelas 18 a 21 são:

$$AIC(\mu; \sigma_0^2) = n \log(2\pi\sigma_0^2) + \frac{ns^3}{\sigma_0^2} + 2 \quad (22)$$

$$AIC(\mu_0; \sigma^2) = n \log[2\pi\{s^2 + (\bar{x} - \mu_0)^2\}] + n + 2 \quad (23)$$

$$AIC(\mu; \sigma^2) = n \log(2\pi\sigma^2) + n + 4 \quad (24)$$

$$AIC(\mu_0; \sigma_0^2) = n \log(2\pi\sigma_0^2) + \frac{ns^2 + n(\bar{x} - \mu_0)^2}{\sigma_0^2} \quad (25)$$

Com as Equações 22 a 24, são obtidos os respectivos limites com as definições a seguir:

$$\log \left\{ \frac{s^2 + (\bar{x} - \mu_0)^2}{\sigma_0^2} \right\} + 1 - \frac{s^2}{\sigma_0^2} = 0 \quad (26)$$

$$n \log \left( \frac{s^2}{\sigma_0^2} \right) + n - \frac{ns^2}{\sigma_0^2} + 2 = 0 \quad (27)$$

$$n \log \left\{ \frac{s^2 + (\bar{x} - \mu_0)^2}{\sigma_0^2} \right\} - 2 = 0 \quad (28)$$

A Equação 25 pode ser derivada nos seguintes limites:

$$\frac{n(\bar{x} - \mu_0)^2}{\sigma_0^2} - 2 = 0 \quad (29)$$

$$n \log \left\{ \frac{s^2 + (\bar{x} - \mu_0)^2}{\sigma_0^2} \right\} + n - \frac{ns^2 + n(\bar{x} - \mu_0)^2}{\sigma_0^2} + 2 = 0 \quad (30)$$

$$n \log \left( \frac{s^2}{\sigma_0^2} \right) + n - \frac{ns^2 + n(\bar{x} - \mu_0)^2}{\sigma_0^2} + 4 = 0 \quad (31)$$

Realizando a projeção das equações 17 e equações 26 a 31 para um conjunto de dados de tamanho 5 e que siga uma distribuição normal  $N(7,1)$  obtém-se a Figura 4 que exemplifica o resultado esperado para o gráfico de controle (X-barra,S).

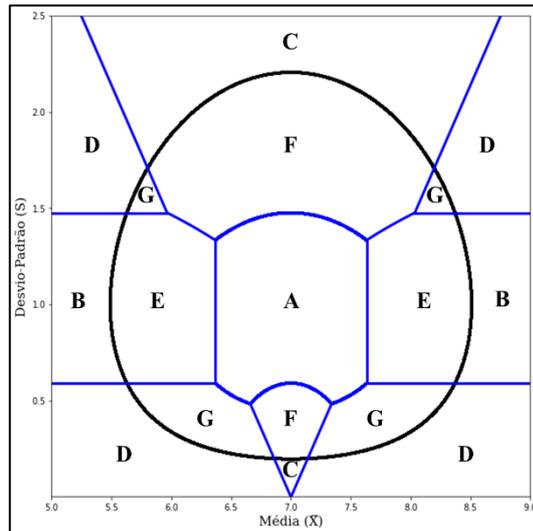


Figura 4 - Representação do gráfico de controle (X-barra,S)

Fonte: Adaptado de Takemoto e Arizono (2018)

Como resultado, tem-se a Figura 4, onde cada área implica nas seguintes condições:

- (A) Condição estável e sob controle;
- (B) Condição fora de controle com mudança na média do processo e sem alteração na variação do processo (alerta);
- (C) Condição fora de controle sem mudança na média do processo e com alteração na variação do processo (alerta);
- (D) Condição fora de controle com mudança em ambas as média e variação (alerta);
- (E) Condição sob controle com possibilidade de alteração na média do processo e sem alteração na variação do processo (atenção);
- (F) Condição sob controle com possibilidade de sem alteração da média do processo e alteração na variação do processo (atenção);
- (G) Condição sob controle com possibilidade de alteração em ambas as média e variação (atenção).

Considerando os trabalhos de Shewhart (1926) e Takemoto e Arizono (2018), é necessário realizar uma comparação quanto a performance dos gráficos a fim de obter parâmetros que auxiliarão na escolha de qual aplicar quando implementar um monitoramento de uma determinada característica de qualidade. Apesar de não estar no escopo deste trabalho,

é necessário apresentar os conceitos do teste de hipóteses aplicados aos gráficos de controle, uma vez que o cálculo dos limites de controle está relacionado ao erro do tipo I ( $\alpha$ ), como será descrito abaixo.

### 2.3 Teste de hipóteses

O teste de hipóteses é uma afirmação sobre os valores dos parâmetros de uma distribuição de probabilidade (MONTGOMERY, 2013). De acordo com Costa, Epprecht e Carpinetti (2012), na aplicação do teste de hipóteses aos gráficos de controle, existem duas possibilidades mutuamente excludentes de hipóteses: a nula ( $H_0$ ), que consiste em considerar que o processo está sob controle; e a alternativa ( $H_1$ ), que considera o processo fora de controle.

Considerando o gráfico de controle de X-barra, o processo é considerado sob controle se a média  $\mu$  de suas saídas é igual a  $\mu_0$ . As equações 32 e 33 mostram a formulação do teste de hipótese.

$$H_0: \mu = \mu_0 \quad (32)$$

$$H_1: \mu \neq \mu_0 \quad (33)$$

Baseando-se na formulação acima, caso a média  $\mu$  for igual à  $\mu_0$ , considera-se que o processo está sob controle e a hipótese  $H_0$  é aceita. Por outro lado, caso  $\mu$  for diferente de  $\mu_0$ , então a hipótese  $H_0$  é rejeitada e, portanto, o processo está fora de controle.

Visto a formulação do teste de hipóteses das equações 32 e 33, é necessário apontar a existência de um risco  $\alpha$  de se recusar a hipótese nula quando ela for verdadeira, denominado erro do tipo I. Por outro lado, existe o risco  $\beta$  de não rejeitar a hipótese nula, mesmo que o processo esteja fora de controle, sendo este o erro do tipo II. Os riscos  $\alpha$  e  $\beta$  representam a probabilidade de obter conclusões incorretas acerca do estado do processo.

Neste trabalho, os limites de controle foram definidos baseando-se apenas na probabilidade do erro do tipo I ( $\alpha$ ), visto que é um dos parâmetros base dos gráficos de Shewhart e as etapas de implementação do gráfico simultâneo (X-barra,S) também constituem no uso  $\alpha$  e não a probabilidade do erro do tipo II ( $\beta$ ).

### 2.4 Fases de aplicação dos gráficos de controle

Os gráficos de controle se baseiam nas estatísticas das saídas de um determinado processo para determinar seu estado. Muitas vezes, não são conhecidos o valor médio e os limites de controle e, portanto, é necessário estimá-los para avaliar se as saídas estão em

conformidade com o projeto. Entretanto, também é possível que o processo não esteja sob controle e, portanto, estimar os parâmetros baseados nesses valores resultaria em gráficos ineficazes.

Dessa forma, Montgomery (2013) define duas fases de aplicação dos gráficos de controle: fase I e fase II. Na fase I, o objetivo é estabelecer um padrão de controle para o processo e, para isso, são coletados dados do processo ao longo do tempo e projetando-os no gráfico de controle, avaliando se as variações observadas estão dentro dos limites aceitáveis para o processo. Durante a fase I, é comum realizar diversos ajustes nos limites de controle e avaliações de presença de causas especiais no processo, visto que é uma etapa de definição dos parâmetros a serem utilizados.

Na fase II, o objetivo é realizar o monitoramento continuamente do processo para detecção de quaisquer desvios fora dos limites especificados e que possam comprometer a sua qualidade. Nesta etapa, os parâmetros do gráfico não são alterados pois considera-se que a sua construção representa um processo sob controle estatístico (MONTGOMERY, 2013).

Considerando as fases de implementação dos gráficos de controle, é ideal que as comparações de performance sejam realizadas na fase II, porque os objetos de estudo são estáveis e estão operando no cenário ideal de suas elaborações.

## 2.5 Cálculo do NMA por simulação de Monte Carlo

A simulação de Monte Carlo se tornou bastante popular entre os pesquisadores quantitativos porque é capaz de gerar aproximações de soluções para diversos problemas matemáticos por meio da realização de experimentos de amostragens estatísticas (FAN, FELSÖVÁLYI, SIVO, KEENAN, 2002).

Conforme Dyer (2016), considerando amostras caracterizadas por distribuições normalmente distribuídas, o cálculo do NMA pode ser aproximado pela equação 34, em que NMA é o número médio de amostras até alarme de um processo monitorado na fase I e  $p$  é a probabilidade de alarme

$$NMA = 1/p \quad (33)$$

Dyer (2016) ainda descreve as etapas da simulação de Monte Carlo para cálculo do NMA:

1. Utilize um gerador numérico aleatório para criar uma série de  $m$  amostras de tamanho  $n$ , sendo  $m$  e  $n$  números inteiros positivos, e que os valores gerados sejam caracterizados por uma distribuição normal  $N(\mu, \sigma)$ ;

2. Transforme a série gerada no gráfico de controle a ser estudado;
3. Estabeleça os limites de controle;
4. Determine a posição de cada amostra no gráfico de estudo, somando o total de amostras que representam um processo sob controle;
5. Obtenha a probabilidade  $p$  pela divisão de amostras sob controle pelo total de amostras comparadas;
6. Calcule o NMA.

Com as etapas apresentadas acima, é obtida o NMA para o gráfico de controle de interesse.

### **3 METODOLOGIA**

Este tópico tem como objetivo apresentar os procedimentos metodológicos que norteiam o trabalho. Primeiramente, a classificação da pesquisa e, em seguida, a descrição das etapas, técnicas e ferramentas utilizadas para atingir os objetivos, justificando as respectivas escolhas.

#### **3.1 Classificação do trabalho**

A classificação do método de pesquisa deste trabalho concorda com o proposto por Miguel *et al.* (2018), quanto à natureza, objetivos, abordagem e método.

Quanto à natureza, a pesquisa pode ser classificada como aplicada, porque é caracterizada por seu interesse prático, em que os resultados sejam utilizados com o intuito de obter a solução de problemas que ocorrem na realidade.

Em relação aos objetivos, a pesquisa exploratória é a que mais se adequa ao trabalho, pois busca proporcionar mais conhecimento sobre o problema estudado, visando torná-lo explícito.

O trabalho pode ser considerado quantitativo, porque utiliza-se da formulação de hipóteses, coleta de dados por meio de observações, análise e geração de resultados.

Por fim, o método de pesquisa é a de modelagem e simulação porque consiste no processamento de parâmetros em um modelo representativo da realidade, sem afetar o ambiente real.

### 3.2 Instrumentos de pesquisa e tratamento dos dados

Todos os procedimentos de modelagem e simulação deste trabalho foram realizados utilizando a linguagem de programação Python, pois é uma linguagem de alto nível (projetada para ser fácil de usar e ser compreendida por usuários de diferentes níveis de conhecimento) e, por possuir uma comunidade extremamente ativa e solidária, existem diversos módulos e ferramentas prontas que facilitam a obtenção dos resultados esperados.

O ambiente de desenvolvimento utilizado para programação e execução do código em Python foi o Spyder, visto que é *open source*, gratuita e foi desenvolvida por pesquisadores, engenheiros e analistas de dados para uso em Python. Desta forma, é comum o seu uso para fins de pesquisa e análises.

Os dados utilizados neste trabalho foram simulados por meio da biblioteca do Python denominada Numpy, uma vez que também é *open source*, gratuita e desenvolvida pela comunidade, além de ser uma das mais aplicadas em trabalhos. O Numpy é capaz de gerar dados aleatórios que sejam caracterizados por qualquer distribuição estatística.

De acordo com a teoria proposta por Shewhart (1926), os dados a serem utilizados na construção dos gráficos de controle necessitam ser independentes, mas não necessariamente precisam ser caracterizados por uma distribuição normal. Entretanto, o trabalho de Takemoto e Arizono (2018) se baseou em dados normalmente distribuídos e, para fins de comparação com os gráficos de Shewhart, os dados aplicados para ambos os gráficos neste trabalho foram simulados com a função do Numpy de forma que fossem caracterizados por uma distribuição normal.

Com o objetivo de realizar comparações de performance, o valor de  $\alpha$  utilizado foi de 0,27% porque é o padrão utilizado desde Shewhart (1926) que atribuiu a regra de deslocamento aceitável da média em três desvios-padrão, com o objetivo de reduzir ajustes desnecessários no processo. Os tamanhos amostrais utilizados foram 4, 6, 8 e 10, e o deslocamento da média e desvio-padrão variou de 0 a 1,5 em intervalos de 0,5. Com a aplicação da combinação desses parâmetros, é possível identificar a performance e definir situações de melhor aplicação de cada gráfico. Como valores da média e desvio-padrão, foram atribuídos 7 e 1, respectivamente, de forma a evitar a obtenção de amostras com valor negativo porque poderia dificultar o desenvolvimento do código em Python para construção dos gráficos e cálculo do NMA.

A Figura 5 é um fluxograma que apresenta a execução da simulação deste trabalho.

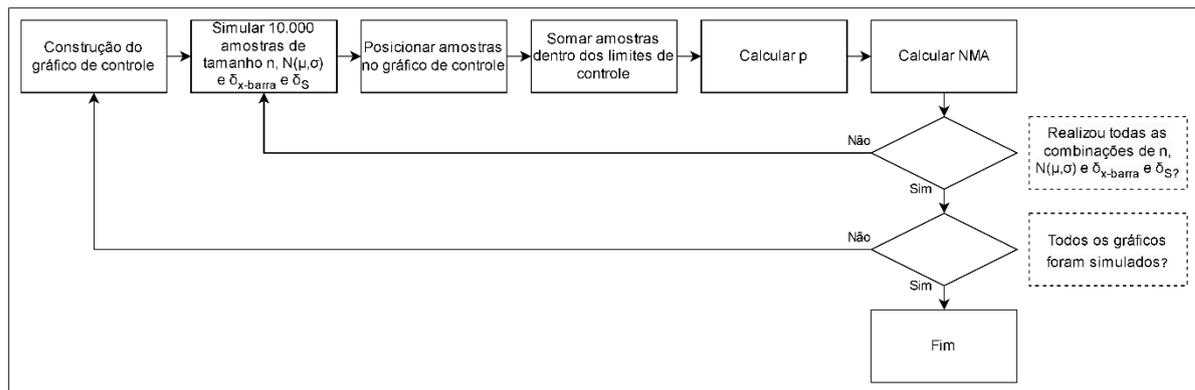


Figura 5 - Fluxograma de execução da simulação deste trabalho

Fonte: Autor do trabalho

A obtenção e tratamento dos dados simulados permitiu dar seguimento ao trabalho para etapa de construção dos gráficos de controle, baseando-se nas equações e procedimentos descritos no tópico de referencial teórico.

Por fim, os limites de controle foram utilizados para determinar a posição das amostras nos gráficos e, assim, calcular o NMA para cada combinação de parâmetros.

## 4 APRESENTAÇÃO E ANÁLISE DE RESULTADOS

Nesta seção estão descritos todos os procedimentos realizados para construção dos gráficos de controle de Shewhart (1926) e Takemoto e Arizono (2018), assim como o cálculo dos respectivos NMA por simulação de Monte Carlo e todos os resultados obtidos com o trabalho.

### 4.1 Construção dos gráficos de controle de Shewhart

Considerando que as amostras sejam caracterizadas por uma distribuição normalmente distribuída  $N(7, 1)$  e foram aplicadas aos gráficos de controle de média ( $\bar{X}$ -barra) e desvio-padrão ( $S$ ) amostral, iniciou-se o processo de construção dos gráficos de Shewhart com o cálculo das médias e dos limites de controle.

Substituindo os valores da média e desvio-padrão nas equações 2 a 7 e utilizando  $c_4$  da Tabela 1 de acordo com o tamanho da amostra  $n$ , foi possível obter os limites de controle e as linhas médias dos gráficos de  $\bar{X}$ -barra e  $S$ , conforme Tabela 2.

Tabela 1 - Valores prefixados para  $c_4$ 

	n = 4	n = 6	n = 8	n = 10
$c_4$	0,921	0,952	0,965	0,973

Fonte: Adaptado de Costa, Epprecht e Carpinetti (2012)

Tabela 2 - Valores médios e limites de controle dos gráficos de Shewhart

n	$LIC_{X\text{-barra}}$	$LM_{X\text{-barra}}$	$LSC_{X\text{-barra}}$	$LIC_S$	$LM_S$	$LSC_S$
4	5,500	7,000	8,500	0,000	1,000	2,089
6	5,775	7,000	8,224	0,034	1,000	1,870
8	5,939	7,000	8,061	0,178	1,000	1,752
10	6,051	7,000	7,949	0,281	1,000	1,667

Fonte: Autor do trabalho

A realização dos cálculos foi feita por meio da execução do código em Python do Apêndice A no Spyder, gerando os valores dos parâmetros apresentados na Tabela 2 e os gráficos da Figura 5 com amostras simuladas.

Baseando-se nos gráficos de controle obtidos na Figura 5, é necessário ressaltar a facilidade em construir os gráficos, pois o cálculo dos valores médios e dos limites de controle são constantes para cada valor de  $n$ .

Outro fator de destaque é que a determinação do estado do processo é bastante visual, pois a avaliação consiste em identificar as amostras que estão além dos limites de controle.

Além disso, supondo que alguma amostra esteja localizada fora dos limites, é rápida a detecção do momento em que houve o desvio, porque os gráficos apresentam a ordem da amostra, garantindo a informação temporal dos estados.

Por outro lado, os gráficos de Shewhart exigem a avaliação dos gráficos de  $X\text{-barra}$  e  $S$  de forma conjunta, porque uma apresenta a centralidade das amostras e a segunda a variabilidade. Visto que é possível variar um dos parâmetros sem necessariamente alterar o outro, o monitoramento do processo com a análise dos dois gráficos em conjunto traz informações mais precisas acerca do sistema.

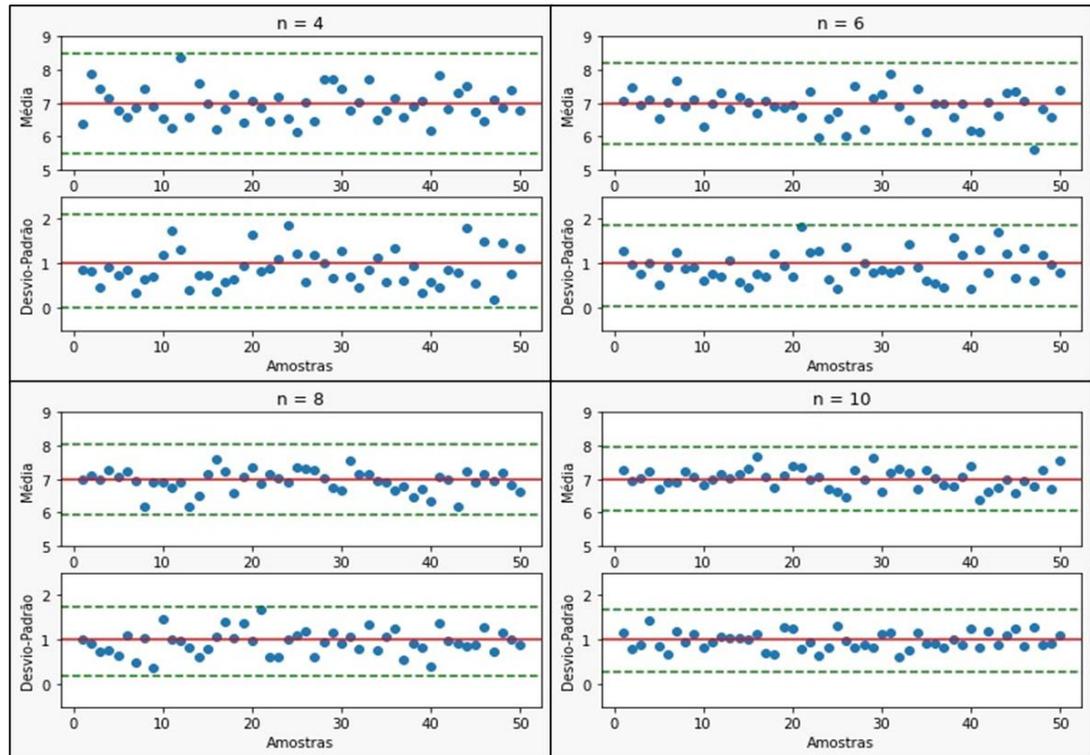


Figura 6 - Gráficos de controle X-barra e S para cada valor de  $n$

Fonte: Autor do trabalho

Baseando-se nos gráficos de controle obtidos na Figura 5, é necessário ressaltar a facilidade em construir os gráficos, pois o cálculo dos valores médios e dos limites de controle são constantes para cada valor de  $n$ .

Outro fator de destaque é que a determinação do estado do processo é bastante visual, pois a avaliação consiste em identificar as amostras que estão além dos limites de controle.

Além disso, supondo que alguma amostra esteja localizada fora dos limites, é rápida a detecção do momento em que houve o desvio, porque os gráficos apresentam a ordem da amostra, garantindo a informação temporal dos estados.

Por outro lado, os gráficos de Shewhart exigem a avaliação dos gráficos de X-barra e S de forma conjunta, porque uma apresenta a centralidade das amostras e a segunda a variabilidade. Visto que é possível variar um dos parâmetros sem necessariamente alterar o outro, o monitoramento do processo com a análise dos dois gráficos em conjunto traz informações mais precisas acerca do sistema.

Com a obtenção dos limites de controle dos gráficos de Shewhart, segue-se para construção do gráfico de controle simultâneo (X-barra,S) que será objeto de comparação.

## 4.2 Construção do gráfico de controle (X-barra,S)

A construção do gráfico de controle simultâneo (X-barra,S) foi realizada baseando-se na distribuição normal  $N(7, 1)$  e  $\alpha = 0,0027$ , iniciando-se pela determinação do limite de controle por meio da Equação 17. Todas os procedimentos descritos nessa sessão foram realizados por meio da execução do código em Python do Apêndice B no Spyder.

Aplica-se a média  $\mu_0$  e variância  $\sigma_0^2$  na Equação 17 e substitui-se o valor de  $\lambda(\alpha)$  conforme  $n$  e a probabilidade do erro Tipo I  $\alpha$  presente na Tabela 3. Dessa forma, é possível resolver a equação em função de  $\mu$  ou  $\sigma$  para obter os infinitos pontos  $(\mu, \sigma)$  que caracterizam o limite de controle do gráfico.

Tabela 3 - Valores de  $\lambda(\alpha)$  para  $n$  e  $\alpha=0,0027$

	$n = 4$	$n = 6$	$n = 8$	$n = 10$
$\alpha = 0,0027$	7,81	7,03	6,71	6,53

Fonte: Adaptado de Kanagawa, Arizono e Ohta (1997)

O mesmo procedimento deve ser realizado para determinar as condições que caracterizam alterações na média e/ou desvio-padrão do sistema, substituindo os mesmos valores de média  $\mu_0$  e variância  $\sigma_0^2$  nas Equações 26 a 31. Entretanto, a Tabela 2 não se faz necessária.

Após a obtenção das respectivas funções derivadas das 26 a 31, ressalta-se a necessidade de realizar ajustes na projeção dos limites, visto que o conjunto de pontos resultará na Figura 6. Como solução, deve-se restringir as projeções das funções obtidas nos pontos de intersecção entre elas, tornando o visual mais limpo e capaz de identificar as subdivisões do gráfico, similar à Figura 4.

Com todos os limites definidos, a projeção das amostras necessita estar como um conjunto de pontos do tipo  $(\mu, \sigma)$  e, por isso, a média e desvio-padrão devem ser calculados para cada subgrupo racional. Assim, tem-se os gráficos de controle simultâneo de (X-barra,S) da Figura 7 com dados simulados.

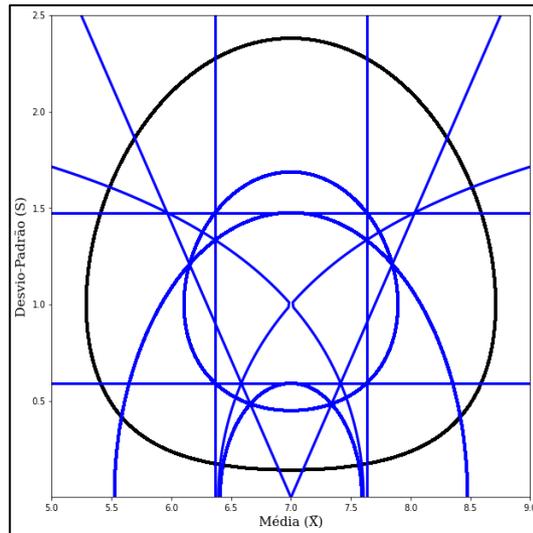


Figura 7 - Projeção dos limites das Equações 26 a 31

Fonte: Autor do trabalho

Após a construção do gráfico simultâneo ( $\bar{X}$ -barra,S), é necessário ressaltar o aumento da complexidade dos procedimentos necessários em comparação com os gráficos de Shewhart. Enquanto os limites de controle dos gráficos de  $\bar{X}$ -barra e S são constantes lineares, os limites do gráfico simultâneo resultam em um formato oval obtidos pela solução da equação 17. Entretanto, a área interna do limite de controle se torna menor à medida que o tamanho amostral aumenta, ou seja, tornando a detecção de desvios menores mais rápida.

O aumento da complexidade também resulta num maior tempo de processamento de máquina para projeção do gráfico. Para gerar os gráficos de Shewhart para cada  $n$  (ou seja, 8 gráficos de controle) demorou em média 0,75 segundo, enquanto os gráficos simultâneos 15 segundos (4 gráficos simultâneos).

Tendo em vista a projeção de 50 amostras em cada gráfico, a análise do gráfico simultâneo demonstra que a identificação dos pontos que se encontram fora de controle não é trivial, sendo visualmente impossível considerando apenas a Figura 7. Essa dificuldade em identificar a amostra individual está relacionada a perda da informação temporal no esquema de projeção do plano  $\bar{X}$ -barra-S. Takemoto e Arizono (2018) tentaram resolver com a aplicação de uma informação visual que difere o tamanho e cor das amostras ao longo do tempo, mas conforme a quantidade de amostras projetadas no gráfico aumenta, mais amostras se sobrepõem e menor é a diferenciação do tamanho do ponto.

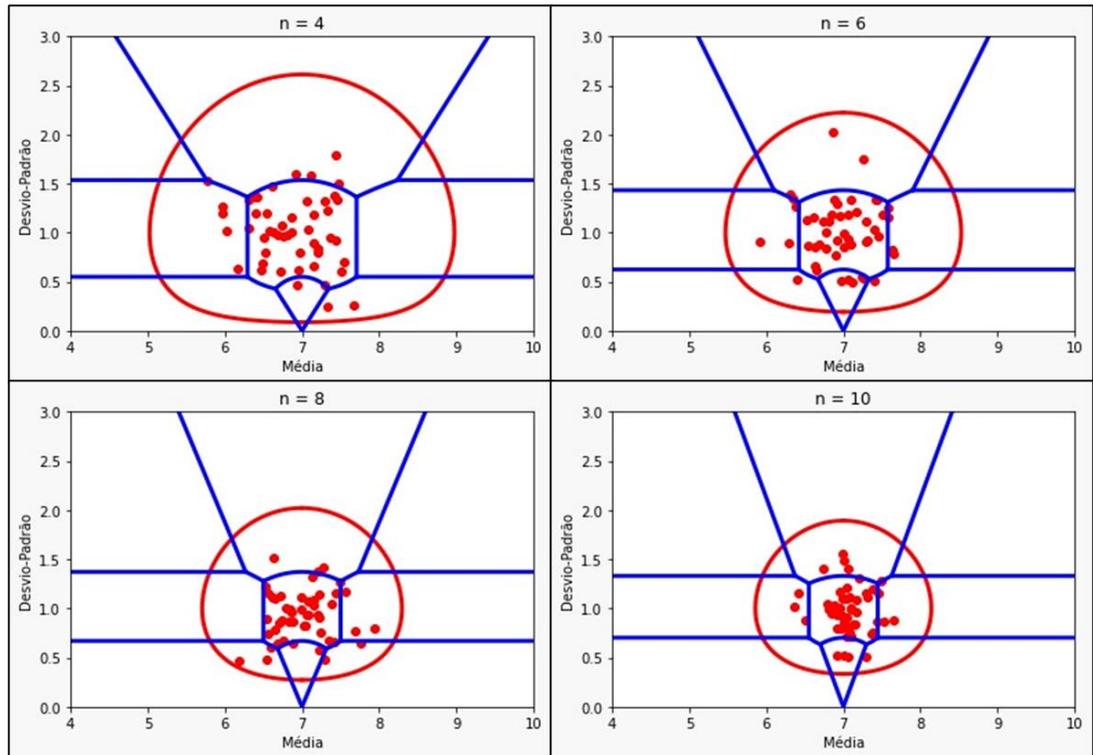


Figura 8 - Gráficos de controle simultâneo (X-barra,S) para cada valor de n

Fonte: Autor do trabalho

Por outro lado, a percepção do estado do processo no momento da projeção da amostra é facilitada no plano X-barra-S, visto que a semelhança com uma folha de tiro ao alvo torna intuitiva a compreensão de que quanto mais centralizada as amostras, mais próxima às especificações de projeto o processo está. Além disso, caso a amostras se encontre em qualquer outra região além da A, é fácil identificar se há alteração na média e/ou desvio-padrão porque as duas informações estão em uma única representação, agilizando o processo de identificação de desvios.

### 4.3 Cálculo do NMA

Após a obtenção dos limites de controle de cada gráfico, foi necessário aplicar a simulação de Monte Carlo descrita por Dyer (2016). Para isso, foram utilizadas as combinações de parâmetros descritas na seção de metodologia deste trabalho, ou seja, os tamanhos amostrais n utilizados foram 4, 6, 8 e 10, o deslocamento da média e desvio-padrão variou de 0 a 1,5 em intervalos de 0,5, e as amostras seguiram uma distribuição normal com média 7 e desvio-padrão

1, valores definidos de forma a evitar valores negativos. Além disso, cada simulação contou com 10.000 amostras.

Os códigos utilizados para simulação são os mesmos da construção dos gráficos de controle e, portanto, estão disponíveis nos Apêndices A e B, sendo o primeiro para os gráficos de Shewhart e o segundo para o gráfico simultâneo (X-barra,S). Os resultados obtidos da simulação foram compilados na Tabela 4.

Tabela 4 - Valores de NMA dos gráficos estudados de acordo com o n e desvios da média e desvio-padrão

dx	ds	n=4		n=6		n=8		n=10	
		$\bar{X}$ e S	( $\bar{X}$ , S)	$\bar{X}$ e S	( $\bar{X}$ , S)	$\bar{X}$ e S	( $\bar{X}$ , S)	$\bar{X}$ e S	( $\bar{X}$ , S)
0,00	0,50	6,236	19,231	4,822	11,494	3,982	6,993	3,302	5,814
	1,00	2,285	3,745	1,771	2,247	1,501	1,887	1,328	1,517
	1,50	1,525	2,020	1,261	1,497	1,139	1,209	1,073	1,114
0,50	0,00	37,157	166,667	24,330	83,333	16,907	27,778	12,295	33,333
	0,50	4,891	13,889	3,710	7,353	3,024	4,717	2,523	3,460
	1,00	2,134	3,311	1,665	2,273	1,425	1,608	1,271	1,435
	1,50	1,489	1,852	1,239	1,395	1,124	1,199	1,064	1,111
1,00	0,00	6,176	18,868	3,404	7,463	2,307	3,704	1,767	2,639
	0,50	2,920	5,405	2,147	3,003	1,746	2,028	1,500	1,623
	1,00	1,800	2,475	1,435	1,669	1,258	1,381	1,152	1,222
	1,50	1,394	1,718	1,182	1,292	1,089	1,124	1,043	1,066
1,50	0,00	1,993	3,690	1,333	1,805	1,119	1,299	1,043	1,109
	0,50	1,782	2,494	1,373	1,664	1,190	1,261	1,099	1,156
	1,00	1,478	1,838	1,227	1,337	1,114	1,166	1,058	1,059
	1,50	1,283	1,484	1,117	1,176	1,051	1,071	1,022	1,028

Fonte: Autor do trabalho

Analisando os valores obtidos na Tabela 4, é possível notar que a performance dos gráficos de Shewhart em conjunto são melhores que o gráfico simultâneo (X-barra,S). Em nenhuma combinação de parâmetros o gráfico (X-barra,S) se mostrou superior, apesar de que a proximidade aumenta conforme o tamanho amostral e os deslocamentos da média e desvio-padrão forem maiores, sendo que o NMA para os gráficos de Shewhart e simultâneo (X-barra,S) é muito próximo quando n=10 e os deslocamentos da média e desvio-padrão forem 1,5.

Dessa forma, o gráfico de controle simultâneo possui uma performance pior que os gráficos de Shewhart em conjunto para alertar desvios no processo.

## 5 CONCLUSÃO

Esse trabalho realizou uma comparação entre os gráficos de controle de Shewhart para média (X-barra) e desvio-padrão (S) amostral em conjunto e o gráfico simultâneo (X-barra,S) de Takemoto e Arizono (2018), por meio da avaliação dos procedimentos de construção, performance do NMA e a facilidade em avaliar o estado do processo durante toda a série temporal.

A construção dos gráficos de controle foi realizada pelos códigos em Python dos Apêndices A e B, sendo o primeiro para os gráficos de Shewhart e o segundo o gráfico simultâneo (X-barra,S). Apenas observando a quantidade de linhas de código, é possível afirmar que os gráficos de Shewhart possuem procedimentos mais simples de construção. O principal fator é que os limites de controle são fixos para cada valor de n. Enquanto o limite de controle do gráfico simultâneo (X-barra,S) é obtido com a equação 17, gerando infinitos pontos que delimitam um estado do processo sob controle. Considerando a complexidade da construção dos gráficos, os gráficos de Shewhart necessitam de um tempo de execução muito menor do que o gráfico simultâneo (X-barra,S).

Com os gráficos construídos e na fase II e as amostras projetadas, pode-se realizar a avaliação do estado do processo. O gráfico simultâneo (X-barra,S) se mostra mais intuitivo, visto que é similar a um alvo de estande de tiro, ou seja, quanto mais centralizada for a projeção da amostra, mais próxima às especificações de projeto. Outro fator que o torna mais rápido de se analisar é que as informações da média e desvio-padrão estão em uma única representação e separada por regiões que definem se há alteração em um ou ambos os parâmetros, conforme Figura 4. Por outro lado, o gráfico simultâneo (X-barra,S) não apresenta a informação temporal das amostras e, por isso, haverá dificuldades em determinar em qual momento o processo apresentou algum desvio. A perda da informação temporal não ocorre nos gráficos de Shewhart.

Por fim, foi avaliada a performance dos gráficos com o valor do NMA para cada uma das combinações de n e deslocamentos da média e desvio-padrão. Os gráficos de Shewhart em conjunto apresentaram um NMA menor em todas as combinações utilizadas. Entretanto, a diferença diminui conforme o tamanho amostral n e os deslocamentos da média e desvio-padrão são maiores, portanto, a menor diferença foi encontrada quando  $n = 10$  e os deslocamentos fossem iguais a 1,5. Assim, considerando que o processo de amostragem seja destrutivo, ocasionando a perda do produto, seria necessário aumentar muito os custos do gráfico simultâneo (X-barra,S) com o objetivo de igualar a velocidade de alerta.

O Quadro 1 apresenta um resumo dos resultados obtidos pela comparação entre o gráfico simultâneo (X-barra,S) e os gráficos de Shewhart para média e desvio-padrão.

Quadro 1 – Resumo dos resultados obtidos pelas comparações

	<b>Gráficos de X-barra e S</b>	<b>Gráfico simultâneo (<math>\bar{X}</math>, S)</b>
Construção	<ul style="list-style-type: none"> <li>• Mais simples: médias e limites de controle são constantes.</li> <li>• Necessário construir dois gráficos.</li> <li>• Tempo de execução para construção dos gráficos: 0,75 segundos.</li> </ul>	<ul style="list-style-type: none"> <li>• Mais complexa: limite de controle não é constante e é necessário resolver várias equações.</li> <li>• Necessário construir apenas um gráfico.</li> <li>• Tempo de execução para construção do gráfico: 15 segundos.</li> </ul>
Avaliação do estado o processo	<ul style="list-style-type: none"> <li>• Menos intuitiva: necessário avaliar dois gráficos para determinar o estado do processo.</li> </ul>	<ul style="list-style-type: none"> <li>• Mais intuitiva: analogia com uma folha de tiro ao alvo.</li> <li>• Avaliação de apenas um gráfico, pois apresenta as informações de centralidade e variabilidade.</li> </ul>
Identificação da amostra	<ul style="list-style-type: none"> <li>• Mais rápida: possui a informação temporal das amostras</li> </ul>	<ul style="list-style-type: none"> <li>• Dificuldade em identificar as amostras: perda da informação temporal</li> </ul>
Performance - NMA	<ul style="list-style-type: none"> <li>• Menor NMA: os gráficos combinados alertam deslocamentos da média e desvio-padrão com menos amostras.</li> </ul>	<ul style="list-style-type: none"> <li>• Maior NMA: o gráfico simultâneo necessita de mais amostras para alertar deslocamentos da média e desvio-padrão.</li> </ul>

Fonte: Autor do trabalho

Conforme os resultados apresentados, conclui-se que não é vantajosa a aplicação do gráfico simultâneo (X-barra,S) para avaliar processos concorrentes dos gráficos de Shewhart para média e desvio-padrão. O gráfico simultâneo (X-barra,S) mostrou-se mais lento para alertar desvios, maior complexidade de construção e implementação, além de apresentar a perda da informação temporal da série amostral. As inovações propostas por Takemoto e Arizono (2018) são interessantes para detectar tendências, mas as desvantagens são superiores aos ganhos apresentados.

Por outro lado, o trabalho de Takemoto e Arizono (2018) mostrou-se bastante promissor para monitoramento de processos multivariados, sendo um forte candidato para comparação com o gráfico de Hotelling ( $T^2$ ), visto que a criação de espaços soluções com a implementação do critério informativo de Akaike torna interessante a avaliação de processo que possuam várias variáveis a serem monitoradas.

Visto que este trabalho não realizou o cálculo teórico do poder do gráfico simultâneo ( $\bar{X}$ -barra,S) e que sua aplicação se mostra bastante promissor para processos multivariados, esses serão temas de trabalhos futuros.

## REFERÊNCIAS

- CORRÊA, Henrique. L.; CORRÊA, Carlos. A. **Administração de produção e operações: manufatura e serviços: uma abordagem estratégica**. 2. ed. São Paulo: Atlas, 2011.
- COSTA, Antonio F. B.; EPPRECHT, Eugenio K.; CARPINETTI, Luiz C. R. **Controle estatístico de qualidade**. 2. ed. São Paulo: Atlas. 2012.
- DYER, John N. Monte Carlo simulation design for evaluating normal-based control chart properties. **Journal of Modern Applied Statistical Methods**, v. 15, n. 2, p. 35, 2016.
- FAN, X.; FELÖVÁLYI, Á.; SIVO, S. A.; KEENAN, S. C. **SAS for Monte Carlo Studies: A Guide for Quantitative Researchers**. North Carolina: SAS Institute Inc, 2002.
- GAN, F. F. Joint monitoring of process mean and variance using exponentially weighted moving average control charts. **Technometrics**, v.37, p. 446–453, 1995
- KANAGAWA, A.; ARIZONO, Ikuo; OHTA, H. Design of the control chart based on Kullback-Leibler information. **Frontiers in Statistical Quality Control**, v. 5, p. 183–192, 1997.
- KNAUER, Whitney. A.; GODDEN, Sandra. M.; DIETRICH, Andrea.; HAWKINS, Douglas. M.; JAMES, Robert. E. Evaluation of applying statistical process control techniques to daily average feeding behaviors to detect disease in automatically fed group-housed preweaned dairy calves. **Journal of Dairy Science**, v. 101, n. 9, p. 8135-8145, 2018.
- LOUZADA, Francisco; DINIZ, Carlos A R.; FERREIRA, Paulo H.; FERREIRA, E. L. **Controle Estatístico de Processos - Uma Abordagem Prática para Cursos de Engenharia e Administração**. Rio de Janeiro: LTC, 2013.
- MCCRACKEN, A. K.; CHAKRABORTI, S. Control charts for joint monitoring of mean and variance: An overview. **Quality Technology & Quantitative Management**, v. 10, p. 17–36, 2013.
- MIGUEL, Paulo A. C. et al. **Metodologia de Pesquisa em Engenharia de Produção e Gestão de Operações**. 3. ed. Rio de Janeiro: Elsevier, 2018.
- MONTGOMERY, Douglas. C. **Introduction to statistical process control**. 7. ed. John Willey & Sons, Inc. 2013.
- PRIYA, Metlapall C.; KANTAM, Rani R. Variable control charts-linear failure rate distribution. **Pakistan Journal of Statistics and Operation Research**, p. 715-726, 2017.
- SCOPUS. **Base de dados bibliográficos**. Disponível em: <https://www.scopus.com>. Acesso em: 15 mar. 2022.
- SHEWHART, Walter A. Quality control charts. **The Bell System Technical Journal**, v. 5, n. 4, p. 593-603, 1926.
- SHEWHART, Walter. A. Some applications of statistical methods to the analysis of physical and engineering data. **Bell System Technical Journal**, v. 3, n. 1, p. 43-87, 1924.

SIQUEIRA, Luiz G. P. **Controle estatístico do processo**. São Paulo: Pioneira, 1997.

SIMÕES, Felipe D.; COSTA, Antonio F. B.; MACHADO, Marcela A. G. The trinomial ATTRIVAR control chart. **International Journal of Production Economics**, v. 224, n. 107559, 2020.

TAKEMOTO, Yasuhiko; ARIZONO, Ikuo. Information visualization about changes of process mean and variance on  $(\bar{X}, S)$  control chart. **Quality Technology & Quantitative Management**, v. 16, n. 4, p. 496-510, 2018.

TAKEMOTO, Yasuhiko; ARIZONO, Ikuo; SATOH, Takanori. Discrimination of out-of-control condition using AIC in  $(\bar{X}, s)$  control chart. **Industrial Engineering and Management Systems**, v. 12, n. 2, p. 112-117, 2013.

YEGANEH, Ali; SHADMAN, Alireza; AMIRI, Amirhossein. A novel run rules based MEWMA scheme for monitoring general linear profiles. **Computers & Industrial Engineering**, v. 152, p. 107031, 2021.

## APÊNDICE A – CÓDIGO PARA CONSTRUÇÃO E CÁLCULO DO NMA DOS GRÁFICOS DE SHEWHART

```

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import itertools

#####
# DEFINE OS PARÂMETROS DOS GRÁFICOS #
#####

n = [4, 6, 8, 10]
size = 1_000_000
x = 7
s = 1
dx = [0, 0.5, 1, 1.5]
ds = [0, 0.5, 1, 1.5]
c4 = {4: 0.921, 6: 0.952, 8: 0.965, 10: 0.973,}

#####
# CALCULA OS LIMITES DE CONTROLE #
#####

def limites_de_controle(n, x, s, c4):
    licx = x - 3*s/np.sqrt(n)
    lscx = x + 3*s/np.sqrt(n)
    lics = ((c4[n]*s) - 3*s*np.sqrt(1-c4[n]**2)) if ((c4[n]*s) - 3*s*np.sqrt(1-c4[n]**2)) >=0 else 0
    lscs = (c4[n]*s) + 3*s*np.sqrt(1-c4[n]**2)

    return licx, lscx, lics, lscs

#####
#     CALCULA O NMA     #
#####

def calcula_nma(n, x, s, dx, ds, c4):
    licx, lscx, lics, lscs = limites_de_controle(n, x, s, c4)

    data = [np.random.normal(x+dx, s+ds, size) for i in range(n)]
    columns = ["x"+str(i+1) for i in range(n)]
    df = pd.DataFrame(data=dict(zip(columns, data)))
    df["mean"] = df[columns].mean(axis=1)
    df["std"] = df[columns].std(axis=1)
    df["out"] = np.where(
        (df["mean"] >= licx) & (df["mean"] <= lscx) &
        (df["std"] >= lics) & (df["std"] <= lscs)
        , True, False)

```

```

t = df[df["out"]==False].shape[0]/size

return 1/t

results = pd.DataFrame(data=[], columns=['n', 'dx', 'ds', 'NMA'])
for n_, dx_, ds_ in itertools.product(n, dx, ds):
    if dx_ == 0 and ds_ == 0:
        pass
    else:
        results = results.append(pd.DataFrame([[n_, dx_, ds_, calcula_nma(n_, x, s, dx_, ds_,
c4)]], columns=['n', 'dx', 'ds', 'NMA']), ignore_index=True)

#####
# REALIZA A PROJEÇÃO DOS GRÁFICOS #
#####

for n_ in n:
    licx, lscx, lics, lscs = limites_de_controle(n_, x, s, c4)
    size = 50
    data = [np.random.normal(x, s, size) for i in range(n_)]
    columns = ["x"+str(i+1) for i in range(n_)]
    df = pd.DataFrame(data=dict(zip(columns, data)))
    df["mean"] = df[columns].mean(axis=1)
    df["std"] = df[columns].std(axis=1)
    amostras = list(range(1, len(data[0])+1))

    # GRÁFICO DA MÉDIA
    plt.subplot(2, 1, 1)
    plt.plot(amostras, df["mean"], 'o')
    plt.axhline(x, color='r', linestyle='-')
    plt.axhline(licx, color='g', linestyle='--')
    plt.axhline(lscx, color='g', linestyle='--')
    plt.title(f'n = {n_}')
    plt.ylabel('Média')
    plt.ylim(5,9)

    # GRÁFICO DO DESVIO-PADRÃO
    plt.subplot(2, 1, 2)
    plt.plot(amostras, df["std"], 'o')
    plt.axhline(s, color='r', linestyle='-')
    plt.axhline(lics, color='g', linestyle='--')
    plt.axhline(lscs, color='g', linestyle='--')
    plt.xlabel('Amostras')
    plt.ylabel('Desvio-Padrão')
    plt.ylim(-0.5,2.5)
    plt.show()

```

## APÊNDICE B – CÓDIGO PARA CONSTRUÇÃO E CÁLCULO DO NMA DO GRÁFICO SIMULTÂNEO ( $\bar{X}, S$ )

```

import matplotlib.pyplot as plt
import numpy as np
import pandas as pd
import sympy as sp
import math
import shapely
import itertools

from sympy import *

#####
# DETERMINA AS EQUAÇÕES DAS RETAS #
#####

def eq10(mu,sigma,n,L):

    v_L, v_mu, v_sigma, v_n = L, mu, sigma, n
    L, mu, sigma, x, s, n = symbols("L mu sigma x s n")
    eq = -L+n/2*(log(sigma**2/s**2)-1+s**2/sigma**2+(x-mu)**2/sigma**2)

    answers = [i.subs({L: v_L, mu: v_mu, sigma: v_sigma, n: v_n}) for i in solve(eq, x)]
    x1 = sp.lambdify(s, answers[0].evalf(), 'numpy')
    x2 = sp.lambdify(s, answers[1].evalf(), 'numpy')

    return x1,x2

def eq17(mu, sigma):

    v_mu, v_sigma = mu, sigma
    mu, sigma, x, s = symbols("mu sigma x s")
    eq = log((s**2+(x-mu)**2)/sigma**2)+1-(s**2/sigma**2)

    answers = [i.subs({mu: v_mu, sigma: v_sigma}) for i in solve(eq, x)]
    x1 = sp.lambdify(s, answers[0].evalf(), 'numpy')
    x2 = sp.lambdify(s, answers[1].evalf(), 'numpy')

    return x1,x2

def eq18(sigma, n):

    v_sigma, v_n = sigma, n
    sigma, x, s, n = symbols("sigma x s n")
    eq = n*log(s**2/sigma**2)+n-(n*s**2/sigma**2)+2

    answers = [i for i in solve(eq.subs({sigma: v_sigma, n: v_n}), s)]

```

```
s1 = sp.lambdify(s, answers[0].evalf(), 'numpy')
s2 = sp.lambdify(s, answers[1].evalf(), 'numpy')
```

```
return s1,s2
```

```
def eq19(mu, n):
```

```
v_mu, v_n = mu, n
mu, x, s, n = symbols("mu x s n")
eq = n*log((s**2+(x-mu)**2)/s**2)-2
```

```
answers = [i.subs({mu: v_mu, n: v_n}) for i in solve(eq, x)]
x1 = sp.lambdify(s, answers[0].evalf(), 'numpy')
x2 = sp.lambdify(s, answers[1].evalf(), 'numpy')
```

```
return x1,x2
```

```
def eq22(mu, sigma, n):
```

```
v_mu, v_sigma, v_n = mu, sigma, n
mu, sigma, x, n = symbols("mu sigma x n")
eq = (n*(x-mu)**2)/sigma**2-2
```

```
answers = [i.subs({mu: v_mu, sigma: v_sigma, n: v_n}) for i in solve(eq, x)]
x1 = sp.lambdify(x, answers[0].evalf(), 'numpy')
x2 = sp.lambdify(x, answers[1].evalf(), 'numpy')
```

```
return x1,x2
```

```
def eq23(mu, sigma, n):
```

```
v_mu, v_sigma, v_n = mu, sigma, n
mu, sigma, x, s, n, y = symbols("mu sigma x s n y")
```

```
eq = 2+n+n*log((s**2+(x-mu)**2)/sigma**2)-(n*s**2+n*(x-mu)**2)/sigma**2
neq = eq.subs({mu: v_mu, sigma: v_sigma, n: v_n})
answers = [si*sqrt(1-x**2) for i in solve(neq.subs(s**2, y - x**2), y) for si in (-1,1)]
s1 = sp.lambdify(x, answers[1].evalf(), 'numpy')
s2 = sp.lambdify(x, answers[3].evalf(), 'numpy')
```

```
return s1,s2
```

```
def eq24(mu, sigma, n):
```

```
v_mu, v_sigma, v_n = mu, sigma, n
mu, sigma, x, s, n = symbols("mu sigma x s n")
```

```
eq = n*log(s**2/sigma**2)+n-(n*s**2+n*(x-mu)**2)/sigma**2+4
answers = [i.subs({mu: v_mu, sigma: v_sigma, n: v_n}) for i in solve(eq, x)]
```

```

x1 = sp.lambdify(s, answers[0].evalf(), 'numpy')
x2 = sp.lambdify(s, answers[1].evalf(), 'numpy')

return x1,x2

#####
# DETERMINA OS PARÂMETROS DAS FUNÇÕES #
#####

mu = x_ = 7
sigma = s_ = 1
size = 1_000
n_ = [4, 6, 8, 10]
L_ = [7.81, 7.03, 6.71, 6.53]
dx = [0, 0.5, 1, 1.5]
ds = [0, 0.5, 1, 1.5]

x = np.linspace(4, 10, 1_000_000)
s = np.linspace(0, 3, 1_000_000)

results = pd.DataFrame(data=[], columns=['n', 'dx', 'ds', 'NMA'])

#####
# DETERMINA OS LIMITES DAS FUNÇÕES #
#####

for n, L in zip(n_, L_):

    # Limites
    x10_left, x10_right = eq10(mu, sigma, n, L)
    x17_left, x17_right = eq17(mu, sigma)
    s18_down, s18_up = eq18(sigma, n)
    x19_left, x19_right = eq19(mu, n)
    x22_left, x22_right = eq22(mu, sigma, n)
    s23_down, s23_up = eq23(mu, sigma, n)
    x24_left, x24_right = eq24(mu, sigma, n)

    # Eq17
    l_x17_left = x17_left(s18_up(1))
    r_x17_left = x22_left(1)
    l_x17_right = x22_right(1)
    r_x17_right = x17_right(s18_up(1))

    left_s17 = s[np.where((x17_left(s) > l_x17_left) & (x17_left(s) < r_x17_left) & (s >
s[np.where((np.absolute(x22_left(s)-x17_left(s))<1e-5) == True)].max()))]
    left_x17 = x17_left(left_s17)
    right_s17 = s[np.where((x17_right(s) > l_x17_right) & (x17_right(s) < r_x17_right) & (s >
s[np.where((np.absolute(x22_right(s)-x17_right(s))<1e-5) == True)].max()))]
    right_x17 = x17_right(right_s17)

```

```

# Eq18
r_x18_up = l_x17_left
l_x18_up = r_x17_right
r_x18_down = x24_left(s18_down(1))
l_x18_down = x24_right(s18_down(1))

l_up_x18 = x[np.where((x < r_x18_up) == True)]
l_up_s18 = np.array([s18_up(i) for i in l_up_x18])
r_up_x18 = x[np.where((x > l_x18_up) == True)]
r_up_s18 = np.array([s18_up(i) for i in r_up_x18])
l_down_x18 = x[np.where((x < r_x18_down) == True)]
l_down_s18 = np.array([s18_down(i) for i in l_down_x18])
r_down_x18 = x[np.where((x > l_x18_down) == True)]
r_down_s18 = np.array([s18_down(i) for i in r_down_x18])

# Eq19
r_s19_aux = np.where((np.absolute(x19_left(s)-x24_left(s))<1e-5) == True)
l_s19_aux = np.where((np.absolute(x19_right(s)-x24_right(s))<1e-5) == True)

r_x19_left = x19_left(s18_up(1))
l_x19_left = x19_left(s[r_s19_aux].min())
r_x19_right = x19_right(s[l_s19_aux].min())
l_x19_right = x19_right(s18_up(1))

l_left_s19 = s[np.where(x19_left(s) < r_x19_left)]
l_left_x19 = x19_left(l_left_s19)
r_left_s19 = s[np.where(x19_left(s) > l_x19_left)]
r_left_x19 = x19_left(r_left_s19)
l_right_s19 = s[np.where(x19_right(s) < r_x19_right)]
l_right_x19 = x19_right(l_right_s19)
r_right_s19 = s[np.where(x19_right(s) > l_x19_right)]
r_right_x19 = x19_right(r_right_s19)

# Eq22
d_s22_left_aux = np.where((np.absolute(x22_left(s)-x24_left(s))<1e-5) == True)
u_s22_left_aux = np.where((np.absolute(x22_left(s)-x17_left(s))<1e-5) == True)
d_s22_right_aux = np.where((np.absolute(x22_right(s)-x24_right(s))<1e-5) == True)
u_s22_right_aux = np.where((np.absolute(x22_right(s)-x17_right(s))<1e-5) == True)

u_s22_left = s[u_s22_left_aux].max()
d_s22_left = s[d_s22_left_aux].min()
u_s22_right = s[u_s22_right_aux].max()
d_s22_right = s[d_s22_right_aux].min()

up_s22 = s[np.where((s < u_s22_left) & (s > d_s22_left))]
up_x22 = np.array([x22_left(i) for i in up_s22])
down_s22 = s[np.where((s < u_s22_right) & (s > d_s22_right))]
down_x22 = np.array([x22_right(i) for i in up_s22])

# Eq23

```

```

r_x23_up = x22_right(1)
l_x23_up = x22_left(1)
r_x23_down = l_x19_left
l_x23_down = r_x19_right

left_x23 = x[np.where((x > l_x23_up) & (x < r_x23_up))]
left_s23 = s23_up(left_x23)
right_x23 = x[np.where((x < l_x23_down) & (x > r_x23_down))]
right_s23 = s23_down(right_x23)

# Eq24
l_x24_left = x24_left(s18_down(1))
r_x24_left = l_x19_left
l_x24_right = r_x19_right
r_x24_right = x24_right(s18_down(1))

l_left_s24 = s[np.where((x24_left(s) > l_x24_left) & (x24_left(s) < r_x24_left) & (s <=
sigma))]
l_left_x24 = x24_left(l_left_s24)
l_right_s24 = s[np.where((x24_right(s) > l_x24_right) & (x24_right(s) < r_x24_right) & (s
<= sigma))]
l_right_x24 = x24_right(l_right_s24)

#####
# REALIZA A PROJEÇÃO DO GRÁFICO #
#####

size = 50
data = [np.random.normal(x_, s_, size) for i in range(n)]
columns = ["x"+str(i+1) for i in range(n)]
df = pd.DataFrame(data=dict(zip(columns, data)))
df["mean"] = df[columns].mean(axis=1)
df["std"] = df[columns].std(axis=1)

plt.plot(df["mean"],df["std"],'ro',linewidth=3)

plt.plot(x10_left(s), s, '-r', linewidth=3)
plt.plot(x10_right(s), s, '-r', linewidth=3)

plt.plot(left_x17, left_s17, '-b', linewidth=3)
plt.plot(right_x17, right_s17, '-b', linewidth=3)

plt.plot(l_up_x18, l_up_s18, '-b', linewidth=3)
plt.plot(r_up_x18, r_up_s18, '-b', linewidth=3)
plt.plot(l_down_x18, l_down_s18, '-b', linewidth=3)
plt.plot(r_down_x18, r_down_s18, '-b', linewidth=3)

plt.plot(l_left_x19, l_left_s19, '-b', linewidth=3)
plt.plot(r_left_x19, r_left_s19, '-b', linewidth=3)
plt.plot(l_right_x19, l_right_s19, '-b', linewidth=3)

```

```

plt.plot(r_right_x19, r_right_s19, '-b', linewidth=3)

plt.plot(up_x22, up_s22, '-b', linewidth=3)
plt.plot(down_x22, down_s22, '-b', linewidth=3)

plt.plot(left_x23, left_s23, '-b', linewidth=3)
plt.plot(right_x23, right_s23, '-b', linewidth=3)

plt.plot(l_left_x24, l_left_s24, '-b', linewidth=3)
plt.plot(l_right_x24, l_right_s24, '-b', linewidth=3)
plt.xlim(min(x),max(x))
plt.ylim(min(s),max(s))
plt.xlabel('Média')
plt.ylabel('Desvio-Padrão')
plt.title(f'n = {n}')
plt.show()

#####
#  CRIA OS ESPAÇOS SOLUÇÕES  #
#####

def sort_points(points):
    centre_x, centre_y = sum([x for x,_ in points])/len(points), sum([y for _,y in
points])/len(points)
    angles = [math.atan2(y - centre_y, x - centre_x) for x,y in points]
    counterclockwise_indices = sorted(range(len(points)), key=lambda i: angles[i])
    counterclockwise_points = [points[i] for i in counterclockwise_indices]

    return counterclockwise_points

limite_controle = zip(np.hstack((x10_left(s),x10_right(s))), np.hstack((s,s)))
limite_controle = [(x,y) for x,y in limite_controle if ~np.isnan(x)]
limite_controle = shapely.Polygon(sort_points(limite_controle))

up_left = zip(np.hstack((l_left_x19,l_up_x18)), np.hstack((l_left_s19,l_up_s18)))
up_left = [(x,y) for x,y in up_left if ~np.isnan(x)]
up_left = shapely.Polygon(sort_points(up_left))

up_right = zip(np.hstack((r_right_x19,r_up_x18)), np.hstack((r_right_s19,r_up_s18)))
up_right = [(x,y) for x,y in up_right if ~np.isnan(x)]
up_right = shapely.Polygon(sort_points(up_right))

up_center = zip(
    np.hstack((l_left_x19, r_right_x19, left_x17, right_x17, left_x23,)),
    np.hstack((l_left_s19, r_right_s19, left_s17, right_s17, left_s23,))
)
up_center = [(x,y) for x,y in up_center if ~np.isnan(x)]
up_center = shapely.Polygon(sort_points(up_center))

center_left = zip(

```

```

    np.hstack((l_up_x18, l_down_x18, left_x17, up_x22)),
    np.hstack((l_up_s18, l_down_s18, left_s17, up_s22))
)
center_left = [(x,y) for x,y in center_left if ~np.isnan(x)]
center_left = shapely.Polygon(sort_points(center_left))

center_right = zip(
    np.hstack((r_up_x18, r_down_x18, right_x17, down_x22)),
    np.hstack((r_up_s18, r_down_s18, right_s17, down_s22))
)
center_right = [(x,y) for x,y in center_right if ~np.isnan(x)]
center_right = shapely.Polygon(sort_points(center_right))

center_center = zip(
    np.hstack((down_x22, up_x22, left_x23, right_x23, l_left_x24, l_right_x24)),
    np.hstack((down_s22, up_s22, left_s23, right_s23, l_left_s24, l_right_s24))
)
center_center = [(x,y) for x,y in center_center if ~np.isnan(x)]
center_center = shapely.Polygon(sort_points(center_center))

down_left = zip(
    np.hstack((l_down_x18, r_left_x19, l_left_x24)),
    np.hstack((l_down_s18, r_left_s19, l_left_s24))
)
down_left = [(x,y) for x,y in down_left if ~np.isnan(x)]
down_left = shapely.Polygon(sort_points(down_left))

down_right = zip(
    np.hstack((r_down_x18, l_right_x19, l_right_x24)),
    np.hstack((r_down_s18, l_right_s19, l_right_s24))
)
down_right = [(x,y) for x,y in down_right if ~np.isnan(x)]
down_right = shapely.Polygon(sort_points(down_right))

down_center = zip(
    np.hstack((l_right_x19, right_x23)),
    np.hstack((l_right_s19, right_s23))
)
down_center = [(x,y) for x,y in down_center if ~np.isnan(x)]
down_center = shapely.Polygon(sort_points(down_center))

#####
#      CALCULA O NMA      #
#####

def calcula_nma(n, x, s, dx, ds):
    data = [np.random.normal(x+dx, s+ds, size) for i in range(n)]
    columns = ["x"+str(i+1) for i in range(n)]
    df = pd.DataFrame(data=dict(zip(columns, data)))
    df["mean"] = df[columns].mean(axis=1)

```

```

df["std"] = df[columns].std(axis=1)
points = np.array([shapely.Point(ij) for ij in zip(df["mean"],df["std"])])
# points
lim = limite_controle.contains(points)
lim_q = len([x for x in filter(lambda i: i==False, lim)])/size
cc = center_center.contains(points)
cc_q = len([x for x in filter(lambda i: i==False, cc)])/size
df['lim'] = lim
df['cc'] = cc
df['ul'] = up_left.contains(points)
df['uc'] = up_center.contains(points)
df['ur'] = up_right.contains(points)
df['cl'] = center_left.contains(points)
df['cr'] = center_right.contains(points)
df['dl'] = down_left.contains(points)
df['dc'] = down_center.contains(points)
df['dr'] = down_right.contains(points)
df.to_csv(f'df{n}_{s}_{dx}_{ds}.csv')

return 1/lim_q if lim_q > 0 else 0, 1/cc_q if cc_q > 0 else 0

for dx_, ds_ in itertools.product(dx, ds):
    if dx_ != 0 or ds_ != 0:
        nma, nma_c = calcula_nma(n, x_, s_, dx_, ds_)
        results = results.append(pd.DataFrame([[n, dx_, ds_, nma, nma_c]], columns=['n', 'dx',
'ds', 'NMA', 'NMA_c']), ignore_index=True)

```