UNIVERSIDADE ESTADUAL DA PARAÍBA
CENTRO DE CIÊNCIAS E TECNOLOGIA
DEPARTAMENTO DE ESTATÍSTICA

CLÁUDIO CRUZ DA SILVA

ANÁLISE DA VARIÂNCIA (ANOVA) NA OTIMIZAÇÃO DE PROCESSOS INDUSTRIAIS

CAMPINA GRANDE - PB

Fevereiro de 2018

CLÁUDIO CRUZ DA SILVA

ANÁLISE DA VARIÂNCIA (ANOVA) NA OTIMIZAÇÃO DE PROCESSOS INDUSTRIAIS

Trabalho de Conclusão de Curso apresentado ao Curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de Bacharel em Estatística.

Orientador: Prof. Dr. Ricardo Alves Olinda

CAMPINA GRANDE - PB

Fevereiro de 2018

S586a Silva, Cláudio Cruz da.
Análise da variância (ANOVA) na otimização de processos industriais [manuscrito] : / Claudio Cruz da Silva. - 2018.
39 p. : il. colorido.

Digitado.
Trabalho de Conclusão de Curso (Graduação em Estatística) - Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia, 2018.
"Orientação : Prof. Dr. Ricardo Alves de Olinda, Departamento de Estatística - CCT."

1. Performance industrial. 2. Otimização estatística. 3. Qualidade dos serviços. 4. Análise estatística.

21. ed. CDD 519.5

CLÁUDIO CRUZ DA SILVA

ANÁLISE DA VARIÂNCIA (ANOVA) NA OTIMIZAÇÃO DE PROCESSOS INDUSTRIAIS

Trabalho de Conclusão de Curso apresentado ao Curso de Bacharelado em Estatística do Departamento de Estatística do Centro de Ciências e Tecnologia da Universidade Estadual da Paraíba em cumprimento às exigências legais para obtenção do título de Bacharel em Estatística.

Trabalho aprovado em 26 de fevereiro de 2018.

BANCA EXAMINADORA

Ricardo Alves de Olinda

Dr. Ricardo Alves de Olinda
Universidade Estadual da Paraíba

Gustavo H. E.

Dr. Gustavo Henrique Esteves
Universidade Estadual da Paraíba

Dr. Sílvio Fernando Alves Xavier Júnior
Universidade Estadual da Paraíba

Agradecimentos

A Deus primeiramente, pois sem a vontade e misericórdia dele não estaria aqui presente.

Ao meus pais Carmelita Cruz da Silva e José Miguel da Silva que sempre estiveram ao meu lado em todos os momentos difíceis dessa vida, aos meus irmãos Delmicio Cruz da Silva, Nilza Cruz da Silva, Roberval Cruz da Silva, Robeilton Cruz da Silva, Claudemir Cruz da Silva e Claudineide Cruz da Silva que me incentivaram ao melhor da vida. Também a Márcia Lopes da Silva e Yasmim da Silva Peres que mudaram o meu sentido de viver. Ao Professor Ricardo Alves de Olinda, pela orientação e incentivo, colaborando de maneira excelente para minha vida acadêmica.

Aos docentes do Curso de Bacharelado em Estatística da Universidade Estadual da Paraíba, que auxiliaram na minha formação acadêmica citarei apenas alguns Ana Patrícia Bastos Peixoto, Thiago Almeida, Gustavo Esteves, Diana Maia, Juarez, João Gil de Luna, Giselly Oliveira, Tanyse Kely, Vítória, Kleber Barros, Sílvio Fernando e outros professores que não são do curso, mas contribuem de forma significativa para a minha formação acadêmica.

Aos meus amigos e colegas de curso citarei apenas alguns deles(as) Sônia Eliane Gonçalves dos Santos, Márcia de Lourdes, Wanessa Isthéwany, Damião Flávio, Arnete Campos, Klaini Clemente, Aline Porto, Filipe, Abraão, Leomir, Rayane Santos, Itamara Júnior, Waleska S. Tavares, Bruno, Regina e assim aos demais que não citei mas que permanece em meus pensamentos. Aos meus familiares e a todos que torceram para meu exito pessoal e profissional.

"No principio era o verbo, e o verbo estava com Deus, e o verbo era Deus. $_{(\textit{João}1:1)}$"

Resumo

Desde o pós-guerra, as características das atividades econômicas experimentaram alterações que impuseram diferentes ritmos de desenvolvimento, até o período atual, em que decididamente a competitividade industrial deixou de ser definida pelos ganhos de escala e da produção seriada, passando a ser decidida nos campos da qualidade e da produtividade. A economia de escala está dando lugar a economia de escopo. Neste cenário a manutenção desponta-se como a única função operacional que influencia e melhora os três eixos determinantes da performance industrial ao mesmo tempo, isto é, custo, prazo e qualidade de produtos e serviços. Diante do exposto, as análises realizadas neste estudo têm como objetivo melhorar a qualidade dos serviços prestados pelos mantenedores de uma indústria têxtil na medida em que facilitam o cumprimento de prazos e cronogramas pré-estabelecidos. Ainda como parte do objetivo, pretende-se diferenciar as gratificações do prêmio de produção em relação ao seu desenvolvimento na otimização dos serviços executados. Com os resultados obtidos através das análises estatísticas, pode-se desenvolver métodos na otimização dos processos industriais.

Palavras-chave: Performance industrial, Otimização estatística, Qualidade dos serviços, Análise estatística

Abstract

Since the post-war period, the characteristics of economic activities have undergone changes that imposed different rhythms of development until the present period, when industrial competitiveness was decidedly no longer defined by gains in scale and serial production, being decided in the fields of quality and productivity. The economy of scale is giving way to economy of scope. In this scenario, maintenance stands out as the only operational function that influences and improves the three determinants of industrial performance at the same time, ie cost, term and quality of products and services. In view of the above, the analyzes carried out in this study aim to improve the quality of the services provided by the maintainers of a textile industry as they facilitate the fulfillment of pre-established deadlines and schedules. Still as part of the objective, it is intended to differentiate the bonuses of the production premium in relation to its development in the optimization of the executed services. With the results obtained through the statistical analyzes, one can develop methods in the optimization of the industrial processes.

Key-words: Industrial performance, Statistical optimization, Analysis of experiments

Lista de ilustrações

Lista de tabelas

Sumário

1 Introdução

A elevação do capital investido na indústria, a necessidade crescente de controlar os custos operacionais para manter a competitividade e o reforço da segurança industrial são apenas alguns dos pontos que tornam imprescindível a gestão de ativos e manutenção. Para alcançar as expectativas, a manutenção passou por consideráveis mudanças técnicas como maiores investimentos em computadores e *softwares* para realização de análises de falha e risco. Com isso proporcionaram maior confiabilidade e disponibilidade dos equipamentos. Cada vez mais a manutenção ganha espaço nas organizações, antes vista apenas como o conserto após a falha, a função manutenção tem hoje a responsabilidade de manter padrões de qualidade estabelecidos para alcançar resultados. A questão é como se deve tratar o mantenedor na organização e como seus esforços devem ser reconhecidos e recompensados (ABRAMAN, 2012).

As organizações precisam estar alerta e conscientes a esse fato de grande importância. Quando se informa ao mantenedor sobre o seu desempenho e o que seu superior espera dele, obtém-se uma ajuda significativa para a realização do treinamento, podendo assim, ficar motivado. Sendo assim, é por meio dessa mensuração, que será possível desenvolver formas de recompensa para os mantenedores como um *feedback* (retorno) para avaliar seu desempenho. A avaliação de desempenho tem como conceito agregar valor ao capital da empresa, fornecendo dados para planejar uma política de ampliação e aprimoramento individual e organizacional. Segundo Campos (2004), o cumprimento de prazos é um dos principais requisitos para a qualidade assegurada. A grande interseção do setor de manutenção com o de produção, influenciam diretamente a qualidade e produtividade, fazendo com que o mesmo desempenhe um papel estratégico fundamental na melhoria dos resultados operacionais e financeiros dos negócios (XENOS, 1998).

Dentro desse contexto existe a necessidade de utilizar uma das técnicas estatísticas que contemple essa questão para uma otimização nos resultados. Entre várias técnicas estatísticas utilizou a de Análise de Variância (ANOVA) num esquema Fatorial. A técnica da ANOVA permite avaliar o impacto que estes fatores provocam na característica de interesse. Segundo Montgomery (1991), as técnicas de planejamento e análise de experimento são utilizadas basicamente para melhorar as características de qualidade dos produtos ou processos de fabricação, reduzir o número de testes e otimizar o uso de recursos da empresa (material, tempo dos funcionários, disponibilidades de equipamentos, etc).

Montgomery (1991) faz algumas recomendações sobre o uso de métodos estatísticos para o planejamento experimental:

I) O conhecimento técnico específico, não estatístico sobre o problema deve ser usado;

II) O delineamento experimental deve ser o mais simples possível;

III) Reconhecer a diferença entre o que é significativo estatisticamente e o que é significativo na prática, seja industrial ou de pesquisa;

IV) Reconhecer que a experimentação é um processo iterativo.

O planejamento de experimentos deve-se a Sir Ronald Aylmer Fisher, que durante alguns anos foi responsável pela estatística e análise de dados na Estação Agrícola Experimental em Londres (Rothamsted Experimental Station) - Inglaterra. Fisher foi quem desenvolveu e usou pela primeira vez a técnica de ANOVA (Análise de Variância) como ferramenta primária para a análise estatística de projetos experimentais. Outro autores que contribuíram de maneira significativa para a evolução das técnicas sobre experimentos em ensaios fatoriais foram: Yates, Box, Bose, Kempthorne e Cochran (MONTGOMERY, 1991).

O experimento projetado ou planejado é um teste ou uma série de testes nos quais se induzem mudanças deliberadas ou estímulos nas variáveis de entrada *(inputs)* do processo ou sistema, de tal forma que seja possível observar e identificar os efeitos na resposta ou nas variáveis de saída *(outputs)*. O processo ou sistema de transformação é representado pela combinação de máquina, método, pessoas e outros recursos que transformam uma entrada em produtos acabados ou semi-acabados, com características ou parâmetros específicos.

Button (2001), descreve que esse objetivo geral pode ser dividido em outros objetivos secundários:

i) Identificar as variáveis (fatores de controle) do processo que mais influenciam nos parâmetros de resposta de interesse;
ii) Atribuir valores às variáveis influentes do processo de modo que a variabilidade da resposta de interesse seja mínima ou que o valor do resultado (parâmetro de qualidade) seja próximo do valor nominal;
iii) Atribuir valores às variáveis influentes do processo de modo que o efeito das variáveis não controláveis seja reduzida.

Diante do exposto, as análises realizadas neste estudo têm como objetivo melhorar a qualidade dos serviços prestados pelos mantenedores de uma indústria têxtil na medida em que facilitam o cumprimento de prazos e cronogramas preestabelecidos. Ainda como parte do objetivo, pretende-se diferenciar as gratificações do prêmio de produção em relação ao seu desenvolvimento na otimização dos serviços executados.

2 Fundamentação Teórica

Um experimento planejado é um teste, ou série de testes, no qual são feitas mudanças propositais nas variáveis de entrada de um processo ou sistema, de modo a observar e identificar mudanças correspondentes na resposta de saída. De acordo com Montgomery (2009), o processo pode ser visualizado como uma combinação de máquinas, métodos e pessoas, que transforma um material de entrada em um produto de saída. Este produto de saída pode ter uma ou mais características da qualidade observáveis. A Figura 1 representa um processo ou experimento qualquer com as variáveis de entrada (Fatores controláveis) e o fatores não controláveis. Algumas vezes, esses fatores não controláveis são chamados de fatores de ruído, que possui características experimentais não controladas, seja pelo desconhecimento de sua existência ou pelo alto custo para controlá-los, ou seja, tais variáveis constituem o erro do experimento (MONTGOMERY; RUNGER, 2011).

É importante evitar que os efeitos produzidos pelos fatores controláveis, fiquem misturados ou mascarados com os efeitos provocados pelos fatores não controláveis (GALDÁMEZ, 2002).

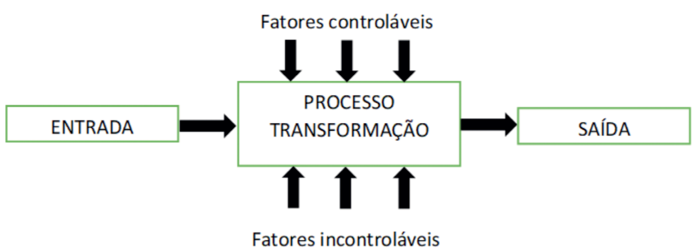

Figura 1 – Modelo geral de um sistema processo ou experimento qualquer (MONTGOMERY; RUNGER, 2011)

Observa-se que uma variável de entrada é considerada controlável se os valores que ela assumir, denominados de níveis, forem definidos antes do início dos experimentos. As variáveis de entrada isto é, controláveis que é de interesse do pesquisador, comumente, são denominadas de fatores, cuja variação pode ou não influenciar a resposta final do processo, seja sozinha ou interagindo com uma ou mais variáveis do mesmo processo. As combinações possíveis entre os níveis dos fatores são denominadas de tratamentos (MONTGOMERY, 2009; BORTOLINI, 2012). É importante ressaltar que nem todos os fatores afetam o desempenho da mesma forma, isto é, alguns fatores podem apresentar fortes influências, enquanto que outros podem nem ter efeito na variável resposta (y) e poderão ser descartados à posteriori (ANTONY, 2003). Atualmente experimentos planejados têm tido aplicação ampla nos negócios de transações, comércio, indústria química, farmacêutica, agrícola e têxtil. As aplicações incluem planejamento de páginas da *Web*, teste de preferências de consumidores e planejamento/melhoria de sistemas de serviços. Algumas vezes, um modelo de simulação computacional é desenvolvido e são realizados testes desse modelo (BARROS NETO 1995).

Os objetivos do experimento podem incluir as seguintes etapas:

1. Determinação de quais variáveis são mais influentes na resposta, y.

2. Determinação dos valores a serem atribuídos aos x's influentes de modo que y esteja perto da exigência nominal.

3. Determinação dos valores a serem atribuídos aos x's influentes de modo que a variabilidade em y seja pequena.

4. Determinação dos valores a serem atribuídos aos x's influentes de modo que os efeitos das variáveis não controláveis sejam minimizados.

Assim, métodos de planejamento experimental podem ser usados tanto no desenvolvimento do processo quanto na solução de problemas do processo, para melhorar o seu desempenho ou obter um processo que seja robusto ou não sensível a fontes externas de variabilidade.

2.1 Diretrizes para o Planejamento de Experimentos

Os experimentos planejados são uma abordagem poderosa para a melhoria de um processo. Para usar essa abordagem, é necessário que todos os envolvidos no experimento tenham uma ideia prévia clara do seu objetivo, de exatamente quais fatores devem ser estudados, de como o experimento deve ser conduzido e, pelo menos, uma compreensão qualitativa de como os dados serão analisados. Em síntese, segundo Costa (2008), Gomes (2007) e Montgomery (2009) para que se possa aplicar a metodologia da estatística experimental em qualquer processo, alguns requisitos básicos tornam-se necessários e serão descritos a seguir:

1. Reconhecimento e relato do problema. Na prática, é, em geral, difícil perceber-se que existe um problema que exige experimentos planejados formais, de modo que pode não ser fácil obter-se um relato do problema, claro e aceito por todos. No entanto, é absolutamente essencial desenvolverem-se completamente todas as ideias sobre o problema e sobre os objetivos específicos do experimento. Usualmente, é importante solicitarem-se entradas de todas as partes envolvidas – engenharia, qualidade, marketing, cliente, gerência e operadores (que, em geral, têm muito discernimento que costuma ser ignorado). Um relato claro do problema e dos objetivos do experimento costuma contribuir substancialmente para uma melhor compreensão do processo e para uma eventual solução do problema.

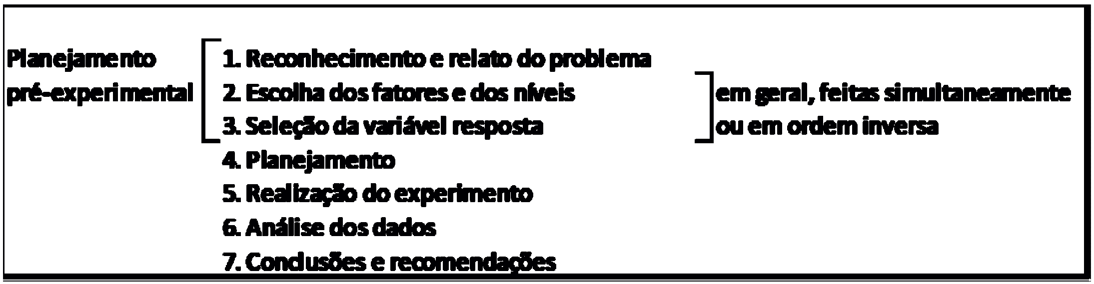

Figura 2 – Procedimento para o planejamento de um experimento (MONTGOMERY; RUNGER, 2011)

2. Escolha dos fatores e dos níveis. A pessoa que conduz o experimento deve escolher os fatores que devem variar, os intervalos sobre os quais esses fatores variarão e os níveis específicos nos quais cada rodada será feita. Exige-se conhecimento do processo para fazer isso. Esse conhecimento é, em geral, uma combinação de experiência prática e conhecimento teórico. É importante investigar todos os fatores que possam ser importantes e evitar ser excessivamente influenciado pela experiência passada, particularmente nos estágios iniciais do experimento ou quando o processo não está ainda muito amadurecido. Quando o objetivo é a varredura dos fatores ou caracterização do processo, é, em geral, melhor manter baixo o número de níveis dos fatores. (Em geral, são usados dois níveis.) Como observado na Figura 2, os passos 2 e 3 são, quase sempre, realizados simultaneamente, ou o passo 3 pode ser feito antes, em algumas aplicações.

3. Seleção da variável resposta. Na seleção da variável resposta, o experimentador deve ter certeza de que aquela variável realmente fornece informação útil sobre o processo em estudo. Muitas vezes, a média ou o desvio-padrão (ou ambos) da característica medida será a variável resposta. Respostas múltiplas não são raras. A capacidade do medidor é, também, um fator importante. Se a capacidade do medidor for baixa, então apenas efeitos grandes de fatores serão detectados pelo experimento, ou será necessária replicação adicional.

4. Escolha do planejamento experimental. Se os três primeiros passos forem feitos corretamente, este passo será relativamente fácil. A escolha do planejamento envolve consideração sobre o tamanho da amostra (número de repetições), seleção de uma ordem adequada de rodadas para as tentativas experimentais, ou se a formação de blocos ou outras restrições de aleatorização estão envolvidas.

5. Realização do experimento. Quanto a realização do experimento, é de vital importância o monitoramento do processo, para garantir que tudo esteja sendo feito de acordo com o planejamento. Erros no procedimento experimental nesse estágio, em geral, destruirão a validade do experimento. O planejamento geral, do início até o fim, é crucial para o sucesso. É fácil subestimarem-se os aspectos logísticos e de planejamento em um ambiente industrial complexo.

6. Análise dos dados. Métodos estatísticos devem ser usados para a análise dos dados, de modo que os resultados e conclusões sejam objetivos e não opiniões. Se o experimento foi planejado corretamente e se foi realizado de acordo com o planejamento, então o tipo de método estatístico exigido não é complicado. Muitos pacotes estatísticos excelentes estão disponíveis para ajudar na análise de dados, e métodos gráficos simples desempenham um papel importante na interpretação dos dados. A análise dos resíduos e a verificação da validade do modelo são também importantes.

7. Conclusões e recomendações. Uma vez analisados os dados, o experimento deve acarretar conclusões práticas sobre os resultados e recomendar um curso de ação. Métodos gráficos são, em geral, úteis nesse estágio, particularmente na apresentação dos resultados para outras pessoas. Sequências de acompanhamento e testes de confirmação também devem ser realizados para validação das conclusões do experimento.

Os passos 1 a 3 são usualmente chamados de planejamento pré-experimental. Para o sucesso do experimento, é vital que esses passos sejam realizados tão bem quanto possível (COLEMAN; MONTGOMERY, 1993). Durante todo esse processo, é importante lembrar que a experimentação é uma parte importante do processo de aprendizagem, em que, por tentativa, formulamos hipóteses sobre um sistema, realizamos experimentos para investigar essas hipóteses e, com base nos resultados, formulamos novas hipóteses e assim por diante. Isso sugere que a experimentação seja iterativa. Usualmente, é um grande erro planejar-se um único, grande e abrangente experimento logo no início do estudo.

Um experimento bem-sucedido exige conhecimento dos fatores importantes, dos intervalos nos quais esses fatores variarão, do número apropriado de níveis a serem usados e das unidades de medidas adequadas a essas variáveis. Em geral, não sabemos perfeitamente as respostas a essas perguntas, mas aprendemos sobre elas na medida em que caminhamos. À medida que avança um programa experimental, em geral tiramos algumas variáveis, acrescentamos outras, mudamos a região de exploração de alguns fatores, ou acrescentamos novas variáveis resposta.

2.2 Seleção do Planejamento

Contudo, como em todo experimento, além da necessidade de seguir um roteiro de execução e diagnóstico, existem princípios básicos que regem sua elaboração, implementação e análise. Segundo Montgomery (2009), os princípios básicos de um bom planejamento experimental são:

- **Repetição:** é o processo de repetição do experimento básico, o que permite ao pesquisador obter estimativas de como o erro experimental afeta os resultados dos ensaios e se esses resultados são estatisticamente diferentes. Outro ponto importante é que se a média da amostra é utilizada para estimar o efeito de um fator no experimento, a repetição também permite obter uma estimativa mais precisa desse efeito. No entanto, quanto mais repetições, maior é o custo do experimento.

- **Aleatorização:** tanto a alocação do material quanto a ordem na qual são realizados os experimentos são determinadas aleatoriamente, o que faz com que as variáveis estudadas e os erros experimentais apresentem caráter aleatório. Contudo, em alguns casos, por restrições de custo na experimentação e pela existência de fatores cuja mudança dos níveis é limitada ou difícil nem sempre isso é possível.

- **Controle local:** em muitas situações não se faz necessário. Este deve ser empregado em situações especiais. Na experimentação agrícola, por exemplo, é muito frequente a instalação de experimentos nas encostas onde o solo, em geral, apresenta um gradiente de fertilidade de menos a mais e de cima para baixo, fato explicado pela ação das chuvas. Neste caso, o controle local é feito, instalando-se um conjunto completo de tratamentos em faixas transversais à declividade do solo. A essa faixa de terra chamaremos de bloco. A ideia é que o bloco seja um ambiente homogêneo contendo todos os tratamentos, de modo que a variabilidade observada na variável resposta possa ser atribuída, em grande parte, aos efeitos dos tratamentos e do acaso e não ao efeito da declividade (fertilidade) do solo.

2.3 Experimentos Fatoriais

Quando vários fatores são de interesse em um experimento, um experimento fatorial deve ser usado. Como notado previamente, fatores são variados conjuntamente nesses experimentos. Assim, se houver dois fatores A e B, com a níveis do fator A e b níveis do fator B, cada repetição conterá todas as ab combinações de tratamentos.

O efeito de um fator é definido como a variação na resposta, produzida pela mudança no nível do fator. Ele é chamado de um efeito principal porque ele se refere a fatores primários no estudo. (MONTGOMERY; RUNGER, 2011). Exemplo abaixo:

Tabela 1 – Exemplo hipotético de um Experimento Fatorial com dois Fatores.

Fator A	Fator B	
	B_{baixo}	B_{alto}
A_{baixo}	10	20
A_{alto}	30	40

Consideramos os dados da Tabela 1. Esse é o experimento fatorial com dois fatores A e B, cada um com dois níveis A_{baixo}, A_{alto} e B_{baixo}, B_{alto}. O efeito principal do fator A é a diferença entre a resposta média no nível alto de A e a resposta média no nível baixo de A,

$$A = \left(\frac{30+40}{2}\right) - \left(\frac{10+20}{2}\right) \quad = \quad 20,$$

ou seja, a variação no fator A do nível baixo para o nível alto faz a resposta média aumentar 20 unidade. Similarmente, o efeito principal de B é.

$$B = \left(\frac{20+40}{2}\right) - \left(\frac{10+30}{2}\right) \quad = \quad 10$$

Em alguns experimentos, a diferença na resposta entre os níveis de um fator não é a mesma em todos os níveis do outro fator. Quando isso ocorre, há uma interação entre os fatores. Exemplo abaixo:

Tabela 2 – Exemplo hipotético de um Experimento Fatorial com dois Fatores e Interação.

Fator A	Fator B	
	B_{baixo}	B_{alto}
A_{baixo}	10	20
A_{alto}	30	0

Consideremos os dados da Tabela 2. No nível baixo do fator B, o efeito de A é

$$A = 30 - 10 \quad = \quad 20.$$

e no nível alto do fator B, o efeito de A é

$$A = 0 - 20 \quad = \quad -20.$$

Uma vez que o efeito de A depende do nível escolhido para o fator B, há interação entre A e B.

Quando uma interação é altamente significativa, os efeitos principais correspondentes têm muito pouco significado prático.

Exemplo:

Usando os dados da Tabela 2, encontramos o efeito principal de A como

$$A = \left(\frac{30+0}{2}\right) - \left(\frac{10+20}{2}\right) \quad = \quad 0.$$

Então podemos ser tentados a concluir que não há efeito do fator A. No entanto, quando examinamos os efeitos de A em diferentes *níveis do fator* B, vimos que esse não foi o caso. O efeito do fator A depende dos níveis do fator B. Assim, o conhecimento da interação AB é mais útil que o conhecimento do efeito principal. Uma interação significativa pode mascarar o significado dos efeitos principais. Consequentemente, quando a interação está presente, os efeitos principais dos fatores envolvidos na interação podem não ter muito significado. É fácil estimar o efeito de interação nos experimentos fatoriais, como aqueles ilustrado nas Tabelas 1 e 2. Nesse tipo de experimento, quando ambos os fatores têm dois níveis, o efeito de interação AB é a diferença nas médias da diagonal. Isso representa metade da diferença entre os efeitos de A nos dois níveis de B. Exemplo abaixo:

Na Tabela 1, encontramos o efeitos de interação AB como sendo

$$AB = \left(\frac{20+30}{2}\right) - \left(\frac{10+40}{2}\right) \quad = \quad 0.$$

Assim, não há interação entre A e B. Na Tabela 2, o efeito de interação AB é

$$AB = \left(\frac{20+30}{2}\right) - \left(\frac{10+0}{2}\right) \quad = \quad 20.$$

Como notamos antes, o efeito de interação nesses dados é muito grande. O conceito de interação pode ser ilustrado graficamente de várias maneiras. A Figura 3 apresenta o gráfico dos dados da Tabela 1 contra os níveis de A para ambos os níveis de B. Note que as linhas de B_{baixo} e B_{alto} são aproximadamente paralelas, indicando que os fatores AB não interagem significativamente. A Figura 4 apresenta um gráfico dos dados da Tabela 2, nesse caso as linhas de B_{baixo} e B_{alto} não são paralelas, indicando a interação entre os fatores A e B.

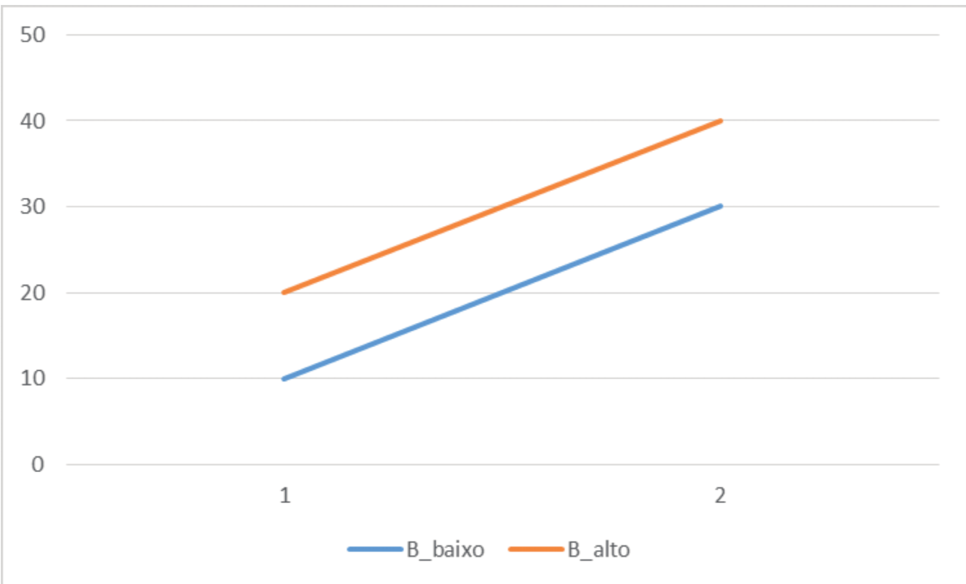

Figura 3 – Experimento fatorial sem interação - Fator A (MONTGOMERY; RUNGER, 2011)

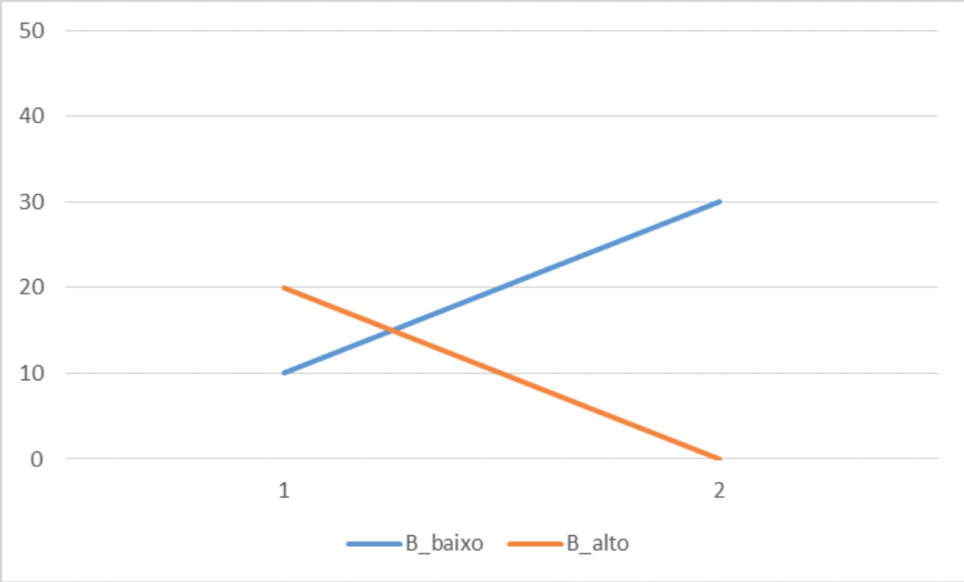

Figura 4 – Experimento fatorial com interação - Fator A (MONTGOMERY; RUNGER, 2011)

Os gráficos de interação entre dois fatores são frequentemente úteis na apresentação dos resultados de experimentos, e muito programas computacionais usados para análise de dados a partir de experimentos planejados fazem esses gráficos automaticamente.

2.3.1 Planejamento Fatorial Generalizado

Ao planejar os experimentos industriais com a técnica fatorial, consideramos que todos os tratamentos da matriz experimental são realizados pela equipe responsável por esta atividade. Segundo Button (2001), o planejamento fatorial é indicado para a fase inicial do procedimento experimental quando há necessidade de se definir os fatores mais importantes e estudar os efeitos sobre a variável resposta escolhida. Para um entendimento melhor dessa técnica, consideramos um experimento com três fatores A, B e C, cada um desses parâmetros serão testados com a níveis para o fator A, com b níveis para o fator B e c níveis para o fator C. Dessa forma obteremos nesse experimentos abc combinações de tratamentos.

Muitos experimentos envolvem mais de dois fatores, que podem ser introduzidos da seguinte forma:

- a níveis do fator A;
- b níveis do fator B;
- c níveis do fator C;
- ... assim por diante.

No caso com R repetições do experimento:

- 1,2,3,...R observações totais.

Assim, como no caso de dois fatores, é necessário que $n \geq 2$ para determinar SQ_{Res}, (se todas as interações entrarem no modelo). Caso os fatores são todos fixos, o modelo é facilmente obtido, e a análise é similar ao de dois fatores (MONTGOMERY; RUNGER, 2011).

2.3.2 Organização dos dados

Para facilitar as análises estatísticas, os dados são organizados em uma matriz (Quadro) de um esquema fatorial num delineamento experimental desejado, conforme Tabela 3.

Tabela 3 – Organização dos dados de um experimento fatorial com três fatores com n replicações (MONTGOMERY; RUNGER, 2011).

Fator A	Fator B	Fator C	Repetições				Total
			1	2	...	r	
A_1	B_1	C_1	y_{1111}	y_{1112}	...	y_{111r}	$y_{111.}$
	B_2	C_2	y_{1221}	y_{1222}	...	y_{122r}	$y_{122.}$
	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	B_J	C_k	y_{1Jk1}	y_{1Jk2}	...	y_{1Jkr}	$y_{1Jk.}$
A_2	B_1	C_1	y_{2111}	y_{2112}	...	y_{211r}	$y_{211.}$
	B_2	C_2	y_{2221}	y_{2222}	...	y_{222r}	$y_{222.}$
	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	B_J	C_k	y_{2Jk1}	y_{2Jk2}	...	y_{2Jkr}	$y_{2Jk.}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
A_I	B_1	C_1	y_{I111}	y_{I112}	...	y_{I11r}	$y_{I11.}$
	B_2	C_2	y_{I221}	y_{I222}	...	y_{I22r}	$y_{I22.}$
	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	B_J	C_k	y_{IJk1}	y_{IJk2}	...	y_{IJkr}	$y_{IJk.}$
Total			$y_{...1}$	$y_{...2}$	...	$y_{...r}$	$y_{....}$

2.3.3 Modelo estatístico para três fatores fixos

Um modelo matemático consiste de um conjunto de equações que representam de uma forma quantitativa, as hipóteses que foram usadas na construção do modelo, as quais se apoiam sobre o sistema real. Tais equações são resolvidas em função de alguns valores conhecidos ou previstos pelo modelo real e podem ser testadas através da comparação com os dados conhecidos ou previstos com as medidas realizadas no mundo real. As equações matemáticas de um modelo não proporcionam a própria explicação científica do modelo, mas simplesmente interpretam as hipóteses de um ponto de vista quantitativo, dando-nos a condição de deduzir consequências e mostrar-nos onde estão os detalhes que deverão ser aceitos ou recusados. A formula (2.1) representa um modelo estatístico para três fatores.

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_k + (\tau\beta)_{ij} + (\tau\gamma)_{ik} + (\beta\gamma)_{jk} + (\tau\beta\gamma)_{ijk} + (\epsilon)_{ijkl} \tag{2.1}$$

com, $i = 1, ..., a;\ j = 1, ..., b;\ k = 1, ..., c$ e $l = 1, ..., n$,
em que

- $y_{ijkl} \longleftarrow$ é o valor da variável resposta obtido da i-ésima unidade experimental que recebeu o i-ésimo nível do fator A, j-ésimo nível do fator B e o k-ésimo nível do fator C;
- $\mu \longleftarrow$ Média geral;
- $\tau_i \longleftarrow$ é o efeito do i-ésimo nível do fator A sobre a variável resposta;
- $\beta_j \longleftarrow$ é o efeito do j-ésimo nível do fator B sobre a variável resposta;
- $\gamma_k \longleftarrow$ é o efeito do k-ésimo nível do fator C sobre a variável resposta;
- $(\tau\beta)_{ij} \longleftarrow$ é o efeito da interação do i-ésimo nível do fator A com o j-ésimo do fator B sobre a variável resposta;
- $(\tau\beta)_{ik} \longleftarrow$ é o efeito da interação do i-ésimo nível do fator A com ok-ésimo do fator C sobre a variável resposta;
- $(\beta\gamma)_{jk} \longleftarrow$ é o efeito da interação do j-ésimo nível do fator B com o k-ésimo do fator C sobre a variável resposta;
- $(\tau\beta\gamma)_{ijk} \longleftarrow$ é o efeito da interação do i-ésimo nível do fator A, o j-ésimo do fator B com o k-ésimo do fator C sobre a variável resposta;
- $(\epsilon)_{ijkl} \longleftarrow$ é o erro casual atribuído aos efeitos de outros fatores não controláveis (Resíduo do modelo)

Inicialmente, deve-se fazer uma análise de variância (ANOVA) preliminar de acordo com o delineamento utilizado para distribuir os tratamentos, em seguida, desdobram-se os graus de liberdade de tratamentos, de acordo com esquema fatorial utilizado, com a finalidade de estudarmos os efeitos principais dos fatores e os efeitos das interações entre os fatores.

2.3.4 Decomposição da variabilidade total e efeitos principais

De acordo com Momtgomery (2011), a decomposição da variabilidade total e dos efeitos principais da Análise da Variância é dado por:

$$
\begin{aligned}
C &= \frac{(y_{....})^2}{abcn} \\
SQ_{Total} &= \sum_{ijkl} y_{ijlk}^2 - C \\
SQ_A &= \frac{1}{bcr}\sum_{i=1}^{a} y_{i...}^2 - C \\
SQ_B &= \frac{1}{acr}\sum_{j=1}^{b} y_{.j..}^2 - C \\
SQ_C &= \frac{1}{abr}\sum_{k=1}^{c} y_{..k.}^2 - C
\end{aligned}
$$

SQ_A, SQ_B e SQ_C são chamados de soma dos quadrados dos efeitos principais onde o r é repetição.

2.3.5 Decomposição da variabilidade da interações entre pares

Segundo Montgomery (2011), a forma para calcular a decomposição da variabilidade da interações entre pares da Análise da Variância é dada por:

$$
\begin{aligned}
SQ_{AB} &= \frac{1}{cr}\sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij..}^2 - C - SQ_A - SQ_B \\
SQ_{AC} &= \frac{1}{br}\sum_{i=1}^{a}\sum_{k=1}^{c} y_{i.k.}^2 - C - SQ_A - SQ_C \\
SQ_{BC} &= \frac{1}{ar}\sum_{j=1}^{b}\sum_{k=1}^{c} y_{.jk.}^2 - C - SQ_B - SQ_C
\end{aligned}
$$

SQ_{AB}, SQ_{AC} e SQ_{BC} são chamados de soma dos quadrados da interação emtre A, B e C em que, o r é repetição.

2.3.6 Decomposição da variabilidade da interações de três fatores

A decomposição da variabilidade da interações de três fatores para análise da variância pode ser obtida através da formula abaixo (MONTGOMERY, 2011):

$$SQ_{ABC} = \frac{1}{r}\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{c} y_{ijk.}^2 - C - SQ_A - SQ_B - SQ_C - SQ_{AB} - SQ_{AC} - SQ_{BC}$$

$$SQ_E = SQ_{Total} - SQ_A - SQ_B - SQ_C - SQ_{AB} - SQ_{AC} - SQ_{BC} - SQ_{ABC}$$

SQ_E é a soma dos quadrados do erro.

2.3.7 Formação dos quadrados médios

Segundo Montgomery (2011), os quadrados médios da análise de variância é dada por:

$$QM_A = \frac{SQ_A}{\textit{(a-1)}}$$

$$QM_B = \frac{SQ_B}{\textit{(b-1)}}$$

$$QM_C = \frac{SQ_C}{\textit{(c-1)}}$$

$$QM_{AB} = \frac{SQ_{AB}}{\textit{(a-1)(b-1)}}$$

$$QM_{AC} = \frac{SQ_{AC}}{\textit{(a-1)(c-1)}}$$

$$QM_{BC} = \frac{SQ_{BC}}{\textit{(b-1)(c-1)}}$$

$$QM_{ABC} = \frac{SQ_{ABC}}{\textit{(a-1)(b-1)(c-1)}}$$

em que QM representa os Quadrados Médios.

2.3.8 Hipóteses à serem testadas

Conforme Montgomery (2011), geralmente num delineamento experimental em esquema fatorial, o maior interesse é testar a hipótese de que os efeitos dos níveis de cada fator sobre a variável resposta sejam estatisticamente nulos, assim como se existe ou não a interação entre os fatores. Então temos:

- Hipótese referente ao fator A

$$\begin{cases} H_0^{(A)} : \mu_1 = ... = \mu_I = \mu \\ H_1^{(A)} : \mu_i \neq \mu_i^{'} \end{cases}$$

- Hipótese referente ao fator B

$$\begin{cases} H_0^{(B)} : \mu_1 = ... = \mu_J = \mu \\ H_1^{(B)} : \mu_j \neq \mu_j^{'} \end{cases}$$

- Hipótese referente ao fator C

$$\begin{cases} H_0^{(C)} : \mu_1 = ... = \mu_K = \mu \\ H_1^{(C)} : \mu_k \neq \mu_k^{'} \end{cases}$$

- Hipótese referente ao fator $A \times B$

$$\begin{cases} H_0^{(A\times B)} : \mu_{11} = ... = \mu_{IJ} = \mu \\ H_1^{(A\times B)} : \mu_{ij} \neq \mu_{ij}^{'} \end{cases}$$

- Hipótese referente ao fator $A \times C$

$$\begin{cases} H_0^{(A\times C)} : \mu_{11} = ... = \mu_{IK} = \mu \\ H_1^{(A\times C)} : \mu_{ik} \neq \mu_{ik}^{'} \end{cases}$$

- Hipótese referente ao fator $B \times C$

$$\begin{cases} H_0^{(B\times C)} : \mu_{11} = ... = \mu_{JK} = \mu \\ H_1^{(B\times C)} : \mu_{jk} \neq \mu_{jk}^{'} \end{cases}$$

- Hipótese referente ao fator $A \times B \times C$

$$\begin{cases} H_0^{(A\times B\times C)} : \mu_{111} = ... = \mu_{IJK} = \mu \\ H_1^{(A\times B\times C)} : \mu_{ijk} \neq \mu_{ijk}^{'} \end{cases}$$

Portanto, o teste para verificar tanto a hipótese nula da igualdade das médias dos fatores A, B, C e da interação $A \times B \times C$, é o teste F, que pode ser calculado por meio do QM_A com QM_{Res}, QM_B com QM_{Res}, QM_C com QM_{Res}, $QM_A \times B$ com QM_{Res}, $QM_A \times C$ com QM_{Res}, $QM_B \times C$ com QM_{Res} e $QM_A \times B \times C$ com QM_{Res}, em que, tais resultados podem ser organizados na tabela (ANOVA).

Tabela 4 – ANOVA - Análise de Variância para o modelo de experimento com três fatores e n repetições (MONTGOMERY; RUNGER, 2011).

Fonte de variação	gl	Soma dos Quadrados	Quadrados Médio	F_0
A	*a-1*	SQ_A	QM_A	$\frac{QM_A}{QM_E}$
B	*b-1*	SQ_B	QM_B	$\frac{QM_B}{QM_E}$
C	*c-1*	SQ_C	QM_C	$\frac{QM_C}{QM_E}$
AB	*(a-1)(b-1)*	SQ_{AB}	QM_{AB}	$\frac{QM_{AB}}{QM_E}$
AC	*(a-1)(c-1)*	SQ_{AC}	QM_{AC}	$\frac{QM_{AC}}{QM_E}$
BC	*(b-1)(c-1)*	SQ_{BC}	QM_{BC}	$\frac{QM_{BC}}{QM_E}$
ABC	*(a-1)(b-1)(c-1)*	SQ_{ABC}	QM_{ABC}	$\frac{QM_{ABC}}{QM_E}$
Erro	*abc(n-1)*	SQ_E	QM_E	-
Total	*abcn-1*	SQ_T	-	-

2.3.9 Desdobramento dos graus de liberdade

Segundo Montgomery (1997), caso as hipóteses $H_0^{(A\times B)}$ seja rejeitada na Tabela 4 devemos desdobrar os graus de liberdade relativos aos efeitos do fator A ou B e da interação, considerando o efeitos de cada fator dentro de cada nível do outro. Então para encontrar as somas dos quadrados temos:

- Soma de quadrados do efeito do fator A dentro dos níveis de fator B

$$
\begin{aligned}
SQ_{(A\ d.\ B_1)} &= \frac{1}{r}\sum_{i=1}^{a} y_{i1..}^2 - \frac{y_{.1..}^2}{Ir} \\
SQ_{(A\ d.\ B_1)} &= \frac{1}{r}\sum_{i=2}^{a} y_{i1..}^2 - \frac{y_{.2..}^2}{Ir} \\
\vdots &= \vdots \\
SQ_{(A\ d.\ B_1)} &= \frac{1}{r}\sum_{i=1}^{a} y_{iJ..}^2 - \frac{y_{.1J..}^2}{Ir}, \quad \text{onde o r é repetição}
\end{aligned}
$$

- Soma de quadrados do efeito do fator B dentro dos níveis de fator A

$$\begin{aligned}
SQ_{(B\ d.\ A_1)} &= \frac{1}{r}\sum_{i=1}^{a} y_{1j..}^2 - \frac{y_{1...}^2}{Jr} \\
SQ_{(B\ d.\ A_2)} &= \frac{1}{r}\sum_{i=2}^{a} y_{2j..}^2 - \frac{y_{2...}^2}{Jr} \\
\vdots &= \vdots \\
SQ_{(B\ d.\ A_I)} &= \frac{1}{r}\sum_{i=1}^{a} y_{iJ..}^2 - \frac{y_{I...}^2}{Jr}, \quad \text{onde o } r \text{ é repetição.}
\end{aligned}$$

As hipóteses a serem testadas no desdobramento dos graus de liberdade são:

- Hipótese do efeito do fator A dentro dos níveis de fator B

$$\left\{ H_0^{(Ad.\ B_j)} : \mu_{1j} = ... = \mu_{Ij} = \mu \right.$$

- Hipótese do efeito do fator B dentro dos níveis de fator A

$$\left\{ H_0^{(B\ d.\ A_i)} : \mu_{i1} = ... = \mu_{iJ} = \mu \right.$$

Depois organizamos os resultados acima em uma nova Tabela (ANOVA). Nessa Tabela (5 e 6) verificamos os comportamentos dos efeitos do fator A dentro dos níveis B e os efeitos do fator B dentro dos níveis A. Com isso vemos se existe diferença significativa entre os efeitos de cada fator.

Tabela 5 – (ANOVA) - Análise de Variância para o efeito do fator A dentro dos níveis de B (MONTGOMERY; RUNGER, 2011).

Fonte de variação	gl	Soma dos Quadrados	Quadrados Médios	F_0
$Ad.B_1$	*(I-1)*	$SQ_{(Ad.B_1)}$	$QM_{A(Ad.B_1)}$	$\frac{QM_{A(Ad.B_1)}}{QM_E}$
$Ad.B_2$	*(I-1)*	$SQ_{(Ad.B_2)}$	$QM_{A(Ad.B_2)}$	$\frac{QM_{A(Ad.B_2)}}{QM_E}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$Ad.B_j$	*(I-1)*	$SQ_{(Ad.B_j)}$	$QM_{A(Ad.B_j)}$	$\frac{QM_{A(Ad.B_j)}}{QM_E}$
Erro	*(IJ)(r-1)*	SQ_E	QM_E	-
Total	*IJr-1*	SQ_T	-	-

Se nas Tabelas 5 e 6 apresentar algumas das hipóteses rejeitadas, ao nível α de significância desejado, então deveremos continuar com as análises estatísticas.

Tabela 6 – (ANOVA) - Análise de Variância para o efeito do fator B dentro dos níveis de A (MONTGOMERY; RUNGER, 2011).

Fonte de variação	gl	Soma dos Quadrados	Quadrados Médios	F_0
$Bd.A_1$	*(J-1)*	$SQ_{(Bd.A_1)}$	$QM_{A(Bd.A_1)}$	$\frac{QM_{A(Bd.B_1)}}{QM_E}$
$Bd.A_2$	*(J-1)*	$SQ_{(Bd.B_2)}$	$QM_{A(Bd.A_2)}$	$\frac{QM_{A(Bd.B_2)}}{QM_E}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$Bd.A_j$	*(J-1)*	$SQ_{(Bd.B_j)}$	$QM_{A(Bd.A_j)}$	$\frac{QM_{A(Bd.B_j)}}{QM_E}$
Erro	*(IJ)(r-1)*	SQ_E	QM_E	-
Total	*IJr-1*	SQ_T	-	-

2.4 Comprovação da idoneidade do modelo

Antes de concluirmos sobre os achados da pesquisa é necessário validar modelo. Ou seja, é preciso comprovar que as suposições impostas às observações y_{ij} e aos erros e_{ij} são válidas. Sem isso, as nossas conclusões serão dúbias e poderemos correr o risco de concluir erroneamente sobre o que estamos pesquisando. Ou melhor, uma pesquisa que não valida o modelo suposto para descrever os dados não é digna de crédito. Os resíduos constituem a principal ferramenta para comprovar as suposições propostas para descrever os dados. Se $y_{ij} = \mu + \tau_i + e_{ij}$ é apropriado, os resíduos observados e_{ij} refletirão as propriedades exigidas para as perturbações e_{ij}. Esta é a ideia básica na análise dos resíduos. Assim, se as hipóteses relativas ao modelo estão corretas os resíduos variarão aleatoriamente. Ao contrário, se descobrirmos que os resíduos apresentam tendências sistemáticas inexplicáveis, deveremos desconfiar da validez do modelo (MONTGOMERY, 1997).

2.4.1 Análise Residual e validade do Modelo

A análise de variância considera que as observações sejam normal e independentemente distribuídas, com mesma variância para cada tratamento ou nível do fator. Essas suposições devem ser verificadas por meio da análises dos resíduos. Um resíduo é a diferença entre uma observação y_{ij} e seu valor estimado (ou ajustado) a partir do modelo estatístico estudado, denotado como $\hat{y}_{ij}$. Para o planejamento completamente aleatorizado $\hat{y}_{ij} = \bar{y}_{i.}$, com cada resíduo $e_{ij} = y_{ij} - \bar{y}_{i.}$; ou seja, a diferença entre uma observação e a média correspondente observada do tratamento (MONTGOMERY, 1997).

A suposição de normalidade pode ser analisada pela construção de um gráfico de probabilidade normal dos resíduos. Para verificar a suposição de igualdade de variância em cada nível do fator, gerar um gráfico dos resíduos contra os níveis do fator e compare a dispersão dos resíduos. É também útil elaborar um gráfico dos resíduos contra $\bar{y}_{i.}$ (algumas vezes chamado de valor ajustado); a variabilidade nos resíduos não deve depender, de jeito algum, do valor de $\bar{y}_{i.}$ A maioria dos pacotes computacionais estatísticos constrói esses gráficos quando requisitado (MONTGOMERY, 1997).

A suposição de independência pode ser analisada, gerando um gráfico dos resíduos contra o tempo ou a ordem na qual o experimento foi realizado. Um padrão de comportamento nesse gráfico, como sequência de resíduos positivos e negativos, pode indicar que as observações não são independentes. Isso sugere que o tempo ou a ordem é importante ou que as variáveis que variam com o tempo são importantes e não foram incluídas no planejamento do experimental. O efeito da dependência entre os resíduos pode ser um problema muito grave e difícil de corrigir, por isso é importante preveni-lo quando os dados estão sendo colhidos. O método mais eficiente para prevenir a dependência é adotar um procedimento apropriado de aleatorização (MONTGOMERY, 1997).

2.4.2 Transformações nos dados

Caso seja identificada a violação de uma ou mais suposições acerca do modelo, faz-se necessário adotar uma das seguintes medidas corretoras: *I)* Modificar o modelo proposto para descrever os dados (esta atitude pode levar a análises estatísticas extremamente complexas); *II)* Fazer uma transformação adequada dos dados; *III)* Quando a dificuldade básica é a falta de normalidade, deveremos aplicar algum teste estatístico não-paramétrico, tal como, o teste da mediana, o teste de Kruskal-Wallis, o teste de Friedman, dentre outros. Abaixo tipos de transformação para corrigir a falta de normalidade e para estabilizar a variância (MONTGOMERY, 1997).

Transformações para estabilizar a variância

$$\begin{aligned}
&\text{a) Quando } \sigma_y^2 \text{ for proporcional a } \mu \\
y^* &= \sqrt{y}; \\
&\text{b) Quando } \sigma_y \text{ for proporcional a } \mu \\
y^* &= \log y; \\
&\text{c) Quando } \sigma_y \text{ for proporcional a } \mu^2 \\
y^* &= \frac{1}{y}; \\
&\text{d) De modo geral, quando } \sigma_y \propto \mu^\alpha \\
y^* &= y^{1-\alpha}.
\end{aligned}$$

Transformações para corrigir a falta de normalidade

a) Se $y_{ij} \sim bin(n_i, p_i)$, então:

$$y_{ij}^* = arcsen\frac{\sqrt{y_{ij}+\frac{3}{8}}}{\sqrt{r_i+\frac{3}{4}}}$$

b) Se $y_{ij} \sim Pisson(\lambda)$, então:

$$y_{ij}^* = \sqrt{y_{ij}}$$

ou

$$y_{ij}^* = \sqrt{y_{ij}+1}$$

c) Se $y_{ij} \sim Log\ normal$, então:

$$y_{ij}^* = \log y_{ij}$$

d)Transformação de Box & Cox:

$$\text{(estabiliza variância)} \longrightarrow y_{ij}^{(\lambda)} = \begin{cases} \frac{(y_{ij}+m)-1}{\lambda}, & se\ \lambda \neq 0 \\ \ln(y_{ij}+m), & se\ \lambda = 0 \end{cases}$$

$$\text{(normaliza os dados)} \longrightarrow y_{ij}^{(\lambda)} = \begin{cases} \frac{y_{ij}^{\lambda}-1}{\lambda}, & se\ \lambda \neq 0 \\ \ln y_{ij}, & se\ \lambda = 0 \end{cases}$$

2.5 Teste de Normalidade

Segundo Montgomery (2011), os teste de normalidade são utilizados para verificar se a distribuição de probabilidade associada a um conjunto de dados pode ser aproximada pela distribuição normal.

2.5.1 Teste de Shapiro-Willk

O teste de Shapiro-Wilk, proposto em 1965, é baseado na estatística

$$W = \frac{b^2}{\sum_{i=1}^{n}(x_{(i)}-\bar{x})^2}$$

em que $x_{(1)}, x_{(2)}, ..., x_{(n)}$ são os valores da amostra $x_1, x_2, ..., x_n$ ordenados ($x_{(1)}$ é o menor e $x_{(n)}$ é o maior). Menores valores de W são evidências de que os dados são normais. A constante b é determinada da seguinte forma:

$$b = \begin{cases} \sum_{i=1}^{n/2} a_{n-i+1}(x_{(n-i+1)-x_i}), \text{se } n \text{ é par} \\ \sum_{i=1}^{(n/2)/2} a_{n-i+1}(x_{(n-i+1)-x_i}), \text{se } n \text{ é impar.} \end{cases}$$

Os coeficientes a_{n-i+1} são constantes geradas pelas médias, variância e covariância das estatística de ordem de uma amostra aleatória de tamanho n de uma distribuição normal.

Para realizar o teste de Shapiro-Wilk pode-se proceder do seguinte modo:
a) Estabelecer as hipóteses:

H_0: A amostra é proveniente de uma distribuição normal

H_1: A amostra não é proveniente de uma distribuição normal
b) Fixar o nível de significância α do teste (em geral $\alpha = 0,05$);
c) Com os dados da amostra avaliar a estatística de teste:

i. Ordenar as n observações da amostra: $x_{(1)}, x_{(2)}, ..., x_{(n)}$;

ii. Calcular $\sum_{i=1}^{n}(x_{(i)} - \bar{x})^2$;

iii. Calcular b;

iv. Calcular W;
d) Adotar a regra de decisão: Rejeita H_0 ao nível de significância α, se $W < W_{(n,\alpha)}$.

2.6 Teste de comparações múltiplas

Uma análise de variância (ANOVA) rejeita ou não a hipótese de igualdade de médias populacionais de diversos grupos, mas não determina quais grupos têm médias estatisticamente diferentes. Por essa razão, o teste F feito na análise de variância é considerado um teste global. Terminada a análise de variância é necessário aplicar um teste de comparação de médias dos tratamentos, daí podendo concluir qual o melhor tratamento.

2.6.1 Teste de Tukey

O teste de Tukey, permite a formação de intervalos de confiança $100(1 - \alpha)$ simultâneos para todas as comparações em pares (WALPOLE; MYERS; YE; YE, 2008). Suas contribuições sobre o assunto encontram-se principalmente no artigo de Tukey (1962) e no seu livro, Tukey (1977). Para Tukey, os problemas da ciência e suas aplicações tecnológicas, incluindo entre estas a engenharia, a agricultura e a medicina, não iniciam nem terminam com respostas ordenadas, daí a reabilitação da estatística descritiva, começando com análises gráficas e visuais. A análise exploratória, com ênfase nos aspectos descritivos, não elimina a análise confirmatória, de cunho inferencial, mas se completam, como escreve Tukey (1980). Para os estatísticos aplicados, Tukey é conhecido por seu teste para comparar todo e qualquer contraste entre duas médias, baseado na amplitude total "estudentizada" **(studentized range)**, chamada na literatura, de teste de Tukey, cuja aplicação requer tabela especial, que pode ser encontrada no livro de Snedecor (1937). Nesse teste apresenta-se algumas característica na sua a aplicação:

I) Não permite comparar grupos de tratamentos entre si;

II) É utilizado para testar toda e qualquer diferença entre duas médias de tratamento;
III) É aplicado quando o teste F para tratamentos da Análise de Variância (ANOVA) for significativo.

2.6.2 Hipóteses à serem testadas

$$\begin{cases} H_0^{(1)} : \mu_i = \mu_i^{'} \\ H_1^{(1)} : \mu_i \neq \mu_i^{'} \end{cases} \forall_i \neq i^{'}, \quad i, i^{'} = 1, 2, ..., I;$$

$$\begin{cases} H_0^{(2)} : \mu_j = \mu_j^{'} \\ H_1^{(2)} : \mu_j \neq \mu_j^{'} \end{cases} \forall_j \neq j^{'}, \quad j, j^{'} = 1, 2, ..., J;$$

$$\begin{cases} H_0^{(3)} : \mu_{ij} = \mu_i j^{'} \\ H_1^{(3)} : \mu_{ij} \neq \mu_i j^{'} \end{cases} \forall_j \neq j^{'}, \quad i = 1, 2, ..., I \;\; e \;\; j, j^{'} = 1, 2, ..., J;$$

$$\begin{cases} H_0^{(4)} : \mu_{ij} = \mu_{i^{'}j} \\ H_1^{(4)} : \mu_{ij} \neq \mu_{i^{'}j} \end{cases} \forall_i \neq i^{'}, \quad i, i^{'} = 1, 2, ..., I \;\; e \;\; j = 1, 2, ..., J.$$

O teste de Tukey tem como base a DMS (diferença mínima significativa), representada no geral por $\triangle$ e calculada da seguinte forma abaixo:

$$|\hat{y}| \;=\; |m_i - m_i^{'}| \quad e \quad \Delta = q_{[I, IJ(R-1)]}\sqrt{\frac{QME}{JR}} \tag{2.2}$$

$$|\hat{y}| \;=\; |m_j - m_j^{'}| \quad e \quad \Delta = q_{[J, IJ(R-1)]}\sqrt{\frac{QME}{IR}} \tag{2.3}$$

Se $|\hat{y}| \geq \Delta$, rejeitamos a hipóteses H_0 e concluirmos que a média do nível i do fator A difere estatisticamente da média do nível $i^{'}$ ou que a média do nível j do fator B difere estatisticamente da média do nível $j^{'}$, ao nível α de significância pelo teste de Tukey.

2.6.3 Teste de Scott-Knott

Scott & Knott (1974) propuseram um procedimento que compara as médias dos tratamentos por conglomerados e sua significância onde é analisada por meio da distribuição de qui-quadrado. A grande vantagem em sua utilização é proveniente do fato de que nenhuma média pode pertencer a mais de um grupo, como ocorre nos anteriores, ou seja, o teste determina a constituição de grupos disjuntos, sempre que for encontrada significância na ANOVA. O teste de Scott-Knott (1974) utiliza a razão de verossimilhança para atestar a significância de que n tratamentos podem ser divididos em grupos que maximizem a soma de quadrados entre grupos. Por exemplo: (3 tratamentos, A, B e C).

- O processo consiste em determinar uma partição, em dois grupos, que maximize a soma de quadrados;

- Nesse caso são possíveis $2^n - 1$ grupos, isto é, A vs B e C, B vs A e C e C vs A e B.

Para atenuar esse problema, basta ordenar as médias dos tratamentos. Nessa situação, o número de partições possíveis passa a ser obtido por $n - 1$. Uma vez ordenada as médias, procede-se do seguinte modo, fazendo inicialmente o número de tratamentos envolvidos no grupo de médias considerado (g) igual ao o número total de tratamentos (n).

Procedimentos para aplicar o Teste de Scott-Knott:
1) Ordenar as médias;
2) Fazer o número de médias no grupo g igual ao número total de tratamentos;
3) Determinar $g - 1$ partições com dois grupos;
4) Para cada partição obter a soma de quadrados B_0 e verificar qual partição maximiza a soma de quadrados.

$$SQB_0 = \frac{T_1^2}{K_1} + \frac{T_2^2}{K_2} - \frac{(T_1 + T_1)^2}{K_1 + K_2}$$

em que:

- $T1$ e $T2$ são os totais das cada média grupo

$$T_1 = \sum_{i=1}^{K_1} \bar{y}_i \quad T_2 = \sum_{k_1+1}^{g} \bar{y}_i$$

- $K1$ e $K2$ é o numero de médias de cada grupo

5) Obter o estimador de máxima verosimilhança de $\sigma_{\bar{y}}^2$

$$\hat{\sigma}_0^2 = \frac{\sum_{i=1}^{K_1}(y_i - \bar{y})^2 + VS^2}{g + V}$$

em que:

- V grau de liberdade do erro;

- $\bar{y}$ média geral de todos os tratamentos envolvidos na comparação;

- S^2 variância de uma média de tratamento (quadrado médio do erro / r)

6) Determinar o valor da estatística λ para a partição que maximizou a soma de quadrados.

$$\lambda = \frac{\pi B_0}{2\sigma^2{}_0(\pi - 2)}$$

- Se $\lambda \geqslant \chi^2_{\alpha,g/(\pi-2)}$ rejeita-se a hipótese de que os dois grupos são idênticos em favor da hipótese alternativa de que os dois grupos diferem;

- No caso de rejeitar esta hipótese, os dois subgrupos formados serão independentemente submetidos aos passos 3 a 6, fazendo respectivamente $g = k_1$ e $g = k_2$.

- O processo em cada subgrupo se encerra ao se aceitar H_0 no passo 6 ou se cada subgrupo contiver apenas uma média.

3 Aplicação

Os dados do tempo de solução (expressos em minutos) dos mantenedores foram gerados no ano de 2009 entre janeiro e dezembro e foram fornecidos por uma empresa têxtil, localizada no município de Campina Grande, Estado da Paraíba, onde foi realizado um estudo descritivo para cada mantenedor (Man) e tipos de soluções (T.S), cujos resultados dessa estatística encontram-se nas Tabelas 7 e 8, onde pode observar os desvio padrão (D.P), coeficiente de variação (C.V) e coeficiente de assimetria (C.A).

Tabela 7 – Estatísticas descritivas da variável aleatória tempos de soluções da manutenção em minutos no ano de 2009 dos mantenedores de uma empresa têxtil no município de Campina Grande-PB

Man.	Média (mês)	Mediana	Variância	*D.P.*	*C.V.*(%)	*C.A.*	Curtose
A1	1484,17	1166,00	678252,15	823,56	55,49	0,75	-0,30
A2	1416,83	1379,00	475181,42	689,33	48,65	0,10	-0,63
A3	1153,33	999,00	498048,97	705,73	61,19	1,15	2,96
A4	3169,08	3436,50	1213225,17	1101,47	34,76	-0,15	-0,42
A5	497,92	372,50	83237,90	288,51	57,94	0,98	0,38
A6	1698,83	1553,00	53492,33	808,39	47,59	-0,10	-0,29

Pode-se observar por meio da Tabela 7, que os mantenedores $A4$ e $A6$ obtiveram, em média, os valores maiores nos tempos de soluções da manutenção e o mantenedor $A5$ o menor valor, já para os mantenedores $A1$ e $A2$ as médias estão com seus valores próximos. Pode-se observar também que os mantenedores $A1$, $A3$ e $A5$ apresentaram maior variabilidade nos tempos de soluções da manutenção e uma assimetria positiva e os demais uma assimetria negativa, o mantenedor $A3$ apresentou uma curva de frequência mais aberta, com os dados fracamente concentrados em torno de seu centro em relação aos demais.

Tabela 8 – Estatísticas descritivas da variável aleatória tempos de soluções da manutenção no ano de 2009 referente ao tipos de solução de uma empresa têxtil no município de Campina Grande-PB

T.S	Média (mês)	Mediana	Variância	*D.P.*	*C.V.*(%)	*C.A.*	Curtose
B1	143,51	84,00	29171,86	170,80	119,01	2,15	4,67
B2	137,19	104,50	19721,79	140,43	102,36	2,78	10,94
B3	147,58	73,00	47554,19	218,07	147,76	3,78	18,46
B4	226,69	163,50	39622,22	199,05	87,81	1,04	0,08
B5	83,33	32,50	13016,48	114,09	136,91	2,26	5,97
B6	149,61	90,00	28608,21	169,14	113,05	1,56	1,52
B7	682,10	483,00	338636,34	581,92	85,31	1,58	2,90

Na Tabela 8, o tipo de solução $B7$ apresentou em média o seu valor maior com maior desvio padrão e seus dados com alta variabilidade quando comparada aos demais (Variância). Todos os tipos de soluções obtiveram uma assimetria positiva onde $B3$ teve o maior valor assimétrico. Os tipos de soluções $B2$ e $B3$ em relação a sua distribuição apresentaram uma curtose mais aberta (curva de frequência), já $B4$ obtive uma curva mais fechada, quando comparados aos demais.

Tabela 9 – (ANOVA) - Análise de Variância para variável resposta original tempos (em minutos) de soluções da manutenção de uma empresa têxtil na cidade de campina Grande.

Fonte de variação	gl	S.Q	Q.M	F_0	$P - Valor$
$Fator A$	5	20198327	4039665	21,660	1,33e−15 ***
$Fator B$	6	55086343	9181057	49,228	< 2e−16 ***
$Interação$(A x B)	30	28540142	951338	5,101	4,02e−11 ***
$Erro$	126	23499197	186502	−	−
Total	127	127324009	−	−	−
W = 0,90046	p-valor = 3,159e−09				

Significânça: ('***' 0,001 - '**' 0,01 - '*' 0,05.)

Dando sequência as análises, foi realizada a Análise da Variância (ANOVA) com os dados originais (Tabela 9), porém observa-se que os resíduos não seguem normalidade na distribuição, consequentemente não atende aos pressupostos da ANOVA. O teste de normalidade Shapiro-Wilk reforça essas informações ($W = 0,90046$, $p - valor < 0,0001$, Tabela 9). Uma alternativa para solucionar o problema dos pressupostos da ANOVA é por meio da transformação ótima de Box-Cox. Sendo assim, foi utilizado a transformação ótima de Box-Cox (Figura 5) a qual maximiza a informação por meio do método da máxima verossimilhança, indicando assim, uma transformação que poderá corrigir os problemas nos pressupostos da ANOVA.

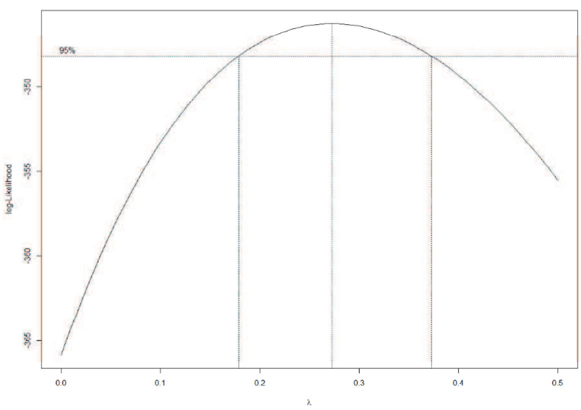

Figura 5 – Transformação ótima de Box-Cox para o modelo ajustado aos dados dos tempos de soluções da manutenção de uma empresa têxtil da cidade de Campiana Grande - PB

Ao observar o gráfico de Box-Cox na Figura 5 verificou o valor de lambda (λ) que maximiza a informação é de aproximadamente $0,25$. É importante ressaltar que, após a transformação dos dados as análises subsequentes deverão ser realizadas com os dados transformados e não com os dados originais. Após a transformação dos dados, deverá observar se os resultados apresentados com os dados transformados têm sentido prático. Isto é, se é possível explicar o fenômeno em estudo com a variável transformada. Em seguida realizou a (ANOVA) com os dados transformados conforme a Tabela 10.

Tabela 10 – (ANOVA) - Análise de Variância para variável resposta transformada tempos (em minutos) de soluções da manutenção de uma empresa têxtil na cidade de Campina Grande.

Fonte de variação	gl	S.Q	Q.M	F_0	$P-Valor$
$FatorA$	5	$62,46$	$12,491$	$21,06$	$2,95e{-}15$ ***
$FatorB$	6	$116,67$	$19,445$	$32,78$	$< 2e{-}16$ ***
$Interação(A$ x $B)$	30	$16,55$	$0,552$	$0,93$	$0,576$
$Erro$	126	$74,75$	$0,593$	–	–
Total	127	$270,43$	–	–	–
$CV.(\%) = 77,20945$	$W = 0,98888$	p-valor $= 0,2092$			

Significânça: ('***' 0,001 - '**' 0,01 - '*' 0,05.)

Após a transformação dos dados pode-se observar, por meio da Tabela 10, que não houve efeito estatisticamente significativo da interação (p-valor>0,05) entre os mantenedores (Fator A) e os tipos de soluções (Fator B), corroborando assim, para a não rejeição da hipótese de igualdade de interação entre os fatores em estudo, ou seja, os mantenedores agem independente do tipo de solução. No entanto, comparando-se os efeitos dos atores isolados, houve diferenças estatisticamente significativas (p-valor<0,05) nos níveis dos mantenedores analisados (Fator A) e nos níveis das soluções estudadas (Fator B), respectivamente.

Uma vez ajustado o modelo da ANOVA, considerando, neste trabalho, um Delineamento Inteiramente Casualisado (DIC) num esquema fatorial com dois fatores (a fundamentação teórica de três fatores abrange todo esse estudo), verificou-se os resíduos graficamente e pelo teste de Shapiro-Wilk para normalidade dos resíduos. De acordo com o teste de normalidade aplicado, pode-se observar (Tabela 10) que os resíduos seguem uma distribuição normal (p-valor >0,05). É importante ressaltar que antes da transformação ótima de Box-Cox, após o ajuste do modelo, considerando os dados originais, os resíduos não seguiram um comportamento normal.

Nem sempre uma transformação nos dados soluciona o problema dos pressupostos da ANOVA. Quando isso acontece, é sugerido ao experimentador/pesquisador ajustar outras distribuições de probabilidade. Uma alternativa que tem se tornado viável na estatística experimental é o ajuste dos modelos pertencente a classe dos Modelos Lineares Generalizados (MLG), atentando para a particularidade de cada situação. Outras classes de modelos têm sido amplamente utilizadas no ajuste da ANOVA, por exemplo, os Modelos Lineares Generalizados Aditivos (GAMLSS - *Generalized Additive Models for Location Scale and Shape*) com parâmetros de locação, escala e forma.

Diante de várias alternativas que o experimentador/pesquisador dispõe atualmente, é importante buscar um modelo mais parcimonioso possível, isto é, o modelo que envolva o mínimo de parâmetros possíveis a serem estimados e que explique bem o comportamento da variável em estudo.

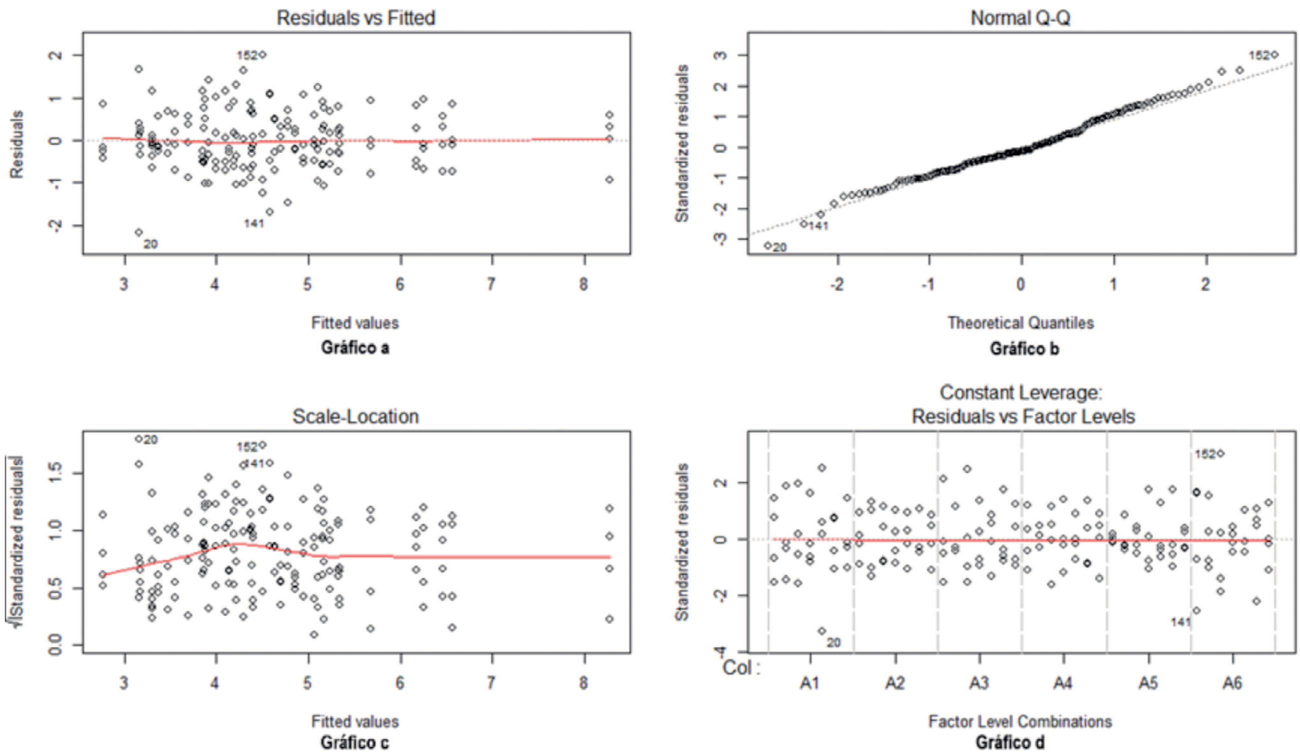

Figura 6 – Gráfico dos resíduos da variável resposta tempos de soluções da manutenção de uma empresa têxtil da cidade de Campiana Grande - PB

Observa-se por meio da Figura 6a os resíduos versus os valores ajustados, visualmente esse gráfico não apresenta tendência nem indicação de suposição de heteroscedasticidade. Na Figura 6b, gráfico superior direito, observa-se o gráfico dos quantis teóricos da distribuição normal versus os resíduos studentizados do modelo ajustado aos dados (Q-Q plot) observou que os resíduos estão distribuídos ao longo de uma reta, dando ao entendimento que há fortes evidências de normalidade dos dados, fato este, corroborado pelo teste de Sharipo-Wilk (p-valor=0,2092, Tabela 10). Na Figura 6c, gráfico inferior esquerdo, observa-se o gráfico da raiz quadrada dos resíduos padronizados versus os valores ajustados, onde os resíduos se apresentam espalhados de forma aleatória, dando ao entendimento que há fortes evidências de homoscedasticidade. Já na Figura 6d, gráfico inferior direito, tem-se o gráfico dos resíduos padronizados versos as combinações de nível de fator, que apresenta a existência de três elementos discrepantes, podendo ser um erro de medida, ou um valor real.

Na sequência, após os resultados apresentados da ANOVA e a verificação dos respectivos pressupostos, aplicou-se um teste de comparações múltiplas para comparar os efeitos das médias, dois a dois, dentro dos níveis de cada fator em estudo. O teste de comparação múltipla utilizado foi o teste de Scott-Knott, conforme os gráficos das Figuras 7 e 8.

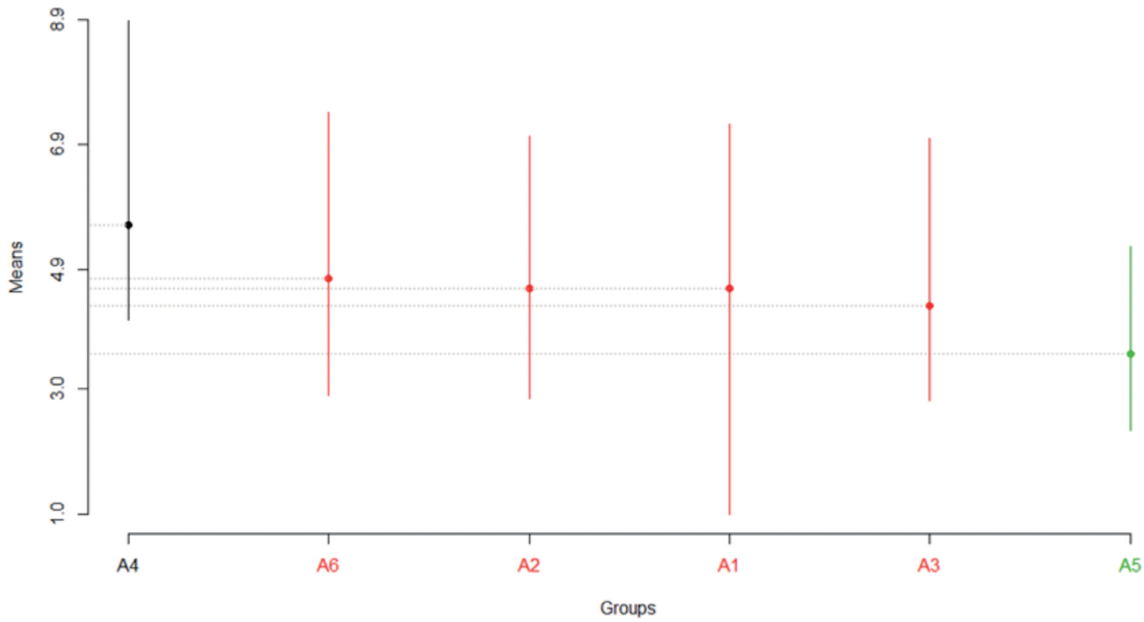

Figura 7 – Gráfico das médias dos tempos de soluções da manutenção de uma empresa têxtil segundo o teste Scott-Knott

Na Figura 7 não houve diferença estatisticamente significativa ao nível de 5% de probabilidade entre os valores médios dos mantenedores $A1$, $A2$, $A3$ e $A6$, ou seja, os tempos de soluções da manutenção não tem diferença entres esses mantenedores. O mantenedor $A4$ apresentou diferença estatisticamente significativa quando comparador aos demais com o maior valor médio nos tempos de soluções da manutenção que significa uma demora maior no tempo para execução da manutenção e o mantenedor $A5$ obteve diferença significativa em relação aos demais com o menor valor médio na execução da manutenção. Com esses resultados pode-se diferenciar as gratificações e desenvolver treinamentos focados para os mantenedores afim de nivelar o conhecimentos e otimizar a manutenção no processos industrial aumentando a qualidade nos serviços prestados e na eficiência da produção.

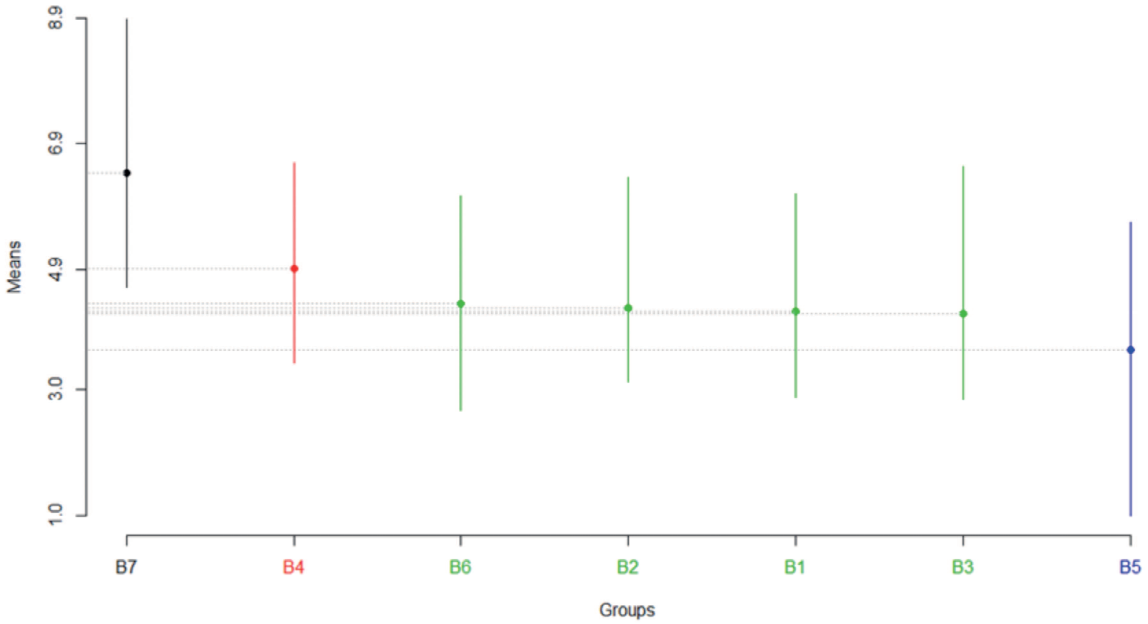

Figura 8 – Gráfico das médias dos tipos de soluções da manutenção de uma empresa têxtil segundo o teste Scott-Knott

No gráfico da Figura 8 observa-se que o valor médio entre os tipos de soluções $B1, B2, B3$ e $B6$ não apresenta diferença estatisticamente significativa, ao nível de 5% de probabilidade conforme o teste de Scott-Knott, levando ao entendimento que não tem diferença de tempo para solucionar esses tipos de soluções entre esses mantenedores. Já os tipos de soluções $B4$ e $B7$ tiveram o valor médio mais alto, ou seja, os mantenedores levaram mais tempo para executar essas manutenções e o tipo de solução $B5$ obteve o menor valor médio no tempo de solução executados pelos mantenedores. Como podemos observar, existe alguns problemas que levam mais tempo para resolver e esse tempo, pode está associado a falta de experiência do mantenedor ou a falta de treinamentos.

4 Conclusão

A técnica de análise da variância (ANOVA) na otimização de processos industriais é uma ferramenta adequada para analisar o comportamental dos tempos de soluções dos mantenedores de uma empresa têxtil da cidade de Campina Grande, onde foram observadas diferenças de tempos nas execução dos serviços de manutenções como nos tipos de soluções.

De acordo com o teste de comparação múltipla os mantenedores apresentaram diferenças na execução dos serviços mostrando três classes diferentes para gratificações, como também três níveis de conhecimentos nos desenvolvimentos dos serviços. Com isso a necessidade de estabelecer diferentes tipos de treinamentos para os mesmos afim de reduzir os tempos de soluções como também otimizar a qualidade dos serviços prestados e custo da manutenção.

5 Referências Bibliográficas

ABRAMAN, Associação brasileiro de manutenção e gestão de ativos, Rio de Janeiro. 2012. Disponível em: http://www.abraman.org.br/sidebar/bibliotecas-e-publicacoes/apostilas-artigos-boletins-e-trabalhos-tecnicos.

ANTONY, J. *Design of experiments for engineers and scientists.* Oxford: Butterworth-Heinemann. 2003. 152p.

BARROS NETO, B.; SCARMÍNIO, I. S.; BRUNS, R. E. Planejamento e otimização de experimentos. Campinas: UNICAMP, 1995. 299p.

BORTOLINI, J. Estudo de experimentos fatoriais 2^k aplicados em um processo industrial. *Dissertação (Mestrado)* - Universidade Federal de Lavras, Lavras. 2012. 143p.

BUTTON, S.T. Metodologia para planejamento experimental e análise de resultado. São Paulo, Universidade Estadual de Campinas, 2001. 88p. /Apostila/

CAMPOS, V. F. Gerenciamento da rotina do trabalho do dia-a-dia. Belo Horizonte: Editora INDG, 8^a edição. 2004. 256p.

COLEMAN, D. E.; MONTEGOMERY, D. C. *A systematic approach to planning for a designed industrial experiment. Technometrics.* 1993. v35.

COSTA, A. F. S. Aplicação de tratamentos biológico e físico-químico em efluentes de lavandeira e tinturaria industriais do município de Toritama no estado de Pernambuco. *Dissertação (Mestrado)*-Universidade Católica de Pernambuco, Recife. 2008. 100p.

GALDÁMEZ, E.V.C. Aplicação das técnicas de planejamento e análise de experimentos na melhoria da qualidade de um processo de fabricação de produtos plásticos.*Dissertação em Mestrado em Engenharia de Produção*, Escola de Engenharia de São Carlos, Universidade de São Paulo. 2002. 133p.

GOMES, U. R. Otimização do Processo de Laminação a Frio através de planejamentos de Experimentos. *Dissertação (Mestrado)* - Pontificia Universidade Católica do Rio de Janeiro, Rio de Janeiro. 2007. 76p.

MONTGOMERY. D.C *Diseno y análisis de experimentos.* Traduzido por Jaime Delgado Saldivar. Mexico, Iberoamérica. 1991.

MONTGOMERY, D. C. *Design and analysis of experiments.* 7. ed. New York: John Wiley & Sons. 2009. 656p.

MONTGOMERY, D.C.; RUNGER, G. C. Estatistica aplicada e probabilidade para engenheiros, quinta edição., p338-400, 2011.

MONTGOMERY, D.C. *Design and Analysis of Experiments* v.5. p.170-387, 1997.

SNEDECOR, G. W. *Statistical methods applied to experiments in agriculture and biology.* Ames: The Iowa Stale College Press, 1937.

SCOTT, A. J.; KNOTT, M. A. *A cluster analysis method for grouping means in the analysis of variance.* Biometrics. Raleigh, v.30, n.3, p.507-512. 1974.

TUKEY, J. W. *The future of data analysis. Annls of Mathematical Statistics*, Baltimore, MD, v.33, 1962. 167p.

TUKEY, J. W. *Exploratory data analysis* Reading: Addison-Wesley. 1977.

TUKEY, J. W. *We nees both exploratory and confirmatory. Tre American Statistician*, Washington, DC, v.34, n. 1, p. 23-25. 1980.

XENOS, H.G Gerenciando a Manutenção Produtiva, Rio de Janeiro: INDG, primeira edição. 1998. 302p.

WALPOLE, R.E.; MYERS, R. H.; MYERS, S. L. Probalidade & Estatística para engenharia e ciências. v.8. p.327-430. 2008.