



JULIANO BORTOLINI

**ESTUDO DE EXPERIMENTOS FATORIAIS
 2^k APLICADOS EM UM PROCESSO
INDUSTRIAL**

LAVRAS - MG

2012

JULIANO BORTOLINI

**ESTUDO DE EXPERIMENTOS FATORIAIS 2^k APLICADOS EM UM
PROCESSO INDUSTRIAL**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de Mestre.

Orientador
Dr. Marcelo Silva de Oliveira

**LAVRAS - MG
2012**

**Ficha Catalográfica Elaborada pela Divisão de Processos Técnicos da
Biblioteca da UFLA**

Bortolini, Juliano

Estudo de experimentos fatoriais 2^k aplicados em um processo industrial / Juliano Bortolini. – Lavras : UFLA, 2012.

143 p. : il.

Dissertação (mestrado) – Universidade Federal de Lavras, 2012.

Orientador: Marcelo Silva de Oliveira.

Bibliografia.

1. Experimento industrial. 2. Experimento não repetido. 3. Um-fator-por-vez. 4. Confundimento. I. Universidade Federal de Lavras.
II. Título.

CDD - 519.57

JULIANO BORTOLINI

**ESTUDO DE EXPERIMENTOS FATORIAIS 2^k APLICADOS EM UM
PROCESSO INDUSTRIAL**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de Mestre.

APROVADA em 24 de fevereiro de 2012.

Dr. Augusto Ramalho de Moraes

UFLA

Dr. Tadayuki Yanagi Junior

UFLA

Dr. Marcelo Silva de Oliveira
Orientador

**LAVRAS - MG
2012**

A meus pais Leda e Jesus,
pelo amor, carinho e educação.

A minha irmã Rafaela,
pelo amor, carinho, educação e travessuras.

DEDICO

AGRADECIMENTOS

A Deus, pelo sol que ilumina os meus dias, pelos corpos celestes que clareiam as minhas noites mais escuras, pelas chuvas que lavam a minha alma, pelas primaveras que colorem a minha vida com as suas flores, pelos pássaros que alegam o meu espírito com os seus cantos, pelos alimentos que fortificam o meu corpo, pela verdade que aos poucos me liberta, pelo trabalho que me faz sentir útil, pela vida que me proporciona novas experiências e oportunidades de crescimento, e por tudo que ainda não tenho consciência que fazes por mim.

À minha querida mãe Marileda e ao meu querido pai José Jesus, pelo amor e dedicação na boa educação, em especial por me ensinarem a como deitar e conseguir dormir em paz.

À minha querida irmã Rafaela, pela companhia fraterna, desde sempre, me alegrando e educando.

A todos os meus familiares e amigos, pelo apoio e carinho.

Ao meu orientador Marcelo Silva de Oliveira, pelos conhecimentos e esclarecimentos intelectuais e morais confiados a mim, pela paciência e compreensão das minhas dificuldades e por me aceitar como seu orientando.

Aos professores Augusto Ramalho de Moraes, Renato Ribeiro de Lima e Tadayuki Yanagi Junior pelas importantes contribuições nesta dissertação, por serem receptivos e gentis ao me receberem em seus gabinetes e por participarem da minha qualificação e defesa.

À Universidade Federal de Lavras (UFLA) e ao Departamento de Ciências Exatas (DEX), pela oportunidade de cursar o mestrado

Aos professores do programa de pós-graduação em estatística e experimentação agropecuária da UFLA, em especial aos professores João D. Scalon, Joel Augusto Muniz, Marcelo Silva de Oliveira, Mário Javier Ferrua Vivanco, Paulo

César Lima, Renato Ribeiro de Lima, Solange Gomes Faria Martins e Thelma Sáfadi, pelas contribuições na minha formação durante as suas disciplinas.

Às funcionárias do DEX: Edila, Josiane Cristina, Josiane Oliveira, Kelly, Maria, Miriam e Selma, pela amizade e por manterem o departamento “em pé”.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão da bolsa de estudos, tornando financeiramente possível a realização do mestrado.

Aos colegas de turma: Adriana, Adriele, André, Danielle, Elayne, Iábita, Juracy, Juscelia, Larissa, Lourenço, Marcelo, Mariele, Tábata e Wederson, pela companhia nos estudos e nos momentos de lazer.

A todos os demais amigos da estatística, em especial à Andressa, ao Danilo, Diogo, Enio, Felipe, Jair, Rossicley, Thalita e Walter, pela amizade e pelos momentos de descontração.

Aos meus companheiros de república: André, Gilberto, Guido Gustavo, Romário, Juracy e Moisés, pela companhia.

A todos os meus professores da Universidade Federal de Mato Grosso, em especial ao Aldi, André, Andréia, Demilson, Gladys, Jones, Luzia, Martinho, Provenzano, Ronie, Sérgio e Vinícius.

Aos amigos da Campanha do Quilo e do Grupo Semeadores, por confiarem o trabalho em mim.

A todos que de alguma forma contribuíram, mesmo inconsciente, para a realização deste trabalho.

“Por vezes sentimos que aquilo que fazemos não é senão uma gota de
água no mar. Mas o mar seria menor se lhe faltasse uma gota.”

Madre Teresa de Calcutá

RESUMO

Experimentos são realizados pelas indústrias com o intuito de aumentar a compreensão de seus processos de fabricação. Desta forma, objetiva-se neste trabalho abordar o planejamento e análise de experimentos fatoriais 2^k no contexto industrial. Especificamente, propôs-se expor aspectos e conceitos básicos de experimentação e sugerir, como alternativa para o uso da técnica de experimentação um-fator-por-vez, o planejamento de experimentos fatoriais 2^k completos, que são úteis quando tem-se por finalidade examinar um grande número de fatores para determinar quais são os mais importantes, ou quando o tempo ou o recurso para o experimento são escassos. As técnicas do confundimento e experimento não repetido também são abordadas neste trabalho. A primeira é uma alternativa para experimentos em que há dificuldade em manter constante alguma característica do processo produtivo, e a segunda quando é inviável realizar repetições. Como resultado ilustra-se as técnicas estatísticas apresentadas, sendo essas aplicações simuladas a partir de um mesmo experimento de verificação da influência de quatro fatores na resistência mecânica de junções adesivas.

Palavras-chave: Confundimento. Experimento industrial. Experimento não repetido. Um-fator-por-vez.

ABSTRACT

Experiments are conducted by industry with the aim of increase understanding of their manufacturing processes. Thus, this study aims to approach the planning and analysis of 2^k factorial experiments in an industrial context. Specifically, we proposed to expose some aspects and basic concepts of experimentation and suggest as a possible alternative to using the technique of testing one-factor-at-a-time, planning full factorial experiments at two levels, which are useful when the objective is to examine a large number of factors to determine which may be the most important, or when the time or resource for the experiment are scarce. The techniques of confounding and unreplicated experiment are also discussed in this work. The first is an alternative to experiments in which there is difficulty in maintaining a constant characteristic of the production process, and the second occur when it is impractical conduct replicates. As a result there is illustrated the statistical techniques presented, and these applications were simulated from the same experiment to verify the influence of four factors on the mechanical strength of adhesive joints.

Keywords: Confounding. Industrial experiment. Unreplicated experiment. One-factor-at-a-time.

LISTA DE FIGURAS

Figura 1	Representação de um modelo geral de processo	20
Figura 2	Representação de um modelo de processo de produção de polímero	21
Figura 3	Gráfico de probabilidade normal para os efeitos fatoriais de um experimento fatorial 2^4	82
Figura 4	Corpo de prova obtido através de sobreposição simples de chapas de aço (Adaptado de Santos, 2007)	91
Figura 5	Gráficos de interações dos fatores <i>A</i> (tratamento superficial) e <i>B</i> (temperatura e tempo de cura) em cada nível do fator <i>C</i> (rugosidade), considerando a codificação “-” sem tratamento e “+” com tratamento, para o fator <i>A</i> ; “-” 22°C e 24 horas e “+” 60°C e 2 horas, para o fator <i>B</i> ; “-” sem jateamento e “+” com jateamento, para o fator <i>C</i>	121
Figura 6	Gráficos de interações do fator <i>C</i> (rugosidade) em cada nível dos fatores <i>A</i> (tratamento superficial) e <i>B</i> (temperatura e tempo de cura), considerando a codificação “-” sem tratamento e “+” com tratamento, para o fator <i>A</i> ; “-” 22°C e 24 horas e “+” 60°C e 2 horas, para o fator <i>B</i> ; “-” sem jateamento e “+” com jateamento, para o fator <i>C</i>	121
Figura 7	Gráfico de probabilidade normal para os efeitos fatoriais do experimento fatorial não repetido de resistência mecânica de junções adesivas	124
Figura 8	Gráficos de efeitos principais dos fatores <i>A</i> (tratamento superficial) e <i>B</i> (temperatura e tempo de cura), considerando a codificação “-” sem tratamento e “+” com tratamento, para o fator <i>A</i> ; “-” 22°C e 24 horas e “+” 60°C e 2 horas, para o fator <i>B</i> . . .	129
Figura 9	Gráficos de efeitos principais dos fatores <i>C</i> (rugosidade) e <i>D</i> (ativador), considerando a codificação “-” sem jateamento e “+” com jateamento, para o fator <i>C</i> ; “-” sem ativador e “+” com ativador, para o fator <i>D</i>	130

LISTA DE TABELAS

Tabela 1	Representações dos níveis dos fatores, A e B , e tratamentos de um experimento 2^2	27
Tabela 2	Tratamentos do experimento referente a análise da força de torque em um eixo de rotação.	31
Tabela 3	Ordem de execução dos testes do experimento referente a análise da força de torque em um eixo de rotação.	32
Tabela 4	Esquema da análise de variância para experimentos no delineamento em blocos casualizados.	51
Tabela 5	Esquema da análise de variância para experimentos no delineamento em blocos casualizados em esquema fatorial 2^k , com J blocos (Adaptado de Montgomery (2009)).	54
Tabela 6	Esquema da análise de variância para experimentos no delineamento em blocos casualizados, com I tratamentos e J blocos. . .	69
Tabela 7	Esquema inicial de uma tabela de sinais para um experimento fatorial 2^3 contendo os tratamentos e os efeitos fatoriais a serem avaliados.	73
Tabela 8	Sinais para calcular os efeitos de um experimento fatorial 2^3 contendo os tratamentos, os efeitos fatoriais a serem avaliados e as médias dos tratamentos.	74
Tabela 9	Esquema da análise de variância de um experimento em esquema fatorial 2^3 , com J blocos.	76
Tabela 10	Esquema da análise de variância do desdobramento da interação AC em efeitos do fator A em cada nível do fator C	79
Tabela 11	Esquema da análise de variância do desdobramento da interação AC em efeitos do fator C em cada nível do fator A	80
Tabela 12	Estimativa dos efeitos fatoriais de um experimento fatorial 2^4 não repetido.	81
Tabela 13	Esquema da análise de variância de um experimento 2^4 não repetido, para a verificação de efeitos principais e de interação de ordem dois.	83
Tabela 14	Constituição dos blocos de um experimento 2^3 para o confundimento da interação tripla.	85
Tabela 15	Esquema de análise de variância com e sem confundimento da interação tripla de um experimento 2^3	86
Tabela 16	Tratamentos dos blocos de cada repetição de um experimento fatorial 2^5 para o confundimento dos efeitos ABE , BCE , CDE , AC , $ABCD$, BD e ADE com o efeito de blocos.	89

Tabela 17	Composição química (%) do aço utilizado na confecção dos corpos de prova (Santos, 2007).	92
Tabela 18	Fatores e níveis analisados codificados.	93
Tabela 19	Tratamentos do experimento de resistência mecânica de junções adesivas.	94
Tabela 20	Dados de tensão média de ruptura (<i>MPa</i>) do experimento de resistência mecânica de junções adesivas.	98
Tabela 21	Esquema da análise de variância preliminar dos dados de tensão média de ruptura de junções adesivas (<i>MPa</i>).	104
Tabela 22	Esquema da análise de variância do experimento de resistência mecânica de junções adesivas em esquema fatorial 2^4	108
Tabela 23	Tabela de sinais para o experimento de resistência mecânica de junções adesivas em esquema fatorial 2^4 , contendo os tratamentos, os efeitos fatoriais a serem avaliados e as médias dos tratamentos.	112
Tabela 24	Esquema da análise de variância para o estudo do efeito do fator <i>A</i> em cada nível do fator <i>C</i>	117
Tabela 25	Esquema da análise de variância para o estudo do efeito do fator <i>B</i> em cada nível do fator <i>C</i>	118
Tabela 26	Esquema da análise de variância para o estudo do efeito do fator <i>C</i> em cada nível do fator <i>A</i>	118
Tabela 27	Esquema da análise de variância para o estudo do efeito do fator <i>C</i> em cada nível do fator <i>B</i>	118
Tabela 28	Médias de tensão média de ruptura nos níveis dos fatores <i>A</i> e <i>C</i>	122
Tabela 29	Médias de tensão média de ruptura nos níveis dos fatores <i>B</i> e <i>C</i>	122
Tabela 30	Dados de tensão média de ruptura (<i>MPa</i>) do experimento fatorial 2^4 não repetido.	123
Tabela 31	Esquema da análise de variância do experimento de resistência mecânica de junções adesivas não repetido.	128
Tabela 32	Médias de tensão média de ruptura nos níveis dos fatores <i>A</i> , <i>B</i> , <i>C</i> e <i>D</i>	130
Tabela 33	Tratamentos designados aos blocos de cada repetição do experimento de resistência mecânica de junções adesivas, para o confundimento dos efeitos <i>ACD</i> , <i>BCD</i> e <i>AB</i> com efeito de blocos.	132
Tabela 34	Esquema da análise de variância do experimento de resistência mecânica de junções adesivas em esquema fatorial 2^4 com os efeitos <i>ACD</i> , <i>BCD</i> e <i>AB</i> confundidos com o efeito de blocos.	135
Tabela 35	Observações sobre as três situações experimentais consideradas.	136

SUMÁRIO

1	INTRODUÇÃO	14
2	REFERENCIAL TEÓRICO	17
2.1	Experimentação na indústria	17
2.2	Planejamento e análise de experimentos industriais	19
2.2.1	Detalhamento de delineamentos experimentais	28
2.2.1.1	Delineamento inteiramente casualizado	28
2.2.1.2	Delineamento em blocos casualizados	34
2.2.2	Teorização da análise de variância	37
2.2.2.1	Análise de variância pela abordagem matricial	44
2.2.2.2	Análise de variância pela abordagem escalar	55
2.2.3	Experimento fatorial 2^k não repetido	80
2.2.4	Confundimento	83
3	METODOLOGIA	91
3.1	Materiais	91
3.2	Métodos	94
4	RESULTADOS E DISCUSSÃO	97
4.1	Situação 1: experimento fatorial 2^4 completo	97
4.1.1	Análise de variância pela abordagem matricial	101
4.1.2	Análise de variância pela abordagem escalar	109
4.2	Situação 2: experimento fatorial 2^4 não repetido	123
4.3	Situação 3: experimento fatorial 2^4 com alguns efeitos fatori- ais confundidos com efeito de blocos	131
4.4	Considerações finais	136
4.4.1	Discussão acerca do desempenho das três situações	136
4.4.2	Sugestões para pesquisas futuras	137
5	CONCLUSÃO	138
	REFERÊNCIAS	139

1 INTRODUÇÃO

O interesse e a necessidade das indústrias de desenvolver e melhorar processos de fabricação — tendo como um dos objetivos a melhoria da qualidade de seus produtos —, tem suscitado a utilização de técnicas estatísticas, em especial o planejamento e análise de experimentos, pelas indústrias.

Esse cenário vem ganhando ênfase no ambiente industrial, motivado pela relação competitiva e preocupação em satisfazer o consumidor.

Por outro lado, de forma preocupante, na literatura há uma reunião de evidências e dúvidas, apresentadas em estudos de casos, sobre a integridade dos planejamentos e análises de experimentos conduzidos pelas indústrias. Destaca-se a utilização da estratégia de experimentação um-fator-por-vez, questionável por não ser possível considerar na análise os efeitos das interações entre os fatores.

Ao planejar um experimento industrial é possível deparar-se com algumas de suas peculiaridades, tais como a dificuldade em construir blocos completos, e até mesmo de realizar repetições. Quando essas, e outras, peculiaridades são relevadas, tem-se o questionamento da veracidade dos resultados das análises estatísticas.

O planejamento de um experimento de forma equívoca compromete, sobretudo, a credibilidade e a confiança dos resultados das análises estatísticas. Resultados enganosos podem prejudicar a estabilidade das indústrias, se utilizados como subsídios para tomadas de decisões.

Assim descritos os desafios de planejar experimentos em processos industriais, o objeto de investigação desta dissertação foi limitado ao planejamento e análise de experimentos fatoriais 2^k completos, e as técnicas de confundimento e experimento sem repetição.

Uma vez limitado o objeto de análise, propôs-se o problema para investi-

gação: quais seriam exemplos de análises básicas para experimentação na indústria que elucidam peculiaridades próprias desta área de aplicação?

Para a pesquisa desenvolvida definiu-se como objetivo geral, discutir o planejamento e análise de experimentos no contexto industrial, com um enfoque didático. Para o seu desenvolvimento, foram fixados três objetivos específicos.

O primeiro, propôs-se expor conceitos básicos de experimentação e sugerir, como possível alternativa para o uso da técnica de experimentação um-fator-por-vez, o planejamento de experimentos fatoriais 2^k completos.

Na sequência, propôs-se apresentar alguns aspectos do planejamento e análise de experimentos em esquema fatorial 2^k , nos delineamentos inteiramente casualizados e em blocos casualizados, e as técnicas do confundimento e experimento não repetido.

Por fim, como terceiro objetivo, propôs-se ilustrar as técnicas estatísticas apresentadas em três condições experimentais. Todas as situações experimentais são baseadas em um mesmo experimento de verificação da influência de quatro fatores na resistência mecânica de junções adesivas.

O marco teórico utilizado nesta dissertação é a análise de variância, desenvolvida por Ronald Aylmer Fisher (1890 - 1962), enquanto trabalhava na estação experimental Rothamsted, localizada na Inglaterra. A análise de variância, que tem por ideia básica comparar a variação entre grupos com a variação dentro de grupos, é um dos métodos estatísticos mais disseminados, principalmente na experimentação agrícola.

Em relação aos aspectos metodológicos, faz-se uso da técnica de revisão bibliográfica da literatura sobre estatística experimental aplicada à indústria, a partir de artigos, livros, teses e dissertações. Para obter-se a massa de dados utilizada na exemplificação das técnicas estatísticas mencionadas, utilizou-se busca na lite-

ratura e relações pessoais.

Para o desenvolvimento do tema suscitado, o plano de trabalho foi estruturado em cinco capítulos, objetivando atingir cada um dos objetivos indicados.

Desse modo, o capítulo primeiro destina-se à introdução desta dissertação, no qual é delimitado o assunto de estudo e são especificados os objetivos de pesquisa.

O segundo capítulo dedica-se ao referencial teórico, no qual são expostos alguns aspectos do planejamento e análise de experimentos fatoriais 2^k completos, e algumas técnicas estatísticas que podem ser úteis ao pesquisador na indústria.

Na sequência, apresentando o método utilizado na pesquisa, desenvolve-se o terceiro capítulo.

O quarto capítulo reserva-se ao desenvolvimento da análise estatística de três situações para o experimento de resistência mecânica de junções adesivas. A primeira situação destina-se à análise do experimento fatorial 2^4 completo; a segunda é considerando, hipoteticamente, a dificuldade em realizar o experimento com repetições; a terceira é supondo a impossibilidade de formar blocos que contenham todos os tratamentos.

Por fim, no último capítulo são apresentadas as principais conclusões deste trabalho.

2 REFERENCIAL TEÓRICO

2.1 Experimentação na indústria

A metodologia planejamento de experimentos é uma técnica estatística que tem sido aplicada por muitos anos na indústria no desenvolvimento de novos produtos, controle e melhoria dos processos de fabricação etc.

Conforme Tanco *et al.* (2008), em pesquisa realizada em três regiões europeias, constataram que 95% das empresas realizam experimentos. No entanto, os métodos utilizados nem sempre são adequados.

Uma das explicações possíveis para esse cenário, planejamento inadequado de experimentos, que não há de ser muito diferente no Brasil, segundo Czirom (1999), é a falta de considerações práticas do planejamento de experimentos nos cursos acadêmicos de engenharia.

Outra explicação cabível é a escassa relação entre universidade e empresa. Segundo Rapini (2007), há pouco interesse das empresas em estabelecer relações com as universidades. E quando as cooperações estão presentes, estas limitam-se a atividades de consultoria.

Ainda no âmbito da interação universidade - empresa, Costa e Cunha (2001) esclarecem que existe uma série de barreiras organizacionais, pessoais, profissionais e culturais, ocasionadas basicamente pelas diferenças de interesses por ambas as partes. A meta principal da universidade é a geração de conhecimentos e tecnologia para o desenvolvimento da sociedade em geral; a da empresa é focada na geração de receitas, sem as quais não sobreviveria e não desempenharia funções sociais (por exemplo, a geração de empregos).

Para um planejamento adequado de um experimento industrial, espera-se que as peculiaridades típicas da indústria sejam consideradas. Algumas peculiari-

dades típicas da indústria são:

- (i) dificuldade de interromper o processo produtivo. Nesse caso são sugeridos os experimentos denominados *online*. Em relação a esses experimentos, Box (1957) propôs o *Evolutionary Operation* (EVOP), um método de controle e melhoria de processos a ser usado durante o processo produtivo;
- (ii) complicação, ou até mesmo a impossibilidade, de modificar constantemente algumas características do processo industrial. Nesta situação, sugere-se o uso de parcelas subdivididas. Jones e Nachtsheim (2009) e Goos, Langhans e Vandebroek (2006), abordam esses experimentos, e ressaltam que muitos experimentos em parcelas subdivididas são analisados como se não o fossem;
- (iii) obstáculo em manter constante alguma característica do processo produtivo – por exemplo, matérias primas de um mesmo fornecedor. Nessa situação é recomendado o uso do delineamento em blocos casualizados, abordado nesta dissertação.
- (iv) inviabilidade em realizar repetições por causa, por exemplo, do alto custo envolvido no processo. Para essa inviabilidade aconselha-se o experimento fatorial não repetido, abordado, também, nesta dissertação.

Ilzarbe *et al.* (2008), em revisão de casos publicados em revistas científicas entre os anos 2001 e 2005, constataram que a maioria dos experimentos são provenientes das indústrias de materiais, mecânica e química. Sendo que a maioria deles são referentes a investigar se certas condições, ou configurações, de um processo influenciam em algumas características dos produtos resultantes desse processo.

Nas próximas seções serão abordados alguns aspectos do planejamento e análise de experimentos, e algumas técnicas da estatística experimental que podem ser úteis para algumas particularidades dos experimentos em processos industriais.

2.2 Planejamento e análise de experimentos industriais

Experimentos são realizados constantemente por pesquisadores de diversas áreas com o intuito de aumentar a compreensão sobre um processo. Experimento é definido por Montgomery (2009) como sendo um teste, ou série de testes, em que são feitas mudanças propositalmente em algumas características do processo, de modo a observar e identificar as razões das mudanças que puderam ser observadas em uma variável resposta, que, segundo Mason, Gunst e Hess (2003), é simplesmente uma observação do experimento. Sucintamente, o objetivo da experimentação é obter uma relação de causa e efeito entre a variável resposta, ou saída, e as variáveis de entrada de um processo.

Um processo pode ser esquematizado conforme o modelo da Figura 1, em que as variáveis de entrada controláveis x_1, x_2, \dots, x_p , segundo Montgomery (2009), podem ser visualizadas como uma combinação de máquinas, métodos, pessoas e outros recursos que transformam entrada (geralmente materiais) em saída, que tem uma ou mais variáveis respostas. As variáveis de entrada não controláveis z_1, z_2, \dots, z_q são características experimentais não controladas, seja por desconhecimento da existência delas ou pelo alto custo para controlá-las, tais variáveis constituem o erro experimental.

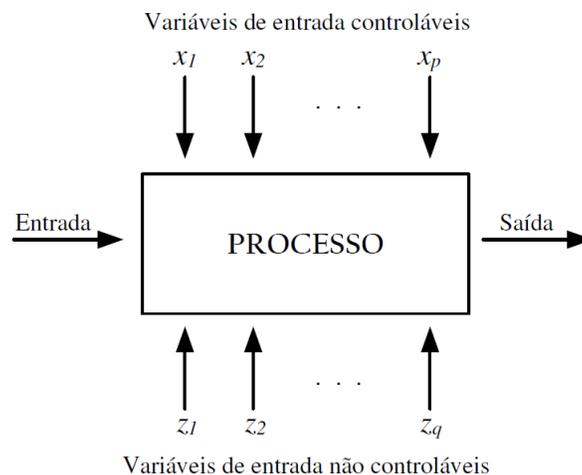


Figura 1 Representação de um modelo geral de processo

É importante notar que uma variável de entrada é considerada controlável se os valores que ela assumir, denominados níveis, podem ser determinados antes do início dos testes. As variáveis de entrada controláveis que são de interesse em serem investigadas pelo pesquisador, geralmente, são denominadas de fatores, e as combinações possíveis entre os níveis dos fatores são chamadas de tratamento.

Como exemplo de experimento, em um processo químico para produzir um polímero, o engenheiro químico tem interesse em estudar o efeito da temperatura do processo e a concentração de um catalisador específico. A temperatura pode ser controlada em 100°C ou 120°C, e a concentração do catalisador em 4% ou 8%. Ao final de cada processo é observada a viscosidade do polímero medido em $Pa \cdot s$ (pascal-segundo). Nesse experimento, há dois fatores: temperatura do processo e concentração do catalisador. As temperaturas 100°C e 120°C são os níveis do fator temperatura; e as concentrações 4% e 8% são os níveis do fator concentração do catalisador. As combinações possíveis entre os níveis dos fato-

res são 100°C com 4%, 100°C com 8%, 120°C com 4% e, por fim, 120°C com 8%, que são os tratamentos possíveis. Nesse experimento, provavelmente, há outras variáveis de entrada envolvidas no processo. Por exemplo, o método utilizado para produzir o polímero pode ser uma variável de entrada controlável. Outras variáveis, como a temperatura ambiente, umidade relativa do ar, teor de oxigênio no ar, etc, podem ser algumas variáveis de entrada não controláveis. O modelo para esse experimento está representado na Figura 2.

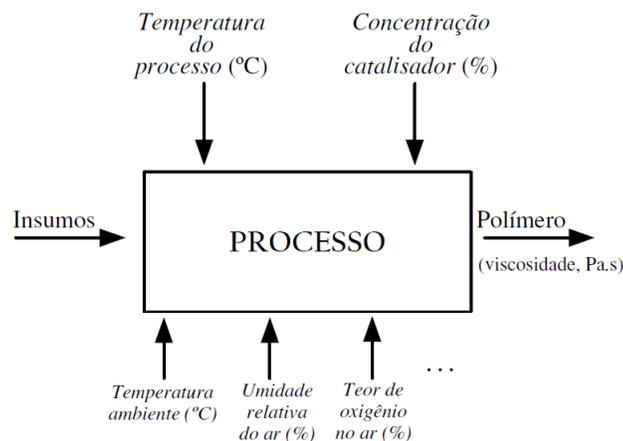


Figura 2 Representação de um modelo de processo de produção de polímero

É importante notar que nem todos os fatores afetam o desempenho da mesma maneira. Alguns podem ter fortes influências, outros podem nem ter efeito na variável resposta (ANTONY, 2003). De tal forma que, no exemplo anterior, pode ser que não haja diferença na viscosidade do polímero entre os tratamentos 100°C com 4% e 120°C com 4%.

Antony (2003), menciona que, em processos de fabricação, é comum o interesse em conduzir experimentos para explorar relações entre variáveis de entrada

e de saída de um processo. No entanto, alerta que, para um experimento ser válido, é necessário valer-se de abordagem científica, a fim de planejá-lo e analisá-lo. Na estatística experimental, o planejamento de experimentos estatísticos refere-se ao processo de realizar um experimento, de tal forma que, os dados obtidos possam ser analisados por meio de métodos estatísticos, e levem a conclusões válidas (MONTGOMERY, 2009).

Conforme Cochran (1947), a análise de variância, desenvolvida por Ronald Aylmer Fisher (1890 - 1962) e comumente utilizada na estatística experimental, depende dos seguintes pressupostos: (i) que os efeitos das variáveis de entrada, controláveis e não controláveis, sejam aditivos; (ii) que os erros experimentais sejam independentes, sejam (iii) de variância igual e, além disso, sejam (iv) normalmente distribuídos.

Para assegurar que as análises estatísticas sejam válidas, há três princípios básicos na experimentação. São eles: realização de repetição, aleatorização e controle local (FISHER, 1971).

De acordo com Montgomery (2009), Storck *et al.* (2006) e Mason, Gunst e Hess (2003), repetição é duas ou mais observações obtidas para determinada combinação específica de níveis dos fatores, os quais são conduzidos, tanto quanto possível, sob idênticas condições experimentais. É importante notar que repetição não é duas ou mais mensurações da mesma observação.

Conforme Montgomery (2009) e Hinkelmann e Kempthorne (2008), o princípio da repetição tem por finalidade propiciar a obtenção de estimativa do erro experimental.

O princípio da aleatorização, segundo Banzatto e Kronka (2008) e Hinkelmann e Kempthorne (2008), consiste em atribuir a todos os tratamentos a mesma probabilidade de serem designados a qualquer unidade experimental, que é a en-

tidade na qual uma observação é feita, e tem por finalidade proporcionar uma estimativa válida para o erro experimental. O modo como é conduzida a aleatorização dos tratamentos nas unidades experimentais estabelecem os delineamentos experimentais (STORCK *et al.*, 2006). Segundo Cochran (1947), muitos problemas de correlação entre os erros experimentais são resolvidos pela aleatorização adequada.

O ideal em um experimento é que todas as unidades experimentais sejam homogêneas, isto é, que não haja diferença sistemática entre elas. Devem ser tão iguais quanto possível em todas as características que possam afetar a resposta (MASON; GUNST; HESS, 2003). No entanto, nem sempre será possível obter unidades experimentais homogêneas. Em tais situações, faz-se uso do terceiro princípio da experimentação, o controle local.

O princípio controle local é frequentemente utilizado, pois, tem por finalidade reduzir o erro experimental e tornar o delineamento mais eficiente (BANZATTO; KRONKA, 2008; HINKELMANN; KEMPTHORNE, 2008). Sua ideia básica consiste em dividir o número total de unidades experimentais em dois ou mais grupos, de maneira que as unidades de cada grupo sejam homogêneas entre si (HINKELMANN; KEMPTHORNE, 2008; GIESBRECHT; GUMPERTZ, 2004; MASON; GUNST; HESS, 2003; COX; REID, 2000). O termo bloco geralmente é usado para se referir a um conjunto de unidades experimentais homogêneas.

Quando os princípios supracitados não são incorporados ao projeto de um experimento, os resultados das análises estatísticas podem ser inconclusivos, ou, o que é pior, enganosos (MASON; GUNST; HESS, 2003).

Em relação aos experimentos conduzidos em indústrias, Galdamez e Carpinetti (2004) esclarecem que o seu intuito é melhorar o desempenho dos produtos e os processos de fabricação. A melhoria de processos pode ser obtida conforme

Montgomery (2004), por:

- (i) determinar quais são as variáveis mais influentes em um processo, por exemplo para planejar produtos novos;
- (ii) determinar o valor a ser atribuído às variáveis influentes de modo que a variável resposta esteja próxima da exigência nominal;
- (iii) determinar o valor a ser atribuído às variáveis influentes de modo que a variabilidade da variável resposta seja pequena;
- (iv) determinar o valor a ser atribuído às variáveis influentes de modo que os efeitos das variáveis não-controláveis sejam minimizados.

Em experimentos industriais, geralmente, o interesse maior é estudar o efeito de mais de um fator, pois são diversas as variáveis que podem influenciar o processo. Segundo Tanco *et al.* (2008) e Czitrom (1999), a estratégia de experimentação um-fator-por-vez é utilizada pelas indústrias quando deseja-se verificar a influência de diversos fatores em um processo. Detalhes do desenvolvimento dessa estratégia para três e quatro fatores, todos de dois níveis, são apresentadas por Daniel (1994).

A estratégia de experimentação um-fator-por-vez consiste em modificar os níveis de um fator específico a cada teste, enquanto os demais permanecem fixos (MONTGOMERY, 2009; WU; HAMADA, 2009; CZITROM, 1999; DANIEL, 1994). Uma crítica a essa estratégia é que ela não estima satisfatoriamente os efeitos fatoriais, principalmente o das interações (SUDARSANAM; FREY, 2011; BOX; HUNTER; HUNTER, 2005; MONTGOMERY, 2004). Segundo Montgomery (2009) e Mason, Gunst e Hess (2003), o efeito de interação ocorre quando o efeito de um fator sobre a variável resposta depende dos níveis de outros fatores.

Alternativamente ao uso da estratégia um-fator-por-vez no estudo da influência de dois ou mais fatores na variável resposta, indica-se o planejamento de experimentos fatoriais (BOX; HUNTER; HUNTER, 2005, MONTGOMERY, 2004, CZITROM, 1999). Montgomery (2009) explica que por experimento fatorial entende-se que em cada repetição do experimento são investigados todas as combinações possíveis dos níveis dos fatores.

Segundo Wu e Hamada (2009) e Czitrom (1999), planejar um experimento fatorial, ao invés de um-fator-por-vez, é a maneira mais eficaz de determinar a influência de dois ou mais fatores sobre a variável resposta, porque:

- (i) requer menos recursos (experimentos, tempo, material) para a quantidade de informação obtida;
- (ii) as estimativas dos efeitos fatoriais são mais precisas. Usando mais observações para estimar um efeito resulta em maior precisão;
- (iii) o efeito das interações entre os fatores podem ser estimadas sistematicamente;
- (iv) obtém-se informações para uma região experimental maior.

De acordo com Wu e Hamada (2009), em experimentos fatoriais obtém-se informações para uma região experimental maior porque cada efeito fatorial é calculado sobre todas as combinações possíveis dos níveis dos outros fatores, o que não necessariamente ocorre com a estratégia um-fator-por-vez.

Giesbrecht e Gumpertz (2004) ratifica a opinião de Czitrom (1999) afirmando que a grande vantagem de experimentos fatoriais é que eles permitem estudar uma série de fatores simultaneamente e, em especial, o efeito de interação.

Alguns tipos especiais de planejamentos fatoriais são muito úteis no desenvolvimento e melhoria de processos. Um deles é o planejamento com k fatores

($k \in \mathbb{N}$), cada um com dois níveis, que podem ser quantitativos - como valores de temperatura, pressão ou tempo - ou qualitativos - tipo de máquina ou operador (MONTGOMERY, 2009). Experimentos com k fatores e todos com dois níveis são representados por 2^k .

Geralmente, os níveis de cada fator em um experimento 2^k são chamados de “baixo” e “alto”, ou “ausente” e “presente”, e podem ser representados por “-” e “+”, ou “0” e “1”, ou “-1” e “+1”, respectivamente (MONTGOMERY, 2009). Outra forma de representar os tratamentos é usando letras minúsculas. Se uma letra está presente, então o fator correspondente é colocado no seu nível alto nesse tratamento; se a letra está ausente, o fator é colocado em seu nível baixo. O tratamento com todos os fatores no nível baixo é representado por “(1)” (MONTGOMERY, 2009). Para o exemplo do processo de produção de um polímero, que é um experimento fatorial 2^2 , os níveis dos fatores temperatura e concentração podem ser codificados em “-” para os níveis baixo, que são 100°C e 4%, e “+” para os níveis altos, que são 120°C e 8%. Na Tabela 1 há uma síntese dessas representações para um experimento 2^2 .

Tabela 1 Representações dos níveis dos fatores, A e B , e tratamentos de um experimento 2^2 .

A	B	A	B	A	B
baixo	baixo	-	-	0	0
alto	baixo	+	-	1	0
baixo	alto	-	+	0	1
alto	alto	+	+	1	1

A	B	A	B	
ausente	ausente	-1	-1	(1)
presente	ausente	+1	-1	a
ausente	presente	-1	+1	b
presente	presente	+1	+1	ab

Os planejamentos fatoriais 2^k , conforme Box, Hunter e Hunter (2005), possuem algumas vantagens, entre elas a capacidade de proporcionar uma análise simplificada e, também, serem a base de muitos outros planejamentos úteis, como os fatoriais fracionados 2^{k-p} . Segundo Giesbrecht e Gumpertz (2004), os experimentos fatoriais fracionados são úteis quando o objetivo é examinar um grande número de fatores para determinar quais podem ser os mais importantes, ou quando o tempo ou recurso disponível para o experimento são escassos.

Nos experimentos fatoriais fracionados, alguns efeitos são confundidos entre si, por isso o seu uso é aconselhável quando acredita-se que certos efeitos podem ser desprezíveis, geralmente de interações de ordem superior (de terceira ordem e superiores). O termo confundido é utilizado quando um ou mais efeitos não podem ser atribuídos de forma inequívoca para um único fator, interação ou bloco. Assim, é possível confundir efeitos entre fatores, como é o caso dos ex-

perimentos fatoriais fracionados e, também, confundir efeitos de fatores com os de blocos, que é o caso de confundimento em blocos. Os primeiros não serão abordados nesta dissertação.

Nas subseções seguintes serão apresentados os delineamentos inteiramente casualizados e em blocos casualizados (ambos em esquema fatorial 2^k), os resultados teóricos importantes para o desenvolvimento da análise de variância e algumas técnicas estatísticas que podem ser úteis na experimentação na indústria.

2.2.1 Detalhamento de delineamentos experimentais

Os modelos experimentais comumente utilizados, no contexto desta dissertação, são os de delineamentos inteiramente casualizados e em blocos casualizados, ambos em esquema fatorial 2^k .

2.2.1.1 Delineamento inteiramente casualizado

O delineamento inteiramente casualizado é o mais simples para implementar e de fácil análise estatística. Nesse delineamento, a atribuição dos tratamentos às unidades experimentais, ou posição em uma sequência de testes, é feita de forma completa, sem nenhuma restrição na aleatorização, de maneira inteiramente casualizada.

O delineamento em destaque é apropriado em situações experimentais, tais como: quando ou toda a matéria prima utilizada em um experimento é proveniente de um mesmo lote, ou há apenas um operador responsável pelo experimento, ou é utilizada somente uma máquina durante os testes. Em suma, o uso desse delineamento é indicado quando todas as unidades experimentais que serão utilizadas no experimento podem ser consideradas homogêneas.

Muitos experimentos de processos industriais, por exemplo os de quimiometria¹, são conduzidos em laboratórios sob condições ambientais controladas. Assim, com a garantia de que não haja efeitos perturbadores, ou não controláveis, o uso do delineamento inteiramente casualizado é aconselhado.

Box; Hunter; Hunter (2005) apresentam um exemplo do delineamento inteiramente casualizado, em que é investigado o quanto a cor esbranquiçada de uma cera de piso é afetada quando certas mudanças são introduzidas na fórmula de preparo. São analisados três fatores: as quantidades dos emulsificantes *A* e *B*, e a concentração do catalisador *C*, sendo que todos os três podem ser classificados em *alto* ou *baixo*, independentes um do outro. Ou seja, o experimento é em esquema fatorial $2 \times 2 \times 2 = 2^3$. No total, são oito combinações diferentes avaliadas, as quais foram aplicadas, em pequena quantidade, em lâminas de vidro, a fim de analisar sua clareza.

Outro exemplo de experimento fatorial conduzido no delineamento inteiramente casualizado é apresentado por Barros Neto; Scarminio; Bruns (2010), em que o experimento se refere ao desenvolvimento de um detergente em pó. Nesse experimento são analisados três ingredientes, todos em dois níveis, sobre o poder de branqueamento e a redeposição da formulação detergente resultante. Cada tratamento foi repetido doze vezes utilizando a mesma quantidade de detergente em pó. As lavagens foram conduzidas em simuladores industriais de lavagem de roupas, e as unidades experimentais homogêneas são peças de tecidos que foram submetidas à mesma solução-padrão, que imita a sujeira doméstica típica. A variável resposta analisada é a densidade óptica que mede a intensidade da luz refletida pela roupa lavada. Os ingredientes e as suas concentrações foram omitidos na

¹Segundo a *International Chemometrics Society* (ICS), quimiometria é a ciência relacionada a medidas realizadas em um sistema ou processo químico, obtendo informações sobre o estado do sistema através da aplicação de métodos matemáticos ou estatísticos.

publicação para garantir o sigilo industrial, já que se trata de uma aplicação real.

A construção de um experimento fatorial no delineamento inteiramente casualizado pode ser feita por meio do procedimento abaixo, adaptado de Mason; Gunst; Hess (2003):

1. enumerar todas as combinações dos níveis dos fatores - tratamentos -, incluindo repetição, sequencialmente de 1 a N ;
2. obter uma sequência de números inteiros aleatórios de 1 a N , sendo que o sorteio dos números aleatórios aconteça sem reposição;
3. atribuir os tratamentos nas unidades experimentais, ou posição na sequência de testes, conforme a ordem especificada pela sequência de números aleatórios. Caso o experimento necessite de unidades experimentais e sequência de testes, utilizar duas sequências de números aleatórios.

O procedimento acima será esclarecido com o exemplo seguinte, adaptado de Mason; Gunst; Hess (2003). Um experimento é conduzido em laboratório com o intuito de analisar a força de torque em um eixo de rotação encontrado em máquinas industriais. Um eixo de rotação será apoiado por uma luva cilíndrica estacionária, e lubrificantes serão aplicados à parede interna da luva com a intenção de reduzir o atrito entre o eixo e a luva. O objetivo desse experimento é estudar o material com que os eixos são feitos (aço ou ligas de alumínio), dois tipos de metal do qual as luvas são feitas (poroso e não poroso), e dois tipos de lubrificantes ($L1$ e $L2$). As unidades experimentais desse experimento são versões pequenas do eixo de rotação, produzidas conforme os materiais de interesse no estudo e em dobro, para ser possível realizar duas repetições. O experimento é planejado no delineamento inteiramente casualizado em esquema fatorial 2^3 com duas repetições. A Tabela 2 lista os oito tratamentos e os dezesseis testes possíveis.

Tabela 2 Tratamentos do experimento referente a análise da força de torque em um eixo de rotação.

Teste	Tratamento	Mat. eixo	Mat. luva	Lub.	Rep.
1	1	aço	poroso	L1	1
2	1	aço	poroso	L1	2
3	2	aço	poroso	L2	1
4	2	aço	poroso	L2	2
5	3	aço	não poroso	L1	1
6	3	aço	não poroso	L1	2
7	4	aço	não poroso	L2	1
8	4	aço	não poroso	L2	2
9	5	alumínio	poroso	L1	1
10	5	alumínio	poroso	L1	2
11	6	alumínio	poroso	L2	1
12	6	alumínio	poroso	L2	2
13	7	alumínio	não poroso	L1	1
14	7	alumínio	não poroso	L1	2
15	8	alumínio	não poroso	L2	1
16	8	alumínio	não poroso	L2	2

Através de um *software* estatístico obteve-se a sequência de números aleatórios referentes ao teste

5 12 4 14 6 7 11 15 13 1 3 9 10 8 2 16

Então, a sequência dos testes a serem executados será conforme a Tabela 3.

Tabela 3 Ordem de execução dos testes do experimento referente a análise da força de torque em um eixo de rotação.

#	Teste	Tratamento	Mat. eixo	Mat. luva	Lub.	Rep.
1	5	3	aço	não poroso	L1	1
2	12	6	alumínio	poroso	L2	2
3	4	2	aço	poroso	L2	2
4	14	7	alumínio	não poroso	L1	2
5	6	3	aço	não poroso	L1	2
6	7	4	aço	não poroso	L2	1
7	11	6	alumínio	poroso	L2	1
8	15	8	alumínio	não poroso	L2	1
9	13	7	alumínio	não poroso	L1	1
10	1	1	aço	poroso	L1	1
11	3	2	aço	poroso	L2	1
12	9	5	alumínio	poroso	L1	1
13	10	5	alumínio	poroso	L1	2
14	8	4	aço	não poroso	L2	2
15	2	1	aço	poroso	L1	2
16	16	8	alumínio	não poroso	L2	2

O modelo estatístico do delineamento inteiramente casualizado com I tratamentos e J repetições, conforme Montgomery (2009) e Wu e Hamada (2009), é dado por

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \quad (2.1)$$

em que,

$i = 1, \dots, I$ e $j = 1, \dots, J$,

Y_{ij} é a j -ésima observação do i -ésimo tratamento,

μ é uma constante comum a todas as observações,

τ_i é o efeito do i -ésimo tratamento,

ε_{ij} é o erro experimental da j -ésima observação do i -ésimo tratamento, com $\varepsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma^2)$.

Para o experimento em esquema fatorial 2^k , o efeito de tratamentos τ_i é decomposto em efeitos principais e de interações, isto é,

$$Y_{i_1 \dots i_k j} = \mu + \underbrace{\alpha_{i_1} + \dots + \kappa_{i_k} + (\alpha\beta)_{i_1 i_2} + \dots + (\alpha \dots \kappa)_{i_1 \dots i_k}}_{\tau_{i'}} + \varepsilon_{i_1 \dots i_k j}, \quad (2.2)$$

ou conforme o modelo (2.1),

$$Y_{i'j} = \mu + \tau_{i'} + \varepsilon_{i'j}, \quad (2.3)$$

em que,

$i_l = 1, 2$ para todo $l = 1, \dots, k$, e $j = 1, \dots, J$;

$Y_{i_1 \dots i_k j}$ é a j -ésima observação do i_1 -ésimo nível do fator A , i_2 -ésimo nível do fator B , \dots e i_k -ésimo nível do fator K ;

μ é uma constante comum a todas as observações;

α_{i_1} é o efeito do i_1 -ésimo nível do fator A ;

\dots ;

κ_{i_k} é o efeito do i_k -ésimo nível do fator K ;

$(\alpha\beta)_{i_1 i_2}$ é o efeito da interação do i_1 -ésimo nível do fator A com o i_2 -ésimo nível do fator B ;

\dots ;

$(\alpha \dots \kappa)_{i_1 \dots i_k}$ é o efeito da interação do i_1 -ésimo nível do fator A com o i_2 -ésimo

nível do fator B , \dots e com o i_k -ésimo nível do fator K ;

$\varepsilon_{i_1 \dots i_k j}$ é o erro experimental da j -ésima observação do i_1 -ésimo nível do fator A , i_2 -ésimo nível do fator B , \dots e i_k -ésimo nível do fator K , com

$$\varepsilon_{i_1 \dots i_k j} \stackrel{iid}{\sim} N(0, \sigma^2).$$

Observe que $\alpha_{i_1} + \dots + \kappa_{i_k} + (\alpha\beta)_{i_1 i_2} + \dots + (\alpha \dots \kappa)_{i_1 \dots i_k}$ é a decomposição dos efeitos de tratamentos, representado por $\tau_{i'}$.

2.2.1.2 Delineamento em blocos casualizados

Quando sabe-se, ou suspeita-se, que existam fontes de variações indesejáveis, fatores que não estão controlados e que influenciarão na variável resposta, o uso do delineamento inteiramente casualizado não é recomendado.

As fontes indesejáveis de variações podem ser causadas de lote para lote, quando o lote não contém unidades experimentais suficientes para conduzir todo o experimento; dia para dia (ou turno para turno), quando não é possível conduzir o experimento em apenas um dia (ou turno); e máquina para máquina (ou operador para operador), quando é necessário utilizar mais do que um tipo de máquina (ou operador) para realizar o experimento.

É possível reduzir ou eliminar as fontes de variações não desejáveis através do que é chamado blocos (BOX; HUNTER; HUNTER, 2005). Pois, conforme Daniel (1976), o uso de blocos permite isolar unidades experimentais que são mais homogêneas do que o todo, sendo essa a sua principal vantagem. Cox e Reid (2000) esclarecem que a ideia central do delineamento em blocos casualizados, além de agrupar as unidades experimentais homogêneas, é realizar comparações entre unidades experimentais semelhantes, de tal forma que todas as unidades experimentais de qualquer bloco são susceptíveis de darem respostas semelhantes na ausência dos efeitos de tratamentos.

No delineamento em destaque, a ordem em que deve ocorrer os tratamentos dentro de cada bloco e a ordem dos blocos são aleatorizadas independentemente. Devido à aleatorização dos tratamentos ocorrer dentro dos blocos, diz-se que houve uma restrição sobre a aleatorização (MONTGOMERY, 2009).

Box, Hunter e Hunter (2005) ilustram o delineamento em blocos casualizados através de um experimento em que são investigados quatro variações no processo de produção de penicilina. As variações do processo são os tratamentos do experimento, e o rendimento da produção é a variável resposta considerada. Infelizmente, as propriedades das matérias primas disponíveis variam consideravelmente e acredita-se que isso, por si só, pode causar diferenças consideráveis na produtividade. Em decorrência desse fato, o uso do delineamento inteiramente casualizado não é aconselhável, e sim a formação de blocos compostos por matérias primas de mesma propriedade. Nesse experimento, é aleatorizada a ordem das matérias primas que serão utilizadas e, posteriormente, dentro de cada matéria prima, a ordem dos tratamentos.

O modelo estatístico para um experimento com I tratamentos e J blocos, conforme Montgomery (2009) e Wu e Hamada (2009), é dado por

$$Y_{ij} = \mu + \tau_i + b_j + \varepsilon_{ij}, \quad (2.4)$$

em que,

$i = 1, \dots, I$ e $j = 1, \dots, J$,

Y_{ij} é a observação Y do j -ésimo bloco e i -ésimo tratamento,

μ é uma constante comum a todas as observações,

τ_i é o efeito do i -ésimo tratamento,

b_j é o efeito do j -ésimo bloco,

ε_{ij} é o erro experimental da observação Y do j -ésimo bloco e i -ésimo tratamento,

com $\varepsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma^2)$.

O delineamento em blocos casualizados também pode ser em esquema fatorial, seguindo o mesmo raciocínio do modelo (2.1) para o modelo (2.2); tem-se que

$$Y_{i_1 \dots i_k j} = \mu + \alpha_{i_1} + \dots + \kappa_{i_k} + (\alpha\beta)_{i_1 i_2} + \dots + (\alpha \dots \kappa)_{i_1 \dots i_k} + b_j + \varepsilon_{i_1 \dots i_k j}, \quad (2.5)$$

em que,

$i_l = 1, 2$ para todo $l = 1, \dots, k$, e $j = 1, \dots, J$;

$Y_{i_1 \dots i_k j}$ é a observação do i_1 -ésimo nível do fator A , i_2 -ésimo nível do fator B , \dots , i_k -ésimo nível do fator K , no j -ésimo bloco;

μ é uma constante comum a todas as observações;

α_{i_1} é o efeito do i_1 -ésimo nível do fator A ;

\dots ;

κ_{i_k} é o efeito do i_k -ésimo nível do fator K ;

$(\alpha\beta)_{i_1 i_2}$ é o efeito da interação do i_1 -ésimo nível do fator A com o i_2 -ésimo nível do fator B ;

\dots ;

$(\alpha \dots \kappa)_{i_1 \dots i_k}$ é o efeito da interação do i_1 -ésimo nível do fator A com o i_2 -ésimo nível do fator B , \dots e com o i_k -ésimo nível do fator K ;

b_j é o efeito do j -ésimo bloco;

$\varepsilon_{i_1 \dots i_k j}$ é o erro experimental da j -ésima observação do i_1 -ésimo nível do fator A , i_2 -ésimo nível do fator B , \dots e i_k -ésimo nível do fator K , com

$\varepsilon_{i_1 \dots i_k j} \stackrel{iid}{\sim} N(0, \sigma^2)$.

2.2.2 Teorização da análise de variância

Nesta subseção serão abordados conceitos importantes para o desenvolvimento matricial e escalar da análise de variância, que, segundo Iversen (2011), é uma coleção de estatísticas usadas para analisar dados obtidos a partir de experimentos.

Conforme Casella e Berger (2010) e Montgomery (2009), a ideia básica da análise de variância é o particionamento da variabilidade total em partes apropriadas, de tal forma que seja possível testar hipóteses.

Na análise de variância, assume-se que os dados, Y_{ij} , são observados de acordo com um modelo linear de efeitos fixo, aleatório ou misto, e que os erros aleatórios são independentes, normalmente distribuídos e homocedásticos.

Um modelo linear pode ser definido, de acordo com Clarke (2008), da seguinte forma

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2.6)$$

sendo \mathbf{Y} um vetor de observações realizáveis, de dimensões $n \times 1$;

\mathbf{X} uma matriz conhecida composta de zeros e uns, ou matriz de planejamento, de dimensões $n \times p$ e posto $\mathcal{P}(\mathbf{X}) = k \leq \min\{n, p\}$;

$\boldsymbol{\beta}$ um vetor de parâmetros desconhecidos, de dimensões $p \times 1$;

$\boldsymbol{\varepsilon}$ um vetor de valores desconhecidos, denominado vetor de resíduos ou de erros, de dimensões $n \times 1$, com a pressuposição básica de que cada componente possui valor esperado zero.

A matriz \mathbf{X} , de dimensões $n \times p$, é denominada e definida por Graybill (1976) de matriz de delineamento se, e somente se, \mathbf{X} puder ser particionado em $\mathbf{X} = \left(\mathbf{X}_1, \dots, \mathbf{X}_q \right)$, em que \mathbf{X}_i é uma submatriz, de dimensões $n \times q_i$, e

satisfaz as seguintes condições:

- (i) os elementos da matriz \mathbf{X}_i , $i = 1, \dots, q$, são apenas os números 0 e 1;
- (ii) para cada matriz \mathbf{X}_i , $i = 1, \dots, q$, cada linha contém exatamente um elemento igual a 1 (os elementos restantes em cada linha são iguais a 0);
- (iii) para cada \mathbf{X}_i , $i = 1, \dots, q$, toda coluna contém no mínimo um elemento igual a 1.

Considerando em (2.6) $\varepsilon \sim (\mathbf{0}, \mathbf{V}\sigma^2)$ em que \mathbf{V} é a matriz de covariâncias, de dimensões $n \times n$, então $\mathbf{Y} = \mathbf{X}\beta + \varepsilon$ é definido como modelo linear de Gauss-Markov. E mais, ao associar a ε a distribuição normal, passa-se a ter um modelo linear de Gauss-Markov normal.

Quando \mathbf{V} é igual a uma matriz identidade \mathbf{I} , ou a uma matriz diagonal \mathbf{D} , ou a uma matriz positiva definida $\mathbf{\Omega}$, todas de dimensões $n \times n$, tem-se, respectivamente, o modelo linear de Gauss-Markov normal ordinário, ponderado e generalizado.

Segundo Gentle (2007), uma matriz simétrica $\mathbf{\Omega}$, de dimensões $n \times n$, é positiva definida se

$$\mathbf{x}'\mathbf{\Omega}\mathbf{x} > 0,$$

para qualquer vetor real não nulo $\mathbf{x} = (x_1, \dots, x_n)$.

Nesta dissertação, todos os modelos considerados serão lineares de Gauss-Markov normal ordinário de efeitos fixo, ou seja,

$$\mathbf{Y} = \mathbf{X}\beta + \varepsilon, \tag{2.7}$$

com $\varepsilon \sim N(\mathbf{0}, \mathbf{I}\sigma^2)$.

Depois de definir o modelo, há interesse em estimar os seus parâmetros, que nesta abordagem é o vetor β . Assim, considerando o modelo linear de Gauss-Markov ordinário, de (2.7), um estimador para o vetor de parâmetros, β , é obtido pelo método dos quadrados mínimos, que consiste em procurar $\tilde{\beta}$ que minimiza

$$\tilde{\epsilon}'\tilde{\epsilon} = (\mathbf{Y} - \mathbf{X}\tilde{\beta})'(\mathbf{Y} - \mathbf{X}\tilde{\beta}) = f(\tilde{\beta}). \quad (2.8)$$

Expandido $\tilde{\epsilon}'\tilde{\epsilon}$ obtem-se $\tilde{\epsilon}'\tilde{\epsilon} = \mathbf{Y}'\mathbf{Y} - 2\tilde{\beta}'\mathbf{X}'\mathbf{Y}' + \tilde{\beta}'\mathbf{X}'\mathbf{X}\tilde{\beta}$, que diferenciando em relação a $\tilde{\beta}$ e igualando a 0, produz a equação normal

$$\mathbf{X}'\mathbf{X}\hat{\beta} = \mathbf{X}'\mathbf{Y} \quad (2.9)$$

que é consistente, ou seja, tem no mínimo um vetor solução (RENCHEER; SCHALJE, 2008; GRAYBILL, 1976; SEARLE, 1971).

Para os modelos de delineamentos, que são os abordados nesta dissertação, o posto de \mathbf{X} é igual a $k < \min\{n, p\}$, isto é, \mathbf{X} é uma matriz de posto coluna incompleto, logo, $\mathbf{X}'\mathbf{X}$ não tem inversa simples e $\mathbf{X}'\mathbf{X}\hat{\beta} = \mathbf{X}'\mathbf{Y}$ não possui solução única. Portanto, para obter uma solução do sistema (2.9) é necessário o conceito de inversa generalizada e do teorema apresentado abaixo.

Segundo Graybill (1976) e Searle (1971), uma matriz \mathbf{G} tal que $\mathbf{AGA} = \mathbf{A}$ é chamada de inversa generalizada de \mathbf{A} , e denotada por \mathbf{A}^- .

Teorema 1. *Se o sistema de equações $\mathbf{Ax} = \mathbf{c}$ é consistente e \mathbf{A}^- é qualquer inversa generalizada de \mathbf{A} , então $\mathbf{x} = \mathbf{A}^-\mathbf{c}$ é uma solução (RENCHEER; SCHALJE, 2008).*

Demonstração:

A partir de $\mathbf{AA}^-\mathbf{A} = \mathbf{A}$, tem-se que $\mathbf{AA}^-\mathbf{Ax} = \mathbf{Ax}$. Substituindo $\mathbf{Ax} = \mathbf{c}$ em ambos os lados, obtém-se $\mathbf{AA}^-\mathbf{c} = \mathbf{c}$, que pode ser escrito como $\mathbf{A}(\mathbf{A}^-\mathbf{c}) = \mathbf{c}$.

Portanto, $\mathbf{x} = \mathbf{A}^{-}\mathbf{c}$ é uma solução para $\mathbf{Ax} = \mathbf{c}$. ■

Como resultado imediato do teorema anterior, tem-se que uma solução para $\hat{\beta}$ em (2.9) é dado por

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y} \quad (2.10)$$

em que $(\mathbf{X}'\mathbf{X})^{-}$ é qualquer inversa generalizada de $(\mathbf{X}'\mathbf{X})$. Para uma particular inversa generalizada $(\mathbf{X}'\mathbf{X})^{-}$, o valor esperado de $\hat{\beta}$ é

$$E(\hat{\beta}) = (\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'E(\mathbf{Y}) = (\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{X}\beta.$$

Note-se que $\hat{\beta}$ é um estimador não viesado de $(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{X}\beta$, mas não de β , pois $(\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{X} \neq \mathbf{I}$. Assim, $\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-}\mathbf{X}'\mathbf{Y}$ não é invariante para escolhas diferentes de $(\mathbf{X}'\mathbf{X})^{-}$. Em outras palavras, o valor esperado de $\hat{\beta}$ é diferente para cada escolha de $(\mathbf{X}'\mathbf{X})^{-}$. O que é feito, na prática, é escolher algumas funções lineares apropriadas de β e estimá-las considerando uma inversa generalizada. Essas funções previamente escolhidas possuem a propriedade de serem invariantes para qualquer escolha da inversa generalizada.

Outro conceito importante para a análise de variância, em especial para a decomposição da variabilidade total dos dados, é a de formas quadráticas. Segundo Rencher e Schaalje (2008), se \mathbf{A} é uma matriz simétrica, de dimensões $n \times n$, isto é, \mathbf{A} é igual a sua matriz transposta, \mathbf{A}' , e \mathbf{Y} é um vetor de dimensões $n \times 1$, o produto $\mathbf{Y}'\mathbf{A}\mathbf{Y}$ é chamado de forma quadrática e a matriz \mathbf{A} é denominada de núcleo da forma quadrática. É comum, principalmente na abordagem escalar da análise de variância, denominar as formas quadráticas de soma de quadrados. O quociente de uma soma de quadrado pelo posto da matriz núcleo da sua forma quadrática é definido como quadrado médio.

Em seqüência, para o desenvolvimento da análise de variância, alguns teoremas, corolários e definições importantes são enunciados, os quais são baseados nas obras de Rencher e Schaalje (2008), Gentle (2007), Graybill (1976) e Searle (1971).

Teorema 2. *Seja $\mathbf{U} = \mathbf{U} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, de dimensões $p \times 1$, $\mathbf{a} = \mathbf{a}$ um vetor de constantes, de dimensões $p \times 1$, e \mathbf{A} uma matriz de constantes, de dimensões $k \times p$, e posto $\mathcal{P}(\mathbf{A}) = k \leq p$. Então,*

$$(i) \quad Z = \mathbf{a}'\mathbf{U} \sim N(\mathbf{a}'\boldsymbol{\mu}, \mathbf{a}'\boldsymbol{\Sigma}\mathbf{a})$$

$$(ii) \quad \mathbf{Z} = \mathbf{A}\mathbf{U} \sim N(\mathbf{A}\boldsymbol{\mu}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}'), \text{ de dimensões } k \times 1.$$

Corolário 1. *Se \mathbf{b} é um vetor de constantes, de dimensões $k \times 1$, então*

$$\mathbf{Z} = \mathbf{A}\mathbf{U} + \mathbf{b} \sim N(\mathbf{A}\boldsymbol{\mu} + \mathbf{b}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}').$$

O teorema e corolário enunciados estão demonstrados em Rencher e Schaalje (2008).

Como resultado do corolário anterior, para $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, com $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{I}\sigma^2)$, tem-se que

$$\mathbf{Y} \sim N(\mathbf{X}\boldsymbol{\beta}, \mathbf{I}\sigma^2)$$

Definição 1. *Sejam Z_1, \dots, Z_n , n variáveis aleatórias independentes e normalmente distribuídas, com média zero e variância um. A variável aleatória $U = \mathbf{Z}'\mathbf{Z} = \sum_{i=1}^n Z_i^2 \sim \chi_{(n)}^2$, isto é, a variável aleatória $U = \sum_{i=1}^n Z_i^2$ tem distribuição qui-quadrado central com n número de graus de liberdade.*

Definição 2. *Sejam Y_1, \dots, Y_n , n variáveis aleatórias independentes e normalmente distribuídas, com média μ_i e variância um, $i = 1, \dots, n$. A variável aleatória $U = \mathbf{Y}'\mathbf{Y} = \sum_{i=1}^n Y_i^2 \sim \chi_{(n, \lambda)}^2$, isto é, a variável aleatória $U = \sum_{i=1}^n Y_i^2$*

tem distribuição qui-quadrado não central com n número de graus de liberdade e parâmetro de não centralidade $\lambda = \frac{1}{2} \sum_{i=1}^n \mu_i^2$.

Teorema 3 (Cochran). Dado $\mathbf{Y} \sim N(0, \mathbf{I}\sigma^2)$, com dimensões $n \times 1$, e $\mathbf{Y}'\mathbf{Y}$ decomposto em k formas quadráticas $Q_i = \mathbf{Y}'\mathbf{A}_i\mathbf{Y}$, $i = 1, \dots, k$, com posto de \mathbf{A}_i igual a r_i . Então, Q_1, \dots, Q_k são independentes entre si e $\frac{Q_i}{\sigma^2} \sim \chi_{(r_i)}^2$ se, e somente se, $\sum_{i=1}^k r_i = n$.

Detalhes do teorema de Cochran estão desenvolvidos em Cochran (1934).

Conforme Gentle (2007), assumindo $\mathbf{Y} \sim N(\boldsymbol{\mu}, \mathbf{I}\sigma^2)$ no teorema de Cochran, tem-se $\frac{Q_i}{\sigma^2} \sim \chi_{(r_i, \lambda_i)}^2$, com $\lambda_i = \frac{1}{2} \boldsymbol{\mu}'\mathbf{A}_i\boldsymbol{\mu}$.

Pelo resultado anterior, $\mathbf{Y}'\mathbf{Y}$ pode ser decomposto em formas quadráticas referentes às fontes de variações do modelo considerado.

Segundo Graybill (1976) e Searle (1971), o número de graus de liberdade (GL) de uma soma de quadrados é definido como o posto do núcleo da forma quadrática correspondente.

O posto de uma matriz \mathbf{A} será denotado por $\mathcal{P}(\mathbf{A})$.

Considerando o modelo (2.7), pelo teorema de Cochran, tem-se que

$$\frac{SQ_i}{\sigma^2} = \frac{\mathbf{Y}'\mathbf{A}_i\mathbf{Y}}{\sigma^2} \sim \chi_{(\mathcal{P}(\mathbf{A}_i), \lambda = \frac{1}{2}\boldsymbol{\beta}'\mathbf{X}'\mathbf{A}_i\mathbf{X}\boldsymbol{\beta})}^2,$$

para $i = 1, \dots, k$.

Para a forma quadrática de matriz núcleo $\mathbf{A}_i = \mathbf{I}_n - \mathbf{B}$, com $\mathbf{B} = \mathbf{X}'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}$, a distribuição qui-quadrado associada a $\frac{\mathbf{Y}'\mathbf{A}_i\mathbf{Y}}{\sigma^2}$ é central, ou seja,

$$\frac{\mathbf{Y}'(\mathbf{I}_n - \mathbf{B})\mathbf{Y}}{\sigma^2} \sim \chi_{(\mathcal{P}(\mathbf{I}_n - \mathbf{B}))}^2,$$

pois,

$$\begin{aligned}
 \lambda &= \frac{\beta' \mathbf{X}' (\mathbf{I}_n - \mathbf{B}) \mathbf{X} \beta}{2} = \frac{\beta' \mathbf{X}' \mathbf{X} \beta - \beta' \mathbf{X}' \mathbf{B} \mathbf{X} \beta}{2} = \\
 &= \frac{\beta' \mathbf{X}' \mathbf{X} \beta - \beta' \mathbf{X}' \mathbf{B} \mathbf{X} \beta}{2} = \frac{\beta' \mathbf{X}' \mathbf{X} \beta - \beta' \mathbf{X}' \mathbf{X} (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{X} \beta}{2} = \\
 &= \frac{\beta' \mathbf{X}' \mathbf{X} \beta - \beta' \mathbf{X}' \mathbf{X} \beta}{2} = 0.
 \end{aligned}$$

A forma quadrática $\mathbf{Y}' (\mathbf{I}_n - \mathbf{B}) \mathbf{Y}$ é definida como soma de quadrados de resíduos, e denotada por SQ_{Res} .

Definição 3 (Distribuição F de Snedecor central). *Seja $V_1 \sim \chi^2_{(m_1)}$, $V_2 \sim \chi^2_{(m_2)}$ com V_1 e V_2 independentes, então*

$$W = \frac{V_1/m_1}{V_2/m_2} \sim F_{(m_1, m_2)},$$

uma distribuição F central com m_1 e m_2 números de graus de liberdade.

Decorre imediatamente da definição anterior que, para duas formas quadráticas, $\mathbf{Y}' \mathbf{A}_j \mathbf{Y}$ e $\mathbf{Y}' \mathbf{A}_k \mathbf{Y}$, com distribuição qui-quadrado central com $\mathcal{P}(\mathbf{A}_j)$ e $\mathcal{P}(\mathbf{A}_k)$ números de graus de liberdade, respectivamente, o quociente

$$\frac{\mathbf{Y}' \mathbf{A}_j \mathbf{Y} / \mathcal{P}(\mathbf{A}_j)}{\mathbf{Y}' \mathbf{A}_k \mathbf{Y} / \mathcal{P}(\mathbf{A}_k)}$$

tem distribuição F com $\mathcal{P}(\mathbf{A}_j)$ e $\mathcal{P}(\mathbf{A}_k)$ números de graus de liberdade.

Teorema 4. *Se \mathbf{Y} é um vetor aleatório com média $\boldsymbol{\mu}$ e matriz de covariâncias $\boldsymbol{\Sigma}$ e, se \mathbf{A} é uma matriz simétrica de constantes, então*

$$E(\mathbf{Y}' \mathbf{A} \mathbf{Y}) = \text{tr}(\mathbf{A} \boldsymbol{\Sigma}) + \boldsymbol{\mu}' \mathbf{A} \boldsymbol{\mu}$$

em que $\text{tr}(\cdot)$ representa o operador traço, isto é, a soma dos elementos da diagonal principal de uma matriz.

Pelo teorema anterior, tem-se

$$\begin{aligned} E(SQ_{\text{Res}}) &= E[\mathbf{Y}'(\mathbf{I}_n - \mathbf{B})\mathbf{Y}] = \\ &= \sigma^2 \text{tr}(\mathbf{I}_n - \mathbf{B}) + \boldsymbol{\beta}'\mathbf{X}'(\mathbf{I}_n - \mathbf{B})\mathbf{X}\boldsymbol{\beta} = \sigma^2 \mathcal{P}(\mathbf{I}_n - \mathbf{B}). \end{aligned}$$

Ou seja, o quadrado médio do resíduo, definido por $QM_{\text{Res}} = \frac{SQ_{\text{Res}}}{\mathcal{P}(\mathbf{I}_n - \mathbf{B})}$, é um estimador não viciado para a variância σ^2 , pois

$$E(QM_{\text{Res}}) = E\left(\frac{SQ_{\text{Res}}}{\mathcal{P}(\mathbf{I}_n - \mathbf{B})}\right) = \frac{E(SQ_{\text{Res}})}{\mathcal{P}(\mathbf{I}_n - \mathbf{B})} = \frac{\mathcal{P}(\mathbf{I}_n - \mathbf{B})}{\mathcal{P}(\mathbf{I}_n - \mathbf{B})}\sigma^2 = \sigma^2.$$

Os resultados desta subseção serão úteis na elaboração de testes que verificarão a influência dos tratamentos na variável resposta de um experimento. Esses testes serão apresentados nas seções seguintes, em momentos oportunos.

2.2.2.1 Análise de variância pela abordagem matricial

Nesta subseção será desenvolvida a análise de variância pela abordagem matricial, em especial, para o modelo do delineamento em blocos casualizados em esquema fatorial 2^k , que é uma generalização do delineamento inteiramente casualizado em esquema fatorial 2^k .

Note-se que, o modelo do delineamento em blocos casualizados, (2.4),

pode ser escrito conforme (2.7); basta considerar

$$\mathbf{Y} = \begin{pmatrix} Y_{11} \\ \vdots \\ Y_{1J} \\ \vdots \\ Y_{I1} \\ \vdots \\ Y_{IJ} \end{pmatrix}, \mathbf{X} = \begin{pmatrix} 1 & 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots \\ 1 & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 \\ \vdots & \vdots \\ 1 & 0 & \cdots & 0 & 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots \\ 1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 1 \end{pmatrix},$$

$$\boldsymbol{\beta} = \begin{pmatrix} \mu \\ \tau_1 \\ \vdots \\ \tau_I \\ b_1 \\ \vdots \\ b_J \end{pmatrix}, \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_{11} \\ \vdots \\ \varepsilon_{1J} \\ \vdots \\ \varepsilon_{I1} \\ \vdots \\ \varepsilon_{IJ} \end{pmatrix},$$

sendo que a matriz de delineamento \mathbf{X} pode ser decomposta em

$$\mathbf{X} = \left(\mathbf{X}_1, \mathbf{X}_t, \mathbf{X}_{bl} \right), \quad (2.11)$$

em que

\mathbf{X}_1 é um vetor contendo apenas o elemento um (1), de dimensões $n \times 1$, referente ao parâmetro μ ;

\mathbf{X}_t é uma matriz referente aos parâmetros de efeito de tratamento $\tau_i, i = 1, \dots, I$, de dimensões $n \times I$;

\mathbf{X}_{bl} é uma matriz referente aos parâmetros de efeito de bloco $b_j, j = 1, \dots, J$, de

dimensões $n \times J$.

Pelo teorema de Cochran, $\mathbf{Y}'\mathbf{Y}$ pode ser decomposto em formas quadráticas referentes às fontes de variações do modelo do delineamento em blocos casualizados, (2.4), com a matriz de delineamento particionada, conforme (2.11),

$$\mathbf{Y}'\mathbf{Y} = \mathbf{Y}'\mathbf{P}_1\mathbf{Y} + \mathbf{Y}'\mathbf{P}_t\mathbf{Y} + \mathbf{Y}'\mathbf{P}_{bl}\mathbf{Y} + \mathbf{Y}'\mathbf{P}_{\text{Res}}\mathbf{Y} \quad (2.12)$$

em que

$$\mathbf{P}_1 = \mathbf{X}_1(\mathbf{X}'_1\mathbf{X}_1)^{-1}\mathbf{X}'_1;$$

$$\mathbf{P}_t = \mathbf{B}_t - \mathbf{P}_1, \text{ com } \mathbf{B}_t = \mathbf{X}_t(\mathbf{X}'_t\mathbf{X}_t)^{-1}\mathbf{X}'_t;$$

$$\mathbf{P}_{bl} = \mathbf{B}_{bl} - \mathbf{P}_1, \text{ com } \mathbf{B}_{bl} = \mathbf{X}_{bl}(\mathbf{X}'_{bl}\mathbf{X}_{bl})^{-1}\mathbf{X}'_{bl};$$

$$\mathbf{P}_{\text{Res}} = \mathbf{I}_n - \mathbf{B}, \text{ com } \mathbf{B} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}';$$

\mathbf{I}_n é a matriz identidade, de dimensões $n \times n$.

Os termos $\mathbf{Y}'\mathbf{Y}$, $\mathbf{Y}'\mathbf{P}_1\mathbf{Y}$, $\mathbf{Y}'\mathbf{P}_t\mathbf{Y}$, $\mathbf{Y}'\mathbf{P}_{bl}\mathbf{Y}$ e $\mathbf{Y}'\mathbf{P}_{\text{Res}}\mathbf{Y}$ são definidos, respectivamente, como somas de quadrados relativos à variação total, à variação do efeito da constante μ , à variação entre tratamentos, à variação entre blocos e à variação residual.

O quociente da soma de quadrados de tratamentos pela variância σ^2 segue distribuição qui-quadrado, com $\mathcal{P}(\mathbf{P}_t)$ número de graus de liberdade e parâmetro de não centralidade

$$\lambda = \frac{\boldsymbol{\beta}'\mathbf{X}'(\mathbf{B}_t - \mathbf{P}_1)\mathbf{X}\boldsymbol{\beta}}{2} = \frac{J}{2I} \sum_{r=1}^{I-1} \sum_{s=r+1}^I (\tau_r - \tau_s)^2.$$

Caso os efeitos de tratamento τ_i , $i = 1, \dots, I$, sejam todos iguais, o que será denominado aqui como H_0 , isto é, a hipótese de nulidade, o parâmetro de

não centralidade $\lambda = \frac{J}{2I} \sum_{r=1}^{I-1} \sum_{s=r+1}^I (\tau_r - \tau_s)^2$ é igual a zero e a distribuição referente ao quociente da soma de quadrados de tratamentos pela variância σ^2 segue distribuição qui-quadrado central, com $\mathcal{P}(\mathbf{P}_t)$ número de graus de liberdade.

Para a esperança da soma de quadrados de tratamentos, tem-se

$$\begin{aligned} E(SQ_t) &= E[\mathbf{Y}'(\mathbf{B}_t - \mathbf{P}_1)\mathbf{Y}] = \\ &= \sigma^2 \text{tr}(\mathbf{B}_t - \mathbf{P}_1) + \beta' \mathbf{X}'(\mathbf{B}_t - \mathbf{P}_1)\mathbf{X}\beta = \\ &= \sigma^2 \mathcal{P}(\mathbf{B}_t - \mathbf{P}_1) + \frac{J}{I} \sum_{r=1}^{I-1} \sum_{s=r+1}^I (\tau_r - \tau_s)^2. \end{aligned}$$

E, para o quadrado médio de tratamento, $QM_t = \frac{SQ_t}{\mathcal{P}(\mathbf{B}_t - \mathbf{P}_1)}$,

$$\begin{aligned} E(QM_t) &= E\left(\frac{SQ_t}{\mathcal{P}(\mathbf{B}_t - \mathbf{P}_1)}\right) = \sigma^2 + \frac{\frac{J}{I} \sum_{r=1}^{I-1} \sum_{s=r+1}^I (\tau_r - \tau_s)^2}{\mathcal{P}(\mathbf{B}_t - \mathbf{P}_1)} = \\ &= \sigma^2 + f(\tau_i), \quad f(\tau_i) \geq 0 \end{aligned}$$

Naturalmente, se $H_0 : \tau_1 = \dots = \tau_I$, então $f(\tau_i) = 0$ e $\frac{\sigma^2 + f(\tau_i)}{\sigma^2} = 1$. Então, é possível estabelecer uma regra para verificar se $\frac{\sigma^2 + f(\tau_i)}{\sigma^2}$ é significativamente diferente de 1, basta que se $\tau_i \neq \tau_j$ para pelo menos um $i \neq j$, com $1 \leq i, j \leq I$. Assim, como já é de conhecimento que o quociente da soma de quadrados de resíduos pela variância σ^2 segue distribuição qui-quadrado central, isto é,

$$\mathcal{P}(\mathbf{I}_n - \mathbf{B}) \frac{QM_{\text{Res}}}{\sigma^2} = \frac{1}{\sigma^2} \mathbf{Y}'(\mathbf{I}_n - \mathbf{B})\mathbf{Y} \sim \chi_{(\mathcal{P}(\mathbf{I}_n - \mathbf{B}))}^2$$

e sob H_0 , isto é, $\tau_1 = \dots = \tau_I$, a distribuição do quociente da soma de quadrados de tratamentos pela variância σ^2 também apresenta distribuição qui-quadrado

central, ou seja,

$$\mathcal{P}(\mathbf{B}_t - \mathbf{P}_1) \frac{QM_t}{\sigma^2} = \frac{1}{\sigma^2} \mathbf{Y}' (\mathbf{B}_t - \mathbf{P}_1) \mathbf{Y} \sim \chi^2_{(\mathcal{P}(\mathbf{B}_t - \mathbf{P}_1))}.$$

Portanto, pela definição de distribuição F de Snedecor central, uma estatística de teste para as hipóteses $H_0 : \tau_1 = \dots = \tau_I$ e $H_1 : \tau_i \neq \tau_j$, para pelo menos um $i \neq j$, com $1 \leq i, j \leq I$, é dada por

$$F_c = \frac{QM_t}{QM_{\text{Res}}}, \quad (2.13)$$

que sob H_0 , tem-se

$$\frac{QM_t}{QM_{\text{Res}}} \sim F_{(\mathcal{P}(\mathbf{P}_t), \mathcal{P}(\mathbf{P}_{\text{Res}}))}.$$

O nível descritivo (valor - p) para o teste supracitado é dado pela probabilidade:

$$\text{valor - p} = P \left(F_{(\mathcal{P}(\mathbf{P}_t), \mathcal{P}(\mathbf{P}_{\text{Res}}))} > \frac{QM_t}{QM_{\text{Res}}} \right) = P(F > F_c).$$

Rejeita-se a hipótese H_0 se valor - p < α , sendo α o nível de significância adotado.

Analogamente, para o efeito de blocos, o quociente da soma de quadrados de blocos pela variância σ^2 segue distribuição qui-quadrado com $\mathcal{P}(\mathbf{P}_{bl})$ número de graus de liberdade e parâmetro de não centralidade

$$\lambda = \frac{\boldsymbol{\beta}' \mathbf{X}' (\mathbf{B}_{bl} - \mathbf{P}_1) \mathbf{X} \boldsymbol{\beta}}{2} = \frac{I}{2J} \sum_{t=1}^{J-1} \sum_{u=t+1}^J (b_t - b_u)^2.$$

Também neste caso, se os efeitos de blocos b_j , $j = 1, \dots, J$ são todos iguais, que também será denominado por H_0 , o parâmetro de não centralidade

$\lambda = \frac{I}{2J} \sum_{t=1}^{J-1} \sum_{u=t+1}^J (b_t - b_u)^2$ é igual a zero e o quociente da soma de quadrados de blocos pela variância σ^2 segue distribuição qui-quadrado central com $\mathcal{P}(\mathbf{P}_{bl})$ graus de liberdade.

Para a esperança da soma de quadrados de blocos, tem-se

$$\begin{aligned} E(SQ_{bl}) &= E[\mathbf{Y}'(\mathbf{B}_{bl} - \mathbf{P}_1)\mathbf{Y}] = \\ &= \sigma^2 \text{tr}(\mathbf{B}_{bl} - \mathbf{P}_1) + \beta' \mathbf{X}'(\mathbf{B}_{bl} - \mathbf{P}_1)\mathbf{X}\beta = \\ &= \sigma^2 \mathcal{P}(\mathbf{B}_{bl} - \mathbf{P}_1) + \frac{I}{J} \sum_{t=1}^{J-1} \sum_{u=t+1}^J (b_t - b_u)^2. \end{aligned}$$

E, para o quadrado médio de blocos, $QM_{bl} = \frac{SQ_{bl}}{\mathcal{P}(\mathbf{B}_{bl} - \mathbf{P}_1)}$,

$$\begin{aligned} E(QM_{bl}) &= E\left(\frac{SQ_{bl}}{\mathcal{P}(\mathbf{B}_{bl} - \mathbf{P}_1)}\right) = \sigma^2 + \frac{\frac{I}{J} \sum_{t=1}^{J-1} \sum_{u=t+1}^J (b_t - b_u)^2}{\mathcal{P}(\mathbf{B}_{bl} - \mathbf{P}_1)} = \\ &= \sigma^2 + g(b_j), \quad g(b_j) \geq 0 \end{aligned}$$

Sob $H_0 : b_1 = \dots = b_J$ tem-se que $g(b_j) = 0$ e $\frac{\sigma^2 + g(b_j)}{\sigma^2} = 1$. Desta forma, também é possível estabelecer uma regra para verificar se $\frac{\sigma^2 + g(b_j)}{\sigma^2}$ é significativamente diferente de 1, ou seja, se $b_i \neq b_j$ para pelo menos um $i \neq j$, com $1 \leq i, j \leq J$. Assim, como é de conhecimento, sob igualdade dos efeitos de blocos, o quociente da soma de quadrados de blocos pela variância σ^2 segue distribuição qui-quadrado central com $\mathcal{P}(\mathbf{B}_{bl} - \mathbf{P}_1)$ número de graus de liberdade, isto é,

$$\mathcal{P}(\mathbf{B}_{bl} - \mathbf{P}_1) \frac{QM_{bl}}{\sigma^2} = \frac{1}{\sigma^2} \mathbf{Y}'(\mathbf{B}_{bl} - \mathbf{P}_1)\mathbf{Y} \sim \chi_{(\mathcal{P}(\mathbf{B}_{bl} - \mathbf{P}_1))}^2.$$

Portanto, pela definição de distribuição F de Snedecor central, uma esta-

tística de teste para as hipóteses $H_0 : b_1 = \dots = b_J$ e $H_1 : b_i \neq b_j$, para pelo menos um $i \neq j$, com $1 \leq i, j \leq J$, é dada por

$$F_c = \frac{QM_{bl}}{QM_{Res}}, \quad (2.14)$$

que sob H_0 , tem-se

$$\frac{QM_{bl}}{QM_{Res}} \sim F_{(\mathcal{P}(\mathbf{P}_{bl}), \mathcal{P}(\mathbf{P}_{Res}))}.$$

O nível descritivo (valor - p) para o teste supracitado é dado pela probabilidade:

$$\text{valor - p} = P \left(F_{(\mathcal{P}(\mathbf{P}_{bl}), \mathcal{P}(\mathbf{P}_{Res}))} > \frac{QM_{bl}}{QM_{Res}} \right) = P(F > F_c).$$

Rejeita-se a hipótese H_0 se valor - p < α , sendo α o nível de significância adotado.

Em relação à decomposição de $\mathbf{Y}'\mathbf{Y}$, em especial o termo referente ao efeito da constante μ , $\mathbf{Y}'\mathbf{P}_1\mathbf{Y}$, nem sempre há interesse prático em estudá-lo, por isso a decomposição de $\mathbf{Y}'\mathbf{Y}$ é comumente corrigida para a constante μ , isto é,

$$\mathbf{Y}'(\mathbf{I}_n - \mathbf{P}_1)\mathbf{Y} = \mathbf{Y}'\mathbf{P}_t\mathbf{Y} + \mathbf{Y}'\mathbf{P}_{bl}\mathbf{Y} + \mathbf{Y}'\mathbf{P}_{Res}\mathbf{Y} \quad (2.15)$$

em que $\mathbf{Y}'(\mathbf{I}_n - \mathbf{P}_1)\mathbf{Y}$ é definido como soma de quadrado total corrigido e denotado por SQ_{T_c} .

É comum resumir os resultados da análise de variância em uma tabela, denominada tabela de análise de variância ou esquema de análise de variância. A tabela de análise de variância, como apresentada na literatura, para os testes de hipóteses

$$H_0 : \tau_1 = \dots = \tau_I$$

$$H_1 : \tau_i \neq \tau_j, \text{ para pelo menos um } i \neq j, \text{ com } 1 \leq i, j \leq I$$

e

$$H_0 : b_1 = \dots = b_J$$

$$H_1 : b_i \neq b_j, \text{ para pelo menos um } i \neq j, \text{ com } 1 \leq i, j \leq J$$

pode ser construída tal como a Tabela 4.

Tabela 4 Esquema da análise de variância para experimentos no delineamento em blocos casualizados.

Fonte de Variação	GL	SQ	QM	F_c
Tratamentos	$\mathcal{P}(\mathbf{P}_t)$	$\mathbf{Y}'\mathbf{P}_t\mathbf{Y}$	$\frac{\mathbf{Y}'\mathbf{P}_t\mathbf{Y}}{\mathcal{P}(\mathbf{P}_t)}$	$\frac{QM_t}{QM_{Res}}$
Blocos	$\mathcal{P}(\mathbf{P}_{bl})$	$\mathbf{Y}'\mathbf{P}_{bl}\mathbf{Y}$	$\frac{\mathbf{Y}'\mathbf{P}_{bl}\mathbf{Y}}{\mathcal{P}(\mathbf{P}_{bl})}$	$\frac{QM_{bl}}{QM_{Res}}$
Resíduo	$\mathcal{P}(\mathbf{P}_{Res})$	$\mathbf{Y}'\mathbf{P}_{Res}\mathbf{Y}$	$\frac{\mathbf{Y}'\mathbf{P}_{Res}\mathbf{Y}}{\mathcal{P}(\mathbf{P}_{Res})}$	
Total	$\mathcal{P}(\mathbf{I}_n - \mathbf{P}_1)$	$\mathbf{Y}'(\mathbf{I}_n - \mathbf{P}_1)\mathbf{Y}$		

Para experimentos fatoriais 2^k a matriz de delineamento \mathbf{X} pode ser decomposta em

$$\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_A, \dots, \mathbf{X}_K, \mathbf{X}_{AB}, \dots, \mathbf{X}_{JK}, \dots, \mathbf{X}_{A \dots K}, \mathbf{X}_{bl}) \quad (2.16)$$

Sendo \mathbf{X}_l referente aos parâmetros de efeito principal, ou interação, ou bloco, $l = A, \dots, K, AB, \dots, JK, \dots, A \dots K, bl$.

Da mesma forma que, pelo teorema de Cochran, $\mathbf{Y}'\mathbf{Y}$ pode ser decomposto em formas quadráticas tal como em (2.12), é possível decompô-la em formas quadráticas referentes às fontes de variações do modelo do delineamento em

blocos casualizados em esquema fatorial 2^k , (2.5), com a matriz de delineamento particionada conforme (2.16),

$$\begin{aligned} \mathbf{Y}'\mathbf{Y} &= \mathbf{Y}'\mathbf{P}_1\mathbf{Y} + \mathbf{Y}'\mathbf{P}_A\mathbf{Y} + \cdots + \mathbf{Y}'\mathbf{P}_K\mathbf{Y} + \\ &+ \mathbf{Y}'\mathbf{P}_{AB}\mathbf{Y} + \cdots + \mathbf{Y}'\mathbf{P}_{JK}\mathbf{Y} + \cdots + \\ &+ \mathbf{Y}'\mathbf{P}_{A\cdots K}\mathbf{Y} + \mathbf{Y}'\mathbf{P}_{bl}\mathbf{Y} + \mathbf{Y}'\mathbf{P}_{\text{Res}}\mathbf{Y} \end{aligned} \quad (2.17)$$

em que

$$\mathbf{P}_1 = \mathbf{X}_1(\mathbf{X}'_1\mathbf{X}_1)^{-1}\mathbf{X}'_1;$$

$$\mathbf{P}_A = \mathbf{B}_A - \mathbf{P}_1, \text{ com } \mathbf{B}_A = \mathbf{X}_A(\mathbf{X}'_A\mathbf{X}_A)^{-1}\mathbf{X}'_A;$$

\cdots ;

$$\mathbf{P}_K = \mathbf{B}_K - \mathbf{P}_1, \text{ com } \mathbf{B}_K = \mathbf{X}_K(\mathbf{X}'_K\mathbf{X}_K)^{-1}\mathbf{X}'_K;$$

$$\mathbf{P}_{AB} = \mathbf{B}_{AB} - \mathbf{P}_A - \mathbf{P}_B - \mathbf{P}_1, \text{ com } \mathbf{B}_{AB} = \mathbf{X}_{AB}(\mathbf{X}'_{AB}\mathbf{X}_{AB})^{-1}\mathbf{X}'_{AB};$$

\cdots ;

$$\mathbf{P}_{JK} = \mathbf{B}_{JK} - \mathbf{P}_J - \mathbf{P}_K - \mathbf{P}_1, \text{ com } \mathbf{B}_{JK} = \mathbf{X}_{JK}(\mathbf{X}'_{JK}\mathbf{X}_{JK})^{-1}\mathbf{X}'_{JK};$$

\cdots ;

$$\mathbf{P}_{A\cdots K} = \mathbf{B}_{A\cdots K} - \mathbf{P}_A - \cdots - \mathbf{P}_K - \mathbf{P}_{AB} - \cdots - \mathbf{P}_{JK} - \cdots - \mathbf{P}_{B\cdots K} - \mathbf{P}_1,$$

$$\text{com } \mathbf{B}_{A\cdots K} = \mathbf{X}_{A\cdots K}(\mathbf{X}'_{A\cdots K}\mathbf{X}_{A\cdots K})^{-1}\mathbf{X}'_{A\cdots K};$$

$$\mathbf{P}_{bl} = \mathbf{B}_{bl} - \mathbf{P}_1, \text{ com } \mathbf{B}_{bl} = \mathbf{X}_{bl}(\mathbf{X}'_{bl}\mathbf{X}_{bl})^{-1}\mathbf{X}'_{bl};$$

$$\mathbf{P}_{\text{Res}} = \mathbf{I}_n - \mathbf{P}_A - \cdots - \mathbf{P}_K - \mathbf{P}_{AB} - \cdots - \mathbf{P}_{JK} - \cdots - \mathbf{P}_{A\cdots K} - \mathbf{P}_{bl} - \mathbf{P}_1,$$

com \mathbf{I}_n a matriz identidade, de dimensões $n \times n$.

Os termos $\mathbf{Y}'\mathbf{P}_A\mathbf{Y}$, $\mathbf{Y}'\mathbf{P}_K\mathbf{Y}$, $\mathbf{Y}'\mathbf{P}_{AB}\mathbf{Y}$, $\mathbf{Y}'\mathbf{P}_{JK}\mathbf{Y}$ e $\mathbf{Y}'\mathbf{P}_{A\cdots K}\mathbf{Y}$ são definidos, respectivamente, como soma de quadrados relativos ao efeito principal do fator A , à interação dupla, ou de ordem 2, dos fatores AB , à interação dupla dos fatores JK e à interação k -ésima, ou de ordem k , dos fatores $AB\cdots K$. Em geral, para $l = 1, \cdots, k$, o termo $\mathbf{Y}'\mathbf{P}_{r_l}\mathbf{Y}$ é definido como soma de quadrados

relativo ao efeito principal do fator r_1 , $r_1 = A, \dots, K$ ou, ao efeito da l -ésima interação dos fatores r_l , $r_2 = AB, \dots, JK, \dots, r_k = A \dots K$.

Em geral, no experimento fatorial 2^k , para verificar as hipóteses referentes a significância dos efeitos principais ou interações ou blocos, basta considerar a estatística de teste

$$F_c = \frac{QM_i}{QM_{\text{Res}}}, \quad (2.18)$$

para $i = A, \dots, K, AB, \dots, JK, \dots, A \dots K, bl$, que sob H_0 , isto é, igualdade entre os efeitos de i , tem-se

$$\frac{QM_i}{QM_{\text{Res}}} \sim F_{(\mathcal{P}(\mathbf{P}_i), \mathcal{P}(\mathbf{P}_{\text{Res}}))}.$$

O nível descritivo (valor - p) para o teste supracitado é dado pela probabilidade:

$$\text{valor - p} = P \left(F_{(\mathcal{P}(\mathbf{P}_i), \mathcal{P}(\mathbf{P}_{\text{Res}}))} > \frac{QM_i}{QM_{\text{Res}}} \right) = P(F > F_c).$$

Rejeita-se a hipótese H_0 se valor - p < α , sendo α o nível de significância adotado.

Note-se que, pelo fato do experimento considerado ser em esquema fatorial 2^k , $\mathcal{P}(\mathbf{P}_i) = 1$, $i = A, \dots, K, AB, \dots, JK, \dots, A \dots K$ e, para J blocos, $\mathcal{P}(\mathbf{P}_{\text{Res}}) = (2^k - 1)(J - 1)$. Caso seja considerado o delineamento inteiramente casualizado ao invés de blocos casualizados, tem-se que $\mathcal{P}(\mathbf{P}_{\text{Res}}) = 2^k(J - 1)$.

Em geral, a tabela da análise de variância para um delineamento em blocos casualizados em esquema fatorial 2^k é semelhante à Tabela 5.

Tabela 5 Esquema da análise de variância para experimentos no delineamento em blocos casualizados em esquema fatorial 2^k , com J blocos (Adaptado de Montgomery (2009)).

Fonte de Variação	GL	SQ
k efeitos principais		
A	1	SQ_A
B	1	SQ_B
\vdots	\vdots	\vdots
K	1	SQ_K
$\binom{k}{2}$ interações 2-fatores		
AB	1	SQ_{AB}
AC	1	SQ_{AC}
\vdots	\vdots	\vdots
JK	1	SQ_{JK}
$\binom{k}{3}$ interações 3-fatores		
ABC	1	SQ_{ABC}
ABD	1	SQ_{ABD}
\vdots	\vdots	\vdots
IJK	1	SQ_{IJK}
\vdots	\vdots	\vdots
$\binom{k}{k}$ interações k-fatores		
$ABC \cdots K$	1	$SQ_{ABC \cdots K}$
Blocos	$J - 1$	SQ_{bl}
Resíduo	$(2^k - 1)(J - 1)$	SQ_{Res}
Total	$J2^k - 1$	SQ_{T_c}

2.2.2.2 Análise de variância pela abordagem escalar

Nesta subseção será desenvolvida a análise de variância pela abordagem escalar, para o modelo do delineamento em blocos casualizados de efeitos fixos. Os objetivos da análise de variância são verificar algumas hipóteses apropriadas sobre os efeitos dos tratamentos e blocos que, para o modelo (2.4), podem ser sintetizadas em

$$H_0 : \tau_1 = \cdots = \tau_I$$

$$H_1 : \tau_i \neq \tau_j, \text{ para pelo menos um } i \neq j, \text{ com } 1 \leq i, j \leq I$$

e

$$H_0 : b_1 = \cdots = b_I$$

$$H_1 : b_i \neq b_j, \text{ para pelo menos um } i \neq j, \text{ com } 1 \leq i, j \leq J.$$

Para verificar as hipóteses, supõe-se que as componentes aleatórias do modelo são variáveis aleatórias independentes com distribuição normal de média zero e variância constante σ^2 , implicando em

$$Y_{ij} \sim N(\mu + \tau_i + b_j, \sigma^2).$$

Considerando as restrições paramétricas

$$\sum_{i=1}^I \tau_i = 0, \quad (2.19)$$

e

$$\sum_{j=1}^J b_j = 0, \quad (2.20)$$

no modelo considerado, (2.4), tem-se que a constante μ é igual à média geral das observações pois,

$$\frac{\sum_{i=1}^I \sum_{j=1}^J E(Y_{ij})}{IJ} = \frac{\sum_{i=1}^I \sum_{j=1}^J (\mu + \tau_i + b_j)}{IJ} = \frac{IJ\mu}{IJ} + \frac{J \sum_{i=1}^I \tau_i}{IJ} + \frac{I \sum_{j=1}^J b_j}{IJ} = \mu.$$

Com as restrições paramétricas (2.19) e (2.20), os efeitos dos tratamentos e de blocos podem ser considerados como desvios em relação à média geral μ .

Uma forma equivalente de escrever as hipóteses anteriores, considerando as restrições paramétricas mencionadas, é

$$H_0 : \tau_1 = \dots = \tau_I = 0$$

$$H_1 : \tau_i \neq 0 \quad \text{para pelo menos um } i$$

e

$$H_0 : b_1 = \dots = b_J = 0$$

$$H_1 : b_j \neq 0 \quad \text{para pelo menos um } j$$

pois, em particular para os efeitos de tratamentos, se $\tau_1 = \dots = \tau_I$ e $\sum_{i=1}^I \tau_i = 0$ então $\tau_1 = \dots = \tau_I = 0$. E, a negação da afirmação anterior é $\tau_i \neq 0$, para pelo menos um i .

Já é de conhecimento que o procedimento apropriado para provar a igualdade dos efeitos de tratamentos ou efeitos de blocos é a análise de variância. Nessa abordagem escalar, a decomposição da variabilidade total em partes apropriadas é feita por meio de somatórios.

A notação utilizada será baseada em Montgomery (2009), Storck *et al.* (2006) e Garcia-Diaz e Phillips (1995), de tal forma que considerar-se-á: Y_{ij} como

a j -ésima observação do i -ésimo tratamento; $Y_{i.}$ o total das observações do i -ésimo tratamento; $\bar{Y}_{i.}$ a média do i -ésimo tratamento; $Y_{.j}$ o total das observações do j -ésimo bloco; $\bar{Y}_{.j}$ a média do j -ésimo bloco; $Y_{..}$ o total de todas as observações; $\bar{Y}_{..}$ a média de todas as observações, ou média geral. Simbolicamente,

$$Y_{i.} = \sum_{j=1}^J Y_{ij}, \quad \bar{Y}_{i.} = \frac{Y_{i.}}{J}, \quad i = 1, \dots, I,$$

$$Y_{.j} = \sum_{i=1}^I Y_{ij}, \quad \bar{Y}_{.j} = \frac{Y_{.j}}{I}, \quad j = 1, \dots, J,$$

$$Y_{..} = \sum_{i=1}^I \sum_{j=1}^J Y_{ij}, \quad \bar{Y}_{..} = \frac{Y_{..}}{IJ}$$

Em algumas situações práticas é preferível utilizar uma notação mais simplificada para representar as médias dos tratamentos. Considerando um experimento fatorial 2^2 dos fatores A e B , as médias dos tratamentos (1), a , b e ab , podem ser representadas por $\bar{Y}_{(1)}$, \bar{Y}_a , \bar{Y}_b e \bar{Y}_{ab} , ao invés de $\bar{Y}_{1.}$, $\bar{Y}_{2.}$, $\bar{Y}_{3.}$ e $\bar{Y}_{4.}$.

Para a decomposição da variabilidade total dos dados, é útil considerar a soma de quadrados total corrigida como

$$SQ_{T_c} = \sum_{i=1}^I \sum_{j=1}^J (Y_{ij} - \bar{Y}_{..})^2, \quad (2.21)$$

que é usada como uma medida de variabilidade total dos dados.

A expressão (2.21) pode ser reescrita decompondo a variabilidade dos da-

dos,

$$\begin{aligned}
SQ_{T_c} &= \sum_{i=1}^I \sum_{j=1}^J (Y_{ij} - \bar{Y}_{..})^2 = \\
&= \sum_{i=1}^I \sum_{j=1}^J [(\bar{Y}_{i.} - \bar{Y}_{..}) + (\bar{Y}_{.j} - \bar{Y}_{..}) + (Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})]^2 = \\
&= \sum_{i=1}^I \sum_{j=1}^J (\bar{Y}_{i.} - \bar{Y}_{..})^2 + \sum_{i=1}^I \sum_{j=1}^J (\bar{Y}_{.j} - \bar{Y}_{..})^2 + \\
&+ \sum_{i=1}^I \sum_{j=1}^J (Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2 + K. \tag{2.22}
\end{aligned}$$

Sendo K o termo referente à soma dos duplos produtos, isto é,

$$\begin{aligned}
K &= 2 \sum_{i=1}^I \sum_{j=1}^J (\bar{Y}_{i.} - \bar{Y}_{..}) (\bar{Y}_{.j} - \bar{Y}_{..}) + \\
&+ 2 \sum_{i=1}^I \sum_{j=1}^J (\bar{Y}_{i.} - \bar{Y}_{..}) (Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..}) + \\
&+ 2 \sum_{i=1}^I \sum_{j=1}^J (\bar{Y}_{.j} - \bar{Y}_{..}) (Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..}).
\end{aligned}$$

O termo K da expressão (2.22) é igual a zero, pois

$$\begin{aligned}
&2 \sum_{i=1}^I \sum_{j=1}^J (\bar{Y}_{i.} - \bar{Y}_{..}) (\bar{Y}_{.j} - \bar{Y}_{..}) = \\
&= 2 \left(\sum_{i=1}^I \bar{Y}_{i.} - \sum_{i=1}^I \bar{Y}_{..} \right) \sum_{j=1}^J (\bar{Y}_{.j} - \bar{Y}_{..}) = \\
&= 2 \left(\frac{Y_{..}}{J} - \frac{IY_{..}}{IJ} \right) \sum_{j=1}^J (\bar{Y}_{.j} - \bar{Y}_{..}) = 0,
\end{aligned}$$

para

$$\begin{aligned}
& 2 \sum_{i=1}^I \sum_{j=1}^J (\bar{Y}_{i.} - \bar{Y}_{..}) (Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..}) = \\
& = 2 \left(\sum_{i=1}^I \bar{Y}_{i.} - \sum_{i=1}^I \bar{Y}_{..} \right) \sum_{i=1}^I \sum_{j=1}^J (Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..}) = \\
& = 2 \left(\frac{Y_{..}}{J} - \frac{IY_{..}}{IJ} \right) \sum_{i=1}^I \sum_{j=1}^J (Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..}) = 0
\end{aligned}$$

e, para

$$\begin{aligned}
& 2 \sum_{i=1}^I \sum_{j=1}^J (\bar{Y}_{.j} - \bar{Y}_{..}) (Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..}) = \\
& = 2 \left(\sum_{j=1}^J \bar{Y}_{.j} - \sum_{j=1}^J \bar{Y}_{..} \right) \sum_{i=1}^I \sum_{j=1}^J (Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..}) = \\
& = 2 \left(\frac{Y_{..}}{I} - \frac{JY_{..}}{IJ} \right) \sum_{i=1}^I \sum_{j=1}^J (Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..}) = 0.
\end{aligned}$$

Portanto,

$$\begin{aligned}
SQ_{T_c} &= \sum_{i=1}^I \sum_{j=1}^J (Y_{ij} - \bar{Y}_{..})^2 = \sum_{i=1}^I \sum_{j=1}^J (\bar{Y}_{i.} - \bar{Y}_{..})^2 + \sum_{i=1}^I \sum_{j=1}^J (\bar{Y}_{.j} - \bar{Y}_{..})^2 + \\
&+ \sum_{i=1}^I \sum_{j=1}^J (Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2. \tag{2.23}
\end{aligned}$$

A soma de quadrados total corrigida em (2.23) está decomposta em três somas de quadrados. Os termos já são conhecidos como soma de quadrados total corrigido (ou, simplesmente soma de quadrados total), soma de quadrados referente aos efeitos de tratamentos (ou, simplesmente soma de quadrados de tratamentos), soma de quadrados referente aos efeitos de blocos (ou, simplesmente

soma de quadrados de blocos) e soma de quadrados de resíduos. Matematicamente, tem-se

$$SQ_t = \sum_{i=1}^I \sum_{j=1}^J (\bar{Y}_{i.} - \bar{Y}_{..})^2 \quad (2.24)$$

$$SQ_{bl} = \sum_{i=1}^I \sum_{j=1}^J (\bar{Y}_{.j} - \bar{Y}_{..})^2 \quad (2.25)$$

$$SQ_{Res} = \sum_{i=1}^I \sum_{j=1}^J (Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2. \quad (2.26)$$

A expressão (2.23) pode ser escrita simbolicamente como

$$SQ_{T_c} = SQ_t + SQ_{bl} + SQ_{Res}. \quad (2.27)$$

A expressão (2.27) é fundamental na análise de variância para o modelo considerado.

Após o cálculo das somas de quadrados previamente citadas, novas estatísticas podem ser obtidas pela divisão de cada soma de quadrados pelo respectivo número de graus de liberdade. Conforme Garcia-Diaz e Phillips (1995), na abordagem escalar, o número de graus de liberdade é considerado como o número de termos independentes na soma de quadrados. Para as somas de quadrados de tratamentos, blocos e resíduos, obtém-se

$$QM_t = \frac{SQ_t}{I-1} = \frac{\sum_{i=1}^I \sum_{j=1}^J (\bar{Y}_{i.} - \bar{Y}_{..})^2}{I-1}$$

$$QM_{bl} = \frac{SQ_{bl}}{J-1} = \frac{\sum_{i=1}^I \sum_{j=1}^J (\bar{Y}_{.j} - \bar{Y}_{..})^2}{J-1}$$

$$QM_{\text{Res}} = \frac{SQ_{\text{Res}}}{(I-1)(J-I)} = \frac{\sum_{i=1}^I \sum_{j=1}^J (Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2}{(I-1)(J-I)},$$

sendo $I-1$, $J-1$ e $(I-1)(J-I)$ os números de graus de liberdade associados, respectivamente, às fontes de variações dos efeitos de tratamentos, blocos e resíduos.

Garcia-Diaz e Phillips (1995) apresentam os cálculos para o valor esperado do quadrado médio de tratamentos, QM_t , que é obtido como segue

$$E(QM_t) = E\left(\frac{SQ_t}{I-1}\right) = \frac{E(SQ_t)}{I-1}$$

Adicionalmente,

$$E(SQ_t) = E\left[\sum_{i=1}^I \sum_{j=1}^J (\bar{Y}_{i.} - \bar{Y}_{..})^2\right]$$

Lembrando que está sendo considerado igual a J o número de repetições de todos os tratamentos. Assim,

$$E(SQ_t) = J \sum_{i=1}^I E\left[(\bar{Y}_{i.} - \bar{Y}_{..})^2\right]$$

Por definição,

$$\bar{Y}_{..} = \frac{1}{I} (\bar{Y}_{1.} + \dots + \bar{Y}_{I.})$$

Portanto,

$$\bar{Y}_{i.} - \bar{Y}_{..} = \frac{I-1}{I} \bar{Y}_{i.} - \frac{1}{I} \sum_{l \neq i}^I \bar{Y}_{l.} \quad (2.28)$$

Lembrando que, para qualquer variável aleatória X ,

$$E(X^2) = Var(X) + [E(X)]^2.$$

Como $\bar{Y}_i. - \bar{Y}..$ é uma variável aleatória, segue que

$$E(SQ_t) = J \sum_{i=1}^I Var(\bar{Y}_i. - \bar{Y}..) + J \sum_{i=1}^I [E(\bar{Y}_i. - \bar{Y}..)]^2. \quad (2.29)$$

Considerando as expressões (2.19) e (2.28) no desenvolvimento de (2.29), obtém-se

$$E(SQ_t) = J \sum_{i=1}^I Var(\bar{Y}_i. - \bar{Y}..) + J \sum_{i=1}^I \tau_i^2.$$

Usando a expressão (2.28), e lembrando que ε_{ij} são variáveis aleatórias independentes com distribuição normal de média zero e variância constante σ^2 , verifica-se que

$$\begin{aligned} Var(\bar{Y}_i. - \bar{Y}..) &= \frac{(I-1)^2}{I^2} Var(\bar{Y}_i.) + \frac{(I-1)}{I^2} Var(\bar{Y}_i.) = \\ &= \frac{(I-1)}{I} Var(\bar{Y}_i.) = \frac{(I-1)}{I} \cdot \frac{\sigma^2}{J}. \end{aligned}$$

E, conseqüentemente

$$\begin{aligned} E(SQ_t) &= J \sum_{i=1}^I \frac{(I-1)}{I} \cdot \frac{\sigma^2}{J} + J \sum_{i=1}^I \tau_i^2 = \\ &= (I-1) \sigma^2 + J \sum_{i=1}^I \tau_i^2. \end{aligned}$$

Finalmente,

$$E(QM_t) = \sigma^2 + \frac{J \sum_{i=1}^I \tau_i^2}{I-1}.$$

O cálculo do valor esperado do quadrado médio de blocos, desenvolvido abaixo, é semelhante ao de tratamentos. Pela expressão (2.25), verifica-se que

$$SQ_{bl} = I \sum_{j=1}^J (\bar{Y}_{.j} - \bar{Y}_{..})^2.$$

Logo,

$$\begin{aligned} E(SQ_{bl}) &= E\left(I \sum_{j=1}^J (\bar{Y}_{.j} - \bar{Y}_{..})^2\right) = \\ &= I \sum_{j=1}^J E\left[(\bar{Y}_{.j} - \bar{Y}_{..})^2\right] = \\ &= I \sum_{j=1}^J [E(\bar{Y}_{.j} - \bar{Y}_{..})]^2 + I \sum_{j=1}^J Var(\bar{Y}_{.j} - \bar{Y}_{..}) = \\ &= I \sum_{j=1}^J b_j^2 + I \sum_{j=1}^J Var(\bar{Y}_{.j} - \bar{Y}_{..}). \end{aligned} \quad (2.30)$$

Para o cálculo da variância do termo $\bar{Y}_{.j} - \bar{Y}_{..}$, mostrado no lado direito da expressão (2.30), note-se que

$$\bar{Y}_{.j} - \bar{Y}_{..} = \bar{Y}_{.j} - \frac{1}{J} (\bar{Y}_{.1} + \dots + \bar{Y}_{.J}) = -\frac{\bar{Y}_{.1}}{J} - \dots + (J-1) \frac{\bar{Y}_{.j}}{J} - \dots - \frac{\bar{Y}_{.J}}{J}$$

e

$$Var(\bar{Y}_{.j}) = Var\left(\frac{\sum_{i=1}^I Y_{ij}}{I}\right) = \frac{\sum_{i=1}^I Var(Y_{ij})}{I^2} = \frac{\sigma^2}{I} \text{ para todo } j = 1, \dots, J.$$

Portanto,

$$\begin{aligned} \text{Var}(\bar{Y}_{.j} - \bar{Y}_{..}) &= \text{Var}\left[-\frac{\bar{Y}_{.1}}{J} - \dots + (J-1)\frac{\bar{Y}_{.j}}{J} - \dots - \frac{\bar{Y}_{.J}}{J}\right] = \\ &= \frac{(J-1)}{J^2} \cdot \frac{\sigma^2}{I} + \frac{(J-1)^2}{J^2} \cdot \frac{\sigma^2}{I} = \frac{(J-1)}{IJ} \sigma^2. \end{aligned}$$

Assim, retomando o cálculo do valor esperado da soma de quadrados de blocos, conclui-se que

$$\begin{aligned} E(SQ_{bl}) &= I \sum_{j=1}^J b_j^2 + I \sum_{j=1}^J \frac{(J-1)}{IJ} \sigma^2 = \\ &= I \sum_{j=1}^J b_j^2 + (J-1) \sigma^2. \end{aligned} \quad (2.31)$$

E, finalmente para o quadrado médio de blocos

$$E(QM_{bl}) = \frac{I \sum_{j=1}^J b_j^2}{J-1} + \sigma^2.$$

O valor esperado da soma de quadrados de resíduos é obtido através da expressão (2.27) isolando o termo referente a resíduos. Para isso, faz-se necessário

o cálculo da esperança da soma de quadrados total, que é obtido como segue

$$\begin{aligned}
 E(SQ_{T_c}) &= E \left[\sum_{i=1}^I \sum_{j=1}^J (Y_{ij} - \bar{Y}_{..})^2 \right] = \\
 &= \sum_{i=1}^I \sum_{j=1}^J E(Y_{ij} - \bar{Y}_{..})^2 = \\
 &= \sum_{i=1}^I \sum_{j=1}^J [E(Y_{ij} - \bar{Y}_{..})]^2 + \sum_{i=1}^I \sum_{j=1}^J Var(Y_{ij} - \bar{Y}_{..}) = \\
 &= \sum_{i=1}^I \sum_{j=1}^J (\tau_i + b_j)^2 + (IJ - 1) \sigma^2. \tag{2.32}
 \end{aligned}$$

Agora, note-se que

$$\sum_{i=1}^I \sum_{j=1}^J (\tau_i + b_j)^2 = J \sum_{i=1}^I \tau_i^2 + I \sum_{j=1}^J b_j^2,$$

pois,

$$\begin{aligned}
 \sum_{i=1}^I \sum_{j=1}^J (\tau_i + b_j)^2 &= \sum_{i=1}^I \sum_{j=1}^J (\tau_i^2 + b_j^2 + 2\tau_i b_j) = \\
 &= J \sum_{i=1}^I \tau_i^2 + I \sum_{j=1}^J b_j^2 + 2 \sum_{i=1}^I \sum_{j=1}^J \tau_i b_j = \\
 &= J \sum_{i=1}^I \tau_i^2 + I \sum_{j=1}^J b_j^2 + 2 \sum_{i=1}^I \tau_i \sum_{j=1}^J b_j
 \end{aligned}$$

e, pelas restrições paramétricas (2.19) e (2.20)

$$2 \sum_{i=1}^I \tau_i \sum_{j=1}^J b_j = 0.$$

Portanto,

$$E(SQ_{T_c}) = J \sum_{i=1}^I \tau_i^2 + I \sum_{j=1}^J b_j^2 + (IJ - 1) \sigma^2.$$

Considerando a expressão (2.27) e isolando SQ_{Res} , obtém-se

$$SQ_{\text{Res}} = SQ_{T_c} - SQ_t - SQ_{bl}$$

e, para o valor esperado

$$\begin{aligned} E(SQ_{\text{Res}}) &= E(SQ_{T_c} - SQ_t - SQ_{bl}) = \\ &= E(SQ_{T_c}) - E(SQ_t) - E(SQ_{bl}) = \\ &= (IJ - 1 - J + 1 - I + 1) \sigma^2 = (I - 1)(J - 1) \sigma^2. \end{aligned}$$

E, finalmente para o quadrado médio de resíduos

$$E(QM_{\text{Res}}) = \sigma^2.$$

Observa-se que, sob as hipóteses de nulidade, H_0 , $\tau_i = 0$, para todo i e $b_j = 0$, para todo j , $E(QM_t) = \sigma^2$ e $E(QM_{bl}) = \sigma^2$. Portanto, nesse caso, tem-se três estimativas não viciadas para σ^2 , QM_t , QM_{bl} e QM_{Res} . Assim, assumindo H_0 verdadeira, pelo teorema de Cochran é possível definir três estatísticas independentes com distribuição qui-quadrado central,

$$\begin{aligned} \frac{SQ_t}{\sigma^2} &\sim \chi_{(I-1)}^2 \\ \frac{SQ_{bl}}{\sigma^2} &\sim \chi_{(J-1)}^2 \\ \frac{SQ_{\text{Res}}}{\sigma^2} &\sim \chi_{((I-1)(J-1))}^2 \end{aligned}$$

pois a soma dos graus de liberdade de SQ_t , SQ_{bl} e SQ_{Res} é igual a

$$(I - 1) + (J - 1) + (I - 1)(J - 1) = IJ - 1,$$

que é o número de graus de liberdade de SQ_{T_c} . Então, $\frac{SQ_t}{\sigma^2}$, $\frac{SQ_{bl}}{\sigma^2}$ e $\frac{SQ_{Res}}{\sigma^2}$ são variáveis aleatórias independentes com distribuição qui-quadrado central. Pela definição da distribuição F de Snedecor central e sob H_0 para o efeito de tratamentos, conclui-se que

$$F_c = \frac{SQ_t}{\sigma^2(I - 1)} \cdot \left[\frac{SQ_{Res}}{\sigma^2(I - 1)(J - 1)} \right]^{-1}$$

que é

$$F_c = \frac{QM_t}{QM_{Res}}. \quad (2.33)$$

Analogamente, sob H_0 para o efeito de blocos, conclui-se que

$$F_c = \frac{SQ_{bl}}{\sigma^2(J - 1)} \cdot \left[\frac{SQ_{Res}}{\sigma^2(I - 1)(J - 1)} \right]^{-1}$$

que é

$$F_c = \frac{QM_{bl}}{QM_{Res}}. \quad (2.34)$$

As somas de quadrados (2.21), (2.24) e (2.25) podem ser calculadas por

expressões equivalentes e computacionalmente mais simples, que são

$$SQ_{T_c} = \sum_{i=1}^I \sum_{j=1}^J Y_{ij}^2 - \frac{\left(\sum_{j=1}^J \sum_{i=1}^I Y_{ij} \right)^2}{IJ} \quad (2.35)$$

$$SQ_t = \frac{\sum_{i=1}^I \left(\sum_{j=1}^J Y_{ij} \right)^2}{J} - \frac{\left(\sum_{i=1}^I \sum_{j=1}^J Y_{ij} \right)^2}{IJ} \quad (2.36)$$

$$SQ_{bl} = \frac{\sum_{j=1}^J \left(\sum_{i=1}^I Y_{ij} \right)^2}{I} - \frac{\left(\sum_{i=1}^I \sum_{j=1}^J Y_{ij} \right)^2}{IJ} \quad (2.37)$$

Uma vez calculadas as somas de quadrados de tratamentos, blocos e total, a soma de quadrados de resíduos é obtida pela subtração indicada na expressão (2.38), que é

$$SQ_{Res} = SQ_{T_c} - SQ_t - SQ_{bl}. \quad (2.38)$$

Em geral, a tabela da análise de variância que sintetiza os testes de hipóteses

$$H_0 : \tau_1 = \dots = \tau_I = 0$$

$$H_1 : \tau_i \neq 0 \quad \text{para pelo menos um } i$$

e

$$H_0 : b_1 = \dots = b_J = 0$$

$$H_1 : b_j \neq 0 \quad \text{para pelo menos um } j$$

é conforme a Tabela 6.

Tabela 6 Esquema da análise de variância para experimentos no delineamento em blocos casualizados, com I tratamentos e J blocos.

Fonte de Variação	GL	SQ	QM	F
Tratamentos	$I - 1$	SQ_t	QM_t	$\frac{QM_t}{QM_{Res}}$
Blocos	$J - 1$	SQ_{bl}	QM_{bl}	$\frac{QM_{bl}}{QM_{Res}}$
Resíduo	$(I - 1)(J - 1)$	SQ_{Res}	QM_{Res}	
Total	$IJ - 1$	SQ_{T_c}		

Detalhes adicionais do desenvolvimento da análise de variância, tal como a demonstração de alguns teoremas, são apresentados em Irwin (1931).

Segundo Casella e Berger (2010), os testes esquematizados na Tabela 6 são, em muitos casos, desinteressantes e não verdadeiros. Em geral, o pesquisador não acredita que os efeitos dos diferentes tratamentos sejam estatisticamente iguais. No entanto, é possível estabelecer testes de maior interesse prático por meio de contrastes ortogonais.

Qualquer combinação linear das médias dos tratamentos de um experimento, com a soma dos coeficientes iguais a zero, é chamada de contraste. Matematicamente,

$$C = \sum_{i=1}^I c_i \bar{Y}_i, \quad \text{com} \quad \sum_{i=1}^I c_i = 0$$

em que,

C é um contraste,

c_i é um número real,

\bar{Y}_i é a média do tratamento i , $i = 1, \dots, I$.

Como $Y_{ij} \sim N(\mu + \tau_i + b_j, \sigma^2)$, então, para um contraste C , segue que

$$C \sim N(\mu_C, \sigma_C^2)$$

sendo,

$$\mu_C = \sum_{i=1}^I c_i (\mu + \tau_i) \text{ e}$$

$$\sigma_C^2 = \frac{\sigma^2}{J} \sum_{i=1}^I c_i^2$$

pois,

$$\begin{aligned} E(C) &= E\left(\sum_{i=1}^I c_i \bar{Y}_i\right) = \frac{\sum_{i=1}^I c_i \left(\sum_{j=1}^J E(Y_{ij})\right)}{J} = \\ &= \frac{\sum_{i=1}^I c_i \left[\sum_{j=1}^J (\mu + \tau_i + b_j)\right]}{J} = \frac{J \sum_{i=1}^I c_i (\mu + \tau_i)}{J} = \sum_{i=1}^I c_i (\mu + \tau_i) \end{aligned}$$

e

$$\begin{aligned} Var(C) &= Var\left(\sum_{i=1}^I c_i \bar{Y}_i\right) = \frac{\sum_{i=1}^I c_i^2 \left(\sum_{j=1}^J Var(Y_{ij})\right)}{J^2} = \\ &= \frac{\sum_{i=1}^I c_i^2 \left[\sum_{j=1}^J \sigma^2\right]}{J^2} = \frac{J \sum_{i=1}^I c_i^2 \sigma^2}{J^2} = \frac{\sum_{i=1}^I c_i^2 \sigma^2}{J}. \end{aligned}$$

Supondo que deseja-se testar $H_0 : \mu_C = 0$. Então, sob esta hipótese, tem-se que o contraste C segue distribuição normal de média zero e variância σ_C^2 .

Logo,

$$\frac{C}{\sqrt{\frac{\sigma^2}{J} \sum_{i=1}^I c_i^2}} \sim N(0, 1)$$

e, pela definição de distribuição qui-quadrado central,

$$\frac{C^2}{\frac{\sigma^2}{J} \sum_{i=1}^I c_i^2} \sim \chi_{(1)}^2.$$

O resultado anterior é útil, pois a soma dos quadrados correspondente aos tratamentos em qualquer análise de variância pode ser particionada em componentes de somas de quadrados, cada uma com um grau de liberdade. Esta partição corresponde a contrastes mutuamente ortogonais (MASON; GUNST; HESS, 2003).

Dois contrastes

$$C_1 = \sum_{i=1}^I a_i \bar{Y}_i, \quad C_2 = \sum_{i=1}^I b_i \bar{Y}_i.$$

são ditos ortogonais se a soma dos produtos dos coeficientes correspondentes aos dois contrastes for igual a zero, isto é,

$$\sum_{i=1}^I a_i b_i = a_1 b_1 + a_2 b_2 + \cdots + a_I b_I = 0.$$

Três ou mais contrastes são ditos mutuamente ortogonais se todos os pares deles resultantes forem ortogonais.

Considerando um experimento com 4 tratamentos, os contrastes

$$C_1 = 3\bar{Y}_1. + 2\bar{Y}_2. - \bar{Y}_3. - 4\bar{Y}_4., \quad C_2 = 4\bar{Y}_1. - 6\bar{Y}_2., \quad C_3 = 8\bar{Y}_3. - 2\bar{Y}_4.$$

são mutuamente ortogonais, pois a soma dos produtos dos coeficientes de cada par

de contrastes é igual a zero, ou seja, para os pares de contrastes

$$C_1 \text{ e } C_2 \text{ tem-se } 3 \cdot 4 + 2 \cdot (-6) + (-1) \cdot 0 + (-4) \cdot 0 = 0;$$

$$C_1 \text{ e } C_3 \text{ tem-se } 3 \cdot 0 + 2 \cdot 0 + (-1) \cdot 8 + (-4) \cdot (-2) = 0;$$

$$C_2 \text{ e } C_3 \text{ tem-se } 4 \cdot 0 + (-6) \cdot 0 + 0 \cdot 8 + 0 \cdot (-2) = 0.$$

Para um conjunto de médias, cada uma com J observações, a soma de quadrados correspondente a um contraste $C = \sum_{i=1}^I c_i \bar{Y}_i$. é dada por

$$SQC = \frac{J \left(\sum_{i=1}^I c_i \bar{Y}_i \right)^2}{\sum_{i=1}^I c_i^2}. \quad (2.39)$$

Em relação à decomposição da soma de quadrados de tratamentos em soma de quadrados de contrastes ortogonais, Casella e Berger (2010) estabelecem que é possível encontrar conjuntos com $I - 1$ vetores de constantes $\mathbf{c}^{(l)} = (c_1^{(l)}, \dots, c_I^{(l)})$, $l = 1, \dots, I - 1$, que satisfazem

$$J \sum_{i=1}^I (\bar{Y}_i - \bar{Y}_{..})^2 = \frac{J \left(\sum_{i=1}^I c_i^{(1)} \bar{Y}_i \right)^2}{\sum_{i=1}^I (c_i^{(1)})^2} + \dots + \frac{J \left(\sum_{i=1}^I c_i^{(I-1)} \bar{Y}_i \right)^2}{\sum_{i=1}^I (c_i^{(I-1)})^2} \quad (2.40)$$

e

$$\sum_{i=1}^I c_i^{(l)} c_i^{(l')} = 0 \quad \text{para todo } l \neq l'.$$

O quociente do lado esquerdo da expressão (2.40) por σ^2 tem distribuição qui-quadrado com $I - 1$ graus de liberdade e o quociente de cada termo do lado direito por σ^2 tem distribuição qui-quadrado de 1 grau de liberdade cada.

Em experimentos fatoriais 2^k , a construção de contrastes mutuamente ortogonais relativos aos efeitos principais e de interações pode ser realizada através de uma tabela de sinais positivos e negativos.

Com o intuito de exemplificar a construção de uma tabela de sinais positivos e negativos, considere-se um experimento em esquema fatorial 2^3 , com J blocos. A tabela de sinais pode ser inicialmente construída como a Tabela 7, e completada através dos itens seguintes até obter a Tabela 8.

Tabela 7 Esquema inicial de uma tabela de sinais para um experimento fatorial 2^3 contendo os tratamentos e os efeitos fatoriais a serem avaliados.

Tratamentos	Efeitos Fatoriais						
	<i>A</i>	<i>B</i>	<i>C</i>	<i>AB</i>	<i>AC</i>	<i>BC</i>	<i>ABC</i>
(1)							
<i>a</i>							
<i>b</i>							
<i>ab</i>							
<i>c</i>							
<i>ac</i>							
<i>bc</i>							
<i>abc</i>							

- (i) atribuir sinais para os fatores de cada tratamento, de tal forma que, se o fator estiver no nível baixo, receberá sinal negativo, caso contrário o positivo;
- (ii) as colunas referente aos efeitos de interação são obtidas pela multiplicação (linha por linha) das colunas de efeitos principais correspondentes;
- (iii) acrescentar uma última coluna com as médias dos tratamentos.

Tabela 8 Sinais para calcular os efeitos de um experimento fatorial 2^3 contendo os tratamentos, os efeitos fatoriais a serem avaliados e as médias dos tratamentos.

Tratamentos	Efeitos Fatoriais							Médias
	A	B	C	AB	AC	BC	ABC	
(1)	-	-	-	+	+	+	-	$\bar{Y}_{1.} = \bar{Y}_{(1)}$
a	+	-	-	-	-	+	+	$\bar{Y}_{2.} = \bar{Y}_a$
b	-	+	-	-	+	-	+	$\bar{Y}_{3.} = \bar{Y}_b$
ab	+	+	-	+	-	-	-	$\bar{Y}_{4.} = \bar{Y}_{ab}$
c	-	-	+	+	-	-	+	$\bar{Y}_{5.} = \bar{Y}_c$
ac	+	-	+	-	+	-	-	$\bar{Y}_{6.} = \bar{Y}_{ac}$
bc	-	+	+	-	-	+	-	$\bar{Y}_{7.} = \bar{Y}_{bc}$
abc	+	+	+	+	+	+	+	$\bar{Y}_{8.} = \bar{Y}_{abc}$

Os contrastes mutuamente ortogonais referentes aos efeitos principais e de interações, são obtidos pelo quociente do somatório dos elementos resultantes da multiplicação (linha por linha) da coluna de cada efeito fatorial pela coluna de médias, pela metade da quantidade de médias consideradas no contraste. Assim, para o experimento fatorial 2^3 , considerado anteriormente, tem-se os seguintes contrastes mutuamente ortogonais referentes aos efeitos principais e de interações.

$$C_A = \frac{1}{4} (\bar{Y}_{2.} + \bar{Y}_{4.} + \bar{Y}_{6.} + \bar{Y}_{8.} - \bar{Y}_{1.} - \bar{Y}_{3.} - \bar{Y}_{5.} - \bar{Y}_{7.}),$$

$$C_B = \frac{1}{4} (\bar{Y}_{3.} + \bar{Y}_{4.} + \bar{Y}_{7.} + \bar{Y}_{8.} - \bar{Y}_{1.} - \bar{Y}_{2.} - \bar{Y}_{5.} - \bar{Y}_{6.}),$$

$$C_C = \frac{1}{4} (\bar{Y}_{5.} + \bar{Y}_{6.} + \bar{Y}_{7.} + \bar{Y}_{8.} - \bar{Y}_{1.} - \bar{Y}_{2.} - \bar{Y}_{3.} - \bar{Y}_{4.}),$$

$$C_{AB} = \frac{1}{4} (\bar{Y}_{1.} + \bar{Y}_{4.} + \bar{Y}_{5.} + \bar{Y}_{8.} - \bar{Y}_{2.} - \bar{Y}_{3.} - \bar{Y}_{6.} - \bar{Y}_{7.}),$$

$$C_{AC} = \frac{1}{4} (\bar{Y}_{1.} + \bar{Y}_{3.} + \bar{Y}_{6.} + \bar{Y}_{8.} - \bar{Y}_{2.} - \bar{Y}_{4.} - \bar{Y}_{5.} - \bar{Y}_{7.}),$$

$$C_{BC} = \frac{1}{4} (\bar{Y}_1. + \bar{Y}_2. + \bar{Y}_7. + \bar{Y}_8. - \bar{Y}_3. - \bar{Y}_4. - \bar{Y}_5. - \bar{Y}_6.)$$

e

$$C_{ABC} = \frac{1}{4} (\bar{Y}_2. + \bar{Y}_3. + \bar{Y}_5. + \bar{Y}_8. - \bar{Y}_1. - \bar{Y}_4. - \bar{Y}_6. - \bar{Y}_7.).$$

Em geral, há interesse em verificar se os efeitos de contrastes mutuamente ortogonais são estatisticamente iguais a zero. Para os contrastes definidos acima é possível estabelecer as hipóteses

$$H_0 : \mu_{C_A} = 0 \text{ vs } H_1 : \mu_{C_A} \neq 0, \quad H_0 : \mu_{C_B} = 0 \text{ vs } H_1 : \mu_{C_B} \neq 0,$$

$$H_0 : \mu_{C_C} = 0 \text{ vs } H_1 : \mu_{C_C} \neq 0, \quad H_0 : \mu_{C_{AB}} = 0 \text{ vs } H_1 : \mu_{C_{AB}} \neq 0,$$

$$H_0 : \mu_{C_{AC}} = 0 \text{ vs } H_1 : \mu_{C_{AC}} \neq 0, \quad H_0 : \mu_{C_{BC}} = 0 \text{ vs } H_1 : \mu_{C_{BC}} \neq 0$$

e

$$H_0 : \mu_{C_{ABC}} = 0 \text{ vs } H_1 : \mu_{C_{ABC}} \neq 0$$

Essas hipóteses podem ser testadas pelo teste F da análise de variância, esquematizado na Tabela 9. Pois, para um contraste $C = \sum_{i=1}^I c_i \bar{Y}_i.$ e sob $H_0 : \mu_C = 0$,

$$\frac{SQC}{\sigma^2} = \frac{1}{\sigma^2} \cdot \frac{J \left(\sum_{i=1}^I c_i \bar{Y}_i. \right)^2}{\sum_{i=1}^I c_i^2} \sim \chi_{(1)}^2.$$

Portanto, pela definição de distribuição F de Snedecor central, uma estatística de teste para as hipóteses $H_0 : \mu_C = 0$ e $H_1 : \mu_C \neq 0$, é dada por

$$F_c = \frac{QMC}{QM_{\text{Res}}},$$

que sob H_0 , tem-se

$$\frac{QMC}{QM_{\text{Res}}} \sim F_{(1, (2^k - 1)(J - 1))},$$

sendo $QMC = \frac{SQC}{1}$ o quadrado médio do contraste C .

O nível descritivo (valor - p) para o teste supracitado é dado pela probabilidade:

$$\text{valor - p} = P \left(F_{(1, (2^k - 1)(J - 1))} > \frac{QMC}{QM_{\text{Res}}} \right) = P(F > F_c).$$

Rejeita-se a hipótese H_0 se valor - p < α , sendo α o nível de significância adotado.

Tabela 9 Esquema da análise de variância de um experimento em esquema fatorial 2^3 , com J blocos.

Fonte de Variação	GL	SQ	QM	F
A	1	SQC_A	QMC_A	$\frac{QMC_A}{QM_{\text{Res}}}$
B	1	SQC_B	QMC_B	$\frac{QMC_B}{QM_{\text{Res}}}$
C	1	SQC_C	QMC_C	$\frac{QMC_C}{QM_{\text{Res}}}$
AB	1	SQC_{AB}	QMC_{AB}	$\frac{QMC_{AB}}{QM_{\text{Res}}}$
AC	1	SQC_{AC}	QMC_{AC}	$\frac{QMC_{AC}}{QM_{\text{Res}}}$
BC	1	SQC_{BC}	QMC_{BC}	$\frac{QMC_{BC}}{QM_{\text{Res}}}$
ABC	1	SQC_{ABC}	QMC_{ABC}	$\frac{QMC_{ABC}}{QM_{\text{Res}}}$
Blocos	$J - 1$	SQ_{bl}	QM_{bl}	$\frac{QM_{bl}}{QM_{\text{Res}}}$
Resíduo	$(2^3 - 1)(J - 1)$	SQ_{Res}	QM_{Res}	
Total	$J2^3 - 1$	SQ_{T_c}		

A soma de quadrados dos resíduos, SQ_{Res} , pode ser obtida pela diferença

entre a SQ_{T_c} e as demais somas de quadrados. Assim, para o experimento fatorial 2^3 com J blocos,

$$SQ_{\text{Res}} = SQ_{T_c} - SQ_{C_A} - \dots - SQ_{C_{ABC}} - SQ_{bl}.$$

Em geral, a tabela da análise de variância para um delineamento em blocos casualizados em esquema fatorial 2^k é conforme a Tabela 5.

Na análise de variância de experimentos fatoriais, quando o efeito de alguma interação é significativo, recomenda-se estudar o desdobramento dessa interação, ou seja, analisar o efeito de um dos fatores dentro dos níveis dos demais. Como é raro, na prática, efeitos de interações de quatro ou mais fatores serem significativos, será abordado apenas o desdobramento das interações de ordem dois e três, que pode ser, sem grandes dificuldades, generalizado para interações de qualquer ordem.

Retomando o experimento fatorial 2^3 com J blocos, e supondo que haja interesse em desdobrar o efeito da interação AC , as fontes de variações A e AC podem ser decompostas em duas novas fontes de variações: $A:C_-$ e $A:C_+$. A primeira refere-se ao efeito do fator A dentro do nível baixo do fator C , e a segunda ao efeito do fator A dentro do nível alto do fator C .

Alternativamente ao desdobramento da interação AC em efeitos do fator A dentro dos níveis do fator C , as fontes de variações C e AC também podem ser decompostas em duas novas fontes de variações: $C:A_-$ e $C:A_+$, referente aos efeitos do fator C dentro dos níveis do fator A .

Os contrastes relativos aos efeitos resultantes dos desdobramentos das interações também podem ser obtidos com o auxílio da tabela de sinais positivos e negativos, que para o experimento fatorial 2^3 está apresentada na Tabela 8.

Para o cálculo do contraste referente ao efeito do fator A dentro do nível

baixo do fator C , $A:C_-$, considera-se, na tabela de sinais, apenas as linhas em que os sinais da coluna do efeito fatorial C são iguais a “-”, e calcula-se o quociente do somatório dos elementos resultantes da multiplicação (linha por linha) da coluna restrita do efeito fatorial A pela coluna restrita de médias, pela metade da quantidade de médias consideradas no contraste.

$$C_{A:C_-} = \frac{1}{2} (\bar{Y}_a + \bar{Y}_{ab} - \bar{Y}_{(1)} - \bar{Y}_b) = \frac{1}{2} (\bar{Y}_2 + \bar{Y}_4 - \bar{Y}_1 - \bar{Y}_3)$$

O contraste referente ao efeito do fator A dentro do nível alto do fator C , $A:C_+$, é obtido de forma semelhante ao caso anterior. Sendo que agora, na tabela de sinais, considera-se apenas as linhas em que os sinais da coluna do efeito fatorial C são iguais a “+”.

$$C_{A:C_+} = \frac{1}{2} (\bar{Y}_{ac} + \bar{Y}_{abc} - \bar{Y}_c - \bar{Y}_{bc})$$

Para o desdobramento da interação AC em efeitos do fator C dentro dos níveis do fator A tem-se os seguintes contrastes:

$$C_{C:A_-} = \frac{1}{2} (\bar{Y}_c + \bar{Y}_{bc} - \bar{Y}_{(1)} - \bar{Y}_b)$$

$$C_{C:A_+} = \frac{1}{2} (\bar{Y}_{ac} + \bar{Y}_{abc} - \bar{Y}_a - \bar{Y}_{ab})$$

Supondo, agora, que o interesse seja desdobrar o efeito da interação ABC , que pode ser realizado de três formas: (i) em efeitos do fator A em cada nível dos fatores B e C , (ii) em efeitos do fator B em cada nível dos fatores A e C , e (iii) em efeitos do fator C em cada nível dos fatores A e B .

Para o desdobramento do efeito da interação ABC em efeitos do fator A em cada nível dos fatores B e C , as fontes de variações A , AB , AC e ABC são

decompostas em quatro novas fontes de variações: $A:B_-C_-$, $A:B_+C_-$, $A:B_-C_+$ e $A:B_+C_+$, que se referem aos efeitos do fator A em cada nível dos fatores B e C .

Em particular, para o cálculo do contraste referente ao efeito $A:B_+C_-$, considera-se, na tabela de sinais, apenas as linhas em que os sinais da coluna do efeito fatorial B são iguais a “+” e os sinais da coluna do efeito fatorial C são iguais a “-”, e calcula-se o quociente do somatório dos elementos resultantes da multiplicação (linha por linha) da coluna restrita do efeito fatorial A pela coluna restrita de médias, pela metade da quantidade de médias consideradas no contraste.

$$C_{A:B_+C_-} = \frac{1}{1} (\bar{Y}_{ab} - \bar{Y}_b) = \bar{Y}_{ab} - \bar{Y}_b$$

As somas de quadrados necessárias para os testes F da análise de variância, no estudo de desdobramento de interações, são calculadas normalmente a partir da expressão (2.39). Geralmente, a análise de variância é esquematizada em uma tabela contendo as somas de quadrados, quadrados médios e os valores F_c dos testes F de cada fonte de variação do desdobramento da interação. As Tabelas 10 e 11 esquematizam a análise de variância para o estudo do desdobramento da interação AC .

Tabela 10 Esquema da análise de variância do desdobramento da interação AC em efeitos do fator A em cada nível do fator C .

Fonte de variação	GL	SQ	QM	F_c
$A:C_-$	1	$SQ_{A:C_-}$	$QM_{A:C_-}$	$\frac{QM_{A:C_-}}{QM_{Res}}$
$A:C_+$	1	$SQ_{A:C_+}$	$QM_{A:C_+}$	$\frac{QM_{A:C_+}}{QM_{Res}}$
Resíduo	$(2^3 - 1)(J - 1)$	SQ_{Res}	QM_{Res}	

Tabela 11 Esquema da análise de variância do desdobramento da interação AC em efeitos do fator C em cada nível do fator A .

Fonte de variação	GL	SQ	QM	F_c
$C:A_-$	1	$SQ_{C:A_-}$	$QM_{C:A_-}$	$\frac{QM_{C:A_-}}{QM_{Res}}$
$C:A_+$	1	$SQ_{C:A_+}$	$QM_{C:A_+}$	$\frac{QM_{C:A_+}}{QM_{Res}}$
Resíduo	$(2^3 - 1)(J - 1)$	SQ_{Res}	QM_{Res}	

2.2.3 Experimento fatorial 2^k não repetido

Em experimentos fatoriais 2^k , o número de tratamentos cresce exponencialmente à medida que aumenta-se o número de fatores considerados. Por exemplo, em um experimento 2^5 tem-se 32 tratamentos; sendo que, com duas repetições haverá necessidade de 64 unidades experimentais para realizar o experimento, que é uma quantidade elevada para determinado tipo de pesquisa. Por isso, algumas vezes o experimento pode ser muito caro, ou até mesmo inviável de ser conduzido quando o número de fatores é elevado.

Conforme Montgomery (2009), os recursos disponíveis para experimentos são, geralmente, escassos, de tal forma que, quando o número de fatores é excessivo, é inviável realizar experimentos com repetições, a menos que o pesquisador esteja disposto a desconsiderar alguns fatores originais.

Experimentos fatoriais com uma só observação são denominados de fatoriais não repetido. Uma abordagem para analisar esses experimentos, fatoriais não repetido, consiste em considerar o princípio da escassez de efeitos, que é a suposição de que os efeitos das interações de ordem superiores são não significativos. Uma estimativa para o resíduo é o somatório das somas de quadrados dos efeitos considerados como insignificantes, e o restante da análise é conduzido

normalmente.

De acordo com Box, Hunter e Hunter (2005), ocasionalmente, efeitos de interações de alta ordem podem ser significativos, por isso mostra-se relevante um método preliminar para selecioná-los. Após a identificação desses efeitos, eles são desconsiderados na estimação do resíduo.

Daniel (1959) propõe a utilização do gráfico de probabilidade semi-normal do módulo das estimativas dos efeitos, para verificá-los quanto a sua significância, em experimento fatorial 2^k não repetido. No entanto, Daniel (1976) e Box e Meyer (1986) admitiram que o gráfico de probabilidade normal é mais útil do que o semi-normal, pois, este último é menos capaz de revelar discrepância de pressuposições, tal como a homocedasticidade.

O gráfico de probabilidade normal pode ser construído com a utilização de *softwares* estatísticos, tal como R (R Development Core Team, 2012).

Suponha um experimento fatorial 2^4 não repetido, com as estimativas dos efeitos fatoriais apresentadas na Tabela 12 e o gráfico de probabilidade normal na Figura 3.

Tabela 12 Estimativa dos efeitos fatoriais de um experimento fatorial 2^4 não repetido.

Efeito	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>AB</i>
Estimativa	-4,28	0,14	5,7	3,01	0,43
Efeito	<i>AC</i>	<i>AD</i>	<i>BC</i>	<i>BD</i>	<i>CD</i>
Estimativa	-0,61	-0,26	0,18	0,68	3,90
Efeito	<i>ABC</i>	<i>ABD</i>	<i>ACD</i>	<i>BCD</i>	<i>ABCD</i>
Estimativa	-0,17	0,16	-0,15	-0,38	-0,11

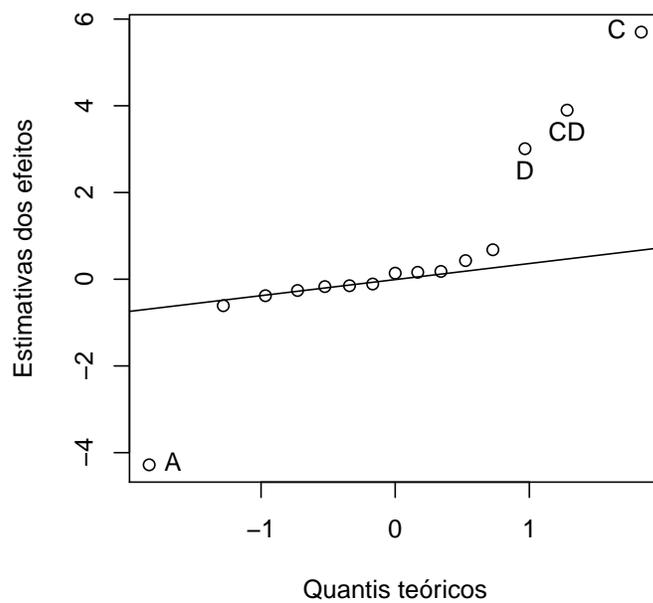


Figura 3 Gráfico de probabilidade normal para os efeitos fatoriais de um experimento fatorial 2^4

Analisando o gráfico de probabilidade normal das estimativas dos efeitos fatoriais, Figura 3, nota-se que os efeitos A , C , D e CD podem ser significativos, uma vez que se localizam distantes da reta que passa pelos outros pontos. Sendo assim, a soma de quadrados de resíduos pode ser obtida pelo somatório das somas de quadrados dos efeitos das interações de três e quatro fatores. O esquema da análise de variância para esse exemplo fictício é apresentado na Tabela 13.

$$SQ_{\text{Res}} = SQ_{ABC} + SQ_{ABD} + SQ_{ACD} + SQ_{BCD} + SQ_{ABCD}$$

ou

$$\begin{aligned}
 SQ_{\text{Res}} &= SQ_{T_c} - SQ_A - SQ_B - SQ_C - SQ_D - SQ_{AB} \\
 &\quad - SQ_{AC} - SQ_{AD} - SQ_{BC} - SQ_{BD} - SQ_{CD}
 \end{aligned}$$

Tabela 13 Esquema da análise de variância de um experimento 2^4 não repetido, para a verificação de efeitos principais e de interação de ordem dois.

Fonte de Variação	GL
<i>A</i>	1
<i>B</i>	1
<i>C</i>	1
<i>D</i>	1
<i>AB</i>	1
<i>AC</i>	1
<i>AD</i>	1
<i>BC</i>	1
<i>BD</i>	1
<i>CD</i>	1
Resíduo	5
Total	15

A análise preliminar de um experimento fatorial não repetido não é realizada apenas pelo gráfico de probabilidade normal, há outros métodos menos triviais, Hamada e Balakrishnan (1998) analisam e discutem alguns desses métodos.

2.2.4 Confundimento

Há experimentos em que é impossível realizar uma repetição completa em blocos, pois a dimensão do bloco é menor do que o número total de tratamen-

tos (MONTGOMERY, 2009). Uma alternativa para essa situação é a técnica do confundimento. Essa técnica faz com que a informação de alguns efeitos do experimento, em geral as interações de ordem superiores, sejam indistinguíveis dos efeitos de blocos.

O confundimento em blocos de um experimento fatorial 2^k pode ser realizado em 2^p blocos em cada repetição, com $p < k$, porém, o confundimento em 2^1 blocos apresenta duas vantagens em relação aos demais. A primeira é que o confundimento em 2^1 blocos em cada repetição acarreta menor prejuízo de informação, já que apenas um efeito será confundido. A outra vantagem é em relação a facilidade do planejamento.

No planejamento de um experimento 2^k confundido em 2^1 blocos, é preciso definir quais tratamentos serão designados para cada bloco. A atribuição dos tratamentos a cada um dos dois blocos dependerá de qual efeito será confundido com o efeito de blocos. Assim, um dos procedimentos para estabelecer quais tratamentos serão considerados para cada bloco consiste em:

- (i) construir a tabela de sinais, ou de efeitos fatoriais;
- (ii) definir qual efeito será confundido com o efeito de blocos;
- (iii) considerando a coluna do efeito definido no item anterior, atribuir a um dos blocos os tratamentos com sinais “+” e ao outro os tratamentos com sinais “-”.

Para ilustrar o procedimento acima, considere-se um experimento 2^3 , ou seja, com três fatores (A , B e C) de dois níveis (“+” e “-”), em que o efeito da interação $A \times B \times C$ será confundido com o efeito de blocos. Pela Tabela 8, verifica-se que os tratamentos atribuídos ao bloco 1 são a , b , c e abc e, ao bloco 2 são (1) , ab , ac e bc . A Tabela 14 contém o esquema dos tratamentos em cada

bloco.

Tabela 14 Constituição dos blocos de um experimento 2^3 para o confundimento da interação tripla.

Bloco 1 (+)	Bloco 2 (-)
<i>a</i>	(1)
<i>b</i>	<i>ab</i>
<i>c</i>	<i>ac</i>
<i>abc</i>	<i>bc</i>

Observa-se que o efeito da interação $A \times B \times C$, $C_{ABC} = \bar{Y}_2 + \bar{Y}_3 + \bar{Y}_5 + \bar{Y}_8 - \bar{Y}_1 - \bar{Y}_4 - \bar{Y}_6 - \bar{Y}_7$, é o mesmo do efeito de blocos. Desta forma, não é possível distinguir a causa do efeito, se é devido à interação ou aos blocos. No entanto, os outros efeitos não são confundidos, pois nos contrastes dos demais efeitos há uma combinação de um tratamento positivo e um negativo de cada bloco. Por exemplo, para o cálculo do efeito principal do fator A , é feita a diferença entre as médias dos tratamentos a , ab , ac e abc com as médias dos tratamentos (1), b , c e bc , no entanto o efeito de blocos é anulado porque os tratamentos a , b , c e abc estão no bloco 1 e os tratamentos (1), ab , ac e bc estão no bloco 2.

Banzatto e Kronka (2008) apresentam esquemas de análise de variância, Tabela 15, com e sem confundimento da interação tripla de um experimento fatorial 2^3 conduzido em blocos casualizados com 4 repetições. Note-se que, de 4 blocos com 8 unidades experimentais cada, passa-se para 8 blocos com 4 unidades experimentais. Verifica-se, também, que o uso do confundimento acarreta em diminuição do número de graus de liberdade do resíduo.

Tabela 15 Esquema de análise de variância com e sem confundimento da interação tripla de um experimento 2^3 .

Sem confundimento		Com confundimento	
Fonte de variação	GL	Fonte de variação	GL
A	1	A	1
B	1	B	1
C	1	C	1
$A \times B$	1	$A \times B$	1
$A \times C$	1	$A \times C$	1
$B \times C$	1	$B \times C$	1
$A \times B \times C$	1	Blocos	7
Blocos	3	Resíduo	18
Resíduo	21	Total	31
Total	31		

Montgomery (2009) descreve outro método de obter quais tratamentos serão designados para cada bloco. O método utiliza a combinação linear

$$L = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_k x_k, \quad (2.41)$$

em que, x_i é o nível do i -ésimo fator que aparece em uma combinação de tratamentos particular e α_i é o nível do i -ésimo fator para qual o efeito será confundido. Para experimentos 2^k tem-se $\alpha_i = 1$ ou 0 e $x_i = 1$ (nível alto) ou $x_i = 0$ (nível baixo). O método consiste em atribuir ao mesmo bloco os tratamentos que produzam o mesmo valor de $L \pmod{2}$.

Considerando o experimento 2^3 com o efeito da interação tripla confundida com o efeito de blocos, tem-se $\alpha_1 = \alpha_2 = \alpha_3 = 1$, logo

$$L = x_1 + x_2 + x_3$$

e, para os tratamentos,

$$\begin{aligned} (1) : L &= 0 + 0 + 0 = 0 \equiv 0 \pmod{2}; & a : L &= 1 + 0 + 0 = 1 \equiv 1 \pmod{2} \\ b : L &= 0 + 1 + 0 = 1 \equiv 1 \pmod{2}; & c : L &= 0 + 0 + 1 = 1 \equiv 1 \pmod{2} \\ ab : L &= 1 + 1 + 0 = 2 \equiv 0 \pmod{2}; & ac : L &= 1 + 0 + 1 = 2 \equiv 0 \pmod{2} \\ bc : L &= 0 + 1 + 1 = 2 \equiv 0 \pmod{2}; & abc : L &= 1 + 1 + 1 = 3 \equiv 1 \pmod{2} \end{aligned}$$

Portanto, os tratamentos a , b , c e abc ocorrem no bloco 1 e os demais no bloco 2, que é o mesmo resultado obtido ao usar a tabela de sinais. A vantagem desse último método é ser aplicável em confundimentos de 2^p blocos, com $p \geq 1$.

No planejamento de um experimento fatorial 2^k confundido em 2^p blocos, com 2^{k-p} tratamentos em cada bloco, selecionam-se p efeitos independentes para que seus efeitos sejam confundidos com o efeito de blocos. Entende-se por efeitos independentes quando nenhum deles é a interação de alguns dos demais. Um método para verificar se um conjunto de efeitos são independentes, é quando, na tabela de sinais, nenhum deles é o resultado do produto das colunas de alguns dos demais. Por exemplo, os efeitos AB , AC e BC não são independentes, pois na Tabela 8 verifica-se que BC é resultado da multiplicação das colunas dos efeitos AB e AC .

Além dos p efeitos escolhidos inicialmente confundir-se-á $2^p - p - 1$ efeitos de interações generalizadas. Por isso, é preciso ter cuidado ao selecionar os efeitos que serão confundidos para não sacrificar informações de interesse (MONTGOMERY, 2009).

Os tratamentos que serão designados para cada um dos p blocos são obtidos através de p combinações lineares L_1, \dots, L_p , tais como a da expressão (2.41). Com as p combinações lineares formam-se 2^p p -uplas e todos os 2^{k-p}

tratamentos de cada bloco produzem os mesmos valores $L_1 \equiv (\text{mod}2), \dots, L_p \equiv (\text{mod}2)$.

Como exemplo, considere-se um experimento 2^5 confundido em $2^{5-2} = 2^3 = 8$ blocos. Escolhendo os efeitos independentes ABE , BCE e CDE , verifica-se que os efeitos AC , $ABCD$, BD e ADE , das interações generalizadas, também serão confundidos. Os tratamentos de cada um dos 8 blocos são obtidos através das seguintes combinações lineares

$$L_{ABE} = x_1 + x_2 + x_5$$

$$L_{BCE} = x_2 + x_3 + x_5$$

$$L_{CDE} = x_3 + x_4 + x_5$$

Considerando a aritmética modular, cada tratamento produzirá uma tripla de valores $L_{ABE} \equiv (\text{mod}2)$, $L_{BCE} \equiv (\text{mod}2)$ e $L_{CDE} \equiv (\text{mod}2)$, sendo que os que resultarem na mesma tripla serão considerados para o mesmo bloco. Os tratamentos de cada bloco estão especificados na Tabela 16.

Tabela 16 Tratamentos dos blocos de cada repetição de um experimento fatorial 2^5 para o confundimento dos efeitos ABE , BCE , CDE , AC , $ABCD$, BD e ADE com o efeito de blocos.

Bloco 1	Bloco 2	Bloco 3	Bloco 4
(0, 0, 0)	(1, 0, 0)	(0, 1, 0)	(0, 0, 1)
(1)	<i>a</i>	<i>ab</i>	<i>d</i>
<i>ace</i>	<i>ce</i>	<i>cd</i>	<i>be</i>
<i>bde</i>	<i>bcd</i>	<i>ade</i>	<i>abc</i>
<i>abcd</i>	<i>abde</i>	<i>bce</i>	<i>acde</i>
Bloco 5	Bloco 6	Bloco 7	Bloco 8
(1, 0, 1)	(1, 1, 0)	(0, 1, 1)	(1, 1, 1)
<i>ad</i>	<i>b</i>	<i>c</i>	<i>e</i>
<i>bc</i>	<i>de</i>	<i>ae</i>	<i>ac</i>
<i>abe</i>	<i>acd</i>	<i>abd</i>	<i>bd</i>
<i>cde</i>	<i>abce</i>	<i>bcde</i>	<i>abcde</i>

Conforme os autores Hinkelmann e Kempthorne (2008), Garcia-Diaz e Phillips (1995) e Winer (1962), a soma de quadrados de blocos pode ser obtida através da expressão (2.42).

$$SQ_{bl} = \frac{\sum_{j=1}^J \left(\sum_{i_{(j)}=1}^{I_{(j)}} Y_{i_{(j)}j} \right)^2}{I'} - \frac{\left(\sum_{j=1}^J \sum_{i_{(j)}=1}^{I_{(j)}} Y_{i_{(j)}j} \right)^2}{I'J} \quad (2.42)$$

em que,

$Y_{i_{(j)}j}$ é a observação do $i_{(j)}$ -ésimo tratamento dentro do j -ésimo bloco do experimento;

I' é o total de tratamentos em cada bloco do experimento, ou seja, $I' = 2^{k-p}$;

J é o total de blocos do experimento, ou seja, $J = n2^p$, com n o número de repetições do experimento.

Para experimentos fatoriais não repetidos, a soma de quadrados do efeito de blocos é obtida somando as somas de quadrados de todos os efeitos confundidos com blocos, e o restante da análise é feita normalmente (Montgomery, 2009).

A soma de quadrados de resíduos é obtida pela diferença da soma de quadrados total e as demais somas de quadrados, com exceção daquelas que são referentes aos efeitos confundidos com blocos. Assim, para um experimento fatorial 2^3 com a interação tripla (ABC) confundida com o efeito de blocos, a soma de quadrados de resíduos é obtida através da seguinte expressão:

$$SQ_{\text{Res}} = SQ_{T_c} - SQ_A - SQ_B - SQ_C - SQ_{AB} - SQ_{AC} - SQ_{BC} - SQ_{bl}.$$

Além desta dissertação, Montgomery (2009), Hinkelmann e Kempthorne (2008), Garcia-Diaz e Phillips (1995) e Winer (1962) abordam a técnica do confundimento para experimentos fatoriais 2^k .

3 METODOLOGIA

Divide-se este capítulo em duas seções: materiais e métodos. Na primeira são descritos os recursos utilizados na aplicação das técnicas experimentais apresentadas e, sucintamente, o experimento considerado na ilustração das técnicas experimentais apresentadas no referencial teórico. Na segunda seção são descritos os métodos utilizados nesta dissertação.

3.1 Materiais

Os dados utilizados na ilustração das técnicas estatísticas de planejamento e análise de experimentos são provenientes de Santos (2007), e referem-se ao estudo da resistência mecânica de junções adesivas através de experimentos fatoriais 2^k .

As unidades básicas do experimento são corpos de prova de junções adesivas que, conforme Santos (2007), foram obtidas por sobreposição simples de chapas de aço unidas por adesivo estrutural de base acrílica. A Figura 4 ilustra um corpo de prova, com as suas respectivas dimensões em *mm*.

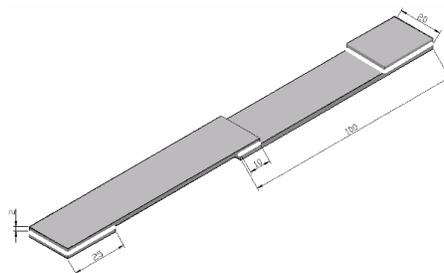


Figura 4 Corpo de prova obtido através de sobreposição simples de chapas de aço (Adaptado de Santos, 2007)

Em relação às chapas de aço utilizadas para a confecção dos corpos de prova, Santos (2007) esclarece que foram fornecidas pela Companhia Siderúrgica de Tubarão, e as especificações são SAE J403 1010 laminado a quente, 2mm de espessura e composição química conforme a Tabela 17.

Tabela 17 Composição química (%) do aço utilizado na confecção dos corpos de prova (Santos, 2007).

<i>C</i>	<i>Si</i>	<i>Mn</i>
0,10	0,021	0,471
<i>P</i>	<i>S</i>	<i>Al</i>
0,016	0,0083	0,041
<i>Cu</i>	<i>Ni</i>	<i>Cr</i>
0,008	0,005	0,017

O adesivo acrílico utilizado no experimento é, segundo Santos (2007), o adesivo Loctite 326 da empresa Henkel, que é considerado pela empresa como adesivo de rápida fixação - cerca de 3 minutos.

Santos (2007) afirma que o ensaio de cisalhamento por tração é o mais utilizado na determinação da resistência de junções adesivas por sobreposição simples de chapas. Também, através de ensaios de cisalhamento, é possível determinar importantes características dos adesivos e junções, como a tensão média de ruptura², que é considerada como a variável resposta do experimento em destaque.

Os ensaios de cisalhamento conduzidos por Santos (2007) foram realizados pelo equipamento EMIC DL5000 com célula de carga de 50kN. Santos

²A tensão média de ruptura é obtida pelo quociente entre a força máxima aplicada e a área de sobreposição das chapas de aço.

(2007) baseou-se em Villenave (2005) para definir os parâmetros de ensaio, tais como a configuração geométrica do corpo de prova, velocidade de tracionamento ($2\text{mm}/\text{min}$) e o comprimento de fixação das chapas nas garras de 25mm .

Santos (2007) considerou os fatores que estão relacionados com a condição de cura do adesivo e de superfície do substrato, no caso as chapas de aço. Os fatores considerados foram tratamento da superfície (A), temperatura e tempo de cura (B), rugosidade da superfície (C) e uso de ativador (D).

Todos os fatores analisados foram considerados em dois níveis, ou seja, em nível baixo e alto. Para o fator A , tratamento da superfície, foram considerados os níveis sem e com tratamento superficial utilizando o produto Bonderite NT1 da empresa Henkel; para o fator B , temperatura e tempo de cura, foram considerados os níveis 22°C por 24 horas e 60°C por 2 horas; para o fator C , rugosidade da superfície, foram considerados os níveis sem e com jateamento de areia; para o fator D , uso de ativador, foram considerados os níveis sem e com ativador Loctite 7649 da empresa Henkel. Os fatores e os seus níveis estão esquematizados e codificados na Tabela 18.

Tabela 18 Fatores e níveis analisados codificados.

Fator	Níveis	
	-	+
A Tratamento superficial	sem tratamento	com tratamento
B Temperatura e tempo de cura	22°C e $24h$	60°C e $2h$
C Rugosidade	sem jateamento	com jateamento
D Ativador	sem ativador	com ativador

Todas as combinações possíveis entre os níveis dos fatores considerados

estão esquematizados na Tabela 19.

Para uma descrição mais detalhada do experimento aqui mencionado recomenda-se a apreciação de Santos (2007).

Tabela 19 Tratamentos do experimento de resistência mecânica de junções adesivas.

Tratamento		Fator			
		<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>
1	(1)	-	-	-	-
2	<i>a</i>	+	-	-	-
3	<i>b</i>	-	+	-	-
4	<i>ab</i>	+	+	-	-
5	<i>c</i>	-	-	+	-
6	<i>ac</i>	+	-	+	-
7	<i>bc</i>	-	+	+	-
8	<i>abc</i>	+	+	+	-
9	<i>d</i>	-	-	-	+
10	<i>ad</i>	+	-	-	+
11	<i>bd</i>	-	+	-	+
12	<i>abd</i>	+	+	-	+
13	<i>cd</i>	-	-	+	+
14	<i>acd</i>	+	-	+	+
15	<i>bcd</i>	-	+	+	+
16	<i>abcd</i>	+	+	+	+

Em relação aos cálculos das análises estatísticas, estes foram realizados com o suporte do *software* R (R Development Core Team, 2012).

3.2 Métodos

O método de pesquisa utilizado no desenvolvimento da dissertação é pesquisa bibliográfica, baseada em revisão bibliográfica da literatura referente à esta-

tística experimental aplicada à indústria, a partir de artigos, livros, teses e dissertações.

Com o propósito de ilustrar as técnicas estatísticas de planejamento e análise de experimentos, abordadas no referencial teórico, propôs-se a pesquisa de dados de experimentos industriais que poderiam ser úteis nas aplicações desta dissertação. Entre os materiais obtidos na literatura, considerou-se como fonte secundária a massa de dados disponível na dissertação de Santos (2007). Esta pesquisa foi feita na biblioteca digital de teses e dissertações da USP.

Os dados de Santos (2007) foram utilizados em três situações, sendo que em duas delas os dados foram adaptados.

A primeira situação refere-se a análise do experimento fatorial 2^4 completo realizado em delineamento inteiramente casualizado com 10 repetições, tal como conduzido por Santos (2007).

A justificativa para analisar o mesmo experimento nas mesmas condições apresentadas por Santos (2007) é fornecer o desenvolvimento detalhado da análise estatística, em especial da explicitação do modelo estatístico e da análise de variância, que não foram abordados diretamente pelo trabalho já citado.

Para a situação em destaque, a análise de variância será desenvolvida tanto pela abordagem matricial, como pela escalar. Os resultados dessas duas abordagens são iguais, porém, com o fim de ilustrar ambos os procedimentos, as duas formas serão apresentadas. Pela abordagem matricial será realizada apenas a análise de variância sem desdobramento dos efeitos de interação caso estes sejam significativos.

Com o objetivo de exemplificar experimentos fatoriais não repetidos, na segunda situação é considerado, hipoteticamente, no experimento de resistência mecânica de junções adesivas, a impossibilidade de realizar mais do que uma re-

petição.

Para a segunda situação os dados foram adaptados, sendo que as médias dos tratamentos da primeira situação foram consideradas como única observação de cada tratamento.

Na terceira situação, o experimento de resistência mecânica de junções adesivas será analisado como se tivesse sido planejado com dez repetições e quatro blocos para cada repetição. Sendo que a dimensão desses blocos é menor que a quantidade total de tratamentos do experimento.

Os blocos serão construídos a partir do confundimento dos efeitos independentes das interações ACD e BCD . Além dos efeitos das interações ACD e BCD , o efeito da interação generalizada AB também confundir-se-á com o efeito de blocos.

A análise estatística para esta situação será apenas para verificar a significância dos efeitos principais, de interações e de blocos, sem considerar o desdobramento das interações caso estas sejam significativas.

4 RESULTADOS E DISCUSSÃO

Neste capítulo serão apresentadas três aplicações do planejamento e análise de experimentos na indústria, baseadas no mesmo experimento de resistência mecânica de junções adesivas.

4.1 Situação 1: experimento fatorial 2^4 completo

Os dados de tensão média de ruptura (MPa) obtidos pelos ensaios de cisalhamento estão disponíveis na Tabela 20.

Tabela 20 Dados de tensão média de ruptura (*MPa*) do experimento de resistência mecânica de junções adesivas.

Tratamento	Repetição									
	1	2	3	4	5	6	7	8	9	10
1 (1)	14,9	14,32	18,39	15,70	16,66	14,04	13,76	14,28	15,28	12,46
2 <i>a</i>	13,03	15,94	15,10	15,55	14,46	14,14	12,98	14,68	13,97	15,93
3 <i>b</i>	18,14	19,15	17,69	17,28	17,29	19,04	19,64	18,52	17,92	17,40
4 <i>ab</i>	15,13	15,31	19,61	15,36	17,56	18,35	15,67	18,30	19,90	19,51
5 <i>c</i>	15,48	17,12	16,36	18,41	19,13	19,51	16,84	13,19	18,25	18,52
6 <i>ac</i>	15,19	13,40	15,38	14,84	12,88	15,02	14,28	13,26	15,38	16,25
7 <i>bc</i>	20,37	18,80	18,82	18,46	17,92	18,42	19,41	19,25	13,79	18,95
8 <i>abc</i>	13,5	18,58	15,91	18,27	16,50	17,64	17,85	18,64	15,86	15,29
9 <i>d</i>	10,81	8,88	11,68	10,45	10,65	11,42	8,84	9,03	10,97	10,11
10 <i>ad</i>	11,38	8,34	10,28	9,22	10,63	9,44	9,75	11,29	11,31	9,57
11 <i>bd</i>	11,47	14,20	16,50	13,19	11,84	13,71	12,14	16,03	13,45	11,71
12 <i>abd</i>	12,06	12,91	9,98	11,10	9,59	12,70	12,10	13,35	14,74	11,96
13 <i>cd</i>	13,03	13,40	14,59	11,86	9,13	11,87	12,16	12,68	13,79	13,23
14 <i>acd</i>	13,24	10,32	9,96	11,85	10,44	9,12	9,84	14,47	11,87	11,85
15 <i>bcd</i>	13,75	12,61	16,29	16,26	14,12	12,13	15,06	12,28	14,52	14,21
16 <i>abcd</i>	13,63	13,79	12,25	13,57	12,40	13,84	12,05	13,07	12,63	12,02

O modelo estatístico para este experimento pode ser conforme (2.1) ou (2.2) em que os efeitos de tratamentos são decompostos em efeitos principais e de interações. Considerando o modelo de efeitos de tratamentos, tem-se

$$Y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \quad (4.1)$$

em que,

$$i = 1, \dots, 16 \text{ e } j = 1, \dots, 10,$$

Y_{ij} é a j -ésima observação do i -ésimo tratamento,

μ é uma constante comum a todas as observações,

τ_i é o efeito do i -ésimo tratamento,

ε_{ij} é o erro experimental da j -ésima observação do i -ésimo tratamento, supondo que $\varepsilon_{ij} \stackrel{iid}{\sim} N(0, \sigma^2)$.

Para a decomposição dos efeitos de tratamentos em efeitos principais e de interações, tem-se o seguinte modelo

$$\begin{aligned} Y_{i_1 i_2 i_3 i_4 j} = & \mu + \alpha_{i_1} + \beta_{i_2} + \gamma_{i_3} + \delta_{i_4} + (\alpha\beta)_{i_1 i_2} + (\alpha\gamma)_{i_1 i_3} + (\alpha\delta)_{i_1 i_4} + \\ & + (\beta\gamma)_{i_2 i_3} + (\beta\delta)_{i_2 i_4} + (\gamma\delta)_{i_3 i_4} + (\alpha\beta\gamma)_{i_1 i_2 i_3} + (\alpha\beta\delta)_{i_1 i_2 i_4} + \\ & + (\alpha\gamma\delta)_{i_1 i_3 i_4} + (\beta\gamma\delta)_{i_2 i_3 i_4} + (\alpha\beta\gamma\delta)_{i_1 i_2 i_3 i_4} + \varepsilon_{i_1 i_2 i_3 i_4 j} \end{aligned} \quad (4.2)$$

em que,

$$i_1 = 1, 2; i_2 = 1, 2; i_3 = 1, 2; i_4 = 1, 2; j = 1, \dots, 10;$$

$Y_{i_1 i_2 j}$ é a j -ésima observação do i_1 -ésimo nível do fator A , i_2 -ésimo nível do fator B , i_3 -ésimo nível do fator C e i_4 -ésimo nível do fator D ;

μ é uma constante comum a todas as observações;

α_{i_1} é o efeito do i_1 -ésimo nível do fator A ;

β_{i_2} é o efeito do i_2 -ésimo nível do fator B ;

γ_{i_3} é o efeito do i_3 -ésimo nível do fator C ;

δ_{i_4} é o efeito do i_4 -ésimo nível do fator D ;

$(\alpha\beta)_{i_1i_2}$ é o efeito da interação do i_1 -ésimo nível do fator A com o i_2 -ésimo nível do fator B ;

$(\alpha\gamma)_{i_1i_3}$ é o efeito da interação do i_1 -ésimo nível do fator A com o i_3 -ésimo nível do fator C ;

$(\alpha\delta)_{i_1i_4}$ é o efeito da interação do i_1 -ésimo nível do fator A com o i_4 -ésimo nível do fator D ;

$(\beta\gamma)_{i_2i_3}$ é o efeito da interação do i_2 -ésimo nível do fator B com o i_3 -ésimo nível do fator C ;

$(\beta\delta)_{i_2i_4}$ é o efeito da interação do i_2 -ésimo nível do fator B com o i_4 -ésimo nível do fator D ;

$(\gamma\delta)_{i_3i_4}$ é o efeito da interação do i_3 -ésimo nível do fator C com o i_4 -ésimo nível do fator D ;

$(\alpha\beta\gamma)_{i_1i_2i_3}$ é o efeito da interação do i_1 -ésimo nível do fator A , i_2 -ésimo nível do fator B e i_3 -ésimo nível do fator C ;

$(\alpha\beta\delta)_{i_1i_2i_4}$ é o efeito da interação do i_1 -ésimo nível do fator A , i_2 -ésimo nível do fator B e i_4 -ésimo nível do fator D ;

$(\alpha\gamma\delta)_{i_1i_3i_4}$ é o efeito da interação do i_1 -ésimo nível do fator A , i_3 -ésimo nível do fator C e i_4 -ésimo nível do fator D ;

$(\beta\gamma\delta)_{i_2i_3i_4}$ é o efeito da interação do i_2 -ésimo nível do fator B , i_3 -ésimo nível do fator C e i_4 -ésimo nível do fator D ;

$(\alpha\beta\gamma\delta)_{i_1i_2i_3i_4}$ é o efeito da interação do i_1 -ésimo nível do fator A , i_2 -ésimo nível do fator B , i_3 -ésimo nível do fator C e i_4 -ésimo nível do fator D ;

$\varepsilon_{i_1i_2i_3i_4j}$ é o erro experimental da j -ésima observação do i_1 -ésimo nível do fator A , i_2 -ésimo nível do fator B , i_3 -ésimo nível do fator C e i_4 -ésimo nível do fator

D , supondo que $\varepsilon_{i_1 i_2 i_3 i_4 j} \stackrel{iid}{\sim} N(0, \sigma^2)$.

A verificação da influência dos fatores de interesse, já mencionados, dar-se-á por meio da análise de variância, a qual será desenvolvida nas subseções seguintes.

4.1.1 Análise de variância pela abordagem matricial

Outra possibilidade de escrever o modelo (4.1) é na forma matricial (2.7), considerando as seguintes matrizes

$$\mathbf{Y} = \begin{pmatrix} Y_{1,1} \\ Y_{1,2} \\ \vdots \\ Y_{1,10} \\ Y_{2,1} \\ \vdots \\ Y_{16,10} \end{pmatrix} = \begin{pmatrix} 14,90 \\ 14,32 \\ \vdots \\ 12,46 \\ 13,03 \\ \vdots \\ 12,02 \end{pmatrix}, \mathbf{X} = \begin{pmatrix} 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & 0 & 0 & \cdots & 1 \end{pmatrix},$$

$$\boldsymbol{\beta} = \begin{pmatrix} \mu \\ \tau_1 \\ \tau_2 \\ \vdots \\ \tau_{16} \end{pmatrix} \text{ e } \boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_{1,1} \\ \varepsilon_{1,2} \\ \vdots \\ \varepsilon_{1,10} \\ \varepsilon_{2,1} \\ \vdots \\ \varepsilon_{16,10} \end{pmatrix},$$

sendo \mathbf{Y} de dimensões 160×1 ; \mathbf{X} de dimensões 160×17 ; $\boldsymbol{\beta}$ de dimensões 17×1 ; $\boldsymbol{\varepsilon}$ de dimensões 160×1 .

A matriz de delineamento \mathbf{X} pode ser particionada semelhante a (2.11),

$$\mathbf{X} = \left(\mathbf{X}_1, \mathbf{X}_t \right),$$

sendo,

$$\mathbf{X}_1 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \text{ e } \mathbf{X}_t = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

A variabilidade total dos dados corrigido para a constante μ pode ser decomposta semelhante a (2.15)

$$\mathbf{Y}'(\mathbf{I}_{160} - \mathbf{P}_1)\mathbf{Y} = \mathbf{Y}'\mathbf{P}_t\mathbf{Y} + \mathbf{Y}'\mathbf{P}_{\text{Res}}\mathbf{Y} \quad (4.3)$$

em que,

$\mathbf{Y}'(\mathbf{I}_{160} - \mathbf{P}_1)\mathbf{Y}$ é a soma de quadrados total corrigido, SQ_{T_c} , com \mathbf{I}_{160} a matriz identidade de ordem 160 e $\mathbf{P}_1 = \mathbf{X}_1(\mathbf{X}'_1\mathbf{X}_1)^{-1}\mathbf{X}'_1$;

$\mathbf{Y}'\mathbf{P}_t\mathbf{Y}$ é a soma de quadrados de tratamentos, SQ_t , com $\mathbf{P}_t = \mathbf{B}_t - \mathbf{P}_1$ e $\mathbf{B}_t = \mathbf{X}_t(\mathbf{X}'_t\mathbf{X}_t)^{-1}\mathbf{X}'_t$;

$\mathbf{Y}'\mathbf{P}_{\text{Res}}\mathbf{Y}$ é a soma de quadrados de resíduo, SQ_{Res} , com $\mathbf{P}_{\text{Res}} = \mathbf{I}_{160} - \mathbf{B}$ e $\mathbf{B} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$.

Em (4.3), o posto da matriz núcleo \mathbf{P}_t é igual a 15, e igual a 144 para \mathbf{P}_{Res} . Naturalmente, o posto da matriz núcleo $\mathbf{I}_{160} - \mathbf{P}_1$ é igual a 159, que é a soma do posto das matrizes núcleos já mencionadas. Pela definição de graus de

liberdade, tem-se que as fontes de variações tratamento, resíduo e total corrigido possuem 15, 144 e 159 graus de liberdade, respectivamente.

As hipóteses para verificar a influência dos tratamentos (Tabela 19) na resistência mecânica das junções adesivas são

$$H_0 : \tau_1 = \dots = \tau_{16}$$

vs

$$H_1 : \tau_i \neq \tau_j \text{ para pelo menos um } i \neq j, \text{ com } 1 \leq i, j \leq 16.$$

Essas hipóteses podem ser verificadas através da análise de variância que será desenvolvida abaixo.

Para as formas quadráticas em (4.3), obtém-se as seguintes somas de quadrados

$$SQ_{T_c} = \mathbf{Y}'(\mathbf{I}_{160} - \mathbf{P}_1)\mathbf{Y} = 1425,4204$$

$$SQ_t = \mathbf{Y}'\mathbf{P}_t\mathbf{Y} = 1114,0260$$

$$SQ_{Res} = \mathbf{Y}'\mathbf{P}_{Res}\mathbf{Y} = 311,3944$$

Para os quadrados médios tem-se

$$QM_t = \frac{SQ_t}{\mathcal{P}(\mathbf{P}_t)} = \frac{\mathbf{Y}'\mathbf{P}_t\mathbf{Y}}{15} = \frac{1114,0260}{15} = 74,2684$$

$$QM_{Res} = \frac{SQ_{Res}}{\mathcal{P}(\mathbf{P}_{Res})} = \frac{\mathbf{Y}'\mathbf{P}_{Res}\mathbf{Y}}{144} = \frac{311,3944}{144} = 2,1625.$$

A estatística de teste para avaliar as hipóteses de igualdade entre os efeitos de tratamento é conforme a expressão (2.14), isto é, o quociente entre o quadrado médio de tratamento e o quadrado médio de resíduo, que sob H_0 segue distribuição

F de Snedecor central. Logo, o valor da estatística de teste, F_c , é

$$F_c = \frac{QM_t}{QM_{\text{Res}}} = \frac{74,2684}{2,1625} = 34,34.$$

Para um nível de 5% de significância, o valor crítico de $F_{(15,144)}$ é 1,74. Como $34,34 > 1,74$ a hipótese nula, H_0 , da análise de variância, esquematizada na Tabela 21, é rejeitada. Desta forma, há pelo menos um tratamento que influencia estatisticamente na resistência mecânica de junções adesivas.

Tabela 21 Esquema da análise de variância preliminar dos dados de tensão média de ruptura de junções adesivas (MPa).

Fonte de Variação	GL	SQ	QM	F_c	valor-p
Tratamento	15	1114,0260	74,2684	34,34	< 0,01
Resíduo	144	311,3944	2,1625		
Total	159	1425,4204			

O resultado da análise de variância para os efeitos de tratamentos, esquematizado na Tabela 21, é pouco informativo, pois não é possível, apenas com o resultado já obtido, concluir qual combinação dos níveis dos fatores considerados proporciona a maior, ou menor, resistência nas junções adesivas.

Às vezes, em especial na abordagem matricial da análise de variância, é preferível considerar o modelo (4.2) a (4.1) quando o estudo envolver mais de 1 fator.

Considerando o modelo (4.2), em que os efeitos de tratamentos foram decompostos em efeitos principais e de interações, também é possível, tal como o modelo (4.1), escrevê-lo na forma matricial e decompor a matriz de delineamento

\mathbf{X} conforme (4.4).

$$\mathbf{X} = \left(\mathbf{X}_1, \mathbf{X}_A, \dots, \mathbf{X}_{AB}, \dots, \mathbf{X}_{ABC}, \dots, \mathbf{X}_{ABCD} \right). \quad (4.4)$$

Tendo em vista o particionamento da matriz de delineamento \mathbf{X} conforme (4.4), a variabilidade total dos dados corrigido para a constante μ pode ser decomposta tal como a expressão abaixo

$$\begin{aligned} \mathbf{Y}'(\mathbf{I}_{160} - \mathbf{P}_1)\mathbf{Y} &= \mathbf{Y}'\mathbf{P}_A\mathbf{Y} + \dots + \mathbf{Y}'\mathbf{P}_{AB}\mathbf{Y} + \dots + \mathbf{Y}'\mathbf{P}_{ABC}\mathbf{Y} + \\ &+ \dots + \mathbf{Y}'\mathbf{P}_{ABCD}\mathbf{Y} + \mathbf{Y}'\mathbf{P}_{\text{Res}}\mathbf{Y}. \end{aligned} \quad (4.5)$$

em que,

$\mathbf{Y}'(\mathbf{I}_{160} - \mathbf{P}_1)\mathbf{Y}$ é a soma de quadrados total corrigido, SQ_{Tc} , com \mathbf{I}_{160} a matriz identidade de ordem 160 e $\mathbf{P}_1 = \mathbf{X}_1(\mathbf{X}'_1\mathbf{X}_1)^{-1}\mathbf{X}'_1$;

$\mathbf{Y}'\mathbf{P}_A\mathbf{Y}$ é a soma de quadrados do efeito principal do fator A , SQ_A ,

com $\mathbf{P}_A = \mathbf{B}_A - \mathbf{P}_1$ e $\mathbf{B}_A = \mathbf{X}_A(\mathbf{X}'_A\mathbf{X}_A)^{-1}\mathbf{X}'_A$;

\dots ;

$\mathbf{Y}'\mathbf{P}_{AB}\mathbf{Y}$ é a soma de quadrados do efeito de interação dos fatores A e B , SQ_{AB} ,

com $\mathbf{P}_{AB} = \mathbf{B}_{AB} - \mathbf{P}_A - \mathbf{P}_B - \mathbf{P}_1$ e $\mathbf{B}_{AB} = \mathbf{X}_{AB}(\mathbf{X}'_{AB}\mathbf{X}_{AB})^{-1}\mathbf{X}'_{AB}$;

\dots ;

$\mathbf{Y}'\mathbf{P}_{ABC}\mathbf{Y}$ é a soma de quadrados do efeito de interação dos fatores A , B e C ,

SQ_{ABC} , com $\mathbf{P}_{ABC} = \mathbf{B}_{ABC} - \mathbf{P}_A - \mathbf{P}_B - \mathbf{P}_{AB} - \mathbf{P}_1$ e

$\mathbf{B}_{ABC} = \mathbf{X}_{ABC}(\mathbf{X}'_{ABC}\mathbf{X}_{ABC})^{-1}\mathbf{X}'_{ABC}$;

\dots ;

$\mathbf{Y}'\mathbf{P}_{ABCD}\mathbf{Y}$ é a soma de quadrados do efeito de interação dos fatores A , B , C e

D , SQ_{ABCD} , com $\mathbf{P}_{ABCD} = \mathbf{B}_{ABCD} - \mathbf{P}_A - \dots - \mathbf{P}_D - \mathbf{P}_{AB} - \dots - \mathbf{P}_{CD} -$

$\mathbf{P}_{ABC} - \dots - \mathbf{P}_{BCD} - \mathbf{P}_1$ e $\mathbf{B}_{ABCD} = \mathbf{X}_{ABCD}(\mathbf{X}'_{ABCD}\mathbf{X}_{ABCD})^{-1}\mathbf{X}'_{ABCD}$;

$\mathbf{Y}'\mathbf{P}_{\text{Res}}\mathbf{Y}$ é a soma de quadrados de resíduos, SQ_{Res} ,

com $\mathbf{P}_{\text{Res}} = \mathbf{I}_{160} - \mathbf{P}_A - \cdots - \mathbf{P}_{AB} - \cdots - \mathbf{P}_{ABC} - \cdots - \mathbf{P}_{ABCD} - \mathbf{P}_1$.

Na decomposição da variabilidade total dos dados corrigido para a constante μ , (4.5), o posto das matrizes núcleos $\mathbf{P}_A, \cdots, \mathbf{P}_{AB}, \cdots, \mathbf{P}_{ABC}, \cdots, \mathbf{P}_{BCD}$ e \mathbf{P}_{ABCD} são todos iguais a 1, e 144 para a matriz núcleo \mathbf{P}_{Res} . Como já esperado, o posto da matriz núcleo $\mathbf{I}_{160} - \mathbf{P}_1$ é igual a 159. Desta forma, as fontes de variações de todos os efeitos principais e de todas as interações possuem 1 grau de liberdade; o resíduo e o total corrigido possuem 144 e 159 graus de liberdade, respectivamente.

Algumas das hipóteses de interesse para verificar a influência dos efeitos principais e de interações são:

$$H_0 : \delta_1 = \delta_2 \text{ vs } H_1 : \delta_1 \neq \delta_2$$

$$H_0 : (\alpha\beta\gamma)_{111} = \cdots = (\alpha\beta\gamma)_{222}$$

vs

H_1 : pelo menos um efeito de interação difere dos demais.

Sendo que, o primeiro par de hipóteses, H_0 e H_1 , é referente ao efeito principal do fator D ; e o segundo par é referente ao efeito de interação dos fatores A , B e C . Essas e outras hipóteses podem ser testadas através da análise de variância.

Algumas somas de quadrados e quadrados médios necessárias para os testes da análise de variância são apresentadas abaixo.

$$SQ_{T_c} = \mathbf{Y}'(\mathbf{I}_{160} - \mathbf{P}_1)\mathbf{Y} = 1425,4204;$$

$$SQ_A = \mathbf{Y}'\mathbf{P}_A\mathbf{Y} = 55,9323;$$

$$\begin{aligned}
 QM_A &= \frac{SQ_A}{\mathcal{P}(\mathbf{P}_A)} = \frac{55,9323}{1} = 55,9323; \\
 SQ_{ABCD} &= \mathbf{Y}'\mathbf{P}_{ABCD}\mathbf{Y} = 0,0093; \\
 QM_{ABCD} &= \frac{SQ_{ABCD}}{\mathcal{P}(\mathbf{P}_{ABCD})} = \frac{0,0093}{1} = 0,0093; \\
 SQ_{\text{Res}} &= \mathbf{Y}'\mathbf{P}_{\text{Res}}\mathbf{Y} = 311,3944. \\
 QM_{\text{Res}} &= \frac{SQ_{\text{Res}}}{\mathcal{P}(\mathbf{P}_{\text{Res}})} = \frac{311,3944}{144} = 2,1625.
 \end{aligned}$$

Os valores de todas as somas de quadrados e quadrados médios estão apresentados na Tabela 22.

É importante ressaltar que os valores dos quadrados médios das fontes de variações $A, B, \dots, ABCD$ são iguais ao valor de sua própria soma de quadrados, pois, as fontes de variações citadas possuem 1 grau de liberdade. Portanto,

$$QM_i = \frac{SQ_i}{\mathcal{P}(\mathbf{P}_i)} = \frac{SQ_i}{1} = SQ_i,$$

para todo $i, i = A, B, \dots, ABCD$.

A estatística de teste para avaliar as hipóteses referentes à influência dos fatores A, B, C e D na resistência mecânica de junções adesivas é

$$F_c = \frac{QM_i}{QM_{\text{Res}}}, \quad \text{para } i = A, B, \dots, ABCD.$$

Alguns valores da estatística de teste F_c são

$$F_c = \frac{QM_A}{QM_{\text{Res}}} = \frac{55,9323}{2,1625} = 25,87$$

para o efeito do fator A ;

$$F_c = \frac{QM_{ABCD}}{QM_{Res}} = \frac{0,0093}{2,1625} = 0,004$$

para o efeito de interação dos fatores A , B , C e D .

Os valores de todas as estatísticas de testes para avaliar as hipóteses referentes à influência dos fatores A , B , C e D na resistência mecânica de junções adesivas estão apresentados na Tabela 22.

Tabela 22 Esquema da análise de variância do experimento de resistência mecânica de junções adesivas em esquema fatorial 2^4 .

Fonte de Variação	GL	SQ	QM	F_c	valor-p
A	1	55,9323	55,9323	25,87	< 0,01
B	1	196,2490	196,2490	90,75	< 0,01
C	1	29,7390	29,7390	13,75	< 0,01
D	1	788,9880	788,9880	364,86	< 0,01
AB	1	0,0951	0,0951	0,04	0,83
AC	1	10,5473	10,5473	4,88	0,03
AD	1	1,2816	1,2816	0,59	0,44
BC	1	13,5490	13,5490	6,27	0,01
BD	1	0,9425	0,9425	0,44	0,51
CD	1	6,3282	6,3282	2,93	0,09
ABC	1	4,5765	4,5765	2,12	0,15
ABD	1	2,1950	2,1950	1,02	0,32
ACD	1	3,1136	3,1136	1,44	0,23
BCD	1	0,4796	0,4796	0,22	0,64
$ABCD$	1	0,0093	0,0093	0,004	0,95
Resíduo	144	311,3944	2,1625		
Total	159	1425,4204			

Através da análise de variância esquematizada na Tabela 22, verifica-se que todos os efeitos principais e as interações AC e BC são estatisticamente sig-

nificativas a um nível de 5% de significância. O valor crítico de $F_{(1,144)}$, para o nível de 5% de significância, é 3,91.

Como resultado da análise de variância, Tabela 22, nota-se que o fator D , uso de ativador, influencia na resistência mecânica das junções adesivas. Pelo fato de nenhuma interação envolvendo o fator D ser significativa, o efeito do uso de ativador independe dos demais fatores.

A média das médias dos tratamentos com o fator D no nível baixo é igual a

$$\frac{\bar{Y}_{(1)} + \bar{Y}_a + \bar{Y}_b + \bar{Y}_c + \bar{Y}_{ab} + \bar{Y}_{ac} + \bar{Y}_{bc} + \bar{Y}_{abc}}{8} = 16,54,$$

e a média das médias dos tratamentos com o fator D no nível alto é igual a

$$\frac{\bar{Y}_d + \bar{Y}_{ad} + \bar{Y}_{bd} + \bar{Y}_{cd} + \bar{Y}_{abd} + \bar{Y}_{acd} + \bar{Y}_{bcd} + \bar{Y}_{abcd}}{8} = 12,10.$$

Em relação ao uso do ativador, tem-se que a maior resistência mecânica das junções adesivas é obtida sem o seu uso pois, $16,54 > 12,10$.

Como os efeitos principais A , B e C , e as interações AC e BC são significativas, é desejável estudar os efeitos das interações ao invés de cada efeito principal individualmente, que poderia acarretar conclusões enganosas.

O estudo dos efeitos das interações significativas será conduzido pela abordagem escalar da análise de variância, a qual será desenvolvida na próxima subseção.

4.1.2 Análise de variância pela abordagem escalar

A princípio, para o desenvolvimento da análise de variância pela abordagem escalar, serão consideradas as hipóteses, já apresentadas na abordagem matricial, referentes à igualdade dos efeitos de tratamentos.

As somas de quadrados total, tratamentos e resíduos serão calculadas a partir das expressões (2.35), (2.36) e (2.38). No cálculo da soma de quadrados de resíduos, o efeito de blocos é desconsiderado, já que o delineamento considerado é o inteiramente casualizado.

$$\begin{aligned}
 SQ_{T_c} &= \sum_{i=1}^{16} \sum_{j=1}^{10} Y_{ij}^2 - \frac{\left(\sum_{j=1}^{10} \sum_{i=1}^{16} Y_{ij} \right)^2}{160} = \\
 &= (14,9^2 + 14,32^2 + \dots + 12,46^2 + 13,03^2 + \dots + 12,02^2) - \\
 &\quad - \frac{(14,9 + 14,32 + \dots + 12,46 + 13,03 + \dots + 12,02)^2}{160} = \\
 &= 34235,9772 - \frac{(2291,2200)^2}{160} = 1425,4204
 \end{aligned}$$

$$\begin{aligned}
 SQ_t &= \frac{\sum_{i=1}^{16} \left(\sum_{j=1}^{10} Y_{ij} \right)^2}{10} - \frac{\left(\sum_{i=1}^{16} \sum_{j=1}^{10} Y_{ij} \right)^2}{160} = \\
 &= \frac{(14,9 + \dots + 12,46)^2 + \dots + (13,63 + \dots + 12,02)^2}{10} - \\
 &\quad - \frac{(14,9 + 14,32 + \dots + 12,46 + 13,03 + \dots + 12,02)^2}{160} = \\
 &= \frac{149,79^2 + \dots + 129,25^2}{10} - \frac{(2291,2200)^2}{160} = \\
 &= 33924,5828 - \frac{(2291,2200)^2}{160} = 1114,0260
 \end{aligned}$$

$$SQ_{Res} = SQ_{T_c} - SQ_t = 1425,4204 - 1114,0260 = 311,3944$$

Para os quadrados médios de tratamentos e resíduos:

$$QM_t = \frac{SQ_t}{16 - 1} = \frac{1114,0260}{15} = 74,2684$$

$$QM_{\text{Res}} = \frac{SQ_{\text{Res}}}{159 - 15} = \frac{311,3944}{144} = 2,1625$$

Para a estatística de teste, F_c , das hipóteses referentes aos efeitos de tratamentos, tem-se que

$$F_c = \frac{QM_t}{QM_{\text{Res}}} = \frac{74,2684}{2,1625} = 34,34$$

O esquema da análise de variância para o teste das hipóteses referentes aos efeitos de tratamentos, independentemente da abordagem considerada, está apresentado na Tabela 21.

Para o desenvolvimento da análise de variância pela abordagem escalar, com a decomposição dos graus de liberdade e somas de quadrados de tratamentos em efeitos principais e de interações, de acordo com o esquema fatorial 2^4 , será considerado o conceito de contrastes mutuamente ortogonais.

Contrastes mutuamente ortogonais, referentes aos efeitos principais e de interações, são construídos a partir da Tabela 23, de sinais positivos e negativos.

Alguns contrastes mutuamente ortogonais são:

$$\begin{aligned}
 C_A &= \frac{1}{8} (\bar{Y}_a + \bar{Y}_{ab} + \bar{Y}_{ac} + \bar{Y}_{abc} + \bar{Y}_{ad} + \bar{Y}_{abd} + \bar{Y}_{acd} + \bar{Y}_{abcd} - \\
 &\quad - \bar{Y}_{(1)} - \bar{Y}_b - \bar{Y}_c - \bar{Y}_{bc} - \bar{Y}_d - \bar{Y}_{bd} - \bar{Y}_{cd} - \bar{Y}_{bcd}) = \\
 &= \frac{1}{8} (14,578 + \dots + 12,925 - 14,979 - \dots - 14,123) = -1,1825
 \end{aligned}$$

$$\begin{aligned}
 C_B &= \frac{1}{8} (\bar{Y}_b + \bar{Y}_{ab} + \bar{Y}_{bc} + \bar{Y}_{abc} + \bar{Y}_{bd} + \bar{Y}_{abd} + \bar{Y}_{bcd} + \bar{Y}_{abcd} - \\
 &\quad - \bar{Y}_{(1)} - \bar{Y}_a - \bar{Y}_c - \bar{Y}_{ac} - \bar{Y}_d - \bar{Y}_{ad} - \bar{Y}_{cd} - \bar{Y}_{acd}) = \\
 &= \frac{1}{8} (18,207 + \dots + 12,925 - 14,979 - \dots - 11,296) = 2,2150
 \end{aligned}$$

$$\begin{aligned}
 C_{ABCD} &= \frac{1}{8} (\bar{Y}_{(1)} + \bar{Y}_{ab} + \bar{Y}_{ac} + \bar{Y}_{bc} + \bar{Y}_{ad} + \bar{Y}_{bd} + \bar{Y}_{cd} + \bar{Y}_{abcd} - \\
 &\quad - \bar{Y}_a - \bar{Y}_b - \bar{Y}_c - \bar{Y}_{abc} - \bar{Y}_d - \bar{Y}_{abd} - \bar{Y}_{acd} - \bar{Y}_{bcd}) = \\
 &= \frac{1}{8} (14,979 + \dots + 12,925 - 14,578 - \dots - 14,123) = \\
 &= -0,0153
 \end{aligned}$$

As estimativas de todos os contrastes mutuamente ortogonais referentes aos efeitos principais e de interações dos fatores A , B , C e D são:

$$\begin{aligned}
 C_A &= -1,1825; & C_B &= 2,2150; & C_C &= 0,8623; \\
 C_D &= -4,4413; & C_{AB} &= -0,0488; & C_{AC} &= -0,5135 \\
 C_{AD} &= 0,1790; & C_{BC} &= -0,5820; & C_{BD} &= -0,1535; \\
 C_{CD} &= 0,3978; & C_{ABC} &= 0,3383; & C_{ABD} &= -0,2343; \\
 C_{ACD} &= 0,2790; & C_{BCD} &= 0,1095; & C_{ABCD} &= -0,0153
 \end{aligned}$$

Algumas somas de quadrados dos contrastes mutuamente ortogonais definidos acima são obtidas da seguinte maneira:

$$SQC_A = \frac{10(C_A)^2}{0,25} = \frac{10(-1,1825)^2}{0,25} = 55,9323$$

$$SQC_B = \frac{10(C_B)^2}{0,25} = \frac{10(2,2150)^2}{0,25} = 196,2490$$

$$SQC_{ABCD} = \frac{10(C_{ABCD})^2}{0,25} = \frac{10(-0,0153)^2}{0,25} = 0,0093.$$

Os valores de todas as somas de quadrados dos contrastes mutuamente ortogonais referentes aos efeitos principais e de interações dos fatores A , B , C e D são os mesmos dos apresentados na Tabela 22.

O cálculo da soma de quadrados total, SQ_{T_c} , já foi obtido no desenvolvimento da análise de variância para as hipóteses referentes aos efeitos de tratamentos. Mesmo utilizando a decomposição dos efeitos de tratamentos em contrastes ortogonais, a soma de quadrados total é calculada através da expressão (2.35).

A soma de quadrados de resíduos, SQ_{Res} , é obtida pela diferença entre a SQ_{T_c} e as demais somas de quadrados.

$$\begin{aligned} SQ_{Res} &= SQ_{T_c} - SQC_A - SQC_B - \dots - SQC_{ABCD} = \\ &= 1425,4204 - 55,9323 - 196,2490 - \dots - 0,0093 = \\ &= 311,3944. \end{aligned}$$

Como já citado na subseção anterior, as fontes de variações A , B , \dots , $ABCD$ possuem 1 grau de liberdade e, por isso, o quadrado médio dessas fontes de variações é igual a sua soma de quadrados. Em relação ao quadrado médio de resíduos, este já foi obtido anteriormente, também, pela abordagem escalar.

Pelo fato de que, para cada fonte de variação, as somas de quadrados e quadrados médios obtidos, tanto pela abordagem matricial como escalar, serem iguais, as estatísticas de teste F_c não serão novamente calculadas, pois seriam iguais às já obtidas.

O esquema da análise de variância para os testes referentes aos efeitos principais e de interações dos fatores A , B , C e D , independentemente da abordagem considerada, está apresentado na Tabela 22.

As primeiras conclusões que se obtém da análise de variância agora conduzida, Tabela 22, são que os efeitos principais A , B , C e D e os efeitos de interações AC e BC são estatisticamente significativos a um nível de 5% de significância.

Pelo contraste C_D facilmente verifica-se que o uso de ativador reduz a resistência mecânica das junções adesivas, pois, $C_D = -4,4413 < 0$.

Para os fatores A , B e C não é recomendado analisar os seus efeitos na resistência mecânica das junções adesivas através dos contrastes C_A , C_B e C_C , porque as interações AC e BC são significativas também. A análise da influência dos fatores A , B e C dar-se-á pelo desdobramento das interações AC e BC , ou seja, para cada interação analisar-se-á o efeito de um fator dentro dos níveis do outro.

Os contrastes dos efeitos do fator A dentro de cada nível do fator C são:

$$\begin{aligned} C_{A:C-} &= \frac{1}{4} (\bar{Y}_a + \bar{Y}_{ab} + \bar{Y}_{ad} + \bar{Y}_{abd} - \bar{Y}_{(1)} - \bar{Y}_b - \bar{Y}_d - \bar{Y}_{bd}) = \\ &= \frac{1}{4} (14,578 + \dots + 12,049 - 14,979 - \dots - 13,424) = -0,6690 \end{aligned}$$

$$\begin{aligned} C_{A:C+} &= \frac{1}{4} (\bar{Y}_{ac} + \bar{Y}_{abc} + \bar{Y}_{acd} + \bar{Y}_{abcd} - \bar{Y}_c - \bar{Y}_{bc} - \bar{Y}_{cd} - \bar{Y}_{bcd}) = \\ &= \frac{1}{4} (14,588 + \dots + 12,925 - 17,281 - \dots - 14,123) = -1,6960 \end{aligned}$$

As estimativas dos contrastes dos efeitos do fator B dentro de cada nível do fator C , e do fator C dentro de cada nível dos fatores A e B são:

$$C_{B:C_-} = 2,7970; \quad C_{B:C_+} = 1,6330;$$

$$C_{C:A_-} = 1,3758; \quad C_{C:A_+} = 0,3488;$$

$$C_{C:B_-} = 1,4443; \quad C_{C:B_+} = 0,2803.$$

As somas de quadrados referentes aos contrastes dos efeitos do fator A dentro de cada nível do fator C são:

$$SQC_{A:C_-} = \frac{10(C_{A:C_-})^2}{0,5} = \frac{10(-0,6690)^2}{0,5} = 8,9512$$

$$SQC_{A:C_+} = \frac{10(C_{A:C_+})^2}{0,5} = \frac{10(-1,6960)^2}{0,5} = 57,5283$$

As somas de quadrados alusivas aos contrastes dos efeitos do fator B dentro de cada nível do fator C , e do fator C dentro de cada nível dos fatores A e B estão apresentadas, respectivamente, nas Tabelas 25, 26 e 27.

Note-se que a soma das somas de quadrados e os graus de liberdade dos efeitos A e AC são decompostos em duas novas somas de quadrados, com um grau de liberdade cada.

$$\begin{aligned} SQC_A + SQC_{AC} &= 55,9323 + 10,5473 = 66,4796 \approx \\ &\approx 66,4795 = 8,9512 + 57,5283 = \\ &= SQC_{A:C_-} + SQC_{A:C_+}. \end{aligned}$$

O mesmo é válido para os efeitos B e BC , C e AC , C e BC .

Pelo fato de que, a cada contraste está associado 1 grau de liberdade, o

quadrado médio de cada efeito citado acima é igual a sua soma de quadrados.

As estatísticas de testes, F_c , para avaliar os efeitos do fator A dentro de cada nível do fator C são:

$$F_c = \frac{QMC_{A:C_-}}{QM_{\text{Res}}} = \frac{SQC_{A:C_-}}{QM_{\text{Res}}} = \frac{8,9512}{2,1625} = 4,14,$$

para o efeito do fator A dentro do nível baixo do fator C ;

$$F_c = \frac{QMC_{A:C_+}}{QM_{\text{Res}}} = \frac{SQC_{A:C_+}}{QM_{\text{Res}}} = \frac{57,5283}{2,1625} = 26,60,$$

para o efeito do fator A dentro do nível alto do fator C .

Os valores das estatísticas de testes, F_c , para avaliar os efeitos do fator B dentro de cada nível do fator C , e C dentro de cada nível de A e B estão apresentadas, respectivamente, nas Tabelas 25, 26 e 27.

As análises de variância para o estudo dos efeitos dos fatores A e B dentro de cada nível do fator C estão esquematizadas nas Tabelas 24 e 25, e para o estudo dos efeitos do fator C dentro de cada nível dos fatores A e B estão esquematizadas nas Tabelas 26 e 27.

Tabela 24 Esquema da análise de variância para o estudo do efeito do fator A em cada nível do fator C .

Fonte de Variação	GL	SQ	QM	F_c	valor-p
$A:C_-$	1	8,9512	8,9512	4,14	0,04
$A:C_+$	1	57,5283	57,5283	26,60	< 0,01
Resíduo	144	311,3944	2,1625		

Tabela 25 Esquema da análise de variância para o estudo do efeito do fator B em cada nível do fator C .

Fonte de Variação	GL	SQ	QM	F_c	valor-p
$B:C_-$	1	156,4642	156,4642	72,35	< 0,01
$B:C_+$	1	53,3338	53,3338	24,66	< 0,01
Resíduo	144	311,3944	2,1625		

Tabela 26 Esquema da análise de variância para o estudo do efeito do fator C em cada nível do fator A .

Fonte de Variação	GL	SQ	QM	F_c	valor-p
$C:A_-$	1	37,8538	37,8538	17,50	< 0,01
$C:A_+$	1	2,4325	2,4325	1,12	0,29
Resíduo	144	311,3944	2,1625		

Tabela 27 Esquema da análise de variância para o estudo do efeito do fator C em cada nível do fator B .

Fonte de Variação	GL	SQ	QM	F_c	valor-p
$C:B_-$	1	41,7172	41,7172	19,29	< 0,01
$C:B_+$	1	1,5708	1,5708	0,73	0,39
Resíduo	144	311,3944	2,1625		

Na análise de variância do desdobramento das interações AC e BC em A dentro de C e B dentro de C , Tabelas 24 e 25, todos os efeitos analisados nos desdobramentos são estatisticamente significativos a um nível de 5% de significância,

$F_{(1,144)} = 3,91$. No entanto, para o desdobramento das interações AC e BC em C dentro de A e B , Tabelas 26 e 27, somente os efeitos de C dentro do nível baixo de A e B são estatisticamente significativos a um nível de 5% de confiança.

Para o efeito da interação AC na resistência mecânica de junções adesivas e o nível do fator C fixado em baixo, a resistência mecânica das junções adesivas é reduzida quando considerado nível alto ao invés de baixo para o fator A , pois, $C_{A:C_-} = -0,6690 < 0$. Fixando o nível do fator C em alto, o efeito do fator A é semelhante ao caso anterior, no entanto, a redução na resistência mecânica das junções adesivas é mais expressiva ao considerar o nível alto ao invés de baixo para o fator A , pois, $C_{A:C_+} = -1,6960 < 0$.

Para o efeito da interação BC na resistência mecânica de junções adesivas, quando o nível do fator C está fixado em baixo, há um expressivo aumento na resistência mecânica das junções adesivas quando considerado o nível alto ao invés de baixo para o fator B , pois, $C_{B:C_-} = 2,7970 > 0$. Fixando o nível do fator C em alto, o efeito do fator B é semelhante ao caso anterior, no entanto, o aumento na resistência mecânica das junções adesivas é menos expressivo ao considerar o nível alto ao invés de baixo para o fator B , pois, $C_{B:C_+} = 1,6330 > 0$.

Para o efeito da interação AC na resistência mecânica de junções adesivas, quando o nível do fator A está fixado em baixo, há um aumento na resistência mecânica das junções adesivas quando considerado o nível alto ao invés de baixo para o fator C , pois, $C_{C:A_-} = 1,3758 > 0$. Estatisticamente, a um nível de significância de 5%, o fator C não influencia na resistência mecânica das junções adesivas quando o nível do fator A está fixado em alto.

Para o efeito da interação BC na resistência mecânica de junções adesivas, quando o nível do fator B está fixado em baixo, há um aumento na resistência mecânica das junções adesivas quando considerado o nível alto ao invés de baixo

para o fator C , pois, $C_{C:B_-} = 1,4443 > 0$. Estatisticamente, a um nível de significância de 5%, o fator C não influencia na resistência mecânica das junções adesivas quando o nível do fator B está fixado em alto.

Em síntese, tem-se um aumento na resistência mecânica quando considerado o fator tratamento superficial (A) no nível sem tratamento (A_-) combinado com qualquer nível do fator rugosidade (C). No entanto, dentro do nível alto do fator rugosidade (com jateamento), o aumento da resistência mecânica é mais expressivo. Para o fator temperatura e tempo de cura (B), aumenta-se a resistência mecânica quando considerado $60^\circ C$ e 2 horas (B_+) dentro de cada nível do fator rugosidade, sendo esse aumento mais expressivo quando considerado o jateamento.

Considerando todas as combinações possíveis entre os níveis dos fatores, e o custo em obter junções adesivas com maiores resistências, é preferível considerar o fator tratamento superficial no nível baixo (sem tratamento), o fator temperatura e tempo de cura no nível alto ($60^\circ C$ e 2 horas), o fator rugosidade no nível baixo (sem jateamento) e o fator ativador no nível baixo (sem ativador). Para essa combinação dos níveis dos fatores tem-se o tratamento b , cuja a média é igual a $\bar{Y}_b = 18,207$, a segunda maior entre os demais tratamentos (Tabela 23). Caso seja necessário maior resistência mecânica, e seja possível dispor de maiores recursos financeiros, aconselha-se o uso do jateamento, obtendo-se o tratamento bc , cuja a média é igual a $\bar{Y}_{bc} = 18,419$, a maior entre os demais tratamentos (Tabela 23).

Os resultados discutidos para as interações AC e BC também podem ser notados analisando-se os gráficos de efeitos de interações, Figuras 5 e 6, e as tabelas de médias, Tabelas 28 e 29.

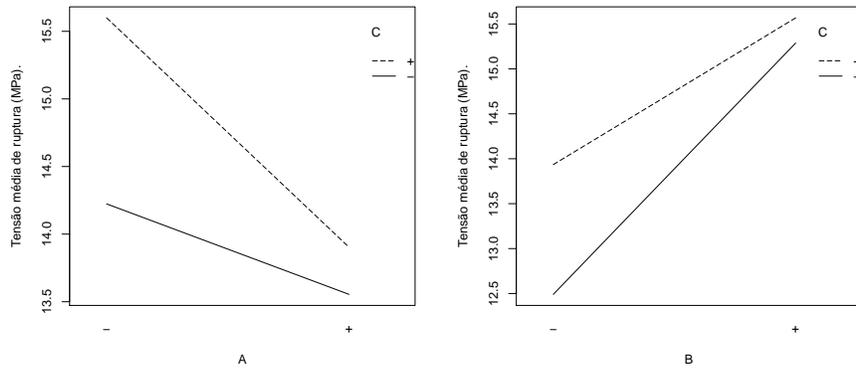


Figura 5 Gráficos de interações dos fatores *A* (tratamento superficial) e *B* (temperatura e tempo de cura) em cada nível do fator *C* (rugosidade), considerando a codificação “-” sem tratamento e “+” com tratamento, para o fator *A*; “-” 22°C e 24 horas e “+” 60°C e 2 horas, para o fator *B*; “-” sem jateamento e “+” com jateamento, para o fator *C*

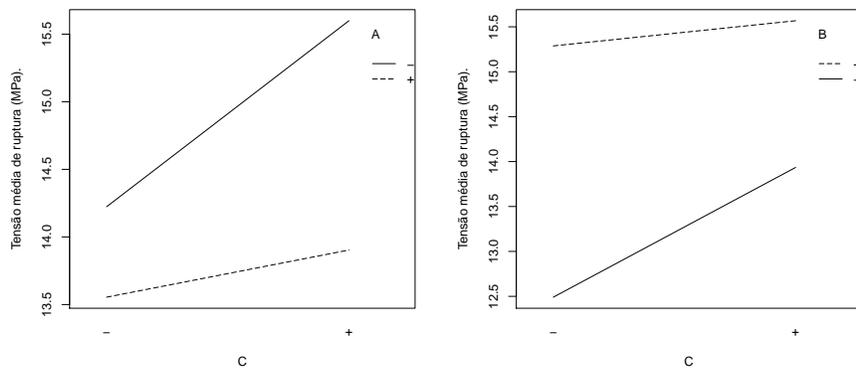


Figura 6 Gráficos de interações do fator *C* (rugosidade) em cada nível dos fatores *A* (tratamento superficial) e *B* (temperatura e tempo de cura), considerando a codificação “-” sem tratamento e “+” com tratamento, para o fator *A*; “-” 22°C e 24 horas e “+” 60°C e 2 horas, para o fator *B*; “-” sem jateamento e “+” com jateamento, para o fator *C*

Tabela 28 Médias de tensão média de ruptura nos níveis dos fatores A e C .

	C_-	C_+
A_-	14,22 u U	15,60 v U
A_+	13,55 u V	13,90 u V

u, v - em cada linha, médias seguidas de mesma letra minúscula não diferem estatisticamente (teste F a 5%);

U, V - em cada coluna, médias seguidas de mesma letra maiúscula não diferem estatisticamente (teste F a 5%).

Tabela 29 Médias de tensão média de ruptura nos níveis dos fatores B e C .

	C_-	C_+
B_-	12,49 u U	13,93 v U
B_+	15,29 u V	15,57 u V

u, v - em cada linha, médias seguidas de mesma letra minúscula não diferem estatisticamente (teste F a 5%);

U, V - em cada coluna, médias seguidas de mesma letra maiúscula não diferem estatisticamente (teste F a 5%).

4.2 Situação 2: experimento fatorial 2^4 não repetido

Na situação agora suposta, as médias dos tratamentos foram consideradas como única observação de cada tratamento. Os dados adaptados para essa situação estão apresentados na Tabela 30.

Tabela 30 Dados de tensão média de ruptura (MPa) do experimento fatorial 2^4 não repetido.

Tratamento	Tensão média de ruptura (MPa)
1 (1)	$Y_{(1)} = 14,979$
2 a	$Y_a = 14,578$
3 b	$Y_b = 18,207$
4 ab	$Y_{ab} = 17,470$
5 c	$Y_c = 17,281$
6 ac	$Y_{ac} = 14,588$
7 bc	$Y_{bc} = 18,419$
8 abc	$Y_{abc} = 16,804$
9 d	$Y_d = 10,284$
10 ad	$Y_{ad} = 10,121$
11 bd	$Y_{bd} = 13,424$
12 abd	$Y_{abd} = 12,049$
13 cd	$Y_{cd} = 12,574$
14 acd	$Y_{acd} = 11,296$
15 bcd	$Y_{bcd} = 14,123$
16 $abcd$	$Y_{abcd} = 12,925$

O gráfico de probabilidade normal para os efeitos fatoriais do experimento agora considerado está ilustrado na Figura 7.

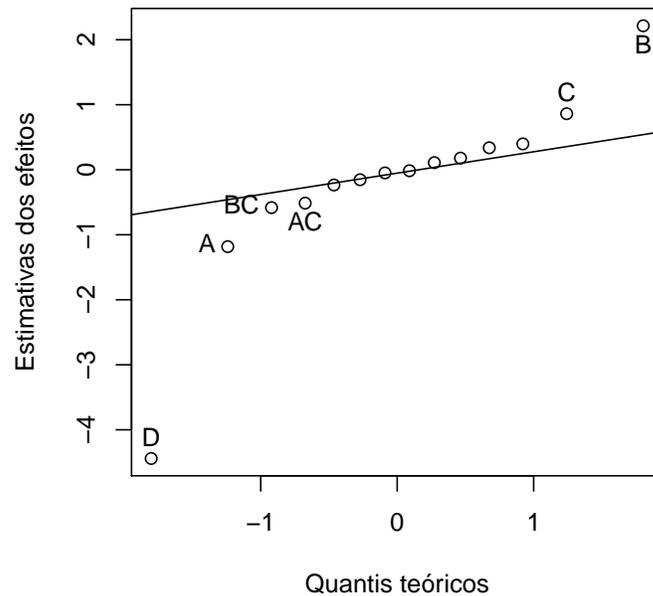


Figura 7 Gráfico de probabilidade normal para os efeitos fatoriais do experimento fatorial não repetido de resistência mecânica de junções adesivas

Analisando o gráfico de probabilidade normal, Figura 7, acredita-se que as interações de ordem superior ou igual a 3 são estatisticamente não significativas, a combinação das somas de quadrados desses efeitos será considerada como uma estimativa da soma de quadrados de resíduos. Naturalmente, estarão associados 5 graus de liberdade ao resíduo, sendo 4 graus de liberdade das interações de ordem 3, e 1 grau de liberdade da interação de ordem 4.

A soma de quadrados total, SQ_{T_c} , é calculada através da expressão (2.35), já usada na seção anterior. É importante observar que, nessa situação, experimento

não repetido, o valor de J é igual a 1, e a média dos tratamentos é igual à única observação. O cálculo da soma de quadrados total é desenvolvido abaixo.

$$\begin{aligned}
 SQ_{TC} &= \left(Y_{(1)}^2 + \dots + Y_{abcd}^2 \right) - \frac{(Y_{(1)} + \dots + Y_{abcd})^2}{16} = \\
 &= (14,979^2 + \dots + 12,925^2) - \frac{(14,979 + \dots + 12,925)^2}{16} = \\
 &= 3392,4580 - \frac{96,796^2}{16} = 111,4026
 \end{aligned}$$

As somas de quadrados dos efeitos principais e de interações de ordem 2 podem ser obtidas através de contrastes mutuamente ortogonais, tal como desenvolvido na seção anterior.

Alguns contrastes mutuamente ortogonais são:

$$\begin{aligned}
 C_A &= \frac{1}{8} (Y_a + Y_{ab} + Y_{ac} + Y_{abc} + Y_{ad} + Y_{abd} + Y_{acd} + Y_{abcd} - \\
 &\quad - Y_{(1)} - Y_b - Y_c - Y_{bc} - Y_d - Y_{bd} - Y_{cd} - Y_{bcd}) = \\
 &= \frac{1}{8} (14,578 + \dots + 12,925 - 14,979 - \dots - 14,123) = -1,1825
 \end{aligned}$$

$$\begin{aligned}
 C_B &= \frac{1}{8} (Y_b + Y_{ab} + Y_{bc} + Y_{abc} + Y_{bd} + Y_{abd} + Y_{bcd} + Y_{abcd} - \\
 &\quad - Y_{(1)} - Y_a - Y_c - Y_{ac} - Y_d - Y_{ad} - Y_{cd} - Y_{acd}) = \\
 &= \frac{1}{8} (18,207 + \dots + 12,925 - 14,979 - \dots - 11,296) = 2,2150
 \end{aligned}$$

$$\begin{aligned}
 C_{CD} &= \frac{1}{8} (Y_{(1)} + Y_a + Y_b + Y_{ab} + Y_{cd} + Y_{acd} + Y_{bcd} + Y_{abcd} - \\
 &\quad - Y_c - Y_{ac} - Y_{bc} - Y_{abc} - Y_d - Y_{ad} - Y_{bd} - Y_{abd}) = \\
 &= \frac{1}{8} (14,979 + \dots + 12,925 - 17,281 - \dots - 12,049) = 0,3978
 \end{aligned}$$

As estimativas de todos os contrastes mutuamente ortogonais referentes aos efeitos principais e de interações de ordem 2 dos fatores A , B , C e D são:

$$\begin{aligned} C_A &= -1,1825; & C_B &= 2,2150; & C_C &= 0,8623; \\ C_D &= -4,4413; & C_{AB} &= -0,0488; & C_{AC} &= -0,5135; \\ C_{AD} &= 0,1790; & C_{BC} &= -0,5820; & C_{BD} &= -0,1535; \\ & & C_{CD} &= 0,3978. \end{aligned}$$

Algumas somas de quadrados dos contrastes mutuamente ortogonais definidos acima são obtidas da seguinte maneira:

$$\begin{aligned} SQ_{C_A} &= \frac{1(C_A)^2}{0,25} = \frac{1(-1,1825)^2}{0,25} = 5,5932; \\ SQ_{C_B} &= \frac{1(C_B)^2}{0,25} = \frac{1(2,2150)^2}{0,25} = 19,6249; \\ SQ_{C_{CD}} &= \frac{1(C_{CD})^2}{0,25} = \frac{1(0,3978)^2}{0,25} = 0,6328. \end{aligned}$$

Os valores de todas as somas de quadrados dos contrastes mutuamente ortogonais referentes aos efeitos principais e de interações de ordem 2 dos fatores A , B , C e D estão apresentados na Tabela 31.

Para o resíduo, a soma de quadrados é obtida pela diferença entre a soma de quadrados total e as somas de quadrados dos efeitos principais e de interações de ordem 2.

$$\begin{aligned} SQ_{\text{Res}} &= SQ_{T_C} - SQ_{C_A} - \dots - SQ_{C_{CD}} = \\ &= 111,4026 - 5,5932 - \dots - 0,6328 = 1,0374. \end{aligned}$$

O quadrado médio dos efeitos principais e interações de ordem 2 é igual a sua soma de quadrados, pois, a cada contraste está associado 1 grau de liberdade.

$$QM_i = \frac{SQ_i}{1} = SQ_i, \quad i = A, \dots, CD.$$

E, para o resíduo,

$$QM_{\text{Res}} = \frac{SQ_{\text{Res}}}{5} = \frac{1,0374}{5} = 0,2075.$$

As estatísticas de teste, F_c , na análise da influência de efeitos principais e interações de ordem 2 na resistência mecânica de junções adesivas são obtidas pela expressão

$$F_c = \frac{QMC_i}{QM_{\text{Res}}}, \quad \text{para } i = A, B, \dots, CD.$$

Alguns valores da estatística de teste F_c são

$$F_c = \frac{QMC_A}{QM_{\text{Res}}} = \frac{SQ_{CA}}{QM_{\text{Res}}} = \frac{5,5932}{0,2075} = 25,96$$

para o efeito do fator A ;

$$F_c = \frac{QMC_B}{QM_{\text{Res}}} = \frac{SQ_{CB}}{QM_{\text{Res}}} = \frac{19,6249}{0,2075} = 94,58$$

para o efeito do fator B ;

$$F_c = \frac{QMC_{CD}}{QM_{\text{Res}}} = \frac{SQ_{CD}}{QM_{\text{Res}}} = \frac{0,6328}{0,2075} = 3,05$$

para o efeito de interação dos fatores C e D .

Os valores de todas as estatísticas de testes para avaliar as hipóteses referentes à influência dos fatores A , B , C e D na resistência mecânica de junções

adesivas estão apresentados na Tabela 31.

Na análise de variância esquematizada na Tabela 31, apenas os efeitos principais são significativos a um nível de 5% de significância, $F_{(1,5)} = 6,61$. No entanto, nota-se que o nível descritivo (valor - p) dos fatores AC e BC são próximos de 0,05, sugerindo resultados semelhantes do experimento completo, que já foram discutidos na subseção anterior.

Com o intuito de acrescentar a ilustração da análise estatística para o caso em que os efeitos de interações são não significativos considerar-se-á, rigorosamente, o nível de significância de 5%. Logo, o efeito de cada fator na tensão média de ruptura das junções adesivas não depende dos níveis dos demais fatores.

Tabela 31 Esquema da análise de variância do experimento de resistência mecânica de junções adesivas não repetido.

Fonte de Variação	GL	SQ	QM	F_c	valor-p
A	1	5,5932	5,5932	25,96	< 0,01
B	1	19,6249	19,6249	94,58	< 0,01
C	1	2,9739	2,9739	14,33	0,01
D	1	78,8988	78,8988	380,24	< 0,01
AB	1	0,0095	0,0095	0,05	0,83
AC	1	1,0547	1,0547	5,08	0,07
AD	1	0,1282	0,1282	0,62	0,47
BC	1	1,3549	1,3549	6,53	0,051
BD	1	0,0943	0,0943	0,45	0,53
CD	1	0,6328	0,6328	3,05	0,14
Resíduo	5	1,0374	0,2075		
Total	15	111,4026			

Como os contrastes referentes aos efeitos principais dos fatores A e D são negativos ($C_A = -1,1825$ e $C_D = -4,4414$), o nível alto desses fatores influenciam na redução da tensão média de ruptura das junções adesivas. Por outro lado, como os contrastes referentes aos efeitos principais dos fatores B e C são positivos ($C_B = 2,2150$ e $C_C = 0,8623$), o nível alto desses fatores acarretam no aumento da resistência mecânica das junções adesivas. Tais resultados também podem ser observados através dos gráficos de efeitos principais, Figuras 8 e 9, e pela tabela de médias de tensão média de ruptura em cada nível dos fatores analisados, Tabela 32.

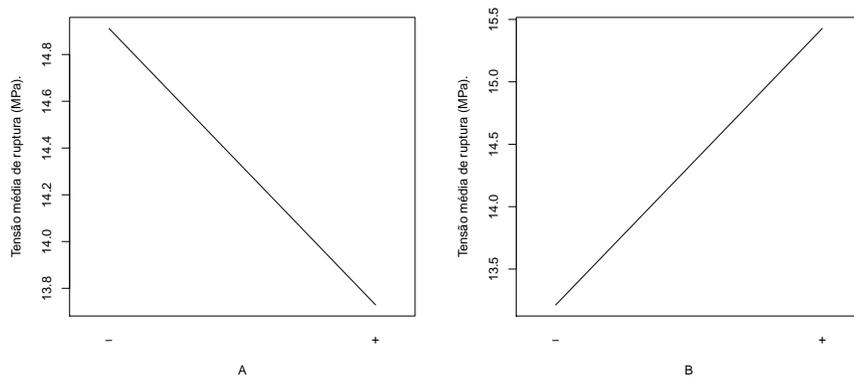


Figura 8 Gráficos de efeitos principais dos fatores A (tratamento superficial) e B (temperatura e tempo de cura), considerando a codificação “-” sem tratamento e “+” com tratamento, para o fator A ; “-” $22^{\circ}C$ e 24 horas e “+” $60^{\circ}C$ e 2 horas, para o fator B

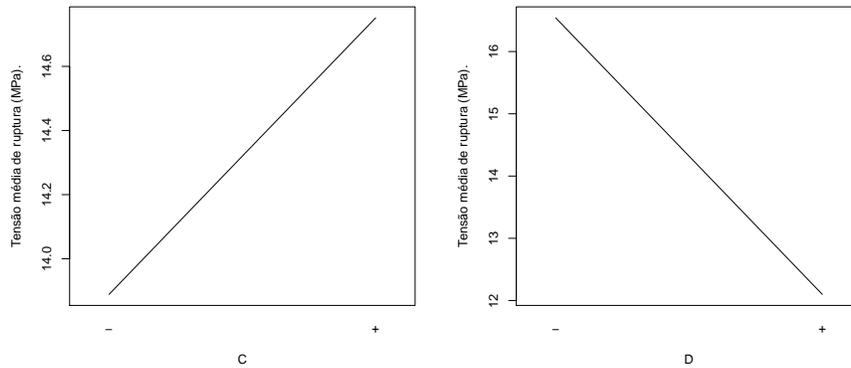


Figura 9 Gráficos de efeitos principais dos fatores C (rugosidade) e D (ativador), considerando a codificação “-” sem jateamento e “+” com jateamento, para o fator C ; “-” sem ativador e “+” com ativador, para o fator D

Tabela 32 Médias de tensão média de ruptura nos níveis dos fatores A , B , C e D .

Fator	Nível	
	-	+
A	14,91 u	13,73 v
B	13,21 u	15,43 v
C	13,89 u	14,75 v
D	16,54 u	12,12 v

u, v - em cada linha, médias seguidas de mesma letra minúscula não diferem estatisticamente (teste F a 5%).

Observando os resultados da análise do experimento não repetido com a do experimento original, nota-se que, a um nível de 5% de significância, alguns efeitos que são significativos na primeira situação experimental não o são na segunda.

A não utilização do princípio da repetição, em especial para o experimento considerado, acarreta em prejuízo na identificação de efeitos que podem ser significativos.

O planejamento de experimento fatorial não repetido é útil na escassez de unidades experimentais, no entanto, a não estimação direta do erro experimental, que é a finalidade da repetição, acarreta na obtenção de pouca informação comparado a experimentos repetidos. Assim, o seu uso deve ser feito com cautela.

4.3 Situação 3: experimento fatorial 2^4 com alguns efeitos fatoriais confundidos com efeito de blocos

Nesta situação, será considerado dez repetições e quatro blocos para cada repetição, em que a dimensão dos blocos é menor que a quantidade total de tratamentos do experimento.

Os efeitos das interações ACD , BCD e AB serão confundidos com os efeitos de blocos, sendo que as interações ACD e BCD são independentes e AB é a interação generalizada obtida de ACD e BCD .

Conforme a expressão (2.41), as combinações lineares para o confundimento dos efeitos das interações ACD , BCD e AB com o efeito de blocos são:

$$L_{ACD} = x_1 + x_3 + x_4;$$

$$L_{BCD} = x_2 + x_3 + x_4,$$

em que, x_i é o nível do i -ésimo fator que aparece em uma combinação de tratamentos particular, considerando a nomenclatura 0, 1 para os níveis dos fatores.

Os valores de $L_{ACD} \equiv (\text{mod}2)$ e $L_{BCD} \equiv (\text{mod}2)$ para os tratamentos (1), a , b e ab estão apresentados abaixo, e na Tabela 33 estão especificados os

tratamentos de cada bloco das repetições.

$$\begin{aligned}
 (1) & \begin{cases} L_{ACD} = 0 + 0 + 0 = 0 \equiv 0 \pmod{2} \\ L_{BCD} = 0 + 0 + 0 = 0 \equiv 0 \pmod{2} \end{cases} \\
 a & \begin{cases} L_{ACD} = 1 + 0 + 0 = 1 \equiv 1 \pmod{2} \\ L_{BCD} = 0 + 0 + 0 = 0 \equiv 0 \pmod{2} \end{cases} \\
 b & \begin{cases} L_{ACD} = 0 + 0 + 0 = 0 \equiv 0 \pmod{2} \\ L_{BCD} = 1 + 0 + 0 = 1 \equiv 1 \pmod{2} \end{cases} \\
 ab & \begin{cases} L_{ACD} = 1 + 0 + 0 = 1 \equiv 1 \pmod{2} \\ L_{BCD} = 1 + 0 + 0 = 1 \equiv 1 \pmod{2} \end{cases}
 \end{aligned}$$

Tabela 33 Tratamentos designados aos blocos de cada repetição do experimento de resistência mecânica de junções adesivas, para o confundimento dos efeitos ACD , BCD e AB com efeito de blocos.

Bloco 1	Bloco 2	Bloco 3	Bloco 4
$(0, 0)$	$(1, 0)$	$(0, 1)$	$(1, 1)$
(1)	a	b	c
cd	bc	ac	d
abc	bd	ad	ab
abd	acd	bcd	$abcd$

A massa de dados para esse experimento é a mesma da Tabela 20, sendo que os tratamentos são designados para um dos quatro blocos, em cada repetição, conforme a Tabela 33.

A análise estatística para esta situação será apenas para verificar a significância dos efeitos principais, de interações e de blocos, sem considerar o desdobramento das interações caso estas sejam significativas.

Os cálculos das somas de quadrados de todas as fontes de variações, exceto blocos e resíduo, são os mesmos do experimento da situação 1. A soma de quadrados de blocos pode ser calculada através da expressão (2.42). É importante lembrar que em cada repetição há 4 blocos, e no experimento 40 blocos.

$$\begin{aligned}
 SQ_{bl} &= \frac{\sum_{j=1}^{40} \left(\sum_{i(j)=1}^4 Y_{i(j)j} \right)^2}{4} - \frac{\left(\sum_{j=1}^{40} \sum_{i(j)=1}^4 Y_{i(j)j} \right)^2}{4 \cdot 40} = \\
 &= \frac{(14,9 + \dots + 12,06)^2 + \dots + (18,52 + \dots + 12,02)^2}{4} - \\
 &= \frac{(14,9 + \dots + 12,06 + \dots + 18,52 + \dots + 12,02)^2}{4 \cdot 40} = \\
 &= \frac{53,49^2 + \dots + 60,16^2}{4} - \frac{2291,22^2}{160} = \\
 &= 32895,4369 - 32810,5568 = 84,8801
 \end{aligned}$$

A soma de quadrados de resíduos é obtida pela diferença da soma de quadrados total, SQ_{T_c} , e as somas de quadrados das demais fontes de variações. As somas de quadrados dos efeitos confundidos com blocos não são utilizadas no cálculo da soma de quadrados de resíduos.

$$\begin{aligned}
 SQ_{Res} &= SQ_{T_c} - SQ_A - \dots - SQ_{ABCD} - SQ_{bl} = \\
 &= 1425,4204 - 55,9323 - \dots - 0,0093 - 84,8801 = \\
 &= 230,2027
 \end{aligned}$$

Em relação aos graus de liberdade das fontes de variações, tem-se que, 1 grau de liberdade para cada efeito principal ou de interação, 39 para o efeito de blocos, 159 para o total corrigido e, por diferença, 108 graus de liberdade para o

resíduo.

Os quadrados médios dos efeitos principais e de interações são iguais as suas somas de quadrados, e dos efeitos de blocos e resíduos são apresentados a seguir.

$$QM_{bl} = \frac{SQ_{bl}}{J - 1} = \frac{84,8801}{39} = 2,1764$$

$$QM_{Res} = \frac{SQ_{Res}}{108} = \frac{230,2027}{108} = 2,1315$$

As estatísticas de teste, F_c , para verificar a influência do efeito principal do fator A e de blocos na resistência mecânica de junções adesivas são:

$$F_c = \frac{QM_A}{QM_{Res}} = \frac{55,9323}{2,1315} = 26,24$$

para o efeito do fator A ;

$$F_c = \frac{QM_{bl}}{QM_{Res}} = \frac{2,1764}{2,1315} = 1,02$$

para o efeito de blocos.

Todas as estatísticas de teste, F_c , do experimento de resistência mecânica de junções adesivas, com os efeitos ACD , BCD e AB confundidos com efeito de blocos, estão apresentadas na Tabela 34, referente ao esquema da análise de variância.

Para um nível de 5% de significância, o valor F_c para os efeitos A , B , C , D , AC e BC é maior que $F_{(1,108)} = 3,93$. Sendo assim, somente os testes referentes aos efeitos citados anteriormente são significativos, isto é, esses fatores são influentes na resistência mecânica de junções adesivas.

O teste F para a influência do efeito de blocos na resistência mecânica das junções adesivas apresentou valor F_c menor que o valor crítico de F para o nível

de 5% de significância, $F_{(39,108)} = 1,51$, ou seja, o fator controlado por blocos não influencia na resistência mecânica das junções adesivas.

Note-se que os resultados agora obtidos são os mesmos da análise do experimento original. Com a ressalva que, quanto menor for a dimensão dos blocos, mais blocos por repetição serão necessários, acarretando em mais efeitos fatoriais confundidos com efeito de blocos e redução dos graus de liberdade do resíduo.

Tabela 34 Esquema da análise de variância do experimento de resistência mecânica de junções adesivas em esquema fatorial 2^4 com os efeitos ACD , BCD e AB confundidos com o efeito de blocos.

Fonte de Variação	GL	SQ	QM	F_c	valor-p
<i>A</i>	1	55,9323	55,9323	26,24	< 0,01
<i>B</i>	1	196,2490	196,2490	92,07	< 0,01
<i>C</i>	1	29,7390	29,7390	13,95	< 0,01
<i>D</i>	1	788,9880	788,9880	370,16	< 0,01
<i>AC</i>	1	10,5473	10,5473	4,95	0,03
<i>AD</i>	1	1,2816	1,2816	0,60	0,44
<i>BC</i>	1	13,5490	13,5490	6,36	0,01
<i>BD</i>	1	0,9425	0,9425	0,44	0,51
<i>CD</i>	1	6,3282	6,3282	2,97	0,09
<i>ABC</i>	1	4,5765	4,5765	2,15	0,15
<i>ABD</i>	1	2,1950	2,1950	1,03	0,31
<i>ABCD</i>	1	0,0093	0,0093	0,004	0,95
Blocos	39	84,8801	2,1764	1,02	0,31
Resíduo	108	230,2027	2,1315		
Total	159	1425,4204			

4.4 Considerações finais

4.4.1 Discussão acerca do desempenho das três situações

Comparando os resultados da análise do experimento da situação 2 com o experimento original (situação 1), tem-se que, o experimento fatorial não repetido necessita de menos esforço e custo para a sua realização, no entanto, os resultados da análise estatística são menos informativos em relação ao experimento original.

Em relação ao experimento da situação 3, em que alguns efeitos fatoriais são confundidos com efeito de blocos, as fontes de variações significativas na análise estatística são iguais a do experimento original (situação 1). O que indica uma boa alternativa para situações de escassez de unidades experimentais homogêneas. No entanto, é importante utilizar a técnica do confundimento com cautela pois, quanto menor for a dimensão dos blocos, mais blocos por repetição serão necessários acarretando em mais efeitos fatoriais confundidos com efeito de blocos e, conseqüentemente, maior redução dos graus de liberdade do resíduo.

Os dois resultados discutidos anteriormente estão sintetizados na Tabela 35.

Tabela 35 Observações sobre as três situações experimentais consideradas.

	Esforço e custo para realizar o experimento	Informação resultante
Situação 1	maior	maior
Situação 2	menor	menor
Situação 3	médio	médio

4.4.2 Sugestões para pesquisas futuras

Para a experimentação aplicada em processos industriais, sugere-se ampliar a sua discussão, em especial para os planejamentos envolvendo arranjos ortogonais de Taguchi, blocos incompletos, delineamentos não ortogonais, efeitos aleatórios, fatoriais fracionados, misturas, parcelas subdivididas, superfícies de respostas etc.

Em relação aos métodos abordados nesta dissertação, em especial ao fatorial não repetido, sugere-se avaliar a eficiência dos métodos de análises preliminares, tais como, os gráficos de probabilidades normais e semi-normais.

5 CONCLUSÃO

Em relação à questão de pesquisa “quais seriam exemplos de análises básicas para experimentação na indústria que elucidam peculiaridades próprias desta área de aplicação?”, tem-se que:

- (i) o experimento fatorial 2^k é sugerido como um dos planejamentos interessantes para a indústria, visto que o seu planejamento e análise é relativamente simples se comparado a experimentos que envolvem mais de dois níveis para cada fator;
- (ii) outra vantagem apresentada dos experimentos fatoriais 2^k é que eles são úteis quando o objetivo é examinar um grande número de fatores para determinar quais podem ser os mais importantes, ou quando o tempo ou o recurso para o experimento são escassos;
- (iii) das três situações estudadas, observa-se que o confundimento de efeitos fatoriais com efeito de blocos mostrou-se uma solução intermediária para a experimentação na indústria quando o problema de esforço e custo é limitante.

REFERÊNCIAS

- ANTONY, J. **Design of experiments for engineers and scientists**. Oxford: Butterworth-Heinemann, 2003. 152 p.
- BANZATTO, D. A.; KRONKA, S. N. **Experimentação agrícola**. 4. ed. Jaboticabal: Funep, 2008. 237 p.
- BARROS NETO, B.; SCARMINIO, I. S.; BRUNS, R. E. **Como fazer experimentos**: pesquisa e desenvolvimento na ciência e na indústria. 4. ed. Porto Alegre: Bookman, 2010. 414 p.
- BOX, G. E. P. Evolutionary operation: a method for increasing industrial productivity. **Journal of the Royal Statistical Society**, v. 6, n. 2, pp. 81-101, 1957.
- BOX, G. E. P.; HUNTER, J. S.; HUNTER, W. G. **Statistics for experimenters**: design, innovation, and discovery. 2. ed. Hoboken: John Wiley & Sons, 2005. 633 p.
- BOX, G. E. P.; MEYER, R. D. An analysis for unreplicated fractional factorials. **Technometrics**, v. 28, n. 1, p. 11-18, 1986.
- CASELLA, G.; BERGER, R. **Inferência estatística**. São Paulo: Cengage Learning, 2010. 588 p.
- CLARKE, B. R. **Linear models**: the theory and application of analysis of variance. Hoboken: John Wiley & Sons, 2008. 241 p.
- COCHRAN, W. G. The distribution of quadratic forms in a normal system, with application to the analysis of covariance. **Proceedings of the Cambridge Philosophical Society**, v. 30, p. 178-191, 1934.
- COCHRAN, W. G. When the assumptions for the analysis of variance are not satisfied. **Biometrics**, v. 3, n. 1, p. 22-38, 1947.

COSTA, V. M. G.; CUNHA, J. C. da. A universidade e a capacitação tecnológica das empresas. **Revista de Administração Contemporânea** v. 5 n. 1, p. 61-81, 2001.

COX, D. R.; REID, N. **The theory of the design of experiments**. Florida: CHAPMAN & HALL/CRC, 2000, 314p.

CZITROM, V. One-factor-at-a-time versus designed experiments. **The American Statistician**, v. 53, n. 2, p. 126-131, 1999.

DANIEL, C. Use of half-normal plots in interpreting factorial two-level experiments. **Technometrics**, v. 1, n. 4, p. 311-341, 1959.

DANIEL, C. **Applications of statistics to industrial experimentation**. New York: John Wiley & Sons, 1976. 294 p.

DANIEL, C. Factorial one-factor-at-a-time experiments. **The American Statistician**, v. 48, n. 2, p. 132-135, 1994.

FISHER, R. A. **The design of experiments**. 9. ed. New York: HAFNER PRESS, 1971. 248 p.

GALDAMEZ, E. V. C.; CARPINETTI, L. C. R. Aplicação das técnicas de planejamento e análise de experimentos no processo de injeção plástica. **Gestão & Produção**, v. 11, n. 1, p. 121-134, 2004.

GARCIA-DIAZ, A.; PHILLIPS, D. T. **Principles of experimental design and analysis**. London: Chapman & Hall, 1995. 409 p.

GENTLE, J. E. **Matrix algebra: theory, computations, and applications in statistics**. New York: Springer, 2007. 528 p.

GIESBRECHT, F. G.; GUMPERTZ, M. L. **Planning, construction, and statistical analysis of comparative experiments**. Hoboken: John Wiley & Sons, 2004. 693 p.

GOOS, P.; LANGHANS, I.; VANDEBROEK, M. Practical inference from industrial split-plot designs. **Journal of Quality Technology**, v. 38, n. 2, p. 162-179, 2006.

GRAYBILL, F. A. **Theory and application of the linear model**. Belmont: Duxbury Press, 1976. 704 p.

HAMADA, M.; BALAKRISHNAN, N. Analyzing unreplicated factorial experiments: a review with some new proposals. With discussions. **Statistica Sinica**, v. 8, n. 1, p. 1-41, 1988.

HINKELMANN, K.; KEMPTHORNE, O. **Design and analysis of experiments: introduction to experimental design**. Hoboken: John Wiley & Sons, 2008. 631 p.

ILZARBE, L. *et al.* Practical applications of design of experiments in the field of engineering: a bibliographical review. **Quality and Reliability Engineering International**, v. 24, p. 417-428, 2008.

IRWIN, J. O. Mathematical theorems involved in the analysis of variance. **Journal of the Royal Statistical Society**, v. 94, n. 2, p. 284-300, 1931.

IVERSEN, G. R. Analysis of variance. In: LOVRIC, M. (Ed.) **International encyclopedia of statistical science**. Springer-Verlag Berlin Heidelberg, 2011. 1673 p.

JONES, B.; NACHTSHEIM, C. J. Split-plot designs: what, why, and how. **Journal of Quality Technology** v. 41, n. 4, p. 340-361, 2009.

MASON, R. L.; GUNST, R. L.; HESS, J. L. **Statistical design and analysis of experiments** with applications to engineering and science. 2. ed. Hoboken: John Wiley & Sons, 2003. 728 p.

MONTGOMERY, D. C. **Introdução ao controle estatístico da qualidade**. Tradução Ana Maria Lima de Farias. 4. ed. Rio de Janeiro: LTC, 2004, 516p.

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

R Development Core Team. **R: A language and environment for statistical computing**. R Foundation for Statistical Computing. Vienna: R Foundation for Statistical Computing, 2011. Disponível em: <<http://www.r-project.org>>. Acesso em: 1 jan. 2012.

RAPINI, M. S. Interação universidade-empresa no Brasil: evidências do Diretório dos Grupos de Pesquisa do CNPq. **Estudos Econômicos** v. 37, n. 1, p. 211-233, 2007.

RENCHER, A. C.; SCHAALJE, G. B. **Linear models in statistics**. 2. ed. Hoboken: John Wiley & Sons, 2008. 672 p.

SANTOS, D. J. **Estudo experimental da resistência mecânica de junções adesivas**. 2007. 140 p. Dissertação (Mestrado em Engenharia Mecânica) - Escola Politécnica da Universidade de São Paulo, São Paulo, 2007.

SEARLE, S. R. **Linear models**. New York: John Wiley & Sons, 1971. 532 p.

STORCK, L. *et al.* **Experimentação vegetal**. 2. ed. Santa Maria: Ed. UFSM, 2006. 198 p.

SUDARSANAM, N.; FREY, D. D. Using ensemble techniques to advance adaptive one-factor-at-a-time experimentation. **Quality and Reliability Engineering International**, v. 27, n. 7, p. 947-957, 2011.

TANCO, M. *et al.* How is experimentation carried out by companies? A survey of three european regions. **Quality and Reliability Engineering International**, v. 24, p. 973-981, 2008.

VILLENAVE, J. **Assemblage par collage**. Paris: DUNOD, 2005. 319 p.

WINER, B. J. **Statistical principles in experimental design**. New York:

McGraw-Hill, 1962. 672 p.

WU, C. F. J.; HAMADA, M. S. **Experiments** planning, analysis, and optimization. 2. ed., Hoboken: John Wiley & Sons, 2009. 716 p.