ESTATÍSTICA EXPERIMENTAL

GUIA DE ESTUDOS

Paulo César Lima

Renato Ribeiro de Lima

Ficha catalográfica preparada pela Divisão de Processos Técnicos da Biblioteca Universitária da UFLA

Espaço a ser preenchido pela biblioteca

[A ser preenchido posteriormente]

Espaço a ser preenchido pelo CEAD

ÍNDICE

UNIDADE 1 - PESQUISA CIENTÍFICA E ESTATÍSTICA

Nesta primeira unidade vamos apresentar os conceitos sobre Ciência, Pesquisa Científica e os Fundamentos da Estatística envolvidos na Experimentação

Anotações:

UNIDADE 1 – PESQUISA CIENTÍFICA E ESTATÍSTICA

1.1 CIÊNCIA E MÉTODO CIENTÍFICO

Existem muitas definições e conceitos para "ciência". Escolhemos um conceito mais geral para que você possa começar a entender a importância da Estatística no desenvolvimento da Ciência.

*"A **Ciência** é um conjunto de conhecimentos obtidos através do Método Científico, organizados e verificáveis".*

As afirmações verificáveis são aquelas que podemos comprovar através de observações. As afirmações que não puderem ser comprovadas em experiências não constituem ciência. De uma forma geral também, podemos conceituar o ***Método Científico*** como:

"Um conjunto de normas e procedimentos básicos necessários para a realização de experiências com o objetivo de produzir conhecimento".

Anotações:

1.2 PESQUISA CIENTÍFICA

Se quisermos realizar uma ***Pesquisa Científica*** precisamos seguir as etapas do método científico. Estas etapas são apresentadas na Figura 1.1.

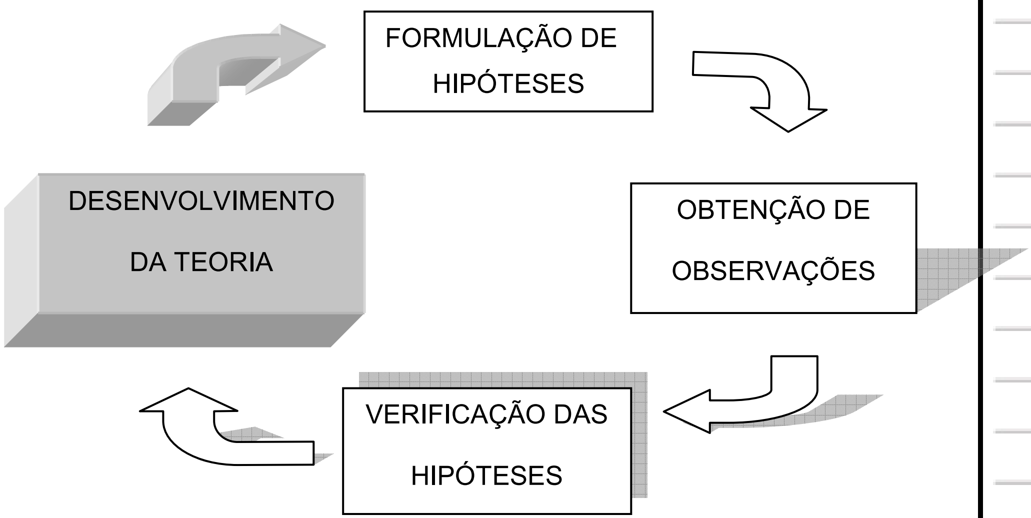

FIGURA 1.1. O Método Científico.

Observando a Figura 1.1 você pode entender como trabalham os pesquisadores:

⇨ Inicialmente, com o conhecimento do pesquisador, com a revisão de literatura e com as discussões sobre o tema, planejam todos os detalhes a serem estudados e definem os objetivos a serem atingidos;

Anotações:

⇨ Em sequência, procuram formular suposições (hipóteses) como possíveis soluções para o problema;

⇨ Através da observação dos processos envolvidos, obtêm informações que serão utilizadas para verificar se as suposições formuladas são explicações para o problema;

⇨ Dessa forma, os resultados das pesquisas podem tornar-se novos conhecimentos.

Podemos resumir os passos de uma pesquisa científica em:

- Definição do Problema
- Formulação das Hipóteses
- Obtenção das Observações
- Análise das Observações
- Interpretação dos Resultados
- Publicação das Conclusões

1.3 A ESTATÍSTICA NA PESQUISA CIENTÍFICA

A ***Estatística*** está inserida no Método Científico. Possui fundamentos teóricos que são comuns a todas as áreas do conhecimento e metodologias de planejamento de pesquisas e de análises de informações aplicadas a todas estas áreas.

Anotações:

Vamos falar dos fundamentos da estatística associados a cada etapa do Método Científico.

1.3.1 Hipóteses Estatísticas

Nas pesquisas científicas estamos interessados em explicar as relações causa-efeito entre características dos elementos envolvidos no estudo em foco.

As hipóteses sobre as relações causa-efeito geralmente surgem do conhecimento teórico relacionado ao problema em estudo, mas podem surgir também com base em literatura ou pela observação do fenômeno.

Lembre-se de que o sucesso de um estudo científico inicia-se com o maior entendimento e definição do problema a ser pesquisado, com a habilidade para formular os objetivos a serem atingidos e com a clareza na formulação das hipóteses a serem testadas.

A ***Hipótese Estatística*** é uma formulação provisória de resposta ao problema investigado.

Toda hipótese estatística deve ser passível de um teste de confirmação em experiências. Se os resultados obtidos em várias experiências não contradizem a hipótese ela será aceita como conhecimento científico.

Anotações:

1.3.2 Amostragem

A maioria das pesquisas envolve populações infinitas ou extremamente grandes em que o estudo de todos os seus integrantes (***censo***) é inviável em função de a população ser infinita ou dos altos custos e do tempo necessário demandados.

A teoria da amostragem estatística permite que tomemos uma parte representativa da população para a obtenção das informações necessárias às pesquisas.

Designamos por ***característica*** à propriedade que distingue ou caracteriza cada elemento de uma amostra e vamos utilizar o termo ***dado*** para designar cada observação ou cada medida efetuada nas características destes elementos.

As características dos elementos de uma população ou amostra podem ser consideradas como na Figura 1.2.

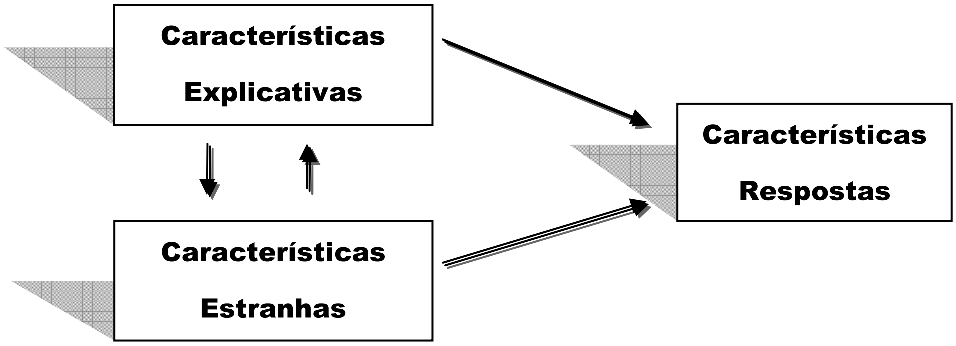

FIGURA 1.2 Grupos de Características dos Indivíduos.

Em uma experiência, as características causais (***explicativas***) são aquelas responsáveis pelos efeitos que procuramos observar ou mensurar (***respostas***). As características estranhas também são características causais, mas desconhecidas ou consideradas como de menor importância em nosso estudo.

1.3.3 Estimação e Testes de Hipóteses

O objetivo das pesquisas é conhecer sobre os parâmetros da população Quando não for possível realizar um censo podemos empregar a amostragem. Através de amostras, não é possível determinar os parâmetros da população então nós usamos a inferência estatística para generalizar resultados obtidos em amostras para a população.

Anotações:

Você sabe quais são as diferenças entre Parâmetro, Estimativa e Estimador?

Anotações:

Em primeiro lugar, vamos relembrar o que é ***parâmetro***: é o valor de uma função dos indivíduos de uma população para descrever uma característica qualquer.

Ao valor obtido de uma amostra para uma característica denominada Y através de uma função qualquer, f(Y), chamamos ***estimativa*** do respectivo parâmetro populacional e, à função f utilizada para obtê-lo chamamos ***estimador*** daquele parâmetro.

1.3.3.1 Estimação

A diferença entre uma estimativa e o parâmetro é chamada de ***erro de estimação***. Cada indivíduo de uma amostra é a expressão da ação simultânea de muitas características, o que implica na existência de variabilidade entre os elementos da amostra. Podemos expressar o erro de estimação como uma função da variabilidade inerente à característica em estudo:

- Se a variabilidade for pequena, repetidas amostras da população, provavelmente, fornecerão estimativas similares implicando em que a estimativa obtida de uma amostra qualquer provavelmente esteja mais próxima do valor do parâmetro;

- Se a variabilidade for grande a estimativa obtida de uma amostra qualquer provavelmente estará mais distante do valor do parâmetro.

Anotações:

1.3.3.2 Testes de Hipóteses

O objetivo da análise estatística é avaliar as relações causa-efeito entre características dos elementos de uma amostra e possibilitar a inferência para a população. Isto é feito através dos **testes de hipóteses**.

Você sabia que a estatística nunca prova nada*?*

A estatística, considerando o erro de estimação, permite dimensionar a confiança que você pode ter no resultado de uma pesquisa ou então, a probabilidade de que você tome uma decisão incorreta ao aceitar ou ao rejeitar uma hipótese estatística.

Para realizar um Teste de Hipótese siga os seguintes passos:

- Formular uma hipótese a ser testada (H_0) e uma hipótese alternativa (H_a);
- Especificar do grau de confiança desejado (γ=1-α);
- Escolher uma estatística de teste que avalie os desvios de H_0 e que tenha distribuição amostral conhecida (D_t);
- Calcular a estimativa do parâmetro através da amostra (D_c);
- Estabelecer uma Regra de Decisão sobre D_c e D_t.

Anotações:

A regra de decisão a ser usada no último passo admite quatro resultados possíveis, conforme apresentado na Tabela 1.1.

TABELA 1.1 Decisões e Tipos de Erros Associados aos Testes de Hipóteses.

H_0	DECISÃO	RESULTADO
Verdadeira	Não rejeitar H_0	CORRETO
Verdadeira	Rejeitar H_0	**ERRO TIPO I**
Falsa	Rejeitar H_0	CORRETO
Falsa	Não rejeitar H_0	**ERRO TIPO II**

A probabilidade de se cometer o Erro Tipo I (rejeitar uma hipótese verdadeira) em um teste de hipótese qualquer é representada pela letra grega α. Usualmente chamamos α de ***nível de significância***. O **nível de confiança** é dado por $1 - \alpha$.

Anotações:

A probabilidade de se cometer o Erro Tipo II (aceitar uma hipótese falsa) é representada por β.

Desafio:

Você é o pesquisador. Então:

a- Descreva um estudo simples a ser pesquisado na sua área. Descreva todos os detalhes conhecidos sobre o problema proposto. Formule os principais objetivos a serem alcançados.

b- Formule uma hipótese a ser verificada no estudo,

c- Indique as observações a serem coletadas visando a verificação da hipótese formulada.

UNIDADE 2 – CONCEITOS E PRINCÍPIOS BÁSICOS DA ESTATÍSTICA EXPERIMENTAL

Nesta unidade são apresentados os principais conceitos e os princípios fundamentais da experimentação

UNIDADE 2 – CONCEITOS E PRINCÍPIOS BÁSICOS DA ESTATÍSTICA EXPERIMENTAL

Anotações:

A ***Estatística Experimental*** trata das técnicas apropriadas ao planejamento e às análises dos dados de Experimentos.

Os fundamentos dessas técnicas foram apresentados por R.A. Fisher (1890-1962) em *Statistical methods for research workers* (1925). Este livro era direcionado às áreas de biologia e agricultura e era essencialmente aplicado.

2.1 CONCEITOS GERAIS

Alguns termos e expressões são características da área experimental e frequentemente utilizadas em todas as etapas da experimentação. Vários termos utilizados têm origem na área agrícola e permanecem em uso até os dias de hoje, mas com conotação mais geral.

2.1.1 Experimento

Anotações:

Um **experimento** é a realização de um procedimento em que algumas características explicativas são controladas pelo experimentador.

2.1.2 Variáveis

A função numérica que estabelece a correspondência um a um entre as manifestações de uma característica e os valores de um conjunto numérico é denominada **variável**.

Quando começarmos o estudo de um problema, iremos encontrar uma grande quantidade de variáveis envolvendo o processo em foco. A definição da relevância dessas variáveis e das hipóteses envolvendo as relações causa-efeito entre elas é uma etapa fundamental nesta fase de planejamento da pesquisa.

As variáveis relevantes consideradas no estudo de um problema podem ser agrupadas em:

- ➔ Variáveis causais (c_i, i=1, 2,..., n): aquelas que influenciam o desempenho dos elementos ou indivíduos;

➔ Variáveis efeitos (e_j, j=1, 2,..., m): aquelas que exprimem o desempenho dos mesmos;

➔ Variáveis irrelevantes (causais ou efeitos): todas as outras variáveis inerentes aos elementos do processo mas sem interesse no estudo.

Anotações:

A Figura 2.1 representa as variáveis em uma pesquisa.

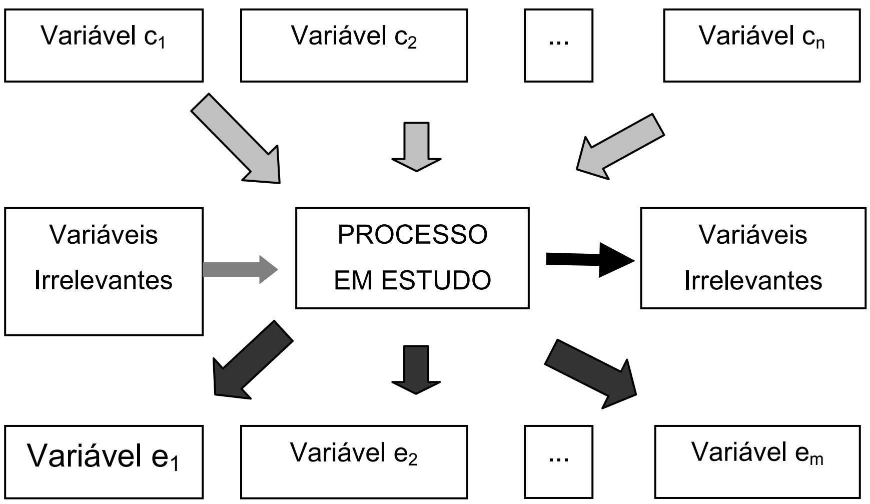

FIGURA 2.1 Variáveis Causais (c_i), Variáveis Efeitos (e_j) e Outras Variáveis.

2.1.3 Variável Fator e Variável Resposta

Anotações:

No planejamento de um experimento devemos escolher a variável causal cujos efeitos queremos observar no experimento e as variáveis efeitos que iremos mensurar conforme os objetivos da pesquisa.

Às variáveis causais cujos efeitos queremos estudar denominamos ***fatores***. Os valores que um fator assume são suas ***categorias*** (tipos, se o fator for uma variável qualitativa e níveis, se o fator for quantitativo).

Às variáveis que iremos mensurar ou observar designamos ***respostas*** e valores obtidos para essas respostas são os ***dados***.

As outras variáveis causais, que podem ser relevantes ou não, deverão ser controladas através das técnicas experimentais para que as respostas obtidas sejam a expressão apenas do fator ou fatores em estudo (Figura 2.2).

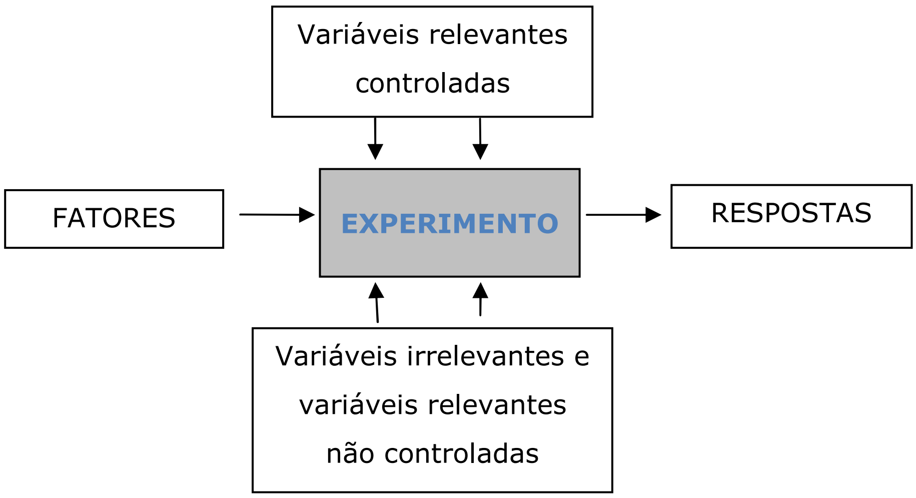

FIGURA 2.2 Representação das Variáveis em um Experimento

Anotações:

2.1.4 Tratamentos

Se quisermos estudar apenas um fator em um experimento, as categorias desse fator são denominadas ***tratamentos***. Para o caso queremos estudar dois ou mais fatores, os tratamentos são as combinações das categorias desses fatores.

2.1.5 Parcela Experimental

Unidade experimental ou ***parcela experimental*** é a quantidade de indivíduos de uma amostra em que aplicamos apenas um tratamento e tomamos uma única medida para cada variável resposta.

A quantidade de material ou de indivíduos que define uma parcela (tamanho da parcela) depende basicamente da variabilidade inerente à variável resposta.

Anotações:

Existem procedimentos estatísticos apropriados para a determinação do tamanho e da forma de parcelas experimentais, fundamentados na teoria da amostragem. No entanto, na prática, o tamanho da parcela de um experimento frequentemente é escolhido por analogia a outros ensaios de mesma natureza e realizado em condições experimentais semelhantes.

2.1.6 Bordadura

Quando existir a possibilidade do tratamento aplicado a uma parcela influenciar ou ser influenciado pelos tratamentos aplicados nas parcelas vizinhas, cada parcela deverá dispor de uma quantidade de material a mais para servir de proteção contra esta interferência. Este parte da parcela que servirá de proteção é denominada ***bordadura da parcela***.

Não podemos incluir a bordadura na obtenção dos dados das respostas. A área da parcela, excetuando-se a bordadura, é denominada ***área útil da parcela*** e apenas nela deverão ser efetuadas as avaliações.

Anotações:

Como exemplo vamos simular uma parcela constituída por 3 linhas de plantio de uma cultura qualquer, cada linha com 5 metros de comprimento. O espaçamento entre linhas de plantio é de 0,50 m e vamos colocar 10 plantas por metro linear. Iremos pesar as produções da linha central, eliminando-se 1 metro em cada extremidade das linhas. A Figura 2.3 mostra o detalhe desta parcela.

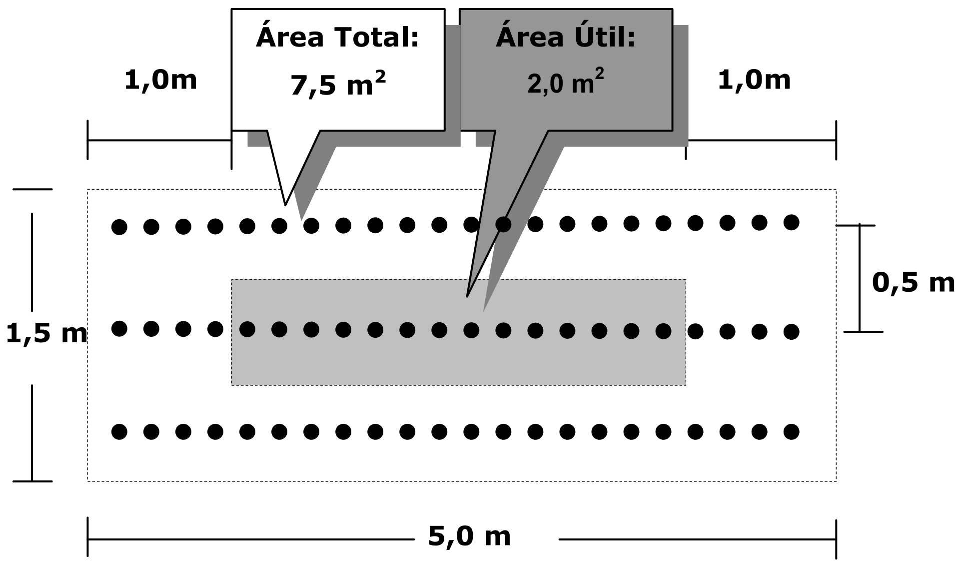

FIGURA 2.3 Detalhe de uma Parcela Experimental com Bordadura.

Anotações:

2.2 PRINCÍPIOS BÁSICOS DA EXPERIMENTAÇÃO

Os princípios básicos da experimentação são: ***repetição*** e ***aleatorização*.** São denominados princípios básicos porque são os fundamentos mínimos necessários a todo experimento.

2.2.1 Repetição

A ***repetição*** consiste na aplicação de cada tratamento a mais de uma parcela experimental.

A função da repetição é permitir a obtenção de uma estimativa da variabilidade atribuída a todas as variáveis consideradas irrelevantes ou não conhecidas e às variáveis explicativas que não foram controladas.

Outra função atribuída ao aumento do número de repetições é a diminuição do erro experimental (veja a seguir, no item 2.3.1).

Alguns procedimentos estatísticos permitem determinar o número necessário de repetições, geralmente considerando o grau de confiança requerido, a variabilidade do material experimental e os objetivos da pesquisa.

Anotações:

2.2.2 Aleatorização

A ***aleatorização*** ou ***casualização*** consiste em um conjunto de regras que define o processo de distribuição das parcelas na área experimental.

A função da casualização é evitar a tendenciosidade ou o viés dos efeitos das variáveis não controladas sobre os resultados obtidos no experimento permitindo que as estimativas e os testes de hipóteses sejam válidos.

Para exemplificar a utilização desses termos e expressões, considere o experimento apresentado no exemplo seguinte.

Exemplo 2.1

Um engenheiro, interessado em estudar a resistência de fibras sintéticas utilizadas na confecção de vestuário, decidiu utilizar diferentes quantidades de algodão já que *é sabido, de pesquisas anteriores, que a resistência de fibras sintéticas aumenta com a inclusão de algodão. Como o produto final deve conter de 10 a 40% de algodão, devido a outras características importantes para a qualidade do produto, foram escolhidas as quantidades de 15, 20, 25, 30 e 35% de algodão. Também, decidiu-se testar cinco amostras de cada nível de algodão tomando, como amostra, um atado de fibras com 10 centímetros de diâmetro. As avaliações foram feitas em apenas uma máquina e por um único técnico.*

Fonte: Montgomery, D.C.

No exemplo 2.1, temos apenas um **fator** que são os teores de algodão na fibra com as **categorias** (ou níveis): 15, 20, 25, 30 e 35%. Como as categorias deste fator são expressas em uma escala intervalar, o fator é denominado quantitativo. Quando as categorias de um fator são expressas em uma escala nominal, o fator é denominado qualitativo.

A tensão ao rompimento corresponde à variável **resposta**. Os valores obtidos para a resistência das amostras são os **dados**.

Os **tratamentos** são: T_1 = 15% de algodão na fibra, T_2 = 20%, T_3 = 25%, T_4 = 30% e T_5 = 35% de algodão na fibra.

Anotações:

Atenção:

Anotações:

Se em outro experimento os fatores fossem Reagentes (A e B) e uso ou não de um Catalisador, os tratamentos seriam: T_1 = reagente A sem catalisador; T_2 = reagente A com catalisador; T_3 = reagente B sem catalisador e T_4 = reagente B com catalisador (correspondentes às combinações dos reagentes com o catalisador).

No Exemplo 2.1, cada **parcela** corresponde a um amarrado de fibras com dez centímetros de diâmetro e, neste caso, não foi necessário usar bordadura porque uma amostra sempre estava separada das outras.

Nesse exemplo foram utilizadas 5 **repetições** (cinco amostras com cada teor de algodão), perfazendo 25 parcelas experimentais.

A aleatorização consistiu em submeter as amostras ao teste de resistência em uma ordem determinada ao acaso.

Anotações:

2.3 CONCEITOS FUNDAMENTAIS

Os ***conceitos fundamentais*** da Estatística Experimental são aqueles necessários à compreensão da metodologia estatística utilizada no planejamento, nas análises e na inferência dos resultados dos experimentos. Estes conceitos estão intimamente relacionados com a teoria da Amostragem, com a Inferência Estatística e com a modelagem do problema em estudo.

2.3.1 Erro Experimental

Vamos lembrar que as alterações nas respostas provocadas pelos fatores são o objeto da nossa pesquisa, mas outras variáveis consideradas de menor importância ou aquelas cujo controle não foi eficiente e, ainda, outras variáveis desconhecidas do pesquisador, também podem provocar mudanças nas respostas que vamos obter.

Estes efeitos estarão confundidos com os efeitos do fator em estudo provocando a variação que vamos observar nas respostas. Essa variação indesejada é denominada ***Erro Experimental***.

O erro experimental pode fazer com que as inferências sejam tendenciosas ou até mesmo inviabilizar a utilização dos resultados das pesquisas.

O erro experimental também pode ser considerado como o desvio de ajuste (erro de estimação) do modelo proposto para explicar o efeito dos fatores sobre as respostas.

Em se tratando de estudos de amostragem, ele sempre estará presente e assim as técnicas de planejamento, controle e de condução dos experimentos devem visar à minimização da quantidade dessas variáveis e de seus efeitos nas respostas, tornando constantes ou irrelevantes suas manifestações nos resultados.

No Exemplo 2.1, algumas variáveis responsáveis pelo **erro experimental** poderiam ser: regulagem da máquina, variações na matéria prima e na preparação das fibras, variações no diâmetro da amostra e muitas outras relacionadas, principalmente com a condução do ensaio.

Quando o pesquisador busca conhecer as variáveis envolvidas em um experimento e homogeneizar as condições experimentais através de, por exemplo, seleção do material experimental, treinamento de

Anotações:

pessoal, aperfeiçoamento da técnica de condução além de outros, ele visa minimizar o erro experimental.

Anotações:

A precisão é função da variabilidade entre as medidas efetuadas nas parcelas. Precisão alta implica em que as observações iram fornecer estimativas mais próximas dos valores da população.

Um processo de mensuração pode apresentar outro tipo de erro – o erro sistemático. Neste caso, as medidas obtidas são tendenciosas ou viesadas, subestimando ou superestimando os valores verdadeiros. Este tipo de erro não contribui para o erro experimental e, portanto, não afeta a precisão do experimento, mas reduz a eficiência.

2.3.2 Controle Local

O ***controle local*** consiste no agrupamento das parcelas de um experimento de maneira que os efeitos de variáveis estranhas, mas conhecidas, não sejam confundidos com os efeitos dos fatores.

Geralmente os grupos de parcelas são denominados blocos e devem ser homogêneos, mas podem variar entre si.

Anotações:

No Exemplo 2.1, supondo que, ao invés de apenas um técnico fossem utilizados quatro pessoas diferentes para conduzir o ensaio, poderíamos supor que a diferença de habilidade entre os quatro técnicos afetaria aleatoriamente a variável resposta. Uma solução para o controle desta variável seria distribuir uma amostra de cada um dos cinco tratamentos para cada técnico. Assim, cada pessoa (ou bloco) estaria avaliando todos os diferentes tratamentos com o mesmo critério pessoal.

Sempre que as categorias de várias variáveis puderem ser combinadas, seus efeitos serão confundidos, mas o controle simultâneo pelos blocos poderá ser efetuado. Para usar o Exemplo 2.1, supondo que as medições não pudessem ser realizadas em um mesmo dia da semana, além de serem necessários os quatro técnicos, o experimento poderia ser realizado em quatro dias diferentes, em cada dia um determinado técnico iria avaliar uma amostra de cada tratamento. Assim, os tratamentos em cada bloco (técnico-dia), estariam sendo avaliados com o mesmo grau de habilidade e sujeitos as mesmas condições climáticas. Os efeitos de diferença de habilidade dos técnicos e diferenças climáticas estariam confundidos entre si, mas não estariam confundidos com os efeitos das outras variáveis.

2.3.3 Interação entre Fatores

Anotações:

Nos experimentos com dois ou mais fatores, além dos efeitos dos fatores, podemos esperar por um outro efeito denominado interação entre os fatores.

A ***interação*** entre dois ou mais fatores significa que os efeitos destas variáveis são relacionados ou que o efeito de um fator depende da categoria do outro fator. Neste caso, os efeitos observados nas variáveis respostas são funções dos efeitos de cada fator e dos efeitos das interações entres eles.

Mais discussões sobre interação entre fatores serão apresentadas na unidade sobre ensaios fatoriais.

UNIDADE 3 – PLANEJAMENTO EXPERIMENTAL

São apresentadas aqui as condições que afetam os resultados dos experimentos e os fatores que devem ser considerados no planejamento dos experimentos

Anotações:

UNIDADE 3 – PLANEJAMENTO EXPERIMENTAL

Os resultados dos experimentos são afetados pela ação dos tratamentos que estamos comparando, mas também, pelo efeito de variáveis estranhas que tendem a mascarar seus efeitos. Estas variações formam o erro experimental.

As causas principais do erro experimental são:

⟹ Variabilidade inerente ao material experimental onde foram aplicados os tratamentos e

⟹ a falta de uniformidade na condução física do experimento.

Para atingir os objetivos propostos em uma pesquisa, devemos procurar minimizar o erro experimental, maximizando a precisão de nossos experimentos com a realização de um bom planejamento e uma condução consciente.

Anotações:

3.1 Requisitos para um bom Experimento

Um experimento fadado ao sucesso deve:

- Ser simples,
- Ter precisão suficiente e ausência de erro sistemático,
- Possibilitar análises estatísticas apropriadas,
- Fornecer conclusões com grande amplitude de validade.

Estes requisitos podem ser satisfeitos atentando pela:

- Escolha do Material Experimental,
- Seleção das Unidades Experimentais,
- Seleção dos Tratamentos,
- Agrupamento de Unidades Experimentais,
- Utilização de técnicas mais refinadas.

3.1.1 Escolha do Material Experimental

Para certos tipos de estudo é desejável um material uniforme, cuidadosamente selecionado. Entretanto, na seleção do material experimental, deve-se ter em mente a população a respeito da qual se deseja obter conclusões. Portanto, é importante empregar os tipos de materiais que serão realmente utilizados na prática.

Anotações:

3.1.2 Seleção das Unidades Experimentais

No planejamento de experimentos de campo, tem-se feito numerosos estudos da variabilidade entre os rendimentos de cultivos em parcelas de diferentes tamanhos e formas submetidas a tratamentos uniformes.

Em geral, a variabilidade decresce com o aumento na precisão, mas uma vez atingido o tamanho ideal, o aumento da precisão diminui rapidamente com tamanhos maiores. As parcelas retangulares são mais eficientes na superação da heterogeneidade do solo quando seu eixo maior está na direção da menor variação do solo. O critério para solucionar o melhor tamanho e forma da parcela é aquele no qual se obtém a máxima exatidão para um dado gasto de tempo e trabalho.

3.1.3 Seleção dos Tratamentos

Em certos casos, a seleção dos tratamentos tem um efeito notável sobre a precisão de um experimento. Por

Anotações:

exemplo, ao se estudar o efeito de um fertilizante, inseticida, fungicida ou herbicida, é mais útil determinar como as parcelas respondem a doses crescentes do produto do que decidir se duas doses sucessivas são ou não significativamente diferentes. Consequentemente, um conjunto apropriado de doses tornará possível planejar testes de significância que são mais sensíveis do que simplesmente comparar médias adjacentes em um conjunto.

3.1.4 Agrupamento de Unidades Experimentais

O agrupamento planejado das unidades experimentais é chamado de controle local. Através de certas restrições na casualização dos tratamentos nas parcelas, é possível remover algumas fontes de variação, tais como variações na fertilidade do solo ou na disponibilidade de água ao longo da área experimental. O agrupamento das parcelas de modos diferentes dá origem aos diferentes delineamentos experimentais.

3.1.5 Utilização de Técnicas mais Refinadas

Uma técnica errônea pode aumentar o erro experimental e distorcer os efeitos dos tratamentos. A técnica é responsabilidade do pesquisador. Uma técnica adequada tem por objetivos:

Anotações:

a) A aplicação uniforme dos tratamentos - em experimentos de adubação, em que se deseja avaliar apenas os níveis de um dado nutriente, os demais deverão ser aplicados de forma uniforme em todas as unidades experimentais. Na prática, em experimentos de campo, consegue-se uma boa aplicação dos tratamentos, planejando-se com antecedência a pesagem dos materiais (adubos, rações, meio de cultura, etc.), ou a confecção de recipientes com peso conhecido.

b) Proporcionar medidas adequadas e não viciadas dos efeitos dos tratamentos - freqüentemente, as medidas apropriadas são logo aparentes, no entanto, algumas vezes, o desenvolvimento e o método satisfatório de medidas requerem anos de investigação, como em pesquisas sociológicas.

c) Prevenir erros grosseiros, dos quais nenhum tipo de experimentação está inteiramente livre - a supervisão e comprovação adequada do trabalho dos ajudantes e um exame dos dados de cada unidade experimental, por parte do pesquisador, muito contribuirá para a descoberta e correção desses erros.

d) Controlar influências externas de forma que cada tratamento produza seu efeito, quando submetidos a condições desejáveis e comparáveis. É difícil generalizar a respeito do grau de controle necessário; pode-se fazer um balanço entre o ganho de precisão obtido e o custo. A produção artificial de enfermidades para experimentos sobre resistência a infecção exemplifica um caso onde a experimentação não pode avançar rapidamente sem controle sobre as condições externas.

Anotações:

Uma técnica deficiente pode introduzir variações adicionais de natureza mais ou menos aleatória. Tais variações adicionais, quando significativas, se revelam na estimação do erro que se calcula na análise de variância. Em casos onde os erros estimados por um pesquisador são consistentemente mais altos que os de outros, os quais utilizam material semelhante, aconselha-se ao pesquisador buscar a razão desta variação, a qual pode ser encontrada nas diferenças de técnicas utilizadas por ambos.

UNIDADE 4 – ANÁLISE DOS DADOS DE UM EXPERIMENTO

Neste capítulo é mostrada a análise de variância e o teste F. As fórmulas de definição e as fórmulas práticas para os cálculos relativos à análise de variância são apresentadas. Através de um exemplo simples é mostrada a interpretação do resultado da análise de variância.

Anotações:

UNIDADE 4 – ANÁLISE DOS DADOS DE UM EXPERIMENTO

Você pode considerar de uma maneira simplista, que o objetivo em um experimento é comparar os tratamentos e saber se eles exercem o mesmo efeito em uma característica avaliada ou se pelo menos um deles tem efeito diferente de outro.

Será que podemos tomar uma decisão sobre o efeito dos tratamentos com base apenas em um resultado observado para cada um deles?

A maneira mais natural parece ser comparar as médias obtidas para cada um dos tratamentos analisando as diferenças entre elas. Mas quando fazemos várias observações em cada tratamento (repetições) para obtermos os valores médios, iremos observar que os valores variam entre si inclusive entre os valores de um mesmo tratamento.

- ➔ **Quais as causas desta variabilidade?**
- ➔ **Quais as causas da variabilidade entre as observações de um mesmo tratamento?**
- ➔ **Isto irá afetar a comparação entre as médias dos tratamentos?**

Anotações:

4.1 Causas da Variabilidade

Todo conjunto de dados obtido em um experimento apresenta variabilidade entre seus componentes. Vamos considerar o exemplo seguinte.

Exemplo 4.1

Foram anotadas as produções de uma variedade de trigo recomendada para Minas Gerais. Foram plantadas no campo experimental da Universidade Federal de Lavras - UFLA, com semeadura realizada no mês de maio de 1997, sob regime de cultivo irrigado. O solo é um Latossolo Vermelho Amarelo, corrigido de acordo com a análise de terra, seguindo as recomendações da Comissão Centro Brasileira de Pesquisa do Trigo. Os tratos culturais e controle de pragas e doenças foram os comuns para a cultura. Foi obtido o peso de grãos na área útil de cada parcela (10 m^2): W = {2,0; 2,2; 2,3; 2,5; 3,0; 3,2; 2,8; 2,9; 2,4; 2,7}.

Através de um cálculo simples podemos ter uma idéia da variabilidade desse conjunto de dados:

- calculamos a média destes dados (no caso, 2,6 kg);
- de cada dado do conjunto subtraímos a média;
- elevamos ao quadrado cada desvio obtido e

- somamos os resultados obtendo Soma dos Quadrados dos Desvios (SQD) que é uma medida de variabilidade:

Anotações:

$$SQD = (2{,}0 - 2{,}6)^2 + (2{,}2 - 2{,}6)^2 + \ldots + (2{,}7 - 2{,}6)^2 = 1{,}32 \text{ kg}^2$$

Nesse exemplo, é razoável supor que toda a variabilidade observada no conjunto de dados W seja devida a:

. heterogeneidade na fertilidade do solo,

. variabilidade genética das sementes,

. variações na condução do experimento (variações na correção do solo, na irrigação, na adubação, na condução, na colheita e pesagem e outros),

. outras causas aleatórias (ataques de pragas e doenças, etc).

Vamos considerar agora que as produções do conjunto W tenham sido devidas a duas variedades ao invés de uma variedade única e que o subconjunto W_A contenha as produções da variedade A e W_B as produções da outra variedade.

$W_A = \{2{,}0; 2{,}2; 2{,}3; 2{,}5; 3{,}0\}$ Média = 2,4 kg

$W_B = \{3{,}2; 2{,}8; 2{,}9; 2{,}4; 2{,}7\}$ Média = 2,8 kg

A variabilidade em cada um desses conjuntos é:

SQD dentro de W_A=

$=(2,0 - 2,4)^2 + (2,2 - 2,4)^2 +... + (3,0 - 2,4)^2 = 0,58\ kg^2$

SQD dentro de W_B=

$=(3,2 - 2,8)^2 + (2,8 - 2,8)^2 +... + (2,7 - 2,8)^2 = 0,34\ kg^2$

Anotações:

Essas variedades são os tratamentos que propositadamente foram incluídos no experimento. Nesse caso nosso interesse consiste em comparar as produções das duas variedades.

Aqui temos um experimento com dois tratamentos (variedades A e B de trigo) e com 5 repetições para cada um. A variabilidade observada nesse experimento provavelmente seja devida a:

- duas variedades de trigo,
- heterogeneidade na fertilidade do solo,
- variabilidade genética das sementes,
- variações na condução do experimento (variações na correção do solo, na irrigação, na adubação, nos tratos culturais, na colheita e pesagem e outros),
- outras causas aleatórias (ataques de pragas e doenças, outras).

Anotações:

Porque a soma da medida da variabilidade da variedade A com a da variabilidade de B não é igual à medida da variabilidade do conjunto todo?

4.2 Análise da Variabilidade

As produções médias das variedades são 2,4 e 2,8 Kg/10m^2, respectivamente. Podemos concluir que a variedade B é mais produtiva? Podemos pensar que no plantio extensivo (grandes áreas, plantio comercial, etc), a variedade B irá produzir mais que a variedade A?

Como existem outras variáveis afetando as médias alem do efeito das variedades, não podemos analisar apenas essas médias para decidir se a variedade B é "realmente" mais produtiva do que a variedade A.

A técnica estatística para tentar respostas para problemas desse tipo foi introduzida por R. A. FISHER, na década de 20 e é chamada Análise de Variância.

O primeiro passo consiste na formalização da hipótese a ser testada. A hipótese de nosso interesse é (releia o 1º parágrafo desta unidade): não existem diferenças entre os efeitos dos tratamentos na variável resposta.

Anotações:

A hipótese que assume que os tratamentos têm o mesmo efeito ou de que os efeitos não diferem entre si, é denominada **hipótese de nulidade** e a notação utilizada é H_0.

Em geral, a hipótese de nulidade para os efeitos de I tratamentos ($\boldsymbol{\tau_i}$) pode ser formalizada do seguinte modo:

$$H_0: \tau_1 = \tau_2 = \cdots = \tau_I$$

A hipótese alternativa em geral é: existe pelo menos uma diferença entre efeitos dos tratamentos.

A metodologia descrita a seguir irá utilizar os dados obtidos no experimento para justificar a aceitação ou a rejeição de H_0.

4.2.1 Tabela da Análise de Variância

A variabilidade presente em um ensaio é analisada com o auxílio de uma tabela padrão denominada: **Tabela da Análise de Variância**, cujo modelo é apresentado a seguir. As colunas dessa tabela referem-se a:

FV – Fontes de Variação

Anotações:

Nessa coluna são descritas as causas de variabilidade dos dados do experimento. Nosso interesse está em conhecer a variabilidade **ENTRE os TRATAMENTOS**. Todas as outras fontes de variabilidade são agrupadas em **RESÍDUO**.

GL – Graus de Liberdade

A cada fonte de variação está associado um número de graus de liberdade.

SQ – Somas de Quadrados

São as somas dos quadrados de desvios ou as medidas de variabilidade calculadas para cada fonte de variação.

QM – Quadrados Médios

São obtidos pela razão entre as Somas de Quadrados e os seus respectivos graus de liberdade. São as medidas de variabilidade para cada fonte de variação, comparáveis entre si.

F_C – valor da estatística F

Anotações:

É o valor obtido para a comparação entre os quadrados médios, dado pela razão entre o QM Entre Tratamentos e o QM do Resíduo. É a estatística de teste apropriada para o teste de hipótese sobre os quadrados médios.

TABELA 4.1 Modelo de Tabela de Análise de Variância.

FV	***GL***	**SQ**	**QM**	**F_c**
Entre Tratamentos	GLEntre	SQEntre	SQEntre/ GLEntre	QMEntre/ QMDentro
Dentro de Tratamentos	GLDentro	SQDentro	SQDentro/ GLDentro	
Total	**GLTotal**	**SQTotal**		

Na Tabela 4.1 observamos que Variabilidade Total existente nos dados do experimento esta dividida em:

Variabilidade Entre Tratamentos – provocada pelos tratamentos e por outras fontes de variabilidade e

Variabilidade Dentro de Tratamentos – provocada por várias fontes de variabilidade, exceto tratamentos.

Anotações:

4.2.2 Procedimento Geral

Para facilitar o seu entendimento da análise de variância, vamos inicialmente considerar o exemplo genérico de um experimento com I tratamentos, cada tratamento com r_i repetições. Y é uma variável resposta qualquer e os dados obtidos serão representados por y_{ij}, onde i refere-se ao tratamento (i=1,2,...I) e j refere-se à repetição (j = r_1, r_2,..., r_I). O número total de parcelas é N = $r_1 + r_2 + ... + r_I$.

Após a coleta dos dados, a organização em uma tabela como a seguir irá facilitar nossos cálculos.

	Repetições				Totais de
Tratamentos	1	2	...	J	Tratamentos
1	y_{11}	y_{12}	...	y_{1r1}	T_1
2	y_{21}	y_{22}	...	y_{2r2}	T_2
...	...	...	...	...	...
I	y_{I1}	y_{I2}	...	y_{IrI}	T_I

Anotações:

Os passos necessários para o preenchimento da tabela da análise de variância são descritos a seguir.

FV: A variação observada entre todos os dados, também chamada de Variação Total, é dividida em Variação Entre Tratamentos (ou simplificadamente Tratamentos) e Variação Dentro de Tratamentos (ou Resíduo).

GL: Para ENTRE TRATAMENTOS: é a quantidade de tratamentos menos uma unidade (I –1);

Para TOTAL: é o número total de parcelas menos um (N – 1);

Para o RESÍDUO: é a soma dos graus de liberdade dentro dos tratamentos. Dentro de cada tratamento o número de graus de liberdade corresponde ao número de repetições do tratamento menos um (r_i - 1). Na prática, o grau de liberdade para o resíduo é obtido pela diferença entre o GLTotal e o GLTratamentos.

SQ:

SQTotal: é a soma dos quadrados das diferenças entre cada observação (y_{ij}) e a média geral do experimento ($\bar{y}_{..}$).

Anotações:

$$SQTOTAL = \sum_{ij} (y_{ij} - \bar{y}_{..})^2$$

(fórmula de definição)

Desenvolvendo o 2º termo da expressão, chegamos a:

$$SQTOTAL = \sum_{ij} y_{ij}^2 - \frac{(\sum_{ij} y_{ij})^2}{N}$$

(fórmula prática)

SQTratamentos: corresponde a soma dos quadrados das diferenças entre as médias de cada tratamento ($\bar{y}_{i.}$) e a média geral lembrando que cada média de tratamento foi obtida de r_i repetições.

$$SQTRATAMENTOS = \sum_{i} r_i (\bar{y}_{i.} - \bar{y})^2$$

A fórmula prática, sendo T_i o total de cada tratamento (somas dos dados das repetições para o tratamento i) e r_i o número de repetições do tratamento i é:

Anotações:

$$SQTRATAMENTOS = \sum_i \frac{T_i^2}{r_i} - \frac{(\sum_{ij} x_{ij})^2}{N}$$

SQResíduo: é o somatório das somas de quadrados dos desvios entre as repetições de cada tratamento e sua média, considerados todos os I tratamentos.

$$SQErro = \sum_i \sum_j (y_{ij} - \bar{y}_{i.})^2$$

Na prática, fazemos: SQErro = SQTotal – SQTratamentos.

QM: Cada Quadrado Médio é obtido dividindo-se a Soma de Quadrados pelo respectivo número de Graus de Liberdade.

F_C: Corresponde à razão entre o QMratamentos e o QMResíduo.

Vamos usar novamente o Exemplo 4.1 determinando as somas de quadrados e apresentando os resultados da análise de variância na Tabela 4.2.

$$SQTotal = 2{,}0^2 + 2{,}2^2 + \dots + 2{,}7^2 - 26^2/10 = 1{,}32$$

$$SQVariedades = 12^2/5 + 14^2/5 - 26^2/10 = 0{,}40$$

$$SQResíduo = 1{,}32 - 0{,}40 = 0{,}92$$

Anotações:

TABELA 4.2 Análise de Variância para as Produções das Duas Variedades de Trigo.

Fontes de Variação	**GL**	**SQ**	**QM**	**F_C**
Entre Variedades	1	0,40	0,40	3,33
Erro Experimental	8	0,92	0,12	
Total	**9**	**1,32**		

Desafio:

Para você calcular uma análise de variância, os dados da Tabela 4.3, são as produções de massa verde (t/ha) de uma cultivar de sorgo plantado em três diferentes espaçamentos. Apresente a análise da variância na tabela 4.4.

Anotações:

TABELA 4.3 Produções de Massa Verde (t/ha) de uma Cultivar de Sorgo.

	ESPAÇAMENTOS		
REPETIÇÕES	**0,50**	**0,75**	**0,90**
I	*186*	*158*	*190*
II	*180*	*173*	*215*
III	*187*	*175*	*221*
IV	*181*	*174*	*195*
V	*184*	*170*	*210*
TOTAIS	**918**	**850**	**1.031**

TABELA 4.4 Análise de Variância para as Produções de Massa Verde.

Fontes de Variação	*GL*	*SQ*	*QM*	F_C
Espaçamentos				
Resíduo				
Total				

4.2.3 Teste F

Anotações:

Podemos considerar a análise de variância como um teste de hipótese sobre os efeitos dos tratamentos: H_0 : não existem diferenças entre os efeitos dos tratamentos.

Para o teste desta hipótese é necessário que os dados experimentais satisfaçam a algumas pressuposições. Estes requisitos são denominados Hipóteses Fundamentais da Análise de Variância e serão discutidos posteriormente. No momento, vamos considerar que estas pressuposições tenham sido satisfeitas.

Se H_0 for falsa, ou seja, existe pelo menos dois efeitos de tratamentos diferentes, as diferenças entre os reais efeitos dos tratamentos aumentarão o valor da SQTratamentos mas não afetarão a SQResíduo fazendo com que a razão entre QMTratamentos e QMResíduo seja maior que 1.

Mas, se H_0 for verdadeira, isto é, os efeitos dos tratamentos são iguais, o QMTratamentos e o QMResíduo serão estimativas do mesmo parâmetro e, portanto, a razão entre eles deverá ser próxima de 1.

A distribuição de probabilidade para a razão entre duas variâncias é conhecida como distribuição F. A estatística F_C = QMTratamentos/QMErro tem distribuição de F com (I – 1) e (N – 1) graus de liberdade.

Anotações:

Com estas considerações e as condições estatísticas asseguradas pelo experimento podemos comparar os valores F_c e F_t, mas como estamos utilizando amostras, sempre nos defrontamos com o erro experimental. Por isto não podemos ter certeza em nossos resultados. Precisamos admitir um nível de confiança nas nossas decisões diferente de 100%. O nível de significância é o complemento do grau de confiança.

Como escolher o nível de significância (α)?

Geralmente toma-se α=5% ou menor. Esta é a probabilidade do erro Tipo I, isto é, a probabilidade de rejeitarmos H_0 quando a mesma for verdadeira.

A maioria dos programas computacionais utilizados para a análise de variância determina o nível de significância exato para cada teste F em cada análise de variância. Quando este recurso não estiver disponível, deveremos utilizar tabelas prontas com os valores da distribuição F como aquela apresentada na Tabela A.1.

Os valores obtidos nas tabelas da distribuição F, de acordo com o nível de significância escolhido, deverão ser comparados com o valor da estatística F_c da tabela da análise de variância.

Anotações:

Qual é a regra de decisão?

Se o valor F_c for maior que o valor F tabelado, rejeita-se H_0 e consideramos o teste F significativo ao nível de α% de probabilidade, isto é, admitimos que, com (100 - α)% de confiança, existe pelo menos uma diferença entre os efeitos dos tratamentos.

Caso o valor F_c seja menor ou igual ao valor F ao nível de α%, não existem evidências para rejeitarmos H_0. Consideramos o teste não-significativo ao nível de α%, isto é, as diferenças numéricas observadas entre as médias dos tratamentos são irrelevantes no contexto daquele experimento.

Para os dados da Tabela 4.2, tomando α = 5% temos que F_t = 5,32 (Tabela A.1). Pela regra de decisão, não existem evidências para rejeitar H_0, portanto esse resultado indica que não existe diferença significativa entre as produtividades médias das duas variedades. A diferença observada entre as duas médias (2,4 para 2,8) é considerada não-significativa.

Anotações:

A Figura 4.1 apresenta uma interpretação gráfica para o resultado do teste F para a Análise de Variância apresentada na Tabela 4.2.

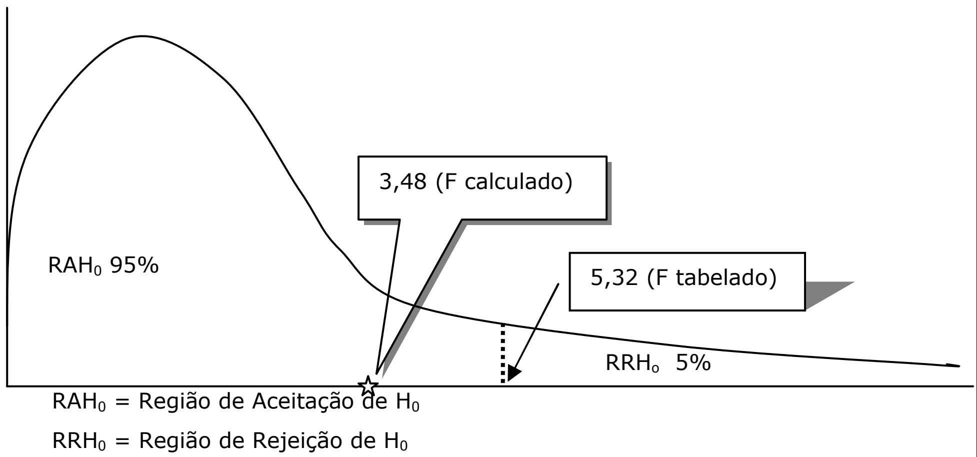

FIGURA 4.1 Interpretação Gráfica do Teste F (Tabela 4.2).

Desafio:

a) *Suponha que o teste F para a Tabela 4.2 tivesse sido significativo. Qual o resultado do experimento neste caso?*
b) *Aplique o teste F na análise de variância da Tabela 4.4. Comente o resultado deste teste.*

UNIDADE 5 – COMPARAÇÕES ENTRE MÉDIAS DE TRATAMENTOS – TESTE TUKEY

Você irá aprender a comparar as médias dos tratamentos para completar a análise de variância dos experimentos. Vamos utilizar o teste Tukey para identificar as diferenças significativas entre as médias dos tratamentos.

Anotações:

UNIDADE 5 – COMPARAÇÕES ENTRE MÉDIAS DE TRATAMENTOS – TESTE TUKEY

Você já sabe que se o teste F para tratamentos na Análise de Variância for significativo, existem evidências para a não aceitação de H_0 como verdadeira, ao nível α% de probabilidade, isto é, **existem efeitos diferenciados para, pelo menos dois tratamentos**.

Agora, o próximo passo será a identificação das diferenças existentes entre os tratamentos (caso sejam mais de dois tratamentos). Se o fator em estudo é uma variável qualitativa (variedades, tipos de adubos, diferentes dietas alimentares) o procedimento apropriado é o das comparações entre as médias dos tratamentos.

5.1 Comparações das Médias Duas a Duas

Se o experimento tem I tratamentos são possíveis $\frac{I!}{2!\,(I-2)!}$ combinações diferentes com as médias desses tratamentos, tomadas duas a duas.

Veja que para o exemplo da Tabela 5.1 são possíveis as comparações:

– m_A comparada com m_B

– m_A comparada com m_C

– m_B comparada com m_C

TABELA 5.1 Alturas de Plantas, em cm, de Três Cultivares de Milho.

	CULTIVARES		
Repetições	**A**	**B**	**C**
I	1,62	1,91	2,15
II	1,75	2,09	2,12
III	1,71	1,88	1,99
IV	1,80	2,10	2,15
V	1,95	2,18	2,10
VI	1,76	2,12	2,14
VII	1,69	2,14	2,11
Médias	**1,75**	**2,06**	**2,11**

A análise de variância para o Exemplo 5.1 mostra que o teste F para tratamentos é significativo a 5% de probabilidade (Tabela 5.2).

Anotações:

TABELA 5.2 Análise de Variância para as Alturas Cultivares de Milho.

Anotações:

Fontes de Variação	GL	SQ	QM	Fc
Entre Cultivares	2	0,5165	0,2582	28,16 *
Resíduo	18	0,1651	0,0092	
Total	**20**	**0,6815**		

** significativo ao nível de 5% de probabilidade

5.2 Teste Tukey

O teste mais utilizado na experimentação para a comparação das médias de tratamentos tomadas duas a duas é o teste Tukey.

Esse teste consiste em, para cada comparação entre duas médias, comparar a diferença entre elas com a diferença mínima significativa (DMS) calculada com o critério de Tukey. A regra de decisão é a de que se a diferença for maior que a DMS, o teste será significativo e as duas médias consideradas estatisticamente diferentes.

A DMS para o teste Tukey é dada por:

$$\text{DMS} = \boldsymbol{q}_{(I,v,\propto\%)}\sqrt{\frac{1}{2}\widehat{\boldsymbol{Var}}(\boldsymbol{m}_i - \boldsymbol{m}_j)}$$

Anotações:

onde $q_{(I,\ v,\ \alpha\%)}$ é a amplitude total estudentizada para uso no teste Tukey ao nível de α% de probabilidade para I tratamentos e v graus de liberdade do Resíduo (Tabela A2).

Se os tratamentos têm o mesmo número de repetições (J), a DMS é:

$$DMS = q_{(I,v)}\sqrt{\frac{QMErro}{J}}.$$

Para o exemplo da Tabela 5.1 obtemos $DMS_{5\%} = 3{,}33\sqrt{\frac{0{,}0092}{7}}$ = 0,13 cm e os resultados são:

Comparação 1: 1,75 – 2,06 = -0,31cm

Como a diferença (em valor absoluto) é maior que a DMS (0,13), os tratamentos A e B têm médias diferentes, ao nível de 5% de probabilidade e, portanto, a altura média da cultivar B é superior à altura média da cultivar A.

Comparação 2: 1,75 – 2,11 = -0,36 cm

A cultivar C tem altura média estatisticamente superior à altura média da cultivar A.

Comparação 3: 2,06 – 2,11 = -0,05 cm

Anotações:

As cultivares B e C tem as mesmas alturas médias ao nível de 5% de probabilidade.

Podemos apresentar esses resultados em uma tabela como a seguir.

TABELA 5.3 Alturas Médias de Cultivares de Milho.

Cultivares	Médias (*)
A	1,75 b
B	2,06 a
C	2,11 a

(*) As médias seguidas da mesma letra não diferem entre si ao nível de 5% de probabilidade.

Um modo mais prático para a aplicação do teste Tukey, principalmente para um número maior de tratamentos é utilizar o algoritmo seguinte.

⇨ 1 Cálcular a DMS;

Anotações:

⇨ 2 Ordenar as médias (ordem decrescente) e colocar a letra a para a primeira média. Esta será a primeira média base.

⇨ 3 Subtrair a DMS da média base, obtendo o intervalo: [(média base); (média base – DMS)]. Toda média contida neste intervalo recebe uma mesma letra. A primeira média fora do intervalo recebe uma letra diferente. Se o intervalo contiver a última média terminou senão, continuar.

⇨ 4 Mudar a base para a próxima média (sem saltar nenhuma). Se a média base for a última média, terminou senão, voltar ao passo 3.

Com as devidas alterações, o algoritmo se aplica analogamente, tomando-se as médias em ordem crescente e somando a DMS em cada passo.

Exemplo 5.1

A Tabela 5.4 apresenta as produções, em Kg/parcela, de seis cultivares de arroz: A – Pratão; B – Dourado Precoce; C – Pérola; D – Batatais; E – IAC-4 e F – IAC-9. A análise de variância esta apresentada na Tabela 5.5.

Anotações:

TABELA 5.4 Produções, em kg/parcela, obtidas de Seis Cultivares de Arroz.

	Tratamentos					
Repetições	**A**	**B**	**C**	**D**	**E**	**F**
I	*2,6*	*2,8*	*2,4*	*1,3*	*1,0*	*3,3*
II	*1,6*	*1,8*	*2,7*	*1,1*	*1,8*	*2,8*
III	*1,4*	*1,8*	*2,1*	*1,3*	*1,2*	*2,3*
IV	*2,4*	*3,0*	*2,4*	*1,4*	*0,8*	*2,6*
V	*2,0*	*2,4*	*3,1*	*1,7*	*1,9*	*2,8*
Totais	**10,0**	**11,8**	**12,7**	**6,8**	**6,7**	**13,8**

TABELA 5.5 Análise de Variância para as Produções das Cultivares de Arroz.

Fontes de Variação	GL	SQ	QM	Fc
Entre Cultivares	*5*	*9,11*	*1,82*	*9,58 *****
Resíduo	*24*	*4,52*	*0,19*	
Total	**29**	**13,63**		

*** significativo ao nível de 1% de probabilidade*

A aplicação do teste Tukey, utilizando o algoritmo descrito é:

1-> DMS = $q_{(6,24)} \sqrt{\frac{0,19}{5}}$ = 4,37 x 0,194936 = 0,8

2->

F	2,8a
C	2,5
B	2,4
A	2,0
D	1,4
E	1,3

Anotações:

3-> **2,8 – 0,8 = 2,0**

Todas as médias no intervalo [2,0 ; 2,8] são iguais à média base; a primeira média fora do intervalo recebe uma letra diferente:

F	2,8a
C	2,5a
B	2,4a
A	2,0a
D	1,4 b
E	1,3

Foram colocadas letras a para as médias C, B e A por serem iguais a F e a letra b para a média D por ser a primeira diferente de F.

Como o intervalo anterior não incluiu a última média e a próxima média base não é a última média, muda-se a média base para a próxima média e repete-se o passo 3:

F	2,8a
C	2,5a
B	2,4a
A	2,0a
D	1,4 b
E	1,3

3-> **2,5 – 0,8 = 1,7**

F	2,8a
C	2,5a
B	2,4a
A	2,0a
D	1,4 b
E	1,3

Observe que todas as médias no intervalo [1,7; 2,5] são iguais à média base e a primeira média fora do intervalo já possui uma letra diferente. Neste caso não há alteração.

Como o intervalo anterior não incluiu a última média e a próxima média base não é a última média, muda-se a média base para a próxima média e repete-se o passo 3:

F	2,8a
C	2,5a
B	2,4a
A	2,0a
D	1,4 b
E	1,3

3-> **2,4 – 0,8 = 1,6**

F	2,8a
C	2,5a
B	2,4a
A	2,0a
D	1,4 b
E	1,3

Anotações:

Observe que todas as médias no intervalo [1,6; 2,4] são iguais à média base e a primeira média fora do intervalo já possui uma letra diferente. Neste caso também não há alteração.

Como o intervalo anterior não incluiu a última média e a próxima média base não é a última média, muda-se a média base para a próxima média e repete-se o passo 3:

F	2,8a
C	2,5a
B	2,4a
A	2,0a
D	1,4 b
E	1,3

3-> **2,0 – 0,8 = 1,2**

F	2,8a
C	2,5a
B	2,4a
A	2,0ab
D	1,4 b
E	1,3 b

Observe agora que as médias no intervalo [1,2; 2,0] são iguais à média base (que já tem a letra) e uma das médias já tem uma letra b. A letra a não pode ser usada, mas a letra b foi usada para todas as médias neste intervalo (A, D e E).

Como o intervalo anterior incluiu a última média o processo terminou e o resultado final é apresentado na Tabela 5.6.

Anotações:

Anotações:

TABELA 5.6 Produções Médias (Kg/parcela) para Seis Cultivares de Arroz.

Cultivares	Produções Médias
Pratão	2,0 ab
Dourado Precoce	2,4 a
Pérola	2,5 a
Batatais	1,4 b
IAC-4	1,3 b
IAC-9	2,8 a

As médias seguidas da mesma letra, não diferem entre si pelo Teste Tukey, ao nível de 5% de probabilidade.

COMO INTERPRETAR ESTES RESULTADOS?

Uma possibilidade é:

As cultivares Dourado Precoce, Pérola e IAC-9 apresentaram as mesmas produtividades e superando as produtividades das cultivares Batatais e IAC-4. As cultivares Pratão, Batatais e IAC-4 apresentaram mesmas produtividades médias.

Desafios:

1. *Abaixo estão os dados de Peso Médio Final (kg) em um experimento com diferentes aditivos (A, B, C e D) utilizados na ração para peixes. Foram utilizados 12 tanques de 500 litros de água com 20 peixes em cada um.*

D 0,93	C 1,40	B 1,12	D 1,21
A 1,04	B 0,98	B 1,14	A 1,14
C 1,22	A 1,33	D 1,16	C 1,24

1.1 Coloque os dados do croqui em uma tabela Tratamentos x Repetições.

1.2 Apresente uma tabela com a análise de variância.

1.3 Apresente uma tabela com as médias dos tratamentos.

1.4 Comente os resultados que julgar relevantes.

Anotações:

Anotações:

2. Um experimento foi conduzido para comparar quatro cultivares de tomate quanto à textura dos frutos. As medidas foram realizadas em uma pequena porção da casca na região equatorial do fruto. Os valores obtidos são apresentados a seguir e estão expressos em Newton (N), onde valores mais altos correspondem a frutos mais firmes. Apresente a análise de variância e comente os resultados.

A 15,1	D 17,5	A 11,4	A 13,7	B 26,5
A 13,5	B 23,5	C 17,7	C 14,6	C 15,3
B 25,6	C 16,3	D 13,7	B 24,2	D 15,9
D 15,3	D 16,5	C 15,6	B 22,3	A 13,2

3. Os dados seguintes são as produções, em Kg/parcela, de 5 cultivares de milho. Foi utilizado do delineamento Inteiramente Casualizado com quatro repetições. Apresente a análise de variância e comente os resultados.

	REPETIÇÕES			
CULTIVARES	**I**	**II**	**III**	**IV**
ESAL-2	2,6	3,2	2,8	2,8
SL 15	2,0	2,2	1,6	1,8
SL 7	1,9	1,8	2,0	2,0
IAC 18	1,2	1,1	1,3	1,4
IAC 38	2,2	2,2	2,2	2,3

Anotações:

4. As produções de repolho ($kg/10m^2$) obtidas em um experimento em Delineamento Inteiramente Casualizado onde foram estudadas diferentes fontes de Nitrogênio estão apresentadas a seguir. Efetue a análise de variância e comente os resultados.

	REPETIÇÕES		
TRATAMENTOS	I	II	III
Nitrocálcio	66,2	61,3	79,5
Esterco de Curral	80,0	50,4	71,3
Uréia	75,5	61,0	65,6
Sulfato de Amônia	88,2	81,8	83,7
Testemunha	39,7	36,6	46,5

UNIDADE 6 – ESTUDO DAS MÉDIAS - REGRESSÃO LINEAR

Nesse tópico são apresentadas as metodologias de Análise de Regressão Linear Simples e de Ajuste de Modelos de Regressão Polinomial na Análise de Variância.

Anotações:

UNIDADE 6 – ESTUDO DAS MÉDIAS – REGRESSÃO LINEAR SIMPLES

6.1 REGRESSÃO

Em muitos estudos o objetivo é identificar a forma de relacionamento entre variáveis. O que procuramos é mensurar **o quanto** variáveis diferentes são relacionadas entre si e conseguir um modelo matemático que explique **de que forma** variáveis se relacionam.

A análise de correlação avalia o grau de associação entre variáveis e a análise de regressão procura verificar a existência de uma relação funcional entre essas variáveis.

A análise de regressão consiste na obtenção de uma função que explique a variação em uma variável (chamada variável dependente) pela variação dos valores de outra (s) variável (is), designada (s) variável (is) independente(s).

Vamos considerar apenas duas variáveis: X como variável independente, ou seja, os valores assumidos só dependem do pesquisador e uma variável dependente (Y) cujos valores são afetados pela variação em X.

Anotações:

O comportamento de Y em relação a X pode ser expresso por diversos modelos matemáticos: linear, quadrático, cúbico, exponencial, logarítmico e muitas outras.

Se fizermos um gráfico simples (X, Y) poderemos ter uma idéia inicial de como se comportam os valores da variável dependente (Y) em função da variação da variável independente (X). Observando este gráfico você poderá tentar estabelecer o tipo de curva ou um modelo matemático que mais se aproxime dos pontos plotados. Veja os dados do Exemplo 6.1 quando apresentados na Figuras 6.1a e 6.1b.

Exemplo 6.1

As produções médias de leite em kg (Y) de um grupo de vacas tratadas com diferentes níveis de proteína em % na ração (X) são apresentadas na Tabela 6.1.

TABELA 6.1 Produções médias de leite em Kg.

X	*10*	*12*	*14*	*16*	*18*	*20*	*22*	*24*	*26*	*28*
Y	*11,8*	*12,0*	*12,1*	*13,2*	*14,1*	*14,4*	*15,6*	*16,0*	*16,4*	*17,0*

Anotações:

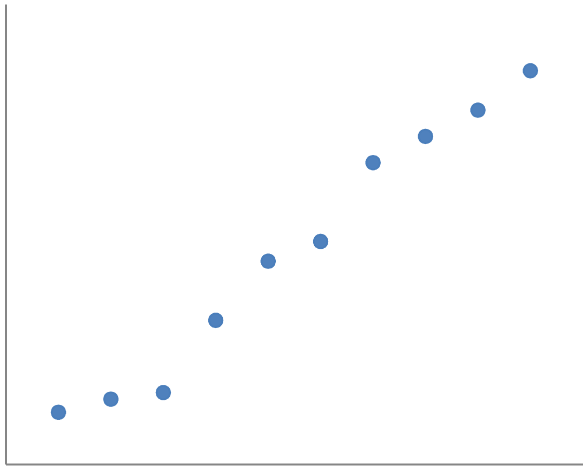

FIGURA 6.1a **Dispersão das Produções Médias de Leite em função dos Níveis de Proteína (%).**

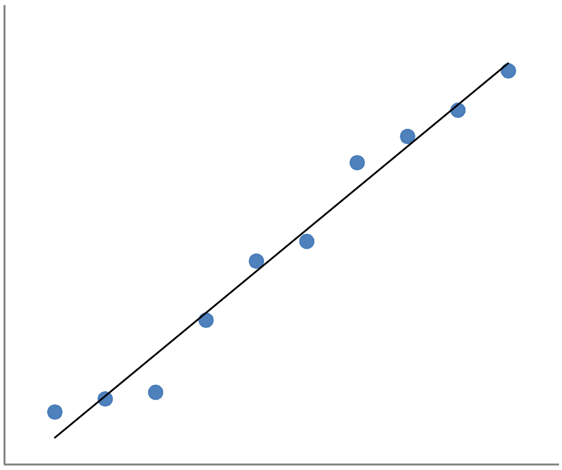

FIGURA 6.1b **Curva de Tendência Linear para as Produções Médias de Leite em Função dos Níveis de Proteína (%).**

Anotações:

Contudo, observamos que os pontos do gráfico (diagrama de dispersão), não vão coincidir perfeitamente com a curva do modelo matemático proposto. Haverá na maior parte dos pontos, uma distância entre os pontos do diagrama e a curva do modelo matemático. Isto acontece, devido ao fato de que o que está em estudo não é um fenômeno matemático e sim um fenômeno natural ou um processo que está sujeito a influências de muitas outras variáveis não consideradas no estudo.

Assim, o objetivo da análise de regressão é obter um modelo matemático que melhor se aproxime dos valores observados de Y em função da variação em X. Chamamos este procedimento de ajuste de um modelo de regressão a um conjunto de dados de observação.

Na escolha do modelo para o estudo de regressão, devemos considerar:

- O modelo selecionado deve ser condizente tanto no grau como no aspecto da curva, para representar em termos práticos, o fenômeno em estudo;

- O modelo deve conter apenas as variáveis que são relevantes para explicar o fenômeno.

Anotações:

6.2 REGRESSÃO LINEAR SIMPLES

A regressão linear simples consiste em utilizar o modelo linear do 1º grau (modelo estatístico) para explicar a relação entre duas variáveis. O polinômio do 1º grau é definido por:

$$\mathbf{Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i}$$

em que:

Y_i corresponde ao i-ésimo valor da variável dependente;
X_i corresponde ao i-ésimo valor da variável independente;

$\boldsymbol{\beta_0}$ e $\boldsymbol{\beta_1}$ são os parâmetros do modelo (coeficiente linear e coeficiente angular, respectivamente);

$\boldsymbol{\varepsilon_i}$ é o erro não observável que corresponde à distância entre o valor observado e o ponto correspondente na curva e i = 1, 2,..., n.

Um dos métodos que utilizamos para conhecer a relação funcional se baseia na obtenção de uma equação em que as distâncias entre os pontos observados e os pontos da curva do modelo matemático, sejam as menores possíveis. Se considerarmos as distâncias ao quadrado, este método é denominado de Método dos Mínimos Quadrados (MMQ).

Anotações:

Para obter a equação de regressão linear ajustada aos seus dados $(\mathbf{x_i}, \mathbf{y_i})$ você pode utilizar as fórmulas seguintes, que são os estimadores de MMQ de $\boldsymbol{\beta_0}$ e $\boldsymbol{\beta_1}$:

$$\hat{\boldsymbol{\beta}}_1 = \frac{\sum_i^n x_i y_i - \frac{\sum_i^n x_i \sum_i^n y_i}{n}}{\sum_i^n x_i^2 - \frac{(\sum_i^n x_i)^2}{n}}$$

$$\hat{\boldsymbol{\beta}}_0 = \frac{\sum_i^n y_i}{n} - \hat{\boldsymbol{\beta}}_1 \frac{\sum_i^n x_i}{n}$$

Uma vez obtida estas estimativas, podemos escrever a equação estimada:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$$

Vamos ajustar o modelo de regressão linear simples aos dados do Exemplo 6.1

➔ da tabela 6.1 obtemos:

$$\sum_i X_i = 10 + 12 + \ldots + 28 = 190$$

$$\sum_i X_i^2 = 10^2 + 12^2 + \ldots + 28^2 = 3940$$

$$\sum_i Y_i = 11{,}8 + \ldots + 17{,}0 = 142{,}6$$

$$\sum_i Y_i^2 = 11{,}8^2 + \ldots + 17{,}0^2 = 2067{,}38$$

$$\sum_i X_i Y_i = (10 \cdot 11{,}8) + (12 \cdot 12{,}0) + \ldots + (28 \cdot 17{,}0) = 2814$$

n = 10

➔ para o modelo proposto $\mathbf{Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i}$ temos:

$$\widehat{\boldsymbol{\beta}}_1 = \frac{\mathbf{2814 - \frac{190\ x\ 142,6}{10}}}{\mathbf{3940 - \frac{(190)^2}{10}}} = \mathbf{0,3170}$$

$$\widehat{\boldsymbol{\beta}}_0 = \frac{\mathbf{142,6}}{\mathbf{10}} - \mathbf{0,3170}\frac{\mathbf{190}}{\mathbf{10}} = \mathbf{8,2370}$$

➔ resultado:

A equação de regressão estimada é: $\hat{Y}_i = 8{,}2370 + 0{,}3170 X_i$.

ATENÇÃO!

A equação estimada apenas estabelece uma relação funcional entre a variável dependente e a variável independente para representar o fenômeno em estudo. Portanto a simples obtenção da equação estimada não responde ao pesquisador se a variação da variável independente influencia significativamente na variação da variável dependente.

Anotações:

6.2.1 Análise de Variância para Regressão

Um teste de hipótese para os parâmetros (coeficientes da equação de regressão) pode nos fornecer um nível de confiança para a adequação do modelo ajustado aos dados.

A metodologia indicada para tal é realizar o teste F da análise de variância para a regressão com os dados observados em função do modelo proposto. Veja o modelo apresentado na Tabela 6.2.

TABELA 6.2 Modelo de Análise de Variância para os Modelos de Regressão.

FV	GL	SQ	QM	F
Regressão	p	SQReg	QMReg	F_c
Resíduo de Regressão	N-1-p	SQRes	QMRes	
Total	**N-1**	**SQTotal**		

Nesta análise, p é o número de coeficientes de regressão (não inclui o β_0) e N é o número de pares de observações.

A soma de quadrados para a regressão varia de acordo com o modelo em teste. Para o modelo de Regressão Linear Simples, p é igual a 1 e você obtém a Soma de

Quadrados para Regressão Linear utilizando a seguinte fórmula:

Anotações:

$$\mathbf{SQREG.LINEAR} = \frac{\left[\sum_i^n x_i y_i - \frac{\sum_i^n x_i \sum_i^n y_i}{n}\right]^2}{\sum_i^n x_i^2 - \frac{(\sum_i^n x_i)^2}{n}}$$

As hipóteses estatísticas para o teste F são as seguintes:

$\boldsymbol{H_0}: \boldsymbol{\beta_1} = \mathbf{0}$, que significa dizer que a variável independente (X_i) não exerce influência na variável dependente, segundo o modelo proposto.

$\boldsymbol{H_0}: \boldsymbol{\beta_1} \neq \mathbf{0}$, o que significa dizer que a variável independente exerce influência na variável dependente, segundo o modelo proposto.

6.2.2 Coeficiente de Determinação

O coeficiente de determinação ($\boldsymbol{R^2}$) fornece uma informação auxiliar ao resultado da análise de variância da regressão, para verificar se o modelo proposto é adequado ou não para descrever o fenômeno.

O R^2 é obtido por: $\mathbf{R^2} = \frac{\mathbf{SQRegressão}}{\mathbf{SQTotal}}$. O valor de R^2 varia no intervalo de 0 a 1. Valores próximos de 1 indicam que o modelo proposto é adequado para descrever o fenômeno.

Anotações:

Vamos testar o ajuste do modelo de regressão linear simples aos dados do Exemplo 6.1

➔ Cálculo das somas de quadrados:

$$\mathbf{SQReg.\,Linerar} = \frac{\left[\mathbf{2814} - \frac{\mathbf{190\ x\ 142,6}}{\mathbf{10}}\right]^2}{\mathbf{3940} - \frac{(\mathbf{190})^2}{\mathbf{10}}} = \mathbf{33,1550}$$

SQTotal = $\sum_{i=1}^{N} y_i^2 - \frac{\left(\sum_{i=1}^{N} y_i\right)^2}{N}$ = 2067,38 – $(142,6)^2/10$ = 33,9040

➔ Análise de variância da regressão

TABELA 6.3 Análise de Variância para Regressão Linear Simples do Exemplo 6.1

FV	GL	SQ	QM	F_c
Regressão	1	33,1550	33,1550	354,22 * *
Resíduo	8	0,7490	0,0936	
Total	**9**	**33,9040**		

* * Significativo ao nível de 1% de probabilidade

Como o teste F foi significativo, rejeita-se H_0 ao nível de 1% de probabilidade. O modelo proposto (1º grau) é adequado para descrever a relação entre o nível de proteína na ração e produção de leite.

Anotações:

➔ Coeficiente de Determinação

$$R^2 = \frac{SQRegressão}{SQTotal} = 33{,}1550/33{,}9040 = 0{,}978$$

Este valor de R^2 indica que 97,80% da variação total é explicada pela regressão linear simples ajustada.

➔ Resultados:

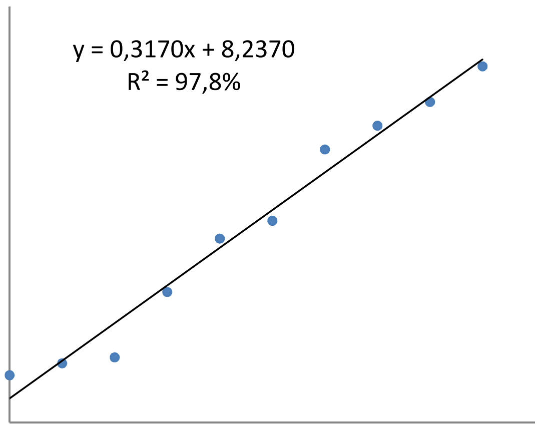

Figura 6.2 Equação de Regressão para Produção de Leite em Função do Nível de Proteína na Ração.

O modelo ajustado esta apresentado na Figura 6.2. Esta equação nos indica que para cada aumento de 1% de proteína na ração, espera-se, em média, um aumento de 0,3170 kg na produção de leite. O ajuste foi considerado muito bom, pois apresentou um R^2 de 97,8% porém, deve ser ressaltado que este resultado é válido apenas no intervalo estudado, ou seja, para nível de proteína na ração de 10 a 28%.

Anotações:

6.3 REGRESSÃO NA ANÁLISE DE VARIÂNCIA

Quando as categorias de um fator são níveis ou doses, significando que o fator é uma variável quantitativa, o procedimento apropriado para o estudo das médias dos tratamentos é a análise de regressão.

Neste caso a variável independente será o fator e a variável dependente será a resposta (médias dos tratamentos).

Neste estudo estaremos analisando apenas o ajuste dos modelos polinomiais para tentar explicar o efeito dos tratamentos na variável reposta. O polinômio de ordem p é da forma:

$$\mathbf{Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \cdots + \beta_p X_i^p}$$

Os modelos polinomiais mais comumente utilizados na análise de regressão das médias de tratamentos são:

Polinômio do 1º grau Regressão Linear Simples:

$$\boldsymbol{Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i}$$

Anotações:

Polinômio do 2º grau ou Regressão Quadrática:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \varepsilon_i$$

Polinômio do 3º grau ou Regressão Cúbica:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \beta_3 X_i^3 + \varepsilon_i$$

O método de regressão polinomial na análise de variância consiste em determinar se um dos polinômios explica satisfatoriamente a relação entre os tratamentos utilizados e as médias dos tratamentos.

Para a escolha do modelo que melhor se ajusta às médias dos tratamentos utilizamos a significância do teste F, a significância dos testes sobre os parâmetros dos modelos e os coeficientes de determinação, como exemplificado a seguir.

6.3.1 Passos para a Análise de Regressão na Análise de Variância

PASSO 1. Definir o fator para o qual será feita a análise de regressão.

PASSO 2. Definir os modelos de regressão a serem testados. A metodologia empregada usa o teste F da análise de variância para testar apenas o coeficiente de regressão associado ao tratamento (X) em seu maior expoente, testando-se assim o efeito de grau p.

Anotações:

Para cada polinômio, as hipóteses do teste F serão:

Efeito Linear: H_0: $\boldsymbol{\beta_1 = 0}$ H_a $\boldsymbol{\beta_1 \neq 0}$

Efeito Quadrático: H_0: $\boldsymbol{\beta_2 = 0}$ H_a $\boldsymbol{\beta_2 \neq 0}$

Efeito Cúbico: H_0: $\boldsymbol{\beta_3 = 0}$ H_a $\boldsymbol{\beta_3 \neq 0}$

. . .

Efeito grau p: H_0: $\boldsymbol{\beta_p = 0}$ H_a $\boldsymbol{\beta_p \neq 0}$

A cada efeito de regressão corresponde 1 grau de liberdade. Assim, é possível testar tantos efeitos de regressão quantos são os graus de liberdade para o fator;

PASSO 3. Determinar as Somas de Quadrados para cada um destes efeitos de regressão. No quadro da Análise de Variância, testar os efeitos de regressão utilizando o Quadrado Médio do Erro Experimental para obter os valores F_c;

PASSO 4. Através do teste F e do coeficiente de determinação (R^2), definir o grau do polinômio que melhor se ajusta às médias da característica observada.

Calculamos o R^2 para cada modelo de regressão através de:

$$\mathbf{R^2} = \frac{\sum(\textbf{Somas de Quadrados dos Efeitos de Regressão})}{\textbf{Soma de Quadrados do Fator}}$$

Observe que no numerador da fórmula do R^2 temos um somatório das somas de quadrados obtidas para cada efeito de regressão.

Observação: Os outros parâmetros dos modelos também devem ser testados para a escolha final do domelo.

PASSO 5. Obter as estimativas dos parâmetros do modelo de Regressão escolhido.

Exemplo 6.2

Um experimento foi realizado para testar o efeito da adubação nitrogenada (0, 100, 200 e 300 kg/ha de Adubo Nitrogenado) na produção de milho. Veja os resultados 6.4.

Anotações:

Anotações:

TABELA 6.4 Produções de Milho (kg/parcela) para doses de adubo Nitrogenado.

Tratamentos	Repetições				Totais	Médias
	I	II	III	IV		
0	49	47	52	50	198	49,50
100	53	58	52	60	223	55,75
200	62	52	74	63	251	62,75
300	72	68	58	67	265	66,25

TABELA 6.5 Análise de Variância das Produções de Milho.

F V	GL	SQ	QM	F_c
Doses de N	3	666,69	222,23	*6,58**
Resíduo	12	405,25	33,77	
Total	**15**	**1.071,94**		

Passo 1. Considere que, neste estudo, o pesquisador esta interessado apenas na regressão linear.

Passo 2. A fórmula para o cálculo da soma de quadrados do efeito linear é:

$$\mathbf{SQEfeito\ linear} = \frac{\left[\sum_i^n x_i y_i - \frac{\sum_i^n x_i \sum_i^n y_i}{n}\right]^2}{\sum_i^n x_i^2 - \frac{(\sum_i^n x_i)^2}{n}} . J$$

Anotações:

onde n é o número de médias, J é o número de repetições, x_i é o tratamento i, y_i é a média do tratamento i.

Utilizando a tabela auxiliar a seguir, facilmente obtemos a Soma de Quadrados Relativa ao Efeito Linear:

Tratamentos (X)	Médias (Y)	XY	X^2
0	49,50	0	0
100	55,75	5.575	10.000
200	62,75	12.550	40.000
300	66,25	19.875	90.000
600	234,25	38.000	140.000

$$SQEL = \frac{\left[3.800 - \frac{(600)(234,25)}{4}\right]^2}{140.000 - \frac{(600)^2}{4}} \cdot 4 = \frac{(2862,5)^2}{50.000} \cdot 4 = 655,51$$

Passo 3. Teste F para os efeitos de regressão:

TABELA 6.6 Análise de Regressão para a Produção de Milho.

Anotações:

Fonte de Variação	GL	SQ	QM	Fc
Doses	(3)	(666,69)		
Efeito Linear	1	655,51	655,51	19,41*
Desvios de Regressão	2	11,18	5,59	<1
Resíduo	12		33,37	

Passo 4. Observa-se pela significância do teste F, que existe uma tendência linear para as produções de milho em função das doses de adubo nitrogenado.

Coeficiente de Determinação:

$$R^2 = 100.\frac{655{,}51}{666{,}69} = 98{,}3\%$$

Passo 5. Modelo de Regressão Linear:

$$\boldsymbol{Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i}$$

As estimativas de β_0 e β_1 são dadas por:

$$\hat{\beta}_1 = \frac{\sum XY - \frac{(\sum X)(\sum Y)}{n}}{\sum X^2 - \frac{(\sum X)^2}{n}} = \frac{2862{,}5}{50.000} = 0{,}0572$$

$$\hat{\beta}_0 = \frac{\sum Y}{n} - \beta_1 \frac{\sum X}{n} = \frac{234,25}{4} - 0,0572 \frac{600}{4} = 49,9825$$

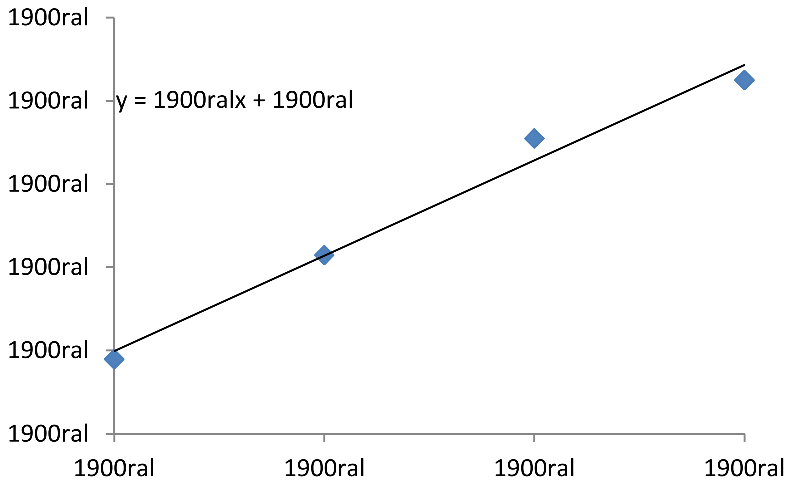

FIGURA 6.3 **Equação de Regressão Linear para as Produções de Milho em Função de Diferentes Doses de Adubo Nitrogenado.**

O ajuste do modelo linear às produções médias de milho foi muito bom (R^2 = 98,3%). A equação ajustada mostra que a cada 1 kg/ha de adubo nitrogenado adicionado, espera-se um aumento médio na produção de 0,0572 kg/parcela.

Anotações:

Anotações:

Exemplo 6.3

Os dados a seguir referem-se às produções de grãos obtidas em um experimento de adubação em milho no qual os tratamentos foram as doses de 25, 50, 75 e 100 kg/ha de P_2O_5 além de uma testemunha que não recebeu a adubação fosfatada.

TABELA 6.7 Produções de Milho (Kg/parcela) para Doses de P_2O_5.

	Doses de P_2O_5				
Repetições	**0**	**25**	**50**	**75**	**100**
I	8,38	7,15	10,07	9,55	9,14
II	5,77	9,78	9,73	8,98	10,17
III	4,90	9,99	7,92	10,24	9,75
IV	4,54	10,70	9,48	8,66	9,50

Nesse exemplo, o pesquisador está interessado em determinar o modelo de regressão que melhor explique o efeito da adubação com P_2O_5 na produção de milho. A análise de variância com regressão é apresentada na Tabela 6.8.

TABELA 6.8 Análise de Variância e Regressão para as Produções de Milho.

Fontes de Variação	GL	SQ	QM	F	R^2
Doses de P_2O_5	(4)	(40,0998)	10,0249	7,17*	-
Efeito Linear	1	22,1266	22,1266	15,82*	55,2%
Efeito Quadrático	1	11,2950	11,2950	8,07*	83,3%
Efeito Cúbico	1	5,8905	5,8950	4,21	-
Efeito de 4º Grau	1	0,7876	0,7876	0,56	-
Erro	15	20,9838	1,3989		
TOTAL	**19**	**61,0836**			

Os resultados do teste F e os valores dos coeficientes de determinação indicam que a regressão quadrática é o modelo apropriado para explicar a relação entre as doses de P_2O_5 e a produção de milho, neste exemplo.

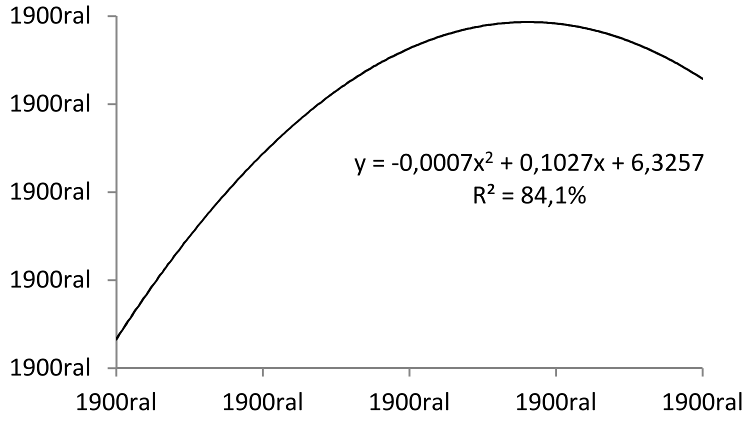

FIGURA 6.4 Equação de Regressão para as Produções de Milho em Função de Diferentes Doses de P_2O_5.

Anotações:

Observamos na Figura 6.4 que a produção de milho cresce segundo uma equação do segundo grau com o aumento das doses de P_2O_5. Inicialmente há um aumento rápido da produção até um máximo de 10,09 kg/parcela para a dose 73,4 kg/ha aproximadamente. A partir desta dose a produção tende a diminuir. O ajuste foi bom (R^2 = 83,1%).

Anotações:

UNIDADE 7 – PRESSUPOSIÇÕES DA ANÁLISE DE VARIÂNCIA

A metodologia de análise dos dados obtidos em experimentos e apresentada na Estatística Experimental só é válida se algumas premissas forem satisfeitas pelos mesmos. Estas premissas são denominadas Hipóteses Fundamentais da Análise de Variância e serão vistas nessa unidade.

Anotações:

UNIDADE 7 – PRESSUPOSIÇÕES DA ANÁLISE DE VARIÂNCIA

Nas décadas de 20 e 30, Ronald A. Fisher foi o pesquisador responsável pela análise de dados da Estação Experimental de Rothamsted de Londres, Inglaterra. Ele foi o pioneiro no uso de métodos estatísticos nos delineamentos experimentais.

Fisher desenvolveu a análise de variância como o primeiro método de análise de dados experimentais. A maioria das aplicações foi feita nas áreas de agricultura e biologia, mas atualmente, constitui uma das principais técnicas utilizadas em todas as áreas do conhecimento.

A utilização da Análise de Variância para um conjunto de dados provenientes de algum experimento pressupõe a verificação de algumas hipóteses.

7.1 HIPÓTESES FUNDAMENTAIS DA ANÁLISE DE VARIÂNCIA

Anotações:

As hipóteses fundamentais estão relacionadas ao modelo estatístico adotado em cada experimento:

1. Os erros têm distribuição Normal (normalidade).

2. Os erros têm a mesma variância (homocedasticidade).

3. Os erros das observações não são correlacionados (independência).

4. Os diferentes efeitos admitidos no modelo estatístico são aditivos (aditividade).

7.1.1 Normalidade

Quando essa hipótese não é satisfeita, além da introdução de erro no nível de significância do teste F e de outros, há uma perda de eficiência na estimação dos efeitos de tratamentos e uma correspondente perda de poder dos testes.

Anotações:

São propostos diversos testes para a verificação de distribuição Normal dos erros, tais como: Kolmogorov–Smirnov, Shapiro–Wilks, assimetria e curtose, entre outros.

Outra ferramenta útil para a verificação da normalidade é o uso do papel normal de probabilidade onde devem ser plotados os resíduos (diferenças entre as observações e a média dos dados). A simples inspeção do gráfico fornece indícios sobre a normalidade.

7.1.2 Homocedasticidade

A falta de homogeneidade de variância é uma das mais graves quebras de suposição básica principalmente para os modelos não balanceados e os modelos de efeitos aleatórios.

Através do gráfico resíduos versus o valor estimado ($\hat{Y}_{ij}$) ou versus a variável X (tratamentos), podemos detectar a não homogeneidade de variância. Em geral, os resíduos não devem ser correlacionados com qualquer outra variável. Os gráficos devem apresentar a ausência de estrutura entre os resíduos e a outra variável plotada.

Também, existem vários testes para a verificação da homocedasticidade: teste de Anscombe e Tukey, teste de Bartlett, etc.

A heterogeneidade dos erros pode ser classificada como irregular e regular. A heterogeneidade é irregular quando aparentemente não existe uma relação entre médias e variâncias.

No caso da heterogeneidade irregular um procedimento empregado é a exclusão de certos tratamentos ou subdividi-los de tal forma que, com os tratamentos restantes ou dentro de cada subdivisão, tenha-se homocedasticidade. Outra alternativa é decompor o quadrado médio do resíduo em componentes apropriados às comparações de interesse.

A heterogeneidade do tipo regular usualmente decorre da não normalidade dos dados, existindo certa relação entre a média e a variância dos vários tratamentos.

Sendo conhecida a distribuição da qual são provenientes os dados, a relação entre a média e a variância dos tratamentos também é conhecida e nestes casos, os dados podem ser transformados de modo que passem a ter distribuição aproximadamente normal e as médias e

Anotações:

variâncias se tornem independentes, resultando também em variâncias homogêneas.

Anotações:

7.1.3 Independência

Como independência dos erros entende-se que a probabilidade do erro de uma observação ter certo valor não depende dos valores dos erros de outras observações.

Quando os erros são correlacionados, os testes de significância não são válidos. Há casos em que, devido a uma correlação positiva entre os erros, o teste de F leva a muitos resultados significativos. Em casos de correlação negativa, o valor da estatística F_c pode ser muito menor que um.

A dependência entre os erros é comum em ensaios em que uma unidade é usada várias vezes como unidade experimental ou quando diferentes parcelas estão em contato físico direto.

Como exemplo, é comum a correlação entre as observações de ensaios de campo onde a semelhança entre as observações de parcelas adjacentes é maior de que entre parcelas distantes ou em ensaios de laboratório, nas observações feitas por uma mesma pessoa ou durante determinado intervalo de tempo.

Anotações:

Em muitos experimentos as unidades experimentais são fisicamente distintas e a hipótese de independência é automaticamente satisfeita. Uma precaução efetiva consiste na aleatorização dos tratamentos.

Plotando os resíduos na ordem em que os dados foram coletados (resíduos versus tempo) podemos verificar facilmente a existência de correlação entre eles. Quando os resíduos se distribuem de maneira desordenada, podemos pensar em não existência de correlação.

7.1.4 Aditividade

Os efeitos admitidos em um modelo estatístico devem ser aditivos.

O modelo estatístico para o delineamento Blocos Casualizados, por exemplo, implica em que o efeito de um tratamento é o mesmo em todos os blocos e o efeito de um bloco é o mesmo em todos os tratamentos. Uma consequência da aditividade é que as diferenças entre os efeitos de dois tratamentos A e B, usualmente é estimada por:

média de todas observações com A – média de todas as observações com B.

Anotações:

A Tabela 7.1 apresenta dois conjuntos de dados supondo um modelo aditivo $(y = t_i + b_j)$ e outro multiplicativo $(y = t_i \cdot b_j)$. Os modelos são apresentados sem erro experimental para facilitar a compreensão.

TABELA 7.1 Modelo Aditivos e Multiplicativos Admitida a Ausência de Erro.

	Modelo Aditivo		Modelo Multiplicativo		Logaritmo do Mod. Multiplicativo	
	Bloco I	Bloco II	Bloco I	Bloco II	Bloco I	Bloco II
Trat. A	10	20	10	20	1,00	1,30
Trat. B	30	40	30	60	1,48	1,78

Fonte: Steel e Torrie (1960)

A não aditividade resulta na heterogeneidade do erro e afeta o nível de significância para comparações entre os tratamentos. Há perda de precisão porque o Erro Experimental é acrescido do componente de não aditividade.

7.2 CASOS DE HIPÓTESES FUNDAMENTAIS NÃO SATISFEITAS

Anotações:

Quando uma destas hipóteses não é satisfeita, a análise de variância não tem validade como técnica de análise estatística e torna-se um simples tratamento matemático dos dados coletados. Para alguns destes casos podem existir alternativas simples.

Na maioria destes casos, as falhas nestas premissas são provocadas por: assimetria extrema, presença de erros grosseiros, comportamento anormal de certos tratamentos ou parte do experimento, não aditividade e variâncias como função das médias.

Alguns dos métodos utilizados nestes casos são: omissão de determinada parte do experimento, subdivisão da variância residual, transformação prévia dos dados e outros.

Em outros casos procura-se empregar outras técnicas de análise de dados, tais como: o método dos Mínimos Quadrados Ponderados para o caso de não homocedasticidade, o método dos Mínimos Quadrados Generalizados para o caso de erros correlacionados, a análise Não-Paramétrica para o caso de não normalidade e análise generalizada.

Anotações:

7.3 Transformação de Dados

Vamos tratar aqui das técnicas de transformação de dados. Uma transformação adequada aos dados é aquela em que:

- A variância da variável transformada não é afetada por mudanças do valor médio;
- A variável transformada é normalmente distribuída;
- A escala de transformação é tal que a média aritmética estime imparcialmente a média verdadeira;
- A escala de transformação é tal que os efeitos reais são lineares e aditivos.

Quando uma transformação de dados é feita, todas as comparações e estimativas de intervalo de confiança devem ser determinadas na nova escala, sendo que as médias podem ser transformadas para a escala original.

A mudança exata da escala é, em geral, difícil e a escolha de uma transformação adequada depende, em parte, da experiência do estatístico.

Anotações:

Devemos lembrar-nos de verificar se as hipóteses fundamentais foram satisfeitas após a escolha e aplicação de uma transformação de dados.

O estudo das relações entre médias e variâncias de tratamentos pode sugerir uma transformação apropriada, veja Box e Cox (1964).

As principais transformações utilizadas são: Raiz Quadrada, Logarítmica e Angular.

7.3.1 Transformação Raiz Quadrada

Esta transformação é utilizada para dados provenientes de contagens como: número de bactérias em uma placa, número de plantas ou insetos em uma dada área, número de defeitos ou acidentes. Geralmente eles se distribuem de acordo com a distribuição de Poisson, em que a média e a variância são iguais. Neste caso, a transformação raiz quadrada dos dados estabiliza a variância, além de torná-la independente da média.

Anotações:

A transformação raiz quadrada pode também ser usada com dados de contagens em que a variância de X é proporcional à média de X, ou seja, $\sigma_x^2 = K\overline{X}$.

Para a distribuição de Poisson tem-se K = 1 mas, frequentemente, encontra-se K > 1, o que indica que a distribuição dos erros tem uma variância maior que aquela de Poisson.

Dados de porcentagem baseados em contagens com um denominador comum, sendo a amplitude de 0% a 20% ou de 80% a 100%, mas não ambas, podem também ser analisados utilizando-se a transformação raiz quadrada.

Quando os dados estão situados entre 80% e 100%, devem ser subtraídos de 100 antes da transformação. A mesma transformação é útil para porcentagens na mesma amplitude quando as observações provêm de uma mesma escala contínua, desde que médias e variâncias sejam aproximadamente iguais.

Quando entre os dados ocorrem valores pequenos, inferiores a 10 e, principalmente, zeros, as transformações $\sqrt{X+1/2}$, $\sqrt{X+1}$ ou $\sqrt{X}+\sqrt{X+1}$ estabilizam a variância mais efetivamente que $\sqrt{X}$, sendo X o valor observado.

Anotações:

A transformação raiz quadrada afeta o tipo de achatamento da distribuição de frequência dos erros e a medida de aditividade. Assim, se os efeitos de blocos e tratamentos são aditivos na escala original, geralmente não o serão na escala raiz quadrada ou vice versa. Contudo, a menos que efeitos de blocos e tratamentos sejam ambos grandes, efeitos que são aditivos em uma escala serão aproximadamente aditivos na escala raiz quadrada.

As médias obtidas com os dados transformados são reconvertidas para a escala original, utilizando-se da operação inversa, ou seja, sendo elevadas ao quadrado. Os valores obtidos, geralmente são ligeiramente menores que as médias originais, porque a média de uma série de raízes quadradas é menor que a raiz quadrada da média original.

7.3.2 Transformação Logarítmica

A transformação logarítmica estabiliza a variância quando o desvio padrão na escala original varia diretamente com a média, ou seja, o coeficiente de variação é constante de tratamento para tratamento. Esse tipo de relação entre média e desvio padrão é encontrado geralmente quando os efeitos são multiplicados em lugar de aditivos. Nessa situação, tal transformação, além de estabilizar a variância, produz aditividade nos efeitos e tende a normalizar a distribuição dos erros. A base 10 para o logaritmo é a mais usada, por conveniência, contudo, qualquer base é satisfatória.

Anotações:

Essa transformação é usada para números inteiros positivos que cobrem uma grande amplitude, sendo que não pode ser usada diretamente quando ocorrem zeros ou quando alguns dos valores são menores que 10. Neste caso, é necessário ter-se uma transformação que equivale à transformação $\sqrt{X}$ para valores pequenos e log X para valores grandes de X. A transformação log(X+1) é a que mais se aproxima da desejada.

As médias obtidas na escala logarítmica são convertidas para a escala original através da operação inversa, ou seja, utilizando-se antilogarítmos dos valores obtidos para essas médias estando, porém afetadas de um erro.

7.3.3 Transformação Angular ou arc sen $\sqrt{p/100}$

Esta transformação é utilizada para homogeneizar a variância residual dos dados de proporção X/N, ou porcentagens 100 (X/N), correspondentes a indivíduos portadores de um dado atributo, em uma amostra de tamanho N e é especialmente recomendada quando as porcentagens cobrem uma grande amplitude de valores.

Anotações:

Admite-se que as proporções têm distribuição binomial com média igual a μ e variância igual a $\mu(1-\mu)/N$. Desde que as proporções têm distribuição binomial, essa variância será máxima para $p = 0{,}5$. As proporções igualmente afastadas de 0,5 terão variâncias iguais e quanto mais afastadas de 0,5, valores menores. A transformação irá, pois, alterar as porcentagens extremas, ou seja, aquelas de menores variâncias.

SNEDECOR e COCHRAN (1976) dizem que essa transformação também pode ser usada para proporções que estão sujeitas a outra causa de variação que não a binomial, sendo porem que a variância dessas proporções deve ser um múltiplo de μ (1- μ). Como, porém, esse produto varia pouco se as porcentagens estiverem todas entre 30% e 70%, a transformação angular será desnecessária. Essa transformação produzirá sensíveis alterações nos valores das porcentagens se estiverem entre 0% e 30% ou 70% e 100%. A transformação arc sen $\sqrt{\%}$ dará melhores resultados quando todas as porcentagens forem baseadas em denominadores iguais, porém, tem sido frequentemente usada quando são diferentes, especialmente, se são aproximadamente iguais.

Pode acontecer que a variável X/N não tenha distribuição binomial e que a transformação angular não atinja seu objetivo, como é o caso, muitas vezes, de dados de controle de pragas e moléstias no campo. Neste caso, deve-se considerar o numerador da proporção como a variável aleatória, podendo ser analisada utilizando-se uma das transformações citadas anteriormente.

A transformação raiz quadrada é recomendada para porcentagens entre 0% e 20% ou 80% e 100% sendo subtraídos de 100 antes da transformação.

Anotações:

UNIDADE 8 – DELINEAMENTOS EXPERIMENTAIS

Vamos estudar nesta unidade as três maneiras de sorteio das parcelas experimentais, chamadas delineamentos experimentais. Vamos apresentar as características de cada um deles, conhecimento que é imprescindível para o planejamento dos experimentos.

Anotações:

UNIDADE 8 – DELINEAMENTOS EXPERIMENTAIS

Os Delineamentos Experimentais são as formas de distribuição das parcelas experimentais na área do experimento.

Os delineamentos experimentais são:

- Delineamento Inteiramente Casualizado
- Delineamento Blocos Casualizados
- Delineamento Quadrado Latino

8.1 Delineamento Inteiramente Casualizado (DIC)

8.1.1 Características

Esse delineamento é utilizado quando a variabilidade entre as parcelas experimentais for muito pequena, isto é, praticamente inexistente.

Devido a esta exigência, você deve usá-lo em locais em que as condições experimentais possam ser bem controladas (laboratórios, casa de vegetação, terrenos com pouca heterogeneidade e outros similares).

As vantagens deste delineamento são:

- o número de graus de liberdade para o Erro Experimental é máximo;
- o número de tratamentos e de repetições depende apenas do número de parcelas experimentais disponíveis;
- é o delineamento mais simples de ser instalado e conduzido.

A maior desvantagem é que toda a variabilidade existente irá compor o erro experimental, exceto apenas a variação entre os efeitos dos tratamentos.

Se as condições do experimento não forem homogêneas podemos encontrar um erro muito grande, o que poderá comprometer os resultados obtidos.

8.1.2 Aleatorização

Nesse delineamento os tratamentos são designados aleatoriamente às parcelas experimentais. Este tipo de

Anotações:

sorteio implica em que todo tratamento tenha a mesma chance de ser aplicado a qualquer parcela na área experimental.

Anotações:

8.1.3 Modelo Estatístico

O modelo linear adequado para este delineamento é dado por:

$$y_{ij} = \mu + t_i + e_{ij}$$

onde y_{ij} é a observação feita na parcela para o tratamento i na repetição j ;

μ representa uma constante inerente a toda parcela ;

t_i representa o efeito do tratamento i ;

e_{ij} representa o erro experimental na parcela i, j.

8.1.3 Modelo Geral de Análise

Considere um experimento com três tratamentos e duas repetições, instalado em DIC. A variável resposta é representada por Y e os dados, observados nas parcelas, representados por y com os índices i e j referentes à, respectivamente, o tratamento e a repetição.

Os dados podem ser colocados em uma tabela como em 8.1 e o modelo geral de análise de variância para os ensaios em DIC é apresentado na Tabela 8.2.

Anotações:

TABELA 8.1 Representação dos Dados de um Experimento em DIC com 3 Tratamentos e 2 Repetições.

Repetições	TRATAMENTOS A	B	C	
I	y_{11}	y_{21}	y_{31}	
II	y_{12}	y_{22}	y_{32}	
TOTAIS	T_1	T_2	T_3	G

TABELA 8.2 Modelo Geral de Análise para ensaios em DIC com I Tratamentos e J repetições.

Fontes de Variação	GL	SQ	QM	Fc
Tratamentos	I-1	SQTrat.(2)	$\frac{\text{SQ Trat.}}{I-1}$	$\frac{\text{QM Trat.}}{\text{QM Erro}}$
Erro Experimental	I(J-1)	SQErro(3)	$\frac{\text{SQ Erro}}{I(J-1)}$	
TOTAL	IJ-1	SQTotal (1)		

- **Cálculos das Somas de Quadrados:**

Anotações:

(1) $SQTotal = \sum y_{ij}^2 - \frac{G^2}{N}$ sendo G = $\sum_i \sum_j y_{ij}$ e N = número total de parcelas (I.J);

(2) $SQTratamentos = \sum \frac{T_i^2}{J} - \frac{G^2}{N}$

sendo $T_i = \sum_j y_{ij}$ (total de cada tratamento);

(3) SQErro = $\sum_i \left(\sum_j y_{ij}^2 - \frac{T_i^2}{J} \right) = SQTotal - SQTratamentos$

8.2 Delineamento Blocos Casualizados (DBC)

No delineamento em blocos casualizados, o material experimental é dividido em grupos homogêneos, cada grupo constituindo uma repetição. Cada repetição ou bloco deve conter uma vez cada tratamento, no caso de blocos completos.

O objetivo em todas as etapas do experimento é manter o erro, dentro de cada bloco, tão pequeno quanto seja possível na prática. Na condução do ensaio deve ser empregada uma técnica uniforme para todas as parcelas de um mesmo bloco. Quaisquer alterações na técnica de condução ou em outras condições que possam afetar os resultados devem ser feitas entre os blocos.

Anotações:

No campo esse agrupamento é feito quando as parcelas são dispostas na área experimental. Cada repetição deverá ser formada por um grupo compacto de parcelas, de forma tão aproximada quanto possível, pois se sabe que as parcelas vizinhas são mais semelhantes em fertilidade do que parcelas distantes.

Em casos em que a colheita do ensaio deva estender-se por algum tempo é aconselhável que seja feita repetição por repetição, sendo que, chuvas e outros fatores podem produzir alterações no peso do material colhido de um dia para o outro.

8.2.1 Características

As principais características e vantagens em relação ao delineamento Inteiramente Casualizado são:

- Permite o controle da influência de uma fonte de variação além do efeito de tratamentos, pelo agrupamento hábil das parcelas (controle local);

Anotações:

- Dentro de cada bloco (repetição), as condições ambientais devem ser homogêneas, podendo variar de bloco para bloco;

- As repetições podem ser distribuídas por uma área maior permitindo conclusões mais gerais.

8.2.2 Aleatorização

Quando as parcelas se acham agrupadas em blocos, os tratamentos são aleatoriamente designados às unidades dentro de cada bloco. Posteriormente, os blocos são sorteados na área experimental.

8.2.3 Modelo Estatístico

A observação da parcela que recebe o tratamento i no bloco j (y_{ij}) é definida, estatisticamente, por:

$$y_{ij} = \mu + t_i + b_j + e_{ij}$$

com i = 1, 2, ..., I e j = 1, 2, ..., J, onde:

μ é uma constante inerente à toda observação;

t_i é o efeito do tratamento i;

b_j é efeito do bloco j;

e_{ij} é o erro na parcela i, j.

Anotações:

8.2.4 Modelo Geral de Análise

As fontes de variação e os respectivos graus de liberdade para o delineamento em blocos casualizados completos são apresentados na Tabela 8.3.

TABELA 8.3 Modelo de Análise de Variância para o Delineamento Blocos Casualizado com I Tratamentos e J Repetições.

Fontes de Variação	GL	SQ
Entre Blocos	J-1	SQBlocos
Entre Tratamentos	I-1	SQTratamentos
Erro	(I-1) (J-1)	SQErro
Total	**IJ-1**	**SQTotal**

As somas de quadrados são computadas através de:

Anotações:

$$SQTotal = \sum_i \sum_j y_{ij}^2 - C$$

$$SQBlocos = \frac{1}{I} \sum_j B_j^2 - C$$

$$SQTratamentos = \frac{1}{J} \sum_i T_i^2 - C$$

com $C = \frac{1}{N} G^2$; B_j o total do bloco j; T_i o total do tratamento i e G o total geral.

Exemplo 5.1

Os dados da Tabela 8.4 referem-se a um ensaio sobre a influência de quatro épocas de corte na produtividade de matéria verde de uma variedade de alfafa. As épocas estudadas foram A, B, C e D sendo A mais precoce e D mais tardia. Foi utilizado o delineamento Blocos Casualizados para controlar um possível gradiente de fertilidade do solo já que a área experimental apresentava uma declividade de 12%.

A análise de variância correspondente é apresentada na Tabela 8.5. As somas de quadrados foram calculadas como:

Correção $(C) = \frac{1}{24}(2{,}89 + 1{,}58 + \ldots + 1{,}00)^2$

Anotações:

$$SQTratamentos = \frac{1}{6}\left[(15{,}40)^2 + (8{,}25)^2 + (11{,}04)^2 + (10{,}91)^2\right] - C$$

$$SQBlocos = \frac{1}{4}[(9{,}32)^2 + (9{,}14)^2 + \ldots + (6{,}43)^2] - C$$

$$SQErro = SQTotal - SQTratamentos - SQBlocos$$

$$SQTotal = (2{,}89)^2 + (1{,}58)^2 + \ldots + (1{,}00)^2 - C$$

TABELA 8.4 Produções em Kg/parcela de matéria verde de alfafa.

Blocos	Épocas de Corte				Totais
	A	B	C	D	
I	2,89	1,58	2,29	2,56	**9,32**
II	2,88	1,28	2,98	2,00	**9,14**
III	1,88	1,22	1,55	1,82	**6,47**
IV	2,90	1,21	1,95	2,20	**8,26**
V	2,20	1,30	1,15	1,33	**5,98**
VI	2,65	1,66	1,12	1,00	**6,43**
Totais	**15,40**	**8,25**	**11,04**	**10,91**	**45,60**

TABELA 8.5 Análise de Variância para a produção de matéria verde (kg/parcela) de alfafa.

Fontes de Variação	GL	SQ	QM	F_c	$F_{5\%}$
Entre Blocos	5	2,7590	0,5518	3,20*	2,90
Entre Épocas	3	4,3820	1,4607	8,47*	3,29
Erro	15	2,5866	0,1724		
Total	**23**				

A DMS para o teste Tukey, ao nível de 5% de probabilidade é igual a 0,69. As médias dos tratamentos são apresentadas na Tabela 8.6. Os resultados mostram que a produção foi maior no corte mais precoce. Para as outras épocas, incluindo a mais tardia, a produção de matéria verde foi a mesma.

Anotações:

TABELA 8.6 Produções Médias de Matéria Verde de Alfafa.

Épocas de Corte	Médias (kg/parcela)
A (mais precoce)	2,57 a
B	1,38 b
C	1,84 b
D	1,82 b

*** As médias seguidas da mesma letra não diferem estatisticamente entre si, pelo teste Tukey, ao nível de 5 % de probabilidade.**

8.3 Delineamento Quadrado Latino (DQL)

8.3.1 Características

Para o delineamento Quadrado Latino, os tratamentos são agrupados nas repetições de duas maneiras distintas.

Essa sistematização dos blocos em duas direções designadas genericamente por “linhas” e “colunas”, permite eliminar os efeitos de duas fontes de variação do erro experimental.

Anotações:

O esquema do delineamento para I tratamentos corresponde a um “quadrado” com I linhas e I colunas, contendo I^2 parcelas. Cada tratamento ocorre uma vez em cada linha e em cada coluna. Um dos possíveis arranjos para um ensaio com quatro tratamentos (A, B, C e D é:

	COLUNAS			
LINHAS	**1**	**2**	**3**	**4**
1	B	C	D	A
2	A	B	C	D
3	C	D	A	B
4	D	A	B	C

Linhas e Colunas são termos gerais para referenciar critérios de classificação e, assim, podem representar uma “espécie” de tratamentos. Se existir interação (dependência) entre os critérios de classificação e os tratamentos, a estatística F_c não tem distribuição de F e o teste não é válido.

O emprego do delineamento Quadrado Latino é muito comum em ensaios industriais, zootécnicos e outras áreas. Em ensaios agronômicos é utilizado geralmente para controlar as diferenças de fertilidade em dois sentidos, nos ensaios de campo.

Anotações:

8.3.2 Aleatorização

Em geral, é satisfatório tomar um quadrado latino qualquer, permutar as linhas e colunas e designar, ao acaso, os tratamentos às letras.

Um processo mais rigoroso para a obtenção de um quadrado latino, aleatoriamente, é dado por FISHER e YATES (1948).

8.3.3 Modelo Estatístico

A observação da parcela na coluna j e na linha k, que recebe o tratamento i, é definida por:

$$y_{ijk} = \mu + t_i + c_j + l_k + e_{jk(i)}$$

com i, j e k =1, 2, ..., I e:

μ é a constante comum a todas as parcelas;

t_i é o efeito do tratamento i;

c_j é o efeito da coluna j;

l_k é o efeito da linha k;

$e_{jk(i)}$ representa o erro aleatório na parcela i,j,k.

Anotações:

8.3.4 Modelo Geral de Análise

Este modelo estatístico leva ao modelo de análise de variância apresentado na Tabela 8.7.

TABELA 8.7 Modelo da Análise de Variância para o delineamento Quadrado Latino com I tratamentos.

Fontes de Variação	GL	SQ
Linhas	I-1	SQLinhas
Colunas	I-1	SQColunas
Tratamentos	I-1	SQTratamentos
Erro	(I-1) (I-2)	SQErro
Total	**I^2-1**	**SQTotal**

Exemplo 8.2

Um experimento foi desenvolvido visando comparar a eficiência de técnicos treinados em amostragem. Uma cultura foi dividida em seis áreas, cada área sendo amostrada por seis técnicos diferentes. O amostrador deveria escolher oito plantas que julgasse com altura representativa da área e anotar a altura média destas plantas. Para as análises estatísticas foi considerada a diferença entre a altura média da amostra e a verdadeira altura média da área correspondente (determinada pela medição de todas as plantas da área). Tais diferenças correspondem aos erros amostrais e são apresentadas na Tabela 8.8. Também foi anotada a ordem em que cada área foi amostrada.

Com auxílio dos totais apresentados na Tabela 8.9 foi computada a análise de variância do exemplo (Tabela 8.10).

TABELA 8.8 Erros Amostrais Referentes às Alturas Médias de Trigo. Área Amostrada por Seis Técnicos (A, B, C, D, E e F).

Ordem de Visita	Áreas					
	I	II	III	IV	V	VI
1ª	3,5(F)	4,2(B)	6,7(A)	6,6(D)	4,1(C)	3,8(E)
2ª	8,9(B)	1,9(F)	5,8(D)	4,5(A)	2,4(E)	5,8(C)
3ª	9,6(C)	3,7(E)	-2,7F)	3,7(B)	6,0(D)	7,0(A)
4ª	10,5D)	10,2(C)	4,6(B)	3,7(E)	5,1(A)	3,8(F)
5ª	3,1(E)	7,2(A)	4,0(C)	-3, (F)	3,5(B)	5,0(D)
6ª	5,9(A)	7,6(D)	-0,7(E)	3,0(C)	4,0(F)	8,6(B)

Anotações:

TABELA 8.9 Totais Para a Amostragem de Plantas de Trigo.

	Totais					
Critérios	**1**	**2**	**3**	**4**	**5**	**6**
Tratamentos	36,4	33,5	36,7	41,5	16,0	7,2
Áreas	41,5	34,8	17,7	18,2	25,1	34,0
Visita	28,9	29,3	27,3	37,9	19,5	28,4

TABELA 8.10 Análise de Variância para Amostragem de Plantas de Trigo.

Fontes de Variação	GL	SQ	QM	F_c	5 %
Tratamentos	5	155,60	31,12	9,35**	2,71
Áreas	5	78,87	15,77	4,74**	
Visita	5	28,60	5,72	1,72	
Erro	20	66,56	3,33		
Total	35	329,63			

As somas de quadrados foram calculadas como:

$$\text{SQTotal} = 3{,}5^2 + 4{,}2^2 + \ldots + 8{,}6^2 - C$$

$$\text{SQTratamentos} = \frac{1}{6}(36{,}4^2 + 33{,}5^2 + \ldots + 7{,}2^2) - C$$

$$\text{SQOrdem Visita} = \frac{1}{6}(28{,}9^2 + 29{,}3^2 + \ldots + 28{,}4^2) - C$$

$$\text{SQÁreas} = \frac{1}{6}(41{,}5^2 + 34{,}8^2 + \ldots + 34{,}0^2) - C$$

Pela simples inspeção nos dados podemos verificar uma tendência consistente de superestimação das alturas das plantas, desde que apenas três dos trinta e seis erros amostrais tiveram valores negativos. O teste de F para

Anotações:

tratamentos, significativo ao nível de 5% mostra que a superestimação varia de técnico para técnico. Com a aplicação do teste de Tukey ($DMS_{5\%}$= 3,1), verifica-se que a tendência é menor para os amostradores E e F, em relação aos outros.

Anotações:

Pelo teste de F observamos que a ordem em que dada área é amostrada não tem efeito na superestimação enquanto que esta tendência varia de área para área.

UNIDADE 9 – ENSAIOS FATORIAIS

Nesta unidade vamos apresentar a análise de variância e o estudo das médias dos tratamentos para experimentos contendo mais de um fator. Apresentamos o importante conceito na estatística experimental denominado interação entre fatores.

Anotações:

UNIDADE 9 – ENSAIOS FATORIAIS

Os ensaios em que é estudado o efeito de apenas um fator são conhecidos como ensaios simples. Mas e se você quiser planejar um experimento para estudar o efeito de dois fatores ou mais fatores ao mesmo tempo como, por exemplo, comparar a produção de 4 cultivares de feijão (A, B, C e D) em três espaçamentos (0,50; 0,75 e 1,00 metros)?

Uma possibilidade seria escolher um espaçamento e plantar as 4 cultivares, separando o mais produtivo. Em outro experimento seria plantada essa cultivar nos 3 diferentes espaçamentos procurando identificar aquele mais favorável.

No entanto, esta não é uma boa alternativa porque, além de menos eficaz, a combinação ótima de cultivar e espaçamento só seria encontrada em uma situação especial. Os fatores em estudo podem se auto relacionar e o melhor nível de um fator poderia depender do nível do outro fator.

A alternativa correta é estudar os dois fatores ao mesmo tempo através dos experimentos denominados **Fatoriais** em que os tratamentos são as combinações das categorias dos fatores.

Anotações:

No exemplo citado seriam plantados as 4 cultivares nos 3 espaçamentos, simultaneamente. Os tratamentos seriam:

Número do Tratamento	Tratamento Cultivar	Espaçamento	Notação 1	Notação 2
1	A	*0,50*	A-0,50	A1
2	*A*	*0,75*	A-0,75	A2
3	*A*	*1,00*	A-1,00	A3
4	*B*	*0,50*	B-0,50	B1
5	*B*	*0,75*	B-0,75	B2
6	*B*	*1,00*	B-1,00	B3
7	*C*	*0,50*	C-0,50	C1
8	*C*	*0,75*	C-0,75	C2
9	*C*	*1,00*	C-1,00	C3
10	*D*	*0,50*	D-0,50	D1
11	*D*	*0,75*	D-0,75	D2
12	*D*	*1,00*	D-1,00	D3

9.1 Notação

A notação utilizada para os experimentos fatoriais é bem variada. No exemplo citado, suponha que tenha sido utilizado o delineamento Blocos Casualizados com 3 repetições. Este experimento seria designado por Ensaio Fatorial 4x3 em DBC com 3 repetições, sendo 4 cultivares e 3 espaçamentos.

Anotações:

Das notações possíveis para os tratamentos, devemos utilizar a mais simples e informativa. No exemplo anterior, a notação 2 é mais concisa.

Considere um ensaio Fatorial 2x2 (ou 2^2) em DIC com 7 repetições sendo 2 cultivares e ausência e presença de calagem. Os tratamentos são:

Tratamento	Níveis dos Fatores	Notação 1	Notação 2	Notação 3
1	cultivar A sem calagem	A-sem	A0	A
2	cultivar B sem calagem	B-sem	B0	B
3	cultivar A com calagem	A-com	A1	A+
4	cultivar B com calagem	B-com	B1	B+

Considere ainda outro exemplo onde serão estudados os nutrientes N, P e K, cada qual estando presente ou ausente. Este ensaio é designado fatorial 2^3 (ou 2x2x2), onde os tratamentos serão:

Anotações:

Tratamento	Níveis dos Fatores	Notação 1	Notação 2
1	sem nutrientes	(T)	000
2	só N	N	100
3	só P	P	010
4	só K	K	001
5	N e P	NP	110
6	N e K	NK	101
7	P e K	PK	011
8	Todos os nutrientes	NPK	111

9.2 Vantagens e Desvantagens

Os ensaios fatoriais permitem economia de tempo e recursos, mas principalmente, possibilitam conclusões mais amplas sobre os fatores incluindo o estudo da interação entre eles e maior precisão para as estimativas dos efeitos principais dos fatores

A desvantagem é o aumento rápido do número de tratamentos a medida que se aumenta o número de fatores ou o número de categorias dos fatores.

Se o delineamento utilizado for o de Blocos Casualizados, o aumento do tamanho do bloco pode acarretar perda de eficiência se houver um aumento da heterogeneidade dentro do bloco.

Anotações:

9.3 Efeitos dos Fatores

O efeito de um fator é a mudança na variável resposta provocada pela mudança de categoria desse fator. Quando estão envolvidos mais de um fator, define-se **efeito simples** como efeito de um fator em cada nível do outro, e **efeito principal** como a média dos efeitos simples do fator.

EXEMPLO 9.1

Veja os dados fictícios da produção de uma cultura em um experimento com duas doses de adubo nitrogenado e duas doses de adubo fosfatado conforme Tabela 9.1.

A Tabela 9.2 é um quadro auxiliar com os totais dos tratamentos para auxiliar os cálculos dos efeitos dos fatores. Ela é construída colocando-se as categorias de um fator nas linhas e as categorias do outro fator nas colunas. Nas células dessa tabela são colocados os totais dos tratamentos correspondentes às somas dos dados das repetições. Nas margens da tabela são colocados os totais de linhas e de colunas.

Anotações:

TABELA 9.1 Dados fictícios de um Ensaio Fatorial com dois Fatores (Nitrogênio e Fósforo) e dois níveis cada.

			Repetições			
Nitrogênio	**Fósforo**	**Trat.**	**I**	**II**	**III**	**Totais**
Nível 1	**Nível 1**	**11**	8	10	6	24
	Nível 2	**12**	10	14	12	36
Nível 2	**Nível 1**	**21**	15	12	15	42
	Nível 2	**22**	16	18	20	54

TABELA 9.2 Totais dos Tratamentos para o Exemplo 9.1.

	Fósforo		
Nitrogênio	**Nível 1**	**Nível 2**	**Totais**
Nível 1	$24_{(3)}$	36	$60_{(6)}$
Nível 2	42	54	96
Totais	$66_{(6)}$	90	$144_{(12)}$

Vamos utilizar a tabela auxiliar de totais para calcular os efeitos dos fatores. Os efeitos para o Nitrogênio são:

Efeito Simples de Nitrogênio no nível 1 de Fósforo

$$= N : P_1 = \frac{42-24}{3} = +6$$

Anotações:

Efeito Simples de Nitrogênio no nível 2 de Fósforo

$$= N : P_2 = \frac{54-36}{3} = +6$$

Efeito Principal de Nitrogênio

$$= \frac{96-60}{6} = +6$$

Estes resultados indicam que, com a quantidade 1 de Fósforo a mudança da dose 1 para a dose 2 de Nitrogênio provoca um aumento médio na produção de 6 unidades. Se a dose de Fósforo for a 2, o efeito simples de Nitrogênio é o mesmo.

Estes dois resultados indicam que o efeito do Nitrogênio não está dependendo do Fósforo. Observe que o efeito de N em geral (considerando todas as categorias do P) é igual à média de seus efeitos simples.

Para o Fósforo, os efeitos são:

Efeito Simples de P no nível 1 de $N = P : N_1 = \frac{36-24}{3} = +4$

Anotações:

Efeito Simples de P no nível 2 de $N = P : N_2 = \frac{54-42}{3} = +4$

Efeito Principal de $P = \frac{96-66}{6} = +4$

Como o efeito do Fósforo não depende do Nitrogênio, efeito simples de Fósforo é o mesmo qualquer que seja o nível de Nitrogênio escolhido.

9.4 Interação Entre os Fatores

Quando os efeitos simples de um fator não são os mesmos em todos os níveis de outro fator, diz-se que existe **interação entre esses fatores**.

Para você fixar este importante conceito, vamos utilizar os dados da Tabela 9.3.

Tabela 9.3 Totais das Produções para um Fatorial 2^2 com 4 Repetições e Ausência e Presença de Calagem e de Adubo com Potássio.

Anotações:

	K		
Calagem	**0**	**50Kg/ha**	**Totais**
Sem	$80_{(4)}$	240	$320_{(8)}$
Com	160	120	280
Totais	$240_{(8)}$	360	$600_{(16)}$

Os efeitos para o Potássio são:

$$\text{Potássio Sem Calagem} = \frac{240 - 80}{4} = 40$$

$$\text{Potássio Com Calagem} = \frac{120 - 160}{4} = -10$$

$$\text{Potássio} = \frac{360 - 240}{8} = 15$$

Verifica-se que o efeito do Potássio depende da calagem: na ausência de Calagem, a adição de Potássio provoca um aumento médio de 40 unidades enquanto que, com a calagem, a adição do potássio provocou uma redução média de 10 unidades na produção.

Devido à existência de interação entre K e Calagem nesse exemplo, o efeito principal do fator representa a médias do que acontece nos diferentes níveis do outro fator (é a média dos efeitos simples).

Nesse exemplo, o efeito principal do Potássio é 15 unidades, o que leva à uma conclusão generalizada sobre seu efeito real.

Vejamos para o outro fator:

Calagem: $K_0 = \dfrac{160 - 80}{40} = 20$

Calagem: $K_{50} = \dfrac{120 - 240}{4} = -30$

Calagem $= \dfrac{280 - 320}{8} = -5$

Como regra geral, quando não existir interação entre os fatores, basta estudar os seus efeitos principais, mas quando existir interação, devemos estudar os efeitos de um fator em cada nível do outro.

Anotações:

A Figura 9.1 ilustra os conceitos de efeitos dos fatores e interação.

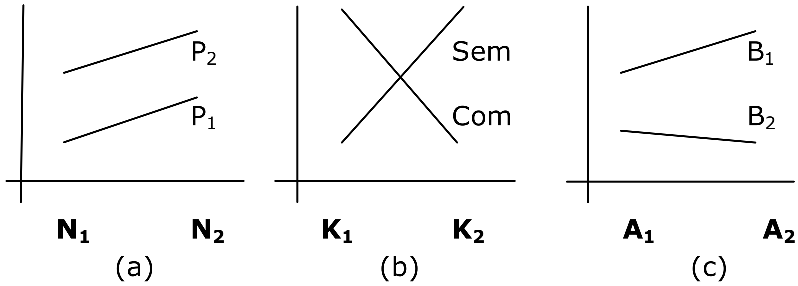

FIGURA 9.1 Ausência de Interação (a), Presença de Interação (b) e (c).

Vemos na Figura 9.1 que, na ausência de interação entre os fatores, as retas são paralelas, significando que o efeito de um fator é o mesmo, independentemente do nível do outro fator.

A interação é positiva quando os fatores apresentam efeito sinérgico. O efeito conjunto dos fatores, nesse caso, é aumentar o resultado da combinação de seus níveis mais altos. Quando os fatores têm efeitos antagônicos, a interação é negativa.

9.5 O Fatorial mais Simples

Anotações:

Um dos ensaios fatoriais mais simples é aquele com dois fatores com dois níveis cada. O Modelo Estatístico para este fatorial, considerando o Delineamento Inteiramente Causualisado é:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk}$$

onde $i = 1,2$ e $j = 1,2$ e:

μ : representa uma constante comum a todas observações

α_i : é o efeito do nível *i* do fator A;

β_j : é o efeito do nível *j* no fator B;

γ_{ij} : é o efeito da interação entre A e B;

ε_{ijk} : é o erro experimental na parcela que recebe o nível *i* do fator A, o nível *j* do fator B e na repetição *k*.

EXEMPLO 9.2

Anotações:

Os dados da Tabela 9.4 referem-se ao tempo em segundos da reação entre duas concentrações de um reagente na presença e na ausência de um catalizador. A análise de variância considerando o delineamento Inteiramente Casualisado é apresentada na Tabela 9.5.

TABELA 9.4 Tempo de Reação (segundos) para um Catalisador (ausência e presença) e duas Concentrações (15% e 25%) de um Reagente.

Tratamentos	Repetições			Totais
	I	II	III	
R15	28	25	27	80
R25	36	32	32	100
R15 + C	18	19	23	60
R25 + C	31	30	29	90

TABELA 9.5 Análise de Variância para o Exemplo 9.2 sem Considerar a Estrutura Fatorial.

FV	GL	SQ	QM	F_c
Tratamentos	3	291,67	97,22	24,80*
Erro	8	31,33	3,92	
Total	**11**	**323,00**		

Anotações:

Considerando a estrutura fatorial dos tratamentos, vamos decompor a soma de Quadrados para Tratamentos nas somas de quadrados relativas ao efeito dos fatores e suas interações, melhorando nossa análise.

Assim, as Fontes de Variação passam a ser todos os fatores e todas as possíveis interações entre eles. Para o cálculo das somas de quadrados dos fatores e das interações é interessante utilizar os quadros auxiliares de totais. A análise de variância apropriada aos ensaios com estrutura fatorial é apresentada na Tabela 9.6.

	Catalisador		
Reagente	**Sem**	**Com**	**Totais**
15%	$80_{(3)}$	100	$180_{(6)}$
25%	60	90	150
Totais	$140_{(6)}$	190	$330_{(12)}$

$$\text{SQConcentração} = \frac{1}{6}\left[180^2 + 150^2\right] - \frac{330^2}{12} = 208{,}33$$

$$\text{SQCatalisador} = \frac{1}{6}\left[140^2 + 190^2\right] - \frac{330^2}{12} = 75{,}00$$

SQ InteraçãoR x C =

$$= \frac{1}{3}\left[80^2 + 100^2 + 60^2 + 90^2\right] - \frac{330^2}{12} - SQR - SQC = 8{,}34$$

Anotações:

TABELA 9.6 Análise de Variância para o Exemplo 9.2

F.V	GL	SQ	QM	F_c
Concentração	1	208,33	208,33	52,14*
Catalisador	1	75,00	75,00	19,13*
R x C	1	8,34	8,34	2,12
Erro	8	31,33	3,92	
Total	**11**	**323,00**		

*** Significativo ao nível de 5% de probabilidade**

Nesse exemplo, embora exista interação entre Concentração e Catalisador, a análise de variância mostra que é muito pequena, isto é, não significativa. Em termos práticos é considerada inexistente e os efeitos dos fatores são representados por seus efeitos principais. Verifique que, nesse caso, os efeitos simples de um fator têm valores muito próximos.

Anotações:

TABELA 9.7 Efeitos Médios de Concentração (a) e de Catalisador (b) sobre os Tempos de Reação.

Reagente	Médias
15%	23,3
25%	31,7

(a)

Catalisador	Médias
Sem	30,0
Com	25,0

(b)

Chegamos aos resultados para cada fator, independentemente do outro e da não existência ou não significância da interação entre eles.

Na Tabela 9.7 observamos que o catalisador diminui o tempo de reação em 5 segundos, para qualquer concentração (15 ou 25%).

Com o aumento da concentração de 15 para 25%, o tempo de reação aumenta em 7,4 segundos em média (com ou sem catalisador).

9.6 Fatoriais p x q

Anotações:

Nestes fatoriais vamos estudar dois fatores sendo um com p categorias e outro com q categorias. O modelo estatístico é semelhante ao modelo apresentado para os fatoriais 2^2 (dois fatores com dois níveis cada) com a diferença apenas no número de níveis dos fatores.

Para ensaio em blocos casualisados com *k* repetições, o modelo estatístico é:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + b_k + e_{ijk}$$

com $i = 1,2,...,p$ e $j = 1,2,...,q$ onde:

μ: representa uma constante inerente a todas as parcelas;

α_i: é o efeito do nível *i* do fator A;

β_j: é o efeito do nível j do fator B;

γ_{ij}: é o efeito da interação entre os fatore A e B;

b_k: é o efeito do bloco *k*;

e_{ijk}: é o erro experimental em cada parcela.

Anotações:

EXEMPLO 9.3

Os dados apresentados na Tabela 9.8 foram adaptados de um ensaio sobre a produção de matéria seca de forrageiras consorciadas com leguminosas. O ensaio foi montado segundo o esquema fatorial 3x4 em blocos casualisados, sendo 3 leguminosas (Azevém, Falaris e Festuca) e 3 doses de calagem além de uma testemunha (0, 1, 2 e 4 toneladas/ha).

TABELA 9.8 Teores de Matéria Seca (t/ha) de Gramíneas Forrageiras Consorciadas com Leguminosas em Diferentes Doses de Calagem.

Calcário (t/ha)	Blocos	Leguminosas		
		Azevém	Falaris	Festuca
0	I	1,97	4,48	6,46
	II	1,90	4,40	7,80
	III	2,02	3,89	6,82
1	I	2,59	5,05	7,64
	II	2,40	5,00	7,80
	III	2,63	4,98	7,82
2	I	2,83	5,55	5,37
	II	2,94	5,60	5,66
	III	3,00	5,78	6,72
4	I	3,32	3,78	5,32
	II	4,80	4,20	5,48
	III	5,00	3,65	4,90

A Tabela 9.9 será utilizada para o cálculo das somas de quadrados dos fatores e da interação. A análise de variância é apresentada na Tabela 9.10.

Anotações:

TABELA 9.9 Quadro Auxiliar de Totais para o Exemplo 9.3.

	Leguminosas			
Calcário	**Azevem**	**Falaris**	**Festuca**	**Totais**
0	5,89$_{(3)}$	12,77	21,08	39,74$_{(9)}$
1	7,62	15,03	23,26	45,91
2	8,77	16,93	17,75	43,45
4	13,12	11,63	15,70	40,45
Totais	**35,40**$_{(12)}$	**56,36**	**77,79**	**169,55**$_{(36)}$

SQ Calcáreo =

$$= \frac{1}{9}\left[39,74^2 + 49,91^2 + 43,45^2 + 40,45^2\right] - \frac{169,55^2}{36} = 2,7000$$

SQ Leguminosas =

$$= \frac{1}{12}\left[35,40^2 + 56,36^2 + 77,79^2\right] - \frac{169,55^2}{36} = 74,8744$$

SQ C x L =

$$= \frac{1}{3}\left[5,85^2 + 12,77^2 + \ldots + 15,70^2\right] - \frac{169,55^2}{36} - 2,7000 - 74,8744 = 23,7608$$

Anotações:

TABELA 9.10 Análise de Variância dos Teores de Matéria Seca (Exemplo 9.3).

Causas de Variação	GL	SQ	QM	F_c	$F_{5\%}$
Calcário (C)	3	2,70	0,90	5,29*	3,05
Leguminosas (L)	2	74,87	37,43	220,18**	
C x L	6	23,76	3,96	23,29**	3,44
Blocos	2	0,61	0,30	1,76	
Erro	22	3,70	0,17		
Total	**35**	**105,64**			

*** significativo ao nível de 5%** **** significativo ao nível de 1 %**

Na Tabela 9.10 observa-se que a interação foi significativa e devemos estudar os efeitos simples dos fatores. Para a interação entre dois fatores, existem duas possibilidades para o seu estudo.

As Tabelas 9.11 e 9.13 apresentam os estudos possíveis para a interação do Exemplo 9.3. Vamos estudar, inicialmente, o efeito das leguminosas em cada dose de calcário. Os cálculos das somas de quadrados são feitos com os totais da Tabela 9.10:

Anotações:

SQ Leguminosas dentro de 0 t/ha =

$=1/3\ [(5,89)^2 + (12,77)^2 + (21,08)^2\ ] - 1/9\ (\ 39,74\ )^2$

SQ Leguminosas dentro de 1 t/ha =

$=1/3\ [(7,62)^2 + (15,03)^2 + (23,26)^2\ ] - 1/9\ (\ 45,91\ )^2$

SQ Leguminosas dentro de 2 t/ha =

$=1/3\ [(8,77)^2 + (16,93)^2 + (17,75)^2\ ] - 1/9\ (\ 43,45\ )^2$

SQ Leguminosas dentro de 4 t/ha =

$= 1/3\ [(13,12)^2 + (11,63)^2 + (15,70)^2\ ] - 1/9\ (\ 40,45\ )^2$

As produções médias são apresentadas na Tabela 9.12 onde foi aplicado o teste de Tukey. A DMS foi igual a 0,85 para a comparação das médias referentes às leguminosas em cada dose de calcário.

Anotações:

TABELA 9.11 Estudos das Leguminosas em cada dose de Calcário do Exemplo 9.3.

Causas de Variação	GL	SQ	QM	F_C
Leguminosas : 0 t/ha	2	38,57	19,28	113,44 **
Leguminosas : 1 t/ha	2	40,80	20,40	120,02 **
Leguminosas : 2 t/ha	2	16,43	8,22	48,33 **
Leguminosas : 4 t/ha	2	2,83	1,41	8,31 **
Erro	22	3,70	0,17	

TABELA 9.12 Produções Médias de Matéria Seca (kg/ha) e Teste de Tukey ($\alpha = 5\%$) referentes ao Exemplo 9.3.

	Leguminosas		
Calcário	Azevém	Falaris	Festuca
0	1,96 c	4,26 b	7,03 a
1	2,54 c	5,01 b	7,75 a
2	2,92 b	5,64 a	5,92 a
4	4,37 b	3,88 c	5,23 a

As médias seguidas da mesma letra nas linhas não diferem entre si, pelo teste de Tukey, ao nível de 5% de probabilidade.

Agora, vamos ver o estudo para comparar os efeitos das doses de calcário em cada leguminosa. As somas de quadrados e a análise dos resultados são:

Anotações:

SQDoses dentro de Azevém =

$$= \frac{1}{3}\left[(5,89)^2 + (7,62)^2 + (8,77)^2 + (13,12)^2\right] - \frac{1}{12}\cdot(35,40)^2$$

SQDoses : Falaris =

$$= \frac{1}{3}\left[(12,77)^2 + (15,03)^2 + (16,93)^2 + (11,63)^2\right] - \frac{1}{12}\cdot(56,36)^2$$

SQDoses : Festuca =

$$= \frac{1}{3}\left[(21,08)^2 + (15,03)^2 + (17,75)^2 + (15,70)^2\right] - \frac{1}{12}\cdot(77,79)^2$$

Anotações:

TABELA 9.13 Desdobramento Doses de Calcário:Leguminosas com Análise de Regressão. Exemplo 9.3.

Causas de Variação	GL	SQ	QM	F_C	R^2
Calcário: Azevém	(3)	(9,50)	3,17	18,63**	
Efeito Linear	1	9,38	9,38	55,71 **	98,7%
Efeito Quadrático	1	0,08	0,08	<1	
Efeito 3^0 grau	1	0,04	0,04	<1	
Calcário: Falaris	(3)	(5,58)	1,86	10,94**	
Efeito Linear	1	0,39	0,39	2,34	7,1%
Efeito Quadrático	1	4,99	4,99	29,65 **	96,5%
Efeito 3^0 grau	1	0,20	0,20	1,15	
Calcário: Festuca	3	(11,37)	3,79	22,30**	
Efeito Linear	1	8,09	8,09	48,05 **	71,1%
Efeito Quadrático	1	0,17	0,17	1,03	72,6%
Efeito 3^0 grau	1	3,11	3,11	18,49 **	100,0%
Erro	**22**	**3,70**	**0,17**		

Para a leguminosa Azevém, o efeito das doses de calcário no teor de matéria seca pode ser explicado por uma reta cuja equação estimada é $\hat{Y} = 1{,}9040 + 0{,}5977X$ com R^2 de 98,7%.

Para a Falaris, o modelo ajustado foi $\hat{Y} = 4{,}1837 + 1{,}3844X - 0{,}3638X^2$, R^2 igual a 96,5 %.

A equação de regressão ajustada para o efeito do calcário na matéria seca da Festuca foi $\hat{Y} = 7{,}0267 + 2{,}8983X - 2{,}6167X^2 + 0{,}4450X^3$.

Anotações:

9.7 Fatoriais p x q x s

São os esquemas fatoriais com três fatores sendo um fator com p níveis, outro com q níveis e o terceiro fator com s níveis. O modelo estatístico para um fatorial pxqxs em blocos casualisados com J repetições é:

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + \delta_{ijk} + b_{lk} + e_{ijkl}$$

onde μ representa uma constante inerente a todas as parcelas; i= 1,2, ..., p, j= 1,2, ... , q e k= 1, 2,..., s e:

α_i : é o efeito do nível *i* do fator A;

β_j : é o efeito do nível j do fator B;

γ_k : é o efeito do nível k do fator C;

$\alpha\beta_{ij}$: é o efeito da interação entre os fatores A e B;

$\alpha\gamma_{ik}$: é o efeito da interação entre os fatores A e C;

$\beta\gamma_{jk}$: é o efeito da interação entre os fatores B e C;

δ_{ijk} : é o efeito da interação entre os fatores A , B e C;

b_l : é o efeito do bloco *l*;

e_{ijkl} : é o erro experimental em cada parcela;

EXEMPLO 9.4

Anotações:

Os dados apresentados na Tabela 9.14 referem-se ao ganho em peso diário de leitões submetidos a dietas com suplementação de Lisina, Metionina e Proteína em um fatorial 3 x 3 x 3 em delineamento Blocos Casualizados.

TABELA 9.14 Ganho em Peso Médio Diário de Suínos em Diferentes Dietas.

Lisina	Metionina	Proteína	Bloco I	Bloco II	Totais
0,00	0,000	8	1,11	0,97	2,08
		12	1,31	1,13	2,44
		14	1,52	1,45	2,97
0.00	0,025	8	1,09	0,99	2,08
		12	1,14	1,12	2,26
		14	1,27	1,22	2,49
0.00	0,050	8	0,85	1,21	2,06
		12	0,98	1,22	2,20
		14	1,67	1,24	2,91
0,05	0,000	8	1,30	1,00	2,30
		12	1,44	1,27	2,71
		14	1,55	1,53	3,08
0.05	0,025	8	1,03	1,21	2,24
		12	1,14	1,30	2,44
		14	1,24	1,34	2,58
0.05	0,050	8	1,12	0,96	2,08
		12	1,44	1,08	2,52
		14	1,76	1,27	3,03
0,10	0,000	8	1,22	1,13	2,35
		12	1,26	1,14	2,40
		14	1,38	1,08	2,46
0.10	0,025	8	1,34	1,41	2,75
		12	1,36	1,30	2,66
		14	1,40	1,21	2,61
0.10	0,050	8	1,34	1,19	2,53
		12	1,48	1,35	2,83
		14	1,46	1,39	2,85
Totais			**35,20**	**32,71**	**67,91**

Com os dados e totais da Tabela 9.14, calculamos as somas de quadrados para "Total" e "Blocos". Usamos os totais da Tabela 9.15 para os cálculos das somas de quadrados dos fatores e interações.

Anotações:

TABELA 9.15 Quadros Auxiliares de Totais para o Exemplo 9.4.

	Lisina			
Metionina	**0,00**	**0,05**	**0,10**	**Marginais**
0,000	7,49$_{(6)}$	8,09	7,21	27,79$_{(18)}$
0,025	6,83	7,26	8,02	22,11
0,050	7,17	7,63	8,21	23,01
Marginais	21,49$_{(18)}$	22,98	23,44	67,91$_{(54)}$

(a)

	Lisina			
Proteína	**0,00**	**0,05**	**0,10**	**Marginais**
8	6,22$_{(6)}$	6,62	7,73	20,47$_{(18)}$
12	6,90	7,67	7,89	22,46
14	8,37	8,69	7,92	24,98
Marginais	21,49	22,98	23,44	67,91

(b)

	Metionina			
Proteína	**0,000**	**0,025**	**0,050**	**Marginais**
8	6,73$_{(6)}$	7,07	6,67	20,47
12	7,55	7,36	7,55	22,46
14	8,51	7,68	8,79	24,98
Marginais	22,79	22,11	23,01	67,91

(c)

Dos totais em **9.15 (a)**:

Anotações:

$$SQLisina = \frac{1}{18}\left[\left(21{,}49^2\right) + (22{,}98)^2 + (23{,}44)^2\right] - \frac{1}{54}\cdot(67{,}91)^2$$

$$SQMetionina = \frac{1}{18}\left[(22{,}79)^2 + (22{,}11)^2 + (23{,}01)^2\right] - \frac{1}{54}\cdot(67{,}91)^2$$

$$SQ(L\times M) = \frac{1}{6}\left[(7{,}49)^2 + (8{,}09)^2 + \ldots + (8{,}21)^2\right] - \frac{1}{54}\cdot(67{,}91)^2 - SQL - SQM$$

Dos totais em **9.15 (b)**, calculamos:

$$SQ\,\mathrm{Pr}\,oteína = \frac{1}{18}\left[(20{,}47)^2 + (22{,}46)^2 + (24{,}98)^2\right] - \frac{1}{54}\cdot(67{,}91)^2$$

$$SQ(L\times P) = \frac{1}{6}\left[(6{,}22)^2 + (6{,}62)^2 + \ldots + (7{,}92)^2\right] - \frac{1}{54}\cdot(67{,}91)^2 - SQL - SQP$$

Dos totais em **9.15 (c)**:

$$SQ(M\times P) = \frac{1}{6}\left[(6{,}73)^2 + (7{,}07)^2 + \ldots + (8{,}79)^2\right] - \frac{1}{54}\cdot(67{,}91)^2 - SQM - SQP$$

Anotações:

Calculamos a soma de quadrados para a interação do três fatores por:

$$SQ(L \times M \times P) = SQ(Tratamento) - [SQL + SQM + SQP + SQLxM + SQLxP + SQMxP]$$

TABELA 9.16 Análise de Variância para o Exemplo 9.4.

Causas de Variação	GL	SQ	QM	F_c
Lisina (L)	2	0,1154	0,0577	3,01
Metionina (M)	2	0,0244	0,0122	< 1
Proteína (P)	2	0,5676	0,2838	14,81**
L x M	4	0,1636	0,0490	2,13
L x P	4	0,2006	0,0501	2,61
M x P	4	0,1062	0,0265	1,38
L x M x P	8	0,0236	0,0029	< 1
Blocos	1	0,1148	0,1148	5,99*
Erro	26	0,4982	0,0192	
Total	**53**	**1,8144**		

Para Proteína, cujo efeito foi significativo, a análise de regressão é apresentada na Tabela 9.17 e a equação ajustada foi $\hat{Y} = 0{,}8072 + 0{,}0397X$. Para os outros fatores, as médias são apresentadas na Tabela 9.18.

TABELA 9.17 Análise de Regressão para o Efeito de Proteína no Ganho em Peso de Leitões.

FV	GL	SQ	QM	F_C	R^2
Proteína	(2)	(0,5676)			
Efeito Linerar	1	0,5307	0,5307	27,70 **	93,5 %
Efeito Quadrático	1	0,0369	0,0369	1,93	
Erro	26	0,4982	0,0192		

TABELA 9.18 Médias para o Ganho em Peso Diário de Leitões.

Metionina	0,000
0,000	1,27 a
0,025	1,23 a
0,050	1,28 a

Lisina	Médias
0,00	1,19 a
0,05	1,28 a
0,10	1,30 a

9.8 Ensaios Fatoriais com uma Repetição

Nos ensaios fatoriais, o número de tratamentos aumenta rapidamente com o número de fatores. Por exemplo, o fatorial 2^5 tem 32 tratamentos, o 2^6 tem 64 combinações, etc. Geralmente, desde que os recursos são limitados, o número de repetições que o experimentador pode empregar é restrito.

Anotações:

Em algumas situações o experimentador usa apenas uma repetição completa do fatorial (a não ser que possa omitir alguns dos fatores originais). Contudo, com apenas uma repetição não é possível computar uma estimativa do erro experimental e as hipóteses sobre os efeitos e interações não podem ser testadas. Nesses casos, uma análise aproximada é feita assumindo-se que algumas interações de ordem mais elevada são desprezíveis e com expectância σ^2, que são combinadas para estimar o erro experimental.

A prática de combinar interações de ordem elevada para estimar o erro experimental é sujeita a crítica. Se algumas dessas interações são significativas o erro estará superestimado e outros efeitos significativos poderão não ser detectados.

As interações que serão combinadas devem ser escolhidas antes de serem examinadas, evitando que a escolha recaia sobre aquelas cujos quadrados médios apresentem-se menores, subestimando assim, o erro experimental. Por exemplo, se na análise de um fatorial 2^5 os efeitos A, B e C bem como as interações AB e AC são bastante grandes, provavelmente o valor da interação ABC será também elevado. Assim, ABC não deverá ser incluída no conjunto que será usado como estimativa do erro. Usualmente, o procedimento é recomendado para fatoriais a partir do 2^4.

Anotações:

9.9 Fatoriais Fracionados

Anotações:

Assumindo que certas interações de ordem elevada são desprezíveis, as informações sobre os efeitos principais e interações de ordem mais baixa podem ser obtidas utilizando-se apenas uma fração do ensaio fatorial completo, isto é, somente certos tratamentos.

Os fatoriais fracionados são bastante empregados em pesquisas industriais e em ensaios preliminares para a identificação dos fatores de maior importância (MONTGOMERY, 1976).

UNIDADE 10 – ENSAIOS FATORIAIS COM PARCELAS DIVIDIDAS

Vamos tratar de modelos fatoriais em que as parcelas experimentais são divididas em subparcelas nas quais são sorteadas as categorias de outro fator.

Anotações:

UNIDADE 10 – ENSAIOS FATORIAIS COM PARCELAS DIVIDIDAS

Nos ensaios fatoriais, os tratamentos são distribuídos nas parcelas de acordo com o procedimento apropriado ao delineamento empregado. Entretanto, outros procedimentos para a aleatorização são possíveis sem alteração da estrutura dos fatores.

Uma das alternativas consiste no sorteio em etapas: inicialmente são sorteadas as categorias de um fator nas parcelas experimentais. Na segunda etapa do sorteio, as parcelas experimentais são subdivididas, formando subparcelas e as categorias do outro fator são sorteadas nessas subparcelas.

Por exemplo, seja um ensaio em que são testadas quatro categorias do fator A em dois blocos. Um segundo fator B com três categorias pode ser incorporado ao ensaio, dividindo-se cada parcela com uma determinada categoria do fator A em três subparcelas.

Após a aleatorização, o croqui do ensaio pode ser representado como o da Figura 10.1.

Bloco I

a_1 b_1	a_1 b_3	a_1 b_2	a_3 b_1	a_3 b_2	a_3 b_3	a_4 b_2	a_4 b_1	a_4 b_3	a_2 b_3	a_2 b_2	a_2 b_1

Bloco II

a_4 b_2	a_4 b_3	a_4 b_1	a_3 b_2	a_3 b_1	a_3 b_3	a_2 b_3	a_2 b_2	a_2 b_1	a_1 b_1	a_1 b_2	a_1 b_3

FIGURA 10.1 Experimentos em Parcelas Subdivididas e Delineamento Blocos Casualizados. Fator A nas Parcelas e B nas Subparcelas.

No planejamento de ensaios com dois ou mais fatores a primeira opção deve ser o esquema fatorial. Você deverá optar pelo esquema de parcelas subdivididas em situações especiais:

1 - Quando as categorias de um dos fatores requerem uma maior quantidade de material experimental do que as categorias de outro fator e for importante reduzir o tamanho dos blocos.

Por exemplo, em um experimento em que os fatores a serem estudados são: dosagens de calagem e diferentes variedades de uma cultura. O primeiro fator pode

Anotações:

requerer uma maior área experimental, dependendo do método de aplicação;

Anotações:

2 - Quando for requerido um maior grau de precisão para as comparações entre as categorias de um fator.

Nesse caso, o fator de deverá ser designado às subparcelas.

3 - Para facilitar a instalação e condução do experimento.

Por exemplo, em um ensaio de comparação de variedades de arroz sob diferentes níveis de irrigação - o fator irrigação deverá ser designado às parcelas para maior facilidade operacional.

4 - Quando um fator adicional deve ser incorporado a um ensaio simples já instalado impossibilitando o sorteio relativo ao esquema fatorial.

Por exemplo, seja um ensaio planejado e já instalado com o objetivo de comparar as produções de massa verde de gramíneas forrageiras. Aproximando-se a época da colheita o pesquisador resolve efetuar três cortes em diferentes épocas para também estudar o efeito das épocas de corte.

Anotações:

Embora o ensaio tenha sido instalado em blocos casualizados com um fator apenas, pode-se analisar os efeitos de gramíneas e de cortes na produção, através do esquema de parcelas subdivididas.

Na Tabela 10.1 são apresentadas as causas de variação com os respectivos graus de liberdade para o esquema de parcelas subdivididas com dois fatores – fator A nas parcelas e fator B nas subparcelas, segundo os delineamentos básicos.

TABELA 10.1 Causas de Variação e Graus de Liberdade para o Esquema de Parcelas Subdivididas em Diferentes Delineamentos.

DIC		DBC		DQL	
FV	**GL**	**FV**	**GL**	**FV**	**GL**
				Linhas	a-1
		A		Colunas	a-1
A	a-1	Blocos	r-1	A	a-1
Erro a	a(r-1)	Erro a	(a-1)(r-1)	Erro a	(a-1) (a-2)
Parcelas	ar-1	Parcelas	ar-1	Parcelas	a^2-1
B	b-1	B	b-1	B	b-1
A x B	(a-1) (b-1)	A x B		A x B	(a-1) (b-1)
Erro b	a(r-1) (b-1)	Erro b	a-1	Erro b	a(a-1)(b-1)
Total	**abr-1**	**Total**	**abr-1**	**Total**	**abr-1**

O modelo estatístico para o esquema de Parcelas Subdivididas, no delineamento blocos Casualizados é:

$$Y_{ijk} = \mu + \alpha_i + b_k + \alpha b_{ik} + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijk}$$

onde:

y_{ijk} : é o valor observado na subparcela i, j, k ;

μ : é uma constante inerente a toda observação;

α_i : é o efeito do *i*-ésimo nível do fator A (*i* = 1, 2, ..., I);

b_k : é o efeito do bloco *k* (*k* = 1, 2, ..., K);

αb_{ik} : representa o erro experimental a nível de parcelas ;

β_j : é o efeito do j-ésimo nível do fator B (*j* = 1, 2, ..., J);

$\alpha\beta_{ij}$: é o efeito da interação entre os fatores A e B;

ε_{ijk} : é o erro experimental a nível de subparcelas.

EXEMPLO 10.1

Um ensaio for realizado em parcelas subdivididas com três variedades de alfafa nas parcelas e quatro épocas de corte final nas subparcelas (SNEDECOR e COX, 1956). As duas primeiras colheitas foram comuns a todas as

Anotações:

parcelas. As épocas de corte foram: A – sem corte; B – corte em 1º de setembro; C – corte em 20 de setembro e D – corte em 7 de outubro. Os dados de produção de matéria verde no ano seguinte aos cortes são apresentados na Tabela 10.2.

Anotações:

TABELA 10.2 Produções de Três Variedades de Alfafa em Quatro Diferentes Épocas de Corte em 1943. Dados em Toneladas por Acre.

		Blocos					
Variedades	**Datas**	**1**	**2**	**3**	**4**	**5**	**6**
Ladak	A	2,17	1,88	1,62	2,34	1,58	1,66
	B	1,58	1,26	1,22	1,59	1,25	0,94
	C	2,29	1,60	1,67	1,91	1,39	1,12
	D	2,23	2,01	1,82	2,10	1,66	1,10
Cossack	A	2,33	2,01	1,70	1,78	1,42	1,35
	B	1,38	1,30	1,85	1,09	1,13	1,06
	C	1,86	1,70	1,81	1,54	1,67	0,88
	D	2,27	1,81	2,01	1,40	1,31	1,06
Ranger	A	1,75	1,95	2,13	1,78	1,31	1,30
	B	1,52	1,47	1,80	1,37	1,01	1,31
	C	1,55	1,61	1,82	1,56	1,23	1,13
	D	1,56	1,72	1,99	1,55	1,51	1,33

Anotações:

10.1 Análise de Variância

As somas de quadrados para os efeitos principais e interação são calculadas com os totais da Tabela 10.3 (a) e para “blocos” e “parcelas” da Tabela 10.3 (b). Os resultados obtidos são apresentados nas Tabelas 10.4 e 10.5.

A soma de quadrados para parcelas é dada por:

$$SQParcelas = \frac{1}{4}(8{,}27)^2 + (7{,}84)^2 + \ldots + (5{,}07)^2 - \frac{1}{72}\cdot(114{,}97)^2 .$$

TABELA 10.3 Tabelas Auxiliares Para os Cálculos das Somas de Quadrados Referentes ao Exemplo 10.1.

	Variedades			
Datas	**L**	**C**	**R**	**Totais**
A	11,25	10,59	10,22	32,06
B	7,84	7,81	8,48	24,13
C	9,98	9,46	8,90	28,34
D	10,92	9,86	9,66	30,44
Totais	39,99	37,72	37,26	114,97

(a)

	Variedades			
Blocos	**L**	**C**	**R**	**Totais**
1	8,27	7,84	6,38	22,49
2	6,75	6,82	6,75	20,32
3	6,33	7,37	7,74	21,44
4	7,94	5,81	6,26	20,01
5	5,88	5,53	5,06	16,47
6	4,82	4,35	5,07	14,24

(b)

Anotações:

TABELA 10.4 Análise de Variância Para o Exemplo 10.1.

Anotações:

Causas de Variação	GL	SQ	QM	F_c
Blocos	5	4,1498	0,8300	6,09*
Variedades (V)	2	0,1780	0,0890	< 1
Erro a	10	1,3623	0,1362	
Parcelas	17	5,6901		
Datas	3	1,9625	0,6542	23,36**
V x D	6	0,2105	0,0351	1,25
Erro b	45	1,2586	0,0280	
Total	**71**	**9,1218**		

TABELA 10.5 Produções Médias (t/acre) de Três Variedades de Alfafa (Exemplo 10.1).

Datas	Médias
Sem Corte	1,78 a
1^0 de Setembro	1,34 c
20 de Setembro	1,57 b
7 de Outubro	1,69 ab

(a)

Variedades	Médias
Ladak	1,67 a
Cossack	1,57 a
Ranger	1,55 a

(b)

Observamos por esses resultados que as diferentes épocas de corte afetam a capacidade de renovação do crescimento e restauração das reservas das raízes. Para uniformidade nas produções seria necessário que o último corte fosse bastante cedo para permitir a recuperação ou bastante tardio para evitar a diminuição das reservas.

Anotações:

As menores produções relativas às datas de corte B e C ($DMS_{TURKEY\ 5\%}$= 0,15 t/acre) também eram esperadas, porém nesta região, o final do mês de setembro é considerado uma época imprópria para o corte da alfafa e, no entanto, a produção relativa a época C superou a relativa a B.

Outro fato não esperado foi a ausência da interação entre variedades e épocas de corte pois a variedade Ladak tem menor capacidade de renovar o crescimento após o corte e deveria ter-se comportado diferentemente das outras variedades.

Esse exemplo também ilustra a necessidade de associação entre os resultados estatísticos, a experiência e o conhecimento do pesquisador na discussão e interpretação dos resultados observados experimentalmente.

10.2 Estudo das Médias

Anotações:

As comparações entre as médias dos tratamentos nos ensaios em parcelas subdivididas envolvem diferentes erros padrões, considerando-se os dois tipos de erros: Erro a e Erro b. As variâncias dos contrastes entre duas médias de tratamentos são:

Entre duas médias do fator A=

$$= 2\left[\frac{QMErroa}{NíveisdeB \times N^{o}\, de\, \mathrm{Re}\, petições}\right]$$

Entre duas médias do fator B=

$$= 2\left[\frac{QMErrob}{NíveisdeA \times N^{o}\, de\, \mathrm{Re}\, petições}\right]$$

Entre duas médias de B em um mesmo nível de A=

$$= 2\left[\frac{QMErrob}{N^{o}\, de\, \mathrm{Re}\, petições}\right]$$

Entre duas médias de A em um mesmo nível de B=

$$= 2\left[\frac{QM\,"ErroCombinado"}{N^{0}\, de\, \mathrm{Re}\, petições}\right]$$

A comparação de duas médias do fator A em um mesmo nível de B envolve uma combinação dos dois erros. O QM e o GL desta combinação são calculados como:

Anotações:

QM "Erro Combinado" =

$$= \frac{[QMErroa + (NíveisdeB - 1)QMErrob]}{NíveisdeB}$$

GL "Erro Combinado" =

$$= \frac{[QMErroa + (NíveisdeB - 1)QMErrob]^2}{\frac{[QMErroa]^2}{GLErroa} + \frac{[(NíveisdeB - 1)QMErrob]^2}{GLErrob}}$$

Para o Exemplo 10.1, as variâncias são:

$\hat{V}ar$ **(Entre duas médias de Variedades)=**

$$= 2\left(\frac{0{,}1362}{4 \times 6}\right) = 0{,}0113$$

$\hat{V}ar$ **(Entre duas médias de Datas)=**

$$= 2\left(\frac{0{,}9280}{3 \times 6}\right) = 0{,}0031$$

Anotações:

$\hat{V}ar$ **(Duas médias de Datas, mesma Variedade)=**

$$= 2\left(\frac{0,0280}{6}\right) = 0,0093$$

QM "Erro Combinado"=

$$= \frac{1}{4}\left[0,1362 + (4-1)0,0280\right] = 0,0551$$

GL "Erro Combinado"=

$$= \frac{\left[0,1362 + (4-1)0,0280\right]^2}{\frac{\left[0,1362\right]^2}{10} + \frac{\left[(4-1)0,0280\right]^2}{45}} = 24,10 \cong 24$$

$\hat{V}ar$ **(Duas médias de variedades, mesma data)=**

$$= 2\left(\frac{0,0551}{6}\right) = 0,1837$$

Lembrando que a DMS do teste Tukey é dada por $DMS = q\sqrt{\frac{1}{2}\hat{V}ar(\hat{Y})}$, onde $\hat{V}ar$ ($\hat{Y}$) representa a estimativa da variância da comparação, pode-se testar qualquer comparação entre duas médias por este teste.

10.3 Ensaios em Parcelas Sub-Subdivididas

Anotações:

Muitas variantes do esquema de parcelas subdivididas podem ser empregadas: uma delas consiste em subdividir cada subparcela em c unidades para a inclusão de um terceiro fator C com c categorias.

O método de análise é uma expansão do método apresentado para o esquema de parcelas subdivididas: o fator A é testado com o erro a; o fator B e a interação AxB com erro b, o fator C e as interações AxC, BxC e AxBxC com o erro c. Para maiores detalhes e outras variantes veja COCHRAN e COX (1950), FEDERER (1955), STEEL e TORRIE (1960).

10.4 Ensaios em Parcelas Subdivididas no Tempo

Verifique que nos exemplos apresentados cada parcela é formada por subparcelas distintas. Em muitos ensaios, entretanto, são feitas observações sucessivas em uma mesma parcela (medidas repetidas) por um período de tempo.

Por exemplo, em um ensaio em blocos casualizados para a comparação de variedades de alfafa, as produções são determinadas por um período de anos, com dois cortes por ano, geralmente.

Anotações:

A análise dos dados de um ensaio em parcelas subdivididas no tempo é semelhante a análise apresentada para parcelas subdivididas no espaço.

Uma diferença consiste na inclusão da interação Fator B x Repetições no modelo. Essa interação é testada com o erro b. Outras diferenças estão relacionadas com as variâncias das comparações entre duas médias de tratamentos (veja Steel, Torrie e Dickey, 1997).

BIBLIOGRAFIA

BANZATTO, D.A. e S. N. KRONKA. *Experimentação Agrícola*. Jaboticabal, SP: FUNEP, 4ª Ed., 237p, 2006.

BARROS NETO, J. C. e Outros. *Planejamento e Otimização de Experimentos*. Unicamp, Campinas, SP, 1995.

BOX, G. E. P. e COX, D. R. *An analisis of transformatios*. Jornal of the Royal Statistical
Society, series B, 26, 211-243, 1964.

BRIEN, C.J. Analysis of variance tables based on experimental structure. *Biometrics*, v.39, p.53-59, 1983.

COCHRAN, W. G. e G. M. COX. *Experimental Designs*. New York, Wiley, 1957.

FISHER, R.A. *The design of experiments*. Edinburgh, Oliver and Boyd, 248p, 1935.

GOMES, F. P. Curso de *Estatística Experimental*. ESALQ, Piracicaba, SP, Editora da USP, 9ª ed, 430p, 1981.

LAPPONI, J. C. *Estatística Usando o Excel*. Versões 4 e 5. São Paulo, SP, Lapponi Trein. e Ed. Ltda, 1995.

MONTGOMERY, D.C. *Design and Analysis of Experiments*. John Wiley & Sons, 649p., 1991.

SNEDECOR, G. W. e W. G. COCHRAN. *Statistical Methods*. Ames, Iowa State Univ. Press, 1976.

STEEL, R.G.D., TORRIE, J.H. e DICKEY, D.A. *Principles and Procedures of Statistics*: *a biometrical approach*. Mc Graw Hill, 666p, 1997.

Tabela A.1. **Quantis superiores da distribuição F ($F_{0,05}$) com ν_1 graus de liberdade do numerador e ν_2 graus de liberdade do denominador, para o valor de 5% da probabilidade α, de acordo com a seguinte afirmativa probabilística: $P(F > F_{0,05}) = 0,05$.**

	ν_1										
ν_2	1	2	3	4	5	6	7	8	9	10	11
1	161,45	199,50	215,70	224,58	230,16	234,0	236,8	238,9	240,5	241,9	242,98
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,40
3	10,13	9,55	9,27	9,11	9,01	8,94	8,88	8,84	8,81	8,78	8,76
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,94
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,70
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,60
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,31
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,10
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,94
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,82
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,72
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,63
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,57
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,51
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,46
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,41
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,37
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,34
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,31
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,28
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,26
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,24
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,22
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,20
26	4,23	3,37	2,97	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,18
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,17
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,15
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,14
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,13
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,04
50	4,03	3,18	2,79	2,56	2,40	2,29	2,20	2,13	2,07	2,03	1,99
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,95
120	3,92	3,07	2,68	2,45	2,29	2,18	2,09	2,02	1,96	1,91	1,87
240	3,88	3,03	2,64	2,41	2,25	2,14	2,05	1,98	1,92	1,87	1,83
480	3,86	3,01	2,62	2,39	2,23	2,12	2,03	1,96	1,90	1,85	1,81
960	3,85	3,01	2,61	2,38	2,22	2,11	2,02	1,95	1,89	1,84	1,80
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,79

Continua ...

Tabela A.1. Continuação ...

ν_2	ν_1: 12	13	14	15	20	30	40	60	120	240	∞
1	243,91	244,69	245,36	245,95	248,0	250,1	251,1	252,2	253,3	253,8	254,31
2	19,41	19,42	19,42	19,43	19,45	19,46	19,47	19,48	19,49	19,49	19,50
3	8,74	8,72	8,71	8,69	8,65	8,60	8,57	8,54	8,49	8,42	8,53
4	5,91	5,89	5,87	5,86	5,80	5,75	5,72	5,69	5,66	5,64	5,63
5	4,68	4,66	4,64	4,62	4,56	4,50	4,46	4,43	4,40	4,39	4,36
6	4,00	3,98	3,96	3,94	3,87	3,81	3,77	3,74	3,70	3,69	3,67
7	3,57	3,55	3,53	3,51	3,44	3,38	3,34	3,30	3,27	3,25	3,23
8	3,28	3,26	3,24	3,22	3,15	3,08	3,04	3,01	2,97	2,95	2,93
9	3,07	3,05	3,03	3,01	2,94	2,86	2,83	2,79	2,75	2,73	2,71
10	2,91	2,89	2,86	2,85	2,77	2,70	2,66	2,62	2,58	2,56	2,54
11	2,79	2,76	2,74	2,72	2,65	2,57	2,53	2,49	2,45	2,43	2,40
12	2,69	2,66	2,64	2,62	2,54	2,47	2,43	2,38	2,34	2,32	2,30
13	2,60	2,58	2,55	2,53	2,46	2,38	2,34	2,30	2,25	2,23	2,21
14	2,53	2,51	2,48	2,46	2,39	2,31	2,27	2,22	2,18	2,15	2,13
15	2,48	2,45	2,42	2,40	2,33	2,25	2,20	2,16	2,11	2,09	2,07
16	2,42	2,40	2,37	2,35	2,28	2,19	2,15	2,11	2,06	2,03	2,01
17	2,38	2,35	2,33	2,31	2,23	2,15	2,10	2,06	2,01	1,99	1,96
18	2,34	2,31	2,29	2,27	2,19	2,11	2,06	2,02	1,97	1,94	1,92
19	2,31	2,28	2,26	2,23	2,16	2,07	2,03	1,98	1,93	1,90	1,88
20	2,28	2,25	2,22	2,20	2,12	2,04	1,99	1,95	1,90	1,87	1,84
21	2,25	2,22	2,20	2,18	2,10	2,01	1,96	1,92	1,87	1,84	1,81
22	2,23	2,20	2,17	2,15	2,07	1,98	1,94	1,89	1,84	1,81	1,78
23	2,20	2,18	2,15	2,13	2,05	1,96	1,91	1,86	1,81	1,79	1,76
24	2,18	2,15	2,13	2,11	2,03	1,94	1,89	1,84	1,79	1,76	1,73
25	2,16	2,14	2,11	2,09	2,01	1,92	1,87	1,82	1,77	1,74	1,71
26	2,15	2,12	2,09	2,07	1,99	1,90	1,85	1,80	1,75	1,72	1,69
27	2,13	2,10	2,08	2,06	1,97	1,88	1,84	1,79	1,73	1,70	1,67
28	2,12	2,09	2,06	2,04	1,96	1,87	1,82	1,77	1,71	1,68	1,65
29	2,10	2,08	2,05	2,03	1,94	1,85	1,81	1,75	1,70	1,67	1,64
30	2,09	2,06	2,04	2,01	1,93	1,84	1,79	1,74	1,68	1,65	1,62
40	2,00	1,97	1,95	1,92	1,84	1,74	1,69	1,64	1,58	1,54	1,51
50	1,95	1,92	1,89	1,87	1,78	1,69	1,63	1,58	1,51	1,48	1,44
60	1,92	1,89	1,86	1,84	1,75	1,65	1,59	1,53	1,47	1,43	1,39
120	1,83	1,80	1,78	1,75	1,66	1,55	1,50	1,43	1,35	1,31	1,25
240	1,79	1,76	1,73	1,71	1,61	1,51	1,44	1,37	1,29	1,24	1,17
480	1,77	1,74	1,71	1,69	1,59	1,48	1,42	1,35	1,26	1,20	1,12
960	1,76	1,73	1,70	1,68	1,58	1,47	1,41	1,33	1,24	1,18	1,08
∞	1,75	1,72	1,69	1,67	1,57	1,46	1,39	1,32	1,22	1,15	1,00

Tabela A.2 Quantil superior da amplitude estudentizada para o teste de Tukey, em função do número de tratamentos (I) e dos graus de liberdade do resíduo (v), ao nível de 5% de probabilidade.

v	Número de tratamentos											
	2	3	4	5	6	7	8	9	10	11	12	13
1	17,97	26,97	32,81	37,06	40,39	43,11	45,40	47,37	49,09	50,62	51,99	53,23
2	6,08	8,33	9,80	10,88	11,73	12,43	13,03	13,54	13,99	14,39	14,75	15,08
3	4,50	5,91	6,82	7,50	8,04	8,48	8,85	9,17	9,46	9,71	9,94	10,15
4	3,93	5,04	5,76	6,29	6,71	7,05	7,35	7,60	7,83	8,03	8,21	8,37
5	3,64	4,60	5,22	5,67	6,03	6,33	6,58	6,80	6,99	7,17	7,32	7,47
6	3,46	4,34	4,90	5,30	5,63	5,90	6,12	6,32	6,49	6,65	6,79	6,92
7	3,34	4,16	4,68	5,06	5,36	5,61	5,82	6,00	6,16	6,30	6,43	6,55
8	3,26	4,04	4,53	4,89	5,17	5,40	5,60	5,77	5,92	6,06	6,18	6,29
9	3,20	3,95	4,42	4,76	5,03	5,25	5,43	5,60	5,74	5,87	5,98	6,09
10	3,15	3,88	4,33	4,66	4,91	5,13	5,31	5,46	5,60	5,72	5,83	5,94
11	3,11	3,82	4,26	4,58	4,82	5,03	5,20	5,35	5,49	5,61	5,71	5,81
12	3,08	3,77	4,20	4,51	4,75	4,95	5,12	5,27	5,40	5,51	5,62	5,71
13	3,06	3,74	4,15	4,45	4,69	4,89	5,05	5,19	5,32	5,43	5,53	5,63
14	3,03	3,70	4,11	4,41	4,64	4,83	4,99	5,13	5,25	5,36	5,46	5,56
15	3,01	3,68	4,08	4,37	4,60	4,78	4,94	5,08	5,20	5,31	5,40	5,49
16	3,00	3,65	4,05	4,33	4,56	4,74	4,90	5,03	5,15	5,26	5,35	5,44
17	2,98	3,63	4,02	4,30	4,52	4,71	4,86	4,99	5,11	5,21	5,31	5,39
18	2,97	3,61	4,00	4,28	4,50	4,67	4,83	4,96	5,07	5,17	5,27	5,35
19	2,96	3,59	3,98	4,25	4,47	4,65	4,80	4,92	5,04	5,14	5,23	5,32
20	2,95	3,58	3,96	4,23	4,45	4,62	4,77	4,90	5,01	5,11	5,20	5,28
25	2,91	3,52	3,89	4,15	4,36	4,53	4,67	4,79	4,90	4,99	5,08	5,16
30	2,89	3,49	3,85	4,10	4,30	4,47	4,60	4,72	4,83	4,92	5,00	5,08
35	2,87	3,46	3,82	4,07	4,26	4,42	4,56	4,67	4,77	4,86	4,95	5,02
40	2,86	3,44	3,79	4,04	4,23	4,39	4,52	4,64	4,74	4,82	4,90	4,98
50	2,84	3,42	3,76	4,00	4,19	4,34	4,47	4,58	4,68	4,77	4,85	4,92
60	2,83	3,40	3,74	3,98	4,16	4,31	4,44	4,55	4,65	4,73	4,81	4,88
70	2,82	3,39	3,72	3,96	4,14	4,29	4,42	4,53	4,62	4,71	4,78	4,85
80	2,81	3,38	3,71	3,95	4,13	4,28	4,40	4,51	4,60	4,69	4,76	4,83
90	2,81	3,37	3,70	3,94	4,12	4,27	4,39	4,50	4,59	4,67	4,75	4,81
100	2,81	3,37	3,70	3,93	4,11	4,26	4,38	4,48	4,58	4,66	4,73	4,80
120	2,80	3,36	3,69	3,92	4,10	4,24	4,36	4,47	4,56	4,64	4,72	4,78
200	2,79	3,34	3,66	3,89	4,07	4,21	4,33	4,44	4,53	4,61	4,68	4,74
240	2,79	3,34	3,66	3,89	4,06	4,21	4,33	4,43	4,52	4,60	4,67	4,73
∞	2,77	3,31	3,63	3,86	4,03	4,17	4,29	4,39	4,47	4,55	4,62	4,68

Continuação

Tabela A.2 Quantil superior da amplitude estudentizada para o teste de Tukey, em função do número de tratamentos (I) e dos graus de liberdade do resíduo (v), ao nível de 5% de probabilidade.

v	Número de tratamentos												
	14	15	16	17	18	19	20	25	30	40	45	50	60
1	54,35	55,38	56,33	57,21	58,03	58,79	59,51	62,52	64,86	68,36	69,74	70,94	72,98
2	15,37	15,65	15,90	16,14	16,36	16,57	16,77	17,60	18,27	19,30	19,71	20,07	20,70
3	10,34	10,52	10,68	10,84	10,98	11,11	11,24	11,78	12,20	12,86	13,13	13,36	13,75
4	8,52	8,66	8,79	8,91	9,03	9,13	9,23	9,66	10,00	10,53	10,74	10,92	11,24
5	7,60	7,72	7,83	7,93	8,03	8,12	8,21	8,58	8,87	9,33	9,51	9,67	9,95
6	7,03	7,14	7,24	7,34	7,43	7,51	7,59	7,92	8,19	8,60	8,77	8,91	9,16
7	6,66	6,76	6,85	6,94	7,02	7,10	7,17	7,48	7,73	8,11	8,26	8,40	8,63
8	6,39	6,48	6,57	6,65	6,73	6,80	6,87	7,16	7,40	7,76	7,90	8,03	8,25
9	6,19	6,28	6,36	6,44	6,51	6,58	6,65	6,92	7,15	7,49	7,63	7,75	7,96
10	6,03	6,12	6,20	6,27	6,34	6,41	6,47	6,74	6,95	7,28	7,41	7,53	7,73
11	5,90	5,99	6,06	6,14	6,20	6,27	6,33	6,58	6,79	7,11	7,24	7,35	7,55
12	5,80	5,88	5,95	6,02	6,09	6,15	6,21	6,46	6,66	6,97	7,10	7,21	7,40
13	5,71	5,79	5,86	5,93	6,00	6,06	6,11	6,36	6,55	6,86	6,98	7,08	7,27
14	5,64	5,72	5,79	5,85	5,92	5,97	6,03	6,27	6,46	6,76	6,87	6,98	7,16
15	5,58	5,65	5,72	5,79	5,85	5,90	5,96	6,19	6,38	6,67	6,79	6,89	7,07
16	5,52	5,59	5,66	5,73	5,79	5,84	5,90	6,13	6,31	6,60	6,71	6,81	6,98
17	5,47	5,55	5,61	5,68	5,74	5,79	5,84	6,07	6,25	6,53	6,64	6,74	6,91
18	5,43	5,50	5,57	5,63	5,69	5,74	5,80	6,02	6,20	6,47	6,58	6,68	6,85
19	5,39	5,46	5,53	5,59	5,65	5,70	5,75	5,97	6,15	6,42	6,53	6,63	6,79
20	5,36	5,43	5,49	5,55	5,61	5,66	5,71	5,93	6,11	6,37	6,48	6,58	6,74
25	5,23	5,30	5,36	5,42	5,47	5,52	5,57	5,78	5,94	6,20	6,30	6,39	6,55
30	5,15	5,21	5,27	5,33	5,38	5,43	5,48	5,68	5,83	6,08	6,18	6,27	6,42
35	5,09	5,15	5,21	5,26	5,32	5,36	5,41	5,60	5,76	6,00	6,09	6,18	6,33
40	5,04	5,11	5,16	5,22	5,27	5,31	5,36	5,55	5,70	5,94	6,03	6,11	6,26
50	4,98	5,04	5,10	5,15	5,20	5,25	5,29	5,47	5,62	5,85	5,94	6,02	6,16
60	4,94	5,00	5,06	5,11	5,16	5,20	5,24	5,42	5,57	5,79	5,88	5,96	6,09
70	4,91	4,97	5,03	5,08	5,12	5,17	5,21	5,39	5,53	5,75	5,84	5,91	6,05
80	4,89	4,95	5,00	5,05	5,10	5,14	5,18	5,36	5,50	5,72	5,80	5,88	6,01
90	4,88	4,93	4,99	5,04	5,08	5,12	5,17	5,34	5,48	5,69	5,78	5,86	5,98
100	4,86	4,92	4,97	5,02	5,07	5,11	5,15	5,32	5,46	5,67	5,76	5,83	5,96
120	4,84	4,90	4,95	5,00	5,04	5,09	5,13	5,30	5,43	5,64	5,73	5,80	5,93
200	4,80	4,86	4,91	4,96	5,00	5,04	5,08	5,25	5,38	5,59	5,67	5,74	5,86
240	4,79	4,85	4,90	4,95	4,99	5,03	5,07	5,24	5,37	5,57	5,65	5,72	5,85
∞	4,74	4,80	4,85	4,89	4,93	4,97	5,01	5,17	5,30	5,50	5,58	5,65	5,76

Continuação

Tabela A.2 Quantil superior da amplitude estudentizada para o teste de Tukey, em função do número de tratamentos (I) e dos graus de liberdade do resíduo (v), ao nível de 5% de probabilidade.

	Número de tratamentos												
v	70	80	90	100	150	200	400	600	800	1000	1400	1800	2000
1	74,65	76,07	77,30	78,39	82,47	85,26	91,69	95,27	97,74	99,62	102,39	104,42	105,26
2	21,22	21,67	22,06	22,41	23,75	24,69	26,90	28,16	29,03	29,70	30,68	31,40	31,70
3	14,08	14,35	14,59	14,81	15,60	16,14	17,37	18,05	18,52	18,87	19,40	19,78	19,94
4	11,50	11,73	11,93	12,10	12,75	13,20	14,25	14,85	15,26	15,58	16,04	16,39	16,53
5	10,18	10,37	10,54	10,70	11,26	11,66	12,57	13,09	13,45	13,72	14,14	14,44	14,56
6	9,37	9,55	9,70	9,84	10,36	10,71	11,54	12,00	12,32	12,57	12,93	13,20	13,31
7	8,83	8,99	9,13	9,26	9,74	10,07	10,84	11,27	11,57	11,79	12,13	12,37	12,48
8	8,43	8,59	8,73	8,84	9,30	9,61	10,34	10,74	11,03	11,24	11,56	11,79	11,88
9	8,13	8,28	8,41	8,53	8,96	9,26	9,96	10,34	10,61	10,82	11,12	11,34	11,44
10	7,90	8,04	8,17	8,28	8,70	8,98	9,65	10,03	10,29	10,49	10,78	11,00	11,08
11	7,71	7,85	7,97	8,08	8,48	8,76	9,41	9,77	10,03	10,22	10,50	10,71	10,80
12	7,55	7,69	7,81	7,91	8,30	8,58	9,21	9,56	9,81	10,00	10,27	10,48	10,56
13	7,42	7,55	7,67	7,77	8,15	8,42	9,04	9,38	9,62	9,81	10,08	10,28	10,36
14	7,31	7,44	7,55	7,65	8,03	8,29	8,89	9,23	9,47	9,65	9,91	10,11	10,19
15	7,21	7,34	7,45	7,55	7,92	8,17	8,76	9,10	9,33	9,51	9,77	9,96	10,04
16	7,13	7,25	7,36	7,46	7,82	8,07	8,65	8,98	9,21	9,39	9,64	9,83	9,91
17	7,06	7,18	7,28	7,38	7,74	7,98	8,56	8,88	9,10	9,28	9,53	9,72	9,79
18	6,99	7,11	7,21	7,31	7,66	7,90	8,47	8,79	9,01	9,18	9,43	9,61	9,69
19	6,93	7,05	7,15	7,24	7,59	7,83	8,39	8,71	8,93	9,09	9,34	9,52	9,60
20	6,88	7,00	7,10	7,19	7,53	7,77	8,32	8,63	8,85	9,02	9,26	9,44	9,51
25	6,68	6,79	6,89	6,97	7,30	7,52	8,05	8,35	8,56	8,71	8,95	9,12	9,19
30	6,54	6,65	6,74	6,83	7,14	7,36	7,87	8,16	8,36	8,51	8,74	8,90	8,97
35	6,45	6,55	6,64	6,72	7,03	7,24	7,74	8,02	8,21	8,36	8,58	8,74	8,81
40	6,38	6,48	6,57	6,65	6,95	7,15	7,64	7,91	8,10	8,25	8,46	8,62	8,69
50	6,27	6,37	6,46	6,54	6,83	7,03	7,50	7,76	7,94	8,09	8,30	8,45	8,51
60	6,21	6,30	6,39	6,46	6,75	6,94	7,40	7,66	7,84	7,98	8,18	8,33	8,39
70	6,16	6,25	6,34	6,41	6,69	6,88	7,33	7,58	7,76	7,89	8,09	8,24	8,30
80	6,12	6,22	6,30	6,37	6,64	6,83	7,28	7,53	7,70	7,83	8,03	8,18	8,24
90	6,09	6,19	6,27	6,34	6,61	6,80	7,23	7,48	7,65	7,78	7,98	8,12	8,18
100	6,07	6,16	6,24	6,31	6,58	6,77	7,20	7,44	7,61	7,75	7,94	8,08	8,14
120	6,04	6,13	6,21	6,28	6,54	6,72	7,15	7,39	7,56	7,69	7,88	8,02	8,07
200	5,97	6,06	6,13	6,20	6,46	6,63	7,05	7,28	7,44	7,56	7,75	7,88	7,94
240	5,95	6,04	6,11	6,18	6,44	6,61	7,02	7,25	7,41	7,53	7,71	7,85	7,90
∞	5,86	5,95	6,02	6,08	6,33	6,50	6,88	7,10	7,25	7,37	7,54	7,67	7,72