10959

EMPRESA BRASILEIRA DE PESQUISA AGROPECUARIA - EMBRAPA

UEPAE DE MANAUS

CURSO DE ESTATÍSTICA EXPERIMENTAL

PARTE II: PLANEJAMENTO E ANÁLISE ESTATÍSTICA

ADROALDO GUIMARÃES ROSSETTI

PARTE II - PLANEJAMENTO E ANÁLISE ESTATÍSTICA

Adroaldo Guimarães Rossetti

5. CONSIDERAÇÕES GERAIS SOBRE REQUISITOS BÁSICOS PARA ANÁLISE ESTATÍSTICA

5.1 - Porque se faz análise estatística

O que exige análise estatística e às vezes até dificulta o trabalho do experimentador, é a presença, em todos os dados obtidos, de efeitos não controlados (que podem ser controlados ou não), tais como:

a) Ligeiras variações na constituição genética do material utilizado nos experimentos.

b) Ligeiras variações no espaçamento, na profundidade dos plantios ou semeaduras.

c) Pequenas diferenças de fertilidade do solo.

d) Maior ou menor disponibilidade de água.

e) Tamanho das parcelas, etc.

Tais efeitos, chamados de variação do acaso ou variação aleatória, não podem ser conhecidos individualmente, portanto podem alterar completamente os resultados experimentais, daí a necessidade de Análise Estatística.

5.1.1 - Medida dos fatores de variação

O coeficiente de variação (CV), que é uma medida absoluta de dispersão, independente da grandeza dos valores e das Unidades em que foram medidas as variáveis dá uma idéia da precisão do experimento. Este porém deve ser utilizado com certa cautela, uma vez pode ser influenciado por determinados fatores.

5.2 - Requisitos básicos para análise estatística

Nem todo tipo de pesquisa requer análise estatística. A pesquisa experimental, baseada em dados de experimentos "controlados" e que envolve fonte (s) de variação, exige, em geral, tal análise, cuja cofiabilidade está bastante ligada às condições da pesquisa, ou experimento (s). Por isso há grande necessidade de serem observados, criteriosamente, no planejamento das pesquisas, aspectos tais como:

5.2.1 - DEFINIÇÃO DO PROBLEMA: Esta é a primeira questão envolvida e deve ser feita de forma clara e concisa. Se o problema não pode ser definido, há pouca chance de ser substancialmente resolvido. Se o problema não foi entendido deve-se ser capaz de formular questões, que respondidas indicam caminhos para solução.

5.2.2 - FORMULAÇÃO DOS OBJETIVOS: Estes podem ser formulados de várias maneiras: sob forma de perguntas e respostas, hipóteses a serem testadas ou efeitos a serem estimados. Devem ser escritos com precisão, sua condição não deve ser vaga ou ambiciosa, permitindo ao experimentador planejar efetivamente os procedimentos experimentais que conduzam a seus alcances. Devem ser listados na ordem de importância, conforme suas situações nos procedimentos experimentais.

5.2.3 - SELEÇÃO DO MATERIAL EXPERIMENTAL: Na seleção do material experimental que abrange desde a área experimental, até o material genético ou botânico, bem como os aspectos que lhes são inerentes, os objetivos do experimento e a população sobre a qual estão sendo feitas inferências, devem ser considerados com bastante ênfase. O material utilizado deve ser representativo da população na qual os tratamentos serão testados.

5.2.4 - SELEÇÃO DOS TRATAMENTOS: O sucesso de um experimento repousa na criteriosa seleção dos tratamentos, cuja evolução responderá às questões propostas.

5.2.5 - SELEÇÃO DO DELINEAMENTO EXPERIMENTAL: Aqui novamente a consideração dos objetivos é importante, mas de modo geral deve-se escolher um delineamento simples que provavelmente fornecerá com maior facilidade a precisão desejada. Aqui devem ser considerados três importantes princípios, inerentes a todos os delineamentos experimentais, indispensáveis aos objetivos da análise estatística. Tais são:

a) REPETIÇÃO: Sua função é proporcionar uma melhor estimativa do erro experimental e maior precisão na medida do efeito do tratamento.

b) ALEATORIZAÇÃO: É a forma de distribuir os tratamentos pelas unidades experimentais, tal que todas as unidades consideradas tenham igual probabilidade de receber um tratamento. Sua função básica é assegurar estimadores não viesados de médias de tratamentos e erro experimental.

c) CONTROLE LOCAL: Este princípio de delineamento experimental, permite, para certas situações, reduzir o erro experimental. Num delineamento de blocos ao acaso, por exemplo, os tratamentos são agrupados nos blocos, de tal modo que se espera uma performance diferente para cada bloco.

5.2.6 - SELEÇÃO DA UNIDADE DE OBSERVAÇÃO (PARCELA): Em experimentos de campo, com plantas, por exemplo, o tamanho e forma das parcelas são importantes e por isso devem ser observados a fim de produzirem a precisão desejada na estimativa dos efeitos dos tratamentos.

5.2.7 - CONTROLE DOS EFEITOS DE UNIDADES ADJACENTES EM CADA UMA DAS OUTRAS

Este é frequentemente considerado completo, pelo uso de linhas de bordaduras e pela aleatorização dos tratamentos.

5.2.8 - CONDUÇÃO DO EXPERIMENTO: Na condução do experimento deve-se usar procedimentos que sejam livres da influência pessoal. Seguir rigorosamente o plano experimental no que concerne a um manejo livre de diferenças entre indivíduos ou que diferenças associadas a ordem da coleção possam ser eliminadas do erro experimental.

5.2.9 - CONSIDERAÇÕES SOBRE OS DADOS COLETADOS: Os dados coletados, se tomadas todas as providências antes expostas, devem avaliar corretamente os efeitos dos tratamentos, alinhados com objetivos do experimento. Evitar distração ou fadiga na coleta de dados. Rechecar imediatamente as observações que apresentem qualquer sintoma de estarem fora do padrão normal. Considerações adicionais devem ser fornecidas para a coleta dos dados, que esclarecerão o desempenho dos tratamentos. O cronograma de coleta deve estar de acordo com o planejado, sua organização deve ser feita de modo a facilitar as análises a serem feitas de modo a evitar a necessidade de transcrição e consequentes erros.

5.2.10 - ESQUEMA DE ANÁLISE ESTATÍSTICA E SUMARIZAÇÃO DOS RESULTADOS

É sempre importante uma exploração detalhada dos dados, antes de qualquer análise mais específica. Nisso as tabelas, gráficos, as medidas simples de tendência central e dispersão, ajudarão muito, além de contribuirem no preparo de resumos e mostrarem os efeitos detectados (esperados?). Esses "resultados" iniciais devem ser comparados com os objetivos do experimento para se ter indicação do seu alcance. A este ponto, às vezes convém uma revisão dos planos iniciais, por um estatístico, por um ou mais colegas da área. A revisão por outros, pode trazer pontos que o experimen tador (autor), passou por cima. Além disso, certas alterações ou ajustamentos podem não só enriquecer o trabalho, como proporcionar um aprendizado consideravelmente maior do trabalho que se está executando.

5.2.11 - ANÁLISE DOS DADOS E INTERPRETAÇÃO DOS RESULTADOS

As análises específicas devem ser feitas rigorosamente de acordo com o que foi planejado: delineamento, testes de comparações múltiplas, etc, levando-se em conta as indicações apresentadas na análise exploratória dos dados. Dessa forma os resultados serão interpretados à luz das condições experimentais, hipóteses testadas e os resultados relacionados aos fatos previamente estabelecidos.

5.2.12 - PREPARAÇÃO DO RELATÓRIO DA PESQUISA: CLARO, COMPLETO E FIEL

Não existe coisa tão ruim como um resultado negativo. Deve-se então examinar com detalhe todos os resultados ou indicações. Por exemplo, se a hipótese nula não foi rejeitada, é possível evidenciar que não há diferenças reais entre os tratamentos testados. Neste caso, checar novamente, se necessário, com outros colegas e rever as conclusões.

Embora haja muito mais passos sobre a parte não estatística, esta é uma parte importante da experimentação pois ajuda o experimentador a delinear seus esperimentos e avaliar objetivamente os dados numéricos por estes produzidos.

6. ANÁLISE DE VARIÂNCIA

6.1. Considerações sobre os dados

A análise de variância, em geral baseada no teste F, cujo objetivo é comparar variâncias ou desvios padrões, exige, entre outras coisas, que os dados utilizados provenham de variáveis aleatórias contínuas. Esta porém, dependendo dos objetivos da pesquisa, de certas propriedades da população, objeto do estudo, da quantidade das observações das variáveis em estudo, de suas distribuições amostrais e de terem sido levados em conta, no planejamento e execução, os princípios básicos da experimentação, pode ser aplicada a dados de amostragem, dados de contagem, ou dados de "scores", entre outros. Nestes dois últimos casos há que se ter bastante cuidado, pois às vezes é preferível a utilização de técnicas não paramétricas, com muito boa eficiência.

6.2. Análise de variância de delineamentos/Transformação de dados

A análise de variância de um experimento qualquer pressupõe um modelo estatístico e a aceitação de algumas hipóteses básicas. Tais hipóteses, uma vez aceitas, dão validade às análises, Estas são:

6.2.1 - Aditividade dos diversos efeitos

6.2.2 - Independência dos erros e_{ij}

6.2.3 - Homogeneidade de variâncias dos erros

6.2.4 - Distribuição normal dos erros.

Dependendo porém, da natureza dos dados, essas hipóteses, quase sempre se verificam, pelo menos aproximadamente. Há, entretanto, necessidade de que elas sejam testadas, pois além da plena segurança dos resultados, ajudam a compreender melhor os dados e consequentemente explorá-los de forma mais completa.

Os teste t e F, em geral não se alteram se a distribuição for apenas aproximadamente normal, ou mesmo que a distribuição se afaste, até certo ponto, da normalidade. A aditividade tem, por certo, importância muito maior, embora quase sempre se verifique. Quando isso não acontece. usa-se a transformação de variáveis para contornar o problema. A transformação logarítimica, quando o desvio padrão é proporcional à média, é apropriada para converter efeitos multiplicativos em aditivos. Esta prática permite, em geral, contornar também as dificuldades de normalidade dos erros.

Se a distribuição é a de Poisson, por exemplo, a transformação para $\sqrt{x}$ é uma boa opção. Se porém, neste caso, os valores são muito baixos $(0 \leq x \leq 15)$, a melhor transformação será $\sqrt{x + 1/2}$.

Para dados com distribuição binomial ou para valores medidos em proporções, uma forma de estabilizar as variâncias é a utilização da transformação angular arc sen $\sqrt{x}$. Nesta caso, as observações iguais a zero devem ser substituídas por $1/2n$ ou $1/4n$ e as iguais a 1, devem ser substituídas por $1 - 1/2n$ ou $1 - 1/4n$.

Dessa forma, antes de se realizar a análise de variância de quaisquer dados, é prudente que se proceda pelo menos, testes de aditividade, homogeneidade e aderência.

Vale ressaltar que em certos casos tem-se usado abusivamente essa prática e o que é pior, sem se assegurar da real necessidade de transformação dos dados. Tem se verificado que transformações em determinados tipos de observações mudam totalmente os resultados, chegando, em certos casos, a invertê-los.

6.3. Interpretações dos resultados

Este é sem dúvidas um dos pontos mais importantes envolvidos na pesquisa experimental. A utilização de certos "pacotes" de análise estatística via computador, por exemplo, exige certas precauções, como um rigoroso exame dos resultados "listados", pois há, quase sempre, a necessidade de se realizar cálculos complementares para serem tiradas determinadas conclusões, uma vez que tais "pacotes" nem sempre são feitos por profissionais da área de estatística. Já se tem constatado, até erros, de certa forma graves, que só uma análise rigorosa ajuda detectá-los. Ao contrário disso, se tem observado certo hábito de transcrever os resultados como saem do computador, sem muitas vezes, fazer sequer uma revisão.

Nesta fase é que mais pesa o planejamento da pesquisa, pois se ela foi bem planejada e o experimento bem manejado já há uma grande margem de segurança dos resultados, daí a necessidade de se procurar um estatístico na fase de planejamento. Tem-se percebido, com frequência, a existência de certo divórcio entre o planejamento, que às vezes nem existe, a condução e a análise, levando assim a interpretações, nem sempre condizentes com a realidade.

Um outro fato importante é que tem se tornado cada vez mais frequente se tirar conclusões sob significações aos níveis de 1% e 5%. Caso haja diferenças de décimos, o pesquisador diz, em geral, que não foi significativo, tira suas conclusões sobre esses níveis, disperdiçando às vezes grandes resultados. Por outro lado, mesmo se trabalhando com materiais ou culturas desconhecidas (heterogêneas), com CV_s bastante altos, ainda se tira conclusões exclusivamente sobre esses níveis.

7. DELINEAMENTOS EXPERIMENTAIS

7.1. Inteiramente casualizados

Este delineamento é caracterizado quando os tratamentos são arranjados completamente ao acaso, em área seguramente uniforme, quando se trata de material experimental de pequena variação. O número de repetições para os tratamentos pode ser variado. Neste caso porém, os testes de comparação de média são apenas aproximados.

Esse delineamento é muito usado em experimentos de laboratório, em estudo de técnicas metodológicas, em alguns estudos em casa de vegetação, em alguns estudos com animais, em raras situação, em pequenos experimentos de campo, quando se tem um terreno bem uniforme, ou quando se tratar de experimentos em "vasos", desde que a posição destes for mudada, com frequência, ao acaso.

As vantagens deste delineamento são:

i) Tem um "layout " bastante simples

ii) Permite estimar as somas de quadrados com um número máximo de graus de liberdade para o erro (resíduo).

iii) A análise estatística de todo o delineamento é bastante simples.

iv) Permite número de repetições diferentes para vários tratamentos, sem induzir complicações em muitos casos nas análises.

A grande desvantagem deste delineamento é que ele é usualmente adequado para experimentos com poucos tratamentos e material experimental homogêneo. Quando se usa um grande número de tratamentos a variação das respostas desses tratamentos geralmente cresce bastante, caindo muito a confiabilidade dos resultados.

7.1.1 - Modelo estatístico

Num delineamento inteiramente casualizado, qualquer observação pode ser expressa como

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

onde μ representa a média geral; t_i é o efeito do i-ésimo tratamento (i=1,2,....I) e e_{ij} é o efeito do erro experimental.

Se as quatro hipóteses (6.2), sobre o modelo, são aceitas, o uso do metodo dos mínimos quadrados, para minimisar os erros, permite obter estimativas para as somas de quadrados, pelas relações:

TOTAL: $SQT_o = \Sigma x_{ij}^2 - C$

onde $C = \frac{(\Sigma x_{ij})^2}{N}$, é o fator de correção; x_{ij} é o valor da observação i, na repetição j e N é o número total de observações.

TRATAMENTOS: $SQT = \frac{\Sigma T_i^2}{r_i} - C$

onde

T_i é o total do i-ésimo tratamento e r_i o número de repetições do tratamento i.

ERRO: A soma de quadrado do erro (SQR), neste caso pode ser obtida por diferença, isto é:

$$SQR = SQT_o - SQT$$

A análise de variância é pois:

Fonte de variação	G.L.	SQ	QM	F
Entre tratamentos	t - 1	SQT	SQT/t-1	$\frac{QMT}{QMR}$
Entre unidades experimentais dentro de tratamentos	t(r - 1)	SQR	SQR/t(r-1)	
Total	tr - 1	SQT_o		

7.1.2 - Exemplo

Considere-se o ganho de peso arredondados para kg de 5 animais, pela aplicação de 4 tipos de rações, sob as mesmas condições (i , é, controladas todas as outras possíveis influências, como tipo de baia empregada, etc.).

Reações

A	B	C	D	
35	40	39	27	
19	35	27	12	
31	46	20	13	
15	41	29	28	
30	33	45	30	
130	195	160	110	Totais
26	39	32	22	Médias

cálculos das somas de quadrados dos diversos componentes.

$\Sigma x^2 = 35^2 + 19^2 + \ldots\ldots\ldots + 30^2 \rightarrow \Sigma x^2 = 19625$

$\Sigma x = 35 + 19 + \ldots\ldots\ldots + 30 \rightarrow \Sigma x = 595$

$C = (595)^2/20 \rightarrow C = 17701,25$

$SQT_o = \Sigma x^2 - C = 19625 - 17701,25 = 1923,75$

$SQT = \frac{1}{5}(130^2 + 195^2 + 160^2 + 110^2) - C = 823,75$

$SQR = SQT_o - SQT = 1923,75 - 823,75 = 1100,00$

com isso obtém-se a análise de variância que se segue:

Fonte de variação	GL	SQ	QM	F
Entre tratamentos (rações)	3	823,75	274,58	3,99*
Dentro de tratamentos (erro)	16	1.100,00	68,75	
Total	19	1.923,75		

* Diferença significativa a 0,05 de probabilidade

Considere-se que por qualquer motivo a primeira ração (A) foi fornecida a apenas 4 animais. Neste caso, todos os cálculos anteriores são efetuados de forma análoga, com os dados existentes, para o que se obteria:

$\Sigma x^2 = 18400, \quad \Sigma x = 560, \quad C = 16505,26$

$SQT_o = 1894,74$ e

$SQT = \frac{95^2}{4} + \frac{195^2 + 160^2 + 110^2}{5} - C \rightarrow SQT = 895,99$

A análise de variância então será:

Fonte de variação	GL	SQ	QM	F
Tratamentos	3	895,99	298,66	4,49**
Resíduo	15	998,75	66,58	

** Diferença significativa a 0,01 de probabilidade

7.1.3 - Exemplo

Em um experimento de laboratório foram testados 4 níveis (A,B,C,D) de um inseticida, para mandarová. Os resultados (nº de lagartas mortas), 12 horas após a aplicação, nas 5 repetiçoes estão no quadro a seguir. Faça a respectiva análise de variância.

a) Sem transformar os dados
b) Transformando-os adequadamente
c) Suponha, no nível A, a segunda observação perdida (sem informação) e no nível C foram perdidas a primeira e a terceira observações. Considerando esses fatos, faça a análise de variância.

Tratamentos

A	B	C	D
7	5	6	5
8	5	7	7
8	6	8	7
10	6	9	8
12	8	10	8

7.2. Blocos casualizados ou blocos ao acaso

Se todo o material experimental, ou parte deste, ou a área, não são homogêneos, é possível fazer-se uma extratificação em grupos homogêneos, ou em subgrupos, de tal modo que haja um controle relativamente homogêneo do material, em cada extrato. Aqui intervém o princípio do **controle local**. Os blocos ao acaso constituem talvez o delineamento experimental mais importante e mais utilizado, principalmente na experimentação de campo, onde, por hipótese, há um maior número de fatores "não controláveis".

As grandes vantagens deste delineamento são:

i) **Precisão**. Tem-se demonstrado ser o delineamento mais preciso para a maioria dos trabalhos experimentais. A eliminação da soma de quadrado de blocos, da soma de quadrado do erro resulta em decréscimo da variância estimada (quadrado médio do erro).

ii) **Flexibilidade.** Não há restrições quanto ao número de tratamentos ou de repetições. Em geral no mínimo duas repetições são requiridas para se obter significância nos testes.

iii) **Facilidade da análise.** A análise estatística é simples e rápida. Além disso, o erro de comparação de qualquer tratamento, pode ser isolado e qualquer número de tratamentos pode ser omitido da análise sem complicação. Essas facilidades podem ser utilizadas quando certas diferenças entre tratamentos forem muito grandes; a ponto de poder comprometer os resultados, ou quando o erro experimental produzir comparações heterogêneas.

As desvantagens do delineamento de blocos ao acaso são: o número de tratamentos não deve ser muito elevado e cada bloco não pode conter grande variabilidade. Pelo contrário, em cada bloco deverá haver a maior homogeneidade possível.

O diagrama seguinte ilustra um "layout" experimental de um experimento com 5 tratamentos e 4 blocos.

Bloco I	E (1)	A (2)	C (3)	B (4)	D (5)
Bloco II	A (10)	D (9)	B (8)	C (7)	E (6)
Bloco III	B (11)	C (12)	A (13)	E (14)	D (15)
Bloco IV	E (20)	D (19)	A (18)	B (17)	C (16)

7.2.1 - Modelo estatístico

Num delineamento de blocos casualizados, qualquer observação pode ser expressa como

$$Y_{ij} = \mu + t_i + b_j + e_{ij},$$

onde Y_{ij} é o valor observado relativo à parcela que recebeu o tratamento i no bloco j; μ é a média geral; t_i mede o efeito do tratamento i; b_j mede o efeito do bloco j e e_{ij} é a contribuição do acaso, isto é, a parte da variação devida a fatores não controlados.

7.2.2 - Cálculo das somas de quadrados

A imediata aplicação do método dos mínimos quadrados ao modelo, conduz às fórmulas que se seguem:

$SQT_o = \Sigma x^2 - C$, onde $C = \frac{(\Sigma x)^2}{b \times t}$, referentes a soma de quadrado total e fator de correção.

$SQT = \frac{1}{b} \Sigma T_i^2 - C$ soma de quadrado devido a tratamentos.

$SQB = \frac{1}{t} \Sigma B_j^2 - C$ soma de quadrado de blocos

$SQR = SQT_o - SQT - SQB$ soma de quadrados dos erros.

No caso do exemplo ilustrado no "layout", tem-se o quadro de análise a seguir:

Fonte de variação	GL		SQ	QM	F
Entre blocos	b-1	3	SQB	SQB/b-1	
Entre tratamentos	t-1	4	SQT	SQT/t-1	QMT/QMR
Erro	(b-1)(t-1)	12	SQR	SQR/(b-1)(t-1)	
Total	bt-1	19	SQT_o		

7.2.3 - Análise para mais de uma observação por unidade experimental.

Considere-se que em cada parcela haja k observações. O modelo estatístico é, neste caso

$$Y_{ijk} = \mu + t_i + b_j + tb_{ij} + l_{ijk},$$

que dá origem ao quadrado de análise de variância seguinte.

Fonte de variação	GL	SQ	QM	F
Entre blocos	b-1			
Entre tratamentos	t-1			
Erro experimental	(b-1)(t-1)			
Erro amostral	bt (k-1)			
Total	btk-1			

Neste caso, os efeitos de blocos e tratamentos são testados com a interação (erro experimental) e este, por sua vez, é testado com o resíduo (erro amostral. Se o maior desses erros for até, 3 ou 4 vezes maior que o menor, eles podem ser juntados em um só, o que recai no quadro anterior.

7.2.4 - Exemplo

Os dados a seguir, referem-se à variável stand, aos sessenta dias, de um experimento de competição de 12 espaçamentos, com 3 repetições, em viveiro de seringueira, já transformados para $\sqrt{x}$. Faça a respectiva análise de variância.

Blocos / Tratam.	I	II	III	Totais
1	7,35	6,85	5,57	
2	6,32	7,21	6,32	
3	6,48	4,47	5,20	
4	6,24	5,38	5,29	
5	6,16	5,48	6,32	
6	5,48	5,48	5,48	
7	5,74	4,47	5,00	
8	5,48	5,48	4,47	
9	6,71	5,74	6,00	
10	6,48	6,63	4,90	
11	6,48	7,00	5,83	
12	6,16	5,74	6,08	
Totais				

7.2.5 - Exemplo

Os dados a seguir referem-se à variável Diâmetro do caule de um experimento em viveiro de seringueira com 4 plantas por parcelas, no qual foram testados os efeitos de 5 fungicidas, para controle do "mal das folhas". O experimento tinha 3 repetições. Faça a análise de variância:

a) Levando em conta as 4 plantas por parcela

b) Juntando os dois "erros"

c) Comente os resultados

	I				II				III			
	P_1	P_2	P_3	P_4	P_1	P_2	P_3	P_4	P_1	P_2	P_3	P_4
F_1	1,20	2,12	0,86	2,01	2,10	0,91	1,86	1,84	0,89	1,25	1,35	1,45
F_2	1,08	1,08	1,18	1,06	1,86	0,86	1,04	1,09	1,98	1,00	0,98	1.29
F_3	0,96	0,56	1,00	0,96	0,90	1,07	1,20	1,64	2,07	1,12	1,75	1.18
F_4	0,74	1,40	0,95	2,04	1,00	1,98	1,86	1,52	1,26	1,04	0,56	1.64
F_5	2,01	1,00	1,08	0,89	1,03	1,00	1,92	1,10	1,30	1.57	1.80	0,99

7.2.6 - Parcela perdida em delineamento de blocos ao acaso

Tem-se observado com certa frequência a tendência de se substituir uma parcela perdida pela média das existentes, ou pelo valor zero. O zero indica que não houve produção, por exemplo, na parcela, o que é diferente de se ter perdido-a. Por outro lado, a substituição pela média das parcelas não é uma prática correta. Uma estimativa desse valor, embora não substitua com precisão o verdadeiro valor, mas dá uma boa aproximação. Demonstra-se que sendo y uma parcela perdida, uma estimativa é encontrada pela relação

$$Y = \frac{b\,B + tT - G}{(b-1)\ (t-1)},$$

onde B é o total das parcelas existentes no bloco em que figura a parcela perdida, b é o número de blocos, t o número de tratamentos, T é o total das parcelas existentes no tratamento onde ocorreu a perda da parcela e G é o total das parcelas disponíveis.

Obtido esse valor, os cálculos são efetuados da maneira usual, com a única diferença da perda de um grau de liberdade para o resíduo.

Cumpre salientar que embora o quadrado médio do resíduo esteja corretamente estimado, o correspondente a tratamentos fica ligeiramente superestimado. Para corrigi-lo, basta subtrair da soma de quadrados para tratamentos o número u, dado pela expressão:

$$u = \frac{t-1}{t}\left(y - \frac{B}{t-1}\right)^2$$

Observa-se, consequentemente, que haverá uma ligeira alteração no quadrado médio para tratamentos e consequentemente no valor de F. Tal alteração em geral, influi pouco, de modo que muitas vezes pode ser dispendada.

7.2.7 - Exemplo

No exemplo 7.2.4, considere perdida a parcela referente a repetição III, do tratamento 9 e faça a análise de variância, comentando o resultado.

7.3. Quadrados latinos

Enquanto no delineamento de blocos ao acaso impõe-se a restrição que todos os tratamentos apareçam no mesmo bloco em igual ou proporcional número de vezes, no delineamento em quadrados latinos, duas restrições são impostas. Quando a área experimental dividida em linhas e colunas, cada tratamento aparece apenas uma vez em cada linha e uma vez em cada coluna.

Esses "layouts" dos delineamentos em quadrados latinos, proporcionam duplo controle local. Sua aplicação, em trabalhos experimentais é bastante variada: Em experimentos industriais, de laboratórios, casa de vegetação, experimentos de campo, visando a eliminar a heterogeneidade do solo em duas direções ou para comparar um grupo de variedades ou tratamentos fertilizantes, entre tantos outros.

7.3.1 - Vantagens e desvantagens

As vantagens do delineamento em quadrado latino são:

i) Com dois rumos de estratificação ou grupos, o quadro latino controla melhor a variação do que o inteiramente casualizado e o de blocos casualizados. Além disso dá um valor menor para o quadrado médio do erro.

ii) A análise é simples, isto é, apenas um pouco mais complicada que no caso dos blocos casualizados.

iii) Existem procedimentos analíticos para avaliar a omissão de um ou mais tratamentos, linhas ou colunas.

As desvantagens do quadrado latino são:

i) O número de tratamentos é limitado ao número de linhas e colunas e deve ser um número de quadrado perfeito. Para mais de 10 tratamentos é raramente usado. Os mais usados são os de 5 x 5 a 8 x 8.

ii) Para menos que 5 tratamentos a alocação dos graus de liberdade para controle da heterogeneidade é desproporcionalmente grande.

7.3.2 - Modelo estatístico

Se uma única observação é feita em cada parcela de um delineamento em quadrado latino, o valor dessa observação pode ser expressa por:

$$Y_{ijk} = \mu + l_i + c_j + t_k + e_{ijk} ,$$

onde μ representa a média da população, l_i é o efeito da i-ésima linha; c_j é o efeito da j-ésima coluna, t_k é o efeito do k-ésima tratamento e e_{ijk} é o efeito aleatório, não controlável pelo experimentador.

A aplicação do método dos quadrados mínimos, permite estimar-se as somas de quadrados dos diversos componentes, como segue:

$SQT_o = \Sigma x^2 - C$ (soma de quadrado total)

$SQL = \frac{1}{k} \Sigma L_i^2 - C$ (soma de quadrados de linhas)

$SQC = \frac{1}{k} \Sigma C_j^2 - C$ (soma de quadrados de colunas)

$SQT = \frac{1}{k} \Sigma T_k^2 - C$ (soma de quadrados de tratamentos)

$SQR = SQT_o - SQL - SQC - SQT$ (soma de quadrado do erro)

A análise de variância correspondente a um quadrado latino k x k, é a que se segue:

Fonte de variação	GL	SQ	QM	F
Linhas	k-1	SQL	SQL/k-1	QML/QMR
Colunas	k-1	SQC	SQC/k-1	QMC/QMR
Tratamentos	k-1	SQT	SQT/k-1	QMT/QMR
Resíduo	(k-1) (k-2)	SQR	SQR/(k-1)(k-2)	
Total	$k^2 - 1$	SQT_0		

7.3.3 - Exemplo

Num experimento de competição de variedade de cana-de-açúcar, foram usadas cinco variedades (A, B, C, D, E), dispostas em quadrado latino 5 x 5. As produções, em kg por parcela, são dadas na tabela seguinte.

	C_1	C_2	C_3	C_4	C_5
L_1	D 432	A 518	B 458	C 583	E 331
L_2	C 724	E 478	A 524	B 550	D 400
L_3	E 489	B 384	C 556	D 297	A 420
L_4	B 494	D 500	E 313	A 486	C 501
L_5	A 515	C 660	D 438	E 394	B 318

Faça a análise de variância dos dados.

8. TESTES DE SIGNIFICÂNCIA (comparações múltiplas)

8.1. Contrastes e contrastes ortogonais

8.1.1 - Contrastes

Sejam, por exemplo, m_1, m_2, m_3 e m_4 as respectivas médias de quatro tratamentos. Diz-se que y_1 é um contraste das médias m_1 e m_2 se e só se $y_1 = m_1 - m_2 = 0$ $\rightarrow m_1 = m_2$. Do mesmo modo se $y_2 = m_1 + m_2 + m_3 - 3m_4 = 0$, diz-se que y_2 é um contraste. Assim o que caracteriza um contraste é que se as médias que nele ocorrem forem todas iguais, o contraste será nulo. Como se vê acima, para que isso aconteça, a soma algébrica dos coeficientes deve ser nula.

8.1.2 - Contrastes ortogonais

Dois contrastes y_1 e y_2, são ditos ortogonais se e somente se:

i) Em cada contraste a soma algébrica dos coeficientes for nula

ii) A soma dos produtos dos coeficientes correspondentes do primeiro pelo segundo contraste for nula, isto é. Seja $Y_1 = m_1 - m_2 + 0.m_3$

$$Y_2 = m_1 - m_2 - 2m_3$$

coeficientes de Y_1: 1 -1 0 = 0
coeficientes de Y_2: 1 +1 -2 = 0 (i)

Produto dos coeficientes : 1 -1 0 = 0 (ii)

Logo Y_1 e Y_2 são contrastes ortogonais

Na análise de variância os contrastes ortogonais têm grande importância. Do ponto de vista prático, a ortogonalidade indica que a variação de um contraste é totalmente independente da variação de qualquer outro que lhe seja ortogonal.

Num experimento com 4 tratamentos, cujas médias sejam m_1, m_2, m_3 e m_4, existem 3 graus de liberdade para tratamentos isso indica que se pode construir três contrastes ortogonais, como os que se seguem, por exemplo:

$$y_1 = m_1 - m_2$$
$$y_2 = m_1 + m_2 - 2m_3$$
$$y_3 = m_1 + m_2 + m_3 - 3m_4$$

Mas poderiam também ser os seguintes:

$$y_1 = 3m, - 2m_2 - m_3$$
$$y_2 = m_2 - 2m_3 + m_4$$
$$y_3 = 3m_1 + 5m_2 - m_3 - 7m_4$$

Como se vê, a construção dos três contrastes ortogonais (3 graus de liberdade) dependem do interesse do experimentador.

Observe-se que m_1, m_2, m_3, m_4 são as verdadeiras médias dos tratamentos que em geral não se conhece. Conhece-se porém, suas estimativas, que são indicadas respectivamente por $\hat{m}_1$, $\hat{m}_2$, $\hat{m}_3$ e $\hat{m}_4$, assim como uma estimativa do desvio padrão residual.

8.2. Prova dos contrastes ortogonais/teste de t

O teste t, chamado "t de Student", usado para comparar conjunto de médias, exige, como requisito, para sua aplicação, as seguintes condições:

i) As comparações feitas pelo teste t devem ser escolhidas antes de serem examinados os dados.

ii) Pode-se fazer no máximo tantas comparações quantos são os graus de liberdade para tratamentos e os contrastes devem ser ortogonais.

Estas condições determinam a plena validade deste teste.

Provar um contraste pelo teste de t implica em verificar se ele difere de zero, ou seja, se o contraste em prova é o $\hat{y}_1$, por exemplo, isto se consegue pela fórmula.

$$t = \frac{\hat{y}_1 - 0}{s(\hat{y}_1)}$$

onde s $(\hat{y}_1)$ é o erro padrão do contraste, que é calculado do seguinte modo:

$$V(\hat{y}_1) = \left(\frac{c_1^2}{r_1} + \frac{c_2^2}{r_2} + \ldots\ldots + \frac{c_n^2}{r_n}\right) s^2 \rightarrow s(\hat{y}_1) = \sqrt{\left(\frac{c_1^2}{r_1} + \ldots\ldots + \frac{c_n^2}{r_n}\right) s^2}$$

onde c_i são os coeficientes das médias que formam o contraste, r_i é o número de repetições para cada tratamento que deu origem à média $\hat{m}_i$.

8.2.1 - Exemplo

Sejam $\hat{m}_1 = 26{,}0$, $\hat{m}_2 = 24{,}8$, $\hat{m}_3 = 22{,}8$ e $\hat{m}_4 = 24{,}0$ estimativas das médias de 4 tratamentos, com 6 repetições por tratamento, cuja análise de variância forneceu um QMR = 1,44 (estimativa da variância), portanto s = 1,20 (estimativa do desvio padrão).

. Considere-se testar o contraste

$\hat{y}_2 = \hat{m}_1 + \hat{m}_2 - 2\hat{m}_3$

$\hat{y}_2 = 26{,}0 + 24{,}8 - 2 \times 22{,}8 \rightarrow \hat{y}_2 = 5{,}2$

$$V(\hat{y}_2) = \left(\frac{1}{6} + \frac{1}{6} + \frac{4}{6}\right) s^2 \rightarrow V(\hat{y}_2) = s^2$$

como $V(\hat{y}_2) = s^2 \rightarrow s(\hat{y}_2) = s = 1{,}20$, logo:

$$t = \frac{5{,}2 - 0}{1{,}20} \rightarrow t = 4{,}33$$

com 20 graus de liberdade para o resíduo a tabela de t fornece os valores para o nível de 0,05, 2,09 e para 0,01, 2,84. Isto é, o contraste difere significativamente de zero, o que equivale, neste caso, a dizer que a média dos dois primeiros tratamentos difere da média do terceiro.

No caso particular de um contraste entre duas médias apenas; $\hat{y}_1 = \hat{m}_1 - \hat{m}_2$, por exemplo, tem-se

$$t = \frac{\hat{m}_1 - \hat{m}_2}{s\sqrt{\frac{1}{r_1} + \frac{1}{r_2}}}$$

8.2.2 - Exemplo

. No exercício 7.1.3, escrever os contrastes possíveis de seu interesse e prová-los pelo teste t. (Fazer pelo menos um contrastes com mais de duas médias para testá-lo). Teste, nesse caso, o contraste $\hat{y} = \hat{m}_1 + \hat{m}_2 - \hat{m}_3 - \hat{m}_4$.

8.3. Teste de Duncan

O teste de Duncan, para comparação de médias de tratamentos, permite chegar-se a resultados detalhados, indica resultados significativos onde o teste de Tukey, por exemplo, não permite obter significação estatística. Este teste exige, para que seja exato, que todos tratamentos tenham o mesmo número de repetições.

O seu uso, como o de t e Tukey, exige tabelas especiais. Ele é baseado na am

SUMÁRIO

plitude total mínima significativa, abrangida por um grupo de médias, dada por:

$$D = z \frac{s}{\sqrt{r}} ,$$

onde r é o número de repetições, s é o desvio padrão e z é tirado das tabelas, para o número de médias ordenadas abrangidas pelo contraste em estudo e o número de graus de liberdade do resíduo.

8.3.1 - Exemplo

Considerando o exemplo dado em 7.1.2, onde $\hat{m}_1 = 26,0$, $\hat{m}_2 = 39,0$, $\hat{m}_3 = 32,0$ e $\hat{m}_4 = 22,0$, de 5 repetições e QMR = 68,75, portanto s = 8,29 e g.l. do resíduo igual 16.

Considere-se em primeiro lugar as 4 médias ordenadas, abrangidas pelo contraste, ou seja, a comparação entre a maior e a menor das médias, que abrange todas as quatro, isto é.

$$z\ (4,\ 16) = 3,23 \rightarrow D_4 = 3,23 \times \frac{8,29}{\sqrt{5}} \rightarrow D_4 = 11,97$$

Compare-se agora a segunda maior com a menor. Esse contraste abrange 3 médias. Então:

$$z\ (3,\ 16) = 3,15 \rightarrow D_3 = 3,15 \times \frac{8,29}{\sqrt{5}} \rightarrow D_3 = 11,68$$

Este valor serve para testar o contraste entre a maior e a penúltima médias.

Compare-se agora a terceira maior com a menor. Esse contraste abrange 2 médias. Então:

$$z\ (2,\ 16) = 3,00 \rightarrow D_2 = 3,00 \times \frac{8,29}{\sqrt{5}} \rightarrow D_2 = 11,12$$

Este último caso corresponde ao que daria o teste t.

Assim tem-se que:

$\hat{m}_2 = 39,0$ $\quad$ $\hat{m}_3 = 32,0$ $\quad$ $\hat{m}_1 = 26,0$ $\quad$ $\hat{m}_4 = 22,0$

Médias sobre a mesma barra não diferem significativamente ao nível de 0,05 de probabilidade;

É usual também, e até mais prático, representar-se por letras, como segue:

$\hat{m}_2 = 39,0$ a

$\hat{m}_3 = 32,0$ ab

$\hat{m}_1 = 26,0$ b

$\hat{m}_4 = 22,0$ b

Assim, médias seguidas da mesma letra não diferem estatisticamente, ao nível de 0,05 de probabilidade pelo teste de Duncan

Note-se que, sempre que num grupo de médias, a maior não difere da menor, pelo teste de Duncan, não se admite diferença significativa, pelo mesmo teste, entre médias intermediárias.

8.3.2 - Exemplo

Considere o exercício 7.1.3 e compare as médias dos tratamentos, pelo teste de Duncan.

8.4. Teste de Tukey

O teste de Tukey, baseado na "amplitude total estudentizada", é frequentemente utilizado para comparar qualquer contraste de duas médias de tratamentos. O número de contrastes a ser utilizado, neste caso, é determinado pela combinação do número de tratamentos, tomados dois a dois $\binom{t}{2}$. No caso de um experimento com 4 tratamentos, ter-se-á 6 comparações ou 6 contrastes a serem testados pelo teste de Tukey. O teste é exato e bastante simples, quando o número de repetições é o mesmo para todos os tratamentos. Quando entretanto o número de repetições é diferente, ele é apenas aproximado.

A diferença mínima significativa é calculada pela expressão.

$$\Delta = q\sqrt{\frac{QMR}{r}} \quad \text{ou} \quad \Delta = q \times \frac{s}{\sqrt{r}}$$

onde q é o valor da amplitude total estudentizada, encontrado na tabela, ao nível α de probabilidade, para n tratamentos e n_1 graus de liberdade para o resíduo ou erro.

8.4.1 - Exemplo

Considerando ainda o exemplo 7.1.2, com 4 tratamentos e 5 repetições, cujas médias dos tratamentos são $\hat{m}_1 = 26{,}0$, $\hat{m}_2 = 39{,}0$, $\hat{m}_3 = 32{,}0$ e $\hat{m}_4 = 22{,}0$, 16 g.l. para o resíduo e QMR = 68,75 ou s = 8,29, tem-se que:

$$q\ (4,\ 16{,}5\%) = 4{,}05 \qquad \Delta_5 = 4{,}05 \times \frac{68{,}75}{5} \qquad \Delta_5 \doteq 15{,}02$$

Portanto

$\hat{m}_2 = 39{,}0$ a

$\hat{m}_3 = 32{,}0$ ab

$\hat{m}_1 = 26{,}0$ ab

$\hat{m}_4 = 22{,}0$ b

Médias seguidas da mesma letra não diferem estatisticamente, ao nível de 0,05 de probabilidade.

Caso os tratamentos tenham números diferentes de repetições, o teste de Tukey pode ainda ser usado, mas então é apenas aproximado. Nesse caso o cálculo de Δ será efetuado por:

$$\Delta = q\sqrt{\left(\frac{1}{2}\right)\hat{V}\ (\hat{y})}$$

Se no caso do exemplo anterior, $\hat{m}_1$ tivesse 5 repetições e $\hat{m}_2$ apenas 4, ter-se-ia

$$q\ (4,\ 15{,}5\%) = 4{,}08 \qquad \Delta = 4{,}08\sqrt{\frac{1}{2} \times \left(\frac{1}{5} + \frac{1}{4}\right) \times 68{,}75} \qquad \Delta = 16{,}05$$

E a diferença observada ainda continuaria sendo significativa.

8.4.2 - Exemplo

Considere ainda o exemplo 7.1.3 e compare as médias dos tratamentos, pelo teste de Tukey.

a) Usando o mesmo número de repetições para todos os tratamentos

b) Considere o tratamento B com apenas 4 repetições e o D com 3 repetições.

9. ALGUNS TIPOS DE EXPERIMENTOS

9.1. Levantamentos por amostragem

Esta é uma prática que dependendo de certas condições dá bons resultados. É usado quando se tem um plantio, por exemplo, e sobre ele se deseja tirar informações sobre determinadas variáveis. Para que se consiga condições experimentais ideais para justificar a análise dos dados, é importante que se leve em conta os princípios básicos da experimentação, na distribuição dos tratamentos ou coleta das informações, de modo que as hipóteses exigidas pela análise de variância sejam satisfeitas de acordo como se discutiu na seção 6.

9.1.1 - Exemplo

Suponha que em um viveiro instalado para produção de mudas (sem fins de pesquisa), após sua instalação precisa-se montar um experimento para testar fungicidas, por exemplo. Por estratificação da área, aloca-se as parcelas ao acaso, levando-se em conta todos os aspectos anteriormente discutidos. A análise dos dados, dependerá do "delineamento" que se optar, em função dos objetivos da pesquisa.

9.2. Experimentos fatoriais: Aplicação e análise

Experimentos fatoriais são aqueles que incluem todas as combinações de vários conjuntos de tratamentos ou fatores. Os experimentos fatoriais são, em geral, mais eficientes do que os experimentos simples, com um só conjunto de tratamentos e permitem tirar conclusões mais gerais.

Sua utilidade é bastante ampla. Em experimentos de cultivares e espaçamentos, por exemplo, permitem tirar conclusões sobre a melhor combinação cultivar x espaçamento. Em experimentos de adubação, sua utilidade é bastante grande, permitindo informações bastante sólidas, traçar curvas de respostas, determinar pontos críticos, etc.

9.2.1 - Vantagens e desvantagens

Entre as vantagens, destacam-se:

i) Todas as unidades experimentais são utilizadas na avaliação dos efeitos, resultando na maior eficiência do uso dos recursos.

ii) Os efeitos são avaliados sobre uma vasta extensão de condições com o mínimo de recursos externos.

iii) Uma estimativa da interação dos fatores é obtida.

iv) Estimadores não viesados de efeitos são obtidos quer haja ou não tendência.

v) Um conjunto de tratamentos fatoriais é ótimo para estimar efeitos e interações

A principal desvantagem dos experimentos fatoriais é que o número de tratamentos aumenta rapidamente, trazendo alguns problemas tanto para a condução do experimento, podendo afetar a precisão, como para a análise dos dados.

9.2.2 - Modelo estatístico

Devido a multiplicidade dos arranjos fatoriais existentes não é razoável escrever um modelo geral para os fatoriais. Isso, além de complicado, pouco contribuiria, para os efeitos deste curso, uma vez que para se escrever um modelo mais geral ter-se-ia que lançar mão dos fatoriais da série p^n (n fatores e p níveis). Daí surgem os casos particulares 2^n, 3^n, p^2, p^3, etc.

Considere-se, para efeito ilustrativo um fatorial da série p x q, em blocos casualizados, respectivamente com j e k níveis cada fator. O modelo estatístico, neste caso é o que se segue.

$$Y_{ijk} = \mu + \rho_i + \alpha_j + \beta_k + \alpha\,\beta_{ik} + e_{ijk},$$

onde μ representa a média da população; ρ_i é o efeito do i-ésimo bloco; α_j é o efeito do j-ésimo nível do fator A; β_k é o efeito do j-ésimo nível do fator B; $\alpha\,\beta_{jk}$ é o efeito da interação do j-ésimo nível do fatorA com o k-ésimo nível do fator B e e_{ijk} e o efeito aleatório.

Os cálculos dos componentes da análise de variância podem ser obtidos, em quaisquer séries de fatoriais pela aplicação do método dos mínimos quadrados. Há porém outros procedimentos práticos, mais simples, que além de chegar diretamente aos resultados, dispensam maiores conhecimentos de matemática.

9.2.3 - Exemplo

Os dados a seguir são produções, em t/ha, de parcelas de um experimento de adubação, de 2^3, em blocos casualizados, N, P, K, em 4 blocos.

Trat. / Blocos	Controle	N	P	K	NP	NK	PK	NPK	Totais
I	1,32	1,80	1,66	2,58	1,72	2,72	2,26	2,95	17,01
II	2,12	2,20	2,66	3,56	3,82	3,20	2,08	3,28	22,95
III	1,75	2,95	1,73	2,86	2,62	2,25	1,95	2,40	18,51
IV	2,35	2,96	2,58	2,75	3,00	2,75	2,70	3,35	22,44
Totais	7,54	9,91	8,63	11,75	11,19	10,92	8,99	11,98	80,91

Nesse experimento tem-se $2^3 = 8$ tratamentos. Se se fizer a análise de variância, do modo usual, tem-se:

Fonte de variação	GL	SQ	QM	F
Blocos	3	3,2011		
Tratamentos	7	4,5246	0,6464	3,88**
Resíduo	21	3,4955	0,1665	
Total	31	11,2212		

** Diferença significativa ao nível de 0,01 de probabilidade

Entretanto, este resultado pouco ou quase nada interessa num esperimento fatorial, onde deve-se isolar os efeitos dos diversos fatores e testá-los isoladamente. Uma técnica bastante simples para calcular esses efeitos isolados é a que se segue. No quadro a seguir estão indicados todos os contrastes a serem estudados, com o sinal (-) indicando -1 e (+), indicando +1.

	(1)	N	P	K	NP	NK	PK	NPK
Efeito de N	-	+	-	-	+	+	-	+
Efeito de P	-	-	+	-	+	-	+	+
Efeito de K	-	-	-	+	-	+	+	+
Interação N x P	+	-	-	+	+	-	-	+
Interação N x K	+	-	+	-	-	+	-	+
Interação P x K	+	+	-	-	-	-	+	+
Interação N x P x K	-	+	+	+	-	-	-	+

Os sinais para as interações são obtidos pelo produto dos sinais dos efeitos principais. Cálculos dos efeitos dos diversos contrastes.

N: O efeito de N é medido pelo contraste $\hat{y}_n$;

(1)	N	P	K	NP	NK	PK	NPK
-1	+1	-1	-1	+1	+1	-1	+1

$$\hat{y}_n = -7{,}54 + 9{,}91 - 8{,}63 - 11{,}75 + 11{,}19 + 10{,}92 - 8{,}99 + 11{,}98$$

$$\hat{y}_n = 7{,}09 \quad \rightarrow \quad \hat{N} = \frac{7{,}09}{16} \quad \rightarrow \quad \hat{N} = 0{,}443 \quad \text{t/ha,}$$

onde 16 é o número de parcelas em cada grupo: 16 parcelas sem N e 16 com N

$$SQN = \frac{(7{,}09)^2}{4 \times 8} \quad \rightarrow \quad SQN = 1{,}5709 \quad \text{com 1 grau de liberdade.}$$

Aí 4 é o número de repetições e 8 é a soma dos quadrados dos coeficientes referentes aos efeitos do nitrogênio.

Por analogia determinam-se as outras estimativas que estão resumidas na tabela a seguir.

Fonte de variação	GL	$\hat{y}$	Estimativa do efeito	S.Q.
Efeito de N	1	7,09	0,443	1,5709
" P	1	0,67	0,042	0,0140
" K	1	6,37	0,398	1,2680
" N x P	1	4,01	0,251	0,5025
" N x K	1	-2,77	-0,173	0,2398
" P x K	1	-4,07	-0,254	0,5177
" N x P x K	1	3,63	-0,227	0,4118

A análise de variância completa está na tabela seguinte:

Fonte de variação		GL	SQ	QM	F
Nitrogênio	N	1	1,5709	1,5709	9,43**
Fósforo	P	1	0,0140	0,0140	0,084
Potássio	K	1	1,2680	1,2680	7,62*
Interação	N x P	1	0,5025	0,5025	3,02
Interação	N x K	1	0,2398	0,2398	1,44
Interação	P x K	1	0,5177	0,5177	3,11
Interação	N x P x K	1	0,4118	0,4118	2,47
(Tratamentos)		(7)	(4,5246)		
Blocos		3	3,2011		
Resíduo		21	3,4955	0,1665	
Total		31	11,2212		

Observa-se que somente os efeitos de N e K foram significativos. Não houve reação a adubação fosfatada.

9.2.4 - Exemplo

Considere-se um fatorial 2^2, em blocos ao acaso, em 4 blocos. Os dados da tabela seguinte fornecem os resultados do experimento.

Trat. / Blocos	(1)	A	B	AB	Totais
I	18,0	20,6	19,6	19,2	77,4
II	8,6	21,0	15,0	19,6	64,2
III	9,4	18,6	14,6	18,4	61,0
IV	11,4	20,6	15,8	20,2	68,0
Totais	47,4	80,8	65,0	77,4	270,6

Procedimento idêntico ao caso anterior, dá o quadro de análise de variância a seguir.

Fonte de variação	GL	SQ	QM	F
Blocos	3	37,83		
Fator A	1	131,11	131,11	31,29**
Fator B	1	12,61	12,61	3,01
Interação (A x B)	1	27,57	27,57	6,58*
(Tratamentos)	(3)	(171,29)		
Resíduo	9	37,68	4,19	
Total	15	246,80		

A interação significativa indica que o comportamento de um fator depende dos níveis do outro. Neste caso a ação de B é influenciada pela presença ou ausência de A. Quando isso acontece a análise acima não é a mais indicada. Pode-se então adotar o sistema a seguir, onde serão estudados, todos os fatores, em relação aos outros.

	GL
Fator A -------------------	1
Fator B na ausência de A	1
Fator B na presença de A	1

Os cálculos podem ser feitos da forma exposta anteriormente, como a seguir:

	(-)	A	B	AB	$\hat{y}$	SQ
Fator A	-	+	-	+	45,8	131,10
Fator B na ausência de A	-	0	+	0	17,6	38,72
Fator B na presença de A	0	-	0	+	-3,4	1,45

A análise de variância é agora:

Fonte de variação	GL	SQ	QM	F
Blocos	3	37,83		
Fator A	1	131,10	131,10	31,29**
Fator B na ausência de A	1	38,72	38,72	9,24*
Fator B na presença de A	1	1,45	1,45	0,346
(Tratamentos)	(3)	(171,27)		
Resíduo	9	37,68	4,19	
Total	15	246,78		

Verifica-se pois que o solo reagiu a adubação com o fator B, mas apenas na ausência de A

9.2.5 - Exemplo

Os dados seguintes são de um experimento com 2 níveis dos fatores P, K, Mg, em viveiro de seringueira, com 3 repetições. Faça a respectiva análise de variância.

	(1)	P	K	Mg	PK	PMg	KMg	PKMg	TOTAIS
I	54,47	53,80	48,52	40,09	38,64	40,96	34,28	36,61	
II	45,32	45,41	41,05	41,25	36,31	35,46	35,09	46,78	
III	51,39	45,27	50,64	39,61	32,92	40,56	32,46	32,62	
Totais									

9.2.6 - Exemplo

Os dados abaixo, de um experimento fatorial, cujos fatores, adubação mineral completa (A) e matéria orgânica (M), foram instalados em 4 blocos.

	(1)	A	M	AM
I	0,430	3,041	3,400	6,815
II	0,725	0,326	5,760	5,784
III	0,038	0,981	5,960	2,940
IV	0,181	2,415	7,342	5,320

Efetue a análise adequada e interprete os resultados.

9.2.6 - Exemplo

Planeje um experimento com 2 macro nutrientes e dois níveis cada um, com 4 repetições. Faça o "layout" do delineamento.

9.3. Experimentos em parcelas divididas e subdivididas.

Os delineamentos em parcelas divididas ou subdivididas constituem mais uma forma de controle local. São bastante usados em experimentos fatoriais, experimentos em casa de vegetação, para estudar temperatura, em pesquisas com inseticidas ou fungicidas, em experimentos com animais, etc.

9.3.1 - Vantagens e desvantagens

As principais vantagens de um delineamento "sptit plot" são:

i) Possuir grandes unidades experimentais por necessidade ou pode ser utilizado para comparar tratamentos subsidiários.

ii) É possível testar um número maior de tratamentos (sub ou sub-sub)

As desvantagens são as seguintes

i) Todos os tratamentos são medidos com menor precisão que nos blocos ao acaso.

ii) Quando ocorre erro de dados (parcelas perdida), por exemplo, cresce a complexidade de análise.

9.2.2. - Modelo estatístico

Como nos experimentos fatoriais os modelos dependem do grau de subdivisão que se fizer. Uma observação de um delineamento "split plot" com p parcelas, r repetições, q subparcelas, em blocos ao acaso, será obtida por:

$$Y_{ijk} = \mu + r_i + \alpha_j + \delta_{ij} + \beta_k + (\alpha\beta)_{jk} + e_{ijk},$$

onde μ é o efeito da média da população, r_i é o efeito da i-ésima repetição, α_j é o efeito do j-ésimo nível do fator A, δ_{ij} é o componente aleatório (erro a), β_k é o efeito do k - ésimo nível do fator B, $(\alpha\beta)_{ij}$ é o efeito e a interação dos fatores A x B e e_{ijk} é o componente aleatório, associado à subparcela ou fator B.

Por processos análogos aos anteriores, estima-se os componentes da análise de variância, cujas fórmulas são semelhantes às já vistas antes, para cada fonte de variação, dando origem ao quadro de análise de variância que se segue:

Fontes de variação	Gl	SQ	QM	Desvio Padrão	F
Blocos	$(r-1)$				
Tratamentos A	$(a-1)$				
Resíduos (a)	$n_a = (a-1)(r-1)$			s_a	
(Parcelas)	$(ar-1)$				
Tratamentos B	$(b-1)$				
Interação A x B	$(a-1)(b-1)$				
Resíduos (b)	$n_b = a(b-1)(r-1)$			s_b	
Subparcelas	$abr-1$				

9.3.3 - Comparação de média dos tratamentos

1º caso: Tratamentos A (parcelas)

A comparação entre as médias dos tratamentos A (parcelas) é feita com o auxílio do desvio padrão s_a, com base no número de graus de liberdade do resíduo (a), (n_a).

Assim por exemplo, para se testar um contraste $\hat{Y} = \hat{A}_i - \hat{A}_u$, $\hat{A}_i$ e $\hat{A}_u$ são médias de tratamentos A, tem-se que:

$$s(\hat{Y}) = \sqrt{\frac{2}{br}}\; s_a \qquad \text{e} \qquad t = \frac{\hat{A}_i - \hat{A}_u}{s(\hat{Y})}$$

As diferenças mínimas significativas pelos testes de Tukey e Duncan são respectivamente estimadas por:

$$\Delta = q \times \frac{s_a}{\sqrt{br}} \qquad \text{e} \qquad D = z \times \frac{s_a}{\sqrt{br}}$$

2º caso: Tratamentos B (subparcelas)

A comparação entre médias dos tratamentos B (subparcelas) se faz com o auxílio do desvio padrão s_b, com base no número de graus de liberdade do resíduos (b), (n_b).

Assim para se testar um contraste $\hat{y} = \hat{B}_i - \hat{B}_u$, tem-se que:

$$s(\hat{y}) = \sqrt{\frac{2}{ar}}\, s_b \qquad \text{e} \qquad t = \frac{\hat{B}_i - \hat{B}_u}{s(\hat{y})}$$

E de forma análoga, as diferenças mínimas significativas, respectivamente pelos testes de Tukey e Duncan são estimadas por:

$$\Delta = q \times \frac{s_b}{\sqrt{ar}} \qquad \text{e} \qquad D = z \times \frac{s_b}{\sqrt{ar}},$$

onde q e z são encontrados nas tabelas respectivas, neste caso, em função do número de tratamentos b e do número de graus de liberdade do resíduo (b)

9.3.4 - Exemplo

Os dados da tabela seguinte são produções de adubo verde (kg de matéria verde por parcela) de um experimento com 8 tratamentos, em blocos ao acaso com 4 repetições, realizado em dois anos sucessivos, nas mesmas parcelas.

Produção obtida (kg de matéria verde)

Tratamentos	1º Bloco		2º Bloco		3º Bloco		4º Bloco	
	1º ano	2º ano	1º ano	2º ano	1º ano	2º ano	1º ano	2º ano
1. Mucuna preta	86,8	90,2	76,8	94,0	88,6	86,4	81,6	82,2
2. Feijão de porco	44,0	83,8	56,6	72,2	52,4	88,6	52,2	83,2
3. *Crot. juncea*	102,4	120,2	90,8	104,6	92,0	112,0	84,8	113,6
4. Guandu	68,4	91,0	55,2	78,8	49,0	83,4	61,2	91,2
5. *Teph. candida*	34,0	57,2	32,4	54,0	24,4	50,8	30,0	46,2
6. Soja	33,0	33,6	34,8	33,2	32,0	33,4	33,6	42,6
7. *Crot. grantiana*	25,8	77,0	21,6	62,4	19,2	63,6	21,0	63,4
8. Milho	138,8	110,2	106,4	80,0	108,0	92,0	81,8	90,6
	533,2	663,2	474,6	579,2	465,6	610,2	446,2	613,0

cálculos: $\Sigma x^2 = 686381,68$ $\Sigma x = 4385,2$

$$C = \frac{(4385,2)^2}{64} = 300468,42$$

SQTo = 686381,6 - 300,42 SQTo = 53348,86

$$SQParcelas = \frac{177,0^2 + 170,8^2 + \ldots\ldots + 172,4^2}{2} - C = 42722,42$$

$$SQB = \frac{1}{8}\left[(1196,4)^2 + (1053,8)^2 + \ldots\ldots + (1059,2)^2\right] - C = 851,43$$

$$SQT = \frac{1}{8}\left[(686,6)^2 + (533,0)^2 + \ldots\ldots + (807,8)^2\right] - C = 40189,61$$

SQR (a) = SQParcelas - SQB-SQT SQR(a) = 1681,38

$$SQAnos = \frac{1}{32}\left[(1919,6)^2 + (2465,6)^2\right] - C = 4658,07$$

$$SQ(A \times T) = \frac{1}{4}\left[(333,8)^2 + \ldots\ldots + (435,0)^2 + (352,8)^2 + \ldots\ldots + (372,8)^2\right] - C - SQA - SQT$$

SQ(A x T) = 49896,32 - 4658,07 - 40189, 61 = 5048,64

Análise de variância

Fonte de variação	G.L.	S.Q.	Q.M.	Desvio Padrão	F
Blocos	3	851,43			
Tratamentos (T)	7	40189,61	5741,37	75,77	71,70**
Resíduo (a)	21	1681,38	80,07	8,95	
(Parcelas)	(31)	(42722,42)			
Anos (A)	1	4658,07	4658,07	68,25	121,88**
Interação A x T	7	5048,64	721,23	26,86	18,87**
Resíduo (b)	24	919,73	38,22	6,81	
Total	63	53348,86			

** Diferença significativa a 0,01 de probabilidade

Como a interação A x T é significativa, indicando que os tratamentos têm comportamento diferente em cada ano, deve-se modificar esse esquema, estudando os efeitos dos anos, sobre cada tratamento, como se segue:

$$\text{SQ Anos dentro do Trat. } 1 = \frac{1}{4}\left[(333,8)^2 + (352,8)^2\right] - \frac{1}{8}(686,6)^2$$

SQ Anos dentro do Trat. 1 = 45,13

De maneira análoga estuda-se os outros casos. Assim se tem a análise de variância do quadro seguinte:

Fonte de variação	GL	SQ	QM	Desvio padrão	F
Blocos	3	851,43			
Tratamentos (T)	7	40189,61	5741,37	75,77	71,70
Resíduo (a)	21	1681,36	80,07	8,95	
(Parcelas)	(31)	(42722,42)			
Anos d. Trat. 1	1	45,13	45,13		1,18
Anos d. Trat. 2	1	1878,85	1878,85		49,03**
Anos d. Trat. 3	1	808,02	808,02		21,09**
Anos d. Trat. 4	1	1529,05	1529,05		39,90**
Anos d. Trat. 5	1	954,85	954,85		24,92**
Anos d. Trat. 6	1	11,05	11,05		0,28
Anos d. Trat. 7	1	3996,18	3996,18		104,28**
Anos d. Trat. 8	1	483,61	483,61		12,62**
Resíduo (b)	24	919,73	38,32		
Total	63				

9.3.5 - Exemplo

Faça a comparação das médias dos tratamentos (A) e (B) utilizando os testes de Tukey e Duncan. Compare os resultados.

9.3.6 - Exemplo

No quadro seguinte estão os dados de produção em t/ha, obtidos de um experimento em 6 blocos casualizados, com 3 clones (parcelas), cortados em 4 épocas diferentes (subparcelas). Faça a análise dos dados, comparando as médias dos tratamentos, por teste adequado e faça uma interpretação minuciosa dos resultados.

Clones	Épocas de corte	Blocos						
		I	II	III	IV	V	VI	TOTAIS
C_1	1	2,17	1,88	1,62	2,34	1,58	1,66	
	2	1,58	1,26	1,22	1,59	1,25	0,94	
	3	2,29	1,60	1,67	1,91	1,39	1,12	
	4	2,23	2,01	1,82	2,10	1,66	1,10	
C_2	1	2,33	2,01	1,70	1,78	1,42	1,35	
	2	1,33	1,30	1,85	1,09	1,13	1,06	
	3	1,86	1,70	1,81	1,54	1,67	0,88	
	4	2,27	1,81	2,01	1,40	1,31	1,06	
C_3	1	1,75	1,95	2,13	1,78	1,31	1,30	
	2	1,52	1,47	1,80	1,37	1,01	1,31	
	3	1,55	1,61	1,82	1,56	1,23	1,13	
	4	1,56	1,72	1,99	1,55	1,51	1,33	

9.4. Experimento de espaçamentos

Como em todo tipo de pesquisa, o delineamento experimental a ser utilizado depende dos objetivos da pesquisa, associados ao material experimental.

Se o objetivo é apenas comparar espaçamentos de determinada cultivar, o delineamento deve ser o mais simples possível, como o de blocos ao acaso, por exemplo. Se se pretende testar espaçamentos e cultivares, por exemplo, um arranjo fatorial ou em parcelas subdivididas, certamente darão ótimos resultados, pois permitem estudar a ação conjunta dos dois fatores (Espaçamentos e Cultivares) ou mais fatores, caso haja interesse. No caso da utilização de parcelas subdivididas é prudente que não hajam muitos fatores.

9.4.1- Exemplo

Os dados que se seguem são produções em t/parcela de um experimento em 3 blocos casualizados com arranjo fatorial de 4 cultivares e 3 espaçamentos.

	I			II			III		
	E_1	E_2	E_3	E_1	E_2	E_3	E_1	E_2	E_3
C_1	3,5	0,98	5,70	3,45	4,33	3,45	1,46	2,30	0,97
C_2	8,4	3,40	3,02	1,09	0,99	2,83	4,31	5,10	3,21
C_3	2,9	3,01	1,06	5,41	2,33	4,73	2,01	3,47	2,41
C_4	5,4	4,60	2,43	2,45	3,02	1,08	3,41	0,89	3,02

Os resultados da análise desse experimento fatorial 4 x 3 estão na tabela abaixo:

Fonte de variação	G	SQ	QM	F
Blocos	2	6,4497	3,2249	0,94
Cultivares	3	2,8586	0,9529	0,28
Espaçamentos	2	5,1526	2,5763	0,75
Interação Cul x Esp	6	5,0836	0,8473	0,25
Resíduo	22	75,7407	3,4428	
TOTAL	35	95,2853		

9.4.2 - Exemplo

Considere os dados de produção, a seguir, de um experimento em parcelas subdivididas com 4 cultivares e 2 espaçamentos, em blocos casualizados, com 3 repetições. Faça a análise conveniente e compare os resultados com o exemplo anterior fazendo um breve comentário.

Tratamentos	I		II		III	
	Esp.1	Esp.2	Esp.1	Esp.2	Esp.1	Esp.2
Cultivar 1	6,8	6,2	6,8	8,6	4,6	6,1
Cultivar 2	4,0	3,8	2,4	5,2	6,8	8,4
Cultivar 3	2,4	7,5	5,2	4,9	4,3	6,3
Cultivar 4	8,4	4,9	7,1	3,4	6,3	2,4

9.5. Experimentos de competição de cultivares

Os experimentos de competição de cultivares, não raro, pela própria natureza dessa pesquisa tem ao longo de seu curso parcelas perdidas. Por isso os delineamentos experimentais devem, além de simples, ter bastante flexibilidade, para nesses casos, ainda oferecerem algúma opção de análise de resultados, com certa confiabilidade.

Foi bastante frequente, principalmente nos estudos de melhoramento, o uso dos "Latices" (reticulados quadrados), sob a alegação de que esse tipo de pesquisa envolve um grande número de tratamentos. Esse delineamento contudo, tem sérias limitações que o tornam totalmente superados, mesmo porque já existem delineamentos mais modernos, mais simples e com muita flexibilidade, para pesquisas dessa natureza. Assim é que os delineamentos de blocos ao acaso com tratamentos comuns são apropriados para tal fim.

9.5.1- Experimentos em blocos ao acaso com alguns tratamentos comuns

Esse delineamento, além de simples oferece as vantagens de poder trabalhar com um número elevado de tratamentos e caso haja omissão de parcelas ou até de tratamentos ou blocos, ele ainda continua com boa validade. Exige apenas que já se conheça bem a performance de pelo menos um ou dois tratamentos que funcionam, neste caso, como controles ou "testemunhas".

9.5.2- Exemplo

Suponha que se queira colocar em competição 40 clones de guaraná, por exemplo, e disponha de 3 clones, cujo comportamento é já, bem conhecido. Forma-se então 4 grupos com 10 clones mais os 3 conhecidos. Os novos clones são chamados de tratamentos regulares e os três conhecidos, de tratamentos comuns. Assim tem-se:

Grupo 1: 1, 2, ..., 10, A, B, C
Grupo 2: 11, 12, ..., 20, A, B, C
Grupo 3: 21, 22, ..., 30, A, B, C
Grupo 4: 31, 32, ..., 40, A, B, C

Cada grupo corresponde a um experimento em blocos casualizados, que se se tiver 3 repetições, por exemplo, dá, para cada grupo o esquema da análise de variância a seguir:

Fonte de variação	GL
Blocos	2
Tratamentos	12
Resíduo	24
TOTAL	38

Por fim o conjunto dos 4 grupos (4 experimentos em blocos ao acaso) poderão ser analisados conjuntamente, dando origem ao quadro a seguir:

Análise conjunta dos 4 grupos:

Fonte de variação	GL
Experimentos	3
Blocos dentro de experimentos	8
Tratamentos ajustados	42
Interação Trat. Comuns x Exp.	6
Resíduo	96
TOTAL	135

Genericamente para g grupos, com z tratamentos regulares e c comuns, com r repetições, o quadro de análise será:

Fonte de variação	GL
Experimentos	g-1
Blocos de experimentos	g (r-1)
Tratamentos (ajustados)	gz + c-1
Interação (trat.comuns x exp.)	(c-1) (g-1)
Resíduo	g(r-1) (z+c-1)
TOTAL	(z+c) gr-1

9.5.3- Exemplo

Considere-se, a título de exercício, 8 clones, colocados em competição em dois grupos de 4 clones, juntamente com dois outros clones de comportamento bem conhecidos. Os dados de cada grupo, em 3 repetições encontram-se abaixo:

Grupo 1 (Experimento 1)

Blocos \ Trat.	Clone 1	Clone 2	Clone 3	Clone 4	A	B	TOTAIS
I	32	26	35	32	32	29	186
II	28	35	30	26	33	30	182
III	27	30	20	23	28	27	155
TOTAIS	87	91	85	81	93	86	523

Grupo 2 (Experimento 2)

Blocos \ Trat.	Clone 5	Clone 6	Clone 7	Clone 8	A	B	TOTAIS
I	24	26	32	26	25	32	165
II	23	20	31	31	28	33	166
III	16	25	32	20	23	20	136
TOTAIS	63	71	95	77	76	85	467

Os cálculos das somas de quadrados, para cada experimento são feitas da forma usual, cujos quadro de análise estão a seguir:

Experimento 1

Fonte de variação	GL	SQ	QN	F
Blocos	2	94,7777	47,3889	
Tratamantos	5	30,9444	6,1889	0,451
Resíduo	10	137,2223	13,7222	
TOTAL	17	262,9444		

Experimento 2

Fonte de variação	GL	SQ		F
Blocos	2	96,7777	48,3886	
Tratamentos	5	205,6111	41,1222	2,93
Resíduo	10	140,5556	14,0556	
TOTAL	17	442,9444		

Análise conjunta

$$C = \frac{(\Sigma x + \Sigma y)^2}{g \times r \times (z+c)} \qquad C = \frac{(523 + 467)^2}{36} \qquad C = 27.225$$

SQTo = $\Sigma x^2 - C$ (Σx^2 dos dois experimentos).

Exper. \ T.Comuns	A	B	TOTAL
E_1 (r = 3)	93	86	179
E^2 (r = 3)	76	85	161
TOTAL	169	171	340

$$SQTc = \frac{169^2 + 171^2}{6} - \frac{(340)^2}{12} \qquad SQTc = 0,3334$$

$$SQEx = \frac{179^2 + 161^2}{6} - \frac{(340)^2}{12} \qquad SQEx = 87,1111$$

$$SQ\ (Ex \times Tc) = \frac{93^2 + 86^2 + 76^2 + 85^2}{r = 3} - \frac{(340)^2}{12} - SQEx - SQTc = 21,3333$$

$$SQB\ d.\ Ex = SQB_1 + SQB_2 = 94,7777 + 96,7777 = 191,5554$$

$$SQR = SQR_1 + SQR_2 \qquad SQR = 277,7779$$

$$SQT\ (ajustados) = SQTc - SQB\ d.Ex - SQ\ (Ex \times Tc) - SQR = 187,8889$$

Análise conjunta

Fonte de variação	GL	SQ	QN	F
Experimentos	1	87,1111	87,1111	6,27 *
Blocos d. Experimentos	4	191,5554	47,8889	3,45 *
Tratamentos (ajustados)	9	215,2223	23,9136	1,72
Interação (Trat. comum x Exp.)	1	21,3333	21,3333	1,54
Resíduo	20	277,7779	13,8889	
TOTAL	35	793,000		

9.5.4. Exemplo

Os dados a seguir são de um grupo de experimentos de competição de variedades de cana de açúcar. Com relação a eles efetue a análise adequada. (pg. seguinte)

9.6. Experimentos hierárquicos

Esse tipo de experimento é bastante utilizado em pesquisas sobre subamostras de amostras, ou em subclasses dentro de classes.

Dados do exemplo 9.5.4.

Experimento 1

TRATAMENTOS	BLOCOS I	II	III
CP 34-120	46,6	45,7	49,6
NA 56-30	61,1	72,1	51,3
NA 56-79	74,2	74,3	100,7
Tuc 5619	50,2	37,3	66,1
CP 44-101	40,3	38,7	35,1
NA 56-35	37,5	40,1	36,9
NA 56-68	69,6	61,0	59,3

Experimento 2

TRATAMENTOS	I	II	III
CP 34-120	55,2	51,2	48,0
NA 56-30	62,7	83,0	75,1
NA 56-79	63,4	81,7	110,3
Tuc 5619	62,0	69,1	77,9
NA 56-62	71,0	72,9	81,6
NA 59-17	59,5	83,0	82,9
NA 59-44	73,2	98,6	81,0
NA 59-60	47,6	37,7	42,3

9.6.1- Exemplo

Numa prospecção de dendezeiro no Estado do Amazonas, o Estado foi dividido em municípios e dentro desses municípios foram escolhidas regiões, em cujas regiões foram coletados materiais nativos. Esta pesquisa caracteriza um experimento hierárquico ou uma classificação hierárquica.

9.6.2- Modelo estatístico

Num experimento hierárquico com p unidades, das quais são escolhidas aleatoriamente f subamostras, cada observação é dada por:

$$y_{ijk} = \mu + p_i + f_{ij} + e_{ijk}$$

onde m é a média geral, p_i é o efeito da pésima unidade e f_{ij} é o efeito da j-ésima subamostra dentro da unidade i e e_{ijk} é o efeito aleatório.

9.6.3- Exemplo

Numa prospecção de dendê em 4 locais, foram coletadas, ao acaso 10 plantas em cada local. Os pesos dos cachos colhidos, estão na tabela a seguir

Locais / Plantas	L_1	L_2	L_3	L_4
1	3	10	12	6
2	5	4	1	2
3	8	3	6	9
4	9	9	7	3
5	4	5	2	5
6	2	7	5	7
7	7	2	4	10
8	1	11	3	4
9	12	6	8	6
10	6	1	6	12
TOTAL	57	58	54	64

$\Sigma x^2 = 1755$ $\Sigma x = 233$ $C = 1357{,}2250$

$SQTo = 397{,}7750$

$$SQLoc = \frac{57^2 + 58^2 + \ldots\ldots + 64^2}{10} - C \qquad SQLoc = 5{,}2750$$

$SQPl\ d.\ loc = SQTo - SQLoc$ $\qquad SQPl\ d.\ Loc = 392{,}5000$

Quadro de análise

Fonte de variação	GL	SQ	QM	E(QM)
Locais	$\ell - 1 = 3$	5,2750	1,7583	$\sigma^2 + 10\sigma^2$
Plantas d. locais	$\ell(p-1) = 36$	392,5000	10,9027	σ^2
TOTAL	$\ell p - 1 = 39$	397,7750		

9.6.4- Exemplo

Na determinação do teor de cálcio de uma forrageira foram selecionadas ao acaso 4 plantas e de cada planta foram tiradas 3 folhas e em cada folha foram feitas 2 análises químicas, cujos resultados encontram-se a seguir:

Faça a respectiva análise, utilizando esses dados.

Plantas	Folhas	Análises químicas		Totais
1	1	3,28	3,09	
	2	3,52	3,48	
	3	3,88	2,80	
2	1	2,46	2,44	
	2	1,87	1,92	
	3	2,19	2,19	
3	1	2,77	2,66	
	2	3,74	3,44	
	3	2,55	2,55	
4	1	3,78	3,87	
	2	4,07	4,12	
	3	3,31	3,31	

10. Regressão e Correlação

10.1. Regressão linear, de 2º grau e múltipla.

A análise de variância, tal como é feita pressupõe independência dos diversos tratamentos utilizados. Quando esta hipótese não se verifica, a análise da variância deve refletir a dependência entre os tratamentos. Isso normalmente acontece em caso de tratamentos de fertilizantes ou inseticidas, quando se usa níveis crescentes, por exemplo.

10.1.1 - Exemplo

Considere-se, por exemplo, os dados de 5 tratamentos e 5 repetições, um experimento inteiramente casualizado. Os valores de X são os níveis de adubo usados e os de T, os totais de produção, de cada tratamento, nas 4 repetições, como na tabela a seguir:

X	0	1	2	3	4
T	2,0	2,0	3,0	4,0	5,0

Uma análise de variância, do modo usual, sem considerar a regressão deu.

FONTE DE VARIAÇÃO	GL	SQ	QM	F
Tratamentos	4	1,70	0,425	3,04
Resíduo	15	2,10	0,140	
T O T A L	19	3,80		

O valor de F não é significatico, pelo fato de não se ter levado em conta a regressão.

10.1.2 - Exemplo

Considere-se os dados a seguir, onde X são níveis de adubo e T os totais de produção.

X	0	1	2	3	4
T	3,0	5,0	2,0	4,0	2,0

Admitamos que se trata de uma regressão linear: $Y = a + bX$. Neste caso as estimativas dos parâmetros b e a serão:

$$\hat{b} = \frac{\Sigma XY - \frac{(\Sigma X)(\Sigma Y)}{N}}{\Sigma X^2 - \frac{(\Sigma X)^2}{N}} \quad \text{e} \quad a = \bar{Y} - \hat{b}\,\bar{X}$$

$\Sigma XY = 40$ $\Sigma X = (4)\ (10) = 40$ $\Sigma Y = 16$ $\Sigma Y^2 = 120$

Por substituição $\hat{b} = 0,20$

$$\bar{x} = \frac{40}{20} = 2,0 \ , \quad \bar{y} = \frac{16}{20} = 0,8 \ , \quad \hat{a} = 0,8 - (0,20)2 = 0,4$$

logo $\hat{Y} = 0,4 + 0,20\ X$

$$SQ\hat{b} = \frac{\left[\Sigma XY - \frac{(\Sigma X)\ (\Sigma Y)}{N}\right]^2}{\Sigma X^2 - \frac{(\Sigma X)^2}{N}} = \frac{(8,0)^2}{40} = 1,60$$

$SQD = 1,70 - 1,60 = 0,10$

Tem-se portanto a seguinte análise de variância:

FONTE DE VARIAÇÃO	GL	SQ	QM	F
Regressão linear	1	1,60	1,600	11,42**
Desvio de regressão	3	0,10	0,033	
(Tratamentos)	(4)	(1,70)		
Resíduo	15	2,10		
T O T A L	19	3,80		

Como se vê é significativa a regressão linear ao mesmo tempo que os desvios de regressão não o são.

Identicamente se poderia fazer para regressões de graus superiores ao primeiro.

10.1.3 - Exemplo

Para se testar a linearidade de uma regressão, tome-se um exemplo com mais de uma observação de Y para cada X. Faça-se este teste usando a análise de variância. Considere-se os dados a seguir:

X	1	2	3	4
Y	8	20	31	29
	10	23	28	28

$$SQTo = 8^2 + 10^2 + \dots + 28^2 - \frac{(8 + 10 + \dots 28)^2}{2} = 546,9$$

$$SQ\ Entre\ grupos = \frac{(8 + 10)^2 + \dots + (29 + 28)^2}{2} = 535,4$$

$$SQ\ d.\ grupos = 546,9 - 535,4 = 11,5$$

SQ entre grupos é dividida em duas partes

SQ Reg. Linear + SQ Desv. de regressão

$$SQ\ Reg.\ Linear = \frac{\left[509 - \frac{(20)\ (177)}{8} \right]^2}{60 - \frac{20^2}{8}} = 442,2$$

SQ Desv. Reg. = 535,4 - 442,2 = 93,2

Análise de Variância

FONTE DE VARIAÇÃO	GL	SQ	QM	F
Entre grupos:	3	535,4	178,47	
Regressão linear	1	442,2	442,2	
Desvio de regressão	2	93,2	46,6	$\frac{46,6}{2,87} = 16,2^*$
Dentro de Grupos (resíduo)	4	11,5	2,87	
T O T A L	7	546,9		

11. Exercícios complementares

11.1- Deseja-se fazer um experimento para testar 5 níveis de um fungicida. Para isso coletou-se o material infectado e o experimento será conduzido em laboratório.

Planeje esse experimento, desde a descrição dos objetivos, faça o "layout", desenvolva o plano de tal sorte que a sua condução atinja os objetivos.

11.2- Considere os dados de nº de folhas de guaranazeiro, do quadro abaixo e faça a análise conveniente.

Tratam. / Bloco	375	392	68	272	263	329	501	505	503
I	9,60	8,00	5,60	7,00	9,00	6,80	7,20	11,60	7,30
II	5,80	7,80	6,20	7,30	6,70	6,80	9,20	6,60	8,40
III	5,70	6,20	7,40	9,00	10,20	5,40	11,50	7,60	6,50

REFERÊNCIAS

ok FEDERER, W.T. Experimental design. 1955. Macmillan Nova York.

COCHRAN, W.G. & COX, G.M. Experimental design. 1957. J. Wiley.

ANDERSON, R.L. & BANCROFT, T.A. Statistical Threory in Research. 1952. McGraw-Hill, Nova York.

ok LITTLE, T.M. & HILLS, F.J. Agricultural experimentation - Design and Analysis, 1978. J. Wiley.

ok PIMENTEL GOMES, F. Curso de Estatística Experimental, 11 edição, 1985. Livraria Nobel, S.P.

-X-X-X-