Universidade Federal de Mato Grosso

ENGENHARIA FLORESTAL UFMT

UNIDADE IV – DELINEAMENTO INTEIRAMENTE CASUALIZADO (DIC)

PROF.: RÔMULO MÔRA
romulomora.webnode.com

CUIABÁ, MT
2015/2

1. INTRODUÇÃO

Este delineamento apresenta como característica principal a necessidade de homogeneidade do ambiente para todas as unidades experimentais.

É o tipo de delineamento mais simples que existe. A distribuição dos tratamentos às unidades experimentais é feita completamente ao acaso, ou seja, não é feita nenhuma restrição na casualização.

Este é o delineamento básico, os demais se originam dele pela imposição de restrições (controle local). Envolve dois princípios básicos da experimentação: repetição e casualização.

Vantagens:

- ✓ Delineamento bastante flexível;
- ✓ Pode-se usar qualquer número de tratamentos ou repetições;
- ✓ Apresenta maior número de graus de liberdade associado ao resíduo;
- ✓ A análise estatística é simples.

Desvantagens:

- ✓ Exige homogeneidade total das condições ambientais;
- ✓ Pode conduzir a uma estimativa de variância residual alta, uma vez que abre mão do uso do principio do controle local.

2. MODELO ESTATÍSTICO

Cada delineamento possui um modelo estatístico que identifica os fatores que estão influenciado a variável em estudo.

Para o DIC tem-se o seguinte modelo:

$$Y_{ij} = m + t_i + e_{ij} \qquad i = 1, 2, ..., I \qquad j = 1, 2, ..., J$$

em que:
Y_{ij} = é o valor observado para a variável em estudo referente ao i-ésimo tratamento na j-ésima repetição;
m = média de todas as unidades experimentais para a variável em estudo;
t_i = é o efeito do particular tratamento i no valor observado Y_{ij};
e_{ij} = é o erro associado a observação Y_{ij} ou seja é o efeito dos fatores não controlados na parcela;

O erro se deve ao fato de não ser possível controlar todas as condições experimentais. O erro experimental refere-se às variações observadas entre as repetições do mesmo tratamento.

3. ESQUEMA DE CASUALIZAÇÃO DOS TRATAMENTOS

Seja um experimento com 5 tratamentos (A, B, C, D e E) e 4 repetições (20 parcelas ou unidades experimentais)

A	B	D	E
B	A	C	D
E	A	B	C
D	E	A	B
C	E	D	C

4. QUADRO DE TABULAÇÃO DE DADOS

Considere um experimento instalado no DIC com 5 tratamentos e 4 repetições. A coleta de dados da pesquisa pode ser resumida, num quadro do tipo a seguir:

	Tratamento				
Repetições	**1**	**2**	**3**	**4**	**5**
1	Y_{11}	Y_{21}	Y_{31}	Y_{41}	Y_{51}
2	Y_{12}	Y_{22}	Y_{32}	Y_{42}	Y_{52}
3	Y_{13}	Y_{23}	Y_{33}	Y_{43}	Y_{53}
4	Y_{14}	Y_{24}	Y_{34}	Y_{44}	Y_{54}
Totais	T_1	T_2	T_3	T_4	T_5

Deste quando pode-se retirar algumas informações de interesse:

- Número de unidades experimentais: N = I x J ou N = número de unidades experimentais

-Total Geral: $G = \sum_{\substack{i=1 \\ j=1}}^{I,J} Y_{ij} = \sum_{i=1}^{I} T_i$

- Total para o tratamento i: $T_i = \sum_{j=1}^{J} Y_{ij}$

- Média para o tratamento i: $\hat{m}_i = \dfrac{T_i}{J}$

- Média geral do experimento: $\hat{m} = \dfrac{G}{I.J}$

5. ANÁLISE DE VARIÂNCIA

A análise de variância foi introduzida por Fisher e é essencialmente um processo baseado na decomposição da variação total existente entre uma série de observações, ou seja, a variação existente entre todas as observações, em partes, na variação devido à diferença entre os efeitos dos tratamentos e na variação devido ao acaso, que também é denominada de erro experimental ou resíduo.

No entanto, para que esta técnica seja empregada é necessário que sejam satisfeitas as seguintes pressuposições:

1º) Os efeitos do modelo estatístico devem ser aditivos

Nos experimentos, os vários efeitos devem ser aditivos, tanto é que para cada delineamento estatístico existe um modelo matemático denominado modelo linear aditivo. Para o delineamento inteiramente casualizado, este modelo é:

$$Y_{ij} = m + t_i + e_{ij}$$

onde expressa que o valor de qualquer unidade experimental é resultante de uma média geral, mais o efeito de tratamento e o mais o efeito do erro experimental.

2º) Os erros experimentais devem ser independentes

Cada observação possui um erro que deve ser independente dos demais. O princípio da casualização assegura a validade da estimativa do erro experimental, pois permite uma distribuição independente do mesmo. A casualização evita que todas as parcelas que recebem o mesmo tratamento sejam favorecidas ou desfavorecidas entre as parcelas experimentais

3º) Os erros experimentais devem ser normalmente distribuídos

A única fonte de variação de amostragem são os erros aleatórios. Estes devem ter distribuição normal (ou aproximadamente normal) com média igual a zero e variância igual s^2. Felizmente, as variações da suposição de normalidade não afetam muito seriamente a validade da análise de variância.

4º) As variâncias das diferentes amostras devem ser homogêneas

Na análise de variância, o valor do Quadrado Médio do Resíduo, que corresponde à estimativa da análise de variância do erro experimental, é utilizada nas fórmulas matemáticas dos testes de hipóteses.

Tais testes são utilizados para verificar se existe ou não diferença significativa entre os tratamentos avaliados.

O Quadrado Médio do Resíduo nada mais é que a média das diferentes variâncias de cada tratamento (amostras). Assim sendo, é importante que as variâncias das diferentes amostras seja homogêneas, de modo que os resultados obtidos dos testes de hipóteses tenham validade.

5.1 TESTE F MÁXIMO - HARTLEY

Uma das exigências do modelo estatístico e, portanto, da validade da análise de variância, é que as variâncias das diferentes amostras devem ser homogêneas. Entre os vários testes estatísticos utilizados para verificar a homogeneidade de variâncias, temos o teste F-máximo proposto por Hartley.

O teste F - máximo: é simples e rápido, porém apresenta menor precisão quando as amostras tem graus de liberdade diferentes.

$$F - Máximo = \frac{s^2 máxima}{s^2 mínima}$$

onde:
s² máxima: maior valor das estimativas das variâncias entre as amostras;
s² mínima: menor valor das estimativas das variâncias entre as amostras;

O valor calculado de F-máximo é confrontado com o valor de F-máximo tabelado, com K (número de estimativas das variâncias das diferentes amostras) e (g – 1) graus de liberdade associados a cada estimativa de variância, sendo N número de observação de cada amostra.

Assim, temos:

- ✓ Se F-máximo calculado ≥ F-máximo tabelado – as estimativas das variâncias são estatisticamente diferentes ao nível α% de probabilidade, isto é, não há homogeneidade de variâncias.
- ✓ Se F-máximo calculado < F-máximo tabelado – as estimativas das variâncias não diferem estatisticamente entre si, ao nível α% de probabilidade, isto é, as variâncias são homogêneas.

Obs.: Para efeito de cálculo manual, quando os graus de liberdade para cada amostra forem diferentes, toma-se a média aritmética dos mesmos.

5.2 TESTE DE LILLIEFORS PARA NORMALIDADE

Uma das exigências do modelo estatístico e, portanto, da validade da análise de variância, é que os erros e_{ij} tenham distribuição normal.

Esta verificação dessa exigência pode ser feita pelo teste de Lilliefors. Este teste é uma adaptação do teste de Kolmogorov-Smirnov, sendo somente usado para a verificação da normalidade a partir das estimativas da média e desvio padrão.

Após calculadas todas as diferenças absolutas entre $F(Z_i)$ e $S(Z_i)$ e entre $F(Z_i)$ e $S(Z_i - 1)$, para i = 1, 2, ..., n, identifica-se a diferença absoluta máxima dada por:

$$D_{cal} = máximo\left\{\left|F(Z_i) - S(Z_i)\right|, \left|F(Z_i) - S(Z_{i-1})\right|\right\}$$

Se $D_{cal} \geq D_{tab}$ (α, n) rejeita-se H_0 ao nível de α% de probabilidade, ou seja, os erros não seguem distribuição normal
Se $D_{cal} < D_{tab}$ (α, n) não rejeita-se H_0 ao nível de α% de probabilidade, ou seja, os erros seguem distribuição normal.

Sendo $F(Z_i)$ as probabilidades da variável normal reduzida, calculadas da seguinte forma:

$$Z_i = \frac{X_i - \hat{m}}{s}$$

em que: X_i – são os erros e_{ij}
$\hat{m}$ – é a estimativa da média do e_{ij} estimados, portanto, é igual a zero
s – é a estimativa do desvio padrão dos e_{ij}.

Dessa forma:

$$Z_i = \frac{e_{ij}}{s} = d_{ij} =$$ desvios padronizados;

- Para o cálculo de $S(Z_i)$

$$S(Z_i) = \frac{K}{n},$$ onde K é o número de observações ≤ X_i, em nosso caso, é o número de desvios ≤ e_{ij}.

Devemos, inicialmente, obter os erros e_{ij}.
Como $y_{ij} = m + t_i + e_{ij}$, temos que: $e_{ij} = y_{ij} - m - t_i$
Mas, não conhecemos m e t_i, logo, devemos trabalhar com suas estimativas:

$$\hat{e}_{ij} = y_{ij} - \hat{m} - \hat{t}_i$$

Temo que: $\hat{m}_i = \hat{m} + \hat{t}_i$ logo $\hat{t}_i = \hat{m}_i - \hat{m}$

EXEMPLO 1

Para comparar a produtividade (kg) de quatro variedades de pequi, um engenheiro florestal tomou vinte parcelas similares e distribuiu, inteiramente ao acaso, cada uma das 4 variedades em 5 parcelas experimentais. A partir dos dados experimentais fornecidos abaixo, é possível concluir que existe diferença significativa entre as variedades com relação a produtividade, utilizando o nível de significância de 5%? (Realizar o teste de Homogeneidade de variância e o teste de normalidade).

	A	B	C	D
	19	32	22	33
	18	29	26	29
	17	23	28	34
	21	27	25	28
	21	25	29	27
Totais	96	136	130	151

Observação: Se os erros não seguirem distribuição normal, realiza-se a transformação dos erros afim de assegurar a distribuição normal dos dados.

6. ANÁLISE DE VARIÂNCIA PARA O DIC

No Delineamento inteiramente casualizado (DIC), a variação total, ou seja, a variação existente entre todas as observações, é decomposta apenas na variação devido à diferença entre os efeitos dos tratamentos e na variação devido ao acaso ou resíduo.

SQTotal = SQTrat + SQRes

O quadro da análise de variância, geralmente denotada por ANOVA (**AN**alysis **O**f **VA**riance) para a análise de um experimento instalado segundo o DIC, com igual número de repetições para todos os tratamentos é do seguinte tipo:

FV	GL	SQ	QM	F_{cal}	F_{tab}
Tratamento	(I-1)	SQTrat	SQTrat/(I-1)	QMTrat/QMRes	[(I-1);I(J-1)]
Resíduo	I(J-1)	SQRes	SQRes/I(J-1)		
Total	IJ-1	SQTotal			

Sendo: FV = Fonte de Variação; GL = Graus de Liberdade; SQ = Soma de Quadrados; QM = Quadrado Médio; F_{cal} = Valor de F calculado; F_{tab} = Valor de F tabelado.

Com as seguintes fórmulas definidas:
- Soma de Quadrado do Tratamento

$$SQTrat = \sum_{i=1}^{I} \frac{T_i^2}{J} - \frac{\left(\sum_{i=1,j=1}^{I,J} Y_{ij}\right)^2}{I.J}$$

Tratamentos com mesmo número de repetições

$$SQTrat = \sum_{i=1}^{I} \frac{T_i^2}{r_i} - \frac{\left(\sum_{i=1,j=1}^{I,r_i} Y_{ij}\right)^2}{N}$$

Tratamentos com números de repetições diferentes

Em que: T_i = Totais dos tratamentos; J = número de repetições (tratamentos balanceados); r_i = número de unidades experimentais do tratamento i; Y_{ij} = valor observado no tratamento i e repetição j; e N = número de unidades experimentais = $\sum_{i=1}^{I} r_i$

Com as seguintes fórmulas definidas:
- Soma de Quadrado Total

$$SQTotal = \sum_{i=1,j=1}^{I,J} Y_{ij}^2 - \frac{\left(\sum_{i=1,j=1}^{I,J} Y_{ij}\right)^2}{I.J}$$

Tratamentos com mesmo número de repetições

$$SQTotal = \sum_{i=1,j=1}^{I,r_i} Y_{ij}^2 - \frac{\left(\sum_{i=1,j=1}^{I,r_i} Y_{ij}\right)^2}{N}$$

Tratamentos com números de repetições diferentes

-Soma de Quadrado do Resíduo

SQRes = SQTotal-SQTrat

- Graus de Liberdade do Tratamento

GLTrat = (I-1), sendo I = número de tratamentos

- Graus de Liberdade do Total

GLTotal = I.J-1, sendo J número de repetições (Tratamentos com mesmo número de repetições)

GLTotal = n – 1 (Tratamentos com número de repetições diferentes)

- Graus de Liberdade do Resíduo

GLRes = I(J-1), sendo J número de repetições (Tratamentos com mesmo número de repetições) ou GLRes = GLTotal - GLTrat

-Quadro Médio do Tratamento

$$QMTrat = \frac{SQTrat}{(I-1)} = \frac{SQTrat}{GLTrat}$$

-Quadrado Médio do Resíduo

$$QM\,\mathrm{Re}s = \frac{SQ\,\mathrm{Re}s}{I(J-1)} = \frac{SQ\,\mathrm{Re}s}{GL\,\mathrm{Re}s}$$

-F calculado

$$F_{cal} = \frac{QMTrat}{QM\,\mathrm{Re}s}$$

- F tabelado

$$F_{tab} = F_{\alpha\%}\left[(I-1); I(J-1)\right]$$

$$F_{tab} = F_{\alpha\%}\left[GLTrat; GL\,\mathrm{Re}s\right]$$

As hipóteses para o teste F da análise de variância para tratamentos são as seguintes:

H_0 : $m_1 = m_2 = ... = m_i = m$, o que equivale a dizer que todos os possíveis contrastes entre as médias dos tratamentos, são estatisticamente nulos, ao nível de probabilidade que foi executado o teste.

H_a : não H_0 , o que equivale a dizer que existe pelo menos um contraste entre as médias dos tratamentos, é estatisticamente diferentes de zero, ao nível de probabilidade que foi realizado o teste.

A regra decisória para o teste F é a seguinte:

- $F_{cal} \geq F_{tab,}$ rejeita-se H_0 ao nível de α% de probabilidade, ou seja, existe pelo menos um contraste entre as médias dos tratamentos estatisticamente diferente de zero.

- $F_{cal} < F_{tab,}$ não rejeita-se H_0 ao nível de α% de probabilidade, ou seja, todos os possíveis contrastes entre as médias dos tratamentos são estatisticamente nulos.

EXEMPLO 2

O resultado das vendas efetuadas por 3 vendedores de uma indústria de pesticidas durante certo período é dado a seguir. Ao nível de 5% de probabilidade e considerando os vendedores como tratamentos de um DIC, verifique se há diferença de eficiência entre os vendedores. (Utilizar α = 5%; considerar variâncias homogêneas e dados seguindo distribuição normal)

	VENDEDORES		
	A	B	C
	29	27	30
	27	27	30
	31	30	31
	29	28	27
	32		29
	30		
Totais	178	112	147

FIM