

UNIVERSIDADE ABERTA

MESTRADO EM ESTATÍSTICA, MATEMÁTICA E COMPUTAÇÃO

(Ramo – Estatística Computacional)

**Técnicas de Experimentação Estatística:
Planos Factoriais e aplicações**

Aluno: **Roberval Monteiro Bezerra de Lima**

Orientadora: **Professora Doutora Teresa Paula Costa Azinheira Oliveira**

Lisboa, 2010

AGRADECIMENTOS

Agradeço em primeiro lugar a Deus por ter me dado força e discernimento durante a realização deste Curso.

Meu sincero agradecimento a minha orientadora, professora doutora Teresa Oliveira pelas suas palavras de incentivo e pela transmissão de seus conhecimentos, contribuindo para minha formação profissional.

Sou grato ao Sr. Magaldi de Maués, AM, pela cessão da área, implantação e manutenção do experimento .

Aos demais professores e professoras do Curso de Matemática, Estatística e Computação da Universidade Aberta de Portugal, expresso minha gratidão pelas vossas paciências e profissionalismo na condução das disciplinas.

ÍNDICE

AGRADECIMENTOS	ii
ÍNDICE	iii
RESUMO	vi
SUMMARY	vii
CAPÍTULO 1	1
1. Introdução	1
CAPÍTULO 2	2
2. Revisão de literatura: Abordagem histórica do Planejamento de Experiências	2
CAPÍTULO 3	3
Princípios e bases da Experimentação Estatística	3
3.1. Análise de variância – pressupostos fundamentais	7
CAPÍTULO 4	8
4.1. Planejamento Completamente Casualizado	8
4.2. Planos de Blocos Completos Casualizados	11
4.3. Quadrado Latino: Modelo em agrupamentos duplos	15
4.3.1. Definição de quadrado latino	15
4.3.2. Análise de variância do Quadrado Latino	17
CAPÍTULO 5	19
Modelos de Planejamento de Experiências com Blocos	19
5.1. Planos em Blocos Incompletos	19
5.2. Planos em Blocos Incompletos Equilibrados	21
5.3. Planos em Blocos Incompletos Parcialmente Equilibrados (PBIBD- Partially Balanced Incomplete Block Designs)	24
5.3.2. PBIBD: Definições	27
CAPÍTULO 6	29
6.1. Introdução aos Planos Factoriais	29
6.2. Planos Factoriais de Base 2	31
6.4. Planos Factoriais Fraccionados	33
6.5. Blocos e Técnica de confundimento no planejamento Fatorial	38
CAPÍTULO 7	41
Técnicas de Experimentação Estatística: uma aplicação a ensaios florestais na Amazônia brasileira	41
7.1. Introdução	41
7.2. Descrição do problema	42
7.3. Material e Métodos	43
7.4. Resultados e Discussão	48
7.4.1. Estudo da altura total	48
7.4.1.1. Análise do ensaio fatorial pelo método da Regressão Múltipla	57
7.4.3. Conclusões	64
CAPÍTULO 8	65
Considerações e conclusões gerais	65
Perspectivas de investigação futura	66
Referências Bibliográficas	67
ANEXOS	

ÍNDICE DE FIGURAS

Figura 1. Circularidade do método científico	3
Figura 2. Localização do experimento no município de Maués-AM, com área de 1,5 hectares.	43
Figura 3. Detalhe da parcela experimental	45
Figura 4. Distribuição das parcelas/tratamentos no campo.	46
Figura 5. Gráfico dos resíduos versus os fatores N, P e K.	55
Figura 6. Gráficos dos Resíduos versus preditos; Resíduos padronizados e da probabilidade Normal dos resíduos	56
Figura 7. Crescimento em altura do pau-rosa, aos 12 meses de idade, em função dos níveis de potássio (0=0 kg/ha; 1=50 kg/ha; 2=100 kg/ha) e Fósforo (0=0 kg/ha; 1=80 kg/ha; 2=160 kg/ha). Manaus, 2010.	59
Figura 8. Crescimento em altura do pau-rosa, aos 12 meses de idade, em função dos níveis de potássio (0=0 kg/ha; 1=50kg/ha; 2=100kg/ha) e Fósforo (0=0kg/ha; 1=80kg/ha; 2=160 kg/ha). Manaus, 2010.	60
Figura 9. Crescimento em altura do pau-rosa, aos 12 meses de idade, em função dos níveis de Fósforo (0=0 kg/ha; 1=80 kg/ha; 2=160 kg/ha) e Nitrogênio (0=0 kg/ha; 1=60 kg/ha; 2=120 kg/ha). Manaus, 2010.	61
Figura 10. Gráficos da análise de resíduos para a variável diâmetro de colo.	63
Figura 10. Pau-rosa (<i>Aniba rosaedora</i> Ducke), aos 12 meses de Idade. Maués, AM,	74
Figura 11. Visão geral do experimento em Fatorial NPK com 3 níveis em Maués, AM, Brasil. 2010.....	75
Figura 12. Outra visão geral do experimento em Fatorial NPK com 3 níveis em Maués, AM, Brasil. 2010.	76
Figura 13. Pau-rosa (<i>Aniba rosaedora</i> Ducke), com 1 ano de idade, Maués, AM, Brasil. 2010.....	77
Figura 14. Medição do diâmetro de colo com paquímetro digital. Maués, AM, Brasil. 2010.....	78

ÍNDICE DE TABELAS

Tabela 1. Quadro de análise de variância para um simples fator, modelo de efeitos fixos.	10
Tabela 2. Disposição das observações em Planos de Blocos Completos Casualizados	13
Tabela 3. Quadro de análise de variância para blocos casualizados, modelo de efeitos fixos.	14
Tabela 4. Quadro de ANOVA: quadrados latinos	18
Tabela 5. Quadro de análise de variância: blocos incompletos casualizados.....	20
Tabela 6. Quadro da ANOVA: Fatorial com 2 fatores fixos.....	32
Tabela 8. Construção de um fatorial fracionário 2^{6-2} , de resolução IV, com relação definição: $I=ACDF$ $I=BCDE$	37
Tabela 9. Arranjo fatorial $3 \times 3 \times 3$ confundido com a interação tripla dos grupos W, X, Y e Z.....	39
Tabela 10. Análise da variância do ensaio fatorial $3 \times 3 \times 3$ ($A \times B \times C$), com confundimento de 2 GL da interação tripla e duas repetições.	39
Tabela 11. Quadro de ANOVA: Fatorial confundido $3 \times 3 \times 3$	40
Tabela 12. Arranjo fatorial $3 \times 3 \times 3$ com interação em confundimento.	44
Tabela 13. Médias da altura total do pau-rosa, em metros, aos 12 meses de idade do experimento factorial de 3^3 de adubação com N, P e K, com duas repetições em Maués, AM, Brasil. Manaus, 2010.....	48
Tabela 14. Análise de variância com o efeito da interação tripla confundido com o resíduo, usando o modelo: altura = bloco + Nitro*Phos*Potas - Nitro:Phos:Potas + Erro. Manaus, 2010	50
Tabela 15. Análise de variância com o efeito da interação tripla confundido com o resíduo, usando o modelo: altura = bloco + Nitro*Phos*Potas - Nitro:Phos:Potas - Nitro:Potas - Phos:Potas + Erro. Manaus, 2010	50
Tabela 16. Análise de variância desdobrando o efeito dos níveis do fator N dentro do fator P. Manaus, 2010	52
Tabela 17. Análise de variância desdobrando o efeito dos níveis do fator P dentro do fator N. Manaus, 2010	52
Tabela 18. Comparação de médias para variável altura, do tratamento Nitrogênio (N) dentro dos níveis de Fósforo(P) pelo teste da Diferença Mínima Significativa (LSD). Manaus, 2010.	53
Tabela 19. Médias do diâmetro do colo das plantas de pau-rosa, em mm, aos 12 meses de idade do experimento factorial de 3^3 de adubação com N, P e K, com duas repetições em Maués, AM, Brasil. Manaus, 2010.....	62

RESUMO

Esta dissertação aborda as técnicas de experimentação estatística com ênfase nos planos factoriais. Faz-se uma descrição das principais técnicas, a partir dos delineamentos mais simples- inteiramente casualizados até delineamentos mais complexos como os fatoriais confundidos. Uma aplicação prática de um ensaio em delineamento em blocos em fatorial confundido foi realizada, utilizando-se a espécie *Aniba rosaedora* Ducke (pau rosa), testando-se três tipos de fertilizantes (nitrogênio, fósforo e potássio) em três diferentes níveis, na região de Maués-AM-Brazil. Os resultados indicaram que a técnica de experimentação foi eficiente na discriminação dos resultados, ao mesmo tempo em que reduziu a área experimental e o custo de implantação. Estudos iniciais da resposta da *Aniba rosaedora* aos tratamentos de adubação com nitrogênio(N), P(fósforo) e potássio (K), indicaram que a espécie responde favoravelmente ao crescimento em altura aos fatores N e P e suas interações

SUMMARY

This dissertation discusses the experimental techniques with emphasis on statistical factorial designs. It is a description of the main techniques, from simple designs, randomized designs even more complex as the factorial with confounding. A practical application of a test in randomized blocks in factorial with confounding was performed using the species *Aniba rosaedora* Ducke (rosewood), testing three types of fertilizers (nitrogen, phosphorus and potassium) on three different levels in the region Maués-AM-Brazil. The results indicated that the experimental technique was efficient in the breakdown of the results while reducing the experimental area and the cost of deployment. Initial studies of response to treatment *Aniba rosaedora* fertilization with nitrogen (N), P (phosphorus) and potassium (K), indicated that the species responds favorably to height growth factors N and P and their interactions.

CAPÍTULO 1

1. Introdução

O tema desta dissertação enquadra-se na área da Estatística experimental, com ênfase nos métodos estatísticos em experimentação agrícola. Trata-se de uma linha de estudo importante para apoiar o setor agrícola/florestal da Amazônia, para ajudar a definir as melhores espécies, os melhores tratamentos silviculturais e conseqüentemente promover a maior produtividade dos cultivos agrícolas e florestais.

Para estudar diferentes fatores de interesse em um experimento, quando não se tem conhecimento prévio de como estes diferentes fatores irão influenciar na variável resposta, uma boa opção para o planejamento experimental é a estruturação dos modelos de acordo com planos fatoriais.

Na definição de F. Yates (1937a), experimentos fatoriais são aqueles que incluem todas as combinações de vários conjuntos de tratamentos ou fatores. Os experimentos fatoriais geralmente são mais eficientes do que os experimentos simples com um só conjunto de tratamentos e permitem tirar conclusões mais gerais.

A principal desvantagem dos experimentos fatoriais é que o número de tratamentos aumenta rapidamente. Um ensaio fatorial, em blocos casualizados, com um grande número de tratamentos, em geral, perde bastante sua eficiência. Somos levados, não raramente, a usar blocos incompletos e a adotar o sistema de confundimento (counfounding), do qual resultam maiores complicações no planejamento e análise dos experimentos. Outra solução é o uso dos fatoriais fracionários ou dos fatoriais incompletos, em que não são usadas todas as combinações dos fatores. Não há dúvida, porém, de que os experimentos fatoriais são, em geral, muito convenientes e de uso bastante comum.

Para elaborar esta dissertação, planejou-se e instalou-se um ensaio experimental com a espécie pau-rosa (*Aniba rosaedora*), estruturando as parcelas no campo de acordo com os planos fatoriais. O objetivo deste ensaio foi estudar o comportamento da espécie quando submetida a 3 fatores de fertilização: nitrogênio (N), fósforo (P) e potássio (K) em três níveis. A variável de interesse nesta fase experimental foi o crescimento em altura e o diâmetro do colo das árvores.

CAPÍTULO 2

2. Revisão de literatura: Abordagem histórica do Planeamento de Experiências

“ Todo o método científico assenta no método experimental.” Arquimedes

A metodologia para o Planeamento e Análise de estudos experimentais teve a sua origem a partir de problemas referentes à agricultura, em Londres “Rothamsted Agricultural Experiment Station”, sob a direcção de Sir R. A. Fisher. Como resultado, muita da terminologia ainda actualmente usada reflete conceitos agrícolas, tais como por exemplo “tratamentos”, “variedades”, “parcelas”.

As primeiras contribuições estatísticas de Fisher (1912) foram feitas, quando o método da máxima verosimilhança foi usado primeiramente para estimar parâmetros desconhecidos. Isto foi seguido por uma série de artigos nos 12 anos seguintes, que assentaram as fundações da estatística como um disciplina científica independente, com um grande potencial para aplicações em todas as ciências (Rao, 1992).

Ronald Fisher além de pioneiro da aplicação da estatística matemática ao planeamento de experiências, desenvolveu inúmeras técnicas, como por exemplo a análise de variância, tendo dominado esta nova área científica na época entre a 1ª e a 2ª guerra mundiais (1918-1939). Pode-se ainda citar-se muitos outros autores cuja contribuição foi notável para o desenvolvimento do planeamento de experiências, por exemplo: F. Yates, R.C. Bose, O. Kempthorne, W.G.Cochran, G.P.E. Cox, dos quais são indicadas algumas referências bibliográficas importantes (ver Bibliografia).

Na actualidade os domínios de aplicação do Planeamento de Experiências são extremamente vastos, destacando-se além da agricultura, a engenharia, a física, a química, a sociologia, a psicologia, a medicina, a biologia e recentemente a genética.

CAPÍTULO 3

Princípios e bases da Experimentação Estatística

Um *planejamento de experiências* não é mais do que um protocolo para selecção de amostras das populações que interessam investigar. O Planejamento de Experimentos (DOE – design of experiments), segue o princípio da circularidade do método científico, conforme a figura 1.

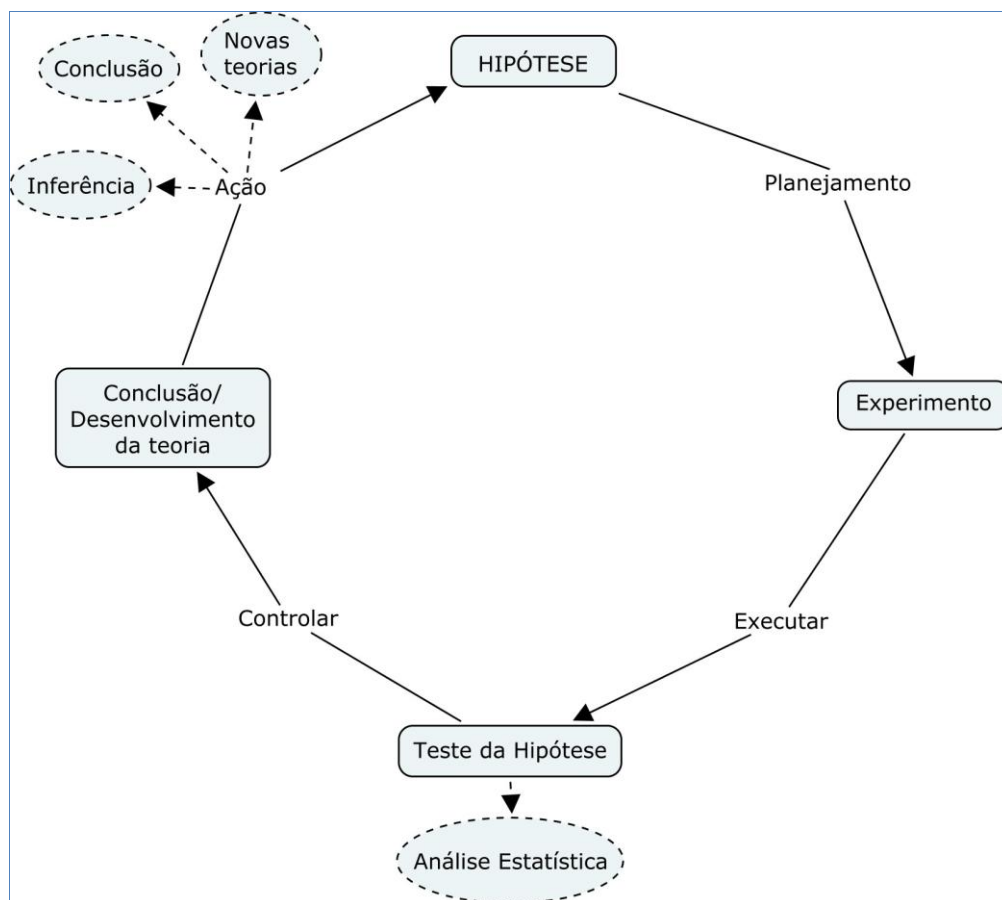


Figura 1. Circularidade do método científico

Todo experimento deve começar com uma hipótese, ou seja, é preciso ter os objetivos do experimento bem definidos. Normalmente quando isso não acontece, não raro, os experimentos não são conclusivos e não fornecem informações úteis.

Dada uma hipótese, ela pode ser avaliada por meio de um experimento. É no planejamento que é definido o arranjo das parcelas, a(s) variável(eis) resposta(s), as unidades de medida etc. Um experimento planejado incorretamente pode resultar na invalidade das informações obtidas e conseqüentemente invalidar as conclusões, além da perda de recursos.

Ao planejar o experimento, uma hipótese é formulada, e após a obtenção dos dados, esta hipótese será testada. Nessa fase é realizada a análise estatística propriamente dita. Testada a hipótese, interpreta-se o resultado, fazendo inferências sobre os resultados. A partir daí, pode-se decidir pela elaboração de uma teoria (conclusões) ou por um novo experimento, completando o ciclo.

A circularidade do método científico pode ser vista como o Ciclo PDCA (Plan, Do, Check and Action). Tanto em um como em outro, busca-se uma resposta para um problema. Na indústria, isso pode ser visto como o desenvolvimento de um novo processo (produto) ou a melhoria de um processo já existente.

Por *programa de investigação* entende-se um esforço organizado por parte de um ou vários cientistas, a fim de adquirir conhecimentos acerca de um determinado processo. Processo esse que pode ser natural ou por exemplo resultante de produção industrial.

O estudo deve ser conduzido no sentido de obter toda a informação necessária da forma mais eficiente, atendendo aos mais diversos problemas, tais como:

- Garantir que o processo de medição usado nas observações seja simples;
- Garantir que o estudo produzirá dados úteis e fiáveis;
- Tentar realizar a experiência minimizando os custos inerentes;
- Garantir a gestão da experiência num período de tempo razoável.

Serão apresentados métodos apropriados para a realização de experiências, de modo a obter estimativas válidas das diferenças entre os tratamentos ensaiados ou testados e, em simultâneo, estabelecem-se estimativas igualmente válidas dos erros cometidos.

Na realização de qualquer experiência é geralmente inevitável a *existência de erros* provenientes das mais diversas causas: a variabilidade do material a ser testado, as condições do meio onde se realiza a experiência, os erros humanos ou instrumentais. São muitos e diversos os factores passíveis de contribuir para aumentar a variabilidade dos resultados finais. Assim, um dos objectivos do planeamento de experiências é o controlo deste erro experimental, quer por meio de medidas tendentes à sua redução, quer pela utilização de métodos específicos para a organização do material em estudo, que permitam a identificação e a separação da variabilidade atribuível a causas definidas.

Serão apresentados em seguida três princípios básicos para a obtenção de um bom Planeamento de Experiências: Uso de réplicas, agrupamentos convenientes e a casualização ou aleatorização.

Assim, uma das condições indispensáveis para a realização de uma experiência é o

recurso a *réplicas* ou *repetições*. A réplica ou repetição consiste na execução da mesma observação ou aplicação do mesmo tratamento mais do que uma vez. Pode repetir-se toda a experiência, ou parte dela, sob alguns conjuntos de condições. A repetição tem duas finalidades importantes: por um lado *aumenta a precisão das observações*, por outro lado *permite a obtenção de estimativas dos erros cometidos*.

Em todo experimento, ocorre sempre uma variação ao acaso entre observações de um mesmo tratamento. Essa variação ao acaso é denominada *erro experimental*. A variação ao acaso pode ocorrer devido a variabilidade do material utilizado no experimento e/ou pelo modo como o experimento foi conduzido. O erro experimental pode ser reduzido se na organização inicial dos dados forem usados *agrupamentos* discriminados em categorias. Com base no tipo de agrupamento os resultados a obter serão certamente mais precisos, conduzindo a uma apreciável redução do erro. Recorrendo à análise de variância, esta permitirá identificar e eliminar o contributo da variação entre repetições, embora não o permita dentro das repetições, principal causa do erro experimental.

As providências a tomar para diminuir o erro experimental são da responsabilidade do planeamento da experiência em si. Tudo o que validamente for feito com o objectivo de uniformizar as condições da realização da experiência, excepto no que respeita aos efeitos dos tratamentos, contribui para uma maior eficiência do teste experimental. Na literatura, há exemplos clássicos de erros graves cometidos na análise de dados, por se ter pretendido à posteriori organizar os resultados de uma forma que não a prevista na planificação.

Com a finalidade de reduzir o *erro experimental* existem os chamados *Delineamentos experimentais*. Um delineamento experimental é a forma como os tratamentos são designados às unidades experimentais ou parcelas. A análise de variância é baseada no delineamento experimental utilizado.

Falou-se na necessidade da repetição, na vantagem do agrupamento e falta finalmente o terceiro princípio básico, a casualização (ou aleatorização).

O princípio da casualização consiste em distribuir aleatoriamente os tratamentos nas unidades experimentais de modo que cada um tenha a mesma chance de ocupar qualquer unidade experimental. A casualização permite uma distribuição independente do erro experimental. A distribuição casual ou aleatória dos tratamentos é condição indispensável para a validação da estimativa do erro.

Por vezes em certas situações experimentais há necessidade de comparar novos métodos (ou tratamentos) com um método (tratamento) já existente e para o qual já são conhecidos os resultados. Um tratamento que não interessa em particular testar, mas que

serve como base de comparação é conhecido por *controlo ou testemunha*. Por exemplo num estudo para avaliar a eficácia de um novo medicamento para a diabetes, é dado um comprimido a cada doente. Alguns destes comprimidos são respeitantes ao novo medicamento e outros são inertes. Este tipo de comprimido inerte é um controlo, também conhecido por placebo.

Numa experiência cada observação (registo de medição ou contagem) é obtida aplicando um tratamento a uma categoria básica de material, conhecida por *unidade experimental*. A unidade experimental pode ser por exemplo um indivíduo, um talhão de terra ou uma máquina. Um grupo homogéneo de unidades experimentais designa-se por bloco. Unidades experimentais homogéneas são unidades tão uniformes quanto possível em todas as características que possam afectar a resposta. Designa-se por *factor* uma variável experimental controlável que se supõe influenciar a resposta (dados observados), e por *nível* o valor específico de um factor. A alteração na resposta média entre duas combinações de factor ou entre duas condições experimentais designa-se por *efeito*.

Em resumo, pode-se afirmar que:

- A aleatorização torna os testes estatísticos válidos
- A repetição faz os testes estatísticos possíveis
- O controle local faz o experimento mais eficiente

3.1. Análise de variância – pressupostos fundamentais

A *análise de variância* (ANOVA) é um processo de análise baseado na decomposição da variação total existente entre uma série de observações. Esta variabilidade pode ser decomposta em causas conhecidas (ex. tratamentos e blocos) e em outra parte devido a causas desconhecidas (erro experimental ou resíduo).

Para que a análise de variância seja considerada válida, algumas pressuposições devem ser obedecidas. As principais delas, segundo Zimmermann (2004), são que os erros devem ser independentes; com distribuição normal; com média $\mu = 0$; variância σ^2 e modelo matemático linear aditivo. Ademais, exige-se que os erros tenham todos a mesma variância e que os dados sejam quantitativos, isto é, que se enquadrem nas escalas de proporção ou de intervalo.

Para Sokal & Rohlf (1979), se os erros não são independentes, a validade do teste F sobre a significância pode ser afetado gravemente. Além disso, para Sokal & Rohlf (1979), as conseqüências da não normalidade dos erros não são demasiados graves; unicamente uma distribuição muito assimétrica teria um efeito marcante sobre o nível de significância do teste F ou sobre a eficiência do desenho experimental.

Portanto, a análise de variância só é possível se forem satisfeitas certas condições, ou seja, certas exigências do modelo matemático (Barbin, 2003):

- 1º.) O modelo deve ser aditivo, isto é, os efeitos devem se somar (não há interação);
- 2º.) Os erros (e_{ij}) devem ter distribuição normal;
- 3º.) Os erros (e_{ij}) devem ser independentes;
- 4º.) Os erros (e_{ij}) devem ter a mesma variância, ou seja, deve existir homocedasticidade.

As pressuposições em relação aos erros podem ser resumidas na seguinte expressão:

$$e_{ij} \sim \text{NID}(0, \sigma^2) \text{ ou } e_{ij} \sim \text{IID}(0, \sigma^2)$$

Todas essas pressuposições podem ser verificadas através da análise de resíduos.

CAPÍTULO 4

Modelos de Planeamento de Experiências com um Factor

4.1 Planeamento Completamente Casualizado

O tipo de planeamento mais simples é aquele em que os tratamentos são atribuídos aleatoriamente às unidades experimentais, conhecido por planeamento completamente casualizado ou experimentos inteiramente ao acaso. Se um tratamento é para ser aplicado a 5 unidades experimentais, então a casualização atribui a cada grupo de 5 unidades igual probabilidade de receber o tratamento.

Segundo Banzatto & Kronka (2008), o delineamento inteiramente casualizado apresenta as seguintes características:

- a) Utiliza apenas os princípios da repetição e da casualização, deixando de lado o princípio do controle local, e, portanto as repetições não são organizadas em blocos;
- b) Os tratamentos são designados às parcelas de forma inteiramente casual, com números iguais ou diferentes por tratamento.

Este tipo de planeamento é apropriado quando o material experimental é homogéneo. Trata-se de um procedimento corrente, por exemplo no âmbito de experiências laboratoriais em química e biologia, onde a quantidade de material é dividida em pequenas amostras às quais são aplicados os tratamentos. É também adequado no caso de experiências de pequena dimensão, ou quando é provável a existência de observações omissas.

Este tipo de planeamento apresenta como principal desvantagem o problema da precisão pelo facto de a variação ser espalhada por todas as unidades, mas apresenta algumas vantagens importantes, tais como:

- É permitida uma flexibilidade total pois pode ser usado qualquer número de tratamentos e réplicas, podendo o número de réplicas variar de tratamento para tratamento, embora tal não seja recomendado, a menos que por uma boa causa. Todo o material

disponível é passível de ser utilizado, o que é uma grande vantagem especialmente quando o material é escasso.

- A análise estatística é simples, mesmo quando o número de réplicas não é o mesmo para todos os tratamentos, ou quando os erros experimentais diferem de tratamento para tratamento.

- O método de análise mantém-se simples quando se perdem ou são rejeitados os resultados de algumas unidades ou de tratamentos completos. Não é tão grave como noutros planeamentos o caso da perda de dados ou observações omissas.

As observações de um experimento inteiramente ao acaso podem ser descritas através do **modelo médio** (Montgomery, 2001):

$$y_{ij} = \mu_i + \varepsilon_{ij} \quad \left\{ \begin{array}{l} i = 1, 2, \dots, a \\ j = 1, 2, \dots, n \end{array} \right.$$

Onde y_{ij} é a ij -ésima observação, μ_i é a média do i -ésimo fator ou tratamento, e ε_{ij} é o erro aleatório, componente que incorpora todas as outras fontes de variação no experimento, incluindo medições, variabilidade de fatores incontroláveis, diferenças entre as unidades experimentais no qual os tratamentos são aplicados, e todos os ruídos gerais do processo (variabilidade no tempo, efeitos das variáveis ambientais, etc).

Uma alternativa ao **modelo médio** é definir $\mu_i = \mu + \tau_i$, com $i = 1, 2, \dots, a$

Então, teríamos:

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij} \quad \left\{ \begin{array}{l} i = 1, 2, \dots, a \\ j = 1, 2, \dots, n \end{array} \right.$$

Nesta forma μ é um parâmetro comum a todos tratamentos, chamado de **média geral**, e τ_i é um parâmetro único para o i -ésimo tratamento chamado de **efeito do i -ésimo tratamento**.

A Análise Estatística destes planos é uma aplicação da análise de variância a um factor. Considerem-se então k níveis do factor a testar e n_i observações por nível e mantenham-se as notações adoptadas na análise de variância a um factor. Assim admitindo o

modelo linear e os usuais pressupostos, tem-se o quadro de análise de variância conforme apresentado na tabela 1:

Tabela1 Quadro de análise de variância para um simples fator, modelo de efeitos fixos.

F.V.	G.L.	S.Q.	Q.M.	F ₀
Entre Amostras (tratamentos)	a-1	$\sum_{i=1}^k \frac{y_{io}^2}{n_i} - \frac{y_{oo}^2}{n}$	QMA	$F_0 = \frac{QMA}{QME}$
Erro ou Resíduo	n-a	SQE=(SQT-SQA)	QME	
Total	n-1	$\sum_{i=1}^k \sum_{j=1}^{n_i} y_{ij}^2 - \frac{y_{oo}^2}{n}$		

Fonte: Montgomery, 2001

4.2 Planos de Blocos Completos Casualizados

Nos Planos de Blocos Completos Casualizados procura-se melhorar a precisão de estimativas de diferenças entre médias de determinados tratamentos ou variedades. O domínio de aplicação mais vasto deste tipo de delineamentos é a Agricultura.

Neste domínio, um Plano de Blocos Completos Casualizados consiste em dividir um campo em b blocos e subdividir cada bloco em v parcelas ou unidades experimentais. A cada um dos b blocos serão aplicados aleatoriamente os v tratamentos ou variedades. Cada bloco, consistirá então numa réplica da experiência e cada variedade ou tratamento repete-se r vezes. O número de unidades experimentais será conseqüentemente “ br ”. No decorrer da experiência deve-se procurar manter os erros experimentais tão pequenos quanto seja praticável para cada grupo, e quaisquer mudanças, quer de condições quer de ordem técnica, susceptíveis de afectar os resultados, deverão ser feitas entre os grupos.

Estes planos são especialmente convenientes para campos experimentais em que o número de tratamentos é relativamente pequeno. O seu objectivo principal será reduzir o erro experimental eliminando a contribuição das conhecidas fontes de variação entre unidades experimentais. A ideia subjacente aos Blocos Casualizados é que, dispondo aleatoriamente os tratamentos nos diferentes blocos, a variabilidade dentro de cada bloco seja minimizada e a variabilidade entre blocos seja maximizada.

Este tipo de delineamento, designado por Plano de Blocos Casualizados é o mais usado no Planeamento de Experiências, desde que cada bloco seja suficientemente largo para conter os v tratamentos, ou seja quando existem unidades experimentais homogéneas em número suficiente, por forma a que todos os tratamentos apareçam juntos em cada bloco. Não há restrições quanto aos tamanho dos blocos nem quanto ao número de tratamentos.

Segundo Banzatto & Kronka (2008), As principais características deste delineamento são:

- a) as parcelas são distribuídas em grupos ou blocos (princípio do controle local), de tal forma que elas sejam o mais uniforme possível, dentro de cada bloco;
- b) o número de parcelas por bloco deve ser um múltiplo do número de tratamentos (geralmente o número de parcelas é igual ao número de tratamentos);

- c) os tratamentos são designados às parcelas de forma casual, sendo esta casualização feita dentro de cada bloco.

Em relação aos outros delineamentos, o delineamento em blocos casualizados apresenta as seguintes desvantagens:

- a) pela utilização do princípio local, há uma redução no número de graus de liberdade do resíduo;
- b) a exigência da homogeneidade das parcelas dentro de cada bloco limita o número de tratamentos, que não pode ser muito elevado.

Para o delineamento em blocos casualizados, o **modelo matemático** é o seguinte (Montgomery, 2001):

$$y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij} \begin{cases} i = 1, 2, \dots, a \\ j = 1, 2, \dots, n \end{cases}$$

Onde μ é a **média geral**, τ_i é o efeito do i -ésimo tratamento, β_j é o efeito do j -ésimo bloco, e ε_{ij} é o usual NID $(0, \sigma^2)$ erro aleatório.

Os resultados podem ser apresentados na forma matricial de r linhas por b colunas, com uma observação por célula, como se segue na tabela 2.

Tabela 2. Disposição das observações em Planos de Blocos Completos Casualizados

	Blocos				Totais	
	1	2	...	B		
Tratamentos (variedades)	1	y_{11}	y_{12}	\dots	y_{1b}	T_1
	2	y_{21}	y_{22}	\dots	y_{2b}	T_2
	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
	v	y_{r1}	y_{r2}	\dots	y_{rb}	T_r
Totais	B_1	B_2	\dots	B_b	G	

A cada valor observado, podemos associar o modelo matemático do delineamento, então para o valor observado no tratamento i no bloco j , teremos:

$$\begin{aligned} x_{ij} &= \mu + t_i + b_j + e_{ij} \\ e_{ij} &= x_{ij} - \mu - t_i - b_j \end{aligned}$$

em que

A Análise Estatística dos planos em blocos casualizados é uma aplicação da análise de um factor acrescido do efeito dos blocos. Considerem-se então A níveis do factor a testar, B blocos e n_i observações por nível. Assim admitindo o modelo linear e os usuais pressupostos, tem-se o quadro de análise de variância conforme apresentado na tabela 3.

Tabela 3. Quadro de análise de variância para blocos casualizados, modelo de efeitos fixos.

F.V.	G.L.	S.Q.	Q.M.	F ₀
Tratamentos	a-1	$\sum_{i=1}^k \frac{y_{i.}^2}{n_i} - \frac{y_{..}^2}{n}$	$\frac{SQT}{a-1}$	$\frac{QM_{tratamentos}}{QM_{erro}}$
Blocos	b-1	$\frac{1}{a} \sum_{i=1}^a y_{i.}^2 - \frac{y_{..}^2}{N}$	$\frac{SQB}{b-1}$	
Erro	(a-1)(b-1)	SQE=(SQT-SQA)	$\frac{SQ_{erro}}{(a-1)(b-1)}$	
Total	n-1	$\sum_{i=1}^k \sum_{j=1}^{n_i} y_{ij}^2 - \frac{y_{..}^2}{n}$		

Fonte: Montgomery, 2001

4.3 Quadrado Latino: Modelo em agrupamentos duplos

4.3.1 Definição de quadrado latino

Os quadrados latinos levam em conta o controle local, aplicado em dois sentidos. É um delineamento usado para eliminar duas fontes de variabilidade; de tal forma que sistematicamente aloca-se blocos em duas direções (Montgomery, 2001; Barbin, 2003).

Define-se um quadrado latino de lado m como uma matriz quadrada onde cada uma das m letras ocorre uma vez em cada linha e uma vez em cada coluna.

São exemplos de quadrados latinos os seguintes:

(1) $m=4$

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

(2) $m=5$

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

Um quadrado latino no qual as letras das primeiras linha e coluna se encontram ordenados alfabeticamente é conhecido por *Quadrado Latino Standard* (caso 1).

No quadrado latino os tratamentos são agrupados em réplicas de dois modos diferentes. O efeito do agrupamento duplo é eliminar nos erros, todas as diferenças entre as linhas e igualmente todas as diferenças entre as colunas. Com este tipo de planejamento, os quadrados latinos permitem ainda uma maior redução de erros do que os blocos casualizados.

A experiência deve ser conduzida e o material experimental deve ser organizado por forma a que as diferenças entre as linhas e as colunas representem as maiores origens de variação. Graças à ortogonalidade podem ainda ser realizadas três tipos de comparações independentes entre os elementos do quadrado latino, isto é, as comparações feitas entre as linhas, entre as colunas e entre as letras (tratamentos) são independentes entre si. Entende-se por comparação entre as letras, comparações de todos os A ou da sua média, com todos os B, os C, os D e os E.

Exemplo:

No caso de uma experiência agrícola, para comparar 5 tratamentos por este método, as linhas representariam 5 faixas de terreno paralelas entre si, as colunas representariam 5 faixas paralelas entre si mas perpendiculares às primeiras de modo a que se definissem 25 talhões. As letras constituiriam os tratamentos que seriam distribuídos pelos 25 talhões de acordo com o esquema, isto é sem repetições na linha ou na coluna.

A desvantagem do quadrado latino consiste no facto de o número de réplicas ser igual ao número de tratamentos. É óbvio que se há muitos tratamentos a comparar o número de réplicas torna-se impraticável. Os Quadrados Latinos mais frequentemente usados variam entre 4×4 e 8×8 . A partir de $m=12$ raramente se usa este tipo de planeamento.

4.3.2 Análise de variância do Quadrado Latino

Seja Y_{ijk} , $i=1,\dots,m$; $j=1,\dots,m$; $k=1,\dots,m$, a variável aleatória representativa da observação da i -ésima linha, j -ésima coluna e k -ésimo tratamento. O modelo linear adequado para representar este tipo de dados será (Montgomery, 2001):

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \tau_k + \varepsilon_{ijk},$$

com $E(Y_{ijk}) = \mu + \alpha_i + \beta_j + \tau_k$ e $V(Y_{ijk}) = \sigma^2$,

onde:

- $\alpha_i \rightarrow$ representa o efeito da i -ésima linha
- $\beta_j \rightarrow$ representa o efeito da j -ésima coluna
- $\tau_k \rightarrow$ representa o efeito do k -ésimo tratamento
- $\varepsilon_{ijk} \rightarrow$ representa o erro aleatório

O modelo em questão é completamente aditivo, ou seja não há interacção entre linhas, colunas e tratamentos. Note-se que $\sum_{i=1}^m \alpha_i = 0$, $\sum_{j=1}^m \beta_j = 0$ e $\sum_{k=1}^m \tau_k = 0$.

A análise de variância consiste numa partição da soma de quadrados total, das m^2 observações, em componentes para linhas, colunas tratamentos e erro, ou seja:

$$SQT = SQL + SQC + SQtrat + SQE$$

Com os graus de liberdade respectivos:

$$m^2 - 1 = m - 1 + m - 1 + m - 1 + (m - 2)(m - 1).$$

Assumindo os usuais pressupostos relativamente ao erro, ou seja ε_{ijk} independentes e identicamente distribuídos de acordo com a distribuição normal de valor médio 0 e variância σ^2 , pode-se escrever que :

$$\frac{SQL}{\sigma^2} \approx \chi_{m-1}^2, \quad \frac{SQC}{\sigma^2} \approx \chi_{m-1}^2, \quad \frac{SQtrat}{\sigma^2} \approx \chi_{m-1}^2 \quad \text{e} \quad \frac{SQE}{\sigma^2} \approx \chi_{(m-2)(m-1)}^2$$

Represente-se por:

- $R_i \rightarrow$ Total das observações da linha i
- $C_j \rightarrow$ Total das observações da coluna j
- $T_k \rightarrow$ Total das observações do tratamento k

$G \rightarrow$ Grande total , soma de todas as observações

$Q \rightarrow$ soma dos quadrados das observações

A tabela de análise de variância apropriada para estes dados é mostrada na tabela 4.

Tabela 4. Quadro de ANOVA: quadrados latinos

F.V.	G.L.	S.Q.	Q.M.	R.V.
Entre Linhas	m-1	$SQL = \sum_{i=1}^m \frac{R_i^2}{m} - \frac{G^2}{m^2}$	$QML = \frac{SQL}{m-1}$	$F_0 = \frac{QML}{QME}$
Entre Colunas	m-1	$SQC = \sum_{j=1}^m \frac{C_j^2}{m} - \frac{G^2}{m^2}$	$QMC = \frac{SQC}{m-1}$	$F_1 = \frac{QMC}{QME}$
Entre Tratamentos	m-1	$SQ_{trat} = \sum_{k=1}^m \frac{T_k^2}{m} - \frac{G^2}{m^2}$	$QM_{trat} = \frac{SQ_{trat}}{m-1}$	$F_{trat} = \frac{QM_{trat}}{QME}$
Erro	(m-1)(m-2)	$SQE = SQT - SQL - SQC - SQ_{trat}$	$QME = \frac{SQE}{(m-1)(m-2)}$	
Total	n-1	$SQT = Q - \frac{G^2}{m^2}$		

A estatística apropriada para testar a hipótese de que não há diferenças entre tratamentos será F_{trat} , obedecendo a uma distribuição $F_{m-1, (m-2)(m-1)}$, sob a hipótese nula (de não haver diferenças entre tratamentos). Valores observados de F_{trat} superiores ao valor tabelado $F_{m-1, (m-2)(m-1)} (\alpha\%)$, conduzem à rejeição da hipótese nula, sendo portanto significativos ao nível de significância α .

Pode ainda ser de interesse averiguar os efeitos das linhas e das colunas, comparando respectivamente F_0 e F_1 com o valor tabelado da distribuição $F_{m-1, (m-2)(m-1)} (\alpha\%)$. No entanto, e uma vez que tanto as linhas como as colunas representam restrições à casualização, estes testes não são de manifesto interesse.

CAPÍTULO 5

Modelos de Planeamento de Experiências com Blocos

5.1 Planos em Blocos Incompletos

Em aplicações experimentais utilizando blocos casualizados, pode não ser possível implementar todos os tratamentos em cada um dos blocos. Para este tipo de problema é possível usar um delineamento em blocos casualizados em que cada um dos tratamentos não está presente em todos os blocos. Este tipo de delineamento é conhecido como blocos incompletos casualizados (Montgomery, 2001).

Os delineamentos em blocos incompletos equilibrados se caracterizam pelo fato de que neles cada tratamento aparece no mesmo bloco, com cada um dos outros tratamentos, e sempre o mesmo número de vezes λ (Pimentel-Gomes & Garcia, 2002).

Supondo que tem-se a tratamentos e que em cada bloco tem-se exatamente k tratamentos ($k < a$). Um Delineamento de bloco incompleto balanceado do inglês *balanced incompleted block design* (BIDB) pode ser construído fazendo $\binom{a}{k}$ blocos e atribuindo uma diferente combinação de tratamentos para cada bloco. Frequentemente, todavia, balanceamento pode ser obtido com menos que $\binom{a}{k}$ blocos (Montgomery, 2001).

Tabelas de BIDBs podem ser verificadas em Fisher & Yates (1953), Davies (1956), e Cochran & Cox (1957).

Na tabela 5 a seguir apresenta-se o quadro da análise de variância para um plano de BIDB.

Tabela 5. Quadro de análise de variância: blocos incompletos casualizados.

Origem de Variação.	G.L.	S.Q.	Q.M.	Razão de Variâncias
Tratamentos (ajustados)	$\frac{k \sum Q_i^2}{\gamma a}$	a-1	$\frac{SS_{tratamentos}}{a-1}$	$F_o = \frac{MS_{tratamentos}}{MS_e}$
Blocos	$\frac{1}{k} \sum y_{.j}^2 - \frac{y_{..}^2}{N}$	b-1	$\frac{SS_{blocos}}{b-1}$	
Erro	SS_e (diferença)	N-a-b+1	$\frac{SS_E}{N-a-b+1}$	
Total	$\sum \sum y_{ij}^2 - \frac{y_{..}^2}{N}$	N-1		

Fonte: Montgomery (2001)

Onde,

$$\gamma = \frac{r(k-1)}{a-1}$$

k = número de tratamentos contido em cada bloco;

r = número de replicações;

a = número de tratamentos;

b = número de blocos.

Se a=b, o delineamento é simétrico. O parâmetro γ deve ser um número inteiro.

Quando todos os tratamentos são igualmente importantes, a combinação de tratamentos em cada bloco dever ser selecionada de maneira balanceada. Assim, um delineamento em bloco incompleto balanceado é um delineamento em bloco incompleto no qual qualquer par de tratamentos ocorrem juntos o mesmo número de vezes em relação aos outros pares.

5.2 Planos em Blocos Incompletos Equilibrados

Uma apresentação detalhada aos BIBD é abordada em Oliveira, T. (1994). Nesse trabalho são definidos os IBD e os BIBD, mencionados os teoremas fundamentais, é descrita a análise estatística dos IBD e dos BIBD e são estudados os BIBD com repetições e a sua construção. Os BIBD com Repetições permitem simplificar a concretização das experiências e a análise dos resultados obtidos. Note-se que, a definição de BIBD não impõe qualquer restrição relativamente à repetição de blocos, desta forma quando há $b^* < b$ blocos distintos num plano, diz-se que há repetições de blocos.

No seguimento deste trabalho, Oliveira, T. (1999) aprofunda os temas anteriores e estuda os PBIBD com aplicações à genética. São também apresentados casos particulares de BIBD com blocos repetidos, com diferentes suportes (número de blocos não repetidos no plano) e diferentes estruturas. Modelos de BIBD com repetições de blocos descobrem-se em Oliveira, T. (2004).

De forma a dar resposta à controversa rigidez das condições impostas aos BIBD, Bose e Nair (1939) admitiram e expandiram os Planos em Blocos Incompletos Parcialmente Equilibrados (PBIBD) e os Esquema de Associação. Estes autores, nos seus trabalhos introduziram implicitamente na definição dos PBIBD os esquemas de associação, no entanto não encontraram expressamente definidos. Na primeira definição de PBIBD todos os parâmetros λ_i eram distintos, contudo Nair e Rao (1942) mostraram que essa restrição era desnecessária. Bose e Nair (1939) defendiam ainda que os contrastes não eram todos estimados com a mesma variância. Posteriormente, Bose e Shimamoto (1952) concluíram também que esta condição não era necessária. Estes autores classificaram ainda os PBIBD para duas classes de associação: divisível em grupos, simples, triangular, quadrados latinos e cíclico.

Bose e Mesner (1959) recorreram à álgebra para incrementar a sua tarefa em delineamento experimental, originando a álgebra conhecida hoje como Álgebra de Bose-Mesner. Entretanto, assiste-se que, já anteriormente, Connor e Clatworthy (1954) tinham alcançado resultados fascinantes neste âmbito.

Os anos 50 e 60 foram produtivos na descoberta de resultados referentes a esquemas de associação, havendo o objectivo de classificar todos os esquemas de associação e catalogar

todos os PBIBD. Por exemplo, nas listas de Bose e Shimamoto (1952) encontram-se novos esquemas de associação. Vartak (1955) introduziu o esquema rectangular com três classes de associação que anteriormente foram estudados por Harshbarger (1946, 1947 e 1949). Roy P. (1953) introduziu os esquemas de associação hierárquicos divisíveis em grupos. Hinkelmann e Kempthorne (1963) reinventaram os esquemas de associação factoriais como uma extensão dos grupos divisíveis, trabalhos depois desenvolvidos por Kusumoto (1967) e Surendran (1968).

John, P. (1966) universalizou os esquemas triangulares, embora esta ampliação tenha também sido efectuada independentemente por Bose e Laskar (1967) e por Kusomoto (1965). Kempthorne (1956) retomou o estudo dos factores de eficiência nos Planos em Blocos Incompletos e Pearce (1968) estudou-os nos planos equirreplicados. James e Wilkinson (1971) admitiram os factores de eficiência canónica e os contrastes foram estudados por Pearce et al. (1974).

Patterson e Williams (1976) introduziram um novo tipo de Planos em Blocos Incompletos: os Planos α . Estes planos podem ser aplicados sempre que o número de tratamentos é múltiplo da dimensão dos blocos, isto é de k , contudo só se aplicam quando os planos são resolúveis. Os Planos α foram descobertos através do recurso a um processo de construção cíclica, tendo uma aplicação bastante vasta. Patterson *et al* (1978) catalogaram os Planos α mais eficientes, com 100 tratamentos e 4 réplicas. A partir desta altura, o recurso a algoritmos informáticos permitiu um avanço significativo na geração de planos com factores de eficiência mais relevantes, não só no que concerne à descoberta de novos Planos α , bem como de todos os outros delineamentos.

Investigação, revisão e sistematização dos PBIBD e dos esquemas de associação aparecem em trabalhos de Bose (1963), Kageyama (1974), John, P. (1971) e Raghavarao (1971). No artigo de Clatworthy (1973) encontram-se tabelas de PBID. Os planos cíclicos foram também tabelados em John *et al* (1972) e são um dos principais temas dos livros John (1987), John e Williams (1995) e Hinkelmann e Kempthorne (2005).

A ocorrência da repetição de blocos nos Planos de Blocos Incompletos Equilibrados (Balanced Incomplete Block Design - BIBD) é desejável, não só por estar provado serem planos optimos, mas também do ponto de vista económico e da facilidade de aplicação prática (Foody and Hedayat, 1977; Hedayat and Hwang, 1984). Para além disso, em casos

problemáticos que por acidente se percam unidades experimentais a existência de blocos repetidos revela-se de primordial importância.

Cochran & Cox (1957), definem três tipos de blocos incompletos equilibrados:

- Tipo I: experimentos em que os blocos possam ser agrupados em repetições.
- Tipo II: experimentos em que os blocos possam ser dispostos em grupos de repetições.
- Tipo III: experimentos cujos blocos não possam ser agrupados em repetições ou grupo de repetições.

Para Pimentel-Gomes & Garcia (2002), a análise dos experimentos em blocos incompletos é mais difícil do que a dos experimentos inteiramente ao acaso ou em blocos completos casualizados. Contudo, essa maior dificuldade da análise, como também a maior perda de graus de liberdade no resíduo, é frequentemente compensada por uma redução grande no Quadrado Médio do Resíduo, proporcionando experimentos mais precisos.

5.3 Planos em Blocos Incompletos Parcialmente Equilibrados (PBIBD- Partially Balanced Incomplete Block Designs)

5.3.1 PBIBD: Resenha histórica

Os Planos em Blocos Incompletos Parcialmente Equilibrados (PBIBD) surgem da necessidade de dar resposta ao facto de determinadas condições impostas aos parâmetros nos Planos em Blocos Incompletos Equilibrados (BIBD) não serem contempladas ou serem de difícil aplicação.

O requisito de λ ser um número inteiro pode implicar que o número de blocos e a sua dimensão atinjam valores impraticáveis para uma experiência. Recorde-se que $\lambda = r(k-1)/(v-1)$ (1.2.2.), então se $v-1$ for um número primo, r terá de ser um número múltiplo desse primo e terá ainda que se verificar a restrição $b = rv/k$ (1.2.1.). Por exemplo, considerando as restrições anteriores, para num plano onde se pretende aplicar 32 variedades, r terá de ser um número múltiplo de 31 e, se por acaso, o número de unidades experimentais for ímpar e primo, b terá de ser múltiplo de 31×32 .

Bose e os seus alunos introduziram e desenvolveram os Planos em Blocos Incompletos Parcialmente Equilibrados (PBIBD) e os Esquema de Associação. Em 1939, Bose e Nair definiram os PBIBD. Neste trabalho, os esquemas de associação foram introduzidos implicitamente na definição dos PBIBD, mas não foram definidos. Os autores defendiam que as concorrências para as diferentes classes de associação deveriam ser diferentes. Em 1952, Bose e Shimamoto concluem que esta condição não era necessária. Definiram também, pela primeira vez, os esquemas de associação e classificaram quatro tipos de esquemas de associação de ordem três: delineamento divisível em grupos, delineamento em quadrados latinos, delineamento cíclico e delineamento triangular. Designaram estes delineamentos como miscelâneas.

Bose e Connor, em 1952, estudaram o delineamento divisível em grupos classificando-o como: regular, semi-regular e singular. Em 1959, Bose e Mesner recorreram à álgebra para desenvolver o seu trabalho no delineamento experimental, dando origem à álgebra hoje reconhecida como Álgebra de Bose-Mesner, embora Connor e Clatworthy (1954) tivessem já chegado a resultados interessantes nesta área.

Os anos 50 e 60 foram profícuos na descoberta de resultados referentes a esquemas de associação, havendo o objectivo de classificar todos os Esquemas de Associação e catalogar todos os PBIBD.

Harshbarger (1946, 1947 e 1949) descobriu os reticulados rectangulares na sequência dos reticulados quadrados. Em 1951, Nair refere que os “Pair (n)” são esquemas de associação e que os reticulados rectangulares duplamente replicados são parcialmente equilibrados. Em 1965, Ishii e Ogawa também referem este esquema de associação.

Vartak (1955, 1959) estudou os esquemas de associação reticulados. P. M. Roy, em 1953, introduziu os esquemas de associação divisíveis em grupos hierárquicos. Hinkelmann e Kempthorne, 1963, reinventaram esquemas de associação factoriais como uma extensão dos grupos divisíveis, tal como Kusumoto (1967) e Surendran (1968). Com alguma precaução podem ser vistos como exemplos de esquemas obtidos a partir de esquemas triviais pelo cruzamento iterativo e encaixes (iterated crossing and nested). Por um processo análogo Tharthare (1963, 1965) generalizou esquemas de ângulo recto (Right-angular schemes) do tipo $\left(\underline{\underline{n_1}} \times \underline{\underline{n_2}}\right) / \underline{\underline{n_3}}$ e Adhikary (1966) generalizou os do tipo $\underline{\underline{n_1}} / \left(\underline{\underline{n_2}} \times \underline{\underline{n_3}}\right)$.

Artigos de revisão dos PBIBD e dos esquemas de associação foram produzidos por Bose (1963), Kageyama (1974), P. John (1971) e Raghavarao (1971). No artigo de Clatwothy (1973) encontram-se tabelas de PBIBD.

Trabalho desenvolvido recentemente

Nos últimos vinte anos, o destaque inicialmente dado aos PBIBD foi suprido pelo estudo dos esquemas de associação. Cada um dos PBIBD caracterizados anteriormente, tais como: delineamento divisível em grupo, delineamento triangular, etc., foram estudados independentemente, assumindo cada um deles grande relevância, dando origem a novos capítulos nesta área.

Os esquemas de associação assumiram um papel preponderante. Este tema suscitou o interesse de estatísticos e matemáticos, na medida em que, envolve assuntos diversificados como o cálculo combinatório, a teoria de grafos, os grupos abelianos e a teoria de permutações de grupos. Salienta-se o recente livro de Bailey R. (2004) que aborda as várias vertentes dos esquemas de associação.

Os esquemas de associação têm uma vasta aplicabilidade na Estatística porque permitem formar estruturas básicas para muitos delineamentos experimentais. Várias experiências, como por exemplo, as estruturas em blocos ortogonais, apresentam mais do que um sistema de blocos, tornando-se complexo estudar a sua interacção. Apesar de este tema ter sido introduzido independentemente dos PBIBD, as estruturas ortogonais são um esquema de associação, o que facilita o seu estudo.

No mundo actual, em que cada vez mais se recorre às novas tecnologias e à codificação de dados e de movimentos, reveste-se da maior importância investigar planos específicos, tais como os PBIBD, que permitem dar resposta às novas problemáticas em áreas tão relevantes como a Genética, a Medicina, a Biometria ou a Criptografia.

Em termos futuros, podemos ainda conjecturar se o facto dos PBIBD terem aplicabilidade na descodificação de esquemas criptográficos, não se virá a revelar de primordial importância na descodificação dos tão emblemáticos esquemas de partículas atómicas; ou se a solução tão procurada para resolver a actual crise energética, não residirá no recurso ao *bio-diesel* - e ser então urgente e necessário rentabilizar a produção de determinadas plantações, recorrendo a experiências agrícolas estruturadas eventualmente sob a forma de PBIBD.

5.3.2 PBIBD: Definições

Esquema de associação: Definição

Dados ν tratamentos $1, 2, \dots, \nu$, uma relação diz-se um esquema de associação com m classes associadas se:

i) Quaisquer dois tratamentos são 1.º, 2.º, ..., ou m -ésimos associados tal que, se o tratamento α é i -ésimo associado do tratamento β , então o tratamento β é i -ésimo associado do tratamento α .

ii) Qualquer tratamento α tem n_i i -ésimos associados, sendo n_i independente de α .

iii) Sejam α e β dois tratamentos quaisquer i -ésimos associados, então o número de tratamentos j -ésimos associados de α e k -ésimos associados de β é p_{jk}^i , sendo p_{jk}^i independente do par de i -ésimos associados.

Um esquema de associação com m classes é definido pelos parâmetros: ν , n_i e p_{jk}^i $i, j, k = 1, 2, \dots, m$.

A definição anterior deve-se a Bose e Shimamoto (1952). Bose e Mesner (1959) mostraram que a condição assumida nos primeiros artigos, $p_{jk}^i = p_{kj}^i$, era redundante na medida em que é uma consequência da definição de esquema de associação. Bose e Clatworthy (1955) mostram que para duas classes de esquemas de associados não é necessário assumir que os parâmetros p_{jk}^i sejam constantes porque assumindo que n_1 , n_2 , p_{11}^1 e p_{12}^1 são constantes, as restantes condições advêm automaticamente.

PBIBD: Definição

Um esquema de associação com m classes, definido pelos respectivos parâmetros, é um Plano Incompleto Parcialmente Equilibrado com m classes de associação (PBIBD (m)) se os ν tratamentos estão dispostos em b blocos de dimensão k e $k < \nu$, tal que:

i) Cada tratamento ocorre no máximo uma única vez em cada bloco (plano binário) de dimensão k (plano próprio).

ii) Cada tratamento ocorre em r blocos (plano equi-replicado).

iii) Dois tratamentos α e β i -ésimos associados ocorrem juntos em λ_i blocos, sendo λ_i independente do par de tratamentos i -ésimos associados escolhidos.

Um PBIBD é definido pelos parâmetros (v, b, r, k, λ_i) , $i = 1, 2, \dots, m$. Os parâmetros anteriores são designados por parâmetros de primeira ordem e os parâmetros p_{jk}^i por parâmetros de segunda ordem.

Os parâmetros p_{jk}^i podem ser escritos sob a forma de m matrizes simétricas P_i , tal que: $P_i = [p_{jk}^i]$, $i = 1, 2, \dots, m$ $i, j, k = 1, 2, \dots, m$ e p_{jk}^i representa o termo de ordem (j, k) na matriz P_i .

CAPÍTULO 6

Experimentação em Planos Factoriais

6.1 Introdução aos Planos Factoriais

Quando há vários factores de interesse em um experimento, um planeamento fatorial deve ser usado. Em tais experimentos, os factores variam juntos. Especificamente, por um experimento fatorial queremos dizer que em cada tentativa completa ou replicação do experimento, são investigadas todas as combinações possíveis dos níveis dos factores. Assim, se há dois factores A e B , com a níveis para o factor A e B níveis para o factor B , então cada replicação contém todas as ab combinações possíveis (Montgomery, 2004).

Nos planos factoriais existe uma variável que se designa por variável resposta, que é dependente de outras variáveis controladas. Estas variáveis controladas são chamadas de factores, sendo que cada factor pode ter duas ou mais versões, designadas por níveis. As combinações dos níveis dos factores designam-se por tratamentos. Segundo Zimmermann (2004), chama-se de factor àquilo que se quer testar e de níveis às suas diferentes manifestações. Por exemplo, em estudos de adubação de plantas com três elementos (nitrogênio, fósforo e potássio) são os macronutrientes, que são considerados os factores e suas diferentes doses de aplicação os níveis. Se apenas um elemento for testado num ensaio, cada nível será chamado de tratamento, mas, se dois ou três elementos forem testados, cada combinação entre os níveis de cada elemento é que será declarada como um tratamento.

O objectivo dos planos factoriais é estudar os efeitos individuais dos factores e as relações entre eles, na variável resposta. Este tipo de delineamento consiste em cruzar cada nível de um factor com cada um dos níveis dos outros factores, pelo que também é usual designar-se por classificação cruzada

É usual apresentar um plano fatorial por uma base e um expoente. O expoente indica o número de factores em análise e a base o número de níveis de cada factor. Assim no plano fatorial " p^n "; existem n factores cada um com p níveis.

Se todos os factores tiverem o mesmo número de níveis, os planos designam-se por planos simétricos. Caso contrário designa-se por planos assimétricos.

O exemplo mais simples de um plano fatorial é aquele em que se pretendem estudar dois factores, cada um com apenas dois níveis. Este plano designa-se por plano fatorial 2×2

ou plano factorial 2^2 , pelo que estes planos também são conhecidos por planos factoriais de base dois.

Os planos factoriais têm muitas aplicações em diversas áreas de investigação. Uma das grandes contribuições para o desenvolvimento dos planos factoriais foi feita em 1937 por F.Yates, na Rothamsted Agricultural Experiment Station. Este autor descreve os planos factoriais bem como a sua análise com considerável detalhe. Desde então vários foram os autores que se dedicaram ao seu estudo. Salientam-se os trabalhos de Kempthorne (1947) e mais recentemente Oliveira(2005).

O principal defeito dos experimentos factoriais é que o número de tratamentos aumenta rapidamente. Um ensaio fatorial em blocos casualizados, com grande número de tratamentos, perde bastante sua eficiência, em geral. Para contornar este problema, pode-se usar blocos incompletos e adotar o sistema de confundimento (*confounding*), do qual resultam maiores complicações no planeamento e análise dos experimentos. Outra solução é o uso dos fatoriais fracionários ou incompletos, em que são usadas todas as combinação dos fatores (Pimentel-Gomes. & Garcia, 2002).

6.2 Planos Factoriais de Base 2

Segundo Neter et al. (1996), muito da discussão sobre o desenho de experimentos tem sido focado no planejamento de experimentos confirmatórios. Geralmente estes experimentos utilizam apenas uma pequena quantidade de fatores explicativos. Tais fatores sob investigação são usualmente utilizados a partir da teoria existente ou por experimentação prévia. Por outro lado, experimentos exploratórios são estudos tipicamente encontrados durante os primeiros estágios de uma nova pesquisa, quando pouco é conhecido sobre o conjunto dos principais fatores explicativos. Neste estágio da investigação, o pesquisador necessita considerar um grande número de fatores, de maneira a identificar aqueles que apresentam os mais importantes efeitos. Uma das maneiras de incluir um grande número de fatores em um experimento é estudar cada fator em apenas dois níveis.

Para Montgomery (2001), o primeiro arranjo na série 2^n é aquele que contém somente 2 fatores, A e B, sendo cada um em dois níveis. Este arranjo é chamado arranjo fatorial 2^2 . Os níveis dos fatores pode ser arbitrariamente chamados de “baixo” ou “-“ e “alto” ou “+“ do fator em estudo.

Para um experimento fatorial com 2 fatores, em um delineamento completamente casualizado, o modelo estatístico é dado da seguinte forma:

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk}$$

Onde,

Y_{ijk} = observação aleatória correspondente ao i-ésimo nível do fator A, j-ésimo nível do fator B, e a k-ésima repetição.

μ = constante;

α_i = efeito do i-ésimo nível do fator A; $i=1, \dots, a$

β_j = efeito do j-ésimo nível do fator B; $j=1, \dots, b$

$(\alpha\beta)_{ij}$ = efeito da interação entre o i-ésimo nível do fator A e j-ésimo nível do fator B;

e_{ijk} = erro associado ao i-ésimo nível do fator A, j-ésimo nível do fator B e a k-ésima repetição; $k=1, \dots, r$

O quadro da análise de variância é construído conforme apresenta-se na tabela 6.

Tabela 6. Quadro da ANOVA: Fatorial com 2 fatores fixos

O.V.	G.L.	S.Q.	Q.M.	R.V.
A	a-1	SQA	SQA/a-1	QMA/QMREs
B	b-1	SQB	SQB/b-1	QMB/QMREs
AB	(a-1)(b-1)	SQAB	SQAB/(a-1)(b-1)	QMAB/QMREs
Resíduo	ab(r-1)	Diferença	SQRes/ab(r-1)	
Total	Ab-1	SQtotal		

Fonte: Montgomery, 2004

6.4 Planos Factoriais Fraccionados

Quando há vários fatores de interesse em um experimento, um planejamento fatorial deve ser usado. Em tais experimentos, os fatores variam juntos. Especificamente, por um experimento fatorial queremos dizer que em cada tentativa completa ou replicação do experimento, são investigadas todas as combinações possíveis dos níveis dos fatores. Assim se há dois fatores A e B, com a níveis para o fator A e b níveis para o fator B, então cada replicação contém todas as ab combinações possíveis.

Quando todas as combinações possíveis estão presentes em um experimento, tem-se um experimento fatorial completo, caso contrário, tem-se um experimento fatorial incompleto ou fracionário.

Os planos factoriais fraccionados de base dois têm uma longa história de sucesso no uso de screening e em muitas investigações científicas.

Quando se pretendem estudar n variáveis cada uma com dois níveis, pode escolher-se um plano factorial completo (PFC) para efectuar a experiência, onde todas as combinações dos níveis dos n factores são investigadas. Este tipo de plano designa-se por PF 2^n . À medida que o número de variáveis em estudo aumenta, e um elevado número de observações necessárias para efectuar um PFC pode tornar a experiência demasiado dispendiosa ou mesmo impossível de realizar. Se for possível assumir que as interações de ordem superior podem ser desprezadas, então informações sobre os efeitos principais e sobre interações de ordem inferior podem ser obtidas efectuando apenas uma fracção do PFC. Este tipo de plano factorial designa-se por plano fraccionário (PFF) ou plano com réplicas fraccionadas e foi introduzido por Finney (1945).

Nos experimentos factoriais, o número de tratamentos aumenta rapidamente quando temos muitos fatores em estudo. Por exemplo: $2^6 = 64$ tratamentos, a divisão dos graus de liberdade é dada por: a) 6 graus de liberdade para os efeitos principais; b) 15 graus de liberdade para as interações de 1ª ordem (interações com dois fatores); c) 42 graus de liberdade para interações de 2ª ordem ou de ordem superior.

Se o pesquisador pode assumir que as interações de maior ordem (2ª ordem ou acima) podem ser desprezáveis, então, informações sobre os efeitos principais e interações de ordem menor podem ser obtidas utilizando apenas uma fração do experimento factorial completo.

Na maioria das situações, o princípio da escassez de efeitos se aplica; isto é, o sistema é, usualmente, dominado pelos efeitos principais e pelas interações de ordem inferior. Interações de três ou mais fatores são, em geral, desprezíveis. Portanto, quando o número de fatores é moderadamente grande – digamos, $k \geq 4$ ou 5 – uma prática comum é rodar apenas uma replicação do planejamento 2^k e, então, combinar as interações de ordem maior como uma estimativa do erro.

Fração 1/2 de um delineamento 2^k

Uma fração meio de um planejamento 2^k contém 2^{k-1} rodadas, e é, em geral, chamada planejamento fatorial fracionado 2^{k-1} .

EXEMPLO:

Vamos considerar o experimento fatorial: 2^{3-1} – isto é, a fração um meio do 2^3 (8 tratamentos). Porém, só podemos realizar 4 tratamentos, assim, temos:

$$\frac{1}{2}2^3 = 2^3 2^{-1} = 2^{3-1} = 2^2 = 4$$

A tabela com sinais de + e - para o fatorial 2^3 é dada na tabela 7 a seguir:

Tabela 7. Sinais mais e menos para o planejamento fatorial 2^3

Tratamentos	Efeitos Fatoriais							
	I	A	B	C	AB	AC	BC	ABC
a	+	+	-	-	-	-	+	+
b	+	-	+	-	-	+	-	+
c	+	-	-	+	+	-	-	+
abc	+	+	+	+	+	+	+	+
ab	+	+	+	-	+	-	-	-
ac	+	+	-	+	-	+	-	-
bc	+	-	+	+	-	-	+	-
(1)	+	-	-	-	+	+	+	-

Fonte: Montgomery (2004)

As rodadas dos planejamentos 2^{3-1} resultam em três graus de liberdade associados aos efeitos principais. Pela tabela 7, obtemos as estimativas dos efeitos principais como

$$A = \frac{1}{2} [a - b - c + abc]$$

$$B = \frac{1}{2} [1a + b - c + abc]$$

$$C = \frac{1}{2} [-a - b + c + abc]$$

As estimativas das interações também podem ser obtidas a partir da tabela 7.

$$BC = \frac{1}{2} [a - b - c + abc], \text{ e assim para AC e AB}$$

Assim, a combinação linear de observações na coluna A, l_A dá a estimativa de A + BC. Analogamente, l_B , dá a estimativa de B + AC e l_C , a estimativa de C + AB. Dois ou mais efeitos que tenham essa propriedade são chamados de **alias**. No planejamento 2^{3-1} , A e BC são aliases, bem como B e AC, e C e AB. Os aliases são resultado direto da replicação fracionada.

Uso dos gráficos de probabilidade normal e resíduos em Fatorial Fracionado

O gráfico de probabilidade normal é muito útil na avaliação da significância dos efeitos de um fatorial fracionado, especialmente quando muito efeitos devem ser estimados. Podem-se obter os resíduos de um fatorial fracionado utilizando-se modelos de regressão. Os resíduos devem ser plotados *versus* os valores preditos, *versus* os níveis de fatores, e em papel (gráfico) de probabilidade normal, tanto para avaliar a validade das hipóteses do modelo subjacente como para adquirir compreensão adicional da situação experimental (Montgomery, 2004).

Projeção do planejamento 2^{k-1}

Se um ou mais fatores de uma fração um meio podem ser omitidos, o planejamento se projetará em um planejamento fatorial completo. Se considerarmos que no máximo dois ou três fatores são importantes, o planejamento 2^{3-1} é um excelente planejamento para identificar

os fatores significantes. Experimentos que buscam identificar relativamente poucos fatores significantes em um grupo maior de fatores são chamados **experimentos de varredura**. Essa propriedade de projeção é útil na varredura de fatores porque permite que os fatores desprezíveis possam ser omitidos, resultando em um experimento mais forte nos fatores ativos que permanecem.

Resolução de planejamento

O conceito de resolução de planejamento é uma maneira útil de catalogar planejamentos fatoriais fracionados de acordo com o padrão de aliases que produzem. Eis as definições de cada um destes:

1. **Planejamento de resolução III.** Nestes planejamentos, nenhum efeito principal é alias de qualquer outro efeito p, principal, mas os efeitos principais são aliases das interações de dois fatores, e as interações de dois fatores podem ser aliases uma das outras. O planejamento 2^{3-1} com $I = ABC$ é de resolução III. Usualmente usa-se numeral romano subscrito para indicar a resolução de planejamento; assim a fração um meio é um planejamento 2^{3-1}_{III} .

2. **Planejamento de resolução IV.** Nesses planejamentos, nenhum efeito principal é alias de qualquer outro efeito principal e nem de interações de dois fatores, mas as interações de dois fatores são aliases uma das outras. O planejamento 2^{4-1} com $I = ABCD$ é de resolução 2^{4-1}_{IV} .

3. **Planejamento de resolução 5.** Neste planejamento, nenhum efeito principal ou interação de dois fatores é alias de qualquer outro efeito principal ou interação de dois fatores, mas as interações de dois fatores são aliases das interações de três fatores.

Frações menores: o planejamento fatorial fracionado 2^{k-p}

Embora o planejamento 2^{k-1} seja de grande valor na redução do número de rodadas necessárias para um experimento, com freqüência vemos que frações menores darão quase a mesma quantidade de informação útil a um custo ainda menor. Em geral, um planejamento 2^k pode ser rodado em uma fração $1/2^p$ chamado um planejamento fatorial fracionado 2^{k-p} . Assim uma fração $1/4$ é chamado um planejamento fracionado 2^{k-2} , uma fração $1/8$ é um planejamento 2^{k-3} , uma fração $1/16$ é um planejamento 2^{k-4} , e assim por diante.

EXEMPLO: FRAÇÃO 1/4

Os efeitos principais estão associados com interações de três e cinco fatores, ao passo que interações com dois fatores estão associados com interações de dois fatores ou mais. Portanto, quando estimamos A, na verdade estamos estimando, $A+CDF+ABCDE+BEF$. Se as interações triplas ou de maior ordem são desprezíveis, então este delineamento dá estimativas dos efeitos principais (tabela 8).

Tabela 8. Construção de um fatorial fracionário 2^{6-2} , de resolução IV, com relação definição: $I=ACDF$ $I=BCDE$

Experimentos	Fatorial 2^4 completo				E=BCD	F=ACD	Combinações	Proteínas
	A	B	C	D				
1	-	-	-	-	-	-	(1)	6
2	+	-	-	-	-	+	<i>af</i>	10
3	-	+	-	-	+	-	<i>be</i>	32
4	+	+	-	-	+	+	<i>abef</i>	60
5	-	-	+	-	+	+	<i>cef</i>	4
6	+	-	+	-	+	-	<i>ace</i>	15
7	-	+	+	-	-	+	<i>bcf</i>	26
8	+	+	+	-	-	-	<i>abc</i>	60
9	-	-	-	+	+	+	<i>def</i>	8
10	+	-	-	+	+	-	<i>ade</i>	12
11	-	+	-	+	-	+	<i>bdf</i>	34
12	+	+	-	+	-	-	<i>abd</i>	60
13	-	-	+	+	-	-	<i>cd</i>	16
14	+	-	+	+	-	+	<i>acdf</i>	5
15	-	+	+	+	+	-	<i>bcde</i>	37
16	+	+	+	+	+	+	<i>abcdef</i>	52

6.5. Blocos e Técnica de confundimento no planejamento Fatorial

Segundo Montgomery (2004), é quase impossível rodar todas as observações em um planejamento fatorial 2^k ou 3^k sob condições constantes ou homogêneas. Quando este problema ocorre, a formação de blocos é uma excelente técnica para eliminar a variação indesejada que poderia ser causada pelas condições não homogêneas. Se o planejamento é replicado, e se o bloco tem tamanho suficiente, então uma abordagem é rodar cada réplica em um bloco.

Em algumas situações não é possível rodar uma replicação completa de um experimento fatorial sob condições experimentais homogêneas. Para estas condições a técnica de confundimento pode ser aplicada, que é a técnica de planejamento para um experimento fatorial em blocos, onde o tamanho do bloco é menor que o número de rodadas em uma replicação completa. A técnica faz com que certas interações sejam indistinguíveis ou confundidas com blocos.

Conforme Yates (1937a), o confundimento com três ($3 \times 3 \times 3$) e quatro ($3 \times 3 \times 3 \times 3$) fatores, em três níveis podem ser arranjados em blocos de nove parcelas ou em 9×9 em quadrados latinos, confundindo somente a interação tripla. Estes desenhos são considerados de grande importância prática na agricultura.

No caso específico do fatorial 3^3 , com 27 tratamentos, seria demasiadamente amplo para um ensaio em blocos casualizados. Como alternativa, pode-se utilizar a técnica de confundimento para reduzir o tamanho dos blocos e subdividir os tratamentos de uma maneira balanceada. Yates (1937a) propõe quatro soluções, designadas pelas letras W, X, Y e Z. Em cada uma delas se confundem dois graus de liberdade da interação tripla. Neste caso tem-se oito graus de liberdade para a interação tripla. Estes podem ser divididos em quatro pares ortogonais, onde cada par é dado pelos contrastes das somas e do conjunto das nove parcelas cada. Os quatro grupos são mostrados na tabela 9.

Tabela 9. Arranjo fatorial 3x3x3 confundido com a interação tripla dos grupos W, X, Y e Z..

Combinação do 1º e 2º. fator	Nível do terceiro fator											
	W ₁	W ₂	W ₃	X ₁	X ₂	X ₃	Y ₁	Y ₂	Y ₃	Z ₁	Z ₂	Z ₃
00	0	2	1	0	1	2	0	2	1	0	1	2
10	1	0	2	2	0	1	1	0	2	2	0	1
20	2	1	0	1	2	0	2	1	0	1	2	0
01	2	1	0	1	2	0	1	0	2	2	0	1
11	0	2	1	0	1	2	2	1	0	1	2	0
21	1	0	2	2	0	1	0	2	1	0	1	2
02	1	0	2	2	0	1	2	1	0	1	2	0
12	2	1	0	1	2	0	0	2	1	0	1	2
22	0	2	1	0	1	2	1	0	2	2	0	1

Nota: adaptado de Yates (1937a)

Para aplicação de um plano fatorial confundido com duas repetições por local, a análise de variância segue o esquema da tabela 10.

Tabela 10. Análise da variância do ensaio fatorial 3x3x3 (AxBxC), com confundimento de 2 GL da interação tripla e duas repetições.

Causa da variação	G.L.
A	2
B	2
C	2
AxB	4
AxC	4
BxC	4
Interação AxBxC(não confundida)	6
Tratamentos (não confundidos)	(24)
Blocos	5
Resíduo	24
Total	53

Para aplicação de um plano fatorial 3x3x3 confundido com com uma repetição por local, em ensaios de pouca precisão, a análise de variância seguirá o esquema mostrado na tabela 11.

Tabela 11. Quadro de ANOVA: Fatorial confundido 3x3x3.

Causa da variação	G.L.
A	2
B	2
C	2
AxB	4
AxC	4
BxC	4
(Tratamentos)	(24)
Blocos	2
Resíduo=Interação AxBxC(não confundida)	6
Total	26

CAPÍTULO 7

Técnicas de Experimentação Estatística: uma aplicação a ensaios florestais na Amazónia brasileira

7.1. Introdução

A Amazônia apresenta grande potencial para plantios florestais homogêneos, mistos ou sistemas agroflorestais, em áreas já alteradas pelas atividades de agricultura e pecuária. O desenvolvimento de tecnologias para estabelecimento de plantios ordenados, de forma ecológica e economicamente embasada em princípios sustentáveis, trarão como consequência benéfica a redução sobre a pressão do desmatamento da floresta tropical nativa. Além disso, para sistemas já estabelecidos, pode-se buscar alternativas de aumento da produtividade, o que elevará a renda dos agricultores, pela comercialização de produtos madeireiros e não-madeireiros.

A espécie florestal na qual foram aplicados os tratamentos e as análises estatísticas foi o pau-rosa (*Aniba rosaedora* Ducke).

O objetivo metodológico deste ensaio foi estudar a aplicação das técnicas fatoriais em ensaios de adubação, visando viabilizar a aplicação desta técnica estatística, com a redução do número de parcelas aplicadas por tratamento e consequentemente reduzindo o tamanho da área experimental; os recursos materiais necessários; e finalmente o custo total do experimento.

Como resultado metodológico espera-se definir o método estatístico adequado ao planejamento de experimento com vários fatores e níveis para o pau-rosa nas condições de clima e solo da região Amazônica. Como resultado prático espera-se estimar a natureza da resposta à aplicação dos nutrientes nitrogênio (N), fósforo (P) e potássio (K), incluindo a interação.

7.2 Descrição do problema

A exploração do pau-rosa remonta ao início do século passado, como principal fonte de fornecimento do linalol para indústria de perfumaria. O processo extrativista, sem a devida reposição do estoque de indivíduos, fez com que a espécie fosse colocada na lista de espécies em extinção da flora brasileira. Um dos agravantes na conservação da espécie é a condição dióica da espécie e a baixa frequência dos indivíduos em populações naturais. Plantações homogêneas que concilia aspectos preservacionistas e econômicos tem tido pouco êxito, em função do desconhecimento do manejo silvicultural da espécie.

Um dos agravantes para o crescimento de espécies florestais em taxas satisfatórias, quando plantadas em espaçamentos adensados é o baixo nível de estoque de nutrientes dos solos da região. A maioria dos solos que ocorrem na Amazônia, são antigos, profundos, altamente lixiviáveis e com baixa fertilidade. Uma das possíveis soluções para tornar estes projetos mais atrativos sob o ponto de vista econômico é incrementar a taxa de crescimento das espécies com a aplicação de fertilizantes.

Existem poucos estudos sobre a espécie relacionados ao seu crescimento na forma de reflorestamentos homogêneos. É necessário estudar quais são os fatores edáficos e climáticos que contribuem para a otimização do plantio de pau-rosa, qual o sistema de produção adequado, definindo-se os melhores espaçamentos e os nutrientes e doses de adubo adequados.

Em razão do pouco conhecimento sobre o pau-rosa em condições de plantios homogêneos, estudos exploratórios fazem-se necessários para identificar quais os principais fatores de clima ou de solo que são condicionantes para o crescimento satisfatório da espécie. Com este objetivo as técnicas fatoriais podem ser aplicadas com bastante propriedade, e em muitos casos, com grandes vantagens sobre os experimentos simples, que estudam apenas um fator de cada vez.

7.3 Material e Métodos

Este ensaio foi instalado na região de Maués-AM, na área da empresa Magaldi Agrocomercial e Industrial Ltda (Figura 2).

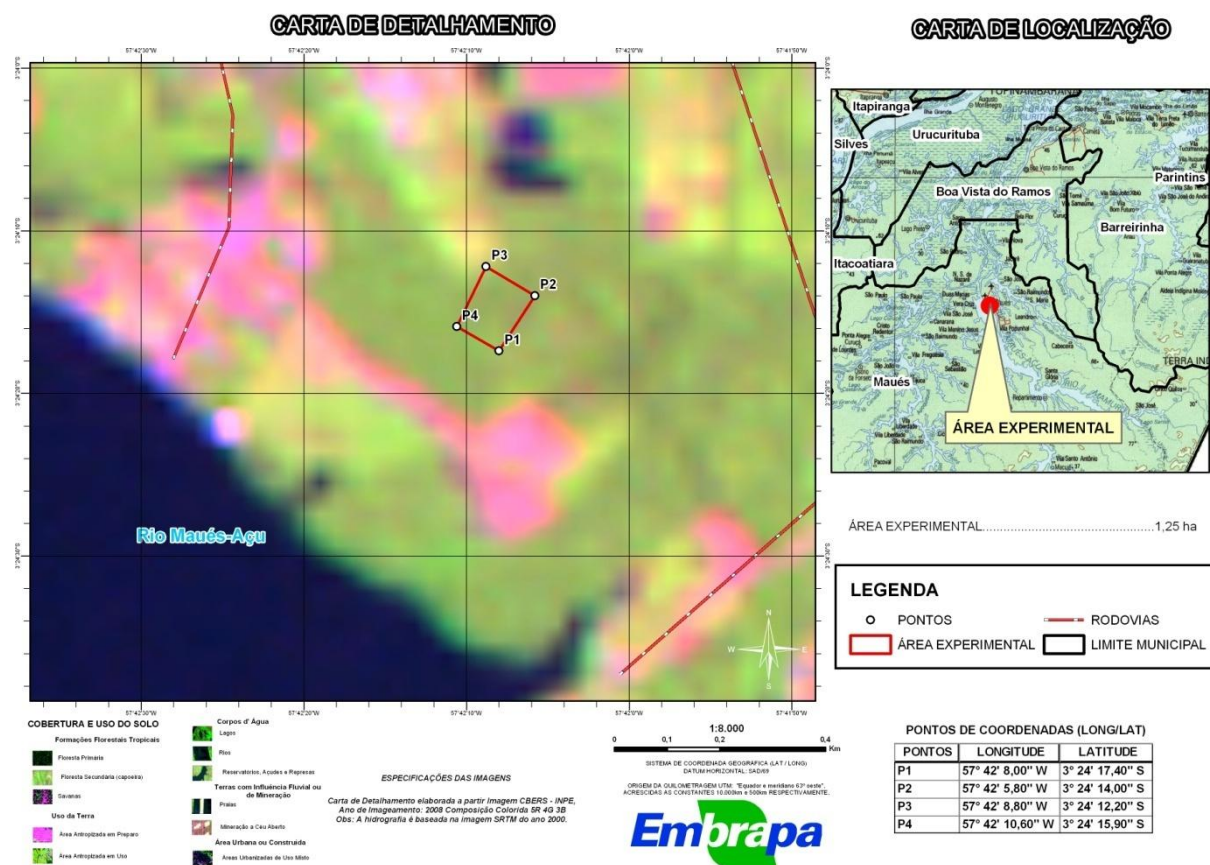


Figura 2. Localização do experimento no município de Maués-AM, com área de 1,5 hectares..

O **objetivo** do experimento é avaliar a resposta em crescimento da espécie florestal-pau-rosa (*Aniba roseodora*), à aplicação de 3 doses (níveis) de 3 nutrientes (fatores) – nitrogênio (N), fósforo (P) e potássio (K).

O **delineamento** dos tratamentos corresponde a um experimento em blocos ao acaso em arranjo fatorial 3 x 3 x 3 com confundimento de dois graus de liberdade da interação tripla, conforme designado por Yates (1937) como grupo W (Tabela 12):

Tabela 12. Arranjo fatorial 3x3x3 com interação em confundimento.

Bloco 1	Bloco 2	Bloco 3
000	001	002
012	010	011
021	022	020
101	102	100
110	111	112
122	120	121
202	200	201
211	212	210
220	221	222

Nota: Adaptado de Yates(1937a)

Fatores: N, P, K

Níveis: N 0 – 0 kg/hectare de N (sob a forma Sulfato Amônio – 20% de N)

1 – 60 kg/hectare de N

2 – 120 kg/hectare de N

Níveis: P 0 – 0 kg/hectare de P_2O_5 (sob a forma SFT – 45% de P_2O_5)

1 – 80 kg/hectare de P_2O_5

2 – 160 kg/hectare de P_2O_5

Níveis: K 0 – 0 kg/hectare de K (sob a forma KCl – 60% de K)

1 – 50 kg/hectare de K

2 – 100 kg/hectare de K

Parcela experimental: foram utilizadas unidades experimentais constituídas por 25 plantas úteis, com bordadura simples, cada uma com 27 tratamentos (combinação dos 3 tipos de adubo com os 3 níveis de adubação). O espaçamento foi de 2,0 x 2,5 metros, conforme figura 3.

0	0	0	0	0	0	0
---	---	---	---	---	---	---

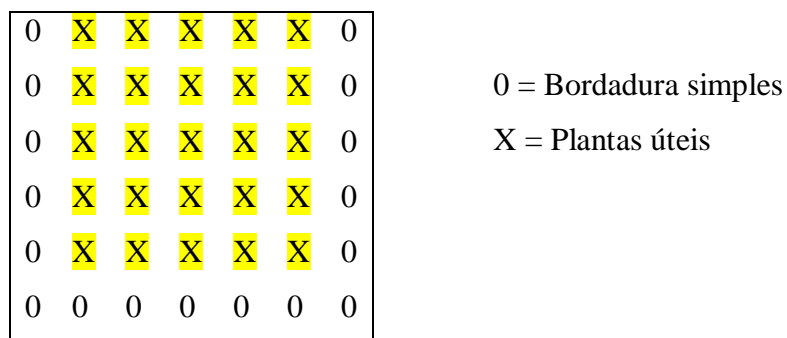


Figura 3. Detalhe da parcela experimental

O fósforo (P) foi aplicado na cova à época do plantio (31/03/2009). O potássio (K) e o nitrogênio (N) foram aplicados em adubação de cobertura, aproximadamente 3 meses depois do plantio (16-17/07/2009), quando a espécie já havia estabelecido seu sistema radicular com melhor capacidade para absorção dos nutrientes.

Para corrigir o pH do solo, aplicou-se uma dosagem de calcário dolomítico na quantidade de 1 ton/ha (500 g/planta).

Na figura 4 apresenta-se o arranjo das parcelas/tratamentos no campo:

REPETIÇÃO 1		
Bloco 1	Bloco 2	Bloco 3
(NPK)	(NPK)	(NPK)
000	001	002
012	010	011
021	022	020
101	102	100
110	111	112
122	120	121
202	200	201
211	212	210
220	221	222

REPETIÇÃO 2		
Bloco 1	Bloco 2	Bloco 3
(NPK)	(NPK)	(NPK)
000	001	002
012	010	011
021	022	020
101	102	100
110	111	112
122	120	121
202	200	201
211	212	210
220	221	222

Figura 4. Distribuição das parcelas/tratamentos no campo.

Informações técnicas

:

Área total da parcela: 49 PL. X 5,0 m² = 245 m²

Área total do bloco: 245 X 9 = 2.205 m²

Área total do experimento: 2205 X 6 = 13230 m² = 1,3230 ha = 1,5 ha

Número plantas úteis parcela: 25

Número plantas total parcela: 49

Número plantas total por bloco: 9 trat X 49 = 441 pl.

Número plantas total por experimento: 441 X 6 = 2646 pl.

O **modelo estatístico** utilizado para a análise de variância é mostrado na equação

1:

$$y_{ijkl} = \mu + \rho_l + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \varepsilon_{ijkl} \quad (1)$$

em que:

$$i = 1, \dots, I$$

$$j = 1, \dots, J$$

$$k = 1, \dots, K$$

$$l = 1, \dots, r$$

e

y_{ijkl} é a observação referente ao efeito do i -ésimo nível do fator α , j -ésimo nível do fator β , k -ésimo nível do fator γ , efeito da interação entre o i -ésimo nível do fator α e j -ésimo nível do fator β , efeito da interação entre $(\alpha\gamma)_{ik}$, $(\beta\gamma)_{jk}$ e $(\alpha\beta\gamma)_{ijk}$, efeito do l -ésimo bloco ρ_l , uma constante μ e ao erro ε_{ijkl} . Os fatores α , β e γ , correspondem aos nutrientes N, P e K, respectivamente.

Inicialmente foram tomadas como variável resposta y_{ijkl} duas observações alométricas das árvores: altura total (alt) e o diâmetro do colo (dc)

7.4 Resultados e Discussão

Foram estudados 27 tratamentos, correspondente às 27 interações entre os 3 níveis (0, 1 e 2) dos fatores N, P e K.

Os resultados foram obtidos e analisados para as variáveis de crescimento: altura total (alt), e diâmetro do colo (dc) aos 12 meses após o plantio. O diâmetro do colo, corresponde a medição do diâmetro à altura de 5cm a partir do solo.

7.4.1. Estudo da altura total

O estudo da altura total do pau-rosa foi realizado sob as seguintes hipóteses:

$$H_0: \mu_{000} = \mu_{012} = \mu_{220} \dots = \mu_{222}$$

H_1 : pelo menos uma média μ_{ijk} é \neq das demais

Os resultados com as médias das observações para a altura são apresentadas na tabela 13.

Tabela 13. Médias da altura total do pau-rosa, em metros, aos 12 meses de idade do experimento factorial de 3^3 de adubação com N, P e K, com duas repetições em Maués, AM, Brasil. Manaus, 2010.

Trat (NPK)	Rep1 Bloco1	Rep2 Bloco4	Trat (NPK)	Rep1 Bloco2	Rep2 Bloco5	Trat (NPK)	Rep1 Bloco3	Rep2 Bloco6
000	115,9	102,7	001	85,4	74,7	002	69,0	84,6
012	58,5	81,5	010	78,9	66,8	011	67,0	85,6
021	84,0	80,0	022	62,7	41,0	020	77,9	83,8
101	60,2	76,4	102	68,7	72,7	100	78,5	NA
110	91,1	78,0	111	85,6	75,0	112	78,5	NA
122	83,8	NA	120	72,5	94,0	121	73,4	NA
202	82,7	NA	200	66,0	NA	201	63,0	NA
211	45,0	NA	212	60,0	NA	210	101,0	NA
220	62,7	NA	221	NA	NA	222	NA	NA

Fonte: dados do autor.

Nota: NA = dados não disponíveis, em função da perda total da parcela.

Analisando pelo R:

```
rm(list=ls())
```

```

paurosa <-
read.table("C:\\Users\\roberval\\Documents\\mestrado_estadistica_portugal\\Tese-
UAB\\dados-paurosa\\ paurosa4-ago_2010_final.txt", head=T)

Paurosa
attach(paurosa)
names(paurosa)
sapply(paurosa, is.numeric)
paurosa$bloco <- as.factor(paurosa$bloco)
paurosa$Nitro <- as.factor(paurosa$N)
paurosa$Phos <- as.factor(paurosa$P)
paurosa$Potas <- as.factor(paurosa$K)
attach(paurosa)

# confundindo o resíduo com a interação NxPxK
modelo2.fat <- lm(prod ~ bloco + Nitro*Phos*Potas - Nitro:Phos:Potas)
anova(modelo2.fat)

# análise das interações
modelo2.fat <- aov(alt ~ N/P) # Phos dentro de cada Nitro
anova(modelo2.fat)

summary(modelo2.fat, split=list("N:P"=list(N0=c(1,4), N1=c(2,5), N2=c(3,6))))
modelo3.fat <- aov(alt ~ P/N) #Nitro dentro de cada Phos
anova(modelo3.fat)

summary(modelo3.fat, split=list("P:N"=list(P0=c(1,4), P1=c(2,5), P2=c(3,6))))

#Realiza as comparações múltiplas LSD.
gl<-df.residual(modelo) # extrai os graus de liberdade do erro
erro<-deviance(modelo)/gl #extrai o QM (erro)
compara.NP<-LSD.test(alt, N:P, gl, erro, group=TRUE) # para N dentro de P
compara.PN<-LSD.test(alt,P:N,gl,erro, group=TRUE) # para P dentro de N
df<-df.residual(modelo)

MSerror<-deviance(modelo)/df

comparison.NP <- HSD.test(alt,N:P,df,MSerror, group=TRUE) #para N dentro de P
comparison.PN <- HSD.test(alt,P:N,df,MSerror, group=TRUE) #para N dentro de P

```

O Resultado da análise de variância, utilizando-se o modelo completo menos a interação tripla é apresentado na tabela 14. Na tabela 11 do subcapítulo 6.5, é mostrado o quadro da ANOVA com confundimento de 2 gl da interação tripla.

Tabela 14. Análise de variância com o efeito da interação tripla confundido com o resíduo, usando o modelo: altura = bloco + Nitro*Phos*Potas - Nitro:Phos:Potas + Erro. Manaus, 2010

	Df	Sum Sq	Mean Sq	F	value Pr(>F) 1
bloco	5	768,5	153,71	0,7039	0,62716
Nitro	2	616,6	308,30	1,4118	0,26697
Phos	2	801,4	400,71	1,8350	0,18549
Potas	2	1824,2	912,08	4,1767	0,03050 *
Nitro:Phos	4	2195,9	548,98	2,5140	0,07397 .
Nitro:Potas	4	696,8	174,20	0,7977	0,54069
Phos:Potas	4	432,2	108,05	0,4948	0,73969
Residuals	20	4367,4	218,37		

¹ Código signif.: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1

Observando os resultados da tabela 14, nota-se o efeito significativo para o potássio (K), com significância de 3,05% e para a interação Nitrogênio(N) vs. Fósforo(P), com significância de 7,4%.

Desprezando-se os efeitos das interações não significativas, foi realizada uma nova análise, utilizando-se o modelo parcial. Este resultado é apresentado na Tabela 15.

Tabela 15. Análise de variância com o efeito da interação tripla confundido com o resíduo, usando o modelo: altura = bloco + Nitro*Phos*Potas - Nitro:Phos:Potas - Nitro:Potas - Phos:Potas + Erro. Manaus, 2010

	Df	Sum Sq	MeanSq	F value	Pr(>F) ¹
bloco	5	768,5	153,71	0,7830	0,57051
Nitro	2	616,6	308,30	1,5705	0,22571
Phos	2	801,4	400,71	2,0413	0,14874
Potas	2	1824,2	912,08	4,6463	0,01809 *
Nitro:Phos	4	2195,9	548,98	2,7966	0,04518 *
Residuals	28	5496,4	196,30		

¹ Código signif.: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1

Consideramos sob o ponto de vista agrônomico que os efeitos significativos do Potássio (probabilidade 1,8%) e da interação (probabilidade de 4,5%) não devem ser desprezados. Procedeu-se, a análise do efeito da interação N x P e o desdobramento dos graus de liberdade da Interação para se verificar o efeito do Fósforo (P) dentro dos níveis de Nitrogênio (N), conforme o seguinte script no R (quadro 1).

Quadro 1. Script para análise utilizando o R.

```

modelo2.fat <- aov(alt ~ bloco + Nitro/Phos) # Phos dentro de cada Nitro
anova(modelo2.fat)

summary(modelo2.fat, split=list("Nitro:Phos"=list(N0=c(1,4), N1=c(2,5),
N2=c(3,6))))

modelo3.fat <- aov(alt ~ bloco + Phos/Nitro) # Nitro dentro de cada Phos
anova(modelo2.fat)

summary(modelo3.fat, split=list("Phos:Nitro"=list(N0=c(1,4), N1=c(2,5),
N2=c(3,6))))

```

Na análise dos desdobramentos, pode-se estruturar os seguintes:

1º. Desdobramento:	G.L.
Níveis de N dentro de P ₀ :	2
Níveis de N dentro de P ₁ :	2
Níveis de N dentro de P ₂ :	2
2º. Desdobramento:	G.L.
Níveis de P dentro de N ₀ :	2
Níveis de P dentro de N ₁ :	2
Níveis de P dentro de N ₂ :	2

Os resultados obtidos com os procedimentos da análise dos desdobramentos são mostrado nas tabelas 16 e 17.

Tabela 16. Análise de variância desdobrando o efeito dos níveis do fator N dentro do fator P. Manaus, 2010

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
bloco	5	768,5	153,71	0,6490	0,66441	
Nitro	2	616,6	308,30	1,3017	0,28699	
Nitro:Phos	6	3212,8	535,46	2,2609	0,06425 .	
Nitro:Phos: N0	2	1083,5	541,77	2,2875	0,11896	
Nitro:Phos: N1	2	385,1	192,56	0,8130	0,45304	
Nitro:Phos: N2	2	1744,1	872,06	3,6821	0,03715 *	
Residuals	30	7105,2	236,84			
¹ Código signif.:	0	****' 0,001	**'0,01	*'0,05	'0,1	' ' 1

O resultado do estudo do desdobramento de N em P, apresentado na tabela 19, mostra que o fator Fósforo (P) só foi significativo no nível 2 de N. O nível 2 de N corresponde a dosagem de 120 kg/hectare de sulfato de amônio, na presença do fósforo

Tabela 17. Análise de variância desdobrando o efeito dos níveis do fator P dentro do fator N. Manaus, 2010

	Df	Sum Sq	Mean Sq	F value	Pr(>F) ¹	
bloco	5	768,5	153,71	0,6490	0,66441	
Phos	2	827,1	413,53	1,7460	0,19173	
Phos:Nitro	6	3002,3	500,38	2,1128	0,08115 .	
Phos:Nitro: P0	2	844,7	422,33	1,7832	0,18546	
Phos:Nitro: P1	2	192,3	96,15	0,4060	0,66994	
Phos:Nitro: P2	2	1965,4	982,68	4,1491	0,02565 *	
Residuals	30	7105,2	236,84			
¹ Código signif.:	0	****' 0,001	**'0,01	*'0,05	'0,1	' ' 1

O resultado do desdobramento de P em N, apresentado na tabela 20, mostra que o fator Nitrogênio (N) só foi significativo no nível 2 de P. O nível 2 de P corresponde a dosagem de 160 kg/hectare de P₂O₅.

Na continuidade das análises realizou-se o teste de comparação múltiplas da diferença mínima significativa (DMS). Este resultado é apresentado no tabela 18.

Tabela 18. Comparação de médias para variável altura, do tratamento Nitrogênio (N) dentro dos níveis de Fósforo(P) pelo teste da Diferença Mínima Significativa (LSD). Manaus, 2010.

Tratamentos (N:P)	Altura (cm)	Grupos ¹
0:0	88,7	a
1:1	81,6	a
1:2	80,9	a
2:1	79,0	a
2:0	76,1	a
0:1	73,2	a
0:2	71,6	ab
1:0	71,3	ab
2:2	53,9	b

¹Médias seguidas pela mesma letra não diferem pelo teste LSD, com 5% de significância.

Observando-se os resultados do desdobramento de N em P, verifica-se uma interação negativa para a aplicação de 120 kg de N/hectare e mais 160 kg de P, implicando no menor crescimento médio em altura do pau-rosa aos 12 meses de idade.

Para comparar as análises realizadas no R com o SAS, procedeu-se a mesma análise utilizando o programa SAS v. 9.1.2 (Quadro 2).

Quadro 2. Script utilizado para Análise pelo SAS

```
proc glm data = work.paurosa;
class bloco N P K;
model alt12=bloco N P K N*P N*K P*K /SS1;
LSMEANS N*P/SLICE=N;
LSMEANS N*P/SLICE=P;
RUN;
```

Os resultados obtidos utilizando-se o programa SAS é mostrado no Quadro 3.

Quadro 3. Resultado das análises utilizando o programa SAS

Variable dependiente: alt12						
	Fuente	DF	Suma de cuadrados	Cuadrado de la media	F-Valor	Pr
> F	Modelo	23	7335.63792	318.94078	1.46	
0.1978	Error	20	4367.43117	218.37156		
	Total correcto	43	11703.06909			
	R-cuadrado		Coef Var	Raiz MSE	alt12 Media	
	0.626813		19.55859	14.77740	75.55455	
	Fuente	DF	Tipo I SS	Cuadrado de la media	F-Valor	Pr
> F	bloco	5	768.532166	153.706433	0.70	
0.6272	N	2	616.591346	308.295673	1.41	
0.2670	P	2	801.425750	400.712875	1.84	
0.1855	K	2	1824.156508	912.078254	4.18	
0.0305	N*P	4	2195.932145	548.983036	2.51	
0.0740	N*K	4	696.811687	174.202922	0.80	
0.5407	P*K	4	432.188313	108.047078	0.49	
0.7397						
		N	P	alt12 LSMEAN		
		0	0	88.7166667		
		0	1	73.0500000		
		0	2	71.5666667		
		1	0	73.1597052		
		1	1	82.5367581		
		1	2	83.6469406		
		2	0	75.9259665		
		2	1	79.4680250		
		2	2	48.8243596		
	N*P Efecto dividido por N for alt12					
		Suma de	Cuadrado de			
	N	DF	cuadrados	la media	F-Valor	Pr > F
	0	2	1083.534444	541.767222	2.48	0.1090
	1	2	287.298955	143.649477	0.66	0.5288
	2	2	1888.753909	944.376954	4.32	0.0275
	N*P Efecto dividido por P for alt12					
		Suma de	Cuadrado de			
	P	DF	cuadrados	la media	F-Valor	Pr > F
	0	2	720.628205	360.314102	1.65	0.2171
	1	2	240.042051	120.021025	0.55	0.5856
	2	2	2189.061410	1094.530705	5.01	0.0172

Observando-se os resultados obtidos com as análises realizadas nos programas R e SAS , verifica-se uma perfeita concordância nos resultados. Ademais, é importante evidenciar

nas análises realizadas com os programas SAS e R, a flexibilidade e a adaptabilidade do R na elaboração das rotinas de acordo com a necessidade do usuário. Esta constatação é principalmente importante para validar os resultados obtidos pelo programa R, face a sua natureza de elaboração dos scripts e para comprovação da robustez do R na resolução deste tipo de análise.

Análise dos resíduos

A análise do gráfico dos resíduos versus cada um dos fatores N, P e k mostrado na figura 5, não indica qualquer problema potencial para o modelo.

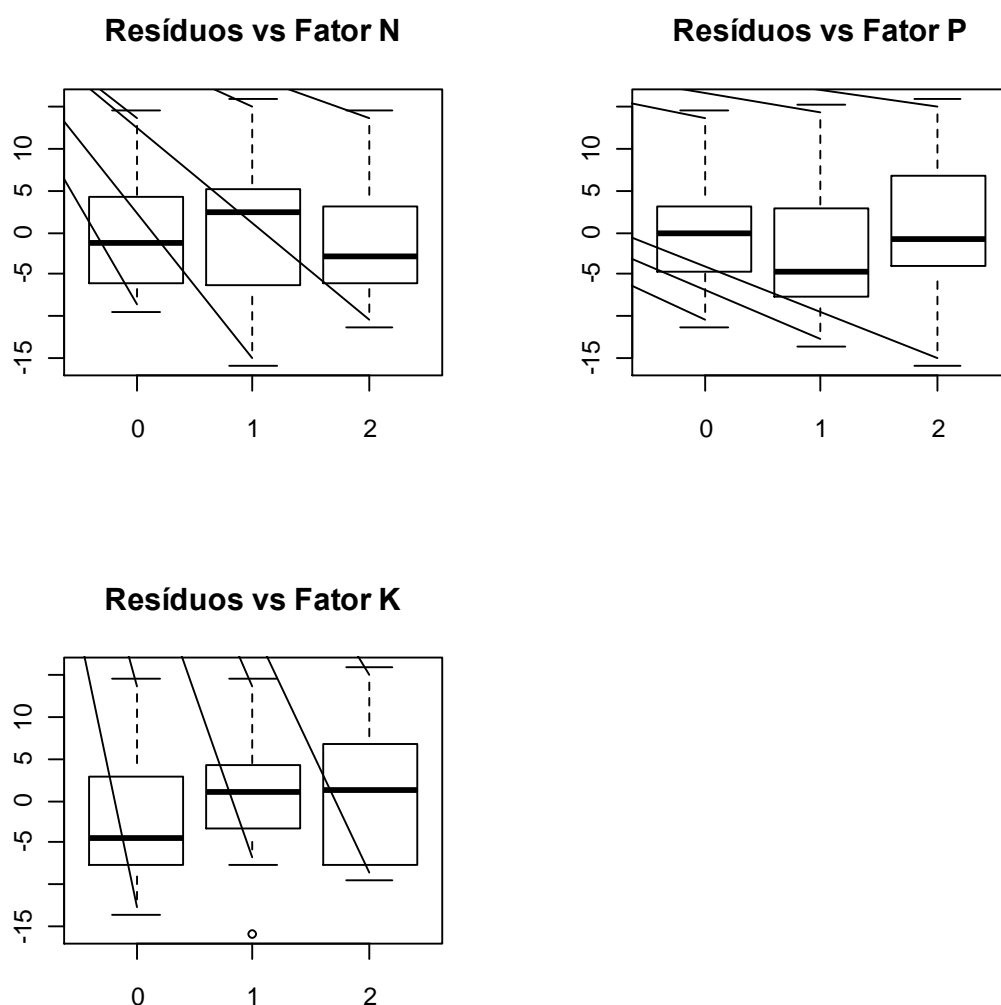


Figura 5. Gráfico dos resíduos versus os fatores N, P e K.

A análise do gráfico dos resíduos *versus* valores preditos e o gráfico de probabilidade normal dos resíduos mostrados na figura 6, indicam alguns pequenos problemas com a hipótese de normalidade, pois as caudas do gráfico normal de probabilidades não se situam

exatamente ao longo da reta central. Contudo, analisa-se que não seja de todo prejudicial para invalidar a ANOVA realizada.

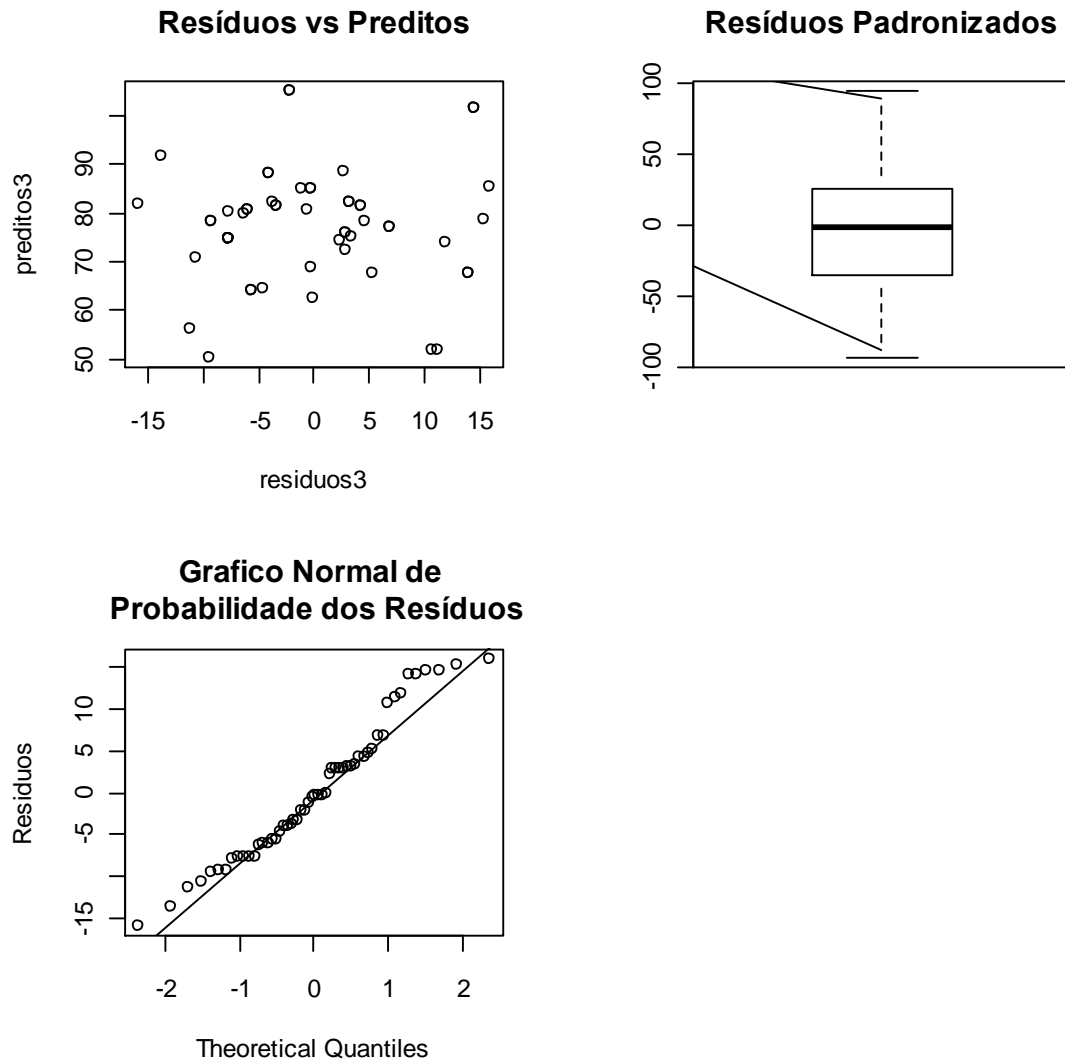


Figura 6. Gráficos dos Resíduos versus preditos; Resíduos padronizados e da probabilidade Normal dos resíduos

Pelos resultados da análise dos resíduos apresentados nas figuras 5 e 6, verifica-se que não houve violação de nenhum dos pressupostos da normalidade, homocedasticidade e da independência.

7.4.1.1 Análise do ensaio fatorial pelo método da Regressão Múltipla.

O ensaio de adubação em fatorial analisado pelo método da regressão é apresentado no quadro 7. O principal objetivo desta análise é obter uma estimativa média das respostas linear e quadrática dos fatores N, P e K. Os resultados mostrados no quadro 7 indicaram o efeito linear significativo para o P e K. Não ficou evidente na análise a resposta quadrática para qualquer um dos fatores.

Quadro7. Resultados obtidos pelo programa R

```

Analysis of Variance Table

Response: alt12.cm

      Df Sum Sq Mean Sq F value    Pr(>F)
bloco      1      4.3      4.28  0.0171 0.896756
paurosa$N  1  435.4  435.45  1.7400 0.196224
paurosa$P  1  755.0  755.02  3.0169 0.091727 .
paurosa$K  1 1975.3 1975.30  7.8929 0.008278 **
I (paurosa$P^2)  1   97.3   97.31  0.3888 0.537191
I (paurosa$N^2)  1  120.5  120.45  0.4813 0.492684
I (paurosa$K^2)  1   11.2   11.22  0.0449 0.833579
paurosa$N:paurosa$P  1    3.3    3.35  0.0134 0.908597
paurosa$N:paurosa$K  1   35.9   35.85  0.1433 0.707481
paurosa$P:paurosa$K  1    6.2    6.18  0.0247 0.876121
Residuals    33 8258.6  250.26
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Call:
lm(formula = alt12.cm ~ bloco + paurosa$N + paurosa$P + paurosa$K
+
  paurosa$N:paurosa$P + paurosa$N:paurosa$K +
paurosa$P:paurosa$K +
  I (paurosa$P^2) + I (paurosa$N^2) + I (paurosa$K^2))

Residuals:
    Min       1Q   Median       3Q      Max
-25.074  -9.330  -3.364   9.751  40.217

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   91.69067   10.00081    9.168 1.36e-10 ***
bloco         -0.09694    1.49897   -0.065  0.949
paurosa$N     1.31762   11.81283    0.112  0.912
paurosa$P     1.52167   11.73045    0.130  0.898
paurosa$K    -11.20913   11.73779   -0.955  0.347
I (paurosa$P^2) -3.14044    5.07560   -0.619  0.540
I (paurosa$N^2) -3.42426    5.22724   -0.655  0.517
I (paurosa$K^2)  1.14956    5.08292    0.226  0.822
paurosa$N:paurosa$P -0.47876    3.63607   -0.132  0.896
paurosa$N:paurosa$K  1.37288    3.63880    0.377  0.708
paurosa$P:paurosa$K -0.56800    3.61549   -0.157  0.876
---

```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 15.82 on 33 degrees of freedom
(10 observations deleted due to missingness)
Multiple R-squared:  0.2943, Adjusted R-squared:  0.08047
F-statistic: 1.376 on 10 and 33 DF,  p-value: 0.2339
```

Estudo gráfico das Interações

As figuras 7, 8 e 9 mostram o comportamento das médias dos fatores N, P e K *versus* os níveis de adubação para cada um destes fatores.

Os gráficos K vs.P e K vs. N apresentam uma forte tendência ao paralelismo das retas, evidenciando a ausência de interação entre os fatores, conforme os resultados analíticos apresentados anteriormente.

A presença de interação é mais evidente no comportamento do gráfico P vs. N conforme se observa na figura 8, um padrão de não paralelismo entre as retas.

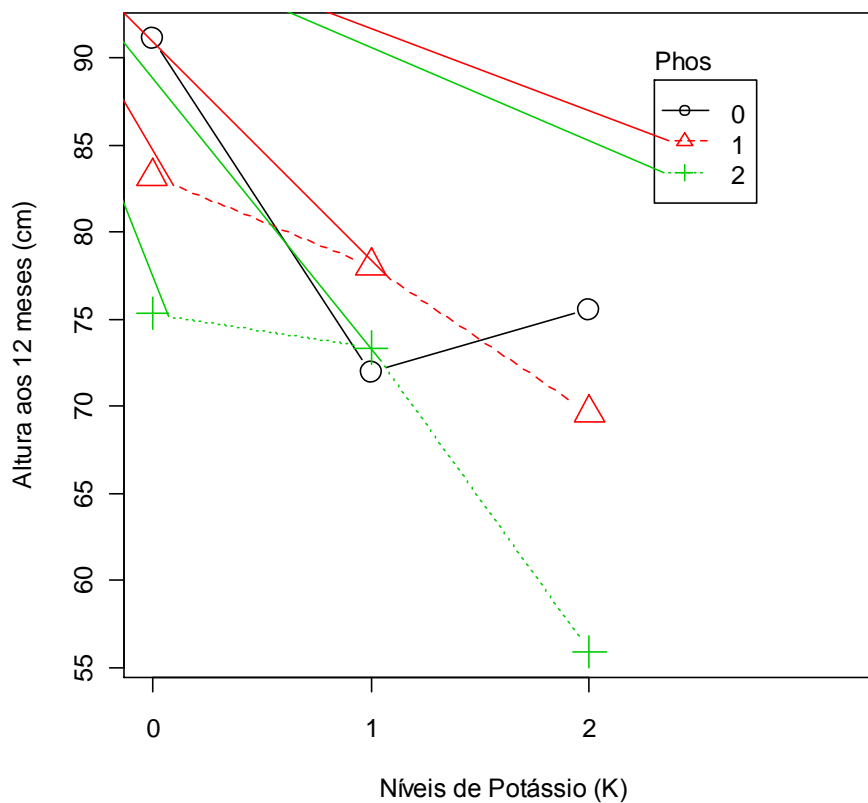


Figura 7. Crescimento em altura do pau-rosa, aos 12 meses de idade, em função dos níveis de potássio (0=0 kg/ha; 1=50 kg/ha; 2=100 kg/ha) e Fósforo (0=0 kg/ha; 1=80 kg/ha; 2=160 kg/ha). Manaus, 2010.

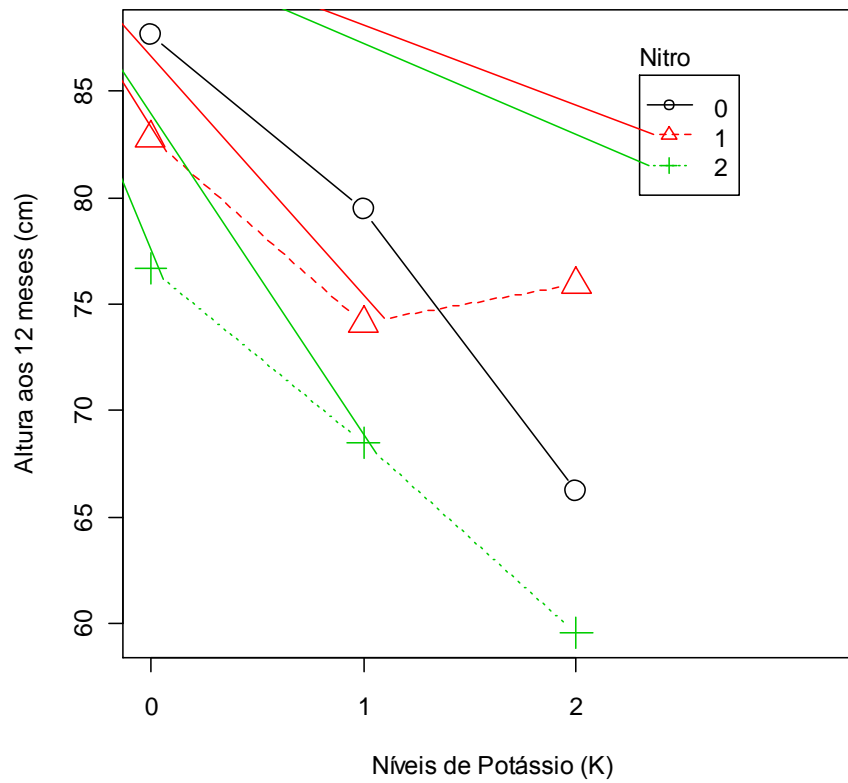


Figura 8. Crescimento em altura do pau-rosa, aos 12 meses de idade, em função dos níveis de potássio (0=0 kg/ha; 1=50kg/ha; 2=100kg/ha) e Fósforo (0=0kg/ha; 1=80kg/ha; 2=160 kg/ha). Manaus, 2010.

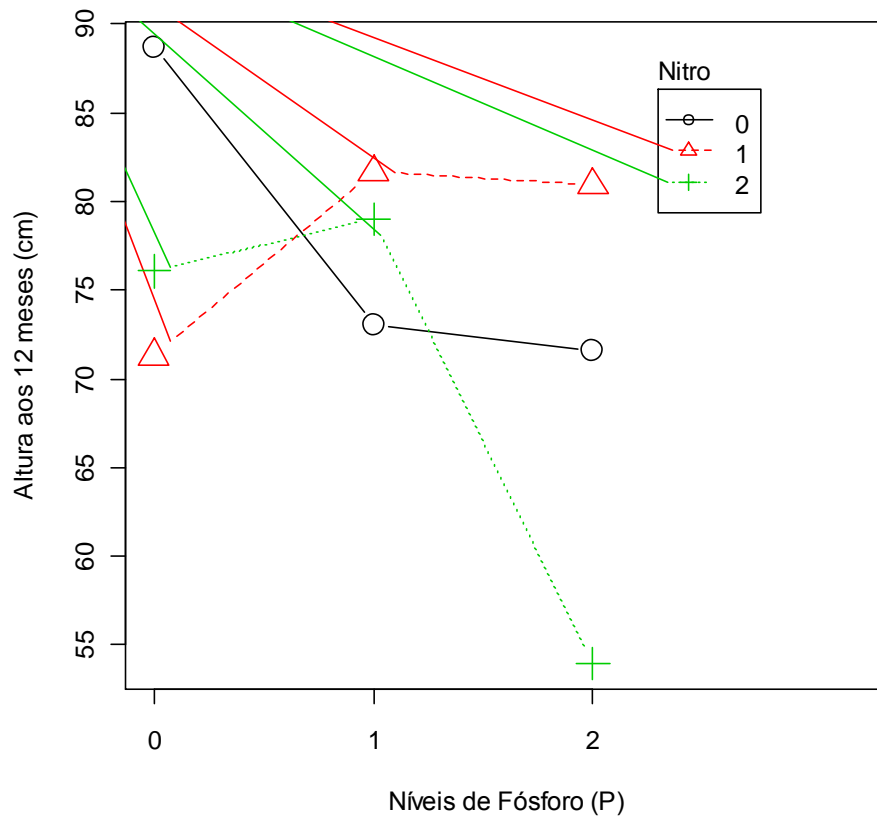


Figura 9. Crescimento em altura do pau-rosa, aos 12 meses de idade, em função dos níveis de Fósforo (0=0 kg/ha; 1=80 kg/ha; 2=160 kg/ha) e Nitrogênio (0=0 kg/ha; 1=60 kg/ha; 2=120 kg/ha). Manaus, 2010.

7.4.2 Estudo do diâmetro do colo (dc)

O estudo do diâmetro de colo das plantas de pau-rosa foi realizado sob as seguintes hipóteses:

$$H_0: \mu_{000} = \mu_{012} = \mu_{220} \dots = \mu_{222}$$

H_1 : pelo menos uma média μ_{ijk} é \neq das demais

Os resultados para a variável diâmetro do colo, medidos em dm, na altura de 10 cm a partir do solo, são apresentados na tabela 19.

Tabela 19. Médias do diâmetro do colo das plantas de pau-rosa, em mm, aos 12 meses de idade do experimento factorial de 3^3 de adubação com N, P e K, com duas repetições em Maués, AM, Brasil. Manaus, 2010.

O.V.	Df	Sum Sq	Mean Sq	F	value Pr(>F)
bloco	5	34,496	6,8991	0,5870	0,7099
Nitro	2	1,959	0,9797	0,0834	0,9204
Phos	2	20,321	10,1605	0,8645	0,4372
Potas	2	32,596	16,2980	1,3867	0,2741
Nitro:Phos	4	70,102	17,5255	1,4911	0,2444
Nitro:Potas	4	22,606	5,6516	0,4808	0,7495
Phos:Potas	4	2,873	0,7184	0,0611	0,9925
Residuals	19	223,313	11,7533		

Na figura 10 apresenta-se os gráficos correspondentes à análise de resíduos.

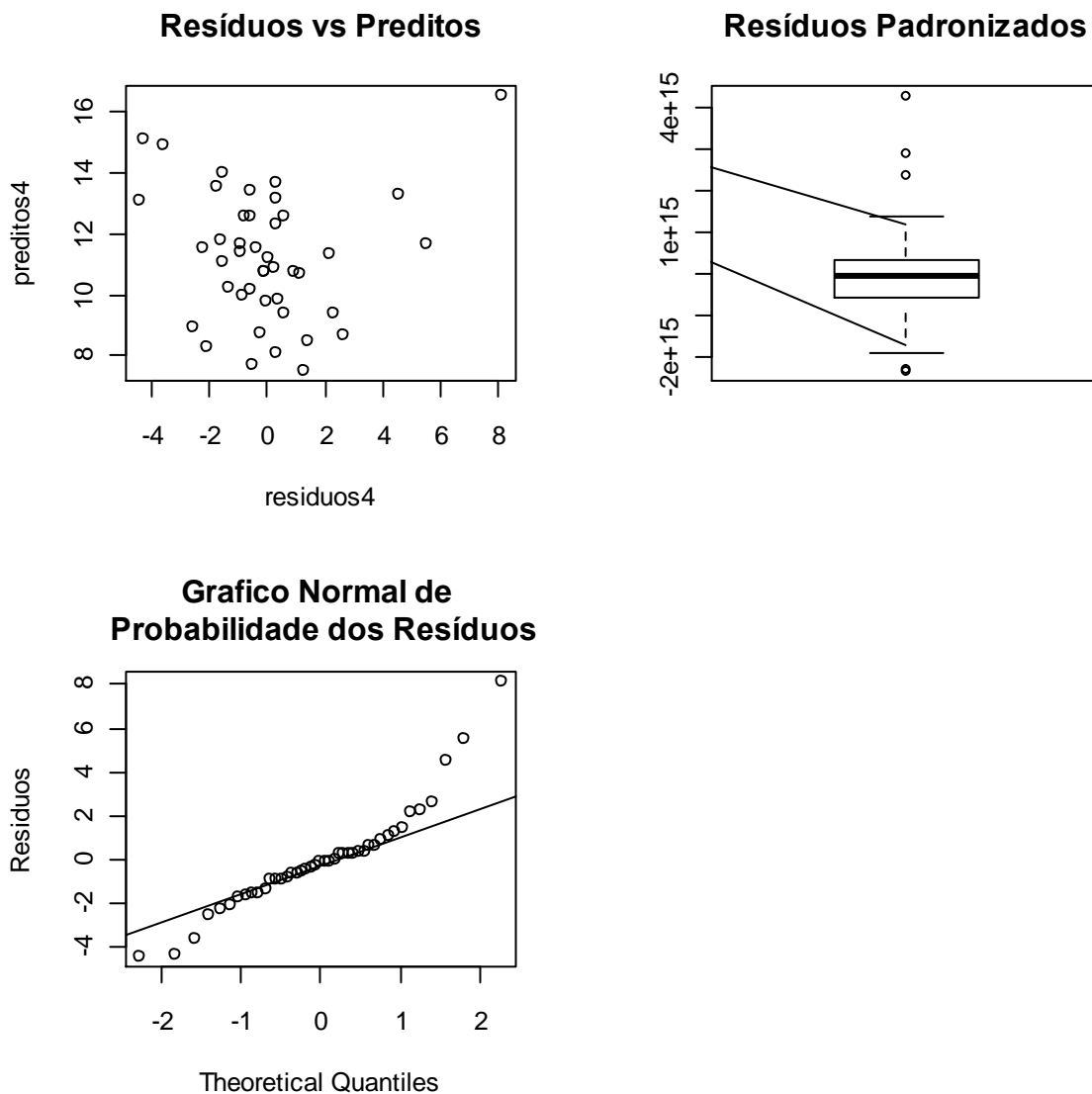


Figura 10. Gráficos da análise de resíduos para a variável diâmetro de colo.

Na figura 10, a análise gráfica da distribuição dos resíduos demonstra-se que os pressupostos da normalidade, homocedasticidade e independência foram atendidos.

Para a variável diâmetro do colo, avaliada aos 12 meses de idade, não se detectou na análise de variância diferenças significativas entre os tratamentos (tabela 19). Este resultado está relacionado a fisiologia de crescimento da espécie, que em função da ausência de competição lateral, está empregando toda a sua energia no crescimento em altura das plantas. A competição por espaço lateral passará a existir quando a copa das plantas começarem a se tocar. Neste ponto passará a haver um crescimento diferenciado no diâmetro das plantas.

7.4.3. Conclusões

- Estudos exploratórios do comportamento da *Aniba rosaedora* quando submetido aos tratamentos de adubação com N, P e K, delineados em blocos ao acaso em arranjo fatorial confundido 3x3x3, indicaram que a espécie responde favoravelmente ao crescimento em altura dos fatores N e P e suas interações.
- O crescimento em diâmetro do colo, aos 12 meses de idade não foi influenciado pelos tratamentos aplicados (N, P e K).
- As interações N vs. k e P vs. K, apresentaram efeito nulo na resposta ao crescimento em altura da espécie, aos 12 meses de idade.
- A análise dos resultados obtidos 12 meses após a aplicação dos tratamentos para uma espécie florestal de ciclo longo (25-30 anos) implica que os mesmos não são conclusivos, recomendando-se novas avaliações durante o período de crescimento da espécie.
- O delineamento e o desenho fatorial confundido foi eficiente para a condução do experimento. O resultado das análises utilizando-se o programa R foi similar aos resultados verificados com o SAS.

CAPÍTULO 8

Considerações e conclusões gerais

- No decorrer deste estudo, foram apresentadas técnicas avançadas de experimentação pouco utilizadas na prática pelos pesquisadores da área agrícola. A popularização de tais técnicas, quando empregadas corretamente, vão permitir a redução do erro experimental com uso eficiente dos recursos físicos e materiais para implantação dos experimentos.
- A validação do desenho experimental com uso do fatorial confundido para estudos exploratórios em diferentes fatores e níveis de adubação em ensaios florestais na região amazônica, é uma técnica que se mostrou adequada no delineamento de experimentos com adubação em espécies florestais.
- O uso de delineamentos e arranjos mais elaborados, como os fatoriais confundidos, implica em que as análises estatísticas devem ser cuidadosamente realizadas, porém tem a vantagem de reduzir o tamanho do experimento e conseqüentemente o custo de implantação.
- O delineamento utilizado na aplicação prática, permitiu uma redução no número de plantas e em área experimental de 66,7%, valor este extremamente significativo para redução dos custos experimentais (insumo, preparo de área, plantio, adubação, manutenção, coleta de dados) na realização dos ensaios.

Perspectivas de investigação futura

- A partir da definição que o fósforo (P) e o nitrogênio (N) são importantes para o crescimento da espécie *Aniba rosaedora*, recomenda-se em investigações futuras, o delineamento de planos experimentais para estudar estes fatores em número maior de níveis para definir a curva de resposta do crescimento da espécie em relação aos respectivos fatores.
- Estudos com outros fatores como o magnésio (Mg) e cálcio (Ca) em diferentes níveis deverão ser realizados, assim como com micronutrientes como o zinco, boro, enchofre e molibdênio.
- Para o fósforo (P) e o nitrogênio (N) novos estudos devem ser realizados para se verificar a influência da época de aplicação dos nutrientes na resposta ao crescimento da espécie.
- Estudo de diferentes delineamentos e arranjos fatoriais poderão ser aplicados em estudos de crescimento de mudas de espécies florestais, considerando o reduzido tempo em que poderá se obter as conclusões, podendo-se validar o uso destas técnicas e tornar o seu uso mais comum para aplicação em ensaios florestais.
- Em futuros plantios experimentais, recomenda-se utilizar parcelas com número menor que 49 plantas . Hipoteticamente, 25 plantas total, com 9 na parcela útil, poderá ser adequado e suficiente, reduzindo-se os custos de implantação e condução com novas experimentações.

Referências Bibliográficas

BAILEY, R.A.; CAMERON, PETER J.; DOBCSÁNYI, PETER; MORGAN, JOHN P. & SOICHER, LEONARD H. *Designs on the web*; **Discrete Mathematics**, Volume 306, Issue 23, 6 December, 3014-3027, 2006.

BAILEY, R. A. **Association Schemes: Designed Experiments, Algebra and Combinatorics**. Cambridge:University Press. 2004.

BARBIN, D. **Planejamento e análise de experimentos agronômicos**. Arapongas: Midas. 2003

BOSE, R. C. & NAIR, K. R. Partially balanced incomplete block designs, **Journal of Combinatorial Theory**, Volume 4, 337-372. 1939.

BOSE, R. C. & SHIMAMOTO, T. Classification and analysis of partially balanced incomplete block designs with two association schemes, **Journal of the American Statistical Association**, Volume 47, 151-184. 1952.

BOSE, R. C. & MESNER, D. M. On linear associative algebras corresponding to association schemes of partially balanced designs, **Annals of Mathematical Statistics**, Volume 30, 31-38. 1959.

BOSE, R. C. Combinatorial properties of partially balanced designs and association schemes, Series A, Volume 25, part 2. 1963.

BOSE, R. C. & LASKAR, R. A characterization of tetrahedral graphs, **Journal of Combinatorial Theory**:2, 266-385. 1967.

BOX, G. E. P.; HUNTER, J.S. **The 2k-p Fractional Factorial Designs Part I. Technometrics**, v.42, Issue 1, p.28-48, 2000.

BOX, G. E. P.; HUNTER, W. G.; HUNTER, J. S. **Statistics for Experiments – An Introduction to Design, Data Analysis and Model Building**. John Wiley & Sons, New York, 1978.

BOX, G. E. P.; MEYER, R. D. Dispersion Effects from Fractional Designs. **Technometrics**, v. 28, Issue 1, p.19-27, 1986.

CALINSKI, T. , KAGEYAMA,S. **Block Designs: A Randomization Approach, Volume I: Analysis**. Lecture Notes in Statistics, Springer, 2000.

CALINSKI, T. , KAGEYAMA,S. **Block Designs: A Randomization Approach, Volume II: Design**. Lecture Notes in Statistics, Springer, 2003.

CLATWORTHY, W. H. Tables of Two-Associate Partially Balanced Designs, National Bureau of Standards, **Applied Mathematics Series 63**, Washington D.C. 1973.

- COCHRAN, W. G.; G. M. COX. **Experimental designs**. 2. Ed. New York:Wiley. 1957.
- COLBOURN, C. J., DINITZ, J.H., **Handbook of Combinatorial Designs**, 2nd Ed., Chapman & Hall/CRC. 2007.
- CONNOR, W. S. & CLATWORTHY, W. H. Some theorems for partially balanced designs, **Annals of Mathematical Statistics**, Volume 25, 100-112. 1954.
- COSTA, S. C. **Modelos Lineares Generalizados Mistos para Dados Longitudinais..** Tese de Doutorado, Escola Superior de Agricultura “Luiz de Queiroz”, USP, Piracicaba, 2003.
- DAGNELIE, P. **Principes d’expérimentation**. Les Presses Agronomiques de Gembloux. 1981.
- DAVIES, O. L.; HAY, W. A. The construction and uses of fractional factorial designs in industrial research. **Biometrics**, v. 6, Issue 3, p. 233-351. 1950.
- DAVIES, O. L. **Design and analysis of industrial experiments**. 2. Ed. New York: Hafner Publishing Company.
- DEAN, A., VOSS, D.. **Design and Analysis of Experiments**. Springer texts in Statistics, 1999.
- FINNEY, D.J. The fractional replications of factorial experiments. **Ann. Eugen.**, 12:291-300. 1945.
- FISHER, R. A. *Statistical methods for research workers*, Oliver and Boyd, Edinburgh. 1925.
- FISHER, R. A. *The design of experiments*, Oliver and Boyd, Edinburgh. 1935.
- FISHER, R. A. An examination of the different possible solutions of a problem in incomplete blocks, **Annals of Eugenics**, 10, 52-75. 1940
- FISHER, R. A.; YATES. **Statistical tables for biological, agricultural, and medical research**. 4. Ed. Edinburgh: Oliver and Boyd.
- FOODY, W AND HEDAYAT, A. On theory and applications of BIB Designs with repeated blocks. **The Annals of Statistics**, 5 (5):932 - 945. 1977.
- GIESBRECHT, F.G., GUMPERTZ, M. L. **Planning, Construction, and Statistical Analysis of Comparative Experiments**. Wiley Series in Probability and Statistics. 2004.
- HARSHBARGER, B. Preliminary report on the rectangular lattices, **Biometrics**: 2, 115-119. 1946.
- HARSHBARGER, B. Rectangular lattices, Virginia Agricultural Experiment Station, Memoir 1. 1947
- HARSHBARGER, B. Triple rectangular lattices, **Biometrics**, 5:1-13. 1949.

HEDAYAT, A.S., AND HWANG, H. L. *BIB(8,56,21,3,6) and BIB(10,30,9,3,2) Designs with Repeated Blocks*. **Journal of Combinatorial Theory**, Series A 36:73 – 91. 1984.

HEDAYAT, A.S., SLOANE, N.J.A., STUFKEN, J. **Orthogonal Arrays: Theory and Applications**. Springer Series in Statistics. 1999.

HINKELMANN, K. & KEMPTHORNE, D. Two classes of group divisible partial diallel crosses, **Biometrika**:50 (3, 4), 281-291. 1963.

HINKELMANN, KLAUS E KEMPTHORNE, OSCAR **Design and Analysis of Experiments, Advanced Experimental Design**, Wiley Series in Probability and Statistics, Wiley-Interscience. 2005

JAMES, A. T. & WILKINSON, G. N. Factorization of the residual operator and canonical decomposition of nonorthogonal factors in the analysis of variance, **Biometrika**:58, 279-294. 1971.

JOHN, P. W. M. **Statistical design and analysis of experiments**, Macmillan, New York. 1971.

JOHN, J. A.; WOLOCK, F. W. & DAVID, H. A. Cyclic Designs, **Applied Mathematics Series**, 62, National Bureau of Standards, Washington D. C. 1972.

JOHN, J. A. **Cyclic Designs**, Chapman and Hall, London. 1987.

JOHN, J. A. & WILLIAMS, E. R. **Cyclic and Computer generated Designs**, Chapman and Hall, London. 1995.

KAGEYAMA, S. Reduction of associate classes of block designs and related combinatorial arrangements, **Hiroshima Mathematical Journal**, 4, 527 – 618. 1974.

KEMPTHORNE O. **The Design an Analysis of Experiments**, Wiley, New York. 1952.

KEMPTHORNE O. The efficiency factor of an incomplete block design, **Annals of Mathematical Statistics**:27, 846-849. 1956.

KUSUMOTO, K. Hyper cubic designs, **Wakayama Medical Reports**:9, 123-132. 1965.

KUSUMOTO, K. Association schemes of new types and necessary conditions for existence of regular and symmetrical PBIB designs with those association schemes, **Annals of Mathematical Statistics**:19, 73-100. 1967.

LIAO, C. T. Identification of dispersion effects from unreplicated $2n-k$ fractional factorial designs. **Computational Statistics & Data Analysis**, v. 33, p. 291-298, 2000.

MASCARENHAS, V.; **Planos em Blocos Incompletos parcialmente Equilibrados: Análise de situações especiais.** Tese de Mestrado em Estatística e Optimização - FCT, Universidade Nova de Lisboa, 2008. (submetida)

MASON, R.L., GUNST, R.F., HESS, J.L.; **Statistical Design and Analysis of Experiments. With Applications to Engineering and Science, 2nd Ed.** Wiley Series in Probability and Statistics. 2003.

MEJZA, I., MEJZA, S.. **Model building and analysis for block design with nested rows and columns.** *Biom. J.* 36, 327–340, 1994.

MONTEIRO, S., OLIVEIRA, T.A.; **Planos Factoriais Fraccionados de Base Dois, com Blocos: A evolução da definição de Resolução e do Critério de Aberração Mínima.** Ferrão, M. E., Nunes, C. e Braumann, C. A., eds, 2007. *Estatística Ciência Interdisciplinar*, Actas do XIV Congresso Anual da SPE, Covilhã, p. 551-559.

MONTEIRO, S.; **Planos Factoriais de Base 2: inteiros e fraccionados.** Tese de Mestrado em Estatística e Optimização - FCT, Universidade Nova de Lisboa, 2007.

MONTEIRO, S., OLIVEIRA, T.A.; **Yate's algorithm for Two-level Factorial and Fractional Designs- an application;** Book of abstracts, ISI 2007, pg 582; paper published on CD Proceedings of ISI 2007, Bulletin of the International Statistical Institute, 56th Session.

MONTGOMERY, D. C. **Introdução ao controle estatístico da qualidade.** 4^a. Ed., Rio de Janeiro: LTC. 2004 (tradução de Ana Maria Lima de Farias; Vera R. Lima de Farias e Flores; Luiz da Costa Laurencel)

MONTGOMERY, D. C. **Design and Analysis of Experiments.** 5^o Ed., John Wiley & Sons, New York. 2001.

MONTGOMERY, D. C.; PECK, E. A. **Introduction to Linear Regression Analysis.** 2^o Ed., John Wiley & Sons, New York. 1991.

MORGAN, J.P., BAILEY, R. . Optimal design with many blocking factors. *The Annals of Statistics*, 2000, Vol 28, 2,553-577.

MYERS R. H.; MONTGOMERY, D. C. A Tutorial on Generalized Linear Models. *Journal of Quality Technology*, v.29, Issue 3, p.274-291, 1997.

MYERS, R. H.; MONTGOMERY, D. C.; VINING, G. G. **Generalized Linear Models With Applications in Engineering and the Sciences.** John Wiley & Sons: New York. 2002.

NAIR, K. R. & RAO, C. R. A note on partially balanced incomplete block designs, *Science and Culture*, 7, 568-569.1942.

NETER, J.; KUTNER, M. H.; NACHTSHEIM, C. J.; WASSERMAN, W. **Applied Linear Statistical Models.** 3rd. Ed. WCB/McGraw-Hill, New York. 1996.

OLIVEIRA, T.A. **Planos de Blocos Equilibrados Incompletos com Repetições**, Tese de Mestrado. Faculdade de Ciências da Universidade de Lisboa. 1994.

OLIVEIRA; T.A. **Planeamento de Experiências: Novas Perspectivas**, Tese de Doutoramento. Universidade de Lisboa. 2000.

OLIVEIRA, A., OLIVEIRA, T.A., MEJZA, S.; **Joint Regression Analysis applied to Genotype Stability Evaluation**. XVIII Summer School in Biometry, Rackova Dolina, Slovakia, 23-27 Junho 2008 (submetido).

PATTERSON, H. D.; WILLIAMS, E. R.; HUNTER E. A. Block designs for variety trials, **Journal of Agricultural Science**, 90, 395 – 400. 1978.

PATTERSON, H. D.; WILLIAMS, E. R.; HUNTER E. A. Block designs for variety trials, **Journal of Agricultural Science**, 90, 395 – 400. 1978.

PEARCE, S. C. The mean efficiency of equi-replicate designs, **Biometrika**:55, 251-253. 1968.

PEARCE, C.; CALINSKI, T. & MARSHALL, T. F. DE C. The basic contrasts of an experimental design with special reference to the analysis of data, **Biometrika**:61 449-460. 1974.

PIMENTEL-GOMES F. ; GARCIA, C. H. **Estatística aplicada a experimentos agronômicos e florestais**. FEALQ. 2002.

PIMENTEL-GOMES, F.; CONAGIN, A. **Experimentos de adubação: planejamento e análise estatística**. UEL. 1987.

RAGHAVARAO, D. **Constructions and Combinatorial Problems in Design of Experiments**, New York:John Wiley and Sons. 1971.

RAO, C. R. General methods of analysis for incomplete block designs, **Journal of the American Statistical Association**, 42, 541-561. 1947.

RAO, R. R. A. Fisher: The Founder of Modern Statistics, **Statistical Science**,7(1): 34-48. 1992.

SOKAL, R. R.; ROHLF, F. J. **Biometria: principios y métodos estadísticos en La investigación biológica**. 1ª. Ed. H. Blume Ediciones: Rosario. 1979 (traduzido do original **Biometry** por Miguel Lahoz Leon)

SURENDRAN, P. U. Association matrices and Kronecker product of designs, **Annals of Mathematical Statistics**:39, 676-680. 1968.

THARTHARE, S. K. Right angular designs, **Annals of Mathematical Statistics**, 36, 1051-1062. 1963.

THARTHARE, S. K. Generalized right angular designs, **Annals of Mathematical Statistics**, 34, 1057-1067. 1965.

YATES, F. Incomplete randomized blocks, **Annals of Eugenics**, 7, 121-140. 1936a.

YATES, F. A new method for arranging variety trials involving a large number of varieties, **Journal of Agricultural Science**, 26, 424-455. 1936b

YATES, F. The design and analysis of factorial experiments, **Technical Communication**, 35, Imperial Bureau of Soil Science, Harpenden. 1937a.

YATES, F. A further note on the arrangement of variety trials:quasi-Latin squares, **Annals of Eugenics**, 7, 319-332. 1937b.

YATES, F. The recovery of inter-block information in variety trials arranged in three-dimensional lattices, **Annals of Eugenics**, 9, 136-156. 1939.

YATES, F. The recovery of inter-block information in balanced incomplete block designs, **Annals of Eugenics**, 10, 317-325. 1940.

VARTAK, M. N. On application of Kronecker product of matrices to statistical designs, **Annals of Mathematical Statistics**, 26, 420-438. 1955.

VARTAK, M.N. The non-existence of certain PBIB designs, **Annals of Mathematical Statistics**, 30, 1051-1062. 1959.

ZIMMERMANN, F. J. P. **Estatística aplicada à pesquisa agrícola**. Santo Antônio de Goiás: Embrapa Arroz e Feijão. 2004

ANEXOS

Registros fotográficos da condução do experimento no município de
Maués, AM, Brasil



Figura 10. Pau-rosa (*Aniba rosaedora* Ducke), aos 12 meses de Idade. Maués, AM, Brasil. 2010



Foto: Roberval Lima

Figura 11. Visão geral do experimento em Fatorial NPK com 3 níveis em Maués, AM, Brasil. 2010.



Figura 12. Outra visão geral do experimento em Fatorial NPK com 3 níveis em Maués, AM, Brasil. 2010.



Figura 13. Pau-rosa (*Aniba rosaedora* Ducke), com 1 ano de idade, Maués, AM, Brasil. 2010.



Foto: Roberval Lima

Figura 14. Medição do diâmetro de colo com paquímetro digital. Maués, AM, Brasil. 2010.