Planejamento de Experimentos Delineamento em Blocos Completamente Aleatorizados

Enrico A. Colosimo/UFMG

Depto. Estatística - ICEx - UFMG

Soluções Enxaguantes

- Um experimento foi realizado para comparar o efeito de 3 enxaguantes em retardar o crescimento de bactérias em containers de leite (20 l).
- A análise é realizada em laboratório e somente é possível realizar 3 testes por dia.
- O dia pode ser fonte de variabilidade?
- O experimentador decidiu aleatorizar os enxaguantes por dia. Você concorda?
- O resultado (em horas) para 4 dias de experimento está mostrado na tabela a seguir.

Tabela: Tempo até a aparecimento de bactérias - Três soluções em 4 dias.

	Dias				
Soluções	1	2	3	4	Médias
1	13	22	18	39	23,0
2	16	24	17	44	25,3
3	5	4	1	22	8,0
Médias	11,3	16,7	12,0	35,0	18,8

Delineamento em Blocos

- Em várias situações experimentais é necessário planejar o experimento de modo que a variabilidade proveniente de um fator de perturbação ("nuisance factor") possa ser controlada.
- O planejamento com blocos aleatorizados é uma extensão do teste t pareado para situações em que o fator de interesse tem mais de dois níveis.
- O planejamento é dito em blocos completos aleatorizados porque cada bloco inclui todos os tratamentos, e os mesmos são alocados de forma aleatória.
- O procedimento geral para um planejamento com blocos completos aleatorizados consiste em selecionar b blocos e correr uma réplica completa do experimento em cada bloco.
- Haverá *a* observações (uma por nível do fator) em cada bloco.

Delineamento em Blocos

- O experimento em blocos completos aleatorizados pode ser representado pelo modelo (linear) estatístico

$$Y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij},$$

  com $i = 1, ..., a$ e $j = 1, ..., b$, em que
  - μ é a média geral, comum a todos os tratamentos;
  - τ_i é o efeito do i-ésimo tratamento;
  - β_j é o efeito do j-ésimo bloco;
  - ϵ_{ij} é um componente estocástico associado a variabilidade não explicada inerente ao processo (erro aleatório) com distribuição $N(0, \sigma^2)$;
- Novamente, os efeitos dos tratamentos e dos blocos são definidos como desvios da média global, de modo que

$$\sum_{i=1}^{a} \tau_i = 0 \quad \text{e} \quad \sum_{j=1}^{b} \beta_j = 0.$$

Delineamento em Blocos

- Para o modelo com blocos completos aleatorizados, assumimos a não existência de interação entre blocos e tratamentos.
- Estamos interessados em testar

$$\begin{cases} H_0 : \tau_1 = \tau_2 = \cdots = \tau_a = 0 \\ H_1 : \tau_i \neq 0 \text{ para pelo menos um } \ i = 1, \ldots, a. \end{cases}$$

- A SQ_{Total} associada a um experimento com blocos completos aleatorizados pode ser escrita da seguinte forma:

$$\sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij} - \bar{y}_{..})^2 = b\sum_{i=1}^{a}(\bar{y}_{i.} - \bar{y}_{..})^2 + a\sum_{j=1}^{b}(\bar{y}_{.j} - \bar{y}_{..})^2 + \sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij} - \bar{y}_{.j} - \bar{y}_{i.} + \bar{y}_{..})^2$$

- Simbolicamente, temos

$$SQ_{Total} = SQ_T + SQ_B + SQ_E.$$

- Pode ser mostrado que os valores esperados de QM_T QM_B e QM_E são, respectivamente, dados por

$$E(QM_T) = E\left(\frac{SQ_T}{a-1}\right) = \sigma^2 + \frac{b\sum_{i=1}^{a}\tau_i^2}{a-1},$$

$$E(QM_B) = E\left(\frac{SQ_B}{b-1}\right) = \sigma^2 + \frac{a\sum_{j=1}^{b}\beta_j^2}{b-1},$$

e

$$E(QM_E) = E\left(\frac{SQ_E}{(a-1)(b-1)}\right) = \sigma^2.$$

Delineamento em Blocos

- Fórmulas de cálculo para ANOVA - bloco aleatorizado:

$$SQ_{Total} = \sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij}^2 - \frac{y_{..}^2}{ab}$$

e

$$SQ_T = \frac{1}{b}\sum_{i=1}^{a} y_{i.}^2 - \frac{y_{..}^2}{ab}.$$

$$SQ_B = \frac{1}{a}\sum_{j=1}^{b} y_{.j}^2 - \frac{y_{..}^2}{ab}.$$

- O cálculo de SQ_E pode ser feito por subtração, isto é,

$$SQ_E = SQ_{Total} - SQ_T - SQ_B$$

Tabela: ANOVA para um experimento com blocos completos aleatorizados.

Fonte de Variação	Soma de Quadrados	Graus de Liberdade	Quadrado Médio	F_0
Tratamento	SQ_T	a-1	QM_T	QM_T/QM_E
Blocos	SQ_B	b-1	QM_B	QM_B/QM_E
Erro	SQ_E	(a-1)(b-1)	QM_E	
Total	SQ_{Total}	ab-1		

Exemplo: Enxaguantes em 4 dias

```
> modelo=aov(resp~trat+bloco,data=dados)
>  summary(modelo)
Df Sum Sq Mean Sq F value   Pr(>F)
trat          2  703.5   351.8   40.72 0.000323 ***
bloco         3 1106.9   369.0   42.71 0.000192 ***
Residuals     6   51.8     8.6
---
```

Análise de resíduos

- As conclusões obtidas a partir da tabela da ANOVA dependem das seguintes suposições:
  - Observações são independentes e normalmente distribuídas.
  - Variância constante e comum para cada nível do fator em estudo.
- A verificação de tais suposições é feita através da chamada análise de resíduos.
- Por definição, um resíduo é a diferença entre a observação y_{ij} e o seu valor estimado (ou ajustado) $\hat{y}_{ij}$, obtido a partir do modelo estatŽsstico sendo estudado, isto é,

$$e_{ij} = y_{ij} - \hat{y}_{ij}.$$

- Para o planejamento com blocos completos aleatorizados, mostra-se que $\hat{y}_{ij} = \bar{y}_{i.} + \bar{y}_{.j} - \bar{y}_{..}$ e, consequentemente $e_{ij} = y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}$.

Análise de resíduos

- Os resíduos assim definidos podem ser interpretados como estimativas do erros ϵ_{ij} não observados.
- A análise dos resíduos é indispensável para a validação do modelo em estudo.
- Se o modelo ajustado é adequado para os dados observados, então os resíduos não devem apresentar nenhum tipo de padrão.
- A suposição de normalidade pode ser verificada pela construção um histograma dos resíduos.
- Infelizmente, para amostras pequenas, uma considerável flutuação na forma do histograma é observada com certa frequência, de tal forma que uma moderada desvio da normalidade não necessariamente implica em uma séria violação desta suposição.
- Entretando, desvios grosseiros (em especial quando o tamanho da amostra é grande), são potencialmente sérios e requerem análises adicionais.
- Outro gráfico bastante utilizado para verificar a suposição de normalidade é o chamado gráfico de probabilidade normal.

Análise de resíduos

- Um gráfico simples (não necessariamente adequado para todas as situções) é o de resíduos na ordem em que os dados foram coletados. Ele possibilita a deteção de possível correlação entre os erros.
- Uma tendência significa obter corridas com resíduos positivos e negativos, indicando correlação positiva, e consequentemente, violação da suposição de independência dos erros.

Análise de resíduos

- Além da análise gráfica, alguns testes encontram-se disponíveis para testar a validade das suposições impostas pelo modelo.
- Testes de normalidade:
  - Teste de Shapiro-Wilk;
  - Teste de Anderson-Darling
- Testes de homogeneidade das variâncias:
  - Teste de Bartlett;
  - Teste de Levene;
- É indicado realizar a análise de resíduos considerando tanto a análise gráfica quanto os testes indicados acima.
- Observe que a hipótese nula é sempre de validade da suposição.

Comparações Múltiplas

- Quando $H_0 : \tau_1 = \cdots = \tau_a = 0$ não é rejeitada, concluímos que não existem diferenças estatisticamente significativas entre os níveis do fator de interesse.

- Caso $H_0 : \tau_1 = \cdots = \tau_a = 0$ seja rejeitada na ANOVA, o próximo passo consiste em verificarmos quais níveis do fator de interesse apresentam diferenças estatisticamente significativas.

- Os métodos para realizar este tipo de análise são chamados métodos de comparações múltiplas.

- Existem diversos métodos disponíveis para a realiação de comparações múltiplas, dentre os quais destacamos:
  - Teste de Bonferroni.
  - Teste de Tukey (HSD).
  - Teste de Scheffè.
  - etc, etc.

- Família: Uma família de testes é um conjunto de testes baseado em alguma medida de erro global.

- Taxa de erro da família: (familywise error rate): é a probabilidade de rejeitarmos incorretamente ao menos umas das hipóteses nulas que compõem a família.

- Taxa de erro por comparação (per-comparison error rate): é a probabilidade de rejeitarmos incorretamente cada uma das hipóteses nulas que compõem a família.

Comparações Múltiplas

- Na presença de a tratamentos (ou níveis), temos $\binom{a}{2} = \frac{a(a-1)}{2}$ comparações de pares de médias.

- A probabilidade de cometermos pelo menos um erro do tipo I é maior que α (note que estamos realizando vários testes simultaneamente).

- É difícil calcular de maneira exata a probabilidade global de erro (do tipo I) envolvendo comparações múltiplas.

- Uma cota superior para a probabilidade global de erro é dada por

$$P(\text{no. erros tipo I} \geq 1) = 1 - (1-\alpha)^{\frac{a(a-1)}{2}}.$$

- OBS: a cota superior apresentada acima é obtida assumindo que as comparações entre os pares de médias são independentes (o que de fato não ocorre na prática).

Contrastes e Contrastes Ortogonais

- Muitos métodos de comparações múltiplas utilizam a ideia de contrastes.
- Em geral, um contraste é uma combinação linear dos parâmetros da forma

$$\Gamma = \sum_{i=1}^{a} c_i \mu_i,$$

tal que $\sum_{i=1}^{a} c_i = 0$.

- Contrastes são utilizados para testarmos hipóteses do tipo

$$\begin{cases} H_0 : \sum_{i=1}^{a} c_i \mu_i = 0 \\ H_1 : \sum_{i=1}^{a} c_i \mu_i \neq 0 \end{cases}$$

- Retornando ao exemplo dos enxaguantes, a comparação entre as médias dos enxaguantes 1 e o 2 pode ser escrita como: $c = (1, -1, 0)$.

Contrastes e Contrastes Ortogonais

- Seja $C = \sum_{i=1}^{a} c_i \bar{y}_{i.}$. Então C tem distribuição normal e

$$E(C) = \sum_{i=1}^{a} c_i \mu_i \quad \text{e} \quad V(C) = \frac{\sigma^2}{n} \sum_{i=1}^{a} c_i^2$$

- Logo, sob H_0 temos que

$$T_0 = \frac{\sum_{i=1}^{a} c_i \bar{y}_{i.}}{\sqrt{\frac{QM_E}{n} \sum_{i=1}^{a} c_i^2}} \sim t_{N-a}.$$

- Alternativamente, temos que

$$F_0 = T_0^2 = \frac{\left(\sum_{i=1}^{a} c_i \bar{y}_{i.}\right)^2}{\frac{QM_E}{n} \sum_{i=1}^{a} c_i^2} \sim F_{1,N-a}.$$

- Note que

$$F_0 = \frac{\left(\sum_{i=1}^{a} c_i \bar{y}_{i.}\right)^2}{\frac{QM_E}{n} \sum_{i=1}^{a} c_i^2} \quad \Rightarrow \quad SQ_C = \frac{\left(\sum_{i=1}^{a} c_i \bar{y}_{i.}\right)^2}{\frac{1}{n} \sum_{i=1}^{a} c_i^2}.$$

n (um fator) pode ser substituído por *b* no caso do experimento em bloco.

Contrastes e Contrastes Ortogonais

- Os contrastes ortogonais são um caso especial de contrastes, são úteis na prática.
- Dizemos que dois contrastes com coeficientes c_i e d_i são ortogonais se

$$\sum_{i=1}^{a} c_i d_i = 0,$$

  ou, para experimentos desbalanceados,

$$\sum_{i=1}^{a} n_i c_i d_i = 0,$$

- Para experimentos com um fator fixo e a níveis, existem apenas $a - 1$ contrastes ortogonais.
- Os contrastes ortogonais particionam a SQ_T em $a - 1$ somas de quadrados indepedentes, cada uma com um grau de liberdade.
- Consequentemente, testes baseados em contrastes ortogonais são independentes.

Método de Bonferroni

- O método de Bonferroni considera significativa uma diferença entre pares de médias se

$$|\bar{y}_{i.} - \bar{y}_{j.}| > t_{\frac{\alpha'}{2};(a-1)(b-1)} \times \sqrt{QM_E\left(\frac{1}{b}+\frac{1}{b}\right)}$$

  em que $\alpha' = \alpha/\left(\frac{a(a-1)}{2}\right)$.

- O método de Bonferroni é um teste de comparações múltiplas convervador, pois

$$P(\text{Erro I global}) \leq \alpha. \tag{1}$$

Teste de Bonferroni (via simulação)

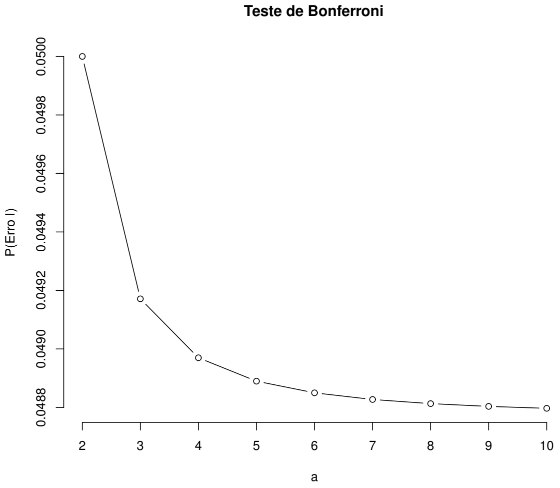

Figura: Probabilidades do Erro I - Tese de Bonferroni.

Método de Tukey

- Tukey propôs um procedimento para comparar todos os pares de médias com as hipóteses nulas $H_0 : \mu_i = \mu_j, \forall\, i \neq j$ usando a distribuição da seguinte estatística

$$Q_0 = \frac{\bar{Y}_{max} - \bar{Y}_{min}}{\sqrt{QM_E\left(\frac{1}{b} + \frac{1}{b}\right)}}$$

- Considerando $H_0 : \mu_i = \mu_j$, o par de médias (μ_i, μ_j) é declarado significativamente diferente se:

$$|\bar{y}_{i.} - \bar{y}_{j.}| > \frac{q_\alpha(a,(a-1)(b-1))}{\sqrt{2}} \times \sqrt{QM_E\left(\frac{1}{b} + \frac{1}{b}\right)} = HSD$$

em que $q_\alpha(a,(a-1)(b-1))$ é um valor tabelado e α é a probabilidade de erro global associada aos $\binom{a}{2}$ testes.

Teste de Tukey - Exemplo Enxaguantes

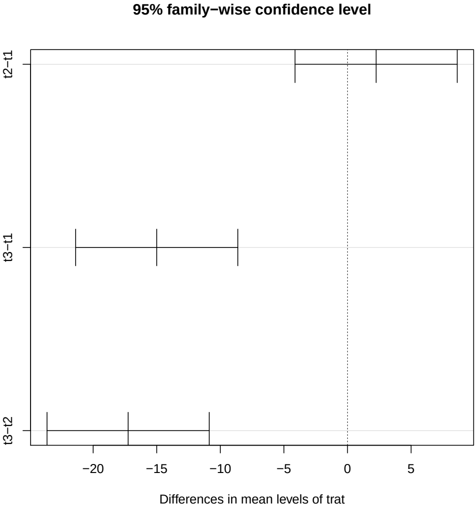

Figura: Teste de Tukey.

Conclusão - Enxaguantes

- Não existe diferença entre os enxaguantes 1 e 2.
- Existe diferença entre as médias dos enxaguantes 1 e 3. O enxaguante 1 prolonga o aparecimento de bactérias, em média, em 15 horas (IC 95% 8; 22) comparado ao 3.
- Existe diferença entre as médias dos enxaguantes 2 e 3. O enxaguante 2 prolonga o aparencimento de bactérias, em média, em cerca de 17 horas (IC 95% 10; 24) comparado ao 3.

Quadrado Latino

- Suponha que um pesquisador quer estudar o efeito de 5 tratamentos.
- Cada lote de materia prima é suficiente para produzir exatamente os 5 tratamentos.
- Os tratamentos são produzidos por diferentes operadores.
- Fator de interesse: Tratamento com níveis 1,2,3,4,5.
- Fatores de perturbação: matéria prima e operador.
- Quadrado Latino: conduzir cada tratamento exatamente uma vez com cada operador e matéria prima.
- Vantagem: remover efeito de ambos fatores de perturbação.

Exemplo: Quadrado Latino

- Um engenheiro está estudando o tempo de montagem de um determinado componente elétrico.
- Existem 4 métodos de montagem (A, B, C e D).
- Quatro operadores forem selecionados para o estudo.
- Sabe-se que a tarefa produz fadiga nos operadores, acarretando possivelmente um maior tempo para as últimas montagens.
- Fatores de perturbação: operador e ordem.
- Conduzir cada método de montagem exatamente uma vez com cada operador e em cada ordem.
- Vantagem: remover efeito de ambos fatores de perturbação.

Exemplo: Quadrado Latino

Tabela: Tempo de Montagem (minutos).

	Operador			
Ordem	1	2	3	4
1	C=10	D=14	A=7	B=8
2	B=7	C-18	D=11	A=8
3	A=5	B=10	C=11	D=9
4	D=10	A=10	B=12	C=14

Quadrado Latino

- O quadrado latino é representado pelo modelo (linear) estatístico

$$Y_{ijk} = \mu + \tau_i + \beta_j + \alpha_k + \epsilon_{ijk},$$

  com $i = 1, ..., p$, $j = 1, ..., p$ e $k = 1, ..., p$, em que
  - μ é a média geral, comum a todos os tratamentos;
  - τ_i é o efeito do i-ésimo tratamento;
  - β_j é o efeito do j-ésimo nível do primeiro bloco;
  - α_k é o efeito do k-ésimo nível do segundo bloco;
  - ϵ_{ijk} é um componente estocástico associado a variabilidade não explicada inerente ao processo (erro aleatório) com distribuição $N(0, \sigma^2)$;
- Novamente, os efeitos dos tratamentos e dos blocos são definidos como desvios da média global, de modo que

$$\sum_{i=1}^{p} \tau_i = 0 \quad \text{e} \quad \sum_{j=1}^{p} \beta_j = \sum_{k=1}^{p} \alpha_j = 0.$$

Quadrado Latino

- As suposições do modelo devem ser verificadas da forma usual, utilizando os resíduos.
- Múltiplas comparações são usadas, caso haja diferenças entre os níveis do fator, utilizando os valores dos blocos como replicações.
- As conclusões são obtidas a partir dos métodos de comparações múltiplas.

Conclusão - Métodos de Montagem

- Não existe diferença entre os métodos de montagem considerando um nível de significância de 0,05.
- Uma diferença marginalmente significante entre os métodos A e C pode ser observada a um nível $\alpha = 0,10$.
- No caso acima, uma diferença média de 5,5 horas a favor do método A (redução do tempo médio de montagem) com um intervalo de 90% de confiança de 0,1 a 11,4 horas.