

1. Os dados a seguir correspondem a um estudo de melhoramento genético em gados de corte. Os pesos ao nascer de oito progênies machos resultantes de cruzamentos envolvendo cinco touros reprodutores são apresentados a seguir.

Touro Reprodutor	Peso ao nascer de progênies								Média	dp
T1	61	100	56	113	99	103	75	62	83,625	22,551
T2	75	102	95	103	98	115	98	94	97,5	11,225
T3	58	60	60	57	57	59	54	100	63,125	15,028
T4	57	56	67	59	58	121	101	101	77,5	25,945
T5	59	46	120	115	115	93	105	75	91	28,025

a) Por que seria apropriado assumir um modelo de efeitos aleatórios na análise destes dados? Sob tal modelagem dos dados, que suposições estão envolvidas? As médias amostrais apresentadas na tabela acima correspondem a uma amostra aleatória de qual distribuição? As observações avaliadas nas progênies dentro de cada touro reprodutor correspondem a uma amostra aleatória de qual distribuição?

b) Obtenha a Tabela de ANOVA destes dados. Interprete os resultados.

c) Estime os componentes de variância (entre respostas de diferentes touros reprodutores e entre progênies do mesmo touro reprodutor) e o coeficiente de correlação intra-classe do modelo proposto. Interprete os resultados. Qual fonte de variabilidade mais influencia no peso dos animais?

2. Um estudo foi conduzido por pesquisadores do IPT para avaliar a qualidade de cápsulas usadas em máquinas de café expresso. Dois supermercados foram escolhidos para fazer parte do estudo. De cada supermercado três marcas foram selecionadas e, de cada uma, três cápsulas foram avaliadas relativamente a uma medida de qualidade. Os seguintes dados foram obtidos:

Supermercado I			Supermercado II		
M1	M2	M3	M1	M2	M3
1,12	0,16	0,15	0,91	0,66	2,17
1,10	0,11	0,12	0,83	0,83	1,52
1,12	0,26	0,12	0,95	0,61	1,58

a) Considere que as marcas M1, M2 e M3, são as únicas disponíveis nos mercados nacionais. Como os dados dos dois Supermercados podem ser combinados? Apresente

a Tabela de ANOVA resultante. Há diferença entre as marcas? Obtenha um intervalo de confiança para a média da resposta.

b) Considere que as marcas avaliadas em cada Supermercado correspondem a uma amostra aleatória das possíveis marcas disponíveis nos mercados nacionais, tal que, o estudo combinado, está avaliando seis dentre o total dessas marcas. Como os dados dos dois Supermercados podem ser combinados? Apresente a Tabela de ANOVA resultante. Há diferença entre as marcas? Obtenha um intervalo de confiança para a média da resposta.

c) Na prática, como decidir se um fator deve ser modelado como fixo ou aleatório?

3. Considere o seguinte modelo linear de efeitos mistos,

$$y_{ijk} = \mu_{jk} + u_{ijk} + e_{ijk}$$

com $u_{ijk} \sim N(0; \sigma_p^2)$ e $e_{ijk} \sim N(0; \sigma_e^2)$ variáveis aleatórias independentes, $Cov(y_{ijk}; y_{i'j'k'}) = \sigma_p^2$ para $i=i', j=j'$ e $k \neq k'$, e 0 caso contrário, $j=1, \dots, a$, $k=1, \dots, b$. Matricialmente, tem-se:

$$Y_{n \times 1} = X_{n \times p} \beta_{p \times 1} + Z_{n \times q} u_{q \times 1} + e_{n \times 1},$$

com Y o vetor de observações, X e Z matrizes de planejamento conhecidas, associadas, respectivamente, com o vetor de parâmetros β e o vetor de variáveis aleatórias u .

a) Para um experimento Split-Plot 2x3, com duas réplicas ($i=1,2$) para a aleatorização do primeiro fator (em $a=2$ níveis), e considerando o modelo acima, apresente as matrizes X e Z correspondentes. Como os vetores β e u estão definidos?

b) Assuma valores para os parâmetros* do modelo definido em a) e gere um conjunto de observações. Use seu número USP como semente na simulação dos dados. Apresente a tabela dos dados obtidos.

*Exemplo dos parâmetros: $\mu_{jk}, \sigma_p^2, \sigma_e^2; j=1,2; k=1,2,3$.

c) Analise os dados gerados. Apresente os resultados obtidos e interprete o efeito dos fatores.

4. Considere os seguintes dados de um experimento industrial em que uma resposta da produção de peças foi avaliada segundo os fatores Tipo de Material (M1 a M5), Dia (D1 e D2) e Turno (T1, T2 e T3) da linha de produção. Há interesse em investigar o efeito desses fatores no padrão de variação da resposta de produção.

Dia	Turno	M1	M2	M3	M4	M5
D1	T1	20	14	17	12	22
	T2	16	19	16	17	21

	T3	25	32	31	24	24
D2	T1	18	22	9	15	17
	T2	21	21	16	20	17
	T3	19	28	22	29	29

Três modelos foram propostos na análise destes dados:

- **Mod1:** Modelo de ANOVA com efeito fixo dos fatores Tipo de Material e Bloco (em seis níveis, definido pelo cruzamento dos fatores Dia e Turno).
- **Mod2:** Modelo de ANOVA com efeito fixo dos fatores Tipo de Material, Dia, Turno, e interação Dia*Turno.
- **Mod3:** Modelo de ANOVA com efeito fixo dos fatores Material e Turno, e efeito aleatório de Bloco (em seis níveis, definido pelo cruzamento dos fatores Dia e Turno).

a) Em cada caso ajuste o modelo proposto, sob normalidade dos dados e faça suposições necessárias sobre os componentes de (co)variância das observações. Apresente a Tabela de ANOVA resultante e as estimativas dos parâmetros envolvidos.

b) Em cada caso, a linha de produção é mais afetada por qual dos fatores sob estudo?

c) Qual modelo você recomendaria na análise dos dados? Justifique.

5. Considere os dados a seguir de um estudo avaliando a mudança no peso de peixes submetidos a ambientes com e sem seus competidores naturais (peixes de uma outra espécie) e densidade da água, nos níveis baixo e alto. Para o experimento, 16 tanques foram aleatorizados a quatro compartimentos, dimensionados para receber quatro tanques. Dentro de cada compartimento os tratamentos foram aleatorizados aos tanques. Após 5 semanas o peso dos peixes foi mensurado. Os dados estão apresentados a seguir.

Competidor	Densidade	Compartimentos			
		1	2	3	4
Ausente	Baixa	0.0	0.4	0.9	-0.4
	Alta	0.9	0.4	0.6	1.2
Presente	Baixa	0.0	-0.4	-0.9	-0.9
	Alta	-1.2	-1.5	-1.1	-0.7

a) Analise estes dados supondo compartimento como fator (fixo) de blocagem.

b) Analise os dados supondo um delineamento Split-Plot.

c) Apresente as estimativas dos componentes de (co)variância sob cada um dos modelos propostos. Como decidir entre os dois modelos acima?

6. Proponha um delineamento experimental para as situações a seguir e justifique:

E1. Um floricultor deseja estudar os efeitos da drenagem do solo e da variedade de bulbos de tulipas na produção de flores. A área de plantação para o experimento está dividida em 12 canteiros de 3x10m com 0,5m de profundidade. A drenagem do solo é definida pela mistura de areia ao solo que, em geral, é argiloso. Mais areia melhora a drenagem. Estando o solo preparado e drenado, os bulbos são então plantados no solo e a produção de flores é mensurada na primavera seguinte. O floricultor deseja avaliar o efeito de quatro diferentes níveis de drenagem do solo e quinze variedades de tulipas na produção final.

E2. É hora de Produzir Pudim! Há interesse em saber como produzir o Melhor Pudim do Mundo. Foram definidas 4 receitas básicas, dois tipos de mistura dos componentes das receitas e dois tipos de cozimento das mesmas. Os alunos de um curso serão responsáveis por produzir os pudins. Para uma melhor condução do experimento, os alunos deverão ser divididos em grupos para que cada grupo confeccione uma receita. Dentro dos grupos o tipo de mistura e o tipo de cozimento deverão ser aleatorizados aos alunos para a produção final do pudim. Uma análise sensorial do pudim pronto será usada para avaliar os vários fatores envolvidos na produção. Considere que há um certo número elevado de alunos que aceitaram participar do experimento.

E3. Um Ensaio Clínico será conduzido para comparar a eficácia de dois medicamentos, A e B, na redução de dor em episódios de enxaqueca. Para fazer parte do estudo, há 22 voluntários do sexo feminino e 22 do sexo masculino.

7. Considere os seguintes dados de um experimento em blocos incompletos:

Blocos											
1	2	3	4	5	6	7	8	9	10	11	12
A	D	G	A	B	C	A	B	C	A	B	C
19	6	21	20	17	15	20	16	13	20	17	14
B	E	H	D	E	F	E	F	D	F	D	E
17	26	19	7	26	23	26	23	7	24	6	24
C	F	I	G	H	I	I	G	H	H	I	G
11	23	28	20	19	31	31	21	20	19	29	21

a) Por que este é um experimento em blocos incompletos? Pode ser considerado balanceado? Justifique.

b) Na análise destes dados, obtenha a tabela de ANOVA e comparações entre médias de interesse.

c) Se há interesse em controlar um segundo fator Bloco, como um experimento poderia ser planejado?

8. Considere um Delineamento Quadrado Latino 4x4 com réplicas. Apresente um exemplo, o modelo estrutural e distribucional, bem como o esqueleto da ANOVA (FV e número de GL) nos seguintes casos:

- a) Ambos os fatores Bloco (Horizontal e Vertical) estão replicados.
- b) Somente um dos fatores Bloco está replicado.
- c) Nenhum dos fatores Bloco está replicado.

9. Considere os seguintes dados de um delineamento fatorial 2^3 :

A	B	C	Y
-1	-1	-1	75.4
-1	-1	1	75.3
-1	1	-1	76.8
-1	1	1	76.5
1	-1	-1	73.9
1	-1	1	71.4
1	1	-1	72.8
1	1	1	72.3

- a) Analise estes dados por meio de um modelo de ANOVA.
- b) A partir deste experimento fatorial, obtenha um Fatorial Fracional $2^{(3-1)}$. Quais efeitos estão confundidos?
- c) Analise os dados do Fatorial Fracional obtido.

10. Considere os dados a seguir de um Delineamento Quadrado Latino 3x3:

QI	Idade		
	Jovem	Adulto	Idoso
Alto	B	A	C
	19	20	25
	16	24	21
Normal	C	B	A
	24	14	14
	22	15	14
Baixo	A	C	B
	10	12	7
	14	13	4

- a) Analise os dados: obtenha a tabela de ANOVA e as comparações de médias de interesse. Apresente também o modelo estrutural e distribucional adotado.
- b) Calcule a eficiência da blocagem realizada neste experimento. Qual das duas estruturas de Bloco contribui mais para aumentar a eficiência do experimento?
- c) Usando este banco de dados, compare a eficiência da blocagem dupla (por QI e Idade) com um suposto Delineamento Completamente Aleatorizado com 6 réplicas que poderia ter gerado os dados.