Delineamento e Análise de Experimentos

Professora Juliana Betini Fachini Gomes
e-mail: jfachini@unb.br

Brasília - 2023

Exemplo

- Suponha que um pesquisador esteja estudando os efeitos de cinco formulações diferentes de um propulsor de foguete usado em sistemas de escape de tripulação na taxa de queima observada;
- Cada formulação é misturada a partir de um lote de matéria-prima que é grande o suficiente apenas para cinco formulações serem testadas;
- As formulações são preparadas por vários operadores

Exemplo

- Logo, parece que há dois fatores de "incômodo" a serem considerado no projeto: lotes de matéria-prima e operadores;
- O projeto apropriado para este problema consiste em testar a exatidão de cada formulação uma vez em cada lote de matéria-prima e que cada formulação seja preparada exatamente uma vez por cada um dos cinco operadores;
- O experimento adequado para esse problema é o Experimento em Quadrado Latino.

Os experimentos em Quadrado Latino consideram os seguintes princípios básicos da experimentação:

- Casualização;
- Repetição;
- Controle local em dois sentindos perpendiculares: chamados de linhas e colunas

Quadrado Latino

- A principal característica nesse tipo de experimento é que o número de linhas é igual ao número de colunas e igual ao número de tratamentos;
- Então, cada tratamento aparece apenas uma vez em cada linha e uma vez em cada coluna;
- Se tivermos p tratamentos, teremos p^2 parcelas.

Quadrado Latino

Tabela 1: Quadrado Latino antes da casualização

	Colunas				
Linhas	1	2	3	4	5
1	A	B	C	D	E
2	B	C	D	E	A
3	C	D	E	A	B
4	D	E	A	B	C
5	E	A	B	C	D

Quadrado Latino

- A casualização pode ser feita em uma ou duas etapas;
- Em uma etapa - faz-se o sorteio das linhas ou das colunas;
- Em duas etapas - primeiro faz-se o sorteio das linhas e em seguida das colunas. Ou vice e versa.

Quadrado Latino

Tabela 2: Quadrado Latino após casualização das colunas

	Colunas				
Linhas	3	5	1	2	4
1	C	E	A	B	D
2	D	A	B	C	E
3	E	B	C	D	A
4	A	C	D	E	B
5	B	D	E	A	C

Quadrado Latino

Tabela 3: Quadrado Latino após casualização das colunas e linhas

	Colunas				
Linhas	3	5	1	2	4
3	E	B	C	D	A
1	C	E	A	B	D
4	A	C	D	E	B
5	B	D	E	A	C
2	D	A	B	C	E

Modelo

O modelo do experimento em Quadrado Latino é definido por:

$$y_{ijk} = \mu + \alpha_i + \tau_j + \beta_k + \varepsilon_{ijk}, \qquad (1)$$

em que $i = 1, 2, ..., p$; $j = 1, 2, ..., p$; $k = 1, 2, ..., p$; y_{ijk} é a observação da i-ésima linha e k-ésima coluna para o j-ésimo tratamento; μ é a média geral; α_i é o efeito de linha; τ_j é o efeito dos tratamentos; β_k é o efeito de coluna e ε_{ijk} componente de erro aleatório com distribuição $N(0, \sigma^2)$.

Modelo

- As hipóteses de interesse para o modelo (1) são definidas por:

$$\begin{cases} H_0 : \tau_1 = \tau_2 = ... = \tau_p = 0, & \text{(O efeito de tratamento é nulo)} \\ H_1 : \exists \tau_j \neq 0 \end{cases}$$

Modelo

A análise de variância pode ser estendida para o modelo (1) ao particionar a soma quadrados total de $N = p^2$ observações em componentes de linhas, colunas, tratamentos e erros, por exemplo

$$SQ_T = SQ_{Linha} + SQ_{Trat} + SQ_{Coluna} + SQ_{Res}; \tag{2}$$

com respectivo grau de liberdade:

$$p^2 - 1 = (p-1) + (p-1) + (p-1) + (p-2)(p-1)$$

Modelo

- As somas de quadrados também podem ser escritas como:

$$SQ_T = \sum_i \sum_j \sum_k y_{ijk}^2 - \frac{y_{...}^2}{N}$$

$$SQ_{Trat} = \frac{1}{p} \sum_{j=1}^{p} y_{.j.}^2 - \frac{y_{...}^2}{N}$$

$$SQ_{Linha} = \frac{1}{p} \sum_{i=1}^{p} y_{i..}^2 - \frac{y_{...}^2}{N}$$

$$SQ_{Coluna} = \frac{1}{p} \sum_{k=1}^{p} y_{..k}^2 - \frac{y_{...}^2}{N}$$

$$SQ_{Res} = SQ_T - (SQ_{Linha} + SQ_{Trat} + SQ_{Coluna})$$

ANOVA

- Sob a suposição de $\varepsilon_{ijk} \sim N(0, \sigma^2)$, cada soma de quadrados do lado direito da equação (1) dividido por σ^2, são variáveis aleatórias independente com distribuição χ^2.
- Sendo assim, para testar a igualdade das médias de tratamento, a estatística de teste é definida por:

$$F_0 = \frac{QM_{Trat}}{QM_{Res}},$$

com distribuição $F_{p-1,(p-2)(p-1)}$ se a hipótese nula for verdadeira.

- A região crítica é a cauda superior da distribuição F, e rejeitamos H_0 se $F_0 > F_{\alpha,p-1,(p-2)(p-1)}$.

ANOVA

Tabela 4: Tabela de Análise de Variância - Quadrado Latino

Fonte de Variação	SQ	g.l.	QM	E[QM]	F
Tratamentos	SQ_{Trat}	p - 1	QM_{Trat}	$\sigma^2 + \frac{p\sum \tau_j^2}{p-1}$	QM_{Trat}/QM_{Res}
Linhas	SQ_{Linha}	p - 1	QM_{Linha}	$\sigma^2 + \frac{p\sum \alpha_i^2}{p-1}$	
Colunas	SQ_{Coluna}	p - 1	QM_{Coluna}	$\sigma^2 + \frac{p\sum \beta_k^2}{p-1}$	
Resíduo	SQ_{Res}	(p-2)(p-1)	QM_{Res}	σ^2	
Total	SQ_T	$p^2 - 1$			

ANOVA

- Como o número de tratamentos define o número de linhas e de colunas ele também define o número de repetições;
- A grande restrição dos experimentos em quadrado latino é que para 2, 3 ou 4 tratamentos, teremos apenas 0, 2 ou 6 grau(s) de liberdade para o resíduo, respectivamente;
- Por outro lado, com 9 ou mais tratamentos, o quadrado latino fica muito grande trazendo dificuldades na instalação;
- Então, os quadrados latinos mais utilizados são de $5 \times 5, 6 \times 6, 7 \times 7$ e 8×8.

É possível fazer experimento em Quadrado Latino 3×3 ou 4×4?

Replicação do Quadrado Latino

- Quando faz-se necessário utilizar pequenos quadrados latinos, a solução para aumentar os graus de liberdade do erro é replicá-los;
- Um quadrado latino pode ser replicado de várias maneiras:

Caso 1: Usar o mesmo controle local em linhas e colunas;

Caso 2: Usar o mesmo controle local em linha, mas controle local diferente em coluna em cada replicação (ou, de forma equivalente, usar o mesmo controle local em coluna e diferente controle local em linha em cada replicação);

Caso 3: Usar diferente controle local em linhas e colunas;

- Então, a análise de variância depende do método de replicação utilizado.

Relembrando

Tabela 4: Tabela de Análise de Variância para Quadrado Latino

Fonte de Variação	SQ	g.l.	QM	F
Tratamentos	$\frac{1}{p}\sum_{j=1}^{p} y_{.j.}^2 - \frac{y_{...}^2}{N}$	p - 1	QM_{Trat}	QM_{Trat}/QM_{Res}
Linhas	$\frac{1}{p}\sum_{i=1}^{p} y_{i..}^2 - \frac{y_{...}^2}{N}$	p - 1	QM_{Linha}	
Colunas	$\frac{1}{p}\sum_{k=1}^{p} y_{..k}^2 - \frac{y_{...}^2}{N}$	p - 1	QM_{Coluna}	
Resíduo	SQ_{Res}	(p-2)(p-1)	QM_{Res}	
Total	$\sum_i \sum_j \sum_k y_{ijk}^2 - \frac{y_{...}^2}{N}$	$p^2 - 1$		

ANOVA

Tabela 5: Tabela de Análise de Variância para Quadrado Latino replicado - Caso 1

Fonte de Variação	SQ	g.l.	QM	F
Tratamentos	$\frac{1}{np}\sum_{j=1}^{p} y_{.j..}^2 - \frac{y_{....}^2}{N}$	p - 1	QM_{Trat}	$\frac{QM_{Trat}}{QM_{Res}}$
Linhas	$\frac{1}{np}\sum_{i=1}^{p} y_{i...}^2 - \frac{y_{....}^2}{N}$	p - 1	QM_{Linha}	
Colunas	$\frac{1}{np}\sum_{k=1}^{p} y_{..k.}^2 - \frac{y_{....}^2}{N}$	p - 1	QM_{Coluna}	
Resíduo	SQ_{Res}	subtração	QM_{Res}	
Total	$\sum_i \sum_j \sum_k \sum_l y_{ijkl}^2 - \frac{y_{....}^2}{N}$	$np^2 - 1$		

ANOVA

Tabela 6: Tabela de Análise de Variância para Quadrado Latino replicado - Caso 2

Fonte de Variação	SQ	g.l.	QM	F
Tratamentos	$\frac{1}{np}\sum_{j=1}^{p} y_{.j..}^2 - \frac{y_{....}^2}{N}$	p - 1	QM_{Trat}	$\frac{QM_{Trat}}{QM_{Res}}$
Linhas	$\frac{1}{p}\sum_{l=1}^{n}\sum_{i=1}^{p} y_{i..l}^2 - \sum_{l=1}^{n}\frac{y_{...l}^2}{p^2}$	np - 1	QM_{Linha}	
Colunas	$\frac{1}{np}\sum_{k=1}^{p} y_{..k.}^2 - \frac{y_{....}^2}{N}$	p - 1	QM_{Coluna}	
Resíduo	SQ_{Res}	subtração	QM_{Res}	
Total	$\sum_i\sum_j\sum_k\sum_l y_{ijkl}^2 - \frac{y_{....}^2}{N}$	$np^2 - 1$		

- Observe que existem p novas linhas dentro de cada réplica e que a fonte de variação para as linhas realmente mede a variação entre as linhas nas n replicações.

ANOVA

TABELA 7: Tabela de Análise de Variância para Quadrado Latino replicado - Caso 3

Fonte de Variação	SQ	g.l.	QM	F
Tratamentos	$\frac{1}{np}\sum_{j=1}^{p} y_{.j..}^2 - \frac{y_{....}^2}{N}$	p - 1	QM_{Trat}	$\frac{QM_{Trat}}{QM_{Res}}$
Linhas	$\frac{1}{p}\sum_{l=1}^{n}\sum_{i=1}^{p} y_{i..l}^2 - \sum_{l=1}^{n}\frac{y_{...l}^2}{p^2}$	np - 1	QM_{Linha}	
Colunas	$\frac{1}{p}\sum_{l=1}^{n}\sum_{k=1}^{p} y_{..kl}^2 - \sum_{l=1}^{n}\frac{y_{...l}^2}{p^2}$	np - 1	QM_{Coluna}	
Resíduo	SQ_{Res}	subtraçao	QM_{Res}	
Total	$\sum_i\sum_j\sum_k\sum_l y_{ijkl}^2 - \frac{y_{....}^2}{N}$	$np^2 - 1$		

- Observe que a variação resultante das linhas e colunas mede a variação resultante desses fatores nas réplicas.

Suposição de Normalidade

- Pode ser inicialmente verificada ao construir um histograma dos resíduos. Se a suposição sobre os erros for satisfeita, este gráfico deve se parecer com uma amostra de uma distribuição normal centrada em zero;
- Outro recurso gráfico importante é o **Gráfico de Probabilidade Normal dos resíduos** - cada resíduo é colocado em um gráfico versus seu valor esperado sob normalidade.
- Se o resultado gráfico for aproximadamente linear, há evidências da suposição de normalidade.
- Se o resultado gráfico afasta substancialmente da linearidade, não há evidências de que a distribuição do erro é normal.
- Teste de hipótese também pode ser utilizado, como o teste de Shapiro-Wilk.

Suposição de independência

- O objetivo é verificar se existe qualquer correlação entre os erros que são próximos um dos outros em uma sequência;
- Quando os erros são independentes espera-se que os resíduos versus a ordem das observações flutuem em um padrão mais ou menos aleatório em torno da linha de referência Zero.

Variância constante

- Gráficos dos resíduos versus valores ajustados podem ser usados para verificar se a variância dos erros é constante. Pois, os sinais dos resíduos não são importantes para examinar a constância da variância do erro;
- Testes de hipóteses: F, Bartlett e Levene podem ser utilizados com certa parcimônia.

Aditividade

- Teste de hipótese: Teste de aditividade de Tukey

Exercício

Um estudo é configurado para determinar se a resistência à ruptura de uma fibra sintética experimental é afetada pela taxa de tração. Três fiandeiras (1, 2 e 3) foram usadas, cada uma fornecendo fio para uma bobina diferente. As três fiandeiras e suas bobinas correspondentes operavam a três diferentes taxas de estiramento: o normal (A), aumento de 5% sobre o normal (B) e aumento de 10% (C). Quando uma bobina estava cheia, foi retirada da máquina e substituída por uma nova (processo remoção). Ao todo foram 12 remoções.