

INTERPRETAÇÃO DOS DADOS DO PLANEJAMENTO FATORIAL 2^3

13
aula

META

Possibilitar ao aluno a interpretação dos resultados do planejamento fatorial 2^3 , culminando na construção do modelo estatístico.

OBJETIVOS

Ao final desta aula, o aluno deverá:

- calcular o erro associado a cada efeito;
- interpretar a significância estatística dos efeitos, permitindo a triagem de variáveis;
- interpretar graficamente os resultados;
- construir o modelo estatístico e verificar a previsibilidade deste modelo.

PRÉ-REQUISITOS

o aluno deve ter habilidade para elaboração do planejamento fatorial 2^3 , com este número de variáveis desde o início ou por expansão de planejamentos mais simples; cálculo da variância em combinações lineares de variáveis aleatórias.

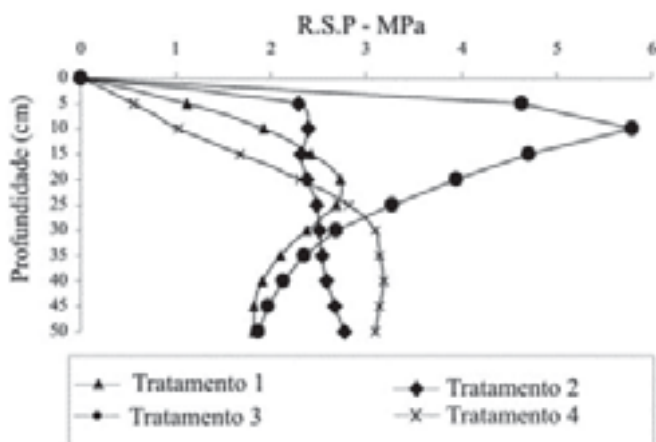


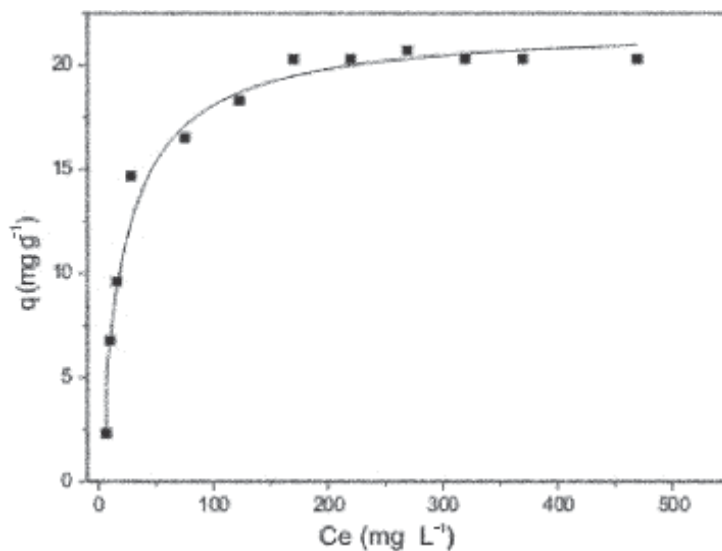
Figura 2. Curvas de resistência do solo a penetração (RSP) para os dois tipos de solo e número de cortes.

Curvas de resistência do solo - reprodução (Fonte: <http://images.google.com.br>)

Deparamo-nos na aula passada com uma situação interessante que foi a possibilidade de que no planejamento 2², que foi feito inicialmente, houvesse outra variável importante e que não havia sido estudada: o tempo. Vimos como é possível, sem ter que realizar todo o estudo novamente, incluir outra variável no pla-

INTRODUÇÃO

nejamento. Nesta aula, vamos avançar na análise dos resultados na direção de verificar se as variáveis estudadas realmente são todas relevantes, o que pode ser feito pelo exame dos erros associados aos efeitos. A importância prática disto é evitar que o pesquisador analise os dados de forma apenas qualitativa “olhando” para os números, o que nem sempre conduz a conclusões confiáveis. Quando comparamos os efeitos calculados com o erro associado a cada um, sempre que o valor de um efeito for próximo ao erro, este não pode ser considerado estatisticamente significativo. Se isso ocorrer para uma das variáveis estudadas, pode ser que esta variável possa ser excluída do estudo, por ser considerada sem importância para o problema.



Isoterma de biossorção do solo – reprodução (Fonte: <http://www.scielo.br>)

Precisamos neste momento obter uma estimativa do erro padrão dos efeitos que foram calculados na última aula. Para isto vamos lembrar que cada ensaio foi realizado em duplicata, de modo que a estimativa conjunta da variância de uma observação individual pode ser calculada com a seguinte expressão:

CÁLCULO DOS ERROS

$$V(y) = s^2 = \sum d_i^2 / 2N$$

onde d_i é a diferença entre as duplicatas de cada ensaio e a média correspondente. Para nossos dados, este valor corresponde a 6,25.

Em um planejamento fatorial 2^3 cada efeito é uma combinação linear de oito valores, com coeficientes $1/4$. Admitindo que estes valores sejam independentes, podemos aplicar a equação:

$$\sigma_y^2 = \sum_i a_i^2 \sigma_i^2$$

assim, os coeficientes a_i^2 ficam $1/16$, correspondentes a 8 observações. Além disso, o valor 6,25 é a variância média de uma observação individual, a variância da média de duas observações será a metade disto e, na somatória, teremos:

$$\sigma_y^2 = 8.(1/16)(6,25/2) = 1,565$$

para cada efeito, o erro padrão será a raiz quadrada deste valor e para a média de todas as observações, será a metade deste último. Os valores obtidos estão na tabela a seguir:

Média	64,35 [±] 0,625
Efeito Principal	
Temperatura (T)	23,0 [±] 1,25
Catalisador (C)	-5,0 [±] 1,25
Tempo (H)	1,5 [±] 1,25
Efeito de Interação (2 Fatores)	
TxH	1,5 [±] 1,25
TxC	10,0 [±] 1,25
HxC	0,0 [±] 1,25
Efeito de Interação (3 Fatores)	
TxHxC	0,5 [±] 1,25

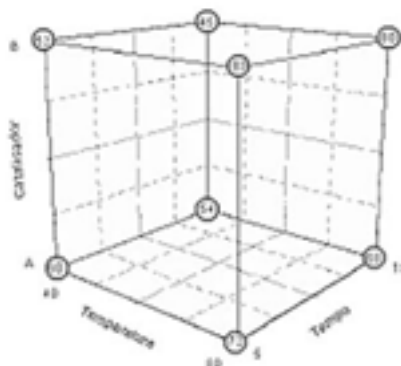
INTERPRETAÇÃO GRÁFICA DO PLANEJAMENTO

Pela observação do gráfico a seguir, podemos destacar que:

Rendimento aumenta com aumento da temperatura, mantendo tempo e catalisadores constantes em ambos os níveis;

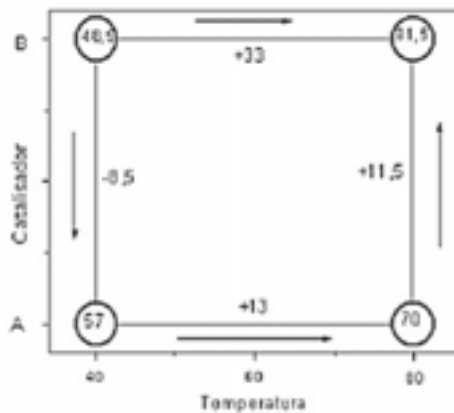
Rendimento diminui com aumento do tempo, mantendo constantes os demais fatores em ambos os níveis;

Rendimento aumenta mudando-se o catalisador de A para B na temperatura mais alta em ambos os níveis de tempo, mas diminui quando se muda o catalisador de A para B na temperatura mais baixa em ambos os níveis de tempo, confirmando que não há interação entre catalisador e tempo.



Além disso, olhando criteriosamente para os dados da tabela anterior contendo os erros, observamos que tanto o efeito principal do tempo quanto os efeitos de interação envolvendo este fator não são estatisticamente significativos, pois são muito próximos aos valores dos erros. A partir desta observação fica claro que o efeito aparente que o pesquisador observou para o tempo acabou por se mostrar irrelevante.

Portanto, somente houve efeito das variáveis que foram estudadas inicialmente no planejamento 2². Na figura abaixo, resumimos graficamente as observações que podem ser realizadas, muito semelhantes às obtidas a partir do próprio planejamento original:



Modelo

Para expressarmos o modelo estatístico que descreve o planejamento, lembre-se de que a forma geral dele é:

$$y(x_1, x_2, x_3) = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 + b_{123}x_1x_2x_3$$

onde:

b_0 : média global

b_1, b_2, b_3 : valores dos efeitos principais

b_{12}, b_{13}, b_{23} : efeitos binários de interação

$$y = 64,25 + 11,5T - 2,5C + 5,0TC$$

Note que este modelo exclui a variável tempo, devido a sua falta de significância estatística. Isso se utilizarmos este modelo, aplicando no lugar das variáveis os valores correspondentes a valo-

res possíveis para as variáveis codificadas (correspondendo ao ensaio 1 e lembrando que o sinal para o cálculo do efeito é dado pelo produto dos sinais dos fatores correspondentes):

$$T = -1; C = -1; TC = +1$$

A estimativa da resposta média segundo o modelo fica $y = 60,25$. Se seguirmos aplicando o modelo para todos os demais ensaios e compararmos o valor previsto pelo modelo com o valor observado experimentalmente, teremos os resultados apresentados na tabela a seguir:

Modelo:

Rendimento estimado (%)	Rendimento (%)	Erro	Erro (%)
60,25	60	0,25	0,42
73,25	72	1,25	1,74
55,25	54	1,25	2,31
68,25	68	0,25	0,37
50,25	52	-1,75	-3,37
83,25	83	0,25	0,30
45,25	45	0,25	0,56
78,25	80	-1,75	-2,19

Podemos perceber que os valores previstos pelo modelo obtido sem a variável tempo, apresenta valores muito próximos aos obtidos experimentalmente. Portanto, considerando que um dos principais objetivos dos métodos quimiométricos é a possibilidade de previsões, fica claro que nosso modelo é muito bem sucedido nesta empreitada.

Concluimos que o cálculo dos erros associados aos efeitos é de extrema importância, pois, ao compararmos os valores de efeitos e de seus erros, é possível que os efeitos sejam considerados estatisticamente insignificantes. Quando isso ocorre para um efeito principal, significa que a variável pode ser descartada, pois não influencia a resposta. Muitas vezes a simples análise qualitativa de um experimento não permite que uma interpretação neste nível possa ser feita com segurança. Portanto, deve ficar destacado como conclusão que somente a análise feita com bases estatísticas sólidas pode permitir que isso seja feito.

CONCLUSÃO

RESUMO



Nesta aula, vamos aprimorar a análise dos resultados na direção de verificar se as variáveis estudadas realmente são todas relevantes, o que pode ser feito pelo exame dos erros associados aos efeitos. A importância prática disto é evitar que o pesquisador analise os dados de forma apenas qualitativa “olhando” para os números, o que nem sempre conduz a conclusões confiáveis. Para isto, vamos comparar o valor dos efeitos calculados com a estimativa do erro associado a cada um, para avaliar sua significância estatística.

Vimos como podemos calcular o erro associado a cada efeito, com segurança dado os ensaios terem sido realizados em duplicata, através da variância. Em um planejamento fatorial 2^3 cada efeito é uma combinação linear de oito valores, com coeficientes $\frac{1}{4}$. Admitindo que estes valores sejam independentes, podemos aplicar a equação que estudamos na aula de combinações lineares de variáveis aleatórias. Assim, no exemplo estudado, vimos que o efeito da variável tempo foi apenas ligeiramente maior do que o erro, mostrando que este efeito não é estatisticamente significativo e, por conseguinte, também não é a variável que ele representa.

Portanto, somente houve efeito das variáveis que foram estudadas inicialmente no planejamento 2². Na figura abaixo resumimos graficamente as observações que podem ser realizadas, muito semelhantes às obtidas a partir do próprio planejamento original.

Com relação ao modelo empírico proposto, este culminou por incluir apenas as variáveis significativas. Finalmente, este modelo foi testado, aplicando-o aos níveis das variáveis significativas e calculando-se respostas previstas. Em seguida, os valores previstos foram comparados com os valores observados, mostrando, no caso, valores bastante próximos. Veremos um pouco mais à frente que este procedimento se chama análise de resíduos, que nada mais é do que construir um modelo a partir de dados experimentais, para em seguida aplicar o modelo às variáveis fixas nos experimentos, calculando-se valores previstos. Finalmente, comparam-se os valores previstos com os observados, para a validação ou não do modelo.

Tabela 2. Otimização da bio sorção de cobre(II) em cascas de nozes pecã. Análise de superfície de resposta empregando planejamento composto central contendo 13 experimentos e 2 fatores (pH, Co). O tempo foi fixado em 2,5 h e a massa do bio sorvente em 30,0 mg

Experimentos	pH	Co	q (mg g ⁻¹)
1	-1	-1	7,647
2	1	-1	15,56
3	-1	1	15,561
4	1	1	10,085
5	-1	0	7,981
6	1	0	6,819
7	0	-1	15,557
8	0	1	9,079
9	0	0	10,173
10	0	0	9,88
11	0	0	10,477
12	0	0	9,115
13	0	0	7,647
		Níveis	
Fatores	-1	0	1
pH	5,0	6,0	7,0
Co (mg L ⁻¹)	20,0	35,0	50,0

Otimização da bio sorção do cobre – reprodução (Fonte: <http://www.scielo.br>)

COMO FAZER UMA REPETIÇÃO LEGÍTIMA DE UM EXPERIMENTO

Quando vamos fazer uma análise química, por exemplo, e sabemos que é imprescindível a realização de repetições, é necessário que estas repetições sejam realmente o que chamamos de repetições verdadeiras. Para isto, devem-se repetir todas as etapas, desde o início. Por exemplo, se for a análise de um material em pó, a partir do qual uma amostra deve ser pesada e dissolvida, para se analisar o líquido, deve-se partir da etapa inicial de pesagem de uma nova amostra. Se a pessoa já dispõe de uma amostra dissolvida, com a qual já realizou uma análise, não basta simplesmente fazer outra análise desta solução. Isto não é uma repetição verdadeira e o resultado não pode ser usado para um cálculo de variância confiável.

PRÓXIMA AULA



Mais adiante estudaremos as limitações e possibilidades de uso de planejamentos fatoriais com números relativamente elevados de variáveis, que exigem numerosos ensaios.

REFERÊNCIAS

- BOX, G. E. P.; HUNTER, W. G.; HUNTER, J. S. Statistics for experimenters. An introduction to design, data analysis and model building. New York: Wiley, 1978.
- BUSSAB, W. O.; MORETIN, P. A. Estatística básica. São Paulo, Ed. Atual, 1985.
- BARROS NETO, B.; SCARMINIO, I. E.; BRUNS, R. E.; Planejamento e otimização de experimentos. Editora da Unicamp: Campinas, 1995.