A Lei de Mitscherlich e a Análise da Variância em Experiências de Adubação

FREDERICO PIMENTEL GOMES

Escola Superior de Agricultura "Luiz de Queiroz",
Universidade de São Paulo

ÍNDICE

1 — INTRODUÇÃO

A análise da variância em certos casos é perturbada pela existência da correlação. Tal se dá, por exemplo, quando se experimentam doses diferentes de um mesmo adubo. Nêste caso, é preciso levar em conta a correlação, sem o que a análise da variância pode conduzir a resultados falsos.

Em geral se supõe que a regressão é linear. Em experiências de adubação, porém, é frequente o caso da existência de regressão não linear, geralmente do tipo exponencial introduzido por Mitscherlich. Êste trabalho tem por fim especial estudar a análise da variância nesse caso.

2 — A CORRELAÇÃO LINEAR E A ANÁLISE DA VARIÂNCIA

Consideremos os seguintes dados (fictícios) onde se supõem 5 tratamentos e 4 repetições. Os valores de x são as doses de adubo usadas e os Σ y são os totais de cada tratamento.

x	0	1	2	3	4
Σ y	2	2	3	4	5

Uma análise da variância segundo o esquema corrente nos daria :

	Soma dos quadrados dos desvios	Grau de liberdade	Variância
Tratamentos	1,7	4	0,425
Resíduo	2,1	15	0,140
Total	3,8	19	

O valor de ϑ será

$$\vartheta = \sqrt{\frac{0,425}{0,14}} = 1,74, \text{ insignificante.}$$

O fato de não levarmos em conta a correlação nos levaria a êsse resultado em desacôrdo com a realidade e equipararia os dados em apreço, onde é evidente a influência da adubação, aos seguintes, por exemplo, onde tal influência não aparece.

x	0	1	2	3	4
Σ y	3	5	2	4	2

Consideremos, porém, a correlação e vamos obter

$r = 0{,}65$, significativo.

A equação de regressão é

$y = 0{,}2x + 0{,}4,$

com a qual calculamos as médias esperadas.

x	0	1	2	3	4
Médias esperadas	0,40	0,60	0,80	1,00	1,20
Médias observadas	0,50	0,50	0,75	1,00	1,25

A soma dos quadrados dos desvios entre as médias esperadas e as observadas dá 0,025, número que multiplicamos por 4 (nº. de repetições). A soma dos quadrados dos desvios relativos aos tratamentos se decompõe, então, em duas partes :

	Soma dos quadrados	Grau de liberdade	Variância
Regressão linear	1,6	1	1,600
Desvios a partir dos valores esperados	0,1	3	0,033
Tratamentos	1,7	4	0,425

Temos agora

$$\vartheta = \sqrt{\frac{1{,}60}{0{,}14}} = 3{,}38, \text{ significativo.}$$

O cálculo numérico poderia ser feito de maneira mais simples (1, pp. 41-43), mas a marcha seguida é mais conveniente para a compreensão do que se segue.

Como exemplo objetivo, damos a seguir a análise dos dados de uma experiência de adubação fosfatada de milho, realizada em Campinas pelo Engenheiro-Agrônomo Glauco Pinto Viegas.

x	0	0,25	0,50	0,75	1,00
y	2,82	3,34	3,90	4,10	4,68
	2,40	4,52	3,64	3,56	6,46
	2,80	2,80	3,36	4,30	4,48
	1,78	4,74	3,68	4,14	3,94
Σ	9,80	15,40	14,58	16,10	19,56

Cada linha horizontal representa um bloco Os dados se referem à produção em quilos por canteiro de 30m². O x está expresso em quintais métricos de P_2O_5 por hectare sob a forma de superfosfato. Além da adubação fosfatada, fez-se também uma adubaçao geral de 25 kg/ha de N sob a forma de salitre do Chile e 45 kg/ha de K_2O sob a forma de KCl.

Temos a seguinte análise da variância, onde não isolamos a parte da variação devida aos blocos por não haver vantagem neste caso, devido à uniformidade notável dos mesmos.

	Soma dos quadrados dos desvios	Grau de liberdade	Quadrado médio
Tratamentos	12,0303	4	3,01
Resíduo	7,6572	15	0,51
Total	19,6875	19	

$$\vartheta = \sqrt{\frac{3{,}01}{0{,}51}} = 2{,}43.$$

Os limites da tabela de ϑ (2, p. 523) são:

1‰ 2,87,
1% 2,21,
5% .. .. 1,75.

Logo o valor obtido é significativo para 1%, mas não para 1‰.

A consideração da correlação nos dá, porém,

r = 0,72, significativo.

E a equação de regressão é

y = 2,02x + 2,76.

A soma dos quadrados dos desvios devidos aos tratamentos x se decompõe como se segue:

	Soma dos quadrados	Grau de liberdade	Quadrado médio
Regressão linear	9,9035	1	9,90
Desvios a partir dos valores esperados	2,1268	3	0,71
Tratamentos	12,0303	4	3,01

Temos agora

$$\vartheta = \sqrt{\frac{9,90}{0,51}} = 4,41.$$

Êste valor de ϑ é superior ao limite de 1‰, que é 4,05. Fica evidente, pois, a vantagem do isolamento da parcela correspondente à correlação.

3 — O CASO DA CORRELAÇÃO NÃO LINEAR

No caso de uma correlação não linear como, por exemplo, a que obedece à lei de Mitscherlich, torna-se, agora, evidente, a necessidade de uma análise da variância que leve em conta êsse fato. Utilizando os dados observados e com o auxílio das fórmulas já conhecidas (3, pp. 200-205) podemos calcular os parâmetros da função de Mitscherlich

$$y = A\,[1 - 10^{-c\,(x + b)}]$$

por meio do método dos quadrados mínimos ou do método dos momentos. Torna-se possível, então, calcular os valores esperados e, a seguir, decompor a soma dos quadrados dos desvios relativos aos tratamentos em duas partes, uma devida à regressão e outra relativa aos desvios a partir da curva de regressão.

Suponhamos um caso de 5 doses diferentes de um mesmo fertilizante e 4 repetições. A soma dos quadrados dos desvios atribuidos aos tratamentos é:

$$4 \sum_{j=1}^{5} (\bar{y}_j - \bar{\bar{y}})^2$$

onde $\bar{\bar{y}}$ é a média geral, e $\bar{y}_j$ é a média de cada tratamento. Seja $\hat{y}_j$ o valor esperado de acôrdo com a equação de regressão. Temos então:

$$\sum_{j=1}^{5} (\bar{y}_j - \bar{\bar{y}})^2 = \sum_{j=1}^{5} (\bar{y}_j - \hat{y}_j + \hat{y}_j - \bar{\bar{y}})^2,$$

$$= \sum_{j=1}^{5} (\bar{y}_j - \hat{y}_j)^2 + \sum_{j=1}^{5} (\hat{y}_j - \bar{\bar{y}})^2 + 2\sum_{j=1}^{5} (\bar{y}_j - \hat{y}_j)(\hat{y}_j - \bar{\bar{y}})$$

Mas o último termo nos dá

$$\sum_{j=1}^{5} (\bar{y}_j - \hat{y}_j)(\hat{y}_j - \bar{\bar{y}}) = \sum_{j=1}^{5} (\bar{y}_j - \hat{y}_j)\hat{y}_j - \bar{\bar{y}} \sum_{j=1}^{5} (\bar{y}_j - \hat{y}_j)$$

Temos porém

$$4 \sum_{j=1}^{5} (\bar{y}_j - \hat{y}_j) = \sum_{i=1}^{20} y_i - \sum_{1}^{20} A\left[1 - 10^{-c(x_i + b)}\right]$$

Mas, quer no caso do método dos quadrados mínimos, quer no do método dos momentos, o segundo membro da última igualdade é nulo. Logo temos:

$$4\sum_{j=1}^{5}(\bar{y}_j-\hat{y}_j)(\hat{y}_j-\bar{y}) = 4\sum_{j=1}^{5}(\bar{y}_j-\hat{y}_j)\hat{y}_j$$

$$(3,1) \qquad = \sum_{i=1}^{20}\left\{y_i - A\left[1-10^{-c(x_i+b)}\right]\right\} A\left[1-10^{-c(x_i+b)}\right]$$

No caso do método dos quadrados mínimos a última expressão se anula (3, p. 201). Logo neste caso temos :

$$(3,2) \quad \sum_{j=1}^{5}(\bar{y}_j-\bar{y})^2 = \sum_{j=1}^{5}(\bar{y}_j-\hat{y}_j)^2 + \sum_{j=1}^{5}(\hat{y}_j-\bar{y})^2$$

Fica, assim, a soma dos quadrados dos desvios das médias parciais dos tratamentos repartida em duas parcelas, uma correspondente aos desvios das médias esperadas em relação à média geral, outra relativa aos desvios entre os valores observados e os valores esperados.

Quando se aplica o método dos momentos, porém, nada nos garante que a somatória de (3,1) se anule. Logo a igualdade expressa em (3,2) não se verifica necessàriamente e a decomposição por ela indicada é falha.

Vejamos um exemplo.

Os dados seguintes se referem a uma experiência de adubação de milho realizada em Ipanema pelos Engenheiros-Agrônomos Glauco Pinto Viegas e Erik Smith, segundo o esquema adotada para a experiência que vimos atrás.

x	0	0,25	0,50	0,75	1,00
y	3,38	7,15	10,07	9,55	9,14
	5,77	9,78	9,73	8,95	10,17
	4,90	9,99	7,92	10,24	9,75
	4,54	10,10	9,48	8,66	9,50
Σ y	18,59	37,02	37,20	37,40	38,56

A análise da variância nos conduziu aos resultados seguintes, onde não se isolou a variação atribuída aos blocos por serem êstes muito uniformes, como o leitor poderá verificar.

	Soma dos quadrados dos desvios	Grau de liberdade	Quadrado médio
Tratamentos	72,2199	4	18,055
Resíduo	13,6482	15	0,910
Total	85,8681	19	

$$\vartheta = \sqrt{\frac{18,055}{0,910}} = 4,45, \text{ significativo.}$$

Se procurarmos o coeficiente de correlação, acharemos
r = 0,688, significativo.

No entanto a correlação linear dá apenas uma aproximação muito grosseira, como mostram os dados e a análise que se seguem.

Correlação Linear

x	0	0,25	0,50	0,75	1,00
Valores esperados	6,422	7,430	8,438	9,446	10,454
Valores observados	4,648	9,255	9,300	9,350	9,640

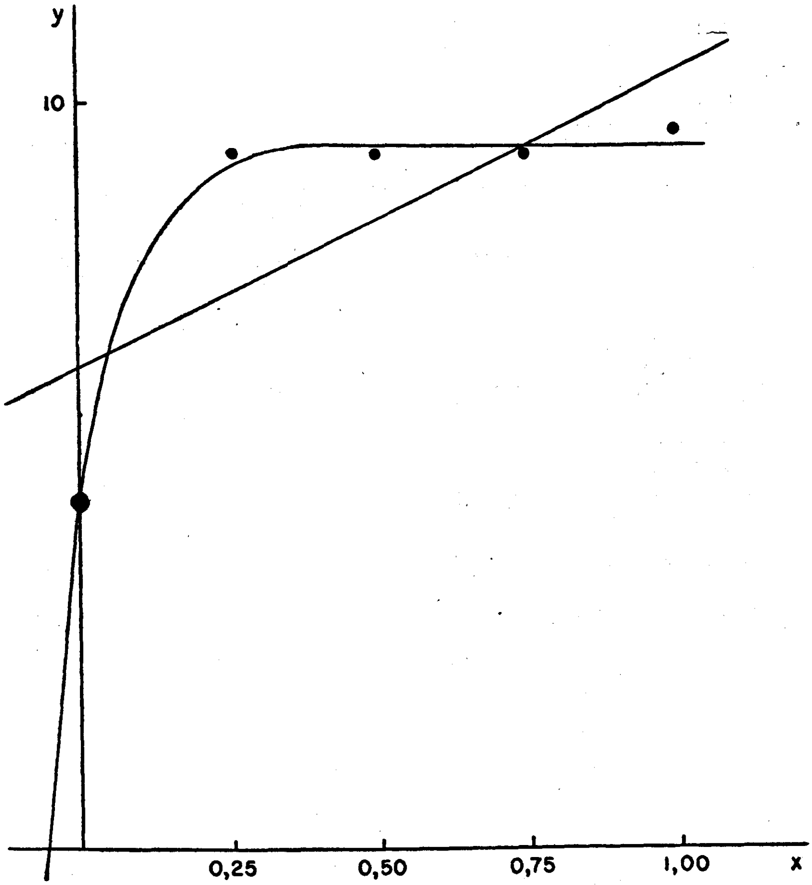
	Soma dos quadrados	Grau de liberdade	Quadrado médio
Regressão linear	40,6499	1	40,650
Desvios a partir dos valores esperados	31,5700	3	10,523
Tratamentos	72,2199	4	

$$\vartheta = \sqrt{\frac{40,650}{0,910}} = 6,68, \text{ significativo.}$$

Porém os desvios a partir dos valores esperados nos dão uma nova estimativa da variância, que não deve diferir significativamente da que nos dá o resíduo. No entanto temos

$$\vartheta = \sqrt{\frac{10,523}{0,910}} = 3,4, \text{ significativo,}$$

o que nos mostra que a correlação linear é, neste caso, pouco satisfatória. Aliás, o gráfico anexo confirma o que aí fica dito.



O gráfico acima representa a reta de regressão calculada, bem como a curva de Mitscherlich interpolada pelo método dos quadrados mínimos, e ainda os valores observados, que estão indicados por pontos.

O método dos quadrados mínimos aplicado à lei de Mitscherlich nos dá, porém, para equação de regressão

$$y = 9{,}436 \left[1 - 10^{-5{,}590\,(x + 0{,}0527)} \right],$$

com a qual calculamos os valores esperados que figuram no quadro seguinte

Método dos quadrados mínimos

x	0	0,25	0,50	0,75	1,00
Valores esperados	4,648	9,244	9,428	9,436	9,436
Valores observados	4,647	9,255	9,300	9,350	9,640

A análise da variância nos dá então o seguinte quadro.

	Soma dos quadrados	Grau de liberdade	Quadrado médio
Regressão pela lei de Mitscherlich	71,9432	2	35,978
Desvios a partir dos valores esperados	0,2621	2	0,131
Tratamentos	72,2199	4	18,055

A soma 71,9432 + 0,2621 difere um pouco do total dos tratamentos 72,2199 devido aos pequenos desvios inerentes às aproximações feitas.

Temos agora

$$\vartheta = \sqrt{\frac{35{,}978}{0{,}910}} = 6{,}29, \text{ significativo,}$$

$$\vartheta = \sqrt{\frac{0{,}131}{0{,}910}} = 0{,}379, \text{ não significativo}$$

A análise da variância nos mostra, pois, que a correlação obedece muito bem à lei de Mitscherlich.

Se quizéssemos isolar a influência da variação entre blocos, teriamos o quadro a seguir. Como já dissemos, isso não convém neste caso, mas em geral é vantajoso.

	Soma dos quadrados	Grau de liberdade	Quadrado médio
Blocos	2,7349	3	0,912
Regressão pela lei de Mitscherlich	71,9432	2	35,978
Desvios a partir dos valores esperados	0,2621	2	0,131
Resíduo	10,9133	12	0,909
Total	85,8681	19	

A validez de (3,2) é um bom argumento a favor do método dos quadrados mínimos. Além disso a interpolação por êsse método nos dá a melhor concordância possível entre os valores esperados e os observados e, se admitirmos que seja normal a distribuição dos desvios entre êsses valores, êle é equivalente à determinação dos parâmetros pelo método da máxima verossimilhança ("maximum likelihood") (4, 2°. vol., p. 59). O método dos momentos, porém, conduz a uma equação muito mais simples. Por exemplo no caso em estudo o método dos quadrados mínimos nos dá uma equação do quarto grau em $z = 10^{-0,25c}$, ao passo que o método dos quadrados mínimos nos leva, depois de operações muito mais laboriosas, e uma equação do décimo grau. E' possível, portanto, que muitos dêem preferência ao método dos momentos que, no exemplo em estudo, nos conduz à equação

$$y = 9{,}553 \left[1 - 10^{-3{,}470\,(x + 0{,}086)} \right]$$

Com esta equação obtivemos os valores esperados do quadro seguinte.

Método dos momentos

x	0	0,25	0,50	0,75	1,00
Valores esperados	4,748	8,901	9,465	9,541	9,552
Valores observados	4,647	9,255	9,300	9,350	9,640

Para avaliar o grau de concordância da curva calculada podemos obter a soma dos quadrados dos desvios entre os valores observados e os esperados

$$\Sigma\ (\bar{y}j \ - \ \hat{y}j)^2.$$

Esta soma, multiplicada pelo número de repetições e dividida pelo grau de liberdade, nos dá uma estimativa da variância, que não deverá diferir significativamente da que se obtém do resíduo.

No exemplo em estudo temos

$$\frac{4\ \Sigma\ (\bar{y}j \ - \ \hat{y}j)^2}{2} = 0.414$$

Logo

$$\vartheta = \sqrt{\frac{0{,}414}{0{,}910}} = 0{,}674, \text{ não significativo.}$$

A interpolação é, portanto, satisfatória.

4 — UMA DIGRESSÃO INTERESSANTE

Em um trabalho anterior (5, pp. 7-9) apresentamos uma fórmula para o cálculo da adubação econômicamente aconselhável. Tal fórmula era

$$\log \frac{f\ t\ 0{,}4343}{s\ c} = \log A - c\ (x + b),$$

onde x é a dose de adubo a ser usada, A, b, c são os parâmetros da equação de Mitscherlich, t é o preço de um quintal métrico de adubo, s é o preço de igual quantidade do produto agrícola obtido, e f é um fator maior que um.

A aplicação dessa fórmula aos dados em estudo, suposto o preço de 750 cruzeiros por quintal-métrico de P_2O_5 e de 100 cruzeiros por quintal de milho e ainda tomando-se $f = 1{,}5$, nos dá

$$x = 0{,}132$$

no caso do método dos quadrados mínimos. Conclui-se, então, que a adubação mais aconselhável seria a de cêrca de 13 quílos de P_2O_5 por hectare, isto é, uns 72 quilos de superfosfato com 18% de P_2O_5 por hectare. Esta adubação é evidentemente pobre. Isso se deve aos seguintes fatores:

I. Elevado preço do adubo;
II. Baixo preço do milho;
III. Valor elevado de c.

Com efeito o valor de c obtido, corresponde, se referido a quintais-métricos de milho por hectare, a 1,677, quando Mitscherlich obtinha em média, na Alemanha, 0,60 e CROWTHER (6, p. 96) conseguia na Grã-Bretanha um valor correspondente a 0,64. Porém SARAIVA (7, p. 26) obteve no Rio de Janeiro um valor de c que, reduzido às unidades usadas por Mitscherlich, corresponde a 1,532, valor que pouco difere do que agora indicamos.

O método dos momentos nos dá para dose mais aconselhável de adubo $x = 0{,}152$, isto é, cêrca de 15 quilos de P_2O_5 por hectare.

5 — CONCLUSÕES

A análise da variância pode ser utilizada com vantagens no estudo das experiências de adubação com o auxílio da lei de Mitscherlich. Ela nos permite avaliar a precisão da concordância entre os valores esperados e os observados e verificar se há ou não vantagem e conveniência na aplicação daquela lei. Uma vez obtida uma concordância razoável, podem-se tirar conclusões bem fundadas sôbre a adubação econômicamente mais aconselhável. Ao contrário a aplicação indiscriminada da lei de Mitscherlich, como de qualquer processo interpolatório, sem os cuidados necessários, é suspeita.

6 — AGRADECIMENTO

Agradecemos as indicações do Prof. W. L. Stevens que, além de examinar a nossa sugestão sôbre a análise da variância no caso de aplicação da lei de Mitscherlich, ainda nos deu conhecimento do importante e moderno trabalho de CROWTHER (6). Agradecemos também cordialmente o auxílio do Engenheiro-Agrônomo Izaías Rangel Nogueira, Assistente de Matemática da E. S. A. "Luiz de Queiroz", nos laboriosos cálculos realizados, e a colaboração de minha esposa no tedioso trabalho datilográfico. E finalmente, "last but not least", ficamos gratos ao Engenheiro-Agrônomo Glauco Pinto Viegas pelos dados que gentilmente nos forneceu.

7 — BIBLIOGRAFIA CITADA

1 — WISHART, J. e H. G. Sanders — "Princípios e Prática de Experimentação de Campo". Tradução de G. P. Viegas. S. Paulo.
2 — BRIEGER, F. G. — "Limites Unilaterais e Bilaterais na Análise Estatística". Bragantia 6, pp. 479-545, 1946. Campinas.
3 — PIMENTEL GOMES, Frederico e Eurípedes Malavolta — "Aspectos Matemáticos e Estatísticos da Lei de Mitscherlich". Anais da Escola Superior de Agricultura "Luiz de Queiroz", vol. 6, (1949), pp. 193-229. Piracicaba.
4 — KENDALL, Maurice G. — "The Advanced Theory of Statistics". 3a. edição, 1947. Londres.
5 — PIMENTEL GOMES, Frederico e Eurípedes Malavolta — "Considerações Matemáticas sôbre a Lei de Mitscherlich". 1949. Piracicaba.
6 — CROWTHER, E. M. e F. Yates — "Fertilizer Policy in War-Time: The Fertilizer Requirements of Arable Crops". The Empire Journal of Experimental Agriculture, vol. IX, 1941, pp. 77 — 97. Oxford.
7 — SARAIVA, Mario, Admar Lopes da Cruz e Carlos del Negro — "Contribuição para o Estudo dos Methodos de Mitscherlich, Wiessmann e Neubauer". 1937. Rio de Janeiro.