ISSN 0034-765

REVISTA DE AGRICULTURA

Vol. 68 Junho/1993 Nº 1

ESTUDO DA DISTRIBUIÇÃO DAS COORDENADAS DO PONTO CRÍTICO DA EQUAÇÃO QUADRÁTICA DE REGRESSÃO A DUAS VARIÁVEIS INDEPENDENTES

Marli de Bem Gomes[1]

INTRODUÇÃO

Na pesquisa científica agronômica é bastante comum o ajustamento de superfície de resposta a dados de experimentos de adubação. Geralmente procura-se determinar o ponto crítico de interesse, mais comumente o ponto de máximo, sendo porém freqüentemente encontrados pontos de sela, e às vezes até de mínimo. Outra dificuldade séria porém é que os estimadores das coordenadas do ponto crítico são variáveis aleatórias de distribuição desconhecida, que pretendemos estudar neste trabalho.

Há muitos artigos que poderíamos citar neste campo porém vamos nos restringir àqueles que têm relação mais direta com este trabalho.

No estudo do ajuste de superfície de resposta a dados de ensaios de adubação, MASON (1956), TEJEDA (1966), CAMPOS (1967) e MORAES (1969) constataram que os intervalos de confiança obtidos para os parâmetros foram muito amplos.

[1] Professora Associada do Departamento de Matemática e Estatística da ESALQ/USP, Piracicaba-SP.

ZAGATTO & PIMENTEL-GOMES (1967), em estudo das fun-
)es de produção sob o ponto de vista econômico, abordaram
ınções com uma, duas ou mais variáveis independentes,
›licando-as a ensaios fatoriais de 3^3. Indicaram uma sé-
ie de dificuldades e sugeriram:

1) Usar sempre grupos de ensaios numerosos ou ensaios
;olados com muitas repetições e boa precisão.

2) Não confiar em doses ótimas obtidas a partir de po
inômios cujos coeficientes para os termos de 2º grau nao
ejam significativamente diferentes de zero.

3) Verificar se o ponto crítico obtido corresponde
ealmente a um máximo.

4) Calcular intervalos de confiança para as doses óti
as encontradas.

PIMENTEL-GOMES & CONAGIN (1987) resumiram com exem-
los agronômicos a teoria de curvas e superfícies de res-
›sta. Citam dois métodos para cálculo do intervalo de con
iança para a dose econômica de adubação, no caso do poli-
›mio de 2º grau, com uma só variável independente. Mas es
es métodos só são satisfatórios quando o coeficiente $\hat{a}_2$,
› termo de segundo grau, difere significativamente de ze-
›.

GOMES (1989) estudou com detalhes as variâncias e a
eterminação de intervalos de confiança para as coordena-
as da dose ótima de adubação. Obteve, entre outras coi-
as, intervalos de confiança para a_{11} e a_{22} e aplicou tes-
t para comparar suas estimativas com zero. Isto para o
juste de uma superfície de 2º grau com duas variáveis in-
ependentes.

Neste trabalho foram simulados experimentos fatoriais
3^2, obtidas as estimativas dos parâmetros do polinômio
segundo grau ajustado, seus pontos críticos, determina-
›s intervalos de confiança para a_{11} e a_{22}, e os coeficien
es de assimetria e de curtose para as estimativas das coor
enadas dos pontos críticos obtidos.

ATERIAL E MÉTODOS

Vamos considerar um experimento fatorial de 3^2, onde tenhamos dois nutrientes em níveis indicados por X_1 e X_2, com três doses para cada um deles: 0, 40, 80 kg/ha. Portanto a produção Y é uma função:

$$\hat{Y} = f\ (X_1,\ X_2)$$

e poderemos escrever que:

$$\hat{Y} = \hat{a}_{33} + 2\hat{a}_{13}X_1 + 2\hat{a}_{23}X_2 + \hat{a}_{11}X_1^2 + \hat{a}_{22}X_2^2 + 2\hat{a}_{12}X_1X_2,$$

no caso do polinômio do segundo grau. Há, então, porém, correlação entre os coeficientes. Por isso, vamos substituir as variáveis por polinômios ortogonais adequados. De acordo com GOMES (1992) os estimadores dos 6 parâmetros, da nova equação, passam a ser independentes, isto é, têm covariâncias nulas.

Usamos as mesmas produções de milho com adubação de N e P, com doses 0, 40 e 80 kg/ha para cada um dos nutrientes, citadas por GOMES (1992). E a simulação foi feita da mesma forma, porém aqui obtivemos uma série de 200 repetições e duas de 50 repetições, para os seguintes valores de σ: 10, 50, 100, 200, 300, 400, 500 e 600. Com esses valores fizemos os cálculos a seguir.

1) Estimação dos parâmetros da equação (1) e obtenção, para cada σ, das médias das 200 estimativas dos parâmetros. As variâncias desses estimadores são conhecidas, pois os dados foram gerados. Com elas calculamos intervalos de confiança para os coeficientes a_{11} e a_{22}.

2) Cálculo dos pontos críticos e verificação da sua natureza, em todos os experimentos gerados.

3) Cálculo dos coeficientes de assimetria e de curtose para $\hat{x}_1^*$ e $\hat{x}_2^*$, que são, respectivamente, as estimativas das coordenadas dos pontos críticos, isto é, $\hat{\gamma}_1$ e $\hat{\gamma}_2$ de acordo com:

$$\hat{\gamma}_1 = \frac{\hat{\mu}_3}{\hat{\sigma}^{\,3}(\hat{x}_i^*)}, \qquad \hat{\gamma}_2 = \frac{\hat{\mu}_4}{\hat{\sigma}^{\,4}(\hat{x}_i^*)} - 3,$$

onde $\hat{\mu}_3$ é o terceiro momento de $\hat{x}_i^*$ em relação à média e $\hat{\mu}_4$ é o quarto momento, e $i=1,\ 2$.

Por KENDALL & STUART (1963), calculamos:

$$V(\hat{\gamma}_1) = \frac{6n(n-1)}{(n-2)\ (n+1)\ (n+3)},$$

$$V(\hat{\gamma}_2) = \frac{24n(n-1)^2}{(n-3)\ (n-2)\ (n+3)\ (n+5)},$$

onde n=200 no nosso caso.

Daí vamos aplicar o teste **t**, sendo:

$$t_1 = \frac{\hat{\gamma}_1 - 0}{\sqrt{V(\hat{\gamma}_1)}}, \qquad t_2 = \frac{\hat{\gamma}_2 - 0}{\sqrt{V(\hat{\gamma}_2)}}.$$

4) Com as duas séries geradas de 50 repetições calculamos apenas as estatísticas:

$$Z_1 = \frac{\hat{a}_{11} - 0}{\sigma(\hat{a}_{11})}, \qquad Z_2 = \frac{\hat{a}_{22} - 0}{\sigma(\hat{a}_{22})}.$$

onde $\sigma(\hat{a}_{11})$ e $\sigma(\hat{a}_{22})$ é o valor teórico obtido por:

$$\sigma(\hat{a}_{11}) = \sqrt{\sigma^2/2} \quad \text{e} \quad \sigma(\hat{a}_{22}) = \sqrt{\sigma^2/2}\ .$$

RESULTADOS E DISCUSSÃO

Estudamos a natureza dos pontos críticos dos dados gerados e obtivemos os resultados da **TABELA I**. Por esta tabela notamos que os pontos de mínimo começam a aparecer a partir de $\sigma \geq 200$, isto é, a partir de um coeficiente de variação de apenas 4,17%. E para $\sigma = 500$ (C.V. = 10,42%) já tivemos apenas 68 pontos de máximo, além de 23 pontos de mínimo, em 200 experimentos gerados.

Saliente-se o fato de que estudar os dados gerados com σ = 10, 50, 100, 200, 300, 400, 500 e 600 equivale a tomar valores provenientes de ensaios com 3600; 144; 36; 9; 4;

2,25; 1,44 e 1 repetições. Portanto, os pontos de mínimo começam a surgir com r ≤ 9 repetições, isto é $\sigma \geq 200$.

TABELA I. Resultados obtidos com relação à natureza dos pontos críticos.

σ	C.V.(%)	P. de Máximo	P. de Sela	P. de Mínimo
10	0,21	200	0	0
50	1,04	172	28	0
100	2,08	142	58	0
200	4,17	107	92	1
300	6,25	83	105	12
400	8,33	73	118	9
500	10,42	68	109	23
600	12,50	67	115	18

Vamos nos preocupar apenas com as médias e intervalos de confiança dos parâmetros a_{11} e a_{22} que, segundo ZAGATTO & PIMENTEL-GOMES (1967), são os mais importantes. Os verdadeiros valores desses parâmetros são:

$$a_{11} = -228,05 \qquad \text{e} \qquad a_{22} = -36,06.$$

Observando as **TABELAS II** e **III**, verifica-se que a estimativa $\bar{a}_{11}$ teve uma pequena dispersão em relação ao seu valor real. Já a estimativa $\bar{a}_{22}$ teve, para $\sigma = 400$ e $\sigma = 600$, valores bem menores que seu valor real.

Embora tenhamos nos preocupado mais com a_{11} e a_{22}, é interessante observar que a estimativa $\bar{a}_{12}$ teve sempre grande dispersão. As estimativas dos demais parâmetros, porém, tiveram pequena dispersão em torno do seu valor real. Os erros padrões das médias estimadas de todos os parâmetros estiveram sempre próximos dos valores verdadeiros.

TABELA II. Médias e intervalos de confiança ao nível de 5% de probabilidade para a_{11} usando $\sigma(\bar{a}_{11}) = \sigma/20$.

σ	C.V.(%)	$\bar{a}_{11}$	I.C. para a_{11}
10	0,21	-228,20	[-229,18; -227,22]
50	1,04	-231,00	[-235,90; -226,10]
100	2,08	-230,70	[-240,50; -220,90]
200	4,17	-222,60	[-242,20; -203,00]
300	6,25	-218,80	[-248,20; -189,40]
400	8,33	-227,20	[-266,40; -188,00]
500	10,42	-208,30	[-257,30; -159,30]
600	12,50	-214,80	[-273,60; -156,00]

TABELA III. Médias e intervalos de confiança ao nível de 5% de probabilidade para a_{22} usando $\sigma(\bar{a}_{22}) = \sigma/20$.

σ	C.V.(%)	$\bar{a}_{22}$	I.C. para a_{22}
10	0,21	-36,06	[-37,04; -35,08]
50	1,04	-33,89	[-38,79; -28,99]
100	2,08	-40,39	[-50,19; -30,59]
200	4,17	-34,16	[-53,76; -14,56]
300	6,25	-36,17	[-65,57; -6,77]
400	8,33	-51,10	[-90,30; -11,90]
500	10,42	-29,54	[-78,54; 19,46]
600	12,50	-63,97	[-122,77; -5,17]

Todos os intervalos de confiança calculados para a_{11} e a_{22} contêm os verdadeiros valores. Porém, para a_{11} os intervalos obtidos com $\sigma = 500$ e $\sigma = 600$ são muito amplos. Para a_{22} estes intervalos se tornam excessivamente amplos a partir de $\sigma = 200$.

Os resultados obtidos para os coeficientes de assimetria ($\hat{\gamma}_1$) e de curtose ($\hat{\gamma}_2$) para x_1^* e x_2^*, assim como os valores de **t** calculados, estão nas **TABELAS IV** (para $\hat{x}_1^*$) e **V** (para $\hat{x}_2^*$).

TABELA IV. Resultados de γ_1 e γ_2 obtidos para o valor de $\hat{x}_1^*$ e teste **t** correspondente.

σ	γ_1	γ_2	$t(\gamma_1)$	$t(\gamma_2)$
10	0,2290	-0,2170	1,33	-0,634
50	10,778	119,50	62,69***	349,22***
100	6,823	66,49	39,69***	194,31***
200	12,924	171,17	75,17***	500,21***
300	12,540	169,65	72,94***	495,76***
400	-13,257	180,46	-77,11***	527,40***
500	-3,222	55,00	-18,74***	160,72***
600	13,717	189,15	79,78***	552,75***

Nota: Foram obtidas $V(\hat{\gamma}_1) = 0{,}029558$ e $V(\hat{\gamma}_2) = 0{,}117172$, usadas para o cálculo do teste **t** desta e da **TABELA V**.

Pela **TABELA IV** verificamos que γ_1 para a coordenada x_1^* só tem sinal negativo para $\sigma = 400$ e 500. Portanto, em 6 casos a assimetria é positiva. Para $\sigma = 10$ a distribuição é praticamente simétrica, pois $\gamma_1 = 0{,}2290$ é bem próximo de zero. Olhando γ_2, coeficiente que mede a curtose, verificamos que para 7 valores de σ a distribuição é leptocúrtica, isto é, γ_2 é positivo. Só para $\sigma = 10$ a distribuição é platicúrtica.

Na **TABELA V**, relativa à coordenada do ponto crítico

x_2^*, obtivemos para γ_1 5 valores negativos e 3 positivos, predominando portanto a assimetria à esquerda. O cálculo de γ_2 nos revelou que a distribuição é leptocúrtica em todos os casos.

TABELA V. Resultados de γ_1 e γ_2 obtidos para o valor de $\hat{x}_2^*$ e teste **t** correspondente.

σ	γ_1	γ_2	$t(\gamma_1)$	$t(\gamma_2)$
10	1,213	2,382	7,06***	6,96***
50	-6,892	117,585	-40,09***	343,61***
100	5,638	64,532	32,79***	188,58***
200	-13,956	103,178	-81,18***	564,52***
300	7,037	193,905	40,93***	303,64***
400	-3,175	35,306	-18,47***	103,17***
500	-3,021	30,228	-17,57***	88,33***
600	-13,956	193,208	-81,18***	564,60***

O teste **t** aplicado tanto para γ_1 como para γ_2 deu sempre resultados altamente significativos tanto para $\hat{x}_1^*$ como para $\hat{x}_2^*$. Fazem exceção apenas os valores de $t(\hat{\gamma}_1)$ e $t(\hat{\gamma}_2)$ para $\hat{x}_1^*$ para $\sigma = 10$, que não foram significativos.

A **TABELA VI** mostra os resultados obtidos para as duas séries geradas de 50 ensaios, em conjunto. Apresentamos o número total de pontos de máximo encontrados para cada σ, indicamos o número desses pontos de máximo para os quais o teste Z_1, com relação ao parâmetro a_{11} foi significativo, e o número de casos em que Z_2, com relação a a_{22}, foi significativo, e, finalmente, o número total de casos em que Z_1 e Z_2 foram significativos. Observamos que até $\sigma = 50$ Z_1 foi significativo em todos os casos de pontos de máximo. Daí para a frente, já é menor o número desses testes significativos, com relação ao número de pontos de máximo obtidos. Já Z_2 só é significativo em todos os casos de máximo para

$\sigma = 10$. Para $\sigma = 50$, com 91 casos de máximo, há apenas 17 casos de Z_2 significativo. Notamos com σ maiores que o parâmetro a_{11} acusa diferença significativa sempre com maior freqüência que a_{22}. O mais importante segundo ZAGATTO & PIMENTEL-GOMES (1967) são os pontos de máximo onde tanto a_{11} como a_{22} são significativos e isto podemos observar pela **TABELA VI** que apenas para $\sigma = 10$ os 100 pontos de máximo tiveram ambos parâmetros com estimativas significativas. Já para $\sigma = 50$ houve apenas 91 casos de máximo, dos quais 17 com ambos significativos. Observamos muito poucos casos para σ = 100, 200 e 300, e para $\sigma \geq 400$, com C.V. = 8,33%, já o número de casos de ambos os parâmetros significativos se torna nulo.

TABELA VI. Avaliação do número de casos onde houve ponto de máximo, número de casos onde Z_1, Z_2 ou ambos foram significativos ao nível de 5% de probabilidade.

σ	Nº de Pontos de máximo	Nº de Valores de Z_1 signif.	Z_2 signif.	Nº de casos com Z_1 e Z_2 significat.
10	100	100	100	100
50	91	91	17	17
100	64	59	6	5
200	48	21	1	1
300	39	6	5	3
400	41	12	2	0
500	28	3	3	0
600	31	7	3	0

CONCLUSÕES

1) A estimativa do parâmetro a_{11} teve pequena dispersão em torno de seu valor real, mas a estimativa do parâme

tro a_{22} subestimou muito o verdadeiro valor para $\sigma = 400$ e $\sigma = 600$. A estimativa do parâmetro a_{12} teve sempre grande dispersão.

2) Os intervalos de confiança calculados para a_{11} e a_{22} contêm o verdadeiro valor dos parâmetros mencionados, em todos os casos.

3) em 75% dos casos a distribuição de $\hat{x}_1^*$ tem assimetria positiva e em 87,5% é leptocúrtica. Na distribuição de $\hat{x}_2^*$ observou-se ligeira tendência para que tenha assimetria negativa. É leptocúrtica em todos os casos.

4) Com $\sigma \geq 50$ é raro que os coeficientes do 2º grau sejam ambos significativamente diferentes de zero.

5) Os testes do parâmetro a_{22} diferem significativamente de zero em uma freqüência muito menor que a_{11}, para $\sigma \geq 50$. Isto se deve ao fato de ser a resposta da lavoura maior ao nitrogênio do que ao fósforo.

6) As distribuições de x_1^* e x_2^* não são normais.

RESUMO

Consideramos neste trabalho um experimento fatorial 3^2 de milho com dois nutrientes, nitrogênio e fósforo, ambos com doses 0, 40, 80 kg/ha, com $q = 40$. Ajustamos uma superfície de segundo grau substituindo-se as variáveis por polinômios ortogonais, para que não haja correlação entre os coeficientes.

Aos dados desse experimento ajustamos a superfície:

$$\hat{Y} = 4975{,}25 + 666{,}50x_1 + 165{,}16x_2 - 228{,}05x_1^2 - 36{,}06x_2^2 + \\ + 3{,}9176x_1x_2,$$

cujos parâmetros serviram de base para a simulação de uma série de 200 experimentos e duas repetições de 50 experimentos com σ: 10, 50, 100, 200, 300, 400, 500 e 600.

Com esses valores:

1) Calculamos estimativas dos parâmetros da equação supracitada para cada σ de geração e intervalos de confiança para os coeficientes a_{11} e a_{22}.

2) Verificamos a natureza dos pontos críticos obtidos em todos os experimentos gerados.

3) Calculamos os coeficientes de assimetria e de curtose para $\hat{x}_1^*$ e $\hat{x}_2^*$, com aplicação de teste **t** para ambos.

4) Calculamos com as duas séries geradas de 50 repetições as estatísticas Z_1 e Z_2 para as estimativas de a_{11} e a_{22}.

Concluímos que:

1) Os intervalos de confiança calculados para a_{11} e a_{22} contêm o verdadeiro valor desses parâmetros.

2) Os pontos de mínimo começam a aparecer para $\sigma \geq 200$.

3) As estimativas do parâmetro a_{22} diferem significativamente de zero com freqüência muito menor que a_{11} para $\sigma \geq 50$.

4) Com $\sigma \geq 50$ é raro que os coeficientes do 2$ grau sejam ambos significativamente diferentes de zero.

5) As distribuições de $\hat{x}_1^*$ e $\hat{x}_2^*$ não são normais.

Palavras-chave: Superfície de Resposta do 2º Grau; Distribuição das Coordenadas do Ponto Crítico; Intervalos de Confiança dos Parâmetros da Superfície de Resposta.

SUMMARY

STUDY OF THE DISTRIBUTION OF THE CRITICAL POINT COORDINATES OF THE QUADRATIC REGRESSION EQUATION WITH TWO INDEPENDENT VARIABLES

In this paper a 3^2 factorial experiment with two nutrients (nitrogen and phosphorus) was used, assuming levels 0, 40, 80 kg/ha. A second degree response surface was fitted, using orthogonal polynomials in order to have estimates independent among one another. The equation obtained was:

$$\hat{Y} = 4975,25 + 666,50x_1 + 165,16x_2 - 228,05x_1^2 - 36,06x_2^2 + 3,9176x_1x_2.$$

the parameters which were used to simulate a series of 200 experiments, and two replications of 50 experiments with σ: 10, 50, 100, 200, 300, 400, 500 and 600.

With these values were calculated:

1) Averages of the estimated coefficients of the response surface above, for each σ of generation, and confidence intervals for a_{11} and a_{22}.

2) The nature of the critical point.

3) The asymetry and curtosis coefficients of $\hat{x}_1^*$ and $\hat{x}_2^*$, both with **t** test application to compare with zero.

4) The Z test for the parameter estimates $\hat{a}_{11}$ and $\hat{a}_{22}$ to compare with zero, for the two sets of 50 experiments.

The conclusions were:

1) The confidence intervals obtained for a_{11} and a_{22} include the real value of these parameters.

2) The minimum points appear for $\sigma \geq 200$.

3) The a_{22} estimates are significantly different from zero with a smaller frequency then a_{11}, for $\sigma \geq 50$.

4) With $\sigma \geq 50$ it is rare that both a_{11} and a_{22} coefficients are significantly different from zero.

5) The $\hat{x}_1^*$ and $\hat{x}_2^*$ distributions are not normal.

Key words: Second degree Response Surface; Coordinates of Critical Points Distribution; Confidence Intervals for Response Surface Parameters.

LITERATURA CITADA

CAMPOS, H., 1967. Aspectos da Aplicação das Superfícies de Resposta a Ensaios Fatoriais 3^3 de Adubação. Piracicaba. 82p. (Livre-Docência - ESALQ/USP).

GOMES, Marli de Bem, 1989. Intervalos de Confiança para Coordenadas do Ponto Crítico da Equação Quadrática de Regressão a Duas Variáveis Independentes. Piracicaba. 70p. (Livre-Docência - ESALQ/USP).

GOMES, Marli de Bem, 1992. Intervalos de Confiança para as Coordenadas do Ponto Crítico da Equação de Regressão a Duas variáveis Independentes. **Revista de Agricultura**, Piracicaba, **67**(3): 205-217.

KENDALL, M.G. & A. STUART, 1963. **The Advanced Theory of Statistics.** Londres, Charles Griffin. Vols. 1 e 2.

MASON, D.D., 1956. Functional models and Experimental Designs for Caracterization Response Curve and Surfaces. **In:** BAUM, E.L.; E.O. HEADY & J. BLACKMORE. **Methodological Procedures in the Economic Analysis of Fertilizer Data.** Ames, Iowa State University Press. p.76-98.

MORAES, R.S., 1969. Superfície Polinomial de Resposta num Ensaio de Adubação com Níveis Não Equidistantes. Piracicaba. 58p. (Doutorado - ESALQ/USP).

PIMENTEL-Gomes, F. & A. CONAGIN, 1987. Experimentos de Adubação: Planejamento e Análise Estatística. Londrina, 2º SEAGRO. 102p.

TEJADA, H., 1966. Evaluación de Alguns Aspectos de la Metodologia para Determinar Funciones de Resposta a la Fertilización y su Utilización Económica. **In:** SEMINÁRIO INTERNACIONAL SOBRE INVESTIGACIÓN ECONÓMICA Y EXPERIMENTACIÓN AGRICOLA. Santiago, Chile.

ZAGATTO, A.G. & F. PIMENTEL-GOMES, 1967. Aspectos Econômicos da Adubação. **In:** MALAVOLTA, E. **Manual de Química Agrícola - Adubos e Adubação.** 2.ed. São Paulo, Ceres. p. 560-586.