

## REGIÕES DE CONFIANÇA PARA LOCALIZAÇÃO DO PONTO ESTACIONÁRIO EM SUPERFÍCIES DE RESPOSTA USANDO OS MÉTODOS “BOOTSTRAP” CLÁSSICO E BAYESIANO

David José MIQUELLUTI<sup>1</sup>  
Silvio Sandoval ZOCCHI<sup>2</sup>  
Clarice Garcia Borges DEMÉTRIO<sup>2</sup>

- RESUMO: Abordaram-se procedimentos para construção de regiões de confiança das coordenadas do ponto estacionário em diferentes situações considerando-se a forma das superfícies analisadas e a distribuição da variância dos erros do modelo. Foram utilizadas a metodologia de Box e Hunter (1954) (BH), “bootstrap” clássico e o “bootstrap” Bayesiano aliados ao cálculo da distância de Mahalanobis entre as coordenadas do ponto estacionário da amostra observada e aquelas obtidas por meio das estimativas “bootstrap” (BCM e BBM), e métodos “bootstrap” clássico e “bootstrap” Bayesiano aliados a métodos não paramétricos de estimação de densidades de probabilidade (BCNP e BBNP). Avaliaram-se as metodologias por meio de simulação e foi analisado um conjunto de dados de produção de amendoim. A metodologia BH apresentou bom desempenho havendo concordância entre as regiões de confiança nominais e reais. O mesmo comportamento foi observado para os métodos BCM e BBM. No entanto, os métodos BCNP e BBNP não tiveram desempenho satisfatório, resultando em nível de significância real menor do que o nominal para os menores autovalores. No caso dos autovalores maiores observou-se situação inversa. Para o conjunto de dados de produção de amendoim, com os métodos BH, BCM e BCNP obtiveram-se regiões de confiança mais amplas comparativamente aos métodos BBM e BBNP. No entanto, os valores das estimativas “bootstrap” Bayesiano foram mais próximas das estimativas de mínimos quadrados e apresentaram menor dispersão.
- PALAVRAS-CHAVE: Métodos computacionalmente intensivos; densidade não paramétrica; regressão linear múltipla.

<sup>1</sup>Departamento de Solos e Recursos Naturais, Centro de Ciências Agroveterinárias, Universidade do Estado de Santa Catarina – UDESC, CEP: 88520-000, Lages, SC, Brasil. E-mail: [dmiquell@cav.udesc.br](mailto:dmiquell@cav.udesc.br)

<sup>2</sup>Departamento de Ciências Exatas, Escola Superior de Agricultura “Luiz de Queiroz” – ESALQ, Universidade de São Paulo – USP, CEP: 13418-900, Piracicaba, São Paulo, Brasil. E-mail: [sszocchi@esalq.usp.br](mailto:sszocchi@esalq.usp.br) / [clarice@esalq.usp.br](mailto:clarice@esalq.usp.br)

## 1 Introdução

Experimentos nos quais uma ou mais variáveis respostas são influenciadas por diversos fatores quantitativos são bastante comuns nas áreas agrícola, química, biológica, dentre outras. Nesse caso, o problema de pesquisa consiste em se estudar essa relação, sendo de grande utilidade o uso da metodologia de superfícies de resposta (MSR).

A MSR compreende um conjunto de técnicas estatísticas que incluem desde o planejamento do experimento até a análise dos dados, com o objetivo de relacionar a(s) resposta(s) com os níveis de fatores quantitativos que a(s) afeta(m). Por exemplo, em agricultura, a resposta pode ser a produtividade de uma cultura e os fatores, as quantidades de diferentes tipos de nutrientes aplicados na lavoura.

O grande impulso na análise de dados experimentais por meio de superfícies de resposta, foi dado por Box e Wilson (1951) que desenvolveram métodos de determinação de condições ótimas em pesquisas industriais. Consideraram, basicamente, que a resposta pode ser aproximada, em uma região de interesse, por um polinômio nos níveis de vários fatores ou variáveis envolvidos. Apresentaram, então, planejamentos para a determinação do ponto de resposta ótima, utilizando o menor número possível de pontos, cujos desempenhos foram comparados com os de outros planejamentos experimentais utilizados na época. Foi preconizada, também, a adoção do método de máxima inclinação ascendente na procura de uma região estacionária em torno do ótimo e o estudo da natureza dessa região reduzindo a equação ajustada à sua forma canônica. Eles estavam interessados, principalmente, em problemas de estimação de condições ótimas em trabalhos na área de química, mas relataram que os métodos descritos poderiam ser de grande valia em outros campos em que a experimentação é seqüencial e o erro experimental, razoavelmente pequeno. Esse e vários outros estudos conduzidos por Box e associados, posteriormente, constituíram a base para os trabalhos de pesquisa em MSR (Box e Draper, 1959, 1963, 1987; Andrade, 1977). A literatura existente específica sobre a MSR, por sua vez, é bastante extensa, apresentando diversas aplicações em diferentes campos de pesquisa sendo que revisões sobre o tema são apresentadas por Mead e Pike (1975), Myers, Khuri e Carter Jr. (1989), Myers (1999) e Myers et al. (2004).

Conforme Khuri e Cornell (1996) e Myers e Montgomery (2002), de uma forma geral, mencionar que uma resposta depende dos níveis de fatores quantitativos é o mesmo que supor a existência de uma função de resposta que represente a verdadeira relação entre estas quantidades. Na prática, não se observa o verdadeiro valor da resposta, mas sim o valor de uma variável aleatória, denominada resposta observada.

Apesar de, em certas ocasiões haver o interesse no modelo global, geralmente deseja-se determinar os valores das variáveis preditoras cuja resposta é um máximo ou mínimo, ou seja, procura-se obter o ponto ótimo de operação (Campos, 1967; Malheiros e Perecin, 1983; Sanches, 1986; Pimentel Gomes e Conagin, 1987; Gomes, 1989; Pimentel Gomes, 1990).

A determinação dos níveis dos fatores que otimizam a resposta consiste, inicialmente, na obtenção das coordenadas do ponto estacionário de um modelo ajustado, geralmente um modelo de segunda ordem completo. No entanto, devido à variação amostral inerente ao processo experimental, é interessante obter uma região de confiança das coordenadas verdadeiras de localização do ponto estacionário para complementar a análise da superfície na qual foi ajustado o polinômio de segunda ordem (Box e Hunter, 1954,1957; Stablein, Carter Jr. e Wampler, 1983; Khuri e Cornell, 1996; Myers e Montgomery, 2002).

O conhecimento da região de confiança da localização do ponto estacionário é importante pois mostra a qualidade da estimativa pontual por meio da proporção da região experimental abrangida pela região de confiança, fornece informação sobre a ação sinérgica dos fatores e também é útil em situações que envolvem experimentos com misturas e respostas múltiplas (Castillo e Cahya, 2001; Peterson, Cahya e Castillo, 2002; Cahya, Castillo e Peterson, 2004).

Várias alternativas têm sido sugeridas para se determinar a região de confiança sendo a primeira delas, a metodologia apresentada por Box e Hunter (1954) que é baseada na pressuposição de normalidade dos erros aleatórios do modelo. Recentemente, Freitas (1993), Zocchi e Silva (2003) e Silva e Zocchi (2004) apresentaram formas diferentes de se construir as regiões de confiança baseadas na metodologia “bootstrap” não paramétrica.

“Bootstrap” é uma técnica computacionalmente intensiva utilizada para se estimar a distribuição amostral de uma estatística (Efron, 1979, 1982; Efron e Tibshirani, 1993; Davison e Hinkley, 1997; Manly, 1997). Dada uma amostra aleatória de tamanho  $n$ , essa técnica consiste na reamostragem, com reposição, a partir de uma função de distribuição não paramétrica que atribui probabilidades iguais a  $1/n$  a todas as  $n$  observações.

Como alternativa, Rubin (1981), baseado em inferência Bayesiana via processo Dirichlet (Ferguson, 1973), apresentou o método “bootstrap” Bayesiano, semelhante ao “bootstrap” clássico. Nesse método, é atribuído, a cada observação, um peso proporcional às probabilidades geradas a partir de uma distribuição a posteriori, considerando-se uma priori imprópria. Essa distribuição a posteriori é centrada em  $\frac{1}{n}$ , mas varia de uma repetição “bootstrap” para outra. Nesse caso, valores da variável resposta que não são observados têm probabilidade igual a zero, da mesma forma que ocorre sob a função de distribuição amostral. A distribuição “bootstrap” Bayesiana do parâmetro é semelhante à distribuição “bootstrap” do seu estimador diferindo, no entanto, na forma com que as probabilidades associadas às observações são geradas e na interpretação das distribuições resultantes. Enquanto o “bootstrap” Bayesiano simula uma distribuição a posteriori para o parâmetro, o “bootstrap” clássico simula a distribuição amostral de uma estatística. Segundo o autor, esse método seria o análogo Bayesiano do “bootstrap” clássico. Dadas as suas características, acredita-se que ele propicie uma maior cobertura do espaço paramétrico e com isso uma determinação mais precisa das regiões de confiança do parâmetro.

Neste trabalho, foram comparadas diferentes metodologias para a determinação de regiões de confiança para as coordenadas de localização do ponto estacionário de uma superfície de resposta representada por um modelo polinomial de segundo grau. O comportamento dos diferentes métodos foram avaliados por meio de estudos de simulação. Foi efetuada uma aplicação a um conjunto de dados de produção de amendoim. Os procedimentos computacionais desenvolvidos foram implementados no programa R (R development core team, 2006).

## 2 Metodologia

### 2.1 Descrição dos métodos para determinação das regiões de confiança

#### 2.1.1 Método de Box e Hunter

Seja  $Y$  uma variável resposta e  $\mathbf{x} = [x_1, \dots, x_k]^T$  o vetor de  $k$  variáveis explicativas. Considere o modelo de superfície de resposta de segundo grau completo dado por:

$$Y = \beta_0 + \mathbf{x}^T \mathbf{b} + \mathbf{x}^T \mathbf{B} \mathbf{x} + \epsilon \quad (1)$$

em que,  $\beta_0$  é um parâmetro;  $\mathbf{b} = [\beta_1, \dots, \beta_k]^T$  é o vetor de parâmetros de primeira ordem;  $\mathbf{B}$  é a matriz de parâmetros de segunda ordem, de dimensões  $k \times k$ , dada por:

$$\mathbf{B} = \begin{bmatrix} \beta_{11} & \frac{\beta_{12}}{2} & \dots & \frac{\beta_{1k}}{2} \\ \frac{\beta_{12}}{2} & \beta_{22} & \dots & \frac{\beta_{2k}}{2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\beta_{1k}}{2} & \frac{\beta_{2k}}{2} & \dots & \beta_{kk} \end{bmatrix}$$

e  $\epsilon$  é o erro aleatório.

Dados  $\mathbf{b}$  e  $\mathbf{B}$ , a localização  $\boldsymbol{\theta}$  do ponto estacionário é o vetor  $\mathbf{x}$  solução do sistema de equações,

$$\frac{\partial Y}{\partial \mathbf{x}} = \mathbf{b} + 2\mathbf{B}\mathbf{x} = \mathbf{0} \quad (2)$$

em que  $\mathbf{0}$  é um vetor de zeros de dimensões  $k \times 1$ , ou seja,  $\boldsymbol{\theta} = -\frac{1}{2}\mathbf{B}^{-1}\mathbf{b}$ , desde que  $\mathbf{B}$  seja não singular.

Seguindo Box e Hunter (1954), considere o conjunto de  $n$  vetores de variáveis explicativas  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$  e respectivas observações da variável resposta  $Y$ ,  $\{y_1, \dots, y_n\}$ . Sejam,  $\hat{\mathbf{B}}$  e  $\hat{\mathbf{b}}$  a matriz e o vetor de estimativas de mínimos quadrados dos parâmetros de  $\mathbf{B}$  e  $\mathbf{b}$ .

Seja, ainda,  $\boldsymbol{\delta} = \hat{\mathbf{b}} + 2\hat{\mathbf{B}}\boldsymbol{\theta}$  e  $\hat{\boldsymbol{\theta}} = -\frac{1}{2}\hat{\mathbf{B}}^{-1}\hat{\mathbf{b}}$ , o estimador de mínimos quadrados de  $\boldsymbol{\theta}$ .

Supondo-se que  $\epsilon_i \sim N(0, \sigma^2)$ ,  $i = 1, \dots, n$ , e que os erros são independentes, tem-se que  $Var(\boldsymbol{\delta}) = \mathbf{V}\sigma^2$ , em que  $\mathbf{V}$  é a matriz de variâncias e covariâncias de  $\boldsymbol{\delta}$ . Conseqüentemente, tem-se que  $\boldsymbol{\delta} \sim N(\mathbf{0}, \mathbf{V}\sigma^2)$ .

Nessas condições, a região de  $100(1-\alpha)\%$  de confiança para  $\boldsymbol{\theta}$  pode ser definida como o conjunto de valores de  $\boldsymbol{\theta}$  que satisfazem a desigualdade:

$$\frac{\boldsymbol{\delta}^T \mathbf{V}^{-1} \boldsymbol{\delta}}{2\hat{\sigma}^2} \leq F_{\alpha, k, \nu}, \quad (3)$$

em que  $\hat{\sigma}^2$  é o quadrado médio residual da análise da variância, com  $\nu$  graus de liberdade e  $F_{\alpha, k, \nu}$  é o percentil de ordem  $100(1-\alpha)\%$  da distribuição  $F$  com  $k$  graus de liberdade dos parâmetros e  $\nu$  graus de liberdade do resíduo.

### 2.1.2 Métodos baseados na técnica “bootstrap” clássica

Zocchi e Silva (2003) apresentaram uma metodologia para determinação das regiões de confiança baseada na técnica “bootstrap”, cujo procedimento, com pequenas modificações, será apresentado a seguir.

Considere o conjunto de observações  $\{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_n, y_n)\}$ , as estimativas  $\hat{\beta}_0$ ,  $\hat{\mathbf{b}}$  e  $\hat{\mathbf{B}}$ , e o conjunto de resíduos padronizados  $\{\zeta_1, \dots, \zeta_n\}$ , em que

$$\zeta_i = \left[ y_i - (\hat{\beta}_0 + \mathbf{x}_i^T \hat{\mathbf{b}} + \mathbf{x}_i^T \hat{\mathbf{B}} \mathbf{x}_i) \right] \left( \frac{n}{n-p} \right)^{\frac{1}{2}}, \quad (i = 1, \dots, n)$$

em que  $p$  é o número total de parâmetros do modelo, dado por  $p = (k+1)(k+2)/2$ .

Em seguida, retire  $M$  (geralmente,  $M > 1000$ ) amostras de tamanho  $n$ , com reposição, do conjunto de resíduos  $\{\zeta_1, \dots, \zeta_n\}$ , representadas por  $\{\zeta_1^{*(m)}, \dots, \zeta_n^{*(m)}\}$ , ( $m = 1, \dots, M$ ). Para cada uma dessas amostras, denominadas amostras “bootstrap”, construa um novo conjunto de dados:

$$\{(\mathbf{x}_1, \hat{\beta}_0 + \mathbf{x}_1^T \hat{\mathbf{b}} + \mathbf{x}_1^T \hat{\mathbf{B}} \mathbf{x}_1 + \zeta_1^{*(m)}), \dots, (\mathbf{x}_n, \hat{\beta}_0 + \mathbf{x}_n^T \hat{\mathbf{b}} + \mathbf{x}_n^T \hat{\mathbf{B}} \mathbf{x}_n + \zeta_n^{*(m)})\}, \quad (m = 1, \dots, M).$$

Para cada um desses  $M$  conjuntos estime  $\beta_0$ ,  $\mathbf{b}$ ,  $\mathbf{B}$  e  $\boldsymbol{\theta}$ , obtendo dessa forma,  $\{\hat{\boldsymbol{\theta}}^{*(1)}, \dots, \hat{\boldsymbol{\theta}}^{*(M)}\}$ , em que  $\hat{\boldsymbol{\theta}}^{*(m)}$  é a estimativa de  $\boldsymbol{\theta}$  referente à  $m$ -ésima amostra “bootstrap”.

Baseando-se em  $\{\hat{\boldsymbol{\theta}}^{*(1)}, \dots, \hat{\boldsymbol{\theta}}^{*(M)}\}$ , estime-se a função densidade de probabilidade conjunta de  $\boldsymbol{\theta} = [\theta_1, \dots, \theta_k]^T$ , por meio de

$$f(\boldsymbol{\theta}) = f(\theta_1, \dots, \theta_k) = \frac{1}{M \prod_{j=1}^k h_j} \sum_{m=1}^M \varphi\left(\frac{\theta_1 - \hat{\theta}_1^{*(m)}}{h_1}\right) \cdots \varphi\left(\frac{\theta_k - \hat{\theta}_k^{*(m)}}{h_k}\right) \quad (4)$$

em que,  $h_j$ ,  $j = 1, \dots, k$ , são as constantes de suavização e  $\varphi$  é a função densidade kernel, no caso, normal padrão, dada por  $\varphi(t) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{1}{2}t^2\}$ , para  $t \in \mathfrak{R}$ . Seguindo a sugestão de Bowman e Azzalini (1997), utiliza-se, como constantes de suavização,

$$h_j = \left[ \frac{4}{M(k+2)} \right]^{\frac{1}{(k+4)}} \sqrt{s_{jj}}$$

em que,

$$s_{jj} = \frac{1}{M} \sum_{m=1}^M (\hat{\theta}_j^{*(m)} - \bar{\theta}_j^*)^2$$

e

$$\bar{\theta}_j^* = \frac{1}{M} \sum_{m=1}^M \hat{\theta}_j^{*(m)}.$$

Em seguida, calcule o valor de  $f(\hat{\theta}^{*(m)})$ , para cada  $m$  ( $m = 1, \dots, M$ ). Seja  $P_\alpha$  o percentil de ordem  $\alpha$  do conjunto de valores  $\{f(\hat{\theta}^{*(1)}), \dots, f(\hat{\theta}^{*(M)})\}$ . A região de  $100 \times (1 - \alpha)\%$  de confiança para  $\theta$  será dada por:

$$\theta \in \mathfrak{R}^k : f(\theta) \leq P_\alpha.$$

$$\text{Sejam } \hat{\theta}^* = [\hat{\theta}_1^{*(m)}, \hat{\theta}_2^{*(m)}, \dots, \hat{\theta}_k^{*(m)}]^T, \bar{\theta}^* = [\bar{\theta}_1^*, \bar{\theta}_2^*, \dots, \bar{\theta}_k^*]^T,$$

$$\Sigma_r = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1k} \\ s_{12} & s_{22} & \cdots & s_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ s_{1k} & s_{2k} & \cdots & s_{kk} \end{bmatrix} \text{ e}$$

$$s_{jj'} = \frac{1}{M} \sum_{m=1}^M (\hat{\theta}_j^{*(m)} - \bar{\theta}_j^*)(\hat{\theta}_{j'}^{*(m)} - \bar{\theta}_{j'}^*).$$

Uma forma alternativa ao uso de densidades não paramétricas para a obtenção das regiões de confiança para as coordenadas do ponto estacionário, é defini-las por meio da distância de Mahalanobis (Freitas, 1993). Seja  $DM(\theta)$  a distância de Mahalanobis do vetor  $\theta$  ao vetor de médias “bootstrap”  $\bar{\theta}^*$  dada por,

$$DM(\theta) = (\theta - \bar{\theta}^*)^T \Sigma_r^{-1} (\theta - \bar{\theta}^*)$$

e seja

$$DM(\theta^{*(m)}) = (\theta^{*(m)} - \bar{\theta}^*)^T \Sigma_r^{-1} (\theta^{*(m)} - \bar{\theta}^*)$$

a distância de Mahalanobis de  $\theta^{*(m)}$  a  $\bar{\theta}^*$ , ( $m = 1, \dots, M$ ).

Seja, ainda,  $Q$  o percentil de ordem  $\alpha$  do conjunto de distâncias  $\{DM(\hat{\theta}^{*(1)}), DM(\hat{\theta}^{*(2)}), \dots, DM(\hat{\theta}^{*(M)})\}$ . A região de  $100 \times (1 - \alpha)\%$  de confiança para  $\theta$  será dada por:

$$\theta \in \mathfrak{R}^k : DM(\theta) \leq Q_\alpha.$$

### 2.1.3 Métodos baseados na técnica “bootstrap” Bayesiana

Um método similar ao “bootstrap” clássico é aquele apresentado por Rubin (1981), baseado em inferência Bayesiana via processo Dirichlet (Ferguson, 1973; Wilks, 1962). Segundo esse método, atribui-se a cada observação um peso igual às probabilidades geradas de uma distribuição a posteriori, obtida com uma priori imprópria. Essa distribuição a posteriori é centrada em  $1/n$ , mas varia de uma repetição “bootstrap” para outra. Valores da variável que não são observados têm probabilidade igual a zero, da mesma forma que ocorre sob a função de distribuição amostral.

Seja  $(\mathbf{x}_1, Y_1), \dots, (\mathbf{x}_n, Y_n)$ , uma amostra aleatória simples de uma população com função de distribuição  $F$ , e  $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_n, y_n)$ , os seus respectivos valores. No caso do “bootstrap” clássico, tendo-se uma função de distribuição empírica ou amostral (fda),  $\hat{F}$ , com probabilidade  $\frac{1}{n}$  para cada uma das observações  $(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_n, y_n)$ , pode-se expressar  $F$  como  $\mathbf{f} = (f_1, f_2, \dots, f_L)$ , em que  $f_l = Prob[(\mathbf{x}, Y) = l]$ , e  $\hat{F}$  como  $\hat{\mathbf{f}} = (\hat{f}_1, \hat{f}_2, \dots, \hat{f}_L)$ , em que  $\hat{f}_l = \frac{\#\{(\mathbf{x}_i, y_i) = l\}}{n}$ , com  $\sum_{l=1}^L \hat{f}_l = 1$ . Assim, a distribuição de  $\hat{f}_l$  é proporcional à distribuição binomial,  $Bi(n, f_l)$ . Considerando que cada tentativa é independente e que qualquer um dos  $n$  elementos amostrais tem a mesma probabilidade,  $1/n$ , de ser selecionado, a distribuição de  $\hat{\mathbf{f}} = (\hat{f}_1, \hat{f}_2, \dots, \hat{f}_L)$  é proporcional à distribuição multinomial,  $Mn_L(n, \mathbf{f})$ .

$$(\hat{\mathbf{f}}|n; \mathbf{f}) \propto \frac{1}{n} Mn_L(n, \mathbf{f}). \quad (5)$$

A distribuição “bootstrap” de  $(\mathbf{x}_i, Y_i)^* \stackrel{i.i.d}{\sim} \hat{F}$  ( $i = 1, \dots, n$ ), pode ser descrita em termos de  $\hat{\mathbf{f}}^* = (\hat{f}_1^*, \hat{f}_2^*, \dots, \hat{f}_L^*)$ , em que  $\hat{f}_l^* = \frac{nr \cdot \{(\mathbf{x}_i, y_i)^* \in l\}}{n}$ , com  $\sum_{l=1}^L \hat{f}_l^* = 1$ . Ela é a mesma como em (5), com  $\hat{\mathbf{f}}$  no lugar de  $\mathbf{f}$ ,

$$(\hat{\mathbf{f}}^*|n; \hat{\mathbf{f}}) \propto \frac{1}{n} Mn_L(n, \hat{\mathbf{f}}). \quad (6)$$

Portanto, o conceito chave do método é que cada repetição “bootstrap” cria uma coleção de pesos, proporcional a  $Mn_L(n, \hat{\mathbf{f}})$ , que representa a proporção de vezes que cada par  $(\mathbf{x}_i, y_i)$  foi amostrado, gerando, portanto, uma função de verossimilhança para os dados.

A idéia de Rubin (1981) foi, a partir dessa função de verossimilhança e de uma distribuição a priori conjugada imprópria, obter uma coleção dos pesos amostrando a partir da posteriori.

Considere como a priori sobre  $\mathbf{f}$  uma distribuição Dirichlet, conforme descrita em Wilks (1962), com parâmetro  $a$ ,  $\mathbf{f} \sim Di_L(a, a, \dots, a)$ , com função densidade dada por:

$$p(\mathbf{f}) = \frac{\Gamma(a + a + \dots + a)}{\prod_{l=1}^L \Gamma(a)} \prod_{l=1}^L f_l^{a-1}. \quad (7)$$

Tendo-se observado  $\hat{\mathbf{f}}$ , a distribuição a posteriori de  $\mathbf{f}$ , é:

$$\mathbf{f}|\hat{\mathbf{f}} \sim Di_L(a + n\hat{f}_1, a + n\hat{f}_2, \dots, a + n\hat{f}_L). \quad (8)$$

com função densidade proporcional a  $\prod_{l=1}^L f_l^{n\hat{f}_l+a-1}$ . Fazendo-se  $a = 0$  para representar a informação a priori, resulta em  $\mathbf{f}|\hat{\mathbf{f}} \sim Di_L(n\hat{f}_1, n\hat{f}_2, \dots, n\hat{f}_L)$ , cuja função densidade é:

$$p(\mathbf{f}|\hat{\mathbf{f}}) = \frac{\Gamma(n\hat{f}_1 + n\hat{f}_2 + \dots + n\hat{f}_L)}{\prod_{l=1}^L \Gamma(n\hat{f}_l)} \prod_{l=1}^L f_l^{n\hat{f}_l-1}. \quad (9)$$

Assim, no “bootstrap” clássico, os pesos atribuídos aos pares  $(\mathbf{x}_i, y_i)$  têm distribuição proporcional a  $Mn_L(n, \hat{\mathbf{f}})$ , e, no “bootstrap” Bayesiano esses pesos apresentam distribuição  $Di_L(1, 1, \dots, 1)$ . Essa duas distribuições são bastante similares, sendo que,

1. Ambas têm suporte sobre aquelas categorias com  $\hat{f}_l > 0$ , isto é, sobre aquelas categorias nas quais os dados foram observados;
2. Ambas têm vetor de esperanças igual a  $\hat{\mathbf{f}}$ ;
3. As matrizes de variâncias e covariâncias são aproximadamente iguais,  $Cov(\mathbf{f}|\hat{\mathbf{f}}) = \frac{\Sigma_{\hat{\mathbf{f}}}}{n+1}$  e  $Cov_*(\hat{\mathbf{f}}^*|\hat{\mathbf{f}}) = \frac{\Sigma_{\hat{\mathbf{f}}}}{n}$ , em que  $\Sigma_{\hat{\mathbf{f}}}$  tem elementos iguais a  $\hat{f}_l(1-\hat{f}_l)$  na diagonal principal e  $-\hat{f}_l\hat{f}_{l'}$  fora da diagonal principal;
4. As matrizes de correlação são iguais com os elementos fora da diagonal principal iguais a  $\frac{-\hat{f}_l\hat{f}_{l'}}{\sqrt{(\hat{f}_l(1-\hat{f}_l))(\hat{f}_{l'}(1-\hat{f}_{l'}))}}$ .

A distribuição Dirichlet, com as características desejadas, pode ser facilmente simulada, considerando-se os seguintes resultados apresentados em Wilks (1962):



1. Seja  $X$  uma variável aleatória com função densidade  $f$  e função de distribuição  $F$ . Seja  $Y = F(X)$ . Então,  $Y$  será uniformemente distribuída no intervalo  $[0, 1]$ .
2. Sejam  $x_{(1)}, \dots, x_{(n)}$  as estatísticas de ordem de uma amostra extraída de uma função de distribuição contínua  $F(x)$ . Então, as diferenças  $g_1 = F(x_1), g_2 = F(x_2) - F(x_1), \dots, g_n = F(x_n) - F(x_{n-1})$ , são variáveis aleatórias com a distribuição Dirichlet, com parâmetros  $Di_{(n-1)}(1, 1, \dots, 1)$ .

Para se implementar o método “bootstrap” Bayesiano, considere o conjunto de observações  $(\mathbf{x}_1, Y_1), \dots, (\mathbf{x}_n, Y_n)$ . Os passos são os seguintes:

1. Gere uma amostra  $(u_1, u_2, \dots, u_{n-1})$  de tamanho  $n - 1$  de uma distribuição uniforme no intervalo  $[0, 1]$ ;
2. Disponha os valores gerados em ordem crescente. Sejam  $u_{(1)}, u_{(2)}, \dots, u_{(n-1)}$  esses valores. Considere  $u_{(0)} = 0$  e  $u_{(n)} = 1$ ;
3. Determine  $g_i = u_{(i)} - u_{(i-1)}$  com  $i = 1, 2, \dots, n$ . Seja  $\mathbf{g} = [g_1, \dots, g_n]$  e  $\mathbf{G} = \text{diag}(\mathbf{g})$ ;
4. Determine o vetor de parâmetros do modelo:

$$\boldsymbol{\beta} = (\mathbf{Z}^T \mathbf{G} \mathbf{Z})^{-1} \mathbf{Z}^T \mathbf{G} \mathbf{Y}$$

em que  $\mathbf{Z}_{[n, (k+1)(k+2)/2]} = [\mathbf{1}, \mathbf{x}_1, \dots, \mathbf{x}_k, \mathbf{x}_1^2, \dots, \mathbf{x}_k^2, \mathbf{x}_1 \mathbf{x}_2, \dots, \mathbf{x}_{(k-1)} \mathbf{x}_k]$  e  $\mathbf{y} = (y_1, \dots, y_n)$ ;

5. Obtenha, a partir do item anterior, os valores para  $\beta_0$ ,  $\mathbf{b}$ ,  $\mathbf{B}$  e  $\boldsymbol{\theta}$ ;
6. Repita  $M$  vezes os passos 1, 2, 3, 4 e 5.

De posse da distribuição “bootstrap” Bayesiana de  $\boldsymbol{\theta}$ , a obtenção das regiões de confiança foi efetuada utilizando-se os métodos da distância de Mahalanobis e método não-paramétrico de estimação de densidade, conforme descrito anteriormente.

## 2.2 Simulação

A fim de avaliar a metodologia proposta neste trabalho, foram construídos, por meio de simulação de Monte Carlo, 1000 conjuntos de dados. Considerou-se um modelo polinomial de segundo grau completo e um delineamento composto central com duas variáveis preditoras, conforme descrito a seguir.

Seja  $k = 2$  e o modelo (1) escrito na sua forma canônica, ou seja,

$$Y = Y_{\boldsymbol{\theta}} + (\mathbf{x} - \boldsymbol{\theta})^T \mathbf{O} \boldsymbol{\Lambda} \mathbf{O}^T (\mathbf{x} - \boldsymbol{\theta}) + \epsilon$$

em que,  $Y_\theta$  é o valor da variável resposta para  $x = \theta$ ,  $O = [o_1 \ o_2]$  a matriz de dimensões  $2 \times 2$  formada pelos autovalores normalizados  $o_1$  e  $o_2$  associados aos autovalores  $\lambda_1$  e  $\lambda_2$  da matriz  $B$  e

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

Estipulou-se, para efeito de simulação o valor de  $Y_\theta = 100$ ,  $\theta = \begin{bmatrix} 0,2 \\ 0,4 \end{bmatrix}$ , de forma que as coordenadas do ponto estacionário pertençam à região experimental, e

$$O = \begin{bmatrix} \cos 30^\circ & -\text{sen } 30^\circ \\ \text{sen } 30^\circ & \cos 30^\circ \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{bmatrix}.$$

Considerou-se ainda  $\epsilon_i$  com distribuições Normal,  $N(0,1)$  e Lognormal  $LN(0,1)$ . Para essa última distribuição os valores da média e do desvio padrão são obtidos a partir dos valores em escala logarítmica. Para cada uma dessas situações foram escolhidos os valores de  $(\lambda_1; \lambda_2)$  dados por:  $(-0,025; -0,017)$ ,  $(-0,25; -0,17)$ ,  $(-0,64; -0,59)$  ou  $(-1,25; -1,17)$ , de forma a se obterem superfícies de resposta mais planas (autovalores com valores absolutos menores), ou com maior grau de curvatura (autovalores com maiores valores absolutos). Em cada uma das situações indicadas foram geradas amostras “bootstrap” de tamanho  $N = 1000$ . A seguir, avaliou-se, a porcentagem de vezes em que a região de  $(1-\alpha)100\%$  de confiança cobriu o ponto estacionário verdadeiro, considerando-se os níveis de significância  $\alpha = 0,05$  e  $\alpha = 0,10$ , pelos métodos de Box e Hunter (1954), distância de Mahalanobis e contornos de densidade não paramétrica bivariada. As porcentagens de cobertura foram analisadas utilizando-se os quantis de uma distribuição binomial de parâmetros 0,95 e 0,90 e um número de repetições igual a 1000. Este procedimento baseia-se no fato de que em cada conjunto de dados simulado, a região de confiança contém ou não o verdadeiro valor do parâmetro, e considerando-se que o número de conjuntos de dados simulados é fixo e que eles são gerados independentemente.

Para ilustrar a metodologia foi analisado um conjunto de dados oriundos de um estudo com fertilizantes descrito em Khuri e Cornell (1996). Este foi conduzido com o objetivo de observar o efeito de dois tipos de fertilizantes, uma formulação NPK padrão (NPK) e um suplemento nutricional (SN), sobre a produtividade de amendoim. Os tratamentos consistiram nas diferentes combinações dos níveis dos dois fertilizantes e foram atribuídos às parcelas (unidades experimentais) de acordo com um delineamento composto central rotacional, com duas repetições.

### 3 Resultados e discussão

#### 3.1 Simulação

Na Tabela 1, são apresentados os valores referentes à utilização da metodologia de Box e Hunter (1954). Se forem considerados os quantis de ordem 2,5% e 97,5% da distribuição binomial com parâmetros 0,90 e 0,95, obtêm-se, respectivamente, os intervalos de confiança (0,881;0,918) e (0,936;0,963). Nota-se que nas situações estudadas os intervalos contêm as porcentagens de cobertura reais. Essa situação não era esperada no caso dos conjuntos de dados gerados com erros com distribuição lognormal, e demonstra que a metodologia clássica de Box e Hunter, baseada na pressuposição de que os erros têm distribuição normal é bastante robusta, mesmo em se tratando de situações em que a distribuição dos erros é bastante assimétrica.

Tabela 1 - Porcentagens de vezes em que as regiões de 90% e 95% cobriram o ponto estacionário verdadeiro, de acordo com os autovalores de  $\mathbf{B}$  e com a distribuição dos erros utilizando a metodologia de Box e Hunter (1954)

$\lambda_1$	$\lambda_2$	Coberturas nominais	
		90%	95%
distribuição lognormal(0;1)			
-0,025	-0,017	89,3	93,7
-0,250	-0,170	90,0	95,1
-0,640	-0,590	90,3	95,0
-0,125	-0,117	89,1	94,1
distribuição normal(0;1)			
-0,025	-0,017	89,9	95,7
-0,250	-0,170	90,3	95,8
-0,640	-0,590	91,5	95,7
-0,125	-0,117	89,4	95,8

Os resultados referentes à utilização dos métodos “bootstrap” aliada ao cálculo da distância de Mahalanobis entre as coordenadas do ponto estacionário da amostra observada e aquelas obtidas através das estimativas “bootstrap” podem ser visualizados na Tabela 2.

Similarmente ao comportamento observado no método anterior, em todas as situações os intervalos de confiança contêm as porcentagens reais de cobertura. Embora fosse esperado um melhor comportamento do método “bootstrap” Bayesiano em relação ao “bootstrap” clássico por, teoricamente, este último propiciar uma melhor cobertura no espaço paramétrico, isto não ocorreu. Os resultados alcançados neste trabalho discordam parcialmente daqueles obtidos por Freitas (1993), que em um estudo similar com o método “bootstrap” clássico verificou que a concordância entre os níveis nominais e reais ocorria apenas quando a superfície era bem definida (valores absolutos dos autovalores maiores do que o

Tabela 2 - Porcentagens de vezes em que as regiões de 90% e 95% cobriram o ponto estacionário verdadeiro, de acordo com os autovalores de  $\mathbf{B}$  e com a distribuição dos erros utilizando a distância de Mahalanobis

$\lambda_1$	$\lambda_2$	Coberturas nominais			
		90%		95%	
		$BB^{(1)}$	$BC^{(2)}$	$BB^{(1)}$	$BC^{(2)}$
distribuição lognormal(0;1)					
-0,025	-0,017	88,3	89,9	94,6	95,3
-0,250	-0,170	90,5	88,9	94,1	93,8
-0,640	-0,590	89,2	90,4	94,2	95,0
-0,125	-0,117	91,3	88,7	96,0	93,6
distribuição normal(0;1)					
-0,025	-0,017	89,4	89,2	94,4	95,1
-0,250	-0,170	87,8	88,9	93,0	94,5
-0,640	-0,590	89,4	89,4	94,0	94,2
-0,125	-0,117	91,5	90,3	95,4	95,1

<sup>(1)</sup>“Bootstrap” Bayesiano      e <sup>(2)</sup>“Bootstrap” clássico

desvio padrão dos erros do modelo). Numa situação de uma superfície plana (valores absolutos dos autovalores menores do que o desvio padrão dos erros do modelo) o nível de significância real tendia a ser menor do que o nominal, em torno de 5%, o que gera regiões de confiança mais amplas, com níveis de confiança reais superiores aos nominais. No entanto, Freitas (1993) salientou que o procedimento adotado gera regiões de confiança para a localização exata do ponto estacionário com razoável precisão.

Os resultados referentes à utilização dos métodos “bootstrap” aliada a métodos não paramétricos de estimação de funções densidade de probabilidade podem ser visualizados na Tabela 3. Observa-se que no caso do “bootstrap” clássico sob distribuição lognormal (0, 1) não houve alteração das porcentagens de cobertura para as quatro situações (autovalores) estudadas. No caso do “bootstrap” Bayesiano sob distribuição lognormal e dos dois métodos para a distribuição normal  $N(0, 1)$  verifica-se que aos autovalores com menor valor absoluto corresponde um nível de significância real menor do que o nominal. Isto origina regiões de confiança reais mais amplas do que as nominais. Por sua vez, à medida que se aumenta o valor absoluto dos autovalores ocorre um cenário diferente, passando-se a ter regiões de confiança menores do que a nominal, principalmente no caso do “bootstrap” Bayesiano. Apenas em poucas situações houve concordância entre os níveis de significância reais e nominais. Provavelmente, isto deve-se ao modo de se selecionar a constante de suavização, a qual, conforme Silverman (1986) e Bowman e Azalini (1997) afeta a contorno da densidade não paramétrica. Resultados similares foram observados por Silva e Zocchi (2004) para o “bootstrap” clássico, usando um procedimento diferente para a seleção da constante de suavização.

Tabela 3 - Porcentagens de vezes em que as regiões de 90% e 95% cobriram o ponto estacionário verdadeiro, de acordo com os autovalores de  $\mathbf{B}$  e com a distribuição dos erros utilizando contorno de densidade não paramétrica bivariada

$\lambda_1$	$\lambda_2$	Coberturas nominais			
		90%		95%	
		$BB^{(1)}$	$BC^{(2)}$	$BB^{(1)}$	$BC^{(2)}$
distribuição lognormal(0;1)					
-0,025	-0,017	89,1	98,6	94,6	99,2
-0,250	-0,170	90,6	98,8	94,2	99,5
-0,640	-0,590	86,6	98,9	91,9	99,6
-0,125	-0,117	82,7	98,1	87,9	99,0
distribuição normal(0;1)					
-0,025	-0,017	94,7	99,4	97,6	99,6
-0,250	-0,170	93,6	98,6	96,2	99,3
-0,640	-0,590	90,6	97,7	94,4	98,2
-0,125	-0,117	72,0	91,3	80,2	95,2

(1) “Bootstrap” Bayesiano      e (2) “Bootstrap” clássico

### 3.1.1 Análise dos dados de produção de amendoim

Para o ajuste do modelo de segundo grau os valores de  $\mathbf{x}_1$  e  $\mathbf{x}_2$  foram codificados por:  $x_{1i} = (a_{1i} - 85)/35$  e  $x_{2i} = (a_{2i} - 20)/5$  ( $i = 1, \dots, 18$ ), produzindo as seguintes estimativas pontuais pelo método de mínimos quadrados:  $\hat{\beta}_0 = 16,365$ ,  $\hat{\mathbf{b}} = [1,6751 \ 2,7650]^T$ ,  $\hat{\mathbf{B}} = \begin{bmatrix} -2,4634 & -0,1669 \\ -0,1669 & -1,9309 \end{bmatrix}$ ,  $\hat{\sigma}^2 = 0,8688$  e  $\hat{\boldsymbol{\theta}} = [0,2932 \ 0,6906]^T$ , o que corresponde às quantidades de fertilizantes  $a_1 = 95,26$  e  $a_2 = 23,45$  libras por parcela. Os autovalores de  $\hat{\mathbf{B}}$ ,  $\lambda_1 = -1,8830$  e  $\lambda_2 = -2,5114$  são menores do que zero e, portanto,  $\hat{\boldsymbol{\theta}}$  é um ponto de máximo.

As estimativas de  $\boldsymbol{\theta}$  e os respectivos erros padrões obtidos por meio dos métodos “bootstrap” e “bootstrap” Bayesiano (Tabela 4) sugerem que os resultados dos dois métodos baseados nas técnicas “bootstrap” são similares, o que é esperado, uma vez que se utilizou uma distribuição a priori não informativa, no caso do “bootstrap” Bayesiano. Pode-se verificar que as regiões de confiança determinadas pelos três métodos analisados estão localizadas dentro da região experimental (Figura 1).

Observa-se que os valores das estimativas do “bootstrap” Bayesiano estão um pouco mais próximas das estimativas pontuais de mínimos quadrados, além de apresentarem menor dispersão; no caso de  $\hat{\theta}_2$  o erro padrão obtido com o método “bootstrap” foi bem superior àquele obtido com o método “bootstrap” Bayesiano; esses fatos propiciaram a maior concentração de pontos em torno da estimativa de  $\theta_2$  apresentada pelo segundo método, diminuindo a área da

Tabela 4 - Estimativas de  $\theta$  e erros padrões obtidos através dos métodos de mínimos quadrados, “bootstrap” clássico e “bootstrap” Bayesiano

Método	$\hat{\theta}_1$	$\hat{\theta}_2$
Mínimos quadrados	0,2932 (0,0855)	0,6902 (0,1824)
“bootstrap” clássico	0,2944 (0,0982)	0,7464 (0,2332)
“bootstrap” Bayesiano	0,2940 (0,0628)	0,7060 (0,0979)

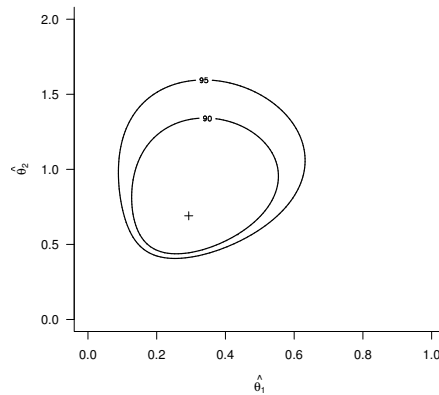


Figura 1 - Regiões de 90% e de 95% de confiança para a localização do ponto estacionário utilizando o método de Box e Hunter (1954).

região de confiança, conforme se constata na análise dos gráficos apresentados nas Figuras 2 e 3. Finalmente, é importante observar que, embora, os métodos “bootstrap” Bayesiano e “bootstrap” clássico dêem respostas similares, a interpretação destas respostas é bastante diferente. Enquanto o “bootstrap” Bayesiano mostra a incerteza a posteriori sobre os parâmetros de acordo com um modelo específico (multinomial/Dirichlet), o “bootstrap” clássico estima a incerteza das estimativas de mínimos quadrados dos parâmetros (Rubin,1981; Lancaster, 2007).

A metodologia proposta baseada no método “bootstrap” Bayesiano, descrito por Rubin (1981), prescinde de qualquer conhecimento a priori. Considerando esse fato, é interessante que se estudem formas de incorporar alguma informação a priori na análise, o que possibilitaria o seu uso em inferências mais precisas. Também, é importante que haja mais pesquisas sobre a forma de seleção da constante de suavização no cálculo de densidades não paramétricas multivariadas.

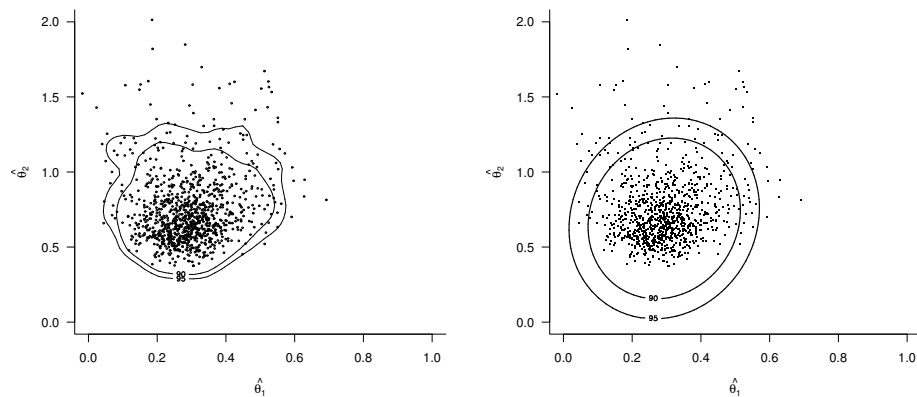


Figura 2 - Regiões de 90% e de 95% de confiança para a localização do ponto estacionário utilizando o método “bootstrap” com contornos da densidade não paramétrica bivariada -  $\mathbf{h} = [0,0311 \ 0,0750]^T$  (esquerda) e da distância de Mahalanobis (direita) considerando-se  $M = 1000$  amostras.

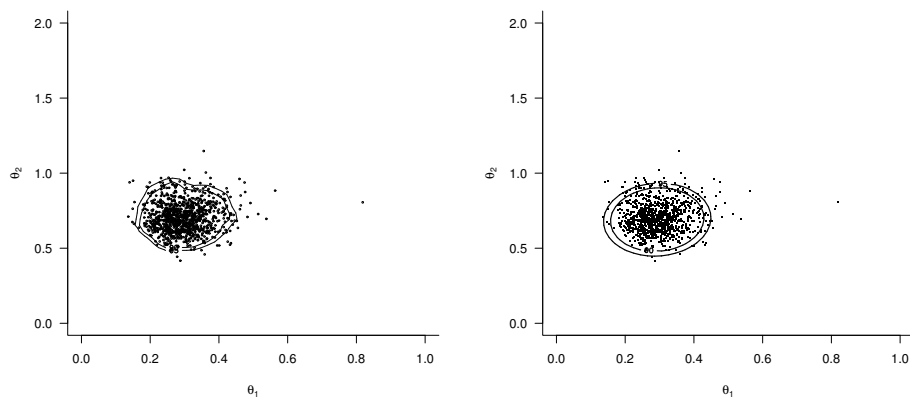


Figura 3 - Regiões de 90% e de 95% de confiança para a localização do ponto estacionário utilizando o método “bootstrap” Bayesiano com contornos da densidade não paramétrica bivariada -  $\mathbf{h} = [0,0199 \ 0,0310]^T$  (esquerda) e da distância de Mahalanobis (direita) considerando-se  $M = 1000$  amostras.

## Conclusões

No estudo de simulação a metodologia de Box e Hunter, baseada na distribuição normal dos erros, apresentou um bom desempenho em todas as situações analisadas, havendo concordância entre as regiões de confiança nominais e reais, mesmo naquelas em que essa distribuição é bastante assimétrica. Este mesmo comportamento foi observado para os métodos BCM e BBM. No entanto, os métodos BCNP e BBNP não apresentaram um desempenho satisfatório, resultando em um nível de significância real menor que o nominal para os autovalores com menor valor absoluto, gerando regiões de confiança maiores. No caso de autovalores com menor valor absoluto observou-se situação inversa. Com base nestes resultados, tem-se que tanto o método BH quanto os métodos BCM e BBM são indicados para a construção das regiões de confiança para as coordenadas do ponto estacionário.

No caso da análise do conjunto de dados de amendoim os métodos BH, BM e BCNP apresentaram regiões de confiança mais amplas comparativamente aos métodos BBM e BBNP. No entanto, os valores das estimativas do “bootstrap” Bayesiano são mais próximas das estimativas de mínimos quadrados e apresentam menor dispersão o que explica a menor área da região de confiança.

## Agradecimentos

Este trabalho é parte da tese de doutorado do primeiro autor no Departamento de Ciências Exatas, Escola Superior de Agricultura “Luiz de Queiroz”, Universidade de São Paulo, Piracicaba, e foi realizado com o apoio da Universidade do Estado de Santa Catarina (UDESC).

MIQUELLUTI, D. J.; ZOCCHI, S. S.; DEMÉTRIO, C. G. B. Confidence region on the location of the stationary point in response surfaces, a Bayesian bootstrap approach. *Rev. Bras. Biom.*, São Paulo, v.26, n.2, p.129-147, 2008.

- **ABSTRACT:** *The procedures for the construction of confidence regions for the coordinates of the stationary point were studied in different situations, considering the shape of the surfaces and the distribution of the variance errors. The methodology of Box and Hunter (1954) (BH), bootstrap and Bayesian bootstrap with Mahalanobis distance among the coordinates of the stationary point of the observed sample and those obtained using bootstrap estimates (BCM and BBM) and bootstrap and Bayesian bootstrap with non-parametric methods for density estimation (BCNP and BBNP) were compared. The methodology evaluation was realized by means of simulation and applied to a peanuts data set. The BH methodology presented a good performance in the analyzed situations, having concordance among the nominal and real confidence regions. This behavior was also observed for the BCM and BBM methods. The BCNP and BBNP methods did not presented a satisfactory performance, resulting in a real significance level lower than the nominal for the lower eigenvalue. The inverse was observed using higher eigenvalue.*



*In the analysis of the peanuts data set the BH, BCM and BCNP methods presented confidence regions larger than the BBM and BBNP methods. The Bayesian bootstrap estimate values were closer of the minimum square estimates and presented less dispersion.*

- **KEYWORDS:** *Respose surface; stationary point; confidence regions; Bayesian bootstrap; non parametric density.*

## Referências

- ANDRADE, D. F. *Alguns aspectos da metodologia de superfície de resposta*. 1977. 87f. Dissertação (Mestrado em Estatística) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 1977.
- BOWMAN, A. W.; AZZALINI, A. *Applied smoothing techniques for data analysis: the Kernel approach with S-Plus illustration*. New York: Oxford Universidade Press, 1997. 210p.
- BOX, G. E. P; DRAPER, N. R. A basis for the selection of response surface design. *Journal of the American Statistical Association*, Washington, v.54, p.622-654, 1959.
- BOX, G. E. P; DRAPER, N. R. The choice of a second order rotatable design. *Biometrika*, London, v.50, p.335-352, 1963.
- BOX, G. E. P; DRAPER, N. R. *Empirical model building and response surfaces*. New York: John Wiley, 1987. 669p.
- BOX, G. E. P; HUNTER, J. S. A confidence region for the solution of a set de simultaneous equations with an application to experimental design. *Biometrika*, Oxford, v.40, p.190-199, 1954.
- BOX, G. E. P; HUNTER, J. S. Multifactor experimental designs for exploring response surfaces. *The Annals of Mathematical Statistics*, Stanford, v.28, p.195-241, 1957.
- BOX, G. E. P; WILSON, K. B. 1951 On the experimental attainment of optimum conditions (with discussion). *Journal of the Royal Statistical Society. Series B*, Oxford, v.13, p.1-45, 1951.
- CAHYA, S.; CASTILLO, E. del; PETERSON, J. J. Computation of confidence regions for optimal factor levels in constrained response surface problems. *Journal of computational and graphical statistics*, New York, v.13, p.499-518, 2004.
- CAMPOS, H. de *Aspectos da aplicação das superfícies de resposta a ensaios fatoriais 3<sup>3</sup> de adubação*. 1967. 82f. Tese (Livre-Docência) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 1967.
- CASTILLO, E. del; CAHYA, S. A tool for computing confidence regions on the stationary point of a response surface. *The American Statistician*, Washington, v.55, p.358-365, 2001.

- DAVISON, A. C.; HINKLEY, D. V. *Bootstrap methods and their application*. Cambridge: Cambridge University Press, 1997. 582p.
- EFRON, B. Bootstrap methods: another look at the jackknife. *Annals of Statistics*, Washington, v.7, p.1-26, 1979.
- EFRON, B. *The Jackknife, the bootstrap and other resampling plans*. Philadelphia: SIAM, 1982. v.38, 93p.
- EFRON, B.; TIBSHIRANI, R. J. *An introduction to the bootstrap*. London: Chapman & Hall, 1993. 436p.
- FERGUSON, T. S. A bayesian analysis of some nonparametric problems. *Annals of Statistics*, Hayjward, v.1, p.209-230, 1973.
- FREITAS, S. M. *Emprego do bootstrap para determinação de regiões de confiança para os localizadores do ponto ótimo de operação em modelos de superfície de respostas*. 1993. 81f. Dissertação (Mestrado em Estatística) - Instituto de Matemática, Estatística e Ciência da Computação, Universidade de Campinas, Campinas, 1993.
- GOMES, M. B. *Intervalos de confiança para as coordenadas do ponto crítico da equação quadrática de regressão a duas variáveis independentes*. 1989. 70f. Tese (Livre-Docência) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 1989.
- KHURI, A. I.; CORNELL, J. A. *Response surfaces - designs and analysis*. 2.ed. New York: Marcel Dekker, 1996. 510p.
- LANCASTER, T. *A note on bootstraps and robustness*. Disponível em: < [http : //www.econ.brown.edu/~ tl/papers/robust.pdf](http://www.econ.brown.edu/~tl/papers/robust.pdf) >. Acesso em: dez. 2007.
- MALHEIROS, E. B.; PERECIN, D. Posição das doses de nutrientes na análise de uma superfície de resposta à adubação. *Revista de Matemática e Estatística*, São Paulo, v.15, p.69-78, 1983.
- MANLY, B. F. J. *Randomization, bootstrap and Monte Carlo methods in biology*. 2.a ed. London: Chapman & Hall, 1997. 399p.
- MEAD, R.; PIKE, D. J. A review of response surface methodology from a biometric viewpoint. *Biometrics*, Washington, v.31, p.803-851, 1975.
- MYERS, R. H.; KHURI; A. L.; CARTER Jr., W. H. Response surface methodology: 1966-1988. *Technometrics*, Alexandria, v.31, p.137-157, 1989.
- MYERS, R. H. Response surface methodology: current status and future directions. *Journal of Quality Technology*, Milwaukee, v.31, p.30-44, 1999.
- MYERS, R. H.; MONTGOMERY, D. C. *Response surface methodology*. 2.ed. New York: John Wiley, 2002. 798p.
- MYERS, R. H.; MONTGOMERY; D. C.; VINING, G. G.; BORROR, C. M.; KOWALSKI, S. M. Response surface methodology: a retrospective and literature survey. *Journal of Quality Technology*, Milwaukee, v.36, p.53-77, 2004.

PETERSON, J. J.; CAHYA, S.; CASTILLO, E. del. A general approach to confidence regions for optimal factor levels of response surfaces. *Biometrics*, Washington, v.58, p.422-431, 2002.

PIMENTEL GOMES, F. *Curso de estatística experimental*. Piracicaba: Nobel, 1990. 468p.

PIMENTEL GOMES, F.; CONAGIN, A. Experimentos de adubação: planejamento e análise estatística. In: SIMPÓSIO DE ESTATÍSTICA APLICADA À EXPERIMENTAÇÃO AGRONÔMICA-REUNIÃO ANUAL DA REGIÃO BRASILEIRA DA SOCIEDADE INTERNACIONAL DE BIOMETRIA, 1987, Londrina. *Mini Curso...* Londrina: UEL, 1987. 102p.

R Development Core Team. *R: A language and environment for statistical computing*. Vienna, 2006. URL. Disponível em: < <http://www.R-project.org> >. Acesso em: out. 2006.

RUBIN, D. B. Bayesian bootstrap. *Annals of Statistics*, Washington, v.9, p.130-134, 1981.

SANCHES, A. *Superfícies de resposta em experimentos de adubação: o problema dos pontos de sela*. 1986. 91f. Tese (Doutorado em Estatística e Experimentação Agronômica) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, 1986.

SILVA, R. R.; ZOCCHI, S. S. Regiões de confiança para a localização do ponto estacionário em superfícies de resposta usando "bootstrap" e métodos "kernel", com escolha automática das constantes de suavização. In: REUNIÃO ANUAL DA REGIÃO BRASILEIRA DA SOCIEDADE INTERNACIONAL DE BIOMETRIA, 49., 2004, Uberlândia. *Resumos...* Uberlândia: UFU, 2004. p.108.

SILVERMAN, B. W. *Density estimation for statistics and data analysis*. London: Chapman & Hall, 1986. 175p.

STABLEIN, D. M.; CARTER, W. H.; WAMPLER, G. L. Confidence regions for constrained optima in response surface experiments. *Biometrics*, Washington, v.39, p.754-763, 1983.

WILKS, S. S. *Mathematical statistics*. New York: John Wiley, 1962. 644p.

ZOCCHI, S. S.; SILVA, R. R. Regiões de confiança para a localização do ponto estacionário em superfícies de resposta usando bootstrap. In: SIMPÓSIO DE ESTATÍSTICA APLICADA À EXPERIMENTAÇÃO AGRONÔMICA, 10., 2003, Lavras. *Resumos...* Lavras: UFLA, 2003. p.305.

Recebido em 23.02.2008.

Aprovado após revisão em 15.07.2008.