MELHOR ESTIMADOR BAYESIANO VIA FUNÇÃO DE RISCO DO PROBLEMA

Osmir José Lavoranti - Embrapa Florestas / Doutorando em Estatística e Experimentação Agronômica na ESALQ/USP(osmir@cnpf.embrapa.br). **José Eduardo Corrente -** Professor Associado, Departamento de Ciências Exatas da ESALQ/USP (jecorren@esalq.usp.br).

RESUMO

Do ponto de vista clássico, diferentes estimadores são comparados via função de risco, o que na maioria das vezes torna-se impraticável. Já para os Bayesianos todos os problemas inferenciais acerca do parâmetro θ poderão ser resolvidos através da densidade a posteriori, que é proporcional ao produto da função de verossimilhança e da distribuição a priori, o que possibilita a obtenção de diferentes posterioris e conseqüentemente diferentes estimadores. Assim, o presente trabalho tem como objetivo apresentar um aspecto teórico para a escolha do melhor estimador Bayesiano via função de risco do problema. Como o teorema de Bayes propicia soluções precisas com distribuição a posteriori exata, a escolha do melhor estimador Bayesiano, via função de risco do problema deve ser sempre preferível.

PALAVRAS-CHAVE: Bayes, risco.

BEST BAYESIAN ESTIMATE VIA RISK PROBLEM FUNCTION

Osmir José Lavoranti - Embrapa Florestas / Doctor student in Statistics and Agronomic Planning at ESALQ/USP (osmir@cnpf.embrapa.br). **Dr. José Eduardo Corrente** - Associate Professor, Department of Mathematics and Statistics at ESALQ/USP (jecorren@esalq.usp.br).

SUMMARY

From the classical point of view, comparing different estimates via risk function is something out of question. Once, from the Bayesian view, inferential problems about the parameter θ can be solved by posteriori density, which is proportional to the product of the likelihood function and the priori distribution, allowing to obtain different posteriori and, however, different estimates. Thus, the aim of this work is to present a theoretical aspect to choose the best Bayesian estimate via risk problem function. As the Bayes theorem provides accurate solutions whit an exact posteriori distribution, the choice of the best Bayesian estimate via risk problem function should always be preferable.

KEY-WORDS: Bayesian, risk.

1. INTRODUÇÃO

A inferência Bayesiana surgiu em 1763, com a publicação da obra “An Essay Towards Solving a Problem in the Doctrine of Chances”, escrita por Thomas Bayes. Desde então, o Teorema de Bayes é um elemento essencial para a análise Bayesiana, pois toda inferência é feita a partir da posteriori, que em termos de distribuição de probabilidade contínua, tem a seguinte formulação:

$$f(\theta \mid y) = \frac{f(\theta) f(y \mid \theta)}{\int_R f(y \mid \theta) f(\theta) d\theta}$$

sendo: θ, um vetor de parâmetros; **y**, o vetor de dados ou de informações obtidas por amostragem; **f(θ|y)**, a distribuição condicional de θ dado y, ou **distribuição a posteriori** (que é a base da estimação e predição Bayesiana); **f(y|θ)**, a função densidade de probabilidade da distribuição condicional (y) dado θ (denomina função de verossimilhança); **f(θ)**, a função densidade de probabilidade da **distribuição a priori**. Esta função denota o grau de conhecimento acumulado sobre θ, antes da observação de y e **f(y|θ) f(θ)**, a função densidade conjunta de y e θ (O´Hagan, 1994).

O problema da abordagem Bayesiana é devido à escolha da distribuição a priori, pois esta, muitas vezes induz a uma distribuição a posteriori imprópria. Como do ponto de vista Bayesiano, todos os problemas inferenciais acerca de θ poderão ser resolvidos através da densidade a posteriori, as inferências sobre θ podem ser errôneas e de baixa qualidade.

Por outro lado, do ponto de vista clássico, diferentes estimadores são comparados via função de risco, o que dificulta a escolha do melhor estimador (Lavoranti, 2001).

O presente trabalho tem como objetivo apresentar um aspecto teórico para a escolha do melhor estimador Bayesiano via análise do risco do problema.

2. MELHOR ESTIMADOR BAYESIANO

Do ponto de vista clássico, diferentes estimadores são comparadas via função de risco.

$$R_{\hat{\theta}}(\theta) = E_{\theta}\left[L\left(\theta, \hat{\theta}(y)\right)\right] \xrightarrow{\text{Perda quadrática}} E_{\theta}\left(\hat{\theta}(y), \theta\right)^2$$

Se θ é aleatório então $r\left(\hat{\theta}\right) = E\ R_{\theta}(\theta) = \int_{\Theta} R_{\hat{\theta}}(\theta)\ f(\theta)\ d\theta$, ou seja,

$$r_{\hat{\theta}} = E_{\theta}\ E_{y|\theta}\left[L\left(\theta, \hat{\theta}(y)\right)\right] = \int_{\Theta} \int_{\theta} L\left(\theta, \hat{\theta}(y)\right) f(y \mid \theta)\ f(\theta)\ dy\ d\theta =$$

$$= E_y \; E_{\theta|y}\left[L\left(\theta,\ \hat{\theta}(y)\right)\right] \;=\; \int_{\lambda}\int_{\Theta} L\left(\theta,\ \hat{\theta}(y)\right)\ f(\theta \mid y)\ f(y)\ d\theta\ dy$$

$$= \int_{\lambda}\left[\int_{\Theta} L\left(\theta,\ \hat{\theta}(y)\right)\ f(\theta \mid y) d\theta\right] f(y)\ dy \;=\; \int_{\lambda}\left[r_{\hat{\theta}}(y)\right]\ f(y)\ dy$$

Observa-se, dessa forma, a dificuldade em se comparar dois estimadores do ponto de vista clássico.

Do ponto de vista Bayesiano, $\hat{\theta}$ é melhor do que $\tilde{\theta}$ se $r_{\hat{\theta}} \leq r_{\tilde{\theta}}$. Ou seja, o estimador $\hat{\theta}$ é ótimo se $r_{\hat{\theta}} \leq r_{\tilde{\theta}}$ para todo outro estimador $\tilde{\theta}$.

A quantidade $r = \min_{\theta}\{ r_{\hat{\theta}} \}$ é chamada de risco do problema. Como $r_{\hat{\theta}} = \int_{\lambda}\left[\int_{\Theta} L\left(\theta,\ \hat{\theta}(y)\right)\ f(\theta \mid y) d\theta\right] f(y)\ dy = \int_{\lambda}\left[r_{\hat{\theta}}(y)\right]\ f(y)\ dy$, e como $\left[r_{\hat{\theta}}(y)\right]$ e f(y) são maiores que zero, queremos escolher $\hat{\theta}(y)$ tal que a integral que define $r_{\hat{\theta}}$ seja mínima.

Para isso, é suficiente que $\hat{\theta}(y)$ minimize $r_{\hat{\theta}} = \int_{\lambda}\left[\int_{\Theta} L\left(\theta,\ \hat{\theta}(y)\right)\ f(\theta \mid y) d\theta\right] f(y)\ dy$.

3. EXEMPLO

Sejam, $Y|\theta \cap B(\theta)$; $Y_1|\theta, Y_2|\theta, \ldots, Y_n|\theta$ são variáveis aleatórias i.i.d.; $Y_1|\theta, Y_2|\theta, \ldots, Y_n \mid \theta \cap$ i.i.d. $B(\theta)$; $\theta \cap U(0,1)$, $0 < \theta < 1$ e $f(\theta) = I_{(0,1)}(\theta)$ (densidade a priori), então,

$$f\left(\underset{\sim}{y} \mid \theta\right) = \theta^{\sum_{i=1}^{n} y_i}(1-\theta)^{n-\sum_{i=1}^{n} y_i}$$

$$f\left(\underset{\sim}{y}\right) = \int_0^1 f\left(\underset{\sim}{y} \mid \theta\right) f(\theta) d\theta = \int_0^1 \theta^{\sum_{i=1}^{n} y_i}(1-\theta)^{n-\sum_{i=1}^{n} y_i}\ I_{(0,1)}(\theta)\ d\theta = \mathrm{Beta}\left(\sum_{i=1}^{n} y_i + 1, n - \sum_{i=1}^{n} y_i + 1\right) =$$

$$= \frac{\Gamma\left(\sum_{i=1}^{n} y_i + 1\right)\Gamma\left(n - \sum_{i=1}^{n} y_i + 1\right)}{\Gamma(n+2)} = \frac{\left(\sum_{i=1}^{n} y_i\right)!\left(n - \sum_{i=1}^{n} y_i\right)!}{(n+1)!} = \frac{1}{n+1}\begin{pmatrix} n \\ \sum_{i=1}^{n} y_i \end{pmatrix}^{-1}$$

$$f\left(\theta \mid \underset{\sim}{y}\right)=\frac{f\left(\underset{\sim}{y} \mid \theta\right)f(\theta)}{f\left(\underset{\sim}{y}\right)}=\frac{\theta^{\sum_{i=1}^{n} y_i}(1-\theta)^{n-\sum_{i=1}^{n} y_i}}{\frac{\Gamma\left(\sum_{i=1}^{n} y_i+1\right)\Gamma\left(n-\sum_{i=1}^{n} y_i+1\right)}{\Gamma(n+2)}}=\frac{\theta^{\sum_{i=1}^{n} y_i}(1-\theta)^{n-\sum_{i=1}^{n} y_i}\quad\Gamma(n+2)}{\Gamma\left(\sum_{i=1}^{n} y_i+1\right)\Gamma\left(n-\sum_{i=1}^{n} y_i+1\right)}\Rightarrow$$

$\Rightarrow\theta \mid \underset{\sim}{y} \cap Beta\left(\sum_{i=1}^{n} y_i+1,\ n-\sum_{i=1}^{n} y_i+1\right)$ (distribuição a posteriori) e $f\left(\theta \mid \underset{\sim}{y}\right)$ é a densidade a posteriori. O que implicaria na $\mathrm{E}\left[\theta \mid \underset{\sim}{\mathrm{Y}}\right]=\frac{\sum_{i=1}^{n} y_i+1}{n+2}$, e conseqüentemente,

$$r(y) \;=\; E\left[\left(\theta-\hat{\theta}(y)\right) \;\mid\; Y=y\right]=\frac{\left(n-\sum_{i=1}^{n} y_i+1\right)\left(\sum_{i=1}^{n} y_i+1\right)}{(n+2)^2(n+3)}.$$

4. CONSIDERAÇÕES FINAIS

Como a distribuição a posteriori é proporcional ao produto da verossimilhança e da distribuição a priori, à obtenção de momentos a posteriori nem sempre é simples e possível. Vê-se contudo, a possibilidade de obtenção de diferentes posteriori e conseqüentemente diferentes estimadores. Muito embora, o pesquisador recorra a métodos numéricos ou de aproximação para obter os momentos a posteriori, o qual permite o cálculo dos momentos a posteriori sem a necessidade do cálculo da distribuição a posteriori marginal, os estimadores obtidos são inferiores aqueles obtidos diretamente da distribuição a posteriori, pois o teorema de Bayes propicia soluções precisas com distribuição a posteriori exata.

Dessa forma, a escolha do melhor estimador Bayesiano, via função de risco do problema deve ser sempre preferível.

5. REFERÊNCIAS BIBLIOGRÁFICAS

O´HAGAN, A. **Bayesian inference. Kendall´s advanced theory of statistic.** London: Cambridge University Press, 1994. 330p.

LAVORANTI, O.J. Inferência bayesiana aplicada na estimação de componentes de variância e predição de valores genéticos. Piracicaba: ESALq, 2001. 31p. (Seminário da Disciplina Inferência Estatística II, do Curso doutorado em Estatística e Experimentação Agronômica).

This document was created with Win2PDF available at http://www.daneprairie.com.
The unregistered version of Win2PDF is for evaluation or non-commercial use only.