

Estimação bayesiana dos parâmetros da TRI

Euler Rorigues de Alencar, Wesley de Jesus Silva e Cibele Queiroz da Silva

Departamento de Estatística

Instituto de Ciências Exatas - Universidade de Brasília

Campus Universitário Darcy Ribeiro - Asa Norte

70910-900 - Brasília - DF - Brasil,

era_alencar@yahoo.com.br, wesley.jesus@gmail.com, cibeles@unb.br

Resumo

Este trabalho envolve o estudo de modelos de teoria de resposta ao item, de processos markovianos em estados contínuos e inferência bayesiana à estimação dos parâmetros da Teoria de Resposta ao Item (TRI).

Na literatura, há diversas formas de estimação dos parâmetros da TRI. O enfoque é dado à estimação bayesiana desses parâmetros, mais especificamente aos métodos bayesianos de estimação utilizando técnicas de Monte Carlo via Cadeias de Markov (MCMC). Serão utilizados algoritmos tradicionais, tais como o amostrador de Gibbs e o método de Metropolis-Hasting, a fim de simular uma amostra da distribuição a posteriori e assim obter as estimativas dos parâmetros de interesse.

Em geral, interessa-se estimar tanto os parâmetros de habilidades quanto os parâmetros dos itens. Porém, o objetivo é obter apenas estimativas dos parâmetros dos itens, e um enfoque maior será dado ao modelo de respostas binárias com dois parâmetros.

Palavras-chaves: *TRI, Inferência Bayesiana, Cadeias de Markov, Algoritmos MCMC, Convergência.*

1 TRI e o modelo logístico de 2 parâmetros

A Teoria da Resposta ao Item, TRI, é uma modelagem estatística utilizada em medidas de traços ou habilidades latentes de indivíduos a partir das suas respostas a um conjunto de itens. O principal enfoque da TRI está na probabilidade de determinada resposta de determinado indivíduo, condicionada à sua proficiência (traço latente). Existem diversas formas de modelar esse processo, mas neste trabalho tratará do *modelo logístico de 2 parâmetros* para respostas binárias (0 ou 1):

$$P(U_{ij} = 1|\theta_j) = \frac{1}{1 + e^{-Da_i(\theta_j - b_i)}}. \quad (1)$$

Onde:

- a_i : Parâmetro de *discriminação* do item i ;
- b_i : Parâmetro de *dificuldade* do item i ;
- θ_j : Traço latente do indivíduo j ;
- D : Fator de escala, constante e igual a 1 ou 1,7.

Motivação

Para a estimação dos parâmetros em (1), utiliza-se métodos de Máxima Verossimilhança Marginal em conjunto com o método de Newton-Raphson, métodos de quadratura, algoritmo EM, entre outros. Porém, em alguns desses métodos, existem limitações (como nos casos em que um indivíduo “acerta” todas as questões) além de problemas com as estimativas (discriminação negativa, por exemplo). Para contornar tais problemas, sugeriu-se uma estimação bayesiana através de simulação estocástica.

2 Estimação através do algoritmo de Metropolis-Hastings com Amostrador de Gibbs

Um método MCMC (*Markov Chain Monte Carlo*) de simulação de uma distribuição f é qualquer método que produz uma Cadeia de Markov ergódica cuja

distribuição estacionária é f . Em inferência Bayesiana, f é definida como sendo a distribuição a posteriori.

Um dos mais tradicionais métodos MCMC é o método do amostrador de Gibbs, que utiliza como kernel de transição as *distribuições condicionais completas*, $\pi(\theta_i|\theta_{-i}, Y)$, conhecidas. E assim obtém-se uma amostra da distribuição à posteriori de $(\theta_1, \dots, \theta_k)$. Quando $\pi(\theta_i|\theta_{-i}, Y)$ é desconhecido, recorre-se ao algoritmo de Metropolis-Hastings (MH), um processo de amostragem baseado no método da aceitação/rejeição, que utiliza uma *distribuição geradora de candidatos* $q(\cdot|\theta)$.

MH com Gibbs na TRI

Por questões computacionais, reparametrizou-se o modelo (1), determinando-se $\beta_{2i} = a_i$ e $\beta_{1i} = \frac{b_i}{a_i}$. Feito isso, implementou-se o MH com Gibbs. Em cada iteração k do algoritmo, procedeu-se da seguinte maneira:

1. para cada j relacionado a um indivíduo, seleciona-se um candidato $\theta_j^{(*)}$ de uma geradora de candidatos, que é aceito com probabilidade α_θ ;
2. em seguida, seleciona-se um candidato $\beta_i^{(*)}$ de outra geradora de candidatos, aceito com probabilidade α_β ;

Formalmente, os α 's são dados por:

- Formalmente, os α 's são dados por:

$$\alpha(\theta_j^{k-1}, \theta_j^*) = \min \{R_\theta, 1\}$$

onde

$$R_\theta = \frac{P(X|\beta^{k-1}, \theta_j^*)P(\theta_j^*)q_\theta [(\theta^*, \theta^{k-1})]}{P(X|\beta^{k-1}, \theta_j^{k-1})P(\theta_j^{k-1})q_\theta [(\theta^{k-1}, \theta^*)]}$$

- e

$$\alpha(\beta_i^{k-1}, \beta_i^*) = \min \{R_\beta, 1\}$$

onde

$$R_\beta = \frac{P(X|\theta^k, \beta_i^*)P(\beta_i^*)q_\beta [(\beta^*, \beta^{k-1})]}{P(X|\theta^k, \beta_i^{k-1})P(\beta_i^{k-1})q_\beta [(\beta^{k-1}, \beta^*)]}$$

- Prioris consideradas:

$$\begin{aligned} p(\theta_j) &= n(\theta_j|0, 1); \\ p(\beta_{1i}) &= n(\beta_{1i}|0, 1); \\ p(\beta_{2i}) &= \text{lognormal}(\beta_{2i}|0, 2), \end{aligned}$$

- Distribuições Geradoras de Candidatos:

$$q_{\theta}(\theta_j^{k-1}, \theta_j) \sim n(\theta_j | \theta_j^{k-1}, 1.2);$$

$$q_{\beta_1}(\beta_{1i}^{k-1}, \beta_{1i}) \sim n(\beta_{1i} | \beta_{1i}^{k-1}, 0.01);$$

$$q_{\beta}(\beta_{2i}^{k-1}, \beta_{2i}) \sim \text{lognormal}(\beta_{2j} | \beta_{2i}^{k-1}, 0.01).$$

3 Metodologia e Resultados

Simulou-se 3000 respostas indivíduos a 10 itens (totalizando 20 parâmetros a serem estimados). Para cada parâmetro, construiu-se uma cadeia com 25000 interações. As cadeias foram construídas através de funções definidas no R. Para verificar a convergência das cadeias, utilizou-se vários testes, entre eles:

- **Geweke:** baseia-se na diferença entre médias de duas partes da cadeia;
- **Gelman e Rubin:** analisa a diferença entre quantis da cadeia;
- **Heidelberger e Welch:** testa convergência e verifica se a média temporal pode ser utilizada como média a posteriori.
- Gráficos do Diagnóstico de Gelman e Rubin para a cadeia Beta(2,9) (β_2 do item 9)

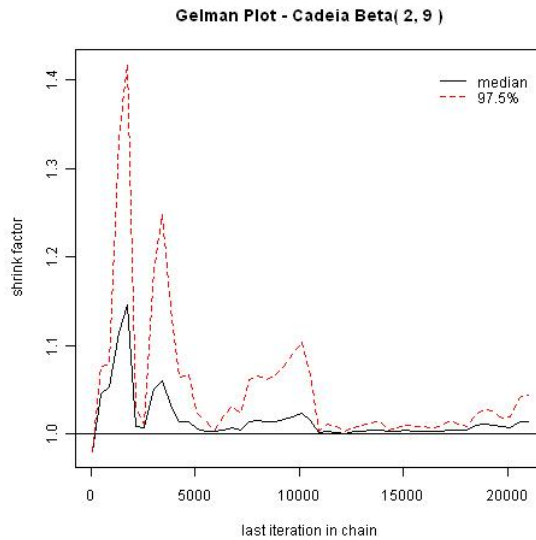


Figura 1: Gráfico de Gelman e Rubin- cadeia Beta(2,9)

As convergências das cadeias foram aceitas pelos diagnósticos de Heidelberger e Welch, e as médias das cadeias podem ser usadas como média a posteriori. Confirmadas as convergências, selecionou-se uma amostra sistemática de cada cadeia com intervalo de seleção igual a 100 (*lag* com autocorrelações pouco significativas). As curvas estimadas do modelo (1) para os 10 itens seguem. A linha vermelha pontilhada é a curva real (esperada), e a linha compacta azul representa a curva estimada.

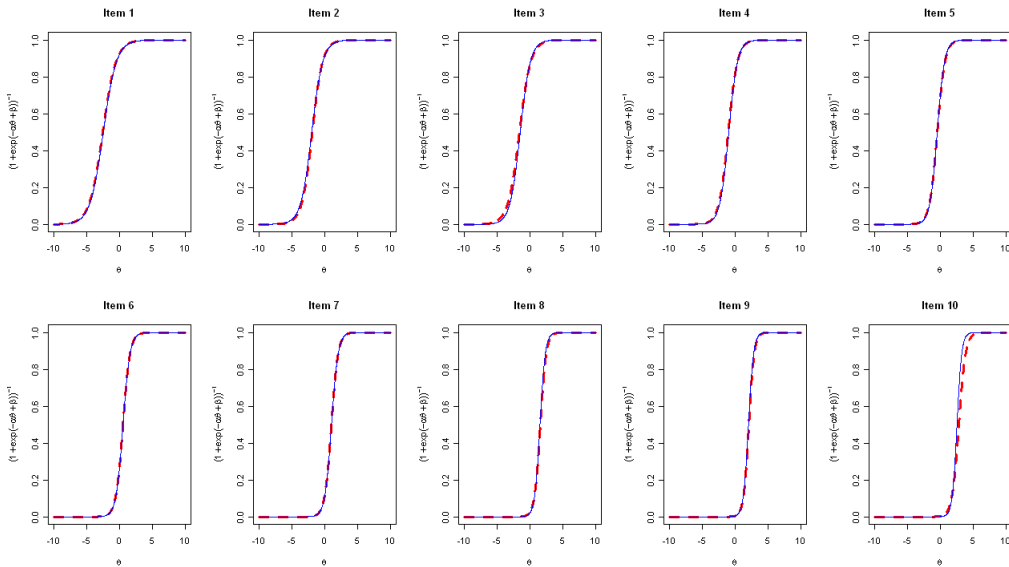


Figura 2: Curvas logísticas estimadas dos itens

Conclusões

Os testes referidos acima, embora tenham apresentado problemas para alguns parâmetros, não indicaram motivos para descartar convergência e estacionariedade de qualquer uma das cadeias. Porém, fica claro que é necessário obter cadeias maiores, pois alguns itens (como o item 9) apresentaram dificuldades para alcançar convergência.

Em geral, as estimativas a posteriori foram bem próximas aos valores reais simulados. Apesar do viés pouco acentuado observado no item 10, todas as outras estimativas foram bem próximas do valor real utilizado na simulação.

Referências

1. ANDRADE, D.F., TAVARES, H.R., VALLE, R.C. (2000) *Teoria de Resposta ao Item: Conceitos e Aplicações*, Associação Brasileira de Estatística - ABE, SINAPE 2000.
2. da-SILVA, C.Q. e Tibúrcio, J.D (2003). *Estimação do Tamanho Populacional de Mamíferos Aquáticos via Método Bayesiano Empírico Baseado em Dados de Foto-identificação*. Universidade Federal de Minas Gerais.
3. GAMERMAN, D. and H. LOPES (2006). *Markov chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Texts in Statistical Science Series. CRC Press.
4. PATZ, Richard J. and JUNKER, Brian W. (1999). *A Straightforward Approach to Markov Chain Monte Carlo Methods for Item Response Models*. Journal of Educational and Behavioral Statistics, Vol. 24, No. 2, pp. 146-178.