

Inferência Bayesiana em modelos de mistura finita: uma aplicação à estimação de densidades

Guilherme Lopes de Oliveira^{1 3}

Rosângela Helena Loschi^{2 3}

Um modelo de mistura finita é uma combinação convexa de densidades cada uma com peso. Misturas de distribuições de probabilidade são muito utilizadas para modelar o comportamento de dados heterogêneos como, por exemplo, análise de agrupamentos, de outliers, de pontos de mudança e na estimação de densidades. Neste estudo, uma mistura de densidades normais foi utilizada no contexto de estimação de densidade. A inferência foi feita sob o paradigma Bayesiano. Com o intuito de garantir a identificabilidade do modelo, incluiu-se uma restrição à média de umas das densidades presentes na mistura. Na estimação dos parâmetros envolvidos no modelo métodos Monte Carlo via cadeia de Markov foram considerados juntamente com a técnica de aumento de dados, i.e; variáveis latentes foram incluídas no modelo visando facilitar a implementação computacional. Dados foram gerados de misturas de distribuições normais considerando-se diversas especificações para os pesos e variâncias das densidades. Considerou-se distribuições a priori pouco informativas para todos os parâmetros. Comparou-se as curvas da densidade preditiva a priori e das densidades estimadas via método plug-in usando média e moda a posteriori com a curva da densidade real de cada modelo. Resumidamente, conclui-se que, a variância das densidades presentes na mistura tem uma influência importante na qualidade das estimativas a posteriori. Estas são piores quando os dados provêm de uma mistura de densidades que coloca maior peso para o componente da mistura que apresenta variância maior. Tem-se ainda que a densidade preditiva a priori e as densidades estimadas via método plug-in são bastante influenciadas pelos dados e apresentam uma estimativa razoável para a densidade real.

Palavras-chave: aumento de dados, distribuição preditiva a priori, métodos plug-in.

¹ Departamento de Estatística - ICEx – Universidade Federal de Minas Gerais. Email: guilherme.foxx@hotmail.com

² Departamento de Estatística - ICEx – Universidade Federal de Minas Gerais. Email: loschi@est.ufmg.br

³ Agradecimento à FAPEMIG pelo apoio financeiro.