46th Annual Iranian Mathematics Conference

25-28 August 2015

Yazd University

Robust mixture regression model fitting by slash distribution with application to musical tones

Hadi Saboori
Department of Statistics,
University of Zabol, Zabol, Iran

Sobhan Shafiei*
Department of Statistics,
Shahid Bahonar University
of Kerman, Kerman, Iran

Afsaneh Sepahdar
Department of Statistics,
Shahid Bahonar University
of Kerman, Kerman, Iran

Abstract

The traditional estimation of mixture regression models is based on the normal assumption of component errors and thus is sensitive to outliers or heavy-tailed errors. A robust mixture regression model based on the slash distribution by extending the mixture of slash distributions to the regression setting is proposed. Using the fact that the slash distribution can be written as a scale mixture of a normal and a latent distribution, this procedure is implemented by an EM algorithm. Finally, the proposed method is compared with other procedures, based on a real data set.

Keywords: EM algorithm, Normal mixture regression, Outliers
Mathematics Subject Classification [2010]: 62J05, 62F35

1 Introduction

Mixture regression models (MRM) are well known as switching regression models in the econometrics literature, which were introduced by Goldfeld and Quandt [4]. These models have been widely used to investigate the relationship between variables coming from several unknown latent homogeneous groups and applied in many fields, such as business, marketing, and social sciences.

In general, a normal mixture regression model $(\mathrm{N-MRM})$ is defined as: let Z be a latent class variable such that given $Z = j$, the response y depends on the p-dimensional predictor $\mathbf{x}$ in a linear way

$$Y = \mathbf{x}^\top \boldsymbol{\beta}_j + \epsilon_j, \quad j = 1, \ldots, m, \tag{1}$$

where m is the number of groups (also called components in mixture models) in the population, the $\boldsymbol{\beta}_j$ are unknown p-dimensional vectors of regression coefficients and $\epsilon_j \sim N(0, \sigma_j^2)$ is independent of $\mathbf{x}$. Suppose $P(Z = j) = \pi_j$ and Z is independent of $\mathbf{x}$, then the conditional density of Y given $\mathbf{x}$, without observing Z, is

*Speaker