Journal of Statistical Planning and Inference 215 (2021) 268–288

Contents lists available at ScienceDirect

Journal of Statistical Planning and Inference

journal homepage: www.elsevier.com/locate/jspi

Density estimation of a mixture distribution with unknown point-mass and normal error

Dang Duc Trong [a,b,c], Nguyen Hoang Thanh [b,c], Nguyen Dang Minh [a,b,e], Nguyen Nhu Lan [d,*]

[a] Faculty of Maths and Computer Science, University of Science, Ho Chi Minh City, Viet Nam
[b] Vietnam National University, Ho Chi Minh City, Viet Nam
[c] Centre for Mathematical Science, University of Science, Ho Chi Minh City, Viet Nam
[d] Faculty of Basic Science, VanLang University, Viet Nam
[e] Department of Fundamental Studies, Ho Chi Minh City Open University, Viet Nam

A R T I C L E I N F O

Article history:
Received 21 December 2019
Received in revised form 5 April 2021
Accepted 10 April 2021
Available online 18 April 2021

MSC:
62G07
45Q05
62G05

Keywords:
Deconvolution
Mixture distribution
Inversion problems
Nonparametric estimation

A B S T R A C T

We consider the model $Y = X + \xi$ where Y is observable, ξ is a noise random variable with density f_ξ, X has an unknown mixed density such that $\mathbb{P}(X = X_c) = 1 - p$, $\mathbb{P}(X = a) = p$ with X_c being continuous and $p \in (0, 1)$, $a \in \mathbb{R}$. Typically, in the last decade, the model has been widely considered in a number of papers for the case of fully known quantities a, f_ξ. In this paper, we relax the assumptions and consider the parametric error $\xi \sim \sigma N(0, 1)$ with an unknown $\sigma > 0$. From i.i.d. copies $Y_1, \ldots, Y_m$ of Y we will estimate (σ, p, a, f_{X_c}) where f_{X_c} is the density of X_c. We also find the lower bound of convergence rate and verify the minimax property of established estimators.

1. Introduction

In applications, we often have the problem of finding the probability distribution of a random variable X from measurements contaminated with a noise ξ. Simplicity, we can assume the additive model $Y = X + \xi$ with Y observed. Estimation of the target density f_X of X from observations of the random variable Y is called the *(statistical) deconvolution problem*. The literature of the specific problem has grown very rapidly in last decades (see Devroye, 1989; Meister, 2009 and references therein). Precisely, in the present paper, we assume the data model

$$Y_j = X_j + \xi_j, \quad j = \overline{1, m}, \tag{1}$$

where the $Y_j's$ are i.i.d. observable copies of the random variable Y, the $X_j's$ and the $\xi_j's$, $j = \overline{1, m}$, are mutually independent and have the same distribution as X, ξ respectively. As is well known, any estimation tool of deconvolution problem is based on a combination of presumptions about the target density function and the density function of error.

We first discuss shortly the presumption on the target density function. In the deconvolution literature, most of the research is focused on continuous distributions. However, in some applications, the target random variable X is not

* Corresponding author.
E-mail address: lan.nn@vlu.edu.vn (N.N. Lan).

https://doi.org/10.1016/j.jspi.2021.04.002
0378-3758/© 2021 Published by Elsevier B.V.