Journal of the Korean Statistical Society 45 (2016) 505–517

Contents lists available at ScienceDirect

Journal of the Korean Statistical Society

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

Uniform distribution width estimation from data observed with Laplace additive error

M. Benšić, K. Sabo *

Department of Mathematics, University of Osijek, Trg Ljudevita Gaja 6, HR-31000 Osijek, Croatia

A R T I C L E I N F O

Article history:
Received 13 August 2015
Accepted 2 March 2016
Available online 25 March 2016

AMS 2000 subject classifications:
primary 46N30
secondary 62H30

Keywords:
Maximum likelihood estimator
Method of moments estimator
Measurement error
Laplace additive error
Uniform distribution

A B S T R A C T

A one-dimensional problem of a uniform distribution width estimation from data observed with a Laplace additive error is analyzed. The error variance is considered as a nuisance parameter and it is supposed to be known or consistently estimated before. It is proved that the maximum likelihood estimator in the described model is consistent and asymptotically efficient and sufficient conditions for its existence are given. The method of moment estimator is also analyzed in this model and compared with the maximum likelihood estimator theoretically and in simulations. Finally, one real-world example illustrates the possibility for applications in two-dimensional problems.

1. Introduction

How to estimate the support of a uniform distribution from the data measured with additive errors is the problem that comes from different applications. Generally, if the object is measured with errors, it is complicated to determine its edges by using the well known edge detection methods (see e.g. Qiu, 2005) and consequently, it is not easy to accurately estimate its dimensions either. Such problems can arise, for example, when the object is observed with a fluorescent microscope (Ruzin, 1999), a ground penetrating radar, ultrasound, etc. The same type of model can be used in the problem of protein secondary structure assignment (Frishman & Argos, 1995; Kabsch & Sander, 1983), detection of shapes in image analysis (Garlipp & Müller, 2006), etc.

To illustrate applications more precisely, let us mention the problem of estimating the size of black fungi colonies on the basis of a monochromatic image (see Fig. 1). Here we have seven colonies captured with errors and we focus our attention on a diameter of each colony (see Garlipp & Müller, 2006). In Section 5.3 we give a complete analysis of this problem and present the results obtained by the method presented in this paper.

The basic statistical model that we are going to use for our purpose is a parametric one and it is built for one-dimensional data in the data set. It is a simple random sample $(X_1, \ldots, X_n)$ from the distribution that is a mixture of two independent random variables U and Y

$$X = U + Y,$$

where U is uniformly distributed on the segment $[-a, a]$, $a > 0$, and Y is the error variable. In this setting, the estimation of the parameter a means also the estimation of the boundary points.

* Corresponding author.
E-mail addresses: mirta@mathos.hr (M. Benšić), ksabo@mathos.hr (K. Sabo).

http://dx.doi.org/10.1016/j.jkss.2016.03.001

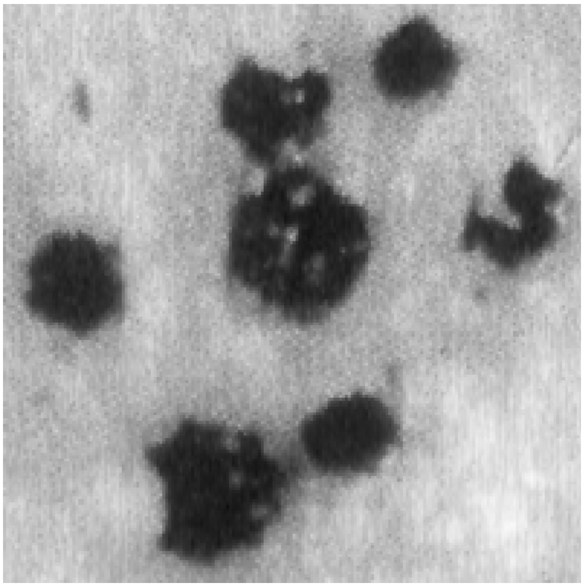

Fig. 1. Real world problem: Estimation the size of colonies of black fungi (Garlipp & Müller, 2006; Sterflinger & Krumbein, 1995).

Boundary estimation in the presence of measurement errors is a problem extensively treated in literature in different contexts (see e.g. Delaigle & Gijbels, 2006, Kneip & Simar, 1996, Meister, 2006 and Neumann, 1997). Although this problem generally can be treated as a classical deconvolution density estimation problem (see e.g. Carroll & Hall, 1988, Fan, 1991, Feuerverger, Kim, & Sun, 2008, Stefanski & Carroll, 1990 and Zhang, 1990), these methods usually face problems at density discontinuity points. That is why the modified kernel estimator has been proposed in Zhang and Karunamuni (2000) if the boundary points are of interest. Also, in Delaigle and Gijbels (2006) a diagnostic function which is proportional to the derivative of the deconvolution kernel density estimator has been optimized in order to estimate the boundary points. However, some computational problems for an easy application remain, for instance the choice of bandwidth which is very important for a good performance of any kernel estimator.

Here we discuss a completely different approach which assumes the error distribution type to be known. Similar models have been analyzed in Benšić and Sabo (2007a,b, 2010) and Schneeweiss (2004). In these papers, the basic model assumes that the distribution of the part U is uniform on an interval and the distribution of the part Y, which is considered an error, is Gaussian. The maximum likelihood approach applied in the aforementioned papers shows that the length of the uniform support can be estimated consistently and in an asymptotically efficient way even if we have the added Gaussian error part in the data. Also, the asymptotic variance of the estimator is calculated so we can construct confidence intervals and statistical tests about the length of the uniform support in a classical parametric way.

However, the estimation procedures based on the ML estimator from the mentioned papers are shown to be very sensitive to outliers in applications. Unfortunately, the outliers are often present in images, so we tried to consider the estimation procedure which is less sensitive to outliers and allows a parametric approach at the same time. As expected, the ML estimator derived from a similar model but with a Laplace error distribution fulfilled our expectation.

In this paper, we present results which are based on the assumptions that U is uniformly distributed on the interval $[-a, a]$ for some $a > 0$, which is to be estimated, and Y is a Laplace random variable with a location parameter $\mu = 0$ and a scale parameter $\lambda > 0$.

If U is uniform with a density function

$$f_U(x|a) = \begin{cases} \frac{1}{2a}, & x \in [-a, a] \\ 0, & \text{else}, \end{cases} \tag{1}$$

and Y has a Laplace distribution with a location parameter $\mu = 0$ and a scale parameter $\lambda > 0$ i.e. a density function

$$f_Y(x|\lambda) = \frac{1}{2\lambda} e^{-\frac{|x|}{\lambda}}, \tag{2}$$

then a density function for X has a form

$$f_X(x|a, \lambda) = \frac{1}{2a}\left(G\left(\frac{x+a}{\lambda}\right) - G\left(\frac{x-a}{\lambda}\right)\right), \tag{3}$$