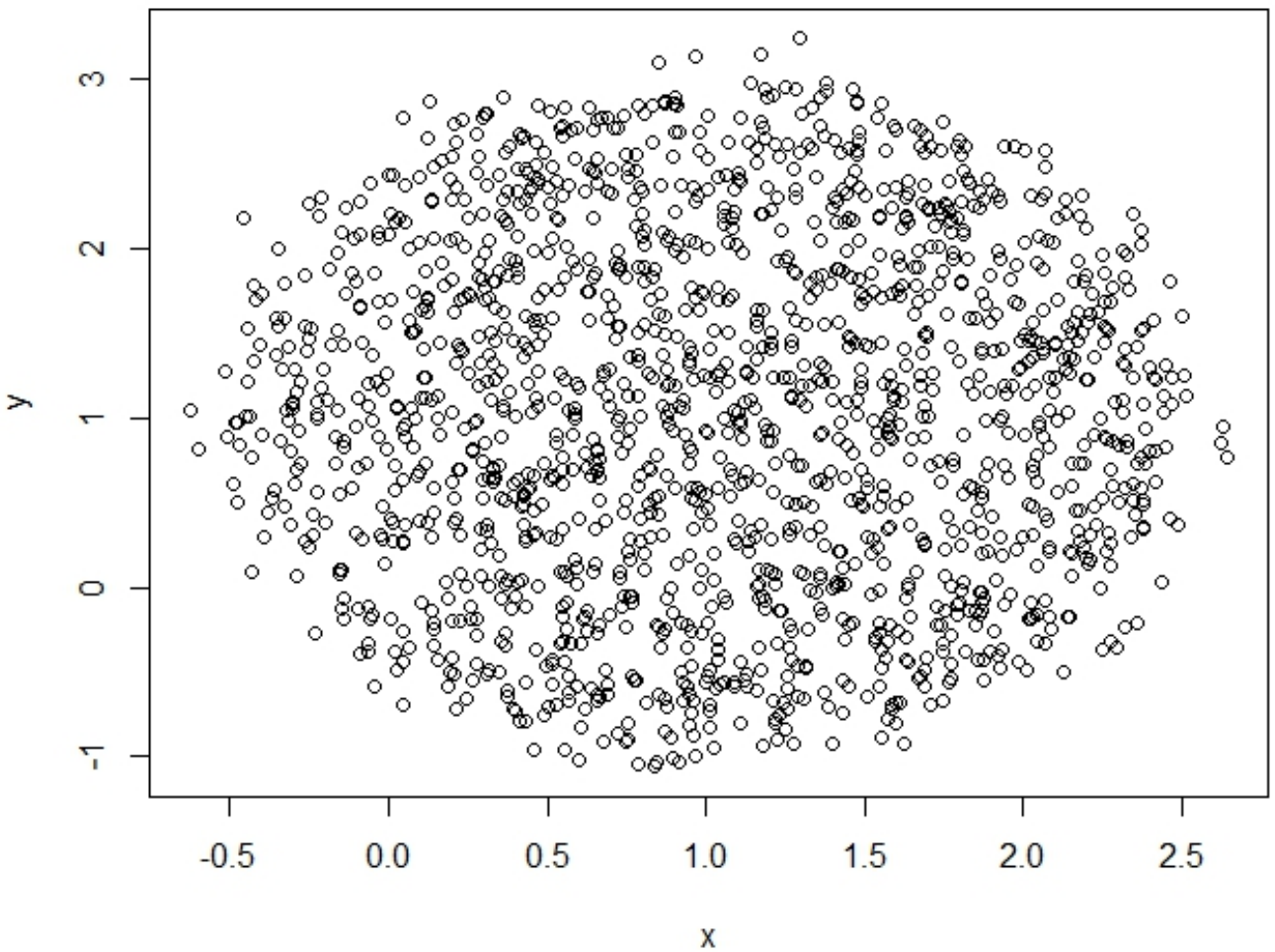
LeArEst - The Software for Border and Area Estimation of Data Measured with Additive Error

Petar Taler, Safet Hamedović, Mirta Benšić, Emmanuel Karlo Nyarko

This work was supported by the Croatian Science Foundation through research grant IP-2016-06-6545

ELMAR 2017, Zadar, September 2017

Example



Problem

$$X = U + \varepsilon$$

- U and ε independent
- U uniform on some region
- ε error

Model

$$X = U + \varepsilon$$

- U and ε independent random variables
  - U uniform on some region
    One-dimensional example:

$$f_U(x; a) = \begin{cases} \frac{1}{2a}, & x \in [-a, a] \\ 0, & otherwise \end{cases}$$

  - One-dimensional examples for ε
    - normal

$$f_\varepsilon^{(\mathcal{N})}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/(2\sigma^2)}$$

    - Laplace

$$f_\varepsilon^{(\mathcal{L})}(t) = \frac{1}{2\lambda} e^{-|x|/\lambda}$$

General one-dimensional model

- $X = U + \varepsilon$
- $$\Rightarrow \quad f_X(x; a) = \frac{1}{2a} \left(F_\varepsilon(x + a) - F_\varepsilon(x - a) \right)$$
- Data x_i, $i = 1, \ldots, m$ from independent replications of the model variable

Goals

To estimate:

- $a > 0$ (or $2a$, a length of the uniform support)
- $\sigma > 0$, error variance.

General one-dimensional model

- $X = U + \varepsilon$
- $$\Rightarrow \quad f_X(x; a) = \frac{1}{2a} \left(F_\varepsilon(x + a) - F_\varepsilon(x - a) \right)$$
- Data x_i, $i = 1, \ldots, m$ from independent replications of the model variable

Goals

To estimate:

- $a > 0$ (or $2a$, a length of the uniform support)
- $\sigma > 0$, error variance.

Two-dimensional model

- reduce the original problem to several corresponding one-dimensional problems
- set of the data points: $D = \{(x_i, y_i),\ i = 1, \ldots, n\}$

Algorithm - Transformation through the y-axis

Step 1 separating through y-axis

Choose an integer $m < n$ and real numbers $\eta_1 < \eta_2 < \cdots < \eta_m$ such that

(i) $\eta_1 \leq \min\{y_i : i = 1, \ldots n\},\ \max\{y_i : i = 1, \ldots n\} \leq \eta_m$ and

(ii) $C_k := \{(x_i, y_i) \in D : y_i \in [\eta_k, \eta_{k+1}]\}$ is a nonempty set.

Two-dimensional model

Algorithm - Transformation through the y-axis; cont'd

Step 2 centering through y-axis

Let us denote

$$c_k := \frac{1}{|C_k|} \sum_{(x_i,y_i)\in C_k} x_i, \quad d_k := \frac{1}{|C_k|} \sum_{(x_i,y_i)\in C_k} y_i,$$

$$k = 1, \ldots, m-1$$

For $k = 1, \ldots, m-1$ define $\overline{C}_k := \{x_i - c_k : (x_i, y_i) \in C_k\}$.

- sets $\overline{C_k}$ represent centered tiny strips $\rightarrow$ one-dimensional model
- $\Rightarrow$ border of the domain

Some references

- H. Schneeweiss, *Estimating the endpoint of a uniform distribution under measurement errors*, CEJOR 12 (2004), 221–231.
- M. Benšić, K. Sabo, *Estimating the width of a uniform distribution when data are measured with additive normal errors with known variance*, Computational Statistics and Data Analysis, 51(2007), 4731-4741
- M. Benšić, K. Sabo, *Border estimation of a Two-dimensional Uniform Distribution if Data are Measured with Additive Error*, Statistics, 41 (2007), 4, 311–319.
- K. Sabo, M. Benšić, *Border estimation of a disc observed with random errors solved in two steps*, Journal of Computational and Applied Mathematics, 229 (2009)
- M. Benšić, K. Sabo, *Estimating a uniform distribution when data are measured with a normal additive error with unknown variance*, Statistics, 44 (2010), 235–246.
- M. Benšić, K. Sabo, *Uniform distribution width estimation from data observed with Laplace additive error*, Journal of the Korean Statistical Society, 45 (2016), 505–517.

LeArEst package

- software for border and area estimation of data measured with additive error
- package for R programming language, available on CRAN: `https://cran.r-project.org/package=LeArEst`
- border and area estimation
- objects may be defined numerically, or recorded in picture

LeArEst package - border estimation, function *lengthest()*

Function input

- vector of input data,
- error distribution (normal, Laplace, Student with 5 degrees of freedom),
- error variance or estimation method (Method of Moments, Maximum Likelihood Method),
- confidence level.

Function output

- estimated half-width of uniform distribution,
- error variance, estimated or given,
- used method for computing a confidence interval (asymptotic distribution of ML or likelihood ratio statistic).

LeArEst package - border estimation, function *lengthest()*

Function input

- vector of input data,
- error distribution (normal, Laplace, Student with 5 degrees of freedom),
- error variance or estimation method (Method of Moments, Maximum Likelihood Method),
- confidence level.

Function output

- estimated half-width of uniform distribution,
- error variance, estimated or given,
- used method for computing a confidence interval (asymptotic distribution of ML or likelihood ratio statistic).

LeArEst package - border estimation, object on the picture

- function *startweb.esttest()* starts a web application for border estimation
- demo

LeArEst package - area estimation, function *areaest()*

Function input

- vector of two-dimensional input data,
- number of *slices* for plain data cutting,
- error distribution (normal, Laplace, Student with 5 degrees of freedom),
- error variance or estimation method (Method of Moments, Maximum Likelihood Method),
- whether to plot input data, calculated edge points and the resulting ellipse.

Function output

- estimated area of the object,
- set of calculated object's edge points,
- resulting ellipse's semi-axes.

LeArEst package - area estimation, function *areaest()*

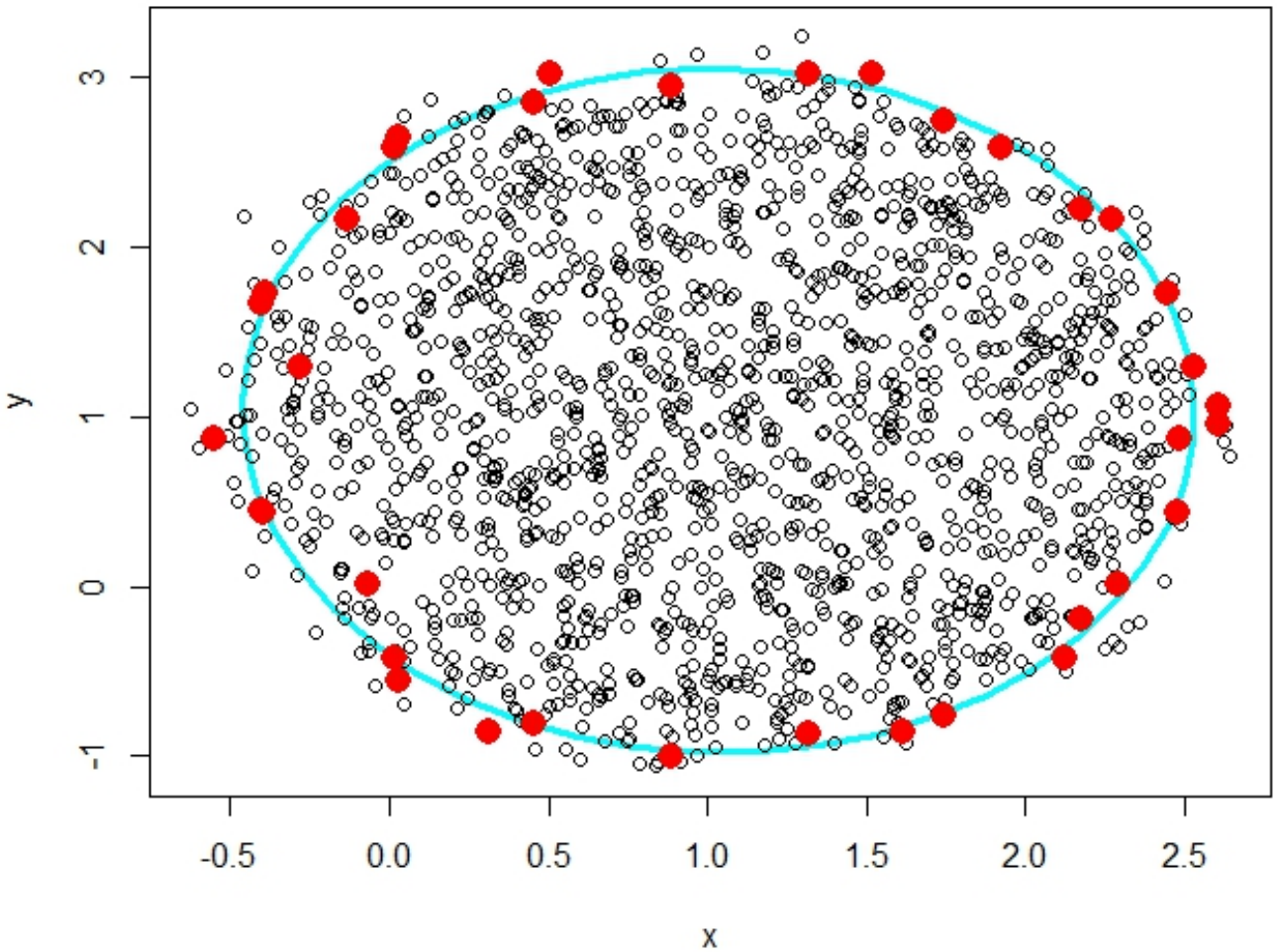
Function input

- vector of two-dimensional input data,
- number of *slices* for plain data cutting,
- error distribution (normal, Laplace, Student with 5 degrees of freedom),
- error variance or estimation method (Method of Moments, Maximum Likelihood Method),
- whether to plot input data, calculated edge points and the resulting ellipse.

Function output

- estimated area of the object,
- set of calculated object's edge points,
- resulting ellipse's semi-axes.

LeArEst package - area estimation, function *areaest()*



LeArEst package - area estimation, object on the picture

- function *startweb.area()* starts a web application for area estimation
- demo

Thank you for your attention!