Moderate and large deviations principles for the hazard rate function kernel estimator under censoring

Amadou DIALLO and Djamal LOUANI

L.S.T.A - Université Pierre et Marie Curie - Paris 6

Rencontres des jeunes statisticien-ne-s

Aussois, August 27^{th} 2013

Summary

Target

- $X_1, X_2, \ldots, X_n$ are nonnegative, i.i.d. random variables with a distribution function F and a density f.
- $Y_1, Y_2, \ldots, Y_n$ are nonnegative, i.i.d. variables with a distribution function G and a density g.
- The observed random variables are then Z_i and δ_i where

$$Z_i = X_i \wedge Y_i \quad \text{and} \quad \delta_i = \mathbb{1}_{\{X_i \leq Y_i\}}.$$

- Let H be the common distribution of the Z_i's.
- The Kernel estimator of $\lambda(x) = \frac{f(x)(1-G(x))}{1-H(x)}$ [Blum and Susarla (1980)] is

$$\lambda_n(x) = \frac{1}{nh_n} \sum_{i=1}^{n} \frac{\delta_i}{1 - H_n(x)} K\left(\frac{x - Z_i}{h_n}\right), \text{ where } H_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{Z_i \leq x\}}.$$

Target: *establish pointwise large and moderate deviation principles for the estimate $\lambda_n(x)$ of $\lambda(x)$.*

Target

- $X_1, X_2, \ldots, X_n$ are nonnegative, i.i.d. random variables with a distribution function F and a density f.
- $Y_1, Y_2, \ldots, Y_n$ are nonnegative, i.i.d. variables with a distribution function G and a density g.
- The observed random variables are then Z_i and δ_i where

$$Z_i = X_i \wedge Y_i \quad \text{and} \quad \delta_i = \mathbb{1}_{\{X_i \leq Y_i\}}.$$

- Let H be the common distribution of the Z_i's.
- The Kernel estimator of $\lambda(x) = \frac{f(x)(1-G(x))}{1-H(x)}$ [Blum and Susarla (1980)] is

$$\lambda_n(x) = \frac{1}{nh_n} \sum_{i=1}^{n} \frac{\delta_i}{1 - H_n(x)} K\left(\frac{x - Z_i}{h_n}\right), \text{ where } H_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{Z_i \leq x\}}.$$

Target: *establish pointwise large and moderate deviation principles for the estimate* $\lambda_n(x)$ *of* $\lambda(x)$.

Definition

$$P(\hat{S}_n \approx x) \asymp \exp\{-v_n I(x)\}.$$

$\{\mu_\epsilon\}$ satisfies a large deviation principle with a rate function I if, for all $A \in \mathcal{B}$,

$$-\inf_{x \in \overset{\circ}{A}} I(x) \leq \liminf_{\epsilon \to 0} \epsilon \log \mu_\epsilon(A) \leq \limsup_{\epsilon \to 0} \epsilon \log \mu_\epsilon(A) \leq -\inf_{x \in \bar{A}} I(x)$$

where $\overset{\circ}{A}$ denote the interior of A and $\bar{A}$ the closure of A.

The logarithmic moment generating function associated with the law μ is defined as $\Phi(\lambda) := \log E[\exp\langle\lambda, X_1\rangle]$.
The Fenchel-Legendre transform of $\Phi(\lambda)$ is

$$\Phi^*(x) := \sup_{\lambda\in\mathbb{R}^d} \{\langle\lambda, x\rangle - \Phi(\lambda)\}.$$

Theorem (Cramér(1938))

Assume that $D_\Phi = \mathbb{R}^d$, then
(a) For any closed set $F \subset \mathbb{R}^d$,

$$\limsup_{n\to\infty} \frac{1}{n} \log P(S_n \in F) \leq -\inf_{x\in F} \Phi^*(x).$$

(b) For any open set $G \subset \mathbb{R}^d$,

$$\liminf_{n\to\infty} \frac{1}{n} \log P(S_n \in G) \geq -\inf_{x\in G} \Phi^*(x).$$

Theorem (Chernoff(1952))

For any $y \in \mathbb{R}$

$$\lim_{n \to \infty} \frac{1}{n} \log P(S_n > y) = - \inf_{x \geq y} \Phi^*(x).$$

Bahadur (1971), Nikitin (1995), Louani $(1998, 2000)$

Consider $X_n \in \mathbb{R}^d$, where X_n possesses the law μ_n and the logarithmic moment generating function

$$\Psi_n(\lambda) := \log E[\exp\langle \lambda, X_n \rangle]$$

Assumptions:

- $\lim_{n\to\infty} \frac{1}{n}\Psi_n(\lambda) = \Phi(\lambda) \in [-\infty, \infty]$ exists.
- $0 \in \overset{\circ}{D}_\Phi$, with $D_\Phi = \{t \in \mathbb{R}^d : \Phi(\lambda) < \infty\}$.

Return The Fenchel-Legendre transform of $\Phi(\lambda)$ is

$$\Phi^*(x) := \sup_{\lambda\in\mathbb{R}^d} \{\langle \lambda, x \rangle - \Phi(\lambda)\},\ x \in \mathbb{R}^d.$$

Theorem (Gärtner-Ellis)

Let Assumption hold.
(a) For any closed set F,

$$\limsup_{n\to\infty} \frac{1}{n} \log P(S_n \in F) \leq - \inf_{x\in F} \Phi^*(x).$$

(b) For any open set G,

$$\liminf_{n\to\infty} \frac{1}{n} \log P(S_n \in G) \geq - \inf_{x\in G\cap\mathcal{F}} \Phi^*(x).$$

where $\mathcal{F}$ is the set exposed points of Φ^* whose exposing hyperplane belong to $\overset{\circ}{D}_\Phi$. See
(c) if Φ is an essentially smooth, lower semicontinuous function, then the LDP hold with the good rate function $\Phi^*(\cdot)$. See

LDP

We consider a vector process taking value in $\mathbb{R} \times \mathbb{R}_+$ and defined by

$$\mathbf{Z}_n(x) = (f_n^1(x), f_n^2(x)),$$

where

$$f_n^1(x) = \frac{1}{nh_n} \sum_{i=1}^{n} \left[\delta_i K\left(\frac{x - Z_i}{h_n}\right) - h_n \lambda(x)(1 - \mathbb{1}_{\{Z_i \leq x\}}) \right]$$

and

$$f_n^2(x) = 1 - H_n(x).$$

A large deviation principle is stated for the process $\mathbf{Z}_n(x)$ and the result on to the estimate $\lambda_n(x)$ is derived as a by-product by the contraction principle.

Hypothesis

- In order to give our results, we gather some notations and assumptions,

(H.1) As $n \to \infty$, $h_n \to 0$ and $nh_n \to \infty$;

(K.1) $K \geq 0$, $\int_{-\infty}^{\infty} K(z)dz = 1$;

(F.1) f is differentiable with bounded derivative $f^{'}$;

(G.1) G is differentiable with bounded derivative g;

Hypothesis

(I.1) $I(t) = \int_{-\infty}^{\infty} (\exp(tK(z)) - 1)dz < \infty, \ \ \forall\ t > 0.$ Set $J \equiv I^{'}$;

(I.2) $\int_{-\infty}^{\infty} |z|(\exp(tK(z)) - 1)dz < \infty, \ \ \forall\ t > 0;$

(I.3) $\int_{-\infty}^{\infty} K(z)\exp(tK(z))dz < \infty, \ \ \forall\ t > 0.$

LDP

Theorem (2.1)

Suppose that conditions **(H1)**, **(K1)**, **(F1)**, **(G1)** *and* **(I1)** - **(I3)** *are satisfied. Then the process* $\mathbf{Z}_n(x)$ *satisfies a large deviation principle in* $\mathbb{R} \times \mathbb{R}_+$ *with the rate* $1/nh_n$ *and the rate function* Υ_x *defined by*

$$\Upsilon_x(u,v) = \sup_{(t,s)\in\mathbb{R}^2} \{ut + vs - \Phi_x(t,s)\}$$

$$= \begin{cases} (u+\lambda(x)(1-H(x)))J^{-1}\left(\frac{u+\lambda(x)(1-H(x))}{f(x)(1-G(x))}\right) \\ \quad -f(x)(1-G(x))I\left(J^{-1}\left(\frac{u+\lambda(x)(1-H(x))}{f(x)(1-G(x))}\right)\right), \\ \qquad\qquad\qquad\qquad\qquad\qquad \textit{if } v \leq (1-H(x)) \\ \infty, \qquad\qquad\qquad\qquad\qquad\qquad \textit{if } v > (1-H(x)). \end{cases}$$

LDP

For any $(r, s) \in \mathbb{R}^2$, the limit normalized logarithm of the Laplace transform of the process $\mathbf{Z}_n(x)$ is given by

$$\Phi_x(t,s) = \lim_{n\to\infty} \frac{1}{nh_n} \log E[\exp\{nh_n \langle (t,s), \mathbf{Z}_n(x)\rangle\}]$$
$$= (s - t\lambda(x))(1 - H(x)) + f(x)(1 - G(x)) \int (\exp\{tK(z)\} - 1)\,dz.$$

We have now to establish that the function $\Phi_x(t,s)$ is essentially smooth. See

Corollaire (2.2)

Suppose that conditions of Theorem (2.1) *are satisfied. Then the process* $\lambda_n(x) - \lambda(x)$ *satisfies a large deviation principle in* $\mathbb{R}$ *with the rate* $1/nh_n$ *and the rate function* $F_x(\alpha)$ *given by*

$$F_x(\alpha) = \begin{cases} (\alpha+\lambda(x))(1-H(x))J^{-1}\left(\frac{(\alpha+\lambda(x))(1-H(x))}{f(x)(1-G(x))}\right) \\ \qquad -f(x)(1-G(x))I\left(J^{-1}\left(\frac{(\alpha+\lambda(x))(1-H(x))}{f(x)(1-G(x))}\right)\right), \\ \qquad\qquad \text{if } 0 < \alpha \leq \frac{f(x)(T_1-(1-F(x))}{(1-F(x))}, \\ \infty, \qquad\qquad \text{elsewhere.} \end{cases}$$

where $T_1 = \sup_{t>0}\{J(t)\}$. See

$$F_x(\alpha) := \inf\{\Upsilon_x(u,v) : u = \alpha v\} = \inf_v \Upsilon_x(\alpha v, v).$$

MDP

Set

$$D_n(x) = \frac{f_n^*(x) - f^*(x)}{1 - H(x)} - \frac{\lambda(x)}{1 - H(x)}[(1 - H_n(x)) - (1 - H(x))],$$

where

$$f_n^*(x) = \frac{1}{nh_n}\sum_{i=1}^{n} \delta_i K\left(\frac{x - Z_i}{h_n}\right) \quad \text{and} \quad f^*(x) = f(x)(1 - G(x)),$$

A moderate deviation principle is stated for the process $D_n(x)$ allowing to derive the result on to the random sequence $\lambda_n(x) - \lambda(x)$ by an exponential equivalence theorem.

Hypothesis

(H.2) Let (b_n) be a sequence of positive real numbers such that $\frac{nh_n}{b_n} \to \infty$, $\frac{nh_n^2}{b_n} \to 0$ and $\frac{nh_n}{b_n^2} \to 0$ as $n \to \infty$;

(K.2) $\int_{-\infty}^{\infty} K^2(z)dz < \infty$;

MDP

Theorem (2.3)

Assume that assumptions **(H1)-(H2)**, **(F1)**, **(G1)**, *and* **(K1) - (K2)** *hold true. Then, the sequence* $\frac{nh_n}{b_n} D_n(x)$ *satisfies a large deviation principle with speed* $\frac{nh_n}{b_n^2}$ *and a good rate function* γ_x, *where*

$$\gamma_x(\alpha) \equiv \sup_t \{\alpha t - \Phi_x(t)\} = \frac{\alpha^2(1 - H(x))^2}{2f(x)(1 - G(x)) \int K^2(z)dz}.$$

For any $t \in \mathbb{R}$, the limit normalized logarithm of the Laplace transform of the process $\mathbf{D}_n(x)$ is given by *

$$\Phi_x(t) = \lim_{n \to \infty} \frac{nh_n}{b_n^2} \log E\Big(\exp\{tb_n D_n(x)\}\Big)$$
$$= \frac{t^2 f(x)(1 - G(x))}{2(1 - H(x))^2} \int K^2(z)dz.$$

Since $\Phi_x(t)$ is a differentiable function.

Theorem (2.4)

Suppose that conditions of Theorem (2.3) are satisfied. Then the sequence $\frac{nh_n}{b_n}(\lambda_n(x) - \lambda(x))$ satisfies a large deviation principle with the speed $\frac{nh_n}{b_n^2}$ and the rate function γ_x. See

We have to show that $b_n(\lambda_n(x) - \lambda(x))$ and $b_n D_n$ are exponentially equivalent, i.e., for any $\alpha > 0$,

$$\limsup_{n\to\infty} \frac{nh_n}{b_n^2} \log P\left(\frac{nh_n}{b_n}|\lambda_n(x) - \lambda(x) - D_n(x)| > \alpha\right) = -\infty.$$

The moderate deviation result for the sequence $\lambda_n(x) - \lambda(x)$ is presented in the following theorem.

Lemma

Whenever $b_n \to \infty$ *and* $b_n/nh_n \to 0$, *as* $n \to \infty$, *for any* $0 < x < x_H$, *we have*

$$\lim_{n\to\infty} \frac{h_n}{b_n} \log P\left(\sqrt{\frac{nh_n}{b_n}}|H_n(x) - H(x)| > \beta\right) = -\frac{\beta^2}{2H(x)(1 - H(x))}.$$

Thank you for your attention!

J. Blum and V. Susarla.
Maximal deviation theory of some estimators of prior distribution functions.
Ann. Inst. Statist. Math., 33(3):425–436, 1981.

Amir Dembo and Ofer Zeitouni.
Large deviations techniques and applications, volume 38 of *Applications of Mathematics (New York)*.
Springer-Verlag, New York, second edition, 1998.

Amadou Oury Korbe Diallo and Djamal Louani.
Moderate and large deviation principles for the hazard rate function kernel estimator under censoring.
Statistics & Probability Letters, 83(3):735–743, 2013.

Abdelkader Mokkadem, Mariane Pelletier, and Julien Worms.
Large and moderate deviations principles for kernel estimation of a multivariate density and its partial derivatives.
Aust. N. Z. J. Stat., 47(4):489–502, 2005.

Definitions

- $y \in \mathbb{R}^d$, is an exposed point of Φ^* if for some $\lambda \in \mathbb{R}^d$ and all $x \neq y$,
$$\langle \lambda, y \rangle - \Phi^*(y) > \langle \lambda, x \rangle - \Phi^*(x)$$

Return

- A convex function $\Phi : \mathbb{R}^d \to (-\infty, \infty]$ is essentially smooth if:
  - $\overset{\circ}{D}_\Phi$ is non-empty.
  - $\Phi(\cdot)$ is differentiable throughout $\overset{\circ}{D}_\Phi$.
  - $\Phi(\cdot)$ is steep, namely, $\lim_{n \to \infty} |\nabla \Phi(\lambda_n)| = \infty$ whenever $\{\lambda_n\}$ is a sequence in $\overset{\circ}{D}_\Phi$ converging to a boundary point of $\overset{\circ}{D}_\Phi$.

Return

Theorem (*Contraction principle*)

Let $\mathcal{X}$ and $\mathcal{Y}$ be topological spaces. Let $\{\mu_n\}_{n=1}^{\infty}$ be a sequence of Borel probability measures on $\mathcal{X}$ and $g : \mathcal{X} \to \mathcal{Y}$ be a continuous function. If $\{\mu_n\}_{n=1}^{\infty}$ satisfies the LDP in $\mathcal{X}$ with good rate function I, then the sequence of induced measures $\{\nu_n = \mu_n g^{-1}\}_{n=1}^{\infty}$ satisfies the LDP in $\mathcal{Y}$ with good rate function

$$I^g(y) = \inf_{x \in g^{-1}(\{y\})} I(x).$$

Return

Definition

Two families of random variables, $\{X_n\}$ and $\{Y_n\}$ taking values in a common metric space, are exponentially equivalent when, for all $\delta \geq 0$,

$$\lim_{n\to\infty} \frac{1}{n} \log P(d(X_n, Y_n) > \delta) = -\infty.$$

Return