Moderate deviations for Grenander estimator near boundaries

Hui Jiang
Nanjing University of Areonautics and Astronautics
(a joint work with Prof. Fuqing Gao)

The 12th Workshop on Markov Processes and Related Topics
Jiangsu Normal University, Xuzhou, China

07.16 2016

Contents

1. Grenander estimator and asymptotic properties

Let f be a decreasing density with support $[0,1]$. Denote by

$$F_n(t) = \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \le t\}}$$

the empirical distribution function of a sample $X_1, ..., X_n$ from f.

Grenander estimator (NPMLE) $\hat{f}_n$ (Grenander, *Skand. Akt.*,1956): the left derivative of $\hat{F}_n$, where $\hat{F}_n$ is the concave majorant of F_n on $[0,1]$, i.e. the smallest concave function such that

$$\hat{F}_n(t) \ge F_n(t), \quad \hat{F}_n(0) = 0, \quad \hat{F}_n(1) = 1, \quad t \in [0,1].$$

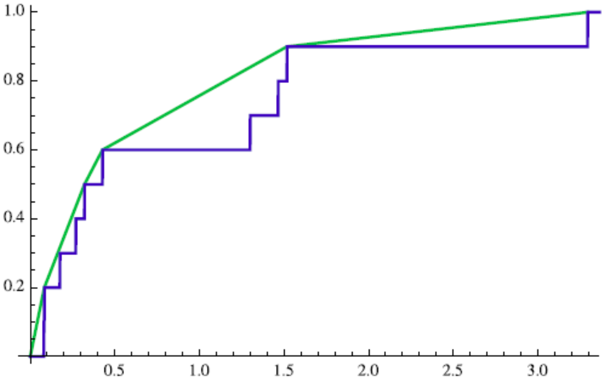

Fig. 1. Empirical distribution and least concave majorant, $n = 10$.

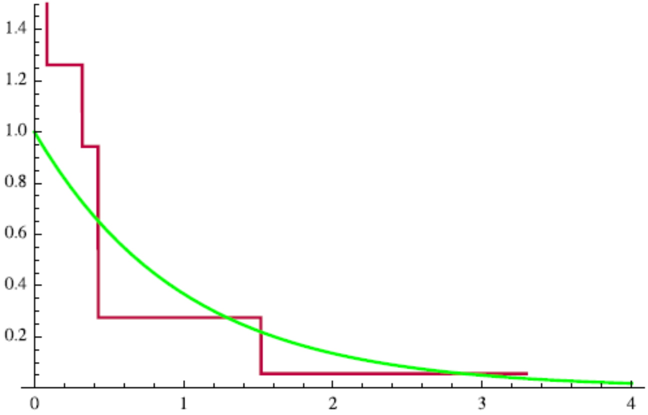

Fig. 2. Grenander estimator and Exp(1) density, $n = 10$.

Asymptotic pointwise properties

Suppose f is differentiable at $t \in (0,1)$ with $f(t) > 0, f'(t) < 0$.

- **Asymptotic distribution** (Prakasa Rao, *Sankhyā*,1969; Groeneboom, *Proc. Berkeley Confer.*,1985):

$$\left|4f(t)f'(t)\right|^{-1/3} n^{1/3}\left(\hat{f}_n(t) - f(t)\right) \xrightarrow{d} V(0),$$

where $V(c) = \operatorname{argmax}_{t\in\mathbb{R}}\left\{W(t) - (t-c)^2\right\}$ and W denotes a two-sided standard Wiener process originating from zero.

Remark (Groeneboom, *PTRF*,1989)

For the density f_V of the $V(0)$, as $|s| \to \infty$

$$f_V(s) \sim \frac{1}{2Ai'(a_1)} 4^{4/3}|s| \exp\left\{-\frac{2}{3}|s|^3 + 2^{1/3}a_1|s|\right\},$$

where $a_1 \approx -2.3381$ is the largest zero of the Airy function Ai and $Ai'(a_1) = 0.7002$.

Using **strong approximate theorem** (Komlós, Major and Tusnády, *Z.W.Verw.Geb.*,1975) and **small ball estimate** (Li and Shao, *Handbook Stat.*,2001), Gao, Zhao and Xiong established the moderate deviations for $\hat{f}_n(t)$:

- **MDP** (Gao, Zhao and Xiong, *preprint*,2014): for any $x > 0$

$$\lim_{n\to\infty} \frac{1}{b_n^3} \log P\left(\frac{n^{1/3}}{b_n} \left| \hat{f}_n(t) - f(t) \right| \geq x \right) = -\frac{4f(t)|f'(t)|}{6} x^3,$$

where b_n satisfies that as $n \to \infty$

$$b_n \to \infty, \quad \frac{b_n^7}{n^{1/3}} \to 0.$$

- **LIL** (Dümbgen, Wellner and Wolff, *SPA*,2016): If f is differentiable at $t \in (0,1)$ with $f(t) > 0, f'(t) < 0$, then

$$\limsup_{n\to\infty} \left(\frac{n}{2\log\log n}\right)^{1/3} \left(\hat{f}_n(t) - f(t)\right) = \left|3f(t)f'(t)\right|^{1/3}, \quad a.s.$$

Asymptotic global properties $(\|\hat{f}_n - f\|_{L^k}, \|\hat{f}_n - f\|_\infty)$

- L_1-**error** (Groeneboom, Hooghiemstra and Lopuhaä, *AOS*,1999; Durot, *PTRF*,2002): Let f is twice continuously differentiable, satisfying

$$0 < f(1) \leq f(0) < \infty, \quad \inf_{t \in (0,1)} |f'(t)| > 0.$$

Then, we have

$$n^{1/6}\left(n^{1/3}\int_0^1 \left|\hat{f}_n(t) - f(t)\right| dt - \mu\right) \xrightarrow{d} N(0, \sigma^2),$$

where $\mu = 2E|V(0)| \int_0^1 |f(t)f'(t)/2|^{1/3} dt$,

$$\sigma_{L_1}^2 = 8\int_0^\infty Cov(|V(0)|, |V(c) - c|)dc,$$

and $V(c) = \text{argmax}_{t \in \mathbb{R}} \left\{W(t) - (t - c)^2\right\}$.

Using the idea in Gao, Zhao and Xiong (*preprint*,2014), Bernstein type inequality for weakly dependent sequences (Merlevède, Peligrad and Rio, *PTRF*, 2011), we have

- **MDP of L_1-error** (Gao, Jiang, *preprint*, 2015): Let f be twice continuously differentiable and $\{\lambda_n\}$ be a sequence of positive numbers, satisfying

$$0 < f(1) \le f(0) < \infty, \quad \inf_{t \in (0,1)} |f'(t)| > 0, \quad \frac{n}{\lambda_n^{13} (\log n)^{16}} \to \infty$$

Then, for any $x > 0$

$$\lim_{n \to \infty} \frac{1}{\lambda_n} \log P\left(\lambda_n^{-1/2} n^{1/6} \left| n^{1/3} \int_0^1 |\hat{f}_n(t) - f(t)| dt - \mu \right| \ge x \right) = -\frac{x^2}{2\sigma_{L_1}^2}.$$

For the boundary points of the support of f, Woodroofe and Sun (*Stat.Sinica*,1993) showed that $\hat{f}_n$ is **not consistent** at 0 and 1.

This inconsistency at the boundaries will dominate the convergence, and then will have an great effect on the global measures of deviation, such as L_k-error, for $k \geq 2.5$, or the L_∞-error.

- L_k**-error** (Kulikov and Lopuhaä, *AOS*, 2005; Durot, *AOS*, 2007): Let f is twice continuously differentiable, satisfying

$$0 < f(1) \leq f(0) < \infty, \quad \inf_{t \in (0,1)} |f'(t)| > 0.$$

For $1 \leq k < 2.5$,

$$n^{1/6}\left(n^{1/3}\int_0^1 \left|\hat{f}_n(t) - f(t)\right|^k dt - \mu_k\right) \xrightarrow{d} N(0, \sigma_k^2),$$

and for $k \geq 2.5$, $1/6 < \varepsilon < (k-1)/(3k-6)$

$$n^{1/6}\left(n^{1/3}\int_{n^{-\varepsilon}}^{1-n^{-\varepsilon}} \left|\hat{f}_n(t) - f(t)\right|^k dt - \mu_k\right) \xrightarrow{d} N(0, \sigma_k^2),$$

where μ_k, σ_k can be formulated explicitly by f and

$$V(c) = \operatorname{argmax}_{t \in \mathbb{R}}\left\{W(t) - (t-c)^2\right\}, \ \xi(c) = V(c) - c.$$

- L_∞**-error** (Durot, Kulikov and Lopuhaä, *AOS*, 2012):
  Consider $0 \le u < v \le 1$ fixed. For real numbers $(\alpha_n)_n$ and $(\beta_n)_n$:

$$\alpha_n \to 0, \quad \beta_n \to 0, \quad 1 - v + \beta_n,\ u + \alpha_n \gg n^{-1/3}(\log n)^{-2/3},$$

we have that for any $x \in \mathbb{R}$, as $n \to \infty$,

$$P\left(\log n \left\{ \left(\frac{n}{\log n}\right)^{1/3} \sup_{t \in (u+\alpha_n, v-\beta_n]} \frac{|\widehat{f}_n(t) - f(t)|}{\left|2f(t)f'(t)\right|^{1/3}} - \nu_n \right\} \le x \right)$$
$$\to \exp\left\{-e^{-x}\right\} \ \textbf{(Gumbel distribution)},$$

where $C_{f,L} = 2\int_u^v \left(\frac{|f'(t)|^2}{f(t)}\right)^{1/3} dt$, and

$$\nu_n = 1 - \frac{\kappa}{2^{1/3}(\log n)^{2/3}} + \frac{1}{\log n}\left(\frac{1}{3}\log\log n + \log(\lambda C_{f,L})\right).$$

To make the properties of $\hat{f}_n$ near boundaries more clear, study the asymptotic convergence of

$$n^{\beta}\left(\hat{f}_n(cn^{-\alpha}) - f(cn^{-\alpha})\right), \ c, \beta > 0, 0 < \alpha < 1.$$

- **Near boundaries** (Kulikov and Lopuhaä, *AOS*,2006): Assume
  ◇ $0 < f(0) = \lim_{t \downarrow 0} f(t) < \infty$;
  ◇ there exists some positive constant ε_0 such that f has k-th order derivative in $(0, \varepsilon_0]$ and $f(\varepsilon_0) \neq 0$. Moreover $0 < |f^{(k)}(0)| < \infty$, with $f^{(k)}(0) = \lim_{t \downarrow 0} f^{(k)}(t)$ and $f^{(i)}(0) = 0$ for $1 \leq i \leq k-1$,
  ■ as $0 < \alpha < 1/(2k+1)$,

$$n^{\frac{1}{3}+\frac{\alpha(k-1)}{3}}\left(\hat{f}_n(cn^{-\alpha}) - f(cn^{-\alpha})\right)$$
$$\xrightarrow{d} 2\left(2(k-1)!\right)^{-1/3}\left(f(0)f^{(k)}(0)c^{k-1}\right)^{1/3} V(0),$$

  where $V(0) = \operatorname{argmax}_{t \in \mathbb{R}}\left\{W(t) - t^2\right\}$.

■ as $1/(2k+1) < \alpha < 1$,

$$n^{(1-\alpha)/2}\left(\hat{f}_n(cn^{-\alpha}) - f(cn^{-\alpha})\right)$$
$$\xrightarrow{d} (f(0)/c)^{1/2}\sqrt{\operatorname{argmax}_{t\in[0,\infty)}\{W(t)-t\}}.$$

- Note that for any $t \in (0,1)$,

$$\left|4f(t)f'(t)\right|^{-1/3} n^{1/3}\left(\hat{f}_n(t) - f(t)\right) \xrightarrow{d} V(0).$$

Question: As $1/(2k+1) < \alpha < 1$, consider the asymptotic convergence rate of $n^{(1-\alpha)/2}\left(\hat{f}_n(cn^{-\alpha}) - f(cn^{-\alpha})\right)$: for any $x > 0$,

$$P\left(n^{(1-\alpha)/2}\left(\hat{f}_n(cn^{-\alpha}) - f(cn^{-\alpha})\right) \geq \lambda_n x\right)$$

and $P\left(n^{(1-\alpha)/2}\left(\hat{f}_n(cn^{-\alpha}) - f(cn^{-\alpha})\right) \leq b_n^{-1} x\right)$, where $\lambda_n, b_n \to \infty$.

2. Inverse process

- $U_n(a)$ is defined as the last time that the process $F(t) - at$ attains its maximum:

$$U_n(a) = \underset{t\in[0,1]}{\operatorname{argmax}} \{F_n(t) - at\}.$$

Then, with probability 1, (Groeneboom (*Proc. Berkeley Confer.*,1985))

$$\hat{f}_n(t) \leq a \Leftrightarrow U_n(a) \leq t.$$

Therefore, $\{U_n(a) : a \in [f(1), f(0)]\}$ is called the inverse process of $\left\{\hat{f}_n(t) : t \in [0,1]\right\}$, which has become a cornerstone in this field.

- For any $x > 0$

$$
\begin{aligned}
& P\left(n^{(1-\alpha)/2}\left(\hat{f}_n(cn^{-\alpha}) - f(cn^{-\alpha})\right) \geq \lambda_n x\right) \\
& = P\left(n^{\alpha} U_n\left(f(cn^{-\alpha}) + \lambda_n n^{(\alpha-1)/2} x\right) \geq c\right) \\
& = P\left(\tau_{n,\lambda_n}^{x} \geq c\right),
\end{aligned}
$$

where $\tau_{n,\lambda_n}^{x} = \operatorname{argmax}_{t\in[0,\infty)} Z_{n,\lambda_n}^{x}(t)$ and

$$
\begin{aligned}
Z_{n,\lambda_n}^{x}(t) = {} & n^{(1+\alpha)/2}(F_n - F)(n^{-\alpha} t) \\
& + n^{(1-\alpha)/2}\left(n^{\alpha} F(n^{-\alpha} t) - f(cn^{-\alpha}) t\right) - \lambda_n x t.
\end{aligned}
$$

- Similarly,

$$P\left(n^{(1-\alpha)/2}\left(\hat{f}_n(cn^{-\alpha}) - f(cn^{-\alpha})\right) \leq b_n^{-1}x\right) = P\left(\varsigma_{n,b_n}^{x} \leq c\right),$$

where $\varsigma_{n,b_n}^{x} = \mathrm{argmax}_{t\in[0,\infty)} T_{n,b_n}^{x}(t)$, and

$$\begin{aligned} T_{n,b_n}^{x}(t) = {} & n^{(1+\alpha)/2}(F_n - F)(n^{-\alpha}t) \\ & + n^{(1-\alpha)/2}\left(n^{\alpha}F(n^{-\alpha}t) - f(cn^{-\alpha})t\right) - b_n^{-1}xt. \end{aligned}$$

3. Moderate deviations for Grenander estimator near boundaries

Let $\{\lambda_n\}, \{b_n\}$ satisfy that

$$\lambda_n \to \infty, \quad \frac{n^{(1-\alpha)}}{\lambda_n^6} \to \infty, \quad \frac{n^{(2k+1)\alpha-1}}{\lambda_n^2} \to \infty$$

and $b_n \to \infty, \quad \frac{n^{(2k+1)\alpha-1}}{b_n^{4(k+1)}(\log b_n)^{2k+3}} \to \infty.$

Theorem

For $1/(2k+1) < \alpha < 1, c > 0$, assume condition (C1), (C2) hold. Then for any $x > 0$,

- $\lim_{n\to\infty} \frac{1}{\lambda_n^2} \log P\left(\frac{n^{(1-\alpha)/2}}{\lambda_n}\left(\hat{f}_n(cn^{-\alpha}) - f(cn^{-\alpha})\right) \geq x\right) = -\frac{cx^2}{2f(0)}$;
- $\lim_{n\to\infty} b_n P\left(b_n n^{(1-\alpha)/2}\left(\hat{f}_n(cn^{-\alpha}) - f(cn^{-\alpha})\right) \leq x\right) = \sqrt{\frac{2c}{e\pi f(0)}}x.$

Key ingredient

- For any $x > 0$

$$P\left(\frac{n^{(1-\alpha)/2}}{\lambda_n}\left(\hat{f}_n(cn^{-\alpha}) - f(cn^{-\alpha})\right) \geq x\right) = P\left(\tau^x_{n,\lambda_n} \geq c\right),$$

where $\tau^x_{n,\lambda_n} = \operatorname{argmax}_{t\in[0,\infty)} Z^x_{n,\lambda_n}(t)$ and

$$Z^x_{n,\lambda_n}(t) = \underbrace{n^{(1+\alpha)/2}(F_n - F)(n^{-\alpha}t)}_{n^{\alpha/2}W_n(F(n^{-\alpha}t)) - n^{\alpha/2}F(n^{-\alpha}t)W_n(1) \text{ KMT strong approximate}}$$

$$+ \underbrace{n^{(1-\alpha)/2}\left(n^{\alpha}F(n^{-\alpha}t) - f(cn^{-\alpha})t\right)}_{0?} - \lambda_n x t.$$

- For any $x > 0$, we have

$$\limsup_{N\to\infty}\limsup_{n\to\infty}\frac{1}{\lambda_n^2}\log P\left(\tau_{n,\lambda_n}^x \geq N\right) = -\infty.$$

- Define $\tilde{\tau}_{n,\lambda_n}^x = \text{argmax}_{t\in[0,\infty)}\tilde{Z}_{n,\lambda_n}^x(t)$, where

$$\tilde{Z}_{n,\lambda_n}^x(t) = n^{\alpha/2}W_n\left(F(n^{-\alpha}t)\right) - \lambda_n xt.$$

Then, we have
(1). $\limsup_{N\to\infty}\limsup_{n\to\infty}\frac{1}{\lambda_n^2}\log P\left(\tilde{\tau}_{n,\lambda_n}^x \geq N\right) = -\infty.$
(2). τ_{n,λ_n}^x is exponential equivalent to $\tilde{\tau}_{n,\lambda_n}^x$: for any $\varepsilon > 0$

$$\limsup_{n\to\infty}\frac{1}{\lambda_n^2}\log P\left(\left|\tau_{n,\lambda_n}^x - \tilde{\tau}_{n,\lambda_n}^x\right| \geq \varepsilon\right) = -\infty.$$

- By some calculations, we also have $\tilde{\tau}^{x}_{n,\lambda_n}$ is exponential equivalent to $\frac{f(0)}{\lambda_n^2 x^2}\tau$, where

$$\tau = \operatorname{argmax}_{u\in[0,\infty)} \left\{ W(u) - u \right\}.$$

Then

$$P\left(\frac{n^{(1-\alpha)/2}}{\lambda_n}\left(\hat{f}_n(cn^{-\alpha}) - f(cn^{-\alpha})\right) \geq x\right) \sim P\left(\tau \geq \frac{cx^2}{f(0)}\lambda_n^2\right).$$

- Similarly, we can obtain

$$P\left(b_n n^{(1-\alpha)/2}\left(\hat{f}_n(cn^{-\alpha}) - f(cn^{-\alpha})\right) \le x\right) = P\left(\varsigma^x_{n,b_n} \le c\right),$$

where $\varsigma^x_{n,b_n} = \text{argmax}_{t\in[0,\infty)} T^x_{n,b_n}(t)$, and

$$\begin{aligned} T^x_{n,b_n}(t) = & \ n^{(1+\alpha)/2}(F_n - F)(n^{-\alpha}t) \\ & + n^{(1-\alpha)/2}\left(n^{\alpha}F(n^{-\alpha}t) - f(cn^{-\alpha})t\right) - b_n^{-1}xt. \end{aligned}$$

Then, (1). $\limsup_{N\to\infty}\limsup_{n\to\infty} b_n P\left(\varsigma^x_{n,b_n} \ge b_n^2 N \log b_n\right) = 0.$

(2). ς^x_{n,b_n} is equivalent to $\frac{f(0)b_n^2}{x^2}\tau$: for any $\varepsilon > 0$

$$\limsup_{n\to\infty} b_n P\left(\left|\varsigma^x_{n,b_n} - \frac{f(0)b_n^2}{x^2}\tau\right| \ge \varepsilon\right) = 0.$$

Therefore

$$P\left(b_n n^{(1-\alpha)/2}\left(\hat{f}_n(cn^{-\alpha}) - f(cn^{-\alpha})\right) \leq x\right) \sim P\left(\tau \leq \frac{cx^2}{f(0)b_n^2}\right).$$

Lemma (Bhattacharya and Brockwell, *Z.W.Verw.Geb.*,1976)

Let $M = \sup_{t\in[0,\infty)}\{W(t) - t\}$. Then (τ, M) has the joint density

$$f(u,v) = 2vu^{-3/2}\phi(vu^{-1/2} + u^{1/2}),$$

where $\phi(\cdot)$ is the standard normal density. In particular, the density of τ satisfies

$$f_\tau(t) \sim Ct^{-1/2}e^{-t/2}, \quad t \to \infty$$

and $f_\tau(t) \sim \frac{1}{\sqrt{2e\pi}}t^{-1/2}, \quad t \to 0$, where C is some positive constant.

Hence, for any $x > 0$,

$$P\left(\frac{n^{(1-\alpha)/2}}{\lambda_n}\left(\hat{f}_n(cn^{-\alpha}) - f(cn^{-\alpha})\right) \geq x\right) \sim P\left(\tau \geq \frac{cx^2}{f(0)}\lambda_n^2\right)$$
$$\sim \exp\left(-\frac{cx^2}{2f(0)}\lambda_n^2\right).$$

$$P\left(b_n n^{(1-\alpha)/2}\left(\hat{f}_n(cn^{-\alpha}) - f(cn^{-\alpha})\right) \leq x\right) \sim P\left(\tau \leq \frac{cx^2}{f(0)b_n^2}\right)$$
$$\sim \sqrt{\frac{2c}{e\pi f(0)}}x.$$

Thank you for your attention!