Shahid Bahonar University of Kerman
Journal of Mahani Mathematical Research
Print ISSN: 2251-7952 Online ISSN: 2645-4505

THE EXTENDED GLIVENKO-CANTELLI PROPERTY FOR KERNEL-SMOOTHED ESTIMATOR OF THE CUMULATIVE DISTRIBUTION FUNCTION IN THE LENGTH-BIASED SAMPLING

M. Ajami ✉, R. Zamini, and SMA. Jahanshahi

Article type: Research Article
(Received: 03 December 2023, Received in revised form 22 May 2024)
(Accepted: 12 July 2024, Published Online: 13 July 2024)

Abstract. When the probability of selecting an individual from a population is proportional to its length, the resulting distribution of observation will exhibit length bias. This distribution is referred to as a length-biased distribution. Let $\{Y_i; i = 1, \dots, n\}$ be a sample from a length-biased population with cumulative distribution function $G(\cdot)$. In this paper we consider Cox's empirical estimator $F_n^c(\cdot)$ and the smoothed kernel-type estimator $F_n^s(\cdot)$ of $F(\cdot)$. Under suitable conditions, the extended Glivenko-Cantelli theorem for $F_n^c(\cdot)$ and $F_n^s(\cdot)$ are proved. Also, the validity of the extended Glivenko-Cantelli property for the smoother estimator $F_n^s(\cdot)$ is investigated using a simulation study.

Keywords: Law of iterated logarithm, Length-biased data, Smoothed estimator, Strong consistency.
2020 MSC: 62G05-62G20.

1. Introduction

For a given i.i.d. sample $X_1, \dots, X_n$ from an unknown continuous cumulative distribution function (CDF) $F(\cdot)$,the empirical distribution function is defined as $F_n(t) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \le t)$ where $I(\cdot)$ is the indicator function. Since X_i's are i.i.d. from Strong Law of Large Numbers one can see that

$$F_n(t) = \frac{1}{n}\sum_{i=1}^{n} I(X_i \le t) \to E[I(X \le t)] = F(t), \qquad \text{as } n \to \infty, \quad w.p.1,$$

and since $\mathrm{Var}[I(X \le t)] = F(t)(1-F(t))$ from Central Limit Theorem we have

$$F_n(t) \sim AN\left(F(t), n^{-1}F(t)(1-F(t))\right), \qquad \text{as } n \to \infty.$$

Also, the classical Glivenko-Cantelli theorem says that F_n converges almost surely (a.s.) to $F(t)$ uniformly in $t \in R$, i.e.,

$$\sup_{t\in R} |F_n(t) - F(t)| \to 0, \text{ a.s.}$$

✉ m.ajami@vru.ac.ir, ORCID: 0000-0002-6045-128x
https://doi.org/10.22103/jmmr.2024.22620.1546

Publisher: Shahid Bahonar University of Kerman
How to cite: M. Ajami, R. Zamini and SMA. Jahanshahi, *The extended Glivenko-Cantelli property for Kernel-Smoothed estimator of the cumulative distribution function in the length-biased sampling*, J. Mahani Math. Res. 2024; 13(2): 535 - 545.

As well as the extended Glivenko-Cantelli lemma (in Fabian and Hannan [5], pp 80-83) states that $\sup n^{\alpha} |F_n(t) - F(t)| \to 0$ a.s., for any $0 < \alpha < 1/2$. Although $F_n(t)$ is a consistent estimator for $F(t)$ and it also has the good properties mentioned above, but the empirical distribution is not smooth as it jumps by $1/n$ at each sample realization point. As an alternative Nadaraya [11] introduced Kernel-type estimators for distribution estimation, based on symmetric kernels as follows:

$$\widehat{F}_n(t) = \int_{-\infty}^{t} f_n(x)\, dx$$

where

$$f_n(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h_n}\right),$$

and $K(x)$ is some density function such that $K(x) < c < \infty, \lim_{x\to\pm\infty} |xK(x)| = 0$, and $h_n \to 0$ with increasing n. On the other hand, the other form of the above formula can also be written as follows

$$\widehat{F}_n(t) = \int_{-\infty}^{t} \hat{f}(x)dx = \frac{1}{n} \sum_{i=1}^{n} L\left(\frac{t - X_i}{h_n}\right),$$

where $L(t) = \int_{-\infty}^{t} K(x)dx$ is a CDF.

A lot of research has been done on the smooth estimator of the cumulative distribution function, among which the following can be considered. Yamato [17] proposed the new smoothed kernel distribution estimator and provided mild necessary and sufficient conditions for the consistency of it in uniform norm. Also, under certain regularity conditions the Chung-Smirnov property has been obtained for it by Winter [16] for the upper bound and Degenhardt [3] for the lower bound, i.e.

$$\limsup_{n\to\infty} \sup_{0<t<\infty} \{2n/\log\log n\}^{1/2} \left|\widehat{F}_n(t) - F(t)\right| = 1 \quad \text{a.s.}$$

Furthermore, it has been shown by Reiss [12] and Falk [6] that asymptotic performance of Yamato's estimator is better than empirical function $\widehat{F}_n(\cdot)$. Also Fernholz [7] gave a rate of order $o(n^{-1/2})$ for strong uniform convergence under Lipschitz condition on $F(\cdot)$ and some regular conditions on kernel function. Also Yukich [18] found sufficient conditions for this convergence with rate $o_p(n^{-1/2})$ in a more general set. Almost sure limit behavior for the maximal deviation between kernel c.d.f. estimator and the true underlying c.d.f. was investigated by Degenhardt [4] under various smoothness conditions on $F(\cdot)$ and the class of kernels. In the continuation of the discussion, because in the next section we need to use the law of iterated logarithm to prove part of theorem 2.1, we will recall the famous classic definition of the law of iterated logarithm.

Definition 1.1. Let $X_1, \ldots, X_n$ from $F(\cdot)$ be independent, identically distributed random variables with means μ and variances σ^2. Let $S_n = X_1 + \ldots + X_n$. Then

$$\limsup_{n\to\infty} \frac{|S_n - n\mu|}{\sqrt{2\sigma^2 n \log\log n}} = 1 \quad \text{a.s.}$$

In this article, based on a special type of sampling called length-biased sampling, we will investigate some properties of CDF estimators. The issue of length-biased data was proposed for the first time by Wicksell [15] in the context of anatomy which he named corpuscle problem. While observing the corpuscles, he found out that only those corpuscles are observable that their size is larger than a certain magnitude and the smaller ones are not visible by microscope. Then, this phenomenon was investigated by Mcfadden [10] and Blumenthal [1] with statistical concept. Cox [2] found out that in an industrial sampling, the longer fibers, the bigger the chance of being chosen. This issue imposed a bias on the results which is later called length-bias. In general, length-biased data sets arise when the sampling mechanism is such that the larger the potential observation, the higher the probability of being included in the final sample.

Now as mentioned before let X be a non-negative continuous random variable with its CDF $F(\cdot)$ and density function $f(\cdot)$. As above mentioned, if the probability of a selected item for the sample is proportional to its length (X), the distribution of the observed length is length-biased. In the case of length-biased data, we consider random variable Y has the length-biased distribution with CDF $G(\cdot)$ and density function $g(\cdot)$. One can see that the following relationship between $F(\cdot)$ and its length-biased distribution $G(\cdot)$ will be held,

$$G(t) = \frac{1}{\mu}\int_0^t x\, dF(x), \quad t \geq 0,$$

where $\mu = E(X)$ is assumed finite. Throughout this paper we assume that $G(\cdot)$ is continuous on

$$\mathbb{R}^+ = [0, \infty).$$

From this it can be concluded that $F(\cdot)$ is also continuous. An elementary calculation shows that $F(\cdot)$ is determined uniquely by $G(\cdot)$, namely

$$F(t) = \mu \int_0^t y^{-1} dG(y), \quad t \geq 0. \tag{1}$$

Cox [2] proposed an estimator of the common underlying CDF $F(\cdot)$ based on a sample $Y_1, \ldots, Y_n$ from $G(\cdot)$. Let $\nu_n = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{Y_i}$, then

$$\begin{aligned} F_n^c(t) &= \nu_n^{-1}\int_0^t y^{-1} dG_n(y) \\ &= \frac{\nu_n^{-1}}{n}\sum_{i=1}^{n}\frac{1}{Y_i} I(Y_i \le t), \end{aligned} \tag{2}$$

where $G_n(\cdot)$ is the empirical estimator of $G(\cdot)$. Vardi [14] showed that the Cox's estimator is a non-parametric maximum likelihood estimator for the target distribution $F(\cdot)$. Horváth [8] established strong uniform consistency without rate of $F_n^c(\cdot)$ and showed that under the condition $E(Y^{-1}) < \infty$,

$$\lim_{n\to\infty}\sup_{0<t<\infty}|F_n^c(t) - F(t)| = 0 \quad a.s. \tag{3}$$

The empirical estimator $F_n^c(\cdot)$ is step function and discontinuous. Since, smoothed estimators have a better performance compared to non-smoothed estimators, we define the kernel smoothed version of $F_n^s(\cdot)$ through the convolution of two functions $F_n^c(\cdot)$ and CDF $L(\cdot)$, i.e.

$$\begin{aligned} F_n^s(t) &= \int_0^\infty L\left(\frac{t-y}{h_n}\right) dF_n^c(y) \\ &= \frac{\nu_n^{-1}}{n}\sum_{i=1}^{n}\frac{1}{Y_i} L\left(\frac{t-Y_i}{h_n}\right), \quad t > 0, \end{aligned} \tag{4}$$

where h_n is an arbitrary sequence of smoothing parameters (or bandwidths) that fulfills $\lim_{n\to\infty} h_n = 0$ and $L(\cdot)$ is a cumulative form of a kernel density function $K(\cdot)$, i.e.

$$L(t) = \int_{-\infty}^{t} K(x)\, dx.$$

Jahanshahi et al. [9] investigated uniform consistency and asymptotic normality of (4) and made a one-sample Kolmogorov type of goodness of fit test with this estimator for length-biased data. Also Zamini et al. [19] established a Berry-Esseen type bound for the smoothed estimator in this setting

The purpose of this paper is to obtain a more general result of (3), as well as the extended Glivenko-Cantelli theorem for the smoothed estimator $F_n^s(\cdot)$. In the following, we prove a limit theorem consisting of strong consistency for $F_n^s(\cdot)$. To this end, some assumptions are required which are presented below:

Assumptions.

A1. The kernel function $K(\cdot)$ is symmetric, of bounded variation on $(-1, 1)$. In addition $K(t) = 0$ if $t \notin (-1, 1)$ and satisfies the following conditions: $\int_{-1}^{1} K(t)dt = 1$, $\int_{-1}^{1} tK(t)dt = 0$, $\int_{-1}^{1} t^2K(t)dt = m < \infty$.

A2. Suppose that there exists a sequence of positive real number γ_n such that $\sum_{n=1}^{\infty} G(\gamma_n) < \infty$.
A3. Suppose that $|f(s)| \leq M$ for s in a neighborhood of t, where M is a constant depending only on t.
A4. Let $\gamma_n = o(n^{-\beta})$ for any $\beta > \alpha \geq 0$, and $\beta + \alpha < \frac{1}{2}$.
A5. Let Y is a non-negative random variable with its CDF $G(\cdot)$ then we suppose that $E(Y^{-1}) < \infty$ and $E(Y^{-2}) < \infty$.

Discussion on the Assumptions. Assumption **A1** is widely used in the literature of kernel density estimators and is assumed for extended Glivenko-Cantelli of $F_n^s(\cdot)$. Assumption **A2** is required to guarantee that for each n enough large, we have $G(y) = 0$ for $y < \gamma_n$. Assumption **A3** is used in Theorems 2.1 and 2.3 to obtain the rate of smoothed and non smoothed estimators that were mentioned before. We note that Assumption **A4** in Theorems 2.1 and 2.3 is needed to find the certain rate of strong consistency of $F_n^c(\cdot)$ and $F_n^s(\cdot)$. Assumption **A5** is used in part I_1 in Theorem 2.1.

2. Main results

In the following theorems, we give the extended Glivenko-Cantelli theorem for the Cox's estimator $F_n^c(\cdot)$ and the smoother estimator $F_n^s(\cdot)$. It is worth noting that all the main theorems and their proofs have been presented and proven by the authors in this section for the first time.

Theorem 2.1. *Let* $0 \leq \alpha < 1/4$. *Suppose that* $0<\mu< \infty$ *and Assumptions* **A2**-**A5** *are satisfied. Then, we have*

$$\lim_{n\to\infty} \sup_{0\leq t<\infty} n^{\alpha}|F_n^c(t) - F(t)| = 0, \quad a.s. \tag{5}$$

Proof. It can clearly be seen that

$$\begin{aligned}
\sup_{0\leq t<\infty} n^{\alpha}|F_n^c(t) - F(t)| \leq & n^{\alpha}|\nu_n^{-1} - \mu| \int_0^{\infty} y^{-1} dG_n(y) \\
& + \mu n^{\alpha} \sup_{0\leq t<\infty} \Big| \int_0^{t} y^{-1} d\left(G_n(y) - G(y)\right) \Big| \\
\leq & n^{\alpha}|\nu_n^{-1} - \mu| \int_0^{\infty} y^{-1} dG_n(y) + \mu n^{\alpha} \int_0^{\gamma_n} y^{-1} dG_n(y) \\
& + \mu n^{\alpha} \int_0^{\gamma_n} y^{-1} dG(y) + 2\mu\gamma_n^{-1} n^{\alpha} \sup_{\gamma_n\leq t<\infty} |G_n(t) - G(t)| \\
=: & I_1 + I_2 + I_3 + I_4.
\end{aligned} \tag{6}$$

But according to (1), we have

$$E(\nu_n) = E(\frac{1}{n}\sum_{1}^{n}\frac{1}{Y_i}) = E(Y_1^{-1})$$

$$= \int_0^\infty y^{-1}g(y)dy = \frac{1}{\mu}\int_0^\infty y^{-1}yf(y)dy = \frac{1}{\mu}. \tag{7}$$

Also with a little calculation it can be shown that

$$Var(\nu_n) = \frac{\gamma^2}{n},$$

where $\gamma^2 = \int_0^\infty y^{-2}dG(y) - \mu^{-2}$ and according to Assumption **A5**, it is obvious that $\gamma^2 < \infty$. On the other hand one can see that

$$|\nu_n^{-1} - \mu| = \left|\frac{1}{\mu^{-1}} - \frac{1}{\nu_n}\right| = \frac{|\nu_n - \mu^{-1}|}{\mu^{-1}\nu_n}$$

$$= \frac{|\sum_{i=1}^n \left(Y_i^{-1} - \mu^{-1}\right)|}{n\mu^{-1}\nu_n}$$

$$= \frac{\sqrt{2n\gamma^2\log\log n}}{n\mu^{-1}\nu_n}\frac{|\sum_{i=1}^n \left(Y_i^{-1} - \mu^{-1}\right)|}{\sqrt{2n\gamma^2\log\log n}}.$$

By the law of the iterated logarithm for independent identically distributed random variables as mentioned in the definition 1.1,

$$\limsup_{n\to\infty}\frac{|\sum_{i=1}^n \left(Y_i^{-1} - \mu^{-1}\right)|}{\sqrt{2n\gamma^2\log\log n}} = 1 \quad \text{a.s.}$$

Furthermore, based on (7) and the strong law of large numbers, it follows that $\nu_n \longrightarrow \mu^{-1}$ a.s, so

$$\limsup_{n\to\infty}\sqrt{\frac{n}{\log\log n}}|\nu_n^{-1} - \mu| = \sqrt{2}\gamma\mu^2 \quad \text{a.s.}$$

Now, based on $0<\mu<\infty$, we have

$$|\nu_n^{-1} - \mu| = O\left(\sqrt{\frac{\log\log n}{n}}\right) \quad a.s.$$

But this is a reasonable condition to assume $\int_0^\infty y^{-1}dG_n = \frac{1}{n}\sum_{i=1}^n \frac{1}{Y_i^{-1}} < \infty$, then one can see that

$$I_1 = O\left(n^{\alpha-1/2}\sqrt{\log\log n}\right), \quad a.s.,$$

and therefore, since $\alpha<1/4$, we obtain that

$$I_1 = n^\alpha|\nu_n^{-1} - \mu|\int_0^\infty y^{-1}dG_n(y) \longrightarrow 0, \quad a.s. \tag{8}$$

To deal with I_2, if the Assumption **A2** is satisfied, it is an immediate consequence of Theorem 1 of Section 10.1 of Shorack and Wellner [13] that $P(min(Y_1, \ldots, Y_n) \leq \gamma_n,\ i.o.)$=0, so

$$\lim_{n\to\infty} n^{\alpha} \int_0^{\gamma_n} \frac{1}{y} dG_n(y) = 0, \quad a.s,$$

it follows that

$$I_2 = \mu n^{\alpha} \int_0^{\gamma_n} \frac{1}{y} dG_n(y) \longrightarrow 0, \quad a.s. \tag{9}$$

On the other hand according to Assumption **A3**, we have

$$I_3 = n^{\alpha} \mu \int_0^{\gamma_n} y^{-1} dG(y) = n^{\alpha} \int_0^{\gamma_n} f(y) dy \leq M n^{\alpha} \gamma_n.$$

Now, according to Assumption **A4**, we have $\gamma_n.n^{\beta} \longrightarrow 0$ for any $\beta > \alpha \geq 0$, that implies $\gamma_n.n^{\alpha} \longrightarrow 0$. So

$$I_3 = n^{\alpha} \mu \int_0^{\gamma_n} y^{-1} dG(y) \longrightarrow 0, \quad a.s. \tag{10}$$

For I_4, we use the extended Glivenko-Cantelli lemma mentioned in Fabian and Hannan [5] [pp. 80-83], thus

$$I_4 = 2\mu \gamma_n^{-1} n^{\alpha} \sup_{\gamma_n \leq t < \infty} |G_n(t) - G(t)| \to 0, \quad a.s. \tag{11}$$

So, by substituting (8)-(11) into (6), the proof is completed.

□

Remark 2.2. Note that considering $\alpha = 0$, (5) reduces to Theorem 1 in Horváth [8].

Theorem 2.3. *Suppose that Assumption* **A1** *is satisfied. Then under the assumptions of Theorem 2.1 and* $\lim_{n\to\infty} n^{\alpha} h_n = 0$ *for* $0 \leq \alpha < 1/4$*, we have that*

$$\lim_{n\to 0} \sup_{0 \leq t < \infty} n^{\alpha} |F_n^s(t) - F_n^c(t)| = 0, \quad a.s. \tag{12}$$

Proof. Using change of variables and integration by parts, one can see that, for any $t > 0$, $u \in [-1, 1]$ and large n,

$$\begin{aligned} |F_n^s(t) - F_n^c(t)| \leq & \int_{-1}^{1} |F_n^c(t - h_n u) - F_n^c(t)| dL(u) \\ \leq & \int_{-1}^{1} |F_n^c(t - h_n u) - F(t - h_n u)| dL(u) + \int_{-1}^{1} |F_n^c(t) - F(t)|\, dL(u) \\ & + \int_{-1}^{1} |F(t - h_n u) - F(t)| dL(u) \\ \leq & \ 2 \sup_{0 \leqslant t < \infty} |F_n^c(t) - F(t)| + \int_{-1}^{1} |F(t - h_n u) - F(t)| dL(u). \qquad (13) \end{aligned}$$

By the Mean Value Theorem and Assumption **A3**, we can obtain

$$|F(t-h_nu)-F(t)|K(u)du \leq Mh_n.$$

So according to $\int_{-1}^{1} K(t)dt = 1$, one can see that

$$\sup_{0\leqslant t<\infty} \int_{-1}^{1} |F(t-h_nu)-F(t)|\, K(u)du = O\left(h_n\right).$$

Now, according to condition $\lim_{n\to 0} n^{\alpha} h_n = 0$, it is obvious that

$$\lim_{n\to 0} \sup_{0\leqslant t<\infty} n^{\alpha} \int_{-1}^{1} |F(t-h_nu)-F(t)|\, K(u)du = 0. \tag{14}$$

On the other hand, according to Theorem 2.1, we have

$$\lim_{n\to\infty} 2 \sup_{0\leq t<\infty} n^{\alpha} |F_n^c(t)-F(t)| = 0, \quad a.s. \tag{15}$$

Hence, substituting (15) and (14) in (13), we obtain

$$\lim_{n\to 0} \sup_{0\leqslant t<\infty} n^{\alpha} |F_n^s(t)-F_n^c(t)| = 0.$$

Now, the proof is completed. □

Remark 2.4. In the following, as an example for Theorem 2.3, we will consider the correctness of (12) based on simulation studies. For this purpose, we assume that the unbiased data have a *gamma*$(1,2)$ distribution. Therefore, it is obvious that according to (1), one can see that the random sample $Y_1,\ldots,Y_n$ must be drawn from the *gamma*$(2,2)$ distribution. Also, $F_n^c(\cdot)$ and $F_n^s(\cdot)$ are obtained based on Formulas (2) and (4) respectively. The kernel used in $F_n^s(\cdot)$ is the Epanechnikov kernel, which also fulfills Assumption **A1**. We have considered $\alpha = 0.1$, which applies to the conditions of this Theorem. Furthermore, the bandwidth parameter h_n is obtained from the cross-validation method.

As can be seen in the Figure 1, with the increase of the sample size n, the desired value tends to zero, that is a confirmation for the correctness of Theorem 2.3.

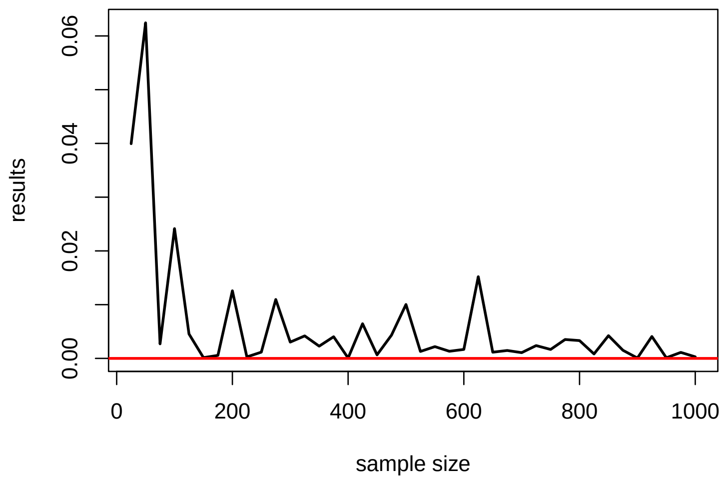

FIGURE 1. Values of $\sup_{0\leq t<\infty} n^{\alpha}|F_n^s(t) - F_n^c(t)|$ for different n.

Corollary 2.5. *Under the Assumptions of Theorems 2.1 and 2.3, we have*

$$\lim_{n\to 0} \sup_{0\leq t<\infty} n^{\alpha}|F_n^s(t) - F(t)| = 0, \quad a.s.$$

Proof. Using the triangle inequality, we have

$$|F_n^s(t) - F(t)| \leq |F_n^s(t) - F_n^c(t)| + |F_n^c(t) - F(t)|. \tag{16}$$

According to Theorem 2.1

$$\lim_{n\to\infty} \sup_{0\leq t<\infty} n^{\alpha}|F_n^c(t) - F(t)| = 0, \quad a.s. \tag{17}$$

On the other, based on Theorem 2.3

$$\lim_{n\to 0} \sup_{0\leq t<\infty} n^{\alpha}|F_n^s(t) - F_n^c(t)| = 0, \quad a.s. \tag{18}$$

So by placing (17) and (18) in (16) the proof is completed. □

3. Concluding

In this paper, we gave an overview of the Asymptotic properties such as the Chung-Smirnov property and the law of the iterated logarithm for the kernel-type estimators of CDF function. Also, based on length-biased sampling, we proved the extended Glivenko-Cantelli theorem for non-smooth estimator $F_n^c(\cdot)$

and the smooth estimator $F_n^s(\cdot)$. In the following, we checked the correctness of Theorem 2.3 in a special case by using the simulation study.

4. Aknowledgement

The authors are grateful to anonymous referees and the associate editor for providing some useful comments on an earlier version of this manuscript.

References

[1] Blumenthal, S. (1967). Proportional sampling in life length studies. Technometrics, 9(2) 205–218. https://doi.org/10.1080/00401706.1967.10490456

[2] Cox, D. R. (1969). Some sampling problems in technology. In: Johnson NL, Smith H Jr New developments in survey sampling, Wiley, New York, 506–527.

[3] Degenhardt, HJA. (1993). Chung—Smirnov property for perturbed empirical distribution functions. Statistics and probability letters, 16(2), 97–101. https://doi.org/10.1016/0167-7152(93)90152-9

[4] Degenhardt, HJA. (1996). Strong limit theorems for the difference of the perturbed empirical distribution function and the classical empirical distribution function. Scandinavian journal of statistics, 23(3), 331–351.

[5] Fabian, V., & Hannan, J. (1985). Introduction to probability and mathematical statistics. John Wiley & Sons Incorporated.

[6] Falk, M. (1983). Relative efficiency and deficiency of kernel type estimators of smooth distribution functions. Statistica Neerlandica, 37(2), 73–83. https://doi.org/10.1111/j.1467-9574.1983.tb00802.x

[7] Fernholz, L. T. (1991). Almost sure convergence of smoothed empirical distribution functions. Scandinavian Journal of Statistics, 22, 255–262.

[8] Horváth, L. (1985). Estimation from a length-biased distribution. Statistics & Risk Modeling, 3(1-2), 91–114. https://doi.org/10.1524/strm.1985.3.12.91

[9] Jahanshahi, SMA., Rad, A., & Fakoor, V. (2017). Goodness-of-fit test under length-biased sampling. Communications in Statistics-Theory and Methods, 46(15), 7580–7592. https://doi.org/10.1080/03610926.2016.1157187

[10] McFadden, J. A. (1962). On the lengths of intervals in a stationary point process. Journal of the Royal Statistical Society. Series B (Methodological), 364–382. https://doi.org/10.1111/j.2517-6161.1962.tb00464.x

[11] Nadaraya, E. (1964). Some new estimates for distribution functions. Theory of Probability & Its Applications, 9(3), 497–500. https://doi.org/10.1137/1109069

[12] Reiss, R. D. (1981). Nonparametric estimation of smooth distribution functions. Scandinavian Journal of Statistics, 12, 116–119.

[13] Shorack, G. R, & Wellner, J. A. (1986). Empirical Processes with Applications to Statistics. John Wiley & Sons Incorporated.

[14] Vardi, Y. (1982). Nonparametric estimation in renewal processes. The Annals of Statistics, 10(3), 772–785.

[15] Wicksell, S. D. (1925). The corpuscle problem: a mathematical study of a biometric problem. Biometrika, 17(1-2), 84–99. https://doi.org/10.2307/2332027

[16] Winter, BB. (1979). Convergence rate of perturbed empirical distribution functions. Journal of Applied Probability, 16(1), 163–173. https://doi.org/10.2307/3213384

[17] Yamato, H. (1973). Uniform convergence of an estimator of a distribution function. Bulletin of Mathematical Statistics, 15(3), 69–78. https://doi.org/10.5109/13073

[18] Yukich, J. E. (1992). Weak convergence of smoothed empirical processes. Scandinavian Journal of Statistics, 19(3), 271–279.

[19] Zamini, R., Ajami, M., & Fakoor, V. (2024). Berry-Esseen bound for smooth estimator of distribution function under length-biased data. Communications in Statistics-Theory and Methods, 53(5), 1800–1809. https://doi.org/10.1080/03610926.2022.2112695

M. Ajami
Orcid number: 0000-0002-6045-128x
Department of Statistics, Faculty of Mathematical Sciences,
Vali-e-Asr University of Rafsanjan
Rafsanjan, Iran
Email address: m.ajami@vru.ac.ir

R. Zamini
Orcid number: 0009-0001-4502-7763
Department of Mathematics, Faculty of Mathematical Sciences and Computer
Kharazmi University
Thehran, Iran
Email address: rahelehzamini@yahoo.com

SMA. Jahanshahi
Orcid number: 0000-0002-3571-886X
Department of Statistics,Faculty of Mathematics
University of Sistan and Baluchistan
Zahedan, Iran
Email address: mjahan@math.usb.ac.ir