ASYMPTOTIC MEAN SQUARED ERROR OF KERNEL ESTIMATOR OF EXCESS DISTRIBUTION FUNCTION

Shimokihara, Atsushi
Graduate School of Mathematics, Kyushu University

Maesono, Yoshihiko
Faculty of Mathematics, Kyushu University

https://doi.org/10.5109/2233859

出版情報：Bulletin of informatics and cybernetics. 50, pp.51-64, 2018-12. Research Association of Statistical Sciences
バージョン：
権利関係：

ASYMPTOTIC MEAN SQUARED ERROR OF KERNEL ESTIMATOR OF EXCESS DISTRIBUTION FUNCTION

by

Atsushi Shimokihara and **Yoshihiko Maesono**

Reprinted from the Bulletin of Informatics and Cybernetics
Research Association of Statistical Sciences, Vol.50

FUKUOKA, JAPAN
2018

Bulletin of Informatics and Cybernetics, Vol. 50, 2018

ASYMPTOTIC MEAN SQUARED ERROR OF KERNEL ESTIMATOR OF EXCESS DISTRIBUTION FUNCTION

By

Atsushi Shimokihara* and **Yoshihiko Maesono†**

Abstract

In this paper we will discuss asymptotic properties of a kernel estimator of excess distribution function (EDF). The excess distribution function takes an important role in extreme value analysis, survival analysis, and so on. The excess distribution function is a conditional distribution function $H_u(x) = P(X - u \leq x | X > u)$ $(x > 0)$. Thus it is a ratio of the functions which relate to the distribution function X. If we can assume a parametric model, we can get an estimator of EDF. Also, using the empirical distribution function, a nonparametric estimator of EDF is obtained. Since the empirical distribution is not smooth, we propose a kernel type estimator of EDF, which gives us a smooth estimate, and discuss its mean squared errors, theoretically.

1. Introduction

Let $X_1, X_2, \ldots, X_n$ be independently and identically distributed random variables with a density and distribution function $f(x)$, $F(x)$. From the definition of the conditional probability, we have the excess distribution function (EDF) as follows: for $x > 0$

$$H_u(x) = P(X_1 - u \leq x | X_> u) = \frac{F(x+u) - F(u)}{1 - F(u)}.$$

The above EDF is a distribution function of residual time after survive until u. If we can assume some parametric model, it is possible to obtain a smooth estimaor of EDF. Using the empirical distribution function, a nonparametric estimator is given by

$$\widetilde{H}_u(x) = \frac{F_n(x+u) - F_n(u)}{1 - F_n(u)}$$

where for the indicator function $I(\cdot)$

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I(X_i \leq x).$$

Since $F_n(\cdot)$ is a discrete function, the estimator $\widetilde{H}_u(x)$ is not smooth enough. On the other hand, the kernel estimator of the distribution function is smooth. In this

* Graduate School of Mathematics, Motooka 744, Nishi-ku Fukuoka 819-0395, Japan.
† Faculty of Mathematics, Kyushu University, Motooka 744, Nishi-ku Fukuoka 819-0395, Japan. Email: maesono@math.kyushu-u.ac.jp

paper we study a kernel estimator of $H_u(x)$ and obtain mean squared errors. Let us consider a kernel function $k(\cdot)$ which satisfies

$$\int_{-\infty}^{\infty} k(v)dv = 1.$$

Using this kernel function, we get a kernel estimator of the distribution function $F(x)$

$$\widehat{F}(x) = \frac{1}{n}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right)$$

where $K(\cdot)$ is an integral of the kernel function $k(\cdot)$, that is

$$K(t) = \int_{-\infty}^{t} k(v)dv.$$

For the sake of simplicity, we assume that the support of the kernel is $[-1,\ 1]$. Then we will discuss mean squared errors of a kernel excess distribution function estimator

$$\widehat{H}_u(x) = \frac{\widehat{F}(x+u) - \widehat{F}(u)}{1 - \widehat{F}(u)} \qquad (x > 0). \tag{1}$$

For this kernel estimator $\widehat{H}_u(x)$, we get an asymptotic bias, variance and mean squared error.

Theorem 1.1. *Assume that $f^{(j)}(x)$ $(j = 1, \dots, 4)$ exists and is bounded around the neighborhood of u.*

$$\int z^j k(z)dz < \infty \qquad (j = 1, 2, 3, 4).$$

Then we have the bias, variance and mean squared error as follows:

$$\begin{aligned}
& Bias\left[\widehat{H}_u(x)\right] \\
= \; & \left\{\frac{\beta_2(x;u)}{1 - F(u)} + \frac{\Delta_u(x)\alpha_2(u)}{(1 - F(u))^2}\right\} h^2 \\
& + \left\{\frac{\beta_4(x;u)}{1 - F(u)} + \frac{\alpha_2(u)\beta_2(x;u) + \alpha_4(u)\Delta_u(x)}{(1 - F(u))^2} + \frac{\alpha_2^2(u)\Delta_u(x)}{(1 - F(u))^3}\right\} h^4 \\
& + o\left(\frac{1}{n}\right) + o(h^4),
\end{aligned}$$

$$\begin{aligned}
& Var\left[\widehat{H}_{n,u}(x)\right] \\
= \; & \frac{1}{n}\frac{(F(x+u) - F(u))(1 - F(x+u))}{(1 - F(u))^3} + o\left(\frac{1}{n}\right) + o(h^4),
\end{aligned}$$

$$\begin{aligned}
& AMSE\left[\widehat{H}_{n,u}(x)\right] \\
= \; & \frac{1}{n}\frac{(F(x+u) - F(u))(1 - F(x+u))}{(1 - F(u))^3} \\
& + \left\{\frac{\beta_2(x;u)}{1 - F(u)} + \frac{\Delta_u(x)\alpha_2(u)}{(1 - F(u))^2}\right\}^2 h^4 + o\left(\frac{1}{n}\right) + o(h^4)
\end{aligned}$$

wehre for $j = 1, \dots, 4$

$$\begin{aligned}\alpha_j(u) &= \frac{1}{j!} f^{(j-1)}(u) \int y^j k(y) dy, \\ \beta_j(x;u) &= \alpha_j(x+u) - \alpha_j(u)\end{aligned}$$

and $\Delta_u(x) = F(x+u) - F(u)$.

PROOF. See the appendix.

2. Simulation

In this section, we will compare the obtained asymptotic mean squared errors and the mean squared errors by simulation. For fix zs, we draw the asymptotic mean squared errors

$$\begin{aligned}& E\left[\left(\widehat{H}_{n,u}(x) - H_u(x)\right)^2\right] \\ \approx\ & \frac{(F(x+u) - F(u))(1 - F(x+u))}{n(1-F(u))^3} + \left\{ \frac{\beta_2(x;u)}{1 - F(u)} + \frac{\Delta_u(x)\alpha_2(u)}{(1-F(u))^2} \right\}^2 h^4 \qquad (2)\end{aligned}$$

and simulated mean squared errors

$$\left(\widehat{H}^N_{n,u}(x) - H_u(x)\right)^2. \qquad (3)$$

Using the Epanechnikov kernel

$$k(u) = \frac{3}{4}(1-u^2) I(|u| \le 1),$$

we simulate the mean squared error. The Epanechnikov kernel satisfies

$$\int z^j k(z) dz < \infty \qquad (j = 1, 2, 3, 4).$$

In these simulations, we put the bandwidth $h = n^{-1/4}$ and repetition number $10,000$. The vertical line represents both mean squared errors and the horizontal line represents $x > 0$. Here we consider the underlying distributions are t-distribution with 3-degrees of freedom and double-exponential. The upper line is the asymptotic mean squared errors (2) and the lower line is the estimated values of the equation (3).It follows from the simulation results that when n goes to infinity the estimated values are closer to the asymptotic mean squared errors. This supports that our derivations are correct.

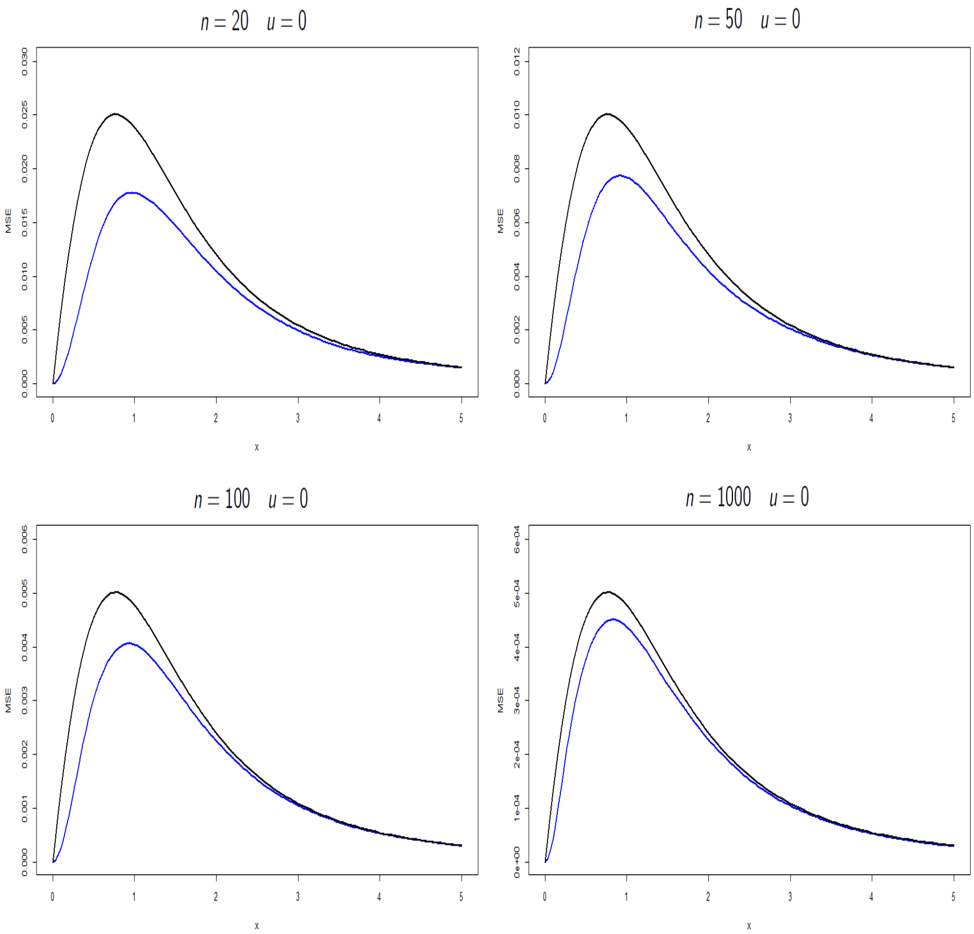

Figure 1: t-distribution with 3-degrees of freedom.

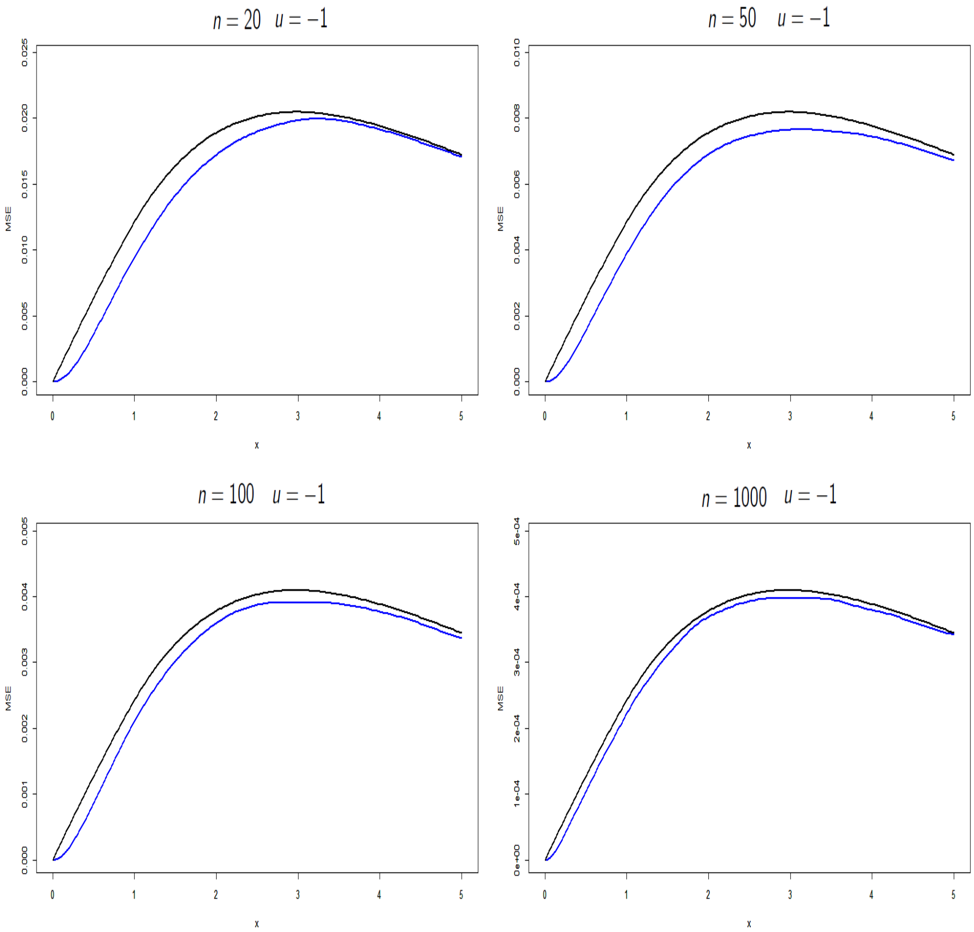

Figure 2: Double exponential distribution

3. Appendix

Let us obtain moment evaluations of the kernel estimator of the distribution function. Motoyama and Maesono (2018) obtain the following evaluation.

LEMMA 3.1. *Assume that* (i)$f'(u)$ *exists and* $|f'(u)| \leq M$ *for* $M > 0$ *and* $u \in \mathbf{R}$, (ii) $k(\cdot)$ *is a symmetric around the origin 0 and satisfies* $\int u^2 k(u)du < \infty$. *Then for* $p \geq 2$ *we have*

$$E\{|\widehat{F}(x) - F(x)|^p\} = O(h^{2p}) + O\left(\frac{1}{n^{p/2}}\right).$$

PROOF. See Motoyama and Maesono (2018).

Using the Taylor expansion of $(1-x)^{-1}$, we have

$$\begin{aligned} H_{n,u}(x) \quad \approx \quad & \frac{1}{1-F(u)}(\widehat{F}_n(x+u) - \widehat{F}_n(u)) \\ & + \frac{1}{(1-F(u))^2}(\widehat{F}_n(u) - F(u))(\widehat{F}_n(x+u) - \widehat{F}_n(u)) \\ & + \frac{1}{(1-F(u))^3}(\widehat{F}_n(u) - F(u))^2(\widehat{F}_n(x+u) - \widehat{F}_n(u)). \end{aligned}$$

In order to get approximations of the moments of

$$\begin{aligned} & E\left[\widehat{F}_n(x+u) - \widehat{F}_n(u)\right], \\ & E\left[(\widehat{F}_n(x) - F(u))(\widehat{F}_n(x+u) - \widehat{F}_n(u))\right], \\ & E\left[(\widehat{F}_n(u) - F(u))^2(\widehat{F}_n(x+u) - \widehat{F}_n(u))\right], \end{aligned}$$

we prepare the following lemma.

LEMMA 3.2. *Under the same assumptions of Theorem 1.1, we have the followings:*
(1) *For* $m \geq 2$, *we have*

$$E\left[K^m\left(\frac{u-X_1}{h}\right)\right] = F(u) + \sum_{j=1}^{5} \alpha_{m,j}(u)h^j + O(h^6) \tag{4}$$

where

$$\alpha_{\ell,j}(u) = \frac{(-1)^j \ell}{j!} f^{(j-1)}(u) \int_{-1}^{1} y^j K^{\ell-1}(y)k(y)dy.$$

(2) *For* $e_K(u) = K\left(\frac{u-X_1}{h}\right) - E\left[K\left(\frac{u-X_1}{h}\right)\right]$, *we have*

$$E\left[e_K^2(u)\right] = F(u)(1-F(u)) + o(1), \tag{5}$$

$$E\left[e_K^m(u)\right] = O(1) \ (m = 3, 4). \tag{6}$$

PROOF. (1) Using the integration by parts and the Taylor expansion, we get

$$
\begin{aligned}
& E\left[K^m\left(\frac{u-X_1}{h}\right)\right] \\
= & \int K^m\left(\frac{u-z}{h}\right) f(z)dz \\
= & \left[K^m\left(\frac{u-z}{h}\right) F(z)\right]_{z=-\infty}^{z=\infty} + \frac{m}{h}\int K^{m-1}\left(\frac{u-z}{h}\right) k\left(\frac{u-z}{h}\right) F(z)dz \\
= & m\int K^{m-1}(y)k(y)F(u-hy)dy \\
= & mF(u)\int_{-1}^{1} K^{m-1}(y)k(y)dy \\
& + m\sum_{j=1}^{5}\frac{(-1)^j h^j}{j!} f^{(j-1)}(u)\int_{-1}^{1} y^j K^{m-1}(y)k(y)dy + O(h^6) \\
= & F(u)\left[K^m(y)\right]_{-1}^{1} + \sum_{j=1}^{5}\alpha_{m,j}(u)h^j + O(h^6) \\
= & F(u) + \sum_{j=1}^{5}\alpha_{m,j}(u)h^j + O(h^6).
\end{aligned}
$$

(2) It follows from the above result (1) that

$$
\begin{aligned}
& E\left[e_K^2(u)\right] \\
= & E\left[K^2\left(\frac{u-X_1}{h}\right)\right] - \left(E\left[K\left(\frac{u-X_1}{h}\right)\right]\right)^2 \\
= & F(u) + \sum_{j=1}^{5}\alpha_{2,j}(u)h^j + O(h^6) \\
& - \left\{F^2(u) + 2F(u)\alpha_2(u)h^2 + \left(\alpha_2^2(u) + 2F(u)\alpha_4(u)\right)h^4\right\} \\
= & F(u)(1-F(u)) + \sum_{j=1}^{5} c_j(u)h^j + O(h^6) \\
= & F(u)(1-F(u)) + o(1)
\end{aligned}
$$

where c_js are coefficient of h^j and do not depend on n.

Proof of Theorem 1.1.
Using Lemma 3.2, we will obtain approximations of the following expectations:

$$
\begin{aligned}
& E\left[\widehat{F}_n^i(u)\right] \quad (i=1,2,3,4), \\
& E\left[\widehat{F}_n^j(u)\widehat{F}_n(x+u)\right] \quad (j=1,2,3), \\
& E\left[\widehat{F}_n^j(u)\widehat{F}_n^2(x+u)\right] \quad (j=1,2).
\end{aligned}
$$

It is easy to show that

$$\begin{aligned}
E\left[\widehat{F}_n(u)\right] &= F(u)+\alpha_2(u)h^2+\alpha_4(u)h^4+O(h^6),\\
E\left[\widehat{F}_n^2(u)\right] &= E\left[\left(\widehat{F}_n(u)-E\left[\widehat{F}_n(u)\right]\right)^2\right]+\left(E\left[\widehat{F}_n(u)\right]\right)^2\\
&= \frac{1}{n}E\left[e_K^2(u)\right]+\left(E\left[\widehat{F}_n(u)\right]\right)^2\\
&= F^2(u)+2F(u)\alpha_2(u)h^2+(\alpha_2^2(u)+2F(u)\alpha_4(u))h^4\\
&\quad +\frac{1}{n}F(u)(1-F(u))+o\left(\frac{1}{n}\right)+o(h^4),
\end{aligned}$$

$$\begin{aligned}
E\left[\widehat{F}_n^3(u)\right] &= E\left[\left(\widehat{F}_n(u)-E\left[\widehat{F}_n(u)\right]\right)^3\right]+3E\left[\widehat{F}_n^2(u)\right]E\left[\widehat{F}_n(u)\right]\\
&\quad -3E\left[\widehat{F}_n(u)\right]\left(E\left[\widehat{F}_n(u)\right]\right)^2+\left(E\left[\widehat{F}_n(u)\right]\right)^3\\
&= \frac{1}{n^2}E\left[e_K^3(u)\right]+3E\left[\widehat{F}_n^2(u)\right]E\left[\widehat{F}_n(u)\right]-2\left(E\left[\widehat{F}_n(u)\right]\right)^3\\
&= F^3(u)+3F^2(u)\alpha_2(u)h^2+3F(u)(\alpha_2^2(u)+\alpha_4(u)F(u))h^4\\
&\quad +\frac{3}{n}F^2(u)(1-F(u))+o\left(\frac{1}{n}\right)+o(h^4)
\end{aligned}$$

and

$$\begin{aligned}
E\left[\widehat{F}_n^4(u)\right] &= E\left[\left(\widehat{F}_n(u)-E\left[\widehat{F}_n(u)\right]\right)^4\right]+4E\left[\widehat{F}_n^3(u)\right]E\left[\widehat{F}_n(u)\right]\\
&\quad -6E\left[\widehat{F}_n^2(u)\right]\left(E\left[\widehat{F}_n(u)\right]\right)^2+3\left(E\left[\widehat{F}_n(u)\right]\right)^4\\
&= \frac{1}{n^3}E\left[e_K^4(u)\right]+4E\left[\widehat{F}_n^3(u)\right]E\left[\widehat{F}_n(u)\right]\\
&\quad -6E\left[\widehat{F}_n^2(u)\right]\left(E\left[\widehat{F}_n(u)\right]\right)^2+3\left(E\left[\widehat{F}_n(u)\right]\right)^4\\
&= F^4(u)+4F^3(u)\alpha_2(u)h^2+2F^2(u)(3\alpha_2^2(u)+2\alpha_4(u)F(u))h^4\\
&\quad +\frac{6}{n}F^3(u)(1-F(u))+o\left(\frac{1}{n}\right)+o(h^4).
\end{aligned}$$

In order to get approximations of $E\left[\widehat{F}_n^j(u)\widehat{F}_n(x+u)\right]$ $(j=1,2,3)$, let us consider

$$E\left[K\left(\frac{u-X_1}{h}\right)K\left(\frac{x+u-X_1}{h}\right)\right].$$

Using the integration by parts, we have

$$\begin{aligned}
& E\left[K\left(\frac{u-X_1}{h}\right)K\left(\frac{x+u-X_1}{h}\right)\right] \\
= & \int K\left(\frac{u-z}{h}\right)K\left(\frac{x+u-z}{h}\right)f(z)dz \\
= & \left[K\left(\frac{u-z}{h}\right)K\left(\frac{x+u-z}{h}\right)F(z)\right]_{z=-\infty}^{z=\infty} \\
& +\frac{1}{h}\int k\left(\frac{u-z}{h}\right)K\left(\frac{x+u-z}{h}\right)F(z)dz \\
& +\frac{1}{h}\int K\left(\frac{u-z}{h}\right)k\left(\frac{x+u-z}{h}\right)F(z)dz \\
= & \int_{-1}^{1} k(y)K\left(y+\frac{x}{h}\right)F(u-hy)dy \\
& +\int_{-1}^{1} K\left(y-\frac{x}{h}\right)k(y)F(x+u-hy)dy.
\end{aligned}$$

Since the support of the kernel $k(\cdot)$ is $[-1,\ 1]$ and $h \to 0$, for a fix $x > 0$ and $y \in [-1, 1]$ we have

$$1 \leq y + \frac{x}{h_n},$$

and then

$$K\left(y+\frac{x}{h}\right) = 1.$$

Thus we get

$$\begin{aligned}
& \int_{-1}^{1} k(y)K\left(y+\frac{x}{h}\right)F(u-hy)dy \\
= & \int_{-1}^{1} k(y)F(u-hy)dy \\
= & F(u)+\alpha_2(u)h^2+\alpha_4(u)h^4+O(h^6).
\end{aligned}$$

Similarly for the fix $x > 0$ and $y \in [-1, 1]$, we get

$$-1 \geq y - \frac{x}{h_n}$$

and

$$K\left(y-\frac{x}{h}\right) = 0.$$

Therefore, we have

$$\int_{-1}^{1} K\left(y-\frac{x}{h}\right)k(y)F(x+u-hy)dy = 0.$$

Using these evaluations, we can show that

$$E\left[K\left(\frac{u-X_1}{h}\right)K\left(\frac{x+u-X_1}{h}\right)\right] = F(u)+\alpha_2(u)h^2+\alpha_4(u)h^4+O(h^6)$$

and

$$E\left[e_K(u)e_K(x+u)\right] = F(u)(1-F(x+u))+o(1).$$

Then for the fix $x > 0$ we get

$$\begin{aligned}
& E\left[\widehat{F}_n(u)\widehat{F}_n(x+u)\right] \\
= & E\left[\left(\widehat{F}_n(u) - E\left[\widehat{F}_n(u)\right]\right)\left(\widehat{F}_n(x+u) - E\left[\widehat{F}_n(x+u)\right]\right)\right] \\
& +E\left[\widehat{F}_n(u)\right] E\left[\widehat{F}_n(x+u)\right] \\
= & \frac{1}{n} E\left[e_K(u)e_K(x+u)\right] + E\left[\widehat{F}_n(u)\right] E\left[\widehat{F}_n(x+u)\right] \\
= & F(u)F(x+u) + \frac{1}{n}F(u)(1-F(x+u)) + \left(\alpha_2(u)F(x+u) + \alpha_2(x+u)F(u)\right)h^2 \\
& +(F(u)\alpha_4(x+u) + \alpha_2(u)\alpha_2(x+u) + \alpha_4(u)F(x+u))h^4 + o\left(\frac{1}{n}\right) + o(h^4).
\end{aligned}$$

Similarly, we can show that

$$\begin{aligned}
& E\left[\widehat{F}_n^2(u)\widehat{F}_n(x+u)\right] \\
= & F^2(u)F(x+u) + \left(2F(u)F(x+u)\alpha_2(u) + F^2(u)\alpha_2(x+u)\right)h^2 \\
& +\Big\{F^2(u)\alpha_4(x+u) + 2F(u)\alpha_2(u)\alpha_2(x+u) \\
& +F(x+u)\alpha_2^2(u) + 2F(u)F(x+u)\alpha_4(u)\Big\}h^4 \\
& +\frac{1}{n}F(u)(F(x+u) + 2F(u) - 3F(u)F(x+u)) + o\left(\frac{1}{n}\right) + o(h^4), \\
& E\left[\widehat{F}_n^3(u)\widehat{F}_n(x+u)\right] \\
= & F^3(u)F(x+u) + \left(3F^2(u)F(x+u)\alpha_2(u) + F^3(u)\alpha_2(x+u)\right)h^2 \\
& +\Big\{F^3(u)\alpha_4(x+u) + 3F^2(u)\alpha_2(u)\alpha_2(x+u) \\
& +3F(u)F(x+u)\alpha_2^2(u) + 3F^2(u)F(x+u)\alpha_4(u)\Big\}h^4 \\
& +\frac{3}{n}F^2(u)(F(x+u) + F(u) - 2F(u)F(x+u)) + o\left(\frac{1}{n}\right) + o(h^4), \\
& E\left[\widehat{F}_n(u)\widehat{F}_n^2(x+u)\right] \\
= & F(u)F^2(x+u) + \left(2F(u)F(x+u)\alpha_2(x+u) + F^2(x+u)\alpha_2(u)\right)h^2 \\
& +\Big\{F^2(x+u)\alpha_4(u) + F(u)\alpha_2^2(x+u) \\
& +2F(x+u)\alpha_2(x+u)\alpha_2(u) + 2F(u)F(x+u)\alpha_4(x+u)\Big\}h^4 \\
& +\frac{3}{n}F(u)F(x+u)(1-F(x+u)) + o\left(\frac{1}{n}\right) + o(h^4)
\end{aligned}$$

and

$$
\begin{aligned}
& E\left[\widehat{F}_n^2(u)\widehat{F}_n^2(x+u)\right] \\
= \ & F^2(u)F^2(x+u) + 2F(u)F(x+u)\left(F(x+u)\alpha_2(u) + F(u)\alpha_2(x+u)\right)h^2 \\
& +\Big\{2\alpha_4(x+u)F^2(u)F(x+u) + 4F(u)F(x+u)\alpha_2(u)\alpha_2(x+u) \\
& +F^2(u)\alpha_2^2(x+u) + F^2(x+u)\alpha_2^2(u) + 2F^2(x+u)F(u)\alpha_4(u)\Big\}h^4 \\
& +\frac{1}{n}F(u)F(x+u)(F(x+u) + 5F(u) - 6F(u)F(x+u)) \\
& +o\left(\frac{1}{n}\right) + o(h^4).
\end{aligned}
$$

Using these evaluations, we can show that

$$
\begin{aligned}
& E\left[\widehat{F}_n(x+u) - \widehat{F}_n(u)\right] \\
= \ & F(x+u) - F(u) + \beta_2(x;u)h^2 + \beta_4(x;u)h^4 + o\left(\frac{1}{n}\right) + o(h^4), \\
& E\left[\left(\widehat{F}_n(x+u) - \widehat{F}_n(u)\right)\left(\widehat{F}_n(u) - F(u)\right)\right] \\
= \ & -\frac{1}{n}F(u)\Delta_u(x) + \Delta_u(x)\alpha_2(u)h^2 \\
& +(\alpha_2(u)\beta_2(x;u) + \alpha_4(u)\Delta_u(x))h^4 + o\left(\frac{1}{n}\right) + o(h^4), \\
& E\left[\left(\widehat{F}_n(x+u) - \widehat{F}_n(u)\right)\left(\widehat{F}_n(u) - F(u)\right)^2\right] \\
= \ & \alpha_2^2(u)\Delta_u(x)h^4 + \frac{1}{n}F(u)\Delta_u(x)(1 - F(u)) + o\left(\frac{1}{n}\right) + o(h^4)
\end{aligned}
$$

where

$$
\begin{aligned}
\beta_j(x;u) &= \alpha_j(x+u) - \alpha_j(u) \\
\Delta_u(x) &= F(x+u) - F(u).
\end{aligned}
$$

Combining the above evaluations, we get

$$
\begin{aligned}
& E\left[\widehat{H}_{n,u}(x)\right] \\
\approx \ & \frac{1}{1-F(u)}E\left[\widehat{F}_n(x+u) - \widehat{F}_n(u)\right] \\
& +\frac{1}{(1-F(u))^2}E\left[\left(\widehat{F}_n(x+u) - \widehat{F}_n(u)\right)\left(\widehat{F}_n(u) - F(u)\right)\right] \\
& +\frac{1}{(1-F(u))^3}E\left[\left(\widehat{F}_n(x+u) - \widehat{F}_n(u)\right)\left(\widehat{F}_n(u) - F(u)\right)^2\right] \\
= \ & H_u(x) + \left\{\frac{\beta_2(x;u)}{1-F(u)} + \frac{\Delta_u(x)\alpha_2(u)}{(1-F(u))^2}\right\}h^2 \\
& +\left\{\frac{\beta_4(x;u)}{1-F(u)} + \frac{\alpha_2(u)\beta_2(x;u) + \alpha_4(u)\Delta_u(x)}{(1-F(u))^2} + \frac{\alpha_2^2(u)\Delta_u(x)}{(1-F(u))^3}\right\}h^4 \\
& +o\left(\frac{1}{n}\right) + o(h^4).
\end{aligned}
$$

Then we have the bias of $\widehat{H}_{n,u}(x)$

$$\begin{aligned} & Bias\left[\widehat{H}_{n,u}(x)\right] \\ \approx\ & \left\{\frac{\beta_2(x;u)}{1-F(u)}+\frac{\Delta_u(x)\alpha_2(u)}{(1-F(u))^2}\right\}h^2 \\ & +\left\{\frac{\beta_4(x;u)}{1-F(u)}+\frac{\alpha_2(u)\beta_2(x;u)+\alpha_4(u)\Delta_u(x)}{(1-F(u))^2}+\frac{\alpha_2^2(u)\Delta_u(x)}{(1-F(u))^3}\right\}h^4 \\ & +o\left(\frac{1}{n}\right)+o(h^4). \end{aligned}$$

Next we will study the variance of $\widehat{H}_{n,u}(x)$. Using the approximation of the ratio statistic, we have

$$\begin{aligned} & \widehat{H}^2_{n,u}(x) \\ \approx\ & \frac{1}{(1-F(u))^2}\left(\widehat{F}_n(x+u)-\widehat{F}_n(u)\right)^2 \\ & +\frac{2}{(1-F(u))^3}\left(\widehat{F}_n(x+u)-\widehat{F}_n(u)\right)^2\left(\widehat{F}_n(u)-F(u)\right) \\ & +\frac{3}{(1-F(u))^4}\left(\widehat{F}_n(x+u)-\widehat{F}_n(u)\right)^2\left(\widehat{F}_n(u)-F(u)\right)^2. \end{aligned}$$

In order to obtain an expectation of $\widehat{H}^2_{n,u}(x)$, we need approximations of the following expectations:

$$\begin{aligned} & E\left[\left(\widehat{F}_n(x+u)-\widehat{F}_n(u)\right)^2\right], \\ & E\left[\left(\widehat{F}_n(x+u)-\widehat{F}_n(u)\right)^2\left(\widehat{F}_n(u)-F(u)\right)\right], \\ & E\left[\left(\widehat{F}_n(x+u)-\widehat{F}_n(u)\right)^2\left(\widehat{F}_n(u)-F(u)\right)^2\right]. \end{aligned}$$

Using the previous evaluations, we have

$$\begin{aligned} & E\left[\left(\widehat{F}_n(x+u)-\widehat{F}_n(u)\right)^2\right] \\ =\ & \Delta_u^2(x)+2\Delta_u(x)\beta_2(x;u)h^2+(\beta_2^2(x;u)+2\beta_4(x;u)\Delta_u(x))h^4 \\ & +\frac{1}{n}\Delta_u(x)(1-\Delta_u(x))+o\left(\frac{1}{n}\right)+o(h^4), \end{aligned}$$

$$\begin{aligned} & E\left[\left(\widehat{F}_n(x+u)-\widehat{F}_n(u)\right)^2\left(\widehat{F}_n(u)-F(u)\right)\right] \\ =\ & \Delta_u^2(x)\alpha_2(u)h^2+(\alpha_4(u)\Delta_u^2(x)+2\alpha_2(u)\Delta_u(x)\beta_2(x;u))h^4 \\ & -\frac{2}{n}F(u)\Delta_u^2(x)+o\left(\frac{1}{n}\right)+o(h^4) \end{aligned}$$

and

$$\begin{aligned} & E\left[\left(\widehat{F}_n(x+u)-\widehat{F}_n(u)\right)^2\left(\widehat{F}_n(u)-F(u)\right)^2\right] \\ =\ & \frac{1}{n}F(u)(1-F(u))\Delta_u^2(x)+\alpha_2^2(u)\Delta_u^2(x)h^4+o\left(\frac{1}{n}\right)+o(h^4). \end{aligned}$$

Therefore, we have

$$\begin{aligned}
& E\left[\widehat{H}^2_{n,u}(x)\right] \\
\approx\ & \frac{1}{(1-F(u))^2}E\left[\left(\widehat{F}_n(x+u)-\widehat{F}_n(u)\right)^2\right] \\
& +\frac{2}{(1-F(u))^3}E\left[\left(\widehat{F}_n(x+u)-\widehat{F}_n(u)\right)^2\left(\widehat{F}_n(u)-F(u)\right)\right] \\
& +\frac{3}{(1-F(u))^4}E\left[\left(\widehat{F}_n(x+u)-\widehat{F}_n(u)\right)^2\left(\widehat{F}_n(u)-F(u)\right)^2\right] \\
=\ & H^2_u(x)+2\left\{\frac{\Delta_u(x)\beta_2(x;u)}{(1-F(u))^2}+\frac{\Delta^2_u(x)\alpha_2(u)}{(1-F(u))^3}\right\}h^2+\left(\sum_{i=1}^{3}\gamma_i(x;u)\right)h^4 \\
& +\frac{1}{n}\frac{(F(x+u)-F(u))(1-F(x+u))}{(1-F(u))^3}+o\left(\frac{1}{n}\right)+o(h^4)
\end{aligned}$$

where

$$\begin{aligned}
\gamma_1(x;u) &= \frac{\beta_2^2(x;u)+2\Delta_u(x)\beta_4(x;u)}{(1-F(u))^2}, \\
\gamma_2(x;u) &= \frac{2\alpha_4(u)\Delta^2_u(x)+4\alpha_2(u)\beta_2(x;u)\Delta_u(x)}{(1-F(u))^3}
\end{aligned}$$

$$\gamma_3(x;u)=\frac{3\alpha_2^2(u)\Delta^2_u(x)}{(1-F(u))^4}.$$

Thus we have

$$\begin{aligned}
& Var\left[\widehat{H}_{n,u}(x)\right] \\
=\ & E\left[\widehat{H}^2_{n,u}(x)\right]-\left(E\left[\widehat{H}_{n,u}(x)\right]\right)^2 \\
\approx\ & \frac{1}{n}\frac{(F(x+u)-F(u))(1-F(x+u))}{(1-F(u))^3}+o\left(\frac{1}{n}\right)+o(h^4)
\end{aligned}$$

and

$$\begin{aligned}
& MSE\left[\widehat{H}_{n,u}(x)\right] \\
\approx\ & \frac{1}{n}\frac{(F(x+u)-F(u))(1-F(x+u))}{(1-F(u))^3} \\
& +\left\{\frac{\beta_2(x;u)}{1-F(u)}+\frac{\Delta_u(x)\alpha_2(u)}{(1-F(u))^2}\right\}^2h^4+o\left(\frac{1}{n}\right)+o(h^4).
\end{aligned}$$

Acknowledgement

This work was supported by JSPS KAKENHI Grant number JP16H02790.

References

Akaike, H. (1954). An approximation to the density function. *Ann. Inst. Statist. Math.*, Vol.6, 127-132

Epanechnikov, V.A. (1969). Non-parametric estimation of a multivariate probability density. *Theory Probab. Appl.*, Vol.14, 153-158

Fix, E. and HodgesJ.L. (1951). Discriminatory analysis nonparametric discrimination: consistency properties. *Report No.4, Project no.21-29-004*

Parzen, E. (1962). On the estimation of a probability density function and the mode. *Ann. Math. Statist.*, Vol.33, 1065-1076

Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. *Ann. Math. Statist.*, Vol.27, 832-837

Alexandre B. Tsybakov (2009). Introduction to Nonparametric Estimation. *Springer Series in Statistics*

Takahashi, R. and Shimura, T. (2016). Kyokuchi Tokeigaku (IMS Series, Shinka suru Tokeisuuri), Kindai-Kagakusha (In Japanese)

Received November 20, 2018
Revised December 17, 2018