Department of Policy and Planning Sciences

Discussion Paper Series

No.1360

Nonparametric direct density ratio estimation

using beta kernel

by

Gaku IGARASHI

May 2019

UNIVERSITY OF TSUKUBA
Tsukuba, Ibaraki 305-8573
JAPAN

Nonparametric direct density ratio estimation using beta kernel

Gaku IGARASHI

Division of Policy and Planning Sciences, Faculty of Engineering, Information and Systems,

University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan

Abstract

A new nonparametric density ratio estimator using the beta kernel is proposed. It is shown that the beta kernel density ratio estimator (KDRE) is free of boundary bias, and the asymptotic properties of the beta KDRE are derived. Simulation studies are conducted to illustrate the finite performance of the beta KDRE.

Keywords: nonparametric density ratio estimator; boundary bias problem; beta kernel;
MSC: 62G07; 62G20

1. Introduction

Let $\{X_1, \ldots, X_n\}$ and $\{Y_1, \ldots, Y_m\}$ be random samples drawn from distribution functions F_X and F_Y with densities f_X and f_Y, respectively. Ćwik and Mielniczuk (1989) suggested, assuming $n = m$, a nonparametric estimator of a density ratio $g(x) = f_X(x)/f_Y(x)$;

$$\widehat{g}_h^{(K)}(x) = \frac{1}{h}\int K\left(\frac{\widehat{F}_{Y,n}(x) - \widehat{F}_{Y,n}(t)}{h}\right)d\widehat{F}_{X,n}(t) = \frac{1}{nh}\sum_{i=1}^n K\left(\frac{\widehat{F}_{Y,n}(x) - \widehat{F}_{Y,n}(X_i)}{h}\right),$$

where K is two times continuously differentiable kernel function with a support $[-A, A]$ for some $A > 0$, $h = h_n$ satisfies $h \to 0$ and $nh^2/\log n \to \infty$ as $n \to \infty$, and $\widehat{F}_{X,n}(x) = n^{-1}\sum_{i=1}^n \chi_{\{X_i \le x\}}$ and $\widehat{F}_{Y,m}(x) = m^{-1}\sum_{j=1}^m \chi_{\{Y_j \le x\}}$ are the empirical distribution functions with the indicator function χ. The Ćwik–Mielniczuk kernel density ratio estimator (KDRE) $\widehat{g}_h^{(K)}$ is motivated by $g(x) = f_X(x)/f_Y(x) = f_{F_Y(X)}(F_Y(x))$ and the kernel density estimator (Rosenblatt (1956) and Parzen (1962)) of f_X;

$$\widehat{f}_{X,h}^{(K)}(x) = \frac{1}{nh}\sum_{i=1}^n K\left(\frac{x - X_i}{h}\right).$$

That is,

$$\widetilde{g}_h^{(K)}(x) = \widehat{f}_{F_Y(X),h}^{(K)}(F_Y(x)) = \frac{1}{nh}\sum_{i=1}^n K\left(\frac{F_Y(x) - F_Y(X_i)}{h}\right)$$

is an (infeasible) estimator of $g(x)$, where unknown F_Y is replaced by $F_{Y,n}$ in the Ćwik–Mielniczuk KDRE. Chen et al. (2009) discussed the asymptotic properties of the Ćwik–Mielniczuk KDRE, but also an indirect KDRE $\widehat{f}_{X,h}^{(K)}(x)/\widehat{f}_{Y,h}^{(K)}(x)$.

If the support of f_X is finite or semi-infinite interval, then the Rosenblatt–Parzen kernel density estimator $\widehat{f}_{X,h}^{(K)}$ is not consistent near the boundary, since its boundary bias is $O(1)$ (e.g.,

Wand and Jones (1995; Section 2.11)). Unfortunately, the Ćwik–Mielniczuk KDRE inherits the boundary problem, regardless of the supports of f_X and f_Y, since $0 \leq F_{Y,n}(X_i) \leq 1$. Gijbels and Mielniczuk (1995) discussed the boundary problem of the Ćwik–Mielniczuk KDRE. In order to avoid the problem, Ćwik and Mielniczuk (1993) and Gijbels and Mielniczuk (1995) considered applying the reflection method which was originally developed for the Rosenblatt–Parzen kernel density estimator (Schuster (1985)), i.e.,

$$\widehat{g}_h^{(K_R)}(x) = \frac{1}{nh} \sum_{i=1}^{n} \left\{ K\left(\frac{\widehat{F}_{Y,m}(x) - \widehat{F}_{Y,m}(X_i)}{h} \right) + K\left(\frac{\widehat{F}_{Y,m}(x) + \widehat{F}_{Y,m}(X_i)}{h} \right) \right.$$
$$\left. + K\left(\frac{\widehat{F}_{Y,m}(x) + \widehat{F}_{Y,m}(X_i) - 2}{h} \right) \right\}.$$

In this paper, we consider a new nonparametric density ratio estimator using the beta kernel

$$K_{b,c_1,c_2}^{(B)}(u_1, u_2) = \frac{u_1^{u_2/b+c_1}(1-u_1)^{(1-u_2)/b+c_2}}{B(u_2/b + c_1 + 1, (1-u_2)/b + c_2 + 1)}, \quad 0 \leq u_1, u_2 \leq 1, \quad c_1, c_2 > -1,$$

which was originally used for estimating a density with support $[0,1]$ by Chen (1999), i.e.,

$$\widehat{f}_{X,b}^{(B)}(x) = \frac{1}{n} \sum_{i=1}^{n} K_{b,0,0}^{(B)}(X_i, x), \quad 0 \leq x \leq 1,$$

where $B(p,q) = \int_0^1 u^{p-1}(1-u)^{q-1}du$ is the beta function, $b = b_n$ is a smoothing parameter.

This paper is organized as follows. In Section 2, the (boundary) bias of Ćwik–Mielniczuk KDRE is derived. In Section 3, the beta KDRE is introduced, and its asymptotic properties are investigated. Simulation studies are conducted to illustrate the finite sample performance of the beta KDRE in Section 4. Section 5 summarize this paper. The proofs are presented in Appendix.

2. Boundary bias of Ćwik–Mielniczuk KDRE

In this section, the boundary bias of the Ćwik–Mielniczuk KDRE is revealed. For this, we impose following assumptions:

A1. $\{X_1, \ldots, X_n\}$ and $\{Y_1, \ldots, Y_m\}$ are random samples from unknown densities f_X and f_Y, respectively, with support $\mathcal{S}$, where $m = m_n$ satisfies $m \to \infty$ as $n \to \infty$.

A2. $g(x) = f_X(x)/f_Y(x)$ is twice continuously differentiable on $\mathcal{S}$.

Theorem 1 *Let $m = n$. Assume that $h = h_n > 0$ is a smoothing parameter satisfying $h \to 0$*

and $nh^3/(\log n)^2 = O(1)$ as $n \to \infty$. The bias of the Ćwik–Mielniczuk KDRE is given by

$$
Bias[\widehat{g}_h^{(K)}(x)]
$$
$$
= \begin{cases}
g(x)\left\{\displaystyle\int_{-A}^{F_Y(x)/h} K(z)du - 1\right\} + O\left(h + \left(\dfrac{\log n}{n}\right)^{1/2} + \dfrac{\log n}{nh^2}\right), & F_Y(x) < Ah, \\[3mm]
\dfrac{h^2}{2}\gamma_2(x)\displaystyle\int_{-A}^{A} z^2 K(z)du + o(h^2) + O\left(\left(\dfrac{\log n}{n}\right)^{1/2} + \dfrac{\log n}{nh^2}\right), & Ah \le F_Y(x) \le 1 - Ah, \\[3mm]
g(x)\left\{\displaystyle\int_{-(1-F_Y(x))/h}^{A} K(z)du - 1\right\} + O\left(h + \left(\dfrac{\log n}{n}\right)^{1/2} + \dfrac{\log n}{nh^2}\right), & 1 - Ah < F_Y(x),
\end{cases}
$$

where

$$
\gamma_2(x) = \frac{f_X''(x)\{f_Y(x)\}^2 - 3f_X'(x)f_Y(x)f_Y'(x) + 3f_X(x)\{f_Y'(x)\}^2 - f_X(x)f_Y(x)f_Y''(x)}{\{f_Y(x)\}^5}.
$$

Proof The stochastic expansion of the Ćwik–Mielniczuk KDRE enables us to see that

$$
\widehat{g}_h^{(K)}(x) = \widetilde{g}_h^{(K)}(x) + \frac{1}{nh^2}\sum_{i=1}^{n}\{\widehat{F}_{Y,n}(x) - \widehat{F}_{Y,n}(X_i) - F_Y(x) + F_Y(X_i)\}K'\left(\frac{F_Y(x) - F_Y(X_i)}{h}\right)
$$
$$
+ \frac{1}{2nh^3}\sum_{i=1}^{n}\{\widehat{F}_{Y,n}(x) - \widehat{F}_{Y,n}(X_i) - F_Y(x) + F_Y(X_i)\}^2
$$
$$
\times K''\left(\frac{(1-\theta)\{F_Y(x) - F_Y(X_i)\} + \theta\{\widehat{F}_{Y,n}(x) - \widehat{F}_{Y,n}(X_i)\}}{h}\right)
$$
$$
= \widetilde{g}_h^{(K)}(x) + J_2 + J_3 \quad \text{(say)}
$$

for some $\theta \in (0,1)$. Chen et al. (2009) showed $E[J_2 + J_3] = O(\{(\log n)/n\}^{1/2} + (\log n)/(nh^2))$. Furthermore, it is easy to see that

$$
E[\widetilde{g}_h^{(K)}(x)] = \frac{1}{h}\int_{\mathcal{S}} K\left(\frac{F_Y(x) - F_Y(s)}{h}\right)f_X(s)ds
$$
$$
= \frac{1}{h}\int_0^1 K\left(\frac{F_Y(x) - u}{h}\right)g(F_Y^{-1}(u))du
$$
$$
= g(x)\int_{-\min\{A,(1-F_Y(x))/h\}}^{\min\{A,F_Y(x)/h\}} K(z)du - h\gamma_1(x)\int_{-\min\{A,(1-F_Y(x))/h\}}^{\min\{A,F_Y(x)/h\}} zK(z)du
$$
$$
+ \frac{h^2}{2}\gamma_2(x)\int_{-\min\{A,(1-F_Y(x))/h\}}^{\min\{A,F_Y(x)/h\}} z^2 K(z)du + o(h^2),
$$

where

$$
\gamma_1(x) = \frac{f_X'(x)f_Y(x) - f_X(x)f_Y'(x)}{\{f_Y(x)\}^3}. \quad \square
$$

Theorem 1 means that the bias of the Ćwik–Mielniczuk KDRE is $O(1)$ near the boundary or tail of the density ratio ($F_Y(x) < Ah$ or $1 - Ah < F_Y(x)$), even if $\mathcal{S} = (-\infty, \infty)$.

3. Beta kernel density ratio estimator

Throughout this paper, in order to derive the asymptotic properties of the beta KDRE defined below, in addition to Assumptions A1 and A2, we impose a following assumption:

3

A2. $b = b_n > 0$ is a smoothing parameter satisfying $b \to 0$ and $mb^2/(\log m)^2 \to \infty$ as $n \to \infty$.

Now, we define the beta KDRE

$$\widehat{g}_b^{(B)}(x) = \frac{1}{n} \sum_{i=1}^{n} \mathcal{K}(\widehat{F}_{Y,m}(X_i), \widehat{F}_{Y,m}(x)), \quad x \in \mathcal{S}, \tag{1}$$

where $\mathcal{K}(u_1, u_2) = K_{b,c,c}^{(B)}(u_1, u_2)$ with $c > -1$. Also, we use the following subsets of $\mathcal{S}$, in order to specify the boundary and interior region;

$$\mathcal{S}_I = \{x \in \mathcal{S} \mid 0 < F_Y(x) < 1\}, \quad \mathcal{S}_B = \{x \in \mathcal{S} \mid F_Y(x) = 0, 1\},$$

$$\mathcal{S}_{I,b} = \{x \in \mathcal{S} \mid F_Y(x)/b \to \infty, (1 - F_Y(x))/b \to \infty\},$$

$$\mathcal{S}_{B,b} = \{x \in \mathcal{S} \mid F_Y(x)/b \to \kappa \text{ or } (1 - F_Y(x))/b \to \kappa\},$$

where $\kappa \geq 0$ is a constant.

The proofs of all Lemmas and Theorems are postponed to Appendix.

3.1. Asymptotic properties

The stochastic expansion of the beta KDRE (1) enables us to see that

$$\widehat{g}_b^{(B)}(x) = \frac{1}{n} \sum_{i=1}^{n} \mathcal{K}(F_Y(X_i), F_Y(x)) + \frac{1}{n} \sum_{i=1}^{n} \{\widehat{F}_{Y,m}(X_i) - F_Y(X_i)\} \mathcal{K}_1(F_Y(X_i), F_Y(x))$$

$$+ \frac{1}{n} \sum_{i=1}^{n} \{\widehat{F}_{Y,m}(x) - F_Y(x)\} \mathcal{K}_2(F_Y(X_i), F_Y(x)) + \frac{1}{n} \sum_{i=1}^{n} \int_0^1 (1 - \theta) \sum_{j=1}^{2} \sum_{k=1}^{2} \mathcal{R}_{j,k} d\theta$$

$$= \widetilde{g}_b^{(B)}(x) + \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{R} \quad \text{(say)},$$

where $\mathcal{K}_j(u_1, u_2) = \partial \mathcal{K}(u_1, u_2)/\partial u_j$ and

$$\mathcal{R}_{j,k} = \{\widehat{F}_{Y,m}(X_i) - F_Y(X_i)\}^{4-j-k} \{\widehat{F}_{Y,m}(x) - F_Y(x)\}^{j+k-2}$$

$$\times \mathcal{K}_{j,k}(F_Y(X_i) + \theta\{\widehat{F}_{Y,m}(X_i) - F_Y(X_i)\}, F_Y(x) + \theta\{\widehat{F}_{Y,m}(x) - F_Y(x)\}),$$

with $\mathcal{K}_{j,k}(u_1, u_2) = \partial \mathcal{K}_j(u_1, u_2)/\partial u_k$. Note that $\widetilde{g}_b^{(B)}(x)$ is an (infeasible) estimator of $g(x)$ (see Lemmas 2 (i) and 4 (i) below for the bias and variance of $\widetilde{g}_b^{(B)}(x)$). Here, it is easy to see that

$$\mathrm{E}[\mathcal{I}_1] = \mathrm{E}[\mathrm{E}[\widehat{F}_{Y,m}(X_1) - F_Y(X_1)|X_1] \mathcal{K}_1(F_Y(X_1), F_Y(x))] = 0,$$

$$\mathrm{E}[\mathcal{I}_2] = \mathrm{E}[\mathrm{E}[\widehat{F}_{Y,m}(x) - F_Y(x)|X_1] \mathcal{K}_2(F_Y(X_1), F_Y(x))] = 0.$$

Therefore, in order to derive the bias of the estimator (1), it is sufficient to derive $\mathrm{E}[\widetilde{g}_b^{(B)}(x)]$ and $\mathrm{E}[\mathcal{R}]$.

Lemma 2 (i). *We have*

$$\mathrm{E}[\widetilde{g}_b^{(B)}(x)] = g(x) + bB(x) + o(b),$$

where

$$B(x) = (c+1)(1 - 2F_Y(x))\gamma_1(x) + \frac{1}{2}F_Y(x)(1 - F_Y(x))\gamma_2(x).$$

4

(ii). *For any constant $c > 1$, we have*

$$\mathrm{E}[\mathcal{R}] = O\left(\frac{\log m}{mb}\right).$$

Lemma 2 immediately yields the following theorem.

Theorem 3 *For any constant $c > 1$, the bias of the estimator (1) is given by*

$$\mathrm{Bias}[\widehat{g}_b^{(B)}(x)] = bB(x) + o(b) + O\left(\frac{\log m}{mb}\right).$$

Theorem 3 means that the estimator (1) is free of boundary bias.

Next, in the spirit of Ćwik and Mielniczuk (1989), we rewrite $\mathcal{I}_1$ and $\mathcal{I}_2$, as follows;

$$\mathcal{I}_1 = \int \{\widehat{F}_{Y,m}(s) - F_Y(s)\}\mathcal{K}_1(F_Y(s), F_Y(x))d\{\widehat{F}_{X,n}(s) - F_X(s)\}$$

$$+ \int \{g(s) - g(x)\}\{\widehat{F}_{Y,m}(s) - F_Y(s)\}\mathcal{K}_1(F_Y(s), F_Y(x))dF_Y(s)$$

$$+ g(x)\int \{\widehat{F}_{Y,m}(s) - F_Y(s)\}\mathcal{K}_1(F_Y(s), F_Y(x))dF_Y(s)$$

$$= \mathcal{I}_{11} + \mathcal{I}_{12} + \mathcal{I}_{13} \quad \text{(say)},$$

$$\mathcal{I}_2 = \{\widehat{F}_{Y,m}(x) - F_Y(x)\}\int \mathcal{K}_2(F_Y(s), F_Y(x))d\{\widehat{F}_{X,n}(s) - F_X(s)\}$$

$$+ \{\widehat{F}_{Y,m}(x) - F_Y(x)\}\int \{g(s) - g(x)\}\mathcal{K}_2(F_Y(s), F_Y(x))dF_Y(s)$$

$$+ g(x)\{\widehat{F}_{Y,m}(x) - F_Y(x)\}\int \mathcal{K}_2(F_Y(s), F_Y(x))dF_Y(s)$$

$$= \mathcal{I}_{21} + \mathcal{I}_{22} + \mathcal{I}_{23} \quad \text{(say)},$$

where

$$\mathcal{I}_{13} = -g(x)\left\{\frac{1}{m}\sum_{j=1}^{m}\mathcal{K}(F_Y(Y_j), F_Y(x)) - \mathrm{E}[\mathcal{K}(F_Y(Y_1), F_Y(x))]\right\},$$

$$\mathcal{I}_{23} = g(x)\{\widehat{F}_{Y,m}(x) - F_Y(x)\}\int_0^1 \mathcal{K}_2(u, F_Y(x))du = 0,$$

since $\int_0^1 \mathcal{K}_2(u, t)du = 0$ for any $t \in [0, 1]$. Note that $\mathcal{I}_{13}$ is independent of $\widetilde{g}_b^{(B)}(x)$. The following lemma is important for deriving the variance of the estimator (1).

Lemma 4 (i). *We have*

$$\mathrm{Var}[\widetilde{g}_b^{(B)}(x)] = \begin{cases} \dfrac{g(x)V(x)}{nb^{1/2}}\{1 + o(1)\} + O\left(\dfrac{1}{n}\right), & x \in \mathcal{S}_{I,b}, \\[2ex] \dfrac{g(x)v(\kappa)}{nb}\{1 + o(1)\} + O\left(\dfrac{1}{n}\right), & x \in \mathcal{S}_{B,b}. \end{cases}$$

where

$$V(x) = 2^{-1}\{\pi F_Y(x)(1 - F_Y(x))\}^{-1/2}, \quad v(\kappa) = \frac{\Gamma(2\kappa + 2c + 1)}{2^{2\kappa+2c+1}\Gamma^2(\kappa + c + 1)}.$$

5

(ii). *We have*

$$
\mathrm{E}[\mathcal{I}_{13}^2] = \mathrm{Var}[\mathcal{I}_{13}] =
\begin{cases}
\dfrac{g^2(x)V(x)}{mb^{1/2}}\{1 + o(1)\} + O\!\left(\dfrac{1}{m}\right), & x \in \mathcal{S}_{I,b}, \\[2.5ex]
\dfrac{g^2(x)v(\kappa)}{mb}\{1 + o(1)\} + O\!\left(\dfrac{1}{m}\right), & x \in \mathcal{S}_{B,b}.
\end{cases}
$$

(iii). *We have*

$$
\mathrm{E}[\mathcal{I}_{11}^2] = O\!\left(\frac{1}{nmb^2}\right), \quad \mathrm{E}[\mathcal{I}_{12}^2] = O\!\left(\frac{1}{m}\right), \quad \mathrm{E}[\mathcal{I}_{21}^2] = O\!\left(\frac{1}{nmb^2}\right), \quad \mathrm{E}[\mathcal{I}_{22}^2] = O\!\left(\frac{1}{m}\right).
$$

(iv). *For any constant $c > 1$, we have*

$$
\mathrm{E}[\mathcal{R}^2] =
\begin{cases}
O\!\left(\dfrac{(\log m)^2}{m^2 b^{5/2}\{F_Y(x)(1 - F_Y(x))\}^{1/2}}\right), & x \in \mathcal{S}_{I,b}, \\[2.5ex]
O\!\left(\dfrac{\{\log m\}^2}{m^2 b^3}\right), & x \in \mathcal{S}_{B,b}.
\end{cases}
$$

Lemma 4 yields the following theorem.

Theorem 5 *For any constant $c > 1$, the variance of the estimator (1) is given by*

$$
\mathrm{Var}[\widehat{g}_b^{(B)}(x)] =
\begin{cases}
\left(\dfrac{1}{nb^{1/2}} + \dfrac{g(x)}{mb^{1/2}}\right)g(x)V(x)\{1 + o(1)\} + O\!\left(\dfrac{1}{n} + \dfrac{1}{m}\right), & x \in \mathcal{S}_{I,b}, \\[2.5ex]
\left(\dfrac{1}{nb} + \dfrac{g(x)}{mb}\right)g(x)v(\kappa)\{1 + o(1)\} + O\!\left(\dfrac{1}{n} + \dfrac{1}{m}\right), & x \in \mathcal{S}_{B,b}.
\end{cases}
$$

The asymptotic normality of the estimator (1) is derived, using the following lemma.

Lemma 6 *We have*

$$
(nb^{1/2})^{1/2}\{\widetilde{g}_b^{(B)}(x) - \mathrm{E}[\widetilde{g}_b^{(B)}(x)]\} \xrightarrow{d} N(0, g(x)V(x)) \quad \text{for fixed } x \in \mathcal{S}_I,
$$
$$
(nb)^{1/2}\{\widetilde{g}_b^{(B)}(x) - \mathrm{E}[\widetilde{g}_b^{(B)}(x)]\} \xrightarrow{d} N(0, g(x)v(0)) \quad \text{for } x \in \mathcal{S}_B,
$$
$$
(mb^{1/2})^{1/2}\mathcal{I}_{13} \xrightarrow{d} N(0, g^2(x)V(x)) \quad \text{for fixed } x \in \mathcal{S}_I,
$$
$$
(mb)^{1/2}\mathcal{I}_{13} \xrightarrow{d} N(0, g^2(x)v(0)) \quad \text{for } x \in \mathcal{S}_B.
$$

Theorem 7 *If $m = \lceil Cn \rceil$ for some constant $C > 0$, then,*

$$
(nb^{1/2})^{1/2}\{\widehat{g}_b^{(B)}(x) - \mathrm{E}[\widehat{g}_b^{(B)}(x)]\} \xrightarrow{d} N(0, (1 + C^{-1}g(x))g(x)V(x)) \quad \text{for fixed } x \in \mathcal{S}_I,
$$
$$
(nb)^{1/2}\{\widehat{g}_b^{(B)}(x) - \mathrm{E}[\widehat{g}_b^{(B)}(x)]\} \xrightarrow{d} N(0, (1 + C^{-1}g(x))g(x)v(0)) \quad \text{for } x \in \mathcal{S}_B.
$$

From Theorems 3 and 5, the mean squared error (MSE) of the estimator (1) is given by

$$
\mathrm{MSE}[\widehat{g}_b^{(B)}(x)] =
\begin{cases}
\mathrm{AMSE}[\widehat{g}_b^{(B)}(x)] + o\!\left(b^2 + \dfrac{1}{nb^{1/2}} + \dfrac{1}{mb^{1/2}}\right) & \text{for fixed } x \in \mathcal{S}_I, \\[2.5ex]
\mathrm{AMSE}[\widehat{g}_b^{(B)}(x)] + o\!\left(b^2 + \dfrac{1}{nb} + \dfrac{1}{mb}\right) & \text{for } x \in \mathcal{S}_B,
\end{cases}
$$

6

where

$$\text{AMSE}[\widehat{g}_b^{(B)}(x)] = \begin{cases} b^2\{B(x)\}^2 + \left(\dfrac{1}{nb^{1/2}} + \dfrac{g(x)}{mb^{1/2}}\right)g(x)V(x) & \text{for fixed } x \in \mathcal{S}_I, \\ b^2\{B(x)\}^2 + \left(\dfrac{1}{nb} + \dfrac{g(x)}{mb}\right)g(x)v(0) & \text{for } x \in \mathcal{S}_B. \end{cases}$$

It follows that $\widehat{g}_b^{(B)}(x)$ converges in mean square to $g(x)$, hence, $\widehat{g}_b^{(B)}(x) \overset{p}{\to} g(x)$. Note that

$$\min_{b>0}\text{AMSE}[\widehat{g}_b^{(B)}(x)] = \begin{cases} \dfrac{5}{4^{4/5}}\{B(x)\}^{2/5}\{V(x)g(x)(1+g(x))\}^{4/5}n^{-4/5} & \text{for fixed } x \in \mathcal{S}_I, \\ \dfrac{3}{2^{2/3}}\{B(x)v(0)g(x)(1+g(x))\}^{2/3}n^{-2/3} & \text{for } x \in \mathcal{S}_B, \end{cases}$$

which means the MSE of the estimator (1) for $x \in \mathcal{S}_B$ is slower than that for $x \in \mathcal{S}_I$. However, such a different rate phenomenon has a negligible impact of the weighted mean integrated squared error (MISE); $\text{MISE}_{f_Y}[\widehat{g}_b^{(B)}] = \int_{\mathcal{S}}\{\widehat{g}_b^{(B)}(x) - g(x)\}^2 f_Y(x)dx$, as follows.

Theorem 8 *the weighted MISE of the estimator* (1) *is given by*

$$\text{MISE}_{f_Y}[\widehat{g}_b^{(B)}] = \text{AMISE}_{f_Y}[\widehat{g}_b^{(B)}] + o\left(b^2 + \frac{1}{nb^{1/2}} + \frac{1}{mb^{1/2}}\right),$$

where

$$\text{AMISE}_{f_Y}[\widehat{g}_b^{(B)}] = b^2 \int_{\mathcal{S}} B^2(x)f_Y(x)dx + b^{-1/2}\int_{\mathcal{S}}\left(\frac{1}{n} + \frac{g(x)}{m}\right)V(x)f_X(x)dx. \qquad (2)$$

The AMISE (2) is minimized by

$$b = \left\{\frac{\int_{\mathcal{S}}\left(\frac{1}{n} + \frac{g(x)}{m}\right)V(x)f_X(x)dx}{4\int_{\mathcal{S}} B^2(x)f_Y(x)dx}\right\}^{2/5},$$

i.e.,

$$\min_{b>0}\text{AMISE}_{f_Y}[\widehat{g}_b^{(B)}] = \frac{5}{4^{4/5}}\left\{\int_{\mathcal{S}} B^2(x)f_Y(x)dx\right\}^{1/5}\left\{\int_{\mathcal{S}}\left(\frac{1}{n} + \frac{g(x)}{m}\right)V(x)f_X(x)dx\right\}^{4/5}$$

$$= O(n^{-4/5} + m^{-4/5}).$$

The optimal convergence rate $O(n^{-4/5} + m^{-4/5})$ is equal to that of the AMISE of $\widehat{g}_h^{(K_R)}$ (Ćwik and Mielniczuk (1993)), when $m = n$.

4. Simulation studies

In this section, by simulation (with 1000 repetitions), the finite sample performance of the beta KDRE $\widehat{g}_b^{(B)}$ with $c = 1.001, 2, 3$ was compared with other KDREs $\widehat{g}_h^{(K)}$ and $\widehat{g}_h^{(K_R)}$. Each smoothing parameter b (or h) was so selected as to minimize the objective function of the least squares cross-validation

$$\text{LSCV}(b) = \frac{1}{m}\sum_{l=1}^{m}\{\widehat{g}_{b,-Y_l}^{(B)}(Y_l)\}^2 - \frac{2}{n}\sum_{k=1}^{n}\widehat{g}_{b,-X_k}^{(B)}(X_k),$$

where

$$\widehat{g}^{(B)}_{b,-X_k}(x) = \frac{1}{n-1}\sum_{i\neq k}^{n}\mathcal{K}(\widehat{F}_{Y,m}(X_i),\widehat{F}_{Y,m}(x)), \quad \widehat{g}^{(B)}_{b,-Y_l}(x) = \frac{1}{n}\sum_{i=1}^{n}\mathcal{K}(\widehat{F}_{Y,m,-Y_l}(X_i),\widehat{F}_{Y,m,-Y_l}(x))$$

are leave-one-out estimators with

$$\widehat{F}_{Y,m,-Y_l}(x) = \frac{1}{m-1}\sum_{j\neq l}^{m}\chi_{\{Y_j\leq x\}}$$

(For choosing h of $\widehat{g}^{(K)}_h$ or $\widehat{g}^{(K_R)}_h$, $\widehat{g}^{(B)}_{b,-X_k}$ and $\widehat{g}^{(B)}_{b,-Y_l}$ are replaced with the corresponding leave-one-out estimators). The weighted integrated squared error (ISE); $\int_{\mathcal{S}}\{\widehat{g}_b(x)-g(x)\}^2 f_Y(x)dx$ was computed for each simulated sample of $n = 100, 200, 300$ according to three density ratios $g(x) = f_X(x)/f_Y(x)$:

A. $f_X(x) = \dfrac{1}{\sqrt{2\pi}}\exp\left\{-\dfrac{(x+0.5)^2}{2}\right\}$,

  $f_Y(x) = \dfrac{1}{2}\left[\dfrac{1}{\sqrt{2\pi}}\exp\left\{-\dfrac{(x+0.5)^2}{2}\right\} + \dfrac{1}{\sqrt{2\pi(0.8)^2}}\exp\left\{-\dfrac{(x-1)^2}{2(0.8)^2}\right\}\right], \quad -\infty < x < \infty,$

B. $f_X(x) = \dfrac{1}{4}\dfrac{x^{0.2}e^{-x/(0.5)}}{(0.5)^{1.2}\Gamma(1.2)} + \dfrac{3}{4}\dfrac{x^4 e^{-x/2}}{2^5\Gamma(5)}, \quad f_Y(x) = \dfrac{1}{2}\left[\dfrac{xe^{-x/(0.5)}}{(0.5)^2\Gamma(2)} + \dfrac{x^4 e^{-x/2}}{2^5\Gamma(5)}\right], \quad 0 \leq x < \infty,$

C. $f_X(x) = \dfrac{x^7(1-x)}{B(8,2)}, \quad f_Y(x) = \dfrac{1}{5}\dfrac{x^7(1-x)}{B(8,2)} + \dfrac{4}{5}\dfrac{x^6(1-x)^{12}}{B(7,13)}, \quad 0 \leq x \leq 1.$

Table 1 shows that the average weighted ISEs decreased, as the sample size n increased. In all cases A–C, the average weighted ISEs of the beta KDRE were smaller, as c was closer to one, except for the small sample size n in the cases A and C. Hence, we pay attention to the comparison among the beta KDRE with $c = 1.001$ and other KDREs. In the case A, the beta KDRE was worse than $\widehat{g}^{(K_R)}_h$ (better than $\widehat{g}^{(K)}_h$). However, in the cases B and C, the beta KDRE outperformed the others when $n = 300$. Therefore, the beta KDRE was comparable to $\widehat{g}^{(K_R)}_h$.

5. Conclusion

The beta KDRE has been proposed and studied. It has been shown that the beta KDRE, different from the KDRE proposed by Ćwik and Mielniczuk (1989), is free of boundary bias, and that the (weighted) MISE of the beta KDRE achieves $O(n^{-4/5}+m^{-4/5})$, choosing optimal smoothing parameter b. Furthermore, through the simulation studies, the finite sample performance of the beta KDRE has been illustrated.

Acknowledgements

Table 1: Average weighted ISEs.

The bold-faced number indicates the smallest average weighted ISE in each row.

	n	$\widehat{g}_b^{(B)}(x)$			$\widehat{g}_h^{(K_R)}(x)$	$\widehat{g}_h^{(K)}(x)$
		$c = 1.001$	$c = 2$	$c = 3$		
A	100	0.1115	0.1015	0.1050	**0.0656**	0.1334
	200	0.0577	0.0562	0.0582	**0.0332**	0.0879
	300	0.0416	0.0437	0.0444	**0.0275**	0.0724
B	100	0.2863	0.2937	0.2973	**0.2806**	0.2967
	200	**0.2239**	0.2321	0.2386	0.2250	0.2492
	300	**0.1958**	0.2020	0.2079	0.2002	0.2255
C	100	0.7868	0.7846	0.7305	0.8792	**0.6930**
	200	0.4529	0.4564	0.4716	0.5603	**0.4392**
	300	**0.2905**	0.3375	0.3484	0.3115	0.3191

Appendix. Proofs

Let $\psi(z) = \Gamma'(z)/\Gamma(z)$ be the digamma function. Note that

- $\psi(z+1) = \psi(z) + 1/z$,

- $-1/2 \le z\{\log(z) - \psi(z+1)\} \le 0$ for $z > 0$ (Theorem 3.1 of Anderson et al. (1995)), and

- $\dfrac{(k-1)!}{z^k} - \dfrac{k!}{2z^{k+1}} < (-1)^{k+1}\psi^{(k)}(z+1) < \dfrac{(k-1)!}{z^k}$ for $z > 0$ and $k \in \mathbb{N}$ (Lemma 3 of Qi et al. (2010) and $\psi^{(k)}(z+1) = \psi^{(k)}(z) + (-1)^k k!/z^{k+1}$)

will be repeatedly used in the proofs below. Also, we have

$$\mathcal{K}_1(u,t) = \mathcal{K}(u,t)\frac{1}{b}\left(\frac{t+bc}{u} - \frac{1-t+bc}{1-u}\right)$$

$$\mathcal{K}_2(u,t) = \mathcal{K}(u,t)\frac{1}{b}\left\{\log(u) - \psi\left(\frac{t}{b}+c+1\right) - \log(1-u) + \psi\left(\frac{1-t}{b}+c+1\right)\right\}$$

$$\mathcal{K}_{1,1}(u,t) = \mathcal{K}(u,t)\frac{1}{b^2}\left\{\left(\frac{t+bc}{u} - \frac{1-t+bc}{1-u}\right)^2 - \frac{b(t+bc)}{u^2} - \frac{b(1-t+bc)}{(1-u)^2}\right\}$$

$$\mathcal{K}_{1,2}(u,t) = \mathcal{K}(u,t)\frac{1}{b^2}\left[\left(\frac{t+bc}{u} - \frac{1-t+bc}{1-u}\right)\left\{\log(u) - \psi\left(\frac{t}{b}+c+1\right) - \log(1-u) + \psi\left(\frac{1-t}{b}+c+1\right)\right\}\right.$$
$$\left. +\frac{b}{u(1-u)}\right]$$

$$\mathcal{K}_{2,2}(u,t) = \mathcal{K}(u,t)\frac{1}{b^2}\left[\left\{\log(u) - \psi\left(\frac{t}{b}+c+1\right) - \log(1-u) + \psi\left(\frac{1-t}{b}+c+1\right)\right\}^2\right.$$
$$\left. -\left\{\psi'\left(\frac{t}{b}+c+1\right) + \psi'\left(\frac{1-t}{b}+c+1\right)\right\}\right],$$

where

$$\int_0^1 \mathcal{K}(u,t)du = 1, \quad \int_0^1 \mathcal{K}_k(u,t)du = 0, \quad \int_0^1 \mathcal{K}_{k,k}(u,t)du = 0 \quad \text{for } k = 1, 2.$$

Lemma A.1 *For any $t \in [0, 1]$, we have*

$$\int (u - t)\mathcal{K}(u, t)du = b(c + 1)(1 - 2t) + O(b^2),$$

$$\int (u - t)^2 \mathcal{K}(u, t)du = bt(1 - t) + O(b^2),$$

$$\int (u - t)^4 \mathcal{K}(u, t)du = O(b^2).$$

Proof Use

$$\int u^k K_{b,c,c}^{(B)}(u, t)du = \frac{\{t + (c + k)b\} \cdots \{t + (c + 1)b\}}{\{1 + (2c + k + 1)b\} \cdots \{1 + (2c + 2)b\}}$$

for any natural number k. □

Lemma A.2 (i) and (ii) are slight generalization of Lemma A.3 (i) and (ii) of Igarashi (2016).

Lemma A.2   (i). *For any constants $c_1, c_2 \geq 0$, we have*

$$\sup_{t \in [0,1]} \sup_{u \in [0,1]} K_{b,c_1,c_2}^{(B)}(u, t) \leq \sup_{t \in [-bc_1, 1+bc_2]} \sup_{u \in [0,1]} K_{b,c_1,c_2}^{(B)}(u, t)$$
$$= b^{-1}\{1 + b(c_1 + c_2 + 1)\}.$$

(ii). *For any $t \in (0, 1)$ and constants $c_1, c_2 \geq 0$, we have*

$$\sup_{u \in [0,1]} K_{b,c_1,c_2}^{(B)}(u, t) \leq \frac{b^{-1/2}(1 + b(c_1 + c_2 + 1))(1 + b(c_1 + c_2))^{1/2}}{\{2\pi(t + bc_1)(1 - t + bc_2)\}^{1/2}}.$$

(iii). *For any constant $c > 0$, we have*

$$\sup_{t \in [0,1]} \sup_{u \in [0,1]} |\mathcal{K}_1(u, t)| \leq b^{-2}\{1 + b(2c + 1)\}(1 + 2bc).$$

(iv). *For any constant $c > 0$, we have*

$$\sup_{t \in [0,1]} \sup_{u \in [0,1]} |\mathcal{K}_2(u, t)| \leq b^{-2}\{1 + b(2c + 1)\}\left\{2\left(1 + \frac{1}{c}\right) + \frac{1 + 2bc}{2c(1 + bc)}\right\}.$$

(v). *For any constant $c > 1$, we have*

$$\sup_{t \in [0,1]} \sup_{u \in [0,1]} |\mathcal{K}_{1,1}(u, t)| \leq 2b^{-3}(1 + 2bc)\{1 + b(2c + 1)\}\{1 + b(2c - 1)\}.$$

(vi). *For any constant $c > 1$, we have*

$$\sup_{t \in [0,1]} \sup_{u \in [0,1]} |\mathcal{K}_{1,2}(u, t)|$$
$$\leq b^{-3}\{1 + b(2c + 1)\}\left[2\left\{1 + \frac{1 + 2bc}{2c(1 + bc)}\right\}(1 + 2bc) + 2\left(1 + \frac{1}{c - 1}\right)\left(1 + \frac{1}{c}\right)\{1 + b(2c - 1)\}\right.$$
$$\left. + \frac{1 + 2bc}{c(1 + bc)}\{1 + b(2c - 1)\}\right],$$

10

(vii). *For any constant $c > 1$, we have*

$$\sup_{t\in[0,1]} \sup_{u\in[0,1]} |\mathcal{K}_{2,2}(u,t)|$$

$$\leq b^{-3}\left\{6\left(1+\frac{1}{c}\right)\left(1+\frac{2}{c}\right)\left(1+\frac{1}{(c-1)}\right) + \frac{3(1+2bc)^2}{4c^2(1+bc)^2} + 2\psi'(c+1)\right\}\{1+b(2c+1)\},$$

Proof Note that, for any constants $c_1, c_2 \geq 0$,

$$\sup_{u\in[0,1]} K^{(B)}_{b,c_1,c_2}(u,t) = K^{(B)}_{b,c_1,c_2}\left(\frac{t+bc_1}{1+b(c_1+c_2)}, t\right),$$

$pB(p,q) = (p+q)B(p+1,q)$, and $|\log(u)| \leq u + u^{-1}$ for $u > 0$. Also, we have

$$\left|\log(u) - \psi\left(\frac{t}{b}+c+1\right) - \log(1-u) + \psi\left(\frac{1-t}{b}+c+1\right)\right|$$

$$= \left|\log\left\{\frac{u(1-t+bc)}{(1-u)(t+bc)}\right\} + \log\left(\frac{t}{b}+c\right) - \psi\left(\frac{t}{b}+c+1\right) - \log\left(\frac{1-t}{b}+c\right) + \psi\left(\frac{1-t}{b}+c+1\right)\right|$$

$$\leq \frac{u(1-t+bc)}{(1-u)(t+bc)} + \frac{(1-u)(t+bc)}{u(1-t+bc)} + \frac{b}{2(t+bc)} + \frac{b}{2(1-t+bc)}$$

$$\leq \frac{u(1-t+bc)}{(1-u)(t+bc)} + \frac{(1-u)(t+bc)}{u(1-t+bc)} + \frac{1+2bc}{2c(1+bc)}.$$

It is well-known that $\psi(z)$ is strictly increasing and concave for $z > 0$ (hence, $\psi'(z)$ is positive and strictly decreasing).

(i). Theorem 3.2 (2) of Anderson et al. (1995) enables us to see that

$$\sup_{u\in[0,1]} K^{(B)}_{b,c_1,c_2}(u,t)$$

$$= \frac{e^{1/b+c_1+c_2}\Gamma(1/b+c_1+c_2+2)}{(1/b+c_1+c_2)^{1/b+c_1+c_2}} \frac{(t/b+c_1)^{t/b+c_1}}{e^{t/b+c_1}\Gamma(t/b+c_1+1)} \frac{((1-t)/b+c_2)^{(1-t)/b+c_2}}{e^{(1-t)/b+c_2}\Gamma((1-t)/b+c_2+1)}$$

is strictly log-convex for $t \in [-bc_1, 1+bc_2]$. Hence,

$$\sup_{t\in[-bc_1,1+bc_2]} \sup_{u\in[0,1]} K^{(B)}_{b,c_1,c_2}(u,t) = b^{-1}\{1+b(c_1+c_2+1)\}.$$

(ii). Note that $R(z) = \sqrt{2\pi}z^{z+1/2}e^{-z}/\Gamma(z+1) < 1$ is strictly increasing for $z > 0$ (see Theorem 3.2 (1) of Anderson et al. (1995)). We have

$$\sup_{u\in[0,1]} K^{(B)}_{b,c_1,c_2}(u,t)$$

$$= \frac{b^{-1/2}(1+b(c_1+c_2+1))(1+b(c_1+c_2))^{1/2}R(t/b+c_1)R((1-t)/b+c_2)}{\{2\pi(t+bc_1)(1-t+bc_2)\}^{1/2}R(1/b+c_1+c_2)}$$

$$\leq \frac{b^{-1/2}(1+b(c_1+c_2+1))(1+b(c_1+c_2))^{1/2}}{\{2\pi(t+bc_1)(1-t+bc_2)\}^{1/2}}.$$

11

(iii). Since

$$\mathcal{K}_1(u,t) = \frac{(t/b+c)B(t/b+c,(1-t)/b+c+1)}{B(t/b+c+1,(1-t)/b+c+1)}K_{b,c-1,c}^{(B)}(u,t)$$
$$-\frac{((1-t)/b+c)B(t/b+c+1,(1-t)/b+c)}{B(t/b+c+1,(1-t)/b+c+1)}K_{b,c,c-1}^{(B)}(u,t)$$
$$= \left(\frac{1}{b}+2c+1\right)\{K_{b,c-1,c}^{(B)}(u,t) - K_{b,c,c-1}^{(B)}(u,t)\},$$

the result follows from (i).

(iv). Since

$$|\mathcal{K}_2(u,t)| \leq \mathcal{K}(u,t)\frac{1}{b}\left\{\frac{u(1-t+bc)}{(1-u)(t+bc)} + \frac{(1-u)(t+bc)}{u(1-t+bc)} + \frac{1+2bc}{2c(1+bc)}\right\}$$
$$= \frac{((1-t)/b+c)B(t/b+c+2,(1-t)/b+c)}{b(t/b+c)B(t/b+c+1,(1-t)/b+c+1)}K_{b,c+1,c-1}^{(B)}(u,t)$$
$$+\frac{(t/b+c)B(t/b+c,(1-t)/b+c+2)}{b((1-t)/b+c)B(t/b+c+1,(1-t)/b+c+1)}K_{b,c-1,c+1}^{(B)}(u,t)$$
$$+\frac{1+2bc}{2bc(1+bc)}\mathcal{K}(u,t)$$
$$= \frac{1}{b}\left(1+\frac{1}{c}\right)\{K_{b,c+1,c-1}^{(B)}(u,t) + K_{b,c-1,c+1}^{(B)}(u,t)\} + \frac{1+2bc}{2bc(1+bc)}\mathcal{K}(u,t),$$

the result follows from (i).

(v). Since

$$\mathcal{K}_{1,1}(u,t)$$
$$= \mathcal{K}(u,t)\left\{\frac{(t/b+c)(t/b+c-1)}{u^2} + \frac{((1-t)/b+c)((1-t)/b+c-1)}{(1-u)^2}\right.$$
$$\left.-2\frac{(t/b+c)((1-t)/b+c)}{u(1-u)}\right\}$$
$$= \frac{(t/b+c)(t/b+c-1)B(t/b+c-1,(1-t)/b+c+1)}{B(t/b+c+1,(1-t)/b+c+1)}K_{b,c-2,c}^{(B)}(u,t)$$
$$+\frac{((1-t)/b+c)((1-t)/b+c-1)B(t/b+c+1,(1-t)/b+c-1)}{B(t/b+c+1,(1-t)/b+c+1)}K_{b,c,c-2}^{(B)}(u,t)$$
$$-2\frac{(t/b+c)((1-t)/b+c)B(t/b+c,(1-t)/b+c)}{B(t/b+c+1,(1-t)/b+c+1)}K_{b,c-1,c-1}^{(B)}(u,t)$$
$$= \left(\frac{1}{b}+2c\right)\left(\frac{1}{b}+2c+1\right)\{K_{b,c-2,c}^{(B)}(u,t) + K_{b,c,c-2}^{(B)}(u,t) - 2K_{b,c-1,c-1}^{(B)}(u,t)\},$$

the result follows from (i).

(vi). Since

$$|\mathcal{K}_{1,2}(u,t)|$$
$$\leq \mathcal{K}(u,t)\frac{1}{b}\left[\left(\frac{t/b+c}{u} + \frac{(1-t)/b+c}{1-u}\right)\left\{\frac{u((1-t)/b+c)}{(1-u)(t/b+c)} + \frac{(1-u)(t/b+c)}{u((1-t)/b+c)} + \frac{1+2bc}{2c(1+bc)}\right\}\right.$$

12

$$+\frac{1}{u(1-u)}\Bigg]$$

$$= \mathcal{K}(u,t)\frac{1}{b}\Bigg[\Big\{\frac{t/b+c}{u}+\frac{(1-t)/b+c}{1-u}\Big\}\Big\{1+\frac{1+2bc}{2c(1+bc)}\Big\}$$

$$+\frac{(1-u)(t/b+c)^2}{u^2((1-t)/b+c)}+\frac{u((1-t)/b+c)^2}{(1-u)^2(t/b+c)}+\frac{1}{u(1-u)}\Bigg]$$

$$= \frac{1}{b}\Big(\frac{1}{b}+2c+1\Big)\Bigg[\Big\{1+\frac{1+2bc}{2c(1+bc)}\Big\}\{K^{(B)}_{b,c-1,c}(u,t)+K^{(B)}_{b,c,c-1}(u,t)\}$$

$$+\Big(1+\frac{b}{t+b(c-1)}\Big)\Big(1+\frac{b}{1-t+bc}\Big)K^{(B)}_{b,c-2,c+1}(u,t)$$

$$+\Big(1+\frac{b}{t+bc}\Big)\Big(1+\frac{b}{1-t+b(c-1)}\Big)K^{(B)}_{b,c+1,c-2}(u,t)$$

$$+\frac{b(1+2bc)}{(t+bc)(1-t+bc)}K^{(B)}_{b,c-1,c-1}(u,t)\Bigg]$$

$$\leq \frac{1}{b}\Big(\frac{1}{b}+2c+1\Big)\Bigg[\Big\{1+\frac{1+2bc}{2c(1+bc)}\Big\}\{K^{(B)}_{b,c-1,c}(u,t)+K^{(B)}_{b,c,c-1}(u,t)\}$$

$$+\Big(1+\frac{1}{c-1}\Big)\Big(1+\frac{1}{c}\Big)\{K^{(B)}_{b,c-2,c+1}(u,t)+K^{(B)}_{b,c+1,c-2}(u,t)\}$$

$$+\frac{1+2bc}{c(1+bc)}K^{(B)}_{b,c-1,c-1}(u,t)\Bigg],$$

the result follows from (i).

(vii). Since

$$|\mathcal{K}_{2,2}(u,t)|$$

$$\leq \mathcal{K}(u,t)\frac{1}{b^2}\Bigg[3\Big\{\frac{u^2((1-t)/b+c)^2}{(1-u)^2(t/b+c)^2}+\frac{(1-u)^2(t/b+c)^2}{u^2((1-t)/b+c)^2}+\frac{(1+2bc)^2}{4c^2(1+bc)^2}\Big\}+2\psi'(c+1)\Bigg]$$

$$= \frac{1}{b^2}\Bigg[3\Big\{\Big(1+\frac{b}{t+bc}\Big)\Big(1+\frac{2b}{t+bc}\Big)\Big(1+\frac{b}{1-t+b(c-1)}\Big)K^{(B)}_{b,c+2,c-2}(u,t)$$

$$+\Big(1+\frac{b}{t+b(c-1)}\Big)\Big(1+\frac{b}{1-t+bc}\Big)\Big(1+\frac{2b}{1-t+bc}\Big)K^{(B)}_{b,c-2,c+2}(u,t)\Big\}$$

$$+\Big\{\frac{3(1+2bc)^2}{4c^2(1+bc)^2}+2\psi'(c+1)\Big\}\mathcal{K}(u,t)\Bigg]$$

$$\leq \frac{1}{b^2}\Bigg[3\Big(1+\frac{1}{c}\Big)\Big(1+\frac{2}{c}\Big)\Big(1+\frac{1}{(c-1)}\Big)\{K^{(B)}_{b,c+2,c-2}(u,t)+K^{(B)}_{b,c-2,c+2}(u,t)\}$$

$$+\Big\{\frac{3(1+2bc)^2}{4c^2(1+bc)^2}+2\psi'(c+1)\Big\}\mathcal{K}(u,t)\Bigg],$$

the result follows from (i). $\square$

Lemma A.3 *Let*

$$t_m(u)=\Bigg[\Big\{u(1-u)+\frac{M\log m}{m}\Big\}\frac{M\log m}{m}\Bigg]^{1/2}$$

for any $u \in [0,1]$ and $M > 0$.

(i). *We have*

$$\Pr(|\widehat{F}_{Y,m}(s) - F_Y(s)| \geq t_m(F_Y(s))) \leq \begin{cases} \dfrac{2}{m^{3M/8}}, & 0 < F_Y(s) < 1, \\ 0, & F_Y(s) = 0, 1. \end{cases}$$

(ii). *We have*

$$\Pr(|\widehat{F}_{Y,m}(X_i) - F_Y(X_i)| \geq t_m(F_Y(X_i))|X_i) \leq \begin{cases} \dfrac{2}{m^{3M/8}}, & 0 < F_Y(X_i) < 1, \\ 0, & F_Y(X_i) = 0, 1. \end{cases}$$

Proof If $0 < F_Y(s) < 1$, Bennett's inequality enables us to see that

$$\Pr(|\widehat{F}_{Y,m}(s) - F_Y(s)| \geq t) \leq 2\exp\left[-\frac{mt_m^2(F_Y(s))}{2\{F_Y(s)(1 - F_Y(s)) + t_m(F_Y(s))/3\}}\right]$$

$$= 2\exp\left[-\frac{M\log m}{2\left\{\frac{F_Y(s)(1-F_Y(s))}{F_Y(s)(1-F_Y(s))+M(\log m)/m} + \frac{1}{3}\sqrt{\frac{M(\log m)/m}{F_Y(s)(1-F_Y(s))+M(\log m)/m}}\right\}}\right]$$

$$\leq 2\exp\left(-\frac{3M}{8}\log m\right).$$

In the same way,

$$\Pr(|\widehat{F}_{Y,m}(X_i) - F_Y(X_i)| \geq t|X_i) \leq 2\exp\left(-\frac{3M}{8}\log m\right)$$

is derived using Bennett's inequality with a slight modification. Also, $F_Y(s) = 0, 1$ means $\widehat{F}_{Y,m}(s) - F_Y(s) = 0$. $\square$

Proof of Lemma 2 (i). Using Lemma A.1, we have

$$\mathrm{E}[\widetilde{g}_b^{(B)}(x)] = \int_{\mathcal{S}} \mathcal{K}(F_Y(s), F_Y(x))f_X(s)ds$$

$$= \int_0^1 \mathcal{K}(u, F_Y(x))h(u)du$$

$$= h(F_Y(x)) + h'(F_Y(x))\int_0^1 \mathcal{K}(u, F_Y(x))(u - F_Y(x))du$$

$$+ \frac{1}{2}h''(F_Y(x))\int_0^1 \mathcal{K}(u, F_Y(x))(u - F_Y(x))^2du + R_h(F_Y(x))$$

$$= g(x) + bB(x) + R_h(F_Y(x)) + O(b^2),$$

where $h(u) = g(F_Y^{-1}(u))$ and

$$R_h(F_Y(x)) = \int_0^1 \mathcal{K}(u, F_Y(x))(u - F_Y(x))^2 \int_0^1 \{h''(F_Y(x) + \theta(u - F_Y(x))) - h''(F_Y(x))\}(1-\theta)d\theta du.$$

Noting that h'' is continuous on $[0, 1]$, for any $\epsilon > 0$, there exists a $\delta > 0$ such that if $|F_Y(x) + \theta(u - F_Y(x)) - F_Y(x)| \leq |u - F_Y(x)| < \delta$, then $|h''(F_Y(x) + \theta(u - F_Y(x))) - h''(F_Y(x))| < \epsilon$.

14

Hence,

$$R_h(F_Y(x)) = \frac{\epsilon}{2} \int_{|u-F_Y(x)|<\delta} \mathcal{K}(u, F_Y(x))(u - F_Y(x))^2 du$$

$$+ \frac{\sup_{u\in[0,1]} h''(u)}{\delta^2} \int_{|u-F_Y(x)|\geq\delta} \mathcal{K}(u, F_Y(x))(u - F_Y(x))^4 du$$

$$= O(\epsilon b + b^2).$$

(ii). Let

$$\widetilde{S} = \{Y_1, \ldots, Y_m \mid |\widehat{F}_{Y,m}(x) - F_Y(x)| < t_m(x), |\widehat{F}_{Y,m}(X_i) - F_Y(X_i)| < t_m(X_i)\},$$

where

$$t_m(s) = \left[\left\{F_Y(s)(1 - F_Y(s)) + \frac{M\log m}{m}\right\} \frac{M\log m}{m}\right]^{1/2}$$

with some constant $M \geq 16/3$. For simplicity, we write $\Delta(u) = \theta\{\widehat{F}_{Y,m}(F_Y^{-1}(u)) - u\}$. Then, for any $t \in [0,1]$, we have

$$\int_0^1 \left\{u(1-u) + \frac{\log m}{m}\right\} \mathcal{K}_{1,1}(u + \Delta(u), t + \Delta(t)) du$$

$$= \frac{1}{b^2} \int_0^1 \left\{u(1-u) + \frac{\log m}{m}\right\} \mathcal{K}(u, t + \Delta(t)) \left\{1 + O\left(\left(\frac{t + bc}{u} + \frac{1 - t + bc}{1 - u}\right)\sqrt{\frac{\log m}{mb^2}}\right)\right\}$$

$$\times \left[\left\{\frac{t + bc + \Delta(t)}{u} - \frac{1 - t + bc - \Delta(t)}{1 - u} + O\left(\left(\frac{t + bc}{u^2} + \frac{1 - t + bc}{(1-u)^2}\right)\sqrt{\frac{\log m}{m}}\right)\right\}^2 \right.$$

$$\left. + \frac{b(t + cb + \Delta(t))}{u^2} + \frac{b(1 - t + cb - \Delta(t))}{(1-u)^2} + O\left(\left(\frac{t + bc}{u^3} + \frac{1 - t + bc}{(1-u)^3}\right)b\sqrt{\frac{\log m}{m}}\right)\right] du$$

$$= O(b^{-1}),$$

$$\int_0^1 \left\{u(1-u) + \frac{\log m}{m}\right\}^{1/2} \mathcal{K}_{1,2}(u + \Delta(u), t + \Delta(t)) du$$

$$= \frac{1}{b^2} \int_0^1 \left\{u(1-u) + \frac{\log m}{m}\right\}^{1/2} \mathcal{K}(u, t) \left\{1 + O\left(\left(\frac{t + bc}{u} + \frac{1 - t + bc}{1 - u}\right)\sqrt{\frac{\log m}{mb^2}}\right)\right\}$$

$$\times \left[\left|\log(u) - \psi\left(\frac{t + \Delta(t)}{b} + c + 1\right) - \log(1 - u) + \psi\left(\frac{1 - t - \Delta(t)}{b} + c + 1\right)\right.\right.$$

$$\left. + O\left(u^{-1}(1 - u)^{-1}\sqrt{\frac{\log m}{m}}\right)\right|$$

$$\times \left|\frac{t + bc + \Delta(t)}{u} - \frac{1 - t + bc - \Delta(t)}{1 - u} + O\left(\left(\frac{t + bc}{u^2} + \frac{1 - t + bc}{(1-u)^2}\right)\sqrt{\frac{\log m}{m}}\right)\right| + \frac{b}{u(1 - u)}\right] du$$

$$= O\left(\frac{b^{-1}}{(t + bc)^{1/2}(1 - t + bc)^{1/2}}\right),$$

$$\int_0^1 \mathcal{K}_{2,2}(u + \Delta(u), t + \Delta(t)) du$$

$$= \frac{1}{b^2} \int_0^1 \mathcal{K}(u, t + \Delta(t)) \left\{1 + O\left(\left(\frac{t + bc}{u} + \frac{1 - t + bc}{1 - u}\right)\sqrt{\frac{\log m}{mb^2}}\right)\right\}$$

$$\times \left[\left\{\log(u) - \psi\left(\frac{t + \Delta(t)}{b} + c + 1\right) - \log(1 - u) + \psi\left(\frac{1 - t - \Delta(t)}{b} + c + 1\right)\right.\right.$$

15

$$+O\left(u^{-1}(1-u)^{-1}\sqrt{\frac{\log m}{m}}\right)\Bigg\}^2 + \psi'\left(\frac{t+\Delta(t)}{b}+c+1\right) + \psi'\left(\frac{1-t-\Delta(t)}{b}+c+1\right)\Bigg]du$$

$$= O\left(\frac{b^{-1}}{(t+bc)(1-t+bc)}\right),$$

noting that $\Delta(t) = O((\log m)/m)$. It follows that

$$\mathrm{E}[|\mathcal{R}_{k,l}|\chi_{\widetilde{S}}] = O\left(\frac{\log m}{mb}\right), \quad k,l = 1,2.$$

Also, using Lemmas A.2 and A.3, we have

$$\mathrm{E}[|\mathcal{R}_{k,l}|\chi_{\widetilde{S}^c}] = o(m^{-1}b^{-1}), \quad k,l = 1,2. \quad \square$$

Proof of Theorem 3 Lemma 2 immediately yields result.

Proof of Lemma 4 (i) and (ii). We have

$$\mathrm{Var}[\widetilde{g}_b^{(B)}(x)] = \frac{1}{n}\mathrm{Var}[\mathcal{K}(F_Y(X_1), F_Y(x))]$$

$$= \frac{1}{n}\int_0^1 \mathcal{K}^2(u, F_Y(x))h(u)du + O(n^{-1})$$

$$= \frac{g(x)}{n}\frac{B(2F_Y(x)/b+2c+1, 2(1-F_Y(x))/b+2c+1)}{B^2(F_Y(x)/b+c+1, (1-F_Y(x))/b+c+1)} + O(n^{-1}),$$

$$\mathrm{Var}[\mathcal{I}_{13}] = \frac{g^2(x)}{m}\mathrm{Var}[\mathcal{K}(F_Y(Y_1), F_Y(x))]$$

$$= \frac{g^2(x)}{m}\int_0^1 \mathcal{K}^2(u, F_Y(x))du + O(m^{-1})$$

$$= \frac{g^2(x)}{m}\frac{B(2F_Y(x)/b+2c+1, 2(1-F_Y(x))/b+2c+1)}{B^2(F_Y(x)/b+c+1, (1-F_Y(x))/b+c+1)} + O(m^{-1}).$$

(iii). Note that

$$\int_0^1 (u-t)^2\mathcal{K}(u,t)du$$
$$= O(bt(1-t) + b^2), \tag{A1}$$

$$\int_0^1 \mathcal{K}(u,t)\left(\frac{t+bc}{u} - \frac{1-t+bc}{1-u}\right)^2 du$$
$$= O\left(\frac{b}{(t+bc)(1-t+bc)}\right), \tag{A2}$$

$$\int_0^1 u(1-u)\mathcal{K}(u,t)\left(\frac{t+bc}{u} - \frac{1-t+bc}{1-u}\right)^2 du$$
$$= O(b), \tag{A3}$$

$$\int_0^1 \mathcal{K}(u,t)\left\{\log(u) - \psi\left(\frac{t}{b}+c+1\right) - \log(1-u) + \psi\left(\frac{1-t}{b}+c+1\right)\right\}^2 du$$
$$= O\left(\psi'\left(\frac{t}{b}+c+1\right) + \psi'\left(\frac{1-t}{b}+c+1\right)\right)$$
$$= O\left(\frac{b}{(t+bc)(1-t+bc)}\right) \tag{A4}$$

16

for any $t \in [0, 1]$. We have

$$
\mathrm{E}[\mathcal{I}_{11}^2] = \mathrm{E}\left[\left[\frac{1}{n}\sum_{i=1}^{n}\left[\{\widehat{F}_{Y,m}(X_i) - F_Y(X_i)\}\mathcal{K}_1(F_Y(X_i), F_Y(x))\right.\right.\right.
$$

$$
\left.\left.\left. -\mathrm{E}[\{\widehat{F}_{Y,m}(X_i) - F_Y(X_i)\}\mathcal{K}_1(F_Y(X_i), F_Y(x))|Y_1, \ldots, Y_m]\right]\right]^2\right]
$$

$$
= \frac{1}{n}\mathrm{E}\left[\mathrm{Var}[\{\widehat{F}_{Y,m}(X_1) - F_Y(X_1)\}\mathcal{K}_1(F_Y(X_1), F_Y(x))|Y_1, \ldots, Y_m]\right]
$$

$$
\leq \frac{1}{n}\mathrm{E}\left[\mathrm{E}[\{\widehat{F}_{Y,m}(X_1) - F_Y(X_1)\}^2\mathcal{K}_1^2(F_Y(X_1), F_Y(x))|Y_1, \ldots, Y_m]\right]
$$

$$
= \frac{1}{n}\mathrm{E}\left[\mathrm{Var}[\widehat{F}_{Y,m}(X_1)|X_1]\mathcal{K}_1^2(F_Y(X_1), F_Y(x))\right]
$$

$$
= \frac{1}{nm}\mathrm{E}[F_Y(X_1)(1 - F_Y(X_1))\mathcal{K}_1^2(F_Y(X_1), F_Y(x))]
$$

$$
= O((nmb^2)^{-1}) \quad \text{(we used (A3))},
$$

$$
\mathrm{E}[\mathcal{I}_{21}^2] = \mathrm{E}\left[\left[\frac{1}{n}\{\widehat{F}_{Y,m}(x) - F_Y(x)\}\sum_{i=1}^{n}\{\mathcal{K}_2(F_Y(X_i), F_Y(x)) - \mathrm{E}[\mathcal{K}_2(F_Y(X_i), F_Y(x))|Y_1, \ldots, Y_m]\}\right]^2\right]
$$

$$
= \frac{1}{n}\mathrm{Var}[\widehat{F}_{Y,m}(x)]\mathrm{Var}[\mathcal{K}_2(F_Y(X_1), F_Y(x))]
$$

$$
= \frac{F_Y(x)(1 - F_Y(x))}{nm}\mathrm{Var}[\mathcal{K}_2(F_Y(X_1), F_Y(x))]
$$

$$
= O((nmb^2)^{-1}) \quad \text{(we used (A4))},
$$

$$
\mathrm{E}[\mathcal{I}_{12}^2] = \mathrm{E}\left[\left[\int_{\mathcal{S}}\{g(s) - g(x)\}\{\widehat{F}_{Y,m}(s) - F_Y(s)\}\mathcal{K}_1(F_Y(s), F_Y(x))f_Y(s)ds\right]^2\right]
$$

$$
\leq \frac{1}{b^2}\mathrm{E}\left[\int_{\mathcal{S}}\{\widehat{F}_{Y,m}(s) - F_Y(s)\}^2\mathcal{K}(F_Y(s), F_Y(x))\left(\frac{F_Y(x) + bc}{F_Y(s)} - \frac{1 - F_Y(x) + bc}{1 - F_Y(s)}\right)^2 f_Y(s)ds\right.
$$

$$
\left. \times \int_{\mathcal{S}}\{g(s) - g(x)\}^2\mathcal{K}(F_Y(s), F_Y(x))f_Y(s)ds\right]
$$

$$
= \frac{1}{mb^2}\left\{\sup_{u \in [0,1]} h'(u)\right\}^2\int_0^1(u - F_Y(x))^2\mathcal{K}(u, F_Y(x))du
$$

$$
\times \int_0^1 u(1 - u)\mathcal{K}(u, F_Y(x))\left(\frac{F_Y(x) + bc}{u} - \frac{1 - F_Y(x) + bc}{1 - u}\right)^2 du
$$

$$
= O(m^{-1}) \quad \text{(we used (A1) and (A3))},
$$

$$
\mathrm{E}[\mathcal{I}_{22}^2] = \mathrm{Var}[\widehat{F}_{Y,m}(x)]\left[\int_{\mathcal{S}}\{g(s) - g(x)\}\mathcal{K}_2(F_Y(s), F_Y(x))f_Y(s)ds\right]^2
$$

$$
\leq \frac{F_Y(x)(1 - F_Y(x))}{mb^2}\left\{\sup_{u \in [0,1]} h'(u)\right\}^2\int_0^1(u - F_Y(x))^2\mathcal{K}(u, F_Y(x))f_Y(s)du
$$

$$
\times \int_0^1\mathcal{K}(u, F_Y(x))\left\{\log(u) - \psi\left(\frac{F_Y(x)}{b} + c + 1\right) - \log(1 - u) + \psi\left(\frac{1 - F_Y(x)}{b} + c + 1\right)\right\}^2 du
$$

$$
= O(m^{-1}) \quad \text{(we used (A1) and (A4))}.
$$

(iv). We can see that

$$\int_0^1 \left\{ u(1-u) + \frac{\log m}{m} \right\}^2 \mathcal{K}_{1,1}^2(u + \Delta(u), t + \Delta(t)) du$$

$$= \frac{1}{b^4} \int_0^1 \left\{ u(1-u) + \frac{\log m}{m} \right\}^2 \mathcal{K}^2(u, t + \Delta(t)) \left\{ 1 + O\left(\left(\frac{t+bc}{u} + \frac{1-t+bc}{1-u} \right) \sqrt{\frac{\log m}{mb^2}} \right) \right\}^2$$

$$\times \left[\left\{ \frac{t+bc+\Delta(t)}{u} - \frac{1-t+bc-\Delta(t)}{1-u} + O\left(\left(\frac{t+bc}{u^2} + \frac{1-t+bc}{(1-u)^2} \right) \sqrt{\frac{\log m}{m}} \right) \right\}^2 \right.$$

$$\left. - \frac{b(t+cb+\Delta(t))}{u^2} - \frac{b(1-t+cb-\Delta(t))}{(1-u)^2} + O\left(\left(\frac{t+bc}{u^3} + \frac{1-t+bc}{(1-u)^3} \right) b\sqrt{\frac{\log m}{m}} \right) \right]^2 du$$

$$= \begin{cases} O\left(\dfrac{b^{-5/2}}{(t+bc)^{1/2}(1-t+bc)^{1/2}} \right), & \dfrac{t}{b} \to \infty, \dfrac{1-t}{b} \to \infty, \\ O(b^{-3}), & \dfrac{t}{b} \to \kappa \text{ or } \dfrac{1-t}{b} \to \kappa, \end{cases}$$

$$\int_0^1 \left\{ u(1-u) + \frac{\log m}{m} \right\} \mathcal{K}_{1,2}^2(u + \Delta(u), t + \Delta(t)) du$$

$$= \frac{1}{b^4} \int_0^1 \left\{ u(1-u) + \frac{\log m}{m} \right\} \mathcal{K}^2(u, t + \Delta(t)) \left\{ 1 + O\left(\left(\frac{t+bc}{u} + \frac{1-t+bc}{1-u} \right) \sqrt{\frac{\log m}{mb^2}} \right) \right\}^2$$

$$\times \left[\left| \log(u) - \psi\left(\frac{t+\Delta(t)}{b} + c + 1 \right) - \log(1-u) + \psi\left(\frac{1-t-\Delta(t)}{b} + c + 1 \right) \right. \right.$$

$$\left. + O\left(u^{-1}(1-u)^{-1} \sqrt{\frac{\log m}{m}} \right) \right|$$

$$\left. \times \left| \frac{t+bc+\Delta(t)}{u} - \frac{1-t+bc-\Delta(t)}{1-u} + O\left(\left(\frac{t+bc}{u^2} + \frac{1-t+bc}{(1-u)^2} \right) \sqrt{\frac{\log m}{m}} \right) \right| + \frac{b}{u(1-u)} \right]^2 du$$

$$= \begin{cases} O\left(\dfrac{b^{-5/2}}{(t+bc)^{3/2}(1-t+bc)^{3/2}} \right), & \dfrac{t}{b} \to \infty, \dfrac{1-t}{b} \to \infty, \\ O\left(\dfrac{b^{-3}}{(t+bc)(1-t+bc)} \right), & \dfrac{t}{b} \to \kappa \text{ or } \dfrac{1-t}{b} \to \kappa, \end{cases}$$

$$\int_0^1 \mathcal{K}_{2,2}^2(u + \Delta(u), t + \Delta(t)) du$$

$$= \frac{1}{b^4} \int_0^1 \mathcal{K}^2(u, t + \Delta(t)) \left\{ 1 + O\left(\left(\frac{t+bc+\Delta(t)}{u} + \frac{1-t+bc-\Delta(t)}{1-u} \right) \sqrt{\frac{\log m}{mb^2}} \right) \right\}$$

$$\times \left[\left\{ \log(u) - \psi\left(\frac{t+\Delta(t)}{b} + c + 1 \right) - \log(1-u) + \psi\left(\frac{1-t-\Delta(t)}{b} + c + 1 \right) \right. \right.$$

$$\left. + O\left(u^{-1}(1-u)^{-1} \sqrt{\frac{\log m}{m}} \right) \right\}^2 - \psi'\left(\frac{t+\Delta(t)}{b} + c + 1 \right) - \psi'\left(\frac{1-t-\Delta(t)}{b} + c + 1 \right) \right]^2 du$$

$$= \begin{cases} O\left(\dfrac{b^{-5/2}}{(t+bc)^{5/2}(1-t+bc)^{5/2}} \right), & \dfrac{t}{b} \to \infty, \dfrac{1-t}{b} \to \infty, \\ O\left(\dfrac{b^{-3}}{(t+bc)^2(1-t+bc)^2} \right), & \dfrac{t}{b} \to \kappa \text{ or } \dfrac{1-t}{b} \to \kappa, \end{cases}$$

noting that $\Delta(t) = O((\log m)/m)$. It follows that

$$\mathrm{E}[\mathcal{R}_{k,l}^2 \chi_{\widetilde{S}}] = \begin{cases} O\left(\dfrac{(mb^{1/2})^{-1}}{(t+bc)^{1/2}(1-t+bc)^{1/2}} \dfrac{(\log m)^2}{mb^2} \right), & \dfrac{t}{b} \to \infty, \dfrac{1-t}{b} \to \infty, \\ O\left((mb)^{-1} \dfrac{(\log m)^2}{mb^2} \right), & \dfrac{t}{b} \to \kappa \text{ or } \dfrac{1-t}{b} \to \kappa, \end{cases} \quad k, l = 1, 2.$$

18

Also, using Lemmas A.2 and A.3 (with $M \geq 10$), we have

$$\mathrm{E}[\mathcal{R}_{k,l}^2 \chi_{\widetilde{\mathcal{S}}^c}] = o((mb^{1/2})^{-1}), \quad k, l = 1, 2. \quad \square$$

Proof of Theorem 5 We have

$$\mathrm{Var}[\widehat{g}_b^{(B)}(x)] = \mathrm{Var}[\widetilde{g}_b^{(B)}(x)] + \mathrm{Var}[\mathcal{I}_{13}] + \mathrm{Var}\left[\sum_{j=1}^{2}\sum_{k=1}^{2}\mathcal{I}_{jk} + \mathcal{R}\right]$$

$$+ 2\mathrm{Cov}\left[\widetilde{g}_b^{(B)}(x) + \mathcal{I}_{13}, \sum_{j=1}^{2}\sum_{k=1}^{2}\mathcal{I}_{jk} + \mathcal{R}\right].$$

Here, Lemma 4 yields

$$\mathrm{Var}\left[\sum_{j=1}^{2}\sum_{k=1}^{2}\mathcal{I}_{jk} + \mathcal{R}\right] \leq 5\left(\sum_{j=1}^{2}\sum_{k=1}^{2}\mathrm{E}[\mathcal{I}_{jk}^2] + \mathrm{E}[\mathcal{R}^2]\right)$$

$$= \begin{cases} o\left(\dfrac{V(x)}{mb^{1/2}}\right) + O\left(\dfrac{1}{n} + \dfrac{1}{m}\right), & x \in \mathcal{S}_{I,b}, \\ o\left(\dfrac{1}{mb}\right) + O\left(\dfrac{1}{n} + \dfrac{1}{m}\right), & x \in \mathcal{S}_{B,b}, \end{cases} \tag{A5}$$

$$\mathrm{Cov}\left[\widetilde{g}_b^{(B)}(x) + \mathcal{I}_{13}, \sum_{j=1}^{2}\sum_{k=1}^{2}\mathcal{I}_{jk} + \mathcal{R}\right] \leq \left\{\mathrm{Var}\left[\widetilde{g}_b^{(B)}(x) + \mathcal{I}_{13}\right]\mathrm{Var}\left[\sum_{j=1}^{2}\sum_{k=1}^{2}\mathcal{I}_{jk} + \mathcal{R}\right]\right\}^{1/2}$$

$$= \begin{cases} o\left(\dfrac{V(x)}{mb^{1/2}}\right) + O\left(\dfrac{1}{n} + \dfrac{1}{m}\right), & x \in \mathcal{S}_{I,b}, \\ o\left(\dfrac{1}{mb}\right) + O\left(\dfrac{1}{n} + \dfrac{1}{m}\right), & x \in \mathcal{S}_{B,b}. \quad \square \end{cases}$$

Proof of Lemma 6 Let

$$\Delta_{b,s}(x) = \mathcal{K}(F_Y(s), F_Y(x)) - \mathrm{E}[\mathcal{K}(F_Y(s), F_Y(x))], \quad s \in \mathcal{S}.$$

It is easy to see that

$$nb^{1/2}\mathrm{Var}[\widetilde{g}_b^{(B)}(x)] = b^{1/2}\mathrm{E}[\Delta_{b,X_i}^2(x)] \to g(x)V(x) \quad \text{for fixed } x \in \mathcal{S}_I,$$

$$nb\mathrm{Var}[\widetilde{g}_b^{(B)}(x)] = b\mathrm{E}[\Delta_{b,X_i}^2(x)] \to g(x)v(0) \quad \text{for } x \in \mathcal{S}_B,$$

$$mb^{1/2}\mathrm{Var}[\mathcal{I}_{13}] = b^{1/2}g^2(x)\mathrm{E}[\Delta_{b,Y_j}^2(x)] \to g^2(x)V(x) \quad \text{for fixed } x \in \mathcal{S}_I,$$

$$mb\mathrm{Var}[\mathcal{I}_{13}] = bg^2(x)\mathrm{E}[\Delta_{b,Y_j}^2(x)] \to g^2(x)v(0) \quad \text{for } x \in \mathcal{S}_B,$$

$$\sup_{s \in \mathcal{S}}|\Delta_{b,s}(x)| \leq \overline{\Delta}_b(x),$$

where

$$\overline{\Delta}_b(x) = \begin{cases} 2^{1/2}b^{-1/2}V(x)\{1 + b(2c+1)\}(1 + 2bc)^{1/2} & \text{for fixed } x \in \mathcal{S}_I, \\ b^{-1}\{1 + b(2c+1)\} & \text{for } x \in \mathcal{S}_B \end{cases}$$

(see Lemma A.2 (i) and (ii)). It follows that

$$
\sum_{i=1}^{n} \mathrm{E}\left[\left|\frac{n^{-1}\Delta_{b,X_i}(x)}{\{\mathrm{Var}[\widetilde{g}_b^{(B)}(x)]\}^{1/2}}\right|^{2+\delta}\right] = \frac{\mathrm{E}[|\Delta_{b,X_1}(x)|^{2+\delta}]}{n^{\delta/2}\{\mathrm{Var}[\Delta_{b,X_1}(x)]\}^{1+\delta/2}}
$$

$$
\leq \frac{\overline{\Delta}_b^{\delta}(x)}{n^{\delta/2}\{\mathrm{E}[\Delta_{b,X_1}^2(x)]\}^{\delta/2}}
$$

$$
= \begin{cases} O((nb^{1/2})^{-\delta/2}) & \text{for fixed } x \in \mathcal{S}_I, \\ O((nb)^{-\delta/2}) & \text{for } x \in \mathcal{S}_B, \end{cases}
$$

$$
\sum_{j=1}^{m} \mathrm{E}\left[\left|\frac{m^{-1}g(x)\Delta_{b,Y_j}(x)}{\{\mathrm{Var}[\mathcal{I}_{13}]\}^{1/2}}\right|^{2+\delta}\right] = \frac{\mathrm{E}[|\Delta_{b,Y_1}(x)|^{2+\delta}]}{m^{\delta/2}\{\mathrm{Var}[\Delta_{b,Y_1}(x)]\}^{1+\delta/2}}
$$

$$
\leq \frac{\overline{\Delta}_b^{\delta}(x)}{m^{\delta/2}\{\mathrm{E}[\Delta_{b,Y_1}^2(x)]\}^{\delta/2}}
$$

$$
= \begin{cases} O((mb^{1/2})^{-\delta/2}) & \text{for fixed } x \in \mathcal{S}_I, \\ O((mb)^{-\delta/2}) & \text{for } x \in \mathcal{S}_B, \end{cases}
$$

where $\delta > 0$ is arbitrary. Consequently, Lyapunov's central limit theorem enables us to see that

$$
\frac{\widetilde{g}_b^{(B)}(x) - \mathrm{E}[\widetilde{g}_b^{(B)}(x)]}{\{\mathrm{Var}[\widetilde{g}_b^{(B)}(x)]\}^{1/2}} \xrightarrow{d} N(0,1), \quad \frac{\mathcal{I}_{13}}{\{\mathrm{Var}[\mathcal{I}_{13}]\}^{1/2}} \xrightarrow{d} N(0,1).
$$

The result follows from Slutsky's theorem. $\square$

Proof of Theorem 7 Letting $m = \lceil Cn \rceil$ for $C > 0$, (A5) yields

$$
nb^{1/2}\mathrm{Var}\left[\sum_{j=1}^{2}\sum_{k=1}^{2}\mathcal{I}_{jk} + \mathcal{R}\right] = O\left(b^{1/2} + \frac{(\log n)^2}{nb^2}\right) = o(1) \quad \text{for fixed } x \in \mathcal{S}_I,
$$

$$
nb\mathrm{Var}\left[\sum_{j=1}^{2}\sum_{k=1}^{2}\mathcal{I}_{jk} + \mathcal{R}\right] = O\left(b + \frac{(\log n)^2}{nb^2}\right) = o(1) \quad \text{for } x \in \mathcal{S}_B.
$$

The Chebyshev inequality enables us to show that

$$
(nb^{1/2})^{1/2}\{\widehat{g}_b^{(B)}(x) - \mathrm{E}[\widehat{g}_b^{(B)}(x)]\} = (nb^{1/2})^{1/2}\{\widetilde{g}_b^{(B)}(x) + \mathcal{I}_{13} - \mathrm{E}[\widetilde{g}_b^{(B)}(x)]\} + o_p(1) \quad \text{for fixed } x \in \mathcal{S}_I,
$$

$$
(nb)^{1/2}\{\widehat{g}_b^{(B)}(x) - \mathrm{E}[\widehat{g}_b^{(B)}(x)]\} = (nb)^{1/2}\{\widetilde{g}_b^{(B)}(x) + \mathcal{I}_{13} - \mathrm{E}[\widetilde{g}_b^{(B)}(x)]\} + o_p(1) \quad \text{for } x \in \mathcal{S}_B,
$$

The results follow from Lemma 6. $\square$

Proof of Theorem 8 It is easy to see that

$$
\left|\int_{\mathcal{S}}\{\mathrm{Bias}[\widehat{g}_b^{(B)}(x)]\}^2 f_Y(x)dx - b^2\int_{\mathcal{S}}B^2(x)f_Y(x)dx\right|
$$

$$
= \left|\int_{\mathcal{S}}\{\mathrm{Bias}[\widehat{g}_b^{(B)}(x)] - bB(x)\}\{\mathrm{Bias}[\widehat{g}_b^{(B)}(x)] + bB(x)\}f_Y(x)dx\right|
$$

$$
= o(b^2).
$$

On the other hand, for any $\tau \in (1/2, 1)$,

$$\left| \int_{\mathcal{S}} \mathrm{Var}[\widehat{g}_b^{(B)}(x)] f_Y(x) dx - b^{-1/2} \int_{\mathcal{S}} \left(\frac{1}{n} + \frac{g(x)}{m} \right) V(x) f_X(x) dx \right|$$

$$= \left| \int_0^1 \mathrm{Var}[\widehat{g}_b^{(B)}(F_Y^{-1}(t))] dt - b^{-1/2} \int_0^1 \left(\frac{1}{n} + \frac{g(F_Y^{-1}(t))}{m} \right) V(F_Y^{-1}(t)) dt \right|$$

$$\leq \left| \int_{b^\tau}^{1-b^\tau} \left\{ \mathrm{Var}[\widehat{g}_b^{(B)}(F_Y^{-1}(t))] - b^{-1/2} \left(\frac{1}{n} + \frac{g(F_Y^{-1}(t))}{m} \right) V(F_Y^{-1}(t)) \right\} dt \right|$$

$$+ \left(\int_0^{b^\tau} + \int_{1-b^\tau}^1 \right) \mathrm{Var}[\widehat{g}_b^{(B)}(F_Y^{-1}(t))] dt + b^{-1/2} \left(\int_0^{b^\tau} + \int_{1-b^\tau}^1 \right) \left(\frac{1}{n} + \frac{g(F_Y^{-1}(t))}{m} \right) V(F_Y^{-1}(t)) dt$$

$$= o\left(\frac{1}{nb^{1/2}} + \frac{1}{mb^{1/2}} \right),$$

noting that $\{F_Y^{-1}(t) \mid b^\tau \leq t \leq 1 - b^\tau\} \subset \mathcal{S}_{I,b}$,

$$\mathrm{Var}[\widehat{g}_b^{(B)}(x)] = O\left(\frac{1}{nb} + \frac{1}{mb} \right), \quad \text{uniformly in } x \in \mathcal{S},$$

$$\left(\int_0^{b^\tau} + \int_{1-b^\tau}^1 \right) \left(\frac{1}{n} + \frac{g(F_Y^{-1}(t))}{m} \right) V(F_Y^{-1}(t)) dt \leq \left(\frac{1}{n} + \frac{||g||_{\mathcal{S}}}{m} \right) \left(\int_0^{b^\tau} + \int_{1-b^\tau}^1 \right) \frac{1}{2\sqrt{\pi t(1-t)}} dt$$

$$= O\left(\frac{b^{\tau/2}}{n} + \frac{b^{\tau/2}}{m} \right). \quad \square$$

References

Anderson, G. D., Barnard, R. W., Richards, K. C., Vamanamurthy, M. K. and Vuorinen, M. (1995) "Inequalities for zero-balanced hypergeometric functions", *Transactions of the American Mathematical Society*, 347, 1713–1723.

Chen, S. X. (1999) "Beta kernel estimators for density functions", *Computational Statistics and Data Analysis*, 31, 131–145.

Chen, S.-M., Hsu, Y.-S. and Liaw, J.-T. (2009) "On kernel estimators of density ratio", *Statistics*, 43, 463–479.

Ćwik, J. and Mielniczuk, J. (1989) "Estimating density ratio with application to discriminant analysis", *Communications in Statistics - Theory and Methods*, 18, 3057–3069.

Ćwik, J. and Mielniczuk, J. (1993) "Data-dependent bandwidth choice for a grade density kernel estimate", *Statistics and Probability Letters*, 16, 397–405.

Gijbels, I. and Mielniczuk, J. (1995) "Asymptotic properties of kernel estimators of the Radon-Nikodym derivative with applications to discriminant analysis", *Statistica Sinica*, 5, 261–278.

Igarashi, G. (2016) "Bias reductions for beta kernel estimation", *Journal of Nonparametric Statistics*, 28, 1–30.

Parzen, E. (1962) "On estimation of a probability density function and mode", *The Annals of Mathematical Statistics*, 33, 1065–1076.

Qi, F., Guo, S. and Guo, B.-N. (2010) "Complete monotonicity of some functions involving polygamma functions", *Journal of Computational and Applied Mathematics*, 233, 2149–2160.

Rosenblatt, M. (1956) "Remarks on some nonparametric estimates of a density function", *The Annals of Mathematical Statistics*, 27, 832–837.

Schuster, E. F. (1985) "Incorporating support constraints into nonparametric estimators of densities", *Communications in Statistics - Theory and Methods*, 14, 1123–1136.

Wand, M. P. and Jones, M. C. (1995) *Kernel Smoothing*, London: Chapman & Hall.