ASYMPTOTIC PROPERTIES OF A KERNEL TYPE ESTIMATOR OF A DENSITY RATIO

Moriyama, Taku
Graduate School of Mathematics, Kyushu University

Maesono, Yoshihiko
Faculty of Mathematics, Kyushu University

https://doi.org/10.5109/2203026

出版情報：Bulletin of informatics and cybernetics. 48, pp.37-46, 2016-12. 統計科学研究会
バージョン：
権利関係：

Bulletin of Informatics and Cybernetics, Vol. 48, 2016

ASYMPTOTIC PROPERTIES OF A KERNEL TYPE ESTIMATOR OF A DENSITY RATIO

By

Taku Moriyama* and **Yoshihiko Maesono**†

Abstract

In statistical inference, estimation of a density ratio takes important roll. The density ratio is a measure of difference of two populations, and then its estimator is vital. The ratio estimator is used for testing equality of two density functions, detecting change points, discriminant analysis etc. Under some parametric assumptions, there are many papers which study asymptotic properties of the ratio estimators. In this paper we will discuss a kernel type estimator of the ratio.

Key Words and Phrases: Asymptotic properties; density ratio; Edgeworth expansion; kernel type estimator

1. Introduction

The density ratio is a measure of difference of two populations, and then its estimator is useful in a statistical inference. The ratio estimator is used for testing equality of two density functions, detecting change points, discriminant analysis etc. Under some parametric assumptions, there are many papers which study asymptotic properties of the ratio estimators. In this paper we will discuss a kernel type estimator of the ratio, and obtain a higher order asymptotic mean squared error and an Edgeworth expansion.

Let $X_1, \cdots, X_m$ be *i.i.d.* random variables with density and distribution functions f and F, and $Y_1, \cdots, Y_n$ be *i.i.d.* random variables with g and G. Let us assume that the two densities $f(x)$ and $g(x)$ are mutually independent and $g(x_0) \neq 0$. Here we will discuss a density ratio at point x_0, that is $f(x_0)/g(x_0)$. A kernel type estimator is given by $\widehat{f}(x_0)/\widehat{g}(x_0)$, where $\widehat{f}(x_0)$ and $\widehat{g}(x_0)$ are kernel estimators of the densities. The asymptotic mean squared error ($AMSE$) of the estimator $\widehat{f}(x_0)/\widehat{g}(x_0)$ is obtained by Chen et al. (2009). This estimator is constructed from two estimators and substituted for numerator and denominator separately. For the density estimator $\widehat{f}(x_0)$, Umeno and Maesono(2013) have obtained an Edgeworth expansion, and for the distribution function estimator

$$\widehat{F}(x_0) = \frac{1}{m}\sum_{i=1}^{m} W_f\left(\frac{x_0 - X_i}{h_{f,m}}\right), \qquad \left(W_f(t) = \int_{-\infty}^{t} K_{f,m}(u)du\right)$$

Huang and Maesono(2014) have obtained the Edgeworth expansion. In this paper, we will obtain an asymptotic representation and the mean squared error of the estimator. We also establish an Edgeworth expansion with residual $o(N^{-1/2})$ where $N = m + n$. All proofs are given in Appendix.

* Graduate School of Mathematics, Kyushu University, Motooka, Fukuoka 819–0395, Japan.
† Faculty of Mathematics, Kyushu University, Motooka, Fukuoka 819–0395, Japan. tel +81–92–802–4480 maesono@math.kyushu-u.ac.jp

2. Density ratio estimator

The kernel estimator of the ratio $f(x_0)/g(x_0)$ is given by

$$\frac{\widehat{f}(x_0)}{\widehat{g}(x_0)}$$

where

$$\begin{aligned}\widehat{f}(x_0) &= \frac{1}{mh_{f,m}}\sum_{i=1}^{m}K_f\left(\frac{x_0-X_i}{h_{f,m}}\right),\\ \widehat{g}(x_0) &= \frac{1}{nh_{g,n}}\sum_{j=1}^{n}K_g\left(\frac{x_0-Y_j}{h_{g,n}}\right)\end{aligned}$$

where $h_{f,m}$ and $h_{g,n}$ are bandwidths, and K_f and K_g are kernel functions. When $m=n, h_{g,n}=h_{f,n}$ and $K_f(\cdot)=K_g(\cdot)$, Ćwik and Mielniczuk(1989) and Chen et al.(2009) have discussed the asymptotic mean squared error. In this paper we will obtain a precise mean squared error and the Edgeworth expansion of this estimator. Here we assume that the kernels satisfy

$$\int_{-\infty}^{\infty}K_\zeta(u)du=1, \qquad \int_{-\infty}^{\infty}K_\zeta(u)udu=0, \tag{1}$$

$$\int_{-\infty}^{\infty}K_\zeta(u)u^2du\neq 0, \qquad \int_{-\infty}^{\infty}K_\zeta(u)u^3du=0 \tag{2}$$

where ζ denotes f or g. If the kernels are symmetric, that is $K_\zeta(-u)=K_\zeta(u)$, then odd moments of $K_\zeta(\cdot)$ are 0.

Let us assume that

$$0<\lambda=\lim_{N\to\infty}\frac{m}{N}<1.$$

Let us define $h_N=\min\{h_{f,m},\ h_{g,n}\}$ and $o_L(N^{-1/2})$ as

$$P\left(|o_L(N^{-1/2})|\geq N^{-1/2}\varepsilon_N\right)=o(N^{-1/2})$$

where $\varepsilon_N\to 0$. It follows from the Markov inequality that if $E|R_N|^p=o(N^{-1/2-p/2})$, we have $R_N=o_L(N^{-1/2})$. Using the Taylor expansion and moment evaluations of a sum of *i.i.d.* random variables, we have a stochastic expansion. Since $\widehat{f}(x_0)$ and $\widehat{g}(x_0)$ are sample means of *i.i.d.* random variables, we can show the following lemma.

Lemma 2.1. *Let us assume that $h_N=O(N^{-c})$ ($\frac{1}{4}\leq c<\frac{1}{3}$), and both density function f and g have 3rd bounded continuous derivatives. If K_f and K_g are symmetric and $\int|K_\zeta(u)|^jdu<\infty$ ($j=1,\cdots,4$; $\zeta=f$ or g), we have*

$$\begin{aligned}\frac{\widehat{f}(x_0)}{\widehat{g}(x_0)} &= \frac{\widehat{f}(x_0)}{E[\widehat{g}(x_0)]}-\frac{\widehat{f}(x_0)}{(E[\widehat{g}(x_0)])^2}\{\widehat{g}(x_0)-E[\widehat{g}(x_0)]\}\\ &\quad+\frac{\widehat{f}(x_0)}{(E[\widehat{g}(x_0)])^3}\{\widehat{g}(x_0)-E[\widehat{g}(x_0)]\}^2\\ &\quad-\frac{\widehat{f}(x_0)}{(E[\widehat{g}(x_0)])^4}\{\widehat{g}(x_0)-E[\widehat{g}(x_0)]\}^3+(Nh_N)^{-1/2}o_L(N^{-1/2}).\end{aligned}$$

For $\zeta = f$ or g, let us define

$$A^{\zeta}_{i,j} = \int K^i_{\zeta}(u)u^j du,$$
$$B_{\zeta,k} = \frac{1}{k!}A^{\zeta}_{1,k}\zeta^{(k)}(x_0).$$

Since $\widehat{f}(x)/\widehat{g}(x)$ is invariant under permutation of each of $(X_1, \cdots, X_m)$ and $(Y_1, \cdots, Y_n)$, we can get the following approximation called the Hoeffding decomposition. Using the Hoeffding decomposition and its moments, the following theorem holds.

THEOREM 2.2. *Let us assume that $h_N = O(N^{-c})$ $(\frac{1}{4} \le c < \frac{1}{3})$, and both density function f and g have 3rd bounded continuous derivatives. If K_f and K_g are symmetric, and $\int |K_\zeta(u)|^j du < \infty$ $(j = 1, \cdots, 4\ ;\ \zeta = f$ or $g)$, we have*

$$\begin{aligned}
&\frac{\widehat{f}(x_0)}{\widehat{g}(x_0)} - \frac{f(x_0)}{g(x_0)} \\
= \ &\frac{h^2_{f,m}B_{f,2}}{g(x_0)} - \frac{f(x_0)h^2_{g,n}B_{g,2}}{g^2(x_0)} + \frac{f(x_0)A^g_{2,0}}{nh_{g,n}g^3(x_0)} \\
&+\frac{1}{m}\sum_{i=1}^{m} d_{1,N}U_{i,m} - \frac{1}{n}\sum_{j=1}^{n} d_{2,N}W_{j,n} \\
&-\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} d_{3,N}U_{i,m}W_{j,n} + \frac{2}{n^2}\sum_{1\le i<j\le n}\frac{f(x_0)}{g^3(x_0)}W_{i,n}W_{j,n} \\
&+\frac{1}{n^2}\sum_{i=1}^{n}\frac{f(x_0)}{g^3(x_0)}\left\{W^2_{i,n} - E[W^2_{i,n}]\right\} + O(N^{-3})\sum\nolimits^{*}\alpha_{i,N}\beta_{j,N}\gamma_{\ell,N} \\
&+(Nh_N)^{-1/2}o_L(N^{-1/2}),
\end{aligned}$$

where

$$\begin{aligned}
d_{1,N} &= \frac{1}{g(x_0)} - \frac{h^2_{g,n}}{g^2(x_0)}B_{g,2}, \\
d_{2,N} &= \frac{f(x_0) + h^2_{f,m}B_{f,2}}{g^2(x_0)} - h^2_{g,n}\frac{2f(x_0)}{g^3(x_0)}B_{g,2}, \\
d_{3,N} &= \frac{1}{g^2(x_0)} - \frac{2h^2_{g,n}}{g^3(x_0)}B_{g,2}, \\
U_{i,m} &= \frac{1}{h_{f,m}}\left[K_f\left(\frac{x_0 - X_i}{h_{f,m}}\right) - E\left\{K_f\left(\frac{x_0 - X_i}{h_{f,m}}\right)\right\}\right], \\
W_{i,n} &= \frac{1}{h_{g,n}}\left[K_g\left(\frac{x_0 - Y_i}{h_{g,n}}\right) - E\left\{K_g\left(\frac{x_0 - Y_i}{h_{g,n}}\right)\right\}\right].
\end{aligned}$$

$\sum^{*}$ indicates a sum of different indices i, j, ℓ, and $\alpha_{i,N}$, $\beta_{j,N}$ and $\gamma_{\ell,N}$ are linear combinations of $U_{\cdot,m}$ and $W_{\cdot,n}$.

Using this asymptotic representation, an asymptotic bias and a variance are given by the following theorem.

Theorem 2.3. *Let us assume that* $h_N = O(N^{-c})$ $(\frac{1}{4} \le c < \frac{1}{3})$, *and both density function* f *and* g *have 3rd bounded continuous derivatives. If* K_f *and* K_g *are symmetric and* $\int |k_\zeta(u)|^j du < \infty$ $(j = 1, \cdots, 4\ ;\ \zeta = f$ *or* $g)$, *we have*

$$E\left[\frac{\widehat{f}(x_0)}{\widehat{g}(x_0)}\right] = \frac{f(x_0)}{g(x_0)} + b_{m,n} + O(N^{-1}),$$

and

$$Var\left[\frac{\widehat{f}(x_0)}{\widehat{g}(x_0)}\right] = \tau^2_{m,n} + o(N^{-7/6})$$

where

$$\begin{aligned} b_{m,n} &= \frac{h^2_{f,m} B_{f,2}}{g(x_0)} - \frac{f(x_0) h^2_{g,n} B_{g,2}}{g^2(x_0)} + \frac{f(x_0) A^g_{2,0}}{n h_{g,n} g^3(x_0)}, \\ \tau^2_{m,n} &= \frac{1}{g^2(x_0)} \left\{ \frac{f(x_0)}{m h_{f,m}} A^f_{2,0} - \frac{f^2(x_0)}{m} \right\} + \frac{f^2(x_0)}{g^4(x_0)} \left\{ \frac{g(x_0)}{n h_{g,n}} A^g_{2,0} - \frac{g^2(x_0)}{n} \right\}. \end{aligned}$$

The asymptotic mean squared error is given by

$$AMSE\left(\frac{\widehat{f}(x_0)}{\widehat{g}(x_0)}\right) = b^2_{m,n} + \tau^2_{m,n} + o(N^{-7/6}).$$

3. Edgeworth expansion

Here we discuss the Edgeworth expansion of the density ratio estimator $\widehat{f(x_0)}/\widehat{g(x_0)}$. Using the Edgeworth expansion for two-sample U-statistics (Maesono(1985)), we have the following expansion. Let us define

$$\begin{aligned} \kappa_{m,n} &= \frac{1}{\tau^3_{m,n}} \left\{ \frac{f(x_0)}{m^2 h^2_{f,m}} A^f_{3,0} - \frac{f^3(x_0)}{n^2 h^2_{g,n} g^5(x_0)} A^g_{3,0} \right. \\ &\quad \left. + \frac{6 f^2(x_0)}{m n h_{f,m} h_{g,m} g^4(x_0)} A^f_{2,0} A^g_{2,0} + \frac{6 f^3(x_0)}{n^2 h^2_{g,n} g^5(x_0)} \left[A^g_{2,0}\right]^2 \right\} \end{aligned}$$

and

$$Q_{m,n}(y) = \Phi(y) - \phi(y) \left\{ \frac{\kappa_{m,n}}{6}(y^2 - 1) + \frac{b_{m,n}}{\tau_{m,n}} \right\}$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ denote the distribution and the density function of the standard normal $N(0,1)$. Similarly as and García-Soidán et al.(1997), we can prove the validity of the expansion.

Theorem 3.1. *Let us assume that* $h_N = O(N^{-c})$ $(\frac{1}{4} \le c < \frac{1}{3})$, *and both density function* f *and* g *have 3rd bounded continuous derivatives. If* K_f *and* K_g *are symmetric, and* $\int |K_\zeta(u)|^j du < \infty$ $(j = 1, \cdots, 4\ ;\ \zeta = f$ *or* $g)$, *we have*

$$\sup_{-\infty < y < \infty} \left| P\left(\tau^{-1}_{m,n} \left\{ \frac{\widehat{f}(x_0)}{\widehat{g}(x_0)} - \frac{f(x_0)}{g(x_0)} \right\} \le y \right) - Q_{m,n}(y) \right| = o\left(N^{-1/2}\right).$$

Remark. $\kappa_{m,n}$ depends on unknown values $f(x_0)$, $f''(x_0)$, $g(x_0)$ and $g''(x_0)$. It is possible to make consistent estimators of them, and so we can improve the normal approximation of a significance probability and a confidence coefficient.

4. Appendices

First we prepare moment evaluations for the Hoeffding decomposition (see Maesono (1997)). For a function $v_r(x_1, \cdots, x_r)$ which is invariant under permutations of its arguments and satisfies

$$E[v_r(X_1, \cdots, X_r)|X_1, \cdots, X_{r-1}] = 0 \quad a.s.$$

Let us define

$$D_{r,M} = \sum_{1 \le i_1 < \cdots < i_r \le M} v_r(X_{i_1}, \cdots, X_{i_r}). \tag{3}$$

Then for $M \ge r$ we have

$$E\,|D_{r,M}|^p \le C_v M^{pr/2} E|v_r(X_{i_1}, \cdots, X_{i_r})|^p \tag{4}$$

where C_v does not depend on M. Hereafter, we use same symbol $D_{r,M}$ which is different in each case.

Note that

$$\begin{aligned} E\left[\frac{1}{h_{f,m}^{\ell}} K_f^{\ell}\left(\frac{x_0 - X_1}{h_{f,m}}\right)\right] &= h_{f,m}^{1-\ell} \int_{-\infty}^{\infty} K_f^{\ell}(u) f(x_0 - u h_{f,m}) du \\ &= h_{f,m}^{1-\ell} f(x_0) A_{\ell,0}^{f} + O(h_{f,m}^{2-\ell}). \end{aligned} \tag{5}$$

Proof of Lemma 2.1

Using the Taylor expansion, we get

$$\begin{aligned} \frac{\widehat{f}(x_0)}{\widehat{g}(x_0)} &= \frac{\widehat{f}(x_0)}{E[\widehat{g}(x_0)]} - \frac{\widehat{f}(x_0)}{(E[\widehat{g}(x_0)])^2} \{\widehat{g}(x_0) - E[\widehat{g}(x_0)]\} \\ &\quad + \frac{\widehat{f}(x_0)}{(E[\widehat{g}(x_0)])^3} \{\widehat{g}(x_0) - E[\widehat{g}(x_0)]\}^2 \\ &\quad - \frac{\widehat{f}(x_0)}{(E[\widehat{g}(x_0)])^4} \{\widehat{g}(x_0) - E[\widehat{g}(x_0)]\}^3 \\ &\quad + \frac{\widehat{f}(x_0)}{(g^*)^5} \{\widehat{g}(x_0) - E[\widehat{g}(x_0)]\}^4 \end{aligned}$$

where $|g^* - E[\widehat{g}(x_0)]| \le |\widehat{g}(x_0) - E[\widehat{g}(x_0)]|$. Let us define

$$\delta_N = \frac{g^* - E[\widehat{g}(x_0)]}{E[\widehat{g}(x_0)]}.$$

Then it follows from the definition of δ_N that

$$\begin{aligned} &\frac{\widehat{f}(x_0)}{(g^*)^5} \{\widehat{g}(x_0) - E[\widehat{g}(x_0)]\}^4 \\ = \; &(1+\delta_N)^{-5} \frac{E[\widehat{f}(x_0)]}{(E[\widehat{g}(x_0)])^5} \{\widehat{g}(x_0) - E[\widehat{g}(x_0)]\}^4 \\ &+ (1+\delta_N)^{-5} \frac{\widehat{f}(x_0) - E[\widehat{f}(x_0)]}{(E[\widehat{g}(x_0)])^5} \{\widehat{g}(x_0) - E[\widehat{g}(x_0)]\}^4. \end{aligned}$$

It is easy to see that

$$
\begin{aligned}
& P\left(\left|(1+\delta_N)^{-5}\frac{E[\widehat{f}(x_0)]}{(E[\widehat{g}(x_0)])^5}\left\{\widehat{g}(x_0)-E[\widehat{g}(x_0)]\right\}^4\right| \geq n^{-1/2}(\log n)^{-1}\right) \\
\leq\ & P\left(|\delta_N| \geq \frac{1}{2}\right) + P\left(\left|\frac{1}{2^5}\frac{E[\widehat{f}(x_0)]}{(E[\widehat{g}(x_0)])^5}\left\{\widehat{g}(x_0)-E[\widehat{g}(x_0)]\right\}^4\right| \geq n^{-1/2}(\log n)^{-1}\right).
\end{aligned}
$$

Using the Hoeffding decomposition, we have

$$
\begin{aligned}
& \left\{\widehat{g}(x_0)-E[\widehat{g}(x_0)]\right\}^4 \\
=\ & O(n^{-4})\left\{nE(W_{1,n}^4) + n^2\left[E(W_{1,n}^2)\right]^2\right\} \\
& + O(n^{-4})D_{1,n} + O(n^{-4})D_{2,n} + O(n^{-4})D_{3,n} + O(n^{-4})D_{4,n}
\end{aligned}
$$

where $D_{j,n}$ is defined in the equation (3). It follows from the evaluation (5) that

$$
\begin{aligned}
& O(n^{-4})\left\{nE(W_{1,n}^4) + n^2\left[E(W_{1,n}^2)\right]^2\right\} \\
=\ & O(N^{-4})\left\{nh_{f,n}^{-3}f(x_0)A_{4,0}^f + n^2h_{f,n}^{-2}f^2(x_0)[A_{2,0}^f]^2\right\} \\
=\ & (Nh_N)^{-1/2}o_L(n^{-1/2}).
\end{aligned}
$$

Further, it follows from the moment evaluations (4) and (5) that

$$
E(D_{4,n}^2) = O(n^4)\left[E(W_{1,n}^2)\right]^4 = O(n^4h_{g,n}^{-4})
$$

and

$$
E\left[\left\{\sqrt{Nh_N}O(n^{-4})D_{4,n}\right\}^2\right] = O(n^{-3}h_N^{-3}) = O(n^{-1/2-1-1/2}).
$$

Similarly, we have

$$
O(n^{-4})D_{1,n} + O(n^{-4})D_{2,n} + O(n^{-4})D_{3,n} = (Nh_N)^{-1/2}o_L(N^{-1/2}).
$$

It is easy to see that

$$
P\left(|\delta_N| \geq \frac{1}{2}\right) \leq P\left(\left|\frac{\widehat{g}(x_0)-E[\widehat{g}(x_0)]}{E[\widehat{g}(x_0)]}\right| \geq \frac{1}{2}\right) = o(N^{-1/2}).
$$

Thus we have

$$
\frac{\widehat{f}(x_0)}{(E[\widehat{g}(x_0)])^5}\left\{\widehat{g}(x_0)-E[\widehat{g}(x_0)]\right\}^4 = (Nh_N)^{-1/2}o_L(N^{-1/2}).
$$

Similarly we can show that

$$
(1+\delta_N)^{-5}\frac{\widehat{f}(x_0)-E[\widehat{f}(x_0)]}{(E[\widehat{g}(x_0)])^5}\left\{\widehat{g}(x_0)-E[\widehat{g}(x_0)]\right\}^4 = (Nh_N)^{-1/2}o_L(N^{-1/2}).
$$

Thus we have the desired result.

Proof of Theorem 2.2

Since the kernels satisfy the conditions (1) and (2), we have

$$E[\widehat{f}(x_0)] = f(x_0) + h_{f,m}^2 B_{f,2} + O(h_N^4), \tag{6}$$

and

$$E[\widehat{g}(x_0)] = g(x_0) + h_{g,n}^2 B_{g,2} + O(h_N^4). \tag{7}$$

It is easy to see that

$$\frac{\widehat{f}(x_0)}{E[\widehat{g}(x_0)]} = \frac{E[\widehat{f}(x_0)]}{E[\widehat{g}(x_0)]} + \frac{\widehat{f}(x_0) - E\widehat{f}(x_0)]}{E[\widehat{g}(x_0)]}$$

and

$$\frac{1}{E[\widehat{g}(x_0)]} = \frac{1}{g(x_0)} - \frac{h_{g,n}^2}{g^2(x_0)} B_{g,2} + O(h_N^4).$$

Since

$$E\left[(Nh_N)^{1/2} O(h_N^4) \frac{\widehat{f}(x_0) - E\widehat{f}(x_0)]}{E[\widehat{g}(x_0)]}\right]^2 = O(N^{-2}) = O(N^{-1/2-1-1/2}),$$

using the Taylor expansion, we have

$$\begin{aligned}
\frac{\widehat{f}(x_0)}{E[\widehat{g}(x_0)]} &= \frac{E[\widehat{f}(x_0)]}{E[\widehat{g}(x_0)]} + \frac{1}{m}\sum_{i=1}^{m}\left(\frac{1}{g(x_0)} - \frac{h_{g,n}^2}{g^2(x_0)} B_{g,2}\right) U_{i,m} \\
&\quad + (Nh_N)^{-1/2} o_L(N^{-1/2}) \\
&= \frac{1}{m}\sum_{i=1}^{m}\left(\frac{1}{g(x_0)} - \frac{h_{g,n}^2}{g^2(x_0)} B_{g,2}\right) U_{i,m} + \frac{f(x_0)}{g(x_0)} + \frac{h_{f,m}^2 B_{f,2}}{g(x_0)} \\
&\quad - \frac{f(x_0) h_{g,n}^2 B_{g,2}}{g^2(x_0)} + (Nh_N)^{-1/2} o_L(N^{-1/2}).
\end{aligned}$$

Similarly, we can show that

$$\begin{aligned}
&-\frac{\widehat{f}(x_0)}{(E[\widehat{g}(x_0)])^2}\left\{\widehat{g}(x_0) - E[\widehat{g}(x_0)]\right\} \\
= \; & -\frac{E[\widehat{f}(x_0)]}{(E[\widehat{g}(x_0)])^2}\frac{1}{n}\sum_{j=1}^{n} W_{j,n} - \frac{1}{(E[\widehat{g}(x_0)])^2}\frac{1}{mn}\sum_{i=1}^{m} U_{i,m}\sum_{j=1}^{n} W_{j,n}.
\end{aligned}$$

Let us consider the approximation of the terms which is multiplied by $1/(E[\widehat{g}(x_0)])^2$. Since $U_{i,m}$ and $W_{j,n}$, it follows from (7) that

$$\begin{aligned}
&E\left((Nh_N)^{1/2} h_{g,n}^2 B_{g,2} \frac{1}{mn}\sum_{i=1}^{m} U_{i,m}\sum_{j=1}^{n} W_{j,n}\right)^2 \\
= \; & Nh_N h_{g,n}^4 B_{g,2}^2 \frac{1}{mn} E(U_{1,m}^2) E(W_{1,n}^2) \\
= \; & O(N^{-1} h_N^3) = O(N^{-7/4}) = O(N^{-1/2-1-1/4}).
\end{aligned}$$

Thus we have

$$\begin{aligned}&\frac{1}{(E[\widehat{g}(x_0)])^2}\frac{1}{mn}\sum_{i=1}^m U_{i,m}\sum_{j=1}^n W_{j,n}\\ =\ &\frac{1}{g^2(x_0)}\frac{1}{mn}\sum_{i=1}^m U_{i,m}\sum_{j=1}^n W_{j,n}+(Nh_N)^{-1/2}o_L(N^{-1/2}).\end{aligned}$$

Further we obtain the following approximation

$$\begin{aligned}&\frac{E[\widehat{f}(x_0)]}{(E[\widehat{g}(x_0)])^2}\frac{1}{n}\sum_{j=1}^n W_{j,n}\\ =\ &\frac{1}{n}\sum_{j=1}^n\left(\frac{f(x_0)+h_{f,m}^2B_{f,2}}{g^2(x_0)}-h_{g,n}^2\frac{2f(x_0)}{g^3(x_0)}B_{g,2}\right)W_{j,n}+(Nh_N)^{-1/2}o_L(N^{-1/2}).\end{aligned}$$

For the next term, it follows from the moment evaluations (4) that

$$\begin{aligned}&\frac{\widehat{f}(x_0)}{(E[\widehat{g}(x_0)])^3}\{\widehat{g}(x_0)-E[\widehat{g}(x_0)]\}^2\\ =\ &\frac{2}{n^2}\sum_{1\le i<j\le n}\frac{f(x_0)}{g^3(x_0)}W_{i,n}W_{j,n}+\frac{1}{n^2}\sum_{i=1}^n\frac{f(x_0)}{g^3(x_0)}\left\{W_{i,n}^2-E[W_{i,n}^2]\right\}\\ &+\frac{f(x_0)A_{2,0}^g}{nh_{g,n}g^3(x_0)}+O(N^{-3})\sum\nolimits^*\alpha_{i,N}\beta_{j,N}\gamma_{k,N}+(Nh_N)^{-1/2}o_L(N^{-1/2}).\end{aligned}$$

Here we use the fact that

$$\frac{f(x_0)}{ng^3(x_0)}E[W_{1,n}^2]=\frac{f(x_0)A_{2,0}^g}{nh_{g,n}g^3(x_0)}+O(N^{-1}).$$

Finally, using the Hoeffding decomposition and the moment evaluations (4), we can show that

$$\frac{\widehat{f}(x_0)}{(E[\widehat{g}(x_0)])^4}\{\widehat{g}(x_0)-E[\widehat{g}(x_0)]\}^3=O(N^{-3})\sum\nolimits^*\alpha_{i,N}\beta_{j,N}\gamma_{k,N}+(Nh_N)^{-1/2}o_L(N^{-1/2}).$$

Combining the above approximations, we have the desired result.

Proof of Theorem 2.3

Here we obtain an asymptotic variance. From the asymptotic representation of the estimator, the asymptotic variance is given by

$$\begin{aligned}&Var\left[\frac{\widehat{f}(x_0)}{\widehat{g}(x_0)}\right]\\ =\ &Var[\widehat{f}(x_0)]\left\{\frac{1}{g^2(x_0)}-2h_{g,n}^2\frac{1}{g^3(x_0)}B_{g,2}\right\}\\ &+Var[\widehat{g}(x_0)]\left\{\frac{f^2(x_0)}{g^4(x_0)}+2h_{f,n}^2\frac{f(x_0)}{g^4(x_0)}B_{f,2}-4h_{g,n}^2\frac{f^2(x_0)}{g^3(x_0)}B_{g,2}\right\}+o(N^{-3/2}h_N^{-1}).\end{aligned}$$

Thus we get the approximation $\tau^2_{m.n}$, and the asymptotic mean squared error.

Proof of Theorem 3.1

Maesono(1998) has obtained a Edgeworth expansion of the two sample U-statistics with residual term $o(N^{-1})$. The following moments

$$h_N E(W^3_{1,n}), \qquad h_N E(W^2_{1,n} U_{1,m}), \qquad h_N^{3/2} E(\alpha^2_{1,N}\beta^2_{2,N}\gamma^2_{3,N})$$

appear in the Edgeworth expansion of the N^{-1} term. It follows from (2) that

$$\begin{aligned} &h_N E(W^3_{1,n}) = O(h_N^{-1}),\\ &h_N E(W^2_{1,n} U_{1,m}) = 0,\\ &h_N^{3/2} E(\alpha^2_{1,N}\beta^2_{2,N}\gamma^2_{3,N}) = h_N^{3/2}[E(\alpha^2_{1,N})]^3 = O(h_N^{-3/2}) = o(N^{1/2}). \end{aligned}$$

Thus from Maesono(1985), the Edgeworth expansion of the $N^{-1/2}$ term is given by

$$\begin{aligned} \widetilde{\kappa}_{m,n} &= \frac{1}{\tau^3_{m,n}}\left\{\frac{1}{m^2} d^3_{1,N} E(U^3_{1,m}) - \frac{1}{n^2} d^3_{2,N} E(W^3_{1,n})\right.\\ &\left. + \frac{6}{mn} d_{1,N} d_{2,N} d_{3,N} E(U^2_{1,m}) E(W^2_{1,n}) + \frac{6}{n^2} d_{2,N} \frac{f(x_0)}{g^3(x_0)} \left[E(W^2_{1,n})\right]^2\right\}. \end{aligned}$$

It follows from (5) that

$$\begin{aligned} E(U^2_{i,m}) &= h^{-1}_{f,m} f(x_0) A^f_{2,0} + o(h^{-1}_{f,m}),\\ E(W^2_{j,n}) &= h^{-1}_{g,n} g(x_0) A^g_{2,0} + o(h^{-1}_{g,n}),\\ E(U^3_{i,m}) &= h^{-2}_{f,m} f(x_0) A^f_{3,0} + o(h^{-2}_{f,m}),\\ E(W^3_{j,n}) &= h^{-2}_{g,n} g(x_0) A^g_{3,0} + o(h^{-2}_{g,n}). \end{aligned}$$

Thus we have

$$\widetilde{\kappa}_{m,n} = \kappa_{m,n} + o(1).$$

This completes the proof of Theorem 3.1.

Acknowledgement

This research is supported by JSPS Grant-in-Aid for Exploratory Research No.15K11995.

References

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.

García-Soidán, P. H., González-Manteiga, W., and Prada-Sánchez, J. (1997). Edgeworth expansions for nonparametric distribution estimation with applications. *Journal of statistical planning and inference*, **65**, 213-231.

Huang, Z. and Maesono, Y. (2014). Edgeworth expansion for kernel estimators of a distribution function. *Bulletin of Informatics and Cybernetics*, **46**, 1-10.

Maesono, Y. (1985). Edgeworth expansion for two-sample U-statistics. *Rep. Fac. Sci. Kagoshima Univ*, **18**, 35-43.

Maesono, Y. (1997). Edgeworth expansions of a studentized U-statistics and a jackknife estimator of variance. *Journal of Statistical Planning and Inference*, **61**, 61-84.

Umeno, S. and Maesono, Y. (2013). Improvement of normal approximation for kernel density estimator. *Bulletin of Informatics and Cybernetics*, **45**, pp.11-24.

Received February 13, 2016
Revised April 5, 2016