IMPROVEMENT OF NORMAL APPROXIMATION FOR KERNEL DENSITY ESTIMATOR

Umeno, Shota
Nissay Information Technology Corporation

Maesono, Yoshihiko
Faculty of Mathematics, Kyushu University

https://doi.org/10.5109/1563528

出版情報：Bulletin of informatics and cybernetics. 45, pp.11-24, 2013-12. Research Association of Statistical Sciences
バージョン：
権利関係：

IMPROVEMENT OF NORMAL APPROXIMATION FOR KERNEL DENSITY ESTIMATOR

by

Shota Umeno and **Yoshihiko Maesono**

Reprinted from the Bulletin of Informatics and Cybernetics
Research Association of Statistical Sciences, Vol.45

FUKUOKA, JAPAN
2013

Bulletin of Informatics and Cybernetics, Vol. 45, 2013

IMPROVEMENT OF NORMAL APPROXIMATION FOR KERNEL DENSITY ESTIMATOR

By

Shota Umeno[*] and **Yoshihiko Maesono**[†]

Abstract

Many papers have studied theoretical properties of a kernel density estimator. Especially mean integrated squared errors are precisely studied. The asymptotic distribution of the estimator is also discussed, and it is easy to show asymptotic normality. In this paper, we will discuss higher order approximation of the distribution of the kernel estimator. We will obtain an Edgeworth expansion, which takes an explicit form. Assuming a bandwidth $h_n = O(n^{-\frac{1}{4}})$, we also prove the validity of the expansion with residual term $o(n^{-\frac{3}{4}})$.

Key Words and Phrases: Kernel estimator, Density function, Edgeworth expansion, Normal approximation.

1. Introduction

Let $X_1, X_2, \cdots, X_n$ be independently and identically distributed (*i.i.d.*) random variables with distribution and density functions $F(x)$, $f(x)$. The kernel type estimator of the density function f is given by

$$\hat{f}_n(x) = \frac{1}{nh_n} \sum_{i=1}^{n} K\Big(\frac{x - X_i}{h_n}\Big)$$

where h_n is a bandwidth parameter and $h_n \to 0 (n \to \infty)$. K is a kernel function which satisfies

$$\int_{-\infty}^{\infty} K(x)dx = 1.$$

The kernel estimator was introduced by Fix and Hodges (1951) and Akaike (1954). Rosenblatt (1956) and Parzen (1962) have discussed basic properties of the estimator. Mean integrated squared errors of the kernel density estimator are discussed in many papers. The best order of the bandwidth which minimizes the mean integrated squared error, is $O(n^{-\frac{1}{5}})$. There are also many papers which studied bias reduction, bandwidth selection, etc. It is easy to show asymptotic normality of a standardized kernel estimator. For improvement of the normal approximation, Hall (1991) has obtained the Edgeworth

[*] Nissay Information Technology Corporation, NISSAY Aroma Square, 5-37-1, Kamata, Ota-ku, Tokyo 144-8721, Japan.

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

expansion for the kernel density estimator with residual term $O(\{nh_n\}^{-\frac{3}{2}} + n^{-1})$. It is possible to obtain a higher order expansion with the general bandwidth h_n. Here we want to discuss an explicit form of the expansion, and so we use the bandwidth with $h_n = cn^{-1/4}$ $(c > 0)$. We also discuss conditions for the validity of the expansion, which are different from those of Hall (1991).

In section 2, we will discuss the asymptotic normality and the Edgeworth expansion. We will also discuss the conditions which ensure the validity of the expansion. In section 3, we will compare the normal approximation and the expansion by simulation.

2. Asymptotic expansion

Since the kernel density estimator is a sample mean of the *i.i.d.* random variables, we have an asymptotic distribution of the satandardized kernel density estimator:

$$P\left(\frac{(\hat{f}_n(x_0) - f(x_0))}{\sqrt{n}\sqrt{Var(\frac{1}{nh_n}K(\frac{x_0 - X_1}{h_n}))}} \le y\right) = \Phi(y) + o(1)$$

where $\Phi(y)$ is a distribution function of the standard normal $N(0,1)$ and $x_0 \in \boldsymbol{R}$ is a fixed value. Here we assume that the kernel is the second order and satisfies

$$0 < \int_{-\infty}^{\infty} K^2(z)dz < \infty.$$

Let us define

$$\begin{aligned}
&\mu_\ell(x_0) = \int_{-\infty}^{\infty} \{k(u)\}^\ell f(x_0 - h_n u)du,\\
&Y_1 = K\Big(\frac{x_0 - X_1}{h_n}\Big) - E\Big[K\Big(\frac{x_0 - X_1}{h_n}\Big)\Big],\\
&Y_2 = K^2\Big(\frac{x_0 - X_1}{h_n}\Big) - E\Big[K^2\Big(\frac{x_0 - X_1}{h_n}\Big)\Big],\\
&\mu_{rs} = h_n^{-1}E[Y_1^r Y_2^s].
\end{aligned}$$

Using $\mu_\ell(x_0)$ and μ_{rs}, which depend on the sample size n, Hall (1991) has discussed the Edgeworth expansion with remainder term $O(\{nh_n\}^{-\frac{3}{2}} + n^{-1})$. Here we obtain explicit forms of these terms and prove the validity of the expansion under different conditions. In order to prove the validity of the expansion, we use the following Esséen's (1945) smoothing lemma. Let $H(x)$ be a bounded non-decreasing function, and $G(x)$ be a differentiable function of bounded variation on the real line and $|G'(x)| \le M < \infty$. Further let $H(\pm\infty) = G(\pm\infty)$. Then for any $T > 0,\ b > \frac{1}{2\pi}$, we have

$$\Delta = \sup_x |H(x) - G(x)| \le b\int_{-T}^{T}\left|\frac{h(t) - g(t)}{t}\right|dt + d(b)\frac{M}{T} \tag{1}$$

where

$$h(t) = \int_{-\infty}^{\infty} e^{itx}dH(x), \qquad g(t) = \int_{-\infty}^{\infty} e^{itx}dG(x)$$

and $d(b)$ depends on b, but not on n.

First we will obtain an Edgeworth expansion of

$$S_n = \frac{\hat{f}_n(x_0) - E[\hat{f}_n(x_0)]}{\sqrt{Var(\hat{f}_n(x_0))}}.$$

For the standardized sample mean of the *i.i.d.* random variables, many papers discussed the approximation of the characteristic function. Let us define

$$Z_{n,j} = \frac{h_n^{-1}K(\frac{x_0-X_j}{h_n}) - h_n^{-1}\int K(\frac{x_0-y}{h_n})f(y)dy}{\sqrt{Var(h_n^{-1}K(\frac{x_0-X_1}{h_n}))}} \quad (\text{for } j = 1, 2, \cdots, n)$$

and

$$\begin{aligned}
P_{n,0}(t) &= 1 \\
P_{n,1}(t) &= \frac{\kappa_{n,3}}{6}t^3, \\
P_{n,2}(t) &= \frac{\kappa_{n,4}}{24}t^4 + \frac{\kappa_{n,3}^2}{72}t^6, \\
P_{n,3}(t) &= \frac{\kappa_{n,5}}{120}t^5 + \frac{\kappa_{n,3}\kappa_{n,4}}{144}t^7 + \frac{\kappa_{n,3}^3}{1296}t^9, \\
\kappa_{n,3} &= \mu_3 - 3\mu_1\mu_2 + 2\mu_1^3, \\
\kappa_{n,4} &= \mu_4 - 3\mu_2^2 - 4\mu_1\mu_3 + 12\mu_1^2\mu_2 - 6\mu_1^4
\end{aligned}$$

where

$$\mu_k = E(Z_{n,j}^k).$$

From the definition of $Z_{n,j}$, we have

$$S_n = \frac{1}{\sqrt{n}}\sum_{j=1}^{n} Z_{n,j}.$$

Let $\varphi_{S_n}(t)$ be a characteristic function of S_n, and then we have

$$\varphi_{S_n}(t) = \left\{ \varphi_{Z_{n,1}}\left(\frac{t}{\sqrt{n}} \right) \right\}^n .$$

Using the evaluation for the characteristic function of the standardized sample mean, we have the following lemma.

LEMMA 2.1. *Assume* $E(|Z_{n,j}|^5) < \infty$. *Then for* $|t| \le T_{5,n} = \frac{\sqrt{n}}{40\mu_5^{3/5}}$, *we have*

$$\left|\varphi_{S_n}(t) - e^{-\frac{t^2}{2}}\left(1 + \sum_{k=1}^{3} P_{n,k}(it)n^{-\frac{k}{2}}\right)\right| \le \frac{\varepsilon(n)}{(T_{5,n})^3}(|t|^5 + |t|^9)e^{-\frac{t^2}{4}} \tag{2}$$

where $\varepsilon(n) \to 0$ *as* $n \to \infty$.

PROOF. See Gnedenko and Kolmogorov (1968).

Inverting this approximation of the characteristic function, we have the following Edgeworth expansion for S_n:

$$G_3(y) = \Phi(y) + n^{-\frac{1}{2}}\phi(y)Q_{n,1}(y) + n^{-1}\phi(y)Q_{n,2}(y) + n^{-\frac{3}{2}}\phi(y)Q_{n,3}(y),$$

$$\begin{aligned} Q_{n,1}(y) &= -\frac{\kappa_{n,3}}{6}H_2(y), \\ Q_{n,2}(y) &= -\frac{\kappa_{n,4}}{24}H_3(y) - \frac{\kappa_{n,3}^2}{72}H_5(y), \\ Q_{n,3}(y) &= -\frac{\kappa_{n,5}}{120}H_5(y) - \frac{\kappa_{n,3}\kappa_{n,4}}{144}H_7(y) - \frac{\kappa_{n,3}^3}{1296}H_9(y) \end{aligned}$$

where $\{H_k(x)\}$ are Hermite polynomials. Further, let us define

$$G_2(y) = \Phi(y) + n^{-\frac{1}{2}}\phi(y)Q_{n,1}(y) + n^{-1}\phi(y)Q_{n,2}(y).$$

It is easy to show that

$$G_3(y) = G_2 + O(n^{-\frac{9}{8}}).$$

We will show that $G_2(y)$ is the Edgeworth expansion of S_n with residual term $o(n^{-\frac{3}{4}})$

Usually, when we prove the validity of the expansion for the sum of *i.i.d.* random variables, we assume the Cramer condition

$$\lim_{|t|\to\infty} |E((itX_j))| < 1.$$

Since $Z_{n,j}$ depends on the bandwidth h_n or n, we cannot apply the previous results, directly. Assuming some regularity conditions for the kernel $K(u)$ and the density $f(x)$, we have the following theorem for the standardized density estimator S_n.

THEOREM 2.2. *Assume that the bandwidth $h_n = cn^{-\frac{1}{4}}$ $(c > 0)$, $f^{(5)}(x)$ exists, $K(u)$ is the second order kernel and $\int_{-\infty}^{\infty} K^\ell(u)du < \infty$, $(\ell = 1, 2, \cdots, 5)$. The kernel function K takes a form*

$$K(u) = \widetilde{K}(u)\chi_{[-1,1]}(u), \tag{3}$$

where $\chi_{[-1,1]}(u)$ is an indicator function. Further, there exist $0 < \alpha < 1, \gamma > 0$ which satisfy

$$P(|X_1 - x_0| \geq h_n) \leq 1 - \frac{\gamma}{n^\alpha} + o(n^{-\alpha}), \tag{4}$$

and for $|t| > \frac{\sqrt{n}}{40\mu_5^{3/5}}$

$$\Big|\int_{|\frac{x_0-y}{h_n}|\leq 1} e^{i\frac{t}{\sqrt{n}}K(\frac{x_0-y}{h_n})} f(y)dy\Big| = o(n^{-\alpha}). \tag{5}$$

Then we have

$$\sup_y \Big|P(S_n \leq y) - G_2(y)\Big| = o(n^{-\frac{3}{4}}).$$

PROOF. See Appendix.

Let us define

$$\bar{f}_n(x_0) = \frac{1}{h_n}\int K(\frac{x_0-y}{h_n})f(y)dy \quad \text{and} \quad b_n = \frac{\bar{f}_n(x_0) - f_n(x_0)}{\sqrt{n}\sqrt{Var(\frac{1}{nh_n}K(\frac{x_0-X_1}{h_n}))}}$$

where b_n is a standardized bias. Since

$$P\left(\frac{\hat{f}_n(x_0) - f_n(x_0)}{\sqrt{n}\sqrt{Var(\frac{1}{nh_n}K(\frac{x_0-X_1}{h_n}))}} \le y\right) = P(S_n \le y - b_n) = G_2(y-b_n) + o(n^{-\frac{3}{4}}),$$

using the Taylor expansion, we have the following theorem.

THEOREM 2.3. *Assume that the conditions of Theorem 2.2 are satisfied. Then we have*

$$P\left(\frac{\hat{f}(x_0) - f(x_0)}{\sqrt{n}\sqrt{Var(\frac{1}{nh_n}K(\frac{x_0-X_1}{h_n}))}} \le y\right) = \Phi(y) + C_1 + C_2 + C_3 + C_4 + o(n^{-\frac{3}{4}}), \qquad (6)$$

where

$$\begin{aligned}
A_{3,-\frac{1}{2}} &= \frac{\int K^3(z)dz}{(f(x_0))^{\frac{1}{2}}(\int K^2(z)dz)^{\frac{3}{2}}},\\
A_{3,\frac{1}{2}} &= \frac{3(f'(x_0)\int zK^2(z)dz + f^2(x_0))\int K^3(z)dz}{2(f(x_0))^{\frac{3}{2}}(\int K^2(z)dz)^{\frac{5}{2}}}\\
&\quad -\frac{f'(x_0)\int zK^3(z)dz + 3f^2(x_0)\int K^2(z)dz}{(f(x_0)\int K^2(z)dz)^{\frac{3}{2}}},\\
A_{4,-1} &= \frac{\int K^4(z)dz}{f(x_0)(\int K^2(z)dz)^2},\\
B_{\frac{5}{2}} &= \frac{\int z^2K(z)f''(x_0)dz}{2\sqrt{f(x_0)\int K^2(z)dz}},\\
B_{\frac{7}{2}} &= -\frac{\int z^3K(z)f^{(3)}(x_0)dz}{6\sqrt{f(x_0)\int K^2(z)dz}}\\
&\quad +\frac{(f'(x_0)\int zK^2(z)dz + f(x_0))\int z^2K(z)f''(x_0)dz}{4(f(x_0)\int K^2(z)dz)^{\frac{3}{2}}},\\
B_{\frac{9}{2}} &= \frac{\int z^4K(z)f^{(4)}(x_0)dz}{24\sqrt{f(x_0)\int K^2(z)dz}}\\
&\quad -\frac{(f'(x_0)\int zK^2(z)dz + f(x_0))\int z^3K(z)f^{(3)}(x_0)dz}{12(f(x_0)\int K^2(z)dz)^{\frac{3}{2}}}\\
&\quad -\frac{f''(x_0)\int z^2K^2(z)dz\int z^2K(z)f''(x_0)dz}{8(f(x_0)\int K^2(z)dz)^{\frac{3}{2}}},
\end{aligned}$$

$$
\begin{aligned}
C_1 &= -\phi(y)n^{\frac{1}{2}}h_n^{\frac{5}{2}}B_{\frac{5}{2}} + \frac{1}{2}\phi'(y)nh_n^5B_{\frac{5}{2}}^2 - \phi(y)n^{\frac{1}{2}}h_n^{\frac{7}{2}}B_{\frac{7}{2}} \\
&\quad -\frac{1}{6}\phi^{(2)}(y)n^{\frac{3}{2}}h_n^{\frac{15}{2}}B_{\frac{5}{2}}^3 + \phi'(y)nh_n^6B_{\frac{5}{2}}B_{\frac{7}{2}} + \frac{1}{24}\phi^{(3)}(y)n^2h_n^{10}B_{\frac{5}{2}}^4 \\
&\quad -\phi(y)n^{\frac{1}{2}}h_n^{\frac{9}{2}}B_{\frac{9}{2}} - \frac{1}{2}\phi^{(2)}(y)n^{\frac{3}{2}}h_n^{\frac{17}{2}}B_{\frac{5}{2}}^2B_{\frac{7}{2}} \\
&\quad -\frac{n^{\frac{5}{2}}h_n^{\frac{25}{2}}}{120}\phi^{(4)}(y)B_{\frac{5}{2}}^5 + \frac{1}{2}\phi'(y)nh_n^7(B_{\frac{7}{2}}^2 + 2B_{\frac{5}{2}}B_{\frac{9}{2}}) \\
&\quad +\frac{1}{6}\phi^{(3)}(y)n^2h_n^{11}B_{\frac{5}{2}}^3B_{\frac{7}{2}} + \frac{n^3h_n^{15}}{720}\phi^{(5)}(y)B_{\frac{5}{2}}^6, \\
C_2 &= -\frac{1}{6}\Big\{n^{-\frac{1}{2}}h_n^{-\frac{1}{2}}\phi(y)(y^2-1)A_{3,-\frac{1}{2}} \\
&\quad -(h_n^2B_{\frac{5}{2}} + h_n^3B_{\frac{7}{2}})A_{3,-\frac{1}{2}}\{\phi'(y)(y^2-1) + 2y\phi(y)\} \\
&\quad +n^{-\frac{1}{2}}h_n^{\frac{1}{2}}\phi(y)(y^2-1)A_{3,\frac{1}{2}} \\
&\quad +n^{\frac{1}{2}}h_n^{\frac{9}{2}}B_{\frac{5}{2}}^2A_{3,-\frac{1}{2}}\{\phi(y) + 2y\phi'(y) + \frac{1}{2}(y^2-1)\phi^{(2)}(y)\} \\
&\quad -h_n^3B_{\frac{5}{2}}A_{3,\frac{1}{2}}\{2y\phi(y) + (y^2-1)\phi'(y)\} \\
&\quad -nh_n^7B_{\frac{5}{2}}^3A_{3,-\frac{1}{2}}\{\phi'(y) + y\phi^{(2)}(y) + \frac{1}{6}(y^2-1)\phi^{(3)}\}\Big\}, \\
C_3 &= -\frac{1}{24}n^{-1}h_n^{-1}\phi(y)(y^3-2y^2)A_{4,-1}, \\
C_4 &= -\frac{1}{72}n^{-1}h_n^{-1}\phi(y)(y^5-10y^3+15y)A_{3,-\frac{1}{2}}^2,
\end{aligned}
$$

and $\phi(y)$ denotes a density function of the standard normal distribution.

PROOF. See Appendix.

If the kernel $K(u)$ is the Epanechnikov (1969) kernel and the density function satisfies regularity conditions, we have the following theorem.

THEOREM 2.4. *Assume that the kernel is the Epanechnikov one and satisfies the conditions of Theorem 2.2. Further we assume that there exists a neighborhood $x_0-\delta \le x \le x_0+\delta$, in which $f(x) \ge a$ for some $a>0,\ \delta>0$. Then the equation (6) in Theorem 2.3 holds.*

PROOF. See Appendix.

If the kernel $K(u)$ is symmetric about the origin and 4-*th* order kernel, the Edgeworth expansion takes the following simple form.

COROLLARY 2.5. *Assume that the conditions of Theorem 2.4, and the kernel $K(u)$ is symmetric about the origin and 4-th order kernel. Then we have the simple form of*

each terms of the Edgeworth expansion as follows:

$$\begin{aligned}
A_{3,-\frac{1}{2}} &= \frac{\int K^3(z)dz}{(f(x_0))^{\frac{1}{2}}(\int K^2(z)dz)^{\frac{3}{2}}},\\
A_{3,\frac{1}{2}} &= \frac{3f^2(x_0)\int K^3(z)dz}{2(f(x_0))^{\frac{3}{2}}(\int K^2(z)dz)^{\frac{5}{2}}} - \frac{3f^2(x_0)\int K^2(z)dz}{(f(x_0)\int K^2(z)dz)^{\frac{3}{2}}},\\
A_{4,-1} &= \frac{\int K^4(z)dz}{f(x_0)(\int K^2(z)dz)^2},\\
B_{\frac{5}{2}} &= B_{\frac{7}{2}} = 0,\\
B_{\frac{9}{2}} &= \frac{\int z^4K(z)f^{(4)}(x_0)dz}{24\sqrt{f(x_0)\int K^2(z)dz}},\\
C_1 &= -\phi(y)n^{\frac{1}{2}}h_n^{\frac{9}{2}}B_{\frac{9}{2}},\\
C_2 &= -\frac{1}{6}\Big\{n^{-\frac{1}{2}}h_n^{-\frac{1}{2}}\phi(y)(y^2-1)A_{3,-\frac{1}{2}} + n^{-\frac{1}{2}}h_n^{\frac{1}{2}}\phi(y)(y^2-1)A_{3,\frac{1}{2}}\Big\},\\
C_3 &= -\frac{1}{24}n^{-1}h_n^{-1}\phi(y)(y^3-2y^2)A_{4,-1},\\
C_4 &= -\frac{1}{72}n^{-1}h_n^{-1}\phi(y)(y^5-10y^3+15y)A_{3,-\frac{1}{2}}^2.
\end{aligned}$$

3. Simulation

In this section, we will compare the simple normal approximation and the Edgeworth expansion by simulation. Here we use the Epanechnikov kernel

$$K(u) = \frac{3}{4}(1-u^2)\chi_{[-1,1]}(u)$$

with bandwidth $h_n = n^{-\frac{1}{4}}$. In the tables, "True" means an estimate of

$$P\left(\frac{(\hat{f}(x_0)-f(x_0))}{\sqrt{n}\sqrt{Var(\frac{1}{nh_n}K(\frac{x_0-X_1}{h_n}))}} \le y\right)$$

based on $1,000,000$ replication of the sample sets $\{x_1,\ldots,x_n\}$. Table 1 denotes the results of the comparison when x_0 is 5% quantile of the standard normal distribution $N(0,1)$.

Table 1. (Normal, x_0 : 5% quantile)

y	Normal	Edgeworth	True
-2.5	0.006209665	**-0.003506837**	0.00000
-2	0.022750132	**-0.004460385**	0.00000
-1.5	0.066807201	**0.030039624**	0.00000
-1	**0.158655254**	0.139079764	0.16776
-0.5	**0.308537539**	0.319048671	0.30392
0	0.500000000	**0.519034876**	0.53508
0.5	0.691462461	**0.694662961**	0.69585
1	0.841344746	**0.829284154**	0.82807
1.5	0.933192799	**0.916289906**	0.91041
2	0.977249868	**0.960553435**	0.95716
2.5	0.993790335	**0.980211679**	0.98201

Similarly, Table 2 denotes the results when x_0 is 25% quantile of $N(0,1)$.

Table 2. (Normal, x_0 : 25% quantile)

y	Normal	Edgeworth	True
-2.5	0.006209665	**0.003716433**	0.00143
-2	0.022750132	**0.017044958**	0.01331
-1.5	0.066807201	**0.063640227**	0.06137
-1	0.158655254	**0.168916606**	0.16682
-0.5	0.308537539	**0.334351260**	0.33662
0	0.500000000	**0.528575325**	0.53068
0.5	0.691462461	**0.710160457**	0.71300
1	0.841344746	**0.849451459**	0.84582
1.5	**0.933192799**	0.935266886	0.92844
2	0.977249868	**0.976106958**	0.97148
2.5	0.993790335	**0.991382658**	0.99016

Tables 3 and 4 denote the results when x_0 are 5% and 25% quantiles of the χ^2-distribution with 2 degrees of freedom.

Table 3. (χ^2, x_0 : 5% quantile)

y	Normal	Edgeworth	True
-2.5	0.006209665	**-0.001857202**	0.00000
-2	0.022750132	**0.006498395**	0.00000
-1.5	0.066807201	**0.059326133**	0.00000
-1	**0.158655254**	0.189813229	0.16937
-0.5	0.308537539	**0.382560488**	0.39447
0	0.500000000	**0.579719202**	0.54887
0.5	0.691462461	**0.741135940**	0.73580
1	**0.841344746**	0.860013287	0.84644
1.5	**0.933192799**	0.935735257	0.92448
2	0.977249868	**0.972805897**	0.96403
2.5	0.993790335	**0.987253348**	0.98469

Table 4. (χ^2, x_0 : 25% quantile)

y	Normal	Edgeworth	True
-2.5	0.006209665	**0.0003166831**	0.00000
-2	0.022750132	**0.0091356232**	0.00000
-1.5	0.066807201	**0.0544712728**	0.05068
-1	0.158655254	**0.1667606621**	0.17876
-0.5	0.308537539	**0.3425255064**	0.33485
0	0.500000000	**0.5393735618**	0.53925
0.5	0.691462461	**0.7146763830**	0.71463
1	**0.841344746**	0.8472006341	0.84124
1.5	0.933192799	**0.9303698805**	0.92115
2	0.977249868	**0.9713245879**	0.96650
2.5	0.993790335	**0.9878142460**	0.98661

From the above simulation study, we can see that the Edgeworth expansion improves the normal approximation in most cases.

4. Appendix

Proof of Theorem 2.2

Let $F_{S_n}(y)$ be a distribution function of S_n. Putting $F(y) = F_{S_n}(y), G(y) = G_3(y)$ and $T = n^{\frac{3}{4}} \log n$, we apply the Esseen's smoothing lemma. Let us define

$$\varphi_{S_n}(t) = \int_{-\infty}^{\infty} e^{ity} dF_{S_n(y)}$$

and

$$\psi(t) = e^{-\frac{t^2}{2}} \left(1 + \sum_{k=1}^{3} P_{n,k}(it) n^{-\frac{k}{2}} \right) = \int_{-\infty}^{\infty} e^{ity} dG_3(y).$$

Using the proof of the validity for the Edgeworth expansion of the sample mean (See Gnedenko and Kolmogorov (1968)), we have

$$\begin{aligned} &\sup_y |F_{S_n}(y) - G_3(y)| \\ &= \int_{|t| \le T_{5,n}} \left| \frac{\varphi_{S_n}(t) - \psi(t)}{t} \right| dt + \int_{T_{5,n} < |t| \le T} \left| \frac{\varphi_{S_n(t)} - \psi(t)}{t} \right| dt + o(n^{-\frac{3}{4}}). \end{aligned}$$

It is easy to show that

$$\begin{aligned} \mu_5 &= E(Z_{n,1}^5) = O((n^2 h_n)^{\frac{5}{2}}) \left\{ \frac{1}{n^5 h_n^5} \left(h_n \int K^5(z) f(x_0 - h_n z) dy + O(h_n^2) \right) \right\} \\ &= O(n^{\frac{3}{8}}). \end{aligned}$$

Then, from the usual evaluation for the sample mean, we can show that

$$\int_{|t| \le T_{5,n}} \left| \frac{\varphi_{T_n}(t) - \psi(t)}{t} \right| dt = o(n^{-\frac{3}{4}}).$$

Note that

$$\int_{T_{5,n}<|t|\leq T}\left|\frac{\varphi_{S_n}(t)-\psi(t)}{t}\right|dt\leq\int_{T_{5,n}<|t|\leq T}\left|\frac{\varphi_{S_n}(t)}{t}\right|dt+\int_{T_{5,n}<|t|\leq T}\left|\frac{\psi(t)}{t}\right|dt.$$

It is easy to show that

$$\int_{T_{5,n}<|t|\leq T}\left|\frac{\psi(t)}{t}\right|dt=o(n^{-\frac{3}{4}}).$$

It follows from the assumption (5) of Theorem 2.2 that

$$\begin{aligned}|\varphi_{Z_{n,1}}(t)| &= \Big|\int_{|\frac{x_0-y}{h_n}|\leq 1}e^{i\frac{t}{\sqrt{n}}K(\frac{x_0-y}{h_n})}f(y)dy+\int_{|\frac{x_0-y}{h_n}|\geq 1}e^{i\frac{t}{\sqrt{n}}K(\frac{x_0-y}{h_n})}f(y)dy\Big|\\ &\leq \Big|\int_{|\frac{x_0-y}{h_n}|\geq 1}e^{i\frac{t}{\sqrt{n}}K(\frac{x_0-y}{h_n})}f(y)dy\Big|+o(n^{-\alpha}).\end{aligned}$$

Further, using the inequality (4), we can show that

$$\begin{aligned}&\Big|\int_{|\frac{x_0-y}{h_n}|\geq 1}e^{i\frac{t}{\sqrt{n}}K(\frac{x_0-y}{h_n})}f(y)dy\Big|\leq\Big|\int_{|\frac{x_0-y}{h_n}|\geq 1}f(y)dy\Big|\\ =\;& P\Big(|\frac{x_0-X_1}{h_n}|\geq 1\Big)=P(|X_1-x_0|\geq h_n)\leq 1-\frac{c}{n^\alpha}+o(n^{-\alpha}).\end{aligned}$$

Using the above evaluations, we have

$$\begin{aligned}|\varphi_{S_n}(t)| &\leq \Big|1-\frac{c}{n^\alpha}+o(n^{-\alpha})\Big|^n\\ &= \left\{\Big(1-\frac{c}{n^\alpha}+o(n^{-\alpha})\Big)^{n^\alpha}\right\}^{n^{1-\alpha}}=O(e^{-cn^{1-\alpha}}).\end{aligned}$$

Thus, under the assumptions of Theorem 2.2, we can show that

$$\int_{T_{5,n}<|t|\leq T}\left|\frac{\varphi_{S_n}(t)}{t}\right|dt=o(n^{-\frac{3}{4}}),$$

and so

$$\sup_y\left|P\left(\frac{(\hat{f}_n(x_0)-\bar{f}_n(x_0))}{\sqrt{n}\sqrt{Var(\frac{1}{nh_n}K(\frac{x_0-X_1}{h_n}))}}\leq y\right)-G_3(y)\right|=o(n^{-\frac{3}{4}}).$$

Proof of Theorem 2.3

At first, we will obtain an approximation of the variance $Var(h_n^{-1}K(\frac{x_0-X_1}{h_n}))$. Using

a transformation $z = \frac{x_0 - y}{h_n}$ and Taylor expansion, we can easily show that

$$\begin{aligned} & Var\Big(h_n^{-1}K\Big(\frac{x_0 - X_1}{h_n}\Big)\Big) \\ = \; & h_n^{-1}\int K^2(z)f(x_0 - h_n z)dz - \Big(\int K(z)f(x_0 - h_n z)dz\Big)^2 \\ = \; & h_n^{-1}\int K^2(z)\{f(x_0) - h_n z f'(x_0) + \frac{1}{2}h_n^2 z f''(x_0) + O(h_n^3))\}dz \\ & -(f(x_0) + O(h_n^2))^2 \\ = \; & h_n^{-1}f(x_0)\int K^2(z)dz - f'(x_0)\int zK^2(z)dz - f^2(x_0) \\ & + \frac{h_n f''(x_0)}{2}\int z^2K^2(z)dz + O(h_n^2). \end{aligned}$$

Furthermore, using the Taylor expansion

$$\frac{1}{(x+a)^{\frac{3}{2}}} = \frac{1}{a^{\frac{3}{2}}} - \frac{3x}{2a^{\frac{5}{2}}} + \cdots,$$

we have

$$\begin{aligned} & \frac{1}{(Var(\frac{1}{h_n}K(\frac{x_0 - X_1}{h_n})))^{\frac{3}{2}}} \\ = \; & h_n^{\frac{3}{2}}\left[\frac{1}{\{f(x_0)\int K^2(z)dz\}^{\frac{3}{2}}} + \frac{3h_n\{f'(x_0)\int zK^2(z)dz + f^2(x_0)\}}{2\{f(x)\int K^2(z)dz\}^{\frac{5}{2}}} + O(h_n^2)\right]. \end{aligned}$$

Similarly, we can show that

$$\begin{aligned} & E\Big[\Big\{h_n^{-1}K(\frac{x_0 - X_i}{h_n}) - h_n^{-1}\int K(\frac{x_0 - y}{h_n})f(y)dy\Big\}^3\Big] \\ = \; & h_n^{-2}\int K^3(z)f(x_0 - h_n z)dz - 3h_n^{-1}\int K^2(z)f(x_0 - h_n z)dz\int K(z)f(x_0 - h_n z)dz \\ & + 2\Big(\int K(z)f(x_0 - h_n z)dz\Big)^3. \end{aligned}$$

Combining these evaluations, we get

$$E(Z_{n,i}^3) = h_n^{-\frac{1}{2}}A_{3,-\frac{1}{2}} + h_n^{\frac{1}{2}}A_{3,\frac{1}{2}} + O(h_n^{\frac{3}{2}}).$$

Using the same method, we can obtain

$$E(Z_{n,i}^4) = h_n^{-1}A_{4,-1} + O(1).$$

Since

$$\begin{aligned} \bar{f}_n(x_0) - f(x_0) \; = \; & \frac{1}{2}h_n^2\int z^2K(z)f''(x_0)dz - \frac{1}{6}h_n^3\int z^3K(z)f^{(3)}(x_0)dz \\ & + \frac{1}{24}h_n^4\int z^4K(z)f^{(4)}(x_0)dz + O(h_n^5), \end{aligned}$$

we can show that

$$b_n = n^{\frac{1}{2}} h_n^{\frac{5}{2}} B_{\frac{5}{2}} + n^{\frac{1}{2}} h_n^{\frac{7}{2}} B_{\frac{7}{2}} + n^{\frac{1}{2}} h_n^{\frac{9}{2}} B_{\frac{9}{2}} + O(n^{-\frac{7}{8}}).$$

Note that

$$\begin{aligned} & G_2(y) \\ = \; & \Phi(y) - \frac{1}{6} n^{-\frac{1}{2}} \phi(y)(y^2-1)\{h_n^{-\frac{1}{2}} A_{3,-\frac{1}{2}} + h_n^{\frac{1}{2}} A_{3,\frac{1}{2}}\} \\ & -\frac{1}{24} n^{-1} h_n^{-1} \phi(y)(y^3-3y) A_{4,-1} \\ & -\frac{1}{72} n^{-1} \phi(y)(y^5-10y^3+15y)\{h_n^{-1} A_{3,-\frac{1}{2}}^2 + 2A_{3,-\frac{1}{2}} A_{3,\frac{1}{2}}\} + O(n^{-\frac{7}{8}}) \end{aligned}$$

and

$$P\left(\frac{(\hat{f}(x_0) - f(x_0))}{\sqrt{n}\sqrt{Var(\frac{1}{nh_n} K(\frac{x_0 - X_1}{h_n}))}} \le y \right) = G_2(y - b_n) + o(n^{-\frac{3}{4}}).$$

First, using the Taylor expansion, we have

$$\begin{aligned} & \Phi(y - b_n) \\ = \; & \Phi(y) - \phi(y) n^{\frac{1}{2}} h_n^{\frac{5}{2}} B_{\frac{5}{2}} + \frac{1}{2}\phi'(y) n h_n^5 B_{\frac{5}{2}}^2 - \left(\phi(y) n^{\frac{1}{2}} h_n^{\frac{7}{2}} B_{\frac{7}{2}} + \frac{1}{6}\phi^{(2)}(y) n^{\frac{3}{2}} h_n^{\frac{15}{2}} B_{\frac{5}{2}}^3 \right) \\ & + \left(\phi'(y) n h_n^6 B_{\frac{5}{2}} B_{\frac{7}{2}} + \frac{1}{24}\phi^{(3)}(y) n^2 h_n^{10} B_{\frac{5}{2}}^4 \right) \\ & - \left(\phi(y) n^{\frac{1}{2}} h_n^{\frac{9}{2}} B_{\frac{9}{2}} + \frac{1}{2}\phi^{(2)}(y) n^{\frac{3}{2}} h_n^{\frac{17}{2}} B_{\frac{5}{2}}^2 B_{\frac{7}{2}} + \frac{n^{\frac{5}{2}} h_n^{\frac{25}{2}}}{120}\phi^{(4)}(y) B_{\frac{5}{2}}^5 \right) \\ & + \left(\frac{1}{2}\phi'(y) n h_n^7 (B_{\frac{7}{2}}^2 + 2B_{\frac{5}{2}} B_{\frac{9}{2}}) + \frac{1}{6}\phi^{(3)}(y) n^2 h_n^{11} B_{\frac{5}{2}}^3 B_{\frac{7}{2}} + \frac{n^3 h_n^{15}}{720}\phi^{(5)}(y) B_{\frac{5}{2}}^6 \right) \\ & + O(n^{-\frac{7}{8}}) \\ = \; & \Phi(y) + C_1 + O(n^{-\frac{7}{8}}). \end{aligned}$$

Similarly, we can show that

$$\begin{aligned} & -\frac{1}{6} n^{-\frac{1}{2}} \phi(y - b_n)\{(y-b_n)^2 - 1\}(h_n^{-\frac{1}{2}} A_{3,-\frac{1}{2}} + h_n^{\frac{1}{2}} A_{3,\frac{1}{2}}) \\ & = C_2 + O(n^{-\frac{7}{8}}), \\ & -\frac{1}{24} n^{-1} \phi(-b_n)\{(y-b_n)^3 - 2(y-b_n)^2\} h_n^{-1} A_{4,-1} \\ & = C_3 + O(n^{-\frac{7}{8}}), \\ & -\frac{1}{72} n^{-1} \phi(y - b_n)\{(y-b_n)^5 - 10(y-b_n)^3 + 15(y-b_n)\}(h_n^{-1} A_{3,-\frac{1}{2}}^2 + 2A_{3,-\frac{1}{2}} A_{3,\frac{1}{2}}) \\ & = C_4 + O(n^{-\frac{7}{8}}). \end{aligned}$$

Thus we have

$$P\left(\frac{(\hat{f}(x_0) - f(x_0))}{\sqrt{n}\sqrt{Var(\frac{1}{nh_n} K(\frac{x_0 - X_1}{h_n}))}} \le y \right) = \Phi(y) + C_1 + C_2 + C_3 + C_4 + o(n^{-\frac{3}{4}}).$$

Proof of Theorem 2.4
Since the Epanechnikov kernel takes a form

$$K(u) = \frac{3}{4}(1-u^2)\chi_{[-1,1]}(u),$$

the condition (3) is satisfied. For sufficiently large n, we have $0 \le h_n \le \delta$. Thus we can show that

$$P(|X_1 - x_0| \ge h_n) = 1 - \int_{x_0-h_n}^{x_0+h_n} f(y)dy \le 1 - 2ah_n,$$

and the condition (4) is satisfied. Further, we get

$$\begin{aligned}
\varphi_{Z_{n,1}}(\frac{t}{\sqrt{n}}) &= \int_{-\infty}^{\infty} \exp\left(i\frac{3t}{4\sqrt{n}}\Big(1 - \frac{(x_0-y)^2}{h_n^2}\Big) I\Big(\Big|\frac{x_0-y}{h_n}\Big| \le 1\Big)\right) f(y)dy \\
&= e^{\frac{3it}{4\sqrt{n}}} \left\{ \int_{x_0-h_n}^{x_0+h_n} \exp\left(-\frac{3it(x_0-y)^2}{4\sqrt{n}h_n^2}\right) f(y)dy + P(|X_1-x_0| \ge h_n) \right\},
\end{aligned}$$

and

$$\begin{aligned}
&\int_{x_0-h_n}^{x_0+h_n} \exp\left(-\frac{3it(x_0-y)^2}{4\sqrt{n}h_n^2}\right) f(y)dy \\
= &\int_{x_0-h_n}^{x_0+h_n} \cos\Big(\frac{3t(x_0-y)^2}{4\sqrt{n}h_n^2}\Big) f(y)dy - i\int_{x_0-h_n}^{x_0+h_n} \sin\Big(\frac{3t(x_0-y)^2}{4\sqrt{n}h_n^2}\Big) f(y)dy. \quad (7)
\end{aligned}$$

For the first term of (7), we have

$$\begin{aligned}
\int_{x_0-h_n}^{x_0+h_n} \cos\Big(\frac{3t(x_0-y)^2}{4\sqrt{n}h_n^2}\Big) f(y)dy &= h_n\int_{-1}^{1} \cos(\frac{3tu^2}{4\sqrt{n}}) f(h_nu+x_0)du \\
&= h_n\int_{-1}^{1} \cos(\frac{3tu^2}{4\sqrt{n}})(f(x_0)+O(h_n))du \\
&= 2h_nf(x_0)\sqrt{\frac{\sqrt{n}}{3t}} \int_{-\sqrt{\frac{3t}{4\sqrt{n}}}}^{\sqrt{\frac{3t}{4\sqrt{n}}}} \cos(w^2)dw + O(h_n^2) \\
&\le 2h_nf(x_0)\sqrt{\frac{\sqrt{n}}{3t}}\sqrt{\frac{\pi}{2}} + O(h_n^2).
\end{aligned}$$

Therefore for $t > \frac{\sqrt{n}}{40\mu_5^{\frac{3}{5}}} = O(n^{\frac{29}{40}})$, we can show

$$\int_{x_0-h_n}^{x_0+h_n} \cos\Big(\frac{3t(x_0-y)^2}{4\sqrt{n}h_n^2}\Big) f(y)dy = o(h_n)$$

and

$$\int_{x_0-h_n}^{x_0+h_n} \sin\Big(\frac{3t(x_0-y)^2}{4\sqrt{n}h_n^2}\Big) f(y)dy = o(h_n).$$

Combining the above evaluations, we get

$$|\varphi_{z_{n,1}}(t)| = o(h_n),$$

and the condition (5) is satisfied.

Acknowledgement

This research was supported by JSPS Grant-in-Aid for Scientific Research (B) No.21340026 and Exploratory Research No.24650151.

References

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

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

Esséen, C. G. (1945). Fourier analysis of distribution functions: A mathematical study of the Laplace-Gaussian law. *Acta Mathematica,* 77, 1-125

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

Gnedenko, B.V. and Kolmogorov, A.N. (1968). *Limit Distributions for Sums of Independent Random Variables.* Addison-Wesley.

Hall, P. (1991). Edgeworth expansions for nonparametric density estimators, with applications. *Statistics*, 22, 215-232

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

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

Received November 25, 2012
Revised February 25, 2013