EDGEWORTH EXPANSION FOR KERNEL ESTIMATORS OF A DISTRIBUTION FUNCTION

Huang, Zhong
Graduate School of Mathematics, Kyushu University

Maesono, Yoshihiko
Faculty of Mathematics, Kyushu University

https://doi.org/10.5109/1798142

出版情報：Bulletin of informatics and cybernetics. 46, pp.1-10, 2014-12. Research Association of Statistical Sciences
バージョン：
権利関係：

EDGEWORTH EXPANSION FOR KERNEL ESTIMATORS OF A DISTRIBUTION FUNCTION

by

Zhong Huang and **Yoshihiko Maesono**

Reprinted from the Bulletin of Informatics and Cybernetics
Research Association of Statistical Sciences, Vol.46

FUKUOKA, JAPAN
2014

Bulletin of Informatics and Cybernetics, Vol. 46, 2014

EDGEWORTH EXPANSION FOR KERNEL ESTIMATORS OF A DISTRIBUTION FUNCTION

By

Zhong Huang* and **Yoshihiko Maesono**†

Abstract

Many papers have studied theoretical properties of a kernel type estimator of a distribution function. Especially mean 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 $h_n = o(n^{-c})$ $(\frac{1}{4} \le c < \frac{1}{2})$, we obtain the explicit form of the expansion with residual term $o(n^{-1})$. We also discuss a bias term precisely.

Key Words and Phrases: Kernel estimator, Distribution function, Edgeworth expansion, Normal approximation, Bias reduction.

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 $f(x_0)$ is

$$\widehat{f}_n(x_0) = \frac{1}{nh_n} \sum_{i=1}^{n} K\left(\frac{x_0 - X_i}{h_n}\right)$$

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 of the distribution function $F(x_0)$ is given by

$$\widehat{F}_n(x_0) = \frac{1}{n} \sum_{i=1}^{n} W\left(\frac{x_0 - X_i}{h_n}\right)$$

where

$$W(t) = \int_{-\infty}^{t} K(u)du.$$

* 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

Mean squared errors and asymptotic normality are precisely studied by many papers. Azzalini (1981) proved that $h_n = O(n^{-1/3})$ attained a minimum mean squared error.

Garsía-Soidán et al. (1997) have obtained Edgeworth expansions of standardized and studentized estimators $\widehat{F}_n(x_0)$, and proved validity of them. Residual terms of the expansions are $O(n^{1/2}h_n^3 + h_n^2 + n^{-1/2}h_n)$. They have also discussed an bias estimator which includes an consistent estimator of $f'(x_0)$. In this paper, we will obtain an explicit form of the expansion with residual terms $o(n^{-1})$.

For the kernel K, let us assume the following conditions. Hereafter, for the sake of simplicity, we use $\int$ which means $\int_{-\infty}^{\infty}$.

$$\text{(k1)} \qquad \int K(z)dz = 1,$$

$$\text{(k2)} \qquad \int zK(z)dz = 0,$$

$$\text{(k3)} \qquad \int z^\ell K(z)dz < \infty \qquad (\ell = 2, 3, 4).$$

The kernel estimator of the distribution F was introduced by Nadaraya (1964), and showed that its asymptotic mean and variance are same as the empirical distribution. Under some regularity conditions, we can easily show the asymptotic normality of the estimator $\widehat{F}_n(x_0)$.

In section 2, we will discuss the asymptotic normality and the Edgeworth expansion. After obtaining an explicit form of the bias term, we will give the expansion, which enable us to make a confidence interval of $F(x_0)$ in section 3. In section 4, we will compare the normal approximation and the expansion by simulation.

2. Asymptotic expansion

Since the kernel estimator of the distribution function is a sample mean of the *i.i.d.* random variables, we have an asymptotic distribution of the estimator. If the bandwidth $h_n = o(n^{-1/4})$ and the conditions (k1) $\sim$ (k3) are satisfied, it is easy to show that

$$P\left(\frac{\sqrt{n}[\widehat{F}_n(x_0) - F(x_0)]}{\sqrt{Var\left[W\left(\frac{x_0 - X_1}{h_n}\right)\right]}} \leq y \right) = \Phi(y) + o(1)$$

where $\Phi(y)$ is a distribution function of the standard normal $N(0,1)$ and $x_0 \in \boldsymbol{R}$ is a fixed value. At first we will discuss the Edgeworth expansion for the standardized

$\widehat{F}_n(x_0)$. Let us define

$$\begin{aligned} W_i &= W\left(\frac{x_0 - X_i}{h_n}\right) \\ \sigma_n^2 &= Var(W_1) \\ \kappa_{3,n} &= \frac{E[\{W_1 - E(W_1)\}^3]}{\sigma_n^3} \\ \kappa_{4,n} &= \frac{E[\{W_1 - E(W_1)\}^4]}{\sigma_n^4} \\ Q_{1,n}(y) &= -\frac{\kappa_{3,n}}{6} H_2(y), \\ Q_{2,n}(y) &= -\frac{\kappa_{4,n}}{24} H_3(y) - \frac{\kappa_{3,n}^2}{72} H_5(y) \end{aligned}$$

where $\{H_k(y)\}$ are Hermite polynomials

$$\begin{aligned} H_2(y) &= y^2 - 1, \\ H_3(y) &= y^3 - 3y, \\ H_5(y) &= y^5 - 10y^3 - 15y. \end{aligned}$$

Then using Lemma 3.1 of Garsía-Soidán et al. (1997), we have the following theorem.

THEOREM 2.1. *Assume that f' exists and is continuous on a neighborhood of x_0, $h_n = cn^{-d}$ ($c > 0$, $\frac{1}{4} \le d < \frac{1}{2}$) and the conditions $(k1) \sim (k3)$ are satisfied. Then we have*

$$P\left(\frac{\sqrt{n}\{\widehat{F}_n(x_0) - E[\widehat{F}_n(x_0)]\}}{\sigma_n} \le y\right) = P_n(y) + o(n^{-1})$$

where

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

PROOF. Since the estimator $\widehat{F}_n$ is the sample mean of $\{W_i\}$, we can obtain the formal Edgeworth expansion. Instead of the Cramer condition, we can apply Lemma 3.1 of Garsía-Soidán et al. (1997), and prove the validity of the expansion.

Next we will obtain approximations of the moments of W_1. Using a transformation $u = W(z)$, we get

$$\int_{-\infty}^{\infty} W(z)K(z)dz = \frac{1}{2},$$
$$\int_{-\infty}^{\infty} W^2(z)K(z)dz = \frac{1}{3},$$
$$\int_{-\infty}^{\infty} W(z)^3K(z)dz = \frac{1}{4}.$$

Then, changing the variable, it follows from (k1) $\sim$ (k3) that

$$
\begin{aligned}
E(W_1) &= \int W\left(\frac{x_0-y}{h_n}\right) f(y)dy \\
&= h_n \int W(z) f(x_0 - h_n z)dz \\
&= [-W(z)F(x_0 - h_n z)]_{-\infty}^{+\infty} + \int K(z)F(x_0 - h_n z)dz \\
&= \int K(z)F(x_0 - h_n z)dz \\
&= \int K(z)\{F(x_0) - h_n z f(x_0) + O(h_n^2)\}dz \\
&= F(x_0)\int K(z)dz - h_n f(x_0)\int zK(z)dz + O(h_n^2) \\
&= F(x_0) + O(h_n^2).
\end{aligned}
$$

Let us define

$$
A_{i,j} = \int W^i(z) z^j K(z)dz. \tag{1}
$$

Similarly, for the second moment, we have

$$
\begin{aligned}
E(W_1^2) &= \int W^2(\frac{x_0-y}{h_n}) f(y)dy \\
&= h_n \int W^2(z) f(x_0 - h_n z)dz \\
&= 2\int W(z)K(z)F(x_0 - h_n z)dz \\
&= 2\int W(z)K(z)\{F(x_0) - h_n z f(x_0) + O(h_n^2)\}dz \\
&= F(x_0) - 2h_n f(x_0)A_{1,1} + O(h_n^2)
\end{aligned}
$$

and for the third and fourth moment, we get

$$
\begin{aligned}
E(W_1^3) &= 3\int W^2(z)K(z)F(x_0 - h_n z)dz \\
&= F(x_0) - 3h_n f(x_0)A_{2,1} + O(h_n^2), \\
E(W_1^4) &= F(x_0) + O(h_n).
\end{aligned}
$$

Combing the above evaluations, we can get the approximations of the cumulants $\kappa_{3,n}$ and $\kappa_{4,n}$. Using the Taylor expansion of $(x+a)^{-3/2}$ and $(x+a)^{-2}$, it is easy to see that

$$\begin{aligned}\sigma_n^2 &= Var(W_1) = E(W_1^2) - \{E(W_1)\}^2 \\ &= F(x_0) - 2h_n f(x_0)A_{1,1} - \{F(x_0)\}^2 + O(h_n^2) \\ &= F(x_0)\{1-F(x_0)\} - 2h_n f(x_0)A_{1,1} + O(h_n^2),\end{aligned}$$

$$\begin{aligned}\sigma_n^{-3/2} &= \frac{1}{[F(x_0)\{1-F(x_0)\}]^{3/2}} + h_n \frac{3f(x_0)A_{1,1}}{[F(x_0)\{1-F(x_0)\}]^{5/2}}, \\ \sigma_n^{-2} &= \frac{1}{[F(x_0)\{1-F(x_0)\}]^2} + O(h_n).\end{aligned}$$

Thus we have the approximations of $\kappa_{3,n}$ and $\kappa_{4,n}$. Since

$$\begin{aligned}& E[\{W_1 - E(W_1)\}^3] = E(W_1^3) - 3E(W_1^2)E(W_1) + 2\{E(W_1)\}^3 \\ = \;& F(x_0)\{1-F(x_0)\}\{1-2F(X_0)\} + 3h_n f(x_0)\{2F(x_0)A_{1,1} - A_{2,1}\} + O(h_n^2),\end{aligned}$$

and

$$E[\{W_1 - E(W_1)\}^4] = F(x_0)\{1-F(x_0)\}\{1-3F(x_0)+3F^2(x_0)\} + O(h_n),$$

we get

$$\begin{aligned}\kappa_{3,n} &= \frac{E[\{W_1 - E(W_1)\}^3]}{\sigma_n^3} \\ &= \frac{1-2F(x_0)}{[F(x_0)\{1-F(x_0)\}]^{1/2}} + \frac{3f(x_0)(A_{1,1}-A_{2,1})}{[F(x_0)\{1-F(x_0)\}]^{3/2}} + O(h_n^2) \\ &= B_{3,0} + h_n B_{3,1} + O(h_n^2),\end{aligned}$$

$$\kappa_{4,n} = \frac{E[\{W_1 - E(W_1)\}^4]}{\sigma_n^4} = B_{4,0} + O(h_n)$$

where

$$B_{3,0} = \frac{1-2F(x_0)}{[F(x_0)\{1-F(x_0)\}]^{1/2}}, \qquad B_{3,1} = \frac{3f(x_0)(A_{1,1}-A_{2,1})}{[F(x_0)\{1-F(x_0)\}]^{3/2}}$$
$$B_{4,0} = \frac{1-3F(x_0)+3F^2(x_0)}{F(x_0)\{1-F(x_0)\}}.$$

Using these approximations, we have the following theorem.

THEOREM 2.2. *Assume that f' exists and is continuous on a neighborhood of x_0, and $h_n = cn^{-d}$ $(c > 0,\ \frac{1}{4} \le d < \frac{1}{2})$. Then we have*

$$P\left(\frac{\sqrt{n}\{\widehat{F}_n(x_0) - E[\widehat{F}_n(x_0)]\}}{\sigma_n} \le y\right) = \widetilde{P}_n(y) + o(n^{-1})$$

where

$$\widetilde{P}_n(y) = \Phi(y) - n^{-1/2}\phi(y)\widetilde{Q}_1(y) - n^{-1/2}h_n\phi(y)\widetilde{Q}_1^*(y) - n^{-1}\phi(y)\widetilde{Q}_2(y)$$

and

$$\widetilde{Q}_1(y) = \frac{B_{3,0}}{6}H_2(y), \quad \widetilde{Q}_1^*(y) = \frac{B_{3,1}}{6}H_2(y), \quad \widetilde{Q}_2(y) = \frac{B_{4,0}}{24}H_3(y) - \frac{B_{3,0}^2}{72}H_5(y).$$

3. Asymptotic representation of bias

In order to construct a confidence interval of $F(x_0)$, we have to obtain an Edgeworth expansion of

$$P\left(\frac{\sqrt{n}\{\widehat{F}_n(x_0) - F(x_0)\}}{\sigma_n} \le y\right).$$

Let us define the bias term

$$\Delta_n = \frac{n^{1/2}\{E[\widehat{F}_n(x_0)] - F(x_0)\}}{\sigma_n}.$$

Then we have

$$\begin{aligned} & P\left(\frac{\sqrt{n}\{\widehat{F}_n(x_0) - F(x_0)\}}{\sigma_n} \le y\right) \\ = \;& P\left(\frac{\sqrt{n}\{\widehat{F}_n(x_0) - E[\widehat{F}_n(x_0)]\}}{\sigma_n} \le y - \Delta_n\right) = \widetilde{P}_n(y - \Delta_n). \end{aligned}$$

Since $h_n = cn^{-d}$ $(c > 0,\ \frac{1}{4} \le d < \frac{1}{2})$, $\Delta_n = O(1)$.

If the density function has bounded 5-*th* derivative $f^{(5)}$, it follows from $(k2)$ that

$$\begin{aligned} & E[\widehat{F}_n(x_0)] - F(x_0) \\ = \;& \frac{h_n^2}{2}f'(x_0)A_{0,2} - \frac{h_n^3}{6}f''(x_0)A_{0,3} + \frac{h_n^4}{24}f^{(3)}(x_0)A_{0,4} - \frac{h_n^5}{120}f^{(4)}(x_0)A_{0,5} + O(h_n^6). \end{aligned}$$

Similarly, using Taylor expansion, we have

$$\begin{aligned} E(W_1) &= F(x_0) + \frac{h_n^2}{2}f'(x_0)A_{0,2} - \frac{h_n^3}{6}f''(x_0)A_{0,3} + O(h_n^4), \\ E(W_1^2) &= F(x_0) - 2h_nf(x_0)A_{1,1} + h_n^2f'(x_0)A_{1,2} - \frac{h_n^3}{3}f''(x_0) + O(h_n^4), \end{aligned}$$

and then

$$\begin{aligned} \sigma_n^2 &= F(x_0)\{1 - F(x_0)\} - 2h_nf(x_0)A_{1,1} + h_n^2f'(x_0)\{A_{1,2} - F(x_0)A_{0,2}\} \\ & \quad - \frac{h_n^3}{3}f''(x_0)\{A_{1,3} - F(x_0)A_{0,3}\} + O(h_n^4). \end{aligned}$$

Further, we can get

$$\begin{aligned}\sigma_n^{-1} &= \frac{1}{[F(x_0)\{1-F(x_0)\}]^{1/2}} + h_n \frac{f(x_0)A_{1,1}}{[F(x_0)\{1-F(x_0)\}]^{3/2}} \\ &\quad + h_n^2 \left(-\frac{f'(x_0)\{A_{1,2} - F(x_0)A_{0,2}\}}{2[F(x_0)\{1-F(x_0)\}]^{3/2}} \frac{3f^2(x_0)A_{1,1}^2}{2[[F(x_0)\{1-F(x_0)\}]^{5/2}} \right) \\ &\quad + h_n^3 \left(\frac{f''(x_0)\{A_{1,3} - F(x_0)A_{0,3}\}}{6[F(x_0)\{1-F(x_0)\}]^{3/2}} - \frac{3f(x_0)f'(x_0)A_{1,1}\{A_{1,2} - F(x_0)A_{0,2}\}}{2[F(x_0)\{1-F(x_0)\}]^{5/2}} \right. \\ &\qquad \left. + \frac{5f^3(x_0)A_{1,1}^3}{2[F(x_0)\{1-F(x_0)\}]^{7/2}} \right) + O(h_n^4).\end{aligned}$$

Combining the above evaluations, we have the asymptotic representation for $h_n = cn^{-d}$ $(c > 0,\ \frac{1}{4} \le d < \frac{1}{2})$

$$n^{-1/2}\Delta_n = h_n^2 b_2 + h_n^3 b_3 + h_n^4 b_4 + h_n^5 b_5 + o(n^{-3/2}) \tag{2}$$

where

$$\begin{aligned} b_2 &= \frac{f'(x_0)A_{0,2}}{2[F(x_0)\{1-F(x_0)\}]^{1/2}}, \\ b_3 &= -\frac{f''(x_0)A_{0,3}}{6[F(x_0)\{1-F(x_0)\}]^{1/2}} + \frac{f(x_0)f'(x_0)A_{1,1}A_{0,2}}{2[F(x_0)\{1-F(x_0)\}]^{3/2}}, \\ b_4 &= \frac{f^{(3)}(x_0)A_{0,4}}{24[F(x_0)\{1-F(x_0)\}]^{1/2}} \\ &\quad - \frac{2f(x_0)f''(x_0)A_{1,1}A_{0,3} + 3[f'(x_0)]^2 A_{0,2}\{A_{1,2} - F(x_0)A_{0,2}\}}{12[F(x_0)\{1-F(x_0)\}]^{3/2}} \\ &\quad + \frac{3[f(x_0)]^2 f'(x_0)A_{1,1}^2 A_{0,2}}{4[F(x_0)\{1-F(x_0)\}]^{5/2}} \end{aligned}$$

and

$$\begin{aligned} b_5 = & \\ & -\frac{f^{(4)}(x_0)A_{0,5}}{120[F(x_0)\{1-F(x_0)\}]^{1/2}} \\ & + \frac{f(x_0)f^{(3)}(x_0)A_{1,1}A_{0,4} + 2f'(x_0)f''(x_0)\{A_{0,3}A_{1,2} + A_{0,2}A_{1,3} - 2F(x_0)A_{0,2}A_{0,3}\}}{24[F(x_0)\{1-F(x_0)\}]^{3/2}} \\ & - \frac{[f(x_0)]^2 f''(x_0)A_{1,1}^2 A_{0,3} + 3f(x_0)[f'(x_0)]^2 A_{0,2}A_{1,1}\{A_{1,2} - F(x_0)A_{0,2}\}}{4[F(x_0)\{1-F(x_0)\}]^{5/2}} \\ & + \frac{5[f(x_0)]^3 f'(x_0)A_{0,2}A_{1,1}^3}{4[F(x_0)\{1-F(x_0)\}]^{7/2}}. \end{aligned}$$

If the kernel is symmetric around 0, we have $A_{0,3} = A_{0,5} = 0$ and then

$$\begin{aligned}
b_2 &= \frac{f'(x_0)A_{0,2}}{2[F(x_0)\{1-F(x_0)\}]^{1/2}}, \\
b_3 &= \frac{f(x_0)f'(x_0)A_{1,1}A_{0,2}}{2[F(x_0)\{1-F(x_0)\}]^{3/2}}, \\
b_4 &= \frac{f^{(3)}(x_0)A_{0,4}}{24[F(x_0)\{1-F(x_0)\}]^{1/2}} - \frac{3[f'(x_0)]^2A_{0,2}\{A_{1,2}-F(x_0)A_{0,2}\}}{12[F(x_0)\{1-F(x_0)\}]^{3/2}} \\
&\quad + \frac{3[f(x_0)]^2f'(x_0)A_{1,1}^2A_{0,2}}{4[F(x_0)\{1-F(x_0)\}]^{5/2}}
\end{aligned}$$

and

$$\begin{aligned}
b_5 &= \frac{f(x_0)f^{(3)}(x_0)A_{1,1}A_{0,4} + 2f'(x_0)f''(x_0)A_{0,2}A_{1,3}}{24[F(x_0)\{1-F(x_0)\}]^{3/2}} \\
&\quad - \frac{3f(x_0)[f'(x_0)]^2A_{0,2}A_{1,1}\{A_{1,2}-F(x_0)A_{0,2}\}}{4[F(x_0)\{1-F(x_0)\}]^{5/2}} \\
&\quad + \frac{5[f(x_0)]^3f'(x_0)A_{0,2}A_{1,1}^3}{4[F(x_0)\{1-F(x_0)\}]^{7/2}}.
\end{aligned}$$

Furthermore, if we use a symmetric and 4-*th* order kernel, that is $A_{0,2} = A_{0,3} = A_{0,5} = 0$, we have a simple form as follows

$$n^{-1/2}\Delta_n = h_n^4\delta_1 + h_n^5\delta_2 + o(n^{-3/2})$$

where

$$\delta_1 = \frac{f^{(3)}(x_0)A_{0,4}}{24[F(x_0)\{1-F(x_0)\}]^{1/2}} \quad \text{and} \quad \delta_2 = \frac{f(x_0)f^{(3)}(x_0)A_{1,1}A_{0,4}}{24[F(x_0)\{1-F(x_0)\}]^{3/2}}.$$

In this case, it is easy to see that

$$\begin{aligned}
&\Phi(y-\Delta_n) = \Phi(y) - \Delta_n\phi(y) + o(n^{-1}) \\
&\qquad\qquad = \Phi(y) - (n^{1/2}h_n^4\delta_1 + n^{1/2}h_n^5\delta_2)\phi(y) + o(n^{-1}), \\
&n^{-1/2}\phi(y-\Delta_n)\widetilde{Q}_1(y-\Delta_n) = n^{-1/2}\phi(y)\widetilde{Q}_1(y) + o(n^{-1}), \\
&n^{-1/2}h_n\phi(y-\Delta_n)\widetilde{Q}_1^*(y-\Delta_n) = n^{-1/2}h_n\phi(y)\widetilde{Q}_1^*(y) + o(n^{-1})
\end{aligned}$$

and

$$n^{-1}\phi(y-\Delta_n)\widetilde{Q}_2(y-\Delta_n) = n^{-1}\phi(y)\widetilde{Q}_2(y) + o(n^{-1}).$$

Thus we get a simple form of the Edgeowrth expansion

$$P\left(\frac{\sqrt{n}\{\widehat{F}_n(x_0) - F(x_0)\}}{\sigma_n} \le y\right) = \widetilde{P}_{4,n}(y) + o(n^{-1})$$

where

$$\begin{aligned}
\widetilde{P}_{4,n}(y) &= \Phi(y) - n^{-1/2}\phi(y)\widetilde{Q}_1(y) - n^{1/2}h_n^4\delta_1\phi(y) - n^{-1/2}h_n\phi(y)\widetilde{Q}_1^*(y) \\
&\quad - h^{1/2}h_n^5\delta_2\phi(y) - n^{-1}\widetilde{Q}_2(y).
\end{aligned}$$

Müller (1984) discussed higher order kernel, and gave the following 4-th order kernel

$$K(u) = \frac{315}{512}(11u^8 - 36u^6 + 42u^4 - 20u^2 + 3)I(|u| \leq 1)$$

where $I(\cdot)$ is an indicator function.

4. Simulation

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

$$K(u) = \frac{3}{4}(1 - u^2)I(|u| \leq 1)$$

with bandwidth $h_n = n^{-\frac{1}{3}}$. In the tables, "True" means an estimate of

$$P\left(\frac{\sqrt{n}\{\widehat{F}_n(x_0) - F(x_0)\}}{\sigma_n} \leq y\right).$$

based on $1,000,000$ replications of the sample sets $\{x_1, \ldots, x_n\}$. Table 1.$\sim$ 3. denote the results of the comparison when $x_0 = 1.645$ and $F(x)$ is the normal distribution.

Table 1. $x_0 = 1.645,\ (n = 20)$

y	Normal	Edgeworth	True
-2.5	0.0062097	**0.0005635**	0.00000
-2	0.0227501	**0.0011811**	0.00001
-1.5	0.0668072	**0.0505320**	0.00002
-1	0.1586553	**0.1872352**	0.19727
-0.5	**0.3085375**	0.3753120	0.33068
0	0.5000000	**0.5421862**	0.55341
0.5	**0.6914625**	0.6780195	0.71257
1	0.8413447	**0.8054994**	0.82570
1.5	0.9331928	**0.9057585**	0.91194
2	0.9772499	**0.9567984**	0.95216
2.5	0.9937903	**0.9755485**	0.98049

Tabel 2. $x_0 = 1.645,\ (n = 50)$

y	Normal	Edgeworth	True
-2.5	0.0062097	**0.0001790**	0.00000
-2	0.0227501	**0.0065772**	0.00001
-1.5	0.0668072	**0.0484223**	0.04581
-1	0.1586553	**0.1569024**	0.15497
-0.5	**0.3085375**	0.3267333	0.31143
0	0.5000000	**0.5148187**	0.52078
0.5	**0.6914625**	0.6839743	0.70333
1	0.8413447	**0.8184529**	0.82738
1.5	0.9331928	**0.9095608**	0.91453
2	0.9772499	**0.9588444**	0.96210
2.5	0.9937903	**0.9813390**	0.98439

Table 3. $x_0 = 1.645$, $(n = 100)$

y	Normal	Edgeworth	True
-2.5	0.0062097	**0.0008896**	0.00000
-2	0.0227501	**0.0098472**	0.00685
-1.5	0.0668072	**0.0500268**	0.04898
-1	0.1586553	**0.1492084**	0.14800
-0.5	0.3085375	**0.3109683**	0.31087
0	0.5000000	**0.5031549**	0.50483
0.5	**0.6914625**	0.6830423	0.69578
1	0.8413447	**0.8227624**	0.83010
1.5	0.9331928	**0.9134912**	0.91947
2	0.9772499	**0.9622814**	0.96439
2.5	0.9937903	**0.9847047**	0.98629

The simulation results when $F(x)$ is χ^2 and Laplace are similar.

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

Acknowledgement

This research was supported by JSPS Grant-in-Aid for Exploratory Research No.24650151.

References

Azzalini, A. (1981). A note on the estimation of a distribution function and quantiles by a kernel method. *Biometrika* 68, 326-328

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

García-Soidán, P.H., González-Manteiga W. and Prada-Sánchez, J.M. (1997). Edgeworth expansions for nonparametric distribution estimation with applications. *Jour. Stat. Plann. Inf.* 65, 213-231

Nadaraya, E.A. (1964). Some new estimates for distribution functions *Theory Prob. Appl.* 15, 497-500

Müller, H.G. (1984). Smooth optimum kernel estimators of densities, regression curves and modes. *Ann. Statist.*, 12, 766-774

Received November 25, 2013
Revised March 25, 2014