

# Random Weighting Estimation of Kernel Density

Shesheng Gao<sup>a</sup> and Yongmin Zhong<sup>b\*</sup>

<sup>a</sup>School of Automatics, Northwestern Polytechnical University, Xi'an 710072, China

<sup>b</sup>Department of Mechanical Engineering, Curtin University of Technology, Australia

## Abstract

Let  $X_1, X_2, \dots, X_n$  be independent and identically distributed random variables with common probability density function  $f(x)$ . The kernel density estimation of  $f(x)$  can be defined as

$f_n(x) = \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right)$ , where  $K(u)$  is a kernel function and  $h_n > 0$  is a series of positive

constants that satisfy  $\lim_{n \rightarrow \infty} h_n = 0$ . A theory is established to approximate kernel density estimation

$f_n(x)$  by using random weighting estimation  $\hat{H}_n(x)$  of  $f(x)$ . Under certain conditions, it

rigorously proves that  $\sqrt{nh_n}(\hat{H}_n(x) - f_n(x))$  and  $\sqrt{nh_n}(f_n(x) - f(x))$  have the same limiting

distribution for any random series  $X_1, X_2, \dots, X_n$ .

**Keywords:** random weighting estimation, kernel density estimation, kernel function, and approximation theory.

## 1. Introduction

The random weighting method is an emerging computing method in statistics [1]. It has received great attention in the recent years, and has been used to solve different problems [1-7]. However, there has been very limited research to use the random weighting method for kernel density estimation. Estimation of kernel density is an important research topic in statistics and computing.

---

\*Corresponding author. Postal address: Department of Mechanical Engineering, Curtin University of Technology, GPO Box U1987, Perth, WA 6845, Australia. Tel: 61 8 92667657 Fax: 61 8 92662681 Email: Y.Zhong@curtin.edu.au

Currently, Bootstrap is a commonly used method for kernel density estimation [8-10]. In comparison with the Bootstrap method, the random weighting method has advantages [1, 4, 5, 11], such as the simplicity in computation, the suitability for large samples, and there is no need to know the distribution function. The random weighting method can also be used to calculate a statistic with a probability density function, since the resultant statistical distribution provides a probability density function.

In this paper, the random weighting method is adopted in the first time to kernel density estimation. A theory is established for random weighting estimation of kernel density. It rigorously proves that kernel density estimation can be precisely approximated by random weighting estimation under certain conditions.

## 2. Theorem

Assume that  $X_1, X_2, \dots, X_n$  are an independent and identically distributed random sample drawn from unknown probability density function  $f(x)$ , and  $K(x)$  is a Borel function in the real domain. Then, the kernel density estimation of  $f(x)$  can be described as

$$f_n(x) = \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right) \quad (1)$$

where  $K(u)$  is a kernel function, and  $h_n > 0$  is a series of positive constants.

We now discuss how to use the random weighting method to approximate kernel density estimation  $f_n(x)$ . Assume that the joint distribution function of random vector  $(V_1, V_2, \dots, V_n)$  obeys Dirichlet distribution  $D(1, 1, \dots, 1)$ , that is,  $\sum_{i=1}^n V_i = 1$  and the joint density function of

$(V_1, V_2, \dots, V_n)$  is  $f(V_1, V_2, \dots, V_n) = \Gamma(n)$ , where  $(V_1, V_2, \dots, V_n) \in D_{n-1}$  and  $D_{n-1} = \{(V_1, V_2, \dots, V_n) : V_i \geq 0, i = 1, \dots, n-1, \sum_{i=1}^{n-1} V_i \leq 1\}$ . Then, the random weighting estimation of  $f(x)$  can be described as

$$\hat{H}_n(x) = \frac{1}{h_n} \sum_{i=1}^n V_i K\left(\frac{x - X_i}{h_n}\right) \quad (2)$$

The proposed random weighting estimation of kernel density aims to use the conditional distribution of  $\sqrt{nh_n}(\hat{H}_n(x) - f_n(x))$  to approximate the distribution of  $\sqrt{nh_n}(f_n(x) - f(x))$ . To achieve this goal, we shall prove that  $\sqrt{nh_n}(\hat{H}_n(x) - f_n(x))$  and  $\sqrt{nh_n}(f_n(x) - f(x))$  have the same limiting distribution for any random series  $X_1, X_2, \dots, X_n$ .

**Theorem** Suppose  $K(u)$  and  $f(x)$  satisfy the following conditions:

- (i)  $f(x) \neq 0$ , and  $f''(x)$  is bounded and continuous at every point on  $(-\infty, +\infty)$ ;
- (ii)  $K(u)$  is bounded on  $(-\infty, +\infty)$ ,  $\int_{-\infty}^{\infty} K(u) du = 1$ ,  $\int_{-\infty}^{\infty} uK(u) du = 0$  and  $\int_{-\infty}^{\infty} u^2 K(u) du < +\infty$ ;
- (iii)  $\lim_{|u| \rightarrow +\infty} uK(u) = 0$  or  $f(x)$  is bounded on  $(-\infty, +\infty)$ ; and
- (iv)  $f(x)$  satisfies the Lipschitz condition of order  $\lambda$ , which implies  $\int_{-\infty}^{\infty} |x|^\lambda f(x) dx < +\infty$ .

Then, choose  $h_n = \left(\frac{(\log n)^{1/2}}{n}\right)^{\frac{1}{1+\lambda}}$  such that

$$\sup_{z \in (-\infty, +\infty)} \left| P^* \left\{ \sqrt{nh_n}(\hat{H}_n(x) - f_n(x)) \leq z \right\} - P \left\{ \sqrt{nh_n}(f_n(x) - f(x)) \leq z \right\} \right| = o \left( \frac{(\log n)^{\frac{1+2\lambda}{4(1+\lambda)}} M_n}{n^{\frac{\lambda}{2(1+\lambda)}}} \right) \quad a.s \quad (3)$$

where  $P$  represents the probability,  $P^*$  represents the conditional probability when  $X_1, X_2, \dots, X_n$  are given, and  $M_n$  is any sufficiently large number that satisfies  $M_n \leq \log \log n$ .

### 3. Proof of Theorem

We introduce the following results, which are to be used in the proof of the theorem.

**Lemma 1.** [12] For  $p > 0$ ,

$$\sup_x |\Phi(x+q) - \Phi(x)| \leq \frac{|q|}{\sqrt{2\pi}} \quad (4)$$

and

$$\sup_x |\Phi(px) - \Phi(x)| \leq \frac{1}{\sqrt{2\pi e}} \left\{ |p-1| + |p^{-1}-1| \right\} \quad (5)$$

where  $\Phi(x)$  is a normal distribution function, and  $q$  is a real number.

**Lemma 2.** [4] Suppose  $K(u)$  and  $f(x)$  satisfy the following conditions: (i)  $K(u)$  is bounded on  $(-\infty, +\infty)$  and  $\int_{-\infty}^{\infty} K(u)du = 1$ ; and (ii)  $\lim_{|u| \rightarrow \infty} uK(u) = 0$  or  $f(x)$  is bounded on  $(-\infty, +\infty)$ .

Let

$$g_n(x) = \frac{1}{h_n} \int_{-\infty}^{\infty} K\left(\frac{u}{h_n}\right) f(x-u) du \quad (6)$$

where  $h_n > 0$  is a series of positive constants, and  $\lim_{n \rightarrow \infty} h_n = 0$ .

Then, when  $f(x)$  is continuous within  $(-\infty, +\infty)$ ,

$$\lim_{n \rightarrow \infty} g_n(x) = f(x) \int_{-\infty}^{\infty} K(u) du = f(x) \quad (7)$$

**Proof of Theorem** For the sake of concise description, we denote  $\sup_z \left\| \cdot \right\|_{\infty}$ .

Considering the left side of (3), the follow expression may be written

$$\begin{aligned} & \left\| P^* \left\{ \sqrt{nh_n} (\hat{H}_n(x) - f_n(x)) \leq z \right\} - P \left\{ \sqrt{nh_n} (f_n(x) - f(x)) \leq z \right\} \right\|_{\infty} \\ & \leq \left\| P^* \left\{ \sqrt{nh_n} (\hat{H}_n(x) - f_n(x)) \leq z \right\} - \Phi \left( \frac{z}{\left( \frac{1}{h_n} V_{ar}^* K \left( \frac{x - X_i}{h_n} \right) \right)^{1/2}} \right) \right\|_{\infty} \\ & \quad + \left\| P \left\{ \sqrt{nh_n} (f_n(x) - Ef_n(x)) \leq z + \sqrt{nh_n} (f(x) - Ef_n(x)) \right\} \right. \\ & \quad \left. - \Phi \left( \frac{z + \sqrt{nh_n} (f(x) - Ef_n(x))}{\left( \frac{1}{h_n} V_{ar} K \left( \frac{x - X_i}{h_n} \right) \right)^{1/2}} \right) \right\|_{\infty} \quad (8) \\ & \quad + \left\| \Phi \left( \frac{z}{\left( \frac{1}{h_n} V_{ar}^* K \left( \frac{x - X_i}{h_n} \right) \right)^{1/2}} \right) - \Phi \left( \frac{z + \sqrt{nh_n} (f(x) - Ef(x))}{\left( \frac{1}{h_n} V_{ar} K \left( \frac{x - X_i}{h_n} \right) \right)^{1/2}} \right) \right\|_{\infty} \\ & = L_1 + L_2 + L_3 \end{aligned}$$

where  $E$  and  $V_{ar}$  represent the expectation and variance of random sample  $X_1, X_2, \dots, X_n$ , respectively, and  $V_{ar}^*$  represents the conditional variance when  $X_1, X_2, \dots, X_n$  are given.

$L_1$  may be further written as

$$L_1 = \left\| P^* \left\{ (\hat{H}_n(x) - f_n(x)) \leq \frac{z}{\sqrt{nh_n}} \right\} + \frac{1}{\sqrt{nh_n}} \Phi \left( \frac{z}{\left( \frac{1}{h_n} \text{Var}^* K \left( \frac{x - X_i}{h_n} \right) \right)^{1/2}} \right) \right\|_{\infty} \quad (9)$$

According to Theorem 1 in [5], it is known that  $\hat{H}_n(x) - f_n(x)$  converges in probability to 0 when  $n \rightarrow \infty$ , under the condition that the first-order moment of  $x$  is finite. Consequently,

$$L_1 = o \left( \frac{1}{(nh_n)^{1/2}} \right) \quad (10)$$

Similar to the derivation process of  $L_1$ , we can also get

$$L_2 = o \left( \frac{1}{(nh_n)^{1/2}} \right) \quad (11)$$

Considering  $L_3$ , the following expression may be written

$$\begin{aligned}
L_3 &\leq \left\| \left\| \Phi \left( \frac{z}{\left( \frac{1}{h_n} \text{Var}^* K \left( \frac{x-X_i}{h_n} \right) \right)^{1/2}} \right) - \Phi \left( \frac{z}{\left( \frac{1}{h_n} \text{Var} K \left( \frac{x-X_i}{h_n} \right) \right)^{1/2}} \right) \right\| \right\|_{\infty} \\
&\quad + \left\| \left\| \Phi \left( \frac{z}{\left( \frac{1}{h_n} \text{Var} K \left( \frac{x-X_i}{h_n} \right) \right)^{1/2}} \right) - \Phi \left( \frac{z + \sqrt{nh_n} (f(x) - Ef(x))}{\left( \frac{1}{h_n} \text{Var} K \left( \frac{x-X_i}{h_n} \right) \right)^{1/2}} \right) \right\| \right\|_{\infty} \\
&= L_3^{(1)} + L_3^{(2)}
\end{aligned} \tag{12}$$

By Lemma 1,

$$\begin{aligned}
L_3^{(1)} &\leq \frac{1}{\sqrt{2\pi e}} \left\{ \left| \frac{\left( \text{Var} K \left( \frac{x-X_i}{h_n} \right) \right)^{1/2}}{\text{Var}^* K \left( \frac{x-X_i}{h_n} \right)} - 1 \right| + \left| \frac{\text{Var}^* K \left( \frac{x-X_i}{h_n} \right)^{1/2}}{\text{Var} K \left( \frac{x-X_i}{h_n} \right)} - 1 \right| \right\} \\
&= \frac{1}{\sqrt{2\pi e}} \left\{ \frac{\left| \frac{1}{h_n} \text{Var}^* K \left( \frac{x-X_i}{h_n} \right) - \frac{1}{h_n} \text{Var} K \left( \frac{x-X_i}{h_n} \right) \right|}{\left( \frac{1}{h_n} \text{Var}^* K \left( \frac{x-X_i}{h_n} \right) \right)^{1/2} \left( \frac{1}{h_n} \text{Var} K \left( \frac{x-X_i}{h_n} \right) \right)^{1/2}} \right\}
\end{aligned} \tag{13}$$

Denote

$$a_n^2(x) = \frac{1}{h_n} EK^2 \left( \frac{x-X_i}{h_n} \right) - h_n \left( \frac{1}{h_n} EK \left( \frac{x-X_i}{h_n} \right) \right)^2 \tag{14}$$

and

$$\tilde{a}_n^2(x) = \frac{1}{h_n} \sum_{i=1}^n V_i K^2 \left( \frac{x - X_i}{h_n} \right) - h_n \left( \frac{1}{h_n} \sum_{i=1}^n V_i K \left( \frac{x - X_i}{h_n} \right) \right)^2 \quad (15)$$

It is noted from the conditions of Theorem that  $K(u)$  is bounded on  $(-\infty, +\infty)$  and  $\int_{-\infty}^{\infty} K(u) du = 1$ , and  $\lim_{|u| \rightarrow \infty} uK(u) = 0$  or  $f(x)$  is bounded on  $(-\infty, +\infty)$ . Therefore, according to Lemma 2, when  $n \rightarrow \infty$ ,

$$\frac{1}{h_n} \text{Var} K \left( \frac{x - X_i}{h_n} \right) = a_n^2(x) \rightarrow f(x) \int_{-\infty}^{\infty} K^2(u) du \quad (16)$$

and

$$\frac{1}{h_n} \text{Var}^* K \left( \frac{x - X_i}{h_n} \right) = \tilde{a}_n^2(x) \rightarrow f(x) \int_{-\infty}^{\infty} K^2(u) du \quad a.s \quad (17)$$

Since  $K(u)$  is bounded on  $(-\infty, +\infty)$ , there is a positive real number  $M$  such that  $|K(u)| \leq M$ .

Thus,

$$\begin{aligned} & \left| \frac{1}{h_n} \text{Var}^* K \left( \frac{x - X_i}{h_n} \right) - \frac{1}{h_n} \text{Var} K \left( \frac{x - X_i}{h_n} \right) \right| \\ & \leq \left| \frac{1}{h_n} \sum_{i=1}^n V_i \left( K^2 \left( \frac{x - X_i}{h_n} \right) - EK^2 \left( \frac{x - X_i}{h_n} \right) \right) \right| \\ & \quad + 2M \left| \frac{1}{h_n} \sum_{i=1}^n V_i \left( K \left( \frac{x - X_i}{h_n} \right) - EK \left( \frac{x - X_i}{h_n} \right) \right) \right| \end{aligned} \quad (18)$$



Using Bernstein's inequality and noticing that  $h_n = \left( \frac{(\log n)^{1/2}}{n} \right)^{\frac{1}{1+\lambda}}$ , we get

$$\sum_{h=N}^{\infty} P \left\{ \left| \frac{1}{h} \sum_{i=1}^n V_i \left( K^2 \left( \frac{x - X_i}{h_n} \right) - EK^2 \left( \frac{x - X_i}{h_n} \right) \right) \right| \geq \varepsilon \left( \frac{(\log n)^{\frac{1+2\lambda}{4(1+\lambda)}}}{n^{\frac{\lambda}{2(1+\lambda)}}} - M_n \right) \right\} < \infty \quad (19)$$

Therefore, by the Borel-Cantelli Lemma,

$$\left| \frac{1}{h_n} \sum_{i=1}^n V_i \left( K^2 \left( \frac{x - X_i}{h_n} \right) - EK^2 \left( \frac{x - X_i}{h_n} \right) \right) \right| = o \left( \frac{(\log n)^{\frac{1+2\lambda}{4(1+\lambda)}}}{n^{\frac{\lambda}{2(1+\lambda)}}} - M_n \right) \quad a.s \quad (20)$$

Similar to the derivation of (20), we can also get

$$\left| \frac{1}{h_n} \sum_{i=1}^n V_i \left( K \left( \frac{x - X_i}{h_n} \right) - EK \left( \frac{x - X_i}{h_n} \right) \right) \right| = o \left( \frac{(\log n)^{\frac{1+2\lambda}{4(1+\lambda)}}}{n^{\frac{\lambda}{2(1+\lambda)}}} - M_n \right) \quad a.s \quad (21)$$

From (13), (20) and (21), we get

$$L_3^{(1)} = o \left( \frac{(\log n)^{\frac{1+2\lambda}{4(1+\lambda)}}}{n^{\frac{\lambda}{2(1+\lambda)}}} - M_n \right) \quad a.s \quad (22)$$

By Lemma 1,

$$L_3^{(2)} \leq \frac{1}{\sqrt{2\pi}} \frac{\sqrt{nh_n} |f(x) - Ef_n(x)|}{\left( \frac{1}{h_n} \text{Var}K \left( \frac{x - X_i}{h_n} \right) \right)^{1/2}} \quad (23)$$

where

$$\begin{aligned} |f(x) - Ef_n(x)| &= \left| f(x) - \frac{1}{h_n} \int K_k \left( \frac{x-u}{h_n} \right) f(u) du \right| \\ &= \left| f(x) - \int K_k(v) f(x - h_n v) dv \right| \\ &= \left| \int [(f(x) - f(x - h_n v)) - K_k(v)] dv \right| \\ &\leq Ch_n^\lambda \int |v|^\lambda K_k(v) dv \end{aligned} \quad (24)$$

Since

$$\frac{1}{h_n} \text{Var}K \left( \frac{x - X_i}{h_n} \right) = a_n^2(x) \rightarrow f(x) \int_{-\infty}^{\infty} K^2(u) du, \quad h_n = \left( \frac{(\log n)^{1/2}}{n} \right)^{\frac{1}{1+\lambda}} \quad (25)$$

we have

$$L_3^{(2)} = o \left( n^{\frac{1}{2}} h_n^{\frac{1}{2} + \lambda} \right) = o \left( \frac{(\log n)^{\frac{1+2\lambda}{4(1+\lambda)}}}{n^{\frac{\lambda}{(1+\lambda)}}} \right) \quad (26)$$

By (12), (22) and (26), we get

$$L_3 = o\left(\frac{(\log n)^{\frac{1+2\lambda}{4(1+\lambda)}}}{n^{\frac{\lambda}{2(1+\lambda)}}} - M_n\right) \quad a.s \quad (27)$$

Noticing that  $h_n = \left(\frac{(\log n)^{1/2}}{n}\right)^{\frac{1}{1+\lambda}}$ , (3) follows from (8), (10), (11) and (27).

The proof of Theorem is completed.

#### 4. Conclusions

This paper adopts the random weighting method in the first time to kernel density estimation. The contribution of the paper is that a theory of random weighting estimation is established for estimation of kernel density. The established theory has extensive applications in many fields such as navigation, signal processing and image processing. Future research work will focus on applications of the theory in engineering practices.

#### Acknowledgements

The work of this paper is supported by the Aeronautical Science Foundation (China) (Project no. 20080818004) and the Natural Science Foundation of Shaanxi Province (China) (Project no. SJ08F04).

#### References

1. S. Gao, Z. Feng, Y. Zhong, B. Shirinzadeh, Random weighting estimation of parameters in generalized Gaussian distribution, *Information Sciences*, Vol. 178, No. 9, 2008, pp2275-2281.
2. A. Barvinok, A.Samorodnitsky, Random weighting, asymptotic counting, and inverse isoperimetry, *Israel Journal of Mathematics*, Vol. 158, No. 1, 2007, pp159-191.

3. Y. Fang, L. Zhao, Approximation to the distribution of LAD estimators for censored regression by random weighting method, *Journal of Statistical Planning and Inference*, Vol. 136, No. 4, 2006, pp1302-1316.
4. S. Gao, Z. Zhang, B. Yang, The random weighting estimation of quantile process, *Information Sciences*, Vol. 164, No. 1-4, 2004, pp139-146.
5. S. Gao, J. Zhang, T. Zhou, Large numbers law for sample mean of random weighting estimation, *Information Sciences*, Vol. 155, No. 1-2, 2003, pp151-156.
6. X.-Y. Wu, Y.-N. Yang, L.-C. Zhao, Approximation by random weighting method for M-test in linear models, *Science in China Series A: Mathematics*, Vol. 50, No. 1, 2007, pp87-99.
7. L. Xue, L. Zhu, L1-norm estimation and random weighting method in a semiparametric model, *Acta Mathematicae Applicatae Sinica*, Vol. 21, No 2, 2005, pp295-30.
8. C. C. Taylor, Bootstrap choice of the smoothing parameter in kernel density estimation, *Biometrika*, Vol. 76, No. 4, 1989, pp705-712.
9. P. Deheuvels, G. Derzko, Asymptotic certainty bands for kernel density estimators based upon a Bootstrap resampling scheme, *Statistical Models and Methods for Biomedical and Technical Systems*, F. Vonta, M. Nikulin, N. Limnios and C. Huber-Carol (*eds.*), Springer-Verlag: New York, 2008, pp171-186.
10. K. Ziegler, On local bootstrap bandwidth choice in kernel density estimation, *Statistics and Decisions*, Vol. 24, No. 2, 2006, pp291-301.
11. Z. Zheng, Random weighting method, *Acta Mathematicae Applicatae Sinica*, Vol. 10, No. 2, 1987, pp247-253.
12. P. Ganssler, W. Stute, *Wahrscheinlichkeitstheorie (Probability Theory)*, Springer-Verlag: Berlin, 1977, pp105-106.