Title	On the estimation of the probability density of a distribution function
Sub Title	
Author	有本, 彰雄(Arimoto, Akio)
Publisher	慶応義塾大学工学部
Publication year	1972
Jtitle	Keio engineering reports Vol.25, No.9 (1972.) ,p.113- 121
JaLC DOI	
Abstract	In this paper we shall deal with nonparametric estimation of a probability density function f(x), restricting to the kernel type estimator which is the linear smoothing version of the empirical distribution function based on a finite sample of independent observations, each distributed according to f(x). Such estimation has been suggested by several authors, e.g., WHITTLE (1958) and PARZEN (1962). We shall, however, use the estimators with slightly more general kernels than those used by Parzen. We give here some variations of the general convergence theorem in Fourier analysis. Using these, we evaluate the rapidity of convergence to the value of f(x) at each point x of the estimate and calculate the order of unbiasedness and consistency of the estimate. Furthermore, we give a method of getting a good estimate in this connection.
Notes	
Genre	Departmental Bulletin Paper
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO50001004-00250009-0113

KEIO ENGINEERING REPORTS
VOL. 25 NO. 9

ON THE ESTIMATION OF THE PROBABILITY DENSITY OF A DISTRIBUTION FUNCTION

BY

AKIO ARIMOTO

FACULTY OF ENGINEERING
KEIO UNIVERSITY
YOKGHAMA 1972

KEIO ENGINEERING REPORTS
VOL. 25, NO. 9, pp. 113-121, 1972

ON THE ESTIMATION OF THE PROBABILITY DENSITY OF A DISTRIBUTION FUNCTION

AKIO ARIMOTO

Dept. of Administration Engineering, Keio University, Yokohama 223, Japan*

(Received Oct. 5, 1972)

ABSTRACT

In this paper we shall deal with nonparametric estimation of a probability density function $f(x)$, restricting to the kernel type estimator which is the linear smoothing version of the empirical distribution function based on a finite sample of independent observations, each distributed according to $f(x)$. Such estimation has been suggested by several authors, e.g., WHITTLE (1958) and PARZEN (1962). We shall, however, use the estimators with slightly more general kernels than those used by Parzen.

We give here some variations of the general convergence theorem in Fourier analysis. Using these, we evaluate the rapidity of convergence to the value of $f(x)$ at each point x of the estimate and calculate the order of unbiasedness and consistency of the estimate. Furthermore, we give a method of getting a good estimate in this connection.

1. Introduction

Let

$$X_1(\omega), X_2(\omega), \cdots, X_n(\omega) \tag{1.1}$$

be n independently and identically distributed random variables with a common distribution function $F(x)$. We desire to estimate the probability density $f(x)$ corresponding to $F(x)$ from the knowledge of (1.1) for $\omega=\omega_0$. But we shall restrict our attention to the estimation of $f(x)$ in terms of the empirical distribution function.

We begin with some definitions. Let us be given (1.1) and write

$$\begin{aligned} Y_k(x, \omega)&=1, \quad \text{if} \quad X_k(\omega)<x, \\ &=0, \quad \text{if} \quad X_k(\omega)\geqq x, \\ &k=1, 2, 3, \cdots, n \end{aligned} \tag{1.2}$$

* Now at Musashi Institute of Technology, Dept. of Industrial Engineering.

$$F_n(x, \omega) = \frac{1}{n} \sum_{k=1}^{n} Y_k(x, \omega). \tag{1.3}$$

This random variable is the so-called empirical distribution function. We then have

$$EF_n(x, \omega) = F(x). \tag{1.4}$$

Now using (1.3), we construct the estimate of $f(x)$:

$$\begin{aligned} p_{n,\lambda}(x, \omega) &= \int_{-\infty}^{\infty} K_\lambda(x-u)dF_n(u, \omega) \\ &= \frac{1}{n} \sum_{k=1}^{n} K_\lambda(x - X_k(\omega)), \end{aligned} \tag{1.5}$$

where $K_\lambda(x)$ is a Borel measurable bounded function such that

$$\int_{-\infty}^{\infty} K_\lambda(x)dx = 1. \tag{1.6}$$

Then, we have

$$Ep_{n,\lambda}(x, \omega) = \int_{-\infty}^{\infty} K_\lambda(x-u)dF(u), \tag{1.7}$$

$$\text{var } p_{n,\lambda}(x, \omega) = \frac{1}{n} \int_{-\infty}^{\infty} K_\lambda^2(x-u)dF(u) - \frac{1}{n} \left\{ \int_{-\infty}^{\infty} K_\lambda(x-u)dF(u) \right\}^2. \tag{1.8}$$

PARZEN (1962) shows that if $\lambda = \lambda_n$ and $\lambda_n \to \infty$ as $n \to \infty$, and $K_\lambda(x) = \lambda K(\lambda x)$, then $p_n(x, \omega) = p_{n,\lambda_n}(x, \omega)$ is an asymptotically unbiased estimate of $f(x)$, that is,

$$Ep_n(x, \omega) = f(x) + O(1) \tag{1.10}$$

for each x, and if $\lambda_n/n \to 0$ as $n \to \infty$, then this is a consistent estimate of $f(x)$, that is,

$$\lim_{n\to\infty} \text{ var } p_n(x, \omega) = 0 \tag{1.11}$$

for each x. More precisely

$$\frac{n}{\lambda_n} \text{ var } p_n(x, \omega) = f(x) \int_{-\infty}^{\infty} K^2(x)dx + O(1) \tag{1.12}$$

holds for large n.

What we are going to do is to discuss about the order of the remainder term $O(1)$ in (1.10) and (1.12).

If, besides the existence of derivatives, the Lipschitz's conditions of derivatives are assumed, then the more precise result than the PARZEN's with respect to the order of consistency can be obtained. PARZEN has shown that if $f(x)$ has the ν-th derivative, then the order of consistency of the estimate is $n^{2\nu/(1+2\nu)}$. If, in addition, $f^{(\nu)}(x)$ satisfies the Lipschitz's condition of order q, where $0<q<1$, then we can show that the order of consistency of the estimate is $n^{2(\nu+q)/(1+2(\nu+q))}$.

In this connection, WAHBA (1971) has shown that if $f^{(\nu+1)}(x) \in L^2(-\infty, \infty)$, then the order of consistency is $n^{(2\nu+1)/2(\nu+1)}$, where he used the polynomial algorithm based on ordinary (Lagrange) interpolation as the method of estimation.

2. Main Results

We first show the following convergence theorem.

THEOREM 1. *Suppose that for some* $\alpha>0$

(i)

$$|K_\lambda(u)|\leqq C_1\lambda, \quad \text{for} \quad |u|\leqq\frac{1}{\lambda}, \tag{2.1}$$

$$\leqq\frac{C_2}{\lambda^\alpha|u|^{1+\alpha}}, \quad \text{for} \quad |u|\geqq\frac{1}{\lambda}, \tag{2.2}$$

for large $\lambda>0$, *some constants* C_1, C_2, *that*

(ii) $F(x)$ *is a bounded, non-decreasing function and has no singular component and that*

(iii) *for any* $\varepsilon>0$, *there exists* $\delta>0$ *such that*

$$\int_{-h}^{h}|d\{F(x+u)-f(x)u\}|\leqq\varepsilon h^{1+\beta} \tag{2.3}$$

for $\beta\geqq0$, $|h|>\delta$ *at a point point* x, *where* $f(x)=F'(x)$ *exists. Then we have*

$$\begin{aligned} R_\lambda&=\int_{-\infty}^{\infty}K_\lambda(x-u)dF(u)-f(x)\int_{-\infty}^{\infty}K_\lambda(u)du \\ &=O\left(\frac{1}{\lambda^\beta}\right), \quad \textit{for} \quad \alpha>\beta, \\ &=O\left(\frac{1}{\lambda^\alpha}\right), \quad \textit{for} \quad \alpha\leqq\beta. \end{aligned} \tag{2.4}$$

It is noted that the condition (2.3) will be thought of as an analogue of the Lipschitz's condition and (2.3) with $\beta=0$ always holds even if $F(x)$ has the singular component (Van der VAART 1967; KAWATA 1972).

The proof of Theorem 1 will be carried out in a way just similar to the one given in KAWATA loc. cit. (Also see Van der VAART loc. cit.) and so it is omitted here.

We also have the following:

COROLLARY 1. *Under the conditions of Theorem* 1, *we have, for* $0\leqq\beta\leqq1$,

$$\begin{aligned} Q_\lambda&=\int_{-\infty}^{\infty}|K_\lambda(x-u)|^2dF(u)-f(x)\int_{-\infty}^{\infty}|K_\lambda(u)|^2du \\ &=O(\lambda^{1-\beta}), \quad \textit{for large } \lambda. \end{aligned} \tag{2.5}$$

Now we say an estimate $p_{n,\lambda}(x,\omega)$ to be asymptotically unbiased of the order λ^{α} if

$$(2.6)\qquad \lim_{\lambda\to\infty}\lambda^{\alpha}|Ep_{n,\lambda}(x,\omega)-f(x)|=\gamma<\infty$$

and to possess an asymptotic variance σ^2 of the order λ^{α} if

$$(2.7)\qquad \lim_{\lambda\to\infty}\lambda^{2\alpha}E|p_{n,\lambda}(x,\omega)-Ep_{n,\lambda}(x,\omega)|^2=\sigma^2.$$

In the case where (2.6) and (2.7) hold simultaneously, we agree to say $p_{n,\lambda}(x,\omega)$ to be consistent of the order $\lambda^{2\alpha}$ with an asymptotic mean square error $\sigma^2+\gamma^2$ (Parzen 1956). With these terminologies we have the following:

Corollary 2. *If the conditions* (i)~(iii) *of Theorem* 1 *are satisfied and*

$$(2.8)\qquad \int_{-\infty}^{\infty}K_{\lambda}(u)du=1,$$

then

$$(2.9)\qquad p_n(x,\omega)=\int_{-\infty}^{\infty}K_{\lambda_n}(x-u)dF_n(u,\omega)$$

is an asymptotic unbiased estimate of the order $\lambda_n^{\theta}(\theta=\min(\alpha,\beta))$ *for the probability density function* $f(x)$.

Moreover if $\lambda_n=O(n^{1-\gamma})$ *as* $n\to\infty$, $0<\gamma<1$, *then* $p_n(x,\omega)$ *is a consistent estimate of the order* n^{γ}, *whenever* $n^{\gamma}=O(\lambda_n^{2\theta})$ *and* $\theta>0$. *Thus we may take* $\lambda_n\sim n^{-(2\theta+1/1)}$.

The latter half of Corollary 2 is easily shown from the facts that for some constant $C_1>0$,

$$(2.10)\qquad \operatorname{var}\ p_n(x,\omega)\leqq\frac{C_1}{n}\left\{f(x)\int_{-\infty}^{\infty}K_{\lambda_n}^2(x)dx+Q_{\lambda_n}\right\},$$

$$(2.11)\qquad \int_{-\infty}^{\infty}K_{\lambda}^2(x)dx=O(\lambda),$$

and the maximan of γ such that $n^{\gamma}\leqq C_2n^{2(1-\gamma)\theta}$ for some $C_2>0$ is $2\theta/(2\theta+1)$.

Furthermore, in order to improve the order of asymptotic unbiasedness, we need the existence of the derivatives of $f(x)$. In this connection we shall prove the following.

Theorem 2. *Suppose that*

(i)

$$(2.12)\qquad \int_{-\infty}^{\infty}|K_{\lambda}(u)|du<\infty,$$

$$(2.13)\qquad \int_{-\infty}^{\infty}|u^rK_{\lambda}(u)|du\leqq\frac{C_1}{\lambda^r},$$

for some constant C_1 *and some* $r\geqq0$,

(ii) *$f(u)$ is bounded ($|f(u)|\leqq M$) and has the ν-th derivative at $u=x$, where $0\leqq\nu\leqq[r]$ and*

(iii) *there exists a $\delta>0$ such that*

$$|f(u)-g(u,x)|\leqq C|u-x|^{q+\nu}(0\leqq q\leqq 1) \tag{2.14}$$

for $|u-x|\leqq\delta$, where C is a constant independent of u and $g(u,x)$ is defined by

$$g(u,x)=\sum_{k=0}^{\nu}\frac{f^{(k)}(x)}{k!}(u-x)^k. \tag{2.15}$$

Then, we have

$$\begin{aligned} R_\lambda&=\int_{-\infty}^{\infty}\{f(x+u)-g(x+u,x)\}K_\lambda(u)du \\ &=O\left(\frac{1}{\lambda^r}\right), \quad \text{for} \quad r\leqq q+\nu, \\ &=O\left(\frac{1}{\lambda^{\nu+q}}\right), \quad \text{for} \quad r>q+\nu, \end{aligned} \tag{2.16}$$

for large λ.

Proof. Write

$$\begin{aligned} R_\lambda&=\int_{|u|\leqq\delta}\{f(x+u)-g(x+u,x)\}K_\lambda(u)du \\ &-\int_{|u|>\delta}g(x+u,x)K_\lambda(u)du+\int_{|u|>\delta}f(x+u)K_\lambda(u)du \\ &=I_1+I_2+I_3, \end{aligned} \tag{2.17}$$

say. From (2.14)

$$|I_1|\leqq C\int_{|u|\leqq\delta}|u^{q+\nu}K_\lambda(u)|du. \tag{2.18}$$

If $q+\nu\geqq r$, then we have, by (2.13),

$$\begin{aligned} |I_1|&\leqq C\int_{|u|\leqq\delta}|u|^{q+\nu-r}|u^rK_\lambda(u)|du \\ &\leqq\delta^{q+\nu-r}CC_1/\lambda^r, \end{aligned} \tag{2.19}$$

that is,

$$I_1=O\left(\frac{1}{\lambda^r}\right). \tag{2.20}$$

If, on the other hand, $q+\nu<r$, then we have

$$I_1=o\left(\frac{1}{\lambda^{q+\nu}}\right), \tag{2.21}$$

since

$$\int_{-\infty}^{\infty}\lambda^a|s|^a|K_\lambda(s)|ds\leqq\int_{|s|\leqq\frac{1}{\lambda}}|K_\lambda(s)|ds+\int_{|s|>\frac{1}{\lambda}}\lambda^r|s|^r|K_\lambda(s)|ds \tag{2.22}$$

for any positive number $a<r$. We also have

$$|I_2|\leqq\sum_{k=0}^{\nu}\frac{f^{(k)}(x)}{k!}\int_{|s|>\delta}|u^kK_\lambda(u)|du \tag{2.23}$$

$$\leqq\sum_{k=0}^{\nu}\frac{f^{(k)}(x)}{k!}\frac{1}{\delta^{r-k}}\int_{|u|>\delta}|u^rK_\lambda(u)|du.$$

Hence

$$I_2=O\left(\frac{1}{\lambda^r}\right). \tag{2.24}$$

Finally we have

$$|I_3|\leqq\int_{|u|>\delta}|f(x+u)||K_\lambda(u)|du \tag{2.25}$$

$$\leqq\frac{M}{\delta}\int_{|u|>\delta}|u^rK_\lambda(u)|du,$$

that is,

$$I_3=O\left(\frac{1}{\lambda^r}\right). \tag{2.26}$$

Hence altogether we get the conclusion (2.16).

We have the following special cases of Theorem 2.

COROLLARY 3. *If the conditions* (ii)~(iii) *of Theorem* 2 *are satisfied and moreover*

(i)′ *$K(u)$ is a bounded measurable function such that*

$$(1+|u|^r)K(u)\in L_1(-\infty,\infty),\qquad r\geqq 0, \tag{2.27}$$

then we have

$$R_\lambda=\lambda\int_{-\infty}^{\infty}f(x+u)K(\lambda u)du-\lambda\int_{-\infty}^{\infty}g(x+u,x)K(\lambda u)du \tag{2.28}$$

$$=O\left(\frac{1}{\lambda^r}\right),\quad\text{for}\quad r\leqq q+\nu,$$

$$=O\left(\frac{1}{\lambda^{q+\nu}}\right),\quad\text{for}\quad r\geqq q+\nu,$$

for large λ.

COROLLARY 4. *If the conditions* (ii)~(iii) *of Theorem* 2 *are satisfied and moreover*

(i)″

(2.29) $$|K_\lambda(u)|\leqq C_1\lambda, \qquad |u|\leqq\frac{1}{\lambda},$$

(2.30) $$\leqq\frac{C_2}{\lambda^\alpha|u|^{1+\alpha}}, \qquad |u|\geqq\frac{1}{\lambda},$$

for large $\lambda, \alpha>r$, then the same conclusion as in Theorem 2 holds.

This can be easily seen if it is noted that $K_\lambda(u)$ of Corollary 4 satisfies the conditions (2.12) and (2.13).

Since $K(u)$ is bounded, $K^2(u)$ also satisfies the condition (i)′ of Corollary 3. Hence, if we apply this corollary to $K^2(u)$, then we have

(2.31) $$\lambda\int_{-\infty}^{\infty}\{f(x+u)-g(x+u,x)\}K^2(\lambda u)du$$
$$=O\left(\frac{1}{\lambda^r}\right), \quad \text{for} \quad r\leqq q+\nu,$$
$$=O\left(\frac{1}{\lambda^{q+\nu}}\right), \quad \text{for} \quad r\geqq q+\nu.$$

As a special case of Corollary 3, we have the following corollary which was given by Parzen (1962).

Corollary 5. *Suppose that*

(i) *$K(u)$ is a bounded measurable function such that*

(2.32) $$\int_{-\infty}^{\infty}K(u)du=1,$$

(2.33) $$\int_{-\infty}^{\infty}u^kK(u)du=0, \qquad k=1,2,\cdots,p$$

and

(2.34) $$(1+|u|^p)K(u)\in L_1(-\infty,\infty)$$

and that

(ii) *the probability density function $f(u)$ has the p-th derivative at $u=x$. Then*

(2.35) $$p_n(x,\omega)=\lambda_n\int_{-\infty}^{\infty}K(\lambda_n(x-u))dF_n(u,\omega)$$

is an asymptotically unbiased estimate of the order λ_n^p for the value of $f(u)$ at $u=x$. Here, if we take $\lambda_n\sim n^{1/(2p+1)}$, then we get the estimate with the order of consistency $n^{2p/(1+2p)}$.

Furthermore, if we apply Theorem 2 to our problem, then we have

Corollary 6. *Suppose that*

(i) *$f(u)$ is the probability density function which satisfies the conditions* (ii) *and* (iii) *of Theorem 2,* *and that*

(ii) $K_\lambda(u)$ *satisfies* (2.29) *and* (2.30),

$$\int_{-\infty}^{\infty} K_\lambda(u)du=1, \tag{2.36}$$

and

$$\int_{-\infty}^{\infty} u^k K_\lambda(u)du=0, \qquad k=1,2,\cdots,\nu. \tag{2.37}$$

Then, the estimate for $f(u)$ *at* $u=x$

$$p_n(x,\omega)=\int_{-\infty}^{\infty} K_{\lambda_n}(x-u)dF_n(u,\omega) \tag{2.38}$$

is a consistent estimate of the order $n^{2\theta/(1+2\theta)}$, *where* $\theta=\min(q+\nu, r)$ *if* $\theta>0$.

In order to prove this corollary, it should be noted that if $K_\lambda(u)$ satisfies (2.29) and (2.30), then

$$\int_{-\infty}^{\infty} K_\lambda^2(u)du=O(\lambda), \tag{2.39}$$

$$\operatorname{var} p_n(x,\omega) \leqq \frac{C_1}{n}\left\{M\int_{-\infty}^{\infty} K_{\lambda_n}^2(x)dx+Q_{\lambda_n}\right\}, \tag{2.40}$$

and

$$Q_\lambda=O(\lambda), \tag{2.41}$$

where Q_λ is the one defined in (2.5).

Acknowledgements

The author would like to express his thanks to Professor T. KAWATA of Keio University for his constant encouragement and many helpful suggestions. He is also grateful to Dr. Y. WASHIO of Keio University for the useful comments and suggestions.

REFERENCES

KAWATA, T., (1972): *Fourier Analysis in Probability Theory*, Acad. Press, New York.

PARZEN, E., (1956): On the Consistent Estimates of the Spectral Density of a Stationary Time Series, *Proc. Nat. Acad. Sci* (U.S.A.), **42**, pp. 154–157.

——— (1962): On Estimation of a Probability Density Function and Mode, *Ann. Math. Stat.* vol. 33, pp. 1065–1076.

ROSENBLATT, M., (1956): Remarks on Some Nonparametric Estimates of a Density Function, *Ann. Math. Stat.* vol. 27, pp. 832–837.

Van der VAART, H. R., (1967): Determining Absolutely Continuous Component of a Probability Distribution from its Fourier Stieltjes Transform, *Ark. Mat.* vol. 7 pp. 331–342.

WAHBA, C., (1971): A Polynomial Algorithm for Density Estimation, *Ann. Math. Stat.* vol. **42**, pp. 1870–1886.

WHITTLE, P., (1958): On the Smoothing of Probability Density Functions, J. *Roy. Stat. Soc.*, Ser. B. **20**, p. 334–343.