Ann. Inst. Statist. Math.
32 (1980), Part A, 187-199

ON THE UNIFORM COMPLETE CONVERGENCE OF ESTIMATES FOR MULTIVARIATE DENSITY FUNCTIONS AND REGRESSION CURVES

K. F. CHENG[1] AND R. L. TAYLOR[2]

(Received Feb. 5, 1979; revised Jan. 28, 1980)

Abstract

Let $(X_1, Y_1), \cdots, (X_n, Y_n)$ be a random sample from the $(k+1)$-dimensional multivariate density function $f^*(x, y)$. Estimates of the k-dimensional density function $f(x)=\int f^*(x, y)dy$ of the form

$$\hat{f}_n(x)=\frac{1}{nb_1(n)\cdots b_k(n)}\sum_{i=1}^{n} W\left(\frac{x_1-X_{i1}}{b_1(n)}, \cdots, \frac{x_k-X_{ik}}{b_k(n)}\right)$$

are considered where $W(x)$ is a bounded, nonnegative weight function and $b_1(n), \cdots, b_k(n)$ and bandwidth sequences depending on the sample size and tending to 0 as $n \to \infty$. For the regression function

$$m(x)=\mathrm{E}\,(Y|X=x)=\frac{h(x)}{f(x)}$$

where $h(x)=\int yf^*(x, y)dy$, estimates of the form

$$\hat{h}_n(x)=\frac{1}{nb_1(n)\cdots b_k(n)}\sum_{i=1}^{n} Y_i W\left(\frac{x_1-X_{i1}}{b_1(n)}, \cdots, \frac{x_k-X_{ik}}{b_k(n)}\right)$$

are considered. In particular, uniform consistency of these estimates is obtained by showing that $||\hat{f}_n(x)-f(x)||_\infty$ and $||\hat{m}_n(x)-m(x)||_\infty$ converge completely to zero for a large class of " good " weight functions and under mild conditions on the bandwidth sequences $b_k(n)$'s.

[1] Research supported by National Institute of Environmental Health Scicences Under Grant 5T32 ES07011.

[2] For this author, revisions made while supported by the Air Force Office of Scientific Research under Contract Number F49620-79-C-0140.

Key words and phrases: Weight function, bandwidth sequences, regression function estimates, and complete convergence.

AMS Classification Number: Primary 62E15 and 62E40, Secondary 60F15.

1. Introduction

Let (X, Y) be a $(k+1)$-dimensional random vector with the joint probability density function $f^*(x, y)$ and let X be a k-dimensional random vector with the continuous marginal density function $f(x)$. Some modified multivariate density function estimates for f have been discussed by Cacoullos [2] which were based on a random sample $(X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n)$ from f^*. In particular, estimates $f_n(x)$ for the multivariate density function $f(x)$ of the form

$$f_n(x)=\frac{1}{nb^k(n)}\sum_{i=1}^{n} W\left(\frac{x-X_i}{b(n)}\right) \tag{1.1}$$

were considered where $W(x)$ is a bounded, nonnegative, integrable weight function such that

$$\int_{R^k} W(x)dx=1$$

and $b(n)$ is a bandwidth sequence depending on the sample size and tending to 0 as $n \to \infty$.

For the regression function $m(x)$ of Y on X,

$$m(x)=\mathrm{E}\,(Y|X=x)=\frac{h(x)}{f(x)}$$

where $h(x)=\int yf^*(x, y)dy$, Watson [9] and Nadaraya [5] independently proposed the following regression function estimates for the case $k=1$:

$$m_n(x)=\frac{h_n(x)}{f_n(x)}=\sum_{i=1}^{n} Y_i \frac{W\left(\frac{x-X_i}{b(n)}\right)}{\sum_{j=1}^{n} W\left(\frac{x-X_j}{b(n)}\right)}\,. \tag{1.2}$$

Thus, the estimates for $h(x)$ are

$$h_n(x)=\frac{1}{nb^k(n)}\sum_{i=1}^{n} Y_i W\left(\frac{x-X_i}{b(n)}\right). \tag{1.3}$$

An heuristic treatment of $m_n(x)$ as a weighted average of the Y_i's can be found in Watson [9]. The local properties of (1.1) and (1.2) have been studied extensively (see Rosenblatt [7]), and global measurements of deviations of $m_n(x)$ from $m(x)$ and $f_n(x)$ from $f(x)$ are given by

$$||m_n(x)-m(x)||_\infty=\sup_{x\in R^k}|m_n(x)-m(x)| \tag{1.4}$$

and

$$\|f_n(\boldsymbol{x})-f(\boldsymbol{x})\|_\infty=\sup_{\boldsymbol{x}\in R^k}|f_n(\boldsymbol{x})-f(\boldsymbol{x})|\,. \tag{1.5}$$

In this paper the important large sample properties of the more general estimates $\hat{f}_n(\boldsymbol{x})$ and $\hat{m}_n(\boldsymbol{x})$ are explored where

$$\hat{f}_n(\boldsymbol{x})=\frac{1}{nb_1(n)\cdots b_k(n)}\sum_{i=1}^{n}W\left(\frac{x_1-X_{i1}}{b_1(n)},\ \ldots,\ \frac{x_k-X_{ik}}{b_k(n)}\right), \tag{1.6}$$

$$\hat{m}_n(\boldsymbol{x})=\frac{\hat{h}_n(\boldsymbol{x})}{\hat{f}_n(\boldsymbol{x})}\,, \tag{1.7}$$

and

$$\hat{h}_n(\boldsymbol{x})=\frac{1}{nb_1(n)\cdots b_k(n)}\sum_{i=1}^{n}Y_iW\left(\frac{x_1-X_{i1}}{b_1(n)},\ \ldots,\ \frac{x_k-X_{ik}}{b_k(n)}\right). \tag{1.8}$$

Here, $W(\boldsymbol{x})$ is a nonnegative, bounded weight function and the bandwidth sequence $b_j(n)$'s depends on the sample size n and tends to zero as $n\to\infty$. The rate of convergence to zero need not be uniform, and the possible weight functions include more than the uniform kernels. Attention will generally be restricted to a density function $f^*(\boldsymbol{x}, y)$ with compact support $[a_1, b_1]\times[a_2, b_2]\times[a_3, b_3]$ (for notation convenience only $k=2$) and weight functions $W(\boldsymbol{x})$ which satisfy

$$\int_{a_1'}^{b_1'}\int_{a_2'}^{b_2'}W(\boldsymbol{x})d\boldsymbol{x}=1$$

and vanish outside $[a_1', b_1']\times[a_2', b_2']$. Also, the continuity of $f(\boldsymbol{x})$ and $h(\boldsymbol{x})$ on compact support $[a_1, b_1]\times[a_2, b_2]$ and on $[a_1', b_1']\times[a_2', b_2']$ are assumed. However the compact support of $f(\boldsymbol{x})$ can be eliminated (Lemma 2) when $f(\boldsymbol{x})$ has a pth moment, $p>0$.

The major results of this paper give a new class of " good " weight functions under mild conditions on the bandwidth sequences $b_i(n)$'s. Uniform consistency of the estimates $\hat{f}_n(\boldsymbol{x})$ and $\hat{m}_n(\boldsymbol{x})$ is obtained since $\|\hat{f}_n(\boldsymbol{x})-f(\boldsymbol{x})\|_\infty$ and $\|\hat{m}_n(\boldsymbol{x})-m(\boldsymbol{x})\|_\infty$ converge completely to zero (which implies convergence with probability one). The main tools in obtaining these results will be approximating polyhedral functions and sub-Gaussian techniques and will parallel the development in Taylor and Cheng [8].

The modulus of continuity, $\omega_g(\delta_1, \delta_2)$, is defined by Billingsley [1] as

$$\omega_g(\delta_1, \delta_2)=\sup_{\substack{|t_1-s_1|\le\delta_1\\ |t_2-s_2|\le\delta_2}}|g(s_1, s_2)-g(t_1, t_2)|$$

for $\delta_1, \delta_2>0$; $(s_1, s_2), (t_1, t_2)\in[a_1, b_1]\times[a_2, b_2]$; and $g\in C([a_1, b_1]\times[a_2, b_2])$, the space of continuous functions with domain on $[a_1, b_1]\times[a_2, b_2]$.

DEFINITION (Chow [3]). A random variable X is said to be *sub-Gaussian* if there exists $\alpha \geqq 0$ such that

$$\mathrm{E}\left[\exp(tX)\right] \leqq \exp\left(\frac{\alpha^2 t^2}{2}\right) \quad \text{for all } t \in R. \tag{1.9}$$

If X is sub-Gaussian, then let

$$\tau(X) = \inf\{\alpha \geqq 0: \ \text{Inequality (1.9) holds}\}.$$

Some basic properties of sub-Gaussian random variables which will be used include:

1. If $\mathrm{P}[|X| \leqq K] = 1$ and $\mathrm{E}\,X = 0$, then

$$\mathrm{E}\left[\exp(tX)\right] \leqq \exp(K^2 t^2).$$

2. If $\tau(X) = \alpha$, then

$$\mathrm{P}[|X| \geqq \lambda] \leqq 2\exp(-\lambda^2/2\alpha^2). \tag{1.10}$$

3. The sum of two independent sub-Gaussian random variables is sub-Gaussian.

A sequence of random variables $\{X_n\}$ is said to *converge completely* to a random variable X if

$$\sum_{n=1}^{\infty} \mathrm{P}[|X_n - X| > \varepsilon] < \infty \tag{1.11}$$

for each $\varepsilon > 0$.

2. Main results

In this section it is shown that $\|\hat{f}_n(\boldsymbol{x}) - f(\boldsymbol{x})\|_\infty$ and $\|\hat{m}_n(\boldsymbol{x}) - m(\boldsymbol{x})\|_\infty$ converge completely to zero under conditions on the modulus of continuity of the weight function $W(\boldsymbol{x})$ and the rate of convergence to zero by the bandwidth sequences $b_j(n)$'s.

LEMMA 1. *If* (i) $nb_1^2(n)b_2^2(n) > n^\delta$ *for some* $\delta > 0$, (ii) $\omega_W\left(\frac{(b_1+b_1')-(a_1+a_1')}{n^{r_1}b_1(n)}, \frac{(b_2+b_2')-(a_2+a_2')}{n^{r_2}b_2(n)}\right) = o(b_1(n)b_2(n))$ *for some integers* $r_1, r_2 > 0$, *and* (iii) $f(\boldsymbol{x})$ *has a compact support* $[a_1, b_1] \times [a_2, b_2]$, *then*

$$\sup_{\boldsymbol{s} \in R^2} \left| \hat{f}_n(\boldsymbol{s}) - \frac{1}{b_1(n)b_2(n)} \mathrm{E}\, W\left(\frac{s_1 - X_{11}}{b_1(n)}, \frac{s_2 - X_{12}}{b_2(n)}\right)\right| \to 0 \tag{2.1}$$

completely as $n \to \infty$ *where* $\hat{f}_n(\boldsymbol{s})$ *is defined in* (1.6).

PROOF. First, $[a_1', b_1'] \times [a_2', b_2']$ may be expanded to include $(0, 0)$ if

$(0, 0) \notin [a_1', b_1'] \times [a_2', b_2']$. For the positive integers r_1 and r_2, let $I_{si} = [t_{s,i-1}, t_{si}]$, $s=1$ or 2, where $t_{si} = (a_s + a_s') + [(b_s + b_s') - (a_s + a_s')]i/n^{r_s}$ and let $I_{ij} = I_{1i} \times I_{2j}$. Thus,

$$[a_1 + a_1', b_1 + b_1'] \times [a_2 + a_2', b_2 + b_2'] = \bigcup_{i=1}^{n^{r_1}} \bigcup_{j=1}^{n^{r_2}} I_{ij} .$$

Since W and f vanish outside $[a_1', b_1'] \times [a_2', b_2']$ and $[a_1, b_1] \times [a_2, b_2]$ respectively and $b_1(n) \to 0$ and $b_2(n) \to 0$, the sup in (2.1) need only be taken over $[a_1 + a_1', b_1 + b_1'] \times [a_2 + a_2', b_2 + b_2']$. Let $\delta_n^s = \dfrac{(b_s + b_s') - (a_s + a_s')}{n^{r_s} b_s(n)}$ for $s=1$ or 2 and let

$$\tilde{W}_k(s_1, s_2) = W\left(s_1 - \frac{X_{k1}}{b_1(n)}, s_2 - \frac{X_{k2}}{b_2(n)}\right) - \mathrm{E}\, W\left(s_1 - \frac{X_{k1}}{b_1(n)}, s_2 - \frac{X_{k2}}{b_2(n)}\right)$$

for each $k = 1, \cdots, n$. Thus, $\mathrm{E}\, \tilde{W}_k(s) = 0$ for each $s \in [a_1 + a_1', b_1 + b_1'] \times [a_2 + a_2', b_2 + b_2']$ and each k. Furthermore,

$$\begin{aligned}
(2.2) \quad \omega_{\tilde{W}_k}(\delta_n^1, \delta_n^2) &= \sup_{\substack{|s_1 - t_1| \le \delta_n^1 \\ |s_2 - t_2| \le \delta_n^2}} |\tilde{W}_k(s_1, s_2) - \tilde{W}_k(t_1, t_2)| \\
&\le \sup_{\substack{|s_1 - t_1| \le \delta_n^1 \\ |s_2 - t_2| \le \delta_n^2}} \left| W\left(s_1 - \frac{X_{k1}}{b_1(n)}, s_2 - \frac{X_{k2}}{b_2(n)}\right) - W\left(t_1 - \frac{X_{k1}}{b_1(n)}, t_2 - \frac{X_{k2}}{b_2(n)}\right) \right| \\
&\quad + \sup_{\substack{|s_1 - t_1| \le \delta_n^1 \\ |s_2 - t_2| \le \delta_n^2}} \left| \mathrm{E}\, W\left(s_1 - \frac{X_{k1}}{b_1(n)}, s_2 - \frac{X_{k2}}{b_2(n)}\right) - \mathrm{E}\, W\left(t_1 - \frac{X_{k1}}{b_1(n)}, t_2 - \frac{X_{k2}}{b_2(n)}\right) \right| \\
&\le 2\omega_W(\delta_n^1, \delta_n^2) .
\end{aligned}$$

Hence, $\omega_{\tilde{W}_k}(\delta_n^1, \delta_n^2) \le 2\omega_W(\delta_n^1, \delta_n^2) = o(b_1(n) b_2(n))$ for each k from Condition (ii). For $\varepsilon > 0$ let

$$\begin{aligned}
(2.3) \quad A_n &= \left[\sup_{\substack{a_1 + a_1' \le s_1 \le b_1 + b_1' \\ a_2 + a_2' \le s_2 \le b_2 + b_2'}} \left| \frac{1}{n b_1(n) b_2(n)} \sum_{k=1}^n \tilde{W}_k\left(\frac{s_1}{b_1(n)}, \frac{s_2}{b_2(n)}\right) \right| > \varepsilon \right] \\
&= \left[\max_{\substack{1 \le i \le n^{r_1} \\ 1 \le j \le n^{r_2}}} \sup_{(s_1, s_2) \in I_{ij}} \left| \frac{1}{n b_1(n) b_2(n)} \sum_{k=1}^n \tilde{W}_k\left(\frac{s_1}{b_1(n)}, \frac{s_2}{b_2(n)}\right) \right| > \varepsilon \right] .
\end{aligned}$$

Hence,

$$(2.4) \quad A_n \subset \left[\max_{\substack{1 \le i \le n^{r_1} \\ 1 \le j \le n^{r_2}}} \left| \frac{1}{n b_1(n) b_2(n)} \sum_{k=1}^n \tilde{W}_k\left(\frac{t_{1i}}{b_1(n)}, \frac{t_{2j}}{b_2(n)}\right) \right| \right.$$

$$+\max_{\substack{1\leq i\leq n^{r_1}\\ 1\leq j\leq n^{r_2}}}\sup_{(s_1,s_2)\in I_{ij}}\left|\frac{1}{nb_1(n)b_2(n)}\sum_{k=1}^{n}\left[\tilde{W}_k\left(\frac{s_1}{b_1(n)},\frac{s_2}{b_2(n)}\right)\right.\right.$$

$$\left.\left.-\tilde{W}_k\left(\frac{t_{1i}}{b_1(n)},\frac{t_{2j}}{b_2(n)}\right)\right]\right|>\varepsilon\Bigg].$$

However,

$$\max_{\substack{1\leq i\leq n^{r_1}\\ 1\leq j\leq n^{r_2}}}\sup_{(s_1,s_2)\in I_{ij}}\frac{1}{nb_1(n)b_2(n)}\sum_{k=1}^{n}\left|\tilde{W}_k\left(\frac{s_1}{b_1(n)},\frac{s_2}{b_2(n)}\right)-\tilde{W}_k\left(\frac{t_{1i}}{b_1(n)},\frac{t_{2j}}{b_2(n)}\right)\right|$$

$$\leq\frac{2\omega_W(\delta_n^1,\delta_n^2)}{b_1(n)b_2(n)}\,.$$

Since $2\omega_W(\delta_n^1,\delta_n^2)=o(b_1(n)b_2(n))$ by Condition (ii), there exists $N(r_1,r_2)$ such that

$$A_n\subset\left[\max_{\substack{1\leq i\leq n^{r_1}\\ 1\leq j\leq n^{r_2}}}\left|\frac{1}{nb_1(n)b_2(n)}\sum_{k=1}^{n}\tilde{W}_k\left(\frac{t_{1i}}{b_1(n)},\frac{t_{2j}}{b_2(n)}\right)\right|>\frac{\varepsilon}{2}\right]$$

for all $n\geq N(r_1,r_2)$. Using the basic properties of sub-Gaussian random variables $\left[\left\{\tilde{W}_k\left(\frac{t_{1i}}{b_1(n)},\frac{t_{2j}}{b_2(n)}\right):k=1,2,\cdots,n\right\}\text{ for each } i,j\right]$, for each $n\geq N(r_1,r_2)$

$$\begin{aligned}\mathrm{P}\,[A_n]&\leq\mathrm{P}\left[\max_{\substack{1\leq i\leq n^{r_1}\\ 1\leq j\leq n^{r_2}}}\left|\frac{1}{nb_1(n)b_2(n)}\sum_{k=1}^{n}\tilde{W}_k\left(\frac{t_{1i}}{b_1(n)},\frac{t_{2j}}{b_2(n)}\right)\right|>\frac{\varepsilon}{2}\right]\\&\leq\sum_{i=1}^{n^{r_1}}\sum_{j=1}^{n^{r_2}}\mathrm{P}\left[\left|\frac{1}{nb_1(n)b_2(n)}\sum_{k=1}^{n}\tilde{W}_k\left(\frac{t_{1i}}{b_1(n)},\frac{t_{2j}}{b_2(n)}\right)\right|>\frac{\varepsilon}{2}\right]\\&\leq n^{r_1+r_2}2\exp\left(-\varepsilon^2/4\|W\|_\infty^2B_n\right)\end{aligned}$$

where

$$\|W\|_\infty=\sup_{(s_1,s_2)\in R^2}|W(s_1,s_2)|\quad\text{and}\quad B_n=\sum_{k=1}^{n}\left(\frac{1}{nb_1(n)b_2(n)}\right)^2=\frac{1}{nb_1^2(n)b_2^2(n)}\,.$$

To obtain the complete convergence of (2.1), consider

$$\begin{aligned}(2.5)\qquad\sum_{n=1}^{\infty}\mathrm{P}\,[A_n]&=\sum_{n=1}^{N(r_1,r_2)}\mathrm{P}\,[A_n]+\sum_{n=N(r_1,r_2)+1}^{\infty}\mathrm{P}\,[A_n]\\&\leq N(r_1,r_2)+\sum_{n=N(r_1,r_2)+1}^{\infty}2n^{r_1+r_2}\exp\left(\frac{-\varepsilon^2nb_1^2(n)b_2^2(n)}{4\|W\|_\infty^2}\right)\\&\leq N(r_1,r_2)+\sum_{n=N(r_1,r_2)+1}^{\infty}2n^{r_1+r_2}\exp\left(-cn^\delta\right)\end{aligned}$$

where $c=\varepsilon^2/4\|W\|_\infty^2$. Thus the series in (2.5) converges by the integral test.

The compact support of the $f(x)$ can be relaxed if any moment $p>0$ exists. Lemma 2 summarizes this result and again is stated only for $k=2$. The proof of Lemma 2 is similar to the univariate case in Taylor and Cheng [8] and is omitted. Recall that $\|(a, b)\|=(a^2+b^2)^{1/2}$.

LEMMA 2. *If* (i) $nb_1^2(n)b_2^2(n)>n^\delta$ *for some* $\delta>0$, (ii) $\int\|x\|^p f(x)dx <\infty$ *for some* $p>0$, *and* (iii) $W\left(\frac{2+(b_1'-a_1')}{n^{r_1}b_1(n)}, \frac{2+(b_2'-a_2')}{n^{r_2}b_2(n)}\right)=o(b_1(n)b_2(n))$ *for some integers* $r_1, r_2>1/p$, *then*

$$\sup_{s\in R^2}\left|\hat{f}_n(s_1, s_2)-\frac{1}{b_1(n)b_2(n)} \mathrm{E}\, W\left(\frac{s_1-X_{11}}{b_1(n)}, \frac{s_2-X_{12}}{b_2(n)}\right)\right|\to 0$$

completely as $n\to\infty$ *where* $\hat{f}_n(s)$ *is defined in* (1.6).

LEMMA 3. *If the underlying density,* f, *is uniformly continuous, then*

$$\sup_{(s_1, s_2)\in R^2}\left|\frac{1}{b_1(n)b_2(n)} \mathrm{E}\, W\left(\frac{s_1-X_{11}}{b_1(n)}, \frac{s_2-X_{12}}{b_2(n)}\right)-f(s_1, s_2)\right|\to 0\,. \tag{2.6}$$

PROOF. Since f is uniformly continuous on R^2, given $\varepsilon>0$ there exists $\delta>0$ such that $|f(s_1', s_2')-f(s_1, s_2)|<\varepsilon$ whenever $\|(s_1', s_2')-(s_1, s_2)\|<\delta$. Let N be sufficiently large so that $\|(b_1(n)y_1, b_2(n)y_2)\|<\delta$ for all $n\geqq N$ and all $(y_1, y_2)\in[a_1', b_1']\times[a_2', b_2']$. Since $W(y_1, y_2)=0$ for $(y_1, y_2)\notin[a_1', b_1']\times[a_2', b_2']$,

$$\begin{aligned}
&\left|\frac{1}{b_1(n)b_2(n)} \mathrm{E}\, W\left(\frac{s_1-X_{11}}{b_1(n)}, \frac{s_2-X_{12}}{b_2(n)}\right)-f(s_1, s_2)\right| \\
&=\left|\frac{1}{b_1(n)b_2(n)}\int_{R^2} W\left(\frac{s_1-x_1}{b_1(n)}, \frac{s_2-x_2}{b_2(n)}\right)f(x_1, x_2)dx_1dx_2-f(s_1, s_2)\right| \\
&=\left|\int_{R^2} W(y_1, y_2)f(s_1-b_1(n)y_1, s_2-b_2(n)y_2)dy_1dy_2-f(s_1, s_2)\right| \\
&=\left|\int_{R^2} W(y_1, y_2)[f(s_1-b_1(n)y_1, s_2-b_2(n)y_2)-f(s_1, s_2)]dy_1dy_2\right| \\
&=\varepsilon\int_{R^2} W(y_1, y_2)dy_1dy_2=\varepsilon
\end{aligned} \tag{2.7}$$

uniformly in (s_1, s_2) for all $n\geqq N$. Hence

$$\sup_{s\in R^2}\left|\frac{1}{b_1(n)b_2(n)} \mathrm{E}\, W\left(\frac{s_1-X_{11}}{b_1(n)}, \frac{s_2-X_{12}}{b_2(n)}\right)-f(s_1, s_2)\right|\to 0$$

as $n\to\infty$.

If the density, f, is continuous on R^2 and has compact support, then it is uniformly continuous and all pth moments exist. Conditions

(ii) and (iii) of Lemma 2 are easily satisfied, and the more general hypotheses of Lemmas 2 and 3 are listed for Theorem 1. First, it should be indicated that the case where f is only known to be continuous on $[a_1, b_1]\times[a_2, b_2]$ is not entirely excluded from consideration.

COROLLARY 1. *If the underlying density function, f, is only known to be continuous on $[a_1, b_1]\times[a_2, b_2]$, then for arbitrarily small $\varepsilon_1, \varepsilon_2>0$*

$$\sup_{\substack{a_1+\varepsilon_1\leq s_1\leq b_1-\varepsilon_1\\ a_2+\varepsilon_2\leq s_2\leq b_2-\varepsilon_2}}\left|\frac{1}{b_1(n)b_2(n)}\,\mathrm{E}\,W\left(\frac{s_1-X_{11}}{b_1(n)}, \frac{s_2-X_{12}}{b_2(n)}\right)-f(s_1, s_2)\right|\to 0$$

as $n\to\infty$.

The ε's in Corollary 1 and Corollary 2 can be considered as functions of n which tend to zero as $n\to\infty$. The proof of the following theorem is immediate from Lemmas 2 and 3 since for each $\varepsilon>0$

$$\begin{aligned}(2.8)\qquad &\mathrm{P}\left[\sup_{s\in R^2}\left|\frac{1}{b_1(n)b_2(n)}\sum_{k=1}^{n}W\left(\frac{s_1-X_{k1}}{b_1(n)}, \frac{s_2-X_{k2}}{b_2(n)}\right)-f(s_1, s_2)\right|>\varepsilon\right]\\ &\leqq\mathrm{P}\left[\sup_{s\in R^2}\left|\frac{1}{b_1(n)b_2(n)}\sum_{k=1}^{n}\left[W\left(\frac{s_1-X_{k1}}{b_1(n)}, \frac{s_2-X_{k2}}{b_2(n)}\right)\right.\right.\right.\\ &\qquad\qquad\left.\left.\left.-\mathrm{E}\,W\left(\frac{s_1-X_{11}}{b_1(n)}, \frac{s_2-X_{12}}{b_2(n)}\right)\right]\right|>\frac{\varepsilon}{2}\right]\\ &\leqq\mathrm{P}\left[\sup_{s\in R^2}\left|\frac{1}{b_1(n)b_2(n)}\,\mathrm{E}\,W\left(\frac{s_1-X_{11}}{b_1(n)}, \frac{s_2-X_{12}}{b_2(n)}\right)-f(s_1, s_2)\right|>\frac{\varepsilon}{2}\right]\end{aligned}$$

and each of the terms in (2.8) is a convergent series in n. All of the conditions will be stated in Theorem 1 for easy reference, and in particular will be stated for arbitrary k-dimension.

THEOREM 1. *Let $\{X_n\}$ be independent random vectors with the same density function $f(\boldsymbol{x})$ which is uniformly continuous on R^k. Let $W(\boldsymbol{x})$ be a nonnegative weight function which is continuous on its compact support $[a_1', b_1']\times\cdots\times[a_k', b_k']$ and integrates to 1. If* (a) *$nb_1^2(n)\cdots b_k^2(n)>n^\delta$ for some $\delta>0$,* (b) *$\int\|\boldsymbol{x}\|^p f(\boldsymbol{x})d\boldsymbol{x}<\infty$ for some $p>0$, and* (c) *$\omega_W\left(\frac{2+(b_1'-a_1')}{n^{r_1}b_1(n)}, \cdots, \frac{2+(b_k'-a_k')}{n^{r_k}b_k(n)}\right)=o(b_1(n)\cdots b_k(n))$ for some integers $r_1, \cdots, r_k>1/p$, then*

$$\sup_{s\in R^k}\left|\frac{1}{nb_1(n)\cdots b_k(n)}\sum_{i=1}^{n}W\left(\frac{s_1-X_{i1}}{b_1(n)}, \cdots, \frac{s_k-X_{ik}}{b_k(n)}\right)-f(s_1, \cdots, s_k)\right|\to 0$$

completely as $n\to\infty$.

COROLLARY 2. *If all of the conditions of Theorem 1 are satisfied* *except it is only known that f is continuous on its compact support, then for arbitrarily small $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_k > 0$,*

$$\sup_{a+\varepsilon \leqq s \leqq b-\varepsilon} \left| \frac{1}{nb_1(n)\cdots b_k(n)} \sum_{i=1}^{n} W\left(\frac{s_1 - X_{i1}}{b_1(n)}, \cdots, \frac{s_k - X_{ik}}{b_k(n)}\right) - f(s_1, \cdots, s_k) \right| \to 0$$

completely as $n \to \infty$ where $\boldsymbol{a}+\boldsymbol{\varepsilon} \leqq \boldsymbol{s} \leqq \boldsymbol{b}-\boldsymbol{\varepsilon}$ indicates that the relationship holds in each coordinate.

The condition $nb_1^2(n)\cdots b_k^2(n) > n^\delta$ need not hold for all n but only eventually. Also, the condition could have been stated as

$$\int_d^\infty x^{r_1+\cdots+r_k} \exp(-cxb_1^2(x)\cdots b_k^2(x))dx < \infty \qquad \text{for some } d>0$$

where $b_i(x)$'s are functions which generate the bandwidth sequences $b_i(1), b_i(2), \cdots, i=1, \cdots, k$, and $c>0$ is a constant.

Attention is now directed to establishing the uniform complete convergence of $||\hat{m}_n(\boldsymbol{x}) - m(\boldsymbol{x})||_\infty$ to zero. Again, for notational convenience $k=2$.

LEMMA 4. *If the regularity conditions in Lemma 1 are satisfied, then*

$$\sup_{s \in R^2} \left| \frac{1}{nb_1(n)b_2(n)} \sum_{i=1}^{n} Y_k W\left(\frac{s_1 - X_{i1}}{b_1(n)}, \frac{s_2 - X_{i2}}{b_2(n)}\right) - \frac{1}{b_1(n)b_2(n)} \mathrm{E}\, Y_1 W\left(\frac{s_1 - X_{11}}{b_1(n)}, \frac{s_2 - X_{12}}{b_2(n)}\right) \right| \to 0$$

completely as $n \to \infty$.

PROOF. The proof follows similarly to the proof of Lemma 1. First, define

$$\tilde{W}_i^*(s_1, s_2) = Y_i W\left(s_1 - \frac{X_{i1}}{b_1(n)}, s_2 - \frac{X_{i2}}{b_2(n)}\right) - \mathrm{E}\left(Y_1 W\left(s_1 - \frac{X_{11}}{b_1(n)}, s_2 - \frac{X_{12}}{b_2(n)}\right)\right)$$

where

$$\mathrm{E}\left(Y_1 W\left(s_1 - \frac{X_{11}}{b_1(n)}, s_2 - \frac{X_{12}}{b_2(n)}\right)\right) = \int_{R^3} y W\left(s_1 - \frac{x_1}{b_1(n)}, s_2 - \frac{x_2}{b_2(n)}\right) f^*(x_1, x_2, y) dx_1 dx_2 dy\,.$$

Thus, $\mathrm{E}\, W_i^*(s_1, s_2) = 0$ for all s_1, s_2 and i and $\omega_{W_i^*}(\delta_n^1, \delta_n^2) \leqq 2M\omega_W(\delta_n^1, \delta_n^2)$ where $||Y_i|| \leqq M$ a.s. Hence, $\omega_{\tilde{W}_i^*}(\delta_n^1, \delta_n^2) = o(b_1(n)b_2(n))$ a.s., and the remainder of the proof follows from the proof of Lemma 1 with the new constant being $c' = \dfrac{\varepsilon^2}{4M^2||W||_\infty^2}$.

A moment condition again could be used to relax the assumption of compact support for Y with the modifications being similar to Lemma 3 but will not be stated. The next result will establish the convergence for the mean.

LEMMA 5. *If $h(\boldsymbol{x})$ is uniformly continuous on R^2, then*

$$(2.9)\qquad \sup_{s\in R^2}\left|\frac{1}{b_1(n)b_2(n)}\,\mathrm{E}\,Y_1W\left(\frac{s_1-X_{11}}{b_1(n)},\ \frac{s_2-X_{12}}{b_2(n)}\right)-h(s_1,s_2)\right|\to 0$$

as $n\to\infty$.

PROOF. First,

$$(2.10)\qquad \frac{1}{b_1(n)b_2(n)}\,\mathrm{E}\,Y_1W\left(\frac{s_1-X_{11}}{b_1(n)},\ \frac{s_2-X_{12}}{b_2(n)}\right)$$
$$=\frac{1}{b_1(n)b_2(n)}\int_{R^3} yW\left(\frac{s_1-x_1}{b_1(n)},\ \frac{s_2-x_2}{b_2(n)}\right)f^*(x_1,x_2,y)dx_1dx_2dy$$
$$=\frac{1}{b_1(n)b_2(n)}\int_{R^2}\mathrm{E}\left(Y_1W\left(\frac{s_1-x_1}{b_1(x)},\ \frac{s_2-x_2}{b_2(x)}\right)\bigg|X_{11}=x_1,\ X_{12}=x_2\right)$$
$$\cdot f(x_1,x_2)dx_1dx_2$$
$$=\frac{1}{b_1(n)b_2(n)}\int_{R^2}W\left(\frac{s_1-x_1}{b_1(n)},\ \frac{s_2-x_2}{b_2(n)}\right)h(x_1,x_2)dx_1dx_2$$
$$=\int_{R^2}W(y_1,y_2)h(s_1-b_1(n)y_1,\ s_2-b_2(n)y_2)dy_1dy_2\,.$$

Next, N can be chosen large enough so that $\|(b_1(n)y_1,\ b_2(n)y_2)\|<\delta$ for all $(y_1,y_2)\in[a_1,b_1]\times[a_2,b_2]$ when $n\geqq N$. Hence, from the uniform continuity of h and (2.10),

$$\left|\frac{1}{b_1(n)b_2(n)}\,\mathrm{E}\,Y_1W\left(\frac{s_1-X_{11}}{b_1(n)},\ \frac{s_2-X_{12}}{b_2(n)}\right)-h(s_1,s_2)\right|$$
$$=\left|\int_{R^2}W(y_1,y_2)[h(s_1-b_1(n)y_1,\ s_2-b_2(n)y_2)-h(s_1,s_2)]dy_1dy_2\right|$$
$$\leqq\varepsilon\int_{R^2}W(y_1,y_2)dy_1dy_2=\varepsilon\ \text{uniformly in }(s_1,s_2)$$

for all $n\geqq N$.

The following theorem on the complete uniform consistency of $\hat{h}_n$ can be proved immediately from Lemmas 4 and 5.

THEOREM 2. *If the regularity conditions of Lemmas 4 and 5 hold, then*

$$(2.11)\qquad \sup_{s\in R^2}|\hat{h}_n(s_1,s_2)-h(s_1,s_2)|\to 0$$

completely as $n\to\infty$.

COROLLARY 3. *If the regularity conditions of Lemma 4 are assumed and $h(s_1, s_2)$ is continuous on its compact support $[a_1, b_1]\times[a_2, b_2]$ only, then for arbitrarily small $\varepsilon_1, \varepsilon_2>0$*

$$\sup_{(s_1,s_2)\in[a_1+\varepsilon_1, b_1-\varepsilon_1]\times[a_2+\varepsilon_2, b_2-\varepsilon_2]} |\hat{h}_n(s_1, s_2)-h(s_1, s_2)|\to 0$$

completely as $n\to\infty$.

The stage is now set to obtain the complete uniform consistency of $\hat{m}_n$ to m.

THEOREM 3. *If the regularity conditions of Theorem 1, Lemma 4 and Lemma 5 are satisfied and if there exist $\varepsilon_0^1, \varepsilon_0^2>0$ such that in $[a_1+\varepsilon_0^1, b_1-\varepsilon_0^1]\times[a_2+\varepsilon_0^2, b_2-\varepsilon_0^2]=C$ $\inf_{s\in C} f(s_1, s_2)=\mu>0$ and $\sup_{s\in C}|m(s_1, s_2)|=v<\infty$, then*

$$\sup_{s\in C}|\hat{m}_n(s_1, s_2)-m(s_1, s_2)|\to 0 \tag{2.12}$$

completely as $n\to\infty$.

PROOF. First

$$\begin{aligned}
(2.13)\quad &\sup_{s\in C}|\hat{m}_n(s_1, s_2)-m(s_1, s_2)| \\
&\leqq \sup_{s\in C}\left|\frac{\hat{h}_n(s_1, s_2)}{\hat{f}_n(s_1, s_2)}-\frac{h(s_1, s_2)}{\hat{f}_n(s_1, s_2)}\right| + \sup_{s\in C}\left|\frac{h(s_1, s_2)}{\hat{f}_n(s_1, s_2)}-\frac{h(s_1, s_2)}{f(s_1, s_2)}\right| \\
&\leqq (\inf_{s\in C}\hat{f}_n(s_1, s_2))^{-1}|\hat{h}_n(s_1, s_2)-h(s_1, s_2)| + \sup_{s\in C}\left|\frac{h(s_1, s_2)}{\hat{f}_n(s_1, s_2)}-\frac{h(s_1, s_2)}{f(s_1, s_2)}\right| \\
&\leqq (\inf_{s\in C}\hat{f}_n(s_1, s_2))^{-1}\{|\hat{h}_n(s_1, s_2)-h(s_1, s_2)| \\
&\qquad + \sup_{s\in C}|m(s_1, s_2)|\sup_{s\in C}|\hat{f}_n(s_1, s_2)-f(s_1, s_2)|\}\,.
\end{aligned}$$

Since $\inf_{s\in C} f(s_1, s_2)=\mu>0$ and $\|\hat{f}_n-f\|_\infty\to 0$ completely by Theorem 1, the result easily follows from Lemmas 4 and 5 and (2.13).

3. Comparisons and useful weight functions

In this part, a few brief comments on Nadaraya's [6] conditions and on useful weight functions which satisfy the results of this paper are listed for comparisons.

To obtain a strong law rather than uniform consistency in probability, the conditions on the weight function and bandwidth sequences are expected to be more stringent. For example in the case $k=1$, let $f(s)$ and $m(x)$ be unknown continuous density and regression functions on R. If

(N1) $W(x)$ is a function of bounded variation such that

$$\sup_{x\in R}|W(x)|<\infty\,,\qquad \lim_{x\to\pm\infty}|xW(x)|=0\,,$$

$$\int|W(x)|dx<\infty\,,\qquad \int W(x)dx=1\,;$$

(N2) $-\infty<A\leqq Y\leqq B<\infty$ with probability one and

$$\min_{-\infty<a\leqq x\leqq b<\infty} f(x)=\mu>0\,;\quad\text{and}$$

(N3) $\sum_{n=1}^{\infty}\exp(-\gamma nb^2(n))$ exists for each $\gamma>0$, then

$\|\hat{m}_n(x)-m(x)\|_\infty\to 0$ with probability one (Nadaraya [6]).

For the results of this paper, the weight function $W(x)$ was required to be nonnegative and continuous on its compact support along with a smoothness condition. The condition on the bandwidths sequence for the case $k=1$ reduces to $nb^2(n)>n^\delta$ for some $\delta>0$. For useful weight functions, Epanechnikov [4] considered multivariate density function estimates of the form

$$f_n(s_1,\cdots,s_k)=\frac{1}{n}\sum_{i=1}^{n}\prod_{j=1}^{k}\frac{1}{b_j(n)}W_j\left(\frac{s_j-X_{ij}}{b_j(n)}\right).$$

Setting $b_j(n)=b(n)$ and $W_j(s)=W(s)$ for $j=1,\cdots,k$, the optimal weight function $W_0(s)$ was found to be

$$W_0(s)=\begin{cases}\dfrac{3}{4(5)^{1/2}}-\dfrac{3s^2}{20(5)^{1/2}} & \text{for } |s|\leqq(5)^{1/2}\\ 0 & \text{otherwise}\end{cases}$$

in minimizing the relative global error. For this case, let $a=-(5)^{1/2}$ and $b=(5)^{1/2}$. Then, $W_0(s_1,s_2)=W_0(s_1)W_0(s_2)$, and $|W_0(s_1,s_2)-W_0(t_1,t_2)|\leqq c_1|s_1-t_1|+c_2|s_2-t_2|$ for constants c_1 and c_2. Moreover, $\omega_{W_0}(\delta_n^1,\delta_n^2)\leqq\dfrac{c_1'}{b_1(n)n^{r_1}}+\dfrac{c_2'}{b_2(n)n^{r_2}}$ for constants c_1' and c_2', and the optimal weight function easily satisfies the smoothness condition of this paper.

If the weight function $W(s_1,s_2)$ satisfies a Lipschitz condition of order α, then

$$|W(s_1,s_2)-W(t_1,t_2)|\leqq M\|(s_1,s_2)-(t_1,t_2)\|^\alpha\quad\text{and}$$

$$\omega_W\left(\frac{2(b_1-a_1)}{b_1(n)n^{r_1}},\frac{2(b_2-a_2)}{b_2(n)n^{r_2}}\right)\leqq M\left(\frac{(2b_1-2a_1)^2}{b_1^2(n)n^{2r_1}}+\frac{(2b_2-2a_2)^2}{b_2^2(n)n^{2r_2}}\right)^{\alpha/2}$$

for some $M>0$. Hence, bandwidth sequences $b_1(n)$ and $b_2(n)$ are to be chosen so that

$$\frac{1}{b_1(n)b_2(n)}\left[\frac{1}{n^{r_1}b_1(n)}+\frac{1}{n^{r_2}b_2(n)}\right]^{\alpha}\to 0$$

as $n\to\infty$ for some integers $r_1, r_2>0$. Let $b_1(n)=n^{-p_1}$ and $b_2(n)=n^{-p_2}$ for $p_1>p_2>0$ and let $r_1=r_2=r$, then

$$\frac{1}{n^{-(p_1+p_2)}}\left(\frac{1}{n^r n^{-p_1}}+\frac{1}{n^r n^{-p_2}}\right)^{\alpha}\leqq\frac{2^{\alpha}}{n^{-(p_1+p_2)}n^{\alpha r}n^{-\alpha p_1}}\,. \tag{3.1}$$

If $\alpha>0$ or $\alpha<-(p_1+p_2)/p_1$, then a positive integer r can be chosen so that (3.1) converges to zero as $n\to\infty$.

For the bandwidth sequences $b_j(n)=b(n)$, Epanechnikov [4] found the optimum bandwidth sequence (minimizing the asymptotic relative global error) to be

$$b_0(n)\sim\left(\frac{kL^k}{nM_0}\right)^{1/(k+4)}$$

where k is the dimension, $L=\int_R W^2(x)dx$, and

$$M_0=\int\cdots\int\left[\sum_{i=1}^{k}\frac{\partial^2 f(x_1,\cdots,x_k)}{\partial x_i^2}\right]^2 dx_1\cdots dx_k\,.$$

When L and M_0 are bounded and $M_0\neq 0$, then there is no difficulty in showing that the optimum bandwidth sequence satisfies the conditions of this paper. Finally, it should be noted that a weight function $W(x)$ exists satisfying a Lipschitz condition of order $0<\alpha<1$ on $[a, b]$ but which is not of bounded variation.

THE FLORIDA STATE UNIVERSITY
UNIVERSITY OF SOUTH CAROLINA

REFERENCES

[1] Billingsley, P. (1968). *Convergence of Probability Measures*, Wiley, New York.
[2] Cacoullos, T. (1966). Estimation of multivariate density, *Ann. Inst. Statist. Math.*, **18**, 179-189.
[3] Chow, Y. S. (1966). Some convergence theorems for independent random variable, *Ann. Math. Statist.*, **37**, 1482-1493.
[4] Epanechinkov, V. A. (1969). Nonparametric estimates of multivariate probability density, *Theory Prob. Appl.*, **14**, 153-158.
[5] Nadaraya, E. A. (1964). On estimating regression, *Theory Prob. Appl.*, **9**, 141-142.
[6] Nadaraya, E. A. (1970). Remarks on nonparametric estimates for density function and regression curve, *Theory Prob. Appl.*, **15**, 134-137.
[7] Rosenblatt, M. (1971). Curve estimates, *Ann. Math. Statist.*, **42**, 1815-1842.
[8] Taylor, R. L. and Cheng, K. F. (1978). On the uniform complete convergence of density function estimates, *Ann. Inst. Statist. Math.*, **30**, A, 397-406.
[9] Watson, G. S. (1964). Smooth regression analysis, *Sankhyā*, A, **26**, 359-372.