J. Austral. Math. Soc. 20 (Series A) (1975), 230–241.

ESTIMATION OF PARTIAL DERIVATIVES OF THE AVERAGE OF DENSITIES BELONGING TO A FAMILY OF DENSITIES

V. SUSARLA and S. KUMAR

(Received 22 October 1973; revised 4 February 1974)

Communicated by C. C. Heyde

1. Introduction and Summary

Recently, attention has been drawn to the problem of estimation of a k-variate probability density and its partial derivatives of various orders. Specifically, let $X_1, \cdots, X_n$ be i.i.d. k-variate random variables with common density f wrt Lebesgue measure μ on the k-dimensional σ-field $\mathscr{B}^k$. Parzen (1962) in the $k = 1$ case and Cacoullos (1966) in the $k \geqq 1$ case gave the asymptotic properties of a class of kernel estimates $f_n(x)$, $x \in R^k$, of $f(x)$ based on $X_1, \cdots, X_n$. The asymptotic properties given in the above two papers concern consistency, asymptotic unbiasedness, bounds for the mean squared error and asymptotic normality of f_n. Also in the context of an empirical Bayes two-action problem, Johns and Van Ryzin (1972) introduced kernel estimates for $f(x)$ and the derivative $f'(x)$ for $x \in R^1$ when f is a mixture of univariate exponential densities wrt Lebesgue measure on $\mathscr{B}^1$. They also investigated the asymptotic unbiasedness and the mean squared error convergence properties of these estimates. Lin (1968) stated some generalizations of the results of Johns and Van Ryzin, with applications to empirical Bayes decision problems.

In a sequence-compound empirical non-Bayes context, one of the authors (Susarla (1974)) exhibited a class of L_2-convergent estimators for the average of the densities (and its first partials) of $X_1, \cdots, X_n$ when these random variables are independent and for each i, X_i has the k-variate normal distribution with covariance matrix I and unknown mean θ_i in the k-sphere of radius α. In Susarla (1974a), similar results have been given when the random variables $X_1, \cdots, X_n$ have gamma densities.

Now we let $X_1, \cdots, X_n$ be independent random variables with densities (wrt Lebesgue measure μ on $\mathscr{B}^k$) $p_1, \cdots, p_n$ belonging to a family of densities $\mathscr{F}$. The purpose of this paper is to exhibit estimators for $n^{-1}\Sigma_{j=1}^n p_j$ and its various partial derivatives which are asymptotically unbiased (Corollary 3.1), quadratic mean consistent (Corollary 3.2), and asymptotically normal (Corollary 3.3) under varying conditions on $\mathscr{F}$ and the kernel function involved in the definition of estimators. Also, we have given (Lemma 3.3) a generalization of the results stated in the previous paragraph. Wherever possible, we compare our results with the corresponding results of Cacoullos (1966) and Parzen (1962).

Section 2 is devoted to describing some of the notation required here. In Section 3, we introduce (and derive the asymptotic properties of) a class of estimators for the vth partial derivative of $n^{-1}\Sigma_{j=1}^{n}p_j$ wrt the argument i, $i = 1,\cdots,k$. We use the well known method of divided differences to define the class of estimators for partial derivatives of the average $n^{-1}\Sigma_{j=1}^{n}p_j$. In Section 4, we exhibit the kernel function which satisfies conditions imposed on it for proving the main results of Section 3. We conclude the paper with a few remarks.

2. Notation

Let μ denote the Lebesgue measure on the k-dimensional Borel σ-field $\mathscr{B}^k$. Let $X_1,\cdots,X_n,\cdots$ be independent random variables with densities $p_1,\cdots,p_n,\cdots$ respectively wrt μ. We abbreviate $n^{-1}\Sigma_{j=1}^{n}p_j$ by $\bar{p}$. For any real valued function f on R^k, $f_m^{(v)}$ denotes the vth order partial derivative of f wrt the mth coordinate. Also for non-negative integers $v_1,\cdots,v_k$, $f^{(v_1),\cdots,(v_k)}$ denotes the mixed partial derivative of f wrt the mth component v_m times for $m = 1,\cdots,k$. Throughout, $1 \leqq v_1 + \cdots + v_k \leqq s$, a fixed positive integer. Throughout the rest of the paper, x will be an arbitrary (but fixed) point in R^k. u and v with coordinates $u_1,\cdots,u_k$ and $v_1,\cdots,v_k$ are generally used as generic points in R^k and as variables of integration in R^k. For u in R^k, $\|u\|$ denotes its distance from the origin. For any real valued function f on R^k and u, v in R^k, $f]_v^u$ denotes $f(u)-f(v)$.

We suppress the arguments of functions involved whenever and wherever it is convenient. E stands for expectation wrt the joint distribution of the random variables involved inside square brackets. Integration will be over all possible values of the variable unless otherwise stated.

3. Results

We first introduce the kernel function K_v which is used in defining the estimators. For a fixed non-negative integer v, let K_v be a $\mathscr{B}^k$-measurable function from R^k to R^1. In the results to be stated later on, this function K_v satisfies one or more of the following conditions:

(A$_1$) $$K_v \text{ is bounded,}$$

(A$_2$) $$\int \left| K_v(u)\pi_{m=1}^{k}u_m^{v_m}\right| d\mu(u) < \infty \text{ for } \Sigma_{m=1}^{k}v_m = 0,1,\cdots,v,$$

and, for a fixed positive integer s ($> v$),

(A$_3$) $$\int K_v(u)\pi_{m=1}^{k}u_m^{v_m}d\mu(u) = 0 \text{ for } \Sigma_{m=1}^{k}v_m = v+1,\cdots,s-1,$$

$$(A_4) \qquad \int \left|K_\nu(u)\pi_{m=1}^{k} u_m^{\nu_m}\right| d\mu(u) < \infty \text{ for } \Sigma_{m=1}^{k} \nu_m = s.$$

For fixed i in $\{1, \cdots, k\}$, let, for u in R^k,

$$(3.1) \qquad K_\nu^*(u) = \Sigma_{r=0}^{\nu}(-1)^r \binom{\nu}{r} (\nu - r + 1)^{-1} K_\nu(J_i(\nu, r)u)$$

where $J_i(\nu, r)$ is the modification of the $k \times k$ identity matrix obtained by replacing the ith diagonal element unity by $(\nu - r + 1)^{-1}$. Motivated by the estimators (2.1) of Cacoullos (1966) and (1.7) of Parzen (1962) and a technique of of divided differences, we define the estimator $\bar{p}_i^{(\nu)}$of $\bar{p}_i^{(\nu)}(x)$ (for each fixed x in R^k) as follows:

$$(3.2) \qquad \bar{p}_i^{(\nu)} = n^{-1} \Sigma_{j=1}^{n} \hat{p}_{ji}^{(\nu)} \text{ where } \hat{p}_{ji}^{(\nu)} = \varepsilon^{-(\nu+k)} K_\nu^*((X_j - x)/\varepsilon)$$

for $j = 1, \cdots, n$ and $0 < \varepsilon = \varepsilon(n) \to 0$ as $n \to \infty$. We assume that the family of densities $\mathscr{F}$ satisfies the following conditions in most of the results to follow.

$$(B_1) \qquad \sup_{p \in \mathscr{F}} \sup_{u} \left| p^{(\nu_1), \ldots, (\nu_k)}(u) \right| < \infty \text{ for } \Sigma_{m=1}^{k} \nu_m = \nu.$$

$$(B_2) \qquad \sup_{p \in \mathscr{F}} p(x) < \infty.$$

As a step towards reducing symbols, the sums $\Sigma_{m=1}^{k}$, $\Sigma_{r=0}^{\nu}$ and the product $\pi_{m=1}^{k}$ will be abbreviated by Σ, Σ' and π respectively. Our purpose now is to show that $\bar{p}_i^{(\nu)}$ is asymptotically unbiased, quadratic mean consistent and asymptotically normal under different sets of conditions on the function K_ν and the family of densities $\mathscr{F}$. Also, we show how one can use conditions (A_3) and (A_4) to obtain a result concerning the rate of convergence for the bias of $\bar{p}_i^{(\nu)}$ as an estimator of $\bar{p}_i^{(\nu)}(x)$.

3a. Asymptotic unbiasedness of $\bar{p}_i^{(\nu)}$

The asymptotic unbiasedness of $\bar{p}_i^{(\nu)}$ can be obtained as a corollary to the following theorem which is similar to Theorem 2.1 of Cacoullos (1966) or Theorem 1A of Parzen (1962).

THEOREM 3.1. *Let f be a real valued function on R^k with νth order partial derivatives continuous at x and*

$$(3.3) \qquad \sup_u \left| f^{(\nu_1), \ldots, (\nu_k)}(u) \right| < \infty \textit{ for } \Sigma \nu_m = \nu.$$

If K_ν satisfies (A_1) and (A_2), then, as $n \to \infty$,

$$(3.4) \qquad \varepsilon^{-(\nu+k)} \int K_\nu^*(u/\varepsilon) f(x+u) d\mu(u) \to f_i^{(\nu)}(x) \int K_\nu(u) u_i^\nu d\mu(u).$$

REMARK. The conditions reduce to those of Theorem 1A of Parzen (1962), except (3.3), when specialized to $\nu = 0$ and $k = 1$. In the $\nu = 0$ and $k = 1$ case, the condition corresponding to (3.3) is that f be integrable. For $\nu = 0$, and any general k, see Theorem 2.1 of Cacoullos (1966).

Proof. We first note that the following properties are satisfied by the function K_ν^* defined by (3.1):

(3.5) $$K_\nu^* \text{ is bounded (follows from } (A_1)),$$

(3.6) $$\int |K_\nu^*(u)\pi u_m^{\nu_m}|\, d\mu(u) < \infty \text{ for } \Sigma\nu_m = 0, 1, \cdots, \nu \text{ (follows from } (A_2)),$$

and

(3.7) $$\int K_\nu^*(u)\pi u_m{}^{\nu_m} d\mu(u) = \begin{cases} \nu! \int K_\nu(u) u_i^\nu\, d\mu(u) & \text{for } \nu_i = \nu,\ \nu_m = 0 \text{ for } m \neq i \\ 0 & \text{for } \nu_i \neq \nu,\ 0 \leqq \Sigma\nu_m \leqq \nu. \end{cases}$$

(Equation (3.7) follows from the following two equalities. The definition of K_ν^* and change of variables imply that

$$\int K_\nu^*(u)\pi u_m^{\nu_m} d\mu(u) = \Sigma'(-1)^r \binom{\nu}{r}\ (\nu - r + 1)^{\nu_i} \int K_\nu(u)\pi u_m^{\nu_m} d\mu(u).$$

(3.8) $$\Sigma'(-1)^r \binom{\nu}{r}\ (\nu - r + 1)^\eta = 0 \text{ or } \nu! \text{ according as}$$

$$\eta = 0, \cdots,\ \nu - 1 \text{ or } = \nu.$$

(3.8) can be obtained from (12.17), page 63 of Feller (1957).)

By (3.7) and change of variables, we obtain that

(3.9) $$\begin{aligned} &\varepsilon^{-(\nu+k)} \int K_\nu^*(u/\varepsilon) f(x+u) d\mu(u) - f_i^{(\nu)}(x) \int K_\nu(u) u_i^\nu d\mu(u) \\ &= \varepsilon^{-(\nu+k)} \int K_\nu^*(u/\varepsilon)\{f(x+u) - (\nu!)^{-1}\Sigma_{L_\nu} f^{(\nu_1),\ldots,(\nu_k)}(x)\pi u_m^{\nu_m}\} d\mu(u) \end{aligned}$$

where L_ν denotes the sum over non-negative integers $\nu_1, \cdots, \nu_k$ such that $\nu_1 + \cdots + \nu_k = \nu$. Replacing $f(x+u)$ by its νth order Taylor expansion with Lagrange's form of remainder, and with the help of (3.7), we obtain that the rhs of (3.9) is

$$\leqq \Sigma_{L_\nu} \varepsilon^{-(\nu+k)} \int |K_\nu^*(u/\varepsilon)\pi u_m^{\nu_m}| \sup_{0<\theta<1} |f^{(\nu_1),\ldots,(\nu_k)}]_x^{x+\theta u}|\, d\mu(u).$$

We complete the proof of the theorem by showing that the summands of the rhs of the last inequality go to zero as $n \to \infty$. Each summand can be bounded by

$$\begin{aligned} &\sup_{\|u\| \leqq \delta} \sup_{0<\theta<1} |f^{(\nu_1),\ldots,(\nu_k)}]_x^{x+\theta u}|\, \varepsilon^{-(\nu+k)} \int |K_\nu^*(u/\varepsilon)\pi u_m{}^{\nu_m}| d\mu(u) \\ &\quad + \sup_{\|u\|>\delta} \sup_{0<\theta<1} |f^{(\nu_1),\ldots,(\nu_k)}]_x^{x+\theta u}| \int_{\|u\|>\delta/\varepsilon} |K_\nu(u)\pi u_m{}^{\nu_m}| d\mu(u) \end{aligned}$$

where $\delta > 0$. Since the second sup is bounded due to (3.3), letting $n \to \infty$ (or $\varepsilon \to 0$) and then $\delta \to 0$, (A_2) and the continuity of νth partials of f imply that the summand goes to zero completing the proof.

As a consequence of the above theorem, we obtain

COROLLARY 3.1. *If the vth order partial derivatives of the densities in* $\mathscr{F}$ *are continuous at x,* $\mathscr{F}$ *satisfies* (B_1), K_ν *satisfies* (A_1) *and* (A_2) *and*

$$\int K_\nu(u)u_i^\nu d\mu(u) = 1,$$

then

$$E[\bar{p}_i^{(\nu)}] - \bar{p}_i^{(\nu)}(x) \to 0 \text{ as } n \to \infty.$$

NOTE. Observe that the condition (B_1) is sufficient to conclude that $\sup_{1 \leqq j \leqq n} |E[\hat{p}_{j_i}^{(\nu)}] - p_{j_i}^{(\nu)}(x)| \to 0$ as $n \to \infty$ which in turn implies the corollary.

3b. Quadratic Mean Consistency of $\bar{\hat{p}}_i^{(\nu)}$

For the purpose of obtaining a result about the quadratic mean consistency of $\bar{\hat{p}}_i^{(\nu)}$, we need the following theorem which is similar to Lemma 2.1 of Cacoullos (1966).

THEOREM 3.2. *Let the conditions of Theorem* 3.1 *be satisfied, then for any* $\delta \geqq 1$, *as* $n \to \infty$,

$$\varepsilon^{\nu\delta+k(\delta-1)} \int (\varepsilon^{-(\nu+k)} |K_\nu^*(u/\varepsilon)|)^\delta f(x+u)d\mu(u) \to f(x) \int |K_\nu^*(u)|^\delta d\mu(u) .$$

PROOF. Since f has νth order partials continuous at x, an application of Taylor's theorem with Lagrange's form of remainder to $f(x+u)$ and a triangle inequality with intermediate function $\Sigma_{L_\nu} f^{(\nu_1),\dots,(\nu_k)}(x)$, give that, for each u in R^k,

$$|f(x+u) - f(x)| \leqq \Sigma_1^\nu \Sigma_{L_r} (\pi \nu_m!)^{-1} |f^{(\nu_1),\dots,(\nu_k)}(x) \pi u_m^{\nu_m}| + \Sigma_{L_\nu} |\pi u_m^{\nu_m}| \sup_{0<\theta<1} |f^{(\nu_1),\dots,(\nu_k)}]_x^{x+\theta u}| \tag{3.10}$$

where L_r denotes the sum over non-negative integers $\nu_1, \cdots, \nu_k$ such that $\Sigma \nu_m = r$. Also, we have, by a change of variable,

$$|\text{difference between the left and right sides of the result}|$$

$$\begin{aligned} &= |\varepsilon^{-k} \int |K_\nu^*(u/\varepsilon)|^\delta (f(x+u) - f(x)) d\mu(u)| \\ &\leqq \Sigma_1^\nu \varepsilon^r \Sigma_{L_r} (\pi \nu_m!)^{-1} |f^{(\nu_1),\dots,(\nu_k)}(x)| \int |K_\nu^*(u)|^\delta |\pi u_m^{\nu_m}| d\mu(u) \\ &\quad + \varepsilon^\nu \varepsilon^{-(\nu+k)} \Sigma_{L_\nu} \int |(K_\nu^*(u/\varepsilon))^\delta \pi u_m^{\nu_m}| \sup_{0<\theta<1} |f^{(\nu_1),\dots,(\nu_k)}]_x^{x+\theta u}| d\mu(u) \end{aligned} \tag{3.11}$$

where the inequality follows from (3.10) and the triangle inequality.

Since (A_2) implies (3.6) which implies that the integrals in the first term of rhs of (3.11) are bounded, the first term of the rhs of (3.11) goes to zero as $n \to \infty$. Since $\delta \geqq 1$ and K_ν^* is bounded by (3.5), the second term on the rhs of (3.11) is exceeded by a constant multiple of

$$\varepsilon^\nu \varepsilon^{-(\nu+k)} \Sigma_{L_\nu} \int \left| K_\nu^*(u/\varepsilon)\pi u_m^{\nu_m} \right| \sup_{0<\theta<1} \left| f^{(\nu_1),\dots,(\nu_k)}\right]_x^{x+\theta u} \left| d\mu(u). \right.$$

This last expression was shown to converge to zero as $n \to \infty$ in the proof of Theorem 3.1. The proof is complete in view of (3.11).

In stating the above theorem, we have the restriction that $\delta \geqq 1$ which is slightly weaker than required for Lemma 2.1 of Cacoullos (1966). Now we state the quadratic mean consistency property of $\bar{\hat{p}}_i^{(\nu)}$.

Corollary 3.2. *Let the conditions of Corollary* 3.1 *be satisfied. If* $n\varepsilon^{2\nu+k} \to \infty$ *and* (B_2) *is satisfied, then* $\bar{\hat{p}}_i^{(\nu)} - \bar{p}_i^{(\nu)}(x) \to 0$ *in quadratic mean.*

Proof. First note that

$$E[|\bar{\hat{p}}_i^{(\nu)} - \bar{p}_i^{(\nu)}(x)|^2] = \operatorname{var}(\bar{\hat{p}}_i^{(\nu)}) + (E[\bar{\hat{p}}_i^{(\nu)}] - \bar{p}_i^{(\nu)}(x))^2. \tag{3.12}$$

By Corollary 3.1, the second term converges to zero as $n \to \infty$. Also, for $j = 1, \cdots, n$,

$$E[|\hat{p}_{j_i}^{(\nu)}|^2] = \int |\varepsilon^{-(\nu+k)} K_\nu^*(u/\varepsilon)|^2 p_j(x+u)d\mu(u). \tag{3.13}$$

But by Theorem 3.2 (with $\delta = 2$) and the condition (B_1), we obtain that $\Sigma_{j=1}^n E[|\hat{p}_{j_i}^{(\nu)}|^2] \sim n\varepsilon^{-(2\nu+k)}\, \bar{p}(x) \int |K_\nu^*(u)|^2 d\mu(u)$ as $n \to \infty$ where $n\bar{p}(x) = \Sigma_{j=1}^n p_j(x)$. Thus $\operatorname{var}(\bar{\hat{p}}_i^{(\nu)}(x)) \to 0$ since $n\varepsilon^{2\nu+k} \to \infty$ and (B_2) is satisfied. The proof of the corollary is complete.

3c. Asymptotic normality of $\bar{\hat{p}}_i^{(\nu)}$

To establish the asymptotic normality result, we need the following asymptotic result about $\operatorname{cov}(\bar{\hat{p}}^{(\nu)}(x_1), \bar{\hat{p}}^{(\nu)}(x_2))$ where x_1 and x_2 are fixed points in R^k at which the νth order partial derivatives of each p ($p \in \mathscr{F}$) are continuous. Since part of the proof of the following lemma is similar to Lemma 2.2 of Cacoullos (1966), we state the lemma and indicate the method of its proof only.

Lemma 3.1. *If the conditions of Corollary* 3.1 *are satisfied at* x_1 *and* x_2, *then* $n\varepsilon^{2\nu+k}\operatorname{cov}(\bar{\hat{p}}_i^{(\nu)}(x_1), \bar{\hat{p}}_i^{(\nu)}(x_2)) \to 0$ *as* $n \to \infty$ *if* $K_\nu(u) \to 0$ *as* $\|u\| \to \infty$.

Indication of Proof. We first observe that the independence of $X_1, \cdots, X_n$ gives the equality

$$n\operatorname{cov}(\bar{\hat{p}}_i^{(\nu)}(x_1), \bar{\hat{p}}_i^{(\nu)}(x_2)) = n^{-1}\, \Sigma_{j=1}^n E[\hat{p}_{j_i}^{(\nu)}(x_1)\hat{p}_{j_i}^{(\nu)}(x_2)] - n^{-1}\, \Sigma_{j=1}^n\, E[\hat{p}_{j_i}^{(\nu)}(x_1)]E[\hat{p}_{j_i}^{(\nu)}(x_2)]. \tag{3.14}$$

https://doi.org/10.1017/S1446788700020541 Published online by Cambridge University Press

Uniformity of the condition (B_1) in p belonging to $\mathscr{F}$ and in u in R^k and an application of Theorem 3.1 to p_j for $j = 1, \cdots, n$ give that, for sufficiently large n, the second term of the rhs of (3.14) is bounded in magnitude by

$$1 + n^{-1} \Sigma_{j=1}^{n} \{ | p_{j_i}^{(\nu)}(x_1) p_{j_i}^{(\nu)}(x_2) | + | p_{j_i}^{(\nu)}(x_1) | + | p_{j_i}^{(\nu)}(x_2) | \}$$

which in turn is bounded due to (B_1). Since $\varepsilon \to 0$, $\varepsilon^{2\nu+k}$ times the second term of the rhs of (3.14) goes to zero.

We also obtain by change of variables that $n^{-1}\varepsilon^{2\nu+k} \Sigma_{j=1}^{n} E[\hat{p}_{j_i}^{(\nu)}(x_1)\hat{p}_{j_i}^{(\nu)}(x_2)]$ can be expressed as $\int K_\nu^*(u)K_\nu^*(u + \varepsilon^{-1}(x_1 - x_2))\bar{p}(x_1 + \varepsilon u)d\mu(u)$. By following an argument similar to Lemma 2.2 of Cacoullos (1966), we can show that the above integral converges to zero since $K_\nu(u) \to 0$ as $\| u \| \to \infty$.

To prove the asymptotic normality result, we need the following result of Hoeffding and Robbins (1948).

LEMMA 3.2. *Given a sequence* $(X_{n\beta}, Y_{n\beta})$ $(n = 1, 2, \cdots, \beta = 1, \cdots, \gamma, \gamma = \gamma(n);$ $\lim_{n\to\infty} \gamma = \infty)$ *of sets of random vectors in* R^2, *independent for each* n *with* $E[X_{n\beta}] = E(Y_{n\beta}] = 0$, *let*

$$E[X_{n\beta}^i Y_{n\beta}^j] = \mu_{ij}^{(n\beta)} \qquad i + j = 2,$$

$$\rho_{n\beta}^3 = \max\{E[|X_{n\beta}|^3], E[|Y_{n\beta}|^3]\}, \rho_n^3 = \Sigma_{\beta=1}^{\gamma} \rho_{n\beta}^3.$$

If

$$\lim_{n\to\infty} \gamma^{-1} \Sigma_{\beta=1}^{\gamma} \mu_{ij}^{(n\beta)} = \mu_{ij} \qquad i + j = 2$$

and

$$\lim_{n\to\infty} \gamma^{-1/2} \rho_n = 0,$$

then $\gamma^{-1/2}(X_{n1} + \cdots + X_{n\gamma}, Y_{n1} + \cdots + Y_{n\gamma})$ *has a limiting normal distribution with mean* (0, 0) *and covariances* μ_{ij}.

COROLLARY 3.3. *If the conditions of Lemma* 3.1 *are satisfied at* x_1 *and* x_2, (B_2) *is satisfied*, $n\varepsilon^{2\nu+k} \to \infty$, *and* $\bar{p}(x_N) \to c(x_N)$ *for* $N = 1, 2$, *then*

$$(n\varepsilon^{2\nu+k})^{1/2} \begin{pmatrix} \bar{\hat{p}}_i^{(\nu)}(x_1) - E[\bar{\hat{p}}_i^{(\nu)}(x_1)] \\ \bar{\hat{p}}_i^{(\nu)}(x_2) - E[\bar{\hat{p}}_i^{(\nu)}(x_2)] \end{pmatrix}$$

converges in law to the bivariate normal distribution with mean (0,0) *and covariance matrix*

$$\begin{bmatrix} c(x_1) \int |K_\nu^*(u)|^2 d\mu(u) & 0 \\ 0 & c(x_2) \int |K_\nu^*(u)|^2 d\mu(u) \end{bmatrix}$$

NOTE: Extension of this result to $N > 2$ is trivial and, therefore, is omitted. The conditions $\bar{p}(x_N) \to c(x_N)$ as $n \to \infty$ for $N = 1, 2$ are trivially satisfied if $X_1, \cdots, X_n, \cdots$ are identically distributed random variables.

PROOF. In the proof, we use without further reference the fact that the conditions (B_1) and (B_2) are uniform in p belonging to $\mathscr{F}$. By definition of $\bar{\hat{p}}_i^{(v)}$ in (3.2), we have, for $N = 1, 2$,

$$(n\varepsilon^{2v+k})^{1/2}(\bar{\hat{p}}_i^{(v)}(x_N) - E[\bar{\hat{p}}_i^{(v)}(x_N)])$$
$$= n^{-1/2} \Sigma_{j=1}^{n} \varepsilon^{(2v+k)/2}(\hat{p}_{j_i}^{(v)}(x_N) - E[\hat{p}_{j_i}^{(v)}(x_N)]).$$

Now we apply Lemma 3.2 of Hoeffding and Robbins to the double sequence of random variables $\{(X_{nj}, Y_{nj});\ n = 1, 2, \cdots,\ j = 1, \cdots, n\}$ where for $j = 1, \cdots, n$,

$$X_{nj} = \varepsilon^{(2v+k)/2}(\hat{p}_{j_i}^{(v)}(x_1) - E[\hat{p}_{j_i}^{(v)}(x_1)])$$

and

$$Y_{nj} = \varepsilon^{(2v+k)/2}(\hat{p}_{j_i}^{(v)}(x_2) - E[\hat{p}_{j_i}^{(v)}(x_2)]).$$

We first show that $n^{-1} \Sigma_{j=1}^{n} \operatorname{cov}(X_{nj}, Y_{nj}) \to 0$ as $n \to \infty$,

$$n^{-1} \Sigma_{j=1}^{n} \operatorname{var}(X_{nj}) \to c(x_1) \int (K_v^*(u))^2 d\mu(u)$$

and

$$n^{-1} \Sigma_{j=1}^{n} \operatorname{var}(Y_{nj}) \to c(x_2) \int (K_v^*(u))^2 d\mu(u).$$

Since

$$n^{-1} \Sigma_{j=1}^{n} \operatorname{cov}(X_{nj}\ Y_{nj}) = n\varepsilon^{2v+k} \operatorname{cov}(\bar{\hat{p}}_i^{(v)}(x_1), \bar{\hat{p}}_i^{(v)}(x_2)),$$

Lemma 3.1 implies that $n^{-1} \Sigma_{j=1}^{n} \operatorname{cov}(X_{nj}, Y_{nj}) \to 0$.

Now we consider $n^{-1} \Sigma_{j=1}^{n} \operatorname{var}(X_{nj})$ which is equal to

$$n^{-1}\varepsilon^{2v+k} \Sigma_{j=1}^{n} E[|\hat{p}_{j_i}^{(v)}(x_1)|^2] - n^{-1}\varepsilon^{2v+k} \Sigma_{j=1}^{n} (E[\hat{p}_{j_i}^{(v)}])^2.$$

The condition (B_1) and an application of Theorem 3.2 show that

$$n^{-1}\varepsilon^{2v+k} \Sigma_{j=1}^{n} E[|\hat{p}_{j_i}^{(v)}(x_1)|^2] \sim \bar{p}(x_1) \int |K_v^*(u)|^2 d\mu(u).$$

Again (B_1) and an application of Theorem 3.1 to p_j for $j = 1, \cdots, n$, show that for large n,

$$n^{-1}\varepsilon^{2v+k} \Sigma_{j=1}^{n} (E[\hat{p}_{j_i}^{(v)}(x_1)])^2 \leqq \varepsilon^{2v+k}(n^{-1} \Sigma_{j=1}^{n} (p_{j_i}^{(v)}(x_1))^2 + 2\bar{p}_i^{(v)}(x_1) + 1).$$

(B_1) implies that the rhs goes to zero as $n \to \infty$. Hence

$$n^{-1} \Sigma_{j=1}^{n} \operatorname{var}(X_{nj}) \to c(x_1) \int (K_v^*(u))^2 d\mu(u).$$

Similarly, the other variance result follows.

We now show that $n^{-1/2}(\Sigma_{j=1}^n E[|X_{nj}|^3])^{1/2} \to 0$ as $n \to \infty$. By definition of X_{nj} and the inequality $E[|R - ER|^3] \leqq 8E[|R|^3]$ for any random variable R, we obtain that $E[|X_{nj}|^3] \leqq 8\varepsilon^{3(2\nu+k)/2}E[|\hat{p}_{ji}^{(\nu)}(x_1)|^3]$. By applying Theorem 3.2 with $\delta = 3$ to p_j, we obtain that

$$n^{-1}\,\Sigma_{j=1}^n E[|\hat{p}_{ji}^{(\nu)}(x_1)|^3] \sim \varepsilon^{-(3\nu+2k)}\bar{p}(x_1)\int |K_\nu^*(u)|^3 d\mu(u)$$

as $n \to \infty$. Hence, $n^{-1/2}(\Sigma_{j=1}^n E[|X_{nj}|^3])^{1/3} \leqq$ a constant multiple of

$$(n\varepsilon^k)^{-1/6}(\bar{p}(x_1))^{1/3}(\int |K_\nu^*(u)|^3 d\mu(u))^{1/3}.$$

This bound goes to zero as $n \to \infty$ since $n\varepsilon^k \to \infty$ and $\bar{p}(x_1) \to c(x_1)$. Similarly, $n^{-1/2}(\Sigma_{j=1}^n E[|Y_{nj}|^3])^{1/3} \to 0$ completing the proof of asymptotic normality result.

3d. Rate of convergence for $E[|\bar{\hat{p}}^{(\nu)} - \bar{p}_i^{(\nu)}|]$.

Lemma 3.3 to follow is a direct generalization of the corresponding results stated in Parzen (1962) for $v = 0$ and $k = 1$ and in Cacoullos (1966) for $v = 0$ and $k \geqq 1$ when $X_1, \cdots, X_n, \cdots$ are i.i.d. random variables. (See the inequalities (4.6) and (4.8) of Susarla (1974) and (2.10) and (2.11) of Susarla (1974a) when $X_1, \cdots, X_n, \cdots$ are independent, but not identically distributed random variables.)

LEMMA 3.3. *If K_ν satisfies* $(A_1), (A_2), (A_3)$ *and* (A_4) *and* $\int K_\nu(u)u_i^\nu d\mu(u) = 1$, *and $\mathscr{F}$ satisfies (B_1) with v replaced by s, then*

$$\sup_x |E[\bar{\hat{p}}_i^{(\nu)}(x)] - \bar{p}_i^{(\nu)}(x)| \leqq c\varepsilon^{s-\nu} \text{ for some constant } c.$$

PROOF. By (3.7) and change of variables,

$$E[\bar{\hat{p}}_i^{(\nu)}(x)] - \bar{p}_i^{(\nu)}(x) = \int K_\nu^*(u)(\varepsilon^{-\nu}\bar{p}(x+\varepsilon u) - \bar{p}_i^{(\nu)}(x)\,\Sigma_{L_\nu}\pi u_m^{\nu_m})d\mu(u)$$

where Σ_{L_ν} denotes the sum over k-tuples of non-negative integers $(\nu_1, \cdots, \nu_k)$ such that $\nu_1 + \cdots + \nu_k = \nu$. By replacing $\bar{p}(x + \varepsilon u)$ by its sth order Taylor expansion with Lagrange's form of remainder and then using (3.7) and the orthogonality condition (A_3), we obtain that

$$|E[\bar{\hat{p}}_i^{(\nu)}(x)] - \bar{p}_i^{(\nu)}(x)| \leqq \varepsilon^{s-\nu}\,\Sigma_{L_s}\int |K_\nu^*(u)\pi u_m^{\nu m}|\,|\bar{p}^{(\nu_1),\ldots,(\nu_k)}(x+\theta u)|\,d\mu(u)$$

where θ is a function of u, x, ε and s. By using (B_1) and (A_4), the rhs can be uniformly bounded by a constant multiple of $\varepsilon^{s-\nu}$. This completes the proof.

NOTE: If the conditions (A_1) through (A_4) and (B_1) and (B_2) are satisfied and $\int K_\nu(u)u_i^\nu d\mu(u) = 1$, then the proof of Corollary 3.2 and the above result show that

$$E[|\bar{\hat{p}}_i^{(\nu)} - \bar{p}^{(\nu)}(x)|^2] \leqq c((n\varepsilon^{2\nu+k})^{-1} + \varepsilon^{2(s-\nu)}) \text{ for some constant } c.$$

https://doi.org/10.1017/S1446788700020541 Published online by Cambridge University Press

Solving for ε in terms of powers of n so that the expression in the bracket goes to zero at an optimal rate, we obtain that $\varepsilon = n^{-1/(2s+k)}$. For this choice of ε, we obtain

$$n^{2(s-v)/(2s+k)}E[|\bar{\tilde{p}}_i^{(v} - \tilde{p}_i^{(v)}(x)|^2]$$

is bounded. Special cases of this order relation are given in Cacoullos (1966), Parzen (1962), Susarla (1974) and Susarla (1974a).

4. Existence of K_v

From various statements of the results obtained in Section 3, it can be seen that the conditions can be divided into two categories, namely, those related to the kernel function K_v and those related to the family of densities $\mathscr{F}$. Our objective in this section is to show that the kernel function K_v satisfying (A_1) through (A_4), $K_v \to 0$ as $\|u\| \to \infty$ and $\int K_v(u)u_i^v d\mu(u) = 1$ does exist. Since discussion of conditions on K_v is similar for all k, we consider only the case $k = 2$ and $i = 1$ for easy exposition.

As in Section 5 of Susarla (1974), for any two positive integers M and N, let $[a_{pq}]_{M,N}$ denote a $M \times N$ matrix whose p, qth element is a_{pq}. Then, it has been noted in Section 5 of Susarla (1974) that, for any two sets of distinct positive integers $a_1, \cdots, a_M$ and $b_1, \cdots, b_N$, the vectors

$$(4.1) \qquad [p^{a_1}q^{b_1}]_{M,N}, \cdots, [p^{a_M}q^{b_N}]_{M,N}$$

form a basis for R^{MN}.

With $R^{p,q}$ representing the indicator function of the set $\{(a,b) \mid 0 \leqq a \leqq p, 0 \leqq b \leqq q\}$ for p, q in $\{1, \cdots, s\}$, we determine $[a_{pq}]_{s,s}$ such that

$$(4.2) \qquad K_v = \Sigma_{p,q=1}^{s} a_{pq} R^{p,q}$$

satisfies conditions (A_1) through (A_4) and the condition $\mu[K_v(u)u_i^v] = 1$. Since all a_{pq} will be finite, K_v of (4.2) satisfies conditions (A_1), (A_2), and (A_4) specialized to $k = 2$ case trivially. The conditions (A_3) and $\mu[K_v(u)u_1^v] = 1$ on K_v in (4.2) specialize to the following requirements on the inner product. (Inner product is taken in R^{s^2} space.)

$$(4.3) \qquad ([a_{pq}]_{s,s}, [p^{v_1}q^{v_2}]_{s,s}) = \begin{cases} 1 & \text{for } v_1 = v+1, v_2 = 1 \\ 0 & \text{for } v_1 + v_2 = v+3, \cdots, s+1. \end{cases}$$

(Here v_1, v_2 are positive integers.)

As a solution of the above Equation (4.3), we take $[a_{pq}]_{s,s}$ to be the projection of $[p^{v+1}q]_{s,s}$ onto

$$\perp \{[p^{v_1}q^{v_2}]_{s,s} \mid 1 \leqq v_1, v_2 \leqq s, (v_1, v_2) \neq (v+1, 1)\}$$

divided by its squared norm.

The squared norm mentioned above is different from zero due to the aforenoted linear independence of vectors in (4.1) for $(M, N) = (s, s)$. This completes the discussion of the existence of the function K_ν.

5. Final Remarks

The straightforward generalization of our results to the problem of estimation of mixed partial derivatives requires k-tuple sums which are similar to (3.8). Below, we describe an estimator $\bar{\bar{p}}^{(a_1),\cdots,(a_k)}$ for $\bar{p}^{(a_1),\cdots,(a_k)}(x)$ where $a_1 + \cdots + a_k = \nu$ and show that it is consistent. Other proofs are similar to the corresponding ones for $\bar{\bar{p}}^{(\nu)}$. With K_ν satisfying the requirements imposed on it in Section 3, let

$$K_\nu^{**}(u) = \Sigma_{r_1=0}^{a_1} \cdots \Sigma_{r_k=0}^{a_k} \pi\{(-1)^{r_m}(a_m - r_m + 1)^{-1}\binom{a_m}{r_m}\} K_\nu(Ju)$$

where J is the $k \times k$ matrix obtained by modifying the identity matrix by replacing the mth diagonal element by $(a_m - r_m + 1)^{-1}$ for $m = 1, \cdots, k$. Then it follows from (3.8) that

$$\int K_\nu^{**}(u)\pi u_m^{\nu_m} d\mu(u) = \begin{cases} \pi a_m! \int K_\nu(u)\pi u_m^{a_m} d\mu(u) & \text{for } \nu_m = a_m \text{ for } m = 1, \cdots, k \\ 0 & \text{for } \nu_m \neq a_m \text{ for some } m \text{ and } 0 \leqq \Sigma \nu_m \leqq \nu. \end{cases}$$

The intent and a result of the above equality is that if f is a function on R^k with $f^{(\nu_1),\cdots,(\nu_k)}(x) < \infty$ for $\nu_1 + \cdots + \nu_k = \nu$, then (3.9) follows with $f_i^{(\nu)}(x) \int K_\nu(u) u_i^\nu d\mu(u)$ replaced by

$$f^{(a_1),\cdots,(a_k)}(x) \int K_\nu(u)\pi u_m{}^{a_m} d\mu(u).$$

Consequently, the argument succeeding (3.9) gives that if f satisfies conditions of Theorem 3.1, then

$$\varepsilon^{-(\nu+k)} \int K_\nu^{**}(u/\varepsilon) f(x+u) d\mu(u) \to f^{(a_1),\cdots,(a_k)}(x) \int K_\nu(u)\pi u_m^{a_m} d\mu(u)$$

as $n \to \infty$. This result then implies (as in Corollary 3.1) that

$$\bar{\bar{p}}^{(a_1),\cdots,(a_k)} = (n\varepsilon^{\nu+k})^{-1} K_\nu^{**}((X_j - x)/\varepsilon)$$

is a consistent estimator of $\bar{p}^{(a_1),\cdots,(a_k)}(x) \int K_\nu(u)\pi u_m{}^{a_m} d\mu(u)$ under the hypothesis of Corollary 3.1.

The estimate $\bar{\bar{p}}_i^{(\nu)}$ exhibited in this paper can be shown to be strongly consistent under certain Lipschitz conditions on K_ν and some restrictions on ε and $\mathscr{F}$. We omit the proof of this result since it is similar to Theorem 1 of Van Ryzin (1969). For this result, we need the following conditions:

(1) K_ν and $\mathscr{F}$ satisfy conditions of Theorem 3.2.

(2) K_ν satisfies Lipschitz conditions (7) and (8) of Van Ryzin (1969). (K_ν of Section 4 satisfies this condition.)

(3) There exists a $\delta > 0$ such that

$$\sup_{\|u\| \leqq \delta} \left| \bar{p} \right]_x^{x+\varepsilon_n u} \Big| \leqq c(x)\,(< \infty) \text{ for all } n \geqq N.$$

(This condition is implied by (B_2)).

(4) $\varepsilon_n/\varepsilon_{n+1} \to 1$ as $n \to \infty$, $\Sigma_{n=1}^{\infty}(n\varepsilon_n^{\nu+k})^{-2} < \infty$ and

$$\Sigma_{n=1}^{\infty}(n\varepsilon_n^{2\nu+k-\beta})^{-1}\left|\varepsilon_{n+1}^{-1} - \varepsilon_n^{-1}\right|^{\beta} < \infty \text{ with } 2\beta = \min\{1, 2\alpha\}$$

where α is a parameter occurring in Condition (2).

Acknowledgements

The authors are grateful to the referee for pointing out that their proposed condition (3.3) can be replaced by the present weaker condition (3.3) and for some suggestions (notation or otherwise) which improved the clarity of the paper.

References

T. Cacoullos (1966), 'Estimation of a multivariate density', *Ann. Inst. Statist. Math.* **18,** 179–189.

W. Feller (1957), *An introduction to probability theory and its applications*, vol. I. (John Wiley & Sons, N.Y. 1957).

W. Hoeffding and H. Robbins (1948), 'The central limit theorem for dependent random variables', *Duke Math. J.* **15,** 773–780.

M. V. Johns Jr., and J. R. Van Ryzin (1972), 'Convergence rates for empirical Bayes two-action problems II. Continuous case', *Ann. Math. Statist.* **43,** 934–947.

Pi-Erh Lin (1968), 'Estimating partial derivatives of an unconditinal multivariate density (preliminary report'), *Ann. Math. Statist.* **39,** 1371.

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

V. Susarla (1974), 'Rate of convergence in the sequence-compound squared-distance loss estimation problem for a family of *m*-variate normal distributions', *Ann. Statist.* **2,** 118–133.

V. Susarla (1974), 'Rates of convergence in sequence-compound squared-distance loss estimation and linear-loss two-action problems for a family of scale parameter exponential distributions', *Ann. Inst. Statist. Math.* **26,** 53–67.

J. R. Van Ryzin (1969), 'On strong consistency of density estimates', *Ann. Math. Statist.* **40,** 1765–1772.

University of Wisconsin-Milwaukee
Milwaukee, U. S. A.
and
Bureau of Indian Affairs
Ottawa, Canada