

EFFICIENCY OF A SEQUENTIAL DENSITY ESTIMATOR UNDER AUTOREGRESSIVE DEPENDENCE MODEL

A.K. HOSNI AND M.M. EI-FAHHAM

Abstract. Using kernel estimates of Yamato type the effect of dependent observations is studied. The mean integrated square error of the Fourier integral estimator is considered.

The construction of a family of estimates of a density function $f(x)$ has been studied by several authors. Rosenblatt (1965) introduced the kernel estimator of the density $f(x)$, that is

$$\hat{f}_n(x) = \frac{1}{nh_n} \sum_{j=1}^n K\left(\frac{x - X_j}{h_n}\right), \quad (1)$$

and Parzen (1962) developed many of the important properties of these estimators. Here, of course, X_1, X_2, \dots, X_n are independent identically distributed (i.i.d.) random variables chosen according to some density f . Recently, Prazed and Singh (1980) gave a good review of kernel estimates of a density function.

A closed related estimator

$$f_n(x) = \frac{1}{n} \sum_{j=1}^n \frac{1}{h_j} K\left(\frac{x - X_j}{h_j}\right), \quad (2)$$

which has a very useful property that it can be calculated recursively, i.e.

$$f_n(x) = \frac{n-1}{n} f_{n-1}(x) + \frac{1}{nh_n} K\left(\frac{x - X_n}{h_n}\right), \quad (3)$$

was introduced by Wolverton and Wagner (1969), and apparently independently by Yamato (1971) Deheuvels (1974), Wegman and Davies, (1979), and Menon, Prazed and Singh (1983) extended techniques developed by Yamato.

Jeffrey Hart (1984) considered estimator of Parzen (1) and the problem of estimating the probability density function of a strictly stationary process is considered.

Received April 19, 1989.

Key words and phrases. Probability density estimation; Stationary processes; Fourier integral estimator.

The aim of this paper is to study the effect of dependent observations on the recursive estimators (2) of a probability density. The mean integrated square error of the Fourier integral estimator (FIE) is derived on the assumption that the observed data are generated by a first-order autorgressive process.

Let $X_j, j = 1, \dots, n$, be a sample from a process $\{X_j : | j | = 0, 1, 2, \dots\}$, which is assumed to be strictly stationary of order 2. Assume that each $X_j (| j | = 0, 1, \dots)$ has the probability density $f(x)$. In order to estimate the unknown density function $f(x)$ which in L_2 consider the following FIE recursive estimator

$$f_n(x) = \frac{1}{n\pi} \sum_{j=1}^n \frac{\sin(\frac{X_j - x}{h_j})}{(X_j - x)}, \tag{4}$$

where $\{h_j\}$ is a sequence of numbers such that

$$h_n > 0, n = 1, 2, \dots, \quad \lim_{n \rightarrow \infty} h_n = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} nh_n = \infty.$$

Suppose that the data are serially related as the first-order autoregressive model

$$X_j = \rho X_{j-1} + Z_j, \quad j = 2, 3, \dots, n, \tag{5}$$

where $|\rho| < 1, X_1, Z_2, \dots, Z_n$ are mutually independent.

Denote by $\Phi_{f_n}(t)$ as the characteristic function of $f_n(x)$. Hence we have

$$\begin{aligned} \Phi_{f_n}(t) &= \int_{-\infty}^{\infty} f_n(x) e^{itx} dx = \frac{1}{\pi n} \sum_{j=1}^n \int_{-\infty}^{\infty} \frac{\sin(\frac{X_j - x}{h_j})}{(X_j - x)} e^{itx} dx \\ &= \frac{1}{n} \sum_{j=1}^n e^{itX_j} \phi(th_j), \end{aligned} \tag{6}$$

where

$$\phi(th_j) = \int_{-\infty}^{\infty} \frac{\sin y}{\pi y} e^{iy(th_j)} dy.$$

We can easy see that : $\phi(th_j) = 1$ if $|t| < \frac{1}{h_j} = \lambda_j$ and equal to zero otherwise.

Then we get

$$\Phi_{f_n}(t) = \begin{cases} \frac{1}{n} \sum_{j=1}^n e^{itX_j} & \text{if } 0 = \lambda_0 < |t| < \lambda_1 \\ \frac{1}{n} \sum_{j=k}^n e^{itX_j} & \text{if } \lambda_{k-1} < |t| < \lambda_k; k = 2, 3, \dots, n \\ 0 & \text{if } |t| > \lambda_n. \end{cases} \tag{7}$$

Let the characteristic function of each X_j be denoted $\log \Phi_f(t)$. Hence using Parseval's theorem the mean integrated square error (MISE) of $f_n(x)$ is given by

$$\begin{aligned} MISE(f_n(x)) &= E \int_{-\infty}^{\infty} |f_n(x) - f(x)|^2 dx \\ &= \frac{1}{2\pi} E \int_{-\infty}^{\infty} |\Psi_{f_n}(t) - \Psi_f(t)|^2 dt \\ &= \frac{1}{\pi} E \left\{ \int_0^{\lambda_1} + \int_{\lambda_1}^{\lambda_2} + \dots + \int_{\lambda_n}^{\infty} |\Psi_{f_n}(t) - \Psi_f(t)|^2 dt \right\}. \end{aligned}$$

Since $E |\Psi_{f_n} - \Psi_f|^2 = \sigma^2 \Psi_{f_n} + b^2 \{\Psi_{f_n}\}$, where $b\{\Psi_{f_n}\} = |E\Psi_{f_n} - \Psi_f|$. Then the MISE given by

$$MISE(f_n(x)) = \frac{1}{\pi} \left\{ \int_{\lambda_0}^{\lambda_1} + \int_{\lambda_1}^{\lambda_2} + \dots + \int_{\lambda_n}^{\infty} [\sigma^2 \Psi_{f_n}(t) + b^2 \{\Psi_{f_n}(t)\}] dt \right\}.$$

Using (7) we can get

$$\begin{aligned} J(\rho) = MISE(f_n(x)) &= \frac{1}{\pi} \left\{ \int_{\lambda_0}^{\lambda_1} + \int_{\lambda_1}^{\lambda_2} + \dots + \int_{\lambda_n}^{\infty} [\sigma^2 \Psi_{f_n}(t) + b^2 \{\Psi_{f_n}(t)\}] dt \right\} \\ &\quad + \frac{1}{\pi} \int_{\lambda_n}^{\infty} |\Psi_f(t)|^2 dt \\ &= \frac{1}{\pi} \sum_{k=1}^n \int_{\lambda_{k-1}}^{\lambda_k} \sigma^2 \left[\frac{1}{n} \sum_{j=k}^n e^{itX_j} \right] dt \\ &\quad + \frac{1}{\pi} \int_{\lambda_0}^{\lambda_n} b^2 \{\Psi_{f_n}(t)\} dt + \frac{1}{\pi} \int_{\lambda_n}^{\infty} |\Psi_f(t)|^2 dt. \end{aligned} \tag{8}$$

First we calculate the first term of (8). Since for $k = 1, 2, \dots, n$

$$\begin{aligned} \sigma^2 \left[\frac{1}{n} \sum_{j=k}^n e^{itX_j} \right] &= E \left[\frac{1}{n} \sum_{j=k}^n e^{itX_j} \right]^2 - \left[E \left(\frac{1}{n} \sum_{j=k}^n e^{itX_j} \right) \right]^2 \\ &= \frac{1}{n^2} \sum_{j=k}^n \sum_{\ell=k}^n E e^{it(X_j - X_\ell)} - \left(\frac{n-k+1}{n} \right)^2 |\Psi_f(t)|^2 \\ &= \frac{n-k+1}{n^2} + \frac{2}{n^2} \text{Real} \left[\sum_{j=k}^{n-1} \sum_{\ell=j}^n E e^{it(X_j - X_\ell)} \right] \\ &\quad - \left(\frac{n-k+1}{n} \right)^2 |\Psi_f(t)|^2. \end{aligned}$$

Since the process $\{X_j\}$ is strictly stationary, we have

$$\begin{aligned} \sigma^2 \left[\frac{1}{n} \sum_{j=k}^n e^{itX_j} \right] &= \frac{n-k+1}{n^2} + \frac{2}{n^2} \operatorname{Real} \left[\sum_{j=k}^{n-1} (n-j) E e^{it(X_{j+1}-X_k)} \right] \\ &\quad - \left(\frac{n-k+1}{n} \right)^2 |\Psi_f(t)|^2. \end{aligned} \quad (9)$$

Since the process $\{X_j\}$ satisfy condition (5), hence

$$X_{j+1} - X_k = (\rho^j - \rho^{k-1})X_1 + \sum_{\ell=1}^{k-1} (\rho^{j-\ell} - \rho^{k-\ell-1})Z_{\ell+1} + \sum_{\ell=k}^j \rho^{j-\ell} Z_{\ell+1}.$$

From the mutual independence of X_1, Z_2, \dots, Z_n , we have

$$E e^{it(X_{j+1}-X_k)} = \Psi_f(t(\rho^j - \rho^{k-1})) \prod_{\ell=1}^{k-1} \Psi_z(t(\rho^{j-\ell} - \rho^{k-\ell-1})) \prod_{\ell=k}^j \Psi_Z(t\rho^{j-\ell}),$$

where $\Psi_Z(t)$ is the characteristic function of each of the random variables Z_i . Clearly we have

$\Psi_Z(t) = \Psi_f(t)/\Psi_f(\rho t)$, and, therefore

$$\begin{aligned} E e^{it(X_{j+1}-X_k)} &= \Psi_f(t(\rho^j - \rho^{k-1})) \prod_{\ell=1}^{k-1} \frac{\Psi_f(t(\rho^{j-\ell} - \rho^{k-\ell-1}))}{\Psi_f(t(\rho^{j-\ell+1} - \rho^{k-\ell}))} \prod_{\ell=k}^j \frac{\Psi_f(t\rho^{j-\ell})}{\Psi_f(t\rho^{j-\ell+1})} \\ &= \Psi_f(t) \frac{\Psi_f(t(\rho^{j-k+1} - 1))}{\Psi_f(t\rho^{j-k+1})} \end{aligned}$$

From (9) and the last expression, we have

$$\begin{aligned} \sigma^2 \left[\frac{1}{n} \sum_{j=k}^n e^{itX_j} \right] &= \frac{n-k+1}{n^2} - \left(\frac{n-k+1}{n} \right)^2 |\Psi_f(t)|^2 \\ &\quad + \frac{2}{n^2} \operatorname{Real} \left[\sum_{j=k}^{n-1} (n-j) \Psi_f(t) \frac{\Psi_f(t(\rho^{j-k+1} - 1))}{\Psi_f(t\rho^{j-k+1})} \right]. \end{aligned}$$

Then we have

$$\begin{aligned} &\frac{1}{n} \sum_{k=1}^n \int_{\lambda_{k-1}}^{\lambda_k} \sigma^2 \left[\frac{1}{n} \sum_{j=k}^n e^{itX_j} \right] dt \\ &= \frac{1}{\pi} \sum_{k=1}^n \frac{n-k+1}{n^2} (\lambda_k - \lambda_{k-1}) - \frac{1}{\pi} \sum_{k=1}^n \left(\frac{n-k+1}{n} \right)^2 \int_{\lambda_{k-1}}^{\lambda_k} |\Psi_f(t)|^2 dt \end{aligned}$$

$$\begin{aligned}
 & + \frac{2}{\pi n^2} \sum_{j=1}^{n-1} (n-j) \sum_{k=1}^j \int_{\lambda_{k-1}}^{\lambda_k} \text{Real} \left[\Psi_f(t) \frac{\Psi_f(t(\rho^{j-k+1} - 1))}{\Psi_f(t\rho^{j-k+1})} \right] dt \\
 & = \frac{1}{\pi n^2} \sum_{k=1}^n \lambda_k - \frac{1}{\pi} \sum_{k=1}^n \left(\frac{n-k+1}{n} \right)^2 \int_{\lambda_{k-1}}^{\lambda_k} |\Psi_f(t)|^2 dt \\
 & + \frac{1}{\pi n^2} \sum_{j=1}^{n-1} (n-j) \sum_{k=1}^j \int_{\lambda_{k-1}}^{\lambda_k} \text{Real} \left[\Psi_f(t) \frac{\Psi_f(t(\rho^{j-k+1} - 1))}{\Psi_f(t\rho^{j-k+1})} \right] dt. \tag{10}
 \end{aligned}$$

The second term of (8) follows from (7), and, we have

$$\begin{aligned}
 \frac{1}{n} \int_{\lambda_0}^{\lambda_n} b^2 \{ \Psi_{f_n}(t) \} dt & = \frac{1}{\pi} \sum_{k=1}^n \int_{\lambda_{k-1}}^{\lambda_k} |E\Psi_{f_n}(t) - \Psi_f(t)|^2 dt \\
 & = \frac{1}{\pi} \sum_{k=1}^n \int_{\lambda_{k-1}}^{\lambda_k} \left| \frac{1}{n} \sum_{j=k}^n Ee^{itX_j} - \Psi_f(t) \right|^2 dt \\
 & = \frac{1}{\pi} \sum_{k=1}^n \int_{\lambda_{k-1}}^{\lambda_k} \left[\frac{n-k+1}{n} - 1 \right]^2 |\Psi_f(t)|^2 dt \tag{11}
 \end{aligned}$$

From (8), (10) and (11) we get

$$\begin{aligned}
 J(\rho) & = \frac{1}{2\pi} \|\Psi_f\|^2 + \frac{1}{\pi n^2} \sum_{k=1}^n \lambda_k - \frac{2}{\pi n} \sum_{k=1}^n \int_{\lambda_0}^{\lambda_k} |\Psi_f(t)|^2 dt \\
 & + \frac{2}{\pi n^2} \sum_{j=1}^{n-1} (n-j) \sum_{k=1}^j \int_{\lambda_{k-1}}^{\lambda_k} \text{Real} \left[\Psi_f(t) \frac{\Psi_f(t(\rho^{j-k+1} - 1))}{\Psi_f(t\rho^{j-k+1})} \right] dt
 \end{aligned}$$

Hence we have the following theorem:

Theorem 1. *Let $f \in L^2$. Then under model (5) we have*

$$\begin{aligned}
 J(\rho) & = \frac{1}{2\pi} \|\Psi_f\|^2 + \frac{1}{\pi n^2} \sum_{j=1}^n \lambda_j - \frac{2}{\pi n} \sum_{j=1}^n \int_{\lambda_0}^{\lambda_j} |\Psi_f(t)|^2 dt \\
 & + \frac{2}{\pi n^2} \sum_{j=1}^{n-1} (j-j) \sum_{k=1}^j \int_{\lambda_{k-1}}^{\lambda_k} \text{Real} \left[\Psi_f(t) \frac{\Psi_f(t(\rho^{j-k+1} - 1))}{\Psi_f(t\rho^{j-k+1})} \right] dt
 \end{aligned}$$

where $\|g\|$ denotes the L^2 norm $(\int_{-\infty}^{\infty} |g(x)|^2 dx)^{1/2}$.

For the independent observations $\rho = 0$, we get the following theorem:

Theorem 2. *Let $f \in L^2$. Then under independent observations*

$$J(\rho = 0) = \frac{1}{2\pi} \|\Psi_f\|^2 + \frac{1}{\pi n^2} \sum_{j=1}^n \lambda_j - \frac{2}{\pi n^2} \sum_{j=1}^n j \int_0^{\lambda_j} |\Psi_f(t)|^2 dt.$$

Remark:

Results of Jeffrey Hart (1984) (Theorem 3.1) represent a special case of our results (Theorem 1) making in mind $\lambda_0 = 0$ and $\lambda_j = \lambda$ for all $j = 1, 2, \dots, n$.

We shall use the ratio $R(\rho) = J(0)/J(\rho)$ to assess the behavior of f_n under the model (5) and to quantify the loss in efficiency of f_n as ρ moves away from zero. Also we shall fix the smoothing parameter λ to obtain a number for $R(\rho)$. The value of λ that is used to obtain the results is the value that minimizes, for a given sample size, *MISE* in the case of independence. Three important densities (Normal, Cauchy and exponential) are used as examples to know what effect do different values of ρ on the behaviour of f_n .

A popular choice for the sequence λ_j is λj^γ and $0 < \gamma < 1$. Here and latter for the numerical results we shall take $\gamma = 1/3$.

Example 1.

For normal density with mean zero and variance one

$$\begin{aligned} J(\rho) = & \frac{1}{2\sqrt{\pi}} + \frac{\lambda}{\pi n^2} \sum_{j=1}^n j^{1/3} - \frac{2}{\pi n} \sum_{j=1}^n \int_0^{\lambda j^{1/3}} e^{-t^2} dt \\ & + \frac{2}{\pi n^2} \sum_{j=1}^{n-1} (n-j) \sum_{k=1}^j \int_{\lambda(k-1)^{1/3}}^{\lambda k^{1/3}} e^{-(1-\rho^{j-k+1})t^2} dt, \end{aligned}$$

and

$$J(\rho = 0) = \frac{1}{2\sqrt{\pi}} + \frac{\lambda}{\pi n^2} \sum_{j=1}^n j^{1/3} - \frac{2}{\pi n^2} \sum_{j=1}^n j \int_0^{\lambda j^{1/3}} e^{-t^2} dt.$$

Example 2.

For Cauchy density with parameters (0,1)

$$\begin{aligned} J(\rho) = & \frac{\lambda}{\pi n^2} \sum_{j=1}^n j^{1/3} + \frac{1}{\pi n} \sum_{j=1}^n e^{-2\lambda j^{1/3}} \\ & + \frac{2}{\pi n^2} \sum_{j=1}^{n-1} (n-j) \sum_{k=1}^j \frac{e^{\lambda C_{j,k}(k-1)^{1/3}} - e^{-\lambda C_{j,k}k^{1/3}}}{C_{j,k}} \end{aligned}$$

where

$$C_{j,k} = 2(1 - |\rho^{j-k+1}|),$$

and

$$J(\rho = 0) = \frac{1}{\pi} + \frac{\lambda}{\pi n^2} \sum_{j=1}^n j^{1/3} - \frac{1}{\pi n^2} \sum_{j=1}^n j(1 - e^{-2\lambda j^{1/3}})$$

Example 3.

For exponential density with parameter one

$$\begin{aligned} J(\rho) &= \frac{1}{2} + \frac{\lambda}{\pi n^2} \sum_{j=1}^n j^{1/3} - \frac{2}{\pi n} \sum_{j=1}^n \tan^{-1}(\lambda j^{1/3}) \\ &+ \frac{2}{\pi n^2} \sum_{j=1}^{n-1} (n-j) \sum_{k=1}^j (2 - \rho^{j-k+1})^{-1} [(1 - \rho^{j-k+1})^{-1} \{ \tan^{-1} [\\ &(1 - \rho^{j-k+1}) \lambda k^{1/3}] - \tan^{-1} [(1 - \rho^{j-k+1}) \lambda (k-1)^{1/3}] \} \\ &+ (1 - \rho^{j-k+1}) \{ \tan^{-1} (\lambda k^{1/3}) - \tan^{-1} [\lambda (k-1)^{1/3}] \}], \end{aligned}$$

and

$$J(\rho = 0) = \frac{1}{2} + \frac{\lambda}{\pi n^2} \sum_{j=1}^n j^{1/3} - \frac{2}{\pi n^2} \sum_{j=1}^n j \tan^{-1}(\lambda j^{1/3}).$$

It is not so easy to get the smoothing parameter λ , that minimizes $J(\rho = 0)$, for the above examples, hence we calculated $J(\rho = 0)$ numerically and table 1 gives the smoothing parameter λ under independence for four different sample sizes.

n	25	50	100	1000
Density				
Normal	.7579	.6771	.5988	.3882
Cauchy	.6787	.6557	.6163	.4465
Exponential	2.1309	2.4091	2.7143	3.9983

Table 1: *MISE* Optimal λ under independence.

Table 2 and 3 contain values of $R(\rho)$, where λ is *MISE* optimal under independence, at different value of ρ .

ρ	Cauchy				Exponential			
	n=25	n=50	n=10 ⁴	n=1000	n=25	n=50	n=100	n=1000
-0.9	0.6588	0.7164	0.7969	0.9839	0.5583	0.5579	0.5922	0.9893
-0.8	0.7860	0.8438	0.8993	0.9932	0.7880	0.8041	0.8339	0.9971
-0.7	0.8593	0.9040	0.9411	0.9964	0.9087	0.9186	0.9336	0.9990
-0.6	0.9055	0.9383	0.9632	0.9977	0.9694	0.9721	0.9772	0.9997
-0.5	0.9365	0.9597	0.9766	0.9986	0.9934	0.9974	0.9972	0.9999
-0.4	0.9582	0.9441	0.9852	0.9991	1.0130	1.0089	1.0081	1.00007
-0.3	0.9738	0.9841	0.9910	0.9995	1.0177	1.0129	1.0093	1.0001
-0.2	0.9853	0.9912	0.9951	0.9997	1.0166	1.0124	1.0090	1.0001
-0.1	0.9938	0.9963	0.9980	0.9998	1.0107	1.0081	1.0059	1.00007
-0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.1	0.9938	0.9963	0.9980	0.9999	0.9835	0.9873	0.9905	0.9998
0.2	0.9853	0.9912	0.9951	0.9997	0.9595	0.9683	0.9761	0.9997
0.3	0.9738	0.9841	0.9910	0.9995	0.9256	0.9408	0.9548	0.9994
0.4	0.9583	0.9741	0.9852	0.9991	0.8786	0.9011	0.9232	0.9989
0.5	0.9365	0.9597	0.9766	0.9986	0.8142	0.8445	0.8763	0.9981
0.6	0.9055	0.9382	0.9632	0.9977	0.7274	0.7637	0.8061	0.9967
0.7	0.8593	0.9040	0.9411	0.9962	0.6133	0.6502	0.7007	0.9940
0.8	0.7860	0.8438	0.8993	0.9932	0.4688	0.4948	0.5437	0.9877
0.9	0.6588	0.7164	0.7969	0.9839	0.2945	0.2957	0.3192	0.9662

Table 2. Values of $R(\rho)$ for Cauchy density.

Table 3. Values of $R(\rho)$ for Exponential density

From the above results we can see that:

- 1) i) For $\rho > 0$, for each density and for all sample size n , $f_n(x)$ is less efficient than it is in the case of independence.
 - ii) The loss of efficiency increases as $\rho \rightarrow 1$.
 - 2) i) For Cauchy density
 $R(\rho)$ diminishes monotonically as $\rho \rightarrow -1$ and $R(\rho) = R(-\rho)$.
 - ii) For exponential density
 $f_n(x)$ is actually more efficient than it is in the i.i.d. case for values $-0.1 < \rho < -0.4$.
- Comparing our results with that of HART, we note that for each density, some sample, size and ρ , $R(\rho)$ is greater than it is in HART's result which indicates that MISE using recursive estimators is more accurate.

References

- [1] Deheuvel's, P. *Estimation de la densité*. Doctorat d'Etat, Université Paris VI, (K), (1974).
- [2] Hart, J.D. Efficiency of kernel density estimator under an autoregressive dependence model. *J. Amer. Stat. Ass.* Vol. 79, 385, (1984), 110-117.
- [3] Menon, V.V., Prasad, B. and Singh, R.S. Non-parametric recursive estimates of a probability density function and its derivatives. *J. Stat. Planning and Inference.* 9, (1983), 71-82.
- [4] Parzen, E. On estimation of a probability density function and mode. *Ann. Math. Stat.* V. 33, 8, (1962), 1065-1076.
- [5] Prasad, B. and Singh, R.S. Nonparametric kernel estimates of a density function along with its derivatives. *Colloquia Math. Soc. Janos, Bolyai, Nonparametric Stat. Inference Budapest*, (1980).
- [6] Rosenblatt, M. Remarks on some nonparametric estimates of a density function, *Ann. Math. Stat.* 27, (1956), 832-837.
- [7] Wegman, E.J. and Davies, M.I. Remarks on some recursive estimators of probability density. *Ann. of Stat.*, 7, (1979), 316-327.
- [8] Wolverton, C. and Wanger, T. Recursive estimates of probability density. *IEEE Trans. Syst. Sci. Cybernet.* 5, (1969), 246-257.
- [9] Yamato, H. Sequential estimation of a continuous probability density function and mode. *Bulletin of Math. Stat.*, 14, (1971), 1-2.

A.K. hosni

Department of Mathematics, Faculty of Science, Banha University, Banha, EGYPT.

M.M. El-Fahham

Department of Mathematics, Faculty of Science, Cairo University, Giza, EGYPT.