ISSN 1028-3358, Doklady Physics, 2011, Vol. 56, No. 9, pp. 471–475. © Pleiades Publishing, Ltd., 2011.
Original Russian Text © V.F. Kravchenko, V.I. Pustovoit, D.V. Churikov, 2011, published in Doklady Akademii Nauk, 2011, Vol. 440, No. 2, pp. 182–186.

THEORETICAL PHYSICS

Atomic Functions and Nonparametric Estimates of the Probability Density

V. F. Kravchenko[a], Academician V. I. Pustovoit[b], and D. V. Churikov[a]

Received May 26, 2011

DOI: 10.1134/S1028335811090060

For the first time, nonparametric estimates of the probability density and its derivatives are considered on the basis of the theory of atomic functions (AFs) [1–4]. New constructions of weight functions (WFs) with compact support are proposed and substantiated that allow building admissible estimates of both the probability density in itself and its first and second derivatives. The simple structure of the nonparametric estimates and the possibility of using them in the cases when a recovered variable is unknown make them suitable for physical applications [5, 6]. The proposed novel mathematical apparatus of nonparametric statistics on the basis of the AFs allows estimating characteristics of series in the absence of a priori parametric information.

ADMISSIBLE ESTIMATES OF THE PROBABILITY DENSITY

Let $X_1, X_2, \ldots, X_n$ be the sampling of n independent observations of random variable X with the unknown probability density function $f(x)$. According to studies [5, 6], the nonparametric estimate is determined as

$$f_n(x) = \frac{1}{nh}\sum_{j=1}^{n} K\left(\frac{X_j - x}{h}\right), \tag{1}$$

where $h = h(n)$ is some series of positive numbers. Here, $\lim_{n\to\infty} h(n) = 0$ and $K(x)$ is the even function satisfying the normalization condition $\int_{-\infty}^{\infty} K(x)dx = 1$ and $K(x) \in L_2$. The weight function $K(x)$ is called admissible if its Fourier transform is nonnegative and does not exceed unity for all the real frequencies. For the first moments of certain function H from statistics x_n, in the case when the characteristics of the statistics moments are determined as the moments of the limiting distribution of the estimate (mean, dispersion, and root-mean-square error), under the condition of weak convergence, the following theorem holds [5].

Theorem 1. If the distribution of a series of s-dimensional random variables at $n \to \infty$ converges to the s-dimensional normal distribution $N_s\{\mu, \sigma\}$ with the vector of means $\mu = (\mu_1, \mu_2, \ldots, \mu_s)$ and the covariance matrix σ $(0 < \sigma_{ij} = \sigma_{ij}(x) < \infty, j = 1, 2, \ldots, s)$, i.e., $d_n(x_n - x) \Rightarrow N_s\{\mu, \sigma\}$, $H(x) \in \mathcal{N}_{1,s}$, and $\nabla H(x) \neq 0$, then

$$d_n(H(x_n) - H(x)) \Rightarrow N_1\left\{\sum_{j=1}^{s} H_j(x)\mu_j, \sum_{j,p=1}^{s} H_j(x)H_p(x)\sigma_{j,p}\right\} = N_1\left\{\nabla H(x)\mu^{\mathrm{T}}, \nabla H(x)\sigma\nabla^{\mathrm{T}}H(x)\right\}.$$

Here $\mathcal{N}_{\nu,s}$ is the class of the functions for which all the partial derivatives up to the νth order exist and $d_n \to \infty$. Let $\nabla^2 H(x)$ be the matrix of the second derivatives with the $\left.\frac{\partial^2 H(z)}{\partial z_i \partial z_j}\right|_{z=x}$, where $i, j = 1, 2, \ldots, s$, $\eta = (\eta_1, \eta_2, \ldots, \eta_s)$ is the s-dimensional random variable; X^2 is the one-dimensional random variable that has the distribution χ^2 with one degree of freedom and N is the random variable distributed following the standard normal law $N_1\{0, 1\}$. The following three cases are possible:

(1) If $d_n(x_n - x) \Rightarrow \eta = (\eta_1, \eta_2, \ldots, \eta_s)$, $H(x) \in \mathcal{N}_{1,s}$, and $\nabla H(x) \neq 0$, then $d_n(H(x_n) - H(x)) \Rightarrow \sum_{j=1}^{s} H_j(x)\eta_j = \nabla H(x)\eta^{\mathrm{T}}$.

[a] *Kotel'nikov Institute of Radio Engineering and Electronics, Russian Academy of Sciences, Building 7, ul. Mokhovaya 11, Moscow, 125009 Russia*
[b] *Scientific and Technological Center of Unique Instrumentation, Russian Academy of Sciences, Moscow, 117342 Russia*
e-mails: kvf-ok@mail.ru, mpio_nice@mail.ru

(2) If $d_n(x_n - x) \Rightarrow \eta = (\eta_1, \eta_2, \ldots, \eta_s)$, $H(x) \in \mathcal{N}_{2,s}$, $\nabla H(x) = 0$, and $\nabla^2 H(x) \neq 0$, then $d_n^2(H(x_n) - H(x)) \Rightarrow \frac{1}{2}\sum_{i,j=1}^{s} H_{i,j}(x)\eta_i\eta_j = \frac{1}{2}\eta\nabla^2 H(x)\eta^{\mathrm{T}}$.

(3) If $d_n(x_n - x) \Rightarrow N_s\{0, I_s\}$, where I_s is the identity diagonal matrix of the sth order, then $d_n^2(H(x_n) - (x))$ converges by the distribution to the sum of weighted sums s of the random variables X^2 and $s(s-1)$ of the products of independent standard normal random variables, i.e., $H(x) \in \mathcal{N}_{2,s}$, $\nabla H(x) = 0$, and $\nabla^2 H(x) \neq 0$, then

$$d_n^2(H(x_n) - H(x)) \Rightarrow \frac{1}{2}\sum_{j=1}^{s} H_{j,j}(x) X_j^2 + \frac{1}{2}\sum_{i\neq j}^{s} H_{i,j}(x) N_i N_j.$$

The criterion of quality of the estimate of $f_n(x)$ is the integral root-mean-square error determined as

$$E = \left\langle \int_{-\infty}^{\infty} (f_n(x) - f(x))^2 dx \right\rangle.$$

Here the angle parenthesis denotes averaging over the ensemble. In the case when the WF is admissible, the integral root-mean-square error of the estimate obtained using this WF can be reduced for all the probability densities $f(x)$ simultaneously.

NONPARAMETRIC KERNEL ESTIMATES

Let us consider constriction of the admissible WFs by the example of the AF $h_a(x)$. As is known [1], the AFs $h_a(x)$ $(a > 1)$ are the finite solutions of the functional differential equation

$$y'(x) = \frac{a^2}{2}(y(ax+1) - y(ax-1)).$$

The main properties of the AF $h_a(x)$ are the following:

(1) $h_a(x) = 0$ at $|x| \geq (a-1)^{-1}$.

(2) $h_a(x) = \frac{a}{2}$ at $|x| \leq \frac{a-2}{a(a-1)}$, $a \geq 2$.

(3) The Fourier transform of $h_a(x)$ is $\varphi_a(\omega) = \prod_{k=1}^{\infty} \frac{\sin(\omega/a^k)}{\omega/a^k}$ and takes on zero value at the points $\omega = 2\pi n$, $n \neq 0$. If the value of ω is small, then in the numerical implementation it is sufficient to limit the consideration to a small number of terms of the product, since they rapidly tend to 1 with increasing k.

Using property (3), we may write the function $h_a(x)$ in the interval $x \in \left[-\frac{1}{a-1}, \frac{1}{a-1}\right]$ as

$$h_a(x) = (a-1)\left(\frac{1}{2} + \sum_{k=1}^{\infty} \varphi_a((a-1)\pi k)\cos((a-1)\pi k x)\right).$$

Let introduce the function $ch_a(x) = h_a(x) * h_a(x)$, which is the convolution of the AF $h_a(x)$ with itself. The function is even, infinitely differentiable, and has the support $\left[-\frac{2}{a-1}, \frac{2}{a-1}\right]$. Repeating the convolution operation for $(l-1)$ times, obtain the function

$$ch_{a,l}(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} [\varphi_a(\omega)]^l \exp(i\omega x)\, d\omega, \quad l = 1, 2, \ldots,$$

$$\mathrm{supp}\, ch_{a,l}(x) = \left[-\frac{l}{a-1}, \frac{l}{a-1}\right],$$

$$\int_{-l(a-1)^{-1}}^{l(a-1)^{-1}} ch_{a,l}(x)\, dx = 1.$$

An important property of this function is that its Fourier transform for the even index

$$\int_{-\infty}^{\infty} ch_{a,2l}(x)\exp(-i\omega x)\, dx = [\varphi_a(\omega)]^{2l}$$

is nonnegative and does not exceed unity. The admissible WFs are built using the expression in the frequency space $\Psi_r(\omega) = 1 - (1 - \varphi_a(\omega))^{r/2}$, $r = 2, 4, \ldots$, which is also positive and does not exceed unity. Then, the admissible WF is calculated as

$$K_{a,r}(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \{1 - (1 - \varphi_a^2(\omega))^{r/2}\}\exp(i\omega x)\, d\omega, \quad r = 2, 4, \ldots \tag{2}$$

For instance, for several values of r, we obtain $K_{a,2}(x) = ch_{a,1}(x)$, $K_{a,4}(x) = 2ch_{a,1}(x) - ch_{a,2}(x)$, and $K_{a,6}(x) = 3ch_{a,1}(x) - 3ch_{a,2}(x) + ch_{a,3}(x)$. The form of the WFs $K_{a,r}(x)$ and their derivatives is presented in Fig. 1.

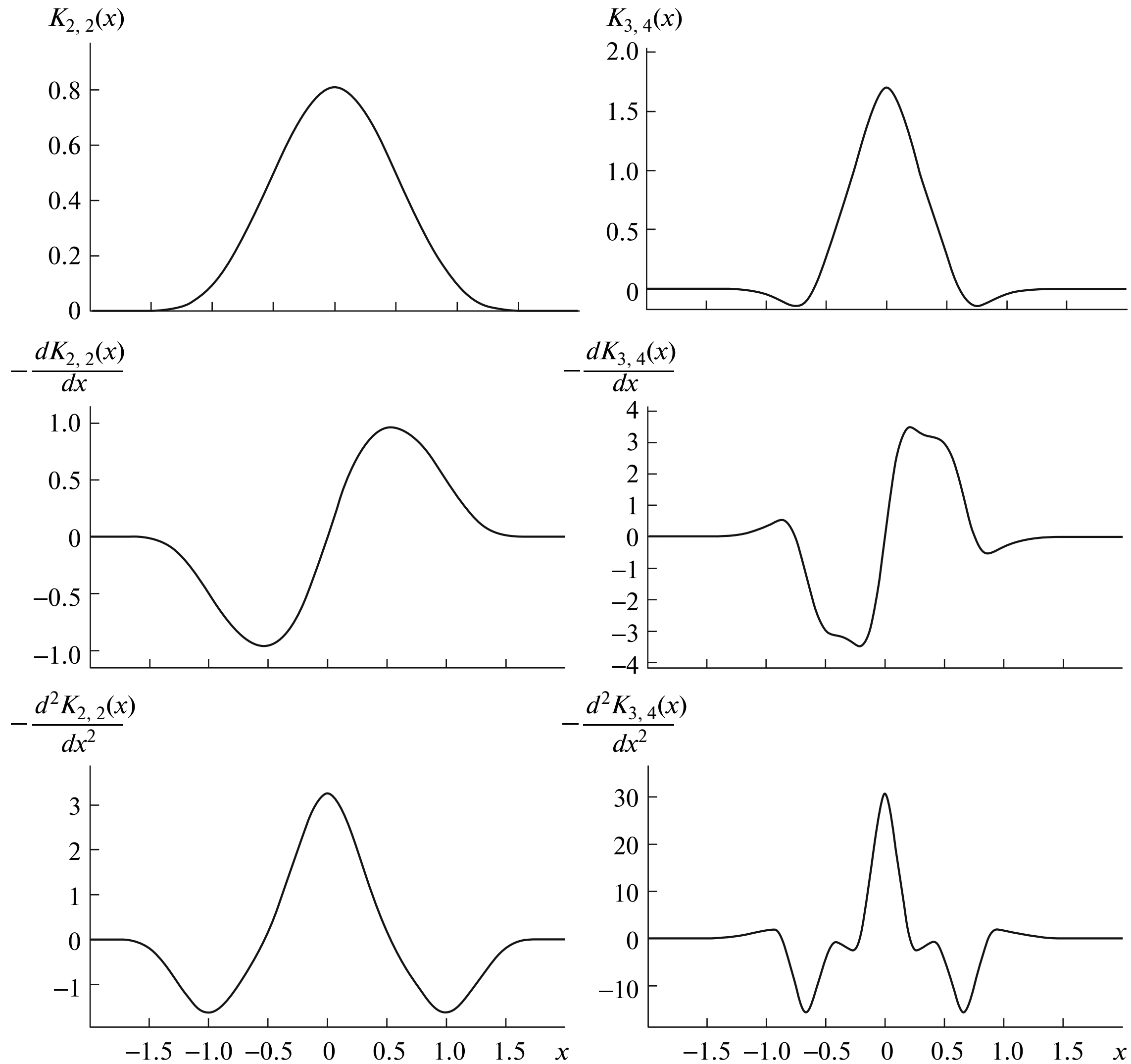

Fig. 1. Form of the weight functions $K_{a,r}(x)$ and their first and second derivatives for $a = 2$, $r = 2$ (on the left) and $a = 3$, $r = 4$ (on the right).

ADMISSIBLE ESTIMATES OF DERIVATIVES OF THE PROBABILITY DENSITY FUNCTION

We write the estimate of the value of $f'(x)$ as

$$Df_n(x) = \frac{1}{nh^2}\sum_{j=1}^{n} N\left(\frac{X_j - x}{h}\right), \quad (3)$$

where $h = h(n)$ is the decreasing sequence of positive numbers, $N(x)$ is the even function $\left(\int_{-\infty}^{\infty} N(x)dx = 1\right)$, and $N(x) \in L_2$. If take $N_r(x) = -K'_{r-1}(x)$, $r = 3,5,\ldots$, then we obtain the admissible weight functions of rth orders. Similarly, for the second derivatives of the probability density function, we have

$$D^2 f_n(x) = \frac{1}{nh^3}\sum_{j=1}^{n} M\left(\frac{X_j - x}{h}\right), \quad (4)$$

where $h = h(n)$ is the decreasing sequence of positive numbers, $M(x)$ is the even function $\left(\int_{-\infty}^{\infty} M(x)dx = 1\right)$, and $M(x) \in L_2$. If take $M_r(x) = -K''_{r-2}(x)$, $r = 3,5,\ldots$, then we obtain the admissible WFs of the rth orders.

PHYSICAL CHARACTERISTICS OF THE ADMISSIBLE WEIGHT FUNCTIONS

To study spectral kernels, we use the modified physical characteristics [1–3, 7–10]: the width of the spectral density function at the level of –3 dB (γ_3), the

Physical characteristics of the WF $K_{a,r}(x)$ for different parameters a and r

r	a	γ_3	γ_4/γ_3	γ_9	P	Δ	supp	supp_E
2	2	1.720	1.391	−46.586	0.945	0.502	4	2.029
	3	2.693	1.444	−34.028	0.735	0.508	2	1.233
	4	3.815	1.412	−30.492	0.622	0.513	4/3	0.898
	5	4.787	1.375	−28.988	0.549	0.518	1	0.708
	6	5.610	1.467	−28.213	0.496	0.522	4/5	0.587
	7	6.732	1.400	−27.755	0.456	0.525	2/3	0.501
4	2	2.581	1.232	−40.585	0.883	0.544	8	2.669
	3	4.114	1.236	−28.094	0.682	0.560	4	1.662
	4	5.610	1.240	−24.602	0.573	0.574	8/3	1.218
	5	7.181	1.229	−23.124	0.503	0.585	2	0.964
	6	8.602	1.239	−22.358	0.454	0.593	8/5	0.796
	7	10.098	1.222	−21.908	0.416	0.599	4/3	0.682
6	2	3.017	1.182	−37.084	0.852	0.588	12	2.669
	3	4.837	1.186	−24.658	0.654	0.614	6	1.674
	4	6.583	1.182	−21.208	0.547	0.638	4	1.225
	5	8.378	1.167	−19.756	0.478	0.656	3	0.972
	6	10.098	1.173	−19.005	0.429	0.670	12/5	0.805
	7	11.819	1.165	−18.565	0.393	0.680	2	0.687

relative width of the spectral density function at the level of −6 dB (γ_4/γ_3), the maximum level of side lobes in dB (γ_9), the L_2 norm of a WF (P), the uncertainty constant (Δ), the support (supp), and the effective support $\text{supp}_E f(x) = \{x \| \|f(x)\|_{L_2} = 0.999P\}$. The physical characteristics of the WF $K_{a,r}(x)$ for different a and r are given in the table.

Let consider the example of estimating the probability density function of a sequence of random variables X_n with the normal distribution law

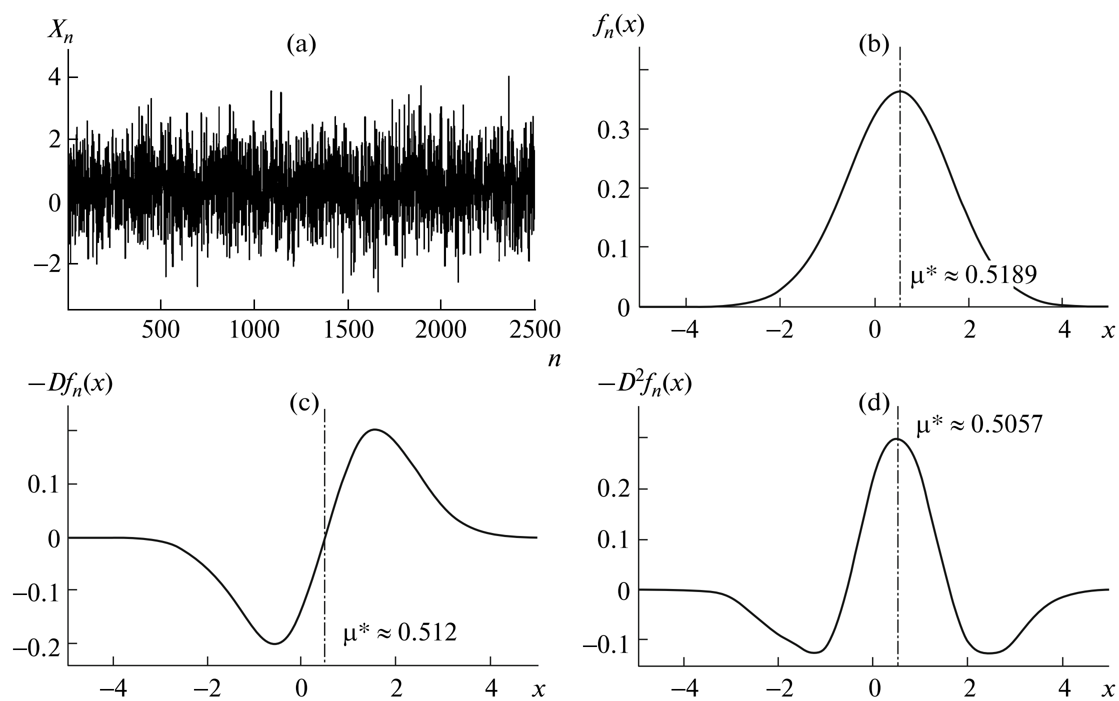

Fig. 2. (a) Sequence of random variables X_n ($n = 2500$) and (b–d) estimate of $f_n(x)$, $Df_n(x)$, $D^2f_n(x)$ at $a = 2$ and $r = 2$.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right),$$

where the average of distribution is $\mu = 0.5$ and the standard deviation is $\sigma = 1$ (Fig. 2a). The results of estimation of the distribution density function $f(x)$ and two of its derivatives for $a = 2$, $r = 2$, and $n = 2500$ are shown in Figs. 2b–2d.

Thus, the new constructions of the WFs with a compact support on the basis of the AF theory are proposed. The admissible nonparametric estimations of the probability density and its first and second derivatives for the sequence of random variables are constructed. The numerical experiment and physical analysis confirm the efficiency of the novel nonparametric estimations of the probability density function.

The results of this work were reported at the International conference "Days on Diffraction 2011" [9, 10].

REFERENCES

1. V. F. Kravchenko, *Lectures on the Theory of Atomic Functions and Some Their Applications* (Radiotekhnika, Moscow, 2003) [in Russian].
2. V. F. Kravchenko and V. L. Rvachev, *Boolean Algebra, Atomic Functions, and Wavelets in Physical Applications* (Fizmatlit, Moscow, 2006) [in Russian].
3. *Digital Signal and Image Processing in Radio Physical Applications*, Ed. by V. F. Kravchenko (Fizmatlit, Moscow, 2007) [in Russian].
4. V. F. Kravchenko, H. M. Perez-Meana, and V. I. Ponomarev, *Adaptive Digital Processing of Multidimensional Signals with Applications* (Fizmatlit, Moscow, 2009).
5. A. V. Dobrovidov and G. M. Koshkin, *Nonparametric Estimation of Signals* (Fizmatlit, Moscow, 1997).
6. A. V. Vasil'ev, A. V. Dobrovidov, and G. M. Koshkin, *Nonparametric Estimation of Functionals of Stationary Sequences* (Nauka, Moscow, 2004) [in Russian].
7. V. F. Kravchenko and D. V. Churikov, in *Proceedings of the IEEE X International Conference on Signal Processing, Taiyangdo, Beijing, China, 2010*, p. 1906.
8. V. F. Kravchenko and D. V. Churikov, in *Proceedings of the International Conference on Ultrawideband and Ultrashort impulse Signals, Sevastopol, Ukraine, 2010*, p. 59.
9. V. F. Kravchenko and D. V. Churikov, in *Proceedings of the XII International Conference and Exhibition on Digital Signal Processing and Its Applications, Moscow, Russia, 2011*, p. 163.
10. V. F. Kravchenko and D. V. Churikov, in *Proceedings of the International Conference Days on Diffraction, St. Petersburg, Russia, 2011*, p. 63.

Translated by E. Bondareva