Seminar of I. Vekua Institute
of Applied Mathematics
REPORTS, Vol. 35, 2009

ON AN INTEGRAL SQUARE DEVIATION MEASURE WITH THE WEIGHT OF "DELTA-FUNCTIONS" OF THE ROSENBLATT–PARZEN PROBABILITY DENSITY ESTIMATOR

Nadaraya E., Babilua P., Sokhadze G.

Abstract. The limit distribution of an integral square deviation with the weight of "delta-functions" of the Rosenblatt–Parzen probability density estimator is defined. Also, the limit power of the goodness-of-fit test constructed by means of this deviation is investigated.

Keywords and phrases: Distribution density, goodness-of-fit test, power, consistency, limit distribution.

AMS subject classification (2000): 62G07; 62G10; 62G20.

1. Limit distributions of some global measures of distributions of estimates $f_n(x)$ of the density $f(x)$ such as, for example, an integral square deviation constructed by means of the so-called weight function $W(x)$ not depending on n were studied in P. Bickel and M. Rosenblatt [1], E. Nadaraya ([2], [3]), P. Hall [4] and other works.

The theory of the asymptotic behavior of an integral mean-square error

$$R(f_n, f; W_n) = E \int W_n(x) \left(f_n(x) - f(x)\right)^2 dx, \tag{1}$$

was developed in the work [5] of T. Tony Cai and Mark G. Low, where $W_n(x) = a_n W(a_n(x - \ell_0))$, $\{a_n\}$ is a sequence of positive numbers, $W(x) \geq 0$ is a Borel-measurable function and ℓ_0 is some fixed point. If in (1) we put $W(x) = \frac{1}{2} I(-1 \leq x \leq 1)$ and pass to the limit as $a_n \to \infty$ then, roughly speaking, $R(f_n, f; W_n) \simeq E(f_n(\ell_0) - f(\ell_0)^2$. If, however, we put $a_n \equiv 1$ in (1) for all n, $\ell_0 = 0$ and assume that $W(x) \geq 0$ is an arbitrary bounded function, then $R(f_n, f; W_n) = E\|f_n - f\|^2_{L_2(W_n)}$. Thus the value $R(f_n, f; W_n)$ can be considered as *a generalization* of a measure of density estimation accuracy which contains a mean-square deviation of the estimate $f_n(x)$ of the density at the point and an integral mean-square deviation. Therefore it is natural to pose the question on the limit distribution of the value $\|f_n - f\|^2_{L_2(W_n)}$, $W_n(x) = a_n W(a_n(x - \ell_0))$. In the present paper this question is considered for the case where $f_n(x)$ is a nonparametric estimate of the Rosenblatt–Parzen density and $a_n \to \infty$ as $n \to \infty$. The case $a_n \to a < \infty$ is of no interest because it follows from the results of the works [1], [2], [3] and [4].

Let $X_1, X_2, \ldots, X_n$ be independent, equally distributed random values having the unknown probability density function $f(x)$ and consider the Rosenblatt–Parzen nonparametric estimator $f_n(x)$ of the density $f(x)$,

$$f_n(x) = \lambda_n / n \sum_{i=1}^{n} K(\lambda_n(x - X_i)),$$

where $K(x)$ is a function belonging to the class

$$H = \Big\{K: \int K(x) = 1,\ K(-x) = K(x),\ \sup_x |K(x)| < \infty,\ x^2K(x) \in L_1(-\infty,\infty)\Big\},$$

and $\{\lambda_n\}$ is a sequence of positive numbers converging to infinity.

Notation.

$$U_n = \frac{n}{\lambda_n}\,\|f_n - f\|^2_{L_2(W_n)},\quad U_n^{(1)} = n\|f_n - Ef\|^2_{L_2(W_n)},\quad \Delta_n = EU_n^{(1)},$$

$$\alpha_n(x,y) = \lambda_n\left[K(\lambda_n(x-y)) - EK(\lambda - X_1))\right],$$

$$\sigma_n^2 = 2\iint \left[E\alpha_n(u_1,X_1)\alpha_n(u_2,X_1)\right]^2 W_n(u_1)W_n(u_2)\,du_1\,du_2,$$

$$\eta_{ij}^{(n)} = 2n^{-1}\sigma_n^{-1}\int \alpha_n(x,X_i)\alpha_n(x,X_j)W_n(x)\,dx,$$

$$\xi_j^{(n)} = \sum_{j=1}^{j-1}\eta_{ij}^{(n)},\ \ j=2,\dots,n,\ \ \xi_1^{(n)}=0,\ \ \xi_j^{(n)}=0,\ \ j>n,\ \ F_k^{(n)}=\sigma(\omega:\ X_1,X_2,\dots,X_k).$$

Lemma 1. *The stochastic sequence* $(\xi_j^{(n)},\mathcal{F}_j^{(n)})_{j\geq 1}$ *is a difference-martingale.*

Lemma 2. *Let* $K(x)\in H$, $f(x)\in F$ *(*F *is the set of bounded functions on* $R=(-\infty,\infty)$ *which have bounded derivatives up to second order inclusive),* $W(x)$ *be bounded and* $W\in L_2(R)$. *If* $\lambda_n\to\infty$, $a_n\to\infty$ *and* $a_n/\lambda_n\to 0$ *as* $n\to\infty$, *then*

$$(\lambda_n a_n)^{-1}\sigma_n^2 \longrightarrow 2f^2(\ell_0)\int K_0^2(z)\,dz\int W^2(v)\,dv,\ \ K_0 = K*K,\ \ f(\ell_0)\neq 0.$$

Theorem 1. *Let* $K(x)\in H$, $f(x)\in F$, $W(x)$ *be bounded and* $W\in L_1(R)$. *If* $a_n\to\infty$, $a_n/\lambda_n\to 0$ *and* $n^{-1}\lambda_n a_n^2\to 0$ *as* $n\to\infty$, *then*

$$\sigma_n^{-1}(U_n^{(1)}-\Delta_n)\xrightarrow{d} N(0,1),$$

where d *denotes the convergence in distribution, and* $N(0,1)$ *is a random value having a normal distribution with a zero mean value and variance* 1.

Proof. We have

$$\sigma_n^{-1}(U_n^{(1)}-\Delta_n) = \sqrt{\frac{n-1}{n}}\,H_n^{(1)} + H_n^{(2)},\ \ H_n^{(1)} = \sum_{j=1}^n \xi_j^{(n)},$$

and also

$$\operatorname{var} H_n^{(2)} = O((\lambda_n a_n)/n) + O(n^{-1}\sigma_n^{-2}),$$

i.e. $H_n^{(2)}\xrightarrow{d} 0$.

The asymptotic normality of $H_n^{(1)}$ takes place [6] if for each $\varepsilon\in(0,1]$ and $n\to\infty$

$$\sum_{k=1}^n E\Big[(\xi_k^{(n)})^2 I\Big(|\xi_k^{(n)}|\geq\varepsilon\Big)/\mathcal{F}_{k-1}^{(n)}\Big]\xrightarrow{d} 0\ \ \text{(the Lindeberg condition)},$$

$$V_n^2 = \sum_{k=1}^n E\Big((\xi_k^{(n)})^2/\mathcal{F}_{k-1}^{(n)}\Big)\xrightarrow{d} 1.$$

First let us verify that $V_n^2 \xrightarrow{d} 1$. For this, taking the definition of $\xi_j^{(n)}$ into account, we can write V_n^2 in the form

$$V_n^2 = \sum_{j=2}^{n} E\Big(\sum_{i=1}^{j-1} (\eta_{ij}^{(n)})^2 |\ X_1, \dots, X_{j-1}\Big) + 2\sum_{j=2}^{n} E \sum_{i=1}^{j-1} \sum_{i=j+1}^{j-1} \Big(\eta_{ij}^{(n)} \eta_{\ell j}^{(n)} |\ X_1, \dots, X_{j-1}\Big)$$
$$= V_{n1} + V_{n2}.$$

It is not difficult to show that

$$\operatorname{Var} V_{n1} = \frac{16\lambda_n^n}{n^4 \sigma_n^4} \Big[\sum_{j=2}^{n} (j-1) E(\varepsilon_1 - \overline{\sigma}_n)^2 + 2\sum_{i=2}^{n} E Z_i^2 (n-i)\Big] = B_{n1} + B_{n2},$$

and also

$$B_{n1} = O\Big(\frac{a_n^2}{n^2}\Big), \quad B_{n2} = O\Big(\frac{a_n^2}{n}\Big),$$
$$\varepsilon_i = \lambda_n^{-2} \iint \alpha_n(x, X_i) \alpha_n(y, X_i) \Phi_n(x,y) W_n(x) W_n(y)\, dx\, dy,$$
$$\Phi_n(x,y) = EK(\lambda_n(x - X_1)) K(\lambda_n(y - X_1)) - EK(\lambda_n(x - X_1)) EK(\lambda_n(y - X_1)),$$
$$Z_j = \sum_{i=1}^{j-1} (\varepsilon_i \overline{\sigma}_n), \quad \overline{\sigma}_n = \iint \Phi_n^2(x,y) W_n(x) W_n(y)\, dx\, dy.$$

Therefore $\operatorname{Var} V_{n1} \to 0$. On the other hand, $EV_{n1} = 1 - 1/n \to 1$. Therefore $V_{n1} \xrightarrow{d} 1$.

Now let us consider V_n^2. Taking into account the inequality

$$E\Big(\sum_{i=1}^{n} Y_j\Big)^2 \le \Big(\sum_{i=1}^{n} (EY_i^2)^{1/2}\Big)^2$$

which is easy to verify and performing some simple calculations, we obtain $EV_{n2}^2 = O(a_n/\lambda_n)$. Therefore $V_n^2 \xrightarrow{d} 1$.

Now we will establish the validity of the Lindeberg condition. For this, it suffices to make sure that $\sum\limits_{j=1}^{n} E(\xi_j^{(n)})^4 \to 0$. Simple calculations show that

$$\sum_{j=1}^{n} E(\xi_j^{(n)})^4 = O((a_n^2 \lambda_n)/n).$$

Therefore

$$\sigma_n^{-1}(U_n^{(1)} - \Delta_n) \xrightarrow{d} N(0,1).$$

Theorem 2. *Let $K(x) \in H$, $f(x) \in F$, $W(x)$ be bounded, $W(-x) = W(x)$, $x \in R$, and $x^2 W(x) \in L_1(R)$. If $\lambda_n \to \infty$, $a_n \to \infty$, $a_n/\lambda_n \to 0$, $(\lambda_n a_n^2)/n \to 0$ and $\lambda_n a_n^{-5} \to 0$, then*

$$(\lambda_n a_n^{-1})^{1/2} \sigma^{-1}(f)(U_n^{(2)} - \Delta(f)) \xrightarrow{d} N(0,1), \quad U_n^{(2)} = \lambda_n^{-1} U_n^{(1)},$$
$$\Delta(f) = f(\ell_0) \int K^2(u)\, du \int W(x)\, dx, \quad \sigma^2(f) = 2f^2(\ell_0) \int K_0^2(z)\, dz \int W^2(v)\, dv, \quad f(\ell_0) \neq 0.$$

Proof. Lemma 2, Theorem 1 and the representation $\Delta_n(f)=\lambda_n[\Delta(f)+O(a_n^{-2})+O(\lambda_n^{-1})]$ provide the proof of the theorem.

Theorem 3. *Let $K(x)$, $f(x)$, $W(x)$ satisfy the conditions of Theorem 2. If $\lambda_n\to\infty$, $a_n\to\infty$, $a_n/\lambda_n\to 0$, $(\lambda_n a_n^2)/n\to 0$ and $\lambda_n a_n^{-5}\to 0$, $\sqrt{na_n}/\lambda_n^{5/2}\to 0$ and $na_n^{-1/2}\lambda_n^{-9/2}\to 0$, then*

$$(\lambda_n a_n^{-1})^{1/2}\sigma^{-1}(f)(U_n-\Delta(f))\xrightarrow{d}N(0,1).$$

Proof. We have

$$(\lambda_n a_n^{-1})^{1/2}(U_n-U_n^{(2)})=\sqrt{\frac{\lambda_n}{a_n}}\,(\Theta_n+R_n),$$
$$\Theta_n=\frac{n}{\lambda_n}\int(Ef_n(x)-f(x))^2W_n(x)\,dx,$$
$$R_n=2\,\frac{n}{\lambda_n}\int(f_n(x)-Ef_n(x))(Ef_n(x)-f(x))W_n(x)\,dx.$$

By virtue of the generalized Minkovskiǐ inequality and

$$\max_x|Ef_n(x)-f(x)|=O(\lambda_n^{-2}),$$

we obtain

$$(\lambda_n a_n^{-1})^{1/2}E|R_n|=O(\sqrt{na_n}\,\lambda_n^{-5/2})$$

and also

$$(\lambda_n a_n^{-1})^{1/2}\Theta_n=O(na_n^{-1/2}\lambda_n^{-9/2}).$$

The theorem is proved.

2. The assertion of Theorem 3 enables us to construct goodness-of-fit tests of the asymptotic level α for testing the hypothesis $H_0:\ f(x)=f_0(x),\ f_0(\ell_0)\neq 0$. For this it is necessary to reject H_0 if

$$U_n\geq d_n(\alpha)=\Delta(f_0)+\Big(\frac{\lambda_n}{a_n}\Big)^{-1/2}\varepsilon_\alpha\sigma(f_0),\tag{2}$$

where ε_α is the quantile of the level α of a standard normal distribution.

Theorem 4. *Let all the conditions of Theorem 3 be fulfilled. Then $\Pi_n(f_1)=P_{H_1}\{U_n\geq d_n(\alpha)\}\to 1$ as $n\to\infty$. Therefore the goodness-of-fit defined in (2) is consistent against any alternative $H_1:\ f(x)=f_1(x)$, $f_1(x)\neq f_0(x)$ on the set of a positive Lebesgue measure $f_1(\ell_0)\neq f_0(\ell_0)$.*

It is not difficult to show that

$$\Pi_n(f_1)=P_{H_1}\Big\{(\lambda_n a_n^{-1})^{-1/2}\sigma^{-1}(f_1)(U_n^*-\Delta(f_1))\geq-\frac{n}{\sqrt{\lambda a_n}}\,\big(\sigma^{-1}(f_1)R_n+o_p(1)\big)\Big\},$$
$$U_n^*=n\lambda_n^{-1}\|f_n-f_1\|^2_{L_2(W_n)}.$$

Since for the hypothesis H_1 we have

$$\sqrt{\frac{\lambda_n}{a_n}}\,\sigma^{-1}(f_1)(U_n^*-\Delta(f_1)) \xrightarrow{d} N(0,1), \quad n\lambda_n^{-1/2}a_n^{-1/2}\to\infty,$$
$$R_n \longrightarrow (f_1(\ell_0)-f_0(\ell_0))^2\int W(x)\,dx>0$$

we conclude that $\Pi_n(f_1)\to 1$.

Now let us introduce into the consideration the sequences of locally close alternatives ([7], [8])

$$H_{1n}:\ f_{1n}(x)=f_0(x)+\alpha_n\varphi\left(\frac{x-\ell_n}{\gamma_n}\right)+o(\alpha_n\gamma_n),$$
$$\ell_n=\ell_0+o(\gamma_n),\ \ \varphi(x)\in F,\ \ \int\varphi(x)\,dx=0.$$

Theorem 5. *Let $K(x)$, $f_{1n}(x)$, $W(x)$, λ_n and a_n satisfy the conditions of Theorem 3. Let, in addition, $W(x)$ be continuous at the point 0 and $W(0)>0$, $\alpha_n\gamma_n=o(n^{-1/2})$, $n\lambda_n^{-1/2}a_n^{1/2}\gamma_n\alpha_n^2\to\gamma_0>0$, $\lambda_na_n^{-1}\alpha_n^2\to 0$, $\lambda_n\gamma_n\to\infty$, $\alpha_n^{-1}\lambda_n^{-2}\to 0$ and $a_n\gamma_n\to 0$ as $n\to\infty$. Then*

$$P_{H_{1n}}\{U_n\ge d_n(\alpha)\}\longrightarrow 1-\Phi\bigg(\varepsilon_\alpha-\gamma_0W(0)\sigma^{-1}(f_0)\int\varphi^2(x)\,dx\bigg).$$

Proof. We have

$$P_{H_{1n}}\{U_n\ge d_n(\alpha)\}=P_{H_{1n}}\Bigg\{\sqrt{\frac{\lambda_n}{a_n}}\,(U_n^{(3)}-\Delta(f_{1n}))\sigma^{-1}(f_{1n})$$
$$\ge\frac{\sigma(f_0)}{\sigma(f_{1n})}\,\varepsilon_\alpha+\sqrt{\frac{\lambda_n}{a_n}}\,\sigma^{-1}(f_{1n})\,[\Delta(f_0)-\Delta(f_{1n})-A_{1n}+A_{2n}]\Bigg\},$$
$$U_n^{(3)}=n\lambda_n^{-1}\|f_n-f_{1n}\|^2_{L_2(W_n)},$$
$$A_{1n}=n\lambda_n^{-1}\|f_{1n}-f_0\|^2_{L_2(W_n)},\quad A_{2n}=n\lambda_n^{-1}\int(f_n(x)-f_{1n}(x))(f_{1n}(x)-f_0(x))W_n(x)\,dx.$$

From Theorem 3 it follows that

$$(\lambda_na_n^{-1})^{1/2}(U_n^{(3)}-\Delta(f_{1n}))\sigma^{-1}(f_{1n})\xrightarrow{d} N(0,1)$$

for the hypothesis H_{1n}. Let us now show that

$$\sqrt{\frac{\lambda_n}{a_n}}\,\sigma^{-1}(f_{1n})A_{2n}\xrightarrow{d} 0.$$

Indeed,

$$\sqrt{\frac{\lambda_n}{a_n}}\,E|A_{2n}|\le L_n^{(1)}+L_n^{(2)};$$

also $L_n^{(2)} = O(\alpha_n^{-1}\lambda_n^{-2})$ and

$$L_n^{(1)} \le cna_n^{1/2}\lambda_n^{-1/2}\alpha_n\bigg\{\frac{1}{n}\int f(u)\varphi^2\Big(\frac{u-\ell_n}{\gamma_n}\Big)\,du$$

$$+\gamma_n^{-2}n^{-1}\lambda_n^{-2}\int f(u)\,du\bigg[\iint\limits_0^1 |t|\,|K(t)|\,\bigg|\varphi^{(1)}\Big(\frac{u-\ell_n}{\gamma_n}\Big)+\frac{zt}{\lambda_n\gamma_n}\bigg|\,dt\,dz\bigg]^2\bigg\}^{1/2}.$$

Hence by virtue of the generalized Minkovskiĭ inequality we obtain

$$L_n^{(1)} = O(\lambda_n^{-1/4}a_n^{1/4}) + O\Big(\gamma_n^{-1}\lambda_n^{-1}\Big(\frac{a_n}{\lambda_n}\Big)^{1/4}\Big).$$

Therefore

$$\sqrt{\frac{\lambda_n}{a_n}}\,E|A_{2n}| = O\Big(\Big(\frac{a_n}{\lambda_n}\Big)^{1/4}\Big) + O(\alpha_n^{-1}\lambda_n^{-2}).$$

Furthermore, using the condition $n\lambda_n^{-1/2}a_n^{1/2}\gamma_n\alpha_n^2 \longrightarrow \gamma_0 > 0$ it is not difficult to establish that

$$\sigma_n^{-1}(f_{1n})\sqrt{\frac{\lambda_n}{a_n}}\,A_{1n} \longrightarrow \gamma_0 W(0)\sigma^{-1}(f_0)\int \varphi^2(u)\,du, \quad W(0)\neq 0.$$

The theorem is proved.

The conditions of the theorem as regards λ_n, a_n, α_n and γ_n are fulfilled if, for example, we assume that $\lambda_n = n^\delta$, $a_n = n^\varepsilon$, $\alpha_n = n^{-\alpha}$, $\gamma_n = n^{-\beta}$ for $\alpha = 9/35$, $\beta = 2/7$, $\delta = 2/5+\varepsilon$, $1/10 < \varepsilon < 1/5$; $\alpha = 11/30$, $\beta = 1/6$, $\delta = 1/5+\varepsilon$, $1/20 < \varepsilon < 1/6$ and so on.

It is well-known that for some α, β and δ, for which $\alpha+\beta > 1/2$, $1-2\alpha-\beta = \delta/2$, the limit power of the Rosenblatt–Bickel goodness-of-fit test ([2], [7], [8])

$$\begin{aligned} T_n &\ge \int f_0(x)W(x)\,dx\int K^2(u)\,du + \lambda_n^{-1/2}\varepsilon_\alpha\sigma_0, \\ T_n &= n\lambda_n^{-1}\int (f_n(x)-f_0(x))^2 w(x)\,dx, \\ \sigma_0^2 &= 2\int f_0^2(x)W^2(x)\,dx\int K_0^2(x)\,dx \end{aligned} \tag{3}$$

used for testing the hypothesis $H_0:\ f(x) = f_0(x)$ against the alternative

$$H_{1n}:\ f_{1n}(x) = f_0(x) + \alpha_n\varphi\Big(\frac{x-\ell_n}{\gamma_n}\Big), \quad \ell_n = \ell_0 + o(\gamma_n)$$

($\lambda_n = n^\delta$, $\alpha_n = n^{-\alpha}$ and $\gamma_n = n^{-\beta}$) is equal to

$$\gamma(T) = 1 - \Phi\Big(\varepsilon_\alpha - \frac{W(\ell_0)}{\sigma_0}\int\varphi^2(u)\,du\Big),$$

while the limit power $\gamma(u)$ of the goodness-of-fit (2) is equal to one for $a_n = n^{\varepsilon}$, $0 < \varepsilon < \delta$. Further, for some α, β, δ and ε, for which $\alpha+\beta > 1/2$, $1-2\alpha-\beta+\varepsilon/2 = \delta/2$, the limit power of the goodness-of-fit (2) is equal by virtue of Theorem 5 to

$$\gamma(u) = 1 - \Phi\left(\varepsilon_\alpha - \frac{W(0)}{\sigma(f_0)} \int \varphi^2(u)\, du\right),$$

while the limit power $\gamma(T)$ of the goodness-of-fit (3) is equal to $1 - \Phi(\varepsilon_\alpha)$. Moreover, the calculation of the right-hand side of (2) becomes essentially simpler as compared with (3) and therefore when choosing between the goodness-of-fit tests we will give preference to the goodness-of-fit test based on U_n.

REFERENCES

1. Bickel P. J., Rosenblatt M. On some global measures of the deviations of density function estimates. *Ann. Statist.*, **1** (1973), 1071–1095.
2. Nadaraya E. Limit distribution of the quadratic deviation of two nonparametric estimators of the density of a distribution. (Russian) *Soobshch. Akad. Nauk Gruz. SSR*, **78** (1975), 25-28.
3. Nadaraya E. A. Application of the central limit theorem for martingales to the investigation of the limit distribution of the square deviation of a kernel-type density estimator. (Russian) *Soobshch. Akad. Nauk Gruzin. SSR*, **113**, 2 (1984), 253-256.
4. Hall P. Central limit theorem for integrated square error of multivariate nonparametric density estimators. *J. Multivariate Anal.*, **14**, 1 (1984), 1-16.
5. Cai T. Tony, Low Mark G., Nonparametric estimation over shrinking neighborhoods: super efficiency and adaptation. *Ann. Statist.*, **33**, 1 (2005), 184-213.
6. Shiryaev A. N., Probability. (Russian) Nauka, Moscow, 1989.
7. Rosenblatt M. A quadratic measure of deviation of two-dimensional density estimates and a test of independence. *Ann. Statist.*, **3** (1975), 1-14.
8. Nadaraya E. A. Nonparametric estimation of probability densities and regression curves. *Kluwer Academic Publishers Group, Dordrecht*, 1989.

Received 27.07.2009; revised 21.09.2009; accepted 23.10.2009.

Authors' address:

E. Nadaraya, P. Babilua and G. Sokhadze
Iv. Javakhishvili Tbilisi State University
2, University St., Tbilisi 0186
Georgia
E-mail: giasokhil@i.ua