

Mathematics

On an Integral Square Deviation Measure with the Weight of “Delta-Functions” of the Rosenblatt-Parzen Probability Density Estimator

Elizbar Nadaraya

Academy Member, I. Javakhishvili Tbilisi State University

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. © 2008 Bull. Georg. Natl. Acad. Sci.

Key words: distribution density, goodness-of-fit test, power, consistency, limit distribution.

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

$$R(f_n, f; W_n) = E \int W_n(x) (f_n(x) - f(x))^2 dx, \quad (1)$$

is developed in the work of T. T. Cai and M. G. Low [5], 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 $a_n \rightarrow \infty$ for fixed n , then, roughly speaking,

$R(f_n, f; W_n) \approx E(f_n(\ell_0) - f(\ell_0))^2$. 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 $R(f_n, f; W_n) = E \|f_n - f\|_{L_2(W_n)}^2$. 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 $\|f_n - f\|_{L_2(W_n)}^2$, $W_n(x) = a_n W(a_n(x - \ell_0))$. 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 $a_n \rightarrow \infty$ as $n \rightarrow \infty$. The case $a_n \rightarrow a_0 < \infty$ is of no interest because it follows from the results of the works [1-4].

Let X_1, X_2, \dots, 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 $f(x)$

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

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

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

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

Notation.

$$\begin{aligned} U_n &= \frac{n}{\lambda_n} \|f_n - f\|_{L_2(W_n)}^2, \quad U_n^{(1)} = n \|f_n - Ef\|_{L_2(W_n)}^2, \quad \Delta_n = EU_n^{(1)}, \\ \alpha_n(x, y) &= \lambda_n [K(\lambda_n(x - y)) - EK(\lambda_n(x - X_1))], \\ \sigma_n^2 &= 2 \iint [E\alpha_n(u_1, X_1)\alpha_n(u_2, X_1)]^2 W_n(u_1)W_n(u_2) du_1 du_2, \\ \eta_{ij}^{(n)} &= 2n^{-1} \sigma_n^{-1} \int \alpha_n(x, X_i)\alpha_n(x, X_j)W_n(x)dx, \\ \xi_j^{(n)} &= \sum_{i=1}^{j-1} \eta_{ij}^{(n)}, \quad j = \overline{2, n}, \quad \xi_1^{(n)} = 0, \quad \xi_j^{(n)} = 0, \quad j > n, \end{aligned}$$

$$\mathcal{F}_k^{(n)} = \sigma(\omega : X_1, X_2, \dots, X_k).$$

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 $\lambda_n \rightarrow \infty$, $a_n \rightarrow \infty$ and $\frac{a_n}{\lambda_n} \rightarrow 0$ as $n \rightarrow \infty$, then

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

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

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

and also

$$\text{var } H_n^{(2)} = O\left(\frac{\lambda_n a_n}{n}\right) + O(n^{-1} \sigma_n^{-2}), \quad \text{i.e. } H_n^{(2)} \xrightarrow{P} 0.$$

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

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

and

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

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

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

It is not difficult to show that

$$\text{Var } V_{n1} = \frac{16 \lambda_n^8}{n^4 \sigma_n^4} \left[\sum_{j=2}^n (j-1) E(\varepsilon_1 - \bar{\sigma}_n)^2 + 2 \sum_{i=2}^n E Z_i^2 (n-i) \right] = B_{n1} + B_{n2},$$

and also

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

Therefore $\text{Var}V_{n1} \rightarrow 0$. On the other hand, $EV_{n1} = 1 - \frac{1}{n} \rightarrow 1$. Therefore $V_{n1} \xrightarrow{P} 1$.

Now let us consider V_{n2} . Taking into account the inequality

$$E\left(\sum_{i=1}^n Y_j\right)^2 \leq \left(\sum_{i=1}^n (EY_i^2)^{1/2}\right)^2$$

which is easy to verify and performing some simple calculations, we obtain

$$EV_{n2}^2 = O\left(\frac{a_n}{\lambda_n}\right).$$

Therefore $V_{n2}^2 \xrightarrow{P} 1$.

Now we will establish the validity of the Lindeberg condition. For this, it suffices to make sure that

$\sum_{j=1}^n E(\xi_j^{(n)})^4 \rightarrow \infty$. Simple calculations show that

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

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

$\lambda_n \rightarrow \infty$, $a_n \rightarrow \infty$, $\frac{a_n}{\lambda_n} \rightarrow 0$, $\frac{\lambda_n a_n^2}{n} \rightarrow 0$ and $\lambda_n a_n^{-5} \rightarrow 0$, then

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

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 $\lambda_n \rightarrow \infty$, $a_n \rightarrow \infty$, $\frac{a_n}{\lambda_n} \rightarrow 0$,

$\frac{\lambda_n a_n^2}{n} \rightarrow 0$, $\lambda_n a_n^{-5} \rightarrow 0$, $\frac{\sqrt{na_n}}{\lambda_n^{5/2}} \rightarrow 0$ and $na_n^{-1/2} \lambda_n^{-9/2} \rightarrow 0$, then

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

Proof. We have

$$(\lambda_n a_n^{-1})^{1/2} (U_n - U_n^{(2)}) = \sqrt{\frac{\lambda_n}{a_n}} (\Theta_n + R_n),$$

$$\Theta_n = \frac{n}{\lambda_n} \int (Ef_n(x) - f(x))^2 W_n(x) dx,$$

$$R_n = 2 \frac{n}{\lambda_n} \int (f_n(x) - Ef_n(x))(Ef_n(x) - f(x)) W_n(x) dx.$$

By virtue of the generalized Minkovski 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

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

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 $\Pi_n(f_1) = P_{H_1}\{U_n \geq d_n(\alpha)\} \rightarrow 1$ as $n \rightarrow \infty$.

Therefore the goodness-of-fit defined in (*) 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

$$\Pi_n(f_1) = P_{H_1} \left\{ (\lambda_n a_n^{-1})^{1/2} \sigma^{-1}(f_1) (U_n^* - \Delta(f_1)) \geq -\frac{n}{\sqrt{\lambda a_n}} (\sigma^{-1}(f_1) R_n + o_p(1)) \right\},$$

$$U_n^* = n\lambda_n^{-1} \|f_n - f_1\|_{L_2(W_n)}^2.$$

Since for the hypothesis H_1 we have

$$\sqrt{\frac{\lambda_n}{a_n}} \sigma^{-1}(f_1) (U_n^* - \Delta(f_1)) \xrightarrow{d} N(0,1), \quad n\lambda_n^{-1/2} a_n^{-1/2} \rightarrow \infty$$

and

$$R_n \rightarrow (f_1(\ell_0) - f_0(\ell_0))^2 \int W(x) dx > 0$$

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

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

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

$$\ell_n = \ell_0 + o(\gamma_n), \quad \varphi(x) \in F, \quad \int \varphi(x) dx = 0.$$

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

$$P_{H_{1n}}\{U_n \geq d_n(\alpha)\} \rightarrow 1 - \Phi\left(\varepsilon_\alpha - \gamma_0 W(0) \sigma^{-1}(f_0) \int \varphi^2(x) dx\right).$$

Proof. We have

$$P_{H_{1n}}\{U_n \geq d_n(\alpha)\} = P_{H_{1n}} \left\{ \sqrt{\frac{\lambda_n}{a_n}} (U_n^{(3)} - \Delta(f_{1n})) \sigma^{-1}(f_{1n}) \geq \frac{\sigma(f_0)}{\sigma(f_{1n})} \varepsilon_\alpha + \sqrt{\frac{\lambda_n}{a_n}} \sigma^{-1}(f_{1n}) [\Delta(f_0) - \Delta(f_{1n}) - A_{1n} + A_{2n}] \right\},$$

$$U_n^{(3)} = n\lambda_n^{-1} \|f_n - f_{1n}\|_{L_2(W_n)}^2,$$

$$A_{1n} = n\lambda_n^{-1} \|f_{1n} - f_0\|_{L_2(W_n)}^2, \quad A_{2n} = n\lambda_n^{-1} \int (f_n(x) - f_{1n}(x))(f_{1n}(x) - f_0(x)) W_n(x) dx.$$

From Theorem 3 it follows that

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

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

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

Indeed,

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

also

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

and

$$L_n^{(1)} \leq c n a_n^{1/2} \lambda_n^{-1/2} \alpha_n \times \left\{ \frac{1}{n} \int f(u) \varphi^2 \left(\frac{u - \ell_n}{\gamma_n} \right) du + \gamma_n^{-2} n^{-1} \lambda_n^{-2} \int f(u) du \left[\int_0^1 |t| |K(t)| \left| \varphi^{(1)} \left(\frac{u - \ell_n}{\gamma_n} \right) + \frac{zt}{\lambda_n \gamma_n} \right| dt dz \right]^2 \right\}^{1/2}.$$

Hence by virtue of the generalized Minkovski inequality we obtain

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

Therefore

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

Furthermore, using the condition

$$n \lambda_n^{-1/2} a_n^{1/2} \gamma_n \alpha_n^2 \rightarrow \gamma_0 > 0$$

it is not difficult to establish that

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

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

$$\lambda_n = n^\delta, \quad a_n = n^\varepsilon, \quad \alpha_n = n^{-\alpha}, \quad \gamma_n = n^{-\beta} \quad \text{for} \quad \alpha = \frac{9}{35}, \quad \beta = \frac{2}{7}, \quad \delta = \frac{2}{5} + \varepsilon, \quad \frac{1}{10} < \varepsilon < \frac{1}{5}; \quad \alpha = \frac{11}{30}, \quad \beta = \frac{1}{6},$$

$$\delta = \frac{1}{5} + \varepsilon, \quad \frac{1}{20} < \varepsilon < \frac{1}{6} \quad \text{and so on.}$$

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 ([2], [7], [8])

$$T_n \geq \int f_0(x) w(x) dx \int K^2(u) du + \lambda_n^{-1/2} \varepsilon_\alpha \sigma_0,$$

$$T_n = n \lambda_n^{-1} \int (f_n(x) - f_0(x))^2 w(x) dx, \quad (**)$$

$$\sigma_0^2 = 2 \int f_0^2(x) w^2(x) dx \int K_0^2(x) dx$$

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

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

$$(\lambda_n = n^\delta, \quad \alpha_n = n^{-\alpha} \quad \text{and} \quad \gamma_n = n^{-\beta})$$

is equal to

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

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

Acknowledgement. This work was supported by the GNSF Grant # GNSF/ST06/3-039.

მათემატიკა

განაწილების სიმკვრივის როზენბლატ-პარზენის შეფასების „დელტა ფუნქციის“ წონით ინტეგრალური კვადრატული გადახრის ზღვართი განაწილების შესახებ

ე. ნადარაია

აკადემიის წევრი, ი. ჯავახიშვილის სახ. თბილისის სახელმწიფო უნივერსიტეტი

მიღებულია განაწილების სიმკვრივის როზენბლატ-პარზენის შეფასების „დელტა ფუნქციის“ წონით ინტეგრალური კვადრატული გადახრის ზღვართი განაწილების კანონი. აგებულია თანხმობის კრიტერიუმი და მოძებნილია მისი ასიმპტოტური სიმძლავრე გარკვეული ტიპის დაახლოებადი ალტერნატივებისთვის. ნაჩვენებია აგრეთვე, რომ აგებული თანხმობის კრიტერიუმი უფრო მძლავრია, ვიდრე როზენბლატ-ბიკელის თანხმობის კრიტერიუმი ხსენებული ტიპის დაახლოებადი ალტერნატივებისათვის.

REFERENCES

1. P. Bickel, M. Rosenblatt (1973), Ann. Statist., 1: 1071-1095.
2. Э. Надарая (1975), Сообщения АН ГССР, 78, 1: 25-28.
3. Э. Надарая (1984), Сообщения АН ГССР, 113, 2: 253-256.
4. P. Hall (1984), Journal of Multivariate Analysis, 14: 1-16.
5. T.T. Cai, G.L. Mark (2005), Ann. Statist., 33: 184-213.
6. А.Н. Ширяев (1989), Вероятность, Москва.
7. M. Rosenblatt (1975), Ann. Math. Statist., 3: 1-14.
8. E. Nadaraya (1989), Nonparametric estimation of probability densities and regression curves. Kluwer Academic Publishers.

Received November, 2008