Reports of Enlarged Session of the
Seminar of I. Vekua Institute
of Applied Mathematics
Volume 29, 2015

ON FUNCTIONALS OF GASSER-MÜLLER ESTIMATORS

Babilua P., Nadaraya E., Sokhadze G.

Abstract. The asymptotic properties of a general functional of the Gasser–Müller estimator are investigated in the Sobolev space. The convergence rate, consistency and the central limit theorem are established.

Keywords and phrases: Nonparametric regression, Gasser–Müller estimator, functionals.

AMS subject classification: 62G08, 62H12.

Let us consider a regression model of the form

$$Y(t)=a(t)+\varepsilon(t), \tag{1}$$

where $t\in[0,1]$, $\varepsilon(\,\cdot\,)$ is noise with $E\varepsilon(t)=0$, $E\varepsilon^2(t)=\sigma^2<\infty$, $Y(t)$ is a random function and $a(t)$ is an unknown function. Suppose we have n numbers

$$0\le t_1\le t_2\le\cdots\le t_n\le 1,$$

where each $t_k, k=1,2,\dots,n$, is dependent on n.

The estimator of an unknown regression function $a(t)$ was introduced by Gasser and Müller and defined by the expression

$$\widehat{a}_n(t)=\frac{1}{h_n}\sum_{i=1}^{n}\int\limits_{s_{i-1}}^{s_i}W\Big(\frac{t-u}{h_n}\Big)\,du\cdot Y(t_i), \tag{2}$$

where $0=s_0\le s_1\le s_2\le\cdots\le s_n=1$, $t_i\le s_i\le t_{i+1}$, $i=1,2,\dots,n-1$ and

$$\max_i|s_i-s_{i-1}|=O\Big(\frac{1}{n}\Big);$$

$\{h_n, n=1,2,\dots\}$ is the sequence of positive numbers which monotonically tend to zero, $W(u)$ is the function with probability density properties.

Gasser and Müller defined also the estimator of the k-th derivative of the regression function $a^{(k)}(t)$ by the formula

$$\widehat{a}_n^{(k)}(t)=\frac{1}{h_n^{k+1}}\sum_{i=1}^{n}\int\limits_{s_{i-1}}^{s_i}W^{(k)}\Big(\frac{t-u}{h_n}\Big)\,du\cdot Y(t_i) \tag{3}$$

for all $k=0,1,\dots,m$. It was assumed that $\widehat{a}_n^{(0)}(t)\doteqdot\widehat{a}_n(t)$.

In the above-mentioned works, the consistency and asymptotic normality theorems for these estimators were obtained by imposing certain conditions.

For some functional $\mathfrak{A}$, here we investigate the asymptotic properties of the expression $\mathfrak{A}(\widehat{\mathfrak{a}}_{\mathfrak{n}})$ as $n\to\infty$.

Let us introduce the notation and conditions which will be used in our argumentation.

Conditions on a***:***

(a1) The function $a = a(t)$ is well defined and continuous on $[0, 1]$ and takes its values in the interval $[-\Bbbk; \Bbbk]$;

(a2) The function $a(t)$ has continuous derivatives up to order m inclusive;

(a3) For any $i = 0, 1, \dots, m$, $a^{(i)}(t)$ takes its values in $[-\Bbbk; \Bbbk]$ and $a^{(i)}(\,\cdot\,) \in L_1([0, 1])$.

Conditions on ε_k***:***

(ε1) Random values $\varepsilon_k = \varepsilon(t_k)$, $k = 1, 2, \dots$, are independent and equally distributed;

(ε2) $E\varepsilon_k = 0$, $E\varepsilon_k^2 = \sigma^2 < \infty$.

Conditions on W***:***

(w1) $\int\limits_{-\infty}^{\infty} W(t)\, dt = 1$;

(w2) Functions $W^{(i)}(t)$, $i = 0, 1, \dots, m$ have the compact support $[-\tau, \tau]$,

$$W^{(i)}(-\tau) = W^{(i)}(\tau) = 0;$$

(w3) The function $W(t)$ has continuous derivatives up to order m, $m \geq 1$;

(w4) There exists a constant $C_W > 0$, for which

$$\sup_{t \in R} |W^{(i)}(t)| \leq C_W < \infty, \quad i = 0, 1, \dots, m;$$

(w5) For any $i = 0, 1, \dots, m$, $W^{(i)} \in L_1([-\tau, \tau])$.

Denote by $a_n(t)$ the mathematical expectation $\widehat{a}_n(t)$:

$$\begin{aligned} a_n(t) = E\widehat{a}_n(t) &= E\,\frac{1}{h_n} \sum_{i=1}^{n} \int\limits_{s_{i-1}}^{s_i} W\Big(\frac{t-u}{h_n}\Big)\, du \cdot Y(t_i) \\ &= \frac{1}{h_n} \sum_{i=1}^{n} \int\limits_{s_{i-1}}^{s_i} W\Big(\frac{t-u}{h_n}\Big)\, du \cdot a(t_i). \end{aligned}$$

Then we obtain

$$a_n^{(k)}(t) = E\widehat{a}_n^{(k)}(t) = \frac{1}{h_n^{i+1}} \sum_{i=1}^{n} \int\limits_{s_{i-1}}^{s_i} W^{(k)}\Big(\frac{t-u}{h_n}\Big)\, du \cdot a(t_i).$$

Let $C^m[0,1]$ denote the space of bounded real functions which are defined and continuous on $[0,1]$, have continuous derivatives of at least m-th order. In this space we introduce the norm

$$\|f\|_m = \left(\sum_{k=0}^{m} \int_0^1 \left(\frac{d^k f}{dt^k} \right)^2 dt \right)^{\frac{1}{2}}, \quad f \in C^m[0,1].$$

The closure of $C^m[0,1]$ in this norm is denoted by W_m^2 and called the Sobolev space. This is a complete separable Hilbert space with the scalar product

$$\langle f, g \rangle_m = \sum_{k=0}^{m} \int_0^1 \frac{d^k f}{dt^k} \frac{d^k g}{dt^k} \, dt, \quad f, g \in W_m^2.$$

Conditions on $\mathfrak{A}$:

($\mathfrak{A}1$) The functional $\mathfrak{A} : \mathfrak{W}_{\mathfrak{m}}^2 \to \mathfrak{R}$ is considered in the space W_m^2. It is assumed that this functional is smooth in a strong sense. This means that there exists a bounded linear functional $T_{\mathfrak{A}}$ such that for any two elements from W_m^2, $f, g \in W_m^2$, we have

$$\mathfrak{A}(\mathfrak{f}) - \mathfrak{A}(\mathfrak{g}) = \mathfrak{T}_{\mathfrak{A}}(\mathfrak{f} - \mathfrak{g}) + \mathfrak{O}(\|\mathfrak{f} - \mathfrak{g}\|_{\mathfrak{m}}^2).$$

By the Riesz theorem there is an element $t_{\mathfrak{A}}$ of the space W_m^2, such that

$$T_{\mathfrak{A}} w = \langle t_{\mathfrak{A}}, w \rangle_m.$$

The formulation of our problem reads as follows: Consider the Gasser–Müller scheme, where the components of (1), (2) and (3) satisfy conditions (a1)–(a3), (ε1)–(ε3), (w1)–(w5) and ($\mathfrak{A}1$). Construct the estimator of the variable $\mathfrak{A}(\mathfrak{a})$ using observations $\{(t_1, Y(t_1)), \ldots, (t_n, Y(t_n))\}$.

Theorem 1. *Assume that the conditions* (a1)–(a3), (ε1)–(ε3), (w1)–(w5) *and* ($\mathfrak{A}1$) *are fulfilled. Then a representation formula holds with the remainder of order*

$$R_n = O\Big(\frac{\log n}{n h_n^{2m+2}} \Big).$$

In this section of the paper we use Theorem 1 to prove the strict consistency of the estimator $\mathfrak{A}(\widehat{\mathfrak{a}}_{\mathfrak{n}})$.

Theorem 2. *Let the conditions of Theorem 1 be fulfilled. Then, if the positive sequence h_n, $n = 1, 2, \ldots$, $0 < h_n < 1$, is chosen so that*

$$\frac{\log n}{n h_n^{2m+2}} \longrightarrow 0. \tag{4}$$

Then with probability 1, *we have*

$$\mathfrak{A}(\widehat{\mathfrak{a}}_{\mathfrak{n}}) \longrightarrow \mathfrak{A}(\mathfrak{a})$$

as $n \to \infty$.

REFERENCES

1. Arabidze D., Babilua P., Nadaraya E., Sokhadze G., Tkeshelashvili A. Integral functionals of the Gasser-Müller Regression Function. *Ukrainian Math. J.* **67**, 4 (2015), 435-446.

2. Babilua P.K., Nadaraya E.A., Patsatsia M.B., Sokhadze G.A. On the integral functionals of a kernel estimator of a distribution density. *Proc. I. Vekua Inst. Appl. Math.*, **58** (2008), 6-14, 110.

3. Goldstein L., Messer K. Optimal plug-in estimators for nonparametric functional estimation. *Ann. Statist.*, **20**, 3 (1992), 1306-1328.

4. Gasser T., Müller H.G. Estimating regression functions and their derivatives by the kernel method. *Scand. J. Statist.*, **11**, 3 (1984), 171-185.

Received 13.05.2015; revised 12.10.2015; accepted 27.11.2015.

Authors' addresses:

P. Babilua, E. Nadaraya
Iv. Javakhishvili Tbilisi State University
Faculty of Exact and Natural Sciences
Department of Mathematics
2, University St., Tbilisi 0186
Georgia
E-mail: petre.babilua@tsu.ge, elizabar.nadaraya@tsu.ge

G. Sokhadze
I. Vekua Institute of Applied Mathematics of
Iv. Javakhishvili Tbilisi State University
2, University St., Tbilisi 0186
Georgia
E-mail: grigol.sokhadze@tsu.ge