Transactions of A. Razmadze
Mathematical Institute
Vol. **176** (2022), issue 3, 429–433

ON THE NADARAYA–WATSON TYPE NONPARAMETRIC ESTIMATE OF POISSON REGRESSION FUNCTION

PETRE BABILUA AND ELIZBAR NADARAYA

Abstract. The Nadaraya–Watson kernel-type nonparametric estimate of Poisson regression function is studied. The uniform consistency conditions are established and the limit theorems are proved for continuous functionals on $C[a,1-a]$, $0<a<1/2$.

Let a random variable Y take values $0,1,2,\dots$ with probabilities $\Pi(k,\lambda)=\mathbf{P}\{Y=k\}=\frac{\lambda^k}{k!}\,e^{-\lambda}$, $\lambda>0$, $k=0,1,\dots$. Assume that the parameter λ is the function of an independent variable $x\in[0,1]$. $\lambda(x)$ is known as a Poisson regression function [2,6]. Let x_i, $i=1,\dots,n$ be the division points of the interval $[0,1]$:

$$x_i=\frac{2i-1}{2n}\,,\ \ i=1,2,\dots,n.$$

Let, further, Y_i, $i=1,2,\dots,n$ be independent Poisson random variables with $\mathbf{P}\{Y_i=k\,|\,x_i\}=\Pi(k,\lambda(x_i))$. The problem consists in estimating the function $\lambda(x)$, $x\in[0,1]$, by the samples $Y_1,Y_2,\dots,Y_n$ [2]. Problems of this kind arise, for example, in medicine [5,10], in astrophysics [7], and so on.

As an estimator for $\lambda(x)$, we consider the following statistics [8,11] which is known as Nadaraya–Watson estimate

$$\widehat{\lambda}_n(x)=\lambda_{1n}(x)\lambda_{2n}^{-1}(x),$$
$$\lambda_{\nu n}(x)=\frac{1}{nb_n}\sum_{i=1}^{n}K\Big(\frac{x-x_i}{b_n}\Big)Y_i^{2-\nu},\ \ \nu=1,2,$$

where $K(x)$ is some distribution density (kernel) and $K(x)=K(-x)$, $x\in(-\infty,+\infty)$, and $\{b_n\}$ is a sequence of positive numbers converging to zero.

The aim of the present paper is to establish uniform convergence of the estimate $\widehat{\lambda}_n(x)$ to the $\lambda(x)$ by probability and also to state the limit theorems for continuous functionals connected with this function on $C[a,1-a]$, $0<a<\frac{1}{2}$.

For obtaining these results, we need the following lemmas given in [1].

Lemma 1. *Assume that:*

1^0. $K(x)$ *is some function with a bounded variation;*
2^0. $\lambda(x)$ *is also a function with a bounded variation on* $[0,1]$.

If $nb_n\to\infty$, *then*

$$\frac{1}{nb_n}\sum_{i=1}^{n}K^{\nu_1}\Big(\frac{x-x_i}{b_n}\Big)p^{\nu_2}(x_i)=\frac{1}{b_n}\int_0^1K^{\nu_1}\Big(\frac{x-u}{b_n}\Big)p^{\nu_2}(u)\,du+O\Big(\frac{1}{nb_n}\Big),\tag{1}$$

uniformly in $x\in[0,1]$; $\nu_i\in N\cup\{0\}$, $i=1,2$.

2020 *Mathematics Subject Classification.* Primary 60F05; Secondary 62G05.

Key words and phrases. Poisson regression function; Consistency; Uniform convergence; Wiener process; Nadaraya–Watson estimate.

Let us introduce for the function $K(x)$ the Fourier transform:

$$\psi(t)=\int\limits_{-\infty}^{\infty} e^{itx}K(x)\,dx$$

and assume that

3^0. $\psi(x)$ is absolutely integrable. Then we can write $\lambda_{1n}(x)-\mathbf{E}\,\lambda_{1n}(x)$ in the form

$$\lambda_{1n}(x)-\mathbf{E}\,\lambda_{1n}(x)=\frac{1}{2\pi}\int\limits_{-\infty}^{\infty} e^{-iux/b_n}\psi(u)\frac{1}{nb_n}\sum_{j=1}^{n}[Y_j-\lambda(x_j)]e^{iux_j/b_n}\,du.$$

Denote

$$d_n=\sup_{x\in\Omega_n}|\widehat{\lambda}_n(x)-\mathbf{E}\,\widehat{p}_n(x)|,\quad \Omega_n=[b_n^\alpha,1-b_n^\alpha],\ \ 0<\alpha<1.$$

Theorem 1. *Let $K(x)$ satisfy conditions 1^0 and 3^0, and let $\lambda(x)$ be continuous and satisfy condition 2^0.*

(a) *Let $nb_n^2\to\infty$, then*

$$D_n=\sup_{x\in\Omega_n}|\widehat{\lambda}_n(x)-\lambda(x)|\xrightarrow{P}0.$$

(b) *If $\sum\limits_{n=1}^{\infty} n^{-s/2}b_n^{-s}<\infty$, $s>2$, then $D_n\to 0$ a.s.*

Corollary 1. *Under the conditions of Theorem 1,*

$$\sup_{x\in[a,b]}|\widehat{\lambda}_n(x)-\lambda(x)|\to 0$$

in probability (almost surely) for any fixed interval $[a,b]\subset[0,1]$.

Assume that $b_n=n^{-\gamma}$, $\gamma>0$. The following conditions of Theorem 1:

$$n^{1/2}b_n\to\infty,\ \text{ if }\ 0<\gamma<\frac{1}{2}$$

and

$$\sum_{n=1}^{\infty} n^{-s/2}b_n^{-s}<\infty,\ \text{ if }\ 0<\gamma<\frac{s-2}{2s},\ \ s>2$$

are fulfilled.

Before we proceed to proving Theorem 2, let us consider two lemmas below.

Lemma 2. *Let the kernel $K(x)\ge 0$ be chosen so that it would be a function of a finite variation and satisfy the conditions $K(x)=K(-x)$, $K(x)=0$ for $|x|\ge 1$, $\int K(u)\,du=1$. Let $g(x)\ge 0$, $x\in[a,1-a]$, $0<a<1/2$, be any measurable bounded function. Let, further, $0<\inf\lambda(x)$, $x\in[0,1]$.*

(a) *If $\lambda(x)$ is continuous and with a bounded variation on $[0,1]$ and $nb_n^2\to\infty$ as $n\to\infty$, then*

$$\overline{\xi}_n=\sqrt{n}\int\limits_a^{1-a} g_1(x)\big[\,\widehat{\lambda}_n(x)-\mathbf{E}\,\widehat{\lambda}_n(x)\big]\,dx\xrightarrow{d}N(0,\sigma^2),\quad g_1(x)=g(x)\varphi(x),\ \ \varphi(x)=\lambda^{-1/2}(x). \tag{2}$$

(b) *If $nb_n^2\to\infty$, $nb_n^4\to 0$, and $\lambda(x)$ has bounded derivatives up to the second order, then*

$$\xi_n=\sqrt{n}\int\limits_a^{1-a} g_1(x)[\widehat{\lambda}_n(x)-\lambda(x)]\,dx\xrightarrow{d}N(0,\sigma^2),\quad \sigma^2=\int\limits_a^{1-a} g^2(x)\,dx,$$

for $n\to\infty$.

Remark 1. We have introduced $a > 0$ in the theorem in order avoid of the boundary effect of the estimate $\widehat{\lambda}_n(x)$. More exactly, near the boundary of the interval $[0,1]$, the estimate $\widehat{\lambda}_n(x)$, being a kernel-type estimate, behaves worse in the sense of order of convergence to zero of the bias $\mathbf{E}\,\widehat{\lambda}_n(x) - \lambda(x)$ than in any interval $[a, 1-a] \subset [0,1]$, $0 < a < 1/2$ [4,9].

Lemma 3. *Under the conditions* (a) *and* (b) *of Lemma* 2*, we, respectively, have*

$$\mathbf{E}\,|\overline{\xi}_n|^s \le c_{13}\Bigg(\int\limits_a^{1-a} g(u)\,du\Bigg)^{s/2}, \quad s > 2, \tag{3}$$

and

$$\mathbf{E}\,|\xi_n|^s \le c_{14}\Bigg(\int\limits_a^{1-a} g(u)\,du\Bigg)^{s/2}, \quad s > 2. \tag{4}$$

Let us introduce the following random processes:

$$\overline{\xi}_n(t) = \sqrt{n}\int\limits_a^t \big(\widehat{\lambda}_n(u) - \mathbf{E}\,\widehat{\lambda}_n(u)\big)\psi(u)\,du, \quad \xi_n(t) = \sqrt{n}\int\limits_a^t \big(\widehat{\lambda}_n(u) - \lambda(u)\big)\psi(u)\,du.$$

Theorem 2. *Let all conditions of Lemma* 2 *be fulfilled. Then for all continuous functionals* $f(\,\cdot\,)$ *on* $C[a, 1-a]$, $0 < a < 1/2$ *the distributions of* $f(\overline{\xi}_n(t))$ *and* $f(\xi_n(t))$ *converge to the distribution of* $f(w(t-a))$, *where* $w(t-a)$, $a \le t \le 1-a$, *is a Wiener process with a correlation function* $r(s,t) = \min(t-a, s-a)$, $w(t-a) = 0$, $t = a$.

Proof. First, we show that the finite-dimensional distributions of the processes $\overline{\xi}_n(t)$ converge to the finite-dimensional distributions of the process, $w(t-a)$, $t \ge a$.

Let us consider one moment of time t_1; we have to show that

$$\overline{\xi}_n(t_1) \xrightarrow{d} w(t_1 - a). \tag{5}$$

To prove (5), it suffices to take $g(x) = I_{[a,t_1]}(x)$ in (2). Then by virtue of Lemma 2,

$$\overline{\xi}_n(t_1) \xrightarrow{d} N(0, t_1 - a).$$

Consider now two moments of time t_1, t_2, $t_1 < t_2$. Towards this end, we have to show that

$$\big(\overline{\xi}_n(t_1), \overline{\xi}_n(t_2)\big) \xrightarrow{d} \big(w(t_1 - a), w(t_2 - a)\big). \tag{6}$$

To prove (6), it suffices to take

$$g(x) = (\lambda_1 + \lambda_2)I_{[a,t_1)}(x) + \lambda_2 I_{[t_1,t_2)}(x)$$

in (2), where λ_1, λ_2 are arbitrary finite numbers. Then by virtue of Lemma 2,

$$\lambda_1\overline{\xi}_n(t_1) + \lambda_2\overline{\xi}_n(t_2) \xrightarrow{d} N\Big(0, (\lambda_1+\lambda_2)^2(t_1 - a) + \lambda_2^2(t_2 - t_1)\Big).$$

On the other hand,

$$\lambda_1 w(t_1 - a) + \lambda_2 w(t_2 - a) = (\lambda_1 + \lambda_2)\big[w(t_1 - a) - w(0)\big] + \lambda_2\big[w(t_2 - a) - w(t_1 - a)\big]$$

is distributed as $N(0, (\lambda_1 + \lambda_2)^2(t_1 - a) + \lambda_2^2(t_2 - t_1))$.

Therefore (6) holds. The case of three and more moments of time is considered analogously. Now, let us show that the sequence $\{\overline{\xi}_n(t)\}$ is dense, i.e., the sequence of respective distributions is dense. To this end, it suffices to show that for any t_1, $t_2 \in [a, 1-a]$ and all n,

$$\mathbf{E}\,\big|\overline{\xi}_n(t_1) - \overline{\xi}_n(t_2)\big|^s \le c_{19}|t_1 - t_2|^{s/2}, \;\; s > 2.$$

Indeed, this inequality is obtained from (3) for $g(x) = I_{[t_1,t_2]}(x)$.

Further, taking into account (4) and statements (b) of Lemma 2, we easily conclude that the finite-dimensional distributions of the processes $\xi_n(t)$ converge to the finite-dimensional distributions of the Wiener process $w(t-a)$, and also

$$\mathbf{E}\,|\xi_n(t_1)-\xi_n(t_2)|^s \le c_{20}|t_1-t_2|^{s/2}, \ \ s>2.$$

Thus the proof of the theorem follows from Theorem 2 of the monograph [9, p. 583]. □

Corollary 2. *By virtue of Theorem 2 and Theorem 1 from* [3, p. 371]*, we can write*

$$p\Big\{\max_{a\le t\le 1-a}\xi_n(t)>\lambda\Big\} \longrightarrow \frac{2}{\sqrt{2\pi(1-2a)}}\int_{\lambda}^{\infty}\exp\Big\{-\frac{x^2}{2(1-2a)}\,dx\Big\}, \ \ 0<a<\frac{1}{2}, \ \ \textit{as}\ \ n\to\infty.$$

This result makes it possible to construct the goodness-of fit test of the level α for testing hypothesis H_0, according to which

$$H_0:\ \lim_{n\to\infty}\mathbf{E}\,\widehat{\lambda}_n(x)=\lambda_0(x), \ \ a\le x\le 1-a,$$

when the alternative hypothesis is

$$H_1:\ \int_a^{1-a}\psi_0(x)\Big(\lim_{n\to\infty}\mathbf{E}\,\widehat{\lambda}_n(x)-\lambda_0(x)\Big)\,dx>0, \quad \psi_0(x)=\lambda_0^{-1/2}(x).$$

Further, we note that the functionals

$$f_1(x(\,\cdot\,))=\sup_{a\le t\le 1-a}|x(t)|, \quad f_2(x(\,\cdot\,))=\int_a^{1-a}x^2(t)\,dt$$

are continuous on $C[a,1-a]$. Therefore Theorem 2 also implies

$$f_1(\xi_n(\,\cdot\,))\xrightarrow{d} f_1(W(\,\cdot\,))$$

and

$$f_2(\xi_n(\,\cdot\,))\xrightarrow{d} f_2(W(\,\cdot\,)).$$

Remark 2. Let t_i be the division points of the interval $[0,1]$ chosen so that

$$H(t_j)=\frac{2j-1}{2n}\,, \ \ j=1,\dots,n,$$

where $H(x)=\int\limits_0^x h(u)\,du$, $h(u)$ is some known continuous distribution density on $[0,1]$. Then, arguing analogously to the above, one can obtain a generalization of the results of this paper.

References

1. P. K. Babilua, E. A. Nadaraya, G. A. Sokhadze, On the square-integrable measure of the divergence of two nuclear estimations of the Bernoulli regression functions. Translation of *Ukraïn. Mat. Zh.* **67** (2015), no. 1, 3–18; *Ukrainian Math. J.* **67** (2015), no. 1, 1–18.
2. S. Efromovich, *Nonparametric Curve Estimation.* Methods, theory, and applications. Springer Series in Statistics. Springer-Verlag, New York, 1999.
3. I. I. Gikhman, A. V. Skorohod, *Introduction to the Theory of Random Processes.* (Russian) Izdat. Nauka, Moscow, 1965.
4. J. D. Hart, T. E. Wehrly, Kernel regression when the boundary region is large, with an application to testing the adequacy of polynomial models. *J. Amer. Statist. Assoc.* **87** (1992), no. 420, 1018–1024.
5. C.-H. Hwang, J.-Y. Shim, Semiparametric kernel Poisson regression for longitudinal count data. *Commun. Stat. Appl. Methods* **15** (2008), no. 6, 1003–1011.
6. M. Kohler, A. Krzyżak, Asymptotic confidence intervals for Poisson regression. *J. Multivariate Anal.* **98** (2007), no. 5, 1072–1094.
7. E. D. Kolaczyk, Nonparametric Estimation of Gamma-Ray Burst intensities using Haar wavelets. *The Astrophysical Journal* **483** (1997), no. 1, 340–349.
8. E. A. Nadaraja, On a regression estimate. (Russian) *Teor. Verojatnost. i Primenen.* **9** (1964), 157–159.
9. E. A. Nadaraya, R. M. Absava, *Some Probems of the Theory of Nonparametric Estimation of Functional Characteristics of the Observations Distribution Law.* (Russian) Tbilisi University Press, Tbilisi, 2008.

10. Y. Pawitan, F. O'Sullivan, Data-dependent bandwidth selection for emission computed tomography reconstruction. *IEEE Transactions on Medical Imaging* **12** (1993), no. 2, 167–172.
11. G. S. Watson, Smooth regression analysis. *Sankhya Ser. A* **26** (1964), 359–372.

(Received 14.09.2022)

Department of Mathematics, Faculty of Exact and Natural Sciences, Ivane Javakhishvili Tbilisi State University, 13 University Str., Tbilisi 0186, Georgia

Email address: `petre.babilua@tsu.ge`

Email address: `elizbar.nadaraya@tsu.ge`