UDC 519.234.22, 519.236.2

Applying the maximum likelihood method for constructing asymptotically effective nonparametrical estimators of functionals from the regression function

Yu. I. Pastukhova[1]

[1] *Central Economics and Mathematics Institute of RAS, Nakhimovsky prosp. 47, Moscow, 117418, Russia*

Abstract. In the case when the sample design is given, and the "noise" conditional distribution function is known, nonparametric estimators of linear and smooth functionals from the regression function, for which minimax lower bound of mean square risk is asymptotically tight, are constructed

Keywords: asymptotically effective nonparametrical estimator, maximum likelihood method, regression function.

1. Introduction

Let the values $X_1(t_1), ..., X_n(t_n)$ of unknown function $f(t)$ be observations at some point $t_1, ..., t_n$, $t \in [0;1]$. The values $X_1(t_1), ..., X_n(t_n)$ are conditionally independent for a given design $t^{(n)} = (t_1, ..., t_n)$.

Let us consider the problem of estimation of linear functional

$$L(f) = \int_0^1 f(t)l(t)dt$$

and the smooth functional $F : L_2 \to R^1$. It is supposed that $f(t) \in K \subset L_2$, where K is known compact in L_2, $l(t)$ is the known function in L_2.

Assume that the regressor's design is given. It means that the random variables $t_1, \ldots, t_n$ are independent and have the common distribution density $p(t)$,

$$0 < c \leq p(t) \leq C < \infty,$$

here c, C are some constants.

2. Main Section

Observation $X_i(t_i)$ can be written in the form

$$X_i(t_i) = f(t_i) + \varepsilon_i(t_i) \quad (i = 1, 2, \ldots, n).$$

Here $\varepsilon_i(t_i) = X_i(t_i) - f(t_i)$ is the "noise",

$$E\{X_i(t_i) - f(t_i)|t_i\} = 0,$$

$\varepsilon_i(t_i), \ldots, \varepsilon_n(t_n)$ are conditionally independent for a given design $t^{(n)}$, $f(t)$ is regression function. Let $f \in K \subset L_2$, $\sup\limits_{f\in K} |f(0)| < \infty$ were K is some compact in L_2.

The distribution of $X_i(t_i) - f(t_i)$, for a given $t_i = t$ is supposed to be known and the density function $q(x|t)$ with respect to the Lebesgue measure in R^1 is absolutely continuous in x, has the finite Fisher's information

$$I(t) = \int\limits_{R^1} \frac{(q'_x(x|\,t))^2}{q(x|\,t)}\, dx$$

and satisfies the following regularity conditions

$$0 < \inf_{[0;1]} I(t) \le \sup_{[0;1]} I(t) < \infty; \qquad \sup_{[0;1]} \int\limits_{R^1} |x| q(x|t) dx < \infty; \tag{1}$$

$$\sup_{[0;1]} \int\limits_{R^1} \left(\frac{q'_x(x+s|t)}{q^{1/2}(x+s|t)} - \frac{q'_x(x|t)}{q^{1/2}(x|t)} \right)^2 dx \to 0 \qquad s \to 0; \tag{2}$$

$$\sup_{[0;1]} \int\limits_{|q'_x(x|t)/q^{1/2}(x|t)|>A} \left(\frac{q'_x(x|t)}{q^{1/2}(x|t)} \right)^2 dx \to 0 \qquad A \to \infty.$$

The functional F is Frechet differentiable in L_2 for $f \in K$ and the derivative $F'(f,t)$ satisfies the Hölder conditional in L_2 with some index $\alpha > 0$, i.e. for $f_1,\ f_2 \in K$

$$\| F'(f_1, \cdot) - F'(f_2, \cdot) \| \le C_1 \| f_1 - f_2 \|^\alpha, \quad (\alpha \le 1).$$

Theorems of asymptotical minimax lower bounds of mean square risk were proved in [1], [2] if conditions (1), (2) are satisfied.

Definition 1. The estimator $\widehat{L_n}$, is asymptotically effective in K non-parametrical estimator of linear functional $L(f)$ for the given experiment design and the known noise conditional density, if

$$\lim_{n\to\infty} \left[n \sup_{f\in K} E \left\{ \widehat{L_n} - L(f) \right\}^2 \right] = \int_0^1 \frac{l^2(t)}{I(t)p(t)} dt,$$

Definition 2. The estimator $\widehat{F_n}$ is asymptotically efficient in K non-parametrical estimator of smooth functional from the regression function, if

$$\lim_{n\to\infty} \sup_K [nE\{\widehat{F_n} - F(f)\}^2 - \int_0^1 \frac{(F'(f,t))^2}{I(t)p(t)}\, dt] = 0.$$

To estimate the linear functional, the interval $[0;1]$ is divided into $N = [\gamma\sqrt{n}](\gamma > 0)$ mutually disjoin parts so that

$$\lim_{N\to\infty}\sup_{f\in K}\left[\int_0^1 f^2(t)l(t)p(t)\,dt - \sum_{k=1}^{N}\frac{\left(\int\limits_{\triangle_k} f(t)l(t)p(t)\,dt\right)^2}{\int\limits_{\triangle_k} l(t)p(t)\,dt}\right] = 0.$$

It is noted that on each set $\triangle_j$ the function $f(t)l(t)$ is approximated successfully by constant

$$S_j = N\int\limits_{\triangle_j} f(t)l(t)\,dt, \quad j = 1,\dots,N$$

Let us denote by A_j the event consisting in the fact that the number of elements from $t^{(n)}$ belonging to $\triangle_j$ greater than $\sqrt{n}$. On each such interval the estimator $\widehat{\delta_j}$ is constructed. It maximizes function

$$Z_j(\delta) = \prod_{t_i\in\triangle_j} q(X_i(t_i) - \tfrac{\delta}{a_j}|t_i)p(t_i) \qquad ,$$

here $a_j = N\int\limits_{\triangle_j} l(t)\,dt.$

then the estimator

$$\widehat{L_n} = \sum_{j=1}^{N}\frac{\widehat{\delta_j}}{N}\,\chi(A_j)$$

($\chi(A)$- is indicator of an event A) is asymptotically efficient in nonparametrical estimator in the sense of definition 1.

To estimate the value $F(f)$ the following conditions will be used

1. $f \in K \subset W_2^{\beta}$ $(\beta > \frac{1}{2})$.
2. The derivative $F'(f,t)$ satisfies the Hölder condition with some index $\alpha > (2\beta)^{-1}$ uniformly in K.
3. The set $\{F'(f,\cdot),\ f \in K\}$ is compact in L_2.
4. $\inf\limits_{t\in[0;1]} I(t) \geq I_c > 0$ (I_c is constant).

Let us divide the sample design $(t_1,\dots,t_n)$ into two parts. The first one has volume n_0 . It is known (see [2]) that for $f \in W_2^{\beta}$ there exists based on $X_1(t_1),\dots,X_{n_0}(t_{n_0})$ an estimator $\widehat{f_{n_0}}$ for which

$$\sup_{f\in K} E \parallel \widehat{f_{n_0}} - f \parallel^2 \leq M n_0^{-\frac{2\beta}{2\beta+1}}.$$

The volume n_0 satisfies the conditions

$$n_0 = [n]^{\nu} \ , \frac{2\beta+1}{2\beta(\alpha+1)} < \nu < 1.$$

The remaining part of the sample is used for the linear functional

$$L_{n_0}(f) = \int_0^1 f(t) F'(\widehat{f_{n_0}}, t)\, dt$$

estimation. The estimator has the form

$$\widehat{L_{n_1}}(f) = \sum_{j=1}^{N} \frac{\widehat{\delta_J}}{N}\, \chi(A_J)\, \chi(|a_j| > N^{-\varepsilon}).$$

Here $N = \gamma\sqrt{n_1}$ ($\gamma > 0$) is the number of equal part $\triangle_j$ which divided $[0;1]$; A_j is the event consisting in the fact that the number of elements from $t^{(n_1)}$ belonging to $\triangle_j$ is greater than $\sqrt{n_1}$; $a_j = N \cdot \int\limits_{\triangle_j} F'(\widehat{f_{n_0}}, t)\, dt$, $\widehat{\delta_J}$ is the value that maximizes function

$$Z_j(\delta) = \prod_{t_i \in \triangle_j} q(X_i(t_i) - \tfrac{\delta}{a_j} | t_i) p(t_i).$$

Than the estimator

$$\widehat{F_n} = F(\widehat{f_{n_0}}) + \widehat{L_{n_1}} - \int_0^1 \widehat{f_{n_0}}(t) F'(\widehat{f_{n_0}}, t)\, dt$$

is asymptotically efficient in K nonparametrical estimator of smooth functional $F(f)$.

3. Conclusions

Thus, the constructed estimators of the linear functionals are the maximum likelihood estimators. This is possible because it is assumed that the function belongs to some compact. It is possible to divide the interval $[0, 1]$ into parts so that on each segment the function is almost constant

References

1. *Pastukhova Yu.I. , Khas'minskii R.Z.* Nonparametric Estimation of a Linear Functional of the Regression Function for a Given Observation

Design // Problems of Information Transmission. — 1988. — Vol. 24, no 3. — P. 215–223.

2. *Pastukhova Yu.I.* A nonparametric estimator of a nonlinear functional of regression under a given plan // Journal of Soviet Mathematics. — 1990. — Vol. 52. — P. 2965–2973.
3. *Pastukhova Yu.I.* Nonparametric Estimation of Linear Functionals of the Regression Function for a Known Conditional Distribution of the Observation Noise // Problems Information Transmission. — 1999. — Vol. 35, no 2. — P. 150–157.
4. *Pastukhova Y.I.* Asymptotically efficient estimation of smooth functionals of the regression function for a known distribution of the observation noise // Problems of Information Transmission. — 2004. — Vol. 40, no. 4. — P. 365–378.