Asymptotic properties of the method of empirical mean for
stationary random processes and homogeneous random fields

Knopov P.S., Gololobov D.

(Kiev)

1. The method of empirical means for discrete and continuous models with dependent observations.
2. The method of empirical means applied to the non-stationary models.
3. The method of empirical means under the restrictions of unknown parameters, described in the form of equalities and inequalities.
4. The method of empirical means for the models, in which the random functions depend on several variables or random fields.
Let \((Y, L(Y))\) be some metric space, where \(L(Y)\) is the minimal \(\sigma\)-algebra on \(Y\), and denote by \(\|\cdot\|\) the norm in \(Y\).

Let \(\{\xi_i, i \in N\}\) be independent identically distributed observations of a random variable defined on a probability space \((\Omega, \mathcal{F}, P)\) with values in \((Y, L(Y))\), and let \(\hat{\xi}\) be a random variable with the same distribution and taking values in the same metric space. We assume that \(I\) is a closed subset in \(\mathbb{R}^l, l \geq 1\), possibly \(I = \mathbb{R}^l\), and \(f: I \times Y \to \mathbb{R}\) is a nonnegative function satisfying the following conditions:

1. \(f(u, z), u \in I,\) is continuous for all \(z \in Y;\)
2. for any \(u \in I,\) the mapping \(f(u, z), z \in Y,\) is \(L(Y)\)-measurable.

The problem consists in finding the minimum point of the function

\[ F(u) = E\left(f(u, \hat{\xi})\right), u \in I, \]

and its minimal value.

This problem is approximated by the following one: find the minimum points of the function

\[ F_n(u) = \frac{1}{n} \sum_{i=1}^{n} f(u, \xi_i), \]

and its minimal value.
We give some examples of regression models, which are widely known to specialists in the field of theoretical and applied statistics.

1. \[ y_i = \sum_{t=1}^{p} x_{it} \alpha^0_t + \varepsilon_i, \quad i = 1, ..., n. \]

Here \( \varepsilon_i, \ i = 1, ..., n \) are independent or stationary dependent random variables, \( x_i = \{x_{it}, i = 1, p\}, \ i = 1, ..., n \) are independent identically distributed random vectors, independent of \( \varepsilon_i, \ i = 1, ..., n \).

The vector \( \alpha^0 = (\alpha^0_1, ..., \alpha^0_p) \) is unknown and would be estimated.

2. \[ y_i = g(x(i), \alpha^0) + \varepsilon_i, \quad x(i) \in \mathbb{R}^p, \]

where the \( p \)-dimensional vector \( x(i) \) and \( \varepsilon_i \) are mutually independent, and each of the sequences \( \{x(i)\} \) and \( \{\varepsilon_i\}, \ i = 1, ..., n \), is the sequence of independent or stationary random vectors or variables.
Some cost functions characterizing the accuracy of the estimate:

1. \( F_n(\alpha) = \frac{1}{n} \sum_{i=1}^{n} \left[ y_i - \sum_{j=1}^{p} x_{ij} \alpha_j \right]^2 \); 

2. \( F_n(\alpha) = \frac{1}{n} \sum_{t=1}^{n} \left[ y_i - g(x(i), \alpha) \right]^2 \); 

3. \( F_n(\alpha) = \frac{1}{n} \sum_{i=1}^{n} \left| y_i - \sum_{t=1}^{p} x_{it} \alpha_t \right| \); 

4. \( F_n(\alpha) = \frac{1}{n} \sum_{i=1}^{n} \left| y_i - g(x(i), \alpha) \right| \).
Theorem 1. Let the following conditions be satisfied:

1. For any $c > 0$, $E \left( \max_{\|u\| \leq c} f(u, \xi) \right) < \infty$, where $\| \cdot \|$ is a norm in $\mathbb{R}^l$;

2. If $P\{ \xi \in Y' \} = 1$, then for all $z \in Y'$ we have $f(u, z) \to \infty$ as $\|z\| \to \infty$;

3. There is a unique point $u_0$, at which the function $F(u)$ attains its minimum.

Then, for any $n$ and $\omega \in \Omega'$, $P(\Omega') = 1$, there is at least one vector $u_n = u_n(\omega) \in I$ for which the minimum value of $F_n(u)$ is attained and, for any $n \geq 1$, the vector $u_n$ can be chosen to be $G'_n$-measurable, where $G'_n = G_n \cap \Omega'$ and $G_n = \sigma\{ \xi_i, i = 1, n \}$. In this case, with probability 1, $u_n \to u_0$ and $F(u_n) \to F(u_0)$.

The proof of this theorem is based on the joint result by A.Dorogovtsev, which in turn is based on the fundamental works by J.Pfanzagl, R.Jenrich and others. Below we state this theorem.
Theorem 2. Let \( (\Omega, U, P) \) be a probability space, \( K \) is a compact subset of some Banach space with a norm \( \| \cdot \| \).

Suppose that

\[
\{Q_n(s) = Q_n(s, \omega) \in K \times \Omega, n \geq 1\}
\]

is a family of real functions satisfying the following conditions:

1) for any \( n, s \) the function \( Q_n(s, \omega), \omega \in \Omega \) is measurable;

2) for fixed \( n \) and \( \omega \) the function \( Q_n(s, \omega), s \in K \) is continuous;

3) for some element \( s_0 \in K \) for each \( s \in K \) one has

\[
P\left\{ \lim_{T \to \infty} Q_n(s, \omega) = \Phi(s, s_0) \right\} = 1,
\]

where \( \Phi(s, s_0) > \Phi(s_0, s_0), s \neq s_0; \)

4) there exist \( \gamma_0 > 0 \) and a function \( c(\gamma), \gamma > 0, c(\gamma) \to 0, \gamma \to 0 \), such that for any element \( s' \in K \) and any \( 0 < \gamma < \gamma_0 \) one has
\[
P\left\{ \lim_{T \to \infty} \sup_{s \neq s'} \left| Q_n(s) - Q_n(s') \right| < c(\gamma) \right\} = 1.
\]

For each \( n \geq 1 \) and \( \omega \in \Omega \) define an element \( s_n = s_n(\omega) \in K \) by the relation

\[
Q_n(s_n) = \min_{s \in K} Q_n(s).
\]

If there is more than one point of minimum of the function \( Q_n \), we will take any point \( s_n \). Then

\[
P\left\{ \|s_n - s_0\| \to 0, Q_n(s_n) \to \Phi(s_0, s_0), n \to \infty \right\} = 1.
\]

Applying the ergodic theorem, it is possible to show that the claim of Theorem 1 holds true, and in this case the random sequence \( \{\xi_i, i \in N\} \) is ergodic and stationary in the restricted sense.

A statement similar to those of Theorem 1 also holds true in the continuous time settings. As well as before, will suppose that \( u \in I \in \mathcal{R}^l \).
Theorem 3. Let \( \{\xi(t), t \in \mathbb{R}\} \) be a random ergodic process stationary in the restricted sense and defined on the probability space \((\Omega, \mathcal{F}, P)\) with values in \(\mathbb{R}\). Let the following conditions be satisfied:

1. for any \(c>0\), \(E\left\{ \max_{\|u\| \leq c} f(u, \xi(0)) \right\} < \infty\);

2. if \(I\) is an unbounded set for any \(z \in Y'\), and \(P\left\{ \xi(t) \in Y' \text{ } \forall t \geq 0 \right\} = 1\), then \(f(u, z) \to \infty\) as \(\|u\| \to \infty\);

3. there is a unique element \(u_0 \in I\) for which the minimal value of the function \(F(u) = Ef(u, \xi(0))\) is attained.

Then, for all \(T>0\) and \(\omega \in \Omega', P(\Omega') = 1\), there is at least one vector \(u(T) \in I\) for which the minimal value of the function

\[
F_T(u) = \frac{1}{T} \int_0^T f(u, \xi(t)) dt
\]

is attained and measurable.

Let \(u_0 = \text{arg min} F(u)\), where \(F(u) = Ef(u, \xi(0))\). Then we have

\[
P\left\{ \lim_{T \to \infty} u(T) = u_0 \right\} = 1, \quad P\left\{ \lim_{T \to \infty} F_T(u_T) = F(u_0) \right\} = 1.
\]
Consider a more general case of non-stationary model. We will suppose that the criterion function depends also on the temporal parameter, i.e., it is a function of three variables. For example, in the discrete time the criterion function has the form

\[ F_n(u) = \frac{1}{n} \sum_{i=1}^{n} f(i, u, \xi_i). \]

As an example, one can take

\[ F_n(u) = \frac{1}{n} \sum_{i=1}^{n} \left[ y_i - g(i, u) \right]^2 \]

or

\[ F_n(u) = \frac{1}{n} \sum_{i=1}^{n} |y_i - g(i, u)| \]

for the model of the observation

\[ y_i = g(i, u_0) + \xi_i. \]

Assume also that the unknown parameter is an element of some functional space. For example one can consider the problem of the estimation of the unknown function \( u(t) \in K \), where \( K \) is the compact set of functions defined on \([0,1]\), by observations

\[ y_i = u\left(\frac{j}{n}\right) + \xi_i, i = 1, ..., n \]

with some criterion function.
**Remark 1.** For the proof of our statements we used the conditions of ergodicity for stationary sequences or processes, which are well known and therefore we will not stop on them. We will only note that if the strong mixing condition is satisfied, then the ergodicity takes place. When the observations are the homogeneous random fields, the problem is much more complicated. In this case finding the ergodicity conditions, or finding the conditions when the law of large numbers takes place, requires a ground. We quote only one of the results, which belongs Yurinski and that can be used for the proof of convergence by the method empirical means, if the observations are the homogeneous random fields.
Theorem 4. Let \( \xi(\vec{t}), \vec{t} \in \mathbb{R}^k \) be a homogeneous in the wide sense random field with values in \( \mathbb{R}^k \): \( M\xi(\vec{t})=0, b(\vec{t}) = M\xi(\vec{t} + \vec{s})\xi(\vec{s}) \). Suppose that the realizations of \( \xi(\vec{t}) \) are Lebesgue measurable in \( \mathbb{R}^k \) with probability 1. Denote

\[
X(\tau) = \int_{\|\vec{t}\|<\tau} \xi(\vec{t})d\vec{t}, \quad B(\tau) = \int_{\|\vec{t}\|<\tau} b(\vec{t})d\vec{t}, \quad \Lambda(\tau) = \int_{\|\vec{t}\|<\tau} d\vec{t}.
\]

Suppose that

\[
\int_{0}^{\infty} \left( \frac{1}{\rho^k} \left( \frac{1}{\tau^2} |B(\tau)| d\tau \right) \right) d\rho < \infty.
\]

Then with probability 1

\[
\frac{X(\tau)}{\Lambda(\tau)} \to 0, \tau \to \infty.
\]

The following statement takes place.
Theorem 5. Suppose that for an ergodic random field $\xi(t) \in Y$ the conditions below are fulfilled:

1) for any $c > 0$

\[
E \left\{ \max \left( \frac{f(u, \xi(0))}{\|u\|} \right)^2 \right\} < \infty;
\]

2) if $J$ is unbounded then for each $z \in Y$

\[
f(u, z) \to \infty, \|u\| \to \infty;
\]

3) the function, which defines a criterion, has a unique minimum point $u_0$;

Then for all $T > 0, \omega \in \Omega, P(\Omega) = 1$ there exists at least one minimum point $\tilde{u}(T) = \tilde{u}(T, \omega)$ of the empirical function of the criteria and

\[
P \{ \tilde{u}(T) \to u_0, F_{\tilde{u}(T)}(\tilde{u}(T)) \to F(u_0), T \to \infty \} = 1.
\]

For the proof of the ergodicity we can use the previous theorem.
Another important property of estimates about which I will speak is their limit distributions. It is important to know if the true value is an interior point of the domain of admissible values, or it belongs to the boundary of this domain. I will not formulate all the conditions under which one can prove the statement on the limit distribution of the estimate, because these conditions are indeed very complicated. But I would like to mention, that under some conditions on smoothness of the criterion function, strong mixing condition on the respective random processes (or random fields), the asymptotic distribution is Gaussian.

For example, if we have the observations of the random in area $\|t\| < T$, then the normed variable has a form

$$\left( \int_{\|\tilde{t}\| < T} \frac{1}{2} dt \right)^{\frac{1}{2}} (\tilde{u}(T) - \overline{u}_0)$$

and

$$\left( \int_{\|\tilde{t}\| < T} \frac{1}{2} dt \right)^{\frac{1}{2}} \left( F_T(\tilde{u}(T)) - F(\overline{u}_0) \right).$$

Further we consider a case where the restrictions are of the form

$$J = \{u : g(u) = (g_1(u),...,g_n(u)) \leq 0 \}.$$
Then the family of vectors \( \eta_T = \left( \int_{|t|<T} \hat{d}t \right)^{1/2} (\overrightarrow{u(T)} - \overrightarrow{u_0}) \) converges weakly to the random vector \( \eta \) which is the solution to the problem

\[
\frac{1}{2} u \Phi(u_0) u + \xi u \rightarrow \min,
\]

\[
\nabla g^T(u_0) u \leq 0,
\]

where \( \xi \) is normal random vector.

Apart from the results about consistency and asymptotic distribution, there are developments concerning large deviations for our models with dependent observations. Such investigations are very important for the estimation of speed of convergence. For some models with dependent observations such results were proved. For this one can use the results of monograph by Deutschel, J.D. Strook D.W.[8], and the paper by Kaniovskii Yu.M., King A., Wets R.J [9].
REFERENCES


Thank you for attention