

2. Akers S., Krishnamurthy B. A group theoretic model for symmetric interconnection networks. Proceedings of the International Conference on Parallel Processing, 1986. P. 216–223.

3. Schibell S., Stafford R. Processor interconnection networks and Cayley graphs. Discrete Applied Mathematics. 1992. Vol. 40. P. 337–357.

4. Even S., Goldreich O. The Minimum Length Generator Sequence is NP-Hard. Journal of Algorithms. 1981. Vol. 2. P. 311–313.

5. Kuznetsov A. A., Kuznetsova A. S. [Relation between growth functions in symmetric groups and tasks of combinatorial optimization]. *Vestnik SibGAU*, 2012, no. 6 (46), p. 93–97.

© Кузнецов А. А., Кузнецова А. С., 2014

УДК 519.7

НЕПАРАМЕТРИЧЕСКАЯ ОЦЕНКА КРИВОЙ РЕГРЕССИИ В УСЛОВИЯХ БОЛЬШИХ ВЫБОРОК*

А. В. Лапко^{1,2}, В. А. Лапко^{1,2}, Д. В. Борисов¹

¹Сибирский государственный аэрокосмический университет имени академика М. Ф. Решетнева
Российская Федерация, 660014, г. Красноярск, просп. им. газ. «Красноярский рабочий», 31
E-mail: valapko@yandex.ru

²Институт вычислительного моделирования СО РАН
Российская Федерация, 660036, г. Красноярск, Академгородок, 50, стр. 44. E-mail: lapko@icm.krasn.ru

Предлагается методика построения непараметрической регрессии в условиях обучающих выборок большого объема. Синтез модели основывается на декомпозиции исходных статистических данных и анализе вероятностных характеристик получаемых множеств случайных величин. Исследуются асимптотические свойства непараметрической регрессии и рассматриваются результаты вычислительного эксперимента. Устанавливается зависимость свойств непараметрической регрессии от количества интервалов дискретизации значений случайной величины и объема исходных данных. Проводится сравнение аппроксимационных свойств предлагаемой модели и традиционной непараметрической регрессии. Результаты исследований имеют важное значение при решении задач доверительного оценивания плотности вероятности и кривой регрессии.

Ключевые слова: непараметрическая регрессия, плотность вероятности, регрессионная оценка, аппроксимационные свойства, методы дискретизации.

NONPARAMETRIC ESTIMATION OF REGRESSION CURVES IN THE CONDITIONS OF LARGE SAMPLES*

A. V. Lapko^{1,2}, V. A. Lapko^{1,2}, D. V. Borisov¹

¹Siberian State Aerospace University named after academician M. F. Reshetnev
31, Krasnoyarsky Rabochny Av., Krasnoyarsk, 660014, Russian Federation
E-mail: valapko@yandex.ru

²Institute of Computational Modeling, Siberian Branch of RAS
50, Akademgorodok, Krasnoyarsk, 660036, Russian Federation
E-mail: lapko@icm.krasn.ru

The technique of construction of a nonparametric regression in the conditions of training samples of large volume is offered. Model synthesis is based on decomposition of initial statistical data and the analysis of probabilistic characteristics of received random variables sets. Asymptotic properties of a nonparametric regression are investigated and results of computing experiment are considered. Association of nonparametric regression properties on an amount of sampling intervals of values of a random variable and volume of input datas is established. Comparison of approximating properties of offered model and a traditional nonparametric regression is spent. The results of researches are important to the solution of problems of a confidential estimation of a probability density and a regression curves.

Keywords: nonparametric regression, probability density, regression estimations, approximating properties, sampling methods.

* Работа выполнена в рамках базовой части государственного задания Минобрнауки РФ (СибГАУ № Б121/14).

Вычислительная эффективность непараметрических алгоритмов обработки информации во многом определяется объемом статистических данных и снижается по мере его увеличения, что затрудняет построение систем принятия решений в условиях больших выборок [1; 2].

В подобной ситуации используются принципы декомпозиции исходных статистических данных по их объему и технологии параллельных вычислений. К этому направлению относится смесь непараметрических оценок плотностей вероятности для одномерных и многомерных случайных величин [3–7]. При этом сокращение времени вычислений сопоставимо с количеством составляющих смеси.

Полученные результаты обобщены при оценивании решающей функции в задаче распознавания образов для условий больших выборок. Разработаны двухуровневые непараметрические системы классификации [8; 9]. Установлены асимптотические свойства непараметрических оценок их уравнений разделяющих поверхностей для одномерного и многомерного случаев [10; 11]. Данный подход при восстановлении стохастических зависимостей предполагает разбиение обучающей выборки по её объёму. На этой основе осуществляется синтез семейства частных непараметрических регрессий с последующей их интеграцией в обобщённой модели [12; 13].

Перспективное направление «обхода» проблем больших выборок связано с использованием регрессионной оценки плотности вероятности. Её синтез основан на декомпозиции исходных статистических данных и последующем анализе количественных характеристик получаемого множества случайных величин [14–17].

Цель настоящего исследования состоит в разработке методики синтеза и анализа непараметрических моделей стохастических зависимостей в условиях больших выборок, которая основана на использовании регрессионной оценки плотности вероятности.

Синтез регрессионной оценки плотности вероятности

Пусть имеется выборка $V = (x^i, i = \overline{1, n})$ из n независимых значений одномерной случайной величины x с неизвестной плотностью вероятности $p(x)$.

Разобьем область определения $p(x)$ на N непересекающихся интервалов длиной 2β и сформируем множество случайных величин $X^j, j = \overline{1, N}$. В качестве характеристик X^j примем частоту \bar{P}^j попадания случайной величины x в j -й интервал и его центр z^j . На основе полученной информации определим элементы выборки $V_1 = (z^j, y^j = \bar{P}^j / (2\beta), j = \overline{1, N})$, где z^j центры введенных интервалов и соответствующих им значений оценок y^j плотности вероятности. Границы области Δ изменения случайной величины x априори неизвестны и определяются минимальным и максимальным значениями исходных статистических

данных $V = (x^i, i = \overline{1, n})$. Поэтому центры интервалов $z^j, j = \overline{1, N}$ являются случайными величинами и имеют равномерный закон распределения $p_1(z) = (2\beta N)^{-1}$. Объем N полученной выборки V_1 значительно меньше объема n исходных статистических данных V . Выборка V_1 позволяет оценивание плотности вероятности $p(x)$ свести к задаче восстановления стохастической зависимости.

В качестве приближения по эмпирическим данным V_1 искомой плотности вероятности $p(x)$ примем статистику [14]

$$\bar{p}(x) = c^{-1} \sum_{j=1}^N \bar{P}^j \Phi\left(\frac{x - z^j}{c}\right), \quad (1)$$

которая является непараметрической оценкой условного математического ожидания

$$2\beta N \int y p(x, y) dy. \quad (2)$$

Здесь и далее бесконечные пределы интегрирования опускаются.

В регрессионной оценке плотности вероятности (1) ядерные функции $\Phi(u)$ удовлетворяют условиям H :

$$\begin{aligned} \Phi(u) &= \Phi(-u), \quad 0 \leq \Phi(u) < \infty, \\ \int \Phi(u) du &= 1, \quad \int u^2 \Phi(u) du = 1. \end{aligned}$$

Коэффициенты размытости $c = c(N)$ ядерных функций характеризуют область их определения.

При синтезе $\bar{p}(x)$ в выражение (2) подставляется непараметрическая оценка

$$\bar{p}(x, y) = \frac{1}{nc c_1} \sum_{j=1}^N \Phi\left(\frac{x - z^j}{c}\right) \Phi\left(\frac{y - y^j}{c_1}\right) \quad (3)$$

совместной плотности вероятности $p(x, y)$.

Интегрируя выражение (2) при

$$\frac{1}{c_1} \int y \Phi\left(\frac{y - y^j}{c_1}\right) dy = y^j \text{ и } y^j = \frac{\bar{P}^j}{2\beta},$$

получим регрессионную оценку плотности вероятности (1).

В многомерном случае $x = (x_v, v = \overline{1, k})$ статистика (1) имеет вид

$$\bar{p}(x) = \frac{1}{c^k} \sum_{j=1}^{N^k} \bar{P}^j \prod_{v=1}^k \Phi\left(\frac{x_v - z_v^j}{c}\right). \quad (4)$$

Регрессионные оценки плотности вероятности (1), (4) обладают свойствами асимптотической несмещённости и состоятельности [14]. Из условия минимума асимптотического выражения среднеквадратического отклонения $\bar{p}(x)$ от $p(x)$ определена процедура оптимального выбора количества интервалов дискретизации [18; 19]

$$\bar{N} = \sqrt{\Delta n \int p^2(x) dx}. \quad (5)$$

Количество интервалов дискретизации зависит от вида восстанавливаемой плотности вероятности, длины Δ интервала значений случайной величины и объёма n исходных статистических данных.

Непараметрическая оценка кривой регрессии

Пусть $V = (x^i, y^i, i = \overline{1, n})$ статистическая выборка независимых наблюдений одномерных случайных величин (x, y) , распределенных с неизвестной плотностью вероятности $p(x, y)$ и $p(x) > 0 \forall x \in \Omega(x)$. Априори вид однозначной стохастической зависимости $y = F(x)$ не задан.

В качестве её модели будем использовать условное математическое ожидание

$$\bar{y} = \bar{\varphi}(x) = \frac{1}{p(x)} \int y p(x, y) dy. \quad (6)$$

Решающая функция (6) является оптимальной в смысле минимума квадратического критерия [20].

При оценивании $p(x, y)$ используем технологию синтеза регрессионной оценки плотности вероятности типа (1). Для упрощения преобразований будем считать, что интервалы значений (x, y) одинаковы.

Тогда

$$\bar{p}(x, y) = c^{-2} \sum_{j=1}^{N^2} P^j \Phi\left(\frac{x-x^j}{c}\right) \Phi\left(\frac{y-y^j}{c}\right), \quad (7)$$

где N^2 – объем массива данных $(x^j, y^j, \bar{P}^j, j = \overline{1, N^2})$, формируемых на основе исходной статистической выборки $V = (x^i, y^i, i = \overline{1, n})$.

Значение \bar{P}^j определяет частоту встречаемости наблюдений из V в элементе $(x^j \pm \beta, y^j \pm \beta)$ равномерной сетки в пространстве (x, y) .

Тогда с учетом

$$p\left(\frac{y}{x}\right) = \frac{p(x, y)}{p_1(x)} = 2 N \beta p(x, y)$$

непараметрическую оценку (6) представим в виде

$$\bar{y} = \bar{\varphi}(x) = \frac{2 N \beta}{c} \sum_{i=1}^{N^2} y^i \bar{P}^i \Phi\left(\frac{x-x^i}{c}\right). \quad (8)$$

Исследуем свойство асимптотической несмещённости статистики (8). Пусть функции $\varphi(x), p(x, y)$ ограничены и непрерывны со всеми своими производными до 2-го порядка включительно. Ядерные функции $\Phi(u)$ удовлетворяют условиям H .

Определим частоту появления случайных величин в виде

$$\bar{P}^j = \frac{1}{n} \sum_{i=1}^n h\left(\frac{x^j - x^i}{\beta}\right) h\left(\frac{y^j - y^i}{\beta}\right),$$

где

$$h(u) = \begin{cases} 1 & \forall |u| \leq 1, \\ 0 & \forall |u| > 1. \end{cases}$$

Наблюдения из выборки $(x^i, y^i), i = \overline{1, n}$ имеют один и тот же закон распределения $p(z_1, z_2)$. Элементы массива данных $(x^j, y^j) j = \overline{1, N^2}$ характеризуются плотностью вероятности $p(t_1, t_2)$.

Тогда

$$M(\bar{\varphi}(x)) = \frac{2 N^3 \beta}{c} \int \dots \int z_1 \Phi\left(\frac{x-z_2}{c}\right) h\left(\frac{z_2-t_2}{\beta}\right) \times \\ \times h\left(\frac{z_1-t_1}{\beta}\right) p(t_1, t_2) p(z_1, z_2) dt_1 \dots dz_2,$$

где M – знак математического ожидания.

Осуществим замену переменных $(z_1-t_1)/\beta = u_1, (z_2-t_2)/\beta = u_2$. С учетом $p(z_1, z_2) = (2\beta N)^2$ получим

$$M(\bar{\varphi}(x)) = \frac{N\beta}{2c} \int \dots \int (\beta u_1 + t_1) \Phi\left(\frac{x-\beta u_2-t_2}{c}\right) \times \\ \times h(u_2) h(u_1) p(t_1, t_2) du_1 \dots dt_2 = \\ = \frac{N\beta}{2c} \iiint \Phi\left(\frac{x-\beta u_2-t_2}{c}\right) \times \\ \times h(u_2) p(t_1, t_2) du_2 dt_1 dt_2. \quad (9)$$

Обозначим $\varphi(t_2) = M(t_1/t_2)$. Проведём замену $(x-\beta u_2-t_2)/c = u$. На этой основе преобразуем (9) к виду

$$M(\bar{\varphi}(x)) = N\beta \iint \varphi(x-\beta u_2-cu) \times \\ \times \Phi(u) h(u_2) p(x-\beta u_2-cu) du du_2.$$

Разложим функции $\varphi(x-\beta u_2-cu), p(x-\beta u_2-cu)$ в ряд Тейлора в точке x . Тогда при достаточно больших n приходим к утверждению

$$M(\bar{\varphi}(x)) \sim \varphi(x) \frac{p(x)}{p_1(x)} + \left(\frac{\beta^2}{3} + c^2\right) \times \\ \times \left(\frac{\varphi(x) p^{(2)}(x)}{2p_1(x)} + \frac{\varphi^{(1)}(x) p^{(1)}(x)}{p_1(x)} + \varphi^{(2)}(x) p(x)\right) + \\ + \frac{\varphi^{(2)}(x) p^{(2)}(x)}{6} \beta^2 + O(c^4, \beta^4, c^2 \beta^2). \quad (10)$$

Статистика (8) обладает свойством асимптотической несмещенности по отношению к оптимальному решающему правилу (6), если закон распределения исходной выборки наблюдений x равномерный, т. е. $p(x) = p_1(x)$. Это возможно при проведении активного эксперимента при исследовании зависимости $y = F(x)$.

Для устранения смещения необходимо умножить статистику (8) на отношение $p_1(x) / \bar{p}(x)$. В результате получим

$$\bar{\bar{\varphi}}(x) = \frac{\sum_{i=1}^{N^2} y^i \alpha^i(x)}{\sum_{i=1}^{N^2} \alpha^i(x)}, \quad \alpha_i(x) = \bar{P}^i \Phi\left(\frac{x-x^i}{c}\right). \quad (11)$$

В многомерном случае при $x = (x_1, \dots, x_k)$ непараметрическая регрессия имеет вид

$$\bar{\bar{\varphi}}(x) = \frac{\sum_{i=1}^{N^{k+1}} y^i \alpha^i(x)}{\sum_{i=1}^{N^{k+1}} \alpha^i(x)}, \quad \alpha^i(x) = \bar{P}^i \prod_{v=1}^k \Phi\left(\frac{x_v - x_v^i}{c_v}\right).$$

**Анализ результатов
вычислительных экспериментов**

Исследовались свойства непараметрических моделей типа (8), (10) и традиционной непараметрической регрессии [20] методом вычислительного эксперимента. Исходные данные $(x^i, y^i, i = \overline{1, n})$ формировались в соответствии с примером работы [21]

$$y = F(x) = 1 - x + \exp(-200(x-0,5)^2). \quad (12)$$

Значения y^i вычислялись по формуле

$$y^i = F(x^i) + 2(0,5 - \varepsilon^i) F(x^i) r, \quad i = \overline{1, n}.$$

Случайная величина $\varepsilon \in [0; 1]$ имеет равномерный закон распределения, а r – уровень помех.

Оценка эффективности изучаемых моделей определялась критерием

$$\tilde{W} = \frac{1}{m} \sum_{j=1}^m \left| \frac{F(x^j) - \bar{\varphi}(x^j)}{F(x^j)} \right|,$$

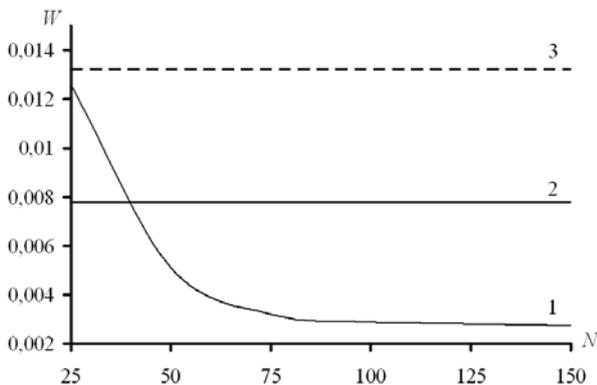
где m – объём контрольной выборки.

При построении непараметрических моделей в качестве ядерных функций использовалось ядро В. А. Епанечникова [22].

$$\Phi(u) = \begin{cases} \frac{3}{4\sqrt{5}} - \frac{3u^2}{20\sqrt{5}} & \forall |u| < \sqrt{5} \\ 0 & \forall |u| \geq \sqrt{5}. \end{cases}$$

Вычислительные эксперименты при фиксированных условиях исследования повторяются 20 раз. По полученным результатам вычислялись ошибки аппроксимации $\tilde{W}^t, t = \overline{1, 20}$ и среднее \bar{W} их значение.

Установлено, что объём исходной выборки n может быть сокращён на порядок и выше. На такую же величину повышается вычислительная эффективность непараметрической регрессии. Обнаружена линейная зависимость между значениями уровня помех r и количеством N интервалов дискретизации.



Зависимость средней относительной ошибки аппроксимации \bar{W} от количества интервалов дискретизации N при $n = 10000, r = 0,05$: кривая 1 соответствует модели $\bar{\varphi}(x)$ (11); прямые 2 и 3 характеризуют традиционную непараметрическую регрессию при объёме обучающей выборки $n = 500$ и $n = 200$

Точность аппроксимации непараметрической регрессии $\bar{\varphi}(x)$ (11) выше в два и более раза, чем со смещением (8). Традиционная непараметрическая регрессия менее устойчива к уровню помех.

С ростом количества N интервалов дискретизации аппроксимационные свойства статистики (11) улучшаются, что согласуется с результатами аналитических исследований (см. рисунок). Данный факт особенно проявляется при высоком уровне помех.

Непараметрическая модель стохастической зависимости (11) является эффективным средством обработки данных большого объема. Ее синтез осуществляется путем декомпозиции исходной статистической информации и анализа количественных характеристик, получаемых множеств случайных величин. На этой основе осуществляется «сжатие» исходных данных. Предлагаемая статистика обладает свойством асимптотической несмещенности.

По своим аппроксимационным свойствам и устойчивости к помехам она значительно превосходит традиционную непараметрическую регрессию. Особенность структуры разработанной непараметрической модели стохастической зависимости позволяет решить проблему её доверительного оценивания.

Библиографические ссылки

1. Лапко А. В., Лапко В. А. Непараметрические системы обработки неоднородной информации. Новосибирск : Наука, 2007. 174 с.
2. Лапко В. А., Капустин А. Н. Синтез нелинейных непараметрических коллективов решающих правил в задачах распознавания образов // Автометрия. 2006. Т. 42, № 6. С. 26–33.
3. Лапко А. В., Лапко В. А., Егорочкин И. А. Непараметрические оценки смеси плотностей вероятности и их применение в задаче распознавания образов // Системы управления и информационные технологии. 2009. № 1 (35). С. 60–64.
4. Лапко А. В., Лапко В. А. Анализ свойств смеси непараметрических оценок плотности вероятности многомерной случайной величины // Вестник СибГАУ. 2010. № 2 (28). С. 32–35.
5. Лапко А. В., Лапко В. А. Анализ свойств непараметрических оценок смеси плотностей вероятности при различных условиях распределения статистических данных // Информатика и системы управления. 2013. № 1 (35). С. 119–126.
6. Лапко А. В., Лапко В. А. Свойства непараметрической оценки плотности вероятности многомерных случайных величин в условиях больших выборок // Информатика и системы управления. 2012. № 2. С. 121–126.
7. Лапко А. В., Лапко В. А. Синтез структуры смеси непараметрических оценок плотности вероятности многомерной случайной величины // Системы управления и информационные технологии. 2011. Т. 43, № 1. С. 12–15.
8. Лапко А. В., Лапко В. А. Разработка и исследование двухуровневых непараметрических систем классификации // Автометрия. 2010. Т. 46, № 1. С. 70–78.

9. Лапко А. В., Лапко В. А. Коллектив непараметрических решающих функций в двальтернативной задаче распознавания образов // Системы управления и информационные технологии. 2009. № 3.1 (37). С. 156–160.

10. Лапко А. В., Лапко В. А. Асимптотические свойства многомерной непараметрической оценки уравнения разделяющей поверхности в двальтернативной задаче распознавания образов // Системы управления и информационные технологии. 2010. № 1 (39). С. 16–19.

11. Лапко А. В., Лапко В. А. Непараметрическая оценка уравнения разделяющей поверхности в условиях больших выборок и её свойства // Системы управления и информационные технологии. 2010. № 1.2 (39). С. 300–304.

12. Лапко А. В., Лапко В. А., Варочкин С. С. Коллектив непараметрических регрессий, основанный на принципе декомпозиции обучающей выборки // Вестник СибГАУ. 2009. № 1 (22). Ч. 2. С. 38–40.

13. Лапко А. В., Лапко В. А. Коллектив многомерных непараметрических регрессий, основанный на декомпозиции обучающей выборки по её объёму // Вестник СибГАУ. 2012. № 3 (43). С. 42–46.

14. Лапко А. В., Лапко В. А. Регрессионная оценка плотности вероятности и её свойства // Системы управления и информационные технологии. 2012. Т. 49, № 3. С. 152–156.

15. Лапко А. В., Лапко В. А. Непараметрические методики анализа множеств случайных величин // Автометрия. 2003. Т. 39, № 1. С. 54–61.

16. Лапко А. В., Лапко В. А. Дискретизация интервала изменения значений случайной величины на основе результатов оптимизации непараметрической оценки плотности вероятности // Информатика и системы управления. 2013. 4 (38). С. 63–69.

17. Лапко А. В., Лапко В. А. Построение доверительных границ для плотности вероятности на основе её регрессионной оценки // Метрология. 2013. № 12. С. 3–9.

18. Лапко А. В., Лапко В. А. Оптимальный выбор количества интервалов дискретизации области изменения одномерной случайной величины при оценивании плотности вероятности // Измерительная техника. 2013. № 7. С. 24–27.

19. Lapko A. V., Lapko V. A. Optimal selection of the number of sampling intervals in domain of variation of a one-dimensional random variable in estimation of the probability density // Measurement Techniques. 2013. Vol. 56, no. 7. P. 24–27 (DOI: 10.1007/s11018-013-0279-x).

20. Надарая Э. А. Замечания о непараметрических оценках плотности вероятности и кривой регрессии // Теория вероятности и её применения. 1970. Т. 15, № 1. С. 139–142.

21. Хардле В. Прикладная непараметрическая регрессия. М. : Мир, 1993. 300 с.

22. Епанечников В. А. Непараметрическая оценка многомерной плотности вероятности // Теория вероятности и её применения. 1969. Т. 14, вып. 1. С. 156–161.

References

1. Lapko A. V., Lapko V. A. *Neparametricheskie sistemy obrabotki neodnorodnoj informacii* (Nonparametric heterogeneous information processing system). Novosibirsk, Nauka, 2007. 174 p.

2. Lapko V. A., Kapustin A. N. *Avtometrija*. 2006, vol. 42, no. 6, p. 26–33.

3. Lapko A. V., Lapko V. A., Egorochkin I. A. *Sistemy upravlenija i informacionnye tehnologii*. 2009, no. 1 (35), p. 60–64.

4. Lapko A. V., Lapko V. A. *Vestnik SibGAU*. 2010, no. 2 (28), p. 32–35.

5. Lapko A. V., Lapko V. A. *Informatika i sistemy upravlenija*. 2013, no. 1 (35), p. 119–126.

6. Lapko A. V., Lapko V. A. *Informatika i sistemy upravlenija*. 2012, no. 2, p. 121–126.

7. Lapko A. V., Lapko V. A. *Sistemy upravlenija i informacionnye tehnologii*. 2011, vol. 43, no. 1, p. 12–15.

8. Lapko A. V., Lapko V. A. *Avtometrija*. 2010, vol. 46, no. 1, p. 70–78.

9. Lapko A. V., Lapko V. A. *Sistemy upravlenija i informacionnye tehnologii*. 2009, no. 3.1 (37), p. 156–160.

10. Lapko A. V., Lapko V. A. *Sistemy upravlenija i informacionnye tehnologii*. 2010, no. 1 (39), p. 16–19.

11. Lapko A. V., Lapko V. A. *Sistemy upravlenija i informacionnye tehnologii*. 2010, no. 1.2 (39), p. 300–304.

12. Lapko A. V., Lapko V. A., Varochkin S. S. *Vestnik SibGAU*. 2009, no. 1 (22), part 2, p. 38–40.

13. Lapko A. V., Lapko V. A. *Vestnik SibGAU*. 2012, no. 3 (43), p. 42–46.

14. Lapko A. V., Lapko V. A. *Sistemy upravlenija i informacionnye tehnologii*. 2012, vol. 49, no. 3, p. 152–156.

15. Lapko A. V., Lapko V. A. *Avtometrija*. 2003, vol. 39, no. 1, p. 54–61.

16. Lapko A. V., Lapko V. A. *Informatika i sistemy upravlenija*. 2013, no. 4 (38), p. 63–69.

17. Lapko A. V., Lapko V. A. *Metrologija*. 2013, no. 12, p. 3–9.

18. Lapko A. V., Lapko V. A. *Izmeritel'naja tehnika*. 2013, no. 7, p. 24–27.

19. Lapko A. V., Lapko V. A. Optimal selection of the number of sampling intervals in domain of variation of a one-dimensional random variable in estimation of the probability density. *Measurement Techniques*, 2013, vol. 56, no. 7, p. 24 – 27. doi: 10.1007/s11018-013-0279-x.

20. Nadaraja Je. A. *Teorija verojatnosti i ejo primenenija*. 1970, vol. 15, no. 1, p. 139–142.

21. Hardle V. *Prikladnaja neparametricheskaja regressija* (Applied nonparametric regression). Moscow, Mir, 1993. 300 p.

22. Epanechnikov V. A. *Teorija verojatnosti i ee primenenija*. 1969, vol. 14, no. 1, p. 156–161.