

УДК 519.237.5

## Оценивание параметров регрессионных моделей методом усеченного максимального правдоподобия\*

А.В. ФАДДЕЕНКОВ<sup>1</sup>, Е.А. ХАЙЛЕНКО<sup>2</sup>

<sup>1</sup> 630073, РФ, г. Новосибирск, пр. Карла Маркса, 20, Новосибирский государственный технический университет, кандидат технических наук, доцент. E-mail: faddeenkov@corp.nstu.ru

<sup>2</sup> 630073, РФ, г. Новосибирск, пр. Карла Маркса, 20, Новосибирский государственный технический университет, кандидат технических наук. E-mail: xajlenko@corp.nstu.ru

В работе представлен новый метод оценивания параметров регрессионных уравнений – метод усеченного максимального правдоподобия. Данный метод основан на объединении подходов максимального правдоподобия и метода усеченных наименьших квадратов. Сформулирована задача построения линейной модели регрессионного анализа. Для этой задачи подробно описан алгоритм метода усеченного максимального правдоподобия, на простом примере продемонстрированы основные принципы, послужившие основой алгоритма. Обоснована задача анализа устойчивости оценок регрессионных моделей. Анализ работоспособности предлагаемого авторами алгоритма проведен с использованием методов вычислительных экспериментов. При проведении вычислительных экспериментов моделирование случайных ошибок проводилось методами Монте-Карло. Распределение случайных ошибок моделировалось как смесь двух распределений. В роли базового распределения случайных ошибок рассматривалось нормальное распределение. Оценивание параметров регрессионной модели проводилось с использованием метода наименьших квадратов, метода усеченных наименьших квадратов и метода усеченного максимального правдоподобия. В качестве меры точности использовалась относительная погрешность оценивания параметров регрессионной модели. Были рассмотрены различные варианты засорения: несмещенное засорение с нулевым средним, симметричное и асимметричное засорения. Доля выбросов варьировалась от 2,5 до 20 %. Для каждого из вариантов засорения исходных данных проводилась серия из 100 экспериментов, результаты которых усреднялись. Результаты вычислительных экспериментов показали, что метод усеченного максимального правдоподобия обладает хорошей устойчивостью к различного рода засорениям исходных данных. В большинстве случаев он практически не уступает методу усеченных наименьших квадратов, а в ряде случаев значительно превосходит его по точности.

**Ключевые слова:** регрессионная модель, устойчивость оценок, метод усеченного максимального правдоподобия, метод усеченных наименьших квадратов, метод Монте-Карло, выбросы, смесь распределений, вычислительный эксперимент

DOI: 10.17212/1814-1196-2016-4-135-145

---

\* Статья получена 01 июля 2016 г.

## ВВЕДЕНИЕ

Регрессионный анализ является одним из самых распространенных методов статистического анализа данных. При этом практическое применение этого инструмента не требует каких-либо уникальных знаний прикладной статистики и доступно широкому кругу исследователей. В подавляющем большинстве статистических пакетов в том или ином объеме реализованы алгоритмы регрессионного анализа [1–4]. Однако не следует забывать, что качество регрессионных моделей существенно зависит от методов, используемых при оценивании параметров. В последние десятилетия большое внимание уделяется проблеме устойчивости оценок регрессионного анализа [5–14]. В случаях нарушений основных предположений регрессионного анализа или наличия выбросов в исходных данных используются методы оценивания параметров, такие как метод наименьших модулей, знаковый метод, метод усеченных наименьших квадратов (LTS-метод), метод наименьших медиан (LMS-метод) [15–17].

В данной работе представлен новый, разработанный авторами метод оценивания параметров регрессионных уравнений – метод усеченного максимального правдоподобия (MTL-метод), основанный на объединении подходов максимального правдоподобия и LTS-метода.

## ПОСТАНОВКА ЗАДАЧИ

Рассмотрим регрессионное уравнение вида

$$y = X\theta + \varepsilon, \quad (1)$$

где  $X = \begin{bmatrix} f_1(x_{11}) & \cdots & f_m(x_{1m}) \\ \vdots & \ddots & \vdots \\ f_1(x_{n1}) & \cdots & f_m(x_{nm}) \end{bmatrix}$  – матрица регрессоров, имеющая полный

столбцовый ранг, т. е.  $rg(X) = m$ ;  $m$  – количество регрессоров;  $n$  – количество испытаний;  $f(x) = (f_1(x), \dots, f_m(x))^T$  – вектор известных действительных функций;  $x_{ij}$  – заданные значения входных факторов в  $n$  наблюдениях;  $y = (y_1, \dots, y_n)^T$  – вектор отклика;  $\theta = (\theta_1, \dots, \theta_m)^T$  – вектор неизвестных параметров, подлежащих оцениванию;  $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)^T$  – вектор ошибок наблюдений, имеющих одинаковое распределение. Имеют место следующие предположения [15]:

$$E(\varepsilon) = 0; \quad D(\varepsilon_i) = \sigma_i^2, \quad \sigma_i^2 < \infty, \quad i = 1, \dots, n. \quad (2)$$

Задача состоит в том, чтобы по имеющимся исходным данным (значениям отклика и входных факторов) как можно точнее оценить вектор неизвестных параметров уравнения регрессии (1).

### АЛГОРИМ МЕТОДА УСЕЧЕННОГО МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ

Как известно, на практике фактическое распределение случайных ошибок может существенно отличаться от нормального закона. Рассмотрим случай, когда распределение случайной величины  $\varepsilon$  представлено в виде смеси двух нормальных распределений:

$$F(x) = (1 - \nu)F_1(x, 0, \sigma_1^2) + \nu F_2(x, \mu, \sigma_2^2), \quad (3)$$

где  $F_1(x, 0, \sigma_1^2)$  – функция нормального распределения с математическим ожиданием 0 и дисперсией  $\sigma_1^2$ ;  $F_2(x, \mu, \sigma_2^2)$  – функция нормального распределения с математическим ожиданием  $\mu$  и дисперсией  $\sigma_2^2$ ;  $\nu \in [0, 1]$ .

На рис. 1, *a* представлена гистограмма, построенная по выборке из 500 случайных величин, соответствующих распределению (3) при  $\sigma_1^2 = 0.01$ ,  $\mu = 0.5$ ,  $\sigma_2^2 = 0.005$ ,  $\nu = 0.1$ . На представленном примере хорошо видно, что распределение является двумодальным и обладает ярко выраженной асимметрией.

Основная идея МТЛ-метода состоит в том, чтобы исключить из исходного набора данных те наблюдения, которые вызывают отклонение распределения от базового нормального. Необходимо сформировать некоторое подмножество исходных данных исходя из информативности каждого наблюдения. Поскольку наблюдения являются независимыми, функция правдоподобия имеет вид

$$L(\varepsilon) = \prod_{i=1}^n g(\varepsilon_i),$$

где  $g(\varepsilon_i)$  – функция плотности нормального распределения. Тогда логарифм функции правдоподобия

$$\ln L(\varepsilon) = \sum_{i=1}^n L(\varepsilon_i) = \sum_{i=1}^n \ln g(\varepsilon_i). \quad (4)$$

Суть алгоритма формирования подмножества состоит в следующем: выбирается размер подмножества  $h$ , для всех наблюдений вычисляется вклад в функцию правдоподобия  $L(\varepsilon_i)$ . Наблюдения сортируются по убыванию величины  $L(\varepsilon_i)$ , и в подмножество вносятся первые  $h$  наблюдений.

Для данного примера возьмем  $h = 0.9n = 450$ . На рис. 1, *б* изображена гистограмма распределения на подмножестве. Как видно, распределение стало унимодальным и симметричным.

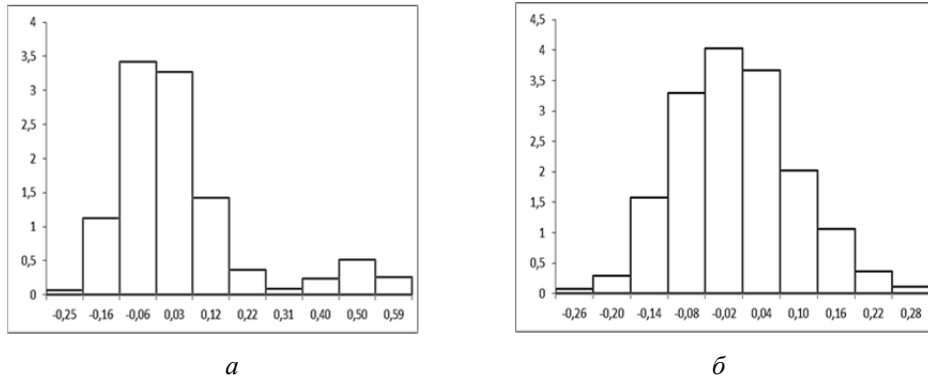


Рис. 1. Гистограмма распределения

Предложенный подход лег в основу метода усеченного максимального правдоподобия для устойчивого оценивания параметров регрессионных зависимостей, алгоритм которого состоит в следующем.

1. Из множества всех индексов наблюдений  $\{1, \dots, n\}$  случайным образом формируется подмножество индексов  $H_1$ , содержащее  $h$  элементов,

$$h \in \left[ \frac{n+m+1}{2}, n \right], \text{ т. е. } \dim(H_1) = h.$$

2. Вычисляется МНК-оценка вектора параметров регрессионных зависимостей  $\hat{\theta}^1$ .

3. Устанавливается счетчик  $k = 0$ , выполняется С-шаг до сходимости.

4. Вычисляются остатки  $r_i = y_i - \hat{\theta}_1^{k-1} f(x_{i1}) - \hat{\theta}_2^{k-1} f(x_{i2}) - \dots - \hat{\theta}_m^{k-1} f(x_{im})$ ,  $i = 1, \dots, n$ , и проводится идентификация распределения остатков.

5. Вычисляется вклад каждого наблюдения в функцию правдоподобия

$$L_i(r_i) = \ln g(r_i), \quad i = 1, \dots, n.$$

6. Вклады каждого наблюдения в функцию правдоподобия  $L_i(r_i)$  сортируются:

$$L_{\pi(1)}(r_{\pi(1)}) \geq L_{\pi(2)}(r_{\pi(2)}) \geq \dots \geq L_{\pi(n)}(r_{\pi(n)}),$$

где  $\pi(i)$  – номер наблюдения в исходной выборке;  $i$  – номер наблюдения в вариационном ряду.

7. Во множество  $H_k$  записываются номера первых  $h$  наблюдений, соответствующих максимальным вкладам в функцию правдоподобия:

$$H_k = \{\pi(1), \pi(2), \dots, \pi(h)\}.$$

8. Вычисляется МНК-оценка  $\hat{\theta}^k$  вектора параметров регрессионной модели по наблюдениям, соответствующим номерам из подмножества  $H_k$ .

$$9. \text{ Вычисляется } Q^k = \sum_{i=1}^h L_i(r_i).$$

Шаги 4–9 повторяются до тех пор, пока не будет достигнута требуемая точность вычислений  $\delta$ , т. е.  $|Q^k - Q^{k-1}| \leq \delta$ .

## РЕЗУЛЬТАТЫ ВЫЧИСЛИТЕЛЬНЫХ ЭКСПЕРИМЕНТОВ

Исследование работоспособности предложенного метода проводилось с использованием вычислительных экспериментов. Для тестирования была использована модель

$$y_i = \theta_1 + \theta_2 x_{1i} + \theta_3 x_{1i}^2 + \varepsilon_i, \quad i=1, \dots, n, \quad (5)$$

где количество регрессоров  $m=3$ , количество наблюдений  $n=1000$ , значения входных факторов  $x_{ij}$  выбирались из отрезка  $[0, 1]$ ,  $\theta_{\text{ист}} = (1, 1.5, 2)^T$ . Случайные ошибки  $\varepsilon_i$ ,  $i=1, \dots, n$ , моделировались независимыми и одинаково распределенными с функцией распределения вида (3), параметры которого варьировались.

Для каждой комбинации параметров распределения (3) (доли выбросов  $\nu$ , смещения  $\mu$  и величины дисперсии  $\sigma_2^2$ ) проводилась серия из 100 вычислительных экспериментов. В каждом эксперименте моделировались исходные данные и проводилось оценивание параметров регрессионной модели с использованием метода усеченного максимального правдоподобия (MTL), быстрого метода усеченных наименьших квадратов (FAST-LTS) [17] и классического метода наименьших квадратов (МНК). Во всех экспериментах размер оценочного подмножества выбирался равным  $h = (1 - \nu)n$ .

В качестве итогового показателя точности оценивания было взято усредненное по 100 экспериментам соотношение

$$\psi = \sum_{i=1}^m \frac{(\theta_i^{\text{ист}} - \hat{\theta}_i)^2}{(\theta_i^{\text{ист}})^2}. \quad (6)$$

В первой серии экспериментов рассматривался случай, когда распределение ошибок наблюдений является симметричным, функция распределения (3) имеет вид  $F(x) = (1 - \nu)N(0, 0.01) + \nu N(0, 0.005)$ .

В табл. 1 представлены результаты оценивания параметров регрессионной зависимости (5), по которой можно отметить, что наиболее точные результаты показывает МНК, поскольку выбросы являются несмещенными и обладают меньшей, чем основная масса наблюдений, дисперсией и, следовательно, более точными. Методы FAST-LTS и MTL проигрывают в данном случае в точности, потому что оценивание проводится не по всем наблюдениям, а только по некоторым.

Таблица 1

**Точность оценивания параметров регрессионной модели при симметричном засорении**

Метод	Доля выбросов, $\nu$				
	0.025	0.05	0.10	0.15	0.20
МНК	1,342E-03	1,511E-03	1,323E-03	1,386E-03	1,297E-03
FAST-LTS	1,805E-03	2,059E-03	2,660E-03	3,097E-03	2,754E-03
MTL	1,754E-03	2,063E-03	2,627E-03	2,980E-03	2,715E-03

Далее рассмотрим два случая: когда засорение имеет нормальное распределение со смещением  $\mu = 0.5$ , т. е. выбросы произошли только справа, и когда смещение  $\mu = \pm 0.5$ , выбросы присутствуют с обеих сторон от итогового распределения, дисперсия  $\sigma_2^2 = 0.005$ ,  $\sigma_1^2 = 0.01$ .

Таблица 2

**Точность оценивания в зависимости от доли выбросов  
при  $\sigma_1^2 = 0.01$ ,  $\sigma_2^2 = 0.005$ ,  $\mu = \pm 0.5$**

Метод	Доля выбросов, $\nu$				
	0.025	0.05	0.10	0.15	0.20
МНК	1,230E-03	1,996E-03	1,628E-03	2,206E-03	2,491E-03
FAST-LTS	2,323E-03	3,506E-03	7,176E-03	1,214E-02	1,556E-02
MTL	1,270E-03	1,969E-03	1,564E-03	2,334E-03	2,355E-03

Таблица 3

**Точность оценивания в зависимости от доли выбросов  
при  $\sigma_1^2 = 0.01$ ,  $\sigma_2^2 = 0.005$ ,  $\mu = \pm 0.5$**

Метод	Доля выбросов, $\nu$				
	0.025	0.05	0.10	0.15	0.20
МНК	2,559E-03	3,239E-03	7,946E-03	1,366E-02	1,929E-02
FAST-LTS	1,791E-03	1,621E-03	1,840E-03	2,197E-03	2,587E-03
MTL	1,815E-03	1,566E-03	1,697E-03	1,959E-03	3,248E-03

Как видно из табл. 2 и 3, FAST-LTS и MTL показали близкие результаты и существенно выигрывают в точности по сравнению с МНК-оценками, что объясняется отклонением от нормального закона итогового распределения ошибок наблюдений и, следовательно, невыполнением необходимых для корректного использования МНК предположений, накладываемых на случайную ошибку.

Далее рассмотрим результаты исследований при фиксированной доле выбросов  $\nu$  и дисперсиях  $\sigma_1^2 = 0.01$  и  $\sigma_2^2 = 0.005$ . Будем изменять значения математического ожидания засорений от 0 до 1. На рис. 2, *a* и *б* представлены результаты вычислительных экспериментов для  $\nu = 0.1$  и  $\nu = 0.15$  соответственно. Можно отметить, при увеличении математического ожидания метод MTL начинает давать более точные результаты оценивания, чем метод FAST-LTS, что связано со способом формирования оценочных подмножеств. Также видно, что при малом смещении засоряющего распределения МНК дает более точные результаты оценивания.

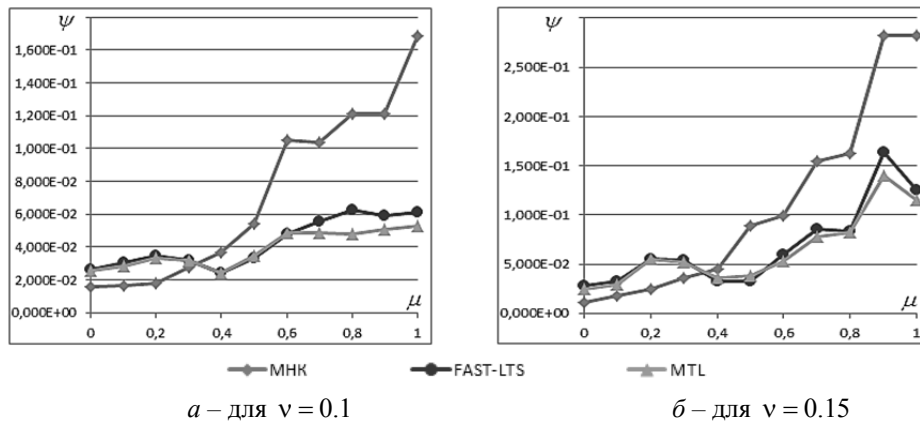


Рис. 2. Точность оценивания в зависимости от математического ожидания  $\mu$  при  $\sigma_1^2 = 0.01$ ,  $\sigma_2^2 = 0.005$ .

Рассмотрим случай, когда грубые ошибки наблюдений имеют смещение и достаточно высокое значение дисперсии по сравнению с основной массой наблюдений,  $\mu = 0.3$ , т. е. выбросы произошли только справа, дисперсия  $\sigma_2^2 = 1$ ,  $\sigma_1^2 = 0.01$ . В табл. 4 приведены результаты этого исследования.

Таблица 4

**Точность оценивания параметров регрессионной модели  $\psi$  при симметричном засорении,  $\sigma_2^2 = 1$ ,  $\mu = 0.3$**

Метод	Доля выбросов, $\nu$		
	0.05	0.1	0.15
МНК	8,055E-03	1,546E-02	2,317E-02
FAST-LTS	1,301E-03	1,598E-03	1,868E-03
MTL	1,304E-03	1,613E-03	1,872E-03

По результатам, приведенным табл. 2, можно отметить, что методы MTL и FAST-LTS показали устойчивые результаты к появлению в выборке такого рода засорений, МНК показал неустойчивые результаты оценивания в данном случае.

## ЗАКЛЮЧЕНИЕ

В данной работе предложен новый алгоритм усеченного максимального правдоподобия. С помощью вычислительных экспериментов показана работоспособность предложенного алгоритма, проведено сравнение точности нахождения оценок параметров регрессионных зависимостей с использованием МНК, алгоритмов метода LTS (FAST-LTS) и MTL.

Результаты вычислительных экспериментов показали, что при унимодальном распределении случайных ошибок и несмещенном засорении с меньшей по сравнению с основной массой наблюдений дисперсией МНК показал более точные результаты оценивания, однако при высоком значении дисперсии засоряющего распределения алгоритмы FAST-LTS и MTL показали более точные результаты.

При не унимодальном распределении случайной ошибки, где выбросы являются смещенными, но обладают меньшей по сравнению с основной массой наблюдений дисперсией, алгоритмы FAST-LTS и MTL показали более устойчивые результаты, чем МНК. Кроме того, стоит отметить, что при высоком значении смещения в распределении выбросов (для исследуемого примера  $\mu \geq 0.6$ ) наиболее точные результаты показал алгоритм MTL.

## СПИСОК ЛИТЕРАТУРЫ

1. Тюрин Ю.Н., Макаров А.А. Статистический анализ данных на компьютере. – М.: Инфра-М, 1998. – 528 с.
2. Наследов А.Д. IBM SPSS Statistics 20 и AMOS: профессиональный статистический анализ данных. – СПб.: Питер, 2013. – 416 с.
3. Халафян А.А. Промышленная статистика: контроль качества, анализ процессов, планирование экспериментов в пакете STATISTICA: учебное пособие. – М.: Либроком, 2013. – 384 с.
4. Dalgaard P. Introductory statistics with R. – 2<sup>nd</sup> ed. – New York: Springer Science Business Media, 2008. – 363 p.
5. Хьюбер П. Робастность в статистике. – М.: Мир, 1984. – 303 с.
6. Тимофеев В.С., Хайленко Е.А. Адаптивное оценивание параметров регрессионных зависимостей при неоднородности случайных ошибок // Доклады Академии наук высшей школы Российской Федерации. – 2014. – № 4 (25). – С. 115–123.
7. Тимофеев В.С. Адаптивное восстановление регрессионных зависимостей на основе полупараметрической оценки плотности случайной компоненты // Научный вестник НГТУ. – 2013. – № 4. – С. 24–30.
8. Тимофеев В.С., Хайленко Е.А. Адаптивное оценивание параметров регрессионных моделей с использованием обобщенного лямбда-распределения // Доклады академии наук высшей школы Российской Федерации. – 2010. – № 2 (15). – С. 25–36.
9. Dovgal S.Yu., Lisitsin D.V. Robust estimation of count response regression models // Applied Methods of Statistical Analysis. Simulations and Statistical Inference, AMSA-2011, Novosibirsk, 20–22 September 2011: proceedings of the international workshop. – Novosibirsk: NSTU Publ., 2011. – P. 318–321.
10. Лисицин Д.В., Гаврилов К.В. Об устойчивом оценивании параметров моделей при асимметричном засорении данных // Научный вестник НГТУ. – 2008. – № 1 (30). – С. 33–40.
11. Денисов В.И., Лисицин Д.В. О свойствах оценок параметров регрессионной модели с эллиптическим распределением и мультипликативной ковариационной структурой ошибок // Сибирский журнал индустриальной математики. – 2003. – № 2. – С. 37–45.
12. Денисов В.И., Лисицин Д.В. Методы построения многофакторных моделей по неоднородным, негауссовским, зависимым наблюдениям: монография. – Новосибирск: Изд-во НГТУ, 2008. – 360 с.



13. Устойчивое оценивание нелинейных структурных зависимостей / В.И. Денисов, А.Ю. Тимофеева, Е.А. Хайленко, О.И. Бузмакова // Сибирский журнал индустриальной математики. – 2013. – № 4. – С. 47–60.

14. Денисов В.И., Тимофеев В.С. Устойчивые распределения и оценивание параметров регрессионных зависимостей // Известия Томского политехнического университета. – 2011. – Т. 318, № 2. – С. 10–15.

15. Мудров В.И., Кушко В.Л. Метод наименьших модулей. – М.: Знание, 1971. – 61 с.

16. Rousseeuw P.J. Least median of squares regression // Journal of the American Statistical Association. – 1984. – N 79. – P. 871–880.

17. Rousseeuw P.J., Driessen K. van. Computing LTS regression for large data sets. Mimeo / University of Antwerpen, Department of Mathematics and Computer Science. – Antwerpen, Belgium, 1999. – 21 p.

*Фаддеенков Андрей Владимирович*, кандидат технических наук, доцент кафедры теоретической и прикладной информатики Новосибирского государственного технического университета. Основное направление научных исследований – разработка и исследование методов и алгоритмов анализа многофакторных объектов со структурированной ошибкой. Имеет более 40 публикаций. E-mail: faddeenkov@corp.nstu.ru

*Хайленко Екатерина Аленксеевна*, кандидат технических наук, доцент кафедры теоретической и прикладной информатики Новосибирского государственного технического университета. Основное направление научных исследований – разработка и исследование методов и алгоритмов оценивания и планирования эксперимента при различных условиях вычислительных экспериментов. Имеет более 20 публикаций. E-mail: hajlenko@corp.nstu.ru

### ***Estimation of regression model parameters by the trimmed maximum likelihood method\****

*A.V. FADDEENKOV<sup>1</sup>, E.A. KHAIENKO<sup>2</sup>*

<sup>1</sup> *Novosibirsk State Technical University, 20 K. Marx Prospekt, Novosibirsk, 630073, Russian Federation, PhD (Eng.), associate professor. E-mail: faddeenkov@corp.nstu.ru*

<sup>2</sup> *Novosibirsk State Technical University, 20 K. Marx Prospekt, Novosibirsk, 630073, Russian Federation, PhD (Eng.), associate professor. E-mail: xajlenko@corp.nstu.ru*

The new method of estimating the parameters of regression equations named a trimmed maximum likelihood method is proposed in this paper. This method is based on combining the maximum likelihood method and the trimmed least squares method. The problem of constructing a linear model of regression analysis is formulated. To do this an algorithm of the trimmed maximum likelihood method is described in detail and the main principles which form the basis of the algorithm are demonstrated on a simple example. The task of robustness analysis of regression model parameter estimates is substantiated. The analysis of the proposed algorithm efficiency is made using computing experiment methods. Simulation of random errors in computing experiments was conducted by Monte-Carlo methods. The distribution of random errors was modeled as a mixture of two distributions. Normal distribution was considered as the base distribution of random errors. The estimation of regression model parameters was conducted using least squares, trimmed least squares and trimmed maximum likelihood methods. A relative error of regression model parameter estimation was used as a precision measure. Such types of contamination, as unbiased with a zero mean, symmetric and asymmetric contaminations were considered. The ratio of outliers varied from 2.5 to 20 percent. A series of 100 experiments whose results were averaged was carried out for each type of original data contami-

---

\* Received 01 July 2016.

nations. The results of computational experiments have shown that the trimmed maximum likelihood method has good robustness to various types of contaminations of original data. In most cases, it is almost as good as the trimmed least squares method, and in some cases significantly exceeds it in accuracy.

**Keywords:** regression model, estimate robustness, trimmed maximum likelihood method, trimmed least squares method, Monte-Carlo method, outliers, mixture of distributions, computing experiment

DOI: 10.17212/1814-1196-2016-4-135-145

## REFERENCES

1. Tyurin Yu.N, Makarov A.A. *Statisticheskii analiz dannykh na komp'yutere* [Statistical analysis of the data on the computer]. Moscow, Infra-M Publ., 1998. 528 p.
2. Nasledov A.D. *IBM SPSS Statistics 20 i AMOS: professional'nyi statisticheskii analiz dannykh* [IBM SPSS Statistics 20 and AMOS: professional statistical analysis]. St. Petersburg, Piter Publ., 2013. 416 p.
3. Khalafyan A.A. *Promyshlennaya statistika: kontrol' kachestva, analiz protsessov, planirovanie eksperimentov v pakete STATISTICA* [Industrial statistics. Quality control, process analysis, design of experiments in the package STATISTICA]. Moscow, Librokom Publ., 2013. 384 p.
4. Dalgaard P. *Introductory statistics with R*. 2<sup>nd</sup> ed. New York, Springer Science Business Media, 2008. 363 p.
5. Huber P.J. *Robust statistics*. New York, Wiley, 1981 (Russ. ed.: Kh'yuber P. *Robastnost' v statistike*. Moscow, Mir Publ., 1984. 303 p.).
6. Timofeev V.S., Khailenko E.A. Adaptivnoe otsenivanie parametrov regressionnykh zavisimostei pri neodnorodnosti sluchainykh oshibok [Adaptive estimation of regression model parameters with error distribution inhomogeneity]. *Doklady Akademii nauk vysshei shkoly Rossiiskoi Federatsii – Proceedings of the Russian higher school Academy of sciences*, 2014, no. 4 (25), pp. 115–123.
7. Timofeev V.S. Adaptivnoe vosstanovlenie regressionnykh zavisimostei na osnove poluparametricheskoi otsenki plotnosti sluchainoi komponenty [Adaptive construction of regression models based on semiparametric estimation of disturbance density function]. *Nauchnyi vestnik Novosibirskogo gosudarstvennogo tekhnicheskogo universiteta – Science bulletin of the Novosibirsk state technical university*, 2013, no. 4, pp. 24–30.
8. Timofeev V.S., Khailenko E.A. Adaptivnoe otsenivanie parametrov regressionnykh modelei s ispol'zovaniem obobshchennogo lyambda-raspredeleniya [Adaptive estimation of regression models parameters using generalized lambda-distribution]. *Doklady Akademii nauk vysshei shkoly Rossiiskoi Federatsii – Proceedings of the Russian higher school Academy of sciences*, 2010, no. 2 (15), pp. 25–36.
9. Dovgal S.Yu., Lisitsin D.V. Robust estimation of count response regression models. *Applied Methods of Statistical Analysis. Simulations and Statistical Inference, AMSA-2011: proceedings of the international workshop*, Novosibirsk, 20–22 September 2011, pp. 318–321.
10. Lisitsin D.V., Gavrilov K.V. Ob ustoychivom otsenivanii parametrov modelei pri asimmetrichnom zasorennii dannykh [On stable estimation of models parameters in presence of asymmetric data contamination]. *Nauchnyi vestnik Novosibirskogo gosudarstvennogo tekhnicheskogo universiteta – Science bulletin of the Novosibirsk state technical university*, 2008, no. 1 (30), pp. 33–40.
11. Denisov V.I., Lisitsin D.V. O svoistvakh otsenok parametrov regressionnoi modeli s ellipticheskim raspredeleniem i mul'tiplikativnoi kovariatsionnoi strukturoi oshibok [About properties of the regression model parameter estimates with elliptic distribution and a multiplicative covariance structure of errors]. *Sibirskii zhurnal industrial'noi matematiki – Journal of Applied and Industrial Mathematics*, 2003, no. 2, pp. 37–45. (In Russian)
12. Denisov V.I., Lisitsin D.V. *Metody postroeniya mnogofaktornykh modelei po neodnorodnym, negaussovskim, zavisimym nablyudeniyam* [Construction methods of multivariate models for heterogeneous, non-Gaussian, dependent observations]. Novosibirsk, NSTU Publ., 2008. 360 p.

13. Denisov V.I., Timofeeva A.Yu., Khailenko E.A., Buzmakova O.I. Ustoichivoe otsenivanie nelineinykh strukturnykh zavisimosti [Sustainable estimation of nonlinear structural dependencies]. *Sibirskii zhurnal industrial'noi matematiki – Journal of Applied and Industrial Mathematics*, 2013, no. 4, pp. 47–60. (In Russian)

14. Denisov V.I., Timofeev V.S. Ustoichivye raspredeleniya i otsenivanie parametrov regressiionnykh zavisimosti [Stable distributions and estimation of regression parameters]. *Izvestiya Tomskogo politekhnicheskogo universiteta – Bulletin of the Tomsk Polytechnic University*, 2011, no. 2, pp. 10–15.

15. Mudrov V.I., Kushko V.L. *Metod naimen'shikh modulei* [The method of least modules]. Moscow, Znaniye Publ., 1971. 61 p.

16. Rousseeuw P.J. Least median of squares regression. *Journal of the American Statistical Association*, 1984, no. 79, pp. 871–880.

17. Rousseeuw P.J., Driessen K. van. *Computing LTS regression for large data sets. Mimeo*. Antwerpen, Belgium, University of Antwerpen, Department of Mathematics and Computer Science, 1999. 21 p.