

УДК 519.242.5

**УСТОЙЧИВОЕ ОЦЕНИВАНИЕ ПАРАМЕТРОВ СТРУКТУРНЫХ  
ЗАВИСИМОСТЕЙ В ЗАДАЧЕ АНАЛИЗА УСПЕВАЕМОСТИ****А.Ю. Тимофеева, Е.А. Хайленко***Новосибирский государственный технический университет*

Рассмотрена задача оценивания параметров полиномиальных структурных зависимостей. Разработаны модификации метода наименьших уравновешенных квадратов (LTS) на основе расстояний Кука, Велша–Куха, Махаланобиса и робастного расстояния для устойчивого оценивания такого рода зависимостей, с помощью вычислительных экспериментов показана работоспособность предложенных модификаций. Проведено исследование этих методов при различных условиях вычислительных экспериментов. Приведены результаты сравнения точности оценивания параметров структурного уравнения, полученных данными модификациями метода LTS, с результатами, полученными скорректированным методом наименьших квадратов (ALS) и разработанным авторами ранее методом LTS на основе метода ALS. Получено, что при появлении в выборке только горизонтальных либо только вертикальных выбросов для получения устойчивых оценок параметров полиномиальных структурных зависимостей целесообразно использовать метод LTS на основе метода ALS; при появлении в выборке как горизонтальных, так и вертикальных выбросов рекомендуемыми методами оценивания являются модификации метода LTS на основе расстояний Махаланобиса и робастного. Все модификации метода LTS применены для анализа успеваемости студентов Новосибирского государственного технического университета на примере факультета механизации и автоматизации и дисциплины «Информатика».

*Ключевые слова:* структурная зависимость, метод наименьших уравновешенных квадратов, метод скорректированных квадратов, расстояние Кука, расстояние Велша–Куха, расстояние Махаланобиса, робастное расстояние.

DOI: 10.17212/1727-2769-2015-2-117-127

**Введение**

В последнее время особое внимание начинает уделяться проблемам оценки качества образовательных услуг. Наряду с формальными показателями работы высших учебных заведений ключевой характеристикой здесь должна являться удовлетворенность основных потребителей – студентов. Одним из индикаторов, тесно связанным с удовлетворенностью, можно считать степень востребованности выпускников [1]. Однако такой индикатор отражает только конечный результат обучения и не позволяет судить об отношении к самому образовательному процессу, хотя именно это было бы наиболее ценной информацией для принятия управленческих решений руководством вуза о корректировке учебных планов, рабочих программ и других аспектов образовательной деятельности.

Некоторая обратная связь в ходе обучения все же присутствует, она выражается в основном эпизодическими жалобами неуспевающих студентов. Иногда проводятся анкетирования студентов на предмет их удовлетворенности обучением конкретным дисциплинам. Однако вся эта информация носит субъективный характер, и на ее основе сложно сделать какие-то выводы и рекомендации для

---

Работа выполнена при финансовой поддержке РФФИ в рамках научного проекта № 14-07-31171 мол\_а.

© 2015 А.Ю. Тимофеева, Е.А. Хайленко

управления образовательным процессом. В целом, есть потребность в более объективном инструменте оценки и сопоставления преподаваемых дисциплин с точки зрения их сложности для студентов.

### 1. Методы статистического анализа успеваемости

Согласно классической теории тестирования [2], фактически наблюдаемая успешность  $x_i$  выполнения заданий  $i$ -м студентом складывается из некоторого истинного ненаблюдаемого уровня успеваемости  $X_i$  и случайной погрешности измерения  $\delta_i$ . Относительно случайной ошибки предполагается равенство нулю математического ожидания, конечность дисперсии. Стандартным также является предположение об отсутствии корреляции ошибки  $\delta_i$  с истинным значением  $X_i$ . Однако многие исследователи [3] не признают его обоснованность в силу того, что, например, успеваемость более сильных студентов может быть измерена более точно, если тест достаточно сложный. В то же время, в соответствии с учебной практикой, слабые студенты тоже достаточно четко выявляются, они обычно не успевают по многим дисциплинам. Следовательно, далее будет предполагаться существование зависимости между истинным уровнем успеваемости  $X_i$  и дисперсией ошибки  $\sigma_{\delta_i}^2$  его измерения. При этом описывающая такую зависимость гладкая функция  $\sigma_{\delta_i}^2(X_i)$  должна убывать в направлении как низких, так и высоких значений  $X_i$ , т. е. иметь максимум при среднем уровне успеваемости.

Успешность освоения  $Y_i$  отдельной дисциплины  $i$ -м студентом в общем случае является некоторой нелинейной функцией от его истинного уровня успеваемости. Степень нелинейности этой функции зависит от используемой преподавателем шкалы для перевода знаний и навыков студентов в оценки. Для иллюстрации и удобства интерпретации такую функцию берут обычно степенной [4]:

$$Y_i = X_i^a, \quad i = \overline{1, n}, \quad (1)$$

где  $X_i \in [0, 1]$ ,  $a > 0$ ,  $n$  – число студентов. При  $a = 1$  получаем линейную зависимость, и оценка прямо пропорциональна успеваемости студентов. В случае если  $a < 1$ , зависимости называются сублинейными, и дифференцирующая способность используемой преподавателем шкалы выше для слабых студентов. При  $a > 1$  зависимости  $Y_i$  от  $X_i$  будут надлинейными, они повышают дифференцирующую способность шкалы для сильных студентов.

По результатам сессии истинный уровень успешности освоения  $Y_i$  отдельной дисциплины измеряется с некоторой погрешностью  $\varepsilon_i$ :

$$y_i = Y_i + \varepsilon_i, \quad i = \overline{1, n}.$$

Предположения относительно данной погрешности аналогичны введенным выше предположениям о случайной ошибке измерения уровня успеваемости.

Для представления нелинейной модели (1) предлагается использовать полиномиальную аппроксимацию:

$$y_i = \theta_1 + \theta_2(x_i - \delta_i) + \theta_3(x_i - \delta_i)^2 + \dots + \theta_m(x_i - \delta_i)^{m-1} + \varepsilon_i, \quad i = \overline{1, n}. \quad (2)$$

В силу наличия ошибок в объясняющей переменной зависимость (2) носит название структурной [5]. Задача состоит в оценивании вектора неизвестных параметров  $\theta$ .

В качестве  $y_i$  предлагается рассматривать оценки по отдельным сессиям. Всеобщее внедрение 100-балльной шкалы оценок и балльно-рейтинговой системы позволяет более четко дифференцировать уровень освоения предметов по сравнению со стандартными оценками от «неудовлетворительно» до «отлично». Однако возникает проблема определения индикатора  $x_i$  истинного уровня успеваемости студентов для сопоставления его с результатами сессии по каждой дисциплине. В качестве такого индикатора предлагается использовать суммарный (средний) балл за сессию.

Согласно рассмотренной постановке использование такого индикатора в моделях вида (2) сопряжено с проблемой наличия погрешностей в уровне успеваемости. Для анализа такого рода структурных зависимостей применяют специальные методы и алгоритмы, активно развиваемые авторами [6]. В данной работе предлагается ряд модификаций методов, направленных на повышение устойчивости результатов к наличию грубых ошибок измерения.

## 2. Модификации метода LTS для оценивания структурных зависимостей

Для поиска оценок параметров полиномиальных зависимостей вида (2) в [7] предлагается скорректированный метод наименьших квадратов (ALS) [7], который основан на использовании момента  $r$ -порядка величины  $X$ . В  $i$ -й точке наблюдений его значение  $t_i^r$  можно представить в виде

$$t_i^r = E(x_i - \delta_i)^r.$$

Суть данного метода состоит в следующем:

- вычисляются оценки  $t^r$  исходя из рекуррентного соотношения

$$\hat{t}_i^{r+1} = x_i \hat{t}_i^r - r \sigma_\delta^2 \hat{t}_i^{r-1}, \quad \hat{t}_i^{-1} = \hat{t}_i^0 = 1;$$

- оценивается вектор

$$\hat{R}_j = \sum_{i=1}^n \hat{t}_i^j y_i, \quad j = \overline{1, m};$$

- находится оценка матрицы

$$\hat{P}_{rs} = \sum_{i=1}^n \hat{t}_i^{r+s}, \quad r, s = \overline{1, m};$$

- вычисляется вектор неизвестных параметров

$$\hat{\theta}_{ALS} = \hat{P}^{-1} \hat{R}.$$

Как известно [6, 7], данный метод не является устойчивым к появлению грубых ошибок наблюдений (выбросов), поэтому в данной работе предлагается использовать робастный метод LTS, который основан на многократном построении оценочных подмножеств размерности  $h$ ,  $h \in \left[ \frac{n+m+1}{2}; n \right]$  и получении оценок методом наименьших квадратов параметров регрессионных моделей по наблюде-

ниям из этих подмножеств [8]. В [9] разработаны и исследованы алгоритмы модификаций метода на основе расстояний Кука, Велша–Кука, Махаланобиса и робастного расстояния [10,11] для оценивания параметров регрессионных моделей. Использование этих расстояний при формировании оценочных подмножеств позволяет учитывать точность и разброс наблюдений, что дает возможность повысить качество оценивания в ряде случаев.

Однако метод LTS и его модификации не учитывают погрешности во входных факторах, поэтому в данной работе предлагается расширить рассмотренные методы для применения их в оценивании параметров полиномиальных структурных зависимостей. Для этого на каждой итерации метода LTS и его модификаций используется метод ALS для нахождения начального приближения и оценивания параметров по оценочным подмножествам.

### 3. Модельный пример

Для проверки работоспособности предложенных модификаций метода LTS для устойчивого оценивания был проведен ряд вычислительных экспериментов. В качестве исследуемой использовалась следующая модель:

$$\eta = -\xi^2 + 2\xi, \quad (3)$$

где  $\xi$  подчиняется стандартному бета-распределению с параметрами  $\alpha_\xi = 3$ ,  $\beta_\xi = 3$ . Таким образом, моделировалась сублинейная зависимость результатов освоения дисциплины от уровня успеваемости, позволяющая в большей мере дифференцировать более слабых студентов.

Наблюдаемые значения объясняющей переменной моделировались как независимые случайные величины  $x_i$ ,  $i = \overline{1,500}$ , подчиненные бета-распределению с параметрами, определяемыми реализациями  $X_i$  случайной величины  $\xi$ , и заданным уровнем дисперсии  $\sigma_{x_i}^2$ .

$$\sigma_{x_i}^2 = 0,05\delta_i,$$

где  $\delta_i$  принимает значения 1, 5 с вероятностями  $(1-\mu_x)$  и  $\mu_x$  ( $\mu_x$  – доля выбросов в объясняющих переменных),

$$\alpha_{x_i} = X_i \left( \frac{1}{\sigma_{x_i}^2} - 1 \right), \quad \beta_{x_i} = \frac{1-X_i}{X_i} \alpha_{x_i}.$$

Параметры подобраны так, чтобы  $E(x_i) = X_i$ ,  $D(x_i) = 0,05 \cdot X_i(1-X_i)\delta_i$ . Следовательно, наибольший уровень погрешности достигается при  $X_i = 0,5$ , а наименьший при крайних значениях истинного уровня успеваемости.

Наблюдаемые значения отклика моделировались как независимые случайные величины  $y_i$ ,  $i = \overline{1,500}$ , подчиненные бета-распределению с параметрами, определяемыми реализациями  $Y_i$  случайной величины  $\eta$ , и заданным уровнем дисперсии  $\sigma_{y_i}^2$ .

$$\sigma_{y_i}^2 = 0,05\epsilon_i,$$

где  $\varepsilon_i$  принимает значения 1, 5 с вероятностями  $(1-\mu_y)$  и  $\mu_y$  ( $\mu_y$  – доля выбросов в отклике),

$$\alpha_{y_i} = Y_i \left( \frac{1}{\sigma_{y_i}^2} - 1 \right), \quad \beta_{y_i} = \frac{1-Y_i}{Y_i} a_{y_i}.$$

Параметры подобраны так, чтобы  $E(y_i) = Y_i$ ,  $D(y_i) = 0,05Y_i(1-Y_i)$ .

В качестве показателя точности оценивания параметров структурных зависимостей был взят показатель *MAD* (Mean Absolute Derivation – среднее абсолютное отклонение), который вычисляется по формуле [12]:

$$MAD = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|,$$

где  $n$  – количество наблюдений, в данном случае  $n = 500$ ;  $\hat{y}$  – оцененный вектор отклика.

Проводилось по 100 вычислительных экспериментов, каждый из которых заключался в моделировании выборки исходных данных, а также ошибок в объясняющих наблюдениях и в отклике, с последующим оцениванием параметров модели (3) с использованием модификаций метода LTS для оценивания структурных зависимостей. В качестве итоговых показателей точности оценивания *MAD* использовалось усредненное значение.

#### 4. Результаты вычислительных экспериментов

Для удобства представления результатов вычислительных экспериментов введем обозначения используемых модификаций метода LTS для оценки структурных зависимостей: LTS с использованием метода ALS – ALTS; на основе расстояния Кука – ALTS\_CD; на основе расстояния Велша–Куха – ALTS\_WKD, на основе расстояния Махаланобиса – ALTS\_MD, на основе робастного расстояния – ALTS\_RD.

Рассмотрим случай появления в выборке только вертикальных выбросов, т. е. доля выбросов в объясняющих переменных  $\mu_x = 0$ , доля выбросов в отклике  $\mu_y$  варьировалась от 0 до 0,2. На рис. 1 представлены результаты вычислительных экспериментов.

Из рис. 1 видно, что наименее точные результаты оценивания показал метод ALS, что объясняется неустойчивостью метода к появлению в выборке грубых ошибок наблюдений. Также недостаточно точные результаты дал метод LTS для оценивания параметров структурных зависимостей на основе расстояния Кука. Причиной этого может являться способ вычисления расстояния Кука [11], поскольку при появлении ошибок в объясняющих переменных наблюдения, отстоящие от основной массы, но не являющиеся выбросом, могут не учитываться в процессе формирования оценочного подмножества. Наиболее точные результаты показали методы ALTS и модификации метода LTS на основе расстояний Махаланобиса и робастного. Такое поведение этих модификаций объясняется особенностью алгоритма формирования оценочного подмножества, где в него добавляются наблюдения с учетом их удаленности от основной массы и величины остатков.

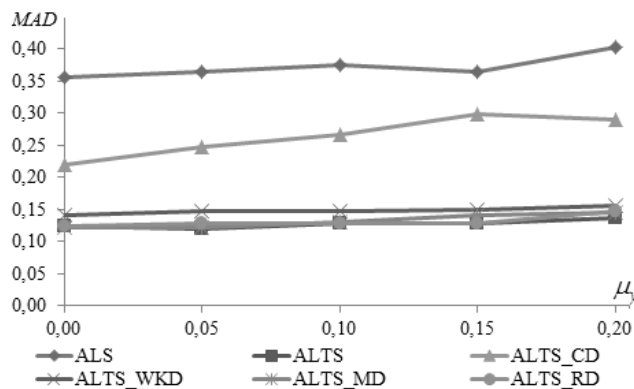


Рис. 1 – Зависимость точности оценивания от доли выбросов  $\mu_y$

Fig. 1 – Relation between estimation accuracy and proportion of outliers  $\mu_y$

Теперь рассмотрим случай появления только горизонтальных выбросов, т. е.  $\mu_y = 0$ , а доля выбросов в объясняющих переменных варьировалась от 0 до 0,1. В табл. 1 представлены результаты оценивания.

Таблица 1 / Table 1

**Точность оценивания параметров структурных зависимостей при наличии горизонтальных выбросов**  
**The estimation accuracy of structural model parameters, when horizontal outliers exist**

$\mu_x$	Параметры	Метод					
		ALS	ALTS	ALTS_CD	ALTS_WKD	ALTS_MD	ALTS_RD
0,00	$\hat{\theta}_1$	0,903	-0,077	0,379	-0,044	-0,035	0,039
	$\hat{\theta}_2$	0,055	2,406	1,126	2,478	2,229	1,992
	$\hat{\theta}_3$	-1,04	-1,375	-0,853	-1,740	-1,225	-1,047
	MAD	0,355	0,124	0,219	0,141	0,122	0,125
0,05	$\hat{\theta}_1$	0,825	-0,011	0,719	0,081	-0,029	0,028
	$\hat{\theta}_2$	0,301	2,183	0,061	1,975	2,278	2,051
	$\hat{\theta}_3$	-1,24	-1,182	-0,263	-1,246	-1,340	-1,129
	MAD	0,343	0,129	0,289	0,149	0,133	0,127
0,10	$\hat{\theta}_1$	0,713	-0,022	0,718	0,079	0,048	0,037
	$\hat{\theta}_2$	0,668	2,275	0,122	1,968	1,988	2,075
	$\hat{\theta}_3$	-1,57	-1,334	-0,365	-1,199	-1,071	-1,195
	MAD	0,340	0,130	0,292	0,162	0,139	0,135

Из табл. 1 видно, что при появлении горизонтальных выбросов методы ALS и LTS на основе расстояния Кука так же, как и в предыдущем случае, показали наименее точные результаты по сравнению с другими рассматриваемыми разновидностями метода LTS. Наиболее точные результаты оценивания показали методы LTS на основе расстояний Махаланобиса и робастного, а также метод ALTS.

Поскольку при появлении как только горизонтальных, так и только вертикальных выбросов методы ALS и модификация метода LTS для оценки структурных зависимостей на основе расстояния Кука показали наименее точные результаты оценивания, то в дальнейших исследованиях будем рассматривать только методы ALTS и модификации метода LTS на основе расстояний Велша–Кука, Махаланобиса и робастного.

Рассмотрим случай появления в выборке и горизонтальных, и вертикальных выбросов. На рис. 2, а, б представлены результаты оценивания, по которым видно, что при доле выбросов 5 % ( $\mu_x = 0,05$ ) методы ALTS и LTS на основе расстояний Махаланобиса и робастного показали близкие результаты оценивания, однако первый немного превосходит по точности. При увеличении доли горизонтальных выбросов до 10 % ( $\mu_x = 0,1$ ) оценки, получаемые модификациями метода LTS на основе расстояний Махаланобиса и робастного, являются более точными.

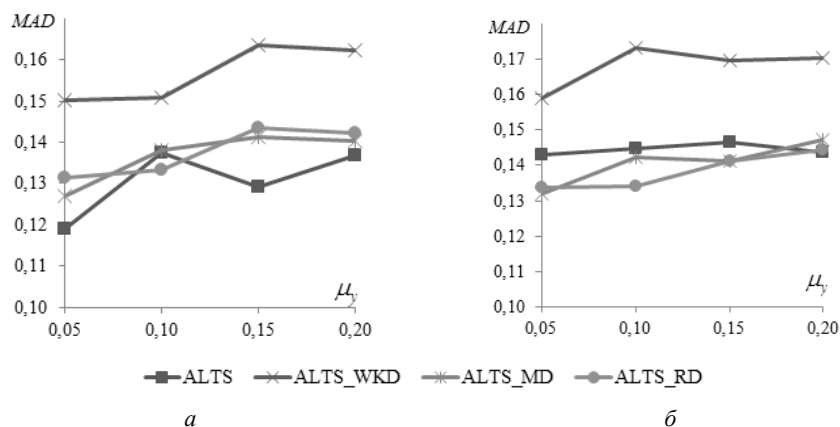


Рис. 2 – Зависимость точности оценивания от доли выбросов  $\mu_y$ :

а –  $\mu_x = 0,05$ ; б –  $\mu_x = 0,1$

Fig. 2 – Relation between estimation accuracy and proportion of outliers  $\mu_y$ :

а –  $\mu_x = 0,05$ ; б –  $\mu_x = 0,1$

По результатам проведенных исследований можно дать следующие рекомендации: при наличии в выборке только горизонтальных либо только вертикальных выбросов для устойчивого оценивания параметров структурных зависимостей лучше использовать метод ALTS; при наличии в выборке как вертикальных, так и горизонтальных выбросов – модификации метода LTS на основе расстояний Махаланобиса и робастного.

## 5. Практическое приложение

Для демонстрации применимости предложенных методов для практического анализа образовательного процесса проведено исследование успеваемости студентов Новосибирского государственного технического университета.

В качестве примера взят один из факультетов (механизации и автоматизации) и дисциплина «Информатика». Объясняющей переменной выступал средний балл студента в первую сессию. Рассматривались студенты 2011, 2012 и 2013 годов поступления. Масштаб переменных приведен к отрезку  $[0,1]$ . Результаты оценивания квадратичной зависимости представлены в табл. 2.

Таблица 2 / Table 2

Результаты оценивания квадратичной модели студенческой успеваемости  
The estimation results of the quadratic model of student achievement

Год поступления	Метод	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	MAD
2011	ALTS	-0,026	1,372	-0,308	0,119
	ALTS CD	-0,325	2,531	-1,343	0,126
	ALTS WKD	-0,506	2,377	-0,819	0,133
	ALTS MD	-0,025	1,369	-0,306	0,119
	ALTS RD	-0,140	1,727	-0,570	0,119
2012	ALTS	-0,953	3,790	-1,853	0,114
	ALTS CD	-0,987	3,910	-2,017	0,112
	ALTS WKD	-0,823	3,232	-1,378	0,114
	ALTS MD	-0,931	3,645	-1,716	0,114
	ALTS RD	-0,805	3,396	-1,599	0,112
2013	ALTS	-0,808	3,156	-1,238	0,109
	ALTS CD	-1,650	5,327	-2,695	0,112
	ALTS WKD	-2,163	7,042	-4,034	0,118
	ALTS MD	-0,661	3,069	-1,420	0,100
	ALTS RD	-1,528	5,268	-2,813	0,105

Для удобства интерпретации различий в освоении дисциплины по годам графически представлены расчетные кривые, полученные с помощью модификации на основе расстояний Махаланобиса (рис. 3).

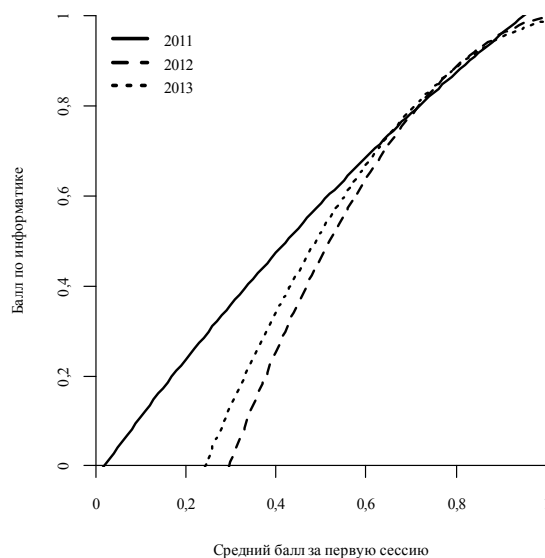


Рис. 3. Расчетные значения успешности освоения информатики в зависимости от средней успеваемости

Fig. 3 – The calculated values of the successful development of computer science depending on the average of students progress



Из рис. 3 видно, что поток студентов, поступивших в 2011 году, в целом более существенно отличается от других рассмотренных потоков. Зависимость успешности освоения дисциплины от средней успеваемости очень близка к линейной. Это свидетельствует о том, что в ходе аттестации установлена шкала оценок, прямо пропорциональная истинному уровню успеваемости студентов. Иначе обстоит дело со студентами более поздних лет поступления. Зависимости здесь носят явный сублинейный характер, поэтому по итогам сессии по дисциплине «Информатика» удастся больше дифференцировать более слабых студентов, чем более сильных. Возможно, это связано с падением общего уровня знаний поступающих в вузы в последние годы. В результате преподаватели вынуждены пересматривать систему оценивания с тем, чтобы дифференцировать слабых студентов в ущерб более сильным.

### Заключение

В статье рассмотрена задача оценивания параметров полиномиальных структурных зависимостей. Предложены и исследованы устойчивые модификации метода LTS для оценки такого рода зависимостей на основе расстояний Кука, Велша–Куха, Махаланобиса и робастного расстояния. С помощью вычислительных экспериментов подтверждена работоспособность предложенных методов. На основе полученных результатов исследований даны рекомендации о применении модификаций метода LTS. Проведено исследование успеваемости студентов Новосибирского государственного технического университета на примере факультета механизации и автоматизации и дисциплины «Информатика».

### ЛИТЕРАТУРА

1. **Борисова А.А., Тимофеева А.Ю.** Выпускники вузов на рынке труда: индикаторы мониторинга и ограничители профильной занятости // Университетское управление: практика и анализ. – 2014. – № 1 (89). – С. 71–80.
2. **Crocker L., Algina J.** Introduction to classical and modern test theory. – New York: Harcourt Brace Jovanovich College Publishers, 1986. – 527 p.
3. **Hambleton R.K., Jones R.W.** An NCME instructional module on comparison of classical test theory and item response theory and their applications to test development // Educational Measurement: Issues and Practice. – 1993. – Vol. 12, iss. 3. – P. 38–47. – doi: 10.1111/j.1745-3992.1993.tb00543.x.
4. **Ким В.С.** Тестирование учебных достижений: монография. – Уссурийск: Изд-во УГПИ, 2007. – 214 с.
5. **Айвазян С.А., Енюков И.С., Мешалкин Л.Д.** Прикладная статистика. Исследование зависимостей. – М.: Финансы и статистика, 1985. – 488 с.
6. Устойчивое оценивание нелинейных структурных зависимостей / В.И. Денисов, А.Ю. Тимофеева, Е.А. Хайленко, О.И. Бузмакова // Сибирский журнал индустриальной математики. – 2013. – Т. 16, № 4 (56). – С. 47–60.
7. **Cheng C.-L., Schneeweiss H.** Polynomial regression with errors in the variables // Journal of the Royal Statistical Society: Series B. – 1998. – Vol. 60, iss. 1. – P. 189–199. – doi: 10.1111/1467-9868.00118.
8. **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.
9. **Хайленко Е.А.** Модификации метода LTS для устойчивого оценивания параметров регрессионных моделей // Сборник научных трудов НГТУ. – 2011. – N 1 (63). – С. 75–82.
10. **Chatterjee S., Hadi A.S.** Influential observations, high leverage points, and outliers in linear regression // Statistical Science. – 1986. – Vol. 1, N 3. – P. 379–416. – doi: 10.1214/ss/1177013627.

11. **Radhakrishna R.C., Toutenburg H.** Linear models: least squares and alternatives. – 2<sup>nd</sup> ed. – New York: Springer-Verlag, 1999. – 428 p.
12. Робастность в статистике: подход на основе функций влияния / Ф. Хампель, Э. Ронchetti, П. Рауссеу, В. Штаэль. – М.: Мир, 1989. – 512 с.

## ROBUST ESTIMATION OF STRUCTURAL MODEL PARAMETERS IN APPLICATION TO STUDENTS' PROGRESS ANALYSIS

**Timofeeva A.Yu., Khailenko E.A.**

*Novosibirsk state technical university, Novosibirsk, Russian Federation*

The problem of estimation of structural model parameters is considered. Modifications of Least Trimmed Squares (LTS) method using Mahalanobis, robust, Cook and Welsch-Kuh distances have been proposed. The efficiency of the proposed modifications was showed using computing experiment. Results of these methods investigation in different conditions of computing experiments were conducted. Results of compare accuracy between these modifications of LTS method, the Adjusted Least Squares method (ALS) and modification of LTS method using ALS method, which was proposed by authors previously, were discussed. Recommendations to application proposed method are following: in case there are only vertical or only horizontal outliers in sample of data better to use modification of LTS method based on ALS method for estimating structural model parameters; when there are both types of outliers in sample – modification of LTS method based on Mahalanobis and robust distances. These modifications of LTS method were applied to problem of students' progress analysis.

*Keywords:* structural relation, the least trimmed squares method, the adjusted least squares method, Mahalanobis distance, robust distance, Cook distance, Welsch-Kuh distance.

DOI: 10.17212/1727-2769-2015-2-117-127

## REFERENCES

1. Borisova A.A., Timofeeva A.Yu. Vypuskniki vuzov na rynke truda: indikatory monitoringa i ogranichiteli profil'noi zanyatosti [Graduates in the labor market: monitoring indicators and limiters of profiled employment]. *Universitetskoe upravlenie: praktika i analiz – Journal University Management: Practice and Analysis*, 2014, no. 1 (89), pp. 71–80.
2. Crocker L., Algina J. *Introduction to classical and modern test theory*. New York, Harcourt Brace Jovanovich College Publishers, 1986. 527 p.
3. Hambleton R.K., Jones R.W. An NCME instructional module on comparison of classical test theory and item response theory and their applications to test development. *Educational Measurement: Issues and Practice*, 1993, vol. 12, iss. 3, pp. 38–47. doi: 10.1111/j.1745-3992.1993.tb00543.x
4. Kim V.S. *Testirovanie uchebnykh dostizhenii*. [Testing of educational achievements]. Usturiisk, UGPI Publ., 2007. 214 p.
5. Aivazyan S.A., Yenyukov I.S., Meshalkin L.D. *Prikladnaya statistika. Issledovanie zavisimosti* [Applied statistics: study of relationships]. Moscow, Finansy i statistika Publ., 1985. 488 p.
6. Denisov V.I., Timofeeva A.Yu., Khailenko E.A., Buzmakova O.I. Ustoichivoe otsenivanie nelineynykh strukturnykh zavisimosti [Robust estimation of nonlinear structural models]. *Sibirskii zhurnal industrial'noi matematiki – Journal of Applied and Industrial Mathematics*, 2014, vol. 8, iss. 1, pp. 28–39. doi: 10.1134/S1990478914010049. Translated from *Sibirskii zhurnal industrial'noi matematiki*, 2013, vol. 16, no. 4 (56), pp. 47–60.
7. Cheng C.-L., Schneeweiss H. Polynomial regression with errors in the variables. *Journal of the Royal Statistical Society: Series B*, 1998, vol. 60, iss. 1, pp. 189–199. doi: 10.1111/1467-9868.00118
8. 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.

9. Hailenko E.A. Modifikatsii metoda LTS dlya ustoichivogo otsenivaniya parametrov regressionnykh modelei [Modifications of method LTS for robust estimation parameters of regression models]. *Sbornik nauchnykh trudov Novosibirskogo gosudarstvennogo tekhnicheskogo universiteta – Transaction of scientific papers of the Novosibirsk state technical university*, 2011, no. 1 (63), pp. 75–82.
10. Chatterjee S., Hadi A.S. Influential observations, high leverage points, and outliers in linear regression. *Statistical Science*, 1986, vol. 1, no. 3, pp. 379–416. doi: 10.1214/ss/1177013627
11. Radhakrishna R.C., Toutenburg H. *Linear models: least squares and alternatives*. 2<sup>nd</sup> ed. New York, Springer-Verlag, 1999. 428 p.
12. Hampel F.R., Ronchetti E.M., Rousseeuw P.J., Stahel W.A. *Robust statistics. Approach based on the influence function*. New York, John Wiley and Sons, 1986. XXI. 502 p. (Russ. ed.: Khampel' F., Ronchetti E., Rausseu P., Shtael' V. *Robastnost' v statistike: podkhod na osnove funktsii vliyaniya*. Moscow, Mir Publ., 1989. 512 p.).

## СВЕДЕНИЯ ОБ АВТОРАХ



**Тимофеева Анастасия Юрьевна** – родилась в 1984 году, канд. экон. наук, доцент, кафедра экономической информатики, НГТУ. Область научных интересов: развитие методов статистического анализа объектов стохастической природы, в том числе социально-экономических явлений. Опубликовано 35 научных работ. (Адрес: 630073, Россия, г. Новосибирск, пр. Карла Маркса, д. 20. Email: a.timofeeva@corp.nstu.ru).

**Timofeeva Anastasia Yurievna** (b. 1984) – Candidate of Science (PhD) in Economics, associate professor of Computer Science in Economics Department of the Novosibirsk State Technical University. Her research interests are currently focused on the development of methods for the statistical analysis of stochastic objects, including socioeconomic phenomena. She is author of 35 scientific papers. (Address: 20, Karl Marx Av., Novosibirsk, 630073, Russia. Email: a.timofeeva@corp.nstu.ru).



**Хайленко Екатерина Алексеевна** – родилась в 1985 году, канд. техн. наук, научный сотрудник, кафедра теоретической и прикладной информатики, НГТУ. Область научных интересов: разработка и исследование алгоритмов устойчивого и адаптивного оценивания параметров регрессионных зависимостей и планирование эксперимента. Опубликовано 18 научных работ. (Адрес: 630073, Россия, Новосибирск, проспект Карла Маркса, 20. Email: xajlenko@corp.nstu.ru).

**Khailenko Ekaterina Alekseevna** (b. 1985) – Candidate of Technical Sciences, research associate, Department of Theoretical and Applied Informatics, NSTU. Her research interests are currently focused on developing and investigating algorithms of robust and adaptive estimation parameters of regression models and design of experiment. He is author of 18 scientific papers. (Address: 20, Karl Marx Av., Novosibirsk, 630073, Russia. Email: xajlenko@corp.nstu.ru).

Статья поступила 5 марта 2015 г.  
Received March 5, 2015

## To Reference:

Timofeeva A.Yu., Khailenko E.A. Ustoichivoe otsenivanie parametrov strukturnykh zavisimostei v zadache analiza uspevaemosti [Robust estimation of structural model parameters in application to students' progress analysis]. *Doklady Akademii nauk vysshei shkoly Rossiiskoi Federatsii – Proceedings of the Russian higher school Academy of sciences*, 2015, no. 2 (27), pp. 117–127. doi: 10.17212/1727-2769-2015-2-117-127