УДК 519.23

Коваленко И. И.,
д-р техн. наук, профессор кафедры програмного обеспечения автоматизированных систем, Национальный университет кораблестроения имени адмирала Макарова, г. Николаев, Украина; igor.kovalenko@nuos.edu.ua
Швед А. В.,
канд. техн. наук, ст. преподаватель кафедры интеллектуальных информационных систем, Черноморский государственный университет имени Петра Могилы, г. Николаев, Украина helenashv@mail.ru

ФОРМАЛИЗОВАННЫЙ ПОДХОД К ВЫБОРУ ОЦЕНОК ТИПА «СРЕДНЕЕ» В АНАЛИЗЕ МАЛЫХ НЕОДНОРОДНЫХ ВЫБОРОК ДАННЫХ

Выборки реальных данных в среднем содержат 5–10 % наблюдений, в некоторой степени выделяющиеся из исходной совокупности, вследствие чего выборка становится неоднородной. Источником выделяющихся наблюдений может служить появление больших ошибок, округление и группировка. Наличие подобных отклонений ведет к искажению строгих параметрических моделей (например, нормального закона распределения), что характеризуется появлением, так называемых, «тяжелых хвостов», ярко выраженными асимметрией и эксцессом. Применение в этих условиях общепринятых процедур оценивания, например, выборочного среднего, влечет за собой увеличение смещенности и снижение эффективности получаемых оценок.

В статье предложен подход к обработке малых неоднородных совокупностей наблюдений и нахождения их усредненных оценок на основе устойчивых статистических процедур. Для оценивания параметра положения искаженных параметрических моделей распределений рассмотрен класс робастных оценок, представленных линейными комбинациями порядковых статистик.

Ключевые слова: *адаптивные робастные оценки; усеченное среднее; «хвосты» распределения; симметричное и несимметричное усечение.*

Постановка проблемы

При обработке реальных данных аналитики нередко сталкиваются с ситуацией, при которой в исследуемых выборках данных имеются наблюдения, в той или иной мере, выделяющиеся из исходной совокупности. Такие наблюдения получили название «резко выделяющиеся», «подозрительные», «засоряющие» и др., их доля обычно составляет 5 %–10 % [4; 5; 7; 8].

Появление неоднородных данных в выборках является причиной искажения строгих параметрических моделей, которые описываются известными законами распределения вероятностей (например, нормальные распределения, распределения Пуассона, распределения Стьюдента и др.), и ведет к появлению так называемых «хвостов», «горбов», характеризуется ярко выраженными асимметрией и эксцессом [5; 7; 8]. В такой ситуации применение стандартных статистических методов обработки данных, в основе которых лежит явное или неявное предположение о нормальности, может привести к существенному искажению полученных результатов.

Для обработки массивов данных, содержащих неоднородные наблюдения применяют следующие подходы [4]:

1) выявление и исключение резко выделяющихся наблюдений;
2) применение методов робастной статистики для их дальнейшего статистического анализа.

Робастная статистика является обобщением классической параметрической статистики, отражая тот факт, что параметрические модели служат лишь приближенному описанию реальности. Предметом изучения робастной статистики является изучение и разработка устойчивых процедур статистического оценивания не только на самих строго параметрических моделях, но и на меньших или больших окрестностях таких моделей.

Анализ публикаций и последних достижений

Современные исследования показали, что процедура проверки анализируемых выборок данных на соответствие нормальному распределению является достаточно сложной задачей, особенно, когда это касается выборок с ограниченным объемом данных ($n \leq 50$). В настоящее время существует достаточно обширный класс критериев согласия [1; 3]; для проверки гипотезы нормальности малых выборок рекомендуется использовать такие критерии как непараметрический критерий Шапиро-Уилка, мощность которого исследовалась в работе [15], критерий Саркади и его модифи-

кации [2]. При этом доказано, что в условиях малых выборок не всегда удается отличить нормальное распределение от других видов распределений [6].

В связи с этим в работе предлагается исследовать отклонение неизвестной функции распределения анализируемой совокупности данных от известного распределения (например, нормального). Отклонения исследуемого распределения от нормального может свидетельствовать о наличии неоднородных наблюдений в общей совокупности данных. И как следствие, невозможности применения для их усреднения выборочного среднего.

Постановка задачи

Целью работы является рассмотрение и обоснование использования ряда усеченных оценок типа «среднее» для оценивания параметра положения искаженных параметрических моделей распределений.

Изложение основного материала

Пусть сформирована некоторая совокупность измерительных данных $X = (x_1, x_2, \ldots, x_n)$ с неизвестной функцией плотности распределения вероятностей $F(x)$. При этом предполагается, что совокупность X рассматривается как множество взаимно независимых одинаково нормально распределенных величин с неизвестными параметрами (a, σ).

Если случайная величина ξ подчиняется нормальному распределению с параметрами (a, σ), где $|a| < \propto$ – параметр положения; $\sigma > 0$ – параметр масштаба, то справедливо следующее утверждение [1]:

$$\delta = \frac{\mathrm{M}|\xi - a|}{\sigma} = \sqrt{2/\pi} = 0{,}79788; \quad \gamma_1 = \frac{\mathrm{M}(\xi - a)^3}{\sigma^{3/2}} = 0; \quad \beta_2 = \frac{\mathrm{M}(\xi - a)^4}{\sigma^4} = 3, \quad (1)$$

где δ – нормированное среднее абсолютное отклонение; γ_1 – коэффициент асимметрии; $(\beta_2 - 3)$ – коэффициент эксцесса.

Для оценки величин δ, γ_1 и β_2 воспользуемся следующими отношениями:

$$d = \frac{1}{ns}\sum_{i=1}^{n}\left|x_i - \bar{x}\right|, \quad As = \frac{1}{ns^3}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^3, \quad Ex = \frac{1}{ns^4}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^4, \quad (2)$$

В приведенных выражениях использованы следующие обозначения: d – выборочное среднее абсолютное отклонение; As – выборочный коэффициент асимметрии; $(Ex - 3)$ – выборочный коэффициент эксцесса; n – объем оцениваемой выборки данных; $\bar{x}$ – выборочное среднее и s^2 – выборочная дисперсия, которые рассчитывается соответственно:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad s^2 = \frac{1}{n}\sum_{i-1}^{n}\left(x_i - \bar{x}\right)^2. \quad (3)$$

Если выборочные оценки рассчитанных моментных отношений отличаются от соответствующих теоретических значений (1), можно считать что неизвестное теоретическое распределение отлично от нормального.

В свою очередь обратное утверждение не является истинным, т. к. близость теоретических и выборочных значений не может свидетельствовать о нормальности теоретического распределения. Можно утверждать лишь о том, что имеет место распределение при котором выборочные оценки d, As и Ex совпадают со значениями, приведенными в (1).

Для выявления отклонения выборочного распределения от нормального могут быть рассмотрены следующие критерии проверки статистических гипотез:

1. Критерий проверки на асимметрию.

Сформулируем нулевую (H_0) и альтернативную (H_1) гипотезы:

– H_0: $F(x) = G(x; a, \sigma, \gamma_1, \beta_2)$, при $\gamma_1 = 0$;

– H_1: $F(x) \neq G(x; a, \sigma, 0, \beta_2)$, при $\gamma_1 = 0$.

Для проверки выдвинутых гипотез необходимо рассчитать выборочный коэффициент асимметрии As, используя (2).

Если значение $|As| \geq As(Q, n)$ при фиксированном значений n и заданном значении Q-процентных точек ($0 < Q < 100\%$), то справедлива гипотеза H_1 – распределение рассматриваемых величин отлично от нормального, поскольку $\gamma_1 \neq 0$. Процентные точки распределений рассматриваемых выборочных моментных отношений найдены в предположении, что распределение результатов наблюдений нормально [1].

Если значение $|As| < As(Q, n)$, то принимается H_0 – неизвестное теоретическое распределение $F(x)$ соответствует распределению $G(x; a, \sigma, \gamma_1, \beta_2)$ с параметром коэффициента асимметрии $\gamma_1 = 0$, как у нормального распределения.

2. Критерий проверки на эксцесс.

Сформулируем нулевую (H_0) и альтернативную (H_1) гипотезы:

– H_0: $F(x) = G(x; a, \sigma, \gamma_1, \beta_2)$, при $\beta_2 = 3$;

– H_1: $F(x) \neq G(x; a, \sigma, \gamma_1, 3)$, при $\beta_2 = 3$.

Для проверки выдвинутых гипотез необходимо рассчитать выборочный коэффициент эксцесса Ex, используя (2).

Если значение Ex не выходит за установленные Q %-ные критические границы, т. е. $Ex(100 - Q, n) < Ex < Ex(Q, n)$, то следует принять H_0 – имеет место распределение с параметром коэффициента эксцесса $\beta_2 = 3$, как у нормального распределения.

Иначе, справедлива гипотеза H_1 – распределение рассматриваемых величин отлично от нормального, поскольку не выполняется условие $\beta_2 = 3$.

3. Критерий проверки на значение статистики d.

Сформулируем нулевую (H_0) и альтернативную (H_1) гипотезы:

– H_0: $F(x) = G(x; a, \sigma, \gamma_1, \beta_2)$, при $\delta = \sqrt{2/\pi}$;

– H_1: $F(x) \neq G(x; a, \sigma, \gamma_1, \beta_2)$, при $\delta \neq \sqrt{2/\pi}$.

Для проверки выдвинутых гипотез необходимо рассчитать статистику d – выборочное среднее абсолютное отклонение, исходя из (2).

Если значение статистики d выходит за установленные Q %-ные критические границы, т. е. не выполняется условие $d(100-Q,n)<d<d(Q,n)$, то следует принять H_1 – распределение рассматриваемых величин отлично от нормального, поскольку не выполняется условие $\delta=\sqrt{2/\pi}$.

4. Комбинация рассмотренных критериев:

– H_0: $F(x)=G(x;a,\sigma,\gamma_1,\beta_2)$, при $\gamma_1=0$, $\beta_2=3$ и $\delta=\sqrt{2/\pi}$;

– H_1: $F(x)\neq G(x;a,\sigma,\gamma_1,\beta_2)$, при $\gamma_1=0$, $\beta_2=3$ и $\delta\neq\sqrt{2/\pi}$.

При справедливости гипотезы H_1 – истинно утверждение: неизвестное теоретическое распределение $F(x)$ отлично от нормального.

Также для выявления отклонения выборочного распределения от нормального могут быть использованы и другие комбинации рассмотренных критериев, например, метод моментов проверки гипотез [4] или критерий D'Agostino [10].

Отклонения выборочного распределения от нормального свидетельствует о наличии неоднородных наблюдений и о невозможности использования выборочного среднего для их усреднения.

В такой ситуации в качестве робастного аналога выборочного среднего, весьма чувствительного к наличию выбросов, особенно в случае, когда распределение ошибок характеризуется наличием «тяжелых хвостов», могут быть использованы усеченные оценки типа «среднее».

В настоящее время развиваются три класса робастных оценок [5; 7; 8; 9; 14]: устойчивые оценки на основе метода максимального правдоподобия (M-оценки); устойчивые оценки на основе ранговых критериев (R-оценки); линейные комбинации порядковых статистик (L-оценки).

Оценки, основанные на линейной комбинации порядковых статистик [5; 7; 8; 14], получили наибольшее распространение вследствие простоты их вычислительной реализации:

$$L_n=\sum_{i=1}^{n}a_i x_{(i)},\quad \sum_{i=1}^{n}a_i=1, \tag{4}$$

где a_i – некоторые постоянные коэффициенты; $x_{(i)}$, $i=\overline{1,n}$ – размещенные в порядке возрастания выборочные данные x_i: $x_{(1)}\le x_{(2)}\le \mathrm{K}\le x_{(i)}\le \mathrm{K}\le x_{(n)}$.

К таким оценкам относятся усеченные, цензурированные, винзорированные средние, выборочная медиана и др.

Рассмотрим некоторые из них для определения параметра положения a выборки $x_{(1)},x_{(2)},...,x_{(i)},...,x_{(n)}$:

1. α-усеченное среднее ($0<\alpha<0{,}5$) определяется формулой [5; 7; 8; 14]:

$$T(\alpha)=\frac{1}{n-2g}\sum_{i=g+1}^{n-g}x_{(i)}\ , \tag{5}$$

где n – объем оцениваемой выборки данных; значение g определяется из соотношения $g=[\alpha n]$; $[\alpha n]$ – наибольшее целое, не превосходящее αn, ($\alpha=\varepsilon$). Данная оценка вычисляется как среднее внутренних значений ряда, оставшихся при удалении с обоих концов ряда крайних $[\alpha n]$ значений.

2. Среднее по Винзору уровня α ($0<\alpha<0{,}5$) определяется формулой [5; 7; 8; 14]:

$$T(\alpha)=\frac{1}{n}(gx_{(g+1)}+\sum_{i=g+2}^{n-g-1}x_{(i)}+gx_{(n-g)})\ . \tag{6}$$

При вычислении данной оценки производится замена $[\alpha n]$ крайних левых наблюдений (значений) ряда на $x_{([\alpha n]+1)}$ и $[\alpha n]$ крайних правых наблюдений ряда на $x_{(n-[\alpha n])}$ с последующим вычислением среднего получившейся модифицированной выборки. Идея, которая стоит за такой последовательностью действий, заключается в том, чтобы не отбрасывая совсем αn крайних левых и αn крайних правых значений выборки, как при вычислении α-усеченного среднего, лишь сократить их влияние на значения более умеренной порядковой статистики.

3. Выборочная медиана может рассматриваться как одна из оценок α-усеченного среднего с максимальным порогом урезания равным $\alpha=\varepsilon=0{,}5$, т. е. медиана является устойчивой оценкой для уровня «засорения» $\varepsilon=0{,}5$. Выражение для медианы будут иметь следующий вид:

$$X_{med}=\begin{cases}x_{((n+1)/2)}, \text{ если } n-\text{нечетное},\\ (x_{(n/2)}+x_{(1+n/2)})/2\ , \text{ если } n-\text{четное}.\end{cases} \tag{7}$$

4. Оценка Гаствирта [9,14], представляет собой среднее по процентилям $33^1/_3$, 50, $66^2/_3$, взятыми соответственно с весами 0,3; 0,4; 0,3 и имеет вид:

$$GAS=0{,}3X_{([n/3+1])}+0{,}4X_{med}+0{,}3X_{(n-[n/3])}\ , \tag{8}$$

где $X_{([n/3+1])}$ и $X_{(n-[n/3])}$ – соответствующие квантили, X_{med} – выборочная медиана.

5. Усеченное среднее Тьюки [5,7,8], представляет собой средневзвешенную оценку первой, второй и третьей квартилей, взятых с соответствующими весами 0,25; 0,5; 0,25 и имеет вид:

$$HL=\begin{cases}x_{(g)}, \text{ если } n-\text{не кратно } 4;\\ (x_{(g)}+x_{(g+1)})/2\ , \text{ если } n-\text{кратно } 4;\end{cases}$$

$$HU=\begin{cases}x_{(n-g+1)}, \text{ если } n-\text{не кратно } 4;,\\ (x_{(n-g)}+x_{(n-g+1)})/2\ , \text{ если } n-\text{кратно } 4;\end{cases} \tag{9}$$

$$TRI=(HL+2X_{(1/2)}+HU)/4;$$

где HL и HU – соответственно первая и третья квартили, $X_{(1/2)}$ – выборочная медиана (вторая квартиль); значение g определяется из соотношения $g=[(n+3)/4]$; $[\cdot]$ – определяет целую часть.

Основной проблемой рассмотренных оценок является выбор величины коэффициента усечения α, которая в достаточной мере может быть решена посредством применения адаптивных робастных процедур статистического оценивания данных [9; 11; 13; 14].

При адаптивном подходе конкретный вид статистической процедуры выбирается на основе вспомогательной оценки какой-либо характеристики (например,

эксцесс, асимметрия) неизвестной функции распределения генеральной совокупности данных.

В общем виде адаптивная робастная оценка ES может быть представлена, как

$$ES = \begin{cases} T(\alpha_1), & \text{если} \quad stat < p_1; \\ T(\alpha_2), & \text{если} \quad p_1 \le stat < p_2; \\ \dots\dots\dots & \dots\dots \quad \dots\dots\dots\dots\dots\dots; \\ T(\alpha_i), & \text{если} \quad p_{i-1} \le stat < p_i; \\ \dots\dots\dots & \dots\dots \quad \dots\dots\dots\dots\dots\dots; \\ T(\alpha_{k-1}) & \text{если} \quad p_{k-2} \le stat \le p_{k-1}; \\ T(\alpha_k) & \text{если} \quad p_{k-1} < stat; \end{cases} \qquad (10)$$

где $T(\alpha_i)$ – усеченное среднее вида (5); $stat$ – выборочная статистика (эксцесс, асимметрия, мера длины хвостов и др.); p_i – значение порогов, равные определенным значениям выборочной статистики $stat$.

Основная идея процедур вида (10) заключается в том, что в зависимости от значения величины $stat$ выбирается соответствующий уровень усечения α_i и подсчитывается одна из перечисленных оценок типа «среднее» $T(\alpha_i)$. В зависимости от типа адаптивной процедуры уровень усечения α_i может принимать одно из следующих значений: $\alpha = 0$ – в этом случае в качестве усредненной оценки рекомендуется использовать стандартное выборочное среднее $\bar{x}$; $\alpha \in (0; 0{,}5)$ – в этом случае в качестве усредненной оценки используется одна из оценок α-усеченного среднего с выбран-ным уровнем усечения α_i; $\alpha = 0{,}5$ – в качестве усредненной оценки рекомендуется использовать стандартную выборочную медиану X_{med}.

Некоторые адаптивные процедуры не предполагают «нулевое» усечение ($\alpha \in (0; 0{,}5)$), т. е. в любом случае часть наблюдений будет отброшена, исключена из дальнейшего статистического анализа. При таком подходе, особенно в случае необоснованного применения рассматриваемых оценок, это приводит к получению некорректных, неправдоподобных результатов.

В настоящее время сформировался достаточно обширный класс адаптивных оценок, среди которых можно выделить оценки, в основе которых лежит как симметричное [9; 11; 14], так и несимметричное усечение [12; 13]. В случае симметричного усечения отбрасывается $[\alpha n]$ младших и $[\alpha n]$ старших членов вариационного ряда. При несимметричном усечении пропорция усечения α дополнительно разбивается на пропорции α_L и α_U ($\alpha = \alpha_L + \alpha_U$), что соответствует уровням усечения $[\alpha_L n]$ младших и $[\alpha_U n]$ старших членов вариационного ряда.

Выводы

В работе рассмотрены критерии отклонения выборочного распределения от нормального. Отклонения выборочного распределения от нормального позволяет сделать вывод о присутствии в анализируемой совокупности данных некоторой доли неоднородных значений. Для обработки подобных совокупностей наблюдений и нахождения их усредненных оценок в работе рассмотрен ряд усеченных оценок, в частности адаптивные робастные процедуры.

ЛИТЕРАТУРА

1. Большев Л. Н. Таблицы математической статистики / Л. Н. Большев, Н. В. Смирнов. – М. : Наука, 1983. – 416 с.
2. Золотухина Л. А. Применение модифицированного критерия Саркади для проверки гипотезы нормальности малых выборок. II / Л. А. Золотухина, Е. В. Винник // Заводская лаборатория. – 1987. – Т. 53. – Вып. 7. – С. 51–54.
3. Кобзарь А. И. Прикладная математическая статистика. Для инженеров и научных работников / А. И. Кобзарь. – М. : ФИЗМАТЛИТ, 2006. – 816 с
4. Орлов А. И. Прикладная статистика / А. И. Орлов. – М. : Экзамен, 2006. – 671 с.
5. Робастность в статистике. Подход на основе функций влияния / Ф. Хампель, Э. Рончетти, П. Рауссеу [и др.]. – М. : Мир, 1989. – 512 с.
6. Селезнев В. Д. Исследование свойств критериев согласия функции распределения данных с гауссовой методом Монте-Карло для малых выборок / В. Д. Селезнев, К. С. Денисов // Заводская лаборатория. Диагностика материалов. – 2005. – Т. 71. – № 1. – С. 68–73.
7. Тьюки Дж. Анализ результатов наблюдений / Дж. Тьюки. – М. : Мир, 1981. – 692 с.
8. Хьюбер П. Робастность в статистике : [пер. с англ.] / П. Хьюбер. – М. : Мир, 1984. – 304 с.
9. Balakrishnan N. Parameters, order statistics, outliers and robustness / N. Balakrishnan // Revista Matematica Complutence. – 2007. – Vol. 20. – No. 1. – P. 7–107.
10. D'Agostino R. B. Transformation to normality of the null distribution of g1 / R. B. D'Agostino // Biometrika. – 1970. – Vol. 57. – P. 679–681.
11. Hogg R. V. A review of some adaptive statistical techniques / R. V. Hogg, R. V. Lenth // Communications in Statistics – Theory and methods. – 1984. – Vol. 13. – P. 1551–1579.
12. Hersgaard D. Distributions of asymmetric trimmed means / D. Hersgaard // Communications in statistics: Simulation and Computation. – 1979. – No 8 (4). – P. 359–367.
13. Reed III J. F. Hinge estimators of location: Robust to asymmetry / J. F. Reed III, D. B. Stark // Computer methods and programs in Biomedicine. – 1996. – Vol. 49. – P. 11–17.
14. Robust Estimates of Locations: Survey and Advances / D. F. Andrews [et al.]. – Princeton, NJ : Princeton University Press, 1972. – 374 p.
15. Shapiro S. S. An analysis of variance test for normality (complete samples) / S. S. Shapiro, M. B. Wilk // Biometrika. – 1965. – Vol. 52. – P. 591–611.

Коваленко І. І., *Національний університет кораблебудування імені адмірала Макарова, м. Миколаїв, Україна*
Швед А. В., *Чорноморський державний університет імені Петра Могили, м. Миколаїв, Україна*

ФОРМАЛІЗОВАНИЙ ПІДХІД ДО ВИБОРУ ОЦІНОК ТИПУ «СЕРЕДНЄ» ПРИ АНАЛІЗІ МАЛИХ НЕОДНОРІДНИХ ВИБІРОК ДАНИХ

Вибірки реальних даних містять 5-10 % спостережень, що в певній мірі виділяються із аналізованої сукупності даних, внаслідок чого вибірка стає неоднорідною. Джерелом виникнення подібних спостережень може служити поява великих помилок, округлення і угрупування. Наявність подібних відхилень веде до спотворення строгих параметричних моделей (наприклад, нормального закону розподілу), що характеризується появою «важких хвостів», значно вираженими асиметрією та ексцесом. Застосування в цих умовах загальноприйнятих процедур оцінювання, наприклад, вибіркового середнього, тягне за собою збільшення зсуву і зниження ефективності одержуваних оцінок.

У статті запропоновано підхід до обробки малих неоднорідних сукупностей спостережень та визначення їх усереднених оцінок на основі робастних статистичних процедур. Для оцінювання параметру положення спотворених параметричних моделей розподілів розглянуто клас робастних оцінок, представлених лінійними комбінаціями порядкових статистик.

Ключові слова: *адаптивні робастні оцінки; усічене середнє; «хвости» розподілу; симетричне та несиметричне усікання.*

Kovalenko I. I., *Admiral Makarov National University of Shipbuilding, Mykolayiv, Ukraine*
Shved A. V., *Petro Mohyla Black Sea State University, Mykolayiv, Ukraine*

FORMALIZED APPROACH TO THE SELECTION OF ESTIMATES LIKE «MEAN» IN THE ANALYSIS OF SMALL NON-UNIFORM SAMPLES

The «real-world» data sets may contain nearly 5-10 % of cases, which can be interpreted as outliers, whereby the samples becomes nonuniform. The presence of outliers leads to a distortion of strict parametric models (e.g., normal), and implies a heavy-tailed distributions, pronounced skewness and kurtosis. Applying in these cases some traditional measures, such as sample mean, can lead to significant distortion of the results. The approach of processing of small non-uniform set of observations and calculation the estimators like traditional mean based on sustainable statistical procedures is proposed in this paper. To identify outliers or extreme observations is proposed to investigate the deviation of the underlying data distribution model from a family of symmetric distributions (e.g., normal). The following normality test were analyzed: the test of skewness, the test of kurtosis, the omnibus test. The deviation of underlying population distribution from the normal distribution model is indicated on outliers in the overall data set and as a conclusion, the impossibility for using the sample mean for averaging. For analysing data sets with single-outlier or multiple-outlier models was considered the class of robust estimators based on linear combinations of order statistics. The proposed set of estimators include a set of linear estimators capable of symmetric and asymmetric trimming and class of adaptive adaptive statistical procedures which used some selector statistics that are measures of skewness, tail length or kurtosis. These wide range of statistics can be used to characterize data sets, even for small samples, and proposed as alternative estimators to the sample mean.

Key words: *adaptive robust procedures; trimmed mean; «tails» of the distribution; symmetric and asymmetric trimming.*

© Коваленко І. І., Швед А. В., 2014 *Дата надходження статті до редколегії 12.12.2014 р.*