

УДК 519.24

DOI: 10.17223/19988605/55/11

В.П. Шуленин

АСИМТОТИЧЕСКИЕ СВОЙСТВА И РОБАСТНОСТЬ ОЦЕНОК УРЕЗАННЫХ ВАРИАНТОВ СТАНДАРТНОГО ОТКЛОНЕНИЯ И СРЕДНЕГО АБСОЛЮТНЫХ ОТКЛОНЕНИЙ

Изучаются свойства робастных оценок масштабного параметра, который характеризует «разброс» случайной величины. Предложенные оценки асимптотически нормально распределены, имеют ограниченные функции влияния и, следовательно, в отличие от оценки стандартного отклонения, «защищены» от наличия выбросов в выборке. Рассматриваемые оценки вычисляются на основе упорядоченной статистики, из которой предварительно удаляется часть наблюдений. Предложен адаптивный вариант оценок, основанный на использовании выборочных оценок функционалов, характеризующих степень «затянутости хвостов» распределений. Приводятся результаты сравнения оценок масштабного параметра в условиях различных моделей наблюдений, в частности в рамках гауссовской модели с масштабным засорением.

Ключевые слова: масштабный параметр; робастные оценки; выбросы; функция влияния; адаптивные оценки.

При обработке результатов измерений $X_1, \dots, X_n$, полученных в эксперименте путем наблюдений за признаком X изучаемого объекта, традиционно вычисляются по статистическим данным $X_1, \dots, X_n$ оценки математического ожидания $M(X)$ и дисперсии $D(X)$ (или стандартного отклонения $[D(X)]^{1/2}$) и коэффициента корреляции, что вполне оправдано для гауссовской модели наблюдений. Однако оценки этих числовых характеристик изучаемой случайной величины X подвержены сильному влиянию грубых ошибок (выбросов), которые обычно присутствуют в реальных данных $X_1, \dots, X_n$ эксперимента [1–5]. Для больших совокупностей данных, особенно при их автоматизированной обработке на ЭВМ, тщательный анализ данных с целью обнаружения «выпадающих» наблюдений и их корректировки трудно осуществим. В таких ситуациях используются робастные процедуры [5–8], которые, с одной стороны, оказываются нечувствительными к наличию умеренного засорения данных грубыми ошибками, а с другой стороны, ведут себя «достаточно хорошо» при идеальных условиях нормальности или какого-либо другого предположения о типе распределения данных.

В настоящее время разработано большое количество робастных оценок параметра положения в качестве альтернативы оценке математического ожидания, которые имеют ограниченные функции влияния и «защищены» от наличия выбросов в выборке. Многие из этих оценок принадлежат к общим классам M-, L-, R- и MD-оценок, которые хорошо изучены и описаны в литературе (см., напр.: [6–12]).

Менее изученными являются робастные оценки масштабного параметра, который используется в качестве «меры», характеризующей степень разброса случайной величины, и определяется в виде функционала $S(F)$ от функции распределения F наблюдений. Общие требования, предъявляемые к таким функционалам, сформулированы в работах [2, 4].

Традиционно используемые на практике оценки масштабного параметра, такие как оценка $\hat{S}_1(0)$ стандартного отклонения $S_1(F)$, и оценка $\hat{S}_2(0)$ среднего абсолютных отклонений $S_2(F)$, имеют неограниченные функции влияния, и они очень чувствительны к наличию выбросов в выборке (см., напр.: [6–8]). В данной работе предложены урезанные варианты этих оценок $\hat{S}_1(\alpha)$ и $\hat{S}_2(\alpha)$, $0 \leq \alpha < 1/2$, которые вычисляются не по исходной выборке $X_1, \dots, X_n$, а на основе упорядоченной статистики $X_{(1)}, \dots, X_{(n)}$, из которой предварительно удаляются $[\alpha n]$ наименьших и наибольших поряд-

ковых статистик. Получены асимптотические формулы для вычисления характеристик качества оценок, основанные на «стандартизованных дисперсиях» оценок. Показано, что эти характеристики существенно зависят от параметра α . На практике это приводит к дополнительным усилиям по выбору данного параметра. По этой причине в работе предложены адаптивные варианты оценок, для которых параметр $\hat{\alpha}(X_1, \dots, X_n)$ доопределяется на основе информации, содержащейся в исходной выборке. Приводятся результаты сравнения оценок масштабного параметра в условиях различных моделей наблюдений.

Полученные результаты, приводят к следующему выводу. Предложенная адаптивная оценка $\hat{S}_1(\hat{\alpha})$ стандартного отклонения имеет ограниченную функцию влияния, она «защищена» от наличия выбросов в выборке, имеет высокую эффективность в условиях гауссовской модели и может быть рекомендована к использованию на практике для оценивания масштабного параметра при умеренных объемах выборки и при возможных отклонениях от модели, вызванных наличием выбросов в выборке.

1. Общий подход к построению оценок масштабного параметра

Введем необходимые понятия и обозначения. Пусть X – изучаемая случайная величина с функцией распределения (ф.р.) $F(x)$, $x \in R^1$, которая абсолютно непрерывна, имеет плотность $f(x)$, $x \in R^1$, и симметрична относительно точки θ_x , т.е. $F \in \mathfrak{S}_{S|0}$, где $\mathfrak{S}_{S|0} = \{F : F(x) = 1 - F(2\theta_x - x), \forall x \in R^1\}$. Везде ниже обратную (квантильную) функцию для ф.р. $F(x)$ будем обозначать через $F^{-1}(t) = \inf\{x : F(x) \geq t\}$, $0 \leq t \leq 1$. Масштабный параметр функции распределения F используется в качестве меры, характеризующей степень разброса случайной величины (с.в.) X с ф.р. F . Рассмотрим такие меры, которые могут быть представлены в виде функционала $S(F)$, $F \in \mathfrak{S}$, заданного на множестве допустимых распределений $\mathfrak{S}$ в условиях эксперимента, связанного с изучением с.в. X по статистическим данным $X_1, \dots, X_n$, полученным в серии n независимых и повторных наблюдений над с.в. X . Выборочная оценка $\hat{S}(X_1, \dots, X_n)$ функционала $S(F)$, $F \in \mathfrak{S}$, построенная методом подстановки, записывается в виде: $\hat{S}(X_1, \dots, X_n) = S(F_n)$, где $F_n(x)$ – эмпирическая функция распределения, построенная по выборке $X_1, \dots, X_n$. Везде ниже порядковые статистики исходной выборки $X_1, \dots, X_n$ будем обозначать через $X_{(1)}, \dots, X_{(n)}$. Общие требования, которым должен удовлетворять функционал $S(F)$, описывающий разброс случайной величины X , сформулированы в работах [2, 4]. Для формулировки этих требований отметим, что о случайных величинах X_1 и X_2 с функциями распределений F_1 и F_2 говорят, что с.в. X_2 стохастически больше чем с.в. X_1 (при этом используют обозначение в виде $F_1 <_{st} F_2$), если выполняется неравенство $P(X_1 > x) \leq P(X_2 > x)$, $\forall x \in R^1$. Отметим также, что $F_1 <_{st} F_2 \Rightarrow F_1(x) \geq F_2(x)$, $\forall x \in R^1$ и $F_1^{-1}(t) \leq F_2^{-1}(t)$, $0 \leq t \leq 1$. Далее, разброс с.в. X относительно θ_x (масштабный параметр с.в. X) определяют в терминах «расстояния» X от θ_x , т.е. с помощью величины $|X - \theta_x|$, при этом говорят, что с.в. X_1 имеет больший разброс относительно θ_{x_1} , чем с.в. X относительно θ_x , если с.в. $|X_1 - \theta_{x_1}|$ стохастически больше с.в. $|X - \theta_x|$. Функционал $S(F)$ определяет масштабный параметр ф.р. F , если он удовлетворяет следующим условиям.

Условие 1. Монотонность функционала $S(F)$ относительно стохастического возрастания распределений, т.е. выполняется выражение $S(F_1) \leq S(F_2)$ для $F_1 <_{st} F_2$, где F_1 и F_2 – функции распределения вероятностей случайных величин $|X_1 - \theta_{x_1}|$ и $|X_2 - \theta_x|$.

Условие 2. Оценка $\hat{S}(X_1, \dots, X_n) = S(F_n)$ функционала $S(F)$ является эквивариантной относительно линейных преобразований наблюдений $X_1, \dots, X_n$, т.е. для нее выполняется равенство $\hat{S}(aX_1 + b, \dots, aX_n + b) = |a| \hat{S}(X_1, \dots, X_n)$.

В литературе описаны различные подходы к построению «мер», характеризующих масштабный параметр (см., напр.: [6, 7]). Мы рассмотрим такие функционалы $S(F)$, которые построены с помощью отклонений каждого члена генеральной совокупности от некоторого «центрального» (типичного значения) с.в. X с ф.р. F . Обычно в качестве такого значения используется параметр положения, определенный функционалом $T(F)$, либо в виде среднего $T_1(F) = \int x dF(x) = M(X)$, либо в виде медианы $T_2(F) = F^{-1}(1/2) = MED(X)$. Обозначим через F_1 ф.р. случайной величины $|X - T(F)|$. Рассмотрим функционалы вида

$$S(F) = \left\{ \int_0^1 [F_1^{-1}(t)]^\gamma dV(t) \right\}^{1/\gamma}, \quad (1)$$

где $V(t)$ – некоторая функция распределения на $[0, 1]$ и $\gamma > 0$. Например, если в (1) в качестве параметра положения выбрать среднее значение $T_1(F)$, и положить $V(t) = t$, $0 \leq t \leq 1$, то при $\gamma = 1$ получим среднее абсолютных отклонений $S_2(F) = M\{|X - M(X)|\}$, оценка которого запишется в виде: $\hat{S}_2(0) = n^{-1} \sum |X_i - \bar{X}|$. При $\gamma = 2$ будем иметь стандартное отклонение $S_1(F) = \{M(X - M(X))^2\}^{1/2}$, оценка которого запишется в виде: $\hat{S}_1(0) = \{n^{-1} \sum (X_i - \bar{X})^2\}^{1/2}$. Если же положить $V(t) = t / (1 - \alpha)$, $0 \leq t \leq 1 - \alpha$, $0 \leq \alpha < 1/2$, то получим α -урезанные варианты $S_i(F, \alpha)$, $i = 1, 2$, $0 \leq \alpha < 1/2$, указанных мер масштабного параметра, оценки которых вычисляются на основе порядковых статистик $X_{(1)}, \dots, X_{(n)}$ (см. ниже).

Отметим еще раз, что для выбранного функционала $S(F)$, $F \in \mathfrak{F}$, описывающего масштабный параметр с.в. X с ф.р. F , его оценка строится методом подстановки и записывается в виде $\hat{S}(X_1, \dots, X_n) = S(F_n)$. Асимптотическая нормальность таких оценок масштабного параметра изучается методом Мизеса [9, 10] с использованием теоремы Слуцкого, центральной предельной теоремы и разложения вида:

$$S(F_n) = S(F) + \frac{1}{n} \sum_{i=1}^n IF(X_i; F, S) + o_p(n^{-1/2}). \quad (2)$$

Использование разложения (2) позволяет доказать асимптотическую нормальность оценки $\hat{S} = S(F_n)$, где $IF(x; F, S)$ – функция влияния Хампеля оценки $\hat{S} = S(F_n)$ функционала $S(F)$, $F \in \mathfrak{F}$ (см., напр.: [6, 13–15]). Функция влияния Хампеля определяется в виде:

$$IF(x; F, S) = \lim_{\lambda \rightarrow 0+} \frac{S[(1-\lambda)F + \lambda\Delta_x] - S(F)}{\lambda}, \quad 0 \leq \lambda \leq 1, \quad x \in R^1, \quad (3)$$

для тех $x \in R^1$, при которых предел существует. Здесь Δ_x обозначает «вырожденную функцию распределения» в точке $x \in R^1$, которая определяется в точке $x \in R^1$ через единичную функцию Хевисайда $C(y - x) = \begin{cases} 1, & y \geq x, \\ 0, & y < x \end{cases}$ в виде: $F_x(y) = C(y - x)$, $(y, x) \in R^1$. Асимптотическая дисперсия

$\sqrt{n}\hat{S}$ -оценки обозначается через $\sigma^2(F, \hat{S})$ и вычисляется по формуле

$$\sigma^2(F, \hat{S}) = \int_{-\infty}^{\infty} IF^2(x; F, S) dF(x). \quad (4)$$

Для сравнения различных оценок масштабного параметра при заданной ф.р. F будем использовать понятие асимптотической относительной эффективности, определенное через обратное отношение стандартизованных асимптотических дисперсий. Асимптотическую эффективность оценки $\hat{S}_1$ относительно $\hat{S}_2$ при заданной ф.р. F обозначим через $AO\mathfrak{E}_F(\hat{S}_1; \hat{S}_2)$ и, следуя работам [3, 7], определим в виде:

$$AO\mathfrak{D}_F(\hat{S}_1 : \hat{S}_2) = \frac{\tilde{\sigma}^2(F, \hat{S}_2)}{\tilde{\sigma}^2(F, \hat{S}_1)}, \quad (5)$$

где $\tilde{\sigma}^2(F, \hat{S})$ – стандартизованная дисперсия $\sqrt{n}\hat{S}$ -оценки, равная отношению асимптотической дисперсии к квадрату оцениваемого функционала, т.е.

$$\tilde{\sigma}^2(F, \hat{S}) = \sigma^2(F, \hat{S}) / S^2(F). \quad (6)$$

2. Урезанные варианты стандартного отклонения и среднего абсолютных отклонений

Пусть имеется последовательность $X_1, \dots, X_n$ н.о.р. случайных величин с непрерывной симметричной ф.р. $F(x)$, $x \in R^1$, т.е. $F \in \mathfrak{S}_{S|0}$, и везде ниже, согласно второму условию эквивариантности оценок, полагаем без потери общности точку симметрии $\theta = M(X) = 0$. Обозначим ф.р. случайной величины $Y = |X_1|$ через $G(x)$, а ф.р. случайной величины $Z = X_1^2$ обозначим через $H(x)$. Определим оценку α -урезанного стандартного отклонения в виде:

$$\hat{S}_1(\alpha) = \left\{ \frac{1}{n - [\alpha n]} \sum_{i=1}^{n - [\alpha n]} Z_{(i)} \right\}^{1/2}, \quad 0 \leq \alpha < 1/2, \quad (7)$$

где $Z_{(i)}$ – i -я порядковая статистика для преобразованной выборки $Z_1, \dots, Z_n$, $Z_i = X_i^2$, $i = 1, \dots, n$. Функционал, соответствующий этой оценке, может быть определен, либо через ф.р. $H(x)$ в виде:

$$S_1(H, \alpha) = \left\{ \frac{1}{1 - \alpha} \int_0^{H^{-1}(1-\alpha)} z dH(z) \right\}^{1/2}, \quad 0 \leq \alpha < 1/2, \quad (8)$$

либо через исходную ф.р. F в виде:

$$S_1(F, \alpha) = \left\{ \frac{1}{1 - \alpha} \int_{F^{-1}(\alpha/2)}^{F^{-1}(1-\alpha/2)} x^2 dF(x) \right\}^{1/2}, \quad 0 \leq \alpha < 1/2. \quad (9)$$

При $\alpha = 0$ оценка (7) превращается в обычную оценку стандартного отклонения

$$\hat{S}_1(0) = \left\{ n^{-1} \sum_{i=1}^n X_i^2 \right\}^{1/2}. \quad (10)$$

Оценку α -урезанного среднего абсолютных отклонений определим в виде:

$$\hat{S}_2(\alpha) = \frac{1}{n - [\alpha n]} \sum_{i=1}^{n - [\alpha n]} Y_{(i)}, \quad 0 \leq \alpha < 1/2, \quad (11)$$

где $Y_{(i)}$ – i -я порядковая статистика для преобразованной выборки $Y_1, \dots, Y_n$, $Y_i = |X_i|$, $i = 1, \dots, n$. Функционал, соответствующий этой оценке, имеет вид:

$$S_2(G, \alpha) = \frac{1}{1 - \alpha} \int_0^{1-\alpha} G^{-1}(t) dt, \quad 0 \leq \alpha < 1/2. \quad (12)$$

При $\alpha = 0$ оценка (11) превращается в обычную оценку среднего абсолютных отклонений

$$\hat{S}_2(0) = n^{-1} \sum_{i=1}^n |X_i|. \quad (13)$$

3. Асимптотические свойства оценок

Асимптотическая нормальность оценок (7) и (11) масштабного параметра изучается методом Мизеса [9, 10] с использованием теоремы Слуцкого, центральной предельной теоремы и теории

U-статистик на основе разложения (2) (см.: [8, 11, 16–19], где доказывалась асимптотическая нормальность оценок функционалов, построенных методом подстановки). Для сравнения рассмотренных оценок масштабного параметра, определенных функционалами (9) и (12) с использованием формулы (5) следует прежде всего вычислить их функции влияния (3), асимптотические дисперсии (4) и стандартизованные дисперсии, определенные в (6). Для удобства записи введем следующие обозначения:

$$I_0(F, \alpha) = 2 \int_0^{1-(\alpha/2)} \xi(t) dt, \quad I_1(F, \alpha) = \int_{\alpha/2}^{1-(\alpha/2)} \xi^2(t) dt, \quad I_2(F, \alpha) = \int_{\alpha/2}^{1-(\alpha/2)} \xi^4(t) dt, \\ \text{где } \xi(t) = F^{-1}(t), \quad 0 \leq t \leq 1. \quad (14)$$

Можно убедиться, что функция влияния $IF(u; H, \hat{S}_1(\alpha))$, $0 \leq u < \infty$, для $\hat{S}_1(\alpha)$ -оценки вида (7) записывается в виде:

$$IF(u; H, \hat{S}_1(\alpha)) = \frac{1}{2(1-\alpha)S_1(H, \alpha)} \begin{cases} u - b(\alpha), & 0 \leq u \leq H^{-1}(1-\alpha) \\ H^{-1}(1-\alpha) - b(\alpha), & H^{-1}(1-\alpha) < u < \infty \end{cases}, \\ b(\alpha) = \int_0^{H^{-1}(1-\alpha)} z dH(z) + \alpha H^{-1}(1-\alpha). \quad (15)$$

Используя приведенную формулу (15) и формулы (4) и (6) получаем, с учетом (14), выражение для вычисления стандартизованной дисперсии $\sqrt{n}\hat{S}_1(\alpha)$ -оценки в виде:

$$\tilde{\sigma}^2(F, \hat{S}_1(\alpha)) = \{I_2(F, \alpha) - I_1^2(F, \alpha) - 2\alpha\xi^2(1-\alpha/2)I_1(F, \alpha) + \alpha(1-\alpha)\xi^4(1-\alpha/2)\} / 4I_1^2(F, \alpha). \quad (16)$$

Далее, функция влияния $IF(y; G, \hat{S}_2(\alpha))$, $0 \leq y < \infty$ для $\hat{S}_2(\alpha)$ -оценки вида (11) записывается в виде:

$$IF(y; G, \hat{S}_2(\alpha)) = \frac{1}{(1-\alpha)} \begin{cases} y - c(\alpha), & 0 \leq y \leq G^{-1}(1-\alpha) \\ G^{-1}(1-\alpha) - c(\alpha), & G^{-1}(1-\alpha) < y < \infty \end{cases}, \quad c(\alpha) = \int_0^{G^{-1}(1-\alpha)} y dG(y) + \alpha G^{-1}(1-\alpha). \quad (17)$$

Используя приведенную формулу (17) и формулы (4) и (6) получаем, с учетом (14), выражение для вычисления стандартизованной дисперсии $\sqrt{n}\hat{S}_2(\alpha)$ -оценки в виде:

$$\tilde{\sigma}^2(F, \hat{S}_2(\alpha)) = \{I_1(F, \alpha) - I_0^2(F, \alpha) - 2\alpha\xi^2(1-\alpha/2)I_0(F, \alpha) + \alpha(1-\alpha)\xi^4(1-\alpha/2)\} / I_0^2(F, \alpha). \quad (18)$$

Отметим, что при $\alpha = 0$ имеем выражения для стандартизованных дисперсий $\sqrt{n}\hat{S}_1(0)$ - и $\sqrt{n}\hat{S}_2(0)$ -оценок в виде

$$\tilde{\sigma}^2(F, \hat{S}_1(0)) = \frac{\int x^4 dF(x) - \left(\int x^2 dF(x)\right)^2}{4\left(\int x^2 dF(x)\right)^2}, \quad \tilde{\sigma}^2(F, \hat{S}_2(0)) = \frac{\int x^2 dF(x) - \left(\int |x| dF(x)\right)^2}{\left(\int |x| dF(x)\right)^2}.$$

Замечание 1. Отметим, что приведенные функции влияния $\hat{S}_1(\alpha)$ -оценки и $\hat{S}_2(\alpha)$ -оценки (см. формулы (15) и (17)) являются *ограниченными* функциями при $0 < \alpha < 1/2$, следовательно, эти оценки «защищены» от наличия выбросов в выборке. Отметим также, что при изучении $\hat{S}_1(\alpha)$ - и $\hat{S}_2(\alpha)$ -оценок мы рассмотрели случай симметричных распределений при условии, что точка симметрии θ известна. Если она неизвестна, то для построения оценок масштабного параметра следует использовать оценки точки симметрии. Отметим, что при достаточно общих условиях асимптотические свойства $\hat{S}_1(\alpha)$ - и $\hat{S}_2(\alpha)$ -оценок не зависят от типа используемой оценки θ (см., напр.: [2, 3]).

Пример 1. Рассмотрим нормальную модель с засорением в виде: $\mathfrak{Z}_{\varepsilon, \tau}(\Phi) = \{F : F(x) = \Phi_{\varepsilon, \tau}(x)\}$, где $\Phi_{\varepsilon, \tau}(x) = (1-\varepsilon)\Phi(x) + \varepsilon\Phi(x/\tau)$, $0 \leq \varepsilon \leq 1$, $\tau \geq 1$. Обозначим через x^* квантиль уровня $1-(\alpha/2)$ для ф.р. $\Phi_{\varepsilon, \tau}(x)$, т.е. $x^* = \Phi_{\varepsilon, \tau}^{-1}(1-\alpha/2)$. С учетом этих обозначений выражения для I_0 , I_1 , I_2 из (14) для $F \in \mathfrak{Z}_{\varepsilon, \tau}(\Phi)$ записываются в виде:

$$\begin{aligned}
 I_0 &= (1-\varepsilon)[\sqrt{2/\pi} - 2\phi(x^*)] + \varepsilon\tau[\sqrt{2/\pi} - 2\phi(x^*/\tau)], \\
 I_1 &= (1-\varepsilon)[2\Phi(x^*) - 1 - 2x^*\phi(x^*)] + \varepsilon\tau^2[2\Phi(x^*/\tau) - 1 - 2\phi(x^*/\tau)x^*/\tau], \\
 I_2 &= (1-\varepsilon)[3\{2\Phi(x^*) - 1 - 2x^*\phi(x^*)\} - 2(x^*)^3\phi(x^*)] + \\
 &+ \varepsilon\tau^4[3\{2\Phi(x^*/\tau) - 1 - 2\phi(x^*/\tau)x^*/\tau\} - 2\phi(x^*/\tau)(x^*)^3/\tau],
 \end{aligned}$$

где Φ и ϕ обозначают соответственно функцию распределения и плотность стандартного нормального распределения и $x^* = \Phi_{\varepsilon,\tau}^{-1}(1-\alpha/2)$ – квантиль уровня $1-(\alpha/2)$ для ф.р. $\Phi_{\varepsilon,\tau}(x)$. Используя приведенные формулы, получаем выражения для стандартизованных дисперсий оценок $\hat{S}_1(\alpha)$ и $\hat{S}_2(\alpha)$ в виде:

$$\tilde{\sigma}^2(\Phi_{\varepsilon,\tau}, \hat{S}_1(\alpha)) = \{I_2 - I_1^2 - 2\alpha(x^*)^2 I_1 + \alpha(1-\alpha)(x^*)^4\} / 4I_1^2, \quad (19)$$

$$\tilde{\sigma}^2(\Phi_{\varepsilon,\tau}, \hat{S}_2(\alpha)) = \{I_1^2 - I_0^2 - 2\alpha(x^*)^2 I_0 + \alpha(1-\alpha)(x^*)^2\} / I_0^2. \quad (20)$$

В частности, при $\alpha \rightarrow 0$ получаем $I_0 = \sqrt{2/\pi}[1 + \varepsilon(\tau - 1)]$, $I_1 = 1 + \varepsilon(\tau^2 - 1)$, $I_2 = 3[1 + \varepsilon(\tau^4 - 1)]$.

Отметим, что для нормального распределения (при $\varepsilon=0$, или при $\tau=1$) имеем: $I_0 = \sqrt{2/\pi}$, $I_1 = 1$, $I_2 = 3$ и, следовательно,

$$\tilde{\sigma}^2(\Phi, \hat{S}_1(0)) = \frac{1}{4} \left(\frac{I_2 - I_1^2}{I_1^2} \right) = \frac{1}{4} \left(\frac{3-1}{1} \right) = 0,500, \quad \tilde{\sigma}^2(\Phi, \hat{S}_2(0)) = \frac{I_1^2 - I_0^2}{I_0^2} = \frac{\pi}{2} - 1 = 0,571.$$

Таким образом, асимптотическая относительная эффективность $\hat{S}_2(0)$ -оценки по отношению к $\hat{S}_1(0)$ для нормального распределения равна $AO\mathcal{E}_\Phi(\hat{S}_2(0) : \hat{S}_1(0)) = 0,500 / 0,572 = 0,876$, т.е. $\hat{S}_2(0)$ -оценка среднего абсолютных отклонений проигрывает $\hat{S}_1(0)$ -оценке стандартного отклонения по эффективности при нормальном распределении примерно 12%.

С учетом приведенных выражений стандартизованные дисперсии стандартного отклонения $\hat{S}_1(0)$ и среднего абсолютных отклонений $\hat{S}_2(0)$ для модели с засорением, т.е. для $F = \Phi_{\varepsilon,\tau}$, вычисляются по формулам

$$\begin{aligned}
 \tilde{\sigma}^2(\Phi_{\varepsilon,\tau}, \hat{S}_1(0)) &= \frac{3[1 + \varepsilon(\tau^4 - 1)] - [1 + \varepsilon(\tau^2 - 1)]}{4[1 + \varepsilon(\tau^2 - 1)]^2}, \\
 \tilde{\sigma}^2(\Phi_{\varepsilon,\tau}, \hat{S}_2(0)) &= \frac{\pi[1 + \varepsilon(\tau^2 - 1)] - 2[1 + \varepsilon(\tau - 1)]^2}{2[1 + \varepsilon(\tau - 1)]^2}.
 \end{aligned} \quad (21)$$

Согласно формулам (21), асимптотическая относительная эффективность $\hat{S}_2(0)$ -оценки по отношению к $\hat{S}_1(0)$ для супермодели $\mathfrak{T}_{\varepsilon,\tau}(\Phi)$ вычисляется по формуле

$$AO\mathcal{E}_F(\hat{S}_2(0) : \hat{S}_1(0)) = \frac{\{3[1 + \varepsilon(\tau^4 - 1)] - [1 + \varepsilon(\tau^2 - 1)]\}[1 + \varepsilon(\tau - 1)]^2}{2\{\pi[1 + \varepsilon(\tau^2 - 1)] - 2[1 + \varepsilon(\tau - 1)]^2\}[1 + \varepsilon(\tau^2 - 1)]^2}, \quad F \in \mathfrak{T}_{\varepsilon,\tau}(\Phi). \quad (22)$$

Из этой формулы также следует при $\varepsilon=0$, что $AO\mathcal{E}_\Phi(\hat{S}_2(0) : \hat{S}_1(0)) = 1/(\pi - 2) = 0,876$.

Приведем вычисленные по формуле (22) численные значения асимптотической относительной эффективности $\hat{S}_2(0)$ -оценки, $0 \leq \alpha < 1/2$, по отношению к $\hat{S}_1(0)$ для супермодели $\mathfrak{T}_{\varepsilon,\tau}(\Phi)$ (табл. 1).

Отметим, что значения эффективности $AO\mathcal{E}_{\Phi_{\varepsilon,\tau}}(\hat{S}_2(0) : \hat{S}_1(0))$, соответствующие большим значениям параметра τ и малым значениям ε , являются неожиданно высокими. По всей вероятности, это является следствием того факта, что функция влияния оценки $\hat{S}_1(0)$ не является ограниченной (см. также: [3]). Отметим также, что среднее абсолютных отклонений $\hat{S}_2(0)$, проигрывая лишь $\approx 12\%$ в эффективности стандартному отклонению $\hat{S}_1(0)$ при нормальном распределении, становится более эффективной при $\tau=3$ уже для $\varepsilon=0,002$.

Таблица 1

Значения $AO\mathfrak{E}_F(\hat{S}_2(0): \hat{S}_1(0))$ для супермодели $\mathfrak{S}_{\varepsilon, \tau}(\Phi)$

$\varepsilon \setminus \tau$	3	5	7	10	20
0,001	0,95	1,53	3,31	9,46	55,0
0,002	1,02	2,07	4,95	31,1	55,8
0,005	1,20	3,20	7,29	14,7	22,8
0,007	1,30	3,67	7,75	13,5	16,0
0,01	1,44	4,10	7,78	11,6	10,6
0,02	1,75	4,40	6,41	7,08	4,60
0,05	2,04	3,39	3,52	3,00	1,75
0,07	2,02	2,81	2,68	2,20	1,33
0,10	1,90	2,24	2,00	1,62	1,04
0,20	1,51	1,41	1,21	1,02	0,76

При умеренных значениях параметра τ ($\tau < 10$) и небольших значениях ε ($\alpha \leq 0,2$) среднее абсолютных отклонений $\hat{S}_2(0)$ предпочтительнее стандартного отклонения $\hat{S}_1(0)$.

Замечание 2. Заметим, что Эддингтон и Фишер придерживались противоположной точки зрения относительно преимуществ оценок $\hat{S}_1(0)$ и $\hat{S}_2(0)$ (см.: [7. С. 10]). Отметим, что $\hat{S}_1(0)$ и $\hat{S}_2(0)$ оценивают *разные* функционалы: $S_1(F) = \{M(X - M(X))^2\}^{1/2}$ и $S_2(F) = M\{|X - M(X)|\}$. И в условиях, например, нормальной модели наблюдений $N(0, \sigma^2)$ оценка $\hat{S}_1(0)$ сходится к σ , а оценка $\hat{S}_2(0)$ сходится к $\sigma\sqrt{2/\pi} \approx 0,80\sigma$. По этой причине, согласно [3, 7], при сравнении различных оценок масштаба используется не асимптотическая дисперсия, а *стандартизованная дисперсия* вида (6), и асимптотическая относительная эффективность определяется в виде (5).

Численные значения $AO\mathfrak{E}_F(\hat{S}_2(\alpha): \hat{S}_1(\alpha))$, для $F = \Phi_{\varepsilon, \tau}$, вычисленные по формулам (19) и (20), приведены в табл. (2).

Таблица 2

Значения $AO\mathfrak{E}_F(\hat{S}_2(\alpha): \hat{S}_1(\alpha))$, для $F = \Phi_{\varepsilon, \tau}$ при различных значениях параметра α

ε	$\tau \setminus \alpha$	$\alpha = 0,05$	$\alpha = 0,10$	$\alpha = 0,20$	$\alpha = 0,30$
$\varepsilon = 0,05$	$\varepsilon = 3$	0,928	0,892	0,876	0,872
	$\varepsilon = 5$	0,988	0,911	0,882	0,875
	$\varepsilon = 10$	1,089	0,929	0,885	0,887
$\varepsilon = 0,10$	$\varepsilon = 3$	1,032	0,930	0,890	0,879
	$\varepsilon = 5$	1,513	0,994	0,906	0,887
	$\varepsilon = 10$	3,401	1,104	0,920	0,907

Из приведенных данных следует, что при малых значениях параметра $\alpha \leq 0,10$, $\hat{S}_2(\alpha)$ -оценка предпочтительнее $\hat{S}_1(\alpha)$ -оценки лишь для распределений с «очень тяжелыми хвостами», в частности при $(\tau = 5, 10)$. В других случаях картина обратная, т.е. в рамках супермодели $\mathfrak{S}_{\varepsilon, \tau}(\Phi)$ урезанный вариант стандартного отклонения предпочтительнее.

4. Адаптивные оценки

Для построения $\hat{S}_1(\alpha)$ - и $\hat{S}_2(\alpha)$ -оценок на практике требуется задавать значения параметра α , $0 \leq \alpha < 1/2$, при изменении которого свойства этих оценок существенно меняются. Например, в условиях семейства распределений Стюдента с числом степеней свободы $r = 5$ при $\alpha = 0,10$ имеем $AO\mathfrak{E}_F(\hat{S}_1(0,10): \hat{S}_1(0)) = 2,35$, а при $\alpha = 0,20$, $AO\mathfrak{E}_F(\hat{S}_1(0,20): \hat{S}_1(0)) = 2,11$. Следуя работам [20, 21], выбор параметра α можно связать с поведением функционала $Q(F; \nu; \mu)$, характеризующего степень «тяжести хвостов» распределений, который определяется в виде:

$$Q(F; \nu, \mu) = \frac{(1/\nu) \left\{ \int_{1-\nu}^1 F^{-1}(t) dt - \int_0^\nu F^{-1}(t) dt \right\}}{(1/\mu) \left\{ \int_{1-\mu}^1 F^{-1}(t) dt - \int_0^\mu F^{-1}(t) dt \right\}}, \quad 0 < \nu < \mu \leq 0,5. \quad (23)$$

Оценка $Q(F_n)$ функционала $Q(F; \nu, \mu)$, построенная по выборке $X_1, \dots, X_n$ методом подстановки, записывается в виде:

$$Q(F_n; \nu, \mu) = \frac{m}{k} \left(\sum_{i=n-k+1}^n X_{(i)} - \sum_{i=1}^k X_{(i)} \right) / \left(\sum_{i=n-m+1}^n X_{(i)} - \sum_{i=1}^m X_{(i)} \right), \quad k = [\nu n], \quad m = [\mu n], \quad (24)$$

где $0 < \nu < \mu \leq 0,5$ и $X_{(1)}, \dots, X_{(n)}$ – порядковые статистики выборки $X_1, \dots, X_n$. Следуя работе [20],

везде ниже полагаем $\nu = 0,2$ и $\mu = 0,5$. Отметим, что $Q(F_n) \xrightarrow{p} Q(F; \nu, \mu)$ при $n \rightarrow \infty$.

Адаптивный параметр $\hat{\alpha}(X_1, \dots, X_n)$ определим в виде:

$$\hat{\alpha}(X_1, \dots, X_n) = \begin{cases} \alpha_1, & Q(F_n) \leq Q_1, \\ \frac{\alpha_2 - \alpha_1}{Q_2 - Q_1} \{Q(F_n) - Q_1\} + \alpha_1, & Q_1 < Q(F_n) < Q_2, \\ \alpha_2, & Q(F_n) \geq Q_2, \end{cases} \quad (25)$$

где параметры α_1 , α_2 , Q_1 и Q_2 задаются в соответствии с рассматриваемым типом супермодели, а выборочная оценка $Q(F_n)$ функционала $Q(F; \nu; \mu)$ определена в (24). В данной работе параметр $\hat{\alpha}(X_1, \dots, X_n)$ вычисляется по формуле (25) при следующих значениях параметров: $\alpha_1 = 0,05$, $\alpha_2 = 0,50$, $Q_1 = 1,75$, $Q_2 = 2,50$.

Для такого адаптивного выбора параметра $\hat{\alpha}(X_1, \dots, X_n)$ сравним адаптивные оценки $\hat{S}_1(\hat{\alpha})$ и $\hat{S}_2(\hat{\alpha})$ с другими оценками, для которых параметр α фиксирован. Сравнение проведено в условиях супермодели $\mathfrak{T}_{\varepsilon, \tau}(\Phi)$. Результаты приведены в табл. 3 в виде отношения стандартизованной дисперсии оценки при заданном распределении к минимальной стандартизованной дисперсии среди сравниваемых оценок.

Таблица 3

Отношения стандартизованных дисперсий оценок к минимальной среди сравниваемых оценок

ε	τ	$\hat{S}_1(0)$	$\hat{S}_1(0,1)$	$\hat{S}_1(\hat{\alpha})$	$\hat{S}_2(0)$	$\hat{S}_2(0,1)$	$\hat{S}_2(\hat{\alpha})$
0,05	3	2,25	1,02	1,00	1,20	1,15	1,12
	5	6,48	1,00	1,01	1,91	1,10	1,10
	10	13,5	1,00	1,09	4,49	1,08	1,17
0,10	3	2,37	1,00	1,00	1,25	1,08	1,06
	5	4,49	1,05	1,00	2,00	1,05	1,09
	10	6,07	1,13	1,00	3,74	1,02	1,10
0,20	3	1,67	1,05	1,00	1,11	1,00	1,01
	5	2,22	2,07	1,00	1,58	1,37	1,03
	10	2,30	7,26	1,00	2,26	3,15	1,00
0,30	3	1,26	1,25	1,03	1,00	1,08	1,06
	5	1,23	2,15	1,05	1,11	1,45	1,00
	10	1,26	3,55	1,00	1,48	2,70	1,03

Данные табл. 3 показывают, что в рамках супермодели $\mathfrak{T}_{\varepsilon, \tau}(\Phi)$ среди сравниваемых $\hat{S}_1(\alpha)$ - и $\hat{S}_2(\alpha)$ -оценок, предпочтение следует отдать адаптивному стандартному отклонению $\hat{S}_1(\hat{\alpha})$, для которого параметр $\hat{\alpha}(X_1, \dots, X_n)$ вычисляется по формуле (25) при следующих значениях параметров: $\alpha_1 = 0,05$, $\alpha_2 = 0,50$, $Q_1 = 1,75$, $Q_2 = 2,50$. Отметим, что при сравнении оценок использовались зна-

чения асимптотических стандартизованных дисперсий оценок, которые, как показали результаты моделирования для различных супермоделей, являются вполне удовлетворительной аппроксимацией стандартизованных дисперсий оценок при конечных объемах выборки $n \geq 40$. Исключением являются лишь некоторые распределения с очень «тяжелыми хвостами». Таким образом, адаптивная оценка $\hat{S}_1(\hat{\alpha})$ стандартного отклонения может быть рекомендована к использованию на практике для оценивания масштабного параметра при умеренных объемах выборки и при возможных отклонениях от гауссовской модели наблюдений, вызванных наличием выбросов в выборке.

5. Результаты сравнения оценок при конечных объемах выборки

Приведем результаты сравнения оценок масштабного параметра при конечных объемах выборки, полученные методом статистических испытаний. В эксперимент были включены оценка $\hat{S}_1(0)$ стандартного отклонения, оценка $\hat{S}_2(0)$ среднего абсолютных отклонений, средняя разность Джини $\hat{\Delta}_n = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n |X_i - X_j|$ и медиана абсолютных разностей $\hat{S}_4 = \text{med}\{|X_i - X_j|, 1 \leq i < j \leq n\}$ [5, 8].

В эксперименте вычислялись выборочные значения средних и дисперсий оценок и значения их стандартизованных дисперсий при различных объемах выборки n и различных статистических моделях исходного распределения. В качестве модели было взято гауссовское распределение с масштабным засорением, т.е. $F \in \mathfrak{Z}_{\varepsilon, \tau}(\Phi)$. Получение псевдовыборки $X_1, \dots, X_n$ для этой модели осуществлялось следующим образом. Датчиком случайных чисел $N(0, 1)$ генерировалась выборка объема $(n-k)$, оставшиеся k элементов выборки генерировались датчиком $N(0, \tau^2)$, при этом число k задавалось в соответствии с пропорцией засорения ε . В проведенных экспериментах объем выборки n менялся от 5 до 40, число экспериментов M изменялось от 1 000 до 10 000. В табл. 4 приведены значения стандартизованных дисперсий оценок при $n=20$ и $M=1000$ для супермодели $\mathfrak{Z}_{\varepsilon, \tau}(\Phi)$ при различных значениях τ и k . Отметим, что первая строка таблицы ($n \rightarrow \infty$) содержит значения асимптотических стандартизованных дисперсий оценок. В строке ($n = \infty / n = 20$) приведены отношения асимптотических значений к значениям, полученным экспериментально при $n = 20$.

Таблица 4

Стандартизованные дисперсии оценок при $n = 20$ и $M = 1\,000$ для супермодели $\mathfrak{Z}_{\varepsilon, \tau}(\Phi)$

Ф.р. $\Phi_{\varepsilon, \tau}$	k	$\hat{S}_1(0)$	$\hat{S}_2(0)$	$\hat{\Delta}_n$	$\hat{S}_4$
$F = \Phi$	$n \rightarrow \infty$	0,500	0,571	0,511	0,579
	$k = 0$	0,508	0,586	0,520	0,630
	$n = \infty / n = 20$	0,984	0,974	0,982	0,919
$\tau = 3$	$k = 2$	1,168	0,742	0,797	0,719
	$k = 4$	1,169	0,805	0,888	0,814
$\tau = 5$	$k = 2$	2,295	1,113	1,254	0,723
	$k = 4$	1,795	1,171	1,289	0,851
$\tau = 10$	$k = 2$	3,859	2,279	2,202	0,722
	$k = 4$	2,331	1,947	1,839	0,851

Из табл. 4 видно (см. строку $n = \infty / n = 20$), что асимптотика является вполне приемлемой аппроксимацией для объема выборки $n > 20$. Медиана абсолютных разностей $\hat{S}_4$ обладает преимуществом перед другими оценками при масштабном засорении. Ее стандартизованная дисперсия меняется лишь незначительно, в отличие от других оценок, при увеличении параметра τ . Это объясняется ограниченной функцией влияния оценки $\hat{S}_4$ (см.: [5]). Для распределений, «близких по знятнутости хвостов» к нормальному распределению, преимуществом обладает среднее абсолютных отклонений $\hat{S}_2(0)$.

Заключение

В работе изучены различные оценки масштабного параметра, характеризующего «разброс» случайной величины. Показано, что традиционно используемые на практике оценки масштабного параметра, такие как оценка $\hat{S}_1(0)$ стандартного отклонения $S_1(F)$ и оценка $\hat{S}_2(0)$ среднего абсолютных отклонений $S_2(F)$, имеют неограниченные функции влияния и очень чувствительны к наличию выбросов в выборке. В работе предложены урезанные варианты этих оценок $\hat{S}_1(\alpha)$ и $\hat{S}_2(\alpha)$, $0 \leq \alpha < 1/2$, которые вычисляются не по исходной выборке $X_1, \dots, X_n$, а на основе упорядоченной статистики $X_{(1)}, \dots, X_{(n)}$, из которой предварительно удаляются $[\alpha n]$ наименьших и наибольших порядковых статистик. Эти оценки «защищены» от наличия выбросов в выборке, они имеют ограниченные функции влияния, и их характеристики существенно зависят от параметра α , что на практике приводит к дополнительным усилиям по выбору этого параметра. Для устранения этой трудности в работе предложены адаптивные варианты оценок, для которых параметр $\hat{\alpha}(X_1, \dots, X_n)$ определяется на основе исходной выборке $X_1, \dots, X_n$. Получены асимптотические формулы для вычисления характеристик качества предложенных оценок масштабного параметра. Приведены результаты сравнения оценок в условиях гауссовской модели с масштабным засорением.

Полученные результаты приводят к следующему выводу. В тех случаях, когда нет уверенности, что исходное распределение гауссовское, либо выборка может содержать грубые ошибки (выбросы), для оценки масштабного параметра целесообразнее использовать адаптивное стандартное отклонение $\hat{S}_1(\hat{\alpha})$ либо оценку $\hat{S}_4$ медианы абсолютных разностей. Эти оценки имеют ограниченные функции влияния и, следовательно, «защищены» от наличия выбросов в выборке, и они предпочтительнее по эффективности перед другими рассмотренными оценками в условиях различных моделей наблюдений.

ЛИТЕРАТУРА

1. Bickel P.J., Lehmann E.L. Measures of location and scale // Proc. Prague Symp. Asymptotic Statist. 1973. Prague Charles Univ. 1974. V. 1. P. 25–36.
2. Bickel P.J., Lehmann E.L. Descriptive Statistics for nonparametric models. I. Introduction. II. Location // Ann. Statist. 1975. V. 3. P. 1038–1044; 1045–1069.
3. Bickel P.J., Lehmann E.L. Descriptive statistics for nonparametric models. III. Dispersion // Ann. Statist. 1976. V. 4. № 6. P. 1139–1158.
4. Bickel P.J., Lehmann E.L. Descriptive statistics for nonparametric models. IV. Spread // Contributions to Statistics. Hajek Memorial Volume / ed. by J. Jureckova. Prague : Academia, 1979. P. 33–40.
5. Шуленин В.П. Робастные альтернативы стандартному отклонению при обработке данных физических экспериментов // Известия вузов. Физика. 2016. Т. 59, № 6. С. 62–69.
6. Хампель Ф., Рончетти Э., Рауссей П., Штаэль В. Робастность в статистике. Подход на основе функций влияния. М. : Мир, 1989. 512 с.
7. Хьюбер П. Робастность в статистике. М. : Мир. 1984. 304 с.
8. Шуленин В.П. Робастные методы математической статистики. Томск : Изд-во НТЛ, 2016. 260 с.
9. Serfling R.J. Approximation Theorems of Mathematical Statistics. New York : Wiley, 1980. 371 p.
10. Serfling R.J. Generalized L-M-R-statistics // Ann. Statist. 1984. V. 12. P. 76–86.
11. Tarasenko F.P., Shulenin V.P. Connection of MD-estimates with classes of robust estimates of location parameter // 12th Prague Conf. Inf. Theory, Stat. Decision Functions, Random Processes. Prague, 1994. P. 220–223.
12. Jureckova J. M-L-R-estimators // Handbook of Statistics / eds. P.R. Krishnaiah, P. Sen. Elsevier Science Publishers, 1984. V. 4. P. 463–485.
13. Hampel F.R. The influence curve and its role in robust estimation // J. Amer. Statist. Assoc. 1974. V. 69, № 346. P. 383–393.
14. Hampel F.R. Contribution to the theory of robust estimation : Ph. D. diss. Berkeley : Univ. California, 1968. 103 p.
15. Hampel F.R. A general qualitative definition of robustness // Ann. Math. Statist. 1971. V. 42. P. 1887–1896.
16. Janssen P., Serfling R., Veraverbeke M. Asymptotic normality for a general class of statistical functions and applications to measures of spread // Ann. Statist. 1984. V. 12, № 4. P. 1369–1379.
17. Janssen P., Serfling R., Veraverbeke M. Asymptotic normality of U-statistics based on trimmed samples // J. Statist. Planning and Inference. 1987. V. 16. P. 63–74.

18. Шуленин В.П. Асимптотические свойства GL и U- статистик // Вестник Томского государственного университета. 2004. Приложение № 9 (11). С. 184–190.
19. Shulenin V.P. Asymptotic Properties of the Trimmed GL-and U-Statistics // PRAGUE STOCHASTICS'98 : Abstracts 6th Prague Symposium on Asymptotic Statistics, Prague, 1998, August 23–28. P. 84.
20. Hogg R.V. Adaptive robust procedures: a partial review and some suggestions for future applications an theory // J. Amer. Statist. Assoc. 1974. V. 69. P. 909–923.
21. Шуленин В.П. Свойства адаптивных оценок Ходжеса–Лемана в асимптотике и при конечных объемах выборки // Вестник Томского государственного университета. Управление, вычислительная техника и информатика. 2010. № 2 (11). С. 96–112.

Поступила в редакцию 5 июня 2020 г.

Shulenin V.P. (2021) ASYMPTOTIC PROPERTIES AND ROBUSTNESS OF TRIMMED VERSIONS ESTIMATES OF STANDARD DEVIATION AND MEAN ABSOLUTE DEVIATIONS. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika*. [Tomsk State University Journal of Control and Computer Science]. 55. pp. 91–102

DOI: 10.17223/19988605/55/11

In the work, various estimates of the scale parameter characterizing the “spread” of a random variable are studied. It is shown that traditionally used in practice estimates of the scale parameter, such as a sample estimate of the standard deviation, and an estimate of the average absolute deviations, have unlimited influence functions, and they are very sensitive to the presence of outliers in the sample. The paper proposes truncated versions of these estimates and, which are calculated not from the initial sample, but based on ordered statistics, from which the smallest and largest ordinal statistics are previously removed. These estimates are “protected” from the presence of outliers in the sample, they have limited influence functions and their characteristics depend significantly on the parameter α , which in practice leads to additional efforts to select this parameter. In the work, adaptive versions of these estimates are proposed, for which the parameter is determined based on the initial sample. The results of comparing estimates under the conditions of various observation models, in particular, under the conditions of the Gaussian model with large-scale contaminating, are presented. The results obtained lead to the following conclusion. In cases where there is no certainty that the initial distribution is Gaussian, or the sample may contain gross errors (outliers), it is more advisable to use the adaptive standard deviation or the estimate in the form of an estimate of the median of the absolute differences. These estimates have limited influence functions and, therefore, are “protected” from the presence of outliers in the sample, and they are preferable in terms of efficiency over the other estimates considered under the conditions of various observation models.

Keywords: scale parameter; robust estimates; outliers; influence function; adaptive estimates.

SHULENIN Valery Petrovich (Candidate of Technical Sciences, National Research Tomsk State University, Tomsk, Russian Federation).

E-mail: shulenin-vp@rambler.ru

REFERENCES

1. Bickel, P.J. & Lehmann, E.L. (1974) Measures of location and scale. In: *Proc. Prague Symp. Asymptotic Statist.* 1973. Vol. 1. Prague: Charles Univ. pp. 25–36.
2. Bickel, P.J. & Lehmann, E.L. (1975) Descriptive Statistics for nonparametric models. I. Introduction. II. Location. *Ann. Statist.* 3, pp. 1038–1044; pp. 1045–1069.
3. Bickel, P.J. & Lehmann, E.L. (1976) Descriptive statistics for nonparametric models. III. Dispersion. *Ann. Statist.* 4(6). pp. 1139–1158.
4. Bickel, P.J. & Lehmann, E.L. (1979) Descriptive statistics for nonparametric models. IV. Spread. In: Jureckova, J. (ed.) *Contributions to Statistics. Hajek Memorial Volume*. Prague: Academia. pp. 33–40.
5. Shulenin, V.P. (2016) Robust alternatives to standard deviation in physical data processing experiments. *Izvestiya vuzov. Fizika – Russian Physics Journal*. 59(6). pp. 62–69.
6. Hampel, F., Ronchetti, E., Raussey, P. & Stael, V. (1989) *Robastnost' v statistike. Podkhod na osnove funktsiy vliyaniya* [Robustness in statistics. Function-based approach influence]. Translated from English. Moscow: Mir.
7. Huber, P. (1984) *Robastnost' v statistike* [Robustness in Statistics]. Translated from English. Moscow: Mir.
8. Shulenin, V.P. (2016) *Robastnye metody matematicheskoy statistiki* [Robust Methods of Mathematical Statistics]. Tomsk: NTL.
9. Serfling, R.J. (1980) *Approximation Theorems of Mathematical Statistics*. New York: Wiley.
10. Serfling, R.J. (1984) Generalized L-M-R-statistics. *Ann. Statist.* 12. pp. 76–86.
11. Tarasenko, F.P. & Shulenin, V.P. (1994) Connection of MD-estimates with classes of robust estimates of location parameter. *12th Prague Conf. Inf. Theory, Stat. Decision Functions, Random Processes*. Prague. pp. 220–223.
12. Jureckova, J. (1984) M-L-R-estimators. In: Krishnaiah, P.R. & Sen, P. (eds) *Handbook of Statistics*. Vol. 4. Elsevier Science Publishers. pp. 463–485.

13. Hampel, F.R. (1974) The influence curve and its role in robust estimation. *Journal of American Statist. Association.* 69(346). pp. 383–393.
14. Hampel, F.R. (1968) *Contribution to the Theory of Robust Estimation*. Ph. D. Diss. Berkeley, Univ. California.
15. Hampel, F.R. (1971) A general qualitative definition of robustness. *Annals of Mathematical Statistics.* 42. pp. 1887–1896. DOI: 10.1214/aoms/1177693054
16. Janssen, P., Serfling, R. & Veraverbeke, M. (1984) Asymptotic normality for a general class of statistical functions and applications to measures of spread. *Annals of Statistics.* 12(4). pp. 1369–1379. DOI: 10.1214/aos/1176346797
17. Janssen, P., Serfling, R. & Veraverbeke, M. (1987) Asymptotic normality of U-statistics based on trimmed samples. *Journal of Statist. Planning and Inference.* 16. pp. 63–74. DOI: 10.1016/0378-3758(87)90056-5
18. Shulenin, V.P. (2004) Asymptotic properties of GL and U-statistics. *Vestnik Tomskogo gosudarstvennogo universiteta – Tomsk State University Journal.* 9(11). Application. pp. 184–190.
19. Shulenin, V.P. (1998) Asymptotic Properties of the Trimmed GL-and U-Statistics. *6th Prague Symposium on Asymptotic Statistics.* Prague. August 23–28. “Prague Stochastics'98”. Abstracts. p. 84.
20. Hogg, R.V. (1974) Adaptive robust procedures: A partial review and some suggestions for future applications an theory. *Journal of American Statist. Association.* 69. pp. 909–923.
21. Shulenin, V.P. (2010) Asymptotic and nonasymptotic properties of Hodges - Lehmann adaptive estimators. *Vestnik Tomskogo gosudarstvennogo universiteta. Upravlenie, vychislitel'naya tekhnika i informatika – Tomsk State University Journal of Control and Computer Science.* 2(11). pp. 96–112.