

Математические методы исследования

УДК 519.24

ОБОБЩЕННАЯ sech^k -ПЛОТНОСТЬ РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ: РОБАСТНЫЕ СВОЙСТВА, ОЦЕНКИ И ПРИМЕНЕНИЕ

© А. В. Овсянников¹*Статья поступила 28 декабря 2016 г.*

Для адекватного вероятностного описания объектов статистических исследований удобно использовать обобщенные плотности распределения вероятностей, позволяющие оставаться в рамках одной и той же математической модели при изменении параметров плотности, влияющих на ее форму. В статье приведено определение обобщенной sech^k -плотности распределения вероятностей, ее основных характеристик: функции распределения, характеристической функции, дисперсии, коэффициента эксцесса, дифференциальной энтропии, информации Фишера относительно параметра смещения. Показано, что обобщенная sech^k -плотность является удобной вероятностной моделью, которую можно использовать в прикладной статистике в интервале значений эксцесса от нуля до трех. Sech^k -плотность является гладкой функцией и минимизирует информацию Фишера на классе плотностей. Проведено исследование робастных свойств обобщенной sech^k -плотности в применении к алгоритмам оценивания смещения при известных и неизвестных параметрах плотности. Введен в рассмотрение подкласс сниженных робастных процедур оценивания — естественно сниженные оценки. Приведены количественные характеристики чувствительности к большой ошибке и чувствительности к изменению асимптотической дисперсии. Рассмотренные некоторые примеры возможного использования sech^k -плотности позволяют говорить о целесообразности ее применения в теоретической и прикладной статистике.

Ключевые слова: sech^k -плотность; робастность; нелинейное преобразование; робастная оценка; сниженная оценка; чувствительность.

Адекватное вероятностное описание объектов статистических исследований, помех в каналах передачи информации, шумов (погрешностей) измерительных экспериментов требует применения разнообразных математических моделей распределений [1]. В ряде случаев удобно использовать обобщенные распределения как универсальный инструментальный, позволяющий при изменении некоторых параметров, влияющих на форму распределения (плотности вероятности, далее — плотности), оставаться в рамках одной и той же математической модели.

В соответствии с новой парадигмой прикладной статистики [2] целесообразно применение «гладких» плотностей, обладающих свойствами регулярности [3] и, соответственно, робастности относительно некоторого их обобщенного класса [4–9]. В сочетании с непараметрическими методами робастно-адаптивные процедуры, ориентированные на обобщенные распределения, являющиеся при этом «наихудшими» в некотором классе, позволяют ставить и решать разнообразные задачи прикладной статистики. К таким задачам можно отнести идентификацию шумов (по-

грешностей) измерений с «утяжеленными хвостами» распределений [1], построение робастных оценок на обобщенном классе плотностей шумов (погрешностей) измерений, в том числе сниженных оценок и их модификаций [8], построение вероятностных моделей динамики временных рядов, например, моделей логарифмической доходности [10] и др.

Потребность в обобщенных распределениях, обладающих, в частности, свойством робастности на классе распределений, отличающихся простотой технической и аппаратно-программной реализации нелинейных преобразований, компактностью аналитических преобразований при работе с ними, вызвана также и тем, что известные на сегодняшний момент обобщенные распределения этими качествами не обладают.

Таким образом, в связи с отмеченными выше обстоятельствами, представляет интерес исследование возможности применения в прикладной статистике обобщенной sech^k -плотности, введенной в работе [11].

Цель работы — исследование робастных свойств обобщенной sech^k -плотности в применении к алгорит-

¹ Белорусский государственный университет, Минск, Республика Беларусь; e-mail: andovs@mail.ru

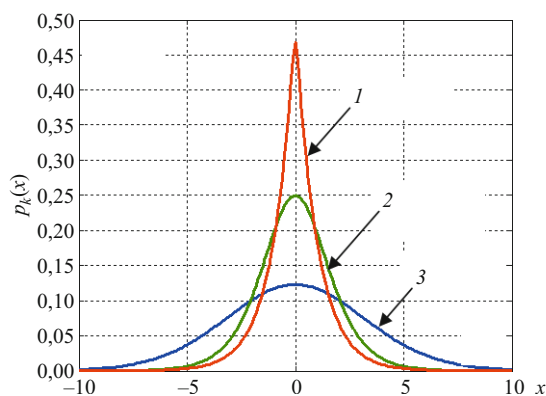


Рис. 1. sech^k -плотность с параметрами $v=1$: 1 — $u=10$, $k=0,1$; 2 — $u=0,5$, $k=2$; 3 — $u=0,1$, $k=10$

мам оценивания смещения при известных и неизвестных параметрах плотности.

Робастность в статье понимается в контексте работ [4–9], в которых исследуется применение методов оценивания, гарантирующих определенное качество, точность оценивания параметров измерительных экспериментов в заданных классах распределений шумов (погрешностей). В этом ключе упоминается о робастности в статистике и в статье [12, с. 62–63].

Обобщенная sech^k -плотность

Обобщенная sech^k -плотность (рис. 1) имеет вид

$$p_k(x) = uC(k)\text{sech}(ux)^k, \quad k = v/u, \quad u, v > 0, \quad (1)$$

где $C(k) = \frac{\Gamma[(1+k)/2]}{\sqrt{\pi}\Gamma(k/2)}$; $\text{sech}(x) = \frac{2}{e^x + e^{-x}}$ — гиперболический секанс [13]; $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} dt$ — гамма-

функция [13]. Модель (1) обобщает полученную в работе [14] плотность для произвольного значения показателя $k > 0$. Обобщенная sech^k -плотность может быть определена как обобщенная гиперболическая секанс-модель степени k (generalized secant hyperbolic of degree k distribution — GSHK модель). Известны и другие модели такого класса: GHS-модель [15] и GSH-модель [16]. Однако практическое использование этих моделей ограничено в основном задачами описательной статистики; существенные трудности возникают при идентификации параметров этих моделей; модели не рассматриваются как распределения, обладающие робастными свойствами на классе, т.е. не могут использоваться в робастных алгоритмах оценивания; нелинейные преобразования моделей, используемые в алгоритмах оценивания, имеют сложный вид.

В то же время для плотности (1) нелинейное преобразование компактно:

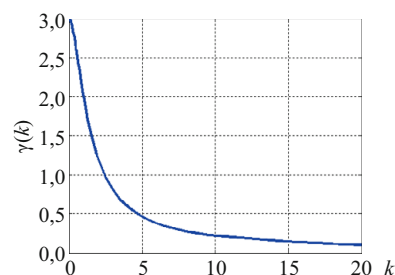


Рис. 2. Коэффициент эксцесса sech^k -плотности

$$z(x) = -\frac{d \ln p_k(x)}{dx} = v \text{th}(ux), \quad (2)$$

где $\text{th}(x) = \frac{1 - e^{-2x}}{1 + e^{-2x}}$ — гиперболический тангенс [13], и

легко может быть реализовано аппаратно-техническими средствами [11]. В аналоговой схемотехнике, например, функция гиперболического тангенса $\text{th}(x)$ реализуется на базе дифференциального усилителя. Плотность (1) включает в себя частные случаи: распределение Чампернауна [17] $p_1(x) = u \text{sech}(ux)/\pi$ и логистическое распределение $p_2(x) = u \text{sech}(ux)^2/2$. Коэффициент эксцесса плотности (1) ограничен (рис. 2; табл. 1, характеристика 4): $\lim_{k \rightarrow 0} \gamma(k) = 3$ [соответствует лапласовской плотности $p_0(x) = (v/2)\exp(-v|x|)$] и $\lim_{k \rightarrow \infty} \gamma(k) = 0$ [соответствует гауссовской плотности $p_\infty(x) = \sqrt{vu/2\pi} \exp(-vux^2/2)$]. Для распределения Чампернауна коэффициент эксцесса $\gamma(k=1) = 2$, а для логистического — $\gamma(k=2) = 1,2$. Основные статистические характеристики sech^k -плотности приведены в табл. 1, где

$$B(x, a, b) = \int_0^x t^{a-1} (1-t)^{b-1} dt$$

неполная бета-функция [13]; ${}_2F_1$, ${}_4F_3$, ${}_6F_5$ — обобщенные гипергеометрические функции [13].

Робастные свойства sech^k -плотности

Рассмотрим типичную в прикладной статистике задачу оценивания неизвестного скалярного параметра $\theta \in \Theta$ (Θ — интервал на действительной оси) в аддитивном шуме. Модель наблюдения: $r = \theta + x$, где x — шум измерительного эксперимента, имеющий плотность распределения $p = p(x)$. Если априорные сведения об оцениваемом параметре отсутствуют (p_θ неизвестна), то в алгоритмах оценивания используют функцию правдоподобия $p_r(r|\theta)$, которая в данном случае будет определяться статистикой шума: $p_r(r|\theta) = p_r(r - \theta) = p(x)$. Далее зависимости функций от аргументов будут опускаться, если они не вызывают сомнений.

Согласно неравенству Крамера – Рао [3] при использовании любых оценок θ_n^* по результатам n однородных независимых статистических регулярных экспериментов [3] нижняя граница дисперсии ошибки оценки имеет вид

$$D_{\theta} = E(\theta - \theta_n^*)^2 \geq \frac{1}{nI_{\theta}}, \quad (3)$$

где

$$I_{\theta} = E \left[\frac{\partial \ln p_r(r|\theta)}{\partial \theta} \right]^2 = \int \left(-\frac{p'_x}{p} \right) dx = \int \frac{p_x'^2}{p} dx \quad (4)$$

информация Фишера, определяющая количество информации об оцениваемом параметре θ , содержащееся в распределении $p_r(r|\theta) = p_r(r - \theta)$; E — символ усреднения по множеству реализаций. В формуле (4) использовано соотношение

$$\frac{\partial \ln p_r(r - \theta)}{\partial \theta} = -\frac{p'_x(x)}{p(x)}.$$

В реальных условиях шум измерительного эксперимента является комбинацией множества составляющих [1]. Это не позволяет дать его четкое вероятностно-статистическое описание в виде конкретного закона распределения. Один из возможных путей преодоления этой неопределенности — применение робастных процедур оценивания на классе распределений шумов P . Такой класс определяется заданием ограничений в виде уравнений связи $\phi_m(x, p, p'_x) = 0$, $m = 0, M - 1$ и/или дополнительных условий

$$\int f_s(x, p, p'_x) dx \leq c_s, \quad s = \overline{0, S - 1}, \quad (5)$$

где c_s — заданные постоянные.

Далее в классе P определяется распределение, доставляющее минимум информации Фишера (4) — p^* , на которое и «настраивается» алгоритм оценивания. При этом согласно неравенству Крамера – Рао (3) гарантируется, что ошибка оценивания неизвестного параметра не превзойдет ту, которая могла бы быть в наихудшем случае: $D_{\theta}(p, P) \leq D_{\theta}(p^*, P)$, где $p, p^* \in P$.

Нахождение распределения, минимизирующего информацию Фишера, происходит путем решения вариационной задачи [4–7, 9, 12]. С учетом уравнений связи и дополнительных условий (5) составляется функция Лагранжа [13]

$$L = \frac{p_x'^2}{p} + \sum_m \mu_m \phi_m + \sum_s \lambda_s f_s, \quad (6)$$

где μ_m, λ_s — множители Лагранжа, и записывается уравнение Эйлера [13]

$$\left[\frac{d}{dx} \left(\frac{\partial L}{\partial p'_x} \right) - \frac{\partial L}{\partial p} \right]_{p=p^*} = 0. \quad (7)$$

Совместное решение обыкновенного дифференциального уравнения (7) с $M + S$ -уравнениями ограничений позволяет найти распределение $p^* \in P$ и определить неизвестные множители μ_m, λ_s .

Так, в классе всех невырожденных распределений $P_{\varepsilon} = \{p: 0 < \varepsilon \leq p(0)\}$ распределением, минимизирующим информацию Фишера, является двойное экспоненциальное (лапласовское) распределение с плотностью $p^* = p_0(x)$ и $I_{\theta}(p^* = p_0) = \nu^2 \leq I_{\theta}(p)$, $p, p^* \in P_{\varepsilon}$. В классе распределений с ограниченной дисперсией

$$P_{\sigma} = \left\{ p: \int x^2 p dx \leq \sigma^2 = \frac{1}{\nu^2} \right\} \quad \text{—}$$

гауссовское распределение с плотностью $p^* = p_{\infty}(x)$ и $I_{\theta}(p^* = p_{\infty}) = 1/\sigma^2 \leq I_{\theta}(p)$, $p, p^* \in P_{\sigma}$. Для хьюберов-

Таблица 1. Основные характеристики sech^k -плотности

1	Функция распределения: $\frac{2^k}{k} C(k) e^{kux} {}_2F_1 \left(\left[\frac{k}{2}, k \right], \left[1 + \frac{k}{2} \right], -e^{2ux} \right)$
2	Характеристическая функция: $2^{k-1} C(k) \exp \frac{\pi(i+juk)}{2u} \left(B \left[-1, \frac{1}{2} \left(k - \frac{j}{u} \right), 1-k \right] + e^{i\pi/4} B \left[-1, \frac{1}{2} \left(k + \frac{j}{u} \right), 1-k \right] \right)$
3	Дисперсия D : $\frac{2^{k+2} C(k)}{\nu^2 k} {}_4F_3 \left(\left[\frac{k}{2}, \frac{k}{2}, \frac{k}{2}, k \right], \left[\frac{k}{2} + 1, \frac{k}{2} + 1, \frac{k}{2} + 1 \right], -1 \right)$
4	Коэффициент эксцесса $\gamma(k) \in [0, \dots, 3]$: $\frac{3k 2^{-k} {}_6F_5 \left(\left[\frac{k}{2}, \frac{k}{2}, \frac{k}{2}, \frac{k}{2}, \frac{k}{2}, k \right], \left[\frac{k}{2} + 1, \frac{k}{2} + 1, \frac{k}{2} + 1, \frac{k}{2} + 1, \frac{k}{2} + 1 \right], -1 \right)}{C(k) {}_4F_3 \left(\left[\frac{k}{2}, \frac{k}{2}, \frac{k}{2}, k \right], \left[\frac{k}{2} + 1, \frac{k}{2} + 1, \frac{k}{2} + 1 \right], -1 \right)^2} - 3$
5	Количество информации Фишера относительно смещения I_{θ} : $\frac{\nu^2}{1+k}$
6	Энтропия $H(x)$: $-\ln[uC(k)] - C(k)J(k)$, $J(k) = \int_{-\infty}^{\infty} \text{sech}(y)^k \ln[\text{sech}(y)^k] dy$

ского класса приближенно нормальных распределений $P_\alpha = \{p: p = (1 - \alpha)N(x) + \alpha G(x)\}$, где $0 < \alpha < 1$ — коэффициент засорения; $N(x)$ — гауссовская плотность; $G(x)$ — произвольная плотность; распределение, минимизирующее информацию Фишера, представляет собой сложную композицию гауссовской и лапласовской компоненты [4, 7].

К настоящему времени определено и исследовано множество классов распределений, некоторые из них приведены в работах [4–9]. Выбор того или иного класса распределений шумов непосредственно связан с особенностями решаемой прикладной задачи и объемом имеющейся относительно этих шумов априорной информации. Однако если измерительный эксперимент протекает в условиях, характеризующихся нестабильностью статистики шума, изменение структуры алгоритма оценивания на другую структуру требует определенных программно-аппаратных или технических преобразований. Причем очевидно, что реализация робастного алгоритма оценивания, «настроенного» на более широкий класс распределений шумов, также неэффективна из-за слишком сильного «огрубления» алгоритма. В этой связи целесообразно использование такого класса распределений шумов, который отвечал бы поставленной задаче оценивания и «наихудшее» распределение в котором имело бы достаточно общий вид. В этом случае вместо перестройки алгоритма оценивания может потребоваться лишь адаптация к изменившимся параметрам шума. Проверим распределение с sech^k -плотностью на выполнение этих требований.

Следующая теорема устанавливает основное свойство sech^k -плотности, которое может быть использовано при построении робастных процедур оценивания скалярного параметра смещения $\theta \in \Theta$ при аддитивном шуме x и модели наблюдения $r = \theta + x$, $p_r(r|\theta) = p(x) = p$.

Теорема 1. Плотность (1) $p_k(x) = p^*$ минимизирует количество информации Фишера (4) на классе P_k регулярных, не финитных на носителе $x \in (-\infty, \infty)$ гладких плотностей p , для которого существует и конечная величина

$$\int z^{*2} p dx \leq c,$$

где $z^* = -(p'_x)/p^*$, $c > 0$ — константа; в условиях $z^*(0) = 0$, $p^*(0) = \varepsilon > 0$.

Доказательство. Функция Лагранжа (6) оптимизационной задачи с учетом нормировки $\int p dx = 1$ принимает вид

$$L = \frac{p'^2_x}{p} + \lambda_0 p + \lambda_1 z^{*2} p,$$

где λ_0, λ_1 — множители Лагранжа. Соответствующее функции L уравнение Эйлера (7): $-2z'_x + z^2 - \lambda_0 -$

$-\lambda_1 z^{*2} = 0$. Поскольку минимум I_θ достигается при $z = z^*$, то из последнего уравнения следует

$$-2(z'_x)^* + (1 - \lambda_1) z^{*2} - \lambda_0 = 0.$$

Рассмотрим, при каких значениях множителей Лагранжа $\lambda_j \neq 0$, $j = 0, 1$, решения этого дифференциального уравнения будут удовлетворять условиям сформулированной теоремы. Варианты значений множителей и соответствующие решения следующие:

- 1) $\lambda_j > 0$ — $z^* = -\sqrt{\frac{\lambda_0}{\lambda_1 - 1}} \text{tg} \frac{x\sqrt{\lambda_0(\lambda_1 - 1)}}{2}$;
- 2) $\lambda_j < 0$ — $z^* = \sqrt{\frac{\lambda_0}{\lambda_1 + 1}} \text{tg} \frac{x\sqrt{\lambda_0(\lambda_1 + 1)}}{2}$;
- 3) $\lambda_0 > 0, \lambda_1 < 0$ — $z^* = -\sqrt{\frac{\lambda_0}{\lambda_1 + 1}} \text{thx} \frac{\sqrt{\lambda_0(\lambda_1 + 1)}}{2}$;
- 4) $\lambda_0 < 0, \lambda_1 > 0$ — $z^* = \sqrt{\frac{\lambda_0}{\lambda_1 - 1}} \text{thx} \frac{\sqrt{\lambda_0(\lambda_1 - 1)}}{2}$.

Поскольку нелинейное преобразование $z^* = -(p'_x)^*/p^* = -d \ln p^*/dx$, то общее выражение для функции p^* имеет вид

$$p^* = \varepsilon \exp\left(-\int z^* dx\right).$$

Исходя из этого, в случаях 1 и 2 получаем

$$p_{1,2}^* = \varepsilon \left[\cos \frac{x\sqrt{\lambda_0(\lambda_1 \mp 1)}}{2} \right]^{\pm 2/(\lambda_1 \mp 1)}.$$

Однако эти функции с учетом нормировки ограничены по переменной x и не удовлетворяют условиям теоремы.

В случаях 3 и 4 имеем общее решение следующего вида:

$$p_{3,4}^* = \varepsilon \left[\text{ch} \frac{x\sqrt{\lambda_0(\lambda_1 \pm 1)}}{2} \right]^{\pm 2/(\lambda_1 \pm 1)},$$

где $\text{ch}(x) = (e^x + e^{-x})/2$ — гиперболический косинус [13].

Выберем значения множителей Лагранжа: $\lambda_0^{3,4} = \mp 2\nu u$, $\lambda_1^{3,4} = \mp 1 \mp 2u/\nu$. В этих случаях получим

$$p_{3,4}^* = \varepsilon \text{ch}(ux)^{-\nu/u} = \varepsilon \text{sech}(ux)^k, \quad k = \nu/u,$$

где $\text{sech}(x) = 1/\text{ch}(x)$ — гиперболический секанс. Эта функция удовлетворяет условиям теоремы. Из условия нормировки находим константу $\varepsilon = uC(k)$. Следовательно, функция $p_{3,4}^* = p_k(x)$ является плотностью распределения, которая минимизирует количество информации Фишера (4) на классе распределений P_k с $I_\theta = \nu^2/(1+k)$. Таким образом, теорема доказана.

Следствие. Пусть имеется n результатов однородных независимых статистических регулярных экспериментов с моделью наблюдения $r = \theta + x$. Тогда любая оценка θ_n^* на классе распределений P_k будет иметь дисперсию ошибки, определяемую неравенством (3):

$$D_{\theta} \leq D_{\max} = \frac{1+k}{nv^2}. \quad (8)$$

Для сравнительного анализа (8) удобно ввести величину эквивалентной дисперсии ошибки оценки по результатам n измерительных экспериментов $D_{\max} = D_{\max}(\sigma^2/n)$, где σ^2/n — дисперсия ошибки оценки в гауссовском шуме. В табл. 2 представлены дисперсии распределений $p_k(x)$ с параметром $k=0, 2$ и соответствующие им эквивалентные дисперсии $D_{\max}(\sigma^2/n)$. По результатам, представленным в табл. 2, отмечаем увеличение эквивалентной дисперсии ошибки с ростом значения k . В предельном случае гауссовского распределения шума дисперсия ошибки оценки по результатам n наблюдений составляет $D_{\max}(\sigma^2/n) = \sigma^2/n$. Такой результат свидетельствует о том, что использование нелинейного преобразования (2) в робастных алгоритмах оценивания позволяет варьировать их точность (8) в достаточно широком диапазоне

$$D_{\theta} \leq D_{\max} \left(\frac{\sigma^2}{n} \right) \in \frac{\sigma^2}{n} \left[\frac{1}{2}, \dots, 1 \right].$$

В сочетании с возможностями реализации робастно-адаптивных процедур оценивания эффективность применения нелинейного преобразования (2) возрастает.

Следует отметить и тот факт, что в отличие от известных классов распределений [4–9] класс P_k имеет наихудшее в информационном смысле распределение, являющееся обобщенным распределением с гладкой функцией нелинейного преобразования и с четкими предельными функциональными границами семейства: от лапласовского $[p_0(x)]$ до гауссовского $[p_{\infty}(x)]$.

Замечание. В условиях теоремы 1 предполагалась неограниченность плотности по переменной x . Если снять это предположение, то «наихудшей» функцией, минимизирующей фишеровскую информацию на классе P_k , будет также и функция $p_{1,2}^*$. Выберем

$\lambda_0^{1,2} = \mp 2\nu u$, $\lambda_1^{1,2} = \pm 1 \mp 2u/\nu$ и, найдя константу ϵ из условия нормировки, получим

$$p_k^{fin}(x) = p_{1,2}^* = \frac{u\Gamma(1+k/2)}{\sqrt{\pi}\Gamma[(1+k)/2]} \cos(ux)^k, \quad k = \nu/u.$$

Эта функция является плотностью при ограниченном значении аргумента $x \in (-\pi/2u; \pi/2u)$. Плотность $p_k^{fin}(x)$ оказывается финитной, гладкой и так же, как и плотность (1), — обобщенной. Частным случаем плотности $p_k^{fin}(x)$ является, например, плотность [9]

$$p_2^{fin}(x) = \frac{1}{l} \cos\left(\frac{\pi x}{2l}\right)^2$$

для $|x| \leq l$ и $p_2^{fin}(x) = 0$ для $|x| > l$; здесь $l = \pi/2u$. Исследование свойств и характеристик полученной обобщенной финитной плотности $p_k^{fin}(x)$ выходит за рамки данной работы и может быть предоставлено заинтересованному читателю.

Оценка параметров sech^k -плотности

Рассмотрим оценку параметра смещения θ sech^k -плотности

$$p_r(r|\theta) = p_k(r - \theta) = uC(k) \text{sech}[u(r - \theta)]^k$$

при известных значениях ν и u . Такая оценка может быть получена с использованием одношаговой процедуры (ОШ-оценка) [18] или метода максимального правдоподобия (МП-оценка). И в том и другом случае нелинейное преобразование при аддитивной модели наблюдения $r_i = \theta + x_i$ определяется функцией (2) в виде $z(r_i - \theta) = \nu \text{th}[u(r_i - \theta)]$. Обозначим

$$Z_n(\theta) = \frac{\partial \ln p_r(\mathbf{r}|\theta)}{\partial \theta} = \sum_{i=1}^n z(r_i - \theta),$$

где $\mathbf{r} = [r_1, \dots, r_n]$ — данные наблюдений, n — объем данных. Тогда МП-оценка

$$\theta_n^* = \theta_{n-1}^* + \frac{z(r_n - \theta_{n-1}^*)}{nI_{\theta}}, \quad \theta_0^* = \theta^0, \quad (9)$$

ОШ-оценка

$$\theta_n^* = \theta^0 + \frac{Z_n(\theta^0)}{nI_{\theta}}, \quad (10)$$

где θ^0 — некоторая начальная (предварительная) оценка. Целесообразность использования той или иной

Таблица 2. Эквивалентные дисперсии ошибки оценки (8)

	$k=0$	$k=1$	$k=2$
Дисперсия распределения $p_k(x)$	$\frac{2}{v^2}$	$\frac{1}{4} \left(\frac{\pi}{u} \right)^2 = \frac{1}{4} \left(\frac{\pi}{v} \right)^2$	$\frac{1}{12} \left(\frac{\pi}{u} \right)^2 = \frac{1}{48} \left(\frac{\pi}{v} \right)^2$
Эквивалентная дисперсия ошибки $D_{\max}(\sigma^2/n)$	$\frac{1}{2} \frac{\sigma^2}{n}$	$\frac{8}{\pi^2} \frac{\sigma^2}{n}$	$\frac{9}{\pi^2} \frac{\sigma^2}{n}$

оценки определяется практическими соображениями в конкретной ситуации и возможностями применения аппаратно-программных средств. Свойства оценок (9), (10) достаточно полно описаны в [3–9, 18]: оценки состоятельны, асимптотически эффективны и асимптотически нормальны. Оценки обладают «сильными» робастными свойствами, следующими из особенностей нелинейного преобразования (2).

Функция влияния [7, 8], определяющая асимптотическое смещение приведенных оценок (9), (10), вызываемое «загрязнениями» в наблюдаемых данных, имеет вид

$$IF(x, p) = \frac{\eta}{\int \eta^2 p dx}, \quad (11)$$

где $\eta = \eta(x)$ — нелинейное преобразование. Для соответствующих оценок на основании (11) получаем МП-оценку

$$IF_{МП}(x, p) = \frac{z(x)}{I_\theta} = \left(\frac{1}{v} + \frac{1}{u} \right) \text{th}(ux), \quad (12)$$

ОШ-оценку

$$IF_{ОШ}(x, p) = \frac{Z(x)}{nI_\theta}. \quad (13)$$

Поскольку функции z , Z_n и, соответственно, функции влияния ограничены (функция гиперболического тангенса ограничена $|\text{th}(ux)| = v$), то приведенные МП- и ОШ-оценки являются B -робастными, т.е. робастными относительно смещения (bias). Предельная чувствительность к большой ошибке определяется верхней границей функции влияния

$$\rho^* = \sup |IF(x, p)|. \quad (14)$$

Для рассматриваемых оценок это одна и та же величина. Подставляя (12), (13) в (14), получаем

$$\rho_{МП}^* = v \frac{\sup |\text{th}(ux)|}{I_\theta} = \frac{1}{v} + \frac{1}{u} = \frac{1+k}{ku},$$

$$\theta_{ОШ}^* = \frac{\sup |Z(x)|}{nI_\theta} = \frac{v}{nI_\theta} \sup \left| \sum_{i=1}^n \text{th}(ux_i) \right| = \frac{1}{v} + \frac{1}{u} = \frac{1+k}{ku}.$$

В робастной статистике, кроме понятия B -робастности, определяется понятие V -робастности (от variance — дисперсия). Согласно работе [8] для неубывающих функций, которыми являются функции z , Z_n , понятия V -робастности и B -робастности эквивалентны. Оценка является V -робастной, когда существует и конечна величина ее чувствительности к изменению асимптотической дисперсии. Эта чувствительность определяется формулой

$$\chi^* = 1 + \rho^{*2} I_\theta = 2 + k.$$

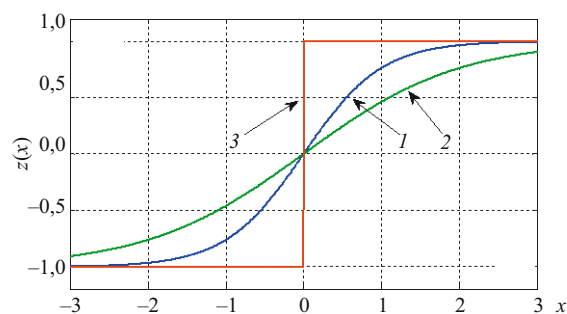


Рис. 3. Нелинейные преобразования с параметром $v = 1$: 1 — $z(x) = \text{th}(x)$, $k = 1$; 2 — $z(x) = \text{th}(0,5x)$, $k = 2$; 3 — $z_{MED}(x) = \text{sign}(x)$

Таким образом, рассматриваемые оценки имеют не только конечные чувствительности ρ^* и χ^* , но и являются наиболее робастными, сходящимися при $u \rightarrow \infty$ к медианной [4, 7, 8] с нелинейным преобразованием: $z_{MED}(x) = v \text{sign}(x)$ (рис. 3), где функция знака $\text{sign}(x) = \{-1, x < 0; 0, x = 0; 1, x > 0\}$. Если параметр sech^k -плотности $k = v/u = 1$ [распределение Чамперна — $p_1(x)$], то характеристики чувствительности составят $\rho^* = 2/u = 2/v$, $\chi^* = 3$. В случае логистического распределения $p_2(x)$ с $k = v/u = 2$ получаем $\rho^* = 3/2u = 3/v$, $\chi^* = 4$.

В случае неизвестности параметров u и v ОШ-оценка вектора неизвестных $\mathbf{Y} = [y_1, y_2, y_3] = [\theta, u, v]$ имеет вид

$$\mathbf{Y}_n^* = \mathbf{Y}^0 + \mathbf{I}_Y^{-1} Z_n(\mathbf{Y}^0)/n, \quad (15)$$

где $\mathbf{Y}^0$ — некоторая начальная (предварительная) оценка вектора неизвестных параметров (в качестве u^0, v^0 могут быть выбраны значения, согласующиеся с некоторой априорной информацией относительно них);

$$Z_n(\mathbf{Y}^0) = \left[\sum_{j=1}^n \frac{\partial \ln p_k(r|y_j)}{\partial y_j} \right], \quad i = \overline{1, 3}, \quad —$$

вектор-столбец с элементами, соответствующими оцениваемому параметру; $\mathbf{I}_Y = [E \partial^2 \ln p_k(x, \mathbf{Y}) / \partial y_i \partial y_j]$, $i, j = \overline{1, 3}$, — информационная матрица Фишера.

Ввиду аналитической громоздкости получающихся точных выражений для информационной матрицы $\mathbf{I}_Y$ и, соответственно, уравнений (15) можно воспользоваться более простым приближенным подходом, заключающимся в следующем. Оценка выборочного коэффициента эксцесса γ_n^* с учетом рис. 2 (см. табл. 1, характеристика 4) позволяет определить оценку параметра k^* . В свою очередь, оценка выборочной дисперсии D_n^* и определенное ранее значение k^* позволяют найти параметр v^* (табл. 1, характеристика 3) и параметр $u^* = v^*/k^*$. Таким образом, в алгоритмах (9), (10) значения u и v следует заменить на u^* и v^* . Качество получающейся оценки смещения θ_n^* будет опреде-

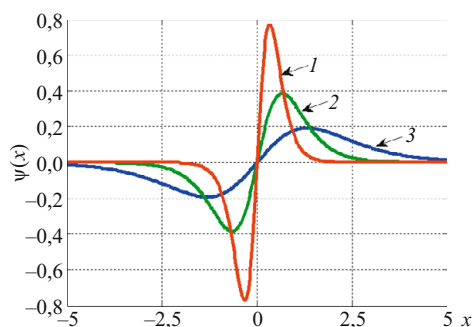


Рис. 4. Нелинейное преобразование ψ с параметрами $v = 1$: 1 — $u = 10, k = 0,1$; 2 — $u = 0,5, k = 2$; 3 — $u = 0,1, k = 10$

ляться точностью вычислений выборочных значений дисперсии D_n^* и эксцесса γ_n^* .

Сниженная оценка смещения sech^k -плотности

Рассмотрим применение sech^k -плотности для построения робастных процедур оценивания в подклассе сниженных оценок [8]. Эффект их использования заключается в снижении влияния на результирующую оценку больших отклонений выборочных данных.

В обширной литературе, посвященной этому вопросу, предлагается большое количество разнообразных эмпирических и искусственно сконструированных нелинейных преобразований, используемых в процедурах, реализующих такие оценки. Однако возникает вопрос: можно ли получать нелинейные преобразования сниженных оценок, формируя их из уже известных, полученных для робастных оценок? Такие оценки с нелинейными преобразованиями, базирующимися на уже существующих, удобно называть *естественно сниженными оценками*.

Определение. Естественной сниженной будем называть такую оценку, для которой нелинейное преобразование в процедуре оценки параметра смещения имеет вид $\psi = f(z)$ и обладает свойствами: $|\psi| \rightarrow 0$ при $|x| \rightarrow \infty$ и $\psi(0) = 0$.

Теорема 2. Оценка смещения с нелинейным преобразованием $\psi = z_x z$ является естественно сниженной, если $|z|$ ограничена и $z(0) = 0$.

Доказательство. Поскольку нелинейное преобразование z ограничено снизу и сверху, то производная $z_x \geq 0$ (ограничена снизу) и $|z'_x| \rightarrow 0$ для $|x| \rightarrow \infty$. Следовательно, справедливо $|\psi| = |z'_x z| \rightarrow 0$ для $|x| \rightarrow \infty$. Тогда решение уравнения $E\psi = 0|_{\theta=\theta^*}$ (E — символ математического ожидания) будет соответствовать уравнению оценки параметра смещения θ . Таким образом, теорема доказана.

Замечание 1. Ограниченность производной функции z'_x снизу означает и V -робастность [8].

Замечание 2. Теорема 2 распространяется на любые B -робастные оценки.

Таким образом, трансформируя приведенные выше МП- и ОШ-оценки параметра смещения (9), (10), получаем МП-оценку

$$\theta_n^* = \theta_{n-1}^* + \frac{\psi(r_n - \theta_{n-1}^*)}{nI_\theta}, \quad (16)$$

ОШ-оценку

$$\theta_n^* = \theta^0 + \frac{\Psi_n(\theta^0)}{nI_\theta}, \quad \Psi_n(\theta) = \sum_{i=1}^n \psi(r_i - \theta). \quad (17)$$

Оценки (16), (17) с функцией ψ обладают теми же свойствами, что и оценки с преобразованием z . Функция влияния с ψ , соответствующем sech^k -плотности, $IF_\psi(x, p) = (k+1)(k+3)\text{sech}(ux)^2 \text{th}(ux)/vk$, чувствительность к большой ошибке

$$\rho^* = \frac{2(k+1)(k+3)}{3\sqrt{3}vk}.$$

Чувствительность к изменению асимптотической дисперсии ввиду громоздкости точной формулы может быть аппроксимирована более простой зависимостью на интервале $k \in [0, \dots, 5]$:

$$\chi^* = 1 + \frac{(k+1)(k+3)(k+5)(1-0,015k)}{v(k+2)}$$

со среднеквадратической ошибкой, не превосходящей 0,15.

На рис. 4 приведена функция нелинейного преобразования $\psi(x) = uv^2 \text{th}(ux)\text{sech}(ux)^2$, полученная для sech^k -плотности с различными значениями параметра k при $v = 1$. Нелинейное преобразование ψ обладает линейным участком в диапазоне изменения аргумента

$\Delta x \in [x_1; x_2]$, где $x_{1,2} = \mp \text{asech} \frac{\sqrt{2/3}}{u}$, и спадом, крутиз-

на которого определяется параметрами u и v . С увеличением параметра u функция ψ стремится к урезанной медиане $v \text{sign}(x)[x_1, x_2]$ [8] как к наиболее B -робастному нелинейному преобразованию.

Если параметры u и v неизвестны, целесообразно использовать методику их определения, описанную выше.

Применение sech^k -плотности

Применение sech^k -плотности с параметром $k = 1, 2$ в практике статистического анализа достаточно известно. Так, плотности $p_1(x)$ (распределение Чампернауна), $p_2(x)$ (логистическое распределение) получили распространение в социальных и финансово-экономических исследованиях. Кроме того, известны применения логистического распределения в описательной статистике биосферы, биотехнологии и технических областях, например, для описания прочностных характеристик усталости металлов.

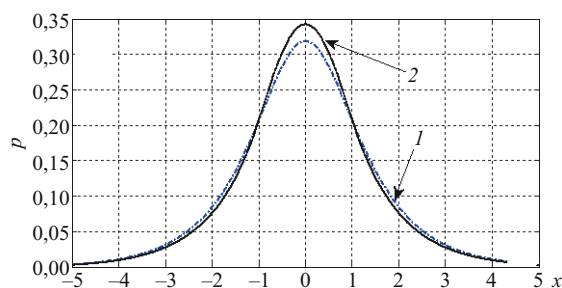


Рис. 5. Плотности распределения: 1 — $p_k(x)$, $k=1$, $u=1$; 2 — $p^*(18)$, $\Delta_x=1$, $\sigma^2=1$

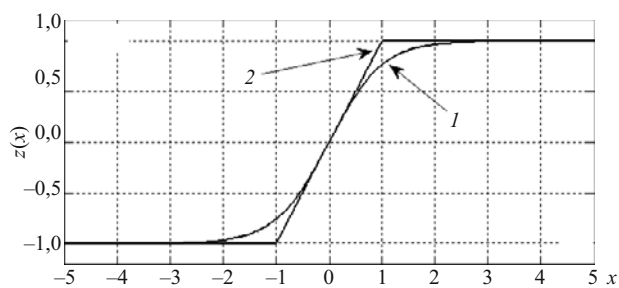


Рис. 6. Нелинейные преобразования: 1 — $z(x) = \text{th}(x)$; 2 — $z_H(x) = \max\{-1; \min(x; 1)\}$

В математической статистике источником возникновения логистического распределения является асимптотическая теория экстремальных значений.

Рассматриваемая в работе обобщенная sech^k -плотность за счет параметра $k = v/u$ представляет более широкий инструментальный в решении задач прикладной и математической статистики. Приведем некоторые области применения свойств и характеристик исследуемой плотности.

1. Как указывалось выше, обобщенная sech^k -плотность может быть отнесена к хьюберовскому классу α -засоренных распределений P_α [7]. Такие распределения представляются в виде линейной комбинации двух плотностей $p(x) = (1 - \alpha)N(x) + \alpha G(x)$ или их переключающихся значений $p(x) = \{(1 - \alpha)N(x), |x| \leq \Delta_x; \alpha G(x), |x| > \Delta_x\}$. Здесь $0 < \alpha < 1$; $N(x) = N(x; 0; \sigma^2)$ — гауссовская плотность с нулевым средним и дисперсией σ^2 ; $G(x)$ — произвольная плотность; $2\Delta_x$ — ширина центрального участка. Эти модели класса α -засоренных распределений $p \in P_\alpha$ получили широкое распространение в практике применения робастных процедур статистической обработки данных при наличии «утяжеленных хвостов» экспериментальных гистограмм. Распределение, минимизирующее информацию Фишера на классе P_α , имеет вид [7]

$$p^* = \begin{cases} \frac{1-\alpha}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right), & |x| \leq \Delta_x, \\ \frac{1-\alpha}{\sqrt{2\pi}} \exp\left(\frac{\Delta_x^2}{2\sigma^2} - \frac{\Delta_x|x|}{\sigma^2}\right), & |x| > \Delta_x, \end{cases} \quad (18)$$

где половина ширины центрального участка Δ_x связана с параметром α соотношением $2N(\Delta_x; 0; 1)/\Delta_x - 2\Phi(-\Delta_x) = \alpha/(1 - \alpha)$, в котором $N(\Delta_x; 0; 1)$ — плотность стандартного нормального распределения;

$$\Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y \exp\left(-\frac{t^2}{2}\right) dt$$

функция распределения.

Нелинейное преобразование, отвечающее распределению (18), имеет вид кусочно-линейной функции

с насыщением $z_H(x) = \sigma^2 \max\{-\Delta_x; \min(x; \Delta_x)\}$. Показано [7, 8], что оценка параметра смещения θ при этом нелинейном преобразовании определяет оптимальную робастную оценку Хьюбера. Однако такое нелинейное преобразование не является гладкой функцией аргумента, а представляет, по сути, удобную аппроксимацию функции (2) $z(x) = v \text{th}(ux)$ с параметрами $k = \Delta_x = v/u$, $\sigma^2 = 1/u$. На рис. 5 приведены две плотности распределения, согласованные по параметрам: кривая 1 соответствует рассматриваемой в работе обобщенной sech^k -плотности $k=1$, $u=1$; кривая 2 — плотности (18) $p^* \in P_\alpha$. На рис. 6 приведены нелинейные преобразования для этих плотностей.

Приведенные результаты показывают, что ОШ- и МП-оценки в случае использования sech^k -плотности оказываются более точными и гибкими с учетом возможности адаптации к параметрам v и u [11].

2. На основании формул табл. 1 получена зависимость энтропийного коэффициента $h = \frac{0,5 \exp[H(x)]}{\sqrt{D}}$

[1] от контрэксцесса $\rho = 1/\sqrt{\gamma}$ в виде линейной функции $h = -0,9 + 3,5\rho$ для значений $k \in [0, 5 \dots 5]$ со среднеквадратической ошибкой, не превосходящей 1,5 %. Такая зависимость полностью согласуется с приближенной топологической классификацией законов распределения погрешностей средств измерений, приведенной в [1].

3. Распределение с плотностью (1) может быть использовано в статистике временных рядов. Так, анализ многочисленных примеров временных рядов (при исключении тренда), приведенных в классической и современной литературе [19–22], показывает лучшую значимость критерия согласия χ^2 (критерия Пирсона) при идентификации для sech^k -плотности в сравнении с гауссовской или лапласовской плотностями.

Похожий результат может наблюдаться и в случаях, когда модель данных измерительного эксперимента не требует оценивания непосредственно самого параметра. Оцениванию подлежат лишь текущие его изменения. Например, это могут быть изменения: уровня шума в канале передачи информации; нагрузок в электросетях промышленного оборудования; переданного или потребленного сетевого трафика; уровня

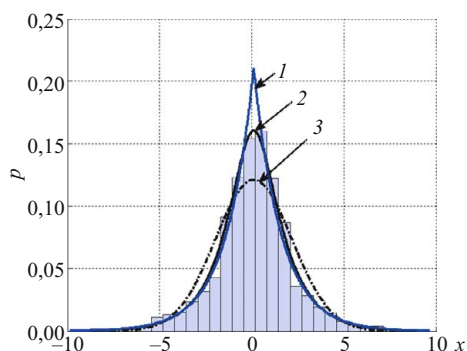


Рис. 7. Аппроксимация гистограммы теоретическими кривыми плотности

социальной активности в группах, и т.п. В частности, в финансовой сфере адекватное вероятностное описание логарифмической доходности ($h_t = \ln(s_t/s_{t-1})$, где s_t — котировки некоторой ценной бумаги) получается именно с использованием sech^k -плотности. На рис. 7 приведен пример гистограммы логарифмической доходности по индексу RTS по 1000 отсчетам временного ряда (источник: <http://www.rts.ru>). Здесь же представлены три «подходящие» кривые плотностей распределения вероятностей: 1 — лапласовская плотность $p_0(x)$; 2 — sech^k -плотность $p_k(x)$; 3 — гауссовская плотность $p_\infty(x)$. Статистический анализ показывает, что для различного числа интервалов группирования [1, 23] выполняется соотношение $\chi_2^2 < \chi_1^2 < \chi_{\text{кр}}^2 < \chi_3^2$, где $\chi_{\text{кр}}^2$ — критическое значение с заданным уровнем значимости при заданном числе степеней свободы. Это соотношение критериев χ^2 свидетельствует о более точном и качественном вероятностном описании экспериментальных статистических данных распределением с sech^k -плотностью.

Таким образом, рассмотренная в работе sech^k -плотность является удобным инструментом, который может быть использован в теоретической и прикладной статистике в интервале значений эксцесса $\gamma \in [0, \dots, 3]$. Sech^k -плотность — обобщенная, гладкая функция на всей области определения $x \in (-\infty, \infty)$, минимизирующая информацию Фишера на классе плотностей (теорема 1).

Построенные с использованием sech^k -плотности алгоритмы МП- и ОШ-оценивания, включая естественно сниженные оценки параметра смещения (теорема 2), являются робастными на классе плотностей. Приведены количественные показатели качества оценок при известных параметрах sech^k -плотности. В случае неизвестности параметров sech^k -плот-

ности могут быть использованы адаптивные схемы их настройки.

Рассмотренные примеры возможного использования sech^k -плотности позволяют говорить о целесообразности ее применения в прикладной статистике.

ЛИТЕРАТУРА

1. Новицкий П. В., Зограф И. А. Оценка погрешностей результатов измерений. Изд. 2-е, перераб. и доп. — Ленинград: Энергоатомиздат, 1991. — 304 с.
2. Орлов А. И. Новая парадигма прикладной статистики / Заводская лаборатория. Диагностика материалов. 2012. Т. 78. № 11. С. 87 — 93.
3. Ибрагимов И. А., Хасьминский Р. З. Асимптотическая теория оценивания. — М.: Наука, 1979. — 528 с.
4. Huber P. J. Robust statistics: A review / The Annals of Mathematical Statistics. 1972. Vol. 43. N 4. P. 1041 — 1067.
5. Ершов А. А. Стабильные методы оценки параметров (обзор) / Автоматика и телемеханика. 1978. № 8. С. 66 — 100.
6. Смоляк С. А., Титаренко Б. П. Устойчивые методы оценивания (Статистическая обработка неоднородных совокупностей). — М.: Статистика, 1980. — 208 с.
7. Хьюбер Дж. П. Робастность в статистике / Пер. с англ. — М.: Мир, 1984. — 304 с.
8. Хампель Ф., Рончетти Э., Рауссеу П., Штаэль В. Робастность в статистике. Подход на основе функций влияния. — М.: Мир, 1989. — 512 с.
9. Цыпкин Я. З. Информационная теория идентификации. — М.: Наука, 1995. — 336 с.
10. Ширяев А. Н. Основы стохастической финансовой математики. Т. 2. Теория. — М.: Фазис, 1998. — 512 с.
11. Овсянников А. В. Робастно-адаптивный усилитель-ограничитель / Радиотехника. 2011. № 3. С. 85 — 89.
12. Орлов А. И. Устойчивые математические методы и модели / Заводская лаборатория. Диагностика материалов. 2010. Т. 76. № 3. С. 59 — 67.
13. Корн Г., Корн Т. Справочник по математике (для научных работников и инженеров) / Пер. с англ. — М.: Наука, 1973. — 831 с.
14. Никитин Я. Ю. Асимптотическая эффективность непараметрических критериев. — М.: Наука, 1995. — 240 с.
15. Harkness W. L., Harkness M. L. Generalized hyperbolic secant distributions / Journal of the American Statistical Association. 1968. N 63. P. 329 — 337.
16. David C. V. The Generalized Secant Hyperbolic distribution and its properties / Theory and Methods. 2002. N 31(2). P. 219 — 238.
17. Вадзинский Р. Н. Справочник по вероятностным распределениям. — СПб.: Наука, 2001. — 295 с.
18. Орлов А. И. Прикладная статистика. — М.: Экзамен, 2006. — 671 с.
19. Андерсон Т. Статистический анализ временных рядов. — М.: Мир, 1976. — 756 с.
20. Martin R. D. Time series: model estimation data analysis and robust procedures / Proceedings of Symposium in Applied Mathematics. 1980. Vol. 23. P. 73 — 110.
21. Лукашин Ю. П. Адаптивные методы краткосрочного прогнозирования временных рядов. — М.: Финансы и статистика, 2003. — 416 с.
22. Box G. E. P., Jenkins J. M., Reinsel G. C. Time Series Analysis: Forecasting and Control. — Wiley, 2008. — 756 p.
23. Орлов Ю. Н. Оптимальное разбиение гистограммы для оценивания выборочной плотности функции распределения нестационарного временного ряда / Препринты ИПМ им. М. В. Келдыша. 2013. № 14. — 26 с.
URL: <http://library.keldysh.ru/preprint.asp?id=2013-14>.

Mathematical methods of investigation

UDC 519.24

GENERALIZED sech^k -PROBABILITY DENSITY: ROBUST PROPERTIES, EVALUATION, AND APPLICATION PROBABILITY DENSITY FUNCTION

© A. V. Ausiannikau

Submitted December 28, 2016.

For an adequate probabilistic description of the objects of statistical research, it is convenient to use a generalized probability density function that provides keeping the same mathematical model when the density parameters that affect the function shape change. We present the definition of a generalized sech^k -density of probability distribution and the main characteristics: distribution function, characteristic function, variance, kurtosis, differential entropy, Fisher information regarding the offset parameter. It is shown that generalized sech^k -density is a convenient probabilistic model that can be used in applied statistics in the range of excess values from zero to three. sech^k -density is a smooth function minimizing Fisher's information on the density class. The robust properties of the generalized sech^k -density are studied as applied to the algorithm of estimating mixing with known and unknown density parameters. A subclass of reduced robust estimation procedures — naturally lowered estimates — is introduced into the consideration. The quantitative characteristics of the sensitivity to a large error and sensitivity to a change in the asymptotic dispersion are presented. The considered examples of possible use of sech^k -density allow us to speak about the expediency of its application in theoretical and applied statistics.

Keywords: sech^k -density; robustness; non-linear transformation; robust estimation; underestimation; sensitivity.

REFERENCES

- Novitskii P. V., Zograf I. A. Evaluation of errors of measurement results. 2nd Edition. — Leningrad: Énergoatomizdat, 1991. — 304 p. [in Russian].
- Orlov A. I. The new paradigm of applied statistics / Zavod. Lab. Diagn. Mater. 2012. Vol. 78. N 11. P. 87 – 93 [in Russian].
- Ibragimov I. A., Khas'minskii R. Z. Asymptotic estimation theory. — Moscow: Nauka, 1979. — 528 p. [in Russian].
- Huber P. J. Robust statistics: A review / The Annals of Mathematical Statistics. 1972. Vol. 43. N 4. P. 1041 – 1067.
- Ershov A. A. Stable methods for estimating parameters (review) / Avtomat. Telemekh. 1978. N 8. P. 66 – 100 [in Russian].
- Stable estimation methods (Statistical processing of heterogeneous populations). — Moscow: Statistika, 1980. — 208 p. [in Russian].
- Huber P. J. Robust Statistics. — Wiley, 1981. — 304 p.
- Khampel' F., Ronchetti É., Rausseu P., Shtáel' V. Robust Statistics. Approach Based on Influence Functions. — Moscow: Mir, 1989. — 512 p. [Russian translation].
- Tsyarkin Ya. Z. Information theory of identification. — Moscow: Nauka, 1995. — 336 p. [in Russian].
- Shiryayev A. N. Fundamentals of stochastic financial mathematics. Vol. 2. Theory. — Moscow: Fazis, 1998. — 512 p. [in Russian].
- Ausiannikau A. V. Robust adaptive amplifier-limiter / Radiotekhnika. 2011. N 3. P. 85 – 89 [in Russian].
- Orlov A. I. Stable mathematical methods and models / Zavod. Lab. Diagn. Mater. 2010. Vol. 76. N 3. P. 59 – 67 [in Russian].
- Korn G. A., Korn T. M. Mathematical handbook for scientists and engineers: definitions, theorems, and formulas for reference and review. — General Publishing Company, 2000. — 1151 p.
- Nikitin Y. Asymptotic efficiency of nonparametric tests. — Cambridge Univ. Press, 1995. — 240 p.
- Harkness W. L., Harkness M. L. Generalized hyperbolic secant distributions / Journal of the American Statistical Association. 1968. N 63. P. 329 – 337.
- David C. V. The Generalized Secant Hyperbolic distribution and its properties / Theory and Methods. 2002. N 31(2). P. 219 – 238.
- Vadzinskii R. N. Handbook of probabilistic distributions. — St. Petersburg: Nauka, 2001. — 295 p. [in Russian].
- Orlov A. I. Applied statistics. — Moscow: Ékzamen, 2006. — 671 p. [in Russian].
- Anderson T. W. The statistical analysis of time series. — John Wiley & Sons, Inc., 1971. — 704 p.
- Martin R. D. Time series: model estimation data analysis and robust procedures / Proc. Symp. Appl. Math. 1980. Vol. 23. P. 73 – 110.
- Lukashin Yu. P. Adaptive methods of short-term forecasting of time series. — Moscow: Finansy i statistika, 2003. — 416 p. [in Russian].
- Box G. E. P., Jenkins J. M., Reinsel G. C. Time Series Analysis: Forecasting and Control. — Wiley, 2008. — 756 p.
- Orlov Yu. N. Optimal splitting of the histogram for estimating the sample density of the distribution function of a non-stationary time series / Preprint IPM im. M. V. Keldysha. 2013. N 14. — 26 p. [in Russian]. URL: <http://library.keldysh.ru/preprint.asp?id=2013-14> [in Russian].