Math-Net.Ru

Общероссийский математический портал

В. Е. Бенинг, В. Ю. Королев, Об использовании распределения Стиюдента в задачах теории вероятностей и математической статистики, *Теория вероятн. и ее примен.*, 2004, том 49, выпуск 3, 417–435

DOI: <https://doi.org/10.4213/tvp201>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 52.90.164.192

26 апреля 2023 г., 07:37:03

ТЕОРИЯ ВЕРОЯТНОСТЕЙ И ЕЕ ПРИМЕНЕНИЯ

Основан в январе 1956 г.
Выходит 4 раза в год
Москва • Наука • ТВП

Том 49
Выпуск 3, 2004 год
июль, август, сентябрь

© 2004 г.

БЕНИНГ В. Е.*, КОРОЛЕВ В. Ю.*

ОБ ИСПОЛЬЗОВАНИИ РАСПРЕДЕЛЕНИЯ СТЬЮДЕНТА В ЗАДАЧАХ ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ¹⁾

Работа посвящена математическому обоснованию возможности использования распределения Стьюдента в задачах описательной статистики. При этом особо выделен случай, когда параметр распределения Стьюдента («число степеней свободы») мал. Показано, что распределение Стьюдента с произвольным «числом степеней свободы» может быть получено в качестве предельного при случайном объеме выборки. Подчеркивается возможность использования семейства распределений Стьюдента в качестве удобной модели распределений с тяжелыми хвостами, так как для него (в отличие от устойчивых законов) многие формулы, в частности, функция правдоподобия, приобретают явный вид.

В качестве иллюстрации возможностей статистического анализа, основанного на стьюдентовском семействе, рассматривается задача статистического оценивания центра распределения Стьюдента в предположении, что параметр формы («число степеней свободы») известен. В работе рассматриваются эквивариантные оценки центра распределения Стьюдента, основанные на порядковых статистиках, М-оценки и оценки максимального правдоподобия. Вычисляется их асимптотическая относительная эффективность и изучается ее поведение при стремлении «числа степеней свободы» к нулю.

Ключевые слова и фразы: асимптотическая нормальность, выборки случайного объема, распределение Стьюдента, асимптотическая относительная эффективность.

* Московский государственный университет им. М.В. Ломоносова, факультет вычислительной математики и кибернетики, Ленинские горы, 119992 Москва, Россия; e-mail: vkorolev@df.ru, vbening@cs.msu.su

¹⁾ Работа выполнена при поддержке Российского фонда фундаментальных исследований, проекты 02-01-00949, 02-01-01080 и 03-01-00428, первый автор поддержан также INTAS, проект 03-51-5018.

1. Введение

Как известно, распределением Стьюдента называется абсолютно непрерывное распределение вероятностей, задаваемое плотностью

$$p_\gamma(x) = \frac{\Gamma((\gamma+1)/2)}{\sqrt{\pi\gamma}\Gamma(\gamma/2)} \left(1 + \frac{x^2}{\gamma}\right)^{-(\gamma+1)/2}, \quad -\infty < x < \infty. \quad (1.1)$$

Здесь $\gamma > 0$ — параметр, $\Gamma(\cdot)$ — эйлерова гамма-функция. В частности, при $\gamma = 1$ плотность (1.1) имеет вид

$$p_1(x) = \frac{1}{\pi(1+x^2)}, \quad -\infty < x < \infty,$$

что соответствует распределению Коши. Несложно видеть, что у распределения Стьюдента с параметром γ отсутствуют моменты порядка $\delta \geq \gamma$.

Если $\gamma = n$ — натуральное число, $X, X_1, \dots, X_n$ — независимые случайные величины (с.в.), имеющие одно и то же стандартное нормальное распределение, то, как известно, с.в.

$$Y = \frac{\sqrt{n} \cdot X}{\sqrt{X_1^2 + \dots + X_n^2}} \quad (1.2)$$

имеет распределение Стьюдента с параметром n , который в таком случае называется *числом степеней свободы*. На представлении (1.2) основан критерий проверки гипотез о среднем значении нормальных выборок, предложенный в 1908 г. У.С. Госсеттом (W.S. Gossett), который подписал свою статью «*On the probable error of the mean*» [1] псевдонимом «Student».

Несложно видеть, что в представлении (1.2) с.в. $(X_1^2 + \dots + X_n^2)/n$ имеет гамма-распределение с параметром формы $\alpha = n/2$ и параметром масштаба $\lambda = n/2$. Таким образом, по теореме Фубини из представления (1.2) вытекает возможность записать функцию распределения Стьюдента $F_n(x)$ с n степенями свободы, $F_n(x) = \int_{-\infty}^x p_n(y) dy$, в виде

$$\begin{aligned} F_n(x) &= \mathbf{P}\{Y < x\} = \mathbf{P}\left\{X < x\sqrt{n^{-1}(X_1^2 + \dots + X_n^2)}\right\} \\ &= \int_0^\infty \Phi(x\sqrt{y}) dG_{n/2, n/2}(u) = \mathbf{E}\Phi\left(x\sqrt{U_{n/2}}\right), \quad -\infty < x < \infty, \end{aligned} \quad (1.3)$$

где $\Phi(x)$ — стандартная нормальная функция распределения, $\mathbf{P}\{U_{n/2} < x\} = G_{n/2, n/2}(x)$, $G_{\alpha, \lambda}(x)$ — функция гамма-распределения с параметром формы α и параметром масштаба λ ,

$$G_{\alpha, \lambda}(x) = \begin{cases} 0 & \text{при } x \leq 0, \\ \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^x e^{-\lambda y} y^{\alpha-1} dy & \text{при } x > 0. \end{cases} \quad (1.4)$$

Таким образом, распределение Стьюдента с целочисленным параметром $\gamma = n$ принадлежит к семейству масштабных смесей нормальных законов.

Общеизвестна важная роль, которую распределение Стьюдента играет в математической статистике при анализе нормальных выборок. Здесь параметр γ тесно связан с объемом выборки и принимает натуральные значения. Однако можно сказать, что в таких задачах роль распределения Стьюдента в значительной мере вспомогательна, оно является в определенном смысле абстрактной, идеальной теоретической моделью.

Вместе с тем, в описательной статистике распределение Стьюдента практически не используется в качестве аналитической модели, «подгоняемой» к экспериментальным данным²⁾. Лишним подтверждением этого служит то обстоятельство, что ни в одном руководстве по теории (или практике) статистического оценивания не рассматривается задача оценивания параметра распределения Стьюдента.

По-видимому, недостаточное доверие прикладных статистиков к распределению Стьюдента как к модели, описывающей статистическое поведение реальных данных, связано с тем, что, в отличие от, скажем, нормального или пуассоновского распределений, фигурирующих в качестве предельных соответственно в центральной предельной теореме и теореме Пуассона о редких событиях, распределение Стьюдента не считается асимптотической аппроксимацией.

Следует особо подчеркнуть, что распределение Стьюдента в силу относительной простоты представления (1.1) могло бы быть удобной аналитической моделью, описывающей вероятностно-статистические свойства больших рисков, так как оно имеет более тяжелые хвосты, нежели нормальный закон. Например, оно могло бы стать удобной альтернативой устойчивым законам, часто применяемым в таком качестве (см., например, [7], [8]). Преимущество распределения Стьюдента перед устойчивыми моделями заключается, например, в том, что статистический анализ стьюдентовских моделей намного проще, так как для них функция правдоподобия выписывается в явном виде в терминах элементарных функций, в то время как для устойчивых законов это невозможно (за четырьмя исключениями). Вместе с тем, для $0 < \gamma \leq 2$ при $|x| \rightarrow \infty$ хвосты распределения Стьюдента убывают степенным образом, что совпадает с асимптотическим (при $|x| \rightarrow \infty$) поведением хвостов устойчивых законов.

²⁾ Лишь относительно недавно появились работы, в которых распределение Стьюдента применяется (впрочем, без надлежащего теоретического обоснования) для описания динамики некоторых финансовых индексов, в частности, приращений логарифмов биржевых цен. В первую очередь здесь следует упомянуть работы П. Прэтца [2] и Р. Блаттберга и Н. Гоундса [3].

Цель данной заметки — указать довольно простую асимптотическую схему, непосредственно приводящую к распределению Стьюдента как к предельному, и, как следствие, дать обоснование возможности более широкого использования распределения Стьюдента в задачах описательной статистики. В качестве примера рассмотрена задача оценивания центра распределения Стьюдента в случае, когда число степеней свободы мало.

В п. 2.1 приводятся вспомогательные результаты, в пп. 2.2–2.4 обосновывается возникновение распределения Стьюдента как предельной аппроксимации, в п. 3.1 приводятся основные результаты, касающиеся асимптотической эффективности эквивариантных оценок центра распределения, в пп. 3.2 и 3.3 рассмотрены линейные комбинации порядковых статистик, М-оценки, оценки Ходжеса–Лемана и оценки максимального правдоподобия как оценки центра распределения Стьюдента с малым числом степеней свободы.

Помимо введенных ранее, используются следующие обозначения: $\mathbf{R}$ — действительная прямая, символ $\Rightarrow$ будет обозначать слабую сходимость, $\Phi(x)$ и $\phi(x)$ — функция распределения и плотность стандартного нормального закона.

2. Распределение Стьюдента как асимптотическая аппроксимация

2.1. Вспомогательные результаты. Наши дальнейшие рассуждения будут основаны на двух следующих леммах.

Рассмотрим с.в. $N_1, N_2, \dots, X_1, X_2, \dots$, определенные на одном и том же измеримом пространстве $(\Omega, \mathcal{A})$. Пусть на $\mathcal{A}$ задано семейство вероятностных мер $\{\mathbf{P}_\theta, \theta \in \Theta\}$. Предположим, что при каждом $n \geq 1$ с.в. N_n принимает только натуральные значения и независима от последовательности $X_1, X_2, \dots$ относительно каждой меры из семейства $\{\mathbf{P}_\theta, \theta \in \Theta\}$. Пусть $T_n = T_n(X_1, \dots, X_n)$ — некоторая статистика, т.е. измеримая функция от с.в. $X_1, \dots, X_n$. Для каждого $n \geq 1$ определим с.в. T_{N_n} , положив $T_{N_n}(\omega) = T_{N_n(\omega)}(X_1(\omega), \dots, X_{N_n(\omega)}(\omega))$ для каждого элементарного исхода $\omega \in \Omega$. Будем говорить, что статистика T_n асимптотически нормальна, если существуют функции $\delta(\theta)$ и $t(\theta)$ такие, что при каждом $\theta \in \Theta$

$$\mathbf{P}_\theta \{ \delta(\theta) \sqrt{n} (T_n - t(\theta)) < x \} \Rightarrow \Phi(x), \quad n \rightarrow \infty. \quad (2.1)$$

Примеры асимптотически нормальных статистик хорошо известны. Свойством асимптотической нормальности обладают, например, выборочное среднее (при условии существования дисперсий), центральные порядковые статистики или оценки максимального правдоподобия (при достаточно общих условиях регулярности) и многие другие статистики.

Лемма 2.1. Пусть $\{d_n\}_{n \geq 1}$ — некоторая неограниченно возрастающая последовательность положительных чисел. Предположим, что $N_n \rightarrow \infty$ по вероятности при $n \rightarrow \infty$ относительно каждой вероятности из семейства $\{\mathbf{P}_\theta, \theta \in \Theta\}$. Пусть статистика T_n асимптотически нормальна в смысле (2.1). Для того чтобы при каждом $\theta \in \Theta$ существовала такая функция распределения $F(x, \theta)$, что

$$\mathbf{P}_\theta \left\{ \delta(\theta) \sqrt{d_n} (T_{N_n} - t(\theta)) < x \right\} \Rightarrow F(x, \theta), \quad n \rightarrow \infty,$$

необходимо и достаточно, чтобы существовало семейство функций распределения $\mathcal{H} = \{H(x, \theta): \theta \in \Theta\}$, удовлетворяющее условиям

$$H(x, \theta) = 0, \quad x < 0, \quad \theta \in \Theta;$$

$$F(x, \theta) = \int_0^\infty \Phi(x\sqrt{y}) d_y H(y, \theta), \quad x \in \mathbf{R}, \quad \theta \in \Theta;$$

$$\mathbf{P}_\theta \{N_n < d_n x\} \Rightarrow H(x, \theta), \quad n \rightarrow \infty, \quad \theta \in \Theta.$$

При этом если функции распределения с.в. N_n не зависят от θ , то не зависят от θ и функции распределения $H(x, \theta)$, т.е. семейство $\mathcal{H}$ состоит из единственного элемента.

Доказательство. Данная лемма по сути лишь переобозначениями отличается от теоремы 3 из работы [5], доказательство которой, в свою очередь, основано на общих теоремах о сходимости суперпозиций независимых случайных последовательностей [4], [6].

Пусть $N_{p,r}$ — с.в., имеющая отрицательное биномиальное распределение,

$$\mathbf{P}\{N_{p,r} = k\} = C_{r+k-2}^{k-1} p^r (1-p)^{k-1}, \quad k = 1, 2, \dots \quad (2.2)$$

Здесь $r > 0$ и $p \in (0, 1)$ — параметры и для нецелых r величина C_{r+k-2}^{k-1} определяется как

$$C_{r+k-2}^{k-1} = \frac{\Gamma(r+k-1)}{(k-1)! \Gamma(r)}.$$

В частности, при $r = 1$ соотношение (2.2) задает геометрическое распределение. Известно, что $\mathbf{E}N_{p,r} = [r(1-p) + p]/p$, так что $\mathbf{E}N_{p,r} \rightarrow \infty$ при $p \rightarrow 0$.

Отрицательное биномиальное распределение с натуральным r допускает наглядную интерпретацию в терминах испытаний Бернулли. А именно, с.в. с распределением (2.2) — это число испытаний Бернулли, проведенных до осуществления r -й по счету неудачи, если вероятность успеха в одном испытании равна $1-p$.

Лемма 2.2. Для любого фиксированного $r > 0$

$$\limsup_{p \rightarrow 0} \sup_{x \in \mathbf{R}} \left| \mathbf{P} \left\{ \frac{N_{p,r}}{\mathbf{E}N_{p,r}} < x \right\} - G_{r,r}(x) \right| = 0,$$

где $G_{r,r}(x)$ — функция гамма-распределения с параметром формы, совпадающим с параметром масштаба и равным r , см. (1.4).

Доказательство. Характеристическая функция с.в. $N_{p,r}$ равна

$$\mathbf{E} \exp\{itN_{p,r}\} = e^{it} \left[\frac{p}{1 - (1-p)e^{it}} \right]^r, \quad t \in \mathbf{R}.$$

Поэтому, используя представление $e^z = 1 + z + o(|z|)$, $|z| \rightarrow 0$, при каждом фиксированном $t \in \mathbf{R}$ мы имеем

$$\begin{aligned} \mathbf{E} \exp \left\{ it \frac{N_{p,r}}{\mathbf{E} N_{p,r}} \right\} &= \mathbf{E} \exp \left\{ \frac{itpN_{p,r}}{r(1-p)+p} \right\} \\ &= \exp \left\{ \frac{itp}{r(1-p)+p} \right\} \left[\frac{p}{1 - (1-p) \exp\{itp/[r(1-p)+p]\}} \right]^r \\ &= \exp \left\{ \frac{itp}{r(1-p)+p} \right\} \\ &\quad \times \left[\frac{1}{p} \left(1 - \exp \left\{ \frac{itp}{r(1-p)+p} \right\} \right) + \exp \left\{ \frac{itp}{r(1-p)+p} \right\} \right]^{-r} \\ &= \exp \left\{ \frac{itp}{r(1-p)+p} \right\} \left[-\frac{1}{p} \left(\frac{itp}{r(1-p)+p} + o(p) \right) + 1 + O(p) \right]^{-r} \\ &\rightarrow \left[1 - \frac{it}{r} \right]^{-r} \end{aligned}$$

при $p \rightarrow 0$. Но правая часть этого соотношения в точности совпадает с характеристической функцией гамма-распределения $G_{r,r}(x)$. Ссылка на теорему о непрерывности соответствия между функциями распределения и соответствующими им характеристическими функциями (см., например, [31, с. 6]) устанавливает сходимость допредельных функций распределения к предельной в каждой точке $x \in \mathbf{R}$. Так как предельная функция распределения непрерывна, то ссылка на теорему 11 главы 1 в [31] завершает доказательство.

2.2. Распределение Стьюдента как предельное при случайном объеме выборки. В подавляющем большинстве ситуаций, связанных с анализом экспериментальных данных, можно признать, что число случайных факторов, влияющих на наблюдаемые величины, само является случайным и изменяется от наблюдения к наблюдению. Поэтому вместо различных версий центральной предельной теоремы, обосновывающих нормальность распределения наблюдаемых с.в. в классической статистике, в таких ситуациях следует опираться на их аналоги для выборок случайного объема (см. лемму 2.1).

Теорема 2.1. Пусть $\gamma > 0$ произвольно и $\{d_n\}_{n \geq 1}$ — некоторая неограниченно возрастающая последовательность положительных чисел. Предположим, что $N_n \rightarrow \infty$ по вероятности при $n \rightarrow \infty$ относительно каждой вероятности из семейства $\{\mathbf{P}_\theta, \theta \in \Theta\}$. Пусть

статистика T_n асимптотически нормальна в смысле (2.1). Для того чтобы при каждом $\theta \in \Theta$

$$P_{\theta} \left\{ \delta(\theta) \sqrt{d_n} (T_{N_n} - t(\theta)) < x \right\} \Rightarrow F_{\gamma}(x), \quad n \rightarrow \infty,$$

где $F_{\gamma}(x)$ — функция распределения Стьюдента с параметром γ , необходимо и достаточно, чтобы

$$P_{\theta} \{N_n < d_n x\} \Rightarrow G_{\gamma/2, \gamma/2}(x), \quad n \rightarrow \infty, \quad \theta \in \Theta.$$

Доказательство. Несложно убедиться в том, что при произвольном $\gamma > 0$ плотность $p_{\gamma}(x)$ распределения Стьюдента с параметром γ (см. (1.1)) представима в виде

$$p_{\gamma}(x) = \mathbf{E} \sqrt{U_{\gamma/2}} \phi \left(x \sqrt{U_{\gamma/2}} \right), \quad (2.3)$$

где $U_{\gamma/2}$ — с.в. с функцией распределения $G_{\gamma/2, \gamma/2}(x)$. Действительно,

$$\begin{aligned} & \mathbf{E} \sqrt{U_{\gamma/2}} \phi \left(x \sqrt{U_{\gamma/2}} \right) \\ &= \frac{\gamma^{\gamma/2}}{2^{(\gamma+1)/2} \sqrt{\pi} \Gamma(\gamma/2)} \int_0^{\infty} \exp \left\{ -u \left(\frac{x^2 + \gamma}{2} \right) \right\} u^{(\gamma-1)/2} du \\ &= \frac{\gamma^{\gamma/2}}{2^{(\gamma+1)/2} \sqrt{\pi} \Gamma(\gamma/2)} \left(\frac{x^2 + \gamma}{2} \right)^{-(\gamma+1)/2} \int_0^{\infty} \exp\{-z\} z^{(\gamma+1)/2-1} dz \\ &= \frac{\gamma^{\gamma/2}}{2^{(\gamma+1)/2} \sqrt{\pi} \Gamma(\gamma/2)} \left(\frac{x^2 + \gamma}{2} \right)^{-(\gamma+1)/2} \Gamma \left(\frac{\gamma+1}{2} \right) \\ &= \frac{\Gamma((\gamma+1)/2)}{\sqrt{\pi} \gamma \Gamma(\gamma/2)} \left(1 + \frac{x^2}{\gamma} \right)^{-(\gamma+1)/2} = p_{\gamma}(x). \end{aligned}$$

Но плотность (2.3) соответствует функции распределения $\mathbf{E}\Phi(x\sqrt{U_{\gamma/2}})$ (для натуральных γ этот факт был отмечен во введении). Теперь требуемое утверждение вытекает из леммы 2.1 с учетом свойства идентифицируемости масштабных смесей нормальных законов, согласно которому если $\mathbf{E}\Phi(xV_1) \equiv \mathbf{E}\Phi(xV_2)$ для некоторых положительных случайных величин V_1 и V_2 , то $V_1 \stackrel{d}{=} V_2$ (см., например, [32]).

Следствие 2.1. Пусть $r > 0$ произвольно. Предположим, что при каждом $n \geq 1$ с.в. N_n имеет отрицательное биномиальное распределение с параметрами $p = 1/n$ и r . Пусть статистика T_n асимптотически нормальна в смысле (2.1). Тогда при каждом $\theta \in \Theta$

$$P_{\theta} \left\{ \delta(\theta) \sqrt{rn} (T_{N_n} - t(\theta)) < x \right\} \Rightarrow P_{2r}(x), \quad n \rightarrow \infty,$$

равномерно по $x \in \mathbf{R}$, где $P_{2r}(x)$ — функция распределения Стьюдента с параметром $\gamma = 2r$.

Доказательство. В силу леммы 2.2 мы имеем

$$\frac{N_n}{nr} = \frac{N_n}{\mathbf{E}N_n} \cdot \frac{\mathbf{E}N_n}{nr} = \frac{N_n}{\mathbf{E}N_n} \cdot \frac{r(n-1)+1}{nr} = \frac{N_n}{\mathbf{E}N_n} \left[1 + O\left(\frac{1}{n}\right) \right] \Rightarrow U_r$$

при $n \rightarrow \infty$, где U_r — с.в., имеющая гамма-распределение с параметром формы, совпадающим с параметром масштаба и равным r . Теперь требуемое утверждение непосредственно вытекает из теоремы 2.1.

З а м е ч а н и е 2.1. Распределение Коши ($\gamma = 1$) возникает в ситуации, описанной в следствии 2.1, когда объем выборки N_n имеет отрицательное биномиальное распределение с параметрами $p = 1/n$, $r = \frac{1}{2}$ и n велико.

З а м е ч а н и е 2.2. В ситуации, когда объем выборки N_n имеет отрицательное биномиальное распределение с параметрами $p = 1/n$, $r = 1$ (т.е. геометрическое распределение с параметром $p = 1/n$), в пределе при $n \rightarrow \infty$ мы получаем распределение Стьюдента с параметром $\gamma = 2$, которому соответствует функция распределения

$$P_2(x) = \frac{1}{2} \left(1 + \frac{x}{\sqrt{2 + x^2}} \right), \quad x \in \mathbf{R}. \quad (2.4)$$

Такое распределение впервые описано как предельное для выборочной медианы, построенной по выборке случайного объема, имеющего геометрическое распределение, по-видимому, в работе [9] (следует отметить, что в этой работе не указано, что функция распределения, стоящая в правой части (2.4), соответствует распределению Стьюдента).

Таким образом, основной вывод из приведенных выше результатов можно сформулировать следующим образом. Если число случайных факторов, определяющих наблюдаемое значение с.в., само является с.в., распределение которой может быть приближено гамма-распределением с одинаковыми параметрами (например, является отрицательным биномиальным с вероятностью успеха, близкой к единице, см. лемму 2.2), то те функции от значений случайных факторов, которые в классической ситуации считаются асимптотически нормальными, в действительности являются асимптотически стьюдентовскими. Следовательно, в силу довольно широкой применимости гамма-моделей с одинаковыми параметрами и отрицательных биномиальных моделей, распределение Стьюдента может рассматриваться в задачах прикладной (описательной) статистики как вполне разумная модель.

2.3. Экстремальные энтропийные свойства распределения Стьюдента. Необходимо отметить, что еще одним обстоятельством, которое могло бы привлечь внимание прикладных статистиков к распределению Стьюдента, является так называемый энтропийный подход, согласно которому в условиях неопределенности математическую модель стохастической ситуации следует выбирать так, чтобы выбранная модель соответствовала максимально возможной (при некоторых разумных условиях) неопределенности. При этом в качестве меры неопределенности выбирается (дифференциальная) энтропия абсолютно непрерывного вероятностного распределения. Хорошо известно, что при со-

ответствующих ограничениях на носитель и моменты плотности $p(x)$ «наиболее неопределенными» в смысле классической дифференциальной энтропии

$$H[p] = - \int_{-\infty}^{\infty} p(x) \ln p(x) dx \quad (2.5)$$

являются, например, равномерное распределение (среди всех распределений с ограниченным носителем), показательное распределение (среди всех распределений, сосредоточенных на неотрицательной оси и имеющих конечное математическое ожидание) и нормальное распределение (среди всех распределений, сосредоточенных на всей числовой оси и имеющих конечный второй момент). Как показано в книге [26], в классе плотностей $p(x)$, положительных на всей числовой оси и удовлетворяющих условию $\int_{-\infty}^{\infty} \ln(1+x^2)p(x)dx = c$, $0 < c < \infty$, максимум функционала $H[p]$ достигается на плотности распределения Стьюдента $p_\gamma(x)$, параметр γ которой зависит от числа c и удовлетворяет уравнению

$$\psi\left(\frac{\gamma+1}{2}\right) - \psi\left(\frac{\gamma}{2}\right) = c,$$

где $\psi(z)$ — дигамма-функция, $\psi(z) = \Gamma'(z)/\Gamma(z)$. (Интересно отметить, что и в цитируемой книге [26] данное распределение не распознано как распределение Стьюдента, а названо «обобщенным распределением Коши».)

Более того, в работе [27] показано, что если вместо (2.5) в качестве меры неопределенности рассмотреть обобщенную q -энтропию (non-extensive entropy)

$$H_q[p] = \frac{1}{q-1} \left[1 - \int_{-\infty}^{\infty} p^q(x) dx \right], \quad q \in \mathbf{R}$$

(для которой функционал (2.5) является предельным случаем при $q \rightarrow 1$), то для $1 < q < 3$ максимум функционала $H_q[p]$ при условиях

$$\int_{-\infty}^{\infty} xp(x) dx = 0 \quad \text{и} \quad \int_{-\infty}^{\infty} x^2 p^q(x) dx = 1$$

доставляет распределение Стьюдента с параметром $\gamma = (3-q)/(q-1)$.

2.4. Случай малого параметра $\gamma = 2r$. Выше мы уже упоминали, что отрицательное биномиальное распределение (как мы убедились, тесно связанное с распределением Стьюдента следствием 2.1) при натуральном r может быть интерпретировано в терминах испытаний Бернулли, проведенных до r -й неудачи. В то же время, особенно в задачах, связанных с анализом больших рисков, существенный интерес представляет изучение распределения Стьюдента с малым параметром r , т.е. с очень тяжелыми хвостами. Более того, можно показать, что при $\gamma = 2r \rightarrow 0$ максимум плотности $p_\gamma(x)$ распределения Стьюдента

(см. (5.1)) стремится к нулю как $O(\sqrt{\gamma})$. Одновременно хвосты распределения Стьюдента становятся все более и более тяжелыми. Поэтому распределение Стьюдента с малым параметром может рассматриваться как некий аналог равномерного распределения на бесконечном интервале.

Чтобы следствие 2.1 можно было использовать и в такой ситуации, следует разобраться, что из себя представляет отрицательное биномиальное распределение, т.е. как оно может быть проинтерпретировано, при $0 < r < 1$. Мы приведем два примера такой интерпретации.

Пример 2.1. Этот пример хорошо знаком. Скажем, в книге [12] он приведен со ссылкой на работу [13]. Рассмотрим с.в. $M_{p,r}$, имеющую смешанное пуассоновское распределение

$$\mathbf{P}\{M_{p,r} = k\} = \mathbf{E} \exp\{-U_{r,p/(1-p)}\} \frac{U_{r,p/(1-p)}^k}{k!}, \quad k = 0, 1, 2, \dots,$$

где $U_{r,p/(1-p)}$ — с.в., имеющая гамма-распределение с параметром формы r и параметром масштаба $p/(1-p)$. Легко убедиться, что безусловное распределение с.в. $M_{p,r}$ имеет вид

$$\mathbf{P}\{M_{p,r} = k\} = C_{r+k-1}^k p^r (1-p)^k, \quad k = 0, 1, 2, \dots \quad (2.6)$$

Несложно видеть, что при этом $M_{p,r} = N_{p,r} - 1$, где, как и ранее, $N_{p,r}$ — с.в., имеющая отрицательное биномиальное распределение с параметрами r и p (см. (2.2)).

Таким образом, если сначала реализуется значение u с.в. $U_{r,p/(1-p)}$ с гамма-распределением $G_{r,p/(1-p)}$, а затем реализуется значение с.в. $M_{p,r}$, имеющей пуассоновское распределение, параметр которого равен полученному значению u , то, прибавив единицу к итоговой реализации с.в. $M_{p,r}$, мы получаем реализацию отрицательной биномиальной с.в. $N_{p,r}$ с параметрами r и p . При этом требуемая асимптотика $p \rightarrow 0$ (гарантирующая применимость следствия 2.1) и $r \rightarrow 0$ для $N_{p,r}$ и $P_{2r}(x)$ естественно возникает как аналогичная асимптотика для $U_{r,p/(1-p)}$.

Пример 2.2. Вновь, наряду со с.в. $N_{p,r}$, введенной выше, рассмотрим с.в. $M_{p,r} = N_{p,r} - 1$, имеющую распределение (2.6). Рассмотрим независимые одинаково распределенные неотрицательные целочисленные с.в. $Z_1, Z_2, \dots$, каждая из которых имеет производящую функцию

$$p(z) = \frac{\ln[1 - (1-p)z]}{\ln p}, \quad |z| \leq 1,$$

где $p \in (0, 1)$. Эта производящая функция задает так называемое логарифмическое распределение Фишера

$$\mathbf{P}\{Z_1 = k\} = \frac{(1-p)^k}{k \ln p}, \quad k = 1, 2, \dots$$

(см. [12]). Пусть N — с.в., имеющая распределение Пуассона с параметром $\mu > 0$, независимая от с.в. $Z_1, Z_2, \dots$. Положим

$$S = Z_1 + \dots + Z_N.$$

Если $N = 0$, то полагаем $S = 0$.

Можно показать (см. [10], [11]), что в таком случае числа q_n в представлении обобщенной пуассоновской производящей функции с.в. S

$$\mathbf{E}z^S = \sum_{n=0}^{\infty} q_n z^n = \exp \{ \mu(p(z) - 1) \}, \quad |z| \leq 1,$$

равны

$$q_n = C_{n+r-1}^n p^r (1-p)^n, \quad n = 0, 1, 2, \dots,$$

где

$$r = -\frac{\mu}{\ln p}. \quad (2.7)$$

Другими словами, в рассматриваемом случае распределение пуассоновской случайной суммы S совпадает с распределением (2.6) с.в. $M_{p,r}$ при r , удовлетворяющем соотношению (2.7).

Таким образом, отрицательное биномиальное распределение с параметрами r и p можно интерпретировать как сдвинутое на единицу распределение суммы случайного числа независимых одинаково распределенных с.в., в которой слагаемые имеют логарифмическое распределение с параметром p , а число слагаемых имеет пуассоновское распределение с параметром $\mu = r \ln(1/p)$. При этом требуемое соотношение $r < 1$ выполняется, если $\mu < \ln(1/p)$, а распределение Стьюдента с $\gamma = 2r \rightarrow 0$ может выступать в качестве асимптотической аппроксимации, основанной на следствии 2.1, если $\mu = \mu(p) = o(\ln(1/p))$ при $p \rightarrow 0$.

3. Статистическое оценивание центра распределения Стьюдента

В качестве иллюстрации тех возможных преимуществ, которые предоставляет семейство распределений Стьюдента, если его использовать в качестве модели распределения с тяжелыми хвостами, мы вычислим некоторые асимптотические характеристики статистических процедур оценивания параметров распределения Стьюдента.

Наиболее очевидной и естественной задачей статистического оценивания, связанной с распределением Стьюдента, является задача оценивания параметра γ — «числа степеней свободы». Для ее решения могут быть использованы различные процедуры, например, метод максимального правдоподобия, метод логарифмических моментов. Особо следует отметить, что так как распределение Стьюдента принадлежит к классу

законов с тяжелыми хвостами, а параметр γ является показателем тяжести хвостов, то для оценивания этого параметра могут также применяться известные процедуры оценивания степени тяжести хвостов, такие, как оценки Хилла [28], Де Хаана (см., например, [29]) и другие (см., например, [30]). Оценки второй группы, как правило, вычисляются намного проще, нежели оценки максимального правдоподобия или оценки метода логарифмических моментов. Подробное сравнительное исследование упомянутых процедур не является целью данной статьи.

Вместо этого, в соответствии с общей направленностью данной статьи, ниже мы рассмотрим другую естественную задачу статистического оценивания, связанную с распределением Стьюдента, — задачу оценивания центра этого распределения, уделяя особое внимание асимптотическим свойствам соответствующих процедур при $\gamma \rightarrow 0$, что важно при использовании распределения Стьюдента для описания больших рисков.

3.1. Асимптотическая эффективность эквивариантных оценок. В этом пункте мы напомним основные сведения из теории эквивариантного оценивания и формулируем необходимые для дальнейшего результаты.

Пусть $X_1, \dots, X_n$ — независимые одинаково распределенные наблюдения, имеющие функцию распределения $F(x - \theta)$, $\theta \in \mathbf{R}$, причем функция распределения $F(x)$ известна, симметрична, $F(x) + F(-x) = 1$, и обладает положительной плотностью $p(x) = F'(x) > 0$. Рассмотрим задачу оценивания параметра сдвига θ . С этой целью будем рассматривать эквивариантные оценки $\delta_n = \delta_n(X_1, \dots, X_n)$, т.е. такие, что при всех $a \in \mathbf{R}$

$$\delta_n(X_1 + a, \dots, X_n + a) = \delta_n(X_1, \dots, X_n) + a.$$

Все традиционно используемые оценки параметра сдвига являются эквивариантными относительно сдвига, например, среднее, выборочная медиана или любое взвешенное среднее порядковых статистик с суммой весов, равной единице. Заметим, что оценка максимального правдоподобия также эквивариантна. Известно, что при квадратичной функции потерь при каждом n наилучшей является оценка Питмэна (см. [15, теорема 3.1.5, с. 148], [16, с. 33])

$$\delta_n^* = \frac{\int_{-\infty}^{+\infty} \theta \prod_{i=1}^n p(X_i - \theta) d\theta}{\int_{-\infty}^{+\infty} \prod_{i=1}^n p(X_i - \theta) d\theta}. \quad (3.1)$$

Однако эти оценки, являющиеся эталоном при сравнении оценок, как правило, являются трудновычислимыми и представляют в основном теоретический интерес.

В дальнейшем мы будем рассматривать асимптотический подход (при $n \rightarrow \infty$) и асимптотически нормальные последовательности эквивариантных оценок δ_n : $\sqrt{n}(\delta_n - \theta) \Rightarrow \mathcal{N}(0, \sigma^2)$, $n \rightarrow \infty$, где $\Rightarrow$ означает

слабую сходимость, а $\mathcal{N}(\mu, \sigma^2)$ — случайная величина, имеющая нормальное распределение с математическим ожиданием μ и дисперсией σ^2 .

Для сравнения предельного качества таких оценок используем понятие *асимптотической относительной эффективности* (АОЭ) (см. [15, с. 305–307]). Для двух последовательностей асимптотически нормальных эквивариантных оценок $\delta_{n,1} = \delta_{n,1}(X_1, \dots, X_n)$ и $\delta_{n,2} = \delta_{n,2}(X_1, \dots, X_n)$

$$\sqrt{n}(\delta_{n,i} - \theta) \Rightarrow \mathcal{N}(0, \sigma_i^2), \quad n \rightarrow \infty, \quad i = 1, 2, \quad (3.2)$$

определим АОЭ оценки $\delta_{n,1}$ относительно оценки $\delta_{n,2}$ как

$$e_{12} = \frac{\sigma_2^2}{\sigma_1^2}. \quad (3.3)$$

Отметим, что если определить число наблюдений $m = m(n)$ так, чтобы $\sqrt{n}(\delta_{m,2} - \theta) \Rightarrow \mathcal{N}(0, \sigma_1^2)$, $n \rightarrow \infty$, то (см. [15, теорема 5.2.1, с. 306])

$$e_{12} = \lim_{n \rightarrow \infty} \frac{m(n)}{n}.$$

Таким образом, чтобы получить предельное распределение, совпадающее с предельным распределением оценки $\delta_{n,1}$, оценке $\delta_{n,2}$ требуется примерно ne_{12} наблюдений. Итак, если $e_{12} > 1$, то оценка $\delta_{n,2}$ асимптотически «хуже» оценки $\delta_{n,1}$. В случае $e_{12} = 1$ эти оценки асимптотически эквивалентны.

Условие асимптотической нормальности (3.2) можно заменить на более общее условие вида (см. [15, теорема 5.2.2, с. 307])

$$n^q(\delta_{n,i} - \theta) \Rightarrow \kappa_i T, \quad n \rightarrow \infty, \quad i = 1, 2, \quad q > 0,$$

где функция распределения случайной величины T непрерывна и строго возрастает. Тогда

$$e_{12} = \left(\frac{\kappa_2}{\kappa_1} \right)^{1/q}$$

с аналогичной интерпретацией в терминах необходимого числа наблюдений.

Рассмотрим теперь абсолютную АОЭ (ААОЭ), т.е. асимптотическую эффективность относительно наилучшей оценки Питмэна δ_n^* (см. (3.1)). В работах [17], [18] показано, что при весьма общих условиях регулярности эта оценка асимптотически нормальна

$$\sqrt{n}(\delta_n^* - \theta) \Rightarrow \mathcal{N}(0, I^{-1}), \quad n \rightarrow \infty,$$

где I — фишеровская информация

$$I = \mathbf{E} \left(\frac{p'(X_1)}{p(X_1)} \right)^2. \quad (3.4)$$

Заметим, что асимптотическая дисперсия I^{-1} оценки Питмэна δ_n^* совпадает с нижней границей дисперсий из неравенства Крамера–Рао. Обозначим ААОЭ последовательности эквивариантных асимптотически нормальных оценок $\delta_{n,1}$ ($\sqrt{n}(\delta_{n,1} - \theta) \Rightarrow \mathcal{N}(0, \sigma_1^2)$, $n \rightarrow \infty$) относительно оценки Питмэна δ_n^* через

$$e_{1*} = \frac{1}{I\sigma_1^2} \quad (\leq 1). \quad (3.5)$$

Рассмотрим теперь задачу оценивания параметра θ по выборке независимых одинаково распределенных случайных величин $X_1, \dots, X_n$, имеющих функцию распределения $F(x - \theta)$. Если функция распределения $F(x)$ нормальна, то выборочное среднее

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \quad (3.6)$$

является несмещенной оценкой параметра θ с *минимальной дисперсией*. Однако в классе *симметричных* функций распределения $F(x)$ оценка (3.6) уже не будет оптимальной (см. [15, с. 128, задача 4.4]). Это связано с тем, что для симметричных распределений класс несмещенных оценок параметра сдвига значительно шире, чем без предположения симметрии.

Свойства оценки (3.6) существенно зависят от поведения хвостов распределения наблюдений. Например, для распределения Коши (имеющего столь тяжелые хвосты, что не существует математическое ожидание), как известно, при каждом $n \geq 1$ оценка (3.6) имеет такое же распределение, как и отдельное наблюдение. Естественный путь получения оценок с регулярным асимптотическим поведением (см. [14]) состоит в отбрасывании крайних наблюдений. Например, можно рассмотреть выборочную медиану

$$M_n = \begin{cases} X_{(m)}, & n = 2m - 1, \\ \frac{1}{2} (X_{(m)} + X_{(m+1)}), & n = 2m, \end{cases} \quad (3.7)$$

где $X_{(1)}, \dots, X_{(n)}$ — вариационный ряд, построенный по выборке $X_1, \dots, X_n$.

Однако такое радикальное отбрасывание крайних членов выборки не всегда приводит к разумным результатам (см. [14], [15, с. 317–319]). Естественным компромиссом между средним и медианой является отбрасывание, например, $[n\alpha]$ наименьших и $[n\alpha]$ наибольших наблюдений, где $0 \leq \alpha < \frac{1}{2}$, т.е. рассмотрение усеченных средних

$$\bar{L}_n = \frac{X_{([n\alpha]+1)} + \dots + X_{(n-[n\alpha])}}{n - 2[n\alpha]} \quad (3.8)$$

или α -уинзоризованных средних (см. [14])

$$\bar{L}_n^* = \frac{1}{n} \left([n\alpha]X_{([n\alpha])} + \sum_{i=[n\alpha]+1}^{n-[n\alpha]} X_{(i)} + [n\alpha]X_{(n-[n\alpha]+1)} \right).$$

Рассмотрим также оценку Ходжеса–Лемана (см. [20] и [15 с. 339–341])

$$W_n = \text{med} \left\{ \frac{X_{(i)} + X_{(j)}}{2}, 1 \leq i \leq j \leq n \right\}, \quad (3.9)$$

т.е. это медиана $n(n+1)/2$ значений $(X_{(i)} + X_{(j)})/2$.

3.2. Оценивание центра распределений Стьюдента с малым числом степеней свободы. Всюду далее будет рассматриваться семейство распределений Стьюдента с параметром $\gamma > 0$, функциями распределения $F_\gamma(x - \theta)$ и плотностями $p_\gamma(x) = F'_\gamma(x)$, определяемыми соотношением (1.1). Параметр $\gamma > 0$ («число степеней свободы») предполагается известным. При $0 < \delta < \gamma$ у этого распределения существует абсолютный момент порядка δ . При $2 < \gamma$ у этого распределения существует дисперсия, которая в таком случае равна $\gamma/(\gamma-2)$. При $0 < \gamma \leq 2$ дисперсии не существует.

Лемма 3.1. Семейство $F_\gamma(x)$ обладает следующими свойствами.

1. Для любого $\gamma > 0$ существует фишеровская информация (см. (3.4))

$$I_\gamma = \mathbf{E} \left(\frac{p'_\gamma(X_1)}{p_\gamma(X_1)} \right)^2 = \frac{\gamma + 1}{\gamma + 3}. \quad (3.10)$$

2. Для любого $x > 0$ справедливы соотношения

$$p_\gamma(x) \sim \frac{\gamma}{2x}, \quad \gamma \rightarrow 0,$$

и

$$p_\gamma(0) = \frac{\sqrt{\gamma} \Gamma((\gamma+1)/2)}{2\sqrt{\pi} \Gamma((\gamma+2)/2)} \sim \frac{\sqrt{\gamma}}{2}, \quad \gamma \rightarrow 0.$$

3. Для любого $x \geq 0$ справедливо равенство

$$F_\gamma(x) = 1 - \frac{x\gamma^{1/2}\Gamma((\gamma+1)/2)}{2\sqrt{\pi}\Gamma((\gamma+2)/2)(\gamma+x^2)^{(\gamma+1)/2}} \left(1 + \frac{\gamma(\gamma+1)}{(\gamma+2)(\gamma+x^2)} + r_\gamma(x) \right),$$

где остаточный член $r_\gamma(x)$ имеет вид

$$r_\gamma(x) = \frac{\gamma^2(\gamma+1)(\gamma+3)(\gamma+x^2)^{(\gamma+1)/2}}{x(\gamma+2)} \int_x^\infty \frac{dt}{(\gamma+t^2)^{(\gamma+5)/2}}.$$

Доказательство леммы непосредственно следует из определения гамма-функции и вида функций $p_\gamma(x)$ и $F_\gamma(x)$.

Теорема 3.1. Для оценок M_n , W_n , $\bar{L}_n$ и δ_n^* (см. (3.7), (3.9), (3.8), (3.1)) справедливы следующие соотношения.

1. Асимптотическая дисперсия оценки M_n имеет вид

$$\sigma_m^2 = \frac{\pi \Gamma^2((\gamma+2)/2)}{\gamma \Gamma^2((\gamma+1)/2)} \sim \frac{1}{\gamma}, \quad \gamma \rightarrow 0.$$

2. Асимптотическая дисперсия оценки W_n имеет вид

$$\sigma_w^2 = \frac{4\pi \Gamma^4((\gamma+2)/2) \Gamma^2(\gamma+1)}{3\gamma^3 \Gamma^4((\gamma+1)/2) \Gamma^2(\gamma+\frac{1}{2})} \sim \frac{4}{3\gamma^3 \pi^2}, \quad \gamma \rightarrow 0.$$

3. АОЭ оценки M_n относительно оценки W_n равна

$$e_{mw} = \frac{\sigma_w^2}{\sigma_m^2} = \frac{4\Gamma^2((\gamma+2)/2) \Gamma^2(\gamma+1)}{3\gamma^2 \Gamma^2((\gamma+1)/2) \Gamma^2(\gamma+\frac{1}{2})} \sim \frac{4}{3\gamma^2 \pi^2}, \quad \gamma \rightarrow 0.$$

4. ААОЭ оценок M_n и W_n относительно оценки Питмэна δ_n^* соответственно равны

$$e_{m*} = \frac{1}{I_\gamma \sigma_m^2} = \frac{(\gamma+3)\gamma \Gamma^2((\gamma+1)/2)}{(\gamma+1)\pi \Gamma^2((\gamma+2)/2)} \sim 3\gamma, \quad \gamma \rightarrow 0,$$

$$e_{w*} = \frac{1}{I_\gamma \sigma_w^2} = \frac{3\gamma^3(\gamma+3) \Gamma^4((\gamma+1)/2) \Gamma^2(\gamma+\frac{1}{2})}{4(\gamma+1)\pi \Gamma^4((\gamma+2)/2) \Gamma^2(\gamma+1)} \sim \frac{9\gamma^3 \pi^2}{4}, \quad \gamma \rightarrow 0.$$

5. Асимптотические дисперсии оценок $\bar{L}_n$ и δ_n^* имеют вид

$$\sigma_l^2 = \frac{4 \int_0^{\xi_{1-\alpha}} x(1-F_\gamma(x)) dx}{(1-2\alpha)^2}$$

$$= \frac{\alpha \sqrt{\gamma}}{(2\alpha)^{1/\gamma} (1-2\alpha)^2} (1+O(\gamma)), \quad \gamma \rightarrow 0, \quad F_\gamma(\xi_{1-\alpha}) = 1-\alpha,$$

$$\sigma_l^{*2} = (1-2\alpha)^2 \sigma_l^2 + \frac{4\alpha^2 \xi_{1-\alpha}}{p_\gamma(\xi_{1-\alpha})} + \frac{2\alpha^3}{p_\gamma^2(\xi_{1-\alpha})} \sim \frac{\alpha}{2\gamma(2\alpha)^{2/\gamma}}, \quad \gamma \rightarrow 0,$$

и АОЭ оценки $\bar{L}_n$ относительно оценок M_n , W_n , δ_n^* и $\bar{L}_n^*$ соответственно равны

$$e_{lm} = \frac{\sigma_m^2}{\sigma_l^2} \sim \frac{(2\alpha)^{1/\gamma} (1-2\alpha)^2}{\alpha \gamma^{3/2}}, \quad e_{lw} = \frac{\sigma_w^2}{\sigma_l^2} \sim \frac{4(2\alpha)^{1/\gamma} (1-2\alpha)^2}{3\alpha \gamma^{7/2} \pi^2},$$

$$e_{l*} = \frac{1}{I_\gamma \sigma_l^2} \sim \frac{3(2\alpha)^{1/\gamma} (1-2\alpha)^2}{\alpha \sqrt{\gamma}}, \quad e_{l*} = \frac{\sigma_l^{*2}}{\sigma_l^2} \sim \frac{(1-2\alpha)^2}{2\gamma^{3/2} (2\alpha)^{1/\gamma}}, \quad \gamma \rightarrow 0.$$

Д о к а з а т е л ь с т в о. Первые четыре утверждения теоремы непосредственно следуют из формул (1.1), (3.3), (3.5), (3.10), теоремы 5.3.2 из [15], устанавливающей асимптотическую нормальность оценки M_n , асимптотической нормальности оценки W_n (см. [15] и [20]) и непрерывности функции $\Gamma(t)$ по $t > 0$. Пятое утверждение теоремы следует из третьего утверждения леммы 3.1 и асимптотической нормальности оценок $\bar{L}_n$ и $\bar{L}_n^*$ (см. [14, теоремы 3.1 и 3.2]).

3.3. М-оценки и оценки максимального правдоподобия. Как уже отмечалось в п. 3.1, усеченное среднее $\bar{L}_n$ (см. (3.8)) является естественным компромиссом между средним $\bar{X}_n$ (см. (3.6)) и медианой M_n (см. (3.7)). Другой компромисс, предложенный Хьюбером [23], подсказан тем фактом, что среднее и медиана являются величинами, минимизирующими по $a \in \mathbf{R}$ соответственно функции $\sum_{i=1}^n (X_i - a)^2$ и $\sum_{i=1}^n |X_i - a|$. Хьюбер предложил минимизировать более общее выражение вида

$$\sum_{i=1}^n \rho(X_i - a) \quad (3.11)$$

с некоторой разумно выбранной функцией $\rho(x)$ (например, выпуклой и четной). Оценки, минимизирующие выражение (3.11), будем называть *М-оценками*. Если функция $\rho(x)$ дифференцируема, то при соответствующих условиях регулярности (см. [23]), М-оценка является решением уравнения

$$\sum_{i=1}^n \rho'(X_i - a) = 0.$$

Заметим также, что М-оценки являются эквивариантными и оценки максимального правдоподобия (в случае параметра сдвига) являются М-оценками с функцией $\rho(x) = -\ln p(x)$. Рассмотрим конкретную функцию $\rho(x)$, предложенную в работе [23]:

$$\rho(x) = \begin{cases} \frac{x^2}{2}, & |x| \leq k, \\ k|x| - \frac{k^2}{2}, & |x| \geq k. \end{cases}$$

Эта функция пропорциональна x^2 при $|x| \leq k$, а вне этого отрезка заменяет ветви параболы на прямые линии. Эти куски согласованы так, что функция $\rho(x)$ и ее производная непрерывны. Оценки, минимизирующие выражение (3.11) с этой функцией, будем обозначать H_n . При растущем k функция $\rho(x)$ будет совпадать с $x^2/2$ на все большей части области ее определения, так что оценка H_n будет сближаться с $\bar{X}_n$. При уменьшении k оценка H_n будет сближаться с медианой M_n .

Обозначим оценку максимального правдоподобия через $\hat{\delta}_n$. Справедлив следующий результат.

Теорема 3.2. Пусть случайные величины $X_1, \dots, X_n$ независимы, имеют одинаковую функцию распределения $F_\gamma(x - \theta)$ и плотность $p_\gamma(x) = F'_\gamma(x) > 0$ (см. (1.1)). Тогда справедливы следующие утверждения.

1. Оценка максимального правдоподобия $\hat{\delta}_n$ удовлетворяет уравнению

$$\sum_{i=1}^n \frac{X_i - a}{\gamma + (X_i - a)^2} = 0,$$

и при каждом $C > 0$ справедливо соотношение

$$\sup_{|x| \leq C} \left| \mathbf{P}_\theta \left\{ \sqrt{n} I_\gamma (\hat{\delta}_n - \theta) < x \right\} - \Phi_n(x) \right| = O(n^{-3/2}),$$

где

$$\begin{aligned} \Phi_n(x) &= \Phi(x) + \frac{x\phi(x)}{n} \left(\frac{x^2}{72} (5a_1 - 12a_2 + 9) - \frac{a_1 - 3}{24} \right), \\ \alpha_1 &= \frac{3\gamma(\gamma+2)(\gamma+3)}{(\gamma+1)(\gamma+5)(\gamma+7)}, \\ \alpha_2 &= \frac{2(\gamma+2)(\gamma+3)(\gamma^2+5\gamma+10)}{\gamma(\gamma+1)(\gamma+5)(\gamma+7)}, \quad I_\gamma = \frac{\gamma+1}{\gamma+3}. \end{aligned}$$

2. Для дисперсии M -оценки H_n справедливо соотношение

$$\sigma_h^2 = \frac{4}{(2F_\gamma(k) - 1)^2} \int_0^k x(1 - F_\gamma(x)) dx \sim \frac{4k^2}{\gamma^2(\ln \gamma - 2 \ln 2k)^2}, \quad \gamma \rightarrow 0,$$

и для АОЭ M -оценки H_n относительно оценок M_n , $\bar{L}_n$, W_n , δ_n^* и $\bar{L}_n^*$ (см. (3.7)–(3.9), (3.1)) соответственно справедливы соотношения

$$\begin{aligned} e_{hm} &= \frac{\sigma_m^2}{\sigma_h^2} \sim \frac{\gamma(\ln \gamma - 2 \ln 2k)^2}{4k^2}, & e_{hl} &= \frac{\sigma_l^2}{\sigma_h^2} \sim \frac{\alpha \gamma^{5/2} (\ln \gamma - 2 \ln 2k)^2}{(2\alpha)^{1/\gamma} (1 - 2\alpha)^{24k^2}}, \\ e_{hw} &= \frac{\sigma_w^2}{\sigma_h^2} \sim \frac{(\ln \gamma - 2 \ln 2k)^2}{3\gamma \pi^2 k^2}, & e_{h*} &= \frac{1}{I_\gamma \sigma_h^2} \sim \frac{3\gamma^2 (\ln \gamma - 2 \ln 2k)^2}{4k^2}, \\ e_{hl*} &= \frac{\sigma_l^{*2}}{\sigma_h^2} \sim \frac{\gamma \alpha (\ln \gamma - 2 \ln 2k)^2}{8k^2 (2\alpha)^{2/\gamma}}, & \gamma &\rightarrow 0. \end{aligned}$$

Доказательство теоремы состоит в непосредственной проверке условий леммы 7.1 из [21], устанавливающей асимптотическую нормальность оценки $\hat{\delta}_n$, и использовании леммы 4 из [23], устанавливающей асимптотическую нормальность оценки H_n .

СПИСОК ЛИТЕРАТУРЫ

1. Student. On the probable error of the mean. — *Biometrika*, 1908, v. 6, № 1, p. 1–25.
2. Praetz P. D. The distribution of share prices changes. — *J. Business*, 1972, v. 45, p. 49–55.
3. Blattberg R., Gonedes N. A comparison of the stable and Student distributions as statistical models of stock prices. — *J. Business*, 1974, v. 47, p. 244–250.
4. Королев В. Ю. Сходимость случайных последовательностей с независимыми случайными индексами. I. — *Теория вероятн. и ее примен.*, 1994, т. 39, в. 2, с. 313–333.
5. Королев В. Ю. Сходимость случайных последовательностей с независимыми случайными индексами. II. — *Теория вероятн. и ее примен.*, 1995, т. 40, в. 4, с. 907–910.
6. Korolev V. Yu. A general theorem on the limit behavior of superpositions of independent random processes with applications to Cox processes. — *J. Math. Sci.*, 1996, v. 81, № 5, p. 2951–2956.

7. Золотарев В. М. Одномерные устойчивые распределения. М.: Наука, 1983, 304 с.
8. Uchaikin V. V., Zolotarev V. M. Chance and Stability. Stable Distributions and Their Applications. Utrecht: VSP, 1999, 570 p.
9. Гнеденко Б. В. Об оценивании неизвестных параметров распределений по случайному числу независимых наблюдений. — Теория вероятностей и математическая статистика. Труды Тбилисского математического института им. А. М. Размадзе, 1989, с. 146–150.
10. Gurland J. Some interrelations among compound and generalized distributions. — Biometrika, 1957, v. 44, p. 265–268.
11. Quenouille M. H. A relation between the logarithmic, Poisson, and negative binomial series. — Biometrics, 1949, v. 5, p. 162–164.
12. Кендалл М. Дж., Стьюарт А. Теория распределений. М.: Наука, 1966, 587 с.
13. Greenwood M., Udney Yule G. An inquiry into the nature of frequency-distributions of multiple happenings, etc. — J. Roy. Statist. Soc., 1920, v. 83, p. 255–279.
14. Bickel P. J. On some robust estimates of location. — Ann. Math. Statist., 1965, v. 36, № 3, p. 847–858.
15. Леман Э. Теория точечного оценивания. М.: Наука, 1991, 443 с.
16. Ибрагимов И. А., Хасьминский Р. З. Асимптотическая теория оценивания. М.: Наука, 1979, 527 с.
17. Stone C. J. Asymptotic properties of estimators of a location parameter. — Ann. Statist., 1974, v. 2, p. 1127–1137.
18. Port S. C., Stone C. J. Fisher information and the Pitman estimator of a location parameter. — Ann. Statist., 1974, v. 2, p. 225–247.
19. Serfling R. J. Approximation Theorems of Mathematical Statistics. New York: Wiley, 1980, 371 p.
20. Hodges J. L., Lehmann E. L. Estimates of location based on rank tests. — Ann. Math. Statist., 1963, v. 34, p. 598–611.
21. Albers W., Bickel P. J., van Zwet W. R. Asymptotic expansions for the power of distribution free tests in the one-sample problem. — Ann. Statist., 1976, v. 4, № 1, p. 108–156.
22. Helmers R. Edgeworth Expansions for Linear Combinations of Order Statistics. Amsterdam: Mathematisch Centrum, 1984, 137 p.
23. Huber P. J. Robust estimation of a location parameter. — Ann. Math. Statist., 1964, v. 35, № 1, p. 73–101.
24. Феллер В. Введение в теорию вероятностей и ее приложения, т. 2. М.: Мир, 1984, 746 с.
25. Двайт Г. Б. Таблицы интегралов и другие математические формулы. М.: Наука, 1977, 224 с.
26. Kapur J. N. Maximum-Entropy Models in Science and Engineering. New York: Wiley, 1989, 635 p.
27. Tsallis C., de Souza A. M. C., Maynard R. Derivation of Lévy-type anomalous superdiffusion from general statistical mechanics. — Lévy Flights and Related Topics in Physics. Ed. by M. F. Shlesinger, G. M. Zaslavsky, and U. Frisch. Berlin: Springer-Verlag, 1995, p. 269–289.
28. Hill B. M. A simple general approach to inference about the tail of a distribution. — Ann. Statist., 1975, v. 3, p. 1163–1174.
29. Resnick S. I. Heavy tail modeling and teletraffic data. — Ann. Statist., 1997, v. 25, № 5, p. 1805–1869.
30. Embrechts P., Klüppelberg C., Mikosch T. Modeling Extremal Events for Insurance and Finance. Berlin: Springer-Verlag, 1997, 645 p.
31. Петров В. В. Суммы независимых случайных величин. М.: Наука, 1972.
32. Teicher H. Identifiability of mixtures. — Ann. Math. Statist., 1961, v. 32, p. 244–248.

Поступила в редакцию
7.V.2003