Записки научных
семинаров ПОМИ
Том 442, 2015 г.

А. В. Чирина

БОЛЬШИЕ УКЛОНЕНИЯ СУММ ОГРАНИЧЕННЫХ ФУНКЦИЙ ОТ НОРМИРОВАННОЙ ВЫБОРКИ ИЗ ГАММА-РАСПРЕДЕЛЕНИЯ

§1. Введение

Пусть $X_1,\ldots,X_n$ – независимые одинаково распределенные случайные величины, имеющие Г-распределение, плотность которого задается формулой

$$f(x) = \frac{x^\alpha e^{-x}}{\Gamma(1+\alpha)}, \quad x > 0.$$

где $\alpha > -1$ – параметр формы, при этом параметр масштаба мы можем, не умаляя общности, считать равным 1. Пусть $\overline{X}$ – их выборочное среднее. Обозначим $Y_i = X_i/\overline{X}$, $i = 1,\ldots,n$. Пусть на неотрицательной полупрямой задана вещественная функция $r(s)$. Определим с ее помощью статистику

$$T_n = n^{-1}\sum_{i=1}^{n} r(Y_i). \tag{1}$$

В [6] изучалась грубая асимптотика больших уклонений статистик вида (1) для монотонных функций $r(s)$, удовлетворяющих ряду ограничений. Ограничения касались не только функции $r(s)$, но и области изменения неотрицательного аргумента функции уклонений. Эти ограничения были обусловлены способом доказательства и являлись неоправданно сильными. В то же время для некоторых приложений (как, например, вычисление больших уклонений статистик типа супремума) требуется знать поведение функции уклонений не только вблизи нуля, но и на заранее известном интервале, а иногда и на всей положительной полуоси. Такая задача была решена в [6] лишь частично – для ограниченных монотонных $r(s)$, причем для возрастающих $r(s)$ требовалось дополнительное условие выпуклости.

Ключевые слова: большие уклонения, нормированная выборка, гамма-распределение, распределение Дирихле, критерий экспоненциальности, преобразование Ханкеля .

Заметим, что при любом $\beta \in \mathbb{R}$

$$T_n = n^{-1} \sum_{i=1}^{n} r(Y_i) = n^{-1} \sum_{i=1}^{n} \big(r(Y_i) + \beta \cdot (Y_i - 1)\big), \tag{2}$$

так что функции $r_\beta(s) = r(s) + \beta \cdot (s-1)$ при любом β определяют одну и ту же статистику. При этом монотонная и ограниченная версии $r_\beta(s)$ могут соответствовать разным значениям β. Настоящая статья посвящена именно этому случаю.

Статья организована следующим образом. В разделе 2 вычисляется функция уклонений

$$k(a) = \lim_{n\to\infty} n^{-1} \ln \mathbf{P}(T_n > a)$$

и описывается ее поведение при малых положительных значениях аргумента. В разделе 3 доказывается справедливость полученного результата при любых положительных значениях a, а также найден интервал, в котором значения функции уклонений отличаются от $-\infty$. В разделе 4 приводится пример применения полученных результатов к недавно предложенной в [2] статистике типа супремума. Также описывается возможный путь дальнейшего ослабления условий на $r(s)$.

§2. Вычисление функции уклонений

Цель этого раздела – вычисление функции

$$k(a) = \lim_{n\to\infty} n^{-1} \ln \mathbf{P}(T_n > a), \quad a > 0, \tag{3}$$

где случайные величины $Y_i, \quad i = 1, \dots, n$ построены по выборке X_i, взятой из Г-распределения с параметром формы $1+\alpha$, а статистика T_n определена в (1). Для этого мы модифицируем теорему 1 из [6] и предшествующие ей леммы. Как и в указанной статье, будем использовать теорему из [5], ключевую роль в которой играет функция

$$\varphi(\mu) = \lim_{n\to\infty} n^{-1} \ln \mathbf{E} \exp\left(\mu \sum_{i=1}^{n} r(Y_i)\right), \quad \mu > 0. \tag{4}$$

Сформулируем условия на $r(x)$, достаточные для того, чтобы предел в (4) существовал при любых положительных μ.

Пусть существуют такие положительные константы ρ и β, что

$$\forall\, s > 0 \quad |r(s)| \leqslant \rho \tag{5}$$

и

$$r_\beta(s) \equiv r(s) + \beta \cdot (s-1) \text{ возрастает на положительной полуоси.} \quad (6)$$

Можем также, не умаляя общности, считать $r(s)$ и $r_\beta(s)$ непрерывными справа (иначе, по ограниченности первой и монотонности второй из них, можем переопределить значения в не более чем счетном числе точек, что не повлияет на распределение статистики). В частности, для выполнения условия (6) достаточно, чтобы функция $r(s)$ удовлетворяла условию Липшица.

Для упрощения вычислений можно также считать, что

$$\int_0^\infty s^\alpha r\Big(\frac{s}{1+\alpha}\Big) e^{-s}\, ds = 0, \quad (7)$$

иначе можно вычесть центрирующую константу.

Введем функцию

$$\begin{aligned} \psi(\mu,\lambda) &= (\lambda-1)(1+\alpha) \\ &\quad + \ln \int_0^\infty s^\alpha \exp\Big(-\lambda s + \mu r\Big(\frac{s}{1+\alpha}\Big)\Big)\, ds - \ln\Gamma(1+\alpha). \end{aligned} \quad (8)$$

Поскольку функция $r(s)$ ограничена, функция $\psi(\mu,\lambda)$ принимает конечные значения при любых положительных λ и μ. Более того, при каждом фиксированном μ эта функция стремится к ∞, если λ стремится к 0 или к ∞, поэтому при неположительных λ можно считать, что $\psi(\mu,\lambda) = \infty$. Функция $\psi(\mu,\lambda)$ является производящей функцией моментов для случайного вектора $(r(\frac{X_i}{1+\alpha}), 1+\alpha-X_i)$ с аргументами $(\mu, \lambda-1)$, поэтому она дифференцируема и выпукла при всех положительных значениях μ и λ. В частности, отсюда следует

Замечание 1. При каждом фиксированном μ инфимум по λ функции $\psi(\mu,\lambda)$ достигается в точке $\lambda(\mu):\ 0 < \lambda(\mu) < \infty$.

Как и в [6], можно доказать, что $\lim_{\mu\to 0+} \lambda(\mu) = 1$, а также найти $\lambda'(\mu)$ в нуле. Эти вычисления почти полностью повторяют аналогичные выкладки в [6], поэтому здесь мы их опускаем.

Введем теперь дополнительно функцию

$$\psi_\beta(\mu,\lambda) = (\lambda-1)(1+\alpha) + \ln\int\limits_0^\infty s^\alpha \exp\left(-\lambda s + \mu r_\beta\left(\frac{s}{1+\alpha}\right)\right) ds - \ln\Gamma(1+\alpha).$$

Замечание 2. В [6] доказано, что для возрастающих функций $r_\beta(s)$

$$\inf_\lambda \psi_\beta(\mu,\lambda)$$

достигается при $\lambda > 1$.

Несложные вычисления показывают, что

$$\psi_\beta(\mu,\lambda) = \psi\Big(\mu, \lambda - \frac{\mu\beta}{1+\alpha}\Big) \tag{9}$$

Из теоремы 1 в [6] непосредственно следует, что при $\mu\beta < 1+\alpha$

$$\varphi(\mu) = \inf_{\lambda>1} \psi_\beta(\mu,\lambda) = \inf_{\lambda>1} \psi\Big(\mu, \lambda - \frac{\mu\beta}{1+\alpha}\Big) = \inf_{\lambda > 1-\frac{\mu\beta}{1+\alpha}} \psi(\mu,\lambda), \tag{10}$$

причем из выпуклости $\psi_\beta(\mu,\lambda)$ и $\psi(\mu,\lambda)$ и замечания 1 следует, что при $\mu\beta<1+\alpha$

$$\varphi(\mu) \equiv \lim_{n\to\infty} n^{-1} \ln \mathbf{E} \exp \mu \sum_{i=1}^n r(Y_i) = \inf_{\lambda>0} \psi(\mu,\lambda) = \inf_{\lambda\in\mathbb{R}} \psi(\mu,\lambda). \tag{11}$$

Наша задача теперь состоит в том, чтобы доказать последнее равенство для всех положительных μ. Для этого докажем две леммы, в которых оценим искомую функцию сверху и снизу.

Лемма 1. *При всех положительных μ выполнено неравенство*

$$\limsup_{n\to\infty} n^{-1} \ln \mathbf{E} \exp\Big(\mu \sum_{i=1}^n r(Y_i)\Big) \leqslant \inf_{\lambda>0} \psi(\mu,\lambda)$$

Доказательство. Перейдем от функции $r(s)$ к возрастающей функции $r_\beta(s)$, для которой, согласно лемме 3 из [6], выполнено неравенство

$$\limsup_{n\to\infty} n^{-1} \ln \mathbf{E} \exp\Big(\mu \sum_{i=1}^n r_\beta(Y_i)\Big) \leqslant \inf_{\lambda>1} \psi_\beta(\mu,\lambda)$$

Но, по замечаниям 1 и 2 и по (11), этот инфимум является глобальным и совпадает с $\inf_{\lambda>0} \psi(\mu,\lambda)$ □

Лемма 2. *При всех положительных μ выполнено неравенство*

$$\liminf_{n\to\infty} n^{-1}\ln\mathbf{E}\exp\biggl(\mu\sum_{i=1}^{n} r(Y_i)\biggr) \geqslant \inf_{\lambda>0}\psi(\mu,\lambda).$$

Доказательство. Имеем

$$\begin{aligned}\ln\mathbf{E}\exp\mu\sum_{i=1}^{n} r_\beta\Bigl(Y_i\Bigr)&\\ =\ln\int_0^\infty\dots\int_0^\infty\prod_{i=1}^{n} s_i^\alpha\exp\biggl(\sum_{i=1}^{n}\Bigl(-s_i+\mu r_\beta\Bigl(\frac{s_i}{\overline{s}}\Bigr)\Bigr)\biggr)ds_1\dots ds_n&\\ -\ln(\Gamma(1+\alpha))\qquad(12)&\end{aligned}$$

Возьмем произвольную константу $M>0$. Зафиксируем μ и введем в рассмотрение независимые одинаково распределенные случайные величины $Z_i^{(\beta,\mu,M)}$, $i=1,\dots,n$, каждая из которых имеет плотность

$$g_{\beta,\mu,M}(s)=\frac{s^\alpha\exp\Bigl(-s+\mu r_\beta\Bigl(\frac{s}{1+\alpha}\Bigr)\Bigr)\cdot\mathbb{I}_{0<s<M}}{\int_0^M t^\alpha\exp\Bigl(-t+\mu r_\beta\Bigl(\frac{t}{1+\alpha}\Bigr)\Bigr)dt}.$$

Обозначим их выборочное среднее $\overline{Z}^{(\beta,\mu,M)}$.

Поскольку подынтегральная функция всюду положительна, с учетом возрастания r_β получаем следующую цепочку неравенств:

$$\begin{aligned}\ln\mathbf{E}\exp\mu\sum_{i=1}^{n} r_\beta\bigl(Y_i\bigr)&\\ \geqslant\ln\mathbf{E}\exp\biggl(\mu\sum_{i=1}^{n}\biggl(r_\beta\Bigl(\frac{Z_i^{(\beta,\mu,M)}}{\overline{Z}^{(\beta,\mu,M)}}\Bigr)-r_\beta\Bigl(\frac{Z_i^{(\beta,\mu,M)}}{1+\alpha}\Bigr)\biggr)\biggr)&\\ +n\ln\mathbf{E}\,\mathbb{I}_{\{X_1<M\}}\exp\biggl(r_\beta\Bigl(\frac{X_1}{1+\alpha}\Bigr)\biggr)&\\ \geqslant\ln\mathbf{P}\bigl(\overline{\mathbf{Z}}^{(\beta,\mu,\mathbf{M})}\leqslant\mathbf{1}+\alpha\bigr)+n\ln\mathbf{E}\,\mathbb{I}_{\{X_1<M\}}\exp\biggl(r_\beta\Bigl(\frac{X_1}{1+\alpha}\Bigr)\biggr).\qquad(13)&\end{aligned}$$

Применим теперь к $\overline{Z}^{(\beta,\mu,M)}$ теорему Чернова ([1, 3, 4]) о больших уклонениях выборочного среднего:

$$\lim_{n\to\infty} \ln \mathbf{P}\big(\overline{Z}^{(\beta,\mu,M)} \leqslant 1+\alpha\big) = \inf_{\nu>0} \ln \mathbf{E} \exp\big(\nu\big(1+\alpha - Z_1^{(\beta,\mu,M)}\big)\big)$$
$$\geqslant \inf_{\nu\in\mathbb{R}} \ln \mathbf{E} \exp\big(\nu\big(1+\alpha - Z_1^{(\beta,\mu,M)}\big)\big). \quad (14)$$

Поскольку

$$\ln \mathbf{E} \exp\big(\nu\big(1+\alpha - Z_1^{(\beta,\mu,M)}\big)\big)$$
$$= \ln \int_0^M s^\alpha \exp\Big(\nu(1+\alpha-s) - s + \mu\, r_\beta\Big(\frac{s}{1+\alpha}\Big)\Big)\, ds$$
$$- \ln \int_0^M s^\alpha \exp\Big(-s + \mu\, r_\beta\Big(\frac{s}{1+\alpha}\Big)\Big) ds,$$

причем второе слагаемое не зависит от ν и представляется в виде

$$\ln \int_0^M s^\alpha \exp\Big(-s + \mu\, r_\beta\Big(\frac{s}{1+\alpha}\Big)\Big) ds$$
$$= \mathbf{E}\mathbb{I}_{\{X_1<M\}} \exp\Big(r_\beta\Big(\frac{X_1}{1+\alpha}\Big)\Big) + \ln\Gamma(1+\alpha).$$

из (12)–(14) получаем оценку

$$\liminf_{n\to\infty}\ n^{-1} \ln \mathbf{E} \exp\Big(\mu \sum_{i=1}^n r_\beta(Y_i)\Big)$$
$$\geqslant \inf_{\nu\in\mathbb{R}} \ln \int_0^M s^\alpha \exp\Big(\nu(1+\alpha-s) - s + \mu r_\beta\Big(\frac{s}{1+\alpha}\Big)\Big) ds - \ln\Gamma(1+\alpha)$$
$$= \inf_{\nu\in\mathbb{R}} \Big((\nu(1+\alpha) - \mu\beta) + \ln \int_0^M s^\alpha \exp\Big(-s\Big(1+\nu - \frac{\mu\beta}{1+\alpha}\Big) + \mu r\Big(\frac{s}{1+\alpha}\Big)\Big) ds\Big)$$
$$- \ln\Gamma(1+\alpha) \quad (15)$$

Возьмем $\lambda = 1 + \nu - \frac{\mu\beta}{1+\alpha}$ и введем

$$\psi(\mu,\lambda,M) = (\lambda-1)(1+\alpha) + \ln\int\limits_0^M s^\alpha \exp\left(-\lambda s + \mu\, r\left(\frac{s}{1+\alpha}\right)\right) ds - \ln\Gamma(1+\alpha).$$

Эта функция при любых фиксированных $\mu > 0$ и $M > 0$ выпукла по λ, поскольку является суммой логарифма производящей функции моментов случайной величины $Z_1^{(\beta,\mu,M)}$ и слагаемого, не зависящего от λ. Тогда при любом $M > 0$ (15) можно переписать следующим образом:

$$\begin{aligned}\liminf_{n\to\infty} n^{-1}\ln \mathbf{E}\exp\left(\mu\sum_{i=1}^n r_\beta(Y_i)\right) &= \liminf_{n\to\infty} n^{-1}\ln \mathbf{E}\exp\left(\mu\sum_{i=1}^n r(Y_i)\right)\\ &\geqslant \inf_{\lambda\in\mathbb{R}} \psi(\mu,\lambda,M)\end{aligned}$$

и, следовательно,

$$\liminf_{n\to\infty} n^{-1}\ln \mathbf{E}\exp\left(\mu\sum_{i=1}^n r(Y_i)\right) \geqslant \sup_{M>M_0}\inf_{\lambda\in\mathbb{R}} \psi(\mu,\lambda,M) \qquad (16)$$

при любом $M_0 > 0$.

Так как при любых фиксированных λ и μ функция $\psi(\mu,\lambda,M)$ возрастает по M и $\lim_{M\to\infty}\psi(\mu,\lambda,M) = \sup_{M>M_0}\psi(\mu,\lambda,M) = \psi(\mu,\lambda)$, для завершения доказательства леммы нам осталось лишь убедиться в том, что можно поменять местами супремум по M и инфимум по λ. Докажем вначале, что при достаточно больших M $\inf_\lambda \psi(\mu,\lambda,M)$ достигается на некотором компакте $[\lambda_1,\lambda_2] \subset (0,\infty)$, который может зависеть от μ, но не от M. Пусть λ_0 – точка, в которой достигается $\inf_\lambda \psi(\mu,\lambda)$. Возьмем такие положительные λ_1 и λ_2, что $\lambda_1 < \lambda_0 < \lambda_2$. Пусть $\delta = \max(\psi(\mu,\lambda_1) - \psi(\mu,\lambda_0), \psi(\mu,\lambda_2) - \psi(\mu,\lambda_0))$. Из сходимости возрастающей последовательности интегралов, определяющих $\psi(\mu,\lambda_M)$, к $\psi(\mu,\lambda)$, следует, что можно определить такое M_0, что для любого $M > M_0$

$$0 < \psi(\mu,\lambda_i,M) - \psi(\mu,\lambda_i,M_0) < \delta/3, \quad i = 0,1,2.$$

Но тогда $\psi(\mu,\lambda_1,M) - \psi(\mu,\lambda_0,M) >$ и $\psi(\mu,\lambda_2,M) - \psi(\mu,\lambda_0,M)$, а это значит, что инфимум по λ функции $\psi(\mu,\lambda,M)$ лежит между λ_1 и λ_2, так как $\psi(\mu,\lambda,M)$ выпукла по λ. Теперь можем применить следствие 2

из теоремы 1 в [7]. Согласно этому следствию,

$$\sup_{M>M_0} \inf_{\lambda\in[\lambda_1,\lambda_2]} \psi(\mu,\lambda,M) = \inf_{\lambda\in[\lambda_1,\lambda_2]} \sup_{M>M_0} \psi(\mu,\lambda,M),$$

что вместе с (16) и завершает доказательство леммы. □

Комбинируя леммы 1 и 2, получим следующее утверждение:

Теорема 1. *При выполнении условий* (5) *и* (6) *предельная функция в* (4) *при любых положительных* μ *существует и равна*

$$\varphi(\mu) \equiv \lim_{n\to\infty} n^{-1} \ln \mathbf{E} \exp \mu \sum_{i=1}^{n} r(Y_i) = \inf_{\lambda>0} \psi(\mu,\lambda) = \inf_{\lambda\in\mathbb{R}} \psi(\mu,\lambda). \quad (17)$$

Как и в [6], можно доказать, что функция $\varphi(\mu)$ является гладкой и выпуклой, и вычислить ее первую и вторую производные справа в нуле. Все выкладки абсолютно аналогичны тем, которые приводятся в указанной статье, поэтому здесь мы не будем их повторять.

Из выпуклости следует, что первая производная функции $\varphi(\mu)$ возрастает, а из ограниченности $r(s)$ вытекает ограниченность $\varphi'(\mu)$. Пусть $a_0 = \lim_{\mu\to\infty} \varphi'(\mu)$. Применяя теорему Плашки–Штайнебаха из [5], получим основной результат этого раздела и всей статьи:

Теорема 2. *При выполнении условий* (5) *и* (6) *функция уклонений статистики* (1) *от выборки из гамма-распределения для любого положительного* $a < a_0$ *может быть найдена по формуле*

$$k(a) = \inf_{\mu>0,\lambda>0} (-a\mu + \ln \psi(\mu,\lambda)), \quad (18)$$

где функция $\psi(\mu,\lambda)$ *определена в* (8). *При этом функция* $k(a)$ *непрерывна при* $a < a_0$, *а в нуле при выполнении дополнительного условия* (7) *справедлива асимптотика*

$$\lim_{a\to 0+} a^{-2} k(a) = \frac{1+\alpha}{2\,\sigma^2 - (1+\alpha)\,\tau^2}\ , \quad (19)$$

где

$$\sigma^2 = \mathbf{E}\, r^2\left(\frac{X_1}{1+\alpha}\right), \quad \tau = \mathbf{E}\, X_1\, r\left(\frac{X_1}{1+\alpha}\right).$$

§3. Область определения функции уклонений

Как отмечалось выше, функция $\varphi'(\mu)$ возрастает и ограничена на $(0,\infty)$. Пусть $a_0 = \lim\limits_{a\to\infty} \varphi'(\mu)$, как и в предыдущем разделе. Тогда a_0 можно вычислить как точную нижнюю границу всех a, для которых $k(a) = -\infty$. Доказательство этого факта мы опускаем, поскольку доказательство аналогичной леммы 9 из [6] без труда переносится на случай функции $r(s)$, удовлетворяющей (5) и (6).

Введем функцию, $r^c(s)$ следующим образом: $-r^c(s)$ – это выпуклая оболочка $-r(s)$ на $[0,\infty)$. Иными словами, подграфик $r^c(s)$ – это пересечение подграфиков всех вогнутых на $[0,\infty)$ функций, превосходящих или равных $r(s)$, а сама $r(s)$ – наименьшая вогнутая функция, превосходящая или равная $r(s)$.

Теорема 3. *При выполнении условий* (5) *и* (6) $a_0 = r^c(1)$.

Доказательство. Так как $\sum\limits_{i=1}^{n} Y_i = n$, по неравенству Йенсена,

$$r^c(1) = r^c\left(\sum_{i=1}^{n} \frac{1}{n} Y_i\right) \geqslant \sum_{i=1}^{n} \frac{1}{n} r^c(Y_i) \geqslant \sum_{i=1}^{n} \frac{1}{n} r(Y_i), \tag{20}$$

а отсюда следует, что $a_0 \leqslant r^c(1)$.

Для доказательства противоположного неравенства нам потребуется следующая лемма.

Лемма 3. *Пусть* $a_0 = r(1)$. *Для любого* $a < a_0$ *существуют такие* $\lambda \in [0,1]$, $\delta > 0$, $s_1 > 0$, $s_2 > 0$, *что* $\lambda(s_1+\delta) + (1-\lambda)(s_2+\delta) \leqslant 1$ *и для любых* $t_1 \in (s_1, s_1+\delta)$ *и* $t_2 \in (s_2, s_2+\delta)$ *выполнено неравенство* $\lambda r(t_1) + (1-\lambda) r(t_2) \geqslant a$.

Доказательство. Обозначим $a_0 - a = \varepsilon$. Поскольку функция $r^c(s)$ выпукла и ограничена, она непрерывна на $(0,\infty)$. Поэтому существует такое $s_0 < 1$, что для любого $s \in [s_0, 1)$ выполнено неравенство $r^c(s) \geqslant a + \varepsilon/2$.

Поскольку точка $(s_0, r^c(s_0))$ находится на границе подграфика функции $r^c(s)$, существуют такое $\lambda \in [0,1]$ и такие $s_1 > 0$, $s_2 > 0$, что $s_0 = \lambda s_1 + (1-\lambda) s_2$ и $r(s_0) = \lambda r(s_1) + (1-\lambda) r(s_2)$.

Обозначим $\delta_0 = 1 - s_0$. Очевидно, для любых $t_1 \in [s_1, s_1+\delta_0]$, $t_2 \in [s_2, s_2+\delta_0]$ выполнено неравенство $\lambda t_1 + (1-\lambda) t_2 \leqslant 1$ и, в частности, $\lambda(s_1+\delta_0) + (1-\lambda)(s_2+\delta_0) = 1$.

По непрерывности функции $r(s)$ справа, существуют такие положительные δ_1 и δ_2, что для любых $t_1 \in [s_1, s_1+\delta_1]$, $t_2 \in [s_2, s_2+\delta_2]$ выполнены неравенства $r(t_1) \geqslant r(s_1) - \varepsilon/2$ и $r(t_2) \geqslant r(s_2) - \varepsilon/2$, так что

$$\begin{aligned}\lambda r(t_1) + (1-\lambda) r(t_2) &\geqslant \lambda(r(s_1) - \varepsilon/2) + (1-\lambda)(r(s_2) - \varepsilon/2)\\ &= \lambda r(s_1) + (1-\lambda) r(s_2) - \varepsilon/2 = r(x_0) - \varepsilon/2.\end{aligned}$$

В качестве δ выберем наименьшее из чисел δ_0, δ_1 и δ_2, что и завершает доказательство леммы 3. □

Замечание 3. Если функция $r^c(s)$ совпадает с $r(s)$ в окрестности единицы, то можно взять $s_1 = s_2$.

Вернемся к доказательству теоремы 3. Как известно, случайный вектор $(Y_1, \dots, Y_{n-1})$ имеет симметричное распределение Дирихле на симплексе

$$\Delta_n = \{t \equiv (t_1, \dots, t_{n-1}) \big|\ t_i \geqslant 0, \quad i = 1, \dots, n-1; \quad \sum_{i=1}^{n-1} t_i \leqslant n\},$$

при этом

$$Y_n = n - \sum_{i=1}^{n-1} Y_i.$$

Докажем, что можно выбрать такое δ и такой $(n-1)$-мерный куб Q_n с ребром δ, полностью содержащийся в Δ_n и отделенный от 0, что для любой точки $t \in Q_n$ и для любого $a < a_0$ выполнено неравенство

$$\sum_{i=1}^{n} r(t_i) \geqslant a, \quad \text{где } t_n = n - \sum_{i=1}^{n-1} t_i > 1. \tag{21}$$

Пусть λ, s_1, s_2 и δ построены по функции $r(s)$ так, как в лемме 3. Возьмем натуральное число $k = [\lambda n]$. Построим куб Q_n как прямое произведение k копий отрезка $(s_1,\ s_1+\delta)$ и $n-k-1$ копий отрезка $(s_2,\ s_2+\delta)$. Тогда, по лемме 3 и с учетом ограниченности $r(x)$, для любого $t \in Q_n$ неравенство (21) выполняется с точностью до слагаемого порядка n^{-1}. Для вероятности попадания в куб выполнено соотношение

$$n^{-1} \ln \mathbf{P}((Y_1, \dots, Y_{n-1}) \in Q_n) > C > -\infty,$$

где константа C зависит от s_1, s_2 и δ (это легко получается оценкой интеграла от плотности распределения Дирихле по кубу, см. аналогичные рассуждения в [6]).

Поскольку $a < a_0$ произвольно и поскольку при достаточно больших n

$$\mathbf{P}\left(n^{-1}\sum_{i=1}^{n} r(Y_i) > a\right) \geqslant \mathbf{P}((Y_1,\dots,Y_{n-1}) \in Q_n) + O\big(n^{-1}\big),$$

теорема доказана полностью. □

§4. Примеры применения

Пример 1. Действительная и мнимая части эмпирической характеристической функции, построенной в точке $t \in \mathbf{R}$ по нормированной выборке Y_i, представляются в виде статистик вида (1), где роль функции $r(s)$ играют $\cos(ts)$ и $\sin(ts)$. Оба семейства функций, очевидно, удовлетворяют условиям (5) и (6).

Пример 2. В недавней статье [2] предложено семейство статистик типа супремума, основанных на эмпирическом аналоге интегрального преобразования Ханкеля. Для произвольного фиксированного положительного T статистика представляется в виде

$$\sup_{0<t<T}\left|\frac{1}{n}\sum_{i=1}^{n} r(Y_i,t)\right|,$$

где

$$r(s,t) = J_0(2\sqrt{st}) - \exp(-s),$$

$J_0(s)$ – функция Бесселя первого рода нулевого порядка. Она ограничена и имеет ограниченную производную, поэтому функции $\pm r(s,t)$ при каждом t удовлетворяют условиям теорем 2 и 3. Исследование этих статистик сейчас готовится к печати.

Пример 3. Наличие разрывов или точек, в которых производная бесконечна, само по себе может не являться препятствием для применения теорем настоящей и предыдущей статей. Например, функция $r_x(s) = \mathbb{I}_{\{s>x\}}$ имеет разрыв, но является неубывающей, поэтому к ней применимы результаты [6]. Однако уже к функции $r_{x,y}(s) = \mathbb{I}_{\{x<s<y\}}$ имеющиеся результаты неприменимы, поскольку она имеет

скачки как вверх, так и вниз и не сводится к монотонной прибавлением линейной функции. Другой пример функции, к которой имеющиеся результаты неприменимы, задает уравнение полуокружности: $r_{a,R}(s) = \mathbb{1}_{\{a-R<s<a+R\}}\sqrt{R^2-(s-a)^2}$.

Возможно, асимптотику больших уклонений статистик, построенных по подобным функциям, удастся изучить следующим образом. Приблизим определяющие их функции сверху и снизу поточечно последовательностями ограниченных функций, имеющих ограниченные производные. Для перехода к пределу нам потребуется знать функции уклонений в каждой точке, а для этого необходимы теоремы 2 и 3.

Автор благодарит Я. Ю. Никитина, А. М. Коточигова, А. Мироненко и Ю. Макарычева за неоценимую помощь в работе над статьей.

Литература

1. R. R. Bahadur, *Some Limit Theorem in Statistics.* Philadelphia: SIAM, 1971.
2. L. Baringhaus, F. Taherizade, *A K-S type test for exponentiality based on empirical Hankel transforms.* — Comm. Stat. Theor. Meth., **42** (2013), 3781–3792.
3. H. Chernoff, *A measure of asymptotic efficiency for tests of a hypothesis based on sums of observations.* — Ann. Math. Statist., **23**, No.4 (1952), 493–507.
4. Ya. Yu. Nikitin, *Asymptotic Efficiency of Nonparametric Tests.* Cambridge University Press, Cambridge 1995.
5. D. Plachky, J. Steinebach, *A theorem about probabilities of large deviations with an application to queuing theory.* — Periodica Mathematica Hungarica, **6**, No. 4 (1975), 343–345.
6. А. В. Чирина, *Большие уклонения некоторых свободных от параметра масштаба функций от выборки из гамма-распределения.* — Зап. научн. семин. ПОМИ, **298** (2003), 252–279.
7. Won Qyu Kim, *A non-compact generalization of Horvath's intersection theorem.* — Bull. Korean Math. Soc., **32**, No. 2 (1995), 153–162.

Tchirina A. V. Large deviations for sums of bounded functions of a normalized sample under gamma distribution.

We study large deviations of a widely used class of scale-free statistics under gamma distribution. We show that the constraints on the functions

defining these statistics can be relaxed with respect to the previously obtained result. The result is applied to a recent exponentiality test.

С.-Петербургский
государственный электротехнический университет
“ЛЭТИ”;
Национальный исследовательский университет
“Высшая школа экономики”
E-mail: antch@mail.ru, achirina@hse.ru

Поступило 5 ноября 2015 г.