Записки научных
семинаров ПОМИ
Том 466, 2017 г.

К. Ю. Волкова, М. С. Каракулов, Я. Ю. Никитин

КРИТЕРИИ СОГЛАСИЯ, ОСНОВАННЫЕ НА ХАРАКТЕРИЗАЦИИ РАВНОМЕРНОСТИ ОТНОШЕНИЕМ ПОРЯДКОВЫХ СТАТИСТИК, И ИХ АСИМПТОТИЧЕСКАЯ ЭФФЕКТИВНОСТЬ

§1. Введение

В настоящей работе мы построим два новых критерия равномерности, основанных на ее характеризации свойствами порядковых статистик, и исследуем их асимптотические свойства. Такие критерии строились изредка и раньше. Один из первых критериев равномерности, основанный на характеризационных свойствах, построен в [7] японскими статистиками Хашимото и Ширахаши. Недавно построены критерии равномерности, основанные на характеризации свойствами ожиданий порядковых статистик, принадлежащей Ту и Лину [12].

Критерии, основанные на характеризациях, и их эффективности в настоящее время сравнительно мало изучены, и в этой области существует лишь небольшое количество публикаций, см., например, работы [5,8,10,13,14,18,22–24], а также обзор [17].

Общая постановка задачи проверки равномерности выглядит так: пусть $X_1, X_2, \ldots, X_n$ – независимые одинаково распределенные случайные величины (н.о.р.с.в.), имеющие непрерывную функцию распределения (ф.р.) F. Требуется проверить гипотезу равномерности $H_0 : F(x) = x$, $x \in [0,1]$.

В 1989 г. известный статистик М. Ахсануллах опубликовал интересную характеризацию равномерного распределения [1, Лемма 1].

Лемма 1. *Пусть* $X_1, \ldots, X_n$*,* $n \geqslant 2$*, – независимые наблюдения с абсолютно непрерывным распределением, сосредоточенным на* $[0,1]$*, и пусть* $X_{1,n} \leqslant X_{2,n} \leqslant \ldots \leqslant X_{n,n}$ *– соответствующий вариационный*

Ключевые слова: равномерное распределение, проверка гипотез, характеризация, U-статистики, большие уклонения, бахадуровская эффективность, локальная оптимальность.

Работа первого и третьего авторов поддержана грантом РФФИ 16-01-00258 и грантом СПбГУ-DFG 6.65.37.2017.

ряд. Выборка имеет равномерное распределение $U[0,1]$ *тогда и только тогда, когда дробь*

$$X_{1,n}/X_{2,n} \sim U[0,1].$$

Воспользуемся этим результатом, чтобы построить два критерия для проверки гипотезы о принадлежности выборки равномерному распределению на $[0,1]$. Для этого построим по исходной выборке так называемую U-эмпирическую функцию распределения

$$G_n(t) = \binom{n}{k}^{-1} \sum_{1\leqslant i_1<\ldots<i_k\leqslant n} \mathbf{1}\{X_{1,\{i_1,\ldots,i_k\}}/X_{2,\{i_1,\ldots,i_k\}} < t\},\ t\in[0,1],$$

где под выражением $X_{s,\{i_1,\ldots,i_k\}}$, $s=1,2$, понимается s-я порядковая статистика выборки $X_{i_1},\ldots,X_{i_k}$.

В соответствии с теоремой Гливенко–Кантелли для U-эмпирических распределений [11] справедливо соотношение: почти наверное

$$\sup_{0\leqslant t\leqslant 1} |G_n(t)-t| \longrightarrow 0 \quad \text{при} \quad n\to\infty.$$

Следовательно, проверку равномерности естественно проводить с помощью следующих двух статистик: интегральной статистики

$$\begin{aligned} IU_n &:= \int_0^1 (t-G_n(t))\,dt \\ &= \binom{n}{k}^{-1} \sum_{1\leqslant i_1<\ldots<i_k\leqslant n} \Big[X_{1,\{i_1,\ldots,i_k\}}/X_{2,\{i_1,\ldots,i_k\}} - \frac{1}{2}\Big] \qquad (1)\end{aligned}$$

и статистики типа Колмогорова

$$DU_n := \sup_{0\leqslant t\leqslant 1} |G_n(t)-t|. \qquad (2)$$

Мы видим, что статистика IU_n является обычной U-статистикой с ядром степени k, тогда как DU_n устроена более сложно и является супремумом целого семейства модулей U-статистик, зависящих от t. Мы обсудим предельные распределения обеих статистик, чтобы использовать их для проверки гипотез в больших выборках. При этом будем пользоваться необходимыми сведениями из теории U-статистик и теории асимптотической эффективности по Бахадуру, изложенными в монографиях [2, 6, 15].

Далее, для выбора наиболее чувствительного критерия и для практических рекомендаций в настоящем исследовании используется асимптотическое сравнение критериев на основе понятия асимптотической относительной эффективности (АОЭ) [4]. Поскольку статистика DU_n имеет предельное распределение, явно отличное от нормального, то питменовская эффективность неприменима, а эффективность по Ходжесу–Леману слишком груба и не различает двусторонние статистики [15]. Поэтому наиболее подходящим средством сравнения статистик представляется локальная асимптотическая эффективность по Бахадуру [2, 4], которая совпадает в большинстве случаев с модифицированными вариантами других эффективностей [3, 15].

§2. Эффективность по Бахадуру

Локальная асимптотическая эффективность (ЛАЭ) по Бахадуру одной последовательности статистик относительно другой вычисляется как предел отношения точных наклонов этих последовательностей статистик при стремлении альтернативы к нулевой гипотезе [3].

Точный наклон характеризует скорость экспоненциального убывания ошибки первого рода при фиксированных мощности и альтернативе и является общепринятой асимптотической характеристикой качества критерия. Для вычисления точных наклонов будем использовать фундаментальную теорему Бахадура [2, 15]:

Теорема 1. *Пусть $\Theta = \Theta_0 \cup \Theta_1$ – параметрическое множество, где значения параметра $\theta \in \Theta_0$ соответствуют гипотезе H_0, а значения $\theta \in \Theta_1$ – альтернативе. Пусть последовательность статистик $\{T_n\}$ удовлетворяет следующим условиям:*

(1) *$T_n \to b(\theta)$ по $\mathbf{P}_\theta$-вероятности, $\theta \in \Theta_1$, где $-\infty < b(\theta) < \infty$, и*

(2) *$\lim\limits_{n\to\infty} n^{-1} \ln \mathbf{P}_\theta(T_n > a) = -k(a)$ для любого $\theta \in \Theta_0$ и некоторого открытого интервала I, где функция k непрерывна на множестве $I = \{b(\theta), \theta \in \Theta_1\}$.*

Тогда при всех $\theta \in \Theta_1$ точный наклон $c_T(\theta)$ существует и вычисляется по формуле:

$$c_T(\theta) = 2\,k(b(\theta)).$$

Под *локальным* точным наклоном понимается главная часть точного наклона $c_T(\theta)$ при $\theta \to 0$. Обычно имеет место соотношение

$$c_T(\theta) \sim l_T\,\theta^2, \quad \theta \to 0,$$

где величину l_T принято называть локальным индексом. Последний, конечно, зависит от альтернативы.

В качестве альтернатив к равномерному распределению на [0,1] мы выберем альтернативы, похожие на те, что использовались недавно Б. Милошевич [12] для исследования другого критерия равномерности, основанного на свойствах ожиданий порядковых статистик. Опишем эти альтернативы и их плотности $f_j(x,\theta)$, $x \in [0,1]$, $j = 1, \dots, 5$:

(1) Лемановская: $f_1(x,\theta) = (\theta+1)\,x^\theta, \quad \theta > 0;$

(2) Линейная: $f_2(x,\theta) = 1 + \theta\,(2x-1), \quad \theta > 0;$

(3) Косинус-альтернатива: $f_3(x,\theta) = 1 - \theta\,\cos\pi x, \quad \theta \in [0, 1/2);$

(4) Альтернатива, основанная на проекции ядра,
$f_4(x,\theta) = 1 + \theta\,(\frac{x}{2} - x\ln x - \frac{1}{2}), \quad \theta > 0;$

(5) Параболическая: $f_5(x,\theta) = 1 - \theta\,(x-1)(x-\frac{1}{3}), \quad \theta > 0.$

Мы будем также пользоваться сведениями о фишеровской информации для этих альтернатив, которые удобно представить в виде таблицы 1. В этой таблице под $I(f_j)$, $j = 1, \dots, 5$, понимается фишеровская информация при $\theta = 0$ для каждой из альтернативных плотностей.

В качестве пояснения напомним, что в проверке гипотез важную роль играет расстояние Кульбака–Лейблера [2,15] между альтернативой с плотностью $f(x,\theta)$ и (в данном контексте) равномерным распределением, а именно:

$$Q(\theta) = \int_0^1 f(x,\theta)\,\ln f(x,\theta)\,dx.$$

Верхняя граница для локального индекса l_T дается, как известно [3, 15], величиной $\lim_{\theta\to 0} 2\,Q(\theta)/\theta^2$. При естественных условиях регулярности [2], которые выполняются для всех пяти альтернативных плотностей, этот предел равен

$$I(f_j) = \int_0^1 \left(\frac{\partial f_j(x,\theta)}{\partial\theta}\Big|_{\theta=0} \right)^2\,dx. \tag{3}$$

Таблица 1. Информация Фишера в нуле для альтернативных плотностей.

Альтернатива	f_1	f_2	f_3	f_4	f_5
$I(f_j)$	1	1/3	1/2	1/54	2/135

Зная эти значения и зная локальные индексы, мы легко можем вычислить локальную бахадуровскую эффективность $e_T^B(g)$ при альтернативной плотности g по формуле

$$e_T^B(g) = l_T/I(g). \tag{4}$$

§3. Интегральная статистика и ее свойства

Вернемся к изучению статистики IU_n из (1). Это представитель классических U-статистик, свойства которых хорошо изучены, см. [11]. Ядро этой статистики степени k имеет вид

$$\Phi(X_1, \dots, X_k) = X_{1,\{1,\dots,k\}}/X_{2,\{1,\dots,k\}} - \frac{1}{2}.$$

Вычислим первую проекцию ядра. Зафиксируем $X_k = x$. Получаем

$$\begin{aligned}\varphi(x) &:= \mathbf{E}(\Phi(X_1, \dots, X_k) \mid X_k = x)\\ &= \mathbf{E}(X_{1,\{1,\dots,k-1,x\}}/X_{2,\{1,\dots,k-1,x\}}) - \frac{1}{2} := E_1 + E_2 + \sum_{l=3}^{k-1} E_l - \frac{1}{2},\end{aligned}$$

где слагаемые определяются положением x в вариационном ряду.

Вычислим отдельно каждое из слагаемых:

(1) Пусть элементы выборки располагаются так: $x < X_1 < \dots < X_{k-1}$. Тогда искомое выражение преобразуется и после некоторых вычислений равно

$$E_1 := (k-1)!\mathbf{E}\Big(\frac{x}{X_1}\mathbf{1}\{x < X_1 < \dots < X_{k-1}\}\Big) = (k-1)x\int_x^1 \frac{(1-x)^{k-2}}{x}\, dx.$$

(2) Если же элементы выборки будут такими: $X_1 < x < X_2 < \dots < X_{k-1}$, тогда имеем

$$E_2 := (k-1)!\,\mathbf{E}\Big(\frac{X_1}{x}\mathbf{1}\{X_1 < x < X_2 < \dots < X_{k-1}\}\Big) = \frac{(k-1)}{2}\, x\,(1-x)^{k-2}.$$

(3) В последнем случае мы рассмотрим все такие варианты, когда x находится на l-том месте в вариационном ряду, $l \geqslant 3 : X_1 < X_2 < \cdots < x < \cdots < X_{k-1}$.

Тогда искомая вероятность для этого случая равна:

$$E_l := \frac{1}{2}\, C_{k-1}^{l-1}\, x^{l-1}\,(1-x)^{k-l}.$$

После объединения всех случаев получаем выражение для проекции

$$\varphi(x) = (k-1)\,x \int\limits_x^1 \frac{(1-x)^{k-2}}{x}\,dx + \frac{1}{2}\sum_{l=2}^{k-1} C_{k-1}^{l-1}\, x^{l-1}\,(1-x)^{k-l} - \frac{1}{2}.$$

Сделаем еще некоторое упрощение, использовав разложение бинома Ньютона, и окончательно получим выражение для проекции:

$$\varphi(x) = (k-1)\,x \int\limits_x^1 \frac{(1-x)^{k-2}}{x}\,dx - \frac{1}{2}\,(1-x)^{k-1}. \tag{5}$$

Нетрудно убедиться, что $\mathbf{E}\,\varphi(x) = 0$, $0 \leqslant x \leqslant 1$.

Важную роль в дальнейшем играет дисперсия этой проекции $\sigma_\varphi^2 = \int\limits_0^1 \varphi^2(x)\,dx$. Эту функцию не удается вычислить аналитически для произвольного k, но она легко вычисляется для конкретных k. Позже мы произведем вычисления в качестве примера для случаев $k = 2$ и $k = 3$.

Из (5) следует, что наша U-статистика и ее ядро степени k оказываются невырождены. Следовательно, статистика IU_n асимптотически нормальна, и по теореме Хёфдинга [11] получаем сходимость по распределению

$$\sqrt{\frac{n}{k^2\sigma_\varphi^2}}\,IU_n \longrightarrow N(0,1), \quad n \to \infty.$$

Заодно мы сформулируем асимптотику больших уклонений нашей статистики при H_0, имея в виду применение к вычислению локальной бахадуровской эффективности по теореме 1. Соответствующий результат мы заимствуем из работы [20], см. также [6, 16]. Он применим, так как наша U-статистика невырождена и имеет ограниченное ядро. Получаем

Теорема 2. *При* $a > 0$

$$\lim_{n\to\infty} n^{-1}\ln \mathbf{P}(IU_n > a) = -r(a),$$

где функция r непрерывна при достаточно малых $a > 0$, причем

$$r(a) \sim \frac{a^2}{2\,k^2\sigma_\varphi^2} \quad \text{при} \ \ a \to 0.$$

Асимптотика $b_{IU}(\theta)$ уже вычислена для широкого класса невырожденных U-статистик в работе [19]. Сформулированные там условия регулярности ND на альтернативу выполнены ввиду ограниченности ядра и гладкой зависимости альтернативных плотностей от параметра θ. В силу этого имеем

$$b_{IU}(f,\theta) \sim k \int_0^1 \varphi(x)\, f_\theta(x,0)\, dx\, \theta, \tag{6}$$

где под $f_\theta(x,0)$ мы понимаем альтернативную плотность, вычисленную в точке $\theta = 0$. Комбинируя последние две формулы, мы получаем окончательно выражение для локального индекса

$$l_{IU} = \frac{1}{\sigma_\varphi^2} \left(\int_0^1 \varphi(x)\, f_\theta(x,0)\, dx \right)^2 .$$

Отметим важное обстоятельство: вычисления показывают, что каждый интеграл, стоящий под скобками и дающий выражение для $b_{IU}(\theta)$ в (6), *положителен*. Это означает, что наш интегральный критерий состоятелен против всех пяти рассматриваемых альтернатив.

Теперь получаем соответствующие значения локальных бахадуровских эффективностей по формуле (4).

§4. Критерий типа Колмогорова

Вернемся к рассмотрению критерия типа Колмогорова, основанного на статистике DU_n из (2), которая является супремумом модуля *семейства* U-статистик, зависящим от $t \in [0,1]$. Ядро нашего семейства имеет вид

$$\Xi(X_1,\dots,X_k;t) = \mathbf{1}\{X_{1,\{1,\dots,k\}}/X_{2,\{1,\dots,k\}} < t\} - t, \quad 0 \leqslant t \leqslant 1.$$

Вычислим проекцию этого ядра при фиксированном t. Ясно, что

$$\begin{aligned}\xi(x;t) &:= \mathbf{E}(\Xi(X_1,\dots,X_k) \mid X_k = x)\\ &= \mathbf{P}\{X_{1,\{1,\dots,k-1,x\}}/X_{2,\{1,\dots,k-1,x\}} < t\} - t, \quad 0 \leqslant t \leqslant 1.\end{aligned}$$

Как и для интегральной статистики, разберем отдельно три случая расположения элементов выборки:

(1) $x < X_1 < \cdots < X_{k-1}$;

(2) $X_1 < x < X_2 < \cdots < X_{k-1}$;

(3) x находится на l-том месте в вариационном ряду, $l \geqslant 3 : X_1 < X_2 < \cdots < x < \cdots < X_{k-1}$.

Опуская проделанные вычисления, получаем окончательное выражение для проекции ядра $\Xi(X_1, \ldots, X_k; t)$:

$$\xi(x;t) = \mathbf{1}\{x < t\} \left(1 - \frac{x}{t}\right)^{k-1} - t\,(1-x)^{k-1}. \tag{7}$$

Тогда функция дисперсии $\sigma_\xi^2(t)$ для каждого конкретного k легко вычисляется. Однако в дальнейшем нам потребуется для описания больших уклонений рассматриваемой статистики значение $\sup\limits_{0\leqslant t\leqslant 1} \sigma_\xi^2(t)$.

Предельное распределение статистики DU_n неизвестно. Опираясь на методы работы [25], можно показать, что U-эмпирический процесс

$$\zeta_n(t) = \sqrt{n}\,(t - G_n(t))\,, \quad t \in (0,1),$$

слабо сходится при $n \to \infty$ к некоторому гауссовскому процессу $\zeta(t)$ с нулевым средним и сложной ковариацией. Тогда последовательность статистик $\sqrt{n}\,DU_n$ сходится по распределению к величине $\sup_t |\zeta(t)|$, найти распределение которой весьма затруднительно. Поэтому критические значения для статистик DU_n естественно искать с помощью моделирования их выборочного распределения.

Семейство ядер $\Xi(X_1, \ldots, X_k; t)$, $t \in (0,1)$, не только центрировано, но и ограничено. Применяя теорему [16] о логарифмической асимптотике вероятностей больших уклонений для U-эмпирических статистик Колмогорова, получаем следующий результат.

Теорема 3. *Для достаточно малых $a > 0$ выполнено*

$$\lim_{n\to\infty} n^{-1} \ln \mathbf{P}(DU_n > a) = -m(a),$$

где функция m непрерывна, причем

$$m(a) = \frac{a^2}{2\,k^2 \sup\limits_{0\leqslant t\leqslant 1} \sigma_\xi^2(t)}(1 + o(1)) \quad \text{\textit{при} } a \to 0.$$

При вычислении локальных бахадуровских наклонов, снова пользуясь теоремой Гливенко–Кантелли для U-эмпирических ф.р. [9], можно показать, что выполняется следующее соотношение

$$b_{DU}(f,\theta) \sim k \sup_{0\leqslant t\leqslant 1}\left|\int_0^1 \xi(x;t)\, f_\theta(x,0)\,dx\right| \theta, \tag{8}$$

где под $f_\theta(x,\theta)$ мы понимаем производную альтернативной плотности, вычисленную при $\theta = 0$.

Вопрос о состоятельности критерия типа Колмогорова здесь не возникает, поскольку мы всегда вычисляем *абсолютное значение* отклонения предела при альтернативе U-эмпирической функции распределения от равномерной функции распределения.

Располагая сведениями о величинах $b_{DU}(f_j,\theta)$ при различных альтернативах, нам остается вычислить локальные точные наклоны по формуле

$$c_{DU}(f_j,\theta) = \frac{b^2_{DU}(f_j,\theta)}{k^2 \sup\limits_{0\leqslant t\leqslant 1} \sigma^2_\xi(t)}, \quad j = 1,\ldots,5.$$

Учитывая фишеровские информации, собранные в таблице 1, мы легко получим теперь локальные абсолютные эффективности критерия Колмогорова для пяти рассматриваемых плотностей.

§5. Таблицы локальных эффективностей и их обсуждение

В этом разделе соберем полученные нами результаты для локальных бахадуровских эффективностей в таблицу 2.

Таблица 2. Сводная таблица локальных эффективностей для IU_n, DU_n

Альтернатива	$IU_n, k=2$	$IU_n, k=3$	$DU_n, k=2$	$DU_n, k=3$
f_1	0.844	0.682	0.685	0.568
f_2	0.501	0.227	0.422	0.197
f_3	0.466	0.177	0.419	0.163
f_4	1	0.859	0.928	0.772
f_5	0.957	0.688	0.877	0.617

Мы видим, что по крайней мере для рассмотренных пяти альтернатив к равномерности, интегральный критерий обычно превосходит по локальной бахадуровской эффективности своего оппонента – критерий типа Колмогорова. Это неудивительно, поскольку и при проверке других гипотез интегральный критерий обычно лучше супремального [15].

Для альтернативы, определяемой проекцией ядра, интегральный критерий оказывается локально оптимальным. Это объясняется тем, что именно такие альтернативы являются наиболее благоприятными [19]. Стоит отметить высокую эффективность новых критериев для альтернатив f_1 и f_5.

Возможно, два новых критерия стоит включать в "батарею" критериев для проверки равномерности, когда требуется высокая уверенность статистика в справедливости этой гипотезы.

§6. Условия локальной асимптотической оптимальности

Выясним условия локальной асимптотической оптимальности для интегральной и супремальной статистик, то есть опишем семейства альтернатив, для которых изучаемая статистика является локально наилучшей в бахадуровском смысле.

Подобные задачи уже рассматривались при изучении классических непараметрических критериев в [15, гл. 6], где использовались вариационные методы.

Пусть $f(x,\theta)$ – альтернативная плотность в точке θ. Для удобства обозначим $h(x) = \frac{\partial}{\partial\theta} f(x,\theta)\mid_{\theta=0}$. Предположим также, что выполнено следующее условие регулярности:

$$\frac{\partial}{\partial\theta}\int\limits_0^\infty x\, f(x,\theta)\,dx\Big|_{\theta=0} = \int\limits_0^\infty x\, h(x)\,dx. \tag{9}$$

Обозначим через $\mathcal{F}$ класс плотностей $f(x,\theta)$ с ф.р. $F(x,\theta)$, удовлетворяющих условию регулярности (9). Выведем условия оптимальности в терминах функции $h(x)$.

Как уже было замечено в (3), для альтернативных плотностей из класса $\mathcal{F}$ выполнено следующее соотношение: $I(f) = \int\limits_0^1 h^2(x)\,dx$.

Сначала рассмотрим интегральную статистику IU_n с проекциями ядра $\varphi(x)$ из (5). Напомним, что в (6) выполнена следующая асимптотика для функции $b_{IU}(\theta)$:

$$b_{IU}(\theta) \sim k\,\theta \int\limits_0^1 \varphi(x)\,h(x)\,dx.$$

Поэтому, подставляя полученное в формулу (4), находим, что локальная асимптотическая эффективность принимает вид

$$e^B_{IU}(f_j) = \Big(\int\limits_0^1 \varphi(x)h(x)\,dx\Big)^2 \Big/ \Big(\int\limits_0^1 \varphi^2(x)\,dx \cdot \int\limits_0^1 h^2(x)\,dx\Big).$$

Локальная асимптотическая оптимальность по Бахадуру статистики IU_n означает, что выражение в правой части равно 1. Из неравенства Коши–Буняковского–Шварца, см. также [19,21], следует, что это выполняется тогда и только тогда, когда $h(x) = C_1\varphi(x)$ для некоторой константы $C_1 \neq 0$. При этом состоятельность при малых $\theta > 0$ будет обеспечиваться при $C_1 > 0$. Распределения, для которых функция $h(x)$ указанного вида, образуют область локальной асимптотической оптимальности (ЛАО) в классе $\mathcal{F}$.

Для случая $k = 2$ примером такой альтернативы может служить уже рассмотренная нами альтернатива с плотностью $f_4(x,\theta)$:

$$f_4(x,\theta) = 1 + \theta\left(\frac{x}{2} - x\,\ln x - \frac{1}{2}\right), \quad \theta > 0.$$

Именно по этой причине эффективность критерия IU_n для этой альтернативы получилась равная 1.

Для случая $k = 3$ одним их примеров наилучшей альтернативы будет плотность:

$$f(x,\theta) = 1 + \theta\left(\frac{3}{2}\,x^2 - 2\,x\,\ln x - x - \frac{1}{2}\right), \quad \theta > 0.$$

Теперь рассмотрим статистику колмогоровского типа DU_n с семействами ядер $\Xi(X_1,\ldots,X_k;t)$ и проекциями $\xi(x;t)$ из (7). В этом случае по формуле (8) верна асимптотика для функции $b_{DU}(\theta)$:

$$b_{DU}(\theta) \sim \theta \sup_{0\leqslant t\leqslant 1}\Big|\,k\int\limits_0^1 \xi(x;t)\,h(x)\,dx\,\Big|.$$

Таблица 3. Примеры плотностей ЛАО альтернатив $f_\theta(x,\theta)$ для статистики DU_n.

	Плотности $f_\theta(x,\theta)$ при $\theta \to 0$, $x \in [0,1]$
$k=2$	$f_\theta(x,\theta) = 1 + \theta\left(\mathbf{1}\{x < t_1\}(1 - \frac{x}{t_1}) - t_1(1-x)\right)$
	$t_1 = \arg\max_{0\leqslant t\leqslant 1}\left(\frac{1}{3}t(1-t)^2\right) = 4/81 \approx 0.0494$
$k=3$	$f_\theta(x,\theta) = 1 + \theta\left(\mathbf{1}\{x < t_2\}(1 - \frac{x}{t_2})^2 - t_2(1-x)^2\right)$
	$t_2 = \arg\max_{0\leqslant t\leqslant 1}\left(\frac{t}{15}t(1-t)^2(3-t)\right) \approx 0.0265$

Поэтому локальная асимптотическая эффективность принимает вид

$$e^B_{DU} = \sup_{0\leqslant t\leqslant 1}\left(\int_0^1 \xi(x;t)\,h(x)\,dx\right)^2 \Big/ \sup_{0\leqslant t\leqslant 1}\left(\int_0^1 \xi^2(x;t)\,dx \cdot \int_0^1 h^2(x)\,dx\right).$$

Мы снова применим неравенство Коши–Буняковского–Шварца ко всей дроби. Тогда получим, что последовательность статистик DU_n является локально асимптотически оптимальной и $e^B_{DU} = 1$ тогда и только тогда, когда

$$h(x) = C_2\,\xi(x;t_0) \qquad \text{для} \quad t_0 = \arg\sup_{0\leqslant t\leqslant 1}\sigma^2_\xi(t)$$

и для некоторой константы $C_2 > 0$. Распределения, для которых функция $h(x)$ указанного вида, образуют область ЛАО в соответствующем $\mathcal{G}$.

Простейшие примеры таких альтернативных плотностей для $k=2$ и $k=3$ указаны в таблице 3.

Литература

1. M. Ahsanullah, *On characterizations of the uniform distribution based on functions of order statistics.* — Aligarh J. Statist. **9** (1989), 1–6.
2. R. R. Bahadur, *Some Limit Theorems in Statistics.* SIAM, Philadelphia, 1971.
3. R. R. Bahadur, *Stochastic comparison of tests.* — Ann. Math. Statist. **31** (1960), 276–295.
4. R. R. Bahadur, *Rates of convergence of estimates and test statistics.* — Ann. Math. Statist. **38** (1967), 303–324.
5. L. Baringhaus, N. Henze, *Tests of fit for exponentiality based on a characterization via the mean residual life function.* — Statist. Papers **41** (2000), 225–236.
6. A. DasGupta, *Asymptotic Theory of Statistics and Probability.* Springer, New York, 2008.

7. T. Hashimoto, S. Shirahata, *A goodness-of-fit test based on a characterization of uniform distribution.* — J. Japan Statist. Soc. **23**, No. 2 (1993), 123–130.
8. N. Henze, S. Meintanis, *Goodness-of-fit tests based on a new characterization of the exponential distribution.* — Commun. Statist. Theor. Meth. **31** (2002), 1479–1497.
9. R. Helmers, P. Janssen, R. Serfling, *Glivenko–Cantelli properties of some generalized empirical DF's and strong convergence of generalized L-statistics.* — Probab. Theor. Relat. Fields, **79** (1988), 75–93.
10. M. Jovanović, B. Milošević, Ya. Yu. Nikitin, M. Obradović, K. Yu. Volkova, *Tests of exponentiality based on Arnold-Villasenor characterization and their efficiencies.* — Computat. Statist. Data Anal. **90** (2015), 100–113.
11. V. S. Korolyuk, Y. V. Borovskich, *Theory of U-statistics.* 2nd ed., Springer Science and Business Media, 2013.
12. B. Milošević, *Asymptotic Efficiency of Goodness-of-fit Tests Based on Too-Lin Characterization.* ArXiv:1508.05314, 2015.
13. B. Milošević, *Asymptotic efficiency of new exponentiality tests based on a characterization.* — Metrika, **79** (2016), 221–236.
14. K. Morris, D. Szynal, *Goodness-of-fit tests using characterizations of continuous distributions.* — Appl. Math. (Warsaw) **28** (2001), 151–168.
15. Y. Nikitin, *Asymptotic Efficiency of Nonparametric Tests.* Cambridge University Press, New York, 1995.
16. Ya. Yu. Nikitin, *Large deviations of U-empirical Kolmogorov-Smirnov tests, and their efficiency.* — J. Nonpar. Statist. **22**(2010), 649–668.
17. Ya. Yu. Nikitin, *Tests based on characterizations, and their efficiencies.* — Acta et Comment. Univ. Tartuens. Matem. **21**, No. 1 (2017), 3–24.
18. P. Muliere, Y. Nikitin, *Scale-invariant test of normality based on Polya's characterization.* — Metron, **60**, No. 1-2 (2002), 21–33.
19. Ya. Yu. Nikitin, I. Peaucelle, *Efficiency and local optimality of distribution-free tests based on U and V-statistics.* — Metron, **LXII** (2004), 185–200.
20. Ya. Yu. Nikitin, E. V. Ponikarov, *Rough large deviation asymptotics of Chernoff type for von Mises functionals and U-statistics.* — Proc. of St. Petersburg Math. Soc. **7** (1999), 124–167. Engl. transl. in AMS Transl., ser. 2, **203** (2001), 107–146.
21. Ya. Yu. Nikitin, A. V. Tchirina, *Lilliefors test for exponentiality: large deviations, asymptotic efficiency, and conditions of local optimality.* — Math. Meth. Statist. **16**, No. 1 (2007), 16–24.
22. Ya. Yu. Nikitin, K. Yu. Volkova, *Asymptotic efficiency of exponentiality tests based on order statistics characterization.* — Georgian Math. J. **17** (2010), 749–763.
23. M. Obradović, *Three characterizations of exponential distribution involving the median of sample of size three.* — J. Statist. Theory Appl. **14** (2015), 257-264.
24. M. Obradovic, M. Jovanovic, B. Milosevic, *Goodness-of-fit tests for Pareto distribution based on a characterization and their asymptotics.* — Statistics **49** (2015), 1026–1041.
25. B. W. Silverman, *Convergence of a class of empirical distribution functions of dependent random variables.* — Ann. Probab. **11** (1983), 745–751.

Volkova K. Yu., Karakulov M. S., Nikitin Ya. Yu. Goodness-of-fit tests based on the characterization of uniformity by the ratio of order statistics, and their efficiencies.

We construct integral and supremum type goodness-of-fit tests for the uniform law based on Ahsanullah's characterization of uniform law. We discuss limiting distributions of new tests and describe the logarithmic large deviation asymptotics of test statistics under null-hypothesis. This enables to calculate their local Bahadur efficiency under some parametric alternatives. Conditions of local optimality of new statistics are given.

С.-Петербургский государственный университет,
Университетская наб. 7/9,
С. Петербург 199034, Россия
E-mail: k.volkova@spbu.ru
E-mail: karakulovms@gmail.com
E-mail: y.nikitin@spbu.ru

Поступило 6 октября 2017 г.