Записки научных
семинаров ПОМИ
Том 294, 2002 г.

А. В. Чирина

АСИМПТОТИЧЕСКАЯ ЭФФЕКТИВНОСТЬ И ЛОКАЛЬНАЯ ОПТИМАЛЬНОСТЬ ПО БАХАДУРУ КРИТЕРИЯ ЭКСПОНЕНЦИАЛЬНОСТИ, ОСНОВАННОГО НА СТАТИСТИКЕ МОРАНА

1. Введение

Пусть $X_1,\ldots,X_n$ – неотрицательные независимые одинаково распределенные случайные величины (н. о. р. с. в.), $\overline{X}$ – их выборочное среднее. Пусть требуется проверить гипотезу H_0 : X_i экспоненциально распределены, т.е. имеют функцию распределения $F(x) = 1 - e^{-\lambda x}$, $x \geq 0$ с некоторым положительным λ, против альтернативы H_1: X_i имеют неэкспоненциальное распределение, сосредоточенное на положительной полуоси. Для проверки H_0 против H_1 предлагается статистика

$$T_n = \left| \mathbf{C} + \frac{1}{n} \sum_{i=1}^{n} \ln \frac{X_i}{\overline{X}} \right| ,$$

а также ее односторонние варианты

$$T_n^+ = \mathbf{C} + \frac{1}{n} \sum_{i=1}^{n} \ln \frac{X_i}{\overline{X}},$$

и

$$T_n^- = -(\mathbf{C} + \frac{1}{n} \sum_{i=1}^{n} \ln \frac{X_i}{\overline{X}}),$$

где $\mathbf{C}$ –постоянная Эйлера.

Очевидно, распределения статистик T_n, T_n^+, T_n^- при H_0 не зависят от параметра λ. Эти статистики обычно называют статистиками Морана, так как критерий, основанный на статистике T_n^+, построен в [7] как локально асимптотически наиболее мощный для проверки экспоненциальности против гамма-альтернативы с положительным параметром. В [6] доказано, что

Работа поддержана грантом РФФИ 00-15-96019.

этот критерий является равномерно наиболее мощным свободным от параметра масштаба , а также равномерно наиболее мощным несмещенным критерием для проверки экспоненциальности против гамма-альтернативы.

Мы предлагаем применять T_n и ее односторонние варианты для проверки H_0 против более широкого класса альтернатив. В частности, T_n^+ состоятельна в классе альтернатив с возрастающей функцией интенсивности отказов (ВФИ), T_n^- – в классе альтернатив с убывающей функцией интенсивности отказов (УФИ), и T_n – в классе альтернатив с монотонной функцией интенсивности отказов. Асимптотическая нормальность и состоятельность изучаемых статистик обсуждаются в разделе 2.

Целью настоящей статьи является вычисление локальной асимптотической эффективности (ЛАЭ) по Бахадуру статистик T_n, T_n^+, T_n^- и выяснение условий их локальной асимптотической оптимальности.

В разделе 3 найдена грубая асимптотика вероятностей больших уклонений T_n , T_n^+ , T_n^- при H_0, необходимая для вычисления точных наклонов. Полученный результат описывает также асимптотическое поведение статистики Дарлинга для проверки равномерности(см. [5]) при нулевой гипотезе. В разделе 4 выведены общие формулы для главных частей точных наклонов и для ЛАЭ по Бахадуру (а для гамма-альтернатив – точные значения точного наклона и асимптотической эффективности). Приведены численные результаты для следующих альтернативных плотностей: плотность с линейной интенсивностью отказов, плотность Макегама, плотность Вейбулла, гамма-плотность. В разделе 5 найдены условия локальной асимптотической оптимальности.

2. Асимптотическая нормальность и состоятельность

Введем обозначения:

$$m = EX_1,\ l = E\ln X_1,\ \sigma^2 = D\left(\frac{X_1}{m} - \ln X_1\right).$$

Теорема 1. *Пусть* m, $|l|$, $\sigma^2 < +\infty$, *и, кроме того,* $\sigma^2 > 0$ *и* $0 < DX_1 < +\infty$.

Тогда справедливы следующие утверждения.

1. *Последовательность случайных величин* $\sqrt{n}(T_n^+ - \mathbf{C} - l + \ln m)$ *сходится по распределению к* $\mathcal{N}(0,\, \sigma^2)$.

В частности, если X_1 *имеет экспоненциальное распределение, то*

$$\sqrt{n}\; T_n^+ \;\xrightarrow{d} \mathcal{N}\left(\, 0,\; \frac{\pi^2}{6} - 1\right).$$

2. *Если при некоторой альтернативе* $\mathbf{C} + l - \ln m \neq 0$ $(> 0, < 0)$, *то статистика, отвергающая* H_0 *при больших положительных значениях* $\sqrt{n}\; T_n$ *(соответственно* $\sqrt{n}\; T_n^+$, $\sqrt{n}\; T_n^-$*), состоятельна для этой альтернативы.*

3. *Пусть при альтернативе* H_1 *случайная величина* X_1 *имеет монотонную функцию интенсивности отказов (ВФИ, УФИ). Тогда статистика* $\sqrt{n}\, T_n$ *(соответственно* $\sqrt{n}\; T_n^+$, $\sqrt{n}\; T_n^-$*) состоятельна для проверки* H_0 *против* H_1.

Доказательство.

1. Заметим, что

$$\begin{aligned} T_n^+ - \mathbf{C} - l + \ln m \\ &= \frac{1}{n}\sum_{i=1}^{n}\left(\ln\frac{X_i}{m} - l - \frac{X_i}{m} + 1\right) + \left(\frac{\overline{X}}{m} - 1 - \ln\frac{\overline{X}}{m}\right) \\ &=: S_{n1} + S_{n2}. \end{aligned} \tag{2.1}$$

По ЦПТ получаем, что

$$\sqrt{n} S_{n1} \xrightarrow{d} \mathcal{N}(0,\, \sigma^2) \tag{2.2}$$

и

$$\sqrt{n}(\overline{X}/m\; - 1) \xrightarrow{d} \mathcal{N}\;(0,\;\; m^{-2}\, DX_1) \tag{2.3}$$

Из (2.3) легко получить, что $\sqrt{n} S_{n2} \to 0$ по вероятности, а это в сочетании с (2.2) и доказывает п. 1 Теоремы 1.

2. Очевидно, что из п. 1 вытекает п. 2.

3. Пусть $G(x) = \mathbf{P}(X_1 < x)$, $\overline{G}(x) = 1 - G(x)$. В условиях Теоремы 1

$$
\begin{aligned}
\mathbf{C}+l-\ln m &= \int \ln\frac{x}{m}dG(x)-\int \ln\frac{x}{m}d(1-e^{-x/m}) \\
&= (G(x)-1+e^{-x/m})\ln\frac{x}{m}\Big|_0^{+\infty} - \int \frac{1}{x}(G(x)-(1-e^{-x/m}))dx \\
&= \int \frac{\overline{G}(x)-e^{-x/m}}{x}dx
\end{aligned}
$$

(Здесь и на протяжении всей статьи все интегралы берутся по положительной полуоси, поэтому мы опускаем пределы интегрирования.)

Так как $1/x$ убывает на $\mathbb{R}^+$, имеем, рассуждая так же, как при доказательстве Теоремы 4.8 Гл. 2 в [1]:

$\mathbf{C}+l-\ln m \geq 0$ для ВФИ-распределений и $\mathbf{C}+l-\ln m \leq 0$ для УФИ-распределений, причем равенство достигается лишь для экспоненциального распределения. Отсюда, с учетом п. 2, получаем п. 3.

3. Вычисление асимптотики вероятностей больших уклонений при H_0

Определим функции $k(t)$, $f^+(a)$, $f^-(a)$, $f(a)$ следующим образом:

$$k(t)=\mathbf{C}t+t+\ln\Gamma(1+t)-(1+t)\ln(1+t), \quad t>-1,$$

$$f^+(a)=\inf_{t\geq 0}\,[-at+k(t)], \quad a>0,$$

$$f^-(a)=\inf_{t\geq 0}\,[-at+k(-t)], \quad a>0,$$

$$f(a)=\max(f^+(a),\ f^-(a)), \quad a>0.$$

Теперь можем сформулировать основной результат.

Теорема 2. *При нулевой гипотезе*

$$\lim_{n\to\infty} n^{-1}\ln\mathbf{P}(T_n^+>a)=\begin{cases} f^+(a), & \text{если } 0<a\leq\mathbf{C}; \\ -\infty, & \text{если } a>\mathbf{C}, \end{cases} \tag{3.1}$$

$$\lim_{n\to\infty} n^{-1}\ln\mathbf{P}(T_n^->a)=f^-(a),\ a>0, \tag{3.2}$$

$$\lim_{n\to\infty} n^{-1}\ln \mathbf{P}(T_n > a) = f(a),\ a > 0, \tag{3.3}$$

причем функции $f^+(a)$, $f^-(a)$, $f(a)$ представимы в виде рядов, сходящихся при достаточно малых a:

$$f(a) = \sum_{j=2}^{\infty} c_j\, a^j, \quad f^+(a) = \sum_{j=2}^{\infty} c_j^+ a^j, \quad f^-(a) = \sum_{j=2}^{\infty} c_j^-\, a^j,$$

где

$$c_2 = c_2^+ = c_2^- = -\Big(2\Big(\frac{\pi^2}{6} - 1\Big)\Big)^{-1}.$$

Доказательство. Прежде всего напомним два известных факта, которые будут использованы при доказательстве Теоремы 2.

Утверждение 1 (см. [9]). *Пусть $\{W_n\}_{n=1,2...}$ – последовательность вещественнозначных случайных величин на вероятностном пространстве с мерой $\mathbf{P}$, удовлетворяющих следующим условиям:*

(i) $m_n(t) = \int e^{tW_n} d\mathbf{P} < \infty$ *для* $\forall t \in [0,\, T_1),\ T_1 > 0$

(ii) $\Psi_n(t)/n \to c_0(t) \in \mathbb{R}$ *для* $\forall t \in (T_0, T_1),\ 0 \le T_0 < T_1$, *где* $\Psi_n(t) = \ln\, m_n(t)$.

Тогда для любой последовательности $\{a_n\}$, $a_n \in \mathbb{R}$, $n = 1, 2\ldots$, где $a_n \to a \in A = \{c_0'(h) \mid c_0'(h)$ существует, непрерывна справа и строго монотонна при $h \in (T_0, T_1)\}$, выполнено $[\mathbf{P}(W_n > na_n)]^{1/n} \to \inf_{t\ge 0}\{\exp[c_0(t) - ta]\}$.[1]

Утверждение 2. *Если $X_1, \ldots, X_n$ – н.о.р.с.в., имеющие гамма-распределение (в частности, экспоненциальное) и $\overline{X}$ – их выборочное среднее, а $p(x)$ и $q(x_1, \ldots, x_n)$ – числовые функции, то $p(\overline{X})$ и $q(X_1/\overline{X}, \ldots, X_n/\overline{X})$ независимы.*

Последний факт широко известен и легко выводится из вида совместной плотности.

Не умаляя общности, можем считать, что все X_i имеют плотность e^{-x}, т.е. $X_i \sim \mathcal{E}(1)$, $i = 1, \ldots, n$. Введем в рассмотрение случайные величины U_n и V_n:

$$U_n = \mathbf{C} + \frac{1}{n}\sum_{i=1}^{n} \ln X_i - \overline{X} + 1 = \frac{1}{n}\sum_{i=1}^{n}(\mathbf{C} + \ln X_i - X_i + 1),$$

[1] Сохранены обозначения оригинала, никак не связанные с обозначениями настоящей статьи.

$$V_n = \ln \overline{X} - (\overline{X} - 1).$$

Пусть $\varphi_n(t)$, $\psi_n(t)$, $\xi_n(t)$ – производящие функции моментов (п.ф.м.) для T_n^+, U_n, V_n соответственно. Несложно убедиться в том, что все п.ф.м. конечны при $t \in [0, +\infty)$. Заметим, что $U_n = T_n^+ + V_n$, причем T_n^+ есть функция от случайного вектора $(X_1/\overline{X}, \dots, X_n/\overline{X})$, а V_n – функция от $\overline{X}$. Поэтому, согласно Утверждению 2, T_n^+ и V_n независимы. Следовательно, $\psi_n(t) = \varphi_n(t) \cdot \xi_n(t)$. Имеем

$$n^{-1} \ln\, \psi_n(nt) = n^{-1} \ln\, \varphi_n(nt) + n^{-1} \ln\, \xi_n(nt). \tag{3.4}$$

Поскольку U_n есть нормированная сумма н.о.р.с.в., функция $n^{-1} \ln\, \psi_n(nt)$ не зависит от n и равна

$$\begin{aligned} &\ln\, \mathbf{E} \exp[t(\mathbf{C} + \ln X_i - X_i + 1)] = \\ &= \mathbf{C}t + t + \ln \int \exp(t(\ln x - x)) e^{-x} dx = \mathbf{C}t + t + \ln \int x^t e^{-(1+t)x} dx = \\ &\qquad = \mathbf{C}t + t + \ln \Gamma(1+t) - (1+t)\ln(1+t). \end{aligned}$$

Окончательно,

$$n^{-1} \ln\, \psi_n(nt) = \mathbf{C}t + t + \ln \Gamma(1+t) - (1+t)\ln(1+t) \tag{3.5}$$

Далее, учитывая, что $n\overline{X}$ имеет плотность $x^{n-1}e^{-x}/\Gamma(n),\ x > 0$, вычислим $n^{-1} \ln \xi_n(nt)$. Имеем

$$\begin{aligned} \xi_n(nt) &= \int \exp\left[nt\Big(\ln \frac{x}{n} - \frac{x}{n} + 1\Big)\right] \cdot \frac{x^{n-1}e^{-x}}{\Gamma(n)} dx \\ &= \frac{e^{nt}}{n^{nt}\Gamma(n)} \int x^{nt+n-1} e^{-x(1+t)} dt = \frac{e^{nt}\Gamma(nt+n)}{n^{nt}(1+t)^{n+nt}\Gamma(n)}. \end{aligned}$$

После преобразования получаем

$$n^{-1} \ln\, \xi_n(nt) = t - t\ln n - (1+t)\ln(1+t) + n^{-1}(\ln \Gamma(nt+n) - \ln \Gamma(n)).$$

Зафиксируем t и перейдем к пределу по n, используя формулу Стирлинга

$$\ln \Gamma(n) = \ln \sqrt{2\pi} + (n - \frac{1}{2}\,)\, \ln n - n + \frac{\theta}{12n}\,,\ (0 < \theta < 1).$$

$$\text{Окончательно,}\quad \lim_{n\to\infty} n^{-1} \ln\, \xi_n(nt) = 0.$$

Комбинируя последнее равенство с (3.4) и (3.5), получаем

$$\lim_{n\to\infty} n^{-1}\ln\,\varphi_n(nt) = t(1+\mathbf{C})+\ln\Gamma(1+t)-(1+t)\ln(1+t) \equiv k(t).$$

Так как функция $k(x)$ является п.ф.м. случайной величины $\mathbf{C}+1+\ln X_1 - X_1$, ее производная $k'(x)$ непрерывна и строго возрастает на $[0,+\infty)$. Учитывая поведение производной гамма-функции в единице и на $+\infty$ (см. [3], Гл. XIV, §5), имеем

$$\lim_{t\to+\infty} k'(t) = \mathbf{C} + \lim_{t\to+\infty}\left(\frac{\Gamma'(1+t)}{\Gamma(1+t)} - \ln(1+t)\right) = \mathbf{C} \text{ и } k'(0) = 0,$$

поэтому $k'(x)$ принимает все значения из промежутка $[0,\ \mathbf{C})$. Таким образом, при $a\in[0,\ \mathbf{C})$ выполнены условия Утверждения 1 и, следовательно,

$$\lim_{n\to\infty} n^{-1}\ln\mathbf{P}(T_n^+ > a) = \inf_{t\ge 0}[-at + k(t)] \equiv f^+(t) \text{ при } 0 < a < \mathbf{C}.$$

Согласно неравенству о среднем геометрическом и среднем арифметическом, $n^{-1}\sum\ln X_i \le \ln\overline{X}$ и, следовательно, $\mathbf{P}(T_n^+ > a) = 0$, если $a\ge\mathbf{C}$. Таким образом, равенство (3.1) полностью доказано.

Функция $k(t)$ аналитична в окрестности нуля, $k'(0) = 0$ и $k''(0) = \pi^2/6-1$. Применяя формулу обращения Лагранжа, получаем для $a>0$

$$f^+(a) = -\frac{a^2}{2(\pi^2/6-1)} + \sum_{j=3}^{\infty} c_j^+ a^j, \text{ где ряд сходится при малых } a.$$

Вывод формулы (3.2) аналогичен выводу (3.1), с теми оговорками, что $\mathbf{E}\exp(ntT_n^-)$ конечно лишь при $t<1$ и область значений $(k(-t))'$ при $0<t<1$ совпадает с $\mathbb{R}^+$.

Поскольку $T_n = \max(T_n^+,\, T_n^-)$ и неравенства $T_n^+\ge a$ и $T_n^-\ge a$ выполняются одновременно лишь с вероятностью 0, верна цепочка равенств:

$$\begin{aligned}
\lim_{n\to\infty} n^{-1}\ln\mathbf{P}(T_n^+\ge a) &= \lim_{n\to\infty} n^{-1}\ln\mathbf{P}(\max(T_n^+,\, T_n^-)\ge a)\\
&= \lim_{n\to\infty} n^{-1}\ln\mathbf{P}\big((T_n^+\ge a)\cup(T_n^-\ge a)\big)\\
&= \lim_{n\to\infty} n^{-1}\ln\mathbf{P}\big((T_n^+\ge a)+\mathbf{P}(T_n^-\ge a)\big)\\
&= \lim_{n\to\infty}\left[\ln\left(1+\frac{\min(P_n^+(a),\, P_n^-(a))}{\max(P_n^+(a),\, P_n^-(a))}\right)+\ln\max(P_n^+(a),\, P_n^-(a))\right]\\
&= \lim_{n\to\infty}\ln\max(P_n^+(a),\, P_n^-(a)) = \max(f^+(a),\, f^-(a)),
\end{aligned}$$

где $P_n^+(a) := \mathbf{P}(T_n^+ \geq a)$, $P_n^-(a) := \mathbf{P}(T_n^- \geq a)$.

Заметим, что из доказательства следует непрерывность $f(a)$, $f^+(a)$, $f^-(a)$ во всех точках $a > 0$, в которых эти функции конечны.

Полученный результат можно применить также к одной статистике для проверки равномерности на отрезке.

Пусть $Z_1, \ldots, Z_{n-1}$ – н.о.р.с.в., $\mathbf{P}(0 \leq Z_j \leq 1) = 1$, $j = 1, \ldots, n$. Пусть $0 = Z_{(0)} \leq Z_{(1)} \leq \ldots \leq Z_{(n-1)} \leq Z_{(n)} = 1$ – соответствующие порядковые статистики, а $Y_j = Z_{(j)} - Z_{(j-1)}, j = 1, \ldots, n$ – спейсинги. В [5] предлагается статистика L_n для проверки равномерности на $[0, 1]$:

$$L_n = n^{-1} \sum_{j=1}^{n} \ln(nY_j).$$

Известно, что распределение случайного вектора $\{nY_j\}_{j=1}^n$ при Z_j, равномерно распределенных на $[0, 1]$, совпадает с распределением случайного вектора $\{X_j / \overline{X}\}_{j=1}^n$, где X_j – независимые случайные величины, имеющие одно и то же экспоненциальное распределение. Поэтому Теорема 2 позволяет вычислить грубую асимптотику больших уклонений статистики L_n при нулевой гипотезе, прежде неизвестную, а именно:

$$\lim_{n\to\infty} n^{-1} \ln \mathbf{P}(\mathbf{C} + L_n > a) = \begin{cases} f^+(a), & \text{если } 0 < a \leq \mathbf{C}; \\ -\infty, & \text{если } a > \mathbf{C}, \end{cases} \tag{3.6}$$

где $f^+(t)$ – функция, определенная в начале этого раздела.

4. Вычисление точных наклонов и ЛАЭ по Бахадуру

Пусть при альтернативе случайные величины X_i имеют функцию распределения $G(x, \theta)$, $\theta \in \Theta \subset \mathbb{R}$, где Θ – интервал, содержащий 0, причем $G(x, \theta) = 0$ при $x < 0$, $G(x, 0) \equiv 1 - e^{-x}$ и существует производная $G'_x(x, \theta) = g(x, \theta)$, почти всюду непрерывная по x. Будем также предполагать, что при любом $\theta \in \Theta$

$$\int x^2 g(x, \theta) dx < +\infty \quad \text{и} \quad \int (\ln x)^2 g(x, \theta) dx < +\infty.$$

Тогда по УЗБЧ имеет место сходимость $\mathbf{P}_\theta$-почти наверное:

$$|T_n| \underset{n\to\infty}{\longrightarrow} b(\theta) := \left| \mathbf{C} + \int \ln x \cdot g(x, \theta) dx - \ln \int x g(x, \theta) dx \right|.$$

(Аналогичные рассуждения для односторонних статистик здесь и далее опускаем).

Поскольку по Теореме 2 функция

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

существует и непрерывна при $\forall\, a>0$, точный наклон $\mathfrak{c}_T(\theta)$ статистики T_n при альтернативе $G(x,\,\theta)$, по теореме 7.2 из [4] , может быть вычислен по формуле

$$\mathfrak{c}_T(\theta)=-2f(b(\theta))=-2f\Big(\Big|\mathbf{C}+\int \ln x\cdot g(x,\,\theta)dx-\ln\int xg(x,\,\theta)dx\Big|\Big).$$

Согласно неравенству Бахадура–Рагавачари [10],

$$\mathfrak{c}_T(\theta)\le 2Q(\theta)=2\inf_{\lambda\ge 0}\int \ln\frac{g(x,\,\theta)}{\lambda e^{-\lambda x}}g(x,\,\theta)dx. \tag{4.1}$$

ЛАЭ по Бахадуру определяется следующим образом:

$$\mathfrak{e}_T^{B,\,loc}=\lim_{\theta\to 0}\frac{\mathfrak{c}_T(\theta)}{2Q(\theta)}.$$

Мы будем вычислять ЛАЭ в терминах функций $h(x)$ и $H(x)$, которые определим так:

$$h(x)=\frac{\partial g(x,\,\theta)}{\partial\theta}\bigg|_{\theta=0},\qquad H(x)=\frac{\partial G(x,\,\theta)}{\partial\theta}\bigg|_{\theta=0}.$$

Предположим, что выполнены следующие условия регулярности:

$$h(x)=H'(x),\ \ \forall x\ge 0 \tag{4.2}$$

$$\int h^2(x)e^x dx<+\infty, \tag{4.3}$$

$$\frac{\partial}{\partial\theta}\bigg|_{\theta=0}\int xg(x,\,\theta)dx=\int xh(x)dx, \tag{4.4}$$

$$\frac{\partial}{\partial\theta}\bigg|_{\theta=0}\int \ln xg(x,\,\theta)dx=\int \ln xh(x)dx. \tag{4.5}$$

(Конечность интегралов в правых частях (4.4) и (4.5) вытекает из (4.3) и неравенства Коши – Буняковского).

Тогда можем выделить главную часть точного наклона при $\theta \to 0$:

$$\mathfrak{c}_T(\theta) \sim \left(\frac{\pi^2}{6} - 1\right)^{-1} \left(\int (\ln x - x)\, h(x)\, dx\right)^2 \cdot \theta^2. \qquad (4.6)$$

Главная часть $2Q(\theta)$ вычислена в [8] и для плотностей $g(x,\theta)$, удовлетворяющих условиям (4.2)–(4.4), выглядит так:

$$2Q(\theta) \sim \Big[\int h^2(x) e^x\, dx - \Big(\int x\, h(x)\, dx\Big)^2\Big]\theta^2, \quad \theta \to 0,$$

если только коэффициент в правой части не равен нулю. Пусть $\beta = \int x h(x) dx$. Из (4.2) и определения $h(x)$ ясно, что $\int h(x) dx = 0$. Поэтому последнее равенство можно переписать в виде

$$2Q(\theta) \sim \int \big(e^x h(x) - \beta(x-1)\big)^2 e^{-x}\, dx \cdot \theta^2, \quad \theta \to 0. \qquad (4.7)$$

Очевидно, $\int \big(e^x h(x) - \beta(x-1)\big)^2 e^{-x}\, dx = 0$ тогда и только тогда, когда $h(x) = \beta(x-1)e^{-x}$ для какого-то $\beta \in \mathbb{R}$. При выполнении этого условия (например, при $g(x,\theta) = (e^{-x} + (1+\theta)e^{-(1+\theta)x})/2$) коэффициент при θ^2 в (4.6) также равен нулю, т.е. $2Q(\theta)$ и $\mathfrak{c}_T(\theta)$ имеют порядок малости выше 2 при $\theta \to 0$. Так как этот случай нетипичен, а вычисления для него чрезвычайно громоздки, мы не будем его здесь рассматривать.

Итак, если для $h(x)$, определенного по $g(x,\theta)$, выполнены условия (4.2)–(4.5) и верно (4.7), то

$$\mathfrak{e}_T^{B,\, loc} = \frac{\left(\int (\ln x - x) h(x) dx\right)^2}{\int h^2(x) e^x\, dx - \left(\int x\, h(x)\, dx\right)^2} \cdot \left(\frac{\pi^2}{6} - 1\right)^{-1}. \qquad (4.8)$$

Вычислим теперь ЛАЭ по Бахадуру статистики T_n для четырех стандартных семейств плотностей: плотность Макегама, плотность с линейной интенсивностью отказов (л.и.о.), плотность Вейбулла и гамма-плотность, положительные при $x > 0$ и равные 0 при $x \le 0$. Все эти плотности определены и имеют ВФИ при $\theta > 0$, а плотность Макегама, плотность Вейбулла и гамма-плотность определены также при $-1 < \theta < 0$ и в этом случае имеют УФИ. При $\theta = 0$ каждая из этих плотностей совпадает со стандартной экспоненциальной плотностью. Для краткости опустим

Вычисление ЛАЭ по Бахадуру.

Название	Плотность Макегама	Плотность с л.и.о.	Плотность Вейбулла	Гамма-плотность
Плотность	$(1+\theta(1-e^{-x}))\cdot$ $\exp(-x-$ $\theta[x-(1-e^{-x})])$	$(1+\theta x)\cdot$ $\exp(-x-\frac{\theta x^2}{2})$	$(1+\theta)x^{\theta}\cdot$ $\exp(-x^{1+\theta})$	$x^{\theta}e^{-x}\cdot$ $\frac{1}{\Gamma(1+\theta)}$
$e^x h(x)$	$2(1-e^{-x})-x$	$x-x^2/2$	$1+\ln x-x\ln x$	$\ln x+\mathbf{C}$
$2Q(\theta)/\theta^2$, $\theta\to 0$	$1/12$	1	$\pi^2/6$	$\pi^2/6-1$
$b(\theta)/\theta$, $\theta\to 0$	$\ln 2-1/2$	$1/2$	1	$\pi^2/6-1$
$\mathfrak{e}_T^{B,\,loc}$	$\frac{18(2\ln 2-1)^2}{\pi^2-6}$ $\approx 0,6941$	$\frac{3}{2(\pi^2-6)}$ $\approx 0,3876$	$\frac{1}{(\frac{\pi^2}{6}(\frac{\pi^2}{6}))}$ ≈ 0.9426	1

доказательство того, что во всех четырех случаях выполнены (4.2)–(4.5). Результаты вычислений приведены в Таблице 1.

Из таблицы видно, что статистика T_n имеет невысокую эффективность для альтернативы с л.и.о. (в частности, существенно уступает статистике Джини [8]), довольно высокую – для альтернативы Вейбулла и является локально асимптотически оптимальной (ЛАО) по Бахадуру для гамма-альтернатив, что согласуется с результатами [6] и [7]. Согласно п. 3 Теоремы 1, односторонние статистики T_n^+ и T_n^- являются состоятельными против гамма-альтернатив с $\theta > 0$ и $\theta < 0$ соответственно. Их ЛАЭ для этих альтернатив также равна 1.

Вычислим точные значения $2Q(\theta)$ и $\mathfrak{c}_{T^+}(\theta)$ при $\theta>0$ (аналогичные вычисления для T_n^- и отрицательного θ опустим для краткости).

$$2Q(\theta) = \inf_{\lambda>0} \int \frac{x^{\theta}e^{-x}}{\Gamma(1+\theta)} \ln \frac{x^{\theta}e^{-x}}{\Gamma(1+\theta)\cdot\lambda\exp(-\lambda x)}dx$$

$$= \inf_{\lambda>0}\left\{\frac{\theta\Gamma'(1+\theta)}{\Gamma(1+\theta)} - \ln\Gamma(1+\theta) - \ln\lambda + (1+\theta)(\lambda-1)\right\}.$$

Инфимум достигается при $\lambda = (1+\theta)^{-1}$ и равен

$$\frac{\theta\,\Gamma'(1+\theta)}{\Gamma(1+\theta)} - \ln\Gamma(1+\theta) + \ln(1+\theta) - \theta.$$

Таким образом,

$$\begin{aligned} b(\theta) &= \mathbf{C} + \int \ln x\,\frac{x^\theta e^{-x}}{\Gamma(1+\theta)}dx - \ln\int x\,\frac{x^\theta e^{-x}}{\Gamma(1+\theta)}dx \\ &= \mathbf{C} + \frac{\Gamma'(1+\theta)}{\Gamma(1+\theta)} - \ln(1+\theta), \end{aligned}$$

т.е. $b(\theta) = k'(\theta)$, где $k(\theta)$ – функция, определенная в разделе 3. По теореме 7.2 из [4] и теореме 2 настоящей статьи,

$$\mathfrak{c}_{T^+}(\theta) = -2f(b(\theta)) = -\inf_{t>0}(-tk'(\theta) + k(t)).$$

Очевидно, инфимум достигается при $t=\theta$ и равен

$$k(\theta) - \theta k'(\theta) = -\frac{\theta\,\Gamma'(1+\theta)}{\Gamma(1+\theta)} + \ln\Gamma(1+\theta) - \ln(1+\theta) + \theta = -2Q(\theta).$$

Таким образом, $\mathfrak{c}_{T^+}(\theta)$ совпадает с $2Q(\theta)$ при любом $\theta>0$, т.е. статистика T_n^+ является асимптотически оптимальной по Бахадуру для гамма-альтернатив с положительным параметром.

Точно так же статистика T_n^- асимптотически оптимальна по Бахадуру для гамма-альтернатив с отрицательным параметром.

Эти результаты не являются неожиданными, так как T_n^+ и T_n^- – статистики отношения правдоподобия для соответствующих гамма-альтернатив при любых разрешенных θ (как известно, статистики отношения правдоподобия являются асимптотически оптимальными по Бахадуру).

5. Область локальной асимптотической оптимальности

В предыдущем разделе мы доказали, что изучаемые статистики локально асимптотически оптимальны для гамма-альтернатив. Опишем все альтернативы, обладающие этим свойством и отвечающие условиям регулярности (4.2)–(4.5) и (4.7). Подобные задачи для классических непараметрических критериев изучались в [2], а для критерия экспоненциальности, основанного на статистике Джини – в [8]. В этих работах использовались методы вариационного исчисления, но в нашем случае можно найти более простое решение.

Пусть плотность $g(x,\theta)$ и функция $h(x)$ определяются так же, как в предыдущем разделе, причем выполнены (4.2)–(4.5) и $h(x)$ не равна $\beta e^{-x}(x-1)$ ни при каком $\beta\in\mathbb{R}$. Не умаляя общности, можем считать, что $\theta>0$. Тогда из (4.7) и (4.8) ясно, что условие ЛАО эквивалентно условию

$$\Big(\int(\ln x-x)h(x)dx\Big)^2 = \int\big(e^x h(x)-\mu(x-1)\big)^2 e^{-x}dx\cdot\Big(\frac{\pi^2}{6}-1\Big), \tag{5.1}$$

где $\mu=\int xh(x)dx=\int(x-1)h(x)dx$. Пользуясь тем, что $\int h(x)dx=\int(x-1)e^{-x}dx=0$ и $\int(\ln x-x)(x-1)e^{-x}dx=0$, перепишем левую часть (5.1) в виде

$$\int(\ln x-x)\big(h(x)-\mu(x-1)e^{-x}\big)dx = \int\big(\ln x-x+1+\mathbf{C}\big)\big(e^x h(x)-\mu(x-1)\big)e^{-x}dx \tag{5.2}$$

По неравенству Коши–Буняковского,

$$\Big(\int(\ln x-x+1+\mathbf{C})(h(x)-\mu(x-1)e^{-x})dx\Big)^2 \le \int\big(\ln x-x+1+\mathbf{C}\big)^2 e^{-x}dx\cdot\int\big(e^x h(x)-\mu(x-1)\big)^2 e^{-x}dx,$$

причем равенство достигается лишь при $e^x h(x)-\mu(x-1)=\eta(\ln x-x+1+\mathbf{C})$, $\eta\in\mathbb{R}$. Так как

$$\int(\ln x-x+1+\mathbf{C})^2\,e^{-x}dx=\frac{\pi^2}{6}-1,$$

условие (5.1) эквивалентно условию

$$h(x)=\big(\mu(x-1)+\eta(\ln x-x+1+\mathbf{C})\big)\,e^{-x},\ \text{ где } \mu,\eta\in\mathbb{R}. \tag{5.3}$$

Кроме того, $b(\theta)$ при малых $\theta>0$ имеет тот же знак, что и η, для T_n^+ и противоположный – для T_n^-, поэтому, по п. 2 теоремы 1, состоятельность односторонних статистик при малых $\theta>0$ определяется знаком η.

Сформулируем окончательный результат.

Теорема 3. *Пусть при альтернативе случайные величины имеют плотность $g(x,\theta)$, и пусть функция $h(x)$ построена по $g(x,\theta)$ так же, как в разделе 4. Предположим также, что выполнены условия (4.2)–(4.5) и $e^x h(x) \neq \beta(x-1), \forall\beta \in \mathbb{R}$. Тогда статистика T_n (T_n^+, T_n^-) является ЛАО по Бахадуру для $g(x,\theta)$ в том и только том случае, если $h(x)$ удовлетворяет условию (5.3) при некоторых вещественных μ и η, $\eta \neq 0$ ($\eta>0$, $\eta<0$ соответственно).*

Автор благодарит Я. Ю. Никитина за помощь в работе над статьей.

Литература

1. Р. Барлоу, Ф. Прошан, *Математическая теория надежности.* М., "Советское радио", 1969.
2. Я. Ю. Никитин, *Асимптотическая эффективность непараметрических критериев.* М., "Наука", 1995.
3. Г. М. Фихтенгольц, *Курс дифференциального и интегрального исчисления, Том II.* М., Физматгиз, 1962.
4. R. R. Bahadur, *Some limit theorem in statistics.* SIAM, Philadelphia, 1971.
5. D. A. Darling, *On a class of problems related to the random division of an interval.* — Ann. Math. Stat. **24** (1953), 239–253.
6. G. R. Shorack, *The best test of exponentiality against gamma alternatives.* — Journal of the American Statistical Association **67** no. 337 (1972), 213–214.
7. P. Moran, *The random division of an interval– Part II.* — Journal of the Royal Statistical Society, Ser. B **13** no. 1 (1951), 147–150.
8. Ya. Yu. Nikitin, A. V. Tchirina, *Bahadur efficiency and local optimality of a test for the exponential distribution based on the Gini statistic.* — J. Ital. Statist. Soc. **5** no. 1 (1996), 163–175.
9. 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.
10. M. Raghavachari, *On a theorem of Bahadur on the rate of convergence of test statistics.* — Ann. Math. Stat. **41** no. 4 (1970), 1695–1699.

Tchirina A. V. Bahadur efficiency and local optimality of a test for the exponentiality based on the Moran statistics.

The scale-free test for exponentiality introduced by Moran is under study. It had been constructed as an optimal test of exponentiality against the gamma alternatives but may also be used to test exponentiality against IFR and DFR classes of alternatives. We obtain the rough large deviations asymptotics of the test under the null hypothesis and find its Bahadur efficiency values for most commonly used alternatives to the exponentiality. We also describe the local Bahadur optimality domain of the test. The large deviations theorem describes as well the asymptotic behaviour of a spacings-based Darling test for the uniformity under the null hypothesis.

С.-Петербургский
государственный университет
anna@ at2706.spb.edu, antch@ mail.ru

Поступило 10 июня 2002г.