А. В. Чирина

БАХАДУРОВСКАЯ ЭФФЕКТИВНОСТЬ КРИТЕРИЯ ЭКСПОНЕНЦИАЛЬНОСТИ, ОСНОВАННОГО НА СВОЙСТВЕ ОТСУТСТВИЯ ПОСЛЕДЕЙСТВИЯ

1. Введение

Рассмотрим класс функций распределения (ф.р.)

$$\mathcal{E} = \{F(x) : F(x) = 1 - e^{-\lambda x}.\lambda > 0\}.$$

Пусть $X_1, \ldots, X_n$ – независимые одинаково распределенные случайные величины (н.о.р.с.в) с ф.р. $F(x)$. Задача проверки экспоненциальности в общем случае формулируется так: проверить истинность гипотезы $H_0 : F \in \mathcal{E}$ против альтернативы $H_1 : F \notin \mathcal{E}$.

Коул [10] предложил для проверки гипотезы экспоненциальности последовательность статистик

$$T_n = \sup_{x,y \geqslant 0} |\bar{F}_n(x+y) - \bar{F}_n(x)\bar{F}_n(y)|,$$

где

$$\bar{F}_n(x) = 1 - F_n(x),$$

а

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}_{\{X_i \leqslant x\}}$$

– эмпирическая функция распределения (э.ф.р.), построенная по выборке $X_1, \ldots, X_n$.

Строго говоря, Коул предложил односторонний вариант этой статистики, состоятельный в классе ф.р. $\{F : \bar{F}(x)\bar{F}(y) - \bar{F}(x+y) \geqslant 0\}$. Однако, в силу знаменитого характеризационного свойства отсутствия последействия, а именно:

$$\bar{F}(x)\bar{F}(y) = \bar{F}(x+y)\ \forall x, y \geqslant 0 \Longleftrightarrow F \in \mathcal{E}$$

и теоремы Гливенко–Кантелли, двусторонняя статистика $\{T_n\}$ оказывается состоятельной для любых альтернатив.

Работа частично поддержана РФФИ, гранты No. 95-01-1260 и 96-01-0852.

В [10] было найдено предельное распределение T_n в виде распределения супремума от некоторого гауссовского поля; найдены также критические значения при $n \leqslant 50$, даны рекомендации по вычислению T_n. "Упрощенный" вариант этой статистики, когда супремум берется лишь при $x = y$ (т.н. статистика Ангуса), был предложен и исследован в [5], где доказана состоятельность соответствующего критерия в некотором весьма широком классе ф.р.; асимптотические свойства этого критерия изучены в [11].

В настоящей статье вычисляется асимптотическая эффективность по Бахадуру последовательности статистик $\{T_n\}$ для часто используемых параметрических альтернатив. Для решения этой задачи в разделе 2 найдена асимптотика вероятностей больших уклонений $\{T_n\}$ при H_0 (заметим, что распределение T_n при H_0 не зависит от масштабного параметра λ).

В разделе 3 найдены значения локальной асимптотической эффективности (ЛАЭ) по Бахадуру. В разделе 4 найдены условия локальной асимптотической оптимальности $\{T_n\}$ при слабых ограничениях на класс альтернатив.

2. Большие уклонения при H_0

При анализе асимптотики вероятностей больших уклонений изучаемой статистики при справедливости гипотезы экспоненциальности важную роль играют следующие функции:

$$\lambda_{s,t}(u) \equiv e^{\frac{stu}{2}}[se^{\frac{u}{2}(1-s-t)} + (t-s)e^{-\frac{us}{2}} + s(1-t)e^{-\frac{u}{2}} + (1-s)(1-t)],$$

где $0 \leqslant s \leqslant t \leqslant 1$, $u \in \mathbb{R}^1$, и

$$k(s) \equiv \sup_{0 \leqslant s \leqslant t \leqslant 1} \inf_{u \in \mathbb{R}} \left[-\frac{au}{2} + \ln \lambda_{s,t}(u)\right] = \max(k^+(a), k^-(a)),$$

где

$$k^{\pm}(a) = \sup_{0 \leqslant s \leqslant t \leqslant 1} \inf_{u \in \mathbb{R}^{\pm}} \left[-\frac{au}{2} + \ln \lambda_{s,t}(u)\right].$$

Теорема 2.1. *Если верна гипотеза экспоненциальности, то для достаточно малых $a > 0$ выполнено равенство*

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

причем $k(a)$ непрерывна при достаточно малых a и

$$k(a) = -2a^2(1 + o(1)), \quad a \to 0.$$

Доказательство. Рассмотрим наряду со статистикой T_n односторонние статистики

$$T_n^+ = \sup_{x,y \geqslant 0} [\bar{F}_n(x+y) - \bar{F}_n(x)\bar{F}_n(y)]$$

и

$$T_n^- = \sup_{x,y \geqslant 0} [\bar{F}_n(x)\bar{F}_n(y) - \bar{F}_n(x+y)].$$

Очевидно, $T_n = \max(T_n^+, T_n^-)$. Поэтому для любого $a > 0$

$$P(T_n^+ \geqslant a) \leqslant P(T_n \geqslant a) \leqslant P(T_n^+ \geqslant a) + P(T_n^- \geqslant a)$$

и

$$P(T_n^- \geqslant a) \leqslant P(T_n \geqslant a) \leqslant P(T_n^+ \geqslant a) + P(T_n^- \geqslant a).$$

Из доказательства Теоремы 2.1 будет видно, что пределы

$$P^+(a) \equiv \lim_{n\to\infty} n^{-1} \ln P(T_n^+ \geqslant a)$$

и

$$P^-(a) \equiv \lim_{n\to\infty} n^{-1} \ln P(T_n^- \geqslant a)$$

существуют, причем

$$P^{\pm}(a) = -2a^2(1+o(1)), \quad a \to 0.$$

Следовательно, предел $P(a) \equiv \lim\limits_{n\to\infty} n^{-1} \ln(T_n \geqslant a)$ также существует и равен $\max(P^+(a), P^-(a))$.

Так как распределение T_n^+ не зависит от масштабного параметра, можем считать, не умаляя общности, что случайные величины $X_1, \dots, X_n$ имеют функцию распределения

$$F(x) = 1 - e^{-x}, \quad x \geqslant 0.$$

Применим преобразование Смирнова:

$$x = F^{-1}(s) = -\ln(1-s), \quad y = -\ln(1-t).$$

Имеем

$$T_n^+ = \sup_{0\leqslant s,t\leqslant 1} [\bar{F}_n(F^{-1}(s) + F^{-1}(t)) - \bar{F}_n(F^{-1}(s)) \cdot \bar{F}_n(F^{-1}(t))] =$$

$$= \sup_{0\leqslant s,t\leqslant 1} [\bar{F}_n(F^{-1}(s+t-st)) - \bar{F}_n(F^{-1}(s)) \cdot \bar{F}_n(F^{-1}(t))] =$$

$$= \sup_{0\leqslant s,t\leqslant 1} [\bar{G}_n(s+t-st) - \bar{G}_n(s)\bar{G}_n(t)] =$$

$$= \sup_{0\leqslant s\leqslant t\leqslant 1} [G_n(s) + G_n(t) - G_n(s+t-st) - G_n(s)G_n(t)],$$

где G_n – э.ф.р., соответствующая выборке из равномерного на $[0,1]$ распределения. Выражение, от которого берется супремум, симметрично относительно s и t, поэтому, не умаляя общности, в дальнейшем можем для удобства считать $s\leqslant t$.

Представим $G_n(s)+G_n(t)-G_n(s+t-st)-G_n(s)G_n(t) \equiv T_n^+(s,t)$ в форме функционала Мизеса.

Для любых $s\leqslant t$ имеем

$$T_n^+(s,t) = \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\Phi_{s,t}(X_i,X_j),$$

где X_i равномерно распределены на $[0,1]$ и независимы, $\Phi_{s,t}(a,b)$ – симметричное ядро,

$$\Phi_{s,t}(a,b) = \frac{1}{2}\cdot\big(\mathbb{I}_{\{a<s\}} + \mathbb{I}_{\{b<s\}} + \mathbb{I}_{\{a<t\}} + \mathbb{I}_{\{b<t\}} -$$

$$-\mathbb{I}_{\{a<s+t-st\}} - \mathbb{I}_{\{b<s+t-st\}} - \mathbb{I}_{\{a<s,b<t\}} - \mathbb{I}_{\{a<t,b<s\}}\big).$$

Легко видеть, что $E\Phi_{s,t}(X_1,X_2)=0$. Заметим, что $|\Phi_{s,t}(a,b)|\leqslant 1$ для любых s,t,a,b, откуда для $\forall n$ с вероятностью 1

$$\left| n^{-2}\sum_{i=1}^{n}\sum_{j=1}^{n}\Phi_{s,t}(X_i,X_j) - (C_n^2)^{-1}\sum_{1\leqslant i<j\leqslant 1}\Phi_{s,t}(X_i,X_j)\right| \leqslant \frac{3}{n}, \qquad (1)$$

т.е. $T_n^+(s,t)$ удается приблизить с помощью U-статистики с ядром $\Phi_{s,t}(X_i,X_j)$.

Введем функцию

$$\Psi_{s,t}(z) = E[\Phi_{s,t}(X_1,X_2)\,|\,X_1=z] =$$

$$= \frac{1}{2}[(1-t)\mathbb{I}_{\{z<s\}} + (1-s)\mathbb{I}_{\{z<t\}} + st - \mathbb{I}_{\{z<s+t-st\}}].$$

При $s\leqslant 1$ имеем

$$E\Psi_{s,t}^2(X_1) = \frac{1}{4}s(1-s)((1-t)(2-t) > 0 \text{ при } 0<s\leqslant t<1$$

$$\sup_{0\leqslant s\leqslant t\leqslant 1} E\Psi^2_{s,t}(X_1) = \frac{1}{16}, \tag{2}$$

поскольку супремум в левой части достигается при $s = t = 1-\frac{1}{\sqrt{2}}$. Для $s \leqslant t$ и $u \geqslant 0$ рассмотрим функцию

$$\lambda_{s,t}(u) \equiv E\exp(u\Psi_{s,t}(X_1)) =$$

$$= e^{\frac{stu}{2}}[se^{\frac{u}{2}(1-s-t)} + (t-s)e^{-\frac{us}{2}} + s(1-t)e^{-\frac{u}{2}} + (1-s)(1-t)].$$

Для доказательства неравенства

$$\varliminf_{n\to\infty} n^{-1}\ln P(T_n^+ \geqslant a) \geqslant k^+(a) \tag{3}$$

воспользуемся результатами Чернова [8] и Дасгупты [9, 3], описывающими асимптотику вероятностей больших уклонений для сумм и невырожденных U-статистик.

Теорема 2.2 ([7, 8]). *Пусть $\{Y_i\}$ – н.о.р.с.в., $P(Y_i > 0) > 0$, $\{\gamma_n\}$ – числовая последовательность, $\gamma_n \to \gamma$. Тогда*

$$\lim_{n\to\infty} n^{-1}\ln P(Y_1 + \ldots + Y_n \geqslant n\gamma_n) = f(\gamma),$$

где $g(\gamma) = \inf\{-t\gamma + \ln\varphi(t) : t \geqslant 0\}$, а $\varphi(t) = E\exp(tY_1)$.

Теорема 2.3 ([3, 9]). *Пусть Φ – функция из $\mathbb{R}^m$ в $\mathbb{R}^1$, симметричная относительно всех перестановок своих переменных. Положим*

$$U_n = (C_n^m)^{-1} \sum_{1\leqslant i_1<\ldots<i_m\leqslant n} \Phi(Y_{i_1}, \ldots, Y_{i_m}),$$

где $Y_1, \ldots, Y_n$ – н.о.р.с.в.;

$$\Psi(t) := E[\Phi(Y_1, \ldots, Y_m) \,|\, Y_1 = t].$$

Если $E\Phi(Y_1, \ldots, Y_m) = 0$, $E\Psi^2(Y_1) > 0$, $E\exp(t\Phi^2(Y_1, \ldots, Y_m)) < \infty$ для $\forall t \in \mathbb{R}^1$ и $\exists\varepsilon_0 > 0$:

$$\lim_{n\to\infty} n^{-1}\ln P\left(n^{-1}\sum_{i=1}^n \Psi(X_i) > \varepsilon_0\right) < \infty,$$

то

$$P(U_n \geqslant m\varepsilon) = P\left(n^{-1}\sum_{i=1}^n \Psi(Y_i) \geqslant \varepsilon\right)\cdot(1+o(1)), \quad n\to\infty.$$

Очевидно, при любых s, t из $[0,1]$

$$P(T_n^+ \geqslant a) \geqslant P(T_n^+(s,t) \geqslant a).$$

По неравенству (1),

$$P(T_n^+(s,t) \geqslant a) = \left(n^{-2} \sum_{i,j=1}^{n} \Phi_{s,t}(X_i, X_j) \geqslant a \right) =$$

$$= P\left((C_n^2)^{-1} \sum_{1 \leqslant i < j \leqslant m} \Phi_{s,t}(X_i, X_j) \geqslant a + 3n^{-1} \right) \cdot (1 + o(1)), \quad n \to \infty.$$

Применяя Теоремы 2.3 и 2.2, имеем

$$\varliminf_{n\to\infty} n^{-1} \ln P(T_n^+ \geqslant a) \geqslant \lim_{n\to\infty} n^{-1} \ln P \left(n^{-1} \sum_{i=1}^{n} \Psi_{s,t}(X_i) \geqslant \frac{a}{2} \right) =$$

$$= \inf_{u \geqslant 0} \left(-\frac{au}{2} + \ln \lambda_{s,t}(u) \right)$$

для любых s, t: $0 \leqslant s \leqslant t \leqslant 1$. По произвольности s и t,

$$\varliminf_{n\to\infty} n^{-1} \ln P(T_n^+ \geqslant a) \geqslant \sup_{0 \leqslant s \leqslant t \leqslant 1} \inf_{u \geqslant 0} \left[-\frac{au}{2} + \ln \lambda_{s,t}(u) \right] \equiv k^+(a).$$

Для доказательства противоположного неравенства

$$\varlimsup_{n\to\infty} n^{-1} \ln P(T_n^+ \geqslant a) \leqslant k^+(a) \tag{4}$$

нам понадобятся две леммы.

Лемма 2.1. $k^+(a) = -2a^2(1 + o(1))$, $a \to 0$.

Лемма 2.2. *При достаточно малых положительных* a *функция* $k^+(a)$ *непрерывна.*

Доказательство Леммы 2.1. Для $\forall u_0 \geqslant 0$

$$\sup_{0 \leqslant s \leqslant t \leqslant 1} \inf_{u \geqslant 0} \left[-\frac{au}{2} + \ln \lambda_{s,t}(u) \right] \leqslant \sup_{0 \leqslant s \leqslant t \leqslant 1} \left[-\frac{au_0}{2} + \ln \lambda_{s,t}(u_0) \right].$$

Положим $u_0 = 8a$, и пусть $0 < a < 1/64$. Воспользуемся стандартным неравенством

$$\ln(1 - b) \leqslant -b - \frac{b^2}{2} + \frac{|b|^3}{2}, \quad |b| < \frac{1}{9}.$$

Для $\forall s,t: 0 \leqslant s \leqslant t \leqslant 1$ справедлива цепочка неравенств:

$$-\frac{au_0}{2} + \ln \lambda_{s,t}(u_0) \leqslant$$

$$\leqslant -\frac{au_0}{2} + \frac{stu_0}{2} + \ln \left[s\left(\frac{u_0}{2}(1-s-t) + \frac{u_0^2}{8}(1-s-t)^2 \right) + \right.$$

$$\left. +(t-s)\left(1 - \frac{u_0 s}{2} + \frac{u_0^2 s^2}{8}\right) + s(1-t)\left(1 - \frac{u_0}{2} + \frac{u_0^2}{8}\right) + (1-s)(1-t) + C_2 u_0^3 \right] \leqslant$$

$$\leqslant -\frac{au_0}{2} + \frac{u_0^2}{8}\, s(1-s)(1-t)(2-t) + C_2 u_0^3 \leqslant$$

$$\leqslant -2a^2(1+o(1)), \quad a \to 0,$$

причем последнее неравенство выполняется в силу (2). Т.о. доказано, что

$$k^+(a) \leqslant -2a^2(1+o(1)), \quad a \to 0. \tag{5}$$

В [11] было доказано неравенство

$$\inf_{u \geqslant 0} \left[-\frac{au}{2} + \lambda_{s_0,t_0}(u) \right] \geqslant -2a^2(1+o(1)), \quad a \to 0,$$

где

$$s_0 = t_0 = 1 - \frac{1}{\sqrt{2}}.$$

Очевидно,

$$\sup_{0 \leqslant s \leqslant t \leqslant 1} \inf_{u \geqslant 0} \left[-\frac{au}{2} + \ln \lambda_{s,t}(u) \right] \geqslant$$

$$\geqslant \inf_{u \geqslant 0} \left[-\frac{au}{2} + \ln \lambda_{s_0,t_0}(u) \right] \text{ для } \forall s_0, t_0.$$

т.о. $k^+(a) \geqslant -2a^2(1+o(1))$, $a \to 0$. Отсюда и из (5) следует утверждение Леммы 2.1.

Доказательство Леммы 2.2. Можно показать, что для достаточно малых положительных a существует такое $B \in (0,+\infty)$, что

$$k^+(a) = \sup_{0 \leqslant s \leqslant t \leqslant 1} \inf_{0 \leqslant u \leqslant B} \left[-\frac{au}{2} + \ln \lambda_{s,t}(u) \right],$$

т.е. можно брать

$$\sup \inf \left[-\frac{au}{2} + \ln \lambda_{s,t}(u) \right]$$

по компактному множеству (например, подходит $B = 64\ln 10 < 200$); далее можем воспользоваться Предложением VIII Приложения III в [2].

Подробное доказательство мы опускаем, т.к. оно почти дословно повторяет рассуждения в [11].

Вернемся теперь к доказательству неравенства (4). Пусть

$$N \in \mathbb{N}, \quad k, m \in \mathbb{Z}, \quad 0 \leqslant m, k \leqslant N.$$

Имеем

$$T_n^+ = \max_{\substack{0\leqslant m\leqslant n-1\\0\leqslant k\leqslant N-1}} \max_{\substack{\frac{m}{N}\leqslant s\leqslant\frac{m+1}{N}\\\frac{k}{N}\leqslant t\leqslant\frac{k+1}{N}}} T_n^+(s,t) \leqslant$$

$$\leqslant \max_{\substack{0\leqslant m\leqslant N-1\\0\leqslant k\leqslant N-1}} \left\{T_n^+\left(\frac{m}{N},\frac{k}{N}\right) + G_n\left(\frac{m+1}{N}\right) - \right.$$

$$\left. -G_n\left(\frac{m}{N}\right) + G_n\left(\frac{k+1}{N}\right) - G_n\left(\frac{k}{N}\right)\right\} \leqslant$$

$$\leqslant \max_{\substack{0\leqslant m\leqslant N-1\\0\leqslant k\leqslant N-1}} T_n^+\left(\frac{m}{N},\frac{k}{N}\right) + 2\max_{0\leqslant m\leqslant N-1}\left[G_n\left(\frac{m+1}{N}\right) - G_n\left(\frac{m}{N}\right)\right] =$$

$$=: R_{1_n} + R_{2_n}.$$

Для любых достаточно малых $\delta_N > 0$, $a > 0$ имеем

$$P(T_n^+ \geqslant a) \leqslant P(R_{1_n} + R_{2_n} \geqslant a) \leqslant P(R_{1_n} \geqslant a - \delta_N) + P(R_{2_n} \geqslant \delta_N).$$

Оценки слагаемых в правой части этого неравенства аналогичны оценкам в [11].

Имеем

$$P(R_{1_n} \geqslant a - \delta_N) \leqslant \frac{N(N+1)}{2}\exp(nk^+(a-\delta_N))\cdot(1+o(1)), \quad n\to\infty$$

и

$$P(R_{2_n} \geqslant \delta_N) \leqslant 2N\exp\big(-n\big(-\ln 2 + \frac{\delta_N}{2}\ln(N-1)\big)\big).$$

Выберем

$$\delta_N = \frac{1}{\ln\ln N}, \quad N \geqslant 3.$$

Для достаточно малых $a > 0$ имеем

$$\overline{\lim_{n\to\infty}}\, n^{-1}\ln P(T^{+})n \geqslant a) \leqslant \max\left[k^{+}(a-\delta_N), \ln 2 - \frac{\delta_N \ln(N-1)}{2}\right] =$$

$$= k^{+}(a-\delta_N)$$

для достаточно больших N.

Так как по Лемме 2.2 $k^{+}(a)$ непрерывна, переходя к пределу при $N \to \infty$, имеем

$$\overline{\lim_{n\to\infty}}\, n^{-1}\ln P(T_n^{+} \geqslant a) \leqslant k^{+}(a).$$

Комбинируя это неравенство с (3), имеем

$$\lim_{n\to\infty} n^{-1}\ln P(T_n^{+} \geqslant a) = k^{+}(a).$$

С помощью аналогичных рассуждений можем получить

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

где

$$k^{-}(a) \equiv \sup_{0\leqslant s,t\leqslant 1}\inf_{u\leqslant 0}\left[-\frac{au}{2} + \ln\lambda_{s,t}(u)\right],$$

причем $k^{-}(a)$ непрерывна при малых a и

$$k^{-}(a) = -2a^2(1+o(1)), \quad a \to 0. \tag{7}$$

Из (6) и (7) следует утверждение Теоремы 2.1.

3. Вычисление точных наклонов на АОЭ по Бахадуру

Зная асимптотику вероятностей больших уклонений изучаемой статистики, мы можем вычислить ее локальную асимптотическую эффективность (ЛАЭ) по Бахадуру. Воспользуемся для этого результатами теории Бахадура ([6, 7, 4]).

Предположим, что при альтернативе наблюдения имеют ф.р. $G(x;\theta)$, совпадающие с $1-e^{-x}$ только при $\theta = 0$. Будем также считать, что $G(x,\theta)$ при почти всех x дифференцируема по θ.

Введем обозначения

$$h(x) \equiv \left.\frac{\partial}{\partial\theta}\right|_{\theta=0} G(x;\theta) \tag{8}$$

и
$$g(x;\theta) \equiv \frac{\partial}{\partial x} G(x;\theta).$$

По теореме Гливенко–Кантелли, с вероятностью 1 при альтернативе $G(x;\theta)$
$$T_n \underset{n\to\infty}{\to} b_T(\theta) \equiv \sup_{x,y\geqslant 0} |\bar{G}(x+y;\theta) - \bar{G}(x;\theta)\bar{G}(y;\theta)| \sim$$
$$\sim \sup_{x,y\geqslant 0} |h(x+y) - e^{-x}h(y) - e^{-y}h(x)| \cdot \theta, \quad \theta \to 0. \tag{9}$$

Используя теорему Бахадура о вычислении точных наклонов (Т.7.2., [7]), получаем, комбинируя Лемму 2.2 и Теорему 2.1, что бахадуровский точный наклон $C_T(\theta)$ последовательности $\{T_n\}$ для достаточно малых θ может быть вычислен по формуле
$$C_T(\theta) = 2k(b_T(\theta)).$$

По Лемме 2.1 и (9), можно извлечь главную часть $C_T(\theta)$ при $\theta \to 0$, т.н. локальный точный наклон:
$$C_T(\theta) \sim 4 \sup_{x,y\geqslant 0} |h(x+y) - e^{-x}h(y) - e^{-y}h(x)|^2\theta^2.$$

Мы будем сравнивать точные наклоны с их потенциальными верхними границами, которые вычисляются в терминах информации Кульбака–Лейблера и определяются из неравенства Бахадура–Рагавачари ([6, 7, 13]). В нашем случае для любого семейства плотностей $g(x,\theta)$ имеем
$$C_T(\theta) \leqslant 2Q(\theta) \equiv 2 \inf_{\lambda>0} \int_0^\infty \ln \frac{g(x;\theta)}{\lambda e^{-\lambda x}}\, g(x;\theta)dx.$$

Вычислить явно $2Q(\theta)$ сложно, но в достаточно регулярных случаях главная часть $2Q(\theta)$ вычисляется по формуле (см. [12])
$$2Q(\theta) \sim \left(\int_0^\infty e^x (h'(x))^2 dx - \left(\int_0^\infty h(x)dx \right)^2 \right) \theta^2, \quad \theta \to 0. \tag{10}$$

Таким образом, ЛАЭ изучаемой статистики вычисляется по формуле

$$e^B(T) \equiv \lim_{\theta \to 0} \frac{C_T(\theta)}{2Q(\theta)} = \frac{4 \sup\limits_{x,y \geqslant 0} |h(x+y) - e^{-x}h(y) - e^{-y}h(x)|^2}{\int\limits_0^\infty e^x (h'(x))^2 dx - \left(\int\limits_0^\infty h(x)dx\right)^2}. \quad (11)$$

Мы найдем значение $e^B(T)$ для следующих стандартных семейств альтернатив:

1°) Плотность Вейбулла:

$$g_1(x;\theta) = x^\theta \exp(-x^{1+\theta}), \quad x \geqslant 0, \quad \theta \geqslant 0.$$

2°) Плотность Макегама:

$$g_2(x;\theta) = (1 + \theta(1 - e^{-x})) \exp(-x - \theta(x - (1 - e^{-x}))), \quad x \geqslant 0, \quad \theta \geqslant 0.$$

3°) Плотность с линейной интенсивностью отказов:

$$g_3(x;\theta) = (1 + \theta x) \exp(-x - \theta x^2/2), \quad x \geqslant 0, \quad \theta \geqslant 0.$$

Заметим, что для этих плотностей супремум

$$|h(x+y) - e^{-x}h(y) - e^{-y}h(x)| \quad (12)$$

достигается при $x = y$.

Поэтому значения $b_T(\theta)$ для них эквивалентны (при $\theta \to 0$) значениям $b_s(\theta)$, где $\{S_n\}$ – т.н. статистика Ангуса ([5]), или "упрощенный" вариант статистики $\{T_n\}$:

$$S_n = \sup_{x \geqslant 0} |\bar{F}_n(2x) - \bar{F}_n^2(x)|.$$

В [11] доказано, что

$$\lim_{n \to \infty} n^{-1} \ln P(S_n \geqslant a) = -2a(1 + o(1)), \quad a \to 0.$$

Поэтому, в силу (12), значения ЛАЭ по Бахадуру статистики $\{T_n\}$ для альтернатив $g_1(x;\theta)$, $g_2(x;\theta)$, $g_3(x;\theta)$ совпадают со значениями ЛАЭ статистики $\{S_n\}$ для этих альтернатив, вычисленными в [11]:

$$\begin{array}{lll} 24\pi^{-2}e^{-2}\ln^2 2 \approx 0.1581 & \text{для} & g_1(x;\theta); \\ 3/16 = 0.1875 & \text{для} & g_2(x;\theta); \\ 4e^{-4} \approx 0.0733 & \text{для} & g_3(x;\theta). \end{array}$$

Автору неизвестно, существуют ли альтернативы, для которых ЛАЭ статистик $\{T_n\}$ и $\{S_n\}$ различны (ясно, во всяком случае, что всегда
$$e^B(T) \geqslant e^B(S),$$
т.е. существуют ли такие $g(x;\theta)$ и соответствующие им $h(x)$, что супремум
$$|h(x+y) - e^{-x}h(y) - e^{-y}h(x)|$$
достигается при $x \neq y$.

4. Условия локальной оптимальности

Для альтернатив g_1, g_2, g_3 эффективность $\{T_n\}$ невысока. Существуют, однако, семейства альтернатив, для которых последовательность $\{T_n\}$ является локально асимптотически оптимальной (ЛАО) по Бахадуру, т.е.
$$\lim_{\theta\to 0} \frac{C_T(\theta)}{2Q(\theta)} = 1.$$
Задача нахождения таких альтернатив сводится к максимизации функционала
$$\mathcal{F}(h) \equiv \sup_{x,y\geqslant 0} |h(x+y) - e^{-x}h(y) - e^{-y}h(x)|$$
на множестве
$$\mathfrak{M} = \{h : \mathcal{G}(h) = C\},$$
где
$$\mathcal{G}(h) \equiv \int\limits_0^\infty e^t (h'(t))^2 dt - \left(\int\limits_0^\infty h(t)dt\right)^2,$$
а C – некоторая константа, которая, в силу однородности $\mathcal{F}$ и $\mathcal{G}$, может быть выбрана произвольно (например, $C = 1$).

Введем функционал
$$\mathcal{F}_{x_0,y_0} \equiv |h(x_0+y_0) - e^{-x_0}h(y_0) - e^{-y_0}h(x_0)|$$
для любых $x_0, y_0 \geqslant 0$.

Заметим, что
$$\max_{\mathfrak{M}} \mathcal{F}(h) = \max_{\mathfrak{M}} \max_{x_0,y_0\geqslant 0} \mathcal{F}_{x_0,y_0}(h)$$

и

$$\mathcal{F}_{x_0,y_0}(h) = \int\limits_0^\infty (\mathbb{I}_{[0,x_0+y_0)}(t) - e^{-x_0}\mathbb{I}_{[0,y_0)}(t) - e^{-y_0}\mathbb{I}_{[0,x_0)}(t))h'(t)dt.$$

Максимизируя $\mathcal{F}_{x_0,y_0}(h)$ при условии $\mathcal{G}(h) = 1$, имеем, из принципа Лагранжа (см., например, [1], 3.2), дифференциальное уравнение для h:

$$a(t) - 2\lambda \cdot (e^t h'(t) + \nu t) = D, \tag{13}$$

где

$$a(t) = \mathbb{I}_{[0,x_0+y_0)}(t) - e^{-x_0}\mathbb{I}_{[0,y_0)}(t) - e^{-y_0}\mathbb{I}_{[0,x_0)}(t),$$

λ и D – некоторые константы, и выполнены условия

$$\begin{cases} \int\limits_0^\infty h(t)dt = \nu \\ h(0) = h(+\infty) = 0 \\ h(t) \text{ непрерывна на } [0,+\infty) \end{cases} \tag{14}$$

Решая (13) и учитывая (14), имеем

$$h(t) = \mu h_0(t) + \nu t e^{-t}, \tag{15}$$

где

$$h_0(t) = \begin{cases} (1-e^{-t}) \cdot e^{-x_0-y_0}, & 0 \leqslant t < x_0 \\ e^{-t} \cdot e^{-y_0}(1-e^{-x_0}) & x_0 \leqslant t < y_0 \\ e^{-x_0-y_0} - e^{-t}(e^{-x_0} + e^{-y_0} - e^{-x_0-y_0}), & y_0 \leqslant t < x_0 + y_0 \\ e^{-t} \cdot (1-e^{-x_0}) \cdot (1-e^{-y_0}), & t \geqslant x_0 + y_0. \end{cases}$$

В результате громоздких, но элементарных вычислений удается показать, что

$$\sup_{x,y \geqslant 0} |h_0(x+y) - e^{-x}h_0(y) - e^{-y}h_0(x)|$$

достигается при

$$x + y = x_0 + y_0, \quad x_0 \leqslant x, y \leqslant y_0 + x_0, \quad y_0 \leqslant x, y \leqslant x_0 + y_0$$

и

$$\max_{x_0,y_0} \max_{\mathfrak{M}} \mathcal{F}_{x_0 y_0}(h)$$

достигается при $x_0 = y_0 = 1 - \frac{1}{\sqrt{2}}$.

В этом случае $h_0(t)$ принимает вид

$$h_0(t) = \begin{cases} 1 - e^{-t}, & 0 \leqslant t < \frac{1}{2}\ln 2 \\ (2\sqrt{2}-1)e^{-t} - 1, & \frac{1}{2}\ln 2 \leqslant t < \ln 2 \\ -(3-2\sqrt{2})e^{-t}, & t \geqslant \ln 2 \end{cases}$$

При таких h_0 значение $e^B(T)$ в (11) равно 1.

Теперь можем сформулировать основной результат этого раздела.

Теорема 4.1. *В классе альтернатив $G(t;\theta)$, для которых верно* (10), *критерий экспоненциальности, основанный на $\{T_n\}$, является ЛАО по Бахадуру тогда и только тогда, когда $h(t)$, введенная в* (8), *имеет вид* (15) *с $h_0(t)$, имеющей вид* (16).

Легко проверить, например, что $e^{-t} + \theta h_0(t)$ при малых θ являются плотностями.

Так как и в этом случае супремум достигается при $x = y$, $\{S_n\}$ также является ЛАО по Бахадуру для альтернатив с ф.р. $G(t;\theta)$, отвечающей Теореме 4.1.

Литература

1. В. М. Алексеев, В. М. Тихомиров, С. В. Фомин, *Оптимальное управление.* М., Наука, 1979.
2. В. Ф. Демьянов, В. Н. Малоземов, *Введение в минимакс.* М., Наука, 1972.
3. В. С. Королюк, Ю. В. Боровских, *Теория U-статистик.* Киев, Наукова Думка, 1989.
4. Я. Ю. Никитин, *Асимптотическая эффективность непараметрических критериев.* М., Физматлит, 1995.
5. J. E. Angus, *Goodness-of-fit tests for exponentiality based on a loss-of-memory type functional equation.* — J. Statist. Plann. Inference **6** (1982), 241–251.
6. R. R. Bahadur, *Rates of convergence of estimates and test statistics.* — Ann. Math Statist. **38** (1967), 3-3–324.
7. R. R. Bahadur, *Some limit theorems in statistics.* SIAM, Philadelphia, 1971.
8. H. A. Chernoff, *A measure of asymptotic efficiency for tests of a hypothesis based on a sum of observations.* — Ann. Math. Statist. **23**, No. **4** (1952), 493–507.
9. R. Dasgupta, *On large deviation probabilities of U-statistics in non-i.i.d. case.* — Sankhyā, Ind. J. Statist. **A46** (1984), 110–116.
10. H. L. Koul, *A test for new better than used.* — Comm. Statist. Theory Meth. **A6**, No. **6** (1977), 563–574.

11. Ya. Yu. Nikitin, *Bahadur efficientcy of a test of exponentiality based on a loss-of-memory type functional equation.* — J. of Nonparametric Statist. **6**, No. **1** (1996), 13–26.
12. Ya. Yu. Nikitin, A. V. Tchirina, *Bahadur efficiency and local optimality of a test for the exponential distribution based on the Gini statistic.* Ricerche, No. 22, Ser. A, 1996, Universita degli Studi di Roma "La Sapienza," pp. 1–12.
13. M. Raghavachari, *On a theorem of Bahadur on the rate of convergence of test statistics.* — Ann. Math. Statist. **41** (1970), 1695–1699.

Tchirina A. V. Bahadur efficiency of a test of exponentiality based on the loss-of-memory property.

A test of exponentiality of Kolmogorov–Smirnov type based on the loss-of-memory characterization property is under study. Large deviation asymptotics under the null-hypothesis is calculated. Local Bahadur efficiencies for commonly used parametric alternatives are found. Conditions of the Bahadur local optimality are obtained.

Санкт-Петербургский
государственный университет

Поступило 8 октября 1997 г.