$$+6400D\{S_2^{(k)}\}-9600\operatorname{cov}\{S_1^{(k)},S_2^{(k)}\},$$
$$-14D_a-\frac{1}{2}D_b-\frac{1}{2}D_c=2880D\{S_1^{(k)}\}+$$
$$+4800D\{S_2^{(k)}\}-7440\operatorname{cov}\{S_1^{(k)},S_2^{(k)}\}$$

Разрешаем эту систему относительно D_a, D_b и D_c и подставляем в формулу для средней интегральной погрешности

$$\mathrm{D}=M\left\{\frac{1}{T}\int_0^T\left[\hat{a}_k+\hat{b}_k\frac{t}{T}+\hat{c}_k\left(\frac{t}{T}\right)^2-a_k-b_k\frac{t}{T}-c_k\times\right.\right.$$
$$\left.\left.\times\left(\frac{t}{T}\right)^2\right]^2dt\right\},$$ которая для стационарного режима принимает вид

$$\mathrm{D}=-M\left\{\frac{1}{N}\right\}\left[-\frac{4}{9}D_a+\frac{1}{12}D_b-\frac{1}{20}D_c\right],$$

или $$\mathrm{D}=M\left\{\frac{1}{N}\right\}\left[\frac{107}{840}\overline{b_k^2}+\frac{23}{84}\overline{b_kc_k}+\frac{53}{315}\overline{c_k^2}+3\sigma^2\right].$$

Отметим, что если считать исходный тренд стационарным случайным процессом, то $\overline{b_k^2}$, $\overline{b_kc_k}$ и $\overline{c_k^2}$ не являются независимыми величинами. Действительно, усредняя уравнение (1), имеем $\overline{a_k}=\overline{a_{k-1}}+\overline{b_{k-1}}+\overline{c_{k-1}}$. В стационарном случае все эти величины не должны зависеть от номера k. Тогда из последнего уравнения следует, что $\overline{b_k}+\overline{c_k}=0$, а величина $\overline{a_k}$ может быть произвольной. Записывая (1) в виде $\overline{a_k}-\overline{a_{k-1}}=\overline{b_{k-1}}+\overline{c_{k-1}}$, возводя его в квадрат и усредняя, получим $\overline{(a_k-a_{k-1})^2}=\overline{(b_{k-1}+c_{k-1})^2}$. Если обозначить $\overline{(a_k-a_{k-1})^2}$ через V, то в силу стационарности сплайна эта величина не зависит от k. Заметим, что это тот же самый параметр $\overline{(x_k-x_{k-1})^2}$, что и для линейных сплайнов первого порядка [3].

Так как $M\left\{\frac{1}{N}\right\}=\frac{1}{\lambda T}+\frac{1}{(\lambda T)^2}-\frac{3}{(\lambda T)^3}+\cdots$, то при $\lambda T\gg 1$ окончательно имеем $D\cong\frac{1}{\lambda T}\left[0{,}16V+3\sigma^2\right]$.

Из полученной формулы имеем, что средняя интегральная погрешность выделения тренда отлична от нуля при отсутствии помех измерения. Это обусловлено тем, что моменты измерений являются случайными величинами и вносят дополнительную погрешность в измеренные значения временного ряда.

ЛИТЕРАТУРА

1. *Андерсон Т.* Статистический анализ временных рядов. М.: Мир, 1976. 755 с.
2. *Кендалл М.Д., Стьюарт А.* Многомерный статистический анализ и временные ряды. М.: Наука, 1976. 736 с.
3. *Тривоженко Б.Е.* Выделение трендов временных рядов и потоков событий. Томск: Изд-во ТГУ, 1989. 285 с.

Статья представлена кафедрой теории вероятностей и математической статистики факультета прикладной математики и кибернетики Томского государственного университета, поступила в научную редакцию 14 марта 2000 г.

УДК 519.2

Б.Е. Тривоженко

АСИМПТОТИЧЕСКИЕ СВОЙСТВА МНК-ОЦЕНОК ПАРАМЕТРОВ ГИПЕРБОЛИЧЕСКОГО ТРЕНДА ИНТЕНСИВНОСТИ ПУАССОНОВСКОГО ПОТОКА

Получены уравнения для МНК-оценок параметров гиперболического тренда пуассоновского потока. Исследованы их статистические свойства. Показано, что эти оценки являются несмещёнными. Получены выражения для подтверждения эффективности МНК-оценок параметров и эффективности МНК-оценки гиперболического тренда.

Пуассоновский поток событий является простейшим для изучения, поскольку его свойства полностью описываются единственной функцией – его интенсивностью $\lambda(t)$ [1]. В нестационарном потоке сохраняются основные свойства, делающие его изучение лёгким, – независимость событий и ординарность.

Обозначим $\int_0^z\lambda(u)du$ через $\Lambda(t)$. Тогда вероятность того, что на интервале $[t_1,t_2]$ наступит i событий, будет

$$p_i=\frac{1}{i!}\left[\Lambda(t_2)-\Lambda(t_1)\right]^ie^{-[\Lambda(t_2)-\Lambda(t_1)]},\qquad(1)$$

а вероятность того, что на бесконечно малом интервале $[t,t+dt]$ наступит одно событие, равна $\lambda(t)dt$.

Будем предполагать, что интенсивность пуассоновского потока описывается функцией

$$\lambda(t)=\frac{1}{a+bt}.\qquad(2)$$

Такой вид зависимости интенсивности от t соответствует линейному тренду средней длительности интервалов между двумя последующими событиями в нестационарном пуассоновском потоке.

По наблюдениям моментов наступления событий $t_1<t_2<\ldots<t_N$ необходимо оценить неизвестные параметры a и b, т.е. выделить тренд интенсивности. Число событий N, по которым оцениваются параметры, фиксировано, и мы в данном случае имеем дело с N-планами эксперимента. Поскольку N фиксировано, то для того, чтобы можно было применять асимптотические методы, представим $\lambda(T)$ в виде

$$\lambda_0(t)=\lambda\left(\frac{t}{N}\right)=\frac{1}{a+bt/N}.\qquad(3)$$

Если t_i – момент наступления i-го события, то в [2] было показано, что асимптотически при $N\to\infty$

$$M\left\{\frac{t_i}{N}\right\}=T\left(\frac{i}{N}\right). \qquad (4)$$

Для функции вида (2) $T(x)=\frac{a}{b}(e^{bx}-1)$.

Оценки неизвестных параметров будем находить методом наименьших квадратов (МНК), т.е. из критерия:

$$Q=\frac{1}{N}\sum_{i=1}^{N}\left[\frac{t_i}{N}-\frac{a}{b}(e^{b\frac{i}{N}})\right]^2 \Rightarrow \min_{a,b}. \qquad (5)$$

Приравнивая нулю производные от Q по a и b, получим систему уравнений для определения оценок $\hat{a}$ и $\hat{b}$:

$$\begin{cases}\frac{\hat{a}}{\hat{b}}\frac{1}{N}\sum_{i=1}^{N}(e^{\hat{b}\frac{i}{N}}-1)^2=\frac{1}{N}\sum_{i=1}^{N}(e^{\hat{b}\frac{i}{N}}-1)\frac{t_i}{N},\\ \frac{\hat{a}}{\hat{b}}\frac{1}{N}\sum_{i=1}^{N}\frac{i}{N}e^{\hat{b}\frac{i}{N}}(e^{\hat{b}\frac{i}{N}}-1)=\frac{1}{N}\sum_{i=1}^{N}e^{\hat{b}\frac{i}{N}}\frac{i}{N}\frac{t_i}{N}.\end{cases} \qquad (6)$$

Найти отсюда оценки $\hat{a}$ и $\hat{b}$ в явном виде не представляется возможным. Однако исследуем асимптотические при $N\to\infty$ свойства оценок $\hat{a}$ и $\hat{b}$. Для этого представим $\hat{a}$ и $\hat{b}$ в виде $\hat{a}=a+\Delta\hat{a}$, $\hat{b}=b+\Delta\hat{b}$, где a и b – истинные значения параметров, а величины t_i/N $-\frac{t_i}{N}=T\left(\frac{i}{N}\right)+\frac{\Delta t_i}{N}$. Разлагая левые и правые части (6) в ряд по $\Delta\hat{a}$ и $\Delta\hat{b}$ и ограничиваясь линейными слагаемыми относительно $\Delta\hat{a}$ и $\Delta\hat{b}$, получим

$$\left[\frac{1}{b}\frac{1}{N}\sum_{i=1}^{N}(e^{b\frac{i}{N}}-1)^2\right]\Delta\hat{a}+$$

$$+\left[\frac{a}{b}\frac{1}{N}\sum_{i=1}^{N}\frac{i}{N}e^{b\frac{i}{N}}(e^{b\frac{i}{N}}-1)^2-\frac{a}{b^2}\frac{1}{N}\sum_{i=1}^{N}(e^{b\frac{i}{N}}-1)^2\right]\Delta\hat{b}=$$

$$=\frac{1}{N}\sum_{i=1}^{N}(e^{b\frac{i}{N}}-1)^2\frac{\Delta t_i}{N},$$

$$\left[\frac{1}{b}\frac{1}{N}\sum_{i=1}^{N}\frac{i}{N}e^{b\frac{i}{N}}(e^{b\frac{i}{N}}-1)\right]\Delta\hat{a}+\left[\frac{a}{b}\frac{1}{N}\sum_{i=1}^{N}\left(\frac{i}{N}\right)^2e^{2b\frac{i}{N}}-\right.$$

$$\left.-\frac{a}{b^2}\frac{1}{N}\sum_{i=1}^{N}\frac{i}{N}e^{b\frac{i}{N}}(e^{b\frac{i}{N}}-1)\right]\Delta\hat{b}=\frac{1}{N}\sum_{i=1}^{N}\frac{i}{N}e^{b\frac{i}{N}}\frac{\Delta t_i}{N}. \qquad (7)$$

При достаточно больших N суммы, стоящие в левых частях (7), можно заменить интегралами. Тогда коэффициенты при $\Delta\hat{a}$ и $\Delta\hat{b}$ можно вычислить:

$$a_{11}=\frac{1}{b}\frac{1}{N}\sum_{i=1}^{N}(e^{b\frac{i}{N}}-1)^2\xrightarrow[N\to\infty]{}$$

$$\xrightarrow[N\to\infty]{}\int_0^1(e^{bx}-1)^2dx=\frac{e^{2b}-4e^b+2b+3}{2b^2},$$

$$a_{12}=\frac{a}{b}\frac{1}{N}\sum_{i=1}^{N}\frac{i}{N}e^{b\frac{i}{N}}(e^{b\frac{i}{N}}-1)-$$

$$-\frac{a}{b^2}\frac{1}{N}\sum_{i=1}^{N}(e^{b\frac{i}{N}}-1)^2\xrightarrow[N\to\infty]{}$$

$$-\frac{a}{b^2}\frac{1}{N}\sum_{i=1}^{N}(e^{b\frac{i}{N}}-1)^2\xrightarrow[N\to\infty]{}$$

$$\xrightarrow[N\to\infty]{}\frac{a}{b}\int_0^1xe^{bx}(e^{bx}-1)dx-\frac{a}{b^2}\int_0^1(e^{bx}-1)^2dx=$$

$$=\frac{a}{4b^3}[(2b-3)e^{2b}-4(b-3)e^b-4b-9],$$

$$a_{21}=\frac{1}{b}\frac{1}{N}\sum_{i=1}^{N}\frac{i}{N}e^{b\frac{i}{N}}(e^{b\frac{i}{N}}-1)\xrightarrow[N\to\infty]{}$$

$$\xrightarrow[N\to\infty]{}\frac{1}{b}\int_0^1xe^{bx}(e^{bx}-1)dx=$$

$$=\frac{1}{4b^3}[(2b-1)e^{2b}-4(b-1)e^b-3],$$

$$a_{22}=\frac{a}{b}\frac{1}{N}\sum_{i=1}^{N}\left(\frac{i}{N}\right)^2e^{2b\frac{i}{N}}-$$

$$-\frac{a}{b^2}\frac{1}{N}\sum_{i=1}^{N}\frac{i}{N}e^{b\frac{i}{N}}(e^{b\frac{i}{N}}-1)\xrightarrow[N\to\infty]{}\frac{a}{b}\int_0^1x^2e^{2bx}dx-$$

$$-\frac{a}{b^2}\int_0^1xe^{bx}(e^{bx}-1)dx=\frac{a}{2b^4}[(b-1)e^b+1]^2. \qquad (8)$$

Разрешая (7) относительно $\Delta\hat{a}$ и $\Delta\hat{b}$, имеем

$$\Delta\hat{a}=\frac{1}{\alpha_{11}\alpha_{22}-\alpha_{12}\alpha_{21}}\left[\alpha_{22}\frac{1}{N}\sum_{i=1}^{N}(e^{b\frac{i}{N}}-1)\frac{\Delta t_i}{N}-\right.$$

$$\left.-\alpha_{12}\frac{1}{N}\sum_{i=1}^{N}\frac{i}{N}e^{b\frac{i}{N}}\frac{\Delta t_i}{N}\right],$$

$$\Delta\hat{b}=\frac{1}{\alpha_{11}\alpha_{22}-\alpha_{12}\alpha_{21}}\left[-\alpha_{21}\frac{1}{N}\sum_{i=1}^{N}(e^{b\frac{i}{N}}-1)\frac{\Delta t_i}{N}+\right.$$

$$\left.+\alpha_{11}\frac{1}{N}\sum_{i=1}^{N}\frac{i}{N}e^{b\frac{i}{N}}\frac{\Delta t_i}{N}\right]. \qquad (9)$$

Так как $\overline{\Delta t_i}=0$, то оценки $\hat{a}$ и $\hat{b}$ – несмещённые. Для вычисления дисперсий оценок $\hat{a}$ и $\hat{b}$ возведём равенства (9) в квадрат и усредним по t_i, а для получения ковариации оценок перемножим равенства (9) и усредним по t_i. В результате получим:

$$ND\{\hat{a}\}=\frac{1}{(\alpha_{11}\alpha_{22}-\alpha_{12}\alpha_{21})^2}\frac{1}{N^2}\times$$

$$\times\sum_{i,j=1}^{N}\left[\alpha_{22}^2(e^{b\frac{i}{N}}-1)(e^{b\frac{j}{N}}-1)-2\alpha_{12}\alpha_{22}(e^{b\frac{i}{N}}-1)\frac{j}{N}e^{b\frac{j}{N}}+\right.$$

$$\left.+\alpha_{12}^2\frac{i}{N}e^{b\frac{i}{N}}\frac{j}{N}e^{b\frac{j}{N}}\right]N\operatorname{cov}\left(\frac{t_i}{N},\frac{t_j}{N}\right),$$

$$ND\{\hat{b}\}=\frac{1}{(\alpha_{11}\alpha_{22}-\alpha_{12}\alpha_{21})^2}\frac{1}{N^2}\times$$

$$\times\sum_{i,j=1}^{N}\left[\alpha_{21}^2(e^{b\frac{i}{N}}-1)(e^{b\frac{j}{N}}-1)-2\alpha_{11}\alpha_{21}(e^{b\frac{i}{N}}-1)\frac{j}{N}e^{b\frac{j}{N}}+\right.$$

$$\left.+\alpha_{11}^2\frac{i}{N}e^{b\frac{i}{N}}\frac{j}{N}e^{b\frac{j}{N}}\right]N\operatorname{cov}\left(\frac{t_i}{N},\frac{t_j}{N}\right),$$

$$N\operatorname{cov}(\hat{a},\hat{b})=\frac{1}{(\alpha_{11}\alpha_{22}-\alpha_{12}\alpha_{21})^2}\frac{1}{N^2}\times$$

$$\times\sum_{i,j=1}^{N}\left[-\alpha_{21}\alpha_{22}(e^{b\frac{i}{N}}-1)(e^{b\frac{j}{N}}-1)+\right.$$

$$+(\alpha_{11}\alpha_{22}+\alpha_{12}\alpha_{21})(e^{b\frac{i}{N}}-1)\frac{j}{N}e^{b\frac{j}{N}}-$$

$$-\alpha_{11}\alpha_{12}\frac{i}{N}e^{b\frac{i}{N}}\frac{j}{N}e^{b\frac{j}{N}}\Bigg]N\operatorname{cov}\left(\frac{t_i}{N},\frac{t_j}{N}\right), \qquad (10)$$

где $\operatorname{cov}\left(\frac{t_i}{N},\frac{t_j}{N}\right)$ определяется формулой

$$N\operatorname{cov}\left(\frac{t_i}{N},\frac{t_j}{N}\right)=\frac{\min(i,j)}{\lambda\left(T\left(\frac{i}{N}\right)\right)\lambda\left(T\left(\frac{j}{N}\right)\right)}.$$

Для рассматриваемого случая

$$N\operatorname{cov}\left(\frac{t_i}{N},\frac{t_j}{N}\right)=a^2e^{b\left(\frac{i}{N}+\frac{j}{N}\right)}\min(i,j).$$

При больших значениях N в формулах (10) суммы можно заменить интегралами, которые нетрудно вычислить, и окончательно получим

$$ND\{\hat{a}\}\sim\frac{\alpha_{22}^2R_1-2\alpha_{12}\alpha_{22}R_2+\alpha_{12}^2R_3}{(\alpha_{11}\alpha_{22}-\alpha_{12}\alpha_{21})^2}=a^2f_a,$$

$$ND\{\hat{b}\}\sim\frac{\alpha_{21}^2R_1-2\alpha_{11}\alpha_{21}R_2+\alpha_{11}^2R_3}{(\alpha_{11}\alpha_{22}-\alpha_{12}\alpha_{21})^2}=a^2f_b,$$

$$N\operatorname{cov}(\hat{a},\hat{b})\sim(-\alpha_{21}\alpha_{22}R_1+$$
$$+(\alpha_{12}\alpha_{21}+\alpha_{11}\alpha_{22})R_2-\alpha_{11}\alpha_{12}R_3)/$$
$$/((\alpha_{11}\alpha_{22}-\alpha_{12}\alpha_{21})^2)=a^2f_{ab}, \qquad (11)$$

где α_{11}, α_{12}, α_{21}, α_{22} определяются формулами (8), а

$$R_1=\frac{a^2}{2b^3}\left(\frac{4b-3}{8}e^{4b}-\frac{6b-7}{3}e^{3b}+\right.$$
$$\left.+\frac{4b-9}{2}e^{2b}+3e^b-\frac{11}{24}\right),$$

$$R_2=\frac{a^2}{4b^4}\left(\frac{16b^2-20b+9}{16}e^{4b}-\right.$$
$$\left.-\frac{18b^2-30b+13}{9}e^{3b}-\frac{6b-1}{4}e^{2b}+e^b-\frac{53}{144}\right),$$

$$R_3=\frac{a^2}{8b^5}\left(\frac{32b^3-56b^2+44b-11}{16}e^{4b}-\right.$$
$$\left.-(2b-1)e^{2b}-\frac{5}{16}\right).$$

Найдём эффективность МНК-оценок параметров неизвестной интенсивности a и b. Для N-планов наблюдений запишем совместную плотность вероятностей моментов появления событий:

$$p(t_1,t_2,\ldots,t_N)=e^{-\Lambda(t_N)}\prod_{i=1}^{N}\lambda(t_i), \qquad (12)$$

где для гиперболического тренда $\lambda(t_i)=\dfrac{1}{a+b\frac{t_i}{N}}$, $\Lambda(t_N)=-\ln a+\ln\left(a+b\frac{t_N}{N}\right)$. Функция правдоподобия для оценок неизвестных параметров a и b имеет вид

$$L(a,b)=\ln p(t_1,t_2,\ldots,t_N)=-\frac{N}{b}\ln\left(a+b\frac{t_N}{N}\right)+$$
$$+\frac{N}{b}\ln a-\sum_{i=1}^{N}\ln\left(a+b\frac{t_i}{N}\right). \qquad (13)$$

Для нахождения эффективности найдём информационную матрицу Фишера. Так как элементы информационной матрицы $I_{aa}=-M\left\{\frac{\partial^2L}{\partial a^2}\right\}$, $I_{ab}=I_{ba}=-M\left\{\frac{\partial^2L}{\partial a\partial b}\right\}$, $I_{bb}=-M\left\{\frac{\partial^2L}{\partial b^2}\right\}$, то, вычисляя соответствующие производные от функции правдоподобия, получим

$$I_{aa}=-\frac{N}{b}\left\langle\frac{1}{(a+bt_N/N)^2}\right\rangle+\frac{N}{a^2b}-\sum_{i=1}^{N}\left\langle\frac{1}{(a+bt_i/N)^2}\right\rangle,$$

$$I_{ab}=-\frac{N}{b^2}\left\langle\frac{1}{a+bt_N/N}\right\rangle-\frac{N}{b}\left\langle\frac{t_N/N}{(a+bt_N/N)^2}\right\rangle+$$
$$+\frac{N}{ab^2}-\sum_{i=1}^{N}\left\langle\frac{t_i/N}{(a+bt_i/N)^2}\right\rangle,$$

$$I_{bb}=\frac{2N}{b^3}\left\langle\ln\left(a+b\frac{t_N}{N}\right)\right\rangle-2\frac{N}{b^2}\left\langle\frac{t_N/N}{a+bt_N/N}\right\rangle+\frac{N}{b}\left\langle\frac{(t_N/N)^2}{(a+bt_N/N)^2}\right\rangle-\frac{2N}{b^3}\ln a-\sum_{i=1}^{N}\left\langle\frac{(t_i/N)^2}{(a+bt_i/N)^2}\right\rangle, \qquad (14)$$

где угловые скобки означают усреднение по моментам наступления событий. Найдём средние значения величин в (14). Так как

$$p(t_i)=\frac{1}{(i-1)!}\left[N\Lambda\left(\frac{t_i}{N}\right)\right]^{i-1}\exp\left\{-N\Lambda\left(\frac{t_i}{N}\right)\right\}\lambda\left(\frac{t_i}{N}\right),$$

то $\left\langle\dfrac{1}{(a+bt_i/N)^2}\right\rangle=\left\langle\lambda^2\left(\dfrac{t_i}{N}\right)\right\rangle=\dfrac{N!}{(i-1)!}\displaystyle\int_0^\infty\frac{\lambda^3(u)}{\Lambda(u)}\times$

$$\times\exp\left\{N\left[\frac{i}{N}\ln\Lambda(u)-\Lambda(u)\right]\right\}du. \qquad (15)$$

Для вычисления интеграла используем метод перевала [3], по которому для унимодальной функции $f(z)$ с максимумом в точке $\xi\in(a,b)$ интеграл асимптотически при $\mu\to\infty$ равен

$$F(\mu)=\int_a^b\varphi(z)e^{\mu f(z)}dz\sim\varphi(\xi)e^{\mu f(\xi)}\sqrt{-\frac{2\pi}{\mu f''(\xi)}}. \qquad (16)$$

Для интеграла в (15) $\varphi(z)=\dfrac{\lambda^3(z)}{\Lambda(z)}$, $f(z)=\dfrac{i}{N}\ln\Lambda(z)-\Lambda(z)$, а стационарная точка ξ, определяемая из условия $f'(z)=0$, является решением уравнения $\Lambda(\xi)=\dfrac{i}{N}$. Тогда по формуле (16)

$$\left\langle\frac{1}{(a+bt_i/N)^2}\right\rangle=\lambda^2\left(T\left(\frac{i}{N}\right)\right)=\frac{1}{a^2}e^{-2b\frac{i}{N}}. \qquad (17)$$

Аналогично можно показать, что

$$\left\langle\frac{t_i/N}{(a+bt_i/N)^2}\right\rangle=T\left(\frac{i}{N}\right)\lambda^2\left(T\left(\frac{i}{N}\right)\right)=$$
$$=\frac{1}{ab}\left(e^{-b\frac{i}{N}}-e^{-2b\frac{i}{N}}\right), \qquad (18)$$

$$\left\langle \frac{(t_i/N)^2}{(a+bt_i/N)^2} \right\rangle = T^2\left(\frac{i}{N}\right)\lambda^2\left(T\left(\frac{i}{N}\right)\right) =$$

$$= \frac{1}{b^2}(e^{b\frac{t}{N}}-1)^2 e^{-2b\frac{t}{N}}. \qquad (19)$$

Подставляя (17)–(19) в формулы (14), заменяя суммы интегралами и приводя подобные, получим

$$I_{aa} = \frac{N}{a^2}\frac{1-e^{-2b}}{2b} = \frac{N}{a^2}F_{aa},$$

$$I_{ab} = \frac{N}{a}\frac{(1-e^{-b})^2}{2b^2} = \frac{N}{a}F_{ab},$$

$$I_{bb} = N\frac{1}{b^2}\left[1-2\frac{1-e^{-b}}{b}+\frac{1-e^{-2b}}{2b}\right] = NF_{bb}, \qquad (20)$$

где $\quad F_{aa} = \frac{1}{2b}(1-e^{-2b}), \quad F_{ab} = \frac{1}{2b^2}(1-e^{-b})^2,$

$$F_{bb} = \frac{1}{b^2}\left[1-2\frac{1}{b}(1-e^{-b})+\frac{1}{2b}(1-e^{-2b})\right]. \qquad (21)$$

Находя матрицу, обратную информационной, получаем, что для дисперсий V_a, V_b и ковариации V_{ab} справедливы формулы

$$\frac{NV_a}{a^2} = \frac{F_{bb}}{F_{aa}F_{bb}-F_{ab}^2}, \quad NV_b = \frac{F_{aa}}{F_{aa}F_{bb}-F_{ab}^2},$$

$$\frac{NV_{ab}}{a} = -\frac{F_{ab}}{F_{aa}F_{bb}-F_{ab}^2}. \qquad (22)$$

Приступим теперь к нахождению основной характеристики выделения тренда – интегральной погрешности. Интегральную погрешность $\mathbf{D}$ для N-планов будем определять по формуле

$$\mathrm{D} = M\left\{\frac{1}{t_N}\int_0^{t_N}\left(\frac{1}{\hat{a}+\hat{b}t/N}-\frac{1}{a+bt/N}\right)^2 dt\right\}. \qquad (23)$$

Представляя оценки $\hat{a}$ и $\hat{b}$ в виде $\hat{a} = a+\Delta a$, $\hat{b} = b+\Delta b$, где a и b – истинные значения параметров, можем записать

$$\frac{1}{\hat{a}+\hat{b}t/N} = \frac{1}{a+bt/N+(\Delta a+\Delta bt/N)} =$$

$$= \frac{1}{a+bt/N}\left(1-\frac{\Delta a+\Delta bt/N}{a+bt/N}+\cdots\right).$$

Ограничиваясь в полученном соотношении линейными слагаемыми относительно Δa и Δb, имеем

$$\left(\frac{1}{\hat{a}+\hat{b}t/N}-\frac{1}{a+bt/N}\right)^2 =$$

$$= \frac{\Delta a^2+2\Delta a\Delta bt/N+\Delta b^2t^2/N^2}{(a+bt/N)^4},$$

$$M\left\{\left(\frac{1}{\hat{a}+\hat{b}t/N}-\frac{1}{a+bt/N}\right)^2\right\} =$$

$$= \frac{D\{\hat{a}\}+2\operatorname{cov}(\hat{a},\hat{b})t/N+D\{\hat{b}\}t^2/N^2}{(a+bt/N)^4}.$$

При $N\to\infty$ величину t_N в пределе интеграла можно заменить её математическим ожиданием $M\{t_N\} = N\frac{a}{b}(e^b-1)$. Поэтому

$$M\left\{\int_0^{t_N}\left(\frac{1}{\hat{a}+\hat{b}t/N}-\frac{1}{a+bt/N}\right)^2 dt\right\} =$$

$$= \int_0^{N\frac{a}{b}(e^b-1)} \frac{D\{\hat{a}\}+2\operatorname{cov}(\hat{a},\hat{b})\frac{t}{N}+D\{\hat{b}\}t^2/N^2}{(a+bt/N)^4}dt.$$

Делая замену переменных в интеграле и обозначая

$$\frac{ND\{\hat{a}\}}{a^2} = A,\ ND\{\hat{b}\} = B,\ \frac{N\operatorname{cov}(\hat{a},\hat{b})}{a} = C, \qquad (24)$$

получим, что интегральная погрешность

$$\mathbf{D} = \frac{1}{Na^2}\frac{b}{e^b-1}\left[\frac{A}{b}\int_0^{e^b-1}\frac{du}{(1+u)^4}+\right.$$

$$\left.+\frac{2C}{b^2}\int_0^{e^b-1}\frac{u\,du}{(1+u)^4}+\frac{B}{b^3}\int_0^{e^b-1}\frac{u^2\,du}{(1+u)^4}\right]. \qquad (25)$$

Вычисляя входящие в (25) интегралы, получим:

$$Na^2\mathbf{D} = \frac{b}{e^b-1}\left[A\frac{1-e^{-3b}}{3b}+\right.$$

$$\left.+C\frac{1-3e^{-2b}+2e^{-3b}}{b^2}+B\frac{(1-e^{-b})^3}{3b^3}\right]. \qquad (26)$$

Эта формула для интегральной погрешности позволяет сравнивать между собой различные методы оценок параметров a и b гиперболического тренда пуассоновского потока.

Минимальную интегральную погрешность выделения тренда дают оценки по методу максимального правдоподобия (МП-оценки). Запишем (25) для МП-оценок:

$$Na^2\mathbf{D}_{\text{МП}} = \frac{b}{e^b-1}\frac{1}{F_{aa}F_{bb}-F_{ab}^2}\left[\frac{1-e^{-3b}}{3b}F_{bb}-\right.$$

$$\left.-\frac{1-3e^{-2b}+2e^{-3b}}{3b^2}F_{ab}+\frac{(1-e^{-b})^3}{3b^3}F_{aa}\right], \qquad (27)$$

где F_{aa}, F_{ab}, F_{bb} определяются выражениями (21). Для МНК-оценок формула интегральной погрешности

$$Na^2\mathbf{D}_{\text{МНК}} = \frac{b}{e^b-1}\left[\frac{1-e^{-3b}}{3b}f_a+\right.$$

$$\left.+\frac{1-3e^{-2b}+2e^{-3b}}{3b^2}f_{ab}+\frac{(1-e^{-b})^3}{3b^3}f_b\right], \qquad (28)$$

где функции f_a, f_{ab}, f_b определяются формулами (11). Выражения для интегральной погрешности позволяют записать наряду с эффективностями МНК-оценок параметров a и b

$$\mathrm{eff}_a = \frac{F_{bb}}{f_a[F_{aa}F_{bb}-F_{ab}{}^2]}, \qquad (29)$$

$$\mathrm{eff}_b = \frac{F_{aa}}{f_b[F_{aa}F_{bb}-F_{ab}{}^2]} \qquad (30)$$

эффективность МНК-оценок гиперболического тренда

$$\mathrm{eff} = \frac{\mathbf{D}_{\text{Мп}}}{\mathbf{D}_{\text{Мнк}}} = \frac{1}{F_{aa}F_{bb}-F_{ab}^2}(b^2(1-e^{-3b})F_{bb}-(1-3e^{-2b}+$$

$$+2e^{-3b})F_{ab}+(1-e^{-b})^3F_{aa})/(b^2(1-e^{-3b})f_a+$$

$$+b(1-3e^{-2b}+2e^{-3b})f_{ab}+(1-e^{-b})^3 f_b). \quad (31)$$

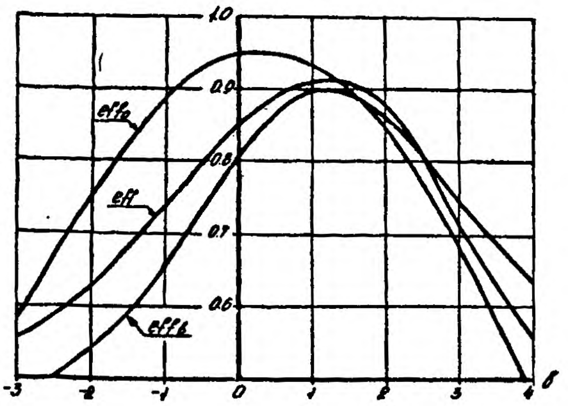

Рис. 1

Из приведённых на рис. 1 графиков зависимости eff_a, eff_b, eff от b следует, что эффективность МНК-оценки параметра a достаточно высока. При $-0{,}9<b<1{,}5$ эффективность не менее 0,9. Эффективность параметра b несколько ниже, чем эффективность параметра a. Так, значения 0,9 она достигает при $0{,}6<b<1{,}7$, а при увеличении или уменьшении b эффективность МНК-оценки параметра b, как и параметра a, убывает. Эффективность МНК-оценки тренда достигает своего наибольшего значения 0,91 при $b=1{,}4$, а с уменьшением или увеличением b она убывает.

ЛИТЕРАТУРА

1. *Кокс Д., Льюис П.* Статистический анализ последовательностей событий. М.: Мир, 1969. 312 с.
2. *Тривоженко Б.Е.* Асимптотические свойства моментов наступления событий нестационарного пуассоновского потока // Математическое моделирование. Кибернетика. Информатика. Томск: Изд-во Том. ун-та, 1999. С. 140–150.
3. *Свешников А.Г., Тихонов А.Н.* Теория функций комплексной переменной. М.: Наука, 1979. 319 с.

Статья представлена кафедрой теории вероятностей и математической статистики факультета прикладной математики и кибернетики Томского государственного университета, поступила в научную редакцию 14 марта 2000 г.

УДК 519.24

С.С. Тарима

СТАТИСТИЧЕСКОЕ ОЦЕНИВАНИЕ ВЕРОЯТНОСТЕЙ С ПРИВЛЕЧЕНИЕМ ДОПОЛНИТЕЛЬНОЙ ИНФОРМАЦИИ

Рассматривается способ учета дополнительной информации о вероятностях одних событий для улучшения эмпирических оценок вероятностей других. Приведен пример такого учета.

1. Пусть $X_1,\cdots,X_n$ – последовательность независимых одинаково распределенных случайных величин, принимающих значения из измеримого пространства (X,B) с неизвестным распределением P. Рассмотрим события $A, B_1,\cdots, B_s$, принадлежащие σ-алгебре B. Предположим, что $P(B_i)=b_i$, где b_i известны ($i=1,\cdots, s$). Рассмотрим задачу статистического оценивания $P(A)$ с учетом данной информации. Результаты наблюдений представим в виде протокола:

	A	B_1	$B_j\ldots$	B_s
X_1	$I_A(X_1)$	$I_{B1}(X_1)$	$I_{Bj}(X_1)\ldots$	$I_{Bs}(X_1)$
X_2	$I_A(X_2)$	$I_{B1}(X_2)$	$I_{Bj}(X_2)\ldots$	$I_{Bs}(X_2)$
…	…	…	… …	…
X_i	$I_A(X_i)$	$I_{B1}(X_i)$	$I_{Bj}(X_i)\ldots$	$I_{Bs}(X_i)$
…	…	…	… …	…
X_n	$I_A(X_n)$	$I_{B1}(X_n)$	$I_{Bj}(X_n)\ldots$	$I_{Bs}(X_n)$

В клетках протокола наблюдений $I_D(X_i)$ – индикатор некоторого события D из σ-алгебры B.

Рассмотрим события $\Lambda_j=U_1\cap U_2\cap\cdots\cap U_s$, где U_i либо B_i, либо его дополнение, $i=1,\cdots, s$. Их совокупность, с возможным количеством до 2^s событий, является разбиением пространства X. В этом случае имеют место равенства

$$\sum_{\Lambda_i\in B_j} P(\Lambda_i)=b_j,\ \ j=1,\cdots,s. \quad (1)$$

2. Для случая известных значений $P(\Lambda_i)$ в [3] была предложена оценка

$$P_n^*(A)=\sum_{i=1}^{2^s} P(\Lambda_i)\frac{P_n(A\Lambda_i)}{P_n(\Lambda_i)},\ \ P_n(\Lambda_i)>0, \quad (2)$$

где
$$P_n(A)=\frac{1}{n}\sum_{i=1}^{n} I_A(X_i).$$

В протоколе наблюдений часто реализуются не все события Λ_i. Предположим, что реализовалось только k из 2^s событий Λ_i, для которых $P_n(\Lambda_i)>0$. Остальные события в формуле (2) при рассмотрении отношения дают неопределенность типа 0/0 (если $P_n(\Lambda_i)=0$, то и $P_n(A\Lambda_i)=0$). Аналогично, из событий $B_1,\cdots, B_s$ также реализуются не все, а только часть, допустим $m\le k$. Для этой ситуации рассмотрим следующую оценку

$$P_n^{**}(A)=\frac{\sum\limits_{i=1,P_n(\Lambda_i)>0}^{2^s} P(\Lambda_i)\frac{P_n(A\Lambda_i)}{P_n(\Lambda_i)}}{\sum\limits_{i=1,P_n(\Lambda_j)>0}^{2^s} P(\Lambda_j)}. \quad (3)$$

Суммирование ведется только по тем индексам, для которых эмпирическая оценка события Λ_i не равна нолю. Эта оценка в отличие от (2) определена при любых $P_n(\Lambda_i)$.

3. В нашем случае значения $P(\Lambda_i)$ неизвестны, но удовлетворяют условиям (1). Осталось оценить k вероятностей $P(\Lambda_i)$. Если оценивать $P(\Lambda_i)$ распределением $P_n(\Lambda_i)$, то из (3) мы получаем $P^{**}_n(A)=P_n(A)$.

Найдем проекцию [ближайшую в смысле информационного расхождения (4)] $P^{**}_n(A)$ в класс, удовлетворя-