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

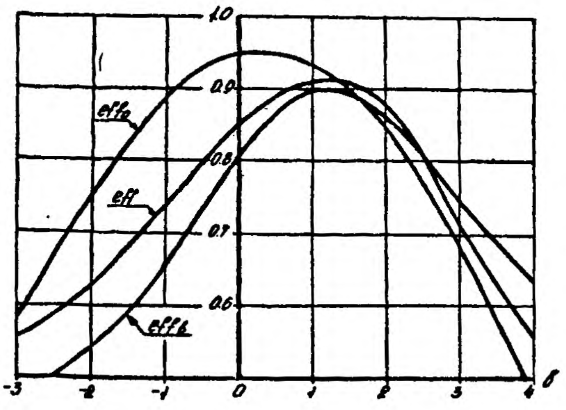

Рис. 1

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

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

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

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

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

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

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, удовлетворяющий условию (1). Информационное расхождение между $\{P(\Lambda_i)\}$ и $\{P_n(\Lambda_i)\}$ будет следующим:

$$d_{kl}(P(\Lambda),P_n(\Lambda))=\sum_{i=1,P_n(\Lambda_i)>0}^{2^s} P(\Lambda_i)\ln\frac{P_n(\Lambda_i)}{P(\Lambda_i)}. \qquad (4)$$

Заметим, что суммирование ведется только по положительным $P_n(\Lambda_i)$ вследствие необходимости абсолютной непрерывности меры P по отношению к мере P_n. Оценим $P(\Lambda_i)$ из условия минимума информационного расхождения при условии (1) и условии нормировки, которое в нашем случае определяется как:

$$\sum_{i=1,P_n(\Lambda_i)>0}^{2^s} P(\Lambda_i)=1. \qquad (5)$$

Нахождение $P(A)$ свелось к задаче нелинейного программирования. Пусть решение поставленной задачи нелинейного программирования будет $\{P^*(\Lambda_i);\ i=1,\cdots,2^s\}$; подставив это решение в (3), получим

$$P_n^{***}(A)=\sum_{i=1,P_n(\Lambda_i)>0}^{2^s} P^*(\Lambda_i)\frac{P_n(A\Lambda_i)}{P_n(\Lambda_i)}. \qquad (6)$$

4. Рассмотрим случай, где решение задачи нелинейного программирования находится в явном виде. При $P(B_j)=b_j$, где $j=1,2$, получим:

$$P(\Lambda_1)=P(B_1\cap B_2);\ P(\Lambda_2)=P(B_1\cap(\Omega\backslash B_2));$$
$$P(\Lambda_3)=P((\Omega\backslash B_1)\cap B_2);\ P(\Lambda_4)=P((\Omega\backslash B_1)\cap(\Omega\backslash B_2));$$
$$P(A)=P(B_1B_2)=?.$$

Если наблюдения распределились по всем четырем множествам, то на вероятности $P(\Lambda_i)$ наложены три ограничения:

$$P(\Lambda_1)+P(\Lambda_2)=b_1;\ P(\Lambda_1)+P(\Lambda_3)=b_2;$$
$$P(\Lambda_1)+P(\Lambda_2)+P(\Lambda_3)+P(\Lambda_4)=1.$$

Определив одну из вероятностей $P(\Lambda_i)$ из наложенных ограничений, можно найти остальные. Оценим $P(\Lambda_1)$ из условия минимума информационного расхождения, которое в рассматриваемом случае примет вид

$$d_{kl}(P(\Lambda_i),P_n(\Lambda_i))=\sum_{i=1}^{4}P(\Lambda_i)\ln\frac{P_n(\Lambda_i)}{P(\Lambda_i)}=$$
$$=P(\Lambda_1)\ln\frac{P_n(\Lambda_1)}{P(\Lambda_1)}+(b_2-P(\Lambda_1))\times$$
$$\times\ln\frac{P_n(\Lambda_2)}{b_2-P(\Lambda_1)}+(b_1-P(\Lambda_1))\times$$
$$\times\ln\frac{P_n(\Lambda_3)}{b_1-P(\Lambda_1)}+(1-b_2-b_1+P(\Lambda_1))\times$$
$$\times\ln\frac{P_n(\Lambda_4)}{1-b_2-b_1+P(\Lambda_1)}. \qquad (7)$$

Минимизация (7) по $P(\Lambda_1)$ приводит к уравнению

$$\left(1-\frac{P_n(\Lambda_1)P_n(\Lambda_4)}{P_n(\Lambda_2)P_n(\Lambda_3)}\right)P^2(\Lambda_1)+$$
$$+\left(\begin{array}{l}1-b_1-b_2+b_1\dfrac{P_n(\Lambda_1)P_n(\Lambda_4)}{P_n(\Lambda_2)P_n(\Lambda_3)}+\\ +b_2\dfrac{P_n(\Lambda_1)P_n(\Lambda_4)}{P_n(\Lambda_2)P_n(\Lambda_3)}\end{array}\right)\times$$
$$\times P(\Lambda_1)-b_1b_2\frac{P_n(\Lambda_1)P_n(\Lambda_4)}{P_n(\Lambda_2)P_n(\Lambda_3)}=0. \qquad (8)$$

Решение уравнения (8) следующее:

$$P_n^*(\Lambda_1)=\left(\frac{b_1+b_2}{2}+\frac{t}{2(1-t)}\right)\pm$$
$$\pm\sqrt{\left(\frac{b_1+b_2}{2}+\frac{t}{2(1-t)}\right)^2-\frac{b_1b_2}{(1-t)}}, \qquad (9)$$

где $t=\dfrac{P_n(\Lambda_2)\cdot P_n(\Lambda_3)}{P_n(\Lambda_1)\cdot P_n(\Lambda_4)}$. Это решение и есть оценка $P(\Lambda_1)$, которая однозначно определяет $P(\Lambda_2)$, $P(\Lambda_3)$, $P(\Lambda_4)$ из вышеупомянутых наложенных ограничений.

В выражении (9), когда $t<1$, используем «минус», а когда $t>1$ – «плюс». При $t=1$ (случай независимости событий B_1 и B_2 между собой) выражение (9) берем равным $b_1\times b_2$. График зависимости $P^*_n(\Lambda_1)$ от t будет следующий:

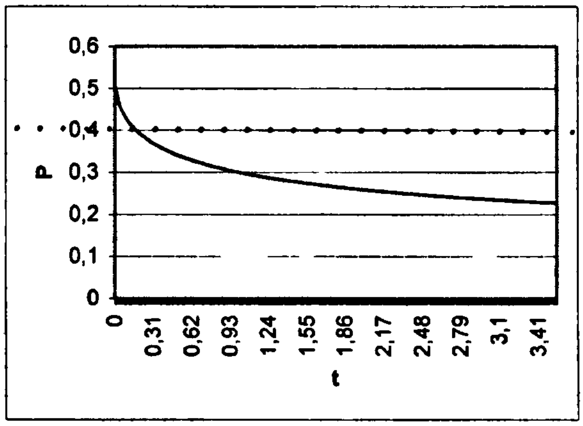

График приведен для $P(B_1)=0{,}5$ и $P(B_2)=0{,}6$

Оценка вероятности $P(B_1\,B_2)$ изменяется в границах от $\max\{0;b_1+b_2-1)\}$ до $\min\{b_1;b_2\}$, а эмпирическая оценка (по комплектным наблюдениям) искомой вероятности изменяется от 0 до 1. Остальные оценки вероятностей определяются из условия нормировки и двух дополнительных ограничений.

Если наблюдения не попали в одно из четырех множеств, то наши три ограничения однозначно фиксируют три оставшихся $P(\Lambda_i)$.

ЛИТЕРАТУРА

1. *Тарима С.С.* О свойствах эмпирического распределения, модифицированного с учетом знания вероятностей // Мат. моделирование и теория вероятностей / Под ред. И.А. Александрова, Томск: Пеленг, 1998. С. 257.
2. *Кульбак С.* Теория информации и статистики. М.: Наука, 1967. 408 с..
3. *Дмитриев Ю.Г., Устинов Ю.К.* Статистическое оценивание распределений вероятностей с использованием дополнительной информации. Томск: Изд-во ТГУ, 1988.

Статья представлена кафедрой теоретической кибернетики факультета прикладной математики и кибернетики Томского государственного университета, поступила в научную редакцию 14 февраля 2000 г.