ВЕСТНИК ТОМСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА
2009 Управление, вычислительная техника и информатика № 4(9)

УДК 519.2

Ю.Г. Дмитриев, А.А. Князева

ОЦЕНКИ ВЕРОЯТНОСТЕЙ СОБЫТИЙ ПРИ НАЛИЧИИ ДАННЫХ С ПРОПУСКАМИ

Рассматривается задача статистического оценивания вероятности произведения двух событий на основе комплектных и некомплектных наблюдений. Предлагаются оценки с привлечением дополнительной информации, содержащейся в некомплектных наблюдениях, и исследуются их свойства.

Ключевые слова: *комплектные и некомплектные наблюдения, дополнительная информация, адаптивные оценки вероятностей.*

В социологических [1] и маркетинговых исследованиях [2], при наблюдении объектов, характеризуемых многомерным вектором признаков, случаются пропуски в компонентах вектора, что приводит к некомплектным наблюдениям. Статистическое оценивание долей объектов с заданными значениями признаков представляет в такой ситуации важную научную и практическую задачу. В статистической практике известны следующие методы статистического анализа данных с пропусками [3,4]:

- исключение некомплектных наблюдений из рассмотрения и построение статистических выводов на основе полных (комплектных) данных;
- методы, основанные на моделировании (строится модель порождения пропусков, параметры модели оцениваются с помощью функции правдоподобия);
- восстановление пропусков;
- методы взвешивания (суть заключается в том, что каждое наблюдение выбирается в выборку с некоторой вероятностью);

Представляет интерес разработка методов статистического анализа данных с одновременным использованием как комплектных, так и некомплектных наблюдении с целью увеличения качества оценивания за счет привлечения дополнительной информации, содержащейся в некомплектных наблюдениях. Рассмотрение этой проблемы на примере оценивания вероятности событий по наблюдениям двумерного вектора признаков приводится в данной работе.

1. Постановка задачи

Пусть X и Y – случайные величины, заданные на вероятностном пространстве $(\Omega, \Im, \mathrm{P})$ и осуществляющие измеримое отображение $(\Omega, \Im)$ на пространство (R^2, Ξ^2). Среди всех наблюдений над парой (X, Y) имеется n пар (X_1, Y_1), $(X_2, Y_2),\ldots,(X_n, Y_n)$, для которых получены значения по обеим компонентам (далее будем именовать такие наблюдения комплектными), и имеются m наблюдений, в которых известны значения только второй компоненты, обозначим их Y_{n+1}, Y_{n+2}, … Y_{n+m} и назовем некомплектными. Предполагается, что все наблюдения Y_1, Y_2, … Y_n, Y_{n+1}, Y_{n+2}, … Y_{n+m}. независимы между собой. На основе этих данных требуется построить оценку вероятности $P(A\times B)$, где событие $A\subset R$, $B\subset R$. Как известно, несмещенной и наилучшей в смысле минимума дисперсии оценкой вероятности

$P(A\times B)$, построенной по комплектным наблюдениям, является эмпирическое распределение

$$P_n(A\times B)=\frac{1}{n}\sum_{i=1}^{n} I_A(X_i) I_B(Y_i)\,, \tag{1}$$

где $$I_A(X_i)=\begin{cases}1, & X_i\in A,\\ 0, & X_i\notin A,\end{cases} \qquad I_B(Y_i)=\begin{cases}1, & Y_i\in B,\\ 0, & Y_i\notin B,\end{cases}$$

являются индикаторными функциями соответствующих событий. Эмпирическое распределение (1) имеет математическое ожидание $EP_n(A\times B)=P_{XY}(A\times B)$ и дисперсию $DP_n(A\times B)=(1/n)\ P_{XY}(A\times B)(1-P_{XY}(A\times B))$. Рассмотрим задачу оценивания $P(A\times B)$, используя наряду с комплектными и некомплектные наблюдения с целью повышения точности оценки.

2. Построение оценок

2.1. Оценка с использованием формулы условной вероятности

Воспользуемся формулой умножения вероятностей $P(A\times B)=P(A|B)P(B)$ и запишем аналогичную формулу для эмпирических вероятностей: $P_n^*(A\times B)=P_n(A|B)P_{n+m}(B)=(P_n(A\times B)/P_n(B))\cdot P_{n+m}(B)$, где $P_{n+m}(B)=(n/(n+m))P_n(B)+$ $+(m/(n+m))P_m(B)$, а

$$\frac{P_n(A\times B)}{P_n(B)}=\begin{cases}0, & P_n(B)=0,\\ \dfrac{P_n(A\times B)}{P_n(B)}, & P_n(B)\neq 0.\end{cases} \tag{2}$$

С учетом (2) и того, что при $P_n(B)=0$ и $P_n(AB)=0$, имеем

$$P_{XY}^*(A\times B)=\begin{cases}0, & P_n(B)=0,\\ \dfrac{n}{n+m}P_n(A\times B)+\dfrac{m}{n+m}\dfrac{P_n(A\times B)}{P_n(B)}P_m(B), & P_n(B)\neq 0.\end{cases} \tag{3}$$

2.2. Оценка по методу коррелированных процессов

В соответствии с методом коррелированных процессов, рассмотрим класс оценок вида

$$P_{XY}^{\lambda}(A\times B)=P_n(A\times B)-\lambda(P_n(B)-P_{n+m}(B)), \tag{4}$$

где параметр λ выбирается из условия минимума дисперсии оценки и имеет вид

$$\lambda=\frac{E\{P_n(A\times B)(P_n(B)-P_{n+m}(B))\}}{E\{(P_n(B)-P_{n+m}(B))^2\}}=\frac{P_{XY}(A\times B)}{P_Y(B)}\,. \tag{5}$$

Поскольку вероятности $P_{XY}(A\times B)$ и $P_Y(B)$ неизвестны, то заменив их на эмпирические вероятности $P_n(A\times B)$ и $P_n(B)$, получим λ^*. Подставив это значение в (5), получим адаптивную оценку по методу коррелированных процессов

$$P_{XY}^{\lambda*}(A\times B)=\frac{n+m-m}{n+m}P_n(A\times B)+\frac{m}{n+m}\frac{P_n(A\times B)}{P_n(B)}P_m(B)\,.$$

Эта оценка совпадает с (3) при выполнении условия (2). Отметим также, что если взять в выражении (4) вместо $(P_n(B)-P_{n+m}(B))$ разность $(P_n(B)-P_m(B))$, то также придем к адаптивной оценке, совпадающей с (3).

3. Свойства оценки

Исследуем свойства оценки (3) в рамках схемы Бернулли. Найдем математическое ожидание

$$EP^*_{XY}(A\times B)=E\left\{\frac{n}{n+m}P_n(A\times B)+\frac{m}{n+m}\frac{P_n(A\times B)}{P_n(B)}P_m(B)\right\}=$$

$$=\frac{n}{n+m}P_{XY}(A\times B)+\frac{m}{n+m}P_Y(B)E\left(\frac{P_n(A\times B)}{P_n(B)}\right).$$

С учетом (2)

$$E\left(\frac{P_n(A\times B)}{P_n(B)}\right)=0+\sum_{j=1}^{n}P\left(P_n(B)=\frac{j}{n}\right)E\left(\frac{P_n(A\times B)}{P_n(B)}\mid P_n(B)=\frac{j}{n}\right)=$$

$$=P_{XY}(A\mid B)\sum_{j=1}^{n}P\left(P_n(B)=\frac{j}{n}\right)=P_{XY}(A\mid B)(1-P(P_n(B)=0)).$$

Следовательно,

$$EP^*_{XY}(A\times B)=\frac{n}{n+m}P_{XY}(A\times B)+\frac{m}{n+m}P_Y(B)P_{XY}(A\mid B)(1-P(P_n(B)=0))=$$

$$=P_{XY}(A\times B)-\frac{m}{n+m}P_Y(B)P_{XY}(A|B)P(P_n(B)=0)=$$

$$=P_{XY}(A\times B)-\frac{m}{n+m}P_{XY}(A\times B)P^n(\overline{B}),$$

где $\overline{B}$ – противоположное событие. Как видно, оценка (3) имеет смещение, равное $\Delta=EP^*_{XY}(A\times B)-P_{XY}(A\times B)=-(m/(n+m))P_{XY}(A\times B)P^n(\overline{B})$.

Найдем среднеквадратическое отклонение (СКО) оценки

$$SP^*_{XY}(A\times B)=E(P^*_{XY}(A\times B)-P_{XY}(A\times B))^2.$$

Имеем

$$SP^*_{XY}(A\times B)=EP^{*2}_{XY}(A\times B)-P^2_{XY}(A\times B)-2\Delta\cdot P_{XY}(A\times B)=$$

$$=\frac{n^2}{(n+m)^2}EP_n^{\,2}(A\times B)+2\frac{nm}{(n+m)^2}E\left(P_n(A\times B)\frac{P_n(A\times B)}{P_n(B)}P_m(B)\right)+$$

$$+\frac{m^2}{(n+m)^2}E\left(\frac{P_n(A\times B)}{P_n(B)}P_m(B)\right)^2-P^2_{XY}(A\times B)-2\Delta\cdot P_{XY}(A\times B).$$

Заметим, что

$$E(P_n^{*2}(A\times B))=P^2_{XY}(A\times B)+(1/n)P_{XY}(A\times B)(1-P_{XY}(A\times B)),$$

$$E\left(P_n(AB)\frac{P_n(A\times B)}{P_n(B)}P_m(B)\right)=P_Y(B)E\left(\frac{P_n^{\,2}(A\times B)}{P_n(B)}\right)=$$

$$=0+P_Y(B)\sum_{j=1}^{n}E\left(\frac{P_n^{\,2}(A\times B)}{P_n(B)}\mid P_n(B)=j/n\right)P(P_n(B)=j/n).$$

Для каждого $j = 1,\ldots,n$

$$E\left(\frac{P_n^{\,2}(A\times B)}{P_n(B)}\mid P_n(B)=\frac{j}{n}\right)=E\left(\frac{(1/n^2)\left(\sum_{i=1}^{n}I_{AB}(X_i,Y_i)\right)^2}{j/n}\mid P_n(B)=\frac{j}{n}\right)=$$

$$=\frac{1}{n}P_{XY}(A\mid B)+\frac{j-1}{n}P_{XY}^2(A\mid B)=\frac{1}{n}P_{XY}(A\mid B)(1-P_{XY}(A\mid B))+\frac{j}{n}P_{XY}^2(A\mid B).$$

С учетом этого

$$E\left(\frac{P_n^{\,2}(A\times B)}{P_n(B)}\right)=\frac{1}{n}\frac{P_{XY}(A\times B)}{P_Y(B)}\left(1-\frac{P_{XY}(A\times B)}{P_Y(B)}\right)\sum_{j=1}^{n}P(P_n(B)=j/n)+$$

$$+\frac{1}{n}\frac{P_{XY}^2(A\times B)}{P^2{}_Y(B)}\sum_{j=1}^{n}jP(P_n(B)=j/n).$$

Поскольку $\sum_{j=1}^{n}jP(P_n(B)=j/n)=nP_Y(B)$, то

$$E\left(P_n(A\times B)\frac{P_n(A\times B)}{P_n(B)}P_m(B)\right)=$$

$$=\frac{1}{n}P_{XY}(A\times B)\left(1-\frac{P_{XY}(A\times B)}{P_Y(B)}\right)(1-P(P_n(B)=0))+P_{XY}^2(A\times B).$$

Аналогично получаем

$$E\left(\frac{P_n(A\times B)}{P_n(B)}P_m(B)\right)^2=0+\sum_{j=1}^{n}E\left(\left(\frac{P_n(A\times B)}{P_n(B)}P_m(B)\right)^2\mid P_n(B)=\frac{j}{n}\right)P(P_n(B)=\frac{j}{n}).$$

Далее, при $j = 1,\ldots,n$

$$E\left(\left(\frac{P_n(A\times B)}{P_n(B)}P_m(B)\right)^2\mid P_n(B)=\frac{j}{n}\right)=$$

$$=E\left(\frac{(1/n^2)\left(\sum_{i=1}^{n}I_{AB}(X_i,Y_i)\right)^2}{j^2/n^2}\middle| P_n(B)=\frac{j}{n}\right)E\left(\frac{1}{m^2}\left(\sum_{i=n+1}^{n+m}I_B(Y_i)\right)^2\right)=$$

$$=\frac{P_{XY}(A\mid B)j+P_{XY}^2(A\mid B)j(j-1)}{j^2}\cdot\frac{mP_Y(B)+m(m-1)P_Y^2(B)}{m^2}=$$

$$=\frac{1}{j}P_{XY}(A\mid B)(1-P_{XY}(A\mid B))\left(\frac{1}{m}P_Y(B)+\frac{m-1}{m}P_Y^2(B)\right)+$$

$$+P^2{}_Y(A\mid B)\big((1/m)P_Y(B)+((m-1)/m)P_Y^2(B)\big).$$

С учетом этого получим

$$\sum_{j=1}^{n} P\left(P_n(B) = \frac{j}{n} \right) E\left(\left(\frac{P_n(A\times B)}{P_n(B)} P_m(B) \right)^2 \mid P_n(B) = \frac{j}{n} \right) =$$

$$= \frac{1}{m} P_{XY}(A\mid B)(1-P_{XY}(A\mid B))P_Y(B)(1-P_Y(B)) \sum_{j=1}^{n} \frac{1}{j} P\left(P_n(B) = \frac{j}{n} \right) +$$

$$+ P_{XY}(A\times B)(P_Y(B)-P_{XY}(A\times B)) \sum_{j=1}^{n} \frac{1}{j} P\left(P_n(B) = \frac{j}{n} \right) + P_{XY}^2(A\times B) +$$

$$+ \frac{1}{m} P_{XY}^2(A\mid B)P_Y(B)(1-P_Y(B)) - P_{XY}^2(A\times B)P(P_n(B)=0) -$$

$$- \frac{1}{m} P_Y(B)(1-P_Y(B))P_{XY}^2(A\mid B)P(P_n(B)=0).$$

Введем обозначение

$$X(B,n) = \sum_{j=1}^{n} \frac{n}{j} \cdot P\left(P_n(B) = \frac{j}{n} \right).$$

Рассмотрим

$$EP_{XY}^{*2}(A\times B) = P_{XY}^2(A\times B) + (1/(n+m))P_{XY}(A\times B)(1-P_{XY}(A\times B)) +$$

$$+ \frac{m}{(n+m)^2} P_{XY}(A\times B)\left(1 - \frac{P_{XY}(A\times B)}{P_Y(B)} \right) +$$

$$+ \frac{m^2}{n(n+m)^2} P_{XY}(A\times B)\left(1 - \frac{P_{XY}(A\times B)}{P_Y(B)} \right) P_Y(B)X(B,n) +$$

$$+ \frac{m}{n(n+m)^2} P_{XY}(A\times B)\left(1 - \frac{P_{XY}(A\times B)}{P_Y(B)} \right)(1-b)X(B,n) + G(n,m)P(P_n(B)=0),$$

где

$$G(n,m) = (m^2/(n+m)^2)P^2_{XY}(A\times B) -$$

$$- (m/(n+m)^2)P_{XY}(A\times B)[2 - P_{XY}(A\times B) - (P_{XY}(A\times B)/P(B))].$$

Отсюда получим окончательное выражение для СКО:

$$SP_{XY}^2(A\times B) = \frac{1}{n} P_{XY}(A\times B)(1-P_{XY}(A\times B)) - \frac{m}{n(n+m)} P_{XY}(A\times B)(1-P_{XY}(A\times B)) +$$

$$+ \frac{m}{(n+m)^2} P_{XY}(A\times B)\left(1 - \frac{P_{XY}(A\times B)}{P_Y(B)} \right) +$$

$$+ \frac{m}{n(n+m)^2} P_{XY}(A\times B)\left(1 - \frac{P_{XY}(A\times B)}{P_Y(B)} \right) P_Y(B)X(B,n) +$$

$$+ \frac{m}{n(n+m)^2} P_{XY}(A\times B)\left(1 - \frac{P_{XY}(A\times B)}{P_Y(B)} \right)(1-P_Y(B))X(B,n) +$$

$$+ G(n,m)P(P_n(B)=0). \qquad (6)$$

Полученное выражение позволяет вычислить СКО оценки (3) при конечных объемах комплектной и некомплектной выборок. Кроме того, представляет интерес случай, когда n и m возрастают.

4. Асимптотическое поведение оценки

Рассмотрим асимптотическое поведение СКО оценки (3) в условиях схемы Бернулли, полагая $m = kn$, $k \geq 0$.

Как было показано в [5]

$$\lim_{n\to\infty} X(B,n) = 1/P_Y(B). \tag{7}$$

Рассмотрим предельный переход в выражении (6). С учетом условия (7), получим

$$\lim_{n\to\infty} nSP^*_{XY}(A\times B) = (1/(k+1))P_{XY}(A\times B)(1-P_{XY}(A\times B)) +$$

$$+\frac{k}{(k+1)^2}P_{XY}(A\times B)\left(1-\frac{P_{XY}(A\times B)}{P_Y(B)}\right)+\frac{k^2}{(k+1)^2}P_{XY}(A\times B)\left(1-\frac{P_{XY}(A\times B)}{P_Y(B)}\right)=$$

$$= P_{XY}(A\times B)(1-P_{XY}(A\times B)) - \frac{k}{k+1}P^2_{XY}(A\times B)\frac{1-P_Y(B)}{P_Y(B)}. \tag{8}$$

Таким образом, $\lim\limits_{n\to\infty} nSP^*_{XY}(A\times B) \leq \lim\limits_{n\to\infty} nDP_n(A\times B)$ на величину

$$(k/(k+1))P^2_{XY}(A\times B)\cdot((1-P_Y(B))/P_Y(B)).$$

5. Иллюстрации асимптотического поведения оценки

Чтобы понять, как соотносятся дисперсия оценки (1) и среднеквадратическое отклонение оценки (3), произведем замену $m=k\cdot n$ и рассмотрим следующий показатель:

$$W = \frac{\lim\limits_{n\to\infty} nSP^2_{XY}(A\times B)}{\lim\limits_{n\to\infty} nDP_n(AB)} = 1-\frac{k}{k+1}\frac{P_{XY}(A\times B)(1-P_Y(B))}{P_Y(B)(1-P_{XY}(A\times B))}. \tag{9}$$

Рассмотрим влияние величины k (соотношение объема некомплектных наблюдений к объему комплектных) на показатель W. Отметим, что коэффициент $k \in [0,\infty)$. Рассмотрим поведение оценки на границах этого промежутка. Значение $k = 0$ означает, что объем некомплектных наблюдений равен нулю. Это равносильно отсутствию дополнительной информации и в этом случае $W = 1$. При $k \to \infty$ показатель W примет вид

$$W = W_1 = \frac{1-P_{XY}(A\times B)/P_Y(B)}{1-P_{XY}(A\times B)}. \tag{10}$$

Заметим, что наименьшее значение величина W_1 принимает в случае, когда $P_{XY}(A\times B) = P_Y(B)$ (легко видеть, что тогда $W_1 = 0$). Чтобы проиллюстрировать соотношение между оценками вида (1) и (3), зафиксируем в выражениях (9) и (10) вероятность $P_Y(B) = 0{,}8$ и рассмотрим значения вероятности $P_{XY}(A\times B)$ из интервала [0,1; 0,8] (рис. 1).

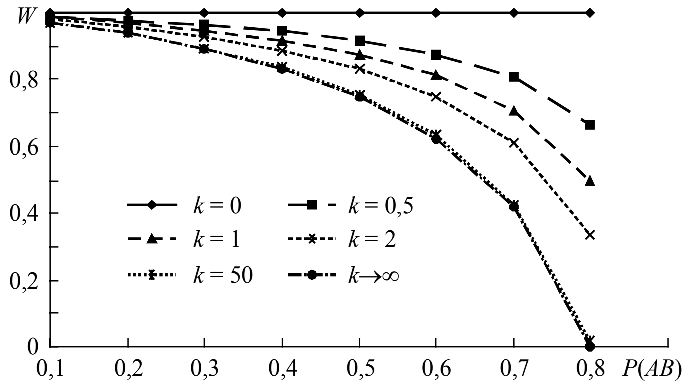

Рис. 1. Параметр W при фиксированной $P_Y(B) = 0,8$
и $P_{XY}(A\times B)\in[0,1; 0,8]$

Результаты расчетов приведены в табл. 1.

Таблица 1

Значения показателя W при $P_Y(B) = 0,8$ и $P_{XY}(A\times B)\in[0,1; 0,8]$

$P_{XY}(A\times B)$	$k = 0,5$	$k = 1$	$k = 2$	$k = 50$	$k = \infty$
0,1	0,99	0,99	0,98	0,97	0,97
0,2	0,98	0,97	0,96	0,94	0,94
0,3	0,96	0,95	0,93	0,90	0,89
0,4	0,94	0,92	0,89	0,84	0,83
0,5	0,92	0,88	0,83	0,76	0,75
0,6	0,88	0,81	0,75	0,63	0,63
0,7	0,81	0,71	0,61	0,43	0,42
0,8	0,67	0,50	0,33	0,02	0,00

Аналогично зафиксируем $P_{XY}(A\times B)$=0,2 и рассмотрим значения вероятности $P_Y(B)$ из промежутка [0,2; 1] (рис. 2). Результаты расчетов приведены в табл. 2.

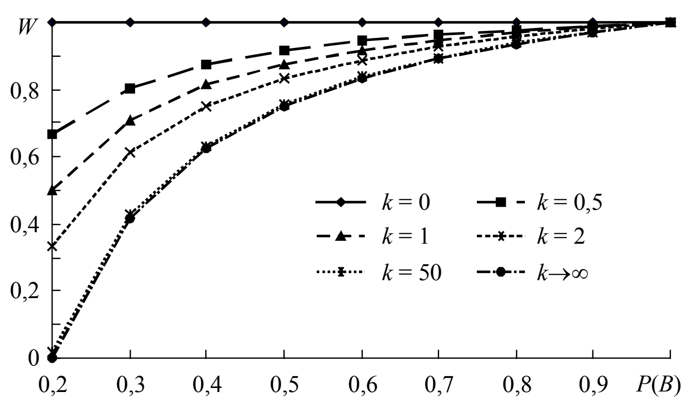

Рис. 2. Параметр W при фиксированной $P_{XY}(A\times B) = 0,2$
и $P_Y(B)\in[0,2; 1]$

Таблица 2

Значения показателя W при $P_{XY}(A\times B) = 0{,}2$ и $P_Y(B) \in [0{,}2;\ 1]$

$P_Y(B)$	$k = 0{,}5$	$k = 1$	$k = 2$	$k = 50$	$k = \infty$
0,2	0,67	0,50	0,33	0,02	0,00
0,3	0,81	0,71	0,61	0,43	0,42
0,4	0,88	0,81	0,75	0,63	0,63
0,5	0,92	0,88	0,83	0,76	0,75
0,6	0,94	0,92	0,89	0,84	0,83
0,7	0,96	0,95	0,93	0,90	0,89
0,8	0,98	0,97	0,96	0,94	0,94
0,9	0,99	0,99	0,98	0,97	0,97
1	1,00	1,00	1,00	1,00	1,00

Заметим, что случай $k = \infty$ эквивалентен ситуации, когда значение вероятности $P_Y(B)$ известно, и это соответствует самому большому выигрышу в СКО.

6. Иллюстрации поведения оценки при конечных объемах выборки

По аналогии с исследованием асимптотического поведения оценки рассмотрим величину

$$V = \frac{SP^*_{XY}(A\times B)}{DP_n(A\times B)}, \qquad (11)$$

зафиксировав конкретные значения вероятностей $P_{XY}(A\times B)$ и $P_Y(B)$ и изменяя объемы комплектной и некомплектной выборок. Для начала рассмотрим, как влияет на поведение оценки привлечение небольшого количества некомплектных наблюдений. В табл. 3 приведены значения выигрыша в СКО, полученные для разных объемов комплектных выборок.

Таблица 3

Значения показателя V при $P_{XY}(A\times B) = 0{,}2$ и $P_Y(B) = 0{,}4$

M	$n = 4$	$n = 10$	$n = 20$	$n = 50$
0	1	1	1	1
1	0,951	0,976	0,985	0,993
2	0,919	0,956	0,971	0,986
3	0,895	0,939	0,958	0,980
4	0,878	0,926	0,946	0,974
5	0,864	0,914	0,936	0,968
6	0,854	0,904	0,926	0,962
7	0,845	0,895	0,917	0,956
8	0,837	0,887	0,908	0,951

Из табл. 3 видно, что привлечение даже одного дополнительного наблюдения уменьшает СКО оценки, при этом чем меньше размерность комплектной выборки, тем значительнее эффект привлечения информации из некомплектной. Далее рассмотрим обратную ситуацию: некомплектная выборка более многочисленна, чем комплектная (табл. 4).

Таблица 4

Значения показателя V при $P_{XY}(A\times B) = 0{,}2$ и $P_Y(B) = 0{,}4$

m	$n = 5$	$n = 10$	$n = 20$	$n = 50$
10	0,7765	0,8016	0,7738	0,8225
20	0,7684	0,7864	0,7370	0,7634
30	0,7656	0,7808	0,7220	0,7339
40	0,7641	0,7779	0,7138	0,7162
50	0,7633	0,7761	0,7086	0,7044
60	0,7627	0,7749	0,7051	0,6960
70	0,7623	0,7740	0,7025	0,6897
80	0,7620	0,7734	0,7005	0,6848
90	0,7617	0,7729	0,6990	0,6808

Таким образом, и при конечных объемах выборки привлечение дополнительной информации позволяет уменьшить СКО оценки вероятности $P_{XY}(A\times B)$.

Заключение

Построены оценка с использованием формулы условной вероятности и адаптивная оценка по методу коррелированных процессов (вид оценок совпадает). Исследованы свойства оценок и асимптотическое поведение. Приведены примеры выигрыша в СКО для конечных объемов комплектной и некомплектной выборок. Проведенное исследование учета дополнительной информации позволяет утверждать, что выигрыш в СКО оценки вероятности $P_{XY}(A\times B)$ вида (3) по сравнению с эмпирической вероятностью (1) зависит от соотношения между комплектной и некомплектной выборками, а также значений вероятностей $P_{XY}(A\times B)$ и $P_Y(B)$.

ЛИТЕРАТУРА

1. *Ядов В.А.* Стратегия социологического исследования. М.: Омега-Л, 2007. 567 с.
2. *Котлер Ф.* Основы маркетинга: пер. с англ. М.: РосИнтер, 1996. 698 с.
3. *Литтл Дж.А.*, *Рубин Д.Б.* Статистический анализ данных с пропусками. М.: Финансы и статистика, 1991. 430 с.
4. *Чурилова А.А.* Корректировка неответов // Материалы матем. семинара «Несплошные статистические исследования». Нижний Новгород, 2000. С. 27.
5. *Тарима С.С.* Использование дополнительной информации при оценке вероятностей и интерпретации натурного эксперимента: дис. ... канд. техн. наук. Томск: ТГУ, 2001. 149 с.

Дмитриев Юрий Глебович
Князева Анна Анатольевна
Томский государственный университет
E-mail: dmit70@mail.ru; amili@mail.ru
Поступила в редакцию 16 апреля 2009 г.