III ЛИТОВСКИЙ МАТЕМАТИЧЕСКИЙ СБОРНИК 1
LIETUVOS MATEMATIKOS RINKINYS

1963

СТАТИСТИЧЕСКАЯ ОЦЕНКА КОЛИЧЕСТВА ИНФОРМАЦИИ В ДИСКРЕТНОМ КАНАЛЕ БЕЗ ПАМЯТИ

А. АКСОМАЙТИС

Настоящая заметка посвящена обобщению одного результата Г. П. Башарина [1] о статистической оценке энтропии.

Через $I=I(\xi, \eta)$ обозначим количество информации случайной величины ξ относительно случайной величины η в дискретном случае, т. е.

$$I(\xi, \eta)=\sum_{i=1}^{k} \sum_{j=1}^{l} p(i, j) \log \frac{p(i, j)}{p(i, \cdot) \cdot p(\cdot, j)}=H(\xi)-\sum_{j=1}^{l} p(\cdot, j) H_j. \tag{1}$$

Здесь

$$H_j=-\sum_{i=1}^{k} p(i \mid j) \log p(i \mid j),$$

$$p(i, j)=P\{\xi=x_i, \eta=y_j\}, \quad p(i, \cdot)=P\{\xi=x_i\}, \quad p(\cdot, j)=P\{\eta=y_i\},$$

$$p(i \mid j)=\mathbf{P}\{\xi=x_i \mid \eta=y_j\} \qquad (i=\overline{1, k}, \quad j=\overline{1, l}).$$

Пусть $\hat{I}_N$ статистика для I, получаема по формуле (1), вместо вероятностей $p(i, j)$, $p(i, \cdot)$, $p(\cdot, j)$ подставив соответствующие частоты $\hat{p}(i, j)$, $\hat{p}(i, \cdot)$, $\hat{p}(\cdot, j)$, получаемые из выборки объема N.

Справедлива следующая теорема:

Теорема. $\hat{I}_N$ *является смещенной асимптотически нормальной оценкой для* I, *причем*

$$\mathbf{M}\hat{I}_N=I+\frac{(k-1)(l-1)}{2N} \log_2 e+O\left(\frac{1}{N^2}\right),$$

$$\begin{aligned}\mathbf{D}\hat{I}_N=\frac{1}{N}\Bigg\{&\sum_{i=1}^{k} p(i, \cdot) \ln^2 p(i, \cdot)-H^2+\\ &+\sum_{i=1}^{k} \sum_{j=1}^{l} p(i, j)\Big[1-p(i \mid j)\Big]\Big[1+\ln p(i \mid j)\Big]^2+\\ &+2 \sum_{i, t=1}^{k} \sum_{j=1}^{l} p(i \mid j) p(t, j)\Big[1+\ln p(i, \cdot)\Big] \times\\ &\times\Big[1+\ln p(t \mid j)\Big]\Bigg\}+O\left(\frac{1}{N^2}\right).\end{aligned}$$

Доказательство теоремы. Разложим функцию

$$\hat{H}_j=-\sum_{i=1}^{k} \hat{p}(i \mid j) \ln \hat{p}(i \mid j)$$

в ряд Тэйлора в окрестности точки $\{p(i, j), p(2, j), \ldots, p(k, j)\}$, ограничиваясь производными третьего порядка. Ясно, что при условии $0 < p(i|j) < 1$, этот ряд сходится.

Имеем разложение

$$\hat{H}_j = H_j - \sum_{i=1}^{k}\left[\hat{p}(i|j) - p(i|j)\right]\left[1 + \ln p(i|j)\right] - \frac{1}{2}\sum_{i=1}^{k}\frac{[\hat{p}(i|j) - p(i|j)]^2}{p(i|j)} +$$

$$+\frac{1}{6}\sum_{i=1}^{k}\frac{[\hat{p}(i|j) - p(i|j)]^3}{\{p(i|j) + \Theta[\hat{p}(i|j) - p(i|j)]\}^2},$$

где $0 < \Theta < 1$.

Элементарно получаем, что

$$\mathbf{M}\hat{p}(i|j) = \sum_{t}\binom{N}{t}p^t(\cdot, j)[1 - p(\cdot, j)]^{N-t}\frac{1}{t}\cdot t\frac{p(i, j)}{p(\cdot, j)}.$$

$$\mathbf{M}\hat{p}(i|j) = p(i|j).$$

Второй центральный момент

$$\mathbf{M}\left\{\hat{p}(i|j) - p(i|j)\right\}^2 = \mathbf{M}\left\{\frac{\hat{p}(i, j)}{\hat{p}(\cdot, j)}\right\}^2 - p^2(i|j). \tag{2}$$

Пусть

$$\hat{h}_{ij} = \hat{p}(i, j) - p(i, j), \quad \hat{h}_{.j} = \hat{p}(\cdot, j) - p(\cdot, j).$$

Так как

$$\mathbf{M}\hat{h}_{ij}^2 = \frac{p(i, j)[1 - p(i, j)]}{N},$$

$$\mathbf{M}\hat{h}_{.j}^2 = \frac{p(\cdot, j)[1 - p(\cdot, j)]}{N},$$

$$\mathbf{M}\{\hat{h}_{ij}\hat{h}_{tr}\} = -\frac{p(i, j)p(t, r)}{N}, \qquad (i \neq t, \quad j \neq r).$$

$$\mathbf{M}\{\hat{h}_{.j}\hat{h}_{ij}\} = \frac{p(i, j)[1 - p(\cdot, j)]}{N},$$

а математическое ожидание остальных членов в (2) порядка не ниже

$$O\left(\frac{1}{N^2}\right),$$

то

$$\mathbf{M}\left\{\hat{p}(i|j) - p(i|j)\right\}^2 = \left\{\frac{p(i, j)}{p(\cdot, j)}\right\}^2\left\{\frac{p(\cdot, j) - p(i, j)}{Np(\cdot, j)p(i, j)} + O\left(\frac{1}{N^2}\right)\right\}.$$

Подробный расчет показывает, что

$$\mathbf{M}\left\{\frac{\hat{p}(i, j)}{\hat{p}(\cdot, j)}\right\}^3 = \left\{\frac{p(i, j)}{p(\cdot, j)}\right\}^3\left\{1 - \frac{3[p(\cdot, j) - p(i, j)]}{Np(\cdot, j)p(i, j)} + O\left(\frac{1}{N^2}\right)\right\}.$$

Тогда

$$\mathbf{M}\left\{\hat{p}(i|j) - p(i|j)\right\}^3 = O\left(\frac{1}{N^2}\right).$$

При помощи найденных оценок имеем:

$$\mathbf{M}\hat{H}_j = H_j - \frac{k-1}{2Np(\cdot, j)} + O\left(\frac{1}{N^2}\right).$$

Обращаясь к подсчету $\mathbf{M}\hat{H}_\eta(\xi)$, получаем:

$$\mathbf{M}\hat{H}_\eta(\xi) = H_\eta(\xi) - \frac{l(k-1)}{2N} + O\left(\frac{1}{N^2}\right).$$

Ясно, что

$$\mathbf{M}\hat{I}_N = \mathbf{M}\hat{H}(\xi) - \mathbf{M}\hat{H}_\eta(\xi) = I + \frac{(k-1)(l-1)}{2N} + O\left(\frac{1}{N^2}\right), \tag{3}$$

так как известно ([1], стр. 361 – 364):

$$\mathbf{M}\hat{H}=H-\frac{k-1}{2N}+O\left(\frac{1}{N^2}\right).$$

Равенство (3) показывает, что оценка смещенная.

Для оценки дисперсии имеем:

$$\mathbf{D}\hat{I}_N=\mathbf{M}\{\hat{I}_N-\mathbf{M}\hat{I}_N\}^2=\mathbf{M}\Big\{-\sum_{i=1}^{k}\Big[\hat{p}(i,\cdot)-p(i,\cdot)\Big]\Big[1+\ln p(i,\cdot)\Big]-$$

$$-\frac{1}{2}\sum_{i=1}^{k}\frac{[\hat{p}(i,\cdot)-p(i,\cdot)]^2}{p(i,\cdot)}+\frac{1}{6}\sum_{i=1}^{k}\frac{[\hat{p}(i,\cdot)-p(i,\cdot)]^3}{\{p(i,\cdot)[1-\Theta]+\Theta\hat{p}(i,\cdot)\}^2}\Big\}+$$

$$+\mathbf{M}\Big\{\sum_{i=1}^{k}\sum_{j=1}^{l}p(\cdot,j)\Big[\hat{p}(i|j)-p(i|j)\Big]\Big[1+\ln p(i|j)\Big]+$$

$$+\frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{l}p(\cdot,j)\frac{[\hat{p}(i|j)-p(i|j)]^2}{p(i|j)}-$$

$$-\frac{1}{6}\sum_{i=1}^{k}\sum_{j=1}^{l}p(\cdot,j)\frac{[\hat{p}(i|j)-p(i|j)]^3}{\{p(i|j)[1-\Theta]+\Theta\hat{p}(i|j)\}^2}-\frac{(k-1)(l-1)}{2N}\Big\}^2+$$

$$+2\mathbf{M}\Big\{-\sum_{i=1}^{k}\Big[\hat{p}(i,\cdot)-p(i,\cdot)\Big]\Big[1+\ln p(i,\cdot)\Big]-$$

$$-\frac{1}{2}\sum_{i=1}^{k}\frac{[\hat{p}(i,\cdot)-p(i,\cdot)]^2}{p(i,\cdot)}+\frac{1}{6}\sum_{i=1}^{k}\frac{[\hat{p}(i,\cdot)-p(i,\cdot)]^3}{\{p(i,\cdot)[1-\Theta]+\Theta\hat{p}(i,\cdot)\}^2}\Big\}\times$$

$$\times\Big\{\sum_{i=1}^{k}\sum_{j=1}^{l}p(\cdot,j)\Big[\hat{p}(i|j)-p(i|j)\Big]\Big[1+\ln p(i|j)\Big]+$$

$$+\frac{1}{2}\sum_{i=1}^{k}\sum_{j=1}^{l}p(\cdot,j)\frac{[\hat{p}(i|j)-p(i|j)]^2}{p(i|j)}-$$

$$-\frac{1}{6}\sum_{i=1}^{k}\sum_{j=1}^{l}p(\cdot,j)\frac{[\hat{p}(i|j)-p(i|j)]^3}{\{p(i|j)[1-\Theta]+\Theta\hat{p}(i|j)\}^2}-\frac{(k-1(l-1)}{2N}\Big\}+O\left(\frac{1}{N^2}\right)$$

Вычисляя центральные моменты, получаем:

$$\mathbf{M}\Big\{-\sum_{i=1}^{k}\Big[\hat{p}(i,\cdot)-p(i,\cdot)\Big]\Big[1+\ln p(i,\cdot)\Big]\Big\}=\frac{1}{N}\Big\{\sum_{i=1}^{k}p(i,\cdot)\ln^2 p(i,\cdot)-H^2\Big\}.$$

$$\mathbf{M}\Big\{\hat{p}(i,\cdot)-p(i,\cdot)\Big\}\Big\{\hat{p}(t|j)-p(t|j)\Big\}=-\frac{p(i,j)p(t,j)}{Np^2(\cdot,j)}+O\left(\frac{1}{N^2}\right),$$

а остальные центральные моменты порядка не ниже $O\left(\frac{1}{N^2}\right)$. Ввиду громоздких выражений автор не дает этих вычислений, но они больших затруднений не составляют. Метод вычисления такой же, как и второго, выше приведенного, центрального момента.

Пренебрегая членами, математическое ожидание которых имеет порядок малости выше второго, имеем:

$$\mathbf{D}\hat{I}_N=\frac{1}{N}\left\{\sum_{i=1}^{k}p(i,\,\cdot)\ln^2 p(i,\,\cdot)-H^2+\sum_{i=1}^{k}\sum_{j=1}^{l}p(i|j)\left[1+\ln p(i|j)\right]^2\left[1-p(i|j)\right]+\right.$$
$$\left.+2\sum_{i,t=1}^{k}\sum_{j=1}^{l}p(i|j)p(t,\,j)\left[1+\ln p(i,\,\cdot)\right]\left[1+\ln p(t|j)\right]\right\}+0\left(\frac{1}{N^2}\right).$$

Для оценки членов, в которых входит Θ, использовали неравенство Шварца (см. [2] 9.5.1).

Асимптотическая нормальность статистики $\hat{I}_N$ следует из легко получаемого разложения

$$\sqrt{N}\,\{\hat{I}_N-I\}=\sqrt{N}\left\{\sum_{i=1}^{k}\sum_{j=1}^{l}p(\cdot,\,j)\left[\hat{p}(i|j)-p(i|j)\right]\left[1+\ln p(i|j)\right]-\right.$$
$$\left.-\sum_{i=1}^{k}\left[p(i,\,\cdot)-p(i,\,\cdot)\right]\left[1+\ln p(i,\,\cdot)\right]\right\}+$$
$$+\frac{1}{2}\sqrt{N}\left\{\sum_{i=1}^{k}\sum_{j=1}^{l}\frac{p(\cdot,\,j)\,[\hat{p}(i|j)-p(i|j)]^2}{p(i|j)\,[1-\Theta]+\Theta\hat{p}(i|j)}-\right.$$
$$\left.-\sum_{i=1}^{k}\frac{[\hat{p}(i,\,\cdot)-p(i,\,\cdot)]^2}{p(i,\,\cdot)\,[1-\Theta]+\Theta\hat{p}(i,\,\cdot)}\right\},$$

применив к случайной величине $\sqrt{N}\,(\hat{I}_N-I)$ теорему 2.8.4 из [2]. Теорема доказана.

Вильнюсский государственный
университет им. В. Капсукаса

Поступила в редакцию
1. X. 1962

ЦИТИРОВАННАЯ ЛИТЕРАТУРА

1. Г. П. Башарин, О статистической оценке энтропии последовательности независимых случайных величин, Теория вероятностей и ее прим., том IV, вып. 3 (1959).
2. Г. Крамер, Математические методы статистики, ИЛ (1948).

DISKRETINIO BE ATMINTIES KANALO INFORMACIJOS KIEKIO STATISTINIS ĮVERTINIMAS

A. AKSOMAITIS

(Reziumė)

Darbe gautas statistinis informacijos kiekio, o tuo pačiu ir kanalo galingumo, įvertinimas. Parodyta, kad įvertinimas turi poslinkį ir yra asimptotiškai normalinis.

STATISTICAL EVALUATION OF INFORMATION RECEIVED BY MEANS OF DISCRETE CANAL WITHOUT MEMORY

A. AKSOMAITIS

(Summary)

Subject of investigation is statistical characteristics of information $\hat{I}_N$. It is proved that the statistics are asymptotically normal but shifted.