

**А. Ю. Кулай** (Москва, ООО «Стэл КС»). Условная вероятность двух реализаций однородной цепи Маркова с неизвестными параметрами распределения.

Пусть  $X$  и  $O$  — реализации случайной последовательности, некоторые параметры распределения которой неизвестны. В ряде приложений, например, при распознавании образов [1], необходимо вычислять вероятность  $X$  при условии наблюдения  $O$ .

Как и в байесовском подходе к оцениванию параметров [2], рассмотрим неизвестные параметры  $\bar{\theta}$  как случайную величину с некоторой плотностью распределения. При отсутствии иной информации будем считать, что на параметрическом пространстве  $\Theta$  задано равномерное распределение с плотностью  $\mu(\bar{\theta}) = \text{const}$ . Тогда

$$P(X|O) = \int_{\bar{\theta} \in \Theta} P(X|\bar{\theta}) dP(\bar{\theta}|O),$$

$$dP(\bar{\theta}|O) = \frac{P(O|\bar{\theta})\mu(\bar{\theta}) d\bar{\theta}}{\int_{\bar{\theta}' \in \Theta} P(O|\bar{\theta}')\mu(\bar{\theta}') d\bar{\theta}'} = \frac{P(O|\bar{\theta}) d\bar{\theta}}{\int_{\bar{\theta}' \in \Theta} P(O|\bar{\theta}') d\bar{\theta}'},$$

$$P(X|O) = \int_{\bar{\theta} \in \Theta} P(X|\bar{\theta}) \frac{P(O|\bar{\theta}) d\bar{\theta}}{\int_{\bar{\theta}' \in \Theta} P(O|\bar{\theta}') d\bar{\theta}'} = \frac{\int_{\bar{\theta} \in \Theta} P(X|\bar{\theta})P(O|\bar{\theta}) d\bar{\theta}}{\int_{\bar{\theta}' \in \Theta} P(O|\bar{\theta}') d\bar{\theta}'}$$

В [3] рассматривался случай выборки из неизвестного полиномиального распределения. Для последовательности случайных величин, связанных в цепь Маркова, вычисление интегралов в последнем равенстве приводит к следующему результату.

**Теорема.** Если  $X$  и  $O$  — реализации неизвестной однородной цепи Маркова (с дискретным временем) с  $n$  исходами, на параметрическом пространстве  $\Theta$  задано равномерное распределение, то

$$P(X|O) = \frac{1 + I\{x_1 = o_1\}}{n + 1} \prod_{i=1}^n \frac{\prod_{j=1}^n A_{v_X(i,j)+v_O(i,j)}^{v_X(i,j)}}{A_{v_X(i,*)+v_O(i,*)+n-1}^{v_X(i,*)}},$$

где  $I_Z$  — индикатор события  $Z$ ,  $y_1$  — первый символ последовательности  $Y$ ,  $Y \in \{O, X\}$ ,  $A_n^m$  — количество размещений из  $n$  по  $m$ ,  $v_Y(i, j)$  — частота встречаемости в последовательности  $Y$  пары символов  $(i, j)$ ,  $Y \in \{O, X\}$ ,  $v_Y(i, *)$  — частота встречаемости в последовательности  $Y$  пар символов с первым символом  $i$ ,  $Y \in \{O, x\}$ .

#### СПИСОК ЛИТЕРАТУРЫ

1. Фукунага К. Введение в статистическую теорию распознавания образов. М.: Наука, 1979.
2. Боровков А. А. Математическая статистика. М.: Наука, 1984.
3. Кулай А. Ю. Условная вероятность двух реализаций случайной последовательности с неизвестными параметрами распределения. — Обозрение прикл. и промышл. матем., 2009, т. 14, в. 6, с. 1087–1088.