

Лекция №4 — 05.10, 2021

1 Дискретные источники сообщений

Для построения математической модели дискретного источника сообщений (ДИС) нам требуется:

- Непустое конечное множество $\mathcal{A} = \{a_1, \dots, a_m\}$ - алфавит источника (элементы a_i - символы алфавита);
- Для любой конечной последовательности $a^n = (a_{i_1}, \dots, a_{i_n})$ букв из алфавита \mathcal{A} определить вероятность $P(a_n)$ появления этой последовательности на выходе источника.

Обозначим через \mathcal{A}^n - n -ю декартову степень множества \mathcal{A} , т.е. множество всех слов длины n , составленных из букв алфавита \mathcal{A} . При этом множество всех бесконечных последовательностей букв обозначим

$$\mathcal{A}^\infty = \{\omega = (\omega_1, \omega_2, \dots) | \omega_i \in \mathcal{A}\}.$$

Определение 1. *Элементарным цилиндрическим множеством $C(n_1, n_2, \dots, n_k; a_{i_1}, \dots, a_{i_k})$ с параметрами $k \geq 1, 1 \leq n_1 < n_2 < \dots < n_k, a_{i_1}, \dots, a_{i_k} \in \mathcal{A}$ называется множество всех последовательностей $\omega \in \mathcal{A}^\infty$, для которых $\omega_{n_1} = a_{i_1}, \dots, \omega_{n_k} = a_{i_k}$.*

Определение 2. *Цилиндрическим множеством общего вида $C(n_1, n_2, \dots, n_k; \mathcal{B})$ с параметрами $k \geq 1, n_1, \dots, n_k, \mathcal{B} \subseteq \mathcal{A}^k$ называется множество всех последовательностей $\omega \in \mathcal{A}^\infty$, для которых $(\omega_{n_1}, \dots, \omega_{n_k}) \in \mathcal{B}$.*

Под дискретным источником сообщений понимается устройство, генерирующее в i -й момент времени (время дискретно, т.е. $i \in \mathbb{N}$) некоторый элемент алфавита $a_i \in \mathcal{A}$. Вероятностное пространство (Ω, \mathcal{F}, P) на множестве всех сообщений, которые может сгенерировать ДИС, задается следующим образом:

- $\Omega := \mathcal{A}^\infty$ - множество элементарных событий;
- \mathcal{F} - наименьшая по включению σ -алгебра подмножеств пространства Ω , включающая все цилиндрические множества общего вида;
- P - определяем через последовательность вероятностных мер на конечных подмножествах \mathcal{A}^k (см. ниже).

Определение 3. Пусть на σ -алгебре \mathcal{F} задана вероятностная мера P . Тогда будем говорить, что задан дискретный случайный процесс (или бесконечная случайная последовательность) $A : \Omega \rightarrow \Omega$, $A = (A_n)_{n=1}^{+\infty} = (A_1, A_2, \dots)$.

Определение 4. Функция P_{n_1, n_2, \dots, n_k} , заданная на классе всех подмножеств множества \mathcal{A}^k формулой

$$P_{n_1, n_2, \dots, n_k}(\mathcal{B}) = P(C(n_1, n_2, \dots, n_k; \mathcal{B})), \mathcal{B} \subseteq \mathcal{A}^k$$

называется конечномерным распределением случайного процесса A на местах n_1, \dots, n_k .

Определение 5. Функция P_{n_1, n_2, \dots, n_k} , заданная на классе всех подмножеств множества \mathcal{A}^k формулой

$$P_{n_1, n_2, \dots, n_k}(\mathcal{B}) = P(C(n_1, n_2, \dots, n_k; \mathcal{B})), \mathcal{B} \subseteq \mathcal{A}^k$$

называется конечномерным распределением случайного процесса A на местах n_1, \dots, n_k .

Определение 6. Пусть для любого $k = 1, 2, \dots$ на классе всех подмножеств множества \mathcal{A}^k задана вероятностная мера $P_{1, 2, \dots, k}$. Последовательность мер $P_{1, 2, \dots, k}$, $k \geq 1$ называется согласованной, если для любых подмножеств $\mathcal{B} \subseteq \mathcal{A}^k$ справедливо равенство

$$P_{1, 2, \dots, k, k+1}(\mathcal{B} \times \mathcal{A}) = P_{1, 2, \dots, k}(\mathcal{B})$$

Теорема 7. Пусть для любого $k = 1, 2, \dots$ на классе всех подмножеств множества \mathcal{A}^k задана последовательность согласованных вероятностных мер $P_{1, 2, \dots, k}$, $k \geq 1$. Тогда на σ -алгебре \mathcal{F} , порожденной классом всех цилиндрических множеств, существует, и притом единственная, вероятностная мера P , продолжающая последовательность мер $P_{1, 2, \dots, k}$, т.е. для всех подмножеств $\mathcal{B} \subseteq \mathcal{A}^k$ справедливо равенство

$$P(C(1, 2, \dots, k; \mathcal{B})) = P_{1, 2, \dots, k}(\mathcal{B})$$

Следствие 8 (Теорема 7). Согласованность конечномерных распределений $P_{1, 2, \dots, k}$ является необходимым и достаточным условием существования вероятностной меры P на вероятностном пространстве $\Omega = \mathcal{A}^\infty, \mathcal{F}, P$.

2 Примеры дискретных источников сообщений

1. **Дискретный источник без памяти.** Задается двумя параметрами (\mathcal{A}, \vec{p}) , где $\mathcal{A} = (a_1, \dots, a_m)$ - алфавит источника, а $\vec{p} = (p_1, \dots, p_m)$ - вероятностный вектор, в котором p_i есть вероятность того, что источник выдаст символ a_i в произвольный момент времени.

Строго говоря, для описания ДИС необходимо построить вероятностное пространство (Ω, \mathcal{F}, P) . Множество элементарных событий Ω и σ -алгебра \mathcal{F} задаются стандартным образом, как описано в *разделе 1*. Вероятностная мера задается через конечномерные распределения $P_{1, 2, \dots, k}$, определенные на словах $(a_{i_1}, \dots, a_{i_k}) \in \mathcal{A}^k$ следующим образом:

$$P_{1, 2, \dots, k}(a_{i_1}, \dots, a_{i_k}) = p_{i_1} p_{i_2} \cdot \dots \cdot p_{i_k}$$

Нетрудно проверить, что определенные таким образом функции $P_{1,\dots,k}$ являются распределением вероятностей, т.е.

$$P_{1,2,\dots,k}(\mathcal{A}^k) = 1$$

и согласованы, т.е.

$$P_{1,2,\dots,k,k+1}(\mathcal{B} \times \mathcal{A}) = P_{1,2,\dots,k}(\mathcal{B})$$

2. **Простой марковский источник** задается тремя параметрами $(\mathcal{A}, \vec{p}, Q)$, где $\mathcal{A} = (a_1, \dots, a_m)$ - алфавит источника, а $\vec{p} = (p_1, \dots, p_m)$ - инициализирующий вектор, в котором p_i есть вероятность того, что источник выдаст символ a_i в начальный момент времени. $Q^{m \times m}$ - стохастическая матрица (матрица переходов), в которой элемент p_{ij} означает вероятность генерации символа a_j после символа a_i .

Строго говоря, для описания ДИС необходимо построить вероятностное пространство (Ω, \mathcal{F}, P) . Множество элементарных событий Ω и σ -алгебра \mathcal{F} задаются стандартным образом, как описано в *разделе 1*. Вероятностная мера задается через конечномерные распределения $P_{1,2,\dots,k}$, определенные на словах $(a_{i_1}, \dots, a_{i_k}) \in \mathcal{A}^k$ следующим образом:

$$P_1(a_i) = p_i \text{ для } k = 1,$$

$$P_{1,2,\dots,k}(a_{i_1}, a_{i_2}, a_{i_3} \dots, a_{i_{k-1}}, a_{i_k}) = p_{i_1} q_{i_1 i_2} q_{i_2 i_3} \cdot \dots \cdot q_{i_{k-1} i_k}.$$

Теорема 9. *Дискретный источник без памяти стационарен.*

Теорема 10. *Простой марковский источник $(\mathcal{A}, \vec{p}, Q)$ стационарен тогда и только тогда, когда $\vec{p} \cdot Q = \vec{p}$.*