Дискретный источник. Дискретный источник без памяти. Цепь Маркова

4 января 2023 г. 2:47

Дискретный источник - это устройство, которое в каждый момент времени выбирает одно сообщение из дискретного множества.
Источник может быть не только дискретным по значениям, но ещё и дискретным по времени.
Если множество значений времени также дискретно, то источник называется дискретным по времени.
Источник считается заданным, если известна его вероятностная модель. Другими словами, мы должны определить вероятностную модель случайного процесса генерирования случайных сообщений на выходе источника. (Т.е. как с дискретным ансамблем: задан алфавит, заданы вероятности, можно сказать, с какой вероятностью источник будет генерировать какую-то последовательность сообщений, тогда мы считаем, что этот источник задан)

Дискретный источник задан, если для n = 1,2,… и i = 0,1,2,… известна вероятность $p\left(x_{i+1},\ x_{i+2},...,x_{i+n}\right)$ случайной последовательности из $\{X_{i+1}X_{i+2}\dots X_{i+n}\}$ (здесь каждый X - это алфавит), которая начинается с индекса i+1 и имеет длину n, где $x_j\in X_j$, j = i+1, …, i+n. Обычно рассматривается случай, когда $X_j=X$ для всех j (то есть каждый раз при передаче сообщения используется один и тот же алфавит).

Последовательность, сгенерированная дискретным источником, может также называться дискретным случайным процессом.
Пусть $\boldsymbol{x}=(x_1,x_2,...,x_n,...)$ - дискретный случайный процесс.
Рассмотрим случайный вектор $\boldsymbol{x}_{j+1}^{i+n}=\left(x_{j+1},x_{j+2},...,x_{j+n}\right)$. (В любой момент времени берем какую-то последовательность значений с индексом от j+1 до j+n)
Стационарность процесса означает, что для любого n и j, $p(\boldsymbol{x}_{j+1}^{i+n}\Big)$ не зависит от сдвига во времени j, т.е. $p(\boldsymbol{x}_{j+1}^{i+n}\Big)=p(\boldsymbol{x}_1^n\Big)$.
Источник называют дискретным источником без памяти, если:
$p\left(x_1,x_2,...,x_n\right)=\prod_{i=1}^n p(x_i)$
(т.е. если источник выдает сообщения, которые не зависят друг от друга, то он без памяти)

Дискретный случайный процесс называется цепью Маркова порядка s (s ещё называется размером памяти), если для любого n и $\boldsymbol{x}=(x_1,...,x_n)\in X^n$ (X^n - всевозможные комбинации иксов) выполняется следующее равенство:
$p\left(\boldsymbol{x}\right)=p\left(x_1,...,x_s\right)p\left(x_{s+1}|x_1,...,x_s\right)p\left(x_{s+2}|x_2,...,x_{s+1}\right)*...*p\left(x_n|x_{n-s},...,x_{n-1}\right)$

(вероятность последовательности сообщений с индексами от 1 до n = вероятность первичного сообщения с индексами от 1 до s - т.е. первые s сообщений, сгенерированных источником, - умноженная на вероятность сообщения с индексом s+1, зависящего от предыдущих s сообщений, и так далее идет умножение с зависимостью от предыдущих s сообщений)
Другими словами, для цепи Маркова справедливо равенство:
$p\left(x_n|x_1,...,x_{n-1}\right)=p\left(x_n|x_{n-s},...,x_{n-1}\right),$
Условная вероятность текущего значения зависит от s предыдущих значений и не зависит от остальных

x_t - состояние марковской цепи в момент времени t. Марковская цепь порядка s задается начальным распределением вероятностей первых s значений (состояний) и условными вероятностями вида $p\left(x_n|x_{n-s},...,x_{n-1}\right)$ для всевозможных последовательностей $x_{n-s},...,x_{n-1}$. Если эти условные вероятности не изменяются при сдвиге во времени, то марковская цепь называется однородной.

Марковская цепь порядка s=1 с состояниями X = { 0, 1, …, M-1 } определяется начальным распределением $\{p\left(x_1\right),\ x_1\in X\}$ и условными вероятностями
$\pi_{ij}=P(x_t=j|x_{t-1}=i)$, i, j = 0, 1, …, M-1
Матрица переходных вероятностей:

$$\Pi=\begin{pmatrix}\pi_{00} & \pi_{01} & \cdots & \pi_{0,M-1}\\ \pi_{10} & \pi_{11} & \cdots & \pi_{1,M-1}\\ \cdots & \cdots & \cdots & \cdots\\ \pi_{M-1,0} & \pi_{M-1,1} & \cdots & \pi_{M-1,M-1}\end{pmatrix}$$

Обозначим через $p_t=\left(p_t\left(0\right),...,\ p_t(M-1)\right)$ стохастический (т.е. вероятностный) вектор, компоненты которого - вероятности состояний цепи Маркова в момент времени t, где $p_t\left(i\right)$, i = 0, 1, …, M-1 - вероятность состояния i в момент времени t.
Из формулы полной вероятности следует:
$p_{t+1}\Big(i\Big)=\sum_{j=0}^{M-1}p_t(j)\pi_{ji}$
(вероятность перехода, например, в букву А = сумме вероятностей перехода из всех букв алфавита в букву А)
В матричном виде: $\boldsymbol{p}_{t+1}=\ \boldsymbol{p}_t\Pi$
(если умножить вероятностный вектор в момент времени t на матрицу переходных вероятностей, получим вероятностный вектор в момент времени t+1)
Для произвольного числа шагов n: $\boldsymbol{p}_{t+n}=\ \boldsymbol{p}_t\Pi^n$
Из формулы следует, что распределение вероятностей в момент времени t зависит от величины t и от начального распределения $\boldsymbol{p}_1$. Отсюда следует, что в общем случае

рассматриваемый случайный процесс нестационарен (т.к. следующий шаг зависит от предыдущего, $\boldsymbol{p}_{t+1}$ зависит от $\boldsymbol{p}_t$). Однако, если существует стохастический вектор $\boldsymbol{p}$, такой что $\boldsymbol{p} = \boldsymbol{p}\Pi$, то, выбрав начальное распределение равное этому $\boldsymbol{p}$ ($\boldsymbol{p}_1 = \boldsymbol{p}$), мы получим стационарный процесс. Вектор $\boldsymbol{p}$, удовлетворяющий $\boldsymbol{p} = \boldsymbol{p}\Pi$, называется стационарным распределением вероятностей для марковской цепи с матрицей переходных вероятностей П.

Энтропия символа x_t (из ансамбля $X_t = X$), сгенерированного в момент времени t, не зависит от t ($H(X_t) = H(X)$) и называется одномерной энтропией источника (процесса). Обозначим её как $H_1(X)$.
$H_1(X)$ не учитывает зависимость между символами, порожденными источником

Рассмотрим $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ из $X_1 X_2 \ldots X_n = X^n$.
Энтропия $H(X_1 X_2 \ldots X_n) = H(X^n)$ называется n-**мерной энтропией** процесса.

энтропия последовательности из n символов

Энтропия на символ для последовательности длины n определяется как:

$$H_n(X) = \frac{H(X^n)}{n},$$

вычисляем как условную энтропию

Другой способ:

$$H(X_n | X_1, \ldots, X_{n-1}) = H(X|X^{n-1}).$$

тоже два способа

Энтропия на сообщение:

$$\lim_{n\to\infty} H_n(X) \text{ и } \lim_{n\to\infty} H(X|X^n)$$

Theorem

Для дискретного стационарного процесса (источника)

A. $H(X|X^n)$ *не возрастает с увеличением* n;
B. $H_n(X)$ *не возрастает с увеличением* n;
C. $H_n(X) \geq H(X|X^{n-1})$;
D. $\lim_{n\to\infty} H_n(X) = \lim_{n\to\infty} H(X|X^n)$.

(док-во в "Энтропия на сообщение")