Теория вероятностей. Лекция пятая
Дискретные случайные величины

Дмитрий Валерьевич Хлопин
glukanat@mail.ru

Институт математики и механики им. Н.Н.Красовского

02.10.2018

Напомним марковские цепи

Последовательность $(X_k)_{k\in\mathbb{N}\cup\{0\}}$ называется стационарной цепью Маркова на пространстве состояний $\{1,\ldots,n\}$, если

$$\begin{aligned} \mathbb{P}(X_k|X_{k-1}) &= \mathbb{P}\big(X_k\big|X_{k-1},X_{k-2},\ldots,X_0\big) \qquad \forall k\in\mathbb{N}, \\ \mathbb{P}(X_k=j|X_{k-1}=i) &= q_{ij}^{(k)} \qquad \forall k\in\mathbb{N} \end{aligned}$$

для некоторой матрицы переходов $Q=(q_{ij})_{n\times n}$.

Зададим распределения–строки $\mu_k \overset{\triangle}{=} \big(\mathbb{P}(X_k=1),\ldots,\mathbb{P}(X_k=n)\big)$.
По начальной строке μ_0 остальные восстанавливаются однозначно:

$$\begin{aligned} \mathbb{P}(X_n=a_n,\ldots,X_0=a_0) &= q_{a_{n-1}a_n}\ldots q_{a_1a_2}q_{a_0a_1}\mathbb{P}(X_0=a_0), \\ \mu_1 &= \mu_0 Q,\ \ \mu_n=\mu_0 Q^n. \end{aligned}$$

Стационарное распределение

Распределение μ называется стационарным (для цепи Маркова с матрицей переходов Q), если $\mu Q = \mu$.

Теорема 4. Всякая цепь Маркова имеет хотя бы одно стационарное распределение.
Доказательство. Рассмотрим множество Δ всевозможных строк-распределений. Отображение $\mu \mapsto \mu Q$ является непрерывным отображением из компакта Δ в него же. По теореме Брауэра это отображение имеет неподвижную точку μ_*, то есть имеется распределение со свойством $\mu_* = \mu_* Q$.

Подумать: единственность стационарного распределения, в общем случае, никто не обещал.

Эргодический случай

Стохастическая матрица $Q = (q_{ij})_{n \times n}$ называется эргодической, если все её элементы положительны.

Теорема 5 [без д-ва] Пусть матрица переходов Q эргодична. Тогда найдется такое стационарное распределение μ_*, что для любого начального распределения μ_0 распределения $\mu_n = \mu_0 Q^n$ сходятся к μ_*.

Важный факт 1: как и в схеме независимых испытаний Бернулли, если достаточно долго подождать, то всё сойдется.
Важный факт 2: такая система забывает прошлое, знание состояния в некоторый момент времени не позволяет восстановить историю.
Подумать: условие эргодичности существенно для сходимости.

Взгляд со стороны динамических систем

Состояние j называют достижимым из i, если $(Q^k)_{ij} > 0$ для некоторого $k > 0$ (существует ненулевая вероятность попасть в x_i из x_j за некоторое число шагов).
Состояния, достижимые друг из друга, называются сообщающимися.
Если x достижимо из y, но не наоборот, то y несущественно для x.
Множество всех существенных для друг друга состояний разбивается на непересекающиеся классы сообщающихся состояний. Если такой класс только один, цепь называется неразложимой.
Множество состояний замкнуто, если вероятность выйти из него равна нулю. Состояние i называется возвратным, если вероятность туда вернуться равна 1, и поглощающим, если $q_{ii} = 1$.

Итог первой части

Задача теории вероятностей — найти вероятность сложных событий.
Рецепты при обработке новой информации:

при выполненной гипотезе — условная вероятность;

объединение по гипотезам — формула полной вероятности;

вероятность *a posteriori* — формула Байеса;

в случае, когда настоящее не зависит от прошлого (обучения нет), использовать независимость;
в случае, когда будущее зависит только от настоящего (текущего состояния), но не от прошлого, использовать марковское свойство.

Замечание в минус: пока всё это работает только при конечном числе состояний и гипотез, время также дискретно.
Еще одно замечание в минус: даже такие подалгебры очень быстро растут со временем, нужны инструменты попроще (да и запросы тоже).

Случайные величины

Инструменты (пока лишь в дискретном случае):

- дискретное распределение;
- совместное распределение случайных величин, маргинальные распределения;
- независимость случайных величин;
- медиана, математическое ожидание;
- дисперсия, ковариация и корреляция;
- условное математическое ожидание.

Дискретные случайные величины

В заданном измеримом пространстве $(\Omega, \mathcal{F})$ отображение $\xi : \Omega \to \mathbb{R}$ называется **дискретной случайной величиной**, если найдутся не более чем счетное разбиение $H_1, \ldots, H_i, \ldots \in \mathcal{F}$ множества Ω и не более чем счетный набор чисел x_i такие, что $\xi(\omega) = x_i$ для всех $\omega \in H_i$.

В заданном измеримом пространстве $(\Omega, \mathcal{F})$ отображение $\xi : \Omega \to \mathbb{R}$ называется **дискретной случайной величиной**, если оно принимает не более чем счетное число значений $x_1, \ldots, x_i, \ldots$, и при этом для всякого числа x_i выполнено

$$H_i = \{\omega \in \Omega \,|\, \xi(\omega) = x_i\} \in \mathcal{F}.$$

Подумать: почему это одно и то же определение.

Операции над случайными величинами

Даны две дискретные величины.

Подумать: докажите, что сумма этих дискретных случайных величин также является дискретной случайной величиной.

Подумать: докажите, что произведение этих дискретных случайных величин также является дискретной случайной величиной.

Подумать: когда можно ввести частное этих дискретных случайных величин?

Подумать: можно ли эти дискретные величины сравнивать, или, к примеру, определить их максимум?

Подумать: что такое предел дискретных случайных величин?

Дискретные распределения

Дискретным распределением называют отображение из счетного подмножества множества $\mathbb{R}$ в $[0,1]$, сумма элементов образа которого равна единице.

Пусть вероятность $\mathbb{P}$ также задана. Теперь каждая случайная величина задает распределение случайной величины ξ, дискретное распределение, правилом:

x_1	x_2	x_3	x_4	$\ldots$
$p_1 \triangleq \mathbb{P}(H_1)$	$p_2 \triangleq \mathbb{P}(H_2)$	$p_3 \triangleq \mathbb{P}(H_3)$	$p_4 \triangleq \mathbb{P}(H_4)$	$\ldots$

Подумать: проверьте формулу $p_i = (\xi\#\mathbb{P})\{x_i\}$.

Функции распределения

Функцией распределения дискретной случайной величины ξ называют отображение $F_\xi : \mathbb{R} \to [0,1]$, заданное по правилу:

$$F_\xi(x) \triangleq \mathbb{P}\{\omega \,|\, \xi(\omega) \le x\} = \sum_{x_i \le x} p_i \qquad \forall x \in \mathbb{R}.$$

Фактически задать функцию распределения можно было и правилом $F_\xi(x) \triangleq (\xi\#\mathbb{P})((-\infty, x])$.

Предупреждение: функцию распределения иногда вводят иначе:

$$F_\xi(x) \triangleq \mathbb{P}\{\omega \,|\, \xi(\omega) < x\} = \sum_{x_i < x} p_i, \;\; F_\xi(x) \triangleq (\xi\#\mathbb{P})((-\infty, x)).$$

Это схоже, но не эквивалентно, определениям выше, в ближайшее время сие различие несущественно.

Функции распределения: подумать...

Функцией распределения дискретной случайной величины ξ называют отображение $F_\xi : \mathbb{R} \to [0,1]$, заданное по правилу:

$$F_\xi(x) \triangleq \mathbb{P}\{\omega \,|\, \xi(\omega) \le x\} = \sum_{x_i \le x} p_i \qquad \forall x \in \mathbb{R}.$$

Подумать: как по функции распределения однозначно восстановить само дискретное распределение.

Подумать: как по дискретному распределению (функции распределения) найти хотя бы одну случайную величину с таким распределением (про вероятностное пространство также забывать не стоит).

Подумать: можно ли по двум дискретным распределениям найти их сумму?

Примеры дискретных распределений: дискретное равномерное

$U\{m,n\}$, где параметры m,n — целые числа $(m \le n)$.

m	$m+1$	$m+2$	$\ldots$	n
$\dfrac{1}{n-m+1}$	$\dfrac{1}{n-m+1}$	$\dfrac{1}{n-m+1}$	$\ldots$	$\dfrac{1}{n-m+1}$

Примеры дискретных распределений: вырожденное

Для каждого числа $a \in \mathbb{R}$ можно ввести распределение, сосредоточенное в a :

$$\begin{array}{|c|}
a \\
\hline
1
\end{array}$$

Иногда его называют дельта-распределением, дельта-функцией Дирака, и обозначают δ_a.

Подумать: во всяком ли вероятностном пространстве имеется случайная величина с таким распределением?

Подумать: почему теперь можно считать действительные числа дискретными распределениями? А как на них теперь умножать?

Подумать: проверьте тождество $\xi + \xi = \delta_2 \xi$.

Примеры дискретных распределений: Бернулли

$Bernulli(p) = Bin(1, p)$, где $p \in [0, 1]$ — вероятность "успеха".

0	1
$q \overset{\triangle}{=} 1 - p$	p

Подумать: почему такое распределение иногда записывают как

$$q\delta_0 + p\delta_1?$$

Примеры дискретных распределений: биномиальное

$B(n,p) = Bin(n,p)$, где $p \in [0,1]$ — вероятность "успеха", $n \in \mathbb{N}$ — число испытаний.

Моделирует число успехов k за ровно n испытаний.

0	$\dots$	k	$\dots$	n
q^n	$\dots$	$C_n^k p^k q^{n-k}$	$\dots$	p^n

Подумать: почему сумма чисел в последней строке равна 1.

Примеры дискретных распределений: геометрическое

$Geom(p)$ (где $p \in (0,1)$ — вероятность "успеха") вводят по-разному. Иногда как "число выстрелов до первого попадания":

1	$\dots$	k	$\dots$
p	$\dots$	pq^{k-1}	$\dots$

Иногда как "число промахов до первого попадания":

0	$\dots$	k	$\dots$
p	$\dots$	pq^{k}	$\dots$

[С-но; 0,6 баллов] Каким может быть заданное на множестве всех натуральных чисел распределение случайной величины ξ, если $\mathbb{P}(\xi > a+b|\xi > a) = \mathbb{P}(\xi > b)$ для всех натуральных a, b?

Примеры дискретных распределений: Пуассона

$Pois(\lambda)$ для $\lambda > 0$. Моделирует число успехов k за единицу времени, если в среднем за эту единицу времени происходит λ успехов.

0	1	2	$\dots$	k	$\dots$
$e^{-\lambda}$	$e^{-\lambda}\lambda$	$\frac{e^{-\lambda}\lambda^2}{2}$	$\dots$	$\frac{e^{-\lambda}\lambda^k}{k!}$	$\dots$

Подумать: почему сумма чисел в последней строке равна 1.

Теорема Пуассона

Теорема. (Теорема Пуассона, теорема о редких событиях) Пусть p_n зависит от n так, что $\lambda_n = np_n \to \lambda > 0$ при $n \to \infty$. Пусть случайная величина $\xi_n \in Bin(n, p_n)$ распределена по биномиальному закону с вероятностью успеха p_n. Тогда для любого $k \in \{0, 1, 2, \dots\}$

$$P_n(\xi_n = k) = C_n^k p_n^k (1 - p_n)^{n-k} \to \frac{\lambda^k}{k!} e^{-\lambda}.$$

К этой теореме также есть оценки точности (Le Cam's theorem): ошибка в теореме не превосходит $2\min(\lambda_n, 1)/n$.

Набросок доказательства:

$$\frac{n(n-1)\dots(n-k+1)}{k!} \cdot \frac{(p_n n)^k}{n^k} (1-p_n)^{n-k} \approx$$

$$\frac{n}{n} \cdot \frac{n-1}{n} \cdot \ldots \cdot \frac{n-k+1}{n} \cdot \frac{\lambda^k}{k!} \Big((1-p_n)^{\frac{1}{p_n}} \Big)^{p_n(n-k)} \quad \to \quad \frac{\lambda^k}{k!} e^{-\lambda}.$$

Что будет дальше?

- распределение случайных величин
- совместное распределение случайных величин, маргинальные распределения
- независимость случайных величин
- медиана, математическое ожидание
- дисперсия, ковариация и корреляция
- условное математическое ожидание

На пять минут...

1. А что такое X_k из определения цепи Маркова? Теперь у Вас имеются нужные определения.

2. Даны две случайные величины, распределенные по Бернулли. Найдите распределение их суммы.

3. В схеме бесконечного числа независимых испытаний Бернулли пусть $f_k(\omega)$ — результат k-го испытания. Напишите отображение $f_k\#\mathbb{P}$.