

Теория вероятностей. Лекция восьмая

Совместное и маргинальное распределения

Дмитрий Валерьевич Хлопин
glukanat@mail.ru

Институт математики и механики им. Н.Н.Красовского

23.10.2018

Что дальше?

- распределение случайных величин
- медиана и математическое ожидание
- независимость случайных величин
- производящие функции как матожидание
- дисперсия, ковариация и корреляция
- векторные случайные величины, их совместные распределения, маргинальные распределения
- энтропия и условная энтропия
- условное математическое ожидание

Для восстановления значения одной случайной величины с помощью значения другой случайной величины, можно воспользоваться ковариацией и корреляцией. Но они дают лишь линейные зависимости. В общем случае нам требуется более тонкие инструменты на большем числе данных.

Напомним о случайных величинах и их распределениях

Пусть в вероятностном пространстве $(\Omega, \mathcal{F}, \mathbb{P})$ задана дискретная случайная величина $\xi : \Omega \rightarrow \mathbb{R}$ с распределением

x_1	x_2	x_3	x_4	\dots
p_1	p_2	p_2	p_4	\dots

Здесь $p_i \triangleq \mathbb{P}(\xi = x_i) = \mathbb{P}\{\omega \in \Omega \mid \xi(\omega) = x_i\}$.

Для мотивации: полиномиальное распределение

Зафиксируем $k, n \in \mathbb{N}$. Примем $\Omega \triangleq \{1, 2, \dots, k\}^n$, то есть элементарные события — строки из n символов алфавита $\{1, 2, \dots, k\}$. Примем вероятность каждого символа $i \in \{1, 2, \dots, k\}$, равной p_i ; пусть, если в строке каждый символ i встречается ровно n_i раз, то вероятность этой строки примем равной $\prod_{i=1}^k p_i^{n_i}$. Тогда, для каждого набора (n_1, n_2, \dots, n_k) со свойствами $n_i \geq 0, n_1 + \dots + n_k = n$ вероятность того, что строка будет иметь ровно n_i вхождений для каждого символа i , равна

$$n! \prod_{i=1}^k \frac{p_i^{n_i}}{n_i!} = n! \frac{p_1^{n_1} p_2^{n_2} \dots p_k^{n_k}}{n_1! n_2! \dots n_k!}.$$

Подумать: нужно ли гарантировать, что $p_1 + p_2 + \dots + p_k = 1$?

Подумать: соответствует ли это какому-то дискретному распределению, а для какой дискретной случайной величины?

Табличка для $k = n = 3$, $(p_1, p_2, p_3) = (1/2, 1/3, 1/6)$

	$n_1 = 0$	$n_1 = 1$	$n_1 = 2$	$n_1 = 3$
$n_2 = 0$	$\frac{1}{6^3}$	$\frac{3}{2 \cdot 6^2}$	$\frac{3}{2^2 \cdot 6}$	$\frac{1}{2^3}$
$n_2 = 1$	$\frac{3}{3 \cdot 6^2}$	$\frac{3 \cdot 2}{2 \cdot 3 \cdot 6}$	$\frac{3}{2^2 \cdot 3}$	0
$n_2 = 2$	$\frac{3}{3^2 \cdot 6}$	$\frac{3}{2 \cdot 3^2}$	0	0
$n_2 = 3$	$\frac{1}{3^3}$	0	0	0

Подумать: это точно та табличка? Составьте табличку для (n_1, n_3) .

Векторные случайные величины

Пусть $m \in \mathbb{N}$.

В заданном измеримом пространстве (Ω, \mathcal{F}) отображение $\vec{\xi} : \Omega \rightarrow \mathbb{R}^m$ называется **дискретной векторной (m -мерной) случайной величиной**, если оно принимает не более чем счетное число значений $\vec{x}_1, \dots, \vec{x}_i, \dots$, и при этом для всякого \vec{x}_i выполнено

$$\{\omega \in \Omega \mid \vec{\xi}(\omega) = \vec{x}_i\} \in \mathcal{F}.$$

Подумать: какие можно ввести операции над такими величинами?

Подумать: чему равно m на предыдущем слайде?

Подумать: пусть $\vec{\xi}$ — векторная случайная величина; верно ли, что первая компонента вектора также (скалярная) случайная величина?

Функция распределения

Пусть также задана вероятность \mathbb{P} .

Совместной функцией распределения векторной дискретной случайной величины $\vec{\xi} = (\xi_1, \dots, \xi_m)$ называют отображение $F_{\vec{\xi}}: \mathbb{R}^m \rightarrow [0, 1]$, заданное по правилу:

$$F_{\vec{\xi}}(\vec{a}) \triangleq \mathbb{P}\{\omega \mid \forall k \in \{1, \dots, m\} \ \xi_k(\omega) \leq a_k\} \quad \forall \vec{a} = (a_1, \dots, a_m) \in \mathbb{R}^m.$$

Предупреждение: как уже говорилось, иногда в определении функции распределения пишут строгое неравенство. При этом часть свойств, по большому счету аналогичных, меняется, причем для многомерных распределений отследить эти изменения существенно сложнее...

Дискретные распределения

Многомерным дискретным распределением называют отображение из счетного подмножества множества \mathbb{R}^m в $[0, 1]$, сумма элементов образа которого равна единице. Совместным распределением случайных величин ξ_1, \dots, ξ_m (распределением в.с.в. $\vec{\xi}$) называют многомерный массив вероятностей векторов $(\xi_1(\omega), \dots, \xi_m(\omega))$.

Совместное распределение двумерной случайной величины $\vec{\xi} = (Y, X)$:

	x_1	x_2	x_3	...
y_1	$p_{11} \triangleq \mathbb{P}(\vec{\xi} = (y_1, x_1))$	$p_{12} \triangleq \mathbb{P}(\vec{\xi} = (y_1, x_2))$	$p_{13} \triangleq \mathbb{P}(\vec{\xi} = (y_1, x_3))$...
y_2	$p_{21} \triangleq \mathbb{P}(\vec{\xi} = (y_2, x_1))$	$p_{22} \triangleq \mathbb{P}(\vec{\xi} = (y_2, x_2))$	$p_{23} \triangleq \mathbb{P}(\vec{\xi} = (y_2, x_3))$...
y_3	$p_{31} \triangleq \mathbb{P}(\vec{\xi} = (y_3, x_1))$	$p_{32} \triangleq \mathbb{P}(\vec{\xi} = (y_3, x_2))$	$p_{33} \triangleq \mathbb{P}(\vec{\xi} = (y_3, x_3))$...
...

Перед Вами двумерное дискретное распределение для $\vec{\xi} = (Y, X)$

(Y, X)	$X = 0$	1	2	3
$Y = 0$	$\frac{1}{6^3}$	$\frac{3}{2 \cdot 6^2}$	$\frac{3}{2^2 \cdot 6}$	$\frac{1}{2^3}$
1	$\frac{3}{3 \cdot 6^2}$	$\frac{3 \cdot 2}{2 \cdot 3 \cdot 6}$	$\frac{3}{2^2 \cdot 3}$	0
2	$\frac{3}{3^2 \cdot 6}$	$\frac{3}{2 \cdot 3^2}$	0	0
3	$\frac{1}{3^3}$	0	0	0

Найдите распределение случайной величины Y .

Перед Вами двумерное дискретное распределение

(Y, X)	$X = 0$	1	2	3	
$Y = 0$	$\frac{1}{6^3}$	$\frac{3}{2 \cdot 6^2}$	$\frac{3}{2^2 \cdot 6}$	$\frac{1}{2^3}$	$\frac{2^3}{3^3}$
1	$\frac{3}{3 \cdot 6^2}$	$\frac{3 \cdot 2}{2 \cdot 3 \cdot 6}$	$\frac{3}{2^2 \cdot 3}$	0	$\frac{2^2 \cdot 3}{3^3}$
2	$\frac{3}{3^2 \cdot 6}$	$\frac{3}{2 \cdot 3^2}$	0	0	$\frac{2 \cdot 3}{3^3}$
3	$\frac{1}{3^3}$	0	0	0	$\frac{1}{3^3}$

Найдите распределение случайной величины X .

Перед Вами двумерное дискретное распределение вместе со всеми его маргиналами:

(Y, X)	$X = 0$	1	2	3	
$Y = 0$	$\frac{1}{6^3}$	$\frac{3}{2 \cdot 6^2}$	$\frac{3}{2^2 \cdot 6}$	$\frac{1}{2^3}$	$\frac{2^3}{3^3}$
1	$\frac{3}{3 \cdot 6^2}$	$\frac{3 \cdot 2}{2 \cdot 3 \cdot 6}$	$\frac{3}{2^2 \cdot 3}$	0	$\frac{2^2 \cdot 3}{3^3}$
2	$\frac{3}{3^2 \cdot 6}$	$\frac{3}{2 \cdot 3^2}$	0	0	$\frac{2 \cdot 3}{3^3}$
3	$\frac{1}{3^3}$	0	0	0	$\frac{1}{3^3}$
	$\frac{1}{2^3}$	$\frac{3}{2^3}$	$\frac{3}{2^3}$	$\frac{1}{2^3}$	

Маргинальные распределения

Для каждой векторной случайной величины $\vec{\xi} = (\xi_1, \dots, \xi_m)$ распределения одной или совместное распределение нескольких (но не всех) ее компонент называют ее **маргинальными распределениями** [от слова *margin* — поле]. Распределения компонент векторной случайной величины — скалярных случайных величин ξ_1, \dots, ξ_m — иногда называют **частными распределениями**.

Подумать: сколько маргинальных распределений у трехмерной случайной величины? Укажите их для полиномиального распределения выше.

Подумать: можно ли восстановить совместное распределение случайной величины $\vec{\xi}$ по всем ее маргинальным распределениям?

Подумать: можно ли для трехмерной случайной величины $\vec{\xi} = (X, Y, Z)$, зная лишь маргинальное распределение (X, Y) , восстановить распределения X и Z ?

Перед Вами двумерное дискретное распределение

(Y, X)	$X = 0$	1	2	3	
$Y = 0$	$\frac{1}{6^3}$	$\frac{3}{2 \cdot 6^2}$	$\frac{3}{2^2 \cdot 6}$	$\frac{1}{2^3}$	$\frac{2^3}{3^3}$
1	$\frac{3}{3 \cdot 6^2}$	$\frac{3 \cdot 2}{2 \cdot 3 \cdot 6}$	$\frac{3}{2^2 \cdot 3}$	0	$\frac{2^2 \cdot 3}{3^3}$
2	$\frac{3}{3^2 \cdot 6}$	$\frac{3}{2 \cdot 3^2}$	0	0	$\frac{2 \cdot 3}{3^3}$
3	$\frac{1}{3^3}$	0	0	0	$\frac{1}{3^3}$
	$\frac{1}{2^3}$	$\frac{3}{2^3}$	$\frac{3}{2^3}$	$\frac{1}{2^3}$	

Подумать: найти распределение случайной величины $X + Y$, как его посчитать, можно ли в данном случае указать его без вычислений?

Подумать: можно ли вычислить распределение случайной величины $X + Y$, зная лишь частные распределения X, Y , но не совместное распределение векторной случайной величины $\vec{\xi} = (Y, X)$?

Пусть известны одномерные распределения. Найдется ли векторная случайная величина, для которой эти распределения — маргинальные?

(Y, X)	x_1	x_2	x_3	\dots	
y_1	?	?	?	\dots	q_1
y_2	?	?	?	\dots	q_2
y_3	?	?	?	\dots	q_3
\dots	\dots	\dots	\dots	\dots	\dots
	p_1	p_2	p_3	\dots	

Совместное распределение с заданными маргинальными распределениями:

(Y, X)	x_1	x_2	x_3	\dots	
y_1	q_1p_1	q_1p_2	q_1p_3	\dots	q_1
y_2	q_2p_1	q_2p_2	q_2p_3	\dots	q_2
y_3	q_3p_1	q_3p_2	q_3p_3	\dots	q_3
\dots	\dots	\dots	\dots	\dots	\dots
	p_1	p_2	p_3	\dots	

Теорема. [С-но] Распределение векторной случайной величины $\vec{\xi} = (\xi_1, \xi_2, \dots, \xi_m)$ однозначно восстанавливается по своим одномерным маргинальным распределениям тогда и только тогда, когда случайные величины $\xi_1, \xi_2, \dots, \xi_m$ независимы в совокупности. В этом случае, например, при $m = 2$:

(Y, X)	x_1	x_2	x_3	\dots	
y_1	q_1p_1	q_1p_2	q_1p_3	\dots	q_1
y_2	q_2p_1	q_2p_2	q_2p_3	\dots	q_2
y_3	q_3p_1	q_3p_2	q_3p_3	\dots	q_3
\dots	\dots	\dots	\dots	\dots	\dots
	p_1	p_2	p_3	\dots	

Задача Монжа-Канторовича

Известны распределения случайных величин X, Y и неотрицательные числа c_{ij} . Найдите среди всех совместных распределений

(Y, X)	x_1	x_2	x_3	\dots	
y_1	p_{11}	p_{12}	p_{13}	\dots	q_1
y_2	p_{21}	p_{22}	p_{23}	\dots	q_2
\dots	\dots	\dots	\dots	\dots	\dots
	p_1	p_2	p_3	\dots	

с заданными маргинальными распределениями X, Y то, для которого минимальна сумма

$$\sum_{i,j} c_{ij} p_{ij}.$$

Задача Монжа-Канторовича и логистика

Известны распределения X, Y и положительные числа c_{ij} . Найдите среди всех совместных распределений с заданными маргинальными распределениями X, Y то, для которого минимальна сумма

$$\sum_{i,j} c_{ij} p_{ij}.$$

Пусть D — минимальная сумма в этой задаче.

Логистическая интерпретация. Пусть известны распределение X некоторого товара по складам, распределение Y его потребности в магазинах и стоимость c_{ij} перевозки (всего товара целиком) из i -го склада в j -й магазин. Перебирая p_{ij} — доли товара, перевезенного из i -го склада в j -й магазин, — находим минимальную стоимость перевозки D .

Метрики на пространстве вероятностей

Пусть Ω — нормированное пространство, $r \geq 1$. Найдите среди всех совместных распределений с заданными маргинальными распределениями X, Y то, для которого минимальна сумма

$$\sum_{i,j} p_{ij} \|x_i - y_j\|^r.$$

Расстояние между распределениями X и Y можно ввести как

$$d_1(X, Y) \triangleq \min \sum_{i,j} p_{ij} \|x_i - y_j\| \quad (\text{метрика Канторовича}),$$

$$d_2(X, Y) \triangleq \min \sqrt{\sum_{i,j} p_{ij} \|x_i - y_j\|^2} \quad (\text{метрика Васерштейна}).$$

[С-но; $\geq 1,5$ балла] Дана цепь Маркова с матрицей переходов Q и длинами ребер $c_{ij} = c_{ji} \geq 0$, $c_{ii} = 0$. Докажите, что если

$$q \triangleq \max_{i \neq j} \frac{d_1(e_i Q, e_j Q)}{c_{ij}} < 1, \text{ то}$$

1) инвариантное распределение π единственно;

2) $d_1(\mu Q^n, \pi) \leq q^n d_1(\mu, \pi)$ для любого начального распределения μ . 

Энтропия

Пусть есть некоторая случайная величина ξ с распределением

x_1	x_2	x_3	x_4	\dots
p_1	p_2	p_3	p_4	\dots

число $H(\xi) \triangleq -\sum_i p_i \log p_i$ называют **энтропией** этого распределения (этой случайной величины).

Для многомерного распределения энтропия вводится также, например для двухмерного случая с $\vec{\xi} = (Y, X)$: $H(\vec{\xi}) \triangleq -\sum_{i,j} p_{ij} \log p_{ij}$.

Терминологическое замечание. Логарифм берут иногда двоичный, иногда — натуральный. Используется также обозначение $Ent X$.

Свойства энтропии [с-но]

$$1^0 \ H(X) \geq 0,$$

причем $H(X) = 0$ тогда и только тогда, когда X — вырожденная случайная величина;

$$2^0 \ H(f(X)) \leq H(X) \text{ для любой функции}$$

$$f: \{x_1, x_2, \dots\} \rightarrow \{x_1, x_2, \dots\},$$

причем $H(f(X)) = H(X)$ тогда и только тогда, когда f — биекция;

$$3^0 \ H(X) \leq \log N, \text{ если } X \text{ принимает не более чем } N \text{ значений,}$$

причем $H(X) = \log N$ только при $p_i = 1/N$;
[воспользоваться 12⁰-м свойством матожидания]

$$4^0 \ [0,5 \text{ балла}] \text{ для } \vec{\xi} = (X, Y) \text{ выполнено}$$

$H(\vec{\xi}) \leq H(X) + H(Y)$, причем $H(\vec{\xi}) = H(X) + H(Y)$ в точности для независимых X, Y .

Теорема Макмиллана [пока без д-ва]

Пусть имеется последовательность независимых в совокупности, одинаково распределенных дискретных случайных величин $\xi_i : \Omega \rightarrow K$, где множество K конечно. Тогда найдется такая последовательность множеств S_n из строк длины n , то есть $S_n \subset K^n$, что

- 1) $\lim_{n \rightarrow \infty} \mathbb{P}((\xi_1, \dots, \xi_n) \in S_n) = 1,$
- 2) $\lim_{n \rightarrow \infty} \frac{\log |S_n|}{n} = H(\xi_1),$
- 3) $\lim_{n \rightarrow \infty} \sup_{s \in S_n} \left| \frac{\log \mathbb{P}((\xi_1, \dots, \xi_n) = s)}{n} + H(\xi_1) \right| = 0.$

Доказательство имеется, например, в [Синай, Теорема 2.6].

Замечание. Теорема верна и в случае, если конечна $\sum_i p_i \log^2 p_i$, например для геометрического распределения.

Подумать: оценить рост S_n , если $\xi_i \in \text{Geom}(1/2)$, если $\xi_i \in U\{1, 10\}$.

Условные энтропия и информация

Пусть $\vec{\xi} \triangleq (Y, X)$, тогда **условная энтропия** $H(X|Y)$ **случайной величины** Y **относительно** X и **совместная информация** $I(X, Y)$ случайных величин X, Y вводятся по правилам:

$$H(X|Y) \triangleq H(\vec{\xi}) - H(Y), \quad I(X, Y) \triangleq H(Y) + H(X) - H(\vec{\xi}).$$

Очевидные из предыдущего свойства:

$$H(X) = H(X|Y) + I(X, Y),$$

$$H(X|Y) \geq 0, \quad I(X, Y) = I(Y, X) \geq 0,$$

$$H(X|Y) = 0 \Leftrightarrow X = f(Y) \text{ для некоторого } f: \Omega \rightarrow \Omega,$$

$$I(X, Y) = 0 \Leftrightarrow H(X|Y) = H(X) \Leftrightarrow X, Y \text{ — независимы.}$$

Подумать: в общем случае, $H(X|Y) \neq H(Y|X)$, хотя $I(X, Y) = I(Y, X)$ всегда.

Посчитаем условную энтропию

$$\begin{aligned}H(X|Y) &= H(\vec{\xi}) - H(Y) \\&= -\sum_{i,j} p_{ij} \log p_{ij} + \sum_j q_j \log q_j \\&= -\sum_{i,j} p_{ij} \log p_{ij} + \sum_{i,j} p_{ij} \log q_j \\&= -\sum_{i,j} p_{ij} \log \frac{p_{ij}}{q_j} = -\sum_j q_j \sum_i \frac{p_{ij}}{q_j} \log \frac{p_{ij}}{q_j} \\&= -\sum_j q_j \sum_i \mathbb{P}(X = x_i | Y = y_j) \log \mathbb{P}(X = x_i | Y = y_j).\end{aligned}$$

Напомним, что $\mathbb{P}(X|Y)$ — условная вероятность случайной величины X относительно Y .

Условное распределение случайной величины X относительно Y :

при	x_1	x_2	x_3	...
$y_1 :$	$\mathbb{P}(X = x_1 Y = y_1) = \frac{p_{11}}{q_1}$	$\mathbb{P}(X = x_2 Y = y_1) = \frac{p_{12}}{q_1}$	$\mathbb{P}(X = x_3 Y = y_1) = \frac{p_{13}}{q_1}$...
$y_2 :$	$\mathbb{P}(X = x_1 Y = y_2) = \frac{p_{21}}{q_2}$	$\mathbb{P}(X = x_2 Y = y_2) = \frac{p_{22}}{q_2}$	$\mathbb{P}(X = x_3 Y = y_2) = \frac{p_{23}}{q_2}$...
$y_3 :$	$\mathbb{P}(X = x_1 Y = y_3) = \frac{p_{31}}{q_3}$	$\mathbb{P}(X = x_2 Y = y_3) = \frac{p_{32}}{q_3}$	$\mathbb{P}(X = x_3 Y = y_3) = \frac{p_{33}}{q_3}$...
...

Что дальше?

- распределение случайных величин
- медиана и математическое ожидание
- независимость случайных величин
- производящие функции как матожидание
- дисперсия, ковариация и корреляция
- векторные случайные величины и их распределения
- энтропия и условная энтропия
- **условное математическое ожидание**

На пять минут...

1. Буфет заказывает пирожки. Распределение случайной величины “число страждущих пирожка в такой день” известно. Непроданный пирожок дает убыток 15 рублей; каждый страждущий пирожка, но ушедший без него, — тоже убыток, но 5 рублей. Сколько пирожков заказать для минимизации убытка?
2. Известно, что ξ и ξ независимы. Найдите все такие случайные величины ξ .
3. Известно, что для $\xi : \Omega \rightarrow [-1, 1]$ выполнено $\mathbb{D}\xi = 1$. Найдите распределение ξ .
4. Введите над \mathbb{N} хотя бы полуалгебру Π , определите на Π конечно-аддитивную функцию из Π в $\{0, 1\}$, не являющуюся счетно-аддитивной (верхний и нижний пределы Вам в помощь...).