Математические основы искусственного интеллекта
Дискретные случайные величины

Солодушкин Святослав Игоревич

Кафедра вычислительной математики и компьютерных наук,
УрФУ имени первого Президента России Б.Н. Ельцина

Октябрь 2021

«Фильм должен начинаться с землетрясения, а потом напряжение должно нарастать»

Альфред Хичкок

Пусть задано вероятностное пространство $\langle \Omega, \mathcal{F}, \mathrm{P} \rangle$.

Определение

Функция $\xi : \Omega \to \mathbb{R}$ называется случайной величиной, если для любого борелевского множества $B \in \mathfrak{B}(\mathbb{R})$ множество $\xi^{-1}(B)$ является событием, т. е. принадлежит σ-алгебре $\mathcal{F}$.

Вводный пример

В группе 6 студентов: Анна, Берта, Вера, Геннадий, Дмитрий и Евргаф. Одного из них случайно по жеребьевке отправляют на стажировку. Какова вероятность того, что поедет Берта? Какова вероятность того, что поедет мужчина?

Василий ищет грибы в лесу, где равновероятно встречатся белые, подосиновики, маслята, поганки, мухоморы, чешуйницы. Какова вероятность того, что Василий найдет подосиновик? Какова вероятность того, что Василий надет ядовитый гриб?

Бросают игральную кость. Какова вероятность того, что выпадет два? Какова вероятность того, что выпадет число большее трех?

Эти три примера показывают, что для многих экспериментов нет никаких различий в подсчете вероятностей событий, в то время как сами элементарные исходы в этих экспериментах сильно различаются. Однако нас и должны интересовать именно вероятности событий, а не структура пространства элементарных исходов.

В свзи с этим целесообразно во всех таких «похожих» экспериментах вместо самых разных исходов использовать, например, числа. Иначе говоря, каждому элементарному исходу поставить в соответствие некоторое вещественное число, и работать только с числами.

Неформальное «определение» случайной величины

Величина, значение котрой заранее неизвестно и зависит от исхода случайного эксперимента, называется случайной...

Если задана случайная величина ξ, нам может потребоваться вычислить вероятности вида

$$\mathsf{P}(\xi = 1) = \mathsf{P}\{\omega \,|\, \xi(\omega) = 1\},$$

$\mathsf{P}(\xi \geq 4)$, $\mathsf{P}(\xi \in [-1, 1])$, или даже $\mathsf{P}(\xi \in \mathbb{Q})$
и вообще самые разные вероятности попадания в борелевские множества на прямой.

Это возможно лишь если множества, стоящие под знаком вероятности, являются событиями, так как вероятность есть функция, определенная только на σ-алгебре событий.

Следовательно, нельзя называть случайной величиной ЛЮБОЕ отображение $\xi : \Omega \to \mathbb{R}$. Можно рассматривать только такие отображение $\xi : \Omega \to \mathbb{R}$, что прообраз любого борелевского множества B является событием $\xi^{-1} \in \mathcal{F}$.

Определение случаной величины

Пусть задано вероятностное пространство $\langle \Omega, \mathcal{F}, \mathrm{P} \rangle$.

Определение

Функция $\xi : \Omega \to \mathbb{R}$ называется случайной величиной, если для любого борелевского множества $B \in \mathfrak{B}(\mathbb{R})$ множество $\xi^{-1}(B)$ является событием, т. е. принадлежит σ-алгебре $\mathcal{F}$.

А можно потребовать в определении чего-нибудь попроще? Например, чтобы событием было попадание в любой интервал: $\{\omega | \xi(\omega) \in (a, b)\} \in \mathcal{F}$, или в полуинтервал: $\{\omega | \xi(\omega) < x\} \in \mathcal{F}$.

Определение

Функция $\xi : \Omega \to \mathbb{R}$ называется случайной величиной, если для любых вещественных $a < b$ множество $\{\omega \,|\, \xi(\omega) \in (a, b)\}$ принадлежит σ-алгебре $\mathcal{F}$.

Определения эквивалентны.

Определение

Распределением случайной величины ξ называется вероятностная мера $\mu(B) = \mathsf{P}(\xi \in B)$ на множестве борелевских подмножеств $\mathbb{R}$.

Событие	Орел	Решка
P	0.5	0.5
ξ	0	1
η	1	2

Вопрос: ξ и η одинаково распределены или нет?

P	0.5	0.5
ξ	0	1

P	0.1	0.9
η	0	1

Вопрос: ξ и η одинаково распределены или нет?

Определение

Распределением случайной величины ξ называется вероятностная мера $\mu(B) = \mathsf{P}(\xi \in B)$ на множестве борелевских подмножеств $\mathbb{R}$.

Событие	Орел	Решка
P	0.5	0.5
ξ	0	1
η	1	0

Для любого множества $B \subseteq \mathbb{R}$ вероятности принадлежать B для ξ и для η одинаковы. Однако, ни для одного элементарного исхода ω значения $\xi(\omega)$ и $\eta(\omega)$ не совпадают. То есть, ξ и η одинаково распределены, но не одинаковы как функции.

Событие	Не сдал экзамен	Сдал экзамен
P	0.5	0.5
ζ	0	1

Определение

Распределением случайной величины ξ называется вероятностная мера $\mu(B) = \mathsf{P}(\xi \in B)$ на множестве борелевских подмножеств $\mathbb{R}$.

Распределения случайных величин — это основные объекты изучения в теории вероятностей.

Мы не будем, как правило, интересоваться тем, из какого множества действует функция ξ и каким именно элементарным исходам сопоставляет свои возможные значения. Нас будет интересовать лишь, с какой вероятностью эти значения принимаются.

Приведите несколько примеров совершенно разных случайных величин, имеющих одно и то же распределение (одинаково распределенных).

Типы распределений случайных величин

Опишем различные типы распределений случайных величин.

Вся вероятностная масса может быть сосредоточена в нескольких точках прямой, а может быть «размазана» по некоторому интервалу или по всей прямой.

В зависимости от типа множества, на котором сосредоточена вся единичная вероятностная масса, распределения делят на дискретные, абсолютно непрерывные, сингулярные и их (линейные) комбинации.

Случайная величина ξ имеет дискретное распределение, если существует конечный или счетный набор чисел $a_1, a_2, \ldots$ такой, что

$$\mathrm{P}(\xi = a_i) > 0 \text{ для всех } i, \qquad \sum_{i=1}^{\infty} \mathrm{P}(\xi = a_i) = 1.$$

Итак, случайная величина ξ имеет дискретное распределение, если она принимает не более чем счетное число значений. Значения эти иначе называют атомами: ξ имеет атом в точке x, если $\mathrm{P}(\xi = x) > 0$.

Если случайная величина ξ имеет дискретное распределение, то для любого $B \subseteq \mathbb{R}$

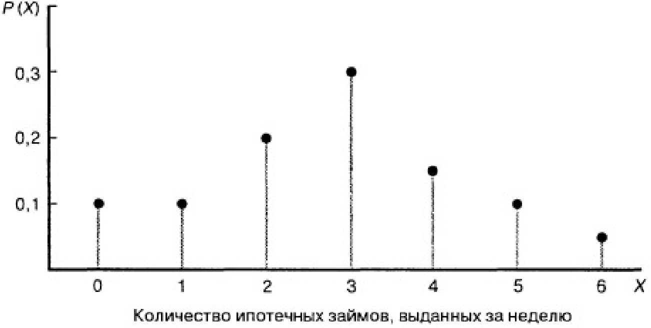
$$\mathrm{P}(\xi \in B) = \sum_{a_i \in B} \mathrm{P}(\xi = a_i).$$

Дискретное распределение удобно задавать следующей таблицей, в которой $p_i = \mathrm{P}(\xi = x_i)$:

ξ	x_1	x_2	x_3	$\dots$
P	p_1	p_2	p_3	$\dots$

График дискретного распределения

Количество ипотечных займов, выданных за неделю	Вероятность
0	0,10
1	0,10
2	0,20
3	0,30
4	0,15
5	0,10
6	0,05



Непрерывное распределение

Определение

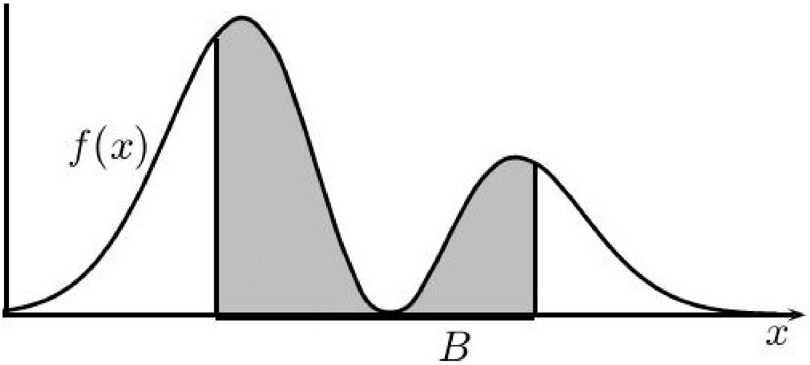
Случайная величина ξ имеет абсолютно непрерывное распределение, если существует неотрицательная функция $f_\xi(x)$ такая, что для любого борелевского множества B имеет место равенство:

$$\mathrm{P}(\xi \in B) = \int\limits_B f_\xi(x)\,dx.$$

Функцию $f_\xi(x)$ называют плотностью распределения величины ξ.

Замечание. Здесь интеграл понимается в смысле Лебега, а не Римана.

График непрерывного распределения



Случайная величина ξ имеет распределение Бернулли с параметром p, если ξ принимает значения 1 и 0 с вероятностями p и $1-p$ соответственно.

Случайная величина ξ с таким распределением равна числу успехов в одном испытании схемы Бернулли с вероятностью успеха p: ни одного успеха или один успех.

Таблица распределения ξ имеет вид:

ξ	0	1
P	$1-p$	p

.

Случайная величина ξ имеет биномиальное распределение с параметрами $n \in \mathbb{N}$ и $p \in (0, 1)$, если ξ принимает значения $k = 0, 1, \ldots, n$ с вероятностями $\mathrm{P}(\xi = k) = C_n^k p^k (1-p)^{n-k}$.

Случайная величина с таким распределением имеет смысл числа успехов в n испытаниях схемы Бернулли с вероятностью успеха p.

Таблица распределения ξ имеет вид

ξ	0	1	$\ldots$	k	$\ldots$	n
P	$(1-p)^n$	$np(1-p)^{n-1}$	$\ldots$	$C_n^k p^k (1-p)^{n-k}$	$\ldots$	p^n

.

Биномиальное распределение. Задача

В среднем из 10 клиентов подписывают контракт 4. В течение дня придут 8 для обсуждения контракта. Случайная величина ξ — число клиентов, подписавших контракт.
Составить распределение случайной величины ξ.
Какова вероятность того, что не менее двух клиентов заключат контракт?

Распределение Пуассона

Случайная величина ξ имеет распределение Пуассона с параметром $\lambda > 0$, если ξ принимает значения $k = 0, 1, 2, \ldots$ с вероятностями $\mathsf{P}(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda}$.

Распределение Пуассона возникло в теореме Пуассона как предельное распределение для числа успехов в n испытаниях схемы Бернулли, когда число испытаний n увеличивается, а вероятность успеха уменьшается обратно пропорционально n. Поэтому распределение Пуассона называют иначе распределением числа редких событий.

Послано 5000 пакетов данных. Вероятность того, что данные в пакете будут повреждены, равна 0.0002. Найти вероятность того, что в трех пакетах данные повредятся.

$$P_n(k) = \frac{n(n-1)\dots(n-(k-1))}{k!} p^k (1-p)^{n-k}.$$

Обозначим $\lambda = pn$, тогда

$$P_n(k) = \frac{n(n-1)\dots(n-(k-1))}{k!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}.$$

$$P_n(k) \approx \lim_{n\to\infty} \frac{n(n-1)\dots(n-(k-1))}{k!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} =$$

$$= \frac{\lambda^k}{k!} \lim_{n\to\infty} \frac{n(n-1)\dots(n-(k-1))}{n^k} \left(1 - \frac{\lambda}{n}\right)^{n} \left(1 - \frac{\lambda}{n}\right)^{-k} = \frac{\lambda^k}{k!} e^{-\lambda}.$$

По условию имеем $n = 5000$, $p = 0.0002$, $k = 3$. Найдем λ :

$$\lambda = np = 5000 \cdot 0.0002 = 1.$$

По формуле Паусcона искомая вероятность приближенно равна

$$P_{5000}(3) = \lambda^k e^{-\lambda}/k! = e^{-1}/3! \approx 0.06.$$

Геометрическое распределение

Случайная величина ξ имеет геометрическое распределение с параметром $p \in (0, 1)$, если ξ принимает значения $k = 1, 2, 3, \ldots$ с вероятностями $\mathsf{P}(\xi = k) = p(1-p)^{k-1}$.

Случайная величина с таким распределением имеет смысл номера первого успешного испытания в схеме Бернулли с вероятностью успеха p. Таблица распределения ξ имеет вид

ξ	1	2	$\ldots$	k	$\ldots$
P	p	$p(1-p)$	$\ldots$	$p(1-p)^{k-1}$	$\ldots$

.

1. Известно, что при выпечке 1000 сладких булочек с изюмом в тесто поладили 10000 изюмин и тщательно перемешали. Какова вероятность того, что в случайно выбранной булочке изюма не окажется?
2. Сколько изюма в предыдущей задаче пекарь можно забрать себе и спать спокойно, зная, что, когда придет проверка и возьмет случайную булочку, с вероятностью 0.999 хотя бы одна изюминка в булочке окажется?
3. При прохождении одного порога байдарка не получает повреждений с вероятностью p_1, полностью ломается с вероятностью p_2, получает серьезные повреждения с вероятностью p_3 ($p_1 + p_2 + p_3 = 1$). Два серьезных повреждения приводят к поломке. Найти вероятность того, что при прохождении n порогов байдарка не будет полностью сломана.

Схему независимых испытаний мы формализуем следующим образом. Всего будет $n = 10000$ испытаний (по числу изюмин). Испытание будет состоять в том, что мы определяем, попала ли очередная изюмина в нашу случайно выбранную булочку.
Поскольку всего булочек 1000, вероятность того, что конкретная изюмина попала именно в нашу булочку, есть $p = 1/1000$. Применяем распределение Пуассона с параметром $\lambda = np = 100001/000 = 10$. Получим:

$$P_{10000}(k = 0) \approx 10^k/k!e^{-10} \approx 4.54\,10^{-5}.$$