Дискретные и непрерывные случайные величины. Условная вероятность.

Леонид Иосипой

Курс «Вероятностные модели и статистика»
Центр непрерывного образования, ВШЭ

1 апреля 2021

Организационная информация

1. О себе
2. Время занятий/Перерывы
3. Теория/Практика/Python
4. Домашние задания/Google Classroom
5. Вопросы по курсу/Обратная связь
6. Сайт

Немного истории

В 1900 г. на Международном Математическом Конгрессе в Париже Давид Гильберт сформулировал проблемы, стоящие перед математикой XX века.

Шестая из 23 знаменитых ныне проблем Гильберта касалась теории вероятностей: «аксиоматизировать те науки, в которых математика играет важную роль, в первую очередь теорию вероятностей и механику».

Эта проблема (в части, касающейся теории вероятностей) была окончательно решена Андреем Николаевичем Колмогоровым в 1933 г.

Случайные величины

Теория вероятностей изучает закономерности, возникающие в случайных экспериментах (точнее, в их математических моделях).

Случайные эксперименты в теории вероятностей моделируются с помощью случайных величин.

Случайные величины

Случайной величиной называется функция, которая каждому возможному исходу в эксперименте ставит в соответствие действительное число. Можно понимать случайную величину как «кодирование» исходов эксперимента.

Нас будет интересовать множество значений случайной величины, то есть с какими вероятностями случайная величина принимает какие-то значения.

Случайные величины

Мы будем говорить о вероятности событий, которые связаны со случайными величинами. Например, если X — случайная величина, то нас могут интересовать вероятности событий:

- $\{X = a\}$ для некоторого $a \in \mathbb{R}$;
- $\{X < a\}$ для некоторого $a \in \mathbb{R}$;
- $\{a \leq X < b\}$ для некоторых $a, b \in \mathbb{R}$, $a < b$.

В самом общем виде любое событие можно записать так:

- $\{X \in A\}$ для некоторого подмножества $A \subset \mathbb{R}$.

Случайные величины

Кроме того, мы будем классифицировать случайные величины по мощности их множества значений.

- Дискретные случайные величины имеют конечное или счетное множество значений. Примеры: $\{1, \ldots, n\}$, $\mathbb{N}$, $\mathbb{Z}$.
- Непрерывные случайные величины имеют несчетное множество значений. Примеры: $[0, 1]$, $\mathbb{R}$.

Дискретные случайные величины

Начнем изучение случайных величин с дискретных моделей.

Как определить дискретную случайную величину?

Чтобы задать распределение дискретной случайной величины, необходимо всем возможным значениям этой случайной величины приписать вероятности. Подробнее:

- перечислить возможные значения $a_1, a_2, a_3, \ldots$;
- задать последовательность действительных чисел $p_1, p_2, p_3, \ldots$, таких что
  1) $p_i \geq 0$ (неотрицательность);
  2) $\sum_{i=1}^{\infty} p_i = 1$ (нормировка);
- положить $\mathbb{P}(X = a_i) = p_i$.

Дискретные случайные величины

Это удобно делать с помощью таблицы:

X	a_1	a_2	a_3	$\ldots$
$\mathbb{P}$	p_1	p_2	p_3	$\ldots$

Такую таблицу мы будем называть таблицей распределения вероятностей X.

Дискретные случайные величины

Если задано распределение дискретной величины, то вероятность любых событий можно посчитать по формуле:

$$\mathbb{P}(X \in A) = \sum_{i\colon a_i \in A} p_i.$$

То есть мы должны просуммировать все вероятности, которые соответствуют исходам, входящим в A.

Примеры дискретных случайных величин

0. Константа.

   Случайная величина X принимает лишь одно значение $c \in \mathbb{R}$ с вероятностью 1, то есть $\mathbb{P}(X = c) = 1$.

   Таблица распределения имеет следующий вид:

X	c
$\mathbb{P}$	1

Примеры дискретных случайных величин

1. Распределение Бернулли $\mathbf{B}_p$, $p \in [0, 1]$.

Случайная величина X имеет распределение Бернулли с параметром $p \in [0, 1]$, если X принимает значение 1 с вероятностью p и значение 0 с вероятностью $1 - p$.

Таблица распределения имеет следующий вид:

X	0	1
$\mathbb{P}$	$1 - p$	p

Параметр p называется вероятностью «успеха».
Будем писать $X \sim \mathbf{B}_{1/2}$, $X \sim \mathbf{B}_{1/4}$, $X \sim \mathbf{B}_1$, ...

Примеры дискретных случайных величин

1. Распределение Бернулли $\mathbf{B}_p$, $p \in [0,1]$.

   Примеры:

   - Орел/Решка;
   - Выиграл/Проиграл;
   - Купил/Не купил;
   - Попал в цель/Не попал в цель;
   - Мальчик/Девочка;
   - Любое другое бинарное отношение.

Примеры дискретных случайных величин

2. Равномерное распределение на конечном множестве.

Случайная величина X имеет равномерное распределение на множестве $\{a_1, a_2, \ldots, a_n\}$, если X принимает каждое значение a_i с вероятностью $1/n$.

Таблица распределения имеет следующий вид:

X	a_1	a_2	$\ldots$	a_n
$\mathbb{P}$	$1/n$	$1/n$	$\ldots$	$1/n$

Примеры дискретных случайных величин

2. Равномерное распределение на конечном множестве.

Примеры:

- Игральная кость;
- Игра в рулетку;
- Принадлежность какой-то группе;
- Априорное распределение на параметры модели в условиях отсутствия априорной информации.

Примеры дискретных случайных величин

3. Биномиальное распределение $\mathbf{B}_{n,p}$, $n \in \mathbb{N}$, $p \in [0, 1]$.

Случайная величина Y имеет биномиальное распределение $\mathbf{B}_{n,p}$ с параметрами $n \in \mathbb{N}$ и $p \in [0, 1]$, если

$$Y = X_1 + X_2 + \ldots + X_n,$$

где $X_1, X_2, \ldots, X_n \sim \mathbf{B}_p$ (независимые).

Другими словами, Y — это количество «успехов» в n независимых испытания Бернулли.

Вероятность получить k «успехов» в n испытаниях равна

$$\mathbb{P}(Y = k) = C_n^k p^k (1 - p)^{n-k}, \quad \text{где } C_n^k = \frac{n!}{k!(n-k)!};$$

здесь $k = 0, 1, \ldots, n$.

Примеры дискретных случайных величин

3. Биномиальное распределение $\mathbf{B}_{n,p}$.

   Примеры:

   - Количество орлов в n бросаниях монеты;
   - Количество покупок, сделанных n посетителями магазина;
   - Количество мальчиков среди n новорожденных;
   - Любое другое количество «успехов» в примерах из распределения Бернулли.

Примеры дискретных случайных величин

4. Распределение Пуассона $\mathbf{\Pi}_\lambda$, $\lambda > 0$.

   Случайная величина X имеет распределение Пуассона с параметром $\lambda > 0$, если X принимает значения $k = 0, 1, 2, \ldots$ (и только их) с вероятностями

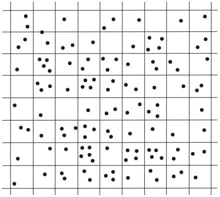
$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Примеры дискретных случайных величин

4. Распределение Пуассона $\mathbf{\Pi}_\lambda$, $\lambda > 0$.

Примеры:

- Является предельным распределением для $\mathbf{B}_{n,p}$ при $p \to 0$, $np \to \lambda$;



n — количество снарядов, N — количетво блоков, $p = 1/N$;

- Количество событий, которые произошли за фиксированный временной интервал.

Непрерывные случайные величины

Перейдем к изучению непрерывных случайных величин.

Задание распределения любых случайных величин по сути аналогично распределению массы по невесомому стержню.

В дискретном случае мы распределяли единичную массу дискретно, помещая в какие-то точки какую-то часть массы.

В непрерывном случае мы будем описывать распределение массы по стержню с помощью функции плотности.

Непрерывные случайные величины

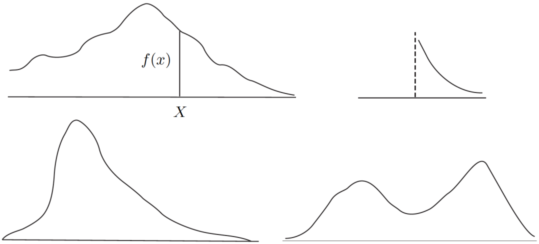
Функция $f : \mathbb{R} \to \mathbb{R}$ называется плотностью на $\mathbb{R}$, если она удовлетворяет двум условиям:

1. $f(u) \geq 0$ для любого $u \in \mathbb{R}$ (неотрицательность);
2. $\int_{-\infty}^{+\infty} f(u)du = 1$ (нормировка).

Непрерывные случайные величины

1. $f(u) \geq 0$ для любого $u \in \mathbb{R}$ (неотрицательность);
2. $\displaystyle\int_{-\infty}^{+\infty} f(u)du = 1$ (нормировка).

Какими должны быть плотности?



Непрерывные случайные величины

Как определить непрерывную случайную величину?

Чтобы задать распределение непрерывной случайной величины X, необходимо задать функцию плотности.

Далее, если случайная величина X имеет плотность распределения $f(u)$, то для (почти) любого подмножества $A \subset \mathbb{R}$ вероятность события $\{X \in A\}$ считается по формуле

$$\mathbb{P}(X \in A) = \int_A f(u)du.$$

Непрерывные случайные величины

Общие правила для формул в дальнейшем:

$$\sum_i \longrightarrow \int_{\mathbb{R}}, \qquad a_i \longrightarrow u \qquad \text{и} \qquad p_i \longrightarrow f(u)du.$$

Пример:

$$\sum_i p_i = 1 \quad \longrightarrow \quad \int_{\mathbb{R}} f(u)du = 1.$$

Непрерывные случайные величины

Частные случаи: для любых $a, b \in \mathbb{R}$, $a < b$,

- $A = (-\infty, a]$, тогда $\mathbb{P}(X \in A) = \mathbb{P}(X \le a) = \int_{-\infty}^{a} f(u)du$;
- $A = [a, +\infty)$, тогда $\mathbb{P}(X \in A) = \mathbb{P}(X \ge a) = \int_{a}^{+\infty} f(u)du$;
- $A = [a, b]$, тогда $\mathbb{P}(X \in A) = \mathbb{P}(a \le X \le b) = \int_{a}^{b} f(u)du$;
- $A = \{a\}$, тогда $\mathbb{P}(X \in A) = \mathbb{P}(X = a) = \int_{a}^{a} f(u)du = 0$.

Непрерывные случайные величины

Заметим, что по свойствам определенного интеграла случайная величина X с плотностью f всегда принимает наперед заданное значение $a \in \mathbb{R}$ с вероятностью 0, то есть

$$\mathbb{P}(X = a) = 0.$$

Непрерывные случайные величины

Пример. Пусть мы выбираем наудачу число из отрезка $[0, 1]$. Запишем теперь произвольную число из $[0, 1]$:

$$0.141592653589793\ldots.$$

- Шансы получить цифру 1 в первом разряде числа равны
$$\frac{1}{10};$$
- Шансы получить совпадающие 15 цифр равны
$$\underbrace{\frac{1}{10}\cdot\frac{1}{10}\cdot\ldots\cdot\frac{1}{10}}_{15\text{ раз}} = \left(\frac{1}{10}\right)^{15}.$$
- Чему равна вероятность, что совпадут все цифры?

Непрерывные случайные величины

Замечание. Мы говорили, что для случайной величины X с плотностью f вероятности $\{X \in A\}$ для почти всех подмножеств $A \subset \mathbb{R}$ считаются по формуле

$$\mathbb{P}(X \in A) = \int_A f(u)du.$$

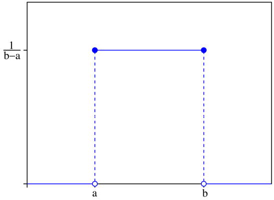
Что значит почти всех?

Предполагается, что A — не более чем счетное объединение интервалов или отрезков, возможно бесконечных и не обязательно непересекающихся. Такие множества называются борелевскими.

Примеры непрерывных случайных величин

1. Равномерное распределение на отрезке $[a, b]$, $a < b$.

Случайная величина X имеет равномерное распределение на отрезке $[a, b]$, если её плотность $f(u)$ имеет вид:



$$f(u) = \begin{cases} \frac{1}{b-a}, & u \in [a, b], \\ 0, & u \notin [a, b]. \end{cases}$$

Примеры непрерывных случайных величин

1. Равномерное распределение на отрезке $[a, b]$, $a < b$.

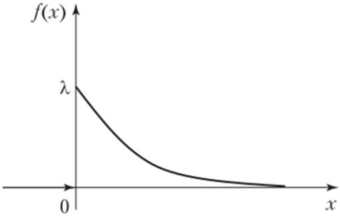
   Примеры:

   - Угол поворота стрелки в ЧГК;
   - Случайный угол выстрела;
   - Шаги человека (довольно грубое предположение);
   - Равномерное распределение на $[0, 1]$ является базовым для генерации псевдослучайных чисел в компьютере;
   - Априорное распределение на параметры модели в условиях отсутствия априорной информации.

Примеры непрерывных случайных величин

2. Экспоненциальное распределение $\mathrm{Exp}(\lambda)$, $\lambda > 0$.

Случайная величина X имеет экспоненциальное (показательное) распределение с параметром $\lambda > 0$, если её плотность имеет вид



$$f(u) = \begin{cases} \lambda e^{-\lambda u}, & u \geq 0, \\ 0, & u < 0. \end{cases}$$

Примеры непрерывных случайных величин

2. Экспоненциальное распределение $\mathrm{Exp}(\lambda)$, $\lambda > 0$.

А плотность ли это?

1) Неотрицательность:

$$f(u) \geq 0.$$

2) Нормировка:

$$\int_{\mathbb{R}} f(u)du = \int_0^{+\infty} \lambda e^{-\lambda u} du = \left(-e^{-\lambda u}\right)\Big|_0^{+\infty} = 1.$$

Примеры непрерывных случайных величин

2. Экспоненциальное распределение $\mathrm{Exp}(\lambda)$, $\lambda > 0$.

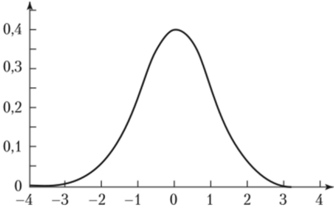
Примеры:

- Время ожидания:
  — время до приезда нужного автобуса;
  — время между покупками в магазине;
  — время между звонками в колл-центре;
- Срок эксплуатации:
  — время работы лампочки до перегорания;
  — время работы других приборов.

Примеры непрерывных случайных величин

3. Нормальное распределение $\mathcal{N}(a, \sigma^2)$, $a \in \mathbb{R}$, $\sigma > 0$.

Случайная величина X имеет стандартное нормальное распределение (или стандартное гауссовское), если её плотность имеет вид:



$$f(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}.$$

Обозначение: $X \sim \mathcal{N}(0, 1)$.

Примеры непрерывных случайных величин

3. Нормальное распределение $\mathcal{N}(a, \sigma^2)$, $a \in \mathbb{R}$, $\sigma > 0$.

А плотность ли это?

1) Неотрицательность:
$$f(u) \geq 0.$$

2) Нормировка:
$$\int_{\mathbb{R}} f(u)du = 1 \quad \text{(доказательство будет на сайте).}$$

Примеры непрерывных случайных величин

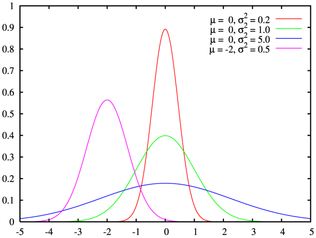
3. Нормальное распределение $\mathcal{N}(a, \sigma^2)$, $a \in \mathbb{R}$, $\sigma > 0$.

В общем случае Y имеет нормальное распределение (или гауссовское) с параметрами $a \in \mathbb{R}$ и $\sigma > 0$, если

$$Y = a + \sigma X,$$

где $X \sim \mathcal{N}(0, 1)$.

Обозначение: $Y \sim \mathcal{N}(a, \sigma^2)$.



Примеры непрерывных случайных величин

3. Нормальное распределение $\mathcal{N}(a, \sigma^2)$, $a \in \mathbb{R}$, $\sigma > 0$.

   Замечание. Явный вид плотности для $\mathcal{N}(a, \sigma^2)$ существует. Мы его выведем позже.

Примеры непрерывных случайных величин

3. Нормальное распределение $\mathcal{N}(a, \sigma^2)$, $a \in \mathbb{R}$, $\sigma > 0$.

Примеры:

- Отклонение снаряда от цели;
- Ошибки измерения;
- Будет дальше — центральная предельная теорема (почти всегда суммы одинаковых независимых случайных величин распределены нормально).

Условная вероятность

Игральная кость подбрасывается один раз. Известно, что выпало более трёх очков. Какова при этом вероятность того, что выпало нечётное число очков?

X	1	2	3	4	5	6
$\mathbb{P}$	1/6	1/6	1/6	1/6	1/6	1/6

Ответ: 1/3.

Условная вероятность

Часто в задачах на условную вероятность все удобно записывать для сразу событий, не упоминая случайные величины (но имея их в виду).

События мы будем обозначать заглавными латинскими буквами: S, H, A, B, и т.д.

Условная вероятность

Условной вероятностью события S при условии, что произошло событие H, называется число

$$\mathbb{P}(S|H) = \frac{\mathbb{P}(SH)}{\mathbb{P}(H)}.$$

В случае $\mathbb{P}(H) = 0$ будем считать, что $\mathbb{P}(S|H)$ не определена.

Условная вероятность

Формула полной вероятности

Разбиением называется конечный или счётный набор попарно непересекающихся событий $H_1, H_2, \ldots$, объединение которых равно тождественному событию (множеству всех возможных результатов эксперимента).

События $H_1, H_2, \ldots$, образующие разбиение, называют гипотезами.

Обычно для некоторого события S можно вычислить $P(S|H_i)$ и $\mathbb{P}(H_i)$. Как, используя эти данные, посчитать вероятность события S?

Условная вероятность

Формула полной вероятности

Разложим событие S в сумму попарно непересекающихся событий:

$$S = S \cap (H_1 \cup H_2 \cup \ldots) = SH_1 \cup SH_2 \cup \ldots$$

Далее

$$\mathbb{P}(S) \overset{\text{св-ва вер.}}{=} \sum_i \mathbb{P}(SH_i) \overset{\text{опр. усл. вер.}}{=} \sum_i \mathbb{P}(H_i)\mathbb{P}(S|H_i).$$

Это и есть формула полной вероятности!

Условная вероятность

Пример

Два стрелка подбрасывают (симметричную) монетку и выбирают, кто из них будет стрелять по мишени. Первый стрелок попадает по мишени с вероятностью 1, второй стрелок — с вероятностью 10^{-5}.

С какой вероятностью в результате эксперимента пуля попадет в мишень?

Условная вероятность

Решение. Введем случайные величины:

- X — номер стрелка, который стреляет по мишени;
- Y — попадание по мишени (зависит от X).

Для удобства введем еще следующие события:

- $S = \{\text{по мишени попали}\} = \{Y = 1\}$;
- $H_1 = \{\text{стрелял первый стрелок}\} = \{X = 1\}$;
- $H_2 = \{\text{стрелял второй стрелок}\} = \{X = 2\}$.

Условная вероятность

По условию задачи:

$$\mathbb{P}(H_1) = \mathbb{P}(H_2) = 1/2,$$

$$\mathbb{P}(S|H_1) = 1, \quad \mathbb{P}(S|H_2) = 10^{-5}.$$

По формуле полной вероятности имеем

$$\mathbb{P}(S) = \mathbb{P}(S|H_1)\mathbb{P}(H_1) + \mathbb{P}(S|H_2)\mathbb{P}(H_2) = \frac{1 + 10^{-5}}{2}.$$

Условная вероятность

Формула Байеса

Формула Байеса позволяет переоценить вероятности гипотез после того, как получено знание о результате эксперимента.

$$\mathbb{P}(H_k|S) \overset{\text{опр. усл. вер.}}{=} \frac{\mathbb{P}(H_k S)}{\mathbb{P}(S)}$$

$$\overset{\text{опр. усл. вер. + форм. полн. вер.}}{=} \frac{\mathbb{P}(H_k)\mathbb{P}(S|H_k)}{\sum_i \mathbb{P}(H_i)\mathbb{P}(S|H_i)}.$$

Условная вероятность

Пример

Два стрелка подбрасывают (симметричную) монетку и выбирают, кто из них будет стрелять по мишени. Первый стрелок попадает по мишени с вероятностью 1, второй стрелок — с вероятностью 10^{-5}.

По мишени попали. С какой вероятностью это сделал первый стрелок? А второй?

Условная вероятность

Решение. В обозначениях, которые у нас были:

$$\mathbb{P}(H_1|S) = \frac{\mathbb{P}(H_1 S)}{\mathbb{P}(S)} = \frac{\mathbb{P}(S|H_1)\mathbb{P}(H_1)}{\mathbb{P}(S)} = \frac{1}{1+10^{-5}}.$$

Аналогично,

$$\mathbb{P}(H_2|S) = \frac{\mathbb{P}(H_2 S)}{\mathbb{P}(S)} = \frac{\mathbb{P}(S|H_2)\mathbb{P}(H_2)}{\mathbb{P}(S)} = \frac{10^{-5}}{1+10^{-5}}.$$

П.С. Обратите внимание, что в этой задаче не нужно использовать формулу Байеса «в лоб», так как вероятность события S мы знаем из предыдущей задачи.

Условная вероятность

Задача (Парадокс Монти Холла)

Представьте, что вы стали участником игры, в которой вам нужно выбрать одну из трёх дверей. За одной из дверей находится автомобиль, за двумя другими дверями — козы. Вы выбираете одну из дверей, например, номер 1, после этого ведущий, который знает, где находится автомобиль, а где — козы, открывает одну из оставшихся дверей, например, номер 3, за которой находится коза. После этого он спрашивает вас — не желаете ли вы изменить свой выбор и выбрать дверь номер 2.

Увеличатся ли ваши шансы выиграть автомобиль, если вы примете предложение ведущего и измените свой выбор?

Спасибо за внимание!