Математика для программистов: теория вероятностей

8 июля 2018 в 11:16, Статьи (https://tproger.ru/category/articles/) 10 минут 7003

Некоторые программисты после работы в области разработки обычных коммерческих приложений задумываются о том, чтобы освоить машинное обучение и стать аналитиком данных. Часто они не понимают, почему те или иные методы работают, и большинство методов машинного обучения кажутся магией. На самом деле, машинное обучение базируется на математической статистике, а та, в свою очередь, основана на теории вероятностей. Поэтому в этой статье мы уделим внимание базовым понятиям теории вероятностей: затронем определения вероятности, распределения и разберем несколько простых примеров.

Возможно, вам известно, что теория вероятностей условно делится на 2 части. Дискретная теория вероятностей изучает явления, которые можно описать распределением с конечным (или счетным) количеством возможных вариантов поведения (бросания игральных костей, монеток). Непрерывная теория вероятностей изучает явления, распределенные на каком-то плотном множестве, например на отрезке или в круге.

Можно рассмотреть предмет теории вероятностей на простом примере. Представьте себя разработчиком шутера. Неотъемлемой частью разработки игр этого жанра является механика стрельбы. Ясно, что шутер в котором всё оружие стреляет абсолютно точно, будет малоинтересен игрокам. Поэтому, обязательно нужно добавлять оружию разброс.

Но простая рандомизация точек попадания оружия не позволит сделать его тонкую настройку, поэтому, корректировка игрового баланса будет сложна. В то же время, используя случайные величины и их распределения можно проанализировать то, как будет работать оружие с заданным разбросом, и поможет внести необходимые корректировки.

Пространство элементарных исходов

Допустим, из некоторого случайного эксперимента, который мы можем многократно повторять (например, бросание монеты), мы можем извлечь некоторую формализуемую информацию (выпал орел или решка). Эта информация называется элементарным исходом, при этом целесообразно рассматривать множество всех элементарных исходов, часто обозначаемое буквой `Ω` (Омега).

Структура этого пространства целиком зависит от природы эксперимента. Например, если рассматривать стрельбу по достаточно большой круговой мишени, — пространством элементарных исходов будет круг, для удобства размещенный с центром в нуле, а исходом — точка в этом круге.

Кроме того, рассматривают множества элементарных исходов — события (например, попадание в «десятку» — это концентрический круг маленького радиуса с мишенью). В дискретном случае всё достаточно просто: мы можем получить любое событие, включая или исключая элементарные исходы за конечное время. В непрерывном же случае всё гораздо сложнее: нам понадобится некоторое достаточно хорошее семейство множеств (https://ru.wikipedia.org/wiki/%D0%A1%D0%B5%D0%BC%D0%B5%D0%B9%D1%81%D1%82%D0% для рассмотрения, называемое алгеброй по аналогии с простыми вещественными числами, которые можно складывать, вычитать, делить и умножать. Множества в алгебре можно пересекать и объединять, при этом результат операции будет находиться в алгебре. Это очень важное свойство для математики, которая лежит за всеми этими понятиями. Минимальное семейство состоит всего из двух множеств — из пустого множества и пространства элементарных исходов.

Мера и вероятность

Вероятность — это способ делать выводы о поведении очень сложных объектов, не вникая в принцип их работы. Таким образом, вероятность определяется как функция от события (из того самого хорошего семейства множеств), которая возвращает число — некоторую характеристику того, насколько часто может происходить такое событие в реальности. Для определённости математики условились, что это число должно лежать между нулем и единицей. Кроме того, к этой функции предъявляются требования: вероятность невозможного события нулевая, вероятность всего множества исходов единичная, и вероятность объединения двух независимых событий (непересекающихся множеств) равна сумме вероятностей. Другое название вероятности — вероятностная мера. Чаще всего используется Лебегова мера (https://ru.wikipedia.org/wiki/%D0%9C%D0%B5%D1%80%D0%B0_%D0%9B%D0%B5%D0%B1%D0(обобщающая понятия длина, площадь, объём на любые размерности (`n`-мерный объем), и таким образом она применима для широкого класса множеств.

Вместе совокупность множества элементарных исходов, семейства множеств и вероятностной меры называется *вероятностным пространством*. Рассмотрим, каким образом можно построить вероятностное пространство для примера со стрельбой в мишень.

Рассмотрим стрельбу в большую круглую мишень радиуса `R`, в которую невозможно промахнуться. Множеством элементарных событий положим круг с центром в начале координат радиуса `R`. Поскольку мы собираемся использовать площадь (меру Лебега для двумерных множеств) для описания вероятности события, то будем использовать семейство измеримых (для которых эта мера существует) множеств.

Примечание На самом деле, это технический момент и в простых задачах процесс определения меры и семейства множеств не играет особой роли. Но понимать, что эти два объекта существуют, необходимо, ведь во многих книгах по теории вероятности теоремы начинаются со слов: «*Пусть* (Ω,Σ,P) *— вероятностное пространство …*».

Как уже сказано выше, вероятность всего пространства элементарных исходов должна равняться единице. Площадь (двумерная мера Лебега, которую мы обозначим $\lambda_2(A)$, где `A` — событие) круга по хорошо известной со школы формуле равна $\pi * R^2$. Тогда мы можем ввести вероятность $P(A) = \lambda_2(A) / (\pi * R^2)$, и эта величина уже будет лежать между `0` и `1` для любого события `A`.

Если предположить, что попадание в любую точку мишени равновероятно, поиск вероятности попадания стрелком в какую-то то область мишени сводится к поиску площади этого множества (отсюда можно сделать вывод, что вероятность попадания в конкретную точку нулевая, ведь площадь точки равна нулю).

Например, мы хотим узнать, какова вероятность того, что стрелок попадёт в «десятку» (событие `A` — стрелок попал в нужное множество). В нашей модели, «десятка» представляется кругом с центром в нуле и радиусом r. Тогда вероятность попадания в этот круг $P(A) = \lambda_2/(A)\pi * R^2 = \pi * r^2/(\pi R^2) = (r/R)^2$.

Это одна из самых простых разновидностей задач на «геометрическую вероятность», — большинство таких задач требуют поиска площади.

Случайные величины

Случайная величина — функция, переводящая элементарные исходы в вещественные числа. К примеру, в рассмотренной задаче мы можем ввести случайную величину $\rho(\omega)$ — расстояние от точки попадания до центра мишени. Простота нашей модели позволяет явно задать пространство элементарных исходов: $\Omega = \{\omega = (x,y) \text{ такие числа, что } x^2+y^2 \le R^2\}$. Тогда случайная величина $\rho(\omega) = \rho(x,y) = x^2+y^2$.

Средства абстракции от вероятностного пространства. Функция распределения и плотность

Хорошо, когда структура пространства хорошо известна, но на самом деле так бывает далеко не всегда. Даже если структура пространства известна, она может быть сложна. Для описания случайных величин, если их выражение неизвестно, существует понятие функции распределения, которую обозначают $F_\xi(x) = P(\xi < x)$ (нижний индекс ξ здесь означает случайную величину). Т.е. это вероятность множества всех таких элементарных исходов, для которых значение случайной величины ξ на этом событии меньше, чем заданный параметр x.

Функция распределения обладает несколькими свойствами:

1. Во-первых, она находится между `0` и `1`.
2. Во-вторых, она не убывает, когда ее аргумент x растёт.
3. В третьих, когда число `-x` очень велико, функция распределения близка к `0`, а когда само `x` большое, функция распределения близка к `1`.

Вероятно, смысл этой конструкции при первом чтении не слишком понятен. Одно из полезных свойств — функция распределения позволяет искать вероятность того, что величина принимает значение из интервала. Итак, `P (случайная величина ξ принимает значения из интервала [a;b]) = Fξ(b)-Fξ(a)`. Исходя из этого равенства, можем исследовать, как изменяется эта величина, если границы a и b интервала близки.

Пусть `d = b-a`, тогда `b = a+d`. А следовательно, `Fξ(b)-Fξ(a) = Fξ(a+d) - Fξ(a)`. При малых значениях `d`, указанная выше разность так же мала (если распределение непрерывное). Имеет смысл рассматривать отношение `pξ(a,d)= (Fξ(a+d) - Fξ(a))/d`. Если при достаточно малых значениях d это отношение мало отличается от некоторой константы `pξ(a)`, не зависящей от d, то в этой точке случайная величина имеет плотность, равную `pξ(a)`.

Примечание Читатели, которые ранее сталкивались понятием производной, могут заметить что `pξ(a)` — производная функции `Fξ(x)` в точке `a`. Во всяком случае, можно изучить понятие производной в посвященной этой теме статье (http://mathprofi.ru/opredelenie_proizvodnoi_smysl_proizvodnoi.html) на сайте Mathprofi.

Теперь смысл функции распределения можно определить так: её производная (плотность `pξ`, которую мы определили выше) в точке a описывает, насколько часто случайная величина будет попадать в небольшой интервал с центром в точке `a` (окрестность точки `a`) по сравнению с окрестностями других точек. Другими словами, чем быстрее растёт функция распределения, тем более вероятно появление такого значения при случайном эксперименте.

Вернемся к примеру. Мы можем вычислить функцию распределения для случайной величины, $\rho(\omega) = \rho(x,y) = x^2+y^2$, которая обозначает расстояние от центра до точки случайного попадания в мишень. По определению $F_\rho(t) = P(\rho(x,y) < t)$. т.е. множество $\{\rho(x,y) < t)\}$ — состоит из таких точек `(x,y)`, расстояние от которых до нуля меньше, чем `t`. Мы уже считали вероятность такого события, когда вычисляли вероятность попадания в «десятку» — она равна t^2/R^2. Таким образом, $F\rho(t) = P(\rho(x,y) < t) = t^2/R^2$, для `0<t`.

Мы можем найти плотность p_ρ этой случайной величины. Сразу заметим, что вне интервала `[0,R]` она нулевая, т.к. функция распределения на этом промежутке неизменна. На концах этого интервала плотность не определена. Внутри интервала её можно найти, используя таблицу производных (например из [PDF] (http://mathprofi.ru/tablica_proizvodnyh.pdf) на сайте Mathprofi) и элементарные правила дифференцирования. Производная от t^2/R^2 равна $2t/R^2$. Значит, плотность мы нашли на всей оси вещественных чисел.

Ещё одно полезное свойство плотности — вероятность того, что функция принимает значение из промежутка, вычисляется при помощи интеграла от плотности по этому промежутку (ознакомиться с тем, что это такое, можно в статьях о собственном (http://mathprofi.ru/opredelennye_integraly_primery_reshenij.html), несобственном (http://mathprofi.ru/nesobstvennye_integraly.html), неопределенном (http://mathprofi.ru/integraly_primery_reshenij.html) интегралах на сайте Mathprofi).

При первом чтении, интеграл по промежутку `[a; b]` от функции `f(x)` можно представлять себе как площадь криволинейной трапеции. Ее сторонами являются фрагмент оси Ох, промежуток `[a,b]` (горизонтальной оси координат), вертикальные отрезки, соединяющие точки `(a,f(a)), (b,f(b))` на кривой с точками `(a,0), (b,0)` на оси `Ox`. Последней стороной является фрагмент графика функции `f` от `(a,f(a))` до `(b,f(b))`. Можно говорить об интеграле по промежутку `(-∞; b]`, когда для достаточно больших отрицательных значений, а значение интеграла по промежутку `[a;b]` будет меняться пренебрежимо мало по сравнению с изменением числа a. Аналогичным образом определяется и интеграл по промежуткам `[a;+∞), (-∞,∞)`.

Следующее важное свойство плотности — интеграл от плотности любой случайной величины равен единице. Трактовка этого свойства такова: вероятность того, что функция принимает любое значение равна единице. Кроме того, при вычислении интегралов от плотностей случайных величин, значения которых лежат в ограниченном промежутке, нужно брать интеграл только по этому промежутку.

Итак, мы разобрались с несколькими важными понятиями: со строгим построением вероятностного пространства и построением случайных величин на нём. Кроме того, мы научились абстрагироваться от конкретного вероятностного пространства при помощи функции распределения и плотности.

Иван Камышан

Типичный программист (/author/tproger)

Математика и теория вероятностей (https://tproger.ru/tag/math-and-probability/)

Также рекомендуем:

1 comment

Your comment...

Send

Artem Nikolaev

"вероятность объединения двух независимых событий (непересекающихся множеств) равна сумме вероятностей"

тут я подзавис. это как? я подбрасываю монетку дважды, два независимых эксперимента. два независимых события - дважды выпала решка. вероятность каждого из этих событий = 0,5, т.е. вероятность выпадения хотя бы одной из двух решек равна сумме, т.е. 1? это очевидно не так. быть может имелись в виду два несовместных события? т.е. я подбрасываю кубик, два несовместных события - выпадение 1 и выпадение 6; вот тут уже действительно будет сумма.

11 Jul at 11:35 am

Andrey Sadkov

Рассматривай это как вероятности попадания в две непересекающиеся площади. В твоём примере с монетками это не так работает: нужно делить число благоприятных исходов на общее число, а отдельный исход в случае двух монеток - произведение вероятностей.

17 Jul at 2:17 am Reply

Leave a comment...

Вероятность встретить машину на пустынном шоссе

(http://tproger.ru/problems/probability-of-observing-a-car-on-a-deserted-highway/?utm_source=grf-eng&utm_medium=partner&utm_campaign=giraff.io)

Создание Minecraft на Unity3D. Часть вторая. Генерация мира

(http://tproger.ru/translations/unity-minecraft-2/?utm_source=grf-eng&utm_medium=partner&utm_campaign=giraff.io)

Блокировка Telegram: развитие событий и реакция общественности

(http://tproger.ru/news/telegram-timeline/?utm_source=grf-eng&utm_medium=partner&utm_campaign=giraff.io)

Python GUI: создаём простое приложение с PyQt и Qt Designer

(http://tproger.ru/translations/python-gui-pyqt/?utm_source=grf-eng&utm_medium=partner&utm_campaign=giraff.io)

Обзор графических библиотек C++

(http://tproger.ru/digest/cpp-best-gui/?utm_source=grf-eng&utm_medium=partner&utm_campaign=giraff.io)

Google убрала возможность использовать свои подсети для обхода блокировки

(http://tproger.ru/news/google-no-proxy/?utm_source=grf-eng&utm_medium=partner&utm_campaign=giraff.io)

Компания InVision представила конкурента Adobe Photoshop

(http://tproger.ru/news/invision-studio/?utm_source=grf-eng&utm_medium=partner&utm_campaign=giraff.io)

Как пройти собеседование в Google: советы по подготовке

(http://tproger.ru/translations/google-interview-tips/?utm_source=grf-eng&utm_medium=partner&utm_campaign=giraff.io)

Бот для Telegram на Python: от первой строчки кода до запуска на Heroku

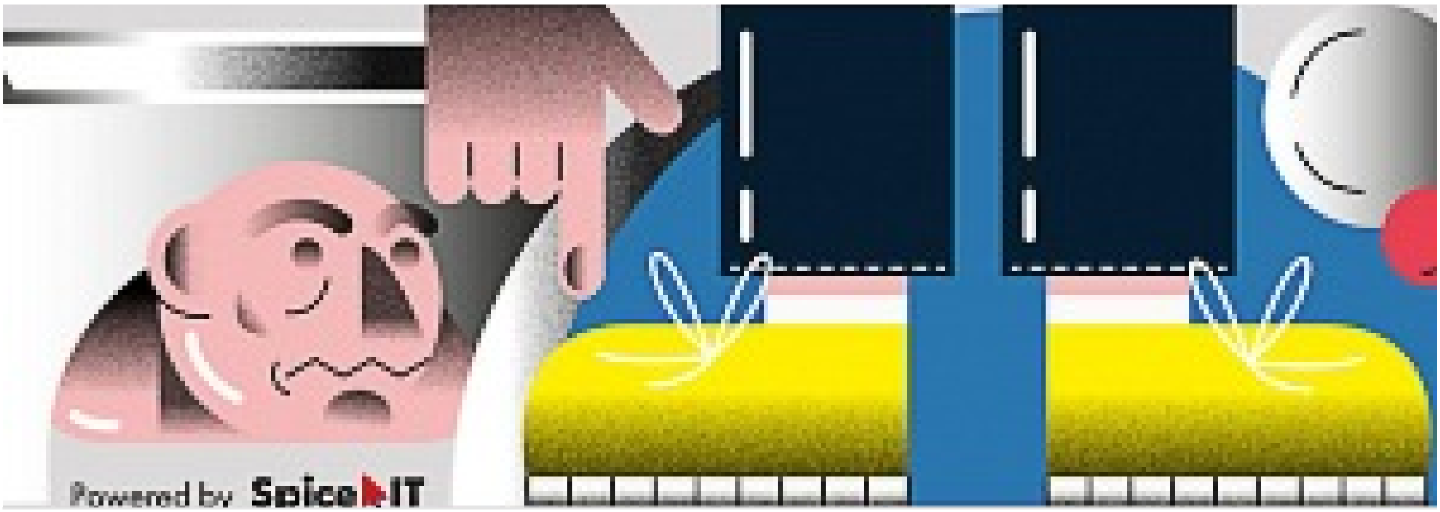
(http://tproger.ru/translations/telegram-bot-create-and-deploy/?utm_source=grf-eng&utm_medium=partner&utm_campaign=giraff.io)

Как вычислить 2 в 64 степени, не пользуясь калькулятором?

(http://tproger.ru/problems/2-64/?utm_source=grf-eng&utm_medium=partner&utm_campaign=giraff.io)

Изучаем нейронные сети: с чего начать

(http://tproger.ru/digest/learning-neuroweb-all-for-begin/?utm_source=grf-eng&utm_medium=partner&utm_campaign=giraff.io)



Плавание, башня и тролли — 3 задачки для разминки мозга

(http://tproger.ru/problems/sailing-tower-trolls/?utm_source=grf-eng&utm_medium=partner&utm_campaign=giraff.io)

(https://g