

Блуждания с Марковым

Александр Гиль и Антон Петрунин

Аннотация

Заметка основана на лекции прочитанной первым автором в школьном математическом кружке в Сиэтле, www.assagames.com/nas/.

1 Вероятность и математические ожидания

Подкидывая кубик, с равными шансами мы можем получить 1, 2, 3, 4, 5 или 6 очков. Результат такого *испытания* — число очков, выпавшее на верхней грани кубика является простым примером *случайной величины*. Естественно предположить, что результат одного испытания (подкидывания кубика) не зависит от других таких же испытаний.

Давайте записывать значения нашей случайной величины, много раз подкидывая кубик, и вычислять долю испытаний, давших конкретный *исход* (например, 5 очков) среди всех наших испытаний. При увеличении числа испытаний до бесконечности, эта доля стремится к числу, которое называется *вероятностью* этого исхода. Поскольку шансы любого из шести исходов равны, вероятность каждого исхода равна $\frac{1}{6}$.

Давайте теперь не только записывать результат каждого испытания, но и считать среднее арифметическое всех записанных на данный момент результатов. Если эта последовательность стремится к определённому числу, то такое число называется *математическим ожиданием* или *средним значением* случайной величины.

В нашем примере искомое среднее значение числа очков можно посчитать по формуле

$$\frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \frac{1}{6} \cdot 3 + \frac{1}{6} \cdot 4 + \frac{1}{6} \cdot 5 + \frac{1}{6} \cdot 6 = 3\frac{1}{2}.$$

Действительно, в пределе на долю каждого исхода 1, 2, 3, 4, 5 и 6 приходится $\frac{1}{6}$ числа всех исходов, и, значит, их среднее арифметическое должно стремиться к левой стороне равенства.

Иначе говоря, для вычисления среднего значения мы должны вычислить *взвешенную сумму* значений нашей величины для каждого исхода (1, 2, 3, 4, 5 и 6), взяв вероятность каждого исхода ($\frac{1}{6}$) в качестве веса.

Несмотря на очевидность высказанных выше утверждений, чтобы придать им точный математический смысл, требуется развить теорию, которая

Авторы признательны Арсению Акопяну за помощь при составлении этой заметки.

останется за рамками нашей заметки. Мы обсудим только способы нахождения вероятностей и средних значений в чуть более сложных ситуациях.

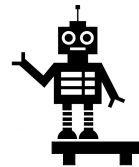
Вообще говоря, может оказаться, что среднее значение (математическое ожидание), вычисляемое как предел для продлеваемой бесконечно серии испытаний, не существует для данной схемы испытаний. Такое может случиться, поскольку у бесконечной последовательности чисел предельного значения может не существовать.

Популярным примером ситуации без предельного значения является так называемый «Санкт-Петербургский парадокс». В наших задачах такого происходить не будет, но мы не будем это строго доказывать.

2 Робот Чебуратор на коротком столе

Электромеханическая игрушка «Робот Чебуратор», будучи включенной, через равные интервалы времени делает шаги одинаковой длины налево или направо, выбирая между этими двумя направлениями с одинаковой вероятностью $\frac{1}{2}$.

Предположим, Чебуратор стоит на левом краю короткого стола: шагая влево, он чебурахнется со стола. Однако, он может шагнуть вправо и остаться на столе (на правом краю) — но он чебурахнется после второго шага направо. Нас интересуют следующие две задачи.



- Какова вероятность того, что робот чебурахнется с левого края стола, и какова — что с правого?
- Чему равно среднее значение числа шагов которые, сделает робот, чтобы чебурахнуться?

При этом мы полагаем, что существование требуемых чисел в обоих вопросах обосновано (это действительно верно, хотя требует доказательства); нам остаётся лишь найти их численные значения.

Решение первой задачи. Пусть p_1 и p_2 обозначают вероятность чебураханья из первого (левого) и второго (правого) положения.

Заметим, что

$$p_1 = \frac{1}{2} \cdot 0 + \frac{1}{2} \cdot p_2.$$

Действительно, после первого шага с вероятностью $\frac{1}{2}$ робот чебурахнется налево, и, значит, шансов чебурахнуться направо у него уже не будет, отсюда слагаемое $0 = \frac{1}{2} \cdot 0$. С той же вероятностью он перейдёт на правый край, откуда у него будет вероятность p_2 чебурахнуться вправо, отсюда слагаемое $\frac{1}{2} \cdot p_2$.

Аналогично получаем уравнение

$$p_2 = \frac{1}{2} \cdot p_1 + \frac{1}{2} \cdot 1.$$

Эти два уравнения говорят, что в последовательности из четырёх чисел

$$0, p_1, p_2, 1$$

каждое является средним арифметическим своих соседей. Иначе говоря, эта последовательность является арифметической, и, значит, $p_1 = \frac{1}{3}$ и $p_2 = \frac{2}{3}$.

Заметим, что если q_1 и q_2 вероятности чебураханья налево, то, повторив те же вычисления, получаем $q_1 = \frac{2}{3}$ и $q_2 = \frac{1}{3}$. Этот же результат можно получить, посмотрев на задачу через зеркало, меняющее местами право и лево.

В частности, из положения 1, робот чебурахаётся направо и налево с вероятностями $\frac{1}{3}$ и $\frac{2}{3}$. Заметим, что

$$p_1 + q_1 = p_2 + q_2 = 1,$$

а, значит, вероятность того, что робот не чебурахнется, равна нулю. \square

Аналогично решается и вторая задача.

Решение второй задачи. Обозначим через s_1 и s_2 средние значения числа шагов, которые сделает робот, чтобы чебурахнуться, если начинает с первой или второй позиции соответственно. Взглянув на робота через зеркало, меняющее местами право и лево, мы убеждаемся, что $s_1 = s_2$, но мы и так это увидим скоро из вычислений.

Заметим, что

$$s_1 = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot (1 + s_2).$$

Действительно, после первого шага с первой позиции с вероятностью $\frac{1}{2}$ робот чебурахнется налево, то есть он проделает всего один шаг; отсюда слагаемое $\frac{1}{2} \cdot 1$. С той же вероятностью он перейдёт на второе место, откуда в среднем он пройдёт ещё s_2 шагов. Учтя первый шаг, получаем второе слагаемое $\frac{1}{2} \cdot (1 + s_2)$.

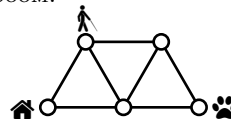
Аналогично получаем равенство

$$s_2 = \frac{1}{2} \cdot (1 + s_1) + \frac{1}{2} \cdot 1.$$

Получаем систему из двух равенств с двумя неизвестными s_1 и s_2 . Решив эту систему из двух уравнений, получаем $s_1 = s_2 = 2$. \square

Попробуйте решить следующую задачу тем же способом.

Задача. На рисунке вы видите план города с отмеченными на нём собакой, её забывчивым хозяином и их домом.



Хозяин ищет свой дом, выбирая с равной вероятности одну из дорог, и идёт до следующего перекрёстка ровно одну минуту. Дойдя до перекрёстка, он опять выбирает с равной вероятностью одну из дорог (возможно, ту же, по которой только что шёл), и идёт до следующего перекрёстка — и так далее.

Собака двигается по тому же принципу, но бежит в полтора раза быстрее.

Докажите, что из данного положения, собака в среднем прибегает домой на 20 секунд раньше хозяина.

3 Решения прямым подсчётом

Две задачи, рассмотренные выше, допускают решения прямым подсчётом вероятностей. Мы опишем их, используя те же обозначения, что и раньше.

Решения. Чтобы решить вторую задачу, Заметим, что робот чебурахается на n -ом шагу либо влево, либо вправо с вероятностью $(\frac{1}{2})^n$. А, значит, среднее значение числа шагов можно найти, просуммировав ряд

$$s_1 = s_2 = \frac{1}{2} \cdot 1 + (\frac{1}{2})^2 \cdot 2 + (\frac{1}{2})^3 \cdot 3 + \dots$$

Чтобы суммировать такой ряд нужно небольшое умение, но его сумма равна 2 — и неудивительно, что это же значение было получено выше.

Далее заметим, что, если робот начинает движение с левой позиции, то он может чебурахнуться направо только на чётных шагах, и, если он начинает с правой позиции, то только на нечётных. То есть,

$$\begin{aligned} p_1 &= (\frac{1}{2})^2 + (\frac{1}{2})^4 + (\frac{1}{2})^6 + \dots \\ p_2 &= (\frac{1}{2})^1 + (\frac{1}{2})^3 + (\frac{1}{2})^5 + \dots \end{aligned}$$

Применив формулу для суммы геометрической прогрессии, получаем те же результаты $p_1 = \frac{1}{3}$ и $p_2 = \frac{2}{3}$. \square

Как видите, этот способ оказался сложнее. Кроме того, он не допускает лёгкого обобщения на случай, когда стол имеет более чем две позиции. Как мы увидим ниже, в этом случае наше первое решение остаётся практически без изменений.

То, что среднее значение числа шагов до чебураханья равно двум, можно также увидеть поставив следующий мысленный эксперимент. Представьте себе, что при каждом чебураханье мы ставим робота обратно на стол. Тогда, вероятность того, что робот чебурахнется на шаге n равна $\frac{1}{2}$, то есть робот будет чебурахаться в среднем на половине шагов. Значит, среднее значение числа шагов между чебураханьями равно 2.

4 Робот Чебуратор на длинном столе

Представим теперь, что Чебуратор стоит на длинном столе, на котором помещается не один, а m шагов. Пусть расстояние от позиции Чебуратора до левого края стола — ровно i шагов (Чебуратор чебурахнется налево, сделав i шагов налево); тогда расстояние от позиции Чебуратора до правого края стола — ровно $m - i$ шагов. Нас интересуют всё те же задачи:

- Какова вероятность того, что робот чебурахнется с левого края стола, и какова — что с правого?
- Чему равно среднее значение числа шагов, которое сделает робот, чтобы чебурахнуться?

Решение первой задачи. Пронумеруем возможные положения Чебуратора от числами от 1 до $m - 1$, по числу шагов налево, которое Чебуратору нужно сделать, чтобы чебурахнуться.

Нам будет удобно добавить ещё два положения под номерами 0 и m : первое соответствует тому, что робот чебурахнулся налево, а второе соответствует тому, что робот чебурахнулся направо (из этих позиций выхода нет).

Обозначим через p_i вероятность того, что Чебуратор чебурахнется направо, если начинает с позиции под номером i . Естественно, мы имеем

$$p_0 = 0 \text{ и } p_m = 1.$$

Применим тот же метод, что и раньше. Стоя на позиции номер i , с вероятностью $\frac{1}{2}$, Чебуратор перейдёт на позицию номер $i + 1$. В этом случае вероятность того, что он чебурахнется вправо будет p_{i+1} . С той же вероятностью $\frac{1}{2}$ он перейдёт на позицию номер $i - 1$. В последнем случае вероятность того, что он чебурахнется вправо будет p_{i-1} . То есть,

$$p_i = \frac{1}{2} \cdot p_{i-1} + \frac{1}{2} \cdot p_{i+1}.$$

Иначе говоря в последовательности из $m + 1$ числа

$$0 = p_0, p_1, p_2, \dots, p_m = 1$$

каждое число является средним арифметическим соседей. Отсюда получаем $p_i = \frac{i}{m}$. \square

Решение второй задачи. Давайте использовать ту же нумерацию позиций, как и в решении первой задачи.

Обозначим через s_i среднее значение числа шагов робота, если он начинает с позиции под номером i . Естественно предположить, что $s_0 = s_m = 0$ — ведь попадание на позиции 0 и m означает, что робот уже чебурахнулся.

Попробуем как и раньше посчитать значение s_i новым способом. После одного шага с i -ой позиции Чебуратор окажется на позиции $i - 1$ или $i + 1$ с равными вероятностями $\frac{1}{2}$. После этого ему останется в среднем пройти s_{i-1} и s_{i+1} шагов соответственно. Не забыв учесть уже пройденный шаг, получаем

$$s_i = \frac{1}{2} \cdot (s_{i-1} + 1) + \frac{1}{2} \cdot (s_{i+1} + 1),$$

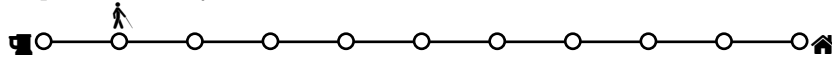
или

$$s_{i+1} = 2 \cdot s_i - s_{i-1} - 2.$$

Применяя эти равенства рекурсивно, получаем $s_i = i \cdot (s_1 - i)$ для любого i . Поскольку $s_m = 0$, получаем $s_1 = m$ и, значит, $s_i = i \cdot (m - i)$. \square

Для закрепления материала мы советуем решить следующую задачу.

Задача. Пьяница вышел из бара, расположенного в 10 кварталах от своего дома на той же улице и идёт домой. Чтобы сделать дорогу веселее, он разнообразит её следующим образом. Дойдя до перекрёстка, он подкидывает монетку — и, если она выпадает орлом, он продолжает путь в том же направлении; если же она выпадает решкой, он разворачивается и идёт в противоположном направлении. Если ему случается вернуться к бару, он всегда разворачивается в сторону дома; если же он дошёл до своего дома, то завершает свой путь.



Докажите, что в среднем, за такую прогулку, пьяница проходит 100 кварталов и среднее число его возвращений к бару равно 10.

5 Бактерии в пробирке

Задача. В пробирке живут 10 бактерий 3 зелёных и 7 жёлтых. Каждую секунду происходит следующее: одна из 10 бактерий (равновероятно выбранная из всех десяти) погибает, и тут же одна из оставшихся девяти (равновероятно выбранная) делится на две своих точных копии. Таким образом, в конце секунды в пробирке остаётся ровно 10 бактерий.

Какова вероятность того, что через некоторое время все станут зелёными?

Решение. Мы сведём задачу к уже решённой.

Предположим, что через несколько секунд число зелёных бактерий стало i . Тогда спустя секунду их число может стать $i - 1$, i или $i + 1$. При этом вероятности первого и последнего исхода равны. Действительно, исход $i - 1$ означает, что первая бактерия оказалась зелёной, а вторая — жёлтой, а в исходе $i + 1$, наоборот, первая — жёлтой а вторая — зелёной. Про исход i можно думать как про пропускание хода.

Таким образом, наша задача становится похожей на задачу про робота Чебуратора — только теперь Чебуратор ходит направо и налево с равными положительными вероятностями, и с оставшейся вероятностью пропускает ход. Однако заметим, что пропускание хода можно вовсе не учитывать. То есть, ответ в задаче можно получить, подставив $i = 3$ и $m = 10$ в задаче про Чебуратора на длинном столе. Таким образом, с вероятностью $\frac{3}{10}$ все бактерии станут зелёными, и с вероятностью $\frac{7}{10}$ все станут жёлтыми.

Эту же задачу (а значит, и первую задачу про Чебуратора) можно решить без вычислений.

Сначала докажем, что с вероятностью 1 через некоторое время все бактерии в пробирке будут потомками одной.

Для начала заметим, что с положительной вероятностью это может случиться за первые 10 секунд. Вероятность этого события мала, но вероятность того, что за достаточно длинный промежуток времени найдутся такие 10 секунд, произвольно близка к 1.

При этом, в самом начале, каждая из десяти бактерий имеет равные шансы на то, чтобы стать прародителем всех выживших. Поэтому в трёх из десяти случаев все станут зелёными, и в семи из десяти все станут жёлтыми.

6 Броуновское движение

Рассмотрим случай, когда Чебуратор стоит на середине стола длиной $2 \cdot n$ — чтобы чебурахнуться, ему необходимо сделать или n шагов вправо, или столько же влево. Как мы выяснили, среднее значение числа шагов, которое он делает для того, чтобы чебурахнуться, равно n^2 . Это наблюдение можно использовать в обратном направлении: повторив это испытание много раз и оценив среднее значение числа шагов до чебураханья, мы, зная размер стола, сможем оценить длину шага Чебуратора. Последнее может оказаться полезным, если шаги Чебуратора настолько мелки, что не поддаются прямому измерению.

Поведение Чебуратора на столе мало чем отличается от хаотического движения малых твёрдых частиц, плавающих в жидкости, под действием ударов молекул жидкости. Это хаотическое движение открыл Роберт Браун в начале XIX-го века, а в начале XX-го века удалось оценить среднее значение числа молекул в единице объёма по параметрам броуновского движения. Этот эксперимент послужил убедительным аргументом в утверждении атомарной теории, при этом оказалось достаточным наблюдений с помощью обычного микроскопа.

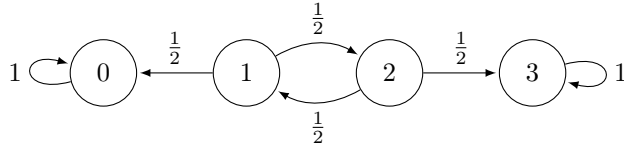
Конечно, эта физическая задача сложнее наших задач о Чебураторе. Тем не менее, идея решения у этих задач одна и та же.

7 Да, а о чём мы говорили?

Мы рассмотрели очень частный случай *цепей Маркова*, названных так в честь Андрея Андреевича Маркова-старшего (его сына тоже звали Андрей Андреевичем и он тоже занимался математикой).

А именно, мы рассматривали *конечные цепи Маркова с поглощающими состояниями*, каждое *поглощающее состояние* соответствует чебураханью нашего Чебуратора.

Цепь Маркова можно задать ориентированным графом на каждом ребре в которого написана вероятность перехода из одного состояния в другое. Например задача о Чебураторе на коротком столе может быть описана таким графом.



Состояниям 0, 1, 2 и 3 соответствуют «чебураханье налево», «левый край стола», «правый край стола» и «чебураханье направо». Заметим, что сумма вероятностей на стрелках исходящих из одного состояния всегда равна единице.

Цепи Маркова оказались очень востребованной моделью. Например ранжирование страниц используемое поисковыми системами основано на цепях Маркова. С помощью этих цепей создаются алгоритмы для автоматического написания музыки. Важные приложения известны в химии, биологии, лингвистике, экономике.

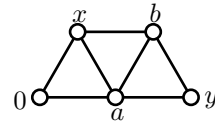
8 Решение

Вот решение на всяк пожарный.

Обозначим через x , y , a и b — среднее число пройденных кварталов с соответственных перекрёстков. Тогда

$$x = 1 + \frac{1}{3} \cdot (a + b)$$

$$y = 1 + \frac{1}{2} \cdot (a + b)$$



Значит время хозяин приходит в среднем за $1 + \frac{1}{3} \cdot (a + b)$ минут, а собака $\frac{2}{3} \cdot (1 + \frac{1}{2} \cdot (a + b)) = \frac{2}{3} + \frac{1}{3} \cdot (a + b)$. То есть, собака прибегает на 20 секунд раньше хозяина.