Схемы Бернулли с конечным алфавитом

! Схемы Бернулли

Здесь нас интересуют последовательности *очень длинных* слов $(x_0, x_1, x_2, ...)$, где каждый x_i берётся из конечного алфавита $\mathcal{A} := \{a, b, ...\}$. Мы будем рассматривать бесконечно длинные слова, что является стандартным подходом, если длина слова значительно превышает число итераций отображения сдвига (определённого ниже). Эти структуры являются базовым строительным блоком для статистической механики, динамических систем и конечных автоматов. Здесь мы рассмотрим простейший пример — случай, когда каждая буква выбирается независимо.
Эта страница содержит интерактивный контент, который вы можете свободно изучать и изменять. Для более сложных изменений рекомендуется запустить и отредактировать интерактивный Jupyter Notebook. Вы можете:

- Запустить его на компьютере с установленным Python/Jupyter-lab (требуются `ipywidgets`, `numpy`, `matplotlib`).
- Запустить его с помощью онлайн-сервиса. Официальный сайт Jupyter предоставляет бесплатный и открытый сервис. Если вам нужна большая вычислительная мощность, вы можете импортировать блокнот в Google Colab (требуется учётная запись Google).

Определения

Пусть $\mathcal{A}$ — конечное или счётное множество, и пусть $M : \mathcal{A} \times \mathcal{A} \to \{0, 1\}$. $\mathcal{A}$ представляет собой алфавит, а M кодирует, какие пары букв могут быть последовательными (более сложные правила для разрешённых последователь-

ностей букв вписываются в эту структуру простым изменением пространства алфавита).

Пространство $\Sigma \equiv \Sigma(\mathcal{A}, M)$ M-совместимых слов — это пространство последовательностей

$$\Sigma := \left\{ x \in \mathcal{A}^{\mathbb{N}} \,:\, M_{x_i, x_{i+1}} = 1,\, \forall i \in \mathbb{N} \right\}$$

Другими словами, M указывает, какие буквы могут идти друг за другом, а Σ — это пространство бесконечно длинных слов, которые содержат только разрешённые последовательные пары букв. Σ является измеримым пространством, будучи измеримым подмножеством $\mathcal{A}^{\mathbb{N}}$ (снабжённого σ-алгеброй произведения). Элементы Σ обычно обозначаются x, y, z, и мы пишем $x = (x_0, x_1 \ldots)$ и так далее.

Определение 0.1 (Отображение сдвига)**.** Отображение

$$\begin{aligned} T &\colon \Sigma \to \Sigma \\ (Tx)_i &= x_{i+1} \end{aligned}$$

называется **отображением сдвига**.

Обычно пара $(\mathcal{A}, M)$ (и, следовательно, получающееся пространство Σ, оснащённое отображением сдвига T) называется **символической динамической системой**.

Заметьте, что T не является обратимым (если только $\mathcal{A}$ не содержит всего один элемент), так как первая буква x исчезает в Tx, вторая буква становится первой и так далее.

Определение 0.2 (Инвариантная мера)**.** Вероятностная мера $\mathbb{P}$ на Σ называется **инвариантной**, если

$$\mathbb{P}(E) = \mathbb{P}(T^{-1}(E)) \tag{1}$$

Пусть $\mu \in \mathcal{P}(\mathcal{A})$ — вероятностная мера на $\mathcal{A}$. Если $M_{a,b} = 1$ для всех $a, b \in \mathcal{A}$, то мера-произведение $\mathbb{P} = \mu^{\otimes \mathbb{N}}$ инвариантна. Это утверждение очевидно, это просто сложный способ сказать, что

$$\mathbb{P}(X_0 \in A_0, \ldots, X_{n-1} \in A_{n-1}) = \mathbb{P}(X_1 \in A_0, \ldots, X_n \in A_{n-1})$$

если X_i являются последовательностью независимых одинаково распределённых (i.i.d.) случайных величин. Действительно, по определению σ-алгебры произведения, достаточно проверить Уравнение 1 для цилиндрических множеств $E = A_0 \times \cdots A_{n-1} \times \mathcal{A} \times \ldots \mathcal{A}$.

Определение 0.3 (Схема Бернулли)**. Схема Бернулли** — это символическая динамическая система, где $\mathcal{A}$ конечно или счётно, $M_{a,b} = 1$ для всех $a, b \in \mathcal{A}$, и $\mathbb{P} = \mu^{\otimes \mathbb{N}}$ — вероятностная мера-произведение на $\Sigma = \mathcal{A}^{\mathbb{N}}$ для некоторой вероятности μ на $\mathcal{A}$. Другими словами, это просто последовательность i.i.d. случайных величин на конечном или счётном пространстве $\mathcal{A}$, каждая с распределением μ. Если $\mathcal{A}$ конечно, то схема называется **конечной**.

Пример 0.1. Рассмотрим схему Бернулли над алфавитом $\mathcal{A} = \{0, 1\}$. Она определяется параметром $p \in [0, 1]$, заданным как $p = \mu(\{1\})$. Чтобы избежать тривиальных ситуаций, мы выбираем $p \in (0, 1)$. С точностью до счётного множества (и, следовательно, множества $\mathbb{P}$-меры 0), $\Sigma = \{0, 1\}^{\mathbb{N}}$ находится в биекции с интервалом $[0, 1]$ через отображение

$$\Phi(x) := \sum_{i \geq 0} 2^{-i-1} x_i$$

Точка $\Phi(x)$ будет находиться в интервале $[0, 1/2)$, если $x_0 = 0$, и в интервале $[1/2, 1]$, если $x_0 = 1$. Аналогично, $\Phi(x) \in [0, 1/4)$, если $x_0 = 0$ и $x_1 = 0$, $\Phi(x) \in [1/4, 1/2)$, если $x_0 = 0$ и $x_1 = 1$, и так далее. Это не совсем точно, так как двоично-рациональные числа допускают два двоичных представления, например, $\Phi(01111111\ldots) = 1/2 = \Phi(10000000\ldots)$. Но это всего лишь счётное множество, которым мы пренебрежём, так как оно имеет меру 0 (можно сделать Φ биективным и измеримым, изменив его определение на счётном множестве, но здесь нам это не понадобится).

Таким образом, мы можем вместо этого рассматривать эту схему Бернулли на $[0, 1]$. Чему соответствует отображение сдвига $x \mapsto Tx$? Другими словами, можем ли мы определить отображение $T' \colon [0, 1] \to [0, 1]$ такое, что $T'(\Phi(x)) = \Phi(Tx)$? Это просто способ сказать: я сначала меняю пространство для представления моих точек, с $\Sigma \ni x$ на $[0, 1] \ni \Phi(x)$. Теперь x меняется на Tx, могу ли я прочитать это изменение *автономно* на $[0, 1]$, просто как функцию от образа $\Phi(x)$? Конечно, можем, так как Φ обратимо почти всюду. Итак, $T'y = \Phi(T\Phi^{-1}y)$. Мы утверждаем, что $T'y = 2y(\mathrm{mod} 1)$; что легко проверить. Действительно,

$$\Phi(Tx) = \sum_{i \geq 0} 2^{-i-1} x_{i+1} = \begin{cases} 2\Phi(x) & \text{if } x_0 = 0 \\ 2\Phi(x) - 1 & \text{if } x_0 = 1 \end{cases}$$

../media/bs_doubling.webm

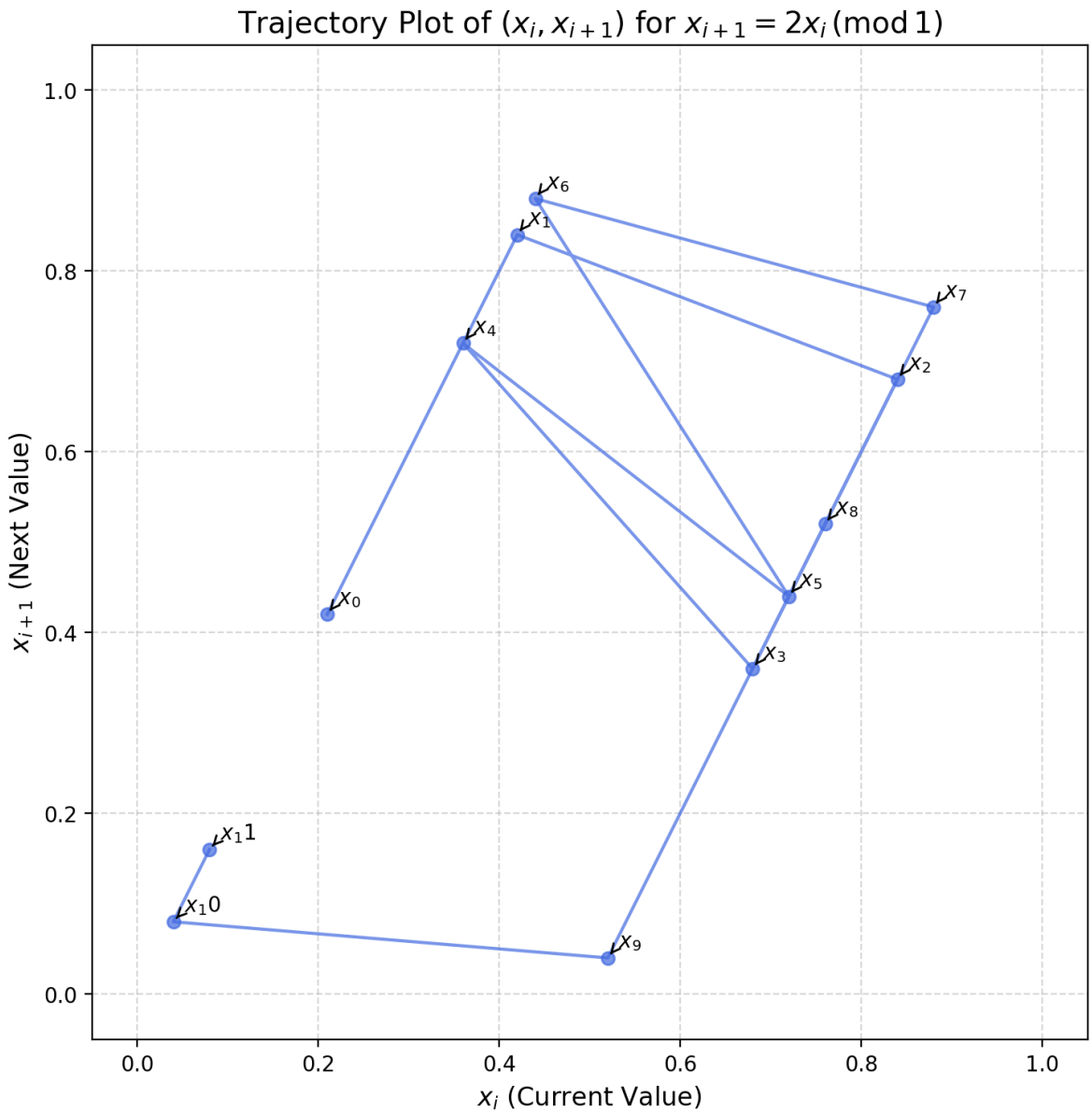

Рисунок 1: График пар (x_i, x_{i+1}), где $x_i = T' x_{i-1}$. По мере итерации отображения точки заполняют график отображения T.

Конечные схемы

Сосредоточимся на конечных схемах Бернулли. Пусть $\mathcal{A} = \{0, 1, \dots, n-1\}$, $\Sigma = \mathcal{A}^{\mathbb{N}}$ и $p_k := \mu(\{k\})$, так что $\sum_i p_i = 1$. Мы предполагаем, что $p_i > 0$ (в противном случае просто игнорируем буквы алфавита с вероятностью 0). Как и в Пример 0.1, рассмотрим n-ичное представление

$$\Phi \colon \Sigma \to [0,1]$$
$$\Phi(x) := \sum_{i \geq 0} n^{-i-1} x_i$$

Упражнение 0.1. Докажите, что в этом случае n-ичное представление переносит отображение сдвига на Σ в отображение $[0,1] \ni y \mapsto ny (\mathrm{mod} 1)$ на $[0,1]$.

Мы хотим ответить на следующий вопрос. Зафиксируем μ, то есть зафиксируем $p_0, \dots, p_{n-1}$. Каков образ меры (pushforward) $\mathbb{P} = \mu^{\otimes \mathbb{N}}$ при отображении на $[0,1]$? Другими словами, если мы возьмём $(X_i)_{i \geq 0}$ i.i.d. с распределением μ (так что $\mathbb{P}(X_i = k) = p_k$), каким будет распределение случайной величины

$$Y = \sum_{i \geq 0} n^{-i-1} X_i$$

Сумма в правой части хорошо определена, так что Y хорошо определена, и мы хотим исследовать её распределение. Обозначим $\mathbb{Q}_\mu = \mu^{\otimes \mathbb{N}} \circ \Phi^{-1}$ закон распределения Y, когда X_i являются i.i.d. с распределением μ.

Утверждение 0.1. *Если $p_k \equiv p = 1/n$, то есть если X_i равномерно распределены на конечном множестве $\mathcal{A}$, то $\mathbb{Q}_\mu$ является мерой Лебега (но определённой на борелевской σ-алгебре) на $[0,1]$.*

Утверждение 0.1. Действительно, зафиксируем интервал вида $I = [kn^{-j}, (k+1)n^{-j}]$, для некоторого $j \geq 1$ и $0 \leq k \leq n^j - 1$. Мы будем иметь, что $Y = \Phi(X) \in I$ тогда и только тогда, когда $\sum_{i=0}^{j-1} n^{-i-1} X_i = kn^{-j}$ (с точностью до счётного множества меры 0). В свою очередь, это верно для единственного значения $(X_0, \dots, X_{j-1})$. Следовательно,

$$\mathbb{Q}_\mu(I) = \mathbb{P}(Y \in I) = \mathbb{P}\left(\sum_{i=0}^{j-1} n^{-i-1} X_i = kn^{-j}\right) = p^j = n^{-j} = \mathrm{Leb}(I)$$

Поскольку интервалы такого типа полностью определяют меру, мы имеем $\mathbb{Q}_\mu = \mathrm{Leb}$. $\square$

Другими словами, мы можем получить меру Лебега, просто подбрасывая монету счётное число раз — довольно интуитивная конструкция для объекта, который был плохо определён до начала 20-го века. Согласно закону больших чисел, Утверждение 0.1 означает, что для почти всех по Лебегу вещественных чисел частота каждой цифры одинакова. Скажем, если мы зафиксируем $n = 10$, тогда определим $\eta_{k,\ell}(x)$ как количество раз, которое цифра $k = 0, 1, 2, 3, 4, 5, 6, 7, 8, 9$ встречается в представлении x в основании 10, в первых ℓ цифрах x. Тогда

$$\mathrm{Leb}\left(\left\{x \in [0,1] \,:\, \limsup_{\ell} \frac{\eta_{k,\ell}}{\ell} = \liminf_{\ell} \frac{\eta_{k,\ell}}{\ell} = \frac{1}{10}\right\}\right) = 1$$

Действительно, как только мы вернёмся к $n = 10$-ичному представлению, $\eta_{k,\ell}$ есть не что иное, как сумма $\sum_{i=0}^{\ell-1} \mathbf{1}_{X_i=k}$, которая является суммой i.i.d. случайных величин с математическим ожиданием $1/n$. Таким образом, усиленный закон больших чисел подразумевает, что предел существует и равен $1/10$ с вероятностью 1.

Теорема 0.1. *Для вероятности $\mu \in \mathcal{P}(\mathcal{A})$, $\mu(\{k\}) := p_k$ на конечном алфавите $\mathcal{A} = \{0, \dots, n-1\}$, определим*

$$E_\mu := \left\{x \in [0,1] \,:\, \limsup_{\ell} \frac{\eta_{k,\ell}}{\ell} = \liminf_{\ell} \frac{\eta_{k,\ell}}{\ell} = p_k, \quad \forall k \in \mathcal{A}\right\}$$

Тогда E_μ являются измеримыми, непересекающимися множествами и $\mathbb{Q}_\mu(E_\mu) = 1$ (и, следовательно, $\mathbb{Q}_\mu(E_{\mu'}) = 0$ для $\mu \neq \mu'$). Таким образом,

A. Если μ равномерна, $p_k = 1/n$, то $\mathbb{Q}_\mu$ является мерой Лебега. B. Если существует $k \in \mathcal{A}$ такое, что $p_k = 1$, то $\mathbb{Q}_\mu$ является мерой Дирака (сосредоточена в одной точке). C. Если $p_k < 1$ для всех k, но мера не является равномерной, то $\mathbb{Q}_\mu$ имеет канторовский тип, а именно, она присваивает меру 0 каждой точке, но сингулярна относительно меры Лебега.

Доказательство оставлено в качестве упражнения с указаниями.

Упражнение 0.2.

a. Заметьте, что по самому своему определению E_μ являются борелевски-измеримыми и непересекающимися.
b. Рассуждая как в тексте после Утверждение 0.1, докажите в деталях, что $\mathbb{Q}_\mu(E_\mu) = 1$.
c. Вспомните Утверждение 0.1 и решите тривиальный случай, когда $p_k = 1$ для некоторого k, чтобы доказать утверждения A) и B) в теореме.
d. В случае C) остаётся доказать, что $\mathbb{Q}_\mu$ присваивает меру 0 каждой точке.

Проверка сходимости

Мы можем посмотреть на траекторию точек карты $x \mapsto nx(mod1)$.

Мы можем проверить, что если начать с N точек, выбранных с заданным распределением μ, то можно численно увидеть, что через несколько шагов они рассредоточатся по интервалу. Идея заключается в том, что если начать с распределения μ, абсолютно непрерывного относительно инвариантной меры, то после множества итераций отображения сдвига T распределение точек приблизится к инвариантной мере (мере Лебега).

Однако мера Лебега — не единственная инвариантная мера, и по мере многократного итерирования отображения мы сталкиваемся с проблемами. Численная точность (тот факт, что мы не можем точно выбрать а.с. относительно Лебега) будет иметь значение при многократном итерировании отображения, например, периодические орбиты (с вероятностью 0 относительно Лебега) имеют положительную вероятность при компьютерной выборке. Эффект виден в этом видео.

../media/bs_simulation.webm

Не стесняйтесь экспериментировать, как различные распределения сходятся к распределению Лебега. Можете ли вы представить себе, как математически выразить результат сходимости?

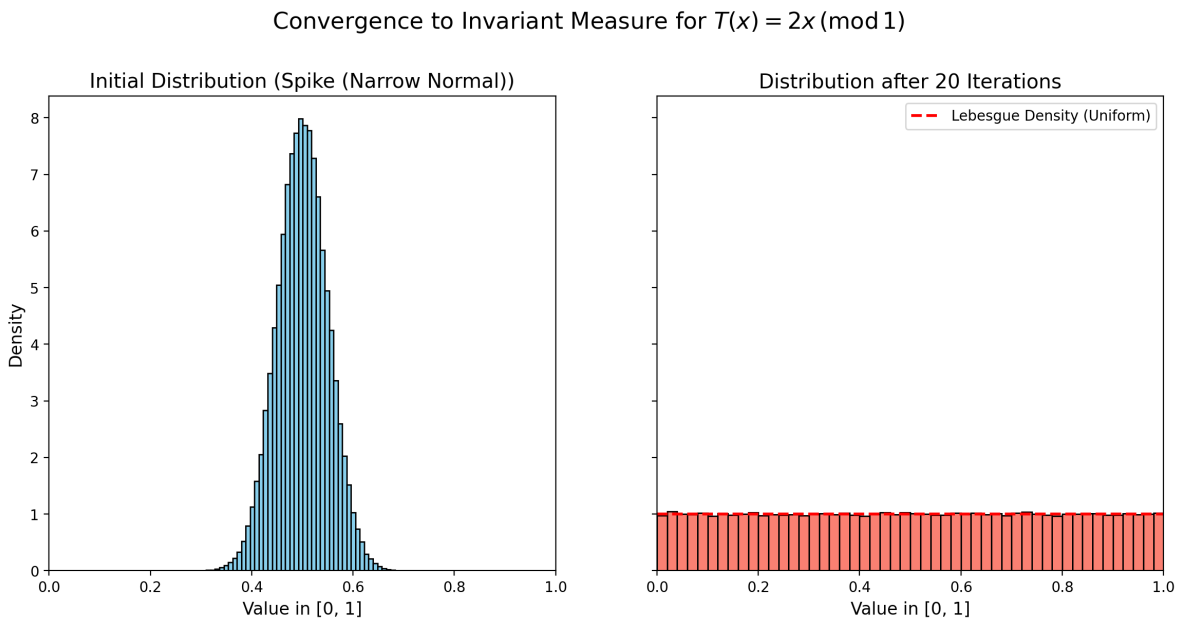

Рисунок 2: Если в момент времени 0 мы выбираем наши точки случайным образом, с мерой, абсолютно непрерывной относительно меры Лебега, их распределение будет экспоненциально быстро сходиться к мере Лебега. Это также верно и для других инвариантных мер.