# Вероятность и вероятностные алгоритмы. Семинар 2.

*Подготовил: Горбунов Э.*

Ключевые слова: вероятностное пространство, случайная величина, формула Байеса, математическое ожидание, дисперсия, неравенство Маркова, неравенство Чебышёва, неравенство Чернова, время работы "в среднем" и в "в худшем" случае, алгоритмы типа "Лас-Вегас" и "Монте-Карло"

## Мотивировка

На первой неделе мы разобрали некоторые примеры алгоритмов и научились оценивать их время работы. Отметим, что все алгоритмы, которые мы разбирали, являются *детерминированными*, то есть шаги, которые они делают, а так же результат, который они выдают, *зависят только от входа* алгоритма. Однако в ряде случаев можно использовать *вероятностные* процедуры, вместо детерминированных. Делается это прежде всего для того, чтобы уменьшить время работы алгоритма. Конечно, за это приходится платить тем, что решение задачи получается неточным или правильный ответ даётся с некоторой вероятностью. Часто бывает, что приходится делать больше итераций, но стоимость одной итерации становится гораздо "дешевле". Стоит отметить, что вероятностные алгоритмы важны не только с теоретической точки зрения, но и с практической, т. к. некоторые вероятностные процедуры очень хорошо себя зарекомендовали на практике[1], так что полезно буквально с первых шагов привыкнуть к языку теории вероятностей, а при разработке и анализе алгоритмов к случайности следует относиться как к потенциально важному ресурсу. Но осознать, в чем, собственно, этот ресурс заключается, можно только решая задачи.

## Базовые понятия теории вероятностей

Большинство приведенных ниже определений справедливы только для случая конечного множества элементарных исходов, хотя их можно обобщить на случай бесконечных вероятностных пространств, но это требует более серьезной техники. В любом случае, следует быть особенно осторожным с определением случайных величин.

1. **Вероятностным пространством** называется[2] множество $\Omega$, элементы которого называются возможными или элементарными исходами $\omega \in \Omega$. На вероятностном пространстве задана функция $\mathbb{P}\{\cdot\} : \Omega \to [0, 1]$, называемая **вероятностным распределением**, для которой выполнено равенство $\sum_{\omega\in\Omega} \mathbb{P}\{\omega\} = 1$. Число $\mathbb{P}\{\omega\}$ понимается как вероятность исхода $\omega$. **Событием** называется произвольное подмножество $\Omega$. Соответственно, вероятность события $A \subseteq \Omega$ определяется формулой $\mathbb{P}\{A\} \stackrel{\text{def}}{=} \sum\limits_{\omega\in A} \mathbb{P}\{\omega\}$.

   **Пример.** Рассмотрим вероятностное пространство $\Omega = \{1, 2, 3, 4, 5, 6\}$, где $\omega \in \Omega$ — исход, который означает, что при бросании игрального кубика выпало число $\omega$. Считаем, что кубик симметричный и выпадение каждой его стороны равновероятно. Тогда $P\{\omega\} = \frac{1}{6}$ для любого $\omega \in \Omega$.

2. Если $\mathbb{P}\{A\} = 0$, то событие $A$ называют **невозможным**. Если $\mathbb{P}\{A\} = 1$, то событие $A$ называют **достоверным**. Если $\mathbb{P}\{A \cap B\} = 0$, то события $A$ и $B$ называются **несовместными**.

3. События $A$ и $B$ называют **независимыми**, если $\mathbb{P}\{A \cap B\} = \mathbb{P}\{A\} \cdot \mathbb{P}\{B\}$.

4. События $A_1, \ldots, A_n$ называются **попарно независимыми**, если любые два из них являются независимыми.

[1] Например, в некоторые люди, которые занимаются глубоким обучением (Deep Learning), под словами "градиентный спуск" (что бы это не значило) часто понимают его стохастическую версию, ввиду её высокой практической ценности.

[2] Везде далее множество считается конечным.

5. События $A_1, \dots, A_n$ называются **независимыми в совокупности**, если для любого набора из этих событий $A_{i_1}, \dots, A_{i_k}$ выполняется $\mathbb{P}\{A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}\} = \mathbb{P}\{A_{i_1}\}\mathbb{P}\{A_{i_2}\} \cdot \dots \cdot \mathbb{P}\{A_{i_k}\}$.

   **Пример.** Из определений следует, что из независимости в совокупности следует попарная независимость. Верно ли, что верно и обратное, то есть что из попарной независимости следует независимость в совокупности? Оказывается, что нет. Продемонстриурем это на простом примере. Пусть $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ и $\mathbb{P}\{\omega_i\} = \frac{1}{4}$ для всех $i$. Рассмотрим события $A = \{\omega_1, \omega_2\}, B = \{\omega_2, \omega_3\}, C = \{\omega_1, \omega_3\}$. Тогда $\mathbb{P}\{A\} = \mathbb{P}\{B\} = \mathbb{P}\{C\} = \frac{1}{2}$ и

$$\mathbb{P}\{A \cap B\} = \mathbb{P}\{\omega_2\} = \frac{1}{4} = \mathbb{P}\{A\}\mathbb{P}\{B\},$$
$$\mathbb{P}\{B \cap C\} = \mathbb{P}\{\omega_3\} = \frac{1}{4} = \mathbb{P}\{B\}\mathbb{P}\{C\},$$
$$\mathbb{P}\{A \cap C\} = \mathbb{P}\{\omega_1\} = \frac{1}{4} = \mathbb{P}\{A\}\mathbb{P}\{C\},$$

   но $\mathbb{P}\{A \cap B \cap C\} = \mathbb{P}\{\varnothing\} = 0 \neq \mathbb{P}\{A\}\mathbb{P}\{B\}\mathbb{P}\{C\} = \frac{1}{8}$.

6. **Условная вероятность** $\mathbb{P}\{A \mid B\}$ события $A$ при условии $B$ определяется из формулы $\mathbb{P}\{A \cap B\} = \mathbb{P}\{A \mid B\} \cdot \mathbb{P}\{B\}$. В частности, если $\mathbb{P}\{B\} > 0$, то справедливо равенство

$$\mathbb{P}\{A \mid B\} = \frac{\mathbb{P}\{A \cap B\}}{\mathbb{P}\{B\}}.$$

   **Пример (Парадокс Монти-Холла).** Представьте, что вы стали участником игры, в которой находитесь перед тремя дверями. Ведущий поместил за одной из трех пронумерованных дверей автомобиль, а за двумя другими дверями — по козе (козы тоже пронумерованы) случайным образом — это значит, что все $3! = 6$ вариантов расположения автомобиля и коз за пронумерованными дверями равновероятны. У вас нет никакой информации о том, что за какой дверью находится. Ведущий говорит: "Сначала вы должны выбрать одну из дверей. После этого я открою одну из оставшихся дверей (при этом если вы выберете дверь, за которой находится автомобиль, то я с вероятностью $1/2$ выберу дверь, за которой находится коза номер 1, и с вероятностью $1 - 1/2 = 1/2$ дверь, за которой находится коза номер 2). Затем я предложу вам изменить свой первоначальный выбор и выбрать оставшуюся закрытую дверь вместо той, которую вы выбрали сначала. Вы можете последовать моему совету и выбрать другую дверь либо подтвердить свой первоначальный выбор. После этого я открою дверь, которую вы выбрали, и вы выиграете то, что находится за этой дверью." Вы выбираете дверь номер 3. Ведущий открывает дверь номер 1 и показывает, что за ней находится коза. Затем ведущий предлагает вам выбрать дверь номер 2. Увеличатся ли ваши шансы выиграть автомобиль, если вы последуете его совету?

   **Решение.** Казалось бы, какая разница, какую дверь выбирать в таком случае: дверей 2 и за одной из них автомобиль, значит, автомобиль за дверью 3 с вероятностью $\frac{1}{2}$. Оказывается, что выгоднее последовать совету ведущего, что для многих людей контринтуитивно. Изначально для нас автомобиль находится за каждой дверью равновероятно. Но когда ведущий показывает, что за одной из невыбранных дверей находится коза, ситуация меняется. Действительно, если мы не поменяем выбор, то вероятность того, что автомобиль находится за дверью номер 3, будет равняться $\frac{1}{3}$, так как это эквивалентно ситуации, когда нам сразу открыли нашу дверь. Ключевая идея: *ведущий всегда открывает дверь с козой.* Вероятность того, что автомобиль находится за дверью 1 или 2 равна $\frac{2}{3}$. Но когда нам показали, что за второй дверью находится коза, эта вероятность не изменилась

$$\mathbb{P}\{\text{за дверью 1 или 2 автомобиль} \mid \text{за одной из дверей 1 или 2 коза}\} = \tfrac{2}{3}$$
$$= \mathbb{P}\{\text{за дверью 1 или 2 автомобиль}\}$$
$$= \mathbb{P}\{\text{за оставшейся дверью среди 1 и 2 автомобиль}\},$$

   где последняя вероятность — это как раз и есть вероятность того, что мы выберем дверь, за которой находится автомобиль, если сменим выбор двери. Таким образом, если менять выбор, то с вероятностью $\frac{2}{3}$ мы выиграем автомобиль, а если не менять — то с вероятностью $\frac{1}{3}$.

7. Если $\mathbb{P}\{A\} > 0$ и $\mathbb{P}\{B\} > 0$, то выполнена **формула Байеса**

$$\mathbb{P}\{A \mid B\} = \mathbb{P}\{A\} \cdot \frac{\mathbb{P}\{B \mid A\}}{\mathbb{P}\{B\}}.$$

8. Если события $B_1, \dots, B_n$ образуют **разбиение** вероятностного пространства $\Omega$, т. е. $\Omega = B_1 \sqcup \dots \sqcup B_n$ и пересечения попарно пусты: $B_i \cap B_j = \emptyset$, $i \neq j$, то для любого события $A$ выполнена формула полной вероятности:

$$\mathbb{P}\{A\} = \sum_{i=1}^{n} \mathbb{P}\{A \mid B_i\} \cdot \mathbb{P}\{B_i\}.$$

**Пример.** Из урны, содержащей $a$ белых и $b$ черных шаров, извлекается наугад один шар и откладывается в сторону. Какова вероятность того, что извлеченный наугад второй шар окажется белым, если:

(a) первый извлеченный шар белый;

(b) цвет первого извлеченного шара остается неизвестным?

**Решение.**

(a) Пусть событие $A = \{\text{первый извлечённый шар белый}\}$ и событие $B = \{\text{второй извлечённый шар белый}\}$. Тогда нас просят найти $\mathbb{P}(B \mid A)$. Легко видеть, что $\mathbb{P}(B \mid A) = \frac{a-1}{a+b-1}$, так как нужно выбрать среди оставшихся $a + b - 1$ шаров один белый шар, которых осталось $a - 1$.

(b) Если цвет первого шара неизвестен, то нужно найти $\mathbb{P}(B)$. Пользуемся формулой полной вероятности:

$$\mathbb{P}(B) = \mathbb{P}(B \mid A) \cdot \mathbb{P}(A) + \mathbb{P}(B \mid \bar{A}) \cdot \mathbb{P}(\bar{A}) = \frac{a-1}{a+b-1} \cdot \frac{a}{a+b} + \frac{a}{a+b-1} \cdot \frac{b}{a+b} = \frac{a(a+b-1)}{(a+b-1)(a+b)} = \frac{a}{a+b}.$$

## Случайные величины

1. **Случайная величина** $\xi : \Omega \to \mathbb{R}$ — это числовая функция на вероятностном пространстве. Иначе говоря, случайная величина — это обычная числовая функция, на аргументах которой задано вероятностное распределение, так что можно говорить о вероятности события, при котором она принимает конкретное значение.

2. **Математическим ожиданием** случайной величины $\xi$ называется ее средневзвешенное значение

$$\mathbb{E}\xi \stackrel{\text{def}}{=} \sum_{\omega \in \Omega} \xi(\omega) \cdot \mathbb{P}\{\omega\}.$$

3. По определению, математическое ожидание линейной комбинации случайных величин равно соответствующей линейной комбинации их матожиданий, т. е. $\mathbb{E}[a\xi + b\psi] = a\mathbb{E}[\xi] + b\mathbb{E}[\psi]$.

**Пример.** Имеется $n$ пронумерованных писем и $n$ пронумерованных конвертов. Письма случайным образом раскладываются по конвертам (все $n!$ способов равновероятны). Найдите математическое ожидание числа совпадений номеров письма и конверта (письмо лежит в конверте с тем же номером).

**Решение.** Пусть $\xi$ — случайная величина, равная числу совпадений номеров письма и конверта. Введём случайную величину[3]:

$$\xi_i = \begin{cases} 1, & i\text{-е письмо попало в } i\text{-й конверт} \\ 0, & \text{иначе.} \end{cases}$$

Пользуясь линейностью математического ожидания и тем фактом, что $\mathbb{E}\xi_i = \mathbb{P}(\xi_i = 1) = \frac{(n-1)!}{n!} = \frac{1}{n}$, получаем:

$$\mathbb{E}\xi = \mathbb{E}\left(\sum_{i=1}^{n} \xi_i\right) = \sum_{i=1}^{n} \mathbb{P}(\xi_i = 1) = n \cdot \frac{1}{n} = 1.$$

[3] Такого рода случайные величины часто называют *индикаторными* случайными величинами.

4. Следующий результат, тривиально вытекающий из определения математического ожидания, имеет обширные алгоритмические, комбинаторные и пр. приложения. Можно даже сказать, что это исходный пункт т. н. **вероятностного метода**.

   *Пусть для случайной величины $\xi$ выполнено $\mathbb{E}\xi = A$, тогда существует исход $\omega$, что $\xi(\omega) \geq A$ (равно как существует исход, при котором выполнено противоположное неравенство).*

   **Пример.** Поверхность некоторой шарообразной планеты состоит из океана и суши (множество мелких островков). Суша занимает больше половины площади планеты. Также известно, что суша есть множество, принадлежащее *борелевской $\sigma$-алгебре* на сфере[4]. На планету хочет совершить посадку космический корабль, сконструированный так, что концы всех шести его ножек лежат на поверхности планеты. Посадка окажется успешной, если не меньше четырех ножек из шести окажутся на суше. Возможна ли успешная посадка корабля на планету?

   **Решение.** Введём индикаторную случайную величину:
$$\xi_i = \begin{cases} 1, & i\text{-я ножка оказалась на суше} \\ 0, & \text{иначе.} \end{cases}$$
   Тогда количество ножек на суше $\xi = \sum\limits_{i=1}^{6} \xi_i$. Отсюда следует, что
$$\mathbb{E}\xi = \sum_{i=1}^{6} \mathbb{E}\xi_i = \sum_{i=1}^{6} \mathbb{P}(\xi_i = 1) > 6 \cdot \frac{1}{2} = 3.$$
   Отсюда следует, что $\mathbb{P}(\xi \geqslant 4) > 0$, так как в противном бы случае $\mathbb{E}\xi$ было бы не больше 3. Значит, успешная посадка корабля возможна.

5. Две случайные величины $\xi$ и $\psi$ независимы, если для любых чисел $c$ и $d$ события $\xi^{-1}(c) = \{\omega \in \Omega : \xi(\omega) = c\}$ и $\psi^{-1}(d)$ независимы.

6. Математическое ожидание произведения независимых случайных величин равно произведению их математический ожиданий, т. е. $\mathbb{E}[\xi \cdot \psi] = \mathbb{E}[\xi] \cdot \mathbb{E}[\psi]$ (что можно проверить непосредственным вычислением).

7. **Условным математическим ожиданием** случайной величины $\xi$ при условии выполнения события $A$ определяется формулой
$$\mathbb{E}[\xi \mid A] \stackrel{\text{def}}{=} \sum_{\omega \in \Omega} \xi(\omega)\mathbb{P}\{\omega \mid A\}.$$

8. Если события $B_1, \ldots, B_n$ образуют **разбиение** вероятностного пространства $\Omega$, т. е. $\Omega = B_1 \sqcup \cdots \sqcup B_n$ и пересечения попарно пусты: $B_i \cap B_j = \emptyset$, $i \neq j$, то
$$\mathbb{E}[\xi] = \sum_{\omega \in \Omega} \xi(\omega)\mathbb{P}\{\omega\} = \sum_{\omega \in \Omega} \xi(\omega) \sum_{i=1}^{n} \mathbb{P}\{\omega \mid B_i\}\mathbb{P}\{B_i\} = \sum_{i=1}^{n} \mathbb{P}\{B_i\} \sum_{\omega \in \Omega} \xi(\omega)\mathbb{P}\{\omega \mid B_i\} = \sum_{i=1}^{n} \mathbb{E}[\xi \mid B_i]\mathbb{P}[B_i].$$

   **Пример (The Coin Flip Conundrum).** Два брата Орвилл и Уилбур[5] спорят за право первым испытать самолёт, который они построили. Для этого Уилбур, более сведующий в теории вероятностей, предложил решить спор следующей игрой. Братья по очереди бросают симметричную монетку (считаем, что сжульничать в броске невозможно). Если раньше выпадут два орла подряд, то выигрывает Орвил, а если раньше выпадет сначала орёл, а потом решка, то выигрывает Уилбур. Так как Орвилл плохо знает теорию вероятностей и полностью доверяет своему брату, он согласился на эту игру. Насколько честна такая игра?

   **Решение.** Покажем, что игра нечестна. Отдельно рассмотрим, сколько раз в среднем нужно бросить монетку, чтобы выпало два подряд орла. Пусть $X$ — число бросков монеты до первого выпадения двух

[4] Данное предложение означает, что можно считать вероятность попадания ножки на сушу.
[5] https://en.wikipedia.org/wiki/Wright_brothers

орлов подряд, $X_1$ — число бросков монеты до первого выпадения орла, $X_2$ — число бросков монеты после выпадения первого орла. Нам нужно найти $\mathbb{E}[X]$. Заметим, что $X = X_1 + X_2 \Rightarrow \mathbb{E}[X] = \mathbb{E}[X_1] + \mathbb{E}[X_2]$.

$$\mathbb{P}\{X_1 = k\} = \tfrac{1}{2} \cdot \tfrac{1}{2^{k-1}} = \tfrac{1}{2^k} \Rightarrow \mathbb{E}[X_1] = \sum_{i=1}^{\infty} \tfrac{k}{2^k} \overset{①}{=} 2,$$

где равенство ① доказывается в конце этого документа (см. пример про алгоритм типа “Лас-Вегас”). Итак, $\mathbb{E}[X] = 2 + \mathbb{E}[X_2]$. Пользуясь формулой для полного математического ожидания, получаем

$$\begin{aligned}\mathbb{E}[X_2] &= \tfrac{1}{2}\mathbb{E}[X_2 \mid \text{на шаге } (X_1+1) \text{ выпала решка}] + \tfrac{1}{2}\mathbb{E}[X_2 \mid \text{на шаге } (X_1+1) \text{ выпал орёл}] \\ &= \tfrac{1}{2}\left(1 + \mathbb{E}[X]\right) + \tfrac{1}{2} = \tfrac{1}{2}\mathbb{E}[X] + 1,\end{aligned}$$

откуда

$$\mathbb{E}[X] = 2 + \frac{1}{2}\mathbb{E}[X] + 1 \Rightarrow \mathbb{E}[X] = 6,$$

то есть среднее число бросков до первого выпадения двух подряд орлов равно 6.

Аналогчино рассмотрим, сколько раз в среднем нужно бросить монетку, чтобы выпал сначала орёл, а затем сразу решка. Пусть $Y$ — число бросков монеты до первого выпадения орла и следом за ним решка, $Y_1$ — число бросков монеты до первого выпадения орла, $Y_2$ — число бросков монеты после выпадения первого орла. Нам нужно найти $\mathbb{E}[Y]$. Заметим, что $Y = Y_1 + Y_2 \Rightarrow \mathbb{E}[Y] = \mathbb{E}[Y_1] + \mathbb{E}[Y_2]$. Так как $\mathbb{E}[Y_1] = \mathbb{E}[X_1] = 2$, то $\mathbb{E}[Y] = 2 + \mathbb{E}[Y_2]$. Пользуясь формулой для полного математического ожидания, получаем

$$\begin{aligned}\mathbb{E}[Y_2] &= \tfrac{1}{2}\mathbb{E}[Y_2 \mid \text{на шаге } (Y_1+1) \text{ выпала решка}] + \tfrac{1}{2}\mathbb{E}[Y_2 \mid \text{на шаге } (Y_1+1) \text{ выпал орёл}] \\ &= \tfrac{1}{2} + \tfrac{1}{2}\left(1 + \mathbb{E}[Y_2]\right) = \tfrac{1}{2}\mathbb{E}[Y_2] + 1 \Rightarrow \mathbb{E}[Y_2] = 2,\end{aligned}$$

откуда

$$\mathbb{E}[Y] = 2 + 2 = 4,$$

то есть среднее число бросков до первого выпадения орла, а затем сразу решки равно 4. Таким образом, Уилбур предложил нечестные правила игры.

9. **Дисперсией** $\mathbb{D}[\cdot]$ случайной величины называется средний квадрат ее отклонения от математического ожидания, т. е. $\mathbb{D}[\xi] = \mathbb{E}[(\xi - \mathbb{E}\xi)^2]$.

10. Если случайные величины $\xi$ и $\eta$ независимы, то дисперсия их суммы равна сумме их дисперсий, т. е. $\mathbb{D}[\xi + \eta] = \mathbb{D}[\xi] + \mathbb{D}[\eta]$.

11. Для оценки вероятности отклонения *неотрицательной* случайной величины $\xi$ от математического ожидания можно использовать **неравенство Маркова**

$$\mathbb{P}\{\xi \geq a\} \leq \frac{\mathbb{E}\xi}{a}.$$

12. Из неравенства Маркова и определения дисперсии немедленно следует **неравенство Чебышева**

$$\mathbb{P}\{\mid \xi - \mathbb{E}\xi \mid \geq a\} \leq \frac{\mathbb{D}\xi}{a^2}.$$

13. **Неравенство больших уклонений.** Обозначим через $\xi_n$ случайную величину, равную количеству выпавших орлов после $n$ подбрасываний “честной” монеты, а через $x_n = \frac{\xi_n}{n}$ — частоту выпавших орлов. При больших $n$ частота с очень большой вероятностью оказывается близкой к $\frac{1}{2}$. Имеет место следующая оценка, которую иногда называют **неравенством Чернова**. Она очень удобна в приложениях вероятностного метода в комбинаторике, а также в теории вычислений.

$$\mathbb{P}\left\{\left|\xi_n - \frac{n}{2}\right| > \varepsilon n\right\} = \mathbb{P}\left\{\left|x_n - \frac{1}{2}\right| > \varepsilon\right\} < 2\exp\left(-2\varepsilon^2 n\right).$$

Имеют место еще несколько неравенств подобного рода, если монетка несимметричная. Пусть орел выпадает с вероятностью $p$. Тогда

$$\begin{aligned}
&\mathbb{P}\{x_n < (1-\delta)p\} < \exp\left(-\frac{\delta^2 p}{2}\right), && 0 \le \delta \le 1,\\
&\mathbb{P}\{x_n > (1+\delta)p\} < \exp\left(-\frac{\delta^2 p}{3}\right), && 0 \le \delta \le 1,\\
&\mathbb{P}\{x_n > (1+\delta)p\} < \exp\left(-\frac{\delta p}{3}\right), && \delta > 1.
\end{aligned}$$

14. **Закон Больших Чисел (ЗБЧ).** Пусть $\{X_i\}_{i=1}^{\infty}$ — бесконечная последовательность случайных величин, имеющих одинаковое распределение и независимых в совокупности. Пусть кроме того, они имеют конечные дисперсии. Тогда среднее арифметическое первых $n$ случайных величин будет стремится *по вероятности* к $\mathbb{E}[X_1]$:

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow[n\to\infty]{\mathbb{P}} \mathbb{E}[X_1] \quad \overset{\text{def}}{\Longleftrightarrow} \quad \forall \varepsilon > 0 \hookrightarrow \lim_{n\to\infty} \mathbb{P}\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}[X_1]\right| < \varepsilon\right\} = 1.$$

**Пример.** Рассмотрим независимые бернуллиевские случайные величины $\xi_1, \xi_2, \ldots$, такие что

$$\forall i \in \mathbb{N} \ \hookrightarrow \ \xi_i = \begin{cases} 1 & \text{с вероятностью } p,\\ 0 & \text{с вероятностью } 1-p.\end{cases}$$

Заметим, что $\mathbb{E}[\xi_i] = 1\cdot\mathbb{P}\{\xi_i = 1\} + 0\cdot\mathbb{P}\{\xi_i = 0\} = \mathbb{P}\{\xi_i = 1\} = p$, $\mathbb{D}[\xi_i] = \mathbb{E}[\xi_i^2] - (\mathbb{E}[\xi_i])^2 = p - p^2 = p(1-p) < \infty$. Часто значение 1 бернуллиевской случайной величины интерпретируют как “успех”, а последовательность бернуллиевских случайных величин называют схемой испытаний Бернулли. Тогда из закона больших чисел следует, что среднее число “успехов” среди первых $n$ испытаний в схеме Бернулли стремится к вероятности “успеха” при $n \to \infty$:

$$\frac{1}{n}\sum_{i=1}^{n} \xi_i \xrightarrow[n\to\infty]{\mathbb{P}} p.$$

В частности, если долго бросать симметричную монетку, то среднее число выпадений орла будет стремится по вероятности к $\frac{1}{2}$.

**Пример (Вероятностный способ вычисления числа $\pi$).** Рассмотрим квадрат со стороной 2 и впишем в него окружность радиуса 1. Как известно площадь такого круга равна $\pi$, а площадь квадрата равна 4. Будем генерировать случайные точки из равномерного распределения на этом квадрате (то есть точка попадает в некоторую область в квадрате с вероятностью $\frac{\text{площадь области}}{\text{площадь квадрата}} = \frac{\text{площадь области}}{4}$). Рассмотрим случайную величину $\xi_i$, которая равна 1, если $i$-я точка попала в круг, и равна 0, если $i$-я точка не попала в круг. Иными словами, $\xi_i$ — бернуллиевская случайная величина с параметром $p = \frac{\pi}{4}$. Тогда из предыдущего примера получаем, что при больших $n$ выражение $\frac{4}{n}\sum\limits_{i=1}^{n}\xi_i$ с большой вероятностью хорошо аппроксимирует число $\pi$.

## Вероятностные алгоритмы. Время работы в среднем и в худшем случае

Как оценивать время вероятностных процедур? Существует 2 основных подхода.

1. Говорят, что алгоритм на входе размера $n$ работает *в худшем* случае время $T(n)$, если на любом входе размера $n$ он работает время $\leqslant T(n)$, причём равенство достигается.

2. Говорят, что алгоритм на входе размера $n$ работает *в среднем* время $T(n)$, если математическое ожидание времени его работы на входе размера $n$ равняется $T(n)$.

Рассмотрим следующую задачу. На вход подаётся массив из $2N$ букв, причём половину из них составляют буквы $a$, а другую половину — буквы $b$. Требуется найти номер любой ячейки, в которой лежит буква $a$. Рассмотрим 2 идеологии вероятностного решения такой задачи.

**Пример (алгоритм типа "Лас-Вегас").** Алгоритмом типа "Лас-Вегас" называется вероятностный алгоритм, который *всегда на выходе даёт корректный результат.* Рассмотрим следующий алгоритм.

**Алгоритм 1** findingA_LV

**Вход:** Массив букв $A$ размера $N$
**повторять**
  Случайно равновероятно выбрать индекс $i \in \{1, 2, \ldots, N\}$
**пока** $A[i] \neq a$
**return** $i$

Время работы в худшем случае Алгоритма 1 **равно бесконечности**. Действительно, есть ненулевая вероятность, что на каждом шаге алгоритм будет выбирать ячейку, в которой лежит буква $b$, а значит, мы не можем ограничить количество итераций. Однако алгоритм останавливается за конечное число шагов с вероятностью 1. Более того, среднее число выполненных итераций равно

$$A \stackrel{\text{def}}{=} \sum_{i=1}^{\infty} \frac{i}{2^i} = \sum_{i=1}^{\infty} \frac{i+1}{2^i} - \sum_{i=1}^{\infty} \frac{1}{2^i} = 2\left(\sum_{i=1}^{\infty} \frac{i}{2^i} - \frac{1}{2}\right) - 1 = 2\left(A - \frac{1}{2}\right) - 1 = 2A - 2 \Rightarrow A = 2,$$

то есть алгоритм *в среднем делает 2 итерации.*

**Пример (алгоритм типа "Монте-Карло").** Алгоритмом типа "Монте-Карло" называется вероятностный алгоритм, который *может давать на выходе неправильный результат с некоторой (как правило маленькой) вероятностью.* Рассмотрим следующий алгоритм.

**Алгоритм 2** findingA_MC

**Вход:** Массив букв $A$ размера $N$; натуральное число $k$
$i := 0$
**повторять**
  Случайно равновероятно выбрать индекс $i \in \{1, 2, \ldots, N\}$
**пока** $k = i$ или $A[i] \neq a$
**return** $i$

Время работы в худшем случае Алгоритма 2 **равно** $k$. Вероятность получить правильный ответ равна $1 - \frac{1}{2^k}$. Более того, среднее число выполненных итераций равно

$$B \stackrel{\text{def}}{=} \sum_{i=1}^{k} \frac{i}{2^i} = \sum_{i=1}^{k} \frac{i+1}{2^i} - \sum_{i=1}^{k} \frac{1}{2^i} = 2\left(\sum_{i=1}^{k} \frac{i}{2^i} - \frac{1}{2} - \frac{k}{2^k}\right) - 1 = 2\left(B - \frac{1}{2} + \frac{k+1}{2^{k+1}}\right) - 1 = 2B - 2 + \frac{k+1}{2^{k+1}} \Rightarrow B = 2 - \frac{k+1}{2^{k+1}},$$

то есть алгоритм *в среднем делает меньше 2-х итераций.*