Общее понятие вероятностного пространства

Высшая школа цифровой культуры
Университет ИТМО
dc@itmo.ru

Содержание

1 Геометрическая вероятность

1.1 Определение геометрической вероятности

До сих пор мы рассматривали пространство элементарных исходов Ω, содержащее конечное число элементарных исходов, однако на практике такие пространства встречаются не то что не всегда, а скорее не часто. Для мотивировки рассмотрения дальнейшего, попробуем решить так называемую задачу о встрече и, с ее помощью, определить понятие «геометрической вероятности».

Пусть двое, скажем, Петя и Маша, условились встретиться в определенном месте между 10 и 11 часами. Пришедший первым ждет другого в течение 10 минут, и, если так и не дожидается, уходит. Какова вероятность того, что Петя и Маша встретятся, если каждый из них приходит независимо от другого в любое время в течение указанного часа?

Интересно, что приведенная формулировка задачи кроме основного вопроса неявно заставляет ответить на еще кучу вопросов. Что значит «случайный момент времени»? Ведь между 10 и 11 часами бесконечное число «моментов времени». А что значит «не зависит от времени прихода другого»? Попробуем ответить на эти вопросы.

Отождествим отрезок времени от 10 до 11 часов с отрезком $[0, 1]$, а время прихода конкретного человека — с точкой на этом отрезке. Тогда под случайным моментом времени логично понимать то, что шанс попасть в любую точку рассматриваемого отрезка одинаков. Классическое определение здесь не подходит, ведь тогда шанс попасть в конкретную точку просто равен нулю (число точек бесконечно), что не дает никакой конкретики. Давайте попробуем поступить несколько иначе.

Разделим отрезок пополам. Тогда логично считать, что шансов попасть в правую часть отрезка столько же, сколько и в левую. Иными словами, вероятности попасть в правую половину отрезка и в левую у точки одинаковы. Дальше каждый из полученных отрезков можно снова поделить пополам и по тем же соображениям получается, что шансы попасть в четверти отрезка одинаковы. Продолжая процесс, получим, что вероятности попасть в десятые, сотые, тысячные и так далее части отрезка должны быть одинаковы. Из этих соображений напрашивается следующее определение.

Определение 1.1.1 (Наводящее определение) *Пусть имеются отрезки L и l, причем $l \subset L$. Вероятность того, что случайно выбранная точка отрезка L принадлежит отрезку l, равна*

$$\mathsf{P}(l) = \frac{\textit{длина}(l)}{\textit{длина}(L)},$$

Легко видеть, что большинство свойств классической вероятности выполняются и в этой ситуации: вероятность находится в пределах от 0 до 1, вероятность пустого множества рана нулю, вероятность суммы (для непересекающихся отрезков) равна сумме вероятностей и др.

Аналогичным образом можно от одномерного случая (прямой и отрезка на ней) перейти к двумерному (плоскость и фигура на ней), к трехмерному (пространство и тело в нем) и так далее. Ясно, что по аналогии с одномерным случаем, вместо длины будем использовать площадь (на плоскости), объем (в пространстве) или просто «меру» $\lambda(A)$ множества A в пространстве $\mathbb{R}^m$.

Определение 1.1.2 *Пусть $\Omega \subset \mathbb{R}^m$, $A \subset \Omega$. Вероятность того, что случайно выбранная точка из Ω принадлежит множеству A, равна*

$$\mathsf{P}(\cdot \in A) = \frac{\lambda(A)}{\lambda(\Omega)}$$

Еще раз отметим, что в одномерном случае, то есть при $m = 1$, функция $\lambda(A)$ дает «длину» множества A, при $m = 2$ площадь, при $m = 3$ объем. К понятию «случайности» нужно относиться с большой осторожностью. Чуть ниже мы это покажем на так называемом парадоксе Бертрана.

Вернемся к нашему примеру. Что значит фраза «не зависит от времени прихода другого»? Математически это означает, что пространство элементарных исходов имеет вид $\Omega = [0, 1] \times [0, 1]$, и все исходы $(x, y) \in \Omega,\ x, y \in [0, 1]$ равновозможны.

Пример 1.1.1 (Задача о встрече) *Пусть двое, скажем, Петя и Маша, условились встретиться в определенном месте между* 10 *и* 11 *часами. Пришедший первым ждет другого в течение* 10 *минут, и, если так и не дожидается, уходит. Какова вероятность того, что Петя и Маша встретятся, если каждый из них приходит независимо от другого в любое время в течение указанного часа?*

Как мы уже сказали, $\Omega = [0, 1] \times [0, 1]$. Пусть Петя пришел в момент времени $x \in [0, 1]$, а Маша пришла в момент времени $y \in [0, 1]$. Для того, чтобы они встретились, нужно, чтобы между временем прихода Пети и Маши прошло не более 10 *минут* (1/6 *часа*)*. Это значит, что встрече Пети и Маши благоприятствуют те элементарные исходы $(x, y) \in \Omega$, для которых*

$$|x - y| \le \frac{1}{6} \ \Leftrightarrow\ -\frac{1}{6} \le x - y \le \frac{1}{6}$$

Ясно, что интересующая нас область ограничена прямыми $y = x \pm \frac{1}{6}$, см. рисунок 1. Осталось вычислить площадь $\lambda(A)$ заштрихованной области.

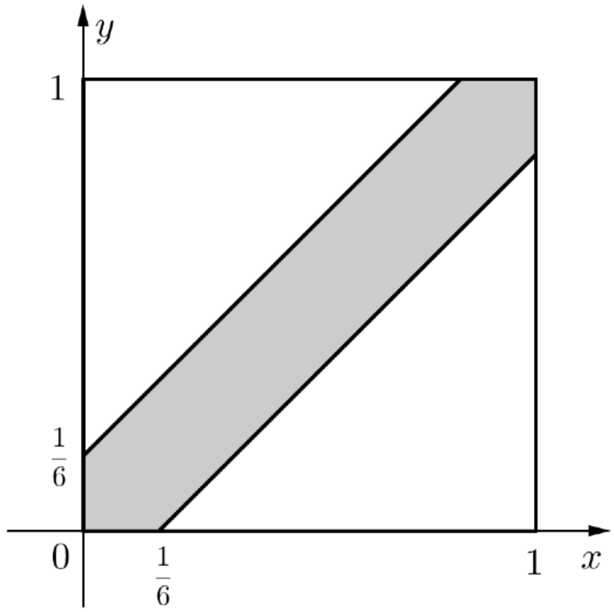

Рис. 1: Задача о встрече

Это можно сделать, например, вычитанием из площади квадрата площадей двух равных (белых) прямоугольных треугольников:

$$\lambda(A) = 1 - \left(\frac{5}{6}\right)^2 = \frac{11}{36}.$$

Значит, так как площадь квадрата $\lambda(\Omega)$ равна единице, то

$$\mathsf{P}(A) = \frac{\lambda(A)}{\lambda(\Omega)} = \frac{11}{36}.$$

Введенное нами определение хоть и обладает свойствами, похожими на классический случай, в то же время имеет существенные особенности. Во-первых, теперь из того, что $\mathsf{P}(A) = 0$ вовсе не следует, что событие A невозможно. Ну, например, так как длина точки равна нулю, то вероятность попасть в конкретную точку на выбранном отрезке равна 0, хотя событие — попасть в точку, не невозможно.

Возникает и еще один немаловажный вопрос. А любое ли подмножество пространства элементарных исходов Ω является событием? Иными словами, у каждого ли подмножества множества Ω мы можем вычислить длину, площадь, объем? Для всех ли этих подмножеств определено понятие меры? Оказывается, что нет. Для особо стойких приведем пример.

Пример 1.1.2 (Множество Витали) *Покажем, что на отрезке существует множество, мера (длина) которого не определена. Из свойств длины нам потребуются следующие: во-первых, длина множества не изменяется, если все его точки передвинуть на какой-либо фиксированный вектор;*

во-вторых, длина счетного объединения попарно непересекающихся множеств есть сумма длин этих множеств.

Рассмотрим окружность единичного радиуса (то же самое, что отрезок $[0, 2\pi]$*) и выберем произвольное иррациональное число* α*. Так как* α *иррационально, то равенство* $2\pi\alpha k = 2\pi\alpha n$ *при* $k, n \in \mathbb{Z}$ *возможно лишь в случае, когда* $k = n$*. Используем это следующим образом. Пусть* x *– произвольная точка окружности. Отнесем в один класс все точки, получающиеся из точки* x *поворотом на углы* $2\pi\alpha n$*,* $n \in \mathbb{Z}$ *(по замечанию, сделанному выше, мы не получим одинаковых точек). С любой другой точкой окружности можно аналогичным образом связать класс точек, получающийся из нее поворотами на углы* $2\pi\alpha n$*,* $n \in \mathbb{Z}$*. Эта точка либо порождает новый класс, либо полностью повторяет какой-то уже имеющийся (все одинаковые классы мы отождествляем). Тем самым, каждая точка окружности входит в какой-то класс, причем только в один, так как разные классы не пересекаются.*

Введем в рассмотрение множество A_0 *– множество, содержащее ровно по одному элементу из каждого класса. Рассмотрим множества* $A_n = A_0 + 2\pi\alpha n$*,* $n \in \mathbb{Z}$*. Так как все точки одного класса можно получить, поворачивая одну точку класса, множество* A_0 *содержит по одной точке каждого класса, а каждая точка окружности входит хотя бы в один класс, то*

$$[0, 2\pi] = \bigcup_{n=-\infty}^{+\infty} A_n.$$

Предположим, что множество A_0 *имеет длину* $\lambda(A_0)$*. Тогда все множества* A_n *имеют ту же длину* $\lambda(A_n)$*, так как они получены из множества* A_0 *поворотом (сдвигом). Все эти множества не пересекаются, а значит*

$$2\pi = \lambda([0, 2\pi]) = \lambda\left(\bigcup_{n=-\infty}^{+\infty} A_n\right) = \sum_{n=-\infty}^{+\infty} \lambda(A_n) = \sum_{n=-\infty}^{+\infty} \lambda(A_0).$$

Последняя сумма либо равна нулю, если $\lambda(A_0) = 0$*, либо равна бесконечности, если* $\lambda(A_0) > 0$*. Это показывает, что множество* A_0 *не имеет длины.*

1.2 Парадокс Бертрана

Рассмотрим достаточно простую задачу, получившую широкую огласку в 19 веке. Пусть дана окружность радиуса 1. Какова вероятность того, что случайная хорда этой окружности имеет длину большую, чем $\sqrt{3}$ (сторона правильного вписанного треугольника)? Оказывается, ответ на этот вопрос в большей степени зависит от того, что понимать под «случайной хордой».

1. Случайную хорду можно определить как отрезок, соединяющий две случайные точки на окружности. Что означают слова «случайная точка» на окружности? Действуя в рамках геометрической вероятности это означает, что $\Omega =$ окружность, $A \subset \Omega$ и

$$\mathsf{P}(\cdot \in A) = \frac{\lambda(A)}{\lambda(\Omega)} = \frac{\lambda(A)}{2\pi},$$

так как $\lambda(\Omega)$ – длина окружности, равна 2π. Не нарушая общности можно считать, что первая точка фиксирована. Отметим пунктиром «крайние случаи», когда длина хорды, проведенной из фиксированной точки, равна $\sqrt{3}$ (рисунок 2). Тогда вторая точка x должна попадать на дугу окружности

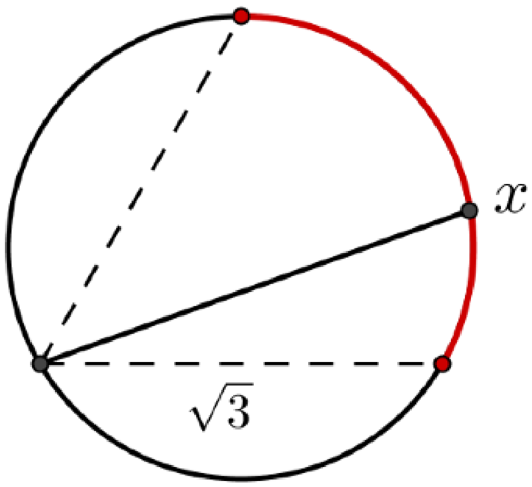

Рис. 2: Первый вариант решения задачи

красного цвета. Этот сегмент составляет треть всей окружности (так как длины сторон равностороннего треугольника вписанного в окружность радиуса 1, равны $\sqrt{3}$). А значит, $\lambda(A) = \frac{2\pi}{3}$ и

$$\mathsf{P}(\text{длина хорды больше, чем } \sqrt{3}) = \frac{1}{3}.$$

2. Определим случайную хорду иначе. Возьмем произвольную случайную точку внутри круга, ограниченного рассматриваемой окружностью, и будем считать, что эта точка является серединой случайной хорды. Что теперь является «случайной точкой» внутри круга? В рамках геометрической вероятности это значит, что $\Omega =$ круг, $A \subset \Omega$ и

$$\mathsf{P}(\cdot \in A) = \frac{\lambda(A)}{\lambda(\Omega)} = \frac{\lambda(A)}{\pi},$$

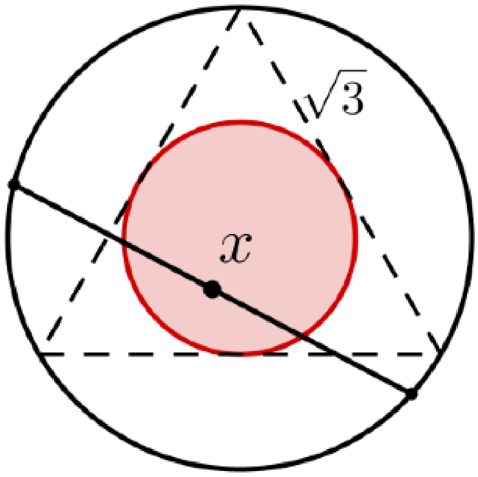

Рис. 3: Второй вариант решения задачи

так как $\lambda(\Omega)$ – площадь круга, равна π. Какое множество точек благоприятствует нашему событию? Это круг (рисунок 3), радиус которого равен $\frac{1}{2}$. Значит,

$$\mathsf{P}(\text{длина хорды больше, чем } \sqrt{3}) = \frac{1}{4}.$$

3. Пусть случайная хорда определяется случайным направлением и случайным расстоянием от центра окружности (хорда перпендикулярна направлению и отстоит от центра на случайное расстояние x). Случайное направление может быть задано центром окружности и случайной точкой на окружности. Не нарушая общности, зафиксируем какое-то направление. Что такое случайное расстояние? Согласно геометрической вероятности, $\Omega =$ радиус (отрезок), $A \subset \Omega$ и

$$\mathsf{P}(\cdot \in A) = \frac{\lambda(A)}{\lambda(\Omega)} = \frac{\lambda(A)}{1},$$

так как радиус окружности равен единице. Длина хорды на расстоянии x от центра окружности вычисляется по теореме Пифагора и равна $2\sqrt{1-x^2}$. Решая неравенство $2\sqrt{1-x^2} > \sqrt{3}$ получаем, что $x \in [0, 1/2]$, откуда

$$\mathsf{P}(\text{длина хорды больше, чем } \sqrt{3}) = \frac{1}{2}.$$

Важно понимать, что все три способа решения рассмотренной задачи являются правильными. Точнее так, каждый рассмотренный способ – это решение своей конкретной задачи, где под случайной хордой понимается некоторый конкретный объект. Итого, мы решили три разные задачи. Так что парадокс Бертрана на самом деле никаким парадоксом не является, а был

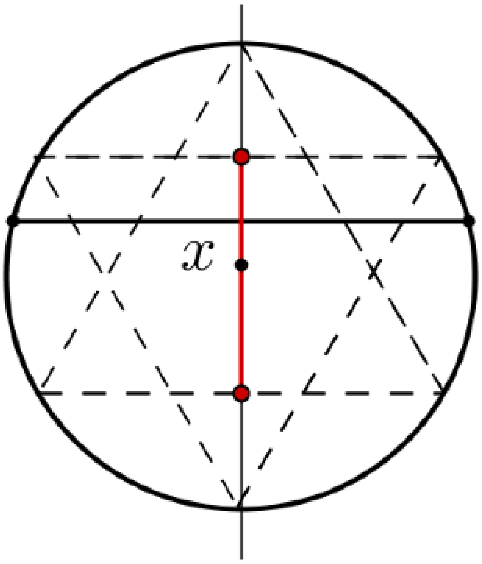

Рис. 4: Третий вариант решения задачи

назван так в то время, когда люди считали, что понятию «случайный» можно дать одно единственно верное определение.

Итак, давайте отметим такую важную мысль. Построение вероятностного пространства – это построение вероятностной модели, на основе которой делаются выводы. Если модель построена неверно, то и выводы будут неверными. Например, если из результата нужно сделать вывод, «пролезет» ли хорда через какое-то отверстие размера $\sqrt{3}$ с вероятностью большей, чем 0.7, то можно получить проблемы.

2 Общее определение вероятностного пространства

Этот пункт является в большей мере теоретическим. Мы дадим общее определение события, вероятности события и вероятностного пространства.

2.1 σ-алгебра событий

Итак, как мы уже неоднократно говорили, в результате некоторого случайного эксперимента возникает пространство элементарных исходов Ω. Оно может быть конечным, и этот случай подробно освещался в первых двух лекциях, а может быть бесконечным (и, как мы видели в примере с геометрической вероятностью, даже несчетным). Приведем пример простого и востребованного в примерах пространства со счетным множеством элементарных исходов.

Пример 2.1.1 *Пусть подбрасывается монета, которая не может встать на ребро, до первого выпадения орла. Тогда элементарных исходов, очевидно,* *счетное число, и их можно описать следующим образом:*

$$\omega_1 = O,\ \omega_2 = PO,\ \omega_3 = PPO,\ ..., \omega_n = PP...PO,\ ...,$$

где исход ω_n означает, что орел выпал первый раз на n-ом бросании.

При определении события в общем случае нужно соблюдать некоторую осторожность. Как уже было отмечено в геометрической вероятности, не всякое подмножество прямой, плоскости имеет длину, площадь, а значит не каждое подмножество может трактоваться, как событие. Забегая вперед, на множестве событий нам нужно будет определить вероятность (или, как говорят математики, вероятностную меру), поэтому задание множества событий — суть задание области определения этой вероятностной меры. На самом деле, все дальнейшие свойства (как и требования к σ-алгебре) не должны выглядеть пугающими. Ведь и правда, говоря о вероятности события, мы хотим иметь в арсенале такие события (произошедшие в результате эксперимента), как: $\varnothing$ – невозможное (условно, ничего не произошло) или Ω – достоверное (что-то обязательно произошло). Кроме того, если произошли события A, B, то мы хотим рассматривать события $A \cup B$ (произошло или A, или B) и $A \cap B$ (произошло и A, и B), или событие $\overline{A}$ – противоположное к A. Операции объединения и пересечения важно уметь распространять и на бесконечное (счетное) число слагаемых, что математики обычно и делают, и что сделано в определении σ-алгебры.

Дадим, наконец, строгое определение понятию σ-алгебры.

Определение 2.1.1 *Система подмножеств Σ пространства элементарных исходов Ω называется σ-алгеброй, если:*

1. *$\varnothing \in \Sigma$ (пустое множество входит в Σ);*
2. *Если $A \in \Sigma$, то и $\overline{A} \in \Sigma$ (вместе с каждым множеством A, входящим в Σ, дополнение $\overline{A}$ тоже входит в Σ);*
3. *Если $A_1, ..., A_n, ... \in \Sigma$, то и*

$$\bigcup_{i=1}^{\infty} A_i \in \Sigma.$$

Иными словами, вместе с последовательностью множеств, принадлежащих Σ, объединение этих множеств тоже принадлежит Σ.

Отметим классические примеры σ-алгебр подмножеств множества Ω.

Пример 2.1.2 (Тривиальная σ-алгебра) *Одна из крайностей – σ-алгебра Σ, состоящая из двух множеств: $\Sigma = \{\varnothing, \Omega\}$. Такую σ-алгебру часто называют тривиальной.*

Пример 2.1.3 (σ-алгебра всех подмножеств множества Ω)
Наоборот, другой крайностью является множество всех подмножеств исходного множества X, то есть $\Sigma = 2^{\Omega}$.

Например, если $\Omega = \{a, b, c\}$, то множество всех подмножеств 2^{Ω} состоит из множеств

$$2^{\Omega} = \{\{\varnothing\}, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}.$$

Легко проверить, что произвольная σ-алгебра (наряду со свойствами из определения) обладает следующими свойствами: она содержит все пространство Ω; вместе с любыми двумя множествами A и B, она содержит их объединение $A \cup B$, пересечение $A \cap B$, разность $A \setminus B$. Итого, справедлива следующая лемма.

Лемма 2.1.1 *Пусть Σ – σ-алгебра подмножеств Ω, $A, B \in \Sigma$, тогда*

1. $\Omega \in \Sigma$;
2. $A \cup B \in \Sigma$;
3. $A \cap B \in \Sigma$;
4. $A \setminus B \in \Sigma$;
5. *Если* $A_1, ..., A_n, ... \in \Sigma$, *то и*

$$\bigcap_{i=1}^{\infty} A_i \in \Sigma.$$

Доказательство. Доказательства моментально следуют из правил теории множеств.
1. Так как $\varnothing \in \Sigma$ (свойство 1 в определении), то $\Omega = \overline{\varnothing} \in \Sigma$ (свойство 2 в определении).
2. Это свойство является частным случаем свойства 3 в определении. Достаточно положить $A_1 = A$, $A_2 = B$, $A_3 = A_4 = ... = \varnothing$.
3. Это свойство немедленно следует из равенства $A \cap B = \overline{\overline{A} \cup \overline{B}}$.
4. Это свойство немедленно следует из равенства $A \setminus B = A \cap \overline{B}$, предыдущего свойства и того, что $\overline{B} \in \Sigma$.
5. Это свойство следует из равенств

$$\bigcap_{i=1}^{\infty} A_i = \overline{\bigcup_{i=1}^{\infty} \overline{A_i}}.$$

□

На самом деле в счетном объединении (пересечении) нет ничего необычного. Возвращаясь к уже рассмотренному примеру с подбрасыванием монеты до первого появления орла, событие A — в результате эксперимента выпал орел, будет записано, как

$$A = \omega_1 \cup \omega_2 \cup \ldots \cup \omega_n \cup \ldots = \bigcup_{i=1}^{\infty} \omega_i.$$

Введенное определение не противоречит тому, что мы рассматривали в случае, когда пространство элементарных исходов Ω конечно, так как множество всех подмножеств произвольного множества всегда образует σ-алгебру. Однако, как уже было отмечено, следует соблюдать некоторую осторожность, так как задать меру на множестве всех подмнождеств рассматриваемого пространства элементарных исходов Ω можно не всегда.

Дадим, наконец, важное для дальнейшего определение события.

Определение 2.1.2 *Пусть дано пространство элементарных исходов Ω и σ-алгебра выделенных подмножеств Σ. Событием называется произвольный элемент σ-алгебры Σ.*

Итак, в качестве событий в дальнейшем мы будем рассматривать только подмножества введенной σ-алгебры.

2.2 Вероятностная мера и вероятностное пространство

До настоящего момента мы определили пару Ω – пространство элементарных исходов и σ-алгебру Σ выделенных подмножеств Ω, элементы которой мы назвали событиями. Покажем, как вводится вероятностная мера, заданная на σ-алгебре, согласно аксиоматике Колмогорова.

Определение 2.2.1 (Аксиоматика Колмогорова) *Вероятностной мерой P, заданной на σ-алгебре Σ выделенных подмножеств Ω, называется такая функция $\mathsf{P} : \Sigma \to [0, 1]$, что*

1. $\mathsf{P}(\varnothing) = 0$;

2. $\mathsf{P}(\Omega) = 1$;

3. *Пусть* $A_1, \ldots, A_n, \ldots \in \Sigma$, $A = \bigcup\limits_{i=1}^{\infty} A_i$, $A_i \cap A_j = \varnothing$ *при* $i \neq j$, *тогда*

$$\mathsf{P}(A) = \mathsf{P}\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathsf{P}(A_i).$$

Поясним введенное определение. Во-первых, вероятность – это число в диапазоне от 0 до 1. Первое условие утверждает, что вероятность невозможного события ($\varnothing$) равна нулю, а второе, что вероятность достоверного события (Ω), то есть события, что что-нибудь да произойдет, равна единице (это условие часто называют условием нормировки). Третье же условие означает, что если целое разбить на непересекающиеся части, то вероятность целого складывается из вероятностей частей (частей может быть и счетное число).

Определение 2.2.2 *Тройка* $(\Omega, \Sigma, \mathsf{P})$*, где* Ω *– некоторое множество (называемое нами пространством элементарных исходов),* Σ *–* σ*-алгебра выделенных подмножеств* Ω*, и* P *– вероятностная мера, заданная на* Σ*, называется вероятностным пространством.*

Вспомнив свойства введенного нами понятия вероятности в случае конечного пространства Ω легко увидеть, что для него аксиоматика Колмогорова справедлива, а значит определением выше мы обобщили введенное ранее понятие.

Для геометрического определения вероятности все аксиомы тоже справедливы. Первые два свойства очевидны. Последнее, в обывательском смысле, тоже очевидно, ведь если некоторую фигуру разбить на непересекающиеся фигуры, имеющие площадь, то сумма площадей составных частей дает площадь всей фигуры.

Приведем еще один пример вероятностного пространства.

Пример 2.2.1 (Геометрическое распределение) *Пусть эксперимент заключается в подбрасывании монеты, которая не может встать на ребро, до первого выпадения орла. Как мы видели, пространство элементарных исходов бесконечно, а элементарный исход* ω_n *заключается в том, что орел выпал ровно на* n*-ом броске, а до этого выпадала решка. Если подбрасывания независимы, монета в каждом броске выпадает орлом с вероятностью* $p \in (0, 1)$ *и решкой с вероятностью* $q = 1 - p$*, то логично положить*

$$\mathsf{P}(\omega_n) = q^{n-1} \cdot p.$$

Пусть Σ *–* σ*-алгебра всех подмножеств* Ω*, тогда определим вероятность события* A *по формуле*

$$\mathsf{P}(A) = \sum_{i:\ \omega_i \in A} \mathsf{P}(\omega_i).$$

Проверим, что выполнены аксиомы Колмогорова.

$$\mathsf{P}(\varnothing) = \sum_{i:\ \omega_i \in \varnothing} \mathsf{P}(\omega_i) = 0,$$

$$\mathsf{P}(\Omega) = \sum_{i=1}^{\infty} \mathsf{P}(\omega_i) = \sum_{i=1}^{\infty} q^{i-1}p = p\sum_{i=1}^{\infty} q^{i-1} = \frac{p}{1-q} = 1,$$

где предпоследнее равенство есть не что иное, как сумма бесконечно убывающей геометрической прогрессии ($q \in (0,1)$)*.*

Проверим выполнение последней аксиомы для двух множеств, тогда для конечного числа множеств равенство может быть доказано по индукции, а для бесконечного – с помощью некоторых фактов из теории рядов. Пусть $A_1, A_2 \in \Sigma$, $A_1 \cap A_2 = \varnothing$, *тогда событие* $A = A_1 \cup A_2$ *состоит из тех* ω_i, *которые входят либо в* A_1, *либо в* A_2, *причем ни один элементарный исход не входит одновременно в* A_1 *и* A_2. *Тогда*

$$\mathsf{P}(A) = \sum_{i:\ \omega_i \in A} \mathsf{P}(\omega_i) = \sum_{i:\ \omega_i \in A_1} \mathsf{P}(\omega_i) + \sum_{i:\ \omega_i \in A_2} \mathsf{P}(\omega_i) = \mathsf{P}(A_1) + \mathsf{P}(A_2).$$

Определенная нами вероятностная мера обладает такими же свойствами, которые были введены и доказаны ранее. Сформулируем их.

Лемма 2.2.1 *Пусть* $(\Omega, \Sigma, \mathsf{P})$ *– вероятностное пространство,* $A, B \in \Sigma$, *тогда:*

1. *Если* $A \subset B$, *то* $\mathsf{P}(A) \le \mathsf{P}(B)$;
2. $0 \le \mathsf{P}(A) \le 1$;
3. *Если* $A \subset B$, *то* $\mathsf{P}(B \setminus A) = \mathsf{P}(B) - \mathsf{P}(A)$;
4. $\mathsf{P}(A \cup B) = \mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}(A \cap B)$;
5. $\mathsf{P}(\overline{A}) = 1 - \mathsf{P}(A)$;
6. *Пусть* $A_1, ..., A_n, ... \in \Sigma$, *тогда*
$$\mathsf{P}\left(\bigcup_{i=1}^{\infty} A_i\right) \le \sum_{i=1}^{\infty} \mathsf{P}(A_i).$$
7. *Пусть* $A_1, ..., A_n \in \Sigma$, *тогда*
$$\mathsf{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathsf{P}(A_i) - \sum_{i<j} \mathsf{P}(A_i \cap A_j) + \sum_{i<j<k} \mathsf{P}(A_i \cap A_j \cap A_k) + ... +$$
$$(-1)^{n-1}\mathsf{P}(A_1 \cap A_2 \cap ... \cap A_n).$$

Доказательство. Многие свойства доказываются аналогично тому, как они доказывались в первой лекции, поэтому докажем только те, для которых доказательство немного поменяется.
1. Легко проверить, что $B = (B \setminus A) \cup A$, причем $(B \setminus A) \cap A = \varnothing$, тогда

$$\mathsf{P}(B) = \mathsf{P}((B \setminus A) \cup A) = \mathsf{P}(B \setminus A) + \mathsf{P}(A) \geq \mathsf{P}(A).$$

3. Так как $B = (B \setminus A) \cup A$, причем $(B \setminus A) \cap A = \varnothing$, то

$$\mathsf{P}(B) = \mathsf{P}(B \setminus A) + \mathsf{P}(A),$$

откуда и следует утверждение.
4. Так как $A \cup B = (A \setminus (A \cap B)) \cup (B \setminus (A \cap B)) \cup (A \cap B)$, и все множества попарно не пересекаются, то

$$\mathsf{P}(A \cup B) = \mathsf{P}\left(A \setminus (A \cap B)\right) + \mathsf{P}\left(B \setminus (A \cap B)\right) + \mathsf{P}\left(A \cap B\right) =$$
$$\mathsf{P}(A) - \mathsf{P}\left(A \cap B\right) + \mathsf{P}(B) - \mathsf{P}\left(A \cap B\right) + \mathsf{P}\left(A \cap B\right) =$$
$$\mathsf{P}(A) + \mathsf{P}(B) - \mathsf{P}\left(A \cap B\right).$$

$\square$

2.3 Условная вероятность, независимость и независимость в совокупности

Нами был не зря подробно изучен и аргументирован случай конечного пространства Ω. Без всяких изменений на общий случай переносятся понятия условной вероятности, независимости событий и независимости событий в совокупности.

Определение 2.3.1 (Понятие условной вероятности в общем случае) *Пусть дано вероятностное пространство* $(\Omega, \Sigma, \mathsf{P})$. *Условной вероятностью события B при условии, что произошло событие A в случае, когда* $P(A) > 0$, *называется вероятность*

$$\mathsf{P}(B|A) = \frac{\mathsf{P}(A \cap B)}{\mathsf{P}(A)}.$$

Отметим также важный момент, который сразу следует из определения. Если $\mathsf{P}(A) > 0$ и $\mathsf{P}(B) > 0$, то

$$\mathsf{P}(A \cap B) = \mathsf{P}(A) \cdot \mathsf{P}(B|A) = \mathsf{P}(B) \cdot \mathsf{P}(A|B).$$

Последнее равенство часто называется теоремой умножения вероятностей. По индукции оно может быть распространено на большее число слагаемых.

Чтобы разрядить обилие теории, давайте разберем так называемый парадокс Монти Холла.

Пример 2.3.1 (Парадокс Монти Холла) *Представьте, что вы стали участником игры, где вам нужно выбрать наугад одну из трех дверей. За одной дверью находится автомобиль, а за двумя другими – козы. Вы выбираете одну из дверей, а затем ведущий, который знает, где находится автомобиль, открывает одну из оставшихся дверей, за которой обязательно находится коза. После этого он спрашивает вас: не желаете ли вы изменить свой выбор и открыть другую, ранее никем не выбранную дверь? Итак, вопрос: увеличатся ли ваши шансы выиграть автомобиль, если вы примите предложение ведущего и поменяете свой выбор?*

Обычно люди отвечают, что шансы не увеличатся, руководствуясь, вероятно, следующей логикой: «дверей осталось две, шансы 50 *на* 50*, а значит нет никакой разницы: менять дверь или не менять». Вы согласны? А я отвечу, что оказывается, если изменить дверь, то шансы выиграть увеличатся аж в два раза! Почему? Давайте разбираться.*

Ясно, что в самом начале вероятность выбрать дверь, за которой находится автомобиль, равна $\frac{1}{3}$*, а дверь, за которой находится коза –* $\frac{2}{3}$*. А что происходит в игре? Давайте разберем варианты.*

Пусть изначально игрок выбрал дверь, за которой находится коза (вероятность этого $\frac{2}{3}$*). Тогда ведущий открывает дверь со второй козой и, если игрок не меняет своего выбора, то он проигрывает, а если меняет, то выигрывает.*

Пусть изначально игрок выбрал дверь с автомобилем (вероятность этого $\frac{1}{3}$*), а ведущий открыл одну из двух дверей с козой. Если игрок не меняет своего выбора, то он выигрывает автомобиль, а если меняет – то проигрывает.*

Смотря на эти варианты видно, что игрок уходит с козой при смене двери только в том случае, когда изначально была выбрана дверь, за которой находится автомобиль (а вероятность этого $\frac{1}{3}$*), а в другом случае, с вероятностью* $\frac{2}{3}$*, он выигрывает автомобиль. Итого, выгоднее дверь менять!*

Аналогично тому, как это было сделано в простейшем случае, мотивируется определение независимых событий.

Определение 2.3.2 (Понятие независимости в общем случае) *События A и B называются независимыми, если*

$$\mathsf{P}(A \cap B) = \mathsf{P}(A)\mathsf{P}(B).$$

Часто приходится рассматривать не одно, два, а куда больше событий. Для них, как мы знаем, вводится понятие независимости в совокупности.

Определение 2.3.3 (Независимость в совокупности в общем случае) *События $A_1, A_2, ..., A_n$ называются независимыми в совокупности, если* *для любого* $1 \le k \le n$ *и любых натуральных чисел* $i_1, ..., i_k$ *таких, что* $1 \le i_1 < i_2 < ... < i_k \le n$ *выполняется*

$$\mathsf{P}(A_{i_1} \cap A_{i_2} \cap ... \cap A_{i_k}) = \mathsf{P}(A_{i_1})\mathsf{P}(A_{i_2}) \cdot ... \cdot \mathsf{P}(A_{i_k}).$$

Конечно, попарная независимость (то есть независимость любых двух взятых событий) следует из независимости в совокупности, достаточно просто взять $k = 2$. Обратное, вообще говоря, неверно, что мы уже показывали на примере с раскрашенными тетраэдрами.

3 Случайные величины и их распределения

3.1 Понятие случайной величины

Как мы уже отмечали ранее, для многих экспериментов нет отличий в способе подсчета вероятностей тех или иных событий, когда как элементарные исходы (и пространство элементарных исходов) существенно отличаются. Так что резонно в «похожих» экспериментах поставить в соответствие элементарным исходам числа, и далее работать только с ними. Пусть дано вероятностное пространство $(\Omega, \Sigma, \mathsf{P})$. Напомним, что в случае конечного пространства элементарных исходов Ω, случайной величиной мы называли произвольную функцию $\xi : \Omega \to \mathbb{R}$. В случае произвольного пространства элементарных исходов понятие случайной величины несколько усложняется. Дело в том, что зачастую нам интересны вероятности «событий», таких, как, например: $A = \{\omega \in \Omega : \xi(\omega) \in (a, b)\}$, или $B = \{\omega \in \Omega : \xi(\omega) > a\}$. Но кто может гарантировать, что A и B являются событиями? Кто гарантирует, что они принадлежат Σ? Если это не так, то мы просто не можем вычислить их вероятность. Отсюда и возникает следующее определение.

Определение 3.1.1 (Понятие случайной величины) *Функция* $\xi : \Omega \to \mathbb{R}$ *называется случайной величиной, если для любого* $a \in \mathbb{R}$ *множество* $\{\omega \in \Omega : \ \xi(\omega) < a\}$ *является событием, то есть принадлежит* Σ.

На самом деле слушатель, уставший на этом моменте от теоретических тонкостей, может немного расслабиться и считать, что любое подмножество Ω является событием. В этом случае можно считать, что любая функция $\xi : \Omega \to \mathbb{R}$ является случайной величиной (на практике это обычно проблем не вызывает). Как мы знаем, последнее утверждение, например, верно, когда Ω конечно или счетно (в этих случаях σ-алгеброй можно считать множество всех подмножеств Ω).

Пример 3.1.1 *Пусть подбрасывается правильный тетраэдр, каждая сторона которого пронумерована цифрой от* 1 *до* 4. *Ясно, что в этом случае*

$\Omega = \{1, 2, 3, 4\}$. *Рассмотрим случайную величину* $\xi(\omega)$*, равную числу выпавших очков. Если* $\Sigma = 2^{\Omega}$*, то любое подмножество* Ω *является событием, а значит* ξ *является случайной величиной*

Пример 3.1.2 (Пример НЕ случайной величины) *Будем рассматривать все тот же пример, но только теперь в качестве* σ*-алгебры рассмотрим множество* $\Sigma = \{\varnothing, \{1, 3\}, \{2, 4\}, \{1, 2, 3, 4\}\}$*. Проверьте, что* Σ *действительно является* σ*-алгеброй. Функция* ξ *из предыдущего примера не будет случайной величиной относительно этой* σ*-алгебры, так как, например, множество*

$$\{\omega \in \Omega : \ \xi(\omega) < 4\} = \{1, 2, 3\}$$

не принадлежит Σ*, а значит событием не является.*

Покажем, что если функция является случайной величиной, то это решает вопрос, поставленный нами выше о том, являются ли те или иные привычные нам множества (отрезки, интервалы, полуинтервалы и проч.) событиями, или нет, то есть можем ли мы вычислять вероятность попадания случайной величины в эти множества.

Лемма 3.1.1 *Пусть* ξ *– случайная величина, тогда для любых чисел* $a, b \in \mathbb{R}$ *справедливо*

1. $\{\omega \in \Omega : \ \xi(\omega) \le a\} \in \Sigma$;
2. $\{\omega \in \Omega : \ \xi(\omega) > a\} \in \Sigma$;
3. $\{\omega \in \Omega : \ \xi(\omega) \ge a\} \in \Sigma$;
4. *Если* I *– отрезок* $[a, b]$*, полуинтервал* $[a, b)$ *или* $(a, b]$*, интервал* (a, b)*, луч* $(a, +\infty)$*,* $(-\infty, a)$*,* $[a, +\infty)$ *или* $(-\infty, a]$*, то* $\{\omega \in \Omega : \ \xi(\omega) \in I\} \in \Sigma$.

Доказательство. 1. Множество $\{\omega \in \Omega : \ \xi(\omega) \le a\}$ может быть представлено, как

$$\{\omega \in \Omega : \ \xi(\omega) \le a\} = \bigcap_{n=1}^{\infty} \left\{\omega \in \Omega : \ \xi(\omega) < a + \frac{1}{n}\right\}.$$

Множества $\left\{\omega \in \Omega : \ \xi(\omega) < a + \frac{1}{n}\right\}$, $n \in \mathbb{N}$ входят в Σ, так как ξ, по условию, является случайной величиной. Так как Σ является σ-алгеброй, то и написанное пересечение сходит в Σ, а значит множество $\{\omega \in \Omega : \ \xi(\omega) \le a\}$ является событием.

2, 3 доказываются аналогично.

4. Достаточно заметить, что, например $\{\omega \in \Omega : \ \xi(\omega) \in [a, b]\} = \{\omega \in \Omega : \ \xi(\omega) \le b\} \cap \{\omega \in \Omega : \ \xi(\omega) \ge a\}$. Так как каждое из множеств является событием, то и их пересечение является событием. Для других типов множеств I доказательство аналогично и опирается на пункты 1-3. $\square$

3.2 Функция распределения случайной величины

В этом пункте возникает новый для нас объект, ранее не обсуждавшийся. Пусть $(\Omega, \Sigma, \mathsf{P})$ – вероятностное пространство и ξ – случайная величина. С каждой случайной величиной мы свяжем так называемую функцию распределения случайной величины.

Определение 3.2.1 *Функцией распределения случайной величины ξ называется функция*

$$F_\xi(x) = \mathsf{P}(\xi < x) = \mathsf{P}\left(\{\omega \in \Omega : \ \xi(\omega) < x\}\right).$$

Заметим, что так как ξ является случайной величиной, то для любого $x \in \mathbb{R}$ множество $\{\omega \in \Omega : \ \xi(\omega) < x\}$ является событием, а значит вероятность этого события определена и определение функции распределения корректно.

Отметим свойства функции распределения, которые интуитивно совершенно понятны.

Лемма 3.2.1 (Основные свойства функции распределения) *Функция распределения $F_\xi(x)$ случайной величины ξ обладает следующими свойствами:*

1. *$F_\xi(x)$ не убывает, то есть*
$$\text{если} \ \ x_1 < x_2, \ \text{то} \ \ F_\xi(x_1) \leq F_\xi(x_2).$$
2. $\lim\limits_{x \to -\infty} F_\xi(x) = 0.$
3. $\lim\limits_{x \to +\infty} F_\xi(x) = 1.$
4. *Функция распределения непрерывна слева, то есть для каждого $x_0 \in \mathbb{R}$ выполняется*
$$\lim_{x \to x_0 - 0} F_\xi(x) = F(x_0).$$

Первое свойство моментально следует из определения, ведь при $x_1 < x_2$ справедливо включение

$$\{\omega \in \Omega : \ \xi(\omega) < x_1\} \subset \{\omega \in \Omega : \ \xi(\omega) < x_2\},$$

то есть первое событие влечет второе (второе множество содержит первое), а значит, в силу монотонности вероятности,

$$F_\xi(x_1) = \mathsf{P}\left(\{\omega \in \Omega : \ \xi(\omega) < x_1\}\right) \leq \mathsf{P}\left(\{\omega \in \Omega : \ \xi(\omega) < x_2\}\right) = F_\xi(x_2).$$

Второе свойство, очень грубо говоря, означает, что мы хотим вычислить вероятность события, состоящего в том, что случайная величина принимает значения меньше, чем $-\infty$. Так как значения случайной величины находятся в промежутке $(-\infty, +\infty)$, то легко понять и поверить в то, что $\mathsf{P}(\xi < -\infty) = 0$, ведь случайная величина не может принимать значения меньшие, чем $-\infty$. Из аналогичных рассуждений следует и третье свойство, в котором, опять же грубо, мы хотим вычислить $\mathsf{P}(\xi < +\infty)$, что выполнено всегда, а значит $\mathsf{P}(\xi < +\infty) = 1$.

Строго свойства 2-4 обосновываются с использованием так называемого свойства непрерывности вероятностной меры. Само свойство, как и доказательства обозначенных пунктов, выходят за рамки нашего курса.

Замечание 3.2.1 *Иногда функцией распределения случайной величины ξ называют функцию*

$$F_\xi(x) = \mathsf{P}(\xi \le x) = \mathsf{P}\left(\{\omega \in \Omega : \ \xi(\omega) \le x\}\right).$$

Обратите внимание, здесь строгое неравенство становится нестрогим. Такая функция распределения обладает свойствами 1-3, а свойство 4 заменяется на свойство непрерывности справа. В нашем курсе лекций мы будем придерживаться ранее данного определения.

Часто остается без внимания важная обратная теорема, которая утверждает, что свойства 1-4 являются характеристическими свойствами функции распределения, а именно.

Теорема 3.2.1 *Пусть функция $F(x)$ удовлетворяет четырем пунктам предыдущей теоремы, тогда существует вероятностное пространство $(\Omega, \Sigma, \mathsf{P})$ и случайная величина ξ, определенная на нем, что $F(x) = F_\xi(x)$, то есть функция F является функцией распределения некоторой случайной величины.*

Зная функцию распределения случайной величины ξ, мы можем находить вероятности попадания случайной величины во многие подмножества множества $\mathbb{R}$ такие, как: точки, отрезки, интервалы и все-все, что обычно приходит в голову. Очень часто для этого достаточно пользоваться следующими правилами (обозначения $F_\xi(x_0 \pm 0)$ стандартно отвечают пределу справа/слева).

Теорема 3.2.2 (Дополнительные свойства функции распределения) *Пусть $F_\xi(x)$ – функция распределения случайной величины ξ, тогда:*

1. $\mathsf{P}(\xi = x_0) = F_\xi(x_0 + 0) - F_\xi(x_0)$.

2. $\mathsf{P}(a \le \xi < b) = F_\xi(b) - F_\xi(a)$.

Первое свойство снова обосновывается с помощью аксиомы непрерывности (хотя на интуитивном уровне оно и понятно), а второе следует из того, что

$$\{a \le \xi < b\} = \{\xi < b\} \setminus \{\xi < a\},$$

и

$$\mathsf{P}\left(a \le \xi < b\right) = \mathsf{P}\left(\{\xi < b\}\right) - \mathsf{P}\left(\{\xi < a\}\right) = F_\xi(b) - F_\xi(a).$$

Используя эти свойства, можно вычислять вероятности попадания случайной величины в различные подмножества вещественной прямой. Например,

$$\mathsf{P}(a \le \xi \le b) = \mathsf{P}(a \le \xi < b) + \mathsf{P}(\xi = b) =$$

$$F_\xi(b) - F_\xi(a) + F_\xi(b+0) - F_\xi(b) = F_\xi(b+0) - F_\xi(a).$$

Пример 3.2.1 *Случайная величина ξ может принимать значения с ненулевой вероятностью только на отрезке $[0,2]$, причем на этом отрезке функция распределения имеет вид ax^2. Найти возможные значения константы a. Написать выражение для функции распределения $F_\xi(x)$ при условии, что в точке $x = 2$ функция непрерывна. Вычислить вероятности $\mathsf{P}(\xi \ge 1)$ и $\mathsf{P}(\xi \in [0.5, 1])$.*

Опираясь на определение и свойства функции распределения можно утверждать, что

$$F_\xi(x) = \begin{cases} 0, & x \le 0 \\ ax^2, & x \in (0,2] \\ 1, & x > 2 \end{cases}.$$

Действительно, так как все значения случайной величины принадлежат отрезку $[0,2]$, то $F_\xi(x) = \mathsf{P}(\xi < x) = 0$ при $x \le 0$. Кроме того, $F_\xi(x) = \mathsf{P}(\xi < x) = 1$ при $x > 2$. Знаки неравенств в кусочном задании выбраны из условий непрерывности слева.

Далее, функция распределения не убывает, а значит $a \ge 0$. Кроме того, из тех же соображений, $4a \le 1$, а значит $a \in \left[0, \frac{1}{4}\right]$.

Функция распределения будет непрерывна в точке 2, если $a = \frac{1}{4}$. В этом случае

$$F_\xi(x) = \begin{cases} 0, & x \le 0 \\ \frac{x^2}{4}, & x \in (0,2] \\ 1, & x > 2 \end{cases}$$

Вычислим требуемые вероятности.

$$\mathsf{P}(\xi \ge 1) = 1 - \mathsf{P}(\xi < 1) = 1 - F_\xi(1) = 1 - \frac{1}{4} = \frac{3}{4}.$$

Далее, по теореме сложения вероятностей,

$$\mathsf{P}(\xi \in [0.5, 1]) = \mathsf{P}(\xi \in [0.5, 1)) + \mathsf{P}(\xi = 1) =$$

$$= F_\xi(1) - F_\xi(0.5) + F_\xi(1+0) - F_\xi(1) = F_\xi(1+0) - F_\xi(0.5) = \frac{1}{8}.$$