Основы теории вероятностей

Содержание

Данный курс посвящен основам теории вероятностей. Теория вероятностей применяется во многих случаях, в том числе и при решении комбинаторных задач. Сначала будет дано классическое определение вероятности, а позже — аксиоматика Колмогорова, которая была предложена как решение проблемы основания теории вероятностей.

1. Классическое определение вероятности

1.1. Историческое отступление: зарождение теории вероятностей

На заре становления теории вероятностей было предложено так называемое классическое определение вероятности, в основе которого лежала проблема нахождения шанса выиграть и правильной стратегии в азартных играх. Позже выяснилось, что понятие вероятности выиграть никим образом помочь не может. Вернее, какие-то алгоритмы всё-таки были созданы, но становились бесполезными почти сразу: казино вскоре узнавало про них и корректировало правила игры.

В настоящее время проблема нахождения выигрышной стратегии в азартных играх не представляет практической ценности. Однако, чтобы правильно понимать историю возникновения теории вероятностей, начинать её изучение стоит именно с данной задачи.

Одним из классических вероятностных объектов является обычная игральная кость. Более простой объект — монетка, которая может упасть на стол «орлом» или «решкой» — слишком просто устроен и слабо подходит для первого знакомства. Поэтому в первую очередь с вероятностной точки зрения будет изучена именно игральная кость.

1.2. Классическое определение вероятностей

Пример. Игральная кость Игральная кость представляет собой это кубик, грани которого пронумерованы от одного до шести, обычно круглыми точками. Кость бросают на стол и фиксируют число на той грани, которая оказывается сверху в результате броска. Можно предположить, что форма кубика идеальная и на ребро он упасть не может. В таком случае всего имеют место 6 различных исходов такого броска, причем для них выполняются следующие свойства:

1. Исходы образуют полную группу событий, то есть хотя бы один из них имеет место в результате броска
2. Исходы попарно несовместны, то есть никакие два одновременно произойти не могут
3. Все исходы равновероятны

В результате рассмотрения игральной кости появилось классическое определение вероятностей.

Исходы, обладающие перечисленными свойствами образовывать полную группу событий, быть попарно несовместными и равновероятными, называются **элементарными исходами**.

Классическое определение вероятностей. Пусть различные исходы $\omega_1, \dots, \omega_n$ некоторого эксперимента являются элементарными. Тогда по определению вероятностью произвольного каждого такого исхода называют величину

$$P(\omega_i) = \frac{1}{n}.$$

Пример. Игральные карты. Пусть дана колода из 36 тщательно перетасованных карт. В этом случае элементарными исходами являются различные последовательности карт внутри колоды, то есть различные варианты их взаимного расположения. Эти исходы, которых всего 36!, образуют полную группу событий и попарно несовместны. Вероятность каждого исхода

$$P(\omega_i) = \frac{1}{36!}.$$

Часто бывает необходимым рассматривать события несколько более сложные, чем элементарные исходы. Например, событию «в результате броска игральной кости выпало четное число очков» соответствуют три благоприятных элементарных исхода $\omega_2, \omega_4, \omega_6$.

Вероятность события. Вероятность некоторого события A полагают равной отношению количества элементарных исходов k, которые благоприятствуют событию, к количеству всех возможных элементарных исходов n:

$$P(A) = \frac{k}{n}.$$

1.3. Основные свойства вероятности

Если принять классическое определение как определение вероятности, то несложно получить следующие свойства.

Под событием удобно понимать множество таких элементарных исходов, которые ему благоприятствуют. Множество же всех элементарных исходов $\Omega = \{\omega_1, \dots, \omega_n\}$ называется пространством элементарных исходов, где n — количество всевозможных элементарных событий.

Свойство 1. Вероятность $P(\Omega)$ события Ω, которое состоит в том, что реализуется любой элементарный исход, равняется 1. Такое событие называется достоверным.

Свойство 2. В классическом определении вероятностей нулевую вероятность имеет событие, которое соответствует пустому множеству элементарных исходов, и только оно: $P(A) = 0 \Leftrightarrow A = \emptyset$.

Свойство 3. Вероятность дизъюнктного объединения двух событий A и B, то есть таких событий, что A и B имеют пустое пересечение, равна сумме их вероятностей $P(A \sqcup B) = P(A) + P(B)$.

Свойство 4. Вероятность объединения двух произвольных событий A и B равна $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Свойство 5. Более того, верна формула включений-исключений, которая была доказана в курсе комбинаторики:

$$\begin{aligned} P(A_1 U \dots U A_k) = & P(A_1) + \dots + P(A_n) - P(A_1 \cap A_2) - P(A_1 \cap A_3) - \dots - \\ & - P(A_{k-1} \cap A_k) + \dots + (-1)^{k-1} P(A_1 \cap \dots \cap A_k). \end{aligned}$$

Естественным следствием является выражение для элементарной оценки сверху вероятности суммы нескольких событий.

Свойство 6. $P(A, \cup \dots \cup A_k) \leq P(A_1) + \dots + P(A_k)$.

Отрицанием события A называется событие $\bar{A}$, которому благоприятствуют только такие исходы, которые не благоприятствуют событию A. Тогда верно следующее.

Свойство 7. Вероятность события $\bar{A} := \Omega \backslash A$ равна $P(\bar{A}) = 1 - P(A)$.

2. Задача о раскрасках

Формулировка задачи

Пусть зафиксированы некоторые 15 подмножеств M_i множества $1, 2, \dots, 30$, причем мощность каждого такого подмножества, то есть количество элементов в нем, равняется 5:

$$M_1, \dots, M_{15} \subset 1, 2, \dots, 30, \ |M_i| = 5.$$

Верно ли, что, каким бы образом не были зафиксированы множества M_i, всегда найдется раскраска в 2 цвета элементов множества $1, 2, \dots, 30$, при которой каждое множество M_i включает как элементы одного цвета, так и элементы другого?

Априори ответ абсолютно не очевиден. Решение можно проводить с помощью принципа Дирихле, но правильнее в данный момент решить эту задачу с помощью вероятностного подхода, чтобы продемонстрировать возможность применения языка теории вероятностей для решения сложных задач.

Решение этой задачи придумал великий венгерский математик Пол Эрдёш в 1961 г., создатель современной школы комбинаторики и теории графов. Он показал, что требуемую раскраску можно найти всегда.

Решение

Пусть дана некоторая случайная раскраска χ множества $\{1, 2, \dots, 30\}$. Количество всех возможных раскрасок равняется 2^{30}, а значит пространство элементарных исходов Ω в данном случае равняется множеством всех раскрасок

$$\Omega = \{\chi_1, \dots, \chi_{2^{30}}\}.$$

Можно проверить, что выбранные таким образом элементарные исходы обладают требуемыми свойствами: образуют полную группу событий, несовместны и равновероятны. Тогда вероятность возникновения каждой раскраски: $P(\chi_i) = \frac{1}{2^{30}}$.

Сопоставим каждому множеству M_i событие A_i, которое состоит в том, что в случайной раскраске множество M_i одноцветно. Вероятность $P(A_i)$ события A_i можно найти по определению как отношение количества раскрасок, в которых множество M_i одноцветно, к общему количеству раскрасок $n = 2^{30}$:

$$P(A_i) = \frac{2 \cdot 2^{25}}{2^{30}} = \frac{1}{16}.$$

Вероятность события, состоящего в том, что в случайной раскраске среди множеств $M_1, \ldots, M_{15}$ найдется одноцветное, равняется

$$P(\cup_{i=1}^{15} A_i) \leqslant \sum_{i=1}^{15} P(A_i) < 1,$$

где была использована оценка сверху для вероятности объединения произвольных событий. Тогда вероятность отрицания этого события, то есть события, состоящего в том, что одноцветных множеств среди M_i нет, строго положительна $P(\frac{M_i}{\cup_{i=1}^{15} A_i}) > 0$. Другими словами, существует раскраска, в которой все множества M_i неодноцветны. Утверждение доказано.

3. Условная вероятность и независимость событий

3.1. Понятие об условной вероятности

Пример. Пусть игральную кость бросают один раз. Событие A состоит в том, что выпало 1 очко, а B — в том, что выпало четное число очков. Требуется найти вероятность $P(A|B)$ события A, если уже известно, что событие B произошло. Как изменится ответ, если в качестве события A взять событие, состоящее в том, что выпало 2 очка.

Ответ на первый вопрос кажется естественным: $P(A|B) = 0$, поскольку 1 является нечетным числом.

Второй вопрос задачи уже не такой тривиальный. Если выпало четное число очков, то пространство элементарных исходов сузилось до множества B. Множество элементарных исходов состоит из 3-х элементов: выпало 2 очка, выпало 4 очка и выпало 6 очков. Причем только один из них благоприятный и состоит в том, что кость выпала стороной 2 кверху. Таким образом, искомая вероятность равняется $P(A|B) = \frac{1}{3}$.

В общем случае можно доказать следующую теорему.

Теорема умножения. Пусть в пространстве элементарных исходов $\Omega = \omega_1, \ldots, \omega$ задано некоторое непустое событие $B = \omega_{i_1}, \ldots, \omega_{i_k}$, $k \geqslant 1$. Пусть также задано другое событие A. Тогда вероятность $P(A|B)$ события A при условии, что событие B уже реализовалось можно определить следующим образом:

$$P(A|B)P(B) = P(A \cap B).$$

Док-во. Если событие B уже реализовалось, то B становится новым множеством элементарных исходов, а множеством благоприятных исходов станет множество $A \cap B$:

$$P(A|B) = \frac{|A \cap B|}{|B|} = \frac{|A \cap B| \frac{1}{|\Omega|}}{|B| \frac{1}{|\Omega|}} = \frac{P(A \cap B)}{P(B)}.$$

Здесь несколько раз было использовано определение вероятности. Из последнего уравнения следует требуемое тождество. Теорема доказана.

Замечание. Если множество $P(B)$ пустое, то по соглашению условную вероятность $P(A|B)$ полагают равной нулю.

3.2. Понятие независимости событий

Пусть в некотором пространстве элементарных исходов Ω заданы два события A и B. Условие того, что событие A не зависит от события B, имеет следующий вид:

$$P(A|B) = P(A).$$

Грубо говоря, знание о том, что произошло непустое событие B не даёт никакой новой информации о том, произошло ли событие A.

Более удобную формулировку условия независимости событий A и B можно получить с помощью теоремы умножения:

$$P(A \cap B) = P(A|B)P(B) = P(A)P(B).$$

То есть события A и B независимы тогда и только тогда, когда $P(A \cap B) = P(A)P(B)$.

Говорят, что $A_1, \dots, A_k$ попарно независимы, если независимы друг относительно друга любые два события из этого множества.

Однако существует более сильное свойство независимости событий — независимость в совокупности. События $A_1, \dots, A_k$ независимы в совокупности, если

$$\forall l \leqslant k \quad \forall i_1, \dots, i_l \quad P(A_{i_1} \cap \dots \cap A_{i_l}) = P(A_{i_1}) \dots P(A_{i_l}).$$

Из совокупной взаимной независимости попарная независимость получается тривиальным образом. Однако, если события независимы попарно, то вовсе не обязательно, что они являются независимыми в совокупности.

Задача. Привести пример таких событий A_1, A_2, A_3, которые попарно независимы, но не являются независимыми в совокупности.

3.3. Формула полной вероятности

Теорема. Пусть некоторое пространство элементарных событий Ω может быть представлено как объединение конечного числа непересекающихся подмножеств $B_1, B_2, \dots, B_k$. Тогда для вероятности произвольного события $A \subseteq \Omega$ верна формула полной вероятности:

$$P(A) = \sum_{i=1}^{k} P(A|B_i)P(B_i).$$

Док-во. Исходя из свойств вероятности:

$$P(A) = P(A \cap B_1) + (A \cap B_2) + \dots + (A \cap B_k).$$

Требуемое утверждение может быть получено с помощью теоремы умножения $(A \cap B_i) = P(A \cap B_i)P(B_i)$:

$$\begin{aligned} P(A) = P(A \cap B_1) + (A \cap B_2) + \dots + (A \cap B_k) = \\ = P(A \cap B_1)P(B_1) + P(A \cap B_2)P(B_2) + \dots + P(A \cap B_k)P(B_k). \end{aligned}$$

Утверждение доказано.

Отступление. Раньше при голосовании использовались шары: если в определенного урну положили черный шар, то это значит голос «ПРОТИВ», а белый — «ЗА» данного кандидата.

Пример. Пусть даны две урны: в первой из них находятся 3 черных и 2 белых шара, а во второй — 1 черный и 3 белых. Из первой урны наугад достают два шара и кладут их внутрь второй урны. После этого шары во второй урне тщательно перемешиваются и из нее извлекается 1 шар. Найти вероятность $P(\bigcirc)$, что этот шар — белый.

В результате первого вытягивания реализуется одно из трех событий B_1, B_2, B_3. Событие $B_1 = \bigcirc\bigcirc$ состоит в том, что выпали два белых шара, $B_2 = \bullet\bullet$ — в том, что выпали два черных, а $B_3 = \bigcirc\bullet$ — в том, что выпал один белый и один черный шар.

Согласно классическому определению вероятности:

$$P(B_1) = \frac{1}{10}, \quad P(B_2) = \frac{3}{10}, \quad P(B_3) = \frac{6}{10}.$$

Если из первой урны были выбраны 2 белых шара, то во второй урне после перемешивания окажутся 5 белых и 1 черный шар. В этом случае вероятность вытянуть шар при условии, что произошло событие B_1 равно $P(A|B_1) = \frac{5}{6}$. Аналогично можно получить вероятности $P(A|B_2) = \frac{3}{6}$ и $P(A|B_3) = \frac{4}{6}$.

Согласно формуле полной вероятности, искомая вероятность равна

$$P(A) = \frac{1}{10}\frac{5}{6} + \frac{3}{10}\frac{3}{6} + \frac{6}{10}\frac{4}{6}.$$

3.4. Формула Байеса

Другая важная на практике формула, связанная с формулой полной вероятности, это формула Байеса. Пусть события $B_1, \dots, B_k$ попарно не пересекаются и в объединении дают все множество элементарных исходов. Тогда вероятность события A по формуле полной вероятности

$$P(A) = \sum_{i=1}^{k} P(A|B_i)P(B_i).$$

Часто бывает необходимо найти вероятность $(B_i|A)$ события B_i при условии, что событие A случилось.

Теорема. Вероятность $(B_i|A)$ события B_i при условии, что событие A случилось:

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{i=1}^{k} P(A|B_i)P(B_i)}.$$

Док-во. Если применить теорему умножения двумя различными способами:

$$P(A \cap B_i) = P(A|B_i)P(B_i) = P(B_i|A)P(A),$$

то с учетом формулы полной вероятности можно получить требуемое выражение:

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{P(A)} = \frac{P(A|B_i)P(B_i)}{\sum_{i=1}^{k} P(A|B_i)P(B_i)}.$$

Утверждение доказано.

Пример. Студент решает задачу с k вариантами ответа. Если студент знает, как решать задачу, он дает правильный ответ. Если нет — может угадать с вероятностью $\frac{1}{k}$.

Какая вероятность $P(\text{Студент знал решение}|\text{Дан верный ответ})$, что студент знает решение задачи, при условии, что был дан правильный ответ.

Пусть $p \in [0, 1]$ — вероятность, что студент знает решение задачи. Событие B_1 соответствует ситуации, в которой студент знает решение задачи, а B_1 — не знает. Тогда:

$$P(B_1) = p, \quad P(B_2) = 1 - p, \qquad P(A|B_1) = 1, \qquad P(A|B_2) = \frac{1}{k}.$$

Теперь можно применить формулу Байеса:

$$P(B_1|A) = \frac{p}{p + \frac{1}{k}(1 - p)}.$$

4. Схема испытаний Бернулли

4.1. Броски монеты со смещенным центром тяжести

Пусть вероятность того, что монета со смещенным центром тяжести при её броске на стол ложится кверху «решкой», равна $p =\in [0,1]$, а «орлом» — соответственно $q = 1 - p$. Монета является идеальной в том смысле, что ни при каких условиях монета не падает на ребро.

Пусть также задано некоторое число $n \in \mathbb{N}$. Монету бросают n раз и каждый раз фиксируют, на какую сторону монета упала.

Успехом считается, если монетка падает кверху «решкой». В этом случае записывают единицу. А если монетка упала «орлом», пишут ноль, который обозначает неудачу. После n бросаний получается последовательность ω из 0 и 1:

$$\omega = (x_1, \dots, x_n), x_i \in \{0, 1\}.$$

Множество всех возможных последовательностей является пространством всех элементарных исходов:

$$\Omega : \{\omega(x_1, \dots, x_n), x_i \in \{0, 1\}\}, \qquad |\Omega| = 2^n.$$

В отличие от классической схемы, различные исходы — получившиеся последовательности ω — это уже не равновероятные события. Тогда необходимо определить вероятность каждого такого элементарного исхода, то есть найти, с какой вероятностью $P(\omega = (x_1, \dots, x_n))$ возникает конкретная последовательность ω из нулей и единиц.

Так как различные броски монеты независимы, в выражении для вероятности появления конкретной последовательности их вероятности перемножаются. Поскольку единица в последовательности встречается $\sum\limits_{i=1}^{n} x_i$ раз, ноль — $n - \sum\limits_{i=1}^{n} x_i$,получается следующее выражение для искомой вероятности:

$$P(\omega = (x_1, \dots, x_n)) = p^{\sum_{i=1}^{n} x_i} \cdot q^{n - \sum_{i=1}^{n}}.$$

В частном случае при $p = \frac{1}{2}$ формула упрощается:

$$P(\omega = (x_1, \dots, x_n)) = \left(\frac{1}{2}\right)^n = \frac{1}{2^n},$$

а вероятностное пространство сводится к классическому вероятностному пространству с $\Omega = 2^n$ исходами. Такой случай соотвествтует монете с несмещенным центром тяжести.

4.2. Вероятность получить k успехов в серии из n испытаний Бернулли

Пусть событие A состоит в том, что среди n испытаний Бернулли было ровно k успехов. Множество элементарных исходов, которые благоприятствуют реализации события, можно отождествить с самим событием, как и в классическом случае:

$$A = \{\omega_{i_1}, \dots, \omega_{i_m}\} \subseteq \Omega.$$

Таковыми являются последовательности из нулей и единиц, в которых количество единиц равняется k:

$$A\{\omega = (x_1, \dots, x_n) : x_i \in \{0, 1\}, x_1 + \dots, x_n = k\}.$$

Количество таких элементарных исходов m равняется количеству способов выбрать k позиции из n:

$$m = |A| = C_n^k.$$

Таким образом вероятность события A равна сумме по всем элементарным исходам, которые благоприятствуют событию A:

$$P(A) := \sum_{\omega \in A} P(\omega) = p^k \cdot q^{n-k} = C_n^k \cdot p^k q^{n-k}.$$

4.3. Задача о пересечении случайных подмножеств

Пусть дано некоторое множество $\{1, 2, \dots, n\}$ с количеством элементов n. Случайное подмножество A этого множества можно извлечь по схеме Бернулли с n испытаниями:

$$\{1, 2, \dots, n\} \longrightarrow A$$

следующим образом: k-ый элемент множества войдет в подмножество A только в случае, если в результате k-го испытания будет успех. Известно, что в каждом испытании вероятность успеха равна p. Таким образом? вероятность того, что произвольный элемент исходного множества попадет в A равняется $P(i \in A) = p$. Независимо от A из исходного множества по той же схеме извлекают другое подмножество

$$\{1, 2, \dots, n\} \longrightarrow B.$$

Необходимо найти вероятность $P(A \cap B = \emptyset)$ того, что множества A и B имеют пустое пересечение.

Принцип решения таких задач следующий. Множествам A и B можно сопоставить последовательности из нулей и единиц длины n. Две данные последовательности можно расположить одна над другой. Тогда множества A и B будут иметь пустое пересечение, если не будет ситуации, когда под единицей в первой последовательности находится единица из второй. Вероятность того, что в конкретной позиции такая ситуация не наблюдается равна $1 - p^2$, а значит ввиду независимости всех испытаний искомая вероятность равна

$$P(A \cap B = \emptyset) = (1 - p^2)^n.$$

5. Задача о раскрасках (обобщение)

Задача, которая была решена в одном из предыдущих разделов, может быть обобщена следующим образом.

Теорема 1. Пусть даны произвольные множества $M_1, \dots, M_m$, причем мощность каждого из них $|M_i| = n$. Если число $m \leq 2^{n-1}$, то существует такая раскраска элементов множества $M_1, \dots, M_m$ в два цвета, при которой каждое из этих множеств неодноцветно.

Действитиельно, утверждение, доказанное ранее является частным случаем данной теоремы. Сама теорема доказана венгерскими математиками в 1961г (Эрдеш, Хайнал). Эта задача является отличным примером того, как тесно связанные чисто вероятностные и чисто комбинаторные задачи.

Теорема 2. Пусть $n \geq 100$, а $m \leq \frac{1}{2}\sqrt[3]{\frac{n}{\ln n}} 2^{n-1}$. Пусть также $M_1, \dots, M_m$ — любые n-элементные множества. Тогда существует раскраска в два цвета элементов $M_1 \cup \dots \cup M_m$, при которой все множества M_i неодноцветны.

Док-во. Доказательство данной теоремы основано на двухкратном применении схемы Бернулли.

Пусть введено обозначение:

$$x = x(n) = \frac{1}{2}\sqrt[3]{\frac{n}{\ln n}}.$$

Тогда значение m можно выразить следующим образом (достаточно доказать для случая равенства: при меньшем m раскраска тоже будет существовать):

$$m = x \cdot 2^{n-1}.$$

Теперь необходимо зафиксировать произвольные множества

$$M_1, \dots, M_m, \quad |M_i| = n$$

и обозначить элементы объединения всех этих множеств:

$$M_1 \cup \dots \cup M_m = \{1, 2, \dots, v\}.$$

Пусть выбрана случайная раскраска. Этот выбор будет происходить в два этапа.

Этап 1. Данный этап соотвествтует переписанной в терминах испытаний Бернулли классической схеме. То есть используется «симметричная монетка» $p = \frac{1}{2}$. Первый этап раскраски строится следующим образом: i-ый элемент множества $\{1, 2, \dots, v\}$ красится в красный цвет, если в i-ом испытании Бернулли выпадает «решка», а если «орел» — то в синий. Т.е. с вероятностью $p = \frac{1}{2}$ любой элемент независимо от других либо красный, либо синий.

Тогда можно ввести обозначение для множества всех элементов $\{1, 2, \dots, v\}$, которые в результате первого этапа оказались принадлежащими к одноцветным множествам M_i. Если правильно перекрасить элементы этого множества, то может получиться требуемая раскраска.

Этап 2. Пусть используется монета с шансом выпадения «решки» $p = p(n)$, который будет указан позднее из некоторых оптимальных соображений. Каждый элемент множества $\mathcal{D}$ теперь необходимо перекрасить с вероятностью p независимо от остальных элементов: если выпадает «решка», то очередной элемент множества $\mathcal{D}$ необходимо перекрасить, если же «орел» — необходимо оставить его текущий цвет.

Пусть теперь событие $\mathcal{F}$ состоит в том, что после двух этапов остались одноцветные множества M_i. Тогда достаточно показать, что $P(\mathcal{F}) < 1$ при правильном подборе значения $p = p(n)$.

Для фиксированного множества M_i есть три события, в результате которых это множество оказывается одноцветным после двух этапов раскраски.

A_i : множество M_i красное после 1-го этапа и красное после 2-го этапа.

A_i' : множество M_i красное после 1-го этапа и синее после 2-го этапа.

C_i : M_i неодноцветное после 1-го этапа и красное после 2-го этапа.

Тогда можно привести следующую оценку (удвоение было сделано, чтобы учесть симметричные приведенным события):

$$P(\mathcal{F}) \leq 2P(\cup_{i=1}^{m}(A_i \cup A_i \cup C_i)) \leq 2\sum_{i=1}^{m}(P(A_i) + P(A_i') + P(C_i)).$$

Для вероятностей событий A_i и A_i' верно следующее:

$$P(A_i) = \frac{1}{2^n} \cdot (1-p)^n, \qquad P(A_i') = \frac{1}{2^n} \cdot p^n.$$

Вероятность события C_i можно оценить следующим образом. Если выполнено событие C_i, то множество M_i было перекрашено на втором шаге, так как существует по крайней мере одно множество M_j, которое пересекается с M_i и является синим после первой раскраски.

$$C_i \implies \exists j \neq i : M_i \cap M_j \neq \emptyset \text{ и выполнено } B_{i,j},$$

которое заключается в том, что M_i — хорошее в 1-м этапе и красное во 2-ом, а M_j — синее в первом этапе и произвольное во втором. Это позволяет получить искому оценку для вероятности события C_i:

$$P(C_i) \leq P(\bigcup_{\substack{i \neq j: \\ M_i \cap M_j \neq \emptyset}} B_{i,j}) \leq \sum_{\substack{i \neq j: \\ M_i \cap M_j \neq \emptyset}} P(B_{i,j}).$$

Теперь необходимо оценить сверху $P(B_{i,j})$: пусть

$$h := |M_i \cap M_j| \geq 1$$

— мощность пересечения множеств M_i и M_j, которая по построению не меньше 1. Вероятность того, что все элементы из пересечения покрашены в первом этапе в синий цвет равна $\frac{1}{2^h}$, а множитель p^h отвечает тому, что при второй перекраске каждый элемент пересечения поменял свой цвет на противоположенный. Оставшиеся элементы множества M_j имеют вероятность $\left(\frac{1}{2}\right)^{n-h}$ быть синими после первой покраски. Оставшиеся элементы множества M_i могут быть как красными на первом этапе, так и синими. Допустим какой-либо элемент из $x \in M_i \backslash M_j$ был красным на первом этапе, тогда вероятность того, что он — красный на втором этапе не превосходит $1/2$. Если $x \in M_i \backslash M_j$

был синим на первом этапе и стал красным на втором, то заведомо произошла перекраска. Этот случай реализуется с вероятностью $\frac{p}{2}$. В таком случае можно сделаеть требуемую оценку сверху:

$$P(B_{i,j}) \leq \frac{1}{2^h} \cdot p^h \cdot \left(\frac{1}{2}\right)^{n-h} \cdot \left(\frac{1}{2} + \frac{p}{2}\right)^{n-h} = p^h(1+p)^{n-h} \cdot 2^{-h+h-n+h-n} =$$
$$= p^h(1+p)^{n-h} \cdot 2^{h-2n} \leq p(1+p)^{n-1} \cdot 2^{1-2n}.$$

Здесь было использовано свойство, что максимум выражения $p^h(1+p)^{n-h} \cdot 2^{h-2n}$ на множестве $h \leq 1$ достигается при $h = 1$.

Тогда можно вернуться к оценке вероятности события C_i

$$P(C_i) \leq m \cdot p(1+p)^{n-1} 2^{1-2n},$$

а после — к оценке вероятности события $\mathcal{F}$:

$$P(\mathcal{F}) \leq 2\left(\sum_{i=1}^{m}(2^{-n}(1-p)^n + 2^{-n}p^n) + m \cdot p(1+p)^n \cdot 2^{1-2n}\right) =$$
$$= 2^{1-n}x \cdot 2^{n-1}((1-p)^n + p^n) + 2^{2-2n} \cdot x^2 \cdot 2^{2n-2} \cdot p(1+p)^n =$$
$$= x(1-p)^n + xp^n + x^2p(1+p)^n.$$

Остается только подобрать такое p, что $x(1-p)^n + xp^n + x^2p(1+p)^n < 1$. Если положить

$$p = \frac{1}{3} \cdot \frac{\ln\left(\frac{n}{\ln n}\right)}{n},$$

то выполняеся следующая оценка для предыдущего выражения

$$x(1-p)^n + xp^n + x^2p(1+p)^n \leq$$
$$\leq \frac{1}{2}\left(\frac{n}{\ln n}\right)^{\frac{1}{3}} \cdot e^{-\frac{1}{3}\ln\left(\frac{n}{\ln n}\right)} + \frac{1}{2}\left(\frac{n}{\ln n}\right)^{\frac{1}{3}}\left(\frac{\ln n}{n}\right)^{n} + \frac{1}{4}\left(\frac{n}{\ln n}\right)^{\frac{2}{3}} \cdot \frac{1}{3} \cdot \frac{\ln n}{n} \cdot e^{\frac{1}{3}\ln\left(\frac{n}{\ln n}\right)} =$$
$$= \frac{1}{2}\left(\frac{n}{\ln n}\right)^{\frac{1}{3}} \cdot \left(\frac{\ln n}{n}\right)^{\frac{1}{3}} + \frac{1}{2}\left(\frac{n}{\ln n}\right)^{\frac{1}{3}}\left(\frac{\ln n}{n}\right)^{4} + \frac{1}{12}\left(\frac{n}{\ln n}\right)^{\frac{2}{3}} \cdot \frac{\ln n}{n} \cdot \left(\frac{n}{\ln n}\right)^{\frac{1}{3}} <$$
$$< \frac{1}{2} + \frac{1}{10} + \frac{1}{12} < 1.$$

Здесь было использовано условие на n для оценки второго слагаемого. Теорема доказана.

6. Случайная величина и её основные свойства

6.1. Обобщение классической вероятности и схемы испытаний Бернулли

На данный момент уже были рассмотрены две вероятностные схемы: классическая вероятность и схема испытаний Бернулли. Однако современная теория вероятностей намного богаче. Изученные схемы являются весьма ограничительными и необходимо сначала построить их естественное обобщение на случай произвольного конечного вероятностного пространства.

Для начала следует напомнить основные положения классической вероятностной схемы и схемы испытаний Бернулли. В обоих схемах присутствует конечное множество всех элементарных исходов Ω, а событием A называлось любое его подмножество $A \subseteq \Omega$.

В классической вероятностной схеме каждому элементарному исходу из $\Omega = \{\omega_1, \ldots, \omega_n\}$ сопоставляется одинаковая вероятность:

$$P(\omega_i) := \frac{1}{n}.$$

В таком случае для вероятности события выражается как отношение числа благоприятных исходов $|A|$ к их общему количеству:

$$P(A) := \frac{|A|}{n}.$$

В случае схемы испытаний Бернулли множество элементарных исходов состояло из 2^n элементарных событий

$$\Omega = \{\omega_1, \dots, \omega_{2^n}\},$$

где n — число испытаний. Это связано с тем, что элементарными событиями в данном случае являются последовательности из нулей и единиц реализаций этих испытаний. Тогда для вероятности элементарного исхода верно соотношение

$$P(\omega = (x_1, \dots, x_n)) = p^{\sum\limits_{i=1}^{n} x_i} \cdot q^{n - \sum\limits_{i=1}^{n} x_i},$$

где p — вероятность успеха в одном испытании, а $q = 1 - p$ — вероятность неудачи. Выражение для вероятности события имеет более сложный вид, поскольку вероятности элементарных исходров могут отличаться:

$$P(A) := \sum_{\omega \in A} P(\omega).$$

Естественное обобщение как классической вероятности, так и схемы Бернулли можно построить следующим образом. Пусть задано произвольное конечное множество объектов:

$$\Omega = \{\omega_1, \dots, \omega_n\},$$

где каждый объект ω_i является элементарным исходов, а Ω — пространство элементарных исходов. Вероятности элементарных исходов есть некоторые числа p_i:

$$P(\omega_1) = p_1, \quad P(\omega_2) = p_2, \dots, \quad P(\omega_n) = p_n,$$

которые должны удовлетворять следующим естественным требованиям:

$$p_i \in [0, 1], \qquad p_1 + \dots + p_n = 1.$$

Эти требования выражают условия того, что вероятность определена корректно.

Событием A будет называться любое подмножество пространства элементарных исходов $A \subseteq \Omega$. Если элементарный исход принадлежит множеству A, то говорят, что он благоприятствует событию A. Вероятность $P(A)$ события A по определению полагается равной сумме вероятностей всех элементарных исходов, которые благоприятствуют данному событию:

$$P(A) := \sum_{\omega \in A} p_i.$$

В таком случае несложно получить выражения для отрицания $\bar{A}$ события A и пересечения двух событий A и B:

$$P(\bar{A}) = 1 - P(A), \qquad P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

В таком случае остаются справедливыми и многие оценки, верные в случаях классической схемы или схемы Бернулли, например:

$$P(A_1 \cup \dots \cup A_m) \le \sum_{i=1}^{m} P(A_i).$$

Пример. Кость со смещенным центром тяжести будет описываться как раз только что введенной обобщенной конечной вероятностной схемой.

6.2. Понятие случайной величины

Понятие случайной величины позволяет более содержательно изучать свойства различных вероятностных пространств. Пусть дано некоторое вероятностное пространство Ω, P, то есть задано множество элементарных исходов Ω, на котором корректно определена заданы вероятности всех элементарных исходов.

Вещественнозначная функция, определенная на множестве Ω, называеся случайной величиной:

$$\xi : \Omega \longrightarrow R.$$

Важно отметить, что на каждом конкретном элементарном исходе ω случайная функция принимает вполне определенное значение $\xi(\omega)$.

Пример. Пусть $\Omega = 1, 2, 3, 4, 5, 6$ — пространство исходов при броске игральной кости. Пусть $\xi(\omega) = \omega^2$. До бросания кости неизвестно, какое число выпадет, а значит и каково будет значение $\xi(\omega)$. Однако после броска функция $\xi(\omega)$ примет определенное значение, поскольку реализуется какой-то один элементарный исход.

6.3. Понятие случайного графа

Случайный граф является примером красивого вероятностного пространства для демонстрации методов, которые будут рассматриваться позднее. Случайный граф, вообще говоря, является частным случаем схемы испытаний Бернулли.

Множество вершин случайного графа V_n отождествим с множеством натуральных чисел от 1 до n:

$$V_n = \{1, 2, \dots, n\}.$$

В случайном графе, как видно из определения множества V_n, множество и количество вершин фиксированно, а значит случайными буду ребра графа.

В полном неориентированном графе на n вершинах без кратных ребер полное количество ребер e_i равно C_n^2:

$$e_1, e_2, \dots, e_{C_n^2}.$$

Теперь пусть зафиксировано число $p \in [0, 1]$, которое принимается равным вероятности каждому отдельному ребру независимо от остальных присутствовать в случайном графе. Фактически имеет место схема Бернулли, где проводятся C_n^2 испытаний: в случае успеха в каком-то из испытаний ребро, которое соотвествует этому испытанию, будет присутствует в случайном графе, а в случае неудачи — будет осутствовать. Таким образом в серии из C_n^2 испытаний реализуется случайный граф $G = (V_n, E)$, где V_n — множество вершин, а E — множество ребер.

Каждый граф можно отождествить с элементарным исходом. Вероятность того, что реализуется в результате данной схемы какой-то определенный граф равна

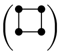
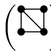
$$P(G = (V_n, E)) = p^{|E|^2} \cdot q^{C_n^2 - |E|},$$

где $|E|$ — число ребер.

Случайные величины на случайном графе

На графах очень удобно задавать случайные величины, поэтому имеет смысл рассмотреть несколько примеров.

Пример. Пусть $\xi(G) :=$ число треугольников. Например:

$\xi\Big($  $\Big) = 0, \qquad \xi\Big($  $\Big) = 2.$

Еще раз стоит отметить, что в зависимости от того, какой элементарный исход реализуется, ξ может принимать то или иное конкретное значение. Случайная величина называется случайной только из-за того, что её аргумент — объект из вероятностного пространства. После того, как элементарный исход реализовался, случайная величина принимает конкретное значение.

Распределение случайной величины

Пусть задана случайная величина $\xi : \Omega \longrightarrow \{y_1, \dots, y_k\}$ на пространстве элементарных событий $\Omega = \{\omega_1, \dots, \omega_n\}$. Это значит, что заданы значения случайной величины для всех элементарных исходов.

В теории вероятностей часто бывает нужно знать не такую функциональную зависимость, а распределение случайной величины. Распределение случайной величины — это центральное понятие в теории вероятностей. Говорят, что распределение известно, если известны все вероятности того, что ξ принимает какое-то возможное значение y_i:

$$P(\xi = y_i) = P(\{\omega_j : \xi(\omega_j) = y_i\}).$$

Распределение может быть как задано по условию задачи, так являться искомомым в задаче.

Пример Пусть ξ на пространстве случайных графов — случайная функция, равная числу треугольников в графе. Если граф был построен на n вершинах, то максимальное число треугольников C_n^3 будет наблюдаться в случае полного графа. Тогда говорят, что известно распределение случайной величины ξ, если известны следующие вероятности:

$$P(\xi = k), \quad k = 0, 1, \ldots, C_n^3.$$

Вычислить точно распределение функции ξ с помощью численного расчета очень сложно, поэтому очень важны асимптотические методы оценки функций распределений случайных величин.

Математическое ожидание случайной величины

Математическое ожидание случайной величины ξ принято обозначать двумя различными способами. Обозначение $E\xi$ восходит к английскому выражению «expected value», а $M\xi$ — «mean value».

Математическое ожидание случайной величины ξ, определенной на конечном вероятностном пространстве, определяется согласно следующему выражению:

$$M\xi = \sum_{\omega \in \Omega} \xi(\omega) \cdot P(\omega).$$

В случае классической вероятностной схемы математическое ожидание соответствует среднему арифметическому последовательности из значений случайной величины. Вообще говоря, получается некое среднее значение, своего рода центр масс как физике, так как $\sum\limits_{\omega \in \Omega} P(\omega) = 1$. Важно, что математическое ожидание может не принадлежать множеству возможных значений случайной величины.

Поскольку несколько элементарных исходов могут диват одно и тоже значение случайной функции, это определение можно переписать, сгруппировова такие элементарные исходы:

$$M\xi = \sum_{\omega \in \Omega} \xi(\omega) \cdot P(\omega) = y_1 \left(\sum_{\omega : \xi(\omega) = y_1} P(\omega) \right) + y_2 \left(\sum_{\omega : \xi(\omega) = y_2} P(\omega) \right) + \ldots + y_k \left(\sum_{\omega : \xi(\omega) = y_k} P(\omega) \right) =$$

$$= y_1 P(\xi = y_1) + y_2 P(\xi = y_2) + \ldots + y_k P(\xi = y_k) = \sum_{j=1}^{k} y_j P(\xi = y_j).$$

Последняя формула является естественным проявление того факта, что математическое ожидание можно вычислить исходя из знаний функции распределения случайной величины.

Замечание. Однако часто бывает наоборот: математическое ожидание считается легче, чем функция распределения, и дает возможность получить некоторые важные свойства этого распределения, не вычисляя его.

Математическое ожидание обладает свойством линейности: пусть $\xi_{1,2}$ — две произвольные случайные величины на конечном вероятностном пространстве, в $_{1,2}$ — произвольные вещественные числа, тогда

$$M(c_1\xi_1 + c_2\xi_2) = c_1 M\xi_1 + c_2 M\xi_2.$$

Это свойство непосредственно следует из определения математического ожидания.

Пример 1. Найти математическое ожидание случайной величины ξ, равной числу треугольников в единичном графе:

$$M\xi = \sum_{k=0}^{C_n^3} k \cdot P(\xi = k).$$

Чтобы решить данную задачу, необходимо обойти вычисление величины $P(\xi = k)$ с помощью свойства линейности математического ожидания. Действительно:

$$\xi(G) = \xi_1(G) + ... + \xi_{c_n^3}(G),$$

где $\xi_i(G)$ — случайная величина, которая равна 1, если треугольник с номером i принадлежит графу G, и равна нулю в иных случаях. Тогда, так как $M\xi_i = P(\xi_i = 1) = p^3$, получаем искомое математическое ожидание:

$$M\xi = M\xi_1 + ... + M\xi_{c_n^3} = C_n^3 p^3.$$

Пример 2. В схеме испытаний Бернулли рассматривается случайная величина, которая равна количеству успехов во всех испытаниях:

$$\mu = \sum_{i=1}^{n} x_i,$$

где ξ_i — случайная величина, которая равна 1, если в i испытании был успех, и равна нулю в случае неудачи. Тогда несложно можно получить выражение для математического ожидания:

$$M\mu = np.$$

Линейность математического ожидания является исключительно полезным свойством, которое, при всей своей тривиальности, позволяет рассчитывать математическое ожидание даже тогда, когда мы не знаем распределение этой случайной величины.

6.4. Неравенство Маркова.

Теорема. Пусть случайная функция ξ принимает только неотрицательные значения. Пусть дано положительное число $a > 0$. Тогда вероятность $P(\xi \geq a)$ того, что случайная функция ξ не меньше a, удовлетворяет неравенству:

$$P(\xi \geq a) \leq \frac{M\xi}{a}.$$

Док-во. Согласно определению математического ожидания:

$$M\xi = \sum_{j=1}^{k} y_j P(\xi = y_j) = \sum_{j: y_j \geq a} y_j P(\xi = y_j) + \sum_{j: y_j < a} y_j P(\xi = y_j).$$

Поскольу значения случайной функции неотрицательны $y_j \geq 0$, второе слагаемое в последнем выражении можно оценить снизу нулем.

$$M\xi \geq \sum_{j: y_j \geq a} y_j P(\xi = y_j) \geq a \sum_{j: y_j \geq a} y_j P(\xi = y_j) = aP(\xi \geq a).$$

Откуда следует требуемое выражение:

$$P(\xi \geq a) \leq \frac{M\xi}{a}.$$

Теорема. Пусть число вершин случайного графа n растет к бесконечности $n \to \infty$, а вероятность $p = p(n)$ такова, что $np(n) \to 0$. Тогда асимптотически в случайном графе нет треугольников: $P(\xi = 0) \to 1$

Док-во. Согласно неравенству Маркова:

$$P(\xi = 0) = 1 - p(\xi \geq 1) \geq 1 - M\xi = 1 - C_n^3 p^3.$$

Так как $np(n) \to 0$, то: $C_n^3 p^3 = \frac{n(n-1)(n-2)}{6} p^3 \sim \frac{n^3 p^3}{6} \to 0$. Откуда следует требуемое утверждение $P(\xi = 0) \to 1$.

6.5. Дисперсия случайной величины

Дисперсия $\mathcal{D}\xi$ случайной величины ξ определяется следующим образом:

$$\mathcal{D}\xi := M(\xi - M\xi)^2.$$

Дисперсия позволяет оценить разброс случайной величины относительно её среднего значения, то есть найти среднеквадратичное отклонение ξ от своего среднего значения.

Замечание. Более простое выражение $M(\xi - M\xi)$ не подходит для оценки разброса распределения вокруг среднего значения, так как:

$$M(\xi - M\xi) = M(\xi) - M(M\xi) = M(\xi) - M(\xi) = 0.$$

Выражение для вычисления дисперсии можно преобразовать с помощью свойства линейности:

$$M(\xi - M\xi)^2 = M(\xi^2 - 2\xi \cdot M\xi + (M\xi)^2) = M\xi^2 - 2M\xi \cdot M\xi + (M\xi)^2 = M\xi^2 - (M\xi)^2.$$

То есть

$$\mathcal{D}\xi := M\xi^2 - (M\xi)^2.$$

Неравенство Чебышёва

Пусть η — любая случайная величина. Пусть $b > 0$, тогда выполняется неравенство

$$P\left(|\eta - M\eta| \geq b\right) \leq \frac{\mathcal{D}\eta}{b^2}.$$

Док-во. Согласно неравенству Маркова при $\xi = (\eta - M\eta)^2 \geq 0$ и $a = b^2$:

$$P(\xi \geq a) \leq \frac{M\xi}{a} \quad \Longrightarrow \quad P\left((\eta - M\eta)^2 \geq b^2\right) \leq \frac{M(\eta - M\eta)^2}{b^2} = \frac{\mathcal{D}\eta}{b^2}.$$

Здесь было использовано определение дисперсии. Таким образом:

$$P(\xi \geq a) \leq \frac{\mathcal{D}\eta}{b^2}.$$

Неравенство доказано.

7. Независимые случайные величины

7.1. Понятие независимости случайных величин

Пусть даны две случайные величины $\xi, \eta : \Omega \to \mathbb{R}$. Причем ξ принимает значения из множества $\{x_1, \dots, x_n\}$, а η — из множества $\{y_1, \dots, y_n\}$:

$$\xi : \Omega \to \{x_1, \dots, x_n\}, \qquad \eta : \Omega \to \{y_1, \dots, y_n\}.$$

Говорят, что случайные величины ξ и η независимы, если:

$$\forall i = 1, \dots, n \quad \forall j = 1, \dots, m \quad P(\xi = x_i, \eta = y_i) = P(\xi = x_i) \cdot P(\eta = y_i).$$

Как и в случае рассмотрения независимости событий, для набора случайных величин $\xi, \dots, \xi_n$ можно говорить о попарной независимости и независимости в совокупности.

Набор случайных величин $\xi, \dots, \xi_n$ называются попарно независимыми в том случае, если любая пара стучайных величин независима.

Набор случайных величин $\xi, \dots, \xi_n$ называется независимым в совокупности тогда и только тогда, когда выполняется

$$\forall k \ \forall_{i_1, \dots, i_k} \ \forall x_{i_1}, \dots, x_{i_k} \quad P(\xi_{i_1} = x_{i_1}, \dots, \xi_{i_k} = x_{i_k}) = P(\xi_{i_1} = x_{i_1}) \cdot \ldots \cdot P(\xi_{i_k} = x_{i_k})$$

Пример. Пусть $A_1, \dots, A_n$ — события, независимые попарно, но зависимые в совокупности. Пример таких событий предлагалось привести в соответствующем упражнении ранее. Тогда случайные величины $\xi_1, \dots, \xi_n$, индикаторы событий $A_1, \dots, A_n$, попарно независимы, но зависимы в совокупности:

$$\xi_i := I_{A_i}(\omega) = \begin{cases} 1, & \text{если } \omega \in A_i \\ 0, & \text{иначе} \end{cases}, \qquad i = 1, \dots, n.$$

7.2. Свойства математического ожидания независимых случайных величин

Можно получить некоторые свойства математического ожидания и дисперсии в случае, если случайные величины независимы. Во-первых, свойство линейности математического ожидания, безусловно, справедливо и в случае независимых случайных величин:

$$M(c_1\xi_1 + ... + c_n\xi_n) = c_1 M\xi_1 + ... + c_n M\xi_n.$$

Теорема. Пусть ξ, η — независимы, тогда:

$$M(\xi \cdot \eta) = (M\xi) \cdot (M\eta)$$

Док-во. Пусть ξ и η принимают на пространстве элементарных исходов множества значений $\{x_1, ..., x_n\}$ и $\{y_1, ..., y_m\}$ соответственно:

$$\xi : \Omega \to \{x_1, ..., x_n\}, \qquad \eta : \Omega \to \{y_1, ..., y_m\}.$$

Математическое ожидания произведения двух случайных величин по определению:

$$M(\xi\eta) = \sum_{\omega\in\Omega} \xi(\omega) \cdot \eta(\omega) \cdot P(\omega).$$

Пространство Ω можно представить в виде объединения по непересекающимся множествам $\Omega_{i,j}$:

$$\Omega = \bigcup_{i=1}^{n} \bigcup_{j=1}^{m} \Omega_{i,j}, \qquad \Omega_{i,j} \subset \Omega : \forall \omega \in \Omega_{i,j} \quad \xi(\omega) = x_i,\ \eta(\omega) = y_i.$$

Тогда исходное выражение для $M(\xi\eta)$ можно преобразовать следующим образом:

$$M(\xi\eta) = \sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{\omega\in\Omega_{i,j}} \xi(\omega) \cdot \eta(\omega) \cdot P(\omega).$$

Так как при $\omega \in \Omega_i, j$ выполняется по построению $\xi(\omega) = x_i$ и $\eta(\omega) = y_i$:

$$\begin{aligned}
M(\xi\eta) &= \sum_{i=1}^{n} \sum_{j=1}^{m} \sum_{\omega\in\Omega_{i,j}} x_i \cdot y_i \cdot P(\omega) = \sum_{i=1}^{n} \sum_{j=1}^{m} x_i \cdot y_j \cdot \sum_{\omega\in\Omega_{i,j}} P(\omega) = \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} x_i \cdot y_j \cdot P(\Omega_{i,j}) = \sum_{i=1}^{n} \sum_{j=1}^{m} x_i \cdot y_j \cdot P(\xi = x_i, \eta = y_j) = \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} x_i \cdot x_j \cdot P(\xi = x_i) \cdot P(\eta = y_i) = \\
&= \left(\sum_{i=0}^{n} x_i \cdot P(\xi = x_i) \right) \cdot \left(\sum_{j=1}^{m} y_i \cdot P(\eta = y_i) \right) = (M\xi) \cdot (M\eta).
\end{aligned}$$

Здесь было использовано определение независимости двух случайных величин. Теорема доказана.

7.3. Свойства дисперсии независимых случайных величин

Полученная свойство математического ожидания позволяет доказать следующее свойство дисперсии для независимых случайных величин. Дисперсия, в отличие от математического ожидания, не является линейной величиной. В частности:

$$\mathcal{D}(c\xi) = c^2 \cdot \mathcal{D}\xi.$$

Теорема. Пусть $\xi_1, ..., \xi_n$ — попарно независимые случайные величины. Тогда

$$\mathcal{D} \cdot (\xi_1 + ... + \xi_n) = \mathcal{D} \cdot \xi_1 + \dots + \mathcal{D} \cdot \xi_n.$$

Док-во. Можно ввести следующие обозначения:

$$\eta_1 = \xi_1 - M\xi_1, \quad ... \quad \eta_n = \xi_n - M\xi_n.$$

Попарная независимость $\eta_1, ..., \eta_n$ есть следствие попарной независимости величин $\xi_1, ..., \xi_n$. Тогда по определению для дисперсии получается:

$$\begin{aligned}\mathcal{D}(\xi_1 + ... + \xi_n) &= M(\xi_1 + ... + \xi_n - M(\xi_1 + ... + \xi_n))^2 = \\ &= M((\xi_1 - M\xi_1) + (xi_2 - M\xi_2) + ... + (\xi_n - M\xi_n)) = \\ &= M\left(\eta^2 + ... + \eta_n^2 + \sum_{i\neq j} \eta_i\eta_j\right) = M\eta_1^2 + ... + M\eta_n^2 + \sum_{i\neq j} M(\eta_i\eta_j) = \\ &= M\eta_1^2 + ... + M\eta_n^2 + \sum_{i\neq j}(M\eta_i)(M\eta_j) = \\ &= M\eta_1^2 + ... + M\eta_n^2 = M(\xi_1 - M\xi_1)^2 + ... + M(\xi_n - M\xi_n)^2 = \mathcal{D}\xi_1 + ... + \mathcal{D}\xi_n.\end{aligned}$$

Здесь было использовано соотношение $M\eta_i = M(\xi_i - M\xi_i) = 0$. Теорема доказана.

Замечание. Можно привести примеры таких зависимых случайных величин ξ, η и, тем не менее:

$$M(\xi\eta) = M(\xi) \cdot (M\eta).$$

Случайные величины ξ и η называются некоррелированными, если $M(\xi\eta) = M(\xi) \cdot (M\eta)$.

Замечание. При доказательстве предыдущей теоремы было использовано только соотношение $M(\xi\eta) = M(\xi) \cdot (M\eta)$. Таким образом теорема остается справедливой даже при более слабых условиях.

Упражнение. Пусть $\xi_1, ..., \xi_n$ — попарно некоррелированные случайные величины. Доказать, что

$$\mathcal{D}(\xi_1 + ... + \xi_n) = \mathcal{D}\xi_1 + \cdots + \mathcal{D}\xi_n.$$

8. Закон больших чисел

8.1. Закон больших чисел в форме Чебышёва

Одним из центральных результатов теории вероятностей является закон больших чисел (ЗБЧ). Вообще говоря, существует множество формулировок ЗБЧ. В следующей формулировке он называется законом больших чисел в форме Чебышёва.

Теорема. Пусть $\xi_1, ..., xi_n, ...$ — некоторая бесконечная последовательность попарно некоррелированных величин. Пусть также их дисперсии не превосходят некоторой константы c:

$$\exists c: \ \forall i \ \mathcal{D}\xi_i \leq c.$$

Тогда $\forall \varepsilon > 0$

$$P\left(\left|\frac{\xi_1 + ... + \xi_n}{n} - \frac{M\xi_1 + ... + M\xi_n}{n}\right| > \varepsilon\right) \overset{n\to\infty}{\to} 0.$$

Замечание. Строго говоря, задать бесконечную последовательность попарно некореллированных величин нельзя на конечном вероятностном пространстве. Однако позже в данном курсе будут введены примеры бесконечных вероятностных пространств и эта теорема станет верной уже без такой формальной неаккуратности. С другой стороны, бесконечное количество случайных величин на практике никогда не бывает известно. Тогда, грубо говоря, следует воспринимать закон больших чисел следующим образом: если рассмотреть некоторое ε и достаточно большое конечное n, рассматриваемая вероятность будет достаточно малой.

Замечание. В частности, если $\forall M\xi_i = a$, утверждение теоремы будет следующим:

$$P\left(\left|\frac{\xi_1 + ... + \xi_n}{n} - a\right| > \varepsilon\right) \overset{n\to\infty}{\to} 0.$$

Док-во. В выражении

$$P\left(\left|\frac{\xi_1+...+\xi_n}{n}-\frac{M\xi_1+...+M\xi_n}{n}\right|>\varepsilon\right)$$

можно ввести обозначение $\eta=\frac{\xi_1+...+\xi_n}{n}$. Тогда можно воспользоваться неравенством Чебышёва:

$$P\left(|\eta-M\eta|>\varepsilon\right)\leq\frac{\mathcal{D}}{\varepsilon^2}=\frac{1}{\varepsilon^2\cdot n^2}\cdot(\mathcal{D}\xi_1+...+\mathcal{D}\xi_n)\leq\frac{c\cdot n}{\varepsilon^2n^2}\overset{n\to\infty}{\to}0.$$

Теорема доказана.

Замечание. Пусть в схеме испытаний Бернулли μ_n — число успехов в n испытаниях, ξ_i — результат i-го испытания:

$$\mu_n=\xi_1+...+\xi_n,\qquad\qquad \xi_i=\begin{cases}1, & \text{при успехе}\\ 0, & \text{при неудаче}\end{cases}.$$

Математическое ожидание и дисперсия случайных величин ξ:

$$M\xi_i=p,\qquad \mathcal{D}\xi_i=M\xi_i^2-(M\xi_i)^2=p\cdot(1-p)=p\cdot q.$$

Тогда для μ_n можно получить ввиду независимости ξ_i:

$$M\mu_n=n\cdot p,\qquad \mathcal{D}\mu_n=\mathcal{D}\xi_1+...+\mathcal{D}\xi_n=npq.$$

8.2. Предельная теорема Пуассона

Вероятность того, что число успехов в схеме Бернулли находится в определенных пределах

$$P(k\leq\mu_n\leq l)=\sum_{i=k}^{l}C_n^i\cdot p^i\cdot q^{n-i}.$$

Такую сумму сложно вычислять при большом количестве испытаний в серии. На практике вычисляются математические аппроксимации требуемых величин.

Теорема. Пусть вероятность успеха в i-м испытании зависит от числа испытаний n. Более того, пусть $p(n)=\frac{\lambda}{n}$, $\lambda>0$. Тогда вероятность $P(\mu_n=k)$ можно аппроксимировать следующим образом:

$$P(\mu_n=k)\to\frac{\lambda^k e^{-\lambda}}{k!},\qquad n\to\infty.$$

Док-во. Для искомой вероятности

$$P(\mu_n=k)=C_n^kp^kq^{n-k}=C_n^k\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}.$$

Для биномиальных коэффициентов, поскольку k зафиксированно, верно следующее асимптотическое равенство:

$$C_n^k=\frac{n!}{k!\cdot(n-k)!}=\frac{n\cdot(n-1)\cdot...\cdot(n-k+1)}{k!}\sim\frac{n^k}{k!},\quad n\to\infty.$$

С учетом этого и того, что

$$\left(1-\frac{\lambda}{n}\right)^{-k}\to1,\quad n\to\infty,$$

можно получить:

$$P(\mu_n=k)=C_n^k\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{nk}\sim\frac{n^k}{k!}\frac{\lambda^k}{n^k}\left(1-\frac{\lambda}{n}\right)^n\sim\frac{\lambda^k}{k!}e^{-\lambda}.$$

Здесь был использован первый замечательный предел $\left(1-\frac{\lambda}{n}\right)^n\sim e^{-\lambda}$. Теорема доказана.

8.3. Интегральная предельная теорема Муавра-Лапласа

Теорема.

$$P(a \leq \frac{\mu_n - n \cdot p}{\sqrt{n \cdot p \cdot q}} \leq b) \Rightarrow \frac{1}{\sqrt{2\pi}} \int\limits_a^b e^{-\frac{x^2}{2}} dx, \qquad n \to \infty.$$

Причем сходимость равномерная по всем $a, b \in \mathbb{R}$.

Замечание. Для случайной величины $\eta = \frac{\mu_n - n \cdot p}{\sqrt{n \cdot p \cdot q}}$ выполняется $M\eta = 0$, $\mathcal{D}\eta = \frac{\mathcal{D}\mu_n}{n \cdot p \cdot q} = 1$. В таком случае принято говорить, что величина μ_n была центрирована и отнормирована.

Док-во этой теоремы выходит за рамки данного курса.

Пример. Театр имеет два входа, около каждого из которых есть гардероб. Гардеробы имеют одинаковую вместимость. Будем считать, что люди выбирают вход взаимно независимо с вероятностью $\frac{1}{2}$. Пусть в театре всего 1000 мест, и каждый день театр посещают 1000 человек. Найти минимальный размер гардероба такой, чтобы вероятность переполнения P(переполн. гардероба) хотя бы одного из них была меньше чем $\frac{1}{300}$.

Пусть x — искомый размер гардероба. Очевидно, что требуемый минимальный размер будет лежать в следующем интервале $500 < x < 1000$. На самом деле правильным ответом будет $x \approx 546$. На самом деле это связано с тем, что практически весь (вернее $\approx 0.997 = 1 - \frac{1}{300}$) интеграл функции $e^{-\frac{x^2}{2}}$ в теореме Муавра-Лапласа набирается в пределах от -3 до 3.

В схеме испытаний Бернулли с $n = 1000$ испытаний и $p = \frac{1}{2}$ условно успехом можно положить то, что человек пошел в какой-то конкретный гардероб:

$$npq = 250, \quad \sqrt{npq} = 5\sqrt{10}.$$

Гардероб не переполняется тогда и только тогда, когда $\mu_n \leq x$ и $1000 - \mu_n \leq x$. Вероятность этого события можно преобразовать к виду, в котором сформулирована теорема Муавра-Лапласа:

$$P(1000 - x \leq \mu_n \leq x) = P\left(\frac{500 - x}{5\sqrt{10}} \leq \frac{\mu_n - np}{\sqrt{npq}} \leq \frac{x - 500}{5 \cdot \sqrt{10}}\right) = \frac{1}{300}.$$

Таким образом,

$$\frac{x - 500}{5\sqrt{10}} = 3 \quad \Longrightarrow \quad x = 500 + 15\sqrt{10} \approx 546.$$

9. Геометрическая вероятность

Ранее разбирались только случаи конечных вероятностных пространств. Но вообще говоря, возможны ситуации, когда элементарных исходов может быть бесконечно много. Примером может служить геометрическая вероятность, классическая конструкция, появившаяся несколько позднее классической вероятностной схемы с конечным числом равновероятных исходов.

9.1. Задача о встрече

Пример. Коля и Вася договорились встретиться на автобусной остановке в промежутке времени от 9:00 до 10:00. Оба друга выбирает время случайно в рамках данного промежутка. Прийдя на остановку, каждый из них ждет на остановке не более 15 минут, после чего уходит. Найти вероятность P(встречи) того, что встреча состоится.

Замечание. На самом деле число возможностей выбрать время прихода каждого из друзей бесконечно. Задачу, конечно, можно решить приближенно с помощью классической схемы вероятности, разбив заданный промежуток времени на небольшие фрагменты, например с точностью до секунд. Однако в таком случае приходится перебирать миллионы вариантов для решения задачи.

Действительно, если каждый из друзей может прийти в любую из 3600 секунд, то вероятностное пространство будет иметь вид:

$$\Omega = \{(,) : \ = 0, 1, \ldots, 3600; \ B = 0, 1, \ldots, 3600\}.$$

Мощность этого пространства: $|\Omega| = 3601^2$. Друзья встречаются в том случае, если величина $|K - B|$ превосходит 15 минут, т.е. 900 секунд:

$$P(\text{встречи}) = P(|K - B| \leq 900).$$

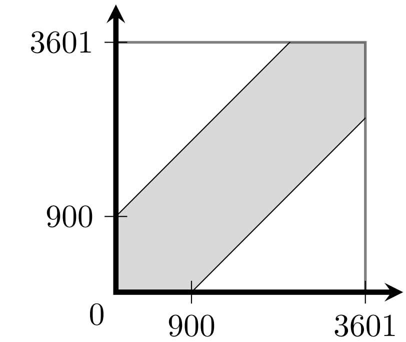


Рис. 1: Графическая иллюстрация для решения задачи в случае классической вероятностной схемы: искомая вероятность равна отношению количества точек с целыми коэффициентами в закрашенной области к общему количеству точек.

Однако такой способ решения содержит даже элемент произвола — на каком основании было выбрано разбиение именно на секунды, а не на какие-нибудь другие отрезки времени. При увеличении мелкости дробления ситуация все больше стремится к непрерывному случаю.

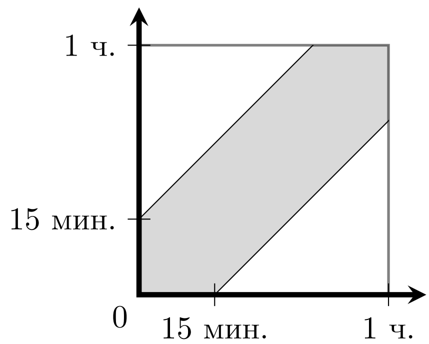


Рис. 2: Графическая иллюстрация для решения задачи: искомая вероятность равняется закрашенной площади.

Площадь фигуры считается значительно проще количества рассматриваемых ранее точек и равняется искомой вероятности встречи:

$$P(\text{встречи}) = 1 - 2\frac{1}{2}\frac{3}{4}\frac{3}{4} = \frac{7}{16}.$$

Здесь искомая площадь была вычислена как разница площади площади квадрата и площади двух треугольников, которые дополняют требуемую область до квадрата.

9.2. Геометрическое определение вероятности

В прошлой задаче вероятность была представлена как площадь некоторой области. Более общее определение соответствующего вероятностного пространства является следующим.

Пусть на плоскости дано некоторое множество $\Omega \subset \mathbb{R}^2$. Внутри этой области выбирают случайно некоторую точку. Пусть также выделено некоторое подмножество $A \subseteq \Omega$. Тогда вероятность того, что точка попадает в данную область A, равняется $P(\text{попасть в } A) = \frac{\mu(A)}{\mu(\Omega)}$, где $\mu(\Omega)$ и $\mu(A)$ — площади множеств Ω и A соответственно. Это соотношение является обобщение формулы вероятности события для классической вероятностной схемы.

Однако такое определение оставляет много неразрешенных вопросов. Во-первых, нет слов о том, что считать площадью и как именно считать площадь. На самом деле площадь не любого подмножества можно вычислить. Эта проблема будет обсуждена немного позже. Также неправильное восприятие геометрической вероятности может приводит к различного рода кажущимся парадоксам. Одним из таковых является парадокс Бертрана.

9.3. Парадокс Бертрана

Задач. Пусть дана единичная окружность. Найти вероятность того, что случайная хорда имеет длину больше, чем сторона правильного вписанного треугольника.

Парадокс состоит в том, что способ определения случайной хорды не задан. Оказывается, что в зависимости от конкретного определения «случайной хорды» меняется ответ в данной задаче. Однако в 19 веке было распространено мнение, что ответ не должен меняться, как бы не была определена эта случайность, что приводило к такого рода парадоксам.

Случайность хорды можно определить следующим образом. Пусть на окружности S единичного радиуса $R = 1$ случайно выбирается одна точка. Множеством элементарных исходов в таком случае становится сама окружность. Вероятность того, что точка принадлежит некоторой дуге $A \subseteq S = \Omega$ этой окружности определяется следующим образом:

$$P(A) := \frac{\mu(A)}{\mu(S)} = \frac{\mu(A)}{2\pi},$$

то есть «площадью» в пространстве S на самом деле является длина. На окружности потом выбираются случайно две точки и в качестве случайной хорды выбирается отрезок xy. Используя формулу полной вероятности, можно получить без ограничения общности, что конкретная позиция точки x не важна и может быть зафиксирована. Вероятность того, что точка y будет таковой, что длина хорды будет больше длины стороны вписанного правильного треугольника, равна вероятности того, что дуга xy будет больше 60°:

$$P(\text{длина хорды} > \text{длина стороны }) = \frac{\mu(A)}{2\pi} = \frac{1}{3}.$$

Другой способ определить случайную хорду заключается в следующем. Пусть случайным образом внутри круга выбрана точка x. Только для одной хорды, проходящей через нее, данная точка будет центральной. Именно эту хорду и можно выбрать в качестве случайной хорды. Тогда вероятность того, что длина хорды будет больше длины стороны вписанного правильного треугольника равна вероятности точке x попасть в круг A радиуса $1/2$.

$$P(\text{длина хорды} > \text{длина стороны }) = \frac{\mu(A)}{\pi} = \frac{1}{4}.$$

В этой задаче явно прослеживается факт, что распределение вероятности меняется в зависимости от конкретного способа эту вероятность задать. А для случайной хорды единственного естественного способа это сделать не существует.

Задача о минимальном количестве пустых треугольников

Пусть дано конечное множество точек на плоскости $X \subset \mathbb{R}^2$, которое состоит из $|X| = n$ элементов. Пусть также никакие 3 точки не лежат на одной прямой. Треугольник $\triangle$, образованный тремя точками этого множества, называется пустым в том случае, если внутрь треугольника никаких точек исходного множества не попадает $\triangle \cap X = \{\text{вершины } \triangle\}$.

Тривиальной является задача поиска максимального количества пустых треугольников в множестве из n точек на плоскости: если в качестве X выбраны вершины выпуклого n-угольника, то все C_n^3 возможные треугольники будут пустыми. В свою очередь задача отыскания минимального количества пустых треугольников $f(n)$ в множестве из n точек на плоскости является одной из классических задач современной комбинаторной геометрии, которая на данный момент полностью не решена.

Верхние оценки для $f(n)$ можно доказывать следующим образом:

$$(f(n) \leq a) \Leftrightarrow (\exists X \in \mathbb{R}^2, |X| = n \text{и в } X \text{ не больше, чем } a \text{пустых } \triangle).$$

Требуемое множество X можно, конечно, строить явно, но такой метод не слишком эффективен. Гораздо лучше работает вероятностный метод. Необходимо, таким образом, выбрать случайное подмножество из n точек на плоскости и показать, что с положительной вероятностью будет не больше a пустых треугольников.

Случайное множество может быть определено неоднозначно (это было хорошо показано на примере случайной хорды). Например, можно взять квадрат определенного размера, внутри которого выбраны n случайных точек. Также можно взять круг и выбирать случайные точки внутри него. Если требуемая вероятность в хотя бы одном случае окажется положительной, то задача будет решена. Поэтому выбор правильного способа определения вероятностного пространства очень важен при решении данной задачи. То есть эти точки нужно стараться выбирать таким образом, чтобы с наибольшей вероятностью пустых треугольников оказалось как можно меньше.

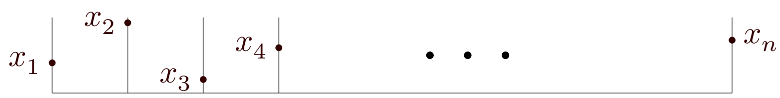


Рис. 3: Каждая случайная точка расположена равновероятно на своем «зубчике».

Подробное решение данной задачи не имеет смысла приводить в рамках данного курса. Стоит лишь привести пример конструкции, с помощью которой были получены наилучшие известные оценки сверху для $f(n)$. Набор $X\{x_1, \dots, x_n\}$ случайных точек располагается на «зубчиках расчески». Тогда можно получить для $\epsilon > 0$:

$$P(\text{в X не больше } (2+\varepsilon)n^2 \text{ пустых } \triangle) > 0.$$

Другими словами $f(n) \leq (2+\varepsilon)n^2$. Также существует другая теорема, которая гласит, что:

$$f(n) \geq (1-\varepsilon)n^2.$$

Это есть пример задачи, когда правильный выбор случайности позволяет с точностью до множителя определить требуемое значение при использовании вероятностного подхода.

10. Несколько слов об аксиоматике Колмогорова

10.1. Переход к аксиоматике Колмогорова

Замечание. В классической вероятностной схеме вероятность события равняется нулю тогда и только тогда, когда это событие есть пустое множество:

$$P(A) = 0 \Leftrightarrow A = \emptyset.$$

В геометрическом определении вероятности это свойство работает только в одну сторону:

$$A = \emptyset \implies P(A) = 0.$$

Например, вероятность попасть в конкретную точку равна нулю. Более того, равна нулю и вероятность попадания на одномерную кривую, например на окружность.

Более того, оказывается, что не у каждого множества можно определить площадь. Существуют различные способы определить площадь, или как это принято называть, меру: мера Жордана и мера Лебега. Эти вопросы также выходят за рамки данного курса.

Однако на данном этапе важно сделать существенный шаг — ввести аксиоматику Колмогорова. В 1933 году тогда еще молодой математик, Колмогоров, взял на себя смелость предложить фундамент теории вероятностей так, чтобы в этот фундамент укладывались и классическая вероятностная схема, и геометрическая вероятность, и схема испытаний Бернулли.

Тот факт, что не у каждого подмножества можно определить площадь, приводит к тому, что не каждое подмножество соотвествует событию с корректно определенной вероятностью. Таким образом, необходимо выделить такие подмножества множества Ω, для которых понятие вероятности можно будет определеить корректным образом.

10.2. Аксиоматическое определение вероятности

Пусть абстрактное множество Ω выбрано как пространство элементарных событий. Пусть также выделена система подмножеств $\mathcal{F}$ такая, что:

1. $\forall F_1, F_2 \in \mathcal{F} \Longrightarrow F_1 \cup F_2 \in \mathcal{F}$
2. $\forall F_1, F_2 \in \mathcal{F} \Longrightarrow F_1 \cap F_2 \in \mathcal{F}$
3. $\forall F \in \mathcal{F} \Longrightarrow \bar{F} := \Omega \setminus F \in \mathcal{F}$
4. $\Omega \in \mathcal{F}, \emptyset \in \mathcal{F}$
5. $\forall F_i \in \mathcal{F}, i = 1, 2, \dots, \bigcup\limits_{i=1}^{\infty} F_i, \bigcap\limits_{i=1}^{\infty} \in \mathcal{F}$

Такая система подмножеств $\mathcal{F}$ называется сигмо-алгеброй событий.

Замечание. Простейшим примером сигма-алгебры событий является тривиальная сигма алгебра

$$\mathcal{F} = \{\emptyset, \Omega\}.$$

Замечание. Для конечного пространства элементарных исходов $\Omega\{\omega_1, \dots, \omega_n\}$ в качестве сигма-алгебры $\mathcal{F} = 2^{\Omega}$ может быть взято множество всех подмножеств множества Ω.

Вероятностной мерой называется такая функция, которая каждому элементу сигма-алгебры сопоставляет некоторое число:

$$P : \mathcal{F} \longrightarrow [0, 1]$$

так, чтобы выполнялись следующие свойства:

1. $P(\bar{F}) = 1 - P(F)$
2. $P(\Omega) = 1$
3. $P(F_1 \bigsqcup F_2) = P(F_1) + P(F_2)$
4. $P(\bigsqcup\limits_{i=1}^{\infty} F_i) = \sum\limits_{i=1}^{\infty} P(F_i)$

Дальнейшее изложение теории вероятности должно быть основано на введенной аксиоматике: можно будет рассматривать различные вероятностные пространства, можно будет доказывать продвинутые законы больших чисел, центральную предельную теорему и так далее. Но это предмет более продвинутого курса теории вероятностей.