Занятие 1.

Основные понятия теории вероятностей.

Концепция вероятностного подхода. Первые примеры: числа Рамсея.

Вопросы о существовании объектов с некоторыми предписанными свойствами являются центральными во многих областях математики. Часто оказывается, что предъявить объект явно трудно, но можно обнаружить, что он является в той или иной степени *типичным.*

Приведем несколько примеров прямо сейчас (в дальнейшем различным, подчас исключительно хитроумным, примерам такого рода будет посвящен весь курс).

1. Рассмотрим случайное число из отрезка [0,1]. "Случайность"здесь означает, что вероятность этого числа попасть в какой-то меньший отрезок равна длине этого отрезка. (Для знакомых с понятием меры Лебега: вероятность числа попасть в некоторое измеримое множество равна по определению мере этого множества). В его двоичной записи цифры 0 и 1 встречаются примерно одинаково часто (то есть доля единиц среди первых N цифр стремится к половине, когда N неограниченно возрастает). То же верно для троичной записи: там поровну нулей, единиц и двоек. Очень сложно (кажется, никто не умеет) предъявлять *явно* хотя бы одно иррациональное число, удовлетворяющего обоим этим требованиям — хотя им удовлетворяет *почти любое* число.

2. Известно, что не всякая непрерывная функция является кусочно дифференцируемой. Первые примеры таких “экзотических” функций были построены Вейерштрассом и описаны во многих учебниках анализа. Однако большинство непрерывных функций, возникающих не искусственно, все же являются кусочно гладкими. Исключение в этом неформальном правиле составляют *случайные* функции — траектории броуновского движения. Случайно движущаяся частица ни в один момент не имеет скорости. Отметим также, что свойство непрерывной функции не иметь ни в одной точки производной является “типичным”: функции, имеющие производную хотя бы в одной точке, образуют “тощее” множество в смысле категории Бэра (то есть представимо как объединение счетного семейства нигде не плотных множеств).

3. Приведем пример из комбинаторики, каковых еще нас ожидает немало. Используемые понятия теории вероятностей интуитивно понятны, определения см. чуть ниже.

Пусть n команд сыграли однокруговой турнир по волейболу. Может ли оказаться, что для любых ста команд найдется команда, выигравшая у всех ста? Оказывается, что может, и проще всего провести такой турнир, определяя победителя каждой встречи, подбрасывая монетку.

Рассмотрим случайный турнир. Какова вероятность, что для некоторых ста команд одна из оставшихся выиграла у них всех? Ответ: 2^{-100}. А какова вероятность, что не у всех? Ответ: $p = 1 - 2^{-100}$. Какова же, наконец, вероятность, что ни одна из оставшихся команд не выиграла у всех этих ста? Ответ: p^{n-100}. Как бы оценить вероятность того, что указанное событие произошло хотя бы для каких-то ста команд? Просто сложим все эти числа по C_n^{100} способам выбрать сто команд — результат будет верхней оценкой на искомую вероятность. Итак, вероятность того, что для каких-то ста команд ни одна из оставшихся не выиграла у них всех, не превосходит $C_n^{100}p^{n-100} \leq n^{100}p^{n-100}$. Так как степенная функция растет медленнее показательной, при больших n это совсем маленькое число — то есть почти все случайные турниры удовлетворяют нашему требованию.

Введем необходимые для дальнейшего понятия из теории вероятностей.

Вероятностное пространство

Пусть Ω — некоторое (обычно конечное) множество, которое мы назовем *вероятностным пространством.* Пусть некоторым (в случае конечного пространства Ω — просто всем) подмножествам $A \subset \Omega$ сопоставляется число $p(A) \in [0, 1]$, называемое *вероятностью* события A (множество $A \subset \Omega$ будем называть событием). так, что

(i) семейство событий, то есть подмножеств пространства Ω, на которых определена вероятность, образует σ-алгебру, то есть содержит $\emptyset$, Ω, вместе с любыми событиями содержит их дополнения и объединение не более чем в счетном количестве.

(ii) $p(\emptyset) = 0$, $p(\Omega = 1)$.

(iii) $p(A \sqcup B) = p(A) + p(B)$ (здесь и далее $A \sqcup B$ означает дизъюнктное объединение событий A и B, то есть объединение непересекающихся множеств A и B. В теории вероятностей говорят не "непересекающиеся множества", а "несовместные события".) Аналогичное равенство верно для счетного дизъюнктного объединения.

Иными словами, вероятность есть мера на пространстве событий такая, что мера всего пространства равна 1. Такие меры называют *вероятностными.*

Приведем несколько примеров:

1. Если Ω — конечное множество, то можно определить вероятность подмножества $A \subset \Omega$ как $|A|/|\Omega|$ ($|\cdot|$ обозначает количество элементов в множестве). Говорят, что элементы Ω равновероятны.

2. Рассмотрим эксперимент с подбрасыванием монеты, которая падет орлом с вероятностью p и решкой с вероятностью $1 - p$. В качестве пространства Ω возьмем двухэлементное множество $\{$орел, решка$\}$ и опре-

делим вероятности указанным образом.

3. Если Ω — пространство с конечной мерой μ, то можно определить вероятностную меру $\mu/\mu(\Omega)$. Так, можно говорить о случайном выборе точки в отрезке — вероятность попасть в меньший отрезок пропорциональна его длине.

4. Если Ω_1, Ω_2 — два вероятностных пространства, можно определить их произведение $\Omega_1 \times \Omega_2$. Как множество это будет просто декартово произведение, а вероятность определим так: для $A \subset \Omega_1$, $B \subset \Omega_2$ положим $p(A \times B) = p(A) \cdot p(B)$. В силу свойств вероятности это равенство однозначно задает вероятности всех событий, которые есть счетные объединения событий типа $A \times B$ (в конечном случае — вообще всех событий). Так, если монетка подбрасывается дважды, то пространство исходов есть произведение пространств исходов для однократного подбрасывания. Часто рассматриваются произведения не двух, а большего конечного числа событий.

Теорию вероятностей отличает от теории меры наличие понятия **независимости**.

События $A, B \subset \Omega$ назовем *независимыми*, если $p(A \cup B) = p(A) \cdot p(B)$. События $A_1, A_2, \ldots, A_k$ называются взаимно независимыми, если для любого множества индексов $I \subset \{1, 2, \ldots, k\}$ выполняется равенство

$$p(\cup_{i \in I} A_i) = \prod_{i \in I} p(A_i).$$

Упражнение: докажите, что если некоторые из набора взаимно независимых событий заменить на их дополнения, то набор событий останется взаимно независимым.

Отметим, что из попарной независимости не следует взаимной.

Типичный пример независимых событий — события типа $A \times \Omega_2$ и $\Omega_1 \times B$ в пространстве $\Omega_1 \times \Omega_2$. Если рассматривается произведение некоторого количества вероятностных пространств, то события, выражающиеся через непересекающиеся семейства этих пространств, заведомо независимы. Например, если монетка подбрасывается десять раз, то событий “первые пять раз выпадет орел” “в седьмой раз выпадет решка” независимы. В основном независимые события, которые нам будут встречаться, имеют именно такую природу.

Случайная величина

Функция (обычно вещественнозначная), заданная на вероятностном пространстве, называется *случайной величиной.* Вообще говоря, требуется измеримость функции относительно сигма-алгебры событий. В конечном случае об измеримости можно не думать.

Примеры случайных величин: количество выпавших орлов при стократном подбрасывании монетки, сумма чисел на гранях трех подброшенных кубиков, площадь треугольника ABC, где точки A, B, C случайно и независимо выбираются в единичном круге.

Случайные величины можно складывать, умножать, вычитать и делить.

Важными характеристиками случайной величины являются ее **математическое ожидание** и **дисперсия**.

Если X — случайная величина, то ее математическим ожиданием $E(X)$ называется число $\int X d\mu$, где μ — вероятностная мера на том вероятностном пространстве, на котором задана случайная величина X (в случае, если интеграл сходится). В случае конечного вероятностного пространства математическое ожидание есть конечная сумма $\sum p_i x_i$, где p_i — вероятность того, что значение случайной величины равно x_i.

Полезным свойство математического ожидания является его линейность: $E(\alpha X + Y) = \alpha E(X) + E(Y)$, где α — число, X и Y — случайные величины.

Простым, но часто используемым свойством математического ожидания неотрицательной случайной величины X является неравенство Чебышева

$$p(X \geq a) \leq E(x)/a, a > 0.$$

Определим *дисперсию* случайной величины X как

$$D(X) = E\left((X - E(x))^2\right) = E(X^2) - (E(X))^2.$$

Для дисперсии неравенство Чебышева приобретает вид

$$P(|X - E(x)| \geq a) \leq D(x)/a^2.$$

Если выбрать случайного человека, то вероятность того, что выбранный — мужчина, чуть меньше половины. Однако если дополнительно известно, что выбранного человека зовут Сережа, то эта вероятность сильно повышается.

Введем понятие **условной вероятности**.

Если B — некоторое событие такое, что $p(B) > 0$, то условной вероятностью события A относительно события B называется число

$$p(A|B) = p(A \cap B)/p(B).$$

Если X — некоторая случайная величина, то ее условным математическим ожиданием относительно события B является $\frac{1}{p(B)} \int_B X d\mu$, где μ

— вероятностная мера (то есть среднее значение на множестве B). В конечном случае условное математическое ожидание есть

$$\sum p(X = x_i|B)x_i,$$

где x_i — возможные значения случайной величины X.

Если вероятностное пространство представлено в виде объединения попарно несовместных событий: $\Omega = \sqcup A_i$, то имеет место *формула полной вероятности* события B:

$$p(B) = \sum p(B|A_i) \cdot p(A_i).$$

Для математического ожидания случайной величины X имеем

$$E(X) = \sum E(X|A_i) \cdot p(A_i).$$

Более подробное знакомство с основами теории вероятностей предлагается получить из какого-нибудь стандартного учебника.

Мы перейдем к применению вероятностных (пока что совсем элементарных) соображений в комбинаторике.

Один из классических примеров — **нижние оценки для чисел Рамсея**.

Напомним, что *числом Рамсея* $R(n, k)$ называется наименьшее из натуральных чисел N таких, что при любой покраске ребер N-вершинного графа в красный и синий цвета найдется либо полный n-вершинный подграф, все ребрак которого окрашены в красный цвет (далее красный $n-$вершинник), либо синий k-вершинник. А приори не известно, существует ли вообще такое N.

Для полноты картины докажем верхнюю оценку $R(n, k) \leq C_{n+k-2}^{n-1}$ (в частности, $R(n, k)$ существует). Индукция по n и k. База $n = 2$ или $k = 2$ очевидна. Пусть n, $k \geq 3$. Положим $N = C_{n+k-2}^{n-1}$, $N_1 = C_{n+k-3}^{n-1}$, $N_2 = C_{n+k-3}^{n-2}$. Заметим, что $N = N_1 + N_2$. Выберем в графе с N вершинами некоторую вершину v. Из нее выходит либо N_1 синих ребер, либо N_2 красных. В первом случае из N_1 концов этих синих ребер можно выбрать по индукционному предположению либо красный n-вершинник, либо синий $(k-1)$-вершинник, который при добавлении вершины v становится синим k-вершинником. Второй случай аналогичен, индукционный переход завершен.

Отсюда, в частности,

$$R(n, n) \leq C_{2n-2}^{n-1}.$$

Последняя величина растет с точностью до множителей степенного роста как 4^n (поскольку это наибольший из биномиальных коэффициентов C_{2n-2}^k, сумма которых равна 4^{n-1}.)

Получим экспоненциальную нижнюю оценку на $R(n,n)$.

Будем случайно и независимо красить ребра графа с N (N выберем потом) вершинами в красный и синий цвета с равной вероятностью обоих цветов (для оценки $R(n,k)$ при $n > k$ следует чаще красить ребра в красный цвет, чем в синий). Оценим вероятность того, что найдется полный одноцветный n-вершинник. Если эта вероятность окажется меньше 1, то $N < R(n,n)$. Для каждого n-вершинника вероятность оказаться одноцветным равна $2^{1-C_n^2}$. Всего n-вершинников имеется C_N^n. Так как вероятность объединения нескольких событий не больше суммы их вероятностей, получаем на искомую вероятность p наличия одноцветного n-вершинника оценку

$$p \le C_N^n 2^{1-C_n^2} < (N^n/n!) \cdot 2^{1-C_n^2}.$$

Последнее выражение меньше 1, если

$$N < (n!)^{1/n} 2^{(n-1)/2-1/n}.$$

Поскольку $(n!)^{1/n}$ ведет себя примерно как n/e (это обстоятельство тут не ключевое, в главном рост определяется вторым множителем), получаем, что $R(n,n) \ge cn2^{n/2}$, где c — некоторая константа.

Для каких λ верна оценка $R(n,n) \ge \lambda^{n+o(n)}$? Мы выяснили, что годится
$lambda = \sqrt{2}$ и не годятся $\lambda > 4$. Устранение зазора между корнем из двух и четырьмя является одной из наиболее вызывающих задач комбинаторики.

Следующий пример основан на линейности математического ожидания.

Турниром будем называть полный ориентированный граф (с помощью таких графов изображаются однокруговые турниры без ничьих).

Напомним, что *гамильтоновым путем* в ориентированном графе называется ориентированный путь, проходящий через все вершины по разу. Как много может быть гамильтоновых путей в турнире с n вершинами? Рассмотрим случайный турнир (в котором ребра ориентируются независимо с равной вероятностью в ту или иную сторону). Вычислим математическое ожидание количества гамильтоновых путей в таком случайном турнире. Заметим, что для каждой перестановки $(x_1, x_2, \dots, x_n)$ вершин турнира путь $x_1 - x_2 - \dots - x_n$ является гамильтоновым с вероятностью 2^{1-n}. Таким образом, математическое ожидание случайной величины “гамильтоновость данной перестановки”, равной 1, если соответствующий путь гамильтонов и 0 в противном случае, есть 2^{1-n}. Количество гамильтоновых путей есть сумма указанных случайных величин по всем перестановкам. Значит, математическое ожидание количества гамильтоновых путей есть $2^{1-n} \cdot n!$. Таким образом, найдется турнир, содержащий хотя бы $2^{1-n} \cdot n!$ гамильтоновых путей.

Этот результат Селе (1943) многократно упоминается как первое использование вероятностного метода в комбинаторике.

Следующий пример интересен тем, что в нем как нижние, так и верхние оценки получаются вероятностным методом.

Пара $H = (V, E)$, где V — конечное множество *вершин*, а $E \subset 2^V$ — некоторое семейство подмножеств множества V, называемых *гиперребрами*, называется *гиперграфом.* Гиперграф называется *n-однородным*, если каждое его ребро содержит ровно n вершин. Говорят, что гиперграф является 2-раскрашиваемым, если множество его вершин можно покрасить в два цвета так, чтобы не было одноцветных ребер.

При каком наименьшем $m(n)$ существует n-однородный гиперграф, не являющийся 2-раскрашиваемым?

Следующее утверждение показывает, что $m(n) \geq 2^{n-1}$:

Теорема. Любой n-однородный гиперграф с $k < 2^{n-1}$ гиперребрами является 2-раскрашиваемым.

Доказательство. Рассмотрим случайную раскраску в два цвета (вершины красятся независимо с вероятностью $1/2$ в красный и $1/2$ в синий). Вероятность того, что некоторое гиперребро является одноцветным, очевидно равна 2^{1-n}. Поскольку вероятность объединения событий не превосходит суммы их вероятностей, получаем, что вероятность наличия одноцветного гиперребра не превосходит $k \cdot 2^{1-n} < 1$.

Оценка на $m(n)$ также получается вероятностными методами. Только теперь следует случайно выбирать гиперребра и оценивать вероятность события, состоящего в том, что при некоторой раскраске найдется одноцветное ребро.

Зафиксируем множество V, содержащее v вершин (значение v будет выбрано позднее). Пусть в некоторой раскраске a вершин имеют синий цвет и $v-a$ — красный. Тогда вероятность того, что случайно выбранное среди всех n-элементных подмножеств V гиперребро при этой раскраске окажется одноцветным, равна

$$\frac{C_a^n + C_{v-a}^n}{C_v^n}.$$

Это выражение как функция от a сначала убывает, а потом возрастает и достигает минимума при $a = [v/2]$. Давайте считать v четным) Тогда получим, что вероятность случайного гиперребра оказаться одноцветным для данной раскраски не меньше

$$p_0 := \frac{2C^n_{v/2}}{C^n_v}.$$

Выберем случайно и независимо m гиперребер. Для каждой фиксированной раскраски вероятность того, что ни одно из них не одноцветно, не больше чем $(1-p_0)^m$. Количество раскрасок равно 2^v, поэтому вероятность того, что для какой-то раскраски ни одно из m случайно выбранных гиперребер не одноцветно, не больше чем

$$2^v(1-p_0)^m.$$

Если эта величина меньше 1, то с положительной вероятностью m гиперребер таковы, что для всякой раскраски одно из них одноцветно — то есть соответствующий гиперграф не 2-раскрашиваемый.

Применим оценку $1-p < e^{-p}$ и получим, что достаточно выбрать v и m так, что $2^v e^{-p_0 m} < 1$, то есть

$$0 > v\log 2 - mp_0 = v\log 2 - 2m\frac{C^n_{v/2}}{C^n_v},$$

таким образом, на интересующую нас величину m получаем оценку

$$m \geq \frac{\log 2}{2}\cdot v \cdot \frac{C^n_v}{C^n_{v/2}}.$$

Если выбрать v порядка n^2, то полученная оценка на m будет порядка $n^2 2^n$ (соответствующие вычисления предлагается провести самостоятельно).

Итак,

$$m(n) \leq Cn^2 2^n.$$