А. Н. ШИРЯЕВ

ВЕРОЯТНОСТЬ
И
КОНЦЕПЦИЯ СЛУЧАЙНОСТИ:

к 75–летию выхода в свет монографии
А. Н. Колмогорова
“Основные понятия теории вероятностей”

26 ноября 2009 г.

§ 1. Введение

§ 2. К истории становления теории вероятностей

§ 3. Частотная теория вероятностей Мизеса

§ 4. Частотоустойчивость (или стохастичность)

§ 5. Типичность (принадлежность к множеству эффективной меры единица)

§ 6. Сложноустроенность (или хаотичность)

§ 7. Непредсказуемость

§ 8. Пример

§ 1. ВВЕДЕНИЕ

1.1. Тема настоящего доклада концентрируется вокруг понятия

СЛУЧАЙНОСТЬ

и, более определенно, вокруг того,

как можно было бы формально определить, что есть

ИНДИВИДУАЛЬНАЯ СЛУЧАЙНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ.

В статье "On logical foundation of Probability Theory" (Fourth USSR–Japan Symposium, 1982) А. Н. Колмогоров писал:

> В повседневной речи мы называем С Л У Ч А Й Н Ы М И те явления, где мы не находим Р Е Г У Л Я Р Н О С Т И, которая позволила бы нам точно предсказать их результаты. Вообще говоря, нет оснований считать, что случайное явление должно обладать какой-либо определенной вероятностью. Следовательно, мы должны были бы различать
>
> С О Б С Т В Е Н Н О случайность
>
> (как отсутствие регулярности)
>
> и
>
> С Т О Х А С Т И Ч Е С К У Ю случайность,
>
> (которая является объектом теории вероятностей).

Возьмем, к примеру, *конечную* бернуллиевскую последовательность

$$\boxed{(\mathrm{I}_{10})}: \qquad \mathbf{0\,1\,1\,1\,0\,1\,0\,0\,1\,0}$$

или *бесконечную* бернуллиевскую последовательность

$$\boxed{(\mathrm{I}_{00})}: \qquad \mathbf{0\,1\,1\,1\,0\,1\,0\,0\,1\,0\,1\,1}\ldots = \boxed{(\mathrm{I}_{10})}\,\mathbf{1\,1}\ldots,$$

образуемые в результате "честного" подбрасывания правильной монеты ($\mathbf{1}$ = орел, $\mathbf{0}$ = решка).

Получив эти последовательности, мы будем склонны называть их

"СЛУЧАЙНЫМИ",

поскольку **не видно закономерности** в порядке следования нулей и единиц.

Получив, однако, последовательность

$\boxed{(\mathrm{II}_{10})}$: **1 1 1 1 1 1 1 1 1 1**, состоящую сплошь из десяти единиц, или

$\boxed{(\mathrm{III}_{10})}$: **1 0 1 0 1 0 1 0 1 0**, состоящую из чередующихся единиц и нулей,

мы вряд ли, следуя нашей интуиции, признаем их “случайными”.

Но с вероятностной точки зрения каждая последовательность

$$(\mathrm{I}_{10}), \quad (\mathrm{II}_{10}) \quad \text{и} \quad (\mathrm{III}_{10})$$

имеет $\boxed{\text{одну и ту же вероятность } (1/2)^{10}}$. *

---

* Рассматриваемая симметричная схема Бернулли предполагает, что результаты подбрасывания на каждом шаге независимы с вероятностями появления единицы и нуля, равными $1/2$.

Наша цель — изложить компактно основные подходы, которые сложились к настоящему времени в понимании того,

какие последовательности было бы естественным называть“случайными”.

В идеале было бы, конечно, хорошо “расщепить” множество всех бинарных последовательностей на два множества — случайных и неслучайных последовательностей.

ЗАМЕЧАНИЕ. Для всего дальнейшего важно подчеркнуть, что

> мы рассматриваем не произвольные последовательности, а только те, которые получены в результате вероятностных экспериментов, описываемых симметричной схемой Бернулли; тем самым, мы исходим из предположения, что мы находимся в рамках аксиоматики Колмогорова, предполагающей заданным измеримое пространство исходов $(\Omega, \mathcal{F})$ с некоторой вероятностной мерой P.

Априори естественно было бы думать, что

вопрос “расщепления” последовательностей
на “случайные” и “неслучайные”

(при том или ином определении, соответствующем нашему неформальному пониманию)

можно решить в рамках
самой же **теории вероятностей**

Однако приведенные примеры последовательностей (I_{10}), (II_{10}) и (III_{10}), имеющих одну и ту же вероятность появления, показывают, что это вряд ли возможно.

Нельзя сказать, что теория вероятностей вовсе неспособна давать ответы на эти вопросы “расщепления”.

В действительности дело обстоит так, что колмогоровская аксиоматика теории вероятностей устроена таким образом, что она позволяет говорить о выполнении тех или иных свойств **лишь P-почти наверное**. Можно сказать, что

> теория вероятностей дает ответ на вопрос
> выполнения тех или иных свойств
> для “подавляющего большинства” объектов
> (бинарных последовательностей, например),
> без детализации того, какой индивидуальный объект
> принадлежит или не принадлежит этому “большинству”.

Оказалось, что разрешению вопроса об отнесении индивидуальных объектов к “случайным” или “неслучайным” помогает обращение к (казалось бы, далекой от теории вероятностей) **теории алгоритмов**.

ВАЖНО: вряд ли в принципе можно провести четкую грань между “случайностью” и “неслучайностью” для *конечных* последовательностей.

Для случая же *бесконечных* последовательностей, как будет видно из дальнейшего, теория алгоритмов приводит к вполне удовлетворяющим интуиции определениям понятия

бесконечная случайная последовательность.

Хорошо известно, что в теории вероятностей есть много результатов

для бесконечных последовательностей

о справедливости разных свойств для “подавляющего большинства”, формулируемых в терминах выполнения свойств

почти наверное.

Например, рассмотрим бернуллиевское вероятностное пространство:

$$(\Omega, \mathcal{F}, \mathsf{P}) \equiv \Big(\{-1,1\}^{\infty},\ \mathcal{B}(\{-1,1\}^{\infty}),\ \mathsf{P}_{\mathsf{Bern}}\Big).$$

Положим для $x = (x_1, x_2, \ldots)$

$$S_n(x) = x_1 + \cdots + x_n, \qquad n \geq 1.$$

Классические результаты теории вероятностей утверждают, что для случайного блуждания $(S_n(x))_{n\geq 1}$ следующие свойства выполнены **$\mathsf{P}_{\mathsf{Bern}}$-почти наверное**:

- $\displaystyle \lim_{n\to\infty} \frac{S_n(x)}{n} = 0$ (усиленный закон больших чисел);
- $\displaystyle \limsup_{n\to\infty} \frac{S_n(x)}{\sqrt{n}} = +\infty, \quad \liminf_{n\to\infty} \frac{S_n(x)}{\sqrt{n}} = -\infty;$
- $\displaystyle \frac{S_n(x)}{\sqrt{n}\log n} \to 0, \qquad n \to \infty;$
- $\displaystyle \limsup_{n} \frac{S_n(x)}{\sqrt{2n\log\log n}} = 1$ (закон повторного логарифма).

Примечательно, что

> алгоритмический подход позволяет конструктивно описать те **индивидуальные** последовательности $x = (x_1, x_2, \ldots) \in \Omega$, для которых сформулированные выше утверждения справедливы (**В. Г. Вовк, А. Шень**).

1.2. Рихард фон Мизес был первым, кто обратился в 1919 г. к понятию ‘бесконечная **случайная** последовательность’. Его замысел был на самом деле в том, чтобы

> заложить основы теории вероятностей, беря в качестве первичного именно понятие
>
> **бесконечной случайной последовательности**

Заметим, что в общепринятой в настоящее время аксиоматике теории вероятностей Колмогорова (1933 г.) первичным является другой объект, а именно,

распределение вероятностей

В §2 своей монографии “Grundbegriffe der Wahrscheinlichkeitsrechnung” (Springer, Berlin, 1933) Колмогоров, отмечая значимость мизесовского подхода к теории вероятностей, писал:

В изложении необходимых предпосылок для приложимости теории вероятностей к миру действительных событий автор в значительной мере следует выводам Мизеса (в частности, ср.

In der Darstellung der notwendigen Voraussetzungen für die Anwendbarkeit der Wahrscheinlichkeitsrechnung auf die Welt der reellen Geschehnisse folgt der Verfasser in hohem Maße den Ausführungen von Herrn von Mises (vgl. insbesondere

R. von Mises. [**Vorlesungen aus dem Gebiete der angewiesen Mathematik. Bd. 1: Wahrscheinlichkeitsrechnung.** Leipzig u. Wien, Fr. Deuticke, 1931], с. 21-27, параграф “Das Verhältnis der Theorie zur Erfahrungswelt”).

Мизес не дал четкого формального математического определения понятия *случайная* последовательность, ограничившись апеллированием к интуитивным идеям

- ‘‘нерегулярности их образования’’,
- ‘‘непредсказуемости их будущих значений по прошлым’’,
- невозможности по таким последовательностям, предъявляемым в казино, построить выигрышные стратегии.

Но, несомненно, сама идея возможного определения ‘‘случайной’’ последовательности привела в дальнейшем к большому циклу работ, приведших к разным и естественным подходам к определению концепции ‘‘случайности’’.

Как уже было отмечено, все эти подходы основаны на, казалось бы, чуждом теории вероятностей понятию ‘алгоритм’. Но именно теория алгоритмов дала возможность уточнения расплывчатого мизесовского понятия ‘допустимых правил выбора’, используемый Мизесом при определении понятия ‘бесконечная случайная последовательность’.

$<$Согласно Мизесу, допустимое правило выбора есть процедура выбора подпоследовательности из данной последовательности $x = (x_1, x_2, \ldots)$ такая, что решение о включении члена x_n в подпоследовательность не может зависеть от значения x_n.$>$

В настоящее время известны следующие четыре основных подхода к определению понятия ‘бесконечная случайная последовательность’, основанные на выполнении одного из четырех требований, интуитивно предъявляемых к тому, что мы называем ‘случайностью’:

ЧАСТОТОУСТОЙЧИВОСТЬ (устойчивость частот), или **стохастичность**

Мизес, Вальд
Чёрч, Колмогоров
Ловеланд

ТИПИЧНОСТЬ (принадлежность к множеству эффективной меры единица)

Мартин-Лёф
Левин
Шнорр

СЛОЖНОУСТРОЕННОСТЬ, или **ХАОТИЧНОСТЬ**

Колмогоров
Левин
Шнорр

НЕПРЕДСКАЗУЕМОСТЬ

Вилль, Успенский

§ 2. К ИСТОРИИ СТАНОВЛЕНИЯ ТЕОРИИ ВЕРОЯТНОСТЕЙ

Для лучшего освещения вопросов, связанных с понятиями “вероятность” и “случайность” целесообразно сейчас напомнить некоторые основные этапы в становлении теории вероятностей как математической науки.

Интуитивные представления о *случайности* и возникновение разного рода рассуждений о возможных *шансах* (в культовой практике, разрешении споров, предсказаниях, ...) уходят в глубь веков и к ним относились как к божественным явлениям, не поддающимся человеческому разуму.

Археологические находки говорят о том, что уже в давние времена * в примитивных играх использовались первые “случайные инструменты” — шестигранные кости (*astragalus* **).

В эпоху Возрождения (конец XIV – начало XVII в.) мы находим следы более или менее серьезных дискуссий, в основном *философского* характера, относительно “вероятностных” рассуждений:

Fra Luca Pacioli (1445–1517(?)), **Celio Calcagnini** (1479–1541)
Nicola Fontana Tartaglea (1500–1557)

---

* во времена Первой династии в Египте (около 3500 до н.э.), затем в Древней Греции и Древнем Риме.

** **astragalus** — запяточная кость у парнокопытных; она имеет такую форму, что при бросании может упасть лишь на одну из (разных) четырех сторон, поскольку две другие имеют закругленную форму.

Одним из первых, кто стал **математически** анализировать игровые шансы, был

Джероламо Кардано (Gerolamo CARDANO, 1501–1576)

решивший уравнение третьей степени и широко известный как изобретатель “карданного вала”. В его манускрипте (ок. 1525 г.), опубликованном лишь в 1663 г. под названием

Liber de Ludo Aleæ (*“Книга об азартных играх”*)

была высказана идея **комбинаций**, с помощью которых удобно описывать множество всех возможных исходов и благоприятных исходов.

Изложенное обычно относят к *Предыстории* теории вероятностей. Вообще же при изложении вопросов истории теории вероятностей принято выделять следующие этапы:

ПРЕДЫСТОРИЯ

1-й ПЕРИОД (XVII век – начало XVIII века)

2-й ПЕРИОД (XVIII век – начало XIX века)

3-й ПЕРИОД (вторая половина XIX века)

4-й ПЕРИОД (начало и середина XX века)

ПЕРВЫЙ ПЕРИОД (XVII в. – начало XVIII в.)

связывают с рождением “исчисления вероятностей”, считая его началом переписку (1654 г.) между Паскалем (Blaise Pascal, 1623–1662) и Ферма (Pierre de Fermat, 1601–1665).

В 1657 г. выходит в свет книга **Х. Гюйгенса** (Christianus Huygens, 1629–1695)

De Ratiociniis in Aleæ Ludo (*“О расчетах в азартных играх”*).

Эта книга, считающаяся первым систематическим текстом по “исчислению вероятностей”, была с интересом встречена современниками и оставалась на протяжении почти полувека единственным введением в теорию вероятностей.

Центральной фигурой рассматриваемого периода является

Я. БЕРНУЛЛИ (1654–1705),

который ввел понятие **'вероятность события'** и первым

- рассмотрел бесконечные последовательности повторных испытаний и

- поставил вопрос о предельном поведении частот появления тех или иных событий в этих испытаниях — кардинально новая ("нефинитная") идея в вероятностных рассмотрениях, ограничивающихся тогда элементарно-арифметическими и простейшими комбинаторными приемами; эта постановка привела Бернулли ("Ars Conjectandi", 1713) к

  закону больших чисел,

  носящему сейчас его имя.

ВТОРОЙ ПЕРИОД (XVIII век – начало XIX века)

Этот период связана с такими именами, как

Pierre-Rémond de **Montmort**	(1678–1719)
Abraham De **Moivre**	(1667–1754)
Thomas **Bayes**	(1702–1761)
Pierre Simon de **Laplace**	(1749–1827)
Carl Friedrich **Gauss**	(1777–1855)
Siméon Denis **Poisson**	(1781–1840)

МОНМОР в книге Essai d'Analyse sur les Jeux de Hasard (1708), уделяет основное внимание именно развитию методов расчетов в разнообразных играх.

МУАВР в книгах Doctrine of Chances (1718) и Miscellanea Analytica Supplementum (1730) дает — в некоторых частных случаях — определение таких понятий, как

- *независимость* событий,
- *ожидание*,
- *условная вероятность*.

Наиболее известно имя Муавра в связи с нормальной аппроксимацией биномиального распределения — называемой по предложению Д. Пойя (1920)

ЦЕНТРАЛЬНОЙ ПРЕДЕЛЬНОЙ ТЕОРЕМОЙ

Работа **БАЙЕСА** "An Essay Towards Solving a Problem in the Doctrine of Chances" (1763) дала

$$\boxed{\textbf{ФОРМУЛУ БАЙЕСА}}$$

— правило пересчета априорных вероятностей в апостериорные по происшествии некоторого события.

ЛАПЛАС, вслед за Бернулли, четко придерживался "классического" определения понятия вероятности (в случае конечного числа возможных исходов).

Однако уже в этот период появляются и "неклассические" вероятности. Так, у Байеса апостериорные вероятности могли быть неравными.

Отметим, что появившиеся у Муавра, Лапласа и Пуассона, как мы сейчас говорим,

> ‘нормальный’ и ‘пуассоновский’ законы рассматривались лишь как некоторые <u>аппроксимации</u>, а не как <u>распределения вероятностей</u> (в современном понимании этого термина).

Из сказанного становится ясным, что рамки “классической” (финитной) теории вероятностей существенно сдерживали возможности ее развития. В этот период в теории вероятностей отсутствовали абстрактные математические конструкции и она не рассматривалась иначе как <u>прикладная</u> математика.

ТРЕТИЙ ПЕРИОД (вторая половина XIX века)

Основным местом, где разрабатывались в это время общие проблемы теории вероятностей, стал **Петербург**, где существенный вклад в расширение и углубление всей системы теории вероятностей внесли:

П. Л. Чебышёв	(1821–1894)
А. А. Марков	(1856–1922)
А. М. Ляпунов	(1857–1918)

Именно благодаря им произошел

отказ от ограничения лишь случаем “классических” вероятностей.

ЧЕБЫШЕВ с полной ясностью оценил роль понятия случайной величины; им был разработан новый прием доказательства предельных теорем — **метод моментов**, позже усовершенствованный

МАРКОВЫМ, который ввел также принципиально новую концепцию — схему зависимых величин, связанных в **“цепь Маркова”**.

Неожиданный шаг в отыскании общих условий справедливости “теоремы Муавра–Лапласа” сделал

ЛЯПУНОВ, разработавший **метод характеристических функций**, который позволил ему доказать эту теорему в предположении, что у независимых слагаемых конечны моменты порядка $2 + \delta$, $\delta > 0$.

В **ЗАПАДНОЙ ЕВРОПЕ** интерес к теории вероятностей во второй половине XIX века стал стремительно возрастать благодаря обнаружившимся глубоким ее связям с

чистой математикой,

статистической физикой

и начавшей бурно развиваться

математической статистикой.

Упомянем здесь лишь имена

Henri **Poincaré** (1854–1912)	возвратность движений в динамических системах
Hugo **Gyldén** (1841–1896)	вопросы планетарной устойчивости и вероятностная теория чисел
James Clerk **Maxwell** (1831–1879)	распределение Максвелла для молекулярных скоростей
Ludwig **Boltzmann** (1844–1906)	*временные средние* и *эргодическая гипотеза*
Josiah Willard **Gibbs** (1839–1903)	понятие ансамбля и “распределение Гиббса”

Для всего последующего развития теории вероятностей и углубления понимания роли вероятностных подходов и концепций важную роль сыграли:

- обнаруженный в 1827 году **Р. Брауном** (1773–1858) феномен, получивший название **броуновского движения**.

  Качественное объяснение и количественное описание броуновского движения были даны впоследствии **А. Эйнштейном** (1879–1955) и **М. Смолуховским** (1872–1917).

- явление **радиоактивного распада**, обнаруженное в 1896 году **А. Беккерелем** (1852–1908) при исследовании свойств урана.

  Явление радиоактивности нашло свое объяснение в рамках **квантовой механики**, создание которой относится к двадцатым годам прошлого века.

ЧЕТВЕРТЫЙ ПЕРИОД (начало и середина XX века)

Выявленные к концу XIX века связи теории вероятностей с чистой математикой привели к постановке **Д. Гильбертом** (David Hilbert, 1862–1943) в его программном докладе 8 августа 1900 г. на Втором математическом конгрессе в Париже задачи

***МАТЕМАТИЗАЦИИ* теории вероятностей**.

Среди его известных проблем *шестая* * формулировалась как проблема

> аксиоматизации тех физических дисциплин, в которых математика играет доминирующую роль.

К этим дисциплинам Д. Гильберт отнес

теорию вероятностей и **механику**.

---

* *Первая* проблема была связана с континуум-гипотезой.

Четвертый период в истории становления теории вероятностей — это период логического обоснования теории вероятностей и становления ее математической дисциплиной.

Вскоре после доклада Д. Гильберта было предпринято несколько попыток построения

математической теории вероятностей

с привлечением элементов теории множеств и теории меры.

В 1904 г. **R. Lämmel** для описания множества исходов обращался к **теории множеств**, однако само понятие вероятности оставалось на интуитивном уровне и ассоциировалось с объемом, площадью, длиной, ...

В 1907 г. **U. Broggi** в диссертации, выполненной под руководством Д. Гильберта, также обращался и к **теории меры** Бореля–Лебега, но само понятие (конечно-аддитивной) вероятности требовало для своего определения обращения (в простейших случаях) к “относительным мерам”, “относительным частотам” и (в общих случаях) к некоторым искусственным предельным процедурам.

В 1917 г. **С. Н. Бернштейн** предложил систему аксиом, основанную на понятии ***качественного* сравнения событий** по степени их большего или меньшего правдоподобия. Само же численное значение вероятности появлялось как некоторое производное понятие.

В конце 1920-х – начале 1930-х **Б. де Финетти** предложил весьма сходный подход, основанный на *субъективных* качественных суждениях (“**системе знаний субъекта**”).

В 1919 г. Р. Мизес предложил так называемый

ЧАСТОТНЫЙ ПОДХОД

[говорят также: **статистический**

или **эмпирический**]

к обоснованию теории вероятностей, положив в основу идею, что

> вероятностные концепции могут применяться только к так называемым **“коллективам”**, т. е. *индивидуальным бесконечным упорядоченным последовательностям*, обладающим некоторым свойством “случайности” их образования.

Мы подошли сейчас собственно к основной теме доклада — изложению мизесовского подхода к понятию “случайность” и дальнейшего его развития, о чем вкратце уже говорилось.

§ 3. Частотная теория вероятностей Мизеса

Рихард фон Мизес (19.04.1883–14.07.1953) был, как бы сейчас сказали, прикладным математиком, широко известным своими работами в механике и особенно гидродинамике, теории полетов. (В Гарвардском университете он был профессором аэродинамики и прикладной математики.) Значителен его вклад и в вопросы оснований теории вероятностей (1919). *

Мизес, прежде всего, интересовался применимостью теории вероятностей к явлениям реального мира. На теорию вероятностей он смотрел как на учение о массовых явлениях и, как следствие, считал ее естественно-научной дисциплиной, определяемой спецификой "массовых явлений". (Ср. с физикой, биологией, ..., имеющими определенную математическую специфику.)

---

* **Grundlagen der Wahrscheinlichkeitsrechnung**, *Math. Z.* 5 (**1919**), 52–99.

Как и всякую естественно-научную дисциплину, учение о “массовых явлениях” можно изучать разными методами (в том числе и математическими), но так или иначе оставаясь в рамках своего предмета. *

Мизес понимал — для того, чтобы заложить основы теории вероятностей, нужна некоторая идеализация предмета, и это он воплощает в виде специальных “бесконечных последовательностей”, называемых *коллективами*. Он полагал, как уже отмечалось, что учение о вероятности должно быть непосредственным образом связано со случайными последовательностями.

---

* Для тех, кто следует аксиоматике Колмогорова, теория вероятностей может считаться математической дисциплиной, являющейся частью общего учения о множествах и функциях.

Более точно схема, принятая Мизесом, может быть описана следующим образом.

Имеется M — некоторое выборочное пространство (“Merkmalraum”) точек (“labels”). Предполагается, что должна иметься возможность проведения бесконечного числа испытаний, приводящих к последовательности $x = (x_1, x_2, \dots)$, где x_n — результат исхода (со значениями в M) в n-м испытании. Пусть A — некоторое подмножество в фазовом пространстве M и

$$\nu_n(A; (x_k)_{k \le n}) = \frac{1}{n} \sum_{k=1}^{n} I_A(x_k)$$

— частота появления “события” A в n испытаниях, приведших к последовательности точек $(x_1, \dots, x_n)$.

Бесконечную последовательность $x = (x_1, x_2, \ldots)$ Мизес назвал

КОЛЛЕКТИВОМ,

если она удовлетворяет следующим двум постулатам:

(A) существует предел $\boxed{\lim_{n\to\infty} \nu_n(A; (x_k)_{k\le n}) \ \ (= P(A;x))}$;

(B) существует предел $\boxed{\lim_{n\to\infty} \nu_n(A; (x'_k)_{k\le n}) \ \ (= P(A;x'))}$

для всех подпоследовательностей $(x'_k)_{k\ge 1} = (x'_1, x'_2, \ldots)$, получаемых из последовательности $x = (x_1, x_2, \ldots)$ посредством *всех* **"допустимых"** выборов элементов $x'_k = x_{n_k}$, $n_1 < n_2 < \cdots$, из последовательности $x = (x_1, x_2, \ldots)$.

Предполагается, что для всех "допустимых" множеств A и всех "допустимых" правил выбора пределы $P(A;x')$ должны совпадать с пределом $P(A;x)$: $P(A;x') = P(A;x)$.

ПРИМЕР образования **допустимой подпоследовательности** x'

из последовательности $\quad x = 0\;0\;1\;0\;1\;0\;1\;1\;1\;0\;1\;1\;0\;1\;0\ldots$

Читаем последовательность x слева направо и отмечаем (знаком $\boxed{0\,1}$) "слова" 0 1: $\quad x = 0\,\boxed{0\,1}\,\boxed{0\,1}\,\boxed{0\,1}\,1\,1\,\boxed{0\,1}\,1\,\boxed{0\,1}\,0\ldots$

$\Downarrow$

Подпоследовательность x' образуют цифры, непосредственно следующие за словами 01: $\quad x = 0\,\boxed{0\,1}\,\boxed{\mathbf{0}\,1}\,\boxed{\mathbf{0}\,1}\,\mathbf{1}\,1\,\boxed{0\,1}\,\mathbf{1}\,\boxed{0\,1}\,\mathbf{0}\ldots$

$\Downarrow$

Допустимая подпоследовательность: $\quad x' = \mathbf{0\;0\;1\;1\;0}\ldots$

Другой пример допустимого правила выбора: последовательность x' состоит из тех элементов $x_{i_1}, x_{i_2}, \ldots$ последовательности x, номера i_k которых являются простыми числами.

Особо важен в мизесовском подходе второй постулат (B), призванный у него отражать идею “случайности”, отсутствия “регулярности” в формировании коллектива.

Из сказанного видно, что при построении вероятностей на выборочном пространстве X Мизес исходит из того, что

> эти вероятности могут быть определены лишь в связи с существованием коллективов, для которых (по постулату (B)) имеет место “случайность” их образования.

Этот постулат (B) о “случайности” вызвал серьезную критику всего (частотного) подхода Мизеса к основаниям теории вероятности, прежде всего, в связи с тем, что Мизесом не было дано формального определения “допустимых” правил выбора. (Существование коллективов он обосновывал, среди прочего, ссылками на существование игральных домов, на невозможность построения выигрышных стратегий против выдаваемых в казино “случайных” последовательностей $x = (x_1, x_2, \ldots)$.)

Но, так или иначе, важнейшая заслуга частотной теории вероятностей Мизеса — в том, что в ней был поставлен вопрос:

Какие бесконечные последовательности отвечают нашим представлениям о “случайности”?

§ 4. ЧАСТОТОУСТОЙЧИВОСТЬ (или СТОХАСТИЧНОСТЬ)

Как уже было сказано, Мизес не дал формально-логического определения понятия “случайной” последовательности, поскольку не было дано определения “допустимых правил выбора” и поэтому было неясно, для каких же подпоследовательностей должен быть выполнен постулате (B).

$\langle$Напомним еще раз, что по Мизесу допустимое правило отбора — это процедура выбора подпоследовательности данной последовательности $x = (x_1, x_2, \dots)$ такая, что решение о включении члена x_n в подпоследовательность не может зависеть от значения x_n.$\rangle$

Не ясны были и теоретико-множественные вопросы (аддитивность, счетная аддитивность) относительно (возникающих как предельные образования) “вероятностей” $P(A, x)$.

Ф. Хаусдорф в своем письме к Д. Пойа (январь 1920) резко высказался против мизесовского подхода к теории вероятностей, выразив сомнение относительно самого *существования* “коллективов” со свойствами инвариантности пределов частот при переходе к допустимым подпоследовательностям.

Одним из первых, кто попытался дать логически состоятельную формулировку "допустимых правил выбора" подпоследовательностей и установить непустоту класса "коллективов", был

А. ВАЛЬД *: Его идея состояла в том, чтобы

> "допустимые" последовательности строить по **функциям** $W = W(x^{(n)})$, определенным на цепочках $x^{(n)} = (x_1, \dots, x_n)$, $n \geq 1$, считая, что каждая из этих функций принимает одно из двух значений, 1 или 0.

(Удобно ввести также "пустую" цепочку $x^{(0)}$, на которой также определено значение $W(x^{(0)})$, равное или единице или нулю.)

---

* **Die Widerspruchsfreiheit des Kollektivbegriffes der Wahrscheinlichkeitsrechnung**, *Ergebnisse eines mathematischen Kolloquiums*, **8**, 38–72; **1937** .

ПОСТРОЕНИЕ ПОДПОСЛЕДОВАТЕЛЬНОСТИ, "W-ДОПУСТИМОЙ" по Вальду:

если $W(x^{(0)}) = 1$, то x_1 включаем в подпоследовательность,
если $W(x^{(0)}) = 0$, то не включаем;

$W(x^{(1)}) \equiv W(x_1)$;

если $W(x_1) = 1$, то x_2 включаем в подпоследовательность,
если $W(x_1) = 0$, то не включаем;

....

То есть

"допустимая" (или "W-допустимая") последовательность есть

$$x_{n(1)}, x_{n(2)}, \ldots \quad \textbf{где} \quad \begin{aligned} &n(1) = \min\{k \geq 0\colon W(x^{(k)}) = 1\}, \\ &n(2) = \min\{k > n(1)\colon W(x^{(k)}) = 1\}, \\ &\ldots \end{aligned}$$

Основной результат А. Вальда состоял в доказательстве **НЕПУСТОТЫ класса “коллективов”**. Именно:

> Пусть S — некоторая <u>счетная</u> система допустимых правил выбора, каждое из которых определяется своим набором функций W. Тогда существует бесконечное число последовательностей, удовлетворяющих постулатам (А) и (В).

! Однако нельзя гарантировать, что предельные функции множеств $P(A;x)$ будут счетно-аддитивными — а без этого нельзя ставить, например, вопрос о наличии свойств типа закона повторного логарифма.

Анализ вальдовских допустимых последовательностей показывает, что они могут быть достаточно стохастически “регулярны” — хотя наша интуиция говорит, что “случайная” последовательность должна быть довольно беспорядочной.

В **1940** г. **Алонзо Чёрч** (Alonzo Church: “On a concept of a random sequence”), один из создателей теории алгоритмов, высказал идею, что привлекаемые к рассмотрению правила отбора должны быть реально осуществимыми. Для реализации этой идеи он предложил при образовании подпоследовательностей использовать вычислимые функции, т.е. функции, для которых существуют вычисляющие их алгоритмы.

Аккуратное определение понятия ‘вычислимая функция’ было дано **А. Чёрчем** еще в 1936 г. Такие функции, определенные для натуральных аргументов и значений, он предложил идентифицировать с ‘общерекурсивными функциями’. (Он также дал первый пример невычислимой функции.) В том же 1936 г., **Э. Пост** и **А. Тьюринг** дали определение понятия ‘алгоритм’, использующее понятие идеализированной вычисляющей машины. Наиболее общее определение понятия алгоритма было предложение в 1953 г. **А. Н. Колмогоровым**. Он доказал, что это общее определение сводится к алгоритму вычисления значений частично рекурсивной функции.

Примеры алгоритмов: правила сложения, вычитания, умножения, деления столбиком; алгоритм Евклида нахождения наибольшего общего делителя двух целых чисел, больших 1; динамическое и линейное программирование.

Класс алгоритмов счетен $\Rightarrow$ класс вычислимых функций счетен $\Rightarrow$ класс $R(\mathrm{MWCh})$ последовательностей, случайных по Мизесу–Вальду–**Чёрчу**, непуст: $\boxed{R(\mathrm{MWCh}) \neq \varnothing}$.

Однако класс $R(\mathrm{MWCh})$ оказался все еще слишком широким:

- **Ж. Вилль** (1939) построил последовательность, случайную в смысле определения Мизеса–Вальда–**Чёрча**, но слишком регулярную, чтобы ее можно было считать случайной — для нее, например, не выполнялся закон повторного логарифма, который естественно предполагать выполненным для последовательностей, призванных именоваться случайными.

- **Д. Ловеланд** (1966) заметил, что класс $R(\mathrm{MWCh})$ содержит последовательности, которые после некоторой вычислимой перестановки членов перестают принадлежать этому классу.

Принимая во внимание эти обстоятельства, Колмогоров в 1963 г. предложил

сузить класс R(MWCh)

ЗА СЧЕТ

расширения множества “допустимых правил выбора” (тестов), которые удовлетворяют мизесовскому постулату (B). *

---

* Заметим, что Д. Ловеланд в 1966 г. пришел к той же конструкции.

Подпоследовательности Чёрча (получаемые при помощи той или иной вычислимой функции φ) имели вид

$$\boxed{x_{\varphi(1)}, x_{\varphi(2)}, \ldots} \qquad \text{где } \varphi(i) < \varphi(j) \text{ при } i < j.$$

Допустимые (обобщенные) подпоследовательности Колмогорова (получаемые при помощи **двух** вычислимых функций ψ и φ) имели вид

$$\boxed{x_{\varphi(1)}, x_{\varphi(2)}, \ldots} \qquad \text{где } \varphi(i) \neq \varphi(j), \text{ если } i \neq j. \qquad (\varphi)$$

Процедура Колмогорова состояла из двух шагов:

(I) сначала (с помощью вычислимой функции ψ) строится обобщенная последовательность

$$\boxed{x_{\psi(1)}, x_{\psi(2)}, \ldots} \qquad (\text{с } \psi(i) \neq \psi(j), \text{ если } i \neq j), \qquad (\psi)$$

(II) затем (повторение процедуры Чёрча) из этой последовательности выбирается (с помощью вычислимой функции φ) подпоследовательность (φ).

Для полученного таким образом класса $R(\mathsf{K})$ верно

$$R(\mathsf{K}) \subset R(\mathsf{MWCh}).$$

К сожалению, этот класс $R(\mathsf{K})$ тоже оказался слишком широким, чтобы взять его в качестве “класса $R(\textbf{?})$” истинно “случайных” последовательностей — в $R(\mathsf{K})$ существует последовательность, в любом начальном отрезке которой число единиц превышает число нулей, что противоречит * как

- нашей интуиции (о “равномерности” появления единиц и нулей в “случайной” последовательности, полученной в симметричной схеме Бернулли), так и
- законам “возвратности” для схемы Бернулли.

---

* Превышение числа единиц над числом нулей соответствовало бы случаю $p > 1/2$ в схеме Бернулли с вероятностью “успеха”, равной p; но, как хорошо известно, в этом случае случайное блуждание невозвратно.

Таким образом, мы имеем цепочку включений

$$\boxed{R(\textbf{?}) \subset R(\mathsf{K}) \subset R(\mathsf{MWCh})}$$

($R(\textbf{?})$ — класс "истинно случайных" последовательностей, полученный с помощью мизесовских "допустимых" правил выбора).

Вышеупомянутые результаты в основном исчерпывают достижения, полученные на пути построения частотоустойчивых последовательностей. Стоит подчеркнуть, что

> **истинно случайный класс** $R(\textbf{?})$ "допустимых" частотоустойчивых последовательностей, для которых выполнены основные законы теории вероятностей,
>
> **по-прежнему НЕ ОПРЕДЕЛЕН**.

§ 5. ТИПИЧНОСТЬ (принадлежность к множеству эффективной меры единица)

Чтобы описать иной (а именно, так называемый "типический", говорят также — "количественный") подход к понятию случайности, предложенный

П. Мартин-Лёфом (**1966**),

напомним сначала усиленный закон больших чисел Бореля.

- Пусть $\boldsymbol{\Omega = [0,1)}$, $\boldsymbol{\mathcal{B}}$ — борелевская система подмножеств Ω, **P** — мера Лебега на $[0,1)$,

$\boxed{\boldsymbol{x = 0.x_1x_2\ldots}}$ — двоичное разложение числа $x \in \Omega$ (с бесконечным числом нулей).

- Определим сл.в. $\boldsymbol{\xi_1(x), \xi_2(x), \ldots}$, полагая $\xi_n(x) = x_n$.

Для любого $n \geq 1$ и всех $b_1, b_2, \ldots$, принимающих значения 0 и 1,

$$\{\omega\colon \xi_1(x) = b_1, \ldots, \xi_n(x) = b_n\} = \left\{ \sum_{i=1}^{n} \frac{b_i}{2^i} \leq x < \sum_{i=1}^{n} \frac{b_i}{2^i} + \frac{1}{2^n} \right\},$$

так что мера P этого множества равна $1/2^n$. Таким образом,

$\xi_1(x), \xi_2(x), \ldots$ — независимые одинаково распределенные сл.в. с $\mathsf{P}(\xi_i = 0) = \mathsf{P}(\xi_i = 1) = 1/2$.

УСИЛЕННЫЙ ЗАКОН БОЛЬШИХ ЧИСЕЛ БОРЕЛЯ

утверждает, что

> **почти все** $x = 0.x_1x_2\dots$ из $[0,1)$ нормальны в том смысле, что **с вероятностью 1** доля нулей (и доля единиц) в двоичном разложении числа x стремится к $1/2$, т.е. для "большинства" x'ов (точнее, P-п.н.)
>
> $$\boxed{\frac{1}{n}\sum_{k=1}^{n} I(x_k = 1) \longrightarrow \frac{1}{2}}. \qquad (*)$$

Тем самым, множество $x = 0.x_1x_2\dots$'ов, для которых предел

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} I(x_k = 1)$$

или не существует, или отличен от $1/2$, имеет P-меру нуль.

Такие "Р-нулевые" множества мы называем **пренебрежимыми**. Интуиция подсказывает, что

> если множество $U \subseteq \Omega$ пренебрежимо ($\mathrm{P}(U) = 0$), то все его элементы $x = 0.x_1x_2\ldots$ (иначе, все последовательности $(x_1, x_2, \ldots)$) должны объявляться **"НЕТИПИЧНЫМИ"** (т.е. не принадлежащими "большинству") — как не имеющие свойства $(*)$.

Но таких множеств U "нетипичных" x'ов, вообще говоря, "много", и поэтому

> в качестве множества "нетипических" x'ов надо брать <u>максимальное</u> множество, содержащее все U с $\mathrm{P}(U) = 0$; т.е.
>
> $$\boxed{\text{множество "нетипичности"} = \bigcup U}$$
>
> где сумма (вообще говоря, несчетная) берется по всем U таким, что $\mathrm{P}(U) = 0$.

ОДНАКО хорошо известно, что вполне может случиться так, что мера P множества $\bigcup U$ определена и P($\bigcup U$) $= 1$. Тем самым,

> **множество всех “нетипических”** $x = 0.x_1x_2\ldots$’ов (или множество последовательностей $(x_1, x_2, \ldots)$) — которое вроде бы естественно выбрать в качестве “множества ‘неслучайности’ ” — **имеет меру 1**.

Множество всех “типических” x’ов есть $\bigcap \overline{U}$ $(= \overline{\bigcup U})$, так что

> **множество “типических”** x**’ов**, которые мы хотим рассматривать как множество “случайности”, вообще говоря, **имеет P-меру нуль**

— что **противоречит как усиленному закону больших чисел, так и здравому смыслу**.

Способ преодолеть трудности в определении “типических” последовательностей, которые мы хотели бы объявить “случайными”, был предложен, как уже говорилось, **П. Мартин-Лёфом**, который дал уточнение понятия

‘пренебрежимое множество’ (т.е. ‘множество меры 0’);

а именно, он ввел новое (<u>алгоритмическое</u>) понятие

‘ЭФФЕКТИВНО пренебрежимое множество’,

позволившее дать естественное определение “типических последовательностей”, которые претендуют на роль “случайных”.

В теории меры множество U называют пренебрежимым или множеством P-меры нуль, если $\mathrm{P}(U) = 0$. Если Ω есть пространство последовательностей, удобно использовать следующее (эквивалентное) определение пренебрежимых множеств.

$\mathbf{\Omega} := \{x \colon x = x_1 x_2 \ldots\}$ есть множество <u>ВСЕХ</u> двоичных (т.е. $x_i = 0$ или 1) последовательностей (бесконечных слов).

$\mathbf{\Xi} := \{\xi \colon \xi = x_1 x_2 \ldots x_{|\xi|}\}$ есть множество <u>конечных</u> двоичных слов ξ с длинами $|\xi|$, принимающими значения в множестве $\{1, 2, \ldots\}$.

$\mathbf{\Omega}_{\boldsymbol{\xi}}$ есть множество <u>бесконечных</u> двоичных слов с начальным отрезком $\xi \in \mathbf{\Xi}$;
другими словами, Ω_ξ — цилиндрическое множество с “основанием” ξ.

ОПРЕДЕЛЕНИЕ I. Множество $U \subseteq \Omega$ называется

пренебрежимым (относительно меры P),

если для любого целого $m \geq 1$ найдется последовательность $\xi_1, \xi_2, \ldots$ двоичных слов (из Ξ) такая, что

$$U \subseteq \bigcup_n \Omega_{\xi_n}, \quad \text{где} \quad \sum_n \mathsf{P}(\Omega_{\xi_n}) = \sum_n 2^{-|\xi_n|} < \frac{1}{m}.$$

ЗАМЕЧАНИЕ.

Любое одноточечное (состоящее из единственной последовательности $x \in \Omega$) множество U, очевидно, пренебрежимо, поскольку начальные отрезки длины n имеют вероятность 2^{-n} и достаточно выбрать n так, что $2^{-n} < 1/m$.

Следующее понятие эффективно пренебрежимого множества, которое нужно для наших целей, уточняет приведенное выше определение пренебрежимого множества.

ОПРЕДЕЛЕНИЕ II. Множество $U^* \subseteq \Omega$ называется

ЭФФЕКТИВНО пренебрежимым

(относительно меры P),

если для любого целого $m \geq 1$ существует m-эффективно вычислимая последовательность $\xi_1^*, \xi_2^*, \ldots$ двоичных слов (из Ξ) такая, что

$$U^* \subseteq \bigcup_n \Omega_{\xi_n^*}, \quad \text{где} \quad \sum_n \mathsf{P}(\Omega_{\xi_n^*}) = \sum_n 2^{-|\xi_n^*|} < \frac{1}{m}.$$

В опр. II "m-эффективно вычислимая" последовательность определяется посредством **АЛГОРИТМИЧЕСКИХ понятий**:

Мы говорим, что последовательность двоичных слов

- $\xi_1, \xi_2, \ldots$ **алгоритмически вычислима**, если существует алгоритм, вычисляющий каждый ее член ξ_n по его номеру n;

- $\xi_1^*, \xi_2^*, \ldots$ **m-эффективно вычислима**, если существует алгоритм, который по числу m, поступившему на его вход, выдает другой алгоритм (программу), создающий алгоритмически вычислимую последовательность.

ОПРЕДЕЛЕНИЕ I. Множество $U \subseteq \Omega$ называют

пренебрежимым (отн-но меры Р),

если $\forall m \in \mathbb{N}$ найдется последовательность $\xi_1, \xi_2, \ldots$ двоичных слов (из Ξ) такая, что

$$U \subseteq \bigcup_n \Omega_{\xi_n}, \quad \text{где} \quad \sum_n \mathsf{P}(\Omega_{\xi_n}) = \sum_n 2^{-|\xi_n|} < \frac{1}{m}.$$

ОПРЕДЕЛЕНИЕ II. Множество $U^* \subseteq \Omega$ называют

ЭФФЕКТИВНО пренебрежимым (отн-но меры Р),

если $\forall m \in \mathbb{N}$ существует m-эффективно вычислимая последовательность $\xi_1^*, \xi_2^*, \ldots$ двоичных слов (из Ξ) такая, что

$$U^* \subseteq \bigcup_n \Omega_{\xi_n^*}, \quad \text{где} \quad \sum_n \mathsf{P}(\Omega_{\xi_n^*}) = \sum_n 2^{-|\xi_n^*|} < \frac{1}{m}.$$

Ценность понятия эффективно пренебрежимого множества раскрывается следующим результатом **П. Мартин-Лёфа**:

> **ТЕОРЕМА.** Существует эффективно пренебрежимое множество, содержащее ВСЕ эффективно пренебрежимые множества.

Таким образом, если дополнение эффективно пренебрежимого множества назвать ‘эффективно большим’, то *пересечение всех эффективно больших множеств снова является эффективно большим множеством и имеет (эффективную) меру 1*.

Упомянутое пересечение и объявляется множеством **“ТИПИЧЕСКИХ” последовательностей**, оно обозначается $R(\mathrm{T})$ и называется множеством

СЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ
по МАРТИН-ЛЁФУ.

Стоит отметить, что

> для “типических” последовательностей (т.е. для последовательностей класса R(T)) выполнены основные “Р-п.н. законы теории вероятностей”, включая закон повторного логарифма.

В цепочку ранее рассмотренных включений

$$R(\textbf{?}) \subseteq R(\mathsf{K}) \subset R(\mathsf{MWCh}),$$

класс R(T) интегрируется следующим образом:

$$\boxed{R(\textbf{?}) \subseteq R(\mathsf{T}) \subseteq R(\mathsf{K}) \subset R(\mathsf{MWCh})}$$

§ 6. СЛОЖНОУСТРОЕННОСТЬ (ХАОТИЧНОСТЬ)

Рассмотренные выше последовательности

$$\underline{\underline{(\mathrm{II}_{10})}}:\quad \mathbf{1\,1\,1\,1\ldots1\,1} \qquad \text{и} \qquad \underline{\underline{(\mathrm{III}_{10})}}:\quad \mathbf{1\,0\,1\,0\ldots1\,0}$$

имеют очень простую структуру. Эта простота проявляется в том, что их можно легко описать, и оправдывает тот факт, что мы склонны считать эти последовательности неслучайными.

С другой стороны, нельзя сказать, что последовательность

$$(\mathrm{I}_{10}):\quad \mathbf{0\,1\,1\,1\,0\,1\,0\,0\,1\,0}$$

имеет простую структуру, ее нелегко описать, и мы склонны считать ее “случайной”.

Напомним, что при описании классов R(MWCh) и R(K) мы интересовались в основном структурой алгоритмов, создающих подпоследовательности.

ПОДХОД КОЛМОГОРОВА, инициированный им в 1960-х, основан на рассмотрении **сложности структуры**

САМИХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ,

конечных или бесконечных. Колмогоров ввел некоторую числовую характеристику (называемую теперь **колмогоровской сложностью**) такую, что

- сложность конечной последовательности измеряется длиной ее кратчайшего "описания";
- бесконечная последовательность объявляется хаотической (как синоним "случайности"), если сложности ее начальных отрезков растут "так быстро, как только возможно".

Перейдем к формальным определениям, включая определение того, что понимается под “описанием”.

Каждая из рассмотренных выше последовательностей (I_{10}), (II_{10}), (III_{10}) может быть словесно описана

РАЗЛИЧНЫМИ СПОСОБАМИ.

Например, можно сказать, что

- “слово (III_{10}): **1 0 1 0 1 0 1 0 1 0** состоит из 10 букв, с 1 на нечетных местах и 0 на четных”.

Однако можно сказать и так:

- “слово (III_{10}): **1 0 1 0 1 0 1 0 1 0** состоит из 10 чередующихся букв 0 и 1, причем на первом месте стоит 1”.

Эти словесные описания имеют **разные длины**, однако ясно, что если мы хотим уметь находить среди описаний кратчайшее, то следует

> привести все эти описания к **единому стандарту**, так чтобы их длины измерялись единообразно.

Самый естественный — и самый простой — путь такого "стандартизованного" описания заключается в том, чтобы

кодировать последовательности в двоичном алфавите,

т.е. представлять их как **двоичные слова**.

Итак, мы будем предполагать, что

> как сами слова x, которыми мы интересуемся, так и их кодовые описания y принадлежат Ξ

Сложность слова $y \in \Xi$ есть, по определению, величина

$$\mathrm{Comp}(y) \equiv \min\{|x| : \ x \text{ есть кодовое описание } y\}.$$

Ясно, что существуют различные “способы описания”. Поэтому следует определиться: что в том или ином случае понимается под “способом описания”**?**

Один из подходов состоит в

рассмотрении так называемых **алгоритмически вычислимых** отображений $f\colon \Xi \to \Xi$ как “способов описания”.

В первом приближении ситуация представляется следующей:

> “машина” по последовательно поступающим значениям $x = (x_1, x_2, \ldots)$ “выдает” некоторые значения $f(x)$, образующие последовательность нулей и единиц.

Двоичное слово x называется

“описанием” (“f-описанием”) конечного слова y,

если y является начальным фрагментом (конечной или бесконечной) последовательности $f(x)$:

$$x = \cdots\cdots \longrightarrow \boxed{\text{МАШИНА}} \longrightarrow f(x) = \underbrace{0111001}_{y}\ldots$$

ОПРЕДЕЛЕНИЕ 1. **Сложность** слова y при заданном алгоритмически вычислимом отображении f есть число

$$K_f(y) = \min\{|x| : x \text{ есть } f\text{-описание } y\},$$

где $|x|$ — длина двоичного слово x; $\min\varnothing := +\infty$.

ОПРЕДЕЛЕНИЕ 2. Алгоритмически вычислимое отображение f называется **оптимальным**, если для любого алгоритмически вычислимого отображения g существует константа $C = C(g)$ такая, что

$$K_f(y) = K_g(y) + C \quad \text{для всех двоичных слов } y.$$

Колмогоров и Соломонов показали, что для некоторых важных семейств $\mathcal{F}$ такие отображения f существуют.

ОПРЕДЕЛЕНИЕ 3. Для произвольного оптимального алгоритмически вычислимого отображения его сложность называется **энтропией**.

Таким образом, энтропия зависит от оптимального отображения. Для каждого оптимального отображения существует своя собственная энтропия. Это может показаться неблагоприятным обстоятельством, однако любые две оптимальные энтропии $K_1(y)$ и $K_2(y)$ отличаются лишь на константу:

$$|K_1(y) - K_2(y)| < C.$$

Как будет видно из дальнейшего, чтобы определить хаотическую последовательность, достаточно выбрать лишь какую-то одну оптимальную энтропию, обозначим ее $K(y)$.

Заметим, что для тождественного отображения $g(y) = y$ сложность слова y, очевидно, равняется его длине. Таким образом, $K(y) \le |y| + C$, и если $y = (y_1, \ldots, y_n)$, то

$$K(y_1, \ldots, y_n) \le n + C.$$

ОПРЕДЕЛЕНИЕ 4. Последовательность $y = (y_1, \ldots, y_n)$ называется **хаотической**, если существует константа C такая, что для любого n

$$K(y_1, \ldots, y_n) > n - C.$$

Это определение показывает, что свойство последовательности быть хаотической <u>не зависит</u> от конкретного выбора оптимальной энтропии.

В 1973 г. Л. Левин предложил уточнение класса $\mathcal{F}$ алгоритмически вычислимых отображений и изучил соответствующее понятие энтропии (называемой монотонной).

К этому пришел и С. Р. Шнорр (1973). Их результаты (с учетом уточнений класса $\mathcal{F}$) приводят к следующей важной теореме.

ТЕОРЕМА (Левин–Шнорр). Класс $R(\mathsf{C})$ хаотических последовательностей совпадает с классом $R(\mathsf{T})$ типических последовательностей:

$$\boxed{R(\mathsf{C}) = R(\mathsf{T})}\,.$$

Итак, обозначая
- через $\boldsymbol{R}(\mathrm{T})$ множество “<u>типических</u>” и
- через $\boldsymbol{R}(\mathrm{C})$ множество “<u>хаотических</u>”

последовательностей, мы имеем следующую диаграмму:

хаотическая случайность, или **сложноустроенность**

частотоустойчивая случайность **по Колмогорову**

$$\boxed{\boldsymbol{R}(\textbf{?})} \subseteq \boldsymbol{R}(\mathrm{C}) = \boldsymbol{R}(\mathrm{T}) \subseteq \boldsymbol{R}(\mathrm{K}) \subset \boldsymbol{R}(\mathrm{MWCh})$$

типическая случайность **по Мартин-Лёфу**

частотоустойчивая случайность **по Мизесу–Вальду–Чёрчу**

Заметим, что для хаотической-типической случайной последовательности (принадлежащей классу $R(\mathrm{C}) = R(\mathrm{T})$) справедливы основные законы теории вероятностей.

§ 7. НЕПРЕДСКАЗУЕМОСТЬ

Как уже было отмечено, Мизес не дал четкого определения понятия “случайная последовательность”, а само их существование он обосновывал ссылкой на то, что такие последовательности имеются в распоряжении казино. Приходящему в казино игроку предлагается угадывать предъявляемые члены “случайной последовательности” и делать при этом денежные ставки. Казино убеждено в “непредсказуемости” таких последовательностей, не позволяющей игроку создать стратегию, которая бы разорила казино.

По-видимому, Ж. Вилль был первым, кто использовал * игровую интерпретацию для определения “непредсказуемости” как синонима “случайности” последовательности.

---

* см. небольшую монографию *“Étude critique de la notion de collectif”*, Gauthier-Villars, Paris, 1939.

Возражение Вилля к мизесовскому понятию коллектива:

(a) Для любого заданного счетного множества правил выбора можно построить последовательность $x = (x_1, x_2, \ldots)$, являющуюся коллективом и обладающую тем свойством, что для всех n, за исключением конечного числа, $\sum_{k \le n} x_k \ge n/2$, что нетипично ввиду закона повторного логарифма (так что последовательность x недостаточно беспорядочна.

(b) Мизесовская формализация игровых стратегий (в защиту понятия коллектива), основанная на использовании допустимых правил отбора, несовершенна, т.к. существует стратегия (**МАРТИНГАЛ**), которая дает бесконечный выигрыш на последовательностях типа (a), в то время как допустимых правил отбора, которые бы делали это, нет. Таким образом, коллективы не дают абсолютно адекватной модели случайных явлений.

---

* M. van Lambalgen, **Randomness and foundations of probability: von Mises axiomatization of random sequence**, Probability, statistics and game theory, Institute for Mathematical Statistics, 1996.

Современные воззрения на

“игровой” подход к понятию “случайности”

через “**НЕПРЕДСКАЗУЕМОСТЬ**”

можно, следуя В.А.Успенскому *, изложить следующим образом.

---

* “Четыре алгоритмических лица случайности”, МЦНМО, 2009.

Пусть приходящий в казино игрок имеет некоторый капитал $V(0)$. Казино также располагает некоторым капиталом $W(0)$, который игроку неизвестен. Пусть предъявляемая игроку последовательность имеет вид $x = (x_1, x_2, \ldots)$, где $x_i = \pm 1$. Перед k-м ходом игрок выбирает величину ставки $\gamma_k = \gamma_k(x_1, \ldots, x_{k-1})$, и тогда его выигрыш/проигрыш на k-м шаге будет равен $\gamma_k x_k$, а общий капитал в момент k будет равен

$$V(k) = V(0) + \sum_{i=1}^{k} \gamma_i x_i.$$

Разумеется, предполагается, что ставка γ_k в каждый момент времени k не может быть больше капитала $V(k-1)$, т.е. на γ_k накладывается условие $\gamma_k \leq V(k-1)$. У игрока есть право брать $\gamma_k = 0$, что означает, что он не делает ставки. В этом случае капитал игрока не меняется.

По определению игрок выигрывает, если

$$\sup_k V(k) = \infty,$$

т.е. если независимо от того, какой капитал W был у казино, наступает момент, когда казино окажется разоренным.

Важно подчеркнуть, что все стратегии (γ_k) игрока считаются вычислимыми, т.е. задаются с помощью некоторого алгоритма.

Предъявляемая последовательность $x = (x_1, x_2, \dots$ называется *предсказуемой*, если для нее существует выигрышная вычислимая стратегия (γ_k). В противном случае такая последовательность называется *непредсказуемой*. Будем обозначать через $R(\mathsf{NP})$ класс непредсказуемых последовательностей. Известно, что

$$R(\mathsf{T}) \subseteq R(\mathsf{NP}) \subset R(\mathsf{K})$$

(подчеркнем, что второе включение строгое (!)).

Итак, мы можем подвести итог современному состоянию наших знаний о соотношении между разными классами $R(\cdot)$ последовательностей:

хаотическая случайность, или **сложноустроенность**

случайность как **непредсказуемость**

частотоустойчивая случайность **по Мизесу–Вальду–Чёрчу**

$$\boldsymbol{R}(\mathbf{?}) \subseteq \boldsymbol{R}(\mathrm{C}) = \boldsymbol{R}(\mathrm{T}) \subseteq \boldsymbol{R}(\mathrm{NP}) \subset \boldsymbol{R}(\mathrm{K}) \subset \boldsymbol{R}(\mathrm{MWCh})$$

типическая случайность **по Мартин-Лёф**

частотоустойчивая случайность **по Колмогорову**

Отметим, что до сих пор неизвестно, имеет ли место равенство $\underline{\underline{\boldsymbol{R}(\mathrm{T}) \stackrel{?}{=} \boldsymbol{R}(\mathrm{NP})}}$. Эта проблема ждет своего решения.

§ 8. ПРИМЕР

Теперь мы хотим проиллюстрировать, как

> **алгоритмическая теория вероятностей** позволяет дать новые доказательства некоторых результатов **классической теории вероятностей**.

В качестве примера возьмем

УСИЛЕННЫЙ ЗАКОН БОЛЬШИХ ЧИСЕЛ:

$$\frac{S_n(x)}{n} \to \frac{1}{2}, \quad \text{где} \quad x = (x_1, x_2, \ldots), \quad x_i = 0 \text{ или } 1$$

Идеи алгоритмического доказательства (**В. Вовк, А. Шень**) таковы.

Возьмем начальный отрезок $x_{(n)} = (x_1, \ldots, x_n)$, и пусть $p_n = \frac{1}{n}(x_1 + \cdots + x_n)$ — частота единиц в этом отрезке.

Со времен Шеннона мы знаем, что "**энтропия на одну букву**" в $x_{(n)}$ [т.е. число битов, необходимых для кодирования одной буквы в $x_{(n)}$] примерно равна $\boxed{H(p_n) \equiv -p_n \log_2 p_n - q_n \log_2 q_n}$.
Таким образом, чтобы закодировать $x_{(n)}$, нужно $nH(p_n)$ битов.

Но для кодирования $x_{(n)}$ мы должны знать вероятность p_n. Так что полный код для $x_{(n)}$ включает также p_n — рациональное число с числителем и знаменателем, не превосходящими n, и, значит, кодирование p_n требует не более $O(\log n)$ битов.

Таким образом, начальный отрезок $x_{(n)}$ любой последовательности $(x_0, x_1, \ldots)$ содержит не более

$$nH(p_n) + O(\log n) \quad \text{битов информации.}$$

Если последовательность $(x_0, x_1, \dots)$ является **хаотической**, то, по Колмогорову–Левину, монотонная энтропия (сложность) ее сегментов длины n должна быть $n + O(1)$. Поэтому

$$H(p_n) = \frac{n + O(1) + O(\log n)}{n} = 1 + O\left(\frac{\log n}{n}\right) \longrightarrow 1, \qquad n \to \infty,$$

и, следовательно, $p_n \to 1/2$. *

Вывод: для любой хаотической (=типической) последовательности $(x_0, x_1, \dots)$

$$p_n = \frac{x_1 + \cdots + x_n}{n} \longrightarrow \frac{1}{2}.$$

Множество таких последовательностей имеет меру 1, что и доказывает **классический** усиленный закон больших чисел. $\square$

---

* поскольку $H(p) = 1 + \mathrm{const}(p - 1/2)^2 + o((p - 1/2)^2)$ вблизи $p = 1/2$, так что $p_n - 1/2 = O(\sqrt{n^{-1} \log n})$.