Конспект лекций В. Л. Калинина по высшей математике студентам физико-технического факультета СПбГТУ (1993-1995г.) с комментариями

О. Л. Шалаев
http://chalaev.com

27 мая 2018 г.

Содержание

Памяти Владимира Леонидовича Калинина.

Последние два класса я учился в “крутой” физико-технической школе номер 566 в Санкт-Петербурге. Преподавание математики там оставляло желать лучшего: снобизма у некоторых учителей было больше, чем профессионализма. Я не мог понять, в чём смысл решения задачек про углы в тетраэдре или многочисленных тригинометрических уравнений, на которые нас натаскивали; математика представлялась мне скучным и заумным предметом. Школьные учителя (да и я сам) почти сумели убедить меня в том, что я неспособен к математике, и, возможно, вообще к точным наукам. Всё изменилось с началом обучения на физико-техническом факультете СПбГТУ, где почти весь (за искл. дифференциальных уравнений) курс лекций по высшей математике нам прочитал В. Л. Калинин. Владимир Леонидович открыл мне красоту математики; в своих лекциях он затрагивал в том числе разделы математики, обычно не преподаваемые будущим физикам или затрагиваемые лишь поверхностно (в том числе в Европе и США). Курс высшей математики для недавно (на тот момент) созданного физико-технического факультета был разработан им самостоятельно. Наивно было бы полагать, что подобное подвижничество будет оценено коллегами: никакими премиями или продвижением по карьерной лестнице “выскочку” не поощрали, и впоследствии Владимир Леонидович ушёл из СПбГТУ в Университет. Ещё во время моей учёбы Владимир Леонидович много курил; через 13 лет это привело его к преждевременной смерти от рака лёгких. После того, как некролог был удалён с сайта Университета, этот текст остаётся, пожалуй, наиболее значимым публичным воспоминанием о моём любимом учителе.

Несмотря на то, что курсы теоретической механики, электродинамики, статистической физики и квантовой механики нам читали хорошие преподаватели (2/3 из которых, к тому же – выдающиеся учёные), бо́льшая часть курса высшей математики ими не использовалась. Практически не используется она и во “взрослой” физике (в научных статьях). Теоретик, который начнёт писать статьи по физике твёрдого тела с применением аппарата теории дифференциальных форм, обречён на игнорирование научным сообществом,[1] для которого этот “новый” математический аппарат незнаком и непривычен. Нужны ли вообще все эти “новшества” в физике? Улучшают ли они решения физических задач, делая их проще, понятнее и короче? Оправдано ли внедрение этих техник в теоретическую физику? Ответ (частично) можно найти в книгах [1, 2], `arXiv/hep-th/0306205`. Ну и я, со своей стороны, кое-где добавляю от себя примеры из физики и вообще из жизни. Например, раздражает меня грубость, с которой сформулирована (обычная, кондовая) теория возмущений у Ландау-Лифшица и в университетском курсе. Я пытался перебороть это дело, используя подход Липпмана-Швингера, и для первых двух порядков это даже работало, но настоящей универсальности и красоты[2] не было!

Курс Калинина хорош, но в нём не хватает жизненных примеров. Характерный пример (и таких – много): при рассмотрении ориентируемых многообразий не упоминается лента Мёбиуса. . .

Текст незавершён, работа продолжается, файл периодически обновляется. . .

[1] Признание “научного сообщества” важно тому, кто хочет сделать карьеру в университете, где “научное сообщество” обладает властью. Я считаю, что большинство представителей этого “научного сообщества” наукой не занимается и является в большей степени инженерами и менеджерами, нежели учёными, см. http://polit.ru/article/2015/05/03/science/ С годами моё уважение к “научному сообществу” постепенно ослабевало и к настоящему моменту окончательно исчезло.

[2] Таких, как с гриновскими функциями.

Обозначения

Используется немецкое нумерование: точка после числа. Например “2. следствие” означает “второе следствие”. $\alpha(x)\nearrow$ – функция $\alpha(x)\nearrow$ монотонно возрастает. $\bar{\mathbb{R}} \equiv \mathbb{R} \cup \{-\infty, \infty\}$ – расширенное множество действительных чисел. нбчс = не более, чем счётное. Черта над знаком подмножества означает дополнение.

Доп. литература (не только на этот семестр, но также и на будущее): [3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25].

1 Предел и непрерывность

§1.вещественные числа §2.предел последовательности §3.монотонные последовательности §5.фундаментальные последовательности §6.функции §7.понятие об асимптотическом разложении §9.основные теоремы о непрерывных на отрезке функциях

2 Производная и её приложения

§1.правила вычисления производных §2.теоремы о среднем и их приложения §3.формула Тейлора §4.Выпуклые функции

3 Линейная алгебра

§1.Системы линейных уравнений. §2.Определители и их свойства. §3.Матрицы. §5.Теорема об определителе произведения матриц. §6.Линейная зависимость и независимость строк (столбцов). §7.Ранг матрицы. Общая теория систем линейных уравнений.

4 Линейные пространства

§1.Линейные пространства. §2.Подпространства линейного пространства. §3.Сопряжённое пространство. §4.Элементы геометрии афинного пространства. §5.Евклидовы и унитарные пространства. §6.Геометрия афинного пространства.

5 Неопределённый интеграл

§1.Элементарные свойства неопределённого интеграла. §2.Комплекснозначные функции. §3.Разложение многочленов на множители. §4.Разложение рациональных дробей на простейшие. §5.Интегрирование рациональных функций. §6.Интегрирование некоторых иррациональных функций.

6 Определённый интеграл

§1.Интеграл Римана-Стильтьеса.

6.1 Определение интеграла Римана-Стильтьеса

$[a, b] \subset \mathbb{R}$, $a < b$, $\alpha : [a, b] \to \mathbb{R}$, $\alpha(x)\nearrow$. Разбиение отрезка: $\lambda([a, b]) = \{x_0, x_1, \dots x_n\}$: $a = x_0 \le x_1 \le \dots \le x_n = b$. $\Delta x_i = x_i - x_{i-1} \ge 0$. На каждом интервале выберем произвольную точку $t_i \in [x_i, x_{i-1}]$; множество $\{t_i\}$ называется оснащением разбиения λ. Обозначим $\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}) \ge 0$, $\sum_{i=1}^n \Delta\alpha_i = \alpha(b) - \alpha(a)$.

Определим теперь интеграл от некоторой ограниченной функции $f : [a, b] \to \mathbb{R}$ относительно неубывающей функции α через вернюю и нижнюю суммы Дарбу-Стильтьеса:

$$m_i = \inf_{x\in[x_i,x_{i-1}]} f(x), \quad M_i = \sup_{x\in[x_i,x_{i-1}]} f(x), \quad L(f, \alpha, \lambda) = \sum_{i=1}^n m_i\Delta\alpha_i, \quad U(f, \alpha, \lambda) = \sum_{i=1}^n M_i\Delta\alpha_i. \tag{1}$$

Далее определяем верхний $\sup_\lambda U(f, \alpha, \lambda)$ и нижний $\inf_\lambda L(f, \alpha, \lambda)$ интегралы Римана-Стильтьеса. Функция f интегрируема по Риману-Стильтьесу относительно неубывающей функции α, если соответствующие верхний и нижний интегралы совпадают.

§2.Свойства и вычисление определённого интеграла.

§3.Интегрирование векторнозначных функций. §4.Функции ограниченной вариации. §5.Дальнейшие обобщения и свойства определённого интеграла. §6.Длина пути. §7.Несобственные интегралы.

7 Линейные операторы и полилинейные формы (тензоры).

§1.Линейные отображения и операторы. §2.Собственные векторы и характеристический многочлен. §3.Структура линейного оператора в векторном пространстве на поле комплексных чисел. §4.Линейные операторы в унитарных (евклидовых) пространствах. §5.Билинейные и квадратичные формы. §6.Гиперповерхности второго порядка в евклидовом точечном пространстве. §7.Кривые и поверхности второго порядка.

§8.Алгебраическая теория тензоров.

7.1 Полилинейные формы

V – некоторое (любое) линейное пространство, а V^* -сопряжённое ему линейное пространство. (Мы знаем, что $V^{**} \equiv V$.)

Определение: Полилинейной формой валентности (p, q) называется отображение

$$T : \underbrace{V \times \ldots \times V}_{p \text{ раз}} \times \underbrace{V^* \times \ldots \times V^*}_{q \text{ раз}} \equiv V^p \times V^{*q} \to \mathbb{R}, \tag{2}$$

линейное по каждому (векторному) аргументу; p называется ковариантной, а q – контрвариантной валентностями. Множество всех полилинейных форм данной валентности обозначим $\mathbb{T}^{(p,q)}$.

Примеры: ковариантные векторы $\mathbb{T}^{(1,0)}$, (обычные) контрвариантные векторы $\mathbb{T}^{(0,1)}$, билинейные формы $\mathbb{T}^{(2,0)}$, линейные операторы $\mathbb{T}^{(1,1)}$, скаляры $\mathbb{T}^{(0,0)}$.

Определение: тензорное произведение полилинейные форм… Замечание: тензорное произведение некоммутативно, но ассоциативно и дистрибутивно.

Литература: [26].

7.2 Кососимметрические тензоры или внешние формы

В этом вопросе будем рассматривать только тензоры из $\mathbb{T}^{(p,0)} \equiv \mathbb{T}^p$. Определим действие подстановки $\sigma \in S_p$ на тензор:

$$(\sigma T)(\vec{x}_1 \ldots \vec{x}_p) \stackrel{\text{df}}{=} T(\vec{x}_{\sigma(1)} \ldots \vec{x}_{\sigma(p)}). \tag{3}$$

– для тензоров из $\mathbb{T}^{(p,q)}$ требовалось бы использовать две различных подстановки.

Определения:

1. полилинейная форма $T \in \mathbb{T}^p$ называется кососимметрическим тензором или внешней формпй если $\forall \sigma \in S_p\ \sigma T = \operatorname{sgn}(\sigma) \cdot T$. Обозначение: $\mathbb{T}^p_a$.
2. полилинейная форма $T \in \mathbb{T}^p$ называется симметрической, если $\forall \sigma \in S_p\ \sigma T = T$ – нам они неинтересны.

Примеры:

1. любой ковектор $\mathbb{T}^{(1,0)}$ одновременно симметричен и кососимметричен.
2. симметричные билинейные формы $\mathbb{T}^{(2,0)}$.
3. для $V = \mathbb{R}^3$ определим смешанное произведение $\Omega \in \mathbb{T}^{(3,0)}$: $\forall \vec{x}, \vec{y}\vec{z} \in V\ \Omega(vecx, \vec{y}\vec{z}) \equiv (vecx, \vec{y}\vec{z}) = \det(X, Y, Z)$, где (X, Y, Z) – матрица, составленная из строк (или столбцов) – координат векторов $\vec{x}, \vec{y}\vec{z}$. Ω называется также дискриминантным тензором Леви-Чивиты, формой объёма и тензором ориентации.

Теорема 7.1 *$n = \dim V$ $T \in \mathbb{T}^p_a$; Если векторы $\vec{x}_1 \ldots \vec{x}_p$ – линейно зависимы, то $T(\vec{x}_1 \ldots \vec{x}_p) = 0$. (Следовательно, при $p > n$ $\forall T \in \mathbb{T}^p_a$ $T = 0$.)*

7.3 Альтернирование и его свойства

Рассмотрим некоторую подгруппу $H \subset S_p$.

Определение 7.2

$$T \in \mathbb{T}^p, \quad \operatorname*{Alt}_H T \stackrel{\text{df}}{=} \frac{1}{|H|} \sum_{\sigma \in H} \operatorname{sgn}(\sigma) \cdot \sigma T, \quad \operatorname*{Sym}_H T \stackrel{\text{df}}{=} \frac{1}{|H|} \sum_{\sigma \in H} \sigma T. \tag{4}$$

– полилинейная форма, полученная из полилинейной формы T альтернированием/симметризацией по H. Обозначение: $\operatorname{Alt}_{S_p} \equiv \operatorname{Alt}$.

7.4 Внешнее умножение полилинейных форм

Определение 7.3

$$T \in \mathbb{T}_a^p, \quad S \in \mathbb{T}_a^q \quad \Longrightarrow T \wedge S \overset{\mathrm{df}}{=} \frac{(p+q)!}{p!q!} \operatorname{Alt}(T \otimes S) \in \mathbb{T}_a^{p+q}. \tag{5}$$

Свойства:

1. $$ {}^1T \wedge {}^2T \wedge \ldots \wedge {}^kT \overset{\mathrm{df}}{=} \frac{(p_1+\ldots+p_k)!}{p_1!\ldots p_k!} \operatorname{Alt}({}^1T \otimes \ldots \otimes {}^k T). \tag{6}$$

2. кососкоммутативность: $T \wedge S = (-1)^{pq} S \wedge T$.

7.5 Базис в пространстве полилинейных форм

Теорема 7.4 *Если $\vec{e}_1 \ldots \vec{e}_n$ и $\vec{e}^{\,1} \ldots \vec{e}^{\,n}$ – сопряжённая пара базисов соответственно в V и V^*, то всевозможные внешние произведения $\vec{e}^{\,i_1} \wedge \ldots \wedge \vec{e}^{\,i_p}$ при $i_1 < i_2 < \ldots < i_p$ образуют базис в $\mathbb{T}_a^p$.*

Следствия и примеры:

1. координаты тензора обладают симметрией при перестановке индексов: $T_{j_1 \ldots j_p} = \operatorname{sgn}(\sigma) T_{j_{\sigma(1)} \ldots j_{\sigma(p)}}$ – равно нулю, если хотя бы два индекса совпадают.

2. $\dim \mathbb{T}_a^p = C_n^p$ при $0 \le p \le n$; при $p > n$ $\dim \mathbb{T}_a^p = 0$.

3. $\omega^1 \wedge \ldots \wedge \omega^p \in \mathbb{T}_a^1 \equiv \mathbb{T}_a$, $\omega^1 \wedge \ldots \wedge \omega^p = 0 \Longleftrightarrow \omega^1 \wedge \ldots \wedge \omega^p$ линейно зависимы.

4. $(\vec{e}^{\,1} \ldots \vec{e}^{\,\mathrm{P}})(\vec{x}_1 \ldots \vec{x}_P) = \det(\langle \vec{e}^{\,i}, \vec{x}_j \rangle)$, $1 \le i, j \le p$.

§9.Тензоры в псевдоевклидовом пространстве.

7.6 Электромагнитное поле

См. стр. 184 из дубровина. Имеем псевдоевклидово пространство и кососимметрическое тензорное поле F^{ik} в нём.

$$F = E_\alpha \mathrm{d}x^0 \wedge \mathrm{d}x^\alpha - H^1 \mathrm{d}x^2 \wedge \mathrm{d}x^3 - H^2 \mathrm{d}x^3 \wedge \mathrm{d}x^1 - H^3 \mathrm{d}x^1 \wedge \mathrm{d}x^2. \tag{7}$$

Известно, что магнитное поле – неполноценный (псевдо)вектор, но про E я как-точки привык думать, что это всё ж таки вектор. А вот и нет! Вектором электрическое поле является лишь в трёхмерном пространстве, но не в нашем псевдоевклидовом! И, кстати, уж не оператотом ли внешнего дифференцирования получается F^{ik} из четырёх-потенциала? Когда-то уравнения Максвелла записывались покомпонентно и выглядели ужасно сложно. Потом люди стали использовать операторы градиента, ротора и дивергенции, и уравнений Максвелла стало всего лишь четыре. Наконец, если использовать внешнюю форму (7), то уравнений станет всего два, причём написаны они будут в универсальном виде, позволяющем легко переходить в произвольную инерциальную систему отсчёта, а также неевклидово пространство.

А ещё похожий антисимметричный тензор возникает в теориии упругости. И вроде как в гидродинамике тоже что-точки такое есть. Как тензор Белинфанте соотносится с тензором энергии-импульса?

Описанные примеры наиболее просты в том смысле, что здесь пространство представляет собой истинное пространство-время. В нерелятивистской квантовой механике в роли пространства будет выступать (вообще говоря, бесконечномерное) функциональное пространство состояний. Хотя, насколько я помню, при изучении фазы Бери в роли пространства тоже выступает обычное трёхмерное пространство.

Понятие “горизонтальности” объяснено у Дубровина на стр.600-601.

8 Числовые и функциональные ряды

§1.Элементарные свойства числовых рядов. §2.Ряды с неотрицательными членами. §3.Абсолютно и условно сходящиеся ряды.

§4.Равномерная сходимость функциональных последовательностей.

8.1 Равномерная сходимость функциональных последовательностей

Рассмотрим на некотором произвольном множестве X функциональную последовательность, т.е. набор функций $f_n : X \to \mathbb{C}$.

Определения:

1. Функциональная последовательность f_n сходится *поточечно* на X к функции f, если $\forall x \in X\ f_n(x) \underset{n\to\infty}{\longrightarrow} f(x)$; аналогично
2. функциональный ряд сходится $\sum\limits_{n\geq 1} u_n(x) = s(x)$, если сходится последовательность его частичных сумм: $s_n(x) \equiv \sum\limits_{i=1}^{n} u_i(x) \underset{n\to\infty}{\longrightarrow}$
3. Последовательность функций f_n сходится *равномерно* на X к функции f, если $\forall \varepsilon > 0\ \exists N : \forall n > N$ и сразу при $\forall x \in X$ $|f_n(x) - f(x)| < \varepsilon$. Обозначение: $f_n \underset{X}{\rightrightarrows} f$.

8.2 Признак Дини

Рассмотрим[3] $X = [a, b]$; допустим, $f_n \in \mathbb{C}$ и $f_n \in \mathbb{C}$.

Теорема 8.1 *Если* $f_n \underset{X}{\searrow} f \implies f_n \underset{X}{\rightrightarrows} f$, $n \to \infty$.

Доказательство: ...

§5.Свойства равномерно сходящихся последовательностей и рядов.

§6.Степенные ряды.

9 Метрические пространства

§1.Метрическое пространство.

9.1 Метрическое пространство

Метрическим пространством (МП) называется пара (X, ρ), где X – некоторое множество, а ρ – функция $\rho : X \times X \to \mathbb{R}$, удовлетворяющая аксиомам

1. $\forall x, y \in X\ \rho(x, y) \geq 0$; $\rho(x, y) = 0 \iff x = y$.
2. $\rho(x, y) = \rho(y, x)$
3. неравенство треугольника $\forall x, y, z \in X\ \rho(x, z) \leq \rho(x, y) + \rho(y, z)$

Примеры МП:

1. все нормированные пространства: по определению $\rho(\vec{x}, \vec{y}) = ||\vec{y} - \vec{x}||$. Подпримеры: а) $\mathbb{R}^n$: $\forall \vec{x} \in \mathbb{R}^n\ ||\vec{x}|| = \sqrt{\sum_{i=1}^n x_i^2}$. б) $\mathbb{C}^n$: $\forall \vec{z} \in \mathbb{C}^n\ ||\vec{z}|| = \sqrt{\sum_{i=1}^n |z|_i^2}$. в) Множество (обозначим его l_2) всех последовательностей $\{x_n\}_{n\geq 1}$ таких, что $\sum_{i=1}^\infty x_i^2 < \infty$. Докажем, что это множество образует линейное пространство. Для любой пары элементов $\vec{x}, \vec{y} \in l_2$ определим их произвольную линейную комбинацию, то есть умножение на скаляр (это очевидно) и сумму. Доказажем, что $\vec{x} + \vec{y} \in l_2$: $\vec{x} + \vec{y} = \{x_n + y_n\}_{n\geq 1}$ пока пропускаю доказательство и другие примеры

9.2 Предел последовательности в метрическом пространстве

9.3 Предельные точки множеств в метрическом пространстве

9.4 Открытые и замкнутые множества в метрическом пространстве

...

Определения:

1. Система τ подмножеств произвольного множества X называется топологией на множестве X, если она удовлетворяет условиям:

[3] Рассуждения останутся верными и в том случае, если X – компактное метрическое пространство.

(a) $X, \emptyset \in \tau$,

(b) объединение любого семейства множеств из τ принадлежит τ,

(c) пересечение любого конечного семейства множеств из τ принадлежит τ.

2. Пара (X, τ) называется топологическим пространством, если на X задана топология τ.

3. Любое множество из τ называется открытым.

4. Множество $V \subset X$ называется замкнутым в данной топологии τ, если $X/V \in \tau$.

9.5 Расположение точки относительно множества

Пусть (X, τ) – топологическое пространство.

Определения:

- Внутренностью ($\operatorname{Int} A \equiv \mathring{A}$) множества $A \subset X$ называется наибольшее открытое множество, содержащееся в A: $\mathring{A} = \cup_{U \in \tau, U \subset A} U$.
- Замыканием ($\operatorname{Cl} A \equiv \bar{A}$) называется[4] наименьшее закрытое множество, содержащее A.
- Границей называется $\operatorname{Fr} A \equiv \partial A = \operatorname{Cl} A \setminus \operatorname{Int} A$.
- Топологическое пространство (X, τ) называется хаусдорфовым, если $\forall x, y \in X : x \neq y \ \exists U(x), U(y) : U(x) \cap U(y) = \emptyset$. Если пространство хаусдорфово, точки предел последовательности в нём единственен.

9.6 Полные МП

§2. Отображения МП

9.7 Непрерывные отображения (НО)

Пусть (X, τ_1) и (Y, τ_2) – топологические пространства, $A \subset X$ и $f : A \to Y$.

Определение 9.1 *Функция f называется непрерывной в точке $a \in X$,*

$$\forall U(f(a)) \subset Y \quad \exists V(a) \in X : \quad f(V(a) \cap A) \subset U(f(a)), \tag{8}$$

где использованы (обычные в дальнейшем) обозначения окрестностей точек символами U и V.

§3. Связность

9.8 Связные множества (СМ)

Связность – математическая интерпретация целого куска.

Определение 9.2 *пусть (X, τ) и – топологическое пространство; оно называется связным, если его нельзя разбить на два непустых открытых множества.*

Следствие: связное пространство нельзя также разбить и на два непустых закрытых множества.
Связность индуцируется на подмножества вместе с топологией.

Свойства СМ:

- Замыкание СМ связно.
- Объединение любого семейства СМ, содержащих общую точку, связно.
- Образ СМ при НО связан.

Определение 9.3 *областью в $\mathbb{R}^n$ называется открытое связное множество.*

§3. Компактность

[4] Обозначение неуниверсальное; напр., кое-где в этом тексте надчёркивание означает дополнение подмножества (т.е. часть пространства за исключением этого подмножества).

9.9 Компактные множества (КМ)

Определение 9.4 *Система множеств I называется открытым покрытием ТП (X, τ), если $X = \bigcup\limits_{\alpha \in I} U_\alpha$ $U_\alpha \in \tau$.*

Определение 9.5 *ТП (X, τ) называется компактным, если из всякого открытого покрытия можно извлечь конечное подпокрытие.*

Компактность индуцируется на подмножества вместе с топологией.

Свойства КМ:

- Замкнутое подмножество компакта компактно.
- Непрерывный образ компакта компактен.
- Подмножество в $\mathbb{R}^n$ или в $\mathbb{C}^n$ компактно, если оно замкнуто и ограничено (верно в обе стороны).

§2.Отображения метрических пространств.

§3.Связность. §4.Компактность. §5.Принцип сжимающихся отображений.

10 Дифференциальное исчисление

§1.Нормированные пространства. §2.Частные производные и дифференциалы высшего порядка. §4.Теорема об обратной функции. §5.Приложения теоремы об обратной функции. §6.Финитные функции и разложение единицы.

11 Гладкие многообразия

Аппарат многообразий полезен в случае, если мы рассматриваем нечто более сложное, чем просто область в евклидовом пространстве. Кстати, вопрос на понимание: является ли (гипер)поверхность в $\mathbb{R}^3$ областью?

§1.Гладкие многообразия.

11.1 Основные определения

Определение: Гладкой структурой класса $\mathbb{C}^r$ (или гладкостью) называется произвольный максимальный атлас класса $\mathbb{C}^r$ на X.

11.2 Гладкие отображения

Дано: два гладких многообразия X и Y, имеющие классы гладкости r_x и r_y. Допустим, имеется отображение $f : X \to Y$; обозначим $r = \min(r_x, r_y)$. Вставить рисуночек. Предположим, что функция f непрерывна[5] в точке $M_0 \in X$. Выберем произвольную карту $(V, \psi) = (V, y^1 \dots y^m)$ в окрестности точки $f(M_0)$. Из непрерывности f следует существование окрестности $U = U(M_0) : f(U) \subset V$. Обозначим соответствующую карту (U, φ).

Определение сквозного отображения: $\mathring{f} \overset{\mathrm{df}}{=} \psi \circ f\Big|_U \circ \varphi^{-1}$. О нём говорят, что оно задаёт (или определяет) отображение f в паре карт (U, φ) и (V, ψ). Переход от f к $\mathring{f}$ означает, что мы записали отображение f в локальных координатах.

§2. Касательное пространство

11.3 Касательное пространство

Наводящие соображения: допустим, имеется гладкая[6] кривая $\vec{x} = \vec{x}(t)$ с координатами $x^i(t)$, $i = 1 \dots n$. Рассмотрим касательный вектор $\vec{v}$ в точке $M_0 \equiv \vec{x}_0 = \vec{x}(t_0)$: $\vec{v} = \dot{\vec{x}}(t_0) \equiv [\dot{x}^1(t_0) \dots \dot{x}^n(t_0)]^{\mathrm{T}}$. Далее, предполагаем, что в некоторой окрестности точки $\vec{x}_0$ существует другая система координат $(\grave{x}^1 \dots \grave{x}^n) \equiv \grave{\vec{x}}$. Посмотрим, как будет выглядеть касательный вектор в новых (штрихованных) координатах: $\grave{\vec{v}}^j = \dfrac{\partial \grave{x}^j}{\partial x^i}(M_0) \cdot \dfrac{\mathrm{d}x^i}{\mathrm{d}t} \dots$

Определение: касательным вектором к гладкому многообразию X в точке M_0 называется такое отображение $\vec{v} : \mathcal{A}(M_0) \to \mathbb{R}^n$, что для любых двух карт (U, φ) и $('U, '\varphi)$, центрированных в точке M_0, $\vec{v}(\grave{U}, \grave{\varphi}) = \left(\dfrac{\partial \grave{\varphi}}{\partial \varphi}\right)_0 \vec{v}(U, \varphi)$

[5] в смысле топологии многообразия.

[6] в смысле существования производной, а не римановской геометрии.

Определение 11.1 *Множество всех касательных векторов в точке M_0 называется касательным пространством к гладкому многообразию X в точке M_0, и обозначается как $\mathbb{T}_{M_0}X$.*

Докажем, что $\dim \mathbb{T}_{M_0}X = n\ldots$

Соглашения и отождествления: имеет место изоморфизм $\mathbb{T}_{M_0}X \longleftrightarrow \mathbb{R}^n$; карта (U,φ) определяет систему координат. Касательные вектора, переходящие при этом изоморфизме в стандартный базис $\vec{e}_1\ldots\vec{e}_n$ в $\mathbb{R}^n$, принято обозначать $\left(\dfrac{\partial}{\partial x^1}\right)_0 \ldots \left(\dfrac{\partial}{\partial x^n}\right)_0$. Таким образом, один и тот же касательный вектор $\vec{v}(U,\varphi)$ можно обозначать двумя способами: как $\vec{v} = (v^1\ldots v^n)^{\mathrm{T}}$ и как $\vec{v} = v^i\left(\dfrac{\partial}{\partial x^n}\right)_0$.

11.4 Дифференциал гладкого отображения

Допустим, X и Y – гладкие многообразия.

Определения:

1. Непрерывное отображение $f : X \to Y$ называется гладким, если оно является гладким в любой точке из X.
2. Отображение $f : X \to Y$ называется диффеоморфизмом гладкого многообразия X на гладкое многообразие Y, если оно взаимооднозначно и отображения f и f^{-1} – гладкие.

Следуя секции 11.2, построим для f сквозное отображение $\mathring{f}$. Теперь сопоставим произвольному касательному вектору $\vec{v} \equiv \vec{v}(U,\varphi) \in \mathbb{T}_{M_0}X$ касательный вектор $\vec{w} \in \mathbb{T}_{f(M_0)}Y$ следующим образом: $\vec{w} = \mathring{f}'(M_0)\vec{v}$. Кстати, интересен частный случай $X \equiv Y$, когда f является отображением замены координат. В этом случае уместно обозначать его матрицу Якоби (якобиан) $f'(M_0) = \left(\frac{\partial y^i}{\partial x^j}\right)_{M_0} \implies w^i = \frac{\partial y^i}{\partial x^j}v^j$. Доказываем, что $\vec{w}$ преобразуется, как полагается касательному вектору…

Определение 11.2 *таким образом, мы получили линейный оператор $\mathrm{d}f_{M_0} : \mathbb{T}_{M_0}X \to \mathbb{T}_{f(M_0)}Y$, называемый дифференциалом отображения f в точке M_0.*

Если $\dim \mathbb{T}_{M_0}X = n$ и $\dim \mathbb{T}_{f(M_0)}Y = m$, то эти (линейные) касательные пространства можно отождествить с $\mathbb{R}^n$ и $\mathbb{R}^m$: $\mathbb{T}_{M_0}X \equiv \mathbb{R}^n$ и $\mathbb{T}_{f(M_0)}Y = \mathbb{R}^m$.

Пример: базис кокасательного пространства Раз уж у нас есть линейное пространство $V = \mathbb{T}_{M_0}X$, давайте рассмотрим сопряжённое ему пространство $V^* = \mathbb{T}^*_{M_0}X$. Базис V^* получается из дифференциала $\mathrm{d}f_{M_0}$ при $f(\vec{x}\,) = x^i$ и обозначается как $\mathrm{d}x^1\ldots\mathrm{d}x^n$. Из $\mathrm{d}x^i_{M_0}\left(\dfrac{\partial}{\partial x^j}\right)_{M_0} = \delta^i_j$ следует, что дифференциалы координатных функций $\mathrm{d}x^1\ldots\mathrm{d}x^n$ образуют базис $\mathbb{T}^*_{M_0}X$, сопряжённый к базису $\left(\dfrac{\partial}{\partial x^1}\right)_{M_0}\ldots\left(\dfrac{\partial}{\partial x^n}\right)_{M_0}$ пространства $\mathbb{T}_{M_0}X$.

§3. Условные экстремумы

§4.Финитные функции и разложение единицы.

11.5 Финитные функции

Рассмотрим произвольное топологическое пространство X и функцию $f : X \to \mathbb{R}$.

Определение 11.3 *Носителем функции $f : X \to \mathbb{C}$ называется* $\operatorname{Supp} f \stackrel{\mathrm{df}}{=} \mathrm{Cl}\{x \in X | f(x) \neq 0\}$.

Определение 11.4 *Непрерывная функция с компактным носителем называется* ***финитной****.*

Обозначим $\mathcal{K}(X)$ – множество финитных функций на X.

Предположим теперь, что X – гладкое многообразие, и $K \subset X$. Семейство функций $\{f_\alpha\}_{\alpha\in A}$ $f_\alpha : X \to \mathbb{R}$ называется **разложением единицы** на K, если

1. $\forall\alpha \in A$ $f_\alpha \in \mathbb{C}^{(2)}(X)$,
2. $\forall x \in X$, $\forall\alpha \in A$ $0 \le f_\alpha(x) \le 1$,
3. $\forall x \in K$ $\exists$ окрестность $U(x)$, в которой только конечное число функций $\{f_\alpha\}$ отлично от нуля,

4. $\forall x \in K \sum\limits_{\alpha \in A} f_\alpha(x) = 1$ (из предыдущего свойства следует, что эта сумма конечна).

Определение 11.5 *допустим, $\{U_i\}$ – открытое покрытие*[7] *K; разложение единицы $\{f_\alpha\}_{\alpha\in A}$ на K называется подчинённым $\{U_i\}$, если $\forall \alpha \in A\ \exists i(\alpha) \in I : \operatorname{Supp} f_\alpha \subset U_{i(\alpha)}$.*

Теорема 11.6 *Допустим, K – компакт на гладком многообразии*[8] *X и $\{U_i\}$ – открытое покрытие K. Тогда существует разложение единицы $\{f_\alpha\}_{\alpha\in A}$, подчинённое покрытию $\{U_i\}$ и состоящее из гладких финитных*[9] *функций. Более того, это будет конечное семейство функций $f_1 \dots f_p$ и $\forall x \in X\ 0 \le \sum\limits_{\alpha=1}^{p} f_\alpha(x) \le 1$.*

Доказательство: ...

12 Введение в общую теорию интегралов

12.1 Элементарные функции и элементарный интеграл

Определение 12.1 ***элементарных функций*** *и решётки Рисса H: рассмотрим некоторое (произвольное) множество X; предположим, что на X задано семейство H вещественнозначных ограниченных функций, удовлетворяющее условиям:*

1. *H – вещественное линейное пространство (в общем случае – бесконечномерное),*
2. *$h \in H \Longrightarrow |h| \in H$.*

Следствия:

$$\begin{gathered}\forall h, g \in H \max(h,g) = \frac{1}{2}(h+g-|h-g|) \in H, \quad \min(h,g) = \frac{1}{2}(h+g-|h-g|) \in H \\ \Longrightarrow h^+ \overset{\text{df}}{=} \max(h,0) \in H, \quad h^- \overset{\text{df}}{=} \max(-h,0) \in H, \quad h^\pm \ge 0, \quad h = h^+ - h^-, \quad |h| = h^+ + h^-.\end{gathered} \tag{9}$$

Определение 12.2 *Элементарным интегралом на пространстве функций H называется отображение $I : H \to \mathbb{R}$, удовлетворяющее трём аксиомам:*

1. *I – линейный функционал на H, т.е. $\forall \alpha_i \in \mathbb{R}, \forall h_i \in H\ (i = 1,2)\ I(\alpha_1 h_1 + \alpha_2 h_2) = \alpha_1 I h_1 + \alpha_2 I h_2$,*
2. *аксиома неотрицательности: $h \in H$, $h \ge 0 \Longrightarrow Ih \ge 0$,*
3. *рассмотрим последовательность функций $h_n \in H$ $h_n \underset{X}{\searrow} 0 \Longrightarrow Ih_n \to 0$.*

Следствие:

$$-|h| \le h \le |h| \Longrightarrow -I|h| \le Ih \le I|h| \Longrightarrow |Ih| \le I|h|. \tag{10}$$

Пример: вспоминая былое (см. сек. 6.1), рассмотрим $X = [a,b]$, $\alpha \underset{X}{\nearrow}$ – вещественнозначная функция; В качестве H возьмём $H = \mathbb{C}([a,b])$. $h \in \mathbb{C} \Longrightarrow \exists \lim\limits_{x\to x_0} h(x) = h(x_0) \Longrightarrow$ Определим интеграл Стильтьеса $Ih \overset{\text{df}}{=} \int_a^b h(x)\mathrm{d}\alpha(x)$. Очевидно выполнение всех условий, кроме 3, которое доказываем: допустим $h_n \in \mathbb{C}([a,b])$ и $h_n \underset{[a,b]}{\searrow} 0$. По теореме Дини (см. сек. 8.2) $h_n \underset{[a,b]}{\rightrightarrows} 0$. По теореме о предельном переходе под интегралом, стр.IV70 $Ih_n = \int_a^b h(x)\mathrm{d}\alpha(x) \to \int_a^b 0\mathrm{d}\alpha(x) = 0 \Longrightarrow 3$. условие выполнено $\Longrightarrow$ интеграл Стильтьеса является элементарным.

[7] определение этого понятия – рядом с определением компактности.

[8] Наши многообразия всегда хаусдорфовы.

[9] Существенно, что эти функции будут гладкими и финитными не только на K, но и на всём X.

12.2 ЭИ на пространствах финитных функций

Пусть X – многообразие, а $H = \mathcal{K}(X)$ – пространство финитных функций. Очевидно, H является линейным пространством вещественнозначных функций. Модуль финитной функции, очевидно, финитен: $\forall h \in H\ |h| \in H$. Получается, что H удовлетворяет аксиомам элементарного интеграла 12.1.

Важнейший пример:

$$X = \mathbb{R}, \quad Ih = \int_{-\infty}^{\infty} h(x)\mathrm{d}x = \int_a^b h(x)\mathrm{d}x, \quad \operatorname{Supp} h \subset [a,b]. \tag{11}$$

Определим норму $||h|| = \sup\limits_{x\in\operatorname{Supp} h} |h(x)| = \sup\limits_{x\in X} |h(x)|$ – очевидно, удовлетворяет аксиомам нормы.[10]

Рассмотрим линейный неотрицательный функционал I на компакте $K \equiv K(X)$; установим его свойства:

1. $\forall K \subset X$: если K – компакт, то $\exists C_K$ =const: $\forall h \in \mathcal{K}$ такой что $\operatorname{Supp} h \subset K \implies |I(h)| \leq C_k ||h||$. Доказательство: по следствию на стр.VI56 $\exists f \in \mathcal{K} : f\Big|_K = 1$ и $f(X) \subset [0,1]$. Допустим $\operatorname{Supp} h \subset K \implies \forall x \in X\ |h(x)| \leq ||h|| \cdot f(x)$ (или, жаргонно, $|h| \leq f \cdot ||h||$). Следовательно, $|I(h)| \leq I(|h|) \leq I(f \cdot ||h||) = ||h|| \cdot I(f)$, так что искомая постоянная есть $C_K = I(f)$.

2. Допустим имеется последовательность функций $h_m \in \mathcal{K}$ такая, что а) существует такой компакт K, что $\forall m \operatorname{Supp} h_m \subset K$ и б) $h_m \rightrightarrows_X h$. Тогда $h \in \mathcal{K}$ и $I(h) = \lim\limits_{m\to\infty} I(h_m)$. Доказательство: ...

3. I есть интеграл на $\mathcal{K}$. Доказательство: ...

12.3 Теорема о плотности

В этом вопросе $X = \mathbb{R}^l$, $Y = \mathbb{R}^m$, $W = X \times Y = \mathbb{R}^n$, $n = l + m$; $\mathcal{K} \equiv \mathcal{K}(W)$ – (компактное[11]) множество финитных функций 11.5. Рассм. финитную функцию $h \in \mathcal{K}$, обозначим $K = \operatorname{Supp} h$, K_x, K_y – проекции компакта K на X и Y. Произвольную точку из W будем обозначать $(\vec{x}, \vec{y})$ имея в виду, что $\vec{x} \in X$, $\vec{y} \in Y$.

Проекция W на X – отображение, переводящее $(\vec{x}, \vec{y})$ в $\vec{x}$; очевидно, это отображение непрерывно. Непрерывный образ компакта компактен, так что K_x, K_y – компактны.

Рассмотрим некоторую последовательность $(\vec{x}_n, \vec{y}_n) \in K_x \times K_y$. Учитывая компактность K_x, из последовательности $\vec{x}_n$ мы можем извлечь подпоследовательность $\vec{x}_{n_k}$, сходящуюся к некоторой точке $\tilde{\vec{x}} \in X$. Аналогично и из $\vec{y}_n$ извлекаем подпоследовательность $\vec{y}_{n_{k_l}} \xrightarrow[l\to\infty]{} \tilde{\vec{y}} \in K_y$. Следовательно, имеется подпоследовательность $(\vec{x}_{n_{k_l}}, \vec{y}_{n_{k_l}}) \xrightarrow[l\to\infty]{} (\tilde{\vec{x}}, \tilde{\vec{y}}) \in K_x \times K_y$.

С использованием этих обозначений будет доказана

Теорема 12.3 *финитная функция на произведении множеств сколь угодно хорошо аппроксимируется линейными комбинациями произведений компактных (на производящих множествах) функций. А именно, ...*

$$\forall\varepsilon > 0 \exists g_\varepsilon(\vec{x}, \vec{y}) = \sum_{i=1}^{p}\sum_{j=1}^{q} c_{ij}\varphi_i(\vec{x})\psi_j(\vec{y}) \tag{12}$$

Доказательство:...

12.4 Теорема Фубини для элементарного интеграла

В этом вопросе также предполагаем, что $X = \mathbb{R}^l$ $Y = \mathbb{R}^m$ $W = X \times Y = \mathbb{R}^n$ $n = l + m$, и также рассм. $\mathcal{K} \equiv \mathcal{K}(W)$ – (компактное) множество финитных функций 11.5.

Кроме того, рассмотрим элементарные интегралы I_X и I_Y на $\mathcal{K}(X)$ и $\mathcal{K}(Y)$ соответственно.

Теорема 12.4 *о прямом произведении интегралов: существуют "кратные" интегралы $I_X I_Y h = I_Y I_X h$; кроме того, $I = I_X \circ I_Y$ есть ...функционал... на K.*

[10] Аксиомы нормы: 1. $||x|| \geq 0$, 2. $||x|| = 0 \iff x = 0$, 3.$||x + y|| \leq ||x|| + ||y||$.

[11] Компактность следует из признака Больцамо-Вееритрасса.

12.5 Теорема о замене переменной в элементарном интеграле Римана

Пусть $E, G \subset \mathbb{R}^n$ – открытые подмножества, а $\Phi : E \to G$ – диффеоморфизм. Тогда для любой функции $h \in \mathcal{K}(G)$ справедлива формула замены переменной в кратном элементарном интеграле

$$\int_G h(\vec{x})\mathrm{d}\mu(\vec{x}) = \int_E \underbrace{h(\Phi(\vec{t})) \cdot |\det \Phi'(\vec{t})|}_{\in\mathcal{K}(E)} \cdot \mathrm{d}\mu(\vec{t}). \tag{13}$$

Далее следует длинющее доказательство в 4 этапа...

Замечания, которые не входят ни в один вопрос:

1. Элементарный интеграл Римана совершенно недостаточен даже для практических занятий; он будет продолжен на более широкий класс функций, чем финитные функции с компактным носителем. Авансом укажем некоторые из этих функций: Рассмотрим компакт $K \subset \mathbb{R}^n$. Функцию f, определённую на K, можно тривиально продолжить нулевым значением на всё $\mathbb{R}^n$. Продолженная функция $\bar{f}$ будет суммируема по Лебегу на $\mathbb{R}^n$. (При этом она не будет суммируема по Риману.) Интеграл по Лебегу от $\bar{f}$ определяется как
$$I\bar{f} = \int_{\mathbb{R}^n} \bar{f}(\vec{x})\mathrm{d}\mu(\vec{x}) \stackrel{\mathrm{df}}{=} \int_K f(\vec{x})\mathrm{d}\mu(\vec{x}), \tag{14}$$
причём остаются справедливыми теоремы а) Фубини и б) о замене переменной.

2. Для любого подмножества $K \subset X$ введём характеристическую функцию $\chi_K : X \to \mathbb{R}$
$$\chi_K(x) = \begin{cases} 1, & x \in K, \\ 0, & x \notin K. \end{cases} \tag{15}$$
Для $K \subset \mathbb{R}^n$ эта функция будет суммируема по Лебегу, а её интеграл называется мерой Лебега или объёмом множества.

§2. Тензорные поля и дифференциальные формы

12.6 Тензорные поля

Рассмотрим гладкое многообразие X класса $\mathbb{C}^2$ размерности $\dim X = n$. Рассмотрим некотоую точку $M \in X$ и касательное пространство $\mathbb{T}_M X$ в этой точке; обозначим $V_M \equiv \mathbb{T}_M X$. V_M представляет собой n-мерное линейное пространство над $\mathbb{R}$. Определим пространство $\mathbb{T}_M^{(p,q)} X \equiv \mathbb{T}^{(p,q)} V_M$ над линейным пространством V_M.

Определение 12.5 *Тензорным полем S типа (p, q) называется соответствие S, сопоставляющее каждой точке $M \in X$ тензор $S(M) \equiv S_M \in \mathbb{T}_M^{(p,q)} X$.*

Рассмотрим некоторую карту $(U, \varphi) = (U, x^1 \dots x^n)$ в точке M. Как мы уже знаем, всевозможные произведения

$$\left(\mathrm{d}x^{i_1}\right)_M \otimes \ldots \otimes \left(\mathrm{d}x^{i_p}\right)_M \otimes \left(\frac{\partial}{\partial x^{j_1}}\right)_M \otimes \ldots \otimes \left(\frac{\partial}{\partial x^{j_q}}\right)_M \tag{16}$$

образуют базис в $\mathbb{T}_M^{(p,q)} X$. Разлагая S по этому базису, мы получаем n^{p+q} штук функций $S_{i_1 \dots i_p}^{j_1 \dots j_q} : U \to \mathbb{R}$, называеных компонентами тензорного поля S в карте (U, φ). При выборе альтернативной (штрихованой) карты $('U, '\varphi)$ в этой же точке M возникает альтернативный (штрихованый) базис в $\mathbb{T}_M^{(p,q)} X$. Координаты тензорного поля в этих двух базисах связаны соотношением

$$M \in U \cap {'U}, \quad {'S}_{l_1 \dots l_p}^{k_1 \dots k_q} = \left(\frac{\partial x^{i_1}}{\partial {'x}^{l_1}}\right)_M \cdots \left(\frac{\partial x^{i_p}}{\partial {'x}^{l_p}}\right)_M \left(\frac{\partial {'x}^{k_1}}{\partial x^{j_1}}\right)_M \cdots \left(\frac{\partial {'x}^{k_q}}{\partial x^{j_q}}\right)_M S_{i_1 \dots i_p}^{j_1 \dots j_q}. \tag{17}$$

В физике (17) является определением тензорного поля.

Определение 12.6 *Говорят, что тензорное поле S принадлежит классу гладкости $\mathbb{C}^{(m)}$ (при $0 \le m \le r-1$, где r – класс гладкости многообразия),если все функции $S_{i_1 \dots i_p}^{j_1 \dots j_q} \in \mathbb{C}^m$.*

Из (17) следует, что класс гладкости не меняется при замене базиса. Поля максимального класса гладкости $m = r - 1$ будем называть гладкими полями. Естественным образом определяется носитель тензорного поля: $\operatorname{Supp} S = \operatorname{Cl}\{M \in X | S(M) \ne 0\}$; непрерывное тензорное поле с компактным носителем называется финитным. Обозначим символом $\mathbb{T}^{(p,q)} X$ множество всех гладких тензорных полей на X.

Примеры:

1. $\mathbb{T}^{(0,0)}X$ – пространство гладких функций; $\forall f \in \mathbb{T}^{(0,0)}X$, $\forall T \in \mathbb{T}^{(p,q)}X$ $f \otimes T = fT \in \mathbb{T}^{(p,q)}X$.
2. $\mathbb{T}^{(0,1)}X$ – пространство гладких векторных полей. Любое такое поле $T \in \mathbb{T}^{(0,1)}X$ можно записать в виде $T = T^i \dfrac{\partial}{\partial x^i}$.
3. $\mathbb{T}^{(1,0)}X$ – пространство ковариантных векторных полей, или дифференциальных один-форм. Любое такое поле $\omega \in \mathbb{T}^{(1,0)}X$ можно записать в виде $\omega = \omega_i \mathrm{d}x^i$.

12.7 Псевдоримановы многообразия

Определение 12.7 *Многообразие X называется псевдоримановым, если на нём определено симметричное невырожденное гладкое тензорное поле $g \in \mathbb{T}^{(2,0)}X$.*

- Симметричность означает, что $\forall M \in X$, $\forall \vec{u}, \vec{v} \in \mathbb{T}_M X$ $g_M(\vec{u}, \vec{v}) = g_M(\vec{v}, \vec{u})$, а
- невырожденность означает, что если при любых $\forall \vec{v} \in \mathbb{T}_M X$ для некоторого $\vec{u}$ $g_M(\vec{u}, \vec{v}) = 0$, точки обязательно $\vec{u} = 0$.

Таким образом, g_M на каждом касательном пространстве задаёт структуру псевдоевклидового пространства. Обозначения $g = g_{ij}\mathrm{d}x^i \otimes \mathrm{d}x^j$. В силу симметричности $g_{ij}\mathrm{d}x^i \otimes \mathrm{d}x^j = g_{ij}\,\mathrm{Sym}_{i,j}\,\mathrm{d}x^i \otimes \mathrm{d}x^j \equiv g_{ij}\,\mathrm{Sym}_{i,j}\,\mathrm{d}x^i \mathrm{d}x^j = \mathrm{d}S^2$.

Определение 12.8 *Многообразие X называется римановым, если в любой его точке $\forall M \in X$ тензор g_M положительно определён: $\forall \vec{v} \neq 0$ $g_M(\vec{v}, \vec{v}) > 0$.*

Определение 12.9 *Многообразие X допускает разложение единицы, если для любого открытого покрытия существует разложение единицы $\{V_i\}_{i \in I}$ $\{f_\alpha\}_{\alpha \in A}$, подчинённое этому покрытию: $\forall \alpha\ \exists i(\alpha) \in I : \mathrm{Supp}\, f_\alpha \subset V_{i(\alpha)}$.*

Теорема 12.10 *Если гладкое многообразие X допускает разложение единицы, точки на нём можно ввести риманову структуру.*

…В процессе доказательства выясняется, что эта риманова структура имеет вид $g = (\mathrm{d}x^1)^2 + \ldots + (\mathrm{d}x^n)^2$…

12.8 Интегрирование плотностей

X – гладкое многообразие размерности n. Каждой карте (U, φ) в X сопоставим функцию $\rho^U : U \to \mathbb{R}$ которая при замене координат $(U, \varphi) \to (\grave{U}, \grave{\varphi})$ преобразуется следующим образом:

$$\text{на } U \cap \grave{U} \quad \rho^{\grave{U}} = \rho^U \left| \det \frac{\partial \varphi}{\partial \grave{\varphi}} \right|. \tag{18}$$

Семейство функций $\{\rho^U\}$ называется плотностью на многообразии X. Заметим, что плотность не является тензорным полем; плотности образуют линейное пространство: по определению $\rho + \Theta = \{\rho^U + \Theta^U\}$. Определим умножение плотности на функцию: $f \cdot \rho = \{f^U \cdot \rho^U\}$. Плотность ρ называется плотностью класса $\mathbb{C}^m$, $0 \le m < r$, если любая функция ρ^U принадлежит этому классу гладкости. Плотность класса $m = 0$ называется непрерывной плотностью. Говорят, что плотность положительна [равна нулю] в данной точке M, если $\rho^U(M) > 0$ [$\rho^U(M) = 0$] в какой-либо карте U. Соответственно, носителем плотности называется $\mathrm{Cl}\{M \in X | \rho(M) \neq 0\}$. Плотность, положительная во всех точках многообразия, называется положительной или объёмной (плотностью объёма); для объёмной плотности используется исторически устоявшееся обозначение $\mathrm{d}V = \{\mathrm{d}V^U\}$. Непрерывная плотность с компактным носителем называется финитной.

Псевдориманова плотность объёма. Рассмотрим псевдориманово многообразие X и псевдометрический тензор g при нём; выберем некоторую карту (U, φ). Псевдоримановой плотностью объёма называется

$$\mathrm{d}V = \sqrt{|\det g|} \equiv \sqrt{g} > 0.$$

Очевидно определение плотности (18) выполняется. Если $(U, x^1 \ldots x^n)$ – исходная карта, в которой метрический тензор имеет координаты g_{ij}, то для некоторой штрихованной системы координат $(\grave{U}, \grave{x}^1 \ldots \grave{x}^n)$ имеем на $U \cap \grave{U}$

$$\forall k, l \quad \grave{g}_{kl} = g_{ij} \frac{\partial x^i}{\partial \grave{x}^k} \frac{\partial x^j}{\partial \grave{x}^l} \quad \text{или} \quad \grave{g} = C^T g C, \quad C = \left\{ \frac{\partial x^j}{\partial \grave{x}^j} \right\}_{j,l=1}^{n} \Longrightarrow \mathrm{d}\grave{V} = \sqrt{|\det \grave{g}|} = |\det C| \sqrt{|\det g|}.$$

В римановом случае плотность имеет прозрачный геометрический смысл. Зафиксируем некоторую систему координат $(U, x^1 \ldots x^n)$. Значения метрического тензора на базисных осях в касательном пространстве

$$g_{ij} = g\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j} \right), \quad \left(\frac{\partial}{\partial x^i} \right)_M, \left(\frac{\partial}{\partial x^j} \right)_M \in \mathbb{T}_M X.$$

g_{ij} – матрица Грамма. То есть $\sqrt{|\det g|}$ есть объём параллелепипеда, построенного на осях $\left(\frac{\partial}{\partial x^1}\right)_M \dots \left(\frac{\partial}{\partial x^n}\right)_M$ евклидового пространства $\mathbb{T}_M X$.

Вывод: на любом многообразии, допускающем разложение единицы, существует плотность объёма, для получения которой можно ввести на многообразии риманову структуру и в качестве плотности выбрать риманову плотность.

Плотность объёма $\mathrm{d}V$ позволяет описать произвольные плотности ρ на многообразии. В самом деле, при замене координат величина $f = \rho/\mathrm{d}V$ не меняется. А потому мы имеем право назвать f функцией $f : X \to \mathbb{R}$ и определить $\rho = f\mathrm{d}V$.

Определим интеграл от произвольной финитной плотности ρ по многообразию X. Из теоремы о разложении единицы следует, что существует открытое покрытие $\{U\}_{i=1}^s$ компакта $K = \operatorname{Supp}\rho$ и такое разложение единицы $\{\alpha_i\}_{i=1}^s$, что $\forall i \operatorname{Supp}\alpha_i \subset U_i$, где U_i – область определения карты (U_i, φ_i). Определение интеграла от плотности по многообразию:

$$\int_X \rho = \sum_{i=1}^{S} \int_{\varphi_i(U_i)} (\rho^{U_i} \circ \alpha_i)[\varphi_i^{-1}(\vec{x}_i)]\mathrm{d}\mu(\vec{x}_i), \quad \mathrm{d}\mu(\vec{x}_i) \equiv \mathrm{d}x^1 \dots \mathrm{d}x^n, \quad \operatorname{Supp}\rho^{U_i} \circ \alpha_i \subset U_i \quad \rho^{U_i} \circ \alpha_i \in \mathcal{K}(U_i), \tag{19}$$

где $\mathrm{d}\mu(\vec{x}_i)$ – элементарный интеграл Римана, а в квадратных скобках указан аргумент сложной функции $\rho^{U_i} \circ \alpha_i$. ($\mathcal{K}(U_i)$ в (19), вероятно, обозначает финитные функции.)

12.9 Поверхностный интеграл I рода

12.10 Дифференциальные формы. Преобразования тензорных полей

Рассмотрим гладкое многообразие X класса $\mathbb{C}^2$ размерности $\dim X = n$. Вообще-то внешние формы уже упоминались в §7.2 на стр.5 и в §12.6 на стр.13, но это не мешает нам поминать их снова и снова...

Определение 12.11 *(Внешней) дифференциальной p-формой называется кососимметрическое тензорное поле типа $(p, 0)$ на X.*

Если $V = \mathbb{T}_M X$, точки обозначим $\Lambda_M{}^{\mathrm{P}} X = \mathbb{T}_a^{(p,0)}(V)$.

Определение 12.12 *Внешней дифференциальной формой называется соответствие $\omega : M \to \omega(M) \equiv \omega_M \in \Lambda_M{}^P X$.*

Как мы уже знаем, в некоторй карте $(U, x^1 \dots x^n)$ внешнюю дифференциальную форму можно записать как

$$\omega = \sum_{1 \le i_1 < \dots < i_p \le n} \omega_{i_1 \dots i_p} \mathrm{d}x^{i_1} \mathrm{d}x^{i_p}. \tag{20}$$

Кстати, вспоминаем операцию внешнего умножения, определённую в §7.4.

Преобразования тензорных полей

Наряду с гладким многообразием X рассмотрим гладкое многообразие Y, $\dim X = n$, $\dim Y = m$. Допустим, $f : X \to Y$ – гладкое отображение, $f : X \to Y \Rightarrow \forall p, q\ \mathrm{d}f : \mathbb{T}^{(p,q)} X \to \mathbb{T}^{(p,q)} Y$; S – ковариантное тензорное поле на Y типа $(p, 0)$. Сопоставим этому полю поле $\tilde{S} = f^* S$ на X следующим образом: для произвольных $\forall M \in X\ \forall \vec{T}_1 \dots \vec{T}_p \in \mathbb{T}_M X$ положим $\tilde{S}_M(\vec{T}_1 \dots \vec{T}_p) = S_{f(M)}(\mathrm{d}f_M(\vec{T}_1) \dots \mathrm{d}f_M(\vec{T}_p))$, где дифференциал определён в секции 11.4.

$$\begin{gathered} f : X \to Y \Rightarrow \forall p, q\ df : \mathbb{T}^{(p,q)} X \to \mathbb{T}^{(p,q)} Y, \\ M \in X, \quad f^* S = \tilde{S} \in \mathbb{T}^{(p,0)} X : \quad \tilde{S}_M\left(\vec{T}_1 \dots \vec{T}_q\right) \stackrel{\mathrm{df}}{=} S_{f(M)}\left[df_M\left(\vec{T}_1\right) \dots df_M\left(\vec{T}_q\right)\right]. \end{gathered} \tag{21}$$

Определение 12.13 *поле $\tilde{S} = f^* S$ называется переносом поля S на гладкое многообразие X при помощи отображения f.*

В координатах

$$\tilde{S}_{j_1 \dots j_p} = S_{i_1 \dots i_p} \frac{\partial f^{i_1}}{\partial x^{j_1}} \dots \frac{\partial f^{i_p}}{\partial x^{j_p}}. \tag{22}$$

Пример: индуцируем метрический тензор g (изначально определённый на Y) на подмногообразие $X \subset Y$ с помощью вложения $f = i^* : X \to Y$. Вложение индуцирует метрический тензор $i^* g$ на X; таким образом можно, например, получить метрический тензор для произвольной поверхности в трёхмерном пространстве:

$$X \subset Y = \mathbb{R}^3, \quad \dim X = 2, \quad g_{ij} = \delta_{ij}, \quad (i^* g)_{j_1 j_2} = g_{i_1 i_2} \frac{\partial f^{i_1}}{\partial x^{j_1}} \frac{\partial f^{i_2}}{\partial x^{j_2}} = \frac{\partial f^i}{\partial x^j} \frac{\partial f^i}{\partial x^j}. \tag{23}$$

For example, consider a situation when $X \equiv S^2 \in \mathbb{R}^3$ is a 2D-sphere, and $f = i : S^2 \to \mathbb{R}^3$ is an insertion function:

$$\begin{gathered} M \equiv (\theta,\varphi) \in S^2, \quad f(M) \equiv i(\theta,\varphi) = \vec{r}(\theta,\varphi) = [x(\theta,\varphi), y(\theta,\varphi), z(\theta,\varphi)]\,, \\ x = \sin\theta\cos\varphi,\ y = \sin\theta\sin\varphi,\ z = \cos\theta, \quad 0 \le \theta \le \pi, \quad -\pi \le \varphi < \pi. \end{gathered} \tag{24}$$

The metric tensor from $\mathbb{R}^3$ induces [through the transmission mechanism (21)] metric tensor on S^2:

$$p = 2,\ g = g_{ij}\mathrm{d}x^i \otimes \mathrm{d}x^j \in \mathbb{T}^{(2,0)}, \quad g \in \mathbb{T}^{(2,0)}\mathbb{R}^3, \quad i^*g \equiv \tilde{g} \in \mathbb{T}^{(2,0)}S^2. \tag{25}$$

The differential df is determined by the Jacobi matrix of f:

$$df = \left(\frac{\partial y_i}{\partial x_j}\right) dx^j, \quad dx^j \frac{\partial}{\partial x^i} = \delta^j_i, \tag{26}$$

and we obtain the induced metric tensor $\tilde{g}$ on a sphere, see (27) below. We see that the corresponding 2D metric tensor (tensor field defined on a sphere) is φ-independent:

$$\tilde{g}(\varphi,\theta) = \begin{pmatrix} \frac{\partial\vec{r}}{\partial\theta}\cdot\frac{\partial\vec{r}}{\partial\theta} & \frac{\partial\vec{r}}{\partial\varphi}\cdot\frac{\partial\vec{r}}{\partial\theta} \\ \frac{\partial\vec{r}}{\partial\theta}\cdot\frac{\partial\vec{r}}{\partial\varphi} & \frac{\partial\vec{r}}{\partial\varphi}\cdot\frac{\partial\vec{r}}{\partial\varphi} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & \sin^2\theta \end{pmatrix} = g(\theta) \equiv \{g_{ij}\}^2_{i,j=1} \in \mathbb{T}^{(2,0)}S^2. \tag{27}$$

Перенос работает и для дифференциальных форм, сохраняя их кососимметричность. Можно показать, что $(g\circ f)^* = f^*\circ g^*$ и $f^*(\omega_1\wedge\omega_2) = f^*\omega_1 \wedge f^*\omega_2$.

12.11 Ориентируемые многообразия

Рассмотрим многообразие X. Карты (U,φ) и $({}'U,{}'\varphi)$ называется положительно согласованными, если либо $U\cap{}'U = \emptyset$, либо на $U\cap{}'U$ $\det\partial\varphi/\partial{}'\varphi > 0$. Если карты согласованы, то наборы векторов $\frac{\partial}{\partial x^1}\dots\frac{\partial}{\partial x^n}$ $\frac{\partial}{\partial\,{}'x^1}\dots\frac{\partial}{\partial\,{}'x^n}$ образуют одноимённые (положительные или отрицательные) базисы в касательных пространствах.

Определение 12.14 *Многообразие называется ориентируемым, если на нём существует атлас, все карты которого попарно положительно согласованы. Такой атлас называется ориентирующим.*

Максимальный ориентирующий атлас называется ориентацией многообразия; ориентация многообразия автоматически определяет ориентацию касательного пространства в каждой точке этого многообразия.

На ориентируемом многообразии имеется естественное соответствие между плотностью и дифференциальной формой, не зависящее от карты. Из плотности дифференциальная форма получается так:

$$(U,\varphi) = (U, x^1\dots x^n), \quad \omega_U \overset{\mathrm{df}}{=} \rho^U \mathrm{d}x^1\wedge\ldots\wedge\mathrm{d}x^n, \tag{28}$$

где (U,φ) – положительная карта. Обратное соответствие: в карте (U,φ) сопоставим форме $\omega \equiv \omega_U = \omega^U\mathrm{d}x^1\wedge\ldots\wedge\mathrm{d}x^n$ плотность ρ^U

$$\rho^U = \begin{cases} \ \ \omega^U, & \text{если } (U,\varphi)\in A^+, \\ -\omega^U, & \text{если } (U,\varphi)\notin A^+. \end{cases} \tag{29}$$

Очевидно, что отображения (28) и (29) – взаимообратны.

Таким образом, на ориентируемом многообразии любую дифференциальную форму можно проинтегрировать (проинтегрировав соответствующую ей плотность), и результат не будет зависеть от выбора карты. Следовательно, дифференциальные формы могут соответствовать физическим (реально существующим) полям; это свойство обеспечивается их кососимметричностью.

12.12 Connection = covariant derivativation

Отрывок из факультатива. Важно научиться определять операции с тензорными полями, особенно – кососимметричными. Мы научились интегрировать внешние формы, умеем их (внешне) умножать; теперь ещё надо научиться их дифференцировать. Одно из преимуществ кососимметрических тензорных полей над произвольными – в том, что для дифференцирования произвольного тензорного поля необходимо задание символа Кристоффеля Γ^k_{sl} (что чаще всего означает задание метрики). При дифференцировании внешних форм Γ^k_{sl} не нужен. – кстати, предстоит доказать совместимость операци дифференцирования для внешних форм с (30).

И ещё: так называемый[12] (в книжках по теории поля) "интеграл по грассмановым переменным" есть всего лишь свёртка по индексам кососимметрического тензора, см. главу 3 в [27].

[12] Определение "интеграла по грассмановым переменным" в физической литературе без упоминания о связи с кососимметрическими формами – наглядная иллюстрация оторванности теоретической физики от математики и происходящего от этого вреда.

A usual derivative operator applied to a tensor field on a manifold does not generally produce a tensor field. One has to use the s.c. *covariant derivativation* operator instead, which always produces a tensor field out of any tensor field; The most general definition is written on p. [27]259. We will study covariant derivative of a vector and a $\mathbb{T}^{(2,0)}$ (metric) tensor:

$$T \in \mathbb{T}^{(0,1)}, \quad \nabla_l T^k = \frac{\partial T^k}{\partial x^l} + \Gamma^k_{sl} T^s, \quad g \in \mathbb{T}^{(2,0)}, \quad \nabla_k g_{ij} = \frac{\partial g_{ij}}{\partial x^k} - \Gamma^l_{ik} g_{lj} - \Gamma^l_{jk} g_{il}. \tag{30}$$

From (30) we see that covariant derivativation operator is determined by the s.c. Christoffel symbol Γ. There is an obvious similarity between Cristoffel symbol Γ and vector potential [also non-abelian one due to spin-orbit interaction (SOI)]. По крайней мере для случая обычного векторного потенциала эта аналогия известна; так можно описывать векторный потенциал в эффекте Аронова-Бома.

Определение 12.15 *Connection is called* ***symmetrical*** *if there exists*[13] *a coordinate system (i.e., a map) where* $\Gamma^l_{jk} = \Gamma^l_{kj}$.

Определение 12.16 *Suppose metric tensor is covariantly constant:* $\nabla_k g_{ij} = 0$. *Then the corresponding connection is called* ***pseudoriem***

Теорема 12.17 *On any pseudoriemannian manifold* $\exists!$ *pseudoriemannian connection:*

$$\Gamma^k_{sl} = g^{ki}\Gamma_{i,sl}, \quad \Gamma_{i,sl} = \frac{1}{2}\left(\frac{\partial g_{il}}{\partial x^s} + \frac{\partial g_{si}}{\partial x^l} - \frac{\partial g_{ls}}{\partial x^i}\right). \tag{31}$$

12.13 Символ Кристоффеля в жизни мирской

Кто-то скажет про символы Кристоффеля, что, мол, “страшно далеки они от народа.” На самом деле Γ^l_{kj} оторван от народа не так уж сильно.

Пример: дифференциальное уравнение для нахождения кратчайшего пути, соединяющего две точки на карте в гористой местности. (Используется вариационное исчисление.)

12.14 Накрытия

Рассмотрим график семизначной функции $f : \mathbb{R} \to \{1, 2, 3, 4, 5, 6, 7\} \times \mathbb{R}$ или просто набор из семи функций $\{f_i\}_{i=1}^7$ $f_i : \mathbb{R} \to \mathbb{R}$. Раскрасим каждую ветвь одним из цветов радуги, так что каждому значению i соответствует один из 7 цветов ($i = 1$ обозначает красный, а $i = 7$ – фиолетовый). Предположим, что при некотором значении $x_1 \in \mathbb{R}$ значения $f_1(x_1) \dots f_7(x_1)$ упорядочены именно так, как в радуге, т.е. $f_1(x_1) \le f_2(x_1) \le \dots \le f_6(x_1) \le f_7(x_1)$. При некотором другом значении аргумента упорядочение, возможно, будет иным, например, $f_2(x_2) \le f_7(x_2) \le f_1(x_2) \le f_4(x_2) \le f_3(x_2) \le f_6(x_2) \le f_5(x_2)$. Это изменение упорядочения соответствует некоторой перестановке, т.е. у нас внезапно вылезает группа перестановок. Если ветви функции представить себе в виде разноцветных волосков, мы получим косу, где волоски перепутаны, но не порваны, так что в любом месте косы число волосков одинаково (и равно семи в нашем простом примере). С этим примером связан английский термин “fiber bundle”. По-русски – накрытие (в случае, если индекс i принимает дискретные значения) или “расслоение” в более общем случае, когда i может принимать непрерывные (несчётные) значения.

12.15 Расслоения

Расслоения могут буть важны, потому что интеграл по расслоению используется в [2], к примеру, для объединения квантовой механики с теорией вероятностей.

Предположим, что карты в атласе можно запараметризовать: …

Примером нетривиального расслоения яляется лента Мёбиуса, параметризуемая уравнениями

$$x = \left(1 + \frac{v}{2}\cos\frac{u}{2}\right)\cos u, \quad y = \left(1 + \frac{v}{2}\cos\frac{u}{2}\right)\sin u, \quad z = \frac{v}{2}\sin\frac{u}{2}. \tag{32}$$

– неориентируемая поверхность с краем. Граница листа Мёбиуса состоит из одной замкнутой кривой. Здесь пространство базы M представляет собой окружность а слои $F = [-1, 1]$. The Möbius strip is also a standard example used to illustrate the mathematical concept of a fiber bundle. Specifically, it is a nontrivial bundle over the circle S1 with a fiber the unit interval, I = [0, 1]. Looking only at the edge of the Möbius strip gives a nontrivial two point (or Z2) bundle over S1.

12.16 Поверхностные интегралы II рода

Пример: поток векторного поля …

[13] Since Γ is not a tensor, the statement $\Gamma^l_{jk} = \Gamma^l_{kj}$ is map-dependent.

12.17 Криволинейные интегралы по пути

13 Введение в общую теорию интегралов

§1.Определение интеграла Лебега

Введём новое обозначение: $\mathcal{K}^+(X)$ – множество неотрицательных финитных функций на X.

13.1 Множества меры нуль и полной меры

Множество $Z \subset X$ называется **множеством меры нуль** (мм0), если $\forall \varepsilon > 0$ $\exists$ такая неубывающая последовательность неотрицательных функций $h_m \in H$, что $\forall m \in \mathbb{N}$ $Ih_m < \varepsilon$ и при этом $\forall x \in Z$ $\sup_m h_m(x) \geq 1$. Дополнение ММ0 называется **множеством полной меры**. В дальнейшем нередко будет использоваться оборот "почти всюду" (a.a., almost anywhere) – означает "на множестве полной меры".

Пример: $H = \mathcal{K}(\mathbb{R})$ для обычного интеграла $Ih = \int_{-\infty}^{\infty} h(x)\mathrm{d}x$ любая точка $x_0 \in \mathbb{R}$ есть множество меры нуль. В этом (простейшем) случае для всех m положим $h_m(x) = 0$ при $|x - x_0| \geq \varepsilon/2$, $h_m(x) = 1 - 2|x - x_0|/\varepsilon$ при $|x - x_0| < \varepsilon/2$ (треугольная горка на графике).

Свойства ММ0:

1. Подмножество ММ0 есть ММ0.
2. НБЧС объединение ММ0 есть ММ0.

Лемма: Если последовательность неотрицательных функций h_n, (всюду) невозрастая, *почти всюду* стремится к нулю, то $Ih_n \underset{n\to\infty}{\longrightarrow} 0$. Доказательство: ...

Следствия:

1. Если *почти всюду* функция $h \in H$ равна нулю, то $Ih = 0$.
2. Если последовательность $h_n \geq 0$ *почти всюду* невозрастает и *почти всюду*[14] стремится к нулю, то $Ih_n \underset{n\to\infty}{\longrightarrow} 0$.

Замечание: в дальнейшем косые стрелки $\nearrow$ и $\searrow$ будут использоваться двояко:

1. для обозначения неубывающих и невозрастающих (числовых) последовательностей, и
2. для *почти всюду* неубывающих и невозрастающих функциональных последовательностей. Например, $h_n \nearrow f$ означает, что на некотором множестве полной меры последовательность неубывает и стремится к f.

13.2 Интеграл в классе L^+

Начиная с этого места нашим функциям разрешается принимать значения $\pm\infty$. Говорят. что функция f принадлежит классу L^+, если существует ПВ неубывающая последовательность элементарных функций, ПВ стремящаяся к f. Доп. условие: интегралы от всех функций последовательности равномерно ограничены: $\exists h_n \in H : h_n \nearrow f$ и $\exists C > 0 : \forall n \geq 1$ $Ih_n < C$.

Свойства функнкций из $f \in L^+$:

1. **1. Функции $f \in L^+$ ПВ конечны.** Так как соответствующая функции f последовательность функций $h_n \in H$ ПВ неубывает, то ПВ $f \neq -\infty$. Остаётся доказать, что ПВ $f \neq +\infty$...
2. Из определения функций в класса L^+ следует, что любая функция, совпадающая на МПМ с некоторой (принадлежащей L^+) функцией, тоже принадлежит L^+. В частности, функция, равная нулю принадлежит L^+; следовательно, любая ф-ция, отличная от нуля только на ММ0, также принадлежит L^+.

Определение интеграла в L^+:

$$h_n \in H, h_n \nearrow f, Ih_n < C \Longrightarrow If \overset{\mathrm{df}}{=} \lim_{n\to\infty} Ih_n. \tag{33}$$

Существование предела в (33) очевидно; надо доказать единственность....

В дальнейшем при рассмотрении одного интеграла и нбчс числа алгебраических операций и/или предельных переходов, мы не будем различать функции, совпадающие на мпм.

Докажем корректность определения (33)...

Следствия:

[14] Если в рассуждении встречается больше одного "*почти всюду*", можно считать, что они совпадают (т.е. относятся к одному и тому же интегралу).

1. ♣ Интеграл определён корректно
2. Если $f, g \in L^+$ и ПВ $f \le g \Longrightarrow If \le Ig$.

Доказательство леммы: ...

13.3 Свойства интегралов в классе L^+

Теорема 13.1 *Класс L^+ замкнут относительно монотонного предельного перехода для функций с ограниченными интегралами. Иначе говоря, для $\forall \{f_n\}_{n\ge1} \subset L^+$: пв $f_n \nearrow f$ $\forall n$ $If_n \le C \Longrightarrow f \in L^+$, $If = \lim\limits_{n\to\infty} If_n$.*

13.4 Интеграл Лебега

Определение суммируемой (или интегрируемой по Лебегу) функции: $\varphi \in L$, если $\exists f, g \in L^+$: пв $\varphi = f - g$.

Аналогия из далёкого прошлого (стр.II133): интеграл относительно функций ограниченной вариации.

§2.Интегрирование функциональных последовательностей

13.5 Теорема Леви

Рассмотрим интеграл от положительного функционального ряда $\sum\limits_{k\ge1}\varphi_k$ с равномерно ограниченными интегралами от частичных сумм:

Теорема 13.2 $\varphi_k \ge 0$, $\varphi_k \in L$, $\forall n$ $I\sum\limits_{k=1}^{n}\varphi_k \le C \Longrightarrow \exists\varphi = \sum\limits_{k\ge1}\varphi_k \in L$ *и* $I\varphi = \sum\limits_{k\ge1} I\varphi_k$.

Доказательство: ...

Результаты этого параграфа (§) позволяют нам в дальнейшем получать пространство L из пространства элементарных функций следующим образом (схема Даниэля[15]):

1. в начале имеем пространство элементарных функций H и элементарный интеграл I.
2. получаем L и L^+.
3. допустим, $\exists\tilde H \subset L$, и это $\tilde H$ удовлетворяет аксиомам пространства элементарных функций. Обозначим ограничение $\tilde I = I\Big|_{\tilde H}$.
4. по теорема Леви, $\tilde I$ – непрерывный функционал на $\tilde H$, т.е. $\tilde I$ есть элементарный интеграл.
5. построим пространство суммируемых функций с $\tilde I$; по одному из следствий т.13.2, любое множество меры ноль относительно $\tilde I$ будет также множеством меры ноль относительно I
6. по одному из следствий т.13.2, ...

13.6 Теорема Лебега

13.7 Варианты теоремы Лебега

Теорема 13.3 *Допустим, пв последовательность функций $\varphi_n \longrightarrow \varphi$ и пв $|\varphi| \le \varphi_0 \in L \Longrightarrow \varphi \in L$ и $|I\varphi| \le |I\varphi_0|$.*

Определение: пв конечная функция f называется измеримой,[16] если она является пределом пв сходящейся последовательности элементарных функций $h_n \in H$.

Очевидно, любая функция из L и из L^+ измерима.

Теорема 13.3 означает, что любая не слишком большая измеримая функция является суммируемой.

[15] Похоже, схемой (и даже интегралом) Даниэля называется построение интеграла через пределы последовательностей элементарных функций; Лебег же танцевал от меры. Калинин, таким образом, приверженец Даниэля. Эти два подхода (интеграла) эквивалентны.

[16] См. также определение на стр. 19.

13.8 Сравнение интеграла Римана и интеграла Лебега

13.9 Измеримость функций и множеств

См. определение измеримой функции на стр. 19.

13.10 Измеримость предела последовательности измеримых функций

Теорема 13.4 *Предел пв сходящейся последовательности измеримых функций измерим.*

13.11 Измеримые множества

Определения:

1. Множество E называется **измеримым**, если его хф χ_E – измеримая функция.
2. Множество E называется **суммируемым**, если его хф $\chi_E \in L$. При этом **мерой множества** E называется $\mu(E) = I\chi_E$; из определения следует, что $\mu(\emptyset) = 0$.
3. Если множество E измеримо, но не суммируемо, будем по определению считать $\mu(E) = \infty$.

Свойства:...

Основная теорема в теории меры:

Теорема 13.5 *о σ-аддитивности меры: объединение последовательности измеримых множеств измеримо; а если эти множества к тому же ещё и взаимно не пересекаются, то $\mu(E) = \sum\limits_{n\geq 1}\mu(E_n)$, причём в обеих частях последнего равенства может стоять бесконечность.*

13.12 Аксиомы Стоуна

Уже имеющихся (см. секцию 12.1) аксиом недостаточно: мы даже не можем утверждать, что множество X измеримо. Начиная с этого места будем предполагать, что H на X, в дополнение к двум аксиомам для элементарного интеграла (см. 1 и 2 в начале вопроса 12.1) выполняются также следующие две:

3. $\forall h \in H\ \min(h, 1) \in H$ – заведомо выполнена, если X – гладкое многообразие, а $\mathcal{K}(X)$ – пространство финитных функций.
4. существует последовательность функций $h_n \in H$, обладающая свойствами: $h_n \geq 0$, $Ih_n \geq 0$, $\forall x \in X\ \sup\limits_n h_n(x) > 0$.

– эти две аксиомы заведомо выполняются, если $1 \in H$ и интеграл (функционал) I ненулевой. Именно это имеет место в теории вероятностей.

Теорема 13.6 *о характеризации измеримых функций: пв конечная функция φ измерима $\iff \forall c \in \mathbb{R}$ множество $E(\varphi, c) = \{x \in X | \varphi(x) > c\}$ измеримо.*

Доказательство: ...

13.13 Определение интеграла Лебега по Лебегу

Рассмотрим $\varphi \in L$; пока будем считать, что $\varphi > 0$. Чему равен $I\varphi$? Возьмем некоторые $k, n \in \mathbb{N}$ и положим

$$E_{kn}(\varphi) \overset{\text{df}}{=} \left\{x \in X | \frac{k}{n} < \varphi(x) \leq \frac{k+1}{n}\right\}. \tag{34}$$

Множество (34) представимо в виде разности $E_{kn}(\varphi) = E(\varphi, \frac{k}{n}) \backslash E(\varphi, \frac{k+1}{n})$; из предыдущего (13.12) вопроса следует, оба множества ($E(\varphi, \frac{k}{n})$ и $E(\varphi, \frac{k+1}{n})$) суммируемы. Следовательно $E_{kn}(\varphi)$ тоже суммируема, т.к. является разностью СМ. Обозначим $\chi_{kn} \equiv \chi_{E_{kn}}$, определим

$$\varphi_n(x) \overset{\text{df}}{=} \sum_{k\geq 1} \frac{k}{n}\chi_{kn}(x) \tag{35}$$

Не исключено, что кое-где $\varphi(x) = \infty$, но эти точки образуют мм0. За их исключением (т.е. пв) (35) равно k/n при $x \in E_{kn}$. Согласно построению, пв $\varphi_n(x) \le \varphi(x)$. По теореме 13.4, φ_n измеримы.

$$I\varphi_n = \sum_{k\ge 1} \frac{k}{n} I\chi_{kn} = \sum_{k\ge 1} \frac{k}{n} \mu_{kn}. \tag{36}$$

По т. Лебега (см. сек. 13.6)

$$\lim_{n\to\infty} I\varphi_n = \lim_{n\to\infty} \frac{k}{n}\mu_{kn} \tag{37}$$

Вставить рисуночек на стр.VII80.

13.14 Интеграл по измеримому подмножеству

13.15 Ступенчатые функции в качестве элементарных

14 Интеграл на многообразии

Замена переменной в интеграле Лебега

14.1 Многообразия со счётным атласом

14.2 Интеграл на многообразии

14.3 Интеграл открытом на подмногообразии

14.4 Теорема о замена переменной в интеграле Лебега

Тетрадь 8

14.5 Теорема Стокса

14.6 Классический вариант формулы Стокса

14.7 Доказательство теоремы Стокса

14.8 Замкнутые и точные дифференциальные формы

– обобщение понятий “потенциальные и соленоидальные векторные поля.”

14.8.1 Вспомогательные утверждения

14.9 Лемма Пуанкаре

14.10 Критерии существования первообразной дифференциальной формы

14.11 Первообразная замкнутой формы вдоль пути

14.12 Первообразная замкнутой формы в односвязной области

14.13 В вопросу применимости дифференциальных форм

По факту, в физике твёрдого тела дифференциальные формы не применяются; я надеюсь, что причина этого – в математической безграмотности физиков. Вот взять, к примеру, дифференциальные формы: в чём полезность их антисимметричности относительно перестановки индексов? Нельзя ли ими моделлировать антикоммутационные соотношения для фермиевских полей, к примеру, в нелинейной сигма-модели? Нельзя ли просто описать дифференциальными формами занудную теорию возмущений из Ландау-Лившица? Ответы на эти вопросы можно поискать в [1, 2].

Энтузиаст: J. Fröhlich, см. квантовый эффект Холла. Ещё дифф. геометрия иногда используется при описании фазы Берри (Berry phase), а также т.н. “топологических изоляторов”. Насчёт последних найти вменяемую литературу особенно тяжело (примерно как золотую серёжку в канализации): тема модная, гранты под неё дают, так что толпы писателей строчат пачки никчёмных статей. Введения этих статей содержат однотипные фразы-заклинания, значения которых авторы зачастую сами не понимают.

При чтении arXiv/hep-th/0306205 возникла у меня такая идея: квантовая механика должна в пределе $\hbar \to 0$ переходить в классическую; в этом же пределе операторы должны коммутировать между собой. Пусть $\hbar$ – кривизна поверхности, а приближённая коммутация сродни аксиомам Евклида...

15 Вероятность и мера

15.1 Аксиомы теории вероятностей

Пусть Ω – некоторое множество, а $\mathcal{A}$ – некоторая система подмножеств $\subset \Omega$.

Определение 15.1 *Система подмножеств $\mathcal{A}$ называется* ***кольцом множеств****, если $\forall A, B \in \mathcal{A}$ $A \cap B \equiv AB \in \mathcal{A}$, $A \cup B \equiv A + B \in \mathcal{A}$ и $A \setminus B \in \mathcal{A}$.*

Определение 15.2 *Кольцо множеств $\mathcal{A}$ называется* ***алгеброй множеств****, если $\Omega \in \mathcal{A}$.*

В теории вероятностей Ω называется **пространством элементарных событий**, а его элементы $\omega \in \Omega$ называются **элементарными событиями**. Элементы $A \in \mathcal{A}$ называются **событиями**. Предположим, что задана функция[17] $P : \mathcal{A} \to [0, +\infty)$.

Определение: Тройка $(\Omega, \mathcal{A}, P)$ называется **вероятностным пространством**, если

1. $\mathcal{A}$ есть алгебра множеств в Ω.
2. Каждому событию $A \in \mathcal{A}$ сопоставлено число $P(A) \geq 0$, называемое **вероятностью события**.
3. $P(\Omega) = 1$.
4. $\forall A, B \in \mathcal{A}$: если $AB = \emptyset$, то $P(A + B) = P(A) + P(B)$.
5. аксиома непрерывности – см. (48) ниже.

Примеры:

1. Пусть на множестве Ω задан интеграл I с мерой μ, причём $\mu(\Omega) = 1$. В таком случае в качестве $\mathcal{A}$ можно выбрать алгебру всех суммируемых[18] подмножеств Ω, и $P(A) = \mu(A)$.
2. Дискретное вероятностное пространство: $\Omega = \{\omega_n\}_{n\in\mathbb{N}}$, $P : \Omega \to [0, 1]$, $P(\omega_n) \overset{\text{df}}{=} p_n$, $\sum_n p_n = 1$. Соответственно $\mathcal{A}$ представляет собой все подмножества Ω.

Интуитивное обоснование: Событие $A \in \mathcal{A}$ переставляет собой некоторую совокупность элементарных событий – элементарных неразложимых исходов эксперимента. Событие есть исход некоторого эксперимента, который можно многократно повторять. Если при одних и тех же условиях один и тот же эксперимент произведён $n \gg 1$ раз, и при этом событие A имело место n_A раз, то вероятность этого события $P(A) \approx n_A/n$. Если $P(A) = 0$, то можно быть практически уверенным в том, что при *однократном* воспроизведении эксперимента событие A не произойдёт. Соответственно, структура алгебры множеств соответствует логике событий:

1. Ω есть достоверное событие, а $\emptyset$ – невозможное событие.
2. $A \cup B$ – событие, заключающееся в происхождении события A или B.
3. $A \cap B \equiv AB$ – событие, заключающееся в происхождении обеих событий A и B.
4. $A \subset B$: если происходит A, то должно произойти также и B.
5. Если $A \cap B \equiv \emptyset$, то события A и B являются взаимоисключающими.

---

[17] Из аксиом следует, что $P : \mathcal{A} \to [0, 1]$, т.е. $\forall A \in \mathcal{A}$ $0 \leq P(A) \leq 1$: Рассмотрим $\bar{A} = \Omega \setminus A \Rightarrow \bar{A}A = \emptyset$, $\bar{A} + A = \Omega \Rightarrow P(A) = 1 - P(\bar{A}) \Rightarrow 0 \leq P(A) \leq 1$.

[18] Возможно, подойдут также измеримые подмножества (см. определение на стр. 19).

15.2 Статистика размещений (различимых и неразличимых объектов)

Пусть у нас имеется n штук неразличимых шариков (частиц) которые нужно произвольным образом разложить по r различным коробочкам (энергетическим уровням[19]) Рассмотрим сначала случай бозонов, когда на любом уровне может сидеть произвольное количество частиц. Таким образом, наше вероятностное пространство описывается множеством $\Omega = \{(n_1 \dots n_r) | \sum_{k=1}^r n_k = n,\ n_k \in \mathbb{N} \cup \{0\}\}$, то есть каждое элементарное событие задаётся набором из r штук неотрицательных целых чисел, сумма которых фиксирована.

Предположим, что все элементарные события имеют одинаковую вероятность P_{nr}, так что P_{nr}^{-1} равно количеству $A(n,r) \equiv |\Omega|$ элементов во множестве Ω. Для подсчёта $A(n,r)$ рассмотрим производящую функцию

$$f_1(z) \overset{\text{df}}{=} \sum_{n \geq 0} A(n,r) z^n = \boxed{z^{n_1} \dots z^{n_r} = z^n} = \prod_{i=1}^r \left(\sum_{n_i \geq 1} z^{n_i} \right) = (1-z)^{-r} \Rightarrow$$

$$\Rightarrow A(n,r) = \frac{r(r+1)\dots(r+n-1)}{n!} = \frac{(r+n-1)!}{n!(r-1)!} = P_{nr}^{-1}.$$

Проделаем теперь то же самое для фермионов, когда на один уровень может попасть только одна частица, так что $n \leq r$:

$$\Omega = \left\{ (n_1 \dots n_r) \middle| \sum_{k=1}^r n_k = n,\ n_k = 0,1 \right\}, \quad f_2(z) = \sum_{n=0}^r B(n,r) z^n = \prod_{i=1}^r \left(\sum_{n_i=0}^1 z^{n_i} \right) = (1+z)^r \Rightarrow$$

$$\Longrightarrow |\Omega| = C_n^r = \frac{r!}{m!(r-n)!} = 1/P(n_1 \dots n_r).$$

Наконец, рассмотрим “классический” случай, когда частицы различимы и количество частиц в одной коробке (на уровне) произвольно. Соответственно теперь маркируем элементарные состояния не числами заполнения, а “в какой коробке (или на каком уровне) находится каждая из частиц”?

$$\Omega = \{(i_1 \dots i_n)\} \Rightarrow |\Omega| = r^n \Rightarrow P(i_1 \dots i_n) = r^{-n}. \tag{38}$$

15.3 Условная вероятность и независимость

Определение: **Условной вероятностью** события B при условии A называется $P(B|A) \overset{\text{df}}{=} P(AB)/P(A)$.

Интуитивный смысл: $P(B|A)$ есть вероятность события B при условии, что событие A произошло.

Очевидные свойства:

1. $P(A|A) = 1 \Rightarrow P(\Omega|A) = 1$,
2. $P(B + C|A) = P(B|A) + P(C|A)$,
3. Из $P(AB) = P(A)P(B|A)$ следует **теорема умножения вероятностей:**
$$P(A_1 A_2 \dots A_n) = P(A_1 \dots A_{n-1}) P(A_n | A_1 \dots A_{n-1}) = \dots = P(A_1)P(A_2|A_1) \cdot P(A_3|A_1A_2) \cdot \ldots \cdot P(A_n|A_1 \dots A_{n-1}).$$
4. $P(A + B) = P(A) + P(B) - P(AB)$. Доказательство:
$$P(A \backslash AB) + P(AB) = P(A) \Longrightarrow P(A \backslash AB) = P(A) - P(AB) \Longrightarrow P(A) + P(B) - 2P(AB) = P(A \backslash AB) + P(B \backslash AB)$$
$$= P(A + B \backslash AB) = P(A+B) - P(AB) \Longrightarrow P(A+B) = P(A) + P(B) - P(AB).$$

Теорема 15.3 *о полной вероятности Если $H_1 \dots H_n$ – полная (т.е. $H_1 + \ldots + H_n = \Omega$) группа попарно несовместимых событий и $\forall k\ P(H_k) > 0$, то*

$$\Longrightarrow \forall A \in \mathcal{A} \quad P(A) = P(A\Omega) = \sum_k P(AH_k) = \sum_k P(A|H_k)P(H_k). \tag{39}$$

Теорема 15.4 *Байеса (Bayes) В условиях теоремы* (39)

$$\sqsupset \exists A : P(A) > 0 \Longrightarrow \forall i = 1 \dots n \quad P(H_i|A) = \frac{P(H_iA)}{P(A)} = \frac{P(H_i)|P(A|H_i)}{\sum_k P(A|H_k)P(H_k)}. \tag{40}$$

Определение: события A и B называются **независимыми**, если $P(AB) = P(A)P(B)$.

[19] Аналогия – так себе: в реальной физической системе число энергетических уровней бесконечно и при конечной температуре вероятность заселённости различных уровней различна (зависит от энергии уровня). STOP Добавить сюда вывод рапределений Ферми и Бозэ.

Пример: (См. стр. [13]29.) Иногда независимость событий неочевидна, и в этом случае для проверки удаётся использовать определение. Например, из карточной колоды, содержащей 36 карт наугад вытягивается одна карта. Пусть событие A состоит в том, что это пика, а событие B – что это дама. Являются ли эти события независимыми?

$$P(A) = 9/36, \quad P(B) = 4/36, \quad P(AB) = 1/36 = P(A)P(B).$$

Следовательно, эти события независимы.

Из независимости A и B следует независимость событий A и $\bar{B} \equiv \Omega - B$, а также $\bar{A}$ и $\bar{B}$. Независимость событий моделирует их причинную независимость: $P(A|B) = P(A)$ не зависит от B. События $A_1, \ldots, A_n$ называются **независимыми в совокупности**, если для любого набора $i_1 \ldots i_k$, $1 \leq k \leq n$, $1 \leq i_1 \leq i_2 \leq \ldots \leq i_k \leq n$ выполняется соотношение $P(A_{i_1} \ldots A_{i_k}) = P(A_{i_1}) \ldots P(A_{i_k})$. При $n \geq 3$ из попарной независимости не обязательно следует совокупная. Из совокупной независимости событий $A_1, \ldots, A_n$ следует совокупная независимость событий $B_1, \ldots, B_n$, где любое событие B_i из списка есть на выбор или A_i, или $\bar{A}_i$.

Пример: предположим, имеется 4 измерительных прибора, один из которых неисправен. Вероятность случайного выбора (любого из трёх) исправного прибора равна $P(H_1) = 0.75$, а единственного неисправного – $P(H_2) = 0.25$. Предположим, что нам известны условные вероятности события A (грубой ошибки измерения): вероятность грубой ошибки на исправном приборе $P(A|H_1) = 0.04$, а на неисправном – $P(A|H_2) = 0.92$. Теорема Байеса позволяет "обратить" условные вероятности, т.е. вычислить $P(H_1|A)$ и $P(H_2|A)$.

Итак, предположим, что мы произвели измерение и результатом его явилось событие A (грубая ошибка). Вероятность того, что выбранный нами прибор (которым мы произвели ошибочное измерение) неисправен, есть

$$P(H_2|A) = \frac{P(H_2)P(A|H_2)}{P(H_1)P(A|H_1) + P(H_2)P(A|H_2)} = \frac{0.23}{0.03 + 0.23} = 0.885.$$

Соответственно, из (39) следует, что вероятность того, что при повторении процедуры (начиная со случайного выбора одного прибора из четырёх) результатом измерения станет грубая ошибка есть $P(A) = 0.26$. Этот результат – демонстрация того, как мы сумели вывести вероятность "сложного" события[20] A (которое, мы, возможно сможем повторить всего лишь несколько раз или вообще однажды из-за его крайне малой вероятности), воспользовавшись знаниями, об "элементарных кирпичиках", приведших к этому событию. В реальной жизни нас очень часто интересуют редкие события, напр., "какова вероятность разрушения этого здания (кототое мы собираемся страховать) в результате стихийного бедствия в течение следующего года?" И тут уже надо брать вероятность каждого из возможных бедствий и устойчивость здания к нему.

Пример: (См. стр. [24]147.) Предположим, что 5% всех мужчин и 0.25% всех женщин страдает дальтонизмом. Наугад выбранное лицо страдает дальтонизмом. Какова вероятность того, что это – мужчина. (Считаем, что число мужчин и женщин одинаковое.)

15.4 Испытания Бернулли

Испытания Бернулли – последовательность независимых (в совокупности – см. определение на стр. 24) испытаний. Каждое испытание имеет два исхода; назовём их "успех" и "неудача". Обозначим (одинаковую для всех испытаний) вероятность успеха $P(1)$. Для одного испытания

$$\Omega_1 = \{0, 1\}, \quad P(1) \equiv p, \quad P(0) \equiv q = 1 - p,$$

а для n испытаний

$$\Omega_n = \underbrace{\{0,1\} \times \ldots \times \{0,1\}}_{n \text{ раз}}, \quad |\Omega| = 2^n, \quad \omega = (\varepsilon_1, \ldots, \varepsilon_n) \in \Omega, \quad P(\omega) = p^{\mu(\omega)} q^{n-\mu(\omega)},$$

где $\mu(\omega)$ есть число успехов (единиц) в ω. Пусть пространство элементарных событий $\mathcal{A}$ – система всевозможных подмножеств[21] $\subset \Omega$. Тогда вероятность некоторого (неэлементарного) события $A \in \mathcal{A}$ равна $P(A) = \sum_{\omega \in A} P(\omega)$. Пример серии таких испытаний: я подпросил монетку 5 раз, какова вероятность того, что 3 раза выпадет "орёл"?

Примеры:

1. (См. стр. [24]154-155.) Играю я, например, в лотерею. Вероятность выигрыша мала. Сколько билетов (n) мне надо купить для её повышения? Вероятность n неудач равна q^n, так что вероятность хотя бы одного успеха есть $1 - q^n$.

---

[20] "Сложное" событие A является следствием совокупности (в этом простейшем примере – всего двух) "простых" событий.

[21] См. пример на стр. 22: "дискретное вероятностное пространство."

Сколько билетов (n) мне надо купить для того, чтоб вероятность хотя бы одного выигрыша была не менее 1/2? Если (допустим) $p = 0.01$, то $1 - (0.99)^n$ и $-n \log 0.99 \geq \log 2$ т.е. $n \geq 70$. (Конечно же, предполагается, что общее число билетов настолько велико, что вероятность выигрыша постоянна. Если бы всего было 100 билетов, и ровно один из них был бы выигрышным, то вероятность была бы непостоянна.)

2. (см. стр. [28]100) Производим ряд независимых опытов до первого успеха (напр., играем в "однорукого бандита" до первого выигрыша). Найти мат. ожидание числа m опытов.

   Решение: мат. ожидание есть

$$Mm = 1 \cdot p + 2 \cdot qp + 3 \cdot q^2 p + \ldots = p(1 + 2q + 3q^2 + \ldots) = p\frac{\mathrm{d}}{\mathrm{d}q}(q + q^2 + q^3 + \ldots) = p\frac{\mathrm{d}}{\mathrm{d}q}\frac{q}{1-q} = \frac{1}{p}. \quad (41)$$

Теорема 15.5

$$P_n(m) \overset{\mathrm{df}}{=} P(\{\omega | \mu_n(\omega) = m\}) = C_n^m p^m q^{n-m}, \quad (42)$$

т.е. число успехов[22] *в испытаниях Бернулли имеет биномиальное распределение.*

Доказательство: $P(\mu = m)$ есть [число точек $\omega \in \Omega$, в которых ровно m единиц] умножить на [вероятность такого события $= p^m q^{n-m}$]. Кстати, нетрудно понять, что $P_n(\Omega_n) = \sum\limits_{m \geq 0} P_n(m) = \sum\limits_{m=0}^{n} C_n^m P(\mu = m) = 1 = P(\Omega)$ как это и должно быть.

<u>Замечания:</u>

1. число успехов в серии испытаний Бернулли представляет собой простейший пример *дискретной* случайной величины (СВ), см. §15.8.

2. в случае, когда каждое испытание имеет больше двух возможных исходов, получаем т.н. полиномиальное распределение, см. стр. [24]174.

Можно рассматривать серии испытаний Бернулли с вероятностью, меняющейся от серии к серии: пусть в первой серии будет одно ($n = 1$) испытание с вероятностью успеха p_1, во второй серии – два ($n = 2$) испытания с вероятностью успеха p_2, и т.д.

Отдельный интерес представляет случай, когда вероятность успеха p мала, а число испытаний n велико, так что $\lambda = np$ – ни мало, ни велико. В этом случае формулой (42) пользоваться неудобно, и нам на помощь приходит следствие:

Теорема 15.6 *Пуассона: пусть (Ω_n, A_n, P_n) – бесконечная последовательность вероятностных пространств, каждое из которых задаёт (всёвозрастающее число) n испытаний Бернулли с вероятностью успеха p_n. (Элементарная) вероятность <u>p_n зависит от n,</u> т.е. меняется от одной серии испытаний к другой – в смысле последующего примера на стр. 26. Предположим, что $\lim\limits_{n\to\infty} np_n = \lambda \geq 0$. Обозначим символом $\mu_n \equiv \mu_n(\omega)$ число единиц (успехов) в точке (элементарном событии) $\omega \in \Omega_n$. Тогда $\forall k \geq 0$ $\lim\limits_{n\to\infty} P(\mu_n = k) = \lambda^k e^{-\lambda}/k!$ (распределение Пуассона).*

То есть сначала мы <u>фиксируем</u> конечное значение k, а потом устремляем $n \to \infty$, так что $n \gg k$.
Доказательство:

$$\lim_{n\to\infty} np_n = \lambda \Longrightarrow p_n = \frac{\lambda}{n} + o\left(\frac{\lambda}{n}\right) \Longrightarrow$$
$$P(\mu_n = k) = C_n^k p_n^k (1 - p_n)^{n-k} = \frac{n(n-1)\ldots(n-k+1)}{k!}\left[\frac{\lambda}{n} + o\left(\frac{\lambda}{n}\right)\right]^k \underbrace{\left[1 - \frac{\lambda}{n} - o\left(\frac{\lambda}{n}\right)\right]^{n-k}}_{\sim e^{-\lambda},\quad n-k\approx n}, \quad (43)$$
$$e^x = \lim_{n\to\infty}\left(1 + \frac{x}{n}\right)^n \Longrightarrow P(\mu_n = k) \sim \frac{1}{k!}\cdot\frac{n}{n}\cdot\frac{n-1}{n}\ldots\frac{n-k+1}{n}\left(\lambda + o(\lambda)\right)^k e^{-\lambda} + O\left(\frac{k}{n}\right) \underset{n\to\infty}{\longrightarrow} \lambda^k e^{-\lambda}/k!$$

– полезная асимптотика формулы (42) для случая малых "чисел заполнения" $k \ll n$. Оценка точности – см. задача про дни рождения на стр. [24]159: вероятность того, что среди 500 наугад выбранных человек никто не родился 1 января, равна $(364/365)^{500} = 0.2537 \approx e^{-\lambda} = 0.2541$, где мы подставили в асимптотическую формулу $p = 1/365$, $n = 500$, $\lambda = 500/365$.

Распределение Пуассона очень часто согласуется с экспериментом: число частиц, зарегистрированное счётчиком Гейгера, количество вызовов, поступившее на телефонную станцию в течение часа, число случаев отказа оборудования.

Теорема (43) показывает, как можно перейти от дискретной СВ к непрерывной, усремляя число испытаний к бесконечности при конечном параметре λ. Параметр этот имеет смысл концентрации. При рассмотрении газа – это самая настоящая

[22] Сдаётся мне, что $\mu_n(\omega)$ представляет собой простейший пример случайной величины, см. определение (52) ниже.

концентрация и есть, то есть число частиц (молекул), поделённое на объём. При рассмотрении событий во времени (вылет электронов, поступление вызовов на АТС, приезд рейсовых автобусов на остановку) λ есть количество событий в единицу времени, то есть "концентрация во времени". Более подробно см. стр. [28]109-111.

Несложно рассчитать мат. ожидание и дисперсию (см. §15.8) для распределения Пуассона и убедиться в том, что обе эти величины равны λ. На практике это считается "фирменным знаком" распределения Пуассона: если для некоторой (дискретной или непрерывной) СВ оценки (см. §15.21) мат. ожидания и дисперсии близки, то мы сразу же начинаем эту СВ подозревать в том, что она распределена по Пуассону.

Важность результата (43) мы сможем осознать лишь позже, когда познакомимся с центральной предельной теоремой (ЦПТ). А именно, рассмотрим ситуацию, когда вероятность успеха в последовательности испытаний непостоянна (зависит от n). Для использования точной формулы (42) нам необходимо владеть большим объёмом информации – знать "элементарные" вероятности успеха $\{p_n\}_{n\geq 1}$ для каждого (n-того) члена последовательности. Следствием же можно пользоваться и в том случае, когда нам известна только средняя вероятность[23] $\lambda = \lim\limits_{n\to\infty}\sum\limits_{n=1}^{N} p_n/N$.

Пример: Имеется комната (область) объёма V, в комнате имеется небольшой аквариум (подобласть) объёма $V_1 \ll V$. Комната заполнена газом, состоящим из одинаковых частиц. Газ весьма разрежён, так что взаимодействиями (столкновениями) между частицами можно пренебречь. (В то же время наш газ пребывает в термодинамическом равновесии, так что исследуемые ниже вероятности постоянны во времени.) То есть имеем идеальный классический газ. Нам известно полное число молекул газа N; определим концентрацию $n = N/V$. Какова вероятность того, что в произвольный момент времени в нашем аквариуме окажется ровно $k \ll N$ частиц?

Сначала предположим, что у нас всего одна частица: $N = 1$ – это соответсвует первому члену последовательности в (42). Вероятность попадания нашей (пока – единственной) частицы в аквариум есть $P_1(1) = p = V_1/V \ll 1$.

Для перехода ко второму члену последовательности в принципе не обязательно добавлять вторую частицу; можно просто зарегистрировать местоположение нашей единственной частицы в другой момент времени. (Временной промежуток между измерениями д.б. достаточен для хаотизации частицы, т.е. для того, чтобы второе измерение было независимым от первого.) Мы верим в справедливость эргодической гипотезы: два разнесённых во времени измерения местоположения одной и той же частицы эквивалентны двум *одновременным* измерениям местоположения двух частиц.

В общем случае ($N > 1$), вероятность попадания k частиц $P_N(k) = C_N^k p^k (1-p)^{N-k}$ при том, что число N – порядка 10^{23}. Вышеприведённое следствие (43) становится полезным в случае, когда $k \ll N$. Осуществим переход к термодинамическому пределу: раздуваем V до $\mathbb{R}^m$, устремляем $N \to \infty$ при постоянной концентрации $n = N/V =$const. Вероятность того, что в объёме V_1 окажется одна частица $= p_n = V_1/V = V_1 n/N$, то есть предел $\lim\limits_{N\to\infty} Np_n = \lambda = V_1 n$ конечен. $\Longrightarrow$ Вероятность того, что в объёме V_1 окажутся $k \ll N$ частиц $= P(k) = (V_1 n)^k \exp(-V_1 n)/k!$

Тутубалин [15] утверждает, что следствие – нечто большее, чем просто асимптотика общей формулы (42). Общая формула неприменима почти никогда из-за того, что она требует постоянства вероятности в серии испытаний, или знания каждой вероятности в случае, если вероятность непостоянна. Более низкая точность асимптотической формулы неожиданно оборачивается её преимуществом, потому что для использования её достаточно знать среднюю вероятность, т.е. обладать куда меньшей информацией о системе. В этом смысле пример Калинина с идеальным газом, судя по всему, не раскрывает всей мощи теоремы (43).

Свою мысль Тутубалин иллюстрирует примером на стр. [15]31: вероятность поломки в течение некоторого времени сложной машины, состоящей из $n \gg 1$ независимых узлов при том, что каждый узел имеет собственную (отличную от других) малую вероятность поломки. Кстати говоря, подобные же оценки справедливы и для вероятности отказа компьютерной программы, состоящей из многочисленных независимых модулей.

Пример: (См. стр. [28]114.) С накалённого катода за единицу времени вылетает в среднем $q(t)$ электронов, где t – время с начала опыта. Найти вероятность того, что в интервал времени $[t_0, t_0+\tau]$ вылетит ровно m электронов.

Решение: эту задачу можно рассматривать как непрерывную версию испытаний Бернулли с переменной вероятностью успеха. Дискретная переменная n (номер испытания) нумерует дискретные времена t_n и превращается в непрерывную переменную t (время); вероятность p_n превращается в $p_n \to p(t_n) \to p(t)$. ... STOP полного понимания пока что нет... среднее число электронов, вылетающее за данный промежуток времени (то бишь концентрация) есть $\lambda = \int_{t_0}^{t_0+\tau} q(t)\mathrm{d}t$. Таким образом, искомая вероятность есть $\lambda^m e^{-\lambda}/m!$.

[23] В реальных задачах вероятность вычисляется с помощью оценки (см. вопрос 15.21), и никакой разницы между вероятностью и средней вероятностью нет. Поэтому если "знающий жизнь" Жирный Тони [29] слышит загачку о монете, которая 99 раз из 100 упала решкой, и при этом вероятность выпадения орла равна 1/2, он скажет, что (в переводе с матерного языка на математический) задача переопределена, противоречива и потому не имеет решения. Потому что Жирный Тони знает, что реальная жизнь – не абстрактная алгебраическая задачка; в жизни истинное значение вероятности почти всегда нам доподлинно неизвестно; мы его расчитываем на основе опыта, который в данном случае говорит нам, что вероятность эта наверняка не равна 1/2. Это в математике и в религии существуют аксиомы и догмы; настоящая жизнь свободна от них.

15.5 Примеры задач с испытаниями Бернулли

15.5.1 Повод для беспокойства

Нередко важен даже не результат расчёта с использованием биномиального распределения, но вывод о том, что биномиальное распределение не работает в случае, когда, казалось бы, работать оно должно. Такие ситуации – повод для волнений, раздумий и действий.

Например, в прошлом году произошла одна авиакатастрофа, а в этом – целых четыре. Означает ли это, что надо срочно искать виновных (их найдут всегда – будьте уверены) и увольнять их, или же ничего делать не нужно, потому что повышение количества катастроф – всего лишь случайная флуктуация, в которой никто не виноват?

Аналогично можно рассуждать про процент бракованных изделий. Если он не подчиняется биномиальному распределению, надо начинать думать и подозревать всякое.

15.5.2 Безошибочный хирург

Как справедливо заметили Хмельницкий и Сталлинга [30], наука ничего не может утверждать, а может только опровергать. В согласии с этим утверждением, мы можем только опровергать статистические гипотезы, но не подтверждать их.

Предположим, некоторый хирург проделал $n = 100$ однотипных операций, и при этом не произошло ни одного смертельного исхода. О вероятности смертельного исхода можно сказать, что она не превышает некоторого числа.

Предположим, что вероятность смерти пациента равна q. Вероятность хотя бы одной смерти есть $1 - q^n$. Надо найти такое (максимальное) значение q^*, при котором со (стандартной) вероятностью 0.997 в сотне операций будет хотя бы одна смерть. Так как мы знаем, что никто не умер, то гипотеза отвергается (с данным коэффициентом доверия). Следовательно, мы можем утверждать, что $q \leq q^*$. Вот если бы умерло не слишком мало пациентов, мы ны могли ограничить q с обеих сторон, но так как никто не умер – то только с одной. (Статистические данные не дают нам возможности утверждать, что $q \neq 0$.)

15.5.3 Задача о совпадениях

15.5.4 Разорение игрока

(См. стр. [13]26-27.) Простейшая игра в орлянку с фиксированным выигрышем/проигрышем (1 рубль). Партнёр (казино) обладает бесконечными финансовыми ресурсами. Начальный капитал: x рублей. Условие прекращения игры: разорение или приумножение капитала до a рублей.

Обозначим $p(x)$ – вероятность разорения при капитале x рублей; B – вероятность однократного выигрыша (в случае честной игры $P(B) = 1/2$.)

Вводим условные вероятности выигрыша

$$P(A|B) = p(x+1), \quad P(A|\bar{B}) = p(x-1), \quad P(B) = P(\bar{B}) = \frac{1}{2} \tag{44}$$

По формуле Байеса (39)

$$p(x) = \frac{1}{2}\left[p(x+1) + p(x-1)\right], \quad p(0) = 1, \quad p(a) = 0. \tag{45}$$

Для решения этого уравнения используем линейный анзац $p(x) = c + dx$. Получаем ответ $p(x) = 1 - x/a$.

15.5.5 Бракованые детали

(См. стр. [13]12-13.) ОТК выборочно контроллирует 10% деталей, если нет ни одного бракованного, вся партия принимается. Какова вероятность того, что партия, из 100 деталей, содержащая 10 бракованных деталей, будет принята?

10 деталей из 100 можно выбрать $N = C_{100}^{10} = 100!/(10!90!)$ способами. Количество способов выбора 90 качественных деталей есть $N(A) = C_{90}^{10} = 90!/(10!80!)$. Следовательно искомая вероятность есть

$$P(A) = \frac{N(A)}{N} = \frac{81 \cdot 82 \cdot \ldots \cdot 90}{91 \cdot 92 \cdot \ldots \cdot 100} \approx \left(1 - \frac{1}{10}\right)^{10} \approx \frac{1}{e}, \quad e \approx 2.78. \tag{46}$$

15.5.6 Оценка для среднего

(См. §[31]6.) Если нет компьютера, или просто нужно провести качественный анализ, можно воспользоваться (почти) гауссовым приближением [см. (1) из §[31]] для биномиального распределения. При $p \approx q$ можем считать, что распределение – гауссово. Соответственно, работает схема с построением доверительного интервала: вооружившись значением дисперсии σ, заявляем, что $P\{|k - np| > 3\sigma\} < 0.997$. Для дисперсии часто возможно воспользоваться её несмещённой оценкой (86).

Например, рассмотрим задачу со стр. [31]43: за некоторый период времени в Швейцарии родилось $b = 1359671$ мальчиков и $g = 1285086$ девочек. Что можно сказать о вероятности рождения мальчиков? Так как число наблюдений велико, да ещё и $p \approx q$

имеем тепличные условия для использования гауссова распределения. Оцениваем дисперсию как $\sigma = \sqrt{pq/n} \le 1/\sqrt{4n} = 0.0003$. Воспользовавшись несмещённой оценкой для среднего (85), пишем

$$n = b + g, \quad p = \frac{b}{n} \pm 3\sigma, \quad 0.5132 \le p \le 0.5150. \tag{47}$$

15.6 Вероятность и мера

(Пятая) **аксиома непрерывности** теории вероятностей: для любой последовательности вложенных множеств $\{A_n\}_{n\ge1} \subset \mathcal{A}$

$$\text{Если } \forall n\ A_{n+1} \subset A_n \text{ и } \bigcap_{n\ge1} A_n = \emptyset \Longrightarrow \lim_{n\to\infty} P(A_n) = 0. \tag{48}$$

Теорема 15.7 *о сигма/раширенной/счётной аддитивности: если A представимо в виде счётного объединения* попарно-несовместных *событий*

$$A = \bigcup_{n\ge1} A_n \in \mathcal{A} \Longrightarrow P(A) = \sum_{n\ge1} P(A_n). \tag{49}$$

Доказательство: условия пятой аксиомы (48) справедливы для

$$R_n = \bigcup_{k>n} A_k = A - \sum_{k=1}^{n} A_k \in \mathcal{A} \Longrightarrow \lim_{n\to\infty} P(R_n) = 0 \Longrightarrow P(A) = P\left(\sum_{k=1}^{n} A_k + R_n\right) = \sum_{k=1}^{n} P(A_k) + P(R_n) \underset{n\to\infty}{\longrightarrow} \sum_{k\ge1} P(A_k).$$

В дальнейшем мы будем предполагать, что условия теоремы (49) выполняются для всего вероятностного пространства Ω, т.е. что существует такая последовательность $\{A_n\}$ вложенных множеств $A_{n+1} \subset A_n$, что $\Omega = \bigcup_{n\ge1} A_n \in \mathcal{A}$. Иначе говоря, предполагаем, что пространство Ω – сигма-конечно.

Определение: **элементарной мерой** на кольце $\mathcal{A}$ называется неотрицательная конечная аддитивная[24] функция μ, заданная на $\mathcal{A}$ и обладающая свойством σ-аддитивности: любая последовательность *непересекающихся* множеств $\{A_n\}_{n\ge1} \subset \mathcal{A}$ таких, что $\bigcup_{n\ge1} A_n \in \mathcal{A}$ обладает свойством $\mu(A) = \sum_{n\ge1} \mu(A_n)$.

Замечание: из теоремы (49) следует, что в качестве вероятности мы имеем право рассматривать любую элементарную меру, обладающую (дополнительным) свойством $P(\Omega) \equiv \mu(\Omega) = 1$. Например, берём любой интеграл (Лебега) и некоторое множество X конечной меры. Тогда по теореме 13.5 все измеримые (см. стр. 19) подмножества образуют алгебру [см. определение 15.2].

Свойства:

1. Если $A = \bigcup_{n\ge1} A_n \in \mathcal{A}$ и $A_{n+1} \supset A_n$, то $\mu(A) = \lim_{n\to\infty} \mu(A_n)$.

   Доказательство: заметим что члены (т.е. множества в) последовательности $A_{n+1} \backslash A_n$ не пересекаются.

$$A = A_1 \bigcup_{n\ge1} (A_{n+1} \backslash A_n) = A_1 + \sum_{n\ge1} (A_{n+1} \backslash A_n) \Longrightarrow \mu(A) = \mu(A_1) + \sum_{n\ge1} [\mu(A_{n+1}) - \mu(A_n)] = \lim_{n\to\infty} \mu(A_n). \tag{50}$$

2. Если $A = \bigcap_{n\ge1} A_n \in \mathcal{A}$ и $\forall n\ A_{n+1} \subset A_n$, то $\mu(A) = \lim_{n\to\infty} \mu(A_n)$. – доказывает вышеприведённое замечание о том, что любой интеграл с конечной мерой порождает вероятностное пространство с выполнением в том числе пятой аксиомы (48).

   Доказательство:

$$A_1 \backslash A = A_1 \backslash \bigcap_{n\ge1} A_n, \quad \mu(A_1 \backslash A) = \mu(A_1) - \mu(A) = \mu\left(\bigcup_{n\ge1} A_1 \backslash A_n\right).$$

   Далее, последовательность $\{A_1 \backslash A_n\}_{n\ge1}$ возрастает с ростом n; следовательно, из предыдущего свойства (50) следует

$$\Longrightarrow \mu(A_1 \backslash A) = \lim_{n\to\infty} [\mu(A_1) - \mu(A_n)] = \mu(A_1) - \lim_{n\to\infty} \mu(A_n).$$

Определение: кольцо множеств $\mathcal{A}$ называется **σ-кольцом**, если (счётные) объединение и пересечение любой последовательности множеств из $\mathcal{A}$ также содержаться в $\mathcal{A}$. σ-кольцо $\mathcal{A}$ называется **σ-алгеброй**, если $\Omega \in \mathcal{A}$.

Определение (неэлементарной) меры: Рассмотрим некоторое кольцо $\mathcal{A}$. **Мерой** назовём σ-аддитивное,[25] отображение $\mu : \mathcal{A} \to [0, \infty]$, т.е. обладающее следующими свойствами:

[24] Аддитивность означает, что мера суммы равна сумме мер.

[25] У В.Л. было "аддитивное и σ-аддитивное", но, на мой взгляд, аддитивность автоматически следует из σ-аддитивности.

1. $\mu(\sum_{n\geq 1} A_n) = \sum_{n\geq 1} \mu(A_n)$ если $\forall n\ A_n \in \mathcal{A}$, $\forall m, n\ A_m \cap A_n = \emptyset$ и $\sum_{n\geq 1} A_n \in \mathcal{A}$,

2. (ради[26] аксиом Стоуна §13.12) множество Ω – σ-конечно, т.е. представимо в виде счётного объединения суммируемых множеств:
$$\Omega = \bigcup_{n\geq 1} \Omega_n, \quad \forall n\ \Omega_n \subset \Omega_{n+1} \in \mathcal{A}, \quad \mu(\Omega_n) < \infty. \tag{51}$$

Определение: (Сравнить со стр. 19.) Если на $\mathcal{A}$ определена мера μ, то множества $A \in \mathcal{A}$ называются **($\mu-$)измеримыми.** А если вдобавок ещё и $\mu(A) < \infty$, то они называются **($\mu-$)суммируемыми.**

Определение: Мера μ, заданная на σ-алгебре $\mathcal{A}$, называется **полной**, если все подмножества меры нуль $[\mu(A) = 0]$ принадлежат $\mathcal{A}$.

Пример: вдобавок к Ω задаём элементарный интеграл I и пространство элементарных функций H. Тогда система всех измеримых (суммируемых) множеств образует σ-кольцо. Предположим, что I удовлетворяет аксиомам Стоуна из §13.12; тогда μ – (полная) мера на σ-алгебре системы измеримых множеств. Если $\mu(\Omega) < \infty$, то суммируемые множества тоже образуют σ-алгебру. ... См. стр.[7]144.

Определение: Пусть имеется два кольца $\mathcal{A}_1$ и $\mathcal{A}_2$, а также меры μ_1 и μ_2 на них. Говорят, что мера μ_1 **продолжает меру** μ_2 ($\mu_1 \supset \mu_2$, $\mu_1 \succ \mu_2$, $\mu_1 > \mu_2$), если $\mathcal{A}_1 \supset \mathcal{A}_2$ и $\forall A \in \mathcal{A}_2\ \mu_1(A) = \mu_2(A)$.

Определение: полная мера $\bar{\mu}$, определённая на σ-алгебре $\bar{\mathcal{A}}$, называется **лебеговым продолжением** меры μ, определённой на кольце $\mathcal{A}$, если выполняются два условия:

1. $\bar{\mu} \supset \mu$;

2. если $\tilde{\mu}$ – некоторая полная мера и $\tilde{\mu} \supset \mu$, то $\tilde{\mu} \supset \bar{\mu}$. (Т.е. $\bar{\mu}$ – наименьшая полная мера такая, что $\bar{\mu} \supset \mu$.)

Из определения следует, что лебегово продолжение меры (если оно вообще существует) единственно.

15.7 Теорема о продолжении меры

Лемма: допустим, что имеется элементарная мера μ на кольце $\mathcal{A}$ на Ω. Рассмотрим пространство элементарных функций $H = \{h(\omega) = \sum_{j=1}^{m} \alpha_j \chi_{E_j}(\omega) | E_j \in \mathcal{A}\}$, и определим интеграл $\forall h \in H\ Ih = \sum_{j=1}^{m} \alpha_j \mu(E_j)$. Доказываемое утверждение: в таком случае (H, I) удовлетворяют всем[27] аксиомам элементарных функций/интегралов, а также аксиомам Стоуна. Доказательство: ...

Теорема 15.8 *Для любой элементарной меры μ на кольце $\mathcal{A}$ существует полная мера $\bar{\mu} > \mu$ продолжающая μ и определённая на σ-алгебре $\bar{\mathcal{A}}$.*

Замечания к теореме 15.8:

1. Любая элементарная мера приводит к интегралу.

2. Продолжений интеграла и меры может быть много; можно доказать,[28] что построенная в теореме 15.8 мера $\bar{\mu}$ есть лебеговское продолжение меры μ.

3. Допустим, β – наименьшая σ-алгебра, порождённая кольцом $\mathcal{A}$. $\beta \subset \bar{\mathcal{A}}$. Можно доказать (теорема Каратеодори), что это продолжение единственно.

4. Можно показать, что $\bar{\mathcal{A}}$ из теоремы 15.8 устроено следующим образом: к любому множеству $B \in \beta$ надо добавить всевозможные подмножества $\tilde{B} \subset B \in \beta$, такое, что ... $\clubsuit \mu(B_0) = 0 \Longrightarrow A = B \cup \tilde{B}$.

§2. Случайные величины

[26] Я пока не понимаю, как из (51) следует какая-либо из аксиом Стоуна.
[27] Вставить ссылки...
[28] См.[7]; в доказательстве используется теория групп.

15.8 Случайные величины

В дальнейшем, рассматривая $(\Omega, \mathcal{A}, P)$, будем считать, что $\mathcal{A}$ – σ-алгебра. Замечания в конце предыдущего вопроса показывают, что этим мы ничуть не теряем в общности.

Определение: Борелевской (σ-)алгеброй $\mathcal{B} \equiv \mathcal{B}(\mathbb{R})$ назовём наименьшую σ-алгебру подмножеств $\subset \mathbb{R}$, содержащую все интервалы $(a, b) \subset \bar{\mathbb{R}}$.

Примеры подмножеств $\subset \mathcal{B}$:

1. Если задана топология: то все открытые подмножества $\tau \subset \mathcal{B}$. Потому что любое открытое множество представимо в виде счётного числа интервалов – см. тему “Многообразия со счётным атласом”.
2. Все закрытые подмножества $\subset \mathcal{B}$ (очевидно из предыдущего примера).
3. Если точка принадлежит $\mathcal{B}$, то и любое нбчс множество точек тоже ему принадлежит.
4. Множество иррациональных чисел $\in \mathcal{B}$, потому что оно является дополнением ко (счётному) множеству рациональных чисел.

Определение: **Случайной величиной** (СВ) называется функция

$$\xi : \Omega \to \mathbb{R} \quad \text{обладающая свойством} \quad \forall B \in \mathcal{B} \quad \xi^{-1}(B) \in \mathcal{A}, \tag{52}$$

т.е. прообраз любого борелевского множества является событием. Является ли функция ξ изоморфизмом? – нет: в этом можно убедиться на простейшем примере неслучайной величины, которая с единичной вероятностью равна некоторому фиксированному числу.

STOP Можно ли считать борелевские множества обобщениями интервалов? И неплохо-бы привести пример функции, для которой не выполняется (52).

Свойства случайных величин:

1. ξ есть СВ $\iff \forall (a, b)\ \xi^{-1}[(a, b)] \in \mathcal{A}$, т.е. если свойство в определении (52) выполняется для интервалов (a, b), то и для прочих $B \in \mathcal{B}$ оно тоже выполняется. Доказательство: $\Longrightarrow$ очевидно, доказываем $\Longleftarrow$.
   Мы пока не уверены в том, что ξ – св, поэтому не можем утверждать, что ξ-прообраз любого подмножества $B \in \mathcal{B}$ является событием. Тем не менее для некоторых $B \in \mathcal{B}$ это справедливо. Обозначим совокупность таких множеств $\tilde{\mathcal{B}} = \{B \in \mathcal{B} | \xi^{-1}(B) \in \mathcal{A}\}$. По определению $\tilde{\mathcal{B}}$ не может быть больше, чем $\mathcal{B}$. Нам надо доказать, что оно не меньше, чем $\mathcal{B}$; эта цель будет достигнута, если мы докажем, что $\tilde{\mathcal{B}}$ есть σ-алгебра.

   К примеру, покажем, что если $B \in \tilde{\mathcal{B}}$, то и $\bar{B} \in \tilde{\mathcal{B}}$. (Заметим, что т.к. $B \in \mathcal{B}$ и по условию $\mathcal{B}$ – σ-алгебра, то $\bar{B} \in \mathcal{B}$.)

   Для доказательства выберем некоторые множества $B, B_1, B_2 \ldots \in \tilde{\mathcal{B}}$ и покажем, что множества $\bar{B}$, $\bigcap\limits_{n\geq 1} B_n$, и $\bigcup\limits_{n\geq 1} B_n$ тоже принадлежат $\tilde{\mathcal{B}}$. В самом деле, $\mathcal{A}$ – σ-алгебра, поэтому если $A = \xi^{-1}(B) \in \mathcal{A}$, то и $\bar{A} = \xi^{-1}(\bar{B}) = \overline{\xi^{-1}(B)} \in \mathcal{A}$:

$$\xi^{-1}(\bar{B}) = \{\omega \in \Omega | \xi(\omega) \notin B\} = \Omega \backslash \xi^{-1}(B) = \overline{\xi^{-1}(B)} \in \mathcal{A}.$$

   Аналогично $\xi^{-1}(\bigcap\limits_{n\geq 1} B_n) = \bigcap\limits_{n\geq 1} \xi^{-1}(B_n) \in \mathcal{A}$ и $\xi^{-1}(\bigcup\limits_{n\geq 1} B_n) = \bigcup\limits_{n\geq 1} \xi^{-1}(B_n) \in \mathcal{A}$. $\Longrightarrow \bar{B}$, $\bigcap\limits_{n\geq 1} B_n$ и $\bigcup\limits_{n\geq 1} B_n$ принадлежат $\tilde{\mathcal{B}}$; следовательно $\tilde{\mathcal{B}}$ есть σ-алгебра. $\Longrightarrow \forall (a, b) \in \mathcal{B} \Longrightarrow \tilde{\mathcal{B}} = \mathcal{B}$, т.к. $\mathcal{B}$ – наименьшая σ-алгебра, содержащая все такие интервалы.

2. $\mathcal{B}$ – наименьшая σ-алгебра, содержащая все полубесконечные интервалы вида $(-\infty, c)$ или (c, ∞). Доказательство для $(-\infty, c)$: Построим $\tilde{\mathcal{B}}$ для $(-\infty, c)$ аналогично тому, как мы строили его для (a, b) в предыдущем (1.) свойстве (несколькими строчками выше): $\tilde{\mathcal{B}} = \{(-\infty, c) \in \mathcal{B} | \xi^{-1}[(-\infty, c)] \in \mathcal{A}\}$. Т.к. $\forall c \in \mathbb{R}$ множество $(-\infty, c)$ – открытое, то $(-\infty, c) \in \mathcal{B} \Longrightarrow \tilde{\mathcal{B}} \subset \mathcal{B}$. Теперь докажем, что $\mathcal{B} \subset \tilde{\mathcal{B}}$. В самом деле, $(a, b) = (-\infty, b) \backslash \bigcap\limits_{n\geq 1} (-\infty, a + 1/n) \in \tilde{\mathcal{B}} \Longrightarrow \mathcal{B} \subset \tilde{\mathcal{B}}$.

3. Из двух предыдущих свойств следует, что если ξ – СВ, то $\forall c\ \xi^{-1}[(-\infty, c)] \in \mathcal{A}$ и $\xi^{-1}[(c, \infty)] \in \mathcal{A}$.

4. если ξ – СВ, то $c\xi$ и $c + \xi$ – также СВ. Доказательство: случай $c = 0$ очевиден, рассм. $c \neq 0$. $\{\omega \in \Omega | c\xi(\omega) < b\} \equiv \{c\xi < b\} = (c\xi)^{-1}[(-\infty, b)] = \{\xi < b/c\}$ при $c > 0$ или $= \{\xi > -b/c\}$ при $c < 0$; Оба множества $\{\xi < b/c\}$ и $\{\xi > -b/c\}$ являются событиями, т.е. принадлежат $\mathcal{A}$.

5. если ξ и η – СВ, то $\{\xi > \eta\} \equiv \{\omega \in \Omega | \xi(\omega) > \eta(\omega)\} \in \mathcal{A}$. Доказательство: пронумеруем множество рациональных чисел $\mathbb{Q} = \{r_k\}_{k\in\mathbb{N}}$. $\forall \alpha, \beta \in \mathbb{R}$, $\alpha < \beta\ \exists p \in \mathbb{Q}$: $\alpha < p < \beta$.

$$\{\xi > \eta\} = \bigcup_{k\in\mathbb{N}} (\{\xi(\omega) > r_k\} \cap \{\eta(\omega) < r_k\}) \in \mathcal{A}. \tag{53}$$

6. если ξ и η – СВ, то $\xi \pm \eta$ – СВ. Доказательство: ...
7. $f \in \mathbb{C}(\mathbb{R}) \Longrightarrow f(\xi) \equiv f \circ \xi$ – СВ. Доказательство: ...
8. ξ^2, $|\xi|$ и $\xi\eta$ – СВ.
9. Если $\forall\omega\ \xi(\omega) \neq 0$ то $1/\xi$ – СВ.
10. Про inf и sup – СВ.
11. Предел последовательности СВ есть СВ.

15.9 Числовые характеристики случайных величин

Случайные величины – как мы уже знаем, функции. Оказывается, их можно интегрировать. Простейший интеграл – т.н. математическое ожидание. Из нашего исходного вероятностного пространства $(\Omega, \mathcal{A}, P)$ с элементарной мерой μ построим по схеме Даниэля,[29] $\bar{\mathcal{A}}$, $\bar{\mu}\Big|_{\mathcal{A}}$, I, L. Тогда если ξ – СВ, то $\{\xi > c\} \in \mathcal{A} \subset \bar{\mathcal{A}}$. Из теоремы 13.6 о характеризации измеримых функций следует, что функция ξ I-измерима:

$$\xi \in L \Rightarrow \exists I\xi = \int_\Omega \xi(\omega)\, \underbrace{\bar{\mu}(\mathrm{d}\omega)}_{\equiv \mathrm{d}\mu(\omega)} \stackrel{\mathrm{df}}{=} M\xi \tag{54}$$

– математическое ожидание ξ. Математическое ожидание обладает обычными свойствами интеграла.

Согласно определению интеграла Лебега по Лебегу,

$$M\xi = \lim_{k\to\infty} \sum_{n\geq 1} \Big[\bar{\mu}(\underbrace{\{\frac{n}{k} < \xi(\omega) \leq \frac{n+1}{k}\}}_{\in\mathcal{A}}) - \bar{\mu}(\underbrace{\{-\frac{n+1}{k} \geq \xi(\omega) < -\frac{n}{k}\}}_{\in\mathcal{A}})\Big]. \tag{55}$$

Т.к. ξ – с.в., то множества эти содержаться не только в $\bar{\mathcal{A}}$, но и в $\mathcal{A}$. Следовательно,

$$M\xi = \lim_{k\to\infty} \sum_{n\geq 1} \frac{n}{k} \left[P\left\{\frac{n}{k} < \xi \leq \frac{n+1}{k}\right\} - P\left\{-\frac{n+1}{k} \geq \xi < -\frac{n}{k}\right\}\right] = \int_\Omega \xi(\omega) P(\mathrm{d}\omega). \tag{56}$$

Упражнение: в качестве случайной величины рассмотрим число успехов μ_n в испытаниях Бернулли (с вероятностями успеха/неудачи p и $q = 1 - p$). Найдём[30] математическое ожидание $\mu_n = np$

$$\clubsuit M\mu_n = np \ldots \tag{57}$$

Из свойств на стр. 30 следует корректность следующих определений: k-тый момент СВ $M\xi^k$, абсолютный k-тый момент СВ $M|\xi|^k$, центральный k-тый момент СВ $M(\xi - M\xi)^k$, абсолютный центральный k-тый момент СВ $M|\xi - M\xi|^k$, дисперсия (второй центральный момент).

Определение: **дисперсией** называется $D\xi \stackrel{\mathrm{df}}{=} M(\xi - M\xi)^2 \geq 0$. Логично величину, равную с единичной вероятностью некоторой постоянной, неслучайной. Соответственно, неслучайная величина обладает нулевой дисперсией $D\xi = 0$.

Свойства:
1. $M(\xi - M\xi)^2 = M\xi^2 - (M\xi)^2$.
2. $D(a\xi + b) = a^2 D\xi$.

Неравенство Маркова: для неотрицательных СВ: $P\{\xi \geq a\} \leq M\xi/a$. Доказательство:

$$\chi\left(\{\omega \in \Omega | \xi(\omega) \leq a\}\right) \leq \xi(\omega)/a \leq 1. \tag{58}$$

Пример: в среднем студенты опаздывают на лекцию на $M\xi = 5$ минут. Грубая оценка опоздания очередного студента на $a =$ 20 минут есть $P\{\xi \geq 20\} \leq 5/20 = 1/4$.

Неравенство Чебышева: $P\{|\xi - M\xi| \geq \varepsilon\} \leq \sigma^2/\varepsilon^2$, где $\sigma = \sqrt{D\xi}$ есть **среднеквадратичное отклонение.** Доказательство: применим неравенство Маркова (58) к СВ $(\xi - M\xi)^2$ и $a = \varepsilon^2$.

Дополнение:
1. Медиана $\mathcal{M}$ определяется в соответствии с уравнением $P\{\xi < \mathcal{M}\} = P\{\xi > \mathcal{M}\}$.
2. Модой СВ называется её наиболее вероятное значение (координата максимума ф.р.).

[29] См. сек. 13.5, а также теорему 15.8.
[30] Вероятно, делать это следует в духе (63).

15.10 Распределения случайных величин

Отныне займёмся многомерными случайными величинами, а потому переопределим некоторые понятия.

Определения:

1. Наименьшая σ-алгебра подмножеств в $\mathbb{R}^n$, содержащая все открытые параллелепипеды $\Pi = (a_1 \dots b_1) \times \dots \times (a_n \dots b_n)$, называется **борелевской σ-алгеброй** $\beta \equiv \beta(\mathbb{R}^n) \equiv \beta^n$.

2. Рассматриваем вероятностное пространство $(\Omega, \mathcal{A}, P)$; **случайной величиной** называется функция

$$\xi : \Omega \to \mathbb{R}^n, \quad \text{обладающая свойством:} \quad \forall B \in \beta^n \subset \mathbb{R}^n \quad \xi^{-1}(B) \in \mathcal{A}. \tag{59}$$

   Естественно записывать наши многомерные случайные величины покомпонентно: $\xi = (\xi_1 \dots \xi_n)$.

3. Как и в одномерном случае, вероятность индуцирует меру на β. По определению $\forall B \in \beta^n$ положим $\mu_\xi(B) = P\{\xi \in B\}$; аксиомы меры легко проверяются.[31] Из-за дополнительного свойства $\mu_\xi(\mathbb{R}^n) = 1$ называем эту меру **вероятностной**. Мера μ_ξ называется **распределением** случайной величины ξ.

4. Функция $f : \mathbb{R}^n \to \mathbb{R}$ называется **измеримой по Борелю**, если $\forall B \in \beta^1$ $f^{-1}(B) \in \beta^n$. Запас таких функций огромен: они получаются счётным числом предельных переходов в последовательностях измеримых функций.

Докажем, что любая непрерывная функция измерима по Борелю: $f \in \mathbb{C}(\mathbb{R}^n) \Longrightarrow f \in \beta^n$. Доказательство:...

Теорема 15.9 *Рассмотрим СВ $\xi : \Omega \to \mathbb{R}^n$, а также измеримую по Борелю функцию (и.ф.) $f : \mathbb{R}^n \to \mathbb{R}$. Предположим, что существует математическое ожидание $Mf(\xi)$. Тогда*

$$Mf(\xi) = \int_{\mathbb{R}^n} f(\vec{x})\mu_\xi(\mathrm{d}\vec{x}), \tag{60}$$

причём обе части (не)существуют одновременно.

Доказательство: по Лебегу

$$\begin{aligned} Mf(\xi) &= \lim_{k\to\infty} \sum_{n\ge1} \frac{n}{k} \left[P\{\frac{n}{k} < f(\xi(\omega)) \le \frac{n+1}{k}\} - P\{-\frac{n+1}{k} \le f(\xi(\omega)) < \frac{n}{k}\} \right] = \\ &= \lim_{k\to\infty} \sum_{n\ge1} \frac{n}{k} \left[P\{\xi(\omega) \in \underbrace{f^{-1}\left((\frac{n}{k}, \frac{n+1}{k}] \right)}_{\in\beta_n}\} - P\{\xi(\omega) \in f^{-1}\left([-\frac{n+1}{k}, -\frac{n}{k}) \right)\} \right] = \\ &= \lim_{k\to\infty} \sum_{n\ge1} \frac{n}{k} \left[\mu_\xi(f^{-1}\left((\frac{n}{k}, \frac{n+1}{k}] \right))\} - \mu_\xi(f^{-1}\left([-\frac{n+1}{k}, -\frac{n}{k}) \right)) \right] = \\ &= \lim_{k\to\infty} \sum_{n\ge1} \frac{n}{k} \left[\mu_\xi(\frac{n}{k} < f(\vec{x}) \le \frac{n+1}{k}) - \mu_\xi(-\frac{n+1}{k} \le< f(\vec{x}) < -\frac{n}{k}) \right] \overset{\mathrm{df}}{=} \int_{\mathbb{R}^n} f(\vec{x})\mu_\xi(\mathrm{d}\vec{x}). \end{aligned} \tag{61}$$

К примеру, из (60) следуют интегральные формулы для моментов:

$$M\xi = \int_{-\infty}^{\infty} x\mu_\xi(\mathrm{d}x), \quad D\xi = \int_{-\infty}^{\infty} (x - M\xi)^2\mu_\xi(\mathrm{d}x), \quad M(\xi - M\xi)^k = \int_{-\infty}^{\infty} (x - M\xi)^k\mu_\xi(\mathrm{d}x). \tag{62}$$

15.11 Частные случаи распределений

Дискретное распределение. Рассмотрим СВ $\xi : \Omega \to \mathbb{R}^n$ и её распределение μ_ξ.

Определение: СВ ξ называется дискретной, если её распределение сосредоточено на нбчс множестве точек $Q = \{\vec{x}_i\} \subset \mathbb{R}^n$. Любое нбчс множество – борелевское[32]; $\mu_\xi(Q) = 1 \Longrightarrow \mu_\xi(\bar{Q}) = 0$. Каждой точке $\vec{x}_i \in Q$ сопоставим число $p_i = \mu_\xi(\vec{x}_i) \equiv \mu_\xi(\{\vec{x}_i\})$.

[31] Возможно, под "асиомами меры" имеется в виду свойство 13.5. Мы определяли интеграл по Даниелю, поэтому свойства меры мы доказывали, как с в случае с теоремой 13.5. Возможно (это подтверждает интернет, хотя Калинин нам этого не рассказывал), в подходе Лебега мера определяется аксиоматически, и тогда 13.5 из теоремы превращается в аксиому.

[32] Как множество может быть борелевским, если оно не содержит интервалов? Вариант ответа: в дискретной топологии все точки являются одновременно открытыми и закрытыми множествами, так что мы можем вместо интервалов использовать множества, состоящие из одной или нескольких точек. Ещё одно замечание: мы определяли борелевскую σ-алгебру, но не борелевское множество. Это – разные понятия?

$\Longrightarrow \forall C \in \beta^n\ \mu_\xi(C) = \mu_\xi(CQ + C\bar{Q}) = \mu_\xi(CQ) = \sum\limits_{\vec{x}_i \in C} \mu_\xi(\{\vec{x}_i\}) = \sum\limits_{\vec{x}_i \in C} p_i$, т.е. СВ с единичной вероятностью принимает значения из Q, причём $p\{\xi = \vec{x}_i\} \overset{\text{df}}{=} \mu_\xi(\{\vec{x}_i\}) = p_i$. То есть СВ может принимать значения не из Q только с нулевой вероятностью, а значит, можно сказать, что никаких иных значений, кроме $\vec{x}_i$, наша СВ не принимает. Для математического ожидания получаем

$$M\xi = \int_{\mathbb{R}^n} f(\vec{x}\,)\mu_\xi(\mathrm{d}\vec{x}\,) = \int_Q f(\vec{x}\,)\mu_\xi(\mathrm{d}\vec{x}\,) + \underbrace{\int_{\bar{Q}} f(\vec{x}\,)\mu_\xi(\mathrm{d}\vec{x}\,)}_{=0 \Longleftarrow \mu_\xi(\bar{Q})=0} = \sum_i \int_{\{\vec{x}_i\}} f(\vec{x}\,)\mu_\xi(\mathrm{d}\vec{x}\,) = \sum_i f(\vec{x}_i)\mu_\xi(\vec{x}_i) = \sum_i f(\vec{x}_i)p_i. \tag{63}$$

Очевидно: существование $M\xi$ эквивалентно ($\Longleftrightarrow$) абсолютной сходимости ряда (63).

Следствие: в одномерном случае

$$n = 1: \quad x_i \in \mathbb{R}, \quad p_i = P\{\xi = x_i\}, \quad M\xi = \sum_{i \ge 1} x_i p_i, \quad D\xi = \sum_{i \ge 1} (x_i - M\xi)^2 p_i. \tag{64}$$

Пример: распределение по закону Пуассона: СВ принимает значения $\xi \in \mathbb{N} \cup \{0\}$ и $P\{\xi = k\} = \lambda^k e^{-\lambda}/k!$, $k > 0$. $\Longrightarrow$ $M\xi = \sum\limits_{k \ge 0} k\lambda^k e^{-\lambda}/k! = \sum\limits_{k \ge 1} \lambda^k e^{-\lambda}/(k-1)! = \lambda e^{-\lambda} \sum\limits_{k \ge 0} \lambda^k/k! = \lambda$. Чтобы посчитать дисперсию, сначала посчитаем $M\xi(\xi - 1) = M\xi^2 - M\xi = \sum\limits_{k \ge 0} k(k-1)\lambda^k e^{-\lambda}/k! = \lambda^2 e^{-\lambda} \sum\limits_{k \ge 0} \lambda^k/k! = \lambda^2 \Longrightarrow D\xi = M\xi^2 - (M\xi)^2 = (\lambda^2 + \lambda) - \lambda^2 = \lambda$.

Плотность распределения. Говорят, что СВ $\xi : \Omega \to \mathbb{R}^n$ имеет плотность распределения или абсолютно непрерывна, если существует β-измеримая функция $p_\xi(\vec{x}\,) : \mathbb{R}^n \to \mathbb{R}$ такая, что

$$\forall B \in \beta^n \quad \mu_\xi(B) \overset{\text{df}}{=} P\{\xi^{-1}(B)\} = \int_B p_\xi(\vec{x}\,)\mathrm{d}\vec{x}. \tag{65}$$

Заметим, что $\mathrm{d}\vec{x}$ – стандартная мера Лебега в $\mathbb{R}^n$, получаемая из элементарного интеграла Римана.

Замечания:

1. (См. §[28]5.4) В то время как функция распределения универсальна (существует для любых величин), плотностью распределения обладают только непрерывные величины.
2. (См. §[28]5.3) Рассмотрим непрерывную СВ ξ принимающую произвольные значения на некотором отрезке $[a, b]$. Для соответствующей меры любая точка $x_0 \in [a, b]$ – ММ0, так что вероятность любого значения этой СВ равна нулю. Таким образом, из того, что вероятность события $\{\xi = x_0\}$ равна нулю, ещё не следует невозможность этого события. Иными словами, вероятность события может быть равна нулю, но его частота – не равна нулю. А всё потому, что частота события при большом количестве опытов *приближается* к вероятности, но не обязательно равна ей. Аналогично, хотя $P\{\xi \ne x_0\} = 1$, но событие это не достоверно.
3. Очевидно, что все измеримые (по Борелю) функции в $\mathbb{R}^n$ измеримы по отношению к стандартному (евклидовому) интегралу Лебега на $\mathbb{R}^n$.
4. Можно доказать (но не нам), что, если плотность p_ξ существует, то она однозначно определяется с точностью до лебеговой меры ноль, если потребовать выполнение (65) не для произвольных $\forall B \in \beta^n$, но для одних лишь только параллелепипедов.

Свойства:

1. Плотность распределения может принимать отрицательные значения только на множестве меры нуль. Доказательство...
2. $\int_{\mathbb{R}^n} p_\xi(\vec{x}\,)\mathrm{d}\vec{x} = P\{\xi \in \mathbb{R}^n\} = 1$. Доказательство...

Теорема: если СВ $\xi : \Omega \to \mathbb{R}^n$ имеет плотность распределения p_ξ и f – измерима по Борелю, то одновременно (не)существуют и равны (если существуют)

$$Mf(\xi) = \int_{\mathbb{R}^n} p_\xi(\vec{x}\,)f(\vec{x}\,)\mathrm{d}\vec{x}. \tag{66}$$

Доказательство...

Например [сравнить с (62) и (64)], в одномерном случае математическое ожидание и дисперсия равны

$$M\xi = \int_{-\infty}^{\infty} x p_\xi(x)\mathrm{d}x, \quad D\xi = \int_{-\infty}^{\infty} (x - M\xi)^2 p_\xi(x)\mathrm{d}x. \tag{67}$$

Определение: одномерная СВ имеет **равномерное распределение** на интервале $[a,b]$, если

$$p_\xi(x) = \begin{cases} c, x \in [a,b] \\ 0, x \notin [a,b], \end{cases} \quad \int_{-\infty}^{\infty} p_\xi(x)\mathrm{d}x = 1 \Longrightarrow c = \frac{1}{b-a}, \quad M\xi = \int_a^b \frac{x}{b-a} = \frac{a+b}{2}, \quad D\xi = \frac{b-a}{12}. \tag{68}$$

Пример: предположим, что распределение μ_ξ СВ ξ имеет вид $\mu_\xi = p\mu_1 + q\mu_2$, $p+q=1$, $p,q \geq 0$, μ_1 и μ_2 – вероятностые меры на β^n. Причём μ_1 сосредоточена на некотором счётном множестве $Q = \{x_i\}$, $\mu_1(\vec{x}_i) = p_i$, а μ_2 – непрерывна и имеет плотность $p(\vec{x})$. Это значит, что $\forall B \in \beta^n$ $\mu_\xi(B) = p\mu_1(B) + q\mu_2(B)$. Из определения интеграла Лебега по Лебегу следует, что

$$Mf(\xi) = p\sum_i f(\vec{x}_i)\mu_1(\vec{x}_i) + q\int_{\mathbb{R}^n} f(\vec{x}_i)p_\xi(\vec{x}_i). \tag{69}$$

15.12 Функция и плотность распределения в одномерном случае

Рассмотрим одномерную СВ $\xi : \Omega \to \mathbb{R}$, а также её распределение μ_ξ.

Определение: **функцией распределения** СВ ξ называется $F_\xi(x) = P\{\xi < x\} = \mu_\xi[(-\infty, x)]$.

Свойства:

1. функция распределения не убывает. Доказательство: ...
2. $$F_\xi(-\infty) = \ldots = 0.$$
   Аналогично $F_\xi(+\infty) = 1$.
3. $F_\xi(x)$ непрерывна слева:$F_\xi(x-0) = F_\xi(x)$.

Замечание 15.9.1 *любая функция, удовлетворяющая трём вышеперечисленным свойствам, является функцей распределения.*

Доказательство (точнее, его набросок): определим интеграл Римана-Стильтьеса $If = \int_{-\infty}^{\infty} f(x)\mathrm{d}F(x)$ следующим образом (вспоминаем вопрос 13.1). В качестве пространства элементарных функций выберем $H = \mathcal{K}(\mathbb{R})$, на H определим элементарный интеграл Римана-Стильтьеса

$$h \in H \Rightarrow Ih = \int_{[a,b]\supset \operatorname{Supp} h} h(x)\mathrm{d}F(x). \tag{70}$$

Продолжим его. Пользуясь свойствами функции F, надо доказать, что $I1 = 1$.
Теперь определяем вероятностное пространство:

$$\Omega = \mathbb{R}, \quad \mathcal{A} = \beta, \quad P(A) = \int_A \mathrm{d}F(x), \quad \xi(x) = x. \tag{71}$$

Отсюда легко проверяется, что для любой β-измеримой функции h

$$\int_{-\infty}^{\infty} h(x)\mathrm{d}F(x) = \int_{-\infty}^{\infty} h(x)\mu_\xi(\mathrm{d}x). \tag{72}$$

Замечание 15.9.2 *Можно также доказать, что через функцию распределения выражаются вероятности всех событий. Перечислим конкретные примеры этого:*

1. *Очевидно, что* $P\{a \leq \xi < b\} = F_\xi(b) - F_\xi(a)$.
2. $P\{\xi = a\} = \ldots = F_\xi(a+0) - F_\xi(a)$.
3. $P\{a < \xi < b\} = \ldots = F_\xi(b) - F_\xi(a+0)$.

Замечания:

1. Если некоторая одномерная СВ ξ имеет плотность распределения $p_\xi(x)$, то из определения следует, что
   $$F_\xi(x) = P\{\xi < x\} = \int_{-\infty}^{x} p_\xi(t)\mathrm{d}t \in \mathbb{C}(\mathbb{R}) \tag{73}$$
   (непрерывность легко доказывается). Непрерывность позволяет восстановить плотность по функции распределения: $F'_\xi(x) = p_\xi(x)$.
   (Проверка осуществляется так же, как и в теореме Барроу.)
   Можно (но не нам) также доказать теорему Лебега: почти всюду по мере Лебега $\exists F'_\xi(x) = p_\xi(x)$.

2. Можно доказать, что если СВ ξ – дискретна с множеством значений $\{c_i\}_{i\geq 1}$, то $P\{\xi = c_i\} = F_\xi(c_i+0) - F_\xi(c_i) = p_i > 0$ ($p_i > 0$ в случае, если $c_i \in \mathbb{R}$), причём $\sum_{i\geq 1} p_i = 1$. Если $c_i < c_{i+1}$, то на полуинтервале $(c_i, c_{i+1}]$ функция распределения постоянна.

Пример: модель радиоактивного распада – **показательное распределение**. (См. также стр. [24]166-167, но там говорится о пуассоновском распределении – непонятно.) Изучаем процесс превращения радия (Ra) в радон (Rn), сопровождающийся излучением α-частицы. В качестве случайной величины рассмотрим момент времени τ, когда происходит превращение конкретного атома Ra в Rn. (Атом рассматривается начиная с момента времени $t = 0$.)

$F_\tau(t) = P(\tau < t)$. Вместо $F_\tau(t)$ удобнее рассматривать $Q_\tau(t) = 1 - F_\tau(t) = P\{\tau \geq t\}$.

Предполагаем

- гладкость $Q_\tau(t) \in \mathbb{C}^{(1)}$ ради дифференцируемости, а также
- справедливость физической модели для условной вероятности $P\{\{\tau \in [a,b]\}|\{\tau \geq a\}\} = \lambda(b-a) + o(b-a)$, $\lambda > 0$. – вероятность распада в заданном промежутке времени при условии, что до этого распад не произошёл.[33]

$$\begin{aligned} P\{t \geq t+\delta t\} &= P\{\tau \geq t\} \cdot P\{\{\tau \geq t+\delta t\}|\{\tau \geq t\}\}, \\ Q_\tau(t+\delta t) &= Q_\tau(t)(1 - P\{\{\tau \in [t, t+\delta t]\}|\{\tau \geq t\}\}) = Q_\tau(t)(1 - \lambda \cdot \delta t + o(\delta t)). \end{aligned} \tag{74}$$

Решаем элементарное дифференциальное уравнение $\Longrightarrow F_\tau(t) = 1 - e^{-\lambda t}$ при $t \geq 0$. Про такие СВ говорят, что они имеют показательное распределение. Легко показать, что $M\tau = 1/\lambda$ и $D\tau = 1/\lambda^2$.

15.13 Плотность распределения в многомерном случае

Рассмотрим случайную величину $\xi = (\xi_1 \dots \xi_n) : \Omega \to \mathbb{R}^n \Longrightarrow$ определим **функцию распределения**

$$F_\xi(x_1 \dots x_n) = P(\xi_1 < x_1 \dots \xi_n < x_n) = P(\prod_{i=1}^n \{\xi_i < x_i\}). \tag{75}$$

Эта функция распределения обладает обычными свойствами; например, по ней можно восстановить μ_ξ. Раз уж мы работаем с многомерным пространством $\mathbb{R}^n$, то встаёт вопрос о том, как меняются наши величины при замене координат. В этом смысле функция распределения неудобна, потому что она привязана к некоторой конкретной системе координат и часто её невозможно пересчитать для других систем координат. Будем предполагать, что рассматриваемая св ξ имеет плотность $p_\xi(\vec{x})$.

Теорема 15.10 *предположим, имеется диффеоморфизм $\Phi : \mathbb{R}^n \to \mathbb{R}^n \Longrightarrow \eta = \Phi \circ \xi \equiv \Phi(\xi)$ – тоже св (где обозначение $\Phi(\xi)$ специфично для теории вероятностей) и эта св имеет плотность распределения $p_\eta(\vec{x}) = p_\xi(\Phi^{-1}(\vec{x})) \cdot |\det(\Phi^{-1})'(\vec{x})|$.*

Доказательство...

Подвопросик: 1.Предположим, св $\xi = (\xi_1 \dots \xi_n)$ имеет плотность распределения $p_\xi(x_1 \dots x_n)$. Как, зная $p_\xi(x_1 \dots x_n)$, получить плотность распределения $p_{\xi_i}(x_i)$ или, к примеру, $p_{\xi_{i_1} \dots \xi_{i_k}}(x_{i_1} \dots x_{i_k})$ имея в виду, что $1 \leq i_1 < \dots < i_k \leq n$ – т.н. **маргинальные плотности распределения**

доказательство ... $\Longrightarrow p_{\xi_1}(x_1) = \int_{\mathbb{R}^{n-1}} p_\xi(x_1 \dots x_n) \mathrm{d}x_1 \dots \mathrm{d}x_n$ STOP Блин, аналогия с сокращённым описанием в стат. физике – просто один в один!

2. Предположим, св $\xi = (\xi_1, \xi_2)$ имеет плотность распределения $p_\xi(x_1, x_2)$. Как распределена сумма $\xi_1 + \xi_2$? Выберем в качестве диффеоморфизма в $\mathbb{R}^2$ $\Phi(x_1, x_2) = (x_1 + x_2, x_2)$ доказательство... $p_{\xi_1+\xi_2} = p_{\eta_1}(x_1) = \int_{-\infty}^{\infty} p_\xi(x_1 - x_2, x_2) \mathrm{d}x_2$.

15.14 Независимость СВ

Рассмотрим СВ св $\xi = (\xi_1 \dots \xi_n)$. Величины св $\xi_1 \dots \xi_n$ называются независимыми (в совокупности), если $\forall A_1 \dots A_n \in \beta_1 \equiv \beta(\mathbb{R})$ события $\{\xi_i \in A_i\}$ $i = 1 \dots n$ независимы.

Лирическое отступление: СВ ξ_i определяет меру μ_{ξ_i} на β_1, а СВ $\xi = (\xi_1 \dots \xi_n)$ определяет меру μ_ξ на β_n. Наших знаний теории меры недостаточно для доказательства того, что независимость соответствует декартовому произведению мер $\mu_\xi = \mu_{\xi_1} \times \dots \times \mu_{\xi_n}$.

Теорема 15.11 *Если СВ $\xi_1 \dots \xi_n$ независимы, и $f_1 \dots f_n$ – некоторые β-измеримые функции $f_i : \mathbb{R} \to \mathbb{R}$, то СВ $f_1(\xi_1) \dots f_n(\xi_n)$ – также независимы.*

Доказательство ...

Теорема 15.12 *Если плотность распределения $p_\xi(x_1 \dots x_n) = \prod_{i=1}^n p_{\xi_i}(x_i)$, то СВ $\xi_1 \dots \xi_n$ независимы. (Если бы мы лучше знали меру, могли бы доказать и обратное.)*

[33] Мы хотим отсчитывать время, начиная с момента $\tau = a$, и предполагаем, что до этого событие распада не происходило.

Доказательство ...

Пример: распределение суммы независимых СВ есть свёртка (см. конец предыдущего вопроса 15.13):

$$p_{\xi_1+\xi_2}(x_1) = \int_{-\infty}^{\infty} p_{(\xi_1,\xi_2)}(x_1 - x_2, x_2)\mathrm{d}x_2 = \int_{-\infty}^{\infty} p_{\xi_1}(x_1 - x_2)p_{\xi_2}(x_2)\mathrm{d}x_2..$$

Это очень напоминает одночастичную матрицу плотности (или функцию распределения).

Самостоятельно доказать аналог последней теоремы для дискретных СВ $\xi_1 \ldots \xi_n$, принимающих значения $\{c_i^{(j)}\}_{i\geq 1}$: Если $\forall i_1 \ldots i_n \ P\{\prod_{j=1}^n \{\xi_j = c_{i_j}^{(j)}\}\} = \{\prod_{j=1}^n P\{\xi_j = c_{i_j}^{(j)}\}\}$, то СВ $\xi_1 \ldots \xi_n$ независимы.

Теорема 15.13 *ожидание произведения двух независимых СВ равно произведению ожиданий.*

Доказательство ...

15.15 Ковариация и коэффициент корреляции

Определение: **ковариацией** двух СВ ξ_1 и ξ_2 называется $\mathrm{cov}(\xi_1, \xi_2) = M[(\xi_1 - M\xi_1)(\xi_2 - M\xi_2)]$. Интеграл от произведения двух функций можно рассматривать как скалярное произведение этих функций.

Свойства ковариации:

1. $\mathrm{cov}(\xi_1, \xi_2) = \mathrm{cov}(\xi_2, \xi_1)$, $\mathrm{cov}(c\xi_1, \xi_2) = c\,\mathrm{cov}(\xi_1, \xi_2)$.
2. $\mathrm{cov}(\xi_1, \xi_2) = M\xi_1\xi_2 - M\xi_1 \cdot M\xi_2$ (ранее мы уже считали это ожидание).
3. Если СВ ξ_1 и ξ_2 – независимы, то $\mathrm{cov}(\xi_1, \xi_2) = 0$ (следует из 2 и теоремы на стр.36).
4. $\mathrm{cov}(\xi, \xi) = D\xi$.

Определение: рассмотрим некоторый случайный вектор $\vec{\xi} = (\xi_1 \ldots \xi_n)^{\mathrm{T}}$; симметричная матрица $D[\vec{\xi}] \equiv \{\sigma_{ij}\} = \{\mathrm{cov}(\xi_i, \xi_j)\}$ называется **ковариационной матрицей**. Нетрудно показать, что $D[\vec{\xi}] = M[(\vec{\xi} - M\vec{\xi})(\vec{\xi} - M\vec{\xi})^{\mathrm{T}}]$ – "диадное" произведение векторов.

Далее из СВ $\vec{\xi}$ и неслучайных матрицы $C \in \mu_{mn}$ и вектора $\vec{a}$ можно образовать линейную комбинацию $\vec{\eta} = C\vec{\xi} + \vec{a} = (\eta_1 \ldots \eta_m)^{\mathrm{T}}$, также являющуюся случайным вектором.

Теорема 15.14 $D[\vec{\eta}] = CD[\vec{\xi}]C^T$.

Доказательство ...

Следствия:

1. При $m = 1$ матрица C превращается в строку $\{C_i\}_{i=1}^n$; получаем $\eta = \sum_{k=1} c_k\xi_k + a$, $D[\eta] = Q(C)$ – квадратичная форма от строки C; $Q(C) = \sum_{k=1}^n \sigma_{ij}C_iC_j$.
2. Если компоненты $\vec{\xi}$ попарно некоррелированы (что заведомо будет иметь место в случае, если они попарно независимы), то из предыдущего следствия при $C_i = 1$ получаем, что $D\sum_{i=1}^n \xi_i = \sum_{i=1}^n D\xi_i$.
3. $D[\vec{\xi}] \geq 0$ (неотрицательно определена); в частности, $\det D[\vec{\xi}] \geq 0$. При $n = 2$ отсюда следует, что $|\,\mathrm{cov}(\xi_2, \xi_1)| \leq \sqrt{D\xi_2 \cdot D\xi_1}$.

Определение: коэффициентом корреляции СВ ξ_1 и ξ_2 называется $\rho(\xi_2, \xi_1) = \mathrm{cov}(\xi_2, \xi_1)/\sqrt{D\xi_2 \cdot D\xi_1}$ – грубая численная характеристика степени (не)зависимости СВ ξ_1 и ξ_2.

Свойства:

1. $\rho(\xi_2, \xi_1)$ имеет смысл косинуса угла в ФП L^2 между функциями $(\xi_1 - M\xi_1)$ и $(\xi_2 - M\xi_2)$.
2. Коэффициент корреляции независимых СВ равен нулю.
3. Если $\rho(\xi_2, \xi_1) = 1$, то $(\xi_1 - M\xi_1) = \pm(\xi_2 - M\xi_2)\cdot\mathrm{const}$.

Пример: рассмотрим схему из n испытаний Бернулли с вероятностью успеха p; Обозначим μ_n – число успехов в проведённом испытании. СВ μ_n представима в виде $\mu_n = \sum_{k=1}^n \mu_n^{(k)}$, $\mu_n^{(k)}(\omega) = \varepsilon_k$, $\omega = \varepsilon_1 \ldots \varepsilon_n$. Более или менее очевидно (Калинину лень это доказывать), что СВ ε_k – независимы. $\Longrightarrow \ldots\ D\mu_n = npq$.

15.16 Закон больших чисел

Теорема 15.15 (Маркова) *: Для последовательности $\{\xi_n\}_{n\geq 1}$ независимых СВ:*

$$\frac{1}{n^2}\sum_{k=1}^{n} D\xi_k \underset{n\to\infty}{\longrightarrow} 0 \Longrightarrow \forall\varepsilon>0 \lim_{n\to\infty} P\{|\frac{1}{n}\sum_{k=1}^{n}\xi_k - \frac{1}{n}\sum_{k=1}^{n} M\xi_k| < \varepsilon\} = 1. \tag{76}$$

– одна из формулировок **закона больших чисел**.
Доказательство (тривиальное): положим $\eta_n = \frac{1}{n}\sum_{k=1}^{n}\xi_k \Longrightarrow M\eta_n = \frac{1}{n}\sum_{k=1}^{n} M\xi_k$. Утверждение (76) эквивалентно

$$\forall\varepsilon>0 \lim_{n\to\infty} P\{|\eta_n - M\eta_n| \geq \varepsilon\} = 0, \quad P\{|\eta_n - M\eta_n| \geq \varepsilon\} \leq \frac{D\eta_n}{\varepsilon^2} = \frac{1}{n^2\varepsilon^2}\sum_{k=1}^{n} D\xi_k \underset{n\to\infty}{\longrightarrow} 0.$$

Определение: Говорят, что последовательность СВ $\{\xi_n\}_{n\geq 1}$ **сходится по вероятности** к СВ ξ, если

$$\forall\varepsilon>0 \lim_{n\to\infty} P\{|\frac{1}{n}\sum_{k=1}^{n}\xi_k - \frac{1}{n}\sum_{k=1}^{n} M\xi_k| < \varepsilon\} = 1. \tag{77}$$

Следствия (простенькие варианты (76)):

1. (Чебышев) Если ξ_n – попарно-некореллированные СВ с равномерно ограниченными дисперсиями, то справедлива теорема Маркова (76). Доказательство:

$$\exists C : \forall n\ D\xi_n \leq C \Longrightarrow \frac{1}{n^2}\sum_{k=1}^{n} D\xi_k \leq \frac{1}{n^2}\sum_{k=1}^{n} C = \frac{C}{n} \underset{n\to\infty}{\longrightarrow} 0.$$

2. Если СВ $\{\xi_n\}_{n\geq 1}$ в (76) одинаково распределены,[34] и имеют конечную дисперсию, то (76) превращается в $\lim_{n\to\infty} P\{|\frac{1}{n}\sum_{k=1}^{n}\xi_k - a| < \varepsilon\} = 1$, где $\forall k\ a = M\xi_k$.

3. (Бернулли) Как обычно, μ_n – число успехов в n испытаниях Бернулли, а p – вероятность успеха в каждом отдельном испытании. $\Longrightarrow \lim_{n\to\infty} P\{|\frac{\mu_n}{n} - p| < \varepsilon\} = 1$. Доказательство:

$$\mu_n = \sum_{k=1}^{n}\mu_n^{(k)}, \quad M\mu_n^{(k)} = p, \quad D\mu_n^{(k)} = pq,$$

после чего остаётся лишь применить следствие 2.

Приятный вывод: частота события по вероятности стремится к вероятности события.

15.17 Центральная предельная теорема (ЦПТ)

Определение: СВ ξ распределена по нормальному закону с параметрами 0 и 1, [обозначение $\xi \sim N(0,1)$], если её плотность и функция распределения равны соответственно

$$p_\xi(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right), \quad \Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp(-\frac{x^2}{2})\mathrm{d}x. \tag{78}$$

$\Phi(x)$ называется функцией Лапласа. Очевидно: если $\xi \sim N(0,1)$, то СВ $\eta = \delta\xi + a \sim N(a,\delta)$.

Теорема 15.16 (ЦПТ в формулировке Ляпунова) *допустим, имеется последовательность $\{\xi_n\}_{n\geq 1}$ независимых СВ и выполняются условия:*

1. *$M\xi_k = a_k$, $D\xi_k = \delta_k^2$, $\exists M|\xi_k - a_k|^3$. Обозначим $B_n^2 = D(\xi_1 + \ldots + \xi_n) = \sum_{k=1}^{n}\delta_k^2$.*

2. *условие Ляпунова: $\sum_{k=1}^{n} M|\xi_k - a_k|^3/B_n^3 \underset{n\to\infty}{\longrightarrow} 0$.*

[34] Это слишком сильное условие.

Определим "нормированную сумму" СВ $\xi_1 \dots \xi_n$ как

$$S_n^* = \frac{1}{B_n}(\sum_{k=1}^{n} M\xi_k - \sum_{k=1}^{n} \xi_k) \Longrightarrow MS_n^* = 0, \quad DS_n^* = 1 \Longrightarrow P(S_n^* < x) \underset{x\in\mathbb{R}}{\rightrightarrows} \Phi(x), \quad n \to \infty. \tag{79}$$

Замечание о том, как можно избавиться от условия Ляпунова: ЦПТ справедлива, если выполнено первое условие и СВ имеют одинаковые моменты. Тогда второе условие выполняется автоматически.

Следствием ЦПТ является

Теорема 15.17 (Муавра-Лапласа) *рассмотрим схему из n испытаний Бернулли с вероятностью успеха p, $q = 1 - p$. Как обычно, μ_n – число успехов; существуют одинаковые третьи моменты.*

$$\mu_n = \sum_{k=1}^{n} \mu_n^{(k)}, \quad M\mu_n = np, \quad B_n^2 = npq \Longrightarrow P\{\frac{\mu_n - np}{\sqrt{npq}} < x\} \underset{x\in\mathbb{R}}{\rightrightarrows} \Phi(x), \quad n \to \infty. \tag{80}$$

Схема приложений ЦПТ. И вот здесь самое время вспомнить о справедливости обсуждения примемнения формулы (43): действительно ли она, как утверждает Тутубалин, в некотором смысле мощнее теоремы (42)?

Пусть $\{\xi_i\}$ удовлетворяют условиям ЦПТ:

$$\begin{aligned} &S_n = \xi_1 + \ldots + \xi_n, \quad MS_n = \sum_{k=1}^{n} a_k, \quad B_n^2 = DS_n \\ &\Longrightarrow P\{A \le S_n \le B\} = P\{\frac{A - MS_n}{\sqrt{DS_n}} \le \frac{S_n - MS_n}{\sqrt{DS_n}} \le \frac{B - MS_n}{\sqrt{DS_n}}\} \approx \Phi(\frac{B - MS_n}{\sqrt{DS_n}}) - \Phi(\frac{A - MS_n}{\sqrt{DS_n}}), \end{aligned} \tag{81}$$

где приближённость происходит из-за того, что в реальной жизни $n \gg 1$ хоть и велико, но не бесконечно.[35]

Следствия формулы (81):

1. "Правило трёх сигм":

$$\begin{aligned} &\frac{B - MS_n}{\sqrt{DS_n}} = 3, \quad \frac{A - MS_n}{\sqrt{DS_n}} = -3, \quad \Phi(3) - \Phi(-3) \approx 0.997 \\ &\Longrightarrow A = MS_n - 3\sqrt{DS_n}, \quad B = MS_n + 3\sqrt{DS_n} \Longrightarrow P\{|S_n - MS_n| \le 3\sqrt{DS_n}\} \approx 0.997. \end{aligned}$$

   Если СВ к тому же ещё и одинаково распределены, то $P\{|\sum_{k=1}^{n} \xi_k - na| \le 3\sigma\sqrt{n}\} \approx 0.997$. Эта формула позволяет с высокой вероятностью выделить из $\sum_{k=1}^{n} \xi_k$ главный член с точностью до поправки $3\sigma\sqrt{n}$.

2. Если СВ одинаково распределены. Мы и раньше из закона больших чисел знали, что $P_n \equiv P\{|\sum_{k=1}^{n} \frac{\xi_k}{n} - a| \ge \varepsilon\} \underset{n\to\infty}{\longrightarrow} 0$. ЦПТ позволяет нам разглядеть процесс этой сходимости в деталях:

$$\begin{aligned} &S_n^* = \frac{\sum_{k=1}^{n} \xi_k - na}{\sigma\sqrt{n}}, \quad P = P\{|S_n^*| \ge \frac{\varepsilon\sqrt{n}}{\sigma}\} = P\{S_n^* \le -\frac{\varepsilon\sqrt{n}}{\sigma}\} + P\{S_n^* \ge \frac{\varepsilon\sqrt{n}}{\sigma}\} = \\ &= (1 - P\{S_n^* < \frac{\varepsilon\sqrt{n}}{\sigma}\}) + P\{S_n^* \le -\frac{\varepsilon\sqrt{n}}{\sigma}\} \approx [1 - \Phi(\frac{\varepsilon\sqrt{n}}{\sigma})] + \Phi(-\frac{\varepsilon\sqrt{n}}{\sigma}) = 2\Phi(-\frac{\varepsilon\sqrt{n}}{\sigma}). \end{aligned}$$

   Теперь положим $\varepsilon\sqrt{n}/\sigma = 3$: Кстати, уместнее везде сменить обозначения $\delta \to \sigma$.

$$\varepsilon = \frac{3\sigma}{\sqrt{n}} \Longrightarrow P\{|\sum_{k=1}^{n} \frac{\xi_k}{n} - a| \le \frac{3\sigma}{\sqrt{n}}\} \approx 0.997. \tag{82}$$

[35] См. стр. [31]41: в мат. статистике нередко 4 – уже большое число, а именно: если *оба* числа np и nq больше четырёх, можно уверенно пользоваться "правилом 3σ".

3. Вероятность и частота – рассмотрим испытания Бернулли; по т. Муавра (80)
$$P\{\frac{\mu_n - np}{\sqrt{npq}} < x\} \underset{x\in\mathbb{R}}{\rightrightarrows} \Phi(x), \quad n\to\infty \Longrightarrow P\{|\frac{\mu_n}{n} - p| > \varepsilon\} \approx 2\Phi(-\frac{\varepsilon\sqrt{n}}{\sqrt{pq}}),$$
$$P\{|\frac{\mu_n}{n} - p| \geq \varepsilon\} \approx 1 - 2\Phi(-\frac{\varepsilon\sqrt{n}}{\sqrt{pq}}) \geq 1 - 2\Phi(-2\varepsilon\sqrt{n}),$$
где мы воспользовались неравенством (?) $pq = p(1-p) \leq 1/4$.

   Выберем некоторое маленькое число $\alpha > 0$. Затем выберем ε_α таким, чтобы $2\Phi(-2\varepsilon_\alpha\sqrt{n}) = \alpha$ (надо численно решить уравнение). $\Longrightarrow P\{|\mu_n/n - p| \leq \varepsilon_\alpha\} \geq 1-\alpha$, то есть интервал $[\mu_n/n - \varepsilon_\alpha, \mu_n/n + \varepsilon_\alpha]$ накрывает значение параметра p с вероятностью $1-\alpha$.

Определение: Случайный интервал (т.е. интервал со случайными концами), накрывающий неизвестное нам значение параметра с вероятностью $1-\alpha$, называется доверительным интервалом для данного параметра с коэффициентом доверия (надёжности) α.

Простенький пример из жизни: задаём $\alpha = 0.05$, бросаем игральную кость 20 раз, хотим выяснить не жулнический ли у нас кубик (кость), т.е. верно ли (с данным коэфф. доверия), что $p = 1/6$. Если $1/6$ не принадлежит этому интервалу, то для $\alpha = 0.05$ гипотеза отвергается.[36]

См. также [**?**] и много где ещё…

15.18 Характеристические функции

Рассмотрим СВ ξ с распределением μ_ξ.

Определение: числовая функция $f_\xi(t) = Me^{it\xi} = \int_{-\infty}^{\infty} e^{itx}\mu_\xi(\mathrm{d}x)$ называется характеристической функцией (ХФ) СВ ξ. Замечание: то, что называлось ХФ в теории множеств, в теории вероятностей называется индикатором множества.

Комплексный интеграл задаётся естественно: $\int(U+iV)\mathrm{d}x = \int U\mathrm{d}x + i\int V\mathrm{d}x$.

Свойства:

1. $|f_\xi(t)| \leq M|e^{it\xi}| = M1 = 1$.
2. Если $\xi_1 \dots \xi_n$ – независимые СВ, то
$$f_{\sum_{k=1}^n \xi_k}(t) = Me^{it\sum_{k=1}^n \xi_k} = M\prod_{k=1}^n e^{it\xi_k} = \prod_{k=1}^n Me^{it\xi_k} = \prod_{k=1}^n f_{\xi_k}(t), \tag{83}$$
где мы учли, что функции независимых СВ являются независимыми СВ.
3. Если k-тый абсолютный момент $M|\xi|^k$ существует, то k-тый момент равен
$$M\xi^k = i^{-k}\frac{\mathrm{d}^k f_\xi(t)}{\mathrm{d}t^k}\Big|_{t=0} \equiv i^k\frac{\mathrm{d}^k f_\xi(0)}{\mathrm{d}t^k}. \tag{84}$$
Доказательство:
$$\frac{\mathrm{d}^k f_\xi(0)}{\mathrm{d}t^k} = \left(\frac{\mathrm{d}^k}{\mathrm{d}t^k}\int_{-\infty}^{\infty} e^{itx}\mu_\xi(\mathrm{d}x)\right)\Big|_{t=0} = i^k\int_{-\infty}^{\infty} x^k e^{itx}\Big|_{t=0}\mu_\xi(\mathrm{d}x) = i^k\int_{-\infty}^{\infty} x\mu_\xi(\mathrm{d}x),$$
причём дифференцирование законно, т.к. $|x|^k$ – суммируемая функция (это справедливо также и для степеней, меньших, чем k).
4. $f_{\delta\xi+a}(t) = e^{ita}f_\xi(t\delta)$.
5. $\sqsupset \xi \sim N(0,1) \Longrightarrow f_\xi(t) = \exp(-t^2/2)$. Доказательство: …
6. $$\sqsupset \varphi, \varphi', \varphi'' \in L \Longrightarrow M\varphi(\xi) = \int_{-\infty}^{\infty}\varphi(x)\mu_\xi(\mathrm{d}x) = \ldots = \frac{1}{2\pi}\int_{-\infty}^{\infty}\tilde{\varphi}(t)f_\xi(-t)\mathrm{d}t,$$
где тильдой обозначен Фурье-образ.
7. $\sqsupset \varphi, \varphi', \varphi'' \in L$ и последовательность ХФ равномерно сходится:
$$\forall T < +\infty \quad f_{\xi_k}(t) \underset{|t|<T}{\rightrightarrows} f_\xi(t),\ k\to\infty \Longrightarrow M\varphi(\xi_n) \underset{n\to\infty}{\rightrightarrows} M\varphi(\xi).$$
Доказательство: …

[36]И я тут же вспоминаю рассказ (см. главу 9) Талеба о “знающем жизнь” (но не знающем математики) толстячке, профессионально кидающем всех этих заумных очкариков, работающих в банках. Очкарику и толстячку даётся задача о многократном кидании монеты; очкарик – балбес, а тостячок пользуется уважением Талеба…

15.19 Теорема о непрерывности

Допустим, $\{\xi_n\}_{n\geq 1}$ – последовательность СВ.

Теорема 15.18

$$\forall T>0 \quad X\Phi f_{\xi_k}(t) \underset{|t|<T}{\rightrightarrows} f_\xi(t) \in \mathbb{C}(\mathbb{R}) \Longrightarrow f_{\xi_n}(x) = P\{\xi_x < x\} \underset{\mathbb{R}}{\rightrightarrows} f_\xi(x).$$

Доказательство: ... длинное...

15.20 Доказательство ЦПТ

Литература: [32, 33, 34, 35, 36]:

15.21 Выборка и оценки

Калинин, вероятно торопился, не рассчитал время и был слишком лаконичен в следующих вопросах. В их разъяснении мне должна помочь книга [31].

Определение: **выборкой** объёма n называются независимые одинаково распределённые СВ $X_1 \dots X_n : \forall i = 1 \dots n$ функция $F_{X_i}(x) = P\{X_i < x\} \equiv F_\xi(x)$ есть ФР некоторой СВ.

Выборка – это математическая модель n независимых измерений (опытов), проведённых в одинаковых условиях. В каком-то смысле СВ ξ "моделлирует" некий прибор (отдельный опыт), а СВ $X_1 \dots X_n$ могут рассматриваться как независимые экземпляры СВ ξ (независимые измерения).

$X_1 \dots X_n$ – числа, полученные в результате эксперимента. Как правило, путём обработки экспериментальных данных мы хотим получить (или хотя бы оценить) неслучайный результат измерения (число) θ. Например, θ может быть параметром функции распределения $F_\xi(\theta, x) \sim N(0, \theta)$.

Оценкой параметра θ называется любая (неслучайная) β-измеримая функция $\hat{\theta}(x_1 \dots x_n)$. В результате эксперимента мы получаем СВ $\hat{\theta} = \hat{\theta}_n(X_1 \dots X_n)$. Видно, что очевидной связи между θ и $\hat{\theta}$ нет.

Оценка $\hat{\theta}_n$ параметра θ называется **состоятельной**, если выборка $\hat{\theta}(x_1 \dots x_n)$ стремится по вероятности к θ:

$$\hat{\theta}(x_1 \dots x_n) \xrightarrow[n\to\infty]{(P)} \theta.$$

Оценка $\hat{\theta}_n$ параметра θ называется **несмещённой**, если $M\hat{\theta}(x_1 \dots x_n) = \theta$. Называя оценку несмещённой, мы имеем в виду, что она не содержит систематической ошибки.

Примеры оценок (т.е. функций $\hat{\theta}$):

1. Эмпирическая функция распределения $F_n(x) =$(число значений X_i, меньших $x)/n$. Тривиально проверяется, что таким образом построенная функция есть функция распределения.

2. Эмпирическое (или выборочное) среднее:
$$\bar{X} = \sum_{k=1}^{n} X_k/n. \tag{85}$$

3. Эмпирическая (выборочная) дисперсия:
$$s^2 = \sum_{k=1}^{n} (X_k - \bar{X})/n. \tag{86}$$

4. k-тый выборочный момент $\alpha_k = \sum_{i=1}^n X_i^k$ – оценка для $a_k = M\xi^k$.

Лемма:...

Теорема 15.19 *В смысле сходимости по вероятности*

1.
$$F_n(X) \xrightarrow[n\to\infty]{(P)} F_\xi(X) \equiv F(X).$$

2. Если $\exists M\xi^{2k} < \infty$, то эмпирический k-тый момент $\alpha_k \xrightarrow[n\to\infty]{(P)} a_k = M\xi^k$. В частности, если $\exists M\xi^2 < \infty$, то $\bar{X} \to a = M\xi$.

3. Если $\exists M\xi^4 < \infty$, то $s^2 \xrightarrow[n\to\infty]{(P)} \delta^2 \equiv D\xi$.

Иными словами, вышеприведённые оценки состоятельны.

Доказательство: ...

Теорема 15.20 *оценка $s^2 = \sum_{k=1}^n (X_k - \bar{X})/(n-1)$ является несмещённой.*

Доказательство: ...

15.22 Многомерное нормальное распределение

Определения – три важнейших распределения мат. статистики:

1. Распределением Пирсона (или χ^2-распределением) называется распределение СВ $\chi_n^2 = \sum_{i=1}^n \xi_i^2$, где ξ_i – независимые и $\xi_i \sim N(0,1)$.

2. Распределением Стьюдента (или t_n-распределением) называется распределение СВ $t_n = \xi / \sqrt{\chi_n^2/n}$, где $\xi \sim N(0,1)$, а СВ χ_n^2 имеет χ^2-распределение; кроме того, СВ ξ и χ_n^2 должны быть независимы. Легко вычислить $M\chi_n^2 = n$ и $D\chi_n^2 = 2n$. Замечание: можно доказать, что $t_n \underset{n\to\infty}{\longrightarrow} N(0,1)$ – не будем доказывать и пользоваться.

3. Распределением Фишера (или $F_{m,n}$-распределением) называется распределение СВ $n\chi_m^2/m\chi_n^2$, где СВ χ_n^2 и χ_m^2 – независимые.

Рассмотрим СВ $\vec{\xi} = (\xi_1 \dots \xi_n)^{\mathrm{T}}$. Выберем некоторую матрицу $C \in \mu_{mn}$, а также неслучайный вектор $\vec{a} = (a_1 \dots a_m)^{\mathrm{T}}$; С их помощью определим СВ $\vec{\eta} = C\vec{\xi} + \vec{a} = (\eta_1 \dots \eta_m)^{\mathrm{T}}$.

Определение: если СВ $\xi_1 \dots \xi_n$ – независимы и $\forall i = 1 \dots n$ $\xi_i \sim N(0,1)$, то говорят, что СВ $\vec{\eta} = C\vec{\xi} + \vec{a}$ **имеет m-мерное нормальное распределение**. Если матрица к тому же квадратная и невырожденная ($m = n$ и $\det C \neq 0$), то говорят, что вектор $\vec{\eta}$ имеет **невырожденное n-мерное нормальное распределение**.

Очевидные свойства:

1. Если $\forall i = 1 \dots m$ $D\eta_i > 0$ (т.е. компоненты вектора $\vec{\eta}$ – действительно случайны), то каждая компонента η_i – нормально рапределена (как ЛК независимых нормально распределённых СВ).

2. $\sqsupset D\eta_i > 0 \Longrightarrow$ СВ $A_1\eta_1 + \dots A_n\eta_n$ – нормально рапределена.

Теорема: рассмотрим (а) случайный вектор $\vec{\xi} = (\xi_1 \dots \xi_n)^{\mathrm{T}}$, состоящий из независимых одинаково распределённых СВ $\xi_i \sim N(0,\delta)$, а также (б) невырожденную квадратную матрицу $C \in \mu_{nn}$, $\det C \equiv |C| \neq 0$ и вдобавок (в) произвольный неслучайный вектор $\vec{a} = (a_1 \dots a_m)^{\mathrm{T}}$.

$\Longrightarrow$ СВ $\vec{\eta} = C\vec{\xi} + \vec{a}$ является абсолютно непрерывной с плотностью распределения[37]

$$P_{\vec{\eta}}(\vec{y}) = \frac{1}{(\sigma\sqrt{2\pi})^n\sqrt{|B|}} \exp\left[-\frac{1}{2\sigma^2}Q(\vec{y}-\vec{a})\right], \quad B = CC^{\mathrm{T}}, \tag{87}$$

где Q – квадратичная форма с матрицей B^{-1}.

Доказательство: ...

Следствия:

1. Положим в условиях теоремы $\sigma = 1 \Longrightarrow$ невырожденная нормальная СВ $\vec{\eta} = C\vec{\xi} + \vec{a}$ имеет плотность распределения
$$P_{\vec{\eta}}(\vec{y}) = \frac{1}{(\sqrt{2\pi})^n\sqrt{|B|}} \exp\left[-\frac{1}{2\sigma^2}Q(\vec{y}-\vec{a})\right], \quad B = CC^{\mathrm{T}} = |D[\vec{\eta}]|, \tag{88}$$
где $D[\vec{\eta}]$ – ковариационная матрица.

2. Известно, что независимые СВ некоррелированы, верно ли обратное? Допустим $\eta_1 \dots \eta_n$ – некоррелировнные компоненты случайного вектора. $\Longrightarrow$ ковариационная матрица диагональна. $\Longrightarrow$ плотность в правой части (88) будет распадаться в произведение одномерных плотностей. $\Longrightarrow \eta_1 \dots \eta_n$ – независимые СВ.

3. В условиях теоремы матрица C – ортогональна $\Longrightarrow CC^{\mathrm{T}} = 1 \Longrightarrow Q$ есть просто сумма квадратов и (87) показывает, что векторы $\vec{\eta}$ и $\vec{\xi}$ имеют одинаковое распределение.

4. Рассмотрим случайный вектор $\vec{\xi}$ с независимыми компонентами, $\xi_i \sim N(0,\sigma)$. Также рассмотрим неслучайный вектор $\vec{b} \in \mathbb{R}^n$, причём $\mathbb{R}^n$ обладает евклидовой структурой. Будем рассматривать подпространства $L_1, L_2 \subset \mathbb{R}^n$.
$$\sqsupset L_1 \perp L_2 \Longrightarrow \text{СВ } \mathrm{pr}_{L_1}(\vec{\xi}+\vec{b}) \text{ и } \mathrm{pr}_{L_2}(\vec{\xi}+\vec{b}) \text{ независимы.} \tag{89}$$
Доказательство: ...

5. $$\vec{b} \perp L \Longrightarrow ||\mathrm{pr}_L(\vec{\xi}+\vec{b})||^2 \sim \sigma^2\chi^2_{\dim L} \tag{90}$$
Доказательство: ...

6. $$\vec{b} \perp L_i, i = 1,2, L_1 \perp L_2 \Longrightarrow \frac{\frac{1}{\dim L_1}||\mathrm{pr}_{L_1}(\vec{\xi}+\vec{b})||^2}{\frac{1}{\dim L_2}||\mathrm{pr}_{L_2}(\vec{\xi}+\vec{b})||^2} \sim F_{m,n}. \tag{91}$$

[37] Выражение (87) вызывает у меня ассоцоации с нелинейной сигма-моделью и вообще с функциональными интегралами. То есть возможно устремить размерность $n \to \infty$ и заявить, что в показателе экспоненты в (87) находятся функции. Не это ли называется случайным процессом?

15.23 Линейная статистическая модель

Допустим, имеется случайный результат n измерений

$$\vec{\eta} = (\eta_1 \dots \eta_n)^{\mathrm{T}} = \vec{a} + \vec{\delta}. \tag{92}$$

Относительно (оцениваемого) вектора $\vec{a}$ в линейной статистической модели предполагается, что 1) $\vec{a} \in L \subset \mathbb{R}^n$, 2) $\vec{\delta} = (\delta_1 \dots \delta_n)^{\mathrm{T}}$ – вектор ошибок. Будем предполагать, что элементы вектора ошибок δ_i – независимы и $\delta_i \sim N(0, \sigma)$. Задача: по результатам наблюдения $\vec{\eta}$ получить оценку $\hat{\vec{a}}(\vec{\eta})$ для $\vec{a}$.

Для нахождения оценки $\hat{\vec{a}}(\vec{\eta})$ воспользуемся предложенным Гауссом **методом максимального правдоподобия**. В предыдущем вопросе посчитана плотность (87) – рассмотрим случай $|B = 1|$. Подставляя в (87) вместо $\vec{y}$ вектор наблюдений, получаем функцию правдоподобия

$$P_{\vec{\eta}}(\vec{a}, \vec{\eta}) = \frac{1}{(\sigma\sqrt{2\pi})^n} \exp\left[-\frac{1}{2\sigma^2}||\vec{x} - \vec{a}||^2\right], \quad \vec{x} - \vec{a} \in L. \tag{93}$$

Гаусс предложил в качестве $\hat{\vec{a}}$ выбрать такую оценку, что

$$P_{\vec{\eta}}(\hat{\vec{a}}(\vec{\eta}), \vec{\eta}) = \max_{\vec{a} \in L} P_{\vec{\eta}}(\vec{a}, \vec{\eta}), \quad \text{что соответствует минимизации} \quad ||\vec{\eta} - \vec{a}(\vec{\eta})||^2 = \min_{\vec{a} \in L} ||\vec{\eta} - \vec{a}||^2. \tag{94}$$

Из линейной алгебры $\Longrightarrow \hat{\vec{a}}(\vec{\eta}) = \mathrm{pr}_L \vec{\eta} = \mathrm{pr}_L \vec{\delta} + \vec{a}$.

Из 15.22 вопроса следует (... см. коммент. в конце тетради...), что $M\hat{\vec{a}}(\vec{\eta}) = \vec{a}$, т.е. оценка несмещённая. Хорошо бы получить также и оценку для дисперсии σ. Для этого поступим следующим образом: возьмём $\vec{\eta} - \hat{\vec{a}}$, обозначим $\Delta^2 = ||\vec{\eta} - \hat{\vec{a}}(\vec{\eta})||^2$. Величина Δ^2 называется **кажущейся ошибкой**. [Настоящая ошибка – это ♣ ошибка в оценке $||\vec{a} - \hat{\vec{a}}(\vec{\eta})||$.]

$$\vec{\eta} = \vec{a} + \vec{\delta}, \quad \vec{a} \perp L^{\perp} \Longrightarrow \Delta^2 = ||\vec{\eta} - \mathrm{pr}_L \vec{\eta}||^2 = ||\mathrm{pr}_{L^{\perp}} \vec{\eta}||^2 \sim \sigma^2 \chi^2_{\dim L^{\perp}} = \sigma^2 \chi^2_{n - \dim L}. \tag{95}$$

В качестве оценки s^2 для σ^2 возьмём

$$s^2 = \frac{1}{n - \dim L} \Delta^2 \sim \frac{\sigma^2 \chi^2_{n - \dim L}}{n - \dim L}. \tag{96}$$

Из того, что

$$M s^2 = \frac{\sigma^2}{n - \dim L} \cdot M \chi^2_{n - \dim L} = \sigma^2 \tag{97}$$

следует, что эта (выбранная) оценка – несмещённая. Предположим,♣ что размерность $\dim L$ фиксирована. Посмотрим на СВ

$$s^2 = \frac{1}{n - \dim L} \sum_{i=1}^{n - \dim L} \sigma^2 \xi_i^2, \quad \xi_i \sim N(0, 1), \tag{98}$$

причём ξ_i – независимы. Из закона больших чисел следует, что $s^2 \xrightarrow{(P)} \sigma^2$.

Замечание: рассмотрим $||\vec{a} - \hat{\vec{a}}(\vec{\eta})||^2 = ||\mathrm{pr}_L \vec{\delta}||^2$. Из 15.22. вопроса ясно, как она распределена: $||\mathrm{pr}_L \vec{\delta}||^2 \sim \sigma^2 \chi^2_{\dim L}$. Заметим, что СВ s^2 и норма $||\vec{a} - \hat{\vec{a}}(\vec{\eta})||^2$ – независимы, т.к. это проекции на разные подпространства. В теории вероятностей нередко используется соотношение

$$\frac{\frac{1}{\dim L}||\vec{a} - \hat{\vec{a}}(\vec{\eta})||^2}{s^2} \sim \frac{\frac{\sigma^2}{\dim L}\chi^2_{\dim L}}{\frac{\sigma^2}{n - \dim L}\chi^2_{n - \dim L}} \sim F_{\dim L, \dim L^{\perp}}. \tag{99}$$

15.24 Нормальная выборка (теория ошибок).

В качестве примера рассмотрим частный случай предыдуєй модели:

$$\sqsupset \vec{a} = (a, a, \dots, a)^{\mathrm{T}} = a\vec{e}\sqrt{n}, \quad \vec{e} = (1, 1, \dots, 1)^{\mathrm{T}}/\sqrt{n}. \tag{100}$$

Имеется вектор ошибок $\delta_i \sim N(0, \sigma) \Longrightarrow \vec{a} \in L = \mathscr{L}(\vec{e})$. Рассмотрим вектор наблюдений $\vec{\eta} = \vec{a}\vec{\delta} \Longrightarrow$ для $\vec{\eta}$ получается такая оценка:

$$\hat{\vec{a}}(\vec{\eta}) = \mathrm{pr}_L \vec{\eta} = (\vec{\eta}, \vec{e})\vec{e} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \eta_i \left(\frac{1}{\sqrt{n}} \dots \frac{1}{\sqrt{n}}\right)^{\mathrm{T}} = (\bar{\eta}, \bar{\eta}, \dots, \bar{\eta})^{\mathrm{T}}, \quad \bar{\eta} = \frac{1}{n} \sum_{i=1}^{n} \eta_i. \tag{101}$$

[В мат. статистике греческие буквы не используются.] $\Longrightarrow$ заменяем

$$\vec{\eta} \to \vec{x}, \quad \hat{\vec{a}} = (\bar{x} \dots \bar{x})^{\mathrm{T}}, \quad \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i,, \quad s^2 = \frac{1}{n - \dim L}||\vec{x} - \hat{\vec{a}}(\vec{x})||^2 = \frac{1}{n - \dim L} \sum_{i=1}^{n} (x_i - \bar{x})^2 \sim \sigma^2 \chi^2_{n-1} \cdot \frac{1}{n-1} \Longrightarrow \frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}. \tag{102}$$

Воспользуемся этим для построения **доверительного интервала** (ДИ) для дисперсии.

Мы хотим построить ДИ для σ^2 с коэффициентом доверия α. Выбираем по таблицам χ^2-распределения числа $x_1(\alpha)$ и $x_2(\alpha)$ так, чтобы

$$P\{x_1(\alpha) \le \chi^2_{n-1} \le x_2(\alpha)\} = 1-\alpha. \tag{103}$$

Соотношения (103) недостаточно для однозначного определения $x_1(\alpha)$ и $x_2(\alpha)$. $\Longrightarrow$ ставим дополнительные условия

$$P\{\chi^2_{n-1} \le x_1(\alpha)\} = \frac{\alpha}{2}, \quad P\{\chi^2_{n-1} \ge x_2(\alpha)\} = \frac{\alpha}{2}. \tag{104}$$

Въ(103) подставляем $(n-1)s^2/\sigma^2$ вместо χ^2_{n-1}, и, решая неравенство относительно σ^2, получаем

$$P\{x_1(\alpha) \le \frac{(n-1)s^2}{\sigma^2} \le x_2(\alpha)\} = P\{\frac{(n-1)s^2}{x_2(\alpha)} \le \sigma^2 \le \frac{(n-1)s^2}{x_1(\alpha)}\} = 1-\alpha. \tag{105}$$

– доверительный интервал по выборке конечного объёма.

Построим ДИ для a. Рассмотрим СВ $\bar{x} - a = \frac{1}{n}\sum_{i=1}^n x_i - a$.

$$M(\bar{x}-a) = 0, \quad D(\bar{x}-a) = \frac{1}{n^2}\sum_{i=1}^n Dx_i - a = \frac{\sigma^2}{n}, \tag{106}$$

где мы учли, что $Dx_i = \sigma^2$. $\Longrightarrow$ СВ $\frac{\bar{x}-a}{\sqrt{\sigma^2/n}} \sim N(0,1)$. Из предыдущего () вопроса $\Longrightarrow$ СВ $\frac{(\bar{x}-a)\sqrt{n}}{\sigma}$ и s^2 – независимы. Далее,

$$\frac{\left(\frac{(\bar{x}-a)}{\sigma}\sqrt{n}\right)}{\sqrt{s^2/\sigma^2}} \sim \frac{\xi}{\sqrt{\chi^2_{n-1}/(n-1)}} \sim t_{n-1} \Longrightarrow \frac{(\bar{x}-a)\sqrt{n}}{s} = t_{n-1}, \tag{107}$$

где мы учли, что $\xi \sim N(0,1)$.

Строим ДИ для среднего: выбираем из таблиц распределения Стьюдента такое число t_α, что $P\{|t_{n-1}| \le t_\alpha\} = 1-\alpha$; подставляем туда $\frac{(\bar{x}-a)\sqrt{n}}{s} = t_{n-1} \Longrightarrow P\{|\frac{(\bar{x}-a)\sqrt{n}}{s}| \le t_\alpha\} = 1-\alpha$, решаем неравенство и получаем, что

$$P\{\bar{x} - \frac{s}{\sqrt{n}}t_\alpha \le a \le \bar{x} + \frac{s}{\sqrt{n}}t_\alpha\} = 1-\alpha. \tag{108}$$

15.25 Нормальная выборка – проверка гипотез.

Гипотеза первая: $a = a_0$. Дано: (неслучайное) чосло a_0, имеется выборка из нормального распределения $x_1 \dots x_n \sim N(a,\sigma)$. Надо проверить гипотезу $H_0 : a = a_0$. Альтернативой нашей гипотезе могут быть две возможности:

1. (односторонняя альтернатива) если $a \ne a_0$, то $a > a_0$.
2. если $a \ne a_0$, то $a > a_0$ или $a < a_0$.

Выбираем довольно-таки малое число $\alpha \ll 1$ – уровень значимости при проверке гипотезы. По числу α строится тн критическая область (КО). Если точка $(x_1 \dots x_n)^{\mathrm{T}} \in$КО, то гипотеза не будет отвергаться, а на практике будет применяться. В случае первой гипотезы оценкой для a является среднее $\bar{x} = \frac{1}{n}\sum_{k=1}^n x_k$. Гипотезу естественно отвергать в случае, когда $\bar{x} - a_0 \ge d_1 > 0$. Число d_1 выбирается через α таким образом, что $P\{\bar{x} - a_0 \ge d_1\} = \alpha$. Как построить d_1? Рассмотрим СВ (или статистику)

$$t_{n-1} = \frac{(\bar{x}-a_0)\sqrt{n}}{s}, \quad s^2 = \frac{1}{n-1}\sum_{k=1}^n (x_k - \bar{x})^2 \tag{109}$$

Из предыдущего (15.24.) вопроса $\Longrightarrow$ СВ t_{n-1} имеет распределение Стьюдента с $n-1$ степенью свободы. $\Longrightarrow$ из таблиц распределения Стьюдента можем выбрать такое число t_α, что $P\{|t_{n-1}| \ge t_\alpha\} = \alpha \Longleftrightarrow P\{\frac{(\bar{x}-a_0)\sqrt{n}}{s} \ge t_\alpha\} = \clubsuit\alpha \Longleftrightarrow P\{\bar{x} - a_0 \ge d_1 = \frac{st_\alpha}{\sqrt{n}}\} = \alpha \Longrightarrow$ при заданном α гипотеза отвергается, если $\bar{x} \ge a_0 + st_\alpha/\sqrt{n}$.

КО выбирается из соображения $|\bar{x} - a_0| \ge d_2$, выбираем $t'_\alpha : P\{|t_{n-1}| \ge t'_\alpha\} = \alpha$ гипотеза отвергается при условии $|\bar{x} - a_0| \ge \frac{st'_\alpha}{\sqrt{n}} = d_2$.

Гипотеза вторая: $\sigma_1 = \sigma_2$. Имеются две независимые выборки $x_1 \dots x_n \sim N(a_1,\sigma_1)$ и $y_1 \dots y_n \sim N(a_2,\sigma_2)$. Будет ли $\sigma_1 = \sigma_2$? Альтернативы могут быть разные: $\sigma_1 \ne \sigma_2$, $\sigma_1 > \sigma_2$, … Как здесь построить КО? $\sigma_1 = \sigma_2 = \sigma \Longrightarrow \sigma_1/\sigma_2 = 1$.

Рассмотрим статистику (про которую мы уже кое-что знаем)

$$\frac{\frac{1}{n-1}\sum_{k=1}^n (x_k - \bar{x})^2}{\frac{1}{m-1}\sum_{s=1}^n (y_s - \bar{y})^2} \sim \frac{\frac{1}{n-1}\sigma^2\chi^2_{n-1}}{\frac{1}{m-1}\sigma^2\chi^2_{m-1}} = F_{n-1,m-1}. \tag{110}$$

– распределение Фишера (которое тоже продублировано), и задача решается аналогично. Выражение (110) равно s_x^2/s_y^2; проверяем $s_x^2/s_y^2 > 1 + d_1$ (для односторонней альтернативы).

Гипотеза третья: $a_1 = a_2$. Имеем две независимые выборки $\sqsupset x_1 \dots x_n \sim N(a_1, \sigma_1)$ и $\sqsupset y_1 \dots y_n \sim N(a_2, \sigma_2)$. Тоже могут быть разные альтернативы. Если $\sigma_1 \neq \sigma_2$, то поставленная задача является проблемой Берена-Фишера. Решения её (пока) нет. Она вполне разрешима при больших m и n. При $m, n \gg 1$, если $a_1 = a_2$

$$\bar{x} - \bar{y} \sim N\left(0, \underbrace{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}}_{\text{неизвестно}}\right) \tag{111}$$

– это выражение нам неизвестно, потому что неизвестен корень. При больших m и n естественно считать, что (подставляем оценки)

$$D(\bar{x} - \bar{y}) \approx \frac{s_x^2}{n} + \frac{s_y^2}{m} \Longrightarrow \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}} \sim \boxed{\text{приближённо}} \sim N(0, 1). \tag{112}$$

После этого с лёгкостью строим КО:

$$\frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}}} \geq d > 0 : P\{\eta \geq d\} = \alpha, \text{ где } \eta \sim N(0, 1). \tag{113}$$

Стьюдент[38] решил эту задачу полностью при условии, что $\sigma_1 = \sigma_2 = \sigma$, и это – основной его результат.

Нам известно, что $\bar{x} - \bar{y} \sim N\left(0, \sigma\sqrt{\frac{1}{n} + \frac{1}{m}}\right)$. В частности, отсюда следует, что СВ $\frac{\bar{x}-\bar{y}}{\sigma\sqrt{\frac{1}{n}+\frac{1}{m}}} \sim N(0, 1)$. Нам неизвестна величина σ, приведём для неё независимую оценку. Оказывается, что величина $s^2 = \frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}$

1. имеет известное распределение и
2. является несмещённой оценкой для σ^2.

Покажем справедливость первого пункта:

$$s^2 \sim \frac{(n-1)\frac{\sigma^2}{n-1}\chi_{n-1}^2 + (m-1)\frac{\sigma^2}{m-1}\chi_{m-1}^2}{n+m-2} \sim \frac{\sigma^2 \chi_{n+m-2}^2}{n+m-2}, \tag{114}$$

где мы учли, что χ_{n-1}^2 и χ_{m-1}^2 – независимы, т.к. они относятся к разным выборкам. Возьмём такую статистику

$$\frac{\bar{x} - \bar{y}}{s\sqrt{\frac{1}{n} + \frac{1}{m}}} = \frac{\frac{\bar{x}-\bar{y}}{\sqrt{\frac{1}{n}+\frac{1}{m}}}}{\underbrace{\sqrt{\sigma^2 \frac{\chi_{n+m-2}^2}{n+m-2}}}_{\sim s}} = \frac{\frac{\bar{x}-\bar{y}}{\sigma\sqrt{\frac{1}{n}+\frac{1}{m}}}}{\sqrt{\frac{\chi_{n+m-2}^2}{n+m-2}}}, \quad \frac{\bar{x} - \bar{y}}{\sigma\sqrt{\frac{1}{n} + \frac{1}{m}}} \sim N(0, 1). \tag{115}$$

Заметим, что СВ $\bar{x}$, $\bar{y}$, s_x^2, s_y^2 – независимы. $\Longrightarrow$ независимыми будут также СВ $\bar{x} - \bar{y}$ и $s^2 = \frac{(n-1)s_x^2 + (m-1)s_y^2}{n+m-2}$ $\Longrightarrow$ в выражении (115) числитель и знаменатель независимы $\Longrightarrow$ (115) $\sim t_{n+m-2}$. Ну а дальше проверка гипотезы происходит так же, как и в случае $a = a_0$.

Обычно сначала проверяют гипотезу $\sigma_1 = \sigma_2$; в случае её подтверждения действуют по Стьюденту.

16 Задачи (в основном по теории вероятностей)

См. также [37].

1. Как вывести двойную группу на краю зоны Бриллюэна в кремнии или германии?
2. В невесомости находится шар из эластичного материала, заполненный несжимаемой жидкостью (водой). Держась за противоположные точки шара, я немного растянул его, так что диаметр шара d увеличился на $\delta \ll d$. Написать выражение для площади поверхности шара в зависимости от δ/d с точностью до первой неисчазающей поправки.
3. Я подбросил монетку 5 раз, какова вероятность того, что 3 раза выпадет "орёл"?

[38] псевдоним Госсета.

4. Главный инженер компании, производящей телевизоры, должен выбрать поставщика электролитических конденсаторов, которых в каждом телевизоре по 200 штук. Каким должен быть средний срок службы конденсаторов, чтобы телевизоры не ломались в течение первых двух лет работы (пока действует гарантия), но в же время чтобы они ломались не позже, чем через 5 лет (чтобы старый телевизор выбросили поскорее на помойку)? Будем считать, что срок службы конденсаторов измеряется с погрешностью ±10%, а также, что другие детали в телевизорах практически никода не ломаются.)

5. Из Нью-Йорка (США) в Петербург (РФ) письмо доходит в среднем за 21 день, если только оно не утеряно. Я отправил письмо 50 дней назад, а адресат её ещё не получил. (Только на основании этих данных) оцените вероятность того, что письмо было утеряно.

6. Из Нью-Йорка (США) в Петербург (РФ) посылка доходит за 30±10 дней, если только она не украдена или утеряна. Я отправил посылку 50 дней назад, а адресат её ещё не получил. Оцените вероятность того, что посылка была украдена или утеряна.

7. Читаем в новостях 2015-07-06 о предстоящем боксёрском поединке "На счету 26-летнего британца 24 победы (18 — нокаутом) в 24 поединках, а 39-летний Кличко выиграл 64 раза (53 — нокаутом) и потерпел три поражения в 67 боях." Исходя из одной лишь только этой информации требуется оценить вероятности каждого из четырёх исходов поединка. Информацию о возрасте боксёров мы для простоты проигнорируем.

8. Рядом со зданием, где находится мой офис, есть парковка, но мне там оставлять машину нельзя. Известно, что после 16:00 всем можно парковаться где угодно. Я предполагаю, что после 15:20 никто не проверяет парковку. Я уже 5 раз оставлял машину после 15:20, и ни разу меня не штрафовали. Как мой личный опыт может помочь оценить вероятность штрафа после 15:20?

9. Придумать задачу на применение (69).

10. Желая сэкономить на страховке, я собираюсь перейсти на страховку с малым километражом. Задача: оценить будущий километраж на основании предыдущего.

11. Читаем в художественной литературе (Дмитрий Рус) про компьютерное опознание людей на случайной фотографии: "Фигура «N1» с вероятностью в 82.13% опознана как <особо разыскиваемый> Глеб Назаров 200X года рождения.". Могут ли эти слова иметь смысл? Имеет ли смысл указывать вероятность с точностью в 4 знака?

12. Там же читаем: "Вероятность поражения цели одной ракетой – 82%, двумя ракетами – 97%." Имеет ли данное высказывание смысл? Возможно ли подобное значительное увеличение вероятности в случае, если события независимы? (На самом деле это не так потому, что лишняя ракета отвлекает защитные устройства цели.)

13. В новостях читаем, о том как россияне любят свои власти: "Опрос проводился 23 августа по телефону среди 1000 граждан в 320 городах и 160 селах. Его статистическая погрешность не превышает 3,8 процента." Что означают эти "3.8 процента"?

14. После катастрофы "Челенджера" выяснилось, что инженеры оценивали вероятность катастрофы как 1/100, а высший менеджмент НАСА — как 1/100000. Судя по всему, даже оценки инженеров были излишне оптимистичны, ну а менеджеры были просто идиотами. (Хотя почему "были", они в большинстве своём такими и остались.) Но вот вопрос: базировались ли оценки инженеров на каких-либо основаниях? На каких? (Кстати, эту аварию тоже, вероятно, можно назвать, хоть и небольшим, но «Чёрным Лебедем».)

15. РВСН время от времени устраивают испытывательные пуски старых (1975г.) советских ракет Р-36М (известных на Западе под именем "Сатана"). Удачный пуск одной ракеты является основанием для продления срока службы оставшихся (напр., до 2026г.). Как производятся соответствующие расчёты?

16. Корреляция между мировыми ценами на нефть и курсом рубля по отношению к другим валютам достигла рекордного показателя с 2003 года, сообщает агентство Bloomberg. Уровень линейной взаимосвязи между показателями достиг 82%... Что сие означает?

17. Реакция разложения соли в химии: $\mathrm{NaCl} \Longrightarrow \mathrm{Na}^+ + \mathrm{Cl}^-$. Подумать, какой смысл имеет т.н. "константа равновесия" этой реакции: $K = [Na^+][Cl^-]/[NaCl]$. Почему в насыщенном растворе $[Na^+][Cl^-] = \mathrm{const}$ при фиксированной температуре? См. также "закон действующих масс".

18. Дигарамная техника усреднения по беспорядку есть пример нахождения усреднённого решения (дифференциального) уравнения для случая, когда параметры этого уравнения (напр., потенциальное поле) неизвестны, а известны лишь их средние. Мне это напоминает мысли об использовании теоремы Пуассона на стр. 26 и [15]31.

19. Модель проводимости двумерного электронного газа. Имеем случайным образом разбросанные одинаковые примесные атомы. Про примесный атом известна вероятность рассеяния им электрона на накоторый угол. Внешнее электрическое поле не ускоряет электроны, а только меняет их траэктории (таким образом избавляемся от расходимостей в интегралах). Получить формулу Друдэ и поправки к ней, обусловленные, к примеру, анизотропией рассеяния на примеси.

20. В [29] приводятся примеры того, как ЦПТ неприменима,[39] напр., в отношении распределения богатства. Представим себе, что нам заранее неизвестно, что один условный Билл Гейтс богаче сотни тысяч обычных людей. Допустим, мы – инопланетяне, только начавшие изучать человечество. Как нам избежать неверного использования ЦПТ, а если это невозможно, какие оговорки к её результатам мы должны написать в нашем отчёте в нашу Далёкую Галактику?

21. Приведите аргументы за и против объектного программирования, в том числе с использованием теории вероятностей.

22. Среднее арифметическое является несмещённой оценкой мат. ожидания. Как насчёт среднего геометрического?

23. В [29] Нассим Талеб рассуждает: счастливчик, выигравший несколько раз подряд в казино, убеждается в существовании ангела-хранителя (или иных помогающих ему сверхъестественных сил), потому что вероятность подобного выигрыша крайне мала (в [29] приводится значение 1/7'000'000). Автор утверждает, что этот счастливчик неправ потому, что "забывает о кладбище", то есть о том, что кроме него миллионы других людей играли в казино и были при этом менее удачливы. Вероятность же того, что хотя бы одному человеку из нескольких миллионов улыбнётся маловероятная удача, не так уж и мала. Вопрос – кто (не)прав: Нассим Талеб, его персонаж-счастливчик, или оба? Почему? Переформулировка: "как же нам повезло, что мы живём именно на Земле, где не слишком жарко/холодно и вообще созданы идеальные условия для жизни!" – так никто не рассуждает, потому что всем очевидно: жизнь, возможно, "пыталась" зародиться и на Марсе, Венере, а также многих других планетах, но развиться там не сумела из-за неподходящих условий. Эти "неподходящие" планеты тоже представляют собой "молчаливое кладбище", о существовании которого не следует забывать, рассуждая о своём "чудесном везении". Формализовать утверждения Талеба можно так: люди часто принимают условные вероятности за обыкновенные.

24. Вот есть у нас, к примеру, электрическое поле 1В/см. С какой скоростью должна двигаться инерциальная система отсчёта, в которой электрическое поле равно нулю?

25. И без кинетики понятно усиление света в лазере: матричный элемент рождения фотона пропорционален числу уже имеющихся фотонов. Казалось, бы напрашивается аналогия с нейтронами в ядерной бомбе, а также с ударной ионизацией в плазме. Проблема в том, что нейтроны и электроны – фермионы, и принцип Паули вроде как не разрешает (точно ли?) этот механизм усиления. Приходится рассматривать кинетику. Случай ядерного распада осложнён тем, что число вылетающих из ядра нейтронов – непостоянно (вроде от 1 до 9). Альтернативный взгляд на проблему: два состояния, разделённые туннельным барьером. Состояние с наименьшей энергией может долго оставаться незаполненным, потому что барьер весьма толст и высок. Но если временно одолжить энергию у пролетающего мимо нейтрона, барьер можно легко перепрыгнуть. Сюда же меня приводят и мысли о неупорядоченных материалах (стёклах). Вопрос: можно ли каким-точки образом отличить электрон в яме, отделённый барьером от более глубокой ямы, от задачки с возбуждённом электроне или атомном ядре? Барьерная аналогия хороша тем, что энергия "вспомогательной" частицы не обязана быть равной разности энергий между уровнями в разделённых барьером ямах. Ещё одна аналогия: установление равновесия в многоэлектронной системе –тоже "вынужденное", т.к. происходит в результате многочастичного взаимодействия. Отсюда, кстати, следующая аналогия: процентная ставка ЦБ подобна температуре.[40] Чем ниже ставка, тем выше температура и тем интенсивнее идут финансовые операции. Почему электрон (в контакте с резервуаром) стремится к минимуму энергии? Да потому, что с минимума у него меньше возможностей сбежать: из возбуждённого уровня он может сбегать и вниз, и вверх, а с основного – только вниз. Такое чисто комбинаторное объяснение. Аналогом энергии в экономике является капитал, а основного состояния – безденежье. Безденежное состояние подобно электрону на дне зоны проводимости. Состояние типичного человека со стабильной работой подобно электрону, захваченному глубоким донором, и вырваться из него ещё сложнее, чем из безденежья. Человек может пытаться "возбудиться", занявшись бизнесом, либо "остепениться", позволив захватить себя какому-нибудь донору (стабильной работе). Электроны не обладают волей и разумом, поэтому они стремятся в состояние с наименьшей энергией. Те люди, которые обладают этими качествами, ведут себя иначе, чем электроны.

Как и в физике, особенно примечательны задачи/примеры, в которых интуитивный ответ неясен или (ещё лучше) неверен.

[39] Неприменимость ЦПТ для данного случая Талеб описывает так, как будто это свидетельствует о её неверности. При этом позволяет себе насмехаться над другими людьми, также допускающими аналогичные (очевидно) ошибочные интерпретации.

[40] Вероятно, эта аналогия общеизвестна (как минимум подсознательно) – вспомним выражение "перегретая экономика".

17 Логические замечания

- Вероятность – друг эмпирика. Как это описано у Ландау-Лифшица, мы не обязаны вникать в подробности сложнейших микроскопишеских взаимодействий во многочастичной системе, мы отказываемся от всей глубины понимания, используя вместо неё вероятностные модели. Другим важным инструментом, позволяющим предсказывать поведение сложных систем, не зная микроскопических подробностей происходящих внутри них процессов, является (групповая) теория симметрии. Добавлю ещё, что стремление фермионной системы к равновесию есть следствие теории вероятностей и принципа запрета Паули.
- Неупорядоченное твёрдое тело (напр., легированный металл или полупроводник), разумеется, никакими пространственными симметриями не обладает; однако же его функция распределения (соответствующая усреднению по ансамблю образцов) однородна и изотропна. Не может [38] ли эта симметрия быть использована, к примеру, при выводе слабой (анти)локализации?
- Интересно, можно ли как-нибудь формализовать тот факт, что некоторые крайне маловероятные события (т.н. «Чёрные Лебеди»[29]) всё равно должны рассматриваться из-за своей важности. Пример: мутация, в результате которой бактерия получает устойчивость к антибиотику – со временем получившийся новый штамм вытеснит остальных. Иначе говоря, малость вероятности события компенсируется его важностью или временем воздействия. Можно ли математически измерить важность?
- Цитата из [29]: "Научитесь в оценке верований исходить не из их видимой убедительности, а из того, сколько вреда они могут принести." При планировании важнее знать не наиболее вероятный вариант развития событий, а наихудший вариант, особенно если вариант этот совершенно неприемлем. Иначе говоря, второй момент (дисперсия) не менее важен, чем первый (ожидание), и это противоречит человеческой интуиции, которая автомтически аппелирует к аналогии с асимптотическим разложением (рядом Тейлора) в малой окрестности некоторой точки, где (обычно) каждый следующий член разложения менее важен, чем предыдущий.
- Обращение вероятности: Талеб не знает вероятности землятресения, но в общем представляет себе, что оно может сделать с городом…
- Любопытно соотнести размышления автора [29] с признаками революционной ситуации, которые вбивали в советских школьников (помню, мы их ещё пошло переиначивали в "верхи хотят, но низы не могут"). Т.е. советский учебник истории считал, что революцию (которая по [29] является классическим «Чёрным Лебедем») можно предсказать, и в этом совершенно противоречит [29].
- Чёрных лебедей можно создавать, причём в некоторых случаях для этого не нужно значительных усилий. Надо только понять, где именно должна взмахнуть крыльями бабочка, чтобы вызвать нужный ураган.
- Мысль Талеба [29] о том, что иногда информация вредна, я бы переформулировал так: информация иногда бывает полезной (правдивой и важной), бесполезной (неважной), и вредной (важной, но ложной). Но бесплатной информации не бывает, потому что любая информация требует усилий по её обработке. В процессе обработки также всегда есть риск ошибки. Таким образом, всегда на обработку информации расходуются ресурсы. Информация полезна, если польза от неё перекрывает стоимость ресурсов, потраченных на её обработку. (А ведь нередко приходится дополнительно платить ещё и за добычу информации!) Если однотипную информацию приходится многократно периодически обрабатывать одним и тем же алгоритмом, стоимость обработки (в пересчёте на мегабайт) снижается и такая информация становится как минимум безвредна.
- Пример контринтуитивного (и уже из-за этого интересного) умозаключения: наблюдение Талеба о том, что влияние неожиданностей на планы одностороннe.
- Швейцария – безопасная и стабильная страна с высоким уровнем жизни. Но, как и в случае родины Талеба, сотни лет стабильности могут внезапно закончиться, если, к примеру, французские мусульмане всерьёз захотят построить халифат.
- Мысль Талеба о вреде чрезмерного применения математики к экономике: точность и подробность информации, которой мы владеем, весьма невысока. Поэтому нет смысла в применении сложных алгоритмов для анализа такой информации; лучше потратить силы и время на что-нибудь более полезное, напр., на разведку. Кстати, это "откровение" встречается в учебниках, процитированных в конце этого конспекта. Нечто вроде "понятно, что на самом деле события не являются независимыми, но связаны множеством неизвестных нам причинных связей. Но из-за недостатка (или неточности) информации об этих связях хреначте Пуассона и не выпендривайтесь."
- Читаю в новостях 2016-04-29: Продав один процент акций Apple, американский миллиардер обвалил курс её акций на три процента.
- Талеб – умнейший из гуманитариев; тем не менее, он считает удивительной мысль о том, что не все случайные величины имеют нормальное распределение! Поразительно.
- О несимметричных распределениях. Известно, что два случайно выбранных писателя в совокупности заработали на своих книгах миллион. Наиболее вероятно, что один из них заработал намного (раз в 20) больше, чем второй. Этакое "спонтанное нарушение симметрии".

- Ещё Талеб. Как правило "добрые" Чёрные Лебеди[41] воздействуют постепенно (годами, иногда – десятилетиями), а злые[42] (землятресения, биржевые обвалы, цунами, "11 сентября") – быстро (в течение одного дня); и это неудивительно: ломать – не строить.
- 2016-12-11: Проиграл в покер, фактически нивелировав предыдущий выигрыш. Главная ошибка: игнорировал фкуктуации. В среднем вероятность удачной комбинации совпадает у всех игроков, но с увеличением числа игроков растёт разброс этой величины. Поэтому вероятность того, что, скажем, 5 конов подряд какой-нибудь (но не один и тот же) из 7 игроков будет удачливее меня, не мала. (Кстати, интересно бы посчитать её.) Я же, думая, что остальным не везёт так же, как и мне, предполагал блеф с их стороны, и рискуя, проигрывал.
- Мои приятели по покеру говорят, что туз и король одной масти на раздаче дают бо̀льшую вероятность выигрыша, чем два туза. Интересно проверить, верно ли это утверждение и зависит ли оно от количества игроков?
- (См. стр. [24]157.) При стократном бросании одной монеты наиболее вероятным является выпадение 50 "решек", но вероятность такого события меньше 0.08. Аналогично, если у меня на раздаче выпала низкая пара, вероятнее всего, других пар не возникнет, но вероятность такого события невелика; поэтому вероятность моего выигрыша не слишком велика.
  Более подробно – см. стр. [24]159: Пётр и Павел бросают монету 10'000 раз. Наивный наблюдатель подумает, что из закона больших чисел следует, что каждый из них выиграет приблизительно одинаковое количество раз.
  На самом же деле вероятность того, что Пётр выиграет не больше 20 раз, намного выше того, что количество его выигрышей расположено в интервале между 4990 и 5010.

18 Психологические замечания

- Талеб делит профессии на "масштабируемые" и "немасштабируемые", обращая при этом внимание на то, что лишь среди "масштабируемых"профессий встречаются интересные. Я полностью с ним в этом согласен, и могу добавить, что по моему разумению, наука должна быть "масштабируемой" профессией, но по факту она перестала ей быть. Я могу писать хорошую статью полтора года, но это не будет оценено. В науке выигрывают стабильные посредственности. Утратив масштабируемость, наука перестала быть мне интересной.
- У итальянцев злоупотребления словом "вероятность" частично основаны на словообразовании (лат. probabilis = внушающий доверие, возможный, правдоподобный). У немцев, вероятно, "Wahrscheinlichkeit", происходит от слова "Wahr" – истинный; настоящий; верный; действительный; подлинный; правильный; правдивый. У американцев такого нет, потому что латыни они не знают. Итальянцы же латынь учат в школе... Русские, подобно итальянцам тоже страдают от близости слов "вероятность" и "вера", "верить". И, наконец, представители любых национальностей нередко забывают аксиомы: теория вероятностей имеет смысл лишь для событий, которые можно (в идеале – неограниченно) много раз повторять. По этой причине, например, фраза "завтра будет дождь с вероятностью 30% " может иметь смысл, в отличие от фразы "завтра Dow Jones упадёт как минимум на 5% с вероятностью 30% ". Да, обвалы индексов случались уже много раз, но каждый обвал уникален и не похож на предыдущие, так что это совершенно разные события, и ни о какой повторяемости здесь речи быть не может. Вопрос о вероятности биржевого краха столь же абсурден, как и вопрос "с какой вероятностью Бог существует?" Замечательно также заявление "Вероятность дефолта Украины в ближайшие три года выросла до 17,8%, говорится в исследовании агентства Bloomberg." – и такие вещи с умным видом говорят миллиардеры!
- Финансовый или политический "эксперт", предсказание которого не сбылось, не испытывает чувства вины, оправдываясь тем, что произошло непредвиденное им событие, неучтённое в его модели.
  Например (см. http://www.gazeta.ru/business/news/2016/08/26/n_9042659.shtml), "Японский государственный пенсионный инвестиционный фонд (GPIF), крупнейший в мире, завершил второй квартал с убытком в $52 млрд, в чем частично винит решение Великобритании о выходе из Евросоюза..." "По мнению президента фонда Норихиро Такахаси, на ситуацию на рынке негативно повлияли два фактора. «Результаты референдума в Великобритании отличаются от тех, что ожидал рынок. И данные по безработице в США в мае было гораздо хуже, чем прогнозировалось», — заявил он."
- Для оценки человека интересно выяснять, достаточно ли у него ума и смелости, чтобы признаться в незнании чего-либо.
- Цитата из Талеба к вопросу о ставшем популярном в последние 10 лет жанре альтернативной истории: "Эмигранты, как правило, становятся пленниками собственных идиллических воспоминаний — они сидят в компании других пленников прошлого и говорят о родине, вкушая традиционные блюда под звуки народной музыки. Они постоянно проигрывают в уме альтернативные сценарии, которые могли бы предотвратить их историческую трагедию — например: «если бы шах не назначил ту бездарь премьер-министром, мы и сейчас были бы дома». Словно исторический перелом имел конкретную причину и катастрофу можно было бы предотвратить, ликвидировав ту конкретную причину."
- Талеб: "Каждый божий день случались неожиданности, опровергавшие их прогнозы, но никто не замечал, что они не были предсказаны." И, словно в подтверждение, через несколько дней (2016-05-11) читаю в новостях "Министерство энергетики США резко повысило прогноз цен на нефть на текущий и следующий годы, сообщается в пресс-релизе ведомства."

[41] Пример: изобретение и постепенное широкое распространение компьютеров, кардинально изменившее общество.

[42] Почти всегда "кому – война, а кому – мать родна." Будем называть "злыми" тех чёрных лебедей, которые вредят (лишают капитала и жизни) большему числу людей, чем помогают.

- “Во время ливанской войны я также заметил, что журналисты имеют тенденцию группироваться, причем не столько вокруг одинаковых мнений, сколько вокруг одинаковых методик анализа. Они придают значение одним и тем же наборам обстоятельств и подразделяют реальность на одинаковые категории (опять проявление платонизма, потребности разложить все по полочкам).” Это свойственно также и учёным, и, вероятно, любым людям, чья профессия сопряжена с продажей собственных мнений. Этот эффект массовости, по-видимому является неотъемлемой частью человеческой психики. И в покере, вероятно, люди охотнее рискуют, если то же самое делают соседи по карточному столу. Добавлю, что увлечение этими вашими фейсбуками должно приводить к усилению “эффект массовости”: люди видят больше чужих мнений и у них возрастает соблазн принять одно из чужих (часто групповых) мнений как своё.
- “Сегодняшний альянс между христианскими фундаменталистами и израильским лобби, безусловно, поставил бы в тупик интеллектуала XIX столетия: христиане были антисемитами, а мусульмане — защитниками евреев, которых они предпочитали христианам.”
- “Доступная публике информация совершенно бесполезна, особенно для бизнесмена, поскольку цены, как правило, уже «включают» всю подобную информацию; то, что известно миллионам, не дает вам реального преимущества. … Поняв это, я полностью отказался от газет и от телевизора…” Я вижу здесь аналогию с покером: книги о нём бесполезны.
- “я с изумлением обнаружил, что финансовые неприятности могут деморализовать сильнее, чем война (вдумайтесь в то, что финансовые потери и сопутствующее унижение могут приводить к самоубийству, а война, насколько мне известно, нет).”
- “капитал … защищает вас от умственной проституции и освобождает от давления извне — любого давления, (Независимость — понятие относительное: меня всегда поражало количество людей, которых астрономические доходы превращают в совершенных лакеев, все усиленнее лебезящих перед клиентами и работодателями и все больше пожираемых страстью к наживе.)”
- Отметил нелогичное до смешного поведение водителей (азарт). Стоим на светофоре, я газанул раньше и обогнал. Но водитель в другом ряду идёт на значительное превышение скорости, лишь бы только обогнать меня. И обгоняет. Я уверен, что если бы он один стоял на светофоре, он не стал бы так разгоняться.
- Опять об индюшке: “…опыт индюшки имеет не нулевую, но отрицательную ценность. Она строила свои заключения на наблюдениях, как нам всем рекомендуют (в конце концов, это считается научным методом).” И тут уместно вспомнить о замечании из грустной заметки Сталинги/Хмельницкого о смерти науки: “наука может лишь опровергать, но не утверждать.” Кстати, этим наверняка пользуются игроки в покер, меняющие стратегии. Они хотят, чтобы о них думали “он никогда не блефует, если уж он сделал большую ставку, наверняка у него очень сильные карты” или, наоборот, “он часто блефует, его ставка ничего не значит”. Вывод: иногда уместнее ради подтверждения своей серьёзной репутации избежать мелкого выигрыша, лишь бы только поменьше открывать свои карты. И ещё об индюшке: “её чувство безопасности росло по мере приближения судного дня” потому, что она не понимала, *почему* её кормят. Она удовлетворялась расплывчатыми ответами вроде “так устроен мир”. Кстати, в этом слабость религиозных людей: они не настаивают на объяснениях. А ещё именно так, по-индюшачи, устроены кредитные рейтинги. Таким образом система уязвима перед мелким жуликом такого типа (я читал описание такой схемы от первого лица): “я подниму свой кредитный рейтинг за несколько лет, а потом возьму в долг побольше и уеду из США в Россию!” На самом деле, конечно, уязвимости нет, потому что почти никто добровольно не вернётся из США в нынешнюю Россию. Даже за 30-40 тысяч долларов, которые теоретически можно набрать в долг с хорошим рейтингом.
- “После краха фондового рынка в 1987 году половина американских трейдеров с ужасом ожидала приближения каждого следующего октября, не принимая во внимание, что у первого кризиса предшественника не было. Мы слишком склонны беспокоиться постфактум. То, что наивное наблюдение в прошлом мы принимаем за нечто окончательное и показательное для будущего, — это единственная причина нашей неспособности понять Черного лебедя.” То есть эти люди уподоблялись азартным игрокам в “однорукого бандита”, которые позволяют себе делать суждения о вероятности выигрыша на основании предыстории. Они думают, что если на каком-то автомате долго не выпадал выигрыш, на нём “накапливается удача”, что удача эта подобна груше, которой надо дать время созреть, а затем сорвать и съесть.
- “Верящим в безусловную полезность прошлого опыта полезно будет ознакомиться с ужасно мудрым высказыванием, якобы принадлежащим одному известному морскому волку: «За всю свою профессиональную жизнь я ни разу не попадал ни в какую хоть сколько-нибудь серьезную аварию. За все свои годы на море я видел только одно судно, терпящее бедствие. Я никогда не видел крушения, не переживал крушения, не оказывался в ситуации, которая грозила катастрофой.» (Э. Дж. Смит, капитан «Титаника», 1907 г.) В 1912 году судно капитана Смита потерпело самое знаменитое кораблекрушение в истории человечества.”
- “Сэр Фрэнсис Гальтон, двоюродный брат Чарльза Дарвина и внук Эразма Дарвина, был наряду со своим кузеном одним из последних независимых ученых-джентльменов, к каковым также принадлежали лорд Кавендиш, лорд Кельвин, Людвиг Витгенштейн (на свой лад) и отчасти наш суперфилософ Бертран Рассел. Хотя Джон Мэйнард Кейнс не вполне вписывался в эту категорию, он мыслил в унисон с ней. Гальтон жил в викторианскую эпоху, когда обладатели наследственного состояния и неограниченного досуга не только упражнялись в верховой езде и стрельбе по дичи, но становились философами, учеными или (менее одаренные) политиками. Как это ни печально, вместе с той эпохой ушло нечто невосполнимое: истинные подвижники, занимающиеся наукой ради науки, не думающие о карьере.”

- “Удивил – наполовину победил, ломай привычную тактику, не действуй по шаблону, не позволяй сопернику просчитать свои действия, и ты сможешь победить. . . ”

19 Заключение

Теория вероятностей не заменяет разведку и может обеспечить преимущество перед плохо знакомыми с ней конкурентами только в случае, если они не обладают большей информацией об объекте предсказаний. Вернёмся, для примера, к задаче номер 7 о предсказаниях итога боксёрского поединка. Если я буду строить прогноз лишь на основании общеизвестной информации, в то время как мои конкуренты добудут развед. данные (инсайд) о том, что от одного из боксёров на этой неделе ушла любовница и теперь он впал в депрессию и каждый вечер нажирается с мордобоем в баре, то и без всякой теории вероятностей прогноз моих соперников будет лучше моего.

19.1 Об индюшке и курице

Излюбленный пример Талеба [29] с индюшкой, ошибочно делающей благоприятные прогнозы будущего на основании того, что в прошлом её заботливо вскармливали, не так уж и очевиден. Для начала предположим, что речь идёт не об индюшке, а об обычной курице.

Курица ошибочно предполагала верность эргодической гипотезы: она думала что усреднение по времени эквивалентно усреднению по курицам. Это, возможно, и так, но только в случае, если период усреднения значительно превышает длительность жизни курицы. Аналогично тому, как при изучении климата некоторой местности нельзя делать выводы на основании слишком короткого промежутка времени (напр., лета). Но что делать, если нам не позволяют подождать год, если мы обязаны сделать хоть какое-нибудь суждение о климате за неделю в июле? Ответ: проанализировать данные о других местах планеты, напр., в другом полушарии. Или найти несколько других планет, похожих на Землю (у которых на данной широте и долготе сейчас разные времена года).

Но индюшка в США – на особом положении: по американской традиции, раз в год для индюшек наступает Судный День (люди называют его Днём Благодарения). Поэтому усредняя по ансамблю куриц (которых убивают равномерно, без этих авралов) я делаю меньшую ошибку, чем усредняя по ансамблю индюшек. Вероятно, многие экономисты ошибочно описывают, скажем, рынок акций так, как будто это была бы стая (или выводок?) кур на большой ферме. На самом же деле более уместна аналогия с индюшками, которых всех режут в один день.

В современном мире формулы теории вероятностей (проверки гипотез) используются повсеместно. Рискну, однако же, предположить (и в этом я не одинок – см. [29]), что в большинстве случаев люди не понимают сути этих формул и области их применимости, и, следовательно, нередко используют их неправильно или вообще в ситуациях, где это недопустимо. Тот же Талеб [29] задавал людям вопросы вроде: “оцените диапазон количества книг в библиотеке Умберто Эко или любовников Екатерины II с вероятностью 98%. . . ” Люди отвечали ему “с вероятностью 98% в постели Екатерины II побывало не менее 10, но не более 25 мужчин”, и никто, судя по всему, не ответил, что вопрос его абсурден, так что никакой ответ не будет иметь смысла. Это иллюстрирует родовую травму Талеба: при всём своём уме он конченный гуманитарий.

Отрывок истории с anekdot.ru . . . После того как рубль совсем упал, а обменники начали удлинять табло обменных курсов, именно Моисеич менял свои франки на рубли. И так происходило еще дважды, менялись только валюты - вместо франков были кроны, какие-то реалы. Он словно чувствовал лавину, пока мы просто созерцали происходящее, наш предел был пойти что-то поменять на доллары или евро, а потом на волне паники с убытком для себя их сдать обратно. Но Моисеич делал все с максимальной и демонстративно холодной расчетливостью и соответственно получал максимальную выгоду. Будь человеком! взмолились мы - объясни, как ты это делаешь? Моисеич поцокал языком, и сказал - ну ребята, это же очевидно: вы никогда не чините ботинки - вы идете и покупаете новые, потому что молодым поколениям навязано мнение, что чинить старые вещи невыгодно. Вот и с деньгами так же, посмотрите на волну – если на главных новостных сайтах появляется много однотипно резких мнений авторитетных экспертов-аналитиков, значит надо делать ровно наоборот. Потому что эти аналитики – именно те, кто вам вбивали в голову, что вместо замены каблука надо купить новую пару.

Предметный указатель

Список литературы

[1] Бернард Шутц. *Геометрические методы математической физики.* М.:Мир, 1984. 3, 21

[2] Robert Hermann. *Cartanian geometry, nonlinear waves, and control theory. Part B.* Math Science Pr, 1980. 3, 17, 21

[3] А. Н. Колмогоров and С. В. Фомин. *Элементы теории функций и функционального анализа.* М.:Физматлит, 7 edition, 2004. 4

[4] А. Д. Александров and Н. Ю. Нецветаев. *Геометрия.* М.:Наука, 1990. 4

[5] В. А. Зорич. *Математический анализ*, volume 1. М.:НЦМНО, 6 edition, 2012. 4

[6] В. А. Зорич. *Математический анализ*, volume 2. М.:НЦМНО, 6 edition, 2012. 4

[7] Б. Л. Гуревич Г. Е. Шилов. *Интеграл, мера и производная. Общая теория.* М.:Наука, 1967. 4, 29

[8] В. С. Владимиров. *Уравнения математической физики.* М.:Наука, 4 edition, 1981. 4

[9] Ю. Н. Бибиков. *Курс обыкновенных дифференциальных уравнений.* М.:ВШ, 1991. 4

[10] Ю. В. Сидоров, М.В. Федорюк, and М. И. Шабунин. *Лекции по теории функций комплексного переменного.* М.:Наука, 3 edition, 1989. 4

[11] А. Картан. *Элементарная теория аналитических функций одного и нескольких комплексных переменных.* М.:ИИЛ, 1963. 4

[12] М. А. Евграфов. *Аналитические функции.* М.:Наука, 3 edition, 1991. 4

[13] Ю. А. Розанов. *Лекции по теории вероятностей.* М.:Наука, 1968. 4, 24, 27

[14] Валерий Николаевич Тутубалин. Курс лекций по теории вероятностей. Краткий (35-стр.), но интересный конспект лекций – хорош для быстрого поверхностного изучения., 2008. 4

[15] Валерий Николаевич Тутубалин. *Теория вероятностей в естествознании.* Знание, 1972. 4, 26, 46

[16] Т. А. Леонтьева, В. С. Панфёров, and В. С. Серов. *Задачи по ТФКП с решениями.* М.:Мир, 2 edition, 2005. 4

[17] Стернберг С. *Лекции по дифференциальной геометрии.* Мир, 1970. 4

[18] С.Стернберг В.Гийемин. *Геометрические асимптотики.* 4

[19] Гулин А.В. Самарский А.А. *Численные методы математической физики.* Научный мир, 2000. 4

[20] Соболев С.Л. *Уравнения математической физики.* Наука, 5 edition, 1992. 4

[21] Ибрагимов Н.Х. *Группы преобразований в математической физике.* 1983. 4

[22] Севастьянов Б. *Курс теории вероятностей и математической статистики.* Наука, 1982. 4

[23] Рашевский. *Риманова геометрия и тензорный анализ.* Наука, 3 edition, 1967. 4

[24] Феллер В.(Feller W.). *Введение в теорию вероятностей и ее приложения.*, volume 1. Мир, 3изд. edition, 1984. 4, 24, 25, 35, 49

[25] Феллер В.(Feller W.). *Введение в теорию вероятностей и ее приложения.*, volume 2. Мир, 3изд. edition, 1984. 4

[26] Булдырев В. С. and Павлов Б. С. *Линейная алгебра и функции многих переменных.* ЛГУ, 2009. 5

[27] Борис А. Дубровин, Сергей Петрович Новиков, and Анатолий Т. Фоменко. *Современная геометрия: методы и приложения.* Наука, 2 edition, 1988. 16, 17

[28] Вентцель Е.С. *Теория вероятностей.* Высш. шк, 10 edition, 2006. 25, 26, 33

[29] Нассим Николас Талеб. *Чёрный лебедь.* М., Колибри, 2009. Обсуждение этой книги – неплохой способ выяснить, понимает ли собеседник теорию вероятностей, или же он всего лишь заучил её формулы. 26, 47, 48, 51

[30] *The state of science in 2014*, volume 2, 2 2014. 27

[31] Ван дер Варден Б.Л. *Математическая статистика.* ИЛ, 1960. 27, 38, 41

[32] Гаральд Крамер. *Методы математической статистики.* М.:Мир, 2 edition, 1975. 40

[33] Г. И. Ивченко and Ю. И. Медведев. *Математическая статистика.* М.:В.ш., 1984. 40

[34] Л. Н. Большев and Н. В. Смирнов. *Таблицы математической статистики.* М.:Наука, 3 edition, 1983. 40

[35] П. Хьюбер. *Рабастность в статистике.* М.:Мир, 1984. 40

[36] Г. И. Ивченко and Ю. И. Медведев. *Введение в математическую статистику.* М.:ЛКИ, 2010. 40

[37] В. А. Ватутин, Г. И. Ивченко, Ю. И. Медведев, and В. П. Чистяков. *Теория вероятностей и математическая статистика в задачах.* М.:Дрофа, 2 edition, 2003. 45

[38] P. Diaconis. *Group Representations in Probability and Statistics.* IoMS, 1988. 48