Конспект к экзамену по билетам (математический анализ) (3-й семестр)

Латыпов Владимир (конспектор)
t.me/donRumata03, github.com/donRumata03, donrumata03@gmail.com

Лимар Иван Александрович (лектор)
https://t.me/limvan

3 мая 2023 г.

Содержание

1. Как работать с этим сжатым конспектом

Составлено в соответствии с лекциями весны 2023

2. Определения

Определение (Веростностное пространство). Это пространство с *вероятностной* (то есть $P(X) = 1$) мерой: мера должна быть счётно-аддитивной функцией $2^X \to [0, \infty)$ на σ-алгебре.

Используется «птичий язык»:

$$AB \stackrel{\text{def}}{=} A \cap B$$
$$A + B \stackrel{\text{def}}{=} A \cup B$$
$$\overline{A} \stackrel{\text{def}}{=} A^{\complement}$$

Почему определяем на какой-то странной сигма-алгебре, а не на полной (2^X)?

В случае с $\mathbb{R}^n$ — на всём не получится сделать адекватную меру, так как, например, если в $\mathbb{R}$ объявим $\mu[0, 1] = 1$, то множество Витали будет неизмеримо.

(Вспомним из матана, что вообще любая мера, инвариантая относительно сдвига, на той же сигма-алгебре — в константу раз отличается от меры Лебега).

Определение (Вероятностное пространство *в широком смысле*). Теперь работаем в алгебре, а мера — счётно-дизъюнктно аддитивна на множествах, объединение которых уже лежит в алгебре.

Теорема 1 (Единственность стандартного распространения). *...веростностной меры с веростностного пространства в широком смысле на вероятностное пространство в обычном, а именно — на .*

Доказательство. Как легко видеть, $\left|\bigoplus_{k\in S}\left(\mathfrak{K}^{\mathbb{F}^{\alpha}(i)}\right)_{i\in\mathcal{U}_k}\right| \preccurlyeq \aleph_1$ при $[\mathfrak{H}]_{\mathcal{W}} \cap \mathbb{F}^{\alpha}(\mathbb{N}) \neq \emptyset$. ■

Замечание. Из матана известно, что достаточно потребовать первоначальное задание меры на полукольце и сигма-конечности, чтобы она совпадала со стандартным распространением на сигма-алгебре измеримых.

Пример. Примеры веростностных пространств:

1. Дискретное: состоит из элементарных исходов, у каждого вес. $\mathbb{A} = 2^\Omega$, $P(A) = \sum_{w \in A} w$

   (a) Броски монеты до первого орла

   (b) Модель классической вероятности: $\forall i : w_i = \frac{1}{n}$. *Колчичество* элементарных исходов в событии считается комбинаторикой.

   Пример: шарики и перегородки кодируют k-элементные мультимножества n объектов или же n-кортежи длины k.

2. Геометрическая вероятность. $\Omega \subset \mathbb{R}^n, \Omega \in \mathbb{A}_n$, $P(A) = \frac{P(A)}{P(\Omega)}$. Пример: вычисление π Монте-Карловскими бросками иголки (считаем меру допустимого множества, интегрируя его сечение по проекции).

Свойство 2.1 (Элементарные свойства веростности). · *Монотонность*

- $P(\overline{A}) = 1 - P(A)$
- *Включения-исключения*
- *Полуаддитивность*

Лекция 2

Теорема 2 (Равносильность непрерывности и счётной аддитивности объёма). *Утверждения равносильны:*

1. P — *мера*
2. P — *объём, непрерывный снизу*
3. P — *объём, непрерывный сверху*

Доказательство. $2 \Leftrightarrow 3$: инвертируем.

$(2, 3) \Leftrightarrow 1$: разбиваем на кольца, остаток сходящегося ряда $\to 0$. ■

Теорема 3 (Формула полной вероятности). *Пусть* $\{A_i\}^n$ *дизъюнктны,* $B \in \bigcup_i A_i$.

Тогда $P(B) = \sum_i P(A_i)P(B|A_i)$.

Теорема 4 (Байеса).

$$\underbrace{P(A|B)}_{\text{likelihood}} = \frac{\overbrace{P(A)}^{\text{prior}}\ \overbrace{P(B|A)}^{\text{likelihood}}}{\underbrace{P(B)}_{\text{marginal}}} \tag{2.1}$$

Можно переписать в виде:

$\{A_i\}$ — система дизъюнктных событий, $B \in \bigcup A_i$. (((Каждое из них „могло вызвать“ B и какое-то точно вызвало))). Вопрос — какое:

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(B)} = \frac{P(A_i)P(B|A_i)}{\sum_i P(A_i)P(B|A_i)} \tag{2.2}$$

То есть при получении информации, что произошло B, ожидания событий скейлятся пропорционально тому, насколько вероятно они вызывают B.

3. Процесс Бернулли, предельные теоремы

Процесс Бернулли: серия экспериментов подбрасывания p-монетки (p может как меняться, так и не меняться).

Предельными теоремами можно аппроксимировать биномиальное (или более извращённое, но порождённое процессом Бернулли) распределение

Теорема Пуассона: аппроксимация $P(S_n = k)$ для $p_n \frac{\lambda}{n}$ распределением Пуассона: $e^{\lambda} \frac{\lambda^k}{k!}$.

(Локальная) теорема Муавра-Лапласа: асимптотическое поведение $P(S_n = k)$ при $n, (n-k) \to \infty$.

Интегральная теорема Муавра-Лапласа (частный случай ЦПТ): аппроксимация биномиального распределения нормальным ($F_{\text{Bin}} \approx \text{erf}$).

4. Переход вероятностному пространству распределения

Случайная величина — $\in \mathcal{B}\left(\Omega \to \mathbb{R}\right)$ (измерима относительно сигма-алгебры этого в.п.).

Распределение с.в.: $P_X : \mathcal{B}_1 \to \mathbb{R}$.

$$P_X(B) \overset{\text{def}}{=} P\left(\{\omega | X\left(\omega\right) \in B\} = P\left(X^{-1}(B)\right) \overset{\text{def}}{=} P\left(X \in B\right)\right) \tag{4.1}$$

Получили веростностную меру на борелевской σ-алгебре $\mathcal{B}_1$.

Вроде и существует какое-то вероятностное пространство с каким-то множеством исходов, но часто будем говорить о некоей «проекции» этой информации — о функции распределения случайной величины: $P_X\left(A\right) = P$

Абсолютно непрерывная с.в., если найдётся p_X, т.ч.: $P_X\left(A\right) = \int_A p_X \,\mathrm{d}\mu$

5. Примеры дискретных распределений

- Одноточечное $I_c : P\left(I_c = c\right) = 1$
- Бернулли: X Bern $(p) \Leftrightarrow \begin{cases} P(X=0) = 1-p \\ P(X=1) = p \end{cases}$
- Бионмиальное: X Bin $(n,p) \Leftrightarrow P(X=k) = \binom{n}{k} p^k q^{n-k}$
- Обратное биномиальное (вероятность, что продолбаем k лишних шагов до достижения r-того успеха): X NB (r,p) $\Leftrightarrow$ $X = \min\{n | S_n \geqslant r\} - r; P(X=k) = \binom{r-1}{k+r-1} p^r q^k$...
- Частный случай — геометрическое распределение: количество неудач до первого выпадения удачи:

6. Примеры непрерывных распределений

Юниформа, автомат и противогаз

Намаааа

Гамма

Пуассон

Экспоненциальное

7. Случайные векторы

Задаётся совместная функция вероятности P_X, уже потом можно из неё получить маргинальные распределения.

$$P_X(B_1, B_2, \dots, B_n) = P(X_1 \in B_1, X_2 \in B_2, \dots, X_n \in B_n), B_i \in \mathcal{B} \quad (7.1)$$

Функция распределения:

$$F_X(x_1, x_2, \dots, x_n) = P(X_1 \leqslant x_1, X_2 \leqslant x_2, \dots, X_n \leqslant x_n) \quad (7.2)$$

Отсюда можно выразить P для отрезков через F. Введём разностный оператор Δ_i: $\Delta_{i,a,b} f(x_1, \dots, x_n) = f(x_1, \dots, x_{i-1}, b, x_{i+1}, \dots, x_n) - f(x_1, \dots, x_{i-1}, a, x_{i+1}, \dots, x_n)$.

Тогда

$$P(X_i \in [a_i, b_i]) = \Delta_{i,a_i,b_i} \dots \Delta_{1,a_1,b_1} F_X(x_1, \dots, x_n) \quad (7.3)$$

Мультиномиа́льное (полиномиа́льное) распределе́ние — обобщение биномиального распределения на случай произвольного числа исходов.

$$P(X_1 = k_1, \dots, X_n = k_n) = \binom{n}{k_1, \dots, k_n} p_1^{k_1} \dots p_n^{k_n} \quad (7.4)$$

(то есть коэфициент при $t_1^{k_1} \dots t_n^{k_n}$ в формальном многочлене $(p_1 t_1 + \dots + p_n t_n)^n$)

Нормальное распределение для случайного вектора $X \sim \mathcal{N}(\mu, \Sigma)$:

$$p_X(x) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\right) \quad (7.5)$$

Расшифровка такая: для стандартного нормального распределения $N(0, \mathbb{I}_n)$:

$$p_X(x) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \tag{7.6}$$

то есть как для n независимых с.в. $X_i \sim \mathcal{N}(0,1)$.

А для произвольных μ, Σ: $Y = \mu + \sqrt{\Sigma}X$. Матрица, на самом деле — матрица ковариации, а вектор μ — матожидание.

8. Свёртки

Нагрузка вероятности для суммы с.в. — свёртка нагрузок:

$$p_{X+Y}(z) = \sum_{x \in \mathbb{R}} p_X(x) p_Y(z-x) \tag{8.1}$$

Абсолютно непрерывные с.в.:

$$F_{X+Y}(z) = \int_{x+y \leq z} p_X(x) p_Y(y) \, \mathrm{d}x \, \mathrm{d}y \tag{8.2}$$

$$p_{X+Y}(z) = \int_{\mathbb{R}} p_X(x) p_Y(z-x) \, \mathrm{d}x \tag{8.3}$$

Маргинальная плотность: $p_{X_i}(x_i) = \int_{\mathbb{R}^{n-1}} p_X(x_1, \dots, x_n) dx_1 \dots \mathrm{d}x_{i-1} \, \mathrm{d}x_{i+1} \dots \mathrm{d}x_n$, интегрировать можно по проекции $\mathbb{Pr}_{1,\dots,i-1,i+1,\dots,n}\mathbb{R}^n$: $\{(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) | \exists x_i \in \mathbb{R}\colon p_X(x_1, \dots, x_n) > 0\}$.

9. Независимые случайные величины

Определение. Случайные величины $X_1, \dots, X_n$ называются независимыми, если $\forall B_1, \dots, B_n \in \mathcal{B}(\mathbb{R})$: $P(X_1 \in B_1, \dots, X_n \in B_n) = \prod_{i=1}^{n} P(X_i \in B_i)$.

Определение. Последовательность случайных величин $\{X_i\}_{i=0}^{\infty}$ называется независимой, если $\forall m \in \mathbb{N}\colon X_1, \dots, X_m$ независимы.

Теорема 1 (Критерий независимости). *Случайные величины $X_1, \dots, X_n$ независимы тогда и только тогда, когда $\forall x_1, \dots, x_n \in \mathbb{R}$:*

$$F_{X_1,\dots,X_n}(x_1, \dots, x_n) = \prod_{i=1}^{n} F_{X_i}(x_i) \tag{9.1}$$

Доказательство. Аналогично теореме о задании случайной величины функцией распределения. ■

Теорема 2 (Критерий независимости с.в. для дискретного случая). *Случайные величины $X_1, \dots, X_n$ независимы тогда и только тогда, когда $\forall x_1, \dots, x_n \in \mathbb{R}$:*

$$p_{X_1,\dots,X_n}(x_1, \dots, x_n) = \prod_{i=1}^{n} p_{X_i}(x_i) \tag{9.2}$$

Доказательство. ($\Rightarrow$) Очевидно из определения

($\Leftarrow$) Пусть $B_1, \dots, B_n \in \mathcal{B}(\mathbb{R})$. Тогда:

$$\begin{aligned}
P(X_1 \in B_1, \dots, X_n \in B_n) &= \\
&\sum_{(x_1,\dots,x_n) \in B_1 \times \dots \times B_n} p_{X_1,\dots,X_n}(x_1, \dots, x_n) = \\
&\sum_{(x_1,\dots,x_n) \in B_1 \times \dots \times B_n} \prod_{i=1}^{n} p_{X_i}(x_i) = \\
&\sum_{x_1 \in B_1} \dots \sum_{x_n \in B_n} \prod_{i=1}^{n} p_{X_i}(x_i) = \\
&\prod_{i=1}^{n} \sum_{x_i \in B_i} p_{X_i}(x_i) = \\
&\prod_{i=1}^{n} P(X_i \in B_i)
\end{aligned} \tag{9.3}$$

■

Теорема 3 (Критерий независимости с.в. для абсолютно непрерывного случая). *Случайные величины $X_1, \dots, X_n$ независимы тогда и только тогда, когда* $\forall x_1, \dots, x_n \in \mathbb{R}$:

$$p_{X_1,\dots,X_n}(x_1, \dots, x_n) = \prod_{i=1}^{n} p_{X_i}(x_i) \tag{9.4}$$

Доказательство. Через общий критерий независимости и ингтегрирование/дифференцирование. ■

9.1. Некоторые распределения в связи с независимостью

Пусть $X_1, \dots, X_n$ независимы и $X_i \sim \text{Bern}(p)$. Тогда $X_1 + \dots + X_n \sim \text{Bin}(n, p)$.

Для Пуассона: параметры складываются.

Нормальное: сумма нормальных независимых нормальных нормальна. Медианы и средние кваадратичные отклонения складываются.

Сумма r независимых геометрических $\sim \text{NB}\,(r, p)$

10. Об интегрировании

Мера образа множества при отображении.

$$P_X(B) = P(X \in B) = P(\{\omega \colon X(\omega) \in B\}) = P(X^{-1}(B)) \tag{10.1}$$

Мера, заданная таким образом, называется:

PushForward measure.

Пусть есть функция от с.в. $g \colon \mathbb{R}^n \to \mathbb{R}$.

Тогда

$$\int_\Omega g(X(w))P(\mathrm{d}w) = \int_{\mathbb{R}^n} g(x)P_X(\mathrm{d}x) = \begin{cases} \sum_i g(x_i)p_X(x_i) & \text{дискретный случай} \\ \int_{\mathbb{R}^n} g(x)p_X(x)\,\mathrm{d}x & \text{абсолютно непрерывный случай} \end{cases} \tag{10.2}$$

Особый случай для $n = 1$: интеграл Лебега-Стилтьеса.

Теорема Фубини/Тонелли: проецируем на часть осей, деля интеграл на два вложенных.

11. Числовые характеристики случайных величин

Определение. Пусть X — случайная величина. Тогда её математическим ожиданием называется число

$$\mathbb{E}X = \int_{\mathbb{R}} x \, \mathrm{d}P_X(x) \tag{11.1}$$