

САНКТ-ПЕТЕРБУРГСКИЙ ПОЛИТЕХНИЧЕСКИЙ
УНИВЕРСИТЕТ им. ПЕТРА ВЕЛИКОГО
Институт прикладной математики и механики

Теория вероятностей. Конспект
лектор: к.б.н. Н. О. Кадырова

Последняя лекция — 15/11/2018

В. А. Тюльпин
Санкт-Петербург
2018 г.

Содержание

Стр.

| | | |
|----|---|----|
| I | Исчисление вероятностей случайных событий | 4 |
| 1 | Соотношение между случайными событиями | 4 |
| 2 | Классическое определение вероятности | 6 |
| 3 | Вероятность и относительная частота | 7 |
| 4 | Геометрические вероятности | 8 |
| 5 | Условная вероятность. Теорема умножения вероятностей. Независимость событий | 11 |
| 6 | Теорема сложения вероятностей | 15 |
| 7 | Формула полной вероятности и формула Байеса | 17 |
| 8 | Последовательность испытаний Бернулли. Биномиальный закон распределения вероятностей. Закон распределения вероятностей Пуассона | 18 |
| 9 | Полиномиальное распределение вероятностей | 20 |
| 10 | Вероятностные производящие функции | 22 |

| | | |
|------------|--|-----------|
| II | Производящее пространство элементарных событий. Случайные величины и векторы | 24 |
| 1 | Аксиомы теории вероятностей. Вероятностное пространство | 24 |
| 2 | Свойства вероятности | 27 |
| 3 | Способы задания вероятностных мер на измеримом пространстве $(\mathbb{R}^1, \mathcal{B}(\mathbb{R}^1))$ | 31 |
| 4 | Случайные величины | 43 |
| 5 | Способы задания вероятности на вероятностном пространстве $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ | 48 |
| 6 | Случайный вектор и его распределение на вероятностном пространстве $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ | 50 |
| 7 | Независимость случайных величин | 53 |
| III | Числовые характеристики распределения случайных величин | 57 |
| 1 | Математическое ожидание случайной величины | 57 |
| 2 | Дисперсия случайной величины | 61 |
| 3 | Матрица ковариаций случайного вектора . . . | 64 |
| 4 | Коэффициент корреляции случайной величины | 66 |
| 5 | Моменты случайных величин произвольных порядков | 69 |

| | | |
|---|--|----|
| 6 | Законы распределения функций случайных величин | 71 |
|---|--|----|

| | | |
|----|--|----|
| IV | Характеристические функции. Предельные теоремы | 81 |
|----|--|----|

| | | |
|---|---|----|
| 1 | Производящие функции. Факториальные моменты | 81 |
|---|---|----|

DRAFT

Сообщить об опечатках или скачать последнюю версию конспекта можно в слаке студентов ИПММ (да, у нас есть Slack) в `#_bachelors_3`. Вас ещё там нет? Жаль, но вот инвайт (действителен до 15/12):
<https://clck.ru/EjCxr>
(актуальный PDF каждое воскресенье)

Глава I

Исчисление вероятностей случайных событий

§1 Соотношение между случайными событиями

Пусть имеется фиксированный комплекс условий, воспроизводимый любое количество раз, в том числе и неограниченное.

Определение. Комплекс испытаний будем называть *экспериментом*.

Определение. Явления в результате испытаний обозначим как *события*.

Рассмотрим связанную с этим комплексом условий *систему событий*.

$$\{A, B, C, \dots\}, \{A_1, A_2, \dots, A_n\}$$

1.1 Соотношения между событиями

1. U — *достоверное событие*, в результате испытаний происходит всегда.
2. V — *невозможное событие*, никогда не происходит.
3. *Сумма событий* $(\sum_{i=1}^n A_i)$ — событие, происходящее тогда и только тогда, когда происходит хотя бы одно из связанных событий.

4. *Произведение событий* $(\prod_{i=1}^n A_i)$ — событие, происходящее тогда и только тогда, когда происходят все из связанных событий.
5. Событие A — *частный случай* события B , если с появлением A появляется и B ($A \subset B$). Если A влечёт B и B влечёт A , то события A и B равносильны ($A = B$).

Доказательство. $A \subset B \Rightarrow A \times B = A, A + B = B. B = A \Rightarrow B + A = A, A + A = A.$ \square

6. \bar{A} — *противоположное событие*, если оно происходит только тогда, когда A не происходит.

$$A, \bar{A} \text{ противоположны} \Leftrightarrow \begin{cases} A \times \bar{A} = V \\ A + \bar{A} = U \end{cases}$$

7. События A и B называются *несовместными*, если их одновременное появление невозможно.

$$A \times B = V$$

8. *Разность событий* A и B состоит в том, что A происходит, а B нет. ($A - B$).

1.2 Свойства событий

Коммуникативность. $A + B = B + A, A \times B = B \times A$

Ассоциативность. $A + (B + C) = (A + B) + C, A \times (B \times C) = (A \times B) \times C$

Дистрибутивность. $(A + B) \times C = A \times C + B \times C, A + B \times C = (A + B) \times (A + C)$

1.3 Некоторые тождества

1. $A + A \times B = A$

Доказательство.

□

2. $A + B = A + \bar{A} \times B$

Доказательство.

□

3. Формула инверсий (законы Де Моргана).

$$\overline{(A \times B)} = \bar{A} + \bar{B}, \text{ можно обобщить до } \overline{\prod_{i=1}^n A_i} = \sum_{i=1}^n \bar{A}_i$$

$$\overline{(A + B)} = \bar{A} \times \bar{B}, \text{ можно обобщить до } \overline{\sum_{i=1}^n A_i} = \prod_{i=1}^n \bar{A}_i$$

События $A_1 \dots A_n$ образуют *полную группу событий*, если при каждом испытании произойдёт хотя бы одно из них.

Пример. A, \bar{A} — полная группа.

Пример. A_0, \dots, A_n — система из $n+1$ событий. Если в серии испытаний событие A_k , где $k = 0, 1, \dots, n$, происходит k раз, то A_0, \dots, A_n — полная группа.

§2 Классическое определение вероятности

Пусть результатом испытания является одно из попарно несовместных образующих полную группу равновозможных событий. Такая группа называется элементарной, а её результат — элементарным исходом.

E_1, \dots, E_n — элементарные события $\Leftrightarrow E_k \times E_i = V$, где $i, k = 0, 1, \dots, n, i \neq k$, при этом $\sum_{i=1}^n E_i = U$

Определение. E_n называется благоприятствующим для некоторого события A , если E_n влечёт A .

Определение. Вероятностью появления события A будем называть отношение числа элементарных событий m к числу исходов n .

$$P(A) = \frac{m}{n} \quad (1)$$

Замечание. При расчёте вероятности события A по формуле (1) выбор системы элементарных событий можно произвести различными способами.

Следствие. $P(U) = 1, P(V) = 0$.

Следствие. $\forall A : 0 \leq P(A) \leq 1$.

§3 Вероятность и относительная частота

Определение. Пусть при одних и тех же условиях произведено N испытаний, в результате которых интересующее событие произошло M раз.

Относительной частотой появления события A назовём число

$$\tilde{P}(A) = \frac{M}{N} \quad (2)$$

Относительная частота является величиной, рассчитываемой по факту произошедшего эксперимента, и, вообще говоря, не является оценкой вероятности, однако при достаточно большом количестве испытаний относительную частоту можно считать приближённо равной вероятности этого события. Очевидно, что

$$\tilde{P}(V) = P(V) = 0,$$

$$\tilde{P}(U) = P(U) = 1.$$

Но из равенства $\tilde{P}(A) = 0$ не следует, что $P(A) = 0$, как и в случае с $\tilde{P}(A) = 1$.

$$P(A) = \cancel{\lim_{n \rightarrow \infty} \tilde{P}(A)} = \lim_{N \rightarrow \infty} \frac{M}{N}.$$

При увеличении числа опытов отклонения (большие) относительной частоты $\tilde{P}(A)$ от вероятности $P(A)$ будут попадаться реже. На практике относительная частота $\tilde{P}(A)$ может быть принята за вероятность $P(A)$ при достаточно большом количестве испытаний.

§4 Геометрические вероятности

Пусть χ_0 — номинальный размер некоторой детали. С учётом погрешности изготовления этой детали реальный размер $\chi \in (\chi - 4; \chi + 4)$. Получено несчётное пространство, и классическое определение вероятности, заданное для счётных пространств событий, не может быть использовано для вычислений в данном случае.

4.1 Геометрический подход

Пусть имеется ограниченное одномерное евклидово множество, имеющее объём.

Рассмотрим S — систему подмножеств исходного множества Ω , имеющих объём.

$$\forall A \in S : P(A) = \frac{|A|}{|\Omega|}$$

Тройка объектов $\langle \Omega, S, P \rangle$ служит моделью задач, в которых точку бросают в область Ω .

При этом вероятность того, что некоторая точка попала в область A , пропорциональна объёму A .

4.2 Парадоксы

- Геометрическая вероятность попадания в какую-либо точку отрезка равна нулю. То есть, вероятность невозможного события равна нулю, однако обратное не является верным.
- Взаимно однозначное преобразование может существенно изменить вероятность.

4.2.1 Парадокс Бертрана

Для некоторой окружности случайно выберем хорду. Найдём вероятность того, что хорда длиннее стороны правильного треугольника, вписанного в эту окружность. A — интересное нас событие.

Метод «случайного центра». Выберем наудачу произвольную точку внутри круга и построим хорду с центром в выбранной точке. Хорда длиннее стороны равностороннего треугольника, если выбранная точка находится внутри круга, вписанного в треугольник. Площадь вписанного круга есть $1/4$ от площади большего, значит, исходная вероятность равна $1/4$.

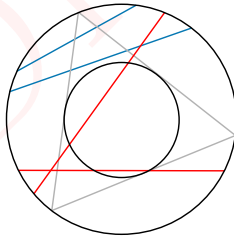


Рис. 1: Случайные хорды, выбранные в случае 1.

$$P(A) = \frac{\pi \frac{R^2}{4}}{\pi R^2} = \frac{1}{4}.$$

Метод «случайных концов». Наудачу выберем две точки на окружности и проведём через них хорду. Чтобы посчитать искомую вероятность, представим, что треугольник повёрнут так, что одна из его вершин совпадает с концом хорды. Заметим, что если другой конец хорды лежит на дуге между двумя другими вершинами треугольника, то длина хорды больше стороны треугольника. Длина рассмотренной дуги равна трети длины окружности, следуя классическому определению, искомая вероятность равна $1/3$.

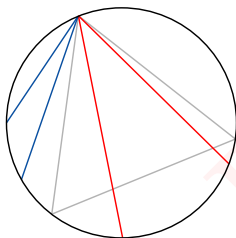


Рис. 2: Случайные хорды, выбранные в случае 2.

$$P(A) = \frac{2\pi \frac{R}{3}}{2\pi R} = \frac{1}{3}.$$

Метод «случайного радиуса». Зафиксируем радиус окружности, наудачу выберем точку на радиусе. Построим хорду, перпендикулярную зафиксированному радиусу, проходящую через выбранную точку. Для нахождения искомой вероятности представим, что треугольник повёрнут так, что одна из его сторон перпендикулярна зафиксированному радиусу. Хорда длиннее стороны треугольника, если её центр ближе к центру, чем точка пересечения треугольника с зафиксированным радиусом. Сторона треугольника делит пополам радиус, следовательно вероятность выбрать хорду длиннее стороны треугольника — $1/2$.

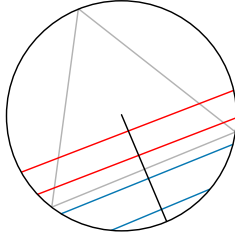


Рис. 3: Случайные хорды, выбранные в случае 3.

$$P(A) = \frac{\frac{R}{2}}{R} = \frac{1}{2}.$$

Возможны различные выборы равномерным образом, и каждый метод использует свой выбор.

§5 Условная вероятность. Теорема умножения вероятностей. Независимость событий

Пусть A, B — события, при этом $B \neq V$.

Определение. Вероятность A , вычисляемая при условии, что B произошло, называется условной вероятностью A относительно B .

$$P(A | B)$$

Если $P(A | B) = P(A)$, то A не зависит от B . В противном случае A зависит от B .

5.1 Теорема умножения вероятностей

Пусть имеется n элементарных событий $\{E_i\}_{i=1,2,\dots,n}$. Событие A происходит m раз, событие B — k раз, $A \times B$ — l раз

$(m, k, l < n)$. Тогда

$$P(A) = \frac{m}{n}, P(B) = \frac{k}{n}, P(A \times B) = \frac{l}{n}.$$

$$P(B | A) = \frac{l}{m} = \frac{\frac{l}{n}}{\frac{m}{n}} = \frac{P(A \times B)}{P(A)},$$

$$P(A | B) = \frac{\frac{l}{n}}{\frac{k}{n}}, \text{ то есть}$$

$$P(A \times B) = P(A) \cdot P(B | A) = P(B) \cdot P(A | B)$$

Установлена теорема умножения двух любых событий, хотя бы одно из которых является возможным.

Пусть A не зависит от B .

$$P(A) \cdot P(B | A) \stackrel{*}{=} P(B) \cdot P(A) \Rightarrow B \text{ не зависит от } A$$

5.2 Независимость событий

Определение. Два события A и B независимы $\Leftrightarrow P(A \times B) = P(A) \times P(B)$ (также верно, если одно из событий является невозможным).

Пример. Пусть имеется колода из 36 карт, из неё случайным образом вытаскиваем одну карту. Пусть событие A – «вытянут А (туз)», событие B – «вытянута карта масти \spadesuit ». Тогда событие $A \times B$ – вытянута карта $\spadesuit A$.

$$P(A) = \frac{4}{36} = \frac{1}{9}, P(B) = \frac{9}{36} = \frac{1}{4},$$

$$P(A \times B) = \frac{1}{36}$$

Приведём теорему об умножении n любых событий, хотя бы одно из которых возможно

$$P(A_1 \times \dots \times A_n) = P(A_1) \times P(A_2 \mid A_1) \cdot P(A_3 \mid A_1 \times A_2) \cdot \dots \\ \dots \cdot P(A_n \mid \prod_{k=1}^{n-1} A_k)$$

Доказательство.

□

$$P(A_1, A_2, A_3) = P(A_1 \times A_2) \times P(A_3 \mid A_1 \times A_2) = \\ P(A_1) \cdot P(A_2 \mid A_1) \cdot P(A_3 \mid A_1 \times A_2)$$

Попарная независимость событий

$$\begin{aligned} P(A_1 \times A_2) &= P(A_1) \cdot P(A_2) \\ P(A_3 \times A_2) &= P(A_3) \cdot P(A_2) \\ P(A_1 \times A_3) &= P(A_1) \cdot P(A_3) \end{aligned} \tag{3}$$

При независимости событий (в совокупности):

$$P(A_1 \times A_2 \times A_3) = P(A_1) \cdot P(A_2) \cdot P(A_3) \tag{4}$$

5.3 Парадокс Бернштейна

Рассмотрим испытание, в результате которого бросаются две монеты. Пусть

- событие A – «на первой монете выпал орёл»,
- событие B – «на второй монете выпал орёл»,
- событие C – «только на одной монете выпал орёл».

Установим систему элементарных событий Ω , равную $\{O-O, P-P, P-O, O-P\}$.

$$P(A) = P(B) = P(C) = \frac{1}{2}$$

$$P(A \times B) = \frac{1}{4} = P(A) \cdot P(B),$$

$$P(A \times C) = \frac{1}{4} = P(A) \cdot P(C),$$

$$P(B \times C) = \frac{1}{4} = P(B) \cdot P(C).$$

$$P(A \times B \times C) = 0 \neq P(A) \cdot P(B) \cdot P(C)$$

Вывод: попарная независимость (3) \nRightarrow независимость в совокупности (4).

События независимы в совокупности \Leftrightarrow (3) и (4) выполняются.

Определение. A_1, \dots, A_n независимы в совокупности, если

$$P(A_{i_1} \times \dots \times A_{i_k}) = P(A_{i_1}) \times \dots \times P(A_{i_k}), \text{ где}$$

$$1 \leq i_1 \leq i_2 \leq \dots \leq i_k \leq n$$

$\boxed{2^n - C_n^1 - C_n^0}$ – число соотношений для проверки независимости событий в совокупности. Если нам заранее известно, что A_1, \dots, A_n независимы в совокупности, то

$$P\left(\prod_{k=1}^n A_k\right) = \prod_{k=1}^n P(A_k)$$

§6 Теорема сложения вероятностей

Рассмотрим события A и B .

- A – m элементарных событий из n ,
- B – n элементарных событий из n ,
- $A \times B$ – l элементарных событий из n .

$$P(A + B) = \frac{m + k - l}{n} = P(A) + P(B) - P(A \times B)$$

Установлена теорема о сложении двух любых событий.

$$A \times B = V \Rightarrow P(A \times B) = P(A) + P(B)$$

Если $B = \bar{A}$, то

$$P(\underbrace{A \times \bar{A}}_{U=1}) = P(A) + P(\bar{A}) \Rightarrow P(\bar{A}) = 1 - P(A)$$

По теореме о сложении:

$$\begin{aligned} P(A_1 + A_2 + A_3) &= P(A_1 + A_2) + P(A_3) - \\ &- P((A_1 + A_2) \times A_3) = P(A_1) + P(A_2) + P(A_3) - \\ &- P(A_1 \times A_2) - P(A_1 \times A_3) - P(A_2 \times A_3) + \\ &+ P(A_1 \times A_2 \times A_3) \end{aligned}$$

Обобщённая формула:

$$\begin{aligned} P\left(\sum_{k=1}^n A_k\right) &= \sum_{k=1}^n P(A_k) - \sum_{k=1}^{n-1} \sum_{j=k+1}^n P(A_k \times A_j) + \\ &+ \sum_{k=1}^{n-2} \sum_{j=k+1}^{n-1} \sum_{i=j+1}^n P(A_k \times A_j \times A_i) - \dots \\ &\dots + (-1)^n \cdot P\left(\prod_{k=1}^n A_k\right) \end{aligned}$$

Для попарно несовместных событий:

$$A_i \times A_j = V \Rightarrow P(\sum_{i \neq j}^n A_k) = \sum_{k=1}^n P(A_k)$$

Если A_1, \dots, A_n независимы в совокупности, то справедлива следующая формула:

$$\begin{aligned} P(\sum_{k=1}^n A_k) &= 1 - P(\overline{\sum_{k=1}^n A_k}) = 1 - P(\prod_{k=1}^n \overline{A_k}) = \\ &= 1 - \prod_{k=1}^n P(\overline{A_k}) = 1 - \prod_{k=1}^n (1 - P(A_k)). \end{aligned}$$

Если вероятности исходных событий не зависят от номеров, а от количества сомножителей, то

$$\begin{aligned} P(\sum_{k=1}^n A_k) &= C_n^1 P(A_1) - C_n^2 P(A_1 \times A_2) + C_n^3 P(A_1 \times A_2 \times \\ &\quad \times A_3) - \dots + (-1)^n - 1 C_n^n P(\prod_{k=1}^n A_k) \end{aligned}$$

Доказательство. □

$$\begin{aligned} P(\prod_{k=1}^n A_k) &= \sum_{k=1}^n P(A_k) - \sum_{k=1}^{n-1} \sum_{j=k+1}^n P(A_k + A_j) + \\ &+ \sum_{k=1}^{n-2} \sum_{j=k+1}^{n-1} \sum_{i=j+1}^n P(A_k + A_j + A_i) - \dots + (-1)^{n-1} P(\sum_{k=1}^n A_k) \end{aligned}$$

Доказательство. □

§7 Формула полной вероятности и формула Байеса

7.1 Формула полной вероятности

Пусть имеется n попарно несовместных событий, составляющих полную группу, H_1, \dots, H_n — назовём их гипотезами.

- $\sum_{i=1}^n H_i = U$, где $i, j = 1, 2, \dots, n$, при этом $i \neq j$,
- $H_i \times H_j = V$, где $i, j = 1, 2, \dots, n$, при этом $i \neq j$,
- $P(\sum_{i=1}^n H_i) = \sum_{i=1}^n P(H_i) = 1$.

Также определим событие A как безусловное:

$$A = U \times A = \sum_{i=1}^n H_i \times A$$

$$P(A) = P\left(\sum_{i=1}^n H_i \times A\right) = \sum_{i=1}^n \underbrace{P(H_i)}_{\substack{\text{a posteriori} \\ \text{зав. от опыта}}} = \sum_{i=1}^n \underbrace{P(H_i) \cdot P(A | H_i)}_{\substack{\text{a priori} \\ \text{не зав. от опыта}}}$$

Формула полной вероятности.

$$P(A) = \sum_{i=1}^n P(H_i) \times P(A | H_i)$$

7.2 Формула Байеса

Получим формулу Байеса:

$$P(A \times H_k) = P(A) \cdot P(H_k | A) = P(H_k) \cdot P(A | H_k)$$

$$P(H_k | A) = \frac{P(H_k) \cdot P(A | H_k)}{P(A)} = \frac{P(H_k) \cdot P(A | H_k)}{\sum_{i=1}^n P(H_i) \cdot P(A | H_i)}$$

Формула Байеса.

$$P(H_k | A) = \frac{P(H_k) \cdot P(A | H_k)}{\sum_{i=1}^n P(H_i) \cdot P(A | H_i)}$$

§8 Последовательность испытаний Бернулли. Биномиальный закон распределения вероятностей. Закон распределения вероятностей Пуассона

8.1 Последовательность испытаний Бернулли

Пусть производится серия из n испытаний, имеется k исходов. Исходы образуют полную группу событий и являются попарно независимыми событиями. В рамках данного параграфа $k = 2$.

Определение. Испытания независимы, если все исходы этих испытаний независимы в совокупности.

Определение. n испытаний образуют последовательность испытаний Бернулли, если

1. Они независимы
2. У каждого из этих испытаний 2 исхода: $\left\{ \underset{\text{успех}}{A}, \underset{\text{неудача}}{\bar{A}} \right\}$

3. Вероятность появления исхода во всех испытаниях не изменяется.

A_i – появление A в i -ом испытании. Тогда

$$\forall i = 1, 2, \dots, n : P(A) = p \Rightarrow P(\bar{A}) = 1 - p = q$$

Пусть имеется n испытаний с $k = 2$ исходами. m – число появления A за n испытаний

$$\begin{aligned} & \overbrace{A_1 \times A_2 \times \dots \times A_m \times A_m \times \bar{A}_{m+1} \times \dots}^{C_n^m p^m q^{n-m}} \\ & \dots \times \bar{A}_n + \underbrace{\bar{A}_1 \times \bar{A}_2 \times \dots \times \bar{A}_{n-m}}_{n-m} \times \overbrace{A_{n-m+1} \times \dots \times A_n}^m. \end{aligned}$$

Формула Бернулли

$$P_m^{(n)} = C_n^m p^m q^{n-m}$$

$$\sum_{m=0}^n P_m^{(n)} = (p + q)^n = 1$$

Рассмотрим, что в n испытаний Бернулли успех A произойдёт не менее m раз.

$$R_m^n = \sum_{k=m}^n P_k^{(n)} = 1 - \sum_{k=0}^{m-1} P_k^{(n)}$$

Пример. Вероятность того, что при n испытаний A произойдёт хотя бы 1 раз:

$$R_1^{(n)} = 1 - P_0^{(n)} = 1 - q^n$$

8.2 Закон распределения вероятностей Пуассона

Рассмотрим последовательность испытаний Бернулли, когда число испытаний $n \rightarrow \infty$, при этом $p = P(A_i) \rightarrow 0$, $np = a$ — константа

$$P_m \equiv \lim_{n \rightarrow \infty} C_n^m p^m (1-p)^{n-m}$$

$$p = \frac{a}{n}$$

$$P_m = \lim_{n \rightarrow \infty} \frac{n!}{(n-m)! m!} \frac{a^m (1 - \frac{a}{n})^n}{n^m (1 - \frac{a}{n})^m} = \frac{a^m}{m!} \cdot e^{-a}, m = 0, 1, 2, \dots$$

Закон распределения вероятностей Пуассона
(редких событий)

$$\sum_{m=0}^{\infty} P_m = \sum_{m=0}^{\infty} \frac{a^m}{m!} \cdot e^{-a} = 1$$

Пример. Пусть n — число атомов в веществе, время распада — 1 секунда. Тогда $p \approx 10^{-12}$.

§9 Полиномиальное распределение вероятностей

Пусть имеется n испытаний с $k \geq 2$ исходами. A_1, \dots, A_k — исходы.

- $A_i \times A_j = V$, где $i \neq j$, $i, j = 1, 2, \dots, k$
- $\sum_{i=1}^k A_i = U$, где $i \neq j$, $i, j = 1, 2, \dots, k$
- p_1, \dots, p_k — вероятности исходов. От испытания к испытанию не изменяются, $\sum_{i=1}^k p_i = 1$, где $\forall i : p_i \geq 0$

m_1, \dots, m_k раз будут происходить A_1, \dots, A_k , $\sum_{i=1}^k m_i = n$, то есть $m_i \geq 0, m_i \in \mathbb{Z}$.

$$P_{m_1, \dots, m_k}^{(n)}$$

$$k = 2, P_{m_1, m_2}^{(n)} : m_1 = m, m_2 = n - m, p_1 = p, p_2 = 1 - p$$

$$P_{m_1, m_2}^{(n)} = \frac{n!}{m_1! m_2!} p_1^{m_1} \cdot p_2^{m_2}$$

$$k \geq 2 : P_{m_1, \dots, m_k}^{(n)} = \frac{n!}{m_1! \dots m_k!} \cdot p_1^{m_1} \cdot p_2^{m_2} \cdot \dots \cdot p_k^{m_k}$$

Любое событие представимо в виде суммы несовместных.
Пусть

$$\begin{aligned} A_k &= A'_k + A'_{k+1}, A'_k \cdot A'_{k+1} = V \\ p'_k + p'_{k+1} &= p_k, \\ m'_k + m'_{k+1} &= m_k, \\ P_k^{m'_k} \cdot P_{k+1}^{m'_{k+1}} & \end{aligned}$$

$$C_{m_k}^{m'_k} \cdot \overbrace{C_{m_k - m'_k}^{m'_{k+1}}}^{=1} = \frac{m_k!}{m'_k! m'_{k+1}!}$$

$$\text{Тогда } \forall i : m_i \geq 0, m_i \in \mathbb{Z}; \sum_{i=1}^k m_i = 0 :$$

Полиномиальный закон распределения вероятностей

$$P_{m_1, \dots, m_k}^{(n)} = \frac{n!}{m'_1! \dots m'_k! m'_{k+1}!} \cdot p_1^{m_1} \cdot p_2^{m_2} \cdot \dots \cdot p_k^{m'_k} \cdot p_{k+1}^{m'_{k+1}}$$

§10 Вероятностные производящие функции

Производящие функции можно определить для любой числовой последовательности (имея в виду счётный набор).

Определение. Производящей функцией числовой последовательности называется сумма ряда:

$$\sum_{k=0}^{\infty} a_k u^k, \text{ где } a_0, \dots, a_k \text{ — числовой ряд, } u \in \mathbb{R}$$

если такой ряд сходится.

$$\{p_i\}_{i=0,1,\dots} : p_i \geq 0, \forall i : \sum_{i=0}^{\infty} p_i = 1$$

Производящая функция:

$$G(u) = \sum_{i=0}^{\infty} p_i u^i \quad (1)$$

$$p_i = \frac{1}{i!} \cdot \frac{d^i G(u)}{du^i} \Big|_{u=0} \quad (2)$$

Между соотношениями закона распределения (1) и (2) устанавливают взаимно однозначное соответствие.

Пример. $B[n, p]$ — биномиальное распределение.

$$G(u) = \sum_{m=0}^n C_n^m \cdot p^m \cdot q^{n-m} \cdot u^m = (pu + q)^n.$$

Пример. $P[a]$ — закон Пуассона.

$$G(u) = \sum_{m=0}^{\infty} \frac{a^m e^{-a}}{m!} u^m = e^{au} \cdot e^{-a}.$$

Пусть имеется n испытаний, $k > 2$ исходов

$$G_n(u_1, \dots, u_k) = \sum_{\substack{u_i \in \mathbb{R} \\ m_1, \dots, m_k: \sum_{i=1}^k m_i = n \\ m_i \geq 0, m_i \in \mathbb{Z}}} P_{m_1, \dots, m_k}^{(n)} \cdot u_1^{m_1} \cdot \dots \cdot u_k^{m_k}$$

Если имеется $n_1 + n_2 = n$ испытаний, то выполняется мультипликативное свойство

$$G_{n_1+n_2}(\bar{u}) = G_{n_1}(\bar{u}) \cdot G_{n_2}(\bar{u}).$$

Пример. $B[n, p]$ — биномиальный закон распределения вероятностей.

$$G_1(u) = q + pu, G_n(u) = (q + pu)^n$$

$$G(u) = \prod_{i=1}^n (p_i u + q_i)$$

Пусть p_1, \dots, p_n — все разные, $q_i = 1 - p_i$. Как определить производящие функции для такой последовательности разных испытаний?

Глава II

Производящее пространство элементарных событий. Случайные величины и векторы

§1 Аксиомы теории вероятностей. Вероятностное пространство

Определение. Пусть $\Omega = \{\omega\}$. Набор подмножества Ω \mathcal{A} называется алгеброй, если

1. $\Omega \in \mathcal{A}$
2. $\left. \begin{array}{l} A \subset \Omega, B \subset \Omega \\ A \in \mathcal{A}, B \in \mathcal{A} \end{array} \right\} \Rightarrow A + B \in \mathcal{A}, A \cdot B \in \mathcal{A}$
3. $\left. \begin{array}{l} A \subset \Omega \\ A \in \mathcal{A} \end{array} \right\} \Rightarrow \bar{A} \in \mathcal{A}$

Определение. Набор подмножеств исходного множества $\Omega = \{\omega\}$ — \mathcal{F} называется σ -алгеброй, если он является алгеброй и:

$$A_1, A_2, \dots, A_i, \dots A_k \subset \Omega \forall k,$$
$$A_i \in \mathcal{F} \Rightarrow \sum_{i=1}^{\infty} A \in \mathcal{F}, \prod_{i=1}^{\infty} A_i \in \mathcal{F}.$$

1.1 Аксиомы Колмогорова

$\langle \Omega, \mathcal{F}, P \rangle$ — измеримое пространство.

1. **Аксиома алгебры событий.** Заданы множества элементарных событий $\Omega = \{\omega\}$. На Ω выделена σ -алгебра \mathcal{F} , её элементы — случайные события.
2. **Аксиома существования алгебры событий.** Любому случайному событию $A \in \mathcal{F}$ сопоставлено неотрицательное число, называемое вероятностью этого события, $\forall A \in \mathcal{F} : P(A) \geq 0$.
3. **Аксиома нормированности.** $P(\Omega) = 1$.
4. **Аксиома аддитивности вероятности.** Если $A, B \in \mathcal{F}, A \times B = \emptyset \Rightarrow P(A + B) = P(A) + P(B)$.

Следствие. $A_1, \dots, A_n \in \mathcal{F} : A_i \times A_j \neq \emptyset, i \neq j, i, j = 1, 2, \dots, n \Rightarrow P(\sum_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$.

5. $A_1, \dots, A_n, \dots, A_i \in \mathcal{F} : A_{ij} = \emptyset$ (попарно несовместны)
 $i \neq j, i, j = 1, 2, \dots$
 $\Rightarrow P(\sum_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

P – нормированная счётно-аддитивная мера

Рассмотрим монотонную случайную последовательность событий A_1, \dots, A_n, \dots

$$A_1 \subset A_2 \subset \dots \subset A_n \subset \dots, A_n \subset A_{n+1}^{\forall n \in \mathbb{N}} \quad (1)$$

$$A_1 \supset A_2 \supset \dots \supset A_n \supset \dots, A_n \supset A_{n+1}^{\forall n \in \mathbb{N}} \quad (2)$$

Тогда

$$A = \sum_{i=1}^{\infty} A_i = \lim_{n \rightarrow \infty} A_n \text{ — предел (1).}$$

$$A = \prod_{i=1}^{\infty} A_i = \lim_{n \rightarrow \infty} A_n \text{ — предел (2).}$$

Определение. Функция событий $Q(A)$ называется непрерывной, если для любой монотонной последовательности случайных событий выполняется равенство

$$\lim_{n \rightarrow \infty} Q(A_n) = Q(\lim_{n \rightarrow \infty} A_n)$$

5'. **Аксиома непрерывности.** Пусть A_n — монотонно убывающая последовательность случайных событий.

$[A_n \downarrow \emptyset] \Leftrightarrow A_{n+1} \subset A_n, n = 1, 2, \dots$ ($\prod_{i=1}^n A_n = \emptyset$ — предел невозможен). Тогда

$$\lim_{n \rightarrow \infty} P(A_n) = 0.$$

Теорема. Вероятность является непрерывной функцией событий.

Доказательство. Из аксиомы (5) \Rightarrow (5')

Пусть $B_n \downarrow \emptyset, A_n = B_n \cdot \overline{B_{n+1}}$, где $n = 1, 2, \dots$

$$B_1 = \sum_{k=1}^n A_k, B_n = \sum_{k=n}^{\infty} A_k,$$

$$P(B_1) = P\left(\sum_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k)$$

$$P(B_n) = P\left(\sum_{k=n}^{\infty} A_k\right) = \underbrace{\sum_{k=n}^{\infty} P(A_k)}_{n \rightarrow \infty} \rightarrow 0$$

$$\lim_{n \rightarrow \infty} P(B_n) = 0$$

Из аксиомы (5') \Rightarrow (5)

Пусть имеется множество попарно несовместных событий $\{A_n\}_{n=1,2,\dots}$.

$$B_n = \sum_{k=n+1}^{\infty} A_k, A = \sum_{k=1}^{\infty} A_k$$

$$A = A_1 + \dots + A_n + B_n$$

$$P(A) = \sum_{k=1}^n P(A_k) + \underbrace{P(B_n)}_{\rightarrow 0}$$

$$B_{n+1} \subset B_n$$

$B_n \Rightarrow A_i, i > n \Rightarrow A_{i+1}$ не наступило \Rightarrow не наступило B_i, \dots

$$\prod_{i=1}^{\infty} B_i = 0$$

□

§2 Свойства вероятности

1. Вероятность невозможного события $P(\emptyset) = 0$

Доказательство. $\emptyset + \Omega = \Omega$

$$1 = P(\Omega) = P(\emptyset + \Omega) \stackrel{(\text{акс. 4})}{=} P(\emptyset) + \underbrace{P(\Omega)}_{=1} = 1$$

□

2. $P(\overline{A}) = 1 - P(A)$

Доказательство.

$$1 = P(\Omega) = P(A + \bar{A}) \stackrel{(\text{акс. 4})}{=} P(A) + P(\bar{A}) = 1$$

□

$$3. A \subset B \Rightarrow P(A) \leq P(B)$$

Доказательство. $B = A \times B + \bar{A} \times B \stackrel{A \subset B}{=} A + \bar{A} \times B$

$$P(B) = P(A + \bar{A} \times B) \stackrel{(\text{акс. 4})}{=} P(A) + P(\bar{A} \times B)$$

□

$$4. P(A + B) = P(A) + P(B) - P(A \times B)$$

Доказательство.

$$A + B = A + \bar{A} \times B, B = A \times B + \bar{A} \times B \Rightarrow$$

$$P(A + B) \stackrel{(\text{акс. 4})}{=} P(A) + P(\bar{A} \times B)$$

$$P(B) = P(A \times B) + P(\bar{A} \times B)$$

□

$$5. P(A + B) \leq P(A) + P(B)$$

Доказательство.

$$P\left(\sum_{k=1}^n A_k\right) \leq \sum_{k=1}^n P(A_k)$$

$$P\left(\sum_{k=1}^{\infty} A_k\right) \leq \sum_{k=1}^{\infty} P(A_k)$$

□

6. A_1, \dots, A_n :

$$\begin{aligned} P\left(\sum_{k=1}^{\infty} A_k\right) &= \sum_{k=1}^{\infty} P(A_k) - \sum_{k=1}^{n-1} \sum_{i=k+1}^n P(A_k \times A_i) + \dots \\ &\dots - (-1)^{n-1} \cdot P\left(\prod_{k=1}^n A_k\right) \end{aligned}$$

Пусть $B = \sum_{k=1}^{n+1} A_k$. Тогда

$$P\left(\sum_{k=1}^{n+1} A_k\right) = P(A_{n+1} + \sum_{k=1}^n A_k)$$

7. **Теорема.** Пусть имеется k попарно несовместных и составляющих полную группу благоприятных к событию A исходов из всех исходов n , $\{E_1, \dots, E_n\}_{E_i \times E_j = \emptyset, i \neq j, i, j=0, 1, \dots}$

$\sum_{i=1}^n E_i = \Omega$. Тогда

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

Доказательство. $A = \sum_{s=1}^k E_{i_s}$

$$1 = P\left(\sum_{i=1}^n E_i\right) \stackrel{(\text{акс. 4})}{=} \sum_{i=1}^n P(E_i) = n \cdot P(E_i), P(E_i) = \frac{1}{n}$$

$$P(A) = P\left(\sum_{s=1}^k E_{i_s}\right) \stackrel{(\text{акс. 4})}{=} \sum_{s=1}^k P(E_{i_s}) = \frac{k}{n}$$

$$\Omega = \{E_1, \dots, E_n\}, \mathcal{F} = \{A = E_{i_1} + \dots + E_{i_k}\}, k \leq n$$

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

□

8. $0 \leq P(A) \leq 1, A \in \mathcal{F}$

$$0 \leq P(\bar{A}) = 1 - P(A) \Rightarrow P(A) \leq 1$$

Для условных вероятностей:

$$P(A|B) \stackrel{\text{def}}{=} \frac{P(A \times B)}{P(B)}$$

$$P(B) > 0$$

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

Аксиома нормированности выполнена.

Пусть A_1, \dots, A_n, \dots — не более чем счётный набор, $A_i \times A_j = \emptyset, i \neq j; i, j = 1, 2, \dots, n$

Пусть также имеется событие $B, P(B) > 0$.

$$\begin{aligned} P\left(\sum_n A_n | B\right) &= \frac{P\left(\left(\sum_n A_n\right) \times B\right)}{P(B)} = \frac{\sum_n P(A_n \times B)}{P(B)} = \\ &\stackrel{(\text{акс. 4,5})}{=} \sum_n P(A_n | B) \end{aligned}$$

9. $\mathcal{F}_* = \{\emptyset, \Omega\}$ — наименьшая σ -алгебра, $\mathcal{F}^* = \{A : A \subseteq \Omega\}$ — наибольшая σ -алгебра.

Пусть имеется n испытаний в эксперименте с подбрасыванием монет. Построим измеримое пространство

$$\Omega = \{O, P\}, \mathcal{F}^* : \{O\}, \{P\}, \{O + P\}, \emptyset$$

Зададим вероятности событий (O — орёл, P — решка)

$$P(O) = p, P(P) = q : p + q = 1.$$

Создано вероятностное пространство.

В случае несчётного пространства элементарных событий набор всех подмножеств элементарных событий и построенная σ -алгебра не будет удовлетворять испытанию, чтобы задать вероятность на этом множестве.

Можем основываться на определённом наборе подмножеств Ω .

Определение. Наименьшая σ -алгебра, содержащая заданный класс подмножеств Ω , называется σ -алгеброй, порождаемой этим классом.

Пример. $\Omega = [0, 1]$. Пусть σ -алгебра \mathcal{F} составляет все подмножества, для которых можно определить длину, $\{\omega\}$ — точки отрезка $[0, 1]$.

$A \in \mathcal{F} : P(A) = \mu(A)$ — длина некоторого отрезка, мера

При не более чем счётном пространстве событий мы говорим о дискретных пространствах. В противном случае — о непрерывно измеримых.

§3 Способы задания вероятностных мер на измеримом пространстве $(\mathbb{R}^1, \mathcal{B}(\mathbb{R}^1))$

$\mathbb{R}^1 = (-\infty; +\infty)$ — вещественная прямая. Рассмотрим на \mathbb{R}^1 интервал такого вида:

$$\{(a, b] = \{x \in \mathbb{R}^1 | a < x \leq b\}\}$$

$$-\infty \leq a < b < +\infty$$

Рассмотрим множество таких интервалов.

Определение. Наименьшая σ -алгебра, содержащая систему вида (*), называется борелевской алгеброй множеств вещественной прямой.

Определение. Борелевские множества – элементы $\mathcal{B}(\mathbb{R}^1)$. Они же и являются событиями.

Заметим, что интервалы

$$(a, b) = \sum_n (a, b - \frac{1}{n}]$$

Отрезки:

$$[a, b] = \prod_n (a - \frac{1}{n}, b]$$

Например, точка a :

$$\{a\} = \prod_n (a - \frac{1}{n}, a]$$

Оставшиеся на рассмотрении $(a, b]$ (и др. виды) ведут к той же борелевской алгебре.

Построим вероятностное пространство на основе этого измеримого пространства. Пусть P – вероятность, заданная на борелевском множестве. Возьмём борелевское множество:

$$B = (-\infty, x], x \in \mathbb{R}^1$$

Положим

$$F(x) = P\{(-\infty, x]\} \quad (3)$$

Так определённая функция, оказывается, обладает следующими свойствами:

1. $F(x)$ монотонно не убывает на \mathbb{R}^1 .

Доказательство. Пусть $x_1 < x_2$. Тогда соответствующее событие

$$(-\infty, x_1] \subset (-\infty, x_2] \Rightarrow P\{(-\infty, x_1]\} \leq P\{(-\infty, x_2]\}$$

Итак, функция монотонно неубывающая

$$F(x_1) \leq F(x_2), \forall x_1, x_2 \in \mathbb{R}^1, x_1 < x_2$$

□

2. Пусть $\lim_{x \rightarrow -\infty} F(x) = F(-\infty)$, $\lim_{x \rightarrow +\infty} F(x) = F(+\infty)$, тогда $F(-\infty) = 0, F(+\infty) = 1$.

Доказательство. В силу монотонности предел функции существует. Докажем, что $F(-n) \xrightarrow[n=1,2,\dots]{n \rightarrow \infty} 0$.

Рассмотрим последовательность событий:

$$\{B_n = (-\infty, -n]\}$$

$$\lim_{n \rightarrow \infty} = \prod_n \left. \begin{matrix} B_{n+1} \subset B_n \\ B_n = \emptyset \end{matrix} \right\} B_n \downarrow \emptyset$$

Для этой последовательности (по акс. непрерывности)

$$\lim_{n \rightarrow \infty} \underbrace{P(B_n)}_{F(-n)} = 0$$

□

3. $F(x)$ непрерывна справа и имеет предел слева $\forall x \in \mathbb{R}^1$

$$\lim_{y \uparrow x} F(y) = F(x-0), \lim_{y \downarrow x} F(y) = F(x+0)$$

Тогда $\forall x \in \mathbb{R}^1 : F(x+0) = F(x), F(x) - F(x-0) \geq 0$.

Доказательство. Рассмотрим монотонную последовательность $x_n \downarrow x$. Введём последовательность событий

$$\{B_n = (-\infty, x_n]\}, B = (-\infty, x]$$

Рассмотрим предел $\{B_n\}_{n=1,2,\dots}$

$$\prod_n B_n = B$$

Перейдём к вероятностям:

$$\lim_{n \rightarrow \infty} P(B_n) = P(\lim_{n \rightarrow \infty} B_n) = P(B)$$

$$P(B) = F(x) \Rightarrow \lim_{n \rightarrow \infty} F(x_n) = F(x)$$

□

Определение. Любая функция, удовлетворяющая условиям (1, 2, 3), называется функцией распределения на вещественной прямой.

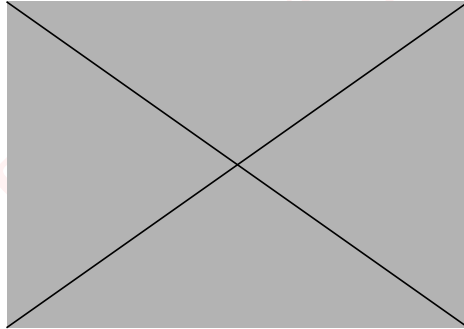


Рис. 4: Какой-то рисунок, который ещё не нарисовали.

Оказывается, что распределение вероятности имеет не более чем счётное количество скачков (разрывов функции). Функция распределения может иметь не более

- 1 скачка размером $\frac{1}{2}$,
- 2 скачков размером $\frac{1}{4}$,
- n скачков размером $\frac{1}{2^n}$.

Функция распределения непрерывна \Leftrightarrow вероятность точечного события равна нулю.

$$\forall x \in \mathbb{R}^1 : F(x) \text{ непрерывна} \Leftrightarrow P(\{x\})$$

Доказательство.

$$\begin{aligned} F(x) - F(x-0) &= \lim_{n \rightarrow \infty} (F(x) - F(x - \frac{1}{n})) = \\ &= \lim_{n \rightarrow \infty} [P(-\infty, x] - P(-\infty, x - \frac{1}{n})] \end{aligned}$$

Представим как сумму двух несовместных событий

$$\begin{aligned} &(-\infty, x - \frac{1}{n}] + (x - \frac{1}{n}, x] \\ P\{(-\infty, x]\} &= P\{(-\infty, x - \frac{1}{n}]\} + P\{(x - \frac{1}{n}, x]\} \\ \lim_{n \rightarrow \infty} P\{(x - \frac{1}{n}, x]\} &= P\{\prod_n (x - \frac{1}{n}, x]\} = P\{x\}. \end{aligned}$$

Итак, $F(x) - F(x-0) = P\{x\}$. □

$$\{(a, b]\}, -\infty \leq a < b < +\infty$$

В зависимости от выбора вида интервала свойства (1, 2, 3) изменяются подобно изменению интервала.

Теорема. Пусть $F(x)$ — некоторая функция распределения. Тогда на измеримом пространстве $(\mathbb{R}^1, \mathcal{B}(\mathbb{R}^1))$ \exists ! вероятностная мера P , такая, что

$$P : -\infty \leq a < b < +\infty$$

$$P\{(a, b]\} = F(b) - F(a) \quad (4)$$

Соотношения (3) и (4) устанавливают взаимно однозначное соответствие между функцией распределения и вероятностной мерой.

Теорема. (Каратеодори). Пусть имеется вероятностное пространство в широком смысле (Ω, \mathcal{A}, P) . Тогда существует единственная вероятностная мера Q на σ -алгебре $\mathcal{F} = \sigma(\mathcal{A})$ — порождённая \mathcal{A} , что:

$$Q(A) = P(A), A \in \mathcal{A}.$$

Следствие. Любое вероятностное пространство в широком смысле автоматически определяет вероятностное пространство.

Из этого следует, что для определения вероятности достаточно задать вероятности интервалов.

Рассмотрим алгебру, элементами которой является сумма непересекающихся интервалов.

$$\mathcal{A} : A = \sum_{k=1}^n (a_k, b_k], a_k < b_k$$

На этих множествах определим функцию множеств

$$P_0(A) = \sum_{k=1}^n [F(b_k) - F(a_k)] \text{ (вероятности интервалов)}$$

Проверим выполнение аксиом

- Акс. 2, 3, 4 выполняются, что почти очевидно.
- Акс. 5 также выполняется, см [1].

$$\sigma(\mathcal{A}) = \mathcal{B}(\mathbb{R}^1)$$

Итак, для любой функции распределения существует единственная вероятностная мера. Таким образом строим вероятностное пространство $(\mathbb{R}^1, \mathcal{B}(\mathbb{R}^1), P)$.

3.1 Меры на измеримых пространствах

3.1.1 Дискретные вероятностные меры

Пусть $F(x)$ — функция распределения $x_1 < x_2 < \dots < \dots$

$$\Delta F(x) = F(x) - F(x-0),$$

$$P(x_k) = \Delta F(x_k), k = 1, 2, \dots$$

$$\sum_k P(x_k) = 1$$

Представим иллюстрацию ситуации

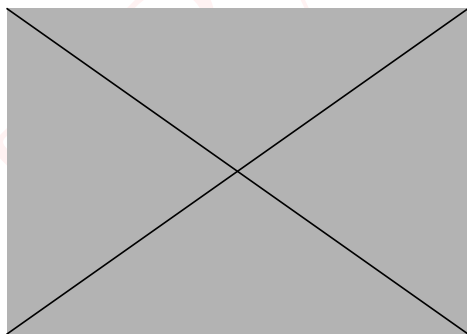


Рис. 5: Какой-то рисунок, который ещё не нарисовали.

Такой набор чисел называется дискретным законом распределения вероятностей на вещественной прямой.

Пример. $\{P_k = \frac{1}{N}\}$ — дискретный равномерный закон, $k = 1, \dots, N$

Пример. $\{p_1, p_2\} : p_1 = 1 - p_2$ — бернуллиевское дискретное распределение (часто обозначают как p, q).

Пример. $\{P_m^{(n)}\}_{m=0,1,\dots,n}$, n — число испытаний, p — вероятность появления успеха в каждом испытании.

$P_m^{(n)} = C_n^m \cdot p^m \cdot (1 - p)^{n-m}$ — биномиальное дискретное распределение.

Пример. $P_m = \frac{a^m}{m!} \cdot e^{-a}$, $a \in \mathbb{R}$, $a > 0$, $m = 0, 1, \dots$ — дискретный закон распределения Пуассона.

3.1.2 Абсолютно непрерывные вероятностные меры

Пусть $F(x)$ непрерывна $\forall x \in \mathbb{R}$, при этом существует вещественная неотрицательная кусочно-непрерывная функция плотности распределения вероятностей $f(x)$:

$$f(x) \geq 0 : F(x) = \int_{-\infty}^x f(t) dt$$

В этом случае

$$P\{(a, b]\} = \int_a^b f(x) dx$$

Очевидно, что если x — точка непрерывности $f(x)$, $x \in \mathbb{R}$, то

$$F'(x) = f(x)$$

Представим данную ситуацию

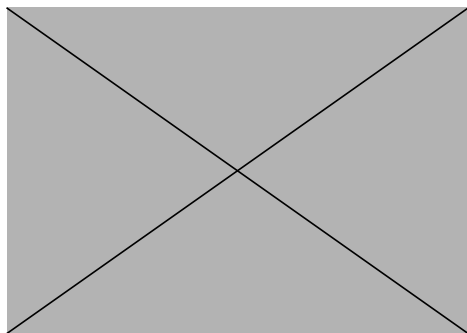


Рис. 6: Какой-то рисунок, который ещё не нарисовали.

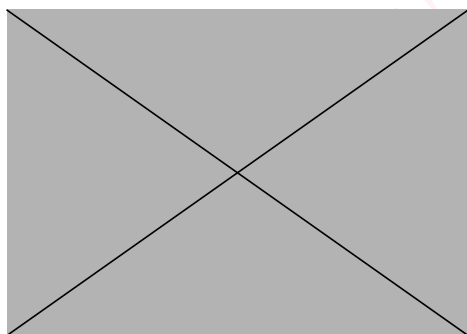


Рис. 7: Какой-то рисунок, который ещё не нарисовали.

Плотностью распределения может быть любая кусочно-непрерывная, вещественно значимая функция $f(x)$, удовлетворяющая условию нормировки

$$\int_{-\infty}^{+\infty} f(x) dx = 1$$

Пример. Равномерное распределение на отрезке $[a, b]$, $a < b$

$$f(x) = \frac{1}{b-a}, a \leq x \leq b$$

Если $x < a$ или $x > b$, тогда

$$f(x) = 0.$$

Итак,

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & x < a \text{ или } x > b \end{cases}$$

Пример. Распределение Гаусса (нормальное)

$$f(x) \sim N(a, \sigma), a \in \mathbb{R}, \sigma > 0$$

Тогда

$$f(x) = \underbrace{\frac{1}{\sqrt{2\pi\sigma}}}_{\text{норм. множитель}} \cdot e^{-\frac{(x-a)^2}{2\sigma^2}}$$

Изобразим на графике

Пример. Г-распределение

Здесь в нормированном множителе для плотности участвует Г-функция

$$f(x) = \begin{cases} 0, & x \leq 0 \\ \frac{\alpha^\lambda}{\Gamma(\lambda)} \cdot x^{\lambda-1} \cdot e^{-\alpha x}, & x > 0 \end{cases}$$

α — параметр масштаба, λ — параметр формы.

Экспоненциальное (показательное) распределение получаем при $\lambda = 1$:

$$f(x) = \begin{cases} \alpha e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Изобразим функцию плотности вероятности $f(x)$ экспоненциального распределения

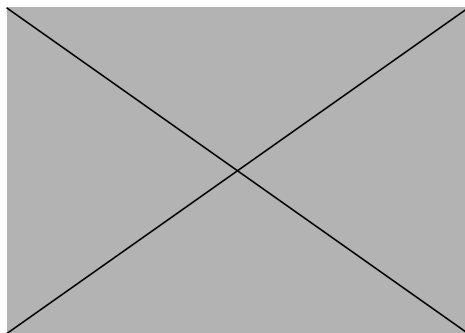


Рис. 8: Какой-то рисунок, который ещё не нарисовали.

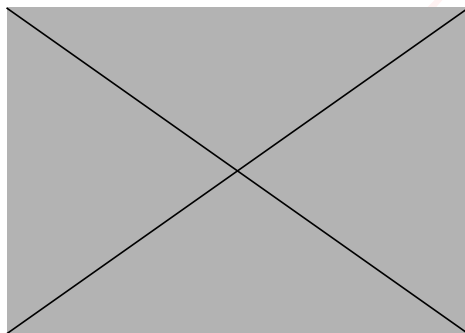


Рис. 9: Какой-то рисунок, который ещё не нарисовали.

3.1.3 Сингулярные вероятностные меры на \mathbb{R}^1

Оказывается, $F(x)$ может быть непрерывной, но не иметь плотности.

$F(x)$ — непрерывная функция распределения, функции плотности вероятности $f(x)$ не существует.

Пример. Канторова функция.

$$F(x) = \begin{cases} 0, & x \leq 0 \\ 1, & x \geq 1 \end{cases}$$

Представим, как ведёт себя функция распределения при $x \in (0, 1)$

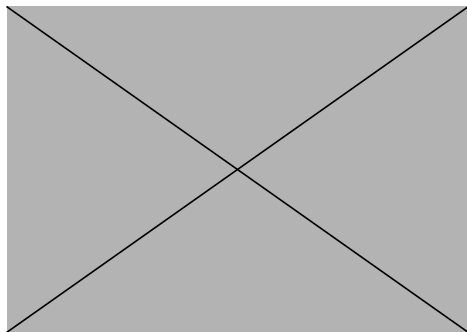


Рис. 10: $F_1(x)$ — первое приближение Канторовой функции.

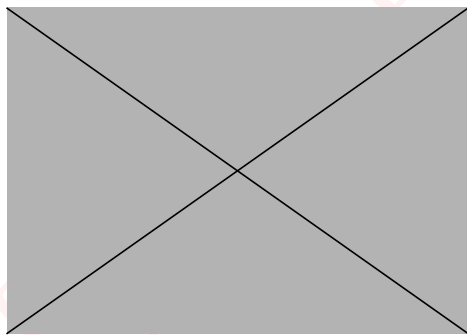


Рис. 11: $F_2(x)$ — второе приближение Канторовой функции.

$F_n(x) \xrightarrow{n \rightarrow \infty} F(x)$ — Канторова функция.

Длина интервалов, на которой Канторова функция должна быть постоянна:

$$\frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \dots = \frac{1}{3} \cdot \sum_{n=0}^{\infty} \left(\frac{2}{3}\right)^n = 1.$$

$f(x)$ почти всюду обращается в 0, за исключением, может быть, множеств меры нуль.

Функция называется сингулярной, поскольку она сингулярна по отношению к мере Лебега.

Теорема. (Лебега.) Любая функция распределение в таком измеримом пространстве может быть представлена в виде:

$$F(x) = p_1 F_1(x) + p_2 F_2(x) + p_3 F_3(x)$$

где $p_i \geq 0$, $\sum_{i=1}^3 p_i = 1$,

- $F_1(x)$ — дискретная функция распределения
- $F_2(x)$ — абсолютно непрерывная функция распределения
- $F_3(x)$ — сингулярная функция распределения

§4 Случайные величины

Определение. Вещественно-значная функция $\xi(\omega)$, определённая на измеримом пространстве (Ω, \mathcal{F}) , называется *случайной величиной*, если

$$\xi : \xi(\omega), \xi : \Omega \rightarrow \mathbb{R}^1$$

$\forall B \in \mathcal{B}(\mathbb{R}^1)$ — борелевское множество из σ -алгебр $\mathcal{B}(\mathbb{R}^1)$

$\{\omega : \xi(\omega) \in B\} \in \mathcal{F}$ — такое множество всех элементов событий, для которых $\xi(\omega) \in B$ является событием.

$$\forall B : \xi^{-1}(B) = \{\omega : \xi(\omega) \in B\} \in \mathcal{F} \quad (1)$$

Определение. Вероятностная мера \mathcal{P}_ξ на измеримом пространстве $(\mathbb{R}^1, \mathcal{B}(\mathbb{R}^1))$ такая, что:

$$\mathcal{P}_\xi(B) \stackrel{\text{def}}{=} \mathbf{P}\{\omega : \xi(\omega) \in B\}$$

называется *распределением вероятностей случайной величины* ξ на измеримом пространстве $(\mathbb{R}^1, \mathcal{B}(\mathbb{R}^1))$.

Эту вероятность далее будем обозначать как

$$\mathbf{P}\{\xi \in B\}.$$

Определение. $\forall x \in \mathbb{R}$ определим

$$F_\xi(x) = \mathbf{P}\{\omega : \xi(\omega) \leq x\} = \mathcal{P}_\xi\{(-\infty, x]\}$$

Здесь в качестве B выступает интервал $(-\infty, x]$.

$F_\xi(x)$ — функция распределения случайной величины ξ . Обозначать далее будем как

$$F_\xi(x) = \mathbf{P}\{\xi \leq x\}$$

Итак, каждая случайная величина порождает своё вероятностное пространство

$$\xi : \Omega \rightarrow \mathbb{R}^1 \rightarrow (\mathbb{R}^1, \mathcal{P}_\xi)$$

Оказывается, множество пробных борелевских множеств можно существенно сузить, так как для определения вероятности достаточно задать вероятности интервалов.

Следовательно, функция распределения случайной величины полностью определяет распределение случайной величины. $F_\xi(x)$ может восстановить меру борелевского множества. Рассмотрим типы случайных величин.

1. Если возможные значения случайной величины образуют не более чем счётное множество, то такая случайная

величина называется *дискретной* случайной величиной. ξ — дискретна, $\{x_j$ — возможное значение $\}_{j=1,2,\dots}$, $P\{\xi = x_i\} = p_i$ — вероятности.

Можно вычислить все вероятности, связанные с ξ . B — борелевское множество, $B \in \mathcal{B}(\mathbb{R}^1)$, тогда

$$P_\xi(B) = P\{\xi \in B\} = P\left(\sum_{j:x_j \in B} (\xi = x_j)\right) = \sum_{j:x_j \in B} P(\xi = x_j) = \sum_{p_j}$$

В этом случае

$$F_\xi(x) = \sum_{j:x_j \leq x} p_j \text{ — ст. функция с разрывами в } x_j$$

2. Случайная величина называется *непрерывной*, если

$$\exists f_\xi(x) \geq 0 : F_\xi(x) = \int_{-\infty}^x f_\xi(t) dt$$

3. Случайная величина называется *сингулярной*, если функция её распределения $F_\xi(x)$ непрерывна, но функция плотности вероятности $f_\xi(x)$ не существует.

Заметим, что в то время как любая случайная величина однозначно определяет свою функцию распределения, существует сколько угодно различных случайных величин, имеющих одну и ту же функцию распределения.

Пример. $x_i : -1, +1$; $p_i : \frac{1}{2}, \frac{1}{2}$; $\xi_i - \xi = \eta$

$$F_\eta = F_\xi = \begin{cases} 0, & x < -1 \\ \frac{1}{2}, & -1 \leq x < 1 \\ 1, & x \geq 1 \end{cases}$$

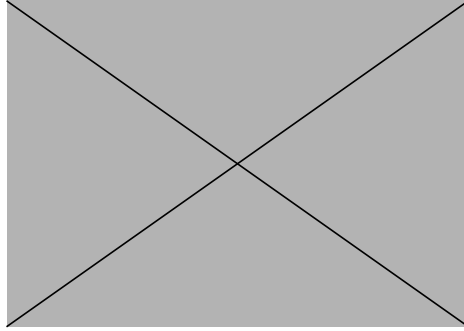


Рис. 12: Какой-то рисунок, который ещё не нарисовали.

Определение. Вещественно-значимая функция $g(x)$, $x \in \mathbb{R}$ называется *борелевской*, если для любого борелевского множества при отображении g его полный прообраз также является борелевским множеством

$$g : \mathbb{R}^1 \rightarrow \mathbb{R}^1, B \in \mathcal{B}(\mathbb{R}^1), g^{-1}(B) = \{x, x \in \mathbb{R}^1 : g(x) \in B\}$$

Теорема. Пусть $\xi(\omega)$ — случайная величина.

$$\xi(\omega) : \Omega \rightarrow \mathbb{R}^1$$

$$g(x) : \mathbb{R}^1 \rightarrow \mathbb{R}^1 \text{ — борелевская функция}$$

$\Omega = \{\omega\}$ — значение элементарного события. Тогда $\eta(\omega) = g(\xi(\omega))$ является случайной величиной.

Доказательство. $B \in \mathcal{B}(\mathbb{R}^1)$

$$\{\omega : \eta(\omega) \in B\} = \{\omega : g(\xi(\omega)) \in B\} = \{\omega : \xi(\omega) \in g^{-1}(B)\} \in \mathcal{F}$$

$$\xi \sim (\Omega, \mathcal{F})$$

Это означает, что $\xi(\omega)$ — случайная величина. □

Также возможно конструировать случайные величины как функции других величин.

4.1 Примеры случайных величин

- Пусть A — событие, $\xi(\omega) \equiv I_A(\omega) = \begin{cases} 1, \omega \in A \\ 0, \omega \notin A \end{cases}$

Задана функция-индикатор.

- Пусть испытание — бросание игральной кости. Случайная величина — число выпавших на верхней грани очков, всего возможно 6 значений, вероятность каждого — $\frac{1}{6}$. Можем задать дискретное распределение с параметром 6.
- Распределение последовательности испытаний Бернулли. Пусть $\xi_i(\omega)$ — число появления «успеха» в i -ом испытании.

$\xi_i(\omega) : I_{A_i}(\omega)$ — Бернуллиевская случайная величина

Пусть A_i — в i -ом испытании «успех».

- Рассмотрим серию из n испытаний Бернулли. μ_n — число появления «успеха» в серии из n испытаний. Пусть p — вероятность появления успеха в каждом из испытаний.

$$\{P\{\mu_n = m\} = C_n^m p^m (1-p)^{n-m}\}_{m=0, \dots, n}$$

- Пусть испытание — «бросание» точки в отрезок случайным образом. $\Omega = \{\omega\}$.

$$\begin{cases} \Omega = [a, b] & a < b \\ \mathcal{F} \end{cases}$$

$$\xi(\omega) = \omega$$

$$F_\xi(x) \stackrel{\text{def}}{=} P\{\xi \leq x\} = \frac{x-a}{b-a}, a \leq x \leq b, \xi \in [a, x]$$

$$F'(x) = \frac{1}{b-a}, x \in [a, b]$$

$$x < a : \mathcal{F}_\xi(x) = 0$$

$$x > b : \mathcal{F}_\xi(x) = 1$$

§5 Способы задания вероятности на вероятностном пространстве $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$

$\mathbb{R}^1 \times \dots \times \mathbb{R}^1 = \mathbb{R}^n$ — прямое произведение n экземпляров вещественных прямых.

$$\{x = \{x_1, \dots, x_n\}\}, \forall i : -\infty < x_i < +\infty$$

Определим пространство элементарных событий. Определим прямоугольник (объект) в \mathbb{R}^n .

$$-\infty \leq a_k < b_k < +\infty, k = 1, \dots, n$$

$$(a_1, b_1] \times \dots \times (a_k, b_k] \times \dots \times (a_n, b_n]$$

$$B_1 \times \dots \times B_n, B_i \in \mathcal{B}(\mathbb{R}^1)$$

$$\sigma\{(B_1 \times \dots \times B_n)\} = \mathcal{B}(\mathbb{R}^n)$$

Построим вероятностное пространство на этом измеримом пространстве. Пусть P — вероятностная мера на $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. Рассмотрим функцию n -мерной точки $F : \mathbb{R}^n \rightarrow \mathbb{R}^1$

$$F(x_1, \dots, x_n) = P\{(-\infty, x_1] \times \dots \times (-\infty, x_n]\} \quad (1)$$

Введём разностный оператор.

$$\begin{aligned} \Delta_{a_i b_i} F(x_1, \dots, x_n) &= F(x_1, \dots, x_{i-1}, b_i, x_{i+1}, \dots, x_n) - \\ &- F(x_1, \dots, x_{i-1}, a_i, x_{i+1}, \dots, x_n), a_i < b_i, i = 1, \dots, n. \end{aligned}$$

Пример. $\Delta_{a_1 b_1} \Delta_{a_2 b_2} \dots \Delta_{a_n b_n} F(x_1, \dots, x_n) = P\{(a_1, b_1] \times \dots \times (a_n, b_n]\}$

Отметим определяющие свойства функции F .

1. $\forall k : -\infty \leq a_k < b_k < +\infty, k = 1, \dots, n$

$$\Delta_{a_1 b_1} \cdot \dots \cdot \Delta_{a_n b_n} F(x_1, \dots, x_n) \geq 0$$

2. Из монотонности вероятности — F монотонно не убывает по любому своему аргументу

Из непрерывности вероятности — F непрерывна справа по совокупности аргументов в каждой точке.

$$x, x^{(k)} \in \mathbb{R}^n \Rightarrow F(x^{(k)}) \xrightarrow{k \rightarrow \infty}$$

3. Устремим каждый аргумент функции к $+\infty$:

$$F(+\infty, \dots, +\infty) = 1$$

4. $\lim_{x \downarrow y} F(x_1, \dots, x_n) = 0, y \in \mathbb{R}^n, y_i = -\infty$.

Определение. Любую функцию n аргументов, удовлетворяющую условиям (1, 2, 3, 4), будем называть *функцией распределения* в \mathbb{R}^n

Пример. $n = 2$

$$F(x_1, x_2) = \begin{cases} 0, x_1 + x_2 < 1 \\ 1, x_1 + x_2 \geq 1 \end{cases}$$

Очевидно, свойства (2, 3, 4) выполнены. Применим разностный оператор.

$$a_1 = 0, b_1 = 0 : \Delta_{01} F(x_1, x_2) = F(1, x_2) - F(0, x_2)$$

$$a_2 = 0, b_2 = 0 : \Delta_{01} [\Delta_{01} F(x_1, x_2)] = [F(1, 1) - F(0, 1)] - [F(1, 0) - F(0, 0)] = 1 - 1 - 1 + 0 = -1$$

Условие (1) не выполнено.

Теорема. Пусть $F = F(x_1, \dots, x_n)$ — функция распределения на \mathbb{R}^n . Тогда на вероятностном пространстве $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ существует единственная вероятностная мера P :

$$P\{(a_1, b_1] \times \dots \times (a_n, b_n]\} = \Delta_{a_1 b_1} \dots \Delta_{a_n b_n} F(x_1, \dots, x_n) \quad (2)$$

(1) и (2) говорят о том, что между вероятностью и функцией распределения в \mathbb{R}^n установлено взаимно однозначное соответствие.

§6 Случайный вектор и его распределение на вероятностном пространстве $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$

Определение. Любой упорядоченный набор случайных величин ξ_1, \dots, ξ_n , заданный на (Ω, \mathcal{F}, P) , называется случайным вектором.

$$\bar{\xi} = (\xi_1, \dots, \xi_n)$$

Если для любого $B \in \mathcal{B}(\mathbb{R}^n)$

$\{\omega : \xi(\omega) \in B\} \in \mathcal{F}$ — событие, то $\bar{\xi}$ — случайный вектор.

$$B = B_1 \times \dots \times B_n, B_i \in \mathcal{B}(\mathbb{R}^1)$$

$$\{\omega : \xi_1(\omega) \in B_1, \dots, \xi_n(\omega) \in B_n\}$$

Определение. Распределением вероятностей n -мерного случайного вектора называется функция множества $\mathcal{P}_{\bar{\xi}}(B)$ такая, что

$$B \in \mathcal{B}(\mathbb{R}^n) : \mathcal{P}_{\bar{\xi}} \stackrel{\text{def}}{=} P\{\omega : \bar{\xi}(\omega) \in B\}$$

Определение. Функцией распределения случайного вектора $\bar{\xi}$ называется функция n -мерной точки

$$F_{\bar{\xi}}(x_1, \dots, x_n) \stackrel{\text{def}}{=} P\{\omega : \xi_1(\omega) \leq x_1, \dots, \xi_n(\omega) \leq x_n\}$$

$$B \in \mathcal{B}(\mathbb{R}^n) : \mathcal{P}_{\bar{\xi}} \stackrel{\text{def}}{=} \mathbf{P}\{\omega : \bar{\xi}(\omega) \in B\} \equiv \mathbf{P}\{\bar{\xi}(\omega)\}$$

Отметим свойство согласованности функции распределения случайного вида.

1. $\lim_{x_n \rightarrow +\infty} F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = F_{\xi_1, \dots, \xi_{n-1}}(x_1, \dots, x_{n-1})$
2. $\lim_{x_n \rightarrow -\infty} F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = 0$

Пусть

$$(\cdot)(\xi_{i_1}, \dots, \xi_{i_k}), 1 \leq i_1 < \dots < i_k \leq n; \quad k < n \text{ — подвектор}$$

Лемма. Распределение любого подвектора исходного вектора (\cdot) полностью определяется распределением исходного вектора $\bar{\xi}$. $\bar{\eta} = (\xi_1, \dots, \xi_k), k < n$ порождает на $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ вероятностную меру $\mathcal{P}_{\bar{\eta}}$. $B \in \mathcal{B}(\mathbb{R}^k)$.

$$\begin{aligned} \mathcal{P}_{\bar{\eta}}(B) &\stackrel{\text{def}}{=} \mathbf{P}(\bar{\eta} \in B) = \mathbf{P}((\xi_1, \dots, \xi_k) \in B) = \\ &\mathbf{P}\{(\xi_1, \dots, \xi_k) \in B, (\xi_{k+1}, \dots, \xi_n) \in \mathbb{R}^{n-k}\} = \\ &= \mathbf{P}\{(\xi_1, \dots, \xi_n) \in B \times \mathbb{R}^{n-k} = \mathcal{P}_{\bar{\xi}}(B \times \mathbb{R}^{n-k}) \end{aligned}$$

Обратное не верно.

Случайный n -мерный вектор имеет дискретное распределение вероятностей, если он может принимать не более чем счётное число возможных значений.

$$\bar{\xi} = (\xi_1, \dots, \xi_n), x^{(j)} = (x_1^{(j)}, \dots, x_n^{(j)}), \mathbf{P}\{\bar{\xi} = x^{(j)}\} = p_j, \sum_{j=1}^n p_j = 1$$

$$\mathcal{P}_{\bar{\xi}}(B) = \mathbf{P}(\bar{\xi} \in B) = \sum_{j: x^{(j)} \in B} p_j =$$

$$\begin{aligned}
&= \sum_{j: x_1^{(j)} \in B} \dots \sum_{j: x_n^{(j)} \in B} P(\xi_1 = x_1^{(j)}, \dots, \xi_n = x_n^{(j)}) \\
&\quad F_{\bar{\xi}}(x_1, \dots, x_n) = P\{\xi_1 \leq x_1, \dots, \xi_n \leq x_n\} = \\
&= \sum_{j: x_1^{(j)} \leq x_1} \dots \sum_{j: x_n^{(j)} \leq x_n} P(\xi_1 = x_1^{(j)}, \dots, \xi_n = x_n^{(j)})
\end{aligned}$$

Будем говорить, что n -мерный случайный вектор имеет абсолютно непрерывное распределение, если существует функция плотности вероятности $f_{\bar{\xi}}(x_1, \dots, x_n) \geq 0$, $\int_{\mathbb{R}^n} f_{\bar{\xi}}(x) dx = 1$, такая, что

$$F_{\bar{\xi}}(B) \stackrel{\text{def}}{=} P(\bar{\xi} \in B) = \int_B f_{\bar{\xi}}(x) dx, B \in \mathcal{B}(\mathbb{R}^n)$$

Пример. $S \in \mathbb{R}^n$.

$$f_{\bar{\xi}}(x) = \begin{cases} \frac{1}{|S|}, & x \in S \\ 0, & x \notin S \end{cases}$$

Тогда

$$\mathcal{P}_{\bar{\xi}} = P\{\xi \in B\} = \frac{|B \cap S|}{|S|}$$

Рассмотрим положительно определённую симметричную матрицу R

$$R = \{r_{ij}\}_{i=1, \dots, n, j=1, \dots, n}, r_{ij} = r_{ji}$$

$$\forall x_i, x_j \in \mathbb{R}^1 : \sum_{i,j=1} r_{ij} \cdot x_i \cdot x_j > 0$$

$$\det R > 0 \Rightarrow \exists R^{-1} = B = \{b_{ij}\}_{i=1, \dots, n, j=1, \dots, n}$$

n -мерная невырожденная функция плотности вероятности:

$$f(x_1, \dots, x_n) = \frac{[\det B]^{\frac{1}{2}}}{2\pi^{\frac{n}{2}}} = e^{-\frac{1}{2}b_{ij}(x_i - a_i)(x_j - a_j)}, a_i, a_j \in \mathbb{R}^1$$

§7 Независимость случайных величин

Пусть имеется упорядоченный набор случайных величин ξ_1, \dots, ξ_n , заданный на вероятностном пространстве $(\Omega, \mathcal{F}, \mathbf{P})$.

Определение. Будем говорить, что ξ_1, \dots, ξ_n независимы, если для любого одномерного борелевского множества $\{B_1, \dots, B_n\}$ $B_i \in \mathcal{B}(\mathbb{R}^1)$ выполняется:

$$\mathbf{P}\{\xi_1 \in B_1, \dots, \xi_n \in B_n\} = \prod_{i=1}^n \underbrace{\mathbf{P}\{\xi_i \in B_i\}}_{\mathcal{P}_{\xi_i}(B_i)} \quad (1)$$

$$\mathcal{P}_{\xi}(B_1 \times \dots \times B_n), B_i = (-\infty, x_i]$$

$$\forall x : F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = F_{\xi_1}(x_1) \cdot \dots \cdot F_{\xi_n}(x_n) \quad (2)$$

$$\underbrace{\Delta_{a_1 b_1}(\dots \Delta_{a_n b_n}(F_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n)))}_{\mathcal{P}_{\xi}\{(a_1, b_1] \times \dots \times (a_n, b_n]\}} = \prod_{i=1}^n \underbrace{\Delta_{a_i b_i} F_{\xi_i}(x_i)}_{\mathcal{P}_{\xi_i}\{(a_i, b_i]\}} \quad (3)$$

(1), (2), (3) эквивалентны.

Теорема. Пусть имеется набор дискретных случайных величин ξ_1, \dots, ξ_n . Они независимы тогда и только тогда, когда

$$\mathbf{P}\{\xi_1 = x_1, \dots, \xi_n = x_n\} = \prod_{i=1}^n \mathbf{P}\{\xi_i = x_i\}, \exists \{x^{(j)}\}_{j=1,2,\dots}$$

Теорема. Пусть имеется набор абсолютно непрерывных случайных величин ξ_1, \dots, ξ_n . Они независимы тогда и только тогда, когда

$$f_{\xi_1, \dots, \xi_n}(x_1, \dots, x_n) = \prod_{i=1}^n f_{\xi_i}(x_i), x \in \mathbb{R}^n$$

Замечание. Пусть имеется набор случайных величин ξ_1, \dots, ξ_n , $g_1(x), \dots, g(x)$ — набор борелевских функций, $x \in \mathbb{R}^1$. Тогда $g_1(\xi_1), \dots, g_n(\xi_n)$ также независимы.

Отступление. Сконструируем интеграл Лебега. Пусть имеется измеримое пространство $[a, b]$, $g(x) = y$ — непрерывная ограниченная функция.

- A — точка верхней границы $g(x)$,
- B — точка нижней границы $g(x)$.

$$A = y_0 < y_1 < \dots < y_n = B$$

$$S_i = \{x\} : y_{i-1} < g(x) < y_i, x \in [a, b]$$

Тогда

$$\lim_{\max(y_i - y_{i-1}) \rightarrow 0} \sum_{i=1}^n y_i \nu(S_i) \stackrel{\text{def}}{=} \int_a^b g(x) dx$$

Замечание. Конструкция интеграла Лебега не зависит от того, на каком измеримом пространстве он задаётся, в отличие от интеграла Римана, не определяющемся на абстрактном пространстве (Ω, \mathcal{F}) .

В конструкции Лебега (когда Ω — отрезок вещественной прямой) точки функции группируются согласно близости значений самой функции, а не как в интеграле Римана по близости на самой оси.

Таким образом, интегралы суммы Римана будут иметь предел для не слишком разрывных функций. Интеграл Лебега имеет смысл для более широкого класса функций.

Пусть интеграл Римана существует в смысле абсолютной сходимости, значит, совпадает с интегралом Лебега.

Для любого не более чем счётного набора попарно несовместных событий B_1, \dots, B_n, \dots интеграл Лебега:

$$\int_{\sum_i B_i} = \sum_i \int_{B_i} \dots$$

Теорема. (*Фубини*). Рассмотрим специальный класс измеримых пространств с определённой на них вероятностной мерой: (Ω, \mathcal{F}, P) . Пусть $\Omega_1 = \{\omega_1\}$, $\Omega_2 = \{\omega_2\}$

$$\Omega = \Omega_1 \times \Omega_2, \mathcal{F} = \mathcal{F}_1 \times \mathcal{F}_2, P = P_1 \times P_2$$

$$A \in \mathcal{F}_1, B \in \mathcal{F}_2, A \times B \in \mathcal{F}$$

$$P(A \times B) = P_1 \times P_2(A \times B) = P_1(A) \times P_2(B)$$

Тогда, если

$$\iint_{\Omega_1 \times \Omega_2} |\xi(\omega_1, \omega_2)| d(P_1 \times P_2) < +\infty \text{ — интеграл конечен}$$

То

$$\begin{aligned} & \exists \int_{\Omega_1} \xi(\omega_1, \omega_2) dP_i, i = 1, 2, \dots \\ & \iint_{\Omega_1 \times \Omega_2} \xi(\omega_1, \omega_2) d(P_1 \times P_2) = \int_{\Omega_1} \left[\int_{\Omega_2} \xi(\omega_1, \omega_2) dP_2 \right] dP_1 = \\ & = \int_{\Omega_2} \left[\int_{\Omega_1} \xi(\omega_1, \omega_2) dP_1 \right] dP_2 \end{aligned}$$

Пример. $n = 2$.

$$(\xi, \eta), (\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \mathcal{P}_{\xi, \eta}, f_{\xi, \eta}(x, y))$$

- $\xi \sim x$ — возможное значение ξ
- $\eta \sim y$ — возможное значение η

Пусть $B \in \mathcal{B}(\mathbb{R}^2)$

$$\mathcal{P}_{\xi, \eta}(B) = \mathbb{P}\{(\xi, \eta) \in B\} = \iint_B f_{\xi, \eta}(x, y) dx dy$$

$$f_{\xi}(x) = \int_{-\infty}^{+\infty} f_{\xi, \eta}(x, y) dy, f_{\eta}(y) = \int_{-\infty}^{+\infty} f_{\xi, \eta}(x, y) dx$$

Выведем $f_{\xi}(x)$. Пусть $B_1 \in \mathcal{B}(\mathbb{R}^1)$,

$$\begin{aligned} \mathcal{P}_{\xi}(B) &= \mathbb{P}\{\xi \in B\} = \mathbb{P}\{(\xi, \eta) \in B \times \mathbb{R}^1\} = \\ &= \iint_{B \times \mathbb{R}^1} f_{\xi, \eta}(x, y) dx dy \stackrel{(\text{т. Фубини})}{=} \int_{B_1} \left[\int_{\mathbb{R}^1} f_{\xi, \eta}(x, y) dy \right] dx \end{aligned}$$

Для $f_{\eta}(x)$ вывод аналогичен.

Глава III

Числовые характеристики распределения случайных величин

§1 Математическое ожидание случайной величины

Определение. Математическим ожиданием случайной величины ξ на вероятностном пространстве (Ω, \mathcal{F}, P) будем называть интеграл Лебега:

$$M_{\xi} = \int_{\Omega} \xi(\omega) dP(\omega), \text{ при этом, если}$$

$$\int_{\Omega} |\xi(\omega)| dP < +\infty$$

Т.е. если интеграл конечен. Иначе случайная величина математического ожидания не имеет.

$$\exists M_{\xi} \Leftrightarrow M_{|\xi|} < +\infty$$

$$\exists M_{\xi} \Rightarrow \forall A \in \mathcal{F} \exists \int_A \xi(\omega) dP = \int_{\Omega} \xi \cdot I_A dP = M[\xi \cdot I_A]$$

$$MI_A = \int_{\Omega} I_A(\omega) dP = \int_A dP = P(A)$$

1.1 Свойства математического ожидания

1. Пусть a, b — константы. Тогда $M_a = a$, $M[b \cdot \xi] = b \cdot M_\xi$
2. $|M_\xi| \leq M_{|\xi|}$
3. $M[\xi + \eta] = M_\xi + M_\eta$
4. $\forall \varepsilon > 0 : |\xi| \leq \varepsilon \Rightarrow |M_\xi| \leq \varepsilon$

Доказательство. (Свойство 4)

$$|M_\xi| \leq M_{|\xi|} = \int_{\Omega} |\xi| dP \leq \varepsilon \int_{\Omega} dP = \varepsilon \cdot P(\Omega) = \varepsilon$$

□

5. **Неравенство Чебышёва.** Пусть $\xi \geq 0$ — случайная величина, $t > 0$, $t \in \mathbb{R}^1$. Тогда

$$P\{\xi \geq t\} \leq \frac{M\xi}{t}$$

Рассмотрим

$$P\{\xi \geq t\} = \int_{\{\omega: \xi(\omega) \geq t\}} dP \leq \int_{\{\xi \geq t\}} \frac{\xi(\omega)}{t} dP \leq \frac{1}{t} \int_{\Omega} \xi(\Omega) dP = \frac{M\xi}{t}$$

Следствие. $g(x)$ — борелевская функция, $x \in (0, +\infty)$. $g(x)$ монотонно возрастает, $g(x) \geq 0$, ξ — случайная величина. $t > 0 : g(t) \neq 0$

$$P\{|\xi| \geq t\} \leq \frac{M[g(|\xi|)]}{g(t)}$$

$$P\{|\xi| \geq t\} = P\{g(|\xi|) \geq g(t)\} \leq \frac{Mg(|\xi|)}{g(t)}$$

Пусть $g(x) = x^2$. $x \in (0, +\infty)$:

$$P\{|\xi| \leq t\} \leq \frac{M\xi^2}{t^2}, t > 0$$

$$g(x_1, \dots, x_n), g: \mathbb{R}^n \rightarrow \mathbb{R}^1$$

$$\forall B \in \mathcal{B}(\mathbb{R}^1) : g^{-1}(B) \in \mathcal{B}(\mathbb{R}^n)$$

$$g^{-1}(B) = \{x, x \in \mathbb{R}^n : g(x) \in B\}$$

$$\bar{\xi} = (\xi_1, \dots, \xi_n), g(\bar{\xi}) : \Omega \rightarrow \mathbb{R}^1, \xi_i : \Omega \rightarrow \mathbb{R}^n$$

Теорема. (О вычислении математического ожидания случайной величины.) Пусть задан вектор случайных величин $\bar{\xi} = (\xi_1, \dots, \xi_n)$ на вероятностном пространстве (Ω, \mathcal{F}, P) .

$g(x) : \mathbb{R}^n \rightarrow \mathbb{R}^1$, $x \in \mathbb{R}^n$, g — борелевская функция. Тогда на конкретном вероятностном пространстве задан интеграл Лебега

$$Mg(\bar{\xi}) = \int_{\mathbb{R}^n} g(x) dP_{\bar{\xi}}(x), x \in \mathbb{R}^n$$

По определению математического ожидания задан также интеграл Лебега, но на абстрактном пространстве этих событий.

$$\int_{\Omega} g(\bar{\xi}(\omega)) dP(\omega)$$

Мы понимаем равенство следующим образом: если первый интеграл существует, то второй интеграл так же существует и ему равен.

Следствие. Пусть задан дискретный вектор случайных величин $\bar{\xi} = (\xi_1, \dots, \xi_n)$, возможные значения — $\{x^{(j)}\}_{j=1,2,\dots}$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^1$ — борелевская функция. Тогда математическое ожидание

$$Mg(\bar{\xi}) = \sum_{x^{(j)} \in \{x^{(i)}\}_{i=1,2,\dots}} g(x^{(j)}) \cdot P\{\bar{\xi} = x^{(j)}\}$$

$$M_{\xi} = \sum_j x_j \cdot P\{\xi = x_j\}$$

Следствие. Пусть задан абсолютно непрерывный вектор случайных величин $\bar{\xi} = (\xi_1, \dots, \xi_n)$, $f_{\bar{\xi}}(x_1, \dots, x_n)$, $x \in \mathbb{R}^n$, g — борелевская функция. Определим:

$$Mg(\bar{\xi}) = \int_{\mathbb{R}^n} g(x) f_{\bar{\xi}}(x) dx$$

$$M_{\xi} = \int_{-\infty}^{+\infty} x \cdot f_{\xi}(x) dx$$

Пример. *Равномерное распределение.* $\xi \sim U[a, b]$.

$$f_{\xi}(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b] \\ 0, & x \notin [a, b] \end{cases}$$

$$M_{\xi} = \int_a^b \frac{x}{b-a} dx = \frac{a+b}{2}$$

Пример. *(Дискретная случайная величина (бернуллиевская), задающая некоторое распределение).*

$$p + q = 1, M_{\xi} = p$$

Рассмотрим биномиальную случайную величину. Пусть μ_n — число появления «успеха» в n испытаниях Бернулли. Также пусть ξ_k — число появления успеха в k -ом испытании Бернулли

$$\mu_n = \sum_{k=1}^n \xi_k$$

$$M_{\mu_n} = M \left[\sum_{k=1}^n \right] = \sum_{k=1}^n M_{\xi_k} = np$$

Пример. (Распределение вероятностей случайных величин Пуассона). $\xi \sim P[a], a > 0$

$$M_{\xi} = \sum_{k=0}^{\infty} k \cdot \frac{a^k}{k!} e^{-a} = a \cdot e^{-a} \sum_{k=0}^{\infty} \frac{a^{k-1}}{(k-1)!} = a$$

Пример. (Распределение вероятностей случайных величин Коши). $a > 0$

$$f_{\xi}(x) = \frac{a}{\pi(x^2 + a^2)}, x \in \mathbb{R}^1$$

$$F_{\xi}(x) = \frac{1}{2} + \frac{1}{\pi} \operatorname{arctg} \frac{x}{a}$$

Закон распределения Коши не имеет моментов математического ожидания. Рассмотрим $M_{|\xi|}$

$$M_{|\xi|} = \int_{-\infty}^{+\infty} |x| \cdot f_{\xi}(x) dx$$

Интеграл расходится. Данное распределение не имеет математического ожидания.

§2 Дисперсия случайной величины

Пусть задана случайная величина ξ на вероятностном пространстве (Ω, \mathcal{F}, P) . Определим дисперсию (центральный момент) случайной величины:

$$D_{\xi} = M[(\xi - M_{\xi})^2] = \int_{-\infty}^{+\infty} (x - M_{\xi})^2 dP_{\xi}$$

$$M[\xi^2 - 2\xi \cdot M_{\xi} + (M_{\xi})^2] = M_{\xi^2} - 2M_{\xi} \cdot M_{\xi} + (M_{\xi})^2 = M_{\xi^2} - (M_{\xi})^2$$

$$D_{\xi} = \min_{a \in \mathbb{R}^1} M[(\xi - a)^2]$$

Минимум достигается в $a = M_{\xi}$.

Также определим среднеквадратическое отклонение (СКО)
 $— \sqrt{D_{\xi}}$

2.1 Свойства дисперсии

1. Пусть a, b — константы, ξ — случайная величина

$$D[a\xi + b] = a^2 D_{\xi}$$

Доказательство.

$$M[((a\xi + b) - M[a\xi + b])^2] = a^2 M[(\xi - M_{\xi})^2] = a^2 D_{\xi}$$

□

2. $g(x) = x^2, x > 0, g(|\eta|), \eta = \xi - M_{\xi}, \varepsilon > 0 :$

$$P\{|\xi - M_{\xi}| \geq \varepsilon\} \leq \frac{D_{\xi}}{\varepsilon^2} — \text{классическое нер-во Чебышёва}$$

Теорема. Пусть ξ_1, \dots, ξ_n — независимые случайные величины.

$$\forall i : \exists M_{\xi_i} : M[\xi_1 \cdot \dots \cdot \xi_n] = M_{\xi_1} \cdot \dots \cdot M_{\xi_n}$$

Обратное не верно.

ξ, η — независимые случайные величины. $\xi \sim x, \eta \sim y, \mathcal{P}_{\xi, \eta} = \mathcal{P}_{\xi} \times \mathcal{P}_{\eta}$

$$M[\xi, \eta] = \iint_{\mathbb{R}^1 \times \mathbb{R}^1} xy d(\mathcal{P}_{\xi, \eta}) \stackrel{\text{по т. Фубини}}{=} \int_{\mathbb{R}^1} x d\mathcal{P}_{\xi} \int_{\mathbb{R}^1} y d\mathcal{P}_{\eta} = M_{\xi} \cdot M_{\eta}$$

Пример. Приведём пример, иллюстрирующий неверность обратного. Пусть ξ, θ — независимые случайные величины, $M_\xi = 0, M_\theta = 0, \eta = \xi \cdot \theta$

$$M[\xi \cdot \eta] = M[\xi^2 \cdot \theta] \stackrel{(\text{по } T.1)}{=} M_{\xi^2} \cdot M_\theta = 0, M_\xi \cdot M_\eta = 0$$

Теорема. Пусть ξ_1, \dots, ξ_n — независимые случайные величины, имеющие дисперсию $\forall i : D_{\xi_i}$

$$D \left[\sum_{i=1}^n \xi_i \right] = \sum_{i=1}^n D_{\xi_i}$$

Доказательство. Пусть ξ, η — независимые.

$$\begin{aligned} D[\xi + \eta] &= M[(\xi + \eta - M[\xi + \eta])^2] = M[((\xi - M_\xi) + (\eta - M_\eta))^2] = \\ &= M[(\xi - M_\xi)^2] + M[(\eta - M_\eta)^2] + 2M[(\xi - M_\xi)(\eta - M_\eta)] \\ M_{\xi\eta} - M_\xi \cdot M_\eta + M_\xi \cdot M_\eta - M_\xi \cdot M_\eta &= 0 \end{aligned}$$

□

Пример. Пусть μ_n — число успеха в n испытаний Бернулли,

$$\mu_n = \sum_{k=1}^n \xi_k$$

$$D_{\xi_k} = M_{\xi_k^2} - (M_{\xi_k})^2$$

Распределение ξ_k^2 совпадает с ξ_k . Тогда

$$D_{\xi_k} = p - p^2 = p \cdot q$$

$$D_{\mu_n} = \sum_{k=1}^n D_{\xi_k} = n \cdot p \cdot q$$

§3 Матрица ковариаций случайного вектора

Определение. Пусть задан вектор случайных величин $\bar{\xi} = (\xi_1, \dots, \xi_n)$. Матрицей ковариаций $\bar{\xi}$ называется матрица, элементами которой являются

$$\underbrace{M[(\xi_i - M_{\xi_i})(\xi_j - M_{\xi_j})]_{i,j=1,\dots,n}}_{cov(\xi_i, \xi_j)} = \sigma_{ij}$$

То есть матрица ковариаций:

$$R = \{\sigma_{ij}\}_{i,j=1,\dots,n}$$

Определение. Корреляционным моментом называют величину $k_{ij} = cov(\xi_i, \xi_j)$

Если случайные величины ξ_i, ξ_j независимы, то

$$\sigma_{ij} = 0$$

Обратное не верно.

Определение. Если $\sigma_{ij} = 0$, то ξ_i, ξ_j называют некоррелированными случайными величинами

Очевидно, что:

- $cov(\xi_i, \xi_j) = cov(\xi_j, \xi_i)$,
- $cov(\xi_i, \xi_j) = D_{\xi_i}$,
- $cov(a\xi_i, \xi_j) = a \cdot cov(\xi_i, \xi_j)$, где a — константа.

Также

$$D[\xi + \eta] = D_\xi + D_\eta + 2cov(\xi, \eta)$$

Можно показать (по индукции), что

$$D \left[\sum_{i=1}^n \xi_i \right] = \sum_{i=1}^n D_{\xi_i} + 2 \sum_{i < j}^n cov(\xi_i, \xi_j)$$

Теорема. Пусть заданы

- вектор случайных величин $\bar{\xi} = (\xi_1, \dots, \xi_n)$,
- матрица C , элементами которой являются константы c :
 $C = \{c_{ij}\}_{m \times n}$, вектор $\bar{a} = \{a_i\}_{i=1, \dots, m}$,
- вектор $\bar{b} = \{b_j\}_{j=1, \dots, m}$.

Рассмотрим вектор $\bar{\eta}$:

$$\bar{\eta} = C\bar{\xi} + \bar{a}$$

Тогда матрица ковариаций вектора $\bar{\eta}$:

$$R_{\bar{\eta}} = C \cdot R_{\bar{\xi}} \cdot C^T$$

$$R_{\bar{\xi} + \bar{b}} = R_{\bar{\xi}}$$

Лемма. Пусть у вектора случайных величин $\bar{\xi} = (\xi_1, \dots, \xi_n)$ существуют все ковариации $cov(\xi_i, \xi_j)$. Тогда при любом наборе вещественных констант $\{c_i\}_{i=1, \dots, n}$

$$D[c_1\xi_1 + \dots + c_n\xi_n] = \sum_{i,j=1}^n \sigma_{ij}c_ic_j$$

Доказательство. Пусть $\eta_n = c_1\xi_1 + \dots + c_n\xi_n$

$$\eta_n - M_{\eta_n} = \sum_{i=1}^n c_i(\xi_i - M_{\xi_i})$$

$$(\eta_n - M_{\eta_n})^2 = \sum_{i,j=1}^n c_i c_j (\xi_i - M_{\xi_i})(\xi_j - M_{\xi_j})$$

Применим к обеим частям математическое ожидание:

$$D_{\eta_n} = \sum_{i,j=1}^n \sigma_{ij} c_i c_j \geq 0$$

Тогда

$$\det R \geq 0$$

□

§4 Коэффициент корреляции случайной величины

Рассмотрим ситуацию, когда в векторе случайных величин содержится два элемента: $\bar{\xi} = (\xi_1, \xi_2)$. R — матрица ковариаций.

$$\det R = \begin{vmatrix} D_{\xi_1} & cov(\xi_1, \xi_2) \\ cov(\xi_2, \xi_1) & D_{\xi_2} \end{vmatrix}$$

Отсюда следует, что

$$|cov(\xi_1, \xi_2)| \leq \sqrt{D_{\xi_1} \cdot D_{\xi_2}}$$

Введём числовую характеристику — степени зависимости двух случайных величин. Назовём её *коэффициентом корреляции* двух случайных величин:

$$\rho(\xi_1, \xi_2) = \frac{cov(\xi_1, \xi_2)}{\sqrt{D_{\xi_1} \cdot D_{\xi_2}}}, \text{ где } D_{\xi_1} \neq 0 \text{ и } D_{\xi_2} \neq 0$$

Отсюда следует, что $|\rho(\xi_1, \xi_2)| \leq 1$, т.е. ρ — безразмерная числовая характеристика.

4.1 Свойства коэффициента корреляции

1. ρ — безразмерная числовая характеристика — $|\rho(\xi_1, \xi_2)| \leq 1$
2. Пусть заданы независимые случайные величины ξ, η . Тогда $\rho(\xi_1, \xi_2) = 0$. Обратное в общем случае не верно.
Но! Если ξ, η — нормальные случайные величины то $\rho(\xi_1, \xi_2) = 0 \Rightarrow \xi, \eta$ — независимые.
3. ξ, η линейно независимы $\Leftrightarrow |\rho(\xi_1, \xi_2)| = 1$.

Доказательство. (Свойство [3] — необходимость) Пусть a, b — константы, $a \neq 0$.

$$\eta = a\xi + b, M_\eta = aM_\xi + b, D_\eta = a^2 D_\xi$$

$$\text{cov}(\xi, \eta) = M[(\xi - M_\xi)a(\xi - M_\xi)] = a D_\xi$$

$$\rho(\xi, \eta) = \frac{a D_\xi}{\sqrt{D_\xi a^2 D_\xi}} = \frac{a}{|a|} = \text{sign } a$$

□

Отступление. Для любой случайной величины, имеющей конечную ненулевую дисперсию, можно рассмотреть такую случайную величину:

$$\xi_1 = \frac{\xi - M_\xi}{\sqrt{D_\xi}}$$

Если

- $M_{\xi_1} = 0$, то ξ_1 — *центрированная* случайная величина

- $D_{\xi_1} = 1$, то ξ_1 — *нормированная* случайная величина
- $M_{\xi_1} = 0$ и $D_{\xi_1} = 1$, то ξ_1 — *стандартизированная* случайная величина

То есть для любой случайной величины ξ можно сопоставить ξ_1 — стандартизованную случайную величину. Тогда

$$\eta : \eta_1 = \frac{\eta - M_\eta}{\sqrt{D_\eta}}$$

$$\rho(\xi, \eta) = M_{\xi_1 \cdot \eta_1},$$

где ξ_1, η_1 — стандартизованные случайные величины.

Рассмотрим такую дисперсию стандартизованных случайных величин:

$$0 \leq D[\xi_1 \pm \eta_1] = M[(\xi_1 \pm \eta_1)^2] = 2 \pm 2\rho(\xi, \eta)$$

$$M_{\xi_1^2} = D_{\xi_1} = 1, M_{\eta_1^2} = D_{\eta_1} = 1$$

Доказательство. (Свойство [3] — достаточность). Рассмотрим $\rho(\xi, \eta) = 1$. Тогда

$$D[\xi_1 - \eta_1] = 2(1 - \rho(\xi, \eta)) = 0 \Rightarrow P\{\xi_1 - \eta_1 = 0\} = 1 \Rightarrow \text{линейно нез.}$$

Рассмотрим $\rho(\xi, \eta) = -1$. Тогда

$$D[\xi_1 + \eta_1] = 2(1 + \rho(\xi, \eta)) = 0 \Rightarrow P\{\xi_1 + \eta_1 = 0\} = 1 \Rightarrow \text{линейно зав.}$$

□

Определение. Две случайные величины называют *положительно коррелированными*, если их коэффициент корреляции положительный.

Так же и с отрицательным коэффициентом корреляции — *отрицательно коррелированными*.

Если коэффициент корреляции случайных величин равен нулю, то такие случайные величины не коррелированы.

§5 Моменты случайных величин произвольных порядков

Кроме коэффициента корреляции существует немало других характеристик, например, корреляционное отношение.

Определение. Моментом (или начальным моментом) порядка $p > 0$ случайной величины ξ на вероятностном пространстве (Ω, \mathcal{F}, P) назовём M_{ξ^p} , при этом предполагается, что существует $M_{|\xi|^p} < +\infty$. По теореме о вычислении математического ожидания:

$$M_{\xi^p} = \int_{-\infty}^{+\infty} x^p dP_{\xi}(x)$$

$$\xi : \Omega \rightarrow \mathbb{R}^1 \quad (\mathbb{R}^1, \mathcal{B}(\mathbb{R}^1), P_{\xi})$$

Определение. Центральным моментом порядка $p > 0$ случайной величины ξ на вероятностном пространстве (Ω, \mathcal{F}, P) назовём

$$M[(\xi - M_{\xi})^p]$$

При этом предполагается, что $M|(\xi - M_{\xi})|^p < +\infty$.

5.1 Свойства моментов случайных величин

1. Если существует абсолютный момент порядка p случайной величины, то существуют все моменты порядка $0 < q < p$, т.е.

$$\forall 0 < q < p : \exists M_{|\xi|^p} \Rightarrow \exists M_{|\xi|^q}$$

Свойство (*): $\xi \leq \eta \Rightarrow M_{\xi} \leq M_{\eta}$ Покажем это:

$$|x|^q \leq 1 + |x|^p$$

$$|\xi|^q \leq 1 + |\xi|^p$$

Тогда

$$M_{|\xi|^q} \leq 1 + M_{|\xi|^p} < +\infty$$

2. **Неравенство Гёльдера.** Пусть $p > 1$, $q > 1$,

$$\frac{1}{p} + \frac{1}{q} = 1$$

Также пусть существует $M_{|\xi|^p} < +\infty$, $M_{|\eta|^q} < +\infty$. Отсюда существует $M_{|\xi \cdot \eta|^p} < +\infty$. Тогда

$$M_{|\xi \cdot \eta|} \leq [M_{|\xi|^p}]^{\frac{1}{p}} \cdot [M_{|\eta|^q}]^{\frac{1}{q}}$$

3. **Неравенство Йенсена.** Пусть существует M_ξ , $g(x)$ — выпуклая вниз борелевская функция. Тогда

$$g(M_\xi) \leq M g(\xi)$$

Доказательство. Так как $g(x)$ выпукла вниз, то

$$\forall y \in \mathbb{R}^1 \exists \lambda(y) : \forall x \in \mathbb{R}^1 g(x) \geq g(y) + g(x - y) \cdot \lambda(y)$$

Положим $x = \xi$, $y = M_\xi$. Тогда

$$g(\xi) \geq g(M_\xi) + (\xi - M_\xi) \cdot \lambda(M_\xi)$$

Воспользуемся свойством $\ast(1)$:

$$M g(\xi) \geq g(M_\xi)$$

□

4. **Неравенство Ляпунова.** Пусть задана случайная величина ξ , имеющая математическое ожидание M_ξ , а также $p > q > 0$. Тогда

$$\left[\underbrace{M_{|\xi|^p}} \right]^{\frac{1}{p}} \geq [M_{|\xi|^q}]^{\frac{1}{q}}$$

Доказательство. Пусть $\frac{p}{q} = r > 1$. Рассмотрим борелевскую функцию $g(x) = |x|^r$. $r > 0$, значит, $g(x)$ выпуклая вниз. Обозначим $\eta = |\xi|^q$. Тогда запишем неравенство Йенсена:

$$[M_\eta]^r \leq M_{|\eta|^r}, \quad [M_{|\xi|^q}]^{\frac{p}{q}} \leq M_{|\xi|^p}$$

Возведём в $\frac{1}{p}$:

$$[M_{|\xi|^q}]^{\frac{1}{q}} \leq [M_{|\xi|^p}]^{\frac{1}{p}}$$

□

Определение. Пусть $k_1, \dots, k_n \geq 0$ — числа, а ξ_1, \dots, ξ_n — упорядоченный набор случайных величин. Тогда

1. *Смешанным* моментом порядка $(k_1 + \dots + k_n)$ назовём

$$M[\xi_1^{k_1} \cdot \xi_2^{k_2} \cdot \dots \cdot \xi_n^{k_n}]$$

2. *Центральным смешанным* моментом порядка $(k_1 + \dots + k_n)$ назовём

$$M[(\xi_1 - M_{\xi_1})_1^k \cdot \dots \cdot (\xi_n - M_{\xi_n})^{k_n}]$$

Если ξ_1, \dots, ξ_n независимы, то 1 и 2 представляются произведением соответствующих моментов.

§6 Законы распределения функций случайных величин

Пусть определена случайная величина ξ , функция распределения этой случайной величины $F_\xi(x)$, а также, если ξ абсолютно непрерывна, то функция плотности вероятности $f_\xi(x)$. Рассмотрим борелевскую функцию одного аргумента:

$$g(\xi) = \eta$$

Тогда

$$\begin{aligned} F_\eta(y) &= \mathbf{P}\{\eta \leq y\} = \mathbf{P}\{g(\xi) \leq y\} = \mathbf{P}\{\xi \in g^{-1}\{(-\infty, y]\}\} = \\ &= \int_{g^{-1}\{(-\infty, y]\}} d\mathcal{P}_\xi \equiv \dots \end{aligned}$$

Замечание. Вероятностная мера \mathcal{P}_ξ однозначно восстанавливается по функции распределения F_ξ . Поэтому часто интеграл Лебега обозначают

$$\begin{aligned} \int_{\mathbb{R}^1} g(x) d\mathcal{P}_\xi &= \int_{\mathbb{R}^1} g(x) dF_\xi \text{ — интеграл Лебега-Стилтьеса} \\ \dots &\equiv \int_{g^{-1}\{(-\infty, y]\}} dF_\xi(x) \end{aligned}$$

Рассмотрим

- $\eta = a\xi + b$, где a, b — константы

— $a > 0$:

$$F_\eta(y) = \mathbf{P}\{a\xi + b \leq y\} = \mathbf{P}\left\{\xi \leq \frac{y-b}{a}\right\} = F_\xi\left(\frac{y-b}{a}\right)$$

— $a < 0$:

$$\begin{aligned} F_\eta(y) &= \mathbf{P}\{a\xi + b \geq y\} = \mathbf{P}\left\{\xi \geq \frac{y-b}{a}\right\} = \\ &= 1 - F_\xi\left(\frac{y-b}{a}\right) + \mathbf{P}\left\{\xi = \frac{y-b}{a}\right\} \end{aligned}$$

- $\eta = \xi^2$

$$\begin{aligned} F_\eta(y) &= \mathbf{P}\{\xi^2 \leq y\} = \mathbf{P}\{-\sqrt{y} \leq \xi \leq \sqrt{y}\} = \\ &= F_\xi(\sqrt{y}) - F_\xi(-\sqrt{y}) + \mathbf{P}\{\xi = -\sqrt{y}\} \end{aligned}$$

Пусть задана абсолютно непрерывная случайная величина ξ , функция распределения этой случайной величины $F_\xi(x)$ и функция плотности вероятности $f_\xi(x)$.

$J = (a, b) \subset \mathbb{R}^1$, $g(x)$ — непрерывная дифференцируемая функция, строго монотонна. $x \in J$, $g'(x) \neq 0$.

$$x \in J, g'(x) = 0 \Rightarrow \exists g^{-1}(y), y \in g(J)$$

Рассмотрим случайную величину $g(\xi)$, пусть $g(\xi)$ строго возрастает.

$$F_{g(\xi)}(y) = P\{g(\xi) \leq y\} = P\{\xi \leq g^{-1}(y)\} = F_\xi(g^{-1}(y))$$

$$\xi \sim \{x\}, g(\xi) \sim \{y\}$$

Если существует функция плотности вероятности, то

$$\begin{aligned} F_{g(\xi)}(y) &= F_\xi(\underbrace{g^{-1}(y)}_{h(y)}) = \int_{-\infty}^{h(y)} f_\xi(x) dx = [x = h(z), z = g(x)] = \\ &= \int_{-\infty}^y \underbrace{f_\xi(h(z))h'(z)}_{f_{g(\xi)}(z)} dz \end{aligned}$$

Пусть $g(x)$ строго убывает

$$\begin{aligned} F_{g(\xi)}(y) &= P\{\xi \geq g^{-1}(y)\} = 1 - F_\xi(\underbrace{g^{-1}(y)}_{h(y)}) = \\ &= \int_{h(y)}^{+\infty} f_\xi(x) dx = [x = h(z), z = g(x)] = \int_y^{-\infty} f_\xi(h(z))h'(z) dz = \\ &= - \int_{-\infty}^y f_\xi(h(z))h'(z) dz \end{aligned}$$

Если $g(x)$ строго монотонна:

$$f_{g(\xi)}(y) = f_{\xi}(y^{-1}(y)) \cdot [|g^{-1}(y)|]$$

Рассмотрим область значений

$$\xi : \sum_{k=1}^n [a_k, b_k], \quad J_k = (a_k, b_k), \quad g(x), \quad x \in J_k$$

$g(x)$ либо (строго) монотонно возрастает, либо (строго) монотонно убывает, непрерывно дифференцируема на каждом J_k , $g'(x) \neq 0$.

Введём обозначение:

$$h_k(y) = g^{-1}(y), \quad y \in g(J_k)$$

Тогда обобщённая формула имеет вид:

$$f_{g(\xi)}(y) = \sum_{k=1}^n f_{\xi}(h_k(y)) \cdot |h'_k(y)| \cdot I_{g(y_k)}(y)$$

Пример. Пусть $\eta = g(\xi)$.

1. $\eta = \xi^2$, ξ — абсолютно непрерывная случайная величина.

$$y \leq 0 : f_{\eta^2}(y) = 0, \quad y > 0 : g(x) = x^2$$

$J_1 = (-\infty, 0)$ — $g(x)$ строго монотонно убывает

$J_1 = (0, +\infty)$ — $g(x)$ строго монотонно возрастает

$$y = x^2, \quad x \in \mathbb{R}^1$$

$$x = \begin{cases} -\sqrt{y}, & x < 0 \\ \sqrt{y}, & x \geq 0 \end{cases}$$

Вычислим $h'(y)$

$$h'(y) = \pm \frac{1}{2\sqrt{y}}, \quad |h'(y)| = \left| \frac{1}{2\sqrt{y}} \right|$$

Таким образом

$$f_{\xi^2}(y) = \frac{1}{2\sqrt{y}} \cdot [f_{\xi}(-\sqrt{y}) + f_{\xi}(\sqrt{y})], \quad y > 0$$

Рассмотрим:

$$\xi \sim N(0, 1), \quad f_{\xi}(x) = \sqrt{12\pi} e^{-\frac{x^2}{2}}$$

$$f_{\xi^2}(y) = \frac{1}{\sqrt{2\pi y}} e^{-\frac{y}{2}}, \quad y > 0$$

2. Пусть $\eta = |\xi|$, $g(x) = |x|$, $y = |x|$,

$$x = \begin{cases} -y, & x < 0 \\ y, & x > 0 \end{cases}$$

$$f_{|\xi|}(y) = f_{\xi}(-y) + f_{\xi}(y), \quad y \geq 0$$

3. Пусть $\eta = \sqrt{|\xi|}$

$$f_{\sqrt{|\xi|}}(y) = 2y \cdot [f_{\xi}(y^2) + f_{\xi}(-y^2)], \quad y > 0$$

4. Рассмотрим пример другого вида. Пусть имеется строго возрастающая функция распределения $F_{\xi}(x)$. ξ — случайная величина. Каков закон распределения?

$$\begin{aligned} F_{F_{\xi}(\xi)}(x) &= P\{F_{\xi}(\xi) \leq x\} = P\{\xi \leq F_{\xi}^{-1}(x)\} = \\ &= F_{\xi}(F_{\xi}^{-1}(x)) = x, \quad 0 \leq x \leq 1 \end{aligned}$$

$$P\{F_{\xi}(\xi) \leq x\} = \begin{cases} 0, & x < 0 \\ x, & 0 \leq x \leq 1 \\ 1, & x > 1 \end{cases} \stackrel{\text{def}}{=} F_{F_{\xi}(\xi)}(x)$$

$$F_{\xi} : \mathbb{R}^1 \rightarrow [0, 1]$$

$F_{F_{\xi}(\xi)} \sim U[0, 1]$ — равномерна на отрезке

5. Пусть $\eta \sim U[0, 1]$, $F(x)$ — непрерывная функция распределения. Тогда

$$F_{F^{-1}(\eta)}(x) = P\{F^{-1}(\eta) \leq x\} = P\{\eta \leq F(x)\} = F_\eta(F(x)) = F(x)$$

$y = F^{-1}(x)$, где $x \in [0, 1]$.

Строим случайную величину с наперёд заданным распределением, используя случайную величину с равномерным распределением.

Рассмотрим ситуацию, когда борелевская функция g — функция двух переменных. Пусть ξ, η — случайные величины, их функция распределения — $F_{\xi, \eta}(x, y)$, $g(\xi, \eta)$, $z \in \mathbb{R}^1$.

$$F_{g(\xi, \eta)}(z) = P\{g(\xi, \eta) \leq z\} = \iint_{\{x, y: g(x, y) \leq z\}} dF_{\xi, \eta}(x, y)$$

Также пусть $g(x, y) = x + y$, ξ, η — независимые случайные величины

$$\begin{aligned} F_{\xi+\eta}(z) &= \iint_{\{x, y: x+y \leq z\}} dF_\xi(x) dF_\eta(y) = \\ &= \iint_{\mathbb{R}^2} I_{\{x+y \leq z\}}(x, y) dF_\xi(x) dF_\eta(y) = \end{aligned}$$

По теореме Фубини:

$$\begin{aligned} &= \int_{-\infty}^{+\infty} dF_\xi(x) \left[\int_{-\infty}^{+\infty} I_{\{x+y \leq z\}}(x, y) dF_\eta(y) \right] = \int_{-\infty}^{+\infty} F_\eta(z - x) dF_\xi(x) \\ F_{\xi+\eta}(z) &= \int_{-\infty}^{+\infty} F_\xi(z - y) dF_\eta(y) \end{aligned}$$

Если имеются функции распределения F, G , то

$$H(z) = \underbrace{\int_{-\infty}^{+\infty} F(z-x)dG(x) = \int_{-\infty}^{+\infty} F(z-y)dF(y)}_{\text{свёртка } F \text{ и } G \text{ (композиция)}}$$

Функция распределения суммы двух случайных (независимых) величин есть свёртка функций распределения слагаемых.

Пусть существует функция плотности вероятности $f_\xi(x), f_\eta(y)$. Тогда

$$\begin{aligned} F_{\xi+\eta}(z) &= \int_{-\infty}^{+\infty} \left[\int_{-\infty}^{z-y} f_\xi(u)du \right] \cdot f_\eta(y)dy = [\nu = u + y] = \\ &= \int_{-\infty}^{+\infty} \left[\int_{-\infty}^z f_\xi(\nu - y)du \right] \cdot f_\eta(y)dy = \end{aligned}$$

По теореме Фубини:

$$\int_{-\infty}^{+\infty} \left[\int_{-\infty}^{+\infty} f_\xi(\nu - y)f_\eta(y)dy \right] d\nu$$

Итак:

$$f_{\xi+\eta}(z) = \int_{-\infty}^{+\infty} f_\xi(z-y)f_\eta(y)dy = \int_{-\infty}^{+\infty} f_\eta(z-x)f_\xi(x)dx$$

Пусть имеется набор независимых случайных величин $\xi_1, \dots, \xi_n, \forall i : \xi_i \sim N(0, 1)$.

Составим случайную величину, подчиняющуюся распределению χ^2

$$\xi_1^2 + \dots + \xi_n^2 = \chi_n^2$$

$$f_{\chi_n^2}(x) = \begin{cases} \frac{x^{\frac{n}{2}} \cdot e^{-\frac{x}{2}}}{2^{\frac{n}{2}} \cdot \Gamma\left(\frac{n}{2}\right)}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

Составим распределение Стьюдента

$$t_n = \frac{\xi_o}{\sqrt{\frac{1}{n} \sum_{i=1}^n \xi_i^2}} = \frac{\varepsilon_0 \sqrt{n}}{\sqrt{\chi_n^2}}$$

где χ_n — χ -распределение с n степенями свободы

$$\chi_n = \sqrt{\chi_n^2} \sim f_{\chi_n}(x) = \begin{cases} \frac{2 \cdot x^{n-1} \cdot e^{-\frac{x^2}{2}}}{2^{\frac{n}{2}} \cdot \Gamma\left(\frac{n}{2}\right)}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

Можно показать, что

$$f_{t_n}(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n \cdot \pi} \cdot \Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}$$

Если $n \rightarrow \infty$, то получим функцию плотности вероятности нормального распределения

$$f_{t_n}(x) \xrightarrow{n \rightarrow \infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Если $n = 1$, тогда получим функцию плотности вероятности распределения Коши с параметром 1

$$f_{t_n}(x) = \frac{1}{\pi(1+x^2)}$$

Рассмотрим случайные величины ξ, η

- $\xi \sim \{x_i\}_{i=1,2,\dots},$
- $\eta \sim \{y_i\}_{i=1,2,\dots},$
- $\xi + \eta \sim \{z_k\}_{k=1,2,\dots}$

$$\begin{aligned} P\{\xi + \eta = z_k\} &= \sum_i P\{\xi = x_i\} P\{\eta = z_k - x_i\} = \\ &= \sum_j P\{\eta = y_j\} P\{\xi = z_k - y_j\} \end{aligned}$$

Тогда функция распределения

$$F_{g(\xi,\eta)}(z) = \iint_{\{(x,y): g(x,y) \leq z\}} dF_{\xi,\eta}(x, y)$$

Рассмотрим произведение этих случайных величин. Пусть также ξ, η — независимые. Тогда функция плотности вероятности

$$\begin{aligned} f_{\xi \cdot \eta}(z) &= \int_{-\infty}^{+\infty} f_{\xi}\left(\frac{z}{y}\right) f_{\eta}(y) \frac{1}{|y|} dy \\ f_{\frac{\xi}{\eta}}(z) &= \int_{-\infty}^{+\infty} f_{\xi}(zy) f_{\eta}(y) |y| dy = \frac{1}{z^2} \int_{-\infty}^{+\infty} f_{\eta}\left(\frac{x}{z}\right) f_{\xi}(x) |x| dx \end{aligned}$$

Пусть задан вектор случайных величин с абсолютно непрерывным распределением

$$\bar{\xi} = (\xi_1, \dots, \xi_n), f_{\xi}(x), x = (x_1, \dots, x_n)$$

Пусть g — борелевская функция

$$g_1 : \mathbb{R}^n \rightarrow \mathbb{R}^1, \quad i = 1, \dots, n$$

Рассмотрим

$$g : \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad g(x) = (g_1(x), \dots, g_n(x))$$

$$J(x) = \left| \frac{\partial(g_1, \dots, g_n)}{\partial(x_1, \dots, x_n)} \right| \neq 0$$

Сконструируем вектор

$$\bar{\eta} = (\eta_1, \dots, \eta_n), \quad \eta_i = g_i(\bar{\xi})$$

Его распределение описывается следующей функцией плотности вероятности

$$f_{\bar{\eta}}(x) = f_{\xi}(g^{-1}(x)) = |J(g^{-1}(x))|^{-1}$$

Глава IV

Характеристические функции. Предельные теоремы

§1 Производящие функции. Фактори- альные моменты

DRAFT