ФИЗИЧЕСКИЙ
ФАКУЛЬТЕТ
МГУ ИМЕНИ
М.В. ЛОМОНОСОВА

teach-in
ЛЕКЦИИ УЧЕНЫХ МГУ

ТЕОРИЯ ВЕРОЯТНОСТЕЙ

ЧУЛИЧКОВ
АЛЕКСЕЙ ИВАНОВИЧ

ФИЗФАК МГУ

КОНСПЕКТ ПОДГОТОВЛЕН СТУДЕНТАМИ, НЕ ПРОХОДИЛ ПРОФ. РЕДАКТУРУ И МОЖЕТ СОДЕРЖАТЬ ОШИБКИ. СЛЕДИТЕ ЗА ОБНОВЛЕНИЯМИ НА VK.COM/TEACHINMSU.

ЕСЛИ ВЫ ОБНАРУЖИЛИ ОШИБКИ ИЛИ ОПЕЧАТКИ, ТО СООБЩИТЕ ОБ ЭТОМ, НАПИСАВ СООБЩЕСТВУ VK.COM/TEACHINMSU.

БЛАГОДАРИМ ЗА ПОДГОТОВКУ КОНСПЕКТА
СТУДЕНТКУ ФИЗИЧЕСКОГО ФАКУЛЬТЕТА МГУ
КУЩЕНКО ЮЛИЮ КОНСТАНТИНОВНУ

ОГЛАВЛЕНИЕ

ЛЕКЦИЯ 1. ОСНОВНЫЕ ПОНЯТИЯ ТЕОРИИ ВЕРОЯТНОСТЕЙ 6

Фундаментальные понятия теории вероятностей 6
Примеры вероятностных пространств 8
События и их свойства 9
Свойства операций над событиями 11
Алгебра событий 12
Вероятность, аксиомы вероятности 13
Предел последовательности событий 14

ЛЕКЦИЯ 2. ПОСЛЕДОВАТЕЛЬНОСТИ СОБЫТИЙ 17

Теорема о сходимости монотонных последовательностей: 17
Теорема о непрерывности вероятности относительно монотонных предельных переходов 19
Теорема о непрерывности вероятности относительно любых типов сходимости . 20
Примеры вероятностных пространств. Дискретное вероятностное пространство 21
Условное вероятностное пространство 24
Независимость событий 26

ЛЕКЦИЯ 3. РАСПРЕДЕЛЕНИЯ ТЕОРИИ ВЕРОЯТНОСТЕЙ 27

Последовательность независимых испытаний 27
Схема Бернулли 29
Распределение Паскаля 29
Полиномиальное распределение 30
Формула Пуассона 30
Теоремы Муавра-Лапласа 31
Модель «случайные величины» 32

ЛЕКЦИЯ 4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ 34

Теория случайных величин 34
Классификация случайных величин по виду их функции распределения 38
Случайные векторы 41

ЛЕКЦИЯ 5. НЕЗАВИСИМОСТЬ СЛУЧАЙНЫХ ВЕЛИЧИН 45

Функции от случайных величин 45
Независимость случайных величин 47
Условные распределения случайных величин 49
Числовые характеристики случайных величин 50

ЛЕКЦИЯ 6. МАТЕМАТИЧЕСКОЕ ОЖИДАНИЕ И ДИСПЕРСИЯ 54

Числовые характеристики случайных величин 54
Неравенства Чебышёва 59

ЛЕКЦИЯ 7. НЕРАВЕНСТВО КОШИ-БУНЯКОВСКОГО 61

Неравенство Коши-Буняковского: 61
Условное распределение случайных величин 62
Условное математическое ожидание случайной величины ξ 63
Задача наилучшего в среднем квадратичном прогноза 64
Задача о наилучшем в среднем квадратичном линейном прогнозе 65
Векторные случайные величины 66

ЛЕКЦИЯ 8. ЗАКОНЫ БОЛЬШИХ ЧИСЕЛ 68

Метод Монте-Карло 70
Сходимость с вероятностью 1 (почти наверное) 71
Лемма Бореля-Кантелли 73
Усиленный закон больших чисел 73

ЛЕКЦИЯ 9. ХАРАКТЕРИСТИЧЕСКИЕ ФУНКЦИИ СЛУЧАЙНЫХ ВЕЛИЧИН 76

Сходимость по распределению 76
Характеристические функции случайных величин 77
Теорема о непрерывности характеристических функций 79
Центральная предельная теорема 80
Восстановление распределения случайных величин по наблюдению реализации (применение ЗБЧ и ЦПТ): 81
Интегральная теорема Муавра-Лапласа 82

ЛЕКЦИЯ 10. ЦЕПИ МАРКОВА 83

Конечные однородные цепи Маркова 83
Матрицы перехода 85
Эргодичность цепей Маркова 87

ЛЕКЦИЯ 11. ТЕОРИЯ АЗАРТНЫХ ИГР 91

Теория азартных игр 91
Многомерное нормальное распределение 93
Ортогональное преобразование 95

ЛЕКЦИЯ 12. ТЕОРИЯ ВОЗМОЖНОСТЕЙ 100

Операции с возможностями 101
Упорядоченность возможностей 105

ЛЕКЦИЯ 13. ОСНОВЫ ТЕОРИИ ВОЗМОЖНОСТЕЙ 107

Повторение материала предыдущей лекции 107
Построение теории возможностей 108
Нечеткие множества 109
P-независимые события. Условная возможность. 109

ЛЕКЦИЯ 1. ОСНОВНЫЕ ПОНЯТИЯ ТЕОРИИ ВЕРОЯТНОСТЕЙ

Введение

Приступим к изучению теории вероятностей как математической дисциплины. В качестве эталона математической науки будем рассматривать аксиоматическую теорию вероятности, в которой заданы аксиомы, а все возникающие при изучении теоретические вопросы решаются с помощью правил логического вывода.

Вероятность – это некоторая величина, которая контролирует частоту появления событий в длинной серии испытаний. Таким образом, теория вероятностей предназначена для моделирования систем, в которых проводятся эксперименты, исходы которых мы не можем предсказать, но можем предсказать их частоту появления в длинной серии независимых испытаний. Вероятность есть предел частоты появления события.

С помощью теории вероятности нельзя моделировать системы, в которых эксперимент проводится один раз, или же эксперимент проводится много раз, но частота возникновения от раза к разу меняется непредсказуемым образом. Для моделирования таких систем существуют другие математические дисциплины, с одной из них - теорией возможностей мы познакомимся в конце курса.

Фундаментальные понятия теории вероятностей

Эксперимент со случайным исходом – эксперимент, исход которого мы не можем однозначно предсказать, однако нам известно множество всех его элементарных исходов. Все исходы определены, и ничего другого произойти не может.

$$\{\omega_\lambda,\ \lambda \in \Lambda\} = \Omega,$$

ω_λ – элементарный исход эксперимента,

Ω – пространство элементарных исходов.

Ω обязательно задано для любого эксперимента со случайным исходом и состоит из некоторого набора элементарных исходов.

Пример: Если подбросить монетку, она может упасть вверх орлом или вверх решкой. Мы не знаем, чем закончится эксперимент, однако ничем другим, он закончиться не может. При желании можно приписать и другие исходы.

Пример: Стреляем в мишень, при этом эксперимент заканчивается при попадании в мишень. Не факт что мы попадем с первого раза. Множество всех элементарных исходов бесконечное, но счетное, потому что можно попасть с первого раза, можно один раз промазать и во второй раз попасть, можно два раза промазать, а в третий раз попасть и т.д. Множество Λ есть натуральный ряд чисел.

Пример: Случайный выбор точки на отрезке [0,1]. Точек на отрезке бесконечно много. Множество элементарных исходов бесконечно и несчетно, каждому элементарному исходу не удается единственным образом сопоставить натуральное число.

Ω удобно отображать как множество точек на плоскости (Рис. 1.1).

$$\omega \in \Omega.$$

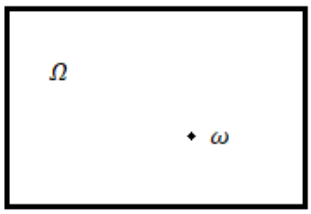


Рис. 1.1. Пространство элементарных исходов.

Также интересны события, которые складываются из нескольких элементарных исходов. *Например,* подбрасываем игральную кость, существует 6 элементарных исходов. Но если мы интересуемся выпадением четного числа очков, то это событие состоит из трех элементарных. Если выпадет 2, то случится событие А (выпадение четного числа), если выпадет 4, событие А тоже случится, если выпадет 6 – тоже.

$$\{1, 2, 3, 4, 5, 6\} = \Omega,$$

$$A = \{2, 4, 6\},$$

$$2 \in A, \quad A \subset \Omega.$$

На языке теории вероятностей элементарные исходы ω_A влекут события А. На языке теории множеств точка ω_A является точкой множества А, и множество А задается указанием всех элементарных исходов, которые его влекут. Так ω_A есть элемент Ω, а событие А есть подмножество Ω. (Рис. 1.2)

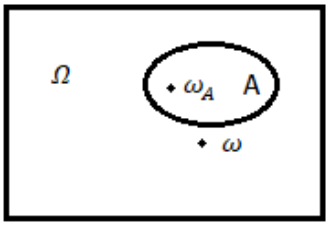


Рис. 1.2. Событие А - подмножество Ω.

Некоторые события могут состоять из одного элементарного. Например, событие В состоит из какого-то одного единственного элементарного события ω и это есть подмножество события Ω.

$$B = \{\omega\} \subset \Omega,$$

$$\omega \in \Omega.$$

ω - элементарное событие, а B – одноточечное событие, но часто разницу между ними не подчеркивают, и считается, что произошло некоторое событие, которое состоит из одного элементарного.

Вероятность P(A), она определена для каждого события A и контролирует частоту появления события A в бесконечной серии независимо проведенных экспериментов со случайным исходом. Важно, что в этой серии условия проведения экспериментов таковы, что вероятность не может меняться произвольно. Такие эксперименты называются экспериментами со стохастической устойчивостью, нет изменения вероятности от эксперимента к эксперименту. В теории вероятности рассматриваются только модели, в которых есть стохастическая устойчивость.

Таким образом, основные понятия теории вероятностей - это пространство элементарных событий Ω, некоторая система его подмножеств F (т.е. множество всех событий) и вероятность $P(\cdot)$, которая задана для каждого события системы F. Эта вероятностная тройка называется вероятностным пространством:

$$(\Omega, F, P(\cdot)).$$

Примеры вероятностных пространств

- ***Классическая вероятность***

Пространство элементарных исходов состоит из конечного числа исходов:

$$\Omega = \{\omega_1, \dots, \omega_N\}.$$

Задается вероятность каждого элементарного исхода, она всегда одна и та же:

$$p(\omega_j) = \frac{1}{N}.$$

Это позволяет задать вероятность события A, как отношение числа благоприятствующих исходов к общему числу исходов:

$$P(A) = \frac{m}{N}.$$

Где m - это число элементарных исходов, которые влекут событие A:

$$m: \{\omega_{j_1}, \dots, \omega_{j_m}\} = A.$$

- ***Геометрическая вероятность***

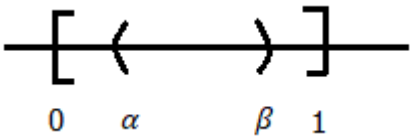


Рис. 1.3. Интервал на отрезке от 0 до 1.

Например, точка наугад бросается на отрезок от 0 до 1, причем нет предпочтения одних точек перед другими. Тогда разумно считать, что вероятность попадания в интервал от α до β пропорциональна длине этого отрезка. В данном случае просто равна:

$$P([\alpha, \beta]) = \beta - \alpha.$$

Таким образом, вероятность можно задать для любого множества, которое представляется в виде конечной суммы или даже бесконечной счетной суммы непересекающихся отрезков. Множество таких отрезков называется Борелевскими.

Борелевское множество – это такое множество, которое может быть получено как конечное или счетное объединение не пересекающихся отрезков. Не все подмножества отрезка $[0,1]$ являются Борелевскими, и эта вероятностная модель задана не на любом подмножестве отрезка. Это принципиальное ограничение.

События и их свойства

Уже обсуждалось, что если какое-то элементарное событие ω с точки зрения теории множеств является элементом события A, то с точки зрения теории вероятности это интерпретируется так, что событие ω влечет события A.

$$\omega \in A.$$

1) События могут быть упорядочены таким образом, что событие A влечет событие B:

$$A \subset B.$$

Это означает что все те элементарные исходы, которые влекут события A, влекут и события B. То есть, множество A является подмножеством множества B. (Рис. 1.4).

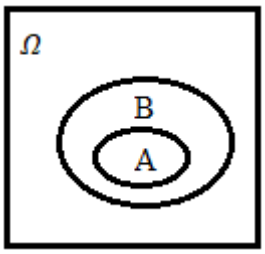


Рис. 1.4. Множество A – подмножество множества B.

Если A влечет B, а B влечет A, то такие события являются эквивалентными, потому что они состоят из одних и тех же элементарных исходов:

$$A \subset B, B \subset A \Leftrightarrow A = B.$$

2) Существует дополнение события A до всего пространства элементарных событий. Дополнение $\bar{A}$ – это событие, состоящее в том, что A не наступает. (Рис. 1.5).

$$\bar{A} = \{\omega \in \Omega : \omega \notin A\}.$$

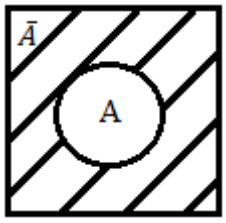


Рис. 1.5. Дополнение события A.

Дважды противоположное событие совпадает самим событиям A, потому что это такие события, которые не входят в те события, которые влекут не A:

$$\bar{\bar{A}} = A.$$

Если A и B и эквивалентные события, согласно приведенному выше определению, дополнения этих двух событий тоже совпадают:

$$A = B \Leftrightarrow \bar{A} = \bar{B}.$$

3) Пространство элементарных исходов можно рассматривать как событие, состоящее из всех элементарных событий:

$$\Omega = \{\Omega\}.$$

Ω – это достоверное событие, поэтому возникает необходимость ввести противоположное событие. В теории множеств это пустое событие, оно не состоит не из одного элемента. Это невозможное событие (не путать с невероятным):

$$\bar{\Omega} = \emptyset, \qquad \bar{\emptyset} = \Omega.$$

4) Возможно пересечение событий:

$$A \cap B.$$

С точки зрения теории множеств пересечение A и B – это множество тех элементарных исходов, которые входят и в A и в B (Рис. 1.6). А с точки зрения теории вероятности события A и B происходят одновременно. Пересечение можно эквивалентно записать без использования знака пересечения, подобно умножению: AB.

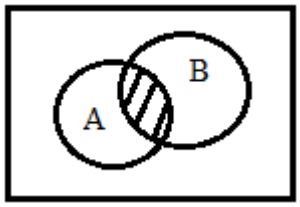


Рис. 1.6. Пересечение событий.

Несовместные события – это события, при пересечении которых получаем пустое множество, события у которых нет общих точек (Рис 1.7).

$$AB = \emptyset.$$

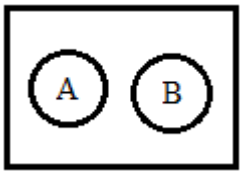


Рис. 1.7. Несовместные события.

5) Можно рассматривать сумму или объединение двух событий:

$$A \cup B.$$

Объединение A с B состоит из тех ω, которые влекут либо A, либо B, либо оба (Рис.1.8). Хотя бы одно событие происходит.

Рис. 1.8. Объединение событий.

Если A и B: $AB = \emptyset$, то A∪B → A+B. (Если A и B несовместные, то знак пересечения заменяется на +.)

6) Можно рассматривать разность 2 событий:

$$A \backslash B = \{\omega \in A, \omega \notin B\}.$$

A без B - это такие точки ω, которые влекут A, но при этом не влекут B (Рис.1.9).

$$A \backslash B = A\bar{B}.$$

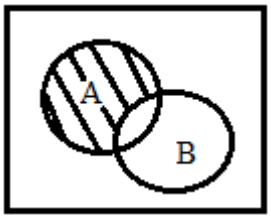


Рис. 1.9. Событие A без B.

Формально равенство доказывается по следующей схеме:

$$C_1 = A \backslash B;\ C_2 = A\bar{B}$$

$$C_1 \subset C_2$$

$$C_2 \subset C_1$$

$$\Leftrightarrow C_1 = C_2$$

Свойства операций над событиями

1) Ассоциативность:

$$(A \cup B) \cup C = A \cup (B \cup C),$$
$$(AB)C = A(BC).$$

2) Коммутативность:

$$A \cup B = B \cup A,$$
$$AB = BA.$$

3) Дистрибутивность:

Предварительно следует дать определения объединению и пересечению счетного или несчетного числа событий.

$$\bigcup_{\lambda \in \Lambda} A_\lambda = \{\omega: \exists \lambda \in \Lambda: \omega \in A_\lambda\},$$

Объединение всех событий A_λ состоит из таких элементарных исходов, для которых найдется хотя бы одно λ из множества Λ, такое что $\omega \in A_\lambda$.

$$\bigcap_{\lambda \in \Lambda} A_\lambda = \{\omega: \forall \lambda \in \Lambda: \omega \in A_\lambda\},$$

Пересечение состоит из элементарных исходов ω таких, что для любого λ из Λ выполняется $\omega \in A_\lambda$. То есть, ω должно содержаться во всех событиях A_λ, когда λ пробегает Λ.

Дистрибутивность:

$$A \cup \left(\bigcap_{\lambda \in \Lambda} A_\lambda\right) = \bigcap_{\lambda \in \Lambda} (A \cup A_\lambda).$$
$$A \cap \left(\bigcup_{\lambda \in \Lambda} A_\lambda\right) = \bigcup_{\lambda \in \Lambda} (A \cap A_\lambda).$$

4) Свойства двойственности (де Моргана):

$$\bar{A} \cap \bar{B} = \overline{A \cup B}, \qquad \bar{A} \cup \bar{B} = \overline{A \cap B}.$$
$$\overline{\bigcup_{\lambda \in \Lambda} A_\lambda} = \bigcap_{\lambda \in \Lambda} \overline{A_\lambda}, \qquad \overline{\bigcap_{\lambda \in \Lambda} A_\lambda} = \bigcup_{\lambda \in \Lambda} \overline{A_\lambda}.$$

Алгебра событий

Опр.: Алгеброй событий F называется система событий, которая удовлетворяет двум свойствам:

1) Если $A \in F$, то $\bar{A} \in F$,

2) Если $A, B \in F$, то $A \cup B \in F$.

Следствия:

1) Достоверное событие всегда является событием из алгебры F:

$$A \in F, \qquad \bar{A} \in F,$$

$$A \cup \bar{A} = \Omega \in F.$$

2) Пустое событие тоже принадлежит алгебре F:

$$\emptyset = \bar{\Omega} \in F.$$

3) Пересечение событий из алгебры F принадлежит алгебре F

$$A, B \in F, \qquad AB = \overline{\bar{A} \cup \bar{B}} \in F.$$

Алгебра событий – это множество событий, замкнутых относительно тех операций, которые мы определили выше.

Опр.: Алгебра F событий называется σ-алгеброй, если из $A_1 \dots A_n \dots \in F$, следует

$$\bigcup_{j=1}^{\infty} A_j \in F.$$

Также для σ-алгебры справедливо:

$$\overline{\bigcup_{j=1}^{\infty} \bar{A}_j} = \bigcap_{j=1}^{\infty} A_j \in F.$$

Опр.: Борелевская алгебра – это алгебра, которая получена конечным или счетным числом объединений и пересечений всех отрезков числовой прямой. (На плоскости она получена конечным числом объединений и пересечений всех прямоугольников, со сторонами параллельными осям координат).

Опр.: Если не разрешено делать счетное объединение и пересечение, то алгебра будет Жордановой.

Вероятность, аксиомы вероятности

Опр.: Вероятностью называется функция P(·):

1) P(·): F → [0,1], где F σ-алгебра,

2) $0 \le P(A) \le 1, \ \forall A \in F.$

3) $P(\Omega) = 1$ (нормировка).

4) Если AB $= \emptyset$, то P(A+B) = P(A) + P(B) (аддитивность).

5) σ-аддитивность. Для несовместных событий $A_1, A_2, \dots, A_n, ..$ выполняется:

$$P\left(\sum_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} P(A_j).$$

Простейшие свойства вероятности

1) Монотонность: если $A \subset B$, то $P(A) \leq P(B)$. Док-во:

$$B = A + B\backslash A,$$

$$P(B) = P(A) + P(B\backslash A) \geq 0,$$

$$P(B) \geq 0.$$

2) Если $A \subset B$, то $P(B\backslash A) = P(B) - P(A)$. Док-во:

$$B = A + B\backslash A,$$

$$P(B\backslash A) = P(B) - P(A).$$

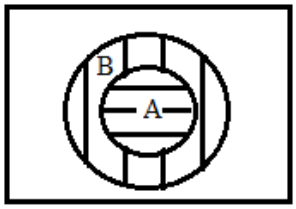


Рис. 1.10. Событие B без A.

3) $P(A) = 1 - P(\bar{A})$ и наоборот. Док-во:

$$A + \bar{A} = \Omega,$$

$$P(A) + P(\bar{A}) = P(\Omega) = 1.$$

4) $P(A \cup B) = P(A) + P(B) - P(AB)$. Док-во:

$$A \cup B = A + B\backslash A = A + B\backslash AB,$$

$$P(A \cup B) = P(A) + P(B\backslash AB) = P(A) + P(B) - P(AB) \leq 1 \quad \forall A \in F.$$

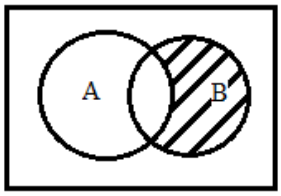


Рис. 1.11. Объединение A и B.

Предел последовательности событий

$$\{A_n\}, \qquad n = 1,2, \ldots$$

Опр.: Верхним пределом последовательности событий $\{A_n\}$ назовем событие:

$$\tilde{A} = \bigcap_{n=1}^{\infty} \bigcup_{k \geq n} A_k.$$

1) Рассмотрим $\omega \in \Omega$ такое, что существует только конечное число A_k: $\omega \in A_k$. Такое ω не попадет в $\tilde{A}$. Док-во:

$$\exists\, N: \forall k > N \quad \omega \notin A_k \implies$$

$$\omega \notin \bigcup_{k \geq N} A_k \implies \omega \notin \tilde{A}.$$

2) Рассмотрим ω, содержащееся в бесконечном числе A_k: $\forall N\ \exists k$: $\omega \in A_k$.

$$\omega \in \bigcup_{k \geq N} A_k \quad \forall n = 1,2, \ldots$$

$$\omega \in \bigcap_{n=1}^{\infty} \bigcup_{k \geq n} A_k.$$

$$\tilde{A} = \{\text{все } \omega \in \Omega, \text{которые влекут } \infty \text{ число } A_k\}.$$

Пример: Верхний предел – точка 0. (Рис. 1.12). $\left[0, \frac{1}{n}\right] = A_n$:

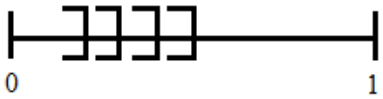


Рис. 1.12. Стремление последовательности к пределу.

Опр.: Нижним пределом последовательности событий $\{A_n\}$ назовем событие:

$$A_{\sim} = \bigcup_{n=1}^{\infty} \bigcap_{k \geq n} A_k.$$

1) Рассмотрим $\omega \in A_k$ $\forall k = N, N+1, \ldots$

$$\text{т.е. } \exists N: \quad A_k \ni \omega \quad k = N, N+1, \ldots$$

$$\bigcap_{k \geq N} A_k \ni \omega,$$

$$\bigcup_{N=1}^{\infty} \bigcap_{k \geq N} A_k \ni \omega.$$

2) Рассмотрим случай, когда $\forall N\ \exists k > N$: $\omega \notin A_k$.

$$\omega \notin \bigcap_{k \geq N} A_k \implies \omega \notin A_{\sim}.$$

Опр.: Последовательность событий называется сходящейся, если верхний предел равен нижнему пределу:

$$\tilde{A} = A_{\sim} = A$$

$$A - \text{предел}\,\{A_k\}, \qquad k \to \infty.$$

Свойство: Если $\{A_k\} \to A$, при $k \to \infty \Longrightarrow \{\overline{A_k}\} \to \bar{A}$, при $k \to \infty$. Док-во:

$$\bar{A}_{\sim} = \bigcup_{n=1}^{\infty} \bigcap_{k \geq n} \bar{A}_k = \bigcup_{n=1}^{\infty} \overline{\bigcup_{k \geq n} A_k} = \overline{\bigcap_{n=1}^{\infty} \bigcup_{k \geq n} A_k} = \bar{A} = \bigcap_{n=1}^{\infty} \bigcup_{k \geq n} \bar{A}_k = \bar{\tilde{A}}.$$

ЛЕКЦИЯ 2. ПОСЛЕДОВАТЕЛЬНОСТИ СОБЫТИЙ

Вспомним определение верхнего и нижнего пределов:

$$\tilde{A} = \bigcap_{n=1}^{\infty} \bigcup_{k \geq n} A_k = \{\omega \in \Omega \colon \forall N \; \exists n > N \colon \omega \in A_n\}.$$

Верхний предел состоит из всех элементарных событий, которые входят в бесконечное число A_k.

$$A_{\sim} = \bigcup_{n=1}^{\infty} \bigcap_{k \geq n} A_k = \{\omega \in \Omega \colon \exists N \colon \; n \geq N \;\; \omega \in A_n \}.$$

Нижний предел это множество тех элементарных исходов, которые влекут все события A_k, начиная с некоторого номера.

Опр.: Последовательность $B_1 \subset B_2 \subset \cdots \subset B_n \subset \cdots$ - монотонно возрастающая.

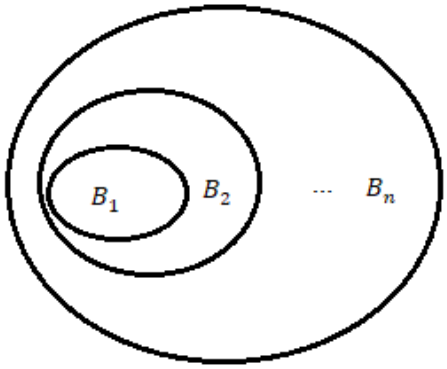


Рис. 2.1. Монотонно возрастающая последовательность.

Опр.: Последовательность $D_1 \supset D_2 \supset \cdots \supset D_n \supset \cdots$ - монотонно убывающая.

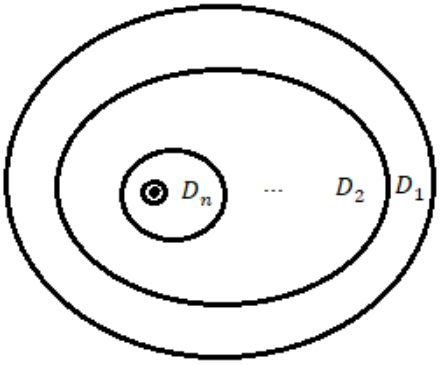


Рис. 2.2. Монотонно убывающая последовательность.

Теорема о сходимости монотонных последовательностей:

Монотонные последовательности сходятся.

$$\lim_{n\to\infty} B_n = \bigcup_{n=1}^{\infty} B_n.$$

$$\lim_{n\to\infty} D_n = \bigcap_{n=1}^{\infty} D_n.$$

Доказательство:

a) для монотонно возрастающей последовательности:

$$\tilde{B} = \bigcap_{n=1}^{\infty} \bigcup_{k\geq n} B_k,$$

$\bigcup_{k\geq n} B_k$ можно заменить на $\bigcup_{k=1}^{\infty} B_k$, останется то же самое множество (см. Рис. 2.1).

$$\tilde{B} = \bigcap_{n=1}^{\infty} \bigcup_{k\geq n} B_k = \bigcap_{n=1}^{\infty} \bigcup_{k=1}^{\infty} B_k = \bigcup_{n=1}^{\infty} B_n.$$

$$B_{\sim} = \bigcup_{n=1}^{\infty} \bigcap_{k\geq n} B_k,$$

$\bigcap_{k\geq n} B_k$ можно заменить на Bn (см. Рис. 2.1).

$$B_{\sim} = \bigcup_{n=1}^{\infty} \bigcap_{k\geq n} B_k = \bigcup_{n=1}^{\infty} B_n.$$

$$\Longrightarrow B_{\sim} = \tilde{B} = \bigcup_{n=1}^{\infty} B_n = \lim_{n\to\infty} B_n.$$

b) Для монотонно убывающей последовательности:

$$D_1 \supset D_2 \supset \cdots \supset D_n \supset \cdots$$

$$\Longrightarrow \overline{D_1} \subset \overline{D_2} \subset \cdots \subset \overline{D_n} \subset \cdots$$

Для таких последовательностей доказательство приведено в пункте a.

$$\Longrightarrow \lim_{n\to\infty} \overline{D_n} = \bigcup_{n=1}^{\infty} \overline{D_n}.$$

$$\Longrightarrow \lim_{n\to\infty} \overline{\overline{D_n}} = \lim_{n\to\infty} D_n = \overline{\bigcup_{n=1}^{\infty} \overline{D_n}}.$$

Необходимо воспользоваться правилом де Моргана:

$$\overline{\bigcup_{n=1}^{\infty} \overline{D_n}} = \bigcap_{n=1}^{\infty} \overline{\overline{D_n}}$$

$$\Longrightarrow \lim_{n\to\infty} D_n = \bigcap_{n=1}^{\infty} D_n.$$

Теорема о непрерывности вероятности относительно монотонных предельных переходов:

$$P\left(\lim_{n\to\infty} B_n\right) = \lim_{n\to\infty} P(B_n).$$

$$P\left(\lim_{n\to\infty} D_n\right) = \lim_{n\to\infty} P(D_n).$$

Доказательство:

a) для монотонно возрастающей последовательности:

$$P\left(\lim_{n\to\infty} B_n\right) = P\left(\bigcup_{n=1}^{\infty} B_n\right).$$

Разобьем множество B_n на множество непересекающихся множеств, для этого введем $B_0 = \emptyset$:

$$P\left(\lim_{n\to\infty} B_n\right) = P(B_1 \backslash B_0 + B_2 \backslash B_1 + \cdots + B_n \backslash B_{n-1} + \cdots) =$$

$$= P(B_1 \backslash B_0) + P(B_2 \backslash B_1) + \cdots + P(B_n \backslash B_{n-1}) + \cdots =$$

Воспользуемся свойством вероятности $P(B \backslash A) = P(B) - P(A)$.

$$= \cancel{P(B_1)} - P(B_0) + \cancel{P(B_2)} - \cancel{P(B_1)} + \cdots + \cancel{P(B_n)} - \cancel{P(B_{n-1})} + P(B_{n+1}) - \cancel{P(B_n)} + \cdots =$$

$$= \lim_{n\to\infty} P(B_n).$$

Этот предел существует, $P(B_n)$ – монотонная неубывающая последовательность (в силу монотонности B_n), ограниченная сверху.

b) Для монотонно убывающей последовательности:

$$\{\overline{D_n}\} - \text{монотонно возрастает},$$

Для монотонно возрастающих последовательностей уже доказали:

$$P\left(\lim_{n\to\infty} \overline{D_n}\right) = \lim_{n\to\infty} P(\overline{D_n}).$$

Левую часть этого уравнения запишем, используя предыдущую теорему, а затем воспользуемся формулой де Моргана:

$$P\left(\lim_{n\to\infty}\overline{D_n}\right) = P\left(\bigcup_{n=1}^{\infty}\overline{D_n}\right) = 1 - P\left(\overline{\bigcup_{n=1}^{\infty}\overline{D_n}}\right) = 1 - P\left(\bigcap_{n=1}^{\infty}D_n\right) = 1 - P\left(\lim_{n\to\infty}D_n\right).$$

Для правой части уравнения верно следующее:

$$\lim_{n\to\infty}P(\overline{D_n}) = \lim_{n\to\infty}\big(1 - P(D_n)\big) = 1 - \lim_{n\to\infty}\big(P(D_n)\big).$$

Соответственно:

$$1 - P\left(\lim_{n\to\infty}D_n\right) = 1 - \lim_{n\to\infty}\big(P(D_n)\big),$$

$$P\left(\lim_{n\to\infty}D_n\right) = \lim_{n\to\infty}P(D_n).$$

Предел существует, так как последовательность D_n монотонно убывающая, и вероятности образуют монотонно убывающую, ограниченную снизу последовательность.

Теорема о непрерывности вероятности относительно любых типов сходимости:

$$\text{Пусть}\quad \lim_{n\to\infty}A_n = A = A_\sim = \tilde{A},$$

$$\text{тогда:}\quad \lim_{n\to\infty}P(A_n) = P(\lim_{n\to\infty}A_n).$$

Непрерывность вероятности распространяется на любые сходящиеся последовательности.

Доказательство:

$$P(A_\sim) = P\left(\bigcup_{n=1}^{\infty}\bigcap_{k\geq n}A_k\right).$$

$$\bigcap_{k\geq n}A_k - \text{чем больше } n, \text{тем меньше пересекающихся сомножителей,}$$

и тем больше это множество. Значит, пересечение – элемент монотонно возрастающей последовательности B_n.

$$\Rightarrow \bigcup_{n=1}^{\infty}\bigcap_{k\geq n}A_k = \bigcup_{n=1}^{\infty}B_n = B = \lim_{n\to\infty}B_n.$$

$$\Rightarrow P(A_\sim) = \lim_{n\to\infty}P\left(\bigcap_{k\geq n}A_k\right) \leq \underline{\lim}_{n\to\infty}P(A_n).$$

Вероятность A_n больше чем вероятность пересечения всех A_k, т.к. $\bigcap_{k\geq n}A_k \subset A_n$. Имеем право записать только нижний предел, т.к. не знаем, сходится ли последовательность.

$$\underline{\lim_{n\to\infty}} P(A_n) \le \overline{\lim_{n\to\infty}} P(A_n) \le \lim_{n\to\infty} P\left(\bigcup_{k\ge n} A_k\right)$$

При замене A_n на объединение уже можем записать предел, потому что объединение – это бесконечно убывающая и ограниченная последовательность:

$$\bigcup_{k\ge n} A_k = D_n, \qquad \lim_{n\to\infty} D_n = \bigcap_{n=1}^{\infty} D_n.$$

В силу непрерывности вероятности относительно монотонной сходимости:

$$\lim_{n\to\infty} P\left(\bigcup_{k\ge n} A_k\right) = P\left(\lim_{n\to\infty} \bigcup_{k\ge n} A_k\right) = P\left(\bigcap_{n=1}^{\infty} \bigcup_{k\ge n} A_k\right) = P(\tilde{A}).$$

Поскольку верхний и нижний пределы совпадают, все неравенства превращаются в равенства:

$$\Longrightarrow \underline{\lim_{k\to\infty}} P(A_k) = \overline{\lim_{k\to\infty}} P(A_k) = P(\lim_{n\to\infty} A_k).$$

Примеры вероятностных пространств. Дискретное вероятностное пространство.

Опр.: (Ω, F, P) – дискретное вероятностное пространство, если $\Omega = \{\omega_1, \omega_2, \dots, \omega_n, \dots\}$ - конечно или счётно, F – σ-алгебра всех подмножеств множества Ω,

$$\forall\{\omega_j\} \subset F\colon\ P(\{\omega_j\}) = p_j$$

$$j = 1, \dots, \infty$$

$$\sum_{j=1}^{\infty} p_j = 1.$$

Вероятность любого события:

$$P(A) = \sum_{j\colon \omega_j \in A} p_j.$$

Сколько событий в σ-алгебре F:

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_N\}$$

Посчитаем число способов выбрать одно событие из N, затем число способов выбрать два события из N и т.д., плюс пустое событие:

$$C_N^{\mathbf{0}} + C_N^{\mathbf{1}} + C_N^{\mathbf{2}} + \dots + C_N^{\mathbf{N}} = 2^{\boldsymbol{n}}.$$

Вспомним бином Ньютона:

$$(a+b)^n = \sum_{j=0}^{n} C_N^j a^j b^{\boldsymbol{n-j}}\Bigg|_{a=b} = \sum_{j=0}^{n} C_N^j = (1+1)^n = 2^n.$$

Если Ω счётно, N=∞

$$\omega_1, \omega_2, \dots, \omega_n, \dots$$

Чтобы задать событие, необходимо указать, какие элементарные исходы его влекут. Будем присваивать 1 тем исходам, которые включены в событие, 0 – другим. Тогда любое событие – бесконечный ряд из 0 и 1.

Множество бесконечных упорядоченных рядов из 0 и 1 эквивалентно множеству всех двоичных дробей на отрезке [0,1]. Тогда множество всех подмножеств Ω эквивалентно множеству точек на отрезке [0,1] – множество мощности континуум, его нельзя пересчитать.

Проверка определения дискретного вероятностного пространства с помощью пяти аксиом вероятности:

Первая аксиома вероятности выполняется, потому что вероятность задана для любого подмножества множества Ω. Множество всех подмножеств есть σ-алгебра, потому что пределы - это тоже события. Третья аксиома выполняется: $\sum_{j=1}^{\infty} p_j = 1$. Вторая аксиома тоже выполняется, вероятность любого подмножества меньше либо равна единице и больше либо равна нулю, потому что в ряде p_j, который нужно просуммировать, слагаемых меньше либо равно чем в достоверном событии. Первые три аксиомы очевидны.

Докажем пятую аксиому: счетная аддитивность.

$$P\left(\sum_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} P(A_j)$$

Событие A_j может состоять из бесконечного числа элементарных исходов:

$$\sum_{\omega_k \in A} p_k = \sum_{j=1}^{\infty} \sum_{l, \omega_l \in A_j} p_l.$$

Лемма о суммировании по блокам:

Пусть I = {1,2,3,…) = ℕ множество натуральных чисел

$I_\alpha,\ \alpha = 1,2,3, \dots, n, \dots$ - бесконечное число подмножеств, каждое подмножество может содержать бесконечное число элементов.

$$I_\propto \subset I\colon\ I_\propto \cap I_\beta = \emptyset$$

$$\sum_{\propto=1}^{\infty} I_\propto = I.$$

Пусть, кроме того:

ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ МГУ ИМЕНИ М.В. ЛОМОНОСОВА
teach-in ЛЕКЦИИ УЧЕНЫХ МГУ

$$\sum_{k=1}^{\infty} C_k = S - \text{абсолютно сходится} \Longrightarrow \text{ можно изменять порядок суммирования.}$$

$$S_\propto = \sum_{j \in I_\alpha} C_j\,, \qquad \alpha = 1,2,\ldots$$

Тогда:

$$\sum_{\propto=1}^{\infty} S_\propto = S - \text{абсолютно сходится}$$

$$\sum_{\propto=1}^{\infty} \sum_{j \in I_\alpha} C_j = \sum_{l=1}^{\infty} C_l$$

Это означает, что ряд можно суммировать по блокам. Как раз то, что нужно для доказательства пятой аксиомы.

Доказательство леммы было в курсе теории рядов.

Замечание 1:

$$\{x_i\}\,, \{y_j\} - \text{бесконечные наборы чисел}$$

$$\sum_i \sum_j x_i y_j = \sum_i x_i \sum_j y_j$$

Однако, если все $x_i = 0,$ а все $y_j = 1$, возникает парадокс:

$$\sum_i \sum_j 0 = \sum_i 0 \sum_j 1,$$

$$0 = 0 \cdot \infty.$$

Выполняется ли равенство: по Лебегу $\infty = \frac{1}{0} \Longrightarrow 0 \cdot \infty \equiv 0$ – по определению.

Замечание 2: Абсолютная сходимость важна

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \cdots - \text{сходящийся ряд}$$

Однако если сначала будем суммировать положительные числа, получим ∞. Т.е. для рядов, которые не являются абсолютно сходящимися, сумма зависит от способа суммирования.

Замечание 3: Обратное не верно, если есть сходимость по блокам, то не обязательно ряд абсолютно сходится.

Замечание 4: Обратное верно при $C_j \geq 0$.

Например, для вероятности, т.к. все слагаемые положительны.

Условное вероятностное пространство

(Ω, F, P) – исходное вероятностное пространство.

Известно, что произошло событие B, значит, в поставленном эксперименте не могут появиться элементарные исходы, которые не содержатся в B. Разумно уменьшить множество всех элементарных исходов до B.

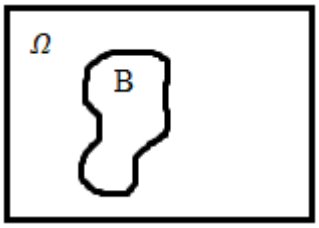


Рис. 2.3. Множество элементарных исходов, уменьшенное до B.

$$\Omega \to B$$

$$A \in F \Longrightarrow AB \in \{F_B\colon\ BA, A \in F\}.$$

Если A пробегает все значения из F, то AB тоже будет содержаться в σ-алгебре.

Необходимо задать вероятность:

$$P(B) > 0$$

$$P(A|B) = \frac{P(AB)}{P(B)}$$

$P(A|B)$- вероятность события A, при условии выполнения события B.

Вероятностное пространство с условной вероятностью: $\big(B, F_B, P(\cdot\,|B)\big)$.

$$P(A|\Omega) = \frac{P-(A\Omega)}{P(\Omega)},$$

$$P(A|B) = \frac{P(AB)}{P(B)}.$$

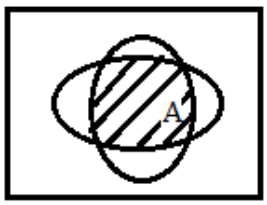


Рис. 2.4. Условная вероятность.

Проверка определения условного вероятностного пространства с помощью пяти аксиом вероятности:

2) $0 \le P(A|B) \le 1$,

т.к. $P(AB) \ge 0,\ P(B) \ge 0$,

$P(AB) \leq P(B)$ – по свойству монотонности вероятности.

3) $P(B|B) = 1 = \frac{P(BB)}{P(B)}$

4) $A = A_1 + A_2$,

$P(A_1 + A_2|B) = \frac{P((A_1+A_2)B)}{P(B)} = \frac{P(A_1B)+P(A_2B)}{P(B)} = P(A_1|B) + P(A_2|B)$.

5) …

Для условной вероятности можно получить полезные формулы:

1) Теорема умножения вероятности:

$$P(A|B) = \frac{P(AB)}{P(B)} \Longrightarrow P(AB) = P(A|B)P(B). \quad (2.1)$$

2) Формула полной вероятности:

Опр.: События $B_1, B_2, \dots, B_n, \dots$ называются полной группой попарно несовместных событий, если:

- $B_i \cap B_j = \emptyset$,
- $\sum B_j = \Omega$.

$$A = A\Omega = A\sum_{j=1}^{\infty} B_j = \sum_{j=1}^{\infty} AB_j,$$

$$P(A) = \sum_{j=1}^{\infty} P(AB_j) = \sum_{j=1}^{\infty} P\left(A\middle|B_j\right)P\left(B_j\right). \quad (2.2)$$

3) Формула Байеса:

$$P(AB) = P(A|B)P(B),$$

$$P(AB) = P(B|A)P(A),$$

$$P(A|B)P(B) = P(B|A)P(A).$$

$B_1, B_2, \dots, B_j, \dots$ - полная группа попарно несовместных событий,

$$P\left(B_j\middle|A\right)P(A) = P\left(A\middle|B_j\right)P\left(B_j\right) \Longrightarrow P\left(B_j\middle|A\right) = \frac{P\left(A\middle|B_j\right)P(B_j)}{P(A)} = \frac{P\left(A\middle|B_j\right)P(B_j)}{\sum_{k=1}^{\infty} P\left(A\middle|B_j\right)P(B_j)}. \quad (2.3)$$

Пример:

А – “студент болен”, В - “студент не сдал экзамен”.

Пусть P(A) = 0,1; P(B) = 0,3.

$$A \subset B.$$

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{0{,}1}{0{,}3} = \frac{1}{3}.$$

$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{0{,}1}{0{,}1} = 1.$$

$$P(\bar{B}|\bar{A}) = \frac{P(\bar{A}\bar{B})}{1 - P(A)} = \frac{P(\bar{B})}{1 - P(A)} = \frac{1 - 0{,}3}{1 - 0{,}1} = \frac{7}{9}.$$

Независимость событий

Опр.: События A и B – независимы, если P(AB) = P(A)P(B).

Свойства:

1) A, B – независимы $\Longrightarrow P(A|B) = \frac{P(AB)}{P(B)} = \frac{\mathrm{P(A)P(B)}}{\mathrm{P(B)}} = \mathrm{P(A)}$.

2) A, B, P(B) = 0 – независимы

$$0 \le P(AB) \le P(B) = 0 \Longrightarrow$$
$$\Longrightarrow 0 = P(AB) = P(A)P(B) = 0.$$

3) A, Ω – независимы

$$P(A\Omega) = P(A) = P(A)P(\Omega).$$

4) A, B – независимые $\Longrightarrow$ A,$\bar{B}$; $\bar{A}, \bar{B}$; B, $\bar{A}$ – независимы.

$$A = A(B + \bar{B}),$$
$$P(A) = P(AB) + P(A\bar{B}),$$
$$\Longrightarrow P(A) - P(A)P(B) = P(A\bar{B}),$$
$$\Longrightarrow P(A)\big(1 - P(B)\big) = P(A\bar{B}),$$
$$\Longrightarrow P(A)P(\bar{B}) = P(A\bar{B}).$$

Аналогично и для случая $B, \bar{A}$.

Опр.: События $A_1, A_2, \dots, A_n$ – независимые в совокупности, если

$$\forall k = 2,3,4, \dots, n$$
$$\forall i_1, i_2, \dots, i_k \in \{1, \dots, n\}$$

Выполнено:

$$P\big(A_{i_1} A_{i_2} \dots A_{i_k}\big) = \prod_{k=1}^{n} P\big(A_{i_k}\big).$$

Такое определение предлагается, чтобы избежать попадания события с нулевой вероятностью в набор событий.

ЛЕКЦИЯ 3. РАСПРЕДЕЛЕНИЯ ТЕОРИИ ВЕРОЯТНОСТЕЙ

Продолжим тему прошлой лекции.

Из независимости в совокупности $\Longrightarrow$ попарная независимость. Обратное неверно.

Рассмотрим тетраэдр Бернштейна (вроде игральной кости):

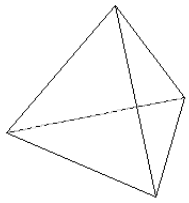


Рис. 3.1. Тетраэдр Бернштейна.

ω_1- красный

ω_2- зеленый

ω_3- синий

ω_4- красный + зеленый + синий

$$P(\omega_j) = \frac{1}{4} - \text{вероятность выпадения одной из граней.}$$

$$\{\text{к}\} = \{\omega_1 + \omega_4\},$$
$$\{\text{з}\} = \{\omega_2 + \omega_4\},$$
$$\{\text{с}\} = \{\omega_3 + \omega_4\},$$
$$\{\text{к}c\} = \{\omega_4\},$$
$$\{c\text{з}\} = \{\omega_4\}.$$

$$P(\text{к}) = P(\text{з}) = \frac{1}{2},$$

$$P(\text{к} \cap \text{з}) = \frac{1}{4} = P(\text{к})P(\text{з}) - \text{попарно независимы,}$$

$$P(\text{к} \cap \text{з} \cap \text{с}) = \frac{1}{4} \neq P(\text{к})P(\text{з})P(\text{с}) = \frac{1}{8} - \text{не независимы в совокупности.}$$

События А и В независимы, когда:

$$P(A|B) = \frac{P(AB)}{P(B)} = P(A) = \frac{P(A\Omega)}{P(\Omega)}.$$

Последовательность независимых испытаний

Пусть (Ω, F, P)-дискретное вероятностное пространство, и $\Omega = \{\omega_1, \omega_2, \dots\}$, тогда σ-алгебра его подмножеств – это множество всех подмножеств, а вероятность можно задать для каждого элементарного исхода.

Проводим эксперимент несколько раз, и эксперимент заканчивается тогда, когда проведено n испытаний. Выпишем исходы всех испытаний:

$\{\langle \omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_n} \rangle,\ \omega_{i_k} \in \Omega\} = \Omega^n = X_{i=1}^{n} \Omega$ – прямое произведение n экземпляров Ω.

Мы получили новый эксперимент со случайным исходом. Этот эксперимент устроен следующим образом: n раз повторяем исходный эксперимент, получаем цепочку из n исходов, каждую такую цепочку считаем новым элементарным исходом. Множество всех таких элементарных исходов называется кратным прямым произведением или n степенью исходного пространства элементарных исходов.

Пространство элементарных исходов Ω конечно или счётно, $\langle \omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_n} \rangle$ - конечное число счетных множеств, значит, пространство Ω^n это тоже конечное или счетное множество.

F^n – σ-алгебра всех подмножеств Ω^n:

$$\{\langle A_{i_1}, A_{i_2}, \dots, A_{i_n} \rangle, \qquad A_{i_k} \in F\} = F^n.$$

Осталось задать вероятность:

$$P_n(\langle \omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_n} \rangle) = \prod_{k=1}^{n} P(\{\omega_{i_n}\}).$$

Опр.: Вероятностное пространство $(\Omega^n, F^n, P_n(\cdot))$ называется последовательностью независимых испытаний.

Пример:

$$n = 2\,,$$

$$(A, B) - \text{независимы.}$$

$$\{A\} = \{\omega_{i_1} \in A, \omega_{i_2} \in \Omega\},$$

$$\{B\} = \{\omega_{i_2} \in B, \omega_{i_1} \in \Omega\}.$$

События $\{A\}, \{B\}$ называются цилиндрическими (Рис. 3.2).

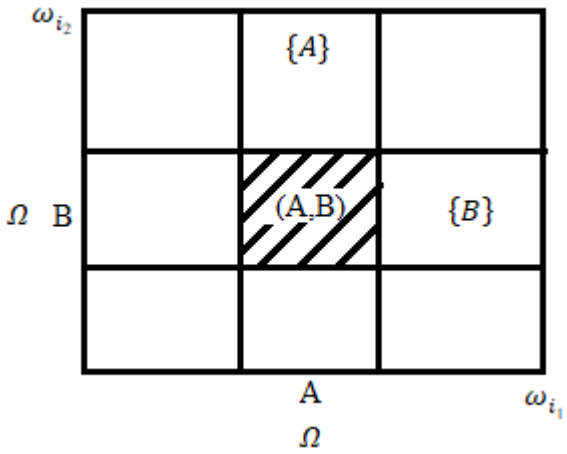


Рис. 3.2. Цилиндрические события.

Нужно просуммировать все элементарные исходы, которые влекут (A,B), а потом вычислить вероятность события $\{A\}$ отдельно, $\{B\}$ отдельно, их перемножить, и в силу леммы о суммировании по блокам мы убедимся, что в этой ситуации P(A,B)=P(A)P(B). (Формально сделать самостоятельно).

Схема Бернулли

Схема Бернулли – последовательность независимых испытаний:

$$\left(\Omega^n, F^n, P_n(\cdot)\right),$$

$$\Omega = \{A, \bar{A}\},$$

$$\text{"A"} - \text{"успех"},$$

$$\text{"}\bar{A}\text{"} - \text{"неудача"}.$$

$$\langle \omega_{i_1}, \omega_{i_2}, \dots, \omega_{i_n} \rangle = \{\text{у}, \text{у}, \text{н}, \dots, \text{н}, \dots, \text{у}, \text{н}\}.$$

$$P(A) = p, \qquad P(\bar{A}) = q = 1 - p.$$

$$P(\langle \text{у}, \text{у}, \text{н}, \dots, \text{н}, \dots, \text{у}, \text{н} \rangle) = p^k q^{n-k},$$

$$k - \text{число успехов из } n \text{ испытаний.}$$

$$P\{\text{из } n \text{ испытаний получить } k \text{ успехов в любом порядке}\} = \mathrm{C}_n^k p^k q^{n-k} = p_n(k).$$

$$p_n(k) = \mathrm{C}_n^k p^k q^{n-k} - \text{распределение Бернулли}$$

$$k = 0,1, \dots, n$$

Распределение Паскаля

Сколько испытаний необходимо провести, для получения n успехов?

Пусть число испытаний n+k, из них n успехов и результат последнего испытания – успех.

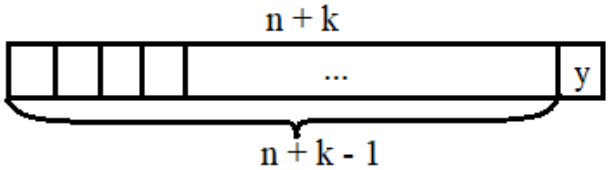


Рис. 3.3. Схема результатов эксперимента.

$$P(\text{для } n \text{ успехов потребовалось } n+k \text{ испытаний}) = \mathrm{C}_{n+k-1}^{n-1} p^{n-1} q^k p =$$
$$= \mathrm{C}_{n+k-1}^{n-1} p^n q^k,$$

$k = 0,1, \ldots -$ распределение Паскаля (отрицательно биномиальное).

$$\mathrm{C}_n^k = \frac{n(n-1)(n-2)\ldots(n-k+1)}{k!},$$

$$\mathrm{C}_{(-n)}^k = \frac{(-n)(-n-1)(-n-2)\ldots(-n-k+1)}{k!},$$

$\mathrm{C}_{(-n)}^k p^n q^k = \mathrm{C}_{n+k-1}^k p^n q^k -$ поэтому называют отрицательно биномиальным.

Полиномиальное распределение

$$\Omega = \{A_1, A_2, \ldots, A_r\}.$$

Цепочка при n кратном повторении этого эксперимента будет состоять из n элементов, на каждом месте из этой из этой цепочки может стоять любой из r исходов.

$$P\big(\langle \omega_{i_1}, \omega_{i_2}, \ldots, \omega_{i_n} \rangle\big) = P_1^{k_1} P_2^{k_2} \ldots P_r^{k_r}, \qquad \text{где}$$
$$P_j = P(A_j), \qquad j = 1, \ldots, r,$$
$$k_j - \text{число исходов } A_j,$$
$$\sum_{j=1}^{r} k_j = n.$$

Вероятность получить в серии из n испытаний k_1 исходов первого сорта, k_2 исходов второго сорта и т.д.:

$$P_n^{(k_1,k_2,\ldots,k_r)} = \frac{n!}{k_1!\,k_2!\ldots k_r!} P_1^{k_1} P_2^{k_2} \ldots P_r^{k_r} - \text{полиномиальное распределение.}$$

Число цепочек, у которых задан состав, и меняется только порядок следования, рассчитать можно так: всего перестановок n элементов – n!, но при этом, если переставлять местами события первого сорта, то это приведет к одному и тому же элементарному исходу, поэтому делим это все на $k_1!$ (число перестановок исходов первого сорта), и т.д..

Формула Пуассона

$$P_n(k) = \begin{cases} C_n^k p^k q^{n-k}, & k = 0,1, \ldots, n \\ 0, & \text{иначе} \end{cases}.$$

$$n \to \infty, \qquad p \to 0,$$

$$p = \frac{\lambda}{n} + o\left(\frac{1}{n}\right).$$

$$P_n(k) = \frac{n(n-1)\dots(n-k+1)}{k!}\left(\frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right)^k \left(1 - \frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right)^{n-k} =$$

$$= \frac{1}{k!}\frac{n}{n}\frac{n-1}{n}\dots\frac{n-k}{n}\left(\lambda + \frac{o\left(\frac{1}{n}\right)}{\frac{1}{n}}\right)^k \left(1 - \frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right)^{-k} \left(1 - \frac{\lambda}{n} + o\left(\frac{1}{n}\right)\right)^n \to$$

$$\to \frac{\lambda^k}{k!}e^{-\lambda}, \qquad k = 0,1,2\dots$$

$p(k) = \frac{\lambda^k}{k!}e^{-\lambda},\ k = 0,1,2\dots$ - распределение Пуассона. Аппроксимирует биномиальное распределение для редких событий.

Если вероятность события близка к 1, то можно использовать распределение Пуассона, поменяв местами успех и неудачу, ведь вероятность неудачи мала.

Теоремы Муавра-Лапласа

n – испытаний, k – успехов, p – вероятность успеха (не близка к 0 или 1).

Локальная теорема:

$$\text{Пусть } x_{n,k} = \frac{k - np}{\sqrt{npq}}, \qquad \text{тогда } \forall a, b - \text{фикс.}$$

$$\lim_{n\to\infty}\left(\sup_{a \le x_{n,k} \le b}\left|\frac{P_n(k)}{\frac{1}{\sqrt{2\pi npq}}e^{-\frac{x_{n,k}{}^2}{2}}} - 1\right|\right) = 0.$$

Интегральная теорема:

$$\lim_{n\to\infty}\left(\sup_{-T \le a < b \le T}\left|p\left(a \le \frac{k - np}{\sqrt{npq}} \le b\right) - \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{z^2}{2}}dz\right|\right) = 0.$$

На практике:

Вероятность получить из n испытаний число успехов k: $m_2 < k < m_1$.

$$P_n(m_2 < k < m_1) = P\left(\frac{m_2 - np}{\sqrt{npq}} \le \frac{k - np}{\sqrt{npq}} \le \frac{m_1 - np}{\sqrt{npq}}\right) \cong$$

$$\cong \frac{1}{\sqrt{2\pi}} \int\limits_{\frac{m_2-np}{\sqrt{npq}}}^{\frac{m_1-np}{\sqrt{npq}}} e^{-\frac{z^2}{2}} dz.$$

Замечание:

Пусть в n испытаниях k раз произошло событие A. Частота A = $\frac{k}{n}$. Попробуем найти вероятность того, что частота отличается от вероятности успеха меньше чем на ε:

$$P\left(\left|\frac{k}{n} - p\right| \le \varepsilon\right)$$

$$P\left(-\varepsilon \le \frac{k}{n} - p \le \varepsilon\right) = P\left(-\frac{\varepsilon\sqrt{n}}{\sqrt{pq}} \le \frac{k-np}{\sqrt{n}\sqrt{pq}} \le \frac{\varepsilon\sqrt{n}}{\sqrt{pq}}\right) \xrightarrow[n\to\infty]{} \frac{1}{\sqrt{2\pi}} \int\limits_{-\frac{\varepsilon\sqrt{n}}{\sqrt{pq}}}^{\frac{\varepsilon\sqrt{n}}{\sqrt{pq}}} e^{-\frac{z^2}{2}} dz$$

$$\xrightarrow[n\to\infty]{} \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz \xrightarrow[n\to\infty]{} 1, \qquad \forall\varepsilon.$$

Модель «случайные величины»

$$(\Omega, F, P) \Longrightarrow \forall A \subset F \text{ можно посчитать } P(A).$$

Опр.: Случайной величиной называется функция $\xi(\cdot): \Omega \to R_1$(числовая ось) и такая, что $\{\text{все } \omega: \xi(\omega) < x\} \in F \quad \forall x \in R_1$. Последнее свойство – измеримость.

Опр.: Функция $F_\xi(x) = P(\xi < x)$ – функция распределения случайной величины.

Пример неизмеримой функции:

$$\Omega = \{(x, y) \in R_2\}, \qquad R_2 - \text{числовая плоскость.}$$

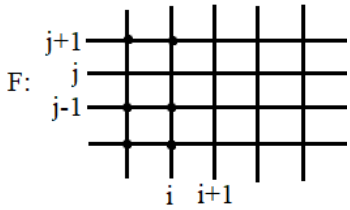


Рис. 3.4. Множество, разбитое на единичные квадраты.

$$j = 0, \pm1, \pm2, \ldots \quad i = 0, \pm1, \pm2, \ldots$$

Квадрат: $A_{ij} = \{(i, j), (i+1, j), (i+1, j+1), (i, j+1)\}$,

$$P(A_{ij}) = p_{ij}.$$

ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ МГУ ИМЕНИ М.В. ЛОМОНОСОВА
teach-in
ЛЕКЦИИ УЧЕНЫХ МГУ

Создали σ-алгебру, состоящую из множеств которые можно разбить на единичные квадраты, и вероятность только таких множеств можно посчитать. Хотя элементарные исходы это любая точка на числовой оси.

$$\xi(\omega) = \xi(x, y) = x + y,$$

$$P(x + y < z) - \text{посчитать не удастся.}$$

Потому что это множество не состоит из целочисленных квадратов, квадратов с целочисленными вершинами. Заданная функция не измерима относительно той σ-алгебры, которую мы построили. Это несколько искусственный пример, потому что от этой неизмеримости легко избавиться, считая, что A_{ij} - это элементарное событие, тогда множество всех подмножеств будет задано.

$$P(y \leq \xi < x) = F_\xi(x) - F_\xi(y).$$

Можно посчитать что-то типа длины, теперь длина отрезка от x и y - это $F_\xi(x) - F_\xi(y)$. Если вместо длины использовать $F_\xi(x)$ и $F_\xi(y)$, можно теперь посчитать вероятность попадания случайной величины в отрезок, а стало быть, в любое борелевское множество.

ЛЕКЦИЯ 4. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ

Теория случайных величин

$$(\Omega, F, P) - \text{вероятностное пространство.}$$

Опр.: Случайная величина $\xi(\cdot)$- это функция, задана $\forall\omega \in \Omega$ и принимает числовые значения $\xi(\omega) \in R_1$, причем $\forall x \in R_1\ \{\omega: \xi(\omega) < x\} \in F$.

Все ω, $\xi(\omega)$ которых попадают в эту полупрямую, должны попадать в алгебру F. Если вместо полупрямой взять ее дополнение от x до ∞, то оно тоже должно принадлежать алгебре F, т.е. $\{\omega: \xi(\omega) > x\} \in F$. Любые пересечения этих полупрямых тоже будут таковы, что ω принадлежат алгебре F. Поскольку для любых элементов из F задана вероятность, то она переносится на все борелевские множества.

$$P(\xi < x) - \text{определена.}$$

Опр.: Функцией распределения случайной величины ξ называется функция

$$\forall x \in R_1: \quad F_\xi(x) = P(\xi < x).$$

Свойства функции распределения:

1) $0 \le F_\xi(x) \le 1$ и монотонно неубывает.

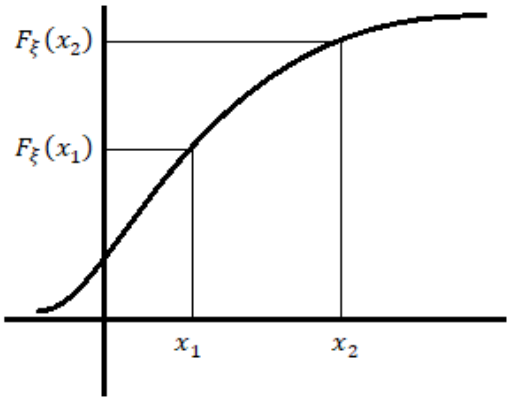


Рис. 4.1. Функция распределения.

Доказательство:

$$F_\xi(x_2) = P(\xi < x_2) = P(\{\xi < x_1\} + \{x_1 \le \xi < x_2\}) =$$

$$= P(\xi < x_1) + P(x_1 \le \xi < x_2).$$

$$F_\xi(x_2) - F_\xi(x_1) = P(x_1 \le \xi < x_2) \ge 0, \qquad x_1 < x_2.$$

2) Непрерывность слева $\forall x_0 \in R_1$.

$$\lim_{x_n \uparrow x_0} F_\xi(x_n) = F_\xi(x_0).$$

ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ МГУ ИМЕНИ М.В. ЛОМОНОСОВА
teach-in ЛЕКЦИИ УЧЕНЫХ МГУ

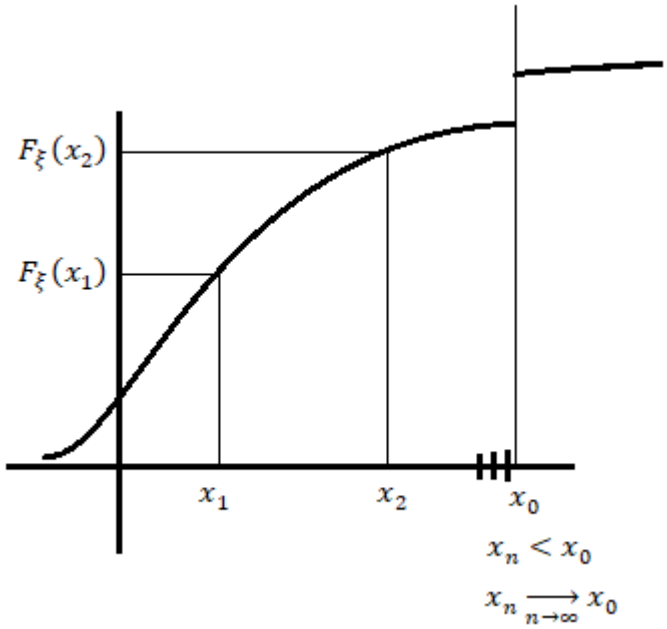


Рис. 4.2. Непрерывность слева.

Доказательство:

$$\lim_{x_n \uparrow x_0} \Big((F_\xi(x_0) - F_\xi(x_n) \Big) = \lim_{x_n \uparrow x_0} P(x_n \le \xi < x_0)$$

Рассмотрим последовательность событий, которые стоят в аргументе нашей вероятности. x_0 не входит ни в один из интервалов. Точка, лежащая правее x_0, не входит ни в один из интервалов. А точка, лежащая левее x_0 (строго меньше чем x_0), рано или поздно не войдёт ни в один из них, потому что x_n стремится к x_0. Начиная с некоторого номера, ни один элемент не войдет ни в один интервал. Значит, все точки либо не входят ни в одно событие, либо входят в конечное число этих событий. Получается:

$$x_n \xrightarrow[n\to\infty]{} \{\xi \in \emptyset\} = \emptyset,$$

$$\lim_{x_n \uparrow x_0} \Big((F_\xi(x_0) - F_\xi(x_n) \Big) = \lim_{x_n \uparrow x_0} P(x_n \le \xi < x_0) = P(\lim_{x_n \uparrow x_0} \{x_n \le \xi < x_0\}) = P(\emptyset) = 0.$$

3)

$$\forall x_0 \in R_1$$

$$\lim_{x_n \downarrow x_0} \Big((F_\xi(x_n) - F_\xi(x_0) \Big) = P(\xi = x_0).$$

Если функция распределения непрерывна в точке x_0, тогда правое предельное значение равно левому, и в точках непрерывности функции распределения вероятность того что ξ примет значение x_0 равна нулю. Но если правое предельное значение отлично от

ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ МГУ ИМЕНИ М.В. ЛОМОНОСОВА
teach-in
ЛЕКЦИИ УЧЕНЫХ МГУ

левого предельного значения, тогда в точке x_0 есть скачок, и величина этого скачка функции распределения есть вероятность того, что ξ=x_0.

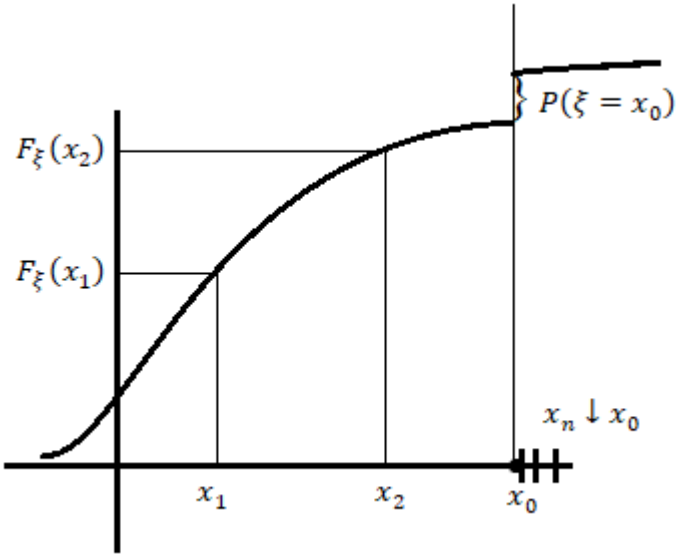


Рис. 4.3. Скачок функции распределения.

Доказательство:

$$\lim_{x_n \downarrow x_0} \left((F_\xi(x_n) - F_\xi(x_0)) \right) = \lim_{x_n \downarrow x_0} P(x_0 \le \xi < x_n) - \text{по пункту 2.}$$

В отличие от предыдущего случая x_0 входит во все интервалы. Ни одна из других точек в этот интервал не входит. Точки, которые лежат левее x_0, не входят. Точки, которые лежат правее x_0, не будут входить в эти интервалы, начиная с некоторого номера в силу сходимости x_n к x_0. Таким образом:

$$x_0 \xrightarrow[n\to\infty]{} \{\xi = x_0\},$$

$$\lim_{x_n \downarrow x_0} \left((F_\xi(x_n) - F_\xi(x_0)) \right) = \lim_{x_n \downarrow x_0} P(x_0 \le \xi < x_n) = P\left(\lim_{x_n \downarrow x_0} P\{x_0 \le \xi < x_n\} \right) = P(\xi = x_0)$$

4)

$$P(x_1 \le \xi \le x_2) = P(\{x_1 \le \xi < x_2\} + \{\xi = x_2\}) =$$
$$= F_\xi(x_2) - F_\xi(x_1) + F_\xi(x_2 + 0) - F_\xi(x_2) = F_\xi(x_2 + 0) - F_\xi(x_1).$$
$$P(x_1 < \xi < x_2) = F_\xi(x_2) - F_\xi(x_1 + 0) - \text{аналогично.}$$

5)

$$F_\xi(+\infty) = \lim_{x\to+\infty} F_\xi(x),$$
$$F_\xi(-\infty) = \lim_{x\to-\infty} F_\xi(x).$$
$$\lim_{x\to+\infty} F_\xi(x) = \lim_{x\to+\infty} \mathrm{P}(\xi < \mathrm{x})$$

Рано или поздно любая точка будет принадлежать множеству. При $x \to +\infty$:

$$\xi \to (-\infty < \xi < +\infty) = \Omega.$$

$$\lim_{x\to+\infty} F_\xi(x) = \lim_{x\to+\infty} \mathrm{P}(\xi < \mathrm{x}) = P\left(\lim_{x\to+\infty}(\xi < \mathrm{x})\right) = P(\Omega) = 1.$$

$$\lim_{x\to-\infty} F_\xi(x) = \lim_{x\to-\infty} \mathrm{P}(\xi < \mathrm{x})$$

Рано или поздно ни одна точка не будет принадлежать множеству. При $x \to -\infty$:

$$\xi \to \emptyset.$$

$$\lim_{x\to-\infty} F_\xi(x) = \lim_{x\to-\infty} \mathrm{P}(\xi < \mathrm{x}) = P(\emptyset) = 0.$$

6) Число точек разрыва $F_\xi(\cdot)$ не более чем счетно.

$$x_0 - \text{т. разрыва} \Leftrightarrow F_\xi(x_0 + 0), F_\xi(x_0).$$

Точке разрыва соответствует следующий интервал:

$$\left(F_\xi(x_0), F_\xi(x_0 + 0)\right).$$

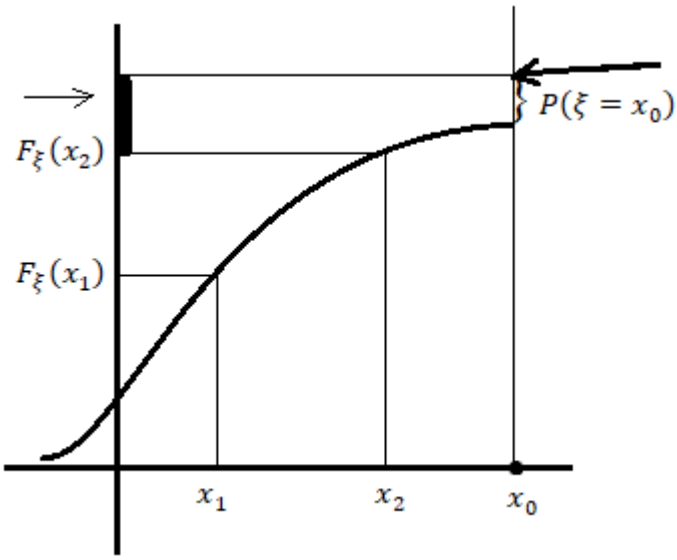


Рис. 4.4. Интервал, соответствующий точке разрыва.

Разным точкам разрыва будут соответствовать непересекающиеся интервалы, потому что функция распределения монотонно не убывает. Интервал не может быть нулевой длины, потому что это точка разрыва. Значит, каждый интервал обязательно содержит рациональную точку:

$$\exists \frac{n}{m}(x) \in \left(F_\xi(x_0), F_\xi(x_0 + 0)\right).$$

Поскольку все рациональные числа можно пересчитать, то можно пересчитать и любое их подмножество, т.е. точки в интервалах. Отсюда следует, что количество точек разрыва счетно.

Классификация случайных величин по виду их функции распределения

Можно ввести три класса случайных величин. И любая случайная может быть представлена суммой трех случайных величин, где каждое слагаемое принадлежит своему классу.

1) Дискретная случайная величина

Опр.: $\xi(\cdot)$ – дискретная случайная величина, если она принимает конечное или счетное число значений.

$$P(\{\xi = x_j\}) = p_j, \qquad j = 1,2,\dots$$

$$\sum_{j=1}^{\infty} p_j = 1.$$

$$(\Omega, F, P) \to \left((x_1, x_2, \dots, x_n, \dots), \bar{F}, \bar{P}\right).$$

В качестве Ω выступает набор чисел, σ-алгебра это множество всех подмножеств, поскольку пространство всех элементарных исходов конечно или счетно. Задана вероятность для каждого элементарного исхода:

$$\bar{P}(\{x_j\}) = p_j.$$

Как определить функцию распределения и построить ее график:

$$F_\xi(x) = P(\xi < x) = \sum_{x_k : x_k < x} p_k.$$

Пример:

$$\xi = \begin{cases} 1, & p = \frac{1}{2} \\ 2, & p = \frac{1}{4} \\ 3, & p = \frac{1}{4} \end{cases}$$

$$F_\xi(x), x \le 1\colon\ P(\xi < x) = 0,$$

$$F_\xi(x), x \le 2\colon\ P(\xi < x) = \frac{1}{2},$$

$$F_\xi(x), x \le 3\colon\ P(\xi < x) = \frac{3}{4},$$

$$F_\xi(x), x > 3\colon\ P(\xi < x) = 1.$$

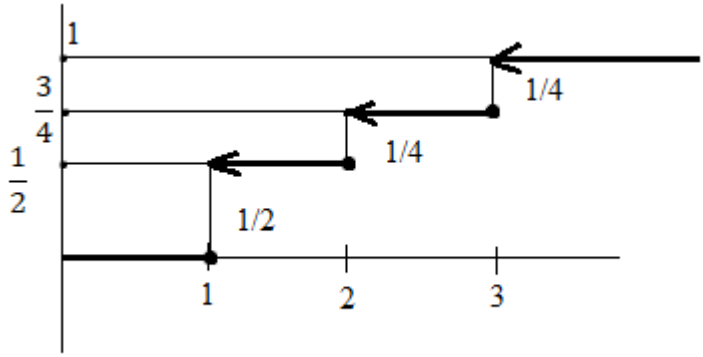


Рис. 4.5. Пример дискретной случайной величины.

2) Абсолютно непрерывная случайная величина

Опр.: ξ - абсолютно непрерывная случайная величина, если:

$$F_\xi(x) = \int_{-\infty}^{x} P_\xi(z)dz, \qquad \text{где}$$

$P_\xi(z) \geq 0 -$ плотность вероятности ξ (кусочно непрерывная функция, имеет не более чем счетное число точек разрыва).

Если x-точка непрерывности $P_\xi(\cdot)$, то $F_\xi(x)$ дифференцируема в точке x, и производная функции распределения есть производная интеграла по верхнему пределу:

$$\frac{dF_\xi(x)}{dx} = P_\xi(x).$$

Для $x \in R_1$ $F_\xi(x) = F_\xi(x+0)$ и $P(\xi = x) = 0$ для $\forall x \in R_1$.

$$P(x < \xi < x + \Delta) = \int_{x}^{x+\Delta} p(z)dz = p(z)\Delta + o(\Delta).$$

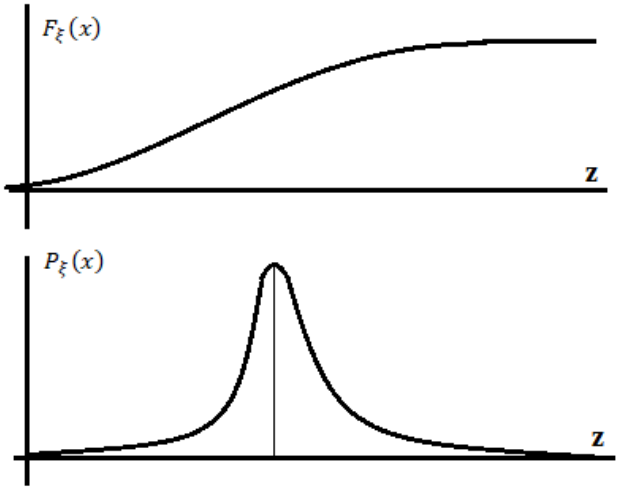


Рис. 4.6. Плотность вероятности и функция распределения.

Производная будет больше там, где скорость роста функции распределения больше. В окрестности выделенной точки наиболее вероятно появление случайной величины.

$$P(a < \xi < b) = \int_a^b p(z)dz - \text{вероятность попадания в интервал.}$$

Если известна вероятность попадания в любой интервал числовой оси, то в силу того что борелевская алгебра, это есть алгебра полученная замыканием множества всех интервалов относительно операций дополнения и пересечения, можно посчитать вероятность попадания в борелевское множество. Тем самым функция распределения абсолютно непрерывно устанавливает взаимно однозначное соответствие:

$$(\Omega, F, P) \to (R_1, B, \bar{P}),$$

B – алгебра борелевских множеств.

$$P(\xi \in B) = \int_B P_\xi(z)\, dz$$

3) Сингулярная случайная величина

Этот тип величин практически никогда не встречается.

Пример:

$$F_\xi(x) - \text{“Канторова лестница”}.$$

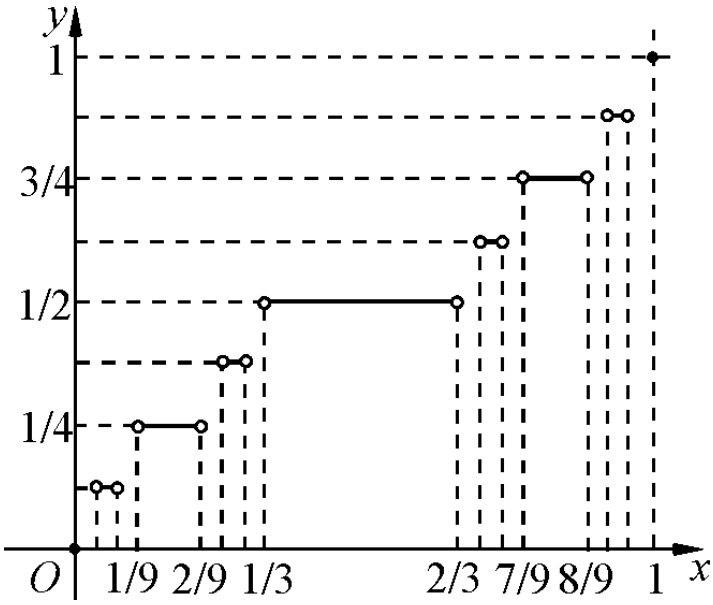


Рис. 4.7. «Канторова лестница».

$$\frac{1}{3}+2\cdot\frac{1}{9}+4\cdot\frac{1}{9\cdot 3}+\cdots+\frac{1}{3}\cdot\left(\frac{2}{3}\right)^{n}+\cdots=\frac{1}{3}\cdot\frac{1}{1-\frac{2}{3}}=1.$$

Справа монотонно убывающая ограниченная снизу последовательность, а слева монотонно возрастающая ограниченная сверху. И слева и справа они должны иметь какие-то пределы, и можно показать, что эти пределы совпадают. То есть, эту функцию можно в оставшихся точках доопределить по непрерывности. Получаем функцию, которая называется «Канторова лестница», которая почти всюду константа и в то же время непрерывна. Заметим, что число точек роста эквивалентно числу точек на отрезке [0,1].

Точки роста – это точки, в троичной записи которых встречаются либо число 0, либо число 2 числа, и 1 не встречается.

Любая случайная может быть представлена суммой трех случайных величин, где каждое слагаемое принадлежит своему классу. F_1, F_2, F_3- дискретная, абсолютно непрерывная и сингулярная случайные величины.

$$\forall\xi:\ F_\xi(x)=k_1F_1(x)+k_2F_2(x)+k_3F_3(x).$$

$$k_j\geq 0,\qquad \sum_{j=1}^{3}k_j=1.$$

Случайные векторы

Опр.: $\vec{\xi}=(\xi_1,\xi_2,\dots,\xi_n)\in R_n$ – случайный вектор, если ξ_j – случайная величина, $j=1,\dots,n$.

$$P(\xi_1(\omega) < x_1, \xi_2(\omega) < x_2, \dots, \xi_n(\omega) < x_n)$$
$$- \text{функция распределения случайного вектора.}$$

Она определена, потому что множества ω, для которых выполнено каждое неравенство, это элементы алгебры F, и пересечение этих множеств также из алгебры F.

$$P(\xi_1(\omega) < x_1, \xi_2(\omega) < x_2, \dots, \xi_n(\omega) < x_n) = F_{\vec{\xi}}(x_1, x_2, \dots, x_n).$$

Свойства $\boldsymbol{F_{\vec{\xi}}(x_1, x_2, \dots, x_n)}$:

$$\vec{\xi} = (\xi_1, \xi_2) \in R_2.$$

1) $F_{\vec{\xi}}(x_1, x_2)$ – не убывает

$$\text{по } x_1 \text{ при } x_2 = const,$$
$$\text{по } x_2 \text{ при } x_1 = const.$$
$$F_{\vec{\xi}}(x_1, x_2) = P(\xi_1 < x_1, \xi_2 < x_2),$$
$$\Delta > 0$$
$$F_{\vec{\xi}}(x_1 + \Delta, x_2) - F_{\vec{\xi}}(x_1, x_2) = P(x_1 \le \xi_1 < x_1 + \Delta, \xi_2 < x_2) \ge 0.$$

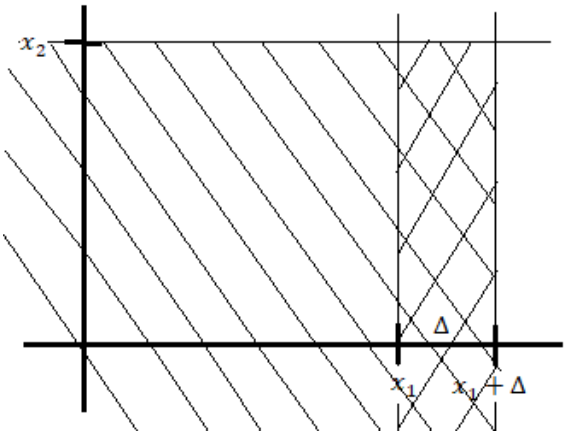


Рис. 4.8. Схема, показывающая, что $F_{\vec{\xi}}(x_1, x_2)$ не убывает.

2) $F_{\vec{\xi}}(x_1, x_2)$ – непрерывна слева по каждой координате.

Доказательство будет аналогично одномерному случаю, только отрезки будут заменены на прямоугольники (Рис. 4.8).

3) $P(x_1 \le \xi_1 < x_2, y_1 \le \xi_2 < y_2) =$
$$= F_{\vec{\xi}}(x_2, y_2) + F_{\vec{\xi}}(x_1, y_1) - F_{\vec{\xi}}(x_1, y_2) - F_{\vec{\xi}}(y_1, x_2).$$

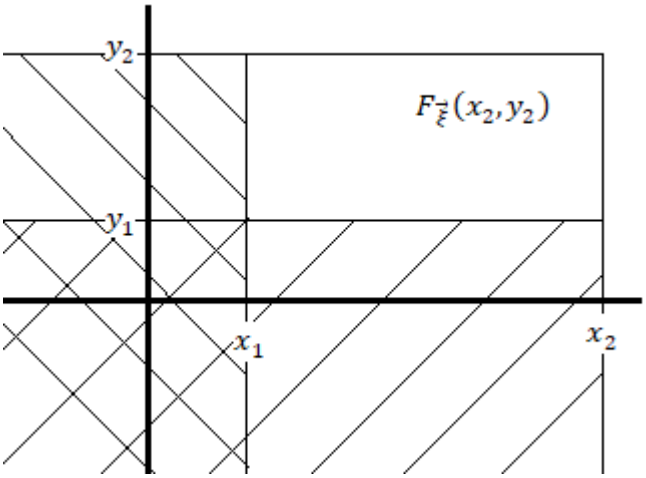


Рис. 4.9. Иллюстрация формулы для $P(x_1 \le \xi_1 < x_2, y_1 \le \xi_2 < y_2)$.

$$\lim_{x_2 \to \infty} F_{\vec{\xi}}(x_1, x_2) = F_{\vec{\xi}}(x_1, +\infty) = F_{\vec{\xi_1}}(x_1),$$

$$F_{\vec{\xi}}(x_1, +\infty) = F_{\vec{\xi_1}}(x_1) = P(\xi_1 < x_1)$$

−маргинальная функция распределения первой координаты вектора ξ

$$F_{\vec{\xi}}(x_1, -\infty) = \lim_{x_2 \to -\infty} F_{\vec{\xi}}(x_1, x_2) = 0.$$

Опр.: Вектор $\vec{\xi}$- называется дискретным случайным вектором, если ξ_j – дискретная случайная величина, $j = 1, \dots, n$.

$$\vec{\xi} = (\xi, \eta),$$

$$\xi \to x_k$$

$$\eta \to y_j$$

$$P(\xi = x_k, \eta = y_j) = P_{ij}, \qquad i, j = 1, \dots, \infty$$

$$\sum_{i,j=1}^{\infty} P_{ij} = 1.$$

Опр.: Вектор $\vec{\xi}$- называется абсолютно непрерывным вектором, если:

$$F_{\vec{\xi}}(x, y) = \int_{-\infty}^{x} dt \int_{-\infty}^{y} dz\, P_{\xi,\eta}(t, z).$$

$$P_{\xi,\eta}(x, y) \ge 0,$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(x, y) dx dy = 1,$$

$$\forall A\colon\ P\left(\vec{\xi} \in A\right) = \int\limits_{A} P_{\vec{\xi}}(\overrightarrow{x})\, d\vec{x}.$$

$$P_{\vec{\xi}}(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}.$$

$$\int\limits_{-\infty}^{\infty} P_{\xi,\eta}\,(x, y)\, dy = P_{\xi}\,(x) - \text{маргинальная плотность вероятности.}$$

ЛЕКЦИЯ 5. НЕЗАВИСИМОСТЬ СЛУЧАЙНЫХ ВЕЛИЧИН

Функции от случайных величин

Пусть ξ – случайная величина, и $\varphi(\cdot)$: $R_1 \to R_1$.

Рассмотрим $\zeta = \varphi(\xi)$, где $\varphi(\cdot)$ - неслучайная функция, а ξ – случайная величина, которая принимает действительные значения. В каком случае ζ будет случайной величиной?

В том случае, когда:

$$\{\omega: \varphi(\xi(\omega)) < x\} \in F.$$

Чтобы это выполнялось, достаточно:

$$\{x: \varphi(x) < y\} \in B.$$

Можно также определить функцию от случайной векторной величины:

$$\varphi(\xi, \eta) = \zeta - \text{случайная величина}$$

Зная распределения случайных величин ξ и η, можем найти функцию распределения случайной величины φ.

$$P\big((\xi, \eta) \in B\big) = \int\limits_{(x,y)\in B} P_{\xi,\eta}\,(x, y)\,dx\,dy.$$

Если $\varphi(\xi, \eta) = \zeta -$ случайная величина, то

$$F_\zeta(z) = P(\zeta < z) = P\big((\xi, \eta) \in \{\varphi(x, y) < z\}\big) = \int\limits_{\varphi(x,y)<z} P_{\xi,\eta}\,(x, y)\,dx\,dy.$$

Пример 1:

Пусть (ξ, η) – двумерный вектор, $P_{\xi,\eta}\,(x, y)$ – плотность вероятности.

$$\zeta = \xi + \eta, \qquad P_\zeta(z) - ?$$

По определению:

$$F_\zeta(z) = P(\zeta < z) = P(\xi + \eta < z) = \int\limits_{x+y<z} P_{\xi,\eta}\,(x, y)\,dx\,dy = \int\limits_{-\infty}^{\infty} dx \int\limits_{-\infty}^{z-x} dy\, P_{\xi,\eta}\,(x, y).$$

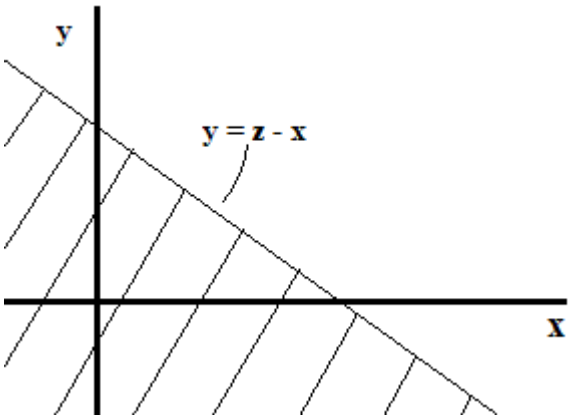


Рис. 5.1. Иллюстрация к примеру 1.

$$P_\zeta(z) = \frac{dF_\zeta(z)}{dz} = \int\limits_{-\infty}^{\infty} dx P_{\xi,\eta}\,(x, z - x)$$

$-$ плотность вероятности случайной величины, равной сумме двух случайных величин.

Пример 2:

Пусть $\vec{\eta} \in R_n$, $\vec{\xi} \in R_n$:

$$\vec{\eta} = A\vec{\xi}, \qquad \exists A^{-1}\colon\ A^{-1}A = AA^{-1} = I.$$

$$P_\xi(x_1, \dots, x_n) - \text{известно}, \qquad P_\eta(y_1, \dots, y_n) - ?$$

Согласно уже известной формуле:

$$P(\vec{\eta} \in B) = \int\limits_{\vec{y} \in B} P_\eta\,(\vec{y})\, d\vec{y}.$$

С другой стороны:

$$P(\vec{\eta} \in B) = P\big(A\vec{\xi} \in B\big) = \int\limits_{A\vec{x} \in B} P_{\vec{\xi}}(\overrightarrow{x})\, d\vec{x} = \begin{bmatrix} \text{Замена переменных} \\ A\vec{x} = \vec{y}, \quad \vec{x} = A^{-1}\vec{y} \end{bmatrix} =$$

$$= \int\limits_{\vec{y} \in B} P_\xi(A^{-1}\vec{y})\, |detA^{-1}| d\vec{y}.$$

Получается:

$$P(\vec{\eta} \in B) = \int\limits_{\vec{y} \in B} P_\eta\,(\vec{y})\, d\vec{y} = \int\limits_{\vec{y} \in B} P_\xi(A^{-1}\vec{y})\, |detA^{-1}| d\vec{y}$$

$$\Longrightarrow P_\eta\,(\vec{y}) = P_\xi(A^{-1}\vec{y}) |detA^{-1}|.$$

ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ МГУ ИМЕНИ М.В. ЛОМОНОСОВА
teach-in
ЛЕКЦИИ УЧЕНЫХ МГУ

Независимость случайных величин

Случайные величины $\xi_1, \xi_2, \dots, \xi_n$ – независимы, если

$$F_{\xi_1,\xi_2,\dots,\xi_n}(x_1, \dots, x_n) = \prod_{j=1}^{n} F_{\xi_j}(x_j).$$

Напомним, что:

$$F_{\xi_1,\xi_2,\dots,\xi_n}(x_1, \dots, x_n) = P(\xi_1 < x_1, \dots, \xi_n < x_n),$$
$$F_{\xi_j}(x_j) = P(\xi_j < x_j).$$

Пусть (ξ, η) – дискретная случайная векторная величина, где ξ, η – независимые:

$$P(\xi = x_i, \eta = y_j) = P(\xi = x_i)P(\eta = y_j).$$

Если $(\xi_1, \xi_2, \dots, \xi_n)$ – абсолютно непрерывная случайная величина, то координаты независимы при условии:

$$P_{\xi_1,\xi_2,\dots,\xi_n}(x_1, \dots, x_n) = \prod_{j=1}^{n} P_{\xi_j}(x_j), \qquad P_j(x_j) - \text{плотность вероятности } \xi_j.$$

Для непрерывных случайных величин (ξ, η), где ξ, η – независимые:

$$P(x_1 \le \xi < x_2, y_1 \le \eta < y_2) = P(x_1 \le \xi < x_2)P(y_1 \le \eta < y_2).$$

Теорема:

ξ, η – независимые случайные величины

$$\tilde{\xi} = \varphi(\xi) - \text{случайная величина},$$
$$\tilde{\eta} = \psi(\eta) - \text{случайная величина}.$$
$$\tilde{\xi}, \tilde{\eta} - \text{независимы}.$$

Любые функции от независимых случайных величин дают независимые случайные величины.

Доказательство (для дискретных случайных величин)**:**

Если ξ – дискретная случайная величина, то известно:

$$P(\xi = x_k) = p_k, \qquad \sum_{k=1}^{\infty} p_k = 1.$$

Значит, $\tilde{\xi} = \varphi(\xi)$ – тоже дискретна:

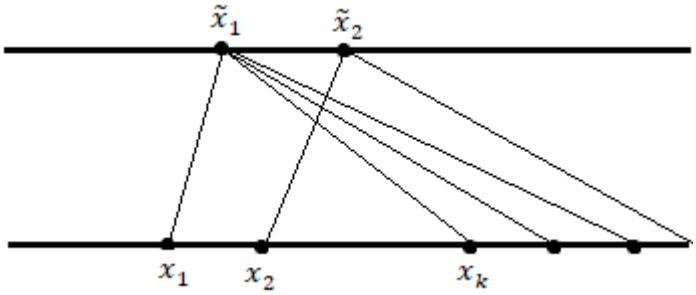


Рис. 5.2 Преобразование x функцией φ.

Сверху изображено значение, которое принимает $\varphi(x_1)$ это $\tilde{x}_1$. Также $\tilde{x}_1$ может принять значение $\varphi(x_k)$. Вообще таких значений x_k, для которых $\varphi(x_k) = \tilde{x}_1$ может быть сколько угодно. Однако если число значений x_k конечно или счетно, то и значений $\varphi(x_k)$ тоже конечно или счетно.

$$P(\tilde{\xi} = \tilde{x}_j) = \sum_{k:\varphi(x_k)=\tilde{x}_j} p_k.$$

Т.е. чтобы найти вероятность события $\tilde{\xi} = \tilde{x}_j$, необходимо просуммировать вероятности всех x_k, для которых значение функции $\varphi(x_k) = \tilde{x}_j$.

Аналогично η – дискретная случайная величина:

$$P(\eta = y_j) = q_j, \qquad \sum_{j=1}^{\infty} q_j = 1.$$

Значит, $\tilde{\eta} = \psi(\eta)$ – дискретная случайная величина:

$$P(\tilde{\eta} = \tilde{y}_k) = \sum_{j:\psi(y_j)=\tilde{y}_k} q_j.$$

Теорема будет доказана, если сможем записать вероятность $P(\tilde{\xi} < \tilde{x}, \tilde{\eta} < \tilde{y})$ в виде произведения:

$$P(\tilde{\xi} < \tilde{x}, \tilde{\eta} < \tilde{y}) = \sum_{\substack{i:\,\varphi(x_i)<\tilde{x} \\ i:\,\psi(y_j)<\tilde{y}}} p_i q_j$$

Получившийся ряд из неотрицательных элементов сходится абсолютно, по лемме о суммировании по блокам можем перегруппировать слагаемые:

$$P(\tilde{\xi} < \tilde{x}, \tilde{\eta} < \tilde{y}) = \sum_{\substack{i:\,\varphi(x_i)<\tilde{x} \\ j:\,\psi(y_j)<\tilde{y}}} p_i q_j = \sum_{i:\,\varphi(x_i)<\tilde{x}} p_i \sum_{j:\,\psi(y_j)<\tilde{y}} q_j = P(\tilde{\xi} < \tilde{x}) P(\tilde{\eta} < \tilde{y}).$$

Условные распределения случайных величин

На вероятностном пространстве (Ω, F, P): $\xi(\omega)$: $\{\xi < x\} \in F$ для $\forall x$

Помимо этого известно, что в эксперименте произошло событие $B \in F$, тогда если $P(B) > 0$, можно рассматривать условную вероятность, заданную на (Ω, F, P).

$$P(\xi < x|B) = \frac{P(\xi < x, B)}{P(B)} = F_\xi(x|B) - \text{функция условного распределения.}$$

В качестве B можно рассмотреть событие:

$$B = \{y \le \eta < y + \Delta\},$$

$$P_{\xi,\eta}(x, y) - \text{плотность совместного распределения.}$$

Можно ли определить условное распределение, при условии, что вероятность события $\mathrm{P(B)} = 0$?

Если:

$$F_\xi(x|B) = \int_{-\infty}^{x} P_\xi(x|B)dx,$$

$$\text{то } P_\xi(x|B) - \text{плотность условной вероятности, } \int_{-\infty}^{\infty} P_\xi(x|B)dx = 1.$$

$$F_\xi(x|y \le \eta < y + \Delta) = \frac{P(\xi < x,\ y \le \eta < y + \Delta)}{P(y \le \eta < y + \Delta)}$$

Если известна плотность совместного распределения, то можно найти маргинальную плотность распределения (см. лекцию 4):

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

$$\Longrightarrow F_\xi(x|y \le \eta < y + \Delta) = \frac{\int_{-\infty}^{x} dt_1 \int_y^{y+\Delta} dt_2 P_{\xi,\eta}\,(t_1, t_2)}{\int_y^{y+\Delta} P_\eta\,(t)\,dt}.$$

$$P_{\xi|\eta}(x|y \le \eta < y + \Delta) = \frac{\int_y^{y+\Delta} P_{\xi,\eta}\,(x|t_2)dt_2}{\int_y^{y+\Delta} P_\eta\,(t_3)\,dt_3}.$$

Получили условную плотность распределения, при том, что событие в условии имеет ненулевую вероятность. Если Δ устремится к нулю, то вероятность условия тоже устремится к нулю. Если точки x и y являются точками непрерывности совместной плотности вероятности, то по теореме о среднем при $\Delta \to 0$:

$$P_{\xi|\eta}(x|y \le \eta < y+\Delta) = \frac{P_{\xi|\eta}(x|y)\Delta + o(\Delta)}{P_\eta(y)\Delta + o(\Delta)} \to \boxed{P_{\xi|\eta}(x|y) = \frac{P_{\xi,\eta}(x,y)}{P_\eta(y)}}.$$

Следствия:

$$P_{\xi,\eta}(x,y) = P_{\xi|\eta}(x|y)P_\eta(y),$$

$$P_\xi(x) = \int_{-\infty}^{\infty} P_{\xi|\eta}(x|y)\,P_\eta(y)dy - \text{аналог формулы полной вероятности.}$$

$$F_{\xi|\eta}(x|y) = \lim_{\Delta\to 0} F_{\xi|\eta}(x|y \le \eta < y+\Delta) = \lim_{\Delta\to 0}\frac{\int_{-\infty}^{x} dt_1 \int_y^{y+\Delta} dt_2 P_{\xi,\eta}(t_1,t_2)}{\int_{-\infty}^{\infty} dt_1 \int_y^{y+\Delta} dt_2 P_{\xi,\eta}(t_1,t_2)} =$$

$$= \frac{\int_{-\infty}^{x} P_{\xi,\eta}(t_1,y)dt_1}{\int_{-\infty}^{\infty} dt_1\, P_{\xi,\eta}(t_1,y)} = \int_{-\infty}^{x} P_{\xi|\eta}(t_1|y)\,dt_1.$$

Числовые характеристики случайных величин

Опр.: Моментом порядка k случайной величины ξ называется

А) для дискретной случайной величины:

$$\sum_{j=1}^{\infty} x_j^k p_j\,, \qquad \text{если } \sum |x_j|^k p_j < \infty \text{ (абс. сх)},$$

Б) для абсолютно непрерывных случайных величин:

$$\int_{-\infty}^{\infty} x^k p_\xi(x)dx\,, \qquad \text{при условии } \int_{-\infty}^{\infty} |x|^k p_\xi(x)dx < \infty \text{ (абс. сх)}.$$

Обобщенная запись:

$$M_\xi^k = \int_{-\infty}^{\infty} x^k dF_\xi(x).$$

Опр.: Момент первого порядка M_ξ называется математическим ожиданием случайной величины ξ.

Пример 1:

Пусть ξ – случайная величина Бернулли (дискретная),

$$P(\xi = k) = C_n^k p^k (1-p)^{n-k}, \qquad k = 0, \dots, n.$$

Свойство абсолютной сходимости ряда здесь выполняется, поскольку он содержит конечное число членов:

$$M_\xi = \sum_{k=0}^{n} k C_n^k p^k q^{n-k} = \sum_{k=0}^{n} p C_n^k \frac{d}{dp} p^k q^{n-k} = p \frac{d}{dp} \sum_{k=0}^{n} C_n^k p^k q^{n-k},$$

тут бином Ньютона

$$\Longrightarrow M_\xi = p \frac{d}{dp} (p+q)^n = pn(p+q)^{n-1}|_{q=1-p} = np \cdot 1 = np.$$

Пример 2:

Пусть ξ – случайная величина Пуассона

$$M_\xi = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \lambda \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \left[\begin{array}{c} \text{замена переменных} \\ k' = k-1, \ \ k = 0,1,2,\dots \end{array} \right] = e^{-\lambda} \lambda e^{\lambda} = \lambda.$$

Сумма оказалась разложением экспоненты.

$$\xi\colon\ P(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \qquad k = 0,1,2,\dots$$

Пример 3:

Пусть ξ – абсолютно непрерывная случайная величина, нормально распределенная (Гауссова).

$$\xi \sim N(\mu, \sigma^2).$$

$$P_{N(\mu,\sigma^2)}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

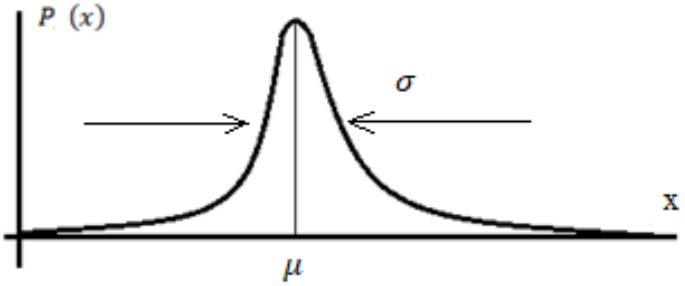


Рис. 5.3 Плотность вероятности.

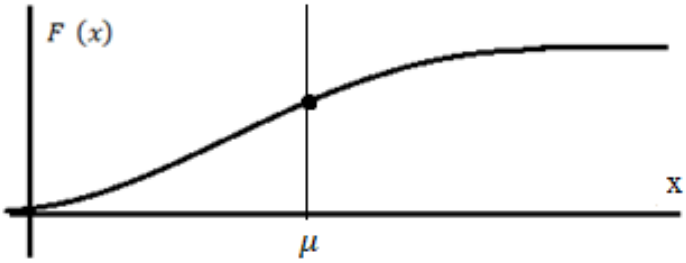


Рис. 5.4 Функция распределения.

$$M_\xi = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{-\infty}^{\infty} (x-\mu+\mu)\, e^{-\frac{(x-\mu)^2}{2\sigma^2}} \frac{dx}{\sqrt{2\pi\sigma^2}} =$$

$$= \int_{-\infty}^{\infty} (x-\mu)\, e^{-\frac{(x-\mu)^2}{2\sigma^2}} \frac{dx}{\sqrt{2\pi\sigma^2}} + \mu \int_{-\infty}^{\infty} P_{N(\mu,\sigma^2)}(x)dx = \mu.$$

$\int_{-\infty}^{\infty} (x-\mu)\, e^{-\frac{(x-\mu)^2}{2\sigma^2}} \frac{dx}{\sqrt{2\pi\sigma^2}} = 0$ – интеграл от нечетной функции в симметричных пределах.

$$\int_{-\infty}^{\infty} P_{N(\mu,\sigma^2)}(x)dx = 1.$$

Теорема о математическом ожидании:

Пусть ξ – случайная величина, и $\varphi(\xi) = \zeta$ – случайная величина, тогда:

$$M_\zeta = M_{\varphi(\xi)} = \int_{-\infty}^{\infty} \varphi(x) dF_\xi(x),$$

и существует тогда и только тогда, когда $\int_{-\infty}^{\infty} |\varphi(x)| dF_\xi(x) < \infty$.

Доказательство (для дискретных случайных величин)**:**

$$P(\xi = x_k) = p_k, \qquad \sum_{k=1}^{\infty} p_k = 1.$$

$$\varphi(\xi) - \text{случайная величина, покажем что:}$$

$$M_{\varphi(\xi)} = \sum_{k=1}^{\infty} \varphi(x_k) p_k,$$

причем $M_{\varphi(\xi)}$ существует тогда и только тогда, когда $\sum_{k=1}^{\infty} |\varphi(x_k)| p_k < \infty$.

Если $\varphi(\xi)$ – случайная величина, то должны быть известны значения, которые она принимает y_j, тогда:

$$P\big(\varphi(\xi) = y_j\big) = \sum_{i:\, \varphi(x_i) = y_j} p_i = q_j.$$

$$M_{\varphi(y)} = \sum_{j=1}^{\infty} y_j q_j$$

Пусть $\sum_{j=1}^{\infty} |y_j| q_j < \infty$, тогда

$$\sum_{j=1}^{\infty}|y_j|q_j = \sum_{j=1}^{\infty}|y_j| \sum_{i:\,\varphi(x_i)=y_j} p_i = \sum_{j=1}^{\infty} \sum_{i:\,\varphi(x_i)=y_j} |\varphi(x_i)|\, p_i = \sum_{j=1}^{\infty} |\varphi(x_i)|\, p_i < \infty.$$

Ряд $\sum_{j=1}^{\infty} |\varphi(x_i)|\, p_i$ абсолютно сходится. Тогда:

$$M_{\varphi(y)} = \sum_{j=1}^{\infty} y_j q_j = \sum_{j=1}^{\infty} y_j \sum_{i:\,\varphi(x_i)=y_j} p_i = \sum_{j=1}^{\infty} \varphi(x_i)\, p_i\,. \blacksquare$$

Для абсолютно непрерывной случайной величины:

$$M_{\varphi(\xi)} = \int_{-\infty}^{\infty} \varphi(x)p(x)dx,$$

причем $M_{\varphi(\xi)}$ существует тогда и только тогда, когда $\int_{-\infty}^{\infty}|\varphi(x)|p(x)dx < \infty$.

Справедливо и для функции многих переменных:

$$\varphi(\xi_1, \xi_2, \dots, \xi_n) = \zeta$$

$$M_\zeta = \int_{R_n} \varphi(x_1, x_2, \dots, x_n)p(x_1, x_2, \dots, x_n)\, d\vec{x},$$

причем M_ζ существует тогда и только тогда, когда $\int_{R_n}|\varphi(\vec{x})|p(\vec{x})\, d\vec{x} < \infty$.

ЛЕКЦИЯ 6. МАТЕМАТИЧЕСКОЕ ОЖИДАНИЕ И ДИСПЕРСИЯ

Числовые характеристики случайных величин

Опр.: Моментом порядка k случайной величины ξ называется

$$M_\xi^k = \begin{cases} \sum\limits_{j=1}^{\infty} x_j^k p_j\,, & \text{если } \xi - \text{дискретная случайная величина} \\ \int\limits_{-\infty}^{\infty} x^k p_\xi(x)dx\,, & \text{если } \xi - \text{абсолютно непрерывная сл. величина} \end{cases},$$

И существует, если соответствующий ряд (интеграл) сходится абсолютно. k=1,2,3…

Обобщенная запись:

$$M_\xi^k = \int\limits_{-\infty}^{\infty} x^k dF_\xi(x).$$

$dF_\xi(x)$ – это приращение функции распределения, вероятность того, что x попадет в малую окрестность. Интеграл это сумма всех вероятностей по всем бесконечно малым отрезкам.

Если $F_\xi(x)$ дифференцируема, то $dF_\xi(x) = p(x)dx$. А если не дифференцируема, то только такие приращения $dF_\xi(x)$ будут неотрицательными, в которые попадают значения дискретной случайной величины, тогда скачок будет уже не бесконечно малым, а равным значению p_j.

Опр.: Момент первого порядка называется математическим ожиданием случайной величины ξ.

$$M_\xi - \text{математическое ожидание } \xi,$$

оно существует если ряд (или интеграл) сходятся абсолютно.

Примеры 1, 2: См. в лекции 5.

Пример 3:

$$p(\xi = const = C) = 1$$
$$\Longrightarrow M_\xi = C.$$

Пример 4:

$$\xi = \pm 2^k \text{ с вероятностью } P_{\pm k} = \frac{1}{2^{k+1}}$$
$$k \neq 0.$$

ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ МГУ ИМЕНИ М.В. ЛОМОНОСОВА
teach-in ЛЕКЦИИ УЧЕНЫХ МГУ

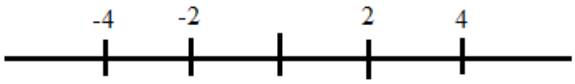


Рис. 6.1. Случайная величина $\xi = \pm 2^k$.

Симметричное распределение, вроде математическое ожидание должно быть равно 0. Но ряд, формально составленный из произведения $x_k p_k$, будет расходящимся, поэтому математического ожидания не существует.

Пример 5:

$$\xi \sim N(\mu, \sigma^2).$$

$$P_{N(\mu,\sigma^2)}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

См. в лекции 5.

Теорема о математическом ожидании: См. в лекции 5.

Следствие:

Пусть M_ξ^k - ∃ для всех k = 1,2,3, …, n, тогда

$$\exists \ M\left(\sum_{k=1}^{n} C_k \xi^k\right) = \sum_{k=1}^{n} C_k M_\xi^k.$$

Доказательство:

По теореме о мат. ожидании:

$$M\left(\sum_{k=1}^{n} C_k \xi^k\right) = \sum_{i=1}^{\infty}\sum_{k=1}^{n} C_k x_i^k \, p_i = \sum_{k=1}^{n} C_k \sum_{i=1}^{\infty} x_i^k p_i = \sum_{k=1}^{n} C_k \, M_\xi^k.$$

Формула верна, если ряд абсолютно сходится. Покажем:

$$\sum_{i=1}^{\infty}\left|\sum_{k=1}^{n} C_k x_i^k\right| p_i \le \sum_{i=1}^{\infty}\sum_{k=1}^{n} |C_k||x_i|^k \, p_i = \sum_{k=1}^{n} |C_k| \sum_{i=1}^{\infty} |x_i|^k p_i < \infty.$$

$$M_\xi^k = \sum_{i=1}^{\infty} |x_i|^k p_i < \infty - \text{по условию}, \qquad \sum_{k=1}^{n} - \text{конечна}.$$

Аналогично для функции многих переменных:

$$\varphi(\xi_1, \xi_2, \dots, \xi_n) = \zeta$$

$$M_\zeta = \sum_{i_1, i_2, \dots, i_n} \varphi\left(x_{i_1}, x_{i_2}, \dots, x_{i_n}\right) p_{i_1, i_2, \dots, i_n} - \text{для дискретных случайных векторов},$$

ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ МГУ ИМЕНИ М.В. ЛОМОНОСОВА
teach-in
ЛЕКЦИИ УЧЕНЫХ МГУ

где $p_{i_1,i_2,\dots,i_n} = p(\xi_1 = x_{i_1}, \dots, \xi_n = x_{i_n})$.

$$M_\zeta = \int\limits_{R_n} \varphi(x_1, x_2, \dots, x_n) p(x_1, x_2, \dots, x_n)\, d\vec{x} - \text{для абсолютно непрерывных сл. в,}$$

если существует $p(x_1, x_2, \dots, x_n)$ - плотность вероятности случайного вектора.

Свойства M_ξ:

1) Пусть (ξ, η) – случайные величины, a, b – константы, тогда:

$$M(\mathrm{a}\xi + \mathrm{b}\eta) = \mathrm{a}M_\xi + \mathrm{b}M_\eta, \qquad \text{если } M_\xi \text{ и } M_\eta \text{ существуют.}$$

Проверим для дискретных случайных величин:

$$M(\mathrm{a}\xi + \mathrm{b}\eta) = \sum_{ij} (\mathrm{a}x_i + \mathrm{b}y_j) p_{i,j} = \sum_{ij} \mathrm{a}x_i p_{i,j} + \sum_{ij} \mathrm{b}y_j p_{i,j} =$$

если ряды сходятся абсолютно, то можно поменять порядок суммирования

$$= \sum_{i=1}^{\infty} \mathrm{a}x_i \sum_{j=1}^{\infty} p_{i,j} + \sum_{j=1}^{\infty} \mathrm{b}y_j \sum_{i=1}^{\infty} p_{i,j} = a \sum_{i=1}^{\infty} x_i p_i + b \sum_{j=1}^{\infty} y_j\, p_j = \mathrm{a}M_\xi + \mathrm{b}M_\eta,$$

$$\text{потому что } \sum_{j=1}^{\infty} p_{i,j} = p(\xi = x_i), \qquad \sum_{i=1}^{\infty} p_{i,j} = p(\eta = y_j).$$

Абсолютная сходимость рядов следует из существования M_ξ и M_η.

Для абсолютно непрерывных случайных величин:

$$\int\limits_{-\infty}^{\infty} \int\limits_{-\infty}^{\infty} (ax + by) p_{\xi,\eta}(x, y) dx dy =$$

$$= a \int\limits_{-\infty}^{\infty} x \left(\int\limits_{-\infty}^{\infty} p_{\xi,\eta}(x, y) dy \right) dx + b \int\limits_{-\infty}^{\infty} y \left(\int\limits_{-\infty}^{\infty} p_{\xi,\eta}(x, y) dx \right) dy =$$

$$= a \int\limits_{-\infty}^{\infty} x p_\xi(x) dx + b \int\limits_{-\infty}^{\infty} y p_\eta(y) dy = \mathrm{a}M_\xi + \mathrm{b}M_\eta,$$

$$\text{потому что } \int\limits_{-\infty}^{\infty} p_{\xi,\eta}(x, y) dy = p_\xi(x), \qquad \int\limits_{-\infty}^{\infty} p_{\xi,\eta}(x, y) dx = p_\eta(y).$$

2) Если ξ, η – независимые случайные величины, тогда

$$M_{\xi,\eta} = M_\xi M_\eta, \qquad \text{если } M_\xi \text{ и } M_\eta \text{ существуют.}$$

Доказательство:

$$M_{\xi,\eta} = \sum_{ij} x_i y_j p_{i,j} = \sum_{ij} x_i p_i y_j q_j = \sum_{i=1}^{\infty} x_i p_i \sum_{j=1}^{\infty} y_j q_j = M_\xi M_\eta.$$

$$p_{i,j} = p_i q_j - \text{в силу независимости.}$$

+ абсолютная сходимость доказывается аналогично.

Опр.: k – м центральным моментом случайной величины называется число:

$$M(\xi - M_\xi)^k, \qquad \text{если } M\left|\xi - M_\xi\right|^k < \infty.$$

Центральный момент отличается от простого момента тем, что из значения ξ вычитается математическое ожидание, и дальше усредняется отклонение ξ от своего математического ожидания (т.е. центра).

Опр.: Центральный момент второго порядка называется дисперсией ξ, если он существует.

$$D_\xi = M\left(\xi - M_\xi\right)^2.$$

Используя свойства математического ожидания можно записать:

$$D_\xi = M_{\xi^2} - M\left(2\xi M_\xi\right) + \left(M_\xi\right)^2 = \left\{M_\xi = const\right\} = M_{\xi^2} - 2M_\xi M(\xi) + \left(M_\xi\right)^2 =$$

$$= M_{\xi^2} - 2\left(M_\xi\right)^2 + \left(M_\xi\right)^2 = M_{\xi^2} - \left(M_\xi\right)^2.$$

Пример 1: Пуассоновская случайная величина.

$$P(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \qquad k = 0,1,2,\ldots$$

$$M_{\xi^2} = \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = \left[k^2 = k\left((k-1)+1\right)\right] = e^{-\lambda}\left[\sum_{k=0}^{\infty} k(k-1)\frac{\lambda^k}{k!} + \sum_{k=0}^{\infty} k\frac{\lambda^k}{k!}\right] =$$

$$= e^{-\lambda}\left[\lambda^2 \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} + \lambda \sum_{k=0}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}\right] = e^{-\lambda}\left[\lambda^2 e^\lambda + \lambda e^\lambda\right] = \lambda^2 + \lambda.$$

$$M_\xi = \lambda - \text{уже было показано.}$$

$$D_\xi = M_{\xi^2} - \left(M_\xi\right)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

Пример 2: Нормально распределенная (Гауссова) случайная величина.

$$\xi \sim N(\mu, \sigma^2).$$

$$D_\xi = M_{\xi^2} - \left(M_\xi\right)^2 = \int_{-\infty}^{\infty} (x-\mu)^2 \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} d(x-\mu) =$$

$$= 2\sigma^2 \int\limits_{-\infty}^{\infty} \frac{(x-\mu)^2}{2\sigma^2}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, \frac{d(x-\mu)}{\sqrt{2\sigma^2}\cdot\sqrt{\pi}} = \left\{\frac{(x-\mu)^2}{2\sigma^2} = z^2\right\} = \frac{2\sigma^2}{\sqrt{\pi}} \int\limits_{-\infty}^{\infty} z^2\, e^{-z^2}\, dz =$$

$$= \left\{2z^2 e^{-z^2} dz = -z de^{-z^2}\right\} = -\frac{\sigma^2}{\sqrt{\pi}} \int\limits_{-\infty}^{\infty} z de^{-z^2} = \{\text{интегрируем по частям}\} =$$

$$= \frac{\sigma^2}{\sqrt{\pi}}\left[-ze^{-z^2}\Big|_{-\infty}^{\infty} + \int\limits_{-\infty}^{\infty} e^{-z^2}\, dz\right] = \left\{e^{-z^2}\Big|_{-\infty}^{\infty} = 0,\ \int\limits_{-\infty}^{\infty} e^{-z^2}\, dz = \sqrt{\pi}\right\} = \sigma^2.$$

Свойства дисперсии:

1) $\xi_1, \xi_2, \dots, \xi_n$ – случайные величины

$$D\sum_{j=1}^{\infty} \xi_j = M\left(\sum_{j=1}^{n}\left(\xi_j - M_{\xi_j}\right)\right)^2 = M\left(\left(\sum_{j=1}^{n}\left(\xi_j - M_{\xi_j}\right)\right)\left(\sum_{k=1}^{n}\left(\xi_k - M_{\xi_k}\right)\right)\right) =$$

$$= \sum_{j=1}^{n} M\left(\xi_j - M_{\xi_j}\right)^2 + \sum_{j\neq k} M\left(\left(\xi_j - M_{\xi_j}\right)\left(\xi_k - M_{\xi_k}\right)\right) = \sum_{j=1}^{n} D_{\xi_j} + \sum_{j\neq k} cov\left(\xi_j, \xi_k\right).$$

$$M\left(\left(\xi_j - M_{\xi_j}\right)\left(\xi_k - M_{\xi_k}\right)\right) = cov\left(\xi_j, \xi_k\right) - \text{ковариация.}$$

Если $cov\left(\xi_i, \xi_j\right) = 0 \ \rightarrow\ \xi_i, \xi_j -$ некоррелированные.

Для некоррелированных случайных величин:

$$D\sum_{j=1}^{\infty} \xi_j = \sum_{j=1}^{n} D_{\xi_j}.$$

Если $\xi_1, \xi_2, \dots, \xi_n$ – независимы $\Longrightarrow \xi_i, \xi_j -$ некоррелированны при $i \neq j$,

$$\text{т. к. } M\left(\left(\xi_i - M_{\xi_i}\right)\left(\xi_j - M_{\xi_j}\right)\right) = \left(M\left(\xi_i - M_{\xi_i}\right)\right)\left(M\left(\xi_j - M_{\xi_j}\right)\right) = 0$$

в силу свойств мат. ожидания.

Однако из некоррелируемости независимость не следует!

Пример:

Дисперсия ξ: $P(\xi = k) = C_n^k p^k (1-p)^{n-k}$.

Пусть ξ – число успехов в серии n испытаний, введем сл. величину

$$\alpha_j = \begin{cases} 1, & \text{если в } j - \text{м испытании успех} \\ 0, & \text{если в } j - \text{м испытании неудача} \end{cases}$$

$$\xi = \sum_{j=1}^{n} \alpha_j.$$

$$M_{\alpha_j} = 0 \cdot q + 1 \cdot p = p,$$

$$M_{\alpha_j{}^2} = 0^2 \cdot q + 1^2 \cdot p = p,$$

$$D_{\alpha_j} = M_{\alpha_j{}^2} - \left(M_{\alpha_j}\right)^2 = p - p^2 = p(1-p) = pq.$$

$$D_\xi = D\sum_{j=1}^{n} \alpha_j = \sum_{j=1}^{n} D_{\alpha_j} = npq.$$

2) $\eta = c \cdot \xi \implies D_\eta = c^2 D_\xi$.

Действительно:

$$D_\eta = M\left(\left(c\xi - M_{c\xi}\right)^2\right) = c^2 M\left(\left(\xi - M_\xi\right)^2\right) = c^2 D_\xi.$$

3) Пусть ξ, η – случайные величины:

$$P(\xi \geq \eta) = 1 \implies \xi \geq \eta.$$

Тогда:

$$\alpha = \xi - \eta \geq 0,$$

$$P_\alpha(x) = 0, \qquad \text{для } x < 0.$$

$$M_\alpha = \int_{-\infty}^{\infty} x p_\alpha(x) dx = \int_{0}^{\infty} x p_\alpha(x) dx \geq 0$$

$$M_{\xi-\eta} \geq 0 \implies M_\xi \geq M_\eta.$$

Неравенства Чебышёва

1) ξ – случайная величина: $M_\xi, M_{|\xi|} < \infty$

$$\eta = \begin{cases} 0, & \text{если } |\xi| \leq \varepsilon \\ \varepsilon, & \text{если } |\xi| > \varepsilon \end{cases}$$

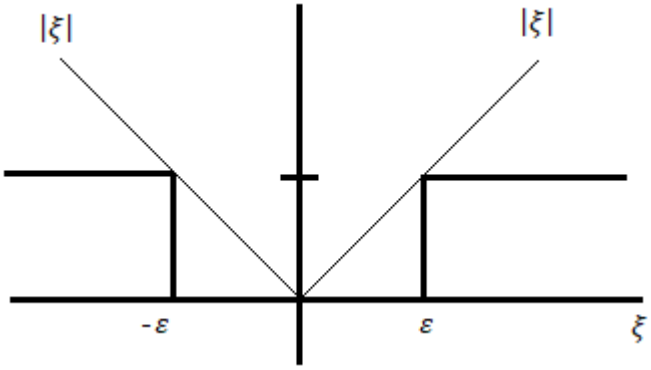


Рис. 6.2. Случайная величина η.

$$\eta \le |\xi| \text{ с вероятностью } 1$$

$$\Longrightarrow M_\eta \le M_{|\xi|}$$

$$\Longrightarrow 0 \cdot P(|\xi| \le \varepsilon) + \varepsilon \cdot P(|\xi| > \varepsilon) \le M_{|\xi|}$$

$$\Longrightarrow \boxed{P(|\xi| > \varepsilon) \le \frac{M_{|\xi|}}{\varepsilon}} \text{ – неравенство Чебышёва.}$$

2) Если $|\xi| \ge \eta \ge 0$, то $\xi^2 \ge \eta^2$

$$\Longrightarrow M_{\xi^2} \ge M_{\eta^2}$$

$$\Longrightarrow M_{\xi^2} \ge 0^2 \cdot P(|\xi| \le \varepsilon) + \varepsilon^2 \cdot P(|\xi| > \varepsilon)$$

$$\Longrightarrow P(|\xi| \ge \varepsilon) \le \frac{M_{\xi^2}}{\varepsilon^2}$$

Заменим ξ на $\xi - M_\xi$:

$$\boxed{P\left(\left|\xi - M_\xi\right| \ge \varepsilon\right) \le \frac{D_\xi}{\varepsilon^2}} \text{ – неравенство Чебышёва.}$$

3) $\xi \ge 0$

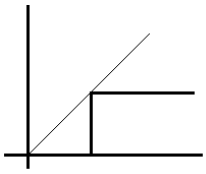


Рис. 6.3.

$$\Longrightarrow \boxed{P(\xi > \varepsilon) \le \frac{M_\xi}{\varepsilon}} \text{ – неравенство Маркова.}$$

ЛЕКЦИЯ 7. НЕРАВЕНСТВО КОШИ-БУНЯКОВСКОГО

Неравенства Чебышева:

$$P(\,|\xi| \geq \varepsilon) \leq \frac{M_{|\xi\,|}}{\varepsilon},$$

$$P\big(\,|\xi - M_\xi| \geq \varepsilon\big) \leq \frac{D_\xi}{\varepsilon^2}.$$

Неравенство Коши-Буняковского:

Возьмем

$$M(\lambda\xi + \eta)^2 \geq 0,$$

Пользуясь свойствами математического ожидания, получим:

$$\lambda^2 M_{\xi^2} + \lambda M_{\xi\eta} + M_{\eta^2} \geq 0.$$

Это выполняется, когда дискриминант

$$\frac{D}{4} = \frac{b}{4} - ac \leq 0.$$

$$\left(M_{\xi\eta}\right)^2 - M_{\xi^2} M_{\eta^2} \leq 0$$

$$\Longrightarrow M_{\xi^2} M_{\eta^2} \geq \left(M_{\xi\eta}\right)^2 - \text{неравенство Коши} - \text{Буняковского.}$$

Заменим $\xi \to \xi - M_\xi, \quad \eta \to \eta - M_\eta$:

$$D_\xi D_\eta \geq cov^2(\xi, \eta).$$

$$-1 \leq \frac{cov^2(\xi,\eta)}{D_\xi D_\eta} = corr \ \leq 1 \ \ - \text{коэффициент корреляции.}$$

Запишем также для случайных векторов (доказательство аналогично):

$$\vec{\xi}, \vec{\eta} - \text{случайные векторы}$$

$$\left(M\left(\vec{\xi}, \vec{\eta}\right)\right)^2 \leq M_{\left\|\vec{\xi}\right\|^2} M_{\|\vec{\eta}\|^2} - \text{неравенство Коши} - \text{Буняковского.}$$

Следствие:

Если существует M_{ξ^2}, то $M_{|\xi\,|} < \infty \ \Longrightarrow M_\xi$ существует.

$$\xi \to |\xi|$$

$$\eta \to 1$$

$$M_{|\xi\,|} \cdot 1 \leq \sqrt{M_{\xi^2} M_{1^2}} = \sqrt{M_{\xi^2} M_{1^2}} < \infty.$$

Лемма: Случайная величина ξ равна константе с вероятностью 1, если и только если $D_\xi = 0$.

Доказательство:

1) Пусть $P(\xi = const) = 1 \implies M_\xi = const \implies D_\xi = 0.$
2) Пусть $D_\xi = 0$, тогда по неравенству Чебышёва:

$$P\left(\left|\xi - M_\xi\right| > \frac{1}{n}\right) \le \frac{D_\xi}{\varepsilon^2} = 0 \ \ \forall n$$

$$\lim_{n\to\infty} P\left(\left|\xi - M_\xi\right| > \frac{1}{n}\right) = 0.$$

Если последовательность A_n сходится, то $\lim\limits_{n\to\infty} P(A_n) = P(\lim\limits_{n\to\infty}(A_n))$.

$$A_n = \left\{\omega \in \Omega\colon \left|\xi - M_\xi\right| > \frac{1}{n}\right\},$$

это монотонная последовательность, она сходится

$$A_n \xrightarrow[n\to\infty]{} \{\omega\colon \xi \ne M_\xi\}.$$

$$\implies P\big(\xi \ne M_\xi\big) = 0$$

$$\implies P\big(\xi = M_\xi = const\big) = 1. \blacksquare$$

Условное распределение случайных величин

$$B \in F, \qquad P(B) \ne 0$$

$$\implies F_\xi(x|B) = P(\xi < x|B).$$

Если $F_\xi(x|B)$ представима в виде интеграла:

$$F_\xi(x|B) = \int\limits_{-\infty}^{x} P_\xi(z|B)\,dz,$$

то $P_\xi(z|B)$ называется условной плотностью вероятности ξ, если она ≥ 0 и кусочно-непрерывна.

Рассмотрение случайных величин позволяет распространить понятие условной вероятности для условий, вероятность которых равна нулю.

$$B = \{y < \eta < (y + \Delta)\}$$

$$\exists\, P_\eta(y) \implies P(B) \xrightarrow[\Delta y\to 0]{} 0.$$

Значит случайная величина η примет некоторое фиксированное значение y, вообще говоря, для абсолютно непрерывных случайных величин вероятность такого события равна 0.

Пусть ξ, η – случайные величины, $P_{\xi,\eta}(x, y)$ –совместная плотность вероятности. Тогда:

$$P\big(\xi < x|y < \eta < (y+\Delta)\big) = F_\xi(x|y < \eta < (y+\Delta)) = \frac{P\big(\xi < x,\ y < \eta < (y+\Delta)\big)}{P\big(y < \eta < (y+\Delta)\big)} =$$

$$= \frac{\int_{-\infty}^{x} dt \int_{y}^{y+\Delta} P_{\xi,\eta}\,(t,z)dz}{\int_{-\infty}^{\infty} dt \int_{y}^{y+\Delta} P_{\xi,\eta}\,(t,z)dz} = \left\{P_\xi(x) = \int\limits_{-\infty}^{\infty} P_{\xi,\eta}\,(x,y)dy\right\} =$$

$$= \frac{\int_{-\infty}^{x} dt P_{\xi,\eta}\,(t,y)\Delta + o(\Delta)}{P_\eta(y)\Delta + o(\Delta)} \xrightarrow[\Delta\to 0]{} \frac{\int_{-\infty}^{x} P_{\xi,\eta}\,(t,y)\Delta\, dt}{P_\eta(y)} = \int\limits_{-\infty}^{x} \frac{P_{\xi,\eta}\,(t,y)}{P_\eta(y)} dt,$$

получается: $\dfrac{P_{\xi,\eta}\,(t,y)}{P_\eta(y)} = P_\xi(x|\eta = y)$.

Аналог формулы умножения вероятности:

$$P_{\xi,\eta}\,(x,y) = P_{\xi|\eta}(x|y)P_\eta(y).$$

Аналог формулы полной вероятности:

(проинтегрируем формулу выше)

$$P_\xi(x) = \int\limits_{-\infty}^{\infty} P_{\xi,\eta}\,(x,y)dy = \int\limits_{-\infty}^{\infty} P_{\xi|\eta}(x|y)P_\eta(y)dy$$

Условное математическое ожидание случайной величины ξ

ξ – дискретная случайная величина, то

$$P(\xi = x_k) = p_k, \qquad \sum_{k=1}^{\infty} p_k = 1$$

$$\Longrightarrow M(x|B) = \sum_{k=1}^{\infty} x_k\, P(\xi = x_k|B), \qquad \text{если ряд сходится абсолютно.}$$

ξ – абсолютно непрерывная случайная величина, то

$$M(\xi|\eta) = \int\limits_{-\infty}^{\infty} xP_{\xi|\eta}(x|y)dx\,, \qquad \text{если интеграл сходится абсолютно.}$$

Утверждение: Если существует безусловное математическое ожидание M_ξ, то существует и условное $M(\xi|B)$ для любого $B \in F$.

Для дискретной случайной величины:

Пусть $P(\xi = x_k|B)$ – заданы, тогда

$$M(\xi|B) = \sum_{k=1}^{\infty} x_k\, P(\xi = x_k|B).$$

Проверим, сходится ли абсолютно этот ряд:

$$\sum_{k=1}^{\infty}|x_k|\,P(\xi = x_k|B) = \{P(B) \neq 0\} = \sum_{k=1}^{\infty}|x_k|\frac{P(\{\xi = x_k\} \cap B)}{P(B)} \leq$$

$$\leq \frac{1}{P(B)}\sum_{k=1}^{\infty}|x_k|P(\xi = x_k) = \frac{1}{P(B)}M_{|\xi|} < \infty - \text{абсолютно сходится.}$$

Рассмотрим: $B_1, \dots, B_s, \dots$ - полная группа попарно несовместных событий,

$$M_\xi = \sum_{k=1}^{\infty} x_k\,P(\xi = x_k) = \sum_{k=1}^{\infty} x_k \sum_{s=1}^{\infty} P(\xi = x_k|B_s)P(B_s) = \begin{Bmatrix}\text{обозначим} \\ P(B_s) = q_s\end{Bmatrix} =$$

$$= \sum_{s=1}^{\infty}\left(\sum_{k=1}^{\infty} x_k\,P(\xi = x_k|B_s)\right)q_s = \left\{\sum_{k=1}^{\infty} x_k\,P(\xi = x_k|B_s) = M(\xi|B_s)\right\} = \sum_{s=1}^{\infty} M(\xi|B_s)\,q_s.$$

Пусть $B_s = \{\eta = y_s\}$, тогда

$$M_\xi = \sum_{s=1}^{\infty}\big(M(\xi|\eta = y_s)\big)q_s\,, \qquad M(\xi|\eta = y_s) = \varphi(\eta) - \text{функция условия.}$$

$$\boxed{M_\xi = M_\eta M(\xi|\eta)}.$$

Следствие:

ξ, η – случайные величины, $\varphi(\eta)$ – случайная величина.

$$M(\varphi(\eta)\xi|\eta = y_s) = \sum_{k=1}^{\infty} x_k\varphi(y_s)P(\xi = x_k|\eta = y_s) = \varphi(y_s)\sum_{k=1}^{\infty} x_k P(\xi = x_k|\eta = y_s) =$$

$$= \varphi(y_s)M(\xi|\eta = y_s).$$

Поскольку это выполняется для всех y_s, можно записать:

$$M(\varphi(\eta)\xi|\eta) = \varphi(\eta)M(\xi|\eta).$$

Если ξ, η – независимы ⟹

$$M(\xi) = M(\xi|\eta).$$

Задача наилучшего в среднем квадратичном прогноза

Пусть ξ, η – случайные величины; ξ – наблюдаемая, а η – не наблюдаемая. Выбрать f(ξ) наиболее близкую к η, т.е.

$$M(\mathrm{f}(\xi) - \eta)^2 - min.$$

Теорема:

Наилучшей в с.к. оценкой η по наблюдению ξ является $M(\eta|\xi)$, ее погрешность равна $M\big(\eta - M(\eta|\xi)\big)^2$.

Доказательство:

На основании 1 − го свойства условного математического ожидания:

$$M\big(\eta - f(\xi)\big)^2 = M_\xi\left(M([\eta - f(\xi)]^2|\xi)\right) =$$

$$= M_\xi\left(M\left(\left[(\eta - M(\eta|\xi)) + \big(M(\eta|\xi) - f(\xi)\big)\right]^2|\xi\right)\right) =$$

$$= M_\xi\left\{M\left[\big(\eta - M(\eta|\xi)\big)^2|\xi\right] + M\left[\big(M(\eta|\xi) - f(\xi)\big)^2|\xi\right]\right.$$
$$\left. + 2M\left[\big(\eta - M(\eta|\xi)\big)\big(M(\eta|\xi) - f(\xi)\big)|\xi\right]\right\}.$$

Подбираем такое значение f, чтобы $M\big(\eta - M(\eta|\xi)\big)^2$ была минимальной:

- Третье слагаемое равно нулю, т.к. $\big(M(\eta|\xi) - f(\xi)\big)|\xi$ можно вытащить перед M, при фиксированных ξ, а $M\big[\big(\eta - M(\eta|\xi)\big)|\xi\big] = M(\eta|\xi) - M(\eta|\xi) = 0$.
- Второе слагаемое может быть сделано равным нулю, если $f(\xi) = M(\eta|\xi)$.
- Первое слагаемое не зависит от f.

Таким образом, $M\big(\eta - f(\xi)\big)^2$ минимальна, когда в 1 слагаемом $\big(\eta - M(\eta|\xi)\big) = 0$. Т.е. $M(\eta|\xi)$ – наилучшая оценка.

Задача о наилучшем в среднем квадратичном линейном прогнозе

Задача выбора линейной функции ξ, чтобы наилучшим в с. к. образом приблизить случайную величину η.

$$M(\eta - a\xi - b)^2 - min$$

Это среднеквадратичная погрешность прогноза, будем искать ее минимум по a и b.

Известно: $M_\xi, M_\eta, D_\xi, D_\eta, cov(\xi,\eta)$.

$$M(\eta - a\xi - b)^2 = M\left(\big(\eta - M_\eta\big) + M_\eta - a\big(\xi - M_\xi\big) - a\,M_\xi - b\right)^2$$

Получили слагаемые, для которых: $M\big(\eta - M_\eta\big) = 0$, $M\left(a\big(\xi - M_\xi\big)\right) = 0$, остальные слагаемые – константа.

$$\Longrightarrow M(\eta - a\xi - b)^2 = D_\eta + a^2 D_\xi - 2a\,cov(\xi,\eta) + M\big(M_\eta - a\,M_\xi - b\big)^2 +$$
$$+2M\left(\left(\big(\eta - M_\eta\big) - a\big(\xi - M_\xi\big)\right)\big(M_\eta - a\,M_\xi - b\big)\right)$$

Подбираем такие значения a, b, чтобы $M(\eta - a\xi - b)^2$ была минимальной:

- Третье слагаемое равно нулю, т.к. $\left(M_\eta - a\,M_\xi - b\right)$ – константа, ее можно вытащить перед M, а $M\left(\left(\eta - M_\eta\right) - a\left(\xi - M_\xi\right)\right) = 0$.
- Второе слагаемое может быть сделано равным нолю выбором b (при любых a). Если $b = M_\eta - a\,M_\xi$.
- Первое слагаемое можно сделать минимальным выбором a, производную по a приравнять к нулю: $a = \frac{cov(\xi,\eta)}{D_\xi}$.

Таким образом, наилучшая оценка:

$$\hat{\eta} = a\xi + b = a\xi + M_\eta - a\,M_\xi = M_\eta + a\left(\xi - M_\xi\right) = \boxed{M_\eta + \frac{cov(\xi,\eta)}{D_\xi}\left(\xi - M_\xi\right)}.$$

$$M(\eta - \hat{\eta})^2 = D_\eta - \frac{cov^2(\xi,\eta)}{D_\xi}.$$

$$-1 < \frac{cov(\xi,\eta)}{\sqrt{D_\xi D_\eta}} = corr < 1$$

При corr равном 1 или -1 мы получаем точный прогноз. Corr –мера линейной связи, связь самая сильна, если коэффициент равен 1, если $corr = 0$, то величины не коррелированы, между ними нет линейной связи. Если величины некоррелированные, то наблюдение ξ никак не уточнит значение η:

$$\hat{\eta} = M_\eta + \frac{cov(\xi,\eta)}{D_\xi}\left(\xi - M_\xi\right) = M_\eta.$$

Если приближать случайную величину η некоторой константой, то нужно брать безусловное мат. ожидание. Погрешность должна быть минимальна:

$$M(\xi - C)^2 - min$$

$$M_{\xi^2} - 2CM_\xi + C^2 - min$$

Производную по C приравняем к 0:

$$C = M_\xi \implies M\left(\xi - M_\xi\right)^2 = D_\xi.$$

Векторные случайные величины

Если $\vec{\xi} \in R_n$ – случайный вектор, то

$$M_{\vec{\xi}} = \begin{pmatrix} M_{\xi_1} \\ M_{\xi_2} \\ \vdots \\ M_{\xi_n} \end{pmatrix};\ D_{\vec{\xi}} = \begin{pmatrix} D_{\xi_1} \\ D_{\xi_2} \\ \vdots \\ D_{\xi_n} \end{pmatrix}.$$

Матрица ковариаций:

$$\left(\sum_{\xi}\right)_{ij} = \begin{cases} D_{\xi_i}, & i = j \\ cov(\xi_i, \xi_j), & i \neq j \end{cases}$$

Если $\vec{\xi} \in R_n$, $\vec{\eta} \in R_m$, можем записать матрицу взаимных ковариаций:

$$\left(\sum_{\xi,\eta}\right)_{ij} = cov(\xi_i, \eta_j)$$

$$\sum_{\xi,\eta}{}^{*} = \sum_{\eta,\xi} - \text{транспонирование.}$$

Если $\vec{\xi}$ – наблюдается, а $\vec{\eta}$ – не наблюдается, то наилучшей в с. к. оценкой $\vec{\eta}$ по наблюдению $\vec{\xi}$ является его условное мат. ожидание $M(\vec{\eta}|\vec{\xi})$.

Наилучшая в с. к. линейная оценка:

$$\hat{\vec{\eta}} = M_{\hat{\vec{\eta}}} + \sum_{\eta,\xi}\sum_{\xi}{}^{-1}(\vec{\xi} - M_{\vec{\xi}}).$$

$$\sum_{\hat{\eta}} = \sum_{\eta} - \sum_{\eta,\xi}\sum_{\xi}{}^{-1}\sum_{\xi,\eta}.$$

Доказывать эту формулу не будем, она написана по аналогии и может быть получена теми же методами.

ЛЕКЦИЯ 8. ЗАКОНЫ БОЛЬШИХ ЧИСЕЛ

З. Б. Ч.

Опр.: последовательность случайных величин $\{\xi_n\}$ называется сходящейся по вероятности к величине ξ_0, если:

$$P(|\xi_n - \xi_0| \geq \varepsilon) \xrightarrow[n\to\infty]{} 0, \qquad \forall\varepsilon > 0.$$

$$P(|\xi_n - \xi_0| < \varepsilon) \xrightarrow[n\to\infty]{} 1 - \text{эквивалентное утверждение.}$$

$$\forall\varepsilon, \delta > 0 \ \ \exists N\colon \ \forall n \geq N$$

$$P(|\xi_n - \xi_0| < \varepsilon) > 1 - \delta,$$

$$P(|\xi_n - \xi_0| > \varepsilon) < \delta.$$

Сходимость по вероятности обозначается следующим образом: $\xi_n \xrightarrow{P} \xi_0$.

Критерий сходимости по вероятности:

Пусть η_n: $M_{\eta_n} = 0$ для всех n, и $D_{\eta_n} \xrightarrow[n\to\infty]{} 0$, тогда

$$\eta_n \xrightarrow{P} 0.$$

Доказательство:

Основано на неравенстве Чебышёва:

$$P\left(\left|\xi_n - M_{\eta_n}\right| > \varepsilon\right) = P(|\xi_n - 0| > \varepsilon) \leq \frac{D_{\eta_n}}{\varepsilon^2} \xrightarrow[n\to\infty]{} 0. \blacksquare$$

Закон больших чисел Чебышёва:

Пусть $\xi_1, \dots, \xi_n, \dots$ - попарно независимые случайные величины, $M_{\xi_n} = \mu_n$, и $D_{\xi_n} \leq C \ \ \forall n = 1,2,3, \dots$ (равномерно ограниченные), тогда

$$\frac{1}{n}\sum_{j=1}^{n}(\xi_j - \mu_j) \xrightarrow{P} 0, \qquad n \to \infty.$$

Доказательство:

Воспользуемся критерием сходимости по вероятности. Обозначим

$$\frac{1}{n}\sum_{j=1}^{n}(\xi_j - \mu_j) = \eta_n.$$

$$M_{\eta_n} = M\left(\frac{1}{n}\sum_{j=1}^{n}(\xi_j - \mu_j)\right) = \frac{1}{n}\sum_{j=1}^{n}\left(M_{\xi_j} - M_{\mu_j}\right) = \left\{M_{\xi_j} = \mu_j, \ M_{\mu_j} = \mu_j = const\right\} = 0.$$

$$D_{\eta_n} = D\left(\frac{1}{n}\sum_{j=1}^{n}(\xi_j - \mu_j)\right) = \frac{1}{n^2}\sum_{j=1}^{n} D(\xi_j - \mu_j) \le \frac{nc}{n^2} \xrightarrow[n\to\infty]{} 0, \qquad \text{т. к } D_{\xi_j},\ D_{\mu_j} \le C$$

Таким образом, выполнен критерий сходимости по вероятности, значит

$$\frac{1}{n}\sum_{j=1}^{n}(\xi_j - \mu_j) \xrightarrow{P} 0, \qquad n \to \infty. \blacksquare$$

Следствие:

Если ξ_j - случайная величина, такая что $M_{\xi_j} = \mu = const$, $D_{\xi_n} \le C$, тогда

$$\eta_n = \left(\frac{1}{n}\sum_{j=1}^{n}\xi_j\right) - \mu \quad \xrightarrow{P} 0$$

$$\text{или} \quad \left(\frac{1}{n}\sum_{j=1}^{n}\xi_j\right) \xrightarrow{P} \mu.$$

Этим часто пользуемся в физических экспериментах: много раз устраиваем эксперимент, наблюдаем ξ_j как независимые случайные величины, затем берем среднее арифметическое и надеемся, что оно хорошо оценивает мат. ожидание.

Закон больших чисел Бернулли:

Пусть ξ_n – число успехов в серии из n испытаний Бернулли, p - вероятность успеха при одном испытании. Тогда частота успеха по вероятности сходится к вероятности успеха:

$$\frac{\xi_n}{n} \xrightarrow{P} p, \qquad n \to \infty.$$

Доказательство:

Введем сл. вел.:

$$\alpha_j = \begin{cases} 1, & \text{если в } j-\text{м испытании успех} \\ 0, & \text{если в } j-\text{м испытании неудача} \end{cases}$$

$$\alpha_1, \dots, \alpha_n - \text{независимы попарно и в совокупности.}$$

$$\xi_n = \sum_{j=1}^{n}\alpha_j.$$

Обозначим:

$$\frac{1}{n}\xi_n - p = \frac{1}{n}\sum_{j=1}^{n}\alpha_j - p = \eta_n.$$

Математическое ожидание α_j считается по формуле математического ожидания для дискретной случайной величины:

$$M_{\alpha_j} = 0 \cdot q + 1 \cdot p = p, \qquad \text{для } j = 1,2,3 \ldots$$

Можем посчитать математическое ожидание η_n:

$$M_{\eta_n} = M\left(\frac{1}{n}\sum_{j=1}^{n}\alpha_j - p\right) = \frac{1}{n}\sum_{j=1}^{n}M\alpha_j - p = \frac{1}{n}\sum_{j=1}^{n}p - p = \frac{np}{n} - p = 0.$$

Теперь посчитаем дисперсию:

$$D_{\alpha_j} = M_{\alpha_j{}^2} - \left(M_{\alpha_j}\right)^2 = 0^2 q + 1^2 p - p^2 = p(1-p) = pq,$$

$$D_{\eta_n} = D\left(\left(\frac{1}{n}\sum_{j=1}^{n}\alpha_j\right) - p\right) = \frac{npq}{n^2} \xrightarrow[n\to\infty]{} 0.$$

Выполняется критерий сходимости, значит

$$\eta_n \xrightarrow{P} 0, \qquad n \to \infty$$

$$\Rightarrow \frac{1}{n}\xi_n - p \xrightarrow{P} 0$$

$$\Rightarrow \frac{1}{n}\xi_n \xrightarrow{P} p, \qquad n \to \infty. \blacksquare$$

Метод Монте-Карло

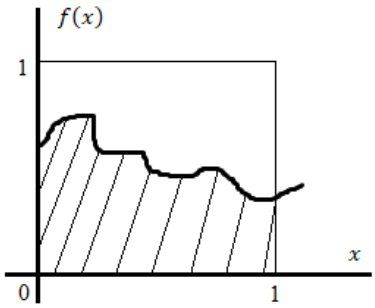


Рис. 8.1. Функция, определенная на квадрате 1×1.

$$\int_0^1 f(x)dx - ?$$

Необходимо вычислить интеграл от функции, заданной на интервале от 0 до 1, и принимающей значения тоже на интервале от 0 до 1. Для этого бросаем точку в единичный квадрат на плоскости: если точка попадает в заштрихованную область, то +1 к счетчику, если нет, то +0. Подсчитав число точек, попавших в заштрихованную область, и поделив его на общее количество точек, надеемся получить площадь под кривой.

Пусть $f(x)$- непрерывная функция: $[0,1] \to [0,1]$

$$(f(x) \in [0,1], \qquad x \in [0,1]).$$

$$\xi_j, \eta_j, \qquad j = 1,2,3 \dots n \quad - \text{независимые случайные величины}$$

$$\xi_j \sim [0,1] - \text{распределена равномерно,}$$

$$\eta_j \sim [0,1] - \text{распределена равномерно.}$$

$$\xi_j, \eta_j \to \rho_j = \begin{cases} 1, & \text{если } \eta_j \le f(\xi_j) \\ 0, & \text{если } \eta_j > f(\xi_j) \end{cases}$$

$$M_{\rho_j} = 1 \cdot p\left(\eta_j \le f(\xi_j)\right) + 0 \cdot p\left(\eta_j > f(\xi_j)\right) = p\left(\eta_j \le f(\xi_j)\right).$$

По следствию из З.Б.Ч Чебышёва:

$$\frac{1}{n}\sum_{j=1}^{n} \rho_j \xrightarrow{P} M_{\rho_j} = p\left(\eta_j \le f(\xi_j)\right) = \int\limits_{\substack{x,y \in [0,1] \\ y \le f(x)}} dxdy = \int\limits_0^1 f(x)dx.$$

$$\frac{1}{n}\sum_{j=1}^{n} \rho_j \xrightarrow{P} \int\limits_0^1 f(x)dx.$$

Сходимость с вероятностью 1 (почти наверное)

Случайная величина это функция, заданная на вероятностном пространстве (Ω, F, P), $\xi\colon \Omega \to R_1$

$$\{\omega\colon \xi(\omega) < x\} \in F, \qquad \forall x \in R_1.$$

Опр.: Последовательность $\{\xi_n\}$ сходится к ξ_0 с вероятностью 1, если

$$P(\{\omega\colon \xi_n(\omega) \to \xi_0(\omega)\}) = 1.$$

Такая сходимость обозначается следующим образом:

$$\xi_n \xrightarrow{\text{п.н.}} \xi_0 \text{ или } \xi_n \to \xi_0 (mod\, P) \text{ или } P\left(\lim \xi_n = \xi_0\right) = 1.$$

Являются ли событием те ω, при которых $\xi_n(\omega)$ является сходящейся последовательностью?

Пусть ω – фиксированное, тогда $\{\xi_n(\omega)\}$ – числовая последовательность, она сходится, если она фундаментальна.

$$\forall \varepsilon = \frac{1}{k} \ \ \exists N: \forall n_1, n_2 \geq N$$

$$\left|\xi_{n_1}(\omega) - \xi_{n_2}(\omega)\right| < \frac{1}{k} = \varepsilon.$$

Событие, которое состоит в том, что последовательность $\{\xi_n(\omega)\}$ сходится, состоит из всех тех ω, для которых выполнено данное соотношение.

$$\{\xi_n \to \xi_0\} = \bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n_1, n_2 \geq N} \left\{\omega: \left|\xi_{n_1}(\omega) - \xi_{n_2}(\omega)\right| < \frac{1}{k}\right\}$$

$$\xi_n \xrightarrow{\text{п.н.}} \xi_0 \Leftrightarrow P\left(\bigcap_{k=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n_1, n_2 \geq N} \left\{\omega: \left|\xi_{n_1}(\omega) - \xi_{n_2}(\omega)\right| < \frac{1}{k}\right\}\right) = 1$$

Так как вероятность пересечения равна 1, то и каждое из пересекающихся событий должно иметь вероятность 1, потому что пересечение не может иметь вероятность большую, чем пересекающиеся члены. Тогда,

$$\forall \varepsilon = \frac{1}{k}: \ P\left(\bigcup_{N=1}^{\infty} \bigcap_{n_1, n_2 \geq N} \left\{\omega: \left|\xi_{n_1}(\omega) - \xi_{n_2}(\omega)\right| < \frac{1}{k}\right\}\right) = 1.$$

А вероятность противоположного события должна быть равна нулю:

$$P\left(\overline{\bigcup_{N=1}^{\infty} \bigcap_{n_1, n_2 \geq N} \left\{\omega: \left|\xi_{n_1}(\omega) - \xi_{n_2}(\omega)\right| < \varepsilon\right\}}\right) =$$

$$= P\left(\bigcap_{N=1}^{\infty} \bigcup_{n_1, n_2 \geq N} \left\{\omega: \left|\xi_{n_1}(\omega) - \xi_{n_2}(\omega)\right| > \varepsilon\right\}\right) = 0.$$

$$\left\{\omega: \left|\xi_{n_1}(\omega) - \xi_{n_2}(\omega)\right| > \varepsilon\right\} = A_n - \text{последовательность}$$

$$\bigcap_{N=1}^{\infty} \bigcup_{n_1, n_2 \geq N} \left\{\omega: \left|\xi_{n_1}(\omega) - \xi_{n_2}(\omega)\right| > \varepsilon\right\} = A^{\sim} - \text{верх. предел } A_n$$

$$A^{\sim} = \{\omega: \exists \text{ бесконечно много } \ k: \omega \in A_k\}$$

Верхний предел состоит в том, что отличие $\xi_{n_1}(\omega)$ от $\xi_{n_2}(\omega)$ должно быть больше ε, для бесконечного количества индексов n. Этот предел должен быть равен нулю для того, чтобы была сходимость почти наверное.

Раз верхний предел должен иметь нулевую вероятность, то невозможно такое, что $\left|\xi_{n_1}(\omega) - \xi_{n_2}(\omega)\right| > \varepsilon$ выполняется бесконечное количество раз.

$$P(A^{\sim}) = 0 \Leftrightarrow P(\overline{A^{\sim}}) = 1.$$

$P(\overline{A^{\sim}}) = 1 \Leftrightarrow$ с вероятностью 1 происходит не более чем конечное число событий $\left|\xi_{n_1}(\omega) - \xi_{n_2}(\omega)\right| > \varepsilon,\ \forall\varepsilon > 0.$

Лемма Бореля-Кантелли

$A_1, A_2, \dots, A_n, \dots$ - последовательность событий, такая что

$$\sum_{\boldsymbol{k=1}}^{\infty} P(A_k) < \infty,$$

Тогда с вероятностью 1 происходит не более чем конечное число событий $\{A_k\}$.

Доказательство:

Надо показать, что $P(A^{\sim}) = 0$

$$P\left(\bigcap_{N=1}^{\infty}\bigcup_{k\geq N} A_k\right) = 0.$$

Если это так, то противоположное событие, состоящее в том, что происходит конечное число A_k, будет иметь вероятность 1, что и нужно доказать.

Надо показать, что верхний предел $P\left(\overline{\lim A_k}\right) = 0$

$$P\left(\bigcap_{N=1}^{\infty}\bigcup_{k\geq N} A_k\right) = 0$$

$$\bigcup_{k\geq N} A_k = D_n - \text{последовательность сужающихся множеств,}$$

$$\bigcap_{N=1}^{\infty}\bigcup_{k\geq N} A_k = \lim D_n.$$

$$P\left(\lim_{n\to\infty} D_n\right) = 0$$

$$\lim_{n\to\infty} P(D_n) = \lim_{n\to\infty} P\left(\bigcup_{k\geq N} A_k\right) = 0$$

Покажем, что последнее справедливо.

Воспользуемся свойством $P(A \cup B) \leq P(A) + P(B)$.

$$\lim_{n\to\infty} P\left(\bigcup_{k\geq N} A_k\right) \leq \lim_{n\to\infty} \sum_{k=n}^{\infty} P(A_k) = 0, \qquad \text{если ряд сходится.}$$

Усиленный закон больших чисел

Пусть $\xi_1, \dots, \xi_n, \dots$ - последовательность попарно независимых случайных величин, $M_{\xi_n} = (\mu = 0),\ D_{\xi_n} = \sigma^2$. Тогда

$$\frac{1}{n}\sum_{k=1}^{n}\xi_k \xrightarrow{\text{п.н.}} (\mu = 0).$$

Доказательство:

На основе леммы Бореля - Кантелли.

Обозначим:

$$\eta_n = \sum_{j=1}^{n}\xi_j.$$

$$n\colon\ m^2 \le n \le (m+1)^2$$

Для каждого n сконструируем:

$$P\left(\left|\frac{\eta_{m^2}}{m^2}\right| > \varepsilon\right) \le \frac{D\left(\frac{\eta_{m^2}}{m^2}\right)}{\varepsilon^2} = \frac{D(\eta_{m^2})}{(m^2)^2\varepsilon^2} = \frac{m^2\sigma^2}{(m^2)^2\varepsilon^2} = \frac{\sigma^2}{m^2\varepsilon^2} = -\text{нерав. Чебышёва.}$$

Вероятность такого события имеет мажоранту $\sim \frac{1}{m^2}$, значит

$$\sum_{m=1}^{\infty} P\left(\left|\frac{\eta_{m^2}}{m^2}\right| > \varepsilon\right) < \infty - \text{ряд сходится.}$$

По лемме Бореля - Кантелли с вероятностью 1 происходит не более чем конечное число событий $\left|\frac{\eta_{m^2}}{m^2}\right| > \varepsilon$. Это означает, что последовательность $\frac{\eta_{m^2}}{m^2} \xrightarrow{\text{п.н.}} 0$.

Теперь рассмотрим случай, когда $n \ne m^2$. Введем:

$$\hat{\eta}_{m^2} = \max_{m^2+1\le k\le (m+1)^2} |\xi_{m^2+1} + \xi_{m^2+2} + \cdots + \xi_k|$$

Рассмотрим следующее событие:

$$\left\{\left|\frac{\hat{\eta}_{m^2}}{m^2}\right| > \varepsilon\right\} = \bigcup\left\{\left|\frac{\xi_{m^2+1} + \xi_{m^2+2} + \cdots + \xi_k}{m^2}\right| > \varepsilon\right\}.$$

Снова воспользуемся неравенством Чебышёва:

$$P\left(\left|\frac{\hat{\eta}_{m^2}}{m^2}\right| > \varepsilon\right) \le \sum_{k=m^2+1}^{(m+1)^2} \frac{D\left(\frac{\hat{\eta}_{m^2}}{m^2}\right)}{\varepsilon^2} = \sum_{k=m^2+1}^{(m+1)^2} \frac{(k-m^2-1)\sigma^2}{\varepsilon^2 m^4} \le \sum_{k=0}^{2m+1} \frac{k\sigma^2}{\varepsilon^2 m^4} \le$$

$$\le \frac{\sigma^2 2m(2m+1)}{\varepsilon^2 m^4} \le \frac{4m^2\sigma^2}{\varepsilon^2 m^4} \le \frac{4\sigma^2}{\varepsilon^2 m^2}$$

$$\sum_{m=1}^{\infty} P\left(\left|\frac{\hat{\eta}_{m^2}}{m^2}\right| > \varepsilon\right) < \infty - \text{ряд сходится.}$$

С вероятностью 1 происходит не более чем конечное число событий $\left|\frac{\hat{\eta}_{m^2}}{m^2}\right| > \varepsilon$ по лемме Бореля – Кантелли.

Нужно показать, что

$$\sum_{n=1}^{\infty} P\left(\left|\frac{\eta_n}{n}\right| > \varepsilon\right) < \infty.$$

$$\frac{\eta_n}{n} \leq \left|\frac{\eta_{m^2}}{m^2}\right| + \left|\frac{\hat{\eta}_{m^2}}{m^2}\right|$$

$$\frac{\eta_{m^2}}{m^2} \xrightarrow{\text{п.н.}} 0, \qquad \frac{\hat{\eta}_{m^2}}{m^2} \xrightarrow{\text{п.н.}} 0$$

$$\Longrightarrow \frac{\eta_n}{n} \xrightarrow{\text{п.н.}} 0. \blacksquare$$

ЛЕКЦИЯ 9. ХАРАКТЕРИСТИЧЕСКИЕ ФУНКЦИИ СЛУЧАЙНЫХ ВЕЛИЧИН

Сходимость в среднем квадратичном:

$$\xi_n \xrightarrow{\text{с.к.}} \xi_0, \qquad \text{при } n \to \infty$$
$$\text{если } M(\xi_n - \xi_0)^2 \xrightarrow[n\to\infty]{} 0.$$

Сходимость по вероятности:

$$\xi_n \xrightarrow{P} \xi_0, \qquad \text{при } n \to \infty$$
$$\text{если } P(|\xi_n - \xi_0| > \varepsilon) \xrightarrow[n\to\infty]{} 0, \qquad \forall \varepsilon > 0.$$

Сходимость почти наверное:

$$\xi_n \xrightarrow{\text{п.н.}} \xi_0,$$
$$\text{если } P\big(\omega\colon \xi_n(\omega) \to \xi_0(\omega)\big) = 1.$$

Сходимость по распределению

Опр.: $\xi_n \xrightarrow[n\to\infty]{} \xi_0$ по распределению, если

$$F_{\xi_n}(x) - \text{функция распределения } \xi_n,$$
$$F_{\xi_0}(x) - \text{функция распределения } \xi_0,$$
$$F_{\xi_n}(x) \xrightarrow[n\to\infty]{} F_{\xi_0}(x) \text{ в каждой точке непрерывности} F_{\xi_0}(x)\ .$$

Сходимость по распределению обозначается следующим образом: $\xi_n \xrightarrow{d} \xi_0$

Связь между четырьмя типами сходимости:

$$\left.\begin{array}{l} \xi_n \xrightarrow{\text{с.к.}} \xi_0 \Longrightarrow \\ \xi_n \xrightarrow{\text{п.н.}} \xi_0 \Longrightarrow \end{array}\right. \xi_n \xrightarrow{P} \xi_0 \Longrightarrow \xi_n \xrightarrow{d} \xi_0$$

Вообще $\xi_n \xrightarrow{d} \xi_0 \not\Rightarrow \xi_n \xrightarrow{P} \xi_0$, но если $P(\xi_0 = m) = 1$:

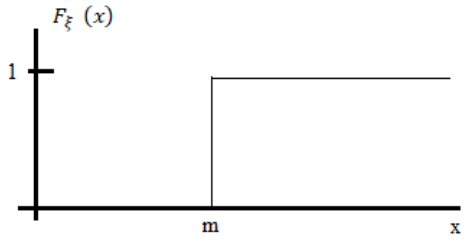


Рис. 9.1. Функция распределения.

В этом случае:

$$\xi_n \xrightarrow{d} \xi_0 = const \Longrightarrow \xi_n \xrightarrow{P} \xi_0 \text{ при } n \to \infty.$$

Характеристические функции случайных величин

Опр.: Характеристической функцией случайной величины ξ называется функция, заданная на всей числовой прямой:

$$f_\xi(t) = Me^{it\xi} = M(\cos t\xi + i\sin t\xi), \qquad t \in (-\infty, \infty).$$

$$\xi = Re\xi + i\,Im\xi, \qquad \text{тогда } M_\xi = M\,Re\xi + iM\,Im\xi.$$

$$f_\xi(t) = \begin{cases} \sum\limits_{k=1}^{\infty} e^{itx_k} p_k, & \text{дискретная} \\ \int\limits_{-\infty}^{\infty} e^{itx} p(x)dx\,, & \text{абсолютно непрерывная} \end{cases}$$

Свойства характеристических функций:

1) Характеристическая функция $\exists\ \forall\xi$.

Чтобы это показать, надо доказать что существует мат ожидание от комплексной экспоненты:

$$M\left|e^{it\xi}\right| = M1 = 1 - \text{всегда существует.}$$

2) $f_\xi(0) = 1, \quad \left|f_\xi(t)\right| \le 1$

Действительно:

$$Me^{it\xi}\Big|_{t=0} = Me^0 = 1.$$

$$\left|Me^{it\xi}\right| \le M\left|e^{it\xi}\right| = 1.$$

3) $\xi_1, \xi_2, \dots, \xi_n$ – независимые в совокупности случайные величины.

$$\alpha = \sum_{j=1}^{n} \xi_j$$

Тогда

$$f_{\xi_1+\xi_2+\dots+\xi_n}(t) = Me^{it(\xi_1+\xi_2+\dots+\xi_n)} = Me^{it\xi_1}e^{it\xi_2}\dots e^{it\xi_n} = \prod_{k=1}^{n} Me^{it\xi_k} = \prod_{k=1}^{n} f_{\xi_k}(t).$$

4)

$$\xi \to f_\xi(t)$$

$$\sigma\xi + \mu \to f_{\sigma\xi+\mu}(t) = e^{it\mu} f_\xi(\sigma t)$$

Это так, потому что по определению:

$$Me^{it(\sigma\xi+\mu)} = Me^{it\mu}e^{i(t\sigma)\xi} = e^{it\mu} f_\xi(\sigma t)$$

Вычислим характеристические функции нескольких случайных величин.

Пример 1:

Пусть $P(\xi = m) = 1$

$$f_\xi(t) = Me^{itm}.$$

Пример 2:

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

$$f_\xi(t) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{itx} e^{-\frac{x^2}{2}} dx = e^{-\frac{t^2}{2}}.$$

Пусть $\xi \sim N(\mu, \sigma^2)$

$$f_\xi(t) = \int_{-\infty}^{\infty} e^{itx} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{-\infty}^{\infty} e^{it(x-\mu)+\mu} \frac{\sigma}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \frac{d(x-\mu)}{\sigma} =$$

$$= e^{it\mu} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{it\sigma\frac{(x-\mu)}{\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \frac{d(x-\mu)}{\sigma} = \left\{ \frac{(x-\mu)}{\sigma} = z \right\} = e^{it\mu} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{it\sigma z} e^{-\frac{z^2}{2}} dz =$$

$$= e^{it\mu} e^{-\frac{t^2\sigma^2}{2}} = e^{it\mu} f_{N(0,1)}(t\sigma)$$

Получается:

$$\xi \sim N(0,1),$$
$$\sigma\xi + \mu \sim N(\mu, \sigma^2).$$

Свойство: Случайные величины, распределенные по разным законам, имеют разные характеристические функции. $\Longrightarrow$ Если характеристические функции случайных величин совпадают, то совпадают и их распределения.

Замечание:

$$\xi_i \sim N(0,1) - \text{независимо в совокупности для } i = 1, \dots, n,$$

$$\sum_{j=1}^{n} (\sigma_j \xi_j + \mu_j) \sim N\left(\sum_{j=1}^{n} \mu_j, \sum_{j=1}^{n} {\sigma_i}^2 \right).$$

$$f_{\sum_{j=1}^{n}(\sigma_j \xi_j + \mu_j)}(t) = \prod_{j=1}^{n} e^{it\mu_j} e^{-\frac{t^2 {\sigma_j}^2}{2}} = e^{it\sum_{j=1}^{n}\mu_j} e^{-\frac{t^2}{2}\sum_{j=1}^{n}{\sigma_j}^2}.$$

Пример 3:

$$P(\xi = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \qquad k = 0,1,2 \dots$$

$$f_\xi(t) = Me^{it\xi} = \sum_{k=0}^{\infty} e^{itk}\frac{\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{\left(e^{it}\lambda\right)^k}{k!} = e^{-\lambda}e^{e^{it}\lambda} = e^{\lambda(e^{it}-1)}.$$

Рассмотрим случайную величину, которая имеет распределение Пуассона

$$\xi_i \sim P(\lambda_j), \qquad \xi_1, \xi_2, \dots, \xi_n \text{ – независимы в совокупности}$$

$$f_{\xi_1+\xi_2+\dots+\xi_n}(t) = \prod_{j=1}^{n} e^{\lambda_j(e^{it}-1)} = e^{(e^{it}-1)\sum_{j=1}^{n}\lambda_j} = f_{P\left(\sum_{j=1}^{n}\lambda_j\right)}(t).$$

Теорема:

Если $\exists\ M_\xi{}^k$, то $\exists$ и непрерывна производная k-го порядка:

$$f_\xi^{(k)}(t), \qquad \boxed{f_\xi^{(k)}(0) = i^k M_\xi{}^k}$$

Покажем дифференцируемость и связь моментов с характеристической функцией:

$$f_\xi(t) = \int_{-\infty}^{\infty} e^{itx}p(x)dx$$

Однократно продифференцируем:

$$f_\xi^{(1)}(t) = \int_{-\infty}^{\infty} ix\, e^{itx}p(x)dx.$$

При дифференцировании k раз:

$$f_\xi^{(k)}(t) = \int_{-\infty}^{\infty} (ix\,)^k e^{itx}p(x)dx$$

$$f_\xi^{(k)}(0) = \int_{-\infty}^{\infty} (ix\,)^k e^{itx}p(x)dx\Big|_{t=0} = i^k\int_{-\infty}^{\infty} x^k p(x)dx = i^k M_\xi{}^k.$$

Теорема о непрерывности характеристических функций

Пусть $\xi_1, \xi_2, \dots, \xi_n, \dots$ - последовательность случайных величин, ξ_0- случайная величина. Каждой случайной величине соответствуют функция распределения и характеристическая функция:

$$F_1(x), F_2(x), \dots, F_n(x);\ \ F_0(x),$$
$$f_1(x), f_2(x), \dots, f_n(x);\ \ f_0(x).$$

Теорема:

Пусть $f_n(t) \xrightarrow[n\to\infty]{} f_0(t) \quad \forall t \in (-\infty, \infty)$, $f_0(t)$ – непрерывна в точке $t = 0$. Это условие необходимо и достаточно для того, чтобы $F_n(t) \xrightarrow[n\to\infty]{} F_0(t)$ в каждой точке непрерывности $F_0(x)$.

Если $F_0(x)$- непрерывна на $(-\infty, \infty)$, то сходимость $F_n(t) \xrightarrow[n\to\infty]{} F_0(t)$ равномерна.

Применение к ЗБЧ

Пусть $\xi_1, \xi_2, \dots, \xi_n, \dots$ - последовательность независимых в совокупности случайных величин и одинаково распределенных. $M_{\xi_j} = \mu$. Покажем, что

$$\frac{1}{n}\sum_{j=1}^{n} \xi_j \xrightarrow{P} \mu.$$

Пусть $\varphi(t) = Me^{it\xi_1}$. У всех ξ_j одинаковое распределение, поэтому характеристическая функция одинакова для всех. Поэтому:

$$\varphi_{\frac{1}{n}\sum_{j=1}^{n}\xi_j}(t) = \left[\varphi\left(\frac{t}{n}\right)\right]^n - \text{из свойств хар. функций 2 и 3.}$$

Характеристическая функция дифференцируема в нуле, т.к. существует момент первого порядка. Первая производная существует и непрерывна, поэтому представим значение функции по формуле Тейлора:

$$\varphi_{\frac{1}{n}\sum_{j=1}^{n}\xi_j}(t) = \left[\varphi\left(\frac{t}{n}\right)\right]^n = \left[\varphi(0) + \varphi'(0)\frac{t}{n} + o\left(\frac{1}{n}\right)\right]^n = \left[0 + \frac{it\mu}{n} + o\left(\frac{1}{n}\right)\right]^n \to e^{it\mu}$$

$$e^{it\mu} - \text{характеристическая функция величины } \zeta, \qquad P(\zeta = \mu) = 1.$$

Значит, последовательность $\frac{1}{n}\sum_{j=1}^{n}\xi_j \xrightarrow{d} \zeta = const \implies \frac{1}{n}\sum_{j=1}^{n}\xi_j \xrightarrow{P} \mu = const$, при n→ ∞.

Центральная предельная теорема

Пусть $\xi_1, \xi_2, \dots, \xi_n, \dots$ - последовательность независимых в совокупности и одинаково распределенных случайных величин. $M_{\xi_j} = \mu$, $D_{\xi_j} = \sigma^2$. Тогда

$$\frac{1}{\sqrt{n}\,\sigma}\sum_{j=1}^{n}(\xi_j - \mu) \xrightarrow{d} \eta \sim N(0,1), \qquad \text{при n} \to \infty.$$

Доказательство:

Основано на теореме о непрерывности. Если удастся доказать сходимость характеристических функций к $f_{N(0,1)}(t) = e^{-\frac{t^2}{2}}$, то теорема будет доказана.

Рассмотрим случайную величину $\xi_j - \mu$:

$$\xi_j - \mu \to \varphi(t) - \text{характеристическая функция.}$$

$$f_{\frac{1}{\sqrt{n}\,\sigma}\sum_{j=1}^{\infty}(\xi_j-\mu)}(t)=\left[\varphi\left(\frac{t}{\sqrt{n}\,\sigma}\right)\right]^n$$

Функция φ при фиксированном t имеет аргумент, стремящийся к нулю. Поэтому при любом фиксированном t можно воспользоваться разложением этой функции в окрестности нуля по формуле Тейлора с остаточным членом в форме Пеано. Первая производная в нуле существует, т.к. есть мат. ожидание. Вторая производная в нуле тоже существует, т.к. есть дисперсия – это момент 2 порядка.

$$f_{\frac{1}{\sqrt{n}\,\sigma}\sum_{j=1}^{\infty}(\xi_j-\mu)}(t)=\left[\varphi\left(\frac{t}{\sqrt{n}\,\sigma}\right)\right]^n=\left[\varphi(0)+\varphi'(0)\frac{t}{\sqrt{n}\,\sigma}+\frac{1}{2}\varphi''(0)\frac{t^2}{n\sigma^2}+o\left(\frac{1}{n}\right)\right]^n=$$

$$=\left[1+i\cdot 0\cdot\frac{t}{\sqrt{n}\,\sigma}+\frac{1}{2}i^2\sigma^2\frac{t^2}{n\sigma^2}+o\left(\frac{1}{n}\right)\right]^n=\left[1-\frac{t^2}{2n}+o\left(\frac{1}{n}\right)\right]^n\xrightarrow[n\to\infty]{}e^{-\frac{t^2}{2}}.$$

Сходится как замечательный предел.

Получили, что в условиях центральной предельной теоремы случайная величина $\frac{1}{\sqrt{n}\,\sigma}\sum_{j=1}^{n}(\xi_j-\mu)$ по распределению сходится к случайной величине, распределенной по нормальному закону. Что позволяет оценивать вероятность попадания случайной величины в любое борелевское множество, пользуясь законом распределения нормальной случайной величины.

Восстановление распределения случайных величин по наблюдению реализации (применение ЗБЧ и ЦПТ):

ξ - случайная величина, ∀B ∈ B-борелевское

$$P(\xi\in B)=P(B).$$

Пусть $\xi_1,\xi_2,\dots,\xi_n,\ \dots$ - последовательность независимых в совокупности случайных величин и одинаково распределенных.

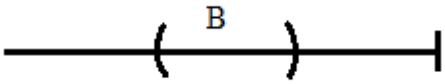


Рис. 9.2. Событие B.

$$\chi_B(x)=\begin{cases}1, & x\in B\\ 0, & x\,\overline{\in}\,B\end{cases}-\text{индикаторная функция,}$$

$$\chi_B(\xi_j)=\begin{cases}1, & \xi\in B\\ 0, & \xi\,\overline{\in}\,B\end{cases}.$$

$$M\chi_B(\xi_j)=1P(\xi\in B)+0P(\xi\,\overline{\in}\,B)=P(B).$$

$$M\chi_B(\xi_j)=P(B)\big(1-P(B)\big).$$

В силу ЗБЧ можем утверждать, что

ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ МГУ ИМЕНИ М.В. ЛОМОНОСОВА
teach-in
ЛЕКЦИИ УЧЕНЫХ МГУ

$$\frac{1}{n}\sum_{j=1}^{n}\chi_B(\xi_j) \xrightarrow{P} P(B), \qquad \text{при } n \to \infty.$$

Обозначим

$$\frac{1}{n}\sum_{j=1}^{n}\chi_B(\xi_j) = P_n^*(B) - \text{выборочная вероятность.}$$

Это частота появления события B в n испытаний.

В силу ЦПТ:

$$\frac{1}{\sqrt{n}\sqrt{P(B)\big(1-P(B)\big)}}\sum_{j=1}^{n}(\chi_B(\xi_j) - P(B)) \sim N(0,1),$$

$$\frac{\sqrt{n}\big(P_n^*(B) - P(B)\big)}{\sqrt{P(B)\big(1-P(B)\big)}} \xrightarrow{d} N(0,1), \qquad \text{при } n \to \infty.$$

Интегральная теорема Муавра-Лапласа

Интегральная теорема Муавра-Лапласа является следствием ЦПТ.

Рассмотрим частоту успехов в серии n испытаний Бернулли:

$$\frac{k}{n} = \frac{\sum_{j=1}^{n}\alpha_j}{n}.$$

В соответствии с ЦПТ

$$\frac{\sum_{j=1}^{n}(\alpha_j - p)}{\sqrt{n}\sqrt{pq}} \xrightarrow{d} N(0,1), \qquad \text{при } n \to \infty$$

$$\Rightarrow \frac{k-np}{\sqrt{npq}} \xrightarrow{d} N(0,1), \qquad \text{при } n \to \infty$$

$$\Rightarrow P\left(a < \frac{k-np}{\sqrt{npq}} < b\right) = \int_a^b P_{N(0,1)}(x)dx.$$

ЛЕКЦИЯ 10. ЦЕПИ МАРКОВА

(Ω, F, P) – вероятностное пространство, которое описывает эксперимент со случайным исходом.

$$\Omega = \{\omega_1, \dots, \omega_n\}$$

$$F - \text{множество всех подмножеств } \Omega$$

$$P(\{\omega_j\}) = p_j, \qquad \sum_{j=1}^{n} p_j = 1$$

Повторим эксперимент n раз:

$$\{\langle \omega_{i_1}, \dots, \omega_{i_n} \rangle\} = \Omega^n$$

$$F_n - \text{алгебра всех подмножеств } \Omega^n$$

$$P(\{\langle \omega_{i_1}, \dots, \omega_{i_n} \rangle\}) = \prod_{k=1}^{n} p_{i_k}. \ (*)$$

$$(\Omega^n, F_n, P_*) - \text{последовательность независимых испытаний.}$$

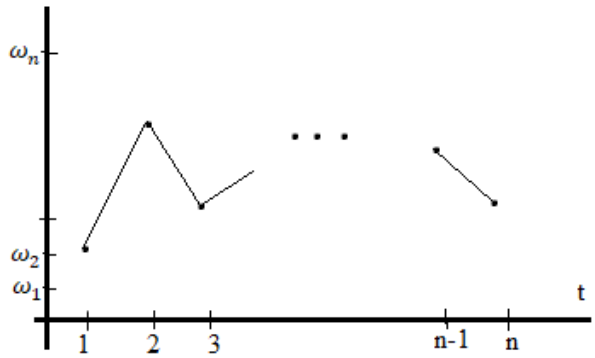


Рис.10.1. Последовательность независимых испытаний.

Конечные однородные цепи Маркова

В каждом испытании (Ω, F, P), $\Omega = \{\omega_1, \dots, \omega_n\}$,

Испытаний n $\{\langle \omega_{i_1}, \dots, \omega_{i_n} \rangle\} = \Omega^n$,

F_n – алгебра всех подмножеств Ω^n,

Вероятность элементарного исхода:

$$P(\langle \omega_{i_1}, \dots, \omega_{i_n} \rangle) = \mathrm{P}(\omega_{i_n} | \omega_{i_{n-1}}, \omega_{i_{n-2}}, \dots, \omega_{i_1}) P(\omega_{i_{n-1}} | \omega_{i_{n-2}}, \dots, \omega_{i_1}) \dots P(\omega_{i_2} | \omega_{i_1}) P(\omega_{i_1}).$$

(с учетом формулы $\mathrm{P(ABC)} = \mathrm{P}(C|AB)\mathrm{P}(B|A)P(A)$)

Положим, что на вероятность влияет только предыдущий исход:

$$\mathrm{P}(\omega_{i_n} | \omega_{i_{n-1}}, \omega_{i_{n-2}}, \dots, \omega_{i_1}) = \mathrm{P}(\omega_{i_n} | \omega_{i_{n-1}}),$$

ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ МГУ ИМЕНИ М.В. ЛОМОНОСОВА
teach-in ЛЕКЦИИ УЧЕНЫХ МГУ

$$P(\omega_{i_{n-1}}|\omega_{i_{n-2}},\dots,\omega_{i_1}) = P(\omega_{i_{n-1}}|\omega_{i_{n-2}}) \text{ и т. д.}$$

$\mathrm{P}(\omega_{i_n}|\omega_{i_{n-1}})$ – переходные вероятности,

$a_1,\dots,a_n$ – начальные вероятности.

$$a_j = P(\omega_j \text{ в нач. момент}).$$

Свойство:

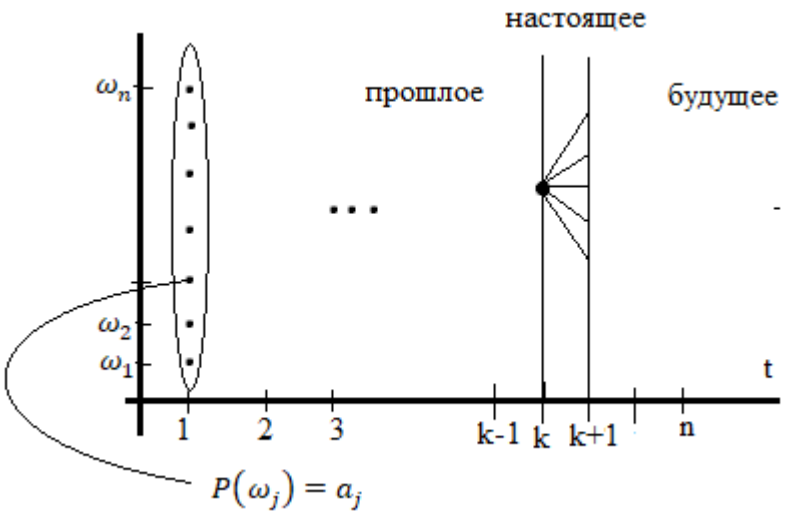


Рис. 10.2. События цепи Маркова.

Если k - момент времени это настоящее, то вероятность попасть в другое состояние в следующий момент времени k+1 не зависит от прошлого.

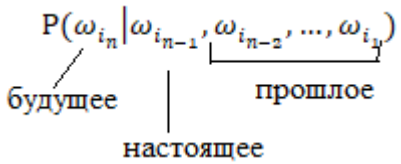


Рис. 10.3

$$\mathrm{P}(\omega_{i_n}|\omega_{i_{n-1}},\omega_{i_{n-2}},\dots,\omega_{i_1}) = \mathrm{P}(\omega_{i_n}|\omega_{i_{n-1}}) \text{ – свойство Маркова.}$$

Опр.: Пусть $\Omega = \{\omega_1,\dots,\omega_n\}$. Вероятностное пространство (Ω^n, F_n, P_{**}) - конечная однородная цепь Маркова, где

$$P(\langle\omega_{i_1},\dots,\omega_{i_n}\rangle) = \mathrm{P}(\omega_{i_n}|\omega_{i_{n-1}})P(\omega_{i_{n-1}}|\omega_{i_{n-2}})\dots P(\omega_{i_2}|\omega_{i_1})a_{n_1}.$$

Где $P(\omega_k|\omega_j)$ – не зависит от номера испытаний (однородность по времени).

Замечание:

Есть последовательность из трех состояний: $\omega_{i_{s-1}}, \omega_{i_s}, \omega_{i_{s+1}}$ (прошлое, настоящее, будущее). Покажем, что событие в прошлом и событие в будущем являются независимыми при известном настоящем для цепи Маркова.

$$P\left(\omega_{i_{s-1}}, \omega_{i_{s+1}} \middle| \omega_{i_s}\right) = \frac{P\left(\omega_{i_{s-1}}, \omega_{i_s}, \omega_{i_{s+1}}\right)}{P\left(\omega_{i_s}\right)} = \frac{P\left(\omega_{i_{s+1}} \middle| \omega_{i_s}\right) P\left(\omega_{i_s}, \omega_{i_{s-1}}\right)}{P\left(\omega_{i_s}\right)} =$$

$$= \frac{P\left(\omega_{i_{s+1}} \middle| \omega_{i_s}\right) P\left(\omega_{i_{s-1}} \middle| \omega_{i_s}\right) P\left(\omega_{i_s}\right)}{P\left(\omega_{i_s}\right)} = P\left(\omega_{i_{s+1}} \middle| \omega_{i_s}\right) P\left(\omega_{i_{s-1}} \middle| \omega_{i_s}\right)$$

События независимы.

Матрицы перехода

Когда цепь Маркова рассматривают как модель системы развивающейся во времени:

Ω – фазовое пространство

$\langle \omega_{i_1}, \dots, \omega_{i_N} \rangle$ – фазовая траектория

$$p_{ij} = P\left(\omega_j | \omega_i\right) - \text{вероятность перехода из } i \text{ в } j$$

$$(\pi)_{ij} = p_{ij} = P\left(\omega_j | \omega_i\right) - \text{элемент матрицы}$$

$$\pi = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1N} \\ p_{21} & p_{22} & \cdots & p_{2N} \\ \vdots & \vdots & p_{ij} & \vdots \\ p_{N1} & p_{N2} & \cdots & p_{NN} \end{pmatrix}$$

Это **матрицы перехода за один шаг** в цепи Маркова. i-я строка – откуда система переходит, j-й столбец – куда переходит система.

$a_j,\ j = 1, \dots, N -$ начальные вероятности.

Свойства матриц перехода:

1) $p_{ij} > 0,\ p_{ij} \leq 1.$
2) $\sum_{j=1}^{N} p_{ij} = 1,$ потому что это равно вероятности перейти хоть куда-нибудь из i, т.к. система обязана перейти в какое-нибудь состояние.

Матрицы, обладающие свойствами 1 и 2 называются *стохастическими.*

Пример 1:

Пусть есть конечное число состояний, и система из состояния k может переходить либо в состояние k+1 с вероятностью p, или же в k-1 с вероятностью q. Если система попадает в стояние 1 или N, то остается в нем.

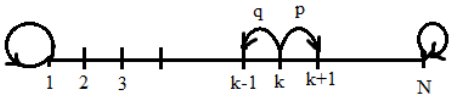


Рис. 10.4. Схема эксперимента.

$$\pi = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ q & 0 & p & 0 & & 0 \\ 0 & q & 0 & p & \cdots & 0 \\ \vdots & & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Пример 2:

Будет ли последовательность независимых испытаний частным случаем цепи Маркова? – Да.

Вероятность оказаться в определенном состоянии не зависит ни от прошлого, ни от будущего:

$$\pi = \begin{pmatrix} p_1 & p_2 & \cdots & p_N \\ p_1 & p_2 & \cdots & p_N \\ \vdots & \vdots & \ddots & \vdots \\ p_1 & p_2 & \cdots & p_N \end{pmatrix}.$$

Переход за n шагов

$$(\pi)_{ij}^{(n)} = P\big(\text{из } \omega_i \text{ перейти в } \omega_j \text{ за } n \text{ шагов}\big)$$

Полная группа попарно независимых событий. Если в момент времени t_k система находится в состоянии i, а в момент t_{k+n}в состоянии j, то вероятность такого перехода $P_{ij}^{(n)}$.

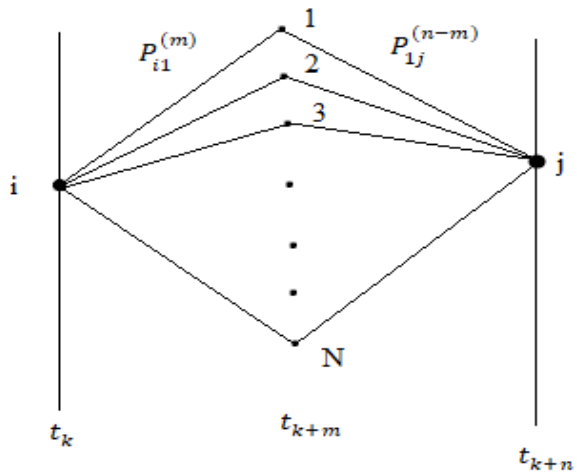


Рис. 10.5. Схема перехода за n шагов.

ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ МГУ ИМЕНИ М.В. ЛОМОНОСОВА
teach-in
ЛЕКЦИИ УЧЕНЫХ МГУ

$$P_{ij}^{(n)} = P_{i1}^{(m)} P_{1j}^{(n-m)} - \text{т. к прошлое и будущее независимы.}$$

$$P_{ij}^{(n)} = \sum_{k=1}^{N} P_{ik}^{(m)} P_{kj}^{(n-m)}.$$

$$\pi^{(n)} = \pi^{(m)} \pi^{(n-m)} = \left(\pi^{(1)}\right)^n.$$

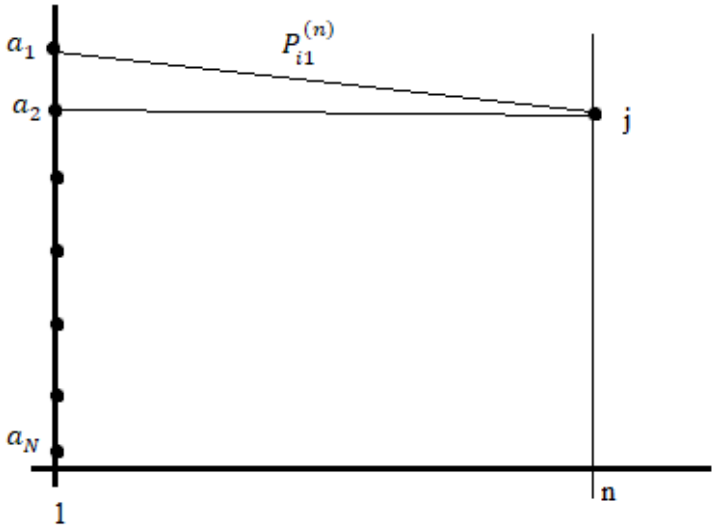


Рис. 10.6. Схема перехода за n шагов.

$$P_j^{(n)} = \sum_{k=1}^{N} a_k P_{kj}^{(n)}.$$

Вероятность того, что система окажется в состоянии j через n шагов после начала наблюдения.

Матрица перехода за n шагов тоже *стохастическая*:

$$\sum_{j=1}^{N} (\pi)_{ij}^{(n)} = \sum_{j=1}^{N} \sum_{k=1}^{N} (\pi)_{ik}^{(n-1)} (\pi)_{kj}^{(1)} = \sum_{k=1}^{N} (\pi)_{ik}^{(n-1)},$$

$$\text{потому что } \sum_{j=1}^{N} (\pi)_{kj}^{(1)} = 1.$$

$$\Rightarrow \sum_{j=1}^{N} (\pi)_{ij}^{(n)} = \sum_{k=1}^{N} (\pi)_{ik}^{(n-1)} = \cdots = \sum_{l=1}^{N} (\pi)_{il}^{(1)} = 1.$$

Эргодичность цепей Маркова

Опр.: Цепь Маркова называется эргодичной, если ∃

ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ МГУ ИМЕНИ М.В. ЛОМОНОСОВА
teach-in ЛЕКЦИИ УЧЕНЫХ МГУ

$$\lim_{n\to\infty} P_{ij}^{(n)} = p_j, \qquad j = 1, \dots, N$$

$$\pi^{(n)} \to \pi^{(\infty)} = \begin{pmatrix} p_1 & p_2 & \cdots & p_N \\ p_1 & p_2 & \cdots & p_N \\ \vdots & \vdots & \ddots & \vdots \\ p_1 & p_2 & \cdots & p_N \end{pmatrix}.$$

Для эргодических цепей переход за n шагов перестает зависеть от того, из какого состояния начался переход. Система со временем забывает свое начальное распределение и стремится к распределению, которое не зависит от времени.

$$P_j^{(n)} = \sum_{m=1}^{N} a_m P_{mj}^{(n)} \xrightarrow[n\to\infty]{} \sum_{m=1}^{N} a_m p_j = p_j.$$

При достижении системой финального распределения еще один шаг по времени не должен ничего поменять.

$$P_j^{(n+1)} = \sum_{k=1}^{N} P_k^{(n)} P_{kj}^{(1)} = \sum_{k=1}^{N} a_k P_{kj}^{(n+1)} \xrightarrow[n\to\infty]{} \sum_{k=1}^{N} a_k p_j = p_j,$$

$$\boxed{p_j = \sum_{k=1}^{N} p_k P_{kj}^{(1)}}.$$

Для финального распределения

$$\vec{p} = (p_1, \dots, p_N)$$

$$\boxed{\vec{p} = \vec{p}\pi^{(1)}}.$$

Однако финальное распределение существует не всегда.

Теорема Перрона:

Если $\pi_{ij}^{(1)} > 0$, для всех i, j, тогда

$$\exists\, \lambda = 1\colon\ \vec{p} = \lambda \vec{p} \pi^{(1)}$$

$$\text{и } \vec{p} = (p_1, \dots, p_N)\colon\ p_j > 0, \qquad j = 1,2, \dots, N\ .$$

Теорема Маркова:

Пусть в конечной однородной цепи Маркова $\pi^{(\vartheta)}$- матрица перехода за ϑ шагов. Если $\exists\, \vartheta$ и номер $j \in \{1, \dots, N\}$ такие, что

$$P_{ij}^{(\vartheta)} \geq \delta > 0\ \ \forall i = 1, \dots, N,$$

$$\text{тогда } \exists \lim_{n\to\infty} P_{ij}^{(n)} = p_j, \qquad j = 1, \dots, N.$$

Т.е. если хотя бы в одной матрице перехода, за какое угодно число шагов найдется хотя бы один столбец, все элементы которого являются ненулевыми, тогда в цепи Маркова существует финальное распределение.

Рассмотрим систему:

$\xi_s = \omega_j$ - в s-тый момент времени система в состоянии ω_j.

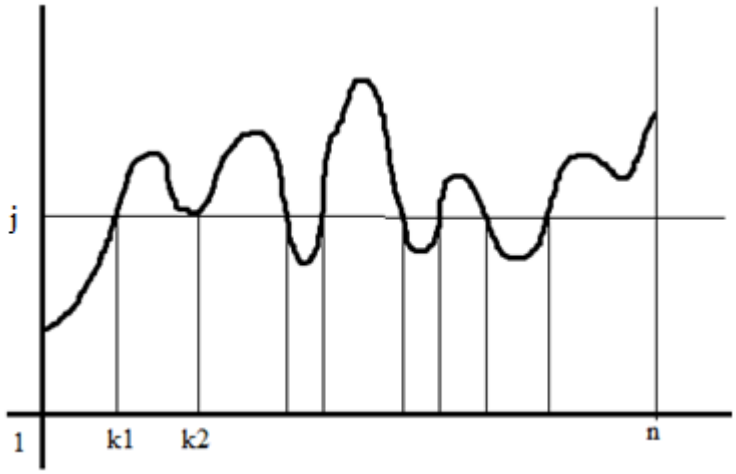


Рис. 10.7. Подсчет состояний, в которых $\xi_s = \omega_j$.

$\tau_j^{(n)}$ – число моментов времени от 1 до n, в которых система была в состоянии ω_j.

$\frac{\tau_j^{(n)}}{n}$ – относительное время, которое система провела в ω_j.

$$\tau_j^{(n)} = \sum_{m=1}^{n} \chi_j^{(m)},$$

$$\chi_j^{(m)} = \begin{cases} 1, & \text{если } \xi_s = \omega_j \\ 0, & \text{если } \xi_s \neq \omega_j \end{cases}.$$

$$M\chi_j^{(m)} = 0 \cdot P(\xi_m \neq \omega_j) + 1 \cdot P(\xi_m = \omega_j) = P(\xi_m = \omega_j).$$

$$M\frac{\tau_j^{(n)}}{n} = \frac{1}{n}\sum_{m=1}^{n} P(\xi_m = \omega_j).$$

Если $P(\xi_m = \omega_j) \xrightarrow[n\to\infty]{} p_j$, то доля времени нахождения в ω_j такое же, как доля систем, находящихся в j-м состоянии среди всех систем.

Лемма:

Пусть $a_n \xrightarrow[n\to\infty]{} a$, тогда

ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ МГУ ИМЕНИ М.В. ЛОМОНОСОВА
teach-in ЛЕКЦИИ УЧЕНЫХ МГУ

$$\frac{1}{n}\sum_{j=1}^{n} a_j \xrightarrow[n\to\infty]{} a.$$

Доказательство:

$$\left|\frac{a_1+a_2+\cdots+a_n}{n}-a\right| = \frac{1}{n}\left|\sum_{j=1}^{n}(a_j-a)\right| \le \frac{1}{n}\left|\sum_{j=1}^{m_0}(a_j-a)\right| + \frac{1}{n}\left|\sum_{j=m_0}^{n}(a_j-a)\right| \le \varepsilon$$

$\frac{1}{n}\left|\sum_{j=m_0}^{n}(a_j-a)\right| < \frac{\varepsilon}{2}$, можем выбрать m_0, для которого выполняется, т. к $a_j \to a$.

$\frac{1}{n}\left|\sum_{j=1}^{n}(a_j-a)\right| < \frac{\varepsilon}{2}$, можем выбрать n, для которого выполняется, т. к. $\frac{1}{n} \to 0$.

Последовательность средних арифметических тоже сходится к пределу a.

$$a_m = P(\xi_m = \omega_j) \xrightarrow[n\to\infty]{} p_j = a.$$

Получается:

$$M\frac{\tau_j^{(n)}}{n} \xrightarrow[n\to\infty]{} p_j.$$

Если $p_j > 0,\ j = 1, \dots, N \implies$ Можно показать, что

$$\frac{\tau_j^{(n)}}{n} \xrightarrow{P} p_j, \qquad \text{при } n \to \infty.$$

ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ МГУ ИМЕНИ М.В. ЛОМОНОСОВА
teach-in
ЛЕКЦИИ УЧЕНЫХ МГУ

ЛЕКЦИЯ 11. ТЕОРИЯ АЗАРТНЫХ ИГР

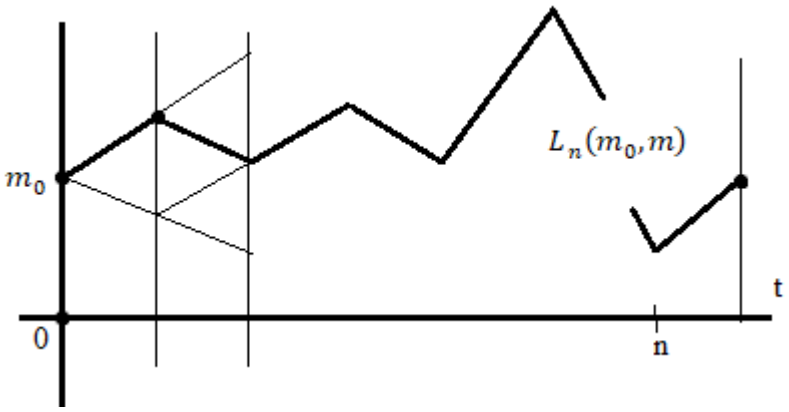


Рис. 11.1. Случайные блуждания.

Случайные блуждания, за один шаг координата может либо увеличиться, либо уменьшиться на единицу.

$$P(\xi_m = k|\xi_{m-1} = l) = \begin{cases} p, & k = l+1 \\ q = 1-p, & k = l-1, \\ 0, & \text{иначе} \end{cases}$$

$$a: P(\xi_0 = m_0) = 1.$$

$L_n(m_0, m)$ – траектория, соединяющая начальный и конечный момент.

Если k шагов сделано вправо, то n-k шагов должно быть сделано влево:

$$P(L: k \text{ вправо}, n-k \text{ влево}) = p^k q^{n-k}.$$

Вероятность всех таких траекторий:

$$P_n(k \text{ вправо}, n-k \text{ влево}) = C_n^k p^k q^{n-k}.$$

Вероятность за n шагов из точки пасть в точку m+d?

$$-n \le d \le n$$

Чему равно d, если сделано k шагов вправо:

$$d = k \cdot 1 + (-1)(n-k) = 2k - n \text{ шагов.}$$

Обратная задача, сколько шагов вправо нужно сделать, чтобы сместиться на d?

$$k = \frac{n+d}{2},$$

$$P_n(m, m+d) = C_n^k p^k q^{n-k}, \qquad k = \frac{n+d}{2}.$$

Теория азартных игр

Если играем в азартную игру: вероятность выигрыша за один кон – p, проигрыша – q. Допустим, мы имеем m рублей, за кон можно выиграть рубль. За n

конов можно остаться в выигрыше или проигрыше. При этом полученная вероятность $P_n(m, m+d) = C_n^k p^k q^{n-k}$ – вероятность выиграть d рублей за k конов.

Если на каком-то шаге мы проиграли все m рублей, то дальше играть не сможем. Поэтому интересна вероятность выиграть d рублей, ни разу не разорившись $\Longrightarrow$ траектории не должны касаться нуля.

Вероятность не прохождения через 0:

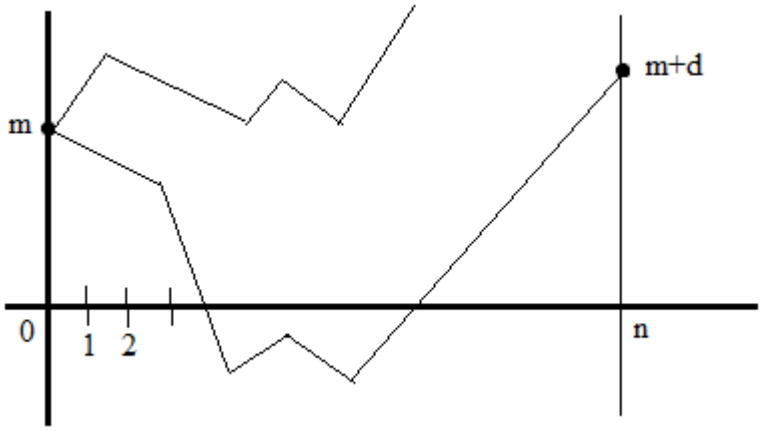


Рис. 11.2. Схема возможного хода игры.

k - вправо, то n-k – влево, $k = \frac{n+d}{2}$ $\Longrightarrow$ чтобы переместиться на четное число шагов, необходимо сделать четное число шагов, чтобы переместиться на нечетное число шагов –нечетное.

$$N_n(m, m+d) - \text{число траекторий, не проходящих через 0.}$$

$$P_n^+(m, m+d) = N_n(m, m+d)p^k q^{n-k} - \text{вероятность без касания 0.}$$

$$\widetilde{N_n}(m, m+d) - \text{число траекторий проходящих через 0.}$$

В таком случае:

$$C_n^k = N_n(m, m+d) + \widetilde{N_n}(m, m+d).$$

Рассмотрим траекторию, которая впервые пересекает ось в точке j:

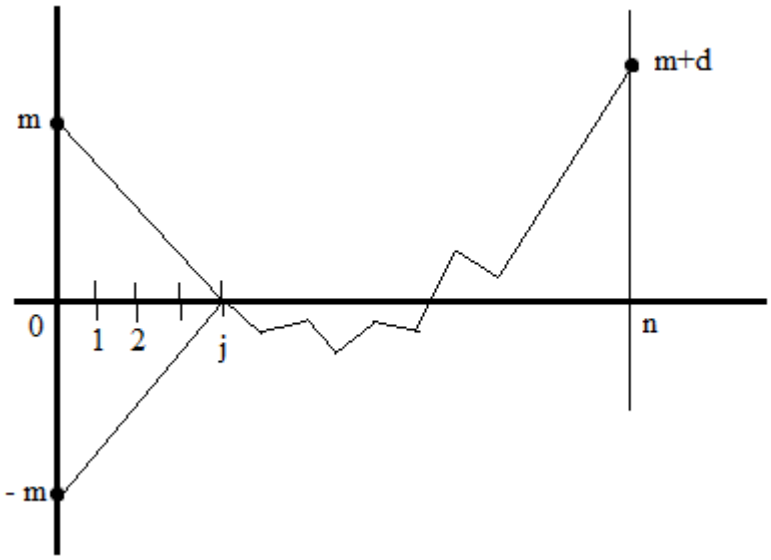


Рис. 11.3. Зеркальное отражение траектории.

Построили траекторию, которая является зеркальным отражением первой. Можно заметить, что число траекторий, которые хоть раз пересекают прямую m=0, равно общему числу траекторий, соединяющих точку –m с точкой m+d за n шагов. Потому что эти траектории обязательно где-то пересекут ось, и каждой траектории $(m, m+d)$ соответствует ровно одна траектория $(-m, m+d)$ и наоборот.

$$\widetilde{N_n}(m, m+d) = N_n(-m, m+d).$$

Выполняя траекторию $(-m, m+d)$, должны сместиться на 2m+d . Число шагов вправо:

$$\tilde{k} + (-1)\left(n - \tilde{k}\right) = 2\mathrm{m} + \mathrm{d}$$

$$\Longrightarrow \tilde{k} = \frac{2m+d+n}{2} = \frac{d+n}{2} + m = k + m.$$

$$C_n^{\tilde{k}} = \widetilde{N_n}(m, m+d).$$

$$P_n^+(m, m+d) = N_n(m, m+d)p^k q^{n-k} = \left(C_n^k - \widetilde{N_n}(m, m+d)\right)p^k q^{n-k} =$$

$$= \left(C_n^k - C_n^{\tilde{k}}\right)p^k q^{n-k}, \qquad \text{где } k = \frac{n+d}{2}, \ \tilde{k} = k + m.$$

Многомерное нормальное распределение

$$\xi \sim N(\mu, \sigma^2) \Longrightarrow P_\xi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$f_\xi(t) = e^{it\mu} e^{-\frac{t^2\sigma^2}{2}}.$$

Опр.: $\vec{\xi} \in R_n$ – случайный вектор, характеристическая функция $\vec{\xi}$:

$$f_{\vec{\xi}}(t_1, \dots, t_n) = Me^{i\left(\vec{t},\vec{\xi}\right)},$$

$$\vec{\xi} = (\xi_1, \dots, \xi_n), \qquad \vec{t} = (t_1, \dots, t_n).$$

Опр.: $\vec{\xi} \sim N(\vec{\mu}, \sum\ )$, если

$$f_{\vec{\xi}}(t_1, \dots, t_n) = e^{i(\vec{t},\vec{\mu})} e^{-\frac{(\vec{t},\sum \vec{t})}{2}},$$

$$\left(\sum_{ij}\right) = \begin{cases} cov(i,j), & i \neq j \\ \sigma_j^2 = D_{\xi_j}, & i = j \end{cases}$$

$$\sum = \sum^* \Longrightarrow$$

$$\exists \text{ ортонормированный базис:} \sum = diag(\sigma_1^2 \dots \sigma_n^2)$$

$$\text{Если} \sum\ - \text{необратима} \Longrightarrow$$

$$\sum = diag(\sigma_1^2 \dots \sigma_k^2, 0, \dots, 0_n)$$

На диагоналях стоят дисперсии координат вектора ξ в заданном базисе. Если дисперсия равна 0, то случайная величина с вероятностью 1 равна своему мат. ожиданию:

$$\left(\xi_j - \mu_j\right) - \text{случайна только для } k \text{ координат вектора } \xi,$$

$$n - k \text{ координат принимают значения 0 с вероятностью 1.}$$

Т.е. вектор принадлежит k мерному линейному подпространству искомого Евклидового пространства.

Если $\sum$ невырожденная $\Longrightarrow$ $\vec{\xi} \sim N(\vec{\mu}, \sum\ )$ имеет плотность вероятности:

$$P_{\vec{\xi}}(x_1, \dots, x_n) = \frac{1}{\sqrt{2\pi}^n \sqrt{det\sum}} \mathrm{e}^{-\frac{1}{2}\left(\vec{x}-\vec{\mu},\sum^{-1}(\vec{x}-\vec{\mu})\right)}.$$

Если $\sum = diag(\sigma_1^2 \dots \sigma_n^2)$ - диагональная матрица$\Longrightarrow$

$$P_{\vec{\xi}}(x_1, \dots, x_n) = \frac{1}{\sqrt{2\pi}^n \sigma_1\sigma_2 \dots \sigma_n} \mathrm{e}^{-\frac{1}{2}\sum_{i=1}^n \frac{(x_i-\mu_i)^2}{\sigma_i^2}} = \prod_{j=1}^n \frac{1}{\sqrt{2\pi\sigma_j^2}} e^{-\frac{\left(x_j-\mu_j\right)^2}{2\sigma_j^2}}.$$

Следствие: для $\vec{\xi} \sim N(\vec{\mu}, \sum\ )$ из некоррелированности координат $\xi_1, \dots, \xi_n$ следует их независимость в совокупности.

Рассмотрим:

$$\vec{\xi} \sim N\left(\vec{0}, \sigma^2 I\right),$$

$$P_{\vec{\xi}}(\vec{x}) = \prod_{j=1}^n \frac{1}{\sqrt{2\pi\sigma_j^2}} e^{-\frac{{x_j}^2}{2\sigma^2}} = \left(\frac{1}{\sqrt{2\pi\sigma_j^2}}\right)^n e^{-\frac{\|\vec{x}\|^2}{2\sigma^2}}.$$

Плотность распределения случайного вектора ξ зависит только от квадрата нормы его аргумента и не зависит от направления вектора в n-мерном Евклидовом пространстве.

Ортогональное преобразование

Опр.: $\cup -$ называется оператором ортогонального преобразования, если он не меняет ни длину, ни угол между векторами:

$$(\vec{x},\vec{y}) = (\cup\,\vec{x},\cup\,\vec{y}),$$

$$(\cup\,\vec{x},\vec{y}) = (\vec{x},\cup^*\,\vec{y}),$$

$$\Longrightarrow (\cup^*\cup\,\vec{x},\vec{y}) = (\vec{x},\vec{y})$$

$$\Longrightarrow \cup^*\cup = I \;\Longrightarrow \cup^* = \cup^{-1}$$

Для ортогональных преобразований сопряжение совпадает с обратным оператором. Оператор ∪ не может быть вырожденным.

$$|\det\cup| = 1$$

R_n: $L -$ линейное подпространство

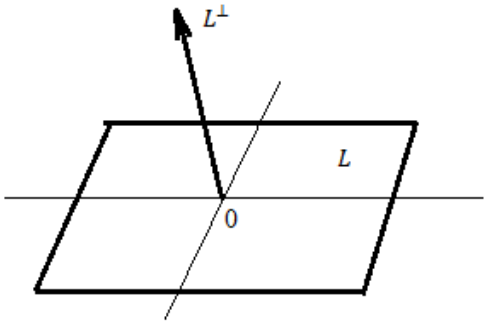


Рис. 11.4. Ортогональное дополнение.

$L^\perp -$ ортогональное дополнение L.

$$\forall x \in R_n \Longrightarrow x = x_1 + x_2,$$

$$x_1 \in L, \qquad x_2 \in L^\perp.$$

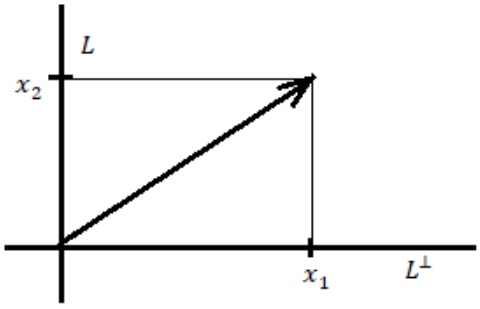


Рис. 11.5. Разложение вектора.

ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ МГУ ИМЕНИ М.В. ЛОМОНОСОВА
teach-in
ЛЕКЦИИ УЧЕНЫХ МГУ

Разложение единственно:

$$\text{Если } x = x_1 + x_2 = y_1 + y_2:$$

$$x_1 \in L, \qquad x_2 \in L^{\perp}, \qquad y_1 \in L, \qquad y_2 \in L^{\perp}, \qquad \text{тогда}$$

$$(x_1 - y_1) + (x_2 - y_2) = 0:$$

$$(x_1 - y_1) \in L, \qquad (x_2 - y_2) \in L^{\perp}.$$

$$\|x_1 - y_1\|^n + \|x_2 - y_2\|^n + 0 = 0,$$

$$\Longrightarrow \|x_1 - y_1\| = 0, \qquad \|x_2 - y_2\| = 0$$

$$\Longrightarrow x_1 = y_1, \qquad x_2 = y_2.$$

Опр.: Оператор П – ортогональный проектор на L:

$$\Pi x = x_1 \in L$$

Свойства П:

1) П – линейный оператор:

$$\Pi(\alpha x + \beta y) = \alpha \Pi x + \beta \Pi y.$$

$$x = x_1 + x_2, \qquad \mathrm{y} = y_1 + y_2$$

$$\Longrightarrow \alpha x + \beta y = \alpha x_1 + \alpha x_2 + \beta y_1 + \beta y_2$$

$$\alpha x_1, \beta y_1 \in L, \qquad \alpha x_2, \beta y_2 \in L^{\perp}.$$

$$\Pi(\alpha x + \beta y) = \alpha x_1 + \beta y_1 = \alpha \Pi x + \beta \Pi y.$$

2) П – самосопряженный оператор:

$$(\Pi x, y) = (x, \Pi y), \qquad \text{т.е.}\, \Pi = \Pi^*$$

$$(\Pi x, y) = (x_1, y) = (x_1, y_1 + y_2) = (x_1, y_1) = (x_1, \Pi y) = (x_1 + x_2, \Pi y) = (x, \Pi y).$$

3) $\Pi^2 = \Pi$:

$$\Pi^2 x = \Pi(\Pi x) = \Pi(x_1) = x_1 = \Pi x.$$

Теорема:

Если для П выполнены свойства 1,2,3, то П – ортогональный проектор.

Доказательство:

$$L = \{x : x = \Pi x\},$$

$$N = \{x \in R_n : \Pi x = 0\}.$$

1) $L \perp N$, т.е. $\forall x \in L, \forall y \in N$:

$$(x, y) = 0.$$

$$(\Pi x, y) = (x, \Pi y) = 0.$$

2) $x = \Pi x + x - \Pi x = \Pi x + (I - \Pi)x.$

$$\Pi x \in L, \qquad \text{т.к.}\, \Pi(\Pi x) = \Pi x \in L.$$

$$(I - \Pi)y \in N, \qquad \text{т.к.}\ \Pi(I - \Pi)y = \Pi y - \Pi^2 y\ = 0.$$

Теорема:

$$\rho^2(x, L) = \min_{y \in L} \|x - y\|^2 = \min_{y \in L} \|x - \Pi x + \Pi x - y\|^2 =$$

$$= \min_{y \in L} \{\|x - \Pi x\|^2 + \|\Pi x - y\|^2 + 0\} = \|x - \Pi x\|^2 + 0, \ \text{при}\ y = \Pi x.$$

$$x - \Pi x = (1 - \Pi)x \in L^{\perp}, \qquad \Pi x - y \in L.$$

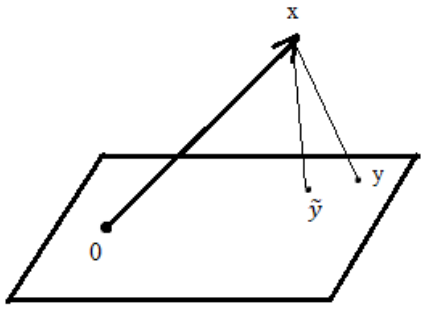


Рис. 11.6. Поиск минимального расстояния.

$$\vec{\xi} \sim N(\vec{0}, \sigma^2 I), \qquad \text{тогда}$$

$$\cup\, \xi \sim N(\vec{0}, \sigma^2 I).$$

Доказано, что если $\eta = A\xi$, и известна плотность распределения $P_\xi(x) \Longrightarrow$

$$P_\eta(y) = P_\xi(A^{-1}x)|\det A^{-1}|.$$

$$A = \cup \Longrightarrow$$

$$P_{\cup_\xi}(x) = P_\xi(\cup^{-1} x)|\det \cup^{-1}| = P_\xi(\cup^{-1} x) \cdot 1.$$

$$P_\xi(\cup\, x) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{\|\cup x\|^2}{2}} = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\frac{\|x\|^2}{2}},$$

$$\text{т.к.}\, (\cup\, x, \cup\, y) = (x, y),$$

$$(\cup\, x, \cup\, x) = \|\cup\, x\|^2 = \|x\|^2.$$

При изменении базиса не изменяется распределение координат:

$$e_j \to \cup\, e_j,$$

$$(\xi, e_j) \to (\xi, \cup\, e_j) = (\cup^*\, \xi, e_j).$$

Рассмотрим:

$$\vec{\xi} \sim N(0, I), \qquad \sigma^2 = 1$$

ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ МГУ ИМЕНИ М.В. ЛОМОНОСОВА
teach-in ЛЕКЦИИ УЧЕНЫХ МГУ

$$\left\|\vec{\xi}\right\|^2 = \sum_{j=1}^{n} \xi_j^2, \qquad \xi_j \sim N(0,1) \text{ и независимы в совокупности.}$$

χ_n^2 - случайная величина, равная сумме квадратов нормально распределенных стандартных независимых случайных величин, называется *распределенной по закону Пирсона* с n степенями свободы.

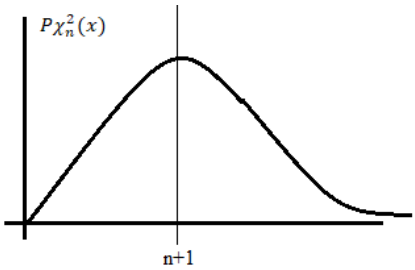


Рис. 11.7. Плотность распределения.

$$M\chi_n^2 = n, \qquad D\chi_n^2 = 2\text{n}, \qquad \text{n} = 1,2,3,\ldots$$

$$\left\|\Pi_k \vec{\xi}\right\|^2 = \chi_k^2.$$

Выберем базис, в котором первые k координат лежат в пространстве L, а оставшиеся n-k в ортогональном дополнении.

$$\left\|\Pi_k \vec{\xi}\right\|^2 = \sum_{\mu=1}^{k} \xi_\mu^2$$

$$\xi_\mu \sim N(0,1), \qquad \mu = 1, \ldots, k,$$

$$\xi_{k+l} = 0, \qquad l = 1, \ldots, n-k.$$

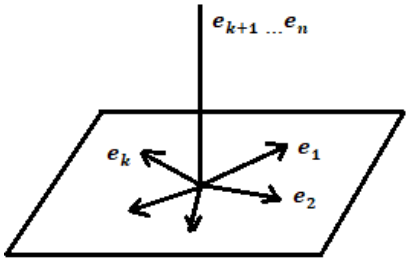


Рис. 11.8.Векторы в пространстве L и ортогональном дополнении.

Опр.:

$$F_{n,m} = \frac{\frac{\chi_n^2}{n}}{\frac{\chi_m^2}{m}} - \text{распределение Снедекора} - \text{Фишера.}$$

ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ МГУ ИМЕНИ М.В. ЛОМОНОСОВА
teach-in
ЛЕКЦИИ УЧЕНЫХ МГУ

$$R_n = L \oplus L^{\perp},$$

$$\Pi_L \xi\,, \quad (I - \Pi_L)\xi,$$

$$\frac{\|\Pi_L \xi\|^2/k}{\|(I - \Pi_L)\xi\|^2/n - k} = F_{k,n-k}.$$

$$\vec{\xi} \sim N(0, \sigma^2 I) \Longrightarrow$$

$$\left\|\vec{\xi}\right\|^2 \sim \sigma^2 \chi_n^2,$$

Если распределение не с единичной дисперсией, то квадрат нормы вектора имеет представленное распределение.

$$\left\|\Pi\vec{\xi}\right\|^2 \sim \sigma^2 \chi_k^2,$$

$$\frac{\left\|\Pi\vec{\xi}\right\|^2/k}{\|(I - \Pi)\xi\|^2/n - k} \sim F_{k,n-k}, \qquad \forall \vec{\xi} \sim N(0, \sigma^2 I), \forall \sigma^2.$$

Опр.:

$$\frac{N(0,1)}{\sqrt{\frac{1}{k}\chi_k^2}} \sim t_k - \text{распределение Стьюдента}$$

Как получит такую дробь? Возьмем:

$$\vec{\xi} \sim N(0, \sigma^2 I), \qquad \vec{e}\colon \|\vec{e}\| = 1.$$

$$\left(\vec{e}, \vec{\xi}\right) \sim N(0, \sigma^2) - \text{координата } \vec{\xi} \text{ в базисе, где } \vec{e} \text{ первый вектор,}$$

$$\left(\vec{\xi} - \left(\vec{e}, \vec{\xi}\right), \vec{e}\right) - \text{ортогональная проекция.}$$

Запишем проектор на вектор $\vec{e}$:

$$\Pi_{\vec{e}}\vec{\xi} = \left(\vec{\xi}, \vec{e}\right)\vec{e}.$$

Проектор на ортогональное дополнение к вектору $\vec{e}$:

$$(1 - \Pi_{\vec{e}})\vec{\xi} = \vec{\xi} - \left(\vec{\xi}, \vec{e}\right)\vec{e}.$$

Тогда:

$$\frac{\left(\vec{\xi}, \vec{e}\right)}{\sqrt{\frac{1}{n-1}\left\|(1 - \Pi_{\vec{e}})\vec{\xi}\right\|^2}} \sim t_{n-1} - \text{распределение Стьюдента } \forall \sigma > 0.$$

Поскольку координата $\vec{\xi}$ вдоль направления вектора $\vec{e}$, и координата $\vec{\xi}$ в ортогональном к $\vec{e}$ направлении независимы.

ЛЕКЦИЯ 12. ТЕОРИЯ ВОЗМОЖНОСТЕЙ

Событие $A \in F$, (Ω, F, P).

Пусть $\chi_j = \begin{cases} 1, \text{ если в } j-\text{м испытании происходит } A \\ 0, \text{ если в } j-\text{м испытании не происходит } A \end{cases}$ – сл. величина.

$\sum_{j=1}^{n} \chi_j$ - число наступлений А в серии из n испытаний. Если испытания независимы, то это сумма независимых случайных величин, одинаково распределенных.

$$\frac{1}{n}\sum_{j=1}^{n} \chi_j = \vartheta_n(A) - \text{частота наступления } A \text{ в длинной серии испытаний.}$$

$$\frac{1}{n}\sum_{j=1}^{n} \chi_j = \vartheta_n(A) \xrightarrow{\text{п.н.}} P(A).$$

Замечание: с этого момента вероятность будем обозначать Pr (probability), чтобы не путать с термином теории возможностей P (possibility).

Если (Ω^n, F^n, Pr_n):

$$Pr_n(\omega_{i_1}, \dots, \omega_{i_n}) = Pr_1(\omega_{i_1}) Pr_2(\omega_{i_2}) \dots Pr_n(\omega_{i_n}).$$

При этом в разных экспериментах разные вероятности. В этом случае:

$$\frac{1}{n}\sum_{j=1}^{n} \chi_j = \vartheta_n(A) \to \frac{1}{n}\sum_{j=1}^{n} Pr_j(A).$$

Частота сходится к некоторой средней вероятности. Но по ней невозможно установить исходную модель вероятностного пространства, данных не хватает. Если вероятность от эксперимента к эксперименту меняется неконтролируемым образом, то теория вероятности оказывается неадекватна в таких задачах.

Возникает интерес к другим способам математического описания эксперимента, исход которого нельзя предсказать заранее. В частности такой теорией является теория возможностей. В теории возможностей мы будем строить теорию на основе стохастической модели, за исходную возьмем вероятностную модель, в которой вероятность каждого исхода от эксперимента к эксперименту меняется, но не произвольным образом.

В теории возможностей возможность контролирует порядок, который характеризует шансы на наступление определенного события, а не частоту. Поэтому содержательными являются утверждения типа: «событие А более возможно, чем событие Б», «событие А менее возможно, чем событие Б», «события равновозможные».

$$\Omega - \text{конечно или счетно,}$$

$$F = P(\Omega) - \text{множество всех возможных подмножеств,}$$

$$pr_j = \Pr(\{\omega_j\}) - \text{вероятнось.}$$

Соглашение:

$$p_j = p(\omega_j) - \text{возможность } \{\omega_j\},$$

$$0 \le p_j \le 1.$$

$$pr_j \ge pr_i \implies p_j \ge p_i\,.$$

Следствие:

$$p_i > p_j \implies pr_i > pr_j\,.$$

$$\text{Пусть } pr_1 \ge pr_2 \ge pr_3 \ge \cdots \ge pr_n \ge \cdots \tag{12.1}$$

$$\sum_{j=1}^{\infty} pr_j = 1.$$

Согласованные с (12.1) возможности:

$$1 = p_1 \ge p_2 \ge p_3 \ge \cdots \ge p_n \ge \cdots \tag{12.2}$$

Это сопоставление не однозначно, в модели есть произвол, так как возможна постановка как знака $\ge$, так и $>$.

Обозначим:

$$p_i = p_{i+1} \implies e_i = 0,$$

$$p_i > p_{i+1} \implies e_i = 1.$$

Запишем двоичную дробь, это параметр, который контролирует упорядоченность:

$$\bar{e} = 0, e_1 e_2 \dots e_n \dots \ \in [0,1].$$

Меняя числовые значения возможностей, но сохраняя при этом упорядоченье, получаем множество эквивалентных возможностных моделей.

Множество способов задания возможности элементарных исходов в классе с определенным упорядоченьем: $P_{\bar{e}}$.

Важно, чтобы при монотонном преобразовании значений возможностей не менялась их упорядоченность.

Операции с возможностями

Если есть значения возможности, равное 0, и значение другой возможности, равное a. Надо пересчитать в значение возможности, которое состоит из одного и другого:

$$0 + a = a;$$
$$1 + a = 1.$$

Рассмотрим возможность одновременного наступления событий:

$$0 \cdot a = 0;$$
$$1 \cdot a = a.$$

Разумно сделать также следующие предположения: (12.3)

$$a + b = b + a;$$
$$a \cdot b = b \cdot a.$$

Монотонное преобразование обладает свойствами:

$$\gamma(a) + \gamma(b) = \gamma(a + b)$$
$$\gamma(a) \cdot \gamma(b) = \gamma(a \cdot b)$$

Теорема:

Если (12..3) выполняется для монотонно возрастающих функций $\gamma: \gamma(0) = 0$, $\gamma(1) = 1$, и сами операции + и · тоже непрерывны, то

$$a + b = \max\{a, b\},$$
$$a \cdot b = min\{a, b\}.$$

Пример 1:

Вы пришли в буфет, но забыли деньги. У вас есть два друга, у которых возможность иметь 100 рублей в кармане у одного 0.8, а у другого 0.9. Вы встречаете их обоих, какова возможность пообедать?

Для возможности нужно брать максимальную возможность из этих двух, если у одного возможность 0.9, у другого 0.8, когда вы встречаетесь вместе, возможность пообедать 0.9.

Пример 2:

Вы пришли домой, но без ключей. На двери 2 замка, у одного вашего друга есть ключ от верхнего замка, у другого - от нижнего. Нужно, чтобы пришли оба ваших друга, и никто из них не забыл про этот ключ. Один не забудет с возможностью 0.1, другой с возможностью 0.05. Возможность попасть домой - это минимум из этих двух возможностей 0.05.

Доказательство:

$$0 \le x < y < 1,$$

$\gamma_n(\cdot)$ - монотонные функции, n=1,2,3,…:

$$\gamma_n(z) \to 0 \ \ \forall\, 0 \le z \le x,$$
$$\gamma_n(z) = z, \ \ \forall\, y \le z \le 1.$$

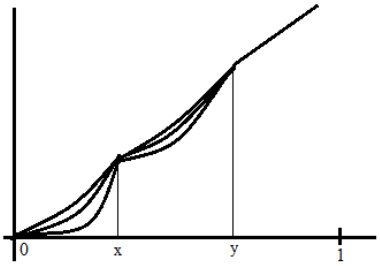


Рис. 12.1. Монотонные функции.

$$\Longrightarrow \gamma_n(x+y) = \gamma_n(x) + \gamma_n(y),$$

$$\text{т. к. } \gamma_n(x) \xrightarrow[n\to\infty]{} 0, \qquad \gamma_n(y) = y.$$

$$\Longrightarrow \gamma_n(x) + \gamma_n(y) \to \gamma_n(y) = y.$$

$$\Longrightarrow \mathrm{x} + \mathrm{y} = \max(x, y).$$

Аналогично можно доказать и для умножения.

L – шкала значения возможностей:

$$L = \{[0,1], "\geq", +_{=max}, -_{=min}\}.$$

Возьмем $A = \{\omega_{i_1}, \dots, \omega_{i_n}, \dots\}$:

$$P(A) = \max_{j:\omega_j \in A} p_j = +_{j:\omega_j \in A} p_j.$$

$$A \neq \emptyset, \qquad P(A \neq \emptyset) = 0.$$

Следствия:

1) $A \subset B,\ P(A) \leq P(B).$
2) $A, B,\ A \cup B$:

$$P(A \cup B) = max\{P(A), P(B)\} = P(A) + P(B).$$

(Знак «+» в значении max.)

3) $A, B,\ A \cap B$:

$$P(A \cap B) \leq min\{P(A), P(B)\}.$$

Не строгое равенство, потому что может случиться следующее: наиболее возможное событие, которое определяет A, и наиболее возможное событие, которое определяет B, не содержатся в пересечении. События, находящиеся в пересечении, имеют возможность меньше чем P(A) и P(B).

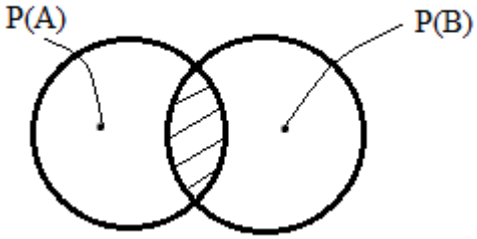


Рис. 12.2. Максимально возможные события из пересекающихся множеств.

В силу того, что:

$$1 = p_1 \geq p_2 \geq p_3 \geq \cdots \geq p_n \geq \cdots$$

$p_1 = 1$ – самый возможный элементарный исход имеет возможность 1.

$$P(A) + P(\Omega \backslash A) = 1$$

Т.к. самое возможное элементарное событие попадает либо в A, либо в $\Omega \backslash A$, а мы ищем максимум, то возможность одного события равна единице, а второго не определена.

В полученной системе возможность противоположного события не определена, поэтому вводим меру необходимости.

Необходимость А:

$$1 - P(\Omega \backslash A).$$

Невозможность события А определяется монотонно убывающей функцией:

$$N(A) = \theta\big(P(\Omega \backslash A)\big),$$

$$\theta(0) = 1, \qquad \theta(1) = 0.$$

Пример:

$$A: P(A) = 1$$

Среди элементарных событий А есть первое элементарное событие, с возможностью 1:

$$A = \{\omega_1, \dots\},$$

$$\Omega \backslash A = \{\omega_k, \dots\}.$$

Возможность события $\Omega \backslash A$ это p_k. Чем p_k меньше, тем больше необходимость А.

$\Longrightarrow$ <u>Чем больше событий с большой возможностью входят в А, тем больше необходимость этого события.</u>

С другой стороны:

$$\text{Если } P(A) < 1 \implies \omega_1 \notin A \implies \Omega \backslash A = \{\omega_1, \dots\},$$

$$P(\Omega \backslash A) = 1 \implies N(A) = 0.$$

Упорядоченность возможностей

Рассмотрим событие:

$$A_1 = \{\omega_1, \dots\},$$
$$P(A_1) = 1, \qquad Pr(A_1) \in [pr_1, 1].$$

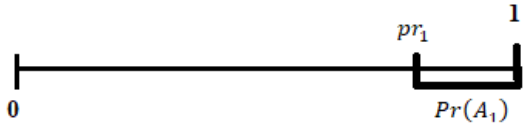


Рис. 12.3 Интервал вероятности для первого события.

$$A_2 = \{\omega_2, \dots\}, \qquad \omega_1 \notin A_2$$
$$P(A_2) = p_2, \qquad Pr(A_2) \in [pr_2, 1 - pr_1].$$

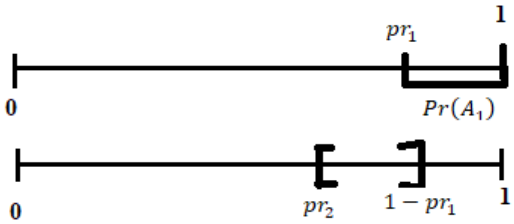


Рис. 12.4. Сопоставление интервалов двух событий.

Интервалы могут пересекаться, или не пересекаться. Если интервалы не пересекаются, тогда вероятность события A_1 всегда больше вероятности A_2. Тогда возможности связаны таким же строгим неравенством.

Чем больше строгих неравенств, тем больше возможность различать по возможности события. В случае нестрогого неравенства события придется считать равновозможными.

$$A_n = \{\omega_n, \dots\}, \qquad \omega_1, \dots, \omega_n \notin A_n$$
$$P(A_n) = p_n, \qquad Pr(A_n) \in [pr_n, 1 - pr_1 - \dots - pr_{n-1}].$$

Таким образом, условие не пересечения интервалов:

$$e_i = 1 \Leftrightarrow pr_1 + \dots + pr_{i-1} + 2pr_i > 1,$$
$$p_i > p_{i-1}.$$

И наоборот:

$$e_i = 0 \Leftrightarrow pr_1 + \dots + pr_{i-1} + 2pr_i \leq 1,$$
$$p_i = p_{i-1}.$$

Возможностная модель может быть установлена точно, с вероятностью 1 за конечное число шагов.

Пример:

$$pr_j = \frac{q}{(1+q)^j}, \qquad j = 1,2,\ldots$$

$$\text{Если } 0 \le q \le 1, \qquad \text{то все } e_i = 0.$$

$$\text{Если } q > 1, \qquad \text{то все } e_i = 1.$$

ЛЕКЦИЯ 13. ОСНОВЫ ТЕОРИИ ВОЗМОЖНОСТЕЙ

Повторение материала предыдущей лекции

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_n, \dots\},$$

$$pr_1 \geq pr_2 \geq pr_3 \geq \dots \geq pr_n \geq \dots, \sum_{j=1}^{\infty} pr_j = 1,$$

$$\Longrightarrow 1 = p_1 \geq p_2 \geq p_3 \geq \dots \geq p_n \geq \dots - \text{по соглашению.}$$

Распоряжаемся возможностями иначе, чем вероятностями:

$$p_1 + p_2 \equiv \max(p_1, p_2),$$

$$p_1 \cdot p_2 \equiv \min(p_1, p_2).$$

Бинарные операции инвариантны относительно монотонного преобразования значений возможностей. То событие, что было максимальным, остается максимальным, минимальное остается минимальным.

$$\Pr(A) = \sum_{j:\, \omega_j \in A}^{\infty} pr_j.$$

$$P(A) = \max_{j:\, \omega_j \in A} p_j.$$

При таком задании возможностей:

$$P(A \cup B) = max\{P(A), P(B)\}.$$

$$P(A \cap B) \leq min\{P(A), P(B)\}.$$

Неравенство, потому что возможна ситуация, как на рис. 13.1.

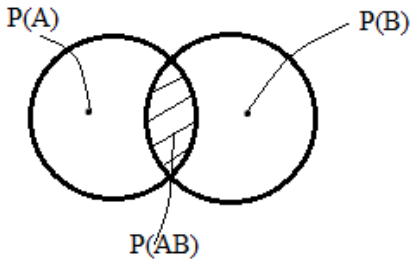


Рис. 13.1. Максимально возможные события из пересекающихся множеств.

Рассматривали классы событий:

$$A_1 = \{\omega_1, \dots\},$$

$$P(A_1) = p_1 = 1, \qquad Pr(A_1) \in [pr_1, 1].$$

$$A_2 = \{\omega_2, \dots\}, \qquad \omega_1 \notin A_2$$

$$P(A_2) = p_2, \qquad Pr(A_2) \in [pr_2, 1 - pr_1].$$

Если интервалы для вероятностей не пересекаются, тогда вероятность события A_1 всегда больше вероятности A_2. Тогда возможности связаны таким же строгим неравенством. Соглашение выполняется.

В случае пересечения вероятностных интервалов события придется считать равновозможными, их нельзя различить.

Построение теории возможностей

Для теории возможностей аналогом случайной величины является нечеткая величина. В теории вероятностей у нас есть:

$$(\Omega, F, Pr).$$

Этому вероятностному пространству сопоставляем:

$$\left(R_1, B, F_\xi\right).$$

Связь между вероятностными пространствами обеспечивается:

$$\xi(\omega) \in R_1,$$

$$\text{если } A \in F, \text{то } \xi(A) \in B.$$

Для этого достаточно, чтобы:

$$\forall x \in R_1 \ \{\omega: \xi(\omega) < x\} \in F.$$

Построение нечеткой величины следует этой же схеме:

$$(\Omega, P(\Omega), P, N).$$

Нечетким элементом называется любая функция, определенная на Ω и принимающая числовые значения:

$$\eta: \ \Omega \to R_1.$$

Зададим распределение возможности для нечеткого элемента:

$$g^\eta(x) \equiv P(\eta = x) = \sup_{\omega:\eta(\omega)=x} P(\{\omega\}), \qquad \forall x \in R_1.$$

$$A: \ \chi_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases} - \text{индикаторная функция } A.$$

$$P(\eta \in A) = \sup_{x \in A} g^\eta(x) = \sup_{x \in R_1} \min\left(\chi_A(x), g^\eta(x)\right) = \sup_{x \in R_1} \chi_A(x) \cdot g^\eta(x) =$$

$$= \underset{x \in R_1}{+} \left(\chi_A(x) \cdot g^\eta(x)\right).$$

$$\Pr(\xi \in A) = \int_{R_1} \chi_A(x) P_\xi(x) dx.$$

Формально записи для вероятности и возможности совпадают, только умножение заменяется на · (min), а вместо интегрирования + в значении max.

Пример:

ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ МГУ ИМЕНИ М.В. ЛОМОНОСОВА
teach-in ЛЕКЦИИ УЧЕНЫХ МГУ

Как задать нечеткий элемент равный длине короткого отрезка? Нужно задать возможность того, что отрезок длины x можно считать коротким:

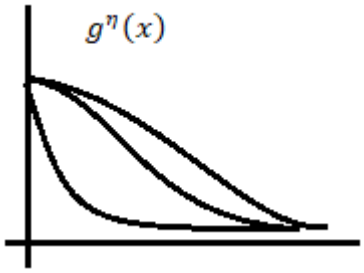


Рис.13.2. Распределение возможности.

Все кривые подходят: чем меньше x, тем больше возможность считать отрезок коротким.

Нечеткие элементы можно преобразовывать по определенному закону.

$$g^{\eta}(x) - \text{распределние возможностей элемента } \eta,$$

$$\alpha = a\eta + b,$$

$$g^{\alpha}(y) = P(\alpha = y) = P(a\eta + b = y) = P\left(\eta = \frac{y-b}{a}\right) = g^{\eta}\left(\frac{y-b}{a}\right).$$

Нечеткие множества

$$\mu^{A}(x) - \text{возможность } x \in A.$$

$$\mu^{A}(x) = \chi_{A}(x) - \text{"четкое" множество},$$

$$\mu^{A}(x) \in [0,1] - \text{"нечеткое" множество}.$$

$$P(\eta \in A) = \sup\{min\{\mu^{A}(x), g^{\eta}(x)\}\}.$$

Р-независимые события. Условная возможность.

События А и В – попарно Р-независимы, если

$$P(AB) = min\{P(A), P(B)\}.$$

События $A_1, A_2, \dots, A_n$ - Р-независимые в совокупности, если

$$\forall k = 2,3, \dots, n \text{ и } \forall i_1, i_2, \dots, i_k \in \{1, \dots, n\}$$

$$P\left(A_{i_1}, A_{i_2}, \dots, A_{i_k}\right) = \min_{j=1,\dots,k}\left\{P(A_{i_j})\right\}.$$

Из попарной независимости не следует Р-независимость в совокупности.

Пример 1:

$$\Omega = \{\omega_1, \omega_2, \omega_3\}, \qquad P(\omega_j) = 1, \qquad j = 1,2,3$$

ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ МГУ ИМЕНИ М.В. ЛОМОНОСОВА
teach-in ЛЕКЦИИ УЧЕНЫХ МГУ

$$A_1 = \{\omega_1, \omega_2\}, \qquad A_2 = \{\omega_2, \omega_3\}, \qquad A_3 = \{\omega_1, \omega_3\}$$

События A_1, A_2, A_3- попарно независимы:

$$P(A_i A_j) = 1 = min\{P(A_i), P(A_j)\}.$$

$$P(A_1 A_2 A_3) = P(\emptyset) = 0 \neq \min_{i=1,2,3}\{P(A_i)\} = 1.$$

Пример 2:

$$P(A_1 A_2 A_3) = \min_{i=1,2,3}\{P(A_i)\},$$

то A_1, A_2, A_3 – не будут независимы в совокупности.

$$\Omega = \{\omega_1, \omega_2, \omega_3\},$$

$$P(\omega_1) = 1, \qquad P(\omega_2) = 1, \qquad P(\omega_3) = 0{,}5.$$

$$A_1 = \{\omega_1, \omega_3\}, \qquad A_2 = \{\omega_3\}, \qquad A_3 = \{\omega_2, \omega_3\}.$$

$$\text{тогда } P(A_1) = 1, \qquad P(A_2) = 0{,}5, \qquad P(A_3) = 1.$$

$$P(A_1 A_2 A_3) = 0{,}5 = \min_{i=1,2,3}\{P(A_i)\},$$

$$P(A_1 A_2) = 0{,}5 = \min\{P(A_1, A_2)\},$$

$$P(A_1 A_3) = 0{,}5 \leq \min\{P(A_1, A_3)\} = 1.$$

Из того, что $P(A_1 A_2 A_3) = \min_{i=1,2,3}\{P(A_i)\}$, не следует даже попарная независимость.

Определим нечеткий вектор:

$$(\xi, \eta)$$

$$g^{\xi,\eta}(x, y) = P(\xi = x, \eta = y).$$

Можно вычислить маргинальные распределения возможностей:

$$g^{\xi}(x) = \sup_{y \in R_1} g^{\xi,\eta}(x, y),$$

т.к. это возможность того, что $\xi = x$, а у любое другое значение из R_1. По правилам комбинирования возможностей нужно выбрать максимально возможное событие.

Опр.: нечеткие элементы ξ, η называются P-независимыми, если

$$g^{\xi,\eta}(x, y) = g^{\xi}(x) \cdot g^{\eta}(y) = \min\left(g^{\xi}(x), g^{\eta}(y)\right).$$

Перейдем к определению условных возможностей.

В теории вероятности:

$$Pr(A|B) = \frac{Pr(AB)}{Pr(B)},$$

Или это решение уравнения: $Pr(AB) = Pr(A|B)Pr(B)$.

В теории возможностей:

Опр.: Вариантом условной возможности $P(A|B)$ называется любое решение уравнения: $P(AB) = P(A|B) \cdot P(B) = min\{P(A|B), P(B)\}$. Это уравнение имеет множество решений, любое решение является вариантом условной возможности.

$$\text{т. к. } P(AB) \leq min\{P(A), P(B)\}:$$

$$P(AB) \leq P(B)\,.$$

$$\text{либо } P(AB) = P(B) \Longrightarrow P(A|B) \geq P(B),$$

$$\text{либо } P(AB) < P(B) \Longrightarrow P(A|B) = P(AB).$$

Пример:

$$\Omega = \{\omega_1, \omega_2, \omega_3\},$$

$$pr_1 = 0.7, \qquad pr_2 = 0.2, \qquad pr_3 = 0.1$$

$$P(A_1) = 1 = p_1, \qquad \mathrm{Pr}(A_1) \in [0.7,1],$$

$$P(A_2) = p_2, \qquad \mathrm{Pr}(A_2) \in [0.2, 0.3],$$

$$P(A_3) = p_3, \qquad \mathrm{Pr}(A_3) \in 0.1.$$

Интервалы не пересекаются. Задана максимально согласованная с вероятностью возможностная модель. Если будем менять вероятности так, чтобы интервалы не пересекались, то возможностная модель будет сохраняться.

Пример:

$$\Omega = \{\omega_1, \omega_2, \omega_3\},$$

$$pr_1 = 0.5, \qquad pr_2 = 0.3, \qquad pr_3 = 0.2$$

$$P(A_1) = 1 = p_1, \qquad \mathrm{Pr}(A_1) \in [0.5,1],$$

$$P(A_2) = p_2, \qquad \mathrm{Pr}(A_2) \in [0.3, 0.5],$$

$$\Longrightarrow\ 1 = p_1 = p_2 > p_3.$$

Пример:

Пусть заданы нечеткие элементы $\varphi,\ \vartheta$,

$$\xi = \varphi + \vartheta - \text{сумма полезного сигнала и погрешности.}$$

$\varphi,\ \vartheta$ – независимы и P-независимы.

ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ МГУ ИМЕНИ М.В. ЛОМОНОСОВА
teach-in ЛЕКЦИИ УЧЕНЫХ МГУ

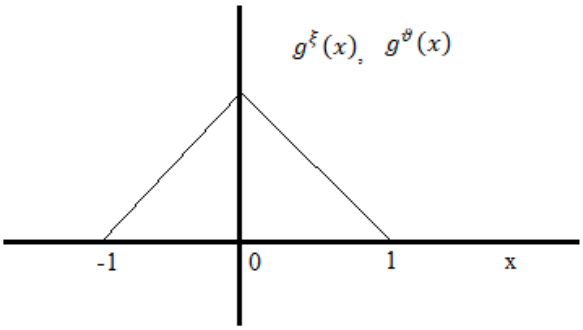


Рис. 13.3. Распределение возможности.

1) $g^{\xi}(z)-?$
2) $g^{\varphi|\xi}(x|1)-?$
3) $\varphi = x$, для которого $g^{\varphi|\xi}(x|1)$ максимальна - ?

$$g^{\varphi,\vartheta,\xi}(x,y,z) = \begin{cases} min\{g^{\varphi}(x), g^{\vartheta}(y)\}, & x+y=z \\ 0, & x+y \neq z \end{cases}$$

Применим формулу для маргинальных распределений.

$$g^{\varphi,\xi}(x,z) = \sup_{y \in R_1}\{g^{\varphi,\vartheta,\xi}(x,y,z)\} = min\{g^{\varphi}(x), g^{\vartheta}(z-x)\},$$

$$g^{\xi}(z) = \sup_{x \in R_1}\{g^{\varphi,\xi}(x,z)\}$$

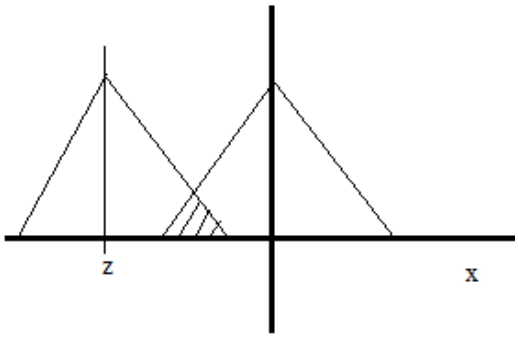
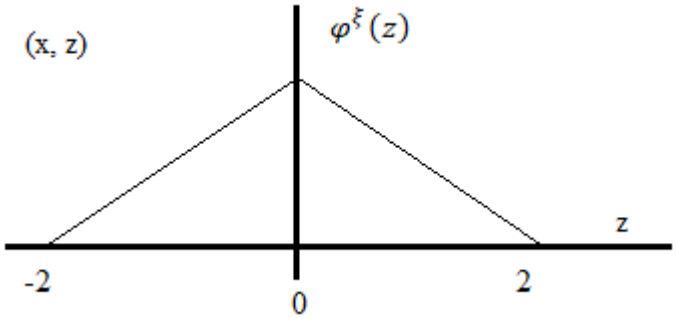


Рис. 13.4. Распределение возможности.



ФИЗИЧЕСКИЙ ФАКУЛЬТЕТ МГУ ИМЕНИ М.В. ЛОМОНОСОВА
teach-in ЛЕКЦИИ УЧЕНЫХ МГУ

Рис. 13.5. Распределение возможности.

$$g^{\varphi,\xi}(x,1) = min\{g^{\varphi|\xi}(x|1), g^{\xi}(z)\}.$$

$$g^{\varphi|\xi}(x|z) = \begin{cases} g^{\varphi,\xi}(x,1), & g^{\xi}(z) > g^{\varphi,\xi}(x,z) \\ \geq g^{\varphi,\xi}(x,1), & g^{\xi}(z) = g^{\varphi,\xi}(x,z) \end{cases}$$

$\boxed{x = \frac{1}{2}}$ – при котором $g^{\varphi|\xi}(x|1)$ максимальна.

ФИЗИЧЕСКИЙ
ФАКУЛЬТЕТ
МГУ ИМЕНИ
М.В. ЛОМОНОСОВА

teach-in
ЛЕКЦИИ УЧЕНЫХ МГУ