Булинский А.В.
Шашкин А.П.

Предельные теоремы для ассоциированных случайных полей и родственных систем

МОСКВА
ФИЗМАТЛИТ ®

УДК 519
ББК 22.171
Б 90

Булинский А. В., Шашкин А. П. **Предельные теоремы для ассоциированных случайных полей и родственных систем.** — М.: ФИЗМАТЛИТ, 2008. — 480 с. — (Теория вероятностей и математическая статистика). — ISBN 978-5-9221-0969-7.

Монография посвящена исследованию асимптотических свойств широкого класса стохастических моделей, возникающих в математической статистике, теории перколяции, статистической физике и теории надежности. В книге содержится множество разнообразных примеров упомянутых моделей, описываемых с помощью марковских процессов, случайных мер, устойчивых законов, ферромагнетиков Изинга, процессов с локальным взаимодействием, случайных графов. Устанавливаются основные предельные теоремы теории вероятностей, а также даются их различные применения. Монография представляет собой первое замкнутое и детальное изложение материала, накопленного в изучаемой области за весь период развития вплоть до настоящего времени. Часть текста основана на лекциях, читавшихся авторами в МГУ им. М. В. Ломоносова.

Для научных работников, профессорско-преподавательского состава, аспирантов, студентов старших курсов математических специальностей университетов.

Ил. 32. Библиогр. 450 назв.

ISBN 978-5-9221-0969-7

ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ

Понятие независимости случайных величин — одно из важнейших в теории вероятностей. Можно сказать, что многочисленные глубокие и красивые результаты, установленные для систем независимых случайных величин, составляют ядро современной теории вероятностей. Однако еще в XIX, а затем в XX веке появились интересные стохастические модели, использующие зависимые величины. Такие модели возникли в физике, химии, биологии, экономике, технических науках. В связи с этим, а также вследствие внутреннего развития математических идей, началось становление общей теории случайных процессов и полей.

Известны такие важные классы случайных процессов и полей, как гауссовские, марковские, мартингалы, системы с перемешиванием разных видов и другие. Для каждого из них разработаны свои методы изучения.

В 1960-е годы появился новый класс положительно (а позднее и отрицательно) зависимых случайных систем, введенный в основополагающих работах Харриса, Лемана, Изери, Прошана, Уолкапа, Фортуина, Кастелейна, Жинибра, Алама, Саксена и Йоаг-Дева. Интерес к этому классу связан с его широкими приложениями в математической статистике, теории надежности, теории перколяции, статистической физике. Самое главное понятие здесь — ассоциированность системы случайных величин, частным случаем которого оказывается независимость.

Начиная с выдающейся статьи Ньюмена (1980), все последующие годы продолжается активная работа по доказательству классических предельных теорем теории вероятностей для ассоциированных систем и их модификаций. Установлены законы больших чисел (ЗБЧ), центральная предельная теорема (ЦПТ), законы повторного логарифма (ЗПЛ), принципы инвариантности и другие предельные закономерности. Удобство работы с положительно и отрицательно ассоциированными случайными величинами заключается в простоте условий, при которых доказывается большинство предельных теорем. Именно, обычно предполагается, что у случайных величин существует абсолютный момент порядка s (как правило, $s \in (2,3]$), и налагаются ограничения на поведение ковариационной функции, например, для стационарных полей – на скорость ее убывания при росте аргументов.

Задача данной книги — послужить введением в эту обширную область исследований, излагающим основные результаты и методы, накопленные за весь период исследований вплоть до наших дней. Дается много примеров разного уровня сложности. Авторы стремились привести детальные доказательства со всеми вспомогательными фактами, по возможности упрощая рассматриваемые работы. В библиографии содержится 450 ссылок, при этом список не претендует на полноту. Слово “родственные” в названии книги имеет смысл “родственные ассоциированным случайным полям”, поскольку исходное понятие ассоциированности допускает различные обобщения, и некоторым из них мы уделяем значительное внимание. В каждой главе даются ссылки для дальнейшего чтения, так как объем книги не позволяет включить ряд интересных теорем.

Полученные результаты применяются в главе 7 при построении приближенных доверительных интервалов для неизвестного среднего стационарного поля, а также при анализе ядерных оценок плотностей. Глава 8 демонстрирует использование развитых методов при изучении асимптотического поведения преобразованных решений многомерного уравнения Бюргерса со случайными начальными данными.

В книге имеется приложение, состоящее из шести разделов, где выводится равенство Хошневисана–Льюиса, которое обобщает классическую формулу Хёфдинга, даются простейшие сведения по марковским процессам и пуассоновским потокам, предназначенные для построения примеров зависимых случайных полей (и процессов), излагаются вспомогательные факты из теории графов, доказывается неравенство Морица и приводятся элементы теории сильного приближения (реконструкции) случайных векторов.

Главы делятся на параграфы и пункты (подпараграфы). В каждом параграфе теоремы, леммы, следствия, определения, примеры и замечания нумеруются последовательно. Ссылки на выключные формулы имеют вид (Параграф.Формула), если указанная формула находится в текущей главе, и (Глава.Параграф.Формула) в других случаях. Так же даются ссылки на теоремы и леммы. Например, теорема 1.3.2 — это теорема 2 в § 3 главы 1. При ссылках на приложение вместо номера главы мы пишем букву П, например, (П.Параграф.Формула). Знак □ отмечает конец доказательства.

Книга адресована математикам, ведущим исследования в области современной теории вероятностей и ее применений. Кроме того, она может быть полезной профессорско-преподавательскому составу университетов для чтения специальных курсов и проведения специальных семинаров. Ряд разделов излагался авторами в лекциях, читавшихся ими на механико-математическом факультете МГУ.

Авторы выражают глубокую признательность своим друзьям и коллегам, с которыми они обсуждали вопросы теории предельного поведения случайных процессов и полей, особенно профессорам В.И.Богачёву, А.А.Боровкову, Р.Брэдли, М.-К.Виано, Ю.А.Давыдову, Ж.Дедекеру, М.Деккингу, П.Дукану, Ж.Жакоду, В.М.Золотареву, И.А.Ибрагимову, М.Иосифеску, М.Кину, Ф.Комецу, В.Ю.Королёву, В.С.Королюку, С.Б.Куксину, И.А.Курковой, Н.Н.Леоненко, М.А.Лифшицу, С.Луиши, П.Матуле, П.Младеновичу, И.С.Молчанову, С.А.Молчанову, Я.Ю.Никитину, О.Пенроузу, В.В.Петрову, В.И.Питербаргу, Ю.В.Прохорову, Дж.Руссасу, Г.Самородницкому, М.Соренсену, Й.Стоянову, Ш.Сюкэ, К.М.Ханину, А.Ю.Хренникову, М.Чёргё, А.Н.Ширяеву, А.Якубовскому.

Работа поддержана грантами РФФИ 08-01-07108-д и 07-01-00373-а.

А.В.Булинский, А.П.Шашкин

Кафедра теории вероятностей
механико-математического факультета
Московского государственного университета
им. М.В.Ломоносова

Нашим родителям

Глава 1

СЛУЧАЙНЫЕ ЭЛЕМЕНТЫ И КОВАРИАЦИОННЫЕ НЕРАВЕНСТВА

В этой главе рассматриваются основные идеи и результаты, относящиеся к свойствам положительной и отрицательной зависимостей и их обобщениям. Также даются разнообразные примеры, иллюстрирующие полезность вводимых понятий. Из результатов следует отметить теорему Питта — критерий ассоциированности гауссовской системы, и теорему Ли, Рачева и Самородницкого, которая устанавливает необходимые и достаточные условия ассоциированности устойчивого случайного вектора в терминах его спектральной меры. Среди примеров особое внимание уделяется марковским процессам и случайным мерам, в частности, решениям стохастических дифференциальных уравнений, полям дробового шума и кластерным мерам. Так, доказывается теорема Бёртона–Уэймира–Эванса об ассоциированности безгранично делимой случайной меры, заданной на польском пространстве. Отдельно рассмотрены векторнозначные случайные поля и случайные элементы со значениями в частично упорядоченных пространствах. Подробно изучаются знаменитые ФКЖ-неравенства Фортуина, Кастелейна и Жинибра, а также общие теоремы Холли и Престона. Они играют большую роль в теории перколяции и статистической физике, из моделей которой мы коснемся ферромагнетиков Изинга и систем с локальным взаимодействием. Отрицательная ассоциированность встречается не только при исследовании таких известных распределений, как полиномиальное, и распределений, связанных с порядковыми статистиками, но и в моделях пространственных электрических сетей. В завершение этой самой крупной главы мы обсуждаем некоторые обобщения введенных понятий, основанные на подходе, развитом в последнее десятилетие. Этот подход использует описание структуры зависимости случайного поля с помощью верхних оценок ковариаций определенных “пробных функций”, берущихся от наборов изучаемых величин.

§ 1. Основные определения и простые примеры

В данном разделе мы вводим основные определения и исследуем простейшие свойства зависимых случайных систем, которые будут использоваться далее. Здесь мы обращаемся к анализу независимых, положительно или отрицательно зависимых систем, мартингалов и демимартингалов. Кроме того, рассмотрен целый ряд несложных примеров.

1°. Ассоциированность. Положительная и отрицательная ассоциированность. Условия зависимости, которые мы будем обсуждать далее, полезно сравнить с классическим понятием независимости действительнозначных случайных величин X и Y, определенных на вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P})$. Такие X и Y называются *независимыми*, если

$$\mathsf{P}(X \in B, Y \in C) = \mathsf{P}(X \in B)\mathsf{P}(Y \in C) \tag{1.1}$$

для любых борелевских множеств $B, C \subset \mathbb{R}$. Напомним, что борелевские множества в $\mathbb{R}$ — это элементы борелевской σ-алгебры $\mathcal{B}(\mathbb{R})$; для топологического (в частности, метрического) пространства S борелевская σ-алгебра $\mathcal{B}(S)$ определяется как наименьшая σ-алгебра, содержащая все открытые множества. Отображение $f : S \to V$, где S и V — топологические пространства, называется *борелевским*, если $f^{-1}(B) := \{x \in S : f(x) \in B\} \in \mathcal{B}(S)$ для каждого $B \in \mathcal{B}(V)$. Стандартное упражнение на использование ступенчатых функций показывает, что (1.1) равносильно выполнению соотношения

$$\mathsf{E}f(X)g(Y) = \mathsf{E}f(X)\mathsf{E}g(Y) \tag{1.2}$$

для любых ограниченных борелевских функций $f, g : \mathbb{R} \to \mathbb{R}$. Как обычно, символ E обозначает математическое ожидание относительно вероятностной меры P. Вместо (1.2) можно написать

$$cov(f(X), g(Y)) = 0,$$

где *ковариация* $cov(W, Z) := \mathsf{E}WZ - \mathsf{E}W\mathsf{E}Z$ для таких действительнозначных случайных величин W и Z, что W, Z и WZ интегрируемы по мере P.

Как известно, если X и Y независимы и интегрируемы, то $cov(X, Y) = 0$. Более того, для любых (возможно, неограниченных) борелевских функций $f, g : \mathbb{R} \to \mathbb{R}$ случайные величины $f(X)$ и $g(Y)$ тоже независимы и $cov(f(X), g(Y)) = 0$, если $\mathsf{E}|f(X)| < \infty$ и $\mathsf{E}|g(Y)| < \infty$. Также легко привести пример *зависимых* (не удовлетворяющих (1.1)) случайных величин X и Y, для которых $cov(X, Y) = 0$. Например, можно взять $Y = X^2$, где $\mathsf{P}(X = -1) = \mathsf{P}(X = 0) = \mathsf{P}(X = 1) = 1/3$.

Во всевозможных приложениях возникает необходимость рассматривать функционал $cov(f(X), g(Y))$ для определенных классов *пробных функций* f и g, предполагая, что его значения принадлежат заданному подмножеству в $\mathbb{R}$, необязательно состоящему из одной точки 0 (например, $[0, +\infty)$). Можно использовать и случайные векторы X, Y со значениями соответственно в $\mathbb{R}^n$ и $\mathbb{R}^m$, а также борелевские функции $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^m \to \mathbb{R}$. Эти естественные идеи можно развить в нескольких направлениях.

После сделанных предварительных замечаний мы введем важные определения, которые помогут построить интересные стохастические модели.

Пусть $\mathcal{M}(n)$ — класс действительнозначных ограниченных покоординатно неубывающих борелевских функций на $\mathbb{R}^n$, где $n \in \mathbb{N}$. Для конечного множества U его мощность обозначим $|U|$, иногда будет использоваться и обозначение $\sharp U$. Рассмотрим семейство $\mathbf{X} = \{X_t,\, t \in T\}$ действительнозначных случайных величин X_t, заданных на вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P})$. Для $I \subset T$ положим $X_I = \{X_t, t \in I\}$.

Приводимые ниже три определения введены в классических работах Харриса, Лемана, Изери, Прошана, Уолкапа, Йоаг-Дева, Ньюмена, Алама, Саксена, Бёртона, Домбровского и Делинга.

Определение 1.1 ([224]). Семейство $\mathbf{X}$ называется *ассоциированным* (сокращенно $\mathbf{A}$), если для каждого конечного множества $I \subset T$ и любых функций $f, g \in \mathcal{M}(|I|)$ выполнено неравенство

$$cov(f(X_I), g(X_I)) \geqslant 0. \tag{1.3}$$

Используя в (1.3) запись $f(X_I)$, мы имеем в виду, что можно брать любой вектор $\overrightarrow{X_I}$ со значениями в $\mathbb{R}^{|I|}$, полученный каким-либо упорядочением набора $\{X_t, t \in I\}$. Если в формулу входят несколько функций $f(X_I), g(X_I), f_1(X_I)$ и т. д., то аргументом каждой из них служит один и тот же вектор $\overrightarrow{X_I}$. Ясно, что в случае конечного T достаточно проверять соотношение (1.3) только для $I = T$ (например, неубывающая функция $f = f(x_1, x_3)$ не убывает и как функция от x_1, x_2, x_3). Удобно также допустить возможность, когда $I = \varnothing$, считая, что в этом случае всегда $f(X_I) := 0$.

Следующее определение вводит более широкий класс случайных величин.

Определение 1.2 ([158, 348]). Семейство $\mathbf{X}$ *слабо ассоциировано*, или *положительно ассоциировано* ($\mathbf{PA}$), если для произвольных конечных непересекающихся множеств $I, J \subset T$ и любых функций $f \in \mathcal{M}(|I|)$, $g \in \mathcal{M}(|J|)$ выполняется неравенство

$$cov(f(X_I), g(X_J)) \geqslant 0. \tag{1.4}$$

С данным определением во многом сходно

Определение 1.3 ([99, 275]). Говорят, что семейство $\mathbf{X}$ *отрицательно ассоциировано* ($\mathbf{NA}$), если для произвольных конечных непересекающихся множеств $I, J \subset T$ и любых функций $f \in \mathcal{M}(|I|)$, $g \in \mathcal{M}(|J|)$ выполнено неравенство

$$cov(f(X_I), g(X_J)) \leqslant 0. \tag{1.5}$$

Отметим, что при $|T| = 1$ в последних двух определениях невозможно разбить T на два непустых подмножества. В соответствии со сделанным выше замечанием в этом случае $\mathbf{X}$ будет $\mathbf{PA}$ и $\mathbf{NA}$.

Все эти определения развивают простую идею положительной (отрицательной) коррелированности. Иногда мы будем называть вероятностную меру на борелевской σ-алгебре $\mathcal{B}(\mathbb{R}^n)$ *ассоциированной* (соответственно *положительно*, *отрицательно ассоциированной*), если она есть распределение случайного вектора, который удовлетворяет соответствующему определению. Иначе говоря, ассоциированность меры Q ([252]) равносильна тому, что

$$\int_{\mathbb{R}^n} f(x)g(x)Q(dx) \geqslant \int_{\mathbb{R}^n} f(x)Q(dx) \int_{\mathbb{R}^n} g(x)Q(dx) \tag{1.6}$$

для всех $f, g \in \mathcal{M}(n)$. В некоторых ситуациях нам будет удобно называть ассоциированными (или положительно, отрицательно ассоциированными) случайные величины X_t, $t \in T$, а не систему $\{X_t,\ t \in T\}$. Также в дальнейшем мы пишем $X \in \mathbf{A}$ (соответственно $X \in \mathbf{PA}$, $X \in \mathbf{NA}$), где $X = (X_{t_1}, \ldots, X_{t_n})$ — случайный вектор, вместо фразы “$X_{t_1}, \ldots, X_{t_n}$ ассоциированы (соответственно положительно, отрицательно ассоциированы)”.

Отметим, что в п. 5° будут изучаться функции и от бесконечного числа случайных величин, а в § 2 — некоторые обобщения ассоциированности на частично упорядоченные пространства, отличные от $\mathbb{R}^n$. Пусть $Law(X)$ обозначает *распределение* случайного элемента или случайной функции X. Далее будет полезно простое

Замечание 1.4. (а) Если $Law\{X_t,\, t \in T\} = Law\{Y_t, t \in T\}$, т.е. конечномерные распределения этих систем случайных величин одинаковы, а система $\{X_t,\, t \in T\}$ ассоциирована (соответственно является **PA**, **NA**), то тем же свойством обладает и $\{Y_t, t \in T\}$.

(б) Если семейство $\{X_t,\, t \in T\} \in \mathbf{A}$ (является **PA** или **NA**) и $L \subset T$, то семейство $\{X_t, t \in L\}$ тоже ассоциировано (соответственно является **PA** или **NA**).

(в) Случайный вектор $X \in \mathbf{A}$ (**PA** или **NA**) тогда и только тогда, когда любая перестановка его координат обладает тем же свойством.

(г) С помощью теоремы о мажорированной сходимости вместо условия ограниченности функций f и g в определениях 1.1—1.3 можно потребовать существования используемых в них ковариаций, т.е. конечности участвующих в них математических ожиданий.

(д) В определениях 1.1—1.3 можно заменить функции f и g на борелевские, ограниченные и покоординатно невозрастающие, поскольку верна формула $cov(-f(X), -g(Y)) = cov(f(X), g(Y))$.

2°. Критерии положительной и отрицательной ассоциированности. Перед обсуждением свойств введенных систем мы получим достаточные условия выполнения определений 1.1—1.3. Они позволят проверять условия ассоциированности (положительной или отрицательной ассоциированности), используя некоторые подклассы функций из $\mathcal{M} = \cup_{n=1}^{\infty}\mathcal{M}(n)$.

Как обычно, для множества $A \subset S$ *индикатор* A есть такая функция $\mathbb{I}_A$, что $\mathbb{I}_A(x) = 1$ при $x \in A$ и $\mathbb{I}_A(x) = 0$ иначе. Иногда мы будем писать $\mathbb{I}\{A\}$. *Бинарной функцией* назовем индикатор измеримого множества. Если $\mathbb{I}_B \in \mathcal{M}(n)$, то говорят, что борелевское множество $B \subset \mathbb{R}^n$ является *возрастающим*.

Теорема 1.5 ([224, 369]). *Случайный вектор $X = (X_1, \ldots, X_n)$ со значениями в $\mathbb{R}^n$ ассоциирован (соответственно* **PA, NA**), *если условие* (1.3) (*соответственно* (1.4), (1.5)) *выполнено для любых f и g, одновременно принадлежащих $\mathcal{M}$ и одному из следующих классов функций:*

(а) *бинарных;*

(б) *непрерывных;*

(в) *имеющих ограниченные частные производные любого порядка.*

Доказательство. Мы рассматриваем только ассоциированность и опускаем совершенно аналогичные рассуждения для **PA** и **NA**. В каждом из случаев (а)—(в) проверим, что $cov(f(X), g(X)) \geqslant 0$ для любых $f, g \in \mathcal{M}(n)$.

(а) Если Y и Z — интегрируемые в квадрате случайные величины, то формула Хёфдинга[1)] ([261]) утверждает, что

$$cov(Y,Z) = \int\limits_{\mathbb{R}^2} (\mathsf{P}(Y \geqslant y,\ Z \geqslant z) - \mathsf{P}(Y \geqslant y)\mathsf{P}(Z \geqslant z))dydz.$$

Поэтому

$$cov(f(X), g(X)) = \int\limits_{\mathbb{R}^2} cov(\mathbb{I}\{f(X) \geqslant y\}, \mathbb{I}\{g(X) \geqslant z\})dydz. \tag{1.7}$$

Выражение под знаком интеграла неотрицательно, так как при любых $y, z \in \mathbb{R}$ функции $\mathbb{I}\{f(x) \geqslant y\}$ и $\mathbb{I}\{g(x) \geqslant z\}$ бинарны и (покоординатно) не убывают по $x \in \mathbb{R}^n$.

(б) В силу (1.7) достаточно установить, что $cov(f_1(X), f_2(X)) \geqslant 0$ при $f_i = \mathbb{I}_{B_i}$, $i = 1, 2$, где B_1, B_2 — произвольные возрастающие борелевские подмножества в $\mathbb{R}^n$. Сначала предположим дополнительно, что B_1 и B_2 замкнуты. Введем для $x \in \mathbb{R}^n$, $i = 1, 2$ и $k \in \mathbb{N}$ функции

$$f_i^k(x) = \begin{cases} 0, & dist(x, B_i) \geqslant k^{-1}, \\ 1 - k \cdot dist(x, B_i), & \text{иначе.} \end{cases}$$

Здесь $dist$ — расстояние, отвечающее норме $|\cdot|$ в $\mathbb{R}^n$, где $|x| = \max_{1 \leqslant r \leqslant n} |x_r|$, т.е. $dist(x, B) := \inf\{|x - y| : y \in B\}$ для $x \in \mathbb{R}^n$ и $B \subset \mathbb{R}^n$. Все функции f_i^k непрерывны. Нетрудно проверить, что из соотношения $x \leqslant y\ (x, y \in \mathbb{R}^n)$ следует неравенство $dist(x, B_i) \geqslant dist(y, B_i)$, $i = 1, 2$; как обычно,

$$x = (x_1, \ldots, x_n) \leqslant y = (y_1, \ldots, y_n) \iff x_i \leqslant y_i,\ i = 1, \ldots, n. \tag{1.8}$$

Следовательно, все функции $f_i^k \in \mathcal{M}(n)$ и $f_i^k(x) \to f_i(x)$ для $x \in \mathbb{R}^n$ и $i = 1, 2$, когда $k \to \infty$. Итак, $cov(f_1(X), f_2(X)) \geqslant 0$ по теореме о мажорированной сходимости.

Пусть Q — это распределение X в $\mathbb{R}^n$, а B_1, B_2 — произвольные возрастающие борелевские множества из $\mathbb{R}^n$, $i = 1, 2$. Ввиду регулярности вероятностной меры (см., напр., [8, § 1]) для любого $\varepsilon > 0$ найдутся такие компакты $C_i \subset B_i$, что $Q(B_i \setminus C_i) < \varepsilon$. Пусть $F_i = C_i + \mathbb{R}^n_+ = \{x + t : x \in C_i, t \in \mathbb{R}^n_+\}$, $i = 1, 2$. Легко видеть, что $C_i \subset F_i \subset B_i$ (рис. 1) и каждое из множеств F_i — возрастающее и замкнутое. Но тогда по уже доказанной части утверждения

$$\begin{aligned} cov(\mathbb{I}_{B_1}(X), \mathbb{I}_{B_2}(X)) &= Q(B_1B_2) - Q(B_1)Q(B_2) = Q(F_1F_2) - Q(F_1)Q(F_2) \\ &+ Q((B_1 \setminus F_1)F_2) + Q(F_1(B_2 \setminus F_2)) + Q((B_1 \setminus F_1)(B_2 \setminus F_2)) - Q(B_1 \setminus F_1)Q(F_2) \\ &- Q(F_1)Q(B_2 \setminus F_2) - Q(B_1 \setminus F_1)Q(B_2 \setminus F_2) \geqslant cov(\mathbb{I}_{F_1}(X), \mathbb{I}_{F_2}(X)) - 3\varepsilon. \end{aligned}$$

В последнем выражении ковариация неотрицательна в силу (а). Поскольку $\varepsilon > 0$ выбиралось произвольно, утверждение доказано.

(в) В силу (б) и теоремы о мажорированной сходимости нам достаточно показать, что для каждой непрерывной функции $f \in \mathcal{M}(n)$ найдется после-

[1)] См. приложение, § 1.

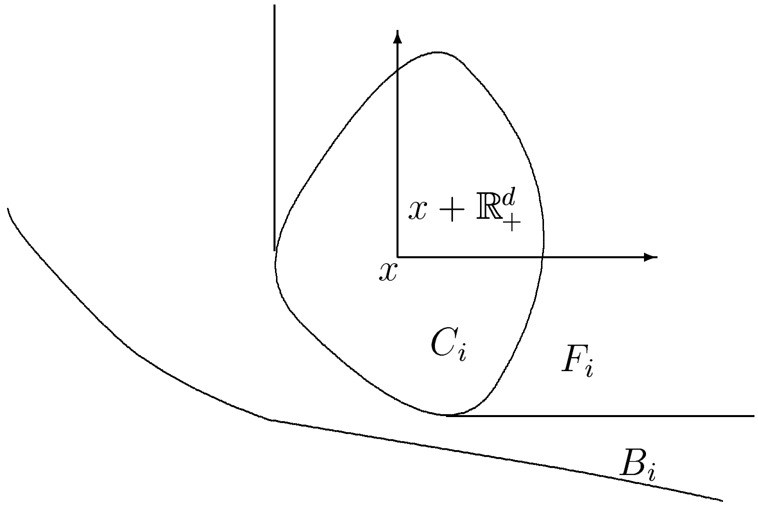

Рис. 1.

довательность ограниченных неубывающих функций f_k $(k \in \mathbb{N})$, имеющих ограниченные производные всех порядков, такая, что $f_k(x) \to f(x)$ для всех $x \in \mathbb{R}^n$ при $k \to \infty$. Если $\chi_k(x) = (k/2\pi)^{n/2} \exp(-\|x\|^2 k/2)$, где $k \in \mathbb{N}$, а $\|\cdot\|$ есть евклидова норма в $\mathbb{R}^n$, то функции

$$f_k(x) := (f * \chi_k)(x) = \int_{\mathbb{R}^n} f(x-y)\chi_k(y)dy, \quad x \in \mathbb{R}^n,$$

удовлетворяют перечисленным требованиям. Поэтому $f_k(X(\omega)) \to f(X(\omega))$ для всех $\omega \in \Omega$, когда $k \to \infty$. $\square$

Замечание 1.6. Если функция $f : \mathbb{R} \to \mathbb{R}$ не убывает, то для каждого $c \in \mathbb{R}$ множество $\{x : f(x) \leqslant c\}$ представляет собой промежуток $(-\infty, a)$ или $(-\infty, a]$, так что f борелевская. Однако при $n > 1$ уже неверно, что каждая покоординатно неубывающая функция на $\mathbb{R}^n$ будет борелевской. Действительно, рассмотрим какое-либо неборелевское подмножество B множества $\{(x,y) \in \mathbb{R}^2 : x = -y\}$. Возьмем функцию $f : \mathbb{R}^2 \to \mathbb{R}$, для которой $f(x,y) = 0$ при $x + y < 0$, $f(x,y) = 1$ при $(x,y) \in B$ и $f(x,y) = 2$ в других случаях. Тогда $\{(x,y) : f(x,y) \leqslant 1\} \notin \mathcal{B}(\mathbb{R}^2)$.

3°. Простейшие примеры. Напомним ключевое

Определение 1.7. Действительнозначные случайные величины X_t, $t \in T$, называются (*взаимно*) *независимыми*, если для любого конечного $I \subset T$ и произвольных $B_t \in \mathcal{B}(\mathbb{R})$, $t \in I$, справедливо соотношение

$$\mathsf{P}\Big(\bigcap_{t \in I} \{X_t \in B_t\}\Big) = \prod_{t \in I} \mathsf{P}(X_t \in B_t).$$

Последнее свойство равносильно независимости σ-алгебр $\sigma\{X_t\}$, $t \in T$, порожденных рассматриваемыми случайными величинами.

Пусть дано *измеримое пространство* $(S, \mathcal{B})$, т.е. множество S, снабженное σ-алгеброй $\mathcal{B}$ его подмножеств. *Случайный элемент* Z на вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P})$ со значениями в S — это отображение $Z : \Omega \to S$, которое $\mathcal{F}|\mathcal{B}$-*измеримо*, т.е. $Z^{-1}(B) \in \mathcal{F}$ для каждого $B \in \mathcal{B}$. В этом случае пишут $Z \in \mathcal{F}|\mathcal{B}$.

Напомним, что если $h : S \to \mathbb{R}$ — некоторая функция и $h \in \mathcal{B}|\mathcal{B}(\mathbb{R})$, то

$$\mathsf{E}h(Z) = \int\limits_{\Omega} h(Z)dP = \int\limits_{S} h(z)P_Z(dz), \tag{1.9}$$

где P_Z — распределение Z в $(S, \mathcal{B})$, т.е. $P_Z(B) := \mathsf{P}(Z^{-1}(B))$, $B \in \mathcal{B}$. Более точно, оба интеграла в (1.9) (по Ω и по S) существуют одновременно и равны.

В следующей теореме собраны элементарные, но важные примеры возникновения (положительной или отрицательной) ассоциированности.

Теорема 1.8 ([224]). *Справедливы следующие утверждения.* (а) *Семейство из одной случайной величины ассоциировано*[1]*.*

(б) *Семейство, образованное объединением независимых друг от друга совокупностей ассоциированных* (*соответственно* **PA**, **NA**) *случайных величин, ассоциировано* (*соответственно* **PA**, **NA**).

(в) *Набор независимых случайных величин ассоциирован и отрицательно ассоциирован, т.е. имеет и свойство* **A**, *и свойство* **NA**.

(г) *Неубывающие функции* (*многих переменных*), *взятые от конечных наборов ассоциированных случайных величин, ассоциированы. Неубывающие функции* (*многих переменных*), *взятые от непересекающихся конечных наборов случайных величин, обладающих свойством* **PA** (**NA**), *сами обладают свойством* **PA** (**NA**).

(д) *Если* $(X_k)_{k\in\mathbb{N}}$ — *последовательность n-мерных ассоциированных* (*соответственно* **PA**, **NA**) *случайных векторов, которая сходится по распределению к случайному вектору* X, *то* X *ассоциирован* (*соответственно* **PA**, **NA**).

Доказательство. (а) Пусть X — произвольная случайная величина, а Y — ее независимая копия. Тогда для неубывающих ограниченных функций $f, g : \mathbb{R} \to \mathbb{R}$ имеем

$$cov(f(X), g(X)) = \mathsf{E}f(X)g(X) - \mathsf{E}f(X)g(Y) = \frac{1}{2}\mathsf{E}(f(X) - f(Y))(g(X) - g(Y)) \geqslant 0,$$

так как под интегралом стоит неотрицательное выражение.

(б) Можно сразу считать, что рассматриваются конечные наборы случайных величин. Предположим, что $X^1, \ldots, X^m$ — независимые случайные векторы, каждый из которых ассоциирован (количества компонент у этих векторов могут быть различными). Если $m = 1$, то утверждение верно. Допустим, что оно справедливо для векторов $X^1, \ldots, X^{m-1}$. Пусть r — размерность X^m и Q обозначает распределение X^m в $\mathbb{R}^r$. Введем $X = (X^1, \ldots, X^{m-1})$. Для функций $f, g \in \mathcal{M}$ в силу независимости, теоремы Фубини и (1.9) имеем

[1] Свойство (а) известно как одно из неравенств Чебышёва (см, напр., [86, с. 59–60]).

$$cov(f(X,X^m),g(X,X^m)) = \int\limits_{\mathbb{R}^r} (\mathsf{E}f(X,x)g(X,x) - \mathsf{E}f(X,x)\mathsf{E}g(X,x))Q(dx)$$
$$+ \int\limits_{\mathbb{R}^r} \mathsf{E}f(X,x)\mathsf{E}g(X,x)Q(dx) - \int\limits_{\mathbb{R}^r} \mathsf{E}f(X,x)Q(dx) \int\limits_{\mathbb{R}^r} \mathsf{E}g(X,x)Q(dx).$$

Рассмотрим правую часть полученного равенства. Интегрируемое выражение в первом слагаемом неотрицательно по предположению индукции. Очевидно, функции $\mathsf{E}f(X,x)$ и $\mathsf{E}g(X,x)$, $x \in \mathbb{R}^r$, принадлежат $\mathcal{M}(r)$. Так как $X^m \in \mathbf{A}$, то разность второго и третьего выражений неотрицательна. Доказательства в случаях $\mathbf{PA}$ и $\mathbf{NA}$ аналогичны приведенному.

(в) Следует из (а) и (б).

(г) Достаточно учесть замкнутость класса $\mathcal{M}$ относительно взятия композиции.

(д) Вытекает из теоремы 1.5(б) и определения сходимости по распределению. $\square$

Следствие 1.9. *Если случайные величины $X_1, \ldots, X_n$ ассоциированы* (*например, независимы*), *то ассоциированы следующие случайные системы:* а) *взвешенные частные суммы c_1X_1, $c_1X_1 + c_2X_2, \ldots, c_1X_1 + \ldots + c_nX_n$, где $c_1, c_2, \ldots, c_n \in \mathbb{R}_+$;* б) $\{\max_{j \leqslant i}(X_1 + \ldots + X_j),\ i = 1, \ldots, n\}$ *или взвешенные частные максимумы с неотрицательными весами;* в) *порядковые статистики*[1] $X_{(1)}, \ldots, X_{(n)}$.

Доказательство заключается в применении теоремы 1.8(г). $\square$

Следствие 1.10. *Пусть $\varphi_k : \mathbb{R} \to \mathbb{R}_+$ — ограниченные неубывающие функции, $k = 1, \ldots, n$. Тогда для любых положительно ассоциированных случайных величин $X_1, \ldots, X_n$ справедливо неравенство*

$$\mathsf{E} \prod_{k=1}^n \varphi_k(X_k) \geqslant \prod_{k=1}^n \mathsf{E}\varphi_k(X_k). \tag{1.10}$$

Если же $X_1, \ldots, X_n$ отрицательно ассоциированы, то (1.10) *выполняется с противоположным знаком неравенства.*

Доказательство состоит в применении теоремы 1.8(г) и индукции. $\square$

Следствие 1.11. *Семейство $\mathbf{X} = \{X_t,\ t \in T\}$ состоит из независимых случайных величин тогда и только тогда, когда одновременно $\mathbf{X} \in \mathbf{PA}$ и $\mathbf{X} \in \mathbf{NA}$.*

Доказательство. Пусть $I = \{t_1, \ldots, t_n\} \subset T$ $(n \in \mathbb{N})$. Если компоненты вектора $X = (X_{t_1}, \ldots, X_{t_n})$ независимы, то по теореме 1.8(б) имеем $X \in \mathbf{PA}$

[1] Для каждого $\omega \in \Omega$ упорядочим числа $X_1(\omega), \ldots, X_n(\omega)$ так, чтобы получить значения $X_{(1)}(\omega) \leqslant \ldots \leqslant X_{(n)}(\omega)$.

и $X \in \mathbf{NA}$. Обратное утверждение немедленно вытекает из предыдущего следствия и условия независимости в терминах функций распределения

$$F_{X_{t_1},\dots,X_{t_n}}(x_1,\dots,x_n) := \mathsf{P}(X_{t_1} \leqslant x_1,\dots,X_{t_n} \leqslant x_n) = \mathsf{E}\prod_{k=1}^{n} \mathbb{I}_{(-\infty,x_k]}(X_{t_k}),$$

где $x_1,\dots,x_n \in \mathbb{R}$. А именно (см., напр., [94, гл. II, § 2.5]), независимость $X_{t_1},\dots,X_{t_n}$ равносильна тому, что для любых $x_1,\dots,x_n \in \mathbb{R}$

$$F_{X_{t_1},\dots,X_{t_n}}(x_1,\dots,x_n) = \prod_{k=1}^{n} F_{X_{t_k}}(x_k).$$

Осталось применить следствие 1.10 и замечание 1.4(д). □

Укажем простое приложение теоремы 1.8. Напомним, что для всякой функции $f : [0,1] \to \mathbb{R}$ ее *полином Бернштейна* степени n — это функция

$$(B_n f)(x) = \sum_{k=0}^{n} \binom{n}{k} f\,(k/n)\, x^k (1-x)^{n-k}, \quad x \in [0,1].$$

Здесь, как обычно, $\binom{n}{k} = n!/(k!(n-k)!)$, $0! := 1$ и $0^0 := 1$.

Пример 1.12 ([402]). Для ограниченных неубывающих $f, g : [0,1] \to \mathbb{R}$ всюду на отрезке $[0,1]$ выполнено неравенство

$$B_n(fg) \geqslant (B_n f)(B_n g). \tag{1.11}$$

В самом деле, достаточно взять биномиальную случайную величину $S_{n,x}$ с параметрами $n \in \mathbb{N}$ и $x \in [0,1]$. Тогда

$$\mathsf{P}(S_{n,x} = k) = \binom{n}{k} x^k (1-x)^{n-k}, \quad k = 0,\dots,n.$$

Следовательно, $(B_n f)(x) = \mathsf{E} f(S_{n,x}/n)$ при всех $n \in \mathbb{N}$ и $x \in [0,1]$. Поэтому (1.11) вытекает из теоремы 1.8(а).

Разные условия зависимости, введенные выше, допускают элементарное описание в простейшем случае, когда имеются две случайные величины, принимающие только два различных значения. Без потери общности мы ограничимся случаем *бинарных случайных величин* X и Y, т.е. величин, принимающих значения 0 и 1 (поскольку можно использовать преобразования $aX + b$, $cY + d$, где $a, c > 0$ и $b, d \in \mathbb{R}$).

Теорема 1.13 ([224]). *Ассоциированность* (**PA**, **NA**, *независимость*) *бинарных случайных величин* X *и* Y *равносильна тому, что* $cov(X,Y) \geqslant 0$ (*соответственно* $cov(X,Y) \geqslant 0$, $cov(X,Y) \leqslant 0$, $cov(X,Y) = 0$).

Доказательство. По теореме 1.5(а), достаточно проверить, что выражение $cov(f(X,Y), g(X,Y))$ неотрицательно (соответственно неположительно, равно нулю) для любых бинарных неубывающих f, g. Но все непостоянные бинарные неубывающие функции из $\{0,1\}^2$ в $\mathbb{R}$ — это функции $x, y, xy, x+$ $+y-xy$. Проверка всех возможных пар приводит к утверждению теоремы. □

Теорема 1.13 не может быть обобщена на случайные векторы более высокой размерности, даже с бинарными компонентами, как показывает следующий

Пример 1.14. Пусть $X = (X_1, X_2, X_3)$ — трехмерный случайный вектор, компоненты которого являются бинарными случайными величинами. Возьмем $f(X_1, X_2) = X_1^+ X_2^+$ и $g(X_3) = X_3$, где $x^+ := \max\{x, 0\}$, $x \in \mathbb{R}$. Пусть

$$p_{ijk} := \mathsf{P}(X_1 = i, X_2 = j, X_3 = k).$$

Тогда при $p_{100} = p_{010} = p_{001} = 1/12$ и $p_{110} = p_{101} = p_{011} = 1/4$ компоненты X положительно коррелированы, но $cov(f(X_1, X_2), g(X_3)) < 0$. Следовательно, $X \notin \mathbf{PA}$. С другой стороны, при $p_{100} = p_{010} = p_{001} = 3/10$, $p_{111} = 1/10$ получаем вектор с отрицательно коррелированными компонентами, для которого $cov(f(X_1, X_2), g(X_3)) > 0$, а значит, $X \notin \mathbf{NA}$.

Итак, для семейства случайных величин $\{X_t,\, t \in T\}$ положительная или отрицательная ассоциированность — вообще говоря, более сильные условия, чем соответственно требования $cov(X_s, X_t) \geqslant 0$ или $cov(X_s, X_t) \leqslant 0$ при всех $s, t \in T$, $s \neq t$. Для еще одной иллюстрации этого обстоятельства рассмотрим

Пример 1.15. Пусть X, Y — такие случайные величины, что

$$\mathsf{P}(X = -1, Y = 1) = \mathsf{P}(X = -1, Y = -1) = 1/4, \quad \mathsf{P}(X = 1, Y = 0) = 1/2.$$

Тогда X и Y зависимы, но $cov(X, Y) = 0$. Для любых $f, g : \mathbb{R} \to \mathbb{R}$ имеем

$$cov(f(X), g(Y)) = f(-1)g(-1)/4 + f(1)g(0)/2 + f(-1)g(1)/4 \\ - (f(1) + f(-1))(g(-1) + 2g(0) + g(1))/8.$$

Пусть $f(-1) = 0$, $f(1) = 1$, $g(-1) = g(0) = 0$, $g(1) = 1$. Ясно, что такие f и g можно выбрать в $\mathcal{M}(1)$. Тогда $cov(f(X), g(Y)) = -1/8$. Если теперь $f(-1) = 0$, $f(1) = 1$, $g(-1) = 0$, $g(0) = 1$, $g(1) = 1$, то $cov(f(X), g(Y)) = 1/8$. Итак, вектор (X, Y) не является ни $\mathbf{PA}$, ни $\mathbf{NA}$.

Положительная ассоциированность очень близка ассоциированности, но они не эквивалентны, как видно из следующего примера Изери, Прошана и Уолкапа.

Пример 1.16 ([224]). Пусть (X, Y) — такой случайный вектор, что для $i, j = 0, 1, 2$ имеем $\mathsf{P}(X = i, Y = j) = p_{ij}$, где p_{ij} заданы матрицей

$$\begin{pmatrix} p_{00} & p_{01} & p_{02} \\ p_{10} & p_{11} & p_{12} \\ p_{20} & p_{21} & p_{22} \end{pmatrix} = \begin{pmatrix} 15/64 & 0 & 1/8 \\ 0 & 9/32 & 0 \\ 1/8 & 0 & 15/64 \end{pmatrix}.$$

Тогда $(X, Y) \in \mathbf{PA}$. Действительно, по теореме 1.5(а), для этого достаточно, чтобы $cov(\mathbb{I}\{X \geqslant x\}, \mathbb{I}\{Y \geqslant y\}) \geqslant 0$ при всех $x, y \in \mathbb{R}$. Последнее неравенство можно проверять лишь для $x, y \in \{1, 2\}$, что делается элементарно. Теперь возьмем $f(x, y) = \mathbb{I}\{x \vee y > 1\}$ и $g(x, y) = \mathbb{I}\{x \wedge y > 0\}$, где $(x, y) \in \mathbb{R}^2$. Здесь и далее $\vee$, $\wedge$ — соответственно знаки максимума и минимума. Очевидно, $f, g \in \mathcal{M}(2)$. Легко видеть, что $\mathsf{E}f(X, Y) = 31/64$, $\mathsf{E}g(X, Y) = 33/64$ и $\mathsf{E}f(X, Y)g(X, Y) = 15/64$, а потому $cov(f(X, Y), g(X, Y)) < 0$. Итак, X и Y не являются ассоциированными величинами.

4°. Демимартингалы. Определения 1.1—1.3 можно видоизменять, налагая дополнительные ограничения на множества I и J или на функции f и g.

Определение 1.17 ([350]). Интегрируемые случайные величины $(S_n)_{n\in T}$, где либо $T=\mathbb{N}$, либо $T=\{1,\dots,N\}\subset\mathbb{N}$, образуют *демимартингал*, если для любого $n\in T$, для которого $n+1\in T$, и всех $f\in\mathcal{M}(n)$ справедливо неравенство

$$\mathsf{E}\left((S_{n+1}-S_n)f(S_1,\dots,S_n)\right)\geqslant 0. \tag{1.12}$$

Они образуют *демисубмартингал*, если то же верно при дополнительном условии, что f неотрицательна.

Например, если случайные величины $X_1,\dots,X_n,\dots$ обладают свойством **PA** и центрированы, то частные суммы $S_n=\sum_{i=1}^n X_i, n\in\mathbb{N}$, дают демимартингал. Напомним, что последовательность $(S_n)_{n\in\mathbb{N}}$ является *мартингалом* (относительно естественной фильтрации $\mathcal{F}_n=\sigma\{S_1,\dots,S_n\}$; здесь, как всегда, $\sigma\{Z_t,t\in T\}$ — наименьшая σ-алгебра, относительно которой измеримы все величины $Z_t,t\in T$), если $\mathsf{E}|S_n|<\infty$ и $\mathsf{E}(S_{n+1}|\mathcal{F}_n)=S_n$ P-п.н. для любого n, "п.н." означает "почти наверное". Несложное упражнение показывает, что для интегрируемых S_n последнее равенство равносильно тому, что для каждой ограниченной борелевской функции f математическое ожидание в (1.12) равно нулю. Итак, мартингал есть демимартингал. Обратное неверно:

Пример 1.18. Пусть случайные величины X и Y те же, что в примере 1.15. Тогда случайные величины (S_1,S_2), где $S_1=Y$ и $S_2=X$, образуют демимартингал, так как для каждой $f\in\mathcal{M}(1)$ имеем

$$\mathsf{E}(X-Y)f(Y)=(f(0)-f(-1))/2\geqslant 0.$$

В то же время (S_1,S_2) — не мартингал, поскольку

$$\mathsf{E}(X|Y=1)=\frac{\mathsf{E}X\mathbb{I}\{Y=1\}}{\mathsf{P}(Y=1)}=-1\neq 1.$$

Заметим, что такой пример нельзя построить, располагая только бинарными случайными величинами X и Y. В самом деле, если $p_{ij}=\mathsf{P}(X=i,X=j)$, $i,j\in\{0,1\}$, то (X,Y) — мартингал тогда и только тогда, когда $p_{01}=p_{10}=0$. Однако последнее условие равносильно и демимартингальности (X,Y).

Покажем, что ассоциированность и демимартингальность — разные (хотя и сходные) свойства.

Пример 1.19. По теореме 1.13 для бинарных векторов из того, что $p_{01}=p_{10}=0$, следует, что $(X,Y)\in\mathbf{A}$, так как в этом случае $cov(X,Y)=$ $=p_{11}(1-p_{11})\geqslant 0$. Пусть $p_{11}=p_{00}=p\in(1/4,1/2)$ и $p_{01}=p_{10}=1/2-p$, тогда $cov(X,Y)=p-1/4>0$. Следовательно, X и Y ассоциированы, но не образуют демимартингала, поскольку $p_{01}=p_{10}>0$. Легко построить и неассоциированный демимартингал. Пусть значения случайных величин X и Y принадлежат соответственно множествам $\{-1,1\}$ и $\{-a,0,a\}$ для некоторого $a>0$. Предположим также, что

$$\mathsf{P}(X=1,Y=-a)=\mathsf{P}(X=-1,Y=a)=p,$$
$$\mathsf{P}(X=1,Y=0)=\mathsf{P}(X=-1,Y=0)=1/2-p,\ \ 0<p<1/2.$$

Возьмем $S_1 = X$, $S_2 = X + Y$, тогда при $f \in \mathcal{M}(1)$ имеем

$$\mathsf{E}Yf(X) = ap(f(1) - f(-1)) \geqslant 0.$$

Таким образом, (S_1, S_2) есть демимартингал. Однако если $a > 1/(2p)$, то[1] $cov(S_1, S_2) = \mathsf{D}X + \mathsf{E}XY = 1 - 2ap < 0$, поэтому $(S_1, S_2) \notin \mathbf{A}$.

В следующей главе будут приведены некоторые примеры того, как элементы мартингальной техники можно применять к демимартингалам. Там мы получим аналог классического неравенства Дуба.

5°. Функции бесконечного числа случайных величин. Перейдем к ситуации, когда изучается ассоциированность функций от бесконечного числа независимых случайных величин. Это потребуется, когда среди примеров ассоциированных систем будет обсуждаться модель перколяции. Для $x = (x_k)_{k\in\mathbb{N}} \in \mathbb{R}^\infty$ и $y = (y_k)_{k\in\mathbb{N}} \in \mathbb{R}^\infty$ полагаем $x \leqslant y$ тогда и только тогда, когда $x_k \leqslant y_k$ при всех $k \in \mathbb{N}$. Рассмотрим в $\mathbb{R}^\infty$ цилиндрическую σ-алгебру, т.е. наименьшую σ-алгебру подмножеств $\mathbb{R}^\infty$, которая содержит все множества вида $\{x_{i_1} \leqslant a_1, \ldots, x_{i_k} \leqslant a_k\}$, где $(a_1, \ldots, a_k) \in \mathbb{R}^k, \{i_1, \ldots, i_k\} \subset \mathbb{N}$, $k \in \mathbb{N}$. Нетрудно показать, что в $\mathbb{R}^\infty$ цилиндрическая σ-алгебра совпадает с борелевской $\mathcal{B}(\mathbb{R}^\infty)$, если наделить $\mathbb{R}^\infty$ топологией покоординатной сходимости.

Определение 1.20. Функция $f : \mathbb{R}^\infty \to \mathbb{R}$ называется *возрастающей*[2], если из соотношения $x \leqslant y$ $(x, y \in \mathbb{R}^\infty)$ следует неравенство $f(x) \leqslant f(y)$. Множество $B \in \mathcal{B}(\mathbb{R}^\infty)$ называется *возрастающим*, если функция $\mathbb{I}_B$ — возрастающая.

Теорема 1.21. *Пусть $X = \{X_t,\, t \in T\}$ — семейство независимых случайных величин, а множество T счетно. Предположим, что функции $f_1, \ldots, f_m : \mathbb{R}^\infty \to \mathbb{R}$ возрастают и $\mathcal{B}(\mathbb{R}^\infty)$-измеримы. Тогда*[3] *случайный вектор $(f_1(X), \ldots, f_m(X)) \in \mathbf{A}$.*

Доказательство. Пусть $T = \{t_1, t_2, \ldots\}$ — некоторая нумерация элементов T. Для произвольных функций $F, G \in \mathcal{M}(m)$ положим

$$Y = F(f_1(X), \ldots, f_m(X)), \quad Z = G(f_1(X), \ldots, f_m(X)).$$

Докажем, что $cov(Y, Z) \geqslant 0$. Обозначим $Y_n = \mathsf{E}(Y|X_{t_1}, \ldots, X_{t_n})$, $n \in \mathbb{N}$. Из теоремы Леви о сходимости мартингалов (см., напр., [94, гл. VII, § 4.3]) следует, что $Y_n \to \mathsf{E}(Y|X_{t_1}, X_{t_2}, \ldots)$ п.н. и в $L^1(\Omega, \mathcal{F}, \mathsf{P})$, когда $n \to \infty$. Но ясно, что $\mathsf{E}(Y|X_{t_1}, X_{t_2}, \ldots) = Y$ п.н. Такое же рассуждение верно для $Z_n = \mathsf{E}(Z|X_{t_1}, \ldots, X_{t_n})$. Итак, $\mathsf{E}Y_n \to \mathsf{E}Y$, $\mathsf{E}Z_n \to \mathsf{E}Z$, $\mathsf{E}Y_nZ_n \to \mathsf{E}YZ$ (последнее соотношение справедливо из-за ограниченности случайных величин Y, Z, Y_n, Z_n, $n \in \mathbb{N}$). Следовательно,

$$cov(Y_n, Z_n) \to cov(Y, Z) \text{ при } n \to \infty.$$

Итак, нам достаточно проверить, что $cov(Y_n, Z_n) \geqslant 0$ при любом $n \in \mathbb{N}$. Воспользуемся простым утверждением (см., напр., [26, гл. 1]).

---

[1] $\mathsf{D}X$ обозначает дисперсию величины X (используется также запись $\mathrm{Var}X$).

[2] Для $f \in \mathcal{M}(n)$ обычно говорят о покоординатном неубывании, а не о возрастании.

[3] Используются соглашения, данные после определений 1.1 и 1.3.

Лемма 1.22. *Пусть $(V, \mathcal{A})$ и $(S, \mathcal{B})$ — некоторые измеримые пространства, а независимые случайные элементы ξ и η, заданные на вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P})$, принимают значения соответственно в V и S. Тогда если функция $H : V \times S \to \mathbb{R}$ ограничена и $(\mathcal{A} \otimes \mathcal{B})|\mathcal{B}(\mathbb{R})$-измерима, то для любого $v \in V$ справедливо равенство*

$$\mathsf{E}(H(\xi, \eta)|\xi = v) = \mathsf{E}H(v, \eta) \quad \text{п.н.} \tag{1.13}$$

Вернемся к доказательству теоремы. Для сокращения записи положим $v_n = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $h_n(v_n) = \mathsf{E}(Y|X_{t_1} = x_1, \ldots, X_{t_n} = x_n)$ и $W_n = (X_{t_{n+1}}, X_{t_{n+2}}, \ldots)$, где $n \in \mathbb{N}$. Тогда из независимости случайных величин X_{t_k} $(k \in \mathbb{N})$ и равенства (1.13) следует, что

$$h_n(v_n) = \mathsf{E}F(f_1(v_n, W_n), \ldots, f_m(v_n, W_n)). \tag{1.14}$$

Функция h_n, возникшая в (1.14), ограничена и покоординатно не убывает. Итак, $Y_n = h_n(X_{t_1}, \ldots, X_{t_n})$. Аналогично $Z_n = g_n(X_{t_1}, \ldots, X_{t_n})$, где $g_n \in \mathcal{M}(n)$. По теореме 1.8(г) величины Y_n и Z_n ассоциированы. Поскольку их ковариация определена, то $cov(Y_n, Z_n) \geqslant 0$, $n \in \mathbb{N}$. $\square$

6°. Условные распределения и перестановки. Много примеров, в которых возникает отрицательная ассоциированность, содержится в следующих двух теоремах. Они основаны на том наблюдении, что отрицательная ассоциированность появляется в задачах типа бесповторной выборки. Иначе говоря, отрицательная зависимость означает, что случайные величины "мешают" друг другу расти (увеличение одних приводит к убыванию других).

Теорема 1.23 ([275]). *Пусть $X_1, \ldots, X_n$ — независимые случайные величины, причем для произвольного $I \subset \{1, \ldots, n\}$ и любой $f \in \mathcal{M}(|I|)$ функция*

$$\mathsf{E}\left(f(X_i, i \in I) \Big| \sum_{i \in I} X_i = t\right) \tag{1.15}$$

не убывает по t. Тогда условное распределение $(X_1, \ldots, X_n)$ при фиксированном значении $\sum_{i=1}^n X_i$ с вероятностью единица обладает свойством **NA**. *Более точно, существует версия регулярного*[1] *условного распределения $Q(\cdot, \omega)$ вектора $(X_1, \ldots, X_n)$ при фиксированном значении $\sum_{i=1}^n X_i$, которая является* **NA** *для почти всех ω.*

Доказательство. Возьмем непустые непересекающиеся $I, J \subset \{1, \ldots, n\}$ так, чтобы $I \cup J = \{1, \ldots, n\}$, и функции $f \in \mathcal{M}(|I|), g \in \mathcal{M}(|J|)$. Определим $S_1 = \sum_{i \in I} X_i$, $S_2 = \sum_{j \in J} X_j$, $S = S_1 + S_2$. По телескопическому свойству условного математического ожидания[2],

$$\begin{aligned} cov(f(X_I), g(X_J)|S) = \mathsf{E}(cov(f(X_I), g(X_J)|S_1, S_2)|S) \\ + cov(\mathsf{E}(f(X_I)|S_1, S_2), \mathsf{E}(g(X_J)|S_1, S_2)|S), \end{aligned} \tag{1.16}$$

[1] См., напр., [94, гл. II, §7.7].

[2] Если $\mathsf{E}|\xi| < \infty$ и σ-алгебры $\mathcal{A} \subset \mathcal{B} \subset \mathcal{F}$, то $\mathsf{E}(\mathsf{E}(\xi|\mathcal{B})|\mathcal{A}) = \mathsf{E}(\xi|\mathcal{A})$; здесь $\mathsf{E}(\xi|\eta_1, \ldots, \eta_n)$ — то же, что $\mathsf{E}(\xi|\sigma\{\eta_1, \ldots, \eta_n\})$, где $\xi, \eta_1, \ldots, \eta_n$ — случайные величины на $(\Omega, \mathcal{F}, \mathsf{P})$.

где использована стандартная запись *условной ковариации*

$$cov(W,Y|Z) := \mathsf{E}(WY|Z) - \mathsf{E}(W|Z)\mathsf{E}(Y|Z) \tag{1.17}$$

для интегрируемых WY, W и Y. Нам потребуется простая

Лемма 1.24. *Пусть ξ_1 и ξ_2 — независимые случайные векторы, принимающие значения в пространствах $\mathbb{R}^{k_1}$ и $\mathbb{R}^{k_2}$ соответственно. Тогда для произвольных ограниченных борелевских функций $F_i:\mathbb{R}^{k_i}\to\mathbb{R}$ и любых борелевских функций $h_i:\mathbb{R}^{k_i}\to\mathbb{R}^{m_i}$ $(m_i,k_i\in\mathbb{N},\ i=1,2)$ верно равенство*

$$\mathsf{E}(F_1(\xi_1)F_2(\xi_2)|h_1(\xi_1),h_2(\xi_2)) = \mathsf{E}(F_1(\xi_1)|h_1(\xi_1))\mathsf{E}(F_2(\xi_2)|h_2(\xi_2)). \tag{1.18}$$

В частности,

$$\mathsf{E}(F_i(\xi_i)|h_1(\xi_1),h_2(\xi_2)) = \mathsf{E}(F_i(\xi_i)|h_i(\xi_i)),\quad i=1,2. \tag{1.19}$$

Доказательство. Рассмотрим σ-алгебру $\mathcal{A}=\sigma\{h_1(\xi_1),h_2(\xi_2)\}$. Случайная величина в правой части (1.18) $\mathcal{A}$-измерима. Пусть событие $A\in\mathcal{A}$ имеет вид $A=A_1A_2$, где $A_i=\{h_i(\xi_i)\in B_i\}$, $B_i\in\mathcal{B}(\mathbb{R}^{m_i})$, $i=1,2$. Тогда легко проверить, что

$$\mathsf{E}F_1(\xi_1)F_2(\xi_2)\mathbb{I}_A = \mathsf{E}\left(\mathsf{E}(F_1(\xi_1)|h_1(\xi_1))\mathsf{E}(F_2(\xi_2)|h_2(\xi_2))\mathbb{I}_A\right), \tag{1.20}$$

поскольку

$$\begin{aligned}\mathsf{E}\mathbb{I}_A\mathsf{E}(F_1(\xi_1)|h_1(\xi_1))\mathsf{E}(F_2(\xi_2)|h_2(\xi_2)) &= \mathsf{E}\mathsf{E}\left(F_1(\xi_1)\mathbb{I}_{A_1}|h_1(\xi_1)\right)\mathsf{E}\left(F_2(\xi_2)\mathbb{I}_{A_2}|h_2(\xi_2)\right)\\ &= \mathsf{E}F_1(\xi_1)\mathbb{I}_{A_1}\mathsf{E}F_2(\xi_2)\mathbb{I}_{A_2} = \mathsf{E}F_1(\xi_1)\mathbb{I}_{A_1}F_2(\xi_2)\mathbb{I}_{A_2} = \mathsf{E}F_1(\xi_1)F_2(\xi_2)\mathbb{I}_A.\end{aligned}$$

Здесь мы воспользовались следующими известными свойствами условных математических ожиданий. Пусть случайная величина ζ измерима относительно σ-алгебры $\mathcal{G}\subset\mathcal{F}$, тогда

$$\mathsf{E}(\zeta\nu|\mathcal{G}) = \zeta\mathsf{E}(\nu|\mathcal{G}),\quad \text{если}\ \ \mathsf{E}|\zeta\nu|<\infty\ \text{и}\ \mathsf{E}|\nu|<\infty.$$

Для интегрируемой случайной величины ξ и случайного вектора η со значениями в $\mathbb{R}^m$ справедливо представление $\mathsf{E}(\xi|\eta)=\Phi(\eta)$, где $\Phi:\mathbb{R}^m\to\mathbb{R}$ — некоторая борелевская функция, вид которой зависит от ξ (см., напр., [94, гл. II, § 7.7]). Из данных свойств следует, что $\mathsf{E}(F_i(\xi_i)\mathbb{I}_{A_i}|h_i(\xi_i))=\Phi_i(h_i(\xi_i))$ для некоторых борелевских функций Φ_i $(i=1,2)$. Значит, случайные величины $\Phi_1(h_1(\xi_1))$ и $\Phi_2(h_2(\xi_2))$ независимы в силу независимости ξ_1 и ξ_2.

Очевидно, соотношение (1.20) выполняется на алгебре $\mathcal{H}$ конечных объединений событий вида $A=A_1A_2$. Алгебра $\mathcal{H}$ порождает $\mathcal{A}$, что позволяет установить (1.20) для всех $A\in\mathcal{A}$, поскольку для любых $\varepsilon>0$ и $A\in\mathcal{A}$ найдется такое $A_\varepsilon\in\mathcal{H}$, что $\mathsf{P}(A\triangle A_\varepsilon)<\varepsilon$ (здесь $A\triangle A_\varepsilon := (A\setminus A_\varepsilon)\cup(A_\varepsilon\setminus A)$). $\square$

Продолжим доказательство теоремы. Первое слагаемое в правой части (1.16) равно нулю, так как по лемме 1.24

$$\mathsf{E}(f(X_I)g(X_J)|S_1,S_2) = \mathsf{E}(f(X_I)|S_1,S_2)\mathsf{E}(g(X_J)|S_1,S_2).$$

Из (1.19) видно, что случайные величины $\mathsf{E}(f(X_I)|S_1,S_2)$ и $\mathsf{E}(g(X_J)|S_1,S_2)$ являются борелевскими функциями соответственно от S_1 и S_2. Эти функции ограничены, обозначим их φ_1 и φ_2. Согласно (1.15)

φ_1 и φ_2 не убывают. Второе слагаемое в (1.16) можно записать как $cov(\varphi_1(S_1), \varphi_2(S_2)|S)$.

Пусть $\mu_z(\cdot)$ — регулярное условное распределение случайного вектора (S_1, S_2) при $S = z$ ($z \in \mathbb{R}$). Так как $S_1 + S_2 = S$, то эта мера сосредоточена на прямой $x + y = z$. Имеем

$$\mathsf{E}(\varphi_1(S_1)\varphi_2(S_2)|S = z) = \int\limits_{\mathbb{R}^2} \varphi_1(x)\varphi_2(y)d\mu_z(x, y)$$
$$= \int\limits_{\{x+y=z\}} \varphi_1(x)\varphi_2(z-x)d\mu_z(x, y) = \int\limits_{\mathbb{R}} \varphi_1(x)\varphi_2(z-x)d\nu_z(x),$$

где мера ν_z — проекция μ_z на ось, отвечающую переменной x. Выражение $\varphi_2(z-x)$ не возрастает по x, а следовательно, по теореме 1.8(а) последний интеграл не превосходит числа

$$\int\limits_{\mathbb{R}} \varphi_1(x)d\nu_z(x) \int\limits_{\mathbb{R}} \varphi_2(z-x)d\nu_z(x) = \mathsf{E}(f(S_1)|S = z)\mathsf{E}(g(S_2)|S = z).$$

Значит, условная ковариация в правой части (1.16) неположительна. □

В качестве иллюстрации выполнения условий установленной теоремы приведем один результат без доказательства. Здесь и далее $\log x$ обозначает натуральный логарифм числа $x > 0$.

Теорема 1.25 ([221]). *Пусть $X_1, \ldots, X_n$ — независимые случайные величины, причем X_i обладает положительной плотностью $p_i(x)$ и функция $\log p_i(x)$ вогнута, $i = 1, \ldots, n$. Тогда функция* (1.15) *не убывает по t.*

Следуя Йоаг-Деву и Прошану [275], будем говорить, что случайный вектор $X = (X_1, \ldots, X_n)$ имеет *распределение перестановки*, если его распределение есть равномерная мера на множестве перестановок действительных чисел $a_1, \ldots, a_n$ ($a_1 \leqslant \ldots \leqslant a_n$; возможно, что некоторые из них совпадают), т.е. каждая перестановка реализуется с вероятностью $1/n!$. Это частный случай ситуации, когда случайный вектор $Y = (Y_1, \ldots, Y_n)$ имеет *перестановочное распределение*. Последнее означает, что для любой перестановки $(i_1, \ldots, i_n)$ чисел $1, \ldots, n$ случайный вектор $(Y_{i_1}, \ldots, Y_{i_n})$ распределен так же, как Y.

Теорема 1.26 ([275]). *Распределение перестановки является* **NA**.

Доказательство. Воспользуемся индукцией по n. Случай $n = 1$ тривиален. Пусть утверждение справедливо для $(n-1)$-мерных случайных векторов. Разобьем $\{1, \ldots, n\}$ на два непустых непересекающихся множества I, J так, чтобы $1 \in I$, и возьмем произвольные функции $f \in \mathcal{M}(|I|), g \in \mathcal{M}(|J|)$. Заметим, что можно вместо f и g подобрать такие симметрические неубывающие функции f_1 и g_1, что

$$\mathsf{E}f(X_I) = \mathsf{E}f_1(X_I), \;\; \mathsf{E}g(X_J) = \mathsf{E}g_1(X_J), \;\; \mathsf{E}f(X_I)g(X_J) = \mathsf{E}f_1(X_I)g_1(X_J).$$

Это легко сделать: пусть $I = \{i_1, \ldots, i_k\}$ и $\mathbf{a}_I = (a_{i_1}, \ldots, a_{i_k})$, тогда положим

$$f_1(\mathbf{a}_I) := \frac{1}{k!} \sum_{\tau \in \mathcal{P}_k} f(\tau(\mathbf{a}_I)),$$

где $\mathcal{P}_k$ есть множество всех перестановок (обозначаемых τ) координат вектора $\mathbf{a}_I$. Вводя аналогично g_1, видим, что f_1 и g_1 обладают нужными свойствами. Например, равенство $\mathsf{E}f(X_I) = \mathsf{E}f_1(X_I)$ следует из того, что $\mathsf{E}f(\tau(\mathbf{X}_I)) = \mathsf{E}f(\mathbf{X}_I)$ для любого подмножества $I \subset \{1, \ldots, n\}$ и каждой перестановки $\tau \in \mathcal{P}_k$. Точнее, если $\tau(X_I) = (X_{j_1}, \ldots, X_{j_k})$, то

$$\mathsf{E}f(\tau(\mathbf{X}_I)) = \sum f(a_{r_1}, \ldots, a_{r_k}) \mathsf{P}(X_{j_1} = a_{r_1}, \ldots, X_{j_k} = a_{r_k}) \\ = \frac{(n-k)!}{n!} \sum f(a_{r_1}, \ldots, a_{r_k}) = \mathsf{E}f(\mathbf{X}_I),$$

где суммы берутся по всем $a_{r_1}, \ldots, a_{r_k} \in \{a_1, \ldots, a_n\}$ таким, что $r_p \neq r_q$ для любой пары различных $p, q \in \{1, \ldots, k\}$.

Из (1.17) легко вывести известную формулу с условной ковариацией

$$cov(f_1(X_I), g_1(X_J)) \\ = \mathsf{E}cov(f_1(X_I), g_1(X_J)|X_1) + cov(\mathsf{E}(f_1(X_I)|X_1), \mathsf{E}(g_1(X_J)|X_1)). \quad (1.21)$$

Заметим, что $\mathsf{P}(X_1 = t) = 1/n$ для любого $t \in \{a_1, \ldots, a_n\}$. Если $|I| = 1$, то $\mathsf{E}(f(X_I)|X_1 = t) = f(t)$. Если же $I = \{1, i_2, \ldots, i_k\}, k > 1$, то

$$\mathsf{E}(f_1(X_I)|X_1 = t) = \frac{\mathsf{E}f_1(X_I)\mathbb{I}\{X_1 = t\}}{\mathsf{P}(X_1 = t)} \\ = n \sum{}' f_1(t, a_{i_2}, \ldots, a_{i_k}) \mathsf{P}(X_1 = t, X_{i_2} = a_{i_2}, \ldots, X_{i_k} = a_{i_k}) \\ = n\frac{(n-k)!}{n!} \sum{}' f_1(t, a_{r_2}, \ldots, a_{r_k}),$$

где сумма $\sum{}'$ берется по всем таким наборам $(a_{r_2}, \ldots, a_{r_k})$, для которых $a_{r_p} \in A_t := \{a_1, \ldots, a_n\} \setminus \{t\}$ и $r_p \neq r_q$ для $p, q \in \{2, \ldots, k\}$, $p \neq q$. В силу симметричности f_1 функция $\mathsf{E}(f_1(X_I)|X_1 = t)$ не убывает по $t \in \{a_1, \ldots, a_n\}$.

Аналогично для каждого множества $J = \{j_1, \ldots, j_m\}$, $m \geqslant 1$, верно равенство

$$\mathsf{E}(g_1(X_J)|X_1 = t) = n\frac{(n-m-1)!}{n!} \sum{}'' g_1(a_{s_1}, \ldots, a_{s_m}),$$

где сумма $\sum{}''$ берется по всем таким наборам $(a_{s_1}, \ldots, a_{s_m})$, что $a_{s_u} \in A_t$ и $s_u \neq s_v$ для $u, v \in \{1, \ldots, m\}$, $u \neq v$. Следовательно, $\mathsf{E}(g_1(X_J)|X_1 = t)$ не возрастает по $t \in \{a_1, \ldots, a_n\}$ в силу симметричности g_1. По теореме 1.8(а) второе слагаемое в правой части (1.21) неположительно.

Далее, для каждого $t \in \{a_1, \ldots, a_n\}$ и любой функции $h : \mathbb{R}^n \to \mathbb{R}$ мы по тем же причинам имеем

$$\mathsf{E}(h(X)|X_1 = t) = \frac{1}{(n-1)!} \sum{}''' h(t, a_{j_2}, \ldots, a_{j_n}) = \mathsf{E}h(t, Y^t), \quad (1.22)$$

где сумма $\sum'''$ — по всем перестановкам $(a_{j_2}, \dots, a_{j_n})$ элементов множества A_t, а случайный вектор Y^t имеет $(n-1)$-мерное распределение перестановки на A_t. Итак, по предположению индукции и ввиду (1.22)

$$\begin{aligned}\mathsf{E}(f_1(X_I)g_1(X_J)|X_1 = t) &= \mathsf{E}(f_1(t, Y^t_{I\setminus\{1\}})g_1(Y^t_J)) \\ &\leqslant \mathsf{E}(f_1(t, Y^t_{I\setminus\{1\}})\mathsf{E}g_1(Y^t_J)) = \mathsf{E}(f_1(X_I)|X_1 = t)\mathsf{E}g_1(X_J|X_1 = t).\end{aligned}$$

Таким образом, условная ковариация в первом слагаемом (1.21) с вероятностью единица неположительна. □

Обобщения двух последних теорем рассматривались в [264].

Теорема 1.27 ([275]). *Следующие многомерные распределения обладают свойством* $\mathbf{NA}$: а) *мультиномиальное,* б) *распределение случайной выборки без возвращения,* в) *многомерное гипергеометрическое.*

Доказательство. а) Предположим, что вектор $X = (X_1, \dots, X_n)$ распределен мультиномиально с параметрами $m \in \mathbb{N}$ и $p = (p_1, \dots, p_n) \in \mathbb{R}^n$. Здесь $p_i \geqslant 0$, $i = 1, \dots, n$ и $\sum_{i=1}^n p_i = 1$. Это означает, что X_i равно числу исходов "i", наступивших в m независимых испытаниях, в каждом из которых вероятность i-го исхода равна p_i. Можно считать, что $p_i > 0$ при $i = 1, \dots, n$, так как в противном случае всегда $X_i = 0$.

Итак, $X = \sum_{i=1}^m \xi^{(i)}$, где случайные векторы $\xi, \xi^{(1)}, \dots, \xi^{(m)}$ независимы, одинаково распределены и отображают Ω в $\mathbb{R}^n$. Вектор $\xi = (\xi_1, \dots, \xi_n)$ принимает с вероятностью p_j значение $(\delta_{1j}, \dots, \delta_{nj})$, где δ_{rj} — символ Кронекера.

По теореме 1.8(г) достаточно проверить, что $\xi \in \mathbf{NA}$. Для этого рассмотрим в $\{1, \dots, n\}$ подмножества $I = \{i_1, \dots, i_k\}$ и $J = \{j_1, \dots, j_s\}$, где $I \cap J = \varnothing$, и покажем, что для любых $f \in \mathcal{M}(k)$ и $g \in \mathcal{M}(s)$ выполнено неравенство $cov(f(\xi_I), g(\xi_J)) \leqslant 0$. Положим

$$\begin{aligned} Q_1 &= f(1, 0, \dots, 0)p_{i_1} + \ldots + f(0, \dots, 0, 1)p_{i_k}, \\ Q_2 &= g(1, 0, \dots, 0)p_{j_1} + \ldots + g(0, \dots, 0, 1)p_{j_s}, \\ f(\mathbf{0}) &= f(0, \dots, 0), \quad g(\mathbf{0}) = g(0, \dots, 0), \quad p_I = \sum_{i \in I} p_i, \quad p_J = \sum_{j \in J} p_j. \end{aligned}$$

Тогда $cov(f(\xi_I), g(\xi_J))$ равна

$$\begin{aligned} f(\mathbf{0})Q_2 + g(\mathbf{0})Q_1 + f(\mathbf{0})g(\mathbf{0})(1 - p_I - p_J) - (Q_1 + f(\mathbf{0})(1 - p_I))(Q_2 + g(\mathbf{0})(1 - p_J)) \\ = -Q_1Q_2 + f(\mathbf{0})Q_2p_I + g(\mathbf{0})Q_1p_J - f(\mathbf{0})g(\mathbf{0})p_Ip_J \\ = -(Q_1 - f(\mathbf{0})p_I)(Q_2 - g(\mathbf{0})p_J) \leqslant 0, \end{aligned}$$

так как $Q_1 \geqslant f(\mathbf{0})p_I$ и $Q_2 \geqslant g(\mathbf{0})p_J$ в силу монотонности f и g.

б) Рассмотрим урну, в которой лежат n различных шаров, отмеченных номерами $a_1, \dots, a_n$. Возьмем без возвращения m шаров; пусть $X_1, \dots, X_m$ — их номера. Тогда если $m = n$, то вектор $(X_1, \dots, X_n)$ имеет распределение перестановки и по теореме 1.26 обладает свойством $\mathbf{NA}$. Если же $m < n$, то $(X_1, \dots, X_m) \in \mathbf{NA}$ по той же теореме и замечанию 1.4(б).

в) Пусть теперь в урне N шаров k разных цветов. Положим $N_0 = 0$, а N_i — равным числу шаров i-го цвета, $i = 1, \dots, k$, $N_1 + \ldots + N_k = N$. Можно перенумеровать шары так, чтобы те из них, цвет которых i-й, имели номера $\sum_{j=1}^{i-1} N_j + 1, \ldots, \sum_{j=1}^{i} N_j$. Теперь выберем без возвращения m шаров. Пусть случайная величина Y_j — это индикатор того, что j-й шар находится в числе извлеченных, $j = 1, \ldots, N$. Вектор $(Y_1, \ldots, Y_N)$ имеет распределение перестановки и поэтому обладает свойством **NA** по теореме 1.26. Значит, $(Y_1, \ldots, Y_m) \in \mathbf{NA}$ при $m \leqslant N$. Введем случайные величины

$$X_i = \sum_{t: N_1 + \ldots + N_{i-1} < t \leqslant N_1 + \ldots + N_i} Y_t, \quad i = 1, \ldots, k.$$

Тогда X_i есть число извлеченных шаров i-го цвета, т.е. $X = (X_1, \ldots, X_k)$ имеет многомерное гипергеометрическое распределение. Теперь $X \in \mathbf{NA}$ по теореме 1.8(г), так как суммы в определении X берутся по непересекающимся множествам индексов. $\square$

7°. Теорема Кимбелла. Подбор элементарных примеров закончим одним статистическим приложением. Пусть $\{X_{ij}, i = 1, \ldots, r, j = 1, \ldots, s\}$ — независимые наблюдения, причем гауссовские величины $X_{ij} \sim N(m_{ij}, \sigma^2)$ с некоторым $\sigma > 0$. В дисперсионном анализе i и j имеют смысл уровней проявления факторов A и B (описание модели см. в [57, гл. 6]). Введем

$$\overline{X} = \frac{1}{rs} \sum_{i=1}^{r} \sum_{j=1}^{s} X_{ij}, \quad \overline{X_{i\cdot}} = \frac{1}{s} \sum_{j=1}^{s} X_{ij}, \quad \overline{X_{\cdot j}} = \frac{1}{r} \sum_{i=1}^{r} X_{ij},$$
$$m = \frac{1}{rs} \sum_{i=1}^{r} \sum_{j=1}^{s} m_{ij}, \quad m_{i\cdot} = \frac{1}{s} \sum_{j=1}^{s} m_{ij}, \quad m_{\cdot j} = \frac{1}{r} \sum_{i=1}^{r} m_{ij},$$

где $i = 1, \ldots, r$, $j = 1, \ldots, s$. Рассмотрим статистические гипотезы

$$H_A = \{m_{1\cdot} = \ldots = m_{r\cdot}\}, \quad H_B = \{m_{\cdot 1} = \ldots = m_{\cdot s}\}.$$

Для их проверки строятся квадратичные формы

$$q_1 = s \sum_{i=1}^{r} (\overline{X_{i\cdot}} - \overline{X})^2, \quad q_2 = r \sum_{j=1}^{s} (\overline{X_{\cdot j}} - \overline{X})^2, \quad q_3 = \sum_{i=1}^{r} \sum_{j=1}^{s} (X_{ij} - \overline{X_{i\cdot}})(X_{ij} - \overline{X_{\cdot j}}).$$

В предположении истинности гипотез H_A и H_B квадратичные формы q_1, q_2, q_3 независимы и имеют распределения χ^2 с числами степеней свободы соответственно $r-1, s-1, m = (r-1)(s-1)$ ([57]). Статистики критериев проверки — дроби Фишера

$$F_A = \frac{(s-1)q_1}{q_3}, \quad F_B = \frac{(r-1)q_2}{q_3},$$

где $F_A := 0$ и $F_B := 0$ в случае (нулевой вероятности), если $q_3 = 0$. С помощью понятия ассоциированности легко доказать теорему Кимбелла.

Теорема 1.28 ([224, 289]). *Вероятность того, что при одновременной проверке гипотез H_A и H_B не произойдет ошибки первого рода, не меньше произведения аналогичных вероятностей, отвечающих независимым проверкам H_A и H_B.*

Доказательство. Случайные величины q_1, q_2, q_3^{-1} независимы, а величины F_A и F_B — неубывающие функции от них. По теореме 1.8(в) и (г), они ассоциированы, поэтому для каждого положительного a выполняется неравенство $\mathsf{P}(F_A \leqslant a, F_B \leqslant a) \geqslant \mathsf{P}(F_A \leqslant a)\mathsf{P}(F_B \leqslant a)$. $\square$

§ 2. Классы ассоциированных и родственных систем

В этом разделе дается ряд важных примеров, показывающих, как свойство ассоциированности (возможно, **PA** или **NA**) удается установить прямой проверкой соответствующего определения. Требуемые неравенства для ковариаций мы установим с помощью разных методов, определяемых специфическими особенностями изучаемого случайного процесса или поля. В § 4 будет рассмотрен более сложный способ получения достаточных условий ассоциированности (или **NA**), который может применяться без явного вычисления конечномерных распределений случайной функции.

1°. Гауссовские системы. Как хорошо известно, семейство действительнозначных случайных величин $\mathbf{X} = \{X_t,\, t \in T\}$ называется *нормальной* (или *гауссовской*) *системой*, если все ее конечномерные распределения гауссовские (см., напр., [94, гл. II, § 13.6]). Гауссовская система $\mathbf{X}$ состоит из независимых случайных величин тогда и только тогда, когда $cov(X_t, X_s) = 0$ для всех $s, t \in T$, $s \neq t$ (см., напр., [94, гл. II, § 13.4]). Мы приведем красивые обобщения этого утверждения, принадлежащие Питту [369], Йоаг-Деву и Прошану [275]. Они дают простые ответы на вопросы, когда гауссовская система $\mathbf{X}$ принадлежит $\mathbf{A}$ или $\mathbf{NA}$. Доказательства этих результатов существенно сложнее, чем упомянутого свойства независимости.

Теорема 2.1 ([369]). *Гауссовский вектор $X = (X_1, \ldots, X_n) \in \mathbf{A}$ тогда и только тогда, когда его компоненты неотрицательно коррелированы.*

Доказательство. Необходимость. Если $X \in \mathbf{A}$, то $cov(X_i, X_j) \geqslant 0$, $i, j = 1, \ldots, n$, поскольку $f(x) = x$ есть неубывающая функция на $\mathbb{R}$ и рассматриваемые ковариации существуют.

Достаточность. Сначала заметим, что если случайный вектор X ассоциирован и $a \in \mathbb{R}^n$, то вектор $X + a$ ассоциирован (по теореме 1.8(г)). Поэтому можно сразу считать, что $\mathsf{E}X = 0$. Пусть $\Sigma = (\sigma_{ij})$ — ковариационная матрица X. Сначала мы рассмотрим случай, когда $\det \Sigma > 0$. Тогда вектор X имеет плотность

$$\phi(x) = (2\pi)^{-n/2} (\det \Sigma)^{-1/2} e^{-(\Sigma^{-1}x, x)/2},$$

где $(\cdot,\cdot)$ есть скалярное произведение в $\mathbb{R}^n$. Теорема 1.5(в) позволяет доказывать только то, что для любой пары функций $f, g \in C_b^1(\mathbb{R}^n) \cap \mathcal{M}(n)$ справедливо неравенство $cov(f(X), g(X)) \geqslant 0$. Возьмем гауссовский вектор $Z \sim N(0, \Sigma)$, не зависящий от X, и введем $Y(\lambda) = \lambda X + (1 - \lambda^2)^{1/2} Z$ для $\lambda \in [0, 1]$. Очевидно, $Y(\lambda) \sim N(0, \Sigma)$ и $cov(X_i, Y_j(\lambda)) = \lambda \sigma_{ij}$, $i, j = 1, \ldots, n$. Рассмотрим функцию

$$F(\lambda) = \mathsf{E} f(X) g(Y(\lambda)), \quad \lambda \in [0, 1].$$

Легко доказать (пользуясь теоремой о мажорированной сходимости), что F непрерывна на отрезке $[0, 1]$ и $F(1) - F(0) = cov(f(X), g(X))$. Поэтому нам достаточно убедиться в наличии производной $F'(\lambda) \geqslant 0$ при $\lambda \in (0, 1)$. Для таких λ можно рассмотреть условную плотность[1])

[1]) См., напр., [94, гл. II, § 7.6].

$$p_{Y(\lambda)|X=x}(y) = \frac{p_{Y(\lambda),X}(y,x)}{p_X(x)} = \frac{\partial^n}{\partial y_1 \ldots \partial y_n} \mathsf{P}(Y_1(\lambda) \leqslant y_1, \ldots, Y_n(\lambda) \leqslant y_n | X = x),$$

где $p_{Y(\lambda),X}$ и p_X — это соответственно плотности случайных векторов $(Y(\lambda), X)$ и X. По лемме 1.22 имеем

$$p(\lambda, x, y) := p_{Y(\lambda)|X=x}(y) = \frac{\partial^n}{\partial y_1 \ldots \partial y_n} \mathsf{P}\left(Z_1 \leqslant \frac{y_1 - \lambda x_1}{\sqrt{1-\lambda^2}}, \ldots, Z_n \leqslant \frac{y_n - \lambda x_n}{\sqrt{1-\lambda^2}} \right)$$
$$= (1-\lambda^2)^{-n/2} \phi((1-\lambda^2)^{-1/2}(\lambda x - y)), \quad x, y \in \mathbb{R}^n.$$

Следовательно, при $\lambda \in (0,1)$

$$F(\lambda) = \int\limits_{\mathbb{R}^n} \int\limits_{\mathbb{R}^n} f(x) g(y) p_{Y(\lambda),X}(y,x) dy dx = \int\limits_{\mathbb{R}^n} f(x) p_X(x) \int\limits_{\mathbb{R}^n} p(\lambda, x, y) g(y) dy dx$$
$$= \int\limits_{\mathbb{R}^n} \phi(x) f(x) \int\limits_{\mathbb{R}^n} p(\lambda, x, y) g(y) dy dx = \int\limits_{\mathbb{R}^n} \phi(x) f(x) h(\lambda, x) dx,$$

где $h(\lambda, x) = \int_{\mathbb{R}^n} p(\lambda, x, y) g(y) dy$. Положим

$$\phi_\lambda(x) = (1-\lambda^2)^{-n/2} \phi((1-\lambda^2)^{-1/2} x),$$

тогда

$$h(\lambda, x) = (\phi_\lambda * g)(\lambda x) = \int\limits_{\mathbb{R}^n} \phi_\lambda(\lambda x - y) g(y) dy = \int\limits_{\mathbb{R}^n} \phi_\lambda(y) g(\lambda x - y) dy.$$

Значит, функция $h(\lambda, x)$ при фиксированном $\lambda \in (0,1)$ имеет ограниченные, непрерывные и неотрицательные частные производные $\partial h / \partial x_k$, $k = 1, \ldots, n$.

Лемма 2.2. *Для любых $x, y \in \mathbb{R}^n$ и $\lambda \in (0,1)$ функция $p(\lambda, x, y)$ удовлетворяет дифференциальному уравнению в частных производных*

$$\frac{\partial p}{\partial \lambda} = -\frac{1}{\lambda} \left(\sum_{j,k=1}^n \sigma_{jk} \frac{\partial^2 p}{\partial x_j \partial x_k} - \sum_{j=1}^n x_j \frac{\partial p}{\partial x_j} \right). \tag{2.1}$$

Доказательство. Для фиксированных y и λ все функции в формуле (2.1) принадлежат пространству $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$. Следовательно, достаточно доказать равенство преобразований Фурье обеих частей (2.1). Здесь мы для простоты определяем преобразование Фурье интегрируемой функции $h : \mathbb{R}^n \to \mathbb{R}$ как

$$F[h](t) = \int\limits_{\mathbb{R}^n} e^{i(t,x)} h(x) dx, \quad i^2 = -1.$$

Равенство, которое мы хотим получить, имеет вид

$$\frac{\partial F[p](t)}{\partial\lambda}=\frac{1}{\lambda}\left(\sum_{j,k=1}^{n}\sigma_{jk}t_jt_kF[p](t)-\sum_{j=1}^{n}\frac{\partial}{\partial t_j}\left(t_jF[p](t)\right)\right). \qquad (2.2)$$

Поскольку при данных $\lambda\in(0,1)$ и $y\in\mathbb{R}^n$ функция $q=q(x)=\lambda^n p(\lambda,x,y)$ есть плотность нормального распределения со средним $\lambda^{-1}y$ и матрицей ковариаций $\lambda^{-2}(1-\lambda^2)\Sigma$, то ее преобразование Фурье — это соответствующая характеристическая функция, т.е.

$$\lambda^n F[p](t)=\exp\left\{i(t,\lambda^{-1}y)-\frac{1}{2}(\lambda^{-2}-1)(\Sigma t,t)\right\},\quad t\in\mathbb{R}^n.$$

Несложная проверка показывает, что эта функция удовлетворяет (2.2). □

Заметим теперь, что так как p экспоненциально убывает на бесконечности, то $\partial h(\lambda,x)/\partial\lambda=\int_{\mathbb{R}^n}g(y)(\partial p(\lambda,x,y)/\partial\lambda)dy$. То же верно при взятии производных по x_j, где $j=1,\dots,n$. Согласно лемме 2.2, учитывая определение $F(\lambda)$, мы получаем, что

$$F'(\lambda)=-\frac{1}{\lambda}\int\limits_{\mathbb{R}^n}\phi(x)f(x)\left(\sum_{j,k=1}^{n}\sigma_{jk}\frac{\partial^2 h(\lambda,x)}{\partial x_j\partial x_k}-\sum_{j=1}^{n}x_j\frac{\partial h(\lambda,x)}{\partial x_j}\right)dx. \qquad (2.3)$$

Лемма 2.3. *Для каждого $\lambda\in(0,1)$ справедливо равенство*

$$F'(\lambda)=\frac{1}{\lambda}\int\limits_{\mathbb{R}^n}\phi(x)\left(\sum_{j,k=1}^{n}\sigma_{jk}\frac{\partial f(x)}{\partial x_j}\frac{\partial h(\lambda,x)}{\partial x_k}\right)dx.$$

Доказательство. Рассмотрим симметричную матрицу $D=(d_{jk})=\Sigma^{-1}$. Имеем

$$\int\limits_{\mathbb{R}^n}\phi(x)\left(\sum_{j,k=1}^{n}\sigma_{jk}f(x)\frac{\partial^2 h(\lambda,x)}{\partial x_j\partial x_k}\right)dx+\int\limits_{\mathbb{R}^n}\phi(x)\left(\sum_{j,k=1}^{n}\sigma_{jk}\frac{\partial f(x)}{\partial x_j}\frac{\partial h(\lambda,x)}{\partial x_k}\right)dx$$

$$=(2\pi)^{-n/2}(\det\Sigma)^{-1/2}\int\limits_{\mathbb{R}^n}e^{-(Dx,x)/2}\sum_{j,k=1}^{n}\sigma_{jk}\frac{\partial}{\partial x_j}\left(f(x)\frac{\partial h(\lambda,x)}{\partial x_k}\right)dx$$

$$=\frac{1}{((2\pi)^n\det\Sigma)^{1/2}}\int\limits_{\mathbb{R}^n}\sum_{j,k=1}^{n}\sigma_{jk}f(x)\frac{\partial h(\lambda,x)}{\partial x_k}e^{-(Dx,x)/2}\frac{\partial}{\partial x_j}\frac{1}{2}\left(\sum_{l,m=1}^{n}d_{lm}x_lx_m\right)dx,$$

где применено интегрирование по частям (по переменной x_j). Используя символ Кронекера δ_{mk}, последнее выражение можно переписать так:

$$\int_{\mathbb{R}^n}\phi(x)f(x)\sum_{j,k=1}^{n}\sigma_{jk}\frac{\partial h(\lambda,x)}{\partial x_k}\sum_{m=1}^{n}d_{mj}x_m dx=\int_{\mathbb{R}^n}\phi(x)f(x)\sum_{k,m=1}^{n}\frac{\partial h(\lambda,x)}{\partial x_k}x_m\sum_{j=1}^{n}d_{mj}\sigma_{jk}dx$$
$$=\int_{\mathbb{R}^n}\phi(x)f(x)\sum_{k=1}^{n}\frac{\partial h(\lambda,x)}{\partial x_k}\sum_{m=1}^{n}x_m\delta_{mk}dx=\int_{\mathbb{R}^n}\phi(x)f(x)\sum_{k=1}^{n}\frac{\partial h(\lambda,x)}{\partial x_k}x_k dx,$$

что вместе с (2.3) приводит к искомому результату. □

Продолжим доказательство теоремы 2.1. Заметим, что если $\sigma_{jk}\geqslant 0$ для всех $j,k\in\{1,\dots,n\}$, то $F'(\lambda)\geqslant 0$. Следовательно, $F(1)-F(0)=cov(f(X),g(X))\geqslant 0$.

Теперь рассмотрим случай $\det\Sigma=0$. Тогда вектор $X_k=X+k^{-1}W$, где случайный вектор $W\sim N(0,I_n)$ не зависит от X (здесь $k\in\mathbb{N}$, а I_n — единичная матрица порядка $n\times n$), ассоциирован по уже доказанной части теоремы. Осталось устремить $k\to\infty$ и применить теорему 1.8(д). Теорема доказана. □

Замечание 2.4. Пусть $X=\{X_t,\,t\in T\}$ — гауссовская система. Тогда свойство $X\in\mathbf{A}$ равносильно $X\in\mathbf{PA}$ (в самом деле, если $X\in\mathbf{PA}$, то $cov(X_t,X_s)\geqslant 0$ для любых $t,s\in T$, и применима теорема 2.1).

Следствие 2.5 ([218]). *Пусть $W=(W_1(t),\dots,W_m(t))_{t\geqslant 0}$ — m-мерное броуновское движение. Предположим, что u_i, $V_{ij}:\mathbb{R}\to\mathbb{R}$ — измеримые функции, для которых $\int_0^t(|u_i(s)|+V_{ij}^2(s))ds<\infty$ при всех $t\geqslant 0$, $i=1,\dots,k$ и $j=1,\dots,m$. Положим*

$$X_i(t)=x_{0,i}+\int_0^t u_i(s)ds+\sum_{j=1}^{m}\int_0^t V_{ij}(s)dW_j(s),$$

где $x_{0,i}\in\mathbb{R}$, $i=1,\dots,k$. Если, кроме того, $V_{ij}(t)V_{lj}(t)\geqslant 0$ для всех $t\geqslant 0$ и $i,l\in\{1,\dots,k\}$, $j\in\{1,\dots,m\}$, то $\{X_i(t),\,i=1,\dots,k,\,t\geqslant 0\}\in\mathbf{A}$.

Доказательство. Рассматриваемая случайная система гауссовская, так что по теореме 2.1 достаточно проверить ее неотрицательную коррелированность. Компоненты броуновского движения независимы, поэтому, используя свойства интеграла Ито (см., напр., [27, гл. 12]), при любых $t,s\geqslant 0$ имеем

$$cov(X_i(t),X_l(s))=\sum_{j=1}^{m}cov\left(\int_0^t V_{ij}(u)dW_j(u),\int_0^s V_{lj}(z)dW_j(z)\right)$$
$$=\sum_{j=1}^{m}\mathsf{E}\int_0^t V_{ij}(u)dW_j(u)\int_0^s V_{lj}(z)dW_j(z)=\sum_{j=1}^{m}\int_0^{s\wedge t}V_{ij}(z)V_{lj}(z)dz\geqslant 0.\ \square$$

Теорема 2.6 ([275]). *Пусть $X=(X_1,\dots,X_n)$ — гауссовский случайный вектор, $n\geqslant 2$. Тогда $X\in\mathbf{NA}$ тогда и только тогда, когда его компоненты неположительно коррелированы.*

Доказательство. Воспользуемся рассуждениями, использованными при доказательстве теоремы 2.1. Единственное отличие: если все σ_{jk}, у которых $j \neq k$, неположительны, а аргументы функций f и g принадлежат непересекающимся множествам, то по лемме 2.3 имеем $F'(\lambda) \leqslant 0$. В самом деле, при каждом $i = 1, \ldots, n$ хотя бы одна из производных $\partial f(x)/\partial x_i$, $\partial g(\lambda, x)/\partial x_i$ равна нулю всюду. $\square$

В [275] дано более простое доказательство теоремы 2.6, основанное на результате [274] (близком к теореме 2.1, но более узком).

Определение 2.7. Вероятностная мера μ на $\mathcal{B}(\mathbb{R}^n)$ называется *корреляционной мерой*, если для всех замкнутых центрально-симметричных выпуклых множеств $A, B \subset \mathbb{R}^n$ справедливо неравенство

$$\mu(A \cap B) \geqslant \mu(A)\mu(B).$$

Пусть $\mu_n = N(0, I_n)$ — стандартная гауссовская мера на $\mathcal{B}(\mathbb{R}^n)$, $n \geqslant 1$. Известная гипотеза утверждает, что μ_n — корреляционная мера при каждом $n \geqslant 1$. При $n = 1$ она, очевидно, верна, а при $n = 2$ ее доказал Питт [368]. При $n \geqslant 3$ решения проблемы пока нет.

Льюис и Притчард [308] показали, что для любых $n \geqslant 3$ и p, принадлежащего интервалу $(0, 2/(n-1))$, на $\mathbb{R}^n$ существует такая сферически симметричная корреляционная мера μ, что

$$\lim_{r \to \infty} r^{-p} \log \mu(B_r(0)^c) = -1,$$

здесь $B_r(0) = \{x \in \mathbb{R}^n : \|x\| \leqslant r\}$, $r > 0$, а A^c обозначает дополнение множества A. Более того, если μ — симметричная корреляционная мера на $\mathcal{B}(\mathbb{R}^n)$, $n \geqslant 2$, и ее носитель не содержится ни в каком одномерном подпространстве $\mathbb{R}^n$, то найдется такое $a > 0$, что

$$\int\limits_{\mathbb{R}^n} e^{a\|x\|^2} d\mu(x) = \infty.$$

Это дает примеры положительно коррелированных индикаторов борелевских множеств A и B, не являющихся возрастающими.

Упомянем еще неравенство Арже, основанное на теореме Каффарелли [162] из теории оптимальной транспортировки масс. В его доказательстве применяются также свойства полугрупп Орнштейна–Уленбека. В следующей теореме подразумевается, что все интегралы берутся по $\mathbb{R}^n$ и существуют.

Теорема 2.8 ([251]). *Пусть f и g — выпуклые функции на $\mathbb{R}^n$ и $\mu = N(0, I_n)$. Тогда*

$$\int fg d\mu \geqslant \Big(1 + (m(f), m(g))\Big) \int f d\mu \int g d\mu,$$

где

$$m(f) = \int x \frac{f(x)}{\int f d\mu} d\mu(x), \quad m(g) = \int x \frac{g(x)}{\int g d\mu} d\mu(x),$$

а $(\cdot, \cdot)$ есть скалярное произведение в $\mathbb{R}^n$.

В частности, если $m(f) = 0$ или $m(g) = 0$, то неравенство принимает вид

$$\int fgd\mu \geqslant \int fd\mu \int gd\mu.$$

Во многих случаях $\int xfd\mu = \int \nabla fd\mu$, где ∇ — градиент, поэтому при $g = f$ имеем

$$\left(\int \nabla fd\mu\right)^2 \leqslant \int f^2 d\mu - \left(\int fd\mu\right)^2.$$

Последнее неравенство интересно сравнить с известным результатом Пуанкаре:

$$\int f^2 d\mu - \left(\int fd\mu\right)^2 \leqslant \int \|\nabla f\|^2 d\mu.$$

Мы не имеем возможности обсуждать здесь связь этой области исследований с логарифмическими неравенствами Соболева и отсылаем, например, к [396].

2°. Ассоциированность мер на частично упорядоченных множествах. Определения 1.1—1.3 по существу опираются только на понятие неубывающей функции. Поэтому их легко обобщить, рассматривая частично упорядоченные пространства, на которых стандартным образом вводятся монотонные функции. Пусть $(S, \mathcal{B})$ — частично упорядоченное пространство (см., напр., [56, гл. I, § 4, п. 1]) с частичным порядком $\leqslant_S$. Символом $\leqslant_S$ мы здесь пользуемся, чтобы избежать смешения с обычным порядком $\leqslant$ в $\mathbb{R}$.

Определение 2.9. Функция $f : S \to \mathbb{R}$ называется $\leqslant_S$-*возрастающей*, если $(x, y \in S$ и $x \leqslant_S y) \Longrightarrow f(x) \leqslant f(y)$.

Определение 2.10. Вероятностная мера μ на $(S, \mathcal{B})$ — *положительно коррелированная*, или *ассоциированная*, если

$$\int_S fgd\mu \geqslant \int_S fd\mu \int_S gd\mu \tag{2.4}$$

для любой пары ограниченных $\leqslant_S$-возрастающих $\mathcal{B}|\mathcal{B}(\mathbb{R})$-измеримых функций $f, g : S \to \mathbb{R}$. Случайный элемент X на вероятностном пространстве $(\Omega, \mathcal{F}, \mathsf{P})$ со значениями в $(S, \mathcal{B})$ ассоциирован, если этим свойством обладает мера $Law(X)$.

Чтобы подчеркнуть, что данное определение зависит от выбора частичного порядка, мы иногда будем писать $(S, \mathcal{B}, \mu, \leqslant_S) \in \mathbf{A}$.

Очевидно, мера $\mu = \delta_x$ (*мера Дирака, сосредоточенная в точке* $x \in S$) ассоциирована при любом выборе частичного порядка в S. Условие (1.6) — частный случай (2.4), соответствующий выбору $S = \mathbb{R}^n$ и частичному порядку (1.8). Заметим, что в противоположность теореме 1.8(а) случайный элемент со значениями в S может не быть ассоциированным (см. далее теорему 2.17). Однако справедлив следующий аналог теоремы 1.5.

Лемма 2.11. $(S, \mathcal{B}, \mu, \leqslant_S) \in \mathbf{A}$ *тогда и только тогда, когда*

$$\mu(A \cap B) \geqslant \mu(A)\mu(B) \tag{2.5}$$

для любых $\leqslant_S$-возрастающих множеств $A, B \in \mathcal{B}$ (т.е. таких, что $\mathbb{I}_A$ и $\mathbb{I}_B$ возрастают на $(S, \leqslant_S)$).

Доказательство. Необходимость очевидна.

Достаточность. Возьмем такую $\leqslant_S$-возрастающую функцию $f : S \to \mathbb{R}$, что $f \in \mathcal{B}|\mathcal{B}(\mathbb{R})$ и $0 \leqslant f < 1$, и пусть $n \in \mathbb{N}$. Положим

$$f_n(x) := \sum_{k=1}^{n} \frac{k}{n}\mathbb{I}\left\{\frac{k-1}{n} \leqslant f(x) < \frac{k}{n}\right\} \equiv \frac{1}{n}\sum_{k=1}^{n}\mathbb{I}\left\{f(x) \geqslant \frac{k-1}{n}\right\},\ x \in S. \quad (2.6)$$

Очевидно, функция $\mathbb{I}\{x : f(x) \geqslant v\}$ является $\leqslant_S$-возрастающей при каждом $v \in [0, 1)$, и $f_n(x) \to f(x)$ для всех $x \in S$ при $n \to \infty$. Общий случай легко сводится к уже разобранному. □

Мы приведем несколько важных примеров ассоциированных мер, для чего сейчас введем новые обозначения и докажем несколько вспомогательных результатов. Теорему 1.8(г) обобщает

Теорема 2.12. *Пусть $(S_i, \mathcal{B}_i, \mathsf{P}_i, \leqslant_i)$ — частично упорядоченные вероятностные пространства, $i = 1, 2$. Пусть функция $h : S_1 \to S_2$ возрастает (т.е. из $x \leqslant_1 y$ следует, что $h(x) \leqslant_2 h(y)$) и $\mathcal{B}_1|\mathcal{B}_2$-измерима. Если при этом $\mathsf{P}_2 = Law(h)$ и $(S_1, \mathcal{B}_1, \mathsf{P}_1, \leqslant_1) \in \mathbf{A}$, то $(S_2, \mathcal{B}_2, \mathsf{P}_2, \leqslant_2) \in \mathbf{A}$.*

Доказательство. Возьмем возрастающие множества $A_2, B_2 \in \mathcal{B}_2$. Тогда множества $A_1 := h^{-1}(A_2)$ и $B_1 := h^{-1}(B_2)$ в $(S_1, \leqslant_1)$ возрастают. Действительно, если, например, $x_1 \in A_1$ и $y_1 \geqslant_1 x_1$, то $h(x_1) \in A_2$ и $h(y_1) \geqslant_2 h(x_1)$, и так как A_2 $\leqslant_2$-возрастает, то $h(y_1) \in A_2$. Следовательно, $y_1 \in h^{-1}(A_2) = A_1$. К тому же $\mathsf{P}_2 = \mathsf{P}_1 h^{-1}$, поэтому

$$\mathsf{P}_2(A_2 \cap B_2) - \mathsf{P}_2(A_2)\mathsf{P}_2(B_2) = \mathsf{P}_1(A_1 \cap B_1) - \mathsf{P}(A_1)\mathsf{P}_1(B_1) \geqslant 0,$$

где последнее неравенство вытекает из леммы 2.11. □

Следствие 2.13. *Пусть $(S_t, \mathcal{B}_t, \preceq_t)_{t\in\mathbb{T}}$ — семейство частично упорядоченных пространств, $S_\mathbb{T} = \prod_{t\in\mathbb{T}} S_t$, $\mathcal{B}_\mathbb{T}$ — цилиндрическая σ-алгебра в $S_\mathbb{T}$, а*

$$x \preceq_\mathbb{T} y \ \text{для}\ x = \{x_t, t \in \mathbb{T}\},\ y = \{y_t, t \in \mathbb{T}\} \in S_\mathbb{T} \iff x_t \preceq_t y_t,\ \ t \in \mathbb{T}. \quad (2.7)$$

Если $\mu_\mathbb{T}$ — ассоциированная мера на $(S_\mathbb{T}, \mathcal{B}_\mathbb{T}, \preceq_\mathbb{T})$, то для каждого множества $U \subset \mathbb{T}$ проекция $\mu_{\mathbb{T},U}$ на $(S_U, \mathcal{B}_U, \preceq_U)$ тоже ассоциированная (пространство $(S_U, \mathcal{B}_U, \preceq_U)$ определяется аналогично $(S_\mathbb{T}, \mathcal{B}_\mathbb{T}, \preceq_\mathbb{T})$, а $\mu_{\mathbb{T},U} = \pi_{\mathbb{T},U}^{-1}\mu_\mathbb{T}$, где $\pi_{\mathbb{T},U}x = x|_U$ — это ограничение функции $x \in S_\mathbb{T}$ на U).

Пусть на S заданы два частичных порядка $\leqslant_1$ и $\leqslant_2$. Говорят, что $\leqslant_2$ *тоньше, чем* $\leqslant_1$, если для любых $x, y \in S$ соотношение $x \leqslant_1 y$ влечет $x \leqslant_2 y$. В частности, если функция $f : S \to \mathbb{R}$ возрастает на $(S, \leqslant_2)$, то она возрастает и на $(S, \leqslant_1)$. Следовательно, если $(S, \mathcal{B}, \mu, \leqslant_1) \in \mathbf{A}$, то $(S, \mathcal{B}, \mu, \leqslant_2) \in \mathbf{A}$.

На пространстве $S = C_0[0, T] = \{f \in C[0, T] : f(0) = 0\}$ введем частичный порядок "inc", основанный на приращениях функций. Будем считать, что

$$x \leqslant_{\mathrm{inc}} y,\ \text{ если }\ x(t) - x(s) \leqslant y(t) - y(s)\ \text{ для любых }\ 0 \leqslant s \leqslant t \leqslant T. \quad (2.8)$$

Кроме того, в S легко ввести обычный порядок $\preceq$ типа (2.7), полагая для $x, y \in S$

$$x \preceq y,\ \text{ если }\ x(t) \leqslant y(t)\ \text{ при всех }\ t \in [0, T]. \quad (2.9)$$

Заметим, что порядок $\preceq$ тоньше, чем $\leqslant_{\mathrm{inc}}$.

Пусть $\mathbb{W}$ — мера Винера на $(S, \mathcal{B}(S))$, т.е. распределение стандартного винеровского процесса W, рассматриваемого на отрезке $[0,T]$. Следующий результат принадлежит Барбато.

Теорема 2.14 ([112]). *Пусть $S = C_0[0,T]$, $\mathcal{B} = \mathcal{B}(C_0[0,T])$ и $\mu = \mathbb{W}$. Тогда мера μ ассоциирована, т.е. $(S, \mathcal{B}, \mu, \leqslant_{\mathrm{inc}}) \in \mathbf{A}$.*

Доказательство. Удобно перейти к каноническому заданию винеровского процесса, т.е. считать, что $(\Omega, \mathcal{F}, \mathsf{P}) = (S, \mathcal{B}, \mu)$ и $W(\omega) = \omega$ для всех $\omega \in S$.

Вначале проверим (2.5) для некоторого специального класса событий A и B. Рассмотрим разбиение $H = \{t_1, \ldots, t_n\}$ отрезка $[0,T]$ точками $0 = t_0 < \ldots < t_n = T$ $(n \in \mathbb{N})$. Обозначим $X_i(\omega) := W_{t_i}(\omega) - W_{t_{i-1}}(\omega)$, $\omega \in \Omega$, $i = 1, \ldots, n$. Тогда $X_H = (X_1, \ldots, X_n) \in \mathbf{A}$ по теореме 1.8(в). Пусть $\mathcal{A}_H := \sigma\{X_H\}$ и введен частичный порядок для $x, y \in S$:

$$x \leqslant_H y \iff x(t_i) - x(t_{i-1}) \leqslant y(t_i) - y(t_{i-1}), \quad i = 1, \ldots, n. \tag{2.10}$$

Теперь $X_H \in \mathbf{A}$ можно переписать в виде $(S, \mathcal{A}_H, \mu, \leqslant_H) \in \mathbf{A}$. Порядок $\leqslant_H$ тоньше $\leqslant_{\mathrm{inc}}$, однако верна

Лемма 2.15. *Множество $A \in \mathcal{A}_H$ является $\leqslant_{\mathrm{inc}}$-возрастающим тогда и только тогда, когда оно $\leqslant_H$-возрастающее.*

Доказательство. Необходимость. Возьмем $\leqslant_{\mathrm{inc}}$-возрастающее A и произвольные $x \in A$ и $y \geqslant_H x$ (последнее означает, что $x \leqslant_H y$). Нужно убедиться, что $y \in A$. Пусть $p_i = y(t_i) - x(t_i)$, $i = 0, \ldots, n$. Тогда ввиду (2.10) соотношение $x \leqslant_H y$ есть то же самое, что

$$0 = p_0 \leqslant p_1 \leqslant \ldots \leqslant p_n.$$

Построим такую $\leqslant_{\mathrm{inc}}$-возрастающую функцию $f \in S$, для которой $f(t_i) = p_i$, $i = 0, \ldots, n$. Пусть $z(t) := x(t) + f(t)$, $t \in [0,T]$. Очевидно, что $z \geqslant_{\mathrm{inc}} x$, т.е. $x \leqslant_{\mathrm{inc}} z$, и потому $z \in A$ (множество A по условию $\leqslant_{\mathrm{inc}}$-возрастающее). В то же время $y(t_i) = z(t_i)$, $i = 0, \ldots, n$. Поскольку $A \in \mathcal{A}_H$, то элемент $y \in A$. Действительно, любое множество A из класса $\mathcal{A}_H$ определяется значениями его элементов в точках $t_i \in H$, $i = 0, \ldots, n$.

Достаточность. Возьмем $\leqslant_H$-возрастающее A, а также элементы $x \in A$ и $y \geqslant_{\mathrm{inc}} x$. Тогда $x \leqslant_H y$, так как порядок $\leqslant_H$ тоньше $\leqslant_{\mathrm{inc}}$. Итак, $y \in A$, и, значит, A есть $\leqslant_{\mathrm{inc}}$-возрастающее множество. $\square$

Рассмотрим алгебру $\mathcal{A} := \cup_H \mathcal{A}_H$, где объединение берется по всем конечным разбиениям $[0,T]$. Возьмем произвольные $\leqslant_{\mathrm{inc}}$-возрастающие множества $A, B \in \mathcal{A}$. Заметим, что $H' \subset H''$ влечет $\mathcal{A}_{H'} \subset \mathcal{A}_{H''}$, поэтому найдется такое разбиение H_0, что $A, B \in \mathcal{A}_{H_0}$. По лемме 2.15 множества A и B являются $\leqslant_{H_0}$-возрастающими. Значит, $(S, \mathcal{A}_{H_0}, \mu, \leqslant_{H_0}) \in \mathbf{A}$. Таким образом, $cov(\mathbb{I}_A, \mathbb{I}_B) \geqslant 0$ для любых $\leqslant_{\mathrm{inc}}$-возрастающих $A, B \in \mathcal{A}$. Нам остается доказать следующую лемму.

Лемма 2.16. *Для любых $\varepsilon > 0$ и $\leqslant_{\mathrm{inc}}$-возрастающего $B \in \mathcal{B}$ существует такое $\leqslant_{\mathrm{inc}}$-возрастающее множество $A \in \mathcal{A}$, что $\mu(A \triangle B) < \varepsilon$.*

Доказательство. Поскольку $\mathcal{B} = \sigma\{\mathcal{A}\}$, найдется такое $C \in \mathcal{A}$, что $\mu(B \triangle C) < \varepsilon$. Тогда $C \in \mathcal{A}_H$ для некоторого разбиения H. Построим $\leqslant_{\mathrm{inc}}$-возрастающее приближающее множество $A \in \mathcal{A}_H$.

Рассмотрим множество $E \subset S$ функций f, линейных на каждом из отрезков $[t_{i-1}, t_i]$, $i = 1, \ldots, n$, разбиения H, и пусть

$$F = \{f \in S : f(t) = 0,\, t \in H\}.$$

Для любой $f \in S$ определим $\Pi_H f$ как функцию, график которой получается линейной интерполяцией точек $(t_i, f(t_i))$, $i = 1, \ldots, n$. Тогда E и F — линейные подпространства S и $S = E \oplus F$. Последнее соотношение означает, что произвольная функция $f \in S$ однозначно представляется в виде $f = \Pi_H f + (f - \Pi_H f)$, где $\Pi_H f \in E$ и $g = f - \Pi_H f \in F$. Пусть σ-алгебры $\mathcal{E}$ и $\mathcal{G}$ — следы $\mathcal{B}$ на E и F соответственно[1]. Положим

$$W(\omega) = \Pi_H W(\omega) + (W(\omega) - \Pi_H W(\omega) =: Y(\omega) + Z(\omega).$$

Тогда случайные элементы Y и Z на вероятностном пространстве $(S, \mathcal{B}, \mu)$ принимают значения соответственно в $(E, \mathcal{E})$ и $(F, \mathcal{G})$. Более того, $\mathcal{A}_H = \sigma\{Y\}$, причем элементы Y и Z независимы. Чтобы это проверить, достаточно заметить, что случайные процессы Y и Z вместе образуют гауссовскую систему и $cov(Y(s), Z(t)) = 0$ для любых $s, t \in [0,1]$. Пусть $\mathbb{P}_1 := Law(Y)$ и $\mathbb{P}_2 := Law(Z)$.

Для функции $\omega = f + g$, где $f \in E$ и $g \in F$, введем

$$\varphi(\omega) := \mathsf{E}(\mathbb{I}\{W \in B\} | Y = f) = \mathsf{E}(\mathbb{I}_B(Y + Z) | Y = f),$$

здесь E — математическое ожидание по мере μ. Согласно лемме 1.22 и (1.9)

$$\varphi(\omega) = \int_F \mathbb{I}_B(f + g)\, d\mathbb{P}_2(g). \tag{2.11}$$

Поэтому для любого множества $D \in \sigma\{Y\}$ с помощью простейших свойств условных математических ожиданий получаем

$$\begin{aligned}
\mu(B \triangle D) &= \mathsf{E}|\mathbb{I}_B - \mathbb{I}_D| = \mathsf{E}(1 - \mathbb{I}\{W \in B\})\mathbb{I}_D + \mathsf{E}\mathbb{I}\{W \in B\}\mathbb{I}_{D^c} \\
&= \mathsf{E}(\mathsf{E}((1 - \mathbb{I}\{W \in B\})\mathbb{I}_D)|Y) + \mathsf{E}(\mathsf{E}(\mathbb{I}\{W \in B\}\mathbb{I}_{D^c}|Y)) \\
&= \mathsf{E}((1 - \varphi)\mathbb{I}_D) + \mathsf{E}(\varphi \mathbb{I}_{D^c}) = \mathsf{E}|\varphi - \mathbb{I}_D|.
\end{aligned}$$

Очевидно, $0 \leqslant \varphi \leqslant 1$, следовательно,

$$\begin{aligned}
|\varphi - \mathbb{I}_D| &= (1 - \varphi)\mathbb{I}_D + \varphi \mathbb{I}_{D^c} \\
&= (1 - \varphi)\mathbb{I}_{D \cap \{\varphi \geqslant 1/2\}} + (1 - \varphi)\mathbb{I}_{D \cap \{\varphi < 1/2\}} + \varphi \mathbb{I}_{D^c \cap \{\varphi \geqslant 1/2\}} + \varphi \mathbb{I}_{D^c \cap \{\varphi < 1/2\}} \\
&\geqslant (1 - \varphi)\mathbb{I}_{D \cap \{\varphi \geqslant 1/2\}} + \varphi \mathbb{I}_{D \cap \{\varphi < 1/2\}} + (1 - \varphi)\mathbb{I}_{D^c \cap \{\varphi \geqslant 1/2\}} + \varphi \mathbb{I}_{D^c \cap \{\varphi < 1/2\}} \\
&= (1 - \varphi)\mathbb{I}_{\{\varphi \geqslant 1/2\}} + \varphi \mathbb{I}_{\{\varphi < 1/2\}} = |\varphi - \mathbb{I}_A|,
\end{aligned}$$

[1] Т.е., например, $\mathcal{E} = \{R \cap E : R \in \mathcal{B}\}$.

где $A = \{\varphi \geqslant 1 - \varphi\} = \{\varphi \geqslant 1/2\} \in \mathcal{A}_H$, D^c обозначает дополнение D в S. Последнее включение верно, потому что $\sigma\{Y\} = \mathcal{A}_H$. В результате получаем

$$\mu(A \triangle B) = \mathsf{E}|\varphi - \mathbb{I}_A| \leqslant \mathsf{E}|\varphi - \mathbb{I}_C| = \mu(B \triangle C) < \varepsilon. \quad \square$$

Возвращаясь к доказательству теоремы, покажем, что если B $\leqslant_{\text{inc}}$-возрастающее, то функция φ возрастает в смысле частичного порядка $\leqslant_{\text{inc}}$. Пусть $\omega_1, \omega_2 \in S$ и $\omega_1 \leqslant_{\text{inc}} \omega_2$. Рассмотрим разложение $\omega_i = f_i + g_i$, $f_i \in E$, $g_i \in F$, $i = 1, 2$. В силу (2.11) имеем

$$\varphi(\omega_i) = \int\limits_F \mathbb{I}_B(f_i + g) d\mathbb{P}_2(g), \quad i = 1, 2.$$

Так как $\omega_1 \leqslant_{\text{inc}} \omega_2$, то $f_1 \leqslant_{\text{inc}} f_2$ и $f_1 + g \leqslant_{\text{inc}} f_2 + g$ при любой $g \in F$. Поскольку множество B $\leqslant_{\text{inc}}$-возрастающее, отсюда вытекает неравенство $\mathbb{I}_B(f_1 + g) \leqslant \mathbb{I}_B(f_2 + g)$. Значит, $\varphi(\omega_1) \leqslant \varphi(\omega_2)$ ввиду (2.11). Итак, множество A является $\leqslant_{\text{inc}}$-возрастающим, что завершает доказательство теоремы. $\square$

Пусть $(S, \mathcal{B}, \leqslant)$ — частично упорядоченное измеримое пространство. Для множества $B \subset S$ определим его *верхнее замыкание*

$$C_B = \{y \in S : y \geqslant x \text{ для некоторого } x \in B\}. \tag{2.12}$$

Мы предполагаем, что *частичный порядок измерим*, т.е. для любого $x \in S$ множества $\{y : y \geqslant x\}$ и $\{y : y \leqslant x\}$ измеримы (принадлежат $\mathcal{B}$). Заметим, что в этом случае $\{x\} \in \mathcal{B}$, так как

$$\{x\} = \{y : y \leqslant x\} \cap \{y : y \geqslant x\} \in \mathcal{B}, \quad \text{и } C_{\{x\}} \in \mathcal{B}.$$

Следующую теорему Линдквиста можно рассматривать как "обобщенное неравенство Чебышёва".

Теорема 2.17 ([316]). *Частично упорядоченное пространство $(S, \leqslant)$, снабженное σ-алгеброй $\mathcal{B}$, содержащей все одноточечные множества, вполне упорядочено тогда и только тогда, когда любая вероятностная мера на $\mathcal{B}$ ассоциирована.*

Доказательство. Необходимость. Пусть множество S не является вполне упорядоченным. Тогда существуют два несравнимых элемента $x, y \in S$. Рассмотрим такую вероятностную меру P, что $\mathsf{P}(\{x\}) = \mathsf{P}(\{y\}) = 1/2$, и возрастающие множества $C_{\{x\}}$ и $C_{\{y\}}$, см. (2.12). Тогда $\mathsf{P}(C_{\{x\}})\mathsf{P}(C_{\{y\}}) > 0$. Но так как $y \notin C_{\{x\}}$ и $x \notin C_{\{y\}}$, то $\mathsf{P}(C_{\{x\}} \setminus \{x\}) = \mathsf{P}(C_{\{y\}} \setminus \{y\}) = 0$, поэтому $\mathsf{P}(C_{\{x\}} C_{\{y\}}) = 0$, и рассматриваемая мера не является ассоциированной.

Достаточность. Пусть S вполне упорядочено. Возьмем на S произвольную вероятностную меру P и любые возрастающие множества $C_1, C_2 \in \mathcal{B}$. Тогда либо $C_1 C_2^c = \varnothing$, либо $C_2 C_1^c = \varnothing$; в противном случае нашлись бы различные точки $x \in C_1 C_2^c$, $y \in C_2 C_1^c$. Но тогда не выполнялось бы ни одно из соотношений $y \geqslant x$ и $x \geqslant y$, что невозможно из-за полной упорядоченности S. Пусть, например, $C_1 C_2^c = \varnothing$. Тогда

$$\mathsf{P}(C_1 C_2) = \mathsf{P}(C_1) - \mathsf{P}(C_1 C_2^c) = \mathsf{P}(C_1) \geqslant \mathsf{P}(C_1)\mathsf{P}(C_2).$$

Аналогичная ситуация имеет место при $C_2 C_1^c = \varnothing$. Теперь утверждение вытекает из леммы 2.11. $\square$

Пример 2.18. Рассмотрим пространство $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$ и в нем обычный частичный порядок $\leqslant$ (см. (1.8)) и лексикографический порядок $\leqslant_{\mathrm{lex}}$. Возьмем бинарный случайный вектор $X = (X_1, X_2)$ со значениями в $\mathbb{R}^2$, для которого $cov(X_1, X_2) < 0$. Разумеется, $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \mathsf{P}_X, \leqslant) \notin \mathbf{A}$, а в силу теоремы 2.17 имеем $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2), \mathsf{P}_X, \leqslant_{\mathrm{lex}}) \in \mathbf{A}$. Итак, определение ассоциированности меры существенно зависит от выбора частичного порядка.

В завершение этого пункта упомянем работы Камае, Кренгеля и О'Брайена [279], а также Алсведе и Дайкина [97], где были установлены первые корреляционные неравенства в частично упорядоченных пространствах. Теорема последних авторов "о четырех функциях", обобщающая теорему Престона, подробно изложена в [1, с. 103] вместе с приложениями к комбинаторным задачам. Ассоциированность мер на таких пространствах с приложениями к байесовской теории риска рассмотрена Линдквистом [316]. В п. 4° теорема 2.14 будет использована для проверки ассоциированности решений стохастических дифференциальных уравнений.

3°. Марковские процессы. Монотонность и ассоциированность. Важный метод построения новых ассоциированных мер с помощью других известных основан на марковских пребразованиях. Используя обозначения из приложения, § 2, дадим

Определение 2.19. Однородный марковский процесс $\mathbf{X} = \{X_t, t \geqslant 0\}$ со значениями в пространстве S (в котором введен частичный порядок $\leqslant$) называется *монотонным*, если для любых $t > 0$ и неубывающей функции $f : S \to \mathbb{R}$ образ $T_t f(\cdot)$ также есть неубывающая функция, и называется *сохраняющим положительные корреляции*, если для любых $t > 0$ и ассоциированной вероятностной меры μ мера μT_t также является ассоциированной.

Другими словами, процесс $\mathbf{X} = \{X_t, t \geqslant 0\}$ сохраняет положительные корреляции, если при любом $t > 0$ и любом ассоциированном начальном распределении $Law(X_0) = \mu$ распределение X_t также ассоциировано. Очевидно, что если процесс $\mathbf{X}$ монотонный и функция $f : S \to \mathbb{R}$ не возрастает (т.е. если $(-f)$ не убывает), то функция $(T_t f)(\cdot)$ — невозрастающая.

Теорема 2.20 ([182, 253]). *Пусть однородная цепь Маркова* $\mathbf{X}$ *со значениями в конечном пространстве* S *монотонна и обладает*[1]) *инфинитезимальной матрицей* A. *Тогда* $\mathbf{X}$ *сохраняет положительные корреляции в том и только том случае, когда*

$$A(x, y) = 0 \quad \textit{для несравнимых } x, y \in S. \tag{2.13}$$

В частности, $\mathbf{X}$ *всегда сохраняет положительные корреляции, если* S *вполне упорядочено.*

Доказательство. Необходимость. Пусть $\mathbf{X}$ сохраняет положительные корреляции, но $A(x, y) > 0$ для некоторой пары несравнимых $x, y \in S$ (отметим, что если x и y несравнимы, то $x \neq y$, поэтому $A(x, y) \geqslant 0$). Возьмем $Law(X_0) = \delta_x$; эта мера ассоциирована. Как обычно, символ P_x означает, что

[1]) Т.е. конечная инфинитезимальная матрица (П.2.10) существует у семейства переходных вероятностей $(p(t, x, y))$, соответствующего $\mathbf{X}$.

мы рассматриваем процесс $\mathbf{X}$, выходящий в момент $t=0$ из точки x. По формуле полной вероятности

$$\mathsf{P}_x(X_t=y)=\sum_{z\in S}\mathsf{P}(X_0=z)\mathsf{P}(X_t=y|X_0=z)=\mathsf{P}(X_t=y|X_0=x)=p(t,x,y).$$

Согласно (П.2.10), при $t\to 0+$ справедливо соотношение

$$\mathsf{P}_x(X_t=y)=\begin{cases}1+A(x,x)t+o(t), & y=x,\\ A(x,y)t+o(t), & \text{иначе.}\end{cases}$$

Следовательно,

$$\mathsf{P}_x(X_t\geqslant x)\geqslant \mathsf{P}_x(X_t=x)\to 1,\quad t\to 0+\,.$$

Если $z\geqslant y$, то $z\neq x$ (иначе x и y сравнимы). Так как S конечно, то

$$\mathsf{P}_x(X_t\geqslant y)=\sum_{z\geqslant y}\mathsf{P}_x(X_t=z)=t\sum_{z\geqslant y}A(x,z)+o(t)$$

при $t\to 0+$, здесь $z>y$ означает, что $z\geqslant y$ и $z\neq y$. Итак,

$$\mathsf{P}_x(X_t\geqslant x)\mathsf{P}_x(X_t\geqslant y)=tA(x,y)+t\sum_{z>y}A(x,z)+o(t),\quad t\to 0+\,. \tag{2.14}$$

Более того,

$$\begin{aligned}\mathsf{P}_x(X_t\geqslant x, X_t\geqslant y)&=\sum_{z\geqslant x, z\geqslant y}\mathsf{P}_x(X_t=z)\\ &=t\sum_{z\geqslant x, z\geqslant y}A(x,z)+o(t)\leqslant t\sum_{z>y}A(x,z)+o(t)\end{aligned} \tag{2.15}$$

при $t\to 0+$, так как если одновременно $z\geqslant x$ и $z\geqslant y$, то $z\neq x$ и $z\neq y$ (следовательно, $\{z\geqslant x, z\geqslant y\}\subset\{z>y\}$). В силу (2.14) и (2.15)

$$q_t(x,y):=\mathsf{P}_x(X_t\geqslant x, X_t\geqslant y)-\mathsf{P}_x(X_t\geqslant x)\mathsf{P}_x(X_t\geqslant y)\leqslant -tA(x,y)+o(t)$$

при $t\to 0+$. Таким образом, существует такое $t_0>0$, что $q_t(x,y)<0$ при $0<t<t_0$.

С другой стороны, при каждом $w\in S$ функция $\mathbb{I}\{z: z\geqslant w\}$ неубывающая (если z и w несравнимы, то полагаем $\mathbb{I}\{z: z\geqslant w\}(z)=0$). Тогда из ассоциированности меры $Law(X_t)$ при каждом $t>0$ и монотонности функций $f(z)=\mathbb{I}\{z\geqslant x\}$ и $g(z)=\mathbb{I}\{z\geqslant y\}$ получаем, что

$$\mathsf{P}_x(X_t\geqslant x, X_t\geqslant y)=\mathsf{E}_x f(X_t)g(X_t)\geqslant \mathsf{E}_x f(X_t)\mathsf{E}_x g(X_t)=\mathsf{P}_x(X_t\geqslant x)\mathsf{P}_x(X_t\geqslant y),$$

здесь символ E_x напоминает, что $\mathsf{E}f(X_t)$ и $\mathsf{E}g(X_t)$ вычисляются для X_t, имеющего $X_0=x$. Однако последнее неравенство противоречит полученной выше оценке для $q_t(x,y)$. Необходимость (2.13) доказана.

Достаточность. Для произвольной функции $h : S \to \mathbb{R}$ и любых $x \in S$ и $u \geqslant 0$, согласно (П.2.12), (П.2.13) и (П.2.15) мы имеем

$$(T_u h)(x) = h(x) + u \sum_{y \in S} A(x, y) h(y) + \theta_1 c_1 u^2 |h|, \tag{2.16}$$

где $\theta_1 = \theta_1(h, x, u)$, $|\theta_1| \leqslant 1$ и $c_1 = \|A\|^2 e^{\|A\|}/2$. Применяя (2.16) и равенство $\sum_{y \in S} A(x, y) = 0$, из (П.2.11) получаем, что для всех $f, g : S \to \mathbb{R}$, $u \in [0, 1]$

$$\begin{aligned} T_u(fg)(x) &- T_u f(x) T_u g(x) \\ &= u \sum_{y \in S} A(x, y)(f(x) - f(y))(g(x) - g(y)) + \theta_2 c_2 u^2 |f||g|, \end{aligned} \tag{2.17}$$

где $\theta_2 = \theta_2(f, g, x, u)$, $|\theta_2| \leqslant 1$ и $c_2 = c_2(\|A\|) > 0$.

Пусть функции f и g неубывающие. Тогда для сравнимых x и y, очевидно,

$$(f(x) - f(y))(g(x) - g(y)) \geqslant 0,$$

а для несравнимых x и y в силу (2.13) имеем $A(x, y) = 0$. Поэтому для таких f, g и $u \in [0, 1]$ из (2.17) получаем неравенство

$$T_u(fg)(x) \geqslant T_u f(x) T_u g(x) - c_2 u^2 |f||g|. \tag{2.18}$$

Оператор T_t линейный и (см. (П.2.7))

$$T_t h = h, \text{ если } h(z) = c \text{ при всех } z \in S \text{и фиксированном } c \in \mathbb{R}, \tag{2.19}$$

$$T_t h \geqslant 0, \text{ если } h(z) \geqslant 0 \text{ для всех } z \in S. \tag{2.20}$$

Кроме того, для каждого $u \geqslant 0$

$$|T_u f| \leqslant \|T_u\| |f| \leqslant |f|. \tag{2.21}$$

Для произвольного фиксированного $k \in \mathbb{N}$ и $u_1, \ldots, u_k \in [0, 1]$ положим $v = u_2 + \ldots + u_k$ и $w = u_2^2 + \ldots + u_k^2$. Так как процесс $\mathbf{X}$ монотонный, то функции $T_u f$ и $T_u g$ не убывают. Поэтому из (2.18)—(2.21) индукцией по k легко вывести соотношение

$$\begin{aligned} T_{u_1+u_2+\ldots+u_k}(fg)(x) &= T_{u_1}(T_v(fg))(x) \geqslant T_{u_1}\Big((T_v f)(T_v g) - c_2 w |f||g|\Big)(x) \\ &\geqslant (T_{u_1}(T_v f))(x)(T_{u_1}(T_v g))(x) - c_2 u_1^2 |T_v f||T_v g| - c_2 w |f||g| \\ &\geqslant (T_{u_1+u_2+\ldots+u_k} f)(x)(T_{u_1+u_2+\ldots+u_k} g)(x) - c_2(u_1^2 + u_2^2 + \ldots + u_k^2)|f||g|. \end{aligned} \tag{2.22}$$

Теперь пусть $t > 0$ и $n > t$ $(n \in \mathbb{N})$. Применяя (2.22), видим, что

$$(T_t fg)(x) = (T_{n(t/n)} fg)(x) \geqslant (T_t f)(x)(T_t g)(x) - c_2 n \left(\frac{t}{n}\right)^2 |f||g|.$$

При $n \to \infty$ отсюда следует, что для любых $t > 0$, $x \in S$ и неубывающих функций $f, g : S \to \mathbb{R}$ верно неравенство

$$(T_t fg)(x) \geqslant (T_t f)(x)(T_t g)(x). \tag{2.23}$$

Если мера μ ассоциирована, то для рассматриваемых f и g из (П.2.7) и (2.23) вытекает, что при всяком $t > 0$

$$\int\limits_S fgd(\mu T_t) = \sum_{y\in S} f(y)g(y)\sum_{x\in S}\mu(x)p(t,x,y) = \int\limits_S (T_t fg)d\mu$$
$$\geqslant \int\limits_S T_t f T_t g d\mu \geqslant \int\limits_S T_t f d\mu \int\limits_S T_t g d\mu = \int\limits_S fd(\mu T_t)\int\limits_S gd(\mu T_t).$$

Итак, μT_t ассоциирована и $\mathbf{X}$ сохраняет положительные корреляции. □

Замечание 2.21. Условие (2.13) допускает интерпретацию в терминах скачков марковской цепи $\mathbf{X}$. Для этого напомним, что траектории процесса $\mathbf{X}$ с инфинитезимальной матрицей A можно построить следующим образом (см., напр., [65, с. 599]). Пусть $X_0 = x$ с вероятностью $\mu_0(\{x\})$, $x \in S$. Далее процесс $\mathbf{X}$ остается в состоянии x случайное время τ_0. Если $A(x,x) = 0$, то $\tau_0 = +\infty$ п.н., т.е. $X(t) = x$ при всех $t \geqslant 0$. Если же $A(x,x) < 0$, то τ_0 распределено показательно с параметром $-A(x,x)$. В момент τ_0 траектория совершает скачок в точку $y \neq x$ с вероятностью $-A(x,y)/A(x,x)$. Далее в состоянии y траектория проводит время τ_1, равное либо $+\infty$ п.н.(при $A(y,y) = 0$), либо показательной случайной величине с параметром $-A(y,y)$. В случае $\tau_0 \vee \tau_1 < \infty$ траектория оказывается в состоянии $z \neq y$ с вероятностью $-A(y,z)/A(y,y)$ и т.д. Итак, условие (2.13) означает, что если x и y несравнимы, то переход из x в y или обратно невозможен, за исключением события нулевой вероятности.

Простое условие на матрицу A, гарантирующее монотонность $\mathbf{X}$, дает

Теорема 2.22 ([315]). *Однородная цепь Маркова $\mathbf{X} = \{X_t, t \geqslant 0\}$ со значениями в конечном пространстве S (с σ-алгеброй $\mathcal{B} = 2^S$ и измеримым частичным порядком "$\leqslant$") монотонна, если неравенство*

$$\sum_{z\in U} A(x,z) \leqslant \sum_{z\in U} A(y,z) \tag{2.24}$$

справедливо для любого возрастающего $U \subset S$ и всех таких пар $x, y \in S$, что $x \leqslant y$ и при этом x, y одновременно лежат либо в U, либо в $S \setminus U$.

Доказательство. Возьмем неубывающую функцию $f : S \to \mathbb{R}$. Требуется показать, что для любых $x, y \in S$ таких, что $x \leqslant y$, имеет место неравенство

$$(T_t f)(x) \leqslant (T_t f)(y), \quad t > 0. \tag{2.25}$$

Поскольку T_t — линейный оператор (см. (П.2.7)), сразу считаем $0 \leqslant f < 1$.

Вначале построим такое $h > 0$ (зависящее лишь от A), что при каждом $t \in (0,h)$ функция $(I + tA)f$ не убывает, т.е. если $x, y \in S$, $x \leqslant y$, то

$$((I + tA)f)(x) - ((I + tA)f)(y) \leqslant 0. \tag{2.26}$$

Ясно, что f — поточечный предел неотрицательных линейных комбинаций индикаторов возрастающих множеств, см. (2.6). Следовательно, достаточно проверить (2.26) для функции $f = \mathbb{I}_U$, где множество U возрастающее.

Определим $h = h(A)$ формулой

$$h = (2 \max_{x \in S} \sum_{z \in S} |A(x,z)|)^{-1}, \tag{2.27}$$

где $0^{-1} := +\infty$. Если $x, y \in S$, $x \leqslant y$ и $t \in (0, h)$, то левую часть (2.26) можно записать в виде

$$f(x) + t \sum_{z \in S} A(x,z) f(z) - f(y) - t \sum_{z \in S} A(y,z) f(z)$$
$$= f(x) - f(y) + t \left(\sum_{z \in U} A(x,z) - \sum_{z \in U} A(y,z) \right) \leqslant 0. \tag{2.28}$$

В самом деле, когда x и y одновременно лежат в U или в $S \setminus U$, неравенство (2.28) вытекает из предположений теоремы (даже при всех $t > 0$). Ситуация $x \in U, y \notin U$ невозможна, так как U возрастает, а $y \geqslant x$. Наконец, если $x \notin U$ и $y \in U$, то имеем $f(x) = 0$ и $f(y) = 1$, а следовательно,

$$f(x) - f(y) + t \left(\sum_{z \in U} A(x,z) - \sum_{z \in U} A(y,z) \right) \leqslant -1 + 2t \max_{x \in S} \sum_{z \in S} |A(x,z)| \leqslant 0$$

в силу выбора h в (2.27).

Итак, если функция $f : S \to \mathbb{R}$ не убывает, то при каждом $n \in \mathbb{N}$ функция $(I + sA)^n f$ неубывающая, здесь $s \in (0, h(f, A))$. В силу (П.2.13) и (П.2.14) для любого $t > 0$ поточечно

$$T_t f = \lim_{n \to \infty} \left(I + \frac{t}{n} A \right)^n f,$$

и при всех n, настолько больших, что $t/n < h(f, A)$, функция под знаком предела не убывает. Поэтому $T_t f$ не убывает. □

Если фазовое пространство $S \subset \mathbb{R}$ (с обычным порядком на прямой), например, $S = \{1, \ldots, n\}$, то любое распределение на S ассоциировано по теореме 1.8(а). Следовательно, для любой марковской цепи $\mathbf{X}$ со значениями в S и всех $t \geqslant 0$ мера $Law(X_t)$ ассоциирована. Нетривиальный вопрос: что можно сказать о всех конечномерных распределениях? Здесь важно снова обратиться к понятию монотонности. Рассмотрим простой пример, когда монотонность марковского процесса имеет место.

Пример 2.23. Пусть $\mathbf{X}$ — цепь Маркова со значениями в пространстве $S = \{s_1, s_2\} \subset \mathbb{R}$, где $s_1 \leqslant s_2$, имеющая инфинитезимальную матрицу A. Тогда, очевидно, выполнено (2.24), и эта цепь Маркова монотонна.

Определение 2.24. Случайный процесс[1] $\{X(t), t \geqslant 0\}$ называется *ассоциированным по времени*, если он представляет собой ассоциированное семейство случайных величин.

[1] Необязательно с конечным числом возможных значений.