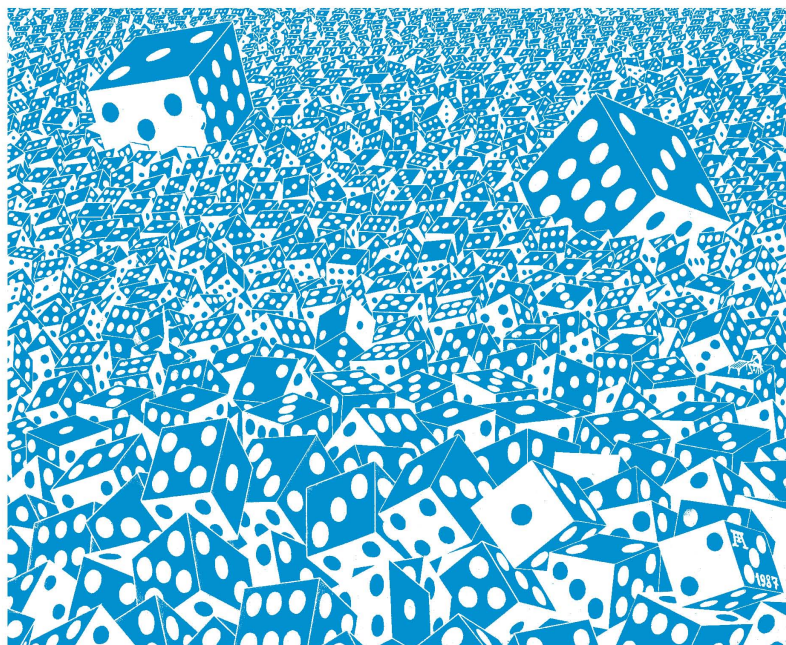


И. СТОЯНОВ

КОНТРИМЕРЫ
в теории вероятностей



Й. СТОЯНОВ

КОНТРПРИМЕРЫ В ТЕОРИИ ВЕРОЯТНОСТЕЙ

Электронное издание

Москва
Издательство МЦНМО
2014

УДК 519.21(075.8)

ББК 22.171

С82

Стоянов Й.

Контрпримеры в теории вероятностей

Электронное издание

М.: МЦНМО, 2014

294 с.

ISBN 978-5-4439-2040-5

Книга содержит около 300 разнообразных контрпримеров и примеров, относящихся к основным разделам теории вероятностей и случайных процессов. Во второе издание добавлен новый материал, расширен список литературы. Книгу можно активно использовать при изучении теории вероятностей и случайных процессов.

Предназначена для студентов, аспирантов и научных сотрудников физико-математических специальностей.

Подготовлено на основе книги: *Й. Стоянов. Контрпримеры в теории вероятностей.* — М.: МЦНМО, 2012.

Издательство Московского центра
непрерывного математического образования

119002, Москва, Большой Власьевский пер., 11. Тел. (499) 241-74-83

<http://www.mccme.ru>

ISBN 978-5-4439-2040-5

© Стоянов Й., 2012.

© МЦНМО, 2014.

ОГЛАВЛЕНИЕ

Предисловие к русскому изданию	13
Предисловие автора	15
Основные обозначения и сокращения	17
Глава 1. Случайные события и их вероятности	18
§ 1. Классы случайных событий	18
1.1. Класс событий, который образует алгебру, но не является σ -алгеброй	19
1.2. О замкнутости класса событий относительно операций объединения, пересечения и дополнения	20
1.3. Каждая алгебра событий является полуалгеброй, но не наоборот	20
1.4. В σ -алгебре подмножеств множества Ω могут не содержаться все его подмножества	20
1.5. Каждая σ -алгебра является d -системой, но обратное неверно	21
1.6. Множества, которые не являются событиями в произведении соответствующих σ -алгебр	22
1.7. Объединение σ -алгебр может не быть σ -алгеброй	22
§ 2. Вероятность	25
2.1. Вероятностная мера, которая является аддитивной, но не σ -аддитивной	26
2.2. Две вероятностные меры могут совпадать на заданном классе событий и не совпадать на σ -алгебре, порожденной этим классом	27
2.3. Теорема Колмогорова о продолжении меры в пространстве $(\mathbb{R}^\infty, \mathcal{B}^\infty)$	28
2.4. О регулярных условных вероятностях	29
§ 3. Независимость случайных событий	31
3.1. О роли вероятностной меры в свойстве независимости	32
3.2. Попарная независимость событий не влечет их независимости в совокупности	33

3.3.	Из соотношения $P(ABC) = P(A)P(B)P(C)$ не следует, что события A, B и C независимы в совокупности . . .	34
3.4.	Система из $n + 1$ зависимых событий, каждые n из которых независимы в совокупности	35
3.5.	О независимости и условной независимости случайных событий	37
3.6.	Об одном выборе условий типа независимости, из которого не следует независимость в совокупности	38
3.7.	Из независимых в совокупности событий можно формировать зависимые классы событий	39
3.8.	Семейства случайных событий с «необычными» свойствами независимости/зависимости	39
§ 4.	Разные свойства случайных событий и их вероятностей .	41
4.1.	Вероятностные пространства без нетривиальных независимых событий	41
4.2.	О лемме Бореля–Кантелли	42
4.3.	Может ли класс событий обладать свойством независимости и быть исчерпывающим?	43
Глава 2.	Распределения. Случайные величины	46
§ 5.	Функции распределения случайных величин	46
5.1.	Эквивалентные случайные величины распределены одинаково, но обратное неверно	49
5.2.	Если случайные величины X, Y и Z заданы на одном вероятностном пространстве, то из $X \stackrel{d}{=} Y$ не следует, что $XZ \stackrel{d}{=} YZ$	50
5.3.	Функция, которая является метрикой в пространстве распределений, но не является метрикой в пространстве случайных величин	50
5.4.	Об n -мерных распределениях	51
5.5.	Отсутствие атомов не влечет непрерывности многомерных функций распределения	51
5.6.	Об абсолютной непрерывности распределений случайного вектора и его компонент	53
5.7.	Многомерные функции распределения не определяются однозначно своими маргинальными распределениями .	54
5.8.	О свертках функций распределений	55
5.9.	О несохранении свойства унимодальности при свертках унимодальных распределений	56
§ 6.	Математическое ожидание	59
6.1.	О линейности математического ожидания	61

6.2.	Об интегрируемости последовательности случайных величин и ограниченности их супремума	63
6.3.	У симметричного распределения все моменты нечетного порядка равны нулю, но обратное неверно	63
6.4.	Об одном свойстве моментов случайных величин, которое не имеет аналога для случайных векторов	64
6.5.	О справедливости теоремы Фубини	65
6.6.	Семейство интегрируемых случайных величин, которое не является равномерно интегрируемым	66
6.7.	Всегда ли справедливо равенство $E\{E(X Y)\} = EX$? . .	67
6.8.	О невозможности расширить одно из свойств условного математического ожидания	68
§ 7.	Независимость случайных величин	69
7.1.	О попарной независимости случайных величин и их независимости в совокупности	69
7.2.	Множество зависимых случайных величин, каждое собственное подмножество которого обладает свойством независимости	70
7.3.	Зависимые случайные величины X и Y , такие что случайные величины X^2 и Y^2 независимы	73
7.4.	Независимость случайных величин в терминах характеристических функций	74
7.5.	Независимость случайных величин в терминах производящих функций	76
7.6.	Свертка функций распределений может быть распределением суммы зависимых случайных величин	77
7.7.	Некоррелированные, но зависимые случайные величины	77
7.8.	Из равенства $E(Y X) = EY$ почти наверное не всегда вытекает, что X и Y независимы	78
7.9.	Существует ли связь между понятиями независимости и условной независимости случайных величин?	79
§ 8.	Характеристические и производящие функции	80
8.1.	Характеристические функции, совпадающие на конечном отрезке, но не на всей числовой прямой	82
8.2.	Дискретное и абсолютно непрерывное распределения с характеристическими функциями, совпадающими на отрезке $[-1, 1]$	83
8.3.	Абсолютно непрерывное распределение может иметь характеристическую функцию, которая не интегрируема абсолютно	84

8.4.	Неинтегрируемая дискретная случайная величина с дифференцируемой характеристической функцией	85
8.5.	Неинтегрируемая абсолютно непрерывная случайная величина с дифференцируемой характеристической функцией	86
8.6.	Производящая функция может не существовать, даже если случайная величина имеет моменты любого порядка r , $r = 1, 2, \dots$	87
§ 9.	Безгранично делимые и устойчивые распределения	89
9.1.	Не обращающаяся в нуль характеристическая функция, которая не является безгранично делимой	90
9.2.	О свойстве безграничной делимости характеристических функций φ и $ \varphi $	91
9.3.	Произведение двух независимых неотрицательных и безгранично делимых случайных величин не всегда сохраняет свойство безграничной делимости	93
9.4.	О свойстве безграничной делимости в многомерном случае	94
9.5.	О безграничной делимости случайного вектора и линейных комбинаций его компонент	95
9.6.	Безгранично делимые распределения, которые не являются устойчивыми	97
9.7.	Устойчивое распределение можно разложить на два безгранично делимых и неустойчивых распределения . . .	99
§ 10.	Нормальное распределение	100
10.1.	Нормальность одномерных распределений не обеспечивает нормальности соответствующего многомерного распределения	101
10.2.	Если (X_1, X_2) имеет двумерное нормальное распределение, то X_1 , X_2 и $X_1 + X_2$ нормально распределены, но обратное неверно	102
10.3.	Свойство нормальности может не иметь места для системы случайных величин, даже если любая ее подсистема обладает этим свойством	104
10.4.	О связи между свойствами нормальности и некоррелированности	105
10.5.	Случайные величины X , Y , $X + Y$ и $X - Y$ могут быть нормально распределенными, причем X и Y некоррелированы, а распределение вектора (X, Y) может не быть нормальным	107

10.6.	Условие, характеризующее нормальное распределение через нормальность линейных комбинаций, не может быть ослаблено	108
10.7.	Распределение, отличающееся от нормального, может иметь условные нормальные распределения	109
§ 11.	Проблема моментов	111
11.1.	Проблема моментов для некоторых функций от нормально распределенной случайной величины	113
11.2.	Логарифмически нормальное распределение и проблема моментов	114
11.3.	Еще один класс абсолютно непрерывных распределений, которые не определяются однозначно своими моментами	117
11.4.	Проблема моментов для степенных преобразований экспоненциального распределения	118
11.5.	Различные дискретные распределения с совпадающими моментами любого порядка	119
11.6.	Как связаны два различных достаточных условия для единственности решения проблемы моментов?	123
11.7.	Условие Карлемана достаточно, но не необходимо для единственности решения проблемы моментов	124
11.8.	Условие Крейна достаточно, но не необходимо для неединственности решения проблемы моментов	126
§ 12.	Характеризационные свойства некоторых вероятностных распределений	126
12.1.	Как распределены случайные величины, если их сумма имеет биномиальное распределение?	127
12.2.	Свойство геометрического распределения, которое не является его характеризационным свойством	128
12.3.	О применимости теоремы Райкова	129
12.4.	О применимости теоремы Крамера	131
12.5.	Об одном свойстве, которое не является характеризационным для нормального распределения	132
12.6.	Еще одно интересное свойство, которое не является характеризационным для нормального распределения	133
12.7.	Об одном свойстве, не являющемся характеризационным для распределения Коши	134
12.8.	Об одном свойстве, которое не является характеризационным для гамма-распределения	135
§ 13.	Разные свойства случайных величин	136

13.1.	О свойстве независимости при суммировании пуассоновских случайных величин	136
13.2.	О симметричности суммы и разности двух случайных величин	137
13.3.	Соотношение между классами распределений NBU и IFR	139
13.4.	Перестановочные и «хвостовые» события, связанные с последовательностями случайных величин	140
13.5.	Перестановочная последовательность случайных величин, не обладающих свойством независимости . .	142
13.6.	О применимости закона «0 или 1» Колмогорова	143
13.7.	О применимости закона «0 или 1» Хьюитта–Сэвиджа .	144
13.8.	Всегда ли применима теорема де Финетти?	145
Глава 3.	Предельные теоремы	146
§ 14.	Разные виды сходимости последовательностей случайных величин	146
14.1.	О сходимости функций случайных величин	148
14.2.	Из сходимости по распределению не вытекает сходимость по вероятности	149
14.3.	Последовательности случайных величин, сходящиеся по вероятности, но не с вероятностью 1	149
14.4.	Из сходимости в среднем порядка r вытекает сходимость по вероятности, но обратное неверно	151
14.5.	Из сходимости в среднем порядка r не всегда вытекает сходимость почти наверное	152
14.6.	Из сходимости почти наверное не всегда следует сходимость в среднем порядка r	152
14.7.	Сходимость плотностей обеспечивает слабую сходимость функций распределений, но обратное неверно	153
14.8.	Из сходимости $X_n \xrightarrow{d} X$ и $Y_n \xrightarrow{d} Y$ не всегда вытекает, что $X_n + Y_n \xrightarrow{d} X + Y$	154
14.9.	Не существует метрики, соответствующей сходимости почти наверное	155
14.10.	О сходимости вполне и сходимости почти наверное последовательностей случайных величин	156
14.11.	О сходимости средних по Чезаро для последовательности случайных величин	156
§ 15.	Законы больших чисел	158
15.1.	Условие Маркова достаточно, но не необходимо для справедливости закона больших чисел	159

15.2.	Условие Колмогорова достаточно, но не необходимо для справедливости усиленного закона больших чисел для неодинаково распределенных случайных величин . . .	160
15.3.	Еще один пример, когда условие Колмогорова не выполнено, но усиленный закон больших чисел имеет место	161
15.4.	О роли условия Колмогорова $\sum_{n=1}^{\infty} (\sigma_n^2/n^2) < \infty$	161
15.5.	Арифметические средние для случайной последовательности могут сходиться по вероятности, даже если теорема Хинчина неприменима	163
15.6.	Последовательность случайных величин, удовлетворяющая слабому, но не усиленному закону больших чисел	164
15.7.	Закон больших чисел может не выполняться, если сходимость с вероятностью 1 заменить сходимостью вполне	166
15.8.	О справедливости обобщенного закона больших чисел	167
§ 16.	Слабая сходимость вероятностных мер и распределений .	168
16.1.	Определяющие классы и классы, определяющие сходимость	169
16.2.	Из сходимости по распределению не вытекает слабая сходимость соответствующих вероятностных мер для всех борелевских множеств	172
16.3.	О слабой сходимости и сходимости моментов	174
16.4.	Два случая, когда теорема непрерывности неприменима	175
16.5.	Когда теорема непрерывности применима для плотностей?	176
16.6.	О слабой сходимости и сходимости по вариации	178
16.7.	Слабая сходимость и метрика Леви	180
§ 17.	Центральная предельная теорема	181
17.1.	Последовательности случайных величин, для которых центральная предельная теорема не справедлива . . .	183
17.2.	Какова роль различных условий для выполнения центральной предельной теоремы?	184
17.3.	Две «эквивалентные» последовательности случайных величин, такие что одна из них удовлетворяет, а другая не удовлетворяет центральной предельной теореме . .	185
17.4.	Условие Ляпунова и центральная предельная теорема	186
17.5.	О центральной предельной теореме и сходимости дисперсии нормированной суммы $S_n/\sqrt{DS_n}$	189

17.6.	Центральная предельная теорема не всегда применима при суммировании случайного числа случайных величин	190
17.7.	Какова связь между двумя центральными предельными теоремами, интегральной и локальной?	191
§ 18.	Разные предельные теоремы	193
18.1.	В теореме Колмогорова о «трех рядах» условие независимости существенно	194
18.2.	Об одном следствии теоремы Колмогорова о «трех рядах»	195
18.3.	О радиусе сходимости случайных степенных рядов	196
18.4.	Всегда ли можно менять порядок операций взятия математического ожидания, бесконечного суммирования и предела	197
18.5.	L^r -сходимость последовательности случайных величин не вытекает из сходимости условных математических ожиданий	199
18.6.	О законе повторного логарифма Човера	200
18.7.	О рекордах и максимумах последовательности независимых случайных величин	201

Глава 4. Случайные процессы **203**

§ 19.	Основные понятия и свойства случайных процессов	203
19.1.	О роли семейства конечномерных распределений при построении случайного процесса с заданными свойствами	205
19.2.	У случайного процесса могут быть модификации с различными свойствами	206
19.3.	О свойстве сепарабельности случайных процессов	207
19.4.	Измеримые и прогрессивно измеримые процессы	208
19.5.	О стохастической непрерывности случайных процессов	210
19.6.	Критерий Колмогорова для непрерывности с вероятностью 1 случайного процесса	211
19.7.	Существует ли связь между непрерывностью случайного процесса и непрерывностью порожденной им фильтрации?	212
§ 20.	Марковские процессы	214
20.1.	Уравнение Колмогорова–Чепмена может иметь место, даже если процесс немарковский	216
20.2.	Об одном следствии марковского свойства	217

20.3.	Связь между различными понятиями эргодичности для марковских цепей	218
20.4.	Эргодичность марковской цепи не всегда обеспечивает сходимость математического ожидания функции от этой цепи	220
20.5.	Две неэквивалентные марковские цепи с непрерывным временем могут быть частично эквивалентными	222
20.6.	Марковские процессы, феллеровские процессы, сильно феллеровские процессы и связи между ними .	223
20.7.	Марковские, но не строго марковские процессы	224
§ 21.	Стационарные процессы и некоторые смежные вопросы .	227
21.1.	О свойстве стационарности случайных процессов . . .	228
21.2.	О стационарности заданного порядка m	229
21.3.	Условие сильного перемешивания не всегда сохраняется при преобразовании стационарного в узком смысле процесса	230
21.4.	Стационарный в узком смысле процесс может быть регулярным и не обладать свойством полной регулярности	232
21.5.	Законы больших чисел для стационарных в широком смысле процессов	233
21.6.	О свойствах эргодичности и перемешивания для сохраняющих меру преобразований	234
21.7.	Стационарные последовательности, для которых центральная предельная теорема не имеет места	236
§ 22.	Мартингалы с дискретным временем	239
22.1.	L^1 -ограниченные мартингалы, не являющиеся L^1 -доминированными	240
22.2.	Мартингалы, для которых теорема Дуба об остановке не имеет места	241
22.3.	Каждый квазимартингал является амартом, но обратное не верно	242
22.4.	Амарты, мартингалы в пределе, эвентуальные мартингалы и связи между ними	243
22.5.	Квазимартингалы, амарты, прогрессивные мартингалы и связи между ними	244
22.6.	О сходимости субмартингалов почти наверное и в пространстве $\mathcal{L}^1$	245
22.7.	Мартингалы могут сходиться по вероятности, но не сходиться почти наверное	246

22.8.	Мартингалы могут расходиться на множестве с заранее заданной вероятностью	248
§ 23.	Мартингалы с непрерывным временем	249
23.1.	Мартингалы, не являющиеся локально квадратично-интегрируемыми	250
23.2.	L^2 -ограниченный локальный мартингал, который не является мартингалом	251
23.3.	Экспоненциальные мартингалы: неухудшаемость условия Новикова и два полезных следствия	253
23.4.	Квадратично-интегрируемый мартингал с неслучайной характеристикой не обязан быть процессом с независимыми приращениями	255
23.5.	О возможности представить мартингал в виде стохастического интеграла относительно другого мартингала	256
23.6.	Гауссовские процессы, которые не являются семимартингалами	257
§ 24.	Разные свойства случайных процессов	259
24.1.	О перестановочности операций пересечения и супремума для последовательности σ -алгебр, связанных со случайными процессами	259
24.2.	О тождествах Вальда для винеровского процесса	261
24.3.	О сходимости квадратической вариации винеровского процесса	262
24.4.	Можно ли характеризовать многомерный гауссовский процесс по его проекциям?	263
24.5.	Представление Крамера, кратность и спектральный тип случайного процесса	265
24.6.	Фрактальный винеровский процесс и марковское свойство	267
Дополнительные замечания		268
Литература		273

ПРЕДИСЛОВИЕ К РУССКОМУ ИЗДАНИЮ

Предметом книги являются *контрпримеры* в традиционном понимании этого термина и *примеры*, раскрывающие разнообразные, стандартные и нестандартные, зачастую причудливые, свойства случайных величин и случайных процессов. Эти примеры и контрпримеры автор собирал на протяжении многих лет. Надо сказать, что построение примеров и контрпримеров, являясь далеко не тривиальной работой, существенно помогает очертить границы применимости тех или иных результатов теории, используемых понятий, конструкций, методов и т. д. Отметим, что имеется еще несколько книг (например, [100, 417, 483]) с небольшим сходством по содержанию с настоящей книгой, уникальной по богатству материала и занимающей достойное место среди книг в области теории вероятностей.

Предлагаемая книга требует от читателя знания основ теории вероятностей, а более конкретно — знакомства с основными определениями и свойствами тех вероятностных объектов, относительно которых строится тот или иной пример. Хотя в книге и даются основные определения, они не являются исчерпывающими и не предназначены для первого знакомства. Назначение книги предполагает определенную краткость изложения. Для более детального ознакомления с основами теории вероятностей можно порекомендовать книги:

Колмогоров А. Н. Основные понятия теории вероятностей. (2-е изд.) — М.: Наука, Физматлит, 1974;

Гнеденко Б. В. Курс теории вероятностей. (7-е изд., испр.). — М.: УРСС, 2001;

Ширяев А. Н. Вероятность. В 2-х книгах. (5-е изд., стер.). — М.: МЦНМО, 2011;

и выходящую почти одновременно книгу

Ширяев А. Н., Эрлих И. Г., Ясков П. А. Вероятность в теоремах и задачах (с доказательствами и решениями). — М.: МЦНМО, 2012,

в которых приведены многие доказательства и решения задач, относящихся к теории вероятностей.

Контрпримеры и примеры в этой книге Й. Стоянова — очень разнообразны. Некоторые из них легкие, в других требуется подумать и кое-что подсчитать, а есть и очень серьезные утверждения. Однако следует помнить, что читатели бывают разной степени подготовленности. И то, что тривиально для специалиста, может быть неожиданно для студента.

Данная книга была подготовлена к публикации много лет назад, но напечатана была только в 1999 г. За это время ее английский вариант был опубликован в издательстве J. Wiley & Sons, Inc. Там же вышло ее второе издание, что подтверждает высокие качества книги. Следует отметить, что Й. Стоянов сам подготовил русский текст книги, и мягкий болгарский акцент придает ее языку определенный шарм. Мы уверены, что данное издание встретит благожелательный интерес у читателей.

*Академик РАН
А. Н. Ширяев*

ПРЕДИСЛОВИЕ АВТОРА

Общие замечания. Начнем с иллюстрации. Предположим, что высказано следующее утверждение:

«В русском языке нет слов, содержащих
последовательность из пяти согласных». (*)

На вопрос «верно ли это?» можно легкомысленно ответить «да». Однако правильным ответом будет «нет». Это означает, что если существует хотя бы одно русское слово, содержащее пять согласных подряд, оно будет служить контрпримером к утверждению (*). Забавно, что в данном случае контрпримером является само слово «контрпример». Любители языка могут поискать и другие контрпримеры.

В настоящей книге термин «контрпример» используется в обычном смысле, как это принято в математике. Книга является собранием около 300 разнообразных контрпримеров и примеров, относящихся к основным разделам теории вероятностей и случайных процессов. В результате моей многолетней работы была издана книга:

Stoyanov J. Counterexamples in Probability. — Chichester, New York: John Wiley and Sons, 1987 (2 ed. 1997).

В 1990 г. мною был подготовлен русский вариант книги. Я высоко ценю труд С. Е. Кузнецова, который в 1991 г., как рецензент русского текста книги, сделал ряд полезных предложений. Позже, благодаря усилиям рано ушедшего из жизни В. В. Калашникова (1942–2001) и В. В. Сенатова, книга была подготовлена к печати в 1995 г. и вышла в свет в издательстве «Факториал» только в 1999 г.

В настоящее издание внесены некоторые дополнения и улучшения. Устранены замеченные опечатки и неточности и учтены предложения читателей. Теперь в книгу включен новый и свежий материал, отражен опыт, накопленный за последние годы. Обновлен список литературы.

Предназначение и содержание книги. Использование контрпримеров помогает понять лучше и глубже все основные понятия и роль условий, при которых доказываются вероятностные утверждения. Выражаясь словами Штирлица, можно сказать, что каждый контрпример и пример в этой книге — это хороший «повод для размышления».

Книга послужит дополнительным пособием для всех, кто изучает или преподаёт теорию вероятностей и случайные процессы. Профессоры в этой области тоже могут извлечь из книги некоторую пользу.

Материал включает основные темы, рассматриваемые в курсах по теории вероятностей и случайным процессам. Среди имеющихся на русском языке изданий, настоящая книга наиболее близка по стилю и содержанию к следующим фундаментальным книгам:

Ш и р я е в А. Н. Вероятность. В 2-х книгах. (5-е изд., стер.). — М.: МЦНМО, 2011;

Ш и р я е в А. Н. Задачи по теории вероятностей. (2-е изд.). — М.: МЦНМО, 2011.

Эти книги и настоящая книга хорошо дополняют друг друга.

История книги. Я начал коллекционировать контрпримеры в 1970 г., когда был студентом МГУ им. М. В. Ломоносова. Моя дальнейшая работа сотрудником Института математики Болгарской АН и преподавателем Софийского университета была благоприятна для этого моего хобби. Со временем накопил значительный материал по объёму и содержанию. Возрастал интерес к идее активно использовать контрпримеры при изучении теории вероятностей. Об этом подходе появился ряд публикаций. Со специальными лекциями на эту тему я выступал во многих европейских и американских университетах. Интернациональный характер настоящей книги очевиден.

Благодарности. При работе над книгой у меня были полезные и плодотворные обсуждения со многими коллегами и друзьями.

В первую очередь я пользуюсь случаем выразить признательность моим учителям Б. В. Гнеденко, Ю. В. Прохорову и А. Н. Ширяеву. Специальной благодарности заслуживают коллеги и друзья: Н. В. Крылов, Р. Ш. Липцер, А. А. Новиков, Ю. М. Кабанов, В. В. Сенатов, С. Е. Кузнецов, А. М. Зубков, А. В. Булинский, Ф. Стойтел, А. Лясоф и С. Гайдов. В списке литературы отмечены имена всех, кто тем или иным образом проявил внимание и интерес к моей работе.

Я благодарен также А. Т. Фоменко, любезно предоставившему оригинальные рисунки, которые очень удачно иллюстрируют, что такое «контрпример». Вот, иногда говорят, что «все игральные кости стандартные». Взгляд на обложки показывает... «да не все»...

Как и прежде, я с вниманием встречу замечания читателей.

Наконец, выражаю благодарность сотрудникам МЦНМО за замечательную работу по подготовке и изданию книги.

*Йордан Стоянов (Данчо)
Весна 2012 г., Newcastle (U. K.)*

ОСНОВНЫЕ ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

$(\Omega, \mathcal{F}, P)$ — вероятностное пространство

EX, DX — математическое ожидание и дисперсия величины X

$\mathbb{R}^n, n \geq 1$, — n -мерное евклидово пространство

$\mathcal{B}^n$ — борелевская σ -алгебра в $\mathbb{R}^n$

$\mathbb{N}$ — множество натуральных чисел: $\mathbb{N} = \{1, 2, \dots\}$

$\mathbb{N}_0$ — множество неотрицательных целых чисел: $\mathbb{N}_0 = \{0, 1, 2, \dots\}$

$\tilde{\mathbb{N}}$ — множество целых чисел: $\tilde{\mathbb{N}} = \{\dots, -2, -1, 0, 1, 2, \dots\}$

$I_A, I(A)$ — индикаторная функция множества A

$\mathcal{N}(a, \sigma^2)$ — нормальное распределение с параметрами a и σ^2

$X \sim F$ — случайная величина X имеет распределение F

Φ — стандартная нормальная функция распределения

$F_1 * F_2$ — свертка функций распределения F_1 и F_2

$x \vee y$ — максимум величин x и y : $\max(x, y)$

$x \wedge y$ — минимум величин x и y : $\min(x, y)$

$\implies, \iff$ — логические импликации

$\downarrow$ — монотонное убывание

п. н. — почти наверное

б. ч. — бесконечно часто

$\xrightarrow{d}$ — сходимость по распределению

$\xrightarrow{P}$ — сходимость по вероятности

$\xrightarrow{\text{п.н.}}$ — сходимость почти наверное (с вероятностью 1)

$\xrightarrow{L^r}$ — сходимость в пространстве $\mathcal{L}^r$

$\xrightarrow{w}$ — слабая сходимость

$\xrightarrow{c}$ — сходимость вполне

$\xrightarrow{v}$ — сходимость по вариации

ЗБЧ — закон больших чисел

УЗБЧ — усиленный закон больших чисел

ЦПТ — центральная предельная теорема

$\binom{n}{k}$ — число сочетаний из n по k

ГЛАВА 1

СЛУЧАЙНЫЕ СОБЫТИЯ И ИХ ВЕРОЯТНОСТИ

§ 1. Классы случайных событий

Пусть задано произвольное непустое множество Ω . Его элементы, обозначаемые ω , будут интерпретироваться как исходы некоторого эксперимента. Как обычно, через $A \cup B$ и $A \cap B$ (а также AB) будем обозначать соответственно объединение и пересечение любых двух подмножеств $A \subseteq \Omega$ и $B \subseteq \Omega$, а через $\bar{A}$ — дополнение множества $A \subseteq \Omega$. В частности, $\bar{\Omega} = \emptyset$, где $\emptyset$ — пустое множество. Если множества A и B не пересекаются, то их объединение будем обозначать $A + B$; соответственно, объединение непересекающихся множеств A_i , $i \geq 1$, обозначается $\sum_i A_i$.

Класс $\mathcal{A}$ подмножеств Ω называется *алгеброй*, если он содержит Ω и замкнут относительно дополнения и конечных объединений, т. е. если:

- а) $\Omega \in \mathcal{A}$;
- б) $A \in \mathcal{A} \implies \bar{A} \in \mathcal{A}$;
- в) $A_1, A_2 \in \mathcal{A} \implies A_1 \cup A_2 \in \mathcal{A}$.

Учитывая законы де Моргана ($\overline{A_1 A_2} = \bar{A}_1 \cup \bar{A}_2$ и $\overline{A_1 \cup A_2} = \bar{A}_1 \bar{A}_2$), легко видеть, что условие в) можно заменить условием

- в') $A_1, A_2 \in \mathcal{A} \implies A_1 A_2 \in \mathcal{A}$.

Это означает, что класс $\mathcal{A}$ замкнут относительно конечных пересечений.

Класс $\mathcal{F}$ подмножеств Ω называется *σ -алгеброй*, если $\mathcal{F}$ — алгебра, которая замкнута относительно счетных объединений, т. е. если

- г) $A_1, A_2, \dots \in \mathcal{F} \implies \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$.

Снова, как и выше, условие г) может быть заменено эквивалентным ему условием

- г') $A_1, A_2, \dots \in \mathcal{F} \implies \bigcap_{n=1}^{\infty} A_n \in \mathcal{F}$.

Таким образом, σ -алгебра $\mathcal{F}$ является замкнутой относительно счетных пересечений.

Напомним, что элементы любой алгебры или σ -алгебры называются *случайными событиями* (или просто *событиями*).

Другие классы событий, такие как полуалгебры, d -системы и произведение σ -алгебр, будут введены и изучены в рассмотренных далее конкретных примерах.

По поводу основных теоретико-вероятностных объектов, идей и результатов читатель может обратиться к любому из хорошо известных учебных пособий, например [9, 26, 60, 75, 81, 92, 106, 116, 117, 118, 145, 179, 330].

1.1. Класс событий, который образует алгебру, но не является σ -алгеброй

(а) Пусть $\Omega = [0, \infty)$, а класс $\mathcal{A}_1$ содержит все интервалы вида $[a, b)$ или $[a, \infty)$, где $0 \leq a < b < \infty$. Далее, пусть класс $\mathcal{A}_2$ состоит из пустого множества $\emptyset$ и всех конечных сумм непересекающихся интервалов из $\mathcal{A}_1$. Покажем, что класс $\mathcal{A}_1$ не является алгеброй, а $\mathcal{A}_2$ — алгебра, которая не является σ -алгеброй.

Действительно, возьмем произвольные действительные числа a и b , $0 < a < b < \infty$. Тогда $A = [a, b) \in \mathcal{A}_1$, но поскольку $\bar{A} = [0, a) \cup [b, \infty) \notin \mathcal{A}_1$, то $\mathcal{A}_1$ не может быть алгеброй.

Далее, легко проверить, что: 1) объединение конечного числа элементов класса $\mathcal{A}_2$ принадлежит $\mathcal{A}_2$; 2) дополнение любого элемента класса $\mathcal{A}_2$ снова принадлежит $\mathcal{A}_2$. Таким образом, мы видим, что $\mathcal{A}_2$ образует алгебру. Однако $\mathcal{A}_2$ не является σ -алгеброй, потому что, например, множество $A_n = \left[0, \frac{1}{n}\right)$ при любом $n \in \mathbb{N}$ принадлежит $\mathcal{A}_2$, в то время как пересечение $\bigcap_{n=1}^{\infty} A_n = \{0\}$ не принадлежит $\mathcal{A}_2$.

(б) Пусть $\mathcal{A}$ — класс подмножеств множества $\Omega = \mathbb{R}^1$, состоящий из конечных сумм непересекающихся интервалов вида $(-\infty, a]$, $(b, c]$ и (d, ∞) , где a, b, c, d — любые числа из $\mathbb{R}^1$. Тогда класс $\mathcal{A}$ образует алгебру. Однако пересечение $\bigcap_{n=1}^{\infty} \left(b - \frac{1}{n}, c\right]$, равное $[b, c]$, не входит в $\mathcal{A}$. Следовательно, $\mathcal{A}$ не является σ -алгеброй.

(в) Возьмем произвольное множество Ω , содержащее бесконечное число элементов. Обозначим через $\mathcal{A}$ совокупность всех таких подмножеств $A \subseteq \Omega$, что конечно либо дополнение $\bar{A}$ подмножества A , либо оно само. Тогда легко видеть, что $\mathcal{A}$ — алгебра, которая σ -алгеброй не является.

1.2. О замкнутости класса событий относительно операций объединения, пересечения и дополнения

Пусть $\Omega = \mathbb{R}^1$, а класс $\mathcal{A}$ состоит из всех интервалов вида (x, ∞) , $x \in \Omega$. Используя обозначения $u = x \wedge y = \min(x, y)$ и $v = x \vee y = \max(x, y)$, мы видим, что

$$(x, \infty) \cup (y, \infty) = (u, \infty) \in \mathcal{A}, \quad (x, \infty) \cap (y, \infty) = (v, \infty) \in \mathcal{A},$$

т. е. класс $\mathcal{A}$ замкнут относительно (конечных) объединений и пересечений. Однако, класс $\mathcal{A}$ не замкнут относительно взятия дополнения. Например, если $A = (x, \infty)$, то $A \in \mathcal{A}$, но $\bar{A} = (-\infty, x] \notin \mathcal{A}$.

1.3. Каждая алгебра событий является полуалгеброй, но не наоборот

Пусть Ω — произвольное множество. Напомним, что класс $\mathcal{S}$ подмножеств Ω называется *полуалгеброй*, если $\Omega \in \mathcal{S}$, $\emptyset \in \mathcal{S}$, класс $\mathcal{S}$ замкнут относительно конечных пересечений и дополнения любого множества в $\mathcal{S}$ является конечной суммой непересекающихся множеств из $\mathcal{S}$.

Легко видеть, что любая алгебра подмножеств множества Ω является и полуалгеброй. Простые примеры показывают, что обратное неверно.

(а) Пусть $\Omega = [-\infty, \infty]$, а класс $\mathcal{S}_1$ содержит Ω , $\{\infty\}$ и все интервалы вида $[a, b)$, где $-\infty \leq a \leq b \leq \infty$. Тогда $\emptyset \in \mathcal{S}_1$, $[a_1, b_1) \cap [a_2, b_2) = [a_1 \vee a_2, b_1 \wedge b_2) \in \mathcal{S}_1$ и $\overline{[a, b)} = [-\infty, a) \cup [b, \infty)$. Следовательно, $\mathcal{S}_1$ — полуалгебра. Но, очевидно, что $\mathcal{S}_1$ — не алгебра.

(б) Возьмем $\Omega = \mathbb{R}^1$ и обозначим через $\mathcal{S}_2$ класс всех подмножеств вида AB , где A — замкнутое, а B — открытое множество в Ω . Тогда $\mathcal{S}_2$ образует полуалгебру, но $\mathcal{S}_2$ не является алгеброй.

1.4. В σ -алгебре подмножеств множества Ω могут не содержаться все его подмножества

В любой теоретико-множественной задаче мы рассматриваем множество Ω , называемое *пространством элементарных событий*, и некоторую σ -алгебру $\mathcal{F}$ случайных событий, т. е. подмножеств Ω . Поэтому для нас важно знать, содержит ли $\mathcal{F}$ все подмножества множества Ω . Приводимый далее пример содержит ответ на этот вопрос.

Напомним сначала, что множество $A \in \Omega$ называется *c-конечным*, если его дополнение $\bar{A}$ конечно.

Пусть класс $\mathcal{F}_1$ состоит из всех конечных и всех *c-конечных* подмножеств множества Ω . Тогда легко проверить, что $\mathcal{F}_1$ является алгеброй.

При этом $\mathcal{F}_1$ будет σ -алгеброй тогда и только тогда, когда Ω — конечное множество.

Далее, множество $A \in \Omega$ называется *c-счетным*, если его дополнение $\bar{A}$ содержит счетное число элементов.

Через $\mathcal{F}_2$ обозначим совокупность всех счетных и всех *c*-счетных подмножеств множества Ω . Нетрудно видеть, что $\mathcal{F}_2$ образует σ -алгебру. Предположим теперь, что множество Ω несчетное. Тогда в Ω имеется такое множество A , что одновременно A и его дополнение $\bar{A}$ — несчетные множества. Это показывает, в частности, что не все подмножества множества Ω входят в σ -алгебру $\mathcal{F}_2$.

1.5. Каждая σ -алгебра является d -системой, но обратное неверно

Напомним, что класс $\mathcal{D}$ подмножеств Ω называется *d-системой*, если:

- 1) $\Omega \in \mathcal{D}$;
- 2) $A, B \in \mathcal{D}$ и $A \subseteq B \implies B \setminus A \in \mathcal{D}$;
- 3) $A_n \in \mathcal{D}$, $n = 1, 2, \dots$, и $A_1 \subseteq A_2 \subseteq \dots \implies \bigcup_{n=1}^{\infty} A_n \in \mathcal{D}$.

Тривиально проверить, что каждая σ -алгебра является также и d -системой. Естественно задаться вопросом о справедливости обратного утверждения. В общем случае ответ отрицательный. Проиллюстрируем это на одном конкретном примере.

Предположим, что Ω — конечное множество, число элементов которого четное (например, $\Omega = \{\omega_1, \omega_2, \dots, \omega_{2n}\}$ для некоторого $n \in \mathbb{N}$). Обозначим через $\mathcal{D}_e$ совокупность всех подмножеств D множества Ω , содержащих четное число элементов. Очевидно, $\Omega \in \mathcal{D}_e$. Пусть $A \in \mathcal{D}_e$ и $B \in \mathcal{D}_e$, причем $A \subseteq B$. Тогда разность $B \setminus A$ тоже содержит четное число элементов множества Ω и, значит, $B \setminus A \in \mathcal{D}_e$.

Наконец, если $A_1, A_2, \dots$ — последовательность подмножеств множества Ω , принадлежащих $\mathcal{D}_e$, т. е. каждое A_j имеет четное число элементов, и $A_1 \subseteq A_2 \subseteq \dots$, то легко заметить, что объединение $A_1 \cup A_2 \cup \dots$ также содержит четное число элементов, т. е. $\bigcup_{n=1}^{\infty} A_n \in \mathcal{D}_e$. Следовательно, условия 1)–3) выполнены и класс $\mathcal{D}_e$ образует d -систему. Однако, положив, например, $A_1 = \{\omega_1, \omega_2\}$ и $B = \{\omega_2, \omega_3\}$, мы видим, что $A \in \mathcal{D}_e$, $B \in \mathcal{D}_e$, но пересечение $AB = \omega_2$ не принадлежит $\mathcal{D}_e$. Это означает, что класс $\mathcal{D}_e$ не является даже алгеброй, а тем более и σ -алгеброй.

Заметим, однако, что если d -система $\mathcal{D}$ подмножеств Ω замкнута относительно конечных пересечений, то тогда $\mathcal{D}$ образует σ -алгебру.

1.6. Множества, которые не являются событиями в произведении соответствующих σ -алгебр

Пусть заданы два произвольных множества, Ω_1 и Ω_2 . Их *произведение*, обозначаемое $\Omega_1 \times \Omega_2$, задается так: $\Omega_1 \times \Omega_2 = \{(\omega_1, \omega_2): \omega_1 \in \Omega_1, \omega_2 \in \Omega_2\}$. Для любого множества $A \subseteq \Omega_1 \times \Omega_2$ через A_{ω_1} обозначим сечение A в точке ω_1 , т.е. $A_{\omega_1} = \{\omega_2 \in \Omega_2: (\omega_1, \omega_2) \in A\}$. Аналогично, $A_{\omega_2} = \{\omega_1 \in \Omega_1: (\omega_1, \omega_2) \in A\}$. Прямоугольник в $\Omega_1 \times \Omega_2$ — это подмножество вида

$$A_1 \times A_2 = \{(\omega_1, \omega_2): \omega_1 \in A_1, \omega_2 \in A_2\}, \quad A_1 \subseteq \Omega_1, A_2 \subseteq \Omega_2.$$

Пусть $\mathcal{F}_1$ и $\mathcal{F}_2$ — σ -алгебры подмножеств множеств Ω_1 и Ω_2 соответственно. Если $A_1 \in \mathcal{F}_1$ и $A_2 \in \mathcal{F}_2$, то множество $A_1 \times A_2$ называется измеримым прямоугольником (относительно $\mathcal{F}_1$ и $\mathcal{F}_2$). Нетрудно видеть, что измеримые прямоугольники образуют полуалгебру подмножеств в произведении $\Omega_1 \times \Omega_2$. Далее, алгебра, порожденная измеримыми прямоугольниками, состоит из всех конечных сумм непересекающихся измеримых прямоугольников. Наконец, σ -алгебра, порожденная этой алгеброй, обозначается $\mathcal{F}_1 \times \mathcal{F}_2$ и называется *произведением σ -алгебр $\mathcal{F}_1$ и $\mathcal{F}_2$* .

Хорошо известен следующий результат [21, 81, 326]: для любого измеримого множества A в $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2)$ и любых фиксированных элементов $\omega_1 \in \Omega_1$ и $\omega_2 \in \Omega_2$ сечения A_{ω_1} и A_{ω_2} являются измеримыми множествами соответственно в $(\Omega_2, \mathcal{F}_2)$ и $(\Omega_1, \mathcal{F}_1)$.

Теперь естественно поставить вопрос: является ли A измеримым множеством в $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2)$, если известно, что его сечения A_{ω_1} , $\omega_1 \in \Omega_1$, и A_{ω_2} , $\omega_2 \in \Omega_2$ — измеримые множества соответственно в $(\Omega_2, \mathcal{F}_2)$ и $(\Omega_1, \mathcal{F}_1)$? Нетрудно видеть, что ответ отрицателен.

Действительно, предположим, что Ω — несчетное множество, а $\mathcal{F}$ — наименьшая σ -алгебра подмножеств множества Ω , содержащая все одноэлементные подмножества Ω . Тогда множество $D = \{(\omega, \omega): \omega \in \Omega\}$, т.е. диагональ произведения $\Omega \times \Omega$, в силу несчетности Ω , не принадлежит произведению $\mathcal{F} \times \mathcal{F}$. В то же время все сечения D принадлежат $\mathcal{F}$. Другими словами, для каждого $\omega \in \Omega$ сечение D_ω принадлежит $\mathcal{F}$, т.е. D_ω — событие, но $D \notin \mathcal{F} \times \mathcal{F}$, и, значит, D не является событием.

1.7. Объединение σ -алгебр может не быть σ -алгеброй

Предположим, что $\mathcal{F}_1, \mathcal{F}_2, \dots$ — последовательность σ -алгебр подмножеств некоторого множества Ω . Тогда их пересечение $\bigcap_{n=1}^{\infty} \mathcal{F}_n$ всегда

является σ -алгеброй. Вопрос: является ли σ -алгеброй объединение $\bigcup_{n=1}^{\infty} \mathcal{F}_n$ заданных σ -алгебр? Покажем, что в общем случае это не так.

(а) Рассмотрим множество $\Omega = \{\omega_1, \omega_2, \omega_3\}$ и следующие два класса его подмножеств: $\mathcal{F}_1 = \{\{\omega_1\}, \{\omega_2, \omega_3\}, \emptyset, \Omega\}$ и $\mathcal{F}_2 = \{\{\omega_2\}, \{\omega_1, \omega_3\}, \emptyset, \Omega\}$. Каждый из этих классов является алгеброй, а также и σ -алгеброй. Очевидно, что пересечение $\mathcal{F}_1 \cap \mathcal{F}_2$ равно тривиальной σ -алгебре $\{\emptyset, \Omega\}$. Однако объединение

$$\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2 = \{\{\omega_1\}, \{\omega_2\}, \{\omega_2, \omega_3\}, \{\omega_1, \omega_3\}, \emptyset, \Omega\}$$

не является σ -алгеброй. В самом деле, достаточно, например, заметить, что множество $\{\omega_1\} \cup \{\omega_2\} = \{\omega_1, \omega_2\}$ не принадлежит классу $\mathcal{F}$.

(б) Пусть Ω — фиксированное множество, а $(\mathcal{A}_n)$, $n \geq 1$, — последовательность алгебр подмножеств множества Ω . Будем говорить, что $(\mathcal{A}_n)$ — возрастающая последовательность, если для каждого $n \geq 1$ имеет место строгое включение $\mathcal{A}_n \subset \mathcal{A}_{n+1}$, и что последовательность $(\mathcal{A}_n)$ — неубывающая, если $\mathcal{A}_n \subseteq \mathcal{A}_{n+1}$ при всех $n \geq 1$, но возможны равенства $\mathcal{A}_n = \mathcal{A}_{n+1}$ для некоторых индексов n .

Для любого класса $\mathcal{A}$ подмножеств Ω и любого фиксированного множества $D \subset \mathcal{A}$ введем новый класс $\mathcal{A}|D = \{A \in \mathcal{A} : A \subset D\}$. Легко заметить, что если $\mathcal{A}$ — алгебра, то $\mathcal{A}|D$ является алгеброй в D , а если $(\mathcal{A}_n)$ — возрастающая последовательность алгебр, то это свойство сохраняется и для новой последовательности $(\mathcal{A}_n|D)$.

Итак, пусть $(\mathcal{A}_n)$ — возрастающая последовательность алгебр. Мы интересуемся свойствами объединения $\mathcal{A} = \bigcup_{n=1}^{\infty} \mathcal{A}_n$. Для этой цели докажем сначала следующие два утверждения:

1. Для каждого фиксированного индекса m существует такое множество $D \in \mathcal{A} \setminus \mathcal{A}_m$, что последовательность $(\mathcal{A}_n|(\Omega \setminus D), n \geq 1)$ возрастает.

2. В классе $\mathcal{A}$ существует последовательность $(C_n, n \geq 1)$, состоящая из таких попарно непересекающихся множеств, что $C_n \notin \mathcal{A}_n$, $n \geq 1$.

Перейдем к доказательству этих утверждений. Поскольку индекс m фиксирован, а последовательность $(\mathcal{A}_n)$ возрастает, то для любых m_1 и m_2 , для которых $m < m_1 < m_2$, найдутся такие множества G_1 и G_2 , что $G_1 \in \mathcal{A}_{m_1} \setminus \mathcal{A}_m$ и $G_2 \in \mathcal{A}_{m_2} \setminus \mathcal{A}_{m_1}$. Тогда среди трех попарно непересекающихся множеств $G_1 \setminus G_2$, $G_1 \cap G_2$ и $G_2 \setminus G_1$ по крайней мере два множества не принадлежат $\mathcal{A}_m$. Таким образом, мы нашли два непересекающиеся множества (обозначим их F_1 и F_2), не входящие в $\mathcal{A}_m$.

Отсюда вытекает, что для любого $n > m_2$ имеет место соотношение

$$\mathcal{A}_n = (\mathcal{A}_n | (\Omega \setminus F_1)) \vee (\mathcal{A}_n | (\Omega \setminus F_2)).$$

(Напомним, что для любых алгебр $\mathcal{D}_1$ и $\mathcal{D}_2$ через $\mathcal{D}_1 \vee \mathcal{D}_2$ принято обозначать алгебру, порожденную объединениями $A \cup B$, $A \in \mathcal{D}_1$, $B \in \mathcal{D}_2$.) Это соотношение и свойство возрастания последовательности $(\mathcal{A}_n)$ позволяют заключить, что хотя бы одна из последовательностей, $(\mathcal{A}_n | (\Omega \setminus F_1))$ или $(\mathcal{A}_n | (\Omega \setminus F_2))$, также возрастает. Следовательно, полагая $D = F_j$, если $(\mathcal{A}_n | (\Omega \setminus F_j))$ возрастает, мы получаем требуемую возрастающую последовательность алгебр $(\mathcal{A}_n | (\Omega \setminus D))$. Этим утверждение 1 доказано.

Далее, из утверждения 1 следует существование такого множества $C_1 \in \mathcal{A}$, что $C_1 \notin \mathcal{A}_1$, причем последовательность $(\mathcal{A}_n | (\Omega \setminus C_1))$ возрастает. Рассуждая индуктивно, предположим, что в классе $\mathcal{A}$ найдены такие попарно непересекающиеся множества $C_1, \dots, C_k$, что $C_i \notin \mathcal{A}_i$, для всех $i = 1, \dots, k$, а последовательность $(\mathcal{A}_n | (\Omega \setminus S_k))$, где $S_k = C_1 \cup \dots \cup C_k$, является возрастающей. Снова, согласно утверждению 1, существует такое множество $C_{k+1} \in \mathcal{A} | (\Omega \setminus S_k)$, что $C_{k+1} \notin \mathcal{A}_{k+1} | (\Omega \setminus S_k)$, а $(\mathcal{A}_n | (\Omega \setminus S_k) | (\Omega \setminus C_{k+1}))$ — возрастающая последовательность. При этом $C_1, \dots, C_{k+1}$ — попарно непересекающиеся множества, а $C_{k+1} \notin \mathcal{A}_{k+1}$. Остается только заметить, что последовательность $(\mathcal{A}_n | (\Omega \setminus S_{k+1}))$ также возрастающая. Это следует из проведенного выше рассуждения и соотношения $\mathcal{A}_n | (\Omega \setminus S_k) | (\Omega \setminus C_{k+1}) = \mathcal{A}_n | (\Omega \setminus S_{k+1})$, где $S_{k+1} = S_k \cup C_{k+1}$. Утверждение 2 доказано.

Теперь не составляет труда показать, что объединение $\mathcal{A} = \bigcup_{n=1}^{\infty} \mathcal{A}_n$ является σ -алгеброй. Предположим обратное, т. е. что $\mathcal{A}$ — σ -алгебра. Возьмем последовательность (C_n) , построенную в утверждении 2. Пусть $\{N_1, N_2, \dots\}$ — разбиение множества натуральных чисел $\mathbb{N}$ на бесконечные подмножества. Обозначим $D_p = \bigcup_{n \in N_p} C_n$. Из сделанного предположения (о том, что $\mathcal{A}$ — σ -алгебра) вытекает, что $D_p \in \mathcal{A}$, и, следовательно, для каждого p существует такое n_p , что $D_p \in \mathcal{A}_{n_p}$. Так как последовательность алгебр $(\mathcal{A}_n)$ возрастает, то можно считать, что последовательность индексов (n_p) также (строго!) возрастает. Далее, для каждого p выберем индекс $m_p \in N_p$, для которого $m_p > n_p$, и положим $E = \bigcup_{p=1}^{\infty} C_{m_p}$. Согласно нашему предположению $E \in \mathcal{A}$, и, следовательно, существует такой индекс q , что $E \in \mathcal{A}_{n_q}$. Отсюда и из конструкции множества D_p вытекает, что $D_q \cap E = C_{m_q} \in \mathcal{A}_{n_q}$. Однако $n_q < m_q$, а это означает, что $C_{m_q} \in \mathcal{A}_{m_q}$. Полученное таким образом

противоречие (согласно утверждению 2, $C_n \notin \mathcal{A}_n$) показывает, что наше предположение (о том, что $\mathcal{A}$ — σ -алгебра) неверно.

Итак, в общем случае объединение счетного числа σ -алгебр не обязательно является σ -алгеброй.

§ 2. Вероятность

Предположим, что Ω — любое множество, а $\mathcal{A}$ — алгебра его подмножеств. Функция P называется *вероятностью* на измеримом пространстве $(\Omega, \mathcal{A})$, если она определена для всех событий $A \in \mathcal{A}$ и удовлетворяет следующим аксиомам:

а) $P(A) \geq 0$ для любого $A \in \mathcal{A}$; $P(\Omega) = 1$;

б) функция P конечно аддитивна, т. е. для любого конечного числа попарно непересекающихся событий $A_1, \dots, A_n \in \mathcal{A}$ выполнено равенство

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i);$$

в) функция P непрерывна в $\emptyset$, т. е. для любых событий $A_1, A_2, \dots \in \mathcal{A}$ таких, что $A_{n+1} \subset A_n$ и $\bigcap_{n=1}^{\infty} A_n = \emptyset$, имеет место соотношение

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

Отметим здесь, что условия б) и в) эквивалентны следующему условию:

г) функция P σ -аддитивна (счетно аддитивна), т. е.

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

для любых попарно непересекающихся событий $A_1, A_2, \dots \in \mathcal{A}$.

Пусть P_0 — σ -аддитивная вероятность на $(\Omega, \mathcal{A})$ и $\mathcal{F} = \sigma(\mathcal{A})$ обозначает наименьшую σ -алгебру, порожденную алгеброй $\mathcal{A}$. Тогда, согласно теореме Каратеодори (см. [21, 60, 75, 81, 116, 117, 139, 145, 330]), на $(\Omega, \mathcal{F})$ существует единственная вероятность P являющаяся *продолжением* P_0 , т. е. такая, что $P(A) = P_0(A)$ для $A \in \mathcal{A}$. В этом случае принято также говорить, что P_0 является *сужением* P на алгебру $\mathcal{A}$, и обозначать это так: $P|_{\mathcal{A}} = P_0$.

Упорядоченная тройка $(\Omega, \mathcal{F}, P)$ называется *вероятностным пространством*, если Ω — произвольное множество точек (объектов), называемых элементарными событиями (исходами) некоторого эксперимента; $\mathcal{F}$ — σ -алгебра подмножеств Ω , причем элементы $\mathcal{F}$ называются

случайными событиями (или просто *событиями*); функция P есть вероятность на $(\Omega, \mathcal{F})$, т. е. P удовлетворяет условиям а), б) и в), или, что эквивалентно, условиям а) и г).

Изложенные далее примеры иллюстрируют свойства вероятностных мер. Пример 2.4 посвящен важному понятию регулярной условной вероятности.

2.1. Вероятностная мера, которая является аддитивной, но не σ -аддитивной

Пусть Ω — множество всех рациональных чисел r , $r \in [0, 1]$, а $\mathcal{F}_1$ — класс подмножеств множества Ω , имеющих вид $[a, b]$, $(a, b]$, $[a, b)$ или (a, b) , где a и b — рациональные числа. Обозначим через $\mathcal{F}_2$ класс всех конечных сумм попарно непересекающихся множеств из $\mathcal{F}_1$. Тогда $\mathcal{F}_2$ является алгеброй. Определим теперь вероятностную меру P так:

$$P(A) = b - a, \quad \text{если } A \in \mathcal{F}_1;$$

$$P(B) = \sum_{i=1}^n P(A_i), \quad \text{если } B \in \mathcal{F}_2.$$

Если $B \in \mathcal{F}_2$, то подразумевается, что $B = \sum_{i=1}^n A_i$, где A_i — попарно не пересекающиеся множества из $\mathcal{F}_1$.

Мы интересуемся свойствами вероятности P . Покажем, например, что P аддитивна. Действительно, возьмем два непересекающихся множества из $\mathcal{F}_2$, $B = \sum_{i=1}^n A_i$ и $B' = \sum_{j=1}^m A'_j$, где $A_i, A'_j \in \mathcal{F}_1$ и все A_i, A'_j попарно непересекающиеся. Тогда сумму $B + B'$ можно представить в виде $B + B' = \sum_{k=1}^{m+n} C_k$, где $C_1, \dots, C_{m+n}$ — это множества $A_1, \dots, A_n$ и $A'_1, \dots, A'_m$, взятые в произвольном порядке. Тогда, согласно определению P , получаем, что

$$\begin{aligned} P(B + B') &= P\left(\sum_{k=1}^{m+n} C_k\right) = \sum_{k=1}^{m+n} P(C_k) = \\ &= \sum_{i=1}^n P(A_i) + \sum_{j=1}^m P(A'_j) = P(B) + P(B'). \end{aligned}$$

Это означает, что вероятностная мера P аддитивна. Далее легко видеть, что каждое одноточечное множество $\{r\}$ принадлежит $\mathcal{F}_2$ и $P(\{r\}) = 0$. Так как Ω — счетное множество и $\Omega = \sum_{i=1}^{\infty} \{r_i\}$, то имеем $P(\Omega) = 1 \neq 0 = \sum_{i=1}^{\infty} P(\{r_i\})$.

Полученное противоречие показывает, что вероятностная мера P не является σ -аддитивной на алгебре $\mathcal{F}_2$.

2.2. Две вероятностные меры могут совпадать на заданном классе событий и не совпадать на σ -алгебре, порожденной этим классом

Напомним сначала следующий результат (см. [116, 145, 157]): пусть Ω — произвольное множество и $\mathcal{C}$ — класс событий в Ω , замкнутый относительно операции пересечения, т. е. $A, B \in \mathcal{C} \implies AB \in \mathcal{C}$. Пусть на измеримом пространстве $(\Omega, \mathcal{F})$, где $\mathcal{F}$ — наименьшая σ -алгебра, порожденная $\mathcal{C}$, заданы две вероятности, P_1 и P_2 , совпадающие на $\mathcal{C}$. Тогда они совпадают и на σ -алгебре $\mathcal{F}$.

Неудивительно, что такие результаты существенно зависят от структуры класса $\mathcal{C}$. Проиллюстрируем это конкретным примером.

Рассмотрим множество $\Omega = \{a, b, c, d\}$ и определим вероятностные меры Q_1 и Q_2 следующим образом:

$$\begin{aligned} Q_1(\{a\}) = Q_1(\{d\}) &= \frac{1}{6}, & Q_1(\{b\}) = Q_1(\{c\}) &= \frac{1}{3}, \\ Q_2(\{a\}) = Q_2(\{d\}) &= \frac{1}{3}, & Q_2(\{b\}) = Q_2(\{c\}) &= \frac{1}{6}. \end{aligned}$$

Выберем класс $\mathcal{C} = \{C_1, C_2, C_3, C_4\}$, где $C_1 = \{a, b\}$, $C_2 = \{c, d\}$, $C_3 = \{a, c\}$, $C_4 = \{b, d\}$. Проверим, что $Q_1(C_i) = Q_2(C_i)$ для каждого $C_i \in \mathcal{C}$, $i = 1, 2, 3, 4$. Например, если $C = C_2$, находим

$$\begin{aligned} Q_1(C_2) &= Q_1(\{c\}) + Q_1(\{d\}) = \frac{1}{3} + \frac{1}{6} = \frac{1}{2}, \\ Q_2(C_2) &= Q_2(\{c\}) + Q_2(\{d\}) = \frac{1}{6} + \frac{1}{3} = \frac{1}{2}. \end{aligned}$$

Так же элементарно проверяется, что $Q_1(\cdot) = Q_2(\cdot)$, и для остальных элементов класса $\mathcal{C}$.

Пусть теперь $\mathcal{F}$ — совокупность всех подмножеств множества Ω . Тогда $\mathcal{F}$ совпадает с σ -алгеброй $\sigma(\mathcal{C})$, порожденной классом $\mathcal{C}$. Спрашивается, совпадают ли Q_1 и Q_2 на $\mathcal{F}$? Ответить на этот вопрос легко. Например, $C_1 \cap C_3 = \{a\} \in \sigma(\mathcal{C})$, однако

$$Q_1(\{a\}) = \frac{1}{6} \neq \frac{1}{3} = Q_2(\{a\}).$$

Несовпадение Q_1 и Q_2 также имеет место для каждого из остальных элементов $\{b\}$, $\{c\}$ и $\{d\}$, принадлежащих σ -алгебре $\mathcal{F}$.

Итак, $Q_1 = Q_2$ на классе $\mathcal{C}$, но $Q_1 \neq Q_2$ на σ -алгебре $\mathcal{F}$. Причина этого — незамкнутость класса $\mathcal{C}$ относительно операции пересечения.

2.3. Теорема Колмогорова о продолжении меры в пространстве $(\mathbb{R}^\infty, \mathcal{B}^\infty)$

Вероятностные меры в измеримом пространстве $(\mathbb{R}^n, \mathcal{B}^n)$, $n \geq 1$, строятся сначала для элементарных множеств (прямоугольников или параллелепипедов вида $(a, b]$, $a, b \in \mathbb{R}^n$), потом для множеств вида $A = \sum_i (a_i, b_i]$ и, наконец, с помощью теоремы Каратеодори, для множеств из $\mathcal{B}^n$. Подобную конструкцию можно использовать и для построения вероятностей в пространстве $(\mathbb{R}^\infty, \mathcal{B}^\infty)$.

Пусть $C_n(B) = \{x \in \mathbb{R}^\infty : (x_1, \dots, x_n) \in B\}$, $B \in \mathcal{B}^n$, обозначает цилиндрическое множество в $\mathbb{R}^\infty$ с основанием $B \in \mathcal{B}^n$. Естественно рассматривать цилиндрические множества как элементарные множества в $\mathbb{R}^\infty$. Вероятностную меру множеств σ -алгебры $\mathcal{B}^\infty$ можно определить через вероятности элементарных множеств.

Пусть P — вероятностная мера на $(\mathbb{R}^\infty, \mathcal{B}^\infty)$. Положим

$$P_n(B) = P(C_n(B)), \quad B \in \mathcal{B}^n, \quad n = 1, 2, \dots$$

Таким образом мы получаем последовательность вероятностных мер $P_1, P_2, \dots$, определенных соответственно на измеримых пространствах $(\mathbb{R}^1, \mathcal{B}^1), (\mathbb{R}^2, \mathcal{B}^2), \dots$. Для каждого $n = 1, 2, \dots$ и любого $B \in \mathcal{B}^n$ имеет место равенство

$$P_{n+1}(B \times \mathbb{R}^1) = P_n(B), \quad (1)$$

называемое *свойством согласованности*. Фундаментальным является следующий обратный результат (теорема Колмогорова о продолжении меры на $(\mathbb{R}^\infty, \mathcal{B}^\infty)$).

Пусть $P_1, P_2, \dots$ — последовательность вероятностных мер на $(\mathbb{R}^1, \mathcal{B}^1), (\mathbb{R}^2, \mathcal{B}^2), \dots$, обладающих свойством согласованности (1). Тогда на $(\mathbb{R}^\infty, \mathcal{B}^\infty)$ существует такая единственная, вероятностная мера P , что $P(C_n(B)) = P_n(B)$ для любого множества $B \in \mathcal{B}^n$, $n = 1, 2, \dots$

Доказательство этой теоремы может быть найдено, например, в книгах [22, 60, 116, 117, 145, 330, 345, 404]. Заметим, что оно использует некоторые специфические свойства евклидовых пространств. Оказывается, что в общем случае, без предположений о топологической структуре измеримых пространств и о структуре семейства мер $\{P_n\}$, эта теорема неверна. Вот конкретный пример.

Рассмотрим пространство $\Omega = (0, 1]$, которое, очевидно, не является полным. В этом пространстве построим последовательность σ -алгебр $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots$ и последовательность вероятностных мер $P_1, P_2, \dots$, где P_n определена на $(\Omega, \mathcal{F}_n)$. Затем мы введем σ -алгебру $\mathcal{F} = \sigma\left(\bigcup_n \mathcal{F}_n\right)$, являющуюся наименьшей σ -алгеброй, содержащей все $\mathcal{F}_n$, и покажем, что

не существует такой вероятностной меры P на $(\Omega, \mathcal{F})$, что ее сужение $P|_{\mathcal{F}_n}$ на $\mathcal{F}_n$ совпадало бы с P_n , $n = 1, 2, \dots$

Действительно, для $n = 1, 2, \dots$ определим функцию $h_n(\omega) = 1$, если $0 < \omega < \frac{1}{n}$, и $h_n(\omega) = 0$, если $\frac{1}{n} \leq \omega \leq 1$. Пусть

$$\mathcal{C}_n = \{A \subset \Omega: A = [\omega: h_n(\omega) \in B], B \in \mathcal{B}^1\}$$

и $\mathcal{F}_n = \sigma(\mathcal{C}_1, \dots, \mathcal{C}_n)$ — наименьшая σ -алгебра, содержащая все классы множеств $\mathcal{C}_1, \dots, \mathcal{C}_n$. Ясно, что $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots$. Введем σ -алгебру $\mathcal{F} = \sigma(\bigcup \mathcal{F}_n)$, содержащую все $\mathcal{F}_n$. На измеримом пространстве $(\Omega, \mathcal{F}_n)$ определим вероятностную меру P_n следующим образом:

$$P_n\{\omega: (h_1(\omega), \dots, h_n(\omega)) \in B^n\} = \begin{cases} 1, & \text{если } (1, \dots, 1) \in B^n, B^n \in \mathcal{B}^n, \\ 0 & \text{в остальных случаях.} \end{cases}$$

Заметим, что $\mathcal{C}_n$ содержит лишь множества $(0, \frac{1}{n})$, $[\frac{1}{n}, 1]$, $\emptyset$ и Ω , и легко описать структуру σ -алгебры $\mathcal{F}_n$. Кроме того, мера P_n сосредоточена на $(0, \frac{1}{n})$. Тогда семейство мер $\{P_n\}$ удовлетворяет свойству согласованности (1), т. е. $P_{n+1}(A) = P_n(A)$ для всех $A \in \mathcal{F}_n$.

Допустим теперь, что на измеримом пространстве $(\Omega, \mathcal{F})$ существует такая вероятностная мера P , что $P|_{\mathcal{F}_n} = P_n$. Тогда для любого $n = 1, 2, \dots$ имело бы место соотношение

$$P\{\omega: h_1(\omega) = \dots = h_n(\omega) = 1\} = P_n\{\omega: h_1(\omega) = \dots = h_n(\omega) = 1\} = 1. \quad (2)$$

Заметим, однако, что монотонное убывание последовательности

$$\{\omega: h_1(\omega) = \dots = h_n(\omega) = 1\} = (0, \frac{1}{n})$$

противоречит равенствам (2) и требованию счетной аддитивности функции P , а значит, и непрерывности этой функции в «нуле» $\emptyset$.

2.4. О регулярных условных вероятностях

Пусть $(\Omega, \mathcal{F}, P)$ — вероятностное пространство и $\mathcal{F}$ — такая σ -алгебра, что $\mathcal{F}_1 \subset \mathcal{F}$. Напомним, что условная вероятность $P(A|\mathcal{F}_1)$, $A \in \mathcal{F}$, определена P -п. н. как такая $\mathcal{F}_1$ -измеримая функция ω , что

$$P(AB) = \int_B P(A|\mathcal{F}_1) dP(\omega) \quad \text{для всех } B \in \mathcal{F}_1.$$

Условная вероятность $P(A|\mathcal{F}_1)$, $A \in \mathcal{F}$, называется *регулярной*, если существует функция $P(A, \omega)$, $A \in \mathcal{F}$, $\omega \in \Omega$, удовлетворяющая условиям:

- а) $P(A, \omega) = P(A|\mathcal{F}_1)$ (Р-п.н.) для любого $A \in \mathcal{F}$;
- б) $P(\cdot, \omega)$ — вероятностная мера на $\mathcal{F}$ при фиксированном ω .

Если выполнено условие б), а условие а) имеет место для всех $\omega \in \Omega$ (а не только для Р-почти всех ω), то $P(A|\mathcal{F}_1)$ называется *собственной регулярной условной вероятностью*. (В терминах распределений естественно говорить о регулярном и о собственном регулярном условном распределении.)

Отметим, что регулярные условные вероятности существуют во многих случаях (см., например, [22, 32, 75, 116, 145, 147, 345]), но собственная регулярная вероятность существует не всегда, как показывает рассмотренный далее пример.

Пусть вероятностное пространство $(\Omega, \mathcal{F}, \lambda)$ имеет следующую структуру: $\Omega = [0, 1]$, $\mathcal{F}$ — σ -алгебра измеримых по Лебегу множеств Ω , и λ — мера Лебега. Хорошо известно, что на отрезке $[0, 1]$ существует такое неизмеримое (в смысле Лебега) множество N , что его внешняя мера $\lambda^*(N)$ равна 1, а его внутренняя мера $\lambda_*(N)$ равна 0 (подробности см. в [61, 108, 114]).

Введем теперь новую σ -алгебру $\tilde{\mathcal{F}}$, порожденную σ -алгеброй $\mathcal{F}$ и множеством N . Это означает, что $\tilde{\mathcal{F}}$ состоит из множеств вида $NB_1 \cup \bar{N}B_2$, где $B_1, B_2 \in \mathcal{F}$. На измеримом пространстве $(\Omega, \tilde{\mathcal{F}})$ определим меру P с помощью равенства

$$P(NB_1 \cup \bar{N}B_2) = \frac{1}{2}[\lambda(B_1) + \lambda(B_2)].$$

Легко убедиться, что мера P определена корректно и, более того, P — вероятность на $\tilde{\mathcal{F}}$. Таким образом, тройка $([0, 1], \tilde{\mathcal{F}}, P)$ задает вероятностное пространство. Для любого $B \in \mathcal{F}$ мы имеем

$$P(NB \cup \bar{N}B) = P(B) = \lambda(B),$$

и, следовательно, на $\mathcal{F}$ мера P совпадает с λ , т. е. $P|_{\mathcal{F}} = \lambda$. Кроме того, $P(N) = \frac{1}{2}$.

Вопрос: существует ли собственная регулярная условная вероятность $P(A|\mathcal{F})$, $A \in \mathcal{F}$, относительно σ -алгебры $\mathcal{F}$? Предположим, что ответ положителен, т. е. что существует функция $P(A, \omega)$, удовлетворяющая указанным выше условиям а) и б). В силу этого, для любого борелевского (и лебеговского) множества A имеем $P(A, \omega) = 1_A(\omega)$ почти наверное. Следовательно, если A — одноточечное множество, $A = \{\omega\}$, то $P(\{\omega\}, \omega) = 1$. Рассмотрим множество N . Из определения

условной вероятности и равенства $P(N) = \frac{1}{2}$ получаем

$$\frac{1}{2} = P(A) = \int_{\Omega} P(N, \omega) \lambda(d\omega). \quad (3)$$

С другой стороны, если $P(\cdot, \omega)$ — мера для каждого ω , тогда

$$P(N, \omega) \geq P(\{\omega\}, \omega) = 1 \quad \text{для} \quad \omega \in N.$$

Следовательно,

$$P(N, \omega) = 1 \quad \text{для всех} \quad \omega \in N.$$

Рассмотрим множество $C = \{\omega : P(\{\omega\}, \omega) = 1\}$. Так как $P(\cdot, \omega)$ — измеримая функция ω , то C — измеримое (борелевское) множество для которого $P(C) = 1$. Пусть $D = \{\omega : P(N, \omega) = 1\}$. Тогда ясно, что D — (борелевское) измеримое множество и $D \supset CN$, откуда следует, что $D \cup \bar{C} \supset N$. Значит, борелевское множество $D \cup \bar{C}$ содержит (неизмеримое!) множество N , которое имеет внешнюю меру $\lambda^*(N) = 1$. Следовательно, $P(D \cup \bar{C}) = 1$ и $P(D) = 1$. Другими словами, для почти всех ω имеем $P(N, \omega) = 1$, которое влечет следующее равенство

$$\int_{\Omega} P(N, \omega) \lambda(d\omega) = 1.$$

Это, очевидно, противоречит полученному выше соотношению (3). Следовательно, наше допущение неверно. Поэтому собственная регулярная условная вероятность $P(A|\mathcal{F})$, $A \in \tilde{\mathcal{F}}$ не существует.

Наконец, отметим, что в этом примере существенна роль неизмеримого множества N . Напомним, что конструкция множества N основывается на аксиоме выбора (см. [61, 108]). По поводу других конструкций и результатов читатель может обратиться к работе [326]. Общие результаты содержатся в работах [22, 32, 75, 81, 116, 145, 157, 345, 404].

§ 3. Независимость случайных событий

Пусть $(\Omega, \mathcal{F}, P)$ — вероятностное пространство. События A и B , $A, B \in \mathcal{F}$, называются *независимыми* (относительно вероятности P), если имеет место следующее мультипликативное свойство:

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

Два класса событий (например, алгебр, σ -алгебр), $\mathcal{A}_1$ и $\mathcal{A}_2$, где $\mathcal{A}_1, \mathcal{A}_2 \subset \mathcal{F}$, называются *независимыми* (относительно P), если независимы любые два события $A_1 \in \mathcal{A}_1$ и $A_2 \in \mathcal{A}_2$.

Мы говорим, что события $A_1, \dots, A_n \in \mathcal{F}$ *независимы в совокупности*, если для любых $k = 2, \dots, n$ и $1 \leq i_1 < i_2 < \dots < i_k \leq n$ имеет место соотношение

$$P(A_{i_1} A_{i_2} \dots A_{i_k}) = P(A_{i_1})P(A_{i_2}) \dots P(A_{i_k}). \quad (1)$$

Число соотношений (1) равно $2^n - n - 1$. Если (1) выполнено только при $k = 2$, то заданные события называются *попарно независимыми*.

Очевидным образом понятие независимости в совокупности вводит-ся и для любого конечного числа классов событий.

Пусть $B \in \mathcal{F}$ — произвольное событие с вероятностью $P(B) > 0$. Условная вероятность любого события $A \in \mathcal{F}$ относительно события B обозначается $P(A|B)$ и задается равенством

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

Свойство независимости двух событий может быть легко выражено в терминах условных вероятностей.

Рассмотренные далее примеры иллюстрируют связь между различными понятиями независимости.

3.1. О роли вероятностной меры в свойстве независимости

Когда мы приводили определение независимости двух событий, мы специально отмечали, что это свойство вводится для фиксированной вероятности P . Это означает, что два события могут быть независимыми относительно одной вероятности и не быть независимыми относительно другой. Проиллюстрируем эту возможность на простом примере, связанном со схемой Бернулли, т.е. с повторными независимыми испытаниями, каждое из которых имеет два исхода, У и Н («успех» и «неудача»). Предположим, что $P(H) = p$, где $0 \leq p \leq 1$. Тогда $P(Y) = 1 - p$. Проведем три независимых испытания и рассмотрим события

$$A = \{\text{не более одного успеха}\}, \quad B = \{\text{исход один и тот же}\}.$$

Очевидно, что

$$A = \{HHN, HNU, HUN, UHH\}, \quad B = \{HHH, UUU\}.$$

Следовательно,

$$P(A) = p^3 + 3p^2(1-p), \quad P(B) = p^3 + (1-p)^3, \quad P(AB) = p^3.$$

Теперь легко видеть, что равенство

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

имеет место в тривиальных случаях $p = 0$ и $p = 1$, а также при $p = \frac{1}{2}$, т. е. в симметрическом случае. Это означает, что события A и B независимы при $p = 0$, $p = 1$ и $p = \frac{1}{2}$. Для остальных значений вероятности успеха p они не являются независимыми.

3.2. Попарная независимость событий не влечет их независимости в совокупности

Приведем несколько примеров, в каждом из которых рассмотренные события (их число больше двух) будут попарно независимыми, но не будут независимыми в совокупности. Естественно начать с классического примера Бернштейна — это случай (а).

(а) Предположим, что в ящике есть четыре билета с номерами 112, 121, 211 и 222 (все номера на билетах различны). Случайным образом выберем один билет и рассмотрим события $A_1 = \{\text{в номере билета на первом месте стоит цифра 1}\}$, $A_2 = \{\text{на втором месте стоит цифра 1}\}$ и $A_3 = \{\text{на третьем месте стоит цифра 1}\}$. Тогда очевидно, что

$$P(A_1) = P(A_2) = P(A_3) = \frac{1}{2} \quad \text{и} \quad P(A_1 A_2) = P(A_1 A_3) = P(A_2 A_3) = \frac{1}{4}.$$

Это означает, что события A_1, A_2, A_3 попарно независимы. Однако

$$P(A_1 A_2 A_3) = P(\emptyset) = 0 \neq \frac{1}{8} = P(A_1)P(A_2)P(A_3),$$

и, следовательно, эти три события не являются независимыми в совокупности.

(б) Пусть в пространство элементарных исходов Ω входят все перестановки, образованные буквами a, b, c , а также тройки (a, a, a) , (b, b, b) и (c, c, c) . Каждому из этих девяти элементов Ω приписываем вероятность $\frac{1}{9}$. Рассмотрим события

$$A_k = \{\text{буква } a \text{ находится на месте с номером } k\}, \quad k = 1, 2, 3.$$

Легко видеть, что

$$P(A_1) = P(A_2) = P(A_3) = \frac{1}{3}, \quad P(A_1 A_2) = P(A_1 A_3) = P(A_2 A_3) = \frac{1}{9}.$$

Итак, события A_1, A_2, A_3 попарно независимы. Однако они не являются независимыми в совокупности, так как $A_1 A_2 \subset A_3$ и

$$P(A_1 A_2 A_3) = P(A_1 A_2) = \frac{1}{9} \neq \frac{1}{27} = P(A_1)P(A_2)P(A_3).$$

Этот пример можно обобщить следующим образом. Предположим, что Ω содержит $n! + n$ элементов, а именно, все $n!$ перестановок символов $a_1, \dots, a_n$, а также элементы $(a_k, \dots, a_k)$, $k = 1, \dots, n$, составленные

каждый раз из одного символа. Пусть каждая перестановка имеет вероятность $\frac{1}{n^2(n-2)!}$, а каждый элемент вида $(a_k, \dots, a_k)$, $k = 1, \dots, n$ имеет вероятность $\frac{1}{n^2}$. (Очевидно, сумма этих вероятностей равна 1.) Рассмотрим события

$$A_k = \{\text{символ } a_1 \text{ находится на месте с номером } k\}, \quad k = 1, \dots, n.$$

Легко проверить, что $A_1, \dots, A_n$ попарно независимы. Далее, для любых трех из них, $A_{i_1}, A_{i_2}, A_{i_3}$, где $1 \leq i_1 < i_2 < i_3 \leq n$, не выполнено равенство $P(A_{i_1}A_{i_2}A_{i_3}) = P(A_{i_1})P(A_{i_2})P(A_{i_3})$. Таким образом, события $A_1, \dots, A_n$ не являются независимыми в совокупности.

(в) Пусть задано множество символов $\{a_1, \dots, a_n\}$, и пусть пространство исходов Ω состоит из всех трехбуквенных слов $a_i a_j a_k$ с возможными повторениями. Предположим, что каждый из элементов Ω , а их число равно n^3 , имеет одну и ту же вероятность, т. е. $\frac{1}{n^3}$ (равномерное дискретное распределение). Рассмотрим следующие три события:

$$A = \{\omega \in \Omega: \omega \text{ начинается фиксированной буквой, например, } a_1\},$$

$$B = \{\omega \in \Omega: \text{средняя буква } \omega \text{ — это } a_2\},$$

$$C = \{\omega \in \Omega: \omega \text{ записывается ровно двумя буквами}\}.$$

Тогда A, B и C — попарно независимые события, не являющиеся независимыми в совокупности.

3.3. Из соотношения $P(ABC) = P(A)P(B)P(C)$ не следует, что события A, B и C независимы в совокупности

(а) Пусть эксперимент состоит в подбрасывании двух различных правильных игральных костей. Тогда пространство элементарных исходов Ω состоит из всех упорядоченных пар ij ($i, j = 1, \dots, 6$) и каждый исход имеет вероятность $\frac{1}{36}$. Рассмотрим события $A = \{i = 1, 2 \text{ или } 3\}$, $B = \{j = 3, 4 \text{ или } 5\}$ и $C = \{i + j = 9\}$, где i — число, выпавшее на первой кости, а j — число, выпавшее на второй кости. Очевидно, что

$$AB = \{31, 32, 33, 34, 35, 36\}, \quad AC = \{36\},$$

$$BC = \{36, 45, 54\}, \quad ABC = \{36\}.$$

Тогда $P(A) = \frac{1}{2}$, $P(B) = \frac{1}{2}$, $P(C) = \frac{1}{9}$ и, значит,

$$P(ABC) = \frac{1}{36} = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{9} = P(A)P(B)P(C).$$

Заметим, однако, что

$$P(AB) = \frac{1}{6} \neq \frac{1}{4} = P(A)P(B),$$

$$P(AC) = \frac{1}{36} \neq \frac{1}{18} = P(A)P(C),$$

$$P(BC) = \frac{1}{12} \neq \frac{1}{18} = P(B)P(C).$$

Итак, равенство $P(ABC) = P(A)P(B)P(C)$ выполнено, но, тем не менее, события A, B, C не являются попарно независимыми, а следовательно, они не являются независимыми в совокупности.

(б) Предположим, что пространство элементарных событий Ω допускает такое разбиение на пять множеств: A_1, A_2, A_3, A_4 и A_5 (т. е. $A_i A_j = \emptyset$ при $i \neq j$ и $A_1 + \dots + A_5 = \Omega$), что

$$P(A_1) = P(A_2) = P(A_3) = \frac{15}{64}, \quad P(A_4) = \frac{1}{64}, \quad P(A_5) = \frac{18}{64}.$$

Рассмотрим три события: $B = A_1 \cup A_4, C = A_2 \cup A_4, D = A_3 \cup A_4$. Тогда

$$P(B) = P(C) = P(D) = \frac{1}{4}, \quad P(BCD) = \frac{1}{64}.$$

Таким образом, $P(BCD) = P(B)P(C)P(D)$. Однако имеет место, например, соотношение

$$P(BC) = P(A_4) = \frac{1}{64} \neq \frac{1}{16} = P(B)P(C).$$

Следовательно, события B, C и D не являются независимыми в совокупности.

3.4. Система из $n + 1$ зависимых событий, каждые n из которых независимы в совокупности

(а) Проводим n независимых бросаний правильной монеты и рассмотрим события $A_k = \{\text{результат } k\text{-го бросания} - \text{«герб»}\}$, где $k = 1, \dots, n$, и $A_{n+1} = \{\text{число «гербов» при этих бросаниях} - \text{четное}\}$. Тогда

$$P(A_1) = \frac{1}{2}, \quad \dots, \quad P(A_n) = \frac{1}{2},$$

$$P(A_{n+1}) = \frac{1}{2^n} \left[\binom{n}{0} + \binom{n}{2} + \dots \right] = \frac{2^{n-1}}{2^n} = \frac{1}{2}.$$

(Для того чтобы найти $P(A_{n+1})$, мы воспользовались известным тождеством для биномиальных коэффициентов: $\binom{n}{0} + \binom{n}{2} + \dots = 2^{n-1}$.) Легко сообразить, что условная вероятность $P(A_{n+1} | A_1 \dots A_n)$ равна 1,

если n четно, и равна 0, если n нечетно. Отсюда следует, что равенство

$$P(A_1 \dots A_n A_{n+1}) = P(A_1) \dots P(A_n) P(A_{n+1})$$

не имеет места, потому что его правая часть равна $2^{-(n+1)}$, в то время как для левой части выполняется равенство

$$P(A_1 \dots A_n A_{n+1}) = P(A_1 \dots A_n) P(A_{n+1} | A_1 \dots A_n),$$

т. е. левая часть равна 2^{-n} , если n четно, и равна 0, если n нечетно. Следовательно, события $A_1, \dots, A_n, A_{n+1}$ не являются независимыми в совокупности.

Покажем теперь, что любые n из этих (зависимых) событий независимы в совокупности. Действительно, для «первых» n событий, т. е. $A_1, \dots, A_n$, это свойство выполнено, так как для любых $A_{i_1}, \dots, A_{i_k}$, где $2 \leq k \leq n$, мы имеем $P(A_{i_1} \dots A_{i_k}) = P(A_{i_1}) \dots P(A_{i_k}) = 2^{-k}$. Поэтому достаточно рассмотреть такой набор n событий, который включает событие A_{n+1} и $n - 1$ из событий $A_1, \dots, A_n$. Не ограничивая общности, можно считать, что мы выбрали для рассмотрения события $A_2, A_3, \dots, A_n, A_{n+1}$. Для их независимости в совокупности достаточно проверить соотношения

$$P(A_{i_1} \dots A_{i_m} A_{n+1}) = P(A_{i_1}) \dots P(A_{i_m}) P(A_{n+1}), \quad (1)$$

где $1 \leq m \leq n - 1$, а индексы $i_1, \dots, i_m$ принадлежат множеству $\{2, \dots, n\}$. Поскольку

$$P(A_{i_1}) = \dots = P(A_{i_m}) = P(A_{n+1}) = \frac{1}{2},$$

правая часть (1) равна $2^{-(m+1)}$. Далее, для левой части (1) получаем

$$\begin{aligned} P(A_{i_1} \dots A_{i_m} A_{n+1}) &= P(A_{i_1} \dots A_{i_m}) P(A_{n+1} | A_{i_1} \dots A_{i_m}) = \\ &= 2^{-m} \cdot 2^{-1} = 2^{-(m+1)}. \end{aligned}$$

Итак, справедливость соотношения (1) установлена, и, следовательно, любые n среди $n + 1$ (зависимых) событий $A_1, \dots, A_n, A_{n+1}$ являются независимыми в совокупности.

(б) Рассмотрим плоскость и в ней $n + 1$ точку, скажем, $M_1, \dots, M_{n+1}$, которые находятся в общем положении (т. е. никакие три из них не лежат на одной прямой). Соединив отрезком каждую пару точек, получим $\binom{n+1}{2}$ отрезков $M_i M_j$, $1 \leq i < j \leq n + 1$. Производим $\binom{n+1}{2}$ независимых бросаний правильной монеты и в зависимости от исхода поставим стрелку на каждом отрезке по следующему правилу. Рассмотрим

отрезок $M_i M_j$; если выпал «герб», то стрелку ставим в направлении от M_i к M_j ; в случае «решка» стрелка направляется от M_j к M_i . Таким образом, в каждую из заданных точек идет некоторое число стрелок. Рассмотрим события $A_1, \dots, A_{n+1}$, где $A_k = \{\text{число стрелок, направленных в точку } M_k, \text{ четное}\}$, $k = 1, \dots, n+1$. Тогда можно проверить, что для каждого k , $2 \leq k \leq n$, и любых $i_1, i_2, \dots, i_k$, $1 \leq i_1 < i_2 < \dots < i_k \leq n+1$, события $A_{i_1}, \dots, A_{i_k}$ независимы в совокупности, но все $n+1$ событий $A_1, \dots, A_{n+1}$ этим свойством не обладают.

Отметим наконец, что, согласно примеру 3.2, в общем случае попарная независимость событий не влечет их независимости в совокупности. Настоящий пример, обобщающий пример 3.2, показывает, что даже если любые n среди $n+1$ событий (в примере 3.2 $n=2$) независимы в совокупности, то вся система из $n+1$ событий может не обладать свойством независимости в совокупности.

3.5. О независимости и условной независимости случайных событий

Напомним, что случайные события A_1 и A_2 называются *условно независимыми* относительно заданного события B , $P(B) > 0$, если

$$P(A_1 A_2 | B) = P(A_1 | B) P(A_2 | B). \quad (1)$$

Интересно выяснить, какова связь между понятиями независимости и условной независимости.

(а) Предположим, что симметричная монета подбрасывается 2 раза. Рассмотрим события $A_k = \{\text{при } k\text{-м бросании выпал «герб»}\}$, $k = 1, 2$ и $B = \{\text{при двух бросаниях выпала по крайней мере одна «решка»}\}$. Тогда $P(A_1) = P(A_2) = \frac{1}{2}$, $P(A_1 A_2) = \frac{1}{4}$, так что $P(A_1 A_2) = P(A_1) P(A_2)$, и, значит, события A_1 и A_2 независимы. Далее, $P(A_1 | B) = P(A_2 | B) = \frac{1}{3}$, но $P(A_1 A_2 | B) = 0$.

Таким образом, условие (1) не выполнено, и, следовательно, независимые события A_1 и A_2 не являются условно независимыми относительно B .

(б) Пусть имеется две монеты, скажем a и b , для которых исход «герб» имеет вероятность соответственно p_a и p_b , $0 \leq p_a \leq 1$, $0 \leq p_b \leq 1$. Предположим, что $p_a \neq p_b$. Случайным образом выберем монету и подбросим ее дважды. Рассмотрим события $A_k = \{\text{при } k\text{-м бросании выпал «герб»}\}$, $k = 1, 2$, и $B = \{\text{для эксперимента выбрана монета } a\}$. Тогда

$$P(A_1 A_2 | B) = p_a p_a, \quad P(A_1 | B) = p_a, \quad P(A_2 | B) = p_a,$$

т.е. соотношение (1) выполнено, и, значит, события A_1 и A_2 условно независимы относительно B . Далее имеем

$$P(A_1 A_2) = \frac{1}{2}p_a^2 + \frac{1}{2}p_b^2,$$

$$P(A_1) = \frac{1}{2}(p_a + p_b), \quad P(A_2) = \frac{1}{2}(p_a + p_b),$$

и, поскольку $p_a \neq p_b$, $P(A_1 A_2) \neq P(A_1)P(A_2)$, т.е. события A_1 и A_2 не являются независимыми, хотя они независимы условно.

3.6. Об одном выборе условий типа независимости, из которого не следует независимость в совокупности

Пусть $A_1, A_2, \dots, A_n$ — события в некотором эксперименте со случайными исходами, и пусть для некоторых чисел $p_1, p_2, \dots, p_n$, где $0 < p_i < 1$, выполнены следующие соотношения:

$$P(A_1) = p_1, \quad P(A_2) = p_2, \quad \dots, \quad P(A_n) = p_n, \quad P(A_1 A_2) = p_1 p_2,$$

$$P(A_1 A_2 A_3) = p_1 p_2 p_3, \quad \dots, \quad P(A_1 A_2 \dots A_n) = p_1 p_2 \dots p_n. \quad (1)$$

Эти условия можно назвать условиями «типа независимости». Если $n = 2$ и соотношения (1) выполнены, то мы имеем два независимых события A_1 и A_2 . Очевидно, что при $n \geq 3$ условия (1), их общее число равно $\frac{n(n+1)}{2}$, «более слабые», чем условия независимости в совокупности, см. соотношения (1) в начале настоящего параграфа. Тем не менее можно поставить вопрос: следует ли из условий (1), что события $A_1, \dots, A_n$ независимы в совокупности? На конкретном примере, при $n = 3$, покажем, что ответ отрицателен.

Пусть A_1, A_2, A_3 — такие события, что

$$P(A_1 A_2 A_3) = P(A_1 A_2 \bar{A}_3) = \frac{1}{8},$$

$$P(A_1 \bar{A}_2 A_3) = P(\bar{A}_1 A_2 A_3) = \frac{1}{8} - \varepsilon,$$

$$P(A_1 \bar{A}_2 \bar{A}_3) = P(\bar{A}_1 A_2 \bar{A}_3) = \frac{1}{8} + \varepsilon,$$

$$P(\bar{A}_1 \bar{A}_2 A_3) = \frac{1}{8} + 2\varepsilon, \quad P(\bar{A}_1 \bar{A}_2 \bar{A}_3) = \frac{1}{8} - 2\varepsilon,$$

где $0 < \varepsilon \leq \frac{1}{16}$. Легко подсчитать, что

$$P(A_1) = P(A_2) = P(A_3) = \frac{1}{2}, \quad P(A_1 A_2) = \frac{1}{4}, \quad P(A_1 A_2 A_3) = \frac{1}{8},$$

т.е. условия (1) выполнены. Чтобы события A_1, A_2, A_3 были независимыми в совокупности, необходимо выполнение еще двух равенств, а именно $P(A_1 A_3) = P(A_1)P(A_3)$ и $P(A_2 A_3) = P(A_2)P(A_3)$. Однако

$$P(A_1 A_3) = \frac{1}{4} - \varepsilon \neq \frac{1}{4} = P(A_1)P(A_3).$$

Следовательно, условия «типа независимости» (1) не могут обеспечить свойство независимости в совокупности.

3.7. Из независимых в совокупности событий можно формировать зависимые классы событий

Из отрезка $[0, 1]$ случайным образом выберем число x и рассмотрим его k -ичное разложение, где $k = 2, 3, \dots$. Пусть $A_m^{(k)}$ — множество всех x , в k -ичном разложении которых на месте с номером m стоит цифра 0. Множества $A_m^{(k)}$, $m = 1, 2, \dots$ образуют класс, который обозначим через $\mathcal{A}^{(k)}$. Если k фиксировано, то нетрудно проверить, что события $A_1^{(k)}, A_2^{(k)}, \dots, A_n^{(k)}$ независимы в совокупности при каждом n , $n = 2, 3, \dots$ (подробности см. в [145, 380]). Вопрос: являются ли классы $\mathcal{A}^{(2)}, \mathcal{A}^{(3)}, \dots, \mathcal{A}^{(k)}$, $k \geq 2$, независимыми в совокупности? Для ответа рассмотрим события $A_1^{(2)}, A_1^{(3)}, \dots, A_1^{(n)}$, принадлежащие соответственно классам $\mathcal{A}^{(2)}, \mathcal{A}^{(3)}, \dots, \mathcal{A}^{(n)}$. С одной стороны,

$$P\left(\bigcap_{k=2}^n A_1^{(k)}\right) = P\left(x < \frac{1}{n}\right) \quad \text{при } n > 2,$$

так как в k -ичном разложении x на первом месте ($m = 1$) стоит цифра 0 тогда и только тогда, когда $x < \frac{1}{k}$ для $k = 2, 3, \dots, n$. Но, с другой стороны,

$$\prod_{k=2}^n P(A_1^{(k)}) = \prod_{k=2}^n \frac{1}{k} = \frac{1}{n!} \neq \frac{1}{n} = P\left(\bigcap_{k=2}^n A_1^{(k)}\right),$$

и, следовательно, классы $\mathcal{A}^{(k)}$, $k = 2, 3, \dots$ не являются независимыми.

3.8. Семейства случайных событий с «необычными» свойствами независимости/зависимости

Рассмотрим несколько моделей экспериментов со случайными исходами. Пространство исходов обозначим через $\Omega = \{1, 2, \dots\}$, а соответствующие вероятности $p_i = P(\{i\})$, $i = 1, 2, \dots$

(а) Предположим, что $\Omega = \{1, 2, 3, 4, 5, 6, 7, 8\}$, а

$$p_1 = \alpha - \frac{1}{8}, \quad p_2 = p_3 = p_4 = \frac{7 - 16\alpha}{24}, \quad p_5 = p_6 = p_7 = \frac{1 + 8\alpha}{24}, \quad p_8 = \frac{1}{8}.$$

Здесь параметр α — любое число из интервала $\left(\frac{1}{8}, \frac{7}{16}\right)$. Рассмотрим следующие три события:

$$A_1 = \{2, 5, 6, 8\}, \quad A_2 = \{3, 5, 6, 8\}, \quad A_3 = \{4, 6, 7, 8\}.$$

Легко найти, что $P(A_1) = P(A_2) = P(A_3) = \frac{1}{2}$, а также, что

$$P(A_1 A_2) = P(A_1 A_3) = P(A_2 A_3) = \frac{1+2\alpha}{6}, \quad P(A_1 A_2 A_3) = \frac{1}{8}.$$

Это означает, что для любого $\alpha \in \left(\frac{1}{8}, \frac{7}{16}\right)$ имеет место равенство

$$P(A_1 A_2 A_3) = P(A_1)P(A_2)P(A_3). \quad (1)$$

Для независимости в совокупности этих трех событий нам нужна и их попарная независимость. Выбор $\alpha = \frac{1}{4}$ — это единственный, когда $P(A_i A_j) = P(A_i)P(A_j)$, $i, j = 1, 2, 3$. В этом случае каждый исход пространства Ω имеет вероятность $\frac{1}{8}$. Если $\alpha \in \left(\frac{1}{8}, \frac{7}{16}\right)$, но $\alpha \neq \frac{1}{4}$, тогда $P(A_i A_j) \neq P(A_i)P(A_j)$.

Итак, события A_1, A_2, A_3 независимы в совокупности только в случае равновероятных исходов Ω , т.е. когда $\alpha = \frac{1}{4}$. Если параметр $\alpha \in \left(\frac{1}{8}, \frac{1}{4}\right) \cup \left(\frac{1}{4}, \frac{7}{16}\right)$, события A_1, A_2, A_3 не являются независимыми в совокупности, несмотря на мультипликативное свойство (1).

(б) Пусть $\Omega = \{1, 2, 3, 4, 5, 6\}$, $p_1 = \frac{1}{16}$, $p_2 = p_3 = p_4 = p_5 = \frac{7}{48}$, $p_6 = \frac{17}{48}$. Рассмотрим следующие четыре события:

$$A_1 = \{1, 2, 3, 4\}, \quad A_2 = \{1, 2, 3, 5\}, \quad A_3 = \{1, 2, 4, 5\}, \quad A_4 = \{1, 3, 4, 5\}.$$

Легко видеть, что

$$P(A_1) = P(A_2) = P(A_3) = P(A_4) = \frac{1}{2}.$$

Тогда

$$P(A_i A_j) = \frac{17}{48} \neq \frac{1}{4} = P(A_i)P(A_j), \quad i < j,$$

$$P(A_i A_j A_k) = \frac{5}{24} \neq \frac{1}{8} = P(A_i)P(A_j)P(A_k), \quad i < j < k.$$

Однако

$$P(A_1 A_2 A_3 A_4) = \frac{1}{16} = P(A_1)P(A_2)P(A_3)P(A_4).$$

Следовательно, $\{A_1, A_2, A_3, A_4\}$ — семейство случайных событий, таких что мультипликативное свойство $P(A_{i_1} A_{i_2} \dots) = P(A_{i_1})P(A_{i_2}) \dots$

не выполнено для любых двух и любых трех из этих событий, но это свойство выполнено для всех четырех событий. Очевидно, что события A_1, A_2, A_3, A_4 не являются независимыми в совокупности.

Здесь уместно отметить следующее. Предположим, что $\mathcal{A} = \{A_1, \dots, A_n\}$ — семейство случайных событий, которые заданы в конкретной вероятностной модели. Для любого набора $A_{i_1}, \dots, A_{i_k}$, где $k = 2, \dots, n$, $1 \leq i_1 < i_2 < \dots < i_k \leq n$, мультипликативное свойство $P(A_{i_1} \dots A_{i_k}) = P(A_{i_1}) \dots P(A_{i_k})$ либо выполнено, либо не выполнено. Выполнение или невыполнение этих соотношений задает структуру независимости/зависимости семейства $\mathcal{A}$. Оказывается, можно построить вероятностную модель и семейство $\mathcal{A}$ такие, что структура независимости/зависимости $\mathcal{A}$ задана заранее. Детали можно найти в работе [453].

§ 4. Разные свойства случайных событий и их вероятностей

В этом параграфе речь будет идти о независимости событий и разбиениях пространства элементарных исходов, а также об одном очень важном для теории вероятностей результате — лемме Бореля–Кантелли. Необходимые определения и формулировки будут даны в самих примерах. Изложенные здесь факты являются естественным продолжением и дополнением материала §§ 1–3.

4.1. Вероятностные пространства без нетривиальных независимых событий

Пусть задано некоторое вероятностное пространство $(\Omega, \mathcal{F}, P)$. Невозможное событие $\emptyset$ и достоверное событие Ω назовем *тривиальными*. События, имеющие вероятность 0 или 1, также назовем тривиальными. Вопрос: в каждом ли вероятностном пространстве существуют нетривиальные независимые события? Более конкретно: всегда ли можно найти два нетривиальных события A и B , которые независимы, т. е. для которых $P(AB) = P(A)P(B)$? Покажем, что это не всегда так.

(а) Пусть пространство Ω конечное, $\Omega = \{\omega_1, \dots, \omega_n\}$, и пусть

$$P(\{\omega_1\}) = 1 - (n-1)\varepsilon, \quad P(\{\omega_2\}) = P(\{\omega_3\}) = \dots P(\{\omega_n\}) = \varepsilon,$$

где ε — произвольное иррациональное число из интервала $\left(0, \frac{1}{n-1}\right)$. Допустим, что в этом пространстве существуют события A и B , которые нетривиальны и независимы. Тогда есть три возможности: 1) $\omega_1 \notin A$, $\omega_1 \notin B$; 2) $\omega_1 \notin A$, $\omega_1 \in B$ или наоборот, $\omega_1 \in A$, $\omega_1 \notin B$; 3) $\omega_1 \in A$, $\omega_1 \in B$.

Проверим, что равенство $P(AB) = P(A)P(B)$ не может выполняться для каждого из случаев 1), 2) и 3). Например, рассмотрим случай 2). Поскольку события A и B нетривиальные, каждое из них содержит по крайней мере по одному элементарному исходу. Пусть A содержит k исходов, взятых из элементов $\omega_2, \dots, \omega_n$, а B содержит ω_1 и еще l исходов, также взятых из $\omega_2, \dots, \omega_n$. Тогда событие AB не содержит элемент ω_1 , и можно предположить, что в AB входят некоторые m из исходов $\omega_2, \dots, \omega_n$. Очевидно, что $1 \leq k \leq n-1$, $0 \leq l \leq n-1$ и $1 \leq m \leq k-1$. Теперь условие независимости A и B записывается так:

$$m\varepsilon = (1 - (n-1)\varepsilon + l\varepsilon)k\varepsilon.$$

Отсюда $\varepsilon = \frac{k-m}{k(n-1-l)}$, т.е. число ε рациональное, что противоречит предположению об иррациональности ε .

К такому же заключению мы приходим и при рассмотрении случаев 1) и 3). Следовательно, описанная модель не содержит нетривиальных и независимых событий.

(б) В предыдущем примере пространство Ω было конечным. Здесь рассмотрим случай, когда Ω — счетное множество. Итак, пусть $\Omega = \{\omega_1, \omega_2, \dots\}$, и пусть

$$P(\{\omega_1\}) = 1 - \sum_{k=2}^{\infty} P(\{\omega_k\}), \quad P(\{\omega_k\}) = 2^{-k!}, \quad k = 2, 3, \dots$$

Как и выше, можно показать, что если $A_1, \dots, A_k$, $k \geq 2$ — нетривиальные события, то они не могут быть независимыми.

4.2. О лемме Бореля–Кантелли

Предположим, что $(\Omega, \mathcal{F}, P)$ — вероятностное пространство, $A_1, A_2, \dots$ — произвольная бесконечная последовательность событий. Каждый элементарный исход $\omega \in \Omega$ принадлежит либо конечному, либо бесконечному числу из этих событий. Обозначим через $\{A_n \text{ б. ч.}\}$ множество тех ω , которые бесконечное число раз встречаются в последовательности $A_1, A_2, \dots$. Для события $\{A_n \text{ б. ч.}\}$, равного $\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$, также используются обозначения A^* и $\overline{\lim} A_n$. Сформулируем теперь *лемму Бореля–Кантелли*, доказательство которой можно найти почти во всех обстоятельных книгах по теории вероятностей (см., например, [9, 60, 106, 116, 117, 145, 157, 177, 179, 330, 345, 404]).

- (а) Если $A_1, A_2, \dots$ — любая последовательность событий, для которой $\sum_{n=1}^{\infty} P(A_n) < \infty$, то $P\{A_n \text{ б. ч.}\} = 0$.
- (б) Если $\sum_{n=1}^{\infty} P(A_n) = \infty$ и события $A_1, A_2, \dots$ независимы в совокупности, то $P\{A_n \text{ б. ч.}\} = 1$.

Итак, лемма Бореля–Кантелли состоит из двух утверждений, и, очевидно, описанные условия достаточны для их справедливости. Вопрос: являются ли эти условия необходимыми? Для того чтобы ответить, рассмотрим пример.

Пусть $\Omega = [0, 1]$, $\mathcal{F} = \mathcal{B}_{[0,1]}$ — борелевская σ -алгебра, порожденная всевозможными подынтервалами отрезка $[0, 1]$, и P — лебеговская мера. Рассмотрим следующую последовательность событий: $A_n = \left[0, \frac{1}{n}\right]$, где $n = 1, 2, \dots$. Очевидно, что события A_n убывают при $n \rightarrow \infty$, и, кроме того, $\{A_n \text{ б. ч.}\} = \bigcap_{n=1}^{\infty} A_n = \emptyset$, так что $P\{A_n \text{ б. ч.}\} = 0$. В то же время $\sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty$. Следовательно, утверждение, обратное к утверждению (а), является неверным, и, значит, условие $\sum_{n=1}^{\infty} P(A_n) < \infty$ не является необходимым для выполнения равенства $P\{A_n \text{ б. ч.}\} = 0$.

Рассмотренный пример позволяет сделать полезный вывод, касающийся и утверждения (б). Действительно, мы видели, что условие $\sum_{n=1}^{\infty} P(A_n) = \infty$ выполнено, и тем не менее, равенство $P\{A_n \text{ б. ч.}\} = 1$ не имеет места. Это означает, что свойство независимости событий $A_1, A_2, \dots$ существенно. В связи с этим заметим, что это свойство не является необходимым. Можно построить последовательность попарно независимых событий $A_1, A_2, \dots$, для которой $\sum_{n=1}^{\infty} P(A_n) = \infty$ и $P\{A_n \text{ б. ч.}\} = 1$ (см. [45, 179]).

4.3. Может ли класс событий обладать свойством независимости и быть исчерпывающим?

Пусть заданы вероятностное пространство $(\Omega, \mathcal{F}, P)$ и класс событий $\mathcal{A} = \{A_i \in \mathcal{F}, i \in \Lambda\}$, где Λ — счетное множество индексов. Класс $\mathcal{A}$ обладает свойством независимости, если для любых k различных индексов $i_1, \dots, i_k \in \Lambda$, $k \geq 2$, выполнено равенство

$$P(A_{i_1} \dots A_{i_k}) = P(A_{i_1}) \dots P(A_{i_k}). \quad (1)$$

Далее, класс $\mathcal{A}$ называется *исчерпывающим*, если

$$P\left(\bigcup_{i \in \Lambda} A_i\right) = 1. \quad (2)$$

Вопрос: может ли одновременно класс $\mathcal{A}$ обладать свойством независимости и быть исчерпывающим? Дадим ответ в двух случаях, а именно, когда число событий, входящих в $\mathcal{A}$, конечно и когда это число бесконечно.

(а) Пусть индексное множество Λ конечно. Предположим, что класс $\mathcal{A} = \{A_i, i \in \Lambda\}$ обладает свойством независимости. Из равенства (1) легко показать, что это свойство имеет место и для класса $\{\bar{A}_i, i \in \Lambda\}$. Следовательно,

$$P\left(\bigcup_{i \in \Lambda} A_i\right) = 1 - P\left(\bigcap_{i \in \Lambda} \bar{A}_i\right) = 1 - \prod_{i \in \Lambda} P(\bar{A}_i). \quad (3)$$

Очевидно, что если $P(\bar{A}_i) > 0$ для всех $i \in \Lambda$, то правая часть равенства (3) будет меньше 1. Но ввиду (2), для исчерпывающего класса левая часть (3) должна быть равна 1. Полученное противоречие показывает, что в этом случае класс $\mathcal{A}$ не может быть исчерпывающим.

Заметим, однако, что если $P(\bar{A}_i) = 0$ для некоторого $j \in \Lambda$, то правая часть (3) равна 1, т. е. класс $\mathcal{A}$ исчерпывающий. Но из $P(\bar{A}_j) = 0$ следует, что $P(A_j) = 1$, т. е. в класс $\mathcal{A}$ должно входить и тривиальное событие Ω (см. пример 4.1). Следовательно, только в этом (тривиальном) случае конечное множество независимых событий может быть исчерпывающим.

(б) Пусть $\Lambda = \mathbb{N}$. Оказывается, что в этом случае можно построить два класса событий, скажем $\mathcal{A}_1$ и $\mathcal{A}_2$, каждый из которых будет обладать свойством независимости, причем $\mathcal{A}_1$ — исчерпывающий класс, а класс $\mathcal{A}_2$ не является исчерпывающим.

Рассмотрим описанное в примере 4.2 (лебеговское) вероятностное пространство $([0, 1], \mathcal{B}_{[0,1]}, P)$. Через A_k обозначим множество тех чисел $x \in [0, 1]$, в двоичном разложении которых на месте с номером k стоит цифра 0. Положим $\mathcal{A}_1 = \{A_1, A_2, \dots\}$. Тогда нетрудно проверить, что класс $\mathcal{A}_1$ обладает свойством независимости и, более того, $P(A_k) = \frac{1}{2}$ для каждого k , $k = 1, 2, \dots$. Таким образом, $\sum_{k=1}^{\infty} P(A_k) = \infty$, и, согласно лемме Бореля–Кантелли (см. пример 4.2), имеем $P\left(\bigcap_{k=1}^{\infty} A_k\right) = 1$. Следовательно, класс $\mathcal{A}_1$ обладает свойством независимости и является исчерпывающим.

Теперь, используя события $A_1, A_2, \dots$ из класса $\mathcal{A}_1$, определим новый класс $\mathcal{A}_2 = \{B_1, B_2, \dots\}$, составленный из событий вида

$$B_i = \bigcap_{j=r}^s A_j, \quad \text{где} \quad r = \frac{1}{2}i(i-1) + 1, \quad s = \frac{1}{2}i(i+1). \quad (4)$$

Очевидно, что событие B_1 означает: первая цифра в двоичном разложении x равна 0; B_2 — вторая и третья цифры — нули; B_3 — следующие три цифры — нули и т. д. Число нулей, о которых идет речь в событии B_i , равно i . Поскольку события $A_1, A_2, \dots$ независимы, из (4) вытекает, что класс $\mathcal{A}_2$ состоит из независимых событий. Кроме того, $P(B_i) = 2^{-i}$, $\sum_{i=1}^{\infty} P(B_i) < \infty$, и, наконец, легко убедиться, что $P\left(\bigcup_{i=1}^{\infty} B_i\right) < 1$. Следовательно, класс $\mathcal{A}_2$ обладает свойством независимости, но не является исчерпывающим.

Г Л А В А 2

РАСПРЕДЕЛЕНИЯ. СЛУЧАЙНЫЕ ВЕЛИЧИНЫ И ИХ ОСНОВНЫЕ ХАРАКТЕРИСТИКИ

§ 5. Функции распределения случайных величин

Пусть $F(x)$, $x \in \mathbb{R}^1$, — функция, удовлетворяющая условиям:

- 1) F — неубывающая функция, т. е. если $x_1 < x_2$, то $F(x_1) \leq F(x_2)$;
- 2) F непрерывна справа, т. е. $F(x) = F(x+) = \lim_{u \downarrow x} F(u)$, $x \in \mathbb{R}^1$;
- 3) $\lim_{x \rightarrow -\infty} F(x) = 0$, $\lim_{x \rightarrow \infty} F(x) = 1$.

Каждая функция F , удовлетворяющая условиям 1)–3), называется *функцией распределения*.

Предположим, что $(\Omega, \mathcal{F}, P)$ — вероятностное пространство. Как обычно, через $\mathcal{B}^1$ обозначим борелевскую σ -алгебру действительной прямой $\mathbb{R}^1$. Функция X , определенная на $(\Omega, \mathcal{F})$ со значениями в $(\mathbb{R}^1, \mathcal{B}^1)$, называется *измеримой функцией* или *случайной величиной*, если для любого $B \in \mathcal{B}^1$ множество $\{\omega: X(\omega) \in B\} \in \mathcal{F}$. Вероятностная мера P_X на $\mathcal{B}^1$, называемая *распределением вероятностей* случайной величины X , задается с помощью равенства $P_X(B) = P(X^{-1}(B))$, $B \in \mathcal{B}^1$. Легко показать, что функция

$$F_X(x) = P_X((-\infty, x]), \quad x \in \mathbb{R}^1,$$

удовлетворяет перечисленным выше условиям 1)–3), и, значит, она является функцией распределения случайной величины X .

Если существует не более чем счетное множество таких чисел $x_1, x_2, \dots$, что $F_X(x_n) - F_X(x_n-) = p_n > 0$ и при этом $\sum_n p_n = 1$, то функция распределения F_X дискретна. В этом случае распределение вероятностей P_X тоже дискретно (это означает, что мера P_X сосредоточена на множестве $\{x_1, x_2, \dots\}$, т. е. $P_X(\{x_n\}) = F_X(x_n) - F_X(x_n-) > 0$, $\sum_n P_X(\{x_n\}) = 1$).

Совокупность $\{p_1, p_2, \dots\}$ называется *дискретным вероятностным распределением*, а X — *дискретной случайной величиной* со значениями

в множестве $\{x_1, x_2, \dots\}$ и с распределением $\{p_1, p_2, \dots\}$. (Очевидно, что $P\{X = x_n\} = p_n$, $n = 1, 2, \dots$)

Мы говорим, что F_X — *абсолютно непрерывная функция распределения*, если существует такая неотрицательная и интегрируемая функция $f(x)$, $x \in \mathbb{R}^1$, что

$$F_X(x) = \int_{-\infty}^x f(u) du, \quad x \in \mathbb{R}^1.$$

Функция f называется *плотностью* функции F_X , а также плотностью случайной величины X и меры P_X . При этом случайная величина X также называется *абсолютно непрерывной*.

(Отметим, что существуют непрерывные распределения, точки роста которых образуют множество нулевой меры Лебега. По поводу этих распределений, называемых сингулярными, читатель может обратиться, например, к книгам [61, 106, 108, 116, 145, 345].)

Многомерные функции распределения могут быть введены следующим образом. Пусть функция $G(x_1, \dots, x_n)$ определена для $(x_1, \dots, x_n) \in \mathbb{R}^n$ и удовлетворяет следующим условиям:

- 1') G — неубывающая функция по каждому из своих аргументов;
- 2') G непрерывна справа по каждому из своих аргументов;
- 3') $G(x_1, \dots, x_n) \rightarrow 0$ при $x_j \rightarrow -\infty$ хотя бы для одного индекса j и $G(x_1, \dots, x_n) \rightarrow 1$, если $x_i \rightarrow \infty$ для всех i ;
- 4) если $a_i \leq b_i$, $i = 1, \dots, n$, и по определению

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

то имеет место неравенство

$$\Delta_{a_1, b_1} \Delta_{a_2, b_2} \dots \Delta_{a_n, b_n} G(x_1, \dots, x_n) \geq 0.$$

Каждая функция, удовлетворяющая условиям 1')–3') и 4), называется *n -мерной функцией распределения*. Уместно сравнить эти условия с условиями 1)–3) в определении одномерной функции распределения. Отметим, что хотя в одномерном случае условие типа 4) выполнено автоматически, в многомерном случае, как будет видно далее, оно существенно.

Напомним следующий важный и хорошо известный результат, называемый теоремой Колмогорова: *если $G(x_1, \dots, x_n)$, $(x_1, \dots, x_n) \in \mathbb{R}^n$, — n -мерная функция распределения, то существует вероятностное пространство $(\Omega, \mathcal{F}, P)$ и определенный на нем случайный вектор $(X_1, \dots, X_n)$, функция распределения которого совпадает с заданной*

функцией G , т. е. $P\{X_1 \leq x_1, \dots, X_n \leq x_n\} = G(x_1, \dots, x_n)$ для всех $(x_1, \dots, x_n) \in \mathbb{R}^n$.

Мы говорим, что функция распределения G и соответствующий ей случайный вектор $(X_1, \dots, X_n)$ абсолютно непрерывны и имеют плотность $g(x_1, \dots, x_n)$, если

$$G(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} g(u_1, \dots, u_n) du_1 \dots du_n$$

для всех $(x_1, \dots, x_n) \in \mathbb{R}^n$. Здесь g — неотрицательная функция, интеграл от которой по всему пространству $\mathbb{R}^n$ равен 1.

Для компоненты X_i случайного вектора $(X_1, \dots, X_n)$ рассмотрим ее функцию распределения $G_i(x_i) = P\{X_i \leq x_i\}$, называемую *маргинальной функцией распределения*. Очевидно, что G_i получается из G после подстановки $x_j = \infty$ для всех $j \neq i$. Если G абсолютно непрерывна, то, проинтегрировав ее плотность $g(x_1, \dots, x_n)$ по всем $x_j \in \mathbb{R}^1$, $j \neq i$, мы получим *маргинальную плотность* $g_i(x_i)$, соответствующую случайной величине X_i .

Понятие дискретной n -мерной функции распределения вводится так же, как и в одномерном случае. Две случайные величины X_1 и X_2 называются независимыми, если для любых борелевских множеств B_1 и B_2 , т. е. для $B_1, B_2 \in \mathcal{B}^1$, выполнено равенство

$$P\{X_1 \in B_1, X_2 \in B_2\} = P\{X_1 \in B_1\}P\{X_2 \in B_2\}.$$

Понятия независимости в совокупности и попарной независимости вводятся так же, как и для случайных событий (см. §3). Действительно, пусть случайные величины $X_1, \dots, X_n$ имеют соответственно функции распределения $G_1, \dots, G_n$, и пусть $G(x_1, \dots, x_n)$ — их совместная функция распределения, т. е. функция распределения случайного вектора $(X_1, \dots, X_n)$. Величины $X_1, \dots, X_n$ называются *независимыми в совокупности* (просто *независимыми*), если

$$G(x_1, \dots, x_n) = G_1(x_1) \dots G_n(x_n) \quad \text{для всех } x_1, \dots, x_n \in \mathbb{R}^1. \quad (1)$$

Если для случайного вектора $(X_1, \dots, X_n)$ выполнено мультипликативное свойство (1), тогда любые $X_{i_1}, \dots, X_{i_k}$, где $1 \leq i_1 < \dots < i_k \leq n$, $k = 2, \dots, n-1$, также удовлетворяют мультипликативному свойству. Поэтому независимость $X_1, \dots, X_n$ есть независимость в совокупности, и достаточно использовать только термин «независимость». Как мы видели в §3, это не так для семейства случайных событий.

В абсолютно непрерывном случае свойство независимости величин $X_1, \dots, X_n$ выражается через соответствующие плотности следующим образом:

$$g(x_1, \dots, x_n) = g_1(x_1) \dots g_n(x_n), \quad x_1, \dots, x_n \in \mathbb{R}^1.$$

Другие понятия, относящиеся к случайным величинам и их распределениям, будут рассмотрены ниже, непосредственно в примерах. Детальное изложение теории вероятностных распределений (1-мерных, n -мерных и т. д.) читатель может найти во многих источниках, см., например, [9, 24, 26, 60, 67, 75, 92, 106, 116, 117, 139, 145, 177, 231, 330, 345, 404, 405].

5.1. Эквивалентные случайные величины распределены одинаково, но обратное неверно

Рассмотрим две случайные величины X и Y , заданные на вероятностном пространстве $(\Omega, \mathcal{F}, P)$, и предположим, что они *эквивалентны*, т. е. $P\{\omega: X(\omega) \neq Y(\omega)\} = 0$. Отсюда вытекает, что

$$F_X(x) = P\{\omega: X(\omega) \leq x\} = P\{\omega: Y(\omega) \leq x\} = F_Y(x) \quad \text{для всех } x \in \mathbb{R}^1.$$

Следовательно, X и Y одинаково распределены. Тот факт, что X и Y имеют одну и ту же функцию распределения, записываем так: $X \stackrel{d}{=} Y$.

Оказывается, однако, что из свойства $X \stackrel{d}{=} Y$ не следует эквивалентность X и Y . Действительно, возьмем случайную величину X , которая абсолютно непрерывна и, кроме того, симметрична относительно 0, т. е. для ее плотности f имеет место равенство $f(-x) = f(x)$, $x \in \mathbb{R}^+$. Пусть $Y = -X$. Из симметричности X легко вытекает, что $F_X = F_Y$, т. е. что $X \stackrel{d}{=} Y$. Поскольку X абсолютно непрерывна, каждое ее значение x имеет нулевую вероятность: $P\{\omega: X(\omega) = x\} = 0$. Таким образом,

$$P\{\omega: X(\omega) = Y(\omega)\} = P\{\omega: X(\omega) = -X(\omega)\} = P\{\omega: X(\omega) = 0\} = 0.$$

Это означает, что случайные величины X и Y не являются эквивалентными, несмотря на то, что они одинаково распределены.

Легко построить аналогичный пример и в дискретном случае. Пусть X — дискретная случайная величина, симметричная относительно 0 и такая, что $P\{X = 0\} < 1$. (Если $P\{X = 0\} = 1$, все тривиально.) Можно считать, что X принимает значения в конечном или счетном числовом множестве $\{\dots, -x_2, -x_1, x_0 = 0, x_1, x_2, \dots\}$ с вероятностями $p_0 = P\{X = 0\}$, $p_j = P\{X = x_j\} = P\{X = -x_j\}$, $j = 1, 2, \dots$, где $0 < x_1 < x_2 < \dots$, и $p_0 + 2 \sum_j p_j = 1$. Определим Y как $-X$. Тогда $X \stackrel{d}{=} Y$, но

$P\{\omega: X(\omega) \neq Y(\omega)\} = P\{X(\omega) = 0\} = p_0 < 1$, т. е. X и Y — неэквивалентные случайные величины.

5.2. Если случайные величины X , Y и Z заданы на одном вероятностном пространстве, то из $X \stackrel{d}{=} Y$ не следует, что $XZ \stackrel{d}{=} YZ$

Напомним, что если $X \stackrel{d}{=} Y$, а функция $h(x)$, $x \in \mathbb{R}^1$, — $\mathcal{B}^1$ -измерима, то $h(X)$ и $h(Y)$ также являются случайными величинами, причем $h(X) \stackrel{d}{=} h(Y)$. Это может послужить мотивом для следующей гипотезы:

$$X \stackrel{d}{=} Y \implies XZ \stackrel{d}{=} YZ \quad \text{для любой случайной величины } Z.$$

Простой пример показывает, что в общем случае это не так.

Пусть случайная величина X симметрична (см. пример 5.1), и пусть $Y = -X$. Тогда $X \stackrel{d}{=} Y$. Возьмем $Z = Y$, т. е. $Z = -X$. Тогда $XZ = -X^2$, а $YZ = X^2$. Таким образом, величина XZ принимает неположительные значения, в то время как YZ принимает неотрицательные значения. Следовательно, равенство $XZ \stackrel{d}{=} YZ$ не может иметь места. Тривиальный случай $P\{X = 0\} = 1$ интереса не представляет.

5.3. Функция, которая является метрикой в пространстве распределений, но не является метрикой в пространстве случайных величин

Определим расстояние $r(X, Y)$ между двумя случайными величинами X и Y следующим образом:

$$r(X, Y) = \sup_{x \in \mathbb{R}^1} |P\{X \leq x\} - P\{Y \leq x\}| = \sup_{x \in \mathbb{R}^1} |F_X(x) - F_Y(x)|.$$

Функция r — так называемое *расстояние Колмогорова*, или *равномерное расстояние*. Вместо $r(X, Y)$ иногда используется обозначение $r(F_X, F_Y)$, где F_X и F_Y — функции распределения соответственно X и Y . Функция r является метрикой в пространстве всех распределений на $\mathbb{R}^1$. Действительно, легко видеть, что:

- 1) $r(X, Y) \geq 0$ и $r(X, Y) = 0 \iff X \stackrel{d}{=} Y$;
- 2) $r(X, Y) = r(Y, X)$;
- 3) $r(X, Y) \leq r(X, Z) + r(Y, Z)$.

В этих условиях X, Y, Z — произвольные случайные величины.

Рассмотрим теперь r как функцию, заданную на пространстве всех случайных величин, определенных на некотором фиксированном пространстве $(\Omega, \mathcal{F}, P)$. Обращаясь к примеру 5.1, легко заключить, что r

не является метрикой, потому что из условия $r(X, Y) = 0$ не вытекает, что $X \stackrel{\text{п.н.}}{=} Y$. (Здесь и далее запись $X \stackrel{\text{п.н.}}{=} Y$ означает, что X и Y — эквивалентные случайные величины, или, как еще говорят, что $X = Y$ п. н.)

5.4. Об n -мерных распределениях

В определение n -мерной функции распределения входит условие 4) (см. с. 47). В одномерном случае ($n = 1$) оно выполнено автоматически. Однако если $n \geq 2$, то условие 4) уже не вытекает из остальных условий. Действительно, пусть $n = 2$, и пусть

$$F(x, y) = \begin{cases} 0, & \text{если } x \leq 0, \text{ или } x + y \leq 1, \text{ или } y \leq 0; \\ 1 & \text{в остальных случаях.} \end{cases}$$

Легко проверить, что функция $F(x, y)$, $x, y \in \mathbb{R}^2$, удовлетворяет условиям 1'–3'). Предположим, что F — двумерная функция распределения, и что она является функцией распределения случайного вектора (X, Y) . Тогда для любого прямоугольника $R = [a_1, b_1] \times [a_2, b_2]$, где $a_1 < b_1$, $a_2 < b_2$, должно выполняться неравенство $P\{(X, Y) \in R\} \geq 0$. Однако, если взять $R = \left[\frac{1}{3}, 1\right] \times \left[\frac{1}{3}, 1\right]$, то

$$P\{(X, Y) \in R\} = F(1, 1) - F\left(1, \frac{1}{3}\right) - F\left(\frac{1}{3}, 1\right) + F\left(\frac{1}{3}, \frac{1}{3}\right) = -1,$$

что невозможно, и, значит, функция F не удовлетворяет условию 4). Следовательно, как и можно было ожидать, условия 1'–3') недостаточны для того, чтобы F была n -мерной функцией распределения. Еще раз отметим, что этот факт имеет место для всех $n \geq 2$.

Приведем еще один пример. Пусть

$$G(x, y) = \begin{cases} 0, & \text{если } x \leq 0 \text{ или } y \leq 0; \\ \min(1, \max(x, y)) & \text{в остальных случаях.} \end{cases}$$

Можно проверить, что G удовлетворяет каждому из условий 1')–3'), но условие 4) не выполнено. (Достаточно взять прямоугольник $R = [x_1, x_2] \times [y_1, y_2]$, где $0 \leq x_1 \leq y_1 < x_2 \leq y_2 \leq 1$, и вычислить вероятность $P\{(X, Y) \in R\}$.)

5.5. Отсутствие атомов не влечет непрерывности многомерных функций распределения

Пусть на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ задана случайная величина X , и $F(x)$, $x \in \mathbb{R}^1$, — ее функция распределения. Предположим,

что X принимает значения в некотором (конечном или бесконечном) интервале в $\mathbb{R}^1$, и что для каждого x из этого интервала выполнено равенство $P\{\omega: X(\omega) = x\} = 0$. В этом случае говорят, что распределение вероятностей P_X не имеет атомов. Отсюда вытекает, что функция распределения F непрерывна всюду в том же интервале. Вопрос: имеет ли место аналогичное свойство в многомерном случае? Иначе говоря, если X есть случайный вектор со значениями в $\mathbb{R}^n$ и $P\{\omega: X(\omega) = x\} = 0$ для всех $x \in \mathbb{R}^n$, то непрерывна ли функция распределения этого вектора?

Конкретный пример для случая $n = 2$ показывает, что ответ отрицателен. Действительно, рассмотрим функцию

$$F(x, y) = \begin{cases} xy, & \text{если } 0 \leq x \leq 1, 0 \leq y \leq \frac{1}{2}; \\ \frac{x}{2}, & \text{если } 0 \leq x \leq 1, \frac{1}{2} < y < \infty; \\ y, & \text{если } 1 < x < \infty, 0 \leq y \leq 1; \\ 1, & \text{если } x > 1, y > 1; \\ 0 & \text{в остальных случаях.} \end{cases}$$

Нетрудно проверить, что F — двумерная функция распределения.

Обозначим через (X, Y) случайный вектор с этой функцией распределения. Для дальнейшего рассуждения удобно воспользоваться рис. 1, где в разных подобластях плоскости $\mathbb{R}^2$ указаны соответствующие значения F . Заметим, что значения вектора (X, Y) — это точки первого квадранта $Q = \{(x, y): 0 \leq x < \infty, 0 \leq y < \infty\}$. Более того, каждая точка $(x, y) \in Q$ имеет нулевую вероятность: $P\{(X, Y) = (x, y)\} = 0$. Можно предположить, что если распределение вероятностей $P_{X,Y}$ не имеет атомов, то функция распределения F непрерывна всюду в Q . Покажем, что это не так. Для этого введем обозначение:

$$\Delta(1, y) = F(1, y) - F(1-, y-).$$

Очевидно, что $\Delta(1, y)$ — величина скачка функции F в точке $(1, y)$. Явный вид F позволяет легко найти, что $\Delta(1, y) = y - \frac{1}{2}$, если $\frac{1}{2} < y < 1$, и $\Delta(1, y) = \frac{1}{2}$, если $1 \leq y < \infty$. Следовательно, каждая точка с координатами $(1, y)$, где $\frac{1}{2} < y < \infty$, является точкой разрыва F . Разрывность F есть следствие того факта, что $P\{X = 1, \frac{1}{2} \leq Y \leq 1\} = \frac{1}{2}$, т. е. что в плоскости Oxy имеется интервал с положительной вероятностью (см. жирный интервал на рис. 1).

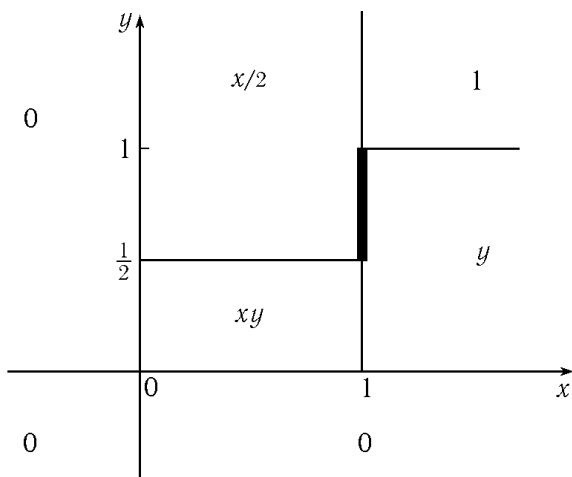


Рис. 1.

5.6. Об абсолютной непрерывности распределений случайного вектора и его компонент

Для простоты рассмотрим двумерный случай. Предположим, что случайный вектор (X_1, X_2) имеет абсолютно непрерывное распределение. Тогда каждая из компонент X_1 и X_2 также имеет абсолютно непрерывное распределение. Вопрос: верно ли обратное утверждение?

Чтобы ответить, возьмем любую абсолютно непрерывную случайную величину X_1 и положим $X_2 \equiv X_1$, т. е. $X_2(\omega) = X_1(\omega)$ для каждого $\omega \in \Omega$. Тогда X_2 также абсолютно непрерывна. Допустим, что вектор (X_1, X_2) абсолютно непрерывен, и имеет, например, плотность $f(x_1, x_2)$, $(x_1, x_2) \in \mathbb{R}^2$. Отсюда следует, что для любого множества $B \subset \mathbb{R}^2$ выполняется равенство

$$P\{(X_1, X_2) \in B\} = \iint_B f(x_1, x_2) dx_1 dx_2. \quad (1)$$

Заметим, что все значения вектора (X_1, X_2) лежат на прямой l , задаваемой уравнением $x_2 = x_1$. Если выбрать $B = l = \{(x_1, x_2) : x_2 = x_1\}$, то левая часть (1) равна 1, а правая часть (1) равна 0, потому что любая прямая на плоскости имеет нулевую меру Лебега. Полученное противоречие показывает, что в общем случае из абсолютной непрерывности компонент случайного вектора не вытекает абсолютная непрерывность самого вектора. Если, однако, компоненты вектора — независимые случайные величины, то ответ на поставленный вопрос положителен.

5.7. Многомерные функции распределения не определяются однозначно своими маргинальными распределениями

Если случайный вектор $(X_1, \dots, X_n)$ имеет функцию распределения $F(x_1, \dots, x_n)$, $(x_1, \dots, x_n) \in \mathbb{R}^n$, то маргинальные функции распределения $F_k(x_k) = P\{X_k \leq x_k\}$, $k = 1, \dots, n$, определяются однозначно. Покажем на конкретных примерах, что обратное утверждение неверно. Ограничимся двумерным случаем.

(а) Рассмотрим двумерное дискретное распределение вероятностей $\mathcal{P} = \{p_{ij} : i, j = 1, 2, \dots\}$. Выберем две точки, (x_1, y_1) и (x_2, y_2) , вероятности которых — соответственно, p_{11} и p_{22} , — положительны. При этом можно считать, что $x_1 \neq x_2$ и $y_1 \neq y_2$. Пусть ε — произвольное число из интервала $(0, \varepsilon_0)$, где $\varepsilon_0 = \min(p_{11}, p_{22})$. Рассмотрим совокупность $\tilde{\mathcal{P}} = \{q_{ij}\}$, где

$$\begin{aligned} q_{11} &= p_{11} - \varepsilon, & q_{12} &= p_{12} + \varepsilon, & q_{21} &= p_{21} + \varepsilon, & q_{22} &= p_{22} - \varepsilon; \\ q_{ij} &= p_{ij} & \text{для всех } i, j &\neq 1, 2. \end{aligned}$$

Легко проверить, что $\tilde{\mathcal{P}}$ — двумерное вероятностное распределение. Более того, при каждом ε , выбранном как указано выше, маргинальные распределения, соответствующие $\mathcal{P}$, совпадают с маргинальными распределениями $\tilde{\mathcal{P}}$. Очевидно, что $\mathcal{P} \neq \tilde{\mathcal{P}}$.

(б) Обозначая отрезок $[-1, 1]$ через I , рассмотрим функции

$$\begin{aligned} f(x_1, x_2) &= \begin{cases} \frac{1}{4}(1 + x_1 x_2), & \text{если } (x_1, x_2) \in I \times I, \\ 0, & \text{если } (x_1, x_2) \notin I \times I, \end{cases} \\ g(x_1, x_2) &= \begin{cases} \frac{1}{4}, & \text{если } (x_1, x_2) \in I \times I, \\ 0, & \text{если } (x_1, x_2) \notin I \times I. \end{cases} \end{aligned}$$

Непосредственно проверяется, что f и g — двумерные плотности. Соответствующие маргинальные плотности выглядят так:

$$\begin{aligned} f_1(x_1) &= \begin{cases} \frac{1}{4}, & \text{если } x_1 \in I, \\ 0, & \text{если } x_1 \notin I, \end{cases} & f_2(x_2) &= \begin{cases} \frac{1}{4}, & \text{если } x_2 \in I, \\ 0, & \text{если } x_2 \notin I, \end{cases} \\ g_1(x_1) &= \begin{cases} \frac{1}{4}, & \text{если } x_1 \in I, \\ 0, & \text{если } x_1 \notin I, \end{cases} & g_2(x_2) &= \begin{cases} \frac{1}{4}, & \text{если } x_2 \in I, \\ 0, & \text{если } x_2 \notin I. \end{cases} \end{aligned}$$

Таким образом, $f_1 = g_1$, $f_2 = g_2$, но очевидно, что $f \neq g$.

(в) Предположим, что функции распределения F_1 и F_2 имеют плотности f_1 и f_2 соответственно. Положим для всех $(x_1, x_2) \in \mathbb{R}^2$

$$f(x_1, x_2) = f_1(x_1)f_2(x_2)[1 + \varepsilon(2F_1(x_1) - 1)(2F_2(x_2) - 1)],$$

где ε — произвольное число, $|\varepsilon| \leq 1$. Нетрудно проверить, что $f(x_1, x_2)$, $(x_1, x_2) \in \mathbb{R}^2$, является двумерной плотностью, а также, что маргинальные плотности, соответствующие f , равны f_1 и f_2 .

Пусть теперь у функций распределения F_1 и F_2 нет плотностей. В этом случае рассмотрим семейство функций

$$F(x_1, x_2) = F_1(x_1)F_2(x_2)[1 + \varepsilon(1 - F_1(x_1))(1 - F_2(x_2))], \quad (x_1, x_2) \in \mathbb{R}^2,$$

где, как и выше, $|\varepsilon| \leq 1$. Легко установить, что (при любом ε) F — двумерная функция распределения, маргинальные распределения которой совпадают с F_1 и F_2 .

Рассмотренные примеры показывают, что различные многомерные распределения могут иметь одни и те же маргинальные распределения. Другими словами, маргинальные распределения не определяют однозначно многомерное распределение. Одно из исключений, когда имеем дело с независимыми случайными величинами (случай $\varepsilon = 0$). Тогда для всех $x_1, x_2 \in \mathbb{R}^1$ верно равенство $F(x_1, x_2) = F(x_1)F(x_2) = F_1(x_1)F_2(x_2)$, а в абсолютно непрерывном случае $f(x_1, x_2) = f_1(x_1)f_2(x_2)$.

5.8. О свертках функций распределений

Пусть $F(x)$, $x \in \mathbb{R}^1$, — произвольная функция распределения. С помощью равенства

$$G(x) = \int_{-\infty}^{\infty} F(x - u) dF(u)$$

определяем новую функцию $G(x)$, $x \in \mathbb{R}^1$, являющуюся *сверткой* F с самой собой. Нетрудно показать, что G также является функцией распределения. Поэтому можно задать следующий вопрос: если свертка $G = F * F$ есть функция распределения, то следует ли отсюда, что сама F — тоже функция распределения? Может быть, это несколько неожиданно, но ответ отрицателен. Чтобы убедиться в этом, рассмотрим функцию $F(x)$, $x \in \mathbb{R}^1$, такую что в промежутках

$$(-\infty, 1); \quad [1, 2); \quad [2, 3); \quad [3, 4); \quad [4, 5); \quad [5, \infty)$$

она принимает соответственно значения

$$0; \quad 0,26; \quad 0,52; \quad 0,48; \quad 0,74; \quad 1.$$

Очевидно, что нарушается только одно из условий, входящих в определение функции распределения, — это условие неубывания. Значит, функция F не является функцией распределения. Прямое вычисление (детали оставляем читателю) показывает, что свертка $G = F * F$ тем не менее есть функция распределения. Отметим, что, по-видимому, представляет интерес найти ответ на поставленный выше вопрос, но в терминах плотностей и их свертки.

5.9. О несохранении свойства унимодальности при свертках унимодальных распределений

Напомним сначала определение понятия унимодальности для абсолютно непрерывных и для дискретных распределений.

Пусть функция распределения $F(x)$, $x \in \mathbb{R}^1$, абсолютно непрерывна с плотностью $f(x)$, $x \in \mathbb{R}^1$. Если существует число x_0 такое, что $f(x)$ не убывает при $x \leq x_0$ и не возрастает при $x \geq x_0$, то f (а также F) называется *унимодальной*. В этом случае точка x_0 является точкой локального максимума функции f и называется *модой*. Легко видеть, что в этом случае $F(x)$ — выпуклая функция при $x < x_0$ и вогнутая — при $x > x_0$.

Пусть далее $\mathcal{P} = \{p_n, n \in \tilde{\mathbb{N}}\}$ — дискретное вероятностное распределение на множестве целых чисел $\tilde{\mathbb{N}}$ или на некотором его подмножестве. (Конечно, $p_n > 0$ и $\sum_n p_n = 1$.) Распределение $\mathcal{P}$ называется *унимодальным*, если существует такое целое $k_0 \in \tilde{\mathbb{N}}$, что числовая последовательность $\{p_k\}$ не убывает при $k \leq k_0$ и не возрастает при $k > k_0$.

Хорошо известно, что операция свертки очень важна для теории вероятностей, поскольку именно через свертки выражаются распределения сумм независимых случайных величин. Поэтому важно знать, сохраняются ли при этой операции некоторые свойства распределений. Рассмотрим вопрос о сохранении свойства унимодальности. Покажем на конкретных примерах, что в общем случае при свертках как дискретных, так и абсолютно непрерывных распределений свойство унимодальности не сохраняется.

(а) Пусть ξ — случайная величина со значениями в множестве $\{0, 1, \dots, m\}$ и с распределением $\mathcal{P}_\xi = \{p_i, i = 0, 1, \dots, m\}$, где $p_0 = \frac{m+2}{2m+2}$, $p_1 = p_2 = \dots = p_m = \frac{1}{2m+2}$. Пусть далее случайная величина η не зависит от ξ , принимает значения в множестве $\{0, 1, \dots, n\}$ и имеет распределение $\mathcal{P}_\eta = \{q_j, j = 0, 1, \dots, n\}$, где $q_0 = \frac{n+2}{2n+2}$, $q_1 = q_2 = \dots = q_n = \frac{1}{2n+2}$.

Легко видеть, что распределения $\mathcal{P}_\xi$ и $\mathcal{P}_\eta$ унимодальны.

Рассмотрим случайную величину $\theta = \xi + \eta$. Очевидно, что θ принимает значения в множестве $\{0, 1, \dots, m+n\}$, а ее распределение $\mathcal{P}_\theta = \{r_k, k = 0, 1, \dots, m+n\}$ ввиду независимости ξ и η есть свертка $\mathcal{P}_\xi$ и $\mathcal{P}_\eta$, т. е. $\mathcal{P}_\theta = \mathcal{P}_\xi * \mathcal{P}_\eta$. Это означает, что для $k = 0, 1, \dots, m+n$

$$r_k = P\{\theta = k\} = \sum_{i=0}^{m \wedge k} P\{\xi = i, \eta = k-i\} = \sum_{i=0}^{m \wedge k} p_i q_{k-i},$$

где, как обычно, $m \wedge k = \min(m, k)$. Элементарное вычисление показывает, что

$$r_0 = \frac{(m+2)(n+2)}{(2m+2)(2n+2)}, \quad r_1 = \frac{m+n+4}{(2m+2)(2n+2)}, \quad r_2 = \frac{m+n+5}{(2m+2)(2n+2)}.$$

Непосредственно проверяется, что $r_0 > r_1$, а $r_1 < r_2$. Этого достаточно, чтобы заключить: распределение $\mathcal{P}_\theta$ унимодальным не является вопреки тому, что $\mathcal{P}_\theta$ есть свертка двух распределений $\mathcal{P}_\xi$ и $\mathcal{P}_\eta$, каждое из которых унимодально.

В частном случае $m = n$ имеет место равенство $\xi \stackrel{d}{=} \eta$, и если обозначить $\mathcal{P}_\xi = \mathcal{P}_\eta = \mathcal{P}$, то $\mathcal{P} * \mathcal{P}$, т. е. свертка $\mathcal{P}$ с собой не сохраняет свойство унимодальности.

Отметим, наконец, следующий результат (см. [201, 482]): *если $\mathcal{P}_\xi$ и $\mathcal{P}_\eta$ — унимодальные и симметричные распределения, то их свертка $\mathcal{P}_\xi * \mathcal{P}_\eta$ задает распределение, которое является унимодальным и симметричным.*

(б) Рассмотрим теперь абсолютно непрерывный случай. Для этого введем следующую функцию:

$$f(x) = \begin{cases} 0, & \text{если } x < -\frac{1}{30} \text{ или } x > \frac{5}{6}, \\ 5, & \text{если } -\frac{1}{30} \leq x \leq 0, \\ 1, & \text{если } 0 < x \leq \frac{5}{6}. \end{cases}$$

Легко проверить, что f есть плотность некоторой функции распределения F , соответствующей случайной величине, X . Более того, F унимодальна, а ее модой может быть любая точка из отрезка $\left[-\frac{1}{30}, 0\right]$.

Пусть X_1 и X_2 — независимые случайные величины с той же плотностью f , что и X . Тогда плотность распределения суммы $X_1 + X_2$ равна свертке f с самой собой, т. е. равна f^{*2} , где

$$f^{*2}(x) = (f * f)(x) = \int_{-\infty}^{\infty} f(x-u)f(u) du.$$

Легко найти f^{*2} :

$$f^{*2}(x) = \begin{cases} 0, & \text{если } x \leq -\frac{1}{15} \text{ или } x > \frac{5}{3}, \\ 25x + \frac{5}{3}, & \text{если } -\frac{1}{15} < x \leq -\frac{1}{30}, \\ -15x + \frac{1}{3}, & \text{если } -\frac{1}{30} < x \leq 0, \\ x + \frac{1}{3}, & \text{если } 0 < x \leq \frac{4}{5}, \\ -9x + \frac{25}{3}, & \text{если } \frac{4}{5} < x \leq \frac{5}{6}, \\ \frac{5}{3} - x, & \text{если } \frac{5}{6} < x \leq \frac{5}{3}. \end{cases}$$

Отсюда видно, что свертка f^{*2} — непрерывная функция, имеющая два локальных максимума и один локальный минимум. Локальные максимумы достигаются в точках $x = -\frac{1}{30}$ и $x = \frac{4}{5}$, причем $f^{*2}\left(-\frac{1}{30}\right) = \frac{5}{6}$ и $f^{*2}\left(\frac{4}{5}\right) = \frac{17}{15}$, а локальный минимум достигается при $x = 0$ и $f^{*2}(0) = \frac{1}{3}$. Следовательно, свойство унимодальности не всегда сохраняется при операции свертки, примененной к унимодальным, абсолютно непрерывным распределениям.

(в) Свойство несохранения унимодальности при свертке унимодальной плотности с собой проиллюстрируем еще одним примером.

Пусть f_a и f_b — равномерные плотности соответственно в интервалах $(0, a)$ и $(0, b)$, где $0 < a < b$. Рассмотрим функцию

$$g(x) = \varepsilon f_a(x) + (1 - \varepsilon) f_b(x), \quad x \in \mathbb{R}^1,$$

где $\varepsilon \in (0, 1)$. Легко видеть, что g является плотностью (сосредоточенной на интервале $(0, b)$), причем g унимодальна с модой в точке $x = 0$. (Более того, мода для g — это любая точка интервала $(0, a)$.) Рассмотрим свертку $h(x) = (g * g)(x) = g^{*2}(x)$ и выясним, обладает ли h свойством унимодальности. Находим, что

$$h(x) = \varepsilon^2 f_a^{*2}(x) + 2\varepsilon(1 - \varepsilon)(f_a * f_b)(x) + (1 - \varepsilon)^2 f_b^{*2}(x).$$

Заметим, что f_a^{*2} имеет форму равнобедренного треугольника с основанием $[0, 2a]$ и высотой $\frac{1}{a}$ (это так называемое треугольное распределение). Поэтому имеют место неравенства:

$$h(a) > \frac{\varepsilon^2}{a}, \quad h(b) > \frac{(1 - \varepsilon)^2}{b}, \quad \int_0^{2a} h(x) dx > \frac{(1 - \varepsilon)^2}{2}.$$

Предположим, что $h(x) > h(a)$ для $x \in (a, b)$. Тогда должны выполняться соотношения

$$\int_a^b h(x) dx \geq (b-a)h(a) > \frac{\varepsilon^2 b}{a} - \varepsilon^2,$$

из которых следовало бы, что

$$\int_a^{2b} h(x) dx > \frac{\varepsilon^2 b}{a} - \varepsilon^2 + \frac{(1-\varepsilon)^2}{2}.$$

Однако, если взять $\varepsilon = \frac{1}{4}$, $a = \frac{1}{16}$ и $b = 1$, то правая часть в этом неравенстве имеет значение большее 1, что невозможно. Это означает, что наше предположение о функции h , т. е. о том, что $h(x) > h(a)$ при $x \in (a, b)$, неверно при выбранных ε, a и b . Другими словами, существует точка x_0 , расположенная между точками a и b , такая что $h(x_0) < h(a)$ и $h(x_0) > h(b)$. Отсюда вытекает, что свертка h не является унимодальной, что и требовалось показать.

Здесь уместно процитировать следующий важный и интересный результат (см. [76, 201]): *пусть вероятностные плотности f и g унимодальны и симметричны. Тогда их свертка $f * g$ унимодальна.* Добавим, однако, что если f унимодальна и симметрична, а g унимодальна и несимметрична, то может оказаться, что их свертка $f * g$ не унимодальна (см. [201]).

§ 6. Математическое ожидание

Пусть случайная величина X задана на вероятностном пространстве $(\Omega, \mathcal{F}, P)$. Одна из самых важных характеристик X — это ее *математическое ожидание* (среднее значение), обозначаемое EX . Если $X \geq 0$ и $P\{X = \infty\} > 0$, то полагаем $EX = \infty$, а если $P\{X = \infty\} = 0$ (что эквивалентно $P\{X < \infty\} = 1$), мы определяем

$$EX = \lim_{n \rightarrow \infty} \sum_{k=1}^{\infty} \frac{k}{2^n} P\left\{\frac{k-1}{2^n} < X \leq \frac{k}{2^n}\right\}. \quad (1)$$

Для любой случайной величины X обозначим $X^+ = \max(X, 0)$ и $X^- = \max(-X, 0)$. Так как X^+ и X^- — неотрицательные случайные величины, то их математические ожидания $E(X^+)$ и $E(X^-)$ могут быть получены из (1), а если $E(X^+) < \infty$ или $E(X^-) < \infty$, то

$$EX = E(X^+) - E(X^-).$$

Математическим ожиданием EX называют также интеграл Лебега от $\mathcal{F}$ -измеримой функции X по вероятностной мере P .

Мы говорим, что случайная величина X имеет конечное математическое ожидание, если $E(X^+) < \infty$ и $E(X^-) < \infty$. Поскольку $|X| = X^+ + X^-$, то конечность EX эквивалентна условию $E|X| < \infty$. В этом случае X называется *интегрируемой случайной величиной*. Если X — абсолютно непрерывная случайная величина с плотностью f , то X интегрируема тогда и только тогда, когда $\int_{-\infty}^{\infty} |x|f(x) dx < \infty$ и $EX = \int_{-\infty}^{\infty} xf(x) dx$.

Если X дискретна и $P\{X = x_n\} = p_n$, $n = 1, 2, \dots$, то интегрируемость X эквивалентна условию $\sum_n |x_n|p_n < \infty$, и в этом случае $EX = \sum_n x_n p_n$.

Напомним также, что интеграл от X по некоторому множеству $A \in \mathcal{F}$ определяется так: $\int_A X dP = \int_{\Omega} X(\omega) I_A(\omega) dP(\omega)$. Через $\mathcal{L}^p = \mathcal{L}^p(\Omega, \mathcal{F}, P)$ ($p > 0$) обозначается *пространство всех p -интегрируемых случайных величин*, т. е. $X \in \mathcal{L}^p \Leftrightarrow E(|X|^p) < \infty$.

Для каждой случайной величины X вводятся и величины (при условии, что они существуют)

$$E\{(X - c)^k\}, \quad E\{|X - c|^k\}, \quad k = 1, 2, \dots, \quad c \in \mathbb{R}^1,$$

называемые соответственно *моментом порядка k* и *абсолютным моментом порядка k относительно константы c* . Если $c = EX$, эти моменты называются *центральными*, а если $c = 0$ — *начальными*. Начальный момент X порядка k обозначается через $m_k = E\{X^k\}$. При $k = 2$ и $c = EX$ получаем *дисперсию* X , обозначаемую обычно через $DX = E\{(X - EX)^2\}$.

Математическое ожидание обладает рядом интересных свойств, среди которых мы чаще всего будем использовать свойство монотонности ($X_1 \leq X_2 \Rightarrow EX_1 \leq EX_2$) и свойство линейности ($E\{c_1 X_1 + c_2 X_2\} = c_1 EX_1 + c_2 EX_2$, $c_1, c_2 \in \mathbb{R}^1$). Эти свойства имеют место только для интегрируемых случайных величин.

Семейство случайных величин $\{X_n, n \geq 1\}$ называется *равномерно интегрируемым*, если

$$\sup_n \int_{|X_n| \geq a} |X_n| dP \rightarrow 0 \quad \text{при} \quad a \rightarrow \infty,$$

или, эквивалентно, если $\sup_n E\{|X_n| I_{[|X_n| \geq a]}\} \rightarrow 0$ при $a \rightarrow \infty$. (Символом I_A обозначаем индикатор множества $A \in \mathcal{F}$, т. е. $I_A(\omega) = 1$, если $\omega \in A$, и $I_A(\omega) = 0$, если $\omega \notin A$.)

Предположим теперь, что кроме случайной величины X , заданной на $(\Omega, \mathcal{F}, P)$, выделена некоторая σ -алгебра $\mathcal{D} \subset \mathcal{F}$. Для интегрируемой случайной величины X можно определить величину $E\{X|\mathcal{D}\}$, называемую *условным математическим ожиданием* X относительно $\mathcal{D}$. Напомним, что $E\{X|\mathcal{D}\}$ определяется как $\mathcal{D}$ -измеримая случайная величина такая, что для всех $A \in \mathcal{D}$

$$\int_A E\{X|\mathcal{D}\} dP = \int_A X dP \quad (P\text{-п. н.}).$$

Из теоремы Радона–Никодима (см. [60, 116, 117, 145, 157]) вытекает существование $E\{X|\mathcal{D}\}$ с точностью до эквивалентности. Перечислим несколько свойств условных математических ожиданий, которые будут использованы далее. Предположим, что X, X_1, X_2 и Y — интегрируемые случайные величины, а c, c_1, c_2 — действительные числа. Тогда:

- (1) $P\{X = c\} = 1 \implies E\{X|\mathcal{D}\} = c$ п. н.;
- (2) $X_1 \leq X_2 \implies E\{X_1|\mathcal{D}\} \leq E\{X_2|\mathcal{D}\}$ п. н. (монотонность);
- (3) $E\{c_1 X_1 + c_2 X_2|\mathcal{D}\} = c_1 E\{X_1|\mathcal{D}\} + c_2 E\{X_2|\mathcal{D}\}$ п. н. (линейность);
- (4) $E\{E\{X|\mathcal{D}\}\} = EX$;
- (5) если $\mathcal{D}_1 \subseteq \mathcal{D}_2 \subset \mathcal{F}$, то $E\{E\{X|\mathcal{D}_2\}|\mathcal{D}_1\} = E\{X|\mathcal{D}_1\}$ п. н.;
- (6) если X и σ -алгебра $\mathcal{D}$ независимы, то $E\{X|\mathcal{D}\} = EX$ п. н.;
- (7) если X — $\mathcal{D}$ -измерима и $E(|XY|) < \infty$, то

$$E\{XY|\mathcal{D}\} = X E\{Y|\mathcal{D}\} \text{ п. н.}$$

Если $X = I_A$, где $A \in \mathcal{F}$, то условное математическое ожидание $E\{I_A|\mathcal{D}\}$ равно условной вероятности $P(A|\mathcal{D})$ события A относительно σ -алгебры $\mathcal{D}$ (см. также пример 2.4).

Рассмотренные далее примеры касаются разных свойств математических ожиданий (безусловных и условных) и раскрывают роль этой характеристики при изучении свойств случайных величин и их распределений.

6.1. О линейности математического ожидания

Если речь идет о математических ожиданиях, то, вообще говоря, принято считать, что

$$E\{X + Y\} = EX + EY \quad \text{и} \quad E\{X + Y + Z\} = EX + EY + EZ.$$

Подобные соотношения, когда они имеют место, выражают *свойство линейности* математического ожидания. Смысл этого свойства

заключается в том, что значение $E\{\cdot\}$ зависит от величин в скобках $\{\cdot\}$ только через их индивидуальные (т. е. через их маргинальные) распределения.

Напомним, что свойство линейности для двух случайных величин имеет место, если, например, EX , EY и $E\{X + Y\}$ существуют. Здесь полезно напомнить следующий более общий результат (см. [436]): *если $E\{X + Y\}$ определено в том смысле, что $E\{(X + Y)^+\}$ и (или) $E\{(X + Y)^-\}$ конечны, то значение $E\{X + Y\}$ зависит от величин X и Y только через их маргинальные распределения.* Для доказательства используются метод усечения и лемма Фату (см. [116, 177]).

Поставим теперь следующий вопрос: имеет ли место свойство линейности при рассмотрении трех величин? Конкретный пример даст нам возможность ответить на этот вопрос.

Пусть случайная величина ξ распределена равномерно на отрезке $[0, 1]$. Тогда случайные величины $1 - \xi$ и $\eta = |2\xi - 1|$ имеют то же самое распределение, что и ξ . Определим три новые величины, притом двумя разными способами:

$$(I) \quad X = Y = \operatorname{tg} \frac{\pi\xi}{2}, \quad Z = -2X;$$

$$(II) \quad X' = \operatorname{tg} \frac{\pi\xi}{2}, \quad Y' = \operatorname{tg} \frac{\pi(1-\xi)}{2}, \quad Z' = -2 \operatorname{tg} \frac{\pi\eta}{2}.$$

Очевидно, что $X \stackrel{d}{=} X'$, $Y \stackrel{d}{=} Y'$, $Z \stackrel{d}{=} Z'$. Наша цель — найти математические ожидания $E\{X + Y + Z\}$ и $E\{X' + Y' + Z'\}$. В случае (I) имеем $X + Y + Z = 0$ и, следовательно, $E\{X + Y + Z\} = 0$. В случае (II) легко находим, что $Y' = \operatorname{ctg} \frac{\pi\xi}{2}$, $Z' = \operatorname{tg} \frac{\pi\xi}{2} - \operatorname{ctg} \frac{\pi\xi}{2}$, если $0 < \xi < \frac{1}{2}$, и $Z' = \operatorname{ctg} \frac{\pi\xi}{2} - \operatorname{tg} \frac{\pi\xi}{2}$, если $\frac{1}{2} < \xi < 1$. Таким образом, $X' + Y' + Z' = 2 \operatorname{tg} \frac{\pi\xi}{2} = 2X'$, если $0 < \xi < \frac{1}{2}$, и $X' + Y' + Z' = 2 \operatorname{ctg} \frac{\pi\xi}{2} = 2Y'$, если $\frac{1}{2} < \xi < 1$. Отсюда вытекает, что $P\{X' + Y' + Z' > 0\} = 1$, а простое вычисление показывает, что $E\{X' + Y' + Z'\} = \frac{4}{\pi} \ln 2$.

Выводы в случаях (I) и (II) показывают, что свойство линейности не имеет места для рассмотренного выше примера трех случайных величин. Заметим, однако, что EX и EY равны ∞ , а EZ равно $-\infty$. Здесь полезно отметить еще следующее. Если обозначить $\tilde{X} = X + Y$ и $\tilde{Y} = Z$, то $E\{X + Y + Z\} = E\{\tilde{X} + \tilde{Y}\}$. Если существует математическое ожидание $E\{\tilde{X} + \tilde{Y}\}$, то оно зависит от $\tilde{X}$ и $\tilde{Y}$ только через распределение $\tilde{X}$ и распределение $\tilde{Y}$. Но $\tilde{X} = X + Y$, и, следовательно, значение $E\{X + Y + Z\}$ в предположении его существования зависит от X , Y и Z через совместное распределение X и Y и через распределение Z .

6.2. Об интегрируемости последовательности случайных величин и ограниченности их супремума

Пусть $\{X_n, n \geq 1\}$ — последовательность неотрицательных случайных величин такая, что $X_n \in \mathcal{L}^p$, $n \geq 1$, для некоторого $p > 0$ и пусть $\sup_n E\{X_n^p\} < \infty$. Вообще говоря, эти условия могут навести на мысль, что такая важная величина, как $\sup_n X_n$, ограничена. Оказывается, что интегрируемость сколь угодно высокой степени p такой ограниченности не обеспечивает. Проиллюстрируем это следующим примером. Рассмотрим последовательность $X_1, X_2, \dots$, состоящую из независимых и одинаково распределенных случайных величин с функцией распределения $F(x) = 0$, если $x \leq 0$, и $F(x) = 1 - e^{-x}$, если $x > 0$ (экспоненциальное распределение с параметром 1). Тогда для любого $p > 0$ мы имеем $E\{X_n^p\} = \Gamma(p+1) < \infty$, где $\Gamma(\cdot)$ — это гамма-функция Эйлера, и, значит, $\sup_n E\{X_n^p\} < \infty$. Далее, для любых $x > 0$ и $m = 1, 2, \dots$ находим, что

$$P\left\{\max_{1 \leq j \leq m} X_j \leq x\right\} = (P\{X_j \leq x\})^m = (1 - e^{-x})^m.$$

Перейдем в этом соотношении к пределу сначала по m , а потом по x . Получаем

$$\lim_{m \rightarrow \infty} P\left\{\max_{1 \leq j \leq m} X_j \leq x\right\} = P\left\{\sup_j X_j \leq x\right\} = 0 \quad \text{для каждого } x > 0$$

и

$$\lim_{x \rightarrow \infty} P\left\{\sup_j X_j \leq x\right\} = P\left\{\sup_j X_j < \infty\right\} = 0.$$

Итак, величины X_n имеют конечные моменты любого порядка $p > 0$, и тем не менее величина $\sup_n X_n$ неограниченна.

6.3. У симметричного распределения все моменты нечетного порядка равны нулю, но обратное неверно

Пусть случайная величина X имеет функцию распределения $F(x)$, $x \in \mathbb{R}^1$, и плотность $f(x)$, $x \in \mathbb{R}^1$. Предположим, что F симметрична, т. е. $F(-x) = 1 - F(x)$, $x \in \mathbb{R}^+$. В терминах плотности это свойство выражается так: $f(-x) = f(x)$, $x \in \mathbb{R}^1$. Предположим далее, что X имеет конечные моменты любого порядка. Тогда при каждом n , $n = 0, 1, \dots$, функция $x^{2n+1}f(x)$ нечетна, и так как интервал $(-\infty, \infty)$ симметричен относительно точки 0, то

$$m_{2n+1} = E\{X^{2n+1}\} = \int_{-\infty}^{\infty} x^{2n+1} f(x) dx = 0, \quad n = 0, 1, \dots$$

Итак, все нечетные моменты симметричного распределения равны 0. Поэтому вполне естественно поставить следующий вопрос: если у распределения $G(x)$, $x \in \mathbb{R}^1$, все моменты существуют, причем все нечетные моменты равны 0, то следует ли отсюда, что распределение G симметричное? Покажем, что ответ на этот вопрос отрицателен. Рассмотрим функцию $g(x)$, $x \in \mathbb{R}^1$, где

$$g(x) = \begin{cases} \frac{1}{48}(1 + \sin \sqrt[4]{|x|}) \exp(-\sqrt[4]{|x|}), & \text{если } x < 0, \\ \frac{1}{48}(1 - \sin \sqrt[4]{x}) \exp(-\sqrt[4]{x}), & \text{если } x \geq 0. \end{cases}$$

Легко проверить, что g является вероятностной плотностью. Пусть ей соответствует случайная величина Y . Конкретный вид плотности g дает возможность вычислить явно моменты $m_n = E\{Y^n\}$:

$$m_{2n+1} = 0, \quad m_{2n} = \frac{1}{6}(8n+3)!, \quad n \geq 0,$$

т. е. у величины Y все центральные моменты нечетного порядка нулевые. Однако распределение G с плотностью g симметричным не является.

6.4. Об одном свойстве моментов случайных величин, которое не имеет аналога для случайных векторов

Пусть $(X_1, \dots, X_n)$ — случайный вектор, заданный на вероятностном пространстве $(\Omega, \mathcal{F}, P)$. Если $k_1, \dots, k_n$ — неотрицательные целые числа и $E\{|X_1|^{k_1} \dots |X_n|^{k_n}\} < \infty$, то величина $m_{k_1, \dots, k_n} = E\{X_1^{k_1} \dots X_n^{k_n}\}$ называется $(k_1, \dots, k_n)$ -смешанным центральным моментом вектора $(X_1, \dots, X_n)$, а число $k = k_1 + \dots + k_n$ — порядком этого момента.

При $n = 1$ мы имеем дело только с одной случайной величиной X_1 . Хорошо известно, что если существует момент $m_k = E\{X_1^k\}$ для некоторого $k > 0$, то существуют моменты m_j для всех $0 \leq j \leq k$. Это вытекает из неравенства Ляпунова $(E\{|X_1|^j\})^{1/j} \leq (E\{|X_1|^k\})^{1/k}$, $0 \leq j \leq k$ (см. [116, 177]), или из элементарного и легко проверяемого неравенства $|x|^j \leq 1 + |x|^k$, $x \in \mathbb{R}^1$, $0 \leq j \leq k$. Отмеченный факт, имеющий место в одномерном случае, приводит нас к вопросу о справедливости аналогичного факта, но уже в многомерном случае. Покажем, что ответ на этот вопрос отрицателен в следующем смысле: если $n \geq 2$ и момент $m_{k_1, \dots, k_n}$ существует, то отсюда не вытекает существование моментов $m_{j_1, \dots, j_n}$ для $0 \leq j_i \leq k_i$, $i = 1, \dots, n$. Это означает, что нет прямого аналога неравенства Ляпунова для многомерного случая.

Итак, пусть $\Omega = (0, 1)$, $\mathcal{F} = \mathcal{B}_{(0,1)}$ и P — мера Лебега. Определим теперь случайные величины X_1 и X_2 так:

$$X_1(\omega) = \begin{cases} \frac{1}{\omega}, & \text{если } 0 < \omega < c_2, \\ 0, & \text{если } c_2 \leq \omega < 1, \end{cases}$$

$$X_2(\omega) = \begin{cases} \frac{1}{1-\omega}, & \text{если } c_1 < \omega < 1, \\ 0, & \text{если } 0 < \omega \leq c_1, \end{cases}$$

где c_1 и c_2 — фиксированные числа, $0 < c_1 \leq c_2 < 1$. Легко видеть, что произведение $X_1 X_2$ интегрируемо, а это означает, что существует момент $m_{1,1}$ для вектора (X_1, X_2) . Более того, если $c_1 < c_2$, то $m_{1,1} > 0$, а при $c_1 = c_2$ имеем $m_{1,1} = 0$. Очевидно также, что ни одна из величин X_1 и X_2 интегрируемой не является. Следовательно, моменты $m_{1,0}$ и $m_{0,1}$ вектора (X_1, X_2) не существуют, хотя, как только что мы видели, момент $m_{1,1}$ существует.

6.5. О справедливости теоремы Фубини

Пусть $(\Omega_1, \mathcal{F}_1, P_1)$ и $(\Omega_2, \mathcal{F}_2, P_2)$ — вероятностные пространства. Тогда существует единственная вероятностная мера P на произведении $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2)$ такая, что

$$P(A_1 \times A_2) = P_1(A_1)P_2(A_2), \quad A_1 \in \mathcal{F}_1, \quad A_2 \in \mathcal{F}_2.$$

Сформулируем теперь один важный результат (*теорема Фубини*), доказательство которого можно найти, например, в [21, 61, 81, 106, 116].

Пусть X — интегрируемая случайная величина, заданная на вероятностном пространстве $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2, P)$, т. е. $X = X(\omega_1, \omega_2)$ измерима относительно σ -алгебры $\mathcal{F}_1 \times \mathcal{F}_2$ и $\int_{\Omega_1 \times \Omega_2} |X(\omega_1, \omega_2)| dP < \infty$.

Тогда интеграл $\int_{\Omega_1} X(\omega_1, \omega_2) dP_1(\omega_1)$ определен для P_2 -почти всех ω_2 , измерим относительно $\mathcal{F}_2$ и

$$P_2 \left\{ \omega_2 : \int_{\Omega_1} |X(\omega_1, \omega_2)| dP_1(\omega_1) < \infty \right\} = 1.$$

Далее, интеграл $\int_{\Omega_2} X(\omega_1, \omega_2) dP_2(\omega_2)$ определен для P_1 -почти всех ω_1 , измерим относительно $\mathcal{F}_1$ и

$$P_1 \left\{ \omega_1 : \int_{\Omega_2} |X(\omega_1, \omega_2)| dP_2(\omega_2) < \infty \right\} = 1.$$

И, наконец, имеет место равенство

$$\begin{aligned} \int_{\Omega_1 \times \Omega_2} X(\omega_1, \omega_2) dP &= \int_{\Omega_1} \left[\int_{\Omega_2} X(\omega_1, \omega_2) dP_2(\omega_2) \right] dP_1(\omega_1) = \\ &= \int_{\Omega_2} \left[\int_{\Omega_1} X(\omega_1, \omega_2) dP_1(\omega_1) \right] dP_2(\omega_2). \quad (1) \end{aligned}$$

Понятно, что для справедливости этого результата (а он включает несколько утверждений) важны все сделанные предположения. Покажем, например, что условие $\int |X| dP < \infty$ является существенным для справедливости равенства (1). Действительно, пусть $Z \geq 0$ — некоторая неинтегрируемая случайная величина на вероятностном пространстве $(\Omega_1, \mathcal{F}_1, P_1)$. На произведении этого пространства с дискретным пространством $\{0, 1\}$, в котором каждый из исходов имеет вероятность $\frac{1}{2}$, задаем случайную величину X следующим образом:

$$X(\omega_1, 0) = Z(\omega_1), \quad X(\omega_1, 1) = -Z(\omega_1).$$

Тогда легко видеть, что второе из равенств в (1) нарушается.

6.6. Семейство интегрируемых случайных величин, которое не является равномерно интегрируемым

Проиллюстрируем связь между введенными понятиями интегрируемости и равномерной интегрируемости.

Рассмотрим последовательность случайных величин $\{X_n, n \geq 1\}$, где

$$P\{X_n = 2^n\} = 2^{-n}, \quad P\{X_n = 0\} = 1 - 2^{-n}.$$

Тогда выполнено следующее равенство:

$$\int_{|X_n| \geq a} |X_n| dP = \begin{cases} 0, & \text{если } a > 2^n, \\ 1, & \text{если } a \leq 2^n. \end{cases}$$

Отсюда вытекает, что величина $\int_{|X_n| \geq a} |X_n| dP$ не стремится к нулю при $a \rightarrow \infty$ равномерно относительно n . Однако все величины X_n интегрируемы, так как $EX_n = 1, n = 1, 2, \dots$. Следовательно, семейство $\{X_n, n \geq 1\}$ состоит из интегрируемых случайных величин, но само семейство не является равномерно интегрируемым. Отметим, что последовательность $\{X_n\}$ в этом примере возникает в так называемом Санкт-Петербургском парадоксе.

6.7. Всегда ли справедливо равенство $E\{E(X|Y)\} = EX$?

Пусть на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ заданы две случайные величины, скажем X и Y , причем X предполагается интегрируемой: $E|X| < \infty$. Если $\mathcal{F}^Y = \sigma\{Y\}$ — σ -алгебра, порожденная Y (очевидно, $\mathcal{F}^Y \subset \mathcal{F}$), то можно ввести величину $E(X|Y)$, равняющуюся $E(X|\mathcal{F}^Y)$ P -п.н., и назвать $E(X|Y)$ условным математическим ожиданием случайной величины X относительно случайной величины Y . Очевидно, что для $E(X|Y)$ будут иметь место все перечисленные в начале этого параграфа свойства для условного математического ожидания относительно заданной σ -алгебры. Остановимся на одном из этих свойств, именно на равенстве $E\{E(X|Y)\} = EX$. Приведем «рассуждение», которое как будто противоречит этому равенству, и выясним соответствующие причины.

Пусть Y — положительная случайная величина с плотностью

$$g_\lambda(y) = \left(\frac{\lambda}{2}\right)^{\lambda/2} \left(\Gamma\left(\frac{\lambda}{2}\right)\right)^{-1} y^{\lambda/2-1} e^{-\lambda y/2}, \quad y > 0, \quad (1)$$

где $\lambda \in \mathbb{N}$ — фиксированный параметр. (Читатель может сравнить g_λ с плотностью гамма-распределения; см. [9, 26, 92, 106, 116].) Условную плотность $f(x|y)$ случайной величины X при фиксированном $Y = y$, где $y > 0$, зададим следующим образом:

$$f(x|y) = (2\pi)^{-1/2} y^{-1/2} e^{-yx^2/2}, \quad x \in \mathbb{R}^1. \quad (2)$$

Очевидно, что функция $f(x|y)$, $x \in \mathbb{R}^1$, симметрична относительно точки $x = 0$ при каждом $y > 0$. Следовательно,

$$E\{X | Y = y\} = \int_{-\infty}^{\infty} x f(x|y) dx = 0,$$

откуда вытекает, что $E\{X|Y\} = 0$, и, значит, $E\{E(X|Y)\} = 0$.

С другой стороны, используя (1) и (2), легко найти плотность h_λ случайной величины X . Именно,

$$h_\lambda(x) = \frac{\Gamma((\lambda+1)/2)}{\sqrt{\lambda\pi}\Gamma(\lambda/2)} \left(1 + \frac{x^2}{\lambda}\right)^{-(\lambda+1)/2}, \quad x \in \mathbb{R}^1, \quad (3)$$

т.е. X имеет t -распределение (распределение Стьюдента; см. [116]) с λ степенями свободы. При $\lambda = 1$ случайная величина X имеет распределение Коши (ее плотность $h_1(x) = \frac{1}{\pi(1+x^2)}$). А в этом случае $E|X| = \infty$, т.е. EX не существует. Так приходим к «заключению», что $E\{E(X|Y)\} \neq EX$.

Причина этого «противоречия» скрывается в выбранном нами ходе рассуждения. Именно, мы начали с плотности (2), а затем из (1) и (2) вывели (3), задающее плотность распределения с математическим ожиданием, равным бесконечности. Достаточно напомнить, что $E|X| < \infty$ — это то условие, с которого начинается определение условного математического ожидания.

6.8. О невозможности расширить одно из свойств условного математического ожидания

Предположим, что X и Z — независимые случайные величины, причем $E|X| < \infty$. Тогда условное математическое ожидание $E(X|Z)$ определено и имеет место равенство $E(X|Z) = EX$ п. н. Теперь хотелось бы безусловное ожидание EX сделать условным, а в условном ожидании $E(X|Z)$ расширить условие. Например, один из возможных вопросов звучит так: имеет ли место равенство типа

$$E(X|Y, Z) = E(X|Y) \text{ п. н.}, \quad (1)$$

где Y — любая случайная величина, заданная на том же пространстве, что и X, Z ?

Покажем, что такое расширение одного из свойств условного математического ожидания не всегда возможно. Действительно, пусть $\Omega = [0, 1]$, $\mathcal{F} = \mathcal{B}_{[0,1]}$ и P — мера Лебега. Рассмотрим три случайные величины X, Y и Z , где

$$X(\omega) = \begin{cases} 1, & \omega \in \left[0, \frac{1}{2}\right), \\ 0, & \omega \in \left[\frac{1}{2}, 1\right], \end{cases} \quad Y(\omega) = \begin{cases} 1, & \omega \in \left[0, \frac{3}{4}\right), \\ 0, & \omega \in \left[\frac{3}{4}, 1\right], \end{cases}$$

$$Z(\omega) = \begin{cases} 1, & \omega \in \left[\frac{1}{4}, \frac{3}{4}\right], \\ 0, & \omega \notin \left[\frac{1}{4}, \frac{3}{4}\right]. \end{cases}$$

Нетрудно проверить, что случайные величины X и Z независимы, а потом вычислить, что

$$E(X|Y) = \begin{cases} \frac{2}{3}, & \omega \in \left[0, \frac{3}{4}\right), \\ 0, & \omega \in \left[\frac{3}{4}, 1\right], \end{cases} \quad E(X|Y, Z) = \begin{cases} 0, & \omega \in \left[\frac{3}{4}, 1\right], \\ \frac{1}{2}, & \omega \in \left[\frac{1}{4}, \frac{3}{4}\right), \\ 1, & \omega \in \left[0, \frac{1}{4}\right). \end{cases}$$

Следовательно, $E(X|Y, Z) \neq E(X|Y)$ и в общем случае соотношение типа (1) не имеет места.

§ 7. Независимость случайных величин

В начале § 5 даны определения независимости двух случайных величин, а также независимости в совокупности (просто независимости) и попарной независимости для любого числа величин. Эти понятия выражаются в общем случае через совместную функцию распределения и маргинальные функции распределения, а в абсолютно непрерывном случае через соответствующие плотности.

Условная независимость случайных величин относительно заданной σ -алгебры вводится также, как это делалось в § 3 при рассмотрении случайных событий.

Для двух случайных величин можно ввести еще одну важную (особенно в приложениях) характеристику. Пусть случайные величины X и Y таковы, что их дисперсии удовлетворяют условиям $0 < DX < \infty$ и $0 < DY < \infty$. Число $\rho(X, Y)$, определяемое по формуле

$$\rho(X, Y) = \frac{E\{(X - EX)(Y - EY)\}}{\sqrt{DX DY}},$$

называется *коэффициентом корреляции* между X и Y . Если $\rho(X, Y) = 0$, то величины X и Y называются *некоррелированными* (или *ортгоналичными*), а если $\rho(X, Y) \neq 0$, — *коррелированными*.

Довольно общие результаты, касающиеся понятий независимости, а также близких вопросов, могут быть найдены во многих книгах и учебных пособиях; см., например, [9, 26, 60, 75, 81, 92, 105, 106, 116, 128, 177, 179, 345, 404, 405].

Выясним далее связь между упомянутыми выше понятиями.

7.1. О попарной независимости случайных величин и их независимости в совокупности

Некоторые из примеров в § 3 могут быть переформулированы в терминах случайных величин. Понятие независимости, однако, настолько важно в теории вероятностей, что ему стоит уделить особое внимание.

(а) Пусть случайный вектор (X, Y, Z) принимает каждое из значений $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ и $(1, 1, 1)$ с вероятностью $\frac{1}{4}$. Тогда для случайных величин X и Z имеем

$$P\{X = i, Z = j\} = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = P\{X = i\}P\{Z = j\}, \quad i = 0, 1; \quad j = 0, 1.$$

Аналогичные равенства выполнены и для каждой из пар X, Y и Y, Z . Поэтому величины X, Y, Z попарно независимы. Но

$$P\{X = 1, Y = 1, Z = 1\} = \frac{1}{4} \neq \frac{1}{8} = P\{X = 1\}P\{Y = 1\}P\{Z = 1\},$$

и, следовательно, эти три величины не являются независимыми в совокупности.

(б) Предположим, что пространство исходов $\Omega = \{\omega_1, \dots, \omega_9\}$, где $\omega_1, \dots, \omega_6$ — перестановки чисел 1, 2 и 3, а $\omega_7 = (1, 1, 1)$, $\omega_8 = (2, 2, 2)$, $\omega_9 = (3, 3, 3)$. Пусть каждый исход ω_i имеет вероятность $\frac{1}{9}$. Проведем одно испытание и введем случайную величину X_k , равную тому числу, которое находится на месте с номером k , $k = 1, 2, 3$. Очевидно, что каждая из этих величин принимает одно из значений 1, 2 и 3. Легко видеть, что

$$P\{X_k = i\} = \frac{1}{3}, \quad P\{X_k = i, X_m = j\} = \frac{1}{9}, \quad k, m = 1, 2, 3, \quad i, j = 1, 2, 3.$$

Отсюда заключаем, что величины X_1 , X_2 и X_3 попарно независимы. Однако они не являются независимыми, потому что, например, X_1 и X_2 определяют однозначно X_3 .

Отметим, что, используя подобную идею, мы сможем построить «длинную» последовательность зависимых случайных величин, которые являются только попарно независимыми.

(в) Пусть ξ и η — независимые случайные величины, равномерно распределенные в интервале $(0, \pi)$. Пусть $X_1 = \operatorname{tg} \xi$, $X_2 = \operatorname{tg} \eta$ и $X_3 = -\operatorname{tg}(\xi + \eta)$. Величины X_1 и X_2 независимы, так как независимы ξ и η . Находя в явном виде распределение X_3 , можно показать, что X_3 не зависит от каждой из величин X_1 и X_2 (детали предоставляем читателю). Следовательно, X_1 , X_2 и X_3 — попарно независимые случайные величины. Поскольку, однако, $X_1 + X_2 + X_3 = X_1 X_2 X_3$, эти три величины не могут быть независимыми.

7.2. Множество зависимых случайных величин, каждое собственное подмножество которого обладает свойством независимости

Напомним, что если случайные величины $X_1, \dots, X_n$, $n \geq 3$, независимы, то любые k из них, $2 \leq k \leq n-1$, также независимы. Покажем, что в общем случае обратное утверждение неверно. Случай $n = 3$ и $k = 2$ был уже рассмотрен в примере 7.1.

(а) В пространстве $\mathbb{R}^{n-1}$, где $n \geq 3$, рассмотрим множество A всех векторов $a = (\alpha_1, \dots, \alpha_{n-1})$ с $\alpha_i = 1$ или $\alpha_i = 0$, $i = 1, \dots, n-1$. Очевидно, что A содержит 2^{n-1} элементов (векторов), т. е. $|A| = 2^{n-1}$. Если обозначить $I(a) = \alpha_1 + \dots + \alpha_{n-1}$, то $I(a)$ равно некоторому из чисел $0, 1, \dots, n-1$. Через $B \subseteq \mathbb{R}^n$ обозначим множество всех векторов b , за-

данных следующим образом:

$$b = \begin{cases} (\alpha_1, \dots, \alpha_{n-2}, 1), & \text{если } I(a) \text{ четное,} \\ (\alpha_1, \dots, \alpha_{n-2}, 0), & \text{если } I(a) \text{ нечетное.} \end{cases}$$

Тогда $T: a \rightarrow b$ является взаимно однозначным отображением множества A в множество B . Таким образом, $|B| = 2^{n-1}$, и, кроме того, B инвариантно относительно перестановок.

Возьмем вектор $b \in B$ и удалим число, находящееся на месте с номером j . Полученный вектор обозначим $a^{(j)}$, а множество всех таких векторов — через $A^{(j)}$. Следовательно, определено отображение $T_j^{-1}: B \rightarrow A^{(j)}$. Ясно, что $A^{(n)} = A$, и так как B инвариантно относительно перестановок, то $A^{(j)} = A^{(n)}$ для каждого j , $j = 1, \dots, n-1$. Тем самым $A^{(j)} = A$ для всех $j = 1, \dots, n$.

Определим теперь случайный вектор $X = (X_1, \dots, X_n) \in \mathbb{R}^n$, принимающий значения в множестве B , причем

$$P\{X = x\} = \begin{cases} 2^{-(n-1)}, & \text{если } x \in B, \\ 0 & \text{в остальных случаях.} \end{cases} \quad (1)$$

Положим $X^{(j)} = T_j^{-1}(X) = (X_1, \dots, X_{j-1}, X_{j+1}, \dots, X_n)$. Так как T_j^{-1} , $j = 1, \dots, n$, являются взаимно однозначными отображениями B в A , мы легко находим распределение вектора $X^{(j)}$:

$$P\{X^{(j)} = x^{(j)}\} = \begin{cases} 2^{-(n-1)}, & \text{если } x^{(j)} \in A, \\ 0, & \text{если } x^{(j)} \notin A. \end{cases} \quad (2)$$

Теперь из (2) нетрудно вычислить маргинальное распределение каждой компоненты X_i вектора X . Получаем

$$P\{X_i = x_i\} = \begin{cases} \frac{1}{2}, & \text{если } x_i = 0 \text{ или } 1, \\ 0 & \text{в остальных случаях.} \end{cases} \quad (3)$$

Из соотношений (1), (2) и (3) следует, что мы построили n дискретных случайных величин $X_1, \dots, X_n$, которые не являются независимыми, но любое их собственное подмножество (т.е. любые k из них, $2 \leq k \leq n-1$) состоит из независимых величин, которые даже одинаково распределены.

(б) Пусть X — абсолютно непрерывная случайная величина с плотностью f и математическим ожиданием $EX = a$. Предположим, что

условие $|x - a| \cdot f(x) < 1$ выполнено для всех $x \in \mathbb{R}^1$. Возьмем n случайных величин $X_1, \dots, X_n$ и в качестве их совместной плотности $g_n(x_1, \dots, x_n)$, $(x_1, \dots, x_n) \in \mathbb{R}^n$ рассмотрим функцию

$$g_n(x_1, \dots, x_n) = \left(\prod_{i=1}^n f(x_i) \right) \left(1 + \prod_{j=1}^n (x_j - a)f(x_j) \right). \quad (4)$$

Эта функция неотрицательна, но, чтобы g_n была вероятностной плотностью, необходимо, чтобы интеграл от g_n по всему пространству $\mathbb{R}^n$ равнялся 1. Это приводит к следующему равенству:

$$\int_{-\infty}^{\infty} (x - a)f^2(x) dx = 0. \quad (5)$$

Отметим, что (5) имеет место, например, для любой плотности f с конечным математическим ожиданием и такой, что f — симметрическая функция относительно числа

$$a = \int_{-\infty}^{\infty} xf(x) dx.$$

Итак, пусть плотность f такова, что $|x - a|f(x) < 1$, $x \in \mathbb{R}^1$, условие (5) выполнено, а совместная плотность g_n величин $X_1, \dots, X_n$ задана равенством (4). Наша цель — установить наличие или отсутствие свойства независимости среди этих n величин.

Из (4) путем интегрирования находим, что каждая из величин $X_1, \dots, X_n$ имеет плотность, равную заданной плотности f . Предположим, что мы выбрали k величин из $X_1, \dots, X_n$. Не ограничивая общности, можно считать, что наш выбор — это $X_1, \dots, X_k$, где $2 \leq k \leq n - 1$. Обозначим совместную плотность $X_1, \dots, X_k$ через $h_k(x_1, \dots, x_k)$, $(x_1, \dots, x_k) \in \mathbb{R}^k$. Тогда из (4) и (5) нетрудно получить, что $h_k(x_1, \dots, x_k) = f(x_1) \dots f(x_k)$, $x_1, \dots, x_k \in \mathbb{R}^1$. Это означает, что величины $X_1, \dots, X_k$ независимы. Такой вывод имеет место и для любого другого подмножества, состоящего из k случайных величин, выбранных из $X_1, \dots, X_n$, где $2 \leq k \leq n - 1$. Тем не менее все n величин $X_1, \dots, X_n$ независимыми не являются, потому что из (4) следует, что $g_n(x_1, \dots, x_n) \neq f(x_1) \dots f(x_n)$.

Отметим, наконец, один частный случай. Пусть случайная величина X распределена равномерно на отрезке $[0, 1]$. Она имеет плотность $f(x) = 1$ при $x \in [0, 1]$ и $f(x) = 0$ при $x \notin [0, 1]$. Тогда $a = EX = \frac{1}{2}$, и можно проверить, что условие (5) выполнено. Рассмотрим случайный

вектор $(X_1, \dots, X_n)$ со следующей плотностью:

$$g_n(x_1, \dots, x_n) = \begin{cases} 1 + \prod_{i=1}^n \left(x_i - \frac{1}{2}\right), & \text{если } 0 \leq x_i \leq 1, \quad i = 1, \dots, n, \\ 0 & \text{в остальных случаях.} \end{cases}$$

Очевидно, что g_n не равняется плотности равномерного распределения в n -мерном кубе $[0, 1]^n \subseteq \mathbb{R}^n$, и, значит, $X_1, \dots, X_n$ не могут быть независимыми. Однако любые k из них, $2 \leq k \leq n-1$, распределены равномерно в k -мерном кубе $[0, 1]^k \subseteq \mathbb{R}^k$, и эти k величин независимы.

7.3. Зависимые случайные величины X и Y , такие что случайные величины X^2 и Y^2 независимы

Хорошо известно, что если X и Y независимые случайные величины, а $g(x)$, $x \in \mathbb{R}^1$, и $h(x)$, $x \in \mathbb{R}^1$, — непрерывные функции, то $g(X)$ и $h(Y)$ являются независимыми случайными величинами (см., например, [26, 75, 106, 116]). Если, например, g и h — взаимно однозначные функции, то верно и обратное утверждение. Оказывается, однако, что функции g и h могут быть выбраны таким образом, что $g(X)$ и $h(Y)$ — независимые случайные величины, даже если сами величины X и Y зависимы.

(а) Пусть (X, Y) — случайный вектор, для которого вероятности $p_{i,j} = P\{X = i, Y = j\}$, где $i, j = -1, 0, 1$, заданы так:

$$\begin{aligned} p_{1,1} = p_{-1,1} &= \frac{1}{32}, & p_{-1,-1} = p_{1,-1} = p_{1,0} = p_{0,1} &= \frac{3}{32}, \\ p_{-1,0} = p_{0,-1} &= \frac{5}{32}, & p_{0,0} &= \frac{8}{32}. \end{aligned}$$

Легко найти распределение каждой из компонент X и Y и убедиться, что X и Y — зависимые случайные величины. Тем не менее (и это также легко проверить) величины X^2 и Y^2 независимы.

(б) Предположим, что X_1 и X_2 — независимые, абсолютно непрерывные случайные величины, и пусть случайная величина Y не зависит от X_1 , X_2 и принимает каждое из значений 1 и -1 с вероятностью $\frac{1}{2}$. Положим

$$Z_1 = YX_1, \quad Z_2 = YX_2.$$

Так как X_1 и X_2 — абсолютно непрерывны, то Z_1 и Z_2 — также абсолютно непрерывные случайные величины. Кроме того, Z_1 и Z_2 не являются независимыми. Наконец, поскольку $Z_1^2 = X_1^2$ и $Z_2^2 = X_2^2$, а, по предположению, X_1 и X_2 независимы, то Z_1^2 и Z_2^2 — независимые случайные величины.

(в) Пусть случайные величины X и Y имеют следующую совместную плотность:

$$f(x, y) = \begin{cases} \frac{1}{4}(1 + xy), & \text{если } |x| < 1 \text{ и } |y| < 1, \\ 0 & \text{в остальных случаях.} \end{cases}$$

Если $f_1(x)$, $x \in \mathbb{R}^1$, и $f_2(y)$, $y \in \mathbb{R}^1$, — маргинальные плотности величин X и Y соответственно, то $f_1(x) = \frac{1}{2}$ при $|x| < 1$ и $f_1(x) = 0$ при $|x| \geq 1$, а $f_2(y) = \frac{1}{2}$ при $|y| < 1$ и $f_2(y) = 0$ при $|y| \geq 1$. Тогда очевидно, что $f(x, y) \neq f_1(x)f_2(y)$, т. е. X и Y независимыми не являются. Заметим далее, что каждая из величин X^2 и Y^2 принимает значения в интервале $(0, 1)$, и для любых $0 < u < 1$ и $0 < v < 1$ находим

$$\begin{aligned} P\{X^2 < u, Y^2 < v\} &= P\{|X| < \sqrt{u}, |Y| < \sqrt{v}\} = \\ &= \frac{1}{4} \int_{-\sqrt{u}}^{\sqrt{u}} \int_{-\sqrt{v}}^{\sqrt{v}} (1 + st) ds dt = (\sqrt{u}) \cdot (\sqrt{v}) = P\{X^2 < u\}P\{Y^2 < v\}. \end{aligned}$$

Следовательно, величины X^2 и Y^2 независимы.

7.4. Независимость случайных величин в терминах характеристических функций

Напомним сначала одно определение. Пусть X — случайная величина на вероятностном пространстве $(\Omega, \mathcal{F}, P)$. Функция $\varphi(t) = E\{e^{itX}\}$, $t \in \mathbb{R}^1$, где $i = \sqrt{-1}$ (мнимая единица), называется *характеристической функцией* случайной величины X . Свойствам характеристических функций целиком посвящен § 8. Здесь мы хотим выяснить, как свойство независимости случайных величин можно выражать и изучать в терминах соответствующих характеристических функций.

Пусть X_1 и X_2 — независимые случайные величины, а φ_1 и φ_2 — их характеристические функции. Тогда характеристическая функция φ суммы $X_1 + X_2$ равна произведению $\varphi_1\varphi_2$, т. е.

$$\varphi(t) = \varphi_1(t)\varphi_2(t) \quad \text{для всех } t \in \mathbb{R}^1. \quad (1)$$

Поставим вопрос: если φ_1 , φ_2 и φ — характеристические функции соответственно случайных величин X_1 , X_2 и $X_1 + X_2$ и выполнено равенство (1), то следует ли отсюда, что X_1 и X_2 независимы. Покажем, что ответ на этот вопрос отрицателен.

(а) Пусть случайный вектор (X_1, X_2) имеет плотность

$$f(x_1, x_2) = \begin{cases} \frac{1}{4}[1 + x_1 x_2(x_1^2 - x_2^2)], & \text{если } |x_1| \leq 1 \text{ и } |x_2| \leq 1, \\ 0 & \text{в остальных случаях.} \end{cases} \quad (2)$$

Во-первых, из (2) легко найти маргинальные плотности f_1 и f_2 , соответствующие величинам X_1 и X_2 :

$$f_1(x_1) = \begin{cases} \frac{1}{2}, & \text{если } |x_1| \leq 1, \\ 0, & \text{если } |x_1| > 1, \end{cases} \quad f_2(x_2) = \begin{cases} \frac{1}{2}, & \text{если } |x_2| \leq 1, \\ 0, & \text{если } |x_2| > 1. \end{cases}$$

Поскольку есть x_1, x_2 , для которых $f(x_1, x_2) \neq f_1(x_1)f_2(x_2)$, то случайные величины X_1 и X_2 зависимы.

Во-вторых, видно, что случайные величины X_1 и X_2 имеют одно и то же распределение (равномерное на $[-1, 1]$), а их характеристические функции φ_1 и φ_2 равны

$$\varphi_1(t) = \varphi_2(t) = \frac{\sin t}{t}, \quad t \in \mathbb{R}^1.$$

В-третьих, найдем плотность g суммы $X_1 + X_2$. Так как функция g связана с f равенством $g(x) = \int_{-\infty}^{\infty} f(x_1, x - x_1) dx_1$ (см. [26, 116]), то с учетом (2) находим

$$g(x) = \begin{cases} \frac{2+x}{4}, & \text{если } -2 \leq x \leq 0, \\ \frac{2-x}{4}, & \text{если } 0 < x \leq 2, \\ 0, & \text{если } |x| > 2. \end{cases}$$

Зная плотность g , нетрудно вычислить характеристическую функцию φ суммы $X_1 + X_2$:

$$\varphi(t) = \frac{\sin^2 t}{t^2}, \quad t \in \mathbb{R}^1.$$

Следовательно, $\varphi(t) = \varphi_1(t)\varphi_2(t)$, т. е. соотношение (1) выполнено, но, как мы видели, величины X_1 и X_2 не являются независимыми.

(б) Рассмотрим случайные величины $X_1 = X$ и $X_2 = X$, где X имеет распределение Коши с плотностью $\frac{1}{\pi(1+x^2)}$, $x \in \mathbb{R}^1$. Если φ_1 , φ_2 и φ — характеристические функции соответственно X_1 , X_2 и $X_1 + X_2$, то $\varphi_1(t) = \varphi_2(t) = e^{-|t|}$, $\varphi(t) = e^{-2|t|}$. Следовательно, $\varphi(t) = \varphi_1(t)\varphi_2(t)$ для всех $t \in \mathbb{R}^1$, т. е. снова выполнено соотношение (1), но величины X_1 и X_2 не являются независимыми.

Наконец, уместно привести следующий общий результат (по этому поводу см., например, [75, 76, 92, 106, 116, 117, 179, 405]): *случайные величины $X_1, \dots, X_n$ с характеристическими функциями $\varphi_1, \dots, \varphi_n$ независимы тогда и только тогда, когда*

$$\mathbf{E}\{\exp[i(t_1 X_1 + \dots + t_n X_n)]\} = \varphi_1(t_1) \dots \varphi_n(t_n)$$

для всех $t_1, \dots, t_n \in \mathbb{R}^1$. Этот результат позволяет нам истолковать правильно рассмотренные выше, а также другие примеры.

7.5. Независимость случайных величин в терминах производящих функций

Если X — целочисленная неотрицательная случайная величина, то функция $p(z) = \mathbf{E}\{z^X\}$ называется *производящей функцией X* (точнее, *вероятностной производящей функцией*). Напомним, что $p(z)$ определена для всех комплексных чисел z с $|z| \leq 1$. В § 8 мы увидим, какова роль производящих функций при изучении случайных величин. Здесь мы выясним, как через эту характеристику можно выразить свойство независимости случайных величин.

Пусть X и Y — независимые случайные величины, принимающие значения в множестве $\mathbb{N}_0$. Через p_X , p_Y и p_{X+Y} обозначим производящие функции, соответствующие X , Y и их сумме $X + Y$. Тогда из свойств математических ожиданий вытекает, что

$$p_{X+Y}(z) = p_X(z)p_Y(z), \quad |z| \leq 1. \quad (1)$$

Ставим вопрос так же, как в примере 7.4: если случайные величины X и Y принимают значения в множестве $\mathbb{N}_0$ и производящие функции p_X , p_Y и p_{X+Y} , отвечающие X , Y и $X + Y$, удовлетворяют соотношению (1), следует ли отсюда, что X и Y независимы? Покажем, что ответ отрицателен.

Пусть случайная величина ξ принимает значения 0, 1 и 2, каждое с вероятностью $\frac{1}{3}$, а случайная величина η не зависит от ξ и принимает значения 0 и 1 с вероятностями соответственно $\frac{1}{3}$ и $\frac{2}{3}$. Введем две новые случайные величины: $X = \xi$ и $Y = \xi + \eta \pmod{3}$. Тогда возможные значения Y — это числа 0, 1 и 2, причем каждое имеет вероятность $\frac{1}{3}$. Далее, сумма $X + Y$ принимает значения 0, 1, 2, 3 и 4 с вероятностями соответственно $\frac{1}{9}$, $\frac{2}{9}$, $\frac{3}{9}$, $\frac{2}{9}$ и $\frac{1}{9}$. Вычисляя производящие функции p_X , p_Y и p_{X+Y} , убеждаемся, что соотношение (1) выполнено. Однако величины X и Y не являются независимыми, так как они связаны

функциональной зависимостью. (Здесь интересно отметить, что X и Y некоррелированы, но об этом см. пример 7.7.)

7.6. Свертка функций распределений может быть распределением суммы зависимых случайных величин

В примере 5.9 уже отмечалось, что если X_1 и X_2 — независимые случайные величины, а F_1 и F_2 — их функции распределения, то сумма $X_1 + X_2$ имеет функцию распределения, равную свертке $F_1 * F_2$. Если X_1 и X_2 абсолютно непрерывны, а f_1 и f_2 — их плотности, то сумма имеет плотность $f_1 * f_2$.

Возникает вопрос: если случайные величины X_1 и X_2 таковы, что их сумма имеет функцию распределения $F_1 * F_2$, или плотность $f_1 * f_2$, следует ли отсюда, что величины X_1 и X_2 независимы? Покажем, что это не всегда так.

(а) Напомним, что функция

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

где $a > 0$, называется плотностью распределения Коши с параметром a . Нетрудно проверить, например, с использованием характеристических функций, что семейство распределений Коши замкнуто относительно операции свертки. Воспользуемся этим свойством следующим образом. Рассмотрим две независимые случайные величины ξ и η , и пусть каждая из них имеет плотность f_a . Положим $X_1 = \alpha\xi + \beta\eta$ и $X_2 = \gamma\xi + \Delta\eta$, где $\alpha, \beta, \gamma, \Delta$ — произвольные положительные числа. Тогда сумма $X_1 + X_2$ имеет плотность $f_{(\alpha+\beta+\gamma+\Delta)a}$, которая равняется свертке плотностей $f_{(\alpha+\beta)a}$ и $f_{(\gamma+\Delta)a}$ соответствующих случайных величин X_1 и X_2 . Очевидно, однако, что случайные величины X_1 и X_2 не являются независимыми.

(б) Вернемся снова к случайному вектору (X_1, X_2) , рассмотренному в примере 7.4 (а). Найденные там в явном виде маргинальные плотности f_1 и f_2 дают возможность очень легко вычислить свертку $f_1 * f_2$. Оказывается, что $(f_1 * f_2)(x) = g(x)$ для всех $x \in \mathbb{R}^1$, где g — плотность суммы $X_1 + X_2$. Здесь вывод такой же, как и в случае (а), а именно: если распределение суммы двух случайных величин равно свертке распределений этих величин, то отсюда не вытекает, что величины независимы.

7.7. Некоррелированные, но зависимые случайные величины

Если случайные величины X и Y интегрируемы и независимы, то они некоррелированы. Покажем, что обратное, вообще говоря, неверно.

(а) Пусть X и Y — случайные величины, вероятности которых $p_{i,j} = P\{X = i, Y = j\}$, заданы следующим образом:

$$p_{1,1} = p_{-1,1} = p_{1,-1} = p_{-1,-1} = \frac{1}{4}\varepsilon, \quad p_{0,1} = p_{0,-1} = p_{1,0} = p_{-1,0} = \frac{1}{4}(1-\varepsilon),$$

где $0 < \varepsilon < 1$. Легко найти распределение каждой из величин X и Y , откуда сразу видно, что $EX = 0$ и $EY = 0$. Кроме того, $E\{XY\} = 0$, и, значит, X и Y некоррелированы. Очевидно, однако, что

$$P\{X = 0, Y = 0\} = 0 \neq \frac{1}{16}(1-\varepsilon)^2 = P\{X = 0\}P\{Y = 0\},$$

и, следовательно, случайные величины X и Y зависимые.

(б) Пусть случайная величина ξ распределена равномерно на отрезке $[0, 2\pi]$. Положим $X = \sin \xi$ и $Y = \cos \xi$. Тогда

$$EX = 0, \quad EY = 0, \quad E\{XY\} = 0.$$

Значит, величины X и Y некоррелированы. Однако они не независимы, так как связаны функционально: $X^2 + Y^2 = 1$.

(в) Напомним, что случайная величина ξ называется нормально распределенной с параметрами a и σ^2 , где $a \in \mathbb{R}^1$, $\sigma > 0$, если ее плотность равняется $(\sqrt{2\pi}\sigma)^{-1} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right)$, $x \in \mathbb{R}^1$. Обозначают это так: $\xi \sim \mathcal{N}(a, \sigma^2)$.

Пусть $X = \xi$, $Y = \xi^2 - 1$, где $\xi \sim \mathcal{N}(a, \sigma^2)$. Тогда $EX = 0$, и легко подсчитать, что $EY = 0$ и $E\{XY\} = 0$. Следовательно, величины X и Y некоррелированы. Поскольку X и Y связаны равенством $Y = X^2 - 1$, то они не могут быть независимыми.

Отметим, наконец, что если X и Y имеют совместное нормальное распределение и, кроме того, некоррелированы, то отсюда будет следовать, что X и Y независимы.

7.8. Из равенства $E(Y|X) = EY$ почти наверное не всегда вытекает, что X и Y независимы

Пусть случайные величины X и Y заданы на вероятностном пространстве $(\Omega, \mathcal{F}, P)$. Предположим, что они независимы, а случайная величина Y интегрируема. Тогда $E(Y|X) = EY$ п. н. (Это одно из свойств условных математических ожиданий; см. начало § 6.) Возникает вопрос о справедливости обратного утверждения. Обратимся сразу к примеру.

Возьмем произвольную случайную величину X такую, что $X \geq 1$ п. н. Пусть, далее, случайная величина Z симметрична относительно

точки 0, причем Z интегрируема и независима от X . Введем величину $Y = \frac{Z}{X}$. Тогда Y — интегрируемая случайная величина и условное математическое ожидание $E(Y|X)$ определено. Очевидно, что

$$EZ = 0, \quad EY = 0, \quad E(Y|X) = 0.$$

Следовательно, для случайных величин X и Y выполнено соотношение $E(Y|X) = EY$ п. н., но величины X и Y зависимы.

7.9. Существует ли связь между понятиями независимости и условной независимости случайных величин?

Связь между этими двумя понятиями проиллюстрируем с помощью двух примеров, из которых станет ясно, что непосредственной связи не существует.

(а) Пусть случайный вектор (X, Y) задан следующим образом:

$$P\{X = i, Y = j\} = p_i q_j, \quad i = 0, 1; \quad j = 0, 1,$$

где $0 < p_0 < 1$, $0 < p_1 < 1$, $p_0 + p_1 = 1$, $0 < q_0 < 1$, $0 < q_1 < 1$, $q_0 + q_1 = 1$. Легко установить, что случайные величины X и Y независимы, потому что для всех возможных пар (i, j) , $i = 0, 1$, $j = 0, 1$, выполнено равенство

$$P\{X = i, Y = j\} = P\{X = i\}P\{Y = j\}.$$

Введем теперь новую случайную величину Z , где $Z = 1$, если $X = 1$ и $Y = 1$, и $Z = 0$ в остальных случаях. Очевидно, что $P\{Z = 0\} = 1 - p_1 q_1$, причем $1 - p_1 q_1 > 0$. Далее находим, что

$$P\{X = 1 | Z = 0\} = \frac{p_1 q_1}{1 - p_1 q_1} > 0, \quad P\{Y = 1 | Z = 0\} = \frac{p_1 q_1}{1 - p_0 q_1} > 0.$$

Поскольку $P\{X = 1, Y = 1 | Z = 0\} = 0$, для рассмотренных условных вероятностей имеет место соотношение

$$P\{X = 1, Y = 1 | Z = 0\} \neq P\{X = 1 | Z = 0\}P\{Y = 1 | Z = 0\}.$$

Это означает, что случайные величины X и Y не являются условно независимыми (здесь относительно случайной величины Z).

(б) Пусть X_n , $n \geq 1$, — независимые целочисленные случайные величины и $S_n = X_1 + \dots + X_n$. Тогда члены последовательности S_n , $n \geq 1$, — это зависимые случайные величины. Покажем тем не менее, что некоторые из этих зависимых величин могут быть условно независимыми относительно конкретно подобранной случайной величины.

Действительно, рассмотрим случайные величины S_1, S_2, S_3 и предположим, что для некоторого k событие $\{S_2 = k\}$ имеет положительную вероятность. Тогда имеют место равенства:

$$\begin{aligned} P\{S_1 = i, S_3 = j \mid S_2 = k\} &= \frac{P\{S_1 = i, S_2 = k, S_3 = j\}}{P\{S_2 = k\}} = \\ &= \frac{P\{S_1 = i\}P\{X_2 = k - i\}P\{X_3 = j - k\}}{P\{S_2 = k\}} = \\ &= P\{S_1 = i \mid S_2 = k\} \frac{P\{X_3 = j - k\}P\{S_2 = k\}}{P\{S_2 = k\}} = \\ &= P\{S_1 = i \mid S_2 = k\}P\{S_3 = j \mid S_2 = k\}. \end{aligned}$$

Отсюда следует, что случайные величины S_1 и S_3 являются условно независимыми относительно случайной величины S_2 .

Итак, рассмотренные примеры показывают, что в общей ситуации из независимости случайных величин не следует их условная независимость, а также, наоборот, из условной независимости случайных величин не вытекает, что эти величины независимы.

§ 8. Характеристические и производящие функции

Пусть случайная величина X определена на вероятностном пространстве $(\Omega, \mathcal{F}, P)$. Как уже упоминалось в примере 7.4, функция

$$\varphi(t) = E\{e^{itX}\}, \quad t \in \mathbb{R}^1, \quad i = \sqrt{-1}, \quad (1)$$

называется *характеристической функцией* случайной величины X . Если $F(x)$, $x \in \mathbb{R}^1$, — функция распределения X , то $\varphi(t) = \int_{-\infty}^{\infty} e^{itx} dF(x)$.

Таким образом, $\varphi(t) = \int_{-\infty}^{\infty} e^{itx} f(x) dx$, если X — абсолютно непрерывная

случайная величина с плотностью f , а если X — дискретная случайная величина с $P\{X = x_n\} = p_n$, $p_n > 0$, $\sum_n p_n = 1$, то $\varphi(t) = \sum_n e^{itx_n} p_n$. Ха-

рактеристическая функция $\varphi(t)$ существует для любого $t \in \mathbb{R}^1$ и обладает рядом полезных свойств, среди которых мы отметим следующие:

- (1) $\varphi(0) = 1$, $\varphi(-t) = \overline{\varphi(t)}$, $|\varphi(t)| \leq 1$, $t \in \mathbb{R}^1$;
- (2) если $E\{X^n\}$ существует, то существует производная $\varphi^{(n)}(t)$ при $t = 0$ и $E\{X^n\} = i^{-n} \varphi^{(n)}(0)$;
- (3) если $\varphi^{(n)}(0)$ существует и n четное, то существует и $E\{X^n\}$, а если n нечетное, то существует $E\{X^{n-1}\}$;
- (4) комплекснозначная функция $\varphi(t)$, $t \in \mathbb{R}^1$, является характеристической функцией тогда и только тогда, когда φ — положительно определенная функция и $\varphi(0) = 1$ (*теорема Бознера*).

В примере 7.5 для неотрицательной целочисленной случайной величины X мы ввели функцию $p(z) = E\{z^X\}$, называемую *вероятностной производящей функцией*. С помощью функции $p(z)$ можно выразить ряд свойств этого класса случайных величин и их распределений. Для любой случайной величины очень важной характеристикой является *производящая функция (моментов)*:

$$M(z) = E\{e^{zX}\}. \quad (2)$$

Очевидно, что можно рассматривать $M(z)$ как функцию комплексной переменной z . Если z — чисто мнимое число, то $M(z)$ — это характеристическая функция φ (см. формулу (1) выше). Для многих целей можно ограничиться рассмотрением $M(t)$ как функции действительной переменной t , т. е. для $t \in \mathbb{R}^1$. Из (1) и (2) ясно, что свойства характеристических функций могут быть переформулированы в терминах производящих функций. Отметим, однако, что характеристическая функция существует для любого распределения, в то время как производящая функция определена далеко не всегда (исключая тривиальный случай, когда $t = 0$).

Характеристическая функция φ называется *аналитической* (а также *r -аналитической*), если для некоторого $r > 0$ она может быть представлена в виде сходящегося ряда в интервале $(-r, r)$, т. е. если $\varphi(t) = \sum_{k=0}^{\infty} \frac{a_k t^k}{k!}$, $t \in (-r, r)$, с некоторыми комплексными коэффициентами a_k . Сформулируем следующий важный результат (см., например, [76, 106, 177]): *пусть F является функцией распределения, а φ — соответствующая ей характеристическая функция. Тогда следующие три условия эквивалентны:*

- (а) *функция φ r -аналитическая;*
- (б) *все моменты $m_k = \int x^k dF(x)$, $k \geq 1$, конечны и φ допускает представление $\varphi(t) = \sum_{k=0}^{\infty} \frac{m_k(it)^k}{k!}$ для $t \in (-r, r)$;*
- (в) *$\int e^{t|x|} dF(x) < \infty$, $0 \leq t < r$ (условие Крамера).*

Отсюда следует, что производящая функция M существует тогда и только тогда, когда характеристическая функция φ аналитична.

Читатель может обратиться ко многим источникам, где теория характеристических и производящих функций изложена очень подробно (см., например, [9, 24, 26, 75, 76, 106, 116, 117, 345, 404, 405, 424]). Из рассмотренных ниже примеров будет видна та важная роль, которую эти две характеристики имеют при изучении случайных величин и их распределений.

8.1. Характеристические функции, совпадающие на конечном отрезке, но не на всей числовой прямой

Предположим, что φ_1 и φ_2 — такие характеристические функции, что $\varphi_1(t) = \varphi_2(t)$ для $t \in [-c, c]$, где c — некоторое фиксированное положительное число (c может быть очень «большим»). Естественно, может возникнуть следующий вопрос: а следует ли отсюда, что φ_1 и φ_2 совпадают для всех $t \in \mathbb{R}^1$?

Впервые эта проблема была рассмотрена Б. В. Гнеденко в его работе [23]. Изложим предложенное им решение.

Рассмотрим функцию $h(x) = 0$, если $|x| > \frac{\pi}{2}$, и $h(x) = x$, если $|x| \leq \frac{\pi}{2}$. Можно показать, что если $b(t) = \int_{-\infty}^{\infty} h(x)h(x+t) dx$, то функция $\varphi_1(t) = \frac{b(t)}{b(0)}$, $t \in \mathbb{R}^1$, является характеристической функцией. Простое вычисление показывает, что

$$\varphi_1(t) = \begin{cases} 1 + 3\pi^{-1}t - 2\pi^{-3}t^3, & \text{если } -\pi \leq t \leq 0, \\ 1 - 3\pi^{-1}t + 2\pi^{-3}t^3, & \text{если } 0 \leq t \leq \pi, \\ 0, & \text{если } |t| > \pi. \end{cases}$$

Введем функцию φ_2 следующим образом:

$$\varphi_2(t) = \varphi_1(t) \quad \text{для } |t| \leq \pi, \quad \varphi_2(t + 2\pi) = \varphi_2(t) \quad \text{для } t \in \mathbb{R}^1.$$

Покажем, что φ_2 — характеристическая функция. Действительно, функция φ_2 четная, а ее ряд Фурье имеет вид

$$\frac{1}{2} a_0 + \sum_{k=1}^{\infty} a_k \cos kt, \quad (1)$$

где

$$a_0 = 0, \quad a_k = 6\pi^2 [k^{-2}(1 + \cos k\pi) + 4\pi^{-2}k^{-4}(1 - \cos k\pi)], \quad k = 1, 2, \dots$$

Таким образом, ряд (1) сходится равномерно, его коэффициенты неотрицательны, а сумма этих коэффициентов равна $\varphi_2(0) = 1$. Следовательно, существует такая функция распределения F , что $\varphi_2(t) = \int_{-\infty}^{\infty} e^{itx} dF(x)$. Отсюда вытекает, что φ_2 является характеристической функцией. Итак, мы построили две характеристические функции φ_1 и φ_2 такие, что $\varphi_1(t) = \varphi_2(t)$ для $t \in [-\pi, \pi]$, но это равенство не имеет места для всех $t \in \mathbb{R}^1$. Подобным же образом можно построить характеристические функции φ_1 и φ_2 , которые совпадают на сколь угодно большом отрезке $[-c, c]$, но не на всей прямой $\mathbb{R}^1$.

Отметим, наконец, что к работе Б. В. Гнеденко [23] имеется короткое, но важное замечание А. Я. Хинчина по поводу изложенного выше результата. Именно, пусть F_1 и F_2 — функции распределения, соответствующие характеристическим функциям φ_1 и φ_2 . Тогда из этого результата вытекает равенство

$$\varphi_1(t)\varphi_1(t) = \varphi_1(t)\varphi_2(t) \quad \text{для всех } t \in \mathbb{R}^1,$$

которое эквивалентно следующему соотношению для сверток:

$$F_1 * F_1 = F_1 * F_2. \quad (2)$$

Соотношение (2) показывает, что существует функция распределения, свертки которой с двумя различными функциями распределения могут привести к одному и тому же результату. Другими словами, из уравнения (2) в общем случае не вытекает, что $F_1 = F_2$.

8.2. Дискретное и абсолютно непрерывное распределения с характеристическими функциями, совпадающими на отрезке $[-1, 1]$

Рассмотрим следующую функцию, принадлежащую классу Пойа:

$$\varphi_1(t) = \begin{cases} 1 - |t|, & \text{если } |t| \leq 1, \\ 0, & \text{если } |t| > 1. \end{cases} \quad (1)$$

Эта функция является характеристической функцией (см. [76, 106]). Следовательно, существует случайная величина X с характеристической функцией, равной φ_1 . Поскольку φ_1 — абсолютно интегрируемая функция на $\mathbb{R}^1$, случайная величина X будет абсолютно непрерывной, а ее плотность f равна обратному преобразованию Фурье от φ_1 , т. е.

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi_1(t) dt = \frac{1 - \cos x}{\pi x^2}, \quad x \in \mathbb{R}^1.$$

Рассмотрим теперь дискретную случайную величину Y , заданную таким образом:

$$P\{Y = 0\} = \frac{1}{2}, \quad P\{Y = (2k - 1)\pi\} = \frac{2}{(2k - 1)^2 \pi^2}, \quad k = 0, \pm 1, \pm 2, \dots$$

Для характеристической функции $\varphi_2(t) = E\{e^{itY}\}$ находим

$$\varphi_2(t) = \frac{1}{2} + \frac{4}{\pi^2} \sum_{k=1}^{\infty} \frac{\cos(2k - 1)\pi t}{(2k - 1)^2}. \quad (2)$$

Мы хотим показать, что функции φ_1 и φ_2 , задаваемые соотношениями (1) и (2), совпадают для каждого $t \in [-1, 1]$. Для этой цели возьмем функцию $h(t) = |t|$. Ее ряд Фурье имеет вид

$$h(t) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos n\pi t, \quad |t| \leq 1,$$

где $a_0 = 1$, $a_n = \frac{2(\cos nt - 1)}{n^2\pi^2}$. Заметим, что при $k = 1, 2, \dots$ имеем $a_{2k} = 0$, $a_{2k-1} = -\frac{4}{(2k-1)^2\pi^2}$. Остается сравнить функции h , φ_2 и φ_1 и заключить, что $\varphi_1(t) = \varphi_2(t)$ для каждого $t \in [-1, 1]$.

Итак, мы описали явно две характеристические функции φ_1 и φ_2 такие, что φ_1 соответствует абсолютно непрерывному распределению, а φ_2 — дискретному распределению, причем $\varphi_1(t) = \varphi_2(t)$ для $t \in [-1, 1]$. Заметим, однако, что $\varphi_1(t) \neq \varphi_2(t)$ при $|t| > 1$.

8.3. Абсолютно непрерывное распределение может иметь характеристическую функцию, которая не интегрируема абсолютно

Пусть φ — характеристическая функция, а F — соответствующая ей функция распределения. Если φ абсолютно непрерывна и интегрируема, то плотность $f = F'$ непрерывна и задается обратным преобразованием Фурье от φ . Этот результат уже упоминался в примере 8.2, а его доказательство может быть найдено, например, в [75, 76, 106]. Вопрос: существуют ли абсолютно непрерывные распределения, у которых характеристические функции не интегрируемы абсолютно. Оказывается, таких примеров много.

Один из самых наглядных примеров следующий. Пусть случайная величина X имеет экспоненциальное распределение с параметром λ ($\lambda > 0$). Это означает, что X — абсолютно непрерывная случайная величина с плотностью $f(x) = \lambda e^{-\lambda x}$, если $x > 0$, и $f(x) = 0$, если $x \leq 0$. Легко найти ее характеристическую функцию φ : $\varphi(t) = \frac{\lambda}{\lambda - it}$, $t \in \mathbb{R}^1$. Записывая φ в виде $\varphi(t) = \frac{\lambda(\lambda + it)}{\lambda^2 + t^2}$, нетрудно убедиться, что $\int_{-\infty}^{\infty} |\varphi(t)| dt = \infty$, т. е. φ не интегрируема абсолютно на $\mathbb{R}^1$. Тем не менее само распределение абсолютно непрерывно.

Другие абсолютно непрерывные распределения с неабсолютно интегрируемыми характеристическими функциями могут быть получены с использованием следующей теоремы Пойа (см. [76, 106, 117]): пусть функция $\psi(t)$, $t \in \mathbb{R}^1$, принимает действительные значения и для нее

выполнены условия: (1) $\psi(0) = 1$; (2) $\psi(-t) = \psi(t)$; (3) $\psi(t)$ выпукла вниз при $t > 0$; (4) $\lim_{t \rightarrow \infty} \psi(t) = 0$. Тогда ψ является характеристической функцией абсолютно непрерывного распределения.

Возьмем, например,

$$\psi_1(t) = \frac{1}{1 + |t|}, \quad t \in \mathbb{R}^1;$$

$$\psi_2(t) = \begin{cases} 1 - |t|, & \text{если } 0 \leq |t| \leq \frac{1}{2}, \\ \frac{1}{4|t|}, & \text{если } |t| \geq \frac{1}{2}. \end{cases}$$

Легко видеть, что для каждой из функций ψ_1 и ψ_2 выполнены условия (1)–(4) в теореме Пойа. Следовательно, ψ_1 и ψ_2 являются характеристическими функциями, и, более того, им соответствуют абсолютно непрерывные распределения. Очевидно, однако, что функции ψ_1 и ψ_2 не интегрируемы абсолютно на $\mathbb{R}^1$.

Итак, абсолютная интегрируемость характеристической функции не является необходимым условием для того, чтобы соответствующее распределение было абсолютно непрерывным.

8.4. Неинтегрируемая дискретная случайная величина с дифференцируемой характеристической функцией

Будет полезно проиллюстрировать свойство (3) характеристических функций, см. начало этого параграфа.

Пусть X — дискретная случайная величина, заданная так:

$$P\{X = n\} = P\{X = -n\} = \frac{c}{n^2 \ln n}, \quad n = 2, 3, \dots,$$

где нормирующая константа $c = \left(2 \sum_{n=2}^{\infty} (n^2 \ln n)^{-1}\right)^{-1}$. Тогда для характеристической функции φ имеем

$$\varphi(t) = 2c \sum_{n=2}^{\infty} \frac{\cos nt}{n^2 \ln n}, \quad t \in \mathbb{R}^1. \quad (1)$$

Вопрос: дифференцируема ли характеристическая функция φ в точке $t = 0$ и существует ли математическое ожидание EX ?

Заметим, что бесконечный ряд $\sum_{n=2}^{\infty} \frac{\sin nt}{n}$ имеет ограниченные частичные суммы. Значит, ряд $\sum_{n=2}^{\infty} \frac{\sin nt}{n \ln n}$ сходится равномерно (см. [38]). Но ряд $\sum_{n=2}^{\infty} \frac{\sin nt}{n \ln n}$ получен путем дифференцирования (1). Следовательно,

характеристическая функция $\varphi(t)$ дифференцируема для всех $t \in \mathbb{R}^1$, и при $t = 0$ имеем $\varphi'(t) = 0$.

С другой стороны, случайная величина X не интегрируема, т. е. математическое ожидание EX не существует, так как ряд $\sum_{n=2}^{\infty} \frac{1}{n \ln n}$ расходится.

8.5. Неинтегрируемая абсолютно непрерывная случайная величина с дифференцируемой характеристической функцией

Пусть случайная величина X имеет плотность

$$f(x) = \begin{cases} \frac{c}{x^2 \ln |x|}, & \text{если } |x| > 2, \\ 0, & \text{если } |x| \leq 2, \end{cases}$$

где c — нормирующая константа, $0 < c < \infty$ (ее точное значение для наших рассуждений неважно). Поскольку $\int_2^{\infty} (x \ln x)^{-1} dx = \infty$, то математическое ожидание EX не существует. Покажем, как и в примере 8.4, что соответствующая характеристическая функция φ дифференцируема. Действительно, поскольку $\varphi(t) = 2c \int_2^{\infty} \frac{\cos tx}{x^2 \ln x} dx$, φ — четная функция и для $t > 0$ выражение $\frac{1 - \varphi(t)}{2c}$ можно представить так:

$$\frac{1 - \varphi(t)}{2c} = \int_2^{1/t} \frac{1 - \cos tx}{x^2 \ln x} dx + \int_{1/t}^{\infty} \frac{1 - \cos tx}{x^2 \ln x} dx.$$

Очевидно, что функция $1 - \varphi(t)$ неотрицательная. Далее, для любого $u \in \mathbb{R}^1$ имеет место неравенство $0 \leq 1 - \cos u \leq \min(2, u^2)$, откуда вытекает, что $1 - \varphi(t)$ не превышает некоторую константу, умноженную на функцию

$$h(t) = t^2 \int_2^{1/t} \frac{1}{\ln x} dx + 2 \int_{1/t}^{\infty} \frac{1}{x^2 \ln x} dx.$$

Однако $h(t) = O\left(-\frac{t}{\ln t}\right) = o(t)$ при $t \rightarrow 0$. Таким образом,

$$\varphi(t) = 1 + o(t) \quad \text{при } t \rightarrow 0.$$

Следовательно, характеристическая функция $\varphi(t)$ дифференцируема в точке $t = 0$ и $\varphi'(0) = 0$.

8.6. Производящая функция может не существовать, даже если случайная величина имеет моменты любого порядка r , $r = 1, 2, \dots$

(а) Рассмотрим случайную величину X с распределением вероятностей $P\{X = 2^n\} = \frac{1}{en!}$, $n = 0, 1, 2, \dots$ Так как

$$E\{X^r\} = \exp(2^r - 1), \quad r = 1, 2, \dots,$$

то X имеет моменты всех порядков. Степенной ряд

$$h(z) = \sum_{n=0}^{\infty} \frac{z^{(2^n)}}{en!}$$

сходится при $|z| \leq 1$ и расходится при $|z| > 1$. Производящая функция $M(t) = E\{e^{tX}\}$ выражается через h с помощью равенства $M(t) = h(e^t)$. Это означает, что $M(t)$ существует для $t \leq 0$, и очевидно, что если $t > 0$, то $M(t)$ не существует. Заметим еще, что $M(t)$ не дифференцируема в любой окрестности точки $t = 0$, и, следовательно, $M(t)$ не может «произвести» моменты случайной величины X , которые, как показано выше, существуют.

(б) Пусть Y — случайная величина, для которой $P\{Y = 2^n\} = P\{Y = -2^n\} = \frac{1}{2en!}$, $n = 0, 1, 2, \dots$ Эта величина имеет конечные моменты любого порядка:

$$E\{Y^{2r+1}\} = 0, \quad E\{Y^{2r}\} = \exp(2^{2r} - 1), \quad r = 0, 1, 2, \dots$$

Если бы производящая функция $M(t) = E\{e^{tY}\}$ существовала, то ее можно было бы записать в виде

$$M(t) = \frac{1}{e} \sum_{n=0}^{\infty} \frac{\operatorname{ch}(2^n t)}{n!}.$$

Однако ряд в правой части сходится только при $t = 0$, что можно легко проверить с помощью критерия Даламбера.

Итак, у случайной величины Y существуют моменты любого порядка, а производящая функция $M(t)$ не существует.

(в) Рассмотрим теперь случайные величины с абсолютно непрерывными распределениями.

Пусть случайная величина Z имеет плотность

$$f(x) = \begin{cases} 0, & \text{если } x < 0, \\ \frac{1}{2} \exp(-\sqrt{x}), & \text{если } x \geq 0. \end{cases}$$

Тогда $m_k = E\{Z^k\} = (2k+1)!$, $k = 0, 1, 2, \dots$, и, значит, Z имеет конечные моменты любого порядка.

Вопрос: существует ли производящая функция $M(t)$, $t \in \mathbb{R}$? В начале настоящего параграфа было отмечено, что существование M эквивалентно аналитичности характеристической функции φ величины Z . Поэтому удобно воспользоваться следующим результатом (см. [76, 345]): *характеристическая функция ψ , отвечающая случайной величине ξ , является аналитической функцией тогда и только тогда, когда выполнены следующие условия: (1) ξ имеет конечные моменты m_k любого порядка k , $k \geq 1$; (2) существует постоянная $c > 0$ такая, что $|m_k| \leq k!c^k$ для каждого $k \geq 1$.*

Выше мы отметили, что $m_k = E\{Z^k\} = (2k+1)!$, и, значит, условие (1) выполнено. Однако условие (2) не выполнено, так как имеет место следующее соотношение:

$$\left(\frac{|m_k|}{k!}\right)^{1/k} = k^{(k+1)/k} \left[\left(1 + \frac{1}{k}\right) \dots \left(1 + \frac{k+1}{k}\right)\right]^{1/k} > k.$$

Следовательно, характеристическая функция φ не является аналитической функцией, т. е. производящая функция M не существует (за исключением тривиального случая, когда $t = 0$).

Несуществование производящей функции M вытекает также из записи $M(t)$ в виде

$$M(t) = \frac{1}{2} \int_0^{\infty} \exp(tx - \sqrt{x}) dx.$$

Если $\varepsilon > 0$ — любое число, то для $t \in (0, \varepsilon)$ имеем $tx - \sqrt{x} \rightarrow \infty$ при $x \rightarrow \infty$. Значит, $\int_0^{\infty} \exp(tx - \sqrt{x}) dx = \infty$. Следовательно, производящая функция M не существует несмотря на то, что существуют все моменты случайной величины Z .

Наконец, сделаем еще одно замечание. Пусть U — абсолютно непрерывная случайная величина, плотность g которой имеет вид $g(x) = c \exp(-|x|^\gamma)$, $x \in \mathbb{R}^1$. Здесь $\gamma \in (0, 1)$, а нормирующая константа c определяется равенством $c^{-1} = \int_{-\infty}^{\infty} \exp(-|x|^\gamma) dx$. Тогда нетрудно показать, что $E\{|U|^k\} < \infty$ для каждого $k \geq 1$, и заключить, что случайная величина U имеет конечные моменты любого порядка. Однако для любого $\gamma \in (0, 1)$ характеристическая функция φ не является аналитической, и производящая функция M не существует, что можно видеть и непосредственно из расходимости соответствующего интеграла.

§ 9. Безгранично делимые и устойчивые распределения

Пусть случайная величина X определена на вероятностном пространстве $(\Omega, \mathcal{F}, P)$. Через F и φ обозначим соответствующие ей функцию распределения и характеристическую функцию. Величина X , а также F и φ называются *безгранично делимыми*, если для каждого $n \geq 2$ существуют независимые и одинаково распределенные случайные величины $X_{n1}, \dots, X_{nn}$ такие, что $X \stackrel{d}{=} X_{n1} + \dots + X_{nn}$. Это эквивалентно соотношениям

$$F = F_n * \dots * F_n = (F_n)^{*n}, \quad \varphi = (\varphi_n)^n,$$

где F_n есть функция распределения, а φ_n — характеристическая функция.

Перечислим здесь только несколько свойств безгранично делимых распределений: (1) распределение F с ограниченным носителем безгранично делимо тогда и только тогда, когда оно вырождено; (2) произведение конечного числа безгранично делимых характеристических функций есть характеристическая функция, которая безгранично делима.

Сформулируем следующий принципиально важный результат (см., например, [24, 75, 76, 116, 177]): *случайная величина X с характеристической функцией φ безгранично делима тогда и только тогда, когда φ допускает следующее каноническое представление, известное как представление Леви–Хинчина:*

$$\varphi(t) = \exp \left(i\gamma t + \int_{-\infty}^{\infty} \left(e^{itu} - 1 - \frac{itu}{1+u^2} \right) \frac{1+u^2}{u^2} dG(u) \right), \quad t \in \mathbb{R}^1, \quad (1)$$

где $\gamma \in \mathbb{R}^1$, а $G(x)$, $x \in \mathbb{R}^1$, — неубывающая, непрерывная слева функция ограниченной вариации с $G(-\infty) = 0$.

Введем еще одно понятие. Случайная величина X , ее функция распределения F и ее характеристическая функция φ называются *устойчивыми*, если для каждого $n \geq 2$ существуют константы $a_n \in \mathbb{R}^1$ и $b_n > 0$ и независимые случайные величины $X_1, \dots, X_n$, распределенные как X , такие что $b_n X + a_n \stackrel{d}{=} X_1 + \dots + X_n$. Последнее эквивалентно каждому из соотношений

$$F\left(\frac{x - a_n}{b_n}\right) = [F(x)]^{*n}, \quad [\varphi(t)]^n = \varphi(b_n t) e^{i t a_n}.$$

Отметим только, что характеристическая функция φ устойчивого распределения также допускает каноническое представление типа (1) (приводить его не будем).

По поводу теории безгранично делимых и устойчивых распределений читатель может обратиться к различным классическим и современным источникам (см., например, [24, 41, 71, 75, 76, 95, 106, 177, 443, 471]).

Рассмотренные далее примеры иллюстрируют свойства безгранично делимых и устойчивых распределений. Два примера посвящены безграничной делимости случайных векторов.

9.1. Не обращающаяся в нуль характеристическая функция, которая не является безгранично делимой

Пусть случайная величина X с характеристической функцией $\varphi(t)$, $t \in \mathbb{R}^1$, безгранично делима. Тогда φ нигде не обращается в нуль. Поэтому можно спросить: если $\varphi(t)$, $t \in \mathbb{R}^1$, — характеристическая функция, которая не обращается в нуль, верно ли, что φ безгранично делима? Ответ получим из следующего примера.

Рассмотрим случайную величину X , принимающую значения -1 , 0 и 1 с вероятностями соответственно $\frac{1}{8}$, $\frac{3}{4}$ и $\frac{1}{8}$. Тогда X имеет характеристическую функцию

$$\varphi(t) = \frac{1}{8} e^{-it} + \frac{3}{4} e^{it \cdot 0} + \frac{1}{8} e^{it} = \frac{1}{4} (3 + \cos t).$$

Очевидно, что $\varphi(t) > 0$ для всех $t \in \mathbb{R}^1$. Тем не менее, X не является безгранично делимой. Достаточно учесть, что значения X лежат в конечном интервале, что эквивалентно ограниченности носителя функции распределения F , и сослаться на упомянутое в начале параграфа свойство (1).

Итак, ответ на поставленный вопрос отрицательный. Однако поучительным является также следующее рассуждение. Предположим, что случайная величина X может быть записана в виде $X \stackrel{d}{=} X_1 + X_2$, где X_1 и X_2 — независимые и одинаково распределенные случайные величины. Поскольку случайная величина X имеет три возможных значения, то каждая из величин X_1 и X_2 может принимать только два значения, скажем, x и y , $x < y$. Пусть $P\{X_1 = x\} = p$ и $P\{X_1 = y\} = 1 - p$ для некоторого p , $0 < p < 1$. Тогда сумма $X_1 + X_2$ принимает значения $2x$, $x + y$ и $2y$ соответственно с вероятностями p^2 , $2p(1 - p)$ и $(1 - p)^2$. Но величина X у нас задана. Таким образом,

$$2x = -1, \quad p^2 = \frac{1}{8}; \quad x + y = 0, \quad 2p(1 - p) = \frac{3}{4}, \quad 2y = 1, \quad (1 - p)^2 = \frac{1}{8}.$$

Очевидно, однако, что эти равенства несовместны. Следовательно, представление $X \stackrel{d}{=} X_1 + X_2$ невозможно, а этого достаточно для за-

ключения, что X безгранично делимой не является несмотря на то, что ее характеристическая функция φ в нуль не обращается.

9.2. О свойстве безграничной делимости характеристических функций φ и $|\varphi|$

Напомним сначала следующий результат (см. [24, 76]): *если φ — безгранично делимая характеристическая функция, то ее абсолютное значение $|\varphi|$ является характеристической функцией, которая также безгранично делима*. В связи с этим результатом возникают следующие два вопроса.

(а) Верно ли утверждение, обратное к сформулированному выше результату? Покажем, что ответ отрицателен, т. е. опишем конкретную характеристическую функцию φ такую, что $|\varphi|$ — безгранично делимая характеристическая функция, причем сама φ свойством безграничной делимости не обладает.

Пусть случайная величина X принимает значения в множестве $\{-1, 0, 1, 2, \dots\}$ с вероятностями

$$P\{X = -1\} = \frac{(1-b)a}{1-a}, \quad P\{X = k\} = \frac{(1-b)(1+ab)b^k}{1-a}, \quad k = 0, 1, 2, \dots,$$

где $0 < a \leq b < 1$. Тогда после некоторого преобразования находим явный вид характеристической функции φ . Именно

$$\varphi(t) = \frac{1-b}{1-a} \cdot \frac{1+ae^{-it}}{1-be^{it}}, \quad t \in \mathbb{R}^1. \quad (1)$$

Покажем, что φ не является безгранично делимой. Для этой цели рассмотрим $\ln \varphi(t)$. Используя (1), можно показать, что

$$\ln \varphi(t) = \sum_{k=1}^{\infty} [(-1)^{k-1} k^{-1} a^k (e^{-itk} - 1) + b^k k^{-1} (e^{itk} - 1)]. \quad (2)$$

Если бы φ была безгранично делимой, то допускала бы представление Леви–Хинчина с параметрами γ и G , где $\gamma = \sum_{k=1}^{\infty} \frac{b^k + (-1)^k a^k}{k^2 + 1}$, а $G(x)$, $x \in \mathbb{R}^1$, — функция ограниченной вариации, имеющая скачок размера $\frac{kb^k}{k^2 + 1}$ при $x = k$ и скачок размера $\frac{(-1)^{k-1} ka^k}{k^2 + 1}$ при $x = -k$, где $k = 1, 2, \dots$. Очевидно, однако, что эта функция G немонотонна, и, следовательно, характеристическая функция φ не может быть безгранично делимой.

Далее, функция

$$\overline{\varphi}(t) = \frac{1-b}{1-a} \cdot \frac{1+ae^{it}}{1-be^{-it}}, \quad t \in \mathbb{R}^1,$$

также является характеристической функцией, не будучи, как и φ , безгранично делимой. Введем еще функцию

$$\psi(t) = |\varphi(t)|^2 = \varphi(t)\overline{\varphi}(t), \quad t \in \mathbb{R}^1,$$

которая является характеристической функцией как произведение двух характеристических функций. Здесь более интересно то, что ψ безгранично делима. Действительно, записывая сначала $\ln \overline{\varphi}(t)$ в форме (2), мы приходим к следующей записи:

$$\begin{aligned} \ln \psi(t) = \sum_{k=1}^{\infty} [k^{-1}(b^k + (-1)^{k-1}a^k)(e^{-itk} - 1)] + \\ + \sum_{k=1}^{\infty} [k^{-1}(b^k + (-1)^{k-1}a^k)(e^{itk} - 1)]. \end{aligned}$$

Следовательно, $\psi(t)$ имеет представление Леви–Хинчина с параметрами γ и G , где $\gamma = 0$, а $G(x)$, $x \in \mathbb{R}^1$, — неубывающая и непрерывная слева функция, которая имеет скачок размера $k \frac{b^k + (-1)^{k-1}a^k}{k^2 + 1}$ в каждой из точек $x = \pm k$, где $k = 1, 2, \dots$. Поскольку найденное выше представление для $\ln \psi(t)$ единственно, то характеристическая функция ψ безгранично делима. Более того, из свойств безгранично делимых распределений (см. [24, 76, 106, 443, 471]) вытекает, что интересующая нас функция $|\varphi|$ (ввиду того, что $|\varphi| = (|\varphi|^2)^{1/2}$) является безгранично делимой характеристической функцией.

Итак, мы показали, что характеристическая функция $|\varphi|$ безгранично делима, в то время как сама функция φ этим свойством не обладает. Мы установили также и следующий интересный факт: существует безгранично делимая характеристическая функция ψ , которая является произведением характеристических функций φ и $\overline{\varphi}$, ни одна из которых не безгранично делима.

(б) По поводу сформулированного вначале результата возникает и такой вопрос: насколько важно условие безграничной делимости φ для того, чтобы $|\varphi|$ была характеристической функцией? Покажем, что это условие существенно.

Рассмотрим случайную величину Y , принимающую каждое из двух значений -1 и 1 с вероятностью $\frac{1}{2}$. Ее характеристическая функция равна $\varphi(t) = \cos t$, $t \in \mathbb{R}^1$. Очевидно, эта функция не является безгранично

делимой. Пусть $\psi(t) = |\varphi(t)|$, $t \in \mathbb{R}^1$, и выясним, какими свойствами обладает $\psi(t)$.

Во-первых, φ имеет ограниченную вторую производную при $t = 0$. Каждая характеристическая функция, удовлетворяющая этому условию, имеет непрерывную первую производную для всех $t \in \mathbb{R}^1$.

Во-вторых, функция $\psi(t) = |\cos t|$ совпадает с $\varphi(t)$ на отрезке $[-\frac{\pi}{2}, \frac{\pi}{2}]$, и если бы ψ была характеристической функцией, то у нее должна была бы существовать непрерывная первая производная для всех $t \in \mathbb{R}^1$. Однако, как нетрудно заметить, в точках $t = \frac{\pi}{2} \pm k\pi$, $k = 0, 1, \dots$, производная $\psi'(t)$ не существует.

Вывод: абсолютное значение $|\varphi|$ характеристической функции φ не обязано быть характеристической функцией. Другими словами, свойство безграничной делимости φ существенно для того, чтобы $|\varphi|$ была безгранично делима.

9.3. Произведение двух независимых неотрицательных и безгранично делимых случайных величин не всегда сохраняет свойство безграничной делимости

В настоящем примере основную роль будет играть распределение Пуассона. Напомним, что случайная величина ξ имеет распределение Пуассона с параметром λ , $\lambda > 0$, если ξ принимает значения в множестве $\mathbb{N}_0 = \{0, 1, 2, \dots\}$ и $P\{\xi = k\} = \frac{\lambda^k e^{-\lambda}}{k!}$. Ее характеристическая функция равна $\varphi(t) = \exp[\lambda(e^{it} - 1)]$. Очевидно, что она безгранично делима.

Для дальнейшего рассуждения нам необходим следующий результат (см. [416]): пусть $\mathcal{P} = \{p_n, n \in \mathbb{N}_0\}$ — вероятностное распределение на множестве $\mathbb{N}_0$, причем $p_0 > 0$, и пусть числа r_k , $k \in \mathbb{N}_0$, определены из соотношений

$$(n+1)p_{n+1} = \sum_{k=0}^n r_k p_{n-k}, \quad n = 0, 1, 2, \dots \quad (1)$$

Распределение $\mathcal{P}$ безгранично делимо тогда и только тогда, когда все числа r_k , $k \in \mathbb{N}_0$, неотрицательные.

Рассмотрим две независимые случайные величины, ξ_1 и ξ_2 , каждая из которых имеет распределение Пуассона с параметром λ . Покажем, что произведение $\eta = \xi_1 \xi_2$ не обладает свойством безграничной делимости.

Пусть целое число $n > 2$ таково, что $n + 1$ — простое число. Тогда

$$\begin{aligned} p_{n+1} &= P\{\eta = n + 1\} = P\{\xi_1 \xi_2 = n + 1\} = \\ &= P\{\xi_1 = 1, \xi_2 = n + 1\} + P\{\xi_1 = n + 1, \xi_2 = 1\} = \\ &= 2P\{\xi_1 = 1\}P\{\xi_2 = n + 1\} = \frac{2\lambda^{n+2}e^{-2\lambda}}{(n+1)!}. \end{aligned}$$

Поскольку само число n четное, оно имеет по крайней мере два разложения: $n = 1 \cdot n = 2 \cdot \frac{n}{2}$. Обозначая $m = \frac{n}{2}$, находим

$$p_n = P\{\eta = n\} > 2 \cdot \frac{\lambda^2 e^{-\lambda}}{2!} \frac{\lambda^m e^{-\lambda}}{m!} = \frac{\lambda^{m+2} e^{-2\lambda}}{m!}. \quad (2)$$

Очевидно, что

$$p_0 = P\{\eta = 0\} = 1 - (1 - e^{-\lambda})^2 > 0, \quad p_1 = \lambda^2 e^{-2\lambda} > 0,$$

и, следовательно, $r_0 = \frac{p_1}{p_0} > 0$. Рассуждая по индукции, допустим, что числа $r_1, \dots, r_{n+1}$ в соотношении (1) неотрицательные. Определим теперь знак числа r_n . Из (1) и (2) нетрудно найти, что

$$\begin{aligned} r_n p_0 &\leq (n+1)p_{n+1} - r_0 p_n < \\ &< \frac{2\lambda^{n+2}e^{-2\lambda}}{n!} - \frac{r_0 \lambda^{m+2}e^{-2\lambda}}{m!} = e^{-2\lambda} \lambda^{m+2} \left(\frac{2\lambda^m}{n!} - \frac{r_0}{m!} \right). \end{aligned}$$

Поскольку $\lambda > 0$ фиксировано, $p_0 > 0$ и $n = 2m$, отсюда вытекает, что для достаточно большого n число r_n будет отрицательным. Таким образом, невозможно, чтобы все числа $r_0, r_1, r_2, \dots$ были неотрицательными, и, следовательно, произведение $\eta = \xi_1 \xi_2$ не может быть безгранично делимым.

9.4. О свойстве безграничной делимости в многомерном случае

Понятие безграничной делимости можно ввести и изучать и для распределений в многомерном пространстве. Основные идеи такие же, как и в одномерном случае.

Пусть на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ задан случайный вектор $X = (X_1, \dots, X_n)$ с функцией распределения $F(x_1, \dots, x_n)$, $(x_1, \dots, x_n) \in \mathbb{R}^n$ и характеристической функцией $\varphi(t_1, \dots, t_n)$, $(t_1, \dots, t_n) \in \mathbb{R}^n$, где $\varphi(t_1, \dots, t_n) = E\{\exp[i(t_1 X_1 + \dots + t_n X_n)]\}$. Вектор X , а также функция распределения F и характеристическая функция φ называются безгранично делимыми, если для каждого $\alpha > 0$ функция

$\varphi^\alpha(t_1, \dots, t_n)$, $(t_1, \dots, t_n) \in \mathbb{R}^n$, снова является характеристической функцией.

Очевидным следствием из этого определения является утверждение о том, что если случайный вектор $(X_1, \dots, X_n)$ безгранично делим, то каждая координата $X_1, \dots, X_n$ и, более того, любой k -мерный вектор $(X_{j_1}, \dots, X_{j_k})$, где $2 \leq k \leq n-1$, $1 \leq j_1 < \dots < j_k \leq n$, также обладают этим свойством. Вполне естественно возникает вопрос о справедливости обратного утверждения. Для простоты рассмотрим случай $n = 3$.

Итак, пусть ξ и η — независимые случайные величины, каждая с распределением $\mathcal{N}(0, 1)$. Положим $X_1 = \xi^2$, $X_2 = 2\xi\eta$, $X_3 = \eta^2$ и рассмотрим 3-мерный вектор (X_1, X_2, X_3) , а также векторы (X_1, X_2) , (X_1, X_3) и (X_2, X_3) .

Во-первых, компоненты X_1 , X_2 и X_3 вектора (X_1, X_2, X_3) безгранично делимы вследствие нормальной распределенности случайных величин ξ и η . Во-вторых, из независимости ξ и η легко следует безграничная делимость вектора (X_1, X_3) . В-третьих, можно установить, что, например, вектор (X_1, X_2) безгранично делим. Для этой цели можно использовать следующий явный вид характеристической функции $\varphi(t_1, t_2, t_3)$ вектора (X_1, X_2, X_3) :

$$\begin{aligned} \varphi(t_1, t_2, t_3) &= \frac{1}{2\pi} \iint_{\mathbb{R}^2} \exp\left(it_1x^2 + 2it_2xy + it_3y^2 - \frac{x^2}{2} - \frac{y^2}{2}\right) dx dy = \\ &= [(1 - 2it_1)(1 - 2it_3) + 4t_2^2]^{-1/2}. \end{aligned}$$

Отсюда следует не только, что каждый из векторов (X_1, X_2) и (X_2, X_3) безгранично делим, но и то, что сам вектор (X_1, X_2, X_3) безгранично делимым не является. Кроме того, любая линейная комбинация $c_1X_1 + c_2X_2 + c_3X_3$, где $c_1, c_2, c_3 \in \mathbb{R}^1$, является безгранично делимой случайной величиной. Подробности читатель может найти в работах [259, 355, 404].

9.5. О безграничной делимости случайного вектора и линейных комбинаций его компонент

Если (X_1, X_2) — безгранично делимый случайный вектор, то любая линейная комбинация $Z = c_1X_1 + c_2X_2$, где $c_1, c_2 \in \mathbb{R}^1$, является безгранично делимой случайной величиной. Возникает вопрос: верно ли обратное утверждение? Оказывается, ответ отрицателен. Вспомним короткое замечание в конце предыдущего примера, где рассматривался 3-мерный случай. Здесь изложен другой пример, для 2-мерного случая.

Для $x = (x_1, x_2) \in \mathbb{R}^2$, $|x| = (x_1^2 + x_2^2)^{1/2}$ и $\varepsilon \in \left(0, \frac{1}{4}\right)$ зададим функцию

$$a_\varepsilon(x) = \begin{cases} 1, & \text{если } |x| \leq \frac{1}{2} - \varepsilon \text{ или } \frac{1}{2} + \varepsilon < |x| \leq 1, \\ 0, & \text{если } |x| > 1, \\ -\varepsilon, & \text{если } \frac{1}{2} - \varepsilon < |x| \leq \frac{1}{2} + \varepsilon. \end{cases}$$

Пусть $A_\varepsilon(U)$, $U \subset \mathbb{R}^2$, — знакопеременная мера с «плотностью», равной a_ε , т. е. $A_\varepsilon(U) = \int_U a_\varepsilon(x) dx$. Введем также функцию

$$\psi_\varepsilon(t) = \exp\left(\int_{\mathbb{R}^2} (e^{i\langle t, x \rangle} - 1) dA_\varepsilon(x)\right),$$

где $t, x \in \mathbb{R}^2$ и $\langle t, x \rangle = t_1 x_1 + t_2 x_2$.

Покажем, что если ε — достаточно малое число, то ψ_ε — положительно определенная функция, и поскольку $\psi_\varepsilon(0) = 1$, то ψ_ε будет характеристической функцией, и, значит, ей будет соответствовать некоторая функция распределения F_ε в пространстве $\mathbb{R}^2$. Действительно, из определения ψ_ε следует, что

$$\psi_\varepsilon(t) = c \left[1 + \sum_{n=1}^{\infty} (n!)^{-1} \left(\int_{\mathbb{R}^2} e^{i\langle t, x \rangle} dA_\varepsilon(x) \right)^n \right],$$

где $c = \exp(-A_\varepsilon(\mathbb{R}^2))$. Таким образом, искомая функция F_ε может быть записана в виде $F_\varepsilon = c(G_0 + G_\varepsilon)$, где G_0 — вырожденная вероятностная мера, $G_0(\{0\}) = 1$, а G_ε — мера с плотностью $\gamma_\varepsilon(x) = \sum_{n=1}^{\infty} \tilde{a}_\varepsilon^{(n)} \frac{x}{n!}$, где $\tilde{a}_\varepsilon^{(n)}(x)$ — n -кратная свертка функции $a_\varepsilon(x)$. Проверим, что $\gamma_\varepsilon(x) \geq 0$ для всех $x \in \mathbb{R}^2$ при малом ε . Оказывается, если число ε выбрано достаточно малым, то

$$\tilde{a}_\varepsilon^{(2)}(x) = a_\varepsilon * a_\varepsilon(x) = \int_{\mathbb{R}^2} a_\varepsilon(x - u) a_\varepsilon(u) du \geq 0,$$

а также

$$\tilde{a}_\varepsilon^{(3)}(x) = \tilde{a}_\varepsilon^{(2)} * a_\varepsilon(x) \geq 0.$$

Следовательно, для $n \geq 4$ мы имеем $\tilde{a}_\varepsilon^{(n)}(x) = \tilde{a}_\varepsilon^{(n-2)} * \tilde{a}_\varepsilon^{(2)}(x) \geq 0$. Кроме того, если ε мало, то функции $a_\varepsilon(x)$ и $\tilde{a}_0^{(2)}(x)$ близки. Легко видеть, что для $\frac{1}{4} \leq x < \frac{3}{4}$ мы имеем $\inf \tilde{a}_0^{(2)}(x) = c_1 > 0$. Таким образом, $\tilde{a}_\varepsilon^{(2)}(x) \geq 2\varepsilon$

для $\frac{1}{2} - \varepsilon < |x| \leq \frac{1}{2} + \varepsilon$. Из этих рассуждений, очевидно, уже вытекает, что $\gamma_\varepsilon(x) \leq 0$ для всех $x \in \mathbb{R}^2$. Следовательно, функция F_ε , введенная выше, действительно является функцией распределения в пространстве $\mathbb{R}^2$.

Пусть теперь (X_1, X_2) — случайный вектор, функция распределения которого равняется F_ε . Поскольку мера A_ε знакопеременная, то функция распределения F_ε не может быть безгранично делимой.

Остается показать, что любая линейная комбинация $c_1 X_1 + c_2 X_2$, $c_1, c_2 \in \mathbb{R}^1$, — это безгранично делимая случайная величина. Действительно, при $s \in \mathbb{R}^1$

$$\begin{aligned} \varphi(s) &= E\{\exp[is(c_1 X_1 + c_2 X_2)]\} = \\ &= \psi_\varepsilon(c_1 s, c_2 s) = \exp\left(\int_{\mathbb{R}^2} (e^{is\langle c, x \rangle} - 1) dA_\varepsilon(x)\right). \end{aligned}$$

Обозначая $\langle c, x \rangle = u$, где $u \in \mathbb{R}^1$, мы можем записать $\varphi(s)$ в виде

$$\varphi(s) = \exp\left(\int_{-\infty}^{\infty} (e^{isu} - 1) dH_c(u)\right)$$

с помощью функции H_c , где

$$dH_c(u) = \left(\int_{\langle c, y \rangle \leq u} dA_\varepsilon(y)\right) du.$$

Так как для достаточно малых ε каждая полоса $\{y: u_1 \leq \langle c, y \rangle \leq u_2\}$ имеет положительную A_ε -меру, то мы можем заключить (подробности см. в работе [288]), что функция $\varphi(s) = \psi_\varepsilon(c_1 s, c_2 s)$, $s \in \mathbb{R}^1$, является характеристической функцией, которая обладает свойством безграничной делимости.

Итак, мы установили, что любая линейная комбинация $c_1 X_1 + c_2 X_2$ безгранично делима, несмотря на то что сам случайный вектор (X_1, X_2) безгранично делимым не является.

9.6. Безгранично делимые распределения, которые не являются устойчивыми

Обычно сначала вводится и изучается класс безгранично делимых распределений, а затем класс устойчивых распределений. Одно из первых устанавливаемых свойств следующее: каждое устойчивое распределение является безгранично делимым. На двух конкретных

примерах (соответственно в дискретном и абсолютно непрерывном случае) покажем, что существуют безгранично делимые распределения, которые не являются устойчивыми.

(а) Пусть случайная величина X имеет *распределение Пуассона* с параметром λ , $\lambda > 0$, т. е. $P\{X = k\} = \frac{\lambda^k e^{-\lambda}}{k!}$, $k = 0, 1, 2, \dots$. Так как X имеет характеристическую функцию

$$\varphi(t) = \exp[\lambda(e^{it} - 1)] = \left(\exp\left[\frac{\lambda}{n}(e^{it} - 1)\right] \right)^n, \quad t \in \mathbb{R}^1, \quad n = 1, 2, \dots,$$

то немедленно заключаем, что X , а также φ безгранично делимы. Хотя имеется более общее утверждение, что любое устойчивое распределение абсолютно непрерывно, полезно показать, почему, например, пуассоновское распределение не является устойчивым. Для этого достаточно проверить, что характеристическая функция φ , выписанная выше, не удовлетворяет соотношению $[(\varphi(t))^n = \varphi(b_n t)e^{ita_n}]$ при любом выборе констант $b_n > 0$ и $a_n \in \mathbb{R}^1$.

Заметим, что здесь можно провести и другое элементарное рассуждение, которое также поучительно. Пусть для простоты $\lambda = 1$. Предположим, что случайная величина X , имеющая распределение Пуассона с параметром 1, является устойчивой. Возьмем тогда две независимые случайные величины X_1 и X_2 , каждая из которых распределена как X . Тогда при подходящем выборе констант $b > 0$ и $a \in \mathbb{R}^1$ случайная величина $Z = \frac{X_1 + X_2}{b + a}$ должна иметь то же самое распределение Пуассона с параметром 1. Отсюда вытекает, что $a = 0$ и $b = 2$. Это, однако, означает, что случайная величина Z принимает и дробные значения с положительной вероятностью, например, $P\left\{Z = \frac{1}{2}\right\} = \frac{1}{e^2}$, что, очевидно, невозможно.

Итак, мы установили, что распределение Пуассона не является устойчивым, несмотря на то что оно безгранично делимое.

(б) Пусть Y — абсолютно непрерывная случайная величина с *распределением Лапласа*, т. е. Y имеет плотность

$$g(x) = \frac{1}{2\lambda} \exp\left(-\frac{|x - \mu|}{\lambda}\right), \quad x \in \mathbb{R}^1,$$

где $\mu \in \mathbb{R}^1$, $\lambda > 0$, и характеристическую функцию

$$\varphi(t) = \frac{e^{it\mu}}{1 + t^2\lambda^2}, \quad t \in \mathbb{R}^1.$$

Нетрудно проверить, что φ безгранично делима. Более того, φ не может удовлетворять соотношению типа $[\varphi(t)]^n = \varphi(b_n t)e^{ita_n}$, $b_n > 0$,

$a_n \in \mathbb{R}^1$. Значит, φ не может быть устойчивой. Следовательно, распределение Лапласа безгранично делимо, но не устойчиво.

Отметим, что еще один случай безгранично делимого, но не устойчивого распределения будет приведен в примере 21.7.

9.7. Устойчивое распределение можно разложить на два безгранично делимых и неустойчивых распределения

Если случайные величины X_1 и X_2 независимы и каждая из них устойчива, то сумма $X_1 + X_2$ также обладает свойством устойчивости. Предположим теперь, что X — устойчивая случайная величина, допускающая разложение $X \stackrel{d}{=} X_1 + X_2$, где X_1 и X_2 — независимые случайные величины. Следует ли отсюда, что X_1 и X_2 устойчивы? Покажем, что ответ отрицателен.

Пусть случайная величина X имеет распределение Коши с плотностью $f(x) = \frac{1}{\pi(1+x^2)}$, $x \in \mathbb{R}^1$. Если φ — характеристическая функция X , то $\varphi(t) = e^{-|t|}$, $t \in \mathbb{R}^1$. Хорошо известно, что φ устойчива. Покажем, что справедливо представление $X \stackrel{d}{=} X_1 + X_2$, в котором X_1 и X_2 — независимые случайные величины, причем обе являются безгранично делимыми, но ни одна не является устойчивой.

Рассмотрим следующие две функции:

$$\varphi_1(t) = \exp(-|t| + 1 - e^{-|t|}), \quad \varphi_2(t) = \exp(e^{-|t|} - 1), \quad t \in \mathbb{R}^1.$$

каждая из этих функций может быть представлена как

$$\exp\left(-\int_0^t \int_0^u \psi(v) dv du\right)$$

с подходящей функцией ψ , где ψ — характеристическая функция (детали см. в [76]). Отсюда вытекает, что φ_1 и φ_2 — характеристические функции и, более того, они безгранично делимы. Непосредственно проверяется, что φ_1 и φ_2 не являются устойчивыми.

Наконец, пусть независимые случайные величины X_1 и X_2 таковы, что X_1 имеет характеристическую функцию φ_1 , а X_2 — характеристическую функцию φ_2 . Тогда мы показали, что

$$\varphi(t) = e^{-|t|} = \varphi_1(t)\varphi_2(t), \quad t \in \mathbb{R}^1,$$

где φ — это характеристическая функция суммы $X_1 + X_2$.

Итак, $X_1 + X_2$ — устойчивая случайная величина, а X_1 и X_2 — безгранично делимые, но не устойчивые случайные величины.

§ 10. Нормальное распределение

Напомним, что через $\mathcal{N}(a, \sigma^2)$ принято обозначать класс нормально распределенных случайных величин с параметрами a и σ^2 , где $a \in \mathbb{R}^1$, $\sigma > 0$. Запись $X \sim \mathcal{N}(a, \sigma^2)$ означает, что X — случайная величина с плотностью

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right), \quad x \in \mathbb{R}^1.$$

Особенно важным является случай, когда $a = 0$ и $\sigma = 1$. Очень часто в дальнейшем изложении будут использоваться функции

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right), \quad x \in \mathbb{R}^1, \quad (1)$$

и

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{1}{2}u^2\right) du, \quad x \in \mathbb{R}^1, \quad (2)$$

которые называются соответственно *стандартной нормальной плотностью* и *стандартной нормальной функцией распределения*. Они отвечают классу $\mathcal{N}(0, 1)$.

Если $X \sim \mathcal{N}(a, \sigma^2)$, то $EX = a$, $DX = \sigma^2$, а характеристическая функция величины X равна $\psi(t) = \exp\left(iat - \frac{\sigma^2 t^2}{2}\right)$. Очевидно, что X — симметричная случайная величина относительно точки $x = a$. При $a = 0$ получаем простые выражения для моментов X , именно

$$m_{2n-1} = E\{X^{2n+1}\} = 0, \quad m_{2n} = E\{X^{2n}\} = \sigma^{2n}(2n-1)!!, \quad n = 1, 2, \dots$$

Рассмотрим теперь случайный вектор $X = (X_1, \dots, X_n)$. Если $EX_j = a_j$, $j = 1, \dots, n$, то вектор $a = (a_1, \dots, a_n)$ называется математическим ожиданием X . Матрица $C = (c_{jk})$, где $c_{jk} = E\{(X_j - a_j)(X_k - a_k)\}$, $j, k = 1, \dots, n$, называется ковариационной матрицей X . Мы говорим, что X имеет *n -мерное нормальное распределение*, если соответствующая совместная плотность $f(x_1, \dots, x_n)$, $(x_1, \dots, x_n) \in \mathbb{R}^n$ имеет вид

$$f(x_1, \dots, x_n) = \frac{|D|^{1/2}}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2} \sum_{j,k=1}^n d_{jk}(x_j - a_j)(x_k - a_k)\right). \quad (3)$$

Здесь $D = (d_{jk})$ — матрица, обратная к матрице C , а $|D| = \det D$. Ясно, что D существует, только если C положительно определена.

Важно отметить, что приведенные выше определения применимы только в случае невырожденных распределений. Однако в терминах

соответствующих характеристических функций можно ввести общее определение нормального распределения, включающее как невырожденный, так и вырожденный случай. Именно случайный вектор X имеет n -мерное нормальное распределение, если его характеристическая функция ψ имеет следующий вид: $\psi(t) = \exp\left(i\langle a, t \rangle - \frac{1}{2}\langle Ct, t \rangle\right)$, $t \in \mathbb{R}^n$, где $a = (a_1, \dots, a_n)$ — числовой вектор с $|a_j| < \infty$, а $C = (c_{jk})$ — симметрическая неотрицательно определенная матрица порядка n . Оказывается, что при этом a и C — соответственно вектор математического ожидания X и ковариационная матрица X .

Для последующих примеров нам нужен явный вид двумерной нормальной плотности:

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(-\frac{1}{2(1-\rho^2)}\left[\frac{(x_1-a_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-a_1)(x_2-a_2)}{\sigma_1\sigma_2} + \frac{(x_2-a_2)^2}{\sigma_2^2}\right]\right), \quad (x_1, x_2) \in \mathbb{R}^2. \quad (4)$$

Здесь $a_1, a_2 \in \mathbb{R}^1$, $\sigma_1 > 0$, $\sigma_2 > 0$ и $|\rho| < 1$. Если (X_1, X_2) — случайный вектор с плотностью (4), то $EX_1 = a_1$, $EX_2 = a_2$, $DX_1 = \sigma_1^2$, $DX_2 = \sigma_2^2$, а ρ равняется коэффициенту корреляции $\rho(X_1, X_2)$.

Из-за важности нормального распределения (и в одномерном, и в многомерном случаях) ему уделяется значительное внимание в любом курсе теории вероятностей и в любой книге в области вероятности и статистики. По этому поводу читатель может обратиться, например, к книгам [1, 9, 13, 26, 75, 92, 96, 106, 116, 117, 145, 231, 177, 330, 345, 374, 386, 391, 392, 404, 405, 463].

Приводимые далее примеры касаются разнообразных свойств нормального распределения.

10.1. Нормальность одномерных распределений не обеспечивает нормальности соответствующего многомерного распределения

(а) Возьмем две независимые случайные величины ξ_1 и ξ_2 , каждая с распределением $\mathcal{N}(0, 1)$ и рассмотрим двумерный случайный вектор

$$(X_1, X_2) = \begin{cases} (\xi_1, |\xi_2|), & \text{если } \xi_1 \geq 0, \\ (\xi_1, -|\xi_2|), & \text{если } \xi_1 < 0. \end{cases}$$

Легко видеть, что распределение вектора (X_1, X_2) не является нормальным, в то время как каждая из компонент X_1 и X_2 имеет нормальное распределение.

(б) Пусть $r(x) = (2\pi e)^{-1/2} x^3 I_{[-1,1]}(x)$, $x \in \mathbb{R}^1$, где $I_{[-1,1]}(\cdot)$ — индикаторная функция отрезка $[-1, 1]$. Очевидно, что $|r(x)| \leq (2\pi e)^{-1/2}$. Пусть

$$f(x_1, x_2) = \varphi(x_1)\varphi(x_2) + r(x_1)r(x_2), \quad x_1, \quad x_2 \in \mathbb{R}^1.$$

Нетрудно проверить, что f — двумерная вероятностная плотность. Очевидно, что f не соответствует двумерному нормальному распределению. Однако, как легко заметить, каждая из маргинальных плотностей $f(x_1)$, $x_1 \in \mathbb{R}^1$, и $f_2(x_2)$, $x_2 \in \mathbb{R}^1$, является нормальной.

(в) Для любого числа ε , $|\varepsilon| \leq 1$, задаем функцию

$$H(x, y) = \Phi(x)\Phi(y)[1 + \varepsilon(1 - \Phi(x))(1 - \Phi(y))], \quad (x, y) \in \mathbb{R}^2,$$

где Φ — стандартная нормальная функция распределения. Тогда H — двумерная функция распределения, и пусть ей соответствует случайный вектор (X, Y) . Легко показать, что X имеет функцию распределения $\Phi(x)$, $x \in \mathbb{R}^1$, а Y имеет функцию распределения $\Phi(y)$, $y \in \mathbb{R}^1$. Очевидно, что если $\varepsilon \neq 0$, то распределение H вектора (X, Y) не является нормальным. Однако каждая из компонент X и Y имеет нормальное распределение.

Такой же вывод получаем, если вместо H рассмотрим плотность

$$h(x, y) = \varphi(x)\varphi(y)[1 + \varepsilon(2\Phi(x) - 1)(2\Phi(y) - 1)], \quad (x, y) \in \mathbb{R}^2,$$

где φ — стандартная нормальная плотность, а $|\varepsilon| \leq 1$. Здесь все выражается в терминах плотностей.

10.2. Если (X_1, X_2) имеет двумерное нормальное распределение, то X_1 , X_2 и $X_1 + X_2$ нормально распределены, но обратное неверно

Пусть случайный вектор (X_1, X_2) имеет нормальное распределение с плотностью $f(x_1, x_2)$, $(x_1, x_2) \in \mathbb{R}^2$ (см. формулу (4) в начале параграфа). Тогда X_1 , X_2 и $X_1 + X_2$ имеют одномерные нормальные распределения. Вопрос: верно ли обратное утверждение? Оказывается, ответ отрицателен.

Предположим, что X_1 и X_2 — независимые случайные величины, каждая с распределением $\mathcal{N}(0, 1)$. Тогда их совместная плотность $f(x_1, x_2) = \varphi(x_1)\varphi(x_2)$, $(x_1, x_2) \in \mathbb{R}^2$, где φ — стандартная нормальная плотность. Воспользуемся теперь свойством симметричности функции $f(x_1, x_2)$ относительно каждого из аргументов x_1 и x_2 , чтобы определить специальным образом новую функцию, которую обозначим g .

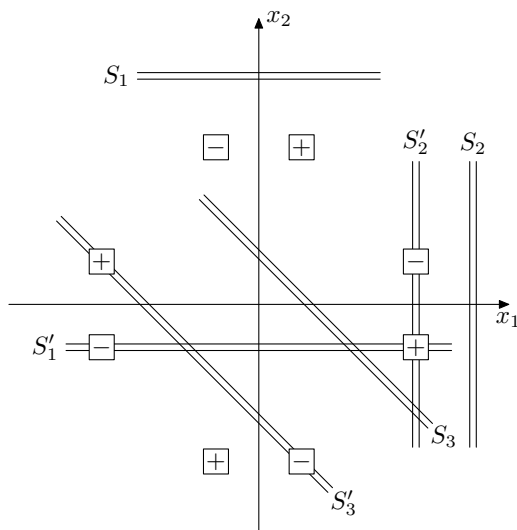


Рис. 2.

Пусть Ox_1x_2 — координатная система на плоскости. На фиксированном расстоянии от начала O нарисуем восемь одинаковых квадратиков, расположенных симметрично относительно осей Ox_1 и Ox_2 и помеченных знаком $(+)$ или $(-)$ так, как показано на рис. 2. Пусть ε — малое положительное число, выражающее ту «массу», которую мы перебрасываем из четырех квадратиков с $(-)$ соответственно в четыре квадратика с $(+)$. Через Q^+ обозначим объединение квадратиков с $(+)$, а через Q^- — объединение квадратиков с $(-)$. Определим функцию $g(x_1, x_2)$ следующим образом:

$$g(x_1, x_2) = \begin{cases} f(x_1, x_2) + \varepsilon, & \text{если } (x_1, x_2) \in Q^+, \\ f(x_1, x_2) - \varepsilon, & \text{если } (x_1, x_2) \in Q^-, \\ f(x_1, x_2) & \text{в остальных случаях.} \end{cases}$$

В зависимости от расстояния квадратиков от начала O мы можем выбрать число ε настолько малым, чтобы $g(x_1, x_2) \geq 0$ для всех $(x_1, x_2) \in \mathbb{R}^2$. Таким образом, учитывая, что интеграл от g по пространству $\mathbb{R}^2$ равен 1, мы заключаем, что g является двумерной вероятностной плотностью. Предположим, что плотность g соответствует случайному вектору (Y_1, Y_2) . Найдем теперь распределение каждой из случайных величин Y_1 , Y_2 и $Y_1 + Y_2$. Для этого удобно рассмотреть полосы $S_1, \dots, S'_3$ так, как они изображены на рис. 2. Каждая из этих

полос может иметь произвольную ширину и быть произвольно расположенной, но обязательно параллельной либо Ox_1 , либо Ox_2 , либо биссектрисе второго и четвертого квадрантов. Очевидно, что или полоса не пересекает ни одного из квадратиков, или она пересекает ровно два из них, причем один помечен знаком $(+)$, а второй — знаком $(-)$. Если S — любая из полос на рис. 2, то из определения g вытекает, что

$$\iint_S f(x_1, x_2) dx_1 dx_2 = \iint_S g(x_1, x_2) dx_1 dx_2.$$

Отсюда, так как $f(x_1, x_2) = \varphi(x_1)\varphi(x_2)$ есть плотность вектора (X_1, X_2) , мы легко заключаем, что $Y_1 \sim \mathcal{N}(0, 1)$, $Y_2 \sim \mathcal{N}(0, 1)$ и $Y_1 + Y_2 \sim \mathcal{N}(0, 2)$. В то же самое время распределение вектора (Y_1, Y_2) , задаваемое плотностью g , не является нормальным.

Итак, мы показали, что нормальность распределения величин Y_1, Y_2 и $Y_1 + Y_2$ не влечет нормальности распределения вектора (Y_1, Y_2) .

10.3. Свойство нормальности может не иметь места для системы случайных величин, даже если любая ее подсистема обладает этим свойством

Если $X_1, \dots, X_n$, $n \geq 3$, — система случайных величин, совместное распределение которых нормально, то любая ее подсистема, т. е. любые $X_{j_1}, \dots, X_{j_k}$, где $2 \leq k \leq n-1$, $1 \leq j_1 < \dots < j_k \leq n$, также имеет совместное нормальное распределение. Проверим, верно ли обратное утверждение.

Пусть случайная величина X имеет распределение $\mathcal{N}(a, \sigma^2)$, а f обозначает ее плотность, $f(x) = \frac{1}{\sqrt{2\pi}\sigma} \cdot \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right)$. Рассмотрим следующую функцию:

$$g_n(x_1, \dots, x_n) = \left(\prod_{j=1}^n f(x_j) \right) \left(1 + \prod_{k=1}^n (x_k - a)f(x_k) \right), \quad (1)$$

$$(x_1, \dots, x_n) \in \mathbb{R}^n.$$

Во-первых, проверим, что g_n — вероятностная плотность. Используя явный вид f , находим, что $\int_{-\infty}^{\infty} (x-a)f^2(x) dx = 0$ (см. формулу (5) в примере 7.2), и легко приходим к заключению, что g_n — неотрицательная функция с интегралом по всему пространству $\mathbb{R}^n$, равным 1. Следовательно, функция g_n , заданная равенством (1), является вероятностной плотностью некоторого случайного вектора $(X_1, \dots, X_n)$. Пусть $n \geq 3$.

Выберем теперь k величин из $X_1, \dots, X_n$, где $2 \leq k \leq n-1$. Не ограничивая общности, можно считать, что мы выбрали случайные величины $X_1, \dots, X_k$. Если g_k — плотность $(X_1, \dots, X_k)$, то из (1) находим

$$g_k(x_1, \dots, x_k) = f(x_1) \dots f(x_k), \quad (x_1, \dots, x_k) \in \mathbb{R}^k. \quad (2)$$

Поскольку f — нормальная плотность, совместное распределение величин $X_1, \dots, X_k$ является нормальным. Более того, из (2) вытекает также, что случайные величины $X_1, \dots, X_k$ независимы.

Итак, мы описали систему из n случайных величин $X_1, \dots, X_n$ такую, что любая подсистема $X_{j_1}, \dots, X_{j_k}$, где $2 \leq k \leq n-1$, имеет нормальное распределение, и тем не менее распределение самой системы $X_1, \dots, X_n$ нормальным не является.

10.4. О связи между свойствами нормальности и некоррелированности

Пусть случайный вектор (X, Y) имеет нормальное распределение. Тогда каждая из компонент X и Y также нормально распределена, и если X и Y некоррелированы, то они независимы. Этот факт хорошо известен. Из рассмотренных далее примеров станет ясно, насколько важна нормальность распределения (X, Y) .

(а) Пусть $X \sim \mathcal{N}(0, 1)$. Для фиксированного $c \geq 0$ определим случайную величину

$$Y = \begin{cases} X, & \text{если } |X| \leq c, \\ -X, & \text{если } |X| > c. \end{cases}$$

Легко проверить, что $Y \sim \mathcal{N}(0, 1)$ для каждого c . Далее,

$$\mathbf{E}\{XY\} = \mathbf{E}\{X^2 I(|X| \leq c)\} - \mathbf{E}\{X^2 I(|X| > c)\}.$$

Отсюда вытекает, что $\mathbf{E}\{XY\} = -1$, если $c = 0$, и $\mathbf{E}\{XY\} \rightarrow 1$, если $c \rightarrow \infty$. Поскольку $\mathbf{E}\{XY\}$ зависит непрерывным образом от c , то существует c_0 , для которого $\mathbf{E}\{XY\} = 0$. При этом c_0 величины X и Y некоррелированы. Однако

$$\mathbf{P}\{X > c, Y > c\} = 0 \neq \mathbf{P}\{X > c\}\mathbf{P}\{Y > c\},$$

и, следовательно, X и Y не являются независимыми.

(б) Пусть $f_1(x, y)$, $(x, y) \in \mathbb{R}^2$, — двумерная нормальная плотность, заданная формулой (4) (см. начало настоящего параграфа), в которой $a_1 = a_2 = 0$, $\sigma_1^2 = \sigma_2^2 = 1$, а коэффициент корреляции $\rho = \rho_1$.

Пусть, далее, $f_2(x, y)$, $(x, y) \in \mathbb{R}^2$, — такая же нормальная плотность, но с коэффициентом корреляции $\rho = \rho_2$. Рассмотрим следующую функцию:

$$g(x, y) = c_1 f_1(x, y) + c_2 f_2(x, y), \quad (x, y) \in \mathbb{R}^2,$$

где c_1 и c_2 — произвольные числа, $c_1 \geq 0$, $c_2 \geq 0$, $c_1 + c_2 = 1$.

Функция g , как это легко заметить, является плотностью, и очевидно, что если $\rho_1 \neq \rho_2$, то g не является нормальной плотностью. Через (X, Y) обозначим случайный вектор, плотность которого равняется g . Тогда $X \sim \mathcal{N}(0, 1)$, $Y \sim \mathcal{N}(0, 1)$. Кроме того, для коэффициента корреляции $\rho(X, Y)$ между X и Y находим, что $\rho(X, Y) = c_1 \rho_1 + c_2 \rho_2$. Теперь понятно, что можно так выбрать параметры c_1 , c_2 , ρ_1 и ρ_2 , чтобы $c_1 \rho_1 + c_2 \rho_2 = 0$. А это означает, что нормально распределенные случайные величины X и Y еще и некоррелированы. Заметим, однако, что величины X и Y не являются независимыми. Подобным образом обстоит дело и в следующем примере (в).

(в) Предположим, что случайный вектор (X, Y) имеет плотность

$$f(x, y) = \frac{1}{2\pi\sqrt{3}} \left(\exp\left[-\frac{2}{3}(x^2 + xy + y^2)\right] + \exp\left[-\frac{2}{3}(x^2 - xy + y^2)\right] \right),$$

$(x, y) \in \mathbb{R}^2$. Очевидно, что эта плотность не является нормальной. Однако прямое вычисление показывает, что $X \sim \mathcal{N}(0, 1)$, $Y \sim \mathcal{N}(0, 1)$ и $E\{XY\} = 0$. Таким образом, X и Y — нормально распределены, некоррелированы и зависимы.

(г) Рассмотрим 4-мерный случайный вектор $(\xi_1, \xi_2, \xi_3, \xi_4)$, который имеет нормальное распределение с нулевым вектором математического ожидания и с ковариационной матрицей

$$C = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}.$$

(В качестве упражнения читателю предлагается проверить, что эта симметрическая матрица C положительно определена.) Рассмотрим две новые случайные величины $X = \xi_1 + i\xi_2$ и $Y = \xi_3 + i\xi_4$, где $i = \sqrt{-1}$. Мы хотим выяснить, имеется ли зависимость между случайными величинами X и Y . Очевидно, $EX = 0$ и $EY = 0$. Далее, поскольку X и Y принимают комплексные значения, то ковариация между ними равняется $E\{X\bar{Y}\}$, где случайная величина $\bar{Y}$ — комплексно сопряженная

к величине Y . Имеем

$$E\{X\bar{Y}\} = E\{\xi_1\xi_3 + \xi_2\xi_4\} + iE\{\xi_2\xi_3 - \xi_1\xi_4\} = 0 + i(-1 + 1) = 0.$$

Следовательно, X и Y — некоррелированные случайные величины. Спрашивается, а являются ли X и Y независимыми? Предположим, что это так. Тогда, как следствие, отсюда вытекало бы, что, например, случайные величины ξ_1 и ξ_4 независимы. Однако, как это видно из матрицы C , величины ξ_1 и ξ_4 даже не являются некоррелированными, и, значит, они не могут быть независимыми. Полученное противоречие показывает, что случайные величины X и Y зависимы.

10.5. Случайные величины X , Y , $X + Y$ и $X - Y$ могут быть нормально распределенными, причем X и Y некоррелированы, а распределение вектора (X, Y) может не быть нормальным

Напомним, что если (X, Y) имеет двумерное нормальное распределение, то каждая из величин X , Y , $X + Y$ и $X - Y$ нормально распределена. Зададим исходную двумерную нормальную плотность так, чтобы $E\{XY\} = 0$, т. е. чтобы компоненты X и Y были некоррелированы. Пусть теперь вектор (X, Y) имеет следующую плотность:

$$f(x, y) = \frac{1}{2\pi} \exp\left(-\frac{1}{2}(x^2 + y^2)\right) \left\{1 + xy(x^2 - y^2) \exp\left(-\frac{1}{2}(x^2 + y^2 + 2\varepsilon)\right)\right\},$$

где $(x, y) \in \mathbb{R}^2$, а константа $\varepsilon > 0$ выбрана так, чтобы выполнялось неравенство

$$\left|xy(x^2 - y^2) \exp\left(-\frac{1}{2}(x^2 + y^2 + 2\varepsilon)\right)\right| \leq 1, \quad (x, y) \in \mathbb{R}^2.$$

Наша цель теперь — установить, что f — двумерная плотность. Для этого найдем функцию ψ , являющуюся преобразованием Фурье от функции f :

$$\begin{aligned} \psi(s, t) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp(isx + ity) f(x, y) dx dy = \\ &= \exp\left(-\frac{1}{2}(s^2 + t^2)\right) + \frac{1}{32}st(s^2 - t^2) \exp\left(-\varepsilon - \frac{1}{4}(s^2 + t^2)\right), \quad (s, t) \in \mathbb{R}^2. \end{aligned}$$

Отсюда можно сделать следующие выводы:

(1) Поскольку $\psi(0, 0) = 1$ и ψ — положительно определенная функция, то ψ — характеристическая функция. Следовательно, f является двумерной плотностью, и пусть это плотность случайного вектора (X, Y) .

- (2) Так как $\psi(t, 0) = \psi(0, t) = e^{-t^2/2}$, то $X \sim \mathcal{N}(0, 1)$ и $Y \sim \mathcal{N}(0, 1)$.
 (3) $\psi(t, t) = e^{-t^2}$, и, значит, сумма $X + Y \sim \mathcal{N}(0, 2)$.
 (4) Разность $X - Y$ также нормально распределена.
 (5) Величины X и Y некоррелированы.

Но, несмотря на все эти свойства, случайный вектор (X, Y) имеет распределение, отличающееся от нормального.

10.6. Условие, характеризующее нормальное распределение через нормальность линейных комбинаций, не может быть ослаблено

Начнем с формулировки следующего результата (его доказательство см., например, в [271]): *пусть Y и X — случайные величины, а $\{(a_k, b_k), k = 1, 2, \dots\}$ — счетная совокупность «различных» последовательностей в пространстве $\mathbb{R}^2$ таких, что $a_k X + b_k Y$ имеет нормальное распределение для каждого k . Тогда случайный вектор (X, Y) имеет двумерное нормальное распределение.* (Поясним, что «различные» означает, что параметрические уравнения $t_1 = a_k t, t_2 = b_k t$ описывают бесконечное число прямых на плоскости $\mathbb{R}^2$.)

Можно ли ослабить условия, приведенные в этом результате? Более точно, пусть выполнено следующее условие:

(C_N) Зафиксируем натуральное число N , и пусть для некоторых N пар (a_k, b_k) , $k = 1, \dots, N$, линейные комбинации $a_k X + b_k Y$, $k = 1, \dots, N$, нормально распределены.

Вопрос: следует ли из условия (C_N) , что (X, Y) имеет двумерное нормальное распределение? Ответ отрицательный, как это следует из примера ниже.

Рассмотрим следующую функцию:

$$\psi(s, t) = \exp\left(-\frac{1}{2}(s^2 + t^2)\right) + \left(\prod_{k=1}^N (b_k^2 s - a_k^2 t)\right) \exp\left(-\varepsilon - \frac{1}{2}c(s^2 + t^2)\right),$$

где $s, t \in \mathbb{R}^1$; $\varepsilon, c \in \mathbb{R}^+$. Покажем сначала, что при подходящем выборе ε и c функция ψ будет характеристической функцией некоторого двумерного распределения. Действительно, если f обозначает обратное преобразование Фурье от функции ψ , то

$$\begin{aligned} f(x, y) &= \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-isx - ity} \psi(s, t) ds dt = \\ &= \frac{1}{2\pi} \exp\left(-\frac{1}{2}(x^2 + y^2)\right) + h_{\varepsilon, c}(x, y), \end{aligned}$$

где для удобства принято обозначение

$$h_{\varepsilon,c}(x,y) = \frac{1}{(2\pi)^2} \cdot \frac{1}{e^\varepsilon} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-isx-it y} e^{-c(s^2+t^2)/2} \prod_{k=1}^N (b_k^2 s^2 - a_k^2 t^2) ds dt.$$

Для дальнейшего рассуждения нам нужна следующая оценка сверху для модуля функции $h_{\varepsilon,c}(x,y)$, которую читатель может вывести самостоятельно, либо найти в работе [271]:

$$|h_{\varepsilon,c}(x,y)| \leq \frac{1}{2\pi} \exp\left(-\frac{1}{2}(x^2+y^2)\right) \quad \text{для всех } (x,y) \in \mathbb{R}^2,$$

и это имеет место при подходяще выбранных ε и c .

Учитывая, что непрерывная функция $\psi(s,t)$ удовлетворяет условию $\psi(0,0) = 1$, а f принимает действительные значения, мы приходим к выводу, что f и ψ — это пара функций, где f — двумерная вероятностная плотность, а ψ — соответствующая ей характеристическая функция. Обозначим через (ξ, η) случайный вектор, плотность которого равна f . Из определения $\psi(s,t)$ немедленно вытекает, что

$$\psi(a_k t, b_k t) = \exp\left(-\frac{1}{2}(a_k^2 + b_k^2)t^2\right), \quad k = 1, \dots, N.$$

Это означает, что случайная величина $a_k \xi + b_k \eta$ имеет распределение $\mathcal{N}(0, a_k^2 + b_k^2)$ для каждого $k = 1, \dots, N$. Заметим, однако, что характеристическая функция ψ не соответствует двумерному нормальному распределению.

Следовательно, мы построили пару случайных величин ξ и η , для которых условие (C_N) выполнено, но тем не менее вектор (ξ, η) не имеет нормального распределения. Таким образом, одного условия (C_N) недостаточно для нормальности распределения (ξ, η) .

Отметим, наконец, еще одно важное обстоятельство. Именно, приведенная выше оценка для $|h_{\varepsilon,c}(s,t)|$ имеет место для всех $s, t \in \mathbb{R}^1$ при подходящих ε и c , но существенно и то, что число N конечно.

10.7. Распределение, отличающееся от нормального, может иметь условные нормальные распределения

(а) Если $f(x,y)$, $(x,y) \in \mathbb{R}^2$, — двумерная нормальная плотность, то легко проверить, что каждая из условных плотностей $f_1(x|y)$ и $f_2(y|x)$ также является нормальной. Этот факт приводит к вопросу о справедливости обратного утверждения. Чтобы найти ответ, достаточно рассмотреть следующий конкретный пример. Пусть

$$g(x,y) = C \exp\left[-(1+x^2)(1+y^2)\right], \quad (x,y) \in \mathbb{R}^2,$$

где нормирующая константа C выбрана так, что интеграл от g по пространству $\mathbb{R}^2$ равен 1. Стандартное вычисление показывает, что условные плотности $g_1(x|y)$ и $g_2(y|x)$ выражаются так:

$$g_1(x|y) = \frac{1}{\sqrt{2\pi\sigma_y^2}} \exp\left(-\frac{x^2}{2\sigma_y^2}\right), \quad g_2(y|x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left(-\frac{y^2}{2\sigma_x^2}\right),$$

где $\sigma_y^2 = \frac{1}{2(1+y^2)}$, $\sigma_x^2 = \frac{1}{2(1+x^2)}$, $x \in \mathbb{R}^1$, $y \in \mathbb{R}^1$.

Очевидно, что $g_1(x|y)$ и $g_2(y|x)$ — это нормальные плотности, соответствующие нормальным распределениям $\mathcal{N}(0, \sigma_y^2)$ и $\mathcal{N}(0, \sigma_x^2)$. Однако очевидно, что двумерная плотность $g(x, y)$ нормальной не является.

Следовательно, в общей ситуации из нормальности условных плотностей не вытекает нормальность соответствующей двумерной плотности. Можно показать (см. [169]), что любая двумерная плотность, которая не является нормальной, но у которой условные плотности нормальные, имеет следующий вид:

$$\tilde{g}(x, y) = \tilde{C} \exp\left(-\sum_{i,j=0}^2 b_{ij} x^i y^j\right), \quad (x, y) \in \mathbb{R}^2,$$

при подходящих коэффициентах b_{ij} и соответствующей нормирующей константе $\tilde{C}$. Частный случай $\tilde{g}$ — это рассмотренная выше функция g .

(б) Рассмотрим еще один интересный пример. Пусть случайная величина ξ имеет равномерное распределение на отрезке $[0, 1]$, а $\eta_1, \dots, \eta_5$ — независимые случайные величины, каждая с распределением $\mathcal{N}(0, 1)$. Предполагается также, что ξ и $\eta_1, \dots, \eta_5$ независимы. Положим

$$X_1 = \sqrt{\xi} \eta_1 + \sqrt{1-\xi} \eta_2, \quad X_2 = \sqrt{\xi} \eta_3 + \sqrt{1-\xi} \eta_2, \quad X_3 = \sqrt{\xi} \eta_4 + \sqrt{1-\xi} \eta_5.$$

Нетрудно проверить, что каждая из случайных величин X_1 , X_2 и X_3 имеет стандартное нормальное распределение $\mathcal{N}(0, 1)$. Далее, если через $\psi_{3|1,2}(t)$ обозначить условную характеристическую функцию величины X_3 при заданных X_1 и X_2 , то мы найдем

$$\psi_{3|1,2}(t) = \mathbb{E}\{e^{itX_3} \mid X_1, X_2\} = \mathbb{E}\{\mathbb{E}\{e^{itX_3} \mid X_1, X_2, \xi\}\} = e^{-t^2/2}.$$

Следовательно, X_3 (условно относительно X_1 и X_2) имеет нормальное распределение $\mathcal{N}(0, 1)$. Вообще говоря, при этих свойствах мы можем предположить, что случайный вектор (X_1, X_2, X_3) имеет трехмерное

нормальное распределение. Чтобы проверить это, найдем, например, характеристическую функцию $\psi(t_1, t_2)$ величин X_1 и X_2 . Имеем

$$\begin{aligned}\psi(t_1, t_2) &= E\{\exp(it_1 X_1 + it_2 X_2)\} = E\{E\{\exp(it_1 X_1 + it_2 X_2) \mid \xi\}\} = \\ &= E\left\{\exp\left(-\frac{1}{2} t_1^2 \xi - \frac{1}{2} t_2^2 \xi - \frac{1}{2} (t_1 + t_2)^2 (1 - \xi)\right)\right\} = \\ &= \exp\left(-\frac{1}{2} (t_1 + t_2)^2\right) E\{\exp(t_1 t_2 \xi)\} = (t_1 t_2)^{-1} (e^{t_1 t_2} - 1) e^{-(t_1 + t_2)^2 / 2}.\end{aligned}$$

Эта форма характеристической функции $\psi(t_1, t_2)$ показывает, что вектор (X_1, X_2) не может иметь двумерное нормальное распределение. Таким образом, распределение вектора (X_1, X_2, X_3) не является нормальным, несмотря на нормальность каждой из компонент X_1 , X_2 и X_3 и условной нормальности X_3 при заданных X_1 и X_2 .

§ 11. Проблема моментов

Пусть F — функция распределения некоторой случайной величины ξ , принимающей значения в интервале I , $I \subset \mathbb{R}^1$. Предположим, что ξ имеет конечные моменты $m_k = E\{\xi^k\} = \int_I x^k dF(x)$ для любого k , $k = 1, 2, \dots$. *Проблему моментов* можно сформулировать следующим образом: определяется ли функция распределения F однозначно по последовательности моментов $\{m_1, m_2, \dots\}$? Другими словами, если F и $\tilde{F}$ — две функции распределения, у которых все моменты существуют и совпадают, т. е. $\int x^k dF = \int x^k d\tilde{F}$ для всех $k \geq 1$, то вытекает ли отсюда совпадение F и $\tilde{F}$? Если $F = \tilde{F}$, то говорим, что имеет место *единственность проблемы моментов* или, что проблема моментов имеет единственное решение. В противном случае проблема моментов имеет неединственное решение.

Если интервал I конечен, то проблема моментов имеет единственное решение (см. [2, 58, 106, 116, 177]). Поэтому в принципе неединственность может возникнуть, если I — бесконечный интервал. В зависимости от I проблема моментов получила специальные названия: если $I = [0, \infty)$, мы говорим о *проблеме моментов Стильтьеса*, а в случае $I = \mathbb{R}^1 = (-\infty, \infty)$ — о *проблеме моментов Хамбургера* (об истории вопроса см. [434]).

Имеются различные условия, обеспечивающие единственность или неединственность проблемы моментов. Сформулируем теперь следующие три критерия.

Критерий (C_1). Пусть $F(x)$, $x \in \mathbb{R}^1$, — функция распределения и ей соответствует характеристическая функция $\psi(t)$, $t \in \mathbb{R}^1$. Если ψ яв-

ляется r -аналитической функцией для некоторого $r > 0$, то F определяется однозначно своими моментами $\{m_k\}$, где $m_k = \int_{-\infty}^{\infty} x^k dF(x)$.

При этом ψ — r -аналитическая функция тогда и только тогда, когда выполнено соотношение

$$\overline{\lim}_{k \rightarrow \infty} \frac{(m_{2k})^{1/(2k)}}{2k} < \infty.$$

Напомним, что r -аналитичность характеристической функции ψ эквивалентна существованию производящей функции моментов $M(t)$, $|t| < t_0$ при некотором $t_0 > 0$. Обычно это называется *условием Крамера*.

Критерий (C₂). Пусть $\{m_0 = 1, m_1, m_2, \dots\}$ — последовательность моментов, соответствующая функции распределения $F(x)$, $x \in \mathbb{R}^1$. Предположим, что

$$\sum_{k=1}^{\infty} (m_{2k})^{-1/(2k)} = \infty \quad (\text{условие Карлемана}).$$

Тогда F определяется однозначно своими моментами $\{m_k\}$.

Если носитель F — это промежуток $[0, \infty)$ (вместо всей прямой $\mathbb{R}^1$), то условие Карлемана, обеспечивающее единственность проблемы моментов, выглядит так: $\sum_{k=1}^{\infty} (m_k)^{-1/(2k)} = \infty$.

Критерий (C₃). (а) Предположим, что $F(x)$, $x \in \mathbb{R}^1$, — абсолютно непрерывная функция распределения с плотностью $f(x) > 0$, $x \in \mathbb{R}^1$. Пусть далее F имеет моменты всех порядков и выполнено следующее условие:

$$\int_{-\infty}^{\infty} \frac{-\ln f(x)}{1+x^2} dx < \infty \quad (\text{условие Крейна}).$$

Тогда F не определяется однозначно последовательностью своих моментов $\left\{m_k = \int_{-\infty}^{\infty} x^k dF(x), k = 1, 2, \dots\right\}$.

(б) Предположим, что $F(x)$, $x \in \mathbb{R}^+$, — абсолютно непрерывная функция распределения на полуоси $[0, \infty)$, т. е. $F(0) = 0$, а ее плотность $f(x) > 0$, $x \in \mathbb{R}^+$. Пусть F имеет моменты всех порядков и пусть для некоторого $a \geq 0$ выполнено следующее условие:

$$\int_a^{\infty} \frac{-\ln f(x^2)}{1+x^2} dx < \infty \quad (\text{условие Крейна}).$$

Тогда F не определяется однозначно последовательностью своих моментов $\left\{m_k = \int_0^\infty x^k dF(x), k = 1, 2, \dots\right\}$.

По поводу критериев (C_1) и (C_2) читатель может обратиться, например, к книгам [106, 434]. Критерий (C_3) , предложенный впервые в [63], обсуждается подробно в [2]; см. также [92] и дополнение к [76]. Полезные результаты, касающихся проблемы моментов и близких вопросов, содержатся в [58, 75, 92, 106, 116, 145, 177, 345, 450, 451].

Далее с помощью конкретных примеров выясним роль различных условий, обеспечивающих единственность решения проблемы моментов, а также связь между различными условиями.

11.1. Проблема моментов для некоторых функций от нормально распределенной случайной величины

Пусть случайная величина $\xi \sim \mathcal{N}(0, 1)$. Для моментов ξ имеем $m_{2k-1} = 0$, $m_{2k} = (2k - 1)!!$, $k = 1, 2, \dots$. Распределение ξ определяется однозначно последовательностью $\{m_k\}$. Это можно установить несколькими способами. Например, легко видеть, что характеристическая функция $\psi(t)$ является аналитической функцией или то, что производящая функция $M(t)$ существует даже для всех $t \in \mathbb{R}^1$, так как $M(t) = e^{t^2/2}$. Тогда вывод о единственности решения проблемы моментов следует из критерия (C_1) .

Рассуждая аналогично, можно доказать, что случайная величина η , где $\eta = \xi^2$, обладает конечными моментами любого порядка и эти моменты определяют однозначно функцию распределения F_η . (F_η — это хи-квадрат распределение с одной степенью свободы.)

Рассмотрим теперь куб нормально распределенной случайной величины. Ради простоты записи предположим, что $\theta \sim \mathcal{N}\left(0, \frac{1}{2}\right)$, и пусть $X = \theta^3$. Тогда θ имеет плотность $\frac{1}{\sqrt{\pi}} e^{-x^2}$, $x \in \mathbb{R}^1$, и очень просто найти плотность f случайной величины X :

$$f(x) = \frac{1}{3\sqrt{\pi}} |x|^{-2/3} \exp(-|x|^{2/3}), \quad x \in \mathbb{R}^1.$$

Непосредственный подсчет показывает, что X имеет конечные моменты любого порядка (их можно выписать явно). Далее оказывается, что

$$\int_{-\infty}^{\infty} \frac{-\ln f(x)}{1+x^2} dx < \infty,$$

т. е. условие Крейна выполнено. Согласно критерию (C_3), распределение случайной величины X не определяется однозначно по последовательности моментов $\{E(X^k), k \geq 1\}$.

Итак, мы установили, что проблема моментов для распределения случайной величины $X = \theta^3$ не имеет единственного решения. А раз так, то было бы интересно найти другое распределение, но с теми же моментами. На самом деле теперь мы построим даже континуальное семейство распределений, все моменты которых совпадают с моментами случайной величины X . Действительно, пусть

$$f_\varepsilon(x) = f(x)\{1 + \varepsilon h(x)\}, \quad x \in \mathbb{R}^1, \quad \varepsilon \in [-1, 1],$$

где f — выписанная выше плотность случайной величины X , а

$$h(x) = \sin\left(\frac{\pi}{6} - \sqrt{3}|x|^{2/3}\right), \quad x \in \mathbb{R}^1.$$

Легко видеть, что $f_\varepsilon(x) \geq 0$, $x \in \mathbb{R}^1$. Далее, можно показать, что имеет место равенство

$$\int_{-\infty}^{\infty} x^k f(x) h(x) dx = 0 \quad \text{для} \quad k = 0, 1, 2, \dots$$

Беря $k = 0$, заключаем, что f_ε является плотностью некоторой случайной величины, X_ε . Более того, $E(X_\varepsilon^k) = E(X^k)$ для каждого $k \geq 1$. Следовательно, случайные величины X_ε и X имеют одни и те же моменты любого порядка, и тем не менее плотности у них разные, т. е. $f_\varepsilon(x) \neq f(x)$, $x \in \mathbb{R}^1$ (исключая тривиальный случай $\varepsilon = 0$).

Семейство распределений $S(f, h) = \{f_\varepsilon = f(1 + \varepsilon h)\}$, $\varepsilon \in [-1, 1]$, называется *классом Стильбеса* (подробности даны далее в примере 11.4).

Отметим, наконец, что для любого n , $n = 1, 2, \dots$, распределение случайной величины θ^{2n+1} не определяется однозначно последовательностью моментов $\{E[(\theta^{2n+1})^k], k \geq 1\}$ (см., например, [92, 142]).

11.2. Логарифмически нормальное распределение и проблема моментов

Напомним, что если $\xi \sim \mathcal{N}(0, 1)$, то распределение случайной величины $X = e^\xi$ называется *логарифмически нормальным распределением*. Плотность f величины X задается так:

$$f(x) = \begin{cases} 0, & \text{если } x \leq 0, \\ \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{x} \cdot \exp\left(-\frac{(\ln x)^2}{2}\right), & \text{если } x > 0. \end{cases} \quad (1)$$

Поскольку $m_k = E(X^k) = e^{k^2/2}$, то X имеет конечные моменты любого порядка. Поэтому вполне естественно спросить: определяется ли однозначно логарифмически нормальное распределение своими моментами $\{m_k\}$? Покажем далее, что ответ отрицателен.

В качестве предварительного шага мы можем проверить, например, что производящая функция X не существует, так как $E\{e^{tX}\}$ определено только при $t = 0$. Кроме того, $\sum_{k=1}^{\infty} (m_k)^{-1/(2k)} < \infty$ для $m_k = e^{k^2/2}$, так что и условие Карлемана не выполнено. Однако мы не можем сделать отсюда никакого вывода, потому что условия, входящие в критерий (C_1) и критерий (C_2) , — это достаточные условия. Таким образом, нам нужно обратиться к критерию (C_3) . Имеют место следующие соотношения:

$$\begin{aligned} \int_0^{\infty} (1+x^2)^{-1} |\ln x|^k dx &= \int_{-\infty}^{\infty} (1+e^{2y})^{-1} |y|^k e^y dy \leq \\ &\leq \int_{-\infty}^0 |y|^k e^y dy + \int_0^{\infty} y^k e^{-y} dy < \infty, \quad k = 0, 1, 2, \dots \end{aligned}$$

Отсюда нетрудно вывести, что условие Крейна выполнено, и, значит, логарифмически нормальное распределение не определяется однозначно своими моментами.

Как и в примере 11.1, здесь также можно описать явно класс Стильеса, т. е. континуальное семейство распределений, имеющих те же самые моменты, что и логарифмически нормальное распределение. Действительно, положим

$$f_{\varepsilon}(x) = f(x) [1 + \varepsilon \sin(2\pi \ln x)], \quad x > 0, \quad (2)$$

где f — плотность, заданная формулой (1), а $\varepsilon \in [-1, 1]$. Очевидно, что $f_{\varepsilon}(x) \geq 0$ для всех $x \in \mathbb{R}^1$. Чтобы вывести другие свойства функции f_{ε} , нужно сначала показать, что

$$J_k = \int_0^{\infty} x^k f(x) \sin(2\pi \ln x) dx = 0 \quad \text{для} \quad k = 0, 1, 2, \dots \quad (3)$$

Замена переменных $\ln x = u = y + k$ позволяет свести J_k к интегралу

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-u^2/2+ku} \sin(2\pi u) du = \frac{1}{\sqrt{2\pi}} e^{k^2/2} \int_{-\infty}^{\infty} e^{-y^2/2} \sin(2\pi y) dy.$$

Однако последний интеграл равен нулю ввиду нечетности подынтегральной функции.

Теперь, опираясь на равенство (3) при $k = 0$ мы заключаем, что при каждом $\varepsilon \in [-1, 1]$ функция f_ε является плотностью некоторой случайной величины, которую мы обозначим X_ε . (Если $\varepsilon = 0$, то f_ε и X_ε совпадают с f и X .) Более того, для каждого k , $k = 1, 2, \dots$, имеет место равенство $E\{X_\varepsilon^k\} = E\{X^k\}$, несмотря на то что $f_\varepsilon \neq f$ при $\varepsilon \neq 0$.

Итак, мы описали явно континуальное семейство $\{X_\varepsilon\}$, состоящее из абсолютно непрерывных случайных величин, моменты любого порядка которых совпадают с соответствующими моментами логарифмически нормально распределенной случайной величины.

Оказывается, мы можем сделать еще один шаг, являющийся, по-видимому, несколько неожиданным. Именно мы построим семейство функций распределения $\{H_a, a > 0\}$, имеющих те же моменты, что и логарифмически нормальное распределение и при этом функция распределения H_a будет соответствовать дискретной случайной величине. Рассмотрим функцию

$$h_a(t) = \sum_{n=-\infty}^{\infty} a^{-n} e^{-n^2/2} \exp(iae^n t), \quad t \in \mathbb{R}^1, \quad a > 0. \quad (4)$$

Легко проверить, что ряд в правой части (4) сходится для всех $t \in \mathbb{R}^1$ и всех $a > 0$. Более того, функция $h_a(t)$, $t \in \mathbb{R}^1$, $a > 0$, непрерывна и положительно определена. Тогда, согласно теореме Бохнера (см. [76, 92, 106, 116, 117, 404]), функция $\psi_a(t) = \frac{h_a(t)}{h_a(0)}$, $t \in \mathbb{R}^1$, является характеристической функцией. Через H_a и Y_a обозначим функцию распределения и случайную величину, соответствующие этой характеристической функции ψ_a . Выражение (4) для функции h_a позволяет описать явно распределение случайной величины Y_a . Мы имеем

$$p_a(k) = P\{Y_a = ae^k\} = a^{-k} \frac{e^{-k^2/2}}{h_a(0)}, \quad k = 0, \pm 1, \pm 2, \dots$$

Следующий шаг состоит в нахождении моментов

$$\tilde{m}_n = E\{Y_a^n\} = \sum_{k=-\infty}^{\infty} (ae^k)^n p_a(k), \quad n = 1, 2, \dots$$

Поскольку

$$\left. \frac{d^n}{dt^n} h_a(t) \right|_{t=0} = \sum_{k=-\infty}^{\infty} a^{-k} (iae^k)^n e^{-k^2/2},$$

мы легко находим, что

$$e^{-n^2/2} \tilde{m}_n = i^{-n} \sum_{k=-\infty}^{\infty} a^{-k} (iae^k)^n \frac{e^{-(n^2+k^2)/2}}{h_a(0)} = \sum_{k=-\infty}^{\infty} a^{-(k-n)} \frac{e^{-(k-n)^2/2}}{h_a(0)} = 1.$$

Отсюда, очевидно, вытекает, что

$$\mathbb{E}\{Y_a^n\} = \tilde{m}_n = e^{n^2/2} = m_n = \mathbb{E}\{X^n\}, \quad n = 1, 2, \dots$$

Таким образом, мы построили бесконечное семейство $\{Y_a\}$, состоящее из дискретных случайных величин таких, что при каждом $a > 0$ случайная величина Y_a имеет те же самые моменты, что и случайная величина X с логарифмически нормальным распределением.

11.3. Еще один класс абсолютно непрерывных распределений, которые не определяются однозначно своими моментами

Рассмотрим случайную величину X со следующей плотностью:

$$f(x) = \begin{cases} c \exp(-\lambda x^\gamma), & \text{если } x > 0, \\ 0, & \text{если } x \leq 0. \end{cases}$$

Здесь $\lambda > 0$, $0 < \gamma < \frac{1}{2}$, а c — нормирующая константа.

Зададим функцию f_ε равенством

$$f_\varepsilon(x) = \begin{cases} c \exp(-\lambda x^\gamma) [1 + \varepsilon \sin(\beta x^\gamma)], & \text{если } x > 0, \\ 0, & \text{если } x \leq 0, \end{cases}$$

где $\varepsilon \in (-1, 1)$, $\beta = \lambda \operatorname{tg}(\gamma\pi)$. Очевидно, что $f_\varepsilon(x) \geq 0$, $x \in \mathbb{R}^1$. Дальнейшее рассуждение основывается на следующем равенстве:

$$\int_0^\infty x^k \exp(-\lambda x^\gamma) \sin(\beta x^\gamma) dx = 0, \quad (1)$$

справедливом для каждого $k = 0, 1, 2, \dots$ (Читатель может самостоятельно доказать это равенство либо обратиться за подробностями к книге [116].) Полагая $k = 0$ в (1), можно увидеть, что функция f_ε является плотностью. Пусть X_ε — случайная величина, плотность которой равняется f_ε . Тогда из связи между функциями f_ε и f и равенства (1) вытекает, что

$$\mathbb{E}\{X_\varepsilon^k\} = \mathbb{E}\{X^k\} \quad \text{для каждого } k, \quad k = 1, 2, \dots$$

Итак, мы построили бесконечно много случайных величин X_ε , моменты которых совпадают с моментами случайной величины X , причем

плотности у них разные (исключая тривиальный случай $\varepsilon = 0$). Другими словами, проблема моментов для распределения случайной величины X не имеет единственного решения. Заметим, однако, что этот факт можно было вывести и из критерия (C_3) , потому что плотность f случайной величины X удовлетворяет условию Крейна.

11.4. Проблема моментов для степенных преобразований экспоненциального распределения

Если ξ — случайная величина с конечным моментом, часто интерес представляют степенные преобразования ξ^r или $|\xi|^r$ с точки зрения теории и их важности в статистических приложениях. В каждом конкретном случае легко найти значения r , для которых эти степени корректно определены. Затем следует вопрос: каково решение проблемы моментов для распределений ξ^r и/или $|\xi|^r$? Увидим далее, что для некоторых r решение единственно, для других — неединственно.

Предположим, что $\xi \sim \text{Exp}(1)$ — экспоненциальное распределение с параметром 1. Его плотность равна e^{-x} , $x > 0$, и так как существует производящая функция моментов, т.е. имеет место условие Крамера, заключаем, что $\text{Exp}(1)$ — это единственное распределение с моментами $m_k = k!$, $k = 1, 2, \dots$

Рассмотрим теперь степени ξ , т.е. ξ^r , где $r \geq 0$. Условие Крамера выполнено для $r \in [0, 1]$, и, следовательно, распределение ξ^r определяется однозначно по своим моментам. (Читатель может выписать явно моменты $(\xi^r)^k$, $k = 1, 2, \dots$, в терминах гамма-функции.) Если $r \in (1, 2]$, условие Крамера не выполнено, но имеет место условие Карлемана, так что распределение ξ^r , $r \in (1, 2]$, — также однозначно определяется своими моментами.

Для любого $r > 0$ случайная величина ξ^r имеет плотность $f_r(x) = \frac{1}{r} x^{-1+1/r} \exp(-x^{1/r})$, $x > 0$. Условие Крейна, как это легко показать, выполнено только тогда, когда $r > 2$. Следовательно, распределение ξ^r , $r > 2$, не определяется единственным образом через свои моменты.

Итак, мы видим, что $r = 3$ — это наименьшая целая степень, такая что распределение ξ^3 неединственно в смысле моментов. В этом случае можно описать явно много других распределений с теми же моментами, что и ξ^3 . Действительно, рассмотрим следующие две функции:

$$f(x) = \frac{1}{3} x^{-2/3} e^{-x^{1/3}}, \quad h(x) = \sin\left(\frac{\pi}{6} - \sqrt{3} x^{1/3}\right), \quad x > 0.$$

Заметим, что $f(x)$, $x > 0$, — это плотность распределения ξ^3 . Моменты ξ^3 выражаются так: $E[(\xi^3)^k] = (3k)!$, $k = 1, 2, \dots$

Введем теперь следующее семейство функций, зависящее от f и h и называемое *классом Стилтъяса* (см. также примеры 11.1–11.4):

$$S = S(f, h) = \{f_\varepsilon(x) = f(x)[1 + \varepsilon h(x)], x > 0, -1 \leq \varepsilon \leq 1\}.$$

Функция h подобрана так, что $\int_0^\infty x^k f(x) h(x) dx = 0$, $k = 0, 1, 2, \dots$. Отсюда, ввиду ограниченности h ($|h(x)| \leq 1$), вытекает, что для каждого $\varepsilon \in [-1, 1]$ функция $f_\varepsilon(x)$, $x > 0$ является плотностью некоторой случайной величины, скажем X_ε . Ясно, что $f_0 = f$ и $X_0 = \xi^3$, и, более того, $f_\varepsilon \neq f$ для $\varepsilon \neq 0$, однако

$$E(X_\varepsilon^k) = E[(\xi^3)^k] = (3k)!, \quad k = 1, 2, \dots, \quad \varepsilon \in [-1, 1].$$

Стоит отметить, что если в общем случае нам удалось построить класс Стилтъяса, мы не только показываем неединственность распределения, но и сразу имеем явно выписанные другие распределения с теми же самыми моментами. За подробностями читатель может обратиться к недавним работам [356, 388, 451, 452].

11.5. Различные дискретные распределения с совпадающими моментами любого порядка

Напомним, что проблема моментов для абсолютно непрерывных распределений изучалась давно (см. [2, 434]). Дискретный случай привлек внимание математиков лишь недавно (см. [121, 240, 429]). Здесь проблема моментов ставится следующим образом.

Пусть $\mathbb{D} = \{x_1, x_2, \dots\}$ — счетное множество действительных чисел. Рассмотрим дискретную случайную величину ξ со значениями в множестве $\mathbb{D}$ и распределением $\mathcal{P}_\xi = \{p_n, n = 1, 2, \dots\}$, где $p_n = P\{\xi = x_n\}$, $p_n > 0$ и $\sum_n p_n = 1$. Предположим, что ξ имеет конечные моменты любого порядка k , $k = 1, 2, \dots$, т. е. $\sum_n |x_n|^k p_n < \infty$, $k = 1, 2, \dots$, и пусть $\{m_k, k = 1, 2, \dots\}$, где $m_k = \sum_n x_n^k p_n$ — это последовательность моментов ξ .

Вопрос: определяется ли распределение $\mathcal{P}_\xi$ единственным образом по последовательности $\{m_k\}$? В зависимости от ответа на этот вопрос проблема моментов имеет или не имеет единственное решение. С использованием критериев (C_1) и (C_2) мы можем показать, что распределения многих дискретных случайных величин определяются однозначно соответствующими моментами. Приведем теперь два примера, когда такой однозначности нет.

(а) Пусть ξ и η — дискретные случайные величины, принимающие значения в множестве $\mathbb{D}$, где $\mathbb{D} = \mathbb{N}_0 = \{0, 1, 2, \dots\}$. Обозначим их распределение через $\mathcal{P}_\xi = \{p_n, n \in \mathbb{N}_0\}$ и $\mathcal{P}_\eta = \{q_n, n \in \mathbb{N}_0\}$. Здесь $p_n = P\{\xi = n\}$, $q_n = P\{\eta = n\}$, $p_n \geq 0$, $q_n \geq 0$, $\sum_n p_n = 1$, $\sum_n q_n = 1$. Предположим, что ξ и η имеют конечные моменты любого порядка, и пусть

$$m_k = E\{\xi^k\} = \sum_n n^k p_n \quad \tilde{m}_k = E\{\eta^k\} = \sum_n n^k q_n, \quad k = 1, 2, \dots$$

Покажем, что случайные величины ξ и η могут иметь одни и те же моменты, т.е. $m_k = \tilde{m}_k$, $k = 1, 2, \dots$, но различные распределения $\mathcal{P}_\xi$ и $\mathcal{P}_\eta$. Это будет означать, что проблема моментов (для обеих величин ξ и η) не имеет единственного решения.

Пусть $h(x)$, $x \in \mathbb{R}^1$, — периодическая функция с периодом 2π и

$$h(x) = \exp\left(-\frac{1}{x(2\pi - x)}\right), \quad x \in [0, 2\pi].$$

Ее ряд Фурье $h(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}$ имеет коэффициенты

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} h(x) e^{-inx} dx = \frac{1}{2\pi} \int_{-\pi}^{\pi} h(x) \cos(nx) dx, \quad n = 0, \pm 1, \pm 2, \dots$$

Положим $a_n = \max(c_n, 0)$, $b_n = \max(-c_n, 0)$, так что $a_n \geq 0$, $b_n \geq 0$ для всех n . Очевидно также, что $c_n = a_n - b_n$. Поскольку функция h бесконечно дифференцируема, то для ее k -й производной имеем $h^{(k)}(x) = \sum_{n=-\infty}^{\infty} c_n (in)^k e^{inx}$. Из того, что $h^{(k)}(0) = 0$, следует равенство $\sum_{n=-\infty}^{\infty} c_n n^k = 0$, справедливое для каждого $k = 0, 1, 2, \dots$. Запишем это равенство в следующем эквивалентном виде:

$$\sum_{n=-\infty}^{\infty} a_n n^k = \sum_{n=-\infty}^{\infty} b_n n^k \quad \text{для каждого } k = 0, 1, 2, \dots$$

Теперь $a_0 = c_0 > 0$, $b_0 = 0$ и, следовательно, $s = \sum_{n=-\infty}^{\infty} a_n = \sum_{n=-\infty}^{\infty} b_n > 0$.

Полагая $p_n = \frac{a_n}{s}$ и $q_n = \frac{b_n}{s}$, получаем два вероятностных распределения, а именно $\{p_n, n \in \tilde{\mathbb{N}}\}$ и $\{q_n, n \in \tilde{\mathbb{N}}\}$. Пусть они соответствуют величинам ξ и η . Естественно для полученных распределений использовать обозначения $\mathcal{P}_\xi$ и $\mathcal{P}_\eta$. Приведенные выше рассуждения показывают, что для $m_k = E\{\xi^k\}$ и $\tilde{m}_k = E\{\eta^k\}$ имеет место

$$m_k = \tilde{m}_k \quad \text{при любом } k, \quad k = 1, 2, \dots, \quad \text{хотя } \mathcal{P}_\xi \neq \mathcal{P}_\eta.$$

Это как раз то, что мы хотели доказать.

(б) Рассмотрим случайную величину X , заданную так:

$$P\{X = e^{8n}\} = ce^{-n^2}, \quad n \in \tilde{\mathbb{N}}, \quad c^{-1} = \sum_{k=-\infty}^{\infty} e^{-k^2}.$$

Легко проверить, что $E\{e^{tX}\}$ определено только для $t = 0$. Значит, не существует производящей функции моментов, т. е. условие Крамера не выполнено. Однако, моменты X любого порядка существуют и вычисляются явно: $m_k = E\{X^k\} = e^{16k^2}$, $k = 1, 2, \dots$. Поскольку $\sum_{k=1}^{\infty} (m_k)^{-1/(2k)} < \infty$, т. е. условие Карлемана не имеет места, мы не можем сделать никакого заключения о единственности распределения X . Заметим, что моменты X имеют тот же самый порядок роста, как и моменты логарифмически нормальной случайной величины (см. пример 11.2). Поэтому можно ожидать, что распределение X неединственно в смысле моментов.

Построим бесконечное семейство случайных величин, каждая из которых будет иметь те же самые моменты, как и X .

Введем функцию $h_\varepsilon(n)$, $n \in \tilde{\mathbb{N}}$, зависящую от параметра ε , где $\varepsilon \in (0, 1)$, следующим образом:

$$h_\varepsilon(n) = \begin{cases} 0, & \text{если } n \equiv 0 \pmod{4}, \\ \varepsilon, & \text{если } n \equiv 1 \pmod{4}, \\ 0, & \text{если } n \equiv 2 \pmod{4}, \\ -\varepsilon, & \text{если } n \equiv 3 \pmod{4}. \end{cases}$$

Нетрудно видеть, что для каждого фиксированного ε , $\varepsilon \in (0, 1)$, функция $h_\varepsilon(n)$ нечетна, т. е. $h_\varepsilon(-n) = -h_\varepsilon(n)$. Кроме того, нечетной является и функция $h_\varepsilon(n + 4k)$ при каждом фиксированном k , $k = 1, 2, \dots$. И, наконец, $h_\varepsilon(n)$ — периодическая функция с периодом 4. Эти свойства позволяют установить соотношение

$$\sum_{n=-\infty}^{\infty} e^{8kn} h_\varepsilon(n) ce^{-n^2} = 0, \quad (1)$$

имеющее место для каждого $k = 0, 1, 2, \dots$ и каждого $\varepsilon \in (0, 1)$ (подробности можно найти в [240, 449]). Положим $q_\varepsilon(n) = ce^{-n^2}(1 - h_\varepsilon(n))$. Из (1) вытекает, что совокупность $\{q_\varepsilon(n), n \in \tilde{\mathbb{N}}\}$ — это дискретное вероятностное распределение, т. е. $q_\varepsilon(n) \geq 0$ для всех $n \in \tilde{\mathbb{N}}$ и $\varepsilon \in (0, 1)$, и, кроме того, $\sum_{n=-\infty}^{\infty} q_\varepsilon(n) = 1$. Обозначим через X_ε случайную величину

ну с этим распределением. Таким образом, мы имеем

$$P\{X_\varepsilon = e^{8n}\} = ce^{-n^2}(1 - h_\varepsilon(n)), \quad n \in \tilde{\mathbb{N}}, \quad \varepsilon \in (0, 1).$$

Так как (1) выполнено для каждого k , $k = 1, 2, \dots$, то

$$E\{X_\varepsilon^k\} = e^{16k^2} = m_k = E\{X^k\}.$$

Следовательно, каждая из случайных величин X_ε , где $\varepsilon \in (0, 1)$, имеет такие же моменты, как и случайная величина X . Очевидно, при этом, что распределения X_ε и X различны.

(в) Зафиксируем натуральное число $q \geq 2$ и рассмотрим множество $M_q = \{q^j : j = 0, 1, 2, \dots\}$. Очевидно, что $M_q \subset \mathbb{N}$. Любое число $n \in M_q$ имеет вид q^j , и мы определяем число p_n так: $p_n = \frac{e^{-q}q^j}{j!}$. Легко видеть, что совокупность чисел $\{p_n, n \in M_q\}$ — это дискретное вероятностное распределение на множестве M_q . Обозначим через X случайную величину со значениями в M_q и распределением $\{p_n\}$. В этом случае принято говорить, что X имеет *логарифмически пуассоновское распределение*. Нетрудно убедиться, что у X нет производящей функции моментов, однако все моменты X конечные и

$$m_k = E\{X^k\} = \sum_{j=0}^{\infty} \frac{e^{-q}q^{kj}q^j}{j!} = \exp[q(q^k - 1)], \quad k = 1, 2, \dots$$

Покажем, что имеется бесконечно много других случайных величин с теми же самыми моментами. Для этой цели рассмотрим функцию

$$h(z) = \prod_{k=1}^{\infty} (1 - zq^{-k}).$$

Поскольку $\sum_{k=1}^{\infty} q^{-k} < \infty$ для любого $q > 1$, функция $h(z)$ аналитическая во всей комплексной плоскости. Пусть $h(z) = \sum_{j=0}^{\infty} c_j z^j$ — ее разложение в ряд Тейлора в окрестности точки 0. С учетом легко проверяемого равенства $h(qz) = (1 - z)h(z)$ мы выводим, что $\frac{c_j}{c_{j-1}} = -(q^j - 1)^{-1}$, где $c_0 = 1$, и находим

$$c_j = (-1)^j [(q - 1)(q^2 - 1) \dots (q^j - 1)]^{-1}, \quad j = 1, 2, \dots$$

Обозначая $a_j = j!c_j$, мы видим, что $|a_j| \leq 1$ для всех $j = 0, 1, 2, \dots$. Так мы приходим к следующему равенству:

$$e^{-q} \sum_{j=1}^{\infty} \frac{q^{kj} a_j q^j}{j!} = e^{-q} h(q^{k+1}) = 0, \quad k = 0, 1, 2, \dots \quad (2)$$

Определим теперь совокупность чисел $\{p_n^{(\varepsilon)}, n \in M_q\}$, где $\varepsilon \in [-1, 1]$, а

$$p_n^{(\varepsilon)} = p_n(1 + \varepsilon a_j) = e^{-q} q^j \left((j!)^{-1} + \varepsilon (-1)^j ((q-1)(q^2-1) \dots (q^j-1))^{-1} \right),$$

$$n = q^j, \quad j = 0, 1, 2, \dots$$

Очевидно, что $p_n^{(\varepsilon)} \geq 0$, и со ссылкой на (2) находим $\sum_{n \in M_q} p_n^{(\varepsilon)} = 1$ для любого $\varepsilon \in [-1, 1]$. Следовательно, $\{p_n^{(\varepsilon)}, n \in M_q\}$ задает дискретное вероятностное распределение на множестве M_q . Обозначим через X_ε случайную величину со значениями в M_q и распределением $\{p_n^{(\varepsilon)}\}$. Ясно, что $X_0 = X$, так как $\{p_n^{(0)}\} = \{p_n\}$. Ссылаясь еще раз на равенство (2), заключаем, что

$$E\{X_\varepsilon^k\} = E\{X^k\}, \quad k = 1, 2, \dots, \quad \varepsilon \in [-1, 1].$$

Иными словами, при $\varepsilon_1 \neq \varepsilon_2$, $\varepsilon_1, \varepsilon_2 \in [-1, 1]$, величины X_{ε_1} и X_{ε_2} имеют одни и те же моменты, однако их распределения различны: $\{p_n^{(\varepsilon_1)}, n \in M_q\} \neq \{p_n^{(\varepsilon_2)}, n \in M_q\}$.

11.6. Как связаны два различных достаточных условия для единственности решения проблемы моментов?

В начале настоящего параграфа мы сформулировали критерий (C_1) и критерий (C_2) . Оба содержат условия, которые достаточны для единственности решения проблемы моментов, и эти условия различны. Поэтому интересно сравнить эти два условия.

(а) Пусть $X = \xi \ln(1 + \eta)$, случайные величины ξ и η независимы и каждая из них экспоненциально распределена с параметром 1. Таким образом, X — абсолютно непрерывная случайная величина с неотрицательными значениями. Нетрудно вычислить момент $m_n = E\{X^n\}$. Именно

$$m_n = n! c_n, \quad \text{где} \quad c_n = \int_0^\infty [\ln(1+x)]^n e^{-x} dx.$$

Можно показать (детали опускаем), что

$$e^{-1} \ln(1+n) < (c_n)^{1/n} < c \ln(1+n), \quad c = \text{const} > 1. \quad (1)$$

Таким образом,

$$\left(\frac{m_n}{n!} \right)^{1/n} = (c_n)^{1/n} > e^{-1} \ln(1+n) \rightarrow \infty \quad \text{при} \quad n \rightarrow \infty,$$

и, следовательно, ряд $\sum_{n=1}^{\infty} \frac{m_n t^n}{n!}$ расходится для любого $t > 0$. Это означает, что характеристическая функция, отвечающая X , не является аналитической, т. е. условие Крамера не выполнено и мы не можем применить критерий (C_1) , чтобы ответить на вопрос о единственности решения проблемы моментов.

Из (1) получаем, что для больших значений n

$$(m_n)^{1/n} \approx e^{-1} n (c_n)^{1/n} < c e^{-1} n \ln(1+n),$$

и, значит, $\sum_{n=1}^{\infty} m_n^{-1/n} = \infty$, так как ряд $\sum_{n=1}^{\infty} \frac{1}{n \ln(1+n)}$ расходится. Тем самым условие Карлемана выполнено и, согласно критерию (C_2) , распределение случайной величины X определяется однозначно своими моментами.

(б) Рассмотрим теперь следующую дискретную случайную величину Y :

$$P\{Y = n\} = c \exp\left(-\frac{n}{\ln n}\right), \quad n = 3, 4, \dots$$

Нормирующая константа c определяется из условия $P\{Y \geq 3\} = 1$.

Отсылая читателя за подробностями к [177, 449], мы сформулируем вывод, к которому приходим: характеристическая функция $E\{e^{itY}\}$ не аналитична, поэтому нельзя применить критерий (C_1) . Однако условие Карлемана оказывается выполненным, и, согласно критерию (C_2) , распределение случайной величины Y определяется однозначно своими моментами.

Итак, для единственности решения проблемы моментов условие Крамера более ограничительно, чем условие Карлемана.

11.7. Условие Карлемана достаточно, но не необходимо для единственности решения проблемы моментов

Покажем, что условие Карлемана не является необходимым для того, чтобы распределение определялось однозначно своими моментами.

Пусть функция распределения $F_H(x)$, $x \in \mathbb{R}^1$, симметрична относительно точки 0. Пусть, далее, F_S — функция распределения, сосредоточенная на полупрямой $[0, \infty)$, причем F_H и F_S связаны соотношением

$$F_H(x) = \begin{cases} \frac{1}{2}[1 + F_S(x^2)], & \text{если } x \geq 0, \\ \frac{1}{2}[1 - F_S(x^2)], & \text{если } x < 0. \end{cases}$$

Заметим, что это соотношение задает взаимно однозначное соответствие между симметричными распределениями на $(-\infty, \infty)$ и распределениями на $[0, \infty)$. (Читатель догадывается, что индексы H и S связаны с проблемой моментов Хамбургера и проблемой моментов Стилтъяеса.) Обозначая

$$\tilde{m}_n = \int_{-\infty}^{\infty} x^n dF_H(x) \quad \text{и} \quad m_n = \int_0^{\infty} x^n dF_S(x),$$

мы видим, что моменты $\{\tilde{m}_n, n \geq 1\}$ существуют тогда и только тогда, когда существуют моменты $\{m_n, n \geq 1\}$. В этом случае

$$\tilde{m}_{2n} = m_n, \quad \tilde{m}_{2n-1} = 0, \quad n = 1, 2, \dots,$$

и можно предположить, что существует связь между решениями проблемы моментов Стилтъяеса и проблемы моментов Хамбургера.

Далее мы воспользуемся следующим результатом (см. [281]): *если проблема моментов Стилтъяеса для последовательности $\{m_n\}$ имеет единственное решение, причем соответствующее распределение не имеет массы в точке 0, то $\{m_n\}$ соответствует также проблеме моментов Хамбургера с единственным решением.*

Рассмотрим случайную величину X со следующей плотностью:

$$f(x) = \begin{cases} \frac{1}{\Gamma(1/\beta)} \exp(-x^\beta), & \text{если } x > 0, \\ 0, & \text{если } x \leq 0, \end{cases}$$

где $0 < \beta < 1$. Можно показать, что

$$m_n = E\{X^n\} = \frac{\Gamma((n+1)/\beta)}{\Gamma(1/\beta)}, \quad n = 0, 1, 2, \dots$$

Так как для больших значений n мы имеем $(m_n)^{1/n} \approx Kn^{1/\beta}$, где K — некоторая константа, то условие Карлемана $\sum_{n=1}^{\infty} m_n^{-1/(2n)} = \infty$ выполнено при $\frac{1}{2} \leq \beta < 1$. Следовательно, при $\frac{1}{2} \leq \beta < 1$ проблема моментов Стилтъяеса имеет единственное решение. Так как плотность f ограничена, то соответствующее ей распределение не имеет массы в точке 0. Согласно приведенному выше результату, последовательности $\{m_n\}$ отвечает проблема моментов Хамбургера с единственным решением. Другими словами, существует единственная функция распределения G , носителем которой является вся числовая прямая, такая что

$$\int_{-\infty}^{\infty} x^n dG(x) = m_n = \frac{\Gamma((n+1)/\beta)}{\Gamma(1/\beta)}, \quad n = 1, 2, \dots, \quad \text{где } \frac{1}{2} \leq \beta < 1.$$

Однако нетрудно проверить, что ряд $\sum_{n=1}^{\infty} m_{2n}^{-1/(2n)}$ сходится для каждого $\beta \in (0, 1)$, т. е. условие Карлемана не выполнено.

11.8. Условие Крейна достаточно, но не необходимо для неединственности решения проблемы моментов

В примере 11.3 мы рассмотрели функцию $f(x) = ce^{-x^\gamma}$ в качестве плотности распределения на $[0, \infty)$. Здесь c — нормирующая константа. Мы показали, что для любого $\gamma \in (0, \frac{1}{2})$, распределение с этой плотностью не определяется однозначно своими моментами.

Рассмотрим теперь следующую функцию:

$$h(x) = \begin{cases} \frac{1}{2}ce^{-x^\gamma}, & \text{если } x \geq 0 \\ \frac{1}{2}e^x, & \text{если } x < 0. \end{cases}$$

Легко убедиться, что $h(x)$, $x \in \mathbb{R}^1$ — вероятностная плотность некоторой случайной величины, скажем Y , со значениями на всей прямой. Заметим, что для $\gamma \in (0, \frac{1}{2})$ у величины Y не существует производящей функции моментов. Однако Y имеет конечные моменты любого порядка. Из примера 11.3 вытекает, что распределение Y не определяется однозначно своими моментами.

Выясним, наконец, какова роль условия Крейна для этого заключения. Из явного вида положительной функции $h(x)$, $x \in \mathbb{R}^1$, находим, что

$$\int_{-\infty}^{\infty} \frac{-\ln h(x)}{1+x^2} dx = \infty.$$

Это означает, что условие Крейна (конечность этого интеграла) не является необходимым для неединственности распределения.

§ 12. Характеризационные свойства некоторых вероятностных распределений

Хорошо известно, что многие классы вероятностных распределений можно характеризовать определенными свойствами, называемыми «характеризационными». Эти свойства дают возможность отнести некоторое распределение к определенному классу. Сформулируем два результата, которые имеют важное значение для теории вероятностей.

Теорема Крамера (см. [187]). *Если X_1 и X_2 — независимые случайные величины, а их сумма $X_1 + X_2$ имеет нормальное распределение, то тогда каждая из величин X_1 и X_2 нормально распределена.*

(По-видимому, это один из первых результатов, содержащих характеристизационные свойства распределений.)

Теорема Райкова (см. [94]). *Пусть X_1 и X_2 — неотрицательные целочисленные случайные величины такие, что сумма $X_1 + X_2$ имеет пуассоновское распределение. Если к тому же X_1 и X_2 независимы, то тогда каждая из этих величин имеет пуассоновское распределение.*

Доказательства этих теорем и их различных обобщений и следствий могут быть найдены, например, в [55, 95, 106, 177, 242, 374].

Далее мы рассмотрим ряд конкретных примеров и выясним, насколько важны условия, при которых имеет место то или иное характеристизационное свойство.

12.1. Как распределены случайные величины, если их сумма имеет биномиальное распределение?

Предположим, что неотрицательные целочисленные случайные величины X и Y таковы, что их сумма $Z = X + Y$ имеет биномиальное распределение с параметрами (n, p) , т. е. $Z \sim \text{Bi}(n, p)$. Тогда Z имеет вероятностную производящую функцию $E\{s^Z\} = (ps + q)^n$, где $q = 1 - p$. Если дополнительно предположить, что X и Y независимы, то тогда $(ps + q)^n = E\{s^X\}E\{s^Y\}$. Однако все делители многочлена $(ps + q)^n$ имеют вид $(ps + q)^k$, $k = 0, 1, \dots, n$, и отсюда вытекает, что каждая из величин X и Y также имеет биномиальное распределение. Вопрос: верно ли это утверждение без предположения о независимости X и Y ? Найдем ответ даже в более общей ситуации.

Пусть неотрицательная целочисленная случайная величина ζ принимает по крайней мере три различных значения. Введем случайные величины ξ и η , где $\xi = \left\lfloor \frac{\zeta}{2} \right\rfloor$, $\eta = \left\lceil \frac{\zeta + 1}{2} \right\rceil$ (как обычно, $[x]$ — это «целая часть» числа x , $x \in \mathbb{R}^1$). Очевидно, что $\zeta = \xi + \eta$.

Зная распределение ζ , легко можно найти $P\{\xi = k\}$ и $P\{\eta = m\}$ для всех возможных значений k и m . Из определения ξ и η вытекает, что значения ξ и η либо совпадают, либо отличаются на 1, а это означает, что $P\{\xi = k, \eta = m\} = 0$ для тех k и m , для которых $|k - m| > 1$. Отсюда, или просто учитывая соотношение $\xi \leq \eta$, заключаем, что величины ξ и η не являются независимыми, и это имеет место при любом распределении случайной величины ζ . Предположим, в частности, что ζ имеет биномиальное распределение. Тогда очевидно, что каждая из величин ξ

и η имеет распределение, отличающееся от биномиального, но их сумма $\xi + \eta$, равняющаяся ζ , биномиально распределена. Напомним еще раз, что ξ и η — зависимы.

12.2. Свойство геометрического распределения, которое не является его характеристическим свойством

Напомним, что случайная величина X имеет *геометрическое распределение* с параметром p , $0 < p < 1$, если $P\{X = n\} = pq^n$, $n = 0, 1, 2, \dots$, $q = 1 - p$.

Пусть X_1 и X_2 — независимые случайные величины, причем каждая из них геометрически распределена с параметром p . Из определения условной вероятности легко вывести следующее свойство:

$$P\{X_1 = k \mid X_1 + X_2 = n\} = \frac{1}{n+1}, \quad k = 0, 1, \dots, n, \quad n = 0, 1, \dots \quad (1)$$

Вопрос: является ли (1) характеристическим свойством геометрического распределения? Более точно вопрос ставится так: предположим, что X_1 и X_2 — неотрицательные целочисленные и независимые случайные величины, удовлетворяющие свойству (1). Следует ли отсюда, что величины X_1 и X_2 имеют геометрическое распределение? Дадим ответ с помощью примера.

Рассмотрим множество $\Omega = \{\omega_{kn} : k = 0, 1, \dots, n, n = 1, 2, \dots\}$ и в нем определим вероятность P так: $P(\omega_{kn}) = \frac{p_n}{n+1}$, где p_n , $n = 0, 1, \dots$, — положительные числа с $\sum_{n=0}^{\infty} p_n = 1$. Тогда

$$\Omega = \bigcup_{n=0}^{\infty} \Omega_n, \quad \text{где} \quad \Omega_n = \{\omega_{kn}, k = 0, 1, \dots, n\} \quad \text{и} \quad P(\Omega_n) = p_n,$$

причем каждый исход $\omega_{kn} \in \Omega_n$ имеет вероятность $\frac{p_n}{n+1}$.

Введем теперь случайные величины ξ_1 и ξ_2 такие, что $\xi_1(\omega_{kn}) = k$, $\xi_2(\omega_{kn}) = n - k$. Тогда ясно, что при $k = 0, 1, \dots, n$

$$\begin{aligned} P\{\xi_1 = k \mid \xi_1 + \xi_2 = n\} &= \frac{P\{\xi_1 = k, \xi_1 + \xi_2 = n\}}{P\{\xi_1 + \xi_2 = n\}} = \\ &= \frac{P(\omega_{kn})}{P(\Omega_n)} = \frac{p_n/(n+1)}{p_n} = \frac{1}{n+1}. \end{aligned}$$

Это означает, что свойство (1) выполнено для пары случайных величин ξ_1 и ξ_2 . Для распределения случайной величины ξ_1 находим

$$P\{\xi_1 = k\} = P\left\{\bigcup_{n=k}^{\infty} \{\omega_{kn}\}\right\} = \sum_{n=k}^{\infty} \frac{p_n}{n+1}.$$

Вспомним, что числа p_n произвольны ($p_n > 0$ и $\sum p_n = 1$), и, значит, вероятность $P\{\xi_1 = k\}$, $k = 0, 1, \dots$, может сильно отличаться от соответствующей вероятности для геометрического распределения. Только при весьма специальном выборе чисел p_n будет иметь место равенство $\sum_{n=k}^{\infty} \frac{p_n}{n+1} = pq^k$. Следовательно, (1) не является характеризационным свойством геометрического распределения. Заметим, что если дополнительно предположить, что случайные величины X_1 и X_2 независимы и одинаково распределены, то тогда из (1) можно вывести, что каждая из величин X_1 и X_2 имеет геометрическое распределение.

12.3. О применимости теоремы Райкова

(а) В теореме Райкова (см. начало этого параграфа) есть требование о независимости случайных величин X_1 и X_2 для того, чтобы из пуассоновости суммы $X_1 + X_2$ вытекала пуассоновость каждого из слагаемых X_1 и X_2 . Интересно выяснить, что будет, если опустить условие независимости X_1 и X_2 . На самом деле, ответ содержится в уже рассмотренном примере 12.1. Действительно, пусть случайная величина ζ имеет пуассоновское распределение с заданным параметром. Положим $X_1 = \left[\frac{\zeta}{2}\right]$, $X_2 = \left[\frac{\zeta+1}{2}\right]$. Тогда X_1 и X_2 — неотрицательные целочисленные случайные величины, причем ни одна из них не имеет распределение Пуассона, и тем не менее их сумма $X_1 + X_2$, равняющаяся ζ , имеет пуассоновское распределение. Остается добавить, что величины X_1 и X_2 зависимы. Следовательно, условие независимости в теореме Райкова существенно.

(б) Выясним теперь вопрос о применимости теоремы Райкова для распределений более общих, чем распределение Пуассона.

Дадим сначала одно определение. Скажем, что неотрицательная целочисленная случайная величина X имеет *обобщенное пуассоновское распределение порядка k* с параметром λ , где $k \in \mathbb{N}$, $\lambda > 0$, если

$$P\{X = n\} = \sum_{*} \frac{e^{ik\lambda} \lambda^{m_1 + \dots + m_k}}{m_1! \dots m_k!}. \quad (1)$$

В (1) сумма $\sum_{*}$ берется по всем неотрицательным целым $m_1, \dots, m_k$ таким, что $m_1 + 2m_2 + \dots + km_k = n$. Если $k = 1$, то (1) задает обычное пуассоновское распределение. Можно показать (см. [107]), что X имеет следующую вероятностную производящую функцию:

$$g(s) = E\{s^X\} = \exp \left[-\lambda \left(k - \sum_{i=1}^k s^i \right) \right], \quad |s| \leq 1. \quad (2)$$

Предположим теперь, что X_1 и X_2 — независимые случайные величины, принимающие значения в множестве $\{0, 1, 2, \dots\}$ и такие, что их сумма $X_1 + X_2$ имеет обобщенное пуассоновское распределение порядка k . Вопрос: следует ли отсюда, что каждая из величин X_1 и X_2 имеет обобщенное пуассоновское распределение порядка k ?

Заметим, что частный случай $k = 1$ соответствует теореме Райкова, так что в этом случае ответ положителен. Оказывается, однако, что если $k \geq 2$, то ответ на поставленный вопрос отрицателен.

Рассмотрим две независимые случайные величины Y_1 и Y_2 , где Y_1 имеет обобщенное пуассоновское распределение порядка $k - 1$ и параметр λ , а Y_2 имеет следующее распределение:

$$P\{Y_2 = m\} = \frac{e^{-\lambda} \lambda^{m/k}}{(m/k)!}, \quad m = 0, k, 2k, 3k, \dots$$

Воспользуемся явными выражениями для производящих функций $g_1(s) = E\{s^{Y_1}\}$ и $g_2(s) = E\{s^{Y_2}\}$. Из (2) находим

$$g_1(s) = \exp\left[-\lambda\left(k - 1 - \sum_{i=1}^{k-1} s^i\right)\right], \quad |s| \leq 1,$$

а непосредственный подсчет показывает, что

$$g_2(s) = \exp[-\lambda(1 - s^k)], \quad |s| \leq 1.$$

Поскольку Y_1 и Y_2 независимы, то сумма $Y = Y_1 + Y_2$ имеет производящую функцию

$$g_3(s) = E\{s^Y\} = g_1(s)g_2(s),$$

т. е.

$$g_3(s) = \exp\left[-\lambda\left(k - \sum_{i=1}^k s^i\right)\right], \quad |s| \leq 1.$$

Из (2) вытекает, что g_3 — производящая функция обобщенного пуассоновского распределения порядка k . Это и есть распределение случайной величины Y .

Таким образом, случайная величина Y с обобщенным пуассоновским распределением порядка k разложена на сумму двух независимых случайных величин Y_1 и Y_2 , ни одна из которых не имеет обобщенного пуассоновского распределения порядка k . Следовательно, теорема Райкова не имеет места для обобщенных пуассоновских распределений порядка $k \geq 2$.

12.4. О применимости теоремы Крамера

Теорема Крамера (см. начало этого параграфа) может быть сформулирована в следующей эквивалентной форме: *пусть $F_1(x)$, $x \in \mathbb{R}^1$, и $F_2(x)$, $x \in \mathbb{R}^1$, — невырожденные функции распределения, удовлетворяющие соотношению*

$$F_1 * F_2(x) = \Phi_{a,\sigma}(x), \quad x \in \mathbb{R}^1, \quad (1)$$

где $\Phi_{a,\sigma}$ — нормальная функция распределения с параметрами a и σ^2 . Тогда каждая из F_1 и F_2 является нормальной функцией распределения.

Возникает вопрос: справедлива ли теорема Крамера, если (1) выполнено не для всех x , а только для $x \leq x_0$ при некотором фиксированном $x_0 \in \mathbb{R}^1$? Покажем, что это не всегда так.

Пусть $\Phi = \Phi_{0,1}$ — стандартная нормальная функция распределения. Рассмотрим следующую функцию:

$$F_1(x) = \begin{cases} 2 \sum_{n=0}^{\infty} (-1)^n \Phi(x-n) = \sqrt{\frac{2}{\pi}} \sum_{n=0}^{\infty} \int_{x-(2n+1)}^{x-2n} e^{-u^2/2} du, & \text{если } x \leq 0, \\ v(x), & \text{если } x > 0, \end{cases}$$

где $v(x)$ — произвольная неубывающая функция, определенная для $x \in (0, \infty)$ и такая, что $v(0+) = F_1(0)$ и $v(\infty) = 1$. Нетрудно убедиться, что F_1 — функция распределения некоторой случайной величины, которую обозначим ξ_1 . Пусть ξ_2 — случайная величина, независимая от ξ_1 и принимающая каждое из значений 0 и 1 с вероятностью $\frac{1}{2}$. Через F_2 обозначим функцию распределения ξ_2 .

Легко проверить, что

$$F_1 * F_2(x) = \frac{1}{2}[F_1(x) + F_1(x-1)] = \Phi(x) \quad \text{для каждого } x \leq 0.$$

Очевидно также, что $F_1 * F_2(x) \neq \Phi(x)$ при $x > 0$. Ясно, что функции распределения F_1 и F_2 нормальными не являются. Следовательно, нельзя ослабить условие (1) и сохранить утверждение о нормальности распределений величин ξ_1 и ξ_2 .

Отметим, что если $F(x)$, $x \in \mathbb{R}^1$, — безгранично делимая функция распределения такая, что $F(x) = \Phi(x)$ для всех $x \leq x_0$ при некотором фиксированном $x_0 \in \mathbb{R}^1$, то $F = \Phi$ (см. [424]).

Легко видеть, однако, что функция распределения $F_1 * F_2$ не является безгранично делимой.

12.5. Об одном свойстве, которое не является характеризационным для нормального распределения

Пусть случайные величины X и Y независимы и каждая имеет распределение $\mathcal{N}(0, 1)$. Тогда частное $\frac{X}{Y}$ имеет распределение Коши с плотностью $\frac{1}{\pi(1+x^2)}$, $x \in \mathbb{R}^1$. Это можно доказать несколькими способами (например, прямым вычислением плотности $\frac{X}{Y}$ или вычислением характеристической функции $E\{e^{itX/Y}\}$). Указанное свойство приводит к следующему: пусть X и Y — независимые и одинаково распределенные случайные величины такие, что их частное $\frac{X}{Y}$ имеет распределение Коши. Следует ли отсюда, что величины X и Y нормально распределены? Покажем, что ответ отрицателен.

(а) Рассмотрим две независимые случайные величины ξ и η , имеющие одну и ту же плотность $f(x) = \frac{\sqrt{2}}{\pi(1+x^4)}$, $x \in \mathbb{R}^1$. Можно показать, что частное $\frac{\xi}{\eta}$ имеет плотность

$$\frac{d}{dz} P\left\{\frac{X}{Y} \leq z\right\} = \frac{d}{dz} \left(\frac{2}{\pi^2} \iint_{x/y \leq z} \frac{dx dy}{(1+x^4)(1+y^4)} \right) = \frac{1}{\pi(1+z^2)}, \quad z \in \mathbb{R}^1.$$

Таким образом, частное $\frac{\xi}{\eta}$ имеет распределение Коши, величины ξ и η независимы и одинаково распределены, но их распределение нормальным не является. Отметим, что частное $\frac{\xi}{\eta}$ распределено по закону Коши и в том случае, когда ξ и η — независимые случайные величины, каждая с плотностью

$$f_1(x) = \frac{2}{\pi} \cdot \frac{1}{(1+x^2)(1+x^4)}, \quad x \in \mathbb{R}^1.$$

(б) Обратим внимание на то, что в рассмотренных выше случаях величины ξ и η были независимыми, и это условие используется существенно для доказательства того, что частное $\frac{\xi}{\eta}$ имеет распределение Коши. Оказывается, это свойство может иметь место и для зависимых случайных величин.

Рассмотрим функцию $\psi(t, u) = (1 - 2iu + t^2)^{-1/2}$, $t, u \in \mathbb{R}^1$. Можно показать, что ψ является характеристической функцией некоторого двумерного случайного вектора (ξ, η) , причем распределение каждой из величин ξ и η отличается от нормального (достаточно посмотреть на явный вид характеристических функций, соответствующих ξ и η),

а величины ξ и η зависимы. Самое интересное здесь то, что частное $\frac{\xi}{\eta}$ имеет распределение Коши (подробности можно найти в [404]).

12.6. Еще одно интересное свойство, которое не является характеристическим для нормального распределения

Начнем с формулировки следующего результата (три различных его доказательства даны в [133], хотя читатель может сам его доказать): пусть X и Y — независимые случайные величины, а $Z = \frac{XY}{\sqrt{X^2 + Y^2}}$. Тогда если $X \sim \mathcal{N}(0, \sigma_1^2)$ и $Y \sim \mathcal{N}(0, \sigma_2^2)$, то $Z \sim \mathcal{N}(0, \sigma^2)$, где $\sigma^2 = \frac{\sigma_1^2 \sigma_2^2}{(\sigma_1 + \sigma_2)^2}$. Так описанное свойство нормального распределения может послужить поводом для следующего вопроса: является ли это свойство характеристическим свойством нормального распределения? Покажем, что ответ отрицателен.

Очевидно, случайные величины X , Y и Z связаны соотношением

$$\frac{1}{X^2} + \frac{1}{Y^2} = \frac{1}{Z^2}.$$

Не ограничивая общности, можно считать, что $Z \sim \mathcal{N}(0, 1)$. Распределение Z^{-2} имеет преобразование Лапласа $\varphi(t) = \exp(-\sqrt{2t})$, $t \geq 0$, а это означает, что распределение Z^{-2} устойчиво с параметром $\frac{1}{2}$. Покажем, что Z^{-2} допускает следующее представление:

$$Z^{-2} \stackrel{d}{=} U_1 + U_2, \quad (1)$$

где U_1 и U_2 — независимые случайные величины с неотрицательными значениями, такие, что распределение каждой из величин U_1 и U_2 не имеет атома в точке 0 и не принадлежит классу устойчивых распределений с параметром $\frac{1}{2}$. Для этой цели представим функцию φ в виде

$$\varphi(t) = \exp\left(-\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (1 - e^{-tx}) x^{-3/2} dx\right), \quad t \geq 0.$$

Введем также следующие две функции для $x \in (0, \infty)$:

$$h_1(x) = \frac{1}{\sqrt{8\pi}} x^{-1/2} I_{(0,1]}(x), \quad h_2(x) = h_1(x) + \frac{1}{\sqrt{2\pi}} x^{-1/2} I_{(1,\infty)}(x)$$

($I_A(x)$ — индикаторная функция множества A). Положим

$$\psi_j(t) = \int_0^{\infty} (1 - e^{-tx}) x^{-1} h_j(x) dx, \quad j = 1, 2,$$

и заметим, что интегралы $\int_1^{\infty} x^{-1} h_j(x) dx$ сходятся. Нетрудно установить, что каждая из функций

$$\varphi_1(t) = \exp[-\psi_1(t)], \quad t \in \mathbb{R}^1, \quad \text{и} \quad \varphi_2(t) = \exp[-\psi_2(t)], \quad t \in \mathbb{R}^1,$$

является преобразованием Лапласа безгранично делимого распределения с носителем $[0, \infty)$ (см. [106]). Поскольку $\psi_j(t) \rightarrow \infty$ при $t \rightarrow \infty$, $j = 1, 2$, эти распределения не имеют атомов в точке 0.

Пусть теперь U_1 и U_2 случайные величины, распределения которых имеют преобразования Лапласа, равные соответственно φ_1 и φ_2 . Предположим также, что U_1 и U_2 независимы. Тогда ввиду очевидного равенства $\varphi(t) = \varphi_1(t)\varphi_2(t)$ и на основе приведенных выше рассуждений приходим к следующему выводу: случайная величина Z^{-2} допускает представление (1), в котором случайные величины U_1 и U_2 независимы, причем каждая из случайных величин U_1^{-1} и U_2^{-1} имеет распределение, отличающееся от нормального. Это как раз то, что мы хотели доказать.

12.7. Об одном свойстве, не являющемся характеристическим для распределения Коши

Если случайная величина X имеет распределение Коши с плотностью $f(x) = \frac{1}{\pi(1+x^2)}$, $x \in \mathbb{R}^1$, то легко проверить, что $\frac{1}{X}$ имеет ту же самую плотность f . Это свойство распределения Коши приводит к следующему вопросу. Пусть X — абсолютно непрерывная случайная величина с функцией распределения $F(x)$, $x \in \mathbb{R}^1$. Предположим, что случайная величина $\frac{1}{X}$ имеет функцию распределения, равную F . Следует ли отсюда, что функция распределения F — это распределение Коши? Если ответ на этот вопрос положителен, то свойство $X \stackrel{d}{=} \frac{1}{X}$ будет характеристическим свойством распределения Коши. Оказывается, однако, что без дополнительных предположений ответ отрицателен. Проиллюстрируем это следующим примером.

Пусть случайная величина X имеет плотность

$$g(x) = \begin{cases} \frac{1}{4}, & \text{если } |x| \leq 1, \\ \frac{1}{4x^2}, & \text{если } |x| > 1. \end{cases}$$

Легко проверить, что случайная величина $\frac{1}{X}$ имеет своей плотностью эту же функцию g . Таким образом, $X \stackrel{d}{=} \frac{1}{X}$, причем X , а значит, и $\frac{1}{X}$ аб-

солютно непрерывны, но тем не менее плотность g отличается от плотности распределения Коши.

Отметим, наконец, что если $X \stackrel{d}{=} \frac{1}{X}$ и мы потребуем, чтобы распределение F было устойчивым, то F будет распределением Коши. Распределение с плотностью g неустойчиво.

12.8. Об одном свойстве, которое не является характеристическим для гамма-распределения

Пусть X и Y — независимые случайные величины, каждая из которых имеет *гамма-распределение* с плотностью

$$f(x) = \begin{cases} 0, & \text{если } x \leq 0, \\ \frac{a^p}{\Gamma(p)} x^{p-1} e^{-ax}, & \text{если } x > 0. \end{cases} \quad (1)$$

(Здесь $a > 0$, $p > 0$.) Хорошо известно, что частное $Z = \frac{X}{Y}$ имеет распределение с плотностью

$$g(y) = \begin{cases} 0, & \text{если } x \leq 0, \\ \frac{1}{B(p, p)} \cdot \frac{y^{p-1}}{(1+y)^{2p}}, & \text{если } x > 0. \end{cases} \quad (2)$$

Распределения, плотность которых имеет вид (2), называются *бета-распределениями* второго рода. Существующая связь между гамма- и бета-распределениями приводит к следующему вопросу. Пусть положительные случайные величины X и Y одинаково распределены, независимы и таковы, что частное $Z = \frac{X}{Y}$ имеет распределение с плотностью (2). Вытекает ли отсюда, что X и Y имеют гамма-распределение?

Покажем, что ответ на поставленный вопрос отрицателен. Для этой цели введем следующие две функции, где $a > 0$, $p > 0$:

$$f_1(x) = \begin{cases} 0, & x \leq 0, \\ \frac{c_1}{e^{a/x} x^{p+1}}, & x > 0, \end{cases} \quad f_2(x) = \begin{cases} 0, & x \leq 0, \\ \frac{c_2 x^p}{(1+x^2)^{p+1/2}}, & x > 0. \end{cases}$$

Нетрудно показать, что при выборе констант $c_1 = \frac{a^p}{\Gamma(p)}$ и $c_2 = 2\Gamma\left(\frac{2p+1}{2}\right) / \left[\Gamma\left(\frac{p}{2}\right)\Gamma\left(\frac{p+1}{2}\right)\right]$ функции f_1 и f_2 будут вероятностными плотностями. Пусть ξ и η — независимые случайные величины, каждая из которых имеет плотность, равную f_1 . Очевидно, что f_1 не имеет

формы (1). Однако непосредственным вычислением с учетом независимости ξ и η можно установить, что плотность g_ζ случайной величины $\zeta = \frac{\xi}{\eta}$ совпадает с плотностью g , задаваемой формулой (2). В точности такой же вывод получается и при рассмотрении функции f_2 в качестве плотности каждой из независимых случайных величин ξ и η .

Следовательно, описанное выше свойство (частное $\frac{X}{Y}$ имеет бета-распределение с плотностью (2)) не является характеризационным свойством гамма-распределения.

§ 13. Разные свойства случайных величин

В этот параграф мы включили примеры, касающиеся различных свойств случайных величин и их характеристик. Некоторые понятия введены в самих примерах. Рассмотренные здесь вопросы являются естественным продолжением и расширением материала предыдущих параграфов.

13.1. О свойстве независимости при суммировании пуассоновских случайных величин

Пусть X и Y — независимые случайные величины, каждая из которых имеет распределение Пуассона. Тогда их сумма $X + Y$ также имеет пуассоновское распределение. В примере 12.3 (а) мы выяснили, что условие независимости величин X и Y существенно для справедливости теоремы Райкова. Найдем теперь ответ на следующий вопрос. Пусть X и Y — неотрицательные целочисленные случайные величины такие, что каждая из величин X , Y и $X + Y$ имеет пуассоновское распределение. Верно ли, что X и Y независимы?

Зафиксируем параметры $\lambda > 0$ и $\mu > 0$ и рассмотрим пуассоновские вероятности $q_i = \frac{\lambda^i e^{-\lambda}}{i!}$ и $r_j = \frac{\mu^j e^{-\mu}}{j!}$, $i = 0, 1, \dots$, $j = 0, 1, \dots$. Введем множества $\Lambda_1 = \{(0, 1), (1, 2), (2, 0)\}$ и $\Lambda_2 = \{(1, 0), (2, 1), (0, 2)\}$ и определим числовую совокупность $\{p_{ij}, i = 0, 1, \dots, j = 0, 1, \dots\}$ следующим образом:

$$p_{ij} = \begin{cases} q_i r_j + \varepsilon, & \text{если } (i, j) \in \Lambda_1, \\ q_i r_j - \varepsilon, & \text{если } (i, j) \in \Lambda_2, \\ q_i r_j & \text{в остальных случаях.} \end{cases} \quad (1)$$

Здесь ε — произвольное число, удовлетворяющее условию

$$|\varepsilon| < \min_{i,j} q_i r_j, \quad \text{где } (i, j) \in \Lambda_1 \cup \Lambda_2.$$

Легко проверить, что $\{p_{ij}\}$ образует двумерное дискретное вероятностное распределение. Предположим, что $\{p_{ij}\}$ — это распределение случайного вектора (ξ, η) . Легко найти, что ξ имеет пуассоновское распределение с параметром λ , а η — пуассоновское распределение с параметром μ . Из (1) вытекает, что случайные величины ξ и η не являются независимыми. Однако (что наиболее интересно здесь) сумма $\xi + \eta$ имеет пуассоновское распределение с параметром $\lambda + \mu$. Это следует из равенства $P\{\xi + \eta = k\} = \frac{(\lambda + \mu)^k e^{-(\lambda + \mu)}}{k!}$. Чтобы установить его, нужно отдельно рассмотреть два случая: (I) $k = 1, 2, 3$ и (II) $k = 0, 4, 5, \dots$ — и использовать явный вид вероятностей p_{ij} из (1).

Следовательно, пуассоновость случайных величин ξ , η и $\xi + \eta$ не влечет независимости ξ и η .

13.2. О симметричности суммы и разности двух случайных величин

Напомним, что случайная величина X называется *симметричной* относительно точки 0, если $X \stackrel{d}{=} (-X)$. В терминах функции распределения F , плотности f и характеристической функции ψ это свойство выражается так: $F(-x) = 1 - F(x)$ для $x \geq 0$; $f(-x) = f(x)$ для всех $x \in \mathbb{R}^1$; $\psi(t)$, $t \in \mathbb{R}^1$, принимает только действительные значения. На двух конкретных примерах проиллюстрируем свойство симметричности при сложении и вычитании случайных величин.

(а) Легко доказать справедливость следующего результата: *пусть X и Y — одинаково распределенные случайные величины. Если X и Y независимы, то их разность $X - Y$ симметрична относительно 0*. Покажем, что условие независимости X и Y существенно для симметричности $X - Y$. Для этой цели рассмотрим случайный вектор (X, Y) , распределение которого задано с помощью следующей таблицы:

	$Y = 1$	$Y = 2$	$Y = 3$
$X = 1$	$\frac{1}{12}$	$\frac{2}{12}$	$\frac{1}{12}$
$X = 2$	$\frac{1}{12}$	0	$\frac{1}{12}$
$X = 3$	$\frac{2}{12}$	0	$\frac{4}{12}$

Легко видеть, что величины X и Y одинаково распределены, причем каждая из них принимает значения 1, 2 и 3 с вероятностью соответственно $\frac{4}{12}$, $\frac{2}{12}$ и $\frac{6}{12}$. Рассмотрим разность $Z = X - Y$. Величина Z

принимает значения $-2, -1, 0, 1$ и 2 соответственно с вероятностью $\frac{2}{12}, \frac{1}{12}, \frac{5}{12}, \frac{3}{12}$ и $\frac{1}{12}$. Очевидно, значения Z симметричны относительно 0 , но Z и $-Z$ имеют разные распределения. Следовательно, случайная величина Z не обладает свойством симметричности, несмотря на то что она есть разность двух одинаково распределенных случайных величин X и Y . Остается заметить, что случайные величины X и Y , заданные с помощью верхней таблицы, не являются независимыми. Следовательно, условие независимости X и Y существенно для симметричности их разности $X - Y$.

(б) Из определения симметричности вытекает, что если X и Y — симметричные и независимые случайные величины, то их сумма $Z = X + Y$ также обладает свойством симметричности. Вопрос: пусть X и Y — независимые случайные величины, такие что X — симметрична и сумма $Z = X + Y$ также симметрична. Имеет ли место свойство симметричности и для величины Y ? Оказывается, это не всегда так.

Пусть ξ — случайная величина с характеристической функцией

$$\psi_{\xi}(t) = \begin{cases} 1 - 2|t|, & \text{если } |t| \leq \frac{1}{2}, \\ 0, & \text{если } |t| > \frac{1}{2}. \end{cases} \quad (1)$$

Очевидно, что величина ξ симметрична.

Рассмотрим еще две характеристические функции:

$$h_1(t) = \begin{cases} 1 - |t|, & |t| \leq \frac{1}{2}, \\ \frac{1}{4|t|}, & |t| > \frac{1}{2}, \end{cases} \quad h_2(t) = \begin{cases} 1 - |t|, & |t| \leq 1, \\ 0, & |t| > 1. \end{cases}$$

Введем случайную величину η с характеристической функцией ψ_{η} , являющейся смесью функций h_1 и h_2 :

$$\psi_{\eta}(t) = \frac{1}{2} e^{it} h_1(t) + \frac{1}{2} e^{-it} h_2(t), \quad t \in \mathbb{R}^1.$$

Элементарное преобразование показывает, что

$$\psi_{\eta}(t) = \begin{cases} (1 - |t|) \cos t, & \text{если } |t| \leq \frac{1}{2}, \\ \frac{e^{it}}{8|t|} + \frac{1}{2} \varepsilon(t) e^{-it} (1 - |t|), & \text{если } |t| > \frac{1}{2}, \end{cases} \quad (2)$$

где $\varepsilon(t) = 1$ при $|t| \leq 1$ и $\varepsilon(t) = 0$ при $|t| > 1$. Из (2) следует, что величина η не является симметричной.

Итак, мы описали две случайные величины ξ и η , причем ξ симметрична, а η — несимметрична. Предположим, что эти величины независимые, и рассмотрим их сумму $\zeta = \xi + \eta$. Поскольку характеристические функции ψ_ζ , ψ_η и ψ_ξ связаны соотношением $\psi_\zeta(t) = \psi_\xi(t)\psi_\eta(t)$, $t \in \mathbb{R}^1$, то из (1) и (2) нетрудно найти, что

$$\psi_\zeta(t) = \begin{cases} (1 - 2|t|)(1 - |t|) \cos t, & \text{если } |t| \leq \frac{1}{2}, \\ 0, & \text{если } |t| > \frac{1}{2}. \end{cases}$$

Очевидно, что ψ_ζ принимает только действительные значения, а это означает, что случайная величина ζ симметрична. Таким образом, мы установили, что из симметричности двух величин, именно ξ и $\zeta = \xi + \eta$, и независимости ξ и η не следует, что случайная величина η симметрична. Возможна, однако, и такая интерпретация: разность, а значит, и сумма двух зависимых случайных величин, которые являются симметричными, не обязательно симметрична.

13.3. Соотношение между классами распределений NBU и IFR

Для ряда прикладных задач теории вероятностей важное значение имеют два класса распределений, обозначаемые NBU и IFR (NBU — New Better than Used; IFR — Increasing Failure Rate).

Рассматриваемые здесь функции распределения будут соответствовать неотрицательным случайным величинам. Они интерпретируются как время жизни или продолжительности работы некоторого прибора. Итак, функция распределения $F(x)$, $x \geq 0$, принадлежит классу NBU, если

$$\bar{F}(x+y) \leq \bar{F}(x)\bar{F}(y) \quad \text{для любых } x, y \geq 0, \quad (1)$$

где $\bar{F} = 1 - F$. В этом случае будем писать $F \in \text{NBU}$. Далее мы говорим, что $F \in \text{IFR}$, если для любого $y > 0$ функция интенсивности отказов $\frac{F(x+y) - F(x)}{\bar{F}(x)}$ возрастает по x , $x > 0$. Предполагается, что $\bar{F}(x) > 0$ для $x > 0$. Это эквивалентно тому, что $h(x) = -\ln \bar{F}(x)$ — выпуклая вниз функция (для тех x , для которых $h(x)$ определена).

Хорошо известно, что $F \in \text{IFR} \implies F \in \text{NBU}$, но в общем случае обратное неверно. Проиллюстрируем это двумя примерами.

Если функция распределения $F \in \text{IFR}$, то легко показать, что F — непрерывная функция на множестве $\{x: F(x) < 1\}$. Поэтому можно спросить: имеет ли место свойство непрерывности для распределений

из класса NBU? Чтобы найти ответ, рассмотрим следующую функцию:

$$G(x) = 1 - 2^{-k} \quad \text{для } x \in (k, k+1], \quad k = 0, 1, 2, \dots$$

Во-первых, эта функция G является функцией распределения, и нетрудно проверить, что для $\bar{G} = 1 - G$ выполнено соотношение (1). Следовательно, $G \in \text{NBU}$. Однако, так как G — разрывная функция, то $G \notin \text{IFR}$.

Ответим теперь на такой вопрос: если функция $H \in \text{NBU}$ и она непрерывна, то верно ли, что $H \in \text{IFR}$? Оказывается, снова ответ отрицателен. Рассмотрим функцию

$$h(x) = \begin{cases} \sin^2 x, & \text{если } x \in \left[0, \frac{\pi}{2}\right], \\ \frac{1}{2} \pi \left(x - \frac{1}{2} \pi\right) + 1, & \text{если } x \in \left(\frac{\pi}{2}, \infty\right). \end{cases}$$

Легко проверить, что

$$H(x) = 1 - \exp[-h(x)], \quad x \geq 0,$$

является функцией распределения и, более того, h удовлетворяет соотношению

$$h(x+y) \geq h(x) + h(y), \quad x, y \geq 0.$$

Отсюда и из (1) вытекает, что $H \in \text{NBU}$. Кроме того, очевидно, что H непрерывна. Тем не менее $H \notin \text{IFR}$, потому что $h(x) = -\ln \bar{H}(x)$ не является выпуклой функцией.

13.4. Перестановочные и «хвостовые» события, связанные с последовательностями случайных величин

Пусть на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ задана последовательность случайных величин $\{X_n, n \geq 1\}$. Через $\mathcal{F}_m^n = \sigma\{X_m, \dots, X_n\}$ обозначим σ -алгебру, порожденную величинами $X_m, \dots, X_n, m \leq n$. Тогда $\bigcup_{k=1}^{\infty} \mathcal{F}_n^{n+k}$ является алгеброй, и пусть $\mathcal{F}_n^{\infty}$ — это порожденная ею σ -алгебра. Последовательность σ -алгебр $\{\mathcal{F}_n^{\infty}, n \geq 1\}$ не возрастает, ее предел существует и является σ -алгеброй. Обозначим этот предел через $\mathcal{G} = \bigcap_{n=1}^{\infty} \mathcal{F}_n^{\infty}$ и назовем $\mathcal{G}$ «хвостовой» (а также «остаточной») σ -алгеброй. Каждое событие $A \in \mathcal{G}$ называется «хвостовым» событием, а каждая функция, заданная на множестве Ω и измеримая относительно σ -алгебры $\mathcal{G}$, — «хвостовой» функцией.

Сформулируем теперь основной результат, касающийся «хвостовых» событий и функций.

Закон «0 или 1» Колмогорова. Пусть $\{X_n, n \geq 1\}$ — последовательность независимых случайных величин и $\mathcal{G}$ — соответствующая ей «хвостовая» σ -алгебра. Тогда для любого «хвостового» события A , $A \in \mathcal{G}$, или $P(A) = 0$, или $P(A) = 1$. Кроме того, любая «хвостовая» случайная величина Y п. н. постоянна, т. е. если $\sigma\{Y\} \in \mathcal{G}$, то $P\{Y = c\} = 1$, $c = \text{const}$.

Введем теперь еще одно понятие. Мы говорим, что случайные величины $X_1, \dots, X_n$ перестановочны (а также симметрично зависимы), если для каждой перестановки $(i_1, i_2, \dots, i_n)$ индексов $(1, 2, \dots, n)$ (число этих перестановок $n!$) случайные векторы $(X_{i_1}, X_{i_2}, \dots, X_{i_n})$ и $(X_1, X_2, \dots, X_n)$ имеют одно и то же распределение. Бесконечная последовательность $\{X_n, n \geq 1\}$ называется перестановочной, если для каждого n случайные величины $X_1, \dots, X_n$ перестановочны. Пусть, функция $g(x_1, x_2, \dots)$ определена для $(x_1, x_2, \dots) \in \mathbb{R}^\infty$ и измерима относительно σ -алгебры $\mathcal{B}^\infty$. Тогда $g(X_1, X_2, \dots)$, являющаяся случайной величиной, называется перестановочной функцией, если $g(\cdot)$ инвариантна относительно всех конечных перестановок своих аргументов, т. е. если $g(X_1, \dots, X_n, X_{n+1}, \dots) = g(X_{i_1}, \dots, X_{i_n}, X_{n+1}, \dots)$. В частности, событие $A \in \mathcal{F}_1^\infty$ называется перестановочным, если его индикатор I_A — перестановочная функция.

Следующий результат играет основную роль при рассмотрении перестановочных событий и функций.

Закон «0 или 1» Хьюитта–Сэвиджа. Пусть $\{X_n, n \geq 1\}$ — последовательность независимых и одинаково распределенных случайных величин. Тогда для любого перестановочного события A , $A \in \mathcal{F}_1^\infty$, или $P(A) = 0$, или $P(A) = 1$.

По поводу введенных выше понятий и доказательств сформулированных результатов читатель может обратиться, например, к книгам [106, 116, 123, 177, 330, 345].

Настоящий и последующие примеры посвящены введенным понятиям и применимости законов «0 или 1».

(а) Первый возникающий здесь вопрос касается «хвостовых и перестановочных» событий. Пусть $\{X_n, n \geq 1\}$ — последовательность случайных величин и $\mathcal{G}$ — ее «хвостовая» σ -алгебра. Если $A \in \mathcal{G}$, то для любой перестановки $(i_1, \dots, i_n)$ индексов $(1, \dots, n)$, $n \geq 1$, можно записать A в виде $A = \{(X_{n+1}, X_{n+2}, \dots) \in B_{n+1}\}$, где B_{n+1} — борелевское множество в $\mathbb{R}^\infty$, т. е. $B_{n+1} \in \mathcal{B}^\infty$. Таким образом, для каждого n

$$A = \{(X_1, X_2, \dots) \in \mathbb{R}^n \times B_{n+1}\} = \{(X_{i_1}, \dots, X_{i_n}, X_{n+1}, \dots) \in \mathbb{R}^n \times B_{n+1}\},$$

и так как $B_\infty = \mathbb{R}^n \times B_{n+1}$ является борелевским множеством в $\mathbb{R}^\infty$, то A — перестановочное событие.

Итак, каждое «хвостовое» событие является перестановочным. Однако, обратное в общем случае неверно. Самый простой пример — это событие $A = \{X_n = 0 \text{ для всех } n \geq 1\}$. Очевидно, A является перестановочным событием. Однако $A \notin \mathcal{F}_n^\infty$ для каждого $n \geq 1$. Следовательно, A «хвостовым» событием не является.

(б) Опишем ситуацию, когда закон «0 или 1» Колмогорова неприменим, а закон «0 или 1» Хьюитта–Сэвиджа применить можно. Пусть $X_1, X_2, \dots$ — независимые и одинаково распределенные случайные величины, и пусть $S_n = \sum_{k=1}^n X_k$. Тогда для любого борелевского множества B , $B \in \mathcal{B}^1$, событие $\{S_n \in B \text{ б. ч.}\}$ является перестановочным и не является «хвостовым». Таким образом, закон «0 или 1» Колмогорова нельзя применить, но с помощью закона «0 или 1» Хьюитта–Сэвиджа мы можем заключить, что $P\{S_n \in B \text{ б. ч.}\} = 0$ или 1.

Продолжим это рассуждение. Пусть $\{\xi_n, n \geq 1\}$ — последовательность независимых бернуллиевских случайных величин с $P\{\xi_n = 1\} = p = 1 - P\{\xi_n = -1\}$, $0 < p < 1$. Здесь важной задачей является изучение поведения случайного блуждания $S_n = \xi_1 + \dots + \xi_n$, $n \geq 1$. В частности, интересно найти вероятность события $\{S_n = 0 \text{ б. ч.}\}$. Как уже отмечалось выше, это событие не является «хвостовым», но оно перестановочно, и, следовательно, $P\{S_n = 0 \text{ б. ч.}\} = 0$ или 1. Оказывается, что значение вероятности $P\{S_n = 0 \text{ б. ч.}\}$ зависит существенно от параметра p . Различными способами можно показать, что если $p = \frac{1}{2}$, то $P\{S_n = 0 \text{ б. ч.}\} = 1$. При $p \neq \frac{1}{2}$ с помощью Бореля–Кантелли легко найти, что $P\{S_n = 0 \text{ б. ч.}\} = 0$.

13.5. Перестановочная последовательность случайных величин, не обладающих свойством независимости

Если последовательность $\{X_n, n \geq 1\}$, состоит из независимых и одинаково распределенных случайных величин, то для любой перестановки $(i_1, \dots, i_n)$ индексов $(1, \dots, n)$ случайные векторы $(X_{i_1}, \dots, X_{i_n})$ и $(X_1, \dots, X_n)$ имеют одно и то же распределение. Следовательно, $\{X_n, n \geq 1\}$ — перестановочная последовательность. С помощью конкретного примера покажем, что обратное, вообще говоря, неверно.

Итак, пусть $\xi_n, n \geq 1$ — независимые и одинаково распределенные случайные величины, а η — другая случайная величина, не зависящая

от ξ_n , $n \geq 1$. Положим

$$Y_n = \xi_n + \eta, \quad n \geq 1. \quad (1)$$

Покажем теперь, что $\{Y_n, n \geq 1\}$ — перестановочная последовательность, которая, однако, не обладает свойством независимости. Не ограничивая общности, можно считать, что $E\xi_n = 0$, $D\xi_n < \infty$, $E\eta = 0$ и $D\eta < \infty$. Тогда для каждого $n \geq 1$ имеем

$$DY_n = E\{(\xi_n + \eta)^2\} = E\{\xi_n^2\} + 2E\{\xi_n\eta\} + E\{\eta^2\} = D\xi_n + D\eta.$$

Легко видеть, что случайные векторы $(Y_{i_1}, \dots, Y_{i_k})$ и $(Y_{j_1}, \dots, Y_{j_k})$, где $k \geq 1$, имеют одно и то же распределение для любых k индексов $(i_1, \dots, i_k)$ и $(j_1, \dots, j_k)$, выбранных из множества натуральных чисел $\mathbb{N}$. Это означает, что последовательность $\{Y_n, n \geq 1\}$ перестановочна. Кроме того, $\{Y_n, n \geq 1\}$ характеризуется общим коэффициентом корреляции, т. е. существует число ρ , $|\rho| \leq 1$, такое, что $\rho(Y_i, Y_j) = \rho$ для любых $i \neq j$. Возьмем, например, величины Y_1 и Y_2 и выразим $\rho(Y_1, Y_2)$ через характеристики ξ_n и η_n , используя равенства (1). С учетом условий на ξ_n и η нетрудно найти, что

$$\rho = \rho(Y_1, Y_2) = \frac{E\{Y_1 Y_2\} - EY_1 EY_2}{(DY_1 DY_2)^{1/2}} = \frac{D\eta}{D\xi_1 + D\eta}.$$

Отсюда следует, что если $D\eta \neq 0$, т. е. если случайная величина η невырождена, то $\rho \neq 0$. А раз $\rho \neq 0$, то величины Y_1 и Y_2 не могут быть независимыми, потому что они даже не являются некоррелированными.

Следовательно, свойство независимости не имеет места для перестановочной последовательности $\{Y_n, n \geq 1\}$.

13.6. О применимости закона «0 или 1» Колмогорова

Пусть $\{X_n, n \geq 1\}$ — последовательность случайных величин, а $\mathcal{G}$ — соответствующая ей «хвостовая» σ -алгебра. Верно ли, что любое «хвостовое» событие A , $A \in \mathcal{G}$, имеет вероятность $P(A) = 0$ или 1? Это так, если величины $\{X_n, n \geq 1\}$ независимы, т. е. в случае, когда применим закон «0 или 1» Колмогорова. А что будет, если последовательность $\{X_n, n \geq 1\}$ удовлетворяет более слабому условию, чем независимость ее членов? Согласно примеру 13.5, в качестве такого более слабого условия можно взять перестановочность.

Итак, наша цель — рассмотреть перестановочную последовательность случайных величин $\{X_n, n \geq 1\}$ и показать, что в ее «хвостовой» σ -алгебре $\mathcal{G}$ можно найти событие A , вероятность $P(A)$ которого отлична от 0 и от 1. Пусть ξ_n , $n \geq 1$, — независимые и одинаково распределенные случайные величины, где $P\{\xi_1 = 1\} = P\{\xi_1 = -1\} = \frac{1}{2}$.

Пусть случайная величина η не зависит от ξ_n , $n \geq 1$, причем $P\{\eta = 1\} = P\{\eta = 0\} = \frac{1}{2}$. Положим $X_n = \xi_n \eta$. Легко доказать, что $\{X_n, n \geq 1\}$ — перестановочная последовательность. Очевидно, что для каждого $m \geq 1$ события $\{X_m = 0\}$ и $\{\eta = 0\}$ эквивалентны. Отсюда вытекает, что $A = \{\eta = 0\} \in \mathcal{F}_n^\infty$ для каждого $n \geq 1$, где $\mathcal{F}_n^\infty = \sigma\{X_n, X_{n+1}, \dots\}$. Следовательно, $A \in \bigcap_{n=1}^{\infty} \mathcal{F}_n^\infty$, т. е. $A \in \mathcal{G}$. Однако по предположению $P\{\eta = 0\} = \frac{1}{2}$. Таким образом, мы нашли «хвостовое» событие A с $P(A) = \frac{1}{2}$. Заметим, что этот факт не так уж удивителен, потому что для последовательности $\{X_n, n \geq 1\}$ не выполнено свойство независимости. Другими словами, в законе «0 или 1» Колмогорова нельзя ослабить требование независимости для последовательности $\{X_n, n \geq 1\}$, заменяя его, например, перестановочностью.

13.7. О применимости закона «0 или 1» Хьюитта–Сэвиджа

С помощью двух примеров покажем, что для справедливости закона «0 или 1» Хьюитта–Сэвиджа важна не только одинаковая распределенность членов последовательности.

(а) Рассмотрим последовательность $\{X_n, n \geq 1\}$, где $P\{X_1 = 1\} = \frac{1}{2} = P\{X_1 = -1\}$ и $P\{X_n = 0\} = 1$ для каждого $n \geq 2$. Очевидно, что все эти величины независимы. Событие

$$A = \left\{ \sum_{j=1}^n X_j > 0 \text{ для бесконечно многих } n \right\}$$

является перестановочным, но не является «хвостовым» для последовательности $\{X_n, n \geq 1\}$. Ясно, что $P(A) = P\{X_1 > 0\} = \frac{1}{2}$, т. е. $P(A)$ не равна ни 0, ни 1, как должно быть, если бы закон «0 или 1» Хьюитта–Сэвиджа имел место. Однако это не так, потому что величины X_n , $n \geq 1$, не распределены одинаково.

(б) Пусть X_n , $n \geq 1$ — независимые случайные величины, для которых $P\{X_n = 1\} = 2^{-n} = 1 - P\{X_n = 0\}$, $n \geq 1$, и пусть $A = \{X_n = 0 \text{ для всех } n \geq 1\}$. Тогда A — перестановочное, но не «хвостовое» событие. Для вероятности $P(A)$ имеем

$$P(A) = \prod_{n=1}^{\infty} P\{X_n = 0\} = \prod_{n=1}^{\infty} (1 - 2^{-n}).$$

Поскольку $\sum_{n=1}^{\infty} 2^{-n} < \infty$, то бесконечное произведение $\prod_{n=1}^{\infty} (1 - 2^{-n})$ сходится к пределу, который больше 0 и меньше 1. Таким образом, и здесь

закон «0 или 1» Хьюитта–Сэвиджа не применим. Подчеркнем еще раз, что рассмотренные случайные величины X_n , $n \geq 1$ независимы, принимают одни и те же значения, но с различными вероятностями, т. е. эти величины не распределены одинаково.

13.8. Всегда ли применима теорема де Финетти?

Сформулируем сначала следующий результат, называемый *теоремой де Финетти* (см., например, [106, 123, 461]): *пусть $\{X_n, n \geq 1\}$ — бесконечная перестановочная последовательность случайных величин, каждая из которых принимает только два значения, 0 и 1. Тогда существует вероятностная мера μ на борелевской σ -алгебре $\mathcal{B}_{[0,1]}$ отрезка $[0, 1]$ такая, что*

$$P\{X_1 = \varepsilon_1, X_2 = \varepsilon_2, \dots, X_n = \varepsilon_n\} = \int_0^1 p^k (1-p)^{n-k} \mu(dp) \quad (1)$$

для каждого n , где $\varepsilon_j = 0$ или 1 и $k = \varepsilon_1 + \dots + \varepsilon_n$. Другими словами, среди первых n членов бесконечной последовательности $X_1, X_2, \dots$ распределение числа исходов 0 или 1 является смесью биномиального распределения с некоторым невырожденным вероятностным распределением на отрезке $[0, 1]$. Поэтому можно задать следующий вопрос: применима ли теорема де Финетти для конечной перестановочной последовательности случайных величин? Покажем, что это не всегда так.

Рассмотрим совокупность из двух случайных величин $\{X_1, X_2\}$, определенных таким образом:

$$\begin{aligned} P\{X_1 = 0, X_2 = 1\} &= P\{X_1 = 1, X_2 = 0\} = \frac{1}{2}, \\ P\{X_1 = 0, X_2 = 0\} &= P\{X_1 = 1, X_2 = 1\} = 0. \end{aligned}$$

Очевидно, что величины X_1 и X_2 перестановочны. Допустим далее, что имеет место представление (1) с некоторой мерой μ . Тогда из равенств

$$\begin{aligned} P\{(X_1, X_2) = (0, 0)\} &= 0, \quad P\{(X_1, X_2) = (1, 1)\} = 0 \quad \text{и} \\ P\{(X_1, X_2) \in \{(0, 1), (1, 0)\}\} &= 1 \end{aligned}$$

легко выводим, что должны выполняться три соотношения:

$$\int_0^1 (1-p)^2 \mu(dp) = 0, \quad \int_0^1 p^2 \mu(dp) = 0, \quad 2 \int_0^1 p(1-p) \mu(dp) = 1.$$

Это, однако, невозможно. Следовательно, теорема де Финетти не применима для конечной совокупности перестановочных случайных величин.

Г Л А В А 3

ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ

§ 14. Разные виды сходимости последовательностей случайных величин

Если $\{X_n, n \geq 1\}$ — последовательность случайных величин, заданных на вероятностном пространстве $(\Omega, \mathcal{F}, P)$, то важной задачей является нахождение предела $\lim_{n \rightarrow \infty} X_n$. Предел можно понимать по-разному, и все зависит от рассматриваемого вида сходимости. Напомним сначала определения четырех основных видов сходимости.

(а) Последовательность случайных величин $\{X_n, n \geq 1\}$ называется *сходящейся с вероятностью 1 (почти наверное, почти всюду)* к случайной величине X при $n \rightarrow \infty$, если

$$P\left\{\omega: \lim_{n \rightarrow \infty} X_n(\omega) = X(\omega)\right\} = 1.$$

Эту сходимость обозначают так: $X_n \rightarrow X$ (P-п.н.) или $X_n \xrightarrow{\text{п.н.}} X$.

(б) Мы говорим, что последовательность $\{X_n, n \geq 1\}$ *сходится по вероятности* к случайной величине X при $n \rightarrow \infty$ ($X_n \xrightarrow{P} X$), если для любого $\varepsilon > 0$ имеет место соотношение

$$\lim_{n \rightarrow \infty} P\left\{\omega: |X_n(\omega) - X(\omega)| \geq \varepsilon\right\} = 0.$$

(в) Последовательность случайных величин $\{X_n, n \geq 1\}$ называется *сходящейся по распределению* к случайной величине X при $n \rightarrow \infty$ ($X_n \xrightarrow{d} X$), если соотношение

$$\lim_{n \rightarrow \infty} E\{g(X_n)\} = E\{g(X)\} \tag{1}$$

выполнено для любой непрерывной и ограниченной функции $g(x)$, $x \in \mathbb{R}^1$. Это эквивалентно следующему: если F_n и F — функции распределения соответственно случайных величин X_n и X , то

$$\lim_{n \rightarrow \infty} F_n(x) = F(x) \tag{2}$$

для всех $x \in \mathbb{R}^1$, являющихся точками непрерывности F .

Отметим, что в общей ситуации соотношением типа (1) задается так называемая *слабая сходимость* последовательности распределений $\{F_n, n \geq 1\}$ к распределению F при $n \rightarrow \infty$ и используется обозначение $F_n \xrightarrow{w} F$. (Вопросам слабой сходимости вероятностных мер и распределений посвящен §16.)

(г) Предположим, что случайные величины X и $X_n, n \geq 1$, принадлежат пространству $\mathcal{L}^r, 0 < r < \infty$ (т. е. $E\{|X|^r\} < \infty, E\{|X_n|^r\} < \infty$). Мы говорим, что последовательность случайных величин $\{X_n, n \geq 1\}$ *сходится в среднем порядка r* к случайной величине X при $n \rightarrow \infty$, и записываем это $X_n \xrightarrow{L^r} X$, если

$$\lim_{n \rightarrow \infty} E\{|X_n - X|^r\} = 0.$$

Особую роль в теории вероятностей играет частный случай L^r -сходимости, когда $r = 2$, и это так называемая *среднеквадратическая сходимость*.

Связь между введенными четырьмя видами сходимости изображена на следующей схеме:



Отметим, что есть еще ряд других видов сходимости последовательностей случайных величин. Некоторые из них будут рассмотрены далее в примерах.

Оказывается, при выяснении связей между разными видами сходимости возникают интересные, а часто и нетривиальные вопросы, ответы на которые требуют специальных конструкций. Тема «предельные теоремы для последовательностей случайных величин» является основной в теории вероятностей, и ей уделяется значительное внимание во многих книгах; см., например, [6, 9, 24, 60, 67, 75, 92, 106, 116, 117, 118, 127, 128, 177, 179, 231, 345, 358, 374, 404, 425].

Рассмотренные в настоящем параграфе примеры иллюстрируют связи между основными видами сходимости. Несколько примеров посвящены близким вопросам.

14.1. О сходимости функций случайных величин

Покажем, что если X_n сходится к X по распределению или по вероятности, то не для любой функции g имеет место сходимость $g(X_n)$ к $g(X)$ в том же смысле.

(а) Пусть последовательность случайных величин $\{X_n, n \geq 1\}$ и случайная величина X таковы, что $X_n \xrightarrow{d} X$ при $n \rightarrow \infty$. Покажем, что существует ограниченная и измеримая функция $g(x)$, $x \in \mathbb{R}^1$, такая что $g(X_n) \not\xrightarrow{d} g(X)$ при $n \rightarrow \infty$. Действительно, рассмотрим (вырожденные) величины $X \equiv 0$ и $X_n = \frac{1}{n}$, $n \geq 1$, и пусть F и F_n , $n \geq 1$, — их функции распределения. Тогда очевидно, что

$$F(x) = \begin{cases} 0, & \text{если } x \leq 0, \\ 1, & \text{если } x > 0, \end{cases} \quad F_n(x) = \begin{cases} 0, & \text{если } x \leq \frac{1}{n}, \\ 1, & \text{если } x > \frac{1}{n}. \end{cases}$$

Легко видеть, что функции распределения F_n , $n \geq 1$, и F удовлетворяют соотношению (2) и, следовательно, $X_n \xrightarrow{d} X$ при $n \rightarrow \infty$. Возьмем теперь функцию

$$g(x) = \begin{cases} 0, & \text{если } x \leq 0, \\ 1, & \text{если } x > 0. \end{cases} \quad (3)$$

Функция g — ограниченная и измеримая (но не непрерывная), и очевидно, что $g(X_n) = 1$ для каждого $n \geq 1$, в то время как $g(X) = 0$. Отсюда следует, что ни в каком смысле $g(X_n)$ не может сходиться к $g(X)$. В частности, $g(X_n) \not\xrightarrow{d} g(X)$ при $n \rightarrow \infty$.

Отметим, наконец, что этот элементарный прием показывает также, что в определении сходимости по распределению с помощью равенства (1) требование непрерывности функции g существенно.

(б) Рассмотрим две последовательности случайных величин $\{Y_n, n \geq 1\}$ и $\{Z_n, n \geq 1\}$, где $P\{Y_n = 1\} = \frac{1}{n} = 1 - P\{Y_n = \frac{1}{n}\}$, а $Z_n \sim \mathcal{N}\left(0, \frac{\sigma^2}{n}\right)$. Тогда для любого $\varepsilon > 0$

$$P\{|Y_n - 0| \geq \varepsilon\} = P\{Y_n = 1\} = \frac{1}{n} \rightarrow 0 \quad \text{при } n \rightarrow \infty,$$

и, значит, $Y_n \xrightarrow{P} Y$ при $n \rightarrow \infty$, где случайная величина $Y = 0$ п. н. Возьмем снова функцию g , задаваемую соотношением (3). Очевидно, что $g(Y_n) = 1$ для каждого $n \geq 1$, а $g(Y) = 0$. Следовательно, $g(Y_n)$ не может сходиться к $g(Y)$ в любом смысле. В частности, $g(Y_n) \not\xrightarrow{P} g(Y)$ при $n \rightarrow \infty$.

Далее, применяя неравенство Чебышёва, мы видим, что

$$P\{|Z_n - 0| \geq \varepsilon\} \leq \frac{1}{\varepsilon^2} DZ_n = \frac{\sigma^2}{\varepsilon^2 n} \rightarrow 0 \quad \text{при } n \rightarrow \infty,$$

и, следовательно, $Z_n \xrightarrow{P} Z$ при $n \rightarrow \infty$, где $Z = 0$ п. н. Однако случайная величина Z_n симметрична. Поэтому $P\{g(Z_n) = 0\} = P\{g(Z_n) = 1\} = \frac{1}{2}$.

Поскольку $g(Z) = 0$ п. н., то приходим к заключению, что $g(Z_n) \not\xrightarrow{P} g(Z)$ при $n \rightarrow \infty$.

14.2. Из сходимости по распределению не вытекает сходимость по вероятности

Хорошо известно, что $X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$. Покажем, что обратное не верно.

(а) Пусть X — бернуллиевская случайная величина, т. е. $P\{X = 1\} = P\{X = 0\} = \frac{1}{2}$. Пусть далее $\{X_n, n \geq 1\}$ — последовательность случайных величин таких, что $X_n = X$ для каждого n . Поскольку $X_n \stackrel{d}{=} X$, то $X_n \xrightarrow{d} X$ при $n \rightarrow \infty$. Введем новую случайную величину $Y = 1 - X$. Очевидно, что $X \stackrel{d}{=} Y$, и, значит, $X_n \xrightarrow{d} Y$. Однако легко видеть, что X_n не может сходиться в любом другом смысле к Y , так как $|X_n - Y| = 1$. В частности, $P\{|X_n - Y| > \varepsilon\} \not\rightarrow 0$ для любого $\varepsilon \in (0, 1)$, и, следовательно, $X_n \not\xrightarrow{P} Y$ при $n \rightarrow \infty$.

(б) Пусть X — любая симметричная случайная величина, например $X \sim \mathcal{N}(0, 1)$. Положим $X_n = -X$ для $n \geq 1$. Тогда $X_n \stackrel{d}{=} X$ и, значит, $X_n \xrightarrow{d} X$ при $n \rightarrow \infty$. Так как для произвольного $\varepsilon > 0$ мы имеем

$$P\{|X_n - X| > \varepsilon\} = P\left\{|X| > \frac{\varepsilon}{2}\right\} = 2\left(1 - \Phi\left(\frac{\varepsilon}{2}\right)\right) > 0,$$

то $X_n \not\xrightarrow{P} X$.

14.3. Последовательности случайных величин, сходящиеся по вероятности, но не с вероятностью 1

(а) Рассмотрим вероятностное пространство $(\Omega, \mathcal{F}, P)$, где $\Omega = [0, 1]$, $\mathcal{F} = \mathcal{B}_{[0,1]}$ и P — лебеговская мера. Для каждого $n \in \mathbb{N}$ существует единственная пар целых чисел k и m , где $m \geq 0$, $0 \leq k \leq 2^m - 1$, таких что $n = 2^m + k$. Определим последовательность событий $A_n = \left[\frac{k}{2^m}, \frac{k+1}{2^m}\right)$ и положим $X_n = X_n(\omega) = 1_{A_n}(\omega)$. Так мы получаем последовательность

случайных величин $\{X_n, n \geq 1\}$. Очевидно, что

$$P\{|X_n| \geq \varepsilon\} = \begin{cases} \frac{1}{2^m}, & \text{если } 0 < \varepsilon < 1, \\ 0, & \text{если } \varepsilon \geq 1. \end{cases}$$

Поскольку $n \rightarrow \infty \Rightarrow m \rightarrow \infty$ и наоборот, то легко заключить, что

$$X_n \xrightarrow{P} 0 \quad \text{при } n \rightarrow \infty. \quad (1)$$

Выясним теперь, можно ли в (1) сходимость по вероятности заменить сходимостью с вероятностью 1.

Нетрудно показать, что для каждого фиксированного $\omega \in \Omega$ существует бесконечно много индексов n , для которых $X_n(\omega) = 1$, а также бесконечно много n , для которых $X_n(\omega) = 0$. Действительно, при фиксированном m имеется единственное значение k , $0 \leq k \leq 2^m - 1$, такое, что интервал $\left[\frac{k}{2^m}, \frac{k+1}{2^m}\right)$ содержит заранее выбранное ω , т.е. $\omega \in A_n$ для $n = 2^m + k$, а тогда $\omega \notin \Omega \setminus A_n$. Можно, однако, так подобрать n , чтобы $\omega \in \Omega \setminus A_n$, а тогда $\omega \notin A_n$. Отсюда можно заключить, что $\omega \in A_n$ бесконечно часто (т.е. для бесконечно большого числа индексов n), а также, что $\omega \in \Omega \setminus A_n$ бесконечно часто. В терминах величин X_n , $n \geq 1$, это записывается так: $\lim_n X_n = 1$, $\underline{\lim}_n X_n = 0$. Следовательно,

$$X_n \not\xrightarrow{\text{п.н.}} 0 \quad \text{при } n \rightarrow \infty.$$

(б) Пусть последовательность $\{X_n, n \geq 1\}$ состоит из независимых случайных величин, каждая из которых принимает два значения: 0 и 1, причем $P\{X_n = 1\} = p_n = 1 - P\{X_n = 0\}$. Используя лемму Бореля–Кантелли, получаем, что

$$X_n \xrightarrow{\text{п.н.}} 0 \quad \text{при } n \rightarrow \infty \iff \sum_{n=1}^{\infty} p_n < \infty. \quad (2)$$

Теперь, основываясь на соотношении (2), легко построить случайную последовательность, сходящуюся по вероятности и не сходящуюся с вероятностью 1. Действительно, пусть числа p_n , $n \geq 1$, выбраны так, что $\lim_{n \rightarrow \infty} p_n = 0$ и $\sum_{n=1}^{\infty} p_n = \infty$ (например, $p_n = \frac{1}{n}$). Поскольку $\sum_{n=1}^{\infty} p_n = \infty$, то X_n при $n \rightarrow \infty$ не может стремиться к 0. С другой стороны, числа $q_n = 1 - p_n = P\{X = 0\}$ таковы, что $\sum_{n=1}^{\infty} q_n = \infty$ (поскольку $q_n \rightarrow 0$). Это означает, что X_n при $n \rightarrow \infty$ не может стремиться к 1. Но 1 и 0 — это единственно возможные значения для членов последовательности X_n ,

$n \geq 1$. Следовательно, $\{X_n, n \geq 1\}$ не сходится с вероятностью 1. Однако при $\varepsilon > 0$

$$P\{|X_n - 0| \geq \varepsilon\} = P\{X_n = 1\} = p_n \rightarrow 0,$$

так что $X_n \xrightarrow{P} 0$ при $n \rightarrow \infty$.

14.4. Из сходимости в среднем порядка r вытекает сходимость по вероятности, но обратное неверно

Если случайные величины X и $X_n, n \geq 1$, принадлежат пространству $\mathcal{L}^r, (r > 0)$ и $X_n \xrightarrow{L^r} X$ при $n \rightarrow \infty$, то с помощью неравенства Чебышёва сразу получаем, что $X_n \xrightarrow{P} X$ при $n \rightarrow \infty$. Покажем, что обратное неверно.

(а) Пусть последовательность случайных величин задана так: $P\{X_n = e^n\} = \frac{1}{n} = 1 - P\{X_n = 0\}, n \geq 1$. Тогда для любого $\varepsilon > 0$

$$P\{|X_n| < \varepsilon\} = P\{X_n = 0\} = 1 - \frac{1}{n} \rightarrow 1 \quad \text{при } n \rightarrow \infty,$$

и, значит, $X_n \xrightarrow{P} 0$ при $n \rightarrow \infty$. Однако для каждого $r > 0$

$$E\{X_n^r\} = \frac{e^{rn}}{n} \rightarrow \infty \quad \text{при } n \rightarrow \infty,$$

и, следовательно, $X_n \not\xrightarrow{L^r} 0$ при $n \rightarrow \infty$.

(б) Рассмотрим последовательность $\{X_n, n \geq 1\}$, где X_n имеет распределение Коши с плотностью $f_n(x) = \frac{n/\pi}{1 + n^2 x^2}, x \in \mathbb{R}^1$. Тогда при $\varepsilon > 0$

$$P\{|X_n| < \varepsilon\} = \int_{-\varepsilon}^{\varepsilon} f_n(x) dx = \frac{2}{\pi} \arctg(n\varepsilon) \rightarrow 1 \quad \text{при } n \rightarrow \infty,$$

и, следовательно, $X_n \xrightarrow{P} 0$ при $n \rightarrow \infty$. Заметим далее, что $E\{|X_n|^r\} = \infty$ для любого $r \geq 1$, так что нет смысла говорить о L^r -сходимости $\{X_n\}$, если $r \geq 1$. Однако можно показать, что если $0 < r < 1$, то $E\{|X_n|^r\} < \infty$

и, более того, $X_n \xrightarrow{L^r} 0$ при $n \rightarrow \infty$.

(в) Последовательности случайных величин $\{Y_n, n \geq 2\}$ и $\{Z_n, n \geq 1\}$ заданы следующим образом:

$$P\{Y_n = 1\} = \frac{1}{\ln n} = 1 - P\{Y_n = 0\}, \quad P\{Z_n = 0\} = 1 - \frac{1}{n^\alpha},$$

$$P\{Z_n = n\} = P\{Z_n = -n\} = \frac{1}{2n^\alpha}, \quad 0 < \alpha \leq 2.$$

Тогда нетрудно показать, что для любого $r > 0$ при $n \rightarrow \infty$: $Y_n \xrightarrow{P} 0$, но $Y_n \not\xrightarrow{L^r} 0$; $Z_n \xrightarrow{P} 0$, но $Z_n \not\xrightarrow{L^r} 0$.

14.5. Из сходимости в среднем порядка r не всегда вытекает сходимость почти наверное

(а) Рассмотрим снова последовательность $\{X_n, n \geq 1\}$ из примера 14.3 (а), т. е. $X_n = X_n(\omega) = 1_{A_n}(\omega)$ для $n = 2^m + k$, где $A_n = \left[\frac{k}{2^m}, \frac{k+1}{2^m}\right)$. Поскольку $\omega \in \Omega = [0, 1]$ и P — мера Лебега, то $E\{|X_n|\} = E\{X_n\} = 2^{-m} \rightarrow 0$ при $n \rightarrow \infty$. Следовательно, $X_n \xrightarrow{L^1} 0$ при $n \rightarrow \infty$, но, как мы установили в цитированном примере, $X_n \not\xrightarrow{п.н.} 0$ при $n \rightarrow \infty$.

(б) Пусть $\{Y_n, n \geq 1\}$ — последовательность независимых случайных величин, где

$$P\{Y_n = 0\} = 1 - \frac{1}{n^{1/4}}, \quad P\{Y_n = 1\} = P\{Y_n = -1\} = \frac{1}{2n^{1/4}}, \quad n \geq 1.$$

Элементарное вычисление показывает, что $Y_n \xrightarrow{L^2} 0$ при $n \rightarrow \infty$. Однако, используя, например, лемму Бореля–Кантелли, нетрудно установить, что $Y_n \not\xrightarrow{п.н.} 0$ при $n \rightarrow \infty$.

14.6. Из сходимости почти наверное не всегда следует сходимость в среднем порядка r

(а) Пусть последовательность случайных величин $\{X_n, n \geq 1\}$ задана следующим образом:

$$P\{X_n = 0\} = 1 - \frac{1}{n^\alpha}, \quad P\{X_n = n\} = P\{X_n = -n\} = \frac{1}{2n^\alpha}, \quad \alpha > 0.$$

Поскольку $E\{|X_n|^{1/2}\} = n^{1/2-\alpha}$, находим, что $\sum_{n=1}^{\infty} E\{|X_n|^{1/2}\} < \infty$ для каждого $\alpha > \frac{3}{2}$. Согласно неравенству Маркова, $P\{|X_n| > \varepsilon\} \leq \varepsilon^{-1/2} E\{|X_n|^{1/2}\}$, и, следовательно, $\sum_{n=1}^{\infty} P\{|X_n| > \varepsilon\} < \infty$ для любого $\varepsilon > 0$. Остается воспользоваться леммой Бореля–Кантелли и заключить, что $X_n \xrightarrow{п.н.} 0$ при $n \rightarrow \infty$. Это верно для $\alpha > \frac{3}{2}$.

Равенство $E\{X_n^2\} = n^{2-\alpha}$ показывает, что $X_n \not\xrightarrow{L^2} 0$ при $n \rightarrow \infty$ для каждого $\alpha \leq 2$. Следовательно, если $\alpha \in \left[\frac{3}{2}, 2\right]$ и $n \rightarrow \infty$, то $X_n \xrightarrow{п.н.} 0$, но $X_n \not\xrightarrow{L^2} 0$.

(б) Пусть $\{Y_n, n \geq 1\}$ — последовательность случайных величин, где Y_n принимает значения e^n и 0 с вероятностями $\frac{1}{n^2}$ и $1 - \frac{1}{n^2}$ соответственно. Тогда для любого $\varepsilon > 0$

$$P\{|X_n| > \varepsilon\} = P\{X_n > \varepsilon\} = P\{X_n = e^n\} = \frac{1}{n^2},$$

и, значит,

$$\sum_{n=1}^{\infty} P\{|X_n|\} = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty.$$

Отсюда заключаем (см. случай (а) выше), что $X_n \xrightarrow{\text{п.н.}} 0$ при $n \rightarrow \infty$.

Заметим, что при каждом $r > 0$

$$E\{|X_n|^r\} = E\{X_n^r\} = \frac{e^{rn}}{n^2} \rightarrow \infty \quad \text{при } n \rightarrow \infty.$$

Следовательно, $X_n \xrightarrow{\text{п.н.}} 0$, но $X_n \not\xrightarrow{L^r} 0$ при $r > 0$.

14.7. Сходимость плотностей обеспечивает слабую сходимость функций распределений, но обратное неверно

Напомним сначала следующий результат, известный как *теорема Шеффе* (см., например, [6, 447]): пусть F и $F_n, n \geq 1$, — абсолютно непрерывные функции распределения, а f и $f_n, n \geq 1$, — соответствующие им плотности. Тогда, если $f_n(x) \rightarrow f(x)$ при $n \rightarrow \infty$ для почти всех $x \in \mathbb{R}^1$ (относительно меры Лебега), то $F_n \xrightarrow{w} F$ при $n \rightarrow \infty$. Вопрос: верно ли обратное утверждение? Оказывается, ответ отрицателен, как будет видно из примера далее.

Рассмотрим функцию $F_n(x), x \in \mathbb{R}^1, n \geq 1$, где

$$F_n(x) = \begin{cases} 0, & \text{если } x \leq 0, \\ x\left(1 - \frac{\sin(2n\pi x)}{2n\pi x}\right), & \text{если } 0 < x \leq 1, \\ 1, & \text{если } x > 1. \end{cases}$$

Тогда F_n является функцией распределения, и, более того, она абсолютно непрерывна с плотностью

$$f_n(x) = 1 - \cos(2n\pi x), \quad x \in \mathbb{R}^1.$$

Введем также функции

$$F(x) = \begin{cases} 0, & \text{если } x \leq 0, \\ x, & \text{если } 0 < x \leq 1, \\ 1, & \text{если } x > 1, \end{cases} \quad f(x) = \begin{cases} 1, & \text{если } x \in (0, 1], \\ 0, & \text{если } x \notin (0, 1]. \end{cases}$$

Очевидно, что F — функция распределения, а f — плотность равномерного распределения на интервале $(0, 1]$. Легко видеть, что

$$\lim_{n \rightarrow \infty} F_n(x) = F(x) \text{ для всех } x \in \mathbb{R}^1, \text{ т. е. } F_n \xrightarrow{w} F.$$

Однако последовательность плотностей $\{f_n, n \geq 1\}$ расходится при $n \rightarrow \infty$ для каждого $x \in \mathbb{R}^1$, т. е. $f_n(x) \not\rightarrow f(x)$ при $n \rightarrow \infty$.

Итак, мы установили, что в общем случае $F_n \xrightarrow{w} F \not\Rightarrow f_n \rightarrow f$.

14.8. Из сходимости $X_n \xrightarrow{d} X$ и $Y_n \xrightarrow{d} Y$ не всегда вытекает, что $X_n + Y_n \xrightarrow{d} X + Y$

Начнем с формулировки следующего результата (см., например, [261]): *пусть случайные величины $X, X_n, n \geq 1$, и $Y, Y_n, n \geq 1$, заданы на одном и том же вероятностном пространстве. Предположим, что X и Y независимы, а также, что X_n и Y_n независимы при всех $n \geq 1$. Тогда из сходимости $X_n \xrightarrow{d} X$ и $Y_n \xrightarrow{d} Y$ при $n \rightarrow \infty$ вытекает, что $X_n + Y_n \xrightarrow{d} X + Y$ при $n \rightarrow \infty$.*

Покажем с помощью двух примеров, что условие независимости существенно для справедливости этого результата.

(а) Пусть $\{X_n, n \geq 1\}$ — независимые и одинаково распределенные случайные величины, где $P\{X_n = 1\} = P\{X_n = 0\} = \frac{1}{2}$. Положим $Y_n = 1 - X_n$. Тогда при $n \rightarrow \infty$ имеем $X_n \xrightarrow{d} X$ и $Y_n \xrightarrow{d} Y$, где каждая из величин X и Y принимает значения 1 или 0 с вероятностью $\frac{1}{2}$. Ясно, что $X_n + Y_n = 1$, а сумма $X + Y$ — это случайная величина с тремя возможными значениями: 0, 1 и 2 с вероятностями соответственно $\frac{1}{4}, \frac{1}{2}$ и $\frac{1}{4}$. Следовательно, сходимость $X_n + Y_n \xrightarrow{d} X + Y$ при $n \rightarrow \infty$ не может иметь места.

(б) Предположим теперь, что последовательности $\{X_n, n \geq 1\}$ и $\{Y_n, n \geq 1\}$ таковы, что $X_n \xrightarrow{d} X$ и $Y_n \xrightarrow{d} Y$, где $X \sim \mathcal{N}(0, 1)$, $Y \sim \mathcal{N}(0, 1)$. Если при каждом n величины X_n и Y_n независимы, то $X_n + Y_n \xrightarrow{d} Z$, где $Z \sim \mathcal{N}(0, 2)$. В этом случае распределение случайного вектора (X_n, Y_n) сходится к двумерному нормальному распределению с нулевым вектором математических ожиданий и матрицей ковариации $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Вопрос: как выглядит этот результат без условия независимости X_n и Y_n ? Пусть снова последовательность $\{X_n, n \geq 1\}$ такова, что $X_n \xrightarrow{d} X$,

где $X \sim \mathcal{N}(0, 1)$. Положим $Y_n = X_n$ для каждого $n \geq 1$. Тогда $Y_n \xrightarrow{d} Y$, где $Y \sim \mathcal{N}(0, 1)$. Поскольку $X_n + Y_n = 2X_n$, то $X_n + Y_n \xrightarrow{d} Z^*$, где $Z^* \sim \mathcal{N}(0, 4)$, т. е. предел Z^* не имеет распределения $\mathcal{N}(0, 2)$, как было выше при выполнении условия независимости.

14.9. Не существует метрики, соответствующей сходимости почти наверное

Начнем со следующего результата (см., например, [129, 210]): в пространстве случайных величин существует метрика относительно каждого из следующих трех видов сходимости: (а) сходимости по распределению; (б) сходимости по вероятности; (в) сходимости в среднем порядка r , $0 \leq r < \infty$. Поскольку в этом результате ничего не говорится о сходимости п. н., то естественно поставить вопрос: существует ли метрика в пространстве случайных величин, соответствующая сходимости п. н.?

Итак, пусть $\mathcal{R}$ — совокупность случайных величин, заданных на вероятностном пространстве $(\Omega, \mathcal{F}, P)$, и пусть $d: \mathcal{R} \times \mathcal{R} \mapsto \mathbb{R}$ — метрика в $\mathcal{R}$, т. е. функция $d(\cdot, \cdot)$, которая неотрицательна, симметрична и удовлетворяет неравенству треугольника. Предположим, что в $\mathcal{R}$ существует метрика d , соответствующая сходимости п. н., т. е. что для $X, X_1, X_2, \dots \in \mathcal{R}$ при $n \rightarrow \infty$

$$d(X_n, X) \rightarrow 0 \iff X_n \xrightarrow{\text{п.н.}} X.$$

Рассмотрим последовательность случайных величин $\{X_n, n \geq 1\}$, которая сходится к случайной величине X по вероятности, но не п. н. (см. пример 14.3). Тогда для некоторого $\delta > 0$ неравенство $d(X_n, X) \geq \delta$ будет выполнено для бесконечно многих индексов n . Через Λ обозначим множество всех таких n . Однако хорошо известно (см. [9, 22, 116, 127, 145]), что если $X_n \xrightarrow{P} X$ при $n \rightarrow \infty$, то существует подпоследовательность $\{X_{n_k}, n_k \in \Lambda\}$ последовательности $\{X_n, n \in \Lambda\}$ такая, что $X_{n_k} \xrightarrow{\text{п.н.}} X$ при $n_k \rightarrow \infty$. Это, однако, означает, что $d(X_{n_k}, X) \rightarrow 0$ при $n_k \rightarrow \infty$, которое, очевидно, невозможно, потому что $d(X_{n_k}, X) \geq \delta$ для каждого $n_k \in \Lambda$.

Полученное противоречие является следствием сделанного нами предположения. Так мы заключаем, что в общем случае в пространстве случайных величин не существует метрики, соответствующей сходимости почти наверное. Отметим, однако, что этот вид сходимости метризуем тогда и только тогда, когда вероятностное пространство атомарно, т. е. когда $\mathcal{R}$ состоит только из дискретных случайных величин (см., например, [210, 463]).

14.10. О сходимости вполне и сходимости почти наверное последовательностей случайных величин

Последовательность случайных величин $\{X_n, n \geq 1\}$ *сходится вполне* к 0 при $n \rightarrow \infty$, если для каждого $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} \sum_{k=n}^{\infty} P\{|X_k| > \varepsilon\} = 0. \quad (1)$$

В этом случае используется обозначение $X_n \xrightarrow{c} 0$.

Наша цель теперь — сравнить этот вид сходимости со сходимостью п. н. Заметим, что

$$X_n \xrightarrow{\text{п.н.}} 0 \text{ при } n \rightarrow \infty \iff \lim_{n \rightarrow \infty} P\left\{\bigcup_{k=n}^{\infty} [|X_k| > \varepsilon]\right\} = 0. \quad (2)$$

Поскольку P — полуаддитивная функция множеств, получаем

$$P\left\{\bigcup_{k=n}^{\infty} [|X_k| > \varepsilon]\right\} \leq \sum_{k=n}^{\infty} P\{|X_k| > \varepsilon\}.$$

Отсюда с учетом (1) и (2) заключаем, что $X_n \xrightarrow{c} 0 \Rightarrow X_n \xrightarrow{\text{п.н.}} 0$. Оказывается, однако, что обратное утверждение неверно. Действительно, рассмотрим вероятностное пространство $(\Omega, \mathcal{F}, P)$, где $\Omega = [0, 1]$, $\mathcal{F} = \mathcal{B}_{[0,1]}$ и P — мера Лебега. Последовательность $\{X_n, n \geq 1\}$ зададим следующим образом:

$$X_n = X_n(\omega) = \begin{cases} 1, & \text{если } 0 \leq \omega < \frac{1}{n}, \\ 0, & \text{если } \frac{1}{n} \leq \omega \leq 1. \end{cases}$$

Тогда легко видеть, что $X_n \xrightarrow{\text{п.н.}} 0$ при $n \rightarrow \infty$. Однако соотношение (1) не выполнено, и, следовательно, эта последовательность не является вполне сходящейся к 0 при $n \rightarrow \infty$.

Отметим, наконец, что сходимость вполне и сходимость п. н. эквивалентны, если последовательность $\{X_n, n \geq 1\}$ состоит из независимых случайных величин (см. [287]).

14.11. О сходимости средних по Чезаро для последовательности случайных величин

Пусть $\{X_n, n \geq 1\}$ — последовательность случайных величин. Тогда

$$X_n \xrightarrow{\text{п.н.}} 0 \text{ при } n \rightarrow \infty \implies \frac{1}{n}(X_1 + \dots + X_n) \xrightarrow{\text{п.н.}} 0 \text{ при } n \rightarrow \infty. \quad (1)$$

Этот результат нетрудно вывести, если использовать некоторые стандартные теоремы из анализа, касающиеся *средних по Чезаро*.

Покажем, что сходимость п. н. в (1) нельзя, вообще говоря, заменить сходимостью по вероятности. Для этой цели рассмотрим последовательность $\{\xi_n, n \geq 1\}$ из независимых случайных величин таких, что ξ_n имеет следующую функцию распределения:

$$F_n(x) = \begin{cases} 0, & \text{если } x \leq 0, \\ 1 - \frac{1}{x+n}, & \text{если } x > 0. \end{cases}$$

Для каждого фиксированного $\varepsilon > 0$ при $n \rightarrow \infty$

$$P\{|\xi_n| > \varepsilon\} = 1 - F_n(\varepsilon) = \frac{1}{\varepsilon + n} \rightarrow 0, \quad \text{т. е.} \quad \xi_n \xrightarrow{P} 0.$$

Покажем, что для средних по Чезаро выполнено соотношение

$$\eta_n = \frac{1}{n} (\xi_1 + \dots + \xi_n) \not\xrightarrow{P} 0 \quad \text{при} \quad n \rightarrow \infty.$$

Обозначая $M_n = \max(\xi_1, \dots, \xi_n)$ и учитывая свойство независимости величин ξ_j , легко получаем, что для каждого $x > 0$

$$P\{M_n \leq x\} = \left(1 - \frac{1}{x+1}\right) \left(1 - \frac{1}{x+2}\right) \dots \left(1 - \frac{1}{x+n}\right) < \left(1 - \frac{1}{x+n}\right)^n.$$

Следовательно,

$$P\left\{\frac{M_n}{n} \leq \varepsilon\right\} < \left(1 - \frac{1}{\varepsilon n + n}\right)^n. \quad (2)$$

Поскольку имеет место включение $\left\{\frac{M_n}{n} > \varepsilon\right\} \subset \{\eta_n > \varepsilon\}$, то

$$P\left\{\frac{M_n}{n} > \varepsilon\right\} \leq P\{\eta_n > \varepsilon\} \quad \text{и} \quad P\left\{\frac{M_n}{n} \leq \varepsilon\right\} \leq P\{\eta_n \leq \varepsilon\}.$$

Из последнего соотношения и неравенства (2) следует, что

$$P\{\eta_n \leq \varepsilon\} < \left(1 - \frac{1}{(\varepsilon + 1)n}\right)^n.$$

Отсюда вытекает, что

$$\lim_{n \rightarrow \infty} P\{\eta_n \leq \varepsilon\} \leq \exp\left(-\frac{1}{\varepsilon + 1}\right),$$

а это эквивалентно соотношению

$$\lim_{n \rightarrow \infty} P\{\eta_n > \varepsilon\} \geq 1 - \exp\left(-\frac{1}{\varepsilon + 1}\right) > 0.$$

Это означает, что случайные величины η_n не сходятся к 0 по вероятности при $n \rightarrow \infty$.

Итак, для последовательности $\{\xi_n, n \geq 1\}$ мы установили, что

$$\xi_n \xrightarrow{P} 0, \quad \text{но} \quad \frac{1}{n} (\xi_1 + \dots + \xi_n) \not\xrightarrow{P} 0 \quad \text{при} \quad n \rightarrow \infty,$$

т. е. утверждение (1) о сходимости средних по Чезаро не распространяется на сходимость по вероятности.

§ 15. Законы больших чисел

Пусть $\{X_n, n \geq 1\}$ — последовательность случайных величин, заданных на вероятностном пространстве $(\Omega, \mathcal{F}, P)$. Обозначим

$$S_n = X_1 + \dots + X_n, \quad a_k = EX_k, \quad A_n = ES_n = a_1 + \dots + a_n.$$

Мы говорим, что последовательность $\{X_n, n \geq 1\}$ удовлетворяет *закону больших чисел* (ЗБЧ), если $\frac{1}{n} S_n - \frac{1}{n} A_n \xrightarrow{P} 0$ при $n \rightarrow \infty$, т. е. если для любого $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} P\left\{\left|\frac{1}{n} S_n - \frac{1}{n} A_n\right| \geq 0\right\} = 0.$$

Далее, если $\frac{1}{n} S_n - \frac{1}{n} A_n \xrightarrow{\text{п.н.}} 0$ при $n \rightarrow \infty$, т. е. если

$$P\left\{\omega: \lim_{n \rightarrow \infty} \left(\frac{1}{n} S_n(\omega) - \frac{1}{n} A_n\right) = 0\right\} = 1,$$

мы говорим, что $\{X_n, n \geq 1\}$ удовлетворяет *усиленному закону больших чисел* (УЗБЧ).

Итак, ЗБЧ, называемый часто обычным или слабым законом, связан со сходимостью по вероятности, а УЗБЧ основывается на сходимости с вероятностью 1. Сформулируем основные результаты, посвященные этим двум законам. Отдельно рассматриваются случаи, когда величины $X_n, n \geq 1$, распределены одинаково и когда эти величины имеют произвольные (неодинаковые) распределения.

Теорема Хинчина. Пусть последовательность $\{X_n, n \geq 1\}$ состоит из независимых и одинаково распределенных случайных величин с $E\{|X_1|\} < \infty$. Тогда $\{X_n, n \geq 1\}$ удовлетворяет ЗБЧ и $\frac{1}{n} S_n \xrightarrow{P} a$ при $n \rightarrow \infty$, где $a = EX_1$.

Теорема Колмогорова. Пусть $\{X_n, n \geq 1\}$ — произвольная последовательность независимых и одинаково распределенных случайных величин. Существование $E\{|X_1|\}$ является необходимым

и достаточным условием для того, чтобы $\{X_n, n \geq 1\}$ удовлетворяла УЗБЧ. В этом случае $\frac{1}{n} S_n \xrightarrow{n.н.} a$ при $n \rightarrow \infty$, где $a = EX_1$.

Теорема Маркова. Предположим, что $\{X_n, n \geq 1\}$ — произвольная последовательность случайных величин такая, что выполнено следующее условие:

$$\frac{1}{n^2} D\{X_1 + \dots + X_n\} \rightarrow 0 \quad \text{при } n \rightarrow \infty \quad (\text{условие Маркова}).$$

Тогда последовательность $\{X_n, n \geq 1\}$ удовлетворяет ЗБЧ.

Теорема Колмогорова. Пусть $\{X_n, n \geq 1\}$ — последовательность независимых случайных величин с конечными дисперсиями, т. е. $\sigma_n^2 = DX_n < \infty, n \geq 1$. Предположим, что выполнено следующее условие:

$$\sum_{n=1}^{\infty} \frac{\sigma_n^2}{n^2} < \infty \quad (\text{условие Колмогорова}).$$

Тогда для последовательности $\{X_n, n \geq 1\}$ имеет место УЗБЧ.

Отметим, что законы больших чисел занимают центральное место во всей теории вероятностей и во многих важных ее приложениях. Доказательства сформулированных выше результатов, как и изложение многих близких вопросов, могут быть найдены, например, в следующих книгах: [9, 13, 26, 32, 60, 75, 87, 92, 106, 116, 117, 145, 177, 179, 231, 330, 345, 407, 405].

Рассмотренные ниже примеры раскрывают роль различных условий для того, чтобы имели место ЗБЧ или УЗБЧ, а также связь между этими двумя законами.

15.1. Условие Маркова достаточно, но не необходимо для справедливости закона больших чисел

Пусть последовательность $\{X_n, n \geq 1\}$ состоит из независимых случайных величин, причем X_n имеет χ_n^2 -распределение с n степенями свободы, т. е. X_n имеет плотность

$$f_n(x) = \begin{cases} \frac{1}{2^{n/2} \Gamma(n/2)} x^{n/2-1} e^{-x/2}, & \text{если } x > 0, \\ 0, & \text{если } x \leq 0. \end{cases}$$

Тогда $EX_n = n, DX_n = 2n$, и легко видеть, что условие Маркова не выполнено. Поскольку это условие является достаточным, то мы не можем применить теорему Маркова, чтобы ответить на вопрос о применимости ЗБЧ к последовательности $\{X_n, n \geq 1\}$. Оказывается, здесь

очень удобно использовать следующий легко устанавливаемый результат (см. например, [106, 116, 145]): если $\{\xi_n, n \geq 1\}$ — последовательность случайных величин и ξ_n имеет характеристическую функцию φ_n , то

$$\xi_n \xrightarrow{P} 0 \text{ при } n \rightarrow \infty \iff \varphi_n(t) \rightarrow 1 \text{ при } n \rightarrow \infty \text{ для всех } t \in \mathbb{R}^1.$$

Заметим теперь, что X_n имеет характеристическую функцию $\psi_n(t) = (1 - 2it)^{-n/2}$, $t \in \mathbb{R}^1$. Тогда, используя свойство независимости величин X_k , нетрудно найти в явном виде характеристическую функцию $\tilde{\psi}_n(t) = E\{e^{it\eta_n}\}$ величины $\eta_n = \frac{S_n - ES_n}{n}$, где $S_n = X_1 + \dots + X_n$, и показать, что $\tilde{\psi}_n(t) \rightarrow 1$ при $n \rightarrow \infty$ для всех $t \in \mathbb{R}^1$. Так заключаем, что $\{X_n, n \geq 1\}$ удовлетворяет ЗБЧ.

Следовательно, условие Маркова не является необходимым для справедливости ЗБЧ (см. также пример 15.6 (б)).

15.2. Условие Колмогорова достаточно, но не необходимо для справедливости усиленного закона больших чисел для неодинаково распределенных случайных величин

Рассмотрим последовательность независимых случайных величин $\{X_n, n \geq 1\}$, где

$$\begin{aligned} P\{X_n = 1\} &= P\{X_n = -1\} = \frac{1}{2}(1 - 2^{-n}), \\ P\{X_n = 2^n\} &= P\{X_n = -2^n\} = 2^{-(n+1)}. \end{aligned}$$

Очевидно, что $EX_n = 0$, $\sigma_n^2 = DX_n = 1 - 2^{-n} + 2^n$, так что ряд $\sum_{n=1}^{\infty} \frac{\sigma_n^2}{n^2}$ расходится. Значит, условие Колмогорова для УЗБЧ (т. е. $\sum_{n=1}^{\infty} \frac{\sigma_n^2}{n^2} < \infty$) не выполнено, и мы не можем сослаться на теорему Колмогорова. А наша цель — проверить, удовлетворяет ли последовательность $\{X_n, n \geq 1\}$ УЗБЧ. Для этого нам потребуется новое понятие. Именно, две последовательности случайных величин $\{\xi_n, n \geq 1\}$ и $\{\eta_n, n \geq 1\}$ называются эквивалентными в смысле Хинчина, если $\sum_{n=1}^{\infty} P\{\xi_n \neq \eta_n\} < \infty$.

Согласно теореме Ревеса (см. [407]), две такие последовательности одновременно удовлетворяют или не удовлетворяют УЗБЧ. Наряду с последовательностью $\{X_n, n \geq 1\}$ введем еще одну, $\{Y_n, n \geq 1\}$, где

$$P\{Y_n = 1\} = P\{Y_n = -1\} = \frac{1}{2}(1 - 2^{-n}), \quad P\{Y_n = 0\} = 2^{-n}.$$

Ясно, что $EX_n = EY_n$ и $P\{X_n \neq Y_n\} = 2^{-n}$, $n \in \mathbb{N}$. Поскольку ряд $\sum_{n=1}^{\infty} P\{X_n \neq Y_n\} = \sum_{n=1}^{\infty} 2^{-n}$ сходится, то последовательности $\{X_n, n \geq 1\}$ и $\{Y_n, n \geq 1\}$ являются эквивалентными в смысле Хинчина. Далее, $DY_n = 1 - 2^{-n}$, так что $\sum_{n=1}^{\infty} \frac{DY_n}{n^2} < \infty$, т. е. для $\{Y_n, n \geq 1\}$ выполнено условие Колмогорова, и, следовательно, $\{Y_n, n \geq 1\}$ удовлетворяет УЗБЧ. Согласно цитированной выше теореме Ревеса, последовательность $\{X_n, n \geq 1\}$ также удовлетворяет УЗБЧ. Следовательно, условие Колмогорова не является необходимым для справедливости УЗБЧ.

15.3. Еще один пример, когда условие Колмогорова не выполнено, но усиленный закон больших чисел имеет место

Рассмотрим последовательность $\{X_n, n \geq 2\}$ из независимых случайных величин, где

$$P\left\{X_n = \frac{n}{\sqrt{\ln n}}\right\} = P\left\{X_n = -\frac{n}{\sqrt{\ln n}}\right\} = \frac{1}{2}.$$

Тогда $\sigma_n^2 = DX_n = \frac{n^2}{\ln n}$ и условие Колмогорова $\sum_{n=2}^{\infty} \frac{\sigma_n^2}{n^2} < \infty$ не выполнено. Мы хотим выяснить, удовлетворяет ли последовательность $\{X_n, n \geq 2\}$ УЗБЧ. Воспользуемся следующей теоремой Ревеса (см. [407]): пусть $\{\xi_n, n \geq 1\}$ — независимые случайные величины с $E\xi_n = 0$, $n \geq 1$, и пусть для некоторого $r \geq 1$

$$E\{|\xi_n|^{2r}\} < \infty \quad \text{и} \quad \sum_{n=1}^{\infty} \frac{E\{|\xi_n|^{2r}\}}{n^{r+1}} < \infty.$$

Тогда последовательность $\{\xi_n, n \geq 1\}$ удовлетворяет УЗБЧ. Теперь легко заметить, что условия теоремы Ревеса выполнены для последовательности $\{X_n, n \geq 2\}$ при $r = 2$. Так мы приходим к заключению, что $\{X_n, n \geq 2\}$ удовлетворяет УЗБЧ.

15.4. О роли условия Колмогорова $\sum_{n=1}^{\infty} (\sigma_n^2/n^2) < \infty$

В примерах 15.2 и 15.3 мы видели, что УЗБЧ может иметь место и для последовательностей случайных величин, для которых не выполнено условие Колмогорова $\sum_{n=1}^{\infty} \frac{\sigma_n^2}{n^2} < \infty$. Полезно, однако, описать и последовательности, для которых и условие Колмогорова не выполнено, и УЗБЧ не справедлив.

Пусть $\{\sigma_n^2, n \geq 1\}$ — такая последовательность положительных чисел, что

$$\sum_{n=1}^{\infty} \frac{\sigma_n^2}{n^2} = \infty. \quad (1)$$

Построим последовательность независимых случайных величин $\{Y_n, n \geq 1\}$ с $DY_n = \sigma_n^2, n \geq 1$, для которой УЗБЧ не имеет места.

Итак, если $\frac{\sigma_n^2}{n^2} \leq 1$, рассмотрим случайную величину Y_n с тремя возможными значениями: $-n, 0$ и n с вероятностями соответственно $\frac{\sigma_n^2}{2n^2}, 1 - \frac{\sigma_n^2}{n^2}$ и $\frac{\sigma_n^2}{2n^2}$. Если же $\frac{\sigma_n^2}{n^2} > 1$, то Y_n принимает каждое из значений σ_n и $-\sigma_n$ с вероятностью $\frac{1}{2}$. Легко видеть, что $EY_n = 0, DY_n = \sigma_n^2$. Далее, для каждого $\varepsilon > 0$

$$P\left\{\frac{|Y_n|}{n} > \varepsilon\right\} = P\{Y_n \neq 0\} = \begin{cases} \frac{\sigma_n^2}{n^2}, & \text{если } \frac{\sigma_n^2}{n^2} \leq 1, \\ 1, & \text{если } \frac{\sigma_n^2}{n^2} > 1. \end{cases} \quad (2)$$

Предположим, что $\{Y_n, n \geq 1\}$ удовлетворяет УЗБЧ, т. е.

$$\frac{1}{n} \sum_{k=1}^n Y_k \xrightarrow{\text{п.н.}} 0 \quad \text{при } n \rightarrow \infty.$$

Отсюда и из равенства

$$\frac{1}{n} Y_n = \frac{1}{n} \sum_{k=1}^n Y_k - \frac{n-1}{n} \frac{1}{n-1} \sum_{k=1}^{n-1} Y_k$$

вытекало бы, что $\frac{1}{n} Y_n \xrightarrow{\text{п.н.}} 0$ при $n \rightarrow \infty$. Из (1) и (2) следует, что $\sum_{n=1}^{\infty} P\{|Y_n| > \varepsilon n\} = \infty$, а поскольку величины $Y_n, n \geq 1$, независимы, то лемма Бореля–Кантелли гарантирует, что событие $\{|Y_n| \geq \varepsilon n\}$ происходит бесконечно часто, т. е.

$$P\{|Y_n| \geq \varepsilon n \text{ б. ч.}\} = 1.$$

Это, однако, не согласуется со сходимостью $\frac{Y_n}{n} \xrightarrow{\text{п.н.}} 0$ при $n \rightarrow \infty$. Противоречие говорит об ошибочности сделанного нами предположения. Итак, последовательность $\{Y_n, n \geq 1\}$ не удовлетворяет УЗБЧ.

Описанный пример и другие подобные примеры показывают, что условие Колмогорова $\sum_{n=1}^{\infty} \frac{\sigma_n^2}{n^2} < \infty$ — наилучшее достаточное условие для применимости УЗБЧ к последовательностям независимых и неодинаково распределенных случайных величин.

**15.5. Арифметические средние для случайной последовательности
могут сходиться по вероятности,
даже если теорема Хинчина неприменима**

Пусть последовательность $\{X_n, n \geq 1\}$ состоит из независимых и одинаково распределенных случайных величин таких, что $E\{|X_1|\} = \infty$. Наша цель — изучить поведение средних арифметических $Y_n = \frac{S_n}{n}$, где $S_n = X_1 + \dots + X_n$ при $n \rightarrow \infty$. Согласно теореме Колмогорова (содержащей необходимое и достаточное условие), $\{X_n, n \geq 1\}$ не удовлетворяет УЗБЧ. Это означает, что при $n \rightarrow \infty$ последовательность $\{Y_n, n \geq 1\}$ не может быть сходящейся с вероятностью 1. Уместно поставить вопрос о сходимости Y_n по вероятности при $n \rightarrow \infty$. Понятно, что если бы $\{X_n, n \geq 1\}$ удовлетворяла обычному ЗБЧ, то $Y_n = \frac{X_1 + \dots + X_n}{n} \xrightarrow{P} a$ при $n \rightarrow \infty$, где $a = EX_1$. Но $E\{|X_1|\} = \infty$, (математическое ожидание EX_1 не существует) и мы не можем применить теорему Хинчина (см. начало настоящего параграфа). Однако в этой теореме условие $E\{|X_1|\} < \infty$ является только достаточным для справедливости (обычного) ЗБЧ. Таким образом, имеет смысл следующий вопрос: существует ли конечная константа c такая, что $Y_n = \frac{X_1 + \dots + X_n}{n} \xrightarrow{P} c$ при $n \rightarrow \infty$? Покажем, что это возможно. Для этой цели рассмотрим последовательность независимых, одинаково распределенных случайных величин $\{\xi_n, n \geq 1\}$, где

$$P\{\xi_1 = (-1)^{k-1}k\} = \frac{6}{\pi^2 k^2}, \quad k = 1, 2, \dots$$

Легко видеть, что $E\{|\xi_1|\} = \infty$. Существенным для дальнейшего рассуждения является следующий результат (см., например, [345, 405, 447]): *пусть независимые и одинаково распределенные случайные величины $\xi_n, n \geq 1$, имеют характеристическую функцию $\psi(t)$, $t \in \mathbb{R}^1$, которая дифференцируема в точке $t = 0$, причем $\psi'(0) = ic$, где $i = \sqrt{-1}$, $c \in \mathbb{R}^1$. Тогда $\frac{\xi_1 + \dots + \xi_n}{n} \xrightarrow{P} c$ при $n \rightarrow \infty$.* Теперь понятно, что нам нужно найти характеристическую функцию $\psi(t) = E\{e^{it\xi_1}\}$ и установить, что она обладает нужными нам свойствами. Поскольку

$$\psi(t) = \frac{6}{\pi^2} \sum_{j=1}^{\infty} \frac{(e^{it})^{2j-1}}{(2j-1)^2} + \frac{6}{\pi^2} \sum_{k=1}^{\infty} \frac{(e^{-it})^{2k}}{(2k)^2},$$

то имеет смысл ввести следующие две функции:

$$h_1(u) = \sum_{j=1}^{\infty} \frac{u^{2j-1}}{(2j-1)^2}, \quad |u| \leq 1, \quad \text{и} \quad h_2(u) = \sum_{k=1}^{\infty} \frac{v^{2k}}{(2k)^2}, \quad |v| \leq 1.$$

Можно показать, что эти функции дифференцируемы, причем

$$h'_1(u) = \frac{1}{2u} \ln \frac{1+u}{1-u}, \quad h'_2(v) = -\frac{1}{2v} \ln(1-v^2).$$

Тогда $h'_1(u) - h'_2(u) = \frac{1}{u} \ln(1+u)$. Следовательно,

$$\psi'(0) = i \frac{6}{\pi^2} [h'_1(1) - h'_2(1)] = i \frac{6 \ln 2}{\pi^2}.$$

Согласно цитированному выше результату, заключаем, что

$$\frac{1}{n} (\xi_1 + \dots + \xi_n) \xrightarrow{P} \frac{6 \ln 2}{\pi^2} \quad \text{при } n \rightarrow \infty,$$

и это как раз то, что мы хотели показать.

Отметим, наконец, что последовательность $\{\xi_n, n \geq 1\}$ удовлетворяет так называемому обобщенному ЗБЧ (см. пример 15.8).

15.6. Последовательность случайных величин, удовлетворяющая слабому, но не усиленному закону больших чисел

Поскольку из сходимости с вероятностью 1 следует сходимость по вероятности, то каждая последовательность случайных величин, удовлетворяющая УЗБЧ, удовлетворяет также и (слабому) ЗБЧ. Поэтому представляет интерес описать конкретные последовательности случайных величин, для которых имеет место ЗБЧ, но не имеет места УЗБЧ.

(а) Пусть $\{X_n, n \geq 2\}$ — независимые случайные величины такие, что

$$\begin{aligned} P\{X_n = n\} &= P\{X_n = -n\} = \frac{1}{2n \ln n}, \\ P\{X_n = 0\} &= 1 - \frac{1}{n \ln n}, \quad n = 2, 3, \dots \end{aligned}$$

Введем события $A_n = \{|X_n| \geq n\}$, $n \geq 2$. Тогда

$$P(A_n) = \frac{1}{n \ln n} \quad \text{и} \quad \sum_{n=2}^{\infty} P(A_n) = \infty.$$

Расходимость ряда $\sum_{n=2}^{\infty} P(A_n)$ и независимость величин X_n позволяют заключить на основании леммы Бореля–Кантелли, что событие $\{A_n \text{ б. ч.}\}$ имеет вероятность 1. Таким образом,

$$P\{|X_n| \geq n \text{ б. ч.}\} = 1 \implies P\left\{\lim_{n \rightarrow \infty} \frac{S_n}{n} \neq 0\right\} = 1,$$

и, значит, $\{X_n, n \geq 2\}$ не может удовлетворять УЗБЧ.

Покажем, что для $\{X_n, n \geq 2\}$ имеет место обычный ЗБЧ. Очевидно, что $DX_k = \frac{k}{\ln k}$. Поскольку функция $\frac{x}{\ln x}$ имеет локальный минимум в точке $x = e$, а ряд $\sum_{k=3}^{\infty} \frac{k}{\ln k}$ является нижней суммой Римана для интеграла $\int_3^{n+1} \frac{x}{\ln x} dx$, легко видеть, что

$$\begin{aligned} \frac{1}{n^2} \sum_{k=2}^n DX_k &\leq \frac{1}{n^2} \left(\frac{2}{\ln 2} + \int_3^{n+1} \frac{x}{\ln x} dx \right) \leq \\ &\leq \frac{2}{n^2 \ln 2} + \frac{(n-2)(n+1)}{n^2 \ln n} \rightarrow 0 \quad \text{при } n \rightarrow \infty. \end{aligned}$$

Итак, выполнено условие Маркова, и, значит, последовательность $\{X_n, n \geq 2\}$ удовлетворяет ЗБЧ.

(б) В случае (а) величины X_n были дискретными. Опишем теперь пример последовательности $\{Y_n, n \geq 2\}$, состоящей из независимых случайных величин, которые абсолютно непрерывны, причем ЗБЧ будет иметь место, а УЗБЧ — нет.

Пусть $Y_n, n \geq 2$, — независимые случайные величины такие, что Y_n имеет следующую плотность (распределение Лапласа):

$$f_n(x) = \frac{1}{\sqrt{2}\sigma_n} \exp\left(-\frac{\sqrt{2}|x|}{\sigma_n}\right), \quad x \in \mathbb{R}^1, \quad \sigma_n > 0.$$

Очевидно, что $EY_n = 0$, и нетрудно показать, что $DY_n = \sigma_n^2$. Теперь выберем σ_n^2 специальным образом: $\sigma_n^2 = \frac{2n^2}{(\ln n)^2}$, $n \geq 2$. Тогда событие $A_n = \{|Y_n| \geq n\}$ имеет вероятность

$$P(A_n) = \frac{\sqrt{2}}{\sigma_n} \int_n^{\infty} \exp\left(-\frac{\sqrt{2}x}{\sigma_n}\right) dx = \exp\left(-\frac{1}{2} \frac{\sqrt{2}(\ln n)^2}{n}\right).$$

Так как $\frac{(\ln n)^2}{n} \rightarrow 0$ при $n \rightarrow \infty$, то $\sum_{n=2}^{\infty} P(A_n) = \infty$. Отсюда, как и в случае (а), заключаем, что $\{Y_n, n \geq 2\}$ не удовлетворяет УЗБЧ.

Можно проверить, что для последовательности $\{Y_n, n \geq 2\}$ условие Маркова не выполнено. Однако можно применить метод урезания и воспользоваться одной теоремой Феллера (см. [106]), чтобы показать справедливость ЗБЧ для $\{Y_n, n \geq 2\}$ (подробности можно найти в [249] или [449]).

Отметим, кстати, что случай (б) — это еще один пример, когда ЗБЧ имеет место, а условие Маркова не выполнено. В отличие от примера 15.1, здесь величины абсолютно непрерывны.

15.7. Закон больших чисел может не выполняться, если сходимость с вероятностью 1 заменить сходимостью вполне

Предположим, что УЗБЧ можно применить для последовательности $\{\xi_n, n \geq 1\}$, состоящей из независимых и одинаково распределенных случайных величин с функцией распределения $F(x)$, $x \in \mathbb{R}^1$, и математическим ожиданием $E\xi_1 = \int_{-\infty}^{\infty} x dF(x) = 0$. Это означает, что

$$\eta_n = \frac{1}{n} (\xi_1 + \dots + \xi_n) \xrightarrow{\text{п.н.}} 0 \quad \text{при } n \rightarrow \infty. \quad (1)$$

Вопрос: можно ли в (1) сходимость п. н. заменить более сильной сходимостью, именно сходимостью вполне (см., пример 14.10)? Покажем, что это не всегда так.

Пусть функция распределения F такова, что выполнены условия

$$\int_{-\infty}^{\infty} x dF(x) = 0, \quad \sigma^2 = \int_{-\infty}^{\infty} x^2 dF(x) < \infty. \quad (2)$$

Можно показать, что для каждого $\varepsilon > 0$ ряд $\sum_{n=1}^{\infty} P\{|\eta_n| > \varepsilon\}$ сходится (подробности см. в [287]). Следовательно, условия (2) обеспечивают сходимость вполне для последовательности $\{\eta_n, n \geq 1\}$. Таким образом, $\eta_n \xrightarrow{c} 0$ при $n \rightarrow \infty$ вместо сходимости с вероятностью 1 в соотношении (1).

Заменим теперь условия (2) следующими:

$$\int_{-\infty}^{\infty} x dF(x) = 0, \quad \int_{-\infty}^{\infty} |x|^{\gamma} dF(x) < \infty, \quad \int_{-\infty}^{\infty} x^2 dF(x) = \infty, \quad (3)$$

где $\gamma = \text{const}$, $\gamma \in \left(\frac{1+\sqrt{5}}{2}, 2\right)$. Тогда $\{\xi_n, n \geq 1\}$ удовлетворяет УЗБЧ. Однако ряд $\sum_{n=1}^{\infty} P\{|\eta_n| > \varepsilon\}$ расходится, и, значит, сходимость $\eta_n \xrightarrow{c} 0$ при $n \rightarrow \infty$ не может иметь места.

Итак, существуют последовательности независимых и одинаково распределенных случайных величин $\{\xi_n, n \geq 1\}$ такие, что их арифметические средние η_n сходятся п. н., но не сходятся вполне.

Понятно, что нужно указать хотя бы один конкретный пример, когда выполнены условия (3). Для этого достаточно рассмотреть абсолютно непрерывную случайную величину ξ_1 с плотностью $f(x) = |x|^{-3}$, если $|x| \geq 1$, и $f(x) = 0$, если $|x| < 1$.

15.8. О справедливости обобщенного закона больших чисел

Пусть $\{X_n, n \geq 1\}$ — последовательность независимых случайных величин, а $S_n = X_1 + \dots + X_n$. Если существуют числовые последовательности $\{a_n, n \geq 1\}$ и $\{b_n, n \geq 1\}$, $b_n > 0$, такие что

$$\frac{S_n - a_n}{b_n} \rightarrow 0 \quad \text{при} \quad n \rightarrow \infty, \quad (1)$$

то говорят, что $\{X_n, n \geq 1\}$ удовлетворяет *обобщенному закону больших чисел*. Он называется слабым, если (1) имеет место относительно сходимости по вероятности, и усиленным — при сходимости с вероятностью 1. Очевидно, при $a_n = ES_n$ и $b_n = n$ выполнены условия классического ЗБЧ. Оказывается, есть последовательности случайных величин, для которых классический ЗБЧ не справедлив, но при подходящем выборе $\{a_n\}$ и $\{b_n\}$ обобщенный ЗБЧ имеет место. Следует конкретный пример.

Рассмотрим последовательность независимых случайных величин $\{X_n, n \geq 1\}$, где

$$P\{X_n = 2^n\} = \frac{1}{2^n}, \quad P\{X_n = 0\} = 1 - \frac{1}{2^n}.$$

Нетрудно проверить, что $\{X_n, n \geq 1\}$ не удовлетворяет обычному ЗБЧ, так как при $a_n = ES_n = n$ и $b_n = n$ соотношение (1) со сходимостью по вероятности места не имеет. Остается, однако, возможность, что (1) может выполняться при некотором специальном выборе $\{a_n, n \geq 1\}$ и $\{b_n, n \geq 1\}$. Действительно, пусть $b_n = n \ln n$, $n \geq 2$. Тогда

$$\frac{S_n}{b_n} \xrightarrow{P} 1 \quad \text{при} \quad n \rightarrow \infty. \quad (2)$$

Проще всего (2) доказывается с использованием характеристических функций: достаточно показать, что $E\left\{\exp\left(\frac{itS_n}{b_n}\right)\right\} \rightarrow e^{it}$ при $n \rightarrow \infty$ для всех $t \in \mathbb{R}^1$. Таким образом,

$$\frac{S_n - a_n}{b_n} \xrightarrow{P} 0 \quad \text{при} \quad n \rightarrow \infty, \quad (3)$$

где $a_n = n \ln n$ и $b_n = n \ln n$. Вопрос: можно ли в соотношении (3) при тех же $\{a_n\}$ и $\{b_n\}$ заменить сходимостью по вероятности сходимостью с вероятностью 1? Для ответа заметим, что $P\{X_n > c\} \geq \frac{1}{c}$ для любого $c > 1$ и всех $n \geq 2$. Тогда

$$P\{X_n > cb_n\} \geq \frac{1}{cb_n} = \frac{1}{cn \ln n} \quad \text{и} \quad \sum_{n=2}^{\infty} P\{X_n > cb_n\} = \infty.$$

Отсюда, используя лемму Бореля–Кантелли, заключаем, что с вероятностью 1 событие $\left\{\frac{X_n}{b_n} > c\right\}$ происходит бесконечно часто, т. е. $P\left\{\frac{X_n}{b_n} > c \text{ б. ч.}\right\} = 1$. Очевидно, однако, что

$$P\left\{\frac{X_n}{b_n} > c \text{ б. ч.}\right\} = 1 \implies P\left\{\frac{S_n}{b_n} > c \text{ б. ч.}\right\} = 1.$$

Следовательно, $P\left\{\lim_{n \rightarrow \infty} \frac{S_n}{b_n} = 1\right\} = 0$, т. е. (3) не имеет места относительно сходимости с вероятностью 1. Итак, $\{X_n, n \geq 1\}$ удовлетворяет обобщенному (слабому) ЗБЧ, но не удовлетворяет обобщенному УЗБЧ.

§ 16. Слабая сходимость вероятностных мер и распределений

Одним из видов сходимости, рассмотренных в § 14, была сходимость по распределению для последовательностей случайных величин. Там же говорилось о ее связи со слабой сходимостью соответствующей последовательности функций распределения. Понятие слабой сходимости вводится для вероятностных мер и распределений в довольно общих пространствах. Мы ограничимся рассмотрением распределений на действительной прямой $\mathbb{R}^1$.

Пусть заданы функции распределения $F(x)$, $x \in \mathbb{R}^1$ и $F_n(x)$, $x \in \mathbb{R}^1$, $n \geq 1$. Обозначим через P и P_n вероятностные меры на измеримом пространстве $(\mathbb{R}^1, \mathcal{B}^1)$, порожденные соответственно F и F_n . Напомним, что P и P_n связаны с F и F_n соотношениями

$$F(x) = P((-\infty, x]), \quad F_n(x) = P_n((-\infty, x]), \quad x \in \mathbb{R}^1.$$

Поскольку F непрерывна в точке x тогда и только тогда, когда $P(\{x\}) = 0$, сходимость $F_n \xrightarrow{d} F$ (см. начало § 14) означает, что $P_n((-\infty, x]) \rightarrow P((-\infty, x])$ при $n \rightarrow \infty$ для каждого x , для которого $P(\{x\}) = 0$. Но $A = (-\infty, x]$ — конкретное борелевское множество в $\mathbb{R}^1$, а $\{x\}$ — граница A . Это рассуждение мотивирует дальнейшее более общее рассмотрение.

Для любого борелевского множества A в $\mathbb{R}^1$, т. е. $A \in \mathcal{B}^1$, через ∂A обозначим границу A . На измеримом пространстве $(\mathbb{R}^1, \mathcal{B}^1)$ заданы вероятностные меры P и P_n , $n \geq 1$. Мы говорим, что последовательность $\{P_n, n \geq 1\}$ слабо сходится к P при $n \rightarrow \infty$, и пишем $P_n \xrightarrow{w} P$, если для любого $A \in \mathcal{B}^1$ с $P(\partial A) = 0$ имеем

$$P_n(A) \rightarrow P(A) \quad \text{при} \quad n \rightarrow \infty.$$

Сформулируем один из основных результатов.

Теорема 1. Следующие утверждения эквивалентны:

- (а) $P_n \xrightarrow{w} P$ при $n \rightarrow \infty$.
 (б) $\lim_{n \rightarrow \infty} P_n(A) \leq P(A)$ для любого замкнутого множества $A \in \mathcal{B}^1$.
 (в) $\lim_{n \rightarrow \infty} P_n(A) \geq P(A)$ для любого открытого множества $A \in \mathcal{B}^1$.
 (г) Для любой непрерывной и ограниченной функции $g(x)$, $x \in \mathbb{R}^1$,

$$\int_{\mathbb{R}^1} g(x) P_n(dx) \rightarrow \int_{\mathbb{R}^1} g(x) P(dx) \quad \text{при } n \rightarrow \infty.$$

Этот результат имеет место для вероятностных мер в гораздо более общих пространствах. Для нас важно, что слабая сходимость в $\mathbb{R}^1$ эквивалентна сходимости по распределению, рассмотренной в § 14. Поэтому принято использовать оба обозначения: $F_n \xrightarrow{d} F$ и $F_n \xrightarrow{w} F$.

Сформулируем еще один результат, связывающий слабую сходимость функций распределения с поточечной сходимостью характеристических функций.

Теорема 2. Пусть $F(x)$ и $F_n(x)$, $x \in \mathbb{R}^1$, $n \geq 1$, — функции распределения, а $\varphi(t)$ и $\varphi_n(t)$, $t \in \mathbb{R}^1$, $n \geq 1$, — соответствующие им характеристические функции.

(а) Если $F_n \xrightarrow{w} F$ при $n \rightarrow \infty$, то $\varphi_n(t) \rightarrow \varphi(t)$ при $n \rightarrow \infty$ для каждого $t \in \mathbb{R}^1$.

(б) Если предел $\lim_{n \rightarrow \infty} \varphi_n(t)$ существует для каждого $t \in \mathbb{R}^1$ и этот предел, обозначаемый $\varphi(t)$, $t \in \mathbb{R}^1$, непрерывен в точке $t = 0$, то $F_n \xrightarrow{w} F$ при $n \rightarrow \infty$, где F — функция распределения, соответствующая характеристической функции φ .

Читатель может найти подробные доказательства теорем 1 и 2, а также множество других полезных и важных результатов в книгах [6, 24, 106, 116, 117, 145, 179, 345, 404].

Рассмотренные далее примеры иллюстрируют условия, при которых слабая сходимость имеет или не имеет места, и связь слабой сходимости с другими видами сходимости.

16.1. Определяющие классы и классы, определяющие сходимость

Пусть P и Q — вероятностные меры, заданные на измеримом пространстве $(E, \mathcal{E})$. Предположим, что $\mathcal{K}_0 \subset \mathcal{E}$ — такой класс подмножеств E , что $P(A) = Q(A)$ для всех $A \in \mathcal{K}_0 \implies P(A) = Q(A)$ для всех $A \in \mathcal{E}$. В этом случае $\mathcal{K}_0$ называется *определяющим классом*.

Пусть далее P и P_n , $n \geq 1$, — вероятностные меры $(E, \mathcal{E})$. Мы говорим, что $\mathcal{K}_1 \subset \mathcal{E}$ является *классом, определяющим сходимость*,

если из того, что

$$P_n(A) \rightarrow P(A) \quad \text{для всех } A \in \mathcal{K}_1 \text{ с } P(\partial A) = 0,$$

вытекает, что

$$P_n(A) \rightarrow P(A) \quad \text{для всех } A \in \mathcal{E} \text{ с } P(\partial A) = 0,$$

т. е. что $P_n \xrightarrow{w} P$ при $n \rightarrow \infty$.

Проиллюстрируем связь между введенными понятиями.

(а) Легко заметить, что каждый класс, определяющий сходимост, является также определяющим классом. Покажем, что обратное не верно. Действительно, пусть $E = [0, 1]$, $\mathcal{E} = \mathcal{B}_{[0,1]}$ и $\mathcal{K} \subset \mathcal{B}_{[0,1]}$ — алгебра всех конечных сумм непересекающихся интервалов вида $[a, b)$, $0 < a < b < 1$. Тогда $\mathcal{K}$ является определяющим классом, но не является классом, определяющим сходимост. В последнем можно убедиться, взяв, например, вероятностные меры P и P_n , сосредоточенные соответственно в точках 0 и $1 - \frac{1}{n}$.

(б) Пусть P, Q и $P_n, n \geq 1$, — вероятностные меры на измеримом пространстве $(\Omega, \mathcal{F})$, где $\Omega = \mathbb{R}^1$, $\mathcal{F} = \mathcal{B}^1$, и пусть $\mathcal{K}_0 \subset \mathcal{F}$ — определяющий класс. Предположим, что выполнены следующие два условия:

$$P_n(A) \rightarrow Q(A) \quad \text{при } n \rightarrow \infty \text{ для всех } A \in \mathcal{K}_0, \quad (1)$$

$$P_n(A) \xrightarrow{w} P \quad \text{при } n \rightarrow \infty. \quad (2)$$

Поскольку класс $\mathcal{K}_0$ определяющий, то, учитывая (1) и (2), можно предположить, что $P = Q$. Оказывается, однако, что это не так. Действительно, определим P, Q и P_n следующим образом:

$$P(\{0\}) = P(\{1\}) = \frac{1}{2}, \quad Q(\{0\}) = 1, \quad P_n\left(\left\{\frac{1}{n}\right\}\right) = P_n\left(\left\{1 + \frac{1}{n}\right\}\right) = \frac{1}{2}.$$

Легко видеть, что $P_n \xrightarrow{w} P$ при $n \rightarrow \infty$. Далее, пусть множество B содержит все точки $0, 1, \frac{1}{n}$ и $1 + \frac{1}{n}$, где $n \geq 1$. Обозначим через $\mathcal{A}$ алгебру, содержащую все множества $A \in \mathcal{F}$ такие, что либо AB имеет конечное число элементов и $0 \notin A$, либо $\bar{A}B$ имеет конечное число элементов и $0 \notin \bar{A}$. Тогда $\mathcal{A}$ является определяющим классом и $P_n(A) \rightarrow Q(A)$ при $n \rightarrow \infty$ для каждого $A \in \mathcal{A}$. Таким образом, условия (1) и (2) выполнены, но очевидно, что $P \neq Q$.

(в) Понятно, что определяющие классы и классы, определяющие сходимост, $\mathcal{K}_0$ и $\mathcal{K}_1$, играют важную роль при изучении вероятностных мер на измеримом пространстве $(E, \mathcal{E})$. Поэтому важно описать эти классы в некоторых часто рассматриваемых случаях. Оказывается, если $E = \mathbb{R}^1$, или $E = \mathbb{R}^n, n \geq 2$, или $E = \mathbb{R}^\infty$ и соответственно $\mathcal{E} = \mathcal{B}^1$,

$\mathcal{E} = \mathcal{B}^n$ и $\mathcal{E} = \mathcal{B}^\infty$, то в качестве $\mathcal{K}_0$ можно взять класс «интервалов», или класс цилиндрических множеств.

Так, например, если $E = \mathbb{R}^1$, то класс $\mathcal{K} = \{(-\infty, x], x \in \mathbb{R}^1\}$ является определяющим, а из упомянутой выше эквивалентности слабой сходимости и сходимости по распределению вытекает, что класс $\mathcal{K}$ определяет сходимость, т. е. в этом случае $\mathcal{K}_0 = \mathcal{K} = \mathcal{K}_1$. Аналогично, если $E = \mathbb{R}^n$, $n \geq 2$ и $\mathcal{K}$ — совокупность всех множеств вида $(-\infty, x] = (-\infty, x_1] \times \dots \times (-\infty, x_n]$, где $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, то $\mathcal{K}$ является как определяющим классом, так и классом, определяющим сходимость, т. е. снова $\mathcal{K}_0 = \mathcal{K} = \mathcal{K}_1$. Наконец, в пространстве $E = \mathbb{R}^\infty$ в качестве $\mathcal{K}$ следует взять совокупность всех цилиндрических множеств, т. е. всех множеств вида $\{x \in \mathbb{R}^\infty : (x_1, \dots, x_n) \in B\}$, $B \in \mathcal{B}^n$, $n \geq 2$. В этом случае класс $\mathcal{K}$ определяющий и, более того, $\mathcal{K}$ определяет сходимость, т. е. $\mathcal{K}_0 = \mathcal{K} = \mathcal{K}_1$.

Итак, цилиндрические множества образуют «хороший» класс в нескольких конкретных измеримых пространствах. А как обстоит дело в более общем случае?

Пусть $E = \mathbb{C}[0, 1]$ — пространство непрерывных функций на отрезке $[0, 1]$ и $\mathcal{E} = \mathcal{C}$ — его борелевская σ -алгебра. Обозначим через $\mathcal{K}$ класс всех цилиндрических множеств в $\mathbb{C}[0, 1]$. Поскольку $\mathcal{K}$ порождает σ -алгебру $\mathcal{C}$, то $\mathcal{K}$ — определяющий класс. Вопрос: является ли $\mathcal{K}$ классом, определяющим сходимость. Покажем, что, вообще говоря, это не так. Действительно, пусть P — вероятностная мера, сосредоточенная на функции $x_t \equiv 0$, $0 \leq t \leq 1$, а P_n — вероятностная мера, сосредоточенная на функции $x^{(n)} = \{x_t^{(n)}, 0 \leq t \leq 1\}$, $n \geq 1$, где

$$x_t^{(n)} = \begin{cases} nt, & \text{если } 0 \leq t \leq \frac{1}{n}, \\ 2 - nt, & \text{если } \frac{1}{n} < t \leq \frac{2}{n}, \\ 0, & \text{если } \frac{2}{n} < t \leq 1. \end{cases}$$

Нетрудно убедиться, что для каждого цилиндрического множества A с $P(\partial A) = 0$ мы имеем $P_n(A) \rightarrow P(A)$ при $n \rightarrow \infty$. Поскольку $x^{(n)}$ не стремятся к 0 относительно равномерной метрики в $\mathbb{C}[0, 1]$, то меры P_n не могут сходиться слабо к мере P при $n \rightarrow \infty$. Например, возьмем множество $A = \left\{x \in \mathbb{C}[0, 1] : |x_t| \leq \frac{1}{2}, 0 \leq t \leq 1\right\}$ — шар в пространстве $\mathbb{C}[0, 1]$ с центром в начале 0 и радиусом $\frac{1}{2}$. Тогда $A \in \mathcal{C}$, $P(\partial A) = 0$, $P_n(A) = 0$, $P(A) = 1$, и, значит, $P_n(A) \not\rightarrow P(A)$ при $n \rightarrow \infty$. Следовательно, определяющий класс $\mathcal{K}$ не определяет сходимость. Другими словами, слабая сходимость в пространстве $\mathbb{C}[0, 1]$ не может быть описана только через

сходимость для всех конечномерных цилиндрических множеств (как это было в $\mathbb{R}^n$ и $\mathbb{R}^\infty$). По поводу слабой сходимости вероятностных мер в пространстве $\mathbb{C}$, а также в других функциональных пространствах читатель может обратиться к книгам [6, 22].

16.2. Из сходимости по распределению не вытекает слабая сходимость соответствующих вероятностных мер для всех борелевских множеств

Пусть $F(x)$, $x \in \mathbb{R}^1$, и $F_n(x)$, $x \in \mathbb{R}^1$, $n \geq 1$, — функции распределения, а P и P_n , $n \geq 1$, — соответствующие им вероятностные меры в измеримом пространстве $(\mathbb{R}^1, \mathcal{B}^1)$. Предположим, что $F_n \xrightarrow{d} F$ при $n \rightarrow \infty$. Ввиду эквивалентности между сходимостью по распределению и слабой сходимостью можно поставить следующий вопрос: верно ли, что $P_n(A) \rightarrow P(A)$ при $n \rightarrow \infty$ для каждого борелевского множества $A \in \mathcal{B}^1$? Конкретными примерами покажем, что это не всегда имеет место и что условие $P(\partial A) = 0$ существенно для слабой сходимости P_n к P .

(а) Проще всего рассмотреть вырожденный случай. Пусть X и X_n — случайные величины такие, что $P\{X = 0\} = 1$, $P\{X_n = \frac{1}{n}\} = 1$, $n \geq 1$. Обозначим через F и F_n их функции распределения, а через P и P_n соответствующие им вероятностные меры. Тогда $F(x) = 1_{[0, \infty)}(x)$, $F_n(x) = 1_{[1/n, \infty)}(x)$ и, очевидно, $F_n \xrightarrow{d} F$ при $n \rightarrow \infty$, т. е. $F_n(x) \rightarrow F(x)$ для всех $x \in \mathbb{R}^1 \setminus \{0\}$ ($x = 0$ — единственная точка скачка предельной функции F). Однако

$$P_n((-\infty, 0]) = F_n(0) = 0 \not\rightarrow P((-\infty, 0]) = F(0) = 1$$

при $n \rightarrow \infty$. Таким образом, соотношение $P_n(B) \rightarrow P(B)$ имеет место не для всех борелевских множеств $B \in \mathcal{B}^1$.

Заметим, что в этом примере предельная функция F разрывна. В последующих примерах предельная функция будет непрерывной, и тем не менее мы придем к подобному же выводу.

(б) Рассмотрим случайные величины $\{X_n, n \geq 1\}$, где

$$P\{X_n = \frac{k}{n}\} = \frac{1}{n}, \quad k = 0, 1, \dots, n-1, \quad n \geq 1.$$

Тогда X_n имеет следующую функцию распределения:

$$F_n(x) = \begin{cases} 0, & \text{если } x \leq 0, \\ \frac{[nx]}{n}, & \text{если } 0 < x \leq 1, \\ 1, & \text{если } x > 1. \end{cases}$$

Возьмем теперь случайную величину X , равномерно распределенную на отрезке $[0, 1]$. Ее функция распределения имеет знакомый вид:

$$F(x) = \begin{cases} 0, & \text{если } x \leq 0, \\ x, & \text{если } 0 < x \leq 1, \\ 1, & \text{если } x > 1. \end{cases}$$

Очень легко проверить, что $\left| \frac{[nx]}{n} - x \right| \leq \frac{1}{n}$ для каждого $x \in \mathbb{R}^1$. Это означает, что $X_n \xrightarrow{d} X$ при $n \rightarrow \infty$, т. е.

$$F_n \xrightarrow{d} F \quad \text{при } n \rightarrow \infty.$$

Пусть теперь Q — множество всех рациональных чисел на отрезке $[0, 1]$. Очевидно, величины X_n , $n \geq 1$, дискретные и все их значения принадлежат Q . Поэтому $P_n(Q) = 1$ для каждого $n \geq 1$. В то же время предельная случайная величина X имеет абсолютно непрерывное распределение (на $[0, 1]$), так что $P(Q) = 0$. Следовательно, для борелевского множества Q имеем

$$\lim_{n \rightarrow \infty} P_n(Q) = 1 \neq 0 = P(Q).$$

Заметим, что здесь $P(\partial Q) \neq 0$.

(в) Рассмотрим последовательность $\{\xi_n, n \geq 1\}$ независимых бернуллиевских случайных величин:

$$P\{\xi_n = 1\} = p, \quad P\{\xi_n = 0\} = 1 - p,$$

где $0 < p < 1$ — фиксированное число. Обозначим

$$S_n = \xi_1 + \dots + \xi_n \quad \text{и} \quad \eta_n = \frac{S_n - np}{\sqrt{np(1-p)}}.$$

Согласно классической теореме Муавра–Лапласа (это частный случай центральной предельной теоремы; см. § 17), $\eta_n \xrightarrow{d} \theta$ при $n \rightarrow \infty$, где случайная величина $\theta \sim \mathcal{N}(0, 1)$. Здесь предельная величина θ имеет абсолютно непрерывное распределение на всей числовой прямой $\mathbb{R}^1$. Как и прежде, обозначим через P_n и P вероятностные меры на $(\mathbb{R}^1, \mathcal{B}^1)$, соответствующие величинам η_n и θ . Рассмотрим множество

$$B = \left\{ \frac{k - np}{\sqrt{np(1-p)}} : k = 0, 1, \dots, n, n \geq 1 \right\}.$$

Очевидно, что B состоит из всех возможных значений случайных величин η_n , $n \geq 1$, и, значит, $P_n(B) = P\{\eta_n \in B\} = 1$. Однако множество B дискретно, и, следовательно, $P(B) = P\{\theta \in B\} = 0$. Таким образом,

$$\lim_{n \rightarrow \infty} P_n(B) = 1 \neq 0 = P(B).$$

Заметим, что здесь, как и в примере (б), $P(\partial B) \neq 0$.

(г) Мы показали в предыдущем примере (в), что соотношение $P_n(B) \rightarrow P(B)$ не имеет места для множества B . При этом B — дискретное множество, состоящее, вообще говоря, из иррациональных чисел. Оказывается, в качестве B можно взять множество $\tilde{B}$ рациональных чисел на прямой $\mathbb{R}^1$. Действительно, пусть $p = \frac{1}{2}$, а $\tilde{\eta}_n = \eta_{n^2}$, т. е. $\tilde{\eta}_n = \frac{1}{n}(2S_{n^2} - n^2)$. Тогда $\tilde{\eta}_n \xrightarrow{d} \theta$ при $n \rightarrow \infty$, но

$$P_n(\tilde{B}) = P\{\tilde{\eta}_n \in \tilde{B}\} = 1 \not\rightarrow 0 = P(\tilde{B}) = P\{\theta \in \tilde{B}\}.$$

16.3. О слабой сходимости и сходимости моментов

Пусть случайные величины X и X_n , $n \geq 1$, имеют функции распределения соответственно F и F_n , $n \geq 1$. Обозначим

$$m_k = \int_{-\infty}^{\infty} x^k dF(x), \quad m_k^{(n)} = \int_{-\infty}^{\infty} x^k dF_n(x), \quad k = 1, 2, \dots$$

Предположим, что $F_n \xrightarrow{d} F$ при $n \rightarrow \infty$, и поставим следующий вопрос: вытекает ли отсюда, что $m_k^{(n)} \rightarrow m_k$ при $n \rightarrow \infty$? Покажем, что без дополнительных предположений ответ отрицателен.

(а) Рассмотрим случайную величину X_n , распределенную равномерно на отрезке $[0, n]$. Через F_n обозначим ее функцию распределения. Пусть F_0 — функция распределения вырожденной случайной величины X_0 , $P\{X_0 = 0\} = 1$. Положим

$$G_n(x) = \frac{1}{n} F_n(x) + \left(1 - \frac{1}{n}\right) F_0(x), \quad x \in \mathbb{R}^1, \quad n \geq 1.$$

Тогда G_n — смесь F_n и F_0 , так что G_n также является функцией распределения. Введем случайную величину Y_n , $Y_n \sim G_n$. Поскольку $\frac{1}{n} F_n(x) \rightarrow 0$ при $n \rightarrow \infty$ для каждого $x \in \mathbb{R}^1$, то

$$\lim_{n \rightarrow \infty} G_n(x) = F_0(x) \quad \text{для всех } x \in \mathbb{R}^1$$

(в том числе в точке $x = 0$, являющейся точкой скачка F_0). Тем самым $G_n \xrightarrow{d} F_0$ при $n \rightarrow \infty$, т. е. $Y_n \xrightarrow{d} X_0$ при $n \rightarrow \infty$. Выясним теперь, сходятся ли моменты $m_k^{(n)} = E\{Y_n^k\}$ к моментам $m_k = E\{X_0^k\}$. Очевидно, что $m_k = 0$ для каждого $k \geq 1$. С другой стороны,

$$m_k^{(n)} = \int_{-\infty}^{\infty} x^k dG_n(x) = \frac{n^{k-1}}{k+1},$$

и понятно, что $\lim_{n \rightarrow \infty} m_k^{(n)} = \infty$ для каждого $k = 2, 3, \dots$. Поэтому соотношение $m_k^{(n)} \rightarrow m_k$ не имеет места ни для какого k .

(б) Рассмотрим следующее семейство функций:

$$F_n(x) = \left(1 - \frac{1}{n}\right) \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du + \frac{1}{2n} (1 + 1_{[n, \infty)}(x)), \quad x \in \mathbb{R}^1.$$

Легко проверить, что F_n — функция распределения, и пусть она соответствует случайной величине, скажем X_n . Очевидно, что

$$\lim_{n \rightarrow \infty} F_n(x) = \Phi(x) \quad \text{для всех } x \in \mathbb{R}^1,$$

где Φ — стандартная нормальная функция распределения. Итак,

$$X_n \xrightarrow{d} Z \quad \text{при } n \rightarrow \infty, \quad \text{где } Z \sim \mathcal{N}(0, 1).$$

Случайная величина Z имеет моменты $m_{2k-1} = 0$, $m_{2k} = (2k-1)!!$. Нетрудно видеть, что моменты $m_k^{(n)} = E\{X_n^k\}$ любого порядка k стремятся к бесконечности. Следовательно, $m_k^{(n)} \not\rightarrow m_k$ при $n \rightarrow \infty$.

16.4. Два случая, когда теорема непрерывности неприменима

Предположим, что заданы функции распределения F_0, F_n , $n \geq 1$, и пусть ψ_0, ψ_n , $n \geq 1$ — соответствующие им характеристические функции. Напомним, что, согласно теореме непрерывности,

$$F_n \xrightarrow{d} F_0 \iff \psi_n(t) \rightarrow \psi_0(t), \quad \text{где } \psi_0 \text{ непрерывна в } 0.$$

Если некоторое из условий не выполнено, то теорему нельзя применять. Полезно, например, показать, что непрерывность $\psi_0(t)$ в точке $t = 0$ существенна.

(а) Рассмотрим последовательность случайных величин $\{X_n, n \geq 1\}$, где $X_n \sim \mathcal{N}(0, n)$. Тогда X_n имеет характеристическую функцию

$\psi_n(t) = \exp\left(-\frac{nt^2}{2}\right)$, и очевидно, что

$$\lim_{n \rightarrow \infty} \psi_n(t) = \tilde{\psi}(t), \quad \text{где} \quad \tilde{\psi}(t) = \begin{cases} 0, & \text{если } t \neq 0, \\ 1, & \text{если } t = 0. \end{cases}$$

Предельная функция $\tilde{\psi}(t)$ разрывна при $t = 0$, и теорема непрерывности не может иметь места. Это видно из соотношения

$$F_n(x) = P\{X_n \leq x\} = P\left\{\frac{X_n}{\sqrt{n}} \leq \frac{x}{\sqrt{n}}\right\} = \Phi\left(\frac{x}{\sqrt{n}}\right) \rightarrow \frac{1}{2} \quad \text{при } n \rightarrow \infty.$$

Предел $\lim_{n \rightarrow \infty} F_n(x) = \tilde{F}(x)$ существует для каждого $x \in \mathbb{R}^1$, однако функция $\tilde{F}(x) \equiv \frac{1}{2}$, $x \in \mathbb{R}^1$, не является функцией распределения.

(б) Рассмотрим семейство функций $\{F_n, n \geq 1\}$, где

$$F_n(x) = \begin{cases} 0, & \text{если } x \leq -n, \\ \frac{n+x}{2n}, & \text{если } -n < x \leq n, \\ 1, & \text{если } x > n. \end{cases}$$

Тогда при каждом n функция F_n является функцией распределения. Более того, $\lim_{n \rightarrow \infty} F_n(x) = \frac{1}{2}$, $x \in \mathbb{R}^1$. Таким образом, последовательность $\{F_n, n \geq 1\}$ сходится при $n \rightarrow \infty$, но ее предел, равняющийся $\frac{1}{2}$, не является функцией распределения. Тот факт, что теорема непрерывности не справедлива в этой ситуации, можно объяснить и следующим образом. Если ψ_n — характеристическая функция, соответствующая функции распределения F_n , то $\psi_n(t) = \frac{\sin nt}{nt}$, и очевидно, что $\lim_{n \rightarrow \infty} \psi_n(t) = \tilde{\psi}(t)$ существует, причем $\tilde{\psi}(t) = 1$, если $t = 0$, и $\tilde{\psi}(t) = 0$, если $t \neq 0$. Снова, как и выше, $\tilde{\psi}(t)$ разрывна в точке $t = 0$.

16.5. Когда теорема непрерывности применима для плотностей?

Пусть случайная величина X_n имеет функцию распределения F_n , плотность f_n и характеристическую функцию ψ_n , $n \geq 1$. Теорема непрерывности содержит условия, при которых слабая сходимость $\{F_n\}$ эквивалентна поточечной сходимости $\{\psi_n\}$. Вполне естественно поставить следующий вопрос: как связаны поточечная сходимость характеристических функций $\{\psi_n\}$ и сходимость плотностей $\{f_n\}$?

Пусть X_0 — случайная величина с функцией распределения F_0 , плотностью f_0 и характеристической функцией ψ_0 . Введем следующие

три условия:

$$\lim_{n \rightarrow \infty} f_n(x) = f_0(x) \text{ для почти всех } x \in \mathbb{R}^1, \quad (1)$$

$$F_n \xrightarrow{w} F_0 \text{ при } n \rightarrow \infty, \quad (2)$$

$$\lim_{n \rightarrow \infty} \psi_n(t) = \psi_0(t) \text{ и } \psi_0(t) \text{ непрерывна при } t = 0. \quad (3)$$

Согласно теореме непрерывности, имеем $(2) \iff (3)$. Далее по теореме Шеффэ (см. пример 14.7) $(1) \implies (2)$. Пример 14.7 показывает, что в общем случае $(2) \not\Rightarrow (1)$. Поэтому $(1) \implies (3)$, однако, вообще говоря, $(3) \not\Rightarrow (1)$. Приведем пример, когда $(3) \not\Rightarrow (1)$.

Рассмотрим стандартную нормальную плотность распределения $\varphi(x) = (2\pi)^{-1/2} \exp\left(-\frac{x^2}{2}\right)$ и ее характеристическую функцию $\psi_0(t) = \exp\left(-\frac{t^2}{2}\right)$. Положим

$$f_\lambda(x) = \frac{\varphi(x)(1 - \cos \lambda x)}{1 - \psi_0(\lambda)}, \quad x \in \mathbb{R}^1, \quad (4)$$

$$\psi_\lambda(t) = \frac{2\psi_0(t) - \psi_0(t + \lambda) - \psi_0(t - \lambda)}{2(1 - \psi_0(\lambda))}, \quad t \in \mathbb{R}^1, \quad (5)$$

где λ — произвольное положительное число (возьмем $\lambda = n$).

Нетрудно проверить, что при каждом λ функция $f_\lambda(x)$, $x \in \mathbb{R}^1$, является плотностью, а функция $\psi_\lambda(t)$, $t \in \mathbb{R}^1$, — ее характеристической функцией. Далее,

$$\lim_{\lambda \rightarrow \infty} \psi_\lambda(t) = \psi_0(t) = \exp\left(-\frac{t^2}{2}\right) \text{ для всех } t \in \mathbb{R}^1. \quad (6)$$

Таким образом, условие (3) выполнено и имеет место слабая сходимость (2). Однако

$$\lim_{\lambda \rightarrow \infty} f_\lambda(x) \neq \varphi(x) = (2\pi)^{-1/2} \exp\left(-\frac{x^2}{2}\right), \quad (7)$$

и, значит, условие (1) не выполнено. Сравнивая (6) и (7), заключаем, что в общем случае поточечная сходимость характеристических функций $\{\psi_n\}$ вместе с непрерывностью предельной функции в точке $t = 0$ не обеспечивает сходимость (1) для плотностей $\{f_n\}$. Итак, теорема непрерывности не имеет места для плотностей.

Здесь полезно привести следующий результат (см. [106]): пусть ψ_n и ψ — абсолютно интегрируемые характеристические функции такие, что

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} |\psi_n(t) - \psi(t)| dt = 0. \quad (8)$$

Тогда соответствующие функции распределения F_n и F имеют ограниченные и непрерывные плотности f_n и f и из (8) вытекает, что

$$\lim_{n \rightarrow \infty} f_n(x) = f(x) \text{ равномерно по } x, \quad x \in \mathbb{R}^1. \quad (9)$$

Очевидно, что в рассмотренном нами примере соотношение (9) не выполнено (см. (7)). Легко заметить также, что поточечная сходимость (6) не достаточна для интегральной сходимости (8).

16.6. О слабой сходимости и сходимости по вариации

Напомним, что для любой пары функций распределения F и G , соответствующих величинам X и Y , можно ввести число, обозначаемое $v(F, G)$ и называемое *расстоянием по вариации* между F и G .

Если X и Y — дискретные случайные величины такие, что $P\{X = x_k\} = p_k$ и $P\{Y = x_k\} = q_k$, где

$$x_k \in \mathbb{R}^1, \quad p_k \geq 0, \quad q_k \geq 0, \quad k = 1, 2, \dots, \quad \sum_k p_k = 1, \quad \sum_k q_k = 1,$$

то

$$v(X, Y) = \sum_k |p_k - q_k|. \quad (1)$$

Если случайные величины X и Y имеют абсолютно непрерывные функции распределения F и G и плотности, соответственно, f и g (относительно меры Лебега), то

$$v(F, G) = \int_{-\infty}^{\infty} |f(x) - g(x)| dx. \quad (2)$$

Отметим, что расстояние по вариации можно ввести для вероятностных мер в довольно общих измеримых пространствах (см. [116]).

Пусть теперь X_0 и X_n , $n \geq 1$, — случайные величины, заданные на некотором вероятностном пространстве, а F_0 и F_n , $n \geq 1$, — соответствующие им функции распределения. Если $v(F_n, F_0) \rightarrow 0$ при $n \rightarrow \infty$, мы пишем $F_n \xrightarrow{v} F_0$ и говорим, что последовательность функций распределения $\{F_n\}$ *сходится по вариации* к функции распределения F_0 . Нетрудно показать, что из сходимости по вариации вытекает слабая сходимость, т. е. $F_n \xrightarrow{v} F_0 \implies F_n \xrightarrow{w} F_0$. Поэтому естественно поставить вопрос о справедливости обратной импликации. Покажем, что ответ отрицательный.

(а) Пусть $F_n(x) = P\{X_n \leq x\}$, $x \in \mathbb{R}^1$, где $P\{X_n = \frac{1}{n}\} = 1$, $n \geq 1$. Тогда очевидно, что $F_n \xrightarrow{w} F_0$ при $n \rightarrow \infty$, где функция распределения F_0 соответствует величине $X_0 = 0$. Однако значение $v(F_n, F_0)$, вычисленное по формуле (1), не стремится к нулю при $n \rightarrow \infty$, т. е. $F_n \not\xrightarrow{v} F_0$.

(б) В примере (а) случайные величины X_0 и X_n , $n \geq 1$, были дискретными (и даже вырожденными). Рассмотрим абсолютно непрерывный случай.

Пусть задана произвольная функция распределения $F(x)$, $x \in \mathbb{R}^1$, с тем единственным условием, что у нее существует плотность, которую обозначим $f(x)$, $x \in \mathbb{R}^1$. Наша цель — построить последовательность функций распределения $\{F_n, n \geq 1\}$ такую, что $v(F_n, F) \rightarrow 0$ при $n \rightarrow \infty$, но тем не менее $F_n \xrightarrow{w} F$.

Итак, обозначим

$$I_n = \int_{-\infty}^{\infty} f(x) \cos^2 nx \, dx, \quad J_n = \int_{-\infty}^{\infty} f(x) \sin^2 nx \, dx, \quad n \geq 1.$$

Из очевидного равенства $\int_{-\infty}^{\infty} f(x) \, dx = I_n + J_n = 1$ вытекает, что числовые последовательности $\{I_n, n \geq 1\}$ и $\{J_n, n \geq 1\}$ не могут одновременно сходиться к нулю при $n \rightarrow \infty$. Предположим, что $I_n \rightarrow 0$ при $n \rightarrow \infty$. В этом случае определим функцию

$$f_n(x) = c_n f(x)(1 + \cos nx), \quad x \in \mathbb{R}^1,$$

где $c_n^{-1} = \int_{-\infty}^{\infty} f(x)(1 + \cos nx) \, dx$. Тогда при каждом n функция f_n является плотностью, и пусть F_n — соответствующая ей функция распределения. Найдем теперь предел последовательности $\{F_n\}$ при $n \rightarrow \infty$. Поскольку f — вероятностная плотность, то, согласно хорошо известной из анализа лемме Римана–Лебега (см., например, [61]), соотношение

$$\int_B f(x) \cos nx \, dx \rightarrow 0 \quad \text{при } n \rightarrow \infty$$

имеет место для любого борелевского множества $B \in \mathcal{B}^1$. Следовательно,

$$\int_B f_n(x) \, dx \rightarrow \int_B f(x) \, dx \quad \text{при } n \rightarrow \infty.$$

Отсюда вытекает слабая сходимость $F_n \xrightarrow{w} F$ при $n \rightarrow \infty$.

Вычислим теперь расстояние по вариации $v(F_n, F)$. Согласно формуле (2) и с учетом связи между плотностями f и f_n , находим

$$v(F_n, F) = \int_{-\infty}^{\infty} |f_n(x) - f(x)| \, dx =$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty} |c_n f(x) \cos nx - (1 - c_n) f(x)| dx \geq \\
&\geq \left| \int_{-\infty}^{\infty} |c_n f(x) \cos nx| dx - \int_{-\infty}^{\infty} |(1 - c_n) f(x)| dx \right|.
\end{aligned}$$

Заметим, что $c_n \rightarrow 1$ при $n \rightarrow \infty$, а

$$\int_{-\infty}^{\infty} f(x) |\cos nx| dx \geq \int_{-\infty}^{\infty} f(x) \cos^2 nx dx = I_n \rightarrow 0.$$

Следовательно, $v(F_n, F) \rightarrow 0$ при $n \rightarrow \infty$. Таким образом, $\{F_n\}$ не сходятся по вариации к F , несмотря на установленную выше слабую сходимость.

16.7. Слабая сходимость и метрика Леви

Пусть заданы функции распределения $F(x)$, $x \in \mathbb{R}^1$, и $G(x)$, $x \in \mathbb{R}^1$. Расстояние (метрика) Леви между F и G обозначается $L(F, G)$ и задается так:

$$L(F, G) = \inf \{ \varepsilon > 0 : F(x - \varepsilon) - \varepsilon \leq G(x) \leq F(x + \varepsilon) + \varepsilon, x \in \mathbb{R}^1 \}.$$

Хорошо известен следующий результат (см. [24, 116]): пусть F и F_n , $n \geq 1$, — функции распределения. Тогда при $n \rightarrow \infty$

$$F_n \xrightarrow{w} F \iff L(F_n, F) \rightarrow 0.$$

Рассмотрим теперь последовательность $\{X_n, n \geq 1\}$ независимых случайных величин. Обозначим $S_n = X_1 + \dots + X_n$ и $s_n^2 = DS_n$, и пусть F_n — функция распределения нормированной суммы $\frac{S_n}{s_n}$. Предположим, что величины X_n таковы, что имеет место слабая сходимость $F_n \xrightarrow{w} G$, где G принадлежит классу безгранично делимых распределений, но здесь для нас это не важно. Это эквивалентно тому, что для каждого $\varepsilon > 0$ найдется такой индекс n_ε , что $L(F_n, G) < \varepsilon$ для всех $n \geq n_\varepsilon$. Поскольку величина $L(F_n, G)$ «мала», то, вообще говоря, можно предположить, что величина $L(\tilde{F}_n, \tilde{G}_n)$ тоже «мала», где $\tilde{F}_n$ — функция распределения суммы S_n (без нормировки!), а $\tilde{G}_n(x) = G(xs_n)$.

Покажем, что из неравенства $L(F_n, G) < \varepsilon$ не всегда вытекает, что $L(\tilde{F}_n, \tilde{G}_n) < \delta(\varepsilon)$, где $\delta(\varepsilon) \rightarrow 0$ при $\varepsilon \rightarrow 0$.

Рассмотрим независимые в совокупности случайные величины X_{nj} , $j = 1, \dots, n$, $n \geq 1$, где

$$P\{X_{nj} = \pm 1\} = \frac{1}{2} \left(1 - \frac{1}{n}\right), \quad P\{X_{nj} = \pm n\sqrt{5}\} = \frac{1}{2n}.$$

Положим $S_n = X_{n1} + \dots + X_{nn}$. Тогда $ES_n = 0$, $DS_n = 5n^2 + n - 1 \rightarrow \infty$ при $n \rightarrow \infty$. Нормированная величина $\eta_n = \frac{S_n}{n\sqrt{5}}$ имеет $E\eta_n = 0$ и $D\eta_n = 1 + \frac{n-1}{5n^2}$, т. е. $D\eta_n \rightarrow 1$ при $n \rightarrow \infty$. Найдем, предел функции распределения $F_n(x) = P\{\eta_n \leq x\}$ при $n \rightarrow \infty$. Для этой цели рассмотрим характеристическую функцию $\psi_n(t) = E\{e^{it\eta_n}\}$. Учитывая конкретную структуру величин X_{nj} , легко находим, что

$$\lim_{n \rightarrow \infty} \psi_n(t) = \psi(t) = \exp(\cos t - 1), \quad t \in \mathbb{R}^1.$$

Оказывается, однако, что функция $\psi(t) = \exp(\cos t - 1)$ является характеристической функцией, отвечающей конкретной случайной величине, а именно величине $\eta_0 = \xi_1 - \xi_2$, где ξ_1 и ξ_2 — независимые случайные величины, каждая имеющая пуассоновское распределение с параметром $\frac{1}{2}$. Тогда по теореме непрерывности

$$F_n \xrightarrow{w} G \quad \text{при } n \rightarrow \infty,$$

где $G(x) = P\{\eta_0 \leq x\}$, что эквивалентно соотношению $\lim_{n \rightarrow \infty} L(F_n, G) = 0$.

Остается найти значение $L(\tilde{F}_n, \tilde{G}_n)$. Легко заметить, что $\tilde{F}_n$ и $\tilde{G}_n$ являются дискретными распределениями. При этом S_n принимает значения в множестве $\{\pm j, \pm k\sqrt{5} \pm l : j, k, l = 1, \dots, n\}$, а функция распределения $\tilde{F}_n$ имеет скачки во всех точках этого множества. Далее, $\tilde{G}_n(x) = P\{\zeta_n \leq x\}$, где $\zeta_n = \eta_0 n\sqrt{5}$, причем очевидно, что случайная величина ζ_n принимает свои значения в множестве $\{0, \pm kn\sqrt{5} : k = 1, 2, \dots\}$, в каждой точке которого $\tilde{G}_n$ имеет скачок. Понятно, что если мы возьмем нечетное n , то существует число $c > 0$ такое, что $L(\tilde{F}_n, \tilde{G}_n) \geq c$ ввиду того, что $P\{\eta_0 = 0\} > 0$.

Таким образом, соотношение $L(\tilde{F}_n, \tilde{G}_n) < \delta(\varepsilon)$ не имеет места.

§ 17. Центральная предельная теорема

Пусть $\{X_n, n \geq 1\}$ — последовательность независимых случайных величин, заданных на вероятностном пространстве $(\Omega, \mathcal{F}, P)$. Обозначим

$$S_n = X_1 + \dots + X_n, \quad a_k = EX_k, \quad A_n = ES_n = a_1 + \dots + a_n, \\ \sigma_k^2 = DX_k, \quad s_n^2 = DS_n = \sigma_1^2 + \dots + \sigma_n^2.$$

Мы говорим, что для последовательности $\{X_n\}$ имеет место (выполнена) *центральная предельная теорема* (ЦПТ), если центрированная и нормированная величина $\frac{S_n - A_n}{s_n}$ сходится по распределению к случайной величине $\xi \sim \mathcal{N}(0, 1)$, т. е. если для каждого $x \in \mathbb{R}^1$

$$\lim_{n \rightarrow \infty} \mathbf{P} \left\{ \frac{S_n - A_n}{s_n} \leq x \right\} = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du.$$

Введем следующие три условия, где F_k — функция распределения величины X_k , $X_{nk} = \frac{X_k - a_k}{s_n}$ и $\varepsilon > 0$ — любое число:

$$\lim_{n \rightarrow \infty} \frac{1}{\sigma_n^2} \sum_{k=1}^n \int_{\{x: |x - a_k| \geq \varepsilon s_n\}} (x - a_k)^2 dF_k(x) = 0 \quad (\text{L})$$

(условие Линдеберга);

$$\lim_{n \rightarrow \infty} \max_{1 \leq k \leq n} \mathbf{E} \{ X_{nk}^2 \} = 0 \quad (\text{F})$$

(условие Феллера);

$$\lim_{n \rightarrow \infty} \max_{1 \leq k \leq n} \mathbf{P} \{ |X_{nk}| \geq \varepsilon \} = 0 \quad (\text{UAN})$$

(условие предельной пренебрегаемости, или, как еще говорят, условие пренебрежимой малости слагаемых; английский термин — «uniform asymptotic negligibility condition»).

Эти условия связаны так: $(\text{L}) \implies (\text{F}) \implies (\text{UAN})$, но оказывается, что каждое из них играет свою специфическую роль для выполнения или невыполнения ЦПТ.

Сформулируем теперь два основных результата.

Теорема Линдеберга. Если последовательность $\{X_n\}$ удовлетворяет условию (L), то для нее имеет место ЦПТ.

Теорема Линдеберга–Феллера. Пусть для последовательности $\{X_n\}$ выполнено условие (F) (или (UAN)). Тогда условие (L) необходимо и достаточно для справедливости ЦПТ.

Эти теоремы, а также ряд их следствий и обобщений составляют существенную часть теории вероятностей. Читатель может обратиться ко многим учебным пособиям и монографиям, например [6, 9, 24, 26, 75, 92, 106, 116, 117, 145, 177, 179, 231, 345, 404, 405, 463].

Рассмотренные далее примеры помогают понять, какова роль отдельных условий в ЦПТ.

17.1. Последовательности случайных величин, для которых центральная предельная теорема не справедлива

(а) Пусть $X_1, X_2, \dots$ — независимые случайные величины такие, что $P\{X_1 = 1\} = P\{X_1 = -1\} = \frac{1}{2}$, а при $k \geq 2$

$$P\{X_k = \pm 1\} = \frac{1}{2c}, \quad P\{X_k = \pm k\} = \frac{1}{2k^2} \left(1 - \frac{1}{c}\right),$$

$$P\{X_k = 0\} = 1 - \frac{1}{c} - \frac{1}{k^2} \left(1 - \frac{1}{c}\right),$$

где $c > 1$. Ответим на вопрос: имеет ли место ЦПТ для последовательности $\{X_k\}$? Начнем с проверки условия Линдеберга (L). Поскольку $a_k = EX_k = 0$, $\sigma_k^2 = EX_k^2 = 1$, $k \geq 1$, то

$$\frac{1}{s_n^2} \sum_{k=1}^n \int_{\{x: |x| \geq \varepsilon s_n\}} x^2 dF_k(x) = \frac{1}{n} \sum_{k=1}^n E\{X_k^2 I(|X_k| \geq \varepsilon \sqrt{n})\}.$$

Если n выбрано так, что $\varepsilon \sqrt{n} > 1$, то

$$\frac{1}{n} \sum_{k=\lceil \varepsilon \sqrt{n} \rceil}^n k^2 P\{|X_k| = k\} \approx \frac{1}{n} (n - \varepsilon \sqrt{n}) \left(1 - \frac{1}{c}\right) \rightarrow 1 - \frac{1}{c} > 0.$$

Это означает, что условие Линдеберга не выполнено. Отсюда, однако, нельзя сделать вывод, что последовательность $\{X_k\}$ не подчиняется ЦПТ, так как условие Линдеберга является только достаточным. Оказывается, ЦПТ действительно не имеет места для $\{X_k\}$, и это можно показать так. Из соотношения

$$P\left\{\frac{|X_k|}{s_n} \geq \varepsilon\right\} = P\{|X_k| \geq \varepsilon \sqrt{n}\} = \begin{cases} 0, & \text{если } k < \varepsilon \sqrt{n}, \\ \frac{1}{k^2} \left(1 - \frac{1}{c}\right), & \text{если } k \geq \varepsilon \sqrt{n}, \end{cases}$$

вытекает, что

$$\max_{1 \leq k \leq n} P\left\{\frac{|X_k|}{s_n} \geq \varepsilon\right\} \leq \frac{1}{\varepsilon^2 n} \left(1 - \frac{1}{c}\right) \rightarrow 0 \quad \text{при } n \rightarrow \infty,$$

т. е. выполнено условие (UAN). Предположим теперь, что для $\{X_k\}$ имеет место ЦПТ, т. е. что $\frac{S_n}{s_n} \xrightarrow{d} \xi$ при $n \rightarrow \infty$, где $\xi \sim \mathcal{N}(0, 1)$. Отсюда вытекало бы выполнение условия Линдеберга (L), которое, как было показано выше, не выполнено в рассматриваемом случае. Поэтому ЦПТ не имеет места для последовательности $\{X_k\}$.

(б) Рассмотрим последовательность $\{\xi_k, k \geq 1\}$ независимых случайных величин, $P\{\xi_k = \pm 1\} = \frac{1}{2}, k = 1, 2, \dots$. Определим новую последовательность $\{X_k, k \geq 1\}$, где $X_k = 4^{-k} \sqrt{15} \xi_k, k \geq 1$, и пусть $S_n = X_1 + \dots + X_n$. Так как $E\xi_k = 0$ и $D\xi_k = 1$, то

$$ES_n = 0, \quad s_n^2 = DS_n = 1 - \left(\frac{1}{16}\right)^n.$$

При больших n мы имеем $s_n^2 \sim 1$ (это и есть причина введения множителя $4^{-k} \sqrt{15}$ при определении X_k). Легко видеть, что

$$P\left\{|S_n| \leq \frac{1}{2}\right\} = 0 \quad \text{для каждого } n \geq 1.$$

Следовательно, при $n \rightarrow \infty$, вероятность $P\{S_n \leq x\}$ не может сходиться к стандартной нормальной функции распределения $\Phi(x)$ для всех $x \in \mathbb{R}^1$ и ЦПТ не имеет места для последовательности $\{X_k\}$.

(в) Рассмотрим серии $\{X_{nk}, k = 1, \dots, n, n \geq 1\}$ из независимых случайных величин, причем все величины в n -й серии $X_{n1}, \dots, X_{nn}$ имеют распределение Пуассона с параметром $\frac{1}{2n}$. Вопрос: имеет ли место ЦПТ для последовательности $\{X_{nk}\}$? Поскольку при больших n

$$P\{X_{nk} = 0\} = e^{-1/(2n)} \approx 1,$$

то каждая величина X_{nk} дает «почти» нулевой вклад в сумму $S_n = X_{n1} + \dots + X_{nn}$. Таким образом, можно предположить, что распределение $\frac{S_n - ES_n}{s_n}$ будет сходиться к стандартной нормальной функции распределения Φ . Это, однако, не так. Заметим, что сумма S_n имеет пуассоновское распределение с параметром $\frac{1}{2}$. В частности, $P\{S_n = 0\} = e^{-1/2}$. Это показывает, что распределение величины $\frac{S_n - ES_n}{s_n}$ не может быть близким к Φ . Другими словами, последовательность $\{X_{nk}, k = 1, \dots, n, n \geq 1\}$ не удовлетворяет ЦПТ.

17.2. Какова роль различных условий для выполнения центральной предельной теоремы?

В начале настоящего параграфа мы сформулировали два основных результата с использованием с использованием условий (L), (F) и (UAN). Выясним теперь, как ЦПТ зависит от этих условий. Для этого рассмотрим последовательность $\{X_n, n \geq 1\}$, состоящую из независимых случайных величин, причем $X_n \sim \mathcal{N}(0, \sigma_n^2)$, где $\sigma_1^2 = 1$ и $\sigma_k^2 = 2^{k-2}$ при $k \geq 2$. Легко видеть, что сумма $S_n = X_1 + \dots + X_n$

имеет математическое ожидание $ES_n = 0$ и дисперсию $s_n^2 = DS_n = 2^{n-1}$. Поскольку $\frac{X_k}{s_n} \sim \mathcal{N}\left(0, \frac{1}{2}\right)$, то

$$\frac{S_n}{s_n} \sim \mathcal{N}(0, 1) \quad \text{для каждого } n.$$

Следовательно, ЦПТ справедлива для последовательности $\{X_n\}$. Далее, очевидно, что

$$\lim_{n \rightarrow \infty} \max_{1 \leq k \leq n} \frac{\sigma_k^2}{s_n} = \lim_{n \rightarrow \infty} \frac{2^{n-2}}{2^{n-1}} = \frac{1}{2} \neq 0$$

и, более того,

$$\max_{1 \leq k \leq n} P\left\{\frac{|X_k|}{s_n} \geq \varepsilon\right\} \geq P\left\{\frac{|X_n|}{s_n} \geq \varepsilon\right\} = 1 - \frac{1}{\sqrt{\pi}} \int_{-\varepsilon}^{\varepsilon} e^{-u^2/2} du > 0.$$

Итак, мы установили, что условие Феллера (F), а также условие (UAN) не выполнены. Отсюда, ввиду того что $\{X_n\}$ удовлетворяет ЦПТ, вытекает, что и условие Линдеберга (L) не выполнено. Следовательно, условие Линдеберга (L) достаточно, но не необходимо для справедливости ЦПТ.

17.3. Две «эквивалентные» последовательности случайных величин, такие что одна из них удовлетворяет, а другая не удовлетворяет центральной предельной теореме

Пусть $\{X_n, n \geq 1\}$ — та же самая последовательность независимых случайных величин, которая введена в примере 17.1. Там мы показали, что $\{X_n\}$ не удовлетворяет ЦПТ. Оказывается, можно построить другую последовательность случайных величин $\{\tilde{X}_n\}$, которая в определенном смысле эквивалентна $\{X_n\}$, и, что интересно, для $\{\tilde{X}_{nk}\}$ уже имеет место ЦПТ.

Рассмотрим последовательность $\{X_{nk}, k = 1, \dots, n, n \geq 1\}$, где

$$\tilde{X}_{nk} = \begin{cases} X_k, & \text{если } |X_k| \geq \sqrt{n}, \\ 0, & \text{если } |X_k| < \sqrt{n}. \end{cases}$$

Тогда $E\tilde{X}_{nk} = 0$, $D\tilde{X}_{nk} = 1$, если $k \leq \sqrt{n}$, и $D\tilde{X}_{nk} = \frac{1}{c}$, если $k > \sqrt{n}$. Следовательно, $E\tilde{S}_n = 0$, где $\tilde{S}_n = \tilde{X}_{n1} + \dots + X_{nn}$, а для дисперсии $\tilde{s}_n^2 = D\tilde{S}_n$ при больших значениях n имеем

$$\tilde{s}_n^2 = \sum_{k=1}^n D\tilde{X}_{nk} = [\sqrt{n}] + \frac{1}{c} (n - [\sqrt{n}]) \approx \frac{n}{c}.$$

Отсюда вытекает, что при $n \rightarrow \infty$

$$\frac{1}{\tilde{s}_n^2} \sum_{k=1}^n \mathbf{E} \left\{ \tilde{X}_{nk}^2 I(|\tilde{X}_{nk}| \geq \varepsilon \tilde{s}_n) \right\} \approx \frac{c}{n} \left(\sqrt{n} - \frac{\varepsilon}{c} \sqrt{n} \right) \left(1 - \frac{1}{c} \right) \rightarrow 0.$$

Это означает, что для последовательности $\{\tilde{X}_{nk}\}$ выполнено условие Линдеберга (L), т. е. ЦПТ имеет место для $\{\tilde{X}_{nk}\}$.

Выясним теперь, какова связь между $\{\tilde{X}_{nk}\}$ и $\{X_{nk}\}$. Покажем, что они «эквивалентны» в следующем смысле:

$$\mathbf{P}\{\tilde{S}_n = S_n\} \rightarrow 1 \quad \text{при } n \rightarrow \infty. \quad (1)$$

Действительно,

$$\mathbf{P}\{\tilde{S}_n \neq S_n\} \leq \sum_{k=1}^n \mathbf{P}\{\tilde{X}_{nk} \neq X_k\} \leq \sum_{k=1}^n \mathbf{P}\{|X_k| > \sqrt{n}\} \leq \sum_{k=[\sqrt{n}]}^n \mathbf{P}\{|X_k| = k\}.$$

Поскольку

$$\mathbf{P}\{|X_k| = k\} = \frac{1}{k^2} \left(1 - \frac{1}{c} \right),$$

то из сходимости ряда $\sum_{k=1}^{\infty} \frac{1}{k^2}$ вытекает, что $\mathbf{P}\{\tilde{S}_n \neq S_n\} \rightarrow 0$, $n \rightarrow \infty$.

Соотношение (1) установлено.

Итак, мы построили две последовательности случайных величин, $\{X_k\}$ и $\{\tilde{X}_{nk}\}$, «эквивалентные» в смысле (1), причем ЦПТ имеет место для $\{\tilde{X}_{nk}\}$ и не имеет места для $\{X_k\}$.

17.4. Условие Ляпунова и центральная предельная теорема

Кроме условия Линдеберга (L) имеется еще одно классическое условие, называемое условием Ляпунова, которое также обеспечивает справедливость ЦПТ. Сформулируем это условие.

Пусть $\{X_n, n \geq 1\}$ — независимые случайные величины, и пусть, как обычно, $a_k = \mathbf{E}X_k$, $\sigma_k^2 = \mathbf{D}X_k$, $S_n = X_1 + \dots + X_n$ и $s_n = (\mathbf{D}S_n)^{1/2}$. Предположим, что для некоторого $\delta > 0$ случайные величины X_k , $k \geq 1$ имеют конечные центральные абсолютные моменты порядка $2 + \delta$, т. е. $\mathbf{E}\{|X_k - a_k|^{2+\delta}\} < \infty$. Условие

$$\lim_{n \rightarrow \infty} \frac{1}{s_n^{2+\delta}} \sum_{k=1}^n \mathbf{E}\{|X_k - a_k|^{2+\delta}\} = 0 \quad (1)$$

называется *условием Ляпунова*. Имеет место

Теорема (Ляпунов). Если выполнено условие (1), то последовательность $\{X_n\}$ удовлетворяет ЦПТ.

Выясним теперь, какова роль условия Ляпунова (1) для ЦПТ, а также какова его связь с условием Линдеберга (L).

(а) Формулировка теоремы Ляпунова подсказывает, что (1) — это достаточное условие для ЦПТ. Легко показать, что (1) не является необходимым для ЦПТ. Обратимся к примеру 17.2, где были рассмотрены независимые случайные величины $X_1, X_2, \dots$ такие, что $X_1 \sim \mathcal{N}(0, 1)$, а при $k \geq 2$ $X_k \sim \mathcal{N}(0, 2^{k-2})$. Легко проверить, что при любом выборе $\delta > 0$ условие (1) не выполняется. Однако, как в примере 17.2, случайная величина $\frac{S_n}{s_n}$ имеет распределение $\mathcal{N}(0, 1)$, и поэтому ЦПТ имеет место для последовательности $\{X_n\}$. Следовательно, условие Ляпунова (1) не является необходимым для справедливости ЦПТ.

(б) Один из способов доказать теорему Ляпунова — проверить, что из условия Ляпунова (1) вытекает условие Линдеберга (L), которое, как уже знаем, обеспечивает ЦПТ. Полезно описать конкретную последовательность случайных величин $\{\xi_n\}$, для которой имеют место ЦПТ и условие Линдеберга (L), а условие Ляпунова (1) не выполнено. Для этого рассмотрим последовательность $\{\xi_n, n \geq 1\}$ независимых и одинаково распределенных случайных величин с плотностью

$$f(x) = \begin{cases} 0, & \text{если } x \leq e, \\ \frac{c}{x^3 \ln^2 x}, & \text{если } x > e. \end{cases}$$

(Здесь $c > 0$ — нормирующая константа: $\int_{-\infty}^{\infty} f(x) dx = 1$.)

Во-первых, $0 < EX_1 < \infty$, $0 < DX_1 < \infty$, и, следовательно, выполнено условие Линдеберга (L). Тем самым ЦПТ справедлива для последовательности $\{\xi_n\}$.

Во-вторых, при произвольном выборе числа $\delta > 0$

$$E\{|X_1|^{2+\delta}\} = \int_e^{\infty} \frac{c dx}{x^{1-\delta} \ln^2 x} = \int_e^{\infty} \frac{cx^{\delta} d(\ln x)}{\ln^2 x} = \int_e^{\infty} \frac{ce^{\delta u}}{u^2} du = \infty.$$

Отсюда вытекает, что условие Ляпунова (1) не может иметь место. Другими словами, для выполнения ЦПТ условие Ляпунова (1) более «жесткое», чем условие Линдеберга (L). Заметим, однако, что при $\delta = 1$, т. е. когда конечен момент третьего порядка, легче проверить условие Ляпунова, чем устанавливать справедливость условия Линдеберга.

(в) Рассмотрим, наконец, последовательность случайных величин $\{\eta_k\}$, для которой не выполнено ни условие Ляпунова, ни условие Линдеберга, ни сама ЦПТ.

Итак, пусть η_k , $k \geq 1$, — независимые случайные величины:

$$P\{\eta_k = \sqrt{k}\} = P\{\eta_k = -\sqrt{k}\} = \frac{1}{2k}, \quad P\{\eta_k = 0\} = 1 - \frac{1}{k}.$$

Очевидно, $E\eta_k = 0$, $D\eta_k = 1$, $E\{|\eta_k|^{2+\delta}\} = k^{\delta/2}$, $s_n^2 = n$, $s_n^{2+\delta} = n^{1+\delta/2}$, а при больших значениях n имеем $\sum_{k=1}^n E\{|\eta_k|^{2+\delta}\} = n^{1+\delta/2}$. Отсюда следует, что условие Ляпунова (1) не выполнено. Поскольку это условие только достаточное, нельзя сказать, имеет место ЦПТ или нет. Чтобы выяснить это, рассуждаем так. Сумма $S_n = \eta_1 + \dots + \eta_n$ имеет математическое ожидание $ES_n = 0$ и дисперсию $s_n^2 = DS_n = n$. Найдем теперь характеристическую функцию $\psi_n(t) = E\{e^{itS_n/\sqrt{n}}\}$ и ее предел при $n \rightarrow \infty$. Записывая $E\{e^{itX_k}\}$ в явном виде, легко получить, что

$$\psi_n(t) = \prod_{k=1}^n \left(1 - \frac{2}{k} \sin^2 \left(\frac{t}{2} \sqrt{\frac{k}{n}} \right) \right).$$

Отсюда нетрудно вывести, что

$$\ln \psi_n(t) \rightarrow -2 \int_0^1 u^{-1} \sin^2 \left(\frac{t}{2} \sqrt{u} \right) du \quad \text{при } n \rightarrow \infty.$$

(Здесь использовано разложение $\ln(1-z)$, $|z| < 1$, и аппроксимация интеграла $\int_0^1 u^{-1} \sin^2 \left(\frac{t\sqrt{u}}{2} \right) du$ соответствующими римановскими суммами.) Таким образом,

$$\lim_{n \rightarrow \infty} \psi_n(t) = \exp \left\{ -2 \int_0^1 u^{-1} \sin^2 \left(\frac{t}{2} \sqrt{u} \right) du \right\} \neq \exp \left(-\frac{t^2}{2} \right).$$

Из теоремы непрерывности следует, что $\frac{S_n}{\sqrt{n}}$ имеет предельное распределение, которое, однако, отличается от нормального. Другими словами, последовательность $\{\eta_k\}$ не удовлетворяет ЦПТ.

Покажем, что для последовательности $\{\eta_k\}$ не выполнено и условие Линдеберга (L). Поскольку $\sigma_k^2 = D\eta_k = 1$ и $s_n^2 = DS_n = n$, то

$$\max_{1 \leq k \leq n} \frac{\sigma_k^2}{s_n^2} = \frac{1}{n} \rightarrow 0 \quad \text{при } n \rightarrow \infty.$$

Это означает, что справедливы условия Феллера (F) и (UAN). Но согласно теореме Линдеберга–Феллера, условие Линдеберга и ЦПТ выполнены или не выполнены одновременно. Поскольку, как было показано выше, $\{\eta_k\}$ не удовлетворяет ЦПТ, условие Линдеберга (L) не выполнено.

То, что для последовательности $\{\eta_k\}$ не выполнено условие Линдеберга, можно проверить и непосредственно.

Полезно сформулировать следующее следствие теоремы Линдеберга: если X_n , $n \geq 1$, — независимые и одинаково распределенные случайные величины с $EX_1 = a$ и $DX_1 = \sigma^2$, то $\frac{S_n - an}{\sigma\sqrt{n}} \xrightarrow{d} \xi$ при $n \rightarrow \infty$, где $\xi \sim \mathcal{N}(0, 1)$, т. е. последовательность $\{X_n\}$ удовлетворяет ЦПТ.

Вернемся теперь к последовательности $\{\eta_k\}$, рассмотренной выше. Знаем, что величины η_k , $k \geq 1$, независимы и все имеют одно и то же математическое ожидание $E\eta_k = 0$ и одну и ту же дисперсию $D\eta_k = 1$. Следовательно, центрированная и нормированная величина $\frac{S_n - A_n}{s_n}$ имеет вид, который встречается в формулировке приведенного выше следствия теоремы Линдеберга. Тем не менее, как было показано, предельное распределение величины $\frac{S_n}{\sqrt{n}}$ отличается от нормального. Вывод такой: в рассмотренной ситуации требование, чтобы случайные величины η_k , $k \geq 1$ были одинаково распределены, существенно для справедливости ЦПТ.

17.5. О центральной предельной теореме и сходимости дисперсии нормированной суммы $S_n/\sqrt{DS_n}$

Рассмотрим две последовательности $\{X_k, k \geq 1\}$ и $\{Y_k, k \geq 1\}$, каждая состоит из независимых случайных величин, причем

$$P\{X_k = \pm 1\} = \frac{1}{2} \left(1 - \frac{1}{k^2}\right), \quad P\{X_k = \pm k\} = \frac{1}{2k^2}, \quad P\{Y_k = \pm 1\} = \frac{1}{2}.$$

Обозначим

$$S_n = Y_1 + \dots + Y_n, \quad \tilde{S}_N = X_1 + \dots + X_n.$$

Поскольку величины Y_k , $k \geq 1$, распределены одинаково, $EY_1 = 0$ и $DY_1 = 1$, для последовательности $\{Y_k\}$ имеет место ЦПТ (см. сформулированное в примере 17.4 следствие теоремы Линдеберга).

Величины X_k , $k \geq 1$, распределены неодинаково, но, например, используя принцип урезания (см. [26, 75, 106]), можно показать, что асимптотическое распределение величин $\frac{\tilde{S}_n}{\sqrt{n}}$ и $\frac{S_n}{\sqrt{n}}$ одно и то же. Следовательно, $\frac{\tilde{S}_n}{\sqrt{n}} \xrightarrow{d} \eta$ при $n \rightarrow \infty$, где $\eta \sim \mathcal{N}(0, 1)$.

Наша цель теперь — найти предел дисперсии каждой из величин $\frac{S_n}{\sqrt{n}}$ и $\frac{\tilde{S}_n}{\sqrt{n}}$. Можно предположить, что в обоих случаях этот предел равен 1. Посмотрим, так ли это.

Для последовательности $\{Y_k\}$, $DY_k = 1$ и $DS_n = n$. Следовательно, $D\left\{\frac{S_n}{\sqrt{n}}\right\} = 1$ при каждом n , и, значит,

$$D\left\{\frac{S_n}{\sqrt{n}}\right\} \rightarrow 1 \quad \text{при } n \rightarrow \infty.$$

Поскольку $EX_k = 0$ и $DX_k = 2 - \frac{1}{k^2}$, то

$$D\left\{\frac{\tilde{S}_n}{\sqrt{n}}\right\} = \frac{1}{n} \sum_{k=1}^n \left(2 - \frac{1}{k^2}\right) = 2 - \frac{1}{n} \sum_{k=1}^n \frac{1}{k^2}.$$

Учитывая сходимость ряда $\sum_{k=1}^n \frac{1}{k^2}$, легко находим, что

$$D\left\{\frac{\tilde{S}_n}{\sqrt{n}}\right\} \rightarrow 2 \quad \text{при } n \rightarrow \infty.$$

Итак, ЦПТ имеет место для последовательностей $\{Y_k\}$ и $\{X_k\}$, т. е. $\frac{S_n}{\sqrt{n}} \xrightarrow{d} \xi$ и $\frac{\tilde{S}_n}{\sqrt{n}} \xrightarrow{d} \eta$, где $\xi \sim \mathcal{N}(0, 1)$ и $\eta \sim \mathcal{N}(0, 1)$. Кроме того, мы показали, что при $n \rightarrow \infty$

$$D\left\{\frac{S_n}{\sqrt{n}}\right\} \rightarrow 1 = D\xi, \quad \text{но} \quad D\left\{\frac{\tilde{S}_n}{\sqrt{n}}\right\} \rightarrow 2 \neq 1 = D\eta.$$

Отсюда вытекает, что ЦПТ не обеспечивает сходимости моментов нормированных сумм $\frac{\tilde{S}_n}{\sqrt{n}}$ к соответствующим моментам нормального распределения $\mathcal{N}(0, 1)$. И это не удивительно, поскольку в общем случае из слабой сходимости не следует сходимость соответствующих моментов (см. пример 16.3). Для сходимости моментов нужно потребовать дополнительные условия интегрируемости рассматриваемых величин.

Например, если $\frac{\tilde{S}_n}{\sqrt{n}} \xrightarrow{d} \eta \sim \mathcal{N}(0, 1)$ и $E\left\{\left|\frac{\tilde{S}_n}{\sqrt{n}}\right|^{2+\delta}\right\} < \infty$ для некоторого $\delta > 0$, то $D\left\{\frac{\tilde{S}_n}{\sqrt{n}}\right\} \rightarrow D\eta$.

17.6. Центральная предельная теорема не всегда применима при суммировании случайного числа случайных величин

Пусть $\{X_n, n \geq 1\}$ — последовательность случайных величин, для которой имеет место ЦПТ. Рассмотрим еще последовательность $\{\nu_n, n \geq 1\}$ положительных целочисленных случайных величин таких, что $\nu_n \xrightarrow{\text{п.н.}} \infty$ при $n \rightarrow \infty$. Положим

$$T_n = S_{\nu_n} = X_1 + \dots + X_{\nu_n}, \quad b_n^2 = DT_n.$$

Если для всех $x \in \mathbb{R}^1$ выполнено соотношение

$$\lim_{n \rightarrow \infty} P\left\{\frac{T_n - ET_n}{b_n} \leq x\right\} = \Phi(x),$$

то мы говорим, что ЦПТ справедлива при суммировании случайного числа случайных величин. Очевидно, что если $\nu_n = n$ п. н. для каждого n , то $T_n = S_n = X_1 + \dots + X_n$, $b_n^2 = s_n^2$ и $\frac{T_n}{b_n} \xrightarrow{d} \xi$, где случайная величина $\xi \sim \mathcal{N}(0, 1)$. Поэтому интересно рассмотреть хотя бы один случай, когда величины ν_n невырождены.

Опишем простой пример, показывающий, что ЦПТ не всегда справедлива при случайном суммировании.

Итак, пусть последовательность $\{X_n, n \geq 1\}$ состоит из независимых одинаково распределенных случайных величин таких, что $P\{X_1 = 1\} = P\{X_1 = -1\} = \frac{1}{2}$. Пусть далее $\{\nu_n, n \geq 1\}$ — независимые случайные величины такие, что $P\{\nu_n = n\} = p$, $P\{\nu_n = 2n\} = q$, где $0 < p < 1$, $q = 1 - p$. Дополнительно предположим, что последовательности $\{X_n\}$ и $\{\nu_n\}$ независимы между собой. Тогда

$$b_n^2 = DT_n = pE\{S_n^2\} + qE\{S_{2n}^2\} = (1 + q)n.$$

Легко видеть, что величина $\frac{T_n}{b_n}$ не сходится по распределению к случайной величине $\xi \sim \mathcal{N}(0, 1)$. Интересующая нас вероятность $P\left\{\frac{T_n}{b_n} \leq x\right\}$ стремится к смеси распределений двух случайных величин ξ_1 и ξ_2 , где $\xi_1 \sim \mathcal{N}(0, (1 + q)^{-2})$, $\xi_2 \sim \mathcal{N}(0, 2(1 + q)^{-2})$ с весами, соответственно равными p и q .

Отметим, что по поводу предельных теорем для случайных сумм читатель может обратиться, например, к работам [25, 64, 253].

17.7. Какова связь между двумя центральными предельными теоремами, интегральной и локальной?

Пусть $\{X_n, n \geq 1\}$ — последовательность независимых случайных величин. Через F_n обозначим функцию распределения величины $\frac{S_n - ES_n}{s_n}$, и пусть f_n — плотность F_n (предполагая, что она существует).

Здесь, как обычно, $S_n = X_1 + \dots + X_n$, $s_n^2 = DS_n$.

Запишем следующие два соотношения:

$$\lim_{n \rightarrow \infty} F_n(x) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du, \quad x \in \mathbb{R}^1, \quad (1)$$

$$\lim_{n \rightarrow \infty} f_n(x) = \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad x \in \mathbb{R}^1. \quad (2)$$

Напомним, что $\{X_n\}$ удовлетворяет *интегральной* ЦПТ, если выполнено соотношение (1), и *локальной* ЦПТ в случае (2). Нетрудно показать, что (2) $\implies$ (1), и эта импликация верна, даже если сходимости в (2) имеет место не для всех $x \in \mathbb{R}^1$, а для почти всех x относительно меры Лебега. По этому поводу полезно обратиться к примеру 14.7, где показано, что в общем случае не имеет места утверждение, обратное к теореме Шеффэ. Поскольку в ЦПТ предел — стандартное нормальное распределение, имеет смысл поставить вопрос о справедливости импликации (1) $\implies$ (2). Приведем конкретный пример, показывающий, что в общем случае это не так.

Пусть последовательность $\{X_n, n \geq 1\}$ состоит из независимых и одинаково распределенных случайных величин с плотностью

$$f(x) = \begin{cases} \frac{1}{2|x| \ln^2 |x|}, & \text{если } |x| < \frac{1}{e}, \\ 0, & \text{если } |x| \geq \frac{1}{e}. \end{cases}$$

Легко видеть, что f — неограничена, $\lim_{|x| \rightarrow 0} f(x) = +\infty$, но сами величины X_n ограничены и невырождены, так что

$$0 < \sigma^2 = DX_n = \int_0^{e^{-1}} \frac{x}{\ln^2 x} dx < \infty.$$

Ввиду того что X_n одинаково распределены, интегральная ЦПТ имеет место для последовательности $\{X_n\}$.

Чтобы установить, имеет ли место локальная ЦПТ для $\{X_n\}$, нужно рассмотреть предельное поведение плотности f_n величины $\frac{S_n}{\sigma\sqrt{n}}$. Для этой цели введем также плотность g_n , суммы $S_n = X_1 + \dots + X_n$ (без какой-либо нормировки). (Понятно, что f_n и g_n связаны соотношением $f_n(x) = s_n g_n(x s_n)$, $s_n = \sigma\sqrt{n}$.)

При $n = 2$ плотность g_2 выражается в виде следующей свертки:

$$g_2(x) = \int_{-e^{-1}}^{e^{-1}} f(u) f(x - u) du.$$

Оценим g_2 снизу. Достаточно рассмотреть значения x в окрестности 0, например, $0 < x < e^{-1}$. В таком случае

$$g_2(x) \geq \int_{-x}^x f(u) f(x - u) du.$$

Поскольку в области $|u| \leq x$ функция $f(x-u)$ достигает своего минимума в точке $u=0$, находим

$$g_2(x) \geq \frac{1}{2x \ln^2 x} \int_{-x}^x \frac{du}{2|u| \ln^2 |u|} = \frac{1}{2x |\ln^3 x|}.$$

Аналогично при $n=3$ плотность g_3 допускает в окрестности 0 следующую оценку:

$$g_3(x) > \frac{c_3}{x \ln^4 x}, \quad c_3 = \text{const} > 0.$$

В общем случае для плотности g_n суммы S_n находим, что в окрестности 0 имеет место оценка

$$g_n(x) > \frac{c_n}{x |\ln^{n+1} x|}, \quad c_n = \text{const} > 0.$$

Это показывает, что для каждого n плотность $g_n(x)$ принимает в точке $x=0$ значение, равное $+\infty$. Так как интересующая нас плотность f_n получается из g_n путем нормировки, то $f_n(x)$ не может сходиться к стандартной нормальной плотности $\varphi(x)$, $x \in \mathbb{R}^1$, при $n \rightarrow \infty$. Тем самым локальная ЦПТ не имеет места для последовательности $\{X_n\}$.

Отметим, что по поводу общих условий для справедливости локальной ЦПТ, читатель может обратиться к монографии [26].

§ 18. Разные предельные теоремы

Здесь собрано несколько разнообразных примеров, посвященных предельному поведению случайных последовательностей. Некоторые определения и необходимые факты приведены в самих примерах. Сформулируем следующий результат:

Теорема Колмогорова о «трех рядах». Пусть $\{X_n, n \geq 1\}$ — последовательность независимых случайных величин. Обозначим $X_n^{(c)} = X_n$, если $|X_n| \leq c$, и $X_n^{(c)} = 0$, если $|X_n| > c$, где c — константа. Тогда для сходимости с вероятностью 1 ряда $\sum_{n=1}^{\infty} X_n$ необходимо, чтобы для любого $c > 0$ сходились ряды

$$\sum_{n=1}^{\infty} \mathbb{E}\{X_n^{(c)}\}, \quad \sum_{n=1}^{\infty} \mathbb{D}\{X_n^{(c)}\}, \quad \sum_{n=1}^{\infty} \mathbb{P}\{|X_n| \geq c\},$$

и достаточно, чтобы эти ряды сходились при некотором $c > 0$.

Доказательство этой теоремы и ее вариантов и следствий можно найти, например, в книгах [60, 116, 157, 177].

18.1. В теореме Колмогорова о «трех рядах» условие независимости существенно

Начнем с формулировки следующего результата, называемого *теоремой о «двух рядах»* (см. [116, 177, 404]), хотя его можно легко вывести из теоремы о «трех рядах»: *пусть $X_n, n \geq 1$, — независимые случайные величины. Обозначим $a_n = EX_n, \sigma_n^2 = DX_n$ и предположим, что ряды $\sum_{n=1}^{\infty} a_n$ и $\sum_{n=1}^{\infty} \sigma_n^2$ сходятся. Тогда ряд $\sum_{n=1}^{\infty} X_n$ сходится с вероятностью 1.* Покажем, что условие независимости величин $X_n, n \geq 1$, существенно для справедливости этого утверждения.

(а) Пусть случайная величина ξ невырождена и $E\xi = 0$ и $0 < D\xi = b^2 < \infty$. Положим $X_n = \frac{\xi}{n}, n \geq 1$. Тогда $a_n = EX_n = 0, \sigma_n^2 = DX_n = \frac{b^2}{n^2}$, так что $\sum_{n=1}^{\infty} a_n = 0$ и $\sum_{n=1}^{\infty} \sigma_n^2 = b^2 \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$. Таким образом, для последовательности $\{X_n, n \geq 1\}$ выполнена часть условий в теореме о «двух рядах». Очевидно, что не выполнено требование о независимости величин $X_n, n \geq 1$. Тем не менее имеет смысл поставить вопрос о сходимости ряда $\sum_{n=1}^{\infty} X_n$. Поскольку

$$\sum_{n=1}^{\infty} X_n(\omega) = \xi(\omega) \sum_{n=1}^{\infty} \frac{1}{n},$$

то ряд $\sum_{n=1}^{\infty} X_n(\omega)$ сходится на множестве $A = \{\omega: \xi(\omega) = 0\}$ и расходится на множестве $\bar{A} = \{\omega: \xi(\omega) \neq 0\}$. Понятно, ввиду невырожденности ξ , что $0 < P(\bar{A}) \leq 1$. Более того, для любого наперед заданного числа $p, 0 < p \leq 1$, можно так задать случайную величину ξ , что ряд $\sum_{n=1}^{\infty} X_n(\omega)$ будет расходиться на множестве $\bar{A}$, вероятность которого $P(\bar{A}) = p$.

(б) Легко видеть, что имеет место следующий вариант упомянутой выше теоремы о «двух рядах»: *пусть случайные величины $X_n, n \geq 1$, независимы и одинаково распределены с $EX_1 = 0$ и $E(X_1^2) < \infty$, а $\{c_n, n \geq 1\}$ — действительные числа, для которых $\sum_{n=1}^{\infty} c_n^2 < \infty$. Тогда ряд $\sum_{n=1}^{\infty} c_n X_n$ сходится с вероятностью 1.* Покажем, что этот результат перестает быть верным, если ослабить требование независимости, заменяя его условием перестановочности последовательности $\{X_n, n \geq 1\}$ (см. пример 13.4).

Итак, пусть $\{\xi_n, n \geq 1\}$ — последовательность независимых и одинаково распределенных случайных величин с $E\xi_1 = 0$ и $E\{\xi_1^2\} < \infty$. Пусть величина η не зависит от $\{\xi_n, n \geq 1\}$, причем $E\eta = 0$ и $E\{\eta^2\} < \infty$,

и положим $X_n = \xi_n + \eta$, $n \geq 1$. Тогда последовательность $\{X_n, n \geq 1\}$ состоит из одинаково распределенных случайных величин с $EX_1 = 0$ и $E\{X_1^2\} < \infty$.

Легко видеть, что для $\{X_n, n \geq 1\}$ не выполнено свойство независимости, но имеет место свойство перестановочности. Исследуем вопрос о сходимости ряда $\sum_{n=1}^{\infty} c_n X_n$, где «весовая» последовательность $\{c_n, n \geq 1\}$ удовлетворяет условию $\sum_{n=1}^{\infty} c_n^2 < \infty$. Выберем эту последовательность так, чтобы $c_n > 0$, $n \geq 1$, и $\sum_{n=1}^{\infty} c_n = \infty$ (например, $c_n = \frac{1}{n}$). Поскольку $c_n X_n = c_n \xi_n + c_n \eta$, имеет место соотношение

$$\sum_{n=1}^{\infty} c_n X_n = \sum_{n=1}^{\infty} c_n \xi_n + \eta \sum_{n=1}^{\infty} c_n.$$

Согласно теореме о «двух рядах», ряд $\sum_{n=1}^{\infty} c_n \xi_n$ сходится с вероятностью 1. Ввиду того, что $\sum_{n=1}^{\infty} c_n = \infty$, ряд $\sum_{n=1}^{\infty} c_n X_n$ будет сходиться только на множестве $A = \{\omega: \eta(\omega) = 0\}$. Но случайная величина η может быть выбрана так, чтобы $P\{\omega: \eta(\omega) \neq 0\} = p$, где p — заранее заданное число, $0 < p \leq 1$. Тогда ряд $\sum_{n=1}^{\infty} c_n X_n$ будет расходиться на множестве $\bar{A}$, вероятность которого $P(\bar{A}) = p$.

Итак, примеры (а) и (б) показывают, что условие независимости существенно для сходимости случайных рядов с вероятностью 1.

Эти же примеры позволяют сделать еще один важный вывод. Согласно закону «0 или 1» Колмогорова, если $\{X_n, n \geq 1\}$ — независимые случайные величины, то множество $\left\{\omega: \sum_{n=1}^{\infty} X_n(\omega) \text{ сходится}\right\}$ имеет вероятность 0 или 1. Мы убедились выше, что без требования независимости это не так. Вероятность $P\left\{\omega: \sum_{n=1}^{\infty} X_n(\omega) \text{ сходится}\right\}$ может равняться любому числу q , где $q \in [0, 1)$. Следовательно, свойство независимости существенно и при применении закона «0 или 1» Колмогорова.

18.2. Об одном следствии теоремы Колмогорова о «трех рядах»

Следствием теоремы Колмогорова о «трех рядах» является следующий результат (см. [177]): *если $X_n, n \geq 1$, — независимые случайные*

величины с $EX_n = 0$, $n \geq 1$, и

$$\sum_{n=1}^{\infty} E\{X_n^2 I(|X_n| \leq 1) + |X_n| I(|X_n| > 1)\} < \infty, \quad (1)$$

то ряд $\sum_{n=1}^{\infty} X_n$ сходится с вероятностью 1.

Выясним роль условия (1) для сходимости ряда $\sum_{n=1}^{\infty} X_n$. Рассмотрим последовательность $\{\xi_n, n \geq 1\}$ независимых и одинаково распределенных случайных величин с $P\{\xi_1 = 1\} = P\{\xi_1 = -1\} = \frac{1}{2}$. Положим $X_n = \frac{\xi_n}{\sqrt{n}}$, $n \geq 1$. Легко проверить, что

$$\sum_{n=1}^{\infty} E\{|X_n|^r\} < \infty \quad \text{при любом } r > 2. \quad (2)$$

Поскольку (2) включает интегрируемость порядка $r > 2$, то может показаться, что условие (1) является следствием (2), что, однако, не так, т. е. условие (1) не выполнено. В то же время ряд $\sum_{n=1}^{\infty} X_n$ расходится с вероятностью 1, что вытекает из теоремы о «трех рядах». Таким образом, при нарушении условия (1) (являющегося только достаточным) можно потерять сходимость случайного ряда $\sum_{n=1}^{\infty} X_n$.

Отметим, наконец, что если условие (2) выполнено при некотором $r \in (0, 2]$, то ряд $\sum_{n=1}^{\infty} X_n$ сходится с вероятностью 1 (см. [75]).

18.3. О радиусе сходимости случайных степенных рядов

Предположим, что задана последовательность $a_n(\omega)$, $n \geq 0$, независимых и одинаково распределенных случайных величин. Случайный степенной ряд, т. е. ряд вида $\sum_{n=0}^{\infty} a_n(\omega)z^n$, где z — комплексное число, определяется обычным образом (см. [358]). Здесь, как и в детерминированном случае (когда a_n , $n \geq 0$, — числа), одна из важных проблем — найти так называемый *радиус сходимости* $r = r(\omega)$. Поясним, что это случайная величина такая, что для всех $|z| \leq r$ ряд $\sum_{n=0}^{\infty} a_n(\omega)z^n$ сходится с вероятностью 1. При этом

$$r(\omega) = \frac{1}{\sup_{n \rightarrow \infty} \sqrt[n]{|a_n(\omega)|}}.$$

Среди множества результатов, касающихся случайных степенных рядов, приведем следующий результат (см. [358]): *пусть независимые*

и одинаково распределенные случайные величины a_n , $n \geq 0$, таковы, что $|a_1|$ имеет функцию распределения $F(x)$, $x \in \mathbb{R}^+$, удовлетворяющую условию

$$\int_1^{\infty} \ln x dF(x) < \infty. \quad (1)$$

Тогда случайный степенной ряд $\sum_{n=0}^{\infty} a_n(\omega) z^n$ имеет радиус сходимости $r(\omega)$ такой, что $P\{r(\omega) \geq 1\} = 1$.

Покажем на конкретном примере, что условие (1) не является необходимым для существования радиуса сходимости r такого, что $P\{r \geq 1\} = 1$. Для этой цели рассмотрим последовательность случайных величин $\{\xi_n, n \geq 0\}$, которые независимы и равномерно распределены на отрезке $[0,1]$. Положим $a_n(\omega) = \exp\left(\frac{1}{\xi_n(\omega)}\right)$. Величины a_n , $n \geq 0$, независимы и одинаково распределены, причем $|a_1| = a_1$ имеет функцию распределения

$$F(x) = \begin{cases} 0, & \text{если } x < e, \\ 1 - \frac{1}{\ln x}, & \text{если } x \geq e. \end{cases}$$

Легко видеть, что

$$\int_1^{\infty} \ln(x) dF(x) = \infty,$$

так что условие (1) не выполнено.

Используя явный вид функции F , определение радиуса сходимости $r(\omega)$ как $\frac{1}{\sup_{n \rightarrow \infty} \sqrt[n]{|a_n(\omega)|}}$ и один результат Лукача (см. [358, теорема 5.2.3]), приходим к соотношению $P\{\omega: r(\omega) < x\} = 0$ для каждого $x \in (0,1]$. Следовательно, $P\{\omega: r(\omega) \geq 1\} = 1$. Таким образом, условие (1) не является необходимым для того, чтобы случайный степенной ряд $\sum_{n=0}^{\infty} a_n(\omega) z^n$ имел радиус сходимости $r \geq 1$.

18.4. Всегда ли можно менять порядок операций взятия математического ожидания, бесконечного суммирования и предела

Сформулируем сначала один из основных результатов о предельном переходе под знаком математического ожидания. Пусть $Y, X_1, X_2, \dots$ —

такие случайные величины, что $X_n \geq Y$ для всех $n \geq 1$, $EY > -\infty$ и $X_n \leq X_{n+1}$ п.н. при всех $n \geq 1$. Тогда

$$\lim_{n \rightarrow \infty} EX_n = E\left\{\lim_{n \rightarrow \infty} X_n\right\}. \quad (1)$$

Доказательство этого результата, называемого теоремой Лебега о мажорируемой сходимости, можно найти во многих книгах (см. [9, 61, 81, 116, 177]). Отметим, что иногда мы имеем дело с частным случаем, когда $Y \equiv 0$. В этом случае X_n , $n \geq 1$, — монотонно возрастающая последовательность неотрицательных случайных величин, и для нее выполнено соотношение (1). Из этого частного случая вытекает как следствие такое утверждение: если X_n , $n \geq 1$, — последовательность неотрицательных случайных величин, то

$$E\left\{\sum_{n=1}^{\infty} X_n\right\} = \sum_{n=1}^{\infty} EX_n. \quad (2)$$

Проиллюстрируем важность условий, обеспечивающих справедливость соотношений (1) и (2).

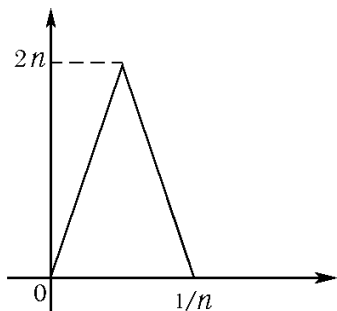


Рис. 3.

(а) Пусть на вероятностном пространстве $(\Omega, \mathcal{F}, P)$, где $\Omega = [0, 1]$, $\mathcal{F} = \mathcal{B}_{[0,1]}$ и P — мера Лебега, задана последовательность случайных величин $\{X_n, n \geq 1\}$, причем X_n имеет вид, изображенный на рис. 3. Очевидно, что

$$P\left\{\omega: \lim_{n \rightarrow \infty} X_n(\omega) = 0\right\} = 1, \quad EX_n = \int_0^1 X_n(\omega) dP = 1.$$

Следовательно,

$$E\left\{\lim_{n \rightarrow \infty} X_n\right\} = 0 \neq 1 = \lim_{n \rightarrow \infty} EX_n.$$

Таким образом, соотношение (2) не выполнено. Заметим, однако, что последовательность $\{X_n, n \geq 1\}$ состоит из неотрицательных величин, которые не обладают свойством монотонности.

(б) Покажем теперь, что соотношение (1) может не выполняться без предположения о неотрицательности величин $X_n, n \geq 1$, даже если ряд $\sum_{n=1}^{\infty} X_n$ сходится с вероятностью 1. Действительно, рассмотрим последовательность $\{\xi_n, n \geq 1\}$ независимых и одинаково распределенных случайных величин с $P\{\xi_1 = 1\} = P\{\xi_1 = -1\} = \frac{1}{2}$. Положим

$$\tau = \inf \left\{ n \geq 1 : \sum_{k=1}^n \xi_k = 1 \right\},$$

где $\inf\{\emptyset\} = \infty$. Легко проверить, что $P\{\tau < \infty\} = 1$. Тогда для последовательности $X_n = \xi_n I(\tau \geq n), n \geq 1$, находим

$$\sum_{n=1}^{\infty} X_n = \sum_{n=1}^{\infty} \xi_n I(\tau \geq n) = \sum_{n=1}^{\tau} \xi_n = 1.$$

Следовательно, $E\left\{\sum_{n=1}^{\infty} X_n\right\} = 1$. Заметим, что событие $\{\tau \geq n\}$ измеримо относительно σ -алгебры $\sigma(\xi_1, \dots, \xi_{n-1})$, а величины ξ_n и $I(\tau \geq n)$ независимы. Из свойств математических ожиданий получаем

$$EX_n = E\xi_n EI(\tau \geq n) = 0.$$

Таким образом,

$$\sum_{n=1}^{\infty} EX_n = 0 \neq 1 = E\left\{\sum_{n=1}^{\infty} X_n\right\}.$$

Это означает, что неотрицательность величин $X_n, n \geq 1$, существенна для выполнения соотношения (1).

18.5. L^r -сходимость последовательности случайных величин не вытекает из сходимости условных математических ожиданий

Пусть на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ заданы случайные величины X и $X_n, n \geq 1$, и пусть они r -интегрируемы при некотором $r \geq 1$. Тогда условные математические ожидания $E(X | \tilde{\mathcal{F}}), E(X_n | \tilde{\mathcal{F}}), n \geq 1$, определены для любой σ -алгебры $\tilde{\mathcal{F}}$, где $\tilde{\mathcal{F}} \subset \mathcal{F}$. Имеется ряд результатов о сходимости X_n к X и сходимости соответствующих условных математических ожиданий. Например, можно показать, что при $n \rightarrow \infty$

$$X_n \xrightarrow{L^r} X \implies E(X_n | \tilde{\mathcal{F}}) \xrightarrow{L^r} E(X | \tilde{\mathcal{F}}) \text{ п. н.} \quad (1)$$

Вопрос: имеет ли место утверждение, обратное к (1)? Покажем, что это не всегда так. Для этой цели рассмотрим последовательность $\{X_n, n \geq 1\}$ независимых и одинаково распределенных случайных величин с $P\{X_1 = 2c\} = P\{X_1 = 0\} = \frac{1}{2}$, где $c \neq 0$ — фиксированное число. Пусть случайная величина $X = c$, т. е. $P\{\omega: X(\omega) = c\} = 1$, а в качестве $\tilde{\mathcal{F}}$ возьмем тривиальную σ -алгебру, т. е. $\tilde{\mathcal{F}} = \{\emptyset, \Omega\}$. Тогда для каждого $n \geq 1$

$$E(X_n | \tilde{\mathcal{F}}) = EX_n = c = E(X | \tilde{\mathcal{F}}).$$

Следовательно, при любом выборе $r \geq 1$ имеем

$$E(X_n | \tilde{\mathcal{F}}) \xrightarrow{L^r} E(X | \tilde{\mathcal{F}}) \quad \text{при } n \rightarrow \infty.$$

Однако при каждом $n \geq 1$

$$E\{|X_n - X|^r\} = E\{|X_n - c|^r\} = c^r.$$

Таким образом, $X_n \not\xrightarrow{L^r} X$ при $n \rightarrow \infty$, что и требовалось показать.

18.6. О законе повторного логарифма Човера

Пусть $X_n, n \geq 1$, — независимые случайные величины, каждая с функцией распределения F . Обозначим

$$S_n = X_1 + \dots + X_n, \quad \xi_n = \frac{S_n}{b_n}, \quad \eta_n = |\xi_n|^{1/\ln \ln n}, \quad n \geq 3,$$

где $b_n > 0, n \geq 1$, — нормирующие константы. При широких предположениях можно описать асимптотическое поведение η_n при $n \rightarrow \infty$. Например, имеет место следующий результат (см. [475]): *пусть существует последовательность $\{b_n\}$ такая, что $\xi_n \xrightarrow{d} \xi$ при $n \rightarrow \infty$, где ξ — устойчивая случайная величина с параметром $\gamma, 0 < \gamma \leq 2$. Тогда $\eta_n \xrightarrow{\text{п.н.}} p$, где p — некоторое число из интервала $[0, \infty)$.*

Отметим, что сходимость с вероятностью 1 последовательности $\eta_n, n \geq 3$, называется *законом повторного логарифма Човера* (см. [475]). Покажем, что хотя бы одно из наложенных условий не является необходимым для справедливости этого результата. Более точно, мы опишем конкретную последовательность случайных величин $\{X_n, n \geq 1\}$ такую, что $\xi_n = \frac{S_n}{b_n}$ при заданных $\{b_n\}$ не сходится по распределению и тем не менее

$$P\left\{\lim_{n \rightarrow \infty} \eta_n = \lim_{n \rightarrow \infty} |\xi_n|^{1/\ln \ln n} = e^{1/\sqrt{2}}\right\} = 1,$$

т. е. выполнен закон повторного логарифма Човера.

Рассмотрим функцию $F(x)$, $x \in \mathbb{R}$, где

$$F(-x) = 1 - F(x) = \begin{cases} \frac{1}{2}, & \text{если } 0 < x < 1, \\ \frac{1}{2}x^{-\sqrt{2}}\left(1 + \frac{1}{12}\sin(\ln x)\right), & \text{если } x \geq 1. \end{cases}$$

Легко проверить, что F — функция распределения. Пусть $\{X_n, n \geq 1\}$ — последовательность независимых и одинаково распределенных случайных величин с общей функцией распределения F . Выберем $b_n = n^{1/\sqrt{2}}$, $n \geq 1$, как нормирующие константы. Заметим, что

$$\frac{11}{12}x^{-\sqrt{2}} \leq 1 - F(x) + F(-x) \leq \frac{13}{12}x^{-\sqrt{2}} \quad \text{при любом } x > 0.$$

Тогда для интересующей нас величины $\eta_n = \left| \frac{S_n}{b_n} \right|^{1/\ln \ln n}$, $n \geq 3$, (при $b_n = n^{1/\sqrt{2}}$) выполнены следующие два соотношения:

$$\begin{aligned} P\left\{|S_n| \geq b_n(\ln n)^{(1-\varepsilon)/\sqrt{2}}\right\} &\geq \frac{c_1}{n(\ln n)^{1-\varepsilon}}, \\ P\left\{|S_n| \geq b_n(\ln n)^{(1+\varepsilon)/\sqrt{2}}\right\} &\leq \frac{c_2}{n(\ln n)^{1+\varepsilon}}. \end{aligned}$$

Эти соотношения, где $c_1 = \text{const} > 0$, $c_2 = \text{const} > 0$, имеют место при произвольном $\varepsilon \in (0, 1)$ и при всех $n \geq n_0$, где n_0 — фиксированное натуральное число. Применяя лемму Бореля–Кантелли и один результат Феллера (см. [226]), заключаем, что

$$\begin{aligned} P\left\{|S_n| \geq b_n(\ln n)^{(1-\varepsilon)/\sqrt{2}} \text{ б. ч.}\right\} &= 1, \\ P\left\{|S_n| \geq b_n(\ln n)^{(1+\varepsilon)/\sqrt{2}} \text{ б. ч.}\right\} &= 0. \end{aligned}$$

Следовательно,

$$P\left\{\lim_{n \rightarrow \infty} \eta_n = e^{1/\sqrt{2}}\right\} = 1.$$

Но, согласно результату Золотарёва и Королюка [42], последовательность $\{\xi_n\}$ не может сходиться по распределению (к невырожденной случайной величине) при любом выборе нормирующих констант $\{b_n\}$.

18.7. О рекордах и максимумах последовательности независимых случайных величин

Пусть последовательность $\{X_n, n \geq 1\}$ состоит из независимых и одинаково распределенных случайных величин с общей функцией распределения F . Напомним, что X_k называется *рекордом* (точнее

k -м рекордом) только тогда, когда $X_k > \max(X_1, \dots, X_{k-1})$. Условимся считать, что X_1 тоже рекорд. Определим целочисленную последовательность $\{\tau_n, n \geq 0\}$ так:

$$\tau_0 = 0, \quad \tau_n = \min\{k: k > \tau_{n-1}, X_k > X_{\tau_{n-1}}\}, \quad n \geq 1.$$

Очевидно, что величины τ_n — это моменты, когда наблюдаются рекордные значения.

Введем также последовательность максимумов $\{M_n, n \geq 1\}$, где $M_n = \max(X_1, \dots, X_n)$. Обсудим некоторые свойства последовательностей $\{X_{\tau_n}, n \geq 1\}$ и $\{M_n, n \geq 1\}$. Напомним сначала одно определение. Последовательность случайных величин $\{\xi_n, n \geq 1\}$ называется *сильно устойчивой*, если существуют нормирующие константы $b_n > 0, n \geq 1$, такие, что $\frac{\xi_n}{b_n} \xrightarrow{\text{п.н.}} 1$ при $n \rightarrow \infty$. Если вместо сходимости п.н. мы имеем $\frac{\xi_n}{b_n} \xrightarrow{P} 1$, то $\{\xi_n\}$ называется *слабо устойчивой*.

Для нас важен следующий результат (см. [406]): *если последовательность рекордов $\{X_{\tau_n}, n \geq 1\}$ слабо устойчива, то этим же свойством обладает и последовательность максимумов $\{M_n, n \geq 1\}$. Более точно, пусть функция распределения $F(x), x \in \mathbb{R}^1$, такова, что $F(x) < 1, x \in \mathbb{R}^1$. Введем функцию $h(x) = -\ln(1 - F(x))$ и обратную к ней функцию $h^{-1}(y) = \inf\{x: h(x) > y\}$. Тогда при $n \rightarrow \infty$*

$$\frac{X_{\tau_n}}{h^{-1}(n)} \xrightarrow{P} 1 \implies \frac{M_n}{h^{-1}(\ln n)} \xrightarrow{P} 1.$$

Естественно возникает вопрос: верно ли обратное утверждение? Покажем, что ответ отрицательный.

Итак, рассмотрим функцию $h(x) = \ln^2 x, x \geq 1$, и пусть F — функция распределения, связанная с h соотношением $h = -\ln(1 - F)$. Возьмем последовательность $\{X_n, n \geq 1\}$ независимых случайных величин, каждая с функцией распределения F , и пусть $\{X_{\tau_n}\}$ и $\{M_n\}$ — соответствующие последовательности рекордов и максимумов. Заметим, что функция $h^{-1}(\ln y) = \exp(\sqrt{\ln y})$ является медленно меняющейся функцией (см. [106]). Этот факт вместе с результатом Гнеденко (см. [252]) позволяют заключить, что последовательность максимумов $\{M_n\}$ слабо устойчива. Более того, из результатов работы [406] следует, что $\{M_n\}$ сильно устойчива. Таким образом, можно предположить, что и последовательность рекордов $\{X_{\tau_n}\}$ сильно устойчива. Оказывается, однако, что $\{X_{\tau_n}\}$ даже не является слабо устойчивой. Это легко обосновать, потому что, согласно [406], для слабой устойчивости $\{X_{\tau_n}\}$ необходимо, чтобы функция h^{-1} была медленно меняющейся. Однако в рассматриваемом нами случае $h^{-1}(\ln^2 y) = y$, так что это необходимое условие не выполнено.

Г Л А В А 4

СЛУЧАЙНЫЕ ПРОЦЕССЫ

§ 19. Основные понятия и свойства случайных процессов

Пусть заданы вероятностное пространство $(\Omega, \mathcal{F}, P)$, множество T , являющееся подмножеством $\mathbb{R}^1$, и измеримое пространство $(E, \mathcal{E})$. Предположим, что при каждом $t \in T$ определена случайная величина X_t со значениями в множестве E . Семейство случайных величин $X = (X_t, t \in T)$ называется *случайным процессом*, T — *временным интервалом* и $(E, \mathcal{E})$ — *пространством состояний* (фазовым пространством) процесса X . Далее мы чаще всего рассматриваем случайные процессы с действительными значениями (т. е. $E \subseteq \mathbb{R}^1$, $\mathcal{E} = \mathcal{B}_E \subseteq \mathcal{B}^1$), причем множество E либо дискретное (подмножество $\mathbb{R}^1$), либо интервал (конечный или бесконечный) полупрямой $\mathbb{R}^+ = [0, \infty)$.

Некоторые классы случайных процессов можно характеризовать с помощью семейства $\mathcal{P}$ всех конечномерных распределений. Напомним, что $\mathcal{P}$ содержит всевозможные распределения (точнее, функции распределения) вида

$$F_{t_1, \dots, t_n}(x_1, \dots, x_n) = P\{X_{t_1} \leq x_1, \dots, X_{t_n} \leq x_n\}$$

при всех $n \geq 1$, $t_1, \dots, t_n \in T$ и $x_1, \dots, x_n \in \mathbb{R}^1$. Мы говорим, что семейство $\mathcal{P}$ обладает *свойством согласованности*, если

$$F_{t_{i_1}, \dots, t_{i_n}}(x_{i_1}, \dots, x_{i_n}) = F_{t_1, \dots, t_n}(x_1, \dots, x_n),$$

где $(i_1, \dots, i_n)$ — произвольная перестановка индексов $(1, \dots, n)$, и, кроме того,

$$F_{t_1, \dots, t_{n-1}, t_n}(x_1, \dots, x_{n-1}, \infty) = F_{t_1, \dots, t_{n-1}}(x_1, \dots, x_{n-1}).$$

В основе теории случайных процессов лежит следующий фундаментальный результат (см. [60, 75, 92, 116, 117, 330]):

Теорема Колмогорова. Пусть задано семейство конечномерных распределений $\mathcal{P}$, удовлетворяющее свойству согласованности. Тогда существуют вероятностное пространство $(\Omega, \mathcal{F}, P)$ и определенный на нем случайный процесс $X = (X_t, t \in T)$ такие, что

$$P\{\omega: X_{t_1} \leq x_1, \dots, X_{t_n} \leq x_n\} = F_{t_1, \dots, t_n}(x_1, \dots, x_n).$$

В общем случае при фиксированном $\omega \in \Omega$ отображение $t \mapsto X_t(\omega)$ множества T в множество E называется *траекторией* (выборочной функцией, реализацией) процесса X . Предположим теперь, что T — некоторый интервал в $\mathbb{R}^+$.

Мы говорим, что $X = (X_t, t \in T)$ — *непрерывный процесс* (более точно, непрерывный Р-п.н.), если почти все его траектории $X_t(\omega)$, $t \in T$, являются непрерывными функциями, т. е. элементами пространства непрерывных функций $\mathcal{C}(T)$.

Процесс X называется *непрерывным справа*, если почти все его траектории — непрерывные справа функции (т. е. если $X_t = X_{t+}$ для каждого t , где $X_{t+} = \lim_{s \downarrow t} X_s$). Если, кроме того, X_t для каждого t имеет предел слева, то мы говорим, что процесс X принадлежит пространству $\mathbb{D}(T)$.

Пусть $(\mathcal{F}_t)_{t \geq 0}$ — возрастающее семейство σ -подалгебр основной σ -алгебры $\mathcal{F}$: $\mathcal{F}_t \subset \mathcal{F}$ для каждого $t \geq 0$ и $\mathcal{F}_s \subset \mathcal{F}_t$ при $s < t$. Семейство $(\mathcal{F}_t)_{t \geq 0}$ называется *фильтрацией*. Положим

$$\mathcal{F}_{t+} = \bigcap_{s > t} \mathcal{F}_s, \quad \mathcal{F}_{t-} = \bigvee_{s < t} \mathcal{F}_s.$$

(Здесь и далее $\bigvee_{s < t} \mathcal{F}_s$ обозначает минимальную σ -алгебру, содержащую все $\mathcal{F}_s$, $s < t$.) При $t = 0$ полагаем $\mathcal{F}_{0-} = \mathcal{F}_0$, а $\mathcal{F}_\infty = \bigvee_{t \geq 0} \mathcal{F}_t$. Фильтрация $(\mathcal{F}_t)_{t \geq 0}$ называется *непрерывной справа*, если $\mathcal{F}_t = \mathcal{F}_{t+}$ для всех $t \geq 0$, и *непрерывной слева*, если $\mathcal{F}_t = \mathcal{F}_{t-}$. Таким образом, $(\mathcal{F}_t)_{t \geq 0}$ — непрерывная фильтрация, если она непрерывна и справа, и слева. Мы говорим, что процесс $X = (X_t, t \geq 0)$ *адаптирован* (согласован) к фильтрации $(\mathcal{F}_t)_{t \geq 0}$ ($(\mathcal{F}_t)$ -адаптирован), если для каждого $t \geq 0$ случайная величина X_t измерима относительно σ -алгебры $\mathcal{F}_t$.

Следующие четыре объекта $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbf{P})$ называются *стохастическим базисом*. Фраза « $X = (X_t, t \geq 0)$ » есть процесс на $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbf{P})$ означает, что процесс X задан на вероятностном пространстве $(\Omega, \mathcal{F}, \mathbf{P})$ и адаптирован к фильтрации $(\mathcal{F}_t)$. Мы говорим, что фильтрация $(\mathcal{F}_t)_{t \geq 0}$ удовлетворяет обычным условиям, если она непрерывна справа, причем $\mathcal{F}_0$, а значит, и каждая $\mathcal{F}_t$ содержит все подмножества Ω из $\mathcal{F}$, имеющие Р-меру нуль.

Несколько понятий и свойств случайных процессов будут введены и проанализированы в примерах.

Читатель может найти систематическое изложение теории случайных процессов во многих книгах, например в [14, 21, 22, 29, 32, 75, 92, 129, 199, 206, 330, 368].

19.1. О роли семейства конечномерных распределений при построении случайного процесса с заданными свойствами

Пусть $\mathcal{P} = \{F_{t_1, \dots, t_n}, n \geq 1, t_i \in T \subset \mathbb{R}^+\}$ — согласованное семейство конечномерных функций распределения. Согласно теореме Колмогорова, существует случайный процесс $X = (X_t, t \in T)$ такой, что его конечномерные функции распределения — это как раз семейство $\mathcal{P}$. Заметим, однако, что о свойствах траекторий процесса X при этом не говорится. А при изучении случайных процессов важно знать, какие свойства X связаны со свойствами семейства $\mathcal{P}$.

Итак, рассмотрим согласованное семейство $\mathcal{P}$ и некоторое двупараметрическое семейство функций $\{X_t(\omega), \omega \in \Omega, t \in T\}$. Обозначим через $\mathcal{A}$ борелевскую алгебру, относительно которой измеримы все $X_t, t \in T$, и пусть $\tilde{\mathcal{A}}$ — минимальная σ -алгебра, порожденная $\mathcal{A}$. Поскольку каждое множество из $\mathcal{A}$ имеет вид

$$A = \{\omega : (X_{t_1}(\omega), \dots, X_{t_n}(\omega)) \in B\}, \quad B \in \mathcal{B}^n,$$

то с помощью равенства

$$P(A) = P\{\omega : (X_{t_1}(\omega), \dots, X_{t_n}(\omega)) \in B\} = \int_B dF_{t_1, \dots, t_n}(x_1, \dots, x_n) \quad (1)$$

мы определяем аддитивную вероятностную меру P на $(\Omega, \mathcal{A})$. Однако может оказаться, что P не является σ -аддитивной, а в этом случае P не может быть продолжена на $(\Omega, \tilde{\mathcal{A}})$. Как следствие, отсюда вытекает, что мы не можем построить процесс X с заранее заданными конечномерными распределениями.

Рассмотрим следующий конкретный пример. Возьмем $T = [0, 1]$, и пусть $\Omega = \mathbb{C}[0, 1]$ (пространство непрерывных функций на $[0, 1]$). Пусть далее $(X_t(\omega))$ — координатная функция, т. е. $X_t(\omega) = \omega(t)$, где $\omega = \{\omega(t), t \in [0, 1]\} \in \mathbb{C}[0, 1]$. Предположим, что семейство $\mathcal{P} = \{F_{t_1, \dots, t_n}\}$ задано следующим образом:

$$F_{t_1, \dots, t_n}(x_1, \dots, x_n) = \prod_{k=1}^n \int_{-\infty}^{x_k} g(u) du, \quad (2)$$

где $g(u), u \in \mathbb{R}^1$, — некоторая вероятностная плотность, симметричная относительно точки 0. Легко видеть, что это семейство $\mathcal{P}$ обладает свойством согласованности. Более того, вероятность P , которую мы хотим найти, должна при каждом $\varepsilon > 0$ удовлетворять соотношению

$$P\{X_t > \varepsilon, X_s < -\varepsilon\} = \left(\int_{\varepsilon}^{\infty} g(u) du \right)^2, \quad s, t \in [0, 1], \quad s \neq t.$$

Поскольку $\Omega = \mathbb{C}[0, 1]$, то для каждого $t \in [0, 1]$ при $n \rightarrow \infty$ множества $A_n = \{\omega: X_t(\omega) > \varepsilon, X_{t+1/n}(\omega) < -\varepsilon\}$ должны сходиться к пустому множеству $\emptyset$ для каждого $\varepsilon > 0$. Однако,

$$\lim_{n \rightarrow \infty} P(A_n) = \left(\int_{\varepsilon}^{\infty} g(u) du \right)^2 \neq 0.$$

Следовательно, мера P , заданная с помощью равенств (1) и (2), не является непрерывной в «точке» $\emptyset$ (это эквивалентно тому, что P не является σ -аддитивной). Таким образом, P не может быть продолжена на σ -алгебру $\tilde{\mathcal{A}}$.

Кроме того, для любой вероятностной меры P на $(\Omega, \mathcal{A})$ имеем $P(\Omega) = 1$, а это означает, что с вероятностью 1 траектории процесса будут непрерывными. Однако, как показано выше, такое свойство несовместимо с семейством $\mathcal{P}$ несмотря на то, что $\mathcal{P}$ удовлетворяет свойству согласованности. В частности, из соотношения (2) вытекает, что $\{X_t, t \in [0, 1]\}$ — совокупность независимых случайных величин, каждая из которых имеет симметричное распределение с плотность g . Следовательно, свойство независимости X_t и X_s для близких моментов времени t и s несовместимо с требуемым свойством непрерывности.

Отметим, наконец, что если семейство $\mathcal{P}$ удовлетворяет некоторым дополнительным условиям, то можно построить случайный процесс с заранее предписанными свойствами (см. [6, 89]).

19.2. У случайного процесса могут быть модификации с различными свойствами

Пусть случайные процессы $X = (X_t, t \in T)$ и $Y = (Y_t, t \in T)$ определены на одном и том же вероятностном пространстве $(\Omega, \mathcal{F}, P)$ и имеют одно и то же пространство состояний $(E, \mathcal{E})$. Процесс Y называется *модификацией* процесса X (и наоборот), если

$$P\{\omega: X_t(\omega) \neq Y_t(\omega)\} = 0 \quad \text{для каждого } t \in T. \quad (1)$$

Если

$$P\left\{ \bigcup_{t \in T} \{\omega: X_t(\omega) \neq Y_t(\omega)\} \right\} = 0, \quad (2)$$

то процессы X и Y называются *неотличимыми*.

С помощью примеров проиллюстрируем связь этих понятий, а также покажем, что у различных модификаций процессов могут быть очень отличающиеся свойства.

(а) Отметим сначала, что если множество T счетно, а процесс Y — модификация процесса X , то X и Y неотличимы (в смысле определения (2)). Поэтому некоторые особенности могут возникнуть, только если T — интервал (конечный или бесконечный).

Пусть $\Omega = [0, 1]$, $\mathcal{F} = \mathcal{B}_{[0,1]}$, P — мера Лебега и $T = \mathbb{R}^+$. Как обычно, символом $[t]$ обозначим целую часть числа t . Рассмотрим два случайных процесса: $X = (X_t, t \in \mathbb{R}^+)$ и $Y = (Y_t, t \in \mathbb{R}^+)$, где

$$X_t(\omega) = 0 \quad \text{для всех } \omega \text{ и всех } t,$$

$$Y_t(\omega) = \begin{cases} 0, & \text{если } t - [t] \neq \omega, \\ 1, & \text{если } t - [t] = \omega. \end{cases}$$

Очевидно, что соотношение (1) выполнено, и, значит, процесс Y является модификацией процесса X . Кроме того, все траектории процесса X непрерывны, в то время как каждая траектория процесса Y разрывна.

(б) Пусть случайная величина τ неотрицательна и имеет абсолютно непрерывное распределение. Определим два процесса: $X = (X_t, t \geq 0)$ и $Y = (Y_t, t \geq 0)$, где

$$X_t = X_t(\omega) = 1_{\{\tau(\omega) \geq t\}}(\omega), \quad X_0 = 0;$$

$$Y_t = Y_t(\omega) = 1_{\{\tau(\omega) < t\}}(\omega), \quad Y_0 = 0.$$

Легко видеть, что для каждого $t \geq 0$

$$P\{\omega: X_t(\omega) \neq Y_t(\omega)\} = P\{\omega: \tau(\omega) = t\} = 0.$$

Следовательно, процесс Y является модификацией процесса X , и наоборот. Однако почти все траектории X непрерывны справа и имеют пределы слева, а почти все траектории Y непрерывны слева и имеют пределы справа.

19.3. О свойстве сепарабельности случайных процессов

Пусть $X = (X_t, t \in T \subset \mathbb{R}^+)$ — случайный процесс, заданный на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ и принимающий значения в измеримом пространстве $(\mathbb{R}^1, \mathcal{B}^1)$. Процесс X называется *сепарабельным*, если в множестве T существует счетное и всюду плотное в нем подмножество S_0 такое, что для каждого замкнутого множества $B \in \mathcal{B}^1$ и каждого открытого множества $I \subset \mathbb{R}^1$ имеет место соотношение

$$\begin{aligned} \{\omega: X_t(\omega) \in B \text{ для всех } t \in T \cap I\} = \\ = \{\omega: X_s(\omega) \in B \text{ для всех } s \in S_0 \cap I\}. \end{aligned}$$

Свойство сепарабельности означает, что любое событие, связанное с процессом X , может быть выражено через счетное число операций типа объединения и пересечения. Однако оказывается, что не всякий процесс сепарабелен.

Пусть τ — случайная величина с равномерным распределением на отрезке $[0, 1]$. Рассмотрим процесс $X = (X_t, t \in [0, 1])$, где

$$X_t = X_t(\omega) = \begin{cases} 1, & \text{если } \tau(\omega) = t, \\ 0, & \text{если } \tau(\omega) \neq t. \end{cases}$$

Ясно, что для любого счетного подмножества $S \subset [0, 1]$

$$\mathbf{P}\{X_t = 0 \text{ для } t \in S\} = 1, \quad \mathbf{P}\{X_t = 0 \text{ для } t \in [0, 1]\} = 0.$$

Следовательно, процесс X не является сепарабельным. Заметим, однако, что у процесса X имеется сепарабельная модификация.

Общие результаты, касающиеся существования сепарабельных модификаций случайных процессов, могут быть найдены, например, в книгах [21, 22, 92, 129, 492].

19.4. Измеримые и прогрессивно измеримые процессы

Пусть на стохастическом базисе $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbf{P})$ задан случайный процесс $X = (X_t, t \geq 0)$, принимающий значения в измеримом пространстве $(E, \mathcal{E})$. Предполагается, что фильтрация $(\mathcal{F}_t)_{t \geq 0}$ удовлетворяет обычным условиям (см. начало настоящего параграфа). Напомним, что при каждом t случайная величина X_t измерима относительно σ -алгебры $\mathcal{F}_t$, и в этом случае мы говорим, что процесс X *адаптирован* относительно фильтрации $(\mathcal{F}_t)$. Процесс X называется *измеримым*, если отображение $(t, \omega) \mapsto X_t(\omega)$ множества $[0, \infty) \times \Omega$ в E является $\mathcal{B}^+ \times \mathcal{F}$ -измеримым. Наконец, процесс X — *прогрессивно измерим*, если это отображение измеримо относительно $\mathcal{B}_{[0, t]} \times \mathcal{F}_t$ для каждого $t \geq 0$.

Приведем примеры, иллюстрирующие введенные понятия.

(а) Пусть случайный процесс $\xi = (\xi_t, t \in [0, 1])$ состоит из независимых случайных величин таких, что $\mathbf{E}\xi_t = 0$ и $\mathbf{E}\{\xi_t^2\} = 1, t \in [0, 1]$. Вопрос: является ли процесс ξ измеримым? Предположим, что ответ положителен, т. е. что $\xi_t(\omega)$ измерима как функция аргументов (t, ω) , причем $\mathbf{E}\xi_t = 0, t \in [0, 1], \mathbf{E}\{\xi_s \xi_t\} = 0$, если $s \neq t$, и $\mathbf{E}\{\xi_s \xi_t\} = 1$, если $s = t$. Тогда для любой подотрезки $I \subset [0, 1]$

$$\int \int_I \int_I |\xi_s(\omega) \xi_t(\omega)| \mathbf{P}(d\omega) ds dt < \infty.$$

Применяя теорему Фубини (см. пример 6.5), находим

$$\mathbb{E} \left\{ \left(\int_I \xi_t(\omega) dt \right)^2 \right\} = \mathbb{E} \left\{ \int_I \int_I \xi_s(\omega) \xi_t(\omega) ds dt \right\} = \int_I \int_I \mathbb{E} \{ \xi_s \xi_t \} ds dt = 0.$$

Таким образом, для множества N_I с $\mathbb{P}(N_I) = 0$ имеем $\int_I \xi_t(\omega) dt = 0$, если $\omega \notin N_I$. Рассмотрим теперь все подынтервалы $I = [r', r'']$, концы которых r' и r'' — рациональные числа, и пусть $N = \bigcup_I N_I$. Тогда $\mathbb{P}(N) = 0$ и для каждого $\omega \in \bar{N}$ имеем $\int_a^b \xi_t(\omega) dt = 0$, где $[a, b]$ — произвольный подотрезок отрезка $[0, 1]$. Это означает, что если $\omega \in \bar{N}$, то $\xi_t(\omega) = 0$ для всех t , за исключением, быть может, некоторого множества нулевой меры Лебега. Снова, применяя теорему Фубини, приходим к равенству

$$\int_{\Omega} \int_0^1 \xi_t^2(\omega) \mathbb{P}(d\omega) dt = 0.$$

Это, однако, невозможно, так как

$$\int_{\Omega} \int_0^1 \xi_t^2(\omega) \mathbb{P}(d\omega) dt = \int_0^1 \mathbb{E} \{ \xi_t^2 \} dt = 1.$$

Полученное противоречие показывает, что процесс ξ не является измеримым. Более того, у ξ нет измеримой модификации.

(б) Если случайный процесс $X = (X_t, t \geq 0)$ прогрессивно измерим, то он $(\mathcal{F}_t)$ -адаптирован и измерим. Вопрос: верно ли, что адаптированный и измеримый процесс является прогрессивно измеримым? Можно догадаться, что ответ отрицателен, но полезно проиллюстрировать это конкретным примером.

Пусть $\Omega = \mathbb{R}^+$, $\mathcal{F} = \mathcal{B}^+$ и $\mathbb{P}(dx) = e^{-x} dx$ (dx соответствует обычной мере Лебега). Рассмотрим множество $\delta = \{(x, x), x \in \mathbb{R}^+\}$, и пусть $\mathcal{F}_t$, $t \in \mathbb{R}^+$, — это σ -алгебра, порожденная точками $\mathbb{R}^+$ (т.е. $A \in \mathcal{F}_t$ тогда и только тогда, когда либо A , либо его дополнение $\bar{A}$ — счетное множество). Введем случайный процесс $\eta = (\eta_t, t \geq 0)$, где $\eta_t(\omega) = 1_\delta(t, \omega)$. Тогда η является $(\mathcal{F}_t)$ -адаптированным и $\mathcal{B}^+ \times \mathcal{F}$ -измеримым процессом, но, как нетрудно заметить, η не обладает свойством прогрессивной измеримости.

Наконец, сформулируем следующий результат: пусть на стохастическом базисе $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ задан случайный процесс $X = (X_t, t \geq 0)$, который адаптирован, непрерывен справа, имеет пределы слева и принимает значения в метрическом пространстве $(E, \mathcal{E})$. Тогда

X является прогрессивно измеримым. По поводу этого результата и ряда других результатов можно обратиться, например, к книгам [32, 120, 199, 206, 209, 480].

19.5. О стохастической непрерывности случайных процессов

Пусть $X = (X_t, t \in T)$ — случайный процесс, где T — интервал (конечный или бесконечный) в $\mathbb{R}^+$. Мы говорим, что X *стохастически непрерывен* (или P -непрерывен) в точке $t_0 \in T$, если для $t, t_0 \in T$ и $t \rightarrow t_0$ имеет место сходимость $X_t \xrightarrow{P} X_{t_0}$. Если это выполнено для всех $t_0 \in T$, то процесс X называется стохастически непрерывным в T .

С помощью двух примеров проиллюстрируем введенное свойство, а также его связь с непрерывностью п. н.

(а) Предположим, что случайный процесс $X = (X_t, t \in [0, 1])$ состоит из независимых и одинаково распределенных случайных величин, каждая с плотностью $g(x)$, $x \in \mathbb{R}^1$. Возьмем два момента времени: t и t_0 ($t, t_0 \in [0, 1]$), причем $t \neq t_0$, и пусть $\varepsilon > 0$ — произвольное число. Тогда

$$P\{|X_t - X_{t_0}| \geq \varepsilon\} = \iint_{|x-y| \geq \varepsilon} g(x)g(y) dx dy.$$

Легко видеть, что последний интеграл стремится к 1 при $\varepsilon \rightarrow 0$. Отсюда вытекает, что при каждом достаточно малом $\varepsilon > 0$

$$P\{|X_t - X_{t_0}| \geq \varepsilon\} \rightarrow 0 \quad \text{при} \quad t \rightarrow t_0,$$

и это имеет место для каждого $t_0 \in [0, 1]$. Следовательно, процесс X не является стохастически непрерывным.

(б) Пусть $\Omega = [0, 1]$, $\mathcal{F} = \mathcal{B}_{[0,1]}$ и P — мера Лебега. Случайный процесс $X = (X_t, t \in [0, 1])$ задан на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ так:

$$X_t = X_t(\omega) = \begin{cases} 1, & \text{если } t > \omega, \\ 0, & \text{если } t \leq \omega. \end{cases}$$

Процесс X имеет своим пространством состояний множество $E = \{0, 1\}$, и легко видеть, что конечномерные распределения X выражаются явно следующим образом:

$$P\{X_{t_1} = 0, \dots, X_{t_{k-1}} = 0, X_{t_k} = 1, \dots, X_{t_n} = 1\} = t_k - t_{k-1},$$

если $t_1 < t_2 < \dots < t_n$ и $1 < k \leq n$;

$$P\{X_{t_1} = 0, \dots, X_{t_n} = 0\} = 1 - t_n, \quad P\{X_{t_1} = 1, \dots, X_{t_n} = 1\} = t_1.$$

В остальных случаях $P\{X_{t_1} = j_1, \dots, X_{t_n} = j_n\} = 0$, где $j_k = 0$ или $j_k = 1$. Тогда для любого $\varepsilon \in (0, 1)$ и произвольных $t_1 < t_2$, находим, что

$$P\{|X_{t_2} - X_{t_1}| \geq \varepsilon\} = P\{X_{t_1} = 0, X_{t_2} = 1\} = t_2 - t_1.$$

Отсюда заключаем, что процесс X стохастически непрерывен на отрезке $[0, 1]$. Поскольку каждая траектория X — ступенчатая функция, то с вероятностью 1 процесс X не является непрерывным.

19.6. Критерий Колмогорова для непрерывности с вероятностью 1 случайного процесса

В начале настоящего параграфа мы привели определение непрерывности с вероятностью 1 случайного процесса. Сформулируем теперь один из самых полезных результатов, связанных с этим свойством.

Критерий Колмогорова. Пусть $X = (X_t, t \in T \subset \mathbb{R}^1)$ — случайный процесс, заданный на вероятностном пространстве $(\Omega, \mathcal{F}, P)$. Предположим, что существуют константы $K > 0$, $p > 0$ и $q > 0$ такие, что для всех $t, s \in T$ выполнено следующее условие (условие Колмогорова):

$$E\{|X_t - X_s|^p\} \leq K|t - s|^{1+q}. \quad (1)$$

Тогда процесс X непрерывен с вероятностью 1. (Это иногда формулируется и так: существует процесс $\tilde{X} = (X_t, t \in T)$, являющийся модификацией X , причем $\tilde{X}$ непрерывен п. н.)

Проиллюстрируем на примерах связь свойства непрерывности случайного процесса и выполнения условия Колмогорова.

(а) Рассмотрим пуассоновский процесс $N = (N_t, t \geq 0)$ с заданным параметром λ ($\lambda > 0$). Напомним, что $N_0 = 0$ п. н., приращения N на непересекающихся интервалах времени независимы, причем $N_t - N_s$ при $t > s$ имеет пуассоновское распределение с параметром $\lambda(t - s)$, т. е.

$$P\{N_t - N_s = k\} = \frac{e^{-\lambda(t-s)}[\lambda(t-s)]^k}{k!}, \quad k = 0, 1, 2, \dots$$

Легко проверить, что для этого процесса условие Колмогорова (1) не выполнено. Отсюда, однако, нельзя сделать вывод, что процесс N не является непрерывным, так как (1) — это достаточное условие для непрерывности. Можно показать, что почти все траектории пуассоновского процесса N являются неубывающими ступенчатыми функциями с единичными скачками, происходящими в случайные моменты времени (см. [14, 32]). Кроме того, пуассоновский процесс N имеет модификацию, непрерывную справа. И, наконец, из определения процесса N сразу вытекает, что $N_t \xrightarrow{P} N_{t_0}$ при $t \rightarrow t_0$ для каждого $t_0 \geq 0$.

Итак, пуассоновский процесс стохастически непрерывен, непрерывен справа, но не непрерывен с вероятностью 1. Условие Колмогорова для него не выполнено.

(б) Пусть $w = (w_t, t \geq 0)$ — стандартный винеровский процесс. Напомним, что $w_0 = 0$ п. н., приращения w независимы, причем $w_t - w_s$ имеет распределение $\mathcal{N}(0, |t - s|)$.

Непосредственно проверяется, что условие Колмогорова (1) выполнено, например, при $p = 4$ и $q = 1$. Следовательно, винеровский процесс w непрерывен с вероятностью 1.

Рассмотрим теперь другой процесс: $\xi = (\xi_t, t \geq 0)$, где $\xi_t = e^{w_t^3}$. Этот процесс также непрерывен с вероятностью 1. Однако при каждом $p > 0$ математическое ожидание $E\{|\xi_t - \xi_s|^p\}$ не существует, так что условие (1) не может выполняться. Таким образом, мы построили непрерывный с вероятностью 1 процесс ξ , для которого условие Колмогорова (1) не выполнено. Другими словами, условие Колмогорова является достаточным, но не необходимым для непрерывности с вероятностью 1 случайных процессов.

19.7. Существует ли связь между непрерывностью случайного процесса и непрерывностью порожденной им фильтрации?

Пусть на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ задан случайный процесс $X = (X_t, t \geq 0)$. Обозначим $\mathcal{F}_t^X = \sigma\{X_s, s \leq t\}$, т. е. $\mathcal{F}_t^X$ — это минимальная σ -алгебра, порожденная процессом X до момента t . Очевидно, что $\mathcal{F}_t^X \subset \mathcal{F}$ для каждого $t \geq 0$ и $\mathcal{F}_s^X \subset \mathcal{F}_t^X$, если $s < t$, т. е. семейство $(\mathcal{F}_t^X)_{t \geq 0}$ является фильтрацией. Естественно говорить также, что $(\mathcal{F}_t^X)_{t \geq 0}$ — это фильтрация, порожденная самим процессом X . Можно считать, что все $\mathcal{F}_t^X$ пополнены всеми множествами из Ω , имеющими P -меру нуль.

Для ряда задач важна непрерывность фильтрации $(\mathcal{F}_t^X)$, и интуитивно можно предположить, что такое свойство фильтрации связано с регулярностью самого процесса X . Ответ на вопрос «Так ли это?» содержится в рассмотренных ниже примерах.

(а) Пусть $w = (w_t, t \geq 0)$ — винеровский процесс. Тогда порожденная им фильтрация $(\mathcal{F}_t^w)_{t \geq 0}$ непрерывна (см. [73, 212]), т. е. из непрерывности винеровского процесса вытекает непрерывность порожденной им фильтрации.

Рассмотрим теперь вероятностное пространство $(\Omega, \mathcal{F}, P)$, где $\Omega = \mathbb{R}^1$, $\mathcal{F} = \mathcal{B}^1$ и P — любая вероятностная мера на $\mathcal{B}^1$. Пусть ξ — случайная величина с нормальным распределением $\mathcal{N}(0, 1)$. Положим $X_t = X_t(\omega) = t\xi(\omega)$, определяя таким образом случайный процесс $X = (X_t,$

$t \geq 0$). Легко видеть, что процесс X непрерывен с вероятностью 1. А что можно сказать о непрерывности его фильтрации $(\mathcal{F}_t^X)_{t \geq 0}$? Заметим, что $\mathcal{F}_0^X$ при $t = 0$ есть тривиальная σ -алгебра $\{\emptyset, \Omega\}$. Если же $t > 0$, то $\mathcal{F}_t^X = \mathcal{B}^1$. Таким образом, $\mathcal{F}_0^X \neq \mathcal{F}_{0+}^X$. Следовательно фильтрация $\{\mathcal{F}_t^X\}$, не являясь непрерывной справа, не может быть непрерывной. Итак, в этом случае из непрерывности процесса X не вытекает непрерывность порожденной им фильтрации $(\mathcal{F}_t^X)$.

(б) Пусть $\Omega = [0, 1]$, $\mathcal{F} = \mathcal{B}^1$ и P — мера Лебега. Возьмем функцию $h \in C^\infty(\mathbb{R}^+)$ такую, что $h(x) = 0$ при $x \leq \frac{1}{2}$ и $h(x) > 0$ при $x > \frac{1}{2}$, причем h строго возрастает в интервале $\left[\frac{1}{2}, \infty\right)$. (Легко привести примеры таких функций.) Рассмотрим случайный процесс $X = (X_t, t \geq 0)$, где $X_t = X_t(\omega) = \omega h(t)$, $\omega \in \Omega$, $t \geq 0$, и пусть $(\mathcal{F}_t^X)$ — порожденная им фильтрация, т. е. $\mathcal{F}_t^X = \sigma\{X_s, s \leq t\}$. Тогда

$$\mathcal{F}_t^X = \begin{cases} \{\emptyset, \Omega\}, & \text{если } 0 \leq t \leq \frac{1}{2}, \\ \mathcal{B}^1, & \text{если } t > \frac{1}{2}. \end{cases}$$

Таким образом, фильтрация $(\mathcal{F}_t^X)$ разрывна несмотря на то, что траектории процесса X принадлежат пространству $C^\infty(\mathbb{R}^+)$.

(в) Покажем теперь, что фильтрация $(\mathcal{F}_t^X)_{t \geq 0}$ может быть непрерывной, даже если процесс X имеет сколь угодно «плохие» траектории, например разрывные. Для этой цели мы воспользуемся следующим утверждением: *пусть $h: \mathbb{R}^+ \mapsto \mathbb{R}^1$ — любая функция. Тогда существует счетное и всюду плотное в $\mathbb{R}^+$ подмножество D такое, что для каждого $t \geq 0$ имеется последовательность $\{t_n, n \geq 1\}$ в D , для которой $t_n \rightarrow t$ и $h(t_n) \rightarrow h(t)$ при $n \rightarrow \infty$. (Очевидно, здесь идет речь о свойстве сепарабельности.)* Рассмотрим одноточечное вероятностное пространство $(\Omega, \mathcal{F}, P)$. Положим $X_t(\omega) = h(t)$ для всех $\omega \in \Omega$ и $t \geq 0$. Поскольку расширенная прямая $\overline{\mathbb{R}^1} = [-\infty, \infty]$ есть компакт, то из общих теорем о сепарабельности (см. [21, 22, 32, 129, 206]) вытекает, что у процесса $X = (X_t, t \geq 0)$ имеется сепарабельная модификация $Y = (Y_t, t \geq 0)$, где $Y_t: \Omega \mapsto \overline{\mathbb{R}^1}$, $t \geq 0$. Однако $Y_t = X_t$ почти наверное, так что $Y_t(\omega) = h(t)$, $t \geq 0$.

Таким образом, можно строить сепарабельные процессы, траектории которых обладают или не обладают любым свойством. Возьмем, например, сепарабельную, но разрывную функцию $h(t)$, $t \geq 0$, и пусть снова $(\Omega, \mathcal{F}, P)$ — одноточечное вероятностное пространство. Рассмотрим случайный процесс $X = (X_t, t \geq 0)$, где $X_t = X_t(\omega) = h(t)$, $\omega \in \Omega$, $t \geq 0$. Тогда для каждого $t \geq 0$ σ -алгебра $\mathcal{F}_t^X = \sigma\{X_s, s \leq t\}$ являет-

ся $\mathcal{P}$ -тривиальной в том смысле, что для любого события $A \in \mathcal{F}_t^X$ либо $\mathcal{P}(A) = 0$, либо $\mathcal{P}(A) = 1$. Отсюда следует, что фильтрация $(\mathcal{F}_t^X)$ непрерывна, несмотря на то, что порождающий ее процесс X разрывен.

Итак, примеры (а), (б) и (в) показывают, что в общем случае не существует прямой связи между свойством непрерывности фильтрации $(\mathcal{F}_t^X)$ и свойствами регулярности процесса X .

Отметим, наконец, что ряд полезных результатов, касающихся свойств случайных процессов и свойств порожденных ими фильтраций, может быть найден, например, в [74, 199, 368, 408].

§ 20. Марковские процессы

Приведем здесь только самые основные определения, касающиеся марковских процессов. Другие будут даны в самих примерах. Читатель, конечно, может обратиться и к существующей литературе.

Пусть $X = (X_t, t \in T \subset \mathbb{R}^+)$ — случайный процесс, заданный на вероятностном пространстве $(\Omega, \mathcal{F}, \mathcal{P})$ и принимающий значения в некотором счетном множестве E . Мы говорим, что X — *марковская цепь*, если выполнено следующее условие, называемое *марковским свойством*: для любых $n \geq 1$, $t_1 < t_2 < \dots < t_n < t$, $t_j, t \in T$, $i_1, \dots, i_n, j \in E$

$$\begin{aligned} \mathcal{P}\{X_t = j \mid X_{t_1} = i_1, \dots, X_{t_{n-1}} = i_{n-1}, X_{t_n} = i_n\} = \\ = \mathcal{P}\{X_t = j \mid X_{t_n} = i_n\}. \end{aligned} \quad (1)$$

Марковская цепь X называется *конечной*, если множество E конечно, а в случае бесконечного (но счетного) E мы говорим, что E — *счетная марковская цепь*. Если $T = \{0, 1, 2, \dots\}$, то принято писать $X = (X_n, n \geq 0)$ или $X = (X_n, n = 0, 1, \dots)$. В этом случае X называется *марковской цепью с дискретным временем*. И наконец, если $T = \mathbb{R}^+$ или $T = [a, b] \subset \mathbb{R}^+$, то $X = (X_t, t \geq 0)$ или $X = (X_t, t \in [a, b])$ называется *марковской цепью с непрерывным временем*.

Вероятностные характеристики любой марковской цепи могут быть найдены, если мы знаем *начальное распределение* $(r_j, j \in E)$, где $r_j = \mathcal{P}\{X_0 = j\}$, $r_j \geq 0$, $\sum_{j \in E} r_j = 1$, и *переходные вероятности* $p_{ij}(s, t) = \mathcal{P}\{X_t = j \mid X_s = i\}$, $s \leq t$. Цепь X называется *однородной*, если $p_{ij}(s, t)$ зависят от s и t только через разность $t - s$. В этом случае для полного описания марковской цепи с дискретным временем нам достаточно знать начальное распределение $(r_j, j \in E)$ и матрицу переходных вероятностей за один шаг $P = (p_{ij})$: $p_{ij} = \mathcal{P}\{X_{n+1} = j \mid X_n = i\}$, $n \geq 0$. Вероятности перехода за n шагов $p_{ij}^{(n)} = \mathcal{P}\{X_n = j \mid X_0 = i\}$, $p_{ij}^{(1)} = p_{ij}$,

образуют матрицу $P^{(n)} = (p_{ij}^{(n)})$, удовлетворяющую следующему *уравнению Колмогорова–Чепмена*:

$$p_{ij}^{(m+n)} = \sum_{k \in E} p_{ik}^{(m)} p_{kj}^{(n)}. \quad (2)$$

Отметим, что переходные вероятности $p_{ij}(t)$ или $p_{ij}(s, t)$ любой марковской цепи с непрерывным временем удовлетворяют так называемым прямым и обратным уравнениям Колмогорова.

Рассмотрим более общую ситуацию. Пусть $X = (X_t, t \geq 0)$ — случайный процесс, заданный на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ и принимающий значения в фазовом пространстве $(\mathbb{R}^1, \mathcal{B}^1)$. Пусть $(\mathcal{F}_t)_{t \geq 0}$ — фильтрация, порожденная процессом X . Мы говорим, что X — *марковский процесс*, если выполнено следующее условие, называемое *марковским свойством* (см. также (1)): для любых $\Gamma \in \mathcal{B}^1$ и $t \geq s$

$$P\{X_t \in \Gamma \mid \mathcal{F}_s\} = P\{X_t \in \Gamma \mid X_s\} \quad (P\text{-п.н.}) \quad (3)$$

Это свойство может быть записано в другой эквивалентной форме.

Функция $P(s, x; t, \Gamma)$, определенная для $s, t \in \mathbb{R}^+$, $s \leq t$, $x \in \mathbb{R}^1$, $\Gamma \in \mathcal{B}^1$, называется *переходной функцией*, если:

(а) при фиксированных s, t и x $P(s, x; t, \cdot)$ — вероятностная мера на σ -алгебре $\mathcal{B}^1$;

(б) $P(s, x; t, \Gamma)$ является $\mathcal{B}^1$ -измеримой по x при фиксированных s, t и Γ ;

(в) $P(s, x; s, \Gamma) = \delta_x(\Gamma)$ — единичная мера, сосредоточенная в точке x ;

(г) имеет место *уравнение Колмогорова–Чепмена* (ср. с соотношением (2)):

$$P(s, x; t, \Gamma) = \int_{-\infty}^{\infty} P(s, x; u, dy) P(u, y; t, \Gamma), \quad s \leq u \leq t. \quad (4)$$

Мы говорим, что $X(X_t, t \geq 0)$ — *марковский процесс с переходной функцией* $P(s, x; t, \Gamma)$, если X удовлетворяет (3) и, кроме того,

$$P\{X_t \in \Gamma \mid X_s\} = P\{s, X_s; t, \Gamma\} \quad (P\text{-п.н.}).$$

Марковский процесс X называется *однородным*, если его переходная функция $P(s, x; t, \Gamma)$ зависит от s и t через $t - s$. Для нахождения вероятностных характеристик марковского процесса X достаточно знать начальное распределение $\{P_0(\Gamma), \Gamma \in \mathcal{B}^1\}$, где $P_0(\Gamma) = P\{X_0 \in \Gamma\}$, и переходную функцию $P(s, x; t, \Gamma)$ (в неоднородном случае) или $P(t, x, \Gamma)$ (в однородном случае).

К настоящему времени имеется хорошо развитая теория марковских цепей и марковских процессов. Читатель может обратиться, например,

к следующим книгам: [14, 21, 22, 32, 33, 34, 88, 106, 111, 151, 180, 290, 293, 420].

В этот параграф мы включили примеры, иллюстрирующие основные свойства марковских цепей и процессов.

20.1. Уравнение Колмогорова–Чепмена может иметь место, даже если процесс немарковский

При определении марковского процесса мы требуем выполнения марковского свойства (1) или (3), а затем как следствие выводим уравнения Колмогорова–Чепмена (2) и (4). Поэтому имеет смысл поставить вопрос: существуют ли немарковские процессы такие, что их переходные вероятности удовлетворяют уравнению Колмогорова–Чепмена? Ответ следует из рассмотренного далее примера.

Предположим, что урна содержит четыре шара, занумерованные числами 1, 2, 3, 4. Случайным образом выбираем один шар, записываем его номер и возвращаем шар обратно в урну. Эту процедуру повторяем много раз. Обозначим через ξ_n число на n -м выбранном шаре. При $j = 1, 2, 3$ вводим события $A_j^{(n)} = \{\text{либо } \xi_n = j, \text{ либо } \xi_n = 4\}$, и пусть $X_{3(m-1)+j} = 1$, если наступает $A_j^{(m)}$, $m = 1, 2, \dots$; иначе мы полагаем $X_{3(m-1)+j} = 0$, т. е. $X_{3(m-1)+j} = I(A_j^{(m)})$ — индикатор события $A_j^{(m)}$. Таким образом, мы определили случайную последовательность $(X_n, n \geq 1)$ и хотим изучить ее свойства.

Легко заметить, что

$$P\{X_n = k_1\} = P\{X_n = k_2 \mid X_m = k_3\} = \frac{1}{2}, \quad n > m,$$

где каждое из k_1, k_2, k_3 принимает значение 1 или 0. Следовательно, при $i < m < n$ имеют место равенства

$$\begin{aligned} \frac{1}{2} &= P\{X_n = k_2 \mid X_i = k_1\} = P\{X_n = k_2 \mid X_m = 0\}P\{X_m = 0 \mid X_i = k_1\} + \\ &+ P\{X_n = k_2 \mid X_m = 1\}P\{X_m = 1 \mid X_i = k_1\} = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2}. \end{aligned}$$

Отсюда вытекает, что переходные вероятности для последовательности $(X_n, n \geq 1)$ удовлетворяют уравнению Колмогорова–Чепмена.

Далее, событие $\{X_{3m-2} = 1, X_{3m-1} = 1\}$ означает, что $\xi_m = 4$, а тогда очевидно, что $X_{3m} = 1$. Таким образом,

$$P\{X_{3m} = 1 \mid X_{3m-2} = 1, X_{3m-1} = 1\} = 1, \quad m = 1, 2, \dots$$

Из этого соотношения следует, что марковское свойство не имеет места для последовательности $(X_n, n \geq 1)$.

20.2. Об одном следствии марковского свойства

Пусть на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ задан марковский процесс $X = (X_t, t \geq 0)$ со значениями в $\mathbb{R}^1$, и пусть $E|X_t| < \infty$ для каждого $t \geq 0$. Возьмем три момента времени: s, t и u , причем $s < t < u$, и назовем t «настоящим» моментом. Тогда σ -алгебры $\mathcal{H}_s = \sigma\{X_v, v \leq s\}$ и $\mathcal{F}_u = \sigma\{X_v, v \geq u\}$ можно назвать «прошлым» и «будущим» процесса X по отношению к моменту t . Мы ввели марковское свойство X с помощью равенства

$$P\{X_t \in \Gamma \mid \mathcal{H}_s\} = P\{X_t \in \Gamma \mid X_s\} \quad (P\text{-п.н.})$$

Оказывается, это свойство можно задать и в следующей эквивалентной форме, в которой «прошлое» и «будущее» симметричны:

$$E\{X_t \mid \mathcal{H}_s \vee \mathcal{F}_u\} = E\{X_t \mid X_s, X_u\} \quad (P\text{-п.н.}). \quad (1)$$

Далее из (1) нетрудно вывести такое следствие:

$$E\{X_t \mid \mathcal{H}_s \vee \sigma(X_u)\} = E\{X_t \mid X_s, X_u\} \quad (P\text{-п.н.}). \quad (2)$$

Выясним теперь, сохраняется ли марковское свойство (2) при небольшом изменении информации, т.е. тех условий, относительно которых вычисляются математические ожидания в (2). Для этой цели рассмотрим винеровский процесс $w = (w_t, t \geq 0)$ и зафиксируем два момента времени s и t , причем $1 < s < t < 2$. Через $\mathcal{H}_s^w = \sigma\{w_v, v \leq s\}$ обозначим «прошлое» процесса w по отношению к настоящему моменту t , и, как обычно, $\sigma(w_1 + w_2)$ — это σ -алгебра, порожденная величиной $w_1 + w_2$. Сравним условные математические ожидания $E\{w_t \mid \mathcal{H}_s^w \vee \sigma(w_1 + w_2)\}$ и $E\{w_t \mid w_s, w_1 + w_2\}$. Вообще говоря, соотношение (2) дает основания предположить, что эти ожидания P -п.н. совпадают. Проверим, так ли это.

С одной стороны, поскольку w — марковский процесс, при выбранных моментах времени $1 < s < t < 2$ имеем P -п.н.

$$E\{w_t \mid \mathcal{H}_s^w \vee \sigma(w_1 + w_2)\} = E\{w_t \mid \mathcal{H}_s^w \vee \sigma(w_2)\} = E\{w_t \mid w_s, w_2\}.$$

(Для второго равенства мы воспользовались соотношением (2).) Ввиду гауссовости процесса w нетрудно подсчитать, что

$$E\{w_t \mid w_s, w_2\} = \frac{(2-t)w_s + (t-s)w_2}{2-s}.$$

С другой стороны, снова используя гауссовость w , находим

$$E\{w_t \mid w_s, w_1 + w_2\} = \frac{(1+s)(1+t)w_s + s(1+t)(w_1 + w_2)}{(1+s)^2 + 5s}.$$

Таким образом,

$$\mathbf{E}\{w_t \mid \mathcal{H}_s^w \vee \sigma(w_1 + w_2)\} \neq \mathbf{E}\{w_t \mid w_s, w_1 + w_2\}.$$

Следовательно, марковское свойство, выраженное равенством (2), будет иметь место, если новая информация, добавляемая к «прошлому» $\mathcal{H}_s^w$, берется только из «будущего».

20.3. Связь между различными понятиями эргодичности для марковских цепей

Рассмотрим марковскую цепь $X = (X_n, n = 0, 1, \dots)$ со значениями в не более чем счетном множестве E , заданную своим начальным распределением $f^{(0)} = (f_j^{(0)}, j \in E)$ и последовательностью матриц переходных вероятностей $(P_n, n \geq 1)$, где $P_n = (p_{ij}(n-1, n))$.

Если пределы $\lim_{n \rightarrow \infty} p_{ij}(k, k+n) = \pi_j$ существуют для всех $j \in E$ и не зависят от i и k , $i \in E, k = 0, 1, \dots$, причем $\pi_j > 0$ и $\sum_{j \in E} \pi_j = 1$, цепь X называется *эргодической*, а $(\pi_j, j \in E)$ — ее *эргодическим распределением*.

Далее используем следующие обозначения:

$$\begin{aligned} f^{(m)} &= f^{(0)} P_1 P_2 \dots P_m, \\ f^{(k,m)} &= f^{(0)} P_{k+1} P_{k+2} \dots P_m, \\ P^{(k,m)} &= P_{k+1} P_{k+2} \dots P_m. \end{aligned}$$

Марковская цепь X называется *w-эргодической* (*эргодической в слабом смысле*), если для всех $k = 0, 1, \dots$

$$\lim_{m \rightarrow \infty} \sup_{f^{(0)}, g^{(0)}} \|f^{(k,m)} - g^{(k,m)}\| = 0, \quad (1)$$

где $f^{(0)} = (f_j^{(0)}, j \in E)$ и $g^{(0)} = (g_j^{(0)}, j \in E)$ — произвольные начальные распределения X .

Марковская цепь X называется *s-эргодической* (*эргодической в сильном смысле*), если существует вероятностное распределение $q = (q_j, j \in E)$ такое, что для всех $k = 0, 1, \dots$

$$\lim_{m \rightarrow \infty} \sup_{f^{(0)}} \|f^{(k,m)} - q\| = 0. \quad (2)$$

В (1) и (2) $\|x\| = \sum_{j \in E} |x_j|$ — норма вектора $x = (x_j, j \in E)$.

При изучении свойства эргодичности важную роль играет так называемый δ -коэффициент

$$\delta(P) = 1 - \inf_{i,k} \sum_{j \in E} \min\{p_{ij}, p_{kj}\},$$

где $P = (p_{ij})$ — матрица переходных вероятностей.

Выясним, какова связь между введенными понятиями эргодичности. Очевидно, например, что из s -эргодичности вытекает w -эргодичность. Таким образом, есть вопрос: справедливо ли обратное утверждение? Можно показать (см. [293]), что если множество E конечно, то эргодичность и w -эргодичность — эквивалентные понятия. Остается вопрос: эквивалентны ли эти понятия, если E — бесконечное множество? Ответы содержатся в рассмотренных ниже примерах.

(а) Пусть $X = (X_n, n \geq 0)$ — неоднородная марковская цепь со следующими матрицами переходных вероятностей:

$$P_{2n-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{pmatrix}, \quad P_{2n} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad n = 1, 2, \dots$$

Легко видеть, что $\delta(P_{2n}) = 1$ и $\delta(P_{2n-1}) = \frac{1}{2}$. Следовательно, при любом $k = 0, 1, \dots$

$$\delta(P^{(k,m)}) \leq \prod_{j=k+1}^m \delta(P_j) \leq \left(\frac{1}{2}\right)^{[(m-k)/2]} \rightarrow 0 \quad \text{при } m \rightarrow \infty.$$

Однако условие $\delta(P^{(k,m)}) \rightarrow 0$ при $m \rightarrow \infty$ для всех k является необходимым и достаточным для w -эргодичности марковской цепи X (см. [293]). Итак, цепь X — w -эргодична. Проверим, обладает ли цепь X свойством s -эргодичности.

Возьмем, например, начальное распределение $f^{(0)} = (0, 1)$. Тогда

$$f^{(2k)} = f^{(0)}(P_1 P_2)(P_3 P_4) \dots (P_{2k-1} P_{2k}) = (0, 1) \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{pmatrix}^k = (0, 1),$$

$$f^{(2k+1)} = f^{(0)}(P_1 P_2)(P_3 P_4) \dots (P_{2k-1} P_{2k}) P_{2k+1} = f^{(2k)} P_{2k+1} = (1, 0).$$

Отсюда вытекает, что

$$\|f^{(2m)} - f^{(2m+1)}\| = 2$$

для $m = k, k+1, \dots$, последовательность $(f^{(k)})$ расходится и соотношение (2) не может иметь места. Значит, цепь X не является s -эргодической.

(б) Пусть $X = (X_n)$ — стационарная марковская цепь с бесконечным множеством значений E и переходной матрицей

$$\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & \dots \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & \dots \\ 0 & \frac{3}{4} & 0 & \frac{1}{4} & 0 & \dots \\ 0 & 0 & \frac{7}{8} & 0 & \frac{1}{8} & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}.$$

Оказывается, цепь X неразложимая, причем все множество E образует один возвратный и положительный класс. Тогда (см. [32, 111, 116]) эта цепь эргодическая, т. е. при любом начальном распределении $f^{(0)} = (f_j^{(0)}, j \in E)$ существует предел $\lim_{n \rightarrow \infty} f^{(0)} P^{(n)} = \pi$, где $\pi = (\pi_j, j \in E)$ — вероятностное распределение. Однако $\delta(P^{(n)}) = 1$ для каждого n . Отсюда следует, что цепь не является w -эргодической, так как для этого свойства необходимо, чтобы $\delta(P^{(n)}) \rightarrow 0$ при $n \rightarrow \infty$.

20.4. Эргодичность марковской цепи не всегда обеспечивает сходимость математического ожидания функции от этой цепи

Пусть $X = (X_n, n = 0, 1, \dots)$ — однородная марковская цепь со счетным множеством состояний E и переходной матрицей за n шагов $(p_{ij}^{(n)})$. Предположим, что эта цепь эргодическая и $\pi_j = \lim_{n \rightarrow \infty} p_{ij}^{(n)}, j \in E$ — ее эргодическое распределение. Нас интересует следующий вопрос: при каких условиях на цепь X и на ограниченную и измеримую функцию $g: E \rightarrow \mathbb{R}^1$ будет иметь место соотношение

$$\lim_{n \rightarrow \infty} \mathbf{E}g(X_n) = \sum_{j \in E} g(j)\pi_j. \quad (1)$$

Сформулируем один из имеющихся результатов (см. [285]): пусть марковская цепь X неразложима, аperiodична и ее множество состояний E образует возвратный и положительный класс. Обозначим через $(\pi_j, j \in E)$ эргодическое распределение X . Предположим, что неотрицательная функция g такова, что выполнено условие $\sum_{j \in E} g(j)\pi_j < \infty$. Пусть, наконец, $\mathbf{P}\{X_0 = i_0\} = 1$ для некоторого фиксированного $i_0 \in E$. Тогда имеет место соотношение (1).

Заметим, что для справедливости этого результата требуется, чтобы цепь X начинала свою эволюцию в момент $n = 0$ с некоторого

фиксированного состояния i_0 . С другой стороны, сама цепь X эргодическая, и, значит, ее эргодическое распределение $(\pi_j, j \in E)$ не зависит от начального состояния. Вопрос: насколько существенно условие $X_0 = i_0$ и, еще более конкретно, можно ли заменить это условие предположением, что в момент $n = 0$ величина X_0 имеет невырожденное распределение на всем множестве E ? Покажем, что это не всегда так.

Рассмотрим однородную марковскую цепь $\xi = (\xi_n, n = 0, 1, \dots)$ с множеством состояний $E = \{0, 1, 2, \dots\}$ и переходными вероятностями: $p_{0j} = q^j p$ для $j = 0, 1, 2, \dots$, $p_{i,i-1} = 1$ для $i = 1, 2, \dots$ и $p_{ij} = 0$ в остальных случаях. Здесь $0 < p < 1$, а $q = 1 - p$. Непосредственным вычислением находим вероятности перехода $p_{ij}^{(n)}$ за n шагов. Именно: $p_{ij}^{(n)} = q^j p$ для $i = 0, 1, \dots, n-1$ и $j = 0, 1, 2, \dots$; $p_{n-1+k, k-1}^{(n)} = 1$ для $k = 1, 2, \dots$; $p_{ij}^{(n)} = 0$ в остальных случаях. Поскольку цепь ξ неразложима, апериодична и ее множество состояний E образует возвратный и положительный класс, то ξ эргодична. Легко видеть, что эргодическое распределение ξ задается следующим образом:

$$\pi_j = \lim_{n \rightarrow \infty} p_{ij}^{(n)} = q^j p, \quad j \in E.$$

Возьмем теперь любую функцию $g: E \mapsto \mathbb{R}^1$, удовлетворяющую условию

$$\sum_{j=0}^{\infty} q^j |g(j)| < \infty,$$

и предположим, что цепь ξ имеет начальное распределение $(r_j, j \in E)$. Тогда

$$\mathbb{E}g(\xi_n) = \sum_{j=0}^{\infty} \left(\sum_{i=0}^{\infty} r_i p_{ij}^{(n)} \right) g(j) = (r_0 + \dots + r_{n-1}) \sum_{j=0}^{\infty} p q^j g(j) + \sum_{j=0}^{\infty} r_{n+j} g(j).$$

Отсюда вытекает, что $\mathbb{E}g(\xi_n)$ будет сходиться к $\sum_{j=0}^{\infty} p q^j g(j)$ при $n \rightarrow \infty$ тогда и только тогда, когда

$$\lim_{n \rightarrow \infty} \sum_{j=0}^{\infty} r_{n+j} g(j) = 0. \quad (2)$$

Пусть выбор функции g и начального распределения $(r_j, j \in E)$ такой:

$$g(j) = j^2 \quad \text{для} \quad j = 0, 1, 2, \dots \quad \text{и}$$

$$r_j = \begin{cases} 0, & \text{если } j = 0, \\ \frac{6}{\pi^2 j^2}, & \text{если } j = 1, 2, \dots \end{cases}$$

Тогда очевидно, что $\sum_{j=0}^{\infty} r_{n+j}g(j) = \infty$ для каждого $n \geq 0$. Этот факт и условие (2) показывают, что соотношение (1) не имеет место для рассматриваемой цепи ξ . Другими словами, в приведенном выше результате нельзя ослабить условие « $X_0 = i_0$ », заменяя его, например, собственным распределением X_0 на всем множестве E .

20.5. Две неэквивалентные марковские цепи с непрерывным временем могут быть частично эквивалентными

Пусть $X = (X_t, t \geq 0)$ и $\tilde{X} = (\tilde{X}_t, t \geq 0)$ — однородные марковские цепи с непрерывным временем и с одним и тем же множеством состояний E . Предположим, что X и $\tilde{X}$ имеют одно и то же начальное распределение и переходные вероятности $p_{ij}(t) = P\{X_t = j \mid X_0 = i\}$, $\tilde{p}_{ij}(t) = P\{\tilde{X}_t = j \mid \tilde{X}_0 = i\}$, $t \geq 0$, $i, j \in E$. Мы говорим, что X и $\tilde{X}$ *частично эквивалентны*, если равенства $p_{ij}(t) = \tilde{p}_{ij}(t)$, $i, j \in E$, выполнены для бесконечно многих t , но не для всех $t \geq 0$. Если же $p_{ij}(t) = \tilde{p}_{ij}(t)$, $i, j \in E$, для всех $t \geq 0$, то процессы X и $\tilde{X}$ эквивалентны в том смысле, что каждый из них является модификацией другого (см. пример 19.2). Понятно, что из эквивалентности вытекает частичная эквивалентность. Поэтому полезно описать пример пары частично эквивалентных процессов X и $\tilde{X}$, которые не являются эквивалентными.

Отметим сначала, что переходные вероятности любой марковской цепи с непрерывным временем удовлетворяют двум системам дифференциальных уравнений, называемых уравнениями Колмогорова (см., например, [21, 22, 111]). Эти уравнения записываются в терминах соответствующей инфинитезимальной матрицы $Q = (q_{ij})$, и при некоторых естественных предположениях переходные вероятности $p_{ij}(s, t)$, $0 \leq s \leq t$ или $p_{ij}(t)$, $t \geq 0$, являются их единственным решением.

Итак, пусть X и $\tilde{X}$ — однородные марковские цепи со значениями в множестве $E = \{1, 2, 3\}$, причем X и $\tilde{X}$ заданы своими инфинитезимальными матрицами $Q = (q_{ij})$ и $\tilde{Q} = (\tilde{q}_{ij})$, где

$$Q = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{pmatrix}, \quad \tilde{Q} = \begin{pmatrix} -1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -1 \end{pmatrix}.$$

Зная Q и $\tilde{Q}$ и используя уравнения Колмогорова, можно найти следующие явные выражения для переходных вероятностей $p_{ij}(t)$ и $\tilde{p}_{ij}(t)$

(детали предоставляем читателю):

$$\begin{aligned}
 p_{11}(t) &= p_{22}(t) = p_{33}(t) = \frac{1}{3} + \frac{2}{3} e^{-3t/2} \cos\left(\frac{t\sqrt{3}}{2}\right), \\
 p_{12}(t) &= p_{23}(t) = p_{31}(t) = \frac{1}{3} + \frac{2}{3} e^{-3t/2} \cos\left(\frac{t\sqrt{3}}{2} - \frac{2\pi}{3}\right), \\
 p_{13}(t) &= p_{21}(t) = p_{32}(t) = \frac{1}{3} + \frac{2}{3} e^{-3t/2} \cos\left(\frac{t\sqrt{3}}{2} + \frac{2\pi}{3}\right), \\
 \tilde{p}_{11}(t) &= \tilde{p}_{22}(t) = \tilde{p}_{33}(t) = \frac{1}{3} + \frac{2}{3} e^{-3t}, \\
 \tilde{p}_{ij} &= \frac{1}{3} - \frac{1}{3} e^{-3t/2}, \quad \text{если } i \neq j, \quad i, j = 1, 2, 3.
 \end{aligned}$$

Легко проверить, что равенства $p_{ij}(t) = \tilde{p}_{ij}(t)$, $i, j = 1, 2, 3$, выполнены для каждого t вида $t = \frac{4k\pi}{\sqrt{3}}$, где $k \in \mathbb{N}$. Однако для всех других t мы имеем $p_{ij}(t) \neq \tilde{p}_{ij}(t)$. Следовательно, марковские цепи X и $\tilde{X}$ не эквивалентны, но являются частично эквивалентными.

20.6. Марковские процессы, феллеровские процессы, сильно феллеровские процессы и связи между ними

Пусть $X = (X_t, t \geq s, \mathbb{P}_{s,x})$ — марковское семейство, т. е. $(X_t, t \geq s)$ есть марковский процесс относительно вероятностной меры $\mathbb{P}_{s,x}$, причем $\mathbb{P}_{s,x}\{X_s = x\} = 1$, и пусть $P(s, x; t, \Gamma)$, $s \leq t$, $x \in \mathbb{R}^1$, $\Gamma \in \mathcal{B}^1$, — переходная функция X . Как обычно, множество ограниченных и измеримых функций на $\mathbb{R}^1$ обозначим через $\mathbb{B} = \mathbb{B}(\mathbb{R}^1)$, а $\mathbb{C} = \mathbb{C}(\mathbb{R}^1)$ — обозначение для множества ограниченных и непрерывных функций на $\mathbb{R}^1$. Равенство

$$P^{st}g(x) = \mathbb{E}_{s,x}g(X_t) = \int_{\mathbb{R}^1} g(y)P(s, x; t, dy)$$

определяет на $\mathbb{B}$ полугруппу операторов $\{P^{st}\}$. Имеет место включение $P^{st}\mathbb{B} \subset \mathbb{B}$, и, более того, $P^{st}\mathbb{C} \subset \mathbb{B}$. Марковский процесс, для которого $P^{st}\mathbb{C} \subset \mathbb{C}$, называется *феллеровским процессом*. Другими словами, переходная функция $P(\cdot)$ такова, что если $g \in \mathbb{C}$, то и $P^{st}g \in \mathbb{C}$. Далее, если для каждой функции $g \in \mathbb{B}$ функция $P^{st}g(x)$ непрерывна по x , то марковский процесс X называется *сильно феллеровским*.

Легко видеть, что сильно феллеровский процесс является феллеровским. Полезно привести примеры марковского процесса, являющегося феллеровским, но не сильно феллеровским, и марковского процесса, не являющегося феллеровским.

(а) Пусть $X = (X_t, t \geq 0)$ — однородный марковский процесс, переходная функция которого задается следующим образом:

$$P(t, x, \Gamma) = I_\Gamma(x + vt), \quad t \geq 0, \quad x \in \mathbb{R}^1, \quad \Gamma \in \mathcal{B}^1, \quad v = \text{const} > 0$$

($I_\Gamma(\cdot)$ — индикатор множества Γ). Нетрудно сообразить, что P описывает движение (например, частицы) с постоянной скоростью v . Легко проверить, что процесс X феллеровский, но не сильно феллеровский.

(б) Предположим, что частица начинает свое движение в момент s из состояния x и что ее движение описывается семейством $X = (X_t, t \geq s, P_{s,x})$. При этом, если $X_s < 0$, частица движется влево со скоростью 1; если $X_s > 0$, она движется вправо со скоростью 1; если $X_s = 0$, то частица движется либо влево, либо вправо и каждое из этих направлений имеет вероятность $\frac{1}{2}$. Формально это можно выразить так:

$$\begin{aligned} P_{s,x}\{X_t = x + (t - s), t \geq s\} &= 1, & \text{если } x > 0, \\ P_{s,x}\{X_t = x - (t - s), t \geq s\} &= 1, & \text{если } x < 0, \\ P_{s,0}\{X_t = t - s, t \geq s\} &= P_{s,0}\{X_t = -(t - s), t \geq s\} = \frac{1}{2}. \end{aligned}$$

Легко видеть, что $X = (X_t, t \geq s, P_{s,x})$ — марковское семейство. Для любой функции $g \in \mathbb{B}$

$$P^{st}g(x) = \begin{cases} g(x + (t - s)), & \text{если } x > 0, \\ g(x - (t - s)), & \text{если } x < 0, \\ \frac{1}{2}g(t - s) + \frac{1}{2}g(-(t - s)), & \text{если } x = 0. \end{cases}$$

Заметим, что функция $P^{st}g(x)$ имеет разрыв в точке $x = 0$. Следовательно, марковский процесс X не является феллеровским.

20.7. Марковские, но не строго марковские процессы

Пусть заданы измеримое пространство $(\Omega, \mathcal{F})$ и на нем семейство вероятностных мер $(P_x, x \in \mathbb{R}^1)$ и случайный процесс $X = (X_t, t \geq 0)$ со значениями в $\mathbb{R}^1$. Пусть фильтрация $(\mathcal{F}_t)_{t \geq 0}$, порожденная процессом X , удовлетворяет обычным условиям. Предположим, что для процесса X выполнено марковское свойство относительно каждой из вероятностных мер P_x . Задана также однородная переходная функция $P(t, x, \Gamma)$, $t \geq 0, x \in \mathbb{R}^1, \Gamma \in \mathcal{B}^1$. Пара (X_t, P_x) называется *однородным марковским семейством с переходной функцией* $P(t, x, \Gamma)$, если

$$P_x\{X_{t+h} \in \Gamma \mid \mathcal{F}\} = P(h, X_t, \Gamma) \quad (P\text{-п.н}).$$

Пусть теперь τ — случайная величина, являющаяся моментом остановки относительно $(\mathcal{F}_t)$. Через $\mathcal{F}_\tau$ обозначим σ -алгебру, содержащую все события $A \in \mathcal{F}$ такие, что $A \cap \{\tau \leq t\} \in \mathcal{F}_t$ для всех $t \geq 0$. Предположим, что процесс X прогрессивно измерим, и пусть η — $\mathcal{F}_\tau$ -измеримая неотрицательная случайная величина, определенная на множестве $\{\omega: \tau(\omega) < \infty\}$. Мы говорим, что (X_t, P_x) — *однородное строгое марковское семейство*, если для любого $\Gamma \in \mathcal{B}^1$

$$P_x\{X_{\tau+\eta} \in \Gamma \mid \mathcal{F}_\tau\} = P_x\{X_{\tau+\eta} \in \Gamma \mid X_\tau\} \quad (P_x\text{-п.н.}).$$

Это соотношение называется *строгим марковским свойством*. Оно может быть записано через переходную функцию $P(t, x, \Gamma)$ так:

$$P_x\{X_{\tau+\eta} \in \Gamma \mid \mathcal{F}_\tau\} = P(\eta, X_\tau, \Gamma) \quad (P_x\text{-п.н.}). \quad (1)$$

Для дальнейшего рассуждения более удобно записать строгое марковское свойство (1) в следующей форме:

$$P_x\{A \cap [X_{\tau+\eta} \in \Gamma]\} = \int_A P(\eta, X_\tau, \Gamma) P_x(d\omega), \quad (2)$$

$$A \subset \mathcal{F} \cap \{\omega: \tau(\omega) < \infty, \eta(\omega) < \infty\}.$$

Перейдем теперь к рассмотрению двух примеров марковских процессов, которые не являются строгими марковскими.

(а) Пусть $w = (w_t, t \geq 0, P_x)$ — винеровский процесс, начинающийся из произвольной точки $x \in \mathbb{R}^1$. Введем процесс $X = (X_t, t \geq 0)$, где

$$X_t = \begin{cases} w_t, & \text{если } w_0 \neq 0, \\ 0, & \text{если } w_0 = 0. \end{cases}$$

Нетрудно видеть, что X — однородный марковский процесс, переходная функция которого имеет следующий вид:

$$P(t, x, \Gamma) = \begin{cases} \frac{1}{\sqrt{2\pi t}} \int_\Gamma \exp\left(-\frac{(u-x)^2}{2t}\right) du, & \text{если } x \neq 0, \\ \delta_0(\Gamma), & \text{если } x = 0. \end{cases}$$

(Здесь $\delta_0(\Gamma)$ — единичная мера, сосредоточенная в точке 0.) Проверим выполнение марковского свойства для процесса X . Ясно, что P удовлетворяет всем требуемым условиям, чтобы быть переходной функцией. Таким образом, остается установить, что соотношение

$$P_x\{A \cap [X_{t+h} \in \Gamma]\} = \int_A P(h, X_t, \Gamma) P_x(d\omega) \quad (3)$$

имеет место для любых $t, h \geq 0$, $A \in \mathcal{F}_t$ и $\Gamma \in \mathcal{B}^1$. Заметим, что (3) выражает марковское свойство семейства $(X_t, t \geq 0, P_x)$, а (2) — его строго марковское свойство. Если $x \neq 0$, то $X_t = w_t$ п. н. и (3) сводится к имеющему место марковскому свойству для винеровского процесса w . Если $x = 0$, то равенство (3) тривиально, потому что обе его части равны одновременно либо 1, либо 0. Следовательно, X — марковский процесс.

Пусть теперь $x \neq 0$, $\tau = \inf\{t: X_t = 0\}$, $\eta = (1 - \tau) \vee 0$, $A = \{\tau < 1\}$, $\Gamma \in \mathbb{R}^1 \setminus \{0\}$. Тогда очевидно, что $\tau < \infty$ п. н. и $\eta < \infty$ п. н. Допустим, что процесс X строго марковский. Из этого вытекало бы, что

$$P_x\{A \cap [X_{\tau+\eta} \in \Gamma]\} = \int_A P(\eta, X_\eta, \Gamma) P_x(d\omega). \quad (4)$$

Левая часть соотношения (4) равна $P_x\{\tau \leq 1, X_1 \neq 0\} = P_x\{\tau \leq 1\} = 2(1 - \Phi(|x|)) > 0$, а его правая часть равна 0. Следовательно, (4) не может иметь места, а это означает, что процесс X не является строго марковским.

(б) Пусть на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ задана экспоненциально распределенная случайная величина τ с параметром 1. Введем случайный процесс

$$X_t = X_t(\omega) = \max(0, t - \tau(\omega)), \quad t \geq 0,$$

и пусть $\mathcal{F}_t = \sigma\{X_u, u \leq t\}$, $t \geq 0$, — порожденная этим процессом фильтрация.

Легко видеть, что если $X_t = a > 0$ для некоторого t , то $X_{t+s} = a + s$ для всех $s \geq 0$. А если при некотором t мы имеем $X_t = 0$, тогда X_s должно равняться нулю для $s \leq t$, т.е. это «прошлое» не влияет на эволюцию процесса X в будущем. Значит, X — марковский процесс. Обозначим через $P(t, x, \Gamma)$ его переходную функцию.

Покажем, что процесс X не является строго марковским. Действительно, соотношение

$$\{\omega: \tau(\omega) \geq t\} = \{\omega: X_t(\omega) = 0\} \in \mathcal{F}_t$$

показывает, что случайная величина τ — это момент остановки относительно фильтрации $(\mathcal{F}_t)$. Поскольку $P\{X_{\tau+s} = s\} = 1$ для заданного τ , то

$$P\{X_{\tau+s} \leq x \mid \mathcal{F}_\tau\} = \begin{cases} 1, & \text{если } x > s, \\ 0, & \text{если } x \leq s. \end{cases} \quad (5)$$

Допустим, что процесс X строго марковский. Тогда должно быть выполнено соотношение

$$P\{X_{\tau+s} \leq x \mid \mathcal{F}_\tau\} = P\{X_{\tau+s} \leq x \mid X_\tau\} \quad (P\text{-п.н.}), \quad (6)$$

причем условную вероятность в правой части можно было записать, согласно (1), через переходную функцию следующим образом:

$$P\{X_{\tau+s} \leq x \mid X_{\tau}\} = P(s, X_{\tau}, \Gamma_x), \quad \text{где } \Gamma_x = (-\infty, x).$$

Однако, с другой стороны,

$$\begin{aligned} P(s, X_{\tau}, \Gamma_x) &= P(s, 0, \Gamma_x) = P\{X_{t+s} \leq x \mid X_t = 0\} = \\ &= \begin{cases} 1, & \text{если } x > s, \\ e^{-(s-x)}, & \text{если } x \leq s. \end{cases} \end{aligned} \quad (7)$$

Равенства (5) и (7) показывают, что соотношение (6) не выполняется. Следовательно, хотя процесс X марковский, он не является строго марковским.

§ 21. Стационарные процессы и некоторые смежные вопросы

Пусть на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ задан случайный процесс $X = (X_t, t \in T \subset \mathbb{R}^1)$ со значениями в $\mathbb{R}^1$. Мы говорим, что X — *стационарный в узком смысле процесс*, если для каждого $n \geq 1$ и $t_k, t_k + h \in T, k = 1, \dots, n$, случайные векторы

$$(X_{t_1}, \dots, X_{t_n}) \quad \text{и} \quad (X_{t_1+h}, \dots, X_{t_n+h})$$

имеют одно и то же распределение.

Предположим теперь, что X является процессом второго порядка, т.е. X принадлежит пространству $\mathcal{L}^2 = \mathcal{L}^2(T)$, где $\mathcal{L}^2(T) = \{X: E(|X_t|^2) < \infty \text{ для каждого } t \in T\}$. Такой процесс X называется *стационарным в широком смысле*, если $EX_t = c = \text{const}$ для всех $t \in T$, а его *ковариационная функция* $C(s, t) = E\{(X_s - c)(X_t - c)\}$ зависит от s и t только через разность $s - t$. Это означает, что существует функция $C(t)$ одного аргумента $t, t \in T$, такая что $C(t) = C(s, s + t)$ для всех $s, s + t \in T$.

Понятия стационарности легко переносятся как на многомерные процессы (в $\mathbb{R}^k, k \geq 2$), так и на комплекснозначные процессы.

Ковариационная функция C любого стационарного в широком смысле процесса допускает следующее *спектральное представление*:

$$C(t) = \int e^{it\lambda} dF(\lambda),$$

где интегрирование в правой части ведется в пределах $-\pi \leq \lambda \leq \pi$ в случае дискретного времени $t \in \mathbb{N}$ и в пределах $-\infty < \lambda < \infty$ в случае непрерывного времени $t \in \mathbb{R}^1$. Функция $F(\lambda)$, где $\lambda \in [-\pi, \pi]$ или $\lambda \in \mathbb{R}^1$,

является ограниченной, неубывающей и непрерывной справа. Она называется *спектральной функцией* процесса X . Если F дифференцируема, то ее производная $F' = f$ называется *спектральной плотностью* процесса X . Отметим, что не только ковариационные функции, но и сами стационарные процессы допускают спектральные представления в форме интегралов $\int e^{it\lambda} dZ(\lambda)$ относительно случайных мер Z с ортогональными приращениями.

Введем символ $\mathcal{M}_a^b$, где $a \leq b$, для обозначения σ -алгебры $\sigma\{X_u, a \leq u \leq b, u \in T\}$, порожденной процессом X на отрезке $[a, b]$. Через вероятности событий, принадлежащих σ -алгебрам $\mathcal{M}_{-\infty}^s$ и $\mathcal{M}_{s+t}^\infty$, $t > 0$, мы определяем важные свойства стационарных в узком смысле процессов, такие, например, как свойство перемешивания и свойство регулярности. Точные определения и анализ этих свойств будут даны в самих примерах ниже.

Отметим, наконец, что стационарным процессам и смежным вопросам посвящена обширная литература. Читатель может обратиться, например, к следующим книгам: [14, 21, 22, 47, 50, 62, 97, 116, 121, 363, 393].

Рассмотренные далее примеры посвящены анализу свойств стационарности, перемешивания, регулярности и эргодичности.

21.1. О свойстве стационарности случайных процессов

Выше мы ввели определения стационарности в узком и широком смыслах. Оказывается, случайный процесс может обладать либо обоими этими свойствами, либо ровно одним из них, либо, наконец, ни одним. Проиллюстрируем все эти возможности.

(а) Если $X = (X_t, t \in \mathbb{R}^1)$ — стационарный в узком смысле процесс и, кроме того, X — процесс второго порядка ($X \in \mathcal{L}^2$), то очевидно, что X является также стационарным в широком смысле процессом. Однако стационарный в узком смысле процесс X может не быть стационарным в широком смысле, и это имеет место для всех процессов, не принадлежащих пространству $\mathcal{L}^2$. Такие примеры можно привести без труда.

(б) Пусть случайная величина θ распределена равномерно на отрезке $[0, 2\pi]$, а ζ — не зависящая от нее случайная величина с произвольным распределением. Рассмотрим случайные процессы:

$$\begin{aligned} X &= (X_t, t \in \mathbb{R}^1), \quad \text{где} \quad X_t = \sin \theta t; \\ Y &= (Y_t, t \in \mathbb{R}^1), \quad \text{где} \quad Y_t = \cos(\zeta t + \theta). \end{aligned}$$

Легко проверить, что процесс X не является ни стационарным в узком, ни стационарным в широком смысле, а процесс Y обладает обоими

свойствами, т. е. Y — стационарный и в узком и в широком смысле процесс.

(в) Возьмем теперь две случайные величины ξ_1 и η_1 , каждая с распределением $\mathcal{N}(0, 1)$, и такие, что ξ_1 и η_1 некоррелированы, причем распределение вектора (ξ_1, η_1) не является двумерным нормальным. Такие примеры были описаны в § 10. Рассмотрим бесконечную последовательность независимых копий (ξ_1, η_1) , т. е.

$$(\xi_1, \eta_1), \quad (\xi_2, \eta_2), \quad \dots$$

Сохраняя тот же порядок элементов этой последовательности, переобозначим их следующим образом:

$$X_1 = \xi_1, \quad X_2 = \eta_1, \quad X_3 = \xi_2, \quad X_4 = \eta_2, \quad \dots$$

Так получаем случайный процесс $X = (X_n, n = 1, 2, \dots)$ с дискретным временем. Нетрудно проверить, что X является стационарным в широком смысле, но не является стационарным в узком смысле процессом.

Наконец, приведем пример процесса с непрерывным временем, подобного процессу X . Предположим, что при $t \geq 0$ случайная величина Y_t имеет распределение $\mathcal{N}(1, 1)$, а при $t < 0$ случайная величина Y_t распределена экспоненциально с параметром 1. Пусть также при $s \neq t$ величины Y_s и Y_t независимы. Тогда легко видеть, что случайный процесс $Y = (Y_t, t \in \mathbb{R}^1)$ является стационарным в широком смысле, но не обладает свойством стационарности в узком смысле.

21.2. О стационарности заданного порядка m

Пусть m — фиксированное натуральное число. Случайный процесс $X = (X_t, t \in T \subset \mathbb{R}^1)$ называется *m -стационарным* в узком смысле (или *стационарным в узком смысле порядка m*), если для любых $t_1, \dots, t_m \in T$ и $t_1 + h, \dots, t_m + h \in T$ векторы $(X_{t_1}, \dots, X_{t_m})$ и $(X_{t_1+h}, \dots, X_{t_m+h})$ имеют одно и то же распределение. Ясно, что стационарный в узком смысле процесс обладает этим свойством для каждого m , $m \geq 1$. Далее, легко видеть, что m -стационарный процесс X является также k -стационарным для каждого $k = 1, \dots, m - 1$, т. е. при уменьшении порядка m свойство стационарности сохраняется. Покажем, что при увеличении порядка m свойство стационарности может не выполняться.

Рассмотрим независимые случайные величины ξ и η , имеющие одну и ту же невырожденную функцию распределения $F(x)$, $x \in \mathbb{R}^1$. Определим последовательность $X = (X_n, n = 1, 2, \dots)$ так:

$$\begin{aligned} X_1 = \xi, \quad X_2 = \xi, \quad X_3 = \xi, \quad X_4 = \eta, \quad X_5 = \eta, \quad X_6 = \xi, \quad X_7 = \xi, \\ X_8 = \xi, \quad X_9 = \eta, \quad X_{10} = \eta, \quad X_{11} = \xi, \quad \dots \end{aligned}$$

Это означает, что для $k = 0, 1, 2, \dots$

$$X_n = \begin{cases} \xi, & \text{если } n = 5k + 1, 5k + 2, 5k + 3, \\ \eta, & \text{если } n = 5k + 4, 5k + 5. \end{cases}$$

Легко убедиться, что X является 1-стационарной в узком смысле последовательностью. Поэтому ставим вопрос о стационарности порядка 2. Заметим, что векторы (X_1, X_2) , (X_2, X_3) , (X_4, X_5) , $(X_6, X_7), \dots$ одинаково распределены. Однако (X_3, X_4) , т. е. (X_{1+2}, X_{2+2}) имеет распределение, отличающееся от распределения (X_1, X_2) . На самом деле, поскольку $X_1 = \xi$, $X_2 = \xi$, $X_3 = \xi$ и $X_4 = \eta$,

$$\begin{aligned} P\{X_1 \leq x_1, X_2 \leq x_2\} &= P\{\xi \leq x_1, \xi \leq x_2\} = \\ &= P\{\xi \leq \min(x_1, x_2)\} = F(\min(x_1, x_2)) \neq \\ &\neq F(x_1)F(x_2) = P\{\xi \leq x_1, \eta \leq x_2\} = P\{X_3 \leq x_1, X_4 \leq x_2\}. \end{aligned}$$

Итак, стационарность порядка 2 не имеет места для последовательности X . Вывод следующий: если задан m -стационарный в узком смысле процесс X , то свойство стационарности порядка k имеет место для каждого $k = 1, \dots, m - 1$, но может не иметь места при $k \geq m + 1$.

21.3. Условие сильного перемешивания не всегда сохраняется при преобразовании стационарного в узком смысле процесса

Пусть стационарный в узком смысле процесс $X = (X_n, n \in \tilde{\mathbb{N}})$ удовлетворяет условию *сильного перемешивания*, т. е. существует числовая последовательность $\alpha(n) \downarrow 0$ при $n \rightarrow \infty$ такая, что

$$\sup_{A, B} |P(AB) - P(A)P(B)| \leq \alpha(n),$$

где $\sup$ берется по всем событиям $A \in \mathcal{M}_{-\infty}^k$ и $B \in \mathcal{M}_{k+n}^\infty$, $k \in \tilde{\mathbb{N}}$, $n \in \mathbb{N}$. Это условие играет существенную роль при доказательстве предельных теорем для «слабо зависимых» случайных величин (см. [6, 46, 47, 419, 422]).

Пусть $g(x)$, $x \in \mathbb{R}^\infty$, — любая ограниченная и $\mathcal{B}^\infty$ -измеримая функция, и пусть $Y_n = g(X_n, X_{n+1}, \dots)$. Естественно поставить вопрос о свойствах процесса $Y = (Y_n, n \in \tilde{\mathbb{N}})$, полученного преобразованием исходного процесса X с помощью функции g . Например, хорошо известно, что свойство стационарности (в узком смысле) имеет место и для процесса Y . Поскольку процесс X удовлетворяет условию сильного перемешивания, то хотелось бы иметь ответ на вопрос: выполнено ли это условие и для процесса Y ?

Рассмотрим последовательность независимых случайных величин $(\varepsilon_j, j \in \tilde{\mathbb{N}})$ таких, что $P\{\varepsilon_j = 1\} = P\{\varepsilon_j = 0\} = \frac{1}{2}$ для каждого $j \in \tilde{\mathbb{N}}$.

Определим новую случайную последовательность $(\xi_j, j \in \tilde{\mathbb{N}})$, где

$$\xi_j = 2^{-1}\varepsilon_j + 2^{-2}\varepsilon_{j+1} + \dots + 2^{-k-1}\varepsilon_{j+k} + \dots$$

Из свойств величин (ε_j) вытекает, что $(\varepsilon_j, j \in \tilde{\mathbb{N}})$ — стационарный в узком смысле процесс. Отсюда заключаем, что процесс $(\xi_j, j \in \tilde{\mathbb{N}})$ также стационарен в узком смысле. Однако (ε_j) удовлетворяет условию сильного перемешивания, и можно предположить, что и последовательность (ξ_j) обладает этим свойством. Допустим, что это предположение верно. Тогда (см. [46, 47]) для любой ограниченной и измеримой функции $u(x)$, $x \in \mathbb{R}^1$, случайная последовательность $(\eta_j = u(\xi_j), j \in \tilde{\mathbb{N}})$ будет удовлетворять ЦПТ, т. е. имеет место соотношение

$$\lim_{n \rightarrow \infty} P\left\{\frac{S_n - ES_n}{b_n} \leq z\right\} = \Phi(z), \quad z \in \mathbb{R}^1,$$

где $S_n = \sum_{j=1}^n \eta_j$, $b_n^2 = DS_n$. При этом необходимо, чтобы дисперсия b_n^2 имела вид $b_n^2 = nh(n)$, где $h(n)$ — медленно меняющаяся функция.

Возьмем теперь последовательность функций Радемахера $r_k(x)$, $k = 1, 2, \dots$, $x \in [0, 1]$ (см. [47, 358]). Так как $r_k(x) = \text{sgn} \sin(2^k \pi x)$, то нетрудно установить, что r_k , $k = 1, 2, \dots$, где $r_k = r_k(\xi_1)$ — последовательность независимых и одинаково распределенных случайных величин таких, что $P\{r_k = \pm 1\} = \frac{1}{2}$. Положим

$$\eta_j = u(\xi_j), \quad \text{где} \quad u(x) = \sum_{k=1}^{\infty} r_k(x) k^{-3/4}.$$

Таким образом, мы построили последовательность случайных величин $\{\eta_j\}$, для которой имеет место ЦПТ и должно выполняться соотношение $b_n^2 = D\{\eta_1 + \dots + \eta_n\} = nh(n)$, где $h(n)$ — медленно меняющаяся функция. Найдем дисперсию b_n^2 . Поскольку $Er_k = 0$ и $Dr_k = 1$, нетрудно видеть, что

$$E\{\eta_1 \eta_{j+1}\} = \sum_{k=1}^{\infty} k^{-3/4} (k+j)^{-3/4} > 2^{-3/4} j^{-3/4},$$

и поэтому

$$b_n^2 = E\{(\eta_1 + \dots + \eta_n)^2\} > n^{5/4} (1 + o(1)).$$

Заметим, однако, что так как $b_n^2 > n^{5/4} (1 + o(1))$, то b_n^2 не может быть представлена в виде $nh(n)$ с медленно меняющейся функцией $h(n)$. Это

противоречие показывает, что процесс $(\xi_j, j \in \tilde{\mathbb{N}})$ не удовлетворяет условию сильного перемешивания. Это и есть ответ на поставленный выше вопрос.

Отметим, наконец, что по поводу условий перемешиваний для случайных полей читатель может обратиться к работам [11, 161].

21.4. Стационарный в узком смысле процесс может быть регулярным и не обладать свойством полной регулярности

Стационарный в узком процесс $X = (X_t, t \in \mathbb{R}^1)$ называется *регулярным*, если σ -алгебра

$$\mathcal{M}_{-\infty} = \bigcap_{t \in \mathbb{R}^1} \mathcal{M}_{-\infty}^t$$

тривиальна: каждое событие из $\mathcal{M}_{-\infty}$ имеет вероятность 0 или 1. Это свойство может быть выражено в следующей эквивалентной форме: для любых $B \in \mathcal{M}_{-\infty}^0$ и $A \in \mathcal{M}_{-\infty}^t$

$$\sup_A |P(AB) - P(A)P(B)| \rightarrow 0 \quad \text{при } t \rightarrow \infty.$$

Положим $\rho(t) = \sup E\{\eta_1 \eta_2\}$, где $\sup$ берется по всем случайным величинам η_1 и η_2 , η_1 измерима относительно $\mathcal{M}_{-\infty}^s$, $E\eta_1 = 0$, $E\{\eta_1^2\} = 1$, а η_2 измерима относительно $\mathcal{M}_{s+t}^\infty$, $E\eta_2 = 0$, $E\{\eta_2^2\} = 1$. Напомним, что $\rho(t)$, $t \geq 0$, называется максимальным коэффициентом корреляции σ -алгебр $\mathcal{M}_{-\infty}^s$ и $\mathcal{M}_{s+t}^\infty$. Случайный процесс X называется *вполне регулярным* (или X обладает *свойством полной регулярности*), если $\rho(t) \rightarrow 0$ при $t \rightarrow \infty$. Для стационарных и гауссовских процессов свойство полной регулярности совпадает со свойством сильного перемешивания, рассмотренного в примере 21.1. Очевидно, что каждый вполне регулярный процесс является также регулярным. Поэтому полезно описать пример стационарного процесса, являющегося регулярным, но не обладающего свойством полной регулярности. Для этой цели воспользуемся следующим результатом (см. [50]): *пусть стационарный в узком смысле процесс $X = (X_t, t \in \mathbb{R}^1)$ имеет спектральную плотность $f(\lambda)$, $\lambda \in \mathbb{R}^1$. Этот процесс является регулярным тогда и только тогда, когда*

$$\int_{-\infty}^{\infty} (1 + \lambda^2) \ln f(\lambda) d\lambda < \infty. \quad (1)$$

Рассмотрим теперь стационарный процесс Y , спектральная плотность которого задается формулой

$$g(\lambda) = (\sin^2 \lambda^2 + 1)(\lambda^{-1} \sin \lambda)^{2p}, \quad \lambda \in \mathbb{R}^1, \quad (2)$$

где p — любое натуральное число. Можно проверить, что функция g удовлетворяет условию (1), и, следовательно, процесс Y регулярный. Оказывается, однако, что спектральная плотность g не удовлетворяет условию (4.4) из книги [50, с. 291], являющемуся необходимым условием для полной регулярности. (Это условие несколько громоздко, и поэтому мы его здесь опускаем.)

Таким образом, мы приходим к выводу, что стационарный процесс Y со спектральной плотностью (2) является регулярным, но не является вполне регулярным.

21.5. Законы больших чисел для стационарных в широком смысле процессов

Стационарная в широком смысле последовательность $X = (X_n, n \in \mathbb{N})$ с $EX_n = 0$, $n \in \mathbb{N}$, удовлетворяет (обычному) ЗБЧ, если

$$\frac{1}{n} \sum_{k=1}^n X_k \xrightarrow{P} 0 \quad \text{при } n \rightarrow \infty. \quad (1)$$

Если же соотношение (1) имеет место с вероятностью 1, то мы говорим, что процесс X удовлетворяет усиленному ЗБЧ.

Если стационарный в широком смысле процесс $X = (X_t, t \in \mathbb{R}^+)$ с непрерывным временем удовлетворяет соотношению

$$\frac{1}{T} \int_0^T X_t dt \xrightarrow{P} 0 \quad \text{при } T \rightarrow \infty, \quad (2)$$

то мы говорим, что X удовлетворяет ЗБЧ, а при выполнении (2) с вероятностью 1 — что X удовлетворяет УЗБЧ.

Отметим, что свойства типа (1) и (2) называются эргодическими свойствами для стационарных в широком смысле процессов. В соответствии с этим, если X удовлетворяет ЗБЧ, мы говорим, что X — *эргодический в слабом смысле процесс*, а если для X имеет место УЗБЧ, то X — *эргодический в сильном смысле процесс*.

Ясно, что из эргодичности в сильном смысле стационарного процесса X вытекает его эргодичность в слабом смысле. Возникает вопрос о справедливости обратного утверждения.

Условия, обеспечивающие эргодичность стационарного процесса, задаются обычно в терминах его ковариационной функции или его спектральной функции распределения. Остановимся на одном конкретном случае.

Предположим, что $X = (X_n, n \geq 1)$ — стационарная в широком смысле последовательность такая, что $EX_n = 0$, $E\{X_n^2\} = 1$, а ковариационная функция $C(n) = E\{X_k X_{k+n}\}$ удовлетворяет условию

$$\lim_{n \rightarrow \infty} C(n) = 0. \quad (3)$$

Тогда процесс X — эргодический в слабом смысле. Доказать это нетрудно, но можно, например, обратиться к книге [21]. Более того, можно показать, что $\frac{1}{n} \sum_{k=1}^n X_k \xrightarrow{L^2} 0$ при $n \rightarrow \infty$. Это означает, что условие (3) достаточно для L^2 -эргодичности процесса X . Если мы дополнительно предположим, что процесс X стационарный в узком смысле, то из условия (3) вытекает, что X эргодический в сильном смысле, т. е. что X удовлетворяет УЗБЧ. Таким образом, один из возможных вопросов здесь следующий: если X — стационарный в широком смысле процесс и выполнено условие (3), то следует ли отсюда, что X эргодический в сильном смысле? Оказывается, в общем случае это не так. Здесь мы не будем приводить подробности, но отметим, что можно построить стационарную в широком смысле последовательность $\xi = (\xi_n, n \geq 1)$ с нулевым средним и ковариационной функцией $C(n)$, удовлетворяющей условию

$$C(n) = O\left(\frac{1}{(\ln \ln n)^2}\right) \quad \text{при } n \rightarrow \infty. \quad (4)$$

Поскольку $C(n) \rightarrow 0$ при $n \rightarrow \infty$, то процесс ξ эргодический в слабом смысле. Однако условие (4) позволяет установить расходимость п. н. для средних, т. е. что $\frac{1}{n} \sum_{k=1}^n \xi_k \rightarrow \infty$ вероятностью 1, когда $n \rightarrow \infty$. Поэтому процесс ξ не может быть эргодичным в сильном смысле. По поводу конструкции и свойств процесса ξ отсылаем читателя к работе [17]. Там же можно найти и описание стационарного процесса с непрерывным временем, удовлетворяющего ЗБЧ и не удовлетворяющего УЗБЧ.

21.6. О свойствах эргодичности и перемешивания для сохраняющих меру преобразований

Сделаем сначала следующее замечание. Стационарные процессы обладают рядом интересных и важных свойств, таких как эргодичность, перемешивание и т. д. Оказывается, их можно изучать единым образом как преобразования вероятностного пространства, на котором эти процессы определены (см. [92, 116, 129, 422]).

Приведем определения нескольких понятий, а потом проиллюстрируем их на одном примере.

Пусть $(\Omega, \mathcal{F}, P)$ — вероятностное пространство и T — преобразование Ω в себя. T называется *измеримым*, если $T^{-1}A = \{\omega: T\omega \in A\} \in \mathcal{F}$ для всех $A \in \mathcal{F}$. Мы говорим, что $T: \Omega \rightarrow \Omega$ — *сохраняющее меру преобразование*, если $P(T^{-1}A) = P(A)$ для каждого $A \in \mathcal{F}$. Если событие $A \in \mathcal{F}$ таково, что $T^{-1}A = A$, то A называется *инвариантным событием*. Класс всех инвариантных событий образует σ -алгебру, которую обозначим $\mathcal{I}$. Если для каждого $A \in \mathcal{I}$ мы имеем $P(A) = 0$ или 1, то сохраняющее меру преобразование T называется *эргодическим*. Функция $g: (\Omega, \mathcal{F}) \mapsto (\mathbb{R}^1, \mathcal{B}^1)$ называется *инвариантной* относительно T тогда и только тогда, когда $g(T\omega) = g(\omega)$ для каждого ω . Наконец, мы говорим, что сохраняющее меру преобразование T удовлетворяет *условию перемешивания*, если для любых $A, B \in \mathcal{F}$ имеет место соотношение

$$\lim_{n \rightarrow \infty} P(A \cap T^{-n}B) = P(A)P(B).$$

Наша цель теперь — сравнить свойства эргодичности и перемешивания, а наши дальнейшие рассуждения основываются на следующем результате (см. [116, 129, 330, 422]): *пусть T — сохраняющее меру преобразование вероятностного пространства $(\Omega, \mathcal{F}, P)$. Тогда:*

(а) *T эргодично только тогда, когда каждая T -инвариантная функция P -н.н. равна константе.*

(б) *Если T удовлетворяет условию перемешивания, то T эргодично.*

Пусть $\Omega = [0, 1)$, $\mathcal{F} = \mathcal{B}_{[0,1)}$ и P — мера Лебега. Рассмотрим преобразование $T\omega = (\omega + \lambda) \bmod 1$, $\omega \in \Omega$. Нетрудно видеть, что T — сохраняющее меру преобразование. Вопрос: обладает ли T свойством эргодичности и свойством перемешивания?

Предположим сначала, что число λ рациональное, т. е. $\lambda = \frac{k}{m}$, где k и m — целые числа. Возьмем множество

$$A = \bigcup_{k=0}^{2m-2} \left(\omega: \frac{k}{2m} \leq \omega < \frac{k+1}{2m} \right).$$

Тогда A — инвариантное множество, причем $P(A) = \frac{1}{2}$. Отсюда вытекает, что в случае рационального λ преобразование T не может быть эргодическим.

Пусть теперь λ — иррациональное число. Покажем, что в этом случае T эргодично. Для этой цели рассмотрим случайную величину $\xi = \xi(\omega)$ с $E\{\xi^2(\omega)\} < \infty$. Известно (см., например, [61, 116, 129, 421]), что ряд Фурье $\sum_{n=-\infty}^{\infty} c_n e^{2\pi i n \omega}$ функции $\xi(\omega)$ сходится в L^2 -смысле и $\sum_{n=-\infty}^{\infty} |c_n|^2 < \infty$. Предположим, что ξ — инвариантная случайная

величина. Поскольку T сохраняет меру, то для коэффициента Фурье c_n находим

$$\begin{aligned} c_n &= \mathbb{E}\{\xi(\omega)e^{-2\pi i n \omega}\} = \mathbb{E}\{\xi(T\omega)e^{-2\pi i n T\omega}\} = \\ &= e^{-2\pi i n \lambda} \mathbb{E}\{\xi(T\omega)e^{-2\pi i n \lambda}\} = e^{-2\pi i n \lambda} \mathbb{E}\{\xi(\omega)e^{-2\pi i n \omega}\} = c_n e^{-2\pi i n \lambda}. \end{aligned}$$

Отсюда следует, что $c_n(1 - e^{-2\pi i n \lambda}) = 0$. Но мы предположили, что λ иррационально, и, значит, $e^{-2\pi i n \lambda} \neq 1$ для каждого $n \neq 0$. Следовательно, $c_n = 0$ при $n \neq 0$ и $\xi(\omega) = c_0$ п. н., где $c_0 = \text{const}$. Ссылаясь на приведенный выше результат, заключаем, что преобразование T эргодическое.

Установим, наконец, что эргодическое преобразование T не удовлетворяет условию перемешивания. Действительно, рассмотрим множество $A = \left\{\omega: 0 \leq \omega \leq \frac{1}{2}\right\}$, и пусть $B = A$. Поскольку T — сохраняющее меру преобразование, то T обратимо и имеют место равенства

$$\mathbb{P}(A \cap T^{-n}B) = \mathbb{P}(A \cap T^{-n}A) = \mathbb{P}(T^n A \cap A) \quad \text{для любого } n.$$

Зафиксируем число $\varepsilon \in (0, 1)$. Так как λ иррационально, то имеется бесконечно много n , для которых разность между $e^{2\pi i n \lambda}$ и $e^{i \cdot 0} = 1$ по абсолютной величине не превышает ε . Множества A и $T^n A$ пересекаются, за исключением некоторого множества, мера которого меньше, чем ε . Таким образом,

$$\mathbb{P}(A \cap T^{-n}B) \geq \mathbb{P}(A) - \varepsilon.$$

Выбирая, например, $0 < \varepsilon < \frac{1}{8}$, мы находим

$$\mathbb{P}(A \cap T^{-n}B) > \frac{3}{8}.$$

Предположим, что T удовлетворяет условию перемешивания. Тогда

$$\mathbb{P}(A \cap T^{-n}B) \rightarrow \mathbb{P}(A)\mathbb{P}(B) \quad \text{при } n \rightarrow \infty,$$

и должно быть $\mathbb{P}(A)\mathbb{P}(B) > \frac{3}{8}$. Однако, поскольку $\mathbb{P}(A) = \mathbb{P}(B) = \frac{1}{2}$, мы имеем

$$\mathbb{P}(A)\mathbb{P}(B) = [\mathbb{P}(A)]^2 = \frac{1}{4}.$$

Итак, мы пришли к противоречию, которое показывает, что эргодическое преобразование T не обладает свойством перемешивания.

21.7. Стационарные последовательности, для которых центральная предельная теорема не имеет места

В классической ЦПТ мы рассматриваем независимые случайные величины (см. § 17). Поэтому, вообще говоря, если последовательность

(X_n , $n \geq 1$) состоит из «слабо» зависимых величин, то мы не можем ожидать, что распределение нормированных сумм $\frac{S_n}{s_n}$ будет сходиться к стандартному нормальному распределению без дополнительных предположений. Здесь, как обычно, $S_n = X_1 + \dots + X_n$, $s_n^2 = DS_n$, и без ограничения общности можно считать, что $EX_n = 0$, $n \geq 1$. Отметим, однако, что существует большое число работ, в которых при подходящих условиях доказывается, что ЦПТ имеет место для широких классов стационарных последовательностей (см., например, [28, 47, 154, 419]) и для стационарных полей (см. [11]).

Приведем примеры, показывающие, что при $n \rightarrow \infty$ нормированные суммы $\frac{S_n}{s_n}$ могут иметь различное поведение. В частности, предельное распределение $\frac{S_n}{s_n}$ может не существовать, а если существует, то может не быть нормальным.

(а) Пусть $(\xi_k, k = 0, \pm 1, \dots)$ — независимые и одинаково распределенные случайные величины с $P\{\xi_k = \pm 1\} = \frac{1}{2}$, и пусть $X_k = \xi_k - \xi_{k-1}$. Так полученная случайная последовательность $(X_k, k \in \tilde{\mathbb{N}})$ является 1-зависимой (т.е. X_k связано с X_{k-1} и X_{k+1} и не зависит от X_j при $|k - j| > 1$), и легко проверяется, что она удовлетворяет условию сильного перемешивания. Изучим асимптотическое поведение $\frac{S_n}{s_n}$. Заметим, что $S_n = \xi_n - \xi_0$ для каждого $n \geq 1$, а это означает, что случайная величина S_n принимает значения $-2, 0$ и 2 соответственно с вероятностями $\frac{1}{4}, \frac{1}{2}$ и $\frac{1}{4}$. Отсюда ясно, что последовательность (X_k) не удовлетворяет ЦПТ.

(б) Рассмотрим такую последовательность случайных величин $(X_n, n = 0, \pm 1, \dots)$, что для произвольного целого n и неотрицательного целого m случайный вектор $(X_n, X_{n+1}, \dots, X_{n+m})$ имеет следующую плотность (смесь двух нормальных плотностей):

$$f(x_n, x_{n+1}, \dots, x_{n+m}) = \frac{1}{2}(2\pi)^{-n/2} \sigma_1^{-n} \exp\left(-\frac{1}{2}\sigma_1^{-2} \sum_{k=0}^m x_{n+k}^2\right) + \\ + \frac{1}{2}(2\pi)^{-n/2} \sigma_2^{-n} \exp\left(-\frac{1}{2}\sigma_2^{-2} \sum_{k=0}^m x_{n+k}^2\right).$$

Здесь $x_j \in \mathbb{R}^1, j \in \tilde{\mathbb{N}}, \sigma_1 > 0, \sigma_2 > 0$ и предполагается, что $\sigma_1 \neq \sigma_2$. Легко проверить, что (X_k) — стационарная в узком смысле последовательность, и если $S_n = X_1 + \dots + X_n$, то

$$G(x) = \lim_{n \rightarrow \infty} P\left\{\frac{S_n}{\sqrt{n}} \leq x\right\} = \frac{1}{2}\Phi(\sigma_1 x) + \frac{1}{2}\Phi(\sigma_2 x), \quad x \in \mathbb{R}^1.$$

Итак, нормированная сумма $\frac{S_n}{\sqrt{n}}$ имеет предельное распределение G , которое, ввиду того, что $\sigma_1 \neq \sigma_2$, отличается от нормального. Очевидно, G является смесью двух нормальных распределений.

(в) Пусть $(X_n, n = 0, \pm 1, \dots)$ — стационарная в узком смысле последовательность с $EX_n = 0$ и $E\{X_n^2\} < \infty$, $n \in \tilde{\mathbb{N}}$, и пусть $\rho(n)$, $n \geq 1$, — связанный с ней максимальный корреляционный коэффициент (см. пример 21.4). Условие $\rho(n) \rightarrow 0$ при $n \rightarrow \infty$ играет существенную роль при доказательстве предельных теорем для стационарных процессов. В частности, из $\rho(n) \rightarrow 0$ вытекает, что (X_n) удовлетворяет условию сильного перемешивания (см. [47]).

Сформулируем следующий результат (см. [48]): *при введенных выше обозначениях, если $\rho(n) \rightarrow 0$ при $n \rightarrow \infty$, то либо $\sup_n s_n^2 < \infty$, либо $s_n^2 = nh(n)$, где $h(n)$ — медленно меняющаяся функция при $n \rightarrow \infty$. Если $s_n^2 \rightarrow \infty$ и $\rho(n) \rightarrow 0$ при $n \rightarrow \infty$ и $E\{|X_0|^{2+\delta}\} < \infty$ при некотором $\delta > 0$, то $\frac{S_n}{\sqrt{n}} \xrightarrow{d} \xi$, где случайная величина $\xi \sim \mathcal{N}(0, 1)$, т. е. стационарная последовательность (X_n) удовлетворяет ЦПТ.*

Возникает вопрос: можно ли ослабить наложенные условия и сохранить асимптотическую нормальность $\frac{S_n}{\sqrt{n}}$? Например, вместо условия $E\{|X_0|^{2+\delta}\} < \infty$ предположим, что $E\{|X_0|^{2+\delta}\} = \infty$ для каждого $\delta > 0$, но $E\{|X_0|^2\} < \infty$. Этот случай изучен подробно в работах [154, 155], где доказан следующий результат: *существует стационарная в узком смысле последовательность $(X_n, n = 0, \pm 1, \dots)$, для которой имеют место следующие свойства:*

- (1) $EX_n = 0$ и $0 < DX_n < \infty$;
- (2) $s_n^2 \rightarrow \infty$ при $n \rightarrow \infty$;
- (3) $\rho(n) \rightarrow 0$ при $n \rightarrow \infty$;
- (4) (X_n) не удовлетворяет ЦПТ в том смысле, что для любого $\lambda > 0$ имеется возрастающая последовательность индексов $\{n(k)\}$ такая, что $\frac{\sqrt{\lambda} S_{n(k)}}{s_{n(k)}} \xrightarrow{d} \xi_\lambda$ при $k \rightarrow \infty$, где ξ_λ — случайная величина со следующей функцией распределения:

$$F_\lambda(x) = e^{-\lambda} \cdot 1_{[0, \infty)}(x) + \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} (2\pi k)^{-1/2} \int_{-\infty}^x e^{-u^2/(2k)} du, \quad x \in \mathbb{R}^1.$$

Доказательство этого результата можно найти в указанных выше работах. Очевидно, что функция распределения F_λ нормальной не является. Она имеет массу, сосредоточенную в точке 0, и содержит слабое, являющееся пуассоновской смесью нормальных распределений.

Вне рамок обсуждаемого здесь вопроса о невыполнении ЦПТ для некоторых стационарных процессов отметим следующий интересный факт: F_λ — безгранично делимое, но не устойчивое распределение (см. также пример 9.7).

§ 22. Мартингалы с дискретным временем

Пусть на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ задана случайная последовательность $(X_n, n \geq 1)$ и фильтрация $(\mathcal{F}_n, n \geq 1)$, т. е. $\mathcal{F}_n \subset \mathcal{F}$ и $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ для каждого $n \geq 1$. Как обычно, запись $(X_n, \mathcal{F}_n, n \geq 1)$ означает, что последовательность (X_n) $(\mathcal{F}_n)$ -адаптирована. Если $\sup_{n \geq 1} E|X_n| < \infty$, мы говорим, что последовательность $(X_n, n \geq 1)$ является L^1 -ограниченной (или L^1 -интегрируемой), а если $E\left\{\sup_{n \geq 1} |X_n|\right\} < \infty$, она называется L^1 -доминированной.

Система $(X_n, \mathcal{F}_n, n \geq 1)$ называется *мартингалом*, если $E|X_n| < \infty$ для каждого $n \geq 1$ и для любых $m \leq n$ выполнено равенство

$$E\{X_n | \mathcal{F}_m\} = X_m \quad (P\text{-п.н.}). \quad (1)$$

Если в (1) вместо равенства мы имеем $E\{X_n | \mathcal{F}_m\} \leq X_m$ п.н. или $E\{X_n | \mathcal{F}_m\} \geq X_m$ п.н., то $(X_n, \mathcal{F}_n, n \geq 1)$ называется соответственно *супермартингалом* и *субмартингалом*.

Марковский момент относительно $(\mathcal{F}_n)$ — это любая функция $\tau: \Omega \mapsto \mathbb{N} \cup \{\infty\}$ такая, что $\{\tau = n\} \in \mathcal{F}_n$ для каждого $n \geq 1$. Если к тому же $P\{\tau < \infty\} = 1$, то τ называется *моментом остановки*. Через T обозначим класс всех моментов остановки. Скажем, что числовое множество $(a_\tau, \tau \in T)$, называемое *сетью*, сходится к числу b , если для любого $\varepsilon > 0$ найдется такой момент $\tau_0 \in T$, что для всех $\tau \geq \tau_0$ при $\tau \in T$ мы имеем $|a_\tau - b| < \varepsilon$.

Введем следующие определения. Мы говорим, что $(X_n, \mathcal{F}_n, n \geq 1)$

(а) *квазимартингал*, если $\sum_{n=1}^{\infty} E\{|X_n - E(X_{n+1} | \mathcal{F}_n)|\} < \infty$;

(б) *амарт*, если сеть $(EX_\tau, \tau \in T)$ сходящаяся;

(в) *мартингал в пределе*, если $\sup_{m > n} |E(X_m | \mathcal{F}_n) - X_n| \xrightarrow{п.н.} 0, n \rightarrow \infty$;

(г) *прогрессивный мартингал*, если для всех $n \geq 1$ события $A_n = \{\omega: E(X_{n+1} | \mathcal{F}_n) = X_n\}$ таковы, что $A_n \subset A_{n+1}$ и $P\left(\bigcup_{n=1}^{\infty} A_n\right) = 1$;

(д) *эвентуальный мартингал*, если $P\{E(X_{n+1} | \mathcal{F}_n) \neq X_n \text{ б. ч.}\} = 0$.

Отметим, что теория мартингалов с дискретным временем изложена во многих книгах; см., например, [32, 116, 177, 381]. По поводу

же различных мартингалных обобщений (а)–(д) читатель может обратиться к статьям [150, 213, 214, 264, 377, 464, 465]. Общие результаты, касающиеся мартингалов как с дискретным, так и с непрерывным временем, содержатся в монографии [74].

Для настоящего параграфа мы подобрали примеры, иллюстрирующие важные свойства и результаты для мартингалов и мартингалных последовательностей.

22.1. L^1 -ограниченные мартингалы, не являющиеся L^1 -доминированными

Пусть задан мартингал $X = (X_n, \mathcal{F}_n, n \geq 1)$. Из соотношения $\sup_{n \geq 1} E|X_n| \leq E\{\sup_{n \geq 1} |X_n|\}$ вытекает, что каждый L^1 -доминированный мартингал является L^1 -ограниченным. Обратное не всегда верно.

(а) Рассмотрим дискретное множество $\Omega = \{1, 2, \dots\}$ с вероятностью P его элементов, заданной так: $P(\{n\}) = \frac{1}{n} - \frac{1}{n+1}$. Обозначим $A_n = \Omega \setminus \{1, \dots, n\}$, и пусть $\mathcal{F}_n$ — это σ -алгебра, порожденная разбиением $\{\{1\}, \dots, \{n\}, A_n\}$ множества Ω . Видно, что σ -алгебры $(\mathcal{F}_n, n \geq 1)$ возрастают. Последовательность случайных величин $(X_n, n \geq 1)$ определим с помощью равенства

$$X_n = X_n(\omega) = (n+1)1_{A_n}(\omega), \quad n \geq 1.$$

Тогда $X = (X_n, \mathcal{F}_n, n \geq 1)$ — неотрицательный мартингал, причем $EX_n = 1$ для каждого $n \geq 1$, так что X — L^1 -ограничен. Однако $\sup_{n \geq 1} X_n(\omega) = \omega$, и ясно, что на рассматриваемом вероятностном пространстве он неинтегрируем. Следовательно, L^1 -ограниченный мартингал X не является L^1 -доминированным.

(б) Пусть $\Omega = [0, 1]$, $\mathcal{F} = \mathcal{B}_{[0,1]}$ и P — мера Лебега. Положим

$$X_n = X_n(\omega) = \begin{cases} 0, & \text{если } \frac{1}{n} \leq \omega \leq 1, \\ -n^2\omega + n, & \text{если } 0 \leq \omega < \frac{1}{n}, \end{cases}$$

и обозначим $\mathcal{F}_n = \sigma\{X_1, \dots, X_n\}$. Тогда $X = (X_n, \mathcal{F}_n, n \geq 1)$ — мартингал. Поскольку $EX_n = \frac{1}{2}$, $n \geq 1$, то X — L^1 -ограничен. Однако его супремум $\sup_{n \geq 1} |X_n|$, как легко видеть, неинтегрируем, и, следовательно, X не является L^1 -доминированным.

(в) Рассмотрим стандартный винеровский процесс $w = (w(t), t \geq 0)$ и порожденную им фильтрацию $(\mathcal{F}_t^w, t \geq 0)$. Возьмем любую числовую

последовательность $\{n_k, k \geq 1\}$, где $0 < n_1 < n_2 < \dots \rightarrow \infty$ при $k \rightarrow \infty$. Обозначим $M_k = \exp\left(w(n_k) - \frac{n_k}{2}\right)$. Тогда нетрудно показать, что $M = (M_k, \mathcal{F}_{n_k}^w, k \geq 1)$ — неотрицательный мартингал, который интегрируем (при этом даже $M_k \xrightarrow{\text{п.н.}} 0$ при $k \rightarrow \infty$), но $E\{\sup_{k \geq 1} M_k\} = \infty$. Следовательно, несмотря на интегрируемость мартингала M , он не обладает свойством L^1 -доминированности.

22.2. Мартингалы, для которых теорема Дуба об остановке не имеет места

Пусть задан мартингал $X = (X_n, \mathcal{F}_n, n \geq 0)$. Из его определения и свойств условных математических ожиданий вытекает, что равенство $EX_n = EX_0$ имеет место для каждого $n \geq 1$. Вопрос: справедливо ли это равенство, если заменить n моментом остановки τ ? По этому поводу сформулируем следующий результат, известный как *теорема Дуба об остановке* (см., например, [32, 57, 116, 381]): *предположим, что мартингал $X = (X_n, \mathcal{F}_n, n \geq 0)$ и момент остановки τ относительно $(\mathcal{F}_n)$ удовлетворяют условиям: (I) $E|X_\tau| < \infty$; (II) $\lim_{n \rightarrow \infty} \int_{\{\tau > n\}} X_n dP = 0$. Тогда $EX_\tau = EX_0$.*

Наша цель далее — показать, что оба условия (I) и (II) существенны для выполнения равенства $EX_\tau = EX_0$.

(а) Пусть последовательность $(\eta_n, n \geq 0)$ состоит из независимых и одинаково распределенных случайных величин таких, что η_1 принимает значения $-1, 0$ и 1 , причем $E\eta_1 = 0$. Положим $X_0 = 0, X_n = \eta_1 + \dots + \eta_n$ и $\mathcal{F}_n = \{\emptyset, \Omega\}$. Ясно, что $X = (X_n, \mathcal{F}_n, n \geq 0)$ — мартингал. Величина $\tau = \inf\{n: X_n = 1\}$ является марковским моментом относительно $\{\mathcal{F}_n\}$, причем $P\{\tau < \infty\} = 1$ и $X_\tau = 1$ п. н. Следовательно, $EX_\tau = 1 \neq 0 = EX_0$, т. е. теорема Дуба об остановке не имеет места для мартингала X и момента остановки τ . Обратимся поэтому к условиям (I) и (II).

Очевидно, что $E|X_\tau| < \infty$, и, значит, условие (I) выполнено. Далее воспользуемся равенством $0 = EX_n = I_n + J_n$, где

$$I_n = \int_{\{\tau \leq n\}} X_n dP, \quad J_n = \int_{\{\tau > n\}} X_n dP.$$

Заметим, что I_n не превышает вероятности того, что мартингал X достигнет уровня 1 до момента времени n , а эта вероятность стремится к 1 при $n \rightarrow \infty$. Так как $I_n + J_n = 0$, то $J_n \rightarrow -1$ при $n \rightarrow \infty$, т. е. $J_n \not\rightarrow 0$, и, значит, условие (II) не выполнено.

(б) Пусть $\xi_1, \xi_2, \dots$ — независимые случайные величины, где $\xi_n \sim \mathcal{N}(0, b_n^2)$. Дисперсии $b_n^2, n \geq 1$, выберем специальным образом.

Возьмем $b_1 = 1$ и положим $b_{n+1}^2 = a_{n+1}^2 - a_n^2$ для $n \geq 1$, где $a_n = \frac{(n-1)^2}{\ln(3+n)}$. (Причина именно такого выбора станет ясна из дальнейшего рассуждения.)

Положим $X_n = \xi_1 + \dots + \xi_n$ и $\mathcal{F}_n = \sigma\{\xi_1, \dots, \xi_n\}$. Тогда $X = (X_n, \mathcal{F}_n, n \geq 1)$ является мартингалом. Пусть g — измеримая функция, отображающая множество $\mathbb{R}^1$ в множество натуральных чисел $\mathbb{N}$, причем $P\{g(\xi_1) = n\} = p_n$, где $p_n = \frac{1}{n^2} - \frac{1}{(n+1)^2}$, $n \geq 1$. Таким образом, величина $\tau = g(\xi_1)$ является моментом остановки с конечным математическим ожиданием $E\tau$. Непосредственно проверяется, что $EX_\tau \neq EX_0$. Поэтому снова обратимся к условиям (I) и (II).

Обозначим через F функцию распределения случайной величины ξ_1 , и пусть $S_1 = 0$, $S_n = \xi_2 + \dots + \xi_n$ для $n \geq 2$. Тогда случайная величина ξ_1 не зависит от $S_1, S_2, \dots$, причем $X_n = \xi_1 + S_n$, где $S_n \sim \mathcal{N}(0, a_n^2)$. Вычислим $E|X_\tau|$ и $\int_{\{\tau > n\}} |X_n| dP$. Мы имеем

$$\begin{aligned} \int_{\{\tau < n\}} |X_n| dP &= \int_{\{g > n\}} E[|y + S_n|] dF(y) \leq \int_{\{g > n\}} (|y| + E|S_n|) dF(y) = \\ &= \int_{\{g > n\}} |y| dF(y) + ca_n P\{g > n\}, \end{aligned}$$

где $c = E|\xi_1|$. Отсюда вытекает, что $\int_{\{g > n\}} |X_n| dP \rightarrow 0$ при $n \rightarrow \infty$, т. е. условие (II) выполнено. Далее

$$\begin{aligned} E|X_\tau| &= \int E\{|y + S_{g(y)}|\} dF(y) \geq \\ &\geq \int E\{|S_{g(y)}|\} dF(y) = c \int a_{g(y)} dF(y) = c \sum_{n=1}^{\infty} p_n a_n = \infty, \end{aligned}$$

а это означает, что условие (I) не выполнено.

Итак, рассмотренные примеры показывают, что оба условия (I) и (II) существенны для справедливости теоремы Дуба об остановке.

22.3. Каждый квазимартингал является амартом, но обратное не верно

Исходя из соответствующих определений, нетрудно показать, что каждый квазимартингал является также амартом (подробности можно найти, например, в работах [213, 264]). Опишем два примера, из которых видно, что в общем случае обратное не верно.

(а) Пусть $a_n = \frac{(-1)^n}{n}$, $n \geq 1$, $X_n = a_n$ — вырожденные случайные величины. Выберем произвольную последовательность $(\tau_n, n \geq 1)$

моментов остановки с единственным условием, чтобы $\tau_n \uparrow \infty$ при $n \rightarrow \infty$. Поскольку $a_n \rightarrow 0$ при $n \rightarrow \infty$, то $X_{\tau_n} \xrightarrow{\text{п.н.}} 0$. Ввиду того, что $|a_n| \leq 1$, заключаем, что $EX_{\tau_n} \rightarrow 0$ при $n \rightarrow \infty$. Следовательно, для любой фильтрации $(\mathcal{F}_n, n \geq 1)$, с которой связана последовательность $(\tau_n, n \geq 1)$, система $(X_n, \mathcal{F}_n, n \geq 1)$ является амартом. Однако

$$\sum_{n=1}^{\infty} E|X_n - E(X_{n+1}|\mathcal{F}_n)| = \sum_{n=1}^{\infty} |a_n - a_{n+1}| = \infty,$$

и, значит, амарт $(X_n, \mathcal{F}_n, n \geq 1)$ не является квазимартингалом.

(б) Рассмотрим последовательность $(X_n, n \geq 1)$, состоящую из независимых и одинаково распределенных случайных величин, где $P\{X_n = \pm 1\} = \frac{1}{2}$. Пусть $c_n, n \geq 1$ — такие положительные числа, что $c_n \downarrow 0$ при $n \rightarrow \infty$ и $\sum_{n=1}^{\infty} c_n = \infty$ (возьмем, например, $c_n = \frac{1}{n}$). Определим новую последовательность $(Y_n, n \geq 1)$, где $Y_n = c_n X_1 \dots X_n$, и пусть $\mathcal{F}_n = \sigma\{X_1, \dots, X_n\}$. Тогда ясно, что Y_n измеримо относительно $\mathcal{F}_n$ при каждом $n \geq 1$. Поскольку $|Y_n| \leq c_n \downarrow 0$, то $Y_{\tau_n} \xrightarrow{\text{п.н.}} 0$ при $n \rightarrow \infty$ для любой последовательности моментов остановки $(\tau_n, n \geq 1)$ такой, что $\tau_n \uparrow \infty$ при $n \rightarrow \infty$. Применяя теорему о мажорируемой сходимости, заключаем, что $EY_n \rightarrow 0$ при $n \rightarrow \infty$. Таким образом, $Y = (Y_n, \mathcal{F}_n, n \geq 1)$ является амартом. Однако

$$\sum_{n=1}^{\infty} E|Y_n - E(Y_{n+1}|\mathcal{F}_n)| = \sum_{n=1}^{\infty} E|Y_n| = \sum_{n=1}^{\infty} c_n = \infty,$$

и, следовательно, амарт Y квазимартингалом не является.

22.4. Амарты, мартингалы в пределе, эвентуальные мартингалы и связи между ними

(а) Пусть $\xi_1, \xi_2, \dots$ — положительные, независимые и одинаково распределенные случайные величины, для которых выполнены условия $E\xi_1 < \infty$ и $E\{\xi_1 \ln^+ \xi_1\} = \infty$ (напомним, что $\ln^+ x = \ln x$ при $x \geq 1$ и $\ln^+ x = 0$ при $0 < x < 1$). Рассмотрим последовательность случайных величин $(X_n, n \geq 1)$ и фильтрацию $(\mathcal{F}_n, n \geq 1)$, где $X_n = \frac{\xi_n}{n}$ и $\mathcal{F}_n = \sigma\{\xi_1, \dots, \xi_n\}$. Легко видеть, что $X_n \xrightarrow{\text{п.н.}} 0$ при $n \rightarrow \infty$. Более того, $EX_n \rightarrow 0$ при $n \rightarrow \infty$, но $E\{\sup_{n \geq 1} X_n\} = \infty$. Отсюда вытекает, что $X = (X_n, \mathcal{F}_n, n \geq 1)$ является мартингалом в пределе. Однако X не является амартом, потому что сеть $(EX_\tau, \tau \in T)$ неограниченна (напомним, что T — множество всех $(\mathcal{F}_n)$ -моментов остановки).

(б) Рассмотрим последовательность независимых случайных величин $(\eta_n, n \geq 1)$, где $P\{\eta_n = 1\} = \frac{1}{n^2} = 1 - P\{\eta_n = 0\}$. Пусть $X_n = \eta_1 + \dots + \eta_n$, $\mathcal{F}_n = \sigma\{\eta_1, \dots, \eta_n\}$. Так как

$$E(X_n | \mathcal{F}_m) - X_m = \sum_{k=m+1}^{\infty} \frac{1}{k^2} \implies \lim_{n \leq m \rightarrow \infty} (E(X_n | \mathcal{F}_m) - X_m) = 0 \text{ п. н.,}$$

то $X = (X_n, \mathcal{F}, n \geq 1)$ — мартингал в пределе. Далее, соотношения

$$E\left\{\sum_{k=1}^{\infty} |\eta_k|\right\} = \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty \quad \text{и} \quad |X_n| \leq \sum_{k=1}^{\infty} |\eta_k|, \quad n \geq 1,$$

показывают, что X равномерно интегрируем. Несмотря на эти свойства, X не является эвентуальным мартингалом. Это вытекает из соотношений $E(X_n | \mathcal{F}_{n-1}) = X_{n-1} + \frac{1}{n^2} \neq X_{n-1}$, справедливых для всех $n \geq 2$, и определения (д) (см. вводные замечания к настоящему параграфу).

22.5. Квазимартингалы, амарты, прогрессивные мартингалы и связи между ними

(а) Пусть $\{\xi_n, n \geq 1\}$ — независимые случайные величины, где $P\{\xi_n = 1\} = \frac{n}{n+1}$, $P\{\xi_n = 0\} = \frac{1}{n}$. Обозначим $\eta_1 = 1$, $\eta_n = (-1)^{n-1} \times \xi_1 \xi_2 \dots \xi_{n-1}$ при $n \geq 2$, $X_n = \eta_1 + \dots + \eta_n$ и $\mathcal{F}_n = \sigma\{\xi_1, \dots, \xi_n\}$. Лемма Бореля–Кантелли показывает, что $P\{\xi_n = 0 \text{ б. ч.}\} = 1$, а тогда $P\{\eta_n \neq 0 \text{ б. ч.}\} = 0$. Легко видеть, что X_n принимает лишь два значения: 0 и 1. Если $\eta_n = 0$ для некоторого n , то $\eta_{n+1} = 0$, а в таком случае $X_{n+1} = X_n$. Отсюда заключаем, что $X = (X_n, \mathcal{F}_n, n \geq 1)$ — прогрессивный мартингал.

Проверим теперь, является ли X квазимартингалом. Мы имеем

$$\begin{aligned} \sum_{n=1}^{\infty} E\{|X_n - E(X_{n+1} | \mathcal{F}_n)|\} &= \sum_{n=1}^{\infty} E\{|\eta_{n+1}| | \mathcal{F}_n|\} = \\ &= \sum_{n=1}^{\infty} E|\eta_{n+1}| = \sum_{n=1}^{\infty} \left(\prod_{k=1}^n \frac{k}{k+1} \right) = \sum_{n=1}^{\infty} \frac{1}{n+1} = \infty. \end{aligned}$$

Следовательно, прогрессивный мартингал X не является квазимартингалом.

(б) Приведем теперь пример прогрессивного мартингала, не являющегося амартом. Для этой цели рассмотрим последовательность независимых случайных величин $(\xi_n, n \geq 1)$, где $P\{\xi_n = 1\} = \frac{n}{n+1} = 1 - P\{\xi_n = 0\}$. Положим $X_1 = 1$, $X_n = n^2 \xi_1 \xi_2 \dots \xi_{n-1}$, $n \geq 2$,

и $\mathcal{F}_n = \sigma\{\xi_1, \dots, \xi_n\}$, $n \geq 1$. Очевидно, что

$$\mathbb{E}\{X_n | \mathcal{F}_{n-1}\} = X_n = \frac{n^2}{(n-1)^2} \xi_{n-1} X_{n-1} \text{ п. н.}$$

Согласно лемме Бореля–Кантелли, $\mathbb{P}\{\xi_n = 0 \text{ б. ч.}\} = 1$, и так как $X_{n-1} = 0$ влечет, что $X_n = 0$, заключаем, что $\mathbb{P}\{X_n \neq 0 \text{ б. ч.}\} = 0$. Следовательно, $X = (X_n, \mathcal{F}_n, n \geq 1)$ — прогрессивный мартингал. Однако, $\mathbb{E}X_n = n \rightarrow \infty$ при $n \rightarrow \infty$, и, значит, X не может быть амартом.

22.6. О сходимости субмартингалов почти наверное и в пространстве $\mathcal{L}^1$

Начнем с формулировки следующего результата (*теорема Дуба*): пусть $(X_n, \mathcal{F}_n, n \geq 1)$ — субмартингал, удовлетворяющий условию

$$\sup_{n \geq 1} \mathbb{E}|X_n| < \infty. \quad (1)$$

Тогда предел $X_\infty = \lim_{n \rightarrow \infty} X_n$ существует п. н. и $\mathbb{E}|X_\infty| < \infty$. Более того, если субмартингал $(X_n, \mathcal{F}_n, n \geq 1)$ равномерно интегрируемый, то кроме сходимости п. н. имеет место и сходимость $X_n \xrightarrow{L^1} X_\infty$ при $n \rightarrow \infty$. Доказательство этого и ряда других близких результатов может быть найдено, например, в [32, 116, 177, 381]. Рассмотренные далее примеры показывают, какова роль наложенных условий для обеспечения одного или другого типа сходимости.

(а) Пусть $(\xi_n, n \geq 1)$ — независимые и одинаково распределенные случайные величины с $\mathbb{P}\{\xi_1 = 0\} = \mathbb{P}\{\xi_1 = 2\} = \frac{1}{2}$. Положим $X_n = \xi_1 \dots \xi_n$ и $\mathcal{F}_n = \sigma\{\xi_1, \dots, \xi_n\}$, $n \geq 1$. Легко видеть, что $(X_n, \mathcal{F}_n, n \geq 1)$ — мартингал, причем $\mathbb{E}X_n = 1, n \geq 1$. Значит, условие (1) выполнено, и существует предел $X_\infty = \lim_{n \rightarrow \infty} X_n$, где X_∞ — случайная величина с $\mathbb{E}|X_\infty| < \infty$. Проверим, имеет ли место сходимость $X_n \xrightarrow{L^1} X_\infty$. Поскольку $\mathbb{P}\{X_n = 2^n\} = 2^{-n}$ и $\mathbb{P}\{X_n = 0\} = 1 - 2^{-n}$, заключаем, что $X_\infty = 0$ п. н. Однако $\mathbb{E}|X_n - X_\infty| = \mathbb{E}X_n = 1$. Отсюда вытекает, что $X_n \not\xrightarrow{L^1} X_\infty$ при $n \rightarrow \infty$.

(б) Вероятностное пространство $(\Omega, \mathcal{F}, \mathbb{P})$ задано следующим образом: $\Omega = [0, 1]$, $\mathcal{F} = \mathcal{B}_{[0,1]}$ и $\mathbb{P}$ — мера Лебега. На этом пространстве рассмотрим случайную последовательность $(X_n, n \geq 1)$:

$$X_n = X_n(\omega) = \begin{cases} 2^n, & \text{если } \omega \in \left[0, \frac{1}{2^n}\right], \\ 0, & \text{если } \omega \in \left(\frac{1}{2^n}, 1\right]. \end{cases}$$

Введем фильтрацию $(\mathcal{F}_n, n \geq 1)$, где $\mathcal{F}_n = \sigma\{X_1, \dots, X_n\}$. Тогда $(X_n, \mathcal{F}_n, n \geq 1)$ — мартингал, причем $\mathbb{E}X_n = 1, n \geq 1$. Снова выполнено условие (1), и, следовательно, $X_n \xrightarrow{\text{п.н.}} X_\infty$ при $n \rightarrow \infty$ с $X_\infty = 0$. Как и в предыдущем примере, мы имеем $X_n \xrightarrow{L^1} X_\infty$ при $n \rightarrow \infty$.

Итак, в обоих примерах (а) и (б) имеет место сходимость п. н., но нет сходимости в пространстве $\mathcal{L}^1$. Существенным является то, что описанные мартингалы не интегрируемы равномерно. Свойство равномерной интегрируемости тесно связано со свойством регулярности, и этому посвящен следующий пример.

(в) Напомним, что мартингал $X = (X_n, \mathcal{F}_n, n \geq 1)$ называется регулярным, если существует интегрируемая случайная величина ξ такая, что $X_n = \mathbb{E}(\xi | \mathcal{F}_n)$ п. н. для каждого $n \geq 1$. Если параметр n принимает только конечное число значений, например, $n = 1, \dots, N$, то мартингал регулярен, поскольку $X_n = \mathbb{E}(X_N | \mathcal{F}_n)$. Оказывается, однако, что если $n \in \mathbb{N}$, то мартингал не обязан быть регулярным. Полезно сформулировать один из важнейших результатов, касающихся свойства регулярности (см. [116]): *мартингал $X = (X_n, \mathcal{F}_n, n \geq 1)$ регулярен тогда и только тогда, когда он равномерно интегрируем. При этом имеет место представление $X_n = \mathbb{E}(X_\infty | \mathcal{F}_n)$ п. н., $n \geq 1$, где $X_\infty = \lim_{n \rightarrow \infty} X_n$.*

Рассмотрим последовательность $(\xi_k, k \geq 1)$ независимых случайных величин, каждая с распределением $\mathcal{N}(0, 1)$. Положим

$$S_n = \xi_1 + \dots + \xi_n, \quad X_n = \exp\left(S_n - \frac{1}{2}n\right), \\ \mathcal{F}_n = \sigma\{\xi_1, \dots, \xi_n\}, \quad n \geq 1.$$

Легко проверить, что $X = (X_n, \mathcal{F}_n, n \geq 1)$ — мартингал. Предел $X_\infty = \lim_{n \rightarrow \infty} X_n$ проще всего найти применением УЗБЧ к последовательности $(\xi_n, k \geq 1)$. Мы имеем

$$X_\infty = \lim_{n \rightarrow \infty} \exp\left[n\left(\frac{1}{n}S_n - \frac{1}{2}\right)\right] = 0 \text{ п. н.}$$

Поскольку $\mathbb{E}(X_\infty | \mathcal{F}_n) = 0$, то

$$X_n \neq \mathbb{E}(X_\infty | \mathcal{F}_n).$$

Следовательно, мартингал X не обладает свойством регулярности. Это связано с фактом, что X не равномерно интегрируем.

22.7. Мартингалы могут сходиться по вероятности, но не сходиться почти наверное

Напомним, что если $\xi_1, \xi_2, \dots$ — независимые случайные величины, то при $n \rightarrow \infty$ сходимость по вероятности для сумм $S_n = \xi_1 + \dots + \xi_n$

эквивалентна сходимости п. н. (см. [116, 177, 294, 404]). С другой стороны, в примере 14.3 мы описали конкретные последовательности случайных величин, сходящихся по вероятности, но не сходящихся п. н. Естественен вопрос о сходимости мартингалов в том или другом смысле.

(а) Пусть задана последовательность $(\xi_n, n \geq 1)$, состоящая из независимых случайных величин, причем

$$P\{\xi_n = 1\} = P\{\xi_n = -1\} = \frac{1}{2n}, \quad P\{\xi_n = 0\} = 1 - \frac{1}{n}.$$

Нетрудно видеть, что $\xi_n \xrightarrow{P} 0$ при $n \rightarrow \infty$, но $\xi_n \not\xrightarrow{п.н.} 0$. Рассмотрим новую последовательность $(X_n, n \geq 1)$, где $X_1 = \xi_1$, а при $n \geq 2$

$$X_n = \begin{cases} \xi_n, & \text{если } X_{n-1} = 0, \\ nX_{n-1}|\xi_n|, & \text{если } X_{n-1} \neq 0. \end{cases}$$

Положим $\mathcal{F}_n = \sigma\{\xi_1, \dots, \xi_n\}$. Тогда имеют место утверждения:

(1) $X = (X_n, \mathcal{F}_n, n \geq 1)$ является мартингалом.

(2) Для каждого $n \geq 1$ $\{X_n = 0\} \Leftrightarrow \{\xi_n = 0\}$.

(3) $P\{X_n = 0\} = P\{\xi_n = 0\} = 1 - \frac{1}{n}$.

(4) $P\{X_n \neq 0 \text{ б. ч.}\} = P\{\xi_n \neq 0 \text{ б. ч.}\} = 1$.

Отметим, что мартингаловое свойство (1) проверяется непосредственно, (2) и (3) — очевидны, а (4) вытекает из соотношения $\sum_{n=1}^{\infty} P\{|\xi_n| = 1\} = \infty$ и леммы Бореля–Кантелли.

Мы интересуемся, каково поведение X_n при $n \rightarrow \infty$. Из (3) следует, что $X_n \xrightarrow{P} 0$. Однако (4) показывает, что

$$P\{\omega: X_n(\omega) \text{ сходится}\} = 0.$$

Значит, мартингал X сходится по вероятности, но не сходится п. н.

(б) Рассмотрим последовательность независимых случайных величин $(\xi_n, n \geq 1)$, где ξ_n принимает значения ± 1 с вероятностью $\frac{1}{2}$. Положим $\mathcal{F}_n = \sigma\{\xi_1, \dots, \xi_n\}$, и пусть $(B_n, n \geq 1)$ — такая последовательность событий, что $B_n \in \mathcal{F}_n$ для каждого $n \geq 1$, причем выполнены условия

$$\lim_{n \rightarrow \infty} P(B_n) = 0 \quad \text{и} \quad P(\sup_{n \rightarrow \infty} B_n) = 1.$$

Введем новую последовательность $(X_n, n \geq 1)$, где $X_1 = 0$, а

$$X_{n+1} = X_n(1 + \xi_{n+1}) + \xi_{n+1} \cdot 1_{B_n}, \quad n \geq 1.$$

Легко проверить, что $X = (X_n, \mathcal{F}_n, n \geq 1)$ — мартингал. Поскольку

$$P\{X_{n+1} \neq 0\} \leq \frac{1}{2}P\{X_n \neq 0\} + P(B_n),$$

закключаем, что

$$\lim_{n \rightarrow \infty} P(X_n = 0) = 1 \quad \text{и} \quad P\{\omega: X_n(\omega) \text{ сходится}\} = 0.$$

Итак, мы снова, как и в примере (а), получаем мартингал, который расходится п. н., хотя он сходится по вероятности к нулю.

(в) Существование мартингалов, обладающих специфическими свойствами, можно доказать с помощью следующего общего результата (см. [152]): *пусть основное вероятностное пространство $(\Omega, \mathcal{F}, P)$ состоит из $\Omega = [0, 1]$, $\mathcal{F} = \mathcal{B}_{[0,1]}$ и P — мера Лебега. Тогда для любой последовательности $(\xi_n, n \geq 1)$, состоящей из элементарных случайных величин, существует мартингал $(X_n, \mathcal{F}_n, n \geq 1)$ такой, что $P\{\omega: \xi_n(\omega) = X_n(\omega) \text{ для всех достаточно больших } n\} = 1$.* (Напомним, что элементарными (простыми) называются случайные величины, принимающие конечное число значений.) Мы видели в § 14, что имеются элементарные случайные величины, сходящиеся по вероятности, но не п. н., а также величины, которые ограничены, но не сходящиеся. Следовательно, будут иметь место следующие два утверждения:

(I) Существует мартингал $(X_n, \mathcal{F}_n, n \geq 1)$, такой что

$$X_n \xrightarrow{P} 0 \text{ при } n \rightarrow \infty, \text{ но } P\{\omega: X_n(\omega) \text{ сходится к } 0\} = 0.$$

(II) Существует мартингал $(X_n, \mathcal{F}_n, n \geq 1)$, такой что

$$P\{\omega: (X_n, n \geq 1) \text{ ограничена}\} = 1, \text{ но } P\{\omega: X_n(\omega) \text{ сходится}\} = 0.$$

22.8. Мартингалы могут расходиться на множестве с заранее заданной вероятностью

В примере 22.7 мы описали мартингалы, которые расходятся п. н., т. е. их множество расходимости имеет вероятность, равную 1 (а тогда множество сходимости имеет вероятность 0). Покажем теперь, что для каждого $p \in [0, 1]$ можно построить мартингал, множество расходимости которого имеет вероятность p .

Пусть $(\xi_n, n \geq 1)$ — независимые и одинаково распределенные случайные величины с $E\xi_1 = 0$ и $E|\xi_1| > 0$, а $(\eta_n, n \geq 1)$ — другая последовательность независимых случайных величин, для которой $E\eta_n = 0$, $E\{\eta_n^2\} = n^{-2}$, $n \geq 1$. Рассмотрим случайные ряды $\sum_{n=1}^{\infty} \xi_n$ и $\sum_{n=1}^{\infty} \eta_n$. Согласно одному из результатов в работе [182], ряд $\sum_{n=1}^{\infty} \xi_n$ расходится п. н.

С другой стороны, из теоремы Колмогорова о трех рядах (см. § 18) следует, что ряд $\sum_{n=1}^{\infty} \eta_n$ сходится п. н.

Предположим, что последовательности (ξ_n) и (η_n) независимы между собой и, более того, что они не зависят от случайной величины X_0 , где $P\{X_0 = 1\} = p$, $P\{X_0 = -1\} = 1 - p$. Здесь p — фиксированное число из отрезка $[0, 1]$. Вводим новую последовательность $(X_n, n \geq 1)$, где

$$X_n = \xi_n I(X_0 = 1) + \eta_n I(X_0 = -1).$$

Положим $S_n = X_1 + \dots + X_n$ и $\mathcal{F}_n = \sigma\{X_1, \dots, X_n\}$. Легко видеть, что $(S_n, \mathcal{F}_n, n \geq 1)$ является мартингалом, причем $ES_n = 0, n \geq 1$. Естественно возникает вопрос о предельном поведении последовательности (S_n) при $n \rightarrow \infty$. Из равенства

$$S_n = I(X_0 = 1) \sum_{k=1}^n \xi_k + I(X_0 = -1) \sum_{k=1}^n \eta_k$$

вытекает, что

$$\begin{aligned} P\{S_n \text{ сходится}\} &= P\{X_0 = -1\} = 1 - p, \\ P\{S_n \text{ расходится}\} &= P\{X_0 = 1\} = p. \end{aligned}$$

Итак, построенный нами мартингал $(S_n, \mathcal{F}_n, n \geq 1)$ расходится на множестве, вероятность которого равна заранее заданному числу p .

§ 23. Мартингалы с непрерывным временем

Предположим, что заданы полное вероятностное пространство $(\Omega, \mathcal{F}, P)$ и фильтрация $(\mathcal{F}_t, t \geq 0)$, удовлетворяющая стандартным условиям: $\mathcal{F}_t \subset \mathcal{F}$ для каждого t ; $\mathcal{F}_s \subset \mathcal{F}_t$ при $s < t$; $\mathcal{F}_t, t \geq 0$, содержат все P -пренебрежимые множества из $\mathcal{F}$; $(\mathcal{F}_t)$ непрерывна справа. Как обычно, запись $(X_t, \mathcal{F}_t, t \geq 0)$ означает, что процесс $(X_t, t \geq 0)$ адаптирован к фильтрации $(\mathcal{F}_t, t \geq 0)$.

Случайный процесс $X = (X_t, \mathcal{F}_t, t \geq 0)$ с $E|X_t| < \infty$ для всех $t \geq 0$ называется *мартингалом*, *субмартингалом* или *супермартингалом*, если для любых $s < t$ соответственно выполнены P -п.н. соотношения

$$E\{X_t | \mathcal{F}_s\} = X_s, \quad E\{X_t | \mathcal{F}_s\} \geq X_s \quad \text{или} \quad E\{X_t | \mathcal{F}_s\} \leq X_s.$$

Мартингал $M = (M_t, \mathcal{F}_t, t \geq 0)$ называется $\mathcal{L}^p$ -мартингалом, $p \geq 1$, если $E\{|X_t|^p\} < \infty$ для всех $t \geq 0$. В случае $p = 2$ используется термин *квадратично-интегрируемый мартингал*.

Марковский момент относительно фильтрации $(\mathcal{F}_t)$ — это случайная величина $T(\omega)$, $\omega \in \Omega$, принимающая значения в множестве

$\mathbb{R}^+ \cup \{\infty\}$ и такая, что $\{T \leq t\} \in \mathcal{F}_t$ при каждом $t \geq 0$. Если $P\{T < \infty\} = 1$, T называется *моментом остановки*.

Непрерывный справа процесс $X = (X_t, \mathcal{F}_t, t \geq 0)$ называется *локальным мартингалом*, если существует возрастающая последовательность $(T_n, n \geq 1)$ марковских моментов относительно $(\mathcal{F}_t)$, $T_n \xrightarrow{п.н.} \infty$ при $n \rightarrow \infty$ и таких, что остановленный процесс $(X_{t \wedge T_n}, \mathcal{F}_t, t \geq 0)$ является равномерно интегрируемым мартингалом. Далее, X *локально квадратично-интегрируем*, если $(X_{t \wedge T_n}, \mathcal{F}_t, t \geq 0)$ — квадратично-интегрируемый мартингал, т.е. если $E\{X_{t \wedge T_n}^2\} < \infty$ для каждого n .

Если $M = (M_t, \mathcal{F}_t, t \geq 0)$ — квадратично-интегрируемый мартингал, то существует единственный предсказуемый возрастающий процесс $\langle M \rangle = (\langle M \rangle_t, \mathcal{F}_t, t \geq 0)$, называемый *квадратической вариацией*, такой что $(M_t^2 - \langle M \rangle_t, \mathcal{F}_t, t \geq 0)$ — мартингал.

Предположим, что процесс $X = (X_t, \mathcal{F}_t, t \geq 0)$ имеет траектории в пространстве $\mathbb{D}$ (т.е. траектории X непрерывны справа и имеют пределы слева), а фильтрация $(\mathcal{F}_t)$ удовлетворяет обычным условиям (для простоты считаем, что $\mathcal{F}_{0-} = \mathcal{F}_0$, $\mathcal{F}_{\infty-} = \mathcal{F}$). Процесс X называется *сесми-мартингалом*, если он допускает следующее представление:

$$X_t = X_0 + M_t + A_t, \quad t \geq 0,$$

где $M = (M_t, \mathcal{F}_t, t \geq 0)$ — локальный мартингал с $M_0 = 0$ и $A = (A_t, \mathcal{F}_t, t \geq 0)$, $A_0 = 0$ — непрерывный справа процесс локально ограниченной вариации.

Отметим, что теории мартингалов и ее приложениям посвящена значительная литература. Читатель может обратиться, например, к следующим книгам: [74, 120, 199, 212, 298, 368].

Мы выбрали несколько примеров, иллюстрирующих связи между близкими классами процессов и их свойства. Эти примеры можно рассматривать параллельно с примерами из § 22.

23.1. Мартингалы, не являющиеся локально квадратично-интегрируемыми

Покажем, что существует мартингал $X = (X_t, \mathcal{F}_t, t \geq 0)$ такой, что $E\{X_T^2\} = \infty$ для любого момента остановки T относительно $(\mathcal{F}_t)$. Очевидно, что такой мартингал не может быть локально квадратично-интегрируемым.

(а) Рассмотрим однородный пуассоновский процесс $N = (N_t, t \geq 0)$ с параметром 1, и пусть τ — момент первого скачка процесса N . Напомним, что случайная величина τ имеет экспоненциальное распределение с параметром 1. Пусть $(\mathcal{F}_t, t \geq 0)$ — фильтрация, порожденная

процессом $N: \mathcal{F}_t = \sigma\{N_s, s \leq t\}$. Введем процесс $m = (m_t, t \geq 0)$, где

$$m_t = \tau^{-1/2} I_{\{\tau \leq t\}} - 2\sqrt{\tau \wedge t}.$$

Согласно одному результату работы [52], процесс m допускает следующее представление в виде интеграла Стильбеса:

$$m_t = \int_0^t s^{-1/2} I_{\{\tau \geq s\}} (dN_s - ds).$$

Из этого представления нетрудно заключить, что $(m_t, \mathcal{F}_t, t \geq 0)$ является мартингалом. Он обладает рядом интересных свойств, но для нас важно установить, является ли процесс m локально квадратично-интегрируемым.

Чтобы ответить на этот вопрос, воспользуемся следующим утверждением (см. [198]): *случайная величина T является моментом остановки относительно фильтрации $(\mathcal{F}_t)$ тогда и только тогда, когда существует число $u \in \mathbb{R}^+ \cup \{\infty\}$ такое, что $T \geq 0$ почти наверное на множестве $\{\tau \leq u\}$ и $T = u$ п. н. на множестве $\{\tau > u\}$* . Возьмем теперь произвольный $(\mathcal{F}_t)$ -момент остановки T . Из цитированного результата вытекает, что $T \wedge \tau = c \wedge \tau$ для некоторой константы c . Более того, для любого $c > 0$ мы имеем $E\{(\tau^{-1/2} I_{\{\tau \leq c\}})^2\} = \infty$. Следовательно, $E\{m_T^2\} = \infty$, и, значит, мартингал m не обладает свойством локальной квадратичной интегрируемости.

(б) Пусть на полном вероятностном пространстве $(\Omega, \mathcal{F}, P)$ задана случайная величина ξ , удовлетворяющая условиям: $E|\xi| < \infty$ и $E\{\xi^2\} = \infty$. (Такие примеры читатель укажет без особого труда.) Определим фильтрацию $(\mathcal{F}_t^0, t \geq 0)$ следующим образом: $\mathcal{F}_t^0 = \{\emptyset, \Omega\}$, если $0 \leq t < 1$, и $\mathcal{F}_t^0 = \mathcal{F}$, если $t \geq 1$, и пусть $\mathcal{F}_t$ — пополнение $\mathcal{F}_t^0$ пренебрежимыми множествами из $\mathcal{F}$. Положим $M_t = E(\xi | \mathcal{F}_t)$, рассматривая непрерывную справа версию этого условного математического ожидания. Тогда $M = (M_t, \mathcal{F}_t, t \geq 0)$ — мартингал, который ввиду выбора случайной величины ξ , очевидно, не является локально квадратично-интегрируемым.

23.2. L^2 -ограниченный локальный мартингал, который не является мартингалом

Пусть $w(t) = (w_1(t), w_2(t), w_3(t))$, $t \geq 0$, — стандартный винеровский процесс в пространстве $\mathbb{R}^3$, начинающийся из точки $O = (0, 0, 0)$. Рассмотрим бесселевский процесс $Z = \left(Z_t = \frac{1}{|w(t)|}, t > 0\right)$, где $|w(t)| = (w_1^2(t) + w_2^2(t) + w_3^2(t))^{1/2}$ (см. [77, 304, 408]). Введем фильтрацию

$\mathcal{F}^w = (\mathcal{F}_t^w)_{t \geq 0}$, порожденную процессом w . Используя явный вид распределения Z_t , мы получаем

$$\begin{aligned} \mathbb{E}\{Z_t^2\} &= \int_0^\infty r^{-2} \cdot 4\pi r^2 (2\pi t)^{-3/2} e^{-r^2/2t} dr = \\ &= 4\pi (2\pi t)^{-1} \int_0^\infty (2\pi t)^{-1/2} e^{-r^2/2t} dr = 4\pi (2\pi t)^{-1} \cdot \frac{1}{2} = t^{-1}. \end{aligned} \quad (1)$$

Вводим новый процесс $Y = (Y_t, t \geq 0)$, а также новую фильтрацию $\mathcal{Y} = (\mathcal{Y}_t, t \geq 0)$:

$$Y_t = Z_{t+1}, \quad \mathcal{Y}_t = \mathcal{F}_{t+1}^w.$$

Очевидно, что процесс Y согласован с фильтрацией $\mathcal{Y}$, и далее мы хотим установить другие его свойства.

Хорошо известно, что функция $\frac{1}{|x|}$, $x \in \mathbb{R}^3 \setminus \{0\}$, является гармонической. Кроме того, винеровский процесс w (в пространстве $\mathbb{R}^3$) хотя и начинается из точки 0 в момент $t = 0$, с вероятностью 1, в 0 не возвращается (см. [34, 70, 157, 238]). Поэтому можно применить формулу Ито (см. [34, 73, 120, 304, 330, 408]), чтобы найти представление для Z_t в виде стохастического дифференциала и потом заключить, что процесс $(Y_t, \mathcal{Y}_t, t \geq 0)$ является локальным мартингалом. Далее, случайная величина Y_0 интегрируема, как это видно из (1), и лемма Фату позволяет показать, что $\mathbb{E}(Y_t | \mathcal{Y}_s) \leq Y_s$ п. н. при $s < t$, т. е. $(Y_t, \mathcal{Y}_t, t \geq 0)$ является также и положительным супермартингалом. Наконец, снова из (1) вытекает L^2 -ограниченность процесса Y . Таким образом, семейство $(Y_t, t \geq 0)$ равномерно интегрируемо. Из того факта, что Y — супермартингал, и из (1) находим, что $Y_t \xrightarrow{\text{п.н.}} 0$ при $t \rightarrow \infty$.

Предположим, что процесс Y — мартингал. Тогда ввиду равномерной интегрируемости мы имели бы

$$Y_t = \mathbb{E}(Y_\infty | \mathcal{Y}_t) = 0 \quad \text{п. н.,}$$

что, очевидно, быть не может. Следовательно, процесс Y мартингалом не является, хотя, повторим это, Y — локальный мартингал, а также положительный супермартингал.

Отметим, наконец, еще один заслуживающий внимания факт. Напомним, что класс (D) (класс Дирихле) — это совокупность всех измеримых L^1 -ограниченных процессов $X = (X_t, \mathcal{F}_t, t \geq 0)$ таких, что семейство случайных величин $\{X_\tau, \tau \in T\}$ равномерно интегрируемо. (Как обычно, T — множество всех $(\mathcal{F}_t)$ -моментов остановки.) По поводу этого класса (D) читатель может обратиться к работам [308, 413].

23.3. Экспоненциальные мартингалы:**неулучшаемость условия Новикова и два полезных следствия**

Пусть заданы вероятностное пространство $(\Omega, \mathcal{F}, P)$, фильтрация $(\mathcal{F}_t, t \geq 0)$, удовлетворяющая обычным условиям, стандартный винеровский процесс $w = (w_t, \mathcal{F}_t, t \geq 0)$ и процесс $\beta = (\beta_t, \mathcal{F}_t, 0 \leq t \leq T)$ такой, что

$$P\left\{\int_0^T \beta_s^2 ds < \infty\right\} = 1.$$

Случайный процесс $X = (X_t, \mathcal{F}_t, 0 \leq t \leq T)$, где

$$X_t = X_t(\beta) = \exp\left(\int_0^t \beta_s dw_s - \frac{1}{2} \int_0^t \beta_s^2 ds\right), \quad (1)$$

называется *экспоненциальным процессом*. Очевидно, что X — неотрицательный супермартингал. Оказывается, что при одном дополнительном условии X будет мартингалом. Имеет место следующий результат (см. [73, 82, 304, 408]): *предположим, при введенных выше обозначениях, что выполнено условие (условие Новикова)*

$$E\left\{\exp\left(\frac{1}{2} \int_0^T \beta_s^2 ds\right)\right\} < \infty. \quad (2)$$

Тогда $EX_t = 1$ для каждого $t \leq T$ и экспоненциальный процесс $X = (X_t, \mathcal{F}_t, 0 \leq t \leq T)$, определяемый равенством (1), является мартингалом. Отметим, что для мартингала X используется термин *экспоненциальный мартингал*. Поскольку условие Новикова (2) достаточно для того, чтобы экспоненциальный процесс (1) был мартингалом, хотелось бы выяснить, можно ли ослабить это условие. Построим далее такой пример, где условие

$$E\left\{\exp\left(\left(\frac{1}{2} - \varepsilon\right) \int_0^T \beta_s^2 ds\right)\right\} < \infty \quad (3)$$

будет выполняться при каждом $\varepsilon > 0$, но при этом $EX_T(\beta) < 1$, и, значит, процесс $X(\beta)$ не будет мартингалом.

Итак, пусть $\tau = \inf\{t < 1: w_t^2 = 1 - t\}$, а процесс $\gamma = (\gamma_t, 0 \leq t \leq 1)$ определим следующим образом:

$$\gamma_t = \begin{cases} -\frac{2w_t}{(1-t)^2} \cdot 1_{\{\tau \geq t\}}, & \text{если } 0 \leq t < 1, \\ 0, & \text{если } t = 1. \end{cases}$$

Поскольку $P\{0 < \tau < 1\} = 1$, то P -п.н.

$$\int_0^1 \gamma_t^2 dt = 4 \int_0^1 \frac{w_t^2}{(1-t)^4} \cdot 1_{\{\tau \geq t\}} dt = 4 \int_0^\tau \frac{w_t^2}{(1-t)^4} dt < \infty, \quad (4)$$

и, следовательно, по формуле (1) можно корректно задать экспоненциальный процесс $X = X(\gamma)$ (с функцией γ). Применяя формулу Ито, легко найти, что при $0 < t < 1$

$$d\left(\frac{w_t^2}{(1-t)^2}\right) = \frac{2w_t^2}{(1-t)^3} dt + \frac{2w_t}{(1-t)^2} dt + \frac{1}{(1-t)^2} dt. \quad (5)$$

Комбинируя (4) и (5), получаем

$$\begin{aligned} \int_0^1 \gamma_s dw_s - \frac{1}{2} \int_0^1 \gamma_s^2 ds &= \\ &= -\frac{w_\tau^2}{(1-\tau)^2} + \int_0^\tau \frac{2w_t^2}{(1-t)^3} dt + \int_0^\tau \frac{dt}{(1-t)^2} - \int_0^\tau \frac{2w_t^2}{(1-t)^4} dt. \end{aligned}$$

После преобразований (детали опускаем) приходим к соотношению

$$\int_0^1 \gamma_s dw_s - \frac{1}{2} \int_0^1 \gamma_s^2 ds = -1 - 2 \int_0^\tau \left(\frac{1}{(1-t)^4} - \frac{1}{(1-t)^3} \right) w_t^2 dt \leq -1.$$

Отсюда вытекает, что $EX_1(\gamma) \leq e^{-1}$. Следовательно, неотрицательный супермартингал $X(\gamma) = (X_t(\gamma), \mathcal{F}_t, 0 \leq t \leq 1)$ не является мартингалом. Условие Новикова (2) не выполнено, но имеет место условие (3). Это говорит о неулущаемости условия Новикова.

Здесь можно сделать еще один шаг. Рассмотрим последовательность моментов остановки $(\tau_n = 1 - n^{-1/2}, n \geq 1)$. Очевидно, что τ_n возрастают при $n \rightarrow \infty$ и стремятся к 1 — это правый конец интервала, где меняется параметр времени t . Легко видеть, что при каждом $n \geq 1$ остановленный процесс $(X_{t \wedge \tau_n}, \mathcal{F}_t, 0 \leq t \leq 1)$ — мартингал. Это означает, что процесс $X(\gamma)$ — локальный мартингал.

Итак, с одной стороны, мы продемонстрировали, что в указанном смысле условие Новикова неулущаемо. С другой стороны, мы построили положительный супермартингал, не являющийся мартингалом. Более того, тот самый процесс, не будучи мартингалом, является локальным мартингалом.

23.4. Квадратично-интегрируемый мартингал с неслучайной характеристикой не обязан быть процессом с независимыми приращениями

Начнем с формулировки следующего результата П. Леви (см. [27, 74, 298]): пусть $X = (X_t, \mathcal{F}_t, t \geq 0)$ — квадратично-интегрируемый мартингал, заданный на полном вероятностном пространстве $(\Omega, \mathcal{F}, \mathbf{P})$, где фильтрация $(\mathcal{F}_t, t \geq 0)$ удовлетворяет обычным условиям. Предположим, что процесс X непрерывный, а его квадратическая характеристика $\langle X \rangle$ неслучайная. Тогда X является гауссовским процессом с независимыми приращениями.

Наша цель — ответить на следующий вопрос: если квадратично-интегрируемый мартингал X имеет неслучайную характеристику $\langle X \rangle$, то вытекает ли отсюда, что у процесса X приращения независимые? Ответ содержится в рассмотренном ниже примере.

Пусть $\Omega = [0, 1]$, $\mathbf{P}$ — мера Лебега и $\mathcal{F}$ — σ -алгебра, порожденная тремя случайными величинами: η_0, η_1 и η_2 , где $\eta_0(\omega) = 0$ для всех $\omega \in \Omega$,

$$\eta_1(\omega) = \begin{cases} -1, & \text{если } \omega \in \left[0, \frac{1}{2}\right), \\ 1, & \text{если } \omega \in \left[\frac{1}{2}, 1\right), \end{cases}$$

$$\eta_2(\omega) = \begin{cases} -2, & \text{если } \omega \in \left[0, \frac{1}{4}\right), \\ 0, & \text{если } \omega \in \left[\frac{1}{4}, \frac{1}{2}\right), \\ 1 - \sqrt{\frac{3}{2}}, & \text{если } \omega \in \left[\frac{1}{2}, \frac{2}{3}\right), \\ 1, & \text{если } \omega \in \left[\frac{2}{3}, \frac{5}{6}\right), \\ 1 + \sqrt{\frac{3}{2}}, & \text{если } \omega \in \left[\frac{5}{6}, 1\right]. \end{cases}$$

Введем σ -алгебры $\tilde{\mathcal{F}}_i = \sigma\{\eta_i\}$, $i = 0, 1, 2$, обозначим $s_0 = 0$, $s_1 = 1$, $s_2 = 2$ и $s_3 = \infty$ и рассмотрим случайный процесс $X = (X_t, t \geq 0)$, $X_t = \sum_{k=0}^2 \eta_k I_{[s_k, s_{k+1})}(t)$, $t \geq 0$. Введем также фильтрацию $(\mathcal{F}_t, t \geq 0)$, где $\mathcal{F}_t = \tilde{\mathcal{F}}_k$ для $t \in [s_k, s_{k+1})$, $k = 0, 1, 2$. Нетрудно видеть, что она удовлетворяет обычным условиям. Следуя работе [93], можно показать, что $X = (X_t, \mathcal{F}_t, t \geq 0)$ является мартингалом, который, очевидно, ограничен. Далее, его квадратическая характеристика имеет следующий

явный вид:

$$\langle X_t \rangle = \begin{cases} 0, & \text{если } 0 \leq t < 1, \\ 1, & \text{если } 1 \leq t < 2, \\ 2, & \text{если } t \geq 2. \end{cases}$$

Очевидно, характеристика $\langle X \rangle$ не случайная. Заметим, что

$$\mathbf{P}\{X_1 - X_0 = 1, X_2 - X_1 = 1\} = 0 \neq \frac{1}{8} = \mathbf{P}\{X_1 - X_0 = 1\}\mathbf{P}\{X_2 - X_1 = 1\},$$

и, следовательно, X не является процессом с независимыми приращениями. Объясняется это тем, что процесс X изменяется скачками, а в теореме Леви требуется непрерывность X . Таким образом, условие непрерывности существенно для справедливости цитированной теоремы Леви. Отметим, наконец, что ряд обобщений теоремы Леви на случай произвольных квадратично-интегрируемых мартингалов (не обязательно непрерывных) можно найти в статье [27] или в книге [74].

23.5. О возможности представить мартингал в виде стохастического интеграла относительно другого мартингала

Пусть процесс $X = (X_t, t \in [0, T])$ является мартингалом относительно собственной порожденной фильтрации $(\mathcal{F}_t^X, t \in [0, T])$. Предположим, что $M = (M_t, t \in [0, T])$ — другой процесс, причем M — мартингал относительно $(\mathcal{F}_t^X)$. Вопрос: можно ли представить M в виде стохастического интеграла по мартингалу X ? Другими словами, существует ли «подходящая» функция $(\varphi_s, s \in [0, T])$ такая, что $M_t = \int_0^t \varphi_s dX_s$? Для такого вопроса есть соответствующий мотив. Достаточно вспомнить, что ответ на этот вопрос положителен в важном частном случае, когда X — стандартный винеровский процесс (см. [73, 185, 138, 211]). Однако, как станет ясно далее, иногда ответ может быть отрицательным.

Рассмотрим два независимых стандартных винеровских процесса $w = (w_t, t \geq 0)$ и $v = (v_t, t \geq 0)$. Пусть $X_t = \int_0^t w_s dv_s$ и $\mathcal{F}_t^X = \sigma\{X_s, s \leq t\}$.

Тогда $\langle X \rangle_t$ измеримо относительно $\mathcal{F}_t^X$, и, поскольку $\langle X \rangle_t = \int_0^t w_s^2 ds$, отсюда следует, что w_t^2 также $\mathcal{F}_t^X$ -измеримо. Следовательно, процесс $M = (M_t, t \geq 0)$, где $M_t = w_t^2 - t$, является квадратично-интегрируемым мартингалом, согласованным с фильтрацией $(\mathcal{F}_t^X, t \geq 0)$. Предположим теперь, что мартингал M может быть представлен в виде $M_t = \int_0^t H_s dX_s$,

$t \geq 0$, где $(H_s, s \geq 0)$ — предсказуемая функция, для которой

$$\mathbb{E} \left\{ \int_0^\infty H_s^2 d\langle X \rangle_s \right\} < \infty$$

(см. [73, 74, 298]). Согласно формуле Ито, мы имеем $M_t = 2 \int_0^t w_s dw_s$, и тогда

$$M_t = 2 \int_0^t w_s dw_s = \int_0^t H_s dX_s = \int_0^t H_s w_s dv_s.$$

Из этих соотношений вытекает справедливость равенства

$$0 = \mathbb{E} \left\{ \left[2 \int_0^t w_s dw_s - \int_0^t H_s w_s dv_s \right]^2 \right\} = 4 \mathbb{E} \left\{ \int_0^t w_s^2 dw_s \right\} + \mathbb{E} \left\{ \int_0^t H_s^2 w_s^2 ds \right\},$$

что, однако, невозможно. Следовательно, мартингал M нельзя представить в виде стохастического интеграла по мартингалу X .

23.6. Гауссовские процессы, которые не являются семимартингалами

Если $X = (X(t), t \geq 0)$ — гауссовский процесс, он квадратично-интегрируем и вполне естественно задаться вопросом: является ли X семимартингалом (относительно собственно порожденной фильтрации)? Оказывается, можно построить семейство $\{X_H\}$ гауссовских процессов, зависящих от параметра H таких, что X_H — семимартингал для одних значений H и не семимартингал для других H . Начнем со следующего определения.

Гауссовский процесс $X_H = (X_H(t), t \geq 0)$ с нулевым средним и ковариационной функцией

$$C_H(s, t) = \mathbb{E}[X_H(s)X_H(t)] = \frac{1}{2}(s^{2H} + t^{2H} - |s - t|^{2H}), \quad s, t \geq 0, \quad H \in (0, 1),$$

называется *фрактальным винеровским процессом* с параметром Харста H . (Используется также термин *фрактальное броуновское движение*.)

Существование процессов $X_H, H \in (0, 1)$ следует из факта, что функция $C_H(s, t), s, t \geq 0$, — положительно определенная.

Заметим, что если $H = \frac{1}{2}$, гауссовский процесс $X_{1/2}$ имеет ковариационную функцию, равную $\min(s, t)$, т. е. $X_{1/2}$ — это стандартный винеровский процесс. Его мартингалное свойство нам хорошо известно.

Пусть теперь $H \in (0, 1)$ и $H \neq \frac{1}{2}$. Приведем основные аргументы, показывающие, что гауссовские процессы X_H , $H \neq \frac{1}{2}$, не являются семимартингалами.

По-видимому, случай $\frac{1}{2} < H < 1$ был впервые рассмотрен в книге [74]. Там показано, что

$$(C_H(s, s) + C_H(t, t) - 2C_H(s, t))^{1/2} = |t - s|^H,$$

и что функцию $|t - s|^H$, $\frac{1}{2} < H < 1$ нельзя представить в виде разности $F(t) - F(s)$, $s < t$ для некоторой неотрицательной и непрерывной справа функции F . Более того, поскольку $E\{(X_H(s) - X_H(t))^2\} = |t - s|^{2H}$, то для любого разбиения $0 = t_0 < t_1 < \dots < t_n = T$ отрезка $[0, T]$ квадратичная вариация процесса X_H сходится к нулю, т. е.

$$V_n^{(2)} = \sum_{k=0}^{n-1} (X_H(t_{k+1}) - X_H(t_k))^2 \xrightarrow{P} 0 \quad \text{при } n \rightarrow \infty \text{ и } \max(t_{k+1} - t_k) \rightarrow 0.$$

При выполнении описанных двух свойств процесс X_H , $H \in (\frac{1}{2}, 1)$ не является семимартингалом.

В случае $0 < H < \frac{1}{2}$ нам нужно снова рассмотреть квадратическую вариацию $V_n^{(2)}$. Используя свойства приращений процесса X_H , можно показать, что $V_n^{(2)} \xrightarrow{P} \infty$ при $n \rightarrow \infty$. (Подробности см., например, в [400, 414].) Однако, чтобы процесс X_H был семимартингалом, его квадратическая вариация должна быть п. н. конечна. Следовательно, и в этом случае, $0 < H < \frac{1}{2}$, процесс X_H не является семимартингалом.

Приведем еще два примера гауссовских процессов, не являющихся семимартингалами. Пусть $B_t = 0$ при $t \leq 0$ и $B_t = w_t$ при $t \geq 0$, где w — стандартный винеровский процесс. Введем два семейства случайных процессов: $\{Y^{(\varepsilon)}, 0 < \varepsilon < 1\}$ и $\{Z^{(\varepsilon)}, 0 < \varepsilon < \frac{1}{2}\}$, где

$$Y_t^{(\varepsilon)} = B_t - B_{t-\varepsilon}, \quad t \geq 0; \quad Z_t^{(\varepsilon)} = \int_0^t B_s ds + B_{t-\varepsilon}, \quad t \geq 0.$$

Тогда $Y^{(\varepsilon)}$ и $Z^{(\varepsilon)}$ — непрерывные гауссовские процессы, которые, однако, не являются семимартингалами (относительно собственно порожденных фильтраций). Подробности читатель найдет в работе [220].

Отметим, наконец, что достигнут значительный прогресс в общей проблеме (по-видимому, поставленной А. Н. Ширяевым) — описать класс гауссовских процессов, являющихся семимартингалами. Результаты в этом направлении можно найти в работах [144, 171, 215, 216, 243, 301, 363, 400].

§ 24. Разные свойства случайных процессов

Мы выбрали еще только несколько примеров, посвященных разным свойствам случайных процессов. Необходимые определения будут даны в самих примерах. Здесь изложение будет более компактным, но везде будут указаны соответствующие источники.

24.1. О перестановочности операций пересечения и супремума для последовательности σ -алгебр, связанных со случайными процессами

Рассмотрим вероятностное пространство $(\Omega, \mathcal{F}, P)$ и в нем последовательность σ -алгебр $\mathcal{F}_n \subset \mathcal{F}$, $n = 0, 1, 2, \dots$ и пусть $\mathcal{A} \subset \mathcal{F}$ — фиксированная σ -алгебра. Для любой σ -алгебры $\tilde{\mathcal{A}} \subset \mathcal{F}$ через $\mathcal{A} \vee \tilde{\mathcal{A}}$ обозначим супремум $\mathcal{A}$ и $\tilde{\mathcal{A}}$ — это наименьшая σ -алгебра, содержащая все множества вида $A \cup \tilde{A}$, где $A \in \mathcal{A}$, $\tilde{A} \in \tilde{\mathcal{A}}$.

Напомним, что пересечение счетного числа σ -алгебр — снова σ -алгебра. Отсюда вытекает, что $\bigcap_n \mathcal{F}_n$ и $\bigcap_n \mathcal{A} \vee \mathcal{F}_n$ — σ -алгебры, и можно предположить, что операции супремума и пересечения перестановочны, т. е. что имеет место соотношение

$$\mathcal{A} \vee \bigcap_n \mathcal{F}_n = \bigcap_n \mathcal{A} \vee \mathcal{F}_n. \quad (1)$$

Оказывается, соотношения типа (1) важны для ряда прикладных задач. Читатель может обратиться, например, к работе [140]. Поэтому, а также с точки зрения теории, важно выяснить, всегда ли имеет место соотношение (1).

Легко видеть, что для любого фиксированного индекса m

$$\mathcal{A} \vee \bigcap_n \mathcal{F}_n \subseteq \mathcal{A} \vee \mathcal{F}_m \implies \mathcal{A} \vee \bigcap_n \mathcal{F}_n \subseteq \bigcap_m \mathcal{A} \vee \mathcal{F}_m.$$

Это означает, что для справедливости (1) нужно показать, что имеет место обратное включение. Однако, как мы увидим дальше, это не всегда так.

Рассмотрим однородную цепь Маркова $X = (X_t, t = 0, 1, 2, \dots)$ с фазовым пространством $\{1, 2, 3, 4\}$ и следующей матрицей переходных вероятностей за 1 шаг $P = (p_{ij})$, где $p_{ij} = P\{X_1 = j \mid X_0 = i\}$:

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix}.$$

Определим новый процесс $Y = (Y_t, t = 0, 1, 2, \dots)$, где $Y_t = 1_{\{X_t=1 \text{ или } 3\}}$, и рассмотрим следующие σ -алгебры:

$$\mathcal{A} = \sigma\{Y_0, Y_1, Y_2, \dots\}, \quad \mathcal{F}_n = \sigma\{X_0, X_1, \dots, X_n\}, \quad \mathcal{F} = \bigvee_n \mathcal{F}_n.$$

Легко проверить, что у P^3 — матрицы переходов цепи X за 3 шага — все элементы положительны. Отсюда вытекает, что цепь X — рекуррентная и апериодическая. Значит, согласно закону 0–1 для марковских цепей (см., например, [148]), остаточная (хвостовая) σ -алгебра $\bigcap_n \mathcal{F}_n$ — тривиальна п. н. Так заключаем, что

$$\mathcal{A} \bigvee \bigcap_n \mathcal{F}_n = \mathcal{A} \quad \text{Р-п.н.}$$

Посмотрим теперь на траектории процессов X и Y . Ясно, что траектории Y составлены из всех переходов $\{1, 3\} \longleftrightarrow \{2, 4\}$. Это, однако, не включает начальное значение X_0 . Имеет место следующее:

$$\begin{aligned} P\{X_0 = i \mid \mathcal{A}\} &= \frac{1}{2} \cdot 1_{\{Y_0=1\}}, \quad \text{если } i = 1 \text{ или } 3, \\ P\{X_0 = i \mid \mathcal{A}\} &= \frac{1}{2} \cdot 1_{\{Y_0=0\}}, \quad \text{если } i = 2 \text{ или } 4. \end{aligned}$$

Отсюда вытекает, что X_0 не измеримо относительно σ -алгебры $\mathcal{A}$, т. е.

$$\mathcal{A} \subset \mathcal{F} \quad (\text{включение строгое}).$$

С другой стороны, допустим, что мы наблюдали всю траекторию процесса Y , т. е. нам доступна σ -алгебра $\mathcal{A}$, и что знаем значение X_m для фиксированного момента времени m . Поскольку нам известны матрицы переходов, это позволяет восстановить всю траекторию цепи X . Следовательно,

$$\mathcal{A} \bigvee \sigma\{X_m\} = \mathcal{F} \quad \text{для каждого } m,$$

и, более того,

$$\mathcal{A} \bigvee \mathcal{F}_m = \mathcal{F} \quad \text{для любого } m.$$

Таким образом, приходим к заключению, что

$$\bigcap_m \mathcal{A} \bigvee \mathcal{F}_m = \mathcal{F} \supset \mathcal{A} = \mathcal{A} \bigvee \bigcap_n \mathcal{F}_n.$$

Это показывает, что в описанном примере нельзя поменять порядок операций пересечения и супремума для последовательности σ -алгебр, т. е. соотношение (1) не имеет места.

Отметим, наконец, что имеются условия, при которых соотношение (1) выполнено. В работах [140] и [474] читатель найдет ряд интересных результатов и их применение в других областях, в частности, в задаче оптимальной фильтрации случайных процессов.

24.2. О тождествах Вальда для винеровского процесса

Пусть $w = (w(t), \mathcal{F}_t^w, t \geq 0)$ — стандартный винеровский процесс и τ — момент остановки относительно $(\mathcal{F}_t^w)$. Объектом нашего внимания будут следующие соотношения, *тождества Вальда*:

$$Ew(\tau) = 0, \quad (1)$$

$$E\{w^2(\tau)\} = E\tau, \quad (2)$$

$$E\left\{\exp\left(w(\tau) - \frac{1}{2}\tau\right)\right\} = 1. \quad (3)$$

Введем также следующие три условия:

$$E\sqrt{\tau} < \infty, \quad (1^*)$$

$$E\tau < \infty, \quad (2^*)$$

$$E\left\{\exp\left(\frac{1}{2}\tau\right)\right\} < \infty. \quad (3^*)$$

Хорошо известно, что (1^*) , (2^*) и (3^*) являются достаточными условиями для справедливости тождеств (1), (2) и (3) (см. [73, 82, 163]). Наша цель здесь — проанализировать эти условия и выяснить, например, можно ли их ослабить.

Рассмотрим момент остановки $\tau_1 = \inf\{t > 0 \mid w(t) \geq 1\}$. Из непрерывности винеровского процесса w следует, что $w(\tau_1) = 1$ и, значит, $Ew(\tau_1) = 1$, т. е. $Ew(\tau_1)$ не равно 0, как в (1). Заметим, что случайная величина τ_1 имеет плотность $(2\pi t^3)^{-1/2} \exp\left(-\frac{1}{2t}\right)$, $t > 0$, и легко проверить, что $E\{\tau_1^\delta\} < \infty$ при $\delta < \frac{1}{2}$, но $E\{\tau_1^\delta\} = \infty$ для каждого $\delta \geq \frac{1}{2}$. Таким образом, оба условия (1^*) и (2^*) нарушены. Тождество (2) также не выполнено, потому что $E\{w^2(\tau_1)\} = 1 \neq E\tau_1 = \infty$.

По поводу тождества (1) доказан следующий общий результат (см. [390]): *пусть положительная, непрерывная и неубывающая функция $f(t)$, $t \geq 0$, удовлетворяет условию: $\int_1^\infty t^{-3/2} f(t) dt = \infty$. Тогда $Ew(\tau) = 0$ для любого $(\mathcal{F}_t^w)$ -момента остановки τ , для которого $Ef(\tau) < \infty$ и $E|w(\tau)| < \infty$. Оказывается, что нельзя изменить требуемое свойство интегрируемости функции f . Более конкретно, пусть $\tilde{f}$ — положительная, непрерывная и неубывающая функция, для которой $\int_1^\infty t^{-3/2} \tilde{f}(t) dt < \infty$, а момент остановки $\tilde{\tau}$ определим так: $\tilde{\tau} = \inf\{t \geq 0 \mid w(t) \geq 1 - \tilde{f}(t)\}$ (здесь важно, что $\tilde{f}(0) > 0$). Тогда можно показать (см. [390]), что условия $E\tilde{f}(\tilde{\tau}) < \infty$ и $E|w(\tilde{\tau})| < \infty$ выполнены,*

но $Ew(\tilde{\tau}) > 0$, т.е. тождество (1) не имеет места для момента остановки $\tilde{\tau}$.

Обратимся теперь к условию (3*) и тождеству (3). Покажем, что (3*) не может быть существенно ослаблено, сохраняя справедливость (3). Для этой цели рассмотрим следующий момент остановки:

$$\tau_a = \inf\{t \geq 0 \mid w(t) \leq -1 + at\},$$

где a — произвольное действительное число. Поскольку случайная величина τ_a имеет плотность $(2\pi t^3)^{-1/2} \exp\left(-\frac{(-1+at)^2}{2t}\right)$, $t > 0$, легко показать, что

$$E\left\{\exp\left(\frac{1}{2} a^2 \tau_a\right)\right\} < \infty \quad \text{для каждого } a. \quad (3^{**})$$

Более того,

$$E\left\{\exp\left(w(\tau_a) - \frac{1}{2} \tau_a\right)\right\} = \begin{cases} 1, & \text{если } a \geq 1, \\ c_a < 1, & \text{если } a < 1. \end{cases} \quad (4)$$

Здесь константа c_a , зависящая от a , может быть найдена явно, но для нас важно то, что $c_a < 1$ для $a < 1$. Таким образом, из (3**) и (4) вытекает, что коэффициент $\frac{1}{2}$ в условии (3*) нельзя уменьшить, сохраняя справедливость тождества (3).

Читатель наверняка заметил, что условие (3*) и тождество (3) тесно связаны с классом экспоненциальных мартингалов, рассмотренных в примере 23.3. Добавим еще, что тождество (3) играет существенную роль в проблеме характеристики класса экспоненциальных непрерывных локальных мартингалов (см. [73, 74, 83, 320, 321, 400, 408]).

24.3. О сходимости квадратической вариации винеровского процесса

Рассмотрим стандартный винеровский процесс $w = (w(t), t \in [0, 1])$, и пусть

$$V_2(w; \pi_n) = \sum_{k=0}^{n-1} (w(t_{k+1}) - w(t_k))^2$$

есть его квадратическая вариация на отрезке $[0, 1]$, порожденная разбиением $\pi_n = \{0 = t_0 < t_1 < \dots < t_n = 1\}$. Число $d_n = \max_k (t_{k+1} - t_k)$ — это диаметр разбиения π_n . Нас интересует поведение последовательности $\{V_2(w; \pi_n)\}$ при $n \rightarrow \infty$ и $d_n \rightarrow 0$. Из элементарных свойств винеровского процесса и неравенства Чебышёва вытекает, что

$$V_2(w; \pi_n) \xrightarrow{P} 1 \quad \text{при } d_n \rightarrow 0. \quad (1)$$

Стоит подчеркнуть, что для справедливости (1) достаточно, чтобы $d_n \rightarrow 0$ при $n \rightarrow \infty$. При этом же условии $d_n \rightarrow 0$ (1) будет иметь место и в среднеквадратическом смысле. Однако, как будет видно из дальнейшего рассуждения, для сходимости $V_2(w; \pi_n) \xrightarrow{\text{п.н.}} 1$ при $n \rightarrow \infty$ существенно скорость, с которой d_n стремится к 0. Напомним сначала результат, известный как *теорема Леви* (см. [70, 73, 304, 354]): рассмотрим следующее разбиение отрезка $[0, 1]$: $\pi_n = \left\{ t_k = \frac{k}{2^n}, k = 0, 1, \dots, 2^n - 1 \right\}$. Тогда при $n \rightarrow \infty$

$$V_2(w; \pi_n) = \sum_{k=0}^{2^n-1} \left[w\left(\frac{k+1}{2^n}\right) - w\left(\frac{k}{2^n}\right) \right]^2 \xrightarrow{\text{п.н.}} 1. \quad (2)$$

Отметим, что предел 1 в соотношениях (1) и (2) — это длина отрезка $[0, 1]$. Очевидно также, что сходимость в (2) более сильная, чем сходимость в (1). Можно, конечно, догадаться, что это обстоятельство связано с тем фактом, что в теореме Леви $d_n = \frac{1}{2^n}$ стремится к 0 «очень» быстро. И действительно, в работе [209] показано, что условие $d_n = o\left(\frac{1}{\ln n}\right)$ влечет сходимость $V_2(w; \pi_n) \xrightarrow{\text{п.н.}} 1$ при $n \rightarrow \infty$. Кроме того, даже для более общего случая указано, что $o\left(\frac{1}{\ln n}\right)$ — это «наилучший» возможный порядок стремления диаметра d_n к 0, когда $V_2(w; \pi_n)$ сходится п. н. к пределу 1 при $n \rightarrow \infty$. В частности, существует разбиение π_n отрезка $[0, 1]$ с диаметром $d_n = O\left(\frac{1}{\ln n}\right)$ такое, что при $n \rightarrow \infty$ квадратическая вариация $V_2(w; \pi_n)$ не сходится к 1, а будет сходиться, например, к некоторому числу, большему 1. Подробную конструкцию разбиения π_n с диаметром $d_n = O\left(\frac{1}{\ln n}\right)$ и доказательство сходимости $V_2(w; \pi_n)$ к пределу $1 + \delta$ с некоторым $\delta > 0$ можно найти в работе [228].

Отметим, наконец, что если разбиение π_n отрезка $[0, 1]$ имеет диаметр d_n , порядок которого меньше, чем $\left(\frac{1}{\ln n}\right)^\alpha$, где $\alpha \in (0, 1)$, то квадратическая вариация $V_2(w; \pi_n)$ винеровского процесса w расходится п. н. при $n \rightarrow \infty$. Это утверждение можно найти в работе [487].

24.4. Можно ли характеризовать многомерный гауссовский процесс по его проекциям?

Пусть задан n -мерный случайный процесс $X = ((X_1(t), \dots, X_n(t)), t \geq 0)$. Для любых действительных чисел $\lambda_1, \dots, \lambda_n$ вводим линейную комбинацию

$$Z(t) = \lambda_1 X_1(t) + \dots + \lambda_n X_n(t), \quad (1)$$

часто называемую *проекцией процесса* X (порожденной числовым вектором $(\lambda_1, \dots, \lambda_n)$).

Рассмотрим теперь n -мерный стандартный винеровский процесс $w = ((w_1(t), \dots, w_n(t), t \geq 0)$ и его проекцию

$$Y(t) = \lambda_1 w_1(t) + \dots + \lambda_n w_n(t). \quad (2)$$

Поскольку $w_j = (w_j(t), t \geq 0)$, $j = 1, \dots, n$, — независимые между собой одномерные стандартные винеровские процессы, то

$$EY(t) = 0, \quad t \geq 0, \quad \text{и} \quad E\{Y(s)Y(t)\} = (\lambda_1^2 + \dots + \lambda_n^2) \min(s, t). \quad (3)$$

Предположим, что процесс X гауссовский, а его проекции (1) удовлетворяют следующим условиям:

$$EZ(t) = 0, \quad t \geq 0,$$

и

$$E\{Z(s)Z(t)\} = (\lambda_1^2 + \dots + \lambda_n^2) \min(s, t), \quad s, t \geq 0. \quad (4)$$

Это означает, что свойство (3) для n -мерного стандартного винеровского процесса имеет место и для процесса X . Вспомним, что гауссовский процесс описывается однозначно двумя характеристиками — математическим ожиданием и ковариационной функцией. Поскольку оба процесса w и X гауссовские, то ввиду совпадения свойств (3) и (4) можно предположить, что процесс X — это тоже n -мерный стандартный винеровский процесс. Описанный ниже пример показывает, однако, что это не всегда так. Для простоты изложения рассмотрим случай $n = 2$.

Итак, пусть $w_1 = (w_1(t), t \geq 0)$ и $w_2 = (w_2(t), t \geq 0)$ — независимые стандартные винеровские процессы. Определим новый процесс $X = ((X_1(t), X_2(t), t \geq 0)$, где

$$X_1(t) = w_1\left(\frac{2t}{3}\right) + w_2\left(\frac{t}{3}\right), \quad X_2(t) = w_1\left(\frac{t}{3}\right) - w_2\left(\frac{2t}{3}\right).$$

По определению, $w = ((w_1(t), w_2(t), t \geq 0)$ — стандартный винеровский процесс в $\mathbb{R}^2$, и легко видеть, что для любых λ_1 и λ_2 его проекция $Y(t) = \lambda_1 w_1(t) + \lambda_2 w_2(t)$ удовлетворяет свойству (3). Заметим далее, что процесс X гауссовский. Беря те же самые λ_1, λ_2 , с помощью непосредственного вычисления убеждаемся, что проекция $Z(t) = \lambda_1 X_1(t) + \lambda_2 X_2(t)$ процесса X удовлетворяет свойству (4). Таким образом, у процессов X и w нулевые средние значения и совпадающие ковариации их проекций. Отсюда, однако, нельзя заключить, что X — двумерный винеровский процесс. Компоненты X_1 и X_2 гауссовского процесса X сами являются гауссовскими, но очевидно, что они не независимы и ни одна из компонент не является стандартным винеровским процессом.

24.5. Представление Крамера, кратность и спектральный тип случайного процесса

Пусть на вероятностном пространстве $(\Omega, \mathcal{F}, P)$ задан $\mathcal{L}^2$ -случайный процесс $X = (X(t), t \geq 0)$, принимающий значения в $\mathbb{R}^1$. Через $\mathcal{H}_t(X)$ обозначим линейное замыкание (в L^2 -смысле), порожденное величинами $\{X(s), s \leq t\}$. Отметим, что имеется значительное число исследований, посвященных процессам из пространства $\mathcal{L}^2$ (см., например, [97, 129, 188, 283, 296]). Здесь мы коснемся только нескольких понятий. (Далее термин «ортогональный» используется как синоним термина «некоррелированный», а символ $\succ$ означает абсолютную непрерывность мер.)

Начнем с формулировки следующего результата (см. [189]): пусть $F_1, F_2, \dots, F_N$ — последовательность мер на σ -алгебре $\mathcal{B}^+$ такая, что

$$F_1 \succ F_2 \succ \dots \succ F_N, \quad (1)$$

где N — либо фиксированное натуральное число, либо бесконечность. Тогда существует непрерывный случайный процесс $X = (X(t), t \geq 0)$ и N взаимно ортогональных процессов $Y_1, \dots, Y_N$, каждый с ортогональными приращениями, $E\{(dY_n(t))^2\} = dF_n(t)$ такие, что $\mathcal{H}_t(X) = \sum_{n=1}^N \mathcal{H}_t(Y_n)$ для каждого $t > 0$. В частности, процесс X допускает представление

$$X(t) = \sum_{n=1}^N \int_0^t g_n(t, u) dY_n(u), \quad t \geq 0, \quad (2)$$

где $g_n, 1 \leq n \leq N$, — неслучайные функции, подчиненные условию

$$\sum_{n=1}^N \int_0^t g_n^2(t, u) dF_n(u) < \infty, \quad t > 0.$$

Соотношение (2) называется *представлением Крамера* для процесса X , а последовательность (1) — его *спектральным типом*. Число N , участвующее в (1) и (2), называется *кратностью* X . Проиллюстрируем теперь введенные понятия.

(а) Предположим, что $\mathcal{L}^2$ -процесс $\eta = (\eta_t, t \geq 0)$ имеет ортогональные приращения и $E\{(d\eta_t)^2\} = dt$. Рассмотрим другой процесс, $\xi = (\xi_t, t \geq 0)$, заданный следующим образом:

$$\xi_t = \int_0^t h(t, u) d\eta_u. \quad (3)$$

Здесь h — некоторая неслучайная функция с $\int_0^t h^2(t, u) du < \infty$, $t > 0$.

Сравнивая (3) и (2), мы видим, что процесс ξ допускает представление Крамера с $N = 1$, т. е. можно предположить, что ξ имеет кратность 1. Оказывается, это не всегда так. Ядро h может быть выбрано таким образом, чтобы кратность процесса ξ была больше 1. Подробное описание подобного примера, в котором кратность ξ равна 2, читатель может найти в [224] или [449].

(б) Пусть $X_1 = (X_1(t), t \geq 0)$ и $X_2 = (X_2(t), t \geq 0)$ — гауссовские процессы, а P_1 и P_2 — вероятностные меры, порожденные этими процессами. Хорошо известно, что меры P_1 и P_2 либо эквивалентны, либо сингулярны (см. [50]). Имеет место, например, следующий результат: *Если меры P_1 и P_2 , соответствующие гауссовским процессам X_1 и X_2 , эквивалентны, то процессы X_1 и X_2 имеют один и тот же спектральный тип.* С помощью конкретного примера ответим на следующий вопрос: если гауссовские процессы X_1 и X_2 имеют один и тот же спектральный тип, то верно ли, что соответствующие им меры эквивалентны?

Рассмотрим винеровский процесс $w = (w_t, t \geq 0)$ и процесс $\xi = (\xi_t, t \geq 0)$, заданный равенством $\xi_t = h(t)w_t$, где h — некоторая неслучайная функция. Таким образом, мы имеем пару гауссовских процессов w и ξ , и пусть P_w и P_ξ — соответствующие им вероятностные меры. Выберем функцию h следующим образом: предположим, что на любом интервале функция h не является абсолютно непрерывной, и пусть для некоторых констант m_1, m_2 , $0 < m_1 < m_2 < \infty$, выполнено условие $m_1 \leq h(t) \leq m_2$ для всех $t > 0$. (Нетрудно привести примеры таких функций.) Из определения процесса ξ вытекает, что $\mathcal{H}_t(\xi) = \mathcal{H}_t(w)$ для каждого $t > 0$. Отсюда можно заключить, что процессы ξ и w имеют один и тот же спектральный тип.

Выясним теперь, являются ли меры P_ξ и P_w эквивалентными. Напомним (см. [50]), что в общем случае гауссовские меры P_1 и P_2 эквивалентны тогда и только тогда, когда разность $\Delta = C_1 - C_2$ их ковариационных функций абсолютно непрерывна. Рассмотрим ковариационные функции C_ξ и C_w , соответствующие процессам ξ и w . Легко найти их разность $\Delta(s, t)$:

$$\Delta(s, t) = C_\xi(s, t) - C_w(s, t) = [h(s)h(t) - 1] \min(s, t).$$

Поскольку функция h выбрана так, что разность $\Delta(s, t)$ не является абсолютно непрерывной, то заключаем, что меры P_ξ и P_w сингулярны.

Итак, в рассмотренном примере меры P_ξ и P_w не являются эквивалентными, несмотря на то, что процессы ξ и w имеют один и тот же спектральный тип.

24.6. Фрактальный винеровский процесс и марковское свойство

Коротко напомним, что фрактальный винеровский процесс $X = (X(t), t \geq 0)$ с параметром Харста H — это гауссовский процесс с нулевым средним и ковариационной функцией

$$C(s, t) = \frac{1}{2} \left(s^{2H} + t^{2H} - |s - t|^{2H} \right), \quad s, t \geq 0, \quad H \in (0, 1)$$

(см. также пример 23.6; здесь мы опускаем индекс H в записи X и C).

Можно показать, со ссылкой на критерий Колмогорова, что для каждого $H \in (0, 1)$ процесс X непрерывный. Он обладает и другими свойствами, некоторые из них были использованы в примере 23.6.

Здесь мы ставим следующий вопрос: является ли X марковским процессом? Ответ положительный по крайней мере в одном случае, когда $H = \frac{1}{2}$. Тогда X — стандартный винеровский процесс, у него приращения независимые, он марковский и т. д. Поэтому нужно рассмотреть процесс X с параметром $H \neq \frac{1}{2}$. Полезно отметить, что приращения X на непересекающихся временных интервалах не являются независимыми, хотя сами приращения стационарны и, например, $C(t, t) = E\{X^2(t)\} = t^{2H}$. Эти свойства полезны при поиске ответа на поставленный выше вопрос.

Напомним, что имеет место следующий общий результат (см., например, [346, 408]): *центрированный гауссовский процесс $X(t), t \geq 0$, является марковским тогда и только тогда, когда его ковариационная функция $C(s, t) = E\{X(s), X(t)\}$ удовлетворяет условию*

$$C(s, u) = \frac{C(s, t)C(t, u)}{C(t, t)} \quad \text{для всех } t \in (s, u). \quad (1)$$

Легко видеть, что равенство (1) выполнено для стандартного винеровского процесса, т. е. при $H = \frac{1}{2}$, как это и должно быть.

Оказывается, однако, что если $H \in (0, 1)$ и $H \neq \frac{1}{2}$, то фрактальный винеровский процесс не является марковским. Чтобы показать это, рассмотрим частное $R(t|s, u) = \frac{C(s, t)C(t, u)}{C(t, t)}$, где $t \in (s, u)$. Очевидно, что $\lim_{t \downarrow s} R(t|s, u) = C(s, u)$ и $\lim_{t \uparrow u} R(t|s, u) = C(s, u)$. Теперь остается установить, что для любого $t \in (s, u)$ имеет место неравенство $R(t|s, u) \neq C(s, u)$. Либо $R(t|s, u) > C(s, u)$, либо $R(t|s, u) < C(s, u)$ в зависимости от значения параметра H ($0 < H < \frac{1}{2}$ или $\frac{1}{2} < H < 1$). Читателю предлагается восполнить технические детали.

ДОПОЛНИТЕЛЬНЫЕ ЗАМЕЧАНИЯ

Некоторые из примеров, включенных в книгу, хорошо известны, и они являются частью так называемого вероятностного фольклора. Большинство примеров принадлежит разным авторам, и эти примеры, иногда просто в форме вопроса, появлялись в книгах и журнальных статьях. Здесь указаны не только соответствующие источники, но по каждой теме даны координаты других источников, где заинтересованный читатель найдет новые (иногда довольно сложные и специальные) примеры.

§ 1. Классы случайных событий. Примеры 1.1, 1.2, 1.3, 1.4, 1.5 и их модификации относятся к вероятностному фольклору. Пример 1.6 основан на рассуждениях, приведенных в [81, 326]. Пример 1.7 составляет основное содержание работы [389]. Другие примеры можно найти в [139, 145, 179].

§ 2. Вероятность. Пример 2.1 хорошо известен. Пример 2.2 использует идею из [157]. Пример 2.3 есть комбинация некоторых рассмотрений в [81] и [116]. По-видимому, пример 2.4 появился впервые в [32], но далее был включен в [75, 92, 114, 145, 345, 404, 480]. Другие примеры содержатся в [117, 118, 145, 147, 149, 330, 394, 401, 404].

§ 3. Независимость случайных событий. Примеры 3.1, 3.2(в), 3.5 и им подобные можно классифицировать как вероятностный фольклор. По-видимому, классический пример С. Н. Бернштейна 3.2(а) (см. [5]) — это первый пример такого рода. Отметим, что этот пример изложен в предыдущем издании того же учебника в 1928 г. А согласно Б. В. Гнеденко, имеется литографское издание лекций по теории вероятностей С. Н. Бернштейна (1916 г.), где можно найти как описание примера 3.2(а), так и анализ свойства независимости случайных величин. Идея для примера 3.2(б) заимствована из [105]. Пример 3.3 можно найти в [116, 127]. Пример 3.4(а) содержится в [448]. Р. Амбарцумян [126] предложил пример 3.4(б). Пример 3.6 описан в [45]. Для примера 3.7 использована идея из [380]. Пример 3.8 основан на работе [453]. Дополнительные примеры можно найти в [117, 118, 145, 146, 177, 261, 272, 326, 335, 347, 374, 380, 391, 402, 453, 479, 485].

§ 4. Разные свойства случайных событий и их вероятностей. Пример 4.2 можно считать фольклором. Идея примера 4.1 содержится в [247]. Пример 4.3 разобран в [335]. Другие примеры можно найти в [45, 239, 274, 352, 401, 405, 449].

§ 5. Функции распределения случайных величин. Примеры 5.1, 5.3, 5.4, 5.6, 5.7(а) и близкие к ним являются вероятностным фольклором. А.Зубков [44] предложил пример 5.2. Идея примера 5.5 высказана в [231]. Пример 5.7(б) содержится в [127]. По поводу примера 5.7(в) можно обратиться к [263] и [237]. Пример 5.9(б) изложен в [178]. Пример 5.9(в) рассмотрен в [41]. Новые примеры можно найти в [62, 106, 109, 186, 201, 217, 231, 235, 375, 386, 463].

§ 6. Математическое ожидание. Примеры 6.2, 6.4, 6.5 и 6.6 можно отнести к вероятностному фольклору. Пример 6.1 описан в [436]. Пример 6.3 заимствован из [183]. Пример 6.7 содержится в работе [223]. В [345] высказана идея примера 6.8. Другие примеры содержатся в [117, 139, 307, 404, 411, 449].

§ 7. Независимость случайных величин. Примеры 7.1, 7.3(б), 7.5, 7.6(а), 7.7 и 7.8 — вероятностный фольклор. Пример 7.2(а) содержится в [478]. Пример 7.2(б) изложен в [272]. В.Рохатти [415] предложил пример 7.3(в). Пример 7.4(а) следует изложению, приведенному в [345] и [231]. Идея примера 7.4(б) имеется в [405]. Пример 7.9(а) содержится в [275], а пример 7.9(б) основан на идее из [177]. Другие примеры можно найти в [75, 106, 116, 117, 118, 129, 145, 160, 179, 207, 234, 244, 246, 247, 261, 306, 309, 345, 383, 425, 449, 462].

§ 8. Характеристические и производящие функции. Пример 8.3 можно считать хорошо известным. Пример 8.1 составляет содержание работы [23]. Пример 8.2 можно найти в [128, 231, 374]. Пример 8.4 предложен в [498], но мы следуем в основном изложению в [67, 405]. Пример 8.5 рассмотрен в [498], но, по-видимому, он предложен впервые в [482]. Примеры 8.6(а) и 8.6(б) заимствованы из [332]. Пример 8.6(в) содержится в [76, 345]. Новые примеры можно найти в [71, 75, 76, 95, 106, 117, 118, 177, 404, 424, 439, 443, 463, 471].

§ 9. Безгранично делимые и устойчивые распределения. Примеры 9.1 и 9.6 относятся к фольклору. Пример 9.2(а) описан в [200]. Пример 9.2(б), принадлежащий В.Сенатову, включен в дополнение, написанное В.Золотарёвым к книге [76]. Пример 9.3 заимствован из [416]. Пример 9.4 основывается на работах [259, 355, 404]. Пример 9.5 описан в [288], а пример 9.7 содержится в [76]. Другие примеры можно найти в [41, 71, 76, 95, 200, 359, 362, 385, 409, 416, 433, 441].

§ 10. Нормальное распределение. Примеры 10.1(а), 10.4(а) и 10.4(б) хорошо известны. Примеры 10.1(б) и 10.1(в) отмечены в [386, 425] и [375].

Идея примера 10.2 высказана в [391]. Пример 10.3 рассмотрен в [397]. Пример 10.4(в) предложен в [495], а пример 10.4(г) можно найти в [129]. Примеры 10.5 и 10.6 основаны на работах [270, 271]. Пример 10.7(а) возник при разговорах с А. Зубковым и Н. Крыловым. Пример 10.7(б) содержится в [122]. Дополнительные примеры можно найти в [1, 106, 158, 169, 200, 246, 266, 303, 312, 367, 418, 463].

§ 11. Проблема моментов. Пример 11.1 отмечен в [92], но мы следуем изложению в [142]. Пример 11.2 основывается на работах [280] и [353], но первую его часть можно найти также в [96, 142, 145] и [106]. Пример 11.3 рассмотрен в [116], а также в [58]. Пример 11.4 описан в [449, 452]. Случаи (а), (б) и (в) в примере 11.5 основаны на идеях в работах [240, 300, 429]. Идея примера 11.6(а) высказана в [201]. Пример 11.6(б) обсуждается в [177], а пример 11.7 рассмотрен в [281]. Пример 11.8 изложен в [449]. Другие примеры можно найти в [58, 200, 231, 300, 356, 380, 388, 434, 449, 450].

§ 12. Характеризационные свойства некоторых вероятностных распределений. Примеры 12.1, 12.3(а) и 12.7 можно отнести к вероятностному фольклору. Пример 12.2 предложен в [59]. Пример 12.3(б) описан в [395] как ответ на поставленный мною вопрос. Пример 12.4 можно найти в [424]. Пример 12.5(а) составляет содержание работы [344], а идея примера 12.5(б) высказана в [404]. Л. Барингхауз и Н. Хенце [134] предложили пример 12.6 как ответ на мой вопрос. Наконец, пример 12.8 рассмотрен в [333]. Дополнительные примеры можно найти в [7, 55, 132, 235, 242, 366, 424, 432, 449].

§ 13. Разные свойства случайных величин. Пример 13.1 изложен в [442]. Пример 13.2(б) содержится в [141]. По поводу примера 13.3 можно обратиться к [86]. В. Рохатги [415] предложил идею примера 13.4. О возможности построить пример 13.5 сказано в [461]. Пример 13.8 можно найти в [202], а также в [106]. Другие примеры содержатся в [13, 15, 123, 137, 162, 167, 177, 230, 236, 239, 248, 310, 341, 345, 349, 387, 397, 399, 403, 404, 405, 449, 458, 461].

§ 14. Разные виды сходимости последовательностей случайных величин. Примеры 14.1, 14.2, 14.3(б), 14.4, 14.5, 14.6, 14.8 и их варианты хорошо известны. Пример 14.3 встречается во многих источниках, например в [139, 145, 345, 425]. Идея примера 14.7 высказана в [106, 358] и [145]. О примере 14.9 упоминается в [129]. При изложении примера 14.10 мы следовали [287]. И, наконец, пример 14.11 рассмотрен в [358]. Дополнительные примеры можно найти в [6, 75, 106, 116, 117, 118, 145, 177, 179, 210, 244, 261, 279, 326, 358, 374, 404, 430, 444, 449, 462].

§ 15. Законы больших чисел. Пример 15.1 хорошо известен. Пример 15.2 содержится в [358] и [139]. Пример 15.3 можно найти в [407].

Пример 15.4 изложен в [358], а также в [231]. В [447] есть подробное описание примера 15.5. Примеры 15.6 и 15.7 содержатся соответственно в [241] и [287]. Идея примера 15.8 указана в [460], а также в [404]. Другие примеры содержатся в [67, 90, 106, 177, 192, 203, 268, 284, 407].

§ 16. Слабая сходимости вероятностных мер и распределений. Примеры 16.2, 16.4 и 16.6(а) можно отнести к вероятностному фольклору. Пример 16.1 основан на рассуждениях в [6] и [116]. Пример 16.3 развивает идею, подсказанную в [324]. Пример 16.5 предложен в [106]. Пример 16.6(б) заимствован из [195]. В. Золотарев [40] предложил пример 16.7. Дополнительные примеры можно найти в [6, 8, 99, 117, 118, 145, 157, 179, 231, 348, 358, 444, 449].

§ 17. Центральная предельная теорема. Примеры 17.1(в), 17.2, 17.4(а) и 17.4(б) — это вероятностный фольклор. Примеры 17.1(а) и 17.3 основаны на работах [128] и [177]. Пример 17.4(в) указан в [404]. Пример 17.5 содержится в [106]. А. Зубков [44] предложил пример 17.6. Наконец, в [24] есть описание примера 17.7. Другие примеры можно найти в [16, 24, 106, 194, 231, 303, 330, 345, 361, 405, 449, 447].

§ 18. Разные предельные теоремы. Примеры 18.4 и 18.5 хорошо известны. Пример 18.1 основывается на работах [404] и [177]. Пример 18.3 рассмотрен в [358]. Пример 18.6 изложен в [475], а идея примера 18.7 имеется в [406]. Другие примеры содержатся в [31, 39, 43, 91, 102, 105, 106, 118, 143, 159, 170, 176, 179, 196, 255, 257, 267, 322, 327, 341, 374, 392, 446, 449, 447, 459, 461, 484].

§ 19. Основные понятия и свойства случайных процессов. К вероятностному фольклору можно отнести примеры 19.1, 19.2, 19.3, 19.5, 19.6, а также их модификации. При изложении примера 19.3 мы используем частично рассуждение в [492]. Пример 19.4(а) описан в [56]. Пример 19.4(б) изложен в [29], а также в [120]. Пример 19.7 основан на результатах в [376]. Новые примеры можно найти в [3, 14, 22, 32, 65, 68, 88, 111, 120, 129, 151, 157, 166, 180, 199, 206, 227, 286, 313, 365, 368, 393, 421, 472, 480].

§ 20. Марковские процессы. Примеры 20.1 и 20.6(а) хорошо известны. Пример 20.2 основывается на [124]. Идея примера 20.3 содержится в [293]. Пример 20.4 описан в [285]. Пример 20.5 рассмотрен в [438], а также в [68]. Пример 20.6(б) частично заимствован из [14] и [34]. Очевидцы (например, Ю. К. Беляев) утверждают, что пример 20.7(а) — это первый пример марковского, но не строго марковского процесса, был предложен А. А. Юшкевичем (см. [35]). Идея примера 20.7(б) имеется в [486]. Дополнительные примеры содержатся в [14, 33, 34, 51, 151, 157, 170, 181, 202, 225, 269, 290, 342, 363, 371, 420, 421, 428, 449, 467, 480].

§ 21. Стационарные процессы и некоторые смежные вопросы. Примеры 21.1, 21.7(а) и 21.7(б) можно классифицировать как вероятностный фольклор. Пример 21.3 основан на работе [46]. Пример 21.4 изложен в [50]. Пример 21.5 обсуждается в [17]. При изложении примера 21.6 мы используем некоторые аргументы из [129] и частично из [116]. Пример 21.7(в) рассмотрен в [154]. Другие примеры (некоторые из них довольно сложные) можно найти в [6, 11, 18, 28, 47, 155, 156, 161, 258, 277, 278, 303, 336, 384, 422, 445, 449, 460].

§ 22. Мартингалы с дискретным временем. Примеры 22.1(а), 22.3(а), 22.6(а) и 22.6(б) хорошо известны. У. Кюхлер [339] предложил пример 22.1(б). Для примера 22.2 использованы аргументы из [57] и [238]. А. Мельников [79] предложил примеры 22.3(б) и 22.4(а). Примеры 22.4(б), 22.5 и 22.6 содержатся в [464]. Пример 22.7(а), принадлежащий Г. Саймонсу, изложен в [444]. Пример 22.7(б) рассмотрен в [381], а случай (б) того же примера обсуждается подробно в [152]. Р. Дж. Томкинс [466] сообщил мне пример 22.8. Дополнительные примеры могут быть найдены в [128, 150, 193, 213, 214, 250, 264, 291, 364, 377, 379, 403, 449, 464, 465].

§ 23. Мартингалы с непрерывным временем. Пример 23.1(б) хорошо известен. Пример 23.1(а) заимствован из [52]. Пример 23.2 основывается на рассуждениях в [112] и [413]. Пример 23.3 описан в [82], а также в [73]. Пример 23.4 можно найти в [93]. Пример 23.5 обсуждается в [56]. При изложении примера 23.6 мы следуем [74, 414] и [220]. Другие примеры содержатся в [29, 73, 103, 125, 130, 131, 138, 165, 172, 198, 199, 205, 212, 219, 225, 233, 292, 299, 301, 304, 305, 316, 317, 318, 319, 321, 329, 338, 340, 368, 369, 370, 372, 408, 454, 476, 480, 481, 493].

§ 24. Разные свойства случайных процессов. Пример 24.1 предложил П. Чигански [173]. Пример 24.2 обсуждается в [73, 82, 84]. Пример 24.3 основывается на рассмотрении в [209] и [228]. Идея примера 24.4 содержится в [273]. Пример 24.5(а) указан в [224]. З. Ивкович [295] предложил пример 24.5(б). Пример 24.6 был темой обсуждения с П. Копановым. Другие примеры можно найти в следующих работах: [3, 14, 20, 22, 37, 54, 56, 74, 77, 110, 131, 135, 136, 175, 180, 191, 197, 221, 222, 256, 262, 283, 296, 297, 304, 305, 316, 351, 357, 365, 368, 373, 400, 413, 427, 449, 455, 456, 457, 472, 477, 480, 488].

ЛИТЕРАТУРА

Для названий часто цитированных журналов приняты краткие обозначения:

ТВП	=	Теория вероятностей и ее применения
AMM	=	American Mathematical Monthly
AMS	=	Annals of Mathematical Statistics
AmS	=	American Statistician
AP	=	Annals of Probability
AS	=	Annals of Statistics
JAP	=	Journal of Applied Probability
LNM	=	Lecture Notes in Mathematics
SPA	=	Stochastic Processes and Their Applications
ZW	=	Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete

- [1] Андерсон Т. Введение в многомерный статистический анализ. — М.: Физматгиз, 1963.
- [2] Ахизер Н.И. Классическая проблема моментов. — М.: Наука, 1961.
- [3] Баласанов Ю.Г., Журбенко И.Г. Заметки по поводу локальных свойств выборочных функций // Мат. заметки. — 1985. — Т. 37, вып. 6. — С. 926–932.
- [4] Беляев Ю.К. Личное сообщение. — 1985.
- [5] Бернштейн С.Н. Теория вероятностей (4-е изд.). — М.: Гостехиздат, 1946.
- [6] Биллингсли П. Сходимость вероятностных мер. — М.: Наука, 1977.
- [7] Бланк Н.М. Об определении функций распределения их асимптотическим поведением при $x \rightarrow \infty$ // Проблемы устойчивости стохастических моделей. — М.: ВНИИСИ, 1981. — С. 10–15.
- [8] Боровков А.А. Сходимость распределений функционалов случайных процессов // Успехи Мат. Наук. — 1972. — Т. 27, вып. 1. — С. 1–41.
- [9] Боровков А.А. Теория вероятностей. (2-е изд.). — М.: Наука, 1986.
- [10] Булинский А.В. Личное сообщение. — 1988.
- [11] Булинский А.В. О различных условиях перемешивания и асимптотической нормальности случайных полей // ДАН СССР. — 1988. — Т. 299, вып. 4. — С. 785–789.
- [12] Ватанабе С., Икеда Н. Стохастические дифференциальные уравнения и диффузионные процессы. — М.: Наука, 1986.
- [13] Вахания Н.Н., Тариеладзе В.И., Чобанян С.А. Вероятностные распределения в банаховых пространствах. — М.: Наука, 1985.
- [14] Вентцель А.Д. Курс теории случайных процессов. (2-е изд.). — М.: Наука, 1996.

- [15] Галамбош Я. Асимптотическая теория экстремальных порядковых статистик. — М.: Наука, 1984.
- [16] Гамкредидзе Н.Г. О связи локальной и интегральной теорем для решетчатых распределений // ТВП. — 1968. — Т. 13, № 1. — С. 175–179.
- [17] Гапошкин В.Ф. К усиленному ЗБЧ для стационарных в широком смысле процессов // ТВП. — 1973. — Т. 18, вып. 2. — С. 388–392.
- [18] Гапошкин В.Ф. О связи эргодичности стационарного процесса с непрерывным временем и квантованного процесса // ТВП. — 1988. — Т. 33, вып. 2. — С. 401–405.
- [19] Гелбаум Б., Олмстед Дж. Контрпримеры в анализе. — М.: Мир, 1967.
- [20] Гирсанов И.В. Пример неединственности решения стохастического интегрального уравнения Ито // ТВП. — 1962. — Т. 7, вып. 2. — С. 336–342.
- [21] Гихман И.И., Скороход А.В. Введение в теорию случайных процессов. (2-е изд.). — М.: Наука, 1977.
- [22] Гихман И.И., Скороход А.В. Теория случайных процессов. — М.: Наука, т. 1, 1971; т. 2, 1973; т. 3, 1975.
- [23] Гнеденко Б.В. О характеристических функциях // Бюллетень МГУ, секц. А. Мат. и Мех. — 1937. — Т. 1. — С. 17–18.
- [24] Гнеденко Б.В., Колмогоров А.Н. Предельные теоремы для сумм независимых случайных величин. — Л. — М.: Гостехиздат, 1949.
- [25] Гнеденко Б.В. Личное сообщение. — 1982.
- [26] Гнеденко Б.В. Курс теории вероятностей. (6-е изд.). М.: Наука, 1988.
- [27] Григелионис Б.И. О мартингальной характеристизации случайных процессов с независимыми приращениями // Литовский Мат. Сборник. — 1977. — Т. 17, вып. 1. — С. 75–86.
- [28] Давыдов Ю.А. Условия перемешивания для цепей Маркова // ТВП. — 1973. — Т. 18, вып. 2. — С. 321–338.
- [29] Деллашери К. Емкости и случайные процессы. — М.: Мир, 1975.
- [30] Димитров Б. Личное сообщение. — 1982.
- [31] Дороговцев А.Я., Сильвестров Д.С., Скороход А.В., Ядренко М.И. Теория вероятностей. Сборник задач. — Киев: Вища школа, 1980.
- [32] Дуб Дж. Вероятностные процессы. — М.: ИЛ, 1956.
- [33] Дынкин Е.Б. Основания теории марковских процессов. — М.: Физматгиз, 1959.
- [34] Дынкин Е.Б. Марковские процессы. — М.: Физматгиз, 1963.
- [35] Дынкин Е.Б., Юшкевич А.А. Строго марковские процессы // ТВП. — 1956. — Т. 1, вып. 1. — С. 149–155.
- [36] Звонкин А.К. Личное сообщение. — 1988.
- [37] Звонкин А.К., Крылов Н.В. О сильных решениях стохастических дифференциальных уравнений // Труды школы-семинара по теории случайных процессов (часть II), — Вильнюс: Инст. физ. мат. АН Лит. ССР, 1975. — С. 9–88.

- [38] Зигмунд А. Тригонометрические ряды. Т. 1,2. — М.: Мир, 1965.
- [39] Золотарёв В.М. Обобщение неравенства Колмогорова // Исследования мех. прикл. мат. (МФТИ). — 1961. Т. 7. — С. 162–166.
- [40] Золотарёв В.М. Личное сообщение. — 1985.
- [41] Золотарёв В.М. Одномерные устойчивые распределения. — М.: Наука, 1986.
- [42] Золотарёв В.М., Корольук В.С. Об одной гипотезе Б. Гнеденко // ТВП. — 1961. — Т. 6, вып. 4. — С. 469–474.
- [43] Золотарёв В.М., Сенатов В.В. Двусторонние оценки метрики Леви // ТВП. — 1975. — Т. 20, вып. 2. — С. 239–250.
- [44] Зубков А.М. Личное сообщение. — 1986.
- [45] Зубков А.М., Севастьянов Б.А., Чистяков В.П. Сборник задач по теории вероятностей. (2-е изд.) — М.: Наука, 1989.
- [46] Ибрагимов И.А. Некоторые предельные теоремы для стационарных процессов // ТВП. — 1962. — Т. 7, вып. 4. — С. 361–392.
- [47] Ибрагимов И.А., Линник Ю.В. Независимые и стационарно связанные величины. — М.: Наука, 1965.
- [48] Ибрагимов И.А. Замечание о ЦПТ для зависимых случайных величин // ТВП. — 1975. — Т. 20, вып. 1. — С. 134–139.
- [49] Ибрагимов И.А. Об условиях гладкости траекторий случайных функций // ТВП. — 1983. — Т. 28, вып. 2. — С. 229–250.
- [50] Ибрагимов И.А., Розанов Ю.А. Гауссовские случайные процессы. — М.: Наука, 1970.
- [51] Ито К. Вероятностные процессы. (Вып. II). — М.: ИЛ, 1963.
- [52] Кабанов Ю.М. Интегральное представление функционалов от процессов с независимыми приращениями // ТВП. — 1974. — Т. 19, вып. 1. — С. 889–893.
- [53] Кабанов Ю.М. Личное сообщение. — 1985.
- [54] Кабанов Ю.М., Липцер Р.Ш., Ширяев А.Н. Слабая и сильная сходимость распределения считающих процессов // ТВП. — 1983. — Т. 28, вып. 2. — С. 288–319.
- [55] Каган А.М., Линник Ю.В., Рао С.Р. Характеризационные задачи математической статистики. — М.: Наука, 1972.
- [56] Каллианпур Г. Стохастическая теория фильтрации. — М.: Наука, 1987.
- [57] Кемени Дж., Снелл Дж., Кнепп А. Счетные цепи Маркова. — М.: Наука, 1987.
- [58] Кендалл М., Стьюарт А. Теория распределений. — М.: Наука, 1966.
- [59] Климов Г.П., Кузьмин А.Д. Вероятность, процессы, статистика. Задачи с решениями. — М.: Изд-во МГУ, 1985.
- [60] Колмогоров А.Н. Основные понятия теории вероятностей. (2-е изд.). — М.: Наука, 1974.
- [61] Колмогоров А.Н., Фомин С.В. Элементы теории функций и функционального анализа (4-е изд.). — М.: Наука, 1988.

- [62] Крамер Х., Лидбеттер М. Стационарные случайные процессы. — М.: Мир, 1969.
- [63] Крейн М.Г. Об одной экстраполяционной проблеме А. Н. Колмогорова // ДАН СССР. — 1944. — Т. 46, вып. 8. — С. 339–342.
- [64] Круглов В.М., Королев В.Ю. Предельные теоремы для случайных сумм. — М.: Изд-во МГУ, 1990.
- [65] Крылов Н.В. Управляемые процессы диффузионного типа. — М.: Наука, 1977.
- [66] Кузнецов С.Е. Личное сообщение. — 1985.
- [67] Ламперти Дж. Вероятность. — М.: Наука, 1973.
- [68] Ламперти Дж. Случайные процессы. — Киев: Вища школа, 1983.
- [69] Лебедев В.А. Личное сообщение. — 1985.
- [70] Леви П. Стохастические процессы и броуновское движение. — М.: Наука, 1972.
- [71] Линник Ю.В., Островский И.В. Разложение случайных величин и векторов. — М.: Наука, 1972.
- [72] Липцер Р.Ш. Личное сообщение. — 1985.
- [73] Липцер Р.Ш., Ширяев А.Н. Статистика случайных процессов. — М.: Наука, 1974.
- [74] Липцер Р.Ш., Ширяев А.Н. Теория мартингалов. — М.: Наука, 1986.
- [75] Лоэв М. Теория вероятностей. — М.: ИЛ, 1962.
- [76] Лукач Е. Характеристические функции. — М.: Наука, 1979.
- [77] Маккин Г. Стохастические интегралы. — М.: Мир, 1972.
- [78] Малиновский В.К. Личное сообщение. — 1988.
- [79] Мельников А.В. Личное сообщение. — 1983.
- [80] Молчанов С.А. Личное сообщение. — 1986.
- [81] Невё Ж. Математические основы теории вероятностей. — М.: Мир, 1969.
- [82] Новиков А.А. Об одном тождестве для стохастических интегралов // ТВП. — 1972. — Т. 17, вып. 4. — С. 761–765.
- [83] Новиков А.А. Об условия равномерной интегрируемости непрерывных неотрицательных мартингалов // ТВП. — 1979. — Т. 24, вып. 4. — С. 821–825.
- [84] Новиков А.А. Мартингалный подход в задаче о времени первого пересечения нелинейных границ // Труды МИАН им. В. А. Стеклова. — 1981. — Т. 158, вып. 3. — С. 130–152.
- [85] Новиков А.А. Личное сообщение. — 1985.
- [86] Павлов Х. Некоторые свойства распределений класса NBU // Математика и мат. образование. (4-я Конф. СБМ). — София: Академия, 1978. — С. 283–285.
- [87] Петров В.В. Суммы независимых случайных величин. — М.: Наука, 1972.
- [88] Портенко Н.И., Скороход А.В., Шуренков В.М. Марковские процессы. — М.: ВИНТИ, 1989.

- [89] Прохоров Ю.В. Сходимость случайных процессов и предельные теоремы теории вероятностей // ТВП. — 1956. — Т. 1, вып. 2. — С. 177–238.
- [90] Прохоров Ю.В. Несколько замечаний об усиленном законе больших чисел // ТВП. — 1959. — Т. 4, вып. 2. — С. 215–220.
- [91] Прохоров Ю.В. О суммах случайных векторов со значениями в гильбертовом пространстве // ТВП. — 1983. — Т. 28, вып. 2. — С. 254–258.
- [92] Прохоров Ю.В., Розанов Ю.А. Теория вероятностей. (3-е изд.). — М.: Наука, 1987.
- [93] Радавичус М. К вопросу об обобщении теоремы П. Леви // Литовский Мат. Сборник. — 1980. — Т. 20, вып. 4. — С. 129–131.
- [94] Райков Д.А. О разложении законов Гаусса и Пуассона // Изв. АН СССР, сер. мат. — 1938. — Т. 2. — С. 91–124.
- [95] Рамачандран Б. Теория характеристических функций. — М.: Наука, 1975.
- [96] Рао С.Р. Линейные статистические методы и их применения. — М.: Наука, 1968.
- [97] Розанов Ю.А. Стационарные случайные процессы. — М.: Физматгиз, 1963.
- [98] Ротарь В.И. Личное сообщение. — 1985.
- [99] Руссас Дж. Контигуальность вероятностных мер. — М.: Мир, 1975.
- [100] Секей Г. Парадоксы в теории вероятностей и математической статистике. — М.: Мир, 1990.
- [101] Соловьев А.Д. Личное сообщение. — 1985.
- [102] Спидер Ф.Л. Принципы случайного блуждания. — М.: Мир, 1969.
- [103] Сургайлис Д. Характеризация супермартингала некоторыми моментами остановки // Литовский Мат. Сборник. — 1974. — Т. 14, вып. 1. — С. 189–193.
- [104] Фалин Г.И. Личное сообщение. — 1985.
- [105] Феллер В. Введение в теорию вероятностей и ее приложения. Т. 1. — М.: Мир, 1984.
- [106] Феллер В. Введение в теорию вероятностей и ее приложения. Т. 2. — М.: Мир, 1984.
- [107] Филиппу А.Н. Пуассоновские и сложные пуассоновские распределения порядка k и некоторые их свойства // — Записки Научных Семинаров ЛОМИ. — 1983. — Т. 130. — С. 175–180.
- [108] Халмош П.Р. Теория меры. — М.: ИЛ, 1953.
- [109] Хенгартнер В., Теодореску Р. Функции концентрации. — М.: Наука, 1980.
- [110] Цирельсон Б.С. Один пример СДУ, не имеющего сильного решения // ТВП. — 1975. — Т. 20, вып. 2. — С. 427–429.
- [111] Чжун К. Однородные цепи Маркова. — М.: Мир, 1964.
- [112] Чжун К., Уильямс Р. Введение в стохастическое интегрирование. — М.: Мир, 1987.
- [113] Шервашидзе Т.Л. Личное сообщение. — 1986.

- [114] Ш и р я е в А. Н. Случайные процессы. — М.: Изд-во МГУ, 1972.
- [115] Ш и р я е в А. Н. Личное сообщение. — 1985.
- [116] Ш и р я е в А. Н. Вероятность. (2-е изд.). — М.: Наука, 1989.
- [117] Ш и р я е в А. Н. Вероятность. В 2-х книгах. Издание 5-е, стер. — М.: МЦНМО, 2011.
- [118] Ш и р я е в А. Н. Задачи по теории вероятностей. (2-е изд.). — М.: МЦНМО, 2011.
- [119] Ш у р М. Г. Личное сообщение. — 1985.
- [120] Э л л и о т т Р. Стохастический анализ и его приложения. — М.: Мир, 1986.
- [121] Я г л о м А. М. Введение в теорию стационарных случайных функций // УМН. — 1953. — Т. 7, вып. 5. — С. 3–168.
- [122] A h s a n u l l a M., S i n h a B. On normality via conditional normality. — Raleigh: Noth Carolina State Univ., 1986.
- [123] A l d o u s D. Exchangeability and related topics // LNM. — 1985. — V. 1117. — P. 1–186.
- [124] A l - H u s s a i n i A., E l l i o t t R. Markov bridges and enlarged filtrations // Canad. J. Statist. — 1989. — V. 17. — P. 329–332.
- [125] A l v o M., C a b i l l o P., F e i g i n P. A class of martingales with non-symmetric distributions // ZW. — 1981. — V. 58. — P. 87–93.
- [126] A m b a r t z u m i a n R. V. Personal communication. — 1982.
- [127] A s h R. Basic Probability Theory. — New York: John Wiley and Sons, 1970.
- [128] A s h R. Real Analysis and Probability. — New York: Academic Press, 1972.
- [129] A s h R., G a r d n e r M. Topics in Stochastic Processes. — New York: Academic Press, 1975.
- [130] A z e m a J., G u n d y R., Y o r M. Sur l'intégrabilité uniforme des martingales continues // LNM. — 1980. — V. 784. — P. 51–61.
- [131] A z e m a J., Y o r M. (Eds). Temps locaux // Asterisque. — 1978. — V. 54–53.
- [132] A z l a r o v T., V o l o d i n N. On the discrete analog of the Marshall–Olkin distribution // LNM. — 1983. — V. 982. — P. 17–23.
- [133] B a r i n g h a u s L., H e n z e N., M o r g e n s t e r n D. Some elementary proofs of the normality of $XY/\sqrt{X^2+Y^2}$ when X and Y are normal // Computers and Math. with Appl. — 1988. — V. 15. — P. 943–944.
- [134] B a r i n g h a u s L., H e n z e N. Personal communication. — 1989.
- [135] B a r l o w M. One-dimensional SDE with no strong solution // J. London Math. Soc. (2). — 1982. — V. 26. — P. 335–347.
- [136] B a r l o w M., P e r k i n s E. One-dimensional SDEs involving a singular increasing process // Stochastics. — 1984. — V. 12. — P. 229–249.
- [137] B a r l o w M., P r o s h a n F. Tolerance limits for classes of distribution based on failure rate // AMS. — 1966. — V. 37. — P. 1593–1601.
- [138] B a r n d o r f f - N i e l s e n O. E., S h i r y a e v A. N. Change of Time and Change of Measure. — Singapore: World Scientific Publishers, 2011.

- [139] Bauer H. Probability Theory and Elements of Measure Theory. — London: Academic Press, 1981.
- [140] Baxendale P., Chigansky P., Liptser R. Asymptotic stability of the Wonham filter: ergodic and nonergodic signals // SIAM J. Control Optimiz. — 2004. — V. 43, № 2. — P. 643–669.
- [141] Behboodian J. Symmetric sum and symmetric product of two independent random variables // J. Theoretical Probability. — 1989. — V. 2. — P. 267–270.
- [142] Berg C. The cube of a normal distribution is indeterminate // AP. — 1988. — V. 16. — P. 910–913.
- [143] Berkes I. The LIL for subsequences of random variables // ZW. — 1974. — V. 30. — P. 209–251.
- [144] Biagini A., Hu Y., Oksendal B., Zhang T. Stochastic Calculus for Fractal Brownian Motion and Applications. — London: Springer, 2008.
- [145] Billingsley P. Probability and Measure. — New York: John Wiley and Sons, 1979.
- [146] Bissinger B. Stochastic independence versus intuitive independence // Two-Year College Math. J. — 1980. — V. 11. — P. 122–123.
- [147] Blackwell D., Dubins L. On existence and non-existence of proper regular conditional probabilities // AP. — 1975. — V. 3. — P. 741–752.
- [148] Blackwell D., Freedman D. The tail sigma-field of a Markov chain and a theorem of Orey // AMS. — 1964. — V. 35, № 3. — P. 1291–1295.
- [149] Blake L. Simple extension of measures and the preservation of regularity of conditional probabilities // Pacific J. Math. — 1973. — V. 46. — P. 355–359.
- [150] Blake L. Every asmart is a martingale in the limit. // J. London Math. Soc. (2). — 1978. — V. 18. — P. 381–384.
- [151] Blumental R., Gettoor R. Markov Processes and Potential Theory. — New York: Academic Press, 1968.
- [152] Bojdecki T. Discrete-time Martingales. — Warszawa: Warsz. Univ. Press, 1977.
- [153] Bojdecki T. Personal communication. — 1985.
- [154] Bradley R. A remark on the central limit question for dependent random variables // JAP. — 1980. — V. 17. — P. 94–101.
- [155] Bradley R. Counterexamples to the CLT under strong mixing conditions // Limit Theorems in Probability and Statistics. — Amsterdam: North-Holland, 1984. — P. 153–172.
- [156] Bradley R. A stationary, pairwise independent, absolutely regular sequence for which the CLT fails // Probab. Theory Related Fields. — 1989. — V. 81. — P. 1–10.
- [157] Breiman L. Probability. — Reading: Addison-Wesley, 1968.
- [158] Breiman L. Probability and Stochastic Processes. — Boston: Houghton Mifflin Co., 1969.
- [159] Brown J. Stochastic metrics // ZW. — 1972. — V. 24. — P. 49–62.
- [160] Buhler W., Mieshke K. On $(n-1)$ -wise and joint independence and normality of n random variables // Commun. in Statist. Theory and Methods. — 1981. — V. 10. — P. 927–930.

- [161] Bulinskii A.V., Shashkin A. Limit Theorems for Associated Random Fields and Related Systems. — New Jersey: World Sci., 2007.
- [162] Burdick D. A note on symmetric random variables // AMS. — 1972. — V. 43. — P. 2039–2040.
- [163] Burkholder D.L., Gundy R.F. Extrapolation and interpolation of quasilinear operators of martingales // Acta Math. — 1970. — V. 124. — P. 249–304.
- [164] Cacoullos Th. Personal communication. — 1985.
- [165] Cairoli R., Walsh J. Stochastic integrals in the plane // Acta Math. — 1975. — V. 134. — P. 111–183.
- [166] Cambanis S. The measurability of a stochastic process of second order and its linear space // Proc. Amer. Math. Soc. — 1975. — V. 47. — P. 467–475.
- [167] Cambanis S., Simons G., Stout W. Inequalities for $E[k(X, Y)]$ when the marginals are fixed // ZW. — 1976. — V. 36. — P. 285–294.
- [168] Capobianco M., Moluzzo J. Examples and Counterexamples in Graph Theory. — Amsterdam: North-Holland, 1978.
- [169] Castillo E., Galambos J. Bivariate distributions with normal conditionals // Simulation, Modelling and Development (Intern. Symp., Cairo'87). — Anaheim: Acta Press, 1987. — P. 59–62.
- [170] Chen R., Shepp L. On the sum of symmetric random variables // AmS. — 1983. — V. 7. — P. 236.
- [171] Cheridito P. Mixed fractional Brownian motion // Bernoulli. — 2001. — V. 7, № 6. — P. 913–934.
- [172] Cherny A.S. Some particular problems of martingale theory // From Stochastic Calculus to Mathematical Finance: The Shiryaev Festschrift. — Berlin: Springer, 2006. — P. 109–124.
- [173] Chigansky P. Private communication. — 2012.
- [174] Chitashvili R. Personal communication. — 1985.
- [175] Chitashvili R., Toronjadze T. On one-dimensional SDEs with unit diffusion coefficient: Structure of solutions // Stochastics. — 1981. — V. 4. — P. 281–315.
- [176] Chow Y., Teicher H. Almost certain summability of independent identically distributed random variables // AMS. — 1971. — V. 42. — P. 401–404.
- [177] Chow Y., Teicher H. Probability Theory: Independence, Interchangeability, Martingales. — New York: Springer-Verlag, 1978.
- [178] Chung K. Sur les lois de probabilité unimodales // C. R. Acad. Sci. Paris. — 1953. — V. 236. — P. 583–584.
- [179] Chung K. A Course in Probability Theory. (2nd edn). — New York: Academic Press, 1974.
- [180] Chung K. Lectures from Markov Processes to Brownian Motion. — New York: Springer-Verlag, 1982.
- [181] Chung K., Walsh J. To reverse a Markov process // Acta Math. — 1969. — V. 123. — P. 225–251.

- [182] Chung K., Fuchs W. On the distribution of values of sums of random variables // *Memoirs Amer. Math. Soc.* — 1951. — № 6.
- [183] Churchill E. Information given by odd moments. // *AMS.* — 1946. — V. 17. — P. 244–246.
- [184] Cinlar E. *Introduction to Stochastic Processes.* — Englewood Cliffs (N.J.): Prentice-Hall, 1975.
- [185] Clark J. The representation of functionals of Brownian motion by stochastic integrals // *AMS.* — 1970. — V. 41. — P. 1282–1295.
- [186] Clarke L. On marginal density functions of continuous densities // *AMM.* — 1975. — V. 82. — P. 845–846.
- [187] Cramér H. Über eine Eigenschaft der normalen Verteilungsfunktion // *Math. Z.* — 1936. — V. 41. — S. 405–414.
- [188] Cramér H. On some classes of non-stationary stochastic processes // *Proc. 4th Berkeley Symp. Math. Statist. Probab.* V. 2. — Berkeley: Univ. Calif. Press, 1972. — P. 57–58.
- [189] Cramér H. Stochastic processes as curves in Hilbert space // *TPA.* — 1964. — V. 9, 1. — P. 169–179.
- [190] Cramér H. *Random Variables and Probability Distributions.* — Cambridge: Cambr. Univ. Press, 1970.
- [191] Csörgő M., Révész P. *Strong Approximations in Probability and Statistics.* — Budapest: Akadémiai Kiadó, 1981.
- [192] Csörgő M., Tandori K., Totik V. On the strong LLN for pairwise independent random variables // *Acta Math. Hungar.* — 1983. — V. 42. — P. 319–330.
- [193] Cuculescu I. Nonuniformly integrable non-negative martingales // *Rev. Roum. Math. Pure Appl.* — 1970. — V. 15. — P. 327–337.
- [194] Cuesta J.A., Matran C. On the asymptotic behavior of sums of pairwise independent random variables // *Statist. Probab. Letters.* — 1991. — V. 11, № 2. — P. 201–210.
- [195] Cuevas A. Density estimation: Robustness versus consistency // *New Perspectives in Theoretical and Applied Statistics.* — New York: John Wiley and Sons, 1986. — P. 259–264.
- [196] Daley D., Hall P. Limit laws for the maximum of weighted and shifted i.i.d. r.v.'s. // *AP.* — 1984. — V. 12. — P. 571–587.
- [197] Delbaen F., Schachermayer W. A simple counterexample to several problems in the theory of asset pricing // *Math. Finance.* — 1998. — V. 8, № 1. — P. 1–11.
- [198] Dellacherie C. Un exemple de la theorie generale des processus // *LMN.* — 1970. — V. 124. — P. 60–70.
- [199] Dellacherie C., Meyer P. *Probabilities and Potential.* — Amsterdam: North-Holland, V. A, 1978; V. B, 1982.
- [200] Devroe L. *Non-uniform Random Variate Generation.* — New York: Springer-Verlag, 1986.
- [201] Dharmadhikari S., Jog-Dev K. *Unimodality, Convexity and Applications.* — Orlando: Academic Press, 1988.

- [202] Diaconis P., Dubins L. Finite exchangeable sequences // AP. — 1980. — V. 8. — P. 745–764.
- [203] Dobrić V. Counterexamples for the law of large numbers in some Banach spaces. — Aarhus: Aarhus Univ., 1983.
- [204] Dodunekova R. D. Personal communication. — 1985.
- [205] Doleans-Dade C. Une martingale uniformement integrable mais non localement de carré intégrable // LNM. — 1971. — V. 191. — P. 138–140.
- [206] Doob J. Classical Potential Theory and its Probabilistic Counterpart. — New York: Springer, 1984.
- [207] Drouot Mari D., Kotz S. Correlation and Dependence. — London: Imperial College Press, 2001.
- [208] Dryginoff M. B. L. Personal communication. — 1998.
- [209] Dudley R. Sample functions of the Gaussian processes // AP. — 1973. — V. 1. — P. 66–103.
- [210] Dudley R. Probabilities and Metrics. — Aarhus: Aarhus Univ., 1976.
- [211] Dudley R. Wiener functionals as Ito integrals // AP. — 1977. — V. 5. — P. 140–141.
- [212] Durrett R. Brownian Motion and Martingales in Analysis. — Monterey (CA): Wadsworth and Brooks, 1984.
- [213] Edgar G., Sucheston L. Amarts: A class of asymptotic martingales. Parts A and B // J. Multivariate Analysis. — 1976. — V. 6. — P. 193–221; — P. 572–591.
- [214] Edgar G., Sucheston L. Martingales in the limit and amarts // Proc. Amer. Math. Soc. — 1977. — V. 67. — P. 315–320.
- [215] Embrechts P., Maejima M. Selfsimilar Processes. — Princeton: Princeton Univ. Press, 2002.
- [216] Emery M. Covariance des semimartingales gaussiennes // C. R. Acad. Sci. Paris. Ser. I. — 1982. — P. 703–705.
- [217] Emmanuele G., Villani A. Problem 6452 // AMM. — 1984. — V. 91. — P. 144.
- [218] Enchev O. Personal communication. — 1985.
- [219] Enchev O. Doob–Meyer decomposition for processes which might not be semimartingales (L^2 -approach) // Statistics and Control of Stochastic Processes (Proc. Steklov Seminar'85), — New York: Optimization Software, 1988.
- [220] Enchev O. Hilbert-space-valued quasimartingales // Bolletino Unione Mat. Italiana. (7). — 1988. — V. 2. — P. 19–39.
- [221] Engelbert H., Schmidt W. On solutions of one-dimensional stochastic equations without drift // ZW. — 1985. — V. 68. — P. 432–433.
- [222] Engelbert H. On uniqueness of solutions to stochastic dimensional equations: a counterexample // AP. — 2002. — V. 30, № 3. — P. 1039–1043.
- [223] Enis P. On the relation $E(E[X|Y]) = EX$ // Biometrika. — 1973. — V. 60. — P. 432–433.
- [224] Ephremides A., Thomas J. On random processes linearly equivalent to white noise // Information Science. — 1974. — V. 7. — P. 133–156.

- [225] Ethier S., Kurtz T. Markov Processes. Characterization and Convergence. — New York: John Wiley and Sons, 1986.
- [226] Feller W. A limit theorem for random variables with infinite moments // Amer. J. Math. — 1946. — V. 68. — P. 257–262.
- [227] Feller W. Non-Markovian processes with the semigroup property // AMS. — 1959. — V. 30. — P. 1252–1253.
- [228] Fernandez De La Vega W. On almost sure convergence of quadratic Brownian variation // AP. — 1974. — V. 2. — P. 551–552.
- [229] Fernique X. Regularite des trajectoires des fonctions aleatoires Gaussiennes // LNM. — 1974. — V. 480. — P. 1–96.
- [230] Fisher L., Walkup D. An example of the difference between the Lévy and Lévy–Prohorov metrics // AMS. — 1969. — V. 40. — P. 322–324.
- [231] Fisz M. Probability Theory and Mathematical Statistics. (3rd ed.). — New York: John Wiley and Sons, 1963.
- [232] Föllmer H. Personal communication. — 1986.
- [233] Föllmer H., Wu C.-T., Yor M. On weak Brownian motion of arbitrary order // Ann. Inst. Henri Poincaré. Probab. et Statist. — 2000. — V. 36, № 4. — P. 447–487.
- [234] Fortet R. Elements of Probability Theory. — London: Gordon and Breach, 1977.
- [235] Foss S., Korshunov D., Zachary S. An Introduction to Heavy-Tailed and Subexponential Distributions. — Berlin, Springer, 2011.
- [236] Franken P., Lisek B. On Wald's identity for dependent variables // ZW. — 1982. — V. 60. — P. 143–150.
- [237] Fréchet M. Sur les tableaux de correlation dont les marges sont donnés // Ann. Univ. Lyon. — 1951. — V. 14. — P. 53–77.
- [238] Freedman D. Brownian Motion and Diffusion. — San Francisco: Holden-Day, 1971.
- [239] Freedman D. A mixture of i.i.d. r.v.'s. need not admit a regular conditional probability given the exchangeable σ -field // ZW. — 1980. — V. 51. — P. 239–248.
- [240] Fu J. The moments do not determine a distribution // AmS. — 1984. — V. 38. — P. 294.
- [241] Gaidov S.D. Personal communication. — 1986.
- [242] Galambos J., Kotz S. Characterization of Probability Distributions. (LNM 675). — Berlin: Springer, 1978.
- [243] Galchuk L. Gaussian semimartingales // Statistics and Control of Stochastic Processes (Proc. Steklov Seminar'84). — New York: Optimization Software, 1985. — P. 102–121.
- [244] Gänsler P., Stute W. Wahrscheinlichkeitstheorie. — Berlin: Springer-Verlag, 1977.
- [245] Gavazki I. Personal communication. — 1989.
- [246] Geisser S., Mantel N. Pairwise independence of jointly dependent variables // AMS. — 1962. — V. 33. — P. 290–291.

- [247] Gelbaum B. Independence of events and of random variables // ZW. — 1976. — V. 36. — P. 333–343.
- [248] Gelbaum B. Some theorems in probability theory // Pacific J. Math. — 1985. — V. 118. — P. 383–391.
- [249] Geller N. Some examples of the weak and strong LLN for averages of mutually independent random variables // AmS. — 1978. — V. 32. — P. 34–36.
- [250] Gilat D. Convergence in distribution, convergence in probability and almost sure convergence of discrete martingales // AMS. — 1972. — V. 43. — P. 1374–1379.
- [251] Gleser L. On the distribution of the number of successes in independent trials // AP. — 1975. — V. 3. — P. 182–188.
- [252] Gnedenko B.V. Sur la distribution limite du terme maximum d'une série aléatoire // Annals of Math. — 1943. — V. 44. — P. 423–453.
- [253] Gnedenko B.V. Limit theorems for sums of random number of positive r.v.'s. // Proc. 6-th Berkeley Symp. Math. Statist. V. II. — Berkeley: Univ. Calif. Press, 1972. — P. 537–549.
- [254] Gnedenko B.V., Solov'yev A.D. On the conditions for existence of final probabilities for a Markov process. // Math. Operationsforsch. Statist. — 1973. — V. 4. — P. 379–390.
- [255] Gnedenko B.V., Korolev V.Yu. Random Summation: Limit Theorems and Applications. — London: CRC Press, 1996.
- [256] Goldman J. Stochastic point processes: Limit theorems // AMS. — 1967. — V. 38. — P. 771–779.
- [257] Greenwood P. Asymptotics of randomly stopped sequences with independent increments // AP. — 1973. V. 1. — P. 317–321.
- [258] Grenander U., Rosenblatt M. Statistical Analysis of Stationary Time Series. — New York: John Wiley and Sons, 1957.
- [259] Griffiths R. Infinitely divisible multivariate gamma distributions // Sankhya, — 1970. — V. A32. — P. 393–404.
- [260] Grigelionis B. Personal communication. — 1985.
- [261] Grimmett G., Stirzaker D. Probability and Stochastic Processes. — Oxford: Clarendon Press, 1982.
- [262] Grunbaum F. An inverse problem for Gaussian processes // Bull. Amer. Math. Soc. — 1972. — V. 78. — P. 615–616.
- [263] Gumbel E. Distributions a plusieurs variables dont les marges sont donnés // C. R. Acad. Sci. Paris. — 1958. — V. 246. — P. 2717–2720.
- [264] Gut A., Schmidt K. Amarts and Set Function Processes. (LNM 1042). — Berlin: Springer, 1983.
- [265] Gyöngy I. Personal communication. — 1985.
- [266] Hahn M., Klass M. The multidimensional CLT for arrays normed by affine transformations // AP. — 1981. — V. 9. — P. 611–623.
- [267] Hall P. A converse to the Spitzer–Rosen theorem // AP. — 1981. — V. 9. — P. 633–641.

- [268] Hall P., Heyde C.C. Martingale Limit Theory and Its Application. — New York: Academic Press, 1980.
- [269] Hamdan D. An ergodic Markov chain is not determined by any p -marginals // Indag. Math. — 2002. — V. 13, № 4. — P. 487–498.
- [270] Hamedani G. On two recent characterizations of multivariate normal distribution // Metrika. — 1988. — V. 35. — P. 41–47.
- [271] Hamedani G., Tata M. On the determination of the bivariate normal distribution from distributions of linear combinations of the variables // AMM. — 1975. — V. 82. — P. 913–915.
- [272] Han C. Dependence of random variables // AmS. — 1971. — V. 25. — № 4. — P. 35.
- [273] Hardin C. A spurious Brownian motion // Proc. Amer. Math. Soc. — 1985. — V. 93. — P. 350.
- [274] Hawkes J. On the covering of small sets by random intervals // Quart. J. Math. — 1973. — V. 24. — P. 427–432.
- [275] Heilmann W.-R., Schroter K. Eine Bemerkung über bedingte Wahrscheinlichkeiten, bedingte Erwartungswerte und bedingte Unabhängigkeit // Blatter. — 1987. — V. 28. — S. 119–126.
- [276] Hendrickx I.M.G. Personal communication. — 2012.
- [277] Herrndorf N. Stationary strongly mixing sequences not satisfying the CLT // AP. — 1983. — V. 11. — P. 809–813.
- [278] Herrndorf N. A functional CLT for weakly dependent sequences of r.v.'s. // AP. — 1984. — V. 12. — P. 141–153.
- [279] Hettmansperger T., Klimko L. A note on the strong convergence of distributions // AS. — 1974. — V. 2. — P. 597–598.
- [280] Heyde C.C. On a property of the lognormal distribution // J. Royal Statist. Soc. — 1963. — V. B29. — P. 392–393.
- [281] Heyde C.C. Some remarks on the moment problem. I and II // Qurt. J. Math. (2). — 1963. — V. 14. — P. 91–96; — P. 97–105.
- [282] Heyde C.C. Personal communication. — 1986.
- [283] Hida T. Canonical representations of Gaussian processes and their applications // Memoirs Colloq. Sci. Univ. Kyoto. — 1960. — V. 23. — P. 109–155.
- [284] Hill B., Lane D., Sudderth W. A strong law for some generalized urn processes // AP. — 1980. — V. 8. — P. 214–226.
- [285] Holewijn P., Hordijk A. On the convergence of moments in stationary Markov chains // SPA. — 1975. — V. 3. — P. 55–64.
- [286] Hoover D., Keisler H. Adapted probability distributions // Trans. Amer. Math. Soc. — 1984. — V. 286. — P. 159–201.
- [287] Hsu P., Robbins H. Complete convergence and the LLN // Proc. Nat. Acad. Sci. USA. — 1947. — V. 33. — P. 25–31.
- [288] Ibragimov I.A. On a problem of C. R. Rao on infinitely divisible laws // Sankhya. — 1972. — V. A34. — P. 447–448.
- [289] Ignatov Zv. Personal communication. — 1988.

- [290] Iosifescu M. Finite Markov Processes and Their Applications. — Chichester: John Wiley and Sons, 1980.
- [291] Iosifescu M. Personal communication. — 1985.
- [292] Isaacson D. Continuous martingales with discontinuous marginal distributions // AMS. — 1971. — V. 42. — P. 2139–2142.
- [293] Isaacson D., Madsen R. Markov Chains. Theory and Applications. — New York: John Wiley and Sons, 1976.
- [294] Ito K. Introduction to Probability Theory. — Cambridge: Cambr. Univ. Press, 1984.
- [295] Ivković Z. Personal communication. — 1985.
- [296] Ivković Z., Bulatović J., Vukmirović J., Zivanović S. Applications of Spectral Multiplicity in Separable Hilbert Space to Stochastic Processes. — Belgrad: Matem. Inst., 1974.
- [297] Jacod J. Two dependent Poisson processes whose sum is still a Poisson process // JAP. — 1975. — V. 12. — P. 170–172.
- [298] Jacod J. Calcul Stochastique et Problème de Martingales. (LNM 714). — Berlin: Springer-Verlag, 1979.
- [299] Jacod J., Shiryaev A.N. Limit Theorems for Stochastic Processes. (2nd ed.). — Berlin: Springer, 2002.
- [300] Jagers A. Problem 650 // Nieuw Archief Voor Wiskunde, Ser. 4. — 1983. — V. 1. — S. 377–378.
- [301] Jain N., Monrad D. Gaussian quasimartingales // ZW. — 1982. — V. 59. — P. 139–159.
- [302] Jain N., Monrad D. Gaussian measures in B_p . // AP. — 1983. — V. 11. — P. 46–57.
- [303] Janson S. Some pairwise independent sequences for which the CLT fails // Stochastics. — 1988. — V. 23. — P. 439–448.
- [304] Jeablanс M., Yor M., Chesney M. Mathematical Methods for Financial Markets. — Berlin: Springer, 2009.
- [305] Jeulin T., Yor M. Inégalité de Hardy, Semimartingales et faux-amis // LNM. — 1979. — V. 721. — P. 332–359.
- [306] Joffe A. On a set of almost deterministic k -dependent random variables // AP. — 1974. — V. 2. — P. 161–162.
- [307] Johnson B. An inequality for conditional distribution // Math. Magazine. — 1974. — V. 47. — P. 281–283.
- [308] Johnson G., Helms L. Class (D) supermartingales // Bull. Amer. Math. Soc. — 1963. — V. 69. — P. 59–62.
- [309] Johnson N., Kotz S. Urn Models and Their Application. New York: John Wiley and Sons, 1977.
- [310] Kadota T., Shepp L. Conditions for absolute continuity between a certain pair of probability measures // ZW. — 1970. — V. 16. — P. 250–260.
- [311] Kaishev V. Personal communication. — 1986.

- [312] K a l e B. Normality of linear combinations of non-normal random variables // AMM. — 1970. — V. 77. — P. 992–995.
- [313] K a l l e n b e r g O. Conditions for the contibuity of random processes without discontinuities of second kind // AP. — 1973. — V. 1. — P. 519–526.
- [314] K a l p a z i d o u S. Personal communication. — 1985.
- [315] K a n t e r M. Stable densities under change and total variation inequalities // AP. — 1975. — V. 3. — P. 697–707.
- [316] K a r a t z a s I., S h r e v e S. Brownian Motion and Stochastic Calculus. — New York: Springer, 1988.
- [317] K a z a m a k i N. Changes of time, stochastic integrals and weak martingales // ZW. — 1972. — V. 22. — P. 25–32.
- [318] K a z a m a k i N. Examples of local martingales // LNM. — 1972. — V. 258. — P. 98–100.
- [319] K a z a m a k i N. On a stochastic integral equation with respect to a weak martingale // Tohoku Math. J. — 1974. — V. 26. — P. 53–63.
- [320] K a z a m a k i N., S e k i g i c h i T. Uniform integrability of continuous exponential martingales // Tohoku Math. J. — 1983. — V. 35. — P. 289–301.
- [321] K a z a m a k i N. Continuous Exponential Martingales and BMO. (LNM 1579). — Berlin: Springer, 1994.
- [322] K e n d a l l D. G. On finite and infinite sequences of exchangeable events // Studia Sci. Math. Hung. — 1967. — V. 2. — P. 319–327.
- [323] K e n d a l l D. G. Personal communication. — 1985.
- [324] K e n d a l l D. G., R a o K. On the generalized second limit theorem in the theory of probability // Biometrika. — 1950. — V. 37. — P. 224–230.
- [325] K h i n t c h i n e A. Korrelationstheorie der stationären stochastischen Prozesse // Math. Annalen. — 1934. — V. 109. — S. 604–615.
- [326] K i n g m a n J., T a y l o r S. Introduction to Measure and Probability. — Cambridge: Cambr. Univ. Press., 1966.
- [327] K l a s s M. Properties of optimal extended-valued stopping rules for s_n/n // AP. — 1973. — V. 1. — P. 719–757.
- [328] K o p a n o v P. Personal communication. — 2011.
- [329] K o p p P. Martingales and Stochastic Integrals. — Cambridge: Cambr. Univ. Press, 1984.
- [330] K o r a l o v L. B., S i n a i Y. G. Theory of Probability and Random Processes (2nd ed.). — Berlin: Springer, 2007.
- [331] K o r d z a h i a N. Personal communication. — 1996.
- [332] K o t l a r s k i I. Some deficiencies of using moment generating functions // AmS. — 1975. — V. 29. — P. 127–128.
- [333] K o t l a r s k i I. On characterizing the gamma and the normal distribution // Pacific J. Math. — 1967. — V. 20. — P. 69–76.
- [334] K o w a l s k i C. Nonnormal bivariate distributions with normal marginals // AmS. — 1973. — V. 27, № 3. — P. 103–106.

- [335] Krewski D., Bickis M. A note on independent and exhaustive events // AmS. — 1984. — V. 38. — P. 290–291.
- [336] Krickeberg K. Strong mixing properties of Markov chains with infinite invariant measure // Proc. 5-th Berkeley Symp. Math. Statist. Probab. V. 2, part II. — Berkeley: Univ. California Press, 1967. — P. 431–446.
- [337] Kronfeld B. Personal communication. — 1982.
- [338] Krylov N.V. A simple proof of a result of A. Novikov. — Preprint. Univ. Minnesota, 2002. — Electronic: arXiv:math.PR/0207013 v1.
- [339] Küchler U. Personal communication. — 1986.
- [340] Küchler U., Sorensen M. Exponential families of stochastic processes: A unified semimartingale approach // Intern. Statist. Review. — 1989. — V. 57. — P. 123–144.
- [341] Kuelbs J. A counterexample for Banach space valued random variables // AP. — 1976. — V. 4. — P. 684–689.
- [342] Kurtz T. A note on sequences of continuous parameter Markov Chains // AMS. — 1969. — V. 40. — P. 1078–1082.
- [343] Kwapien S. Personal communication. — 1983.
- [344] Laha R. An example of a nonnormal distribution where the quotient follows the Cauchy law // Proc. Nat. Acad. Sci. USA. — 1958. — V. 44. — P. 222–223.
- [345] Laha R., Rohatgi V. Probability Theory. — New York: John Wiley and Sons, 1979.
- [346] Lamperti J. Stochastic Processes: A Survey of the Mathematical Theory. — Berlin: Springer, 1997.
- [347] Lancaster H. Pairwise statistical independence // AMS. 1965. — V. 36. — P. 1313–1317.
- [348] Laube G. Weak convergence in the mean of distribution functions // Metrika. — 1973. — V. 20. — P. 103–105.
- [349] Laue G. Existence and representation of density functions // Math. Nachrichten. — 1983. — V. 114. — P. 7–21.
- [350] Le Breton A. Personal communication. — 1996.
- [351] Le Gall J., Yor M. Sur l'équation stochastique de Tsirelson // LNM. — 1983. — V. 986. — P. 81–88.
- [352] Lehmann E. Some concepts of dependence // AMS. — 1966. — V. 37. — P. 1137–1153.
- [353] Leipnik R. The lognormal distribution and strong non-uniqueness of the moment problem // ТБП. — 1981. — Т. 26, вып. 4. — С. 863–865.
- [354] Lévy P. Le Mouvement brownien plan // Amer. J. Math. — 1940. — V. 62. — P. 487–550.
- [355] Lévy P. The arithmetic character of Wishart's distribution // Proc. Cambr. Philos. Soc. — 1948. — V. 44. — P. 295–297.
- [356] Lin G.D., Stoyanov J. The logarithmic skew normal distributions are moment-indeterminate // JAP. — 2009. — V. 49, № 3. — P. 909–916.

- [357] Liptser R.Sh., Shiryaev A.N. On a problem of necessary and sufficient conditions in the functional CLT for local martingales // ZW. — 1982. — V. 59. — S. 311–318.
- [358] Lukacs E. Stochastic Convergence. (2nd ed.). — New York: Academic Press, 1975.
- [359] Lukacs E. Developments in Characteristic Function Theory. — London: Griffin, 1983.
- [360] Lyasoff A. Personal communication. — 2012.
- [361] Mališić J. Collection of Exercises in Probability Theory with Applications. — Belgrade: Gradjevinska Kniga, 1970.
- [362] Marcus D. Non-stable laws with all projections stable // ZW. — 1983. — V. 64. — P. 139–156.
- [363] Marcus M.B., Rosen J. Markov Processes, Gaussian Processes and Local Times. — Cambridge: Cambr. Univ. Press, 2006.
- [364] Marinescu E. Personal communication. — 1985.
- [365] Masry E., Cambanis S. The representation of stochastic processes without loss of information // SIAM J. Appl. Math. — 1973. — V. 25. — P. 628–633.
- [366] Mauldon J. Characterizing properties of statistical distributions // Quart. J. Math.(2). — 1956. — V. 7. — P. 155–160.
- [367] Melnick E., Tenenbein A. Misspecification of the normal distribution // AmS. — 1982. — V. 36. — P. 372–373.
- [368] Métivier M. Semimartingales. A Course on Stochastic Processes. — Berlin: Walter de Gruyter, 1982.
- [369] Métivier M., Pellaumail J. Stochastic Integration. — New York: Academic Press, 1980.
- [370] Meyer P-A., Zheng W. Tightness criteria for laws of Semimartingales // Ann. Inst. H. Poincaré. — 1984. — V. B20. — P. 353–372.
- [371] Monrad D. Lévy processes: absolute continuity of hitting times for points // ZW. — 1976. — V. 37. — P. 43–49.
- [372] Monroe I. A. s. convergence of the quadratic variation of martingales: a counterexample // AP. — 1976. — V. 4. — P. 133–138.
- [373] Moran P. A non-Markovian quasi-Poisson process // Studia Sci. Math. Hungar. — 1967. — V. 2 — P. 425–429.
- [374] Moran P. An Introduction to Probability Theory. — Oxford Univ. Press, 1968.
- [375] Morgenstern D. Einfache Beispiele zweidimensionaler Verteilungen // Mitt. Math. Statist. — 1956. — V. 8. — S. 234–245.
- [376] Morrison J., Wise G. Continuity of filtrations of sigma algebras // Statistics and Probability Letters. — 1987. — V. 6. — P. 55–60.
- [377] Mucci A. Limits for martingale-like sequences // Pacific J. Math. — 1973. — V. 48. — P. 197–202.
- [378] Mutafchiev L. Personal communication. — 1986.
- [379] Nelson P. A class of orthogonal series related to martingales // AMS. — 1970. — V. 41. — P. 1684–1694.

- [380] Ne u t s M. Probability. — Boston: Allyn and Bacon, 1973.
- [381] Ne v e u J. Discrete Parameter Martingales. — Amsterdam: North-Holland, 1975.
- [382] N i v e n I. Diophantine Approximations. — New York: John Wiley and Sons, 1963.
- [383] O' B r i e n G. Pairwise independent random variables. // AP. — 1980. — V. 8. — P. 170–175.
- [384] O' B r i e n G. The occurrence of large values in stationary sequences // ZW. — 1982. — V. 61. — P. 347–353.
- [385] O' C o n n o r T. Infinitely divisible distributions with unimodal Lévy spectral functions // AP. — 1979. — V. 7. — P. 494–499.
- [386] O l k i n I., G l e s e r L., D e r m a n C. Probability Models and Applications. — New York: Macmillan, 1980.
- [387] O r d J. The discrete Student's distribution // AMS. — 1968. — V. 39. — P. 1513–1516.
- [388] O s t r o v s k a S., S t o y a n o v J. Stieltjes classes for M -indeterminate powers of inverse Gaussian distributions // Statist. Probab. Letters. — 2005. — V. 71, № 2. — P. 165–171.
- [389] O v e r d i j k D., S i m o n s F., T h i e m a n n J. A comment on unions of rings // Indagationes Math. — 1979. — V. 41. — P. 439–441.
- [390] P a p a g e o r g i o u H. Personal communication. — 1985.
- [391] P a p o u l i s A. Probability, Random Variables and Stochastic Processes. — New York: McGraw-Hill, 1965.
- [392] P a r z e n E. Modern Probability Theory and Its Applications. — New York: John Wiley and Sons, 1960.
- [393] P a r z e n E. Stochastic Processes. — San Francisco: Holden-Day, 1962.
- [394] P f a n z a g l J. On the existence of regular conditional probabilities // ZW. — 1969. — V. 11. — P. 244–256.
- [395] P h i l i p p o u A., H a d j i c h r i s t o s J. A note on the Poisson distribution of order k and a result of Raikov. — Patras: Univ. Patras, 1985.
- [396] P i e g o r s c h W., C a s e l l a G. The existence of the first negative moments // AmS. — 1985. — V. 39. — P. 60–62.
- [397] P i e r c e D., D y k s t r a R. Independence and the normal distribution // AmS. — 1969. — V. 23, № 4. — P. 39.
- [398] P i r i n s k y Ch. Personal communication. — 2008.
- [399] P i t m a n E., W i l l i a m s E. Cauchy-distributed functions of Cauchy variates // AMS. — 1967. — V. 38. — P. 916–918.
- [400] P r o t t e r P. E. Stochastic Integration and Differential Equations. (2nd ed.). — Berlin: Springer, 2004.
- [401] R a m a c h a n d r a n D. Mixtures of perfect probability measures // AP. — 1974. — V. 2. — P. 496–500.
- [402] R a m a c h a n d r a n D. On the two definitions of independence // Colloq. Math. — 1975. — V. 32. — P. 227–231.
- [403] R a o M. Stochastic Processes and Integration. — Alphen (NL): Sijthoff and Noordhoff, 1979.

- [404] Rao M., Swift R. Probability Theory with Applications. (2nd ed.). — Berlin: Springer, 2006.
- [405] Rényi A. Probability Theory. — Budapest: Akadémiai Kiadó, 1967.
- [406] Resnik S. Record values and maxima // AP. — 1973. — V. 1. — P. 650–662.
- [407] Révész P. The Laws of Large Numbers. — Budapest: Akadémiai Kiadó, 1967.
- [408] Revuz D., Yor M. Continuous Martingales and Brownian Motion. (3rd ed.). — Berlin: Springer, 2003.
- [409] Riedel M. On the one-sized tails of infinitely divisible distributions // Math. Nachrichten. — 1975. — V. 70. — P. 115–163.
- [410] Robbins H. Convergence of distributions // AMS. — 1948. — V. 19. — P. 72–76.
- [411] Robertson T. A smoothing property for conditional expectation given σ -lattices // AMM. — 1968. — V. 75. — P. 515–518.
- [412] Robinson P.M. Personal communication. — 1985.
- [413] Rogers L., Williams D. Diffusions, Markov Processes and Martingales. Vol. 2: Ito Calculus. — Chichester: John Wiley and Sons, 1987.
- [414] Rogers C. Arbitrage with fractional Brownian motion // Math. Finance. — 1997. — V. 7, № 1. — P. 95–105.
- [415] Rohatgi V. Personal communication. — 1986.
- [416] Rohatgi V., Steutel F., Székely G. Infinite divisibility of products and quotients of i.i.d. r.v.'s. // The Math. Scientist. — 1990. — V. 15. — P. 53–59.
- [417] Romano J., Siegel A. Counterexamples in Probability and Statistics. — Monterey (CA): Wadsworth and Brooks, 1986.
- [418] Rosenberg L. Nonnormality of linear combinations of normally distributed r.v.'s. // AMM. — 1965. — V. 72 — P. 888–890.
- [419] Rosenblatt M. A CLT and strong mixing condition // Proc. Nat. Acad. Sci. USA. — 1956. — V. 42. — P. 43–47.
- [420] Rosenblatt M. Markov Processes. Structure and Asymptotic Behavior. — Berlin: Springer-Verlag, 1971.
- [421] Rosenblatt M. Random Processes. — New York: Springer-Verlag, 1974.
- [422] Rosenblatt M. Dependence and asymptotic independence for random processes // Studies in Probability Theory. Washington, Math. Assoc. Amer., 1978. — P. 24–45.
- [423] Rossberg H.-J. Personal communication. — 1989.
- [424] Rossberg H.-J., Jesiak B., Siegel G. Analytic Methods of Probability Theory. — Berlin: Akademie-Verlag, 1985.
- [425] Roussas G. A First Course in Mathematical Statistics. — Reading (MA): Addison-Wesley, 1973.
- [426] Rozovskiĭ B.L. Personal communication. — 1985.
- [427] Rutkowski M. Fundamental solutions of SDEs with drift // Stochastic. — 1989. — V. 26. — P. 193–204.
- [428] Salisbury T. An increasing diffusion // Seminar in Stochastic Processes. — Basel: Birkhäuser, 1986. — P. 173–194.

- [429] Schoenberg I. Problem 650 // *Nieuw Archiff Vor Wiskunde*, Ser. 4. — 1983. — V. 1. — P. 377–378.
- [430] Serfling R. *Approximation Theorems of Mathematical Statistics*. — New York: John Wiley and Sons, 1980.
- [431] Seshadri V. Personal communication. — 1986.
- [432] Shaked M., Shanthikumar J. G. *Stochastic Orders*. — Berlin, Springer, 2007.
- [433] Shanbhad D., Pestana D., Sreehari M. Some further results in infinite divisibility // *Math. Proc. Cambr. Philos. Soc.* — 1977. — V. 82. — P. 289–295.
- [434] Shohat J., Tamarkin J. *The Problem of Moments*. — New York: Amer. Math. Soc., 1943.
- [435] Sibley D. A metric for weak convergence of distribution functions // *Rocky Mountain J. Math.* — 1971. — V. 1. — P. 437–440.
- [436] Simons G. An unexpected expectation // *AP*. — 1977. — V. 5. — P. 157–158.
- [437] Solovay R. A model of set theory in which every set of reals is Lebesgue measurable // *An. Math.* — 1970. — V. 92. — P. 1–56.
- [438] Speakman J. Two Markov chains with common skeleton // *ZW*. — 1967. — V. 7. — S. 224.
- [439] Staudte R., Tata M. Complex roots of real characteristic functions // *Proc. Amer. Math. Soc.* — 1970. — V. 25. — P. 238–246.
- [440] Steen L., Seebach J. *Counterexamples in Topology*. (2nd ed.). — New York: Springer, 1978.
- [441] Steutel F. W. *Preservation of Infinite Divisibility Under Mixing*. — Amsterdam: Math. Centre, 1970.
- [442] Steutel F. W. Problem 153 and its solution // *Statistica Neerlandica*. — 1984. — V. 38. — P. 215.
- [443] Steutel F. W., Van Harn K. *Infinite Divisibility of Probability Distributions on the Real Line*. — Hoboken (NJ): Marcel Dekker, 2003.
- [444] Stout W. *Almost Sure Convergence*. — New York: Academic Press, 1974.
- [445] Stout W. On convergence of ϕ -mixing sequences of random variables // *ZW*. — 1974. — V. 31. — P. 69–70.
- [446] Stout W. Almost sure invariance principle when $EX_1^2 = \infty$ // *ZW*. — 1979. — V. 49. — S. 23–32.
- [447] Stoyanov J., Mirazchiiski I., Ignatov Z., Tanushev M. *Exercise Manual in Probability Theory*. — Dordrecht: Kluwer Academic Publishers, 1989.
- [448] Stoyanov J. The counterexamples are needed for the effective teaching and learning Probability theory // *Studies in Mathematics Education*. V. 7. — Paris: UNESCO, 1989. — P. 165–174.
- [449] Stoyanov J. *Counterexamples in Probability*. (2nd ed.). — Chichester: John Wiley and Sons, 1997.
- [450] Stoyanov J. Krein condition in probabilistic moment problems // *Bernoulli*. — 2000. — V. 6, № 5. — P. 939–949.

- [451] Stoyanov J. Stieltjes classes for moment indeterminate probability distributions // JAP.— 2004. — V. 41A. — P. 281–294.
- [452] Stoyanov J., Tolmatz L. New Stieltjes classes involving generalized gamma distributions // Statist. Probab. Letters. — 2004. — V. 69, № 2. — P. 213–219.
- [453] Stoyanov J. Sets of binary random variables with a prescribed independence dependence structure // The Math. Scientist. — 2005. — V. 28, № 1. — P. 19–27.
- [454] Stricker C. Quasimartingales, martingales locales, semimartingales et filtrations naturelles // ZW. — 1977. — V. 39. — P. 55–64.
- [455] Stricker C. Integral representation in the theory of continuous trading // Stochastic. — 1984. — V. 13. — P. 249–265.
- [456] Stroock D., Varadhan S. Multidimensional Diffusion Processes. — New York: Springer, 1979.
- [457] Szasz D. Once more on the Poisson process // Studia Sei. Math. Hungar. — 1970. — V. 5. — P. 441–444.
- [458] Takacs L. Solution to Problem 6452 // AMM. — 1985. — V. 92. — P. 515.
- [459] Takahasi K. An example of a sequence of frequency function which converges to a frequency function in the mean of order 1 but nowhere // J. Japan Statist. Soc. — 1971/72/ — V. 2. — P. 33–34.
- [460] Tanny D. A zero-one law for stationary sequences // ZW. — 1974. — V. 30. — P. 139–148.
- [461] Taylor R., Daffer P., Patterson R. Limit Theorems for Sums of Exchangeable Random Variables. — Totowa (N.J.): Rowman and Allanheld Publishers, 1985.
- [462] Thomas J. An Introduction to Applied Probability and Random Processes. — New York: John Wiley and Sons, 1971.
- [463] Thomasian A. The Structure of Probability Theory with Applications. — New York: McGraw-Hill, 1969.
- [464] Tomkins R. Properties of martingale-like sequences // Pacific J. Math. — 1975. — V. 61. — P. 521–525.
- [465] Tomkins R. Martingale generalizations // Topics in Applied Statistics. — Montreal: Concordia Univ., 1984. — P. 537–548.
- [466] Tomkins R. Personal communication. — 1985.
- [467] Tweedie R. Sufficient conditions for ergodicity and recurrence of Markov chains on a general state space // SPA. — 1975. — V. 3. — P. 385–403.
- [468] Tzankova L.E. Personal communication. — 2011.
- [469] Urumova V.I. Personal communication. — 2011.
- [470] Urusov M.A. Personal communication. — 2012.
- [471] Ushakov N.G. Selected Topics in Characteristic Functions. — Utrecht: VSP, 1999.
- [472] Van der Hoeven P.C.T. On Point Processes. — Amsterdam: Math. Centre, 1983.
- [473] Vandev D. Personal communication. — 1986.

- [474] Van Handel R. On the exchange of intersection and supremum of σ -fields in filtering theory // Israel J. Math. — 2012. — To appear.
- [475] Vasudeva R. Chover's LIL and weak convergence // Acta Math. Hungar. — 1984. — V. 44. — P. 215–221.
- [476] Walsh J. A non-reversible semimartingale // LNM. — 1982. — V. 920. — P. 212.
- [477] Wang A. Quadratic variation of functionals of Brownian motion // AP. — 1977. — V. 5. — P. 756–769.
- [478] Wang Y.H. Dependent random variables with independent subsets // AMM. — 1979. — V. 86. — P. 290–292.
- [479] Wang Y.H., Stoyanov J., Shao Q.-M. On independence and dependence properties of a set of random events // AmS. — 1993. — V. 47. — P. 112–115.
- [480] Williams D. Diffusions, Markov Processes and Martingals. V. 1: Foundations. — Chichester: John Wiley and Sons, 1979.
- [481] Williams R. Reflected Brownian motion in a wedge: semimartingale property // ZW. — 1985. — V. 69. — P. 161–176.
- [482] Wintner A. The Fourier Transforms of Probability Distributions. — Baltimore (MD), 1947.
- [483] Wise G., Hall E. Counterexamples in Probability and Real Analysis. — Oxford: Oxford Univ. Press, 1993
- [484] Wittmann R. A general law of iterated logarithm // ZW. — 1985. — V. 68. — P. 521–543.
- [485] Wong C. A note on mutually independent events // AmS. — 1972. — V. 26. — P. 27–28.
- [486] Wong E. Stochastic Processes in Information and Dynamical Systems. — New York: McGraw-Hill, 1971.
- [487] Wrobel A. On the almost sure convergence of the square variation of the Brownian motion // Probab. and Math. Statistics. — 1982. — V. 3. — P. 97–101.
- [488] Yamada T., Watanabe S. On the uniqueness of solutions of SDEs. I and II // J. Math. Kyoto Univ. — 1971. — V. 11. — P. 115–167; — P. 553–563.
- [489] Yanev N.M. Personal communication. — 1985.
- [490] Yanev G.P. Personal communication. — 2012.
- [491] Yarho-Zvonkin A.-A. Private communication. — 2012.
- [492] Yeh J. Stochastic Processes and the Wiener Integrals. — New York: Marcel Dekker, 1973.
- [493] Yor M. Un exemple de processus qui n'est pas une semi-martingale // Asterisque. — 1978. — V. 52–53. — P. 219–221.
- [494] Yor M. Personal communication. — 1986.
- [495] Yzeren J. A bivariate distribution with instructive properties as to normality correlation and dependence // Statistica Neerlandica. — 1972. — V. 26. — P. 55–56.
- [496] Zabczyk J. Personal communication. — 1985.
- [497] Zbaganu G. Personal communication. — 1985.
- [498] Zygmund A. A remark on characteristic functions // AMS. — 1947. — V. 18. — P. 272–276.