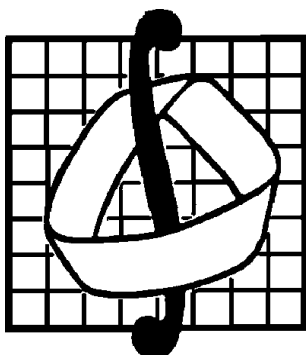


МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М.В. ЛОМОНОСОВА



Механико-математический факультет

СОВРЕМЕННЫЕ ПРОБЛЕМЫ МАТЕМАТИКИ И МЕХАНИКИ

Том IV

Математика

Выпуск 1

Теория вероятностей и математическая статистика



Издательство Московского университета
2009 год

УДК 519.2

ББК 22.17

С 56

СОВРЕМЕННЫЕ ПРОБЛЕМЫ МАТЕМАТИКИ И МЕХАНИ-
С 56 **КИ. Том IV. Математика. Выпуск 3. Теория вероятностей и ма-**
тематическая статистика / Под редакцией А.Н. Ширяева. — М.: Изд-во
МГУ, 2009. — 180 с.

ISBN 978—5—211—05652—7

Выпуск посвящен теории вероятностей, математической статистике и их приложениям.

Механико-математический факультет МГУ выражает
благодарность Попечительскому совету факультета
и выпускникам О.Д. Звягину и А.В. Чеглакову
за помощь в издании сборника.

ISBN 978—5—211—05652—7

© Механико-математический
факультет МГУ, 2009 г.

Сборник посвящается 70-летию со дня рождения

Виктора Антоновича Садовниченко

Раздел 1

Исследования кафедры теории вероятностей

Введение

Кафедра теории вероятностей была образована на механико-математическом факультете МГУ в 1935 году. Ее возглавил великий русский ученый А.Н. Колмогоров. Впоследствии кафедра стала одним из основных центров по подготовке специалистов и по научным исследованиям в области теории вероятностей и математической статистики в нашей стране. Многие поколения советских и российских ученых в этой области науки считают себя питомцами этой кафедры. А.Н. Колмогоров руководил кафедрой с 1935 по 1966 год. С 1966 по 1995 год ею заведовал Б.В. Гнеденко. С 1996 года заведующим кафедрой теории вероятностей является чл.-корр. РАН, профессор А.Н. Ширяев.

На кафедре также работали такие всемирно известные ученые, как А.Я. Хинчин, В.И. Гливенко, Е.Е. Слуцкий, Е.Б. Дынкин, Р.Л. Добрушин, И.В. Гирсанов, А.Д. Соловьев, Я.Г. Синай, А.Д. Вентцель, М.И. Фрейдлин, С.А. Молчанов и др.

В настоящее время непосредственно на кафедре работают (в том числе по совместительству) 15 профессоров, 13 доцентов, 2 старших преподавателя, 5 ассистентов. На кафедре проводится 26 спецкурсов и 17 спецсеминаров.

В частности, в 2007 году появились новые спецкурсы: «Мартингалы и методы в граничных задачах для броуновского движения» (А.Н. Ширяев), «Углубленный курс теории вероятностей» (А.В. Булинский), «Функциональные предельные теоремы для зависимых случайных полей» (А.П. Шашкин), а также спецсеминар «Исследование асимптотического поведения и устойчивости стохастических моделей» (Л.Г. Афанасьева, Е.В. Булинская, Е.Б. Яровая).

Особо следует отметить общекафедральный научно-исследовательский («Большой») семинар под руководством А.Н. Ширяева, проходящий регулярно по средам, где делают доклады как сотрудники кафедры, так и приглашенные гости — ученые из России и из-за рубежа.

На кафедре ведутся обширные и разнообразные исследования в области теории вероятностей и ее приложений. Эти исследования регулярно поддерживаются грантами РФФИ (6 в 2008 году). В качестве основных направлений можно назвать следующие.

1. Теория вероятностей и математическая статистика (общие вопросы, предельные теоремы и их уточнения, теория статистических выводов).
2. Теория случайных процессов и стохастическое исчисление (предельные теоремы, семимартингалы, стохастические дифференциальные уравнения, броуновское движение, процессы Леви, ветвящиеся процессы, случайные блуждания, граничные задачи, теория экстремальных значений).

3. Применения к задачам математической физики, математической биологии, теории информации, теории массового обслуживания, актуарной и финансовой математики и др.

Сотрудники кафедры регулярно участвуют во всероссийских и международных конференциях, многие входили в оргкомитеты и программные комитеты конференций.

16–21 июня 2003 года под эгидой Российской Академии наук и Московского университета проводилась Международная конференция «Колмогоров и современная математика», посвященная 100-летию со дня рождения А.Н. Колмогорова. Основная нагрузка по организации конференции легла на сотрудников кафедры теории вероятностей во главе с А.Н. Ширяевым (вице-председателем оргкомитета конференции). Число участников конференции составило почти 1000 человек. Сотрудниками кафедры было сделано 15 докладов. Сборник тезисов конференции составил 916 страниц (ответственные — профессора Л.Г. Афанасьева и Е.В. Булинская). В секретариат оргкомитета конференции входили и активно работали также профессора А.В. Булинский, Г.И. Фалин, доценты Е.Б. Яровая, С.В. Жуленев и многие другие преподаватели кафедры.

В 2007 году было издано 8 книг, в том числе: А.Н. Ширяев «Мартингалльные методы в граничных задачах для броуновского движения», Е.Б. Яровая «Ветвящиеся случайные блуждания в неоднородной среде», А.В. Bulinskii, А.Р. Shashkin «Limit theorems for associated random fields and related topic», учебники и учебные пособия: А.Н. Ширяев «Вероятность» (четвертое издание, переработанное и дополненное), А.В. Лебедев «Сборник задач по математической экономике», С.В. Жуленев «Стохастическая финансовая математика. Финансовые рынки в дискретном случае», Г.И. Фалин «Математические основы страхования жизни и пенсионных схем» (третье издание), Г.И. Фалин, А.И. Фалин «Тригонометрия на вступительных экзаменах в МГУ».

В 2008 году были изданы две монографии: А.В. Булинский, А.П. Шашкин «Предельные теоремы для ассоциированных случайных полей и родственных систем», В.В. Сенатов «Центральная предельная теорема. Точность аппроксимации и асимптотические разложения», а также новый учебник В.Н. Тутубалина «Теория вероятностей» и переиздания учебников Ю.Н. Тюрина и др. «Теория вероятностей и статистика», «Анализ данных на компьютере».

А.Н. Ширяев состоял и состоит в редколлегиях научных журналов: «Теория вероятностей и ее применения» (заместитель главного редактора), «Успехи математических наук», «Фундаментальная и прикладная математика», «Вестник МГУ, Серия 1. Математика. Механика», «Markov Processes and Related Fields», «Analysis Mathematicae», «Stochastics and stochastic reports», «Sequential Analysis», «Probability and Mathematical Statistics», «Finance and Stochastics».

В состав редколлегий различных научных журналов входят также профессора Л.Г. Афанасьева, А.В. Булинский, Ю.Н. Тюрин.

Следует отметить участие сотрудников кафедры в международной научной жизни. Ее заведующий чл.-корр. РАН профессор А.Н. Ширяев был вице-президентом (1987–1989) и президентом (1989–1991) Общества Бернулли, президентом Российского общества актуариев (1994–1998), Финансового общества Базелье (1998–1999). Он является также действительным членом Академии Европы

(1990), почетным членом Королевского статистического общества Великобритании (1985), членом Международного статистического института, Института математической статистики (США), Московского математического общества, удостоен звания «Человек года» Американским биографическим институтом (1994), он лауреат Гумбольдтовской научной премии (1996). В 2001 году А.Н. Ширяеву вручен почетный диплом «*Doctor rerum naturalium honoris causa*» университета г. Фрейбурга (Германия).

Кафедра ведет активную работу со студентами, привлекая их к науке. На кафедре регулярно проводятся Колмогоровские студенческие олимпиады.

19 апреля 2008 года в ознаменование 105-летия со дня рождения А.Н. Колмогорова и 75-летия со дня публикации его монографии «Основные понятия теории вероятностей» была проведена Седьмая Колмогоровская студенческая олимпиада по теории вероятностей. В олимпиаде приняли участие 90 студентов механико-математического факультета, факультета ВМиК и факультета наук о материалах МГУ, а также Государственного университета «Высшая школа экономики», физико-технического факультета Киевского политехнического института, механико-математического факультета Киевского национального университета, математико-механического факультета Санкт-Петербургского государственного университета, факультета прикладной математики и компьютерных технологий Вологодского государственного педагогического университета.

За последние 5 лет активную научную работу вели студенты под руководством А.Н. Ширяева, А.В. Булинского, Е.В. Булинской, А.Г. Дьячкова, А.Д. Маниты, Д.Д. Соколова, Г.И. Фалина, Е.В. Чепурина, А.С. Черного, Е.Б. Яровой и др. Помимо ежегодных Колмогоровских олимпиад и Конференций молодых ученых мехмата МГУ студенты кафедры участвовали в Воронежской зимней математической школе имени С.Г. Крейна (2006), Российско-японском симпозиуме «Complex stochastic models: asymptotics and applications» (Москва, 2007), XV Встрече молодых европейских статистиков (Испания, Кастро-Урдиалис, 2007), XXVI Международном семинаре «Stability problems for stochastic models» (Израиль, Нахария, 2007), VI Колмогоровских чтениях (Ярославль, 2008) и др.

Студенты Г.И. Фалина в 2004–2006 годах участвовали в конкурсах студенческих научных работ, проводимых страховыми компаниями «Никойл» и «Ингосстрах», и получили там премии и поощрительные призы. Студенты А.Д. Маниты были лауреатами Федеральной стипендиальной программы В. Потанина в 2005 и 2006 годах.

Кафедра заботится о пополнении квалифицированных кадров. Так, после защиты кандидатских диссертаций были приняты на должности ассистентов Е.Е. Баштова (2007), Д.А. Шабанов (2008), Е.М. Суханова (2009). В 2009 году Д.А.Шабанову была выделена стипендия для молодых преподавателей и научных сотрудников, добившихся значительных результатов в педагогической и научно-исследовательской деятельности.

В 2008 году Л.Г. Афанасьевой присвоено звание заслуженного профессора МГУ.

С более подробной и текущей информацией о кафедре теории вероятностей можно ознакомиться на ее сайте (<http://www.math.msu.su/probab>).

О НЕСТАНДАРТНЫХ ПРОБЛЕМАХ СТОХАСТИЧЕСКОЙ ОПТИМИЗАЦИИ: РЕДУКЦИЯ К ЗАДАЧАМ В МАРКОВСКОМ ПРЕДСТАВЛЕНИИ И ИХ РЕШЕНИЕ

А. Н. Ширяев

В статье рассматриваются стохастические задачи, в которых оптимизация состоит в выборе рационального момента («момента остановки») для принятия того или иного управляющего воздействия.

С точки зрения общей «теории оптимальных правил остановки» [1, 2] приводимые задачи имеют нестандартную формулировку. Показывается, как эти задачи редуцируются к стандартным задачам в «марковском представлении», для решения которых затем используются результаты и методы, развитые в теории оптимальных правил остановки для марковских процессов.

1. Введение

Везде далее $W = (W_t)_{t \geq 0}$ обозначает стандартный винеровский процесс (броуновское движение). Через $\tau = \tau(\omega)$ будут обозначаться марковские моменты («моменты остановки») относительно потока σ -алгебр $(\mathcal{F}_t^W)_{t \geq 0}$, где $\mathcal{F}_t^W = \sigma(W_s, s \leq t)$. (Все вероятностные объекты предполагаются определенными на фильтрованном вероятностном пространстве $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$, являющемся стохастическим базисом, то есть таким, что $\mathcal{F}$ пополнено множествами P -нулевой меры, $\mathcal{F}_t$ пополнены множествами P -нулевой меры из $\mathcal{F}$, где $\mathcal{F}_t = \mathcal{F}_{t+} (= \bigcap_{s>t} \mathcal{F}_s)$.)

Дадим формулировку рассматриваемых задач, в которых «стохастика» будет моделироваться броуновским движением.

Задача I (скорейшее обнаружение момента *появления арбитража*).

Пусть $\theta = \theta(\omega)$ — случайная величина со значениями в $[0, \infty)$, не зависящая от процесса $W = (W_t)_{t \geq 0}$. Предположим, что наблюдению подлежит случайный процесс $X = (X_t)_{t \geq 0}$, такой, что

$$X_t = \mu(t - \theta)^+ + \sigma W_t, \quad (1)$$

где μ и σ — известные константы, $\mu \in \mathbb{R} \setminus \{0\}$, $\sigma > 0$.

Из (1) следует, что до момента θ процесс X совпадает с броуновским движением W , и, следовательно, на интервале $[0, \theta)$ мы имеем безарбитражную ситуацию, поскольку броуновское движение является мартингалом. После момента θ у наблюдаемого процесса появляется снос $\mu \neq 0$, и, следовательно, наблюдаемый процесс X будет субмартингалом, если $\mu > 0$, и супермартингалом, если $\mu < 0$. В каждом из этих случаев имеет место арбитражная ситуация [3], и с точки зрения оперирования такими объектами, имеющими, скажем, смысл логарифмов цен некоторых финансовых инструментов, значительный интерес представляет задача скорейшего обнаружения момента появления θ .

Одна из постановок (*вариант А*) такой задачи может состоять в следующем.

Пусть $\mathfrak{M}$ — класс конечных моментов остановки τ и $\mathfrak{M}_{(\alpha)} = \{\tau \in \mathfrak{M} : P(\tau < \theta) \leq \alpha\}$ — класс моментов остановки, для которых вероятность ложной тревоги

$P(\tau < \theta)$ меньше или равна некоторой заданной константе $0 < \alpha < 1$. Требуется найти оптимальный момент $\tau_{(\alpha)}^*$, для которого

$$E(\tau_{(\alpha)}^* - \theta)^+ = \inf_{\tau \in \mathfrak{M}_{(\alpha)}} E(\tau - \theta)^+. \quad (2)$$

Иначе говоря, требуется найти в классе $\mathfrak{M}_{(\alpha)}$ момент, на котором достигается минимум среднего времени обнаружения момента появления арбитража.

Другая интересная постановка (*вариант В*) задачи скорейшего обнаружения состоит в следующем.

Пусть $\mathfrak{M}_{(T)} = \{\tau : E_\infty \tau \geq T\}$ — класс тех моментов остановки, для которых среднее время $E_\infty \tau$ до объявления ложной тревоги (то есть когда на самом деле $\theta = \infty$) больше или равно заданной константе $T > 0$.

Задача состоит в отыскании такого момента τ_T^* , для которого

$$\int_0^\infty E_\theta(\tau_T^* - \theta)^+ d\theta = \inf_{\tau \in \mathfrak{M}_{(T)}} \int_0^\infty E_\theta(\tau - \theta)^+ d\theta, \quad (3)$$

где $E_\theta(\cdot)$ — усреднение по мере P_θ , являющейся распределением процесса $X = (X_t)_{t \geq 0}$, описываемого соотношением (1) с заданным значением θ .

Задача II (обнаружение момента изменения *тенденции в движении* броуновского движения).

Пусть $W = (W_t)_{0 \leq t \leq T}$ — броуновское движение, рассматриваемое на временном интервале $[0, T]$. Обозначим θ тот момент на $[0, T]$, для которого $W_\theta = \max_{t \in [0, T]} W_t$ (такой момент с вероятностью единица определяется однозначно).

Представляет интерес отыскание того момента τ_T^* , для которого

$$E[W_{\tau_T^*} - W_\theta]^2 = \inf_{\tau \in \mathfrak{M}_T} E[W_\tau - W_\theta]^2, \quad (4)$$

где $\mathfrak{M}_T = \{\tau : \tau \leq T\}$ — класс моментов остановки со значениями в $[0, T]$. С точки зрения технического анализа из финансовой инженерии момент τ_T^* может рассматриваться как наилучшее (в среднеквадратическом смысле) приближение значения $W_{\tau_T^*}$ к максимальному значению W_θ .

Другая интересная задача состоит в отыскании момента τ_T^0 , такого, что

$$E|\tau_T^0 - \theta| = \inf_{\tau \in \mathfrak{M}_T} E|\tau - \theta|. \quad (5)$$

Как будет показано ниже, оптимальные моменты τ_T^0 и τ_T^* совпадают.

Задача III (условно-экстремальные задачи для *непредсказуемых* моментов θ и g в броуновском движении).

Как и в задаче II, будем обозначать через θ тот момент, для которого $W_\theta = \max_{t \in [0, T]} W_t$. Пусть также g — момент последнего (на $[0, T]$) обращения броуновского движения в нуль. Аналогичным образом обозначим через θ^μ и g^μ соответствующие моменты для броуновского движения со сносом $W^\mu = (W_t^\mu)_{0 \leq t \leq T}$, где $W_t^\mu = \mu t + W_t$.

Интересующие нас задачи (условно-экстремальные) состоят в отыскании моментов $\tau_T^\mu(\alpha)$ и $\sigma_T^\mu(\alpha)$, для которых

$$\inf_{\tau \in \mathfrak{M}_T^\mu(\alpha, \theta)} E(\tau - \theta)^+ = E(\tau_T^\mu(\alpha) - \theta)^+ \quad (6)$$

и

$$\inf_{\sigma \in \mathfrak{M}_T^\mu(\alpha, g)} \mathbb{E}(\sigma - g)^+ = \mathbb{E}(\sigma_T^\mu(\alpha) - g)^+, \quad (7)$$

где

$$\mathfrak{M}_T^\mu(\alpha, \theta) = \{ \tau \leq T : \mathbb{P}(\tau < \theta) \leq \alpha \}$$

и

$$\mathfrak{M}_T^\mu(\alpha, g) = \{ \sigma \leq T : \mathbb{P}(\sigma < g) \leq \alpha \}.$$

Здесь $0 < \alpha < 1$, а $\mathbb{P}(\tau < \theta)$ и $\mathbb{P}(\sigma < g)$ являются вероятностями ложных тревог относительно моментов θ и g .

Задача IV (оптимальность *трейдингового правила* «Buy and Hold»).

Рассматривается (B, S) -рынок Блэка и Шоулса, состоящий из банковского счёта B и акции S , динамика которых $B = (B_t)_{t \geq 0}$ и $S = (S_t)_{t \geq 0}$ описывается соотношениями

$$\begin{aligned} dB_t &= rB_t dt, \quad B_0 = 1, \\ dS_t &= S_t \cdot (\mu dt + \sigma dW_t), \quad S_0 = 1. \end{aligned} \quad (8)$$

Пусть $U = U(x)$, $x \geq 0$, есть функция полезности (скажем, $U(x) = x$ или $U(x) = \log x$). Обозначим через $P_t = \frac{S_t}{B_t}$ относительные цены, $M_T = \max_{t \in [0, T]} P_t$. Будем отождествлять моменты τ из $\mathfrak{M}_T$ как моменты продажи акции на интервале $[0, T]$. Интерес представляет отыскание момента τ_T^* , для которого

$$\mathbb{E}U\left(\frac{P_{\tau_T^*}}{M_T}\right) = \sup_{\tau \in \mathfrak{M}_T} \mathbb{E}U\left(\frac{P_\tau}{M_T}\right). \quad (9)$$

Ниже будет установлено, что в случае логарифмической и линейной функций полезности оптимальный момент τ_T^* является *вырожденным*:

$$\tau_T^* = \begin{cases} 0, & \text{если } \mu - r - \frac{1}{2}\sigma^2 \leq 0; \\ T, & \text{если } \mu - r - \frac{1}{2}\sigma^2 > 0. \end{cases} \quad (10)$$

Определенное этим моментом правило продажи акции носит название правила «Buy and Hold».

Задача V (оптимальность *стохастического трейдингового правила* «Buy and Hold»).

Предположим, что в (B, S) -модели процесс $S = (S_t)_{t \geq 0}$ подчиняется уравнению

$$dS_t = S_t \cdot (\mu(t, \theta)dt + \sigma dW_t), \quad (11)$$

где θ является случайной величиной со значениями в $[0, \infty)$, не зависящей от броуновского движения W . Пусть

$$\mu(t, \theta) = \begin{cases} \mu_1, & t < \theta, \\ \mu_2, & t \geq \theta, \end{cases} \quad (12)$$

где $\mu_1 > \mu_2$ и к тому же

$$\nu_1 \equiv \mu_1 - r - \sigma^2/2 > 0; \quad (13)$$

$$\nu_2 \equiv \mu_2 - r - \sigma^2/2 < 0. \quad (14)$$

Иначе говоря, пусть

$$\mu_2 - \sigma^2/2 < r < \mu_1 - \sigma^2/2. \quad (15)$$

Таким образом, начиная с момента $t = 0$ «идет» параметр μ_1 . Если интересоваться моментом продажи акции, то правило «Buy and Hold» предыдущей задачи говорит о том, что акцию надо было бы держать до момента времени T . Но если момент θ наступает до этого момента T , то же самое правило рекомендует акцию продать в момент θ . Поскольку этот момент явно ненаблюдаем, то естественным образом возникает задача о том, как по наблюдениям за процессом S «оптимальным образом» оценить θ и затем акцию продать. Как будет показано далее, если θ имеет экспоненциальное распределение и функция полезности является логарифмической, то оптимальный момент, когда надо продавать акцию, имеет вид

$$\tau^* = T \wedge \inf\{0 \leq t \leq T : \pi_t \geq g_T(t)\},$$

где $\pi_t = P(\theta \leq t \mid \mathcal{F}_t^s)$ и $g_T(t)$ — некоторая граница, по достижении которой процессом π_t принимается решение о продаже акции (если $\pi_t < g_T(t)$ при всех $t \leq T$, то акция продается в момент времени T).

Решение задачи V дает также решение следующей задачи о максимизации капитала к моменту времени T .

Рассмотрим (B, S) -модель, изложенную в формулировке предыдущей задачи. Пусть сначала «идет» параметр μ_1 , который в момент θ меняется на параметр $\mu_2 < \mu_1$. Предположим, что в момент $\tau \leq T$ совершается ребалансирование портфеля, состоящее в том, что средства S_τ кладутся на банковский счет. Тогда в момент T капитал будет равен $S_\tau \frac{B_T}{B_\tau}$. Поскольку

$$S_\tau \frac{B_T}{B_\tau} = \frac{S_\tau}{B_\tau} B_T = P_\tau B_T,$$

то видим, что

$$\sup_{\tau \in \mathcal{M}_T} E \log \left(S_\tau \frac{B_T}{B_\tau} \right) = \sup_{\tau \in \mathcal{M}_T} E \log P_\tau + E \log B_T. \quad (16)$$

Таким образом, задача отыскания оптимального момента ребалансирования портфеля ценных бумаг равносильна рассмотренной выше задаче (9) для логарифмической функции полезности.

Основная мотивация сформулированных задач состоит в желании получить математическое объяснение некоторых методов технического анализа (типа «когда покупать» — «когда продавать»), широко известных на практике и носящих, как правило, лишь описательный характер.

Известно (см., например, [3, гл. I, § 2с]), что в финансовой экономике принято участников и исследователей рынка ценных бумаг относить к одной из трех групп — фундаменталисты, количественные аналитики, техники (fundamentalists, quants, technicians).

Фундаменталисты исходят в своих решениях из поведения экономики *в целом*, для них важны представления о перспективах развития. Представители технического анализа руководствуются в своих действиях *локальным* поведением рынка.

Количественные аналитики, являющиеся последователями Л. Башелье, опираются в своих рассуждениях на самый разнообразный аппарат теории вероятностей, математической статистики, теории случайных процессов, теории стохастической оптимизации.

Рассматриваемые нами задачи обязаны прежде всего финансовому техническому анализу. Как мы увидим далее, методы, применяемые для их решения, в значительной степени используют аппарат количественной теории финансов, которую обычно принято называть финансовой математикой.

В заголовке статьи присутствуют слова о *нестандартных* задачах.

В следующем разделе мы сформулируем те *стандартные* задачи, для которых существует широко развитая теория их решения. Затем будет показано, как задачи I–V можно редуцировать к рассмотрению соответствующих стандартных (марковских) задач.

2. О стандартных задачах теории оптимальных правил остановки

1. Под *стандартными* принято понимать задачи следующего типа.

Пусть $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ — фильтрованное вероятностное пространство, $G = (G_t)_{t \geq 0}$ — случайный процесс, *согласованный* с потоком $(\mathcal{F}_t)_{t \geq 0}$, то есть такой, что при каждом $t \geq 0$ величины G_t являются $\mathcal{F}_t$ -измеримыми. Обозначим

$$V = \sup_{\tau \in \mathfrak{M}} \mathbb{E} G_\tau, \quad (17)$$

где, как и ранее, $\mathfrak{M}$ — класс моментов остановки (относительно потока $(\mathcal{F}_t)_{t \geq 0}$) со значениями в $[0, \infty)$.

Важным обстоятельством здесь является то, что величины G_t являются $\mathcal{F}_t$ -измеримыми, $t \geq 0$. Именно это свойство характеризует *стандартность* формулировки задачи об оптимальной остановке (17). Нестандартными называются задачи, в которых G_t не являются, вообще говоря, $\mathcal{F}_t$ -измеримыми.

По поводу общей теории решения стандартных задач, в основе которой лежит понятие *огИБающей Снелла*, см. [4, гл. 1, § 2.1].

Все рассмотренные в разделе 1 задачи относятся к *нестандартным*. Однако оказывается, что их можно свести к стандартным задачам, в которых для функций $G_t(\omega)$ имеет место *марковское* представление:

$$G_t(\omega) = G(t, X_t(\omega)) + \int_0^t F(s, X_s(\omega)) ds, \quad (18)$$

где $X = (X_t)_{t \geq 0}$ — некоторый *марковский* процесс, что дает возможность для решения задач об оптимальной остановке (17) воспользоваться широко известной и глубоко продвинутой теорией марковских процессов.

Приведем некоторые результаты из общей теории оптимальных правил остановки для марковского случая (18), основываясь на изложении в монографиях [1] и [4].

Будем предполагать, что $X = (X_t, \mathcal{F}_t, \mathbb{P}_x)_{t \geq 0}$ есть однородный марковский процесс в фазовом пространстве E , которое будет (для простоты) пространством $\mathbb{R}$; $\mathbb{P}_x$ — распределение вероятностей $(X_t)_{t \geq 0}$ в предположении, что $\mathbb{P}_x(X_0 = x) = 1$.

Пусть «цена»

$$V(x) = \sup_{\tau \in \mathfrak{M}} \mathbb{E}_x G(X_\tau), \quad x \in E. \quad (19)$$

Введем два множества

$$C = \{x \in E : V(x) > G(x)\} \quad (20)$$

и

$$D = \{x \in E : V(x) = G(x)\}. \quad (21)$$

Эти множества называются множествами *продолжения* и *остановки* наблюдений. Объясняется это тем, что при достаточно общих условиях момент

$$\tau_D = \inf\{t \geq 0 : X_t \in D\}, \quad (22)$$

то есть момент первого попадания в множество D , является оптимальным моментом.

Точные утверждения содержатся в приводимых ниже двух теоремах, для формулировки которых введем следующие определения.

Определение 1. Измеримая функция $F: E \rightarrow \mathbb{R}$ называется *супергармонической* (для марковского процесса X), если для любого момента остановки σ и всех $x \in E$

$$\mathbb{E}_x F(X_\sigma) \leq F(x).$$

(Предполагается, что $F(X_\sigma) \in L^1(\mathbb{P}_x)$ для всех $x \in E$. Для того чтобы функция F была супергармонической, необходимо и достаточно, чтобы процесс $(F(X_t))_{t \geq 0}$ являлся непрерывным справа супермартингалом относительно $\mathbb{P}_x$ для каждого $x \in E$, такого, что $\mathbb{E}_x |F(X_t)| < \infty$, $t \geq 0$, $\mathbb{E}_x(F(X_t) | \mathcal{F}_s) \leq F(X_s)$ ($\mathbb{P}_x$ -п. н.), $s \leq t$.)

Определение 2. Функция $F: E \rightarrow \mathbb{R}$ называется *полу непрерывной снизу* (сверху) в точке x_0 , если

$$\lim_{x \rightarrow x_0} F(x) = F(x_0) \quad \left(\overline{\lim}_{x \rightarrow x_0} F(x) = F(x_0) \right).$$

Предложение 1 (необходимость). Предположим, что в задаче (19) существует оптимальный момент $\tau_* \in \mathfrak{M}$, то есть пусть

$$V(x) = \mathbb{E}_x G(X_{\tau_*}), \quad x \in E.$$

Тогда функция $V = V(x)$ является наименьшей супергармонической функцией, мажорирующей функцию $G = G(x)$.

Пусть также функция G является полу непрерывной сверху.

Тогда

(а) момент $\tau_D \leq \tau_*$ ($\mathbb{P}_x$ -п. н., $x \in E$) и является оптимальным;

(б) остановленный процесс $(V(X_{t \wedge \tau_D}))_{t \geq 0}$ является непрерывным справа мартингалом по любой из мер $\mathbb{P}_x$, $x \in E$.

Следующее предложение дает достаточные условия для существования оптимального момента остановки в задаче (19), называемой задачей в марковском представлении.

Предложение 2 (достаточность). Пусть в задаче (19) выполнено условие

$$\mathbb{E}_x \sup_{t \geq 0} |G(X_t)| < \infty, \quad x \in E. \quad (23)$$

Предположим, что существует наименьшая супергармоническая мажоранта $\hat{V} = \hat{V}(x)$ функции $G = G(x)$. Предположим также, что функция $G = G(x)$ является полунепрерывной сверху, а функция $\hat{V} = \hat{V}(x)$ — полунепрерывной снизу.

Пусть $\hat{D} = \{x \in E : \hat{V}(x) = G(x)\}$ и $\tau_{\hat{D}} = \inf\{t \geq 0 : X_t \in \hat{D}\}$.

Тогда

(а) если $\mathbb{P}_x(\tau_{\hat{D}} < \infty) = 1$ для всех $x \in E$, то $\hat{V} = V$ и момент $\tau_{\hat{D}}$ является оптимальным в задаче (19);

(б) если для некоторого $x \in E$ вероятность $\mathbb{P}_x(\tau_{\hat{D}} < \infty) < 1$, то оптимального момента в задаче (19) не существует.

Следствие 1. Пусть функция $G = G(x)$ полунепрерывна сверху и функция $V = V(x)$ ($= \sup_{\tau \in \mathfrak{M}} \mathbb{E}_x G(X_\tau)$) является полунепрерывной снизу.

Если $\mathbb{P}_x(\tau_D < \infty) = 1$, $x \in E$, то момент τ_D является в задаче (19) оптимальным.

Сформулированные утверждения, доказательство которых содержится в монографиях автора [1] и [4], являются ключевыми результатами теории оптимальных правил остановки для марковских процессов, и именно на них основываются решения конкретных задач в марковском представлении.

2. Для отыскания наименьших супергармонических мажорант $\hat{V} = \hat{V}(x)$ можно указать следующие два метода — итеративный и основанный на решении задачи Стефана [1, 4].

Итеративный метод состоит в следующем. Пусть функция $G = G(x)$ является полунепрерывной снизу и $\mathbb{E}_x \inf_{t \geq 0} G(X_t) > -\infty$, $x \in E$. Тогда

$$\hat{V}(x) = \lim_n \lim_N Q_n^N G(x), \quad (24)$$

где

$$Q_n G(x) = \max\{G(x), \mathbb{E}_x G(X_{2^{-n}})\} \quad (25)$$

и Q_n^N — N -я степень оператора Q_n . При этом функция $\hat{V}(x)$ совпадает с ценой $V(x) = \sup_{\tau \in \mathfrak{M}} \mathbb{E}_x G(X_\tau)$.

Метод, основанный на решении задачи Стефана, называемой также задачей со свободными границами, состоит в том, чтобы отыскать такую минимальную функцию $\hat{V} = \hat{V}(x)$ и области $\hat{C} = \{x : \hat{V}(x) > G(x)\}$ и $\hat{D} = \{x : \hat{V}(x) = G(x)\}$, для которых при всех $x \in E$ выполнялось бы условие (супергармоничности функции $\hat{V}$)

$$L_X \hat{V} \leq 0, \quad (26)$$

где L_X — характеристический (инфинитезимальный) оператор процесса X .

Предполагая априори, что функция $G = G(x)$ является достаточно гладкой в окрестности границы $\partial \hat{C}$, этот метод можно переформулировать как метод отыскания таких функций $\hat{V}$ и областей $\hat{C}$ и $\hat{D}$, что

$$L_x \hat{V} = 0 \text{ на } \hat{C}, \quad (27)$$

$$\widehat{V} = G \text{ на } \widehat{D} \quad (28)$$

и

$$\frac{\partial \widehat{V}}{\partial x} = \frac{\partial G}{\partial x} \text{ на } \partial \widehat{C}. \quad (29)$$

Последнее условие (29) совпадения производных на границе $\partial \widehat{C}$ области $\widehat{C}$ принято называть условием *гладкого склеивания*. В тех случаях, когда не приходится рассчитывать на выполнение этого условия, часто бывает выполнено условие *непрерывного склеивания*

$$\widehat{V} = G \text{ на } \partial \widehat{C}. \quad (30)$$

Если найдено некоторое решение $(\widehat{V}, \widehat{C}, \widehat{D})$ задачи Стефана (27)–(29) или (27)–(28), то это еще не означает, что функция $\widehat{V} = \widehat{V}(x)$ совпадает с искомой «ценой» $V = V(x)$ и момент $\widehat{\tau}$ первого попадания в множество $\widehat{D}$ является оптимальным.

Стандартный метод доказательства равенства $\widehat{V}(x) = V(x)$ и оптимальности момента $\widehat{\tau} = \inf\{t \geq 0 : X_t \in \widehat{D}\}$ основан на следующем предложении

Предложение 3 («проверочная теорема»).

Пусть выполнены следующие («проверочные») условия:

- (1) $\widehat{V}(x) \geq G(x)$, $x \in E$;
- (2) момент $\widehat{\tau} \in \mathfrak{M}$ и $\widehat{V}(x) = \mathbb{E}_x G(X_{\widehat{\tau}})$, $x \in E$;
- (3) для любого момента $\tau \in \mathfrak{M}$ $\widehat{V}(x) \geq \mathbb{E}_x \widehat{V}(X_{\tau})$ (супермартингалность $\widehat{V} = \widehat{V}(x)$).

Тогда $\widehat{V}(x)$ совпадает с ценой $V(x) = \sup_{\tau \in \mathfrak{M}} \mathbb{E}_x G(X_{\tau})$ и момент $\widehat{\tau}$ является оптимальным в классе $\mathfrak{M}$.

Действительно, если эти условия выполнены, то из (1) и (3) получаем

$$\widehat{V}(x) \geq \mathbb{E}_x G(X_{\tau}), \quad x \in E$$

для любого $\tau \in \mathfrak{M}$ и, значит,

$$\widehat{V}(x) \geq \sup_{\tau \in \mathfrak{M}} \mathbb{E}_x G(X_{\tau}) = V(x),$$

что вместе с условием (2) показывает, что на самом деле $\widehat{V}(x) = V(x)$, и, следовательно, момент $\widehat{\tau}$ является в классе $\mathfrak{M}$ оптимальным.

В монографии [4] приведено большое количество примеров, показывающих, как находятся $\widehat{V}$, $\widehat{C}$ и $\widehat{D}$ при дополнительных условиях (29) или (30).

Во многих интересных задачах об оптимальной остановке рассматриваемые моменты остановки являются ограниченными. Скажем, пусть $\mathfrak{M}_T = \{\tau : \tau \leq T\}$ есть класс моментов остановки τ , принимающих значения в множестве $[0, T]$. В этом случае вместо задачи (19), в которой фазовая переменная $x \in E$ и играет роль начального состояния процесса $(X_t)_{t \geq 0}$, естественно (следуя идеям обратной индукции в случае дискретного времени [4, гл. 1, § 1.2]) рассматривать серию «вложенных» задач отыскания функций

$$V(t, x) = \sup_{0 \leq \tau \leq T-t} \mathbb{E}_x G(t + \tau, X_{t+\tau}), \quad (31)$$

где $x \in E$, $t \in [0, T]$ и $G = G(t, x)$ — функция выигрыша при остановке в момент t в состоянии $x \in E$.

С принципиальной точки зрения *неоднородная* задача (31) может быть сведена к однородной задаче (19) введением двумерного марковского процесса $X'_t = (t, X_t)$ с фазовым пространством состояний $[0, T] \times E$. Понятно при этом, что ограниченность временного параметра ($t \in [0, T]$) автоматически приводит к выполнению краевого условия

$$V(T, x) = G(T, x). \quad (32)$$

Рассмотренные выше результаты легко переносятся (см. [4]) и на несколько более широкий класс стандартных задач, в которых вместо критерия (19) рассматривается критерий

$$V(x) = \sup_{\tau \in \mathfrak{M}} \mathbb{E}_x \left[G(X_\tau) + \int_0^\tau F(X_s) ds \right] \quad (19')$$

с некоторыми функциями $F = F(x)$, $x \in E$, для которых существуют математические ожидания $\mathbb{E}_x \int_0^\tau |F(X_s)| ds$ для всех $\tau \in \mathfrak{M}$ и $x \in E$.

Точно так же вместо критерия (31) можно рассматривать критерий

$$V(t, x) = \sup_{0 \leq \tau \leq T-t} \mathbb{E}_x \left[G(t + \tau, X_{t+\tau}) + \int_0^\tau F(t + s, X_{t+s}) ds \right]. \quad (31')$$

В следующих разделах будут описаны основные идеи и этапы решения сформулированных задач I–VI. Рассматриваемые задачи таковы, что их формулировки носят *нестандартный* характер в том смысле, что изначально они не имеют вид (19), (19') или (31), (31').

Тем не менее оказывается, что все эти задачи могут быть переформулированы в *стандартной* форме для подходящим образом построенных марковских процессов и функционалов, входящих в (19), (19') или (31), (31'). Это обстоятельство является определяющим для возможности применения результатов общей теории оптимальных правил остановки к решению рассматриваемых нестандартных задач.

3. Задача I

1. В *варианте A* эта задача состоит, как описано в разделе 1, в отыскании в классе $\mathfrak{M}_{(\alpha)} = \{ \tau \in \mathfrak{M} : \mathbb{P}(\tau < \theta) \leq \alpha \}$ такого момента $\tau_{(\alpha)}^*$, для которого выполнено условие (2).

Будем предполагать, что случайная величина $\theta = \theta(\omega)$ имеет экспоненциальное распределение:

$$\begin{aligned} \mathbb{P}(\theta = 0) &= \pi, \quad 0 \leq \pi < 1, \\ \mathbb{P}(\theta > t | \theta > 0) &= e^{-\lambda t}, \quad t > 0, \end{aligned} \quad (33)$$

где λ — известный параметр ($\lambda > 0$).

Для решения поставленной *условно-вариационной* задачи отыскания момента $\tau_{(\alpha)}^*$ рассмотрим, следуя методу множителей Лагранжа, задачу отыскания оптимального момента τ^* для критерия

$$V(\pi) = \inf_{\tau \in \mathfrak{M}} [\mathbb{P}_\pi(\tau < \theta) + c \mathbb{E}_\pi(\tau - \theta)^+], \quad (34)$$

где индекс π у P_π означает, что рассматривается распределение вероятностей в предположении, что $P(\theta = 0) = \pi$.

Введем процесс апостериорных вероятностей $(\pi_t)_{t \geq 0}$ с $\pi_0 = \pi$ и для $t > 0$

$$\pi_t = P(\theta \leq t | \mathcal{F}_t^X). \quad (35)$$

Также пусть

$$L_t = \frac{d(P^0 | \mathcal{F}_t^X)}{d(P^\infty | \mathcal{F}_t^X)},$$

где P^0 и P^∞ — распределения вероятностей процесса $(X_t)_{t \geq 0}$, определяемого формулой (1) при $\theta = 0$ и $\theta = \infty$ соответственно, и $P^0 | \mathcal{F}_t^X$ и $P^\infty | \mathcal{F}_t^X$ — сужения этих мер P^0 и P^∞ на σ -алгебру $\mathcal{F}_t^X = \sigma(X_s, s \leq t)$.

Хорошо известно [5], что

$$L_t = e^{\frac{\mu}{\sigma^2} X_t - \frac{1}{2} \frac{\mu^2}{\sigma^2} t}. \quad (36)$$

По формуле Ито [4, 5] отсюда заключаем, что

$$dL_t = \frac{\mu}{\sigma^2} L_t dX_t, \quad L_0 = 1. \quad (37)$$

Полагая

$$\varphi_t = \frac{\pi_t}{1 - \pi_t},$$

по формуле Байеса [5] находим, что

$$\varphi_t = \varphi_0 e^{\lambda t} L_t + \lambda \int_0^t \frac{e^{\lambda u} L_u}{e^{\lambda u} L_u} du. \quad (38)$$

Снова применяя формулу Ито, получаем, что

$$d\varphi_t = \lambda(1 + \varphi_t)dt + \frac{\mu}{\sigma^2} \varphi_t dX_t. \quad (39)$$

Следовательно, поскольку

$$\pi_t = \frac{\varphi_t}{1 + \varphi_t},$$

то из (39) и формулы Ито имеем

$$d\pi_t = (\lambda - \frac{\mu^2}{\sigma^2} \pi_t^2)(1 - \pi_t)dt + \frac{\mu}{\sigma^2} \pi_t(1 - \pi_t)dX_t \quad (40)$$

с $\pi_0 = \pi$.

Процесс

$$\overline{W}_t = \frac{1}{\sigma} \left(X_t - \mu \int_0^t \pi_s ds \right) \quad (41)$$

является [1] броуновским движением («обновляющим» процессом с $\mathcal{F}_t^{\overline{W}} = \mathcal{F}_t^X$, $t \geq 0$).

Из (40) и (41) находим, что

$$d\pi_t = \lambda(1 - \pi_t)dt + \frac{\mu}{\sigma} \pi_t(1 - \pi_t)d\overline{W}_t, \quad \pi_0 = \pi. \quad (42)$$

Поскольку величины $\overline{W}_t$ являются $\mathcal{F}_t^X$ -измеримыми, то из (42) заключаем, что процесс $(\pi_t)_{t \geq 0}$ является *диффузионным марковским процессом* с характеристическим (инфинитезимальным) оператором

$$L_\pi = \lambda(1 - \pi) \frac{\partial}{\partial \pi} + \frac{1}{2} \left(\frac{\mu}{\sigma} \right)^2 \pi^2 (1 - \pi)^2 \frac{\partial^2}{\partial \pi^2}. \quad (43)$$

Для каждого марковского момента τ с $E_\pi \tau < \infty$ [1, 6]

$$P_\pi(\tau < \theta) + c E_\pi(\tau - \theta)^+ = E_\pi \left[(1 - \pi_t) + c \int_0^\tau \pi_s ds \right]$$

и, следовательно, критерий (34) принимает *стандартный* вид типа (31'):

$$V(\pi) = \inf_{\tau \in \mathfrak{M}} E_\pi \left[(1 - \pi_\tau) + c \int_0^\tau \pi_s ds \right], \quad (44)$$

где $(\pi_t)_{t \geq 0}$ есть диффузионный марковский процесс (относительно потока $(\mathcal{F}_t^X)_{t \geq 0}$ и распределений вероятностей P_π , $\pi \in [0, 1)$).

Полагая $G(\pi) = 1 - \pi$, находим, что, согласно общей теории оптимальных правил остановки для марковских процессов, область продолжения наблюдений C должна иметь вид

$$C = \{ \pi \in [0, 1] : V(\pi) > 1 - \pi \},$$

а область остановки — вид

$$D = \{ \pi \in [0, 1] : V(\pi) = 1 - \pi \},$$

Из выпуклости (вверх) функции $V = V(\pi)$ вытекает, что

$$C = \{ \pi \in [0, 1] : \pi < A \}$$

и

$$D = \{ \pi \in [0, 1] : \pi \geq A \},$$

где A — некоторое подлежащее определению «пороговое» значение, с помощью которого оптимальный момент остановки τ^* в задачах (44) и (34) определяется формулой

$$\tau^* = \inf \{ t : \pi_t \geq A \}. \quad (45)$$

Для отыскания как функции $V(\pi)$, так и «порогового» значения A рассмотрим (вспомогательную) задачу Стефана: найти функцию $\hat{V} = \hat{V}(\pi)$ и константу $\hat{A} \in [0, 1]$, такие, что

$$L_\pi \hat{V}(\pi) = -\pi, \quad \pi \in (0, \hat{A}), \quad (46)$$

$$\hat{V}(\pi)|_{\pi \uparrow \hat{A}} = 1 - \hat{A}, \quad (47)$$

$$\frac{\partial \hat{V}}{\partial \pi} \Big|_{\pi \uparrow \hat{A}} = -1. \quad (48)$$

Общее решение $\widehat{V} = \widehat{V}(\pi)$ уравнения (46) содержит две неопределенные константы. Вместе с неопределенной (пока) константой $\widehat{A}$ нужно иметь *три* дополнительных условия, которые бы дали возможность определить эти константы. Двумя условиями являются условия непрерывного и гладкого склеивания (47) и (48). Третьим условием является условие

$$\left. \frac{\partial \widehat{V}}{\partial \pi} \right|_{\pi \downarrow 0} = 0, \quad (49)$$

выделяющее среди найденных решений то, которое (по смыслу задачи) не уходит в $+\infty$ или $-\infty$ при $\pi \downarrow 0$. (Подробнее по поводу *необходимости* выполнения этого условия см. [1].)

В результате решения задачи Стефана (46)–(49) находим, что «пороговое» значение $\widehat{A}$ и функция $\widehat{V} = \widehat{V}(\pi)$ определяются следующим образом:

$\widehat{A}$ является единственным корнем уравнения

$$\Psi(\widehat{A}) = 1 \quad (50)$$

и

$$\widehat{V}(\pi) = \begin{cases} 1 - \widehat{A} + \int_{\pi}^{\widehat{A}} \Psi(x) dx, & \pi \in [0, \widehat{A}]; \\ 1 - \pi, & \pi \in [\widehat{A}, 1], \end{cases}$$

где

$$\Psi(x) = \frac{\lambda}{\rho} \int_0^x e^{-\frac{\lambda}{\rho}[H(x)-H(y)]} \frac{dy}{y(1-y)^2}$$

с $\rho = \frac{\mu^2}{2\sigma^2}$ и $H(x) = \log \frac{x}{1-x} - \frac{1}{x}$. Основываясь на «проверочной теореме» (предложение 3), устанавливаем, что найденное решение $(\widehat{V}, \widehat{A})$ задачи Стефана (46)–(48) таково, что $\widehat{V} = V$ и момент

$$\tau^* = \inf \{ t \geq 0 : \pi_t \geq A \}$$

с $A = \widehat{A}$ является в задаче (44) (*вариант A*) оптимальным. (Детали, связанные с анализом «проверочных» условий предложения 3, см. в [1].)

Найденное решение для задач (44), (34) позволяет получить и решение условно-вариационной задачи в *варианте A*. Действительно, найденный порог $\widehat{A} = \widehat{A}(c)$ непрерывным образом зависит от c и можно найти такое $c = c(\alpha)$, что $\widehat{A}(c(\alpha)) = 1 - \alpha$. Тогда для всех $\pi \leq 1 - \alpha$ для момента $\tau_{(\alpha)}^* = \inf \{ t \geq 0 : \pi_t \geq 1 - \alpha \}$ находим, что

$$\mathbb{E}_{\pi}(1 - \pi_{\tau_{(\alpha)}^*}) = 1 - (1 - \alpha) = \alpha.$$

Согласно найденному решению задач (44), (34) с $c = c(\alpha)$, для любого $\tau \in \mathfrak{M}$

$$\mathbb{P}_{\pi}(\tau < \theta) + c(\alpha)\mathbb{E}_{\pi}(\tau - \theta)^+ \geq \mathbb{P}_{\pi}(\tau_{(\alpha)}^* < \theta) + c(\alpha)\mathbb{E}_{\pi}(\tau_{(\alpha)}^* - \theta)^+.$$

Отсюда следует, что если $\tau \in \mathfrak{M}_{(\alpha)}$, то, поскольку $\mathbb{P}_{\pi}(\tau_{(\alpha)}^* < \theta) = \alpha$, имеем

$$c(\alpha)\mathbb{E}_{\pi}(\tau - \theta)^+ \geq c(\alpha)\mathbb{E}_{\pi}(\tau_{(\alpha)}^* - \theta)^+$$

(для всех $\pi \leq 1 - \alpha$). При $0 \leq \alpha < 1$ константы $c(\alpha) \neq 0$.

Тем самым доказываем, что в классе $\mathfrak{M}_{(\alpha)}$ момент $\tau_{(\alpha)}^* = \inf\{t \geq 0 : \pi_t \geq 1 - \alpha\}$ является оптимальным для всех начальных состояний π и α , таких, что $\pi \leq 1 - \alpha$. (Заметим, что если начальное состояние $\pi_0 = \pi$ и $\pi \geq 1 - \alpha$, то $\tau_{(\alpha)}^* = 0$.)

2. В варианте B считается, что θ является параметром со значениями в $[0, \infty]$, и в классе $\mathfrak{M}_{(T)} = \{\tau : E_\infty \tau \geq T\}$ требуется найти тот момент τ_T^* (если таковой существует), что выполнено свойство (3):

$$\int_0^\infty E_\theta(\tau_T^* - \theta)^+ d\theta = \inf_{\tau \in \mathfrak{M}_{(T)}} \int_0^\infty E_\theta(\tau - \theta)^+ d\theta.$$

Замечательным обстоятельством здесь явилось то, что (нестандартное) выражение $\int_0^\infty E_\theta(\tau - \theta)^+ d\theta$ допускает для любого марковского момента $\tau \in \mathfrak{M}_{(T)}$ следующее представление:

$$\int_0^\infty E_\theta(\tau - \theta)^+ d\theta = E_\infty \int_0^\tau \psi_t dt, \quad (51)$$

где $(\psi_t)_{t \geq 0}$ — такой процесс, что

$$\psi_t = \int_0^t \frac{L_s}{L_t} ds, \quad (52)$$

и отношение правдоподобия L_t задается формулой (36). По формуле Ито из (52) и (37) находим, что

$$d\psi_t = dt + \frac{\mu}{\sigma^2} \psi_t dX_t \quad (53)$$

с $\psi_0 = 0$. (Доказательство представления (51) см. в [6, 7].)

Процесс $(\psi_t)_{t \geq 0}$ является (по каждой из мер $P_\theta \equiv \text{Law}(X|\theta)$) диффузионным марковским процессом. При этом относительно меры P^∞

$$d\psi_t = dt + \frac{\mu}{\sigma} \psi_t dW_t. \quad (54)$$

Для решения условно-вариационной задачи в варианте B надо, опять же согласно методу Лагранжа, отыскать оптимальный момент остановки $\tau^*(c)$, для которого

$$E_\infty \int_0^{\tau^*(c)} (\psi_t - c) dt = \inf_{\tau} E_\infty \int_0^\tau (\psi_t - c) dt, \quad (55)$$

где $\inf$ берется по всем моментам остановки $\tau \in \mathfrak{M}$.

Данная задача относится к классу стандартных задач типа (19'), и ее решение можно получить, решая соответствующую задачу Стефана (см. [6]). Важным следствием этого решения является то, что оптимальный момент $\tau^*(c)$ существует и имеет вид

$$\tau^*(c) = \inf\{t : \psi_t \geq B^*(c)\},$$

где $B^*(c)$ — некоторое пороговое значение (см. §4 в [6]).

Как и в случае варианта A , отсюда выводится, что $\inf_{\tau \in \mathfrak{M}_{(T)}} E_\infty \int_0^\tau \psi_t dt$ достигается на моменте остановки τ_T^* , который имеет следующий вид:

$$\tau_T^* = \inf\{t : \psi_t \geq B_T^*\} \quad (56)$$

и для которого $E_\infty \tau_T^*$ в точности равно T .

Значение этой константы B_T^* нетрудно найти. Действительно, из формулы (53) получаем

$$\psi_t = t + \frac{\mu}{\sigma} \int_0^t \psi_s dW_s.$$

Поэтому $\psi_{\tau_T^*} = \tau_T^* + \frac{\mu}{\sigma} \int_0^{\tau_T^*} \psi_s dW_s$, и, значит,

$$\mathbb{E}_\infty \psi_{\tau_T^*} = \mathbb{E} \tau_T^*, \quad (57)$$

поскольку для мартингала $(\int_0^t \psi_s dW_s)_{t \geq 0}$ и момента τ_T^* , определенного в (56), математическое ожидание $\mathbb{E}_\infty \int_0^{\tau_T^*} \psi_s dW_s = 0$.

С учетом того, что $\mathbb{E}_\infty \psi_{\tau_T^*} = B_T^*$ и $\mathbb{E} \tau_T^* = T$, видим, что пороговое значение $B_T^* = T$ и тем самым момент τ_T^* , определенный в (56), имеет весьма простую структуру:

$$\tau_T^* = \inf \{ t : \psi_t \geq T \} \quad (58)$$

и является оптимальным моментом остановки в варианте В.

4. Задача II

В отличие от модели (1), где в момент θ меняется снос броуновского движения, в рассматриваемой сейчас задаче (4) параметр θ играет роль момента, где броуновское движение меняет *тенденцию* своего движения.

Более точно: пусть $\theta = \theta(\omega)$ — момент, когда броуновское движение $W = (W_t)_{t \leq T}$ достигает своего максимального значения ($W_\theta = \max_{t \in [0, T]} W_t$). Важно подчеркнуть, что момент θ , будучи «внутренним» параметром броуновского движения, является *непредсказуемым* моментом, поскольку зависит от значений W_t при всех $t \in [0, T]$.

Обе задачи, и (4) и (5), по своей формулировке являются *нестандартными* задачами, если их рассматривать с точки зрения теории оптимальных правил остановки для марковских процессов. Однако, их снова можно свести к стандартной форме и затем воспользоваться результатами указанной теории.

Прежде всего отметим, что критерии (4) и (5) эквивалентны. Следовательно $[4, 9]$, для каждого момента $\tau \in \mathfrak{M}_T$

$$\mathbb{E} |W_\theta - W_\tau|^2 = \frac{1}{2} + \mathbb{E} |\tau - \theta| \quad (59)$$

и, значит, оптимальный момент τ_T^* в задаче (4) совпадает с оптимальным моментом τ_T^0 (похоже, в (5)).

Для решения задачи (4) заметим, что для любого момента $\tau \in \mathfrak{M}_T$ имеет место следующее представление:

$$\mathbb{E} |W_\theta - W_\tau|^2 = T + \mathbb{E} \int_0^\tau [1 - 2H_t] dt, \quad (60)$$

где

$$H_t = 2 \left(1 - \Phi \left(\frac{S_t - W_t}{\sqrt{T - t}} \right) \right), \quad (61)$$

$$S_t = \max_{u \leq t} W_u,$$

$$\Phi(x) = \int_{-\infty}^x \varphi(u) du, \quad \varphi(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}.$$

Для получения этого представления можно поступить, например, следующим образом. Рассмотрим величину $S_T = \max_{u \leq T} W_u$.

Как функционал от броуновского движения, S_T допускает стохастическое интегральное представление [10]

$$S_T = \mathbb{E} S_T + \int_0^T H_t dW_t, \quad (62)$$

где H_t определяется формулой (61).

Введя квадратично-интегрируемый мартингал $M = (M_t)_{t \leq T}$ с

$$M_t = \int_0^t H_s dW_s, \quad (63)$$

находим, что

$$\begin{aligned} \mathbb{E}[W_\theta - W_\tau]^2 &= \mathbb{E}[S_T - W_\tau]^2 = \mathbb{E} S_T^2 + \mathbb{E} W_\tau^2 - 2\mathbb{E} S_T W_\tau \\ &= T - 2\mathbb{E}(M_\tau W_\tau) + \mathbb{E} \tau = T + \mathbb{E} \int_0^\tau (1 - 2H_t) dt, \end{aligned}$$

где мы воспользовались тем результатом (Л. Башелье, П. Леви), что $S_T \stackrel{\text{law}}{=} |W_T|$ (то есть распределения S_T и $|W_T|$ совпадают) и что величины M_t допускают представление (63). Таким образом, из (60), (61) следует, что исходная (нестандартная) задача отыскания

$$V \equiv \inf_{\tau \in \mathfrak{M}_T} \mathbb{E}[W_\tau - W_\theta]^2 \quad (64)$$

сводится к стандартной задаче отыскания величины

$$U \equiv \inf_{\tau \in \mathfrak{M}_T} \mathbb{E} \int_0^\tau F\left(\frac{S_t - W_t}{\sqrt{T-t}}\right) dt, \quad (65)$$

где $F(x) = 4\Phi(x) - 3$, поскольку $V = T + U$.

Если положить $X_t = S_t - W_t$, $t \leq T$, то из (65) видим, что рассматриваемая (нестандартная) задача об оптимальной остановке (64) свелась к стандартной задаче типа (31) для процесса $X = (X_t)_{t \leq T}$, который является диффузионным марковским процессом с отражением в нуле.

Задача (65) является неоднородной стандартной марковской задачей, поскольку подынтегральное выражение в (65) зависит и от X_t , и от t . Однако эту задачу можно свести к некоторой неоднородной.

С этой целью заметим прежде всего, что по распределению процесс $(S_t - W_t)_{t \leq T}$ совпадает с марковским процессом $(|W_t|)_{t \leq T}$ («теорема П. Леви») и

$$U = \inf_{\tau \in \mathfrak{M}_T} \mathbb{E} \int_0^\tau F\left(\frac{|W_t|}{\sqrt{T-t}}\right) dt \quad (66)$$

(подробнее см. [9]).

Полагая для простоты $T = 1$, введем процесс $Z = (Z_s)_{s \geq 0}$ с

$$Z_s = e^s W_{1-e^{-2s}}. \quad (67)$$

По формуле Ито

$$dZ_s = Z_s ds + \sqrt{2} d\beta_s, \quad (68)$$

где процесс $\beta = (\beta_s)_{s \geq 0}$ с

$$\beta_s = \frac{1}{\sqrt{2}} \int_0^s e^u dW_{1-e^{-2u}} = \frac{1}{\sqrt{2}} \int_0^{1-e^{-2s}} \frac{1}{\sqrt{1-u}} dW_u \quad (69)$$

является броуновским движением.

Из (68), (69) вытекает, что характеристический (инфинитезимальный) оператор L_Z процесса $Z = (Z_s)_{s \geq 0}$ задается формулой

$$L_Z = z \frac{d}{dz} + \frac{d^2}{dz^2}. \quad (70)$$

Заменой переменных $t = 1 - e^{-2s}$ в (66) находим, используя (67), что

$$U = 2 \inf_{\tau \in \mathfrak{M}_1} \mathbb{E} \left(\int_0^{\sigma_\tau} e^{-2s} F(|Z_s|) ds \right), \quad (71)$$

где $\sigma_\tau = \log(1/\sqrt{1-\tau})$.

Момент σ_τ является марковским относительно потока $(\mathcal{F}_s^Z)_{s \geq 0}$, $\mathcal{F}_s^Z = \sigma(Z_u, u \leq t)$.

Тем самым

$$U = \inf_{\sigma} \mathbb{E} \int_0^{\sigma} e^{-2s} F(|Z_s|) ds, \quad (72)$$

где σ — момент остановки для класса $\mathfrak{M}$. Полагая

$$U(z) = \inf_{\sigma \in \mathfrak{M}} \mathbb{E}_z \int_0^{\sigma} e^{-2s} F(|Z_s|) ds, \quad (73)$$

где $\mathbb{E}_z$ — усреднение по мере $\mathbb{P}_z$ процесса $Z = (Z_s)_{s \geq 0}$ с $Z_0 = z$, получаем (с точностью до множителя e^{-2s} при $F(|Z_s|)$) однородную стандартную задачу об оптимальной остановке для однородного марковского процесса $Z = (Z_s)_{s \geq 0}$.

Решение задачи (73) можно получить, рассматривая соответствующую задачу Стефана.

Из соображений симметрии и общих результатов об оптимальной остановке, приведенных в предложениях 1 и 2 в разделе 1, получаем, что оптимальный момент остановки σ^* должен иметь следующий вид:

$$\sigma^* = \inf \{ s \geq 0 : |Z_s| > z^* \}$$

где z^* есть некоторая пороговая константа, подлежащая определению. С целью ее отыскания рассмотрим (вспомогательную) задачу Стефана относительно неизвестной функции $\widehat{U}(z)$ и порога $\widehat{z}$, таких, что

$$(L_Z - 2)\widehat{U}(z) = -F(|z|), \quad |z| < \widehat{z}, \quad (74)$$

$$\widehat{U}(\pm\widehat{z}) = 0 \text{ (мгновенная остановка),} \quad (75)$$

$$\widehat{U}'(\pm\widehat{z}) = 0 \text{ (условие гладкого склеивания).} \quad (76)$$

С учетом (70) и (74) находим, что в области $|z| < \widehat{z}$

$$\widehat{U}''(z) + z\widehat{U}'(z) - 2\widehat{U}(z) = -F(|z|). \quad (77)$$

Две константы, входящие в определение общего решения уравнения (77), и пороговая константа $\widehat{z}$ определяются из условий (75), (76) и дополнительного условия $\widehat{U}'(0) = 0$, естественно следующего из предположения $\widehat{U} \in C^2$ и симметрии этой функции $\widehat{U} = \widehat{U}(z)$. Как показано в [9], $\widehat{z}$ определяется как (единственный) корень уравнения

$$4\Phi(z) - 2z\varphi(z) - 3 = 0 \quad (78)$$

(что дает $\widehat{z} = 1, 12\dots$), а функция

$$\widehat{U}(z) = \Phi(\widehat{z})(1 + z^2) - z\varphi(z) + (1 - z^2)\Phi(z) - \frac{3}{2}. \quad (79)$$

Применением «проверочной теоремы» (предложение 3) устанавливаем, что найденное решение $(\widehat{U}, \widehat{z})$ задачи Стефана (74)–(76) дает искомую функцию U , определенную в (73), и оптимальный момент $\sigma^* = \inf\{s \geq 0 : |Z_s| > z^*\}$ с $z^* = \widehat{z}$, на котором достигается $\inf$ в задаче (73).

Процесс $Z = (Z_s)_{s \geq 0}$ и исходный процесс $W = (W_t)_{t \leq 1}$, входящий в (66) (с $T = 1$), связаны соотношением (67), которое показывает, что оптимальный в задаче (66) момент τ^* определяется по оптимальному моменту (в задаче (72)) σ^* по формуле $\sigma^* = \log(1/\sqrt{1 - \tau^*})$.

Наконец, учитывая отмеченное выше совпадение функционалов U в (65) и (66), а также то, что процессы $(S_t - W_t)_{t \leq T}$ и $(|W_t|)_{t \leq T}$ совпадают по распределению, находим, что (для любого $T > 0$) оптимальный (в задаче (64)) момент τ_T^* имеет следующий вид:

$$\tau_T^* = \inf\{t \leq T : S_t - B_t \geq z^*\sqrt{T - t}\}, \quad (80)$$

где $z^* = \widehat{z}$ ($= 1, 12\dots$). (О многих интересных свойствах момента τ_T^* см. [9].)

5. Задача III

1. В этой задаче, близкой по своей постановке к условно-вариационному варианту A задачи I, рассматриваются следующие два *непредсказуемых момента* для броуновского движения $W = (W_t)_{t \leq T}$:

$$\theta = \inf\left\{t \leq T : W_t = \max_{s \in [0, T]} W_s\right\}$$

и

$$g = \sup\{t \leq T : W_t = 0\},$$

где, как обычно, $\inf \emptyset = T$, $\sup \emptyset = 0$. Момент θ — это (P-п.н. единственный) момент, в который на $[0, T]$ броуновское движение достигает максимального значения; g — это момент последнего на $[0, T]$ обращения в нуль броуновского движения $W = (W_t)_{t \leq T}$.

Пусть τ и σ — моменты остановки, $0 < \alpha < 1$ и

$$\mathfrak{M}_\alpha(\theta) = \{ \tau \leq T : P(\tau < \theta) \leq \alpha \},$$

$$\mathfrak{M}_\alpha(g) = \{ \sigma \leq T : P(\sigma < g) \leq \alpha \}.$$

Рассматриваемые задачи состоят в том, чтобы в этих классах найти оптимальные моменты $\tau_T(\alpha)$ и $\sigma_T(\alpha)$, для которых

$$\inf_{\tau \in \mathfrak{M}_\alpha(\theta)} E(\tau - \theta)^+ = E(\tau_T(\alpha) - \theta)^+ \quad (81)$$

и

$$\inf_{\sigma \in \mathfrak{M}_\alpha(g)} E(\sigma - g)^+ = E(\sigma_T(\alpha) - g)^+. \quad (82)$$

Как и в варианте А задачи I, рассмотрим критерии

$$K(c, \theta) = \inf_{\tau \in \mathfrak{M}_T} [P(\tau < \theta) + cE(\tau - \theta)^+] \quad (83)$$

и

$$K(c, g) = \inf_{\sigma \in \mathfrak{M}_T} [P(\sigma < g) + cE(\sigma - g)^+], \quad (84)$$

где $\mathfrak{M}_T$ — класс марковских моментов (относительно $(\mathcal{F}_t^W)_{t \leq T}$, $\mathcal{F}_t^W = \sigma(W_s, s \leq t)$) со значениями в $[0, T]$.

Положим

$$\pi_t = P(\theta \leq t | \mathcal{F}_t^W)$$

и

$$\gamma_t = P(g \leq t | \mathcal{F}_t^W).$$

Имеем

$$\begin{aligned} \pi_t &= P(\theta \leq t | \mathcal{F}_t^W) = P\left(\max_{u \leq t} W_u \geq \max_{t \leq u \leq T} W_u \mid \mathcal{F}_t^W\right) = \\ &= P\left(\max_{u \leq t} W_u - W_t \geq \max_{t \leq u \leq T} (W_u - W_t) \mid \mathcal{F}_t^W\right). \end{aligned} \quad (85)$$

Поскольку величины $\max_{u \leq t} W_u - W_t$ являются $\mathcal{F}_t^W$ -измеримыми, а приращения $W_u - W_t$, $t \leq u \leq T$, не зависят от $\mathcal{F}_t^W$ и $\text{Law}(W_u - W_t, t \leq u \leq T) = \text{Law}(W_u, 0 \leq u \leq T - t)$, то из (85) находим, что

$$\pi_t = P\left(\max_{0 \leq u \leq T-t} W_u \leq b\right) \Big|_{b=S_t-W_t}, \quad (86)$$

где $S_t = \max_{u \leq t} W_u - W_t$.

Согласно уже отмечавшемуся в разделе 4 результату (Л. Башелье, П. Леви)

$$\text{Law}\left(\max_{0 \leq u \leq T-t} W_u\right) = \text{Law}(|W_{T-t}|). \quad (87)$$

Поскольку

$$P(|W_{T-t}| \leq x) = 2\Phi\left(\frac{x}{\sqrt{T-t}}\right) - 1, \quad (88)$$

то из (86)–(88) находим

$$\pi_t = 2\Phi\left(\frac{S_t - W_t}{\sqrt{T-t}}\right) - 1. \quad (89)$$

Аналогичным образом (см. детали в [1]) получаем

$$\gamma_t = 2\Phi\left(\frac{|W_t|}{\sqrt{T-t}}\right) - 1. \quad (90)$$

2. Замечание. Учитывая, что по известной теореме П. Леви

$$\text{Law}(S_t - W_t, t \leq t) = \text{Law}(|W_t|, t \leq T), \quad (91)$$

находим

$$\text{Law}(\pi_t, t \leq T) = \text{Law}(\gamma_t, t \leq T). \quad (92)$$

В частности, отсюда видим, что

$$\mathbb{P}(\theta \leq t) = \mathbb{E}\pi_t = 2\mathbb{E}\Phi\left(\frac{S_t - W_t}{\sqrt{T-t}}\right) - 1 = 2\mathbb{E}\Phi\left(\frac{|W_t|}{\sqrt{T-t}}\right) - 1 = \mathbb{P}(g \leq t). \quad (93)$$

Полагая здесь для простоты $T = 1$, находим

$$\mathbb{E}\Phi\left(\frac{|W_t|}{\sqrt{T-t}}\right) = 1 - \frac{1}{\pi} \arctg \sqrt{\frac{1-t}{t}}.$$

Поэтому

$$\begin{aligned} \mathbb{P}(\theta \leq t) &= 2\left(1 - \frac{1}{\pi} \arctg \sqrt{\frac{1-t}{t}}\right) - 1 = \\ &= 1 - \frac{2}{\pi} \arctg \sqrt{\frac{1-t}{t}} = 1 - \frac{2}{\pi} \arccos \sqrt{t} = \\ &= 1 - \frac{2}{\pi} \left(\frac{\pi}{2} - \arcsin \sqrt{t}\right) = \frac{2}{\pi} \arcsin \sqrt{t}. \end{aligned} \quad (94)$$

Таким образом, обращение к условным вероятностям $\mathbb{P}(\theta \leq t | \mathcal{F}_t^W)$ и $\mathbb{P}(g \leq t | \mathcal{F}_t^W)$ дает не только простое и наглядное доказательство того, что θ и g имеют одно и то же распределение, но и то, что это распределение задается *законом арксинуса* (94).

3. Аналогично рассмотренной в варианте A задаче I находим, что

$$K(c, \theta) = \inf_{\tau \in \mathfrak{M}_T} \mathbb{E}\left((1 - \pi_\tau) + c \int_0^\tau \pi_t dt\right)$$

и

$$K(c, g) = \inf_{\sigma \in \mathfrak{M}_T} \mathbb{E}\left((1 - \gamma_\sigma) + c \int_0^\sigma \gamma_t dt\right).$$

Используя представления (89) и (90), получаем

$$K(c, \theta) = \inf_{\tau \in \mathfrak{M}_T} \mathbb{E}\left[\left(1 - F\left(\frac{S_\tau - W_\tau}{\sqrt{T-\tau}}\right)\right) + c \int_0^\tau F\left(\frac{S_t - W_t}{\sqrt{T-t}}\right) dt\right] \quad (95)$$

и

$$K(c, g) = \inf_{\sigma \in \mathfrak{M}_T} \mathbb{E} \left[\left(1 - F \left(\frac{|W_\sigma|}{\sqrt{T-\sigma}} \right) \right) + c \int_0^\sigma F \left(\frac{|W_\sigma|}{\sqrt{T-t}} \right) dt \right], \quad (96)$$

где $F(x) = 2\Phi(x) - 1$.

Учитывая марковость процессов $(S_t - W_t)_{t \leq T}$ и $(|W_t|)_{t \leq T}$, видим, что исходные нестандартные задачи (83) и (84) сводятся к стандартным задачам (95) и (96).

Решения этих задач находятся аналогично решениям задач (65) и (66), обращением к рассмотрению соответствующих задач Стефана и привлечением «проверочной теоремы». В результате приходим к следующему результату: в классе $\mathfrak{M}_T$ оптимальные моменты $\tau^*(c)$ и $\sigma^*(c)$ в задачах (83) и (84) задаются следующими формулами:

$$\tau^*(c) = \inf \{ t \leq T : S_t - W_t \geq z^*(c) \sqrt{T-t} \}$$

и

$$\sigma^*(c) = \inf \{ t \leq T : |W_t| \geq z^*(c) \sqrt{T-t} \},$$

где $z^*(c)$ есть (единственный) корень уравнения

$$(1+c) \left(\Phi(z) - \frac{1}{2} \right) = \frac{1}{2} \varphi(z) \left(\frac{1}{z} + z - zc \right) \quad (97)$$

(см. детали в [11]).

Перейдем теперь к рассмотрению условно-экстремальных задач (81) и (82).

Для момента $\tau^*(c)$ вероятность ложной тревоги

$$\begin{aligned} \mathbb{P}(\tau^*(c) < \theta) &= \mathbb{E}(1 - \pi_{\tau^*(c)}) = \\ &= \mathbb{E} \left[2 \left(1 - \Phi \left(\frac{S_{\tau^*(c)} - W_{\tau^*(c)}}{\sqrt{T - \tau^*(c)}} \right) \right) \right] = 2[1 - \Phi(z^*(c))]. \end{aligned}$$

Отсюда ясно, что если желать иметь вероятность ложной тревоги равной α , то константу c надо выбрать так, чтобы

$$2[1 - \Phi(z^*(c))] = \alpha.$$

Если $0 < \alpha < 1$, то для z_α , такого, что

$$2[1 - \Phi(z_\alpha)] = \alpha$$

(то есть $z_\alpha = \Phi^{-1}(1 - \alpha/2)$), находим, что соответствующее значение c надо выбрать так, чтобы $z^*(c) = z_\alpha$. Функция $z^*(c)$ непрерывно зависит от c (см. (97)), и найдется c_α , такое, что $z^*(c_\alpha) = z_\alpha$. С таким значением $c = c_\alpha$ из доказанной выше оптимальности момента $\tau^*(c_\alpha)$ находим, что для всякого $\tau \in \mathfrak{M}_\alpha(\theta)$

$$\mathbb{P}(\tau < \theta) + c_\alpha \mathbb{E}(\tau - \theta)^+ \geq \mathbb{P}(\tau^*(c_\alpha) < \theta) + c_\alpha \mathbb{E}(\tau^*(c_\alpha) - \theta)^+. \quad (98)$$

Поскольку здесь $\mathbb{P}(\tau < \theta) \leq \alpha$ и $0 < c_\alpha < \infty$, то из (98) находим, что для $\tau \in \mathfrak{M}_\alpha(\theta)$

$$\mathbb{E}(\tau - \theta)^+ \geq \mathbb{E}(\tau^*(c_\alpha) - \theta)^+,$$

что и доказывает, что момент $\tau_T(\alpha) = \tau^*(c_\alpha)$ является оптимальным в задаче (81).

Аналогичным образом показывается, что момент $\sigma_T(\alpha) = \sigma^*(c_\alpha)$ с тем же самым значением c_α является оптимальным в задаче (82).

Итак, в задачах (81) и (82) оптимальные моменты имеют следующий вид:

$$\tau_T(\alpha) = \inf \{t \leq T : S_t - W_t \geq z_\alpha \sqrt{T-t}\}$$

и

$$\sigma_T(\alpha) = \inf \{t \leq T : |W_t| \geq z_\alpha \sqrt{T-t}\}$$

с z_α , являющимся корнем уравнения $1 - \alpha/2 = \Phi(z)$.

4. В том случае, когда вместо броуновского движения рассматривается броуновское движение со сносом $W^\mu = (W_t^\mu)_{t \leq T}$, $W_T^\mu = \mu T + W_T$, соответствующие оптимальные моменты $\tau_T^\mu(\alpha)$ и $\sigma_T^\mu(\alpha)$ имеют (как это можно вывести из результатов [12]) следующую структуру:

$$\tau_T^\mu(\alpha) = \inf \{t \leq T : S_t^\mu - W_t^\mu \geq b_\alpha^\mu(t)\}$$

и

$$\sigma_T^\mu(\alpha) = \inf \{t \leq T : W_t^\mu \geq \bar{b}_\alpha^\mu(t) \text{ или } \underline{b}_\alpha^\mu(t)\},$$

где $S_t^\mu = \max_{u \leq t} W_u^\mu$, $b_\alpha^\mu(t)$ и $\bar{b}_\alpha^\mu(t)$, $\underline{b}_\alpha^\mu(t)$ — некоторые граничные функции, по достижении которых или процессом $(S_t^\mu - W_t^\mu)_{t \leq T}$, или процессом $(W_t^\mu)_{t \leq T}$ происходит остановка наблюдений.

Точный вид этих граничных функций неизвестен. В нашей работе [11] указаны граничные функции $b(t) = b(t; \mu, \alpha)$ и $\bar{b}(t) = \bar{b}(t; \mu, \alpha)$, $\underline{b}(t) = \underline{b}(t; \mu, \alpha)$, для которых моменты

$$\tilde{\tau}_T^\mu(\alpha) = \inf \{t \leq T : S_t^\mu - B_t^\mu \geq b(t)\}$$

и

$$\tilde{\sigma}_T^\mu(\alpha) = \inf \{t \leq T : W_t^\mu \geq \bar{b}(t) \text{ или } W_t^\mu \leq \underline{b}(t)\}$$

таковы, что вероятности ложной тревоги $P(\tilde{\tau}_T^\mu(\alpha) < \theta)$ и $P(\tilde{\sigma}_T^\mu(\alpha) < g)$ в точности равны α . В [11] приведены графики этих функций для некоторых значений μ и α . По своей форме эти графики соответствуют тому качественному представлению, каким должны быть оптимальные граничные функции $b_\alpha^\mu(t)$, $\bar{b}_\alpha^\mu(t)$ и $\underline{b}_\alpha^\mu(t)$.

6. Задача IV

1. Описанная в разделе 1, эта задача состоит в том, чтобы найти такой момент τ_T^* , что

$$EU\left(\frac{P_{\tau_T^*}}{M_T}\right) = \sup_{\tau \in \mathfrak{M}_T} EU\left(\frac{P_\tau}{M_T}\right), \quad (99)$$

где $P_t = S_t/B_t$, процесс $S = (S_t)_{t \leq T}$ является геометрическим броуновским движением с

$$dS_t = S_t(\mu dt + \sigma dW_t), \quad S_0 = 1$$

и $B = (B_t)_{t \leq T}$ есть банковский счет с

$$dB_t = rB_t dt, \quad B_0 = 1.$$

Задача (99) является, вообще говоря, *нестандартной*, поскольку величина $M_T = \max_{t \leq T} P_t$ зависит от поведения P_t при всех $t \in [0, T]$.

В том же специальном случае, когда функция потерь является *логарифмической*: $U(x) = \log x$, задача отыскания оптимального момента τ_T^* , такого, что

$$\mathbb{E} \log \frac{P_{\tau_T^*}}{M_T} = \sup_{\tau \in \mathfrak{M}_T} \mathbb{E} \log \frac{P_\tau}{M_T},$$

существенно упрощается, поскольку для отыскания оптимального τ_T^* можно не учитывать значение M_T .

Действительно, поскольку

$$\mathbb{E} \log \frac{P_\tau}{M_T} = \mathbb{E} \log P_\tau - \mathbb{E} \log M_T,$$

то видим, что оптимальным здесь является тот момент τ_T^* , для которого

$$\mathbb{E} \log P_{\tau_T^*} = \sup_{\tau \in \mathfrak{M}_T} \mathbb{E} \log P_\tau.$$

Так как

$$P_t = e^{(\mu - r - \sigma^2/2)t + \sigma W_t},$$

то

$$\mathbb{E} \log P_\tau = \mathbb{E}[(\mu - r - \sigma^2/2)\tau + \sigma W_\tau].$$

Для каждого $\tau \in \mathfrak{M}_T$ с $T < \infty$ математическое ожидание $\mathbb{E} W_\tau = 0$ и, значит,

$$\sup_{\tau \in \mathfrak{M}_T} \mathbb{E} \log P_\tau = \sup_{\tau \in \mathfrak{M}_T} \mathbb{E}[(\mu - r - \sigma^2/2)\tau].$$

Отсюда видно, что оптимальным является *вырожденный* (!) момент

$$\tau_T^* = \begin{cases} 0, & \text{если } \mu - r - \sigma^2/2 \leq 0; \\ T, & \text{если } \mu - r - \sigma^2/2 > 0. \end{cases} \quad (100)$$

Оказывается, что этот момент будет оптимальным и в более интересном случае *линейной* функции полезности $U(x) = x$. В этом случае приходится учитывать наличие $M_T = \sup_{t \leq T} P_t$ в $\mathbb{E}(P_\tau/M_T)$, что делает рассматриваемую задачу действительно нестандартной, и ее исследование становится значительно более сложным [13, 14].

2. Прежде всего займемся сведением нестандартной задачи

$$\mathbb{E} \frac{P_{\tau_T^*}}{M_T} = \sup_{\tau \in \mathfrak{M}_T} \mathbb{E} \frac{P_\tau}{M_T} \quad (101)$$

к стандартной для некоторого марковского процесса.

Запишем $P_t = S_t/B_t$ в виде

$$P_t = e^{W_t^\nu}, \quad (102)$$

где $W_t^\nu = \nu t + \sigma W_t$, $\nu = \mu - r - \sigma^2/2$.

Тогда, полагая, $M_t^\nu = \max_{0 \leq u \leq t} W_u^\nu$, находим, что

$$\begin{aligned} \mathbb{E} \frac{P_t}{M_T} &= \mathbb{E} \exp \left[\frac{e^{W_t^\nu}}{\max_{0 \leq t \leq T} e^{W_t^\nu}} \right] = \mathbb{E} \exp \left[\frac{e^{W_t^\nu}}{e^{M_t^\nu} \vee e^{\max_{t \leq u \leq T} W_u^\nu}} \right] = \\ &= \mathbb{E} \left[e^{-(M_t^\nu - W_t^\nu)} \wedge e^{-\max_{t \leq u \leq T} (W_u^\nu - W_t^\nu)} \right] = \\ &= \mathbb{E} \left\{ \mathbb{E} \left[e^{-(M_t^\nu - W_t^\nu)} \wedge e^{-\max_{t \leq u \leq T} (W_u^\nu - W_t^\nu)} \mid \mathcal{F}_t^W \right] \right\} = \\ &= \mathbb{E} \left\{ \mathbb{E} \left[e^{-x} \wedge e^{-M_{T-t}^\nu} \right]_{x=M_t^\nu - W_t^\nu} \right\} = \mathbb{E} G(t, M_t^\nu - W_t^\nu), \quad (103) \end{aligned}$$

где мы воспользовались независимостью $M_t^\nu - W_t^\nu$ от $\mathcal{F}_t^W$, тем, что $\text{Law}(W_u^\nu - W_t^\nu, t \leq u \leq T) = \text{Law}(W_u^\nu, 0 \leq u \leq T - t)$, и положили

$$G(t, x) = \mathbb{E} \left[e^{-x} \wedge e^{-M_{T-t}^\nu} \right], \quad t \in [0, T], \quad x \in [0, \infty). \quad (104)$$

Таким образом, если

$$X_t^\nu = M_t^\nu - W_t^\nu,$$

то

$$\sup_{\tau \in \mathfrak{M}_T} \mathbb{E} \frac{P_\tau}{M_T} = \sup_{\tau \in \mathfrak{M}_T} \mathbb{E} G(\tau, X_\tau^\nu). \quad (105)$$

Процесс $(X_t^\nu)_{t \leq T}$ является марковским, и, следовательно, исходная нестандартная задача (101) свелась к стандартной (неоднородной) задаче отыскания момента τ_T^* , такого, что

$$\mathbb{E} G(\tau_T^*, X_{\tau_T^*}^\nu) = \sup_{\tau \in \mathfrak{M}_T} \mathbb{E} G(\tau, X_\tau^\nu) \quad (106)$$

для марковского процесса $(X_t^\nu)_{t \leq T}$.

Полагая для простоты рассуждений $\sigma = 1$, находим (см. подробнее [13]) следующие представления для $G(t, x)$:

$$\begin{aligned} G(t, x) &= \frac{2(\nu - 1)}{2\nu - 1} e^{-(\nu - 1/2)(T-t)} \Phi \left(\frac{-x + (\nu - 1)(T-t)}{\sqrt{T-t}} \right) + \\ &+ \frac{1}{2\nu - 1} e^{-(1-2\nu)x} \Phi \left(\frac{-x - \nu(T-t)}{\sqrt{T-t}} \right) + e^{-x} \Phi \left(\frac{x - \nu(T-t)}{\sqrt{T-t}} \right), \quad (107) \end{aligned}$$

если $\nu \neq 1/2$, и

$$\begin{aligned} G(t, x) &= (1 + x + (T-t)/2) \Phi \left(\frac{-x - (T-t)/2}{\sqrt{T-t}} \right) - \\ &- \sqrt{\frac{T-t}{2\pi}} e^{-\frac{(x+(T-t)/2)^2}{2(T-t)}} + e^{-x} \Phi \left(\frac{-x - (T-t)/2}{\sqrt{T-t}} \right), \quad (108) \end{aligned}$$

если $\nu = 1/2$.

Для решения задачи (106) воспользуемся методом решения таких (неоднородных) задач, изложенным в разделе 2.

Положим для $0 \leq t \leq T$, $x \in [0, \infty)$

$$V(t, x) = \sup_{0 \leq \tau \leq T-t} \mathbb{E}_x G(t + \tau, X_{t+\tau}^x), \quad (109)$$

где $(X_t^x)_{0 \leq t \leq T}$ — марковский процесс с

$$X_t^x = x \vee X_t^\nu (= x \vee (M_t^\nu - W_t^\nu)) \quad (110)$$

и $\mathbb{E}_x$ — усреднение по мере $\mathbf{P}_x$ этого процесса, когда $X_0^x = x$.

Из (109) видим, что интересующая нас задача (105) сводится к отысканию значения $V(0, 0)$.

Введем множества

$$C = \{(t, x) \in [0, T) \times [0, \infty) : V(t, x) > G(t, x)\}$$

и

$$D = \{(t, x) \in [0, T) \times [0, \infty) : V(t, x) = G(t, x)\}.$$

Согласно общей теории оптимальных правил остановки (раздел 2), естественно ожидать, что множества C и D являются множествами продолжения и прекращения наблюдений и момент первого попадания в множество D является оптимальным.

Один из методов отыскания функции $V = V(t, x)$ и областей C и D состоит в решении соответствующей задачи Стефана: найти $\widehat{V}$, $\widehat{C}$ и $\widehat{D}$, такие, что

$$L_X \widehat{V}(t, x) = 0, \quad (t, x) \in \widehat{C}, \quad (111)$$

$$\widehat{V}(t, x) = G(t, x), \quad (t, x) \in \widehat{D}. \quad (112)$$

Здесь L_X — характеристический (инфинитезимальный) оператор процесса $X^x = (X_t^x)_{t \geq 0}$. Хорошо известно (см., например, [4, 15]), что распределение $\text{Law}(X_t^x, t \leq T | X_0 = x)$ совпадает с распределением $\text{Law}(|Y_t^x|, t \leq T)$, где $Y^x = (Y_t^x)_{t \leq T}$ есть так называемый «bang-bang» процесс, удовлетворяющий стохастическому дифференциальному уравнению

$$dY_t^x = -\nu \operatorname{sgn} Y_t^x dt + d\widetilde{W}_t, \quad Y_0^x = x,$$

где $\widetilde{W} = (\widetilde{W}_t)_{t \leq T}$ — некоторый винеровский процесс (броуновское движение).

По формуле Танака

$$|Y_t^x| = |x| + \int_0^t \operatorname{sgn} Y_s^x dY_s^x + L_t(Y^x) = |x| - \nu t + \int_0^t \operatorname{sgn} Y_s^x d\widetilde{W}_s + L_t(Y^x), \quad (113)$$

где $L_t(Y^x)$ — локальное время процесса Y^x в нуле на интервале $[0, t]$. Отсюда, ввиду совпадения распределений вероятностей процессов X^x и $|Y^x|$, находим, что характеристический оператор процесса X^x имеет следующий вид: для функции $f = f(t, x) \in C^{1,2}$

$$L_X f(t, x) = \frac{\partial f}{\partial t} - \nu \frac{\partial f}{\partial x} + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}, \quad t \in [0, T), \quad x \in (0, \infty), \quad (114)$$

с условием отражения в нуле

$$\frac{\partial f}{\partial x}(t, 0+) = 0. \quad (115)$$

Далее мы не будем искать соответствующее решение задачи Стефана, а применим другой метод, который можно рассматривать как третий метод (в дополнение к «итеративному» и «стефановскому» методам, изложенным в разделе 2) решения задач об оптимальной остановке типа (109).

Суть этого метода состоит в следующем. Если функция $G(t, x) \in C^{1,2}$, то по формуле Ито

$$\begin{aligned} G(t+s, |Y_s^x|) &= G(t, x) + \int_0^s L_X G(t+u, |Y_u^x|) du + \\ &+ \int_0^s \frac{\partial G}{\partial x}(t+u, |Y_u^x|) d\widetilde{W}_u + \int_0^s \frac{\partial G}{\partial x}(t+u, |Y_u^x|) dL_u(Y^x). \end{aligned} \quad (116)$$

В силу (115) последний член в (116) обращается в нуль, и

$$M_s = \int_0^s \frac{\partial G}{\partial x}(t+u, |Y_u^x|) d\widetilde{W}_u, \quad 0 \leq s \leq T$$

является мартингалом, где

$$-1 \leq \frac{\partial G}{\partial x} \leq 0.$$

Тем самым из (116) заключаем, что

$$\begin{aligned} V(t, x) &= \sup_{\tau \in \mathfrak{M}_{T-t}} \mathbb{E}_x G(t+\tau, X_\tau^x) = \sup_{\tau \in \mathfrak{M}_{T-t}} \mathbb{E} G(t+\tau, |Y_\tau^x|) = \\ &= G(t, x) + \sup_{\tau \in \mathfrak{M}_{T-t}} \mathbb{E}_x \int_0^\tau H(t+u, |Y_u^x|) du = \\ &= G(t, x) + \sup_{\tau \in \mathfrak{M}_{T-t}} \mathbb{E}_x \int_0^\tau H(t+u, X_u^x) du, \end{aligned} \quad (117)$$

где

$$H(t, x) \equiv L_x G(t, x) = \frac{\partial G}{\partial t} - \nu \frac{\partial G}{\partial x} + \frac{1}{2} \frac{\partial^2 G}{\partial x^2}. \quad (118)$$

Имея представления (107) и (108) для $G = G(t, x)$, находим

$$H(t, x) = (\nu - 1/2)G(t, x) - \frac{\partial G}{\partial x}(t, x). \quad (119)$$

Итак, для $V(t, x)$ получено представление

$$V(t, x) = G(t, x) + \sup_{\tau \in \mathfrak{M}_{T-t}} \mathbb{E}_x \int_0^\tau H(t+u, X_u^x) du, \quad (120)$$

играющее ключевую роль в третьем методе отыскания оптимальных моментов остановки.

Если $\nu \geq 1/2$, то прямые вычисления, основанные на представлениях (107) и (108) (см. детали в [13]), показывают, что $\partial G/\partial x \leq 0$. Поэтому из (119) заключаем, что $H(t, x) \geq 0$ для всех $(t, x) \in [0, T] \times [0, \infty)$, и из (120) видим, что в рассматриваемом случае $\nu \geq 1/2$

$$V(t, x) \geq G(t, x),$$

причем если $\nu > 1/2$, то $V(t, x) > G(t, x)$, $0 \leq t < T$, $x > 0$.

Отсюда следует, что в случае $\nu > 1/2$ не надо останавливаться до момента времени T с автоматической остановкой в момент $t = T$, когда $V(T, x) = G(T, x)$. Таким образом, при $\nu > 1/2$ оптимальным моментом является детерминированный момент $\tau_T^* = T$.

Если $\nu = 1/2$, то любой детерминированный момент $\tau_T^* = t$ является оптимальным.

Если $\nu < -1/2$, то поскольку

$$H(t, x) = (\nu + 1/2)G(t, x) - \left(G(t, x) + \frac{\partial G}{\partial x} \right),$$

и, как показывают прямые вычисления,

$$G(t, x) + \frac{\partial G}{\partial x} > 0,$$

из (120) следует, что

$$V(t, x) \leq G(t, x),$$

и, значит, оптимальным является момент $\tau_T^* = 0$.

Рассмотрение случаев $-1/2 < \nu < 1/2$ несколько более сложное (см. [13, 14]). В целом же получаем, что, как и в случае логарифмической функции потерь, оптимальным является момент

$$\tau_T^* = \begin{cases} 0, & \text{если } \nu \leq 0; \\ T, & \text{если } \nu > 0, \end{cases}$$

где $\nu = \mu - r - \sigma^2/2$.

7. Задача V

1. Будем, рассматривая модель (11)–(15), интересоваться отысканием оптимального момента τ^* , для которого

$$\mathbb{E} \log \frac{P_{\tau^*}}{M_T} = \sup_{\tau \in \mathfrak{M}_T} \mathbb{E} \log \frac{P_\tau}{M_T} (\equiv V_T), \quad (121)$$

то есть будем рассматривать задачу (9) для логарифмической функции полезности $U(x) = \log x$.

Интересно отметить, что решение задачи (121) дает в то же самое время решение задачи, рассмотренной в [16], о том, как в предположениях (11)–(15) производить ребалансировку портфеля с тем, чтобы получить в момент времени T максимальный капитал (на банковском счете).

Действительно, если на интервале $[0, \tau]$ все средства вложены в акцию и в момент τ происходит продажа акции с размещением ее средств на банковском счете, то капитал в момент времени T будет равен $S_\tau \cdot \frac{B_T}{B_\tau} = \frac{S_\tau}{B_\tau} \cdot B_T = P_\tau \cdot B_T$.

Тем самым

$$W_T \equiv \sup_{\tau \in \mathfrak{M}_T} \mathbb{E} \log \left(S_\tau \cdot \frac{B_T}{B_\tau} \right) = \sup_{\tau \in \mathfrak{M}_T} \mathbb{E} \log P_\tau + \log B_T = V_T + \mathbb{E} \log M_T + \log B_T, \quad (122)$$

и оптимальные моменты в V_T -задаче (121) и W_T -задаче (122) совпадают.

2. Для отыскания оптимального момента в V_T -задаче запишем P_t в виде

$$P_t = \exp \left\{ \int_0^t \nu(s, \theta) ds + \sigma W_t \right\}, \quad (123)$$

где

$$\nu(s, \theta) = \mu(s, \theta) - r - \sigma^2/2$$

с

$$\mu(s, \theta) = \mu_1 I(s < \theta) + \mu_2 I(s \geq \theta).$$

Тем самым

$$\mathbb{E} \log P_\tau = \mathbb{E} \int_0^\tau \nu(s, \theta) ds + \sigma \mathbb{E} W_\tau = \mathbb{E} \int_0^\tau \nu(s, \theta) ds,$$

поскольку $\mathbb{E} W_\tau = 0$ для любых марковских моментов $\tau \leq T$.

Если обозначить $\pi_t = \mathbb{P}(\theta \leq t | \mathcal{F}_t^S)$, то найдем

$$\mathbb{E} \int_0^\tau \nu(s, \theta) ds = \mathbb{E} \left[\nu_1 \tau + (\nu_2 - \nu_1) (\tau \pi_\tau - \mathbb{E}(\theta I(\theta < \tau) | \mathcal{F}_\tau^S)) \right]. \quad (124)$$

Сделаем теперь предположение о том, что распределение для θ является, как и в задаче I (раздел 3), экспоненциальным, описываемым формулами (33).

Подобно тому, как в разделе 3 выводилось стохастическое дифференциальное уравнение (42), для рассматриваемой сейчас модели (11), (12) находим

$$d\pi_t = \lambda(1 - \pi_t) dt + \frac{\nu_2 - \nu_1}{\sigma} \pi_t(1 - \pi_t) d\overline{W}_t, \quad \pi_0 = 0, \quad (125)$$

где $\overline{W}_t = (\overline{W}_t)_{t \leq T}$ является (обновляющим) броуновским движением.

Точно так же для процесса $(\rho_t)_{t \leq T}$ с $\rho_t = \mathbb{E}(\theta I(\theta \leq t) | \mathcal{F}_t^S)$ найдем

$$d\rho_t = \lambda t(1 - \pi_t) dt + \frac{\nu_2 - \nu_1}{\sigma} \gamma_t(1 - \pi_t) d\overline{W}_t, \quad \rho_0 = 0, \quad (126)$$

где

$$\gamma_t = \lambda e^{\lambda t} \int_0^t s e^{-\lambda s} \frac{L_t}{L_s} ds, \quad L_t = \frac{d\mu_t^{(\nu_1)}}{d\mu_t^{(\nu_2)}}$$

и $\mu_t^{(\nu_i)}$ — мера процесса $H^i = (H_s^i)_{s \leq t}$ с $dH_s^i = \nu_i ds + \sigma dW_s$, $i = 1, 2$.

Учитывая также, что

$$d(t\pi_t) = \pi_t dt + t \left\{ \lambda(1 - \pi_t) dt + \frac{\nu_2 - \nu_1}{\sigma} \pi_t(1 - \pi_t) d\overline{W}_t \right\}, \quad (127)$$

из (124)–(126) получаем, что

$$\begin{aligned} \mathbb{E} \int_0^\tau \nu(s, \theta) ds &= \mathbb{E} [\nu_1 \tau + (\nu_2 - \nu_1)(\tau \pi_\tau - \rho_\tau)] = \\ &= \mathbb{E} \int_0^\tau \left\{ \nu_1 + (\nu_2 - \nu_1) [\pi_t + \lambda t(1 - \pi_t)] - (\nu_2 - \nu_1) \lambda t(1 - \pi_t) \right\} dt = \\ &= \mathbb{E} \int_0^\tau [\nu_1 + (\nu_2 - \nu_1) \pi_t] dt. \end{aligned} \quad (128)$$

Тем самым

$$\sup_{\tau \in \mathfrak{M}_T} \mathbb{E} \log P_\tau = \sup_{\tau \in \mathfrak{M}_T} \mathbb{E} \int_0^\tau [\nu_1 + (\nu_2 - \nu_1) \pi_t] dt, \quad (129)$$

и, значит, исходная нестандартная задача (121) сведена к стандартной задаче (типа (19')) для диффузионного марковского процесса $(\pi_t)_{t \leq T}$ с дифференциалом (125).

Преобразуем правую часть в (129). С этой целью заметим, что в силу (125)

$$\lambda t = \pi_t + \lambda \int_0^t \pi_s ds - \int_0^t \frac{\nu_2 - \nu_1}{\sigma} \pi_s(1 - \pi_s) d\overline{W}_s. \quad (130)$$

Тем самым

$$t = \frac{\pi_t}{\lambda} + \int_0^t \pi_s ds - \int_0^t \frac{\nu_2 - \nu_1}{\lambda \sigma} \pi_s(1 - \pi_s) d\overline{W}_s$$

и

$$\nu_1 t = \frac{\nu_1}{\lambda} \pi_t + \nu_1 \int_0^t \pi_s ds - \int_0^t \frac{\nu_1(\nu_2 - \nu_1)}{\lambda \sigma} \pi_s(1 - \pi_s) d\overline{W}_s.$$

Поэтому

$$\begin{aligned} \mathbb{E} \int_0^\tau [\nu_1 + (\nu_2 - \nu_1) \pi_t] dt &= \mathbb{E} \left[\frac{\nu_1}{\lambda} \pi_\tau + \nu_2 \int_0^\tau \pi_t dt \right] = \\ &= \frac{\nu_1}{\lambda} \mathbb{E} \left[\pi_\tau + \frac{\nu_2}{\nu_1} \lambda \int_0^\tau \pi_t dt \right] = -\frac{\nu_1}{\lambda} \mathbb{E} \left\{ -\pi_\tau + \frac{|\nu_2|}{\nu_1} \lambda \int_0^\tau \pi_t dt \right\}. \end{aligned}$$

Отсюда, полагая $c = \frac{|\nu_2|}{\nu_1} \lambda$, находим, что

$$\mathbb{E} \int_0^\tau [\nu_1 + (\nu_2 - \nu_1) \pi_t] dt = \frac{\nu_1}{\lambda} - \frac{\nu_1}{\lambda} \mathbb{E} \left\{ (1 - \pi_\tau) + c \int_0^\tau \pi_t dt \right\} \quad (131)$$

и

$$\sup_{\tau \in \mathfrak{M}_T} \mathbb{E} \int_0^\tau [\nu_1 + (\nu_2 - \nu_1) \pi_t] dt = \frac{\nu_1}{\lambda} - \inf_{\tau \in \mathfrak{M}_T} \frac{\nu_1}{\lambda} \mathbb{E} \left\{ (1 - \pi_\tau) + c \int_0^\tau \pi_t dt \right\}.$$

Из этой формулы мы видим, что исходная задача (121), равносильная задаче

$$\tau \rightsquigarrow \sup_{\tau \in \mathfrak{M}_T} \mathbb{E} \log P_\tau, \quad (132)$$

свелась к задаче

$$\tau \rightsquigarrow \inf_{\tau \in \mathfrak{M}_T} \left\{ (1 - \pi_\tau) + c \int_0^\tau \pi_t dt \right\}. \quad (133)$$

Для случая $T = \infty$ эта задача была рассмотрена в разделе 3, из которого следует, что оптимальный момент

$$\tau_\infty^* = \inf \{ t \geq 0 : \pi_t \geq A_\infty^* \}, \quad (134)$$

где A_∞^* является корнем уравнения (50) (с заменой $\rho = r^2/(2\sigma^2)$ на $\rho = (\mu_1 - \mu_2)^2/(2\sigma^2)$). Тем самым при *больших* T этот момент может быть использован в качестве *почти оптимального момента* и в задаче (132).

Интересно отметить, что из представления

$$\mathbb{E} \log P_\tau = \mathbb{E} \int_0^\tau [\nu_1 + (\nu_2 - \nu_1)\pi_t] dt \quad (135)$$

видно, что если начинать наблюдения с $\pi_0 = 0$, то, поскольку в (135) величина $(\nu_2 - \nu_1)\pi_t < 0$ при $t < 0$ и $(\nu_2 - \nu_1)\pi_0 = 0$, не следует останавливаться до достижения порога $\tilde{A}_\infty = \nu_1/(\nu_1 - \nu_2) = \nu_1/(\nu_1 + |\nu_2|) = \lambda/(\lambda + c)$. Можно показать, что $A_\infty^* > \tilde{A}_\infty$.

В случае *конечного* T решение задачи (133) (а значит, и задачи (132)) намного сложнее. Дело в том, что устанавливается [4, 17], что оптимальный момент τ_T^* имеет следующий вид:

$$\tau_T^* = \inf \{ t \leq T : \pi_t \geq g_T(t) \}, \quad (136)$$

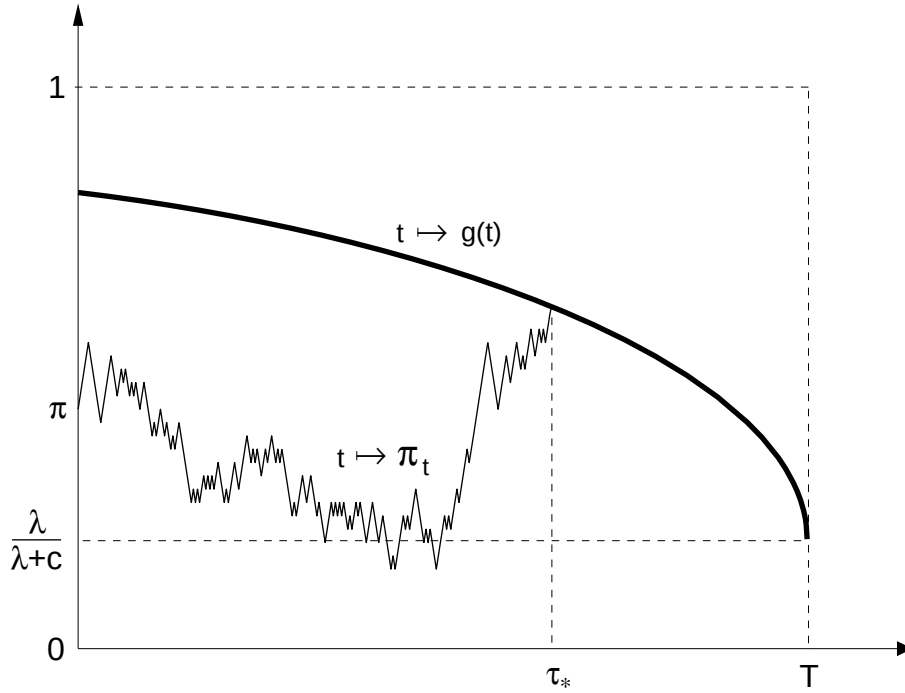
где оптимальная граница $g(t) = g_T(t)$, $0 \leq t \leq T$, является единственным непрерывным решением интегрального уравнения

$$\begin{aligned} \mathbb{E}_{t,g(t)} \pi_T = g(t) + c \int_0^{T-t} \mathbb{E}_{t,g(t)} [\pi_{t+u} I(\pi_{t+u} < g(t+u))] du + \\ + \lambda \int_0^{T-t} \mathbb{E}_{t,g(t)} [(1 - \pi_{t+u}) I(\pi_{t+u} < g(t+u))] du. \end{aligned} \quad (137)$$

В явном виде решить это уравнение не удастся прежде всего потому, что не удастся найти математическое ожидание $\mathbb{E}_{t,g(t)}[\cdot]$, входящее в правую часть (137). Однако это уравнение допускает численное решение, изображенное на приводимом рисунке.

3. Завершая рассмотрение задач, в котором проиллюстрирован метод редукции нестандартных задач к стандартным для соответствующих марковских процессов, отметим (помимо уже изложенных «итеративных» и «стефановских») следующие часто используемые методы решения собственно задач об оптимальной остановке вида

$$V_T = \sup_{\tau \in \mathfrak{M}_T} \mathbb{E} G_\tau, \quad (138)$$



где величины G_t являются $\mathcal{F}_t$ -измеримыми и τ — моменты остановки относительно потока $(\mathcal{F}_t)_{t \leq T}$.

Пусть Q — некоторая мера, локально абсолютно непрерывная относительно меры P (то есть $Q_t \ll P_t$, где $Q_t = Q|\mathcal{F}_t$, $P_t = P|\mathcal{F}_t$) и $Z_t = dQ_t/dP_t$. Предположим, что найдутся такие функции $g = g(x)$ и процесс $X = (X_t)_{t \leq T}$, что G_t может быть представлена в виде

$$G_t = Z_t g(X_t). \quad (139)$$

Тогда для любого $\tau \leq T$

$$EG_\tau = EQg(X_\tau). \quad (140)$$

Действительно,

$$EQg(X_\tau) = Eg(X_\tau) \frac{dQ_T}{dP_T} = Eg(X_\tau) Z_\tau = EG_\tau. \quad (141)$$

Если $(X_t)_{t \leq T}$ — марковский процесс (по мере Q_T), то равенство (140) будет означать, что задача (138) свелась к стандартной марковской задаче отыскания $\sup_{\tau \in \mathfrak{M}_T} EQg(X_\tau)$. Представление (140) приводит к следующему методу (систематически используемому в [18]) отыскания оптимального момента в задаче (138).

Предположим, что функция $g(x) \geq 0$ и $\sup_x g(x)$ достигается в точке x^* . Тогда из (141) имеем

$$EG_\tau = Eg(X_\tau) Z_\tau \leq g(x^*) EZ_\tau = g(x^*), \quad (142)$$

поскольку $EZ_\tau = 1$ для $\tau \leq T$.

Если на временном интервале $[0, T]$ процесс $(X_t)_{t \leq T}$ с вероятностью единица достигает значения x^* , то из (142) будет следовать, что момент

$$\tau_T^* = \inf \{t \leq T : X_t = x^*\}$$

является в задаче $\tau \rightsquigarrow \sup_{\tau \in \mathfrak{M}_T} EQg(X_\tau)$ оптимальным.

Наконец, отметим еще один метод отыскания в марковском случае функции

$$V(x) = \sup_{\tau \in \mathfrak{M}} \mathbb{E}_x G(X_\tau).$$

Из общей теории (раздел 2) мы знаем, что функция $V(x)$ является наименьшей супергармонической мажорантой функции $G(x)$.

Рассмотрим специальный класс $\widetilde{\mathfrak{M}} \subseteq \mathfrak{M}$ моментов остановки, скажем моментов типа моментов первого достижения уровней $a \geq 0$:

$$\tau_a = \inf \{t \geq 0 : X_t \geq a\}.$$

Введем функцию

$$\widetilde{V}(x) = \sup_{\tau \in \widetilde{\mathfrak{M}}} \mathbb{E}_x G(X_\tau).$$

Пользуясь специальной структурой моментов $\tau_a \in \widetilde{\mathfrak{M}}$, в ряде задач (см., например, [19, 20]) удастся установить, что функция $\widetilde{V}(x)$ является супергармонической функцией, мажорирующей функцию $G(x)$. Поскольку функция $V(x)$ является *наименьшей* супергармонической функцией, мажорирующей функцию $G(x)$, то $V(x) \leq \widetilde{V}(x)$. Но с другой стороны, очевидно, $V(x) \geq \widetilde{V}(x)$. Тем самым $\widetilde{V}(x) = V(x)$ и оптимальный момент в классе $\widetilde{\mathfrak{M}}$ будет оптимальным и в классе $\mathfrak{M}$.

Список литературы

- [1] *Ширяев А. Н.* Статистический последовательный анализ. М.: Наука, 1969.
- [2] *Роббинс Г., Сигмунд Д., Чао И.* Теория оптимальных правил остановки. М.: Наука, 1977.
- [3] *Ширяев А. Н.* Основы стохастической финансовой математики: в 2 т. Т. 1: Факты. Модели, Т. 2: Теория. М.: Фазис, 1998.
- [4] *Peskir G., Shiryaev A. N.* Optimal stopping and free-boundary problems. Basel: Birkhäuser, 2006.
- [5] *Лунцер Р. Ш., Ширяев А. Н.* Статистика случайных процессов. М.: Наука, 1974.
- [6] *Feinberg E. A., Shiryaev A. N.* Quickest detection of drift change for Brownian motion in generalized Bayesian and minimax settings // Statistics and Decisions, 2006. **24**, № 4. P. 445–470.
- [7] *Ширяев А. Н.* От «разладки» к нелинейной фильтрации и теории мартингалов // Математические события XX века. М.: Фазис, 2003. С. 491–518.
- [8] *Урусов М.* Об одном свойстве момента достижения максимума броуновским движением и некоторых задачах оптимальной остановки // Теория вероятностей и ее применения. 2004. **49**, № 1. С. 184–190.
- [9] *Graversen S. E., Peskir G., Shiryaev A. N.* Stopping Brownian motion without anticipation as close as possible to its ultimate maximum // Теория вероятностей и ее применения. 2000. **45**, № 1. С. 125–136.
- [10] *Ширяев А. Н., Йор М.* К вопросу о стохастических интегральных представлениях функционалов от броуновского движения. I // Теория вероятностей и ее применения. 2003. **48**, № 2. С. 375–385.

- [11] Ширяев А. Н. Об условно-экстремальных задачах скорейшего обнаружения непредсказуемых моментов у наблюдаемого броуновского движения // Теория вероятностей и ее применения. 2008. **53**, № 4. С. 751–768.
- [12] Du Toit J., Peskir G., Shiryaev A. N. Predicting of last zero of Brownian motion with drift // Stochastics. 2008. **80**, № 2–3. P. 229–245.
- [13] Shiryaev A. N., Xu Z., Zhou X. Y. Thou shalt buy and hold. // Quantitative Finance. 2008. **8**, № 8. P. 765–776.
- [14] Du Toit J., Peskir G. Selling a Stock at the Ultimate Maximum. Research Report № 10. Manchester: Probab. Statist. Group, 2008.
- [15] Graverson S. E., Shiryaev A. N. An extension of P. Lévy's distributional properties to the case of a Brownian motion with drift // Bernoulli. 2000. **6**, № 4. P. 615–620.
- [16] Blanchet-Scalliet Ch., Diop A., Gibson R., Talay D., Tanré E. Technical analysis compared to mathematical models based methods under parameters mis-specification // J. of Banking and Finance. 2007. **31**. P. 1351–1373.
- [17] Gapeev P. V., Peskir G. The Wiener disorder problem with finite horizon // Stochastic Processes and Appl. 2006. **116**. P. 1770–1791.
- [18] Beibel M., Lerche H. R. A new look at optimal stopping problems related to mathematical finance // Statistica Sinica. 1997. **7**, № 1. P. 93–108.
- [19] Novikov A. A., Shiryaev A. N. On a solution of the optimal stopping problem for processes with independent increments // Stochastics. 2007. **79**, № 3–4. P. 393–406.
- [20] Новиков А. А. Несколько замечаний о распределении времени первого выхода и оптимальной остановке AR(1)-последовательностей // Теория вероятностей и ее применения. 2008. **53**, № 3. С. 458–471.

ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ ДЛЯ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ В УСЛОВИЯХ ВЫСОКОЙ ЗАГРУЗКИ

Л. Г. Афанасьева, Е. Е. Баштова

Статья посвящена асимптотическому анализу бесконечноканальных и одноканальных систем обслуживания. В первой части рассматриваются системы с большим числом каналов и входящим потоком высокой интенсивности, возникающей вследствие сжатия времени. Доказывается теорема о гауссовской аппроксимации, условия которой выражены в терминах моментов входящего потока. Во второй части решается задача о сходимости к диффузионному процессу характеристик одноканальных систем с регенерирующим входным потоком.

1. Введение

Значительная часть исследований по теории массового обслуживания проводится в предположении, что входящий поток пуассоновский. Это объясняется не только относительной простотой соответствующих математических моделей, но и адекватностью данного предположения для широкого круга прикладных задач. Дело в том, что наблюдаемые потоки часто являются суммами большого числа независимых потоков, а они, как показано в [1], в достаточно общих предположениях хорошо аппроксимируются пуассоновским процессом. Кроме того, при описании реальных объектов возникают модели, в которых первоначальный поток, проходя ряд последовательных систем, теряет часть своих требований. Как следует из [2] и других работ, такой разреженный поток также будет приближаться к пуассоновскому.

Взросший в последнее время интерес к более сложным входящим потокам в первую очередь связан с существованием практически важных задач, для которых предположение о том, что входящий поток пуассоновский, приводит к искажению операционных характеристик системы, например занижению среднего числа находящихся в ней требований.

Получение явного вида характеристик систем обслуживания со сложными потоками возможно лишь для узкого класса моделей, поэтому исследование, как правило, ведется в двух направлениях. Первое — построение оценок для этих характеристик. Второе — изучение условий высокой загрузки, когда в системе скапливается большое число требований, и малой, когда прибор значительную долю времени простаивает.

Данная статья посвящена анализу систем обслуживания с достаточно общими входящими потоками в асимптотике сжатия времени, когда число требований в системе растет.

В первой части изучаются модели с бесконечным или достаточно быстро растущим числом каналов. Это направление представлено в монографии А.А. Боровкова [3], где развита теория предельного поведения таких систем в достаточно общих предположениях относительно входящего потока, времен обслуживания и структуры системы. Входящий поток после некоторой нормировки предполагается сходящимся в том или ином смысле к случайному процессу $\xi(t)$. Тогда нормированное

число требований $q(t)$, находящихся в системе, в определенных условиях сходится к процессу $\zeta(t)$, связанному с $\xi(t)$. Применение этих результатов к системам с входными потоками, обычно используемыми в теории массового обслуживания (такими, как дважды стохастический пуассоновский, марковски-модулированный, регенерирующий, полумарковский и др.), сталкивается со значительными трудностями. Во-первых, необходимо найти условия сходимости входного потока к некоторому предельному процессу, во-вторых, выразить нормирующие коэффициенты в терминах параметров, определяющих входной поток и время обслуживания, и, наконец, предложить статистическую процедуру оценки этих параметров.

В настоящей работе устанавливаются достаточные условия сходимости одномерных распределений нормированного специальным образом процесса $q(t)$ к гауссовскому в терминах моментов входного потока (математического ожидания, корреляционной функции и функции, определяющей третий смешанный момент). Нормирующие коэффициенты также выражены через эти характеристики. Изучается асимптотика сжатия времени, что для бесконечноканальных систем принципиально, поскольку результаты верны и для асимптотики растущего времени обслуживания [4]. В качестве иллюстрации рассматриваются системы с дважды стохастическим пуассоновским и регенерирующим входными потоками.

Для одноканальных систем обслуживания сжатие времени производится таким образом, что коэффициент загрузки стремится к единице. Изучаются системы с регенерирующим входным потоком. Устанавливается C -сходимость виртуального времени ожидания к процессу диффузии с отражением в нуле. Доказательство опирается на теорему А.А. Боровкова [3], сформулированную в весьма общих предположениях. Однако, как и для бесконечноканальных систем, эта общность делает достаточно затруднительным применение полученных результатов к конкретным моделям, поскольку приходится проверять большое число условий и интерпретировать их в терминах, удобных для приложений.

В статье условия сходимости и нормирующие коэффициенты находятся с помощью моментов периода регенерации и числа требований, поступивших за этот период. Сравнение коэффициентов диффузии предельных процессов в асимптотике сжатия времени и в рассмотренной в [5] асимптотике растущей интенсивности для дважды стохастического пуассоновского входящего потока показывает, что во втором случае этот коэффициент больше, чем в первом.

2. Бесконечноканальные системы

Описание модели. Входящий в систему массового обслуживания поток $\{X(t), t \in \mathbb{R}\}$ — целочисленный случайный процесс с неубывающими траекториями, единичными скачками и $X(0) = 0$. Приращения $X(s, t) = X(t) - X(s) (s < t)$ этого процесса определяют количество требований, поступивших в систему на интервале (s, t) . Моменты скачков процесса $X(t)$ образуют возрастающую последовательность $\{t_j\}_{j=-\infty}^{+\infty}$, в которой $t_0 \leq 0, t_1 > 0$.

Времена обслуживания $\{\beta_i\}_{i=-\infty}^{+\infty}$ — независимые, одинаково распределенные случайные величины, не зависящие от входящего потока, с функцией распределения $B(x)$ и средним $b = \int_0^{\infty} \bar{B}(x) dx < \infty$ (здесь $\bar{B}(x) = 1 - B(x)$ и $B(0) = 0$), причем $P(\beta_j = b) < 1$.

В предположении, что система начала работу в бесконечно далеком прошлом, изучается асимптотическое поведение процесса $q(t)$, представляющего собой количество требований в системе в момент t . Рассматривается асимптотика сжатия времени, при которой входящий поток определяется соотношением

$$X_T(t) = X(Tt) \quad \text{при} \quad T \rightarrow \infty.$$

Предположим, что $E|X(t)|^3 < \infty$, и введем функции $H(t) = EX(t)$, $R(t, s) = E\hat{X}(t)\hat{X}(s)$, $G(t, y, u) = E\hat{X}(t)\hat{X}(y)\hat{X}(u)$, где $\hat{X}(t) = X(t) - H(t)$. Будем считать выполненным следующее условие:

$$\int_0^\infty \int_0^\infty |R(s, t)| dB(t) dB(s) < \infty, \quad (1)$$

которое обеспечивает существование в среднеквадратическом всех вводимых далее интегралов. Обозначим через $q_T(t)$ процесс $q(t)$ в системе с входящим потоком $X_T(t)$. Следующая теорема показывает, что для каждого t случайная величина $q_T(t)$ асимптотически нормальна при $T \rightarrow \infty$.

Основная теорема для бесконечноканальных систем. Нами доказана следующая

Теорема 1. Пусть выполнены условия:

- 1) $\sqrt{|t|} \left(\frac{H(t)}{t} - \lambda \right) \rightarrow 0$ при $|t| \rightarrow \infty$ для некоторого $\lambda \in (0, \infty)$;
- 2) существует $\lim_{T \rightarrow \infty} T^{-1}R(Tt, Ts) = g(t, s)$ и $g(t, s)$ удовлетворяет (1);
- 3) $\lim_{T \rightarrow \infty} T^{-3/2}G(Tt, Tu, Tv) = 0$.

Тогда для каждого t распределение

$$\hat{q}_T(t) = \frac{q_T(t) - \lambda bT}{\sqrt{T}\sigma_q(t)}$$

слабо сходится к нормальному с параметрами $(0, 1)$. При этом

$$\sigma_q^2(t) = \lambda b - \lambda \int_0^\infty [\bar{B}(y)]^2 dy + u(t),$$

где

$$u(t) = g(t, t) - 2 \int_0^\infty g(t, t-y) dB(y) + 2 \int_0^\infty \int_0^y g(t-y, t-s) dB(s) dB(y). \quad (2)$$

Доказательство. Рассмотрим производящую функцию

$$P(z, t) = Ez^{q(t)} = E \prod_{t_j \leq t} (1 + (z-1)\bar{B}(t-t_j)) \equiv Ef(z), \quad |z| < 1. \quad (3)$$

Введем следующие интегралы:

$$J_k^T(t) = \int_{-\infty}^t (\bar{B}(t-y))^k dX_T(y) = X(tT) - k \int_0^{\infty} X(T(t-y))(\bar{B}(y))^{k-1} dB(y), \quad k = 1, 2, 3. \quad (4)$$

Сначала докажем вспомогательное утверждение.

Лемма 1. Пусть при $T \rightarrow \infty$ выполняются условия:

- 1) $\sqrt{T}(T^{-1}\mathbf{E}J_1^T(t) - \lambda b) \rightarrow 0$;
- 2) $T^{-1}\mathbf{D}J_1^T(t) \rightarrow \sigma_J^2(t) > 0$;
- 3) $T^{-1}\mathbf{E}J_2^T(t) \rightarrow a_2$;
- 4) (а) $T^{-3/2}(\mathbf{E}(J_1^T(t) - \mathbf{E}J_1^T(t))^3) \rightarrow 0$, (б) $T^{-3/2}\text{cov}(J_1^T(t), J_2^T(t)) \rightarrow 0$.

Тогда при каждом t распределение

$$\hat{q}_T(t) = \frac{q_T(t) - \lambda bT}{\sqrt{T}\sigma_q(t)}$$

слабо сходится к нормальному с параметрами $(0, 1)$. При этом

$$\sigma_q^2(t) = \lambda b + \sigma_J^2(t) - a_2. \quad (5)$$

Доказательство. Поскольку t фиксировано, в некоторых записях мы его опустим, например напомним J_k^T вместо $J_k^T(t)$. Из определения производящей функции (3) находим, что

$$P_T(e^{is/\sqrt{T}}, t) = \mathbf{E}e^{isq_T/\sqrt{T}} = 1 - (1 - e^{is/\sqrt{T}})\mathbf{E}J_1^T + \\ + \frac{(1 - e^{is/\sqrt{T}})^2}{2}(\mathbf{E}(J_1^T)^2 - \mathbf{E}J_2^T) - \frac{(1 - e^{is/\sqrt{T}})^3}{3!}\mathbf{E}f_T'''(\theta e^{is/\sqrt{T}}), \quad |\theta| \leq 1,$$

где

$$f_T'''(z) = \sum_{t_j, t_k, t_m \leq Tt} P_j^T P_k^T P_m^T \mathbf{1}\{A_{j,k,m}\} \prod_{t_s \leq Tt} (1 + (z-1)P_s^T) \mathbf{1}\{s \neq j, k, m\}$$

и $P_j^T = \bar{B}(t - t_j/T)$. В записи $A_{j,k,m}$ имеется в виду, что все индексы j, k, m различны.

В результате несложных, но довольно длинных вычислений получается оценка

$$\mathbf{E}f_T'''(\theta e^{is/\sqrt{T}}) = (\mathbf{E}(J_1^T)^3 - \mathbf{E}J_1^T J_2^T + 2\mathbf{E}J_3^T)(1 + \varepsilon(\theta s/\sqrt{T})), \quad (6)$$

где $\varepsilon(s) \rightarrow 0$ при $s \rightarrow 0$.

Теперь запишем разложение функции

$$\varphi_T(s) = \ln \mathbf{E}e^{is \frac{q_T - \lambda bT}{\sqrt{T}}}$$

следующим образом:

$$-is\lambda b\sqrt{T} + is \frac{H_T(t)}{\sqrt{T}} - \frac{s^2}{2T}(\mathbf{D}J_1^T - \mathbf{E}J_2^T + \mathbf{E}J_1^T) - \frac{s^3}{T^{3/2}}r_T(\theta s/\sqrt{T}). \quad (7)$$

Оценка $r_T(\theta s/\sqrt{T})$ основана на соотношении (6) и неравенствах

$$|1 - e^{-\beta} + \beta| \leq \frac{|\beta|^2}{2}, \quad |1 - e^{-\beta} + \beta - \frac{\beta^2}{2}| \leq \frac{|\beta|^3}{6} \quad (\operatorname{Re} \beta \geq 0).$$

Приведем здесь окончательный результат:

$$|r_T(s/\sqrt{T})| \leq C(|\mathbf{E}(J_1^T - \mathbf{E}J_1^T)^3| + 3|\operatorname{cov}(J_1^T, J_2^T) + 2\mathbf{E}J_3^T|)(1 + \delta_T(s/\sqrt{T})), \quad (8)$$

где $\delta_T(s/\sqrt{T}) \rightarrow 0$ при $T \rightarrow \infty$. Из (7), (8) и условий леммы 1 получаем

$$\varphi_T(s) \rightarrow -\frac{s^2}{2\sigma_q^2(t)} \quad (T \rightarrow \infty),$$

где $\sigma_q^2(t)$ задается формулой (5).

Следующие три леммы завершают доказательство теоремы 1.

Лемма 2. При выполнении условия 1 теоремы 1

$$\lim_{T \rightarrow \infty} \sqrt{T} (T^{-1} \mathbf{E}J_1^T - \lambda b) = 0; \quad (9)$$

$$\lim_{T \rightarrow \infty} T^{-1} \mathbf{E}J_2^T = \lambda \int_0^\infty [\bar{B}(y)]^2 dy. \quad (10)$$

Доказательство. Справедливость утверждения следует из неравенства

$$\sqrt{T} \left| \frac{\mathbf{E}J_1^T}{T} - \lambda b \right| \leq \int_0^\infty y \sqrt{T} |T^{-1} H(-yT) - \lambda| dB(y)$$

и формулы (4) для J_2^T .

Лемма 3. При выполнении условия 2 теоремы 1

$$\lim_{T \rightarrow \infty} T^{-1} \mathbf{D}J_1^T(t) = u(t),$$

где $u(t)$ задается соотношением (2).

Доказательство. Из (4) и условия 2 теоремы 1 имеем при $T \rightarrow \infty$

$$\begin{aligned} T^{-1} \mathbf{D}J_1^T &= T^{-1} [R(Tt, Tt) - 2 \int_0^\infty R(Tt, T(t-y)) dB(y) + \\ &+ 2 \int_0^\infty \int_0^y R(T(t-s), T(t-y)) dB(s) dB(y)] \rightarrow u(t). \end{aligned}$$

Лемма 4. Если выполнены условия 1–3 теоремы 1, то

$$T^{-3/2} r_T(s/\sqrt{T}) \rightarrow 0 \quad (T \rightarrow \infty).$$

Доказательство. Учитывая (9), (10) и (8), достаточно доказать, что

$$\lim_{T \rightarrow \infty} T^{-3/2} \mathbb{E} (J_1^T - \mathbb{E} J_1^T)^3 = 0. \quad (11)$$

Поскольку

$$\begin{aligned} \mathbb{E} (J_1^T - \mathbb{E} J_1^T)^3 &= G(Tt, Tt, Tt) - 3 \int_0^\infty G(Tt, Tt, T(t-y)) dB(y) + \\ &+ 6 \int_0^\infty \int_0^y G(Tt, T(t-y), T(t-s)) dB(s) dB(y) + \\ &- 6 \int_0^\infty \int_0^y \int_0^u G(T(t-y), T(t-v), T(t-u)) dB(u) dB(v) dB(y), \end{aligned}$$

то (11) следует из условия 3) теоремы 1.

Следствие 1. Если выполняются условия теоремы 1 и

$$g(t, s) = (\sigma_X^2 \min(|t|, |s|) + C) \mathbf{1}\{ts \geq 0\}, \quad (12)$$

то $\hat{q}_T(t)$ асимптотически нормальна с

$$\sigma_q^2 = \lambda b + (\sigma_X^2 - \lambda) \int_0^\infty (\bar{B}(y))^2 dy. \quad (13)$$

Доказательство. Достаточно вычислить $u(t)$, используя (2) и (12).

Предельная теорема для систем с регенерирующим входящим потоком. Начнем с определения регенерирующего потока.

Будем говорить, что входящий процесс $\{X(t), t \in (-\infty, +\infty)\}$ регенерирующий, если существует неубывающая последовательность случайных величин $\{\theta_j\}_{j=-\infty}^{+\infty}$ ($\theta_0 = 0$), таких, что

$$\{\theta_{j+1} - \theta_j, X(\theta_j + t) - X(\theta_j), \quad t \in [0, \theta_{j+1} - \theta_j]\}_{j=-\infty}^{+\infty}$$

— последовательность независимых, одинаково распределенных случайных элементов. Как обычно, введем последовательности случайных величин

$$\tau_j = \theta_{j+1} - \theta_j, \quad \varkappa_j(t) = X(\theta_j + t) - X(\theta_j), \quad \xi_j = \varkappa_j(\tau_j)$$

и обозначим

$$a = \mathbb{E} \xi_i, \quad \mu = \mathbb{E} \tau_i, \quad \sigma_\xi^2 = \mathbb{D} \xi_i, \quad \sigma_\tau^2 = \mathbb{D} \tau_i, \quad r_{\xi, \tau} = \text{cov}(\xi_i, \tau_i).$$

Этот класс процессов довольно широк. В частности, в него входят рекуррентные, марковски-модулированные, марковские, полумарковские и другие входящие

процессы. В дальнейшем будем считать, что распределение τ_i имеет абсолютно непрерывную компоненту.

Теорема 2. Пусть $\{X(t), t \in (-\infty, +\infty)\}$ — регенерирующий процесс и $E\xi_j^4 < \infty$, $E\tau_j^4 < \infty$. Тогда распределение $\hat{q}_T(t)$ слабо сходится к нормальному с параметрами $(0, 1)$ и σ_q^2 задается формулой (13), где

$$\lambda = \frac{a}{\mu}, \quad \sigma_X^2 = \frac{\sigma_\xi^2}{\mu} + \frac{a^2 \sigma_\tau^2}{\mu^3} - \frac{2a r_{\xi, \tau}}{\mu^2}. \quad (14)$$

Доказательство. Достаточно проверить условия 1–3 теоремы 1 для $\{X(t), t \geq 0\}$. Рассмотрим процесс

$$N(t) = \max\{j \geq 0: \theta_j \leq t\}, \quad \gamma_t = t - \theta_{N(t)}.$$

Процесс $X_T(t)$ можно представить следующим образом:

$$X_T(t) = Z(Tt) + \kappa_{N(Tt)}(\gamma_{Tt}), \quad (15)$$

где $Z(u) = \sum_{j=0}^{N(u)} \xi_j$. Известно (см., например, [6]), что случайные величины γ_t , $\kappa_{N(t)}(\gamma_t)$ и $\tau_{N(t)}$ имеют предельное распределение при $t \rightarrow \infty$. Если $F(x, y) = P(\xi_i \leq x, \tau_i \leq y)$, то

$$\lim_{t \rightarrow \infty} P(\xi_{N(t)} \leq x) = \mu^{-1} \int_0^\infty y F(x, dy) = F(x). \quad (16)$$

Поскольку $\kappa_{N(t)}(\gamma_t) \leq \xi_{N(t)}$, то из (15) получаем

$$EZ(t) \leq H(t) = EX(t) \leq EZ(t) + E\xi_{N(t)}.$$

Теперь справедливость условия 1 теоремы 1 следует из асимптотических результатов для $EZ(t)$ (см. [6]) и (16). Для оценки функции корреляции запишем

$$\begin{aligned} R(Tt, Ts) &= \text{cov}(Z(Ts), Z(Tt)) + \text{cov}(Z(Tt), \kappa_{N(Ts)}(\gamma_{Ts})) + \text{cov}(Z(Ts), \kappa_{N(Tt)}(\gamma_{Tt})) + \\ &+ \text{cov}(\kappa_{N(Ts)}(\gamma_{Ts}), \kappa_{N(Tt)}(\gamma_{Tt})) = g_1 + g_2 + g_3 + g_4. \end{aligned}$$

Так как $\text{cov}(Z(Ts), Z(Tt) - Z(Ts)) \leq C$ ($t \geq s$), то

$$\lim_{T \rightarrow \infty} T^{-1} \text{cov}(Z(Ts), Z(Tt)) = s \sigma_X^2 \quad (t \geq s \geq 0),$$

где σ_X^2 определяется (14) (см. [6]), C — некоторая константа. Кроме того,

$$|g_2| = |\text{cov}(\xi_{N(Ts)}, \kappa_{N(Ts)}(\gamma_{Ts}))| \leq C_2 < \infty$$

и аналогично $|g_3| \leq C_3 < \infty$, $|g_4| \leq C_4 < \infty$. Значит, условие 2 выполнено и $g(t, s)$ задается (12).

Что касается условия 3, заметим, что

$$G(y, u, v) = G_z(y, u, v) + d(y, u, v) + d(u, y, v) + d(v, y, u) + E\hat{\eta}_u \hat{\eta}_v \hat{\eta}_y,$$

где $G_z(y, u, v)$ — функция $G(y, u, v)$ для процесса $Z(t)$ и

$$\hat{\eta}_y = \varkappa_{N(y)}(\gamma_y) - \mathbf{E}\varkappa_{N(y)}(\gamma_y), \quad \hat{Z}(y) = Z(y) - \mathbf{E}Z(y),$$

$$d(y, u, v) = \mathbf{E}[\hat{\eta}_y \hat{Z}(u) \hat{Z}(v) + \hat{\eta}_y \hat{\eta}_v \hat{Z}(v)].$$

Таким образом, имеем

$$|\mathbf{E}\hat{\eta}_y \hat{Z}(u) \hat{Z}(v)| \leq \sqrt{\mathbf{E}\hat{\eta}_y^2} \sqrt{\mathbf{E}(\hat{Z}(u))^4 \mathbf{E}(\hat{Z}(v))^4} \leq C_5 \sqrt{uv} + \tilde{C}_6,$$

и условие 3 следует из асимптотического поведения $\hat{Z}(t)$ при $t \rightarrow \infty$ (см. [6, 7]).

Система с дважды стохастическим пуассоновским процессом (ДСПП). Напомним определение ДСПП. Пусть $\{\Lambda(t), t \in (-\infty, +\infty)\}$ — случайный процесс с неубывающими, непрерывными слева и имеющими предел справа траекториями и значениями в $\mathbb{R}$. Пусть к тому же $\Lambda(0) = 0$. Дважды стохастический пуассоновский процесс определяется при помощи случайной замены времени (см. [8])

$$X(t) = A(\Lambda(t)),$$

где $\{A(t), t \in (-\infty, +\infty)\}$ — стандартный пуассоновский процесс, не зависящий от $\Lambda(t)$. В предположении, что $\mathbf{E}|\Lambda(t)|^3 < \infty$, введем функции

$$H_\Lambda(t) = \mathbf{E}\Lambda(t), \quad R_\Lambda(t, s) = \mathbf{E}\hat{\Lambda}(t)\hat{\Lambda}(s), \quad G_\Lambda(t, y, u) = \mathbf{E}\hat{\Lambda}(t)\hat{\Lambda}(y)\hat{\Lambda}(u).$$

Теорема 3. Пусть $\{X(t), t \in (-\infty, +\infty)\}$ — ДСПП с ведущей функцией $\Lambda(t)$. Если для $H_\Lambda(t)$, $R_\Lambda(t, s)$ и $G_\Lambda(t, y, u)$ выполняются условия 1–3 теоремы 1, то распределение $\hat{q}_T(t)$ при каждом t слабо сходится к нормальному. Нормирующий множитель задается формулой

$$\sigma^2(t) = \lambda b + u_\Lambda(t) \int_0^\infty [\bar{B}(y)]^2 dy,$$

где $\lambda = \lim_{t \rightarrow \infty} \frac{H_\Lambda(t)}{t}$ и $u_\Lambda(t)$ — функция $u(t)$ для процесса $\Lambda(t)$.

Доказательство. Теорема 3 следует из теоремы 1, так как существуют явные соотношения между $H_\Lambda(t)$, $R_\Lambda(t, s)$, $G_\Lambda(t, y, u)$ для $\Lambda(t)$ и соответствующими функциями для процесса $X(t)$. Например, $H_\Lambda(t) = H(t)$ и

$$R(t, s) = R_\Lambda(t, s) + H(\min(|s|, |t|)) \mathbf{1}\{ts \geq 0\}.$$

Выражение для $G(t, y, u)$ более сложное и здесь не приводится.

Далее будем предполагать, что

$$\Lambda(t) = \int_0^t \lambda(y) dy,$$

где $\lambda(y)$ — неотрицательный, локально интегрируемый, стационарный в широком смысле случайный процесс со средним λ . Более того, считаем, что $P(\lambda(t) \leq \lambda^*, t \in (-\infty, +\infty)) = 1$, $\lambda^* < \infty$.

Обозначим $\lambda = \mathbf{E}\lambda(y)$, $\hat{\lambda}(y) = \lambda(y) - \lambda$ и $r(y) = \mathbf{E}\hat{\lambda}(0)\hat{\lambda}(y)$.

Следствие 2. Если при $T \rightarrow \infty$

$$|r(T)| \leq C|T|^{-\alpha} \quad \text{для} \quad \alpha > 2, \quad 0 < C < \infty, \quad (17)$$

а функция $G_\Lambda(t, s, y)$ удовлетворяет условию 3 теоремы 1, то $\hat{q}_T(t)$ асимптотически нормальна и

$$\sigma_q^2 = \lambda b + 2 \int_0^\infty (\bar{B}(y))^2 dy \int_0^\infty r(u) du. \quad (18)$$

Доказательство. Для доказательства этого утверждения используем лемму 1. Поскольку $H_\lambda(t) = \lambda t$ и условие 4а следует из условия, наложенного на $G_\Lambda(t, s, y)$, достаточно проверить выполнение условия 2. Для этого запишем

$$T^{-1} \mathbf{D} J_1^T(t) = 2T \int_{-\infty}^t \int_{-\infty}^y \bar{B}(t-u) \bar{B}(t-y) r(T(y-u)) du dy + T^{-1} \mathbf{E} J_2^T.$$

Первое слагаемое представим в виде

$$\begin{aligned} & 2T \int_{-\infty}^t \bar{B}(t-y) \int_{y-T^{-\alpha}}^y \bar{B}(t-u) r(T(y-u)) du dy + \\ & + 2T \int_{-\infty}^t \bar{B}(t-y) \int_{-\infty}^{y-T^{-\alpha}} \bar{B}(t-u) r(T(y-u)) du dy = I_1^\alpha + I_2^\alpha. \end{aligned}$$

При $\alpha \in (0, 1)$ в силу условия (17) $I_2^\alpha \rightarrow 0$ ($T \rightarrow \infty$), а для I_1^α получаем

$$I_1^\alpha = 2T \int_{-\infty}^t \bar{B}(t-y) \bar{B}(t-y-\theta T^{-\alpha}) \int_{y-T^{-\alpha}}^y r(T(y-u)) du dy,$$

где $|\theta| < 1$. Замена переменных $v = T(y-u)$ и переход к пределу при $T \rightarrow \infty$ доказывают условие 2 леммы 1 и соотношение (18).

В качестве примера рассмотрим марковски-модулированный процесс (см. например, [9]), для которого

$$\lambda(t) = \sum_{k=1}^{\infty} \lambda_k \mathbf{1}\{U(t) = k\}, \quad (19)$$

где $U(t)$ — однородная цепь Маркова с конечным или счетным числом состояний, а $\{\lambda_k, k = 0, 1, \dots\}$ — набор неотрицательных чисел, таких, что $\lambda_k \leq \lambda^*$ для всех k . Обозначим $P_{ij}(t) = P(U(t) = j | U(0) = i)$ для $t \geq 0$.

Если цепь $U(t)$ стационарна и π_j — ее стационарное распределение, то корреляционная функция стационарного процесса $\lambda(t)$ задается соотношением

$$r(y) = \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \lambda_k \lambda_j \pi_k (P_{kj}(y) - \pi_j).$$

Для выполнения (17) достаточно потребовать, чтобы при $t \rightarrow \infty$

$$|P_{ij}(t) - \pi_j| \leq \frac{C\pi_j}{t^\alpha}, \quad \alpha > 2, \quad 0 < C < \infty.$$

Нормирующий множитель, как следует из (18), задается соотношением

$$\sigma_q^2 = \lambda b + 2 \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} \lambda_k \lambda_j \pi_k \int_0^{\infty} (P_{kj}(y) - \pi_j) dy \int_0^{\infty} (\bar{B}(u))^2 du. \quad (20)$$

Рассмотрим теперь систему с потерями, и пусть n_T — количество обслуживающих приборов в ней. Обозначим через $q_T^0(t)$ количество требований в системе в момент t . Справедливо следующее утверждение.

Теорема 4. Пусть выполнены условия 1–3 теоремы 1 и

$$\frac{n_T - mT}{\sqrt{T}} \rightarrow \infty \quad (T \rightarrow \infty) \quad (21)$$

для некоторого $m \in (0, \infty)$. Тогда

$$\hat{q}_T^0(t) = \frac{q_T^0(t) - \lambda b T}{\sqrt{T} \sigma_q(t)}$$

асимптотически нормальна с параметрами $(0, 1)$.

Доказательство следует из того факта, что $q_T(t)$ и $q_T^0(t)$ асимптотически эквивалентны при выполнении условия (21).

3. Одноканальные системы. Диффузионная аппроксимация в условиях высокой загрузки

Описание модели. В этой части мы установим диффузионную аппроксимацию одноканальной системы обслуживания с регенерирующим входным потоком в условиях высокой загрузки. В отличие от бесконечноканальных систем, где обслуживание начинается сразу в момент поступления, здесь требования становятся в очередь и главными характеристиками системы являются длина очереди и время ожидания. Существенную роль в их поведении играет параметр, называемый коэффициентом загрузки:

$$\rho = b \lim_{t \rightarrow \infty} \frac{EX(t)}{t} = \lambda b, \quad (22)$$

который служит показателем стохастической ограниченности процессов, описывающих функционирование системы. Высокой загрузкой естественно считать не только такой режим работы, когда $\rho > 1$ (как рассматривалось в [10]), но и такой, при котором ρ стремится к 1, оставаясь меньше ее.

Пусть $X(t)$ — регенерирующий поток, описанный в разделе 2. Рассмотрим серию систем $\mathcal{S}_T$ на интервалах времени $[0, \nu T]$ ($\nu > 0, T \rightarrow \infty$), в которых входящий поток определяется соотношением

$$X_T(t) = X\left(\frac{\sqrt{T} - 1}{\rho\sqrt{T}}t\right). \quad (23)$$

При этом, как следует из (22),

$$\rho_T = 1 - \frac{1}{\sqrt{T}} \xrightarrow{T \rightarrow \infty} 1.$$

Введем обозначения:

$N_T(t)$ и $M(t)$ — процессы восстановления, образованные соответственно последовательностями $\{\tau_i^T\}$ и $\{\beta_i\}$;

$q_T(t)$ и $W_T(t)$ — соответственно количество требований и процесс виртуального времени ожидания в системе $\mathcal{S}_T$.

Отметим, что, согласно, например, [11],

$$W(t) = W(0) + Z(t) - \inf\{Z(s) : 0 \leq s \leq t\}, \quad (24)$$

где

$$Z(t) = Y(t) - t, \text{ а } Y(t) = \beta_1 + \beta_2 + \dots + \beta_{X(t)}.$$

Основная теорема для одноканальных систем. Справедлива следующая

Теорема 5. Пусть $E\beta_k^{2+\delta} < \infty$, $E\tau_k^{2+\delta} < \infty$, $E\xi_k^{2+\delta} < \infty$ для некоторого $\delta > 0$ и $W(0) = 0$. Тогда на каждом конечном отрезке $[0, \nu]$ для процесса $\widehat{W}_T(t) = W_T(Tt)/\sqrt{T}$ имеет место C -сходимость к диффузионному процессу с отражением от нулевой границы и коэффициентами $(-1, \sigma_W^2)$. При этом

$$\sigma_W^2 = \frac{1}{b}\sigma_\beta^2 + \frac{b}{\lambda}\sigma_X^2, \quad (25)$$

где σ_X^2 определяется (14), а $\sigma_\beta^2 = D\beta$.

Доказательство. Первый шаг состоит в построении двух процессов, которые оценивают процесс $W(t)$ сверху и снизу, и разность между которыми стремится к нулю при $\rho \rightarrow 1$. Далее устанавливается, что в сделанных предположениях оценивающие процессы удовлетворяют условиям теоремы, доказанной А.А. Боровковым [3, с. 55–56; 77–78].

Рассмотрим систему $\mathcal{S}^-$, в которой в моменты θ_n поступают партии требований объема $\xi_n = X(\theta_{n+1}) - X(\theta_n)$. Времена обслуживания требований — независимые, одинаково распределенные случайные величины с функцией распределения $B(x)$. Обозначим через $W^-(t)$ виртуальное время ожидания в системе $\mathcal{S}^-$.

Первое утверждение устанавливает асимптотическую эквивалентность нормированных процессов времени ожидания в системах $\mathcal{S}_T$ и $\mathcal{S}_T^-$.

Лемма 5. Для $t \in [0, \nu]$

$$\left| \frac{W_T^-(Tt) - W_T(Tt)}{\sqrt{T}} \right| \xrightarrow{C} 0 \quad \text{при } T \rightarrow \infty$$

Доказательство. Как показано в [5], процессы $W(t)$ и $W^-(t)$ связаны между собой следующими неравенствами:

$$W^-(t) - \gamma_n \leq W(t) \leq W^-(t) + \tau_{k(t)},$$

где $t \in [\theta_{n-1}, \theta_n]$ и $k(t)$ — номер последнего (до t) из интервалов τ_j , на которых было освобождение системы, а $\gamma_n = Y(\theta_{n+1}) - Y(\theta_n)$. Поэтому можно записать

$$\left| \frac{W_T^-(Tt) - W_T(Tt)}{\sqrt{T}} \right| \stackrel{\text{н.н.}}{\leq} \frac{\gamma_{N_T(tT)} + \tau_{k(tT)}}{\sqrt{T}} \equiv \frac{r_T(tT)}{\sqrt{T}}.$$

Согласно критерию С-сходимости (см. [12]), для доказательства леммы 5 достаточно проверить, что для любого $\alpha > 0$

$$P\left(\sup_{t \leq \nu} |r_T(tT)/\sqrt{T}| > \alpha\right) \rightarrow 0 \text{ при } T \rightarrow \infty. \quad (26)$$

Зафиксируем $\Delta > 0$. В соответствии с законом больших чисел для процессов восстановления (см. [13]) можно выбрать s , такое, что

$$P(|N_T(\nu T) - \nu T/\mu| > s\nu T) < \Delta/2.$$

Далее,

$$\begin{aligned} P(\sup_{t \leq \nu} |r_T(tT)/\sqrt{T}| > \alpha) &= P(\max_{i \leq N(\nu T)+1} |r_T(\theta_i)/\sqrt{T}| > \alpha) \leq \\ &\leq \frac{\nu T(1 + s\mu)}{\mu} \frac{\mathbf{E}(r_T(\theta_i))^{2+\delta}}{T^{1+\delta/2}\alpha^{2+\delta}} + \Delta/2 = C/T^{\delta/2} + \Delta/2, \end{aligned}$$

откуда следует (26).

Теперь достаточно доказать диффузионную аппроксимацию для процесса $W_T^-(t)$. Несмотря на наличие формулы (24), представляющей процесс $W^-(t)$ в виде граничного функционала на траекториях процесса $X^-(t)$, с точки зрения получения диффузионной аппроксимации более удобным оказывается процесс $q^-(t)$.

Лемма 6. В условиях теоремы 5 для процесса $\hat{q}_T(t) = q_T(Tt)/\sqrt{T}$ имеет место С-сходимость на каждом конечном отрезке $[0, \nu]$ к диффузионному процессу с отражением от нулевой границы и коэффициентами $(-1/b, \sigma_W^2/b^2)$, где σ_W^2 определяется формулой (25).

Доказательство состоит в проверке условий теоремы о диффузионной аппроксимации из [3] и содержит довольно много технических моментов. Для дважды стохастического пуассоновского входящего потока с растущей (при $T \rightarrow \infty$) интенсивностью полное доказательство имеется в [5]. Здесь мы ограничимся изложением идеи доказательства. Она состоит в том, что в ситуации высокой загрузки система почти всегда работает в режиме полной занятости, то есть без свободных периодов, и тогда процесс $q_T^-(t)$ представляет собой разность процессов $X_T^-(t)$ и $M(t)$. Таким образом, появляется возможность использовать результаты о диффузионной аппроксимации $X_T^-(t)$ и $M(t)$. Эти процессы являются процессами накопления, (а $M(t)$ — классический процесс восстановления), и поэтому можно утверждать [3], что процессы $\hat{X}_T^-(t) = \frac{X_T^-(tT) - \lambda Tt}{\sqrt{T}}$ и $\hat{M}_T^-(t) = \frac{M(tT) - b^{-1}Tt}{\sqrt{T}}$ на каждом конечном интервале С-сходятся к процессам неограниченной диффузии с коэффициентами соответственно $(0, \frac{b}{\lambda}\sigma_X^2)$ и $(0, \sigma_M^2)$, где σ_X^2 определяется выражением (14), а $\sigma_M^2 = \sigma_\beta^2/b^3$.

При проверке условий упомянутой теоремы для процесса $\hat{q}_T(t)$ оказывается, что коэффициент сноса равен $\sqrt{T}(\lim_{t \rightarrow \infty} \mathbf{E}X_T(t)/t - 1/b) = -1/b$, а коэффициент диффузии равен сумме $\frac{1}{\rho}\sigma_X^2 + \sigma_M^2$.

Следующая лемма дает возможность перейти от длины очереди в момент t к процессу виртуального времени ожидания.

Лемма 7. Для любого $\alpha > 0$

$$P\left(\sup_{0 \leq t \leq \nu} \left| \frac{W_T(Tt) - b q_T(Tt)}{\sqrt{T}} \right| > \alpha\right) \rightarrow 0 \text{ при } T \rightarrow \infty. \quad (27)$$

Доказательство. Зафиксируем некоторое $d < \alpha/6b$. Поскольку для всякого t

$$W(t) = \tilde{\beta}_0(t) + \beta_1 + \dots + \beta_{q(t)},$$

где $\tilde{\beta}_0(t)$ — остаточное время обслуживания требования, занимающего прибор в момент t , то выполнено следующее неравенство:

$$\begin{aligned} P\left(\sup_{0 \leq t \leq \nu} |W_T(tT) - b q_T(tT)| > \alpha\sqrt{T}\right) &\leq P\left(\sup_{0 \leq t \leq \nu} \tilde{\beta}_0(tT) > \alpha\sqrt{T}/3\right) + \\ &+ P\left(\sup_{0 \leq t \leq \nu} |\beta_1 + \dots + \beta_{q_T(tT)} - b q_T(tT)| \mathbf{1}_{q_T(tT) \leq d\sqrt{T}} > \alpha\sqrt{T}/3\right) + \\ &+ P\left(\sup_{0 \leq t \leq \nu} |\beta_1 + \dots + \beta_{q_T(tT)} - b q_T(tT)| \mathbf{1}_{q_T(tT) > d\sqrt{T}} > \alpha\sqrt{T}/3\right) = \\ &= g_1(T, t) + g_2(T, t) + g_3(T, t). \end{aligned}$$

Здесь g_1 стремится к нулю, так как процесс $\tilde{\beta}_0(tT)$ стохастически ограничен при $T \rightarrow \infty$,

$$g_2(T, t) \leq P\left(\frac{\beta_1 + \dots + \beta_{[d\sqrt{T}]} - b}{d\sqrt{T}} > \alpha/3\delta - 2b\right) \rightarrow 0 \text{ при } T \rightarrow \infty$$

в силу закона больших чисел.

Для $g_3(T, t)$ имеем

$$\begin{aligned} g_3(T, t) &\leq P\left(\sup_{0 \leq t \leq \nu} q_T(tT) \sup_{k \geq d\sqrt{T}} \left| \frac{\beta_1 + \dots + \beta_k}{k} - b \right| > \alpha\sqrt{T}/3\right) \leq \\ &\leq P(T^{-\delta/(8+8\delta)} \sup_{0 \leq t \leq \nu} q_T(tT)/\sqrt{T} > \alpha/3) + P\left(\sup_{k \geq \delta\sqrt{T}} \left| \frac{\beta_1 + \dots + \beta_k}{k} - b \right| > T^{-\delta/(8+8\delta)}\right). \end{aligned}$$

Здесь снова первое слагаемое стремится к нулю в силу стохастической ограниченности процесса $\hat{q}_T(t)$, а для оценки второго используем результат работы [14]:

$$|B_n(x) - \Phi(x)| \leq C \frac{\mathbf{E}|\beta_i|^{2+\delta}}{n^{\delta/2}(1 + |x|^{2+\delta})},$$

где $B_n(x)$ — функция распределения нормированной суммы

$$\frac{\beta_1 + \dots + \beta_n - bn}{\sigma_\beta \sqrt{n}},$$

а $\Phi(x)$ — функция распределения стандартного нормального закона.

Применяя эту оценку скорости сходимости в центральной предельной теореме, получаем

$$P \left(\sup_{k \geq \delta \sqrt{T}} \left| \frac{\beta_1 + \dots + \beta_k}{k} - b \right| > T^{-\delta/(8+8\delta)} \right) \leq \\ \leq \sum_{k \geq \delta \sqrt{T}} \frac{C_1}{k^{1+\delta} T^{-\delta/4}} + \sum_{k \geq \delta \sqrt{T}} \frac{C_2 E|\xi_i|^{2+\delta}}{k^{\delta/2} (1 + k^{1+\delta/2} T^{-\delta/4})} \leq CT^{-\delta/4} \rightarrow 0 \text{ при } T \rightarrow \infty.$$

Таким образом, доказано соотношение (27).

Системы с дважды стохастическим пуассоновским входящим потоком. Предположим, что $X(t)$ — дважды стохастический пуассоновский процесс и его интенсивность $\lambda(t)$ — регенерирующий, локально интегрируемый случайный процесс. Тогда $X(t)$ будет регенерирующим потоком в смысле раздела 2: в качестве последовательности $\{\theta_n\}$ нужно взять последовательность точек регенерации процесса $\lambda(t)$.

При этом, согласно формуле (25),

$$\sigma^2 = \frac{b_2}{b} + \frac{b}{a} D(\Lambda(\tau) - \lambda\tau).$$

Отметим, что для ДСПП существуют различные способы построения асимптотики увеличения загрузки. Так, например, в [5] рассмотрена асимптотика увеличения интенсивности входящего потока, а именно последовательность $\mathcal{S}_T$ предполагалась такой, что

$$\lambda_T(t) = \frac{\sqrt{T} - 1}{\rho\sqrt{T}} \lambda(t).$$

Там же доказана теорема о сходимости $\widehat{W}_T(t)$ к диффузионному процессу с отражением от нулевой границы и коэффициентами сноса и диффузии $(-1, \sigma_1^2)$, где

$$\sigma_1^2 = \frac{b_2}{b} + \frac{\mu}{a^2} D(\Lambda(\tau) - \lambda\tau).$$

Таким образом, можно видеть, что в различных постановках задачи о высокой загрузке имеется сходимость к диффузионному процессу с отражением от нулевой границы, однако параметр диффузии предельного процесса зависит от способа увеличения загрузки. При этом (поскольку $b < \mu/a$) выполняются неравенства

$$\frac{b_2}{b} \leq \sigma^2 \leq \sigma_1^2,$$

в которых равенство достигается только для пуассоновского процесса с неслучайной (но, быть может, непостоянной) интенсивностью.

Использование доказанных теорем при асимптотическом анализе нагруженных систем обслуживания ставит задачу получения статистических оценок для нормирующих коэффициентов. Если входной поток регенерирующий, то для вычисления σ_q^2 и σ_W^2 в бесконечноканальной и одноканальной системах по формулам (13) и (25) соответственно необходимо найти $\sigma_X^2 = \lim_{t \rightarrow \infty} DX(t)/t$.

Это — традиционная задача математической статистики, когда наблюдаются оба процесса $X(t)$ и $N(t)$, поскольку для построения оценки можно использовать формулу (13). Если наблюдается только процесс $X(t)$, так что моменты регенераций неизвестны, возникает проблема разумного разделения данных на группы таким образом, чтобы, с одной стороны, последние имели малую корреляцию, а с другой — число этих групп было велико при достаточно большом объеме наблюдений. Некоторые из стохастических задач такого сорта рассмотрены в [15].

Работа выполнена при поддержке гранта РФФИ № 07-01-00362.

Список литературы

- [1] Григелионис Б. И. О сходимости сумм ступенчатых случайных процессов к пуассоновскому // Теория вероятностей и ее применения. 1963. **8**, № 2. С. 189–194.
- [2] Rényi A. Poisson-folyamat egy jellemzése // Тр. Матем. ин-та АН Венгрии. 1956. **1**, № 4. С. 519–529 (in Hungarian).
- [3] Боровков А. А. Асимптотические методы в теории массового обслуживания. М.: Наука, 1980.
- [4] Afanaseva L. G. Gaussian approximation for multichannel queueing systems // Proc. XIIth Int. Conf. on Applied Stochastic Models and Data Analysis. Greece, 2007. P. 1–8.
- [5] Афанасьева Л. Г., Баштова Е. Е. Предельные теоремы для систем обслуживания с дважды стохастическим пуассоновским потоком (условия высокой загрузки) // Проблемы передачи информации. 2008. **44**. № 4. С. 81–100.
- [6] Кокс Д. Р., Смит В. Л. Процессы восстановления. М.: Сов. радио, 1965.
- [7] Афанасьева Л. Г., Булинская Е. В. Случайные процессы в теории массового обслуживания и управления запасами. М.: Изд-во МГУ, 1980.
- [8] Grandell J. Doubly stochastic Poisson processes // Lect. Notes Math. 1976. **529**.
- [9] Asmussen S. Ladder heights and Markov-modulated $M|G|1$ queue // Stochastic Processes and Appl. 1991. **37**. P. 313–326.
- [10] Iglehart D. L., Whitt W. Multiple channel queues in heavy traffic // Adv. Appl. Probab. 1970. **2**. P. 150–177.
- [11] Takács L. Investigation of waiting-time problems by reduction to Markov process // Acta math. Acad. sci. hung. 1955. **6**. P. 101–129.
- [12] Биллингсли П. Сходимость вероятностных мер. М.: Наука, 1977.
- [13] Феллер В. Введение в теорию вероятностей и ее приложения. М.: Мир, 1965.
- [14] Бижалис А. Оценки остаточного члена в центральной предельной теореме // Лит. матем. сб. 1966. **6**, № 1. С. 323–346.
- [15] Afanaseva L. Asymptotic estimations for multichannel queueing systems // Abstracts of Barselona Conference on Asymptotic Statistics. Barselona Centre of research Mathematics Bellaterra, 2008. P. 43–44.

НЕКОТОРЫЕ ЗАДАЧИ ДЛЯ ПОТОКОВ ВЗАИМОДЕЙСТВУЮЩИХ ЧАСТИЦ

Л. Г. Афанасьева, Е. В. Булинская

В первой части работы исследуется плотность потока частиц двух типов, движущихся с разными скоростями. Получены оценки сверху и снизу, а также предложен алгоритм вычисления плотности с заданной точностью. Во второй части учитывается наличие на трассе светофоров (точек задержки) и производится оптимизация их функционирования.

1. Введение

Интерес к изучению пропускной способности автотрасс возник еще до второй мировой войны. В этой связи отметим статью Гриншилдса [1] 1935 года, одну из первых работ в этом направлении. Целый ряд работ появился в 50-е годы XX века в журналах по исследованию операций (см., например, [2]). Модели с взаимодействием частиц, возникающие в различных приложениях (транспортные потоки, развитие биологических систем, химические реакции и др.), привлекли в последние годы внимание многих исследователей. Для их изучения применяются разнообразные подходы. Наряду с теорией клеточных автоматов используются методы статистической механики и математической физики, марковские процессы и многие другие (см., например, [3–13] и ссылки в них).

В данной работе основное внимание уделяется стохастическим марковским моделям, анализ которых производится с помощью методов теории массового обслуживания, теории восстановления и регенерирующих процессов.

Работа состоит из двух частей. В первой части рассматриваются частицы двух типов (медленные и быстрые), движущиеся в одном направлении вдоль действительной прямой. Предполагается, что частицы возникают на прямой независимо друг от друга, при этом частица типа i , $i = 1, 2$, появляется в интервале $(x, x + dx)$ в промежутке времени $(t, t + dt)$ с вероятностью $\lambda_i dx dt$. Каждая частица должна пройти некоторое случайное расстояние прежде, чем покинуть прямую. Расстояния являются независимыми, показательно распределенными случайными величинами с параметром μ_i для частиц i -го типа. Пусть V_i — скорость частицы i -го типа, причем $V_1 < V_2$. Очевидно, что $\tau_i = (\mu_i V_i)^{-1}$ — это среднее время, проводимое частицей i -го типа на прямой в том случае, если она не взаимодействует с другими частицами.

Поведение быстрой частицы в тот момент, когда она догоняет медленную, может быть различным. В зависимости от него мы получаем три различные модели.

Одна из основных задач, решаемых в первой части работы, состоит в оценке плотности частиц в модели 1. *Плотностью* (или концентрацией) частиц в момент t в точке x называется такая функция $\Lambda(t, x)$, что $\Lambda(t, x) dx$ — это среднее число частиц в интервале $(x, x + dx)$ в момент t . Вначале мы выводим бесконечную систему нелинейных уравнений для величин, через которые выражается упомянутая плотность в стационарном режиме. Так как получить решение в явном виде не представляется возможным, мы строим модели 2 и 3, позволяющие оценить снизу и сверху искомую плотность. Предлагается также алгоритм для вычисления плотности с любой заранее заданной точностью.

Вторая часть посвящена исследованию транспортных потоков, движущихся по трассе с перекрестками, на которых могут находиться светофоры. Сначала анализируется работа одного светофора, расположенного на пересечении двух перпендикулярных дорог с односторонним движением. Случайные величины, характеризующие функционирование светофора, предполагаются независимыми и показательно распределенными, так что изучаемые процессы оказываются марковскими. Основным интерес представляет среднее число частиц (автомобилей), находящихся в очереди перед светофором в стационарном режиме. Решается задача оптимизации моментов переключения светофора по критерию минимизации среднего суммарного числа ожидающих автомобилей (по обоим направлениям). Далее предлагается модель, в которой светофоры расположены случайным образом, а автомобили поступают как в точках расположения светофоров, так и между ними. Найдена зависимость плотности автомобилей на трассе от параметров системы.

2. Плотность потока частиц, движущихся с различными скоростями

2.1. Описание моделей. Рассмотрим следующие три модели.

Модель 1. Предполагается, что быстрая частица, догнавшая медленную, начинает двигаться со скоростью V_1 до тех пор, пока одна из этих частиц не покинет прямую. При таком поведении могут образовываться группы частиц произвольного размера. Когда исчезает блокирующая медленная частица, все следовавшие за ней быстрые частицы снова начинают двигаться со скоростью V_2 .

Размер частиц мы не принимаем во внимание. Следовательно, при фиксированном t конфигурация частиц на прямой описывается последовательностью (x_s, n_s, e_s) , где $x_s \in \mathbb{R}^1$ — положение s -й группы на прямой, n_s — число частиц в этой группе, а $e_s = 1$, если в группе имеется медленная частица, и $e_s = 0$, если имеются только быстрые частицы. При этом $s = 0, \pm 1, \pm 2, \dots$, а $x_s \leq 0$, если $s \leq 0$, и $x_s > 0$, если $s > 0$. Таким образом, мы изучаем случайный процесс $X(t)$, принимающий значения в пространстве конфигураций $\mathcal{X} = \{(x_s, n_s, e_s)_{s=-\infty}^{s=+\infty}\}$. В сделанных предположениях это однородный марковский процесс. Для любых непересекающихся интервалов из $\mathbb{R}^1$ события, порожденные течением процесса в этих интервалах, независимы. Как будет показано, если $\mu_i > 0$, $i = 1, 2$, то существует стационарное распределение $X(t)$, которое мы будем изучать.

Заметим, что также возможна другая интерпретация рассматриваемого объекта. А именно мы можем рассматривать векторнозначное случайное поле, зависящее от t и x , со значениями в $\mathbb{Z}_+^1 \times \{0, 1\}$.

Модель 2 отличается от предыдущей наличием следующего условия: для заданного k при $e_s = 1$ всегда $n_s \leq k$. Иными словами, если группа частиц с $n \leq k$, $e = 1$, то есть движущаяся со скоростью V_1 , настигается другой группой (состоящей только из быстрых частиц), то не более $k - n$ частиц могут присоединиться к медленной группе, а остальные мгновенно исчезают. Нетрудно видеть, что плотность частиц в данной модели будет меньше, чем в модели 1.

Модель 3. Предполагается, что быстрая частица, догнавшая медленную, навсегда становится медленной. Это значит, что частицы второго типа (быстрые) не

образуют групп. Если $\tau_2 < \tau_1$, то плотность частиц в этой модели больше, чем в модели 1.

2.2. Эргодическая теорема. Мы начнем с теоремы, устанавливающей существование предельного распределения у процесса $X(t)$.

Теорема 1. Если $V_i > 0$, $\mu_i > 0$ при $i = 1, 2$, то марковский процесс $X(t)$ эргодический.

Доказательство проводится в несколько этапов. Прежде всего строится мажорирующий процесс $\tilde{X}(t)$. Предполагается, что все частицы движутся со скоростью $V = V_1$, а интенсивность их появления на участке $(x, x+dx)$ равна $\lambda = \lambda_1 + \lambda_2$. Каждая частица проходит по прямой случайное расстояние, имеющее показательное распределение с параметром $\mu = \min(\mu_1, \mu_2)$. Обозначим $N_A(t)$ (соответственно $\tilde{N}_A(t)$) число частиц в интервале $(-A, A)$ в момент t для процесса $X(\cdot)$ (соответственно $\tilde{X}(\cdot)$). Тогда $\tilde{N}_A(t)$ стохастически не меньше $N_A(t)$, что записывается в виде $\tilde{N}_A(t) \geq_{st} N_A(t)$ и означает

$$P(\tilde{N}_A(t) > K) \geq P(N_A(t) > K) \quad \text{для любого } K.$$

В свою очередь процесс $\tilde{X}(t)$ эквивалентен процессу $\hat{X}(t)$, в котором частицы не движутся, а, появившись в некоторой точке, находятся в ней случайное время, имеющее показательное распределение с параметром μV , и потом исчезают. Процесс $\hat{X}(t)$ является пуассоновским точечным процессом, и для него $\hat{N}_A(t)$ имеет предельное (при $t \rightarrow \infty$) пуассоновское распределение с параметром $2\lambda A(\mu V)^{-1}$.

Это позволяет проверить выполнение условий эргодической теоремы Боровкова (см. [14, с. 184, теорема 1]) для процесса с дискретным временем $\{X(nh), n \geq 1\}$.

Наконец, устанавливается асимптотическая стохастическая непрерывность $X(t)$, то есть $\|P(x, t + \Delta t, \cdot) - P(x, t, \cdot)\| \rightarrow 0$ для любого x при $\Delta t \rightarrow 0$, $t \rightarrow \infty$. Здесь $P(x, t, \cdot)$ — переходная вероятность однородного марковского процесса $X(t)$, а $\|\cdot\|$ — расстояние по вариации. Детали опускаются ввиду громоздкости. $\square$

Процесс $X(t)$ в дальнейшем считается стационарным.

2.3. Вспомогательная система уравнений. Пусть $p_j(t, x) dx$, $j \geq 0$, — вероятность того, что в момент t в интервале $(x, x + dx)$ содержится группа из j быстрых частиц, а соответствующее $e = 1$, в то время как $q_j(t, x) dx$, $j \geq 1$, — аналогичная вероятность при $e = 0$. Следовательно, $\alpha(t, x) = \sum_{j=0}^{\infty} p_j(t, x)$ — это плотность медленных частиц в момент t в точке x .

Теорема 2. В стационарном режиме медленные частицы образуют пуассоновский точечный процесс на $\mathbb{R}^1$ с интенсивностью $\alpha = \lambda_1 \tau_1$, а $p_j(t, x)$, $q_j(t, x)$ не зависят от t и x и равны соответственно p_j , q_j . Для них справедлива следующая система уравнений:

$$\begin{aligned} \{\mu_1 V_1 + \beta(V_2 - V_1)\} p_0 &= \mu_2 V_1 p_1 + \lambda_1, \\ \{j\mu_2 V_1 + \mu_1 V_1 + \beta(V_2 - V_1)\} p_j &= \mu_2 V_1 (j+1) p_{j+1} + (V_2 - V_1) \sum_{i=0}^{j-1} p_i q_{j-i}, \quad j > 0, \end{aligned} \tag{1}$$

где $\beta = \sum_{j=1}^{\infty} q_j$. Кроме того,

$$\begin{aligned} \{\mu_2 V_2 + \alpha(V_2 - V_1)\} q_1 &= 2\mu_2 V_2 q_2 + \mu_1 V_1 p_1 + \lambda_2, \\ \{j\mu_2 V_2 + \alpha(V_2 - V_1)\} q_j &= \mu_2 V_2 (j+1) q_{j+1} + \mu_1 V_1 p_j, \quad j > 1. \end{aligned} \tag{2}$$

Доказательство. Нетрудно проверить, что справедливо следующее соотношение:

$$\alpha(t+h, x) = (1 - \mu_1 V_1 h) \alpha(t, x - V_1 h) + \lambda_1 h + o(h).$$

Устремляя h к нулю, получаем

$$\frac{\partial \alpha}{\partial t} + V_1 \frac{\partial \alpha}{\partial x} = -\mu_1 V_1 \alpha + \lambda_1. \quad (3)$$

Принимая во внимание однородность по времени и пространству в стационарном состоянии, приходим к выводу, что из (3) следует при всех t и x

$$\alpha(t, x) = \alpha = \lambda_1 (\mu_1 V_1)^{-1} = \lambda_1 \tau_1,$$

что и заканчивает доказательство первого утверждения теоремы.

Далее, в силу независимости в любой момент t чисел частиц на непересекающихся интервалах прямой имеем

$$\begin{aligned} p_j(t+h, x) = & (1 - j\mu_2 V_1 h - \mu_1 V_1 h - (V_2 - V_1)h \sum_{i=1}^{\infty} q_i) p_j(t, x) + \\ & + \mu_2 h V_1 (j+1) p_{j+1}(t, x) + (V_2 - V_1)h \sum_{i=0}^{j-1} p_i(t, x) q_{j-i}(t, x) + o(h). \end{aligned}$$

Устремляя к нулю h и пользуясь тем, что $X(t)$ однороден по времени и пространству, получаем соотношения (1). Соотношения (2) устанавливаются аналогично. $\square$

Таким образом, для модели 1 мы получили бесконечную систему нелинейных уравнений для величин, задающих плотность потоков частиц в стационарном режиме. С другой стороны, для модели 2 система уравнений будет конечна, так как $p_j = q_j = 0$ при $j > k$. Более того, при $k \rightarrow \infty$ последовательность решений для модели 2 будет стремиться к решению для модели 1. Однако при любом k соответствующая система уравнений все равно остается нелинейной. Поэтому для получения решения для модели 1 мы будем использовать метод последовательных приближений.

2.4. Аппроксимационная процедура. Для получения решения систем (1), (2) мы предлагаем следующий

Алгоритм. Обозначим $S^{(k)}$ систему первых $(k+1)$ уравнений соответственно из (1) и (2). Идея заключается в том, чтобы получать системы линейных уравнений, используя результаты решения системы $S^{(k-1)}$. Более точно процедура последовательных приближений состоит из следующих шагов.

Шаг 1. Рассмотрим $k = 0$, тогда

$$\alpha^{(0)} = p_0^{(0)} = \lambda_1 (\mu_1 V_1)^{-1}, \quad \beta^{(0)} = q_1^{(0)} = \lambda_1 (\mu_2 V_2 + \alpha^{(0)} (V_2 - V_1))^{-1}.$$

Шаг 2. Положим $\alpha = \alpha^{(0)}$, $\beta = \beta^{(0)}$, а вместо q_1 в первых двух уравнениях из (1) для p_0 и p_1 возьмем $q_1^{(0)}$. В результате мы получим систему

$$[\mu_1 V_1 + \beta^{(0)} (V_2 - V_1)] p_0 = \mu_2 V_1 p_1 + \lambda_1,$$

$$\begin{aligned} [\mu_2 V_1 + \mu_1 V_1 + \beta^{(0)}(V_2 - V_1)]p_1 &= (V_2 - V_1)p_0 q_1^{(0)}, \\ [\mu_2 V_2 + \alpha^{(0)}(V_2 - V_1)]q_1 &= \mu_1 V_1 p_1 + \lambda_2, \\ [2\mu_2 V_2 + \alpha^{(0)}(V_2 - V_1)]q_2 &= \mu_1 V_1 p_1, \end{aligned}$$

Эта система линейных уравнений имеет решение $(p_0^{(1)}, p_1^{(1)}, q_1^{(1)}, q_2^{(1)})$.

Шаг 3. Теперь рассматривается система $S^{(2)}$, в которой вместо α и β берутся соответственно $\alpha^{(1)} = p_0^{(1)} + p_1^{(1)}$ и $\beta^{(1)} = q_1^{(1)} + q_2^{(1)}$. Как и для шага 2, мы используем $q_1^{(1)}$ и $q_2^{(1)}$ соответственно вместо q_1 и q_2 в уравнениях для p_j , $j = 0, 1, 2$.

Следующие шаги с очевидными изменениями повторяют шаг 3.

Если известны (p_j, q_j) , то можно выписать в явном виде плотность частиц

$$\Lambda = \lambda_1 \tau_1 + \sum_{j=1}^{\infty} j(p_j + q_j)$$

и сосчитать вероятности различных конфигураций частиц на конечных интервалах $\mathbb{R}^1$.

2.5. Оценка сверху плотности частиц. Обратимся теперь к модели 3, где существуют группы только медленных частиц и одиночные быстрые частицы. Это означает, что $q_j = 0$ при $j > 1$, и вместо уравнений (1), (2) мы получаем следующую систему:

$$[\mu_1 V_1 j + q_1(V_2 - V_1)]\tilde{p}_j = (j+1)\mu_1 V_1 \tilde{p}_{j+1} + (V_2 - V_1)q_1 \tilde{p}_{j-1}, \quad j > 1, \quad (4)$$

и

$$[\mu_1 V_1 + q_1(V_2 - V_1)]\tilde{p}_1 = 2\mu_1 V_1 \tilde{p}_2 + \lambda_1, \quad (5)$$

$$[\mu_2 V_2 + \tilde{\alpha}(V_2 - V_1)]q_1 = \lambda_2, \quad (6)$$

где $\tilde{\alpha} = \sum_{j=1}^{\infty} \tilde{p}_j$, а $\tilde{p}_j dx$ означает вероятность того, что имеется группа из j медленных частиц в интервале $(x, x+dx)$ для любых $x \in \mathbb{R}^1$ и $t \in \mathbb{R}^1$.

Теорема 3. *Решение уравнений (4)–(6) имеет вид*

$$q_1 = \frac{\delta_0}{\tau_1(V_2 - V_1)}, \quad \tilde{p}_j = \frac{\delta_0^{j-1}}{j!} \tilde{p}_1, \quad j \geq 1,$$

где $\tilde{p}_1 = \lambda_1 \tau_1$ и δ_0 удовлетворяет следующему уравнению:

$$\lambda_1 \tau_1 (e^{\delta} - 1) = \lambda_2 \tau_1 - \delta[(V_2 - V_1)\tau_2]^{-1}, \quad (7)$$

кроме того,

$$\tilde{\Lambda} = \tau_1(\lambda_1 + \lambda_2) + \frac{\tau_2 - \tau_1}{(V_2 - V_1)\tau_1\tau_2} \delta_0. \quad (8)$$

Доказательство. Просуммировав уравнения (4) по j и прибавив уравнение (5), получаем

$$\tilde{p}_1 = \lambda_1(\mu_1 V_1)^{-1} = \lambda_1 \tau_1. \quad (9)$$

Чтобы найти решение уравнений (4) и (5), положим

$$u_j = j\mu_1 V_1 \tilde{p}_j - q_1(V_2 - V_1)\tilde{p}_{j-1}.$$

Тогда (4) можно переписать в виде

$$u_j = u_{j+1} \quad \text{для} \quad j > 1. \quad (10)$$

Подстановка (9) в (5) дает

$$2\mu_1 V_1 \tilde{p}_2 = [\mu_1 V_1 + q_1(V_2 - V_1)]\tilde{p}_1 - \lambda_1 = q_1(V_2 - V_1)\lambda_1(\mu_1 V_1)^{-1},$$

поэтому

$$u_2 = 2\mu_1 V_1 \tilde{p}_2 - q_1(V_2 - V_1)\tilde{p}_1 = 0.$$

Откуда, принимая во внимание (10), легко видеть, что $u_j = 0$ при $j > 1$ и

$$\tilde{p}_{j+1} = \frac{q_1(V_2 - V_1)}{\mu_1 V_1(j+1)}\tilde{p}_j, \quad j \geq 1.$$

Обозначив $\delta = q_1\tau_1(V_2 - V_1)$, мы получаем

$$\tilde{p}_j = \frac{\delta^{j-1}}{j!}\tilde{p}_1, \quad j \geq 1. \quad (11)$$

Используя (6), (9) и (11), устанавливаем справедливость следующих соотношений:

$$\tilde{\alpha} = \frac{\lambda_1(e^\delta - 1)}{\mu_1 V_1 \delta} = \frac{\lambda_2}{\mu_1 V_1 \delta} - \frac{\mu_2 V_2}{V_2 - V_1}, \quad (12)$$

а значит, и тот факт, что δ удовлетворяет (7). Это уравнение имеет единственное решение δ_0 и

$$\delta_0 < \delta_1 = \lambda_2(V_2 - V_1)\tau_1\tau_2. \quad (13)$$

Наконец, мы имеем

$$\tilde{p}_j = \frac{\lambda_1\tau_1\delta_0^{j-1}}{j!}, \quad j \geq 1,$$

в то время как среднее число медленных частиц в группе равно

$$m = \lambda_1\tau_1 \sum_{j=1}^{\infty} j \frac{\delta_0^{j-1}}{j!} = \lambda_1\tau_1 e^{\delta_0} = (\lambda_1 + \lambda_2)\tau_1 - \frac{\delta_0}{\tau_2(V_2 - V_1)}.$$

Согласно (6) и (12), получаем $q_1 = \delta_0\mu_1 V_1/(V_2 - V_1)$. Следовательно, плотность потока частиц, равная $\tilde{\Lambda} = q_1 + m$, имеет вид (8). Тем самым доказательство закончено. $\square$

Если δ_1 в (13) мало, то, используя разложение Тейлора для e^δ только с линейной частью, получим вместо (7) простое уравнение

$$\lambda_1\tau_1\delta = \lambda_2\tau_1 - \delta[\tau_2(V_2 - V_1)]^{-1},$$

из которого вытекает следующее приближение для δ_0 :

$$\tilde{\delta} = \frac{\lambda_2}{\lambda_1 + [\tau_1\tau_2(V_2 - V_1)]^{-1}}.$$

Нетрудно понять, что $\tilde{\Lambda} \geq \Lambda$ при $\tau_1 > \tau_2$.

3. Светофор

Приложение полученных результатов к движению автомобилей по скоростной трассе требует рассмотрения возможности пересечения трассы автомобилями, следующими в перпендикулярном направлении, а значит, и наличия светофоров. В этом разделе будет дан анализ функционирования одного светофора.

3.1. Описание модели. Движение частиц осуществляется в двух (перпендикулярных) направлениях. Потоки пуассоновские с интенсивностью a_i для i -го направления, $i = 1, 2$. В точке пересечения трасс имеется светофор, разрешающий (зеленый свет) или запрещающий (красный свет) движение в первом направлении. Длительности промежутков, когда горит зеленый (красный) свет в первом направлении, — показательно распределенные случайные величины с параметром γ_1 (γ_2). Очевидно, что когда в первом направлении горит зеленый свет, то во втором красный, и наоборот.

Если в момент подхода частицы (автомобиля) к светофору в направлении ее движения горит зеленый свет и нет очереди, она беспрепятственно следует далее. Если горит красный светофор или есть очередь ожидающих частиц, то останавливается. В таком случае время прохождения светофора имеет показательное распределение с параметром ρ . Все случайные величины, определяющие функционирование светофора, а также прохождение перекрестка, независимы и не зависят от входящих потоков.

Введем случайный процесс $X(t) = (X_1(t), X_2(t), e(t))$, где $X_i(t)$ — количество автомобилей в момент t перед светофором в i -м направлении, $i = 1, 2$, а $e(t) = 1$, если зеленый свет горит для первого направления, и $e(t) = 0$ в противном случае. В силу сделанных предположений процесс $X(t)$ является цепью Маркова. Следующая теорема дает необходимые и достаточные условия ее эргодичности.

Теорема 4. *Цепь Маркова $X(t)$ эргодична тогда и только тогда, когда*

$$a_i \theta < \rho \gamma_i^{-1}, \quad i = 1, 2, \quad (14)$$

где $\theta = \gamma_1^{-1} + \gamma_2^{-1}$.

Доказательство. Пусть $\{t_n, n \geq 1\}$ — последовательность моментов включения зеленого света для первого направления. Для i -го направления, $i = 1, 2$, обозначим через $\xi_n^{(i)}$ количество автомобилей, поступающих за время (t_{n-1}, t_n) , а через $\eta_n^{(i)}$ количество автомобилей, которые могли бы проехать через светофор за время, пока для них горит зеленый свет в интервале (t_{n-1}, t_n) , если бы все время была очередь. Введем цепи Маркова $Y_n^{(i)}$ и $Z_n^{(i)}$, $i = 1, 2$, с помощью рекуррентных соотношений

$$\begin{aligned} Z_n^{(i)} &= [Z_{n-1}^{(i)} - \eta_n^{(i)}]^+ + \xi_n^{(i)}, \\ Y_n^{(i)} &= [Y_{n-1}^{(i)} + \xi_n^{(i)} - \eta_n^{(i)}]^+, \quad i = 1, 2. \end{aligned}$$

Если $Z_0^{(i)} = Y_0^{(i)} = X_i(t_0)$, то выполняются стохастические неравенства

$$Y_n^{(i)} \leq_{st} X_i(t_n) \leq_{st} Z_n^{(i)}, \quad i = 1, 2.$$

Поскольку

$$E\xi_n^{(i)} = a_i \theta, \quad E\eta_n^{(i)} = \rho \gamma_i^{-1}, \quad i = 1, 2,$$

процессы $\{Z_n^{(i)}\}$ и $\{Y_n^{(i)}\}$ стохастически ограничены тогда и только тогда, когда выполняется (14) (см., например, [15]). Процесс $X(t)$ – регенерирующий, так что стохастическая ограниченность его координат эквивалентна его эргодичности, что сразу следует из теоремы Смита (см., например, [16]). $\square$

3.2. Средняя очередь перед светофором и оптимизация. Далее предполагается, что условие (14) выполнено и цепь Маркова $X(t)$ стационарна. Рассмотрим цепь Маркова $\{X_1(t), e(t)\}$. Введем обозначения:

$$p_j = P(X_1(t) = j, e(t) = 1), \quad r_j = P(X_1(t), e(t) = 0),$$

$$P(z) = \sum_{j=0}^{\infty} z^j p_j, \quad R(z) = \sum_{j=1}^{\infty} z^j r_j, \quad |z| \leq 1.$$

Из системы уравнений для стационарного распределения процесса $\{X_1(t), e(t)\}$ для производящих функций $P(z)$ и $R(z)$ получаем соотношения

$$[\gamma_1 + \rho(1 - z^{-1}) + a_1(1 - z) - \gamma_2 R(z)]P(z) = p_0[\rho(1 - z) + a_1(1 - z)],$$

$$(\gamma_2 + a_1(1 - z))R(z) = \gamma_1 P(z),$$

откуда

$$P(z) = p_0 \frac{\rho - a_1 z}{\rho - a_1 z - a_1 z \gamma_1 (\gamma_2 + a_1(1 - z))^{-1}}. \quad (15)$$

Поскольку $P(1) = P(e(t) = 1) = \gamma_2(\gamma_1 + \gamma_2)^{-1}$, находим

$$p_0 = \frac{\rho \gamma_2 - a_1(\gamma_1 + \gamma_2)}{(\gamma_1 + \gamma_2)(\rho - a_1)}. \quad (16)$$

Для удобства введем следующие параметры: $c_i = a_i \rho^{-1}$, $x = (\gamma_1 \theta)^{-1}$, $d = \rho \theta$, так что x — доля времени, в течение которого для первого направления горит зеленый свет. Условия эргодичности (14) в этих терминах принимают вид

$$c_i \geq 0, \quad i = 1, 2, \quad c_1 + c_2 < 1, \quad x \in (c_1, 1 - c_2).$$

Среднее число автомобилей $m_1(x)$, находящихся перед светофором (или проезжающих его) по первому направлению, определяется из (15) и (16). Соответствующая формула имеет вид

$$m_1(x) = P'(1) + R'(1) = \frac{c_1(1 - x)(1 + d(1 - c_1)x(1 - x))}{(1 - c_1)(x - c_1)}.$$

Аналогично для второго направления

$$m_2(x) = \frac{c_2 x(1 + d(1 - c_2)x(1 - x))}{(1 - c_2)(1 - x - c_2)}.$$

Предположим, что среднее время θ между последовательными моментами включения зеленого света фиксировано. Мы хотим выбрать $x \in (c_1, 1 - c_2)$ так, чтобы значение функции

$$f(x) = m_1(x) + m_2(x)$$

было минимальным. Поскольку в указанном интервале функция $m_1(x)$ строго убывает, а $m_2(x)$ строго возрастает, существует единственная точка минимума, которая является решением уравнения

$$m'_1(x) = -m'_2(x).$$

В симметричном случае, когда $c_1 = c_2 = c$, минимум достигается в точке $x_{\min} = 0,5$, а минимальное значение равно

$$f(0,5) = 2 \frac{c(1+0,25d(1-c))}{(1-c)(1-2c)}.$$

4. Плотность потока частиц при наличии светофора

В разделе 2 исследовано влияние на плотность потока движущихся частиц различия в их скоростях. Наличие точек задержки (светофоров) способствует образованию больших скоплений частиц и увеличению их плотности, даже если они движутся с одной и той же скоростью. Анализ этого явления посвящен данный раздел.

4.1. Описание модели. Предположим, что расстояния между соседними светофорами образуют последовательность $\{\eta_i\}_{-\infty}^{+\infty}$ независимых случайных величин с общей функцией распределения $F(x)$. Режим работы каждого светофора такой же, как в разделе 3. Автомобили на трассе появляются как в точках расположения светофоров, так и между ними. Пусть $\lambda dx dt$ — вероятность возникновения автомобиля в интервале $(x, x+dx)$ за время dt . В точку расположения светофора поступают два независимых пуассоновских потока. Первый из них (интенсивности a_2) сворачивает на трассу после прохождения светофора, а второй (интенсивности a_3) следует в перпендикулярном направлении. Все автомобили движутся по трассе с постоянной скоростью V и проходят по ней случайное расстояние, имеющее показательное распределение с параметром μ .

4.2. Интенсивность потока частиц и эргодическая теорема. Выберем один из светофоров и введем наборы случайных процессов: $Y_1(t)$ — количество автомобилей, подошедших к светофору по основной трассе за время t ; $Z_1(t)$ — количество автомобилей, прошедших через светофор за то же время; $Q_1(t)$ — число автомобилей, стоящих перед светофором в момент t . Соответственно, процессы $Y_2(t), Z_2(t), Q_2(t)$ будут относиться к автомобилям, сворачивающим на трассу с перпендикулярного направления, а $Y_3(t), Z_3(t), Q_3(t)$ — к пересекающим трассу, не сворачивая на нее.

Под интенсивностью процесса $Y_1(t)$ мы будем понимать (существующую в сделанных предположениях) величину

$$a_1 = \lim_{t \rightarrow \infty} \frac{Y_1(t)}{t}.$$

Первый результат касается этого предела.

Теорема 5. Если процессы $Q_i(t)$, $i = 1, 2, 3$, стохастически ограничены, то

$$a_1 = \frac{a_2 f^*(\mu)}{1 - f^*(\mu)} + \frac{\lambda}{\mu}, \quad f^*(\mu) = \int_0^\infty e^{-\mu x} dF(x). \quad (17)$$

Доказательство. Поскольку $t^{-1}Q_i(t) \rightarrow 0$ при $t \rightarrow \infty$ и $Z_i(t) = Y_i(t) - Q_i(t)$, $i = 1, 2, 3$, то

$$\lim_{t \rightarrow \infty} \frac{Z_i(t)}{t} = \lim_{t \rightarrow \infty} \frac{Y_i(t)}{t} = a_i, \quad i = 1, 2, 3.$$

Интенсивность потока автомобилей, проходящих через светофор и следующих по трассе, таким образом, равна $a_1 + a_2$. Некоторые из этих автомобилей покинут трассу до того, как они достигнут следующего светофора. Вероятность этого события равна $\int_0^\infty [1 - e^{-\mu x}] dF(x)$, так что интенсивность потока автомобилей на следующем светофоре, из числа прошедших предыдущий, составит

$$(a_1 + a_2) \int_0^\infty e^{-\mu x} dF(x) = (a_1 + a_2)f^*(\mu).$$

К следующему светофору также придут некоторые автомобили из числа появившихся на трассе между светофорами. Интенсивность этого потока равна

$$\lambda \int_0^\infty \int_0^y e^{-\mu x} dx dF(y) = \frac{\lambda}{\mu}(1 - f^*(\mu)).$$

Из соображений симметрии интенсивности потоков автомобилей, следующих по трассе, равны для всех светофоров, поэтому получаем уравнение

$$a_1 = (a_1 + a_2)f^*(\mu) + (\lambda/\mu)(1 - f^*(\mu)), \quad (18)$$

откуда немедленно вытекает (17). $\square$

Возникает вопрос, при каких условиях на параметры системы $a_2, a_3, \gamma_1, \gamma_2, \rho, \mu, \lambda, V$ процессы $Q_i(t)$, $i = 1, 2, 3$, будут стохастически ограничены.

Теорема 6. Если $V > 0$, $\mu > 0$, процессы $Q_i(t)$, $i = 1, 2, 3$, имеют собственное предельное распределение тогда и только тогда, когда

$$a_1\theta < \rho\gamma_1^{-1}, \quad (a_2 + a_3)\theta < \rho\gamma_2^{-1},$$

где a_1 задается равенством (17), а параметры γ_i , $i = 1, 2$, θ и ρ определены в разделе 3.

Доказательство. Совокупность светофоров представляет собой своеобразную сеть Джексона с бесконечным числом узлов. Особенность этой сети в том, что во время перехода из узла i в узел $i+1$ требования могут исчезать и появляться. Вероятность исчезновения требования равна $1 - f^*(\mu)$, а интенсивность пуассоновского потока вновь поступающих требований составляет $(\lambda/\mu)(1 - f^*(\mu))$. Следовательно, соотношение (18) представляет собой уравнение баланса для рассматриваемой сети.

Далее доказательство эргодичности проводится обычным для сетей Джексона методом (см., например, [17]). $\square$

4.3. Плотность потока частиц при наличии светофора. Введем случайный процесс $X(t)$ со значениями в фазовом пространстве $\mathcal{X} = \{x\}$, где

$$x = (y_i, l_i, n_i, k_i, z_1^{(i)}, \dots, z_{k_i}^{(i)}, e_i)_{i=-\infty}^{+\infty},$$

y_i — координата i -го светофора,
 l_i — число автомобилей на трассе перед i -м светофором,
 n_i — число автомобилей в перпендикулярном направлении перед i -м светофором,
 k_i — число автомобилей на трассе между i -м и $i + 1$ -м светофорами,
 $z_j^{(i)}$ — расстояние j -го из указанных автомобилей до i -го светофора, $j = 1, \dots, k_i$,
наконец, $e_i = 1$, если у i -го светофора для трассы горит зеленый свет, и $e_i = 0$, если красный.

В сделанных предположениях процесс $X(t)$ марковский, и, если выполнены условия теоремы 6, у него есть стационарное распределение. Далее предполагается, что процесс $X(t)$ стационарен.

Под плотностью потока частиц будем понимать

$$\Lambda = \lim_{A \rightarrow \infty} \frac{N_A}{2A},$$

где, как и раньше, N_A — число частиц в промежутке $(-A, A)$.

Пусть αdx — вероятность наличия частиц в интервале $(x, x + dx)$, причем x лежит между светофорами. Если $\varkappa = \int_0^\infty (1 - F(x)) dx$ — среднее расстояние между светофорами, то

$$\Lambda = \frac{m_1}{\varkappa} + \alpha, \quad (19)$$

где m_1 — среднее число частиц перед светофором на трассе.

Чтобы найти α , введем функцию $\beta(x)$, представляющую собой условную плотность частиц при условии, что ближайший слева светофор находится на расстоянии x . Обычным образом получаем уравнение

$$V \frac{d\beta}{dx} = -\mu V \beta(x) + \lambda \quad (20)$$

с начальным условием

$$\beta(0) = \frac{a_1 + a_2}{V}. \quad (21)$$

Решение (20) при условии (21) имеет вид

$$\beta(x) = \frac{\lambda}{V\mu} (1 - e^{-\mu x}) + \frac{a_1 + a_2}{V} e^{-\mu x}. \quad (22)$$

Теперь из теории восстановления (см., например, [16]) получаем

$$\alpha = (\varkappa)^{-1} \int_0^\infty \beta(y) [1 - F(y)] dy.$$

Используя (22) и (17) после несложных выкладок, находим

$$\alpha = \frac{a_2 + \lambda \varkappa}{V \varkappa \mu} = \frac{\lambda}{V \mu} + \frac{a_2}{V \mu \varkappa},$$

откуда с учетом (19) получаем формулу

$$\Lambda = \frac{\lambda}{V_{\mu}} + \frac{a_2}{V_{\mu}\varkappa} + \frac{m_1}{\varkappa}.$$

Основная сложность состоит в вычислении m_1 . Если $\varkappa$ велико, то поток автомобилей перед светофором близок к пуассоновскому и можно воспользоваться результатами раздела 3.

5. Дальнейшие направления исследований

Конечно, реальные автомобили имеют некоторую конечную длину b , поэтому полученные в предыдущем разделе результаты могут применяться в предположении, что $\varkappa$ много больше, чем $m_1 b$. Естественно, возникает вопрос, что делать, если это условие не выполнено. Исследование влияния длины автомобилей на образование пробок является одной из проблем для будущих исследований. Также желательно принять во внимание при построении модели необходимость поддержания определенного «интервала безопасности» между автомобилями и невозможность мгновенного уменьшения или увеличения скорости движения.

Оптимизацию режима работы светофора можно также проводить в рамках других моделей. Один из возможных подходов состоит в том, что переключение светофора осуществляется в те моменты, когда количество частиц в одном из направлений достигает некоторого уровня K . Этот уровень надо выбрать так, чтобы, например, среднее число ожидающих частиц было минимальным. В такой ситуации следует рассматривать модель, задаваемую процессом $(X_1(t), X_2(t), e(t))$, где $X_i(t)$, $i = 1, 2$, — количество автомобилей перед светофором в момент t . При этом, если $e(t) = 0$ (горит красный свет для первого направления) и процесс $X_1(t)$ достигает уровня K (снизу), значение $e(t)$ становится равным 1, то есть загорается зеленый свет. Момент переключения зеленого света на красный можно также задавать различными способами, например, когда на втором направлении накопится достаточно длинная очередь либо когда на первом направлении очередь окажется мала. Для таких моделей можно выписать уравнения для вероятностей состояний, с помощью которых затем построить вычислительную процедуру выбора оптимального уровня K или каких-то других параметров, характеризующих работу светофора.

Тот факт, что мы рассматриваем лишь одностороннее движение, не является существенным ограничением, так как для скоростных трасс характерно разделение полос для движения в различных направлениях. Однако на таких трассах, как правило, имеется несколько полос движения в одном направлении. Заметим также, что во многих городах, таких, как Париж и Мадрид, существуют специально выделенные полосы для общественного транспорта, въехать на которые или съехать с которых возможно лишь на перекрестке. Таким образом, еще одно направление для исследования — это изучение многополосного движения с возможностью обгона медленно движущегося транспорта.

Интересно также исследовать влияние не только светофоров, но и других «точек задержки» на трассе. Это могут быть, например, дорожно-транспортные происшествия, а также пункты оплаты за проезд или знаки дорожного движения, требующие уменьшения скорости.

Наконец, планируется изучить модели с более сложными входными потоками и правилами функционирования светофоров, что может означать немарковость рассматриваемых процессов.

Работа выполнена при частичной поддержке РФФИ, грант № 07-01-00362.

Список литературы

- [1] *Greenshields B. D.* A study of highway capacity // Proc. Highway Res. 1935. **14**. P. 448–477.
- [2] *Greenberg H.* An analysis of traffic flows // Oper. Res. 1959. **7**. P. 79–85.
- [3] *Blank M.* Ergodic properties of a simple deterministic traffic flow model // J. Stat. Phys. 2003. **111**. P. 903–930.
- [4] *Cáceres F. C., Ferrari P. A., Pechersky E.* A slow to start traffic model related to a M/M/1 queue // arXiv: cond-mat/0703709 v2 [cond-mat.stat-mech] 31 May 2007.
- [5] *Fuks H., Boccara N.* Convergence to equilibrium in a class of interacting particle system evolving in discrete time // Phys. Rev. E. 2001. **64**, № 016117.
- [6] *Maerivoet S., De Moor B.* Cellular automata models of road traffic // Phys. Rept. 2005. **419**. P. 1–64.
- [7] *May A. D.* Traffic Flow Fundamentals. Englewood Cliffs, New Jersey: Prentice-Hall, 1990.
- [8] *Helbing D.* Traffic and related self-driven many-particle systems // Rev. Modern Phys. 2001. **73**. P. 1067–1141.
- [9] *Chowdhury D., Santen L., Schadschneider A.* Vehicular traffic: A system of interacting particles driven far from equilibrium // arXiv: cond-mat/9910173 v1 [cond-mat.stat-mech] 12 October 1999.
- [10] *Schadschneider A.* Statistical physics of traffic flow // arXiv: cond-mat/0007418 v1 [cond-mat.stat-mech] 26 July 2000.
- [11] *Иносе Х., Хамада Т.* Управление дорожным транспортом. М.: Транспорт, 1983.
- [12] *Семенов В. В.* Математическое моделирование динамики транспортных потоков мегаполиса // Препринт. 2004. http://www.keldysh.ru/papers/2004/prepr34/prepr2004_34.html .
- [13] *Малышев В. А., Манита А. Д., Замятин А. А.* Явление гомеостаза в сетях химических реакций // Теория вероятностей и ее применения. 2006. **51**, № 4. С. 793–801.
- [14] *Боровков А. А.* Эргодичность и устойчивость случайных процессов. М.: УРСС, 1999.
- [15] *Афанасьева Л. Г.* Системы массового обслуживания с циклическими управляющими процессами. // Кибернетика и системный анализ. 2005. № 1. С. 54–68.
- [16] *Афанасьева Л. Г., Булинская Е. В.* Случайные процессы в теории массового обслуживания и управления запасами. М.: Изд-во МГУ, 1980.
- [17] *Kelly F. P., Williams J. R.* Stochastic Networks. The IMA Volumes in Mathematics and Applications. Vol. 71. N.Y.: Springer, 1995.

СТОХАСТИЧЕСКИЕ МОДЕЛИ В РАДИОБИОЛОГИИ

А. В. Булинский

Начиная с 90-х годов XX века в радиобиологии стали активно применяться модели, основанные на идее независимого поведения функциональных единиц, составляющих биологическую систему, подверженную облучению. Наряду с рассмотрением таких моделей предлагается более общий подход, использующий различные формы зависимости случайных величин, характеризующих воздействие на клетки и составные элементы системы. При этом существенную роль играют предельные теоремы, устанавливаемые для сумм зависимых мультииндексированных величин. Кроме того, обсуждаются варианты сочетания классической модели «критического объема» и кластерных моделей, вовлекающих геометрические аспекты расположения поврежденных частей органов и тканей.

1. Введение

Проблемы, связанные с онкологическими заболеваниями, относятся к числу важнейших для современного естествознания. На сегодняшний день достигнуты значительные успехи в диагностике и лечении рака. При этом существенную роль играют как практические, так и теоретические исследования, ведущиеся в крупнейших отечественных и зарубежных научных центрах.

Интересно отметить, что В. Рентген объявил об открытии X-лучей, то есть фотонов (определенного частотного диапазона), 28 декабря 1895 года, а уже 12 января 1896 года Э. Груббе начал их использование для лечения раковых опухолей (см. [1]). В настоящее время в клиниках также применяется облучение потоками электронов, протонов, нейтронов и других элементарных частиц. Ныне планирование радиотерапии начинается с построения трехмерной модели опухоли с помощью компьютерной томографии или магниторезонансной техники. Сложность лечения обусловлена множеством разнообразных факторов, которые необходимо учитывать одновременно. Так, наряду с непосредственным воздействием радиации возникают сопутствующие эффекты (например, образование свободных радикалов), результат воздействия зависит от кислородного окружения, а также от индивидуальной радиочувствительности пациента.

Более того, усилия, направленные на разрушение клеток опухоли (клоногенов), приводят к повреждению нормальных органов и тканей, из-за чего возникают осложнения. Поэтому само облучение, как правило, разделено во времени на ряд этапов, и начинает действовать недостаточно изученный механизм взаимодействия различных клеток, органов и тканей, обеспечивающий восстановление пораженных радиацией частей биологической системы. Таким образом, трудности выбора оптимальной схемы облучения связаны и с тем, что приходится одновременно решать две задачи: поражения всех клоногенов и нанесения по возможности минимального ущерба окружающим опухоль нормальным органам и тканям. Соответственно при разработке плана лечения и оценке его эффективности используются такие показатели, как вероятность поражения опухоли и вероятность осложнений, обусловленных повреждением нормальной ткани, то есть TCP (tumour control probability) и NTCP (normal tissue complication probability). Ука-

занные вероятности сами могут по-разному учитываться при задании целевых функций (см., например, [2, 3]).

Существуют различные стохастические модели, описывающие реакцию ткани или органа на облучение. Распространенной является модель ударов и мишеней (hits and targets), предполагающая, что для поражения клетки, содержащей m мишеней (чувствительных элементов), из них необходимо уничтожить не менее k , где $k \in \{1, \dots, m\}$. Часто применяется LQ (linear quadratic)-модель, в которой вероятность выживания клетки дается формулой

$$p_s(D) = \exp\{-(\alpha D + \beta D^2)\}, \quad (1)$$

здесь α и β — положительные параметры, а D — доза облучения, задаваемая в греях ($1 \text{ Гр} = 1 \text{ Дж/кг}$). Считается, что для опухолей $\alpha/\beta \approx 10$, а для нормальных тканей $\alpha/\beta \approx 3$. В частности, (1) при $\beta = 0$ приводит к распространенной модели «одного удара, одной мишени» (single hit, single target). Заметим, что сравнительно недавно обнаружился эффект гиперчувствительности клетки к малым дозам облучения (см., например, [4]), математически это отражается в том, что в (1) величина $\alpha = \alpha(D)$, причем существует $\lim_{D \rightarrow 0} \alpha(D) = D_0$.

В 1988 году Уизерс и др. [5] предложили¹⁾ применять для биологического моделирования идею **независимых функциональных единиц** (FSU — functional subunit). Подход, основанный на этой идее, с применением биномиального распределения использовался в работах [6–8], см. также [9] и данные там ссылки. Имеются различные трактовки понятия FSU. Считается, например, что деление на части органа или ткани должно учитывать выполняемые ими функции, принимаются во внимание и возможности контроля за поведением составных частей биологической системы. Для нас будет существенно, что модель получается двух-уровневой: орган (ткань) состоит из функциональных единиц, а каждая из них образована достаточно большим набором клеток.

Предположим, что опухоль представляет собой N_0 клоногенов, независимо друг от друга реагирующих на облучение. Если вероятность выживания каждого из них под действием дозы облучения D есть $\exp\{-\alpha D\}$, где $\alpha > 0$ — параметр радиочувствительности биологической системы, то вероятность поражения опухоли при дозе облучения D равна

$$p_{\text{tumour}}(D) = (1 - \exp\{-\alpha D\})^{N_0}, \quad (2)$$

поскольку требуется уничтожить все клоногены.

Отдельного упоминания заслуживают усилия, направленные на разработку подхода к классификации «архитектуры» облучаемых органов (тканей). Так, различают последовательные, параллельные, смешанные структуры, а также фиброзные образования (см. [10]). Например, спинной мозг позвоночника часто описывается системой последовательных элементов, и для повреждения такой структуры считается достаточным разрушить хотя бы один элемент этой системы. Большое внимание уделяется также проблеме неравномерного облучения (см., например, [11, 12]).

¹⁾ Аналогичная идея прослеживается и у некоторых других авторов.

Классическая модель критического объема базируется на предположении, что нормальная ткань (или орган) имеет функциональный резерв. Это значит, что для органа, состоящего из N независимо действующих функциональных единиц, существует такое M , что осложнения начинаются в том случае, когда число погибших FSUs больше либо равно M . При $M = 1$ получаем модель критического элемента, а при $M = N$ — модель, относящуюся к поражению всех элементов (в частности, всех клоногенов, образующих опухоль). Пусть $p_{\text{FSU}} = p_{\text{FSU}}(\vec{p}\vec{a}r, D)$ — вероятность поражения данной FSU дозой облучения D , а $\vec{p}\vec{a}r$ — совокупность параметров, характеризующих реакцию FSU на радиацию. Тогда параметры модели — это N, M и $p_{\text{FSU}}(\vec{p}\vec{a}r, D)$. Следовательно, NTCP для «индивидуальной» ткани (органа), соответствующая модели критического объема, имеет вид

$$p_{\text{ind}}(N, M, p_{\text{FSU}}(\vec{p}\vec{a}r, D)) = \sum_{k=M}^N C_N^k p_{\text{FSU}}^k (1 - p_{\text{FSU}})^{N-k}, \quad (3)$$

где $C_N^k = N!/(k!(N-k)!)$.

При облучении $N\nu$ функциональных единиц ($1 \leq N\nu \leq N$) используется следующее, хорошо известное приближение для биномиального распределения:

$$\sum_{k=M}^{N\nu} C_{N\nu}^k p_{\text{FSU}}^k (1 - p_{\text{FSU}})^{N\nu-k} \approx \Phi \left(\frac{\sqrt{N}(\nu p_{\text{FSU}} - \mu_{\text{cr}})}{\sqrt{\nu p_{\text{FSU}}(1 - p_{\text{FSU}})}} \right), \quad (4)$$

здесь Φ — функция распределения стандартного нормального закона и $\mu_{\text{cr}} = M/N$. Об оценке погрешности этого приближенного равенства см., например, [19].

Напомним, что для курса лечения, состоящего из набора процедур облучения, разделенных во времени, для подсчета TCP используется формула Зайдера и Минербо [13] и ее модификация, предложенная в [14] (допускающая различные временные интервалы между сеансами облучения):

$$\text{TCP} = \left[1 - \frac{p_s(T_n) e^{\lambda T_n}}{1 - p_s(T_n) e^{\lambda T_n} \sum_{k=1}^{n-1} \frac{1}{p_s(T_k)} [e^{-\lambda T_k} - e^{-\lambda T_{k+1}}]} \right]^N, \quad (5)$$

где N — начальное число клоногенов, λ — скорость их восстановления, T_k — время между k -м и первым сеансами облучения, а вероятность выживания клеток после k -го сеанса дается формулой

$$p_s(T_k) = \exp \left\{ -\alpha \left(\frac{k}{n} D \right) - \beta \frac{(\frac{k}{n} D)^2}{k} \right\},$$

здесь D — общая доза, полученная за n сеансов, причем считается, что в каждом сеансе облучения дается одна и та же доза; α и β — параметры радиочувствительности.

Объем статьи не позволяет остановиться на других распространенных моделях облучения. Более того, ван Луйк и др. [15] проанализировали некоторые данные, касающиеся эффекта дозы-объема при облучении спинного мозга крыс, и

пришли к выводу, что ни одна из широко известных моделей NTCP (модель Лимана, повреждения смежных единиц, критического элемента с учетом миграции и др.), базирующихся на предположении о независимости поведения FSUs, *не поддерживает эти данные*²⁾. Поэтому естественно обсудить различные обобщения известных моделей, которые могут оказаться полезными для применения. В данной работе мы сосредоточимся на анализе NTCP. При этом ограничимся индивидуальными моделями (применимыми, например, к практически идентичным животным, изучаемым в лабораторных условиях), для популяционных моделей необходимы некоторые предположения, касающиеся стохастической природы используемых параметров (например, радиочувствительности пациентов), см. [9].

Отметим также, что реальные данные экспериментов, как правило, недоступны, а в статьях приводятся лишь графики или таблицы, построенные на основании этих данных. Кроме того, если не учитывать эксперименты над животными, то объемы имеющихся у исследователей выборок обычно весьма малы, см. также статью [16].

Меркс и Глазьер [17] подчеркивают, что для правильного использования имеющейся информации и выдвижения новых гипотез системная биология переходит от традиционных концептуальных моделей к математическим. Поскольку они сложны и не поддаются аналитическому решению, их заменяют вычислительными моделями или алгоритмами. При этом они часто реализуются в виде симуляции (компьютерного моделирования), см. [10, 14]. Отметим также, что ныне для построения и анализа стохастических моделей используется разнообразный аппарат. Например, в статье [18] для изучения эволюции уровня определенного белка, связанного с развитием рака предстательной железы, применяется стохастическое дифференциальное уравнение, приводящее к изучению соответствующего уравнения Колмогорова (Фоккера – Планка).

Справедливость слишком сложных моделей, вообще говоря, трудно установить. Здесь уместно вспомнить классическое высказывание, приписываемое разным людям, в том числе фон Нейману: «Все модели неверны, но некоторые из них полезны».

2. Случайные поля на $\mathbb{Z}^d$ и их частные суммы

В данном разделе мы развиваем подход, предложенный в работах [19] и [20] для подсчета NTCP путем моделирования облученного органа (или ткани) как системы **зависимых** функциональных единиц. Тем самым появляется возможность описывать коллективные эффекты поведения облученных клеток и учитывать взаимодействие ряда факторов. Для описания совместного поведения клеток и FSUs используются различные типы зависимости (о зависимых случайных полях см. [21]). Независимые FSUs представляют частный случай изучаемой нами модели.

Предположим, что облучаемый орган (ткань) разделен на набор одинаковых частей (единиц) с помощью некоторого параллелепипеда разбиения пространства $\mathbb{R}^d$. Удобно считать, что состояние каждой такой FSU описывается с помощью случайной величины X_j , где $j \in \mathbb{Z}^d$ (мультииндекс j представляет собой центр со-

²⁾ Авторы не утверждают, что рассмотренные ими модели не годятся для любых наблюдений.

ответствующего параллелепипеда, а выбор масштаба по каждой координатной оси позволяет свести рассмотрение к решетке $\mathbb{Z}^d$). Далее мы обратимся к обобщениям этой схемы, используя системы подмножеств $\mathbb{R}^d$.

Начнем со следующей простой модели. Пусть $X = \{X_i, i \in \mathbb{Z}^d\}$ — случайное поле, состоящее из независимых одинаково распределенных (н.о.р.) случайных величин, принимающих значения 1 и 0 соответственно с вероятностями p и $1 - p$. Значение 1 показывает, что клетка i убита, а 0 — жива, $0 \leq p \leq 1$, $p = p(D)$, где D — доза радиации. Например, в популярной модели «один удар, одна мишень» $p(D) = 1 - \exp\{-\alpha D\}$, здесь $\alpha > 0$ — (индивидуальная) радиочувствительность. Разумеется, $X_i = X_i(D)$, $i \in \mathbb{Z}^d$.

Коллективное поведение ансамблей клеток зависит от целого ряда факторов (механизмы восстановления, миграция клеток и др.). Поэтому обратимся к более сложной модели и допустим, что строго стационарное случайное поле $Y = \{Y_j, j \in \mathbb{Z}^d\}$ состоит из ассоциированных величин (см., например, [21, с. 11]), принимающих значения 1 и 0. События $\{Y_j = 1\}$ и $\{Y_j = 0\}$ означают соответственно, что FSU с номером j убита или жива. Положим

$$P(Y_j = 1) = p_{\text{FSU}}, \quad P(Y_j = 0) = 1 - p_{\text{FSU}}. \quad (6)$$

Предположим также, что для j -й FSU ее (итоговое) состояние W_j зависит от «окружающих значений» поля Y . А именно пусть I_Δ — индикатор множества $\Delta \subset \mathbb{R}$, и для интервалов $\Delta_0 := [0, u)$, $\Delta_1 = [u, \infty)$, $0 < u < \infty$, введем

$$W_j = I_{\Delta_1}(Z_j), \quad Z_j = \sum_{i \in \mathbb{Z}^d} \varphi(j - i) Y_i, \quad j \in \mathbb{Z}^d, \quad (7)$$

функция $\varphi : \mathbb{Z}^d \rightarrow \mathbb{R}_+$, а ряд сходится в пространстве $L^2(\Omega, \mathcal{F}, P)$. Тогда $W = \{W_j, j \in \mathbb{Z}^d\}$ — ассоциированное случайное поле, поскольку этим свойством обладает $Z = \{Z_j, j \in \mathbb{Z}^d\}$ (см., например, [21, с. 15]).

Граница множества $U \subset \mathbb{Z}^d$ определяется как совокупность точек

$$\delta U = \{j \in \mathbb{Z}^d \setminus U : \rho(j, U) = 1\},$$

где $\rho(j, U) := \inf\{\|j - i\| : i \in U\}$, а $\|u\| = \max_{q=1, \dots, d} |u_q|$.

Напомним (см., например, [21, с. 15]), что последовательность $(U_n)_{n \in \mathbb{N}}$ неограниченно растет в смысле ван Хова (или растет регулярно), если

$$|U_n| \rightarrow \infty \quad \text{и} \quad |\delta U_n|/|U_n| \rightarrow 0 \quad \text{при} \quad n \rightarrow \infty, \quad (8)$$

здесь $|U|$ обозначает мощность конечного множества U . Простейший пример множеств, удовлетворяющих условиям (8), — это совокупность «целочисленных параллелепипедов»

$$U_n = ((a_1^{(n)}, b_1^{(n)}] \times \dots \times (a_d^{(n)}, b_d^{(n)}]) \cap \mathbb{Z}^d$$

с $b_k^{(n)} - a_k^{(n)} \rightarrow \infty$ при $n \rightarrow \infty$ для $k = 1, \dots, d$.

В силу теоремы 1.12 [21, гл. 3] справедлива

Теорема 1. Пусть случайное поле W , введенное в (7), невырождено (это можно гарантировать выбором порога u , когда функция φ не равна тождественно нулю). Тогда для последовательности конечных множеств U_n , растущих в

смысле ван Хова, и величин $S(U_n) = \sum_{j \in U_n} W_j$, представляющих количество «поврежденных» FSUs, имеет место сходимость по распределению:

$$\frac{S(U_n) - \mathbb{E}S(U_n)}{\sigma|U_n|^{1/2}} \xrightarrow{\text{law}} N(0, 1), \quad (9)$$

если

$$\sigma^2 := \sum_{j \in \mathbb{Z}^d} \text{cov}(W_0, W_j) < \infty,$$

а $N(0, 1)$ обозначает стандартное нормальное распределение.

3. Статистическая версия ЦПТ

Обычно, если слагаемые зависимы, величина σ^2 , использованная в (9), неизвестна. Поэтому важно получить «статистическую версию» доказанной центральной предельной теоремы (ЦПТ), включающую случайную (само)нормализацию.

Для любых $r > 0$ и $j \in \mathbb{Z}^d$ обозначим

$$K_j(r) = \{i \in \mathbb{Z}^d : \|i - j\| \leq r\}.$$

Введем для конечного $U \subset \mathbb{Z}^d$

$$Q_j(U) = U \cap K_j(r), \quad r = r(U) > 0, \quad j \in U, \quad (10)$$

$$B(U) = \frac{1}{|U|} \sum_{j \in U} |Q_j(U)| \left(\frac{S(Q_j(U))}{|Q_j(U)|} - \frac{S(U)}{|U|} \right)^2, \quad (11)$$

если последнее выражение не равно нулю, и $B(U) = 1$ в противном случае. Очевидно, $Q_j(U)$ и $B(U)$ зависят и от r , но мы не станем усложнять запись.

Применяя следствие 1.6 [21, гл. 7], приходим к следующему утверждению.

Теорема 2. Пусть выполнены условия теоремы 1 и $(U_n)_{n \in \mathbb{N}}$ — последовательность регулярно растущих конечных подмножеств $\mathbb{Z}^d$. Предположим также, что $(r_n)_{n \in \mathbb{N}}$ — это такая последовательность положительных чисел, для которой $r_n \rightarrow \infty$, $r_n = O(|U_n|^{1/(2d)})$ и $r_n^d |\delta U_n| / |U_n| \rightarrow 0$ при $n \rightarrow \infty$. Тогда, выбирая $r(U_n) = r_n$ в (10) и (11), имеем

$$\frac{S(U_n) - |U_n| p_{\text{FSU}}}{(|U_n| B(U_n))^{1/2}} \xrightarrow{\text{law}} N(0, 1), \quad n \rightarrow \infty. \quad (12)$$

Соотношение (12) позволяет строить приближенные доверительные интервалы для p_{FSU} даже в более общей ситуации, чем (1), а именно при каждом $\varepsilon \in (0, 1)$ с вероятностью, близкой к $1 - \varepsilon$, при всех достаточно больших n имеем

$$|U_n|^{-1} S(U_n) - a_\varepsilon |U_n|^{-1/2} B(U_n)^{1/2} \leq p_{\text{FSU}} \leq |U_n|^{-1} S(U_n) + a_\varepsilon |U_n|^{-1/2} B(U_n)^{1/2}, \quad (13)$$

где Φ — функция распределения стандартного нормального закона, а $\Phi(a_\varepsilon) = 1 - \frac{\varepsilon}{2}$, $p_{\text{FSU}} = p_{\text{FSU}}(D)$, D — доза облучения. При дополнительных предположениях о стохастической модели результаты работы [22] позволяют получить оценки скорости сходимости к нормальному закону в (12) и тем самым уточнить (13). Кроме того, теория, развитая в [21], позволяет модифицировать теорему 1 так, чтобы она

относилась не к полю W , а к полю Z (то есть элементами поля принимались бы не только значения 0 и 1). Это дает возможность лучше описывать состояния биологической системы, подвергнутой облучению, поскольку наряду с состояниями «FSU убита» или «FSU жива» имеет смысл рассматривать промежуточные состояния (часть из которых даст FSU восстановиться).

4. Случайные поля на $\mathbb{R}^d$, двухмасштабные процедуры

В этом разделе, опираясь на подход, развитый в [20], мы изучаем случайные поля, заданные на пространстве $\mathbb{R}^d$ или подмножествах этого пространства. Наряду с множествами $V_n \subset \mathbb{R}^d$ ($n \in \mathbb{N}$), растущими в смысле ван Хова, будут рассматриваться наблюдения, образующие все более плотные (с увеличением n) дискретные подмножества $U_n \subset V_n$ (иначе говоря, берутся «infill asymptotics», см. [23]). Таким образом будет исследоваться двухмасштабная процедура.

Пусть $X = \{X_t, t \in \mathbb{R}^d\}$ — действительное, стационарное в широком смысле, случайное поле с ковариационной функцией

$$R(u) = \text{cov}(X_t, X_{t+u}), \quad u, t \in \mathbb{R}^d.$$

Для положительного Δ введем решетку $T(\Delta) := (\mathbb{Z}/\Delta)^d$, то есть совокупность точек вида $(j_1/\Delta, \dots, j_d/\Delta)$, где $j = (j_1, \dots, j_d) \in \mathbb{Z}^d$. Определим

$$S(U) = \sum_{t \in U} X_t, \quad U \subset \mathbb{R}^d, \quad |U| < \infty.$$

Для $\Delta_n > 0$ положим $T_n = T(\Delta_n)$, $n \in \mathbb{N}$.

Лемма. Пусть ковариационная функция процесса X непосредственно абсолютно интегрируема по Риману, а также

$$\sigma^2 := \int_{\mathbb{R}^d} R(u) du \neq 0. \quad (14)$$

Тогда для любой последовательности множеств $V_n \subset \mathbb{R}^d$, $V_n \rightarrow \infty$ по ван Хову, и любой последовательности положительных чисел $\Delta_n \rightarrow \infty$ имеем

$$\frac{\text{Var } S(U_n)}{c_n |U_n|} \rightarrow 1 \quad \text{при } n \rightarrow \infty, \quad (15)$$

где $U_n = V_n \cap T_n$ и $c_n = \sigma^2 \Delta_n^d$ ($n \in \mathbb{N}$).

Доказательство. Положим

$$b_n := \sum_{t \in T_n} \text{cov}(X_0, X_t) = \sum_{t \in T_n} R(t).$$

В силу гипотезы об интегрируемости R мы получаем $b_n \sim c_n$ при $n \rightarrow \infty$. Для $a > 0$ рассмотрим разбиение $\mathbb{R}^d$ с помощью кубов $B_j(a) = B_0(a) + (aj_1, \dots, aj_d)$, $j = (j_1, \dots, j_d) \in \mathbb{Z}^d$, где

$$B_0(a) = \{x \in \mathbb{R}^d : 0 < x_k \leq a, k = 1, \dots, d\}.$$

Введем $J_n^-(a) = \{j \in \mathbb{Z}^d : B_j(a) \subset V_n\}$, $J_n^+(a) = \{j \in \mathbb{Z}^d : B_j(a) \cap V_n \neq \emptyset\}$ и

$$V_n^-(a) = \bigcup_{j \in J_n^-(a)} B_j(a), \quad V_n^+(a) = \bigcup_{j \in J_n^+(a)} B_j(a).$$

Поскольку $V_n \rightarrow \infty$ в смысле ван Хова

$$\text{mes}(V_n^-(a)) \rightarrow \infty \quad \text{и} \quad \text{mes}(V_n^+(a))/\text{mes}(V_n^-(a)) \rightarrow 1, \quad n \rightarrow \infty. \quad (16)$$

Как обычно, мера Лебега (измеримого) множества B обозначается $\text{mes}(B)$. Для $U_n^-(a) = V_n^-(a) \cap T_n$ получаем

$$A_n := b_n |U_n^-(a)| - \text{Var } S(U_n^-(a)) = \sum_{s \in U_n^-(a), t \notin U_n^-(a)} \text{cov}(X_s, X_t).$$

Возьмем $p \in (0, a/2)$, и пусть $\Lambda_j(a, p) = (u_{j,1}, v_{j,1}] \times \dots \times (u_{j,d}, v_{j,d}]$ — это куб в $\mathbb{R}^d$ с тем же центром, что у $B_j(a)$ и длиной ребра $a - 2p$, $j \in \mathbb{Z}^d$. Положим

$$G_n = \bigcup_{j \in J_n^-(a)} ((B_j(a) \setminus \Lambda_j(a, p)) \cap T_n), \quad W_n = \bigcup_{j \in J_n^-(a)} (\Lambda_j(a, p) \cap T_n). \quad (17)$$

Очевидно, $G_n = G_n(a, p)$ и $W_n = W_n(a, p)$. Тогда для всех достаточно больших n $|A_n|(b_n |U_n^-(a)|)^{-1}$ допускает верхнюю оценку

$$\begin{aligned} \frac{1}{b_n |U_n^-(a)|} \left(\sum_{s \in G_n} \sum_{t \notin U_n^-(a)} |\text{cov}(X_s, X_t)| + \sum_{s \in W_n} \sum_{t \notin U_n^-(a)} |\text{cov}(X_s, X_t)| \right) \leq \\ \leq \frac{4}{\sigma^2} \left(\frac{pd}{a} \int_{\mathbb{R}^d} |R(u)| du + \int_{\|u\| \geq p} |R(u)| du \right), \end{aligned}$$

здесь $\|\cdot\|$ — максимальная норма в $\mathbb{R}^d$. Для любого $\varepsilon > 0$ можно выбрать достаточно большое p , а затем и a , чтобы получить $|A_n|(b_n |U_n^-(a)|)^{-1} < \varepsilon$, когда $n \geq N = N(\varepsilon, p, a)$. Далее,

$$\text{Var } S(U_n \setminus U_n^-(a)) \leq |U_n \setminus U_n^-(a)| \sum_{t \in T_n} |R(t)| \leq 2|U_n \setminus U_n^-(a)| \Delta_n^d \int_{\mathbb{R}^d} |R(u)| du \quad (18)$$

при всех достаточно больших n . Пусть B — куб (вида $(u, v]$, $u, v \in \mathbb{R}^d$) с длиной ребра $a > 0$. Тогда легко видеть, что $|B \cap T_n| \sim (a \Delta_n)^d$ при $n \rightarrow \infty$. В силу (16) и (18) получаем

$$\frac{\text{Var } S(U_n^-(a))}{b_n |U_n^-(a)|} \rightarrow 1, \quad n \rightarrow \infty,$$

то есть верно (15), что завершает доказательство. $\square$

Замечание 1. Соотношение (15) представляет собой аналог леммы, установленной Болтхаузенем [24] для стационарного случайного поля $X = \{X_j, j \in \mathbb{Z}^d\}$ и регулярно растущих множеств $U_n \subset \mathbb{Z}^d$.

Теперь мы обобщим понятие (BL, θ) -зависимости, введенное в статье [25] для случайных полей на решетке $\mathbb{Z}^d$, чтобы охватить случайные поля, заданные на $\mathbb{R}^d$.

Определение 1. Случайное поле $X = \{X_t, t \in \mathbb{R}^d\}$ называется (BL, θ) -зависимым, если найдется последовательность чисел $\theta_n \searrow 0$ при $n \rightarrow \infty$, такая, что для всех достаточно больших Δ , любых конечных непересекающихся множеств $I, J \subset T(\Delta)$ и произвольных ограниченных липшицевых функций $f : \mathbb{R}^{|I|} \rightarrow \mathbb{R}$, $g : \mathbb{R}^{|J|} \rightarrow \mathbb{R}$

$$|\text{cov}(f(X_s, s \in I), g(X_t, t \in J))| \leq \text{Lip}(f)\text{Lip}(g)(|I| \wedge |J|)\Delta^d \theta_r, \quad (19)$$

где $r = \text{dist}(I, J) := \min\{\|s - t\|, s \in I, t \in J\}$ и

$$\text{Lip}(f) := \sup_{x \neq y, x, y \in \mathbb{R}^d} \frac{|f(x) - f(y)|}{\sum_{k=1}^d |x_k - y_k|}.$$

В работе [26] было показано, что для положительно или отрицательно ассоциированного (РА- или НА-) поля $X = \{X_t, t \in T\}$ с конечным вторым моментом, любых непересекающихся конечных множеств $I, J \subset T$ и упомянутых выше функций f и g левая часть (19) допускает верхнюю оценку

$$\text{Lip}(f)\text{Lip}(g) \sum_{s \in I, t \in J} |\text{cov}(X_s, X_t)|. \quad (20)$$

В частности, для стационарного в широком смысле РА- или НА-поля X , имеющего непосредственно абсолютно интегрируемую в смысле Римана ковариационную функцию, в качестве θ_r в (19) можно использовать следующий аналог коэффициента Кокса–Гримметта:

$$\theta_r = 2 \int_{\{u \in \mathbb{R}^d : \|u\| \geq r\}} |R(u)| du, \quad r > 0.$$

Отметим, что целый ряд важных стохастических моделей, обладающих свойством РА или НА, а также модификациями этих свойств, рассмотрен в [21].

Теорема 3. Пусть X — строго стационарное, (BL, θ) -зависимое случайное поле с непрерывной ковариационной функцией R , удовлетворяющей условиям леммы. Тогда для любых множеств $V_n \rightarrow \infty$ в смысле ван Хова ($V_n \subset \mathbb{R}^d$) и произвольной последовательности $(\Delta_n)_{n \in \mathbb{N}}$, где $0 < \Delta_n \rightarrow \infty$, имеем

$$\frac{S(U_n) - |U_n| \mathbb{E} X_0}{\sqrt{\Delta_n^d |U_n|}} \xrightarrow{\text{law}} N(0, \sigma^2), \quad n \rightarrow \infty, \quad (21)$$

где $U_n = V_n \cap T_n$ и σ^2 определено согласно (13).

Доказательство. Без потери общности мы можем считать поле X центрированным. Покажем, что для каждого $\lambda \in \mathbb{R}$

$$\mathbb{E} \exp\{i\lambda (\Delta_n^d |U_n|)^{-1/2} S(U_n)\} \rightarrow \exp\{-\sigma^2 \lambda^2 / 2\} \quad \text{при } n \rightarrow \infty, \quad (22)$$

здесь $i^2 = -1$.

Фиксируем $a > 0$. Используя те же обозначения, что при доказательстве леммы, приходим к неравенству

$$\begin{aligned} & |\mathbb{E}\{i\lambda(\Delta_n^d|U_n|)^{-1/2}S(U_n)\} - \mathbb{E}\exp\{i\lambda(\Delta_n^d|U_n^-(a)|)^{-1/2}S(U_n^-(a))\}| \leq \\ & \leq |\lambda|\Delta_n^{-d/2}\left(|U_n|^{-1/2}\mathbb{E}|S(U_n) - S(U_n^-(a))| + \delta_n(a)\mathbb{E}|S(U_n^-(a))|\right), \end{aligned}$$

где $\delta_n(a) = ||U_n|^{-1/2} - |U_n^-(a)|^{-1/2}|$. Для всех достаточно больших n

$$(\mathbb{E}|S(U_n^-(a))|)^2 \leq \mathbb{E}(S(U_n^-(a)))^2 \leq 2\Delta_n^d|U_n^-(a)| \int_{\mathbb{R}^d} |R(u)|du. \quad (23)$$

В силу (16), (18) и (23) достаточно убедиться, что при каждом $\lambda \in \mathbb{R}$,

$$\mathbb{E}\exp\{i\lambda(\Delta_n^d|U_n^-(a)|)^{-1/2}S(U_n^-(a))\} \rightarrow \exp\{-\sigma^2\lambda^2/2\}, \quad n \rightarrow \infty. \quad (24)$$

Возьмем $p \in (0, a/2)$ и положим $Y_n(a, p) = (\Delta_n^d|W_n(a, p)|)^{-1/2}S(W_n(a, p))$, где $W_n(a, p)$ фигурирует в (17), $n \in \mathbb{N}$. Теперь мы будем рассматривать нормированные суммы, берущиеся по кубам $\Lambda_j(a, p)$, $j \in J_n^-(a)$. Те же самые рассуждения, что были использованы при доказательстве леммы, приводят для любого $\varepsilon > 0$ и произвольного фиксированного $\lambda \in \mathbb{R}$ к оценке

$$\left| \mathbb{E}\exp\{i\lambda(\Delta_n^d|U_n^-(a)|)^{-1/2}S(U_n^-(a))\} - \mathbb{E}\exp\{i\lambda Y_n(a, p)\} \right| < \varepsilon,$$

выполняющейся, если p/a достаточно мало и n достаточно велико.

Для любого $L = (v, v + a - 2p] \subset \mathbb{R}$ и всех достаточно больших n найдется $M_n = [(a - 2p)\Delta_n]$ точек $v_m = z + m\Delta_n^{-1} \in L$, $m = 1, \dots, M_n$, где $z = q\Delta_n^{-1}$ при некотором $q \in \mathbb{Z}$ (здесь $[\cdot]$ обозначает целую часть числа). Для каждого $j \in \mathbb{Z}^d$ возьмем куб $\Gamma_j(a, p) = \prod_{l=1}^d (z_l, z_l + M_n\Delta_n^{-1}] \subset \Lambda_j(a, p)$, $(z_1, \dots, z_d) \in T_n$. Тогда $|\Gamma_j(a, p) \cap T_n| = M_n^d$, $j \in \mathbb{Z}^d$. Положим $N_n = |J_n^-(a)|$ и перенумеруем множество случайных величин

$$\left\{ (\Delta_n^d|\Gamma_j(a, p) \cap T_n|)^{-1/2}S(\Gamma_j(a, p) \cap T_n), \quad j \in J_n^-(a) \right\}.$$

Тем самым задается набор $\zeta_{n,1}, \dots, \zeta_{n,N_n}$ ($N_n = N_n(a, V_n)$, $\zeta_{n,k} = \zeta_{n,k}(a, p, \Delta_n, V_n)$). Для любого фиксированного $\lambda \in \mathbb{R}^d$ получаем

$$\mathbb{E}\exp\{i\lambda Y_n\} - \mathbb{E}\{i\lambda N_n^{-1/2}Z_n\} \rightarrow 0, \quad n \rightarrow \infty,$$

где $Z_n := \sum_{k=1}^{N_n} \zeta_{n,k}$. В силу (19) при всех $\varepsilon > 0$, $n \in \mathbb{N}$ и $\lambda \in \mathbb{R}$ имеем

$$\begin{aligned} & \left| \mathbb{E}\{i\lambda N_n^{-1/2}Z_n\} - \prod_{k=1}^{N_n} \mathbb{E}\exp\{i\lambda N_n^{-1/2}\zeta_{n,k}\} \right| \leq \\ & \leq \sum_{q=1}^{N_n} \left| \text{cov}(\exp\{i\lambda N_n^{-1/2}\zeta_{n,q}\}, \exp\{-i\lambda N_n^{-1/2} \sum_{k=q+1}^{N_n} \zeta_{n,k}\}) \right| \leq 4\lambda^2\theta_p < \varepsilon, \end{aligned}$$

если p достаточно велико. Пользуясь строгой стационарностью X , мы видим, что для каждого $n \in \mathbb{N}$ случайные величины $\zeta_{n,1}, \dots, \zeta_{n,N_n}$ одинаково распределены. Применяя условие Линдеберга к независимым копиям этих величин, можно удостоверить, что $N_n^{-1/2}Z_n$ сходятся по распределению при $n \rightarrow \infty$ к гауссовской величине с нулевым средним и дисперсией, достаточно близкой к σ^2 при подходящем выборе a и p . Отсюда вытекает требуемый результат. $\square$

Если найдется последовательность $(\hat{\sigma}_n^2(U_n))_{n \in \mathbb{N}}$ состоятельных оценок величины $\sigma^2 \neq 0$, где $\hat{\sigma}_n^2(U_n)$ строится по наблюдениям $X_t, t \in U_n$, то при выполнении условий теоремы 1 вместо (21) можно использовать следующий (статистический) вариант ЦПТ:

$$\frac{S(U_n) - nEX_0}{\hat{\sigma}(U_n)\Delta_n^{d/2}|U_n|^{1/2}} \xrightarrow{\text{law}} N(0, 1), \quad n \rightarrow \infty \quad (25)$$

(для $\hat{\sigma}(U_n) = 0$ формально полагаем $z/0 := 0, z \in \mathbb{R}$). Таким образом, возникает возможность строить приближенный доверительный интервал для неизвестного среднего значения EX_0 .

Для точки $t \in T_n$ ($n \in \mathbb{N}$) и любого $r > 0$ введем

$$K_t(r) = \{s \in T_n : \|s - t\| \leq r\}.$$

Конечному множеству $U_n \subset T_n$ и положительному r_n сопоставим статистику

$$\hat{\sigma}(U_n)^2 = (\Delta_n^d|U_n|)^{-1} \sum_{t \in U_n} |Q_t| \left(\frac{S(Q_t)}{|Q_t|} - \frac{S(U_n)}{|U_n|} \right)^2, \quad (26)$$

где $Q_t = K_t(r_n) \cap U_n$ (эти Q_t и $\hat{\sigma}(U_n)$ зависят от U_n и r_n), $n \in \mathbb{N}$.

Теорема 4. Пусть X — строго стационарное, (BL, θ) -зависимое поле, имеющее непрерывную ковариационную функцию R и удовлетворяющее условиям леммы. Тогда для любых множеств $V_n \rightarrow \infty$ в смысле ван Хова ($V_n \subset \mathbb{R}^d$) и для любой последовательности $(\Delta_n)_{n \in \mathbb{N}}$, $0 < \Delta_n \rightarrow \infty$, найдется последовательность положительных чисел $(r_n)_{n \in \mathbb{N}}$, таких, что соотношение (25) справедливо с $\hat{\sigma}(U_n)$, заданными в (26).

Доказательство. В силу (16) мы можем найти последовательность $(a_n)_{n \in \mathbb{N}}$, такую, что $0 < a_n \rightarrow \infty$ при $n \rightarrow \infty$ и

$$\text{mes}(V_n^-(a_n)) \rightarrow \infty, \quad \text{mes}(V_n^+(a_n))/\text{mes}(V_n^-(a_n)) \rightarrow 1, \quad n \rightarrow \infty.$$

Рассмотрим вначале $V_n = V_n^-(a_n)$ и соответственно $U_n = U_n^-(a_n)$, где $U_n^-(a_n) = V_n^-(a_n) \cap T_n$, а чтобы упростить формулы, пишем $\hat{\sigma}_n$ вместо $\hat{\sigma}(U_n^-(a_n))$, $n \in \mathbb{N}$.

Заметим, что

$$\begin{aligned} \mathbb{E}|\hat{\sigma}_n^2 - \sigma^2| &\leq (\Delta_n^d|U_n|)^{-1} \mathbb{E} \left| \sum_{t \in U_n} |Q_t| \left(\left(\frac{S(Q_t)}{|Q_t|} - \frac{S(U_n)}{|U_n|} \right)^2 - \left(\frac{S(Q_t)}{|Q_t|} \right)^2 \right) \right| + \\ &+ (\Delta_n^d|U_n|)^{-1} \mathbb{E} \left| \sum_{t \in U_n} |Q_t|^{-1} (S(Q_t)^2 - \mathbb{E}(S(Q_t)^2)) \right| + \end{aligned}$$

$$+ \left| (\Delta_n^d |U_n|)^{-1} \sum_{t \in U_n} |Q_t|^{-1} \mathbb{E}(S(Q_t))^2 - \sigma^2 \right| =: R_{n,1} + R_{n,2} + R_{n,3}.$$

Для $r > 0$ положим $U_n(r) = U_n \setminus (T_n \setminus U_n)^{(r)}$, где $G^{(r)}$ обозначает r -окрестность конечного множества $G \subset T_n$, то есть $G^{(r)} := \{t \in T_n : \min_{s \in G} \|t - s\| < r\}$.

Наиболее простая часть доказательства состоит в проверке того, что $R_{n,3} \rightarrow 0$, $n \rightarrow \infty$. Пусть $r_n < a_n/4$, $n \in \mathbb{N}$. В силу стационарности X получаем, что $\mathbb{E}(S(Q_t(r_n)))^2 = \mathbb{E}(S(K_0(r_n)))^2$ для $t \in U_n(2r_n)$. Следовательно,

$$\begin{aligned} & (\Delta_n^d |U_n|)^{-1} \sum_{t \in U_n} |Q_t|^{-1} \mathbb{E}(S(Q_t))^2 \leq \\ & \leq \frac{\mathbb{E}(S(K_0(r_n)))^2}{|\Delta_n^d |K_0(r_n)||} + (\Delta_n^d |U_n|)^{-1} \sum_{t \in U_n \setminus U_n(2r_n)} |Q_t|^{-1} \mathbb{E}(S(Q_t))^2 =: L_{n,1} + L_{n,2}. \end{aligned}$$

По лемме имеем $L_{n,1} \rightarrow \sigma^2$ при $r_n \rightarrow \infty$. Для всех достаточно больших n

$$L_{n,2} \leq 2|U_n|^{-1} |U_n \setminus U_n(2r_n)| \int_{\mathbb{R}^d} |R(u)| du.$$

Применяя рассуждения, использованные при доказательстве теоремы 1, мы видим, что $|U_n|^{-1} |U_n \setminus U_n(2r_n)| \leq 4dr_n/a_n$ при всех достаточно больших n . Теперь мы выберем $r_n = o(a_n)$, $n \rightarrow \infty$.

Для $M > 0$ и $t \in T_n$ ($n \in \mathbb{N}$) введем $\xi_t = (\Delta_n^d |Q_t|)^{-1/2} S(Q_t)$ и определим вспомогательные случайные величины $\nu_t = H_M^2(\xi_t)$, $Y_t = \nu_t - \mathbb{E}\nu_t$, $Z_t = \xi_t^2 - \mathbb{E}\xi_t^2 - Y_t$, где

$$H_M(x) = -M\mathbb{I}\{x < -M\} + x\mathbb{I}\{|x| \leq M\} + M\mathbb{I}\{x > M\}, \quad x \in \mathbb{R},$$

и $\mathbb{I}\{A\}$ – индикатор множества A . Тогда

$$R_{n,2} \leq |U_n|^{-1} \mathbb{E} \left| \sum_{t \in U_n} Y_t \right| + |U_n|^{-1} \mathbb{E} \left| \sum_{t \in U_n} Z_t \right| =: R_{n,2}^{(1)} + R_{n,2}^{(2)}.$$

Далее,

$$\begin{aligned} & (R_{n,2}^{(1)})^2 \leq |U_n|^{-2} \sum_{s, t \in U_n} \text{cov}(\nu_s, \nu_t) \leq \\ & \leq |U_n|^{-2} \sum_{s \in U_n} \sum_{t \in U_n, \|s-t\| \leq 3r_n} |\text{cov}(\nu_s, \nu_t)| + |U_n|^{-2} \sum_{s \in U_n} \sum_{t \in U_n, \|s-t\| > 3r_n} |\text{cov}(\nu_s, \nu_t)|. \end{aligned}$$

Первое слагаемое в правой части последнего неравенства допускает верхнюю оценку $M^4 |U_n|^{-1} (6r_n + 1)^d \rightarrow 0$ при $r_n = o(a_n)$, $n \rightarrow \infty$. Чтобы оценить второе слагаемое, мы используем соотношение (20). Имеем

$$\begin{aligned} & J_n := \Delta_n^{-d} |U_n|^{-2} \sum_{s \in U_n} \sum_{t \in U_n: \|s-t\| > 3r_n} |\text{cov}(\nu_s, \nu_t)| \leq \\ & \leq |U_n|^{-2} \sum_{s \in U_n} \sum_{t \in U_n: \|s-t\| > 3r_n} 4M^2 |Q_s|^{-1/2} |Q_t|^{-1/2} (|Q_s| \wedge |Q_t|) \theta_{\|s-t\|-2r_n} \leq 4M^2 \theta_{r_n} \rightarrow 0. \end{aligned}$$

Для $M > 0$ и всех достаточно больших n

$$\begin{aligned} R_{n,2}^{(2)} &\leq |U_n|^{-1} \sum_{t \in U_n(2r_n)} \mathbb{E}|Z_t| + |U_n|^{-1} \sum_{t \in U_n \setminus U_n(2r_n)} \mathbb{E}|Z_t| \leq \\ &\leq 2\mathbb{E} \left(\frac{S(K_0(r_n))^2}{\Delta_n^d |K_0(r_n)|} \mathbb{I} \left\{ \frac{S(K_0(r_n))^2}{\Delta_n^d |K_0(r_n)|} \geq M^2 \right\} \right) + 4 \frac{|U_n \setminus U_n(2r_n)|}{|U_n|} \int_{\mathbb{R}^d} |R(u)| du. \end{aligned}$$

Из теоремы 2 вытекает равномерная интегрируемость семейства случайных величин $\{S(K_0(r_n))^2/|K_0(r_n)|, n \in \mathbb{N}\}$. Поэтому $R_{n,2}^{(2)}$ может быть сделано сколь угодно малым при соответствующем выборе M .

В силу леммы для всех достаточно больших n имеем

$$\begin{aligned} R_{n,1} &\leq \Delta_n^{-d} |U_n|^{-3} \mathbb{E} S(U_n)^2 \sum_{t \in U_n} |Q_t| + 2\Delta_n^{-d} |U_n|^{-2} \mathbb{E} |S(U_n)| \sum_{t \in U_n} |S(Q_t)| \leq \\ &\leq 2\sigma^2 |U_n|^{-2} \sum_{t \in U_n} |Q_t| + 2\Delta_n^{-d} |U_n|^{-2} \sum_{t \in U_n} (\mathbb{E}(S(U_n))^2 \mathbb{E}(Q_t))^2)^{1/2} \leq \\ &\leq 2\sigma^2 |U_n|^{-1} |K_0(r_n)| + 4 \left(\sigma^2 |U_n|^{-1} |K_0(r_n)| \int_{\mathbb{R}^d} |R(u)| du \right)^{1/2}. \end{aligned}$$

Итак, $R_{n,1} + R_{n,2} + R_{n,3} \rightarrow 0$ при $n \rightarrow \infty$.

Чтобы завершить доказательство в общем случае для $V_n \rightarrow \infty$ в смысле ван Хофа, нетрудно проверить, что

$$\mathbb{E} |\hat{\sigma}(U_n)^2 - \hat{\sigma}(U_n^-(a_n))^2| \rightarrow 0 \text{ при } n \rightarrow \infty,$$

когда a_n и r_n выбраны должным образом. $\square$

Замечание 2. Теорема 4 показывает, что для зависимых FSUs аппроксимация (4), вообще говоря, заменится другой, приводящей к аналогу соотношения (13). Если выполнены условия этой теоремы, то в важном частном случае, когда $V_n = (u_n, v_n]^d$ и $v_n - u_n \rightarrow \infty$, можно выбрать $r_n = o(v_n - u_n)$ при $n \rightarrow \infty$.

5. Кластерные модели

В последние несколько лет было сделано важное наблюдение, касающееся пространственной конфигурации пораженных клеток или функциональных единиц, образующих облучаемый орган (ткань) (см., например, [27, 28]). Основная идея заключается в том, чтобы учитывать размер максимального кластера, образованного (в рассматриваемом объеме) исследуемыми клетками (или FSUs).

Имеются различные способы описания геометрических характеристик пораженных клеточных образований. Возьмем разбиение $\mathbb{R}^d$, порожденное кубами $C_j := (j - 1/2, j + 1/2], j \in \mathbb{Z}^d$. Говорят, что кубы C_i и C_j являются соседями, если у них имеется общая $(d - 1)$ -мерная грань, и что путь — это цепочка кубов $C(1), \dots, C(r)$, таких, что $C(m)$ и $C(m + 1)$ — соседи для $m = 1, \dots, r - 1$. Множество $V = \cup_{j \in J} C_j$, где $J \subset \mathbb{Z}^d$, называется связным, если для любых $i, j \in J$ найдется путь в V , соединяющий кубы C_i и C_j . Максимальный кластер в V — это его компонента (компоненты) связности, имеющая наибольшую мощность.

Для $R \in \mathbb{N}$, $j \in \mathbb{Z}^d$ и $B_R(j) = \{i \in \mathbb{Z}^d : \|i - j\| \leq R\}$ введем множество $V_R(j) := \cup_{i \in B_R(j)} C_i$. Пусть независимые случайные величины $Y_j, j \in \mathbb{Z}^d$, имеют распределение, указанное в (6), а $L(V_R(j))$ — мощность максимального кластера множества $\cup_{\{i: C_i \in V_R(j), Y_i=1\}} C_i$. Теперь мы можем предположить, что случайная величина W_j (описывающая состояние j -й FSU) принимает значение 1, если для $u \geq 0$ и $R \in \mathbb{N}$

$$S(V_R(j)) = \sum_{i: C_i \in V_R(j)} Y_i \geq u \text{ и } L(V_R(j)) \geq K, \quad (27)$$

где $0 \leq K \leq |V_R(j)| = (2R+1)^d$; в противном случае полагаем $W_j = 0$. Очевидно, $S(V_R(j)) \geq L(V_R(j))$. Мы одновременно учитываем пораженный объем множества V_R (то есть рассматриваем кубы $C_j \subset V_R$, для которых $Y_j = 1$), а также объем максимального кластера, образованного этими элементами. Точнее говоря, мы предлагаем «локальный геометрический подход», основанный на учете как общего «пораженного объема» в множестве V_R , так и объема максимального кластера, возникающего в окрестности функциональной единицы.

Для так построенного поля W теперь можно использовать модель критического объема. Другой путь состоит в применении ренорм-групп. Для простоты предположим, что V — параллелепипед, состоящий из введенных выше единичных кубов. Пусть длина каждого из ребер V равна $(2R+1)^{m_q}$, где $m_q \in \mathbb{N}$ и $q = 1, \dots, d$. Разделим каждое ребро V на интервалы длины R и образуем разбиение этого параллелепипеда на кубы с ребрами длины $2R+1$. Затем образуем новую конфигурацию (новое случайное поле), если в центр каждого из кубов разбиения B_i , имеющего объем $(2R+1)^d$, поместим случайную величину W_i , значение которой определяется согласно (27) с B_i вместо $V_R(j)$. Описанную процедуру можно повторять (изменяя масштаб координатных осей) неоднократно.

Теперь мы обратимся к частному случаю, когда $R = 1$, то есть базовая окрестность состоит из 3^d единичных кубов. Для $d = 2$ получаем следующую таблицу:

L	S									
	0	1	2	3	4	5	6	7	8	9
0	1	0	0	0	0	0	0	0	0	0
1	0	9	24	22	6	1	0	0	0	0
2	0	0	12	40	40	8	0	0	0	0
3	0	0	0	22	44	32	4	0	0	0
4	0	0	0	0	36	36	12	0	0	0
5	0	0	0	0	0	49	20	0	0	0
6	0	0	0	0	0	0	48	4	0	0
7	0	0	0	0	0	0	0	32	0	0
8	0	0	0	0	0	0	0	0	9	0
9	0	0	0	0	0	0	0	0	0	1

Элемент $a_{l,k}$ этой таблицы представляет собой коэффициент, фигурирующий в формуле

$$P(L(V_1(j)) = l, S(V_1(j)) = k) = a_{l,k} p_{\text{FSU}}^k (1 - p_{\text{FSU}})^{9-k}.$$

Интересно сравнить, например, $P(S(V_1(j)) \geq 6, L(V_1(j)) \geq 4)$ и $P(L(V_1(j)) \geq 5)$. Существует $p_{\text{FSU}} = 0,8032\dots$, при котором эти вероятности совпадают. Мы видим, что даже в простейшем случае оказывается нелегко отдать предпочтение какой-то одной из введенных «локальных моделей».

6. Заключительные замечания

Рассмотренная выше модель зависимых FSUs допускает взаимосвязь коллективного поведения изучаемых клеток (клеточных образований), но при этом, в частности, охватывает и случай их независимого функционирования. При этом не предполагается, что радиус взаимодействия частей локальный (финитный). Привлекательность этого подхода состоит в том, что при весьма широких условиях зависимости изучаемого случайного поля (полей) удастся строить приближенные доверительные интервалы для $p_{\text{FSU}}(D)$. Кроме того, с помощью предельных теорем для сумм зависимых мультииндексированных слагаемых возможно рассматривать ситуации, описываемые не только бинарными величинами. Более того, соотношение (25) при определенных условиях допускает различные уточнения.

Новое направление исследований, оперирующее максимальными кластерами и фактически базирующееся на идеях теории перколяции, представляется весьма перспективным. Однако даже в простейших ситуациях, как показано выше, бывает нелегко сказать, следует принимать во внимание лишь размер максимального кластера или учитывать объем (площадь) пораженной части облучаемого органа или ткани. Далее можно было бы исследовать другие показатели, описывающие итоговое состояние биологической системы, занимающей множество $V \subset \mathbb{R}^d$, например, принимать во внимание $\max_{j \in V} L(V_R(j))$ или «глобальные показатели» (весь «пораженный объем», максимальный кластер, характеристики Эйлера и др.).

Отметим также, что имеются возможности строить иные стохастические модели «дефектных частей» множества $V \subset \mathbb{R}^d$, используя теорию случайных множеств, теорию графов, а также более сложные модели перколяции, например, модель Поттса или модель цветной перколяции. Кроме того, нерешенной проблемой является построение динамической стохастической модели, дающей адекватное пространственно-временное описание поведения системы, подвергнутой облучению. Речь идет об одновременном учете множества сложных факторов, характеризующих (взаимосвязанные) процессы гибели, миграции и восстановления клеток биологической системы, который позволяет применять эти модели к неоднородным группам пациентов. Более того, до настоящего времени остается весьма широкий простор в выборе целевых функций, описывающих ту или иную схему воздействия на живые организмы.

Работа выполнена при частичной поддержке гранта РФФИ № 07-01-00373а.

Список литературы

- [1] Radiology Centennial, Inc., A Century of Radiobiology. [http : //www.xray.hmc.psu.edu/rci/centennial.html](http://www.xray.hmc.psu.edu/rci/centennial.html), 1993.
- [2] Alber M., Nüsslin F. An objective function for radiation treatment optimization based on local biological measures // Phys. Med. Biol. 1999. **44**. P. 479–493.

- [3] Alber M., Nüsslin F. Tools for the analysis of dose optimization: I. Effect-volume histogram // *Phys. Med. Biol.* 1999. **47**. P. 2451–2458.
- [4] Honoré H. B., Bentzen S. M. A modelling study of the potential influence of low dose hypersensitivity on radiation treatment planning // *Radiotherapy and Oncology*. 2006. **79**. P. 115–121.
- [5] Withers H. R., Taylor J. M., Maciejewski B. Treatment volume and tissue tolerance // *Int. J. Radiat. Oncol. Biol. Phys.* 1988. **14**. P. 751–759.
- [6] Jackson A., Kutcher G. J., Yorke E. D. Probability of radiation-induced complications for normal tissues with parallel architecture subject to nonuniform irradiation // *Med. Phys.* 1993. **20**. P. 613–625.
- [7] Niemierko A., Goitein M. Modeling of normal tissue-response to radiation — the critical volume model // *Int. J. Radiat. Oncol. Biol. Phys.* 1993. **25**. P. 135–145.
- [8] Yorke E. D., Kutcher G. J., Jackson A., Ling C. C. Probability of radiation-induced complications in normal tissues with parallel architecture under conditions of uniform whole or partial organ irradiation // *Radiotherapy Oncology*. 1993. **26**. P. 226–237.
- [9] Stavrev P., Stavreva N., Niemierko A., Goitein M. Generalization of a model of tissue response to radiation based on the idea of functional subunits and binomial statistics // *Phys. Med. Biol.* 2001. **46**, 1501–1518.
- [10] Stavreva N., Niemierko A., Stavrev P., Goitein M. Modelling the dose-volume response of the spinal cord, based on the idea of damage to contiguous functional subunits // *Int. J. Radiat. Oncol. Biol. Phys.* 2001. **77**. P. 695–702.
- [11] Miles E. A., Clark C. H., Urbano T. G., Bidmead M., Dearnaley D. P., Harrington K. J., A'Hern R., Nutting C. M. The impact of introducing intensity modulated radiotherapy into routine clinical practice // *Radiotherapy and Oncology*. 2005. **77**. P. 241–246.
- [12] Bakin C. R., Clements R., Gately A., Butgell G. J. A separated primary and scatter model for independent dose calculation of intensity modulated radiotherapy // *Radiotherapy and Oncology*. 2006. **80**. P. 385–390.
- [13] Zaider M., Minerbo G. N. Tumor control probability. A formulation applicable to any temporal protocol of dose delivery // *Phys. Med. Biol.* 2000. **45**. P. 279–293.
- [14] Warkentin B., Stavrev P., Stavreva N., Fleid C., Fallone B. G. A TCP-NTCP estimation module using DVH and known radiobiological models and parameter sets // *J. App. Clinical Med. Phys.* 2004. **6**. P. 50–63.
- [15] Van Luijk P., Bul H. P., Konings A. W. T., van der Kogel A. J., Schippers J. M. Data on dose-volume effects in the rat spinal cord do not support existing models // *Int. J. Radiat. Oncol. Biol. Phys.* 2005. **61**. P. 892–900.
- [16] Tanderup K., Olsen D. R., Grau C. Dose painting: Art or science? // *Radiotherapy and Oncology*. 2006. **79**. P. 246–248.
- [17] Merks R. M. H., Glazier J. A. A cell-centered approach to developmental biology // *Physica A*. 2005. **352**. P. 113–130.

- [18] *Bėlik P., Dayananda P. W. A., Kemper J. T., Shvartsman M. M.* A stochastic model for PSA levels: behavior of solutions and population statistics // *J. Math. Biol.* 2006. **53**. P. 437–463.
- [19] *Bulinski A., Khrennikov A.* Generalization of the Critical Volume NTCP Model in the Radiobiology. 2005. <http://arxiv.org/abs/math/0504225>.
- [20] *Bulinski A.* CLT for random fields with applications // *Advances in Data Analysis* / Ed. by C.H. Skiadas. Boston: Birkhäuser, 2008. P. 1–10 (to appear).
- [21] *Булинский А. В., Шапкин А. П.* Предельные теоремы для ассоциированных случайных полей и родственных систем. М.: ФИЗМАТЛИТ 2008.
- [22] *Bulinski A., Kryzhanovskaya N.* Convergence rate in CLT for vector-valued random fields with self-normalization // *Probab. Math. Statist.* 2006. **26**. P. 261–281.
- [23] *Cressie N. A. C.* Statistics for Spatial Data. N.Y.: Wiley, 1991.
- [24] *Bolthausen E.* On the central limit theorem for stationary mixing random fields // *Ann. Probab.* 1982. **10**. P. 1047–1050.
- [25] *Bulinski A., Suquet C.* Normal approximation for quasi associated random fields // *Stat. Probab. Lett.* 2001. **54**. P. 215–226.
- [26] *Булинский А., Шабанович Э.* Асимптотическое поведение некоторых функционалов от положительно и отрицательно зависимых случайных полей // *Фунд. и прикл. матем.* 1998. **4**. P. 479–492.
- [27] *Thamse H. D., Zhang M., Tucker S. L., Liu H. H., Dong L., Mohan R.* Cluster models of dose-volume effects // *Int. J. Radiat. Oncol. Biol. Phys.* 2004. **59**. P. 1491–1504.
- [28] *Tucker S. L., Zhang M., Dong L., Mohan R., Kuban D., Thames H. D.* Cluster model analysis of late rectal bleeding after IMPT of prostate cancer: a case-control study // *Int. J. Radiat. Oncol. Biol. Phys.* 2006. **64**. P. 1255–1264.

О НЕКОТОРЫХ ЗАДАЧАХ ТЕОРИИ РИСКА

О. П. Виноградов

Устанавливается связь между теорией риска, теорией ветвящихся процессов и задачей о баллотировке, что дает возможность переносить теоремы из одного из этих разделов теории вероятностей в другой. Например, классическая теорема Крамера–Лундберга утверждает, что условие Крамера является достаточным для асимптотически экспоненциального убывания вероятности разорения. Соответствующая теорема из теории ветвящихся процессов позволяет утверждать, что это условие является и необходимым. Предложен метод получения неравенств для некоторых характеристик. Введено понятие неоднородного процесса и получены некоторые явные формулы. Рандомизируя параметры в этих формулах, можно получить результаты для достаточно широкого класса процессов риска.

1. Введение

В современной теории вероятностей значительное внимание уделяется задачам, которые возникают при анализе функционирования финансовых институтов (страховых компаний, бирж, банков и др.). Мы будем рассматривать ряд вероятностных задач, связанных с работой страховых компаний, в частности нас будут интересовать вероятность разорения страховой компании, распределение момента разорения и другие вероятностные характеристики. Этим задачам посвящена очень большая литература. Часть прикладной теории вероятностей, которая изучает эту проблематику, принято называть теорией риска.

2. Основные определения

Предположим, что $x \geq 0$ — начальный капитал страховой компании. Пусть $T_1, T_2, \dots$ ($0 \leq T_1 \leq T_2 \leq \dots$) — моменты, в которые происходят страховые случаи (аварии, смерти и т.д.), и пусть Z_i — размер выплаты в момент T_i , то есть сумма, которую выплачивает страховая компания клиентам для возмещения их потерь.

Положим $\Delta_1 = [0, T_1]$, $|\Delta_1| = T_1$; $\Delta_i = [T_{i-1}, T_i]$, $|\Delta_i| = T_i - T_{i-1}$, $i \geq 2$.

За счет покупки у страховой компании клиентами страховых полисов ее капитал с течением времени растет. Обычно предполагается [1], что на каждом из отрезков Δ_i рост ее капитала линеен с угловым коэффициентом $c \geq 0$, одинаковым для всех интервалов Δ_i , хотя в настоящей работе будет рассмотрен и более общий случай. Важной характеристикой качества функционирования страховой компании является вероятность $\psi(x, t)$ — вероятность ее разорения на отрезке времени $[0, t]$ при начальном капитале x (разорение страховой компании происходит тогда и только тогда, когда ее капитал $X(t)$ становится отрицательным). Введем случайную величину $\tau(x)$ — момент разорения страховой компании, которая имеет начальный капитал x .

Пусть $X(y)$ — капитал компании в момент y . Из предыдущих предположений вытекает, что $X(y) = x + cy - \sum_{j \leq N(y)} Z_j$, где $N(y)$ — число выплат компании

клиентам на отрезке $[0, y]$. Таким образом,

$$\psi(x, t) = P(\tau(x) < t) = P\left(\inf_{0 \leq y \leq t} X(y) < 0 | X(0) = x\right).$$

В теории риска вероятность $\psi(x, \infty) = \psi(x)$ называют вероятностью разорения.

Пусть $\eta(x) = \min(k : X(T_k + 0) < 0)$ — номер той выплаты, после которой капитал страховой компании впервые становится отрицательным, то есть компания разоряется. Таким образом,

$$\tau(x) = T_{\eta(x)} = |\Delta_1| + |\Delta_2| + \dots + |\Delta_{\eta(x)}|.$$

3. Связь между теорией риска и теорией ветвящихся процессов

Заметим, что между некоторыми задачами в различных областях теории вероятности существует тесная связь. Например, задача нахождения вероятности разорения эквивалентна задаче нахождения распределения времени ожидания в теории массового обслуживания, а также задаче нахождения распределения максимума случайного блуждания [1]. В работе автора [2] установлена связь между теорией риска и теорией ветвящихся процессов. Остановимся на этой связи более подробно.

Предположим, что моменты наступления страховых случаев $\{T_n\}$ образуют пуассоновский поток с параметром λ , и пусть величины выплат Z_n — независимые, одинаково распределенные случайные величины с произвольной функцией распределения $F(t)$. Пусть также $\alpha = \frac{\lambda m}{c}$, где $m = MZ_n$ — математическое ожидание величины произвольной выплаты.

Если $\alpha \geq 1$, то $\psi(x) \equiv 1$, если же $\alpha < 1$, то $\psi(x) \rightarrow 0$ при $x \rightarrow \infty$. Это следует из того факта, что математическое ожидание прироста капитала страховой компании между двумя последовательными выплатами равно $\frac{c}{\lambda} - m = \frac{c}{\lambda}(1 - \alpha)$.

В дальнейшем всегда предполагается, что $\alpha < 1$. Очевидно, что функция $g(x) = \frac{1-F(x)}{m}$ является вероятностной плотностью некоторой неотрицательной случайной величины. Обозначим соответствующую ей функцию распределения через $G(x)$. Известно (см., например, [3]), что вероятность разорения $\psi(x)$ удовлетворяет интегральному уравнению типа свертки

$$\psi(x) = \alpha \int_0^x \psi(x-u) g(u) du + \alpha(1-G(x)). \quad (1)$$

Дадим теперь краткое описание ветвящегося процесса с превращениями, зависящими от возраста. Пусть частица, существующая в момент $x = 0$, имеет распределение продолжительности жизни, задаваемое функцией распределения $A(x)$ и плотностью $a(x)$. В конце своей жизни она умирает и порождает n ($n \geq 0$) частиц с вероятностью q_n . Эти новые частицы являются вероятностными копиями первоначальной частицы, и процесс продолжается до тех пор, пока существует хотя бы одна частица. Со строгим определением этого процесса можно ознакомиться в монографии [4].

Пусть $a = \sum_{n=0}^{\infty} nq_n$ — математическое ожидание числа частиц, которое порождает одна частица, и пусть $\xi(x)$ — число частиц в момент x . Положим $M(x) = M\xi(x)$. Известно [4], что $M(x)$ удовлетворяет интегральному уравнению

$$M(x) = a \int_0^x M(x-u) a(u) du + 1 - A(x). \quad (2)$$

Если рассмотреть ветвящийся процесс с зависящими от возраста превращениями, в котором выполнены равенства $A(x) = G(x)$, $a = \alpha$, то нетрудно видеть [5], что из сравнения (1) и (2) вытекает равенство

$$\psi(x) = \alpha M(x). \quad (3)$$

Это равенство является «мостиком», который связывает теорию риска и теорию ветвящихся процессов. Поэтому некоторые результаты, полученные в теории ветвящихся процессов, могут быть в силу равенства (3) перенесены в теорию риска.

4. Условие Крамера является необходимым условием для асимптотической экспоненциальности вероятности разорения

Одним из самых известных результатов в теории риска является так называемая аппроксимация Крамера–Лундберга. Известный шведский математик Асмуссен в своей обширной монографии, посвященной теории риска [6], пишет: «Аппроксимация Крамера–Лундберга является одним из наиболее знаменитых результатов теории риска (и в целом теории вероятностей)».

Поясним, что понимается под аппроксимацией Крамера–Лундберга.

Пусть $B(x)$ — функция распределения некоторой неотрицательной случайной величины, $b(x)$ — соответствующая вероятностная плотность и k — некоторая постоянная ($0 < k < 1$).

Определение 1. Будем говорить, что пара $(k, B(x))$ удовлетворяет условию Крамера, если существует такое $R > 0$, для которого выполнено равенство

$$1 = k \int_0^{\infty} e^{Ru} b(u) du.$$

Условие Крамера, грубо говоря, означает, что $b(x)$ убывает достаточно быстро на бесконечности.

Постоянную R иногда называют постоянной Крамера–Лундберга или подстроеным коэффициентом.

Под аппроксимацией Крамера–Лундберга понимается (см., например, [6]) следующая

Теорема 1 (Крамера–Лундберга). Если пара $(\alpha, G(x))$ удовлетворяет условию Крамера, то $\psi(x) \sim Ce^{-Rx}$ при $x \rightarrow \infty$, причем

$$C = \frac{1 - \alpha}{\alpha R \int_0^{\infty} ue^{Ru} g(u) du}.$$

Заметим, что теорема, аналогичная теореме 1, касающаяся ветвящихся процессов, была доказана автором в [5].

Таким образом, теорема 1 устанавливает, что условие Крамера является достаточным условием для экспоненциальной асимптотики $\psi(x)$.

Возникает вопрос: условие Крамера является необходимым условием для экспоненциальной асимптотики $\psi(x)$? Оказывается, ответ на этот вопрос положителен и дан в работе автора [5], посвященной ветвящимся процессам.

Как следует из соотношения (3), достаточно получить положительный ответ на поставленный вопрос в рамках теории ветвящихся процессов. Переформулируем результат, полученный в [5], на «язык» теории риска (см. теорему 2 из [5]).

Теорема 2. *Если для некоторого процесса риска найдутся постоянные $C > 0$, $R > 0$, такие, что имеет место асимптотическое равенство $\psi(x) \sim Ce^{-Rx}$ при $x \rightarrow \infty$, то имеет место равенство $1 = \alpha \int_0^\infty e^{Ru} g(u) du$, то есть пара $(\alpha, G(x))$ удовлетворяет условию Крамера.*

Следовательно, теорему Крамера–Лундберга можно сформулировать следующим образом.

Теорема 3 (обобщенная теорема Крамера–Лундберга). *Условие Крамера является необходимым и достаточным условием для того, чтобы вероятность разорения $\psi(x)$ при $x \rightarrow \infty$ имела экспоненциальную асимптотику.*

Замечание. Как следует из теоремы 3, в случае невыполнения условия Крамера, асимптотика вероятности разорения $\psi(x)$ отлична от экспоненциальной, и возникает вопрос о характере ее поведения на бесконечности. В силу равенства (3) этот вопрос эквивалентен вопросу об асимптотике математического ожидания числа частиц в ветвящемся процессе с превращениями, зависящими от возраста. Он был рассмотрен в 1964 году в работе В.П. Чистякова [7]. Им было выделено некоторое подмножество из множества функций распределения, для которых не выполнено условие Крамера. Основным результатом этой работы заключается в следующем: для функций распределения времени жизни частицы, принадлежащих этому подмножеству, была найдена асимптотика математического ожидания числа частиц в ветвящемся процессе с превращениями, зависящими от возраста.

Гораздо позже вопрос об асимптотике вероятности разорения $\psi(x)$ в случае невыполнения условия Крамера рассматривался во многих работах по теории риска (см. библиографию в [6]).

5. Об оценке решений интегральных уравнений и ее приложениях в теории риска

Наряду с аппроксимацией Крамера–Лундберга, к фундаментальным результатам теории риска можно отнести также неравенство Лундберга, которое заключается в том, что если пара $(\alpha, G(x))$ удовлетворяет условию Крамера, то для всех $x \geq 0$ справедливо неравенство

$$\psi(x) \leq e^{-Rx}. \quad (4)$$

Этому неравенству и его обобщениям посвящено много работ (см. библиографию в [6]). Известно несколько доказательств неравенства Лундберга. В работе [8] предлагается метод доказательства этого неравенства и его различных обобщений

и уточнений, основанный на оценке решений интегральных уравнений. Одним из результатов этой работы является

Теорема 4. Если выполнено условие Крамера, то есть пара $(\alpha, G(x))$ удовлетворяет условию Крамера, то для всех $x \geq 0$ справедливо неравенство

$$\psi(x) \leq e^{-Rx} - \alpha \left(e^{-Rx} \int_x^\infty e^{Rz} g(z) dz - 1 + G(x) \right). \quad (5)$$

Из этого неравенства вытекает неравенство Лундберга, так как

$$e^{-Rx} \int_x^\infty e^{Rz} g(z) dz - 1 + G(x) \geq \int_x^\infty g(z) dz - 1 + G(x) = 0$$

для всех $x \geq 0$.

Приведем еще один результат этой работы.

Теорема 5. Если выполнено условие Крамера, то для $0 \leq u \leq x$ справедливо неравенство

$$A_1 e^{-Ru} \leq \psi(u) \leq A_2 e^{-Ru},$$

где

$$A_1 = \inf_{0 \leq u \leq x} \frac{\alpha(1 - G(u)) e^{Ru}}{1 - \alpha \int_0^u e^{R(x)z} g(z) dz}, \quad A_2 = \sup_{0 \leq u \leq x} \frac{\alpha(1 - G(u)) e^{Ru}}{1 - \alpha \int_0^u e^{R(x)z} g(z) dz}.$$

Доказательство теорем 4 и 5, а также других результатов работы [8] основано на следующих соображениях

Рассмотрим уравнение Вольтерра 2-го рода

$$\varphi(x) = \nu \int_0^x k(x-u) \varphi(u) du + f(x).$$

Для $0 \leq u \leq x$ зададим некоторую функцию $\bar{\varphi}(u)$. Нас будет интересовать следующий вопрос: каким условиям должны удовлетворять функции $k(u)$, $f(u)$ и постоянная $\nu > 0$, чтобы для $0 \leq u \leq x$ выполнялось неравенство $\varphi(u) \leq \bar{\varphi}(u)$? Нетрудно показать [8], что имеет место следующая

Лемма. Если $\nu > 0$, функции $k(u) \geq 0$ и $f(u) \geq 0$ непрерывны и ограничены для $0 \leq u \leq x$ и, кроме того, выполнено неравенство

$$\bar{\varphi}(u) \geq \nu \int_0^u k(u-v) \bar{\varphi}(v) dv + f(u),$$

то для $0 \leq u \leq x$ справедливо неравенство $\varphi(u) \leq \bar{\varphi}(u)$.

Замечание. Аналогичный результат справедлив и для нижней оценки функции $\varphi(u)$.

Основываясь на этой лемме, можно получать также различные оценки функции восстановления в теории восстановления.

6. Связь между теорией риска и классической задачей о баллотировке

Остановимся на связи классической задачи о баллотировке с задачей о разорении страховой компании. Хорошо известен следующий результат, доказанный в 1878 году У. Уитвортом, который носит название классической теоремы о баллотировке [1].

Теорема 6. *Предположим, что на выборах первый кандидат набрал a голосов, а второй — b голосов, причем $a > b$. Тогда вероятность того, что при последовательном подсчете голосов первый кандидат все время был впереди второго, равна $(a - b)/(a + b)$.*

Оказывается, существует также связь между теорией риска и задачей о баллотировке для двух точечных потоков [9]. Поясним, что мы понимаем под этой задачей. Пусть на положительной полуоси имеются два случайных точечных потока A и B . Пусть $a_1, a_2, \dots$ — точки потока A , а $b_1, b_2, \dots$ — точки потока B . Пусть $\nu_A(t)$ — число точек потока A на отрезке $[0, t]$. Нас будет интересовать событие, которое заключается в том, что либо на отрезке $[0, t]$ нет точек как потока A , так и потока B , либо на отрезке $[0, t]$ есть хотя бы одна точка потока A и для всех k ($1 \leq k \leq \nu_A(t)$) справедливы неравенства $a_k \leq b_k$. В случае наступления этого события будем говорить, что поток A опережает поток B на отрезке $[0, t]$, и будем писать $A(t) \prec B(t)$.

Положим $P(t) = P(A(t) \prec B(t))$. Для некоторых потоков A и B можно найти явное выражение для вероятности $P(t)$. Например, справедлива следующая

Теорема 7. Пусть A и B — два независимых пуассоновских потока с интенсивностями μ и λ соответственно. Тогда

$$P(t) = e^{-(\lambda+\mu)t} \left[I_0 \left(2\sqrt{\lambda\mu}t \right) + \sqrt{\frac{\mu}{\lambda}} I_1 \left(2\sqrt{\lambda\mu}t \right) + \left(1 - \frac{\lambda}{\mu} \right) \sum_{j=2}^{\infty} \left(\sqrt{\frac{\mu}{\lambda}} \right)^j I_j \left(2\sqrt{\lambda\mu}t \right) \right].$$

Здесь $I_j(z) = \sum_{k=0}^{\infty} \frac{z^{j+2k}}{2^{j+2k} k! (j+k)!}$ — обобщенная функция Бесселя первого рода.

Рассмотрим некоторый поток Π . Если $t_1, t_2, \dots$ — точки потока Π , то поток Π^x будем называть сдвигом потока Π на величину $x \geq 0$, если его точками являются точки $t_1 + x, t_2 + x, t_3 + x, \dots$

Пусть $r(x, t) = 1 - \psi(x, t)$ и $r(x) = 1 - \psi(x)$ — вероятности неразорения страховой компании на отрезке $[0, t]$ и на $[0, \infty]$ соответственно. Без ограничения общности в дальнейшем будем предполагать, что $c = 1$.

Нетрудно видеть, что событие, которое заключается в том, что компания не разорится до n -й выплаты включительно, произойдет тогда и только тогда, когда справедлива следующая система неравенств:

$$\left\{ \begin{array}{l} x + T_1 \geq Z_1, \\ x + T_2 \geq Z_1 + Z_2, \\ \\ x + T_n \geq Z_1 + Z_2 + + Z_n. \end{array} \right. \quad (6)$$

Случайные потоки $\{Z_1, Z_1 + Z_2, \dots, Z_1 + Z_2 + \dots + Z_n, \dots\}$ и $\{T_1 + x, T_2 + x, \dots, T_n + x, \dots\}$ будем обозначать соответственно через A_Z и A_{T+x} ($A_{T+0} = A_T$). Из системы неравенств (6) вытекает

Теорема 8. Вероятность разорения страховой компании с начальным капиталом x на отрезке $[0, t]$ равна вероятности того, что случайный поток A_Z опередит случайный поток A_{T+x} на отрезке $[0, x + t]$, то есть справедливо равенство $r(x, t) = P(A_Z(t) < A_{T+x}(t))$, в частности, $r(0, t) = P(A_Z(t) < A_T(t))$.

Основываясь на этой теореме, можно получить ряд результатов из теории риска [9].

Приведем один из них.

Теорема 9. Если начальный капитал страховой компании равен нулю, угловой коэффициент роста капитала $s = 1$, моменты наступления страховых случаев $\{T_n\}$ образуют пуассоновский поток с параметром λ , а величины выплат Z_n — независимые, одинаково распределенные случайные величины с функцией распределения $F(t)$, то

$$r(0, t) = \sum_{m=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^{m-1}}{m!} \int_0^t \lambda(t-y) dF^{*(m)}(y),$$

где $F^{*(m)}(y)$ — m -кратная свертка функции $F(t)$ с тождественной ей функцией ($m \geq 1$).

7. Неоднородный процесс риска

Поводом к рассмотрению неоднородного процесса риска послужило замечание Я. Галамбоша, сделанное им в его монографии [10]:

«Возьмем n клиентов страховой компании, принадлежащих одной страховой группе. Промежутки времени до первого несчастного случая для отдельных представителей этой группы можно рассматривать как независимые одинаково распределенные случайные величины X_j ($1 \leq j \leq n$) с общей непрерывной функцией распределения $F(x)$. Сообщения о несчастных случаях (или исках) поступают в страховую компанию в порядке возрастания этих промежутков времени, то есть компания «наблюдает» величины $X_{(j)}$, (где $X_{(j)}$ — j -я порядковая статистика величин среди величин $X_1, X_2, \dots, X_n$)».

Обычно в литературе предполагается, что длины интервалов времени между соседними несчастными случаями (или исками) являются независимыми, одинаково распределенными случайными величинами. Как показывает вышеприведенное замечание, представляет интерес отказаться от требования одинаковой распределенности интервалов между исками, и рассматриваемый процесс становится в некотором смысле неоднородным процессом. В работе [11] получены явные выражения для совместного распределения момента разорения и номера выплаты, приводящей к разорению в неоднородном процессе риска, определение которого дано в [11]. В частности, модель Я. Галамбоша является частным случаем рассматриваемого неоднородного процесса риска, если предположить, что промежутки времени до первого несчастного случая для отдельных представителей страховой группы имеют показательное распределение, а величины выплат также

имеют показательное распределение. Доказана теорема, которая устанавливает (в некотором смысле) эквивалентность модели Галамбоша и простейшей модели теории риска с конечным числом выплат, что позволяет выписать явное выражение для вероятности разорения в модели Галамбоша. Обратим внимание на то, что помимо этого полученные результаты позволяют рассмотреть ряд достаточно общих моделей функционирования страховой компании (например, рассмотреть случай зависимых и неодинаково распределенных интервалов между выплатами, причем величины выплат также неодинаково распределены и имеют достаточно общие функции распределения). Эти результаты могут также найти приложения в теории массового обслуживания. Более подробно о возможностях вышеупомянутых обобщений см. в [12].

Список литературы

- [1] *Феллер В.* Введение в теорию вероятностей и ее приложения. Т. 2.: Мир, 1984.
- [2] *Виноградов О. П.* Условие Крамера является необходимым и достаточным для экспоненциального убывания вероятности разорения // Теория вероятностей и ее применения. 1997. **42**, № 2. С. 388–389.
- [3] *Эмбрехтс П., Клоппельберг К.* Некоторые аспекты страховой математики // Теория вероятностей и ее применения 1993. **38**, № 2. С. 374–416.
- [4] *Севастьянов Б. А.* Ветвящиеся процессы. М.: Наука, 1971.
- [5] *Виноградов О. П.* Ветвящийся процесс с превращениями, зависящими от возраста // Теория вероятностей и ее применения. 1964. **9**, № 1. С. 146–152.
- [6] *Asmussen S.* Ruin probabilities. Singapore: World Scientific Publisher, 2000.
- [7] *Чистяков В. П.* Теорема о суммах независимых положительных случайных величин и ее приложения к ветвящимся случайным процессам // Теория вероятностей и ее применения. 1964. **9**, № 4. С. 710–718.
- [8] *Виноградов О. П.* Об одном элементарном методе получения оценок вероятности разорения // Обзорение прикл. и промышл. матем. 1998. **5**, № 1. С. 134–140.
- [9] *Виноградов О. П.* Задача о баллотировке для случайных потоков и ее связь с задачами теории риска, теории массового обслуживания и суммирования // Теория вероятностей и ее применения (в печати).
- [10] *Галамбош Я.* Асимптотическая теория экстремальных порядковых статистик. М.: Наука, 1984.
- [11] *Виноградов О. П.* О совместном распределении момента разорения и номера выплаты, приводящей к разорению, в неоднородном процессе риска // Теория вероятностей и ее применения. 2006. **51**, № 3. С. 465–475.
- [12] *Виноградов О. П.* Вероятность разорения страховой компании в случае, когда интервалы между моментами выплат имеют неодинаковые показательные распределения // Теория вероятностей и ее применения. 1998. **43**, № 2. С. 352–357.

МАКСИМАЛЬНЫЕ ВЕТВЯЩИЕСЯ ПРОЦЕССЫ

А. В. Лебедев

В статье дан обзор результатов автора в теории максимальных ветвящихся процессов. Отмечены приложения для вентильных бесконечнолинейных систем массового обслуживания. Основное внимание уделяется стационарным распределениям эргодических максимальных ветвящихся процессов и их предельным свойствам.

1. Введение

Параллели между теорией суммирования и теорией экстремумов случайных величин давно и хорошо известны.

Так, центральной предельной теореме (и ее аналогам) соответствует теорема об экстремальных типах, α -устойчивым распределениям — максимум-устойчивые [1, гл. 1], устойчивым процессам Леви — экстремальные процессы [2, § 6.5]. При отказе от независимости случайных величин в обоих случаях вводятся некоторые условия перемешивания, позволяющие получать предельные теоремы и в ситуации зависимости.

Классическими объектами исследования в теории случайных процессов являются ветвящиеся процессы Гальтона–Ватсона (с одним типом частиц и дискретным временем) [3, 4]. Оказывается, для них также возможно разумным образом построить «экстремальные» аналоги, которые называются максимальными ветвящимися процессами (МВП)¹⁾ и определяются следующим образом.

Рассматриваются случайные процессы со значениями в $\mathbf{Z}_+$, заданные стохастически рекуррентными формулами вида

$$Z_{n+1} = \bigvee_{m=1}^{Z_n} \xi_{m,n}, \quad (1)$$

где через $\bigvee$ обозначена операция взятия максимума и $\xi_{m,n}$, $m \geq 1$, $n \geq 0$, — независимые случайные величины с общим распределением F на $\mathbf{Z}_+$. Полагаем (как и в случае суммирования), что результат взятия максимума «нуль раз» (при $Z_n = 0$) равен нулю.

Можно сказать (по аналогии с процессами Гальтона–Ватсона), что в максимальном ветвящемся процессе выживают потомки только одной частицы, у которой их больше всего. Понятно также, что множество возможных значений МВП (при $n \geq 1$) совпадает с множеством возможных значений числа потомков. Из (1) следует, что процесс является однородной цепью Маркова на этом множестве.

Другую интерпретацию МВП можно предложить в теории массового обслуживания, рассмотрев вентильные бесконечнолинейные системы. Так называют системы с бесконечным числом приборов, в которых доступ заявок к обслуживанию регулируется вентилем. Предполагается, что вентиль открыт только в том случае, когда все приборы свободны. Заявки поступают в очередь с бесконечным числом

¹⁾ Термин «максимальный ветвящийся процесс» используется также в теории сетей Петри, однако там он имеет совершенно иное значение, как и термин «ветвящийся процесс».

мест ожидания, а обслуживание происходит по стадиям. В начале стадии, когда клапан открывается, все заявки из очереди мгновенно получают доступ к приборам и далее обслуживаются параллельно и независимо, до полного освобождения всех приборов. В момент освобождения всех приборов клапан вновь открывается для новой партии заявок (пришедших за это время) и следующей стадии.

Рассмотрим подобную систему с дискретным временем, и пусть в каждый момент времени поступает ровно по одной заявке. Тогда в силу параллельности работы приборов время обслуживания очередной партии заявок (а значит, и количество заявок в следующей) равно максимуму из времен их обслуживания. Таким образом, обозначая через Z_n длительность n -й стадии, а через $\xi_{m,n}$ — времена обслуживания заявок на ней, получаем в точности (1).

Клапанные бесконечнолинейные системы с непрерывным временем и пуассоновским входным потоком рассматривались, например, в [5–8] (другими методами). В этом случае необходимо оговорить, что происходит, если клапан открывается при пустой очереди. Наиболее естественно предположить, что система ждет поступления новой заявки, с которой и начинается следующая стадия.

МВП были введены в [9] (в связи с моделями дальнедействующей перколяции), и там же были получены критерии их возвратности. Предположим, что $F(0) = 0$. Тогда при выполнении условия

$$\limsup_{x \rightarrow +\infty} x(1 - F(x)) < e^{-\gamma}, \quad (2)$$

где $\gamma = 0,577\dots$ — константа Эйлера, цепь $\{Z_n\}$ положительно возвратна, и напротив, при

$$\liminf_{x \rightarrow +\infty} x(1 - F(x)) > e^{-\gamma}$$

имеет место $Z_n \rightarrow +\infty$, $n \rightarrow \infty$ почти наверное (п.н.).

Далее, в [10] рассмотрен критический случай

$$\lim_{x \rightarrow +\infty} x(1 - F(x)) = e^{-\gamma} \quad (3)$$

с учетом дальнейших членов разложения хвоста на бесконечности. Показано, что если $(e^\gamma x(1 - F(x)) - 1) \ln x \rightarrow d$, $x \rightarrow +\infty$, то процесс возвратен при $d < \pi^2/12$ и уходит на бесконечность п.н. при $d > \pi^2/12$. Более того, при условии $d < -(2p - 1)\pi^2/12$ времена возвращения в каждое состояние имеют конечные p -е моменты (где p целое), так что при $d < -\pi^2/12$ существует стационарное распределение.

Автором в [11, 12] было проведено обобщение МВП с $\mathbf{Z}_+$ на произвольные измеримые множества $T \subset \mathbf{R}_+$. Рассматривается цепь Маркова на T с переходными вероятностями

$$\mathbf{P}(Z_{n+1} \leq y | Z_n = x) = F(y)^x, \quad x, y \in T, \quad (4)$$

где распределение F сосредоточено на T (в целочисленном случае это распределение числа непосредственных потомков, в общем случае его тоже можно так называть, для определенности). Упомянутое обобщение аналогично переходу от процессов Гальтона–Ватсона к процессам Иржины [13, 14] (с непрерывным множеством состояний и дискретным временем), однако в данном случае T может быть любым измеримым подмножеством $\mathbf{R}_+$.

Подобные процессы могут рассматриваться как самостоятельно, так и в качестве предельных (в каком-либо смысле) для МВП на $\mathbf{Z}_+$ (нормированных определенным образом). Например, они могут использоваться для описания поведения вентильных бесконечнолинейных систем, в том числе предельного при большой загрузке и т.п.

В частности, для вентильной бесконечнолинейной системы с непрерывным временем и пуассоновским входным потоком последовательность длительностей стадий на периоде занятости удовлетворяет (4) с $F(x) = \exp\{-\lambda \bar{B}(x)\}$, $x \geq 0$, где λ — интенсивность входного потока, $B(x)$ — функция распределения времени обслуживания одной заявки, $\bar{B}(x) = 1 - B(x)$. Для описания системы на всем протяжении времени требуются уже МВП с иммиграцией в момент обнуления и т.п.

Далее будем говорить о максимальных ветвящихся процессах на T , которые будем обозначать МВП(T). Если $T_1 \subset T_2$, то всякий МВП(T_1) является также МВП(T_2). Таким образом, любой МВП(T), в частности МВП($\mathbf{Z}_+$), является МВП($\mathbf{R}_+$).

Формула (1) для таких МВП в общем случае уже не имеет места, но в соответствии с (4) они допускают эквивалентное представление стохастической рекуррентной последовательностью вида

$$Z_{n+1} = \begin{cases} F^{-1}(U_{n+1}^{1/Z_n}), & Z_n > 0; \\ 0, & Z_n = 0, \end{cases}, \quad n \geq 0, \quad (5)$$

где $F^{-1}(y) = \sup\{x : F(x) \leq y\}$ и U_n , $n \geq 1$, — независимые, равномерно распределенные на $(0, 1)$, случайные величины, а $Z_0 \geq 0$ и не зависит от них.

В [12] была доказана эргодическая теорема для МВП на $(0, +\infty)$. К условию Ламперти (2), исключающему уход процесса на бесконечность, при этом добавляется условие

$$\liminf_{x \rightarrow 0} x(-\ln F(x)) > e^{-\gamma}, \quad (6)$$

исключающее вырождение процесса в нуль.

Следует отметить, что базовые результаты для МВП (раздел 2) были получены и анонсированы автором еще в 2001 году [11], однако в развернутой форме их удалось опубликовать только в 2005 году [12].

Для эргодического МВП возникает вопрос о его стационарном распределении, которое обозначим через Ψ . В [15–17] были получены предельные теоремы для семейств Ψ , порожденных некоторыми семействами F (разделы 3–5). Разработанные методы также использовались при изучении вентильных бесконечнолинейных систем массового обслуживания с большой загрузкой [18, 19], где возникают аналогичные процессы.

Такой «коллективный» подход был связан с тем, что в общем случае найти Ψ для отдельно взятого F затруднительно, кроме как с помощью компьютерного моделирования. Однако оказалось возможным исследовать асимптотику хвостов стационарных распределений, связав ее с асимптотикой хвостов F . Новые результаты об этом (раздел 6) были получены относительно недавно и ранее не публиковались.

Интересными представляются также точка зрения на сложение (числа потомков) у процессов Гальтона–Ватсона и взятие максимума у МВП как на частные

случаи обобщенного суммирования [20] и изучение процессов, порожденных подобными операциями. Такие процессы рассматривались автором в [21–23].

2. Основные свойства. Эргодическая теорема. Примеры

Рассмотрим некоторые простые свойства МВП, такие, как свойство преобразования подобия, ассоциированность [24] и монотонность по параметрам [25].

Свойство 1. Если $\{Z_n\}$ является МВП(T) с $F(x)$, то $\{\lambda Z_n\}$ при любом $\lambda > 0$ является МВП(λT) с $F(x/\lambda)^{1/\lambda}$.

Из данного свойства следует замкнутость класса МВП($\mathbf{R}_+$) относительно умножения на $\lambda > 0$ и замкнутость МВП($\mathbf{Z}_+$) при $\lambda \in \mathbf{N}$.

Напомним понятие ассоциированности случайных величин [24]. Заметим, что свойство ассоциированности бывает очень полезным и имеет разнообразные применения в изучении случайных процессов и полей [26].

Функцию многих переменных $f(x)$, где $x = (x_1, \dots, x_n)$, назовем монотонно неубывающей, если из $x'_i \leq x''_i$, $1 \leq i \leq n$, следует $f(x') \leq f(x'')$. Случайные величины набора $\zeta = (\zeta_1, \dots, \zeta_n)$ называются ассоциированными, если $\text{cov}(f(\zeta), g(\zeta)) \geq 0$ для всех тех монотонно неубывающих f и g , для которых эта ковариация существует. Говорят, что случайный процесс или поле $\{\zeta(t) : t \in \mathcal{T}\}$ ассоциированы, если ассоциированы их значения $\zeta(t_1), \dots, \zeta(t_n)$ для любого конечного множества $\{t_1, \dots, t_n\} \subset \mathcal{T}$.

Согласно [24], независимые случайные величины ассоциированы; монотонно неубывающие функции от ассоциированных случайных величин также обладают этим свойством.

Свойство 2. Любой МВП ассоциирован.

Для установления монотонности по параметрам [25] введем отношение (частичного) порядка между распределениями: $F_1 \prec F_2$, если $F_1(x) \geq F_2(x)$, $\forall x$. Заметим, что из $F_1 \prec F_2$ следует $F_1^{-1}(y) \leq F_2^{-1}(y)$, $y \in (0, 1)$.

Обозначим через $Z = \mathcal{Z}(F, G)$ МВП с распределением «числа потомков» F и начальным распределением G для Z_0 .

Свойство 3. Если $F' \prec F''$ и $G' \prec G''$, то можно построить процессы $Z' = \mathcal{Z}(F', G')$ и $Z'' = \mathcal{Z}(F'', G'')$ на одном вероятностном пространстве так, что $Z'_n \leq Z''_n$ для всех $n \geq 0$ п.н.

Заметим, что для МВП нуль всегда является поглощающим состоянием. Так, для МВП($\mathbf{Z}_+$) при условии (2) и $F(0) > 0$ это приводит к вырождению п.н. [9]. Для МВП(T), если $F(0) = 0$, нуль можно просто исключить из множества состояний, рассматривая процесс с ненулевым начальным условием. Однако если нуль является предельной точкой T , остается возможность асимптотической сходимости к нему при $n \rightarrow \infty$. Следующая теорема дает достаточные условия для того, чтобы исключить уход процесса как в нуль, так и на бесконечность и сделать его эргодическим. Здесь и далее имеется в виду эргодичность по Харрису [27, гл. 1].

Теорема 1. Если для МВП(T), $T \subset (0, +\infty)$, выполнены условия (2) и (6), то процесс эргодический.

Достаточное условие Ламперти для положительной возвратности МВП на $\mathbf{Z}_+$ [9] получается из теоремы 1 как частный случай.

Далее будем обозначать эргодическое распределение $\{Z_n\}$ через Ψ , а случайную величину с таким распределением — через $\tilde{Z}$.

Свойство 4. Если для двух эргодических МВП(T) с F' и F'' верно $F' \prec F''$, то $\Psi' \prec \Psi''$.

Данное свойство верно и в тех случаях, когда одно или оба предельных распределения сосредоточены в нуле.

Следствие 1. Если для МВП(R_+) выполнено (2) и $F(0) > 0$, то процесс вырождается п.н.

Здесь под вырождением понимается обращение процесса в нуль начиная с некоторого (случайного) момента.

Пример 1. Переходные вероятности цепи Маркова для МВП($\mathbf{N}$) имеют вид

$$p_{i,j} = F(j)^i - F(j-1)^i, \quad i, j \in \mathbf{N}.$$

Таким образом, в случае конечного $T \subset \mathbf{N}$ по матрице $P = (p_{i,j})$ не составляет труда рассчитать стационарное распределение.

Например, пусть число потомков принимает значения K и L с вероятностями p и $1-p$ соответственно, $1 \leq K < L$, $p \in [0, 1]$. Тогда стационарные вероятности определяются по формулам

$$\mathbf{P}(\tilde{Z} = K) = \frac{p^L}{1 - p^K + p^L}, \quad \mathbf{P}(\tilde{Z} = L) = \frac{1 - p^K}{1 - p^K + p^L}.$$

Пример 2. Пусть $T = \{1/k : k \in \mathbf{N}\}$ и $F(1/k) = p^{k-1}$, $p \in (0, 1)$. В данном случае распределение F и множество T дискретны, однако нуль является предельной точкой T . Имеем $F(x) \leq p^{(1/x)-1}$, $x \in (0, 1]$. По теореме 1 такой МВП эргодический при $p < \exp\{-e^{-\gamma}\} = 0,570\dots$

Пример 3. Если функция распределения $F(x)$ непрерывна и строго возрастает на $(0, +\infty)$, а $F(0) = 0$, то (5) преобразованием $\zeta_n = -\ln(-\ln F(Z_n))$ приводится к форме общей (нелинейной) авторегрессии первого порядка:

$$\zeta_{n+1} = f(\zeta_n) + \eta_{n+1}, \quad n \geq 0, \quad (7)$$

где $f(u) = \ln F^{-1}(\exp\{-e^{-u}\})$, и независимые случайные величины $\eta_n = -\ln(-\ln U_n)$, $n \geq 1$, имеют функцию распределения Гумбеля $\Lambda(x) = \exp\{-e^{-x}\}$.

Пусть $F(x) = \exp\{-(x/c)^{-\alpha}\}$, $x, c, \alpha > 0$ (распределение Фреше), тогда МВП допускает конструктивное представление

$$Z_{n+1} = \nu_{n+1} Z_n^{1/\alpha},$$

где ν_n , $n \geq 1$, независимы и имеют распределение F , а (7) переписывается в форме линейной авторегрессии

$$\zeta_{n+1} = \zeta_n/\alpha + \ln c + \eta_{n+1}, \quad n \geq 0.$$

Отметим, что $\mathbf{M}\eta_1 = \gamma$. При $\alpha < 1$ процесс $\{\zeta_n\}$ уходит на $\pm\infty$ в зависимости от знака начального условия ζ_0 . При $\alpha > 1$ процесс $\{\zeta_n\}$ эргодический. При $\alpha = 1$ имеем простое случайное блуждание, уходящее на $+\infty$ при $c > e^{-\gamma}$, на $-\infty$ при $c <$

$e^{-\gamma}$ и осциллирующее между $\pm\infty$ при $c = e^{-\gamma}$. Отсюда легко получить результаты для $\{Z_n\}$, если учесть, что $Z_n = F^{-1}(\Lambda(\zeta_n))$, откуда $Z_n \rightarrow 0$ при $\zeta_n \rightarrow -\infty$ и $Z_n \rightarrow +\infty$ при $\zeta_n \rightarrow +\infty$, а эргодичность сохраняется.

Очевидно, условия теоремы 1 выполняются только при $\alpha > 1$, поскольку тогда $x(1 - F(x)) \rightarrow 0$, $x \rightarrow +\infty$ и $-x \ln F(x) \rightarrow +\infty$, $x \rightarrow 0$. В этом случае МВП эргодический.

Пример 4. Пусть $F(x) = \exp\{c(1 - 1/x)\}$, $x \in T = (0, 1)$, $c > 0$, тогда $F^{-1}(y) = 1/(1 - (\ln y)/c)$, $y \in (0, 1)$, и (5) принимает форму

$$Z_{n+1} = \frac{Z_n}{Z_n - (\ln U_{n+1})/c},$$

которая заменой $\rho_n = 1/Z_n$ и $\sigma_n = -(\ln U_n)/c$ приводится к виду обобщенной авторегрессии первого порядка

$$\rho_{n+1} = 1 + \sigma_{n+1}\rho_n.$$

Известно [27, § 8.4], что поведение такой последовательности определяется знаком $\mathbf{M} \ln \sigma_1 = -\gamma - \ln c$. При $c < e^{-\gamma}$ получаем $\rho_n \rightarrow +\infty$ и $Z_n \rightarrow 0$, $n \rightarrow \infty$ (п.н.). При $c > e^{-\gamma}$ процессы $\{\rho_n\}$ и $\{Z_n\}$ эргодические (в согласии с теоремой 1).

3. МВП на ограниченных множествах

Здесь изучается подкласс $\{Z_n\}$ с $T \subset [\alpha, \beta]$, где $0 < \alpha < \beta < \infty$. По теореме 1 такие цепи Маркова являются эргодическими.

Рассмотрим семейство процессов $\{Z_n^{(\lambda)}\}$ с $F^{(\lambda)}$ и $T^{(\lambda)} \subset [1, \lambda]$, $\lambda \geq 1$. Для каждого $\{Z_n^{(\lambda)}\}$ существует и единственно стационарное распределение $\Psi^{(\lambda)}$. Будем обозначать случайную величину с таким распределением через $\tilde{Z}^{(\lambda)}$.

Теорема 2. Если

$$\delta(q) = \limsup_{\lambda \rightarrow \infty} F^{(\lambda)}(q\lambda) < 1, \quad \forall q \in (0, 1) \quad (8)$$

и для некоторой функции $u : [1, +\infty) \rightarrow [1, +\infty)$ существует предел

$$\lim_{\lambda \rightarrow \infty} \lambda(1 - F^{(\lambda)}(u(\lambda))) = \tau, \quad 0 \leq \tau \leq \infty,$$

то

$$\lim_{\lambda \rightarrow \infty} \Psi^{(\lambda)}(u(\lambda)) = e^{-\tau}.$$

Замечание. В силу свойства 1 (подобия) теорема 2 обобщается и на случай множеств $T^{(\lambda)} \subset [\alpha, \lambda]$, $\lambda \rightarrow \infty$, с любым $\alpha > 0$.

Предельное распределение может быть разным: непрерывным или дискретным, вырожденным или невырожденным. Важный подкласс возможных предельных распределений составляют распределения экстремальных значений [1, 2].

Будем писать $F \in D(G)$, если существуют числовые последовательности $a_n > 0$, b_n , $n \geq 1$, такие, что

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\bigvee_{k=1}^n \xi_k \leq a_n x + b_n \right) = G(x), \quad x \in \mathbf{R},$$

где случайные величины ξ_k , $k \geq 1$, независимы и имеют одинаковое распределение F ; G — невырожденное распределение, относящееся к одному из экстремальных типов [1, теорема 1.4.2].

Следствие 2. Если верно условие (8), каждая функция $F^{(\lambda)}$ непрерывна на $T^{(\lambda)}$ и выполнено условие

$$\limsup_{\lambda \rightarrow \infty} F^{(\lambda)}(\inf T^{(\lambda)}) < 1,$$

то существует функция $v : [1, +\infty) \times [0, +\infty) \rightarrow [1, +\infty)$, такая, что

$$\lim_{\lambda \rightarrow \infty} \Psi^{(\lambda)}(v(\lambda, \tau)) = e^{-\tau}$$

для любого $\tau \geq 0$.

Следствие 3. Если выполнено условие (8) и существует распределение F_0 на $(-\infty, 1]$, $F_0 \in D(G)$, такое, что для некоторой функции $\varphi : [1, +\infty) \times (-\infty, 1] \rightarrow [1, +\infty)$

$$\frac{1 - F^{(\lambda)}(\varphi(\lambda, c_\lambda(x)))}{1 - F_0(c_\lambda(x))} \rightarrow 1, \quad \lambda \rightarrow \infty,$$

где $c_\lambda(x) = a_{[\lambda]}x + b_{[\lambda]}$, $x \in \mathbf{R}$, то

$$\lim_{\lambda \rightarrow \infty} \Psi^{(\lambda)}(\varphi(\lambda, c_\lambda(x))) = G(x).$$

Пример 5. Пусть $F^{(N)}(k) = (k/N)^\sigma$, $T^{(N)} = \{1, 2, \dots, N\}$, $\sigma > 0$, тогда для любого $k \in \mathbf{Z}_+$ получаем $\Psi^{(N)}(N - k) \rightarrow e^{-\sigma k}$, $N \rightarrow \infty$, так что распределение случайной величины $N - \tilde{Z}^{(N)}$ при $N \rightarrow \infty$ слабо сходится к геометрическому с параметром $e^{-\sigma}$.

Пример 6. Пусть $F^{(N)}(k) = 2^{k-N}$, $T^{(N)} = \{1, 2, \dots, N\}$, тогда $\Psi^{(N)}(N - 1) \rightarrow 0$, $N \rightarrow \infty$, так что $N - \tilde{Z}^{(N)} \rightarrow 0$ по вероятности.

Пример 7. Пусть

$$F^{(\lambda)}(x) = 1 - \left(\frac{\lambda - x}{\lambda - 1} \right)^\gamma, \quad x \in T^{(\lambda)} = [1, \lambda], \quad \gamma > 0,$$

тогда для любого $x > 0$ получаем $\Psi^{(\lambda)}(\lambda - x\lambda^{1-1/\gamma}) \rightarrow e^{-x^\gamma}$, так что распределение случайной величины $(\lambda - \tilde{Z}^{(\lambda)})/\lambda^{1-1/\gamma}$ при $\lambda \rightarrow \infty$ слабо сходится к распределению Вейбулла с показателем γ .

Пример 8. Пусть

$$F^{(\lambda)}(x) = 1 - \exp \left\{ -\frac{x-1}{\lambda-x} \right\}, \quad x \in T^{(\lambda)} = [1, \lambda],$$

тогда

$$\Psi^{(\lambda)} \left(\lambda - \frac{\lambda-1}{\ln \lambda} + \frac{\lambda-1}{(\ln \lambda)^2} (x+1) \right) \rightarrow \exp\{-e^{-x}\}, \quad \forall x \in \mathbf{R},$$

так что распределение случайной величины

$$(\ln \lambda)^2 \frac{\tilde{Z}^{(\lambda)} - \lambda}{\lambda - 1} + \ln \lambda - 1$$

при $\lambda \rightarrow \infty$ слабо сходится к распределению Гумбеля.

4. Двойной показательный закон для МВП

В [16] был рассмотрен случай дискретных F , принадлежащих области притяжения двойного показательного закона (Гумбеля). Однако все доказательства проходят и для непрерывных распределений на $[1, +\infty]$, а в силу свойства 1 (подобия) и для распределений с любой положительной крайней левой точкой.

Далее, полагая $F(0) = 0$, будем рассматривать МВП при более сильном, чем (2), условии конечности математического ожидания «числа потомков»

$$\mu_F = \int_0^{\infty} (1 - F(u)) du < \infty,$$

откуда следует

$$\lim_{x \rightarrow \infty} x(1 - F(x)) = 0. \quad (9)$$

Рассмотрим последовательность МВП $\{Z_n^{(N)}\}$ с $F^{(N)}(x) = F(x)^N$, $N \geq 1$. Пусть F удовлетворяет (9), тогда в силу асимптотики $\bar{F}^{(N)}(x) \sim N\bar{F}(x)$, $x \rightarrow \infty$, для всех $F^{(N)}$ также выполнено (9). Из положительной возвратности $\{Z_n^{(N)}\}$ в сочетании с неприводимостью и аperiodичностью следует эргодичность и существование стационарного предельного распределения $\Psi^{(N)}$ у данной цепи Маркова (при каждом $N \geq 1$). Случайные величины с распределениями $\Psi^{(N)}$ будем обозначать через $\tilde{Z}^{(N)}$.

Пусть F принадлежит области притяжения двойного показательного закона $\Lambda(x) = \exp\{-e^{-x}\}$ при линейной нормировке, то есть существуют числовые последовательности $a_n > 0$, b_n , $n \geq 1$, такие, что

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\bigvee_{k=1}^n \xi_k \leq a_n x + b_n \right) = \Lambda(x), \quad \forall x \in R. \quad (10)$$

Критерии выполнения (10) можно найти в [1, 2]. В частности, хвост F должен убывать быстрее любой степени. В случае дискретных F этот хвост также должен убывать медленнее любой экспоненты, так как в противном случае не выполняется условие существования невырожденного предельного распределения максимумов $\bar{F}(x)/\bar{F}(x-0) \rightarrow 1$, $x \rightarrow \infty$ [1, § 1.7].

Определим функцию $u(s) = \inf\{x : s\bar{F}(x) \leq 1\}$, $s > 0$. По теореме 2.1.3 [2] значения $u(k)$, $k \geq 1$, можно выбрать в качестве b_k , что и будем предполагать в дальнейшем. Заметим, что $u(s)$ монотонно не убывает и стремится к бесконечности при $s \rightarrow \infty$.

Введем дополнительное условие: пусть для любого $\theta > 1$ существует $\varepsilon > 0$, такое, что

$$su(s^{1+\varepsilon})\bar{F}(\theta u(s)) \rightarrow 0, \quad s \rightarrow \infty. \quad (11)$$

Заметим, что (11) имеет смысл ограничения «тяжести» хвоста F . Условиям (10) и (11) удовлетворяют распределения: Вейбулла, показательное, гамма-, нормальное и др. Разницу можно увидеть на примере класса распределений с хвостами $\bar{F}(x) \sim \exp\{-(\ln x)^\alpha\}$, $x \rightarrow \infty$: тогда (10) выполняется при $\alpha > 1$, а (11) — при $\alpha > 2$.

Теорема 3. Если выполнены (10) и (11), то

$$\tilde{Z}^{(N)}/u(N) \xrightarrow{P} 1, \quad N \rightarrow \infty.$$

Следствие 4. Если выполнены (10) и (11), то

$$\lim_{N \rightarrow \infty} \mathbf{P}(\tilde{Z}^{(N)} \leq a_{k(N)}x + b_{k(N)}) = \Lambda(x), \quad \forall x \in R,$$

для любых $k(N) \sim Nu(N)$, $N \rightarrow \infty$.

Пример 9. Пусть $F(n) = 1 - \exp\{-\sqrt{n}\}$, тогда $a_k = 2 \ln k$, $b_k = (\ln k)^2$, $k \geq 1$, и, следовательно,

$$\mathbf{P} \left(\frac{\tilde{Z}^{(N)} - (\ln N + 2 \ln \ln N)^2}{2(\ln N + 2 \ln \ln N)} \leq x \right) \rightarrow \Lambda(x), \quad N \rightarrow \infty.$$

5. МВП в случае степенных хвостов

Рассмотрим случай, когда $F^{(\lambda)}(x) = F(x)^\lambda$, где распределение F сосредоточено на $(0, +\infty)$ и имеет степенной хвост, а следовательно, принадлежит области притяжения предельного закона Фреше для максимумов.

Введем трехпараметрическое семейство распределений Фреше

$$\Phi_{\alpha,b,c}(x) = \begin{cases} \exp\{-c(x-b)^{-\alpha}\}, & x > b; \\ 0, & x \leq b, \end{cases}$$

где $\alpha, c > 0$. Из определения следует

$$\Phi_{\alpha,b,c}^r(sx) \equiv \Phi_{\alpha,b/s,cr/s^\alpha}(x), \quad r, s > 0. \quad (12)$$

По теореме 1 любой МВП с $F = \Phi_{\alpha,b,c}$ при $\alpha > 1$, $b \geq 0$, $c > 0$ эргодический. Его стационарное распределение обозначим через $\Psi_{\alpha,b,c}$.

Теорема 4. Пусть существуют такие числа $\alpha > 1$, $c, c_1 > 0$, что $F \succ \Phi_{\alpha,0,c_1}$ и $\bar{F}(x) \sim cx^{-\alpha}$, $x \rightarrow \infty$. Тогда

$$\lim_{\lambda \rightarrow \infty} \Psi^{(\lambda)}(x\lambda^{1/(\alpha-1)}) = \Psi_{\alpha,0,c}(x). \quad (13)$$

Условию теоремы удовлетворяет, например, распределение Парето

$$F(x) = \begin{cases} 1 - (x/x_0)^{-\alpha}, & x > x_0; \\ 0, & x \leq x_0, \end{cases}$$

при $\alpha > 1$, $x_0 > 0$, поскольку $F(x) \leq \exp\{-(x/x_0)^{-\alpha}\}$ при всех $x > 0$. Также можно подобрать необходимую оценку для любого распределения со степенным хвостом и положительной крайней левой точкой.

Замечание. Если $F = \Phi_{\alpha,0,1}$, $\alpha > 1$, то предел (13) обращается в тождество: получаем формулу

$$\Psi_{\alpha,0,\lambda}(\mu x) \equiv \Psi_{\alpha,0,1}(x), \quad (14)$$

где $\mu = \lambda^{1/(\alpha-1)}$, так как $\Phi_{\alpha,0,\lambda}^\mu(\mu x) \equiv \Phi_{\alpha,0,1}(x)$, согласно (12).

Заметим, что рекуррентная случайная последовательность

$$Z_{n+1} = Z_n^{1/\alpha} \xi_{n+1}, \quad (15)$$

где ξ_n , $n \geq 0$, независимы и имеют распределение $\Phi_{\alpha,0,1}$, удовлетворяет (4) с $F = \Phi_{\alpha,0,1}$. Отсюда следует, что случайная величина $\tilde{Z}$, заданная бесконечным произведением

$$\tilde{Z} = \prod_{n=0}^{\infty} \xi_n^{\alpha^{-n}} \quad (16)$$

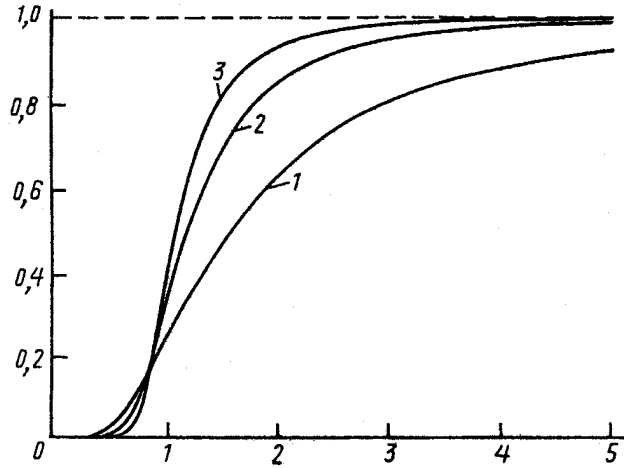
(сходящимся при $\alpha > 1$ почти наверное), будет иметь распределение $\Psi_{\alpha,0,1}$. Случайная величина $\tilde{Y} = \ln \tilde{Z}$ представима рядом

$$\tilde{Y} = \sum_{n=1}^{\infty} \eta_n \alpha^{-n}, \quad (17)$$

где $\eta_n = (\ln \xi_{n-1})/\alpha$, $n \geq 1$, независимы и имеют функцию распределения Гумбеля $\Lambda(x) = \exp\{-e^{-x}\}$. Из (16) и (17) получаем формулу моментов $\tilde{Z}$ и характеристическую функцию $\tilde{Y}$:

$$\mathbf{M}\tilde{Z}^s = \prod_{n=1}^{\infty} \Gamma(1 - s\alpha^{-n}), \quad 0 < s < \alpha; \quad \mathbf{M}e^{it\tilde{Y}} = \prod_{n=1}^{\infty} \Gamma(1 - it\alpha^{-n}).$$

Кроме того, из (14) следует $\Psi_{\alpha,0,c}(x) = \Psi_{\alpha,0,1}(xc^{-1/(\alpha-1)})$ при любом $c > 0$.



На рисунке представлены приближенные графики функций распределения $\Psi_{\alpha,0,1}$ при $\alpha = 2, 3, 4$ (кривые 1, 2, 3 соответственно), построенные по результатам моделирования последовательности (15) за 10000 шагов.

6. Асимптотика хвостов стационарных распределений

Как уже было отмечено, в общем случае найти Ψ для отдельно взятого F затруднительно. Как ни странно, гораздо проще решается обратная задача: найти F по Ψ .

Действительно, в силу (4) должно выполняться равенство

$$\Psi(x) = \int_0^{+\infty} F(x)^u d\Psi(u),$$

что эквивалентно

$$\Psi(x) = \psi(-\ln F(x)),$$

где $\psi(s)$ — преобразование Лапласа–Стилтьеса для Ψ , откуда

$$F(x) = \exp\{-\psi^{-1}(\Psi(x))\}. \quad (18)$$

Из (18) сразу следуют два простых, но немаловажных свойства.

Свойство 5. *Монотонность F по Ψ : из $\Psi_1 \prec \Psi_2$ следует $F_1 \prec F_2$.*

Ранее в [12] была доказана монотонность Ψ по F (свойство 4).

Свойство 6. *Функции распределения F и Ψ имеют одни и те же области постоянства, точки непрерывности, точки скачков (если они есть), левую и правую крайние точки.*

Далее будем полагать распределение F невырожденным, откуда также следует невырожденность Ψ .

Пример 10. Рассмотрим показательное распределение $\Psi(x) = 1 - e^{-x}$, $x > 0$, тогда $\psi(s) = 1/(1+s)$ и

$$F(x) = \exp\left\{-\frac{e^{-x}}{1 - e^{-x}}\right\}, \quad x > 0,$$

так что $\bar{F}(x) \sim e^{-x}$, $x \rightarrow \infty$.

Пример 11. Рассмотрим одностороннее устойчивое распределение с показателем $\alpha = 1/2$: $\Psi(x) = 2(1 - \Phi(x^{-1/2}))$, $x > 0$, где Φ — функция стандартного нормального распределения, тогда $\psi(s) = e^{-\sqrt{2}s}$ и

$$F(x) = \exp\left\{-\frac{1}{2}(\ln(2(1 - \Phi(x^{-1/2})))^2\right\}, \quad x > 0,$$

так что $\bar{F}(x) \sim 1/(\pi x)$, $x \rightarrow \infty$.

Заметим, что в первом примере хвосты F и Ψ имеют один и тот же порядок убывания, а во втором — различные. Оказывается, это зависит от того, имеет F конечное или бесконечное среднее. Далее будут разобраны оба случая.

Пусть сначала F имеет конечное среднее.

Обозначим математические ожидания F и Ψ через μ_F и μ_Ψ соответственно, а также через x_0 и x_ω крайние левую и правую точки распределений F и Ψ (общие по свойству 6).

Используем функцию

$$\varphi(x) = \mathbf{M}(Z_{n+1}|Z_n = x) = \int_0^{+\infty} (1 - F(u)^x) du,$$

которая определена при $\mu_F < \infty$ для всех $x > 0$. Положим $\varphi(0) = x_0$.

Лемма 1. Если $\mu_F < \infty$, то $\varphi(x)$ — непрерывная, строго возрастающая, строго вогнутая функция на $\mathbf{R}_+$, и $\varphi(x) = o(x)$, $x \rightarrow \infty$.

Лемма 2. Положительное решение уравнения $\varphi(x) = x$ существует и единственно; числовая последовательность $v_{n+1} = \varphi(v_n)$ сходится к нему при любом начальном условии $v_0 > 0$.

Обозначим это решение через μ^* .

Теорема 5. $\mu_\Psi < \infty$ тогда и только тогда, когда $\mu_F < \infty$. При этом $\mu_\Psi \leq \mu^*$ и верно

$$\bar{\Psi}(x) \sim \mu_\Psi \bar{F}(x), \quad x \rightarrow x_\omega - 0. \quad (19)$$

Таким образом, в случае конечного среднего хвосты F и Ψ имеют один и тот же порядок убывания с точностью до постоянного множителя.

К сожалению, μ_Ψ удастся вычислить лишь в редких случаях. Например, как показано в [17], если $F(x) = \exp\{-x^{-\alpha}\}$, $x > 0$, $\alpha > 1$ (распределение Фреше), то

$$\mu_\Psi = \prod_{n=1}^{\infty} \Gamma(1 - \alpha^{-n}).$$

При этом $\varphi(x) = x^{1/\alpha} \Gamma(1 - \alpha^{-1})$ и

$$\mu_* = \Gamma(1 - \alpha^{-1})^{\frac{\alpha}{\alpha-1}}.$$

Пусть теперь F имеет бесконечное среднее.

Заметим, что речь идет о довольно узком классе распределений F , поскольку по своей «тяжести» хвостов они ограничены условием (2). Однако этот случай более сложный.

Теорема 6. Пусть $\bar{\Psi}(x) \sim x^{-\alpha}$, $0 < \alpha < 1$, $x \rightarrow \infty$, тогда $\bar{F}(x) \sim c(\alpha)/x$, $x \rightarrow \infty$, где $c(\alpha) = \Gamma(1 - \alpha)^{-1/\alpha}$.

Заметим, что на интервале $(0, 1)$ функция $c(\alpha)$ строго убывает от $e^{-\gamma}$ до 0. Таким образом, все более тяжелым хвостам Ψ (при $\alpha \rightarrow 0$) соответствуют хвосты F , все более близкие к критическому (3).

Теорема 6 допускает следующее частичное обращение.

Теорема 7. Если $\bar{F}(x) \sim c(\alpha)/x$, $x \rightarrow \infty$, то $\limsup_{x \rightarrow \infty} \log_x \bar{\Psi}(x) = -\alpha$. Если, кроме того, $x_0 > 0$, то имеет место точный предел $\lim_{x \rightarrow \infty} \log_x \bar{\Psi}(x) = -\alpha$.

Следствие 5. Если $\bar{F}(x) \sim e^{-\gamma}/x$, $x \rightarrow \infty$, но существует стационарное распределение Ψ , то $\lim_{x \rightarrow \infty} \log_x \bar{\Psi}(x) = 0$.

Таким образом, в критическом случае могут возникать «сверхтяжелые» хвосты, убывающие медленнее любой степени.

Следствие 6. Если $\bar{F}(x) = o(1/x)$, $x \rightarrow \infty$, но $\mu_F = +\infty$, то $\limsup_{x \rightarrow \infty} \log_x \bar{\Psi}(x) = -1$.

Пример 12. Рассмотрим распределение Парето $\Psi(x) = 1 - x^{-1}$, $x \geq 1$, тогда

$$1 - \psi(s) = \int_1^{+\infty} \frac{1 - e^{-sx}}{x^2} dx = s \int_0^{1/s} (1 - e^{-1/u}) du \sim s(-\ln s), \quad s \rightarrow 0 + 0,$$

откуда $\psi^{-1}(1 - u) \sim u/(-\ln u)$, $u \rightarrow 0 + 0$, и $\bar{F}(x) \sim 1/(x \ln x)$, $x \rightarrow \infty$.

В этом примере хвосты F и Ψ различаются не постоянным множителем и не показателем степени, а медленно меняющейся функцией.

Работа выполнена при поддержке грантов РФФИ № 07-01-00077, 07-01-00373.

Список литературы

- [1] Лидбеттер М., Линдгрэн Г., Ротсен Х. Экстремумы случайных последовательностей и процессов. М.: Мир, 1989.
- [2] Галамбош Я. И. Асимптотическая теория экстремальных порядковых статистик. М.: Наука, 1984.
- [3] Севастьянов Б. А. Ветвящиеся процессы. М.: Наука, 1971.
- [4] Ватутин В. А., Зубков А. М. Ветвящиеся процессы. I. // Итоги науки и техники. Сер. Теор. вероятн. Матем. статистика. Теор. кибернетика. Т. 23. М.: ВИНТИ, 1985. С. 3–67.
- [5] Browne S., Coffman E. G., Gilbert E. N., Wright P. E. Gated, exhaustive, parallel service // Probab. Eng. Inf. Sci. 1992. **2**, № 2. P. 217–239.
- [6] Browne S., Coffman E. G., Gilbert E. N., Wright P. E. The gated infinite-server queue: uniform service times // SIAM J. Appl. Math. 1992. **52**, № 6. P. 1751–1762.
- [7] Knessl Ch., Tan X. Heavy traffic asymptotics for a gated, infinite-server queue with uniform service times // SIAM J. Appl. Math. 1994. **54**, № 6. P. 1768–1779.
- [8] Pinotsi D., Zazanis M. A. Stability conditions for gated $M|G|\infty$ queues // Probab. Eng. Inf. Sci. 2004. **18**, № 1. P. 103–110.
- [9] Lamperti J. Maximal branching processes and long-range percolation // J. Appl. Probab. 1970. **7**, № 1. P. 89–96.
- [10] Lamperti J. Remarks on maximal branching processes // Теория вероятностей и ее применения. 1972. **17**, № 1. С. 46–54.
- [11] Лебедев А. В. Максимальные ветвящиеся процессы и их обобщения // Обзорение прикладной и промышленной математики. 2001. **8**, № 2. С. 785–786.
- [12] Лебедев А. В. Максимальные ветвящиеся процессы с неотрицательными значениями // Теория вероятностей и ее применения. 2005. **50**, № 3. С. 564–570.
- [13] Jirina M. Stochastic branching processes with continuous state space // Czech. Math. J. 1958. **8**, № 2. P. 292–313.
- [14] Иржина М. Ветвящиеся случайные процессы с непрерывным пространством состояний // Теория вероятностей и ее применения. 1959. **4**, № 4. С. 482–484.
- [15] Лебедев А. В. Обобщенные максимальные ветвящиеся процессы на ограниченных множествах // Вестн. Моск. ун-та. Матем. Механ. 2002. № 6. С. 55–57.
- [16] Лебедев А. В. Двойной показательный закон для максимальных ветвящихся процессов // Дискретная математика. 2002. **14**, № 3. С. 143–148.
- [17] Лебедев А. В. Обобщенные максимальные ветвящиеся процессы в случае степенных хвостов // Вестн. Моск. ун-та. Матем. Механ. 2005. № 2. С. 47–49.
- [18] Лебедев А. В. Вентильная бесконечнолинейная система с неограниченными временами обслуживания и большой загрузкой // Проблемы передачи информации. 2003. **39**, № 3. С. 87–94.

- [19] *Лебедев А. В.* Вентильная бесконечнолинейная система с большой загрузкой и степенным хвостом // Проблемы передачи информации. 2004. **40**, № 3. С. 62–68.
- [20] *Золотарев В. М.* Современная теория суммирования независимых случайных величин. М.: Наука, 1986.
- [21] *Лебедев А. В.* Ветвящиеся процессы с обобщенной операцией суммирования. Случай умножения // Вестн. Моск. ун-та. Матем. Механ. 2002. № 2. С. 65–68.
- [22] *Лебедев А. В.* Случайные процессы с обобщенными операциями суммирования // Вестн. Моск. ун-та. Матем. Механ. 2005. № 4. С. 3–5.
- [23] *Лебедев А. В.* Ветвящиеся процессы с конкуренцией частиц // Вестн. Моск. ун-та. Матем. Механ. 2006. № 4. С. 56–58.
- [24] *Esary J., Prochan F., Walkup D.* Association of random variables with applications // Ann. Math. Statist. 1967. **38**, №5. P. 1466–1474.
- [25] *Штойян Д.* Качественные свойства и оценки стохастических моделей. М.: Мир, 1979.
- [26] *Булинский А. В., Шашкин А. П.* Предельные теоремы для ассоциированных случайных полей и родственных систем. М.: Физматлит, 2008.
- [27] *Боровков А. А.* Эргодичность и устойчивость случайных процессов. М.: УРСС, 1999.

МНОГОМЕРНЫЕ СТАТИСТИЧЕСКИЕ МОДЕЛИ В ГЕОМЕТРИЧЕСКОМ ИЗЛОЖЕНИИ

Ю. Н. Тюрин

Многомерные случайные величины мы записываем в виде столбцов, а их последовательность — в виде строки (позиции в ней заняты столбцами). В пространстве таких строк мы вводим внутреннее произведение, результатом которого оказываются квадратные матрицы. Это умножение мы толкуем и используем как скалярное умножение, в частности, чтобы определить понятия ортогональности, линейного подпространства и пр. С помощью этой алгебры удастся определить многомерные линейные статистические модели, матричный метод наименьших квадратов, линейные статистические гипотезы и дать способы их проверки в полном сходстве с классической теорией линейного статистического анализа.

1. Метрическое пространство таблиц

Таблицы. Многомерное наблюдение — это несколько одновременно наблюдаемых случайных величин. Приняв для этих случайных величин определенную нумерацию, многомерное наблюдение мы записываем в виде столбца. Наблюдения будем снабжать индексами. В зависимости от структуры эксперимента индексация может быть как простой (линейной, то есть нумерацией), так и более сложной. Например, при двухфакторной классификации наблюдений индекс может состоять из трех номеров: два номера соответствуют уровням двух признаков, а третий номер — это номер наблюдения в группе.

Совокупности многомерных наблюдений образуют таблицы. Пусть X_α обозначает многомерное наблюдение, то есть вектор-столбец. Его размерность обозначим p (это натуральное число). Пусть индекс α пробегает некоторое множество A . Таблицей $\mathbf{T}$ называется совокупность

$$\mathbf{T} := \{X_\alpha, \alpha \in A\}. \quad (1)$$

В зависимости от способа нумерации (индексации) наблюдений таблицы получают разные формы. При линейной нумерации наблюдений таблица имеет форму строки, позиции которой заняты векторами-столбцами:

$$\mathbf{T} = \{X_i, i = \overline{1, n}\}.$$

Таблицы одной формы естественно образуют векторное пространство относительно операций поэлементного сложения и умножения элементов таблицы на числа. Помимо умножений на числа будем рассматривать поэлементное умножение таблиц на квадратные матрицы соответствующей размерности. Так, если K — квадратная матрица размерности p , то определим таблицу $K\mathbf{T}$, положив

$$K\mathbf{T} = \{KX_\alpha, \alpha \in A\}.$$

Для целей последующего статистического анализа в пространствах таблиц мы определим внутреннее произведение, которое за его свойства будем называть скалярным произведением. При задании скалярного произведения мы будем учитывать статистические свойства совокупности многомерных наблюдений $X_\alpha, \alpha \in A$.

Если заполняющие таблицу случайные наблюдения независимы и имеют одинаковые матрицы ковариаций (как в классическом многомерном статистическом анализе), то наиболее подходящим скалярным произведением двух таблиц является обобщенное евклидовское произведение. Более подробно: пусть $\mathbf{T}_X = \{X_\alpha, \alpha \in A\}$ и $\mathbf{T}_Y = \{Y_\alpha, \alpha \in A\}$ — две таблицы, формы которых одинаковы. Скалярным произведением таблиц $\mathbf{T}_X$ и $\mathbf{T}_Y$ назовем

$$\langle \mathbf{T}_X, \mathbf{T}_Y \rangle = \sum_{\alpha \in A} X_\alpha Y_\alpha^T. \quad (2)$$

Заметим, что скалярное произведение (2) — это $(p \times p)$ -матрица. Отсюда естественно следует, что для многомерных наблюдений квадратные матрицы соответствующей размерности должны играть роль скаляров. С помощью такого обобщения понятия скалярного произведения и скаляров в данной статье будет построена теория линейных моделей многомерного гауссовского статистического анализа, в основных чертах повторяющая сложившуюся теорию для одномерных линейных моделей.

Скалярное произведение, очевидно, некоммутативно: если $\mathbf{T}_1$ и $\mathbf{T}_2$ — две таблицы одной формы, то

$$\langle \mathbf{T}_1, \mathbf{T}_2 \rangle = \langle \mathbf{T}_2, \mathbf{T}_1 \rangle^T.$$

Скалярный квадрат таблицы $\mathbf{T}$ есть симметричная и неотрицательно определенная $(p \times p)$ -матрица

$$\langle \mathbf{T}, \mathbf{T} \rangle = \sum_{\alpha \in A} X_\alpha X_\alpha^T.$$

Свойства скалярного произведения в пространствах таблиц напоминают свойства традиционного скалярного произведения в евклидовых векторных пространствах: если $\mathbf{T}_1, \mathbf{T}_2, \mathbf{T}_3$ — это таблицы общей формы, то

$$\langle \mathbf{T}_1 + \mathbf{T}_2, \mathbf{T}_3 \rangle = \langle \mathbf{T}_1, \mathbf{T}_3 \rangle + \langle \mathbf{T}_2, \mathbf{T}_3 \rangle;$$

$$\langle K\mathbf{T}_1, \mathbf{T}_2 \rangle = K\langle \mathbf{T}_1, \mathbf{T}_2 \rangle, \text{ если } K \text{ — квадратная } (p \times p)\text{-матрица};$$

$$\langle \mathbf{T}, \mathbf{T} \rangle \geq 0 \text{ в смысле сравнения квадратных симметричных матриц};$$

$$\langle \mathbf{T}, \mathbf{T} \rangle = 0 \text{ тогда и только тогда, когда } \mathbf{T} = 0.$$

Таблицу $\mathbf{T}_1$ назовем *ортогональной* таблице $\mathbf{T}_2$, если

$$\langle \mathbf{T}_1, \mathbf{T}_2 \rangle = 0.$$

Заметим, что если $\langle \mathbf{T}_1, \mathbf{T}_2 \rangle = 0$, то и $\langle \mathbf{T}_2, \mathbf{T}_1 \rangle = 0$. Поэтому свойство ортогональности таблиц является взаимным.

Отметим форму теоремы Пифагора: если таблицы $\mathbf{T}_1$ и $\mathbf{T}_2$ ортогональны, то

$$\langle \mathbf{T}_1 + \mathbf{T}_2, \mathbf{T}_1 + \mathbf{T}_2 \rangle = \langle \mathbf{T}_1, \mathbf{T}_1 \rangle + \langle \mathbf{T}_2, \mathbf{T}_2 \rangle.$$

Можно сказать, что пространство таблиц — это результат тензорного умножения p -мерного векторного пространства столбцов на n -мерное векторное пространство строк. С алгебраической точки зрения пространство таблиц вместе с введенными выше операциями — это модуль над кольцом квадратных матриц соответствующей размерности.

На этот модуль естественно переносятся многие понятия, известные для евклидовых пространств. В частности, подмодуль — это совокупность таблиц, замкнутая относительно сложения и умножения на скаляры (то есть квадратные матрицы). Для подмодуля $\mathcal{L}$ определим *ортогональное дополнение* $\mathcal{L}^\perp$ как множество таблиц, ортогональных $\mathcal{L}$. Ясно, что $\mathcal{L}^\perp$ тоже является подмодулем.

Линейная нумерация. Всякую совокупность наблюдений можно занумеровать линейно, возможно, пожертвовав ради этой простоты более естественной в данной задаче системой индексации. Для теоретического анализа линейная нумерация представляет определенные удобства. Далее мы будем обсуждать статистические задачи с линейной нумерацией. Исходные данные мы будем представлять себе в виде строки, элементами которой служат p -столбцы:

$$\mathbf{T} = \mathbf{T}_X = \{X_i, i = \overline{1, n}\}. \quad (3)$$

Вместе с таблицами вида (3) будем рассматривать соответствующие им матрицы вида

$$\mathbf{X} = ||X_1, X_2, \dots, X_n||. \quad (4)$$

Матрица $\mathbf{X}$ — это матрица с p строками и n столбцами. Многие операции с таблицами можно проводить в их матричных формах. В частности, сложению таблиц взаимно однозначно соответствует сложение матриц, умножению элементов таблицы на квадратную $(p \times p)$ -матрицу K соответствует матричное произведение $K\mathbf{X}$; для таблиц

$$\mathbf{T}_X = \{X_i, i = \overline{1, n}\}, \quad \mathbf{T}_Y = \{Y_i, i = \overline{1, n}\}$$

скалярное произведение равно произведению соответствующих им матриц $\mathbf{X}$ и $\mathbf{Y}$:

$$\langle \mathbf{T}_X, \mathbf{T}_Y \rangle = \mathbf{X}\mathbf{Y}^T. \quad (5)$$

Для таблиц с линейной нумерацией (3) *линейные преобразования* пространства таблиц в себя имеют вид

$$\mathbf{T} \rightarrow \mathbf{T}\mathbf{A}, \quad (6)$$

где $\mathbf{A}$ — произвольная $(n \times n)$ -матрица. Произведение $\mathbf{T}\mathbf{A}$ определяем обычным для матриц способом («строка на столбец»), с той разницей, что элементами строки $\mathbf{T}$ служат не числа, но столбцы $X_1, \dots, X_n$. По определению

$$\mathbf{T}_X\mathbf{A} = \left\{ \sum_{j=1}^n X_j a_{ji}, i = \overline{1, n} \right\},$$

если $\mathbf{A} = \{a_{ij} \mid i, j = \overline{1, n}\}$.

Ортогональным преобразованием естественно называть такое линейное преобразование $\mathbf{A}$, которое сохраняет скалярное произведение:

$$\langle \mathbf{T}_X, \mathbf{T}_Y \rangle = \langle \mathbf{T}_X\mathbf{A}, \mathbf{T}_Y\mathbf{A} \rangle. \quad (7)$$

Очевидно, что ортогональные преобразования осуществляются ортогональными матрицами.

Если таблицу (3) рассматривать в матричной форме (4), это упомянутое линейное (6) преобразование с матрицей $\mathbf{A}$ представляется в виде произведения матриц

$$\mathbf{X} \rightarrow \mathbf{X}\mathbf{A}. \quad (8)$$

Порождающие базисы. В последующем изложении мы будем рассматривать таблицы вида (3) (с линейной нумерацией элементов), представляя их в основном в виде соответствующих матриц (4).

Матрицы, у которых p строк и n столбцов, будем называть $(p \times n)$ -матрицами. Совокупность $(p \times n)$ -матриц обозначаем $\mathbb{R}_n^p$. Матрицы размера $(p \times 1)$ будем называть p -столбцами или просто столбцами. Совокупность p -столбцов обозначим $\mathbb{R}_1^p$. Матрицы $(1 \times n)$ называем n -строками или просто строками. Совокупность n -строк обозначаем $\mathbb{R}_n^1$.

Особую роль в пространстве таблиц (3), когда они представлены в виде $(p \times n)$ -матриц (4), играют $(p \times n)$ -матрицы вида αx , где α — это p -столбец, x — это n -строка, символ αx обозначает произведение матриц $\alpha \in \mathbb{R}_1^p$, $x \in \mathbb{R}_n^1$.

Пусть n -строки $e_1, e_2, \dots, e_n \in \mathbb{R}_n^1$ и образуют базис пространства $\mathbb{R}_n^1$. Пусть $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{R}_1^p$ — произвольные p -столбцы. Рассмотрим $(p \times n)$ -матрицы $\alpha_1 e_1, \alpha_2 e_2, \dots, \alpha_n e_n$.

Теорема 1. *Всякую $(p \times n)$ -матрицу $\mathbf{X}$ (4) можно представить как*

$$\mathbf{X} = \sum_{i=1}^n \alpha_i e_i \quad (9)$$

при некотором выборе $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{R}_1^p$, и притом единственным образом.

Доказательство. Определим матрицу $\mathbf{E}$ размера $(n \times n)$, образованную n -строками $e_1, e_2, \dots, e_n$. Введем также переменную матрицу $\mathbf{A}$ размера $(p \times n)$, образованную p -столбцами $\alpha_1, \alpha_2, \dots, \alpha_n$. С помощью матриц $\mathbf{A}$ и $\mathbf{E}$ сумма (9) может быть записана как

$$\sum_{i=1}^n \alpha_i e_i = \mathbf{A}\mathbf{E}.$$

Теорема будет доказана, если мы покажем, что уравнение

$$\mathbf{X} = \mathbf{A}\mathbf{E} \quad (10)$$

имеет решение относительно $(p \times n)$ -матрицы $\mathbf{A}$, и притом единственное решение. Это решение очевидно:

$$\mathbf{A} = \mathbf{X}\mathbf{E}^{-1}. \quad (11)$$

Теорема позволяет говорить, что базис в $\mathbb{R}_n^1$ порождает (указанным способом) все пространство $\mathbb{R}_n^p$. Поэтому базисы в $\mathbb{R}_n^1$ мы будем применительно к $\mathbb{R}_n^p$ называть *порождающими базисами*. p -Столбцы $\alpha_1, \alpha_2, \dots, \alpha_n$ из разложения (9) $(p \times n)$ -матрицы $\mathbf{X}$ по порождающему базису $e_1, e_2, \dots, e_n \in \mathbb{R}_n^1$ можно рассматривать как координаты $\mathbf{X}$ в этом базисе. Для канонического базиса пространства $\mathbb{R}_n^1$ (когда e_i — это n -строка, в которой единица стоит на месте i , а прочие элементы равны нулю) координаты $\mathbf{X}$ относительно этого базиса совпадают с p -столбцами $X_1, \dots, X_n \in \mathbb{R}_1^p$, образующими $(p \times n)$ -матрицу $\mathbf{X}$.

Координаты $(p \times n)$ -матрицы $\mathbf{X}$ относительно разных порождающих базисов связаны линейным преобразованием. Пусть, например, n -строки $f_1, \dots, f_n$ образуют базис $\mathbb{R}_n^1$. Пусть $\mathbf{F}$ — $(n \times n)$ -матрица, составленная из этих n -строк. Согласно теореме 1, существует, и притом единственная, совокупность p -столбцов

$\beta_1, \beta_2, \dots, \beta_n$ — координат $\mathbf{X}$ относительно порождающего базиса $f_1, \dots, f_n$. Матрицы $\mathbf{B} = \|\beta_1, \dots, \beta_n\|$ и $\mathbf{F}$ связаны с $(p \times n)$ -матрицей $\mathbf{X}$ соотношением

$$\mathbf{X} = \mathbf{BF}. \quad (12)$$

Вместе с (10) это дает $\mathbf{BF} = \mathbf{AE}$. Отсюда $\mathbf{B} = \mathbf{AEF}^{-1}$, $\mathbf{A} = \mathbf{BFE}^{-1}$.

Следствие 1. Если порождающие базисы $e_1, \dots, e_n$ и $f_1, \dots, f_n$ ортогональные, то переход от координат относительно одного базиса к координатам относительно другого базиса осуществляется умножением на ортогональную матрицу.

Подмодули. Многие понятия классической линейной алгебры почти автоматически переносятся на пространства таблиц (3) и соответствующих им $(p \times n)$ -матриц с естественным расширением запаса скаляров до кольца $(p \times p)$ -матриц. В частности, подмодулем следует называть такое множество (таблиц или матриц), которое замкнуто относительно линейных операций.

Множество $\mathcal{L} \subset \mathbb{R}_n^p$ назовем подмодулем пространства $\mathbb{R}_n^p$, если для любых $\mathbf{X}_1, \mathbf{X}_2 \in \mathcal{L}$

$$K_1\mathbf{X}_1 + K_2\mathbf{X}_2 \in \mathcal{L}$$

при произвольных $(p \times p)$ -матрицах K_1, K_2 .

Теорема 2. Всякий подмодуль $\mathcal{L}, \mathcal{L} \subset \mathbb{R}_n^p$, порождается некоторой линейно-независимой системой n -строк. Число элементов в этой системе единственным образом определяется $\mathcal{L}$. Это число может быть названо размерностью подмодуля $\mathcal{L}$.

Доказательство. Пусть $\mathbf{X} \in \mathcal{L}$. Множество $(p \times n)$ -матриц вида $K\mathbf{X}$, где K — произвольная $(p \times p)$ -матрица, образует подмодуль. Обозначим его $\mathcal{L}(\mathbf{X})$. Пусть $x_1, \dots, x_p$ — это n -строки $(p \times n)$ -матрицы $\mathbf{X}$. Выберем среди этих n -строк максимальную, линейно независимую подсистему, скажем $y_1, \dots, y_k$. Очевидно, что

$$\mathcal{L}(\mathbf{X}) = \{\mathbf{Y} \mid \mathbf{Y} = \sum_{i=1}^k \beta_i y_i, \text{ где } \beta_1, \dots, \beta_k \in \mathbb{R}_1^p \text{ независимо пробегают } \mathbb{R}_1^p\}.$$

Если $\mathcal{L}(\mathbf{X}) = \mathcal{L}$, то $y_1, \dots, y_k$ составляют для $\mathcal{L} \subset \mathbb{R}_n^p$ образующий базис. Если $\mathcal{L}(\mathbf{X}) \neq \mathcal{L}$, то находим в $\mathcal{L}$ элемент, скажем $\mathbf{Z}$, не принадлежащий $\mathcal{L}(\mathbf{X})$. Пополняем систему $y_1, \dots, y_k$ n -строками $z_1, \dots, z_p$ $(p \times n)$ -матрицы $\mathbf{Z}$, находим в этой совокупности n -строк максимальную, линейно независимую подсистему и т.д. На некотором шаге процесс остановится.

Следствие 2. Всякий подмодуль $\mathcal{L} \subset \mathbb{R}_n^p$ может быть представлен как прямая сумма некоторых одномерных подмодулей $\mathcal{L}_i \subset \mathbb{R}_n^p$:

$$\mathcal{L} = \mathcal{L}_1 \oplus \mathcal{L}_2 \oplus \dots \oplus \mathcal{L}_l, \quad (13)$$

где

$$\mathcal{L}_i = \{\mathbf{X} \mid \mathbf{X} = \alpha y_i, y_i \in \mathbb{R}_n^1, \alpha \text{ пробегает } \mathbb{R}_1^p\}.$$

Число l — одно и то же в любом представлении (13). Это число можно называть размерностью подмодуля $\mathcal{L}$: $l = \dim \mathcal{L}$.

Замечание 1. Линейно-независимую систему n -строк, которая порождает $\mathcal{L}$, можно выбрать ортогональной. Для доказательства достаточно заметить, что порождающую систему можно превратить в ортогональную процессом ортогонализации.

Замечание 2. Теорема 2 устанавливает взаимно однозначное соответствие между линейными подпространствами векторного пространства $\mathbb{R}_n^1$ и подмодулями пространства $(p \times n)$ -матриц $\mathbb{R}_n^p$.

Проекция на подмодуль. Для всякого подмодуля $\mathcal{L} \in \mathbb{R}_n^p$ определим его ортогональное дополнение (до всего подпространства $\mathbb{R}_n^p$) $\mathcal{L}^\perp$ как

$$\mathcal{L}^\perp = \{\mathbf{Y} \mid \mathbf{Y} \in \mathbb{R}_n^p \text{ и } \langle \mathbf{Y}, \mathbf{X} \rangle = 0 \text{ для всякого } \mathbf{X} \in \mathcal{L}\}. \quad (14)$$

Очевидно, что $\mathcal{L}^\perp$ является подмодулем и что

$$\mathcal{L} \oplus \mathcal{L}^\perp = \mathbb{R}_n^p.$$

Назовем $(p \times n)$ -матрицу $\mathbf{Z}$ проекцией $(p \times n)$ -матрицы $\mathbf{X}$ на подмодуль $\mathcal{L} \subset \mathbb{R}_n^p$, если

- а) $\mathbf{Z} \in \mathcal{L}$;
- б) $\mathbf{X} - \mathbf{Z} \in \mathcal{L}^\perp$.

Проекцию $\mathbf{X}$ на $\mathcal{L}$ далее обозначаем $\text{proj}_{\mathcal{L}} \mathbf{X}$.

Проецирование на заданный подмодуль, очевидно, является линейной операцией. Важное свойство проецирования $(p \times n)$ -матриц повторяет известное свойство проецирования в евклидовых пространствах: $\text{proj}_{\mathcal{L}} \mathbf{X}$ есть ближайший к $\mathbf{X}$ элемент $\mathcal{L}$. Сформулируем это в виде теоремы:

Теорема 3. Для любой $(p \times n)$ -матрицы $\mathbf{Z} \in \mathcal{L}$ выполняется соотношение

$$\langle \mathbf{X} - \mathbf{Z}, \mathbf{X} - \mathbf{Z} \rangle \geq \langle \mathbf{X} - \text{proj}_{\mathcal{L}} \mathbf{X}, \mathbf{X} - \text{proj}_{\mathcal{L}} \mathbf{X} \rangle \quad (15)$$

в смысле сравнения квадратных симметричных матриц.

Доказательство. Неравенство (15) будет доказано, если мы покажем, что для всякого p -столбца выполняется неравенство между квадратичными формами

$$\lambda^T \langle \mathbf{X} - \mathbf{Z}, \mathbf{X} - \mathbf{Z} \rangle \lambda \geq \lambda^T \langle \mathbf{X} - \text{proj}_{\mathcal{L}} \mathbf{X}, \mathbf{X} - \text{proj}_{\mathcal{L}} \mathbf{X} \rangle \lambda.$$

Рассмотрим n -строки $\xi = \lambda^T \mathbf{X}$, $\zeta = \lambda^T \mathbf{Z}$ и линейное подпространство n -строк $L = \lambda^T \mathcal{L}$. Очевидно, что

$$\lambda^T \text{proj}_{\mathcal{L}} \mathbf{X} = \text{proj}_L \xi$$

и что для n -строк $\xi, \zeta \in \mathbb{R}_n^1$ и линейного подпространства $L \subset \mathbb{R}_n^1$ выполняется неравенство

$$|\xi - \zeta|^2 \geq |\xi - \text{proj}_L \xi|^2. \quad (16)$$

Остается заметить, что (16) является всего лишь другой записью (15).

2. Многомерный статистический анализ

Многомерные линейные модели. Предположим, что составляющие таблицу $\mathbf{T}_X$ (3) элементы являются p -мерными случайными величинами. Математическое ожидание таблиц, их вариации и ковариации вводятся естественно:

$$\begin{aligned} \mathbf{E}\mathbf{T}_X &= \{\mathbf{E}X_i, i = \overline{1, n}\}, \\ \text{Cov}(\mathbf{T}_X, \mathbf{T}_Y) &= \{\mathbf{E}(X_i - \mathbf{E}X_i)(Y_j - \mathbf{E}Y_j)^T \mid i, j = \overline{1, n}\}. \end{aligned}$$

Ковариацию таблиц $\mathbf{T}_X$ и $\mathbf{T}_Y$ мы будем рассматривать как квадратную матрицу размера $(n \times n)$, элементами которой служат $(p \times p)$ -матрицы $\text{Cov}(X_i, Y_j) = \mathbf{E}(X_i - \mathbf{E}X_i)(Y_j - \mathbf{E}Y_j)^T$.

Если элементы случайной таблицы $\mathbf{T}_X$ независимы и имеют общую матрицу ковариаций Σ , то

$$\text{Var } \mathbf{T}_X = \{\mathbf{E}(X_i - \mathbf{E}X_i)(X_j - \mathbf{E}X_j)^T \mid i, j = \overline{1, n}\} = I \otimes \Sigma, \quad (17)$$

где символ $\otimes$ означает кронекеровское произведение. В дальнейшем объектом анализа будут случайные таблицы именно с такой матрицей ковариаций.

Приведем известную теорему об ортогональных преобразованиях таблиц со случайными элементами (см., например, [3, теорема 3.3.1]).

Лемма. Пусть элементы таблицы (3) суть независимые, нормально распределенные p -столбцы с общей матрицей ковариаций Σ . Пусть $\mathbf{C}$ — ортогональная $(n \times n)$ -матрица. Тогда

$$\text{Var } \mathbf{T}\mathbf{C} = \text{Var } \mathbf{T} = I \otimes \Sigma.$$

Доказательство. Из условий леммы и (17) следует, что $\text{Var } \mathbf{T} = I \otimes \Sigma$. Для упрощения записи можем предположить, что $\mathbf{E}\mathbf{T} = 0$. Тогда $\text{Var } \mathbf{T}\mathbf{C} = \mathbf{E}(\mathbf{T}\mathbf{C})(\mathbf{T}\mathbf{C})^T = \mathbf{E}\mathbf{T}\mathbf{T}^T = I \otimes \Sigma$.

Определение 1. Мы скажем, что таблица $\mathbf{T}_X$ (3) со случайными элементами $X_1, \dots, X_n$ и соответствующая ей матрица $\mathbf{X}$ (4) следуют линейной гауссовской модели, если

- (а) $\mathbf{E}\mathbf{X} \in \mathcal{L}$ для некоторого заданного подмодуля $\mathcal{L}$;
- (б) наблюдения $X_1, \dots, X_n$ независимы и распределены по нормальному закону с общей матрицей ковариаций Σ .

Параметры модели $\mathbf{E}\mathbf{X}$ и Σ обычно неизвестны; матрица Σ предполагается невырожденной.

$(p \times n)$ -матрицу $\mathbf{X}$, если она следует линейной модели, мы будем представлять в виде

$$\mathbf{X} = \mathbf{M} + \mathcal{E}, \quad (18)$$

где $\mathbf{M} = \mathbf{E}\mathbf{X}$,

$$\mathcal{E} = \|E_1, E_2, \dots, E_n\|$$

— это $(p \times n)$ -матрица, p -столбцы которой $E_1, E_2, \dots, E_n$ суть независимые, одинаково распределенные $N_p(0, \Sigma)$. Матрица Σ предполагается в дальнейшем невырожденной. Параметры $\mathbf{M}$, Σ неизвестны.

Отметим простую, но идейно важную теорему.

Теорема 4. *Каждой многомерной линейной модели (18) соответствует некоторая линейная модель с одномерными откликами, и обратно.*

Теорема справедлива в силу установленного ранее взаимно однозначного соответствия между подмодулями пространства $(p \times n)$ -матриц $\mathbb{R}_n^p$ и линейными подпространствами векторного пространства $\mathbb{R}_n^1$ (см. замечание 2 к теореме 2).

Таким образом, «запас» многомерных линейных моделей таков же, как и одномерных.

Нетрудно убедиться, что обычно рассматриваемые многомерным линейным анализом задачи (факторные классификации, линейная регрессия) являются в нашем смысле линейными моделями.

3. Достаточные статистики, наилучшие несмещенные оценки

Рассмотрим гауссовскую линейную модель (18). Пусть $\mathbf{M} = \|M_1, M_2, \dots, M_n\|$. Правдоподобие пары $\mathbf{M}, \Sigma$ по наблюдению $\mathbf{X}$ равно

$$\begin{aligned} & \prod_{i=1}^n \left(\frac{1}{\sqrt{2\pi}} \right)^p \frac{1}{\sqrt{\det \Sigma}} \exp \left\{ -\frac{1}{2} (X_i - M_i)^T \Sigma^{-1} (X_i - M_i) \right\} = \\ & = \left(\frac{1}{\sqrt{2\pi}} \right)^{np} \left(\frac{1}{\sqrt{\det \Sigma}} \right)^n \exp \left\{ -\frac{1}{2} \operatorname{tr} \Sigma^{-1} \left[\sum_{i=1}^n (X_i - M_i)(X_i - M_i)^T \right] \right\}. \end{aligned} \quad (19)$$

Сумма в квадратных скобках — это $\langle \mathbf{X} - \mathbf{M}, \mathbf{X} - \mathbf{M} \rangle$. Представим $\mathbf{X} - \mathbf{M}$ как

$$\mathbf{X} - \mathbf{M} = (\mathbf{X} - \operatorname{proj}_{\mathcal{L}} \mathbf{X}) + (\operatorname{proj}_{\mathcal{L}} \mathbf{X} - \mathbf{M}) = (1) + (2)$$

и заметим, что

$$(1) = \operatorname{proj}_{\mathcal{L}^\perp} \mathbf{X} \in \mathcal{L}^\perp, \quad (2) \in \mathcal{L}.$$

Поэтому (теорема Пифагора)

$$\langle \mathbf{X} - \mathbf{M}, \mathbf{X} - \mathbf{M} \rangle = \langle \operatorname{proj}_{\mathcal{L}^\perp} \mathbf{X}, \operatorname{proj}_{\mathcal{L}^\perp} \mathbf{X} \rangle + \langle \operatorname{proj}_{\mathcal{L}} \mathbf{X} - \mathbf{M}, \operatorname{proj}_{\mathcal{L}} \mathbf{X} - \mathbf{M} \rangle. \quad (20)$$

Закключаем отсюда, что правдоподобие (19) выражается через статистики $\operatorname{proj}_{\mathcal{L}} \mathbf{X}$ и $\langle \operatorname{proj}_{\mathcal{L}^\perp} \mathbf{X}, \operatorname{proj}_{\mathcal{L}^\perp} \mathbf{X} \rangle$, которые вместе достаточны для $\mathbf{M}, \Sigma$.

Статистика $\operatorname{proj}_{\mathcal{L}} \mathbf{X}$ очевидно является несмещенной оценкой $\mathbf{M}$. Как функция достаточной статистики она является наилучшей несмещенной оценкой $\mathbf{M}$. То, что наилучшей несмещенной оценкой Σ служит статистика

$$\frac{1}{\dim \mathcal{L}^\perp} \langle \operatorname{proj}_{\mathcal{L}^\perp} \mathbf{X}, \operatorname{proj}_{\mathcal{L}^\perp} \mathbf{X} \rangle, \quad (21)$$

мы установим после доказательства следующей далее теоремы.

Теорема об ортогональном разложении. Справедлива следующая

Теорема 5. *Пусть $\mathbf{X} = \|X_1, X_2, \dots, X_n\|$ — гауссовская $(p \times n)$ -матрица с независимыми p -столбцами $X_1, X_2, \dots, X_n \in \mathbb{R}_1^p$, причем $\operatorname{Var} X_i = \Sigma$ для всех $i = 1, \dots, n$. Пусть $\mathcal{L}_1, \mathcal{L}_2, \dots$ — попарно ортогональные подмодули $\mathbb{R}_n^p$, прямая сумма которых составляет $\mathbb{R}_n^p$:*

$$\mathbb{R}_n^p = \mathcal{L}_1 \oplus \mathcal{L}_2 \oplus \dots$$

Рассмотрим разложение $(p \times n)$ -матрицы $\mathbf{X}$ на попарно ортогональные составляющие

$$\mathbf{X} = \text{proj}_{\mathcal{L}_1} \mathbf{X} + \text{proj}_{\mathcal{L}_2} \mathbf{X} + \dots$$

Тогда:

а) случайные $(p \times n)$ -матрицы $\text{proj}_{\mathcal{L}_1} \mathbf{X}, \text{proj}_{\mathcal{L}_2} \mathbf{X}, \dots$ независимы, распределены нормально и $E \text{proj}_{\mathcal{L}_i} \mathbf{X} = \text{proj}_{\mathcal{L}_i} E \mathbf{X}$;

б) $\langle \text{proj}_{\mathcal{L}_i} \mathbf{X}, \text{proj}_{\mathcal{L}_j} \mathbf{X} \rangle = W_p(\dim L_i, \Sigma, \Delta_i)$, где $W_p(\dim L_i, \Sigma, \Delta_i)$ обозначает случайную матрицу (размера $p \times p$), распределенную по Уишарту, с ν степенями свободы и параметром нецентральности Δ . В данном случае

$$\Delta_i = \langle \text{proj}_{\mathcal{L}_i} E \mathbf{X}, \text{proj}_{\mathcal{L}_i} E \mathbf{X} \rangle.$$

Доказательство. Для каждого из подмодулей $\mathcal{L}_1, \mathcal{L}_2, \dots \subset \mathbb{R}_n^p$, размерности которых обозначаем $m_1, m_2, \dots$, в пространстве $\mathbb{R}_n^1$ находим ортонормированные системы n -строк, которые порождают эти подпространства. Пусть для $\mathcal{L}_1$ — это n -строки $f_1, \dots, f_{m_1}$; для $\mathcal{L}_2$ — это n -строки $f_{m_1+1}, \dots, f_{m_1+m_2}$ и т.д. С помощью этих n -строк каждый подмодуль $\mathcal{L}_1, \mathcal{L}_2, \dots$ можно представить как прямую сумму одномерных подмодулей из $\mathbb{R}_n^p$. Например, $\mathcal{L}_1 = \Phi_1 \oplus \Phi_2 \oplus \dots \oplus \Phi_{m_1}$, где

$$\Phi_1 = \{\mathbf{Y} \mid \mathbf{Y} = \alpha f_1, \alpha \text{ пробегает } \mathbb{R}_1^p\},$$

$$\Phi_2 = \{\mathbf{Y} \mid \mathbf{Y} = \alpha f_2, \alpha \text{ пробегает } \mathbb{R}_1^p\},$$

...

$$\Phi_{m_1} = \{\mathbf{Y} \mid \mathbf{Y} = \alpha f_{m_1}, \alpha \text{ пробегает } \mathbb{R}_1^p\}.$$

Совокупность всех n -строк $f_1, f_2, \dots, f_n$ дает ортонормальный базис $\mathbb{R}_n^1$ и одновременно образующий базис пространства $\mathbb{R}_n^p$. Поэтому всякую $(p \times n)$ -матрицу $\mathbf{X} \in \mathbb{R}_n^p$ можно представить в виде

$$\mathbf{X} = \sum_{i=1}^n Y_i f_i, \quad Y_1, \dots, Y_n \in \mathbb{R}_1^p.$$

При этом

$$\text{proj}_{\mathcal{L}_1} \mathbf{X} = \sum_{i=1}^{m_1} Y_i f_i, \quad \text{proj}_{\mathcal{L}_2} \mathbf{X} = \sum_{i=m_1+1}^{m_1+m_2} Y_i f_i \quad \text{и т.д.}$$

p -Столбцы $Y_1, Y_2, \dots, Y_n$ — это координаты $(p \times n)$ -матрицы $\mathbf{X}$ относительно порождающего базиса $f_1, \dots, f_n$. При этом p -столбцы $X_1, X_2, \dots, X_n$ — это координаты той же $(p \times n)$ -матрицы $\mathbf{X}$ относительно ортогонального канонического базиса $\mathbb{R}_n^1$: $e_1 = (1, 0, \dots)$, $e_2 = (0, 1, 0, \dots)$ и т.д. Как было отмечено ранее (п. 1), переход от одних координат к другим осуществляется умножением $(p \times n)$ -матрицы $\mathbf{X}$ справа на $(n \times n)$ -матрицу перехода, в данном случае на некоторую ортогональную $(n \times n)$ -матрицу $\mathbf{C}$:

$$\|Y_1, Y_2, \dots, Y_n\| = \|X_1, X_2, \dots, X_n\| \mathbf{C}, \quad \text{или} \quad \mathbf{Y} = \mathbf{X} \mathbf{C}.$$

В силу леммы p -столбцы $Y_1, \dots, Y_n$ в совокупности независимы и распределены нормально с общей для всех них матрицей ковариаций Σ .

Рассмотрим случайные $(p \times p)$ -матрицы

$$\langle \text{proj}_{\mathcal{L}_1} \mathbf{X}, \text{proj}_{\mathcal{L}_1} \mathbf{X} \rangle, \langle \text{proj}_{\mathcal{L}_2} \mathbf{X}, \text{proj}_{\mathcal{L}_2} \mathbf{X} \rangle, \dots$$

Например,

$$\langle \text{proj}_{\mathcal{L}_1} \mathbf{X}, \text{proj}_{\mathcal{L}_1} \mathbf{X} \rangle = \sum_{i=1}^{m_1} Y_i Y_i^T.$$

Распределения таких случайных матриц называют распределениями Уишарта. Если при этом $\mathbf{E}Y_1 = \mathbf{E}Y_2 = \dots = \mathbf{E}Y_{m_1} = 0$, мы получаем так называемое центральное распределение Уишарта $W_p(m_1, \Sigma)$. Заметим, что если использовать обозначение $W_p(m, \Sigma)$ не только для закона распределения, но и для случайной матрицы с таким распределением, то можно сказать, что

$$W_p(m, \Sigma) = \Sigma^{\frac{1}{2}} W_p(m, I) \Sigma^{\frac{1}{2}},$$

если обозначить $\Sigma^{\frac{1}{2}}$ симметрическую матрицу, — единственное симметричное решение матричного уравнения $Z^2 = \Sigma$.

Говорят, что случайная $(p \times p)$ -матрица W имеет нецентральное распределение Уишарта, если

$$W = \sum_{i=1}^m (\xi_i + a_i)(\xi_i + a_i)^T,$$

где p -столбцы $\xi_1, \xi_2, \dots, \xi_m$ суть независимые $N_p(0, \Sigma)$, $a_1, a_2, \dots, a_m$ — некоторые неслучайные p -столбцы, вообще говоря, отличные от нуля. Распределение W каким-то образом зависит от p -столбцов $a_1, a_2, \dots, a_m$. Покажем, что распределение W зависит от упомянутых p -столбцов через так называемый параметр нецентральности: $(p \times p)$ -матрицу

$$\Delta = \sum_{i=1}^m a_i a_i^T.$$

Введем $(p \times m)$ -матрицы

$$\begin{aligned} \xi &= \|\xi_1, \xi_2, \dots, \xi_m\|, \\ \mathbf{A} &= \|a_1, a_2, \dots, a_m\|. \end{aligned}$$

В этих обозначениях

$$W = \langle \xi + \mathbf{A}, \xi + \mathbf{A} \rangle.$$

Пусть $\mathbf{C}$ — произвольная ортогональная $(m \times m)$ -матрица. Положим $\eta = \xi \mathbf{C}$. При этом по распределению $\eta = \xi$ и

$$W = \langle \eta + \mathbf{A} \mathbf{C}, \eta + \mathbf{A} \mathbf{C} \rangle.$$

Мы видим, что нецентральное распределение Уишарта зависит от $\mathbf{A} = \|a_1, \dots, a_m\|$ не непосредственно, а через максимальный инвариант $\mathbf{A}$ при ортогональных преобразованиях, то есть через $\langle \mathbf{A}, \mathbf{A} \rangle = \sum_{i=1}^m a_i a_i^T$.

Следовательно, в общем случае

$$\langle \text{proj}_{\mathcal{L}_i} \mathbf{X}, \text{proj}_{\mathcal{L}_i} \mathbf{X} \rangle = W_p(m_i, Q, \Delta_i),$$

где $\Delta_i = \langle \text{proj}_{\mathcal{L}_i} \mathbf{E}\mathbf{X}, \text{proj}_{\mathcal{L}_i} \mathbf{E}\mathbf{X} \rangle$. Доказательство теоремы закончено.

Вернемся к несмещенным оценкам параметров линейной модели. В линейной модели (18) $\text{proj}_{\mathcal{L}^\perp} \mathbf{E}\mathbf{X} = 0$. Поэтому статистика (21) есть

$$\frac{1}{n-m} \Sigma^{\frac{1}{2}} W_p(n-m, I) \Sigma^{\frac{1}{2}}.$$

Очевидно, что ее математическое ожидание равно Σ .

Проверка линейных гипотез. Копируя одномерную линейную модель, линейной гипотезой в многомерной линейной модели (18) назовем гипотезу

$$H : \mathbf{E}\mathbf{X} \in \mathcal{L}_1, \quad (22)$$

где $\mathcal{L}_1$ — заданный подмодуль, причем $\mathcal{L}_1 \subset \mathcal{L}$.

В этом разделе мы предложим статистики, которые могут служить основой для построения статистических критериев для проверки H (22), свободных (при H) от параметров $\mathbf{M}$, Σ .

Введем подмодуль $\mathcal{L}_2$, ортогонально дополняющий $\mathcal{L}_1$ до $\mathcal{L}$:

$$\mathcal{L} = \mathcal{L}_1 \oplus \mathcal{L}_2. \quad (23)$$

Рассмотрим разложение пространства $\mathbb{R}_n^p$ на три попарно ортогональных подмодуля

$$\mathbb{R}_n^p = \mathcal{L}_1 \oplus \mathcal{L}_2 \oplus \mathcal{L}^\perp.$$

В силу теоремы 5 случайные матрицы

$$S_1 := \langle \text{proj}_{\mathcal{L}^\perp} \mathbf{X}, \text{proj}_{\mathcal{L}^\perp} \mathbf{X} \rangle \quad \text{и} \quad S_2 := \langle \text{proj}_{\mathcal{L}_2} \mathbf{X}, \text{proj}_{\mathcal{L}_2} \mathbf{X} \rangle$$

статистически независимы и распределены по Уишарту. При этом

$$S_1 = \langle \text{proj}_{\mathcal{L}^\perp} \mathbf{X}, \text{proj}_{\mathcal{L}^\perp} \mathbf{X} \rangle = W_p(n-m, \Sigma). \quad (24)$$

Если гипотеза H (22) верна, то

$$S_2 = \langle \text{proj}_{\mathcal{L}_2} \mathbf{X}, \text{proj}_{\mathcal{L}_2} \mathbf{X} \rangle = W_p(m_2, \Sigma). \quad (25)$$

(Здесь и далее приняты обозначения $m = \dim \mathcal{L}$, $m_1 = \dim \mathcal{L}_1$, $m_2 = \dim \mathcal{L}_2$.)

При нарушении гипотезы H (22) распределение Уишарта статистики (25) становится нецентральным, приобретая параметр нецентральности

$$\Delta = \langle \text{proj}_{\mathcal{L}_2} \mathbf{E}\mathbf{X}, \text{proj}_{\mathcal{L}_2} \mathbf{E}\mathbf{X} \rangle,$$

который тем больше, чем «сильнее» нарушена гипотеза H (22).

В одномерном случае (когда $p = 1$) статистики (24), (25) превращаются в случайные величины, которые распределены как $\sigma^2 \chi^2(n-m)$ и $\sigma^2 \chi^2(m_2)$, и отношение

которых распределено (при гипотезе) свободно и потому может быть использовано как статистический критерий для проверки H .

В многомерном случае аналог эф-отношения — это отношение $(p \times p)$ -матриц S_2 и S_1 . При $n - m \geq p$ матрица S_1 невырождена, и потому существует статистика ($p \times p$ -матрица)

$$\langle \text{proj}_{\mathcal{L}_2} \mathbf{X}, \text{proj}_{\mathcal{L}_2} \mathbf{X} \rangle \langle \text{proj}_{\mathcal{L}^\perp} \mathbf{X}, \text{proj}_{\mathcal{L}^\perp} \mathbf{X} \rangle^{-1}. \quad (26)$$

В отличие от одномерного случая ($p = 1$) статистика (26) не распределена свободно: по распределению она равна

$$\Sigma^{\frac{1}{2}} W_p(m_2, I) W_p^{-1}(n - m, I) \Sigma^{-\frac{1}{2}}. \quad (27)$$

Однако собственные значения матрицы (26) при гипотезе H (22) распределены свободно от $\mathbf{M}$, Σ . Эти собственные значения совпадают с корнями уравнения (относительно λ)

$$\det(W_p(m_2, I) - \lambda W_p(n - m, I)) = 0. \quad (28)$$

Поэтому некоторые функции от корней уравнения (28) традиционно используют в качестве критических статистик для проверки линейных гипотез.

Здесь наше исследование выходит на традиционную дорогу многомерного статистического анализа и потому должно быть закончено.

Приношу благодарность Е.М. Сухановой за обсуждения и помощь в работе.

Список литературы

- [1] Колмогоров А.Н. К обоснованию методов наименьших квадратов // Успехи матем. наук. 1946. **1**, № 1. С. 57–70 (См. также: Колмогоров А.Н. Теория вероятностей и математическая статистика (сб. статей). М.: Наука, 1986. С. 267–283.).
- [2] Durbin J., Kendall M.G. The geometry of estimation // Biometrika. Jun. 1951. **38**, № 1/2. P. 150–158.
- [3] Андерсон Т. Введение в многомерный статистический анализ. М.: Физматлит, 1963.
- [4] Шеффе Г. Дисперсионный анализ. М.: Физматгиз, 1963.
- [5] Roy S.N. Some Aspects of Multivariate Analysis. Wiley, 1957.
- [6] Bilodeau M., Brenner D. Theory of Multivariate Statistics. Springer-Verlag, 1999.

ОБ ИССЛЕДОВАНИИ ВЕТВЯЩИХСЯ СЛУЧАЙНЫХ БЛУЖДЕНИЙ ПО МНОГОМЕРНЫМ РЕШЕТКАМ

Е. Б. Яровая

В статье рассматриваются две модели случайного блуждания с непрерывным временем на $\mathbb{Z}^d$ с размножением и гибелью частиц в одном из узлов решетки (источнике ветвления). Основное различие между моделями определяется поведением процесса в источнике. В модели I случайное блуждание предполагается симметричным на всей решетке, а в модели II вводится дополнительный параметр, позволяющий «искусственно» усиливать степень преобладания свойства ветвления или блуждания в источнике и в результате нарушающий симметричность случайного блуждания. Получены дифференциальные уравнения для производящих функций численностей частиц как в произвольном узле решетки, так и на всей решетке в модели II. Установлены достаточные условия, при которых в модели II будет наблюдаться экспоненциальный рост процесса. Предложены общие методы исследования моделей в надкритическом случае.

1. Введение

В различных областях естественных наук все более широкое применение находят модели, которые описываются в терминах ветвящихся процессов с блужданием частиц. Подобные эволюционные процессы существенным образом зависят от структуры среды, в которой происходит блуждание. Как известно, неоднородность среды играет определяющую роль в формировании аномальных свойств транспорта частиц. Такие эффекты наблюдаются также и в случайных средах, для которых характерно возникновение нерегулярных структур с выраженной неоднородностью пространственного распределения, вызванное наличием «сильных центров», в окрестности которых происходит основной рост процесса [1, 2]. Это объясняет особый интерес к простейшим ветвящимся случайным блужданиям в неоднородных средах, позволяющих детально изучить эффекты, связанные с неоднородностью, некомпактностью, а также размерностью пространства, в котором происходит блуждание.

Одной из таких моделей служит симметричное ветвящееся случайное блуждание с непрерывным временем на решетке $\mathbb{Z}^d$, в которой размножение и гибель частиц происходят в единственной точке (*источнике*). Эволюция частиц в этой системе определяется структурой спектра линейного оператора $H = A + \beta \delta_0(x)$, где генератор случайного блуждания A является ограниченным самосопряженным оператором в $l^2(\mathbb{Z}^d)$, а параметр β характеризует интенсивность источника. Подробному исследованию и развитию этой модели посвящены работы [3–9].

Аналогичные постановки возникли при описании систем массового обслуживания со случайным числом независимых серверов [10], в которых рассматривалось критическое ветвящееся случайное блуждание на $\mathbb{Z}$. Отличительной особенностью данной модели стало введение параметра α , управляющего поведением процесса в источнике ветвления и приводящего к нарушению симметричности матрицы переходных интенсивностей A . Основной идеей доказательств для такой модификации

(см. [10–12]) явилось представление ветвящегося случайного блуждания как ветвящегося процесса Беллмана-Харриса с двумя типами частиц.

Как следует из [8, 10, 13], на $\mathbb{Z}$ и $\mathbb{Z}^2$ асимптотическое поведение вероятности продолжения критического ветвящегося случайного блуждания как в источнике, так и на всей решетке совпадает в обеих моделях с точностью до умножения на некоторую константу. В то же время, несмотря на близость постановок, эти модели все же различны. Более того, они сформулированы в трудносопоставимых терминах, отличается также и техника доказательств. В связи с этим актуальным становится вопрос формального сравнения упомянутых моделей и развития обобщающего подхода к их исследованию. В статье предложены общие методы исследования надкритических ветвящихся случайных блужданий по D -мерным решеткам, при которых наблюдается экспоненциальный рост численностей частиц как в произвольных узлах, так и на всей решетке, и установлены достаточные условия существования таких процессов в модели, введенной в [10].

2. Описание моделей

В обеих моделях исследуется эволюция системы частиц на $\mathbb{Z}^d$, состояние которой описывается числом $\mu_t(y)$ частиц в момент времени t в каждой точке $y \in \mathbb{Z}^d$, в предположении, что в момент времени $t = 0$ (начальный момент) система состоит из одной частицы, находящейся в точке x , то есть $\mu_0(y) = \delta_x(y)$. Частицы совершают случайное блуждание по $\mathbb{Z}^d$, причем в одной из точек решетки x_0 частицы могут к тому же «размножаться и гибнуть» — точное значение этих слов будет видно из нижеследующего описания моделей. Предполагается также, что каждая из новых частиц эволюционирует по тому же закону независимо от остальных частиц и от всей предыстории.

Модель I. Случайное блуждание частиц полностью определяется матрицей переходных интенсивностей $A = \|a(x, y)\|_{x, y \in \mathbb{Z}^d}$ и предполагается

- 1) симметричным: $a(x, y) = a(y, x)$,
- 2) однородным: $a(x, y) = a(0, y - x) = a(y - x)$,
- 3) неприводимым (то есть, любая точка $y \in \mathbb{Z}^d$ достижима),
- 4) регулярным: $\sum_{x \in \mathbb{Z}^d} a(x) = 0$ с $a(x) \geq 0$, $x \neq 0$, $a(0) < 0$,
- 5) с конечной дисперсией скачков: $\sum_{x \in \mathbb{Z}^d} x^2 a(x) < \infty$.

В силу свойств симметричности и однородности случайного блуждания для матрицы A выполнены условия $\sum_{y \in \mathbb{Z}} a(x, y) = 0$ и $\sum_{x \in \mathbb{Z}} a(x, y) = 0$.

Ветвящийся процесс в источнике $x_0 = 0$ задается с помощью инфинитезимальной производящей функции

$$f(u) = \sum_{n=0}^{\infty} b_n u^n, \quad (1)$$

где $b_n \geq 0$ при $n \neq 1$, $b_1 < 0$ и $\sum_n b_n = 0$. Будем полагать, что

$$\beta_r = f^{(r)}(1) < \infty, \quad r \in \mathbb{N}, \quad \beta = \beta_1 = f'(1). \quad (2)$$

Таким образом, если в момент времени t в точке $x_0 = 0$ находилось $\mu_t(0) > 0$ частиц, то каждая из частиц (независимо от прочих), находящихся в точке $x_0 = 0$,

может за время $[t, t + h)$ либо перейти с вероятностью $p(h, 0, y) = a(y)h + o(h)$ в точку $y \neq 0$, либо произвести потомство из $n \neq 1$ частиц (считается, что и сама частица входит в эти n) или погибнуть (случай, когда $n = 0$) с вероятностью $p_*(h, n) = b_n h + o(h)$, либо сохраниться (никаких изменений не происходит) с вероятностью

$$1 - \sum_{y \neq 0} a(y)h - \sum_{n \neq 1} b_n h + o(h).$$

Стандартным способом [15] можно показать, что время пребывания частицы в источнике имеет экспоненциальное распределение с параметром $a(0) + b_1$. В данной модели не предполагается независимость законов блуждания и ветвления в источнике, поскольку «гибель и размножение» в точке x_0 происходит лишь при условии, что частица находится в источнике и не совершила скачка в какую-нибудь другую точку решетки за время $[t, t + h)$.

Модель II. В модели II перемещение частиц по решетке $\mathbb{Z}^d$ (за исключением тех, которые выходят из нуля, не произведя потомства) также происходит в соответствии с матрицей интенсивностей переходных вероятностей $A = \|a(x, y)\|_{x \in \mathbb{Z}^d \setminus \{0\}, y \in \mathbb{Z}^d}$. Таким образом, поведение системы частиц вне источника в обеих моделях совпадает.

Интенсивности переходов из нуля в модели II пропорциональны интенсивностям переходов из других состояний и имеют вид

$$a(0, x) = -\frac{a(x)}{a(0)}, \quad x \neq 0.$$

Это означает, что вероятности переходов из одного узла решетки в другой для блуждания однородны по пространству и времени, причем время пребывания в точке $x_0 = 0$ имеет экспоненциальное распределение со средним 1, тогда как время пребывания в любом другом состоянии $x_0 \neq 0$ имеет экспоненциальное распределение со средним $-\frac{1}{a(0)}$.

Развитие популяции начинается в момент времени $t = 0$ с одной частицы, расположенной в произвольном узле x (в статьях [10–13] [10, 11, 12, 13] рассмотрен случай старта процесса из нуля). Начальная частица, так же как и любая другая частица, оказавшаяся в нуле позднее, проводит в этой точке случайное время, имеющее экспоненциальное распределение с единичным средним, а затем либо переходит в точку $y \neq 0$ с вероятностью $(1 - \alpha)a(0, y)$, либо гибнет с вероятностью α , порождая при этом случайное количество потомков ξ в соответствии с производящей функцией¹⁾

$$\tilde{\mathcal{F}}(u) = E s^\xi = \sum_{n=0}^{\infty} f_n u^n,$$

а также предполагается, что выполнены следующие условия:

$$\mathbf{E}\xi < \infty, \quad \sigma^2 = \mathbf{D}\xi^2 \in (0, \infty).$$

Отметим, что в статьях [10–13] данная модель исследована при $\mathbf{E}\xi = 1$ на $\mathbb{Z}$ и $\mathbb{Z}^2$.

¹⁾Здесь и ниже символ $\tilde{}$ (check) означает, что рассматриваемая функция относится к модели II. Во избежание недоразумений отметим, что в [10, 12] функция $\tilde{\mathcal{F}}(u)$ обозначается через $f(u)$.

Сравнение моделей. При сравнении моделей I и II возникает ряд трудностей.

Во-первых, эти модели описываются в различных терминах. Так, модель I описывается в терминах инфинитезимальных эволюций частиц — то есть в терминах того, что происходит с частицами за некое малое время h . В модели II применяется «смешанное» описание эволюции частиц — в одних ситуациях используется инфинитезимальное описание, как и в модели I, в других — описание в терминах «времени жизни» отдельных частиц, и, наконец, в третьих, — в терминах «мгновенного», «вневременного» изменения состояния частиц — «смерти одних и рождения новых частиц».

Во-вторых, некоторые термины используются в этих моделях в разном смысле. Так, в модели I под состоянием системы частиц понимается количество частиц в точках решетки, а в модели II в понятие состояния неявно включается «время жизни» частицы. Это приводит к тому, что такое событие, как «дожитие частицы до конца своей жизни, следующая за этим смерть частицы и рождение потом одного потомка», во второй модели рассматривается как изменение состояния — вместо умершей частицы появилась частица нового поколения. В модели I аналогичная ситуация рассматривается как неизменное состояние, так как в результате подобного сценария ни общее число частиц, ни число частиц в отдельных точках не изменилось.

Наконец, существует общая для обеих моделей «двусмысленность», связанная с тем, что в точке ветвления x_0 «блуждание» и «ветвление» частицы могут привести к одному и тому же результату с точки зрения изменения как общего числа частиц, так и числа частиц в отдельных точках.

Если основной целью анализа моделей является исследование эволюции количества частиц в узлах решетки, то, вероятно, более продуктивно было бы проводить описание эволюции системы частиц в рассматриваемых моделях не в классических терминах «блуждания» и «ветвления», как это сделано в обеих моделях, а в терминах неких «обобщенных блужданий» и «обобщенных ветвлений», относя, например, к «обобщенному блужданию» все те сценарии эволюции частиц, *которые не приводят к изменению общего числа частиц на решетке*, и, относя к «обобщенному ветвлению» все те сценарии эволюции частиц, *в результате которых общее число частиц на решетке изменяется*. В предлагаемом ниже описании тогда к «обобщенному блужданию» следовало бы отнести случаи (а), (б) и (г), а к «обобщенному ветвлению» — случай (в).

В обеих моделях предполагается, что в момент времени $t = 0$ система состоит из одной частицы, находящейся в точке x . Рассмотрим всевозможные изменения системы на интервале $(0, h)$. Для сравнения моделей I и II опишем их характеристики в «сопоставимых терминах» — через вероятности следующих взаимоисключающих событий: (а) частица, находящаяся в точке $x \neq x_0$, за малое время h совершает «прыжок» в произвольную точку y — в этом случае общее число частиц на решетке не меняется; (б) частица, находящаяся в точке ветвления $x_0 = 0$, за время h совершает «прыжок» в произвольную точку y — в этом случае, как и в случае (а), общее число частиц на решетке также не меняется; (в) частица, находящаяся в точке ветвления $x_0 = 0$, за время h производит $n \neq 1$ потомков ($n \geq 0$) — в этом случае, в отличие от случаев (а) и (б), общее число частиц на

решетке изменяется; (г) число частиц, находящихся в точке ветвления $x_0 = 0$, за малое время h не изменится, то есть либо частица не совершила из точки $x_0 = 0$ прыжка в какую-либо точку решетки или совершила «псевдопрыжок» в эту же самую точку, либо процесса ветвления не случилось вовсе или частица дожила до своей «смерти», но в соответствии с законом ветвления произвела в точности одного потомка — в обеих ситуациях общее число частиц в источнике $x_0 = 0$, а также на всей решетке не изменилось; (д) произошло более одного изменения численности или положения частиц за время h с вероятностью $o(h)$.

Отметим, что событие (г) с точки зрения процессов блуждания и ветвления носит двойственный характер — его можно рассматривать и как «прыжок» из точки $x = 0$ в себя, и как некое «вырожденное ветвление», когда частица, умирая, рождает одного потомка. В дальнейшем из технических соображений удобнее рассматривать событие (г) как часть процесса блуждания, так как в этом случае общее число частиц на решетке не меняется.

Вернемся к рассмотренным выше случаям.

(а) В модели I вероятность $p(h, x, y)$ события, что частица, находящаяся в точке $x \neq x_0$, за малое время h совершает «прыжок» в произвольную точку y , определяется равенством

$$p(h, x, y) = a(y - x)h + o(h),$$

а в модели II вероятность $\check{p}(h, x, y)$ этого события определяется равенством

$$\check{p}(h, x, y) = a(y - x)h + o(h),$$

правая часть которого с точностью до $o(h)$ совпадает с выражением для $p(h, x, y)$. Другими словами, можно считать, что в обеих моделях «прыжок» из точки $x \neq x_0$ в произвольную точку y за малое время h совершается по одному и тому же закону.

(б) Частица, находящаяся в точке ветвления $x_0 = 0$, за время h совершает «прыжок» в произвольную точку $y \neq x_0$. В этом случае в модели I вероятность $p(h, 0, y)$ соответствующего события определяется равенством

$$p(h, 0, y) = a(y)h + o(h),$$

а в модели II вероятность $\check{p}(h, 0, y)$ того же события определяется равенством

$$\check{p}(h, 0, y) = (1 - e^{-h})(1 - \alpha)a(0, y) = -(1 - \alpha)\frac{a(y)}{a(0)}h + o(h).$$

В дальнейшем генератор случайного блуждания в модели II будем обозначать через

$$A_2 = A - \delta_0(\cdot)A + \frac{(\alpha - 1)}{a(0)}\delta_0(\cdot)A,$$

где $\delta_0(\cdot)$ представляет собой матрицу с ненулевым элементом.

Пример 1. Матрица переходных интенсивностей $A_2 = \|a(x, y)\|_{x, y \in \mathbb{Z}}$ на $\mathbb{Z}$ может быть представлена в следующем виде:

$$A_2 = \left\| \begin{array}{ccc} \cdots a(0) & a(1) & a(2) \cdots \\ \cdots - (1 - \alpha)\frac{a(1)}{a(0)} & -(1 - \alpha) & -(1 - \alpha)\frac{a(1)}{a(0)} \cdots \\ \cdots a(2) & a(1) & a(0) \cdots \end{array} \right\|.$$

В отличие от матрицы A для матрицы A_2 выполняется условие $\sum_{y \in \mathbb{Z}} a(x, y) = 0$, но при этом $\sum_{x \in \mathbb{Z}} a(x, y) \neq 0$.

(в) Частица, находящаяся в точке ветвления $x_0 = 0$, за время h производит $n \neq 1$ потомков. В модели I вероятность $p_*(h, n)$ этого события равна

$$p_*(h, n) = b_n h + o(h), \quad n \neq 1, \quad (3)$$

а в модели II вероятность $\check{p}_*(h, n)$ этого же события равна

$$\check{p}_*(h, n) = (1 - e^{-h})\alpha f_n = \alpha f_n h + o(h), \quad n \neq 1. \quad (4)$$

Отметим, что при $n = 0$ соответствующее событие трактуется как «смерть» частицы. В этом случае, как и в случае (б), число частиц в точке $x_0 = 0$ уменьшается на 1.

(г) В точке ветвления $x_0 = 0$ за малое время h число частиц не меняется. Это событие понимается в том смысле, что частица за малое время h либо не совершила прыжок в какую-либо точку решетки, отличную от x_0 , либо процесса ветвления в источнике x_0 вовсе не произошло или в процессе ветвления произошла «смерть» частицы с воспроизводством одного потомка нового поколения. Так как число частиц в источнике x_0 в такой ситуации не изменилось, то по классификации, приведенной выше, этот случай может быть отнесен к «обобщенному блужданию». В модели I вероятность соответствующего события равна

$$1 - \sum_{y \neq 0} a(y)h - \sum_{n \neq 1} b_n h + o(h),$$

что равносильно

$$1 + a(0)h + b_1 h + o(h), \quad (5)$$

а в модели II вероятность этого события равна

$$(1 - \alpha)e^{-h} + \alpha e^{-h} + \alpha(1 - e^{-h})f_1$$

или, что то же,

$$1 + (\alpha - 1)h + \alpha(f_1 - 1)h + o(h) \iff 1 - (1 - \alpha f_1)h + o(h). \quad (6)$$

Теперь, когда модели определены в сопоставимых терминах, их сравнение не представляет трудностей.

Сравним вначале вероятности отсутствия изменений числа частиц в источнике ветвления в этих моделях. Основное отличие моделей состоит в том, что время, проведенное до первого изменения числа частиц в источнике ветвления в $x_0 = 0$, в модели II имеет экспоненциальное распределение со средним $\frac{1}{1 - \alpha f_1}$, а в модели I со средним $-\frac{1}{a(0) + b_1}$.

Заметим, что при выполнении условия

$$\alpha - 1 = a(0) \quad (7)$$

совпадут генераторы случайного блуждания (см. пример 1, но при этом механизмы ветвления в источнике могут различаться).

Теперь выясним, при каких же условиях будут совпадать механизмы воспроизводства потомства в источнике с точностью до $o(h)$. Из соотношений (3) и (4) вытекает, что при

$$b_1 = \alpha(f_1 - 1), \quad b_n = \alpha f_n, \quad n \neq 1, \quad (8)$$

механизмы ветвления в источнике в обеих моделях совпадают с точностью до $o(h)$. Если дополнительно предположить, что выражения (5) и (6) совпадают с точностью до $o(h)$, то $a(0) + b_1 = (\alpha - 1) + \alpha(f_1 - 1)$. Отсюда в силу (8) получаем (7), и тогда $A = A_2$.

Удобным инструментом описания процесса ветвления с непрерывным временем является инфинитезимальная производящая функция. Поэтому продолжим сравнение моделей, представив и инфинитезимальные функции для них. Для модели I эта функция представляется как

$$f(u) = \sum_{n=0}^{\infty} b_n u^n,$$

а для модели II, согласно (4), аналогичная функция будет иметь следующий вид:

$$\check{f}(u) = \alpha \left(\sum_{n \neq 1}^{\infty} f_n u^n + (f_1 - 1)u \right).$$

Если в модели II дополнительно ввести условие, аналогичное условию (2) в модели I: $\check{\beta}_r = \check{f}^{(r)}(1) < \infty$ при всех r , то есть число потомков частицы имеет конечные моменты всех порядков, то в предположениях (8) ветвящийся процесс в источнике $x_0 = 0$ в модели II может быть задан с помощью инфинитезимальной производящей функции (1) как

$$\check{f}(u) = f(u).$$

Будем обозначать далее $\check{\beta} = \check{\beta}_1 = \check{f}'(1)$.

Полученные представления для вероятностей прыжков из одной точки в другую и вероятностей воспроизведения определенного числа потомков, равно как и выражения для соответствующих инфинитезимальных производящих функций, показывают, что модель II может быть описана в тех же терминах, что и модель I, как набор неких «инфинитезимальных» характеристик. Это дает возможность применить для исследования модели II те же методы описания процесса в виде бесконечномерных дифференциально-разностных уравнений, которые были развиты в работах [3–9] для исследования модели I. Соответствующие уравнения для модели II для дальнейшего сравнения свойств моделей I и II будут выписаны и исследованы в следующих разделах. Отметим в заключение настоящего раздела, что, конечно же, интерес представляет и «обратная» задача — переформулировать модель I в терминах, использовавшихся при описании модели II. Этому будут посвящены дальнейшие публикации.

3. Производящие функции

Остановимся вначале на известных фактах для модели I. Пусть $p(t, x, y)$ — переходная вероятность случайного блуждания. Из описания модели I при малых h вытекают соотношения

$$p(h, x, y) = a(x, y)h + o(h) \quad \text{при } y \neq x, \quad (9)$$

$$p(h, x, x) = 1 + a(x, x)h + o(h). \quad (10)$$

Как показано, например, в [16], из (9), (10) следует, что переходные вероятности $p(t, x, y)$ удовлетворяет системе дифференциально-разностных уравнений (*обратные уравнения Колмогорова*)

$$\frac{\partial p(t, x, y)}{\partial t} = Ap(t, x, y), \quad p(0, x, y) = \delta_y(x),$$

где $\delta_y(\cdot)$ — дискретная δ -функция Кронекера на $\mathbb{Z}^d$:

$$\delta_y(x) = \begin{cases} 1 & \text{при } x = y, \\ 0 & \text{при } x \neq y, \end{cases}$$

а выражение $Ap(t, x, y)$ определяется как

$$Ap(t, x, y) = \sum_{x'} a(x, x')p(t, x', y).$$

Как показано в [17, 9], в модели I производящие функции при каждом $0 \leq z \leq \infty$ удовлетворяют неравенствам $0 \leq F(z, t, x), F(z, t, x, y) \leq 1$ и дифференциальным уравнениям

$$\partial_t F(z, t, x) = (AF(z, t, \cdot))(x) + \delta_0(x)f(F(z, t, x)), \quad (11)$$

$$\partial_t F(z, t, x, y) = (AF(z, t, \cdot, y))(x) + \delta_0(x)f(F(z, t, x, y)) \quad (12)$$

с начальными условиями $F(z, 0, x) = e^{-z}$ и $F(z, 0, x, y) = e^{-z\delta_y(x)}$ соответственно²⁾.

Теперь перейдем к выводу аналогичных (11), (12) дифференциальных уравнений для производящих функций в модели II. Введем вероятность $P_n(t, x, y)$ того, что в момент времени t в точке $y \in \mathbb{Z}^d$ будет n частиц при условии старта процесса в начальный момент времени $t = 0$ из произвольной точки $x \in \mathbb{Z}^d$:

$$P_n(t, x, y) = P\{\mu_t(y) = n \mid \mu_t(z) = \delta_x(z)\} \quad \text{при } z \in \mathbb{Z}^d. \quad (13)$$

Определим производящие функции для случайных величин $\mu_t(y)$ и μ_t при $z \geq 0$ в модели II равенствами

$$\check{F}(z, t, x, y) = E_x e^{-z\mu_t(y)}, \quad \check{F}(z, t, x) = E_x e^{-z\mu_t}, \quad (14)$$

где E_x обозначает математическое ожидание при условии $\mu_0(\cdot) = \delta_x(\cdot)$.

²⁾Выражение $AF(z, t, \cdot)$ следует понимать как единый символ, обозначающий зависящую от параметра z и времени t функцию переменной $x \in \mathbb{Z}^d$, определяемую равенством $(AF(z, t, \cdot))(x) = \sum_{x' \in \mathbb{Z}^d} a(x, x')F(z, t, x')$. Аналогично определяется и выражение $AF(z, t, \cdot, y)$ в формуле (12).

Теорема 1. При каждом $0 \leq z \leq \infty$ функции $\check{F}(z, t, x)$ и $\check{F}(z, t, x, y)$ непрерывно дифференцируемы по переменной t равномерно по $x, y \in \mathbf{Z}^d$, удовлетворяют неравенствам $0 \leq \check{F}(z, t, x), \check{F}(z, t, x, y) \leq 1$ и дифференциальным уравнениям

$$\partial_t \check{F}(z, t, x) = (A_2 \check{F}(z, t, \cdot))(x) + \delta_0(x) \check{f}(\check{F}(z, t, x)), \quad (15)$$

$$\partial_t \check{F}(z, t, x, y) = (A_2 \check{F}(z, t, \cdot, y))(x) + \delta_0(x) \check{f}(\check{F}(z, t, x, y)) \quad (16)$$

с начальными условиями $\check{F}(z, 0, x) = e^{-z}$ и $\check{F}(z, 0, x, y) = e^{-z\delta_y(x)}$ соответственно.

Доказательство. Неравенства $0 \leq \check{F}(z, t, x), \check{F}(z, t, x, y) \leq 1$ следуют из неотрицательности случайных величин μ_t и $\mu_t(y)$ и определения производящих функций.

Начнем с вывода уравнения (16). Продолжим временной интервал $(0, t + h)$ в направлении прошлого, и рассмотрим всевозможные изменения на $(0, h)$. Таким образом, получается система дифференциальных уравнений, в которой остается фиксированным n . Действительно, в начальный момент времени в произвольной точке x находится одна частица, для которой существуют четыре взаимоисключающие возможности (а)–(д), перечисленные в разделе 2:

(а)–(б) если $x \neq 0$, то перейти в произвольную точку решетки $x' \neq x$ с вероятностью $\check{p}(h, x, x') = a(x' - x)h + o(h)$; если же $x = 0$, то перейти в произвольную точку решетки $x' \neq x$ с вероятностью $\check{p}(h, 0, x') = \frac{\alpha-1}{a(0)}a(x')h + o(h)$. Эта возможность осуществляется с вероятностью $(1 - \delta_0(x))(a(x' - x)h + o(h)) + \delta_0(x)((\alpha - 1)\frac{a(x')}{a(0)}h + o(h))$;

(в) произвести n' потомков ($n' \neq 1, n' \geq 0$), если частица находится в источнике ветвления, с вероятностью $\check{p}_*(h, n') = \alpha f_{n'}h + o(h)$;

(г) остаться в точке x , а также если $x = 0$, то остаться в источнике, или не подвергнуться ветвлению, или в результате ветвления произвести только одного потомка с вероятностью $(1 - \delta_0(x))(1 + a(0)h + o(h)) + \delta_0(x)(1 - (1 - \alpha)h - \alpha(1 - f_1)h + o(h))$;

(д) иметь более одного изменения состояния за время h с вероятностью $o(h)$.

В рамках модели II это означает, что

$$\begin{aligned} \mathbf{P}\{\mu_{h+t}(y) = n \mid \mu_0(z) = \delta_x(z)\} &= \mathbf{P}\{\mu_t(y) = n \mid \mu_0(z) = \delta_x(z)\} \times \\ &\times ((1 - \delta_0(x))(1 + a(0)h + o(h)) + \delta_0(x)(1 - (1 - \alpha)h - \alpha(1 - f_1)h + o(h))) + \\ &+ (1 - \delta_0(x)) \sum_{x' \neq x} (a(x' - x)h + o(h)) \mathbf{P}\{\mu_t(y) = n \mid \mu_0(z) = \delta_{x'}(z)\} + \\ &+ \delta_0(x)(\alpha - 1) \sum_{x' \neq 0} \left(\frac{a(x')}{a(0)}h + o(h) \right) \mathbf{P}\{\mu_t(y) = n, \mid \mu_0(z) = \delta_{x'}(z)\} + \\ &+ \delta_0(x)\alpha \sum_{n' \neq 1} (f_{n'}h + o(h)) \mathbf{P}\{\mu_t(y) = n \mid \mu_0(z) = n'\delta_x(z)\}. \end{aligned}$$

Отсюда стандартными методами [16], используя обозначение (13) и определение ветвящегося процесса [14], при $h \rightarrow 0$ получаем обратное уравнение Колмогорова

$$\partial_t \mathbf{P}_n(t, x, y) =$$

$$\begin{aligned}
 &= (1 - \delta_0(x)) \sum_{x'} a(x' - x) P_n(t, x', y) + \delta_0(x) \frac{(\alpha - 1)}{a(0)} \sum_{x'} a(x' - x) P_n(t, x', y) + \\
 &+ \delta_0(x) \alpha \sum_{n'} f_{n'} \left(\sum_{j_1 + \dots + j_{n'} = n'} P_{j_1}(t, x, y) \cdot \dots \cdot P_{j_{n'}}(t, x, y) \right) - \delta_0(x) \alpha P_n(t, x, y).
 \end{aligned}$$

Пользуясь введенными выше обозначениями, получаем

$$\begin{aligned}
 \partial_t P_n(t, x, y) &= \left(1 - \delta_0(x) + \delta_0(x) \frac{(\alpha - 1)}{a(0)} \right) (A P_n(t, \cdot, y))(x) + \\
 &+ \delta_0(x) \alpha \sum_{n'} f_{n'} \left(\sum_{j_1 + \dots + j_{n'} = n'} P_{j_1}(t, x, y) \cdot \dots \cdot P_{j_{n'}}(t, x, y) \right) - \delta_0(x) \alpha P_n(t, x, y),
 \end{aligned}$$

или, что равносильно,

$$\begin{aligned}
 \partial_t P_n(t, x, y) &= (A_2 P_n(t, \cdot, y))(x) + \\
 &+ \delta_0(x) \alpha \sum_{n'} f_{n'} \left(\sum_{j_1 + \dots + j_{n'} = n'} P_{j_1}(t, x, y) \cdot \dots \cdot P_{j_{n'}}(t, x, y) \right) - \delta_0(x) \alpha P_n(t, x, y)
 \end{aligned}$$

с начальным условием $P_n(0, x, y) = \delta_1(n) \delta_x(y)$. В силу определения (14), домножая левую и правую части на e^{-zn} и суммируя по n , получаем обратное уравнение для производящей функции $\check{F}(z, t, x, y)$

$$\partial_t \check{F}(z, t, x, y) = (A_2 \check{F}(z, t, \cdot, y))(x) + \alpha \delta_0(x) \sum_{n'} f_{n'} (\check{F})^{n'}(z, t, x, y) - \alpha \delta_0(x) \check{F}(z, t, x, y)$$

с начальным условием $\check{F}(z, 0, x, y) = e^{-z\delta_x(y)}$. В силу введенных выше определений последнее уравнение может быть представлено в форме

$$\partial_t \check{F}(z, t, x, y) = (A_2 \check{F}(z, t, \cdot, y))(x) + \delta_0(x) \check{f}(\check{F}(z, t, x, y))$$

с начальным условием $\check{F}(z, 0, x, y) = e^{-z\delta_x(y)}$.

Теперь из оценки $0 \leq \check{F}(z, t, x, y) \leq 1$, а также из (16) и ограниченности операторов A_2 и δ_0 в пространстве $l^\infty(\mathbf{Z}^d)$ вытекает равномерная по z, x ограниченность и непрерывность производной по t функции $\check{F}(z, t, x, y)$. А отсюда и из уже доказанного представления (12) следует равномерная по z, x и y непрерывная дифференцируемость по t функции $\check{F}(z, t, x, y)$.

Аналогично доказывается утверждение теоремы для (15).

Сравним дифференциальные уравнения для производящих функций (11), (12) и (15), (16). Сходство заключается в том, что операторы A и A_2 являются линейными ограниченными операторами, действующими из $l^q(\mathbf{Z}^d)$ в $l^q(\mathbf{Z}^d)$. Однако в силу симметрии случайного блуждания оператор A в модели I является самосопряженным, оператор A_2 в модели II этим свойством не обладает.

4. Дифференциальные уравнения для моментов

Напомним, что одним из объектов исследования в модели I являются моменты случайных величин μ_t и $\mu_t(y)$.

Сформулируем ряд утверждений, справедливых и для модели II. Согласно известному свойству преобразования Лапласа [15], производящие функции $\check{F}(z, t, x) = \mathbb{E}_x e^{-z\mu_t}$ и $\check{F}(z, t, x, y) = \mathbb{E}_x e^{-z\mu_t(y)}$ бесконечно дифференцируемы по z при $z > 0$ и их производные имеют вид

$$\begin{aligned}\partial_z^n \check{F}(z, t, x) &= (-1)^n \mathbb{E}_x \mu_t^n e^{-z\mu_t}, \\ \partial_z^n \check{F}(z, t, x, y) &= (-1)^n \mathbb{E}_x \mu_t^n(y) e^{-z\mu_t(y)}.\end{aligned}$$

Поэтому моменты m_n могут быть представлены в виде:

$$\begin{aligned}m_n(t, x) &= (-1)^n \lim_{z \rightarrow 0+} \partial_z^n \check{F}(z, t, x), \\ m_n(t, x, y) &= (-1)^n \lim_{z \rightarrow 0+} \partial_z^n \check{F}(z, t, x, y).\end{aligned}$$

Определим операторы $H = A + \beta\delta_0$ и $H_2 = A_2 + \check{\beta}\delta_0 = A - \left(1 - \frac{\alpha-1}{a(0)}\right)\delta_0 A + \check{\beta}\delta_0$. Как показано в [9], H , как и оператор A , действует в каждом из пространств $l^q(\mathbb{Z}^d)$, $1 \leq q \leq \infty$. Поэтому аналогичное утверждение справедливо и для оператора H_2 . Таким образом, для модели II справедлив аналог теоремы 1.3.1 из [9]:

Теорема 2. При каждом целом $k \geq 1$ моменты $m_k(t, x, y)$ и $m_k(t, x)$ в модели II удовлетворяют дифференциальным уравнениям

$$\begin{aligned}\partial_t m_1 &= H_2 m_1, \\ \partial_t m_k &= H_2 m_k + \delta_0(x) g_k(m_1, m_2, \dots, m_{k-1}), \quad k \geq 2,\end{aligned}$$

с начальными условиями $m_n(0, \cdot, y) = \delta_y(\cdot)$ и $m_n(0, \cdot) \equiv 1$ соответственно. Здесь $H_2 m_k$ есть сокращенная запись для обозначения $(H_2 m_k(t, \cdot, y))(x)$ или $(H_2 m_k(t, \cdot))(x)$ соответственно и

$$g_k(m_1, m_2, \dots, m_{k-1}) = \sum_{r=2}^k \frac{\check{\beta}_r}{r!} \sum_{\substack{i_1, \dots, i_r > 0 \\ i_1 + \dots + i_r = k}} \frac{k!}{i_1! \dots i_r!} m_{i_1} \dots m_{i_r}.$$

В [6, 4, 9] показано, что если в спектре оператора H существует старшее изолированное простое положительное собственное значение λ , то при $t \rightarrow \infty$ происходит экспоненциальный рост как локальных численностей частиц, так и их полного числа:

$$\lim_{t \rightarrow \infty} \mu_t(y) e^{-\lambda t} = \xi \psi(y), \quad \lim_{t \rightarrow \infty} \mu_t e^{-\lambda t} = \xi, \quad (17)$$

где $\psi(y)$ — некоторая функция, а ξ — невырожденная случайная величина. Такой случай назван в [6, 4, 9] надкритическим. Соотношения (17) справедливы также в смысле сходимости по распределению. В [6, 5] выделены условия, при которых распределение ξ однозначно определяется моментами.

Дальнейшая часть статьи посвящена нахождению условий, при которых в модели II у оператора H_2 будет существовать старшее положительное собственное значение λ .

5. Условия существования положительных собственных значений оператора H_2

Как показано в [9], переходные вероятности случайного блуждания в модели I имеют степенную асимптотику по времени³⁾:

$$p(t, x, y) \sim \gamma_d \cdot t^{-d/2} \quad \text{при } t \rightarrow \infty, \quad (18)$$

где $\gamma_d > 0$ — некоторая константа.

Рассмотрим функцию Грина $G_\lambda(x, y) = \int_0^\infty e^{-\lambda t} p(t, x, y) dt$. Из соотношения (18) вытекает, что $G_\lambda(0, 0)|_{\lambda=0} < \infty$ при $d \geq 3$. Следовательно, случайное блуждание транзиентно, если $d \geq 3$, и возвратно, если $d = 1, 2$. Положим

$$\beta_c = \frac{1}{G_0(0, 0)}.$$

В [3–6, 9] установлены условия критичности для модели I. Надкритическими ветвящимися случайными блужданиями будем называть такие, при которых наблюдается экспоненциальный рост процесса. В модели I для того, чтобы рассматриваемый процесс стал надкритическим, необходимо и достаточно, чтобы выполнялось условие $\beta > \beta_c$ [9]. Как будет установлено в этом разделе, таким достаточным условием для модели II служит условие

$$\check{\beta} > \frac{(\alpha - 1)}{a(0)} \beta_c. \quad (19)$$

Перейдем к более удобным обозначениям:

$$G_\lambda = G_\lambda(0, 0), \quad s = 1 - \frac{a(0)}{\alpha - 1}, \quad p(t) = p(t, 0, 0).$$

Тогда с учетом предположений $a(0) < 0$ и $0 < \alpha < 1$ условие (19) примет вид

$$\check{\beta} > \frac{\beta_c}{(1 - s)}, \quad s < 1.$$

Лемма. Число λ является положительным собственным значением оператора H_2 , а $g \in l^2(\mathbf{Z}^d)$ — отвечающим ему собственным вектором тогда и только тогда, когда выполнены следующие условия:

$$g(0) \neq 0, \quad \check{\beta} \neq 0, \quad 0 < \alpha < 1, \quad s < 1, \quad (20)$$

$$\int_{[-\pi, \pi]^d} |\lambda - \phi(\theta)|^{-2} d\theta < \infty; \quad (21)$$

$$((1 - s)\check{\beta} + s\lambda) G_\lambda = 1. \quad (22)$$

³⁾ Отметим, что первый член асимптотического разложения переходной вероятности $p(t, x, y)$ не зависит от x и y .

При этом собственное значение λ является изолированной точкой спектра оператора H_2 , а собственный вектор g , отвечающий собственному значению λ , определяется равенством

$$g(x) = \left(\check{\beta}g(0) + \frac{s}{1-s} \sum_{x'} a(x')g(x') \right) G_\lambda(x, 0), \quad x \in \mathbf{Z}^d, \quad (23)$$

где

$$G_\lambda(x, 0) = \frac{1}{(2\pi)^d} \int_{[-\pi, \pi]^d} \frac{e^{-i(\theta, x)}}{\lambda - \phi(\theta)} d\theta, \quad x \in \mathbf{Z}^d,$$

и поэтому каждое собственное значение оператора H_2 простое.

Доказательство. По определению, для того, чтобы число λ было собственным значением оператора H_2 , необходимо и достаточно, чтобы нашелся ненулевой элемент $g \in l^2(\mathbf{Z}^d)$, удовлетворяющий уравнению

$$(H_2 - \lambda I)g = (A_2 + \check{\beta}\delta_0 - \lambda I)g = 0, \quad (24)$$

которое в силу равенства $\delta_0 g = g(0)\delta_0$ можно представить в виде

$$(A_2 g)(x) + \check{\beta}g(0)\delta_0(x) = \lambda g(x), \quad x \in \mathbf{Z}^d. \quad (25)$$

В силу определения оператора A_2 последнее уравнение может быть переписано следующим образом:

$$(Ag)(x) + \frac{s}{1-s} \left(\sum_{x'} a(x')g(x') \right) \delta_0(x) + \check{\beta}g(0)\delta_0(x) = \lambda g(x), \quad x \in \mathbf{Z}^d. \quad (26)$$

Как уже отмечалось выше, ограниченность оператора A вытекает из леммы Шура [9, 18], а его самосопряженность — из симметричности матрицы $\|a(x, y)\|_{x, y \in \mathbf{Z}^d}$.

В силу ограниченности оператора A в $l^2(\mathbf{Z}^d)$ при каждом $g \in l^2(\mathbf{Z}^d)$ элемент $\psi = Ag$ принадлежит $l^2(\mathbf{Z}^d)$ и для него определено преобразование Фурье $\tilde{\psi}(\theta)$, которое в силу симметричности и однородности матрицы $a(x, y)$ имеет вид

$$\begin{aligned} \tilde{\psi}(\theta) &= \sum_{x \in \mathbf{Z}^d} \psi(x) e^{i(x, \theta)} = \sum_{x \in \mathbf{Z}^d} \left(\sum_{x' \in \mathbf{Z}^d} a(x, x')g(x') \right) e^{i(x, \theta)} = \\ &= \sum_{x \in \mathbf{Z}^d} \left(\sum_{x' \in \mathbf{Z}^d} a(x - x', 0)g(x') \right) e^{i(x - x' + x', \theta)} = \\ &= \left(\sum_{x \in \mathbf{Z}^d} a(x, 0)e^{i(x, \theta)} \right) \left(\sum_{x' \in \mathbf{Z}^d} g(x')e^{i(x', \theta)} \right) = \phi(\theta)\tilde{g}(\theta), \end{aligned}$$

где $\phi(\theta) = \sum_x a(x, 0) \cos(x, \theta)$ и $\tilde{g}(\theta)$ — преобразование Фурье функции $g(x)$.

Применим к уравнению (26) преобразование Фурье:

$$\phi(\theta)\tilde{g}(\theta) + \frac{s}{1-s} \sum_{x'} a(x')g(x') + \check{\beta}g(0) = \lambda\tilde{g}(\theta), \quad x \in [-\pi, \pi]^d. \quad (27)$$

Из разрешимости уравнения (27) в $l^2([-\pi, \pi]^d)$ относительно функции $\tilde{g}(\theta)$ и из обратимости преобразования Фурье в l^2 получаем, что уравнение (25), а с ним и уравнение (24), разрешимы, если разрешимо уравнение (27).

Итак, уравнение (25) имеет ненулевое решение в $l^2(\mathbf{Z}^d)$ тогда и только тогда, когда уравнение (27) имеет ненулевое решение в $l^2([-\pi, \pi]^d)$ или, что то же самое, тогда и только тогда, когда выполнены следующие условия:

$$\tilde{g}(\theta) \neq 0, \quad \tilde{g}(\theta) \in l^2([-\pi, \pi]^d), \quad \tilde{g}(\theta) = \frac{\check{\beta}g(0) + \frac{s}{1-s} \sum_{x'} a(x')g(x')}{\lambda - \phi(\theta)}. \quad (28)$$

Здесь первое из условий в совокупности с третьим равносильно (20). Второе условие (28) в силу (20) равносильно (21). Наконец, для получения условий разрешимости третьего уравнения в (28) применим к нему обратное преобразование Фурье. Тогда получаем, что функция $g(x)$ должна удовлетворять уравнению (23), откуда $g(0) = \check{\beta}g(0)G_\lambda + \frac{s}{1-s} \sum_{x'} a(x')g(x')$, что в силу (20) равносильно условию

$$g(x) = \frac{1}{(2\pi)^d} \int_{[-\pi, \pi]^d} \frac{(\check{\beta}g(0) + \frac{s}{1-s} \sum_{x'} a(x')g(x')) e^{-i(\theta, x)}}{\lambda - \phi(\theta)} d\theta, \quad x \in \mathbf{Z}^d. \quad (29)$$

Это уравнение может быть переписано в виде

$$\begin{aligned} \check{\beta}g(0) + \frac{s}{1-s} \sum_x a(x)g(x) &= \\ &= \check{\beta} \left(\check{\beta}g(0) + \frac{s}{1-s} \sum_{x'} a(x')g(x') \right) \frac{1}{(2\pi)^d} \int_{[-\pi, \pi]^d} \frac{1}{\lambda - \phi(\theta)} d\theta + \\ &+ \left(\check{\beta}g(0) + \frac{s}{1-s} \sum_{x'} a(x')g(x') \right) \frac{1}{(2\pi)^d} \int_{[-\pi, \pi]^d} \frac{\frac{s}{1-s} \sum_x a(x)e^{-i(\theta, x)}}{\lambda - \phi(\theta)} d\theta = \\ &= \left(\check{\beta}g(0) + \frac{s}{1-s} \sum_{x'} a(x')g(x') \right) \frac{1}{(2\pi)^d} \int_{[-\pi, \pi]^d} \frac{\check{\beta} + \frac{s}{1-s} \phi(\theta)}{\lambda - \phi(\theta)} d\theta. \end{aligned}$$

Если константа $\check{\beta}g(0) + \frac{s}{1-s} \sum_{x'} a(x')g(x')$ равна нулю, то в силу (29) $g(x) = 0$ при всех x , что противоречит предположению леммы $g(0) \neq 0$. Разделим левую и правую часть последнего уравнения на $\check{\beta}g(0) + \frac{s}{1-s} \sum_{x'} a(x')g(x')$ и получим

$$\frac{1}{(2\pi)^d} \int_{[-\pi, \pi]^d} \frac{\check{\beta} + \frac{s}{1-s} \phi(\theta)}{\lambda - \phi(\theta)} d\theta = 1. \quad (30)$$

Теперь представим левую часть этого уравнения как

$$\frac{1}{(2\pi)^d} \int_{[-\pi, \pi]^d} \frac{\check{\beta} - \frac{s}{1-s}(\lambda - \phi(\theta) - \lambda)}{\lambda - \phi(\theta)} d\theta = \left(\check{\beta} + \frac{s}{1-s} \lambda \right) G_\lambda - \frac{s}{1-s}.$$

Таким образом, уравнение (30) принимает вид

$$(1-s) \left(\check{\beta} + \frac{s}{1-s} \lambda \right) G_\lambda = 1,$$

что равносильно (22). Лемма доказана. $\square$

Как свидетельствуют приводимые ниже теоремы, анализ условий существования собственного значения оператора H_2 оказывается более сложным, чем аналогичный анализ для оператора H , проведенный в [9].

Теорема 3. *Если*

$$\check{\beta} > \frac{\beta_c}{(1-s)}, \quad (31)$$

то оператор H_2 обладает положительным собственным значением λ .

Доказательство. Представим уравнение (22) в виде

$$\frac{1}{(1-s)\check{\beta} + s\lambda} = G_\lambda. \quad (32)$$

По лемме задача о существовании собственного значения $\lambda > 0$ оператора H_2 сводится к нахождению положительного решения (22). Из определения функции G_λ с учетом (18) вытекают следующие свойства: $G_\lambda > 0$ и строго монотонно убывает по λ , причем $\lim_{\lambda \rightarrow \infty} G_\lambda = 0$; если же $\lambda \rightarrow 0$, то функция G_λ неограниченно возрастает при $d = 1, 2$ и $\lim_{\lambda \rightarrow 0} G_\lambda = \frac{1}{\beta_c}$ при $d \geq 3$.

По условию теоремы $\check{\beta} > (1-s)^{-1}\beta_c$. Откуда следует, что при $s \leq 0$ левая часть уравнения (32) на интервале $\lambda \in (0, \frac{s-1}{s}\check{\beta})$ положительна и строго монотонно возрастает по λ ; а при всех $\lambda > \frac{s-1}{s}\check{\beta}$ — отрицательна. Таким образом, при $s \leq 0$ существует и единственно решение конечной кратности уравнения (32) на интервале $(0, \frac{s-1}{s}\check{\beta})$.

Теперь рассмотрим оставшийся случай $0 < s < 1$. В этом случае левая часть уравнения (32) положительна и строго убывает при всех $\lambda > 0$. Отметим, что по предположению теоремы при $d \geq 3$ левая часть уравнения (32) при $\lambda = 0$ принимает значение $\frac{1}{(1-s)\check{\beta}} < \frac{1}{\beta_c}$. Откуда по непрерывности сразу вытекает, что в сколь угодно малой окрестности нуля выполняется неравенство

$$G_\lambda > \frac{1}{(1-s)\check{\beta} + s\lambda}.$$

Случай $d = 1, 2$ в этом плане несложен, так как G_λ неограниченно возрастает, а при этом левая часть уравнения (32) ограничена в окрестности нуля. Для завершения доказательства достаточно показать, что существуют такие значения $\lambda > 0$, при которых справедливо неравенство

$$G_\lambda < \frac{1}{(1-s)\check{\beta} + s\lambda}. \quad (33)$$

Для этого представим функцию G_λ в виде

$$G_\lambda = \int_0^\infty e^{-\lambda t} p(t) dt = \frac{1}{\lambda} + \frac{1}{\lambda} \int_0^\infty e^{-\lambda t} p'(t) dt.$$

Как известно [10, 9], вероятность $p(t)$ не возрастает, а ее производная $p'(t) < 0$ и не убывает по t . Таким образом, справедливо утверждение

$$G_\lambda = \frac{1}{\lambda} - r(\lambda), \quad r(\lambda) = \frac{1}{\lambda} \int_0^\infty e^{-\lambda t} p'(t) dt < 0, \quad (34)$$

из которого сразу следует, что неравенство (33) уже будет выполнено при таких значениях λ , для которых справедливо соотношение

$$\frac{1}{\lambda} < \frac{1}{(1-s)\check{\beta} + s\lambda}.$$

Очевидно, неравенство выполняется при любых $\lambda > \check{\beta}$. Таким образом, положительное решение уравнения (22) существует и принадлежит интервалу $(0, \check{\beta})$. Теорема доказана.

Справедливо ли утверждение теоремы в обратную сторону в общем случае, пока не ясно.

Теорема 4. Пусть $s \leq 0$. Оператор H_2 обладает единственным положительным собственным значением λ тогда и только тогда, когда выполняется условие (31).

Доказательство. Предположение $\frac{a(0)}{\alpha-1} = 1 - s \geq 1$ может быть представлено как $s \leq 0$. Доказательство достаточности при $s \leq 0$ полностью завершено в теореме 3. Теперь покажем, что (31) является также необходимым условием, то есть если выполняется неравенство

$$\check{\beta} \leq \frac{\beta_c}{(1-s)},$$

то не существует положительных решений уравнения (32). Предположим, что такие решения существуют. Тогда в силу $1 - s > 0$ получаем

$$(G_\lambda)^{-1} = (1-s)\check{\beta} + s\lambda \leq \beta_c + s\lambda,$$

что равносильно

$$(G_\lambda)^{-1} - \beta_c \leq s\lambda. \quad (35)$$

Из свойств функции G_λ следует, что левая часть неравенства положительна при любых $\lambda > 0$, а правая – отрицательна, и поэтому неравенство не выполняется. Приходим к противоречию с предположением. Теорема доказана. $\square$

Теорема 5. Пусть $d = 1, 2$. Оператор H_2 обладает положительным собственным значением λ тогда и только тогда, когда выполняется условие (31).

Доказательство. Случай $s \leq 0$ разобран в теореме 4. Покажем, что при $0 < s < 1$ неравенство (35) также не выполняется. В этом случае

$$(G_\lambda)^{-1} - \beta_c \leq c\lambda, \quad 0 < c < 1.$$

Это неравенство в размерностях $d = 1, 2$ в силу соотношения (34) и равенства $\beta_c = 0$ принимает вид $\frac{1}{G_\lambda} = \frac{\lambda}{1-r(\lambda)}c\lambda$, где $0 < c < 1$, $0 < 1 - r(\lambda) < 1$. Отсюда сразу следует, что неравенство не выполняется (35) ни при каких $\lambda > 0$. Доказательство завершено. $\square$

Работа выполнена при поддержке гранта РФФИ № 07-01-00362-а и программы «Междисциплинарные научные проекты МГУ имени М. В. Ломоносова».

Список литературы

- [1] *Gartner J., Molchanov S.* Parabolic problems for the Anderson model. I: Intermittency and related topics // *Commun. Math. Phys.* 1990. **132**. P. 613–655.
- [2] *Molchanov S.* Lectures on random media // *Lect. Notes Math.* 1994. **1581**. P. 242–411.
- [3] Яровая Е. Б. Применение спектральных методов в изучении ветвящихся процессов с диффузией в некомпактном фазовом пространстве // *Теор. матем. физ.* 1991. **88**, № 1. С. 25–30.
- [4] Богачев Л. В., Яровая Е. Б. Моментный анализ ветвящегося случайного блуждания на решетке с одним источником // *Докл. РАН.* 1998. **363**, № 4. С. 439–442.
- [5] Богачев Л. В., Яровая Е. Б. Предельная теорема для надкритического ветвящегося случайного блуждания на $\mathbb{Z}^d$ с одним источником // *Успехи матем. наук.* 1998. **53**, № 5. С. 229–230.
- [6] *Albeverio S., Bogachev L. V., Yarovaia E. B.* Asymptotics of branching symmetric random walk on the lattice with a single source // *C. r. Acad. sci. Paris. Sér. I.* 1998. **326**, № 9. P. 975–980.
- [7] *Yarovaia E. B.* About limit theorems for branching symmetric random walk on $\mathbb{Z}^d$ // *Kolmogorov and contemporary mathematics (Moscow, 16–21 of June, 2003)*, М.: МГУ, 2003. P. 592–593.
- [8] Яровая Е. Б. Предельная теорема для критического ветвящегося случайного блуждания на $\mathbb{Z}^d$ с одним источником // *Успехи матем. наук.* 2005. **60**, № 1. С. 175–176.
- [9] Яровая Е. Б. Ветвящиеся случайные блуждания в неоднородной среде. М.: Изд-во ЦПИ при механико-математическом факультете МГУ, 2007.
- [10] *Vatutin V. A., Topchii V. A., Yarovaia E. B.* Catalytic branching random walk and queueing systems with random number of independent servers // *Teoriya Imovirnostej ta Matematichna Statistika.* 2003. **69**. P. 158–172.
- [11] *Topchii V. A., Vatutin V. A.* Individuals at the origin in the critical catalytic branching random walk // *Discrete Math. Theor. Comput. Sci. (electronic).* 2003. **6**. P. 325–332.
- [12] *Ватутин В. А., Топчий В. А.* Предельная теорема для критических каталитических ветвящихся случайных блужданий // *Теория вероятностей и ее применения.* 2004. **49**, № 3. С. 461–484.
- [13] *Захарьева Е. В.* О вероятностях выживания частиц на $\mathbb{Z}^2$ в одной из моделей критического ветвящегося случайного блуждания // *Тр. VI Колмогоровских чтений (Ярославль, 2008)*. Ярославль: ЯГПУ, 2008. С. 218–229.
- [14] Севастьянов Б. А. Ветвящиеся процессы. М.: Наука, 1971.
- [15] *Феллер В.* Введение в теорию вероятностей и ее приложения. Т. 2. М.: Мир, 1984.
- [16] *Гихман И. И., Скороход А. В.* Теория случайных процессов. Т. 2. М.: Наука, 1973.

- [17] *Albeverio S., Bogachev L. V.* Branching random walk in a catalytic medium. I. Basic equations // *Positivity*. 2000. **4**. P. 41–100.
- [18] *Шубин М. А.* Псевдоразностные операторы и их функция Грина // *Изв. АН СССР. Серия матем.* 1985. **49**, № 3. С. 652–671.

Раздел 2

Исследования лабораторий

Введение

При кафедре теории вероятностей функционируют три лаборатории.

Лаборатория теории вероятностей существует с 1976 года. С 1997 года ее заведующим является профессор В.И. Питербарг.

На протяжении многих лет в лаборатории работали два научных семинара: вероятностных методов в технике и случайных процессов и полей, которые оказали значительное влияние на развитие соответствующих областей теории вероятностей и приложений. На их базе было организовано несколько всесоюзных и международных конференций, школ и симпозиумов.

В настоящее время основные исследования в лаборатории ведутся в области разработки асимптотических методов исследования локальных свойств и глобального поведения траекторий гауссовских случайных процессов и полей, а также статистической теории экстремумов. Работает новый учебно-исследовательский семинар «Статистика экстремумов».

В.И. Питербарг является членом редколлегии международного научного журнала «Extremes».

Лаборатория вычислительных средств (математико-экономического моделирования) работает при кафедре теории вероятностей с 1976 года. Заведующим лаборатории с момента ее основания является доцент Е.В. Чепурин.

В настоящее время при лаборатории функционирует компьютерный класс, где студенты и аспиранты в режиме свободного доступа выполняют свои работы. Там же проводятся статистические практикумы для студентов кафедры теории вероятностей.

С 1993 года, после создания на механико-математическом факультете по инициативе Б.В. Гнеденко математико-экономической специализации (экономического потока), эта лаборатория стала базовой для студентов данной специализации.

Лаборатория больших случайных систем основана в 1991 году. Инициатором ее создания и заведующим с момента основания является профессор В.А. Малышев.

Основные направления научной деятельности лаборатории — математическая статистическая физика и математика больших социальных сетей (связи, транспорта, обслуживания). С точки зрения теории вероятностей — это бесконечномерные динамические системы (классические и квантовые), гиббсовские случайные поля и процессы с локальным взаимодействием.

С 1995 года на базе лаборатории выходит научный журнал «Markov Processes and Related Fields». Журнал выходит на английском языке, имеет международную редколлегию, и на него подписаны многие зарубежные библиотеки.

ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ В МОДЕЛЯХ РАЗОРЕНИЯ С ГАУССОВСКИМИ УБЫТКАМИ

П. Булонь, С. Г. Кобельков, В. И. Питербарг¹⁾

Рассматривается финансово-актуарный процесс риска с большим начальным капиталом, линейным ростом доходов и убытками, моделируемые гауссовским процессом с растущей дисперсией. Рассмотрены две модели убытков — дробное броуновское движение и проинтегрированный гауссовский стационарный процесс. Для обеих моделей приводятся асимптотики вероятностей разорения и предельные теоремы о распределении момента разорения. Для первой модели изучено также асимптотическое поведение ожидаемых убытков в случае разорения.

1. Введение

Мы рассматриваем финансово-актуарный процесс риска

$$S_t = u + ct - Y_t, \quad (1)$$

параметр $u > 0$ интерпретируется как начальный капитал некоторой компании, $c > 0$ — скорость поступления доходов, наконец, случайный процесс Y_t интерпретируется как суммарные расходы. Таким образом, S_t представляет собой суммарный капитал компании в момент t . Разорение происходит, если в какой-либо момент выполняется соотношение $S_t < 0$. Таким образом, вероятность разорения есть

$$\psi(u) = P(\inf_{t \geq 0} (u + ct - Y_t) < 0) = P(\sup_{t \geq 0} (Y_t - ct) > u). \quad (2)$$

Классическим примером задачи о разорении является модель страхования Крамера–Лундберга (см. [1]), в которой $c > 0$ — скорость поступления взносов, $Y_t = \sum_{j=1}^{N_t} \eta_j$, N_t — пуассоновский процесс интенсивности $\lambda > 0$, а η_j — независимые, одинаково распределенные случайные величины, представляющие индивидуальные страховые выплаты такие, что $Ee^{v\eta_1} < \infty$ для любого $v > 0$ и $c > \lambda E\eta_1$. Тогда для вероятности разорения (2) выполняется оценка

$$P(\sup_{t \geq 0} Y_t - ct > u) \leq e^{-uv_0},$$

где $v_0 > 0$ — (единственный) корень уравнения $\lambda(Ee^{v\eta_1} - 1) = vc$.

Данная модель и ее модификации остаются актуальны их исследования продолжают с различных точек зрения. Так, в случае, когда выплаты имеют распределение фазового типа (то есть вероятностные распределения матрично-экспоненциального вида, которые можно рассматривать как времена перехода однородного марковского процесса с конечным числом состояний в особое поглощающее состояние), для преобразования Лапласа вероятности разорения можно выписать явное выражение через параметры модели [2, 3]. Ряд вопросов, связанных со случайным блужданием с отрицательным сносом, допускает интерпретацию в

¹⁾Исследования двух последних авторов поддержаны грантом РФФИ № 07-01-00077, исследования последнего автора также поддержаны грантом РФФИ № 06-01-00454.

рамках задачи о разорении [4]. Также важными для приложений оказываются случаи зависимых выплат, выплат с тяжелыми хвостами распределений [5].

В теории очередей рассматривается величина

$$Q(t) = \sup_{0 \leq s \leq t} (Q(0) + \xi(t) - \mu t, \xi(t) - \xi(s) - \mu(t - s)),$$

называемая загруженностью очереди, где $\xi(t)$ — суммарный объем входящего трафика на интервале $[0, t]$, а μ — скорость обработки данных; $Q(0)$ — загруженность в начальный момент времени.

В литературе в качестве $\xi(t)$ рассматриваются следующие процессы:

$\xi(t) = \int_0^t Z(s) ds$, где $Z(s)$ — марковская цепь с непрерывным временем [6, 7];

$\xi(t) = \int_0^t Z(s) ds$, где $Z(s)$ — суперпозиция он-офф-процессов [8];

$\xi(t) = \int_0^t Z(s) ds$, где $Z(s)$ — гауссовский процесс [9, 10].

Процесс $Z(t)$ — это скорость, с которой приходит в систему информация. Если $Z(t)$ — суперпозиция большого числа он-офф-процессов (или процессов, моделирующих скорость потока информации из одного источника), то центральная предельная теорема позволяет использовать гауссовскую аппроксимацию для $Z(t)$. Следовательно, процесс $\xi(t)$ также будет гауссовским и мы можем описывать входной поток в терминах кумулятивного процесса. Данный подход также позволяет рассматривать входные потоки с нигде не дифференцируемыми траекториями.

Предположим, что $\xi(t)$ может быть представлен в виде суммы $\zeta(t) + \tilde{d}(t)$, где $\zeta(t)$ — процесс со стационарными приращениями, а $\tilde{d}(t)$ — неслучайная непрерывная функция. Тогда процесс

$$\eta(v) = \zeta(t) - \zeta(t - v)$$

распределен так же, как и $\zeta(v)$, $v \in [0, T]$. Следовательно, если

$$Q(T) = \sup_{0 \leq s \leq T} (\zeta(T) - \zeta(s) + \tilde{d}(T) - \tilde{d}(s) - \mu(T - s))$$

— загруженность очереди, то для вероятности переполнения $\psi_T(u)$ выполнено равенство

$$\psi_T(u) = P(Q(T) > u) = P\left(\sup_{0 \leq v \leq T} \xi(v) - d_T(v) > u\right),$$

где $d_T(s) = -(\tilde{d}(T) - \tilde{d}(T - s)) + \mu s$. Если $d_T(s) = \lambda t$, то $d_T(s) = (\mu - \lambda)s$. Таким образом, мы снова приходим к задаче о разорении.

Часто исследователей интересует не кратковременное поведение системы, а долгосрочное. Для этого необходимо сделать еще несколько предположений: $\xi(t) = \zeta(t) + \lambda t$, где $\zeta(t)$ — центрированный гауссовский процесс с непрерывными траекториями и $\zeta(0) = 0$ почти наверное. Тогда, вводя $c = \mu - \lambda$ и предполагая $c > 0$, получим

$$Q(t) = \sup_{s \leq t} (\zeta(t) - \zeta(s) - c(t - s)).$$

Результаты исследований процесса передачи данных в Интернет показали, что для процесса $\xi(t)$ характерна автомодельность [11]. С другой стороны, в задачах о разорении важным является понятие интенсивности расходов X_t , то есть

$Y_t = \int_0^t X_s ds$; такое представление невозможно для автомодельных процессов, таких, как дробное броуновское движение. В данной работе мы рассматриваем две модели гауссовских убытков. Первой моделью является $Y_t = B_H(t)$, где $B_H(t)$ — дробное броуновское движение с показателем Хёрста H . Вторая модель представляет собой гауссовский проинтегрированный процесс $Y_t = \int_0^t X_s ds$, где X_t — гауссовский стационарный процесс. Заметим, что обе модели являются процессами со стационарными приращениями.

Отметим, что в работе [12] найдено выражение для асимптотики вероятности разорения $\psi_\infty(u)$ для достаточно общего класса гауссовских процессов, включающего обе рассматриваемые модели, однако общность постановки задачи усложнила формулировки условий применимости доказанных утверждений, а получение асимптотики для конкретных моделей требует даже более сложных вычислений, чем приведенные ниже. Ранее эта задача рассматривалась также в [13, 14].

2. Модель разорения с дробным броуновским движением в качестве процесса убытков

Точная асимптотика для вероятности разорения $\psi_\infty(u)$ следует из автомодельности этого процесса и общих результатов об асимптотиках больших уклонов гауссовских процессов. Дробным броуновским движением B_t^H , $t \geq 0$, с показателем Хёрста $H \in (0, 1)$ называется почти наверное (п.н.) непрерывный гауссовский процесс с нулевым средним, стационарными приращениями и дисперсией $E(B_t^H)^2 = t^{2H}$. Из определения следует свойство автомодельности этого процесса: для любого $a > 0$ случайные процессы $\{B_{at}^H, t \geq 0\}$ и $\{a^H B_t^H, t \geq 0\}$ распределены одинаково. Делая замену переменных $t = us$, получаем

$$\begin{aligned} P(\sup_{t \geq 0} B_t^H - ct > u) &= P(\sup_{s \geq 0} u^H B_s^H - cus > u) = \\ &= P\left(\sup_{s \geq 0} \frac{B_s^H}{1 + cs} > u^{1-H}\right). \end{aligned} \quad (3)$$

Для нахождения асимптотики вероятности в правой части замечаем, во-первых, что дисперсия гауссовского случайного процесса под знаком супремума равна $s^{2H}/(1 + cs)^2$, стремится к нулю при $s \rightarrow \infty$ и достигает своего максимального значения, равного

$$A^2 = (c/H)^{-2H} (1 - H)^{2-2H}, \quad (4)$$

в единственной точке $s_0 = \frac{H}{c(1-H)}$.

Во-вторых, для любого $s_1 > s_0$ вероятность

$$P\left(\sup_{s \geq s_1} \frac{B_s^H}{1 + cs} > u^{1-H}\right)$$

может быть оценена сверху суммой

$$\sum_{k=1}^{\infty} P\left(\sup_{s_1+k \geq s \geq s_1+k-1} \frac{B_s^H}{1 + cs} > u^{1-H}\right),$$

каждый член которой в силу неравенства теоремы 8.1 [15] не превосходит для некоторой абсолютной константы C величины

$$Cu^{1/H-2} \exp \left\{ -\frac{1}{2V_k} u^{2-2H} \right\},$$

где V_k — максимум дисперсии процесса $\frac{B_s^H}{1+cs}$ на отрезке $[s_1 + k - 1, s_1 + k]$. Элементарные вычисления показывают, что вся сумма поэтому экспоненциально меньше, чем вероятность

$$\mathbb{P} \left(\frac{B_{s_0}^H}{1 + cs_0} > u^{1-H} \right),$$

которая не превосходит искомую. Это позволяет заключить, что достаточно исследовать асимптотику рассматриваемой вероятности только на конечном отрезке $[0, s_1]$. Асимптотика этой вероятности в свою очередь найдена в теореме D3 [15]. Достаточно исследовать локальное поведение дисперсии процесса в окрестности точки ее абсолютного максимума, чтобы получить окончательно следующий результат. В более общем виде, но несколько другим методом этот результат доказан в [16].

Теорема 1. *Для модели разорения с дробным броуновским движением на входе имеем*

$$\mathbb{P}(\inf_{t \geq 0} S_t < 0) \sim \frac{\mathcal{H}_{2H} A^{1/H-3/2}}{B^{1/2} 2^{1/H}} u^{(1-H)(1-2H)/H} \exp \left(-\frac{1}{2} A^2 u^{2-2H} \right) \quad (5)$$

при $u \rightarrow \infty$, где A определено выше, $B = (c(1-H))^{H+2} H^{-H+1} u$

$$\mathcal{H}_{2H} = \lim_{\lambda \rightarrow \infty} \frac{1}{\lambda} \mathbb{E} e^{\max_{t \in [0, \lambda]} \sqrt{2} B_t^H - t^{2H}} \in (0, \infty)$$

— константа Пикандса.

О константе Пикандса известно, что $\mathcal{H}_1 = 1$, $\mathcal{H}_2 = \pi^{-1/2}$ (см. [15]) имеются некоторые оценки для других значений H [17]. Изложенный выше переход к вспомогательному гауссовскому процессу $B_s^H/(1+cs)$ можно использовать и для нахождения асимптотики вероятности разорения на конечном временном интервале $[0, T]$, а также на интервале, стремящемся к бесконечности. Нетрудно заметить, что из вышеприведенных (слегка модифицированных) оценок вероятности высокого максимума на интервале $[s_1, \infty)$ следует, что если $T = T_u \geq u(s_0 + \epsilon)$, $\epsilon > 0$, то асимптотика вероятности разорения на интервале $[0, T_u]$ такова:

$$\mathbb{P}(\inf_{t \in [0, T_u]} S_t < 0) \sim \mathbb{P}(\inf_{t \geq 0} S_t < 0).$$

Тот же прием позволяет получить точную асимптотику поведения вероятности разорения и для более коротких временных интервалов. В работе [18] доказан следующий результат:

Теорема 2. *Пусть для процесса риска S_t выполнены условия теоремы 1. Пусть $T \geq \varepsilon > 0$ для некоторого $\varepsilon > 0$, и $T/u \rightarrow b$ для некоторого $b \in [0, s_0)$ при $u \rightarrow \infty$. Тогда*

$$\mathbb{P}(\max_{[0, T]} S_t < 0) = \quad (6)$$

$$= \begin{cases} \frac{2^{2H+1/2} \mathcal{H}_{2H}}{(H(bc)\sqrt{\pi}} u^{-3+1/H} T^{3H-1} \exp\left(-\frac{1}{2} \left(\frac{u+cT}{T^H}\right)^2\right) (1+o(1)) & \text{для } H < 1/2, \\ \frac{kT^H}{\sqrt{2\pi}u} \exp\left(-\frac{1}{2} \left(\frac{u+cT}{T^H}\right)^2\right) (1+o(1)) & \text{для } H \geq 1/2 \end{cases}$$

при $u \rightarrow \infty$, где $H(a) = H - a/(1+a)$, $k = 2$ для $H = 1/2$ и $k = 1$ при $H > 1/2$.

Далее, при помощи перехода к вспомогательному процессу $B_s^H/(1+cs)$ может быть также исследовано предельное поведение условного распределения момента разорения

$$\tau_u = \inf\{t > 0 : S_t < 0\} \quad (7)$$

при условии, что разорение произошло. Для этого необходимо модифицировать доказательство теоремы D3 из [15], введя в нее локализацию точки достижения абсолютного максимума, то есть изучив асимптотики вероятности высокого максимума на подмножествах рассматриваемого временного интервала при условии, что высокий максимум на всем временном интервале имел место. Пусть $X(t)$, $t \in [0, T]$, — гауссовский процесс с непрерывными траекториями, нулевым средним, дисперсией $\sigma^2(t)$ и корреляционной функцией $r(t, s)$, которые удовлетворяют следующим трем условиям:

E1) функция $\sigma(t)$, $t \in [0, T]$, имеет единственную точку абсолютного максимума $t = t_0$, $0 < t_0 < T$, при этом для некоторых $a > 0$ и ζ

$$\sigma(t) = 1 - a|t - t_0|^\zeta(1 + o(1)) \quad \text{при } t \rightarrow t_0;$$

E2) для корреляционной функции имеет место следующее поведение в окрестности точки (t_0, t_0) :

$$r(t, s) = 1 - b|t - s|^\alpha(1 + o(1)) \quad \text{при } t, s \rightarrow t_0$$

для некоторых b и $\alpha > 0$;

E3) найдутся $\gamma > 0$ и G , такие, что $E(X(t) - X(s))^2 \leq G|t - s|^\gamma$.

Для случайного процесса $\xi(t)$, $t \in R$, обозначим момент первого достижения

$$\tau_u(\xi, S) = \inf\{t : t \in S, \xi(t) > u\}.$$

Будем обозначать той же буквой B_t^H , $t \in R$, двустороннее дробное броуновское движение, то есть гауссовский процесс с нулевым средним, непрерывными траекториями и ковариационной функцией, равной $|s|^{2H} + |t|^{2H} - |s - t|^{2H}$, $s, t \in R$. Для произвольного $a > 0$ обозначим $\chi_a(t) := \sqrt{2}B_t^{\alpha/2} - (1+a)|t|^\alpha$, $t \in R$. Введем аддитивную функцию на ограниченных борелевских подмножествах S действительной прямой:

$$\mathcal{H}_\alpha^a(S, \lambda) = \int_0^\infty e^{wP(\tau_w(\chi_a, [-\lambda, \lambda]) \in S)} dw.$$

В силу ограниченности S , функция $\mathcal{H}_\alpha^a(S, \lambda)$ не возрастает по λ при всех достаточно больших λ , следовательно, предел

$$\mathcal{H}_\alpha^a(S) = \lim_{\lambda \rightarrow \infty} \mathcal{H}_\alpha^a(S, \lambda)$$

существует и также является аддитивной неотрицательной функцией ограниченных борелевских множеств. В [15] показано, что константа

$$\mathcal{H}_\alpha^a = \mathcal{H}_\alpha^a(R).$$

конечна и положительна.

Теорема 3. Пусть выполнены условия **E1**, **E2** и **E3**. Тогда если $\zeta > \alpha$, то условное распределение

$$P(u^{2/\zeta}(\tau - t_0) \in (\cdot) \mid \max_{t \in [0, T]} X(t) > u)$$

слабо сходится при $u \rightarrow \infty$ к симметрическому распределению Леви, то есть к распределению с плотностью, пропорциональной $\exp(-a|x|^\zeta)$.

Если $\zeta = \alpha$, то это распределение сходится к распределению $\mathcal{H}_\alpha^a(\cdot)/\mathcal{H}_\alpha^a$.

Если $\zeta < \alpha$, то это распределение сходится к δ -мере точки 0.

Доказательство. Первое и второе утверждения следуют из слабой сходимости условных распределений случайного процесса $X_u(t) := u(X(u^{-2/\alpha}(t-t_0)-u)+w)$ при условии $X(t_0) = u - w/u$ при $u \rightarrow \infty$ к распределению случайного процесса $\chi_a(t)$, доказанной в лемме D1 и теореме D3 [15] и из стандартных рассуждений, аналогичных приведенным в доказательствах этих результатов. Для доказательства третьего утверждения достаточно показать, что для любых $d > c > 0$

$$P(u^{2/\zeta}(\tau - t_0) \in [c, d]) = o(P(X(t_0) > u)).$$

Но в силу теоремы D3 [15]

$$P(\max_{[0, T]} X(t) > u) \sim P(X(t_0) > u).$$

Отсюда следует, что исследуемая условная вероятность отрезка $[c, d]$ стремится к нулю, что дает утверждение теоремы. Обозначая $\kappa = 2/\alpha - 2/\zeta$ (в данном случае $\kappa < 0$), имеем

$$\begin{aligned} P(u^{2/\zeta}(\tau - t_0) \in [c, d]) &\leq P\left(X(t_0) < u, \max_{[cu^{-2/\zeta}, du^{-2/\zeta}]} X(t) > u\right) = \\ &= u^{-1} \int_0^\infty \varphi(u - w/u) P\left(\max_{[cu^\kappa, du^\kappa]} X_u(t) > w \mid X(t_0) = u - w/u\right) dw \leq \\ &\leq u^{-1} e^{-u^2/2} \int_0^\infty e^w P\left(\max_{[0, \epsilon]} \chi_a(t) > w\right) dw \end{aligned}$$

для произвольного $\epsilon > 0$ и всех достаточно больших u . Через φ обозначена гауссовская стандартная плотность распределения. Последний интеграл, как показано в доказательстве теоремы D3 [15] может быть сделан сколь угодно малым путем выбора достаточно малого ϵ . Теорема доказана.

Из первого утверждения этой теоремы при помощи перехода к вспомогательному процессу легко получаем предельную теорему для момента разорения.

Теорема 4. Для момента разорения в модели риска с дробным броуновским движением на входе имеем для всех x

$$P(Cu^{-H}(\tau_u - s_0u) < x \mid \tau_u < \infty) \rightarrow \Phi(x)$$

при $u \rightarrow \infty$, где

$$C = c^{H+1} \left(\frac{1-H}{H} \right)^{H+1/2}, \quad s_0 = \frac{H}{c(1-H)} \quad (8)$$

и Φ — стандартная гауссовская функция распределения.

Доказательство. Действительно, для нашего вспомогательного процесса $\zeta = 2$, $\alpha = 2H$, остальные константы пересчитываются при помощи линейной замены времени.

Рассмотрим еще один важный функционал в модели разорения — ожидаемые потери в случае разорения. Он определяется как

$$L(u, T; S) := \mathbb{E} \left(- \inf_{t \in [0, T]} S_t \mid \tau_u \leq T \right). \quad (9)$$

Мы изучим его поведение при больших значениях начального капитала, то есть при $u \rightarrow \infty$. Для простоты, как и ранее, будем считать, что временной горизонт T либо равен бесконечности, либо достаточно большой: $T = T_u \geq u(s_0 + \epsilon)$ для всех достаточно больших u и некоторого $\epsilon > 0$.

Теорема 5. Для ожидаемых потерь в случае разорения (9) в модели разорения с дробным броуновским движением на входе выполнено

$$L(u, T; S) \sim \frac{2H^{2H} u^{2H-1}}{c^{2H} (1-H)^{2H-1}} \quad (10)$$

при $u \rightarrow \infty$ и $T \geq u(s_0 + \epsilon)$ для всех достаточно больших u и некоторого $\epsilon > 0$.

Утверждение этой теоремы следует из асимптотики вероятности разорения, если принять во внимание следующую простую лемму из анализа, доказательство которой в свою очередь очевидно.

Лемма. Пусть о случайной величине X известно, что

$$\mathbb{P}(X > x) = A(x) \exp(-Cx^h)(1 + o(1)) \text{ и } x\mathbb{P}(X > x) \rightarrow 0$$

при $x \rightarrow \infty$, причем константы $h > 0$, $C > 0$ и функция $A(x)$ удовлетворяют следующим предположениям.

1. Для всех v существует предел

$$\lim_{x \rightarrow \infty} \frac{A(x(1 + v/Cx^h)^{1/h})}{A(x)} =: B(v).$$

2. Найдется мажорирующая функция $B_1(v)$, такая, что

$$\int_0^\infty (1+v)^{(1/h-1)+} B_1(v) e^{-v} dv < \infty$$

$(a_+ := \max(0, a))$ и

$$\frac{A(x(1 + v/Cx^h)^{1/h})}{A(x)} \leq B_1(v),$$

для всех достаточно больших x и всех $v \geq 0$. Тогда

$$\mathbb{E}(X \mid X > x) = x + ax^{1-h} + o(x^{1-h}), \quad x \rightarrow \infty,$$

где

$$a = \frac{1}{Ch} \int_0^\infty B(v) e^{-v} dv.$$

Комбинация этой леммы и теоремы 1 дает утверждение теоремы.

Заметим, что если $H > 1/2$, то ожидаемые потери в случае разорения растут вместе с u со скоростью u^{2H-1} ! В случае $H = 1/2$ дробное броуновское движение превращается в просто броуновское движение и ожидаемые потери нейтральны к начальному размеру капитала, что вполне соответствует марковскому характеру броуновского движения. В этом случае нетрудно вывести точную формулу для ожидаемых потерь, которые действительно не зависят от u . В случае $H < 1/2$ риск оказывается с менее тяжелыми последствиями, поскольку ожидаемые потери стремятся к нулю при росте начального капитала. Это связано с тем фактом, что приращения дробного броуновского движения положительно коррелированы для $H > 1/2$ и отрицательно коррелированы при $H < 1/2$. Это наблюдение показывает, в частности, что необходимо тщательно изучать корреляционную структуру движения цен и других финансовых показателей.

Оптимизация портфеля из двух независимых рисков на основе ожидаемых потерь. Рассмотрим пример применения полученных приближенных формул для ожидаемых потерь в случае разорения. Пусть имеется портфель из двух независимых активов

$$\begin{aligned} S_t^1 &= u_1 + c_1 t - \sigma_1 B_t^H, \\ S_t^2 &= u_2 + c_2 t - \sigma_2 \tilde{B}_t^H, \end{aligned}$$

где $c_1, c_2, \sigma_1, \sigma_2 > 0$, $u_1 + u_2 = u$, — начальный капитал, подлежащий оптимальному делению на две части, $B_t^H, \tilde{B}_t^H$ — два независимых дробных броуновских процесса с одним и тем же показателем Хёрста $H \in (0, 1)$, моделирующие убытки в каждом из активов. Пусть рассматриваемый временной горизонт T удовлетворяет условиям теоремы 5 относительно обеих частей капитала u_1 и u_2 . Перейдем к стандартизованным активам

$$\begin{aligned} \sigma_1^{-1} S_t^1 &= u_1 \sigma_1^{-1} + c_1 \sigma_1^{-1} t - B_t^H, \\ \sigma_2^{-1} S_t^2 &= u_2 \sigma_2^{-1} + c_2 \sigma_2^{-1} t - \tilde{B}_t^H, \end{aligned}$$

и применим к ним теорему 5. После несложных вычислений получаем, что

$$L(u_i, T; S_i) \sim p_H \sigma_i^2 c_i^{-2H} u_i^{2H-1}$$

при $u_i \rightarrow \infty$, $i = 1, 2$, где через p_H обозначена константа, не зависящая от номера актива i .

На основании этих асимптотических приближений и заданной (малой) допустимой вероятности разорения ε можно определить уровень риска (value at risk, $V@R$) объединенного актива $S_t^1 + S_t^2$ как минимальное положительное решение u_ε уравнения относительно u

$$P(\inf_{T \geq t \geq 0} (S_t^1 + S_t^2) < 0) = \varepsilon.$$

Поскольку имеет место равенство по распределениям

$$\sigma_1 B_t^H + \sigma_2 \tilde{B}_t^H \stackrel{d}{=} \sqrt{\sigma_1^2 + \sigma_2^2} B_t^H,$$

получаем для u_ε асимптотическое уравнение

$$C'(u')^{(1-H)(1-2H)/H} \exp\left(-\frac{1}{2}A'(u')^{2-2H}\right) (1 + o(1)) = \varepsilon, \quad (11)$$

где $\varepsilon \rightarrow 0$, $u' = u/\sqrt{\sigma_1^2 + \sigma_2^2}$, и C' , A' вычисляются на основании формул (4) и (8) с $c = (c_1 + c_2)/\sqrt{\sigma_1^2 + \sigma_2^2}$. Асимптотическое решение этого уравнения есть

$$\begin{aligned} u_\varepsilon &= (\ln(-\varepsilon)/A')^{1/(2-2H)} + \\ &+ \left(\frac{2 \ln C''}{A'(2-2H)} + \frac{1-2H}{2H} \ln(2 \ln(-\varepsilon)/A') \right) (2 \ln(-\varepsilon)/A')^{(2H-1)/(2-2H)} + \\ &+ O(1/\ln(-\varepsilon)) \end{aligned}$$

при $\varepsilon \rightarrow 0$, где $C'' = C' \Phi((u'_\varepsilon)^{-H}(T - s_0 u'_\varepsilon))$, $u'_\varepsilon = u_\varepsilon/\sqrt{\sigma_1^2 + \sigma_2^2}$, ниже мы приведем соответствующие вычисления.

Теперь можно определить, как разделить оптимальным образом начальный капитал u_ε , обеспечивающий вероятность разорения ε , на две части: $u_\varepsilon = u_1 + u_2$, так, чтобы минимизировать суммарные потери $L(u_1, T; S_1) + L(u_2, T; S_2)$ в случае разорения объединенного актива $S_t^1 + S_t^2$. В нашей асимптотической постановке следует минимизировать функцию

$$\sigma_1^2 c_1^{-2H} u_1^{2H-1} + \sigma_2^2 c_2^{-2H} u_2^{2H-1}$$

при условии $u_1 + u_2 = u_\varepsilon$. Если $H < 1/2$, то минимум достигается при

$$u_1 = u_\varepsilon \frac{(\sigma_1^2 c_1^{-2H})^{1/(2H-2)}}{(\sigma_1^2 c_1^{-2H})^{1/(2H-2)} + (\sigma_2^2 c_2^{-2H})^{1/(2H-2)}}.$$

В броуновском случае ($H = 1/2$) нет асимптотических предпочтений, как разделять начальные капиталы в портфеле. В случае $H > 1/2$ следует вложиться в один из активов с минимальным $\sigma_i^2 c_i^{-2H}$, $i = 1, 2$, то есть

$$u_1 = u_\varepsilon \mathbf{I}_{\sigma_1^2 c_1^{-2H} \leq \sigma_2^2 c_2^{-2H}}.$$

Аналогичным образом можно рассмотреть другие возможные соотношения между T , u_1 и u_2 .

Решение асимптотического уравнения (11). Найдем асимптотику минимального положительного решения асимптотического уравнения

$$Cx^n \exp(-Ax^m) = \varepsilon$$

при $\varepsilon \rightarrow \infty$, где $C, A, m > 0$. Логарифмируя, получаем, что

$$\ln C + n \ln x - Ax^m = -\ln(-\varepsilon).$$

Делая замену переменных

$$x = (A^{-1} \ln(-\varepsilon))^{1/m} + y (A^{-1} \ln(-\varepsilon))^{(1-m)/m}$$

и применяя формулу Тейлора, имеем

$$Ax^m = \ln(-\varepsilon) + my + O(1/\ln(-\varepsilon))$$

при $\varepsilon \rightarrow 0$. Еще раз разлагая в ряд Тейлора, получаем

$$y = \frac{\ln C}{Am} + \frac{n}{m} \ln(A^{-1} \ln(-\varepsilon)) + O(1/\ln(-\varepsilon)).$$

Отсюда

$$\begin{aligned} x_\varepsilon &= (A^{-1} \ln(-\varepsilon))^{1/2} + \\ &+ \left(\frac{\ln C}{Am} + \frac{n}{m} \ln(A^{-1} \ln(-\varepsilon)) \right) (A^{-1} \ln(-\varepsilon))^{(1-m)/m} + O(1/\ln(-\varepsilon)). \end{aligned} \quad (12)$$

3. Модель разорения с проинтегрированным гауссовским процессом в качестве процесса убытков

В этом разделе мы рассмотрим случайный процесс

$$Y_t = \int_0^t X_s ds - ct,$$

где X_t — стационарный действительнзначный центрированный гауссовский процесс с п.н. непрерывными траекториями, $c > 0$.

Пусть $R(t)$ — действительная непрерывная ковариационная функция процесса X_t , удовлетворяющая условиям: (i) $G = \int_0^\infty R(s) ds > 0$, (ii) $H = \int_0^\infty sR(s) ds$ конечен, (iii) $\int_0^\delta \frac{R(t) - R(0)}{t} dt < \infty$ для некоторого $\delta > 0$.

Положим $\kappa(x) = u/c + \sqrt{2Gu}c^{-3/2}x$.

Теорема 6. В вышеприведенных условиях для любого $x > 0$

$$P(\tau_u < \kappa(x) | \tau_u < \infty) \rightarrow \Phi(x)$$

при $u \rightarrow \infty$.

Доказательство. По определению условной вероятности

$$P(\tau_u < \kappa(x) | \tau_u < \infty) = \frac{P(\tau_u < \kappa(x))}{P(\tau_u < \infty)}. \quad (13)$$

Рассмотрим вероятность $P(\tau_u < T)$, где T — это $\kappa(x)$ либо ∞ . Тогда

$$P(\tau_u < T) = P\left(\sup_{[0,T)} \frac{1}{(1+ct/u)} \int_0^t X_s ds > u\right).$$

Обозначим $V_t = \frac{1}{(1+ct/u)} \int_0^t X_s ds$ — аналог вспомогательного процесса из первой части статьи. Поскольку X_s не является автомодельным, исследование вспомогательного процесса существенно усложняется.

Положим $\alpha = \frac{u - \ln u \sqrt{u}}{c}$, $\beta = \frac{u + \ln u \sqrt{u}}{c}$. Тогда, если $T > \beta$, то

$$\begin{aligned} P(\tau_u < T) &= P(\sup_{[0,\alpha]} V_t \geq u) + P(\sup_{[\alpha,\beta]} V_t \geq u) + P(\sup_{(\beta,T)} V_t \geq u) - \\ &- P(\sup_{[0,\alpha]} V_t \geq u, \sup_{[\alpha,\beta]} V_t \geq u) - P(\sup_{[0,\alpha]} V_t \geq u, \sup_{(\beta,T)} V_t \geq u) - \\ &- P(\sup_{[\alpha,\beta]} V_t \geq u, \sup_{(\beta,T)} V_t \geq u) + P(\sup_{[0,\alpha]} V_t \geq u, \sup_{[\alpha,\beta]} V_t \geq u, \sup_{(\beta,T)} V_t \geq u) \leq \\ &\leq 2P(\sup_{[0,\alpha]} V_t \geq u) + P(\sup_{[\alpha,\beta]} V_t \geq u) + P(\sup_{(\beta,T)} V_t \geq u). \end{aligned} \quad (14)$$

Если же $T < \beta$, то

$$\begin{aligned} P(\tau_u < T) &= P(\sup_{[0,\alpha]} V_t \geq u) + P(\sup_{(\alpha,T)} V_t \geq u) - \\ &- P(\sup_{[0,\alpha]} V_t \geq u, \sup_{(\alpha,T)} V_t \geq u) \leq \\ &\leq P(\sup_{[0,\alpha]} V_t \geq u) + P(\sup_{(\alpha,T)} V_t \geq u). \end{aligned} \quad (15)$$

Кроме того,

$$P(\tau_u < T) \geq P(\sup_{[\alpha,\beta]} V_t \geq u). \quad (16)$$

Оценим $P(\sup_{[0,\alpha]} Y_t \geq u)$. Для этого воспользуемся неравенством из теоремы 8.1 [15]:

$$P(\max_{t \in S} V_t > u) < C_1 \max_{t \in S} (\text{Var} V_t) \exp \left\{ -\frac{u^2}{2 \max_{t \in S} \text{Var} V_t} \right\},$$

где C_1 может зависеть только от $s_1 - s_0$.

Дисперсия процесса V_t равна

$$\text{Var} V_t = \frac{2Gt - 2H}{(1+ct/u)^2} - \frac{2Gt \int_t^\infty R(s) ds - 2 \int_t^\infty s R(s) ds}{(1+ct/u)^2} = S_1(t) + R_1(t). \quad (17)$$

Функция $S_1(t)$ достигает максимума в точке $t = u/c$ и принимает значение $S_1(u/c) = \frac{uG - cH}{2c}$.

Обозначим (какую-нибудь) точку достижения максимума дисперсией процесса V_t на промежутке $\delta_k = [k, \min(\alpha, k+1))$ через t_k^* , $k = 0, \dots, [\alpha]$. Тогда

$$\begin{aligned} \exp \left\{ -\frac{u^2}{2 \sup_{\delta_k} \text{Var} V_t} \right\} &= \exp \left\{ -\frac{u^2}{2(S_1(t_k^*) + R_1(t_k^*))} \right\} = \\ &= \exp \left\{ -\frac{u^2}{2(S_1(\alpha) + (S_1(t_k^*) + R_1(t_k^*) - S_1(\alpha)))} \right\} \leq C_2 \exp \left\{ -\frac{u^2}{2S_1(\alpha)} \right\} \end{aligned}$$

для некоторого $C_2 > 0$ и достаточно больших u , так как $S_1(t^*) - S_1(\alpha) \leq 0$, а $R_1(t^*)$ является ограниченной функцией по u . Поскольку $2S_1(\alpha) < S_1(u/c)$ для достаточно больших u и

$$P(\sup_{[0,\alpha]} V_t \geq u) \leq \sum_k P(\sup_{\delta_k} V_t \geq u) \leq C_2 \alpha \exp \left\{ -\frac{u^2}{2S_1(\alpha)} \right\},$$

то

$$P(\sup_{[0,\alpha]} V_t > u) = o \left(\exp \left\{ -\frac{u^2}{2S_1(u/c)} \right\} \right). \quad (18)$$

Перейдем к оценке $P(\sup_{(\beta,\infty)} V_t \geq u)$. Рассмотрим теперь интервалы $\delta_k = [\beta + k, \beta + (k+1))$, $k = 0, 1, \dots$. Очевидно, что

$$P(\sup_{(\beta,\infty)} V_t \geq u) \leq \sum_{k=0}^{\infty} P(\sup_{\delta_k} V_t \geq u).$$

Далее, для $t \geq \beta$

$$\exp \left\{ -\frac{u^2}{2DV_t} \right\} = \exp \left\{ -\frac{u^2}{2S_1(t)} + \frac{u^2 R_1(t)}{2S_1(t)(S_1(t) + R_1(t))} \right\} \leq C_3 \exp \left\{ -\frac{u^2}{2S_1(t)} \right\}$$

для некоторого $C_3 > 0$ в силу условия (ii). Кроме того,

$$S_1(\beta + k) \leq S_1(\beta) - \frac{G\beta c u^{-1}(\beta + k)^2}{(1 + c(\beta + k)/u)^2(1 + c\beta/u)^2}.$$

Следовательно,

$$\sum_{k=0}^{\infty} P(\sup_{\delta_k} V_t \geq u) \leq \sum_{k=0}^{\infty} C_3 \exp \left\{ -\frac{u^2}{\sup_{\delta_k} S_1(t)} \right\} \leq C_4 \exp \left\{ -\frac{u^2}{2S_1(\beta)} \right\}$$

для некоторого $C_4 > 0$. Так как $2S_1(\beta) < S_1(u/c)$ для достаточно больших u , то

$$P(\sup_{(\beta,\infty)} V_t > u) = o \left(\exp \left\{ -\frac{u^2}{2S_1(u/c)} \right\} \right). \quad (19)$$

Перейдем теперь к изучению асимптотики вероятности $P(\sup_{[\alpha,\beta]} V_t > u)$.

Определим случайную величину $N_u([0, T])$, равную числу пересечений уровня u процессом Y_t на отрезке $[0, T]$. В [15, 19] показано, что для случайного процесса с непрерывно дифференцируемыми траекториями и произвольного отрезка S имеет место соотношение

$$0 \leq EN_u(S) - P(\max_{t \in S} Y_t \geq u) \leq \frac{1}{2} EN_u(S)(N_u(S) - 1). \quad (20)$$

Среднее число пересечений уровня u процессом Y_t на отрезке $[0, T]$ вычисляется по формуле Райса (см. [15]):

$$EN_u([0, T]) = \int_0^T \int_0^\infty |y| p_t(u, y) dy dt, \quad (21)$$

где $p_t(u, y)$ — совместная плотность распределения случайных величин Y_t, Y'_t . Введем обозначения: $a(t) = \text{Var}(V_t)$, $b(t) = \text{cov}(V_t, V'_t)$, $\sigma(t) = \sqrt{a(t)R(0) - b(t)^2}$. Обозначим $\gamma = \gamma(t, u) = c\sqrt{a(t)}/\sigma(t) - b(t)(u + ct)/(\sqrt{a(t)}\sigma(t))$,

$$g(t) = \frac{1}{\sqrt{2\pi}} \int_{\gamma}^{\infty} ye^{-y^2/2} dy - \gamma(1 - \Phi(\gamma)),$$

где $\Phi(t)$ — функция распределения стандартной нормальной случайной величины. Применим (21) для вычисления $\mathbb{E}N_u([\alpha, u/c])$ и произведем замену переменной $t = t(\tau) = u\tau/c + u/c$. Тогда после взятия интеграла по переменной y получим

$$\mathbb{E}N_u([\alpha, u/c]) = \frac{u}{c\sqrt{2\pi}} \int_{-\ln u/\sqrt{u}}^0 \frac{\sigma(t(\tau))}{a(\tau)} g(t(\tau)) \exp \left\{ -\frac{u^2(1+\tau)^2}{2a(\tau)} \right\} d\tau,$$

где $a(\tau) = a(t(\tau))$.

Заметим, что $a(\tau)/(1+\tau)^2 = S_1(t(\tau)) + R_1(t(\tau))$. Таким образом,

$$\frac{u^2}{2(S_1(t(\tau)) + R_1(t(\tau)))} = S_3(\tau) - \frac{u^2 R_1(t(\tau))}{2S_1(t(\tau))(S_1(t(\tau)) + R_1(t(\tau)))}, \quad (22)$$

где $S_3(\tau) = \frac{u^2}{2S_1(t(\tau))}$. В силу предположений об интегрируемости ковариационной функции $R(t)$ и того, что на рассматриваемом отрезке $S_1(t(\tau)) = O(u)$, последнее слагаемое в (22) стремится к нулю. Следовательно,

$$\mathbb{E}N_u([\alpha, u/c]) = \frac{u}{c\sqrt{2\pi}} \int_{-\ln u/\sqrt{u}}^0 \frac{\sigma(t(\tau))}{a(\tau)} g(t(\tau)) \exp \{-S_3(\tau)\} d\tau (1 + o(1)). \quad (23)$$

Имеют место равномерные по τ соотношения при $u \rightarrow \infty$

$$S_3''(\tau) = S_3''(0)(1 + o(1)), \quad \frac{\sigma(t(\tau))}{a(\tau)} g(t(\tau)) = \frac{\sigma(t(0))}{a(0)} g(t(0))(1 + o(1)). \quad (24)$$

Поскольку $(S_3''(0))^{-1/2} = o(\ln u/\sqrt{u})$ при $u \rightarrow \infty$, можем применить метод Лапласа [20, теорема 2.2] оценки интегралов к (23):

$$\mathbb{E}N_u([\alpha, u/c]) = \frac{u}{2c\sqrt{S_3''(0)}} \frac{\sigma(t(0))}{a(t(0))} g(t(0)) \exp \{-S_3(0)\}.$$

Нетрудно вычислить, что при $\tau = 0$, $u \rightarrow \infty$ выполнено

$$\begin{aligned} \gamma &\rightarrow 0, \quad g(t(0)) \rightarrow 1/\sqrt{2\pi}, \\ \sigma(t(0))/a(0) &= \frac{\sqrt{R(0)}}{\sqrt{2G}} u^{-1/2} c^{1/2} (1 + o(1)), \\ S_3''(0) &= \frac{cu}{2G} (1 + o(1)). \end{aligned} \quad (25)$$

Таким образом,

$$\mathbb{E}N_u([\alpha, u/c]) = \frac{\sqrt{R(0)}}{2\sqrt{2\pi}} c^{-1} \exp \{-S_3(0)\} (1 + o(1)). \quad (26)$$

Аналогично

$$\mathbf{E}N_u([u/c, \beta]) = \frac{\sqrt{R(0)}}{2\sqrt{2\pi}} c^{-1} \exp \{-S_3(0)\} (1 + o(1)). \quad (27)$$

Вычислим $\mathbf{E}N_u([u/c, \kappa(x)])$. Так как $S_3(\tau) - S_3(0)$ является неотрицательной монотонной функцией на отрезке с концами в точках $0, \kappa(x)c/u - 1$, то можно произвести замену переменной $y^2(\tau) = 2(S_3(\tau) - S_3(0))$ в интеграле (23). Поскольку функция $S_3(\tau)$ имеет минимум в точке $\tau = 0$ и $\kappa(x)c/u - 1 \rightarrow 0$, то $S_3(\tau) - S_3(0) = S_3''(0)\tau^2/2(1 + o(1))$ при $u \rightarrow \infty$. Тогда

$$\begin{aligned} y^2(\kappa(x)c/u - 1) &= 2(S_3(\kappa(x)c/u - 1) - S_3(0)) = \\ &= S_3''(0)(\kappa(x)c/u - 1)^2(1 + o(1)) = x^2(1 + o(1)) \end{aligned}$$

в силу выбора $\kappa(x)$.

Поскольку

$$\frac{dy}{d\tau} = \frac{1}{\sqrt{2}} \sqrt{\frac{d^2(y^2)}{d\tau^2}} (1 + o(1))$$

при $\tau \rightarrow 0$ и

$$\frac{d^2(y^2)}{d\tau^2} = 2S_3''(\tau),$$

то

$$\frac{dy}{d\tau} = \sqrt{S_3''(0)} (1 + o(1)).$$

Отсюда, воспользовавшись равномерной сходимостью (24) и соотношениями (25), получим

$$\mathbf{E}N_u([u/c, \kappa(x)]) = \frac{\sqrt{R(0)}}{\sqrt{2\pi}} c^{-1} \exp \{-S_3(0)\} \left(\frac{1}{\sqrt{2\pi}} \int_0^x e^{-y^2/2} dy \right) (1 + o(1)).$$

Так как $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 e^{-x^2/2} dx = 1/2$, то в силу равенств

$$\begin{aligned} \mathbf{E}N_u([\alpha, \beta]) &= \mathbf{E}N_u([\alpha, 0]) + \mathbf{E}N_u((0, \beta]), \\ \mathbf{E}N_u([\alpha, \kappa(x)]) &= \mathbf{E}N_u([\alpha, 0]) + \mathbf{E}N_u((0, \kappa(x)]) \end{aligned}$$

имеем

$$\mathbf{E}N_u([\alpha, \kappa(x)]) = \frac{\sqrt{R(0)}}{\sqrt{2\pi}} c^{-1} \exp \{-S_3(0)\} \Phi(x) (1 + o(1)) \quad (28)$$

и

$$\mathbf{E}N_u([\alpha, \beta]) = \frac{\sqrt{R(0)}}{\sqrt{2\pi}} c^{-1} \exp \{-S_3(0)\} (1 + o(1)). \quad (29)$$

Перейдем к оценке второго факториального момента $\mathbf{E}N_u([\alpha, \beta])(N_u([\alpha, \beta]) - 1)$. Сразу же заметим, что $\mathbf{E}N_u([\alpha, \kappa(x)])(N_u([\alpha, \kappa(x)]) - 1) \leq \mathbf{E}N_u([\alpha, \beta])(N_u([\alpha, \beta]) - 1)$.

Положим

$$Z_t = \frac{u^{-1/2}}{1+t} \int_0^{ut/c} X_s ds.$$

Тогда

$$P(\sup_{[\alpha, \beta]} V_t > u) = P(\sup_{[\tilde{\alpha}, \tilde{\beta}]} Z_t > \sqrt{u}),$$

где $\tilde{\alpha} = \alpha c/u, \tilde{\beta} = \beta c/u$, причем пересечению уровня u процессом V_t в точке t_0 соответствует пересечение уровня $\sqrt{u}$ процессом Z_t в точке $t_0 c/u$. Следовательно, $EN_u([\alpha, \beta])(N_u([\alpha, \beta]) - 1)$ для процесса V_t совпадает с соответствующим вторым факториальным моментом процесса Z_t на отрезке $[\tilde{\alpha}, \tilde{\beta}]$ уровня $\sqrt{u}$.

Воспользуемся выражением второго факториального момента через совместную плотность процесса Z_t и его производной [15]:

$$\begin{aligned} & EN_{\sqrt{u}}([\tilde{\alpha}, \tilde{\beta}])(N_{\sqrt{u}}([\tilde{\alpha}, \tilde{\beta}]) - 1) = \\ & = \int_{\tilde{\alpha}}^{\tilde{\beta}} \int_{\tilde{\alpha}}^{\tilde{\beta}} \int_0^\infty \int_0^\infty y_1 y_2 \varphi_{t,s,t,s}(\sqrt{u}, \sqrt{u}, y_1, y_2) dy_1 dy_2 ds dt, \end{aligned}$$

где $\varphi_{t,s,t,s}(\sqrt{u}, \sqrt{u}, y_1, y_2)$ — совместная плотность случайных величин Z_t, Z_s, Z'_t, Z'_s . Обозначим $\Omega = \{(s, t) : x \in [\tilde{\alpha}, \tilde{\beta}], y \in [\tilde{\alpha}, \tilde{\beta}]\}$. Тогда

$$\begin{aligned} & \int_{\tilde{\alpha}}^{\tilde{\beta}} \int_{\tilde{\alpha}}^{\tilde{\beta}} \int_0^\infty \int_0^\infty y_1 y_2 \varphi_{t,s,t,s}(\sqrt{u}, \sqrt{u}, y_1, y_2) dy_1 dy_2 ds dt \leq \\ & \leq \iint_{\Omega \cap \{t \geq s\}} ds dt \int_0^\infty y^2 \varphi_{t,s,t}(y) dy + \iint_{\Omega \cap \{t \leq s\}} ds dt \int_0^\infty y^2 \varphi_{t,s,s}(y) dy, \end{aligned}$$

здесь $\varphi_{t,s,t}(y)$ — совместная плотность гауссовских величин Z_t, Z_s, Z'_t в точке $(\sqrt{u}, \sqrt{u}, y)$, а $\varphi_{t,s,s}(y)$ — совместная плотность величин Z_t, Z_s, Z'_s в той же точке. Оценим первый интеграл; оценка второго интеграла проводится аналогично.

Переходя к условной плотности, получим

$$\begin{aligned} & \iint_{\Omega \cap \{t \geq s\}} ds dt \int_0^\infty y^2 \varphi_{t,s,t}(y) dy = \\ & = \iint_{\Omega \cap \{t \geq s\}} \varphi_{t,s}(\sqrt{u}, \sqrt{u}) ds dt \int_0^\infty y^2 \varphi_{t,s}(y) dy ds dt. \end{aligned}$$

Таким образом,

$$\begin{aligned} & EN_{\sqrt{u}}([\tilde{\alpha}, \tilde{\beta}])(N_{\sqrt{u}}([\tilde{\alpha}, \tilde{\beta}]) - 1) \leq \\ & \leq 2 \iint_{\Omega \cap \{t \geq s\}} \varphi_{t,s}(\sqrt{u}, \sqrt{u}) ds dt \int_0^\infty y^2 \varphi_{t,s}(y) dy. \end{aligned} \quad (30)$$

Здесь через $\varphi_{t,s}(\sqrt{u}, \sqrt{u})$ обозначена совместная плотность величин Z_t, Z_s , а через $\varphi_{t,s}(y)$ — условная плотность случайной величины Z'_t при условии $Z_t = \sqrt{u}, Z_s = \sqrt{u}$:

$$\begin{aligned} \varphi_{t,s}(\sqrt{u}, \sqrt{u}) &= \frac{1}{2\pi\sigma} \exp\left\{-\frac{u}{2\sigma^2}(\text{Var}Z_s + \text{Var}Z_t - 2\text{cov}(Z_s, Z_t))\right\}, \\ \varphi_{t,s}(y) &= \frac{1}{\sqrt{2\pi}\Delta^2} \exp\left\{-\frac{1}{2\Delta^2}(y - \sqrt{u}m)^2\right\}, \end{aligned}$$

где

$$\begin{aligned}\sigma^2 &= \text{Var} Z_t \mathbf{D} Z_s - \text{cov}(Z_t, Z_s)^2, \\ m &= \frac{1}{\sigma^2} (\text{Var} Z_s b(t) - \text{cov}(Z_t, Z_s)(b(t) + b(s)) + \text{Var} Z_t b(s)), \\ b(s) &= \text{cov}(Z'_t, Z_s), \\ \Delta^2 &= \text{Var} Z'_t - \frac{\text{Var} Z_s b(t)^2 - 2 \text{cov}(Z_t, Z_s) b(s) b(t) + \text{Var} Z_t b(s)^2}{\sigma^2}.\end{aligned}\quad (31)$$

Функции $g(s) = \text{Var} Z_s$, $h(s) = \text{cov}(Z_t, Z_s)$ дважды непрерывно дифференцируемы по s , а функция $v(s) = \text{cov}(Z'_t, Z_s)$ непрерывно дифференцируема один раз по s . Применим формулу Тейлора к этим функциям при $s \rightarrow t$:

$$\begin{aligned}v(s) &= v(t) + (t-s)v'(t) + o((t-s)), \\ g(s) &= g(t) + (t-s)g'(t) + \frac{(t-s)^2}{2}g''(t) + o((t-s)^2), \\ h(s) &= h(t) + (t-s)h'(t) + \frac{(t-s)^2}{2}h''(t) + o((t-s)^2).\end{aligned}\quad (32)$$

В силу условия (iii) интегралы

$$\int_{s \leq t} \frac{v'(s) - v'(t)}{s-t} ds, \int_{s \leq t} \frac{g''(s) - g''(t)}{(s-t)^2} ds, \int_{s \leq t} \frac{h''(s) - h''(t)}{(s-t)^2} ds$$

конечны. Интегрируя по частям, получим, что выражения

$$\begin{aligned}\frac{v(s) - v(t) - (t-s)v'(t)}{(t-s)^2}, \frac{g(s) - g(t) - (t-s)g'(t) - \frac{(t-s)^2}{2}g''(t)}{(t-s)^3}, \\ \frac{h(s) - h(t) - (t-s)h'(t) - \frac{(t-s)^2}{2}h''(t)}{(t-s)^3}\end{aligned}\quad (33)$$

также интегрируемы по s .

Подставим выражения (32) в определения функций m, Δ, σ^2 . Получим

$$\begin{aligned}m\sigma^2 &= o((t-s)^2), \\ \sigma^2 &= (v'(t)h(t) - v(t)^2)(s-t)^2 + o((s-t)^2), \\ \Delta^2\sigma^2 &= o((s-t)^2),\end{aligned}$$

где $s \rightarrow t$, причем в силу (33) выражения $m^2/\sigma, \Delta^2/\sigma$ интегрируемы по s .

Далее,

$$\int_0^\infty \frac{y^2}{\sigma} \varphi_{t,s}(y) dy < \int_{-\infty}^\infty \frac{y^2}{\sigma} \varphi_{t,s}(y) dy = \frac{\Delta^2}{\sigma} + \frac{um^2}{\Delta}.$$

Правая часть последнего равенства интегрируема по s . Более того, так как все функции $v(s), g(s), h(s)$ и их производные зависят от u полиномиальным образом,

то интеграл $\int_{t \geq s} \frac{\Delta^2}{\sigma} + \frac{um^2}{\Delta} ds$ зависит от u также полиномиальным образом.

Далее, отметим, что

$$\text{Var} Z_s + \text{Var} Z_t - 2 \text{cov}(Z_s, Z_t) = g(s) + g(t) - 2h(s),$$

$$\frac{g(s) + g(t) - 2h(s)}{g(s)g(t) - h^2(s)} = \frac{1}{g(t)} + \frac{(h(s) - h(t))^2}{(g(s)g(t) - h^2(s))g(t)}.$$

Подставляя разложения Тейлора для функций $h(s), \sigma^2$, получим

$$\frac{(h(s) - h(t))^2}{(g(s)g(t) - h^2(s))g(t)} \geq \frac{v^2(t)(s - t)^2}{4(v'(t)h(t) - v(t)^2)(s - t)^2}.$$

Отсюда

$$\exp \left\{ -u \frac{(h(s) - h(t))^2}{2(g(s)g(t) - h^2(s))g(t)} \right\} \leq \exp \left\{ -u \frac{v^2(t)}{4(v'(t)h(t) - v(t)^2)g(t)} \right\}. \quad (34)$$

Применим полученное соотношение (34) к двойному интегралу (30):

$$\begin{aligned} & \iint_{\Omega \cap \{t \geq s\}} \varphi_{t,s}(\sqrt{u}, \sqrt{u}) \int_0^\infty y^2 \varphi_{t,s}(y) dy ds dt = \\ &= \iint_{\Omega \cap \{t \geq s\}} \frac{1}{2\pi} \exp \left\{ -\frac{u}{2\sigma^2} (\text{Var} Z_s + \text{Var} Z_t - 2 \text{cov}(Z_s, Z_t)) \right\} \times \\ & \times \int_0^\infty \frac{y^2}{\sigma} \varphi_{t,s}(y) dy dt ds \leq \\ & \leq \iint_{\Omega \cap \{t \geq s\}} \frac{1}{2\pi} \exp \left\{ -\frac{u}{2g(t)} - u \frac{(h(s) - h(t))^2}{2(g(s)g(t) - h^2(s))g(t)} \right\} \times \\ & \times \int_0^\infty \frac{y^2}{\sigma} \varphi_{t,s}(y) dy dt ds \leq \\ & \leq \sup_{t \geq s} \exp \left\{ -u \frac{(h(s) - h(t))^2}{2(g(s)g(t) - h^2(s))g(t)} \right\} \times \\ & \times \iint_{\Omega \cap \{t \geq s\}} \frac{1}{2\pi} \exp \left\{ -\frac{u}{2g(t)} \right\} \int_0^\infty \frac{y^2}{\sigma} \varphi_{t,s}(y) dy ds dt. \end{aligned} \quad (35)$$

Проинтегрируем (35) по s и оценим оставшийся интеграл по t методом Лапласа:

$$\begin{aligned} & \iint_{\Omega \cap \{t \geq s\}} \frac{1}{2\pi} \exp \left\{ -\frac{u}{2g(t)} \right\} \int_0^\infty \frac{y^2}{\sigma} \varphi_{t,s}(y) dy dt ds \leq \\ & \leq \int_{\tilde{\alpha}}^{\tilde{\beta}} \exp \left\{ -\frac{u}{2g(t)} \right\} \int_0^t \int_0^\infty \frac{y^2}{\sigma} \varphi_{t,s}(y) dy ds dt \leq \\ & \leq \int_{\tilde{\alpha}}^{\tilde{\beta}} \int_0^\infty \frac{y^2}{\sigma} \varphi_{t,s}(y) dy ds \sqrt{2\pi u}^{-1/2} \sqrt{2g''(1)} \exp \left\{ -\frac{u}{2g(1)} \right\} (1 + o(1)). \end{aligned} \quad (36)$$

Из очевидного равенства $\frac{u}{2g(t)} = S_3(t)$ получаем, что $g''(1) = \frac{G}{c}(1 + o(1))$.

Так как (34) убывает экспоненциально, а интеграл $\int_{\tilde{\alpha}}^{\tilde{\beta}} \int_0^\infty \frac{y^2}{\sigma} \varphi_{t,s}(y) dy ds$ зависит от u полиномиальным образом, то из (36) следует соотношение $\mathbf{E}N_u([\alpha, \beta])(N_u([\alpha, \beta]) - 1) = o(\mathbf{E}N_u([\alpha, \beta]))$. В силу неравенства (20), а также неравенств (14,15,16) и асимптотических оценок (18,19) справедливы соотношения

$$\mathbf{P}(\tau_u < \infty) = \mathbf{E}N_u([\alpha, \beta]), \quad \mathbf{P}(\tau_u < \kappa(x)) = \mathbf{E}N_u([\alpha, \kappa(x)]).$$

Подставляя в (13) выражения (28) и (29) вместо соответственно $\mathbf{P}(\tau_u < \infty)$ и $\mathbf{P}(\tau_u < \kappa(x))$, получим утверждение теоремы.

Список литературы

- [1] *Cramer H.* Collective Risk Theory. Stockholm: Esselte, 1955.
- [2] *Asmussen S., Højgaard B.* Ruin probability approximations for Markov-modulated risk processes with heavy tails // *Theory Random Proc.* 1996. **2**. P. 96–107.
- [3] *Asmussen S.* Risk theory in a Markovian environment // *Scand. Act. J.* 1989. **89**. P. 69–100.
- [4] *Mikosch T., Samorodnitsky G.* The supremum of a negative drift random walk with dependent heavy-tailed steps // *Ann. Appl. Probab.* 2000. **10**. P. 1025–1064.
- [5] *Embrechts P., Veraverbeke N.* Estimates for the probability of ruin with special emphasis on the possibility of large claims // *Insurance: Mathematics and Economics.* 1982. **1**, № 1. P. 55–72.
- [6] *Kosten L.* Liquid models for a type of information buffer problem // *Delft Progress Report* 11. 1986. P. 71–86.
- [7] *Anick D., Mitra D., Sondhi M.* Stochastic theory of a data-handling system with multiple sources // *Bell System Techn. J.* 1982. **61**, № 8. P. 1871–1894.
- [8] *Palmowski Z., Rolski T.* The superposition of alternating on-off flows and a fluid model // *Ann. Appl. Probab.* 1998. **8**, № 2. P. 524–540.
- [9] *Simonian A.* Stationary analysis of a fluid queue with input rate varying as an Ornstein-Uhlenbeck process // *SIAM J. Appl. Math.* June 1991. **51**, № 3. P. 823–842.
- [10] *Kulkarni V., Rolski T.* Fluid model driven by an Ornstein-Uhlenbeck process // *Prob. Eng. Inf. Sci.* 1994. **8**. P. 403–417.
- [11] *Norros I.* A storage model with self-similar input // *Queueing Systems.* 1994. **16**. P. 387–396.
- [12] *Dieker A. B.* Extremes and fluid queues. Ph.D.Thesis. University of Amsterdam press, 2006. ISBN 90-5776-151-3.
- [13] *Debicki K.* Ruin probability for Gaussian integrated process // *Stochastic Processes and Appl.* 2002. **98**. P. 151–174.
- [14] *Кобельков С. Г.* О задаче разорения для гауссовского стационарного процесса // *Теория вероятностей и ее применения.* 2004. **49**, № 1. С. 171–178.
- [15] *Питербарг В. И.* Асимптотические методы в теории гауссовских случайных процессов и полей, М.: МГУ, 1988 (англ. испр. и доп. изд.: *Piterbarg V. I.* Asymptotic methods in the theory of Gaussian processes and fields. AMS, Providence, Rhode Island, 1996).
- [16] *Hüsler J., Piterbarg V.* Extremes of a certain class of Gaussian processes // *Stochastic Processes and Appl.* 1999. **83**. P. 257–271.
- [17] *Shao Qi-M.* Bounds and estimators of a basic constant in extreme value theory of Gaussian processes // *Statistica Sinica.* 1996. **6**. P. 245–257.
- [18] *Boulongne P., Pierre-Loti-Viaud D., Piterbarg V.* On average losses in the ruin problem with fractional Brownian motion as input // *Extremes*, 2009. DOI 10.1007/s10687-008-0069-z.

- [19] *Питербарг В.И.* Метод Райса для гауссовских случайных полей // Фунд. и прикл. матем. 1996. **2**, №1. С. 187–204.
- [20] *Федорюк М.В.* Метод перевала. М.: Наука, 1977.
- [21] *Narayan O.* Exact asymptotic queue length distribution for fractional Brownian traffic // Advances in Performance Analysis. 1998. **1**. P. 39–63.

ОСТРОВКИ МАТЕМАТИКИ В ДРУГИХ НАУКАХ (О НЕКОТОРЫХ ПОСЛЕДНИХ РАБОТАХ ЛАБОРАТОРИИ БОЛЬШИХ СЛУЧАЙНЫХ СИСТЕМ)

В. А. Малышев

Дается обзор нескольких последних работ лаборатории больших случайных систем.

1. Введение

Лаборатория больших случайных систем, созданная на механико-математическом факультете в 1992 году приказом ректора, занимается созданием и разработкой математических моделей явлений качественного характера в других науках. Скажем сначала несколько слов о философии выбора задач и тем для исследования.

Математика могла бы существовать сама по себе, с увлечением занимаясь магическими свойствами чисел, комбинаторными головоломками и постепенными обобщениями и уточнениями при построении архитектурных шедевров собственных областей. Такая независимость математики от жизни имеет, конечно, свои плюсы, но должна в конечном счете создавать состояние тоскливости. Поэтому естественны попытки математика вторгнуться в другие науки. В зависимости от возможностей и психики математика эти вторжения могут быть совершенно различны по своим целям. Либо он верит другим математикам, которые уже «знают, что надо делать», и берет задачи из их вполне математических статей. Либо он «идет на поклон» к тому же физику, «берет у него уравнение» и пробует исследовать его своими возможностями. Либо создает компьютерные программы.

Совершенно другой уровень получается, когда математик сам создает модели явлений, изначально мало связанных с математикой. Здесь за отсутствием экспериментальных знаний и опыта математику остается уповать на аксиомы и логику. Проблема всегда состоит в том, что аксиом как таковых почти никогда нет и их надо придумывать самому, исходя из уже аксиоматизированных областей и изучения литературы. При этом чтение литературы часто оказывается более полезным, чем разговор с живым представителем прикладной науки. Полная аксиоматизация даже такой математизированной науки, как физика, невозможна, и приходится ограничиваться островками, где возможна четкая логика, в море интуиции и расплывчатых формулировок.

Мы приведем здесь три примера. Они показывают, что первые две задачи представляют фундаментальный интерес, а третья, которая также интересна, но менее фундаментальна, должна решаться под конкретный заказ в рамках большого проекта.

2. О стохастических микромоделях твердого тела и жидкости

В шестидесятых годах прошлого века произошел фантастический всплеск математической статистической физики и квантовой теории поля. Это привело к аксиоматизации их равновесной части, основанной на понятии гиббсовского случайного поля, а также к математическим теоремам, сравнимым по сложности с

самыми глубокими математическими достижениями. Однако, что касается неравновесной части, дело обстоит значительно хуже. Более того, многие важнейшие куски даже школьного курса общей физики остались непонятыми на математическом (а некоторые и на физическом) уровне.

В моделях разреженного газа молекулы основное время проводят на достаточном расстоянии друг от друга, двигаясь свободно, а их взаимодействие сводится к столкновениям. Характер столкновений мало влияет на качественную картину макроскопического поведения, и часто молекулы считаются твердыми шариками. В классических моделях твердого тела, наоборот, молекулы совершают малые колебания около своих состояний равновесия, все время взаимодействуя с одними и теми же соседями. Так, в каждом учебнике по общей физике можно увидеть задачу о колебаниях однородной системы гармонических осцилляторов, однако вывода закона Гука из микромоделей ни в одном учебнике найти невозможно. Такой вывод на основе равновесных гиббсовских мер дан в [1] для простейшей одномерной модели, однако, многомерный случай пока открыт, а тем более неясно, как выглядит динамика расширения.

В качестве микромоделей жидкости можно представить себе уплотненный газ твердых шариков. Такое представление стало основным направлением в теоретической и математической физике, целью которого было вывести из такой микромоделей уравнения Навье–Стокса и успокоиться. Явный математический неуспех этой программы требует признания того, что сильные школы имеют и свои минусы, подавляя создание других моделей своим авторитетом.

Однако уже довольно давно в монографии [4], вышедшей впервые в 1943 году, был развит другой подход к моделированию жидкостей, где жидкость рассматривается как твердое тело с менее жесткой структурой. Эта монография, будучи физической, основана тем не менее на хорошей вероятностной интуиции и может служить основой самых разных математических моделей. Одна из таких моделей введена в [3]. Она состоит в следующем.

Пусть задан конечный или счетный граф G , например решетка, с множеством вершин $V = V(G)$ и множеством ребер $L = L(G)$. Конфигурацией называется функция $x_v, v \in V$, на множестве вершин со значениями в некотором множестве X , называемом пространством спинов или внутренних степеней свободы. На множестве конфигураций X^V задается марковский процесс с непрерывным временем. Переходы определяются так: для каждого ребра $l \in L(G)$ независимо от остальных ребер с интенсивностью $\lambda_l = \lambda_l(x_v, x_{v'})$ происходит одновременное изменение спинов x_v и $x_{v'}$, где v, v' — вершины ребра l :

$$(x_v(t), x_{v'}(t)) \rightarrow (x_v(t + dt), x_{v'}(t + dt)) = F(x_v(t), x_{v'}(t))$$

здесь $F : S = X \times X \rightarrow S = X \times X$ — некоторое отображение (бинарная реакция).

Частным случаем являются процессы Кавасаки, где пары точек обмениваются своими спинами, то есть F является перестановкой, а также еще более популярный процесс с запретами, где $X = \{0, 1\}$. Введенный процесс относится к процессам с локальным взаимодействием, которые играют в настоящий момент большую роль в построении физических моделей. Такие преобразования соответствуют преобразованию пар скоростей и, возможно, других степеней свободы при столкновениях

частиц или их химических реакциях. В работе [3] для конечных X классифицируются бернуллиевские инвариантные меры для таких цепей.

Пространство X даже для молекул воды НОН может быть довольно сложным. Простейшая классическая трактовка вводит как длины связей ОН, так и угол между ними, а квантово-механическая модель определяет даже некоторый тетраэдр, исходя из молекулярных орбиталей. Более того, как известно, между близко расположенными молекулами воды могут образовываться водородные связи, причем одна молекула может образовать до 4 таких связей. Они во многом обуславливают основные физические свойства воды, в том числе и структуру ее связей, которые интенсивно экспериментально исследуются в последнее время (см., например, [5]).

В работе [2] пространство X состоит из трех элементов, которые соответствуют отсутствию частицы или ее наличию с двумя возможными скоростями. Уже с таким бедным пространством X удастся моделировать фазовый переход от ламинарного (линейного) профиля скоростей к турбулентному (постоянному) в течении Куэтта.

3. Почему течет ток

Чтобы что-то двигалось, нужна движущая сила. Макроскопически ток определяется законом Ома

$$I = \frac{U}{R},$$

где U — потенциал (искомой силы). Имеется в виду большее: между концами отрезка проводника длины ΔL разность потенциалов равна

$$\Delta U = \frac{\Delta L}{L} U,$$

где L — длина всего проводника. Из однородности по длине проводника

$$U(x) = -Fx,$$

где F и есть искомая сила. Эта сила F во всех физических теориях (без объяснений) переносится из макро- на микрошкалу.

Первой микроскопической моделью электрического тока была модель Друде 1900 года, излагаемая во всех курсах физики твердого тела (см., например, [7]). Она основана на представлении об электрическом токе как о движении заряженных частиц, например электронов. В простейшей точной математической формулировке она может быть представлена так. На прямой в момент t в точках $x_i(t)$ находятся одинаковые частицы, динамика которых основана на следующих предположениях: частицы не взаимодействуют между собой, а в случайные моменты времени $t_{j,1} < t_{j,2} < \dots < t_{j,i} < \dots$ частица j останавливается по причине взаимодействия с внешней средой. При этом все $t_{j,i} - t_{j,i-1}$ независимы и одинаково распределены. В интервалах между столкновениями частица j разгоняется согласно закону Ньютона

$$m \frac{d^2 x_j}{dt^2} = F,$$

где сила F является внешней, то есть не зависит от расположения $x_i(t)$ других частиц.

Закон больших чисел говорит, что при широких предположениях относительно распределений случайных времен для каждой частицы с вероятностью единица существует предел

$$\lim_{t \rightarrow \infty} \frac{x_j(t)}{t} = v = \frac{F\tau}{2m}, \tau = \langle t_{j,i} - t_{j,i-1} \rangle.$$

Если частицы одинаковы и плотность ρ частиц постоянна, то в среднем за единицу времени через данную точку прямой будет проходить $v\rho = \frac{F\tau\rho}{2m}$ частиц, что дает закон Ома $U = FL, I = v\rho, R = \frac{2mL}{\tau\rho}$.

Таким образом, модель Друде — это модель трения (Эйнштейн, Стокс и др.), которая при определенном скейлинге даст закон Ньютона с трением

$$m \frac{d^2 x(t)}{dt^2} = F - A \frac{dx}{dt}.$$

Над обобщением **тормозящей** силы потрудились математики (около 15 статей в журнале «Communications in Math. Physics») — от независимости к (асимптотически слабой) зависимости от прошлого и более общим случайным средам. Основная же проблема состоит, конечно, в том, что никакого ускоряющего внешнего поля или внешней силы нет на пассивной части проводника (то есть вне генератора, батареи и т.д.). Нельзя сказать, что она совсем осталась без внимания. Конечно, есть и было понимание, что сила идет от самих зарядов, но вопрос «КАК?» оставался без ответа даже на физическом уровне строгости. Этому есть ряд причин:

1) простота, наглядность и достаточность модели Друде для прикладных целей. Вычислительно и удобно, и естественно считать макросилу также и микросилой;

2) физика электрического тока в начале XX века сразу ушла в квантовую область, где сконцентрировалась на других проблемах. Зоммерфельд добавил другое распределение по скоростям, Блох добавил внешний периодический потенциал (от кристаллической решетки), в ферми-жидкостях Ландау вместо модели свободных электронов возникла модель свободных или слабо взаимодействующих квазичастиц, Андерсон добавил почти периодический потенциал.

3) общая тенденция физики не обращать внимания на логическую структуру теорий.

Однако имеется долгая дискуссия в «American Journal of Physics» насчет этого, откуда полезно привести цитату «Common students misconception — drifting electrons push each other through a wire just like water molecules push each other through a pipe (despite charge neutrality inside the metal)». Однако студенты, возможно, правы — нейтральность, конечно, имеет место на макрошкале, но не на микрошкале, а тем более на еще более мелкой шкале, определяемой ниже.

Введенная в [6] модель максимально упрощена: она не квантовая, нет кристаллической решетки, бесконечно тонкий провод. Однако положительный фон зарядов для макронейтральности, которого нет в модели, можно ввести без потери результата.

В момент 0 на интервале $[0, L)$ есть $N(L)$ частиц в точках

$$0 = x_{N(L)}(0) < x_{N(L)-1}(0) < \dots < x_1(0) < L.$$

Энергия этой конфигурации частиц

$$U = \sum_{\langle i, i-1 \rangle} V(x_i - x_{i-1}).$$

где суммирование ведется по парам ближайших соседей (чисто техническое предположение). Потенциал предполагается отталкивающим (около нуля):

$$V(x) = V(-x) > 0$$

с ростом в нуле:

$$f(r) = -\frac{dV(r)}{dr} \sim c_1 r^{-a}, a \geq 2, r = |x| \rightarrow 0$$

и убыванием в бесконечности: $V(x) \rightarrow 0$ при $x \rightarrow \infty$.

Динамика $x_i(t) = x_i^{(N)}(t)$ дается системой N зацепляющихся уравнений Ньютона с диссипацией энергии

$$\frac{d^2 x_i}{dt^2} = -\frac{\partial U}{\partial x_i} - A \frac{dx_i}{dt}, i = 1, \dots, N \quad (1)$$

(массы $m = 1$). Внешней силы нет, но есть сила взаимодействия $F_i = -\frac{\partial U}{\partial x_i}$ и граничные условия на концах пассивного участка. Как возможно движение без внешней силы? Попробуем расположить частицы так, чтобы на каждую действовала сила, направленная вправо. Для этого достаточно (и необходимо), чтобы для всех i

$$x_i - x_{i+1} < x_{i-1} - x_i$$

Тогда на каждую частицу действуют силы: с каждой из сторон порядка N^a , а результирующая сила

$$G_i^{(N)}(t) = f(x_i(t) - x_{i+1}(t)) - f(x_{i-1}(t) - x_i(t)) > 0$$

должна быть порядка 1. Она соответствует ускоряющей силе Друде.

Но надо соблюсти еще много требований. Основные из них:

1) почему эта картина сохраняется во времени?

2) почему плотность, а также скорости частиц макроскопически (то есть при $N \rightarrow \infty$) однородны?

В модели есть параметр N (число частиц) и 3 шкалы относительно этого параметра:

1) первая шкала — макрошкала — порядка 1 — скорости частиц;

2) вторая шкала — микрошкала (обратная плотность) — порядка $\frac{1}{N}$ — расстояния между соседними частицами

$$x_{i+1}(t) - x_i(t);$$

3) третья шкала — тонкая (субмикро) шкала — порядка $\frac{1}{N^2}$ — разности между последовательными расстояниями

$$\delta_i = x_{i-1}(t) + x_{i+1}(t) - 2x_i(t),$$

которую надо подбирать так, чтобы общая сила была порядка 1. Сила,двигающая ток, возникает именно на тонкой шкале.

Перейдем к строгим формулировкам. Данная система $N(L)$ уравнений имеет единственное решение, но оно зависит от начальной конфигурации частиц и от выбранных граничных условий. Построено простейшее семейство явных решений этой системы.

Решение называется квазистационарным (то есть стремящимся к стационарному для любого $-\infty < t < \infty$ при $N \rightarrow \infty$), если оно обладает следующими свойствами:

1) оно периодически с периодом $s = \frac{1}{N}$, где $N = \frac{N(L)}{L}$ — плотность, то есть

$$x_i(s) = x_{i+1}(0), \quad i = 1, \dots, N(L) - 1,$$

при этом новые частицы входят (вбрасываются) в интервал в моменты, кратные s , а крайне правые частицы покидают интервал в точке L в моменты, кратные s ;

2) частицы движутся только направо;

3) (граничные условия) есть внешние силы, которые действуют на крайне левую и крайне правую частицы так, что

$$F_{\text{left}}(x_{N(L)}(t)) > |f(x_{N(L)-1}(t) - x_{N(L)}(t))|$$

$$F_{\text{right}}(x_1(t)) < |f(x_1(t) - x_2(t))|$$

Теорема. Для любого потенциала V и любых граничных условий из введенного класса, существует $L_0 > 0$, такая, что для любого $L < L_0$ и для любой достаточно большой плотности N существует квазистационарное решение. При этом при $N \rightarrow \infty$ имеют место свойства МАКРОскопической однородности по времени и по пространству без МИКРОоднородности: действующие на каждую частицу i силы

$$G_i^{(N)}(t) \rightarrow \text{const} > 0,$$

скорости всех частиц

$$v_i^{(N)}(t) \rightarrow \text{const} > 0,$$

однородность плотности частиц: для любого t и любого подынтервала $l \subset [0, L]$

$$\frac{N(l, t)}{N} \rightarrow \frac{|l|}{L}, \quad N = N([0, L], t),$$

где $N(l, t)$ — число частиц в интервале I в момент t .

Схема доказательства и построения решения — сведение системы N уравнений к системе двух уравнений: линейному ОДУ и нелинейному функциональному уравнению.

Будем искать гладкую возрастающую функцию $x(t)$ на интервале времени $[0, T_L]$ для некоторого $T_L > 0$, такую, что $x(0) = 0$, $x(T_L) = L$, и удовлетворяющую одновременно двум уравнениям: линейному ОДУ

$$\frac{d^2x}{dt^2} = G(t) - A \frac{dx}{dt} \quad (2)$$

и нелинейному функциональному уравнению

$$f(x(t) - x(t-s)) - f(x(t+s) - x(t)) = G(t), t \in [0, T_L].$$

Тогда

$$x_k(t) = x(t + (k-1)s), t \in (0, s), k = 1, \dots, N(L), \quad (3)$$

будет искомым решением. Это означает, что соседи отстоят друг от друга на равные времена (но НЕ на равные расстояния).

Линейное ОДУ на $[0, T_L]$ с начальными условиями

$$x(0) = 0, \frac{dx}{dt}(0) = v_0 = A^{-1}w \quad (4)$$

может быть решено явно, если считать «силу Друде» $G(t)$ заданной в виде

$$G(t) = w + \nu t + g(t)$$

для некоторой константы $w > 0$, имеющей смысл базисной постоянной силы, которая движет частицы вправо (аналог силы Друде). «Голая» функция νt с малым параметром

$$\nu = \nu(N) = s^{a-2} a^{-1} A^{-a-1} w^{a+2}$$

определяет слабый рост этой силы в некотором интервале времени порядка $O(1)$. Поправочный член $g(t)$, необходимый для удовлетворения уравнений, ищется постановкой в функциональное уравнение и его решением. Полное доказательство см. в [6].

4. Пробки на автотрассах

Пробки бывают не только в городах, но и на автотрассах, за МКАД. Анализ движения автомобилей по автотрассам более доступен точному математическому анализу, чем по переплетениям московских улиц. Есть два подхода к такому анализу: макроподход с уравнениями в частных производных, подобными уравнениям для жидкости (см., например, [10]), и микроподход, где есть много элементарных единиц — машин. Мы предложим одну модель микроподхода, показывающую, какие параметры играют существенную роль.

Пусть сначала по трассе с одной полосой в одном направлении движется машина, которая представляется точкой $x(t)$ на прямой R . В точках

$$\dots < Y_k < Y_{k+1} < \dots$$

прямой стоят светофоры. На каждом светофоре k в детерминированные моменты времени $\tau_{k,n}$ зеленый свет включается, а в $\sigma_{k,n}$ выключается, причем

$$\dots < \tau_{k,n} < \sigma_{k,n} < \tau_{k,n+1} < \dots$$

Движение машины описывается так. Постояв перед светофором, машина набирает скорость по определенному закону $v(t)$ и далее движется с постоянной, максимально разрешенной скоростью $V_{\max}$ до следующего красного света. Задача состоит в вычислении средней скорости автомобиля на большом промежутке времени при разумных предположениях о множестве имеющихся параметров. Параметров действительно много: длительности красного и зеленого света для каждого из светофоров и сдвиги по времени для разных светофоров, расстояния между светофорами, $V_{\max}$ и закон набора скорости. Предположение состоит в том, что ответ будет мало зависеть от дисперсий этих величин, а в основном от их средних.

Рассмотрим теперь поток автомашин, где будем представлять машины точками

$$\dots < x_i(t) < x_{i-1}(t) < \dots$$

на прямой R . При этом добавляются следующие параметры:

1) расстояние между последовательными точками не может быть менее некоторого числа d , которое можно считать, например, средней длиной машин;

2) учитывается запаздывание s начала движения каждого следующего водителя при начале движения на зеленый свет.

При этом содержательна даже задача с одним светофором. Если на трассу поступает слева поток машин с постоянной плотностью $\rho < d^{-1}$, то аналогично задачам теории массового обслуживания, существует критическая плотность $\rho_{\text{кр}}$, выше которой образующаяся перед светофором пробка растет. Но, как мы знаем из опыта, пробки постепенно рассасываются и снова возникают. Это может быть только по причине переменной плотности $\rho(t)$ машин в течение дня. Это единственный макропараметр в задаче. Все остальные — микропараметры, то есть на гораздо меньшей временной шкале. Ситуация вполне аналогична возникающей при выводе уравнений в частных производных из микродинамики.

Одна такая задача о динамике роста пробки была решена в [11].

Список литературы

- [1] *Malyshev V. A.* One-dimensional mechanical networks and crystals // Moscow Math. J. 2006. **6**, № 2. P. 263–268.
- [2] *Малышев В. А., Манита А. Д.* Стохастическая микромодель течения Куэтта // Теория вероятностей и ее применения. 2008. **53**, № 4. С. 798–809.
- [3] *Malyshev V. A., Zamyatin A. A.* Exchange processes with a local interaction: invariant Bernoulli measures // Markov Processes and Related Fields. 2009. **15**, № 1. P. 125–133.
- [4] *Френкель Я.* Кинетическая теория жидкостей. Л.: Наука, 1975.
- [5] *Moro R., Rabinovitch R., Xia Ch, Kresin V.* Electric dipole moments of water clusters from a beam deflection measurement // Phys. Rev. Lett. 2006. **97**. № 123401.
- [6] *Малышев В. А.* Почему течет ток: многочастичная одномерная модель // Теор. и матем. физ. 2008. **155**, № 2. P. 301–311.
- [7] *Ашкрофт Н., Мермин Н.* Физика твердого тела. Т. 1, 2. М.: Мир, 1979.
- [8] *Jackson J.* Surface charges on circuit wires and resistors play three roles // Amer. J. Phys. 1996. **64**, № 7. P. 855–870.

- [9] *Preyer N.* Transient behaviour of simple RC circuits // Amer. J. Phys. 2002. **70**. P. 1187–1193.
- [10] *Prigogine I., Herman R.* Kinetic Theory of Vehicular Traffic. Elsevier, 1971.
- [11] *Замятин А. А., Малышев В. А.* Накопление на границе для одномерной стохастической системы частиц // Проблемы передачи информации. 2008. **43**, № 4. P. 68–82.

О СТАБИЛЬНОСТИ ПРОЦЕССОВ АДсорбЦИИ

С. Е. Волков, В. В. Щербakov

В этой работе мы анонсируем и обсуждаем результаты об асимптотическом поведении случайного процесса, моделирующего процессы адсорбции в физике. Рассматриваемый процесс можно также интерпретировать как некоторую обобщенную урновую схему Пойа.

1. Введение

Рассмотрим целочисленный отрезок $D_N = [1, \dots, N + 1]$ с периодическими граничными условиями для некоторого фиксированного $N \geq 2$. Рассмотрим последовательность $(X_t \in D_N, t \geq 1)$ случайных величин, таких, что X_1 равномерно распределена в D_N и условное распределение X_{t+1} относительно $X(t) = (X_1, \dots, X_t), t \geq 1$, задается следующей формулой:

$$\mathbb{P}\{X_{t+1} = i | X(t)\} = \frac{\beta^{V_i(t)}}{\sum_{j=1}^{N+1} \beta^{V_j(t)}}, \quad i \in D_N,$$

где

$$\begin{aligned} V_i(t) &= \sum_{j \in U_i} \xi_j(t), \quad i = 1, \dots, N + 1, \\ \xi_i(t) &= \sum_{s=1}^t 1_{\{X_s=i\}}, \quad i = 1, 2, \dots, N + 1, \end{aligned} \tag{1}$$

и где U_i — это некоторая окрестность точки $i \in D_N$. В этой работе мы рассматриваем три следующих варианта окрестности U_i :

$$(\mathbf{A1}) : U_i = \{i\}; (\mathbf{A2}) : U_i = \{i, i + 1\}; (\mathbf{A3}) : U_i = \{i - 1, i, i + 1\}.$$

Случайный процесс $(X_t, t \geq 0)$ является одним из вариантов вероятностных моделей процессов адсорбции. Узлы решетки рассматриваются как дискретная аппроксимация непрерывной поверхности некоторого материала. Случайная точка X_t интерпретируется как точка поверхности, где адсорбирована t -я частица. Одно из основных свойств физических моделей адсорбции состоит в том, что адсорбированные частицы меняют поглощающие свойства поверхности. В терминах вероятностной модели это означает, что вероятностное распределение положения следующей адсорбируемой частицы зависит от пространственной конфигурации, образованной положениями прежде адсорбированных частиц. Для более детального описания моделей адсорбции и обзора соответствующей литературы мы отсылаем к работе [1] (см. также [2]). Следует отметить, что строгий анализ физических моделей процессов адсорбции приводит к постановке содержательных математических задач [3, 4]. Кроме того, аналогичные вероятностные модели могут быть использованы в других естественно-научных приложениях (например, таких, как геофизика и экология) для моделирования пространственных временных рядов [5, 6].

Рассмотрим процесс роста $\xi(t) = (\xi_1(t), \dots, \xi_{N+1}(t))$, компоненты которого определены уравнением (1). Нетрудно видеть, что $(\xi(t), t \in \mathbb{Z}_+)$ является марковской цепью со значениями в $\mathbb{Z}_+^{N+1}$ со следующими переходными вероятностями:

$$\mathbb{P}\{\xi_i(t+1) = \xi_i(t) + 1 | \xi(t)\} = \mathbb{P}\{X_{t+1} = i | \xi(t)\} = \frac{\beta^{V_i(t)}}{\sum_{j=1}^{N+1} \beta^{V_j(t)}}.$$

В этой заметке мы интересуемся таким свойством процесса роста, как стабильность, которое на интуитивном уровне можно объяснить следующим образом. Очевидно, что какие-то компоненты процесса роста должны расти неограниченно, когда число размещаемых частиц стремится к бесконечности. И под стабильностью процесса роста мы понимаем ситуацию, когда все компоненты процесса растут примерно с одинаковой скоростью. Процесс роста нестабилен, если мы наблюдаем неравномерный рост компонент процесса. Ниже мы формализуем понятие стабильности и приводим полученные нами результаты, кратко обсуждая основные идеи доказательств и возможные усиления некоторых утверждений. За деталями доказательств мы отсылаем к работе [7].

Следует заметить, что в случае (A1) наша модель является частным вариантом урновой схемы Пойя, для которой некоторые интересующие нас результаты известны (см. ниже). Также в случаях (A2) и (A3) модель может быть интерпретирована как урновая схема Пойя с взаимодействием.

2. Определение стабильности

Для того чтобы формализовать интуитивное понятие стабильности процесса роста, рассмотрим следующий процесс разностей $\zeta_t = (\zeta_1(t), \dots, \zeta_N(t)) \in \mathbb{Z}^N$, $t \in \mathbb{Z}_+$, где

$$\zeta_i(t) = \xi_i(t) - \xi_{N+1}(t), \quad i = 1, \dots, N.$$

Заметим, что во всех трех указанных случаях (A1), (A2) и (A3) процесс $(\zeta_t, t \in \mathbb{Z}_+)$ также является марковской цепью со следующими переходными вероятностями:

$$\mathbb{P}\{\zeta_i(t+1) = \zeta_i(t) + 1 | \zeta(t)\} = \frac{\beta^{\sum_{j \in U_i} \zeta_j(t)}}{Z(\zeta_t)}$$

и

$$\mathbb{P}\{\zeta_i(t+1) = \zeta_i(t) - 1, \forall i \in D_N | \zeta(t)\} = \frac{1}{Z(\zeta(t))},$$

где

$$Z(\zeta(t)) = \sum_{j=1}^N \beta^{\sum_{j \in U_i} \zeta_j(t)}.$$

В терминах процесса разностей можно дать следующее строгое определение стабильности.

Определение 1. Назовем процесс роста стабильным, если соответствующий процесс разностей является эргодической цепью Маркова. В противном случае процесс роста нестабилен.

3. Модель без взаимодействия

Теорема 1. Пусть $U_i = \{i\}$, $i = 1, \dots, N + 1$.

1) Если $0 < \beta < 1$, то марковская цепь $(\zeta(t), t \in \mathbb{Z}_+)$ эргодична.

2) Если $\beta > 1$, то марковская цепь $(\zeta(t), t \in \mathbb{Z}_+)$ транзиентна.

Если $U_i = \{i\}$, то, как было отмечено выше, соответствующая модель является частным случаем хорошо изученной урновой схемы Пойя с $N + 1$ типом шаров [8, 9]. В урновой схеме Пойя предполагается, что в урне могут находиться шары $N + 1$ типа, в момент времени t из урны случайным образом извлекается шар, причем вероятность вынуть шар типа i пропорциональна некоторой функции $w(\xi_i(t))$, где $\xi_i(t)$ — это число шаров типа i в урне в момент t . В рассматриваемом нами случае $w(x) = \beta^x$.

Если $\beta > 1$, то $\sum_{k=1}^{\infty} w(k)^{-1} < \infty$. В этом случае, как хорошо известно (см., например, [9]),

$$\mathbb{P} \left(\bigcup_{i=1}^{N+1} A_i \right) = 1,$$

где

$$A_i = \left\{ \lim_{t \rightarrow \infty} \xi_i(t) = \infty, \sup_t \xi_j(t) < \infty \text{ для всех } j \neq i \right\}, \quad i = 1, \dots, N + 1.$$

То есть с вероятностью 1 начиная с какого-то момента времени из урны вынимаются шары только одного типа. В терминах процесса роста это означает, что начиная с некоторого (случайного) момента времени будет расти только одна из компонент процесса. Нетрудно видеть, что поскольку $\zeta_i(t) = \xi_i(t) - \xi_{N+1}(t)$, то на событии A_i , $i = 1, 2, \dots, N$, $\zeta_i(t) \rightarrow \infty$, и на событии A_{N+1} все $\zeta_i(t) \rightarrow -\infty$. Таким образом, мы получаем, что процесс разностей транзиентен, то есть процесс роста нестабилен.

Если $\beta < 1$, то снова используя интерпретацию процесса роста в терминах урновой схемы Пойя и замечая, что $\sum_{k=1}^{\infty} w(k)^{-1} = \infty$, мы можем применить известный для этого случая результат [9], который состоит в том, что с вероятностью 1 число извлеченных шаров любого типа стремится к бесконечности. Таким образом, в случае $\beta < 1$ все компоненты процесса роста возрастают бесконечно. Наш результат усиливает это утверждение. А именно мы доказываем, что все компоненты процесса растут к бесконечности *согласованным* образом в том смысле, что распределение процесса разностей сходится к стационарному распределению (как это следует из его эргодичности). Мы доказываем эргодичность процесса разностей, используя следующий классический критерий Фостера [10].

Теорема 2 (Фостер). Счетная марковская цепь η_k , $k \in \mathbb{Z}_+$, является эргодической тогда и только тогда, когда существуют конечное множество состояний M , положительное число $\varepsilon > 0$ и положительная функция $f(\eta_k)$, такие, что

$$\mathbb{E} [f(\eta_{k+1}) - f(\eta_k) \mid \eta_k = v] \leq -\varepsilon \quad \text{для всех } v \notin M;$$

$$\mathbb{E} [f(\eta_{k+1}) \mid \eta_k = v] < \infty \quad \text{для всех } v.$$

Покажем, что в качестве функции Ляпунова может быть взята квадратическая функция

$$f(\mathbf{x}) = \sum_{i=1}^N x_i^2, \quad \mathbf{x} = (x_1, \dots, x_N) \in \mathbb{Z}^N,$$

ε может быть любым положительным числом, а множество M может быть выбрано следующим образом:

$$M := \left\{ \sum_{i=1}^N |x_i| \leq CN + \varepsilon \right\},$$

где C есть некоторая (явно вычисляемая) константа. Положим

$$\Delta := \mathbb{E}(f(\zeta(t+1)) - f(\zeta(t)) \mid \zeta_t = \mathbf{x}) = \frac{\sum_{i=1}^N (2x_i(\beta^{x_i} - 1) + 1 + \beta^{x_i})}{1 + \sum_{i=1}^N \beta^{x_i}}.$$

Тогда для любого фиксированного $\varepsilon > 0$

$$\Delta + \varepsilon = \frac{\varepsilon - \sum_{i=1}^N (2x_i(1 - \beta^{x_i}) - 1 - (\varepsilon + 1)\beta^{x_i})}{1 + \sum_{i=1}^N \beta^{x_i}}.$$

Рассмотрим следующую функцию: $h(x) = 2x(1 - \beta^x) - 1 - (\varepsilon + 1)\beta^x$. Нетрудно видеть, что существует такое $A' > 0$, что для $x \geq A'$ выполняется неравенство $h(x) \geq x$. Также если $x = -a < 0$, то

$$h(x) = h(-a) = \frac{1}{\beta^a}(2a - \varepsilon - 1) - 1 - 2a.$$

Таким образом, функция $h(-a)$ растет приблизительно экспоненциально, следовательно, существует такое $A'' > 0$, что $h(-a) \geq a$, когда $a \geq A''$. Выберем

$$-C := \min_{-A'' < x < A'} (h(x) - |x|) = \inf_{x \in \mathbb{R}} (h(x) - |x|)$$

(заметим, что $\varepsilon + 2 \leq C < \infty$). Тогда

$$\begin{aligned} \sum_{i=1}^N (2x_i(1 - \beta^{x_i}) - 1 - (\varepsilon + 1)\beta^{x_i}) &= \sum_{i=1}^N h(x_i) = \sum_{i=1}^N |x_i| + \sum_{i=1}^N (h(x_i) - |x_i|) \geq \\ &\geq \sum_{i=1}^N |x_i| - CN \end{aligned}$$

и, следовательно,

$$\Delta + \varepsilon \leq \frac{CN + \varepsilon - \sum_{i=1}^N |x_i|}{1 + \sum_{i=1}^N \beta^{x_i}} \leq 0,$$

кроме, быть может, конечного множества состояний $(x_1, \dots, x_N)$, принадлежащих множеству

$$M := \left\{ \sum_{i=1}^N |x_i| \leq CN + \varepsilon \right\}.$$

Таким образом, условия критерия Фостера выполнены и марковская цепь $(\zeta(t), t \in \mathbb{Z}_+)$ эргодична.

Следует также заметить, что выполнение условия Фостера влечет также экспоненциальную сходимость к стационарному распределению (см. [10]).

Модель с $\beta = 1$. Если $\beta = 1$, то это специальный и вырожденный в очевидном смысле случай, когда на каждом шаге частицы размещаются по узлам решетки равновероятно. Нетрудно видеть, что процесс разностей является специальным случаем однородного случайного блуждания с ограниченными скачками и нулевым сносом. Такое блуждание возвратно (но не эргодично) в размерностях 1, 2 и транзитивно в размерностях 3 и выше.

4. Модель с несимметричным взаимодействием

Теорема 3. Пусть $U_i = \{i, i + 1\}$, $i = 1, \dots, N + 1$.

1) Если $N = 2$ и $\beta < 1$, то марковская цепь $(\zeta(t), t \in \mathbb{Z}_+)$ эргодична.

2) Если $\beta > 1$, то марковская цепь $(\zeta(t), t \in \mathbb{Z}_+)$ транзитивна.

В случае $N = 2$ процесс разностей $\zeta(t) = (\zeta_1(t), \zeta_1(t))$ является марковской цепью на $\mathbb{Z}^2$ со следующими переходными вероятностями:

$$\begin{aligned}\mathbb{P}\{\zeta(t+1) = (x+1, y) | \zeta(t) = (x, y)\} &= \frac{\beta^{x+y}}{\beta^{x+y} + \beta^y + \beta^x}, \\ \mathbb{P}\{\zeta(t+1) = (x, y+1) | \zeta(t) = (x, y)\} &= \frac{\beta^y}{\beta^{x+y} + \beta^y + \beta^x}, \\ \mathbb{P}\{\zeta(t+1) = (x-1, y-1) | \zeta(t) = (x, y)\} &= \frac{\beta^x}{\beta^{x+y} + \beta^y + \beta^x}.\end{aligned}$$

Можно показать, что в случае $\beta < 1, N = 2$ для Марковской цепи $\zeta(t)$ критерий Фостера выполняется со следующей функцией Ляпунова:

$$f(\mathbf{x}) = \beta^{1-x_1-x_2} + \beta^{-3x_1+x_2} + \beta^{3x_1-4x_2} + \beta^{x_1+4x_2}. \quad (2)$$

Не останавливаясь на деталях вычислений, заметим, что эргодичность процесса разностей становится интуитивно понятной уже после достаточно грубого анализа асимптотического поведения векторного поля процесса вдали от начала координат. И сравнительно несложно предъявить кусочно-линейную функцию Ляпунова, например такую:

$$\tilde{f}(\mathbf{x}) = \max \left\{ \frac{x_1 + x_2}{10}, \frac{8x_1 - 3x_2}{25}, \frac{8x_1 - 7x_2}{20}, \frac{-x_1 - 3x_2}{7} \right\},$$

однако проверка критерия сводится в этом случае к очень большому перебору вариантов начальной точки в вычислении среднего скачка $\tilde{f}(\zeta(t))$. Конструкция же гладкой функции Ляпунова (2) совсем не очевидна. Следует также заметить, что траектории рассматриваемой марковской цепи ведут себя асимптотически приблизительно как траектории следующего обыкновенного дифференциального уравнения:

$$\frac{dx_2}{dx_1} = \frac{\beta^{x_2} - \beta^{x_1}}{\beta^{x_1+x_2} - \beta^{x_1}}.$$

Применяя подстановку $u = \beta^{-x_1}$, $v = \beta^{-x_2}$, это уравнение можно переписать в следующем виде:

$$\frac{dv}{du} = \frac{v(u-v)}{u(1-v)}, \quad (3)$$

Аналитическое решение этого уравнения нам неизвестно. Траектории решения, полученные с помощью численных вычислений, ведут себя как спирали, скручивающиеся в точку $(1, 1)$ с возрастанием времени (рис. 1). Транзиентность в

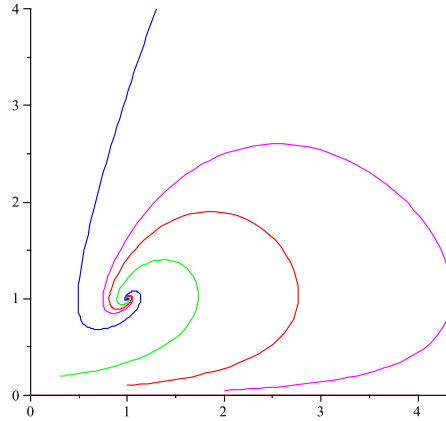


Рис. 1. Траектории динамической системы (3)

случае несимметричного взаимодействия и $\beta > 1$, $N \geq 2$ доказывается следующим образом. Рассмотрим Марковскую цепь $\eta(t) = (\eta_1(t), \dots, \eta_{N+1}(t))$, где $\eta_i(t) = \xi_i(t) - \xi_{i-1}(t)$, $i = 1, \dots, N+1$, с пространством состояний $\{\mathbf{x} = (x_1, \dots, x_{N+1}) : x_1 + \dots + x_{N+1} = 0\}$. Нетрудно видеть, что транзиентность $(\eta(t), t \in \mathbb{Z}_+)$ влечет транзиентность процесса разностей $(\zeta(t), t \in \mathbb{Z}_+)$. Действительно, если предположить, что марковская цепь $(\zeta(t), t \in \mathbb{Z}_+)$ возвратна, то $\xi_1(t) = \dots = \xi_{N+1}(t)$ почти наверное для бесконечно многих t . Но это означало бы, что тогда $\eta_1(t) = \dots = \eta_{N+1}(t) = 0$ почти наверное для бесконечно многих t , что противоречит транзиентности процесса η . Транзиентность же марковской цепи $(\eta(t), t \in \mathbb{Z}_+)$ доказывается применением следующего критерия [10].

Теорема 4 (критерий транзиентности). *Счетная цепь Маркова η_k , $k \in \mathbb{Z}_+$, транзиентна тогда и только тогда, когда существуют такое множество состояний $M \subset \mathcal{A}$ и положительная функция $f(\eta_k)$, такие, что*

$$\mathbb{E} [f(\eta_{k+1}) - f(\eta_k) \mid \eta_k = v] \leq 0 \quad \text{для всех } v \notin M;$$

$$f(v_1) < \inf_{v \in M} f(v), \quad \text{по крайней мере для одного } v_1 \notin M.$$

Можно показать, что условия этой теоремы выполняются с функцией

$$f(\mathbf{x}) = \begin{cases} 1, & \text{если } x_1 = x_2 = \dots = x_{N+1} = 0; \\ \frac{1}{\max\{|x_1|, |x_2|, \dots, |x_{N+1}|\}}, & \text{в противном случае.} \end{cases}$$

и множеством M , которое есть некоторая окрестность начала координат.

5. Модель с симметричным взаимодействием

Теорема 5. Если $U_i = \{i-1, i, i+1\}$, $i = 1, \dots, N+1$, и $N \geq 3$, то марковская цепь $(\zeta(t), t \in \mathbb{Z}_+)$ транзиентна при любых значениях $0 < \beta < 1$ или $\beta > 1$.

В случае $0 < \beta < 1$ мы показываем, что если $\xi_1(T) = \xi_2(T) = \dots = \xi_{N+1}(T)$ в некоторый момент времени T , то с положительной вероятностью, не зависящей от T , происходит следующее событие:

$$A := \left\{ \lim_{t \rightarrow \infty} \xi_i(t)/t = 1/M \text{ для четных } i \text{ и } \sup_{t \geq T} \xi_i(t) = \xi_i(T) \text{ для нечетных } i \right\},$$

что очевидным образом означает транзиентность процесса разностей ζ . Опуская технические детали (см. [7]), этот факт можно объяснить следующим образом. Предположим, что $T = 0$ и $\xi_i(T) = 0$ для всех i . При условии, что частицы поглощаются только в четных узлах решетки, процесс роста в четных узлах ведет себя как процесс роста в модели без взаимодействия (см. теорему 1). Поэтому для достаточно больших времен t мы должны наблюдать $\xi_i(t) \approx t/M$ для четных i . Типичная наблюдаемая конфигурация в этом случае представлена на рис. 2.

Случай $N+1 = 2M+1$ анализируется аналогично, только сценарий развития с периодическими максимумами и минимумами высот нужно несколько изменить. А именно в этом случае мы доказываем, что следующее событие происходит с положительной вероятностью:

$$A := \left\{ \lim_{t \rightarrow \infty} \frac{\xi_1(t)}{t} = \lim_{t \rightarrow \infty} \frac{\xi_2(t)}{t} = \frac{1}{2M}, \lim_{t \rightarrow \infty} \frac{\xi_i(t)}{t} = \frac{1}{M} \text{ для четных } i \geq 4 \right. \\ \left. \text{и } \sup_{t \geq T} \xi_i(t) = \xi_i(T) \text{ для нечетных } i \geq 3 \right\},$$

что и влечет транзиентность процесса ζ . Вероятностная техника получения оценок заимствована из работы [11].

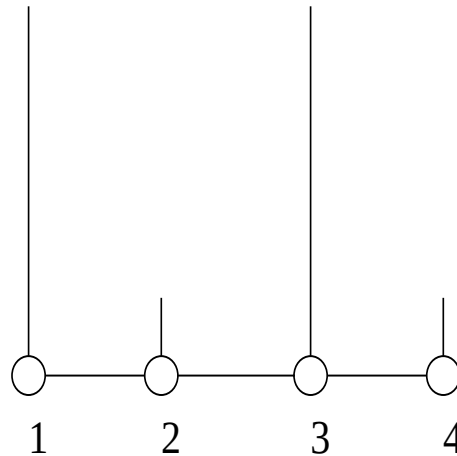


Рис. 2. Типичная наблюдаемая конфигурация ($N = 3$, $0 < \beta < 1$)

В случае $\beta > 1$ транзиентность процесса разностей вытекает из транзиентности следующего процесса разностей: $v(t) = (v_1(t), \dots, v_N(t))$, где $v_k(t) = \xi_{k-1}(t) + \xi_k(t) + \xi_{k+1}(t) - (\xi_N(t) + \xi_{N+1}(t) + \xi_1(t))$, $k = 1, \dots, N$. Нетрудно убедиться,

что процесс $(v(t), t \in \mathbb{Z}_+)$ является цепью Маркова и его транзиентность влечет транзиентность процесса разностей координат. Доказательство основано на применении уже упоминавшегося критерия транзиентности. А именно если $N \geq 4$, то для произвольного $a > 1$ определим следующее подмножество решетки $\mathbb{Z}^N$:

$$\mathcal{C}_a = \{\mathbf{y} \in \mathbb{Z}^N : y_1 > a, y_2 - y_k > a, k = 4, \dots, N\}$$

и рассмотрим функцию

$$g(\mathbf{y}) = \beta^{-y_1} + \sum_{k \neq 1, 2, 3} \beta^{-y_2 + y_k}.$$

Тогда если $a > 1$, то для любого $\mathbf{y} \in \mathcal{C}_a$

$$\mathbb{E}(g(v(t+1)) - g(v(t)) | v(t) = \mathbf{y}) \leq 0.$$

Если $N = 3$, то для проверки критерия транзиентности можно взять множество $\mathcal{D}_a = \{\mathbf{y} \in \mathbb{Z}^3 : y_2 > a\}$ и функцию $g(\mathbf{y}) = \beta^{-y_2}$.

Для модели с симметричным взаимодействием и $\beta > 1$ справедлив более сильный результат, а именно следующая теорема, дающая асимптотическое описание всех возможных траекторий процесса роста.

Теорема 6. Если $U_i = \{i-1, i, i+1\}$, $i = 1, \dots, N+1$, $N \geq 3$, и $\beta > 1$, то с вероятностью 1 существует $k \in \{1, \dots, N+1\}$, такое, что $\lim_{t \rightarrow \infty} \xi_i(t) = \infty$, если и только если $i \in \{k-1, k\}$ и

$$\lim_{t \rightarrow \infty} \frac{\xi_k(t)}{\xi_{k-1}(t)} = \beta^c,$$

где $c = \lim_{t \rightarrow \infty} [\xi_{k+1}(t) - \xi_{k-2}(t)] \in \mathbb{Z}$.

Список литературы

- [1] Evans J. W. Random and cooperative sequential adsorption // Rev. Modern Phys. 1993. **65**, № 4. P. 1281–1329.
- [2] A special issue of Colloids and Surfaces A. Vol. 165 / Ed. by V. Privman. 2000.
- [3] Penrose M. D., Yukich J. E. Limit theory for random sequential packing and deposition // Ann. Appl. Probab. 2002. **12**, № 1. P. 272–301.
- [4] Shcherbakov V. Limit theorems for random point measures generated by cooperative sequential adsorption // J. Statist. Phys. 2006. **124**, № 6. P. 1425–1441.
- [5] Shcherbakov V. On a model of sequential point patterns // Annals of Institute of Statistical Mathematics. 2007. DOI:10.1007/s10463-007-0147-z. <http://www.springerlink.com/content/a4ln86435j2pj4x2/>
- [6] Penrose M. D., Shcherbakov V. Maximum likelihood estimation for cooperative sequential adsorption. Preprint 4/08 of Bath Institute for Complex Systems. 2008 (сдано в печать).
- [7] Shcherbakov V., Volkov S. Stability of a growth process generated by monomer filling with cooperative effects. Препринт. 2008 (готовится к печати).

- [8] *Pemantle R.* A survey of random processes with reinforcement // Probability Surveys. 2007. **4**. P. 1-79.
- [9] *Davis B.* Reinforced random walk // Probability Theory and Related Fields. 1990. **84**. P. 203–229.
- [10] *Fayolle G., Malyshev V. A., Menshikov M. V.* Topics in the Constructive Theory of Countable Markov Chains. Cambridge University Press, 1995.
- [11] *Volkov S.* Vertex-reinforced random walk on arbitrary graphs // Ann. Probab. 2001. **29**, № 1. P. 66–91.

On the nonstandard problems of the stochastic optimization: reduction to the problems with Markov representation and their solution

Shiryayev A.N., Professor, Faculty of Mechanics and Mathematics MSU,
e-mail: albertsh@mi.ras.ru

The paper deals with the stochastic problems where optimization consists of the finding a rational time («stopping time») to make an optimal control decision.

From point of view of the «general theory of optimal stopping», the problems formulated in the paper have nonstandard formulation. We show how these problems can be reduced to the standard ones using the idea of the markov representation and how these new problems can be solved on base of the general theory of optimal stopping.

Limit theorem for queueing systems in the heavy traffic situation

Afanas'eva L.G., Professor, Faculty of Mechanics and Mathematics MSU,
e-mail: afanas@mech.math.msu.su

Bashtova E.E., Assistant, Faculty of Mechanics and Mathematics MSU,
e-mail: bashtovaelena@rambler.ru

The first part of the paper deals with queueing systems with a large number of service channels and input flow of high intensity. The convergence of the distribution of the number of clients to the Gaussian one is proved. Convergence conditions are given in terms of moments of the input flow, which is of a rather general character. In the second part a single-server queue in the asymptotic of compression time is considered. The diffusion approximation for the workload process in the heavy traffic situation is established. A regenerative process is regarded as the input flow.

Some problems for interacting particles flows

Afanas'eva L.G., Professor, Faculty of Mechanics and Mathematics MSU,
e-mail: afanas@mech.math.msu.su

Bulinskaya E.V., Professor, Faculty of Mechanics and Mathematics MSU,
e-mail: ebulinsk@mech.math.msu.su

At first we investigate the density of the flow of two types of particles moving with different velocities. The upper and lower bounds for this density are established and an algorithm is proposed for the density calculation with a given precision. In the second part of the paper the points of particles delay on a highway (traffic lights) are taken into account. The optimization of the traffic lights performance is also treated.

Stochastic models in radiobiology

Bulinski A.V., Professor, Faculty of Mechanics and Mathematics MSU,
e-mail: bulinski@mech.math.msu.su

Starting from the 90-s of the 20-th century, one actively uses in radiobiology the models based on the concept of independent behavior of functional subunits constituting an organ or a tissue. Along with the considering such models we develop a more general approach employing various forms of dependence for families of random variables describing the impact on cells and elements of biological system. Limit theorems for sums of dependent multiindexed random variables play here an essential role. Some versions of combining the classical «critical volume» model and the cluster models involving the geometrical aspects of location of damaged tissues and organs is discussed as well.

On a several problems of the risk theory

Vinogradov O.P., Professor, Faculty of Mechanics and Mathematics MSU,
e-mail: ovinogradov@mail.ru

The connection between the risk theory, the theory of the branching processes and the ballot problem is established. It makes possible to transfer the theorems from one of these domains of the probability theory to another. For example the classical Cramer-Lundberg theorem states that the Cramer condition is sufficient for asymptotic decrease of the ruin probability. One of the theorems of the branching processes allows us to state that this condition is also necessary. A method of getting inequalities for several characteristics is proposed. The notion of inhomogeneous process is introduced and some explicit formulas are derived. Randomizing parameters in these formulas one may receive results for a rather broad class of risk processes.

Maximal branching processes

Lebedev A.V., Associate Professor,
Faculty of Mechanics and Mathematics MSU,
e-mail: alebedev@mech.math.msu.su

In the paper we summarize the author's results on the maximal branching processes theory. We point out some applications to gated infinite-server queues. We consider to the stationary distributions of ergodic maximal branching processes and their limit properties.

Multivariate statistical models in geometric presentation

Tyurin Yu.N., Professor, Faculty of Mechanics and Mathematics MSU,
e-mail: ytjurin@gmail.com

Multivariate random variables we write as columns. Sets (sequences) of multivariate random variables we represent as rows positions of which are occupied by the columns. It give us the array. In the space of arrays we introduce the inner product. Results of such products are square matrix. This multiplication we interpret and use as a scalar multiplication. In particular,

we define with this product the orthogonality of arrays, linear subspaces in the space of arrays etc. In the frame of his concept in the space of arrays with random elements we linear models and linear hypotheses, a matrix method of least squares, the method of testing linear hypotheses etc. This new theory of multivariate statistical analysis goes in parallel with the classical and well-known theory of linear statistical analysis.

About research of branching random walks on multidimensional lattices

Yarovaya E.B., Associate Professor,
Faculty of Mechanics and Mathematics MSU,
e-mail: yarovaya@mech.math.msu.su

The paper discusses two models of continuous-time random walk on $\mathbb{Z}^d$ with birth and death of particles at a single site (i.e., the source of branching). The main difference between the models lies in the behavior of the process at the source. Model I. assumes that the random walk is symmetrical. In Model II. it is entered an additional parameter enabling *artificial* intensification of prevalence of branching or walk at the source and, as the result, violating symmetry of the random walk. In Model II the differential equations for the generating functions of the numbers of particles both at an arbitrary point of the lattice and on the entire lattice were established. Sufficient conditions for exponential growth of the process are established. General methods to study the models in the supercritical case are proposed.

Limit theorems in ruin models with Gaussian losses

Boulongne P., Associate Professor, Université Paris 8, 2, rue de la Liberté,
Saint-Denis, France, e-mail: pbp88@free.fr

Kobelkov S.G., Researcher, Laboratory of Probability Theory,
Faculty of Mechanics and Mathematics MSU,
e-mail: sergey@dodo.inm.ras.ru

Piterbarg V.I., Professor, Laboratory of Probability Theory,
Faculty of Mechanics and Mathematics MSU,
e-mail: piter@mech.math.msu.su

An actuarial-financial model with high initial wealth, linear incomes and Gaussian losses are considered. Two models of Gaussian losses are studied, the fractional Brownian motion (FBM) and an integral of Gaussian stationary process. Asymptotic behaviors of ruin probabilities are given. Limit theorems for the time of ruin are proven. In the case of FBM the asymptotic behavior of expected losses in case of ruin is studied.

**Mathematical islands in other sciences
(review of some recent results
in the laboratory of large random systems)**

Malyshev V.A., Professor, Laboratory of Large Random Systems,
Faculty of Mechanics and Mathematics MSU,
e-mail: malyshev@mech.math.msu.su

We give review of some recent results in the Laboratory of Large Random Systems

On stability of adsorption processes

Volkov S.E., Reader, Department of Mathematics, University of Bristol,
United Kingdom,
e-mail: S.Volkov@bristol.ac.uk

Shcherbakov V.V., Researcher, Laboratory of Large Random Systems,
Faculty of Mechanics and Mathematics MSU,
e-mail: VV.Shcherbakov@gmail.com

In this paper we announce and discuss results on asymptotic behaviour of a random growth process modelling adsorption processes in physics. The process can also be interpreted as a variant of a generalised Polya urn scheme with interaction.

Содержание

Раздел 1. Исследования кафедры теории вероятностей	5
Введение	5
<i>Ширяев А. Н.</i> О нестандартных проблемах стохастической оптимизации: редукция к задачам в марковском представлении и их решение . . .	8
<i>Афанасьева Л. Г., Баштова Е. Е.</i> Предельные теоремы для систем массового обслуживания в условиях высокой загрузки	40
<i>Афанасьева Л. Г., Булинская Е. В.</i> Некоторые задачи для потоков взаимодействующих частиц	55
<i>Булинский А. В.</i> Стохастические модели в радиобиологии	68
<i>Виноградов О. П.</i> О некоторых задачах теории риска	85
<i>Лебедев А. В.</i> Максимальные ветвящиеся процессы	93
<i>Тюрин Ю. Н.</i> Многомерные статистические модели в геометрическом изложении	107
<i>Яровая Е. Б.</i> Об исследовании ветвящихся случайных блужданий по многомерным решеткам	119
Раздел 2. Исследования лабораторий	137
Введение	137
<i>Булонь П., Кобельков С. Г., Питербарг В. И.</i> Предельные теоремы в моделях разорения с гауссовскими убытками	138
<i>Мальшев В. А.</i> Островки математики в других науках (о некоторых последних работах лаборатории больших случайных систем)	157
<i>Волков С. Е., Щербаков В. В.</i> О стабильности процессов адсорбции	166
Аннотации статей	175

Научное издание

СОВРЕМЕННЫЕ ПРОБЛЕМЫ МАТЕМАТИКИ И МЕХАНИКИ

Том IV. Математика

Выпуск 3. Теория вероятностей и математическая статистика /

Под редакцией чл.-корр. РАН, профессора А.Н. Ширяева

М.: Изд-во Моск. ун-та, 180 с.

*Оригинал-макет изготовлен издательской группой механико-математического
факультета МГУ*

Редактор Н.А. Леонтьева

Подписано в печать 27.04.2009

Формат 60 x 90 /8 Бумага офс. № 1. Усл. печ. л. 18, 0.

Заказ Тираж 150 экз.

Ордена «Знак Почета» Издательство Московского университета
125009, Москва, ул. Б.Никитская, 5/7

Отпечатано на типографском оборудовании механико-математического
факультета

119992, Москва, Ленинские горы.