

ИНФОРМАТИКА И ЕЁ ПРИМЕНЕНИЯ

**Научный журнал Отделения нанотехнологий
и информационных технологий Российской академии наук**

Издается с 2007 года
Журнал выходит ежеквартально

Учредители:
Российская академия наук
Институт проблем информатики Российской академии наук

РЕДАКЦИОННАЯ КОЛЛЕГИЯ

академик С. В. Емельянов (главный редактор, член Редакционного совета)
академик Ю. И. Журавлев (председатель Редакционного совета)
академик С. К. Коровин
академик Г. И. Савин
академик И. А. Соколов (зам. главного редактора, член Редакционного совета)
академик А. Л. Стемпковский
академик Ю. И. Шокин (член Редакционного совета)
член-корреспондент РАН В. Л. Арлазаров
член-корреспондент РАН А. Б. Жижченко
член-корреспондент РАН И. А. Каляев
член-корреспондент РАН Ю. С. Попков
член-корреспондент РАН К. В. Рудаков
член-корреспондент РАН Ю. А. Флеров
член-корреспондент РАН Б. Н. Четверушкин
член-корреспондент РАН Р. М. Юсупов
профессор, д.т.н. В. И. Будзко
профессор, д.т.н. А. А. Зацаринный
профессор, д.ф.-м.н. В. Ю. Королёв
профессор, д.ф.-м.н. А. В. Печинкин
профессор, д.т.н. И. Н. Синицын
профессор, д.ф.-м.н. С. Я. Шоргин (ответственный секретарь)

Редакция

профессор, д.г.-м.н. Р. Б. Сейфуль-Мулюков;
к.ф.-м.н. Е. Н. Арутюнов;
О. В. Ломакина

© Институт проблем информатики Российской академии наук, 2009

Адрес редакции:

Москва 119333, ул. Вавилова 44, корп. 2, ИПИ РАН,
редакция журнала «Информатика и её применения»
Тел. 8(499)135-86-92, e-mail rust@ipiran.ru

**Журнал «Информатика и её применения» включен в «Перечень ведущих
рецензируемых научных журналов и изданий, в которых должны быть опубликованы
основные научные результаты диссертации на соискание ученой степени
доктора и кандидата наук», утвержденный ВАК**

Журнал зарегистрирован в Федеральной службе по надзору
в сфере связи и массовых коммуникаций 31 марта 2009 г.
Свидетельство о регистрации средства массовой информации ПИ № ФС77-35837
Подписной индекс журнала в каталоге «Пресса России» 88018 (годовая подписка)

Информатика и её применения

Том 3 Выпуск 3 Год 2009

Тематический выпуск

*Вероятностно-статистические методы
и задачи информатики и информационных технологий*

СОДЕРЖАНИЕ

Предисловие И. А. Соколов, В. Ю. Королёв	2
Многолинейная система массового обслуживания с групповым отказом приборов А. В. Печинкин, И. А. Соколов, В. В. Чаплыгин	4
О предельных характеристиках системы обслуживания $M(t)/M(t)/S$ с катастрофами А. И. Зейфман, Я. А. Сатин, А. В. Коротышева, Н. А. Терёшина	16
Асимптотики вероятностей больших уклонений стационарной очереди Е. В. Морозов	23
Вероятностная модель обслуживания трафика в системе сетевидного типа В. Ю. Бородакий	35
An approach to actuarial modeling with Quasi-Monte Carlo: Simulation of random sums depending on stochastic factors G. Temnov, S. Kucherenko	40
Об устойчивости реконструкции изображений в задачах эмиссионной томографии О. В. Шестаков	47
О вероятности исправления ошибок при помехоустойчивом кодировании, когда число ошибок принадлежит некоторому конечному множеству А. Н. Чупрунов, Б. И. Хамдеев	52
О распределении размеров частиц при дроблении В. Ю. Королёв	60
Некоторые оценки для характеристических функций с применением к уточнению неравенства Мизеса И. Г. Шевцова	69
О мощности критериев в случае обобщенного распределения Лапласа В. Е. Бенинг, О. О. Лямин	79
Abstracts	86
Об авторах	89
About Authors	90

Предисловие

Вниманию читателей журнала «Информатика и её применения» предлагается очередной тематический выпуск «Вероятностно-статистические методы и задачи информатики и информационных технологий». Первый тематический выпуск журнала по данному направлению вышел в 2008 году (т. 2, вып. 2).

Статьи, собранные в данном журнале, посвящены разработке новых вероятностно-статистических методов, ориентированных на применение к решению конкретных задач информатики и информационных технологий, а также — в ряде случаев — и других прикладных задач. Проблематика, охватываемая публикуемыми работами, развивается в рамках научного сотрудничества между Институтом проблем информатики Российской академии наук (ИПИ РАН) и Факультетом вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова, в частности в ходе работ над совместными научными проектами. Многие из авторов статей, включенных в данный номер журнала, являются активными участниками традиционного международного семинара по проблемам устойчивости стохастических моделей, руководимого В. М. Золотарёвым и В. Ю. Королёвым; ежегодные сессии этого семинара проводятся под эгидой МГУ и ИПИ РАН.

Наряду с представителями ИПИ РАН и МГУ, в число авторов данного выпуска журнала входят ученые из Карельского научного центра РАН, Национального исследовательского ядерного университета «МИФИ», Вологодского государственного педагогического университета, Казанского государственного университета, University College (Корк, Ирландия) и Centre for Process Systems Engineering (Лондон, Великобритания).

Ряд статей журнала посвящен тематике, весьма активно разрабатываемой в течение многих лет специалистами ИПИ РАН и МГУ, — математическим методам моделирования информационно-телекоммуникационных систем, основанных на теории массового обслуживания и теории телетрафика.

В статье А. В. Печинкина, И. А. Соколова и В. В. Чаплыгина рассматривается важная для приложений задача анализа многолинейной системы массового обслуживания с ненадежными приборами. Предложены методы расчета стационарного распределения числа заявок в системе при различных вариантах функционирования системы.

Статья А. И. Зейфмана, Я. А. Сатина, А. В. Коротышевой и Н. А. Тершиной посвящена изучению предельных характеристик системы обслуживания с катастрофами в предположении, что интенсивности катастроф зависят от числа требований в системе. Получены достаточные условия слабой эргодичности процесса, описывающего число требований в системе, и соответствующие оценки.

В работе Е. В. Морозова исследуются асимптотики вероятностей больших отклонений стационарной очереди в системах обслуживания для различных классов распределений времен обслуживания.

В статье В. Ю. Бородакия на основе теории телетрафика построена модель сетцентрической системы; проведен анализ модели отдельного звена такой системы, получены формулы для вычисления вероятности блокировки запроса из-за отсутствия достаточной для передачи блока данных ширины полосы пропускания.

Следующая группа статей выпуска посвящена разработке и применению стохастических методов для решения ряда прикладных задач.

Статья Г. Темнова (Ирландия) и С. Кучеренко (Великобритания), публикуемая на английском языке, посвящена моделированию случайных сумм применительно к одной задаче актуарной математики.

В работе О. В. Шестакова развиваются вероятностные методы анализа устойчивости реконструкции изображений в задачах эмиссионной томографии.

Статья А. Н. Чупрунова и Б. И. Хамдеева посвящена асимптотическому анализу вероятности исправления ошибок при помехоустойчивом кодировании.

Заключительная группа статей посвящена развитию перспективных теоретических вероятностно-статистических методов, которые могут найти широкое применение в различных задачах информатики и информационных технологий.

В статье В. Ю. Королёва рассматриваются и развиваются математические модели, описывающие распределение физических размеров частиц при дроблении. Подобные модели могут довольно успешно использоваться (и используются) при описании самых разных объектов, в том числе объемов сообщений в вычислительных или телекоммуникационных системах, капиталов фирм, размеров доходов и др.

Работа И. Г. Шевцовой посвящена уточнению оценок для характеристических функций — основного аппарата, используемого при построении конкретных числовых оценок точности нормальной аппроксимации (такая аппроксимация весьма важна при математическом моделировании и анализе многих реальных систем, в том числе информационных и телекоммуникационных).

В статье В. Е. Бенинга и О. О. Лямина рассмотрены вопросы проверки гипотез о параметрах обобщенного распределения Лапласа, которое находит широкое применение при математическом моделировании многих процессов в телекоммуникационных системах, экономике, финансовом деле, технике и других областях. Получена формула для предела отклонения мощности асимптотически оптимального критерия от мощности наилучшего критерия в случае обобщенного распределения Лапласа.

Редакционная коллегия журнала выражает надежду, что данный тематический выпуск будет интересен специалистам в области теории вероятностей и математической статистики и их применения к решению задач информатики и информационных технологий.

Заместитель главного редактора журнала «Информатика и её применения»,
директор ИПИ РАН, академик

И. А. Соколов

Редактор-составитель тематического выпуска,
профессор кафедры математической статистики
факультета вычислительной математики и кибернетики МГУ им. М. В. Ломоносова,
ведущий научный сотрудник ИПИ РАН,
доктор физико-математических наук

В. Ю. Королёв

МНОГОЛИНЕЙНАЯ СИСТЕМА МАССОВОГО ОБСЛУЖИВАНИЯ С ГРУППОВЫМ ОТКАЗОМ ПРИБОРОВ*

А. В. Печинкин¹, И. А. Соколов², В. В. Чаплыгин³

Аннотация: Рассматривается многолинейная система массового обслуживания (СМО) $SM/PH/n/r$ ($r \leq \infty$) с ненадежными приборами, отказывающимися группами. Отказы и восстановления групп приборов происходят с постоянной интенсивностью, число отказывающихся одновременно приборов является случайной величиной, а заявки с прерванным обслуживанием после восстановления прибора начинают обслуживаться заново. Предложены методы расчета стационарного распределения числа заявок в системе при различных вариантах функционирования системы.

Ключевые слова: многолинейные системы массового обслуживания; ненадежные приборы; отказ и восстановление случайного числа приборов

1 Введение

Системам массового обслуживания с ненадежными приборами уже более полувека уделяется значительное внимание. Среди публикаций по этой тематике подавляющее число работ посвящено однолинейным СМО [1–48] и, в частности, системе типа $M/G/1$ и различным ее обобщениям, когда в системе находится один обслуживающий прибор, поступающий поток заявок является пуассоновским (в некоторых случаях марковским или даже групповым марковским), время обслуживания распределено по произвольному закону, причем времена работы прибора в исправном и неисправном состояниях могут иметь различные распределения, а сами отказы могут оказывать самое разнообразное воздействие на нормальное функционирование системы. Так, например, в [33, 35, 37] рассмотрена однолинейная СМО в стационарном и нестационарном режимах с пуассоновским входящим потоком, гиперэрланговским потоком отказов прибора и временами обслуживания и восстановления, распределенными по произвольному закону. Работа [15] посвящена исследованию однолинейной СМО $M/G/1$ с пуассоновским входящим потоком, интенсивность которого зависит от того, в каком (исправном или неисправном) состоянии находится обслуживающий прибор, время между отказами прибора распределено по экспоненциальному закону, а время обслуживания и время восстановления прибора независимы и имеют произвольные функции распределения.

Весьма популярны модели, в которых вследствие ремонта сервера образуется очередь повтор-

ных заявок. В частности, экспоненциальные системы типа $M/M/1$ с ненадежным прибором и повторными заявками рассмотрены в [38, 39, 46]. Системы типа $M/G/1$ с ненадежным прибором и повторными заявками рассмотрены в [11, 25, 27, 28, 40, 42, 47].

Однолинейная система с групповым пуассоновским потоком, рекуррентным обслуживанием, отказами обслуживающего прибора и повторными заявками рассмотрена в [48]. В [44] рассматривается однолинейная СМО $VMAP/G/1$ с групповым марковским потоком, в которой время жизни (lifetime) сервера распределено по экспоненциальному закону, а время восстановления — уже по произвольному закону. Ненадежные однолинейные СМО с групповым марковским потоком, рекуррентным и полумарковским обслуживанием исследовались соответственно в [36] и [29].

В работах [6, 9, 12, 17, 23, 30] рассмотрены системы с ненадежными приборами и упорядоченным опросом нескольких очередей (поллингом), которые получили широкое распространение в рамках развития теоретических основ проектирования широкополосных беспроводных сетей. Подробную библиографию по системам поллинга и, в частности, по системам поллинга с ненадежными приборами, можно найти в обзорной статье [49] и в монографии [50].

Системы массового обслуживания с ненадежными приборами в дискретном времени рассмотрены в [2, 20, 41, 43].

В гораздо меньшей степени исследовались СМО с несколькими ненадежными приборами и сети

* Работа выполнена при поддержке РФФИ (гранты 08-07-00152 и 09-07-12032-офи-м).

¹ Институт проблем информатики Российской академии наук, apchinkin@ipiran.ru

² Институт проблем информатики Российской академии наук, isokolov@ipiran.ru

³ Институт проблем информатики Российской академии наук, VasilyChaplygin@mail.ru

массового обслуживания с ненадежными серверами. В [51–54] рассмотрены различные варианты экспоненциальных систем с ненадежными серверами. В [55] рассмотрена многолинейная СМО $M/G/n/c$ с ненадежными приборами в стационарном и нестационарном режимах, заявки в которой имеют ограниченное распределенное по экспоненциальному закону время ожидания в системе. В [56] исследована модель с двумя ненадежными экспоненциально отказывающимися и восстанавливающимися серверами, разделенными конечным буфером. В [57, 58] рассмотрена сеть массового обслуживания с одним экспоненциальным источником, образованная параллельными однолинейными экспоненциальными СМО.

Заметим, что кроме перечисленных работ исследования по СМО с ненадежными приборами можно также найти в работах [59–72].

В указанных выше публикациях многолинейные системы с полумарковским потоком заявок, обслуживанием фазового типа на каждом приборе и отказами приборов не рассматривались. Настоящая работа посвящена исследованию многолинейной системы $SM/PH/n/r$ с ненадежными приборами, отказы которых могут происходить группами. В [73–75] рассмотрены различные варианты функционирования СМО $SM/PH/n/r$ с независимыми и одновременными отказами приборов.

Поскольку дальнейшие наши исследования будут опираться на базовые модели СМО $SM/MSP/n/r$ с конечным и бесконечным накопителем, рассмотренные в [74, 76], приведем их краткое описание. Модели отличаются только емкостью накопителя (конечная или бесконечная), поэтому будем рассматривать одну модель, отмечая при необходимости различия.

Всюду в дальнейшем будем обозначать через $\vec{1}$ вектор-столбец из единиц, размерность которого определяется либо из контекста, либо нижним индексом.

Полумарковский входящий поток заявок определяется полумарковским процессом с конечным множеством состояний $\{1, 2, \dots, I\}$, $1 \leq I < \infty$, поведение которого описывается квадратной матрицей $A(x)$. Элемент $A_{ij}(x)$, $i, j = \overline{1, I}$, этой матрицы представляет собой вероятность того, что первый переход из состояния i полумарковский процесс совершит в состояние j и произойдет этот переход за время меньше x . Предполагается, что среднее время a между поступлениями заявок в стационарном режиме функционирования системы удовлетворяет условию

$$a = \vec{\pi}_a \int_0^{\infty} x dA(x) \vec{1} < \infty,$$

где $\vec{\pi}_a$ — вектор-строка стационарных вероятностей цепи Маркова с матрицей $A = A(\infty)$ переходных вероятностей (вложенной цепи Маркова). Более подробное описание полумарковского входящего потока, а также некоторые естественные дополнительные предположения относительно параметров, которые также будут предполагаться выполненными, приведены в [76].

Перейдем теперь к марковскому процессу обслуживания. Если в системе имеется k , $0 \leq k \leq n+r$, заявок (в этом случае также будем говорить, что процесс обслуживания находится на одной из l_k , $l_k < \infty$, фаз обслуживания, причем интенсивность изменения фазы марковского процесса равна соответствующему элементу матрицы Λ_k , $k = \overline{0, n+r}$, если ни одна заявка не обслужилась, и матрицы N_k , $k = \overline{1, n+r}$, если одна из заявок обслужилась. Предполагается, что $l_k = l$ при $k = \overline{n, n+r}$, матрицы $\Lambda_k = \Lambda$ совпадают при $k = \overline{n, n+r}$, матрицы $N_k = N$ совпадают при $k = \overline{n+1, n+r}$, матрица $\Lambda + N$ является неразложимой, а матрица N — ненулевой. Кроме того, если в момент поступления очередной заявки в системе находится k , $k = \overline{0, n-1}$, заявок, то вероятность изменения фазы обслуживания определяется элементом матрицы Ω_k , а если не меньше n заявок, то процесс обслуживания просто переходит на следующий слой с сохранением фазы обслуживания. Обозначим через $\mu = \vec{\pi}_s N \vec{1}$, где $\vec{\pi}_s$ — вектор-строка стационарных вероятностей марковского процесса с инфинитезимальной матрицей $\Lambda + N$, стационарную интенсивность обслуживания при бесконечной очереди. В случае накопителя бесконечной емкости предполагается, что $\rho < 1$, где $\rho = (a\mu)^{-1}$ — нагрузка на систему. Если же накопитель имеет конечную емкость r , то поступающая заявка теряется тогда, когда застает в системе $n+r$ заявок (т. е. процесс обслуживания на слое $n+r$), не изменяя при этом фазы обслуживания. Более подробное описание марковского процесса обслуживания, а также способ формирования его инфинитезимальной матрицы по заданному РН-распределению обслуживания заявки на каждом приборе можно найти, например, в [76, 77].

2 Описание системы

Рассмотрим многолинейную СМО $SM/PH/n/r$ ($r \leq \infty$) с полумарковским входящим потоком заявок, распределением фазового типа времени обслуживания каждой заявки, накопителем конечной или бесконечной емкости и ненадежными приборами.

Распределение фазового типа (РН-распределение) времени обслуживания заявки с числом фаз обслуживания J , $1 \leq J < \infty$, описывается квадратной матрицей H порядка J с элементами h_{ij} , $i, j = \overline{1, J}$, и вектор-строкой \vec{h} размерности J с элементами h_i , $i = \overline{1, J}$. Функцию распределения фазового типа времени обслуживания заявки можно записать в виде

$$H(x) = 1 - \vec{h} e^{Hx} \vec{1}.$$

Далее понадобится также вектор $\vec{h}^* = -H\vec{1}$, координатой h_i^* которого является интенсивность окончания обслуживания заявки при фазе обслуживания i . Более подробное описание распределения фазового типа и его свойств можно найти в [78], а применительно к распределению времени обслуживания в системах с ненадежными приборами — в [73–75].

Каждый из n имеющихся в системе однотипных приборов может находиться либо в исправном, либо в неисправном состоянии. Состояние прибора будем считать исправным, если на приборе находится заявка и прибор занят ее обслуживанием или если прибор свободен, готов принять заявку и немедленно начать ее обслуживание. Состояние прибора будем считать неисправным, если на приборе находится заявка, но прибор ее не обслуживает, или если на приборе нет заявки, но он не может немедленно начать обслуживание заявки, если таковая на него поступает. Если при отказе прибора (переходе прибора из исправного состояния в неисправное) на нем находится заявка, то она остается на приборе до момента восстановления (перехода прибора из неисправного состояния в исправное) и затем обслуживается заново.

Будем называть прибор занятым, если на нем находится заявка, и свободным — в противном случае. Если в некоторый момент времени в систему поступает очередная заявка, но ни один прибор не может принять ее на обслуживание, то эта заявка попадает в накопитель, становясь в очередь на обслуживание (в случае конечной емкости накопителя заявка становится в очередь при наличии в накопителе свободных мест или теряется при их отсутствии). Заявки из очереди на обслуживание выбираются в порядке их поступления в накопитель.

В следующих разделах будут рассмотрены три модели СМО с ненадежными приборами, которые отказывают и восстанавливаются группами случайного размера. Будет показано, как эти модели можно привести к базовым, что позволит вычислять стационарные вероятности состояний с помощью ранее разработанных алгоритмов. Поскольку вся

сложность такого подхода заключается в замене процесса обслуживания СМО с ненадежными приборами марковским процессом обслуживания базовых моделей, то ограничимся только построением матриц Λ_k , $k = \overline{1, n-1}$, Λ , N_k , $k = \overline{1, n}$, и N (матрица Λ_0 имеет порядок 1 и состоит из одного элемента — 0).

В настоящей работе, так же как и в [73–75], вместо линейной нумерации состояний процесса обслуживания будем использовать мультииндексную нумерацию, при которой номер состояния определяется мультииндексом (i_1, \dots, i_r) или объединением таких мультииндексов, где каждому числу k заявок в системе соответствует одно или несколько значений r , а $i_1, \dots, i_r = \overline{1, J}$ (напомним, что J — число фаз РН-распределения времени обслуживания заявки), причем в случае $k \geq n$ подмножества мультииндексов совпадают. Перейти от мультииндексной нумерации состояний процесса обслуживания к линейной можно различными способами. В частности, рациональный способ такого перехода (на примере СМО $SM/PH/n/\infty$ с надежными приборами) использован в [76, 77].

3 Отказы только занятых приборов

Первая рассматриваемая здесь СМО с ненадежными приборами характеризуется следующими особенностями:

- (1) если в некоторый момент времени среди n приборов ровно k , $k = \overline{1, n}$, приборов исправны и заняты обслуживанием заявок, то за «малое» время Δ с вероятностью $\alpha_i(k)\Delta + o(\Delta)$, $i = \overline{1, k}$, произойдет отказ ровно i приборов с заявками;
- (2) если в некоторый момент времени среди n приборов ровно k , $k = \overline{1, n}$, приборов с заявками неисправны, то за «малое» время Δ с вероятностью $\beta_i(k)\Delta + o(\Delta)$, $i = \overline{1, k}$, произойдет восстановление ровно i приборов с заявками;
- (3) при отказе или восстановлении i приборов из k возможных ($1 \leq i \leq k$) все варианты выбора i конкретных приборов из k равновероятны (вероятность выбора каждого варианта $q_k^i = 1/C_k^i$, где C_k^i — число сочетаний из k элементов по i);
- (4) свободные приборы находятся только в исправном состоянии;
- (5) после восстановления прибора заявка, на нем находящаяся, обслуживается заново.

Эту СМО можно привести к базовым моделям, определив марковский процесс обслуживания следующим образом (напомним, что слой k , где $k = \overline{0, n+r}$ при $r < \infty$ и $k \geq 0$ при $r = \infty$, состоит из всех состояний с k заявками в системе).

Множество \mathcal{X}_k состояний слоя k при $k = \overline{0, n-1}$ марковского процесса обслуживания имеет вид

$$\mathcal{X}_k = \{(i_1, \dots, i_k) \cup (i_1, \dots, i_{k-1}) \cup \dots \cup (0)\},$$

где состояние (i_1, \dots, i_m) , $i_1, \dots, i_m = \overline{1, J}$, $m = \overline{1, k}$, означает, что m приборов обслуживают заявки на фазах i_1, \dots, i_m , а $(k - m)$ приборов с заявками неисправны и восстанавливаются, состояние (0) — все k занятых приборов находятся в неисправном состоянии.

Множество \mathcal{X}_k состояний слоя k при $k \geq n$ марковского процесса обслуживания имеет вид

$$\mathcal{X}_k = \{(i_1, \dots, i_n) \cup (i_1, \dots, i_{n-1}) \cup \dots \cup (0)\},$$

где состояние (i_1, \dots, i_m) , $i_1, \dots, i_m = \overline{1, J}$, $m = \overline{1, n}$, означает, что m приборов обслуживают заявки на фазах i_1, \dots, i_m , остальные $(n - m)$ приборов с заявками неисправны и восстанавливаются и еще $(k - n)$ заявок находятся в накопителе. Состояние (0) означает, что все приборы заняты и неисправны, а $(k - n)$ заявок находятся в накопителе.

Далее для единообразия записи состояние (0) (т.е. состояние (i_1, \dots, i_m) при $m = 0$) при всех k будем отождествлять с состоянием (\cdot) .

Найдем интенсивности возможных переходов внутри каждого слоя (элементы матриц Λ_k , $k = \overline{1, n-1}$, и Λ) и между слоями (элементы матриц N_k , $k = \overline{1, n}$, и N) марковского процесса обслуживания базовой модели, соответствующей рассматриваемой СМО. Но прежде договоримся, рассматривая состояние марковского процесса, записывать сначала фазы обслуживания выделяемых приборов, указывая в виде верхнего индекса номер прибора в круглых скобках. Например, состояние $(j_1^{(l_1)}, j_2^{(l_2)}, i_1, \dots, i_{m-2})$ с двумя выделенными приборами, стоящими на местах l_1 и l_2 , $l_1 \neq l_2$ (для определенности пусть $l_1 < l_2$, причем l_1 и l_2 могут принимать любые значения от 1 до m), означает, что первый прибор находится на фазе $i_1, \dots, (l_1 - 1)$ -й прибор находится на фазе i_{l_1-1} , l_1 -й прибор находится на фазе j_1 , $(l_1 + 1)$ -й прибор находится на фазе $i_{l_1}, \dots, (l_2 - 1)$ -й прибор находится на фазе i_{l_2-2} , l_2 -й прибор находится на фазе j_2 , $(l_2 + 1)$ -й прибор находится на фазе i_{l_2-1}, \dots, m -й прибор находится на фазе i_{m-2} .

Рассмотрим вначале переходы внутри слоя, возникающие при изменении фазы обслуживания:

- из состояния $(i^{(l)}, i_1, \dots, i_{m-1})$, $m = \overline{1, k}$, $l = \overline{1, m}$, слоя k , $k = \overline{1, n-1}$, возможен переход в состояние $(j^{(l)}, i_1, \dots, i_{m-1})$ того же слоя с интенсивностью h_{ij} при изменении с i -й на j -ю, $j \neq i$, фазы обслуживания на l -м приборе;
- аналогично возможен переход из состояния $(i^{(l)}, i_1, \dots, i_{m-1})$, $m = \overline{1, n}$, $l = \overline{1, m}$, слоя k , $k \geq n$, в состояние $(j^{(l)}, i_1, \dots, i_{m-1})$, $j \neq i$, того же слоя с интенсивностью h_{ij} .

Следующие переходы внутри слоя происходят при отказе занятых приборов:

- из состояния $(j_1^{(l_1)}, \dots, j_i^{(l_i)}, i_1, \dots, i_{m-i})$, $i = \overline{1, m}$, $m = \overline{1, k}$, $1 \leq l_1 < \dots < l_i \leq m$, слоя k , $k = \overline{1, n-1}$, возможен переход в состояние (i_1, \dots, i_{m-i}) того же слоя с интенсивностью $q_m^i \alpha_i(m)$ при отказе i приборов из m ;
- из состояния $(j_1^{(l_1)}, \dots, j_i^{(l_i)}, i_1, \dots, i_{m-i})$, $i = \overline{1, m}$, $m = \overline{1, n}$, $1 \leq l_1 < \dots < l_i \leq m$, слоя k , $k \geq n$, возможен переход в состояние (i_1, \dots, i_{m-i}) того же слоя с интенсивностью $q_m^i \alpha_i(m)$.

Последний возможный тип переходов внутри слоя образуют переходы при восстановлении занятых приборов, причем предполагается, что восстановившимся приборам присваиваются порядковые номера, следующие за номерами исправных до момента восстановления приборов:

- из состояния (i_1, \dots, i_m) , $m = \overline{0, k-1}$, слоя k , $k = \overline{1, n-1}$, возможен с интенсивностью $h_{i_{m+1}} \dots h_{i_{m+j}} \beta_j(k - m)$ переход в состояние $(i_1, \dots, i_m, i_{m+1}, \dots, i_{m+j})$, $j = \overline{1, k-m}$, того же слоя при восстановлении j приборов из $(k - m)$ и возобновлении обслуживания находящихся на них заявок с новых фаз i_{m+1}, \dots, i_{m+j} ;
- из состояния (i_1, \dots, i_m) , $m = \overline{0, n-1}$, слоя k , $k \geq n$, возможен переход в состояние $(i_1, \dots, i_m, i_{m+1}, \dots, i_{m+j})$, $j = \overline{1, n-m}$, того же слоя с интенсивностью $h_{i_{m+1}} \dots h_{i_{m+j}} \beta_j(n - m)$.

Найденные интенсивности позволяют сформировать матрицы Λ_k , $k = \overline{1, n-1}$, и Λ базовых моделей.

Теперь рассмотрим переходы из слоя k в слой $(k - 1)$, которые могут происходить только при окончании обслуживания заявок на приборах, и их интенсивности задают матрицы N_k , $k = \overline{1, n}$, и N :

- из состояния $(i^{(l)}, i_1, \dots, i_{m-1})$, $m = \overline{1, k}$, $l = \overline{1, m}$, слоя k , $k = \overline{1, n}$, возможен переход в

состояние (i_1, \dots, i_{m-1}) слоя $(k-1)$ с интенсивностью h_i^* при окончании обслуживания заявки на l -м приборе;

- из состояния $(i^{(l)}, i_1, \dots, i_{m-1})$, $m = \overline{1, n}$, $l = \overline{1, m}$, слоя k , $k \geq n+1$, возможен переход в состояние (i_1, \dots, i_{m-1}, j) слоя $(k-1)$ с интенсивностью $h_j h_i^*$ при окончании обслуживания заявки на l -м приборе и поступлении на него заявки из накопителя, которая начинает обслуживаться на фазе j .

Наконец, определим ненулевые элементы матриц Ω_k , $k = \overline{0, n-1}$, связанных с поступлением заявок в систему:

- из состояния (i_1, \dots, i_m) , $m = \overline{0, k}$, слоя k , $k = \overline{0, n-1}$, при поступлении новой заявки на свободный прибор и начале ее обслуживания на фазе j с вероятностью h_j происходит переход в состояние (i_1, \dots, i_m, j) слоя $k+1$.

Перечислив все возможные переходы между слоями марковского процесса обслуживания с указанием их интенсивностей и вероятностей и определив тем самым ненулевые элементы матриц Λ_k , $k = \overline{1, n-1}$, Λ , N_k , $k = \overline{1, n}$, N и Ω_k , $k = \overline{0, n-1}$, можно воспользоваться полученными для базовых моделей в [74, 76] формулами расчета стационарных характеристик, связанных с числом заявок в системе.

4 Отказы всех приборов: заявки могут поступать на свободные неисправные приборы

Вторая рассматриваемая здесь СМО с ненадежными приборами отличается от разобранный в предыдущем разделе тем, что наряду с занятыми могут отказывать также и свободные приборы. А именно выполнены следующие предположения:

- (1) если в некоторый момент времени среди n приборов ровно k , $k = \overline{1, n}$, приборов исправны и заняты обслуживанием заявок, то за «малое» время Δ с вероятностью $\alpha_i(k)\Delta + o(\Delta)$, $i = \overline{1, k}$, произойдет отказ ровно i приборов с заявками;
- (2) если в некоторый момент времени среди n приборов ровно k , $k = \overline{1, n}$, приборов с заявками неисправны, то за «малое» время Δ с вероятностью $\beta_i(k)\Delta + o(\Delta)$, $i = \overline{1, k}$, произойдет восстановление ровно i приборов с заявками;

- (3) если в некоторый момент времени среди n приборов ровно k , $k = \overline{1, n}$, свободных приборов исправны, то за «малое» время Δ с вероятностью $\alpha_i^*(k)\Delta + o(\Delta)$, $i = \overline{1, k}$, произойдет отказ ровно i свободных приборов;
- (4) если в некоторый момент времени среди n приборов ровно k , $k = \overline{1, n}$, свободных приборов неисправны, то за «малое» время Δ с вероятностью $\beta_i^*(k)\Delta + o(\Delta)$, $i = \overline{1, k}$, произойдет восстановление ровно i свободных неисправных приборов;
- (5) при отказе или восстановлении i приборов из k возможных ($1 \leq i \leq k$) все варианты выбора i конкретных приборов из k равновероятны (вероятность выбора каждого варианта $q_k^i = 1/C_k^i$);
- (6) заявка, заставшая в момент поступления все исправные приборы занятыми, поступает на один из свободных неисправных приборов, а если таких нет, то попадает в накопитель;
- (7) после восстановления прибора заявка, на нем находящаяся, обслуживается заново.

Марковский процесс обслуживания, приводящий эту СМО к базовой модели, мало отличается от аналогичного процесса из предыдущего пункта.

В частности, множество \mathcal{X}_k состояний слоя k при $k \geq n$ имеет такой же вид

$$\mathcal{X}_k = \{(i_1, \dots, i_k) \cup (i_1, \dots, i_{k-1}) \cup \dots \cup (0)\},$$

как и раньше, с теми же самыми комментариями относительно самих состояний.

Множество \mathcal{X}_k состояний слоя k при $k = \overline{0, n-1}$ имеет вид

$$\mathcal{X}_k = \{(i_1, \dots, i_k; s) \cup (i_1, \dots, i_{k-1}; s) \cup \dots \cup (0; s)\},$$

где состояние $(i_1, \dots, i_m; s)$, $i_1, \dots, i_m = \overline{1, J}$, $m = \overline{1, k}$, $s = \overline{0, n-k}$, означает, что m приборов обслуживают заявки на фазах i_1, \dots, i_m , а $(k-m)$ приборов с заявками и s свободных приборов неисправны и восстанавливаются. Состояние $(0; s)$, $s = \overline{0, n-k}$, соответствует тому случаю, когда все занятые приборы и s свободных приборов находятся в неисправном состоянии.

Используя далее, как и в предыдущем разделе, обозначения $()$ и $(; s)$ наряду с обозначениями (0) и $(0; s)$ и приняв прежнее соглашение о том, что фазы выделенных приборов помечаются верхним индексом с номером прибора в круглых скобках, видим, что при $k \geq n$ возможные переходы внутри каждого слоя, возникающие при изменении фазы обслуживания, а также при отказе и восстановлении занятых приборов, полностью совпадают с переходами для модели предыдущего раздела.

Те же переходы при $k = \overline{1, n-1}$ определяются следующим образом:

- из состояния $(i^{(l)}, i_1, \dots, i_{m-1}; s)$, $m = \overline{1, k}$, $l = \overline{1, m}$, $s = \overline{0, n-k}$, слоя k , $k = \overline{1, n-1}$, возможен переход в состояние $(j^{(l)}, i_1, \dots, i_{m-1}; s)$ того же слоя с интенсивностью h_{ij} при изменении с i -й на j -ю, $j \neq i$, фазы обслуживания на l -м приборе;
- из состояния $(j_1^{(l_1)}, \dots, j_i^{(l_i)}, i_1, \dots, i_{m-i}; s)$, $i = \overline{1, m}$, $m = \overline{1, k}$, $1 \leq l_1 < \dots < l_i \leq m$, $s = \overline{0, n-k}$, слоя k , $k = \overline{1, n-1}$, возможен переход в состояние $(i_1, \dots, i_{m-i}; s)$ того же слоя с интенсивностью $q_m^i \alpha_i(m)$ при отказе i занятых приборов из m ;
- из состояния $(i_1, \dots, i_m; s)$, $m = \overline{0, k-1}$, $s = \overline{0, n-k}$, слоя k , $k = \overline{1, n-1}$, возможен переход в состояние $(i_1, \dots, i_m, i_{m+1}, \dots, i_{m+j}; s)$, $j = \overline{1, k-m}$, того же слоя с интенсивностью $h_{i_{m+1}} \dots h_{i_{m+j}} \beta_j(k-m)$ при восстановлении j занятых приборов из $(k-m)$.

Однако теперь могут происходить дополнительные переходы внутри слоя при отказе и восстановлении свободных приборов:

- из состояния $(i_1, \dots, i_m; s)$, $m = \overline{0, k}$, $s = \overline{0, n-k-1}$, слоя k , $k = \overline{1, n-1}$, возможен переход в состояние $(i_1, \dots, i_m; s+j)$, $j = \overline{1, n-k-s}$, того же слоя с интенсивностью $\alpha_j^*(n-k-s)$ при отказе j из $(n-k-s)$ исправных свободных приборов;
- из состояния $(i_1, \dots, i_m; s)$, $m = \overline{0, k}$, $s = \overline{1, n-k}$, слоя k , $k = \overline{1, n-1}$, возможен переход в состояние $(i_1, \dots, i_m; s-j)$, $j = \overline{1, s}$, того же слоя с интенсивностью $\beta_s^*(j)$ при восстановлении j из s неисправных свободных приборов.

Так же как и в предыдущем разделе, переходы из слоя k в слой $(k-1)$ могут происходить только при окончании обслуживания заявки на приборе. Если при $k \geq n+1$ переходы те же самые, что и раньше, то при $k = \overline{1, n}$ они отличаются лишь в силу изменения самого определения состояния:

- из состояния $(i^{(l)}, i_1, \dots, i_{m-1}; s)$, $m = \overline{1, k}$, $l = \overline{1, m}$, $s = \overline{1, n-k}$, слоя k , $k = \overline{1, n}$, возможен переход в состояние $(i_1, \dots, i_{m-1}; s)$ слоя $k-1$ с интенсивностью h_i^* при окончании обслуживания заявки на l -м приборе.

Наконец, определим ненулевые элементы матрицы Ω_k , $k = \overline{0, n-1}$, определяющих вероятности изменения состояний при поступлении заявок в систему:

- из состояния $(i_1, \dots, i_m; s)$, $m = \overline{0, k}$, $s = \overline{0, n-k-1}$, слоя k , $k = \overline{0, n-1}$, возможен

переход на слой $k+1$ в состояние $(i_1, \dots, i_m, i; s)$ с вероятностью h_i при поступлении новой заявки на исправный свободный прибор и начале ее обслуживания на фазе i ;

- из состояния $(i_1, \dots, i_m; n-k)$, $m = \overline{0, k}$, слоя k , $k = \overline{0, n-1}$, возможен переход на слой $k+1$ в состояние $(i_1, \dots, i_m; n-k-1)$ с вероятностью 1 при поступлении новой заявки на неисправный свободный прибор.

Теперь осталось воспользоваться формулами расчета стационарного распределения вероятностей состояний, полученными для базовой модели.

5 Отказы всех приборов: заявки могут поступать только на свободные исправные приборы

Последняя СМО, которая будет здесь рассмотрена, отличается от СМО предыдущего раздела лишь тем, что заявка может поступать только на свободный исправный прибор; в противном случае она ожидает появления таких приборов в общей очереди в накопителе. Однако, как будет видно далее, множество состояний теперь будет несколько другим.

Отметим, что в случае накопителя конечной емкости эту СМО нельзя привести к предложенной ранее базовой модели. Поэтому ограничимся здесь только случаем накопителя бесконечной емкости.

Для приведения этой СМО к базовой модели определим марковский процесс обслуживания следующим образом.

Множество состояний \mathcal{X}_k слоя k при $k = \overline{0, n-1}$ имеет вид

$$\mathcal{X}_k = \{(0; i_1, \dots, i_m; m_1) \cup (1; i_1, \dots, i_m; m_2)\},$$

где:

- состояние $(0; i_1, \dots, i_m; m_1)$, $m = \overline{0, k}$, $i_1, \dots, i_m = \overline{1, J}$, $m_1 = \overline{0, n-k}$, означает, что все k заявок находятся на приборах, причем m из них обслуживаются на фазах i_1, \dots, i_m , а остальные $(k-m)$ находятся на неисправных приборах, и m_1 свободных приборов неисправны;
- состояние $(1; i_1, \dots, i_m; m_2)$, $m = \overline{0, k-1}$, $i_1, \dots, i_m = \overline{1, J}$, $m_2 = \overline{0, k-1-m}$, означает, что все свободные приборы неисправны, m приборов обслуживают заявки на фазах i_1, \dots, i_m , m_2 приборов с заявками восстанавливаются, и $k-m-m_2$ ($k-m-m_2 > 0$) заявок находятся в накопителе.

Подчеркнем, что смысл индексов m_1 и m_2 для этих двух типов состояний различен.

Множество состояний \mathcal{X}_k слоя k при $k \geq n$ марковского процесса обслуживания имеет вид

$$\mathcal{X}_k = \{(i_1, \dots, i_m; m_2)\},$$

где состояние $(i_1, \dots, i_m; m_2)$, $m = \overline{0, n}$, $i_1, \dots, i_m = \overline{1, J}$, $m_2 = \overline{0, n - m}$, означает, что m приборов обслуживают заявки на фазах i_1, \dots, i_m , m_2 приборов с заявками неисправны и восстанавливаются, все (если таковые имеются) $(n - m - m_2)$ свободных приборов неисправны и еще $(k - m - m_2)$ заявок находятся в накопителе.

Как и прежде, для единообразия записи приемем соглашение, что состояния $(0; ; m_1)$, $(1; ; m_2)$ и $(; m_1)$ соответствуют тому случаю, когда все занятые заявками приборы неисправны.

Рассмотрим возможные переходы марковского процесса обслуживания, договорившись, как и ранее, верхним индексом в круглых скобках обозначать фазы выделенных приборов.

Первый тип переходов — из слоя k на тот же самый слой k .

Обратимся сначала к переходам на слое k при смене фазы обслуживания:

- из состояния $(0; i^{(l)}, i_1, \dots, i_{m-1}; m_1)$, $l = \overline{1, m}$, $m = \overline{1, k}$, $m_1 = \overline{0, n - k}$, слоя k , $k = \overline{1, n - 1}$, возможен переход в состояние $(0; j^{(l)}, i_1, \dots, i_{m-1}; m_1)$ того же слоя с интенсивностью h_{ij} при изменении с i -й на j -ю, $j \neq i$, фазы обслуживания на l -м приборе;
- из состояния $(1; i^{(l)}, i_1, \dots, i_{m-1}; m_2)$, $l = \overline{1, m}$, $m = \overline{1, k - 1}$, $m_2 = \overline{0, k - 1 - m}$, слоя k , $k = \overline{1, n - 1}$, возможен переход в состояние $(1; j^{(l)}, i_1, \dots, i_{m-1}; m_2)$ того же слоя с интенсивностью h_{ij} при изменении с i -й на j -ю, $j \neq i$, фазы обслуживания на l -м приборе;
- из состояния $(i^{(l)}, i_1, \dots, i_{m-1}; m_2)$, $l = \overline{1, m}$, $m = \overline{1, n}$, $m_2 = \overline{0, n - m}$, слоя k , $k \geq n$, возможен переход в состояние $(j^{(l)}, i_1, \dots, i_{m-1}; m_2)$ того же слоя с интенсивностью h_{ij} при изменении с i -й на j -ю, $j \neq i$, фазы обслуживания на l -м приборе.

Теперь рассмотрим переходы при отказе занятых приборов:

- из состояния $(0; j_1^{(l_1)}, \dots, j_i^{(l_i)}, i_1, \dots, i_{m-i}; m_1)$, $i = \overline{1, m}$, $1 \leq l_1 < \dots < l_i \leq m$, $m = \overline{1, k}$, $m_1 = \overline{0, n - k}$, слоя k , $k = \overline{1, n - 1}$, возможен переход в состояние $(0; i_1, \dots, i_{m-i}; m_1)$ того же слоя с интенсивностью $q_m^i \alpha_i(m)$ при отказе i занятых приборов из m ;

– из состояния $(1; j_1^{(l_1)}, \dots, j_i^{(l_i)}, i_1, \dots, i_{m-i}; m_2)$, $i = \overline{1, m}$, $1 \leq l_1 < \dots < l_i \leq m$, $m = \overline{1, k - 1}$, $m_2 = \overline{0, k - 1 - m}$, слоя k , $k = \overline{1, n - 1}$, возможен переход в состояние $(1; i_1, \dots, i_{m-i}; m_2 + i)$ того же слоя с интенсивностью $q_m^i \alpha_i(m)$ при отказе i занятых приборов из m ;

– из состояния $(j_1^{(l_1)}, \dots, j_i^{(l_i)}, i_1, \dots, i_{m-i}; m_2)$, $i = \overline{1, m}$, $1 \leq l_1 < \dots < l_i \leq m$, $m = \overline{1, n}$, $m_2 = \overline{0, n - m}$, слоя k , $k \geq n$, возможен переход в состояние $(i_1, \dots, i_{m-i}; m_2 + i)$ того же слоя с интенсивностью $q_m^i \alpha_i(m)$ при отказе i занятых приборов из m .

Перейдем к переходам при восстановлении занятых приборов (предполагается, что восстановившимся приборам присваиваются порядковые номера, следующие за номерами исправных до момента восстановления приборов):

- из состояния $(0; i_1, \dots, i_m; m_1)$, $m = \overline{0, k - 1}$, $m_1 = \overline{0, n - k}$, слоя k , $k = \overline{1, n - 1}$, возможен переход в состояние $(0; i_1, \dots, i_m, i_{m+1}, \dots, i_{m+j}; m_1)$, $j = \overline{1, k - m}$, того же слоя с интенсивностью $h_{i_{m+1}} \dots h_{i_{m+j}} \beta_j(k - m)$ при восстановлении j занятых приборов из $(k - m)$ и начале обслуживания заново на фазах i_{m+1}, \dots, i_{m+j} находящихся на них заявок;
- из состояния $(1; i_1, \dots, i_m; m_2)$, $m = \overline{0, k - 2}$, $m_2 = \overline{1, k - 1 - m}$, слоя k , $k = \overline{1, n - 1}$, возможен переход в состояние $(1; i_1, \dots, i_m, i_{m+1}, \dots, i_{m+j}; m_2)$, $j = \overline{1, k - 1 - m}$, того же слоя с интенсивностью $h_{i_{m+1}} \dots h_{i_{m+j}} \beta_j(k - 1 - m)$ при восстановлении j занятых приборов из $(k - 1 - m)$ и начале обслуживания заново на фазах i_{m+1}, \dots, i_{m+j} находящихся на них заявок;
- из состояния $(i_1, \dots, i_m; m_2)$, $m = \overline{0, k - 2}$, $m_2 = \overline{1, k - 1 - m}$, слоя k , $k \geq n$, возможен переход в состояние $(i_1, \dots, i_m, i_{m+1}, \dots, i_{m+j}; m_2)$, $j = \overline{1, k - 1 - m}$, того же слоя с интенсивностью $h_{i_{m+1}} \dots h_{i_{m+j}} \beta_j(k - 1 - m)$ при восстановлении j занятых приборов из $(k - 1 - m)$ и начале обслуживания заново на фазах i_{m+1}, \dots, i_{m+j} находящихся на них заявок.

Далее, обратимся к переходам при отказе свободных приборов:

- на слое k , $k = \overline{0, n - 1}$, переходы возможны только из состояний $(0; i_1, \dots, i_m; m_1)$, $m = \overline{0, k}$, $m_1 = \overline{0, n - k - 1}$, в состояния $(0; i_1, \dots, i_m; m_1 + j)$ того же слоя с интенсивностью $q_{n-k-m_1}^j \alpha_j^*(n - k - m_1)$ при отказе j свободных приборов из $n - k - m_1$;

– на слое k , $k \geq n$, переходы из-за отказа свободных приборов не происходят.

Наконец, рассмотрим переходы при восстановлении свободных приборов:

- из состояния $(0; i_1, \dots, i_m; m_1)$, $m = \overline{0, k}$, $m_1 = \overline{1, n - k}$, слоя k , $k = \overline{0, n - 1}$, возможен переход в состояние $(0; i_1, \dots, i_m; m_1 - j)$, $j = \overline{1, m_1}$, того же слоя с интенсивностью $\beta_j^*(m_1)$ при восстановлении j свободных приборов из m_1 ;
- из состояния $(1; i_1, \dots, i_m; m_2)$, $m = \overline{0, k - 1}$, $m_2 = \overline{1, k - 1 - m}$, слоя k , $k = \overline{1, n - 1}$, возможен переход в состояние $(1; i_1, \dots, i_m, i_{m+1}, \dots, i_{m+j}; m_2)$, $j = \overline{1, k - 1 - m - m_2}$, того же слоя с интенсивностью $h_{i_{m+1}} \dots h_{i_{m+j}} \beta_j^*(n - m - m_2)$ при восстановлении j свободных приборов из $(n - m - m_2)$ и начале обслуживания части (а именно j) заявок из очереди на фазах i_{m+1}, \dots, i_{m+j} ;
- из состояния $(1; i_1, \dots, i_m; m_2)$, $m = \overline{0, k - 1}$, $m_2 = \overline{1, k - 1 - m}$, слоя k , $k = \overline{1, n - 1}$, возможен также переход в состояние $(0; i_1, \dots, i_m, i_{m+1}, \dots, i_{k-m-2}; n - m - m_2 - j)$, $j = k - m - m_2, n - m - m_2$, того же слоя с интенсивностью $h_{i_{m+1}} \dots h_{i_{k-m-2}} \beta_j^*(n - m - m_2)$ при восстановлении j свободных приборов из $(n - m - m_2)$ и начале обслуживания всех $(k - m - m_2)$ заявок из очереди на фазах $i_{m+1}, \dots, i_{k-m-2}$;
- из состояния $(i_1, \dots, i_m; m_2)$, $m = \overline{0, n - 1}$, $m_2 = \overline{1, n - 1 - m}$, слоя k , $k \geq n$, возможен переход в состояние $(i_1, \dots, i_m, i_{m+1}, \dots, i_{m+j}; m_2)$, $j = \overline{1, n - m - m_2}$, того же слоя с интенсивностью $h_{i_{m+1}} \dots h_{i_{m+j}} \beta_j^*(n - m - m_2)$ при восстановлении j свободных приборов из $(n - m - m_2)$ и начале обслуживания j заявок из очереди на фазах i_{m+1}, \dots, i_{m+j} .

Второй тип переходов — из слоя k на слой $(k - 1)$ — образуют переходы при окончании обслуживания заявки на приборе:

- из состояния $(0; i^{(l)}, i_1, \dots, i_{m-1}; m_1)$, $l = \overline{1, m}$, $m = \overline{1, k}$, $m_1 = \overline{0, n - k}$, слоя k , $k = \overline{1, n - 1}$, возможен переход в состояние $(0; i_1, \dots, i_{m-1}; m_1)$ слоя $(k - 1)$ с интенсивностью h_i^* при окончании обслуживания заявки на l -м приборе;
- из состояния $(1; i^{(l)}, i_1, \dots, i_{m-1}; m_2)$, $l = \overline{1, m}$, $m = \overline{1, k - 2}$, $m_2 = \overline{0, k - 2 - m}$, слоя k , $k = \overline{1, n - 1}$, (в накопителе находится две или более заявок) возможен переход в состояние $(1; i_1, \dots, i_{m-1}, i_m; m_2)$ слоя $(k - 1)$ с интенсивностью $h_i^* h_{i_m}$ при окончании обслуживания заявки на l -м приборе и поступлении на него

заявки из накопителя, которая немедленно начинает обслуживаться на фазе i_m ;

- из состояния $(1; i^{(l)}, i_1, \dots, i_{m-1}; k - 1 - m)$, $l = \overline{1, m}$, $m = \overline{1, k - 1}$, слоя k , $k = \overline{1, n - 1}$, (в накопителе находится ровно одна заявка) возможен переход в состояние $(0; i_1, \dots, i_{m-1}, i_m; n - k + 1)$ слоя $(k - 1)$ (все заявки находятся на приборах, а все свободные приборы неисправны) с интенсивностью $h_i^* h_{i_m}$ при окончании обслуживания заявки на l -м приборе и поступлении на него единственной заявки из накопителя, которая немедленно начинает обслуживаться на фазе i_m ;
- из состояния $(i^{(l)}, i_1, \dots, i_{m-1}; n - 1 - m)$, $l = \overline{1, m}$, $m = \overline{1, n - 1}$, слоя n (в накопителе находится ровно одна заявка) возможен переход в состояние $(0, i_1, \dots, i_{m-1}, i_m; 1)$ слоя $(n - 1)$ (все заявки находятся на приборах, а единственный свободный прибор неисправен) с интенсивностью $h_i^* h_{i_m}$ при окончании обслуживания заявки на l -м приборе и поступлении на него единственной заявки из накопителя, которая немедленно начинает обслуживаться на фазе i_m ;
- из состояния $(i^{(l)}, i_1, \dots, i_{m-1}; m_2)$, $l = \overline{1, m}$, $m = \overline{1, n - 2}$, $m_2 = \overline{0, n - 2 - m}$, слоя n (в накопителе находится две или более заявок) возможен переход в состояние $(1, i_1, \dots, i_{m-1}, i_m; m_2)$ слоя $(n - 1)$ с интенсивностью $h_i^* h_{i_m}$ при окончании обслуживания заявки на l -м приборе и поступлении на него заявки из накопителя, которая немедленно начинает обслуживаться на фазе i_m ;
- из состояния $(i^{(l)}, i_1, \dots, i_{m-1}; m_2)$, $l = \overline{1, m}$, $m = \overline{1, n}$, $m_2 = \overline{0, n - m}$, слоя k , $k > n$, возможен переход в состояние $(i_1, \dots, i_{m-1}, i_m; m_2)$ слоя $(k - 1)$ с интенсивностью $h_i^* h_{i_m}$ при окончании обслуживания заявки на l -м приборе и поступлении на него заявки из накопителя, которая немедленно начинает обслуживаться на фазе i_m .

Осталось определить возможные переходы последнего типа, связанные с поступлением заявки, и ненулевые элементы матриц Ω_k , $k = \overline{0, n - 1}$:

- из состояния $(0; i_1, \dots, i_m; m_1)$, $m = \overline{1, k}$, $m_1 = \overline{0, n - k - 1}$, слоя k , $k = \overline{0, n - 2}$, возможен переход на слой $(k + 1)$ в состояние $(0; i_1, \dots, i_m, i_{m+1}; m_1)$ с вероятностью $h_{i_{m+1}}$ при поступлении новой заявки на свободный прибор и начале ее обслуживания на фазе i_{m+1} ;
- из состояния $(0; i_1, \dots, i_m; 0)$, $m = \overline{1, n - 1}$, слоя $(n - 1)$ возможен переход на слой n в состояние $(i_1, \dots, i_m, i_{m+1}; n - m - 1)$ с вероятностью

$h_{i_{m+1}}$ при поступлении новой заявки на единственный свободный (исправный) прибор и начале ее обслуживания на фазе i_{m+1} ;

- из состояния $(0; i_1, \dots, i_m; n - k)$, $m = \overline{1, k}$, слоя k , $k = \overline{0, n - 2}$, возможен переход на слой $k + 1$ в состояние $(1; i_1, \dots, i_m; n - k - m)$ с вероятностью $h_{i_{m+1}}$ при поступлении новой заявки в накопитель;
- из состояния $(0; i_1, \dots, i_m; 1)$, $m = \overline{1, n - 1}$, слоя $(n - 1)$ (единственный свободный прибор находится в неисправном состоянии) с вероятностью 1 происходит переход на слой n в состояние $(i_1, \dots, i_m; n - m - 1)$ при поступлении новой заявки в накопитель;
- из состояния $(1; i_1, \dots, i_m; m_2)$, $m = \overline{0, k - 1}$, $m_2 = 0, k - 1 - m$, слоя k , $k = \overline{0, n - 2}$, (в накопителе есть заявки, и все свободные приборы неисправны) возможен переход на слой $(k + 1)$ в состояние $(1; i_1, \dots, i_m; m_2)$ при поступлении новой заявки в накопитель;
- из состояния $(1; i_1, \dots, i_m; m_2)$, $m = \overline{1, n - 2}$, $m_2 = 0, n - 2 - m$, слоя $(n - 1)$ возможен переход на слой n в состояние $(i_1, \dots, i_m; m_2)$ при поступлении новой заявки в накопитель.

Перечислив все возможные переходы между слоями марковского процесса обслуживания с указанием их интенсивностей или вероятностей и определив тем самым матрицы Λ_k , $k = \overline{1, n - 1}$, Λ , N_k , $k = \overline{1, n}$, N и Ω_k , $k = \overline{0, n - 1}$, можно теперь для расчета стационарного распределения числа заявок в системе воспользоваться соотношениями, полученными для базовой модели.

6 Заключение

Таким образом, в настоящей статье получены соотношения, позволяющие с помощью разработанной ранее базовой модели вычислять стационарные характеристики очереди для многолинейных СМО с ненадежными приборами, полумарковским входящим потоком, РН-распределением времени обслуживания заявок, накопителем конечной или бесконечной емкости, обслуживанием заново заявок после восстановления приборов и тремя типами групповых отказов и восстановлений приборов:

- (1) отказывают только приборы, на которых находятся заявки;
- (2) отказывают все приборы, а заявки могут поступать как на исправные, так и на неисправные свободные приборы;

- (3) отказывают все приборы, но заявки могут поступать только на исправные свободные приборы.

К сожалению, задача нахождения математических соотношений для расчета стационарных распределений, связанных с временем пребывания заявки в рассматриваемых системах, не может быть решена с помощью базовой модели и требует отдельного изучения за рамками данной статьи.

Литература

1. *Dimitrov B., Petrov P.* The minimal blocking time by unreliable server and latent failures // 14th European Meeting of Statisticians, Wroclaw, 1981. P. 126–127.
2. *Dimitrov B., Dokev Ch.* The single server queue system with non-reliable server in discrete time. Non-stationary characteristics // Ann. of Univ. of Sofia Ser. Math., Sofia, 1981. Vol. 70. P. 175–190.
3. *Cao J., Cheng K.* Analysis of $M/G/1$ queueing system with repairable service station // Acta Mathematicae Applicatae Sinica, 1982. Vol. 5. P. 113–127.
4. *Nicola V. F.* A single-server queue with mixed types of interruptions // Acta Inform., 1986. Vol. 23. No. 4. P. 465–486.
5. *Kulkarni V. G., Choi B. D.* Retrial queues with server subject to breakdowns and repairs // Queueing Syst., 1990. Vol. 7. No. 2. P. 191–208.
6. *Ibe O. C., Trivedi K. S.* Two queues with alternating service and server breakdown // Queueing Syst., 1990. Vol. 7. No. 3. P. 253–268.
7. *Dimitrov B., Khalil Z.* On a new characterization of the exponential distribution related to a queueing system with an unreliable server // J. Appl. Probab., 1990. Vol. 27. P. 221–226.
8. *Dimitrov B., Khalil Z.* Some characterizations of the exponential distribution based on the service time properties of an unreliable server // Lect. Notes Math. Stability Problems for Stochastic Models, 1993. Vol. 1546. P. 17–25.
9. *Boxma O. J., Weststrate J. A., Yechiali U.* A globally gated polling system with server interruptions, and applications to the repairman problem // Probab. Eng. Inform. Sci., 1993. Vol. 7. No. 2. P. 187–208.
10. *Khalil Z., Dimitrov B.* The service time properties of an unreliable server characterize the exponential distribution // Adv. Appl. Probab., 1994. Vol. 26. P. 172–182.
11. *Yang T., Li H.* The $M/G/1$ retrial queue with the server subject to starting failures // Queueing Syst. 1994. Vol. 16. Nos. 1–2. P. 83–96.
12. *Blanc J. P. C., van der Mei R. D.* The power-series algorithm applied to polling systems with a Dormant server // The fundamental role of teletraffic in the evolution of

- telecommunication networks / Eds. J. Labetoulle and J. W. Roberts. — Amsterdam: Elsevier, 1994. P. 865–874.
13. *Aissani A.* A retrial queue with redundancy and unreliable server // *Queueing Syst.*, 1994. Vol. 17. No. 3–4. P. 431–449.
 14. *Hsieh Y., Andersland M. S.* Repairable single-server systems with multiple breakdown modes // *Microelectron. Reliab.*, 1995. Vol. 35. No. 2. P. 309–318.
 15. *Kotlyar V. Yu.* Queueing system with an absolutely unreliable server and a variable stream of customers // *Cybern. Syst. Anal.*, 1995. Vol. 31. No. 2. P. 285–292.
 16. *Chukova S., Dimitrov B.* Execution time on an unreliable server with latent breakdowns // *Matrix-analytics methods in stochastic models* / Eds. S. Chakravarthy and A. Alfa. — New York, Basel, Hong Kong: Marcel Dekker, Inc., 1996. P. 225–239.
 17. *Kofman D., Yechiali U.* Polling with stations breakdowns // *Perform. Eval.*, 1996. Vol. 27–28. No. 4. P. 647–672.
 18. *Li W., Shi D., Chao X.* Reliability analysis of $M/G/1$ queueing systems with server breakdowns and vacations // *J. Appl. Probab.*, 1997. Vol. 34. No. 2. P. 546–555.
 19. *Tang Y. H.* A single-server $M/G/1$ queueing system subject to breakdowns — some reliability and queueing problems // *Microelectron. Reliab.*, 1997. Vol. 37. No. 2. P. 315–321.
 20. *Lee D.-S.* Analysis of a single server queue with semi-Markovian service interruption // *Queueing Syst.*, 1997. Vol. 27. No. 1–2. P. 153–178.
 21. *Aissani A., Artalejo J. R.* On the single server retrial queue subject to breakdowns // *Queueing Syst.*, 1998. Vol. 30. No. 3–4. P. 309–321.
 22. *Атенюа И. М., Бочаров П. П., Пузикова Д. А.* Матрично-мультипликативное решение для однолинейной системы с отключениями прибора, конечной очередью повторных заявок и распределениями фазового типа // *Автоматика и телемеханика*, 1999. № 9. С. 72–91.
 23. *Almasi B.* A queueing model for a non-homogeneous polling system subject to breakdowns // *Ann. Univ. Sci. Budapest. Sect. Comp.*, 1999. Vol. 18. P. 11–23.
 24. *Chakravarthy S. R., Krishnamoorthy A., Ushakumari P. V.* A k -out-of- n reliability system with an unreliable server and phase type repairs and services: The (N, T) policy // *J. Appl. Math. Stochastic Anal.*, 2001. Vol. 14. P. 361–380.
 25. *Wang J., Cao J., Li Q.* Reliability analysis of the retrial queue with server breakdowns and repairs // *Queueing Syst.*, 2001. Vol. 38. No. 4. P. 363–380.
 26. *Krishna Kumar B., Arivudainambi D., Vijayakumar A.* An $M/G/1/1$ queue with unreliable server and no waiting capacity // *Inf. Manage. Sci.*, 2002. Vol. 13. P. 35–50.
 27. *Krishna Kumar B., Pavai M. S., Vijayakumar A.* The $M/G/1$ retrial queue with feedback and starting failures // *Appl. Math. Model.*, 2002. Vol. 26. P. 1057–1075.
 28. *Djellab N. V.* On the $M/G/1$ retrial queue subjected to breakdowns // *RAIRO Oper. Res.*, 2002. Vol. 36. P. 299–310.
 29. *Дудин А. Н.* Оптимальное гистерезисное управление ненадежной системой $AP/SM/1$ с двумя режимами работы // *Автоматика и телемеханика*, 2002. № 10. С. 58–72.
 30. *Nakdimon O., Yechiali U.* Polling systems with breakdowns and repairs // *Eur. J. Oper. Res.*, 2003. Vol. 149. No. 3. P. 588–613.
 31. *Chakravarthy S. R., Agarwal A.* Analysis of a machine repair problem with an unreliable server and phase type repairs and services // *Nav. Res. Log.*, 2003. Vol. 50. P. 462–480.
 32. *Xueming Y., Li W.* Availability analysis of the queueing system $GI/PH/1$ with server breakdowns // *J. Syst. Sci. Complexity*, 2003. Vol. 16. No. 2. P. 177–183.
 33. *Хуродзе П. А., Хочолава В. В., Микадзе И. С.* Об одной системе массового обслуживания с ненадежной обслуживающей системой // *Проблемы прикладной механики*, 2003. № 3(12). С. 9–18.
 34. *Микадзе И. С., Хочолава В. В.* Об одной модели передачи информации по ненадежному каналу связи // *Автоматика и телемеханика*, 2004. № 8. С. 85–90.
 35. *Микадзе И. С., Хочолава В. В., Хуродзе П. А.* Виртуальное время ожидания в однолинейной СМО с ненадежным прибором // *Автоматика и телемеханика*, 2004. № 12. С. 119–128.
 36. *Dudin A. N., Kazimirsky A. V., Klimenok V. I.* $BMAP/G/1$ system unreliable in an idle state // *Bull. Kerala Math. Assoc.*, 2004. No. 2. P. 1–19.
 37. *Микадзе И. С., Хочолава В. В.* Исследование длины очереди в однолинейной СМО с ненадежным прибором // *Автоматика и телемеханика*, 2005. № 1. С. 72–81.
 38. *Li H., Zhao Y. Q.* A retrial queue with a constant retrial rate, server downs and impatient customers // *Stoch. Models*, 2005. Vol. 21. P. 531–550.
 39. *Sherman N. P., Kharoufeh J. P.* An $M/M/1$ retrial queue with unreliable server // *Oper. Res. Lett.*, 2006. Vol. 34. No. 6. P. 697–705.
 40. *Li J., Wang J.* An $M/G/1$ retrial queue with second multi-optional service, feedback and unreliable server // *Appl. Math.* — *J. Chin. Univ.*, 2006. Vol. 21. No. 3. P. 252–262.
 41. *Moreno P.* A discrete-time retrial queue with unreliable server and general server lifetime // *J. Math. Sci.*, 2006. Vol. 132. No. 5. P. 643–655.
 42. *Sztrik J., Almasi B., Roszik J.* Heterogeneous finite-source retrial queues with server subject to breakdowns and repairs // *J. Math. Sci.*, 2006. Vol. 132, No. 5. P. 677–685.
 43. *Atencia I., Moreno P.* A discrete-time $Geo/G/1$ retrial queue with the server subject to starting failures // *Ann. Oper. Res.*, 2006. Vol. 141. No. 1. P. 85–107.

44. *Quan-Lin Li, Ying Yu, Yiqiang Zhao Q.* A BMAP/G/1 retrial queue with a server subject to breakdowns and repairs // *Ann. Oper. Res.*, 2006. Vol. 141. No. 1. P. 233–270.
45. *Krishna Kumar B., Krishnamoorthy A., Pavai Madheswari S., Sadiq Basha S.* Transient analysis of a single server queue with catastrophes, failures and repairs // *Queueing Syst.*, 2007. Vol. 56. No. 3–4. P. 133–141.
46. *Falin G. I.* The M/M/1 retrial queue with retrials due to server failures // *Queueing Syst.*, 2008. Vol. 58. No. 3. P. 155–160.
47. *Wang J.* On the single server retrial queue with priority subscribers and server breakdowns // *J. Syst. Sci. Complexity*, 2008. Vol. 21. No. 2. P. 304–315.
48. *Atencia I., Bouza G., Moreno P.* An $M^{[X]}/G/1$ retrial queue with server breakdowns and constant rate of repeated attempts // *Ann. Oper. Res.*, 2008. Vol. 157. No. 1. P. 225–243.
49. *Вишневецкий В. М., Семёнова О. В.* Математические методы исследования систем поллинга // *Автоматика и телемеханика*, 2006. № 2. С. 3–56.
50. *Вишневецкий В. М., Семёнова О. В.* Системы поллинга: теория и применение в широкополосных беспроводных сетях. — М.: Техносфера, 2007. — 312 с.
51. *Mytrany I. L., Avi-Itzhak B.* A many-server queue with service interruptions // *Oper. Res.*, 1968. Vol. 16. P. 628–638.
52. *Kabak I. V.* Blocking and delays in $M^{(n)}/M/c$ queueing systems // *Oper. Res.*, 1968. Vol. 16. P. 830–840.
53. *Neuts M. F., Lucantoni D. M.* A Markovian queue with N servers subject to breakdowns and repairs // *Mgmt. Sci.*, 1979. Vol. 25. P. 849–861.
54. *Roszik J., Sztrik J.* Performance analysis of finite-source retrial queues with nonreliable heterogenous servers // *J. Math. Sci.*, 2007. Vol. 146. No. 4. P. 6033–6038.
55. *Микадзе З. И., Микадзе И. С., Хочолава В. В.* Об одной многоканальной смешанной системе массового обслуживания с ограниченным временем ожидания // *Автоматика и телемеханика*, 2007. № 7. С. 44–51.
56. *Chakravarthy S. R.* Analysis of production line systems with two unreliable machines with phase type processing times and a finite storage buffer // *Stoch. Models*, 1987. Vol. 3. P. 369–391.
57. *Тананко И. Е., Юдаева Н. В.* Исследование сети массового обслуживания с ненадежными системами и задержкой информации // *Тезисы докл. VIII Всероссийского симпозиума по прикладной и промышленной математике. Обзорение прикладной и промышленной математики*, 2007. Т. 14. Вып. 6. С. 1137–1138.
58. *Тананко И. Е., Юдаева Н. В.* Моделирование сети массового обслуживания с ненадежными системами и задержкой информации // *Компьютерные науки и информационные технологии: Тез. докл. Междунар. науч. конф., посвященной памяти проф. А. М. Богомолова.* — Саратов: Изд-во Саратов. ун-та, 2007. С. 119–120.
59. *White H., Christie L. S.* Queueing with preemptive priorities or with breakdowns // *Oper. Res.*, 1958. Vol. 6. P. 79–95.
60. *Thiruvengadam K.* Queueing with breakdowns // *Oper. Res.*, 1963. Vol. 11. P. 62–71.
61. *Avi-Itzhak B., Naor P. P.* Some queueing problems with the service station subject to breakdown // *Oper. Res.*, 1963. Vol. 11. P. 303–320.
62. *Jaiswal N. K., Thiruvengadam K.* Simple machine interference with two types of failure // *Oper. Res.*, 1963. Vol. 11. No. 4. P. 624–636.
63. *Elsayed E. A., Proctor C. L.* Two repair policies for a machine interference problem with two types of failures // *Proceedings of the Annual Pittsburgh Conference “Modeling and Simulation,”* 1979. P. 197.
64. *Federgruen A., Green L.* Queueing systems with service interruptions // *Oper. Res.*, 1986. Vol. 34. P. 752–768.
65. *Sztrik J., Gal T.* A recursive solution of a queueing model for a multi-terminal system subject to breakdowns // *Perform. Eval.*, 1990. Vol. 11. No. 1. P. 1–7.
66. *Artalejo J. R.* New results in retrial queueing systems with breakdown of the servers // *Stat. Neerl.*, 1994. Vol. 48. No. 1. P. 23–36.
67. *Бабицкий А. В., Дудин А. Н., Клименок В. И.* К расчету характеристик ненадежной системы массового обслуживания с конечным источником // *Автоматика и телемеханика*. 1996. № 1. С. 92–103.
68. *Krishnamoorthy A., Ushakumari P. V.* Reliability of a k -out-of- n system with repair and retrial of failed units // *TOP*, 1999. Vol. 7. No. 2. P. 293–304.
69. *Dimitrov B., Chukova S., Chakravarthy S.* A simple unreliable service model characterizes exponential distribution // *Kuwait J. Sci.*, 2001. Vol. 28 No. 2. P. 203–212.
70. *Gray W. J., Wang P. P., Scott M.* A queueing model with service breakdowns and multiple stages of repair // *J. Appl. Statistical Sci.*, 2003. Vol. 12. No. 1. P. 75–89.
71. *Gray W. J., Wang P. P., Scott M.* A queueing model with multiple types of server breakdowns // *Qual. Technol. Quant. Manage.*, 2004. Vol. 1. No. 2. P. 245–255.
72. *Martin S. P., Mitrani I.* Analysis of job transfer policies in systems with unreliable servers // *Ann. Oper. Res.*, 2008. Vol. 162. No. 1. P. 127–141.
73. *Печинкин А. В., Соколов И. А., Чаплыгин В. В.* Многолинейные системы массового обслуживания с независимыми отказами и восстановлениями приборов // *Системы и средства информатики. Спец. выпуск «Математическое и алгоритмическое обеспечение информационно-телекоммуникационных систем».* — М.: ИПИ РАН, 2006. С. 99–123.

74. Печинкин А. В., Соколов И. А., Чаплыгин В. В. Многолинейная система массового обслуживания с конечным накопителем и ненадежными приборами // Информатика и её применения, 2007. Т. 1. Вып. 1. С. 27–39.
75. Печинкин А. В., Соколов И. А., Чаплыгин В. В. Стационарные характеристики многолинейной системы массового обслуживания с одновременными отказами приборов // Информатика и её применения, 2007. Т. 1. Вып. 2. С. 28–38.
76. Печинкин А. В., Чаплыгин В. В. Стационарные характеристики системы массового обслуживания $SM/MSP/n/r$ // Автоматика и телемеханика, 2004. №9. С. 85–100.
77. Печинкин А. В., Чаплыгин В. В. Стационарные характеристики системы массового обслуживания $G/MSP/n/r$ // Вестник РУДН. Сер. «Прикладная математика и информатика», 2003. № 1. С. 119–143.
78. Бочаров П. П., Печинкин А. В. Теория массового обслуживания. — М.: Изд-во РУДН, 1995.

О ПРЕДЕЛЬНЫХ ХАРАКТЕРИСТИКАХ СИСТЕМЫ ОБСЛУЖИВАНИЯ $M(t)/M(t)/S$ С КАТАСТРОФАМИ

А. И. Зейфман¹, Я. А. Сатин², А. В. Коротышева³, Н. А. Терёшина⁴

Аннотация: Рассмотрена модель системы обслуживания $M(t)/M(t)/S$ с катастрофами в общем случае, когда интенсивности катастроф зависят от числа требований в системе. Получены достаточные условия слабой эргодичности процесса, описывающего число требований в системе, и соответствующие оценки. Рассмотрено несколько примеров построения предельных характеристик системы.

Ключевые слова: нестационарные марковские системы обслуживания; процесс рождения и гибели с катастрофами; слабая эргодичность; оценки; предельные характеристики; аппроксимация

1 Введение

Простейшие (стационарные) модели систем обслуживания с катастрофами начали рассматриваться не очень давно (см., например, [1–4]). В таких моделях предполагается, что если в системе обслуживания имеется ненулевое число требований, то с ненулевой интенсивностью возможна катастрофа, т. е. потеря всех требований, с дальнейшим продолжением функционирования системы обслуживания. Нестационарные марковские модели (процессы рождения и гибели) с катастрофами изучались в работах [5, 6] для случая, когда интенсивности катастроф не зависят от числа требований в системе. В настоящей работе изучается модель системы обслуживания $M(t)/M(t)/S$ с катастрофами в более общем случае, когда интенсивности катастроф зависят от числа требований в системе обслуживания. При этом удастся получить достаточно общие условия, гарантирующие наличие слабой эргодичности для процесса, описывающего число требований в такой системе обслуживания, и получить оценки скорости сходимости, гарантирующие возможность приближенного построения предельных характеристик системы.

Обозначим через $X = X(t)$, $t \geq 0$, число требований в момент t для описываемой модели. Тогда $X = X(t)$ является процессом рождения и гибели с катастрофами. Интенсивности рождения, гибели и катастрофы для процесса в случае, если $X(t) = n$, есть $\lambda_n(t) = \lambda(t)$, $\mu_n(t) = \min(n, S)\mu(t)$ и $\xi_n(t) = \zeta_n \xi(t)$ соответственно.

Обозначим через $p_{ij}(s, t)$ вероятности перехода, а через $p_i(t)$ — вероятности состояний для процесса $X = X(t)$:

$$\begin{aligned} p_{ij}(s, t) &= Pr \{X(t) = j | X(s) = i\}; \\ p_i(t) &= Pr \{X(t) = i\} \end{aligned}$$

Тогда для описания процесса получаем прямую систему Колмогорова

$$\left. \begin{aligned} \frac{dp_0}{dt} &= -\lambda_0(t)p_0 + \mu_1(t)p_1 + \sum_{k \geq 1} \xi_k(t)p_k; \\ \frac{dp_k}{dt} &= \lambda_{k-1}(t)p_{k-1} - (\lambda_k(t) + \mu_k(t) + \xi_k(t))p_k + \mu_{k+1}(t)p_{k+1}, \quad k \geq 1. \end{aligned} \right\} (1)$$

Пусть $\mathbf{p}(t) = (p_0(t), p_1(t), \dots)^T$, $t > 0$, — вектор-столбец вероятностей состояний процесса, а $A(t) = \{a_{ij}(t), t \geq 0\}$ — матрица системы (1).

В работе будет предполагаться, что интенсивности поступления и обслуживания требований $\lambda(t)$ и $\mu(t)$ локально интегрируемы на $[0; \infty)$. Будем считать «базисную» интенсивность катастрофы $\xi(t)$ локально интегрируемой на $[0; \infty)$, а коэффициенты состояний — ограниченными, т. е. $0 \leq \zeta_n \leq M$ при некотором $M < \infty$. Тогда систему (1) можно рассматривать как дифференциальное уравнение

$$\frac{d\mathbf{p}}{dt} = A(t) \mathbf{p}, \quad t \geq 0,$$

¹Вологодский государственный педагогический университет, Институт проблем информатики РАН и ВНКЦ ЦЭМИ РАН, a_zeifman@mail.ru

²Вологодский государственный педагогический университет, yacovi@mail.ru

³Вологодский государственный педагогический университет, a_korotyshva@mail.ru

⁴Вологодский государственный педагогический университет, nataliya_tereshi@mail.ru

в пространстве последовательностей l_1 с ограниченной почти при всех $t \geq 0$ локально интегрируемой оператор-функцией $A(t)$. Следовательно, можно применять общий подход, предложенный впервые в заметке [7] и развитый затем в [8, 9]. Метод опирается на две основные составляющие: понятие и оценки, связанные с логарифмической нормой операторной функции, и специальные преобразования редуцированной матрицы интенсивностей рассматриваемой марковской цепи, и позволяет получать явные и точные оценки.

2 Слабая эргодичность

Напомним основные определения.

Марковскую цепь $X(t)$ назовем *слабо эргодичной*, если $\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\| \rightarrow 0$ при $t \rightarrow \infty$ для любых начальных распределений вероятностей состояний $\mathbf{p}^*(0), \mathbf{p}^{**}(0)$ (здесь и далее через $\|\mathbf{x}\|$ обозначена l_1 -норма).

Положим $E_k(t) = E\{X(t) | X(0) = k\}$ — математическое ожидание процесса в момент t при условии, что в нулевой момент времени он находится в состоянии k , иногда будет встречаться также несколько более общее выражение

$$E_{\mathbf{p}}(t) = \sum_{k \geq 0} E\{X(t) | X(0) = k\} p_k(0).$$

Будем говорить, что марковская цепь $X(t)$ имеет *предельное среднее* $\varphi(t)$, если

$$\lim_{t \rightarrow \infty} (\varphi(t) - E_k(t)) = 0$$

при любом k .

В настоящей работе будет изучаться слабая эргодичность и сопутствующие свойства числа требований в рассматриваемой системе обслуживания для следующих важных ситуаций:

- (1) интенсивности катастроф существенны при любой длине очереди;

- (2) интенсивность обслуживания требований достаточно велика;
- (3) достаточно велика интенсивность поступления требований, а интенсивности катастроф существенны для ситуаций, когда длина очереди (количество требований в системе) пропорционально некоторому натуральному числу.

Теорема 1. Пусть

$$\inf_n \zeta_n = \zeta > 0 \tag{2}$$

и, кроме того, найдется $\varepsilon > 0$ такое, что

$$\int_0^{\infty} (\zeta \xi(t) - \varepsilon \lambda(t)) dt = +\infty. \tag{3}$$

Тогда процесс $X(t)$ слабо эргодичен и имеет предельное среднее. Если при этом в качестве предельного режима и предельного среднего выбрать режим $\pi(t)$ и среднее $\phi(t)$, соответствующие начальному условию $X(0) = 0$, то при любом начальном условии вида $X(0) = k$ справедливы следующие оценки:

$$\|\mathbf{p}(t) - \pi(t)\| \leq 4(1 + \varepsilon)^k \varepsilon^{-1} e^{-\int_0^t (\zeta \xi(\tau) - \varepsilon \lambda(\tau)) d\tau} \tag{4}$$

и

$$|E_k(t) - E_0(t)| \leq \frac{4(1 + \varepsilon)^k}{\varepsilon \omega} e^{-\int_0^t (\zeta \xi(\tau) - \varepsilon \lambda(\tau)) d\tau}.$$

Доказательство. Полагая

$$p_0(t) = 1 - \sum_{i \geq 1} p_i(t),$$

получаем аналогично тому, как это проделано в [6], следующую систему:

$$\frac{d\mathbf{z}(t)}{dt} = B(t)\mathbf{z}(t) + \mathbf{f}(t), \tag{5}$$

в которой

$$B(t) = \begin{pmatrix} -(\lambda_0 + \lambda_1 + \mu_1 + \xi_1) & \mu_2 - \lambda_0 & -\lambda_0 & -\lambda_0 & \dots & \dots \\ \lambda_1 & -(\lambda_2 + \mu_2 + \xi_2) & \mu_3 & 0 & 0 & \dots \\ 0 & \lambda_2 & -(\lambda_3 + \mu_3 + \xi_3) & \mu_4 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix};$$

$$\mathbf{z}(t) = (p_1(t), p_2(t), \dots)^T, \quad \mathbf{f}(t) = (\lambda_0(t), 0, 0, \dots)^T.$$

Обозначим оператор Коши линейной неоднородной системы (5) через $V(t, z)$, введем в рассмотрение матрицу

$$D = \begin{pmatrix} d_0 & d_0 & d_0 & \cdots \\ 0 & d_1 & d_1 & \cdots \\ 0 & 0 & d_2 & \cdots \\ \vdots & \vdots & \cdots & \ddots \end{pmatrix}$$

и пространство последовательностей $\ell_{1D} = \{z = (p_1, p_2, \dots)\}$ таких, что $\|z\|_{1D} = \|Dz\|_1 < \infty$, где d_i — пока не выбранные положительные числа.

Теперь рассматриваемый процесс можно исследовать таким же образом, как и в [6].

Положим

$$\alpha_k(t) = \lambda_k(t) + \mu_{k+1}(t) + \xi_{k+1}(t) - \frac{d_{k+1}}{d_k} \lambda_{k+1}(t) - \frac{d_{k-1}}{d_k} \mu_k(t), \quad k \geq 0,$$

и

$$\alpha(t) = \inf_{k \geq 0} \alpha_k(t).$$

Тогда получаем

$$\gamma(B(t))_{1D} = \sup_{i \geq 0} \left(\frac{d_{i+1}}{d_i} \lambda_{i+1}(t) - (\lambda_i(t) + \mu_{i+1}(t) + \xi_{i+1}(t)) + \frac{d_{i-1}}{d_i} \mu_i(t) \right) = -\alpha(t). \quad (6)$$

Полагая теперь $d_{-1} = d_0 = 1$, $d_{k+1} = (1 + \varepsilon)d_k$, $k \geq 0$, получаем с учетом (6) следующую оценку:

$$\begin{aligned} \gamma(B(t))_{1D} &\leq \\ &\leq - \left(\zeta \xi(t) + \frac{\varepsilon}{1 + \varepsilon} (S\mu(t) - (1 + \varepsilon)\lambda(t)) \right) \leq \\ &\leq - (\zeta \xi(t) - \varepsilon \lambda(t)) = -\alpha_*(t), \quad (7) \end{aligned}$$

причем из (2) вытекает, что

$$\int_0^\infty \alpha_*(t) dt = +\infty.$$

Теперь с учетом проведенного ранее (см., например, [8]) сравнения норм получаем слабую эргодичность процесса и оценку

$$\begin{aligned} \|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\| &\leq \\ &\leq 4e^{-\int_0^t (\zeta \xi(\tau) - \varepsilon \lambda(\tau)) d\tau} \|\mathbf{p}^*(0) - \mathbf{p}^{**}(0)\|_{1D}. \end{aligned}$$

Выбирая теперь $\mathbf{p}^*(0) = \pi(0) = \mathbf{e}_0$, $\mathbf{p}^{**}(0) = \mathbf{p}(0) = \mathbf{e}_k$, получаем неравенство (4).

Рассматривая величину $\omega = \inf_{k \geq 1} d_k/k$ и убеждаясь, что $\omega = \inf_{k \geq 1} (1 + \varepsilon)^k / k > 0$, получаем снова из сравнения норм существование предельного среднего и оценку

$$\begin{aligned} |E_{\mathbf{p}^*}(t) - E_{\mathbf{p}^{**}}(t)| &\leq \\ &\leq \frac{4}{\omega} e^{-\int_0^t (\zeta \xi(\tau) - \varepsilon \lambda(\tau)) d\tau} \|\mathbf{p}^*(0) - \mathbf{p}^{**}(0)\|_{1D}. \end{aligned}$$

Для завершения доказательства достаточно выбрать снова $\mathbf{p}^*(0) = \pi(0) = \mathbf{e}_0$, $\mathbf{p}^{**}(0) = \mathbf{p}(0) = \mathbf{e}_k$.

Замечание 1. Пусть все интенсивности (поступления, обслуживания требований, катастроф) 1-периодичны. Тогда:

- (1) вместо условия (3) достаточно потребовать, чтобы $\int_0^1 \xi(t) dt > 0$;
- (2) можно выбрать «особые» предельные характеристики, а именно существует 1-периодический предельный режим $\pi(t)$ и соответствующее ему 1-периодическое предельное среднее $\phi(t)$;
- (3) эти предельные характеристики можно приближенно построить, пользуясь методикой, разработанной в [6, 8], с использованием усеченного процесса и оценок скорости сходимости.

Теорема 2. Пусть при некотором $\varepsilon > 0$ выполняется условие

$$\int_0^\infty (S\mu(t) - (1 + \varepsilon)\lambda(t)) dt = +\infty. \quad (8)$$

Тогда процесс $X(t)$ слабо эргодичен и имеет предельное среднее. Если при этом в качестве предельного режима и предельного среднего выбрать режим $\pi(t)$ и среднее $\phi(t)$, соответствующие начальному условию $X(0) = 0$, то при любом начальном условии вида $X(0) = k$ справедливы следующие оценки:

$$\begin{aligned} \|\mathbf{p}(t) - \pi(t)\| &\leq \\ &\leq 4(1 + \varepsilon)^k e^{-1} e^{-\int_0^t \varepsilon(1 + \varepsilon)^{-1} (S\mu(\tau) - (1 + \varepsilon)\lambda(\tau)) d\tau} \end{aligned}$$

и

$$\begin{aligned} |E_k(t) - E_0(t)| &\leq \\ &\leq \frac{4(1 + \varepsilon)^k}{\varepsilon \omega} e^{-\int_0^t \varepsilon(1 + \varepsilon)^{-1} (S\mu(\tau) - (1 + \varepsilon)\lambda(\tau)) d\tau}. \end{aligned}$$

Доказательство проводится так же, как в предыдущей теореме. При этом, строя то же вспомогательное пространство ℓ_{1D} , вместо (7) получаем неравенство

$$\begin{aligned} \gamma(B(t))_{1D} &\leq \\ &\leq - \left(\zeta \xi(t) + \frac{\varepsilon}{1+\varepsilon} (S\mu(t) - (1+\varepsilon)\lambda(t)) \right) \leq \\ &\leq - \frac{\varepsilon}{1+\varepsilon} (S\mu(t) - (1+\varepsilon)\lambda(t)) = -\alpha_*(t), \end{aligned}$$

и далее оценки получаются тем же образом.

Замечание 2. Если все интенсивности (поступления, обслуживания требований, катастроф) 1-периодичны, то:

- (1) вместо условия (8) достаточно потребовать, чтобы $\int_0^1 (S\mu(t) - \lambda(t)) dt > 0$;
- (2) можно выбрать «особые» предельные характеристики, а именно существует 1-периодический предельный режим $\pi(t)$ и соответствующее ему 1-периодическое предельное среднее $\phi(t)$;
- (3) эти предельные характеристики можно приближенно построить, пользуясь методикой, разработанной в [6, 8], с использованием усеченного процесса и оценок скорости сходимости.

Рассмотрим, наконец, ситуацию, когда условия предыдущих теорем не выполняются.

Теорема 3. Пусть при некотором натуральном N выполняется условие

$$\inf_n \zeta_{nN} = \zeta > 0$$

вместо (2). Пусть, кроме того, найдется $\delta > 0$ такое, что

$$\int_0^\infty g(t) dt = +\infty,$$

где

$$g(t) = \min (\zeta \xi(t) - \delta \lambda(t) - \delta S\mu(t), \lambda(t) - (1 + \delta) S\mu(t)).$$

Тогда процесс $X(t)$ слабо эргодичен и имеет предельное среднее.

Доказательство проводится по той же схеме, что и в теореме 1. При этом вспомогательная

последовательность $\{d_k\}$ выбирается следующим образом: $d_{-1} = d_0 = 1$, $d_{k+1} = (1 + \varepsilon)^{-1} d_k$, если $k \neq iN - 1$, $d_{k+1} = (1 + \varepsilon)^N d_k$ при $k = iN - 1$, где положительное число $\varepsilon < \delta$ таково, что $(1 + \varepsilon)^N - 1 < \delta$. Тогда получаем

$$\begin{aligned} \alpha_k(t) &\geq \lambda(t) + S\mu(t) + \zeta_{k+1} \xi(t) - \\ &- S\mu(t)(1 + \varepsilon) - (1 + \varepsilon)^{-1} \lambda(t) \geq \\ &\geq \frac{\varepsilon}{1 + \varepsilon} (\lambda(t) - (1 + \varepsilon) S\mu(t)) \end{aligned}$$

при $k \neq iN - 1$ и

$$\alpha_k(t) \geq \zeta \xi(t) - \varepsilon S\mu(t) - ((1 + \varepsilon)^N - 1) \lambda(t)$$

при $k = iN - 1$.

Отсюда вытекает, что

$$\gamma(B(t))_{1D} \leq -g(t),$$

причем последовательность $\{d_k\}$ построена так, что выполняется и условие $\omega = \inf_{k \geq 1} d_k/k > 0$. Отсюда и следует утверждение теоремы.

Следствие 1. Если при выполнении условий теоремы 3 в качестве предельного режима и предельного среднего выбрать режим $\pi(t)$ и среднее $\phi(t)$, соответствующие начальному условию $X(0) = 0$, то при любом начальном условии вида $X(0) = k$ справедливы следующие оценки:

$$\|\mathbf{p}(t) - \pi(t)\| \leq \frac{4r_k}{d} e^{-\int_0^t G(\tau) d\tau}$$

и

$$|E_k(t) - E_0(t)| \leq \frac{4r_k}{d\omega} e^{-\int_0^t G(\tau) d\tau},$$

где $d = \min d_k$, $r_k = \sum_{i \leq k-1} d_i$, а

$$\begin{aligned} G(t) &= \min \left(\zeta \xi(t) - \varepsilon S\mu(t) - ((1 + \varepsilon)^N - 1) \lambda(t), \right. \\ &\left. \frac{\varepsilon}{1 + \varepsilon} (\lambda(t) - (1 + \varepsilon) S\mu(t)) \right) \geq g(t). \end{aligned}$$

Замечание 3. В случае 1-периодических интенсивностей за счет более сложного выбора вспомогательной последовательности $\{d_k\}$ условия теоремы можно существенно ослабить.

3 Пример

Рассмотрим систему обслуживания $M(t)/M(t)/100$ с катастрофами в случае периодических интенсивностей в разных ситуациях, гарантирующих слабую эргодичность модели, и вычислим предельное среднее $\phi(t)$, а также асимптотику величины $J_0(t)$ — вероятности того, что в момент t очередь пуста, т. е. в системе обслуживания нет ни одного требования, и величины $J_{10}(t) = \Pr(X(t) \leq 10)$.

При вычислениях используется следствие 4 из работы [6], которое с учетом конкретной модели и зависимости интенсивностей катастроф от состояния выглядит следующим образом.

Следствие 2. Пусть выполнены условия теоремы 1, $X_n(t)$ — соответствующий усеченный процесс, а $X(0) = X_n(0) = 0$. Тогда при всех $t \geq 0$ и любом n справедливы оценки

$$\|\pi(t) - \mathbf{p}_n(t)\|_1 \leq \limsup_{t \rightarrow \infty} \|\pi(t)\|_{1D} \left(4 \frac{(1 + \varepsilon)^k}{\varepsilon} e^{-\int_0^t (\zeta \xi(\tau) - \varepsilon \lambda(\tau)) d\tau} + 6Lw_n^1 t \right)$$

и

$$|\phi(t) - E_{0,n}(t)| \leq \limsup_{t \rightarrow \infty} \|\pi(t)\|_{1D} \left(\frac{4(1 + \varepsilon)^k}{\omega \varepsilon} e^{-\int_0^t (\zeta \xi(\tau) - \varepsilon \lambda(\tau)) d\tau} + 6Lw_n^2 t \right),$$

где

$$L = \sup_{[0,1]} (\lambda(t) + S\mu(t) + M\xi(t));$$

$$w_n^1 = \sup_{k \geq n} \frac{1}{d_k};$$

$$w_n^2 = \sup_{k \geq n} \frac{k}{d_k};$$

$$E_{0,n} = E_k(t) = E\{X_n(t) | X_n(0) = 0\},$$

а $\pi(t)$ — существующий 1-периодический предельный режим.

Замечание 4. Аналогичные оценки (с изменением первого слагаемого в скобках) получаются при выполнении условий теорем 2 и 3.

1. Выбираем вначале интенсивности следующим образом: $\lambda(t) = 240 + \cos 2\pi t$, $\mu(t) = 1 + \sin 4\pi t$, $\xi_n(t) = 100 + \sin 4\pi t$, $n \geq 1$. Тогда при $\varepsilon = 0,4$

выполнены условия теоремы 1 и для построения предельных 1-периодических характеристик (с точностью до 10^{-5}) достаточно выбрать размерность усеченного процесса $n = 100$ и построить нужные характеристики для усеченного процесса с нулевым начальным условием на отрезке [8, 9]. Соответствующие графики приведены на рис. 1.

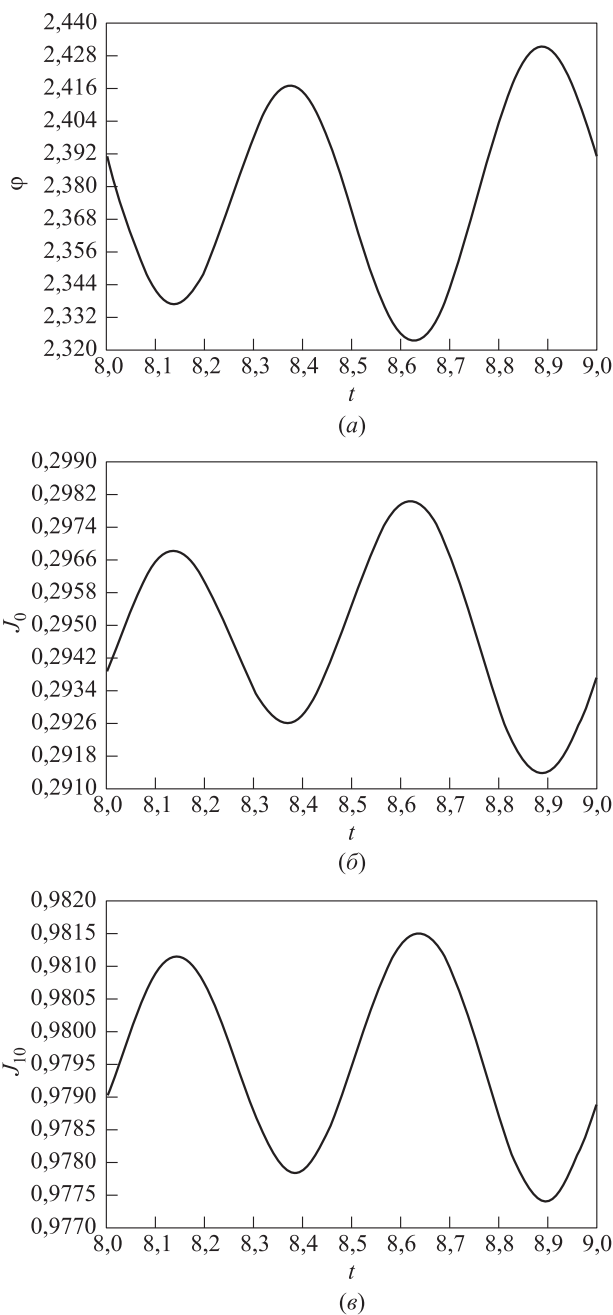


Рис. 1 Случай 1: (а) приближенное значение предельного среднего; (б) приближенное значение предельной величины $J_0(t) = \Pr(X(t) = 0)$; (в) приближенное значение предельной величины $J_{10}(t) = \Pr(X(t) \leq 10)$

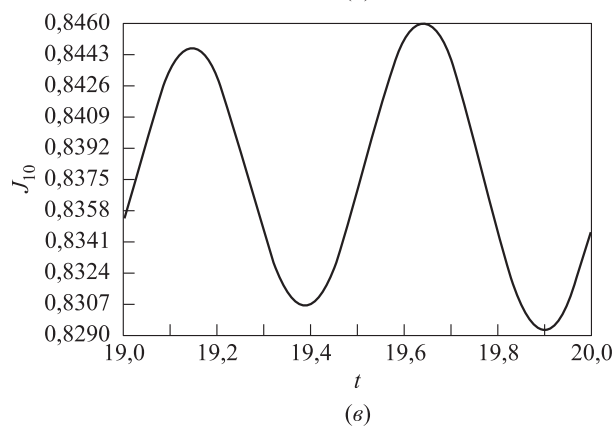
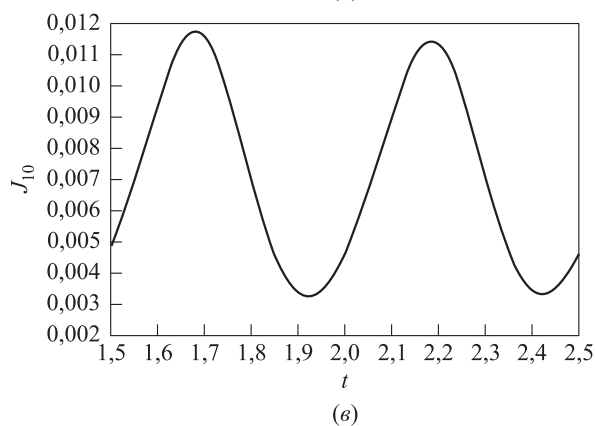
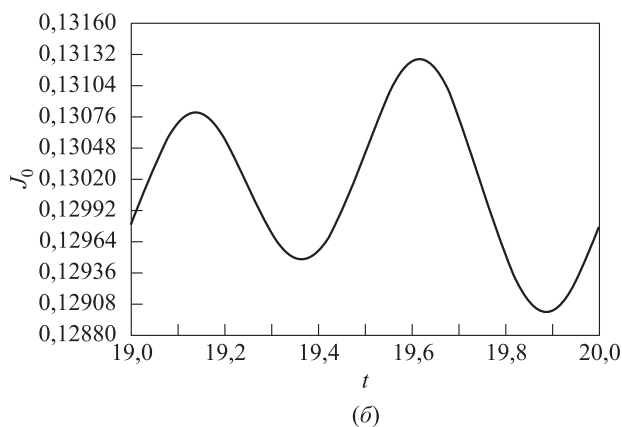
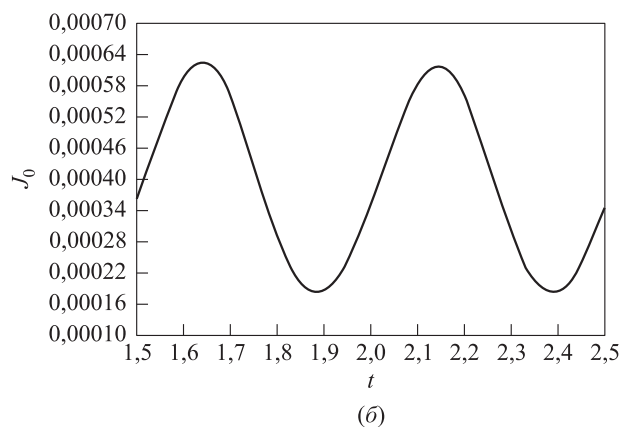
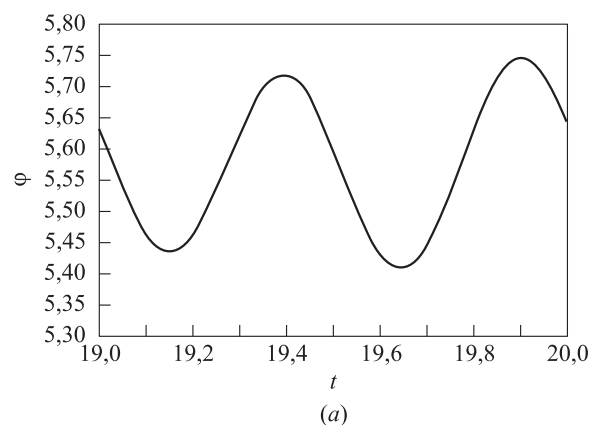
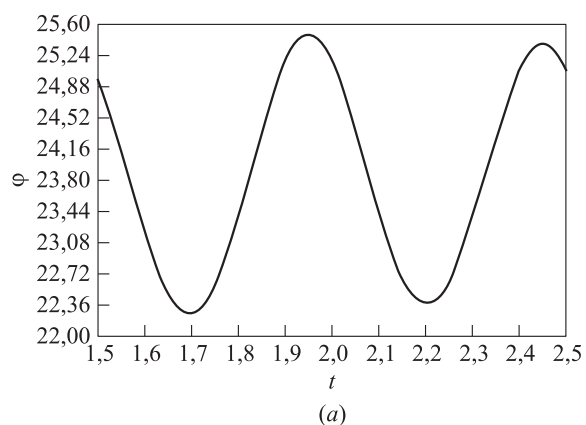


Рис. 2 Случай 2: (а) приближенное значение предельного среднего; (б) приближенное значение предельной величины $J_0(t) = \Pr(X(t) = 0)$; (в) приближенное значение предельной величины $J_{10}(t) = \Pr(X(t) \leq 0)$

Рис. 3 Случай 3: (а) приближенное значение предельного среднего; (б) приближенное значение предельной величины $J_0(t) = \Pr(X(t) = 0)$; (в) приближенное значение предельной величины $J_{10}(t) = \Pr(X(t) \leq 0)$

2. Выбирая теперь

$$\begin{aligned} \lambda(t) &= 240 + \cos 2\pi t; \\ \mu(t) &= 10 + \sin 4\pi t; \\ \xi_n(t) &= \frac{2 + \sin 4\pi t}{n}, \end{aligned}$$

получаем выполнение условий теоремы 2 при $\varepsilon = 3,0$, при этом $n = 65$, а в качестве необходимого отрезка будет $[1.5, 2.5]$ (рис. 2).

3. Если, наконец, взять

$$\begin{aligned} \lambda(t) &= 240 + \cos 2\pi t; \\ \mu(t) &= 1 + \sin 4\pi t; \\ \xi_n(t) &= \begin{cases} 0, & \text{если } n \neq 3k, \\ 153 + \sin 4\pi t, & \text{если } n = 3k, \end{cases} \end{aligned}$$

то выполнены условия теоремы 3, при этом $\varepsilon = 0,15$, $n = 250$, $t \in [19, 20]$ (рис. 3).

Литература

1. *Krishna Kumar B., Arivudainambi D.* Transient solution of an $M/M/1$ queue with catastrophes // *Comput. Math. Appl.*, 2000. Vol. 40. P. 1233–1240.
2. *Di Crescenzo A., Giorno V., Nobile A. G., Ricciardi L. M.* On the $M/M/1$ queue with catastrophes and its continuous approximation // *Queueing Syst.*, 2003. Vol. 43. P. 329–347.
3. *Van Doorn E. A., Zeifman A.* Extinction probability in a birth–death process with killing // *J. Appl. Probab.*, 2005. Vol. 42. P. 185–198.
4. *Di Crescenzo A., Giorno V., Nobile A. G., Ricciardi L. M.* A note on birth–death processes with catastrophes // *Statist. Probab. Lett.*, 2008. Vol. 78. P. 2248–2257.
5. *Zeifman A., Satin Ya., Chegodaev A., Bening V., Shorgin V.* Some bounds for $M(t)/M(t)/S$ queue with catastrophes // *SMCTools08 Proceedings*. Athens, Greece, 2008.
6. *Зейфман А. И., Сатин Я. А., Чегодаев А. В.* О нестационарных системах обслуживания с катастрофами // *Информатика и её применения*, 2009. Т. 3. Вып. 1. С. 47–54.
7. *Гнеденко Б. В., Макаров И. П.* Свойства решений задачи с потерями в случае периодических интенсивностей // *Дифф. уравнения*, 1971. Т. 7. С. 1696–1698.
8. *Zeifman A., Leorato S., Orsingher E., Satin Ya., Shilova G.* Some universal limits for nonhomogeneous birth and death processes // *Queueing Syst.*, 2006. Vol. 52. P. 139–151.
9. *Зейфман А. И., Бенинг В. Е., Соколов И. А.* Марковские цепи и модели с непрерывным временем. — М.: ЭЛЕКС-КМ, 2008.

АСИМПТОТИКИ ВЕРОЯТНОСТЕЙ БОЛЬШИХ УКЛОНЕНИЙ СТАЦИОНАРНОЙ ОЧЕРЕДИ*

Е. В. Морозов¹

Аннотация: Точные аналитические результаты доступны только для сравнительно узкого класса систем обслуживания, и поэтому асимптотические методы анализа оказываются полезным инструментом исследования и оптимизации современных коммуникационных сетей со сложной структурой потоков данных. Статья представляет собой обзор основных асимптотических результатов, играющих важную роль в оценке качества обслуживания (QoS) стационарных систем. Рассматриваются асимптотики вероятности превышения процессом нагрузки/очереди растущего уровня в случае, когда время обслуживания имеет тяжелый хвост. Аналогичные результаты даны для систем с входным процессом Леви, где время обслуживания имеет легкий хвост. Доказательства базируются на методах теории больших уклонений (ТБУ), которые подробно иллюстрируются на примере системы $M/M/1$. Рассмотрена асимптотика вероятности переполнения на цикле регенерации, в том числе и для многоканальной системы. Приведен асимптотический анализ систем, где входной процесс обладает долговременной зависимостью (долгой памятью), причем основное внимание уделено фрактальному броуновскому процессу. Обсуждаются связи между долговременной зависимостью процесса очереди и моментными свойствами вложенного процесса регенераций.

Ключевые слова: стационарная очередь; вероятности больших уклонений; асимптотический анализ; распределения с легким хвостом; фрактальный броуновский процесс; процесс с долговременной зависимостью; регенерация

1 Введение

Статья содержит обзор основных результатов по асимптотике вероятностей превышения растущего уровня стационарным процессом нагрузки W (времени ожидания в очереди) и стационарной очередью ν в системах вида $GI/G/1$, где либо (1) время обслуживания S имеет так называемый тяжелый хвост, либо (2) S имеет легкий хвост, а также (3) где входной процесс обладает долговременной зависимостью.

Случаи 1 и 3 относительно недавно привлекли особый интерес исследователей ввиду обнаружения тесных связей со свойствами трафиков современных телекоммуникационных сетей [1]. В обоих этих случаях асимптотики имеют, как правило, достаточно простой вид, а в случае 2 требуется привлечение методов ТБУ и параметры асимптотического представления не выражаются непосредственно в терминах заданных распределений (в отличие от случаев 1 и 3). Асимптотические результаты для случая 1 были использованы для оценки качества статистического оценивания вероятностей редких событий, получаемых в рамках регенеративной модификации метода расщепления [2]. Численное моделирование дало хорошее согласие с этими (и

известными аналитическими) результатами и позволяет утверждать, что данный метод дает надежные оценки (в том числе интервальные) для вероятностей больших уклонений в регенеративных системах обслуживания. В случае 2 асимптотический анализ связан с вычислением так называемой *функции интенсивности* методами ТБУ и является в типичных ситуациях весьма трудным. Возможность расширения предложенного метода для прямого оценивания параметров асимптотик в регенеративных системах является важным мотивом данной работы и одной из целей дальнейшего исследования.

Основное внимание в статье уделено математическим аспектам асимптотического анализа, однако с трактовкой результатов в терминах конкретных сетевых характеристик. Главная цель работы состоит в том, чтобы представить наиболее важные асимптотические результаты в форме, удобной для специалистов, работающих в области анализа сетей и телекоммуникационных технологий. Это является актуальным, поскольку многочисленные результаты, накопленные к настоящему времени в разных источниках, зачастую представлены в несогласованном виде и в трудной для практического использования форме.

* Работа поддерживается РФФИ, грант 07-07-00088.

¹ Институт прикладных математических исследований КарНЦ РАН, emorozov@karelia.ru

В разд. 2 рассмотрены асимптотики вероятности большого значения процесса нагрузки/очереди в системе $GI/G/1$, когда время обслуживания имеет тяжелый хвост. Раздел 3 посвящен системам, где обслуживание имеет легкий хвост, а входной процесс — либо процесс Леви, либо процесс восстановления. Подход на основе ТБУ подробно иллюстрируется на примере системы $M/M/1$. Здесь же приведена асимптотика вероятности превышения очередью растущего уровня на цикле регенерации, в том числе и для многоканальной стационарной системы $GI/G/m$. В разд. 4 рассматриваются системы, где входной поток обладает долговременной зависимостью, в частности является фрактальным броуновским процессом. В разд. 5 обсуждается связь между долговременной зависимостью процесса очереди/нагрузки и моментными свойствами вложенного процесса регенераций в связи с возможностью применения регенеративного метода для оценивания параметров таких процессов.

2 Время обслуживания с тяжелым хвостом

Рассмотрим случайную величину (с.в.) X с функцией распределения (ф.р.) F и обозначим через $\bar{F} = 1 - F$ ее хвост. Через $F^{(n)}$ обозначим n -кратную свертку ф.р. F с собой. Будем писать $F(x) \sim G(x)$, если $F(x)/G(x) \rightarrow 1$ при $x \rightarrow \infty$. Функция распределения F (и сама с.в. X) имеет *тяжелый хвост* (записывается $F \in \mathcal{H}$), если $\bar{F}(x) > 0$, $x \geq 0$, и для любого фиксированного $y > 0$

$$\lim_{x \rightarrow \infty} P(X > x + y | X > x) = \lim_{x \rightarrow \infty} \frac{\bar{F}(x + y)}{\bar{F}(x)} = 1.$$

Заметим, что если $F \in \mathcal{H}$, то производящая функция моментов $Ee^{\varepsilon X} = \infty$ для любого $\varepsilon > 0$. Функция распределения F называется *субэкспоненциальной* ($F \in \mathcal{S} \subseteq \mathcal{H}$), если $\bar{F}(x) > 0$, $x \geq 0$, и для любого $n \geq 2$

$$\bar{F}^{(n)}(x) \sim n\bar{F}(x), \quad x \rightarrow \infty. \quad (1)$$

(На самом деле достаточно установить (1) для $n = 2$.) В действительности $e^{\varepsilon x} \bar{F}(x) \rightarrow \infty$, $x \rightarrow \infty$, для любого $\varepsilon > 0$, что поясняет название класса \mathcal{S} [3].

Рассмотрим систему $GI/G/1$ с входным потоком восстановления с независимыми, одинаково распределенными (н.о.р) интервалами $\{\tau_n\}$ с интенсивностью $\lambda = 1/E\tau$ и с н.о.р. временами обслуживания $\{S_n\}$ в предположении стационарности, $\rho := \lambda ES < 1$. (Здесь и далее опущены

индексы, обозначающие типичный элемент последовательности н.о.р.с.в.) Последовательность $\{W_n\}$ времен ожидания заявок в очереди является апериодической марковской цепью, удовлетворяющей рекурсии Линдли $W_{n+1} = (W_n + X_n)^+$, где $X_n = S_n - \tau_n$, $n \geq 1$. Поэтому существует предел $W_n \Rightarrow W$ (по распределению), причем *стационарная нагрузка* W распределена как максимум случайного блуждания, $W =_{st} \max_{n \geq 1} (X_1 + \dots + X_n)$ ($=_{st}$ означает равенство по распределению). Пусть с.в. S_e имеет ф.р. F_e

$$F_e(x) = \frac{1}{ES} \int_0^x \bar{F}(y) dy. \quad (2)$$

(Это ф.р. стационарного незавершенного времени восстановления в процессе, порожденном с.в. S_n .) Обозначим $G(x) = P(W \leq x)$. Основной результат формулируется следующим образом [5, 4]: если функция $F \in \mathcal{S}$, то

$$\bar{G}(x) := P(W > x) \sim \frac{\rho}{1 - \rho} \bar{F}_e(x), \quad x \rightarrow \infty. \quad (3)$$

Этот важный результат имеет простое качественное объяснение на примере системы $M/G/1$, где (в очевидных обозначениях) плотность G' удовлетворяет уравнению Такача

$$G'(x) = \lambda \int_{0^-}^x \bar{F}(x - y) dG(y), \quad x > 0 \quad (G(0) = 1 - \rho > 0). \quad (4)$$

Применяя преобразование Лапласа–Стилтьеса к обеим частям уравнения (4) и обозначая $g(\theta) = Ee^{-\theta W}$, $f_e(\theta) = Ee^{-\theta S_e}$ ($\theta > 0$), получим $g(\theta) = 1 - \rho + \rho f_e(\theta)g(\theta)$, или

$$g(\theta) = \frac{1 - \rho}{1 - \rho f_e(\theta)} = (1 - \rho) \sum_{n=0}^{\infty} (\rho f_e(\theta))^n.$$

Используя обратное преобразование и свойство свертки, получаем (3):

$$\begin{aligned} \bar{G}(x) &= (1 - \rho) \sum_{n=1}^{\infty} \rho^n \bar{F}_e^{(n)}(x) \sim \\ &\sim (1 - \rho) \rho \bar{F}_e(x) \sum_{n=1}^{\infty} \rho^{n-1} n = \\ &= \frac{\rho}{1 - \rho} \bar{F}_e(x), \quad x \rightarrow \infty. \quad (5) \end{aligned}$$

В (5) первое равенство является (*обобщенной*) формулой Поллачека–Хинчина, а эквивалентность (при

$x \rightarrow \infty$) опирается на (1). Таким образом, хвост ф.р. W можно выразить через частичные суммы слагаемых $\rho \bar{F}_e$, а затем через единственное слагаемое \bar{F}_e . В действительности \bar{F}_e является условным распределением *возрастающей строгой лестничной высоты* в случайном блуждании с шагом $S - \tau$ при условии выполнения события $A = \{\text{достижение очередной лестничной высоты}\}$, причем $P(A) = \rho$ [6]. В целом (5) отражает тот ключевой факт, что очень большое значение суммы с.в., имеющих распределение с тяжелым хвостом, вызвано очень большим значением одного слагаемого [3, 5, 7, 8].

Опираясь на [9], рассмотрим стационарную очередь ν (общее число заявок) в системе $M/G/1$, где по свойству PASTA ν может рассматриваться как в моменты прихода, так и в произвольный момент времени. (То же самое верно и для W .) В данном случае используется *обобщенный закон Литтла*, согласно которому стационарная очередь ν распределена как число приходов за время W [5]. (Это далеко идущий аналог известного *закона Литтла* для средних, $\lambda EW = E\nu$.) Если ф.р. времени обслуживания $F \in \mathcal{S}$ и, кроме того,

$$\lim_{x \rightarrow \infty} \frac{\bar{F}(xe^{y/\sqrt{x}})}{\bar{F}(x)} = 1, \quad \text{локально равномерно по } y \geq 0, \quad (6)$$

то вероятность превышения уровня k стационарной очередью ν удовлетворяет асимптотическому соотношению [9]

$$P(\nu \geq k) \sim \frac{\rho}{1 - \rho} \bar{F}_e\left(\frac{k}{\lambda}\right), \quad k \rightarrow \infty. \quad (7)$$

Отметим, что (3) имеет место и для стационарного времени пребывания заявки в системе $D =_{st} W + S$, т. е. $P(D > x) \sim P(W > x)$. (Это следствие того, что хвост \bar{F}_e *тяжелее*, чем \bar{F} [3].) Соотношение (6) справедливо для распределения Парето, имеющего (стандартный) вид $\bar{F}(x) = x^{-\alpha}$, $\alpha > 0$, $x \geq 1$, а также для распределения Вейбулла $\bar{F}(x) = \exp\{-x^\beta\}$, $x \geq 0$, с *тяжелым хвостом* $\beta \in (0, 1/2)$. Однако (6) не верно, если $F \in \mathcal{H} \setminus \mathcal{S}$, например для распределения Вейбулла с *умеренно тяжелым хвостом* $1/2 \leq \beta < 1$, [9]. Тем не менее, в этом случае

$$P(\nu > k) \sim \frac{1}{\beta} \left(\frac{k}{\lambda}\right)^{1-\beta} \exp\left\{-\beta \left(\frac{k}{\lambda}\right) - (1-\beta)t^\beta\right\}, \quad k \rightarrow \infty,$$

где $t = t(k)$ есть решение уравнения $\beta t^\beta + \lambda t = k$. (При $\beta = 1/2$ явное решение этого квадратного уравнения приведено в [9].)

В [10] доказано, что в системе $GI/G/1$, где ф.р. времени обслуживания $F \in \mathcal{S}$, максимум $W^* := \max_{0 \leq k < \Delta} (W_k)$ стационарной нагрузки на цикле регенерации удовлетворяет асимптотическому соотношению

$$P(W^* > x) \sim \bar{F}(x)E\Delta, \quad x \rightarrow \infty, \quad (8)$$

где Δ — длина цикла регенерации. (Регенеративные аргументы показывают также, что (8) влечет (3).) В [11] этот результат перенесен на более сложный процесс обслуживания. Заметим, что в [12] предложены несколько иные условия для справедливости (7).

В заключение этого раздела сделаем несколько замечаний. Результаты об асимптотике стационарного времени ожидания в системе $GI/G/2$ с субэкспоненциальным временем обслуживания, приведенные в работе [13], достаточно громоздки и (за единственным исключением) имеют вид неравенств.

В *ациклической сети* для каждого узла i можно естественным образом определить множество Ω_i всех ему *предшествующих узлов*. Тогда (в очевидных обозначениях) если время обслуживания $S^{(i)}$ субэкспоненциально, то для стационарной нагрузки $W^{(i)}$ имеет место асимптотика вида (3)

$$P(W^{(i)} > x) \sim \frac{\rho_i}{1 - \rho_i} P(S_e^{(i)} > x),$$

если все времена обслуживания $S^{(j)} \in \Omega_i$ имеют более легкие хвосты, чем $S^{(i)}$ [14]. (Заметим, что это интуитивно ожидаемый результат.) Укажем также работу [15], где исследуется асимптотика сетей с субэкспоненциальными временами обслуживания, однако асимптотика сетей в данном обзоре не затрагивается.

В [16] показано, что если хвост ф.р. времени обслуживания в системе типа $M/G/1$ с *повторными вызовами* (и экспоненциальным временем ожидания на орбите) имеет вид $x^{-\alpha}L(x)$, где $\alpha > 1$, а функция L медленно меняется на бесконечности, то хвост ф.р. стационарного числа заявок в системе асимптотически эквивалентен хвосту ф.р. очереди в системе $M/G/1$ и имеет вид $x^{-(\alpha-1)}L(x)$, $x \rightarrow \infty$, см. (7).

3 Время обслуживания с легким хвостом

Опираясь в основном на работу [17], рассмотрим стационарную систему $GI/G/1$, в которой время

обслуживания S имеет легкий хвост, т. е. (в отличие от $S \in \mathcal{S}$) *производящая функция моментов*

$$Ee^{\theta S} < \infty \quad (9)$$

в некоторой положительной окрестности параметра $\theta = 0$. Приведем условия, при которых стационарная нагрузка W удовлетворяет соотношению

$$x^{-1} \log P(W > x) \rightarrow -\theta^*, \quad x \rightarrow \infty, \quad (10)$$

где показатель $\theta^* > 0$ является искомой величиной. Напомним, что типичное приращение процесса нагрузки имеет вид $X = S - \tau$. Учитывая независимость τ и S , введем *логарифмическую производящую функцию моментов*

$$\begin{aligned} \Lambda(\theta) &:= \log Ee^{\theta X} := \Lambda_S(\theta) + \Lambda_\tau(-\theta) = \\ &= \log Ee^{\theta S} + \log Ee^{-\theta\tau}, \quad \theta \geq 0. \end{aligned}$$

Условие стационарности $\rho = ES/E\tau < 1$ влечет $P(\tau > 0) > 0$, и поэтому $0 < Ee^{-\theta\tau} < 1$ и $-\infty < \log Ee^{-\theta\tau} < 0$. Так что условие $-\infty < \Lambda(\theta) < \infty$ обеспечивается условием (9). Основным результатом является следующий. Параметр θ^* в (10) оказывается единственным положительным решением уравнения

$$\Lambda(\theta) = 0.$$

(Параметр θ^* часто определяют как решение уравнения $Ee^{\theta X} = 1$ [17].)

Определенная проблема при применении ТБУ (кроме вычисления собственно параметра θ^*) состоит в том, что этот подход не позволяет выяснить, является ли асимптотика хвоста $P(W > x)$ в точности экспоненциальной или содержит корректирующие множители, скажем, вида

$$P(W > x) = g(x)e^{-\theta^*x+f(x)},$$

где функции g, f таковы, что $\log g(x) = o(x)$, $f(x) = o(x)$, $x \rightarrow \infty$. Например, в [17] упоминается система $M/G/1$ (из [18]), где $P(W > x) \sim cx^{-3/2} \exp\{-\theta^*x\}$. В этом случае (10) имеет место, а ожидаемая сходимост $e^{\theta^*x}P(W > x) \rightarrow const > 0$ не верна. Это обстоятельство демонстрируется ниже на примере системы $M/M/1$, и его надо иметь в виду при оценивании показателя θ^* .

3.1 Асимптотика в системе $M/M/1$

В качестве иллюстрации применения методов ТБУ рассмотрим асимптотику процессов нагрузки W и очереди ν в стационарной системе $M/M/1$ с интенсивностью входного потока λ и интенсивностью обслуживания μ , полагая $\rho := \lambda/\mu < 1$. Хорошо известно, что

$$P(W \geq x) = \rho e^{-(\mu-\lambda)x}, \quad x \geq 0. \quad (11)$$

Поэтому логарифмическая асимптотика (10) принимает вид

$$\lim_{x \rightarrow \infty} \frac{1}{x} \log P(W \geq x) := -\theta^* = -(\mu - \lambda), \quad (12)$$

разумеется, не улавливая постоянный множитель ρ . Теперь используем ТБУ, не опираясь на аналитический вид (11). Предполагая $\theta < \mu$, после несложных вычислений имеем

$$\Lambda(\theta) = \log Ee^{\theta(S-\tau)} = \log \left(\frac{\lambda\mu}{(\lambda+\theta)(\mu-\theta)} \right).$$

Решение уравнения $\Lambda(\theta) = 0$, равное $\theta^* = \mu - \lambda$, согласуется с (12). Такой результат можно получить, рассматривая динамику процесса нагрузки в дискретном времени [19]:

$$W(t+1) = (W(t) + a(t) - 1)^+, \quad t = 0, 1, \dots,$$

где $a(t)$ — нагрузка (суммарное время обслуживания заявок), поступившая в систему в интервале $[t, t+1)$. Поскольку входной поток пуассоновский, то $a(t) \stackrel{st}{=} a(0)$, $t \geq 0$, и, кроме того, $Ee^{\theta S} = \mu(\mu - \theta)^{-1}$. Поэтому

$$\begin{aligned} \Lambda(\theta) &= \log Ee^{\theta(a(0)-1)} = \log Ee^{\theta a(0)} - \theta = \\ &= \log \left\{ \sum_{k \geq 0} (Ee^{\theta S})^k e^{-\lambda} \frac{\lambda^k}{k!} \right\} - \theta = \\ &= \log \left\{ \exp \left\{ \frac{\mu\lambda}{\mu-\theta} - \lambda \right\} \right\} - \theta = \frac{\mu\lambda}{\mu-\theta} - \lambda - \theta. \end{aligned}$$

Уравнение $\Lambda(\theta) = 0$ имеет единственное решение $\theta^* = \mu - \lambda > 0$, которое согласуется как с (12), так и с результатом для системы с постоянной скоростью обслуживания $C = 1$ [19, с. 10–12].

Рассмотрим теперь вложенный (по моментам прихода) процесс очереди $\{\nu_n\}$, удовлетворяющий рекурсии

$$\nu_{n+1} = (\nu_n + 1 - D(n))^+, \quad n \geq 1, \quad (13)$$

где $D(n)$ — (потенциальное) число заявок, которое прибор мог бы обслужить в течение n -го интервала входного потока. Очевидно, с.в. $\{D(n), n \geq 1\}$ являются н.о.р. и $\nu_n \Rightarrow \nu, n \rightarrow \infty$. Понятно, что (13) можно трактовать как вариант рекурсии Линдли, где n -й шаг случайного блуждания равен $1 - D(n)$. Таким образом,

$$\begin{aligned} \Lambda(\theta) &:= \log Ee^{\theta(1-D(1))} = \\ &= \theta + \log \left\{ \lambda \sum_{k \geq 0} \int_0^\infty e^{-(\mu+\lambda)x} \frac{(\mu x e^{-\theta})^k}{k!} dx \right\} = \\ &= \theta + \log \frac{\lambda}{\mu + \lambda - \mu e^{-\theta}}. \end{aligned}$$

Уравнение $\Lambda(\theta) = 0$ имеет единственное положительное решение $\theta^* = -\log \rho > 0$, согласующееся с известным результатом $P(\nu \geq k) = \rho^k$, $k \geq 0$. Заметим, что здесь асимптотика $k^{-1} \log P(\nu > k)$ (при $k \rightarrow \infty$) точна в отмеченном выше смысле. Используя дискретизацию времени, можно аналогично рассмотреть процесс очереди $\nu(t)$ в непрерывном времени

$$\nu(t+1) = (\nu(t) + a(t) - d(t))^+, \quad t = 0, 1, \dots,$$

где $a(t)$ — число приходов, а $d(t)$ — потенциальное число уходов в интервале $[t, t+1)$. Очевидно, $a(t) - d(t) =_{st} a(0) - d(0)$, $t \geq 0$, и

$$\begin{aligned} \Lambda(\theta) &= \log E e^{\theta(a(0)-d(0))} = \\ &= \lambda(e^\theta - 1) - \mu(1 - e^{-\theta}), \end{aligned} \quad (14)$$

что снова дает решение $\theta^* = -\log \rho$.

3.2 Входной процесс Леви

В данном разделе рассматривается система типа $M/G/1$, где входной процесс имеет стационарные независимые приращения (*процесс Леви*), включая пуассоновский поток.

Рассмотрим систему $M/G/1$ с постоянной скоростью обслуживания $C > \lambda$, в которой стационарная нагрузка задается следующим образом [19]:

$$W =_{st} \sup_{t \geq 0} (A(t) - Ct), \quad (15)$$

где $A(t) = \sum_{i=0}^{t-1} a(i)$ — суммарная нагрузка, поступившая в интервале $[0, t)$. Соотношение (15) аналогично представлению W в системе $GI/G/1$ как максимума случайного блуждания, однако теперь независимые приращения на единичных интервалах $[t, t+1)$ имеют вид $a(t) - C$. Решение θ^* уравнения

$$\log E e^{\theta a(0)} = \theta C \quad (16)$$

имеет тот же вид $\theta^* = \mu - \lambda > 0$, что и в системе $M/M/1$.

В случае общего входного процесса Леви $A(t) = \sum_{i=0}^{t-1} a(i)$ предположим, что существует (конечный в некоторой окрестности $\theta = 0$) предел

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log E e^{\theta A(t)} := \Lambda_A(\theta) = \log E e^{\theta a(0)}.$$

(То есть входной поток удовлетворяет *принципу больших уклонений*.) Из (16) следует, что в данной системе параметр θ^* определяется из условия

$$\theta^* = \sup(\theta > 0 : \Lambda_A(\theta) \leq \theta C). \quad (17)$$

Приведем уточнение хвоста $P(W > x)$ в предположении, что решение θ^* уравнения $\Lambda_A(\theta) = \theta C$ принадлежит внутренности интервала, где $\Lambda_A(\theta) < \infty$ [20]. (Тогда производная $\Lambda'(\theta^*) < \infty$.) Обозначая $\phi(\theta) = \Lambda_A(\theta)/\theta$, получим

$$P(W > x) \sim \frac{C - \phi(0)}{\theta^* \Lambda'_A(\theta^*)} e^{-\theta^* x}, \quad x \rightarrow \infty,$$

где $C - \phi(0) = C - E a(0) > 0$ ввиду стационарности.

Рассмотрим входной процесс $A(t) = \lambda t + G(t)$ с детерминированной компонентой λt и гауссовской компонентой $G(t) = N(0, \sigma^2 t)$. (Как обычно, отрицательные значения $A(t)$ игнорируются [20].) Поскольку $E e^{\theta G(t)} = \exp\{(1/2)\theta^2 \sigma^2 t\}$, то для любого $t > 0$

$$\frac{1}{t} \log E e^{\theta A(t)} := \Lambda_A(\theta) = \lambda \theta + \frac{\theta^2 \sigma^2}{2}.$$

Поэтому показатель θ^* в (10) и (17) (напомним, что $C > \lambda$) равен

$$\theta^* = \frac{2(C - \lambda)}{\sigma^2}.$$

Другой полезный пример дает входной *обобщенный пуассоновский процесс* [6]

$$A(t) = \sum_{n=0}^{N(t)} Y_n,$$

где н.о.р. положительные с.в. $\{Y_n\}$ ($Y_0 = 0$) имеют распределение F , а $N(\cdot)$ — независимый пуассоновский процесс с параметром λ . Обозначая $E e^{\theta Y_1} = b(\theta)$ и используя равенство

$$1 - b(\theta) = \int (1 - e^{\theta x}) dF(x),$$

легко получить

$$E e^{\theta A(t)} = e^{-\lambda t(1-b(\theta))}.$$

Поэтому

$$\Lambda_A(\theta) = \frac{\log E e^{\theta A(t)}}{t} = \lambda \int_0^\infty (e^{\theta x} - 1) dF(x) \quad (t > 0).$$

Возможность решения уравнения $\Lambda_A(\theta) = \theta C$ определяется видом ф.р. F . Равенство

$$\frac{\Lambda_A(\theta^*)}{\theta^*} = C$$

определяет *эффективную пропускную способность* прибора (скорость) C , если ее необходимо найти,

например, чтобы гарантировать заданное качество обслуживания (QoS) в системе с *конечным буфером*. Такая гарантия состоит в том, что (предельная) доля потерь не должна превышать заранее заданной вероятности потери. С ростом размера буфера такая система сближается с системой с неограниченным буфером, на чем и основана возможность применения приведенной выше асимптотики [19–21]. Актуальной проблемой является построение достоверной оценки параметра C в случае, когда аналитическое выражение функции Λ_A получить трудно (или невозможно). Эта проблема рассматривается в работе [22].

Замечание 1. Асимптотический результат (10) верен (при дополнительных предположениях) и для более общего входного потока с эргодическими (зависимыми) приращениями [17].

3.3 Многоканальная система

Опираясь на [23], рассмотрим асимптотику вероятности того, что стационарная очередь в m -канальной системе $GI/G/m$ превысит высокий уровень *на цикле регенерации*. Рассмотрим входной процесс восстановления с интервалами τ , и пусть S — время обслуживания (в каждом из идентичных каналов). Предположим, что

$$\rho := \frac{ES}{E\tau} < m, \quad P(\tau > S) > 0$$

и что логарифмическая функция моментов

$$\Lambda_S(\theta) := \log Ee^{\theta S} < \infty$$

в некоторой окрестности $\theta = 0$. (То есть время обслуживания имеет легкий хвост.) Пусть ν_n означает число заявок в (стационарной) очереди в момент прихода заявки n . Обозначим через Δ длину цикла регенерации системы, т.е. интервал между двумя последовательными заявками, приходящими в полностью пустую систему. Пусть $\gamma_c(k) = P(\max_{1 \leq l < \Delta} \nu_l \geq k)$ — вероятность превышения (стационарной) очередью (целочисленного) уровня k на цикле регенерации. Обозначим $\Lambda(\theta) = \Lambda_\tau(-\theta m) + \Lambda_S(\theta)$. В [23] показано, что существует единственное решение $\theta^* > 0$ уравнения $\Lambda(\theta) = 0$ и что

$$\lim_{k \rightarrow \infty} \frac{1}{k} \log \gamma_c(k) = \Lambda_\tau(-\theta^* m).$$

Например, в системе $M/M/1$ имеем (используя технику ветвящихся процессов [2])

$$\gamma_c(k) = \frac{\rho^{k-1} - \rho^k}{1 - \rho^k} := \rho^{k-1} \phi(k), \quad k \geq 1,$$

где функция $\phi(k) \rightarrow 1 - \rho > 0$ при $k \rightarrow \infty$. Поэтому $\log \gamma_c(k) = (k-1) \log \rho + o(k)$. Следовательно, $\Lambda_\tau(-\theta^*) = \log \rho$, что совпадает с асимптотикой (14) превышения растущего уровня стационарной очереди в системе $M/M/1$.

Рассмотрим стационарную m -канальную систему $M/M/m$ при $\rho := \lambda/(m\mu) < 1$. Решая уравнение $\Lambda(\theta) = 0$, получим $\theta^* = (\mu m - \lambda)/m$, и поэтому

$$\lim_{k \rightarrow \infty} \frac{1}{k} \log \gamma_c(k) = \log Ee^{-\theta^* m \tau} = \log \rho.$$

В данной системе стационарная вероятность

$$P(\nu \geq k) = D\rho^k$$

для $k \geq m$ (где константа D хорошо известна в явном виде (см., например, [24])). Поэтому снова получаем совпадение логарифмических асимптотик вероятностей $\gamma_c(k)$ и $P(\nu \geq k)$ при $k \rightarrow \infty$.

Замечание 2. В более общей ситуации как нахождение аналитического решения θ^* уравнения $\Lambda(\theta) = 0$, так и последующее вычисление функции $\Lambda_\tau(-m\theta^*)$ является чрезвычайно сложной (или невозможной) задачей. Это ставит вопрос об эффективных методах оценивания параметров асимптотического представления, возможно через прямое оценивание малых вероятностей $\gamma_c(k)$. Отметим, что именно при оценивании вероятности γ_c ускоренный метод регенеративного расщепления показал свою эффективность [2].

4 Входной процесс с долговременной зависимостью

Рассмотрим асимптотику вероятностей больших уклонений для системы, где входной поток обладает *долговременной зависимостью*. В этом обзоре анализ ограничивается наиболее важным в современных сетевых исследованиях *фрактальным входным процессом*, определяемым следующим образом:

$$A(t) = \lambda t + B_H(t), \quad t \geq 0, \quad (18)$$

где $A(t)$ — суммарная нагрузка, поступившая в систему в интервале $[0, t)$, константа $\lambda > 0$, а $B_H(t)$ — *фрактальный броуновский процесс (ФБП)* с параметром Херста $H \in (1/2, 1)$ [25]. (Заметим, что процесс B_H можно задать на всей вещественной прямой.) Процесс B_H имеет непрерывные траектории, нормально распределенные стационарные приращения, причем

$$B_H(0) = 0, \quad EB_H(t) = 0, \\ E\{B_H(t)\}^2 = \sigma^2 t^{2H}, \quad t \geq 0.$$

Более того, процесс B_H обладает свойством *самоподобия*: $B_H(tT) \stackrel{st}{=} T^H B_H(t)$ для любого $T > 0$. Далее положим $\sigma^2 = E\{B_H(1)\}^2 = 1$ и отметим, что $B_{1/2}$ является стандартным броуновским движением. Хорошо известно (и легко подсчитать), что приращения ФБП на единичных интервалах $B_H^*(n) := B_H(n+1) - B_H(n)$, $n \geq 0$, образующие *фрактальный броуновский шум*, имеют (авто)ковариационную функцию

$$r(n) := \text{cov}(B_H^*(0), B_H^*(n)) \sim H(2H-1)n^{-2(1-H)}, \\ n \rightarrow \infty.$$

Поэтому

$$\sum_{n \geq 1} r(n) = \infty$$

и по определению процесс B_H обладает *долгой временной зависимостью*. Используя дискретную шкалу времени, положим далее $t = n$. Для случая независимых приращений дисперсия входного процесса растет линейно со временем: $DA(n) = nDa(1)$, в то время как для процесса (18) имеем $DA(n) = n^{2H}$, $n \geq 1$. Разложим (нормированную) логарифмическую производящую функцию моментов входного потока $\Lambda_n(\theta)$ в ряд Тейлора

$$\Lambda_n(\theta) := \frac{1}{n} \log Ee^{\theta A(n)} = \lambda + n^{2H-1} \frac{\theta^2}{2} + o(\theta^2), \\ \theta \rightarrow 0,$$

где учтено, что $\Lambda_n''(0) = DA(n)/n$. Поэтому $\Lambda_n(\theta) \rightarrow \infty$, $n \rightarrow \infty$, и ТБУ более не применима. Для решения возникшей проблемы используем другую нормировку, полагая

$$\hat{\Lambda}_n(\theta) = \frac{1}{a(n)} \log Ee^{a(n)\theta A(n)/n},$$

где искомая неслучайная последовательность $a(n) \uparrow \infty$. Для выбора этой последовательности используем разложение Тейлора

$$\hat{\Lambda}_n(\theta) = \lambda + \frac{a(n)}{n^2} D\{A(n)\} \frac{\theta^2}{2} + o(\theta^2), \quad \theta \rightarrow 0.$$

Поскольку $DA(n) = n^{2H}$, то можно предположить, что (при малом θ) требование сходимости $\hat{\Lambda}_n(\theta)$ к конечному пределу $\hat{\Lambda}(\theta)$ эквивалентно выбору последовательности $a(n) \sim n^{2(1-H)}$. Действительно, такой выбор приводит к следующему обобщению асимптотики хвоста распределения стационарной нагрузки:

$$\lim_{x \rightarrow \infty} \frac{1}{x^{2-2H}} \log P(W > x) = -\theta^*, \quad (19)$$

где

$$\theta^* = \sup\{\theta > 0 : \hat{\Lambda}(\theta) \leq 0\},$$

а

$$\hat{\Lambda}(\theta) = \lim_{n \rightarrow \infty} \hat{\Lambda}_n(\theta).$$

Важнейший вывод, следующий из этого анализа, состоит в том, что, в отличие от классических систем, асимптотика имеет вид *распределения Вейбулла с умеренно тяжелым хвостом* (см. разд. 2)

$$P(W > x) \asymp e^{-\theta^* x^{2-2H}}, \quad H \in \left(\frac{1}{2}, 1\right). \quad (20)$$

(Здесь \asymp означает логарифмическую эквивалентность.) Этот результат верен и для более общего входного процесса, а в системе с ФБП (18) и скоростью обслуживания $C > \lambda$ параметр θ^* в (19) получен в явном виде:

$$\theta^* = \frac{1}{2\sigma^2} \left(\frac{C-\lambda}{H}\right)^{2H} (1-H)^{-2(1-H)}. \quad (21)$$

Этот результат (как нижняя граница) был найден в [25] с использованием соотношений

$$P(W > x) = P\left(\max_{t \geq 0} (A(t) - Ct) > x\right) \geq \\ \geq \max_{t \geq 0} P(A(t) > x + Ct) = \\ = \max_t P\left(N(0, 1) > (x + t(C - \lambda))t^{-H}\right),$$

где $N(0, 1)$ — стандартная нормальная с.в., и учтен явный вид (18) процесса $A(t)$. Этот результат был обобщен и уточнен (в форме (20)) в работах [26, 27] (см. также [20]).

Важность ФБП связана с тем, что этот процесс возникает, например, как результат суммирования (и надлежащего нормирования) трафиков от растущего числа так называемых ON/OFF-источников, где периоды простоя (OFF) и/или периоды работы (ON) имеют распределения с тяжелым хвостом, конечным средним и бесконечной дисперсией [28–31]. Иными словами, ON/OFF-периоды порождают *альтернирующий процесс восстановления* с бесконечной дисперсией длины (обобщенного) интервала. Такой процесс хорошо описывает огромную вариативность современных трафиков, от кратких электронных сообщений до видеоконференций [1]. Для пояснения этого фундаментального результата рассмотрим трафик, порождаемый m независимыми (идентичными) ON/OFF-источниками, считая, что хвост по крайней

мере одной из ф.р. ON- или OFF-периода имеет вид

$$\bar{F}(x) \sim cx^{-\beta}L(x), \quad 1 < \beta < 2, \quad (22)$$

где константа $c > 0$, а функция L медленно меняется на бесконечности [29, 32]. Условие (22) обеспечивает упомянутые выше моментные свойства. Для источника i определим индикатор $I_i(t) = 1$, если момент t принадлежит ON-периоду ($I_i(t) = 0$, иначе), и пусть $\gamma = P(I_i(t) = 1)$ (стационарная вероятность ON-периода). Рассмотрим процесс, равный суммарному ON-времени от всех источников в интервале $[0, Tt)$ (здесь $T > 0$ — масштабный множитель)

$$W_m(Tt) = \int_0^{Tt} \sum_{i=1}^m I_i(u) du, \quad t \geq 0. \quad (23)$$

В работе [29] доказано, что если $m \rightarrow \infty$, а затем $T \rightarrow \infty$ (порядок взятия пределов важен), то имеет место следующая функциональная центральная предельная теорема

$$\left\{ \frac{W_m(Tt) - Ttm\gamma}{T^H \sqrt{L_1(T)m}}, \quad t \geq 0 \right\} \Rightarrow c_1 B_H, \quad (24)$$

где параметр Херста $H \in (1/2, 1)$ (предельного ФБП B_H), постоянная $c_1 > 0$ и (медленно меняющаяся на бесконечности) функция L_1 полностью определяются через заданные параметры ф.р. ON/OFF-периодов. Таким образом, конечномерные распределения нормированного и центрированного процесса (23) сходятся к соответствующим распределениям ФБП. (Взятие пределов в обратном порядке приводит к процессу Леви.) В работе [30] для данной модели найдены условия согласованного роста параметров m и T , при которых пределы можно переставлять, получая ФБП. Проведенный анализ мотивирует значительное внимание, уделенное выше входному ФБП, а также введение тесно связанного с пределом (24) *фрактального броуновского трафика*, определяемого следующим образом [25, 33]:

$$A(t) = mt + \sigma \sqrt{m} B_H(t), \quad H \in (1/2, 1), \quad t \geq 0,$$

где $m > 0$ — интенсивность трафика, а постоянная $\sigma > 0$ отражает его вариативность.

Входной процесс с долговременной зависимостью возникает также при суммировании ON/OFF-потоков от m независимых источников, где каждый поток имеет интенсивность $r > 0$ в течение ON-периода, имеющего распределение с тяжелым хвостом, а OFF-период, когда $r = 0$, распределен экспоненциально с параметром λ [34–37].

В пределе при $m \rightarrow \infty$, $\lambda m \rightarrow \Lambda$ возникает входной процесс типа $M/G/\infty$, который в каждый момент времени равен текущему числу ON-периодов (*активных сессий*). Этот входной процесс можно представлять как стационарное число β занятых каналов в системе $M/G/\infty$ с интенсивностью пуассоновского потока Λ и *временем обслуживания*, равным длине τ ON-периода. В частности, в [35] показано, что $P(\beta = 0) = \exp\{-\Lambda E\tau\}$ (что соответствует системе $M/G/\infty$ [38, с. 258]) и исследована асимптотика процесса нагрузки в системе, куда поступает такой поток. (Обычно такой входной процесс рассматривается в рамках *жидкостной* модели [35].)

Много полезных асимптотических результатов содержится в диссертациях [39, 40] и недавней книге [41], где большое внимание уделено анализу систем с входным процессом с долговременной зависимостью. Дополнительные детали можно найти в [42].

Замечание 3. Параметр θ^* из (21) возникает также при исследовании асимптотики максимума стационарного виртуального времени ожидания в системе с входным процессом вида (18), см. [43].

5 Долговременная зависимость и регенерация

Теперь рассмотрим процесс обслуживания, который обладает свойством долговременной зависимости, однако имеет положительно возвратный вложенный процесс регенераций. (Последнее означает конечность средней длины цикла регенерации.) Отметим, что собственно *процесс восстановления обладает долговременной зависимостью тогда и только тогда, когда длина интервала восстановления имеет бесконечную дисперсию* [44]. (Этот эффект связан с так называемым парадоксом времени ожидания [6].) Поэтому существование стационарных процессов очереди/нагрузки с долговременной зависимостью, и имеющих одновременно вложенный процесс регенераций с *хорошими моментными свойствами*, на первый взгляд представляется парадоксальным. В частности, такой процесс допускает исследование и оценивание на основе регенеративного метода, что является основной мотивировкой включения данного раздела.

Рассмотрим стационарную систему $GI/G/1$ в предположении, что время обслуживания S обладает следующими моментными свойствами:

$$ES^3 < \infty, \quad ES^4 = \infty. \quad (25)$$

Обозначим длину цикла регенерации через

$$\Delta =_{st} B + I,$$

где B — длина периода занятости, а I — длина интервала простоя системы. В [45] доказано, что при (25) стационарный процесс нагрузки $W = \{W_n\}$ обладает долговременной зависимостью, т. е.

$$\sum_{n=0}^{\infty} \text{cov}(W_0, W_n) = \infty. \quad (26)$$

Известно, что если $ES^3 < \infty$, то также $EB^3 < \infty$, и что $E\tau^3 < \infty$ влечет $EI^3 < \infty$, где τ — интервал входного потока [46]. Предположим, что $E\tau^3 < \infty$, тогда по неравенству Минковского $[EI^3]^{1/3} \leq [EB^3]^{1/3} + [E\tau^3]^{1/3}$, и поэтому $E\Delta^3 < \infty$. Пусть $G(x) = P(\Delta \leq x)$, и рассмотрим незавершенную в момент t длину цикла регенерации $\beta(t)$. Пусть с.в. τ нерешетчатая, тогда длина цикла Δ тоже нерешетчатая и существует (слабый) предел $\beta(t) \Rightarrow \beta$, причем [5] (см. также (2))

$$P(\beta \leq x) = \frac{1}{E\Delta} \int_0^x (1 - G(y)) dy.$$

Поэтому стационарное незавершенное время восстановления β имеет конечный второй момент

$$\begin{aligned} E\beta^2 &= \frac{1}{E\Delta} \int_0^{\infty} x^2 (1 - G(x)) dx = \\ &= \frac{1}{3E\Delta} \int_0^{\infty} x^3 G(dx) = \frac{E\Delta^3}{3E\Delta} < \infty \end{aligned}$$

и, в частности,

$$P(\beta > x) = o(x^{-2}) \quad \text{при } x \rightarrow \infty.$$

Таким образом, моментные свойства Δ и β допускают статистическое исследование процесса нагрузки с долговременной зависимостью регенеративным методом [47], (см. также [48, 49]. Подчеркнем, что у самого процесса восстановления с долговременной зависимостью $E\beta = \infty$, что исключает применение регенеративного имитационного моделирования. (Конечно, в рассматриваемой системе $E\beta < \infty$.) Дополнительные детали проведенного выше анализа можно найти в [48, 50].

Отметим близкий результат, касающийся асимптотики хвоста ф.р. периода занятости B стационарной системы $M/G/1$, где $\rho := \lambda ES < 1$ и ф.р. F времени обслуживания S имеет регулярное изменение на бесконечности, т. е. для любого $t > 0$ и некоторого $\alpha > 0$ верно $\bar{F}(tx) \sim \bar{F}(x)t^{-\alpha}$, $x \rightarrow \infty$. Тогда [51]

$$\bar{G}(x) := P(B > x) \sim \frac{1}{1-\rho} \bar{F}((1-\rho)x), \quad x \rightarrow \infty. \quad (27)$$

Например, пусть S имеет распределение Парето с показателем $\alpha \in (3, 4)$. Поскольку хвост времени простоя I легче, чем \bar{G} , то он не играет роли в асимптотике ([3], Предложение 2.7), и соотношение (27) дает

$$\begin{aligned} P(\Delta > x) \sim \bar{G}(x) &\sim (1-\rho)^{-(\alpha+1)} x^{-\alpha} = \\ &= o(x^{-3}), \quad x \rightarrow \infty. \end{aligned}$$

Обсудим другие недавние результаты, касающиеся долговременной зависимости регенеративных процессов обслуживания. Рассмотрим стационарный процесс $N(t)$, равный числу восстановлений в интервале $[0, t)$ с интервалами τ_n , и пусть $R_n = \tau_1 + \dots + \tau_n$, $R := \{R_n, n \geq 1\}$. Согласно [52], процессы N и R обладают долговременной зависимостью, если соответственно

$$\limsup_{t \rightarrow \infty} \frac{DN(t)}{t} = \infty, \quad \limsup_{n \rightarrow \infty} \frac{DR_n}{n} = \infty. \quad (28)$$

Достаточным условием для второго соотношения в (28) является (см. (26))

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n \text{cov}(\tau_0, \tau_i) = \infty.$$

Также в [52] дается следующее определение параметра Херста H для данных процессов:

$$\begin{aligned} H_N &:= \inf \left\{ h : \limsup_{t \rightarrow \infty} \frac{DN(t)}{t^{2h}} < \infty \right\}, \\ H_R &:= \inf \left\{ h : \limsup_{n \rightarrow \infty} \frac{DR_n}{n^{2h}} < \infty \right\}. \end{aligned} \quad (29)$$

В [52] показано, в частности, что в стационарной системе $GI/G/1$ при входном процессе восстановления R с условиями $E\tau^2 < \infty$, $ES^2 < \infty$ выходной процесс \tilde{R} (последовательность интервалов между уходами) также не обладает долговременной зависимостью. Кроме того, процесс \tilde{R} обладает долговременной зависимостью тогда и только тогда, когда $H_R \in (1/2, 1)$ в (29). Более того, если в стационарной системе $GI/M/1$

$$P(\tau > x) = x^{-c} L(x), \quad c \in (1, 2), \quad x > 0, \quad (30)$$

где функция L медленно меняется на бесконечности, то выходной процесс $\tilde{N}(t)$ (число уходов в $[0, t)$) обладает долговременной зависимостью в смысле (29). Этот же результат верен для процесса $\tilde{N}(t)$ в системе $M/G/1$, если хвост $P(S > x)$ удовлетворяет условию (30) (вместо хвоста $P(\tau > x)$).

Рассмотрим с.в. $T > 0$ с *максимальным конечным моментом* $\gamma := \sup(k > 0 : ET^k < \infty) \in (1, 2)$. Тогда процесс восстановления

$$N(t) = \max(n : T_1 + \dots + T_n \leq t)$$

(где $T_k \stackrel{st}{=} T$) имеет параметр Херста $H_N = (3 - \gamma)/2 \in (1/2, 1)$ [44].

В работе [53] с использованием техники случайных блужданий и теории восстановления (по аналогии с основополагающей работой [45]) получено следующее обобщение последнего результата. Если в системе $GI/G/1$ время обслуживания имеет максимальный конечный момент $\gamma \in (3, 4)$, а его ф.р. регулярно изменяется на бесконечности, то стационарный процесс нагрузки $\{W_n\}$ обладает долговременной зависимостью и имеет параметр Херста

$$H_W := \inf \left\{ h : \limsup_{n \rightarrow \infty} \frac{D \left(\sum_{i=1}^n W_i \right)}{n^{2h}} < \infty \right\} = \frac{5 - \gamma}{2} \in \left(\frac{1}{2}, 1 \right). \quad (31)$$

В случае же стационарной системы $M/G/1$, если (авто)корреляции процесса очереди ν_n (в моменты ухода) удовлетворяют условию $\text{cor}(\nu_0, \nu_n) \sim cn^{-\alpha}$, $\alpha \in (0, 1)$, то этот процесс с долговременной зависимостью также имеет параметр Херста, удовлетворяющий (31) [53].

Приведенные результаты демонстрируют присутствие долговременной зависимости у процессов обслуживания, у которых одновременно длина цикла регенерации имеет конечные моменты выше 3-го, и (как отмечалось) открывают возможность статистического исследования таких систем на основе регенеративного метода. Подробное обсуждение данной проблемы предполагается провести в отдельной работе.

6 Заключение

В работе приведен обзор асимптотик вероятностей больших отклонений процессов нагрузки/очереди в стационарной системе обслуживания. Рассмотрены (1) системы, в которых время обслуживания имеет тяжелый хвост; (2) регенеративные системы с входным процессом Леви и временем обслуживания, имеющим легкий хвост; и (3) системы с входным потоком с долговременной зависимостью. В качестве иллюстрации применения ТБУ подробно рассмотрена система $M/M/1$.

Кроме того, рассмотрены процессы с долговременной зависимостью, допускающие применение регенеративного имитационного моделирования. За исключением п. 3.3, рассматриваются лишь одноканальные системы обслуживания. Важным аспектом работы является указание возможных областей применения методов оценивания параметров асимптотик, учитывая известную сложность такого оценивания на основе ТБУ. Например, в результате проведенного анализа естественной областью применения регенеративного метода оказываются системы вида 2, а также регенеративные процессы с долговременной зависимостью, рассмотренные в разд. 5. Представляется перспективным оценивание параметров асимптотик в таких системах с помощью недавно предложенного варианта метода расщепления [2]. В остальных случаях в основном параметры асимптотик выражаются через заданные параметры, и такие модели могут быть использованы для тестирования методов оценивания. С другой стороны, эти методы могут быть использованы для проверки качества асимптотик.

Обзор не мог вместить многих важных моделей. Например, за его пределами остались асимптотические результаты для систем с разделением процессора [40], для жидкостных моделей [34, 35], для стохастических сетей [14, 15, 54]. В частности, работа [34] содержит великолепный обзор жидкостных моделей с входным процессом с долговременной зависимостью, а в книге [54] даны асимптотические результаты, касающиеся больших отклонений некоторых сетевых процессов с легкими хвостами (в том числе в системах поллинга). Все эти темы требуют отдельного обсуждения.

Литература

1. Willinger W., Taqqu M., Sherman R., Wilson D. Self-similarity through high-variability: Statistical analysis of Ethernet LAN traffic at the source level // IEEE/ACM Transactions on Networking. 1997. Vol. 5. No. 1. P. 71–86.
2. Бородин А. В., Морозов Е. В. Ускоренное регенеративное моделирование вероятности перегрузки односторонней очереди // ОПИПМ, 2007. Т. 14. Вып. 3. С. 385–397.
3. Sigman K. Appendix: A primer on heavy-tailed distributions // Queueing Systems, 1999. Vol. 33. P. 261–275.
4. Embrechts P., Veraverbeke N. Estimates for the probability of ruin with special emphasis on the possibility of large claims // Insurance: Mathematics and Economics, 1982. Vol. 1. P. 55–72.
5. Asmussen S. Applied probability and queues. — NY: Springer, 2003. 2nd ed.
6. Феллер В. Введение в теорию вероятностей и ее приложения. Т. 2. — М.: Мир, 1984.

7. *Asmussen S., Kluppelberg C.* Stationary $M/G/1$ excursions in the presence of heavy tails // *J. Appl. Probab.*, 1997. Vol. 34. P. 208–212.
8. *Greiner M., Jobmann M., Kluppelberg C.* Telecommunication traffic, queueing models and subexponential distributions // *Queueing Systems*, 1999. Vol. 33. P. 125–152.
9. *Asmussen S., Kluppelberg C., Sigman K.* Sampling at subexponential times, with queueing applications // *Stochastic Process. Appl.*, 1999. Vol. 79. P. 265–286.
10. *Asmussen S.* Subexponential asymptotics for stochastic processes: Extremal behavior, stationary distributions and first passage probabilities // *Ann. Appl. Probab.*, 1998. Vol. 8. P. 354–374.
11. *Asmussen S., Møller J. R.* Tail asymptotics for $M/G/1$ type queueing processes with subexponential increments // *Queueing Systems*, 1999. Vol. 33. P. 153–176.
12. *Foss S., Korshunov D.* Sampling at random time with a heavy-tailed distribution // *Markov Processes Relat. Fields*, 2000. Vol. 6. P. 543–568.
13. *Foss S., Korshunov D.* Heavy tails in multiserver queue // *Queueing Systems*, 2006. Vol. 52. P. 31–48.
14. *Huang T., Sigman K.* Steady-state asymptotics for tandem, split-match and other feedforward queues with heavy tailed service // *Queueing Systems*, 1999. Vol. 33. P. 233–259.
15. *Baccelli F., Foss S.* Moments and tails in monotone-separable stochastic networks // *Ann. Appl. Probab.*, 2004. Vol. 14. P. 612–650.
16. *Shang W., Liu L., Li Q.-L.* Tail asymptotics for the queue length in an $M/G/1$ retrial queue // *Queueing Systems*, 2006. Vol. 52. P. 193–198.
17. *Glynn P. W., Whitt W.* Logarithmic asymptotics for steady-state tail probabilities in a single-server queue // *Adv. Appl. Probab.*, 1994. P. 131–156.
18. *Abate J., Choudhury G. L., Whitt W.* Exponential approximations for tail probabilities in queues. I: Waiting times // *Operations Research*, 1995. Vol. 43. P. 885–901.
19. *Ganesh A., O'Connell N., Wischik D.* Big queues. — Berlin: Springer-Verlag, 2004.
20. *Kelly F. P.* Notes on effective bandwidths / In: *Stochastic networks, theory and applications* // F. Kelly, S. Zachary, I. Ziedins, eds. — Oxford: Clarendon Press, 1996. P. 141–168.
21. *Chang C.-S.* Performance guarantees in communications networks. — London: Springer-Verlag, 2000.
22. *Vorobieva I., Morozov E., Pagano M., Prociassi G.* A new regenerative estimator for effective bandwidth prediction // *AMICT'2007 Proceedings*. — Petrozavodsk: Petrozavodsk University Press, 2008. Vol. 9. P. 175–187.
23. *Sadowsky J.* Large deviations theory and efficient simulation of excessive backlogs in a $GI/GI/m$ queue // *IEEE Transactions on Automatic Control*, 1991. Vol. 36. No. 12. P. 1383–1394.
24. *Гнеденко Б. В., Коваленко И. Н.* Введение в теорию массового обслуживания. — М.: Наука, 1987.
25. *Norros I.* A storage model with self-similar input // *Queueing Systems*, 1994. Vol. 16. P. 387–396.
26. *Duffield N. G., O'Connell N.* Large deviations and overflow probability for the general single-server queue with applications // *Math. Proc. Cambridge Philos. Soc.*, 1995. Vol. 118. No. 2. P. 363–374.
27. *Duffy K., Lewis T., Sullivan W. G.* Logarithmic asymptotics for the supremum of a stochastic process // *Annals of Appl. Probab.*, 2003. Vol. 13. No. 2. P. 430–445.
28. *Leland W., Taqqu M., Willinger W., Wilson D.* On the self-similar nature of ethernet traffic // *IEEE/ACM Transactions on Networking*, 1994. Vol. 2. No. 1. P. 1–15.
29. *Taqqu M. S., Willinger W., Sherman R.* Proof of a fundamental result in self-similar traffic modeling // *Computer Communication Review*, 1997. Vol. 27. P. 5–23.
30. *Kaj I.* Convergence of scaled renewal processes to fractional Brownian motion. Sci. report. — Dept. of Mathematics, Uppsala University, 1999. No. 11. P. 1–28.
31. *Samorodnitsky G.* Long range dependence, heavy tails and rare events. — Aarhus: Aarhus University, MaPhySto, 2002. P. 1–84.
32. *Resnick S.* Heavy tail modeling and teletraffic data // *The Annals of Statistics*, 1997. Vol. 25. No. 5. P. 1805–1869.
33. *Norros I.* Studies on a model for connectionless traffic, based on fractional brownian motion // *Conference on Applied Probability in Engineering, Computer and Communications Sciences*. — Paris: INRIA/ ORSA/ TIMS/ SMAI, 1993. P. 1–13.
34. *Boxma O. J., Dumas V.* Fluid queues with long-tailed activity periods // *Computer Communications*, 1998. Vol. 21. P. 1509–1529.
35. *Jelenković P. R., Lazar A. A.* Subexponential asymptotics of a Markov-modulated random walk with queueing applications // *J. Appl. Probab.*, 1998. Vol. 35. P. 338–339.
36. *Resnick S., Samorodnitsky G.* Activity periods of an infinite server queue and performance of certain heavy tailed fluid queues // *Queueing Systems*, 1999. Vol. 33. P. 43–71.
37. *Dumas V., Simonian A.* Asymptotic bounds for the fluid queue fed by sub-exponential on/off sources // *Adv. Appl. Probab.*, 2000. Vol. 32. P. 244–255.
38. *Боровков А. А.* Вероятностные процессы в теории массового обслуживания. — М.: Наука, 1972.
39. *Deng Q.* Queues with regular variation. Ph.D. Thesis. — Eindhoven: Eindhoven University of Technology, 2001.
40. *Zwart A.* Queueing systems with heavy tails. Ph.D. Thesis. — Eindhoven: Eindhoven University of Technology, 2001.
41. *Mandjes M.* Large deviations of Gaussian queues. — Chichester: Wiley, 2007.
42. *Morozov E. V.* Self-similarity and long-range dependence in network traffic modeling // *FDPW'99 "Developments in Distributed Systems and data Communications" Proceedings*, 1999. Vol. 2. P. 32–40.
43. *Zeevi A., Glynn P.* On the maximum workload of a queue fed by fractional Brownian motion // *Annals of Appl. Probab.*, 2000. Vol. 10. No. 4. P. 1084–1099.
44. *Daley D. J.* The Hurst index of long-range dependent renewal process // *Annals of Probability*, 1999. Vol. 27. P. 2035–2041.

45. *Daley D. J.* The serial correlation coefficients of waiting times in a stationary single server queue // *Austr. Math. Society*, 1968. Vol. 8. P. 683–699.
46. *Wolff R. W.* Stochastic modeling and the theory of queues. — Prentice-Hall, 1989.
47. *Glynn P., Iglehart D.* Conditions for the applicability of the regenerative method // *Management Sci.*, 1993. Vol. 39. P. 1108–1111.
48. *Morozov E. V.* Weak regeneration in modeling of queuing processes // *Queueing Systems*, 2000. Vol. 46. P. 293–313.
49. *Морозов Е. В., Белый А. В., Боденов Д. В.* Расширенная регенерация: применения к анализу сетевого графика // *ОПиПМ*, 2007. Т. 14. Вып. 6. С. 1022–1042.
50. *Morozov E.* Communications systems: Rare event simulation and effective bandwidths. — Pamplona: Universidad Publica de Navarra, 2004.
51. *De Meyer A., Teugels J. L. T.* On the asymptotic behaviour of the distribution of the busy period and service time in M/G/1 // *J. Appl. Probab.*, 1980. Vol. 17. P. 802–813.
52. *Daley D. J., Vesilo R.* Long range dependence of point processes, with queuing examples // *Stochastic Processes Appl.*, 1997. Vol. 70. P. 265–282.
53. *Carpio K. J. E.* Long-range dependence of stationary processes in single-server queues // *Queueing Systems*, 2007. Vol. 55. P. 123–130.
54. *Боровков А. А.* Эргодичность и устойчивость случайных процессов. — М.: Эдиториал УРСС, 1999.

ВЕРОЯТНОСТНАЯ МОДЕЛЬ ОБСЛУЖИВАНИЯ ТРАФИКА В СИСТЕМЕ СЕТЕЦЕНТРИЧЕСКОГО ТИПА

В. Ю. Бородакий¹

Аннотация: В терминах теории телетрафика построена модель сетецентрической системы, в которой из центров обработки данных по запросам абонентов передается информация в виде блоков данных различной длины. Передача блоков описывается в терминах «эластичного» трафика, их длина распределена по экспоненциальному закону, а обслуживание осуществляется по дисциплине разделения процессора. Предполагается, что поток запросов абонентов является пуассоновским, а блок данных, независимо от его длины, характеризуется минимальным требованием к ширине полосы пропускания. Проведен анализ модели отдельного звена сети сетецентрической системы (СС), получено аналитическое выражение для вычисления вероятности блокировки запроса из-за отсутствия достаточной для передачи блока данных ширины полосы пропускания.

Ключевые слова: сетецентрическая система; эластичный трафик; вероятность блокировки; звено сети

1 Введение

В настоящее время активно развиваются распределенные автоматизированные системы управления специального назначения, построенные на базе сетецентрической концепции [1]. От выбора структуры подобных систем и поддержки заданных показателей качества существенно зависит эффективность решения возложенных на них задач с точки зрения соответствия заданным требованиям по режиму работы и производительности. Известны системы [2], где невыполнение этих требований не только снижает производительность, но и приводит к деградации системы вплоть до ее уничтожения. В состав СС входят: абонентские узлы, обменивающиеся управляющими сигналами с объектами управления и обладающие потребностью информационного обмена с управляющими компонентами системы; центры обработки данных (ЦОД), принимающие информацию от абонентских узлов и посылающие ответное воздействие; телекоммуникационное оборудование, обеспечивающее передачу информации между абонентскими узлами и ЦОД.

Проектирование СС, как правило, осуществляется в автономном (off-line) режиме. В этом случае исходными данными для проектирования являются: расположение и конфигурация абонентских узлов, структура и топология сети передачи данных, пропускные способности звеньев сети, объемы информационных потребностей абонентских узлов. Одним из важнейших показателей качества

функционирования СС является величина обслуженной нагрузки (объем переданной информации), поскольку в условиях ограниченной пропускной способности звеньев сети часть запросов с некоторой вероятностью может получить отказ в обслуживании из-за отсутствия достаточных ресурсов на маршрутах между абонентскими узлами и ЦОД.

Для дальнейшего изложения понадобятся следующие обозначения:

\mathcal{V}_s — множество абонентских узлов (s -абоненты);

\mathcal{V}_T — множество ЦОД (t -центры);

\mathcal{N} — множество предоставляемых ЦОД типов данных ($\mathcal{N} := \{1, \dots, N\}$);

θ_n — длина блока данных n -типа (n -блока);

b — требование к минимальному значению ширины полосы пропускания для передачи блока данных любого типа;

$\lambda_n(s, t)$ — интенсивность потока запросов s -абонента в t -центр на передачу n -блоков;

$a_n(s, t) := \lambda_n(s, t)\theta_n$ — нагрузка, создаваемая n -блоками между s -абонентом и t -центром (предложенная нагрузка);

$\mathcal{L}(s, t)$ — множество маршрутов между s -абонентом и t -центром;

$l_i(s, t)$ — i -маршрут между s -абонентом и t -центром;

$\mathcal{E}_l(s, t)$ — множество звеньев $l_i(s, t)$ -маршрута.

¹Национальный исследовательский ядерный университет «МИФИ», vladbor@inbox.ru

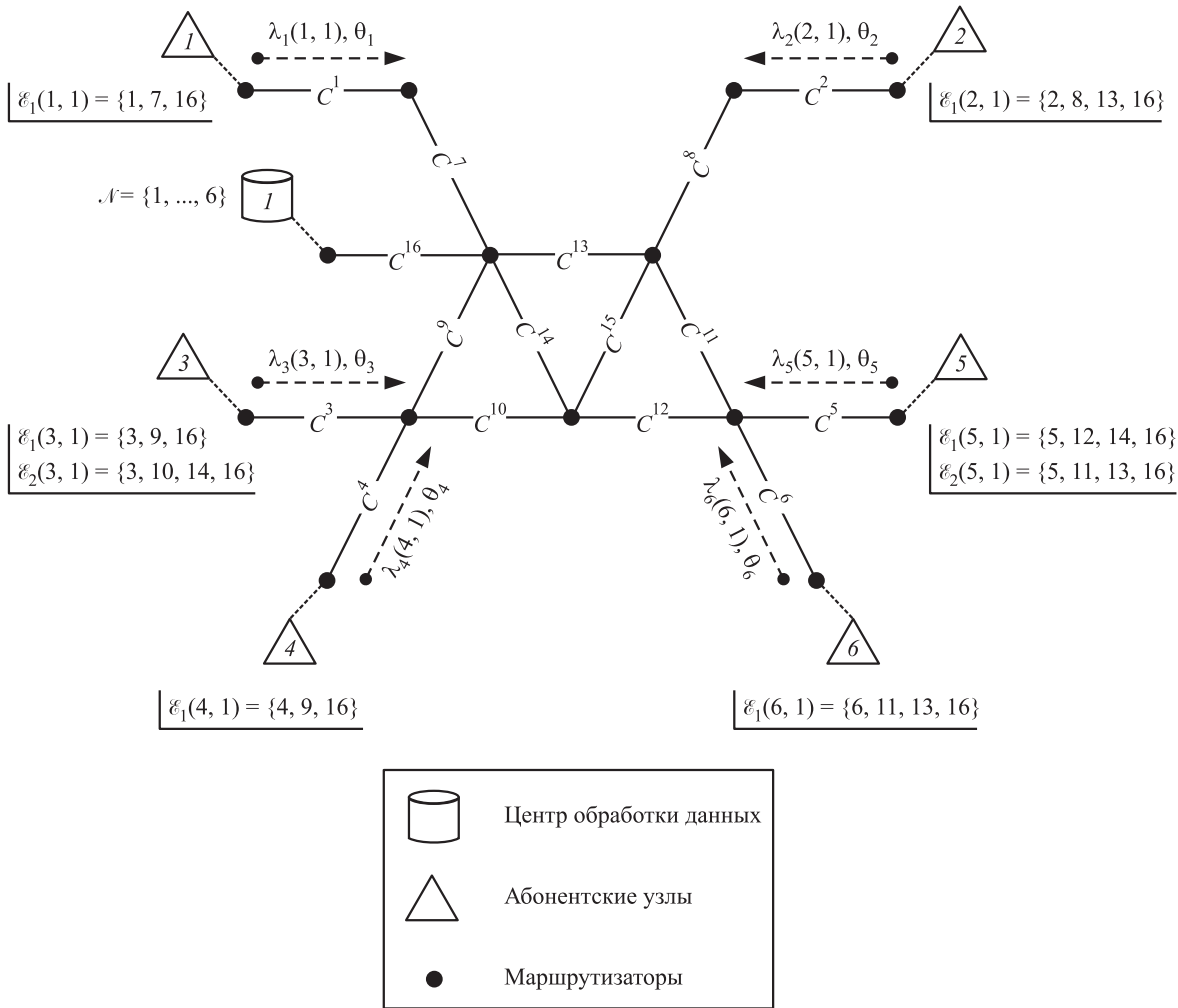


Рис. 1 Схема модели сетецентрической системы

Пример схемы модели СС показан на рис. 1, где проиллюстрированы введенные обозначения.

Величина обслуженной нагрузки в СС определяется формулой

$$\tilde{a} := \sum_{s \in \mathcal{V}_s} \sum_{t \in \mathcal{V}_T} \sum_{i: l_i(s,t) \in \mathcal{L}(s,t)} \sum_{n \in \mathcal{N}} a_{in}(s,t) (1 - B_{in}(s,t)),$$

где $a_{in}(s,t)$ — предложенная n -блоками нагрузка и $B_{in}(s,t)$ — вероятность блокировки n -блока на $l_i(s,t)$ -маршруте. Если считать известными величины предложенных нагрузок и множества маршрутов $\mathcal{L}(s,t)$, то задача оценки величины обслуженной нагрузки \tilde{a} сводится к вычислению вероятностей $B_{in}(s,t)$. Эти вероятности в предположении, что блокировки на всех звеньях $l_i(s,t)$ -маршрута происходят независимо в совокупности, можно представить в виде

$$B_{in}(s,t) \approx 1 - \prod_{e \in \mathcal{E}_i(s,t)} (1 - B_{in}^e(s,t)),$$

$$n \in \mathcal{N}, i: l_i(s,t) \in \mathcal{L}(s,t), t \in \mathcal{V}_T, s \in \mathcal{V}_s,$$

где $B_{in}^e(s,t)$ — вероятность блокировки запроса на передачу n -блока на e -звене $l_i(s,t)$ -маршрута.

Предположение о независимости блокировок лежит в основе известного метода приближенного расчета вероятностей блокировок — метода просеянной нагрузки (reduced load approximation), изначально разработанного для сетей с коммутацией каналов [3] и развитого в [4] для сетей с многоадресными соединениями. Данный подход предполагает наличие точного метода для расчета вероятностей блокировок $B_{in}^e(s,t)$ на отдельном звене сети. Применение метода просеянной нагрузки к анализу вероятностных характеристик СС исследовано в [5], а в данной статье построена модель отдельного звена сети и проведен анализ ее вероятностных характеристик.

2 Модель звена сети сетцентрической системы

Предположим, что в модели сети СС (например, показанной на рис. 1), все звенья, кроме некоторого звена e^* , имеют неограниченные ресурсы для обслуживания запросов абонентов, т. е. $C^e = \infty$ для $e \neq e^*$. Задача анализа блокировок в такой системе сводится к анализу сети, состоящей из одного звена e^* , с одним ЦОД t^* , который передает данные из множества $\mathcal{N} = \{1, \dots, N\}$. Для удобства записи далее индексы e^* и t^* опускаются.

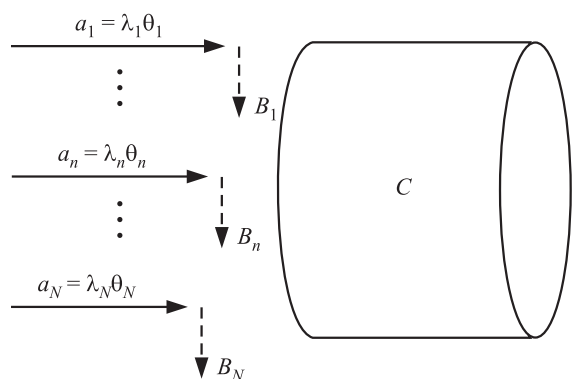


Рис. 2 Модель звена сети СС

Итак, рассмотрим модель звена сети емкостью C условных единиц, схематично изображенную на рис. 2. По звену передаются блоки данных N типов, запросы на передачу n -блока образуют

пуассоновский поток интенсивности λ_n , размер блока является случайной величиной, распределенной экспоненциально со средним θ_n , а величина предложенной n -блоками нагрузки вычисляется по формуле $a_n := \lambda_n \theta_n$. Блоки любого типа характеризуются минимальным требованием b к емкости звена, а их обслуживание производится по дисциплине разделения процессора (PS — Processor Sharing). С вероятностью B_n n -блок может получить отказ в обслуживании, если в момент его поступления свободны менее b из C единиц емкости звена.

Пусть $X_n(t)$ — число передаваемых в момент t n -блоков. Тогда составной случайный процесс (СП) $\{\mathbf{X}(t) = (X_1(t), \dots, X_N(t)), t \geq 0\}$ описывает функционирование модели над пространством состояний

$$\mathcal{X} := \left\{ \mathbf{x} := (x_1, \dots, x_N) : b \sum_{n \in \mathcal{N}} x_n \leq C \right\}.$$

Схема обслуживания в рассматриваемой модели показана на рис. 3, где использовано понятие «эластичный трафик» (elastic traffic) [7, 6]. Такой режим известен также под названием передачи по принципу “best effort”, суть которого заключается в том, что в процессе одновременной передачи по звену сети с ограниченной пропускной способностью нескольких потоков (блоков данных) все потоки получают одинаковую долю пропускной способности вне зависимости от своего объема. В сделанных выше предположениях СП $\mathbf{X}(t)$ является марковским и, как нетрудно убедиться, равновесное распреде-

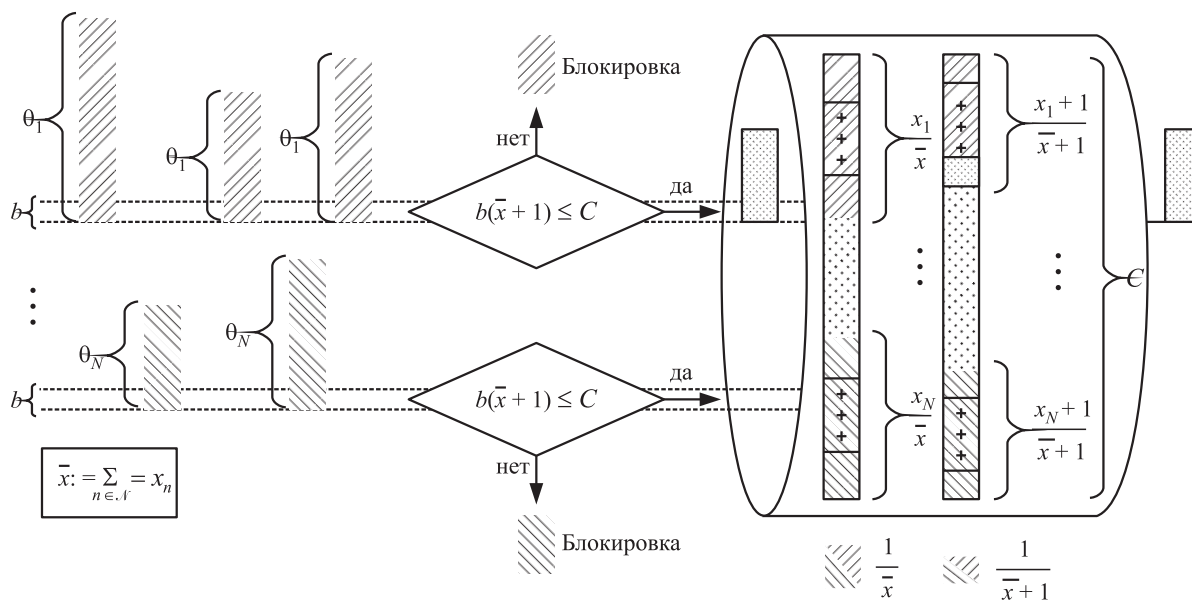


Рис. 3 Схема обслуживания эластичного трафика

ление вероятностей его состояний $\pi(\mathbf{x})$ определяется из системы уравнений локального баланса

$$\pi(\mathbf{x}) \frac{C}{\theta_n} x_n \left(\sum_{j \in \mathcal{N}} x_j \right)^{-1} = \pi(\mathbf{x} - \mathbf{e}_n) \lambda_n, \quad \mathbf{x} \in \mathcal{X}, x_n > 0, n \in \mathcal{N}, \quad (1)$$

где \mathbf{e}_n — вектор, n -я компонента которого равна 1, а остальные — 0. Заметим, что в системе уравнений (1) величина $\left(\sum_{j \in \mathcal{N}} x_j \right)^{-1} C x_n / \theta_n$ соответствует интенсивности обслуживания n -блока в состоянии $\mathbf{x} \in \mathcal{X}$, поскольку обслуживание производится по дисциплине PS.

Используя известное для систем массового обслуживания с дисциплиной PS решение [8], получаем равновесное распределение состояний СП $\mathbf{X}(t)$ в мультипликативном виде:

$$\pi(\mathbf{x}) = G^{-1}(\mathcal{X}) \left(\sum_{n \in \mathcal{N}} x_n \right)! \prod_{n \in \mathcal{N}} \frac{1}{x_n!} \left(\frac{a_n}{C} \right)^{x_n} \quad \mathbf{x} \in \mathcal{X},$$

где $G(\mathcal{X}) = \pi^{-1}(0)$ — нормировочная константа.

Искомой характеристикой является вероятность $B_n := \mathcal{P}(\mathbf{x} \in \mathcal{B}_n)$ блокировки запроса на передачу n -блока, где

$$\mathcal{B}_n := \left\{ \mathbf{x} \in \mathcal{X} : b \left(\sum_{j \in \mathcal{N}} x_j + 1 \right) > C \right\}.$$

В следующем разделе будет получено аналитическое выражения для вычисления вероятностей B_n , $n \in \mathcal{N}$.

3 Анализ вероятностных характеристик модели

Представим пространство состояний \mathcal{X} в виде

$$\mathcal{X} := \bigcup_{c=0}^{C_b} \mathcal{X}(c),$$

где

$$C_b := \left\lfloor \frac{C}{b} \right\rfloor, \quad \mathcal{X}(c) := \left\{ \mathbf{x} \in \mathcal{X} : \sum_{n \in \mathcal{N}} x_n = c \right\},$$

и обозначим

$$q(c) := \mathcal{P}(\mathbf{x} \in \mathcal{X}(c)), \quad c = 0, \dots, C_b.$$

Используя уравнения локального баланса (1) и приняв $x_j > 0$ для некоторого $j \in \mathcal{N}$, получаем

$$\begin{aligned} q(c) &= \sum_{\mathbf{x} \in \mathcal{X}(c)} \pi(\mathbf{x}) = \\ &= \sum_{\mathbf{x} \in \mathcal{X}(c)} \frac{\lambda_j \theta_j}{C x_j} \left(\sum_{n \in \mathcal{N}} x_n \right) \pi(\mathbf{x} - \mathbf{e}_j), \\ & \quad c = 1, \dots, C_b. \end{aligned}$$

Пусть также найдется $n \in \mathcal{N}$ такое, что $x_n > 0$, и положим $j = n$. Тогда из (1) следует, что

$$\begin{aligned} q(c) &= \sum_{\mathbf{x} \in \mathcal{X}(c)} \sum_{n \in \mathcal{N}} \frac{a_n}{C} \pi(\mathbf{x} - \mathbf{e}_n) = \\ &= \sum_{n \in \mathcal{N}} \frac{a_n}{C} \sum_{\mathbf{x} \in \mathcal{X}(c)} \pi(\mathbf{x} - \mathbf{e}) = \\ &= q(c-1) \sum_{n \in \mathcal{N}} \frac{a_n}{C}, \quad c = 1, \dots, C_b. \end{aligned}$$

Обозначив $a := \sum_{n \in \mathcal{N}} a_n$, окончательно получаем

$$q(c) = q(0) \left(\frac{a}{C} \right)^c, \quad c = 1, \dots, C_b, \quad (2)$$

где

$$q(0) = \left(\sum_{c=0}^{C_b} \left(\frac{a}{C} \right)^c \right)^{-1}. \quad (3)$$

Таким образом, если $a \neq C$, тогда

$$q(c) = \frac{a^c C^{C_b-c} (C-a)}{C^{C_b+1} - a^{C_b+1}}, \quad c = 0, \dots, C_b,$$

и вероятность блокировки запроса на передачу n -блока данных имеет вид

$$B_n = q(C_b) = \frac{a^{C_b} (C-a)}{C^{C_b+1} - a^{C_b+1}} =: B, \quad (4)$$

а обслуженная нагрузка, создаваемая n -блоками, вычисляется по формуле

$$\tilde{a}_n = a_n \left(1 - \frac{a^{C_b} (C-a)}{C^{C_b+1} - a^{C_b+1}} \right), \quad n \in \mathcal{N}.$$

Заметим, что вероятность блокировки блока данных не зависит от его длины, что объясняется одинаковым для блоков любого типа требованием b к емкости звена. В этих условиях формулу (4) можно получить, проведя анализ системы массового обслуживания типа $M|G|1|PS$ с нагрузочным параметром a , емкостью прибора C и максимальным числом находящихся на обслуживании заявок C_b . Распределение вероятностей состояний такой системы имеет вид (2), (3).

4 Заключение

В статье проведен анализ модели отдельного звена сети СС в терминах систем, обслуживающих потоки «эластичного трафика» с заданным требованием к минимальному значению выделяемой при его передаче ширины полосы пропускания. В отличие от известных работ в этой области [7, 6], формулы для вычисления вероятностей блокировок получены в аналитическом виде, что дает преимущество при вычислениях вероятностей блокировок для сети в целом при использовании метода просеянной нагрузки или решения задач оптимизации параметров СС при заданных ограничениях на показатели качества их функционирования. Эти вопросы, также как и численный анализ, не являются предметом изложения в данной статье, но частично исследованы и решены в [5].

Литература

1. *Alberts D. S., Garstka J. J., Stein F. P.* Network centric warfare: Developing and leveraging information superiority / DoD C4ISR Cooperative Research Program publication series. 2nd ed. (revised), 2000. 284 p.
2. *Котенко И. В., Боговик А. В., Ковалев И. С., Загоруйко С. С., Масановец В. В.* Теория управления в системах военного назначения / Под ред. И. В. Котенко. — М.: Изд-во МО, 2001. 320 с.
3. *Ross K. W.* Multiservice loss models for broadband telecommunication networks. — Springer, 1995. 343 p.
4. *Наумов В. А., Самуйлов К. Е., Яркина Н. В.* Теория телетрафика мультисервисных сетей. — М.: Изд-во РУДН, 2007. 191 с.
5. *Бородакий В. Ю.* К решению задачи размещения центров обработки данных в сетевидной системе // Вестник РУДН, сер. «Математика, информатика, физика», 2009. № 3.
6. *Iversen V. B.* Teletraffic engineering: Handbook. — ITU-D, June 2006. 354 p. Адрес в Интернете: <http://www.com.dtu.dk/teletraffic/handbook/telenook.pdf>.
7. *Меликов А. З., Пономаренко Л. А., Паладюк В. В.* Телетрафик: модели, методы, оптимизация. — Киев: ИПК «Политехника», 2007. 285 с.
8. *Бочаров П. П., Печинкин А. В.* Теория массового обслуживания. — М.: Изд-во РУДН, 1995. 529 с.

AN APPROACH TO ACTUARIAL MODELING WITH QUASI-MONTE CARLO: SIMULATION OF RANDOM SUMS DEPENDING ON STOCHASTIC FACTORS*

G. Temnov¹ and S. Kucherenko²

Abstract: The problem of estimating the characteristics of a random sum, when the number of summands is also random, is addressed. The considered case includes an additional stochastic factor: although the summed random variables come from a distribution of a known form, the parameters of this distribution are stochastic and can themselves be viewed as random variables (with known distributions). The Quasi-Monte-Carlo (QMC) techniques are used to handle this problem and to analyze its efficiency relative to the regular Monte-Carlo (MC) simulation methods. The typical area of the application of the investigations is actuarial practice which often deals with random sums of financial losses. Besides actuarial applications, the proposed method may be useful in application to certain problems in informatics, related to the aggregation of heavy-tailed data.

Keywords: actuarial modeling; quasi-Monte-Carlo simulation; random sums

1 Introduction

1.1 Setting up the problem

Summation of random number of random variables is a well known problem that has many applications. One of them is the so-called *loss aggregation* problem in insurance. Usually, one needs to compute with sufficient precision the cumulative distribution of the random variables (r.v.'s), having a sense of, for example, financial losses aggregated for some fixed period (normally, one year).

Specifically, we are interested in the distribution of the r.v.

$$S_N = \sum_{k=1}^N X_k \quad (1)$$

where N is the number of events within a selected period, generated by a process $N(t)$ of occurrences, usually called *counting process*. The crucial point is that the summed r.v.'s X_k are assumed to be mutually independent and also independent of the counting process $N(t)$.

Another important point is that usually in practical applications r.v.'s X_k having in insurance the sense of single losses can be viewed as **identically distributed** r.v.'s having a distribution $\mathbf{P}(X_k < x) =: F_X(x)$. This assumption allows to apply some particular *deterministic techniques*, making the loss aggregation a relatively simple computational task.

In the current work, the problem of loss aggregation supposing that the assumption of identical distribution

of single losses X_k may be violated is addressed. As will be remarked below, deviations from the assumption of identical distribution of losses is quite a natural situation in actuarial modeling, making the application of analytical techniques impossible.

1.2 Aggregation with stochastic parameters

In practical applications, compound distributions should often be modeled with respect to the uncertainty of the parameters of initial distribution (F_X in our terms). This problem can be handled with the help of Bayesian inference. The basic idea of Bayesian modeling for taking into account parameters' uncertainty (see, e.g., [1]) is to consider the vector of the parameters (of initial distribution) as a random vector. Using some *a priori* knowledge about the distribution of this random vector, one can form its *prior distribution* probability density function (pdf) $\pi(\theta)$. If an additional information comes into play in the form of observations \mathbf{X} , the *posterior distribution* pdf with respect to this information can be calculated:

$$\pi_{\Theta|\mathbf{X}}(\theta|\mathbf{x}) \propto f_{\mathbf{X}|\Theta}(\mathbf{x}|\theta)\pi(\theta). \quad (2)$$

In the absence of a relevant prior information about the prior distribution, $\pi(\theta)$ can be chosen to be a uniform distribution (the case of so called *noninformative priors*).

Often, the expression (2) cannot be used for the direct computation of the posterior distribution and the

* Part of the work was performed within "PRisMa Lab" in Vienna University of Technology.

¹Edgeworth Centre for Financial Mathematics, University College Cork, Ireland, g.temnov@ucc.ie

²CPSE, Imperial College, London, UK, s.kucherenko@ic.ac.uk

stochastic modeling has to be used to produce a pseudo-random sample from $\pi_{\Theta|\mathbf{X}}(\theta|\mathbf{x})$.

The sample from posterior distribution can therefore be used to get the sample from the corresponding compound distribution. Specifically, if the $g(z|\theta)$ is the pdf of the compound distribution (r.v. S in our terms) given a value of the parameter θ , then the corresponding *full predictive distribution* is defined as $h(z|\mathbf{X}) = \int g(z|\theta)\pi(\theta|\mathbf{X})d\theta$ (which is the weighted average with the respect to the distribution of θ as an r.v.).

Obviously, this task is not compatible with *deterministic techniques* (i.e., the ones based on analytical representations of distribution functions), as each time the realization of the predictive distribution is modeled, different values of the parameters θ and λ should be used. Thus, one needs to use MC simulation methods. The common scheme for modeling of the predictive distribution using MC method could be as follows:

1. Simulate the realization of the severity parameters' vector θ and frequency λ from their joint distribution $\pi(\gamma)$ where $\gamma = (\theta, \lambda)$.
2. Given θ and λ generate yearly losses, i.e., (i) generate the number N of yearly losses $N \sim \text{Pois}(\lambda)$ and (ii) generate the sample $\{X_j\} = (X_1, \dots, X_N)$.
3. Given N and $\{X_j\}$ calculate the annual loss $S = \sum_{i=1}^N X_i$.
4. Repeat Steps 1–3 K times to get: $\{S_j\}_{j=1}^K$.
5. Estimate α -quantile, \widehat{Q}_B , of annual loss ($Q_\alpha \sim \text{sort}(S_j)[\alpha]$).

However, there are at least two aspects in this context that make the use of regular MC simulation techniques very time-demanding. These aspects are (i) actuarial losses are often *heavy tailed* and (ii) one usually needs to estimate a sufficiently high quantile of the aggregate loss distribution. In the case of operational risk measurement, the rules prescript to estimate the α -quantile of aggregate loss called *Value-at-Risk*, i.e., $VaR_\alpha := \sup_x (x : F(x) < \alpha)$, at the level $\alpha = 0.999$.

Thus, the precision of the modeled predictive distribution should be high enough to obtain a reliable estimate of the upper quantile. In practical applications dealing with operational risk modeling, it is rather usual that not less than $K = 10^6$ repetitions should be made to ensure necessary precision for 0.999-quantile estimate (see, e.g., [1, 2]).

That motivated authors' search for techniques that would reduce the number of claimed repetitions in the modeling scheme above and lead to QMC methods.

2 Methods

2.1 Deterministic techniques

A short comment on deterministic approaches should be made. Some of the deterministic techniques are based on a passage from probability distributions to characteristic functions or probability generating functions (pgf). This approach is applicable to the task (1) in the case of *iid* summands which will be summarized below.

For a random variable N taking only nonnegative integer values, consider the pgf $P_N(z) = \mathbf{E}[z^N] = \sum_{n=0}^{\infty} \mathbf{P}[N = n]z^n$ which is defined and analytic at least for $|z| \leq 1$. Considering the power series expansion of this function $P_N(z) = \sum_{n=0}^{\infty} p_n z^n$, one is able to retrieve the distribution $\mathbf{P}[N = n] = p_n$ for $n \geq 0$ by calculating the coefficients of $P_N(z)$. Denote the pgf of a compound sum of the form (1) by $P_S(z)$ (*considering integer valued loss sizes*) and using the independence assumption and find

$$P_S(z) = \mathbf{E}[z^S] = \sum_{k=1}^{\infty} \mathbf{P}[N = k] P_X(z)^k.$$

One has the well known representation

$$P_S(z) = P_N(P_X(z)) \tag{3}$$

where $P_N(z)$ is the pgf of the distribution of loss occurrences and $P_X(z)$ corresponds to loss sizes. For Poisson distributed occurrences,

$$P_S(z) = \exp(\lambda(P_X(z) - 1)).$$

Exactly the same representation is valid in terms of characteristic function (chf).

Concerning the calculation of compound distributions **with fixed parameters**, deterministic methods are well developed and include, besides the techniques based of pgf and chf, also, e.g., recursive techniques related to Panjer recursion. For detailed discussion of deterministic techniques in actuarial modeling, see, e.g., [3] and [4]. Deterministic methods are usually more effective than the ones based on MC modeling from the point of precision and speed of calculations. For comparison of the effectiveness of different techniques, see, e.g., [2].

However, when parameters of the distribution of random summands are also random coming from posterior distribution (2) as in the case we deal with, deterministic techniques cannot be used. Indeed, in this case, the pgf $P_X(z)$ in right-hand side of (3) will have no closed form, as it will depend on the distribution of random parameters, which cannot be included into pgf explicitly. Therefore, to handle this problem, one has to turn to MC simulation methods.

2.2 Monte Carlo modeling in heavy tailed cases

Handling heavy tailed distributions, which implies simulation of rare and severe events, has been a challenging task in applied statistics (see, e. g., [5] or [6] for detailed overviews).

Methods usually proposed for solving this problem allow to reduce the computational effort within using the standard MC modeling. According to [5], algorithms involving order statistics and methods using importance sampling are among the most effective techniques for handling random sums. Moreover, various *variance reduction* techniques can be used to increase the efficiency of MC simulation (see, e. g., [7, 8]). However, in certain cases when the resulting distribution depends on random factors, standard variance reduction methods cannot be used, as these techniques rely on explicit representations of distributions used for the modeling. That is indeed the case of our problem, as we need to calculate the random sum of varying volume, and, in addition, the distribution of each variable includes random parameters.

2.3 Quasi-Monte Carlo in risk management

From the previous subsections, motivation for QMC becomes apparent: Clearly, simulation techniques remain a basic tool for modeling compound sums of the form (1) when the distributions involved have stochastic parameters, but dealing with heavy-tailed distributions one needs to handle the variance of simulations in one or another way, and QMC is one of the effective and stable methods to reduce the variance of simulations.

However, QMC methods are not widely used in modeling random sums, as there are certain natural restrictions for using QMC for that particular task. Application of QMC in risk management was studied [9], including of problem of the summation of random variables. In [9], the high-dimensional Sobol' sequences were apply to the problem of risk aggregation for a portfolio of individual losses, when the dimension of the portfolio is fixed. Thus, the problem reduces to the summation of a specified (fixed) number of random variables. It relates the methodology described in [9] to this work, but there are two specific aspects:

- (1) in our case, the number of random variables to be summed up is a (discrete) random variable itself; and
- (2) we are interested particularly in the summation of heavy tailed r.v.'s.

These aspects motivated the authors to make a separate study in order to find out the efficiency in the QMC scheme in the frame of this problem.

3 Using Quasi-Monte Carlo for the Random Loss Aggregation

In the present section, it will be discussed how Sobol' sequence can be used for the summation of random number of random variables.

Constructing Sobol' sequence. The Sobol' sequence is one of the standard quasi-random sequences and is widely used in QMC applications. The construction of Sobol' sequence will not be described here, referring to [10, 11] and related works for technical details.

3.1 Role of independence

Obviously, the advantage of MC techniques in application to a statistical problem is that it allows to model *independent* random variables. As already mentioned, the assumption of independence plays a major role in the modeling of compound distributions.

Recall that the cumulative probability distribution (cdf) of the compound sum is

$$\begin{aligned} F_S(x) &= \mathbf{P}(S \leq x) \\ &= \sum_{k=1}^{\infty} \mathbf{P}[N = k] \mathbf{P}(X_1 + \dots + X_k \leq x) \\ &= \sum_{k=1}^{\infty} p_k F_X^{*k}(x) \end{aligned} \quad (4)$$

where $F_X^{*k}(x)$ is the k -fold convolution of the pdf with itself, i.e.,

$$F_X^{*k}(x) = \int_0^x F_X^{*(k-1)}(x-u) dF_X(u), \quad (5)$$

and $F_X^0(x) \equiv 1$ ($x > 0$).

Once the assumption of independence is dropped, the representation of the compound sum (4) does not reduce to the sum of convolutions (5) any longer. In case of the summation of dependent sequences, in order to calculate the cdf $F_S(s) = \mathbf{P}(X_1 + \dots + X_d \leq s)$ we would have to deal with the integrals of the form $\int_{\Omega_s := [u_1 + \dots + u_d \leq s]} dF(u_1, \dots, u_d)$ instead of multiple convolutions.

Independence and multidimensional Sobol' sequences. Clearly, it is already the scheme of the construction of the low-discrepancy sequences that claims the QMC sequences to be dependent. However, if one uses the elements of the sequences from different dimensions, their relation would "imitate" relation between independent random variables.

The issues related to the independence of different dimensions of Sobol' sequences were discussed in [9] where the tests based on rank correlations were used. For brevity, the results of independence tests are not indicated here, but refer to [9] and related literature stating that spatial distribution of multidimensional Sobol' sequences relates to the distribution of independent random variables.

Note that the notions of "randomness" is understood in the studied case certainly not in its usual way. The observed "distributions" of low-discrepancy sequences would not be empirical probability distributions in its general sense. Nevertheless, observing the spatial structure formed by the sequences, one is able to judge how good is the resulting "imitation" of the independence between values in simulated sequences, due to independence between different dimensions in QMC.

Summarizing the paragraph, note that the right approach for modeling the sums of independent random variables would be to use different (sequential) dimensions for the generation of each of the r.v.'s

3.2 Quasi-Monte Carlo: the modeling set

In operational risk framework, before modeling of the compound sums, one should use historical data to estimate (single-loss) severity and frequency distributions. The historical data of operational risk losses is classified by business lines (BL), and this division is important for the calculation of regulatory capital for OpRisk. Particularly, according to the Basel II recommendations, the Value-at-Risk (VaR) estimators, defining the regulatory capital, should be done for each BL, L_1, \dots, L_K , separately.

In order to estimate the severity and frequency distributions using historical data $\{X_j^{L_j}\}$ in an i.i.d.-case, standard methods such as maximum likelihood estimate (MLE) can be used to estimate the parameters of the distribution $F(x) = \mathbf{P}(X_1^{L_j} < x)$. In the case of uncertain parameters, the distribution parameters are themselves random variables; hence, the "parameters of the parameters" should be estimated. As soon as the estimates are found, the scheme outlined here can be applied to each BL, using either the regular MC, or QMC techniques.

The following notation for Sobol' sequences is used: **Sobol**($\mathbf{i}, \mathbf{n}, \mathbf{d}$) = $S_n^d(i)$ in a unit hypercube $[0, 1]^d$ (here, i is the initial index, n is the length of the sequence, and d is the number of dimensions). Then, if the inverse pdf of an r.v. X , F^{-1} , is known analytically, the realizations of X can be generated via $F^{-1}(S_n^d(i))$.

Then, the whole algorithm can be represented in the following way. Suppose for simplicity that there

are only one parameter of the severity distribution and one frequency parameter (which is often the case in applications).

- I. Simulate an N -length sequence of the severity parameter θ by $I_\theta = F_\theta^{-1}(S_n^{d_\theta}(i_1))$ with **fixed** d_j (i. e., using a one-dimensional QMC sequence), and a chosen initial index i_1 ; an N -length sequence of the frequency parameter λ by $I_\lambda = F_\lambda^{-1}(S_n^{d_\lambda}(i_1))$ likewise.
- II. Simulate N yearly frequencies using λ_j from I_λ , the quantile function of the Poisson distribution and a new QMC-sequence $\{\Lambda_j\} \sim q\text{Pois}_{\lambda_j}^{-1}(S_n^{d_\Lambda}(i_\Lambda))$ with fixed d_Λ and a chosen index i_Λ .
- III. Simulate N sequences of yearly losses $\{X\}_1, \dots, \{X\}_N$ by
 1. $\{X_j^{(1)}\}_{i \leq \Lambda_1} = F_X^{-1}(S_1^{d_1})$
 2. $\{X_j^{(2)}\}_{i \leq \Lambda_2} = F_X^{-1}(S_2^{d_2})$
 - ...
 - N . $\{X_j^{(N)}\}_{i \leq \Lambda_N} = F_X^{-1}(S_N^{d_N})$

The choice of i_Λ can be arbitrary in the initial simulation set, but one repeats the simulation cycle to obtain a different estimate of the quantile, then the initial index of next simulation should be chosen such that the QMC sequence used in the previous cycle is not used again.

Here, d_i changes sequentially in each case (i. e., from 1 to Λ_i for each i).

IV. Loss aggregation.

Consequently, one gets n sums $S_j = \sum_{i=1}^{N_j} X_i^{(j)}$ ($j = 1, \dots, N$) which are the realizations of the aggregate loss that we are interested in. Find the quantile:

- put the obtained sample in increasing order to get the order statistics $L_{1:n} \leq \dots \leq L_{n:n}$ where $L_{1:n}$ denotes the smallest of the n simulations and $L_{n:n}$ the biggest simulated loss; and
- the element at position $[\alpha n + 1]$ of the ordered sample, where $[\cdot]$ denotes rounding downwards, is the estimator of the quantile (i.e., of VaR) to the level α (e.g., choose $\alpha = 0.999$).

Note that in the above construction, two fundamental properties of the multidimensional QMC sequence have been used: (1) the projection of multidimensional QMC into lower dimensions is again a low-discrepancy sequence; and (2) in each multidimensional sequence, the elements corresponding to different dimensions are independent. That is, taking the elements of Sobol'

sequence sequentially from different dimensions, one keeps the properties of a low-discrepancy sequence, still advancing in having the properties of the QMC sequence.

As a result of the above simulation cycle, a single estimate of the 0.999-quantile of the aggregate loss distribution is obtained. To obtain a different realization of 0.999-quantile, one may repeat the whole cycle, correspondingly changing the initial indices i_1, i_λ, i_θ , etc.

3.3 Results and rates of convergence

Next, overview the obtained results. The *SobolSeq* generator (see www.broda.co.uk, 2009, for the full reference) was used for the QMC simulation, while the general modeling was made in SPLus. We were interested in estimating the 0.999-quantile of the aggregate loss distribution.

To analyze the rate of convergence to the true quantile, the performance of the algorithm was traced for a range of the number of simulation N from 10^5 to $2 \cdot 10^6$, for both QMC and pseudo-random realizations.

Remark 3.1. The benchmark for the true quantile we are referring to can be estimated using, e. g., MC Markov Chain (MCMC) modeling, which is often used as an efficient tool for taking parameter uncertainty into account (see, e. g., [1] or [12]). Using MCMC, it is possible to model the so-called full predictive distribution of aggregate losses with random parameters (though modeling with MCMC is as time demanding as regular MC). The true quantile that can be used as a benchmark to estimate the absolute error of MC and QMC modeling is the quantile of the full predictive distribution.

Note that 10^6 simulations required 525 s of the CPU on a Pentium 2.0 GHz, 1 GB memory, for the whole scheme described above including generation of Sobol' sequence and loss aggregation. The same scheme with the MC requires roughly the same time for the same number of simulations.

The model used for the modeling set was GPD-Poisson, i. e., single losses are supposed to have Generalized Pareto distribution (GPD)

$$G(x) = 1 - \left(1 + \xi \frac{x - \mu}{\sigma}\right)^{-1/\xi}$$

while the number of yearly losses follows a Poisson process with intensity λ .

Among GPD parameters, the location parameter μ is usually fixed (it plays a role of the threshold), while shape and scale parameters (ξ, σ) are the ones to be estimated. As Maximum likelihood estimation is often

used, it is natural to assume that the parameters' vector (ξ, σ) has the bivariate normal distribution and also assume normal distribution for the intensity λ . Furthermore, the following parameters for the distributions of the model parameters were used (the choice of the parameters' values reflects typical values in operational risk framework):

- the threshold $\mu = 7000$;
- vector of mean values for the parameters $(\xi, \sigma) = (1, 12000)$, vector of their variance values $(0.18, 1645.0)$, and the covariance value 0.64; and
- mean and variance of the Poisson intensity λ distribution was taken as $(12, 1.7)$.

The plots summarizing the results of the 0.999-quantile modeling illustrating the convergence to the true quantile are presented in Figs. 1 and 2. Figure 1 gives the picture on the log-log scale for the whole range of the simulation numbers, while Fig. 2 concentrates on the segment of the larger values of N on a regular scale where local smooth trend lines show the rate of convergence.

Interpretation of the results indicated on the plots should be made with respect to specific properties of multidimensional Sobol' sequences. Analyzing empirical errors of numerical integration via QMC, [13] shows that actual rates of convergence may differ sufficiently for different types of QMC sequences, depending on the dimension. Main theoretical result on integration errors with QMC known as the *Koksma–Hlawka inequality* states

$$\left| \int_{I^s} f(x) dx - \frac{1}{N} \sum_{i=1}^N f(x_i) \right| \leq V(f) D_N$$

where D_N is the discrepancy and s is the dimension of the integration domain I^s . The bound on the discrepancy of a random sequence indicates $N^{-1/2}$, suggesting that a sequence with smaller discrepancy could give smaller errors. For low-discrepancy sequences, $D_N \sim N^{-\alpha}$ with α depending on a particular type of sequence. Empirically studying the power α of the rate $N^{-\alpha}$, [13] states quite wide range from 1 to 0.45 depending on the dimension of the function.

In the studied case, the dimensions are “floating,” as each of the random sums consists of random number of variables and is therefore generated via a QMC sequence of different (random) dimensions. That fact can at least partially explain the changing rate of convergence to the true quantile observed in Fig. 1. At the same time, as seen from Fig. 2, the rate of convergence for QMC sequences is more stable than the one via pseudo-random numbers simulation, for higher values N corresponding

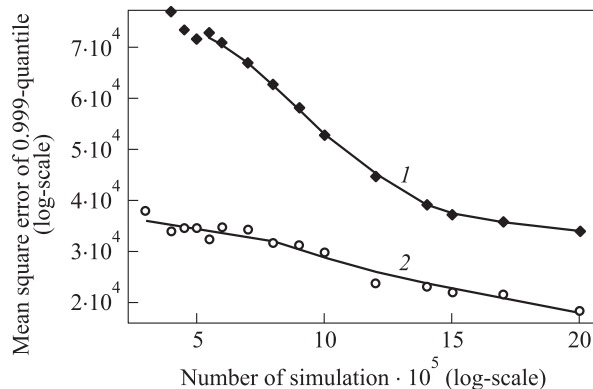


Figure 1 Comparison of the performance of MC (points) with the one obtained by QMC (solid line) for the aggregate–loss distribution of the generalized Pareto

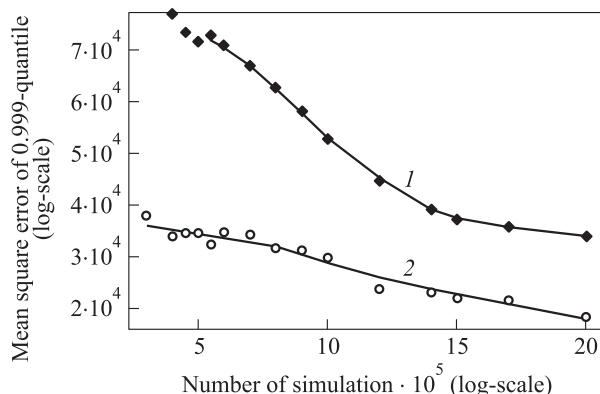


Figure 2 Comparison of the MC precision (1) with the precision of QMC (2) in terms of the mean square error for the aggregate–loss distribution of the generalized Pareto

to higher precision. As estimated by ordinary linear regression, for the range of higher N starting from $8 \cdot 10^5$, the rate of convergence via QMC sequences is $N^{-0.8}$.

Note that not only the rate of convergence can play a role, but also the absolute value of the error. According to [13], for integration problems with discontinuous in high dimensions, the following result is valid

$$|\text{error}| = C_s N^{-s/(2s-1)} \quad (6)$$

where C_s changes depending of the particular type of the low discrepancy sequence and the number of dimensions. Thus, the rate of convergence much better than that of a random sequence cannot be expected; however, the precision still can be improved regarding the constant C_s .

In our case, as illustrated by both Figs. 1 and 2, the coefficient C_s is significantly lower than the one associated with the random sequence, allowing to require much lower number of simulations for obtaining

the same precision of 0.999-quantile, regarding Sobol’ sequence.

References

1. Shevchenko, P. 2008. Estimation of operational risk capital charge under parameter uncertainty. *J. Operational Risk* 3(1):51–63.
2. Temnov, G., and R. Warnung. 2008. A comparison of loss aggregation methods for operational risk. *J. Operational Risk* 3(1):3–23.
3. Klugman, S., H. Panjer, and G. Willmot. 2004. *Loss models. From data to decisions*. Hoboken, NJ, USA: Wiley–Interscience. 2nd ed.
4. Panjer, H. 2006. *Operational risk. Modelling analytics*. Wiley Series in Probability and Statistics. Hoboken, NJ: Wiley-Interscience. 2nd ed.
5. Asmussen, S., K. Binswanger, and B. Hojgaard. 2000. Rare event simulation for heavy tailed distributions. *Bernoulli* 6(2):303–22.

6. Huang, Z., and P. Shahabuddin. 2004. A unified approach for finite-dimensional, rare-event Monte Carlo simulation. *2004 Winter Simulation Conference Proceedings*.
 7. Glasserman, P. 2003. *Monte Carlo methods in financial engineering*. Springer.
 8. Robert, C., and G. Casella. 2004. *Monte Carlo statistical methods*. Springer Texts in Statistics, New York. 2nd ed.
 9. Prakash, S. 2005. *On the use of high dimensional quasi-random sequences for risk measurement*. Master Thesis, ETH Zurich.
 10. Sobol', I. M. 1976. Uniformly distributed sequences with additional uniformity properties. *USSR J. Comput. Math. Phys.* 16(5):236–42.
 11. Sobol', I. M. 1998. On quasi-Monte Carlo integrations. *Math. Computers Simulation* 47:103–12.
 12. Shevchenko, P., and G. Temnov. 2009 (in press.). Modelling operational risk data reported above a time varying threshold. *J. Operational Risk*.
 13. Morokoff, W., and R. Caflisch. 1995. Quasi-Monte Carlo integration. *J. Comput. Physics* 122(2):218–30.
-

ПОДХОД К АКТУАРНОМУ МОДЕЛИРОВАНИЮ НА ОСНОВЕ ПРИМЕНЕНИЯ МЕТОДА КВАЗИ-МОНТЕ-КАРЛО ДЛЯ СЛУЧАЙНЫХ СУММ, ЗАВИСЯЩИХ ОТ СТОХАСТИЧЕСКИХ ФАКТОРОВ

Г. Темнов¹, С. Кучеренко²

¹Корский университет, Ирландия, g.temnov@ucc.ie

²Империял Колледж, Лондон, Великобритания, s.kucherenko@ic.ac.uk

Аннотация: Рассматривается задача оценивания характеристик случайной суммы, в которой число слагаемых также случайно. Рассматриваемый случай включает дополнительный случайный фактор: хотя тип распределения слагаемых известен, параметры этого распределения рассматриваются как случайные величины с известным распределением. Рассматриваемая задача решается с помощью метода квази-Монте-Карло. Анализируется эффективность данного подхода по сравнению с обычным методом Монте-Карло. Рассматриваемые методы имеют применение в актуарной практике, а также при решении некоторых задач информатики, связанных с агрегированием данных с тяжелыми хвостами.

Ключевые слова: актуарное моделирование, метод квази-Монте-Карло, случайные суммы

ОБ УСТОЙЧИВОСТИ РЕКОНСТРУКЦИИ ИЗОБРАЖЕНИЙ В ЗАДАЧАХ ЭМИССИОННОЙ ТОМОГРАФИИ*

О. В. Шестаков¹

Аннотация: В работе рассматривается задача реконструкции изображений по проекционным данным в условиях математической модели эмиссионной томографии. Приводятся оценки точности реконструкции при использовании конечного числа проекций.

Ключевые слова: эмиссионная томография; преобразование Радона; проекции; оценки близости

1 Введение

В задачах эмиссионной томографии возникает проблема обращения преобразования Радона с поглощением, отличающегося от классического преобразования Радона наличием весовой функции, имеющей специальный вид [1]. Вопрос о возможности обращения преобразования Радона с поглощением долгое время оставался открытым, однако в последние годы проблема была успешно решена и были получены различные формулы обращения [1–3], использующие полный набор проекционных данных. Однако на практике известно лишь конечное число проекций, а в этом случае задача обращения уже не имеет единственного решения даже в случае классического преобразования Радона, что приводит к так называемому парадоксу компьютерной томографии [4]. Этот парадокс рассматривался в работах [4–6] для классического и экспоненциального преобразований Радона, и были получены оценки близости между функциями, имеющими одинаковые или близкие проекции по конечному числу заданных направлений. В данной работе эти результаты будут обобщены на случай преобразования Радона с поглощением.

2 Преобразование Радона с поглощением

Пусть $a(x)$ — достаточно гладкая известная функция с компактным носителем. Положим

$$Da(\theta, x) = \int_0^{\infty} a(x + t\theta) dt, \quad x \in \mathbf{R}^2, \quad \theta \in S^1,$$

где S^1 — множество направлений, задаваемых единичными векторами в \mathbf{R}^2 с центром в начале координат. Функция $a(x)$ имеет смысл коэффициента поглощения в данной точке x . Преобразование Радона с поглощением определяется выражением

$$R_a f(\theta, s) = \int_{x\theta=s} f(x) e^{-Da(x, \theta^\perp)} dx, \quad \theta \in S^1, \quad s \in \mathbf{R}, \quad (1)$$

где $f(x)$ — функция изображения, описывающая распределение и интенсивность источников излучения [7], а интеграл берется вдоль прямой $x\theta = s$. Здесь $\theta = (\cos \varphi, \sin \varphi)$, а $\theta^\perp = (-\sin \varphi, \cos \varphi)$. По аналогии с классическим преобразованием Радона будем называть интегральные преобразования вида (1) проекциями. Если в (1) $a(x) = 0$, то $R_a f(\theta, s)$ превращается в классическое преобразование Радона, а если $a(x)$ равна константе μ на носителе $f(x)$, $R_a f(\theta, s)$ превращается в экспоненциальное преобразование Радона.

Всюду в дальнейшем будем предполагать выполнение следующих условий:

- (1) функции $f(x)$ и $a(x)$ неотрицательны, а их носителем является круг U единичного радиуса с центром в начале координат;
- (2) функция $f(x)$ нормирована:

$$\iint_U f(x) dx = 1$$

(т. е. представляет собой вероятностную плотность распределения);

* Работа выполнена при финансовой поддержке РФФИ, грант 08-01-00567.

¹ Московский государственный университет им. М. В. Ломоносова, кафедра математической статистики факультета вычислительной математики и кибернетики, oshestakov@cs.msu.ru

- (3) функция $a(x)$ непрерывно дифференцируема, $\sup_U a(x) \leq \mu$ для некоторой константы μ , $a \sup_{\mathbf{R}^2} |\text{grad } a(x)| = G_a$.

Соответствующие классы всех функций $f(x)$ и $a(x)$, удовлетворяющих этим условиям, будем обозначать через F_U и A_U .

Формула обращения для преобразования Радона с поглощением выглядит следующим образом [2]:

$$f(x) = \frac{1}{4\pi} \times \int_{S^1} \text{Re div } \theta e^{Da(x, \theta^\perp)} (e^{-h} H e^h R_a f) (\theta, x\theta) d\theta. \quad (2)$$

Здесь

$$h = \frac{1}{2} (Ra + iH Ra),$$

$Ra(\theta, s)$ — преобразование Радона от функции $a(x)$, а H — преобразование Гильберта

$$Hg(s) = \frac{1}{\pi} \int_{\mathbf{R}} \frac{g(t)}{s-t} dt,$$

действующее на вторую переменную $Ra(\theta, s)$. В пространстве Фурье для преобразования Гильберта справедливо следующее соотношение [2]:

$$\widehat{Hg}(\omega) = \frac{\text{sgn}(\omega)}{i} \hat{g}(\omega).$$

Формулу (2) можно преобразовать к более удобному виду. Введем функцию $Q_a f = \text{Re } e^{-h} H e^h R_a f$. Полагая $h = h_1 + ih_2$, где $h_1 = (1/2)Ra$, $h_2 = (1/2)H Ra$, имеем

$$Q_a f = e^{-h_1} \cos h_2 H (e^{h_1} \cos h_2 R_a f) + e^{-h_1} \sin h_2 H (e^{h_1} \sin h_2 R_a f).$$

Тогда формула (2) преобразуется к виду

$$f(x) = \frac{1}{4\pi} \text{div} \int_{S^1} \theta e^{Da(x, \theta^\perp)} Q_a f(\theta, x\theta) d\theta.$$

Внося оператор div под знак интеграла, получаем

$$f(x) = \frac{1}{4\pi} \int_{S^1} e^{Da(x, \theta^\perp)} Q_a f(\theta, x\theta) d\theta + \frac{1}{4\pi} \int_{S^1} \frac{\partial}{\partial \theta} Da(x, \theta^\perp) e^{Da(x, \theta^\perp)} Q_a f(\theta, x\theta) d\theta, \quad (3)$$

где $Q_a f(\theta, s)$ — производная функции $Q_a f(\theta, s)$ по второй переменной, а $\partial/(\partial\theta)$ обозначает производную функции $Da(x, \theta^\perp)$ по направлению θ по первой переменной.

3 Оценки точности реконструкции по конечному числу проекций

Задача реконструкции является некорректно поставленной [7], и кроме того, в практических приложениях проекции регистрируются только по конечному числу направлений. Поэтому вместо точной формулы (2) или (3) нужно использовать ее регуляризованный вариант, в котором преобразование Гильберта H заменяется на преобразование H_σ , для которого

$$\widehat{H_\sigma g}(\omega) = \frac{\text{sgn}(\omega)}{i} \hat{g}(\omega) W_\sigma(\omega),$$

где множитель $W_\sigma(\omega)$ играет роль регуляризатора (окна). В результате вместо точной функции $f(x)$ получается приближенная функция $f_\sigma(x)$, но метод реконструкции становится устойчивым к погрешностям в проекционных данных. В дальнейшем будем использовать регуляризатор вида

$$W_\sigma(\omega) = e^{-\omega^2 \sigma^2 / 2}.$$

При выполнении условий (1)–(3), сформулированных в предыдущем разделе, для регуляризованной формулы обращения можно получить оценку точности реконструкции функции по конечному числу проекций. Справедлива следующая теорема.

Теорема 1. Пусть $N = 2n$, где n — натуральное число, и $\theta_1, \dots, \theta_N$ — направления на плоскости \mathbf{R}^2 , выбранные следующим образом:

$$\theta_j = \frac{(\nu_j, -1)}{(\nu_j^2 + 1)^{1/2}}, \quad j = 1, \dots, n,$$

$$\theta_j = \frac{(1, \nu_{j-n})}{(\nu_{j-n}^2 + 1)^{1/2}}, \quad j = n + 1, \dots, 2n,$$

где

$$\nu_k = \cos \frac{\pi(2k-1)}{2n}, \quad k = 1, \dots, n.$$

Если $a \in A_U$, $f, g \in F_U$ и проекции функций f и g по прямым в направлениях θ_j и $-\theta_j$ ($j = 1, \dots, N$) совпадают, т. е.

$$R_a f(\pm\theta_j, s) = R_a g(\pm\theta_j, s), \quad s \in \mathbf{R}, \quad j = 1, \dots, N,$$

тогда

$$\sup_{x \in \mathbf{R}^2} |f_\sigma(x) - g_\sigma(x)| \leq \frac{e^{3\mu}}{n\sigma^2} \left(4G_a + \frac{\mu}{\sigma^2} + \frac{\sqrt{\pi}}{\sigma\sqrt{2}} \right) +$$

$$\begin{aligned}
 & + \frac{e^{3\mu}}{n\sigma} \left(\frac{3\sqrt{\pi}}{\sqrt{2}} G_a + \frac{\sqrt{\pi}\mu}{2\sqrt{2}\sigma^2} \right) \left(4G_a + \right. \\
 & \qquad \qquad \qquad \left. + \frac{\mu}{\sigma^2} + \frac{\sqrt{2}}{\sigma\sqrt{\pi}} \right). \quad (4)
 \end{aligned}
 = \left| \frac{1}{\sqrt{2\pi}} \int_{\mathbf{R}} |\omega| e^{-\sigma^2\omega^2/2} (\hat{q}_{\text{mod}}(\theta, \omega) - \hat{q}_{\text{mod}}(\theta_j, \omega)) e^{i\omega s} d\omega \right|.$$

Доказательство. Используя формулу (3), имеем

$$\begin{aligned}
 |f_\sigma(x) - g_\sigma(x)| &= \left| \frac{1}{4\pi} \int_{S^1} e^{Da(x, \theta^\perp)} Q'_a(\theta, x\theta) d\theta + \right. \\
 & \left. + \frac{1}{4\pi} \int_{S^1} \frac{\partial}{\partial \theta} Da(x, \theta^\perp) e^{Da(x, \theta^\perp)} Q_a(\theta, x\theta) d\theta \right|, \quad (5)
 \end{aligned}$$

где $Q_a = Q_af - Q_ag$, и в определениях Q_af и Q_ag вместо преобразования Гильберта используется его регуляризованный аналог. Обозначим первое слагаемое в (5) через I_1 , а второе — через I_2 и оценим отдельно каждое из них. Для всех $\theta \in [\theta_j - \Delta\theta, \theta_j + \Delta\theta]$, $j = 1, \dots, N$, где $\theta_j - \Delta\theta$ и $\theta_j + \Delta\theta$ — направления, отстоящие на угол $\pi/(2n)$ влево и вправо от направления θ_j , имеем

$$\begin{aligned}
 |Q'_a(\theta, s)| &= \left| (e^{-h_1} \cos h_2 H_\sigma(e^{h_1} \cos h_2 q_a)(\theta, s) + \right. \\
 & \left. + e^{-h_1} \sin h_2 H_\sigma(e^{h_1} \sin h_2 q_a)(\theta, s))'_s \right| \leq \\
 & \leq \left| (e^{-h_1} \cos h_2 H_\sigma(e^{h_1} \cos h_2 q_a)(\theta, s))'_s \right| + \\
 & \quad + \left| (e^{-h_1} \sin h_2 H_\sigma(e^{h_1} \sin h_2 q_a)(\theta, s))'_s \right|,
 \end{aligned}$$

где $q_a = R_af - R_ag$. Оценим первое слагаемое (второе оценивается аналогично):

$$\begin{aligned}
 & \left| (e^{-h_1} \cos h_2 H_\sigma(e^{h_1} \cos h_2 q_a)(\theta, s))'_s \right| = \\
 & = \left| ((e^{-h_1} \cos h_2)(\theta, s) [H_\sigma(e^{h_1} \cos h_2 q_a)(\theta, s) - \right. \\
 & \quad \left. - H_\sigma(e^{h_1} \cos h_2 q_a)(\theta_j, s)])'_s \right| = \\
 & = \left| [(e^{-h_1} \cos h_2)(\theta, s)]'_s [H_\sigma(e^{h_1} \cos h_2 q_a)(\theta, s) - \right. \\
 & \quad \left. - H_\sigma(e^{h_1} \cos h_2 q_a)(\theta_j, s)] \right| + \\
 & + \left| (e^{-h_1} \cos h_2)(\theta, s) [H_\sigma(e^{h_1} \cos h_2 q_a)(\theta, s) - \right. \\
 & \quad \left. - H_\sigma(e^{h_1} \cos h_2 q_a)(\theta_j, s)]'_s \right|.
 \end{aligned}$$

Введем обозначение $q_{\text{mod}}(\theta, s) = e^{h_1} \cos h_2 q_a(\theta, s)$.

Тогда

$$|[H_\sigma q_{\text{mod}}(\theta, s) - H_\sigma q_{\text{mod}}(\theta_j, s)]'_s| =$$

Учитывая, что $a \in A_U$ и $f, g \in F_U$, можно показать, что для всех $\theta \in [\theta_j - \Delta\theta, \theta_j + \Delta\theta]$ и $\theta \in [-\theta_j - \Delta\theta, -\theta_j + \Delta\theta]$, $j = 1, \dots, N$, (а значит, для всех $\theta \in S^1$) при использовании регуляризованного преобразования Гильберта

$$\begin{aligned}
 |\hat{q}_{\text{mod}}(\theta, \omega) - \hat{q}_{\text{mod}}(\theta_j, \omega)| &\leq \\
 &\leq \frac{\sqrt{\pi}e^\mu}{\sqrt{2n}} \left(4G_a + \frac{\mu}{\sigma^2} + |\omega| \right). \quad (6)
 \end{aligned}$$

Следовательно,

$$\begin{aligned}
 |[H_\sigma q_{\text{mod}}(\theta, s) - H_\sigma q_{\text{mod}}(\theta_j, s)]'_s| &\leq \\
 &\leq \frac{e^\mu}{n\sigma^2} \left(4G_a + \frac{\mu}{\sigma^2} + \frac{\sqrt{\pi}}{\sigma\sqrt{2}} \right).
 \end{aligned}$$

Далее

$$\begin{aligned}
 |H_\sigma q_{\text{mod}}(\theta, s) - H_\sigma q_{\text{mod}}(\theta_j, s)| &= \\
 &= \left| \frac{1}{\sqrt{2\pi}} \int_{\mathbf{R}} \frac{\text{sgn}(\omega)}{i} e^{-\sigma^2\omega^2/2} (\hat{q}_{\text{mod}}(\theta, \omega) - \right. \\
 & \quad \left. - \hat{q}_{\text{mod}}(\theta_j, \omega)) e^{i\omega s} d\omega \right| \leq \\
 &\leq \frac{\sqrt{2\pi}e^\mu}{2n\sigma} \left(4G_a + \frac{\mu}{\sigma^2} + \frac{\sqrt{2}}{\sigma\sqrt{\pi}} \right). \quad (7)
 \end{aligned}$$

Также можно показать, что

$$\left| [(e^{-h_1} \cos h_2)(\theta, s)]'_s \right| \leq G_a + \frac{\mu}{2\sigma^2}. \quad (8)$$

Учитывая (6)–(8), получаем

$$\begin{aligned}
 |Q'_a(\theta, s)| &\leq \frac{2e^\mu}{n\sigma^2} \left(4G_a + \frac{\mu}{\sigma^2} + \frac{\sqrt{\pi}}{\sigma\sqrt{2}} \right) + \\
 &+ \frac{2e^\mu}{n\sigma} \left(\frac{\sqrt{\pi}}{\sqrt{2}} G_a + \frac{\sqrt{\pi}\mu}{2\sqrt{2}\sigma^2} \right) \left(4G_a + \frac{\mu}{\sigma^2} + \frac{\sqrt{2}}{\sigma\sqrt{\pi}} \right).
 \end{aligned}$$

Следовательно,

$$\begin{aligned}
 I_1 &\leq \frac{1}{4\pi} \int_{S^1} e^{Da(x, \theta^\perp)} \max_{\theta, s} (Q'_a(\theta, s)) d\theta \leq \\
 &\leq \frac{e^{3\mu}}{n\sigma^2} \left(4G_a + \frac{\mu}{\sigma^2} + \frac{\sqrt{\pi}}{\sigma\sqrt{2}} \right) +
 \end{aligned}$$

$$+ \frac{e^{3\mu}}{n\sigma} \left(\frac{\sqrt{\pi}}{\sqrt{2}} G_a + \frac{\sqrt{\pi}\mu}{2\sqrt{2}\sigma^2} \right) \left(4G_a + \frac{\mu}{\sigma^2} + \frac{\sqrt{2}}{\sigma\sqrt{\pi}} \right). \quad (9)$$

Оценим теперь слагаемое I_2 . Учитывая условия, которым удовлетворяет функция $a(x)$, можно показать, что

$$\frac{\partial}{\partial \theta} Da(x, \theta^\perp) \leq 2G_a.$$

Следовательно, с учетом (7) имеем

$$I_2 \leq \frac{\sqrt{2\pi}e^{3\mu}G_a}{n\sigma} \left(4G_a + \frac{\mu}{\sigma^2} + \frac{\sqrt{2}}{\sigma\sqrt{\pi}} \right). \quad (10)$$

Объединяя (9) и (10), получаем (4). Теорема доказана.

В практических ситуациях в силу несовершенства оборудования и наличия шума проекции регистрируются с некоторой погрешностью. Если предположить, что погрешность не превосходит заданного уровня ε , то можно получить оценку точности реконструкции с учетом этой погрешности. Справедлива следующая теорема.

Теорема 2. Пусть проекции от $f, g \in \mathbf{F}_U$ по прямым в направлениях θ_j и $-\theta_j$ ($j = 1, \dots, N$), где θ_j те же, что и в предыдущей теореме, отличаются не более чем на ε , т. е. для некоторого фиксированного $\varepsilon \in (0, 1)$

$$|R_a f(\pm\theta_j, s) - R_a g(\pm\theta_j, s)| \leq \varepsilon, \quad s \in \mathbf{R}, \quad j = 1, \dots, N,$$

и $a \in A_U$, тогда

$$\begin{aligned} \sup_{x \in \mathbf{R}^2} |f_\sigma(x) - g_\sigma(x)| &\leq \frac{e^{3\mu}}{n\sigma^2} \left(4G_a + \frac{\mu}{\sigma^2} + \frac{\sqrt{\pi}}{\sigma\sqrt{2}} \right) + \\ &+ \frac{e^{3\mu}}{n\sigma} \left(\frac{3\sqrt{\pi}}{\sqrt{2}} G_a + \frac{\sqrt{\pi}\mu}{2\sqrt{2}\sigma^2} \right) \left(4G_a + \frac{\mu}{\sigma^2} + \frac{\sqrt{2}}{\sigma\sqrt{\pi}} \right) + \\ &+ \frac{2\varepsilon e^{3\mu}}{\sigma\sqrt{2\pi}} \left(5G_a + \frac{3\mu}{2\sigma^2} + \frac{\sqrt{2}}{\sigma\sqrt{\pi}} + \right. \\ &\left. + \frac{1}{n} \left(G_a + \frac{\mu}{2\sigma^2} \right) \left(\pi G_a + \frac{\sqrt{\pi}}{\sigma\sqrt{2}} \right) \right). \quad (11) \end{aligned}$$

Доказательство. Поступая так же, как в предыдущей теореме при оценивании I_1 , имеем

$$|Q'_a(\theta, s)| \leq |Q'_a(\theta_j, s)| + |Q'_a(\theta, s) - Q'_a(\theta_j, s)|. \quad (12)$$

Можно показать, что

$$|Q'_a(\theta_j, s)| \leq \frac{4\varepsilon e^\mu}{\sigma\sqrt{2\pi}} \left(G_a + \frac{\mu}{2\sigma^2} \right) + \frac{4\varepsilon e^\mu}{\sigma^2\pi}. \quad (13)$$

Далее

$$\begin{aligned} |Q'_a(\theta, s) - Q'_a(\theta_j, s)| &\leq \\ &\leq \left| (e^{-h_1} \cos h_2 H_\sigma(e^{h_1} \cos h_2 q_a)(\theta, s) - \right. \\ &\quad \left. - e^{-h_1} \cos h_2 H_\sigma(e^{h_1} \cos h_2 q_a)(\theta_j, s))'_s \right| + \\ &+ \left| (e^{-h_1} \sin h_2 H_\sigma(e^{h_1} \sin h_2 q_a)(\theta, s) - \right. \\ &\quad \left. - e^{-h_1} \sin h_2 H_\sigma(e^{h_1} \sin h_2 q_a)(\theta_j, s))'_s \right|. \end{aligned}$$

Оценим первое слагаемое (второе оценивается аналогично):

$$\begin{aligned} &\left| (e^{-h_1} \cos h_2 H_\sigma(e^{h_1} \cos h_2 q_a)(\theta, s) - \right. \\ &\quad \left. - e^{-h_1} \cos h_2 H_\sigma(e^{h_1} \cos h_2 q_a)(\theta_j, s))'_s \right| \leq \\ &\leq \left| [e^{-h_1} \cos h_2(\theta, s) (H_\sigma(e^{h_1} \cos h_2 q_a)(\theta, s) - \right. \\ &\quad \left. - H_\sigma(e^{h_1} \cos h_2 q_a)(\theta_j, s))]'_s \right| + \\ &+ \left| [H_\sigma(e^{h_1} \cos h_2 q_a)(\theta_j, s) (e^{-h_1} \cos h_2(\theta, s) - \right. \\ &\quad \left. - e^{-h_1} \cos h_2(\theta_j, s))]'_s \right|. \end{aligned}$$

Так же, как в предыдущей теореме, убеждаемся, что

$$\begin{aligned} &\left| [e^{-h_1} \cos h_2(\theta, s) (H_\sigma(e^{h_1} \cos h_2 q_a)(\theta, s) - \right. \\ &\quad \left. - H_\sigma(e^{h_1} \cos h_2 q_a)(\theta_j, s))]'_s \right| \leq \\ &\leq \frac{e^\mu}{n\sigma^2} \left(4G_a + \frac{\mu}{\sigma^2} + \frac{\sqrt{\pi}}{\sigma\sqrt{2}} \right) + \\ &+ \frac{e^\mu}{n\sigma} \left(\frac{\sqrt{\pi}}{\sqrt{2}} G_a + \frac{\sqrt{\pi}\mu}{2\sqrt{2}\sigma^2} \right) \left(4G_a + \right. \\ &\quad \left. + \frac{\mu}{\sigma^2} + \frac{\sqrt{2}}{\sigma\sqrt{\pi}} \right). \quad (14) \end{aligned}$$

Далее можно показать, что при выполнении условий теоремы

$$\begin{aligned} |H_\sigma(e^{h_1} \cos h_2 q_a)(\theta_j, s)| &\leq \frac{2\varepsilon e^\mu}{\sigma\sqrt{2\pi}}, \\ |[H_\sigma(e^{h_1} \cos h_2 q_a)(\theta_j, s)]'_s| &\leq \frac{2\varepsilon e^\mu}{\sigma^2\pi}, \\ |[e^{-h_1} \cos h_2(\theta, s)]'_s| &\leq G_a + \frac{\mu}{2\sigma^2} \end{aligned}$$

и

$$\begin{aligned} |(e^{-h_1} \cos h_2(\theta, s) - e^{-h_1} \cos h_2(\theta_j, s))| &\leq \\ &\leq \frac{\pi}{2n} \left(G_a + \frac{\mu}{2\sigma^2} \right). \end{aligned}$$

Следовательно,

$$\begin{aligned} & \left| [H_\sigma(e^{h_1} \cos h_2 q_a)(\theta_j, s) (e^{-h_1} \cos h_2(\theta, s) - \right. \\ & \quad \left. - e^{-h_1} \cos h_2(\theta_j, s))]'_s \right| \leq \\ & \leq 2\varepsilon e^{\mu} \left(\frac{\sqrt{2}}{\sigma\sqrt{\pi}} + \frac{1}{2n\sigma^2} \right) \left(G_a + \frac{\mu}{2\sigma^2} \right). \end{aligned} \quad (15)$$

При оценивании I_2 имеем

$$|Q_a(\theta, s)| \leq |Q_a(\theta_j, s)| + |Q_a(\theta, s) - Q_a(\theta_j, s)|. \quad (16)$$

Можно показать, что

$$|Q_a(\theta_j, s)| \leq \frac{4\varepsilon e^{\mu}}{\sigma\sqrt{2\pi}}. \quad (17)$$

Далее

$$\begin{aligned} |Q_a(\theta, s) - Q_a(\theta_j, s)| &\leq \\ &\leq |(e^{-h_1} \cos h_2 H_\sigma(e^{h_1} \cos h_2 q_a)(\theta, s) - \\ & \quad - e^{-h_1} \cos h_2 H_\sigma(e^{h_1} \cos h_2 q_a)(\theta_j, s))| + \\ & \quad + |(e^{-h_1} \sin h_2 H_\sigma(e^{h_1} \sin h_2 q_a)(\theta, s) - \\ & \quad - e^{-h_1} \sin h_2 H_\sigma(e^{h_1} \sin h_2 q_a)(\theta_j, s))|. \end{aligned}$$

Оценим первое слагаемое (второе оценивается аналогично):

$$\begin{aligned} & |(e^{-h_1} \cos h_2 H_\sigma(e^{h_1} \cos h_2 q_a)(\theta, s) - \\ & \quad - e^{-h_1} \cos h_2 H_\sigma(e^{h_1} \cos h_2 q_a)(\theta_j, s))| \leq \\ & \leq |[e^{-h_1} \cos h_2(\theta, s) (H_\sigma(e^{h_1} \cos h_2 q_a)(\theta, s) - \\ & \quad - H_\sigma(e^{h_1} \cos h_2 q_a)(\theta_j, s))] + \\ & \quad + |[H_\sigma(e^{h_1} \cos h_2 q_a)(\theta_j, s) (e^{-h_1} \cos h_2(\theta, s) - \\ & \quad - e^{-h_1} \cos h_2(\theta_j, s))]|. \end{aligned}$$

Так же, как в предыдущей теореме, убеждаемся, что

$$\begin{aligned} & |[e^{-h_1} \cos h_2(\theta, s) (H_\sigma(e^{h_1} \cos h_2 q_a)(\theta, s) - \\ & \quad - H_\sigma(e^{h_1} \cos h_2 q_a)(\theta_j, s))]| \leq \\ & \leq \frac{\sqrt{2\pi}e^{\mu}}{2n\sigma} \left(4G_a + \frac{\mu}{\sigma^2} + \frac{\sqrt{2}}{\sigma\sqrt{\pi}} \right). \end{aligned}$$

Следовательно,

$$\begin{aligned} |Q_a(\theta, s) - Q_a(\theta_j, s)| &\leq \\ &\leq \frac{\sqrt{2\pi}e^{\mu}}{n\sigma} \left(4G_a + \frac{\mu}{\sigma^2} + \frac{\sqrt{2}}{\sigma\sqrt{\pi}} \right) + \\ & \quad + \frac{\varepsilon\sqrt{2\pi}e^{\mu}}{n\sigma} \left(G_a + \frac{\mu}{2\sigma^2} \right). \end{aligned} \quad (18)$$

Объединяя (12)–(18) и используя (5), получаем (11). Теорема доказана.

Правая часть в оценке (4) из теоремы 1 с ростом n убывает со скоростью $O(1/n)$, а правая часть в оценке (11) из теоремы 2 с ростом n и уменьшением ε убывает со скоростью $O(\varepsilon) + O(1/n)$. Это означает, что использование регуляризованной формулы обращения (5) приводит к устойчивому методу реконструкции.

Литература

1. *Arbuzov E. V., Bukhgeim A. L., Kazantsev S. G.* Two-dimensional tomography problems and the theory of A-analytic functions // *Siberian Adv. Math.*, 1998. Vol. 8. P. 1–20.
2. *Natterer F.* Inversion of the attenuated Radon transform // *Inverse Problems*, 2001. Vol. 17. P. 113–119.
3. *Novikov R. G.* An inversion formula for the attenuated X-ray transformation // *Ark. Mat.*, 2002. Vol. 40. P. 145–167.
4. *Khalfin L. A., Klebanov L. B.* A solution of the computer tomography paradox and estimating the distances between the densities of measures with the same marginals // *The Annals of Probability*, 1994. Vol. 22. No. 4. P. 2235–2241.
5. *Шестаков О. В., Савенков Т. Ю.* Оценка расстояния между плотностями вероятностных мер, имеющих близкие проекции // *Вестн. Моск. ун-та. Сер. 15. Вычисл. матем. и киберн.*, 2001. № 4. С. 44–46.
6. *Шестаков О. В.* Оценка точности восстановления функции по ее экспоненциальному преобразованию Радона при использовании конечного числа проекций // *Вестн. Моск. ун-та. Сер. 15. Вычисл. матем. и киберн.*, 2006. № 4. С. 22–25.
7. *Федоров Г. А.* Вычислительная эмиссионная томография. — М.: Энергоатомиздат, 1990.

О ВЕРОЯТНОСТИ ИСПРАВЛЕНИЯ ОШИБОК ПРИ ПОМЕХОУСТОЙЧИВОМ КОДИРОВАНИИ, КОГДА ЧИСЛО ОШИБОК ПРИНАДЛЕЖИТ НЕКОТОРОМУ КОНЕЧНОМУ МНОЖЕСТВУ

А. Н. Чупрунов¹, Б. И. Хамдеев²

Аннотация: Рассматриваются n сообщений, каждое из которых состоит из N блоков. Каждый блок кодируется помехоустойчивым кодом, который может исправить не более q ошибок. При этом предполагается, что число ошибок в каждом сообщении принадлежит некоторому конечному подмножеству множества натуральных чисел. В работе изучается вероятность $\mathbf{P}(A)$ события A , состоящего в том, что все ошибки будут исправлены. Вероятность $\mathbf{P}(A)$ формулируется в терминах условных вероятностей. Показано, что при $n, N \rightarrow \infty$ так, что $\alpha = n/N \rightarrow \alpha_0 < \infty$, при $q = 1$ вероятности $\mathbf{P}(A)$ сходятся и найдено значение этого предела, $\mathbf{P}(A) \rightarrow 1$, при $q > 1$.

Ключевые слова: условная вероятность; обобщенная схема размещения; код Хемминга

1 Введение

Будем рассматривать код, который позволяет исправить не больше q ошибок типа замещения, т. е. сферически упакованный код. Частным случаем такого кода является код Хемминга (см. о коде Хемминга, например, в [1]), который позволяет исправлять единичные ошибки. Работа посвящена изучению вероятности события A , состоящего в том, что в n сообщениях, каждое из которых состоит из N блоков и каждый блок кодируется помехоустойчивым кодированием, все ошибки будут исправлены.

Пусть случайная величина ξ_{ij} — количество ошибок в j -м блоке i -го сообщения. Будем предполагать, что ξ_{ij} — независимые неотрицательные целочисленные случайные величины, распределенные так же, как случайная величина ξ . Кроме того, будем предполагать, что число ошибок в сообщении принадлежит некоторому (одному и тому же для всех сообщений) подмножеству \mathbf{N}' множества натуральных чисел \mathbf{N} . Тогда $A = \bigcap_{i=1}^n A_i$, где событие

$$A_i = \{ \xi_{i1} \leq q, \xi_{i2} \leq q, \dots, \xi_{iN} \leq q \mid \xi_{i1} + \dots + \xi_{iN} = m \}$$

для некоторого $m \in \mathbf{N}'$

состоит в том, что в i -м сообщении каждый блок имеет не более q ошибок при условии, что число ошибок в сообщении принадлежит множеству \mathbf{N}' . Поэтому вероятность события A равна

$$\mathbf{P}(A) = (\mathbf{P}\{ \xi_1 \leq q, \xi_2 \leq q, \dots, \xi_N \leq q \mid \xi_1 + \dots + \xi_N = m \text{ для некоторого } m \in \mathbf{N}' \})^n,$$

где случайные величины $\xi_1, \xi_2, \dots, \xi_N$ независимы и распределены так же, как случайная величина ξ .

Будем предполагать, что распределение случайной величины ξ имеет следующий вид, зависящий от параметра θ . Пусть существует последовательность неотрицательных чисел b_0, b_1, \dots такая, что радиус сходимости R ряда

$$B(\theta) = \sum_{k=0}^{\infty} \frac{b_k \theta^k}{k!}$$

положителен. Тогда случайная величина $\xi = \xi(\theta)$, $0 < \theta < R$, распределена по следующему закону:

$$\mathbf{P}\{\xi = k\} = \frac{b_k \theta^k}{k! B(\theta)}, \quad k = 0, 1, 2, \dots$$

Везде далее будем предполагать, что $b_0, b_1 > 0$.

Пусть $\alpha = n/N$. В работе изучается асимптотическое поведение вероятности $\mathbf{P}(A)$ при $n, N \rightarrow \infty$ так, что $\alpha \rightarrow \alpha_0$, $\alpha_0 < \infty$ и множество \mathbf{N}' конечно. В первой части предполагается, что случайные величины $\xi = \xi(\theta)$ и последовательность $(b_1/b_0)N\theta$ сходятся к конечному пределу. Показано, что для

¹Научно-исследовательский институт математики и механики им. Н. Г. Чеботарева, achuprunov@mail.ru

²Научно-исследовательский институт математики и механики им. Н. Г. Чеботарева, khamdeyevbi@mail.ru

$q = 1$ при этих условиях вероятность $\mathbf{P}(A)$ сходится, и найдено значение ее предела; $\mathbf{P}(A) \rightarrow 0$ для $q > 1$, и найдена скорость этой сходимости.

Частным случаем случайных величин $\xi_i(\theta)$ являются пуассоновские случайные величины с параметром $\lambda = \theta$ (случай: $b_0 = b_1 = \dots$). Во второй части работы получены аналоги результатов первой части для случая, когда ξ_i — независимые одинаково распределенные пуассоновские случайные величины. При этом предполагается, что параметр $\lambda = \theta$ — константа. При $q = 1$ показано, что $\mathbf{P}(A)$ сходится, и найдено значение этого предела, которое, как оказалось, зависит только от максимального элемента множества \mathbf{N}' и α_0 и не зависит от $\lambda = \theta$; $\mathbf{P}(A) \rightarrow 1$ для $q > 1$, и найдена скорость этой сходимости.

Если множество \mathbf{N}' состоит из одного элемента ($\mathbf{N}' = \{m\}$), то событие A_i является событием обобщенной схемы размещения: событие A_i состоит в том, что в обобщенной схеме размещения m частиц по N ячейкам в каждой ячейке оказалось не более q частиц. Обобщенную схему размещения ввел В. Ф. Колчин в [2] (см. также о ней монографию В. Ф. Колчина [3]). Случайные величины $\xi_i(\theta)$ были введены А. В. Колчиным в [4]. В работах В. Ф. Колчина и А. В. Колчина [4–6] получены предельные теоремы для сумм случайных величин $\xi_i(\theta)$. Теорема 2 этой работы основана на уточнении пуассоновской предельной теоремы из [4]. Результаты настоящей работы являются обобщением результатов работы [7], полученных для схемы размещения различных частиц по ячейкам, на обобщенную схему размещения.

2 Асимптотическое поведения вероятности $\mathbf{P}(A)$ в случае

$$\xi = \xi(\theta)$$

Рассмотрим вероятность $B(m, N) = \mathbf{P}\{\xi_j \leq 1, 1 \leq j \leq N, \xi_1 + \xi_2 + \xi_3 + \dots + \xi_N = m\}$. Легко заметить, что

$$B(m, N) = C_N^m \left(\frac{b_1 \theta}{B(\theta)} \right)^m \left(\frac{b_0}{B(\theta)} \right)^{N-m}.$$

Будем обозначать через $\pi_\lambda(m) = e^{-\lambda}(\lambda^m/m!)$, $m = 0, 1, \dots$, пуассоновские вероятности. Пусть $(b_1/b_0)N\theta = \lambda$. Тогда

$$B(m, N) = \pi_\lambda(m) \left(1 + \frac{O(1)}{N} \right) \text{ для всех } m = 0, 1, 2, \dots, \quad (1)$$

где $O(1)$ ограничена по N и $\lambda > 0$, принадлежащей ограниченному множеству. Будем использовать неравенства:

$$C_N^k \frac{1}{N^k} \leq \frac{1}{k!} \quad \frac{b_0}{B(\theta)} \leq 1. \quad (2)$$

В работе А. Ф. Колчина [4] (теорема 2) доказано, что если $(b_1/b_0)N\theta \rightarrow \lambda$ при $N \rightarrow \infty$, то

$$\mathbf{P}\{\xi_1 + \xi_2 + \dots + \xi_N = m\} \rightarrow \pi_\lambda(m) \quad (3)$$

для всех $m = 0, 1, \dots$. Сделаем следующее уточнение теоремы А. Ф. Колчина.

Лемма 1. Пусть $\theta = \theta(N) = b_0\lambda/(b_1N) = C_1/N$ и $C_2 = (b_2b_0\lambda^2)/(2(b_1)^2)$. Выберем $N_0 \in \mathbf{N}$ такое, что

$$C_3 = \sum_{i=3}^{\infty} \frac{b_i}{i!} (\theta(N_0))^{i-3} < \infty.$$

Тогда при $N > N_0$ и $m \leq N$ имеем:

$$\mathbf{P}\{\xi_1 + \xi_2 + \dots + \xi_N = m\} = B(m, N) \quad (4)$$

при $m = 0, 1$;

$$\begin{aligned} \mathbf{P}\{\xi_1 + \xi_2 + \dots + \xi_N = m\} = \\ = B(m, N) + \frac{1}{N} \frac{b_2(C_1)^2}{2B(\theta)} B(m-2, N-1) \end{aligned} \quad (5)$$

при $m = 2$;

$$\begin{aligned} \mathbf{P}\{\xi_1 + \xi_2 + \dots + \xi_N = m\} = \\ = B(m, N) + \frac{1}{N} \frac{b_2(C_1)^2}{2B(\theta)} B(m-2, N-1) + \frac{C}{N^2}, \end{aligned} \quad (6)$$

где $0 \leq C \leq C_4$ и $C_4 = e^\lambda(b_3/(6b_0))C_1^3$, при $m = 3$;

$$\begin{aligned} \mathbf{P}\{\xi_1 + \xi_2 + \dots + \xi_N = m\} = \\ = B(m, N) + \frac{1}{N} \frac{b_2(C_1)^2}{2B(\theta)} B(m-2, N-1) + \frac{C}{N^2}, \end{aligned} \quad (7)$$

где $0 \leq C \leq C_4$ и $C_4 = e^\lambda(b_3/(6b_0))C_1^3 + e^{(C_2/N)+\lambda}(C_3C_1^3 + C_2^2/2)$, при $m > 3$.

Доказательство. Обозначим

$$\nu_1 = \sum_{j=1}^N \xi_j I_{\{\xi_j \leq 1\}}, \quad \nu_2(k) = \sum_{j=1}^k \xi_j I_{\{\xi_j \geq 2\}}.$$

При $k = N$ будем обозначать случайную величину $\nu_2(k)$ через ν_2 . Рассмотрим выражение

$$\begin{aligned} \mathbf{P}\{\xi_1 + \xi_2 + \dots + \xi_N = m\} = \\ = \sum_{i=0}^m \mathbf{P}(\nu_1 = m-i, \nu_2 = i). \end{aligned} \quad (8)$$

Заметим, что

$$\begin{aligned} \mathbf{P}(\nu_1 = m, \nu_2 = 0) &= B(m, N); \\ \mathbf{P}(\nu_1 = m - 1, \nu_2 = 1) &= 0 \end{aligned} \quad (9)$$

$$\begin{aligned} \mathbf{P}(\nu_1 = m - 2, \nu_2 = 2) &= \\ &= \sum_{j=1}^N \mathbf{P}(\nu_1 = m - 2, \xi_j = 2, \xi_i < 2, i \neq j) = \\ &= N\mathbf{P}\left(\sum_{j=1}^N \xi_j I_{\{\xi_j \leq 1\}} = m - 2, \right. \\ &\quad \left. \xi_1 = 2, \xi_i < 2, 2 \leq i \leq N\right) = \\ &= N\mathbf{P}(\xi_1 = 2)\mathbf{P}\left(\sum_{j=2}^N \xi_j I_{\{\xi_j \leq 1\}} = m - 2, \xi_i < 2, \right. \\ &\quad \left. 2 \leq i \leq N\right) = N \frac{b_2 \theta^2}{2B(\theta)} B(m - 2, N - 1), \end{aligned} \quad (10)$$

и, так как $\mathbf{P}(\xi = 1) \leq \lambda/N$, имеем

$$\begin{aligned} \mathbf{P}(\nu_1 = m - 3, \nu_2 = 3) &= \\ &= C_N^{m-3} (\mathbf{P}(\xi = 1))^{m-3} \mathbf{P}(\nu_2(N - (m - 3)) = 3) \leq \\ &\leq C_N^{m-3} \frac{\lambda^{m-3}}{N^{m-3}} (N - (m - 3)) \frac{b_3}{6B(\theta)} \theta^3 \leq \\ &\leq \frac{e^\lambda}{N^2} \frac{b_3}{6b_0} (C_1)^3. \end{aligned} \quad (11)$$

Равенства (8) и (9) влекут (4), равенства (8)–(10) влекут (5), равенства (8)–(10) и оценка (11) влекут (6).

Рассмотрим

$$\begin{aligned} \sum_{i=4}^m \mathbf{P}(\nu_1 = m - i, \nu_2 = i) &= \\ &= \sum_{i=4}^m C_N^{m-i} (\mathbf{P}(\xi = 1))^{m-i} \mathbf{P}(\nu_2(N - (m - i)) = i). \end{aligned}$$

Обозначим $\nu_3(k) = \sum_{j=1}^k \xi_j I_{\{\xi_j \geq 3\}}$. Так как

$$\mathbf{P}(\xi = 2) \leq \frac{C_2}{N^2},$$

то

$$\begin{aligned} \mathbf{P}(\nu_2(N - (m - i)) = i) &= \\ &= \sum_{k=0}^{i/2} C_{N-(m-i)}^k (\mathbf{P}(\xi = 2))^k \times \\ &\quad \times \mathbf{P}(\nu_3(N - (m - i) - k) = i - 2k) \leq \end{aligned}$$

$$\begin{aligned} &\leq \sum_{k=0}^{\lfloor i/2 \rfloor} C_{N-(m-i)}^k \left(\frac{C_2}{N^2}\right)^k \times \\ &\quad \times \mathbf{P}(\nu_3(N - (m - i) - k) = i - 2k) \leq \\ &\leq \sum_{k=0}^{\lfloor i/2 \rfloor} \frac{1}{k!} \left(\frac{C_2}{N}\right)^k \times \\ &\quad \times \mathbf{P}(\nu_3(N - (m - i) - k) = i - 2k). \end{aligned} \quad (12)$$

Заметим, что

$$\mathbf{P}(\nu_3(k) = j) \leq \mathbf{P}(\nu_3(k) \geq j) \leq k\mathbf{P}(\xi \geq 3) \leq \frac{C_3}{N^2}$$

при $k \leq N$ и $j \geq 3$. Поэтому имеем

$$\begin{aligned} \mathbf{P}(\nu_2(N - (m - i)) = i) &\leq \\ &\leq \sum_{k=0}^{(i-1)/2} \frac{(C_2/N)^k}{k!} \frac{C_3 C_1^3}{N^2} \leq \frac{C_3^3}{N^2} \exp\left(\frac{C_2}{N}\right), \end{aligned} \quad (13)$$

если i нечетно, и

$$\begin{aligned} \mathbf{P}(\nu_2(N - (m - i)) = i) &\leq \\ &\leq \sum_{k=0}^{i/2-1} \frac{(C_2/N)^k}{k!} \frac{C_3 C_1^3}{N^2} + \frac{(C_2/N)^{i/2}}{(i/2)!} \leq \\ &\leq \frac{C_3 C_1^3}{N^2} \exp\left(\frac{C_2}{N}\right) + \frac{(C_2/N)^{i/2}}{(i/2)!}, \end{aligned} \quad (14)$$

если i четно. Из (12)–(14) вытекает

$$\begin{aligned} \sum_{i=4}^m \mathbf{P}(\nu_1 = m - i, \nu_2 = i) &\leq \\ &\leq \frac{C_3 C_1^3}{N^2} e^{C_2/N} \sum_{i=4}^m C_N^{m-i} \left(\frac{\lambda}{N}\right)^{m-i} + \\ &+ \sum_{4 \leq i \leq m, i \in 2\mathbf{N}} C_N^{m-i} \left(\frac{\lambda}{N}\right)^{m-i} \frac{(C_2/N)^{i/2}}{(i/2)!} \leq \\ &\leq \frac{C_3 C_1^3}{N^2} e^{\lambda + C_2/N} + \left(\frac{C_2}{N}\right)^2 \times \\ &\times \sum_{4 \leq i \leq m, i \in 2\mathbf{N}} C_N^{m-i} \left(\frac{\lambda}{N}\right)^{m-i} \frac{(C_2/N)^{i/2-2}}{(i/2-2)!} \leq \\ &\leq \frac{C_3 C_1^3}{N^2} e^{C_2/N + \lambda} + \frac{(C_2)^2}{2N^2} e^{C_2/N} \times \\ &\quad \times \sum_{i=4}^m C_N^{m-i} \left(\frac{\lambda}{N}\right)^{m-i} \leq \\ &\leq \frac{1}{N^2} e^{C_2/N + \lambda} \left(C_3 C_1^3 + \frac{C_2^2}{2}\right), \end{aligned} \quad (15)$$

где $2\mathbf{N}$ — множество четных чисел.

Тогда (8)–(11) и (15) влекут (7). Доказательство закончено.

Замечание 1. Из леммы 1 следует, что

$$\begin{aligned} \mathbf{P}(\xi_1 + \xi_2 \cdots + \xi_N = m) &= \\ &= \pi_\lambda(m) \left(1 + \frac{O(1)}{N}\right) \quad \text{для всех } m = 0, 1, 2, \dots, \end{aligned}$$

где $O(1)$ ограничена по N и $\lambda > 0$, принадлежащей ограниченному множеству. Таким образом, получено уточнение теоремы А. Ф. Колчина.

Теорема 1. Пусть $\lambda = (b_1/b_0)N\theta$. Предположим, что $q = 1$, множество \mathbf{N}' конечно и $\max\{m : m \in \mathbf{N}'\} \geq 2$. Тогда

$$\begin{aligned} \mathbf{P}(A) &= \exp \left(-\frac{b_2\lambda^2}{2b_0} \left(\frac{b_0}{b_1}\right)^2 \times \right. \\ &\times \left. \left(\frac{\sum_{k \in \mathbf{N}', k \geq 2} \pi_\lambda(k-2)}{\sum_{k \in \mathbf{N}'} \pi_\lambda(k)} \right) \alpha \left(1 + \frac{O(1)}{N}\right) \right), \quad (16) \end{aligned}$$

где $O(1)$ ограничена по N и $\lambda > 0$, принадлежащей ограниченному множеству.

Доказательство. Используя (1) и лемму 1, получаем

$$\begin{aligned} \mathbf{P}(A) &= (\mathbf{P}(A_i))^n = \left(\frac{\mathbf{P}\{\xi_j \leq 1, 1 \leq j \leq N, \xi_1 + \xi_2 + \xi_3 + \cdots + \xi_N = k \text{ для некоторого } k \in \mathbf{N}'\}}{\mathbf{P}\{\xi_1 + \xi_2 + \xi_3 + \cdots + \xi_N = k \text{ для некоторого } k \in \mathbf{N}'\}} \right)^n = \\ &= \left(\frac{\sum_{k \in \mathbf{N}'} \mathbf{P}\{\xi_j \leq 1, 1 \leq j \leq N, \xi_2 + \xi_3 + \cdots + \xi_N = k\}}{\sum_{k \in \mathbf{N}'} \mathbf{P}\{\xi_1 + \xi_2 + \xi_3 + \cdots + \xi_N = k\}} \right)^n = \\ &= \left(\sum_{k \in \mathbf{N}'} B(k, N) / \left(\sum_{k \in \mathbf{N}'} B(k, N) + \frac{1}{N} \frac{b_2\lambda^2}{2B(\theta)} \left(\frac{b_0}{b_1}\right)^2 \left(\sum_{k \in \mathbf{N}', k \geq 2} B(k-2, N-1) \right) \left(1 + \frac{O'(1)}{N}\right) \right) \right)^n = \\ &= \left(1 / \left(1 + \frac{1}{N} \frac{b_2\lambda^2}{2B(\theta)} \left(\frac{b_0}{b_1}\right)^2 \frac{\sum_{k \in \mathbf{N}', k \geq 2} B(k-2, N-1)}{\sum_{k \in \mathbf{N}'} B(k, N)} \left(1 + \frac{O'(1)}{N}\right) \right) \right)^{N\alpha} = \\ &= \exp \left(-\frac{b_2\lambda^2}{2b_0} \left(\frac{b_0}{b_1}\right)^2 \left(\frac{\sum_{k \in \mathbf{N}', k \geq 2} \pi_\lambda(k-2)}{\sum_{k \in \mathbf{N}'} \pi_\lambda(k)} \right) \alpha \left(1 + \frac{O(1)}{N}\right) \right). \end{aligned}$$

Доказательство закончено.

Следствие 1. Пусть $\lambda = (b_1/b_0)N\theta$. Предположим, что $q = 1$, множество \mathbf{N}' конечно и $\max\{m : m \in \mathbf{N}'\} \geq 2$. Пусть $\lambda \rightarrow \lambda'$, $\alpha \rightarrow \alpha_0$ при $n, N \rightarrow \infty$. Тогда

$$\begin{aligned} \mathbf{P}(A) &\rightarrow \\ &\rightarrow \exp \left(-\frac{b_2\lambda'^2}{2b_0} \left(\frac{b_0}{b_1}\right)^2 \left(\frac{\sum_{k \in \mathbf{N}', k \geq 2} \pi_{\lambda'}(k-2)}{\sum_{k \in \mathbf{N}'} \pi_{\lambda'}(k)} \right) \alpha_0 \right). \end{aligned}$$

Доказательство. Условие $\lambda \rightarrow \lambda'$, $\alpha \rightarrow \alpha_0$ при $n, N \rightarrow \infty$, влечет ограниченность семейств $\lambda = \lambda(N)$ и $\alpha = n/N$. Поэтому применима теорема 1.

Пусть $F_\lambda(x) = \sum_{i \leq x, i \in \mathbf{N}} \pi_\lambda(i)$, $x \in \mathbf{R}$, — функция распределения пуассоновской случайной величины с параметром λ .

Следствие 2. Пусть $\lambda = (b_1/b_0)N\theta$. Предположим, что $q = 1$, множество $\mathbf{N}' = \{1, 2, \dots, m\}$, где $2 \leq m < \infty$. Пусть $\lambda \rightarrow \lambda'$, $\alpha \rightarrow \alpha_0$ при $n, N \rightarrow \infty$. Тогда

$$\mathbf{P}(A) \rightarrow \exp \left(-\frac{b_2\lambda'^2}{2b_0} \left(\frac{b_0}{b_1}\right)^2 \frac{F_{\lambda'}(m-2)}{F_{\lambda'}(m)} \alpha_0 \right).$$

Доказательство. Так как множество $\mathbf{N}' = \{1, 2, \dots, m\}$ конечно, применимо следствие 1.

Из следствия 1 вытекает

Следствие 3. Пусть $\lambda = (b_1/b_0)N\theta$. Предположим, что $q = 1$, множество $\mathbf{N}' = \{m\}$, где $2 \leq m < \infty$. Пусть $\lambda \rightarrow \lambda'$, $\alpha \rightarrow \alpha_0$ при $n, N \rightarrow \infty$. Тогда

$$\mathbf{P}(A) \rightarrow \exp\left(-\frac{b_2}{2b_0} \left(\frac{b_0}{b_1}\right)^2 \frac{m(m-1)}{2} \alpha_0\right).$$

Следующая теорема показывает, что случай $q > 1$ существенно отличается от случая $q = 1$.

Теорема 2. Пусть $\lambda = (b_1/b_0)N\theta$. Предположим, что множество \mathbf{N}' конечно и $m = \max\{i : i \in \mathbf{N}'\} > q > 1$. Тогда

$$1 - \mathbf{P}(A) = C \frac{\alpha}{N^{q-1}} \left(1 + \frac{O(1)}{N}\right), \quad (17)$$

где

$$C = \left(\frac{b_{q+1}}{(q+1)!b_0} \left(\frac{b_0\lambda}{b_1}\right)^{q+1}\right) / \sum_{k \in \mathbf{N}'} \pi_\lambda(k)$$

и $O(1)$ ограничена по N и $\lambda > 0$, принадлежащей ограниченному множеству. В частности, если $\alpha = o(N^{q-1})$ при $n, N \rightarrow \infty$, то

$$\mathbf{P}(A) \rightarrow 1 \text{ при } N \rightarrow \infty \quad (18)$$

равномерно относительно $\lambda > 0$, принадлежащей ограниченному множеству.

Доказательство. Пусть событие

$$A_{ij} = \{\xi_{ij} \leq q \mid \xi_{i1} + \dots + \xi_{iN} = m \text{ для некоторого } m \in \mathbf{N}'\},$$

состоит в том, что в j -м блоке i -го сообщения имеется не больше q ошибок, A_{ij}^c — его дополнение. Используя (2), (3) и замечание 1, получаем

$$\begin{aligned} \mathbf{P}(A_{ij}^c) &= \frac{\mathbf{P}\{\xi_1 \geq q+1, \xi_2 + \xi_3 + \dots + \xi_N = k \text{ для некоторого } k \in \mathbf{N}'\}}{\mathbf{P}\{\xi_1 + \xi_2 + \xi_3 + \dots + \xi_N = k \text{ для некоторого } k \in \mathbf{N}'\}} \leq \\ &\leq \frac{\mathbf{P}\{\xi \geq q+1\}}{\sum_{k \in \mathbf{N}'} \mathbf{P}\{\xi_1 + \xi_2 + \xi_3 + \dots + \xi_N = k\}} \leq \\ &\leq \left(\frac{b_{q+1}}{(q+1)!b_0} \left(\frac{b_0\lambda}{b_1}\right)^{q+1} \frac{1}{N^{q+1}} \left(1 + \frac{O'(1)}{N}\right)\right) / \left(\left(\sum_{k \in \mathbf{N}'} \pi_\lambda(k)\right) \left(1 + \frac{O''(1)}{N}\right)\right) = \\ &= \frac{1}{N^{q+1}b_0} \left(\frac{b_{q+1}}{(q+1)!} \left(\frac{b_0\lambda}{b_1}\right)^{q+1} / \sum_{k \in \mathbf{N}'} \pi_\lambda(k)\right) \left(1 + \frac{O(1)}{N}\right) = \frac{C}{N^{q+1}} \left(1 + \frac{O(1)}{N}\right). \end{aligned}$$

Поэтому

$$1 - \mathbf{P}(A) = \mathbf{P}\left(\bigcup_{i=1}^n \bigcup_{j=1}^N A_{ij}^c\right) \leq \sum_{i=1}^n \sum_{j=1}^N \mathbf{P}(A_{ij}^c) \leq nN \frac{C}{N^{q+1}} (1 + o(1)) = C \frac{\alpha}{N^{q-1}} \left(1 + \frac{O(1)}{N}\right).$$

Итак, утверждение (17) доказано.

Так как $C(\alpha/N^{q-1})(1 + o(1)) \rightarrow 0$ при $N \rightarrow \infty$, (17) влечет (18). Теорема доказана.

Теперь рассмотрим событие A без условия $\xi_1 + \xi_2 + \dots + \xi_N = k$ для некоторого $k \in \mathbf{N}'$. Это соответствует случаю $\mathbf{N} = \mathbf{N}'$. Для него получим выражение для вероятности $\mathbf{P}(A)$ такое, что аналогии (16) и (17) будут его следствиями.

Теорема 3. Пусть $\lambda = (b_1/b_0)N\theta$. Предположим, что $q \geq 1$, множество $\mathbf{N} = \mathbf{N}'$. Тогда

$$\mathbf{P}(A) = \exp\left(-\frac{1}{N^{q-1}} \frac{b_{q+1}}{(q+1)!b_0} \left(\frac{b_0}{b_1}\right)^{q+1} \times \lambda^{q+1} \alpha \left(1 + \frac{O(1)}{N}\right)\right), \quad (19)$$

где $O(1)$ ограничена по N и $\lambda > 0$, принадлежащей ограниченному множеству.

Доказательство. Имеем

$$\begin{aligned}
 \mathbf{P}(A) &= (\mathbf{P}\{\xi_i \leq q, 1 \leq i \leq N\})^n = (\mathbf{P}\{\xi_1 \leq q\})^{Nn} = \left(\frac{b_0}{B(\theta)} + \frac{b_0}{B(\theta)}\theta + \dots + \frac{b_0}{B(\theta)}\theta^q \right)^{Nn} = \\
 &= \left(\left(b_0 + \frac{b_1}{1!} \left(\frac{b_0}{b_1} \right) \frac{\lambda}{N} + \dots + \frac{b_q}{q!} \left(\frac{b_0}{b_1} \right)^q \frac{\lambda^q}{N^q} \right) / \left(b_0 + \frac{b_1}{1!} \left(\frac{b_0}{b_1} \right) \frac{\lambda}{N} + \dots \right. \right. \\
 &\quad \left. \left. \dots + \frac{b_q}{q!} \left(\frac{b_0}{b_1} \right)^q \frac{\lambda^q}{N^q} + \frac{b_{q+1}}{(q+1)!} \left(\frac{b_0}{b_1} \right)^{q+1} \frac{\lambda^{q+1}}{N^{q+1}} + \dots \right) \right)^{N^2\alpha} = \\
 &= \left(1 / \left(1 + \frac{b_{q+1}}{(q+1)!} \left(\frac{b_0}{b_1} \right)^{q+1} \frac{\lambda^{q+1}}{N^{q+1}} \left(b_0 + \frac{b_1}{1!} \left(\frac{b_0}{b_1} \right) \frac{\lambda}{N} + \dots + \frac{b_q}{q!} \left(\frac{b_0}{b_1} \right)^q \frac{\lambda^q}{N^q} \right)^{-1} \left(1 + \frac{O(1)}{N} \right) \right) \right)^{N^2\alpha} = \\
 &= \exp \left(-\frac{1}{N^{q-1}} \frac{b_{q+1}}{(q+1)!b_0} \left(\frac{b_0}{b_1} \right)^{q+1} \lambda^{q+1} \alpha \left(1 + \frac{O(1)}{N} \right) \right).
 \end{aligned}$$

Доказательство закончено.

Следствие 1. Пусть выполнены условия теоремы 3 и $q \geq 2$. Тогда

$$1 - \mathbf{P}(A) \leq \frac{1}{N^{q-1}} \frac{b_{q+1}}{(q+1)!b_0} \left(\frac{b_0}{b_1} \right)^{q+1} \times \lambda^{q+1} \alpha \left(1 + \frac{O(1)}{N} \right), \quad (20)$$

где $O(1)$ ограничена по N и $\lambda > 0$, принадлежащей ограниченному множеству.

Доказательство. Применяя к правой части (19) элементарное неравенство $1 - e^{-x} < x$, $x > 0$, получаем (20).

Следствие 2. Пусть выполнены условия теоремы 3 и $q = 1$. Тогда

$$\begin{aligned}
 \mathbf{P}(A) &= \\
 &= \exp \left(-\frac{b_{q+1}}{(q+1)!b_0} \left(\frac{b_0}{b_1} \right)^{q+1} \lambda^{q+1} \alpha \left(1 + \frac{O(1)}{N} \right) \right),
 \end{aligned}$$

где $O(1)$ ограничена по N и $\lambda > 0$, принадлежащей ограниченному множеству.

Следующее предложение верно без каких-либо условий на распределение случайной величины ξ .

Предложение 1. Пусть $q \geq 1$. Предположим, что $\mathbf{P}\{\xi_1 + \xi_2 + \xi_3 + \dots + \xi_N = k \text{ для некоторого } k \in \mathbf{N}'\} \geq \beta > 0$ и $\mathbf{p}(\xi < q) = \delta$. Тогда

$$\mathbf{P}(A) \leq \frac{\delta^{nN}}{\beta^n}.$$

Доказательство. Имеем

$$\mathbf{P}(A) = (\mathbf{P}(A_i))^n \leq \left(\frac{\mathbf{P}\{\xi_1 \leq q, \xi_2 \leq q, \dots, \xi_N \leq q\}}{\mathbf{P}\{\xi_1 + \xi_2 + \xi_3 + \dots + \xi_N = k \text{ для некоторого } k \in \mathbf{N}'\}} \right)^n \leq \frac{\delta^{nN}}{\beta^n}.$$

Предложение доказано.

3 Асимптотическое поведение вероятности $\mathbf{P}(A)$ в случае пуассоновской случайной величины ξ

В этой части будем предполагать, что ξ_i — независимые случайные величины, распределенные по закону Пуассона с параметром λ .

Теорема 4. Пусть $q = 1$. Предположим, что \mathbf{N}' — конечное множество и $m = \sup\{i, i \in \mathbf{N}'\} \geq 2$. Тогда

$$\mathbf{P}(A) = e^{(m(m-1)/2)\alpha(1+O(1)/N)}$$

и величина $O(1)$ ограничена, если λ принадлежит ограниченному множеству.

Доказательство. Имеем

$$\begin{aligned}
 \mathbf{P}(A) &= (\mathbf{P}(A_i))^n = \left(\frac{\sum_{k \in \mathbf{N}'} \mathbf{P}\{\xi_j \leq 1, 1 \leq j \leq N, \xi_1 + \xi_2 + \xi_3 + \dots + \xi_N = k\}}{\sum_{k \in \mathbf{N}'} \mathbf{P}\{\xi_1 + \xi_2 + \xi_3 + \dots + \xi_N = k\}} \right)^n = \\
 &= \left(\left(\sum_{k \in \mathbf{N}'} \frac{N(N-1) \dots (N-k+1)}{k!} (e^{-\lambda})^k (e^{-\lambda})^{N-k} \right) / \sum_{k \in \mathbf{N}'} e^{-N\lambda} \frac{(N\lambda)^k}{k!} \right)^n = \\
 &= \left(\left(\sum_{k \in \mathbf{N}'} \frac{N(N-1) \dots (N-k+1)}{k!} \right) \lambda^k / \sum_{k \in \mathbf{N}'} \frac{(N\lambda)^k}{k!} \right)^{N\alpha} = \\
 &= \left(\left(N^m \lambda^m m! \sum_{k \in \mathbf{N}'} \frac{N(N-1) \dots (N-k+1)m!}{k! N^m} \lambda^{k-m} \right) / \left(N^m \lambda^m m! \sum_{k \in \mathbf{N}'} \frac{(N\lambda)^{k-m} m!}{k!} \right) \right)^{N\alpha} = \\
 &= \left(\frac{N(N-1) \dots (N-m+1)m!}{m! N^m} + \sum_{k \in \mathbf{N}', k < m} \frac{N(N-1) \dots (N-k+1)m!}{k! N^m} \lambda^{k-m} \right)^{N\alpha} / \left(1 + \sum_{k \in \mathbf{N}', k < m} \frac{N^{k-m} \lambda^{k-m} m!}{k!} \right)^{N\alpha}.
 \end{aligned}$$

Пусть $m-1 \notin \mathbf{N}'$. Так как

$$\frac{N(N-1) \dots (N-m+1)m!}{m! N^m} = \left(1 - \frac{1}{N}\right) \left(1 - \frac{2}{N}\right) \dots \left(1 - \frac{m-1}{N}\right) = 1 - \frac{m(m-1)}{2N} + \frac{O(1)}{N^2}, \quad (21)$$

то

$$\mathbf{P}(A) = \left(1 - \frac{m(m-1)}{2N} + \frac{O(1)}{N^2}\right)^{N\alpha} / \left(1 + \frac{O(1)}{N^2}\right)^{N\alpha} = \exp\left(-\frac{m(m-1)}{2} \alpha \left(1 + \frac{O(1)}{N}\right)\right). \quad (22)$$

Это соответствует теореме.

Пусть $m-1 \in \mathbf{N}'$. Используя (21), получаем

$$\begin{aligned}
 \mathbf{P}(A) &= \left(\frac{N(N-1) \dots (N-m+1)m!}{m! N^m} + \frac{N(N-1) \dots (N-m+2)m!}{(m-1)! N^m} \lambda^{-1} + \right. \\
 &+ \left. \sum_{k \in \mathbf{N}', k < m-1} \frac{N(N-1) \dots (N-k+1)m!}{k! N^m} \lambda^{k-m} \right)^{N\alpha} / \left(1 + \frac{N^{-1} \lambda^{-1} m!}{(m-1)!} + \sum_{k \in \mathbf{N}', k < m-1} \frac{N^{k-m} \lambda^{k-m} m!}{k!} \right)^{N\alpha} = \\
 &= \left(\left(1 - \frac{m(m-1)}{2N} + \frac{\lambda^{-1} m}{N} + \frac{O(1)}{N^2}\right) / \left(1 + \frac{\lambda^{-1} m}{N} + \frac{O(1)}{N^2}\right) \right)^{N\alpha} = \\
 &= \exp\left(-\frac{m(m-1)}{2} \alpha \left(1 + \frac{O(1)}{N}\right)\right). \quad (23)
 \end{aligned}$$

Из (22) и (23) вытекает теорема 5. Доказательство закончено.

Теорема 5. *Предположим, что \mathbf{N}' — конечное множество и $m = \sup\{i, i \in \mathbf{N}'\} \geq 2$. Пусть $m > q > 1$. Тогда*

$$1 - \mathbf{P}(A) \leq \frac{C_m^{q+1} \alpha}{N^{q-1}} \left(1 + \frac{O(1)}{N}\right),$$

где $O(1)$ ограничена по N и $\lambda > 0$, принадлежащей ограниченному множеству.

Доказательство. Пусть $m > q$. Имеем

$$\begin{aligned}
 \mathbf{P}(A_{ij}^c) &= \frac{\mathbf{P}\{\xi_1 \geq q+1, \xi_1 + \xi_2 + \dots + \xi_N = k \text{ для некоторого } k \in \mathbf{N}'\}}{\mathbf{P}\{\xi_1 + \xi_2 + \xi_3 + \dots + \xi_N = k \text{ для некоторого } k \in \mathbf{N}'\}} = \\
 &= \frac{\sum_{k \in \mathbf{N}'} \mathbf{P}\{\xi_1 \geq q+1, \xi_1 + \xi_2 + \xi_3 + \dots + \xi_N = k\}}{\sum_{k \in \mathbf{N}'} \mathbf{P}\{\xi_1 + \xi_2 + \xi_3 + \dots + \xi_N = k\}} = \frac{\sum_{k \in \mathbf{N}', k > q} \sum_{i=q+1}^k \mathbf{P}\{\xi_1 = i, \xi_2 + \xi_3 + \dots + \xi_N = k-i\}}{\sum_{k \in \mathbf{N}'} \mathbf{P}\{\xi_1 + \xi_2 + \xi_3 + \dots + \xi_N = k\}} = \\
 &= \left(\sum_{k \in \mathbf{N}', k > q} \sum_{i=q+1}^k e^{-\lambda} \frac{\lambda^i}{i!} e^{-(N-1)\lambda} \frac{((N-1)\lambda)^{k-i}}{(k-i)!} \right) / \left(\sum_{k \in \mathbf{N}'} e^{-N\lambda} \frac{(N\lambda)^k}{k!} \right) \leq \\
 &\leq \left((N-1)^{m-(q+1)} \lambda^m \sum_{k \in \mathbf{N}', k > q} \lambda^{k-m} \sum_{i=q+1}^k \frac{1}{i!} \frac{(N-1)^{k-i-(m-(q+1))}}{(k-i)!} \right) / \left(N^m \lambda^m \sum_{k \in \mathbf{N}'} \frac{(N\lambda)^{k-m}}{k!} \right) = \\
 &= \frac{1}{N^{q+1}} \frac{m!}{(q+1)!(m-(q+1))!} \left(1 + \frac{O(1)}{N} \right) = \frac{1}{N^{q+1}} C_m^{q+1} \left(1 + \frac{O(1)}{N} \right).
 \end{aligned}$$

Следовательно,

$$1 - \mathbf{P}(A) = \mathbf{P} \left(\bigcup_{i=1}^n \bigcup_{j=1}^N A_{ij}^c \right) \leq \sum_{i=1}^n \sum_{j=1}^N \mathbf{P} \left(A_{ij}^c \right) \leq n \cdot N \cdot \frac{C_m^{q+1}}{N^{q+1}} \left(1 + \frac{O(1)}{N} \right) = \frac{\alpha}{N^{q-1}} C_m^{q+1} \left(1 + \frac{O(1)}{N} \right).$$

Доказательство закончено.

Заметим, что если $\mathbf{N}' = \{m\}$ — одноэлементное множество, то из доказательств теорем 4 и 5 следует, что оценки, полученные в них, не зависят от λ . Это естественно, так как событие A_i является событием схемы размещения m различных частиц по N ячейкам [3]. В случае одноэлементного множества \mathbf{N}' теорема 4 и теорема 5 получены в [7].

Литература

1. Новиков Ф. А. Дискретная математика для программистов. 2-е изд. — СПб.: Питер, 2004.
2. Колчин В. Ф. Один класс предельных теорем для условных распределений // Литовск. матем. сб., 1968. Т. 8(1). С. 53–63.
3. Колчин В. Ф. Случайные графы. — М.: Физматгиз, 2000.
4. Колчин А. В. Предельные теоремы для обобщенной схемы размещения // Дискрет. матем., 2003. Т. 15(4). С. 143–157.
5. Колчин А. В., Колчин В. Ф. О переходе распределений сумм независимых одинаково распределенных случайных величин с одной решетки на другую в обобщенной схеме размещения // Дискрет. матем., 2006. Т. 18(4). С. 113–127.
6. Колчин А. В., Колчин В. Ф. Переход с одной решетки на другую распределений сумм случайных величин, встречающихся в обобщенной схеме размещения // Дискрет. матем., 2007. Т. 19(3). С. 15–21.
7. Avkhadiev F. G., Chuprunov A. N. The probability of a successful allocation of ball groups by boxes // Lobachevskii J. of Math., 2007. Vol. 25. P. 3–5.

О РАСПРЕДЕЛЕНИИ РАЗМЕРОВ ЧАСТИЦ ПРИ ДРОБЛЕНИИ*

В. Ю. Королёв¹

Аннотация: Предложена новая модель для распределения размера дробящейся частицы, учитывающая непостоянство или случайный характер интенсивности потока соударений. В рамках этой модели сформулирован критерий логнормальности указанного распределения и описан класс возможных распределений размера частиц при дроблении. Наряду со многими известными моделями этот класс содержит масштабные смеси логнормальных законов.

Ключевые слова: логнормальное распределение; смеси нормальных законов; обобщенный процесс Кокса

1 Введение

В данной статье рассматриваются математические модели, описывающие распределение физических размеров частиц при дроблении. Подобные модели могут довольно успешно использоваться (и используются) при описании самых разных объектов от величины капиталов фирм или размера доходов до объемов сообщений в вычислительных или телекоммуникационных системах.

В опубликованной в 1940 г. работе Н. К. Разумовского указано много случаев, в которых логарифмы размеров частиц (золотин в золотоносных россыпях, частиц горных пород при их дроблении и т. п.) имеют примерно нормальное (гауссовское) распределение [1]. На эту работу обратил внимание А. Н. Колмогоров, который предложил математическую модель процесса дробления частиц, аналитически объясняющую возникновение логнормального распределения размеров частиц при дроблении, а также содержания минералов в отдельных пробах [2]. Результат Колмогорова справедлив при довольно сильных предположениях. В частности, в рамках модели Колмогорова для логнормальности распределения частиц при дроблении необходимо, чтобы скорость дробления (уменьшения размеров частиц) была постоянной, т. е. не зависела от размеров дробящихся частиц. В то же время практически очевидно, что с уменьшением размера частицы интенсивность ее соударений с другими частицами или деталями дробильного агрегата может изменяться, например уменьшаться в силу того, что вероятность столкновения с другими частицами, очевидно, в определенном смысле пропорциональна размерам частицы, и чем частица меньше, тем меньше и эта вероятность. На это обстоятельство обратил внимание сам

Колмогоров, который в конце своей статьи [2], упоминавшейся выше, написал: «Было бы интересно изучить математические схемы, в которых скорость дробления частиц уменьшается (или увеличивается) с уменьшением их размеров. Естественно рассмотреть при этом в первую очередь случаи, в которых скорость дробления пропорциональна той или иной степени размеров частицы. Если эта степень отлична от нуля, то, по-видимому, логарифмически нормальный закон будет уже неприменим». В данной работе сделана попытка исследования этого предположения А. Н. Колмогорова.

Предложена довольно естественная модель процесса дробления, учитывающая зависимость интенсивности процесса дробления от параметров распределения размера частицы (в частности, от его среднего значения). В рамках этой модели оказывается возможным сформулировать критерий (т. е. необходимые и достаточные условия) логнормальности распределения размера частиц при дроблении.

В некоторых работах приводятся эмпирические свидетельства того, что в некоторых случаях логнормальная модель для распределения размера частиц при дроблении неадекватна. Так, в 1941 г. Р. А. Багнольд в своей книге [3] заметил, что логарифм функции плотности распределения для логарифма размера частицы в естественных запасах песка больше похож на гиперболу, чем на параболу. Это означает, что, согласно наблюдениям Багнольда, распределение размера не логнормально, а скорее имеет экспоненциально уменьшающиеся хвосты.

Последнее обстоятельство побудило некоторых исследователей обратить внимание на модели типа лог-несимметричного распределения Лапласа, так называемого двойного (двустороннего) Парето-логнормального распределения, а также

* Работа выполнена при поддержке РФФИ, гранты 08-01-00345, 08-01-00363, 08-07-00152, 09-07-12032-офи-м.

¹ Факультет вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова; Институт проблем информатики РАН, vkorolev@comtv.ru

лог-гауссовского//обратного гауссовского распределения. В частности, в работах [4, 5] рассмотрен следующий механизм формирования распределения размера дробящихся частиц.

Базовая посылка этих моделей заключается в том, что частица при перемещении из одного места в другое может разделиться на несколько меньших частиц вследствие соударения или другого воздействия, что обуславливает случайность масс частиц после разделения. Обозначим число разбиений (дроблений) изначально одной частицы к моменту времени t за $N(t)$. Первоначальный размер частицы обозначим s_0 . Пусть D_i — часть (доля) частицы, отделившаяся при i -м соударении. Тогда размер частицы в момент времени t имеет вид

$$Z(t) = s_0 \prod_{i=1}^{N(t)} (1 - D_i). \quad (1)$$

Следовательно,

$$S(t) \equiv \ln Z(t) = \mu + \sum_{i=1}^{N(t)} X_i,$$

где $\mu = \ln(s_0)$ и $X_i = \ln(1 - D_i)$. Считается, что X_i , $i \geq 1$, — независимые и одинаково распределенные случайные величины с математическими ожиданиями a , дисперсиями $\sigma^2 \in (0, \infty)$ и, соответственно, вторыми моментами $b_2 = a^2 + \sigma^2$. Предполагается, что процесс $N(t)$ — пуассоновский с некоторым параметром $\lambda > 0$ и стохастически независим от последовательности X_i . Это означает, что число соударений, вследствие которых частица разбивается, на временном участке длиной s имеет распределение Пуассона с математическим ожиданием λs , а числа соударений на непересекающихся временных интервалах являются независимыми случайными величинами. В данных предположениях вследствие центральной предельной теоремы при достаточно большом λ распределение случайной величины

$$U_\lambda(t) = \frac{1}{\sqrt{\lambda}} \left(\sum_{i=1}^{N(t)} X_i - \lambda t a \right)$$

является приблизительно нормальным с нулевым математическим ожиданием и дисперсией tb_2 . В результате

$$P(S(t) < x) \approx \Phi \left(\frac{x - \mu - \lambda t a}{\sqrt{\lambda t b_2}} \right) \quad (2)$$

для больших λ , где $\Phi(x)$ — стандартная нормальная функция распределения, $x \in \mathbb{R}$. Не все частицы одновременно добираются до места назначения.

Некоторые могут поменять направление и застрять надолго, тогда как другие могут, не задев других, пройти весь путь намного быстрее. Очевидно, состояние частицы зависит от того, на каком расстоянии от конечного пункта она находится. Это учитывается в моделях Рида–Йоргенсена и Соренсена за счет случайности времени, в которое наблюдается частица. По сути, Рид, Йоргенсен и Соренсен ввели рандомизацию в модель, предложенную Колмогоровым в 1940 г., добавив в нее учет того обстоятельства, что частицы затрачивают разное время на преодоление пути от одного источника до другого.

В модели Рида–Йоргенсена время считается показательно распределенной случайной величиной, т. е. итоговое безусловное распределение логарифма размера частицы представляет собой смесь распределения (2) по параметру t , имеющему показательное распределение. Эта смесь как раз и является двусторонним распределением Парето (в работе [4] к величине $S(t)$ дополнительно прибавляется независимая от нее нормально распределенная случайная величина, характеризующая «случайный выбор» начального состояния частицы, так что итоговое распределение оказывается так называемым двусторонним Парето-логнормальным законом).

В модели Соренсена движение частицы считается броуновским со сносом ν и коэффициентом диффузии ω^2 . Это означает, что дистанция, пройденная за временной промежуток длины s , нормально распределена с математическим ожиданием νs и дисперсией $\omega^2 s$, а расстояния, пройденные за непересекающиеся интервалы времени, являются независимыми случайными величинами. Предполагается, что броуновское движение частиц независимо от процессов дробления (разделения), т. е. от пуассоновского процесса $N(t)$ и случайных величин b_i . Модель одномерна, т. е. рассматривается только проекция частицы на прямую. Если будет много движений в других направлениях, это вызовет больший износ частицы и может быть учтено за счет увеличения интенсивности λ пуассоновского процесса разделения частиц. Обозначим через τ (случайное) время прибытия определенной частицы к месту назначения. Для рассматриваемого процесса броуновского движения, как известно, распределение случайной величины τ является обратным гауссовским $IG(a/\omega, \nu/\omega)$, где a означает расстояние между точкой отправления и конечной точкой. Плотность обратного гауссовского распределения $IG(\delta, \gamma)$ равна

$$\frac{\delta}{\sqrt{2\pi x^3}} \exp \left\{ -\frac{\gamma^2}{2x} \left(x - \frac{\delta}{\gamma^2} \right) \right\}, \quad x > 0.$$

Логарифм размера частицы, прибывшей в пункт назначения в момент $\tau = t$, очевидно, равен $S(t)$. Следовательно, условное распределение логарифма размера частицы, прибывшей в момент $\tau = t$, приблизительно нормально:

$$P(S(t) < x) \approx \Phi \left(\frac{x - \mu - \beta \lambda b_2 t}{\sqrt{\lambda b_2 t}} \right),$$

где

$$\beta = \frac{a}{b_2}.$$

Распределение случайной величины $\zeta = \lambda b_2 \tau$ является обратным гауссовским:

$$\zeta \sim IG \left(\frac{a\sqrt{\lambda b_2}}{\omega}, \frac{\nu}{\omega\sqrt{\lambda b_2}} \right).$$

Безусловное же распределение логарифма размера частицы, т. е. безусловное распределение случайной величины $S(\tau) = \ln Z(\tau)$, является приблизительно смесью приведенных выше нормальных распределений по распределению параметра ζ . Эта смесь по определению является гауссовским//обратным гауссовским распределением $GIG(\alpha, \beta, \delta, \mu)$, которому соответствует характеристическая функция

$$f_{\alpha, \beta, \delta, \mu}(s) = \exp \left\{ \delta \sqrt{\alpha^2 - \beta^2} \left[1 - \sqrt{1 + \frac{s^2 - is2\beta}{\alpha^2 - \beta^2}} \right] + is\mu \right\}.$$

В рассматриваемом случае параметры α , β и δ соответственно равны

$$\begin{aligned} \alpha &= \sqrt{\frac{a^2}{b_2^2} + \frac{\nu^2}{\omega^2 \lambda b_2}}; \\ \beta &= \frac{a}{b_2}; \\ \delta &= \frac{a}{\omega} \sqrt{\lambda b_2}. \end{aligned}$$

Распределения, логарифм функции плотности которых представляет собой гиперболу (гиперболические распределения), введены Барндорфф-Нильсеном в 1977 г. [6]. В ряде работ гиперболические распределения дали хорошие результаты при их подгонке к эмпирическим распределениям реальных данных (см., например, [7–9]). В работе [6] фактически введен более общий класс распределений вероятностей — класс так называемых обобщенных гиперболических распределений, которые

во многом подобны гиперболическим распределениям. Одним из представителей этого класса является гауссовское//обратное гауссовское распределение, которое демонстрирует еще лучшее согласие с эмпирическими распределениями размера частиц, нежели гиперболические распределения. Это обусловлено тем, что у GIG -распределений хвосты являются более тяжелыми по сравнению с обычным гиперболическим распределением, что характерно для естественных запасов песка. Другое преимущество GIG -распределений в том, что они имеют много благоприятных математических свойств. Например, их моменты могут быть вычислены явно.

Функция плотности распределения $GIG(\alpha, \beta, \delta, \mu)$ равна

$$\frac{\alpha \delta}{\pi} e^{\delta \gamma} \cdot \frac{K_1(\alpha \sqrt{\delta^2 + (x - \mu)^2})}{\sqrt{\delta^2 + (x - \mu)^2}} \cdot e^{\beta(x - \mu)}, \quad x \in \mathbb{R}.$$

Здесь K_1 — модифицированная функция Бесселя третьего порядка. Возможные значения параметров: $\alpha > 0$, $\delta > 0$, $\beta < |\alpha|$, $\mu \in \mathbb{R}$. Если случайная величина X имеет $GIG(\alpha, \beta, \delta, \mu)$ -распределение, то ее математическое ожидание и дисперсия соответственно равны:

$$EX = \mu + \frac{\delta \beta}{\gamma}, \quad DX = \frac{\delta \alpha^2}{\gamma^3},$$

где $\gamma = \sqrt{\alpha^2 - \beta^2}$. Коэффициент асимметрии равен $3\beta/(\alpha\sqrt{\delta\gamma})$, а коэффициент эксцесса равен $3(1 + 4\beta^2/\alpha^2)/(\delta\gamma)$.

Модели Рида–Йоргенсена и Барндорфф–Нильсена–Соренсена демонстрируют хорошее согласие с экспериментальными данными, отражающими размеры частиц в природных залежах или размеры капиталов фирм. Однако в ее конструкции, тем не менее, не учитывается отмеченное Колмогоровым обстоятельство: интенсивность процесса дробления остается постоянной.

В следующем разделе данной работы будет предложена модель, в которой интенсивность процесса дробления может быть переменной и даже случайной.

2 Модель процесса дробления частиц, основанная на обобщенных процессах Кокса

2.1 Базовая модель

Базой для рассматриваемых конструкций служит структурная модель [1]. Основное отличие предлагаемой модели от моделей, рассмотренных

выше, в том, что интенсивность соударений будет считаться случайной.

Пусть в базовом представлении (2) X_1, X_2, \dots — одинаково распределенные случайные величины, $s_0 = 1$ (так что $\mu = 0$), а $N(t)$ — дважды стохастический пуассоновский процесс (процесс Кокса), управляемый процессом $\Lambda(t)$, т. е.

$$N(t) = N_1(\Lambda(t)), \quad t \geq 0,$$

где $\Lambda(t)$ — случайный процесс с неубывающими, непрерывными справа траекториями, выходящими из нуля, $N_1(t)$ — стандартный пуассоновский процесс (однородный пуассоновский процесс с единичной интенсивностью), причем процессы $N_1(t)$ и $\Lambda(t)$ независимы. Такая модель процесса, подсчитывающего число соударений (дроблений) частицы к моменту t , учитывает возможную изменчивость и даже случайность интенсивности соударений дробящейся частицы, которая на практике, например при дроблении частиц грунта речного русла, может быть вызвана периодической сменной времен года, погодой и другими причинами. Во многих приложениях можно считать, что процесс $\Lambda(t)$ имеет вид

$$\Lambda(t) = \int_0^t \lambda(\tau) d\tau, \quad t \geq 0,$$

где $\lambda(t)$ — положительный случайный процесс с интегрируемыми траекториями. При этом $\lambda(t)$ можно интерпретировать как мгновенную стохастическую интенсивность процесса $N(t)$. Поэтому иногда процесс $\Lambda(t)$, управляющий процессом Кокса $N(t)$, называют *накопленной интенсивностью* процесса $N(t)$.

Предположим, что при каждом $t \geq 0$ случайные величины $N(t), X_1, X_2, \dots$ независимы. Процесс

$$S(t) = \sum_{j=1}^{N(t)} X_j, \quad t \geq 0, \quad (3)$$

назовем обобщенным процессом Кокса (при этом для определенности будем считать, что $\sum_{j=1}^0 = 0$).

Процессы вида (3) играют чрезвычайно важную роль во многих прикладных задачах, будучи естественными моделями неоднородных хаотических потоков (см., например, [10, 11]). Достаточно сказать, что при $\Lambda(t) \equiv \ell t$ с $\ell > 0$ процесс $S(t)$ превращается в классический обобщенный пуассоновский процесс, широко используемый при моделировании многих явлений в физике, теории надежности, финансовой и актуарной деятельности, биологии и т. д., где возникают потоки событий, абсолютно хаотично, беспорядочно распределенных

во времени, или размещения точек, абсолютно хаотично, беспорядочно распределенных в пространстве. Большое число разнообразных прикладных задач, приводящих к обобщенным пуассоновским процессам, описано в книгах [10, 12]. Обобщенные процессы Кокса играют важную роль при моделировании характеристик неоднородных хаотических стохастических потоков случайных событий. При этом особую важность при использовании обобщенных процессов Кокса вида (3), скажем в страховой математике при моделировании неоднородных потоков страховых выплат или в теории управления запасами при моделировании неоднородных потоков заявок на поставку некоторого продукта, имеет тот случай, когда математическое ожидание слагаемых X_j в сумме (3) отлично от нуля (в частности, в условиях приведенных примеров — положительно). В рассматриваемом в данной работе случае по смыслу задачи $0 \leq D_i \leq 1$, так что $a = EX_1 = E \ln(1 - D_i) \leq 0$. Для определенности всюду далее считаем, что $a < 0$. Обозначим $DX_1 = \sigma^2$ и будем считать, что $0 < \sigma^2 < \infty$.

Символ \implies будет обозначать сходимость по распределению, которая, как известно, в конечномерном случае эквивалентна слабой сходимости.

Расстояние (метрика) Леви $\mathcal{L}(F_1, F_2)$ между функциями распределения F_1 и F_2 определяется как

$$\mathcal{L}(F_1, F_2) = \inf\{h > 0 : F_1(x - h) - h \leq F_2(x) \leq F_1(x + h) + h \text{ для всех } x \in \mathbb{R}\}.$$

Если X_1 и X_2 — случайные величины с функциями распределения F_1 и F_2 соответственно, то будем считать, что $\mathcal{L}(X_1, X_2) = \mathcal{L}(F_1, F_2)$. Сходимость в метрике Леви эквивалентна сходимости по распределению (см., например, [13, гл. 2, разд. 9]).

2.2 Предельные теоремы. Обоснование возможности аппроксимации распределения логарифма размера частиц при дроблении сдвигowymi смесями нормальных законов

Теорема 1. Пусть $S(t)$ — логарифм текущего размера дробящейся частицы в момент времени t — представляет собой обобщенный процесс Кокса, управляемый процессом накопленной интенсивности $\Lambda(t)$ и порожденный последовательностью $\{X_i\}_{i \geq 1}$. Предположим, что $EX_1 = a < 0$ и $\Lambda(t) \rightarrow \infty$ по вероятности при $t \rightarrow \infty$. Пусть $D(t) > 0$ — некоторая функция такая, что $D(t) \rightarrow \infty$ при $t \rightarrow \infty$. Тогда одномерные распределения неслучайно центрированного

и нормированного обобщенного процесса Кокса $S(t)$ и слабо сходятся при $t \rightarrow \infty$ к распределениям некоторой случайной величины Z , т. е.

$$\frac{S(t) - C(t)}{D(t)} \Rightarrow Z \quad (t \rightarrow \infty)$$

при некоторой вещественной функции $C(t)$ тогда и только тогда, когда

$$\limsup_{t \rightarrow \infty} \frac{|C(t)|}{D^2(t)} \equiv k^2 < \infty$$

и существует случайная величина V такая, что

$$Z \stackrel{d}{=} k \sqrt{\frac{a^2 + \sigma^2}{|a|}} \cdot W + V, \quad (4)$$

где W — случайная величина со стандартным нормальным распределением, независимая от случайной величины V , и

$$\mathcal{L} \left(\frac{a\Lambda(t) - C(t)}{D(t)}, V(t) \right) \rightarrow 0 \quad (t \rightarrow \infty),$$

где распределение случайной величины $V(t)$ определяется ее характеристической функцией

$$\begin{aligned} \mathbb{E} \exp\{isV(t)\} = \exp \left\{ -\frac{s^2(a^2 + \sigma^2)}{2|a|} \left[k^2 - \frac{|C(t)|}{D^2(t)} \right] \right\} \mathbb{E} \exp\{isV\}, \quad s \in \mathbb{R}. \quad (5) \end{aligned}$$

Доказательство этого результата, основанное на общем утверждении о сходимости распределений суперпозиций независимых случайных процессов, доказанном в [14], можно найти, например, в книге [10] (см. теорему 9.2.1 там).

Из теоремы 1 вытекает следующий критерий логнормальности распределения размера частицы при дроблении.

Следствие 1. В условиях теоремы 1 распределение логарифма текущего размера дробящейся частицы $S(t)$ асимптотически нормально с некоторой асимптотической дисперсией $\delta^2 > 0$:

$$\mathbb{P} \left(\frac{S(t) - C(t)}{D(t)} < x \right) \Rightarrow \Phi \left(\frac{x}{\delta} \right) \quad (t \rightarrow \infty)$$

в том и только в том случае, когда

$$\limsup_{t \rightarrow \infty} \frac{|C(t)|}{D^2(t)} \leq \frac{|a|\delta^2}{a^2 + \sigma^2}$$

$$\begin{aligned} \lim_{t \rightarrow \infty} \mathcal{L} \left(\mathbb{P} \left(\frac{a\Lambda(t) - C(t)}{D(t)} < x \right), \right. \\ \left. \Phi \left(\frac{\sqrt{|a|}D(t)x}{\sqrt{|a|\delta^2 D^2(t) - (a^2 + \sigma^2)|C(t)|}} \right) \right) = 0. \end{aligned}$$

Доказательство. Это утверждение вытекает из теоремы 1 и теоремы Леви–Крамера о разложимости нормального закона только на нормальные компоненты, согласно которой любая случайная величина V , удовлетворяющая (4), должна быть нормально распределенной с нулевым средним и дисперсией $\delta^2 - k^2(a^2 + \sigma^2)/|a|$. Отсюда вытекает, что любая случайная величина $V(t)$, удовлетворяющая (5), неизбежно должна иметь нормальное распределение с нулевым средним и дисперсией $\delta^2 - (a^2 + \sigma^2)|C(t)|/(|a|D^2(t))$.

Другими словами, в условиях теоремы 1 распределение текущего размера дробящейся частицы $Z(t)$ асимптотически логнормально тогда и только тогда, когда процесс накопленной интенсивности $\Lambda(t)$ имеет асимптотически нормальные (в частности, вырожденные) одномерные распределения.

Из теоремы 1 вытекает, что в рамках рассматриваемой модели распределение логарифма дробящейся частицы разумно аппроксимировать сдвиговыми смесями нормальных законов. При этом вид смешивающего распределения полностью определяется асимптотическим поведением накопленной интенсивности процесса соударений. Другими словами, если $Z(t) = \exp\{S(t)\}$ — размер дробящейся частицы, то

$$S(t) \approx X' + V',$$

где X' — случайная величина, имеющая нормальное распределение с нулевым средним и некоторой дисперсией σ_*^2 , а V' — независимая от X' случайная величина с, вообще говоря, произвольным распределением, определяемым статистическими закономерностями в поведении интенсивности столкновений. При этом приближенное равенство понимается как близость распределений левой и правой частей. Иными словами,

$$Z(t) \approx \xi \eta,$$

где $\xi = e^{X'}$ — случайная величина с логнормальным распределением, а $\eta = e^{V'}$ — независимая от ξ положительная случайная величина. Таким образом, в рамках рассматриваемой модели распределение размера частиц при дроблении является масштабной смесью логнормальных распределений.

Необходимо отметить, что в моделях Рида–Йоргенсена и Соренсена чисто сдвиговые смеси нормальных законов не могут возникнуть в качестве модели для распределения логарифма размера частиц (и, следовательно, чисто масштабные смеси логнормальных законов не могут возникнуть в качестве модели для распределения самого размера частиц).

На практике при определении распределения логарифма размера дробящейся частицы разумно использовать дискретную аппроксимацию для смешивающего закона, что позволит искать оценки параметров смеси с помощью статистических процедур, традиционно используемых для этих целей, например с использованием EM-алгоритма или его модификаций.

Если в теореме 1 задать конкретный вид функций $C(t)$ и $D(t)$, то ситуация намного упростится. Предположим, что существуют такие числа $\ell \in (0, \infty)$ и $s \in (0, \infty)$, что

$$E\Lambda(t) \equiv \ell t, \quad D\Lambda(t) \equiv s^2 t, \quad t > 0.$$

Тогда, как несложно убедиться,

$$ES(t) = alt, \quad DS(t) = [\ell(a^2 + \sigma^2) + a^2 s^2] t.$$

Линейная зависимость $ES(t)$ и $DS(t)$ от t обусловлена тем, что в рамках указанного предположения о виде моментов управляющего процесса согласно теореме 1 с $C(t) = ES(t)$ и $D(t) = \sqrt{DS(t)}$ только при совпадении показателей степеней возможно существование собственного невырожденного предельного распределения. С другой стороны, если процесс $\Lambda(t)$ имеет интегральное представление, то равенство показателя степени t единице оказывается вполне естественным.

Из теоремы 1 при этом непосредственно вытекает следующий результат.

Теорема 2. *Одномерные распределения неслучайно центрированных и нормированных обобщенных процессов Кокса слабо сходятся к распределению некоторой случайной величины Z при $t \rightarrow \infty$, т. е.*

$$\frac{S(t) - alt}{\sqrt{t[\ell(a^2 + \sigma^2) + a^2 s^2]}} \Rightarrow Z \quad (t \rightarrow \infty),$$

тогда и только тогда, когда существует случайная величина V такая, что

$$\frac{\Lambda(t) - \ell t}{s\sqrt{t}} \Rightarrow V \quad (t \rightarrow \infty).$$

При этом

$$P(Z < x) = E\Phi \left(x \sqrt{1 + \frac{a^2 s^2}{(a^2 + \sigma^2)\ell}} - \frac{asV}{\sqrt{(\sigma^2 + a^2)\ell}} \right), \quad x \in \mathbb{R}.$$

Несложно видеть, что предельная случайная величина Z допускает представление

$$Z \stackrel{d}{=} \left[1 + \frac{a^2 s^2}{(a^2 + \sigma^2)\ell} \right]^{-1/2} X + \frac{as}{\sqrt{(a^2 + \sigma^2)\ell + a^2 s^2}} \cdot V,$$

где X — случайная величина со стандартным нормальным распределением, независимая от случайной величины V .

Доказательство теоремы 2 совпадает с доказательством соответствующего результата из книги [10] с точностью до переобозначений.

Сравнивая предложенную модель с моделями Рида–Йоргенсена и Барндорфф–Нильсена–Соренсена, можно заметить, что теорема 1 (или теорема 2) играет ту же роль, какую в рассуждениях Рида–Йоргенсена и Барндорфф–Нильсена–Соренсена играет центральная предельная теорема для пуассоновских случайных сумм, обосновывающая переход от (1) к (2). С использованием их логики, приведенные выше рассуждения могут быть продолжены следующим образом. При каждом фиксированном достаточно большом t из теоремы 1 вытекает приближенное равенство

$$P(S(t) < x) \approx E\Phi \left(\frac{x - C(t)}{D(t)k\sqrt{b_2}} - V \frac{\sqrt{|a|}}{k\sqrt{b_2}} \right). \quad (6)$$

Теперь, например, привлекая соображения «случайности времени достижения» из рассуждений Рида–Йоргенсена или Соренсена, приведенных во введении, можно считать время t , а вместе с ним и функции $C(t)$ и $D(t)$ случайными, так что в итоге получается совсем общая модель для распределения логарифма размера частицы при дроблении, имеющая вид сдвиг-масштабной смеси нормальных законов. В частности, обозначив

$$\zeta = C(t), \quad \chi = V \frac{\sqrt{|a|}}{k\sqrt{b_2}},$$

получаем, что при $|C(t)| = k^2 D^2(t)$ эта смесь принимает вид

$$P(S(t) < x) \approx E\Phi \left(\frac{x - \zeta}{\sqrt{b_2|\zeta|}} - \chi \right).$$

Теперь, наделяя величины ζ и χ теми или иными распределениями, можно получить ту или иную конкретную модель для распределения размера частицы при дроблении. В частности, пусть ζ имеет показательное распределение. Тогда итоговая модель распределения размера частицы имеет вид масштабной смеси лог-лапласовских распределений, т.е. получается обобщение модели Рида–Йоргенсена (в рамках которой предполагается, что χ независима от ζ и имеет нормальное распределение). Если же ζ имеет IG -распределение, то итоговая модель имеет вид масштабной смеси лог- GIG -распределений, т.е. получается обобщение модели Барндорфф-Нильсена–Соренсена (в рамках которой предполагается, что χ вырождена).

2.3 Оценки точности асимптотических аппроксимаций распределения логарифмов размеров частиц при дроблении сдвиговыми смесями нормальных законов

Рассмотрим скорость сходимости в теореме 2, что позволит получить оценку точности аппроксимации (6) при специальном виде центрирующих и нормирующих констант.

Прежде чем сформулировать основной результат данного раздела, введем дополнительные обозначения:

$$\begin{aligned} \beta^3 &= E|X_1|^3; \\ L_3 &= \frac{\beta^3}{(a^2 + \sigma^2)^{3/2}}; \\ F_t(x) &= P\left(\frac{S(t) - alt}{\sqrt{t[\ell(a^2 + \sigma^2) + a^2s^2]}} < x\right); \\ \rho_t &= \sup_x \left| F_t(x) - E\Phi\left(x\sqrt{1 + \frac{a^2s^2}{(a^2 + \sigma^2)\ell}} - \frac{asV}{\sqrt{(\sigma^2 + a^2)\ell}}\right) \right|; \\ \Delta_t &= \sup_v \left| P\left(\frac{\Lambda(t) - \ell t}{s\sqrt{t}} < v\right) - P(V < v) \right|. \end{aligned}$$

Теорема 3. Пусть $\beta^3 < \infty$, $E|V| < \infty$. Тогда справедлива оценка

$$\rho_t \leq \Delta_t + \frac{1}{\sqrt{t}} \inf_{\epsilon \in (0,1)} \left\{ \frac{C_0 L_3}{\sqrt{(1-\epsilon)\ell}} + \frac{s}{\ell} \left(\frac{E|V|}{\epsilon} + Q(\epsilon) \right) \right\},$$

где C_0 — абсолютная постоянная в неравенстве Берри–Эссеена, $C_0 \leq 0,7005$,

$$Q(\epsilon) = \max \left\{ \frac{1}{\epsilon}, \frac{\sqrt{1+\epsilon}}{(1+\sqrt{1-\epsilon})\sqrt{2\pi\epsilon(1-\epsilon)}} \right\}.$$

Доказательство этой теоремы приведено в работе [15].

Если дополнительно к условиям теоремы 3 предположить, что семейство случайных величин

$$\left\{ \left| \frac{\Lambda(t) - \ell t}{s\sqrt{t}} \right| \right\}_{t>0}$$

равномерно интегрируемо, то с помощью неравенства Ляпунова можно получить неравенство

$$\begin{aligned} E|V| &= \lim_{t \rightarrow \infty} E \left| \frac{\Lambda(t) - \ell t}{s\sqrt{t}} \right| \leq \\ &\leq \lim_{n \rightarrow \infty} \sqrt{D\left(\frac{\Lambda(t) - \ell t}{s\sqrt{t}}\right)} = 1. \end{aligned} \quad (7)$$

При этом из теоремы 3 очевидным образом вытекает следующий результат.

Следствие 2. Пусть в дополнение к условиям теоремы 3 при каждом $t > 0$ выполнено (7). Тогда справедлива оценка

$$\rho_t \leq \Delta_t + \frac{1}{\sqrt{t}} \inf_{\epsilon \in (0,1)} \left\{ \frac{C_0 L_3}{\sqrt{(1-\epsilon)\ell}} + \frac{s}{\ell} \left(\frac{1}{\epsilon} + Q(\epsilon) \right) \right\},$$

где $C_0 \leq 0,7005$,

$$Q(\epsilon) = \max \left\{ \frac{1}{\epsilon}, \frac{\sqrt{1+\epsilon}}{(1+\sqrt{1-\epsilon})\sqrt{2\pi\epsilon(1-\epsilon)}} \right\}.$$

Замечание 1. Теорема 3 справедлива при любых фиксированных значениях параметров t , ℓ и s . Поэтому ее можно использовать для формализации возможной зависимости интенсивности потока дроблений от текущего размера частицы. Действительно, обозначим средний (ожидаемый) логарифм размера частицы в момент времени t через $m(t)$,

$$m(t) = ES(t) = alt.$$

Тогда

$$\ell = \ell(t) = \frac{m(t)}{at}.$$

Соответственно, тогда и $s = s(t)$. Конкретный вид функций $m(t)$ и $s(t)$ зависит от конкретных физических, химических, технологических и/или геологических условий. В таком случае, как следует из теоремы 3, для возможности аппроксимации

распределения логарифма частицы при дроблении сдвиговыми смесями нормальных законов достаточно, чтобы при $t \rightarrow \infty$:

- 1) $\Delta_t \rightarrow 0$;
- 2) $|m(t)| \rightarrow \infty$;
- 3) $t^{1/2}s(t) = o(|m(t)|)$.

Замечание 2. Пусть, к примеру, процесс $\Lambda(t)$ является неоднородным пуассоновским с накопленной интенсивностью $\ell t = s^2 t$. В таком случае распределение случайной величины $N(t)$ при каждом $t > 0$ является так называемым пуассон-пуассоновским или (инфекционным) распределением Неймана типа А. Класс таких распределений введен Ю. Нейманом в работе [16] в связи с некоторыми задачами из области бактериологии и энтомологии. В этом случае величина Δ_t также имеет порядок $O(t^{-1/2})$, поскольку в условиях примера справедлива оценка

$$\Delta_t \leq \frac{C_0}{\sqrt{\ell t}} = C_0 \sqrt{\frac{a}{m(t)}},$$

причем случайная величина V имеет стандартное нормальное распределение (см., например, [11]), а следовательно, распределение случайной величины Z также является стандартным нормальным. В таком случае имеет место оценка

$$\rho_t \leq \sqrt{\frac{a}{m(t)}} \inf_{\epsilon \in (0,1)} \left\{ C_0 \left(1 + \frac{L_3}{\sqrt{1-\epsilon}} \right) + \frac{1}{\epsilon} + Q(\epsilon) \right\}.$$

Тогда, как следует из сказанного в замечании 1, для возможности аппроксимации распределения логарифма частицы при дроблении сдвиговыми смесями нормальных законов достаточно, чтобы $|m(t)| \rightarrow \infty$ при $t \rightarrow \infty$.

Замечание 3. Другим примером управляющего процесса $\Lambda(t)$, при котором величина Δ_t также имеет порядок $O(t^{-1/2})$, является случай, когда $\Lambda(t) = G(t)$, где $G(t)$ — гамма-процесс Леви, приращение которого на единичном интервале имеет гамма-распределение с параметром масштаба $\lambda = \ell/s^2$ и параметром формы $\alpha = \ell^2/s^2$. В таком случае распределение случайной величины $N(t)$ при каждом $t > 0$ является отрицательным биномиальным (см., например, [11]). При этом случайная величина V имеет стандартное нормальное распределение (см., например, [11]), а следовательно, распределение случайной величины Z опять-таки является стандартным нормальным.

Замечание 4. Если случайный процесс $N(t)$ является однородным пуассоновским (с интенсивностью ℓ),

то, очевидно, $s = 0$. При этом логично считать, что распределение случайной величины V является вырожденным в нуле и $\Delta_t = 0$. Поэтому в таком случае оценка, приведенная в теореме 1, естественно переходит в оценку скорости сходимости распределений пуассоновских случайных сумм к нормальному закону, приведенную, например, в работе [17] с константой C_0 , уточненной в работе [18].

Литература

1. Разумовский Н. К. ДАН СССР, 1940. Т. 28. № 8. С. 55–57.
2. Колмогоров А. Н. О логарифмически нормальном законе распределения размеров частиц при дроблении // ДАН СССР, 1941. Т. 31. С. 99–101.
3. Bagnold R. A. The physics of blown sand and desert dunes. — London: Methuen, 1941.
4. Reed W. J., Jorgensen M. The double Pareto-Lognormal distribution — a new parametric model for size distribution // Communications in Statistics — Theory and Methods, 2004. Vol. 33. No. 8. P. 1733–1753.
5. Sørensen M. On the size distribution of sand // Dept. of Applied Mathematics and Statistics, University of Copenhagen. Working paper. 2006.
6. Barndorff-Nielsen O. Exponentially decreasing distributions for the logarithm of particle size // Proc. Roy. Soc. L. A, 1977. Vol. 353. P. 401–419.
7. Vincent P. Differentiation of modern beach and coastal dune sands — a logistic regression approach using the parameters of the hyperbolic function // Sediment. Geology, 1986. Vol. 49. P. 167–176.
8. McArthur D. S. Distinctions between grain-size distribution of accretion and encroachment deposits in an inland dune // Sediment. Geology, 1987. Vol. 54. P. 147–163.
9. Hartmann D. Cross-shore selective sorting process and grain size distributional shape // Acta Mech. [Suppl.], 1991. Vol. 2. P. 49–63.
10. Bening V. E., Korolev V. Yu. Generalized Poisson models and their applications in insurance and finance. — Utrecht: VSP, 2002.
11. Королёв В. Ю., Бенинг В. Е., Шоргин С. Я. Математические основы теории риска. — М.: Физматлит, 2007.
12. Gnedenko B. V., Korolev V. Yu. Random summation: Limit theorems and applications. — Boca Raton: CRC Press, 1996.
13. Гнеденко Б. В., Колмогоров А. Н. Предельные распределения для сумм независимых случайных величин. — М.: ГИТТЛ, 1949.
14. Korolev V. Yu. A general theorem on the limit behavior of superpositions of independent random processes with applications to Cox processes // J. Math. Sci., 1996. Vol. 81. No. 5. P. 2951–2956.

15. Артохов С. В., Королёв В. Ю. Оценки скорости сходимости распределений обобщенных дважды стохастических пуассоновских процессов с ненулевым средним к сдвиговым смесям нормальных законов // Обозрение промышленной и прикладной математики, 2008. Т. 15. Вып. 6. С. 988–998.
16. Neyman J. On a new class of “contagious” distributions, applicable in entomology and bacteriology // Ann. Math. Statist., 1939. Vol. 10. P. 35–57.
17. Шевцова И. Г. Уточнение структуры оценок скорости сходимости в центральной предельной теореме для сумм независимых случайных величин. Дисс. на соискание ученой степени кандидата физ.-мат. наук. — М.: Московский государственный университет, 2006.
18. Шевцова И. Г. Об абсолютной постоянной в неравенстве Берри–Эссеена // Сборник статей молодых ученых факультета ВМиК МГУ. Вып. 5. — М.: Изд-во факультета ВМиК МГУ, 2008. С. 101–110.

НЕКОТОРЫЕ ОЦЕНКИ ДЛЯ ХАРАКТЕРИСТИЧЕСКИХ ФУНКЦИЙ С ПРИМЕНЕНИЕМ К УТОЧНЕНИЮ НЕРАВЕНСТВА МИЗЕСА*

И. Г. Шевцова¹

Аннотация: Построены новые оценки для модулей характеристических функций распределений, имеющих моменты порядка $2 + \delta$, $0 < \delta \leq 1$. Уточнено моментное неравенство Мизеса для решетчатых распределений.

Ключевые слова: преобразование Фурье; характеристическая функция; симметризация; свертка; решетчатое распределение; арифметическое распределение; шаг распределения

1 Введение

При построении вероятностных моделей стохастических ситуаций часто в качестве аналитической модели распределения вероятностей используется его асимптотическая аппроксимация, возникающая в соответствующей предельной теореме. Примером может служить повсеместное применение нормальной аппроксимации, основанное на центральной предельной теореме. Однако при этом столь же часто обходится стороной вопрос о точности этой аппроксимации, играющий важную роль при обосновании адекватности вероятностной модели. Данная работа посвящена уточнению оценок для характеристических функций — основного аппарата, используемого при построении конкретных числовых оценок точности нормальной аппроксимации.

Пусть $F(x)$, $x \in \mathbb{R}$, — функция распределения с характеристической функцией $f(t)$ такая, что

$$\left. \begin{aligned} \int_{-\infty}^{\infty} x dF(x) &= 0; \\ \sigma^2 &\equiv \int_{-\infty}^{\infty} x^2 dF(x) > 0; \\ \beta_{2+\delta} &\equiv \int_{-\infty}^{\infty} |x|^{2+\delta} dF(x) < \infty \end{aligned} \right\} \quad (1)$$

для некоторого $0 < \delta \leq 1$.

В 1965 г. В. М. Золотарёв [1] доказал, что при $\delta = 1$ для $|f(t)|$ справедлива оценка

$$\ln |f(t)| \leq -\frac{\sigma^2 t^2}{2} + \frac{\beta_3 |t|^3}{3}, \quad t \in \mathbb{R}. \quad (2)$$

Два года спустя он уточнил [2] свой результат, показав, что коэффициент при $\beta_3 |t|^3$ можно снизить примерно до $0.1983 < 1/3$:

$$\ln |f(t)| \leq -\frac{\sigma^2 t^2}{2} + 2\kappa \beta_3 |t|^3, \quad t \in \mathbb{R}, \quad (3)$$

где

$$\kappa = \sup_{x>0} \frac{|\cos x - 1 + x^2/2|}{x^3} = 0,099161 \dots$$

(если быть совсем точными, в работе [2] неравенство (3) упомянуто без доказательства). Отношение коэффициентов при $|t|^3$ в (2) и (3) составляет

$$\frac{1}{6\kappa} = 1,6807 \dots$$

Шесть лет спустя Х. Правитц [3] доказал неравенство

$$|f(t)|^2 \leq 1 - \sigma^2 t^2 + 2\kappa (\beta_3 + \sigma^3) |t|^3, \quad t \in \mathbb{R}, \quad (4)$$

из которого вытекает еще одно уточнение оценки Золотарёва:

$$\ln |f(t)| \leq -\frac{\sigma^2 t^2}{2} + \kappa (\beta_3 + \sigma^3) |t|^3, \quad t \in \mathbb{R}. \quad (5)$$

Отношение коэффициентов при $|t|^3$ в (3) и (5) в силу неравенства Ляпунова не меньше единицы и может быть сколь угодно близко к двум, если отношение Ляпунова β_3/σ^3 велико, а именно справедливы следующие соотношения:

$$1 = \inf_F \frac{2\beta_3}{\beta_3 + \sigma^3} \leq \sup_F \frac{2\beta_3}{\beta_3 + \sigma^3} = 2,$$

* Работа выполнена при поддержке Российского фонда фундаментальных исследований, проекты 08-01-00563, 08-01-00567, 08-07-00152 и 09-07-12032-офи-м, а также Совета по грантам президента Российской Федерации для поддержки молодых российских ученых — кандидатов наук, проект МК-654.2008.1.

¹ Факультет вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова, ishevtsova@cs.msu.su

где инфимум и супремум берутся по всем распределениям F , удовлетворяющим условиям (1) с $\delta = 1$.

Появление оценки (3) оказало значительное влияние на результаты вычисления абсолютной константы C в широко известном неравенстве Берри–Эссеена [4, 5], согласно которому для равномерного расстояния между n -кратной нормированной сверткой распределений F , удовлетворяющих условиям (1) с $\delta = 1$, и стандартным нормальным законом Φ справедливо соотношение

$$\sup_x |F^{*n}(x\sigma\sqrt{n}) - \Phi(x)| \leq CL_n^3; \quad L_n^3 = \frac{\beta_3}{\sigma^3\sqrt{n}}.$$

А именно использование более точной оценки (3) позволило Золотарёву существенно улучшить мажоранту константы C в этом неравенстве с 1,301 [6] до 0,82 [2, 7], при этом в работе [2] сам автор отмечает: «В настоящее время удалось осуществить дальнейшее снижение верхних оценок константы C за счет уточнения весьма важной для выбранного метода леммы 4 работы [1]» (суть упомянутой леммы — утверждение неравенства (2)).

В 1975 г. при решении аналогичной задачи об отыскании верхней асимптотической оценки константы C , справедливой для достаточно малых значений ляпуновской дроби L_n^3 , Правитц [8] получил результат

$$C \leq 0,5152, \text{ если } L_n^3 \leq 0,0985. \quad (6)$$

При этом существенную роль сыграла возможность использования оценки (4) и приводимой ниже оценки (8).

Неравенство Берри–Эссеена является наиболее популярным среди оценок скорости сходимости в центральной предельной теореме и играет значительную роль при решении конкретных задач, возникающих в теории надежности, управлении запасами, страховой и финансовой математике и многих других важных областях прикладной науки. Успех в применении вероятностных моделей к решению этих задач во многом зависит от точности используемых методов. Таким образом, актуальность задачи несомненна.

В то же время при решении задач, связанных с большими рисками, приходится отказываться от условия $\beta_3 < \infty$ и каким-то образом переносить полученные результаты на случай $\delta < 1$.

К одной из первых работ в этой области следует отнести диссертацию В. Тысиака, в которой фактически была обобщена оценка Правитца (4) для $0 < \delta < 1$:

$$|f(t)|^2 \leq 1 - \sigma^2 t^2 + 2\kappa(\delta) (\beta_{2+\delta} + \sigma^{2+\delta}) |t|^{2+\delta}, \quad t \in \mathbb{R}, \quad (7)$$

где

$$\kappa(\delta) = \sup_{x>0} \frac{|\cos x - 1 + x^2/2|}{x^{2+\delta}}, \quad 0 < \delta \leq 1,$$

хотя автор анонсирует и использует полученный результат лишь в форме Золотарёва (3):

$$\ln |f(t)| \leq -\frac{\sigma^2 t^2}{2} + 2\kappa(\delta)\beta_{2+\delta}|t|^{2+\delta}, \quad t \in \mathbb{R}.$$

Применяя последний к неравенству сглаживания Золотарёва, Тысиак [9] в 1983 г. получил верхние оценки константы в аналоге неравенства Берри–Эссеена для распределений, не имеющих третьего момента. Доказательство неравенства (7) также можно найти в монографии Ушакова [10].

В упомянутой выше работе [3] Правитц получил еще одну оценку, на которую существенно опирается доказательство асимптотического результата (6):

$$|f(t)|^2 \leq 1 - 2 \left(\frac{\sigma^2}{\beta_3 + \sigma^3} \right)^2 \left(1 - \cos \frac{(\beta_3 + \sigma^3)t}{\sigma^2} \right), \quad (8)$$

$$\theta_0 \leq \left(\frac{\beta_3 + \sigma^3}{\sigma^2} \right) |t| \leq 2\pi,$$

где $\theta_0 = 3,5995896 \dots$ — единственный корень уравнения

$$3(1 - \cos \theta) - \theta \sin \theta - \frac{\theta^2}{2} = 0,$$

лежащий в интервале $(\pi, 2\pi)$. Кроме того, в указанной работе при каждом $t \in \mathbb{R}$ был найден минимум из трех оценок для $|f(t)|$, устанавливаемых неравенствами (4), (8) и $|f(t)| \leq 1$, и показано, что наилучшая в этом смысле оценка имеет вид (для упрощения обозначений здесь полагаем $\sigma^2 = 1$):

$$|f(t)| \leq \begin{cases} 1 - \frac{t^2}{2} + \kappa(\delta) (\beta_3 + 1) |t|^3, & |t| \leq \frac{\theta_0}{\beta_3 + 1}; \\ 1 - \frac{1 - \cos((\beta_3 + 1)t)}{(\beta_3 + 1)^2}, & \theta_0 \leq (\beta_3 + 1)|t| \leq 2\pi; \\ 1, & |t| \geq \frac{2\pi}{\beta_3 + 1}. \end{cases}$$

Отсюда, кстати, вытекает (см. теорему 5), что максимальный шаг h решетчатых распределений удовлетворяет соотношению:

$$\sigma^2 h \leq \beta_3 + \sigma^3, \quad (9)$$

в то время как широко известное неравенство Мизеса [11] устанавливает лишь более грубую оценку

$$\sigma^2 h \leq 2\beta_3. \quad (10)$$

Как известно, неравенство Мизеса является в определенном смысле точным: знак равенства в нем достигается на симметричном распределении Бернулли, для которого $\beta_3/\sigma^3 = 1$, — и это единственный случай, когда правые части (9) и (10) совпадают. В остальных же случаях неравенство (10) строгое.

В данной работе обобщен результат Правитца (8) на случай $0 < \delta < 1$ (здесь для упрощения обозначений снова полагаем $\sigma^2 = 1$):

$$|f(t)|^2 \leq 1 - 2 \frac{1 - \cos((\beta_{2+\delta} + 1)^{1/\delta} t)}{(\beta_{2+\delta} + 1)^{2/\delta}}; \quad (11)$$

$$\theta_0(\delta) \leq (\beta_{2+\delta} + 1)^{1/\delta} |t| \leq 2\pi,$$

где $\theta_0(\delta)$ — единственный корень уравнения

$$\frac{\delta\theta^2}{2} + \theta \sin \theta + (2 + \delta)(\cos \theta - 1) = 0,$$

лежащий в интервале $(\pi, 2\pi)$. Также в данной работе показано, что минимум из трех оценок для $|f(t)|$, устанавливаемых неравенствами (7), (11) и $|f(t)| \leq 1$, имеет вид

$$|f(t)| \leq \begin{cases} 1 - \frac{t^2}{2} + \varkappa(\delta) (\beta_{2+\delta} + 1) |t|^{2+\delta}, & |t| \leq \frac{\theta_0(\delta)}{(\beta_{2+\delta} + 1)^{1/\delta}}; \\ 1 - \frac{1 - \cos((\beta_{2+\delta} + 1)^{1/\delta} t)}{(\beta_{2+\delta} + 1)^{2/\delta}}, & \theta_0(\delta) \leq (\beta_{2+\delta} + 1)^{1/\delta} |t| \leq 2\pi; \\ 1, & |t| \geq \frac{2\pi}{(\beta_{2+\delta} + 1)^{1/\delta}}. \end{cases}$$

В качестве следствия полученной оценки для $|f(t)|$ в данной работе доказано более общее по сравнению с (9) неравенство:

$$\frac{h}{\sigma} \leq \left(\frac{\beta_{2+\delta}}{\sigma^{2+\delta}} + 1 \right)^{1/\delta}, \quad (12)$$

справедливое для любых решетчатых распределений, удовлетворяющих условиям (1) с произвольным $0 < \delta \leq 1$. Также показано, что неравенство (12) не может быть улучшено ни при каком $0 < \delta \leq 1$.

2 Вспомогательные утверждения

При построении верхних оценок для вещественных характеристических функций успех во многом

зависит от того, насколько точную мажоранту удастся построить для функции $\cos x$, $x \in \mathbb{R}$. Если при этом предполагается существование у рассматриваемого распределения абсолютного момента некоторого порядка $s > 0$ (возможно, дробного), то мажоранта для $\cos x$ может включать степенные члены порядка не выше s . Особый интерес с точки зрения центральной предельной теоремы представляет случай $s \geq 2$. Из разложения функции $\cos x$ в ряд Маклорена вытекает, что ее поведение в окрестности нуля с точностью до бесконечно малых более высокого порядка совпадает с поведением функции $(1 - x^2/2)$. Поэтому при конструировании верхней оценки, имеющей вид полинома, необходимо включить в нее слагаемое вида $(1 - ax^2)$ с $0 \leq a \leq 1/2$ (случай $a < 0$ приводит к слишком грубой оценке в окрестности нуля, и потому интереса не представляет). С другой стороны, при больших x добиться выполнения неравенства $\cos x \leq 1 - ax^2$ с $a > 0$ не представляется возможным, поэтому в мажоранту необходимо включить еще слагаемое вида x^s с неотрицательным коэффициентом.

Пусть теперь $s = 2 + \delta$, где $0 < \delta \leq 1$. Тогда при каждом $0 \leq a \leq 1/2$ найдется число $b = b(a) \geq 0$ такое, что для всех $x \in \mathbb{R}$ выполняется неравенство

$$\cos x \leq 1 - ax^2 + b|x|^{2+\delta}.$$

При этом минимально возможное значение b , гарантирующее выполнение указанного неравенства, имеет вид

$$b(a) = \sup_{x \in \mathbb{R}} \frac{|\cos x - 1 + ax^2|}{|x|^{2+\delta}}, \quad 0 \leq a \leq \frac{1}{2},$$

другими словами, для любого $0 \leq a \leq 1/2$

$$\inf_{x \in \mathbb{R}} (1 - ax^2 + b(a)|x|^{2+\delta} - \cos x) = 0.$$

Следующее утверждение дает более удобное параметрическое представление пары $(a, b(a))$, $0 \leq a \leq 1/2$, являющейся решением описанной задачи.

Лемма 1. Для произвольного $0 < \delta \leq 1$ и всех $x \in \mathbb{R}$ справедливо неравенство

$$\cos x \leq 1 - a(\delta, \theta)x^2 + b(\delta, \theta)|x|^{2+\delta}, \quad \theta_0(\delta) \leq \theta \leq 2\pi, \quad (13)$$

где

$$a(\delta, \theta) = \frac{2 + \delta}{\delta} \frac{1 - \cos \theta}{\theta^2} - \frac{1}{\delta} \frac{\sin \theta}{\theta}; \quad (14)$$

$$b(\delta, \theta) = \frac{2}{\delta} \frac{1 - \cos \theta}{\theta^{2+\delta}} - \frac{1}{\delta} \frac{\sin \theta}{\theta^{1+\delta}}; \quad (15)$$

$\theta_0(\delta)$ — единственный корень уравнения $a(\delta, \theta) = 1/2$, лежащий в интервале $(\pi, 2\pi)$.

Данное утверждение впервые было получено при $\delta = 1$ в работе Правитца [3]. Для других значений δ лемма 1 была доказана лишь в тривиальном случае $\theta = \theta_0(\delta)$ (с $a(\delta, \theta) = 1/2$) в диссертации Тысиака [9], а также, в несколько ином виде, в монографии Ушакова [10]. Для произвольных $0 < \delta < 1$ и $\theta \in (\theta_0(\delta), 2\pi]$ данное утверждение публикуется, по-видимому, впервые. Доказательству леммы 1 будет предпослано несколько вспомогательных утверждений.

Лемма 2. Функция

$$\varphi(x) = \sin x - x \cos x, \quad 0 < x \leq 2\pi,$$

имеет единственный корень x_0 , лежащий в интервале $(\pi, 3\pi/2)$, причем

$$\begin{aligned} \varphi(x) &> 0, & 0 < x < x_0; \\ \varphi(x) &< 0, & x_0 < x \leq 2\pi. \end{aligned}$$

Данное утверждение является очевидным следствием свойств функции $\operatorname{tg} x$.

Лемма 3. Функция

$$\varphi(x) = \frac{\sin x}{x} - \frac{1 + \cos x}{2}, \quad 0 < x \leq 2\pi,$$

имеет единственный корень $x_1 = \pi$, причем

$$\begin{aligned} \varphi(x) &> 0, & 0 < x < \pi; \\ \varphi(x) &< 0, & \pi < x \leq 2\pi. \end{aligned}$$

Доказательство. Знак $\varphi(x)$ определяется знаком выражения

$$\begin{aligned} 2 \sin x - x(1 + \cos x) &= 4 \sin \frac{x}{2} \cos \frac{x}{2} - \\ - 2x \cos^2 \frac{x}{2} &= 4 \cos \frac{x}{2} \left(\sin \frac{x}{2} - \frac{x}{2} \cos \frac{x}{2} \right). \end{aligned} \quad (16)$$

Согласно лемме 2,

$$\sin \frac{x}{2} - \frac{x}{2} \cos \frac{x}{2} > 0, \quad 0 < x \leq 2\pi,$$

поэтому знак (16) определяется множителем $\cos(x/2)$, который положителен при $x \in (0, \pi)$ и отрицателен при $x \in (\pi, 2\pi)$.

Лемма 4. Для произвольного $0 \leq p \leq 1/2$ функция

$$\varphi_p(x) = \frac{\sin x}{x} - (p + (1-p) \cos x), \quad 0 < x \leq 2\pi,$$

имеет единственный корень $x_p \in [\pi, 3\pi/2)$, причем

$$\begin{aligned} \varphi_p(x) &> 0, & 0 < x < x_p; \\ \varphi_p(x) &< 0, & x_p < x \leq 2\pi. \end{aligned}$$

Доказательство. Случаи $p = 0$ и $p = 1/2$ рассмотрены в леммах 2 и 3 соответственно, поэтому будем считать, что $0 < p < 1/2$. Разобьем промежутки $(0, 2\pi]$ на три части и рассмотрим поведение функции $\varphi_p(x)$ на каждом из них.

Для $x \in (0, \pi)$ в силу леммы 3 имеем

$$\begin{aligned} p + (1-p) \cos x &= \cos x + p(1 - \cos x) \leq \\ &\leq \cos x + \frac{1}{2}(1 - \cos x) = \frac{1 + \cos x}{2} < \frac{\sin x}{x}; \end{aligned}$$

следовательно, $\varphi_p(x) > 0$ для $0 < x < \pi$. Заметим, что $\varphi_p(\pi) = 1 - 2p > 0$, поэтому

$$\varphi_p(x) > 0, \quad 0 < x \leq \pi.$$

Пусть x_0 — единственный корень уравнения

$$\sin x = x \cos x, \quad 0 < x \leq \pi,$$

согласно лемме 2 лежащий в интервале $(\pi, 3\pi/2)$. Тогда $\sin x/x < \cos x$ для $x_0 < x \leq 2\pi$ и

$$\begin{aligned} p + (1-p) \cos x &= \cos x + p(1 - \cos x) \geq \\ &\geq \cos x > \frac{\sin x}{x}; \end{aligned}$$

следовательно, $\varphi_p(x) < 0$ для $x_0 < x \leq 2\pi$. В точке x_0 имеем

$$\begin{aligned} \varphi_p(x_0) &= \frac{\sin x_0}{x_0} - (p + (1-p) \cos x_0) = \\ &= \cos x_0 - (p + (1-p) \cos x_0) = p(\cos x_0 - 1) < 0, \end{aligned}$$

поскольку $x_0 \in (\pi, 3\pi/2)$, а значит,

$$\varphi_p(x) < 0, \quad x_0 \leq x \leq 2\pi.$$

На концах интервала $\pi < x < x_0$ непрерывная функция φ_p принимает значения разных знаков, поэтому хотя бы в одной точке обращается в нуль. Единственность нуля вытекает из того, что функция $\sin x/x$ монотонно убывает при $\pi < x < x_0$ (поскольку ее производная

$$\left(\frac{\sin x}{x} \right)' = \frac{x \cos x - \sin x}{x^2}$$

в силу леммы 2 отрицательна при $x < x_0$), а функция $p + (1-p) \cos x$ монотонно возрастает при $\pi < x < x_0$.

Доказательство леммы 1. В силу четности фигурирующих в неравенстве (13) функций можно считать, что $x > 0$ (для $x = 0$ неравенство (13) очевидно).

Рассмотрим сначала $0 < x \leq 2\pi$. Введем функцию

$$P_{a,b}(x) = 1 - ax^2 + bx^{2+\delta} - \cos x, \quad 0 < x \leq 2\pi.$$

Тогда неравенство (13) для указанных x равносильно уравнению

$$\inf_{0 < x \leq 2\pi} P_{a,b}(x) = 0,$$

решением которого является множество пар $(a(\theta), b(\theta))$, $\theta \in (0, 2\pi]$, удовлетворяющих следующим условиям:

$$\begin{aligned} P_{a,b}(\theta) &= 0; \\ \frac{d}{d\theta} P_{a,b}(\theta) &= 0; \\ \frac{d^2}{d\theta^2} P_{a,b}(\theta) &> 0; \end{aligned}$$

$$0 \leq a \leq 1/2, \quad b \geq 0.$$

Первые два уравнения равносильны системе

$$a\theta^2 = 1 - \cos \theta + b\theta^{2+\delta}; \quad (17)$$

$$b = \frac{2a\theta - \sin \theta}{(2 + \delta)\theta^{1+\delta}}, \quad (18)$$

решение которой задается формулами (14) и (15). Поскольку δ фиксировано, для краткости будем опускать аргумент δ у функций $a = a(\delta, \theta)$ и $b = b(\delta, \theta)$, используя обозначения $a(\theta)$ и $b(\theta)$.

Проверим условие неотрицательности $b(\theta)$ и $a(\theta)$. Поскольку $\theta > 0$, знак $b(\theta)$ определяется знаком выражения

$$\begin{aligned} \delta\theta^{2+\delta}b(\theta) &= 2(1 - \cos \theta) - \theta \sin \theta = \\ &= 4 \sin^2 \frac{\theta}{2} - 2\theta \sin \frac{\theta}{2} \cos \frac{\theta}{2} = \\ &= 4 \sin \frac{\theta}{2} \left(\sin \frac{\theta}{2} - \frac{\theta}{2} \cos \frac{\theta}{2} \right), \end{aligned}$$

которое, в свою очередь, принимает только неотрицательные значения, так как $\sin(\theta/2) \geq 0$ при всех $0 < \theta \leq 2\pi$, а выражение в скобках, согласно лемме 2, положительно для указанных θ . Таким образом, $b(\theta) \geq 0$ для всех $0 < \theta \leq 2\pi$. Отсюда немедленно вытекает неотрицательность $a(\theta)$, поскольку в силу (17) имеем

$$\begin{aligned} \theta^2 a(\theta) &= 1 - \cos \theta + b(\theta)\theta^{2+\delta} \geq \\ &\geq 1 - \cos \theta \geq 0, \quad 0 < \theta \leq 2\pi. \end{aligned}$$

Условие $a(\theta) \leq 1/2$ равносильно неравенству

$$\begin{aligned} g(\theta) &\equiv \delta\theta^2(a(\theta) - 1/2) = \\ &= (2 + \delta)(1 - \cos \theta) - \theta \sin \theta - \frac{\delta\theta^2}{2} \leq 0. \end{aligned}$$

Заметим, что

$$g(0) = 0; \quad g(2\pi) = -2\pi^2\delta < 0;$$

$$\begin{aligned} g'(\theta) &= (2 + \delta) \sin \theta - \sin \theta - \theta \cos \theta - \delta\theta = \\ &= (1 + \delta) \left[\frac{\sin \theta}{\theta} - \left(\frac{\delta}{1 + \delta} + \frac{1}{1 + \delta} \cos \theta \right) \right]. \end{aligned}$$

Поскольку $0 < \delta/(1 + \delta) \leq 1/2$, в силу леммы 4 заключаем, что на интервале $(0, 2\pi)$ функция $g'(\theta)$ имеет единственный нуль $\theta_1(\delta) \in [\pi, 3\pi/2)$, причем

$$\begin{aligned} g'(\theta) &> 0, \quad 0 < \theta < \theta_1(\delta); \\ g'(\theta) &< 0, \quad \theta_1(\delta) < \theta \leq 2\pi. \end{aligned}$$

Следовательно, $g(\theta)$ возрастает на $(0, \theta_1(\delta))$ и убывает на $(\theta_1(\delta), 2\pi)$, меняя знак в единственной точке $\theta_0(\delta) > \theta_1(\delta)$. Отсюда вытекает, что $a(\delta, \theta) \leq 1/2$ для $\theta \in [\theta_0(\delta), 2\pi]$, где $\theta_0(\delta)$ — единственный корень уравнения $g(\theta) = 0$ на интервале $0 < \theta \leq 2\pi$, равносильного уравнению $a(\delta, \theta) = 1/2$. При этом $\theta_0(\delta) \in (\theta_1(\delta), 2\pi) \subset [\pi, 2\pi)$ для всех $\delta \in (0, 1]$.

Проверим положительность второй производной функции $P_{a,b}(\theta)$ для задаваемых формулами (14) и (15) значений $a = a(\theta)$, $b = b(\theta)$ на отрезке $[\theta_0(\delta), 2\pi]$. Рассмотрим функцию

$$\begin{aligned} h(\theta) &\equiv \theta^2 \frac{d^2}{d\theta^2} P_{a,b}(\theta) \Big|_{\substack{a=a(\theta) \\ b=b(\theta)}} = \\ &= \theta^2(-2a(\theta) + (1 + \delta)(2 + \delta)\theta^\delta b(\theta) + \cos \theta) = \\ &= 2(2 + \delta)(1 - \cos \theta) - (3 + \delta)\theta \sin \theta + \\ &\quad + \theta^2 \cos \theta, \quad \theta_0(\delta) \leq \theta \leq 2\pi. \end{aligned}$$

Покажем сначала, что $h(\theta)$ принимает положительные значения на концах рассматриваемого отрезка. Для краткости будем опускать аргумент δ у функции $\theta_0(\delta)$. Имеем

$$h(2\pi) = (2\pi)^2 > 0,$$

$$h(\theta_0) = \theta_0^2(-2a(\theta_0) + (1 + \delta)(2 + \delta)\theta_0^\delta b(\theta_0) + \cos \theta_0).$$

Учитывая условие $a(\theta_0) = 1/2$ и соотношение (18), запишем $h(\theta_0)$ в виде

$$\begin{aligned} h(\theta_0) &= \theta_0^2 \left(-1 + (1 + \delta) \frac{\theta_0 - \sin \theta_0}{\theta_0} + \cos \theta_0 \right) = \\ &= -(1 + \delta)\theta_0^2 \left[\frac{\sin \theta_0}{\theta_0} - \left(\frac{\delta}{1 + \delta} + \frac{1}{1 + \delta} \cos \theta_0 \right) \right] \equiv \\ &\equiv -\theta_0^2 g'(\theta_0). \end{aligned}$$

Как было замечено выше, $g'(\theta) < 0$ при $\theta_1(\delta) < \theta \leq 2\pi$, а также $\theta_0(\delta) > \theta_1(\delta)$ для всех $0 < \delta \leq 1$. Следовательно, $h(\theta_0) > 0$.

Рассмотрим поведение $h(\theta)$ внутри отрезка $[\theta_0(\delta), 2\pi]$. Вычисляя производную

$$h'(\theta) = (1 + \delta) \sin \theta - (1 + \delta)\theta \cos \theta - \theta^2 \sin \theta,$$

замечаем, что условие $h'(\theta) = 0$ равносильно уравнению

$$\operatorname{tg} \theta = \frac{\theta}{1 - \theta^2/(1 + \delta)},$$

которое на отрезке $[\theta_0(\delta), 2\pi] \subset (\pi, 2\pi]$ имеет единственный корень $\theta_2(\delta) \in (3\pi/2, 2\pi)$, причем

$$\begin{aligned} h'(\theta) > 0, \quad \theta_0(\delta) \leq \theta < \theta_2(\delta); \\ h'(\theta) < 0, \quad \theta_2(\delta) < \theta \leq 2\pi. \end{aligned}$$

Таким образом, $h(\theta)$ возрастает при $\theta_0(\delta) \leq \theta < \theta_2(\delta)$, убывает при $\theta_2(\delta) < \theta \leq 2\pi$ и принимает положительные значения на концах отрезка $\theta \in [\theta_0(\delta), 2\pi]$; следовательно,

$$\begin{aligned} h(\theta) = 2(2 + \delta)(1 - \cos \theta) - (3 + \delta)\theta \sin \theta + \\ + \theta^2 \cos \theta > 0, \quad \theta_0(\delta) \leq \theta \leq 2\pi. \end{aligned} \quad (19)$$

Для завершения доказательства леммы 1 осталось показать справедливость неравенства (13) при $x > 2\pi$. В нижеследующей лемме 5 будет показано, что функция

$$\varphi_x(\theta) = x^\delta b(\theta) - a(\theta), \quad \theta_0(\delta) \leq \theta \leq 2\pi,$$

монотонно убывает по θ при $x > \theta$, поэтому для всех $x \geq 2\pi \geq \theta$

$$\inf_{\theta_0(\delta) \leq \theta \leq 2\pi} \varphi_x(\theta) = \varphi_x(2\pi) = 0,$$

а следовательно,

$$\begin{aligned} 1 - a(\theta)x^2 + b(\theta)x^{2+\delta} = \\ = 1 + x^2 \varphi_x(\theta) \geq 1, \quad \theta_0(\delta) \leq \theta \leq 2\pi, \end{aligned}$$

в то время как $\cos x \leq 1$ для всех $x \in \mathbb{R}$. Таким образом, при $x > 2\pi$ неравенство (13) тривиально.

Замечание 1. Из построения функций $a(\delta, \theta)$ и $b(\delta, \theta)$ видно, что при каждом $\delta \in (0, 1]$ и $\theta \in [\theta_0(\delta), 2\pi]$

$$\begin{aligned} b(\delta, \theta) &= \sup_{x>0} \frac{|\cos x - 1 + a(\delta, \theta)x^2|}{x^{2+\delta}}, \\ a(\delta, \theta) &= \inf_{x>0} \frac{1 - \cos x + b(\delta, \theta)x^{2+\delta}}{x^2}, \end{aligned}$$

причем супремум и инфимум достигаются в точке $x = \theta$. Кроме того, из построения $\theta_0 = \theta_0(\delta)$

как корня уравнения $a(\delta, \theta) = 1/2$ на интервале $0 < \theta \leq 2\pi$ вытекает, что

$$\begin{aligned} a(\delta, \theta_0(\delta)) &= \frac{1}{2}; \\ b(\delta, \theta_0(\delta)) &= \sup_{x>0} \frac{|\cos x - 1 + x^2/2|}{x^{2+\delta}} \equiv \varkappa(\delta), \end{aligned}$$

причем супремум достигается в точке $x = \theta_0(\delta)$. Из (18) также вытекает представление

$$b(\delta, \theta_0(\delta)) = \frac{\theta_0(\delta) - \sin \theta_0(\delta)}{(2 + \delta)(\theta_0(\delta))^{1+\delta}}.$$

Лемма 5. Для всех $0 < \delta \leq 1$ и $t > 0$ функция

$$\varphi_t(\theta) = t^\delta b(\delta, \theta) - a(\delta, \theta), \quad \theta_0(\delta) \leq \theta \leq 2\pi,$$

монотонно возрастает по θ при $t < \theta$ и монотонно убывает по θ при $t > \theta$.

Доказательство. Для $\theta_0(\delta) \leq \theta \leq 2\pi$ введем функции

$$\begin{aligned} u(\theta) &\equiv -\frac{a(\delta, \theta)}{\delta} = \frac{\sin \theta}{\theta} - (2 + \delta) \frac{1 - \cos \theta}{\theta^2}; \\ v(\theta) &\equiv -\frac{\theta^\delta b(\delta, \theta)}{\delta} = \frac{\sin \theta}{\theta} - 2 \frac{1 - \cos \theta}{\theta^2} \equiv \\ &\equiv u(\theta) + \delta w(\theta); \\ w(\theta) &= \frac{1 - \cos \theta}{\theta^2}. \end{aligned}$$

Тогда можно записать

$$\varphi_t(\theta) = \delta \left(u(\theta) - \left(\frac{t}{\theta} \right)^\delta v(\theta) \right), \quad \theta_0(\delta) \leq \theta \leq 2\pi.$$

Найдем условия, при которых производная $\varphi_t(\theta)$ по θ меняет знак. Имеем

$$\begin{aligned} w'(\theta) &= \frac{\sin \theta}{\theta^2} - 2 \frac{1 - \cos \theta}{\theta^3} = \frac{1}{\theta} v(\theta); \\ v'(\theta) &= u'(\theta) + \delta w'(\theta) = u'(\theta) + \frac{\delta}{\theta} v(\theta); \end{aligned}$$

$$\begin{aligned} u'(\theta) &= \frac{1}{\theta^3} (2(2 + \delta)(1 - \cos \theta) - \\ &- (3 + \delta)\theta \sin \theta + \theta^2 \cos \theta). \end{aligned}$$

Заметим, что $u'(\theta)$ с точностью до множителя θ^{-3} совпадает с функцией $h(\theta)$, введенной при доказательстве леммы 1 (см. (19)). Там же было показано, что $h(\theta) > 0$ при $\theta_0(\delta) \leq \theta \leq 2\pi$, поэтому

$$u'(\theta) > 0, \quad \theta_0(\delta) \leq \theta \leq 2\pi.$$

Учитывая найденные соотношения для производных, получаем

$$\frac{d}{d\theta} \varphi_t(\theta) = \delta u'(\theta) + \frac{\delta t^\delta}{\theta^{1+\delta}} (\delta v(\theta) - \theta v'(\theta)) = \delta u'(\theta) \left(1 - \left(\frac{t}{\theta} \right)^\delta \right).$$

Замечание о том, что знак производной $\varphi_t(\theta)$ определяется знаком выражения в скобках, завершает доказательство леммы.

Лемма 6. Пусть X и Y — независимые одинаково распределенные случайные величины такие, что

$$EX = 0; EX^2 = \sigma^2; E|X|^{2+\delta} = \beta_{2+\delta} < \infty$$

для некоторого $0 < \delta \leq 1$. Тогда

$$E(X - Y)^2 = 2\sigma^2; \\ E|X - Y|^{2+\delta} \leq 2(\beta_{2+\delta} + \sigma^{2+\delta}).$$

Доказательство. Первое соотношение вытекает непосредственно из условий леммы:

$$E(X - Y)^2 = E(X^2 + 2XY + Y^2) = \\ = EX^2 + 2EXEY + EY^2 = 2\sigma^2.$$

Рассмотрим второе соотношение. В силу неравенства $|x - y|^\delta \leq |x|^\delta + |y|^\delta$, справедливого для всех $x, y \in \mathbb{R}$ и $\delta \in (0, 1]$, имеем

$$E|X - Y|^{2+\delta} = E(X - Y)^2 |X - Y|^\delta \leq \\ \leq E(X - Y)^2 (|X|^\delta + |Y|^\delta).$$

В силу независимости X и Y ,

$$EXY (|X|^\delta + |Y|^\delta) = E|X|^{2+\delta} EY + E|Y|^{2+\delta} EX = 0,$$

а следовательно,

$$E|X - Y|^{2+\delta} \leq E(X^2 + XY + Y^2) (|X|^\delta + |Y|^\delta) = \\ = E(X^2 + Y^2) (|X|^\delta + |Y|^\delta) = \\ = E|X|^{2+\delta} + EY^2 E|X|^\delta + EX^2 E|Y|^\delta + E|Y|^{2+\delta} = \\ = 2(\beta_{2+\delta} + \sigma^2 E|X|^\delta).$$

Теперь для доказательства леммы осталось заметить, что

$$E|X|^\delta \leq (EX^2)^{\delta/2} = \sigma^\delta$$

в силу неравенства Ляпунова.

3 Оценки для характеристических функций

Теорема 1. Пусть $f(t)$ — характеристическая функция некоторого распределения F , удовлетворяющего условиям (1). Тогда для всех $t \in \mathbb{R}$ справедливы оценки

$$|f(t)|^2 \leq 1 - 2a(\delta, \theta)\sigma^2 t^2 + \\ + 2b(\delta, \theta)(\beta_{2+\delta} + \sigma^{2+\delta})|t|^{2+\delta}, \theta_0(\delta) \leq \theta \leq 2\pi; \quad (20)$$

$$|f(t)| \leq 1 - 2a(\delta, \theta)\sigma^2 t^2 + \\ + 2b(\delta, \theta)(\beta_{2+\delta} + \sigma^{2+\delta})|t|^{2+\delta} \leq \quad (21)$$

$$\leq \exp \{ -2a(\delta, \theta)\sigma^2 t^2 + \\ + 2b(\delta, \theta)(\beta_{2+\delta} + \sigma^{2+\delta})|t|^{2+\delta} \}, \theta_0(\delta) \leq \theta \leq 2\pi \quad (22)$$

с $a(\delta, \theta)$, $b(\delta, \theta)$ и $\theta_0(\delta)$, определенными в формулировке леммы 1.

Доказательство. Пусть X_1 и X_2 — независимые случайные величины с одинаковой функцией распределения F . Тогда из леммы 1 получаем оценку

$$|f(t)|^2 = E \cos t(X_1 - X_2) \leq \\ \leq 1 - 2a(\delta, \theta)t^2 E(X_1 - X_2)^2 + \\ + 2b(\delta, \theta)|t|^{2+\delta} E|X_1 - X_2|^{2+\delta}, \\ t \in \mathbb{R}, \quad \theta_0(\delta) \leq \theta \leq 2\pi,$$

откуда в силу леммы 6 вытекает (20).

Для доказательства (21) достаточно применить элементарное неравенство

$$1 + 2x \leq (1 + x)^2, \quad x \in \mathbb{R},$$

к правой части (20) и заметить, что $1 + x \geq 0$ всякий раз, когда $1 + 2x \geq 0$.

Неравенство (22) вытекает из (21) с учетом оценки $1 + x \leq e^x$, справедливой для всех $x \in \mathbb{R}$.

Выбирая в теореме 1 величину θ равной $\theta_0(\delta)$ и замечая, что для всех $0 < \delta \leq 1$

$$a(\delta, \theta_0(\delta)) = \frac{1}{2},$$

$$b(\delta, \theta_0(\delta)) = \sup_{x>0} \frac{|\cos x - 1 + x^2/2|}{x^{2+\delta}} \equiv \varkappa(\delta),$$

получаем следующий результат, приведенный в работах Тысиака [9] и Ушакова [10].

Теорема 2. Пусть $f(t)$ — характеристическая функция некоторого распределения F , удовлетворяющего условиям (1). Тогда для всех $t \in \mathbb{R}$ справедливы оценки

$$|f(t)|^2 \leq 1 - \sigma^2 t^2 + 2\kappa(\delta) (\beta_{2+\delta} + \sigma^{2+\delta}) |t|^{2+\delta},$$

$$|f(t)| \leq 1 - \frac{\sigma^2 t^2}{2} + \kappa(\delta) (\beta_{2+\delta} + \sigma^{2+\delta}) |t|^{2+\delta} \leq \exp \left\{ -\frac{\sigma^2 t^2}{2} + \kappa(\delta) (\beta_{2+\delta} + \sigma^{2+\delta}) |t|^{2+\delta} \right\}.$$

Значения величины $\kappa(\delta)$ для некоторых δ приведены в табл. 1.

Таблица 1 Значения величины $\kappa(\delta)$ при некоторых δ

δ	$\kappa(\delta)$	δ	$\kappa(\delta)$	δ	$\kappa(\delta)$
0+	0,5	0,35	0,2702	0,70	0,1538
0,05	0,4564	0,40	0,2485	0,75	0,1425
0,10	0,4171	0,45	0,2288	0,80	0,1322
0,15	0,3816	0,50	0,2109	0,85	0,1228
0,20	0,3495	0,55	0,1946	0,90	0,1143
0,25	0,3204	0,60	0,1797	0,95	0,1064
0,30	0,2941	0,65	0,1662	1,00	0,0992

Выбирая в теореме 1 величину θ минимизирующей правую часть (20), получаем следующее утверждение.

Теорема 3. *Предположим, что $f(t)$ — характеристическая функция некоторого распределения F , удовлетворяющего условиям (1) с $\sigma^2 = 1$. Пусть $\theta_0(\delta)$ — единственный корень уравнения*

$$\frac{\delta \theta^2}{2} + \theta \sin \theta + (2 + \delta)(\cos \theta - 1) = 0,$$

лежащий в интервале $(\pi, 2\pi)$. Тогда для всех t таких, что

$$\theta_0(\delta) \leq (\beta_{2+\delta} + 1)^{1/\delta} |t| \leq 2\pi,$$

справедливы оценки

$$|f(t)|^2 \leq 1 - 2 \frac{1 - \cos \left((\beta_{2+\delta} + 1)^{1/\delta} t \right)}{(\beta_{2+\delta} + 1)^{2/\delta}}; \quad (23)$$

$$|f(t)| \leq 1 - \frac{1 - \cos \left((\beta_{2+\delta} + 1)^{1/\delta} t \right)}{(\beta_{2+\delta} + 1)^{2/\delta}} \leq \exp \left\{ -\frac{1 - \cos \left((\beta_{2+\delta} + 1)^{1/\delta} t \right)}{(\beta_{2+\delta} + 1)^{2/\delta}} \right\}. \quad (24)$$

Замечание 2. Теорема 3 справедлива для любых $\sigma^2 > 0$, в этом случае оценки становятся более громоздкими. Например, (23) примет вид

$$|f(t)|^2 \leq 1 - 2 \left(\frac{\sigma^2}{\beta_{2+\delta} + \sigma^{2+\delta}} \right)^{2/\delta} \times \left(1 - \cos \left(\left(\frac{\beta_{2+\delta} + \sigma^{2+\delta}}{\sigma^2} \right)^{1/\delta} t \right) \right);$$

$$\theta_0(\delta) \leq |t| \left(\frac{\beta_{2+\delta} + \sigma^{2+\delta}}{\sigma^2} \right)^{1/\delta} \leq 2\pi.$$

Доказательство теоремы 3. Не ограничивая общности, будем считать, что $t > 0$. Действительно, характеристическая функция $|f(t)|^2$ и правая часть (23) являются четными функциями, а при $t = 0$ неравенство (23) очевидно. Обозначим

$$\rho = (\beta_{2+\delta} + 1)^{1/\delta}.$$

Из теоремы 1 имеем

$$|f(t)|^2 \leq 1 - 2at^2 + 2b(\beta_{2+\delta} + 1) |t|^{2+\delta} = 1 + 2t^2 ((\rho t)^\delta b - a), \quad (25)$$

где $a = a(\delta, \theta)$ и $b = b(\delta, \theta)$, $\theta_0(\delta) \leq \theta \leq 2\pi$, определены в формулировке леммы 1. Введем функцию

$$g_t(\theta) = (\rho t)^\delta b(\delta, \theta) - a(\delta, \theta), \quad \theta_0(\delta) \leq \theta \leq 2\pi.$$

Тогда последнее неравенство примет вид

$$|f(t)|^2 \leq 1 + 2t^2 g_t(\theta), \quad \theta_0(\delta) \leq \theta \leq 2\pi.$$

Найдем минимум правой части по θ . Из леммы 5 следует, что $g_t(\theta)$ монотонно возрастает на $\theta_0(\delta) \leq \theta \leq 2\pi$ при $\rho t < \theta$ и монотонно убывает при $\rho t > \theta$, поэтому

$$\inf_{\theta_0(\delta) \leq \theta \leq 2\pi} g_t(\theta) = \begin{cases} g_t(\theta_*), & \rho t < \theta; \\ g_t(\theta^*), & \rho t > \theta, \end{cases}$$

где

$$\theta_* = \max\{\theta_0(\delta), \rho t\}, \quad \theta^* = \min\{2\pi, \rho t\}.$$

Если $\rho t \in [\theta_0(\delta), 2\pi]$, то

$$\inf_{\theta_0(\delta) \leq \theta \leq 2\pi} g_t(\theta) = g_t(\rho t) = (\rho t)^\delta b(\rho t) - a(\rho t),$$

откуда с учетом соотношения

$$\theta^\delta b(\delta, \theta) - a(\delta, \theta) = -\frac{1 - \cos \theta}{\theta^2},$$

$$\theta_0(\delta) \leq \theta \leq 2\pi,$$

вытекающего из определения функций $a(\delta, \theta)$ и $b(\delta, \theta)$, получаем

$$|f(t)|^2 \leq 1 + 2t^2 \inf_{\theta_0(\delta) \leq \theta \leq 2\pi} g_t(\theta) = 1 - 2t^2 \frac{1 - \cos \theta}{\theta^2} \Big|_{\theta=\rho t} = 1 - 2 \frac{1 - \cos \rho t}{\rho^2}.$$

Неравенства (24) доказываются аналогично (21) и (22).

Замечание 3. Детальный анализ проведенного доказательства позволяет заключить, что неравенства из теоремы 3 являются более точными в своих интервалах, чем неравенства из теоремы 2. Другими словами, минимум из трех оценок для $|f(t)|$, устанавливаемых теоремами 2, 3 и неравенством $|f(t)| \leq 1$, имеет следующий вид (для упрощения обозначений полагаем $\sigma^2 = 1$):

$$|f(t)| \leq \exp\{-\psi_\delta(t)\},$$

где

$$\psi_\delta(t) = \begin{cases} \frac{t^2}{2} - \varkappa(\delta) (\beta_{2+\delta} + 1) |t|^{2+\delta}, & |t| \leq \frac{\theta_0(\delta)}{(\beta_{2+\delta} + 1)^{1/\delta}}; \\ \frac{1 - \cos\left((\beta_{2+\delta} + 1)^{1/\delta} t\right)}{(\beta_{2+\delta} + 1)^{2/\delta}}, & \theta_0(\delta) \leq (\beta_{2+\delta} + 1)^{1/\delta} |t| \leq 2\pi; \\ 0, & |t| \geq \frac{2\pi}{(\beta_{2+\delta} + 1)^{1/\delta}}. \end{cases}$$

Значения величины $\theta_0(\delta)$ при некоторых δ приведены в табл. 2.

Таблица 2 Значения величины $\theta_0(\delta)$ при некоторых δ

δ	$\theta_0(\delta)$	δ	$\theta_0(\delta)$	δ	$\theta_0(\delta)$
0+	6,2831	0,35	5,3887	0,70	4,6374
0,05	6,1331	0,40	5,2778	0,75	4,5320
0,10	5,9941	0,45	5,1686	0,80	4,4263
0,15	5,8631	0,50	5,0609	0,85	4,3200
0,20	5,7384	0,55	4,9542	0,90	4,2131
0,25	5,6183	0,60	4,8483	0,95	4,1051
0,30	5,5021	0,65	4,7427	1,00	3,9959

Как несложно убедиться,

$$\lim_{\delta \rightarrow 0+} \theta_0(\delta) = 2\pi = 6,2831 \dots$$

Замечание 4. Если есть необходимость использования оценок вида (23) и (24) для

$$|t| < \theta_0(\delta) (\beta_{2+\delta} + 1)^{-1/\delta},$$

то можно заменить неравенство (25) на более грубое

$$|f(t)|^2 \leq 1 - 2at^2 + 2b(\beta_{2+\delta} + 1)\alpha^\delta |t|^{2+\delta}, \quad \alpha \geq 1,$$

и получить оценки

$$|f(t)|^2 \leq 1 - 2 \frac{1 - \cos\left((\beta_{2+\delta} + 1)^{1/\delta} \alpha t\right)}{(\beta_{2+\delta} + 1)^{2/\delta} \alpha^2};$$

$$|f(t)| \leq 1 - \frac{1 - \cos\left((\beta_{2+\delta} + 1)^{1/\delta} \alpha t\right)}{(\beta_{2+\delta} + 1)^{2/\delta} \alpha^2} \leq \exp\left\{-\frac{1 - \cos\left((\beta_{2+\delta} + 1)^{1/\delta} \alpha t\right)}{(\beta_{2+\delta} + 1)^{2/\delta} \alpha^2}\right\},$$

справедливые для

$$\theta_0(\delta) \leq (\beta_{2+\delta} + 1)^{1/\delta} \alpha |t| \leq 2\pi.$$

4 Уточнение неравенства Мизеса

Пусть случайная величина X имеет решетчатое распределение с шагом h . Обозначим

$$\beta_s = E|X - EX|^s, \quad s > 1, \quad \sigma^2 = \beta_2.$$

В 1939 г. Р. фон Мизес [11] получил следующий результат.

Теорема 4. В сделанных выше предположениях для целых $s \geq 2$ имеют место неравенства

$$\beta_{s-1} \leq \frac{2}{h} \beta_s.$$

Из этой теоремы при $s = 3$ получаем

$$\frac{h}{\sigma} \leq 2 \frac{\beta_3}{\sigma^3}.$$

Следующее утверждение, уточняющее неравенство Мизеса для $s = 3$, является простым следствием теоремы 3.

Теорема 5. В сделанных выше предположениях для $s = 2 + \delta$ с $0 < \delta \leq 1$ справедлива оценка

$$\frac{h}{\sigma} \leq \left(\frac{\beta_{2+\delta}}{\sigma^{2+\delta}} + 1\right)^{1/\delta}, \quad (26)$$

причем

$$\sup \left(\frac{h}{\sigma} - \left(\frac{\beta_{2+\delta}}{\sigma^{2+\delta}} + 1\right)^{1/\delta}\right) = 0, \quad (27)$$

где супремум берется по всем распределениям, удовлетворяющим условиям теоремы.

Доказательство. Без ограничения общности будем считать, что $EX = 0$ и $\sigma^2 = EX^2 = 1$. Пусть $f(t)$ — характеристическая функция случайной величины X . Как известно, характеристическая

функция $f(t)$ задает решетчатое распределение тогда и только тогда, когда существует такое $t_0 \neq 0$, что $|f(t_0)| = 1$. При этом в качестве шага распределения h можно взять

$$h = \frac{2\pi}{t_0}.$$

С другой стороны, из замечания 3 к теореме 3 вытекает, что $|f(t)| < 1$ для всех t таких, что

$$|t| < \frac{2\pi}{(\beta_{2+\delta} + 1)^{1/\delta}};$$

следовательно,

$$t_0 \geq \frac{2\pi}{(\beta_{2+\delta} + 1)^{1/\delta}},$$

откуда с учетом соотношения $h = 2\pi/t_0$ получаем (26).

Для доказательства второй части теоремы рассмотрим случайную величину X , имеющую следующее распределение:

$$\begin{aligned} P\left(X = \frac{h}{1+u}\right) &= \frac{u}{1+u}; \\ P\left(X = -\frac{uh}{1+u}\right) &= \frac{1}{1+u}, \quad u, h > 0. \end{aligned}$$

Шаг этого распределения равняется h , $EX = 0$,

$$\begin{aligned} \sigma^2 &= EX^2 = \frac{uh^2}{(1+u)^2}, \\ \beta_{2+\delta} &= E|X|^{2+\delta} = \frac{u(1+u^{1+\delta})h^{2+\delta}}{(1+u)^{3+\delta}}, \\ \frac{\beta_{2+\delta}}{\sigma^2 h^\delta} + \left(\frac{\sigma}{h}\right)^\delta &= \frac{1+u^{1+\delta}}{(1+u)^{1+\delta}} + \frac{u^{\delta/2}}{(1+u)^\delta} \equiv g(u, \delta). \end{aligned}$$

Тогда

$$\begin{aligned} \inf \left(\frac{\beta_{2+\delta}}{\sigma^2 h^\delta} + \left(\frac{\sigma}{h}\right)^\delta \right) &\leq \inf_{u>0} g(u, \delta) \leq \\ &\leq \lim_{u \rightarrow \infty} g(u, \delta) = 1, \quad 0 < \delta \leq 1, \end{aligned}$$

что равносильно (27). Дополнительно заметим, что при $\delta = 1$ инфимум функции $g(u, \delta)$ также достигается в точке $u = 1$, что соответствует симметричному распределению Бернулли, для которого $\beta_3/\sigma^3 = 1$.

В заключение автор выражает благодарность В. Ю. Королёву за идею использования теоремы 3 в целях уточнения неравенства Мизеса и за постоянное внимание к работе.

Литература

1. *Золотарёв В. М.* О близости распределений двух сумм независимых случайных величин // Теория вероятн. и ее примен., 1965. Т. 10. Вып. 3. С. 519–526.
2. *Золотарёв В. М.* Некоторые неравенства теории вероятностей и их применение к уточнению теоремы А. М. Ляпунова // ДАН СССР, 1967. Т. 177. № 3. С. 501–504.
3. *Prawitz H.* Ungleichungen für den absoluten Betrag einer charakteristischen funktion // Skand. Aktuarietidskr., 1973. No. 1. P. 11–16.
4. *Berry A. C.* The accuracy of the Gaussian approximation to the sum of independent variates // Trans. Amer. Math. Soc., 1941. Vol. 49. P. 122–139.
5. *Esseen C.-G.* On the Liapunoff limit of error in the theory of probability // Ark. Mat. Astron. Fys., 1942. Vol. A28. No. 9. P. 1–19.
6. *Золотарёв В. М.* Абсолютная оценка остаточного члена в центральной предельной теореме // Теория вероятн. и ее примен., 1966. Т. 11. Вып. 1. С. 108–119.
7. *Zolotarev V. M.* A sharpening of the inequality of Berry–Esseen // Wahrsch. verw. Geb., 1967. Bd. 8. S. 332–342.
8. *Prawitz H.* On the remainder in the central limit theorem // Scand. Actuarial J., 1975. No. 3. P. 145–156.
9. *Tysiak W.* Gleichmäßige und nicht-gleichmäßige Berry–Esseen–Abschätzungen. Dissertation, Wuppertal, 1983.
10. *Ushakov N. G.* Selected topics in characteristic functions. — Utrecht: VSP, 1999.
11. *Von Mises R.* An inequality for the moments of a discontinuous distribution // Skand. Aktuarietidskr., 1939. Vol. 22. No. 1. P. 32–36.

О МОЩНОСТИ КРИТЕРИЕВ В СЛУЧАЕ ОБОБЩЕННОГО РАСПРЕДЕЛЕНИЯ ЛАПЛАСА*

В. Е. Бенинг¹, О. О. Лямин²

Аннотация: В работе на эвристическом уровне получена формула для предела отклонения мощности асимптотически оптимального критерия от мощности наилучшего критерия в случае обобщенного распределения Лапласа. Это отклонение в силу нерегулярности этого распределения имеет порядок $n^{-1/2}$, в отличие от обычных регулярных семейств, для которых этот порядок равен n^{-1} .

Ключевые слова: обобщенное распределение Лапласа; функция мощности; дефект; асимптотическое разложение

1 Введение

Хорошо известно, что распределение Лапласа находит широкое применение при математическом моделировании многих процессов в телекоммуникационных системах, в экономике, финансовом деле, технике и других областях, например в задачах выделения полезного сигнала на фоне помех (см., например, работы [1–4]). Естественность возникновения этого распределения обоснована в работе [5]. Назовем обобщенным распределением Лапласа распределение на действительной прямой с плотностью вида

$$p(x, \theta) = C(a, b) e^{-a(x-\theta)^2 - b|x-\theta|},$$

$$a \geq 0, b > 0, x \in \mathbb{R}^1, \quad (1)$$

где $C(a, b)$ — нормировочная константа такая, что

$$C(a, b) = \begin{cases} \frac{b}{2}, & a = 0; \\ \frac{\sqrt{a}}{\sqrt{\pi} \exp\{b^2/4a\} \operatorname{erfc}(b/(2\sqrt{a}))}, & a > 0 \end{cases}$$

и

$$\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-y^2} dy.$$

Обычное распределение Лапласа получается при $a = 0$. Обобщенное распределение Лапласа может оказаться полезным в тех случаях, когда необходим более тонкий контроль за поведением функции плотности, чем может быть предоставлен однопараметрическими по параметру формы нормальным

и лапласовским распределениями. Так, правильно подобрав параметры, можно получить распределение с хвостами менее тяжелыми, чем у соответствующего распределения Лапласа, при этом сохранив существенную особенность негладкости функции плотности, которая отсутствует у нормального распределения.

Рассмотрим задачу проверки простой гипотезы

$$H_0 : \theta = 0$$

против последовательности близких сложных альтернатив вида

$$H_{n,1} : \theta = \frac{t}{\sqrt{n}}, \quad 0 < t \leq C, \quad C > 0,$$

на основе выборки (X_1, \dots, X_n) — независимых одинаково распределенных наблюдений, имеющих распределение с плотностью вида (1). Для любого фиксированного $t \in (0, C]$ наилучший критерий всегда существует согласно фундаментальной лемме Неймана–Пирсона и основан на логарифме отношения правдоподобия $\Lambda_n(t)$

$$\Lambda_n(\theta) = \sum_{i=1}^n (l(X_i, \theta) - l(X_i, 0)); \quad (2)$$

$$\Lambda_n(t) \equiv \Lambda_n(tn^{-1/2}),$$

где $l(x, \theta) = \log p(x, \theta)$, $\theta > 0$.

Обозначим через $\beta_n^*(t)$ мощность такого критерия уровня $\alpha \in (0, 1)$. Заметим, что, поскольку t

*Работа выполнена при финансовой поддержке РФФИ, проекты 08-01-00567 и 08-07-00152.

¹Факультет вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова, bening@yandex.ru

²Факультет вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова

неизвестно, мы не можем использовать статистику $\Lambda_n(t)$ для построения критерия проверки гипотезы H_0 против альтернативы $H_{n,1}$. Однако $\beta_n^*(t)$ дает верхнюю границу для мощности любого критерия при проверке H_0 против фиксированной альтернативы

$$H_{n,t} : \theta = \frac{t}{\sqrt{n}}, \quad t > 0,$$

и может служить стандартом при сравнении различных критериев.

В работе [6] было показано, что в случае регулярного семейства плотностей справедливо соотношение

$$\beta_n^*(t) \rightarrow \beta^*(t) = \Phi(t\sqrt{I} - u_\alpha),$$

где $\Phi(x)$ — функция распределения стандартного нормального закона и $\Phi(u_\alpha) = 1 - \alpha$, $\alpha \in (0, 1)$, I — фишеровская информация. Там же было показано, что для проверки H_0 против $H_{n,1}$ существуют критерии, основанные на статистиках, отличных от $\Lambda_n(t)$, и имеющие ту же предельную мощность $\beta^*(t)$. Такие критерии называются асимптотически наиболее мощными (АНМ). Среди таких критериев существуют критерии, основанные на статистиках, которые не зависят от неизвестного параметра t , и поэтому могут быть использованы при проверке гипотезы H_0 против сложной альтернативы $H_{n,1}$. При этом мощности таких критериев типичным образом отличаются от $\beta_n^*(t)$ на величину порядка n^{-1} [6].

В этой статье для распределения (1) рассмотрим АНМ критерий, основанный на статистике вида

$$T_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n [2aX_i + b \operatorname{sign}(X_i)].$$

Обозначим через $\beta_n(t)$ мощность этого критерия уровня $\alpha \in (0, 1)$, через $I_{a,b}$ — фишеровскую информацию распределения (1) и пусть $\varphi(x)$ — плотность стандартного нормального распределения.

Заметим, что семейство (1) не является регулярным, поскольку у $p(x, \theta)$ не существует производной по θ в точке $\theta = x$. Покажем, что это отсутствие регулярности выражается в нарушении естественного порядка n^{-1} разности $\beta_n^*(t) - \beta_n(t)$ и приводит к порядку $n^{-1/2}$.

Далее всюду предполагается $a > 0$. Случай $a = 0$ подробно рассмотрен в работе [7]. Цель настоящей статьи получить результаты, аналогичные случаю распределения Лапласа (см. [7]), для случая обобщенного распределения Лапласа (1).

2 Асимптотическое поведение логарифма отношения правдоподобия

В этом разделе доказываются результаты, описывающие асимптотическое ($n \rightarrow \infty$) поведение логарифма отношения правдоподобия $\Lambda_n(t)$ как при гипотезе H_0 , так и при альтернативе $H_{n,t}$.

Лемма 2.1. В случае распределения (1) справедливы следующие соотношения

$$\begin{aligned} I_{a,b} &= 2(bC(a, b) + a); \\ \mathcal{L}(\Lambda_n(t) | H_0) &\rightarrow \mathcal{N}\left(-\frac{1}{2}t^2 I_{a,b}, t^2 I_{a,b}\right); \\ \mathcal{L}(\Lambda_n(t) | H_{n,t}) &\rightarrow \mathcal{N}\left(\frac{1}{2}t^2 I_{a,b}, t^2 I_{a,b}\right), \quad n \rightarrow \infty. \end{aligned}$$

Доказательство. В силу свойств интеграла Лебега имеем

$$\begin{aligned} I_{a,b} &= \int_{\mathbb{R}^1 \setminus \{0\}} p(x) \left(\frac{p'(x)}{p(x)}\right)^2 dx = \\ &= 2C(a, b) \int_0^\infty (2ax + b)^2 e^{-ax^2 - bx} dx = \\ &= 2(bC(a, b) + a). \end{aligned}$$

Отсюда следует, что

$$C(a, b) = \frac{I_{a,b} - 2a}{2b}. \quad (3)$$

Рассмотрим логарифм отношения правдоподобия $\Lambda_n(\theta)$, $\theta > 0$ более подробно. Из определения (2) следует, что

$$\begin{aligned} \Lambda_n(\theta) &= \sum_{i=1}^n (l(X_i, \theta) - l(X_i, 0)) = \\ &= \sum_{i=1}^n [a\theta(2X_i - \theta) + b(|X_i| - |X_i - \theta|)] \equiv \\ &\equiv \sum_{i=1}^n Z_i(\theta), \quad (4) \end{aligned}$$

где

$$\begin{aligned} Z_i(\theta) &= \\ &= \begin{cases} a\theta(2X_i - \theta) + b\theta, & X_i > \theta; \\ a\theta(2X_i - \theta) + b(2X_i - \theta) = \\ = a\theta(2X_i - \theta) + 2bX_i \mathbf{1}_{[0, \theta]}(X_i) + \\ + 2b\theta \mathbf{1}_{(\theta, \infty)}(X_i) - b\theta, & 0 \leq X_i \leq \theta; \\ a\theta(2X_i - \theta) - b\theta, & X_i < 0. \end{cases} \quad (5) \end{aligned}$$

Найдем характеристические функции случайной величины $Z_1(\theta)$ при распределениях P_0 и P_θ . Для распределения P_0 имеем

$$\begin{aligned}
 f_0(s) &\equiv E_0 e^{isZ_1(\theta)} = \\
 &= \int_{-\infty}^0 e^{is(a\theta(2x-\theta)-b\theta)} C(a,b) e^{-ax^2+bx} dx + \\
 &+ \int_0^\theta e^{is(a\theta(2x-\theta)+b(2x-\theta))} C(a,b) e^{-ax^2-bx} dx + \\
 &+ \int_\theta^\infty e^{is(a\theta(2x-\theta)+b\theta)} C(a,b) e^{-ax^2-bx} dx = \\
 &= C(a,b) e^{-is\theta(a\theta+b)} \int_{-\infty}^0 e^{-ax^2+(2isa\theta+b)x} dx + \\
 &+ C(a,b) e^{-is\theta(a\theta+b)} \int_0^\theta e^{-ax^2+(2is(a\theta+b)-b)x} dx + \\
 &+ C(a,b) e^{-is\theta(a\theta-b)} \int_\theta^\infty e^{-ax^2+(2isa\theta-b)x} dx. \quad (6)
 \end{aligned}$$

Далее нам понадобятся следующие соотношения:

$$\begin{aligned}
 \int_{-\infty}^0 e^{-ax^2+(\omega+b)x} dx &= \\
 &= \frac{\sqrt{\pi} e^{(\omega+b)^2/(4a)} \operatorname{erfc}((\omega+b)/(2\sqrt{a}))}{2\sqrt{a}}; \quad (7)
 \end{aligned}$$

$$\begin{aligned}
 \int_0^\theta e^{-ax^2+(\delta-b)x} dx &= \\
 &= \sqrt{\pi} e^{(\delta-b)^2/(4a)} \left[\operatorname{erfc}\left(\frac{(\delta-b)-2a\theta}{2\sqrt{a}}\right) - \right. \\
 &\quad \left. - \operatorname{erfc}\left(\frac{\delta-b}{2\sqrt{a}}\right) \right] / (2\sqrt{a}), \quad (8)
 \end{aligned}$$

$$\begin{aligned}
 \int_\theta^\infty e^{-ax^2+(\sigma-b)x} dx &= \\
 &= \frac{\sqrt{\pi} e^{(\sigma-b)^2/(4a)} \operatorname{erfc}((2a\theta - (\sigma-b))/(2\sqrt{a}))}{2\sqrt{a}}, \quad (9)
 \end{aligned}$$

где ω, δ, σ — некоторые константы (в том числе комплексные), а также равенства

$$\left. \begin{aligned}
 \operatorname{erfc}'(x) &= -\frac{2e^{-x^2}}{\sqrt{\pi}}; \\
 \operatorname{erfc}''(x) &= \frac{4xe^{-x^2}}{\sqrt{\pi}}; \\
 \operatorname{erfc}'''(x) &= \frac{4xe^{-x^2}(1-2x^2)}{\sqrt{\pi}}.
 \end{aligned} \right\} \quad (10)$$

Из соотношений (6)–(9) и (10) с $\omega = 2isa\theta$, $\delta = 2is(a\theta+b)$, $\sigma = 2isa\theta$ получаем, что в случае близких альтернатив $\theta = \theta_n \rightarrow 0$ при каждом фиксированном s справедливы равенства

$$\begin{aligned}
 f_0(s) &= 1 + \frac{is(is-1)I_{a,b}}{2} \theta_n^2 + o(\theta_n^2); \\
 E_0 Z_1(\theta_n) &= -\frac{I_{a,b}}{2} \theta_n^2 + o(\theta_n^2); \\
 D_0 Z_1(\theta_n) &= I_{a,b} \theta_n^2 + o(\theta_n^2).
 \end{aligned}$$

Аналогично находя характеристическую функцию случайной величины $Z_1(\theta)$ в случае распределения P_θ , получим, что в случае близких альтернатив $\theta = \theta_n \rightarrow 0$ при каждом фиксированном s справедливы равенства

$$\begin{aligned}
 f_{\theta_n}(s) &= 1 + \frac{is(is+1)I_{a,b}}{2} \theta_n^2 + o(\theta_n^2); \\
 E_{\theta_n} Z_1(\theta_n) &= \frac{I_{a,b}}{2} \theta_n^2 + o(\theta_n^2); \\
 D_{\theta_n} Z_1(\theta_n) &= I_{a,b} \theta_n^2 + o(\theta_n^2).
 \end{aligned}$$

Теперь, если $\theta_n = t/\sqrt{n}$, то для любого фиксированного s имеем

$$\begin{aligned}
 E_{n,0} e^{is\Lambda_n(t)} = f_0^n(s) &= \left(1 + \frac{is(is-1)t^2 I_{a,b}}{2n} + \right. \\
 &\quad \left. + o\left(\frac{1}{n}\right) \right)^n \rightarrow e^{-s^2 t^2 I_{a,b}/2 - ist^2 I_{a,b}/2}, \quad n \rightarrow \infty.
 \end{aligned}$$

Аналогично

$$\begin{aligned}
 E_{n,t/\sqrt{n}} e^{is\Lambda_n(t)} = f_{t/\sqrt{n}}^n(s) &\rightarrow \\
 &\rightarrow e^{-s^2 t^2 I_{a,b}/2 + ist^2 I_{a,b}/2}, \quad n \rightarrow \infty.
 \end{aligned}$$

Отсюда и из теоремы непрерывности (см., например, [8, теорема 7.3.2]) следует утверждение леммы. \square

Из этой леммы и работы [7] (см. введение) непосредственно получаем

Следствие 2.1. В случае распределения (1) справедливо соотношение

$$\beta_n^*(t) \rightarrow \beta^*(t) = \Phi\left(t\sqrt{I_{a,b}} - u_\alpha\right), \quad n \rightarrow \infty.$$

3 Формула для предельного отклонения мощностей

В этом разделе на эвристическом уровне будет показано, что справедлива формула для предельного отклонения разностей мощностей

$$r(t) = \lim_{n \rightarrow \infty} \sqrt{n} (\beta_n^*(t) - \beta_n(t)) = \frac{(I_{a,b} - 2a)bt^2}{3\sqrt{I_{a,b}}} \varphi(t\sqrt{I_{a,b}} - u_\alpha), \quad (11)$$

где $\beta_n(t)$ — функция мощности АНМ критерия, основанного на знаковой статистике

$$T_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n [2aX_i + b \text{sign}(X_i)]. \quad (12)$$

Формула (11) показывает, что отсутствие регулярности у распределения (1) приводит к нарушению естественного порядка разности $\beta_n^*(t) - \beta_n(t)$, равного n^{-1} (см. [6]). Из формулы (11) следует, что этот порядок равен $n^{-1/2}$.

Получим сначала стохастическое разложение для $\Lambda_n(t)$. Имеем (см. (4), (5)),

$$\begin{aligned} \Lambda_n(t) &= a \frac{t}{\sqrt{n}} \sum_{i=1}^n \left(2X_i - \frac{t}{\sqrt{n}} \right) + \\ &\quad + 2b \sum_{i=1}^n X_i \mathbf{1}_{[0, t/\sqrt{n}]}(X_i) + \\ &\quad + 2b \frac{t}{\sqrt{n}} \sum_{i=1}^n \mathbf{1}_{(t/\sqrt{n}, \infty)}(X_i) - b \frac{t}{\sqrt{n}} = \\ &= 2a \frac{t}{\sqrt{n}} \sum_{i=1}^n X_i - at^2 + 2b \sum_{i=1}^n X_i \mathbf{1}_{[0, t/\sqrt{n}]}(X_i) + \\ &+ 2b \frac{t}{\sqrt{n}} \sum_{i=1}^n \mathbf{1}_{(0, \infty)}(X_i) - 2b \frac{t}{\sqrt{n}} \sum_{i=1}^n \mathbf{1}_{[0, t/\sqrt{n}]}(X_i) - \\ &- b \frac{t}{\sqrt{n}} \sum_{i=1}^n \mathbf{1}_{(0, \infty)}(X_i) - b \frac{t}{\sqrt{n}} \sum_{i=1}^n \mathbf{1}_{(-\infty, 0]}(X_i) = \\ &= 2a \frac{t}{\sqrt{n}} \sum_{i=1}^n X_i - at^2 + \\ &+ 2b \sum_{i=1}^n \left(X_i - \frac{t}{\sqrt{n}} \right) \mathbf{1}_{[0, t/\sqrt{n}]}(X_i) + \\ &+ b \frac{t}{\sqrt{n}} \sum_{i=1}^n \text{sign}(X_i) + b \frac{t}{\sqrt{n}} \sum_{i=1}^n \mathbf{1}_{\{0\}}(X_i). \end{aligned}$$

Поскольку распределение X_i непрерывно, то

$$P_{n,\theta} \left(\sum_{i=1}^n \mathbf{1}_{\{0\}}(X_i) > 0 \right) = 0, \quad \theta > 0,$$

и поэтому почти всюду справедливо представление

$$\begin{aligned} \Lambda_n(t) &= 2a \frac{t}{\sqrt{n}} \sum_{i=1}^n X_i - at^2 + \\ &+ 2b \sum_{i=1}^n \left(X_i - \frac{t}{\sqrt{n}} \right) \mathbf{1}_{[0, t/\sqrt{n}]}(X_i) + \\ &+ b \frac{t}{\sqrt{n}} \sum_{i=1}^n \text{sign}(X_i) = \\ &= \frac{t}{\sqrt{n}} \sum_{i=1}^n [2aX_i + b \text{sign}(X_i)] - at^2 + \\ &+ 2b \sum_{i=1}^n \left(X_i - \frac{t}{\sqrt{n}} \right) \mathbf{1}_{[0, t/\sqrt{n}]}(X_i). \quad (13) \end{aligned}$$

Рассмотрим следующее монотонное (не меняющее мощность критерия) преобразование ($t > 0$) статистики T_n (см. (12))

$$S_n(t) = tT_n - \frac{1}{2}I_{a,b}t^2.$$

Тогда

$$\begin{aligned} \Delta_n(t) \equiv S_n(t) - \Lambda_n(t) &= at^2 - \frac{1}{2}I_{a,b}t^2 - \\ &- 2b \sum_{i=1}^n \left(X_i - \frac{t}{\sqrt{n}} \right) \mathbf{1}_{[0, t/\sqrt{n}]}(X_i). \quad (14) \end{aligned}$$

Лемма 3.1. В случае распределения (1) справедливы следующие соотношения:

$$\mathcal{L}(\sqrt[4]{n}\Delta_n(t) | H_0) \rightarrow \mathcal{N}\left(0, \frac{2}{3}b(I_{a,b} - 2a)t^3\right);$$

$$\begin{aligned} \mathcal{L}(\sqrt[4]{n}\Delta_n(t), \Lambda_n(t) | H_0) &\rightarrow \\ \rightarrow \mathcal{N}_2\left(0, \frac{2}{3}b(I_{a,b} - 2a)t^3, 0, -\frac{1}{2}I_{a,b}t^2, I_{a,b}t^2\right), \end{aligned}$$

где \mathcal{N}_2 — двумерный нормальный закон с соответствующими параметрами.

Доказательство. Докажем первое соотношение методом характеристических функций. С этой целью найдем характеристическую функцию случайной величины

$$(X_1 - \theta) \mathbf{1}_{[0, \theta]}(X_1), \quad \theta > 0.$$

Имеем

$$\begin{aligned} g_\theta(s) \equiv E_0 e^{is(X_1 - \theta)\mathbf{1}_{[0, \theta]}(X_1)} &= P_0(X_1 < 0) + \\ &+ C(a, b) e^{-is\theta} \int_0^\theta e^{-ax^2 + (is-b)x} dx + \end{aligned}$$

$$\begin{aligned}
 &+ P_0(X_1 > \theta) = \\
 &= \frac{1}{2} + C(a, b) e^{-is\theta} \int_0^\theta e^{-ax^2 + (is-b)x} dx + \\
 &\quad + C(a, b) \int_\theta^\infty e^{-ax^2 - bx} dx.
 \end{aligned}$$

Используя соотношения (8) с $\delta = is$, (9) с $\sigma = 0$ и (10), получим, что если $\theta = \theta_n \rightarrow 0$, то при каждом фиксированном s справедливо представление

$$\begin{aligned}
 g_{\theta_n}(s) &= 1 - is \frac{C(a, b)}{2} \theta_n^2 + \\
 &\quad + is(is + b) \frac{C(a, b)}{6} \theta_n^3 + o(\theta_n^3), \quad (15)
 \end{aligned}$$

поэтому с учетом (3) имеем

$$\begin{aligned}
 E_{n,0} e^{is \sqrt[4]{n} \Delta_n(t)} &= e^{is \sqrt[4]{nt^2(a - I_{a,b}/2)} g_{t/\sqrt{n}}^n(-2b \sqrt[4]{ns}) = \\
 &= \exp \left[-is \sqrt[4]{nt^2} b C(a, b) + n \log \left(1 + is \frac{b C(a, b) t^2}{n^{3/4}} + \right. \right. \\
 &\quad \left. \left. + \frac{b^2 (-4\sqrt{ns}^2 - 2\sqrt[4]{n} is) C(a, b) t^3}{6n^{3/2}} + o\left(\frac{1}{n}\right) \right) \right] = \\
 &= \exp \left(-\frac{2b^2 C(a, b) t^3 s^2}{3} + o\left(\frac{1}{\sqrt[4]{n}}\right) \right) \rightarrow \\
 &\quad \rightarrow e^{-2b^2 C(a, b) t^3 s^2 / 3}. \quad (16)
 \end{aligned}$$

Отсюда следует первое утверждение леммы.

Докажем второе утверждение леммы также методом характеристических функций, хотя его можно доказать и с помощью двумерной центральной предельной теоремы для схемы серий (см., например, [8], теорема 8.7.11). С этой целью найдем двумерную характеристическую функцию случайных величин (см. (13), (14))

$$((X_1 - \theta) \mathbf{1}_{[0, \theta]}(X_1), Z_1(\theta)), \quad \theta > 0;$$

$$\begin{aligned}
 g_\theta(u, v) &\equiv \\
 &\equiv E_0 \exp(iu(X_1 - \theta) \mathbf{1}_{[0, \theta]}(X_1) + iv Z_1(\theta)) = \\
 &= C(a, b) e^{-iv\theta(a\theta+b)} \int_{-\infty}^0 e^{-ax^2 + (2iva\theta+b)x} dx + \\
 &\quad + C(a, b) e^{-iu\theta - iv\theta(a\theta+b)} \times \\
 &\quad \times \int_0^\theta e^{-ax^2 + (iu + 2iv(a\theta+b) - b)x} dx + \\
 &\quad + C(a, b) e^{-iv\theta(a\theta-b)} \int_0^\infty e^{-ax^2 + (2iva\theta - b)x} dx.
 \end{aligned}$$

Используя опять формулы (7)–(9) с соответствующими значениями параметров, а также соотношение (10), получим, что при $\theta = \theta_n \rightarrow 0$ и фиксированных u и v имеем асимптотическое представление:

$$\begin{aligned}
 g_{\theta_n}(u, v) &= 1 + \left(aiv(iv - 1) + \right. \\
 &\quad \left. + \frac{C(a, b)(-iu + 2biv(iv - 1))}{2} \right) \theta_n^2 + \\
 &\quad + \frac{C(a, b)(ibu - u^2 + 3b^2v(i + v))}{6} \theta_n^3 + o(\theta_n^3) = \\
 &= 1 + iv(iv - 1)(bC(a, b) + a) \theta_n^2 - \\
 &\quad - \frac{iuC(a, b)}{2} \theta_n^2 + \frac{ibuC(a, b)}{6} \theta_n^3 - \frac{u^2C(a, b)}{6} \theta_n^3 - \\
 &\quad - \frac{iv(iv - 1)b^2C(a, b)}{2} \theta_n^3 + o(\theta_n^3) = \\
 &= 1 + iv(iv - 1) \frac{I_{a,b}}{2} \theta_n^2 - \frac{iuC(a, b)}{2} \theta_n^2 + \\
 &\quad + \frac{ibuC(a, b)}{6} \theta_n^3 - \frac{u^2C(a, b)}{6} \theta_n^3 - \\
 &\quad - \frac{iv(iv - 1)b^2C(a, b)}{2} \theta_n^3 + o(\theta_n^3).
 \end{aligned}$$

Теперь (см. (13), (14))

$$\begin{aligned}
 E_{n,0} \exp(iu \sqrt[4]{n} \Delta_n(t) + iv \Lambda_n(t)) &= \\
 &= e^{iu \sqrt[4]{nt^2(a - I_{a,b}/2)} g_{t/\sqrt{n}}^n(-2b \sqrt[4]{nu}, v) = \\
 &= \exp \left(\frac{I_{a,b} t^2}{2} (-iv - v^2) - \frac{2b^2 C(a, b) t^3 u^2}{3} + \right. \\
 &\quad \left. + o\left(\frac{1}{\sqrt[4]{n}}\right) \right) \rightarrow \exp \left(-iv \frac{I_{a,b} t^2}{2} - v^2 \frac{I_{a,b} t^2}{2} - \right. \\
 &\quad \left. - u^2 \frac{2b^2 C(a, b) t^3}{3} \right).
 \end{aligned}$$

Теперь из (3) и многомерной теоремы непрерывности (см. [8], теорема 7.6.2А) следует утверждение леммы. \square

Из этой леммы следует, что случайные величины $\sqrt[4]{n} \Delta_n(t)$ и $\Lambda_n(t)$ асимптотически независимы, и поэтому формула для $r(t)$ приобретает вид (см. (11) и работы [6, 9, 7]):

$$\begin{aligned}
 r(t) &\equiv \lim_{n \rightarrow \infty} \sqrt{n} (\beta_n^*(t) - \beta_n(t)) = \\
 &= \frac{1}{2t \sqrt{I_{a,b}}} \varphi(t \sqrt{I_{a,b}} - u_\alpha) D(\Delta(t) | \Lambda(t) = c_t) = \\
 &= \frac{1}{2t \sqrt{I_{a,b}}} \varphi(t \sqrt{I_{a,b}} - u_\alpha) D\Delta(t) = \\
 &= \frac{(I_{a,b} - 2a)bt^2}{3\sqrt{I_{a,b}}} \varphi(t \sqrt{I_{a,b}} - u_\alpha), \quad (17)
 \end{aligned}$$

где $\Lambda(t)$, $\Delta(t)$ — независимые нормальные случайные величины, распределенные соответственно с параметрами $(-(1/2)I_{a,b}t^2, I_{a,b}t^2)$, $(0, (2/3)(I_{a,b} - 2a)bt^3)$ и $c_t = t\sqrt{I_{a,b}u_\alpha} - (1/2)t^2I_{a,b}$.

Формула (17) была получена с помощью общей теоремы 3.2.1 из работы [6] (см. также [7], введение). Формальное доказательство полученной формулы, состоящее в проверке условий этой теоремы, не является целью этой статьи и будет приведено в другой статье. В этой работе будет доказано более слабое утверждение, составляющее содержание следующей леммы.

Лемма 3.2. *В случае распределения (1) для любого $0 \leq \delta < 1/2$ справедливо соотношение*

$$n^\delta (\beta_n^*(t) - \beta_n(t)) \rightarrow 0, \quad n \rightarrow \infty, \quad 0 < t \leq C, \quad C > 0.$$

Доказательство. Для доказательства применим общую теорему из работы [10]. Условия этой теоремы в нашем случае сводятся к проверке следующих соотношений:

существует константа $A > 0$ такая, что для любого $x_0 \in \mathbb{R}^1$, любого $\gamma > 0$ и любого $t \in (0, C]$, $C > 0$ справедливы равенства

$$\sup_{x \leq x_0} P_{n,0} \left(x - n^{-\delta/2} \leq \Lambda_n(t) \leq x \right) = O(n^{-\delta/2});$$

$$E_{n,0} |\Delta_n(t)| \mathbf{1}_{(\gamma n^{-\delta/2}, A)}(|\Delta_n(t)|) = o(n^{-\delta}); \quad (18)$$

$$P_{n,0} (\Delta_n(t) \geq A) = o(n^{-\delta}); \quad (19)$$

$$P_{n,tn^{-1/2}} (\Delta_n(t) \leq -A) = o(n^{-\delta}). \quad (20)$$

Найдем сначала выражения для $E_{n,0}\Delta_n(t)$ и $E_{n,0}\Delta_n^2(t)$. Из соотношения (14) непосредственно получаем

$$E_{n,0}\Delta_n(t) = -bC(a, b)t^2 - 2bnE_0 \left(X_1 - tn^{-1/2} \right) \mathbf{1}_{[0,tn^{-1/2}]}(X_1); \quad (21)$$

$$E_{n,0}\Delta_n^2(t) = D_{n,0}\Delta_n(t) + (E_{n,0}\Delta_n(t))^2 = 4b^2nD_0(X_1 - tn^{-1/2})\mathbf{1}_{[0,tn^{-1/2}]}(X_1) + (E_{n,0}\Delta_n(t))^2. \quad (22)$$

Найдем теперь математическое ожидание и дисперсию случайной величины $(X_1 - \theta)\mathbf{1}_{[0,\theta]}(X_1)$. Используя выражение для характеристической функции (см. (21)) $g_\theta(s)$ и ее разложение при $\theta = \theta_n \rightarrow 0$ (см. (15)), после дифференцирования функции $g_{\theta_n}(s)$ по s получим

$$E_0(X_1 - \theta_n)\mathbf{1}_{[0,\theta_n]}(X_1) = \frac{g_{\theta_n}^{(1)}(0)}{i} = -\frac{C(a, b)}{2}\theta_n^2 + \frac{bC(a, b)}{6}\theta_n^3 + O(\theta_n^4); \quad (23)$$

$$E_0(X_1 - \theta_n)^2\mathbf{1}_{[0,\theta_n]}(X_1) = -g_{\theta_n}^{(2)}(0) = \frac{C(a, b)}{3}\theta_n^3 + O(\theta_n^4). \quad (24)$$

Полагая в выражениях (23) и (24) $\theta_n = tn^{-1/2}$, из (21) и (22), получаем

$$E_{n,0}\Delta_n(t) = -\frac{C(a, b)b^2t^3}{3\sqrt{n}} + O(n^{-1}); \quad (25)$$

$$E_{n,0}\Delta_n^2(t) = \frac{4C(a, b)b^2t^3}{3\sqrt{n}} + O(n^{-1}). \quad (26)$$

Теперь соотношение (19) следует из неравенства Чебышева, поскольку при $0 \leq \delta < 1/2$ из (26) получаем, что

$$P_{n,0} (\Delta_n(t) \geq A) \leq \frac{E_{n,0}\Delta_n^2(t)}{A^2} = O(n^{-1/2}) = o(n^{-\delta}).$$

Доказательство соотношения (20) аналогично:

$$P_{n,tn^{-1/2}} (\Delta_n(t) \leq -A) \leq P_{n,tn^{-1/2}} (|\Delta_n(t)| \geq A) \leq \frac{E_{n,tn^{-1/2}}\Delta_n^2(t)}{A^2} = O(n^{-1/2}) = o(n^{-\delta}).$$

Докажем теперь соотношение (18). Нам понадобится вспомогательная лемма.

Лемма 3.3. *Для любого $x > 0$ существует константа $C > 0$, не зависящая от x и такая, что справедливо следующее неравенство*

$$P_{n,0} (\sqrt[4]{n}|\Delta_n(t)| \geq x) \leq Ce^{-x}, \quad C > 0.$$

Доказательство. Учитывая неравенство Чебышева, рассматриваемую вероятность можно представить в виде

$$P_{n,0} (\sqrt[4]{n}|\Delta_n(t)| \geq x) = P_{n,0} (\sqrt[4]{n}\Delta_n(t) \geq x) + P_{n,0} (-\sqrt[4]{n}\Delta_n(t) \geq x) = P_{n,0} (\exp\{\sqrt[4]{n}\Delta_n(t)\} \geq \exp\{x\}) + P_{n,0} (\exp\{-\sqrt[4]{n}\Delta_n(t)\} \geq \exp\{x\}) \leq e^{-x} (E_{n,0} \exp\{\sqrt[4]{n}\Delta_n(t)\} + E_{n,0} \exp\{-\sqrt[4]{n}\Delta_n(t)\}).$$

Теперь, полагая в формуле (16) $s = \pm i$, получим

$$E_{n,0} \exp\{\pm\sqrt[4]{n}\Delta_n(t)\} = \exp\left\{\frac{2b^2C(a, b)t^3}{3} + o(1)\right\}.$$

Отсюда следует утверждение леммы. \square

Проверим теперь условие (18) леммы 3.2. Имеем

$$\begin{aligned} E_{n,0}|\Delta_n(t)|\mathbf{1}_{(\gamma n^{-\delta/2}, A)}(|\Delta_n(t)|) &\leq \\ &\leq AE_{n,0}\mathbf{1}_{(\gamma n^{-\delta/2}, A)}(|\Delta_n(t)|) \leq \\ &\leq AP_{n,0}\left(|\Delta_n(t)| \geq \gamma n^{-\delta/2}\right) = \\ &= AP_{n,0}\left(\sqrt[4]{n}|\Delta_n(t)| \geq \gamma n^{1/4-\delta/2}\right). \end{aligned}$$

Для оценки этой вероятности применим лемму 3.3, из которой следует, что она не превосходит величины

$$C \exp\{-\gamma n^{1/4-\delta/2}\} = o(n^{-\delta}), \quad C > 0, \quad 0 \leq \delta < \frac{1}{2}.$$

Осталось доказать соотношение (19). Для его доказательства применим неравенство Берри–Эссеена (см. [8, приложение 5, с. 449]). Из соотношений (13), (23) и (24) получаем

$$\begin{aligned} \mu_n \equiv E_{n,0}\Lambda_n(t) &= -at^2 + 2bnE_0\left(X_1 - \frac{t}{\sqrt{n}}\right) \mathbf{1}_{[0, t/\sqrt{n}]}(X_1) = -\frac{1}{2}I_{a,b}t^2 + O(n^{-1/2}); \quad (27) \end{aligned}$$

$$\begin{aligned} \sigma_n^2 \equiv D_{n,0}\Lambda_n(t) &= nD_0\left[\frac{t}{\sqrt{n}}(2aX_1 + b \operatorname{sign}(X_1)) + \right. \\ &\quad \left. + 2b(X_1 - t/\sqrt{n}) \mathbf{1}_{[0, tn^{-1/2}]}(X_1)\right] = \\ &= I_{a,b}t^2 + O(n^{-1/2}). \quad (28) \end{aligned}$$

Теперь в силу неравенства Берри–Эссеена и (27), (28) имеем

$$\begin{aligned} P_{n,0}\left(x - n^{-\delta/2} \leq \Lambda_n(t) \leq x\right) &\leq \\ &\leq \left|P_{n,0}\left(\Lambda_n(t) \leq x\right) - \Phi\left(\frac{x - \mu_n}{\sigma_n}\right)\right| + \\ + \left|P_{n,0}\left(\Lambda_n(t) < x - n^{-\delta/2}\right) - \Phi\left(\frac{x - n^{-\delta/2} - \mu_n}{\sigma_n}\right)\right| + \\ &+ \left|\Phi\left(\frac{x - \mu_n}{\sigma_n}\right) - \Phi\left(\frac{x - n^{-\delta/2} - \mu_n}{\sigma_n}\right)\right| \leq \\ &\leq \frac{C}{\sqrt{n}\sigma_n^3}E_0\left|t\left(2aX_1 + b \operatorname{sign}(X_1)\right) + \right. \\ &+ 2b\sqrt{n}\left(X_1 - \frac{t}{\sqrt{n}}\right)\mathbf{1}_{[0, t/\sqrt{n}]}(X_1) + bC(a, b)\frac{t^2}{\sqrt{n}}\left|^3 + \right. \\ &+ \frac{1}{\sqrt{2\pi}\sigma_n n^{\delta/2}} = O\left(n^{-1/2}\right) + \\ &+ O\left(n^{-\delta/2}\right) = O\left(n^{-\delta/2}\right), \quad C > 0. \quad (29) \end{aligned}$$

Лемма доказана. \square

Из леммы 3.3 непосредственно получаем следующее

Следствие 3.1. Для любого $m > 0$ справедливо соотношение

$$E_{n,0}|\Delta_n(t)|^m = O(n^{-m/4}).$$

Доказательство. Учитывая формулу интегрирования по частям, (см., например, [11, лемма 5.6.1, с. 178]), имеем

$$E_{n,0}|\sqrt[4]{n}\Delta_n(t)|^m = \int_0^\infty P_{n,0}\left(|\sqrt[4]{n}\Delta_n(t)| \geq x^{1/m}\right) dx.$$

Теперь согласно лемме 3.3 последний интеграл не превосходит выражения

$$C \int_0^\infty \exp\{-x^{1/m}\} dx \leq C_1 < \infty, \quad C > 0, \quad C_1 > 0.$$

Отсюда следует утверждение следствия. \square

Литература

1. *Takeuchi K.* Asymptotic theory of statistical estimation. — Tokyo, 1974. (In Japanese.)
2. *Asrabadi B. R.* The exact confidence interval for the scale parameter and the MVUE of the Laplace distribution // Communications in Statistics. Theory and Methods, 1985. Vol. 14, No. 3. P. 713–733.
3. *Бурнашев М. В.* Асимптотические разложения для медианной оценки параметра // Теория вероятности и её применения, 1996. Т. 41. Вып. 4. С. 738–753.
4. *Kotz S., Kozubowski T. J., Podgorski K.* The Laplace distribution and generalizations: A revisit with applications to communications, economics, engineering, and finance. — Birkhauser, 2001. 349 p.
5. *Бенинг В. Е., Королёв В. Ю.* Некоторые статистические задачи, связанные с распределением Лапласа // Информатика и её применения, 2008. Т. 2. Вып. 2. С. 19–34.
6. *Bening V. E.* Asymptotic theory of testing statistical hypotheses. — VSP, Utrecht, 2000. 277 p.
7. *Королёв Р. А., Тестова А. В., Бенинг В. Е.* О мощности асимптотически оптимального критерия в случае распределения Лапласа // Вестник Тверского государственного университета, сер. Прикладная математика, 2008. Вып. 8. № 4(64). С. 5–23.
8. *Боровков А. А.* Теория вероятностей. — М.: УРСС, 2003. 470 с.
9. *Королёв Р. А., Бенинг В. Е.* Асимптотические разложения для мощностей критериев в случае распределения Лапласа // Вестник Тверского государственного университета, сер. Прикладная математика, 2008. Вып. 3(10), № 26(86). С. 97–107.
10. *Bickel P. J., Chibisov D. M., Van Zwet W. R.* On efficiency of first and second order // Intern. Statist. Review, 1981. Vol. 49. P. 169–175.
11. *Феллер В.* Введение в теорию вероятностей и ее приложения. Т. 2. — М.: Мир, 1984. 751 с.

MULTICHANNEL QUEUEING SYSTEM WITH REFUSALS OF SERVERS GROUPS

A. Pechinkin¹, I. Sokolov², and V. Chaplygin³

¹IPI RAN, apechinkin@ipiran.ru

²IPI RAN, isokolov@ipiran.ru

³IPI RAN, VasilyChaplygin@mail.ru

The multichannel queueing system SM/PH/n/r ($r \leq \infty$) with unreliable servers and their group refusals is under consideration. The refusals and the restorations of servers groups occur with a constant intensity, the number of servers refusing simultaneously is a stochastic value, and customers with the interrupted servicing begin its servicing anew after server restoration. The methods are offered to calculate the stationary distribution of the number of the customers in the system under different variants of the functioning of the system.

Keywords: multichannel queueing systems; unreliable servers; refusals and restorations of servers groups

ON THE LIMITING CHARACTERISTICS FOR $M(t)/M(t)/S$ QUEUE WITH CATASTROPHES

A. I. Zeifman¹, Ya. A. Satin², A. V. Korotysheva³, and N. A. Tereshina⁴

¹Vologda State Pedagogical University, a_zeifman@mail.ru

²Vologda State Pedagogical University, yacovi@mail.ru

³Vologda State Pedagogical University, a_korotysheva@mail.ru

⁴Vologda State Pedagogical University, nataliya_tereshi@mail.ru

An $M(t)/M(t)/S$ queue with catastrophes is considered and it is supposed that the catastrophes rates depend on the length of the queue. Sufficient conditions for weak ergodicity of the respective queue-length process and the respective bounds are obtained. Some examples are also considered.

Keywords: nonstationary Markovian queueing system; birth and death process with catastrophes; weak ergodicity; bounds; limiting characteristics; approximation

LARGE DEVIATION ASYMPTOTICS OF STATIONARY QUEUES

E. V. Morozov

Institute of Applied Mathematical Research, Karelian Research Center of the Russian Academy of Sciences, emorozov@krc.karelia.ru

The paper is a survey of main asymptotic results playing an important role in the quality of service estimation (QoS) of stationary systems. The asymptotics of the probability that the workload/queue-size process with heavy tail exceeds an increasing level is considered. Similar results for the systems with Levy input process and light-tailed service time are given. The proofs are based on the methods of large deviations theory and illustrated in detail by the $M/M/1$ system. The asymptotics of the overflow probability within regeneration cycle is considered, including the multiserver systems. An asymptotic analysis of system with the long-range dependent input is discussed, with focus on fractional Brownian process. The ties between the long-range dependence of a queue-size process and the moment properties of the embedded process of the regenerations are discussed.

Keywords: stationary queue; large deviation probabilities; asymptotic analysis; light-tailed distributions; fractional Brownian process; long-range dependent process; regeneration

TRAFFIC-LEVEL PROBABILITY MODEL FOR THE NETWORK CENTRIC SYSTEM

V. Y. Borodakiy

National Research Nuclear University "MEPHI," vladbor@inbox.ru

A network centric system model with datacenters serving subscribers' requests for a data transfer is considered. Data flows, so called elastic traffic, are characterized by a class independent bandwidth requirement and have exponentially distributed size. Flow request arriving is modeled by a Poisson process and data flows are served in accordance with the processor sharing discipline. A single link probability model analysis is presented and the exact expression of the flow blocking probability is derived.

Keywords: network centric system; elastic traffic; blocking probability; single link

AN APPROACH TO ACTUARIAL MODELING WITH QUASI-MONTE CARLO:
SIMULATION OF RANDOM SUMS DEPENDING ON STOCHASTIC FACTORSG. Temnov¹ and S. Kucherenko²¹Edgeworth Centre for Financial Mathematics, University College Cork, Ireland, g.temnov@ucc.ie²CPSE, Imperial College, London, UK, s.kucherenko@ic.ac.uk

The problem of estimating the characteristics of a random sum, when the number of summands is also random, is addressed. The considered case includes an additional stochastic factor: although the summed random variables come from a distribution of a known form, the parameters of this distribution are stochastic and can themselves be viewed as random variables (with known distributions). The Quasi-Monte-Carlo techniques are used to handle this problem and to analyze its efficiency relative to the regular Monte-Carlo simulation methods. The typical area of the application of the investigations is actuarial practice which often deals with random sums of financial losses. Besides actuarial applications, the proposed method may be useful in application to certain problems in informatics, related to the aggregation of heavy-tailed data.

Keywords: actuarial modeling; quasi-Monte-Carlo simulation; random sums

ON STABILITY OF IMAGE RECONSTRUCTION IN THE PROBLEMS OF EMISSION TOMOGRAPHY

O. V. Shestakov

Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics,
M. V. Lomonosov Moscow State University, oshestakov@cs.msu.su

This paper deals with the problem of reconstructing images from projections in emission tomography settings. Within the frames of given mathematical model, the closeness estimates are derived for reconstructed images when using finite number of projections.

Keywords: emission tomography; Radon transform; projections; closeness estimates

ON PROBABILISTIC ASPECTS OF ERROR CORRECTION CODES WHEN THE NUMBER OF ERRORS
IS A RANDOM SETA. N. Chuprunov¹ and B. I. Khamdeyev²¹Research Institute of Mathematics and Mechanics, Kazan State University, achuprunov@mail.ru²Research Institute of Mathematics and Mechanics, Kazan State University, Khamdeyevbi@mail.ru

In the paper, n messages each containing N blocks are considered. Each block is encoded with some antinoise coding method, which can correct not more than q mistakes. Here, it is assumed that the number of mistakes lies

in some random subset $N_i(\omega_1)$, $\omega_1 \in \Omega_1$ of integer numbers. The probability $\mathbf{P}(A)$ of the event A is studied which means that all the mistakes would be corrected. Probability $\mathbf{P}(A)$ is formulated in terms of conditional probabilities. It is shown that as $n, N \rightarrow \infty$ so that $\alpha = n/N \rightarrow \alpha_0 < \infty$, at $q = 1$, probabilities $\mathbf{P}(A)$ converge at almost all $\omega_1 \in \Omega_1$. The limit is obtained.

Keywords: generalized allocation scheme; convergence almost sure; Hamming code

ON THE DISTRIBUTION OF PARTICLE SIZE UNDER FRACTURING

V. Y. Korolev

Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics,
M. V. Lomonosov Moscow State University; Institute of Informatics Problems, Russian Academy of Sciences

A new model for the distribution of particle size under fracturing with the account of nonconstant or random character of the intensity of the flow of impacts. Within the framework of this model, a criterion of the log-normality of the distribution under consideration is formulated and the class of possible distributions of particle size under fracturing is described. Along with many known models, this class contains scale mixtures of log-normal laws.

Keywords: log-normal distribution; mixtures of normal distributions; compound Cox process

SOME ESTIMATES FOR CHARACTERISTIC FUNCTIONS WITH AN APPLICATION TO SHARPENING THE MISES INEQUALITY

I. G. Shevtsova

Faculty of Computational Mathematics and Cybernetics, M.V. Lomonosov Moscow State University,
ishevtsova@cs.msu.su

New estimates are constructed for characteristic functions of distributions with finite absolute moments of order $2 + \delta$, $0 \leq \delta \leq 1$. The Mises moment inequality for lattice distributions is also improved.

Keywords: Fourier transform; characteristic function; symmetrization; convolution; lattice distribution; arithmetic distribution; span

ON THE POWER OF THE TESTS IN THE CASE OF GENERALIZED LAPLACE DISTRIBUTION

V. E. Bening¹ and O. O. Lyamin²

¹Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University,
bening@yandex.ru

²Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University

In the paper, the formula for the normalized limit difference between the power of asymptotically most powerful test and that of a locally most powerful test in the case of the generalized Laplace distribution is obtained.

Keywords: generalized Laplace or generalized double exponential distribution; power function; deficiency; asymptotic expansion

Об авторах

Бенинг Владимир Евгеньевич (р. 1954) — доктор физико-математических наук, профессор кафедры математической статистики факультета вычислительной математики и кибернетики МГУ им. М. В. Ломоносова, старший научный сотрудник ИПИ РАН

Бородакий Владимир Юрьевич (р. 1983) — аспирант факультета кибернетики, Национальный исследовательский ядерный университет «МИФИ»

Зейфман Александр Израилевич (р. 1954) — доктор физико-математических наук, профессор, декан факультета прикладной математики и компьютерных технологий Вологодского государственного педагогического университета; старший научный сотрудник ИПИ РАН; ведущий научный сотрудник ВНКЦ ЦЭМИ РАН

Королёв Виктор Юрьевич (р. 1954) — доктор физико-математических наук, профессор кафедры математической статистики факультета вычислительной математики и кибернетики МГУ им. М. В. Ломоносова, ведущий научный сотрудник ИПИ РАН

Коротышева Анна Владимировна (р. 1988) — студентка-дипломница Вологодского государственного педагогического университета

Кучеренко Сергей Станиславович (р. 1960) — кандидат физико-математических наук, старший научный сотрудник, Империял Колледж, Лондон, Великобритания

Лямин Олег Олегович (р. 1984) — аспирант факультета вычислительной математики и кибернетики МГУ им. М. В. Ломоносова

Морозов Евсей Викторович (р. 1947) — доктор физико-математических наук, ведущий научный сотрудник Института прикладных математических исследований, КарНЦ РАН

Печинкин Александр Владимирович (р. 1946) — доктор физико-математических наук, профессор, глав-

ный научный сотрудник ИПИ РАН; профессор Российского университета дружбы народов (РУДН)

Сатин Яков Александрович (р. 1978) — кандидат физико-математических наук, старший преподаватель кафедры математического анализа Вологодского государственного педагогического университета

Соколов Игорь Анатольевич (р. 1954) — академик (действительный член) Российской академии наук, доктор технических наук, директор ИПИ РАН

Темнов Григорий Олегович (р. 1980) — кандидат физико-математических наук, научный сотрудник и преподаватель Коркского университета, Ирландия

Терёшина Наталия Александровна (р. 1988) — студентка-дипломница Вологодского государственного педагогического университета

Хамдеев Булат Ильдарович (р. 1986) — аспирант, НИИММ им. Н. Г. Чеботарева Казанского государственного университета

Чаплыгин Василий Васильевич (р. 1978) — кандидат физико-математических наук, старший научный сотрудник ИПИ РАН

Чупрунов Алексей Николаевич (р. 1953) — доктор физико-математических наук, ведущий научный сотрудник, НИИММ им. Н. Г. Чеботарева Казанского государственного университета

Шевцова Ирина Геннадьевна (р. 1983) — кандидат физико-математических наук, ассистент факультета Вычислительной математики и кибернетики Московского государственного университета им. М. В. Ломоносова

Шестаков Олег Владимирович (р. 1976) — кандидат физико-математических наук, старший преподаватель кафедры математической статистики факультета вычислительной математики и кибернетики МГУ им. М. В. Ломоносова

About Authors

Bening Vladimir E. (b. 1954) — Doctor of Science in physics and mathematics; professor, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University; senior researcher, Institute of Informatics Problems, Russian Academy of Sciences

Borodakiy Vladimir Y. (b. 1983) — PhD student; Faculty of Cybernetics, National Research Nuclear University “MEPHI”

Chaplygin Vasily V. (b. 1978) — Candidate of Physical and Mathematical Sciences (PhD); senior scientist, Institute of Informatics Problems, Russian Academy of Sciences

Chuprunov Alexey N. (b. 1953) — Doctor of Science in physics and mathematics; leading researcher, Research Institute of Mathematics and Mechanics, Kazan State University

Khamdeyev Bulat I. (b. 1986) — PhD student, Research Institute of Mathematics and Mechanics, Kazan State University

Korolev Victor Y. (b. 1954) — Doctor of Science in physics and mathematics; professor, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University; leading researcher, Institute of Informatics Problems, Russian Academy of Sciences

Korotysheva Anna V. (b. 1988) — graduate student, Vologda State Pedagogical University

Kucherenko Sergey S. (b. 1980) — Candidate of Science (PhD) in physics and mathematics, senior research associate, Imperial College, London, UK

Lyamin Oleg. O. (b. 1984) — PhD student, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University

Morozov Evsey V. (b. 1947) — Doctor of Science in physics and mathematics; leading researcher, Institute of Applied Mathematical Research, Karelian Research Center Russian Academy of Sciences

Pechinkin Alexander V. (b. 1946) — Doctor of Science in physics and mathematics; principal scientist, Institute of Informatics Problems, Russian Academy of Sciences; professor, Peoples’ Friendship University of Russia

Satin Yakov A. (b. 1978) — Candidate of Science in physics and mathematics, senior lecturer, Vologda State Pedagogical University

Shestakov Oleg V. (b. 1976) — Candidate of Science (PhD) in physics and mathematics; senior lecturer, Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University

Shevtsova Irina G. (p. 1983) — Candidate of Science (PhD) in physics and mathematics; assistant, Faculty of Computational Mathematics and Cybernetics, M. V. Lomonosov Moscow State University

Sokolov Igor A. (b. 1954) — Academician of the Russian Academy of Sciences; Doctor of Technical Sciences; director, Institute of Informatics Problems, Russian Academy of Sciences

Temnov Grigory O. (b. 1980) — Candidate of Science (PhD) in physics and mathematics; research fellow and part-time lecturer, University College Cork, Ireland

Tereshina Natalia A. (b. 1988) — graduate student, Vologda State Pedagogical University

Zeifman Alexander I. (b. 1954) — Doctor of Science in physics and mathematics; professor, Dean of the Faculty of Applied Mathematics, Vologda State Pedagogical University; senior researcher, Institute of Informatics Problems, Russian Academy of Sciences; leading researcher, Vologda Science-Coordination Center, Central Economy and Mathematics Institute, Russian Academy of Sciences

Правила подготовки рукописей статей для публикации в журнале «Информатика и её применения»

Журнал «Информатика и её применения» публикует теоретические, обзорные и дискуссионные статьи, посвященные научным исследованиям и разработкам в области информатики и ее приложений. Журнал издается на русском языке. Тематика журнала охватывает следующие направления:

- теоретические основы информатики;
- математические методы исследования сложных систем и процессов;
- информационные системы и сети;
- информационные технологии;
- архитектура и программное обеспечение вычислительных комплексов и сетей.

1. В журнале печатаются результаты, ранее не опубликованные и не предназначенные к одновременной публикации в других изданиях. Публикация не должна нарушать закон об авторских правах. Направляя свою рукопись в редакцию, авторы автоматически передают учредителям и редколлегии неисключительные права на издание данной статьи на русском языке и на ее распространение в России и за рубежом. При этом за авторами сохраняются все права как собственников данной рукописи. В связи с этим авторами должно быть представлено в редакцию письмо в следующей форме: Соглашение о передаче права на публикацию:

«Мы, нижеподписавшиеся, авторы рукописи « _____ », передаем учредителям и редколлегии журнала «Информатика и её применения» неисключительное право опубликовать данную рукопись статьи на русском языке как в печатной, так и в электронной версиях журнала. Мы подтверждаем, что данная публикация не нарушает авторского права других лиц или организаций. Подписи авторов: (ф. и. о., дата, адрес)».

Редколлегия вправе запросить у авторов экспертное заключение о возможности опубликования представленной статьи в открытой печати.

2. Статья подписывается всеми авторами. На отдельном листе представляются данные автора (или всех авторов): фамилия, полное имя и отчество, телефон, факс, e-mail, почтовый адрес. Если работа выполнена несколькими авторами, указывается фамилия одного из них, ответственного за переписку с редакцией.
3. Редакция журнала осуществляет самостоятельную экспертизу присланных статей. Возвращение рукописи на доработку не означает, что статья уже принята к печати. Доработанный вариант с ответом на замечания рецензента необходимо прислать в редакцию.
4. Решение редакционной коллегии о принятии статьи к печати или ее отклонении сообщается авторам. Редколлегия не обязуется направлять рецензию авторам отклоненной статьи.
5. Корректур статей высылаются авторам для просмотра. Редакция просит авторов присылать свои замечания в кратчайшие сроки.
6. При подготовке рукописи в MS Word рекомендуется использовать следующие настройки. Параметры страницы: формат — А4; ориентация — книжная; поля (см): внутри — 2,5, снаружи — 1,5, сверху — 2, снизу — 2, от края до нижнего колонтитула — 1,3. Основной текст: стиль — «Обычный»: шрифт Times New Roman, размер 14 пунктов, абзацный отступ — 0,5 см, 1,5 интервала, выравнивание — по ширине. Рекомендуемый объем рукописи — не свыше 25 страниц указанного формата. Ознакомиться с шаблонами, содержащими примеры оформления, можно по адресу в Интернете: <http://www.ipiran.ru/journal/template.doc>.
7. К рукописи, предоставляемой в 2-х экземплярах, обязательно прилагается электронная версия статьи (как правило, в форматах MS WORD (.doc) или LaTeX (.tex), а также — дополнительно — в формате .pdf) на дискете, лазерном диске или по электронной почте. Сокращения слов, кроме стандартных, не применяются. Все страницы рукописи должны быть пронумерованы.
8. Статья должна содержать следующую информацию на русском и английском языках: название, Ф.И.О. авторов, места работы авторов и их электронные адреса, аннотация (не более 100 слов), ключевые слова. Ссылки на литературу в тексте статьи нумеруются (в квадратных скобках) и располагаются в порядке их первого упоминания. Все фамилии авторов, заглавия статей, названия книг, конференций и т. п. даются на языке оригинала, если этот язык использует кириллический или латинский алфавит.

9. Присланные в редакцию материалы авторам не возвращаются.
10. При отправке файлов по электронной почте просим придерживаться следующих правил:
- указывать в поле subject (тема) название журнала и фамилию автора;
 - использовать attach (присоединение);
 - в случае больших объемов информации возможно использование общеизвестных архиваторов (ZIP, RAR);
 - в состав электронной версии статьи должны входить: файл, содержащий текст статьи, и файл(ы), содержащий(е) иллюстрации.
11. Журнал «Информатика и её применения» является некоммерческим изданием, и гонорар авторам не выплачивается.

Адрес редакции: Москва 119333, ул. Вавилова, д. 44, корп. 2, ИПИ РАН
Тел.: +7 (499) 135-86-92 Факс: +7 (495) 930-45-05 E-mail: rust@ipiran.ru

Выпускающий редактор Л. Кокушкина
Художественный редактор М. Седакова
Сдано в набор 19.07.09. Подписано в печать 15.09.09. Формат 60 × 84/8
Бумага офсетная. Печать цифровая. Усл.-печ. л. 11,5. Уч.-изд. л. 10,5. Тираж 200 экз.

Заказ №

Издательство «ТОРУС ПРЕСС», Москва 119991, ул. Косыгина, д. 4
torus@torus-press.ru; <http://www.torus-press.ru>

Отпечатано в ППП «Типография «Наука» с готовых диапозитивов
Москва 121099, Шубинский пер., д. 6