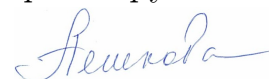


Федеральное государственное бюджетное
образовательное учреждение высшего образования
«Петрозаводский государственный университет»

На правах рукописи



Пешкова Ирина Валерьевна

Монотонность, экстремальное поведение и
регенеративное моделирование систем
обслуживания со смешанными
распределениями

Специальность 1.2.2. Математическое моделирование, численные
методы и комплексы программ

Диссертация на соискание ученой степени
доктора физико-математических наук

Научный консультант:
доктор физико-математических наук,
профессор Морозов Евсей Викторович

Петрозаводск – 2026

Оглавление

Введение	6
Глава 1. Отношения порядка конечных смесей распределений ..	22
1.1. Основные термины и определения	23
1.2. Стохастический порядок и упорядоченность по интенсивности от- казов	28
1.3. Каплинг случайных величин и стохастическая упорядоченность .	30
1.4. Достаточные условия упорядоченности по интенсивности отказов	32
1.5. Функция интенсивности отказов распределений с длинными хво- стами	36
1.6. Функция интенсивности отказов конечной смеси распределений .	39
1.7. Стохастический порядок и упорядоченность по интенсивности от- казов двухкомпонентных смесей	42
1.8. Показательное-Парето распределение	46
1.9. Сравнение некоторых двухкомпонентных смесей	53
1.10. Равномерная метрика для двухкомпонентных смесей распределения	57
1.11. Стохастический порядок и упорядоченность по интенсивности от- казов многокомпонентных смесей	65
1.12. Многокомпонентное показательное-Парето распределение	69
1.13. Выводы по результатам главы 1	71
Глава 2. Монотонность характеристик качества обслуживания .	73
2.1. Понятие монотонности систем обслуживания	74
2.2. Монотонность многосерверных систем обслуживания $GI/G/N/\infty$ с неограниченным буфером	76
2.3. Монотонность многосерверных систем обслуживания $GI/G/N/K$ с конечным буфером	78
2.4. Монотонность в системах со смешанным обслуживанием	82
2.5. Монотонность в системах с повторными вызовами	92

2.6. Монотонность в системах со смешанным входным процессом . . .	109
2.7. Монотонность модели высокопроизводительного кластера	114
2.8. Выводы по результатам главы 2	119

Глава 3. Системы обслуживания со смешанным входным процессом 121

3.1. Аппроксимация суперпозиции процессов восстановления	121
3.2. Аппроксимация входного процесса, заданного суперпозицией показательного и Парето распределений	125
3.3. Аппроксимация входного процесса, заданного суперпозицией распределений Вейбулла	134
3.4. Выводы по результатам главы 3	137

Глава 4. Экстремумы стационарных последовательностей и их применение к системам обслуживания 139

4.1. Асимптотическое распределение экстремальных значений	140
4.2. Области притяжения распределений экстремальных значений . .	143
4.3. Линейные преобразования над максимумами	147
4.4. Асимптотическое поведение функции распределения величины превышения уровня t	150
4.5. Экстремальное поведение строго стационарных последовательностей	152
4.6. Экстремальное поведение в системах $GI/G/1/\infty$	155
4.7. Монотонность систем $GI/G/1/\infty$ по экстремальному индексу . .	159
4.8. Монотонность систем $GI/M/N$ по экстремальному индексу . . .	164
4.9. Монотонность систем $GI/G/1$ со смешанным обслуживанием по экстремальному индексу	165
4.10. Экстремальное поведение в системах со смешанным входным процессом	174
4.11. Сравнение экстремальных индексов стохастически упорядоченных строго стационарных последовательностей	176
4.12. Выводы по результатам главы 4	179

Глава 5. Экстремальное поведение систем с субъэкспоненциальным обслуживанием	180
5.1. Экстремальное поведение в системах $GI/G/1/\infty$	180
5.2. Система $GI/G/1$ со смешанным обслуживанием с субъэкспоненциальными компонентами	189
5.3. Система $M/G/1$ с показательным-Парето временем обслуживания	193
5.4. Система с m классами повторных вызовов и временем обслуживания, заданным смесью субъэкспоненциальных распределений .	198
5.5. Экстремальное поведение в системе $M/G/2$	202
5.6. Выводы по результатам главы 5	208
Глава 6. Методы и алгоритмы регенеративного моделирования систем обслуживания	210
6.1. Классическая регенерация	212
6.2. Метод регенеративных огибающих	214
6.3. Метод построения регенеративных огибающих для систем обслуживания со смешанными распределениями	219
6.4. Искусственная регенерация на основе метода расщепления	220
6.5. Метод экспоненциального расщепления	236
6.6. Метод искусственной регенерации в системах обслуживания с суперпозицией входных потоков	239
6.7. Регенеративное оценивание	248
6.8. Регенеративная оценка экстремального индекса	250
6.9. Описание комплекса программ	251
6.10. Выводы по результатам главы 6	252
Заключение	253
Список сокращений и условных обозначений	255
Список использованной литературы	258
Список рисунков	280

Список таблиц	284
---------------------	-----

Приложение А. Свидетельства о регистрации программы для ЭВМ	287
--	-----

Введение

Актуальность темы исследования. Работа посвящена исследованию нового класса стохастических систем обслуживания, в которых входной поток или обслуживание на серверах задаются смешанными распределениями, т. е. представляют собой конечную смесь распределений. Применение смеси распределений при моделировании систем массового обслуживания мотивировано тем, что процессы прибытия и обслуживания могут фактически представлять собой потоки неоднородных данных и генерироваться разными распределениями. Стремительное развитие телекоммуникационных систем привело, с одной стороны, к росту сложности систем, а, следовательно, и усложнению корреляционной структуры, описывающих их процессы, а с другой – к повышению требований к качеству обслуживания. Это вызывает необходимость в разработке новых методов исследования, позволяющих как аналитически исследовать модели реальных систем, так и получать приемлемые верхние и нижние оценки стационарных характеристик систем.

Одним из важнейших свойств систем (и не только стохастических) является свойство монотонности. В контексте данной работы оно лежит в основе сравнения стационарных характеристик и позволяет вместо исследования характеристик системы со смешанными распределениями получать их оценки, используя аналитические результаты и методы моделирования более простых систем. Поскольку свойства монотонности опираются на теорию отношений порядка (стохастический порядок, упорядоченность по интенсивности отказов, отношения с вероятностью 1), то для исследования свойств монотонности систем со смешанными распределениями требуется получить новые и обобщить известные результаты об отношениях порядка для конечных смесей распределений. Упорядоченность по интенсивности отказов является более ограничительной (по сравнению со стохастическим порядком), но она позволяет в явном виде найти соотношения между параметрами сравниваемых распределений, гарантирующими искомый порядок и, более того, позволяет сравнивать между собой системы, описываемые различными распределениями (времени обслуживания и/или входного потока). Получение оценок характеристик качества обслуживания (стационарного времени ожидания, стационарной длины очереди) опи-

рается на построение минорантной и мажорантной систем для исходной системы со смешанными распределениями, таким образом, чтобы входные потоки и времена обслуживания находились в отношении порядка (стохастического или по интенсивности отказов). Это позволяет, используя оценки в более простых мажорантной и минорантной системах, получить оценки характеристик более сложных систем, таких, как многосерверные системы со смешанным обслуживанием или смешанным входным процессом, многоклассовые системы с повторными вызовами, высокопроизводительные системы со случайным числом обслуживающих серверов. Более того, свойства монотонности позволяют получать условия стационарности сложных систем на основе известных условий стационарности более простых (мажорантных) систем.

Требования к надежности систем тесно связаны с исследованием экстремумов их стационарных характеристик, с которыми, в свою очередь, тесно связаны свойства превышения характеристиками высоких пороговых значений. Так, превышение процессами таких порогов временем ожидания заявки в очереди или наличие коротких интервалов между превышениями или появление превышений группами могут оказывать значительное влияние на работоспособность системы. В этой связи исследование свойств монотонности по экстремальному индексу и экстремального поведения различных систем, в том числе систем со смешанными распределениями, является актуальным и перспективным направлением исследований. Поскольку в сложных стохастических системах исследуемые процессы, как правило, сильно коррелированы, то предельное распределение максимумов стационарного времени ожидания выражается через так называемый экстремальный индекс (он отражает кластерную структуру данных). Кластеризация влияет на последовательность превышения высоких пороговых значений: благодаря коррелированной структуре данных, превышения могут группироваться, а не появляться как изолированные события. Важно подчеркнуть, что для применения теоретических результатов к конкретным распределениям требуется найти нормализующие последовательности, гарантирующие существование предельного распределения нормализованных максимумов. Нахождение каждой такой последовательности, вообще говоря, является отдельной математической задачей. Таким образом, оценка экстремального индекса стационарных характеристик, а также исследование свойств монотонности раз-

личных систем обслуживания по экстремальному индексу, в том числе систем со смешанными распределениями, является важной задачей при прогнозировании экстремальных событий (перегрузок в системах) и, как следствие, предупреждения их последствий.

Часто распределения случайных процессов, описывающих динамику современных систем обслуживания, представляют собой смеси распределений с так называемыми легкими или тяжелыми хвостами. Например, случайно выбранная заявка в сети связи может быть из видеопотока, из онлайн-игр или из аудиопотока. В сетях промышленного Интернета и Интернета вещей некоторые данные являются «скачкообразными», например, температура, давление или свет, считываемые с датчика. Другие устройства могут создавать огромные объемы трафика данных, например, обработка мультимедийного контента или сложных расчетов. Таким образом, для времени обработки данных (моделируемого временем обслуживания в моделях массового обслуживания) наиболее подходящей является модель, учитывающая смешанные данные. Поскольку, как правило, «легкие» данные аппроксимируются показательным распределением, а «тяжелые» данные в коммуникационных сетях – распределением Парето, то в диссертационной работе в качестве основного типа смешанных распределений используется смесь показательного и Парето распределений, а именно многокомпонентное показательное-Парето распределение.

Регенеративное оценивание позволяет исследовать системы обслуживания, для которых отсутствуют аналитические результаты или трудно найти численные решения. Системы со смешанными распределениями как раз относятся к таким сложным системам. Более того, существуют нерегенерирующие в классическом смысле системы, для которых требуется разработка новых или модификация существующих методов регенерации, предназначенных для оценки стационарных характеристик. Так, система со смешанным входным процессом, задаваемым суперпозицией независимых входных потоков, не является регенерирующей.

Эффективность применения регенеративного метода зависит от частоты моментов регенерации. В такие моменты система попадает в определенное состояние, после которого ее поведение не зависит от предыстории. Регенеративный подход сталкивается с трудностями, когда число регенераций растёт

слишком медленно в ходе моделирования, что ведет к неприемлемо длительному времени моделирования для получения оценок с заданной точностью. Более того, частота регенераций влияет на точность оценивания (дисперсию в доверительном интервале). Для ускорения регенерации необходимо разработать новые или модифицировать существующие методы получения искусственных моментов регенерации.

Степень разработанности темы. Впервые систематический анализ монотонности некоторых систем обслуживания был развит Д. Штойяном в работе [24]. Свойства внутренней и внешней монотонности систем обслуживания изучались в работах В. Витта, Д. Сондермана, С. Асмуссена [34, 47, 167, 168, 177]. Так, например, в работе [177] исследуются свойства монотонности многосерверных систем с ограниченным и неограниченным буфером. В работе [47] рассмотрена монотонность систем с групповым входным потоком. Тем не менее, свойства монотонности систем со смешанными потоками ранее не изучались указанными авторами, и, тем более, ранее не рассматривались системы, в которых входные потоки или времена обслуживания представляют собой смеси различных распределений.

Экстремальные значения характеристик производительности систем массового обслуживания изучались Х. Ротсеном, Д. Иглхартом, С. Асмуссеном [34, 35, 95, 146]. В работах этих авторов исследуется предельное поведение максимума стационарного времени ожидания и стационарной длины очереди односерверных систем. Применение теории экстремальных значений для независимых и однозависимых (когда зависимы два соседних цикла регенерации) процессов регенерации изучалось в статье [146]. Исследование асимптотического поведения максимумов времени ожидания, виртуального времени ожидания и максимального размера очереди в системе $GI/G/1$ является основной целью статьи [95]. С. Канг и Р. Серфозо в работе [103] показали, что в ациклической сети Джексона (состоящей из m систем вида $M/M/1$) асимптотическое экстремальное поведение времени пребывания заявки описывается распределением Гумбеля. Алгоритм вычисления экстремального индекса (аппроксимирующего размер кластера больших значений) стационарного времени ожидания системы $G/G/1$ с распределением, принадлежащим области притяжения распределения Гумбеля, приведен в [94]. Свойства оценок экстремального индекса обсуждают-

ся в работе И. Вейсмана, С. Новака [186]. В статье Р. Разумчика и Н. Маркович [112] получены распределения кластерных и межкластерных размеров стационарных времен ожидания в системе $G/G/1$. В работе С. Фосса [77] обсуждается асимптотическое распределение максимума стационарного времени ожидания для системы $M/G/2$, в которой время обслуживания задано распределением с тяжелым хвостом. Однако монотонность систем обслуживания по экстремальному индексу времени ожидания в упомянутых работах не изучалась.

Метод рандомизации переходного ядра с целью ускорения получения моментов регенерации в системе $M/M/1$ впервые был предложен С. Андрадоттир, П. Глином, Дж. Калвином в работе [29]. Метод искусственной регенерации, основанный на экспоненциальном расщеплении, предложен А. Андроновым в [28], где рассматривается система с одним входным потоком. Независимо в работе [120] Р. Мока и С. Джанея рассматривают более общую конструкцию декомпозиции распределений. Для систем со смешанным входным процессом К. Сигман предложил теоретическую конструкцию однозависимых циклов регенерации [161], основная проблема применения которой связана с наличием в конструкции, вообще говоря, неизвестных констант.

Объектом исследования являются системы обслуживания со смешанными распределениями: системы со смешанным обслуживанием и системы со смешанным входным процессом.

Предметом исследования выступают условия монотонности, экстремальное поведение и оценки стационарных характеристик систем обслуживания со смешанными распределениями.

Цель исследования состоит в разработке и обосновании математических моделей и методов оценки стационарных характеристик стохастических систем обслуживания со смешанными распределениями входного процесса и времени обслуживания.

Достижение этой цели требует решения следующих **основных задач**:

1. Построение математических моделей стохастических систем обслуживания со смешанными распределениями.
2. Получение условий монотонности (стохастического порядка или упорядоченности по интенсивности отказов) стационарных характеристик систем со смешанными распределениями.

3. Получение условий монотонности по экстремальному индексу систем обслуживания со смешанными распределениями.

4. Разработка методов искусственной регенерации для систем обслуживания со смешанными распределениями для ускорения оценивания их стационарных характеристик.

5. Разработка комплекса программ регенеративного имитационного моделирования систем обслуживания со смешанными распределениями для оценки стационарных характеристик.

Научная новизна результатов, изложенных в диссертации, состоит в следующем.

1. Доказано, что если компоненты конечной смеси распределений упорядочены (стохастически или по интенсивности отказов), то компоненты и смесь упорядочены соответствующим образом. Получены достаточные условия стохастического порядка двух смесей распределения с разными коэффициентами пропорции смеси.

2. Получены достаточные условия монотонности системы со смешанным обслуживанием, а также достаточные условия монотонности системы со смешанным входным процессом.

3. Доказаны теоремы о монотонности по экстремальному индексу систем $GI/M/N$ и $M/G/1$, а также системы со смешанным обслуживанием.

4. Получено аппроксимирующее распределение входного процесса, заданного многомерной суперпозицией показательного и Парето распределений, а также аппроксимирующее распределение входного потока, заданного суперпозицией распределений Вейбулла. Результаты моделирования подтверждают хорошую точность аппроксимации при коэффициенте загрузки меньше 0.6.

5. Разработан метод регенеративных огибающих для построения минорантной и мажорантной регенерирующих систем для исходной нерегенерирующей системы.

6. Разработан метод искусственной регенерации на основе экспоненциального расщепления плотности системы обслуживания с суперпозицией независимых входных потоков, содержащих распределения с тяжелым хвостом.

7. Получены достаточные условия применимости метода экспоненциального расщепления для системы обслуживания, в которой смешанный входной

процесс включает потоки, заданные распределением Парето.

8. Разработан метод искусственной регенерации на основе рандомизации переходного ядра вложенной Цепи Маркова для систем, в которых входной поток имеет распределение Парето или гиперэкспоненциальное распределение, а время обслуживания имеет показательное распределение.

9. Разработан комплекс проблемно-ориентированных программ и алгоритмов для регенеративного моделирования систем обслуживания со смешанными распределениями.

Методы исследования. Для проведения диссертационного исследования использовался аппарат следующих дисциплин: математический анализ, теория вероятностей, теория случайных процессов, теория массового обслуживания, математическая статистика, теория экстремальных значений, теория надежности.

Для исследования отношений порядка случайных величин использован метод каплинга [109, 171].

Для аппроксимации смешанного входного процесса (суперпозиции независимых входных потоков) применен метод Пальма [27, 180, 181, 179].

Для оценивания характеристик систем обслуживания использован регенеративный метод [121, 124, 127, 157, 158].

Для построения моментов регенерации в системах обслуживания вида $G/M/1$ применен метод расщепления переходного ядра соответствующей вложенной цепи Маркова по моментам прихода заявок [29].

Для построения моментов регенерации в системах со смешанным входным процессом применен метод искусственной регенерации, основанный на экспоненциальном расщеплении [28].

Теоретическая и практическая значимость работы.

Полученные в диссертационной работе теоретические результаты относятся к теории массового обслуживания. Систематически исследованы свойства монотонности широкого и важного класса систем обслуживания со смешанными распределениями. Впервые исследованы свойства внешней монотонности по экстремальному индексу систем обслуживания, в том числе систем со смешанным обслуживанием. Для моделирования нерегенерирующих систем обслуживания со смешанным входным процессом, содержащим распределения с

тяжелым хвостом, разработан метод искусственной регенерации, основанный на экспоненциальном расщеплении.

Проведенное исследование является значимым, поскольку существенно расширяет возможности использования свойств монотонности для исследования стационарности, оценки стационарных характеристик, подбора параметров реальных систем обслуживания со смешанными распределениями. Такие системы адекватно описывают динамику многих современных телекоммуникационных систем и сетей, обрабатывающих большие наборы неоднородных данных.

Практическая значимость работы определяется применимостью полученных свойств монотонности и регенеративного метода оценивания для получения надежных оценок стационарных характеристик систем обслуживания со смешанными распределениями.

Положения, выносимые на защиту:

1. Модели систем обслуживания со смешанными распределениями, которые позволяют учитывать потоки неоднородных данных.
2. Достаточные условия монотонности систем со смешанными распределениями и систем со смешанным входным процессом, а также достаточные условия монотонности по экстремальному индексу стационарного времени ожидания в системах вида $GI/M/N$ и $M/G/1$, в том числе системах со смешанным обслуживанием.
3. Асимптотическое распределение максимума стационарного времени ожидания для систем вида $GI/G/1$ со смешанным обслуживанием и доминирующей субэкспоненциальной компонентой.
4. Численный метод искусственной регенерации, основанный на экспоненциальном расщеплении, для систем обслуживания со смешанным входным процессом, содержащим распределения с тяжелым хвостом, который позволяет применять регенеративное оценивание и получать состоятельные оценки стационарных характеристик систем.
5. Алгоритм искусственной регенерации систем обслуживания со смешанным входным процессом, включающим потоки, заданные распределением Парето.
6. Численный метод искусственной регенерации на основе рандомизации переходного ядра для систем, в которых входной поток имеет распределение

Парето или гиперэкспоненциальное распределение, а время обслуживания имеет показательное распределение.

7. Комплекс программ для регенеративного моделирования систем обслуживания со смешанными распределениями для оценки стационарных характеристик систем обслуживания.

Соответствие паспорту специальности 1.2.2. Математическое моделирование численные методы и комплексы программ. Положения 4 и 6 соответствуют п. 1 «Разработка новых математических методов моделирования объектов и явлений». Положения 5 и 6 соответствуют п. 3 «Реализация эффективных численных методов и алгоритмов в виде комплексов проблемно-ориентированных программ для проведения вычислительного эксперимента». Положения 1, 2, 3 соответствуют п. 8 «Комплексные исследования научных и технических проблем с применением современной технологии математического моделирования и вычислительного эксперимента».

Достоверность полученных результатов подтверждается математически корректными формулировками и доказательствами теорем, представленными в диссертационной работе, а также согласованностью численных результатов с теоретическими выводами.

Личный вклад автора в получении результатов, изложенных в диссертации. Автор лично участвовал в получении всех результатов, представленных в диссертационной работе, а именно в разработке и применении методов исследования монотонности, экстремального поведения систем со смешанными распределениями, выводе всех формул, доказательстве всех представленных в диссертации теорем, разработке представленного комплекса проблемно-ориентированных программ и написании алгоритмов построения моделирования, выполнении анализа полученных результатов.

Направления исследований диссертационной работы и постановки задач обсуждались с научным консультантом доктором физико-математических наук, профессором Морозовым Е.В., что отражено в ряде совместных публикаций с автором диссертационной работы.

Связь работы с крупными научными проектами. Большая часть работ выполнялась в рамках грантов:

1. В качестве руководителя по гранту РФФИ 19-07-00303 «Разработка моде-

лей и методов регенеративного моделирования программно конфигурируемых сетей для анализа показателей качества их обслуживания» (2019-2021 г.г.).

2. В качестве соисполнителя по гранту РФФ 21-71-10135 «Разработка и исследование методов, алгоритмов и программных средств согласованного трехуровневого моделирования производительности и энергоэффективности систем и сетей хранения и обработки данных» (2022-2024 г.г., руководитель – Румянцев А.С.).

3. В качестве соисполнителя по гранту РФФИ 19-57-45022 ИНДа «Оценка производительности стохастических моделей систем высокопроизводительных, распределенных и повсеместных вычислений, а также систем с автоматическим запросом повторной передачи и контролем ошибок матрично-аналитическим и регенеративным методом» (2019-2021 г.г., руководитель – Румянцев А.С.).

4. В качестве соисполнителя по гранту РФФИ 18-07-00147 «Разработка моделей, методов асимптотического анализа и ускоренного оценивания показателей эффективности и качества обслуживания высокопроизводительных, распределенных вычислительных систем и коммуникационных сетей» (2018-2020 г.г., руководитель – Румянцев А.С.).

Публикации.

По материалам диссертации опубликовано 37 работ, в том числе 5 в рецензируемых научных изданиях, из них 3 в журналах Перечня ВАК по специальности 1.2.2 (физ.-мат.) категории К1 [194, 198, 199], 2 статьи в журналах Перечня ВАК по специальности 1.2.2 (физ.-мат.) категории К2 [189, 192], 8 статей в научных журналах, индексируемых в международных базах данных (приравнены к журналам Перечня ВАК, категория К1) [190, 191, 193, 195, 196, 197, 200, 201], 13 статей в сборниках материалов конференций, индексируемых в международных базах данных [202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214], 11 публикаций в прочих научных изданиях [215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225]. Получено 2 свидетельства о государственной регистрации программы для ЭВМ (Приложение А).

Апробация работы. Основные результаты работы и отдельные ее положения докладывались и обсуждались на 23 международных научных конференциях и семинарах: XXIV Международной конференции имени А. Ф. Терпугова «Информационные технологии и математическое моделирование», ИТММ

– 2025 (Россия, Томск, 20-25 октября 2025), The Fourth International Workshop on Stochastic Modeling and Applied Research of Technology, SMARTY'2024 (Россия, Петрозаводск, 26-30 августа 2024), XI Международная Петрозаводская конференция «Вероятностные методы в дискретной математике», посвященная 90-летию со дня рождения В. Ф. Колчина (Россия, Петрозаводск, 27-31 мая, 2024), 26th International Conference on Distributed Computer and Communication Networks: Control, Computation, Communications DCCN-2023. (Москва, Россия, 25-29 сентября 2023), Stochastic Modeling and Applied Research of Technology Youth Camp: SMARTY Youth Camp 2023 (Россия, Петрозаводск, 21-25 августа, 2023), 25th International Conference on Distributed Computer and Communication Networks: Control, Computation, Communications DCCN-2022 (Москва, Россия, 26-29 сентября 2022), The Third International Workshop on Stochastic Modeling and Applied Research of Technology: SMARTY 2022 (Россия, Петрозаводск, 21-25 августа 2022), Stochastic Modeling and Applied Research of Technology Youth Camp: SMARTY 2021 Youth Camp (Россия, Петрозаводск, 2–3 сентября, 2021), XXXVI Международный семинар по проблемам устойчивости стохастических моделей ISSPSM 2021 (Россия, Петрозаводск, 21-25 июня 2021), 24th International Conference on Distributed Computer and Communication Networks: Control, Computation, Communications DCCN-2021. (Россия, Москва, 20-24 сентября 2021), 20th International Conference named after A. F. Terpugov Information technologies and mathematical modelling ITMM – 2021 (Россия, Томск, Россия, 1-5 декабря 2021), 27th International Conference Computer Networks CN 2020. (Польша, Гданьск, 23-24 июня 2020), 23rd International Conference on Distributed Computer and Communication Networks: Control, Computation, Communications DCCN-2020. (Россия, Москва, 14-18 сентября 2020), The Second International Workshop on Stochastic Modeling and Applied Research of Technology SMARTY 2020 (Россия, Петрозаводск, 16-20 августа 2020), 22nd International Conference on Distributed Computer and Communication Networks: Control, Computation, Communications (Россия, Москва, 23-27 сентября, 2019), 21st International Conference on Distributed Computer and Communication Networks: Control, Computation, Communications (Россия, Москва, 17-21 сентября, 2018), First International Workshop on Stochastic Modeling and Applied Research of Technology SMARTY 2020 (Россия, Петрозаводск, 21-25 сентября, 2018), 20th International

Conference on Distributed Computer and Communication Networks: Control, Computation, Communications DCCN-2017 (Россия, Москва, 25-29 сентября, 2017), 23rd International Conference, CN 2016 (Польша, Бруноу, 14-17 июня, 2016), IX Международная Петрозаводская конференция «Вероятностные методы в дискретной математике» (Россия, Петрозаводск, 30 мая- 3 июня, 2016), 8th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops, ICUMT (Португалия, Лиссабон, 18 – 20 October, 2016), 19th International Conference on Distributed Computer and Communication Networks: Control, Computation, Communications DCCN-2016 (Россия, Москва, 21-25 ноября, 2016), Eighth International Workshop on Simulation (Австрия, Вена, 21-25 сентября, 2015).

Структура работы. Диссертация изложена на 288 страницах, состоит из введения, списка обозначений и сокращений, шести глав, заключения и списка использованной литературы, включающего 188 наименований, содержит 18 таблиц, 40 рисунков.

Основное содержание работы. В Главе 1 введены основные термины и определения, используемые далее в диссертации, рассмотрены понятия и свойства стохастического порядка, упорядоченности по интенсивности отказов, приведено понятие каплинга и его связь со стохастическим порядком и упорядоченностью с вероятностью 1. Получено достаточное условие упорядоченности двух случайных величин по интенсивности отказов в зависимости от класса распределения. Рассмотрены вопросы взаимосвязи классов распределений с длинными хвостами и распределений с возрастающей функцией интенсивности отказов (ВФИ распределений). Доказана Теорема 1.1 о том, что если компоненты смеси упорядочены (стохастически или по интенсивности отказов), то компоненты и смесь упорядочены (стохастически или по интенсивности отказов). В Теореме 1.2 утверждается, что если компоненты двух смесей стохастически упорядочены, то смеси стохастически упорядочены. Доказательство проведено с использованием техники каплинга. В Теореме 1.3 доказаны достаточные условия упорядоченности по интенсивности отказов смесей распределения. Введено показательное-Парето распределение, для которого найдены функция интенсивности отказов, характеристическая функция, математическое ожидание и дисперсия, средняя величина превышения уровня t . Показано, что это распе-

деление относится к классу распределений с длинными хвостами и классу распределений с убывающей функцией интенсивности отказов (УФИ распределений). Получена функция интенсивности отказов величины превышения уровня t . В Теореме 1.5 выведены неравенства между параметрами показательного-Парето распределения, при которых компоненты смеси и смесь упорядочены по интенсивности отказов. Такая же упорядоченность сохраняется для распределений величины превышения уровня u и его среднего. Доказаны Теоремы 1.9 и 1.10, в которых обобщены предыдущие результаты для стохастического порядка и упорядоченности по интенсивности отказов для m -компонентных смесей распределений. В Теореме 1.11 найдены условия, при которых стохастически упорядочены две m -компонентные смеси с различными коэффициентами пропорции смесей.

В Главе 2 приведены понятия внутренней и внешней монотонности систем обслуживания. Доказана Теорема 2.3 о внешней монотонности стационарной длины очереди и стационарного времени ожидания в многосерверных системах со смешанными распределениями и неограниченной очередью. В Теореме 2.4 получены условия монотонности процесса нагрузки в системах с различными смесями распределений. В Теореме 2.5 доказаны условия, при которых выполняется свойство внешней монотонности для стационарной длины очереди в многосерверной системе с конечным буфером и смешанным обслуживанием. В Теореме 2.6 получены верхняя и нижняя границы незавершенной работы для многоклассовой системы с повторными вызовами. Также найдены выражения для преобразования Лапласа-Стилтьеса стационарной незавершенной нагрузки в минорантной и мажорантной системах (Леммы 2.1 и 2.2). В Теореме 2.7 получено в явном виде распределение стационарной незавершенной работы в минорантной системе с распределением Парето времени обслуживания. Доказаны Теоремы 2.8 и 2.9 о монотонности стационарного времени ожидания в системе со смешанным входным процессом (суперпозицией независимых стационарных входных потоков восстановления). Доказана Теорема 2.10 о монотонности стационарной незавершенной работы высокопроизводительного кластера. Также в Лемме 2.3 получена оценка близости (верхняя и нижняя границы незавершенной работы) между двумя высокопроизводительными кластерами для случая, когда времена обслуживания и интервалы между приходами заявок отличаются

ся в этих системах на случайную величину.

В Главе 3 получено аппроксимирующее распределение входного потока, заданного многомерной суперпозицией показательного и Парето распределений. Для случая одного класса заявок (когда время обслуживания задано одним и тем же показательным распределением) получено ПЛС стационарного времени ожидания, что позволяет получить функцию распределения для стационарного времени ожидания в явном виде для аппроксимирующей системы. Приведены численные примеры, подтверждающие теоретические выводы. Построено аппроксимирующее распределение входного потока, заданного многомерной суперпозицией входных потоков с распределением Вейбулла. Это распределение имеет вид конечной смеси распределений Вейбулла и обобщенных Стейси-Гамма распределений. Результаты моделирования подтверждают хорошую точность аппроксимации при коэффициенте загрузки меньше 0.6.

В Главе 4 получена нормализующая последовательность для максимума последовательности с.в., имеющих $k + m$ -показательное-Парето распределение, а также доказано, что распределение максимума сходится к распределению Фреше (Теорема 4.1). В разделе 4.6. получены соотношения для расчета экстремального индекса стационарного времени ожидания в системах $M/H_m/1$, $M/We/1$, $H_m/M/1$, $Pareto/M/1$. В разделе 4.7. доказана Теорема 4.2 о монотонности по экстремальному индексу стационарного времени ожидания систем типа $M/G/1$, а также доказана Теорема 4.3 о монотонности по экстремальному индексу стационарного времени ожидания систем типа $GI/M/1$. Для систем $M/H_m/1$ и $H_m/M/1$ доказаны Леммы 4.1 и 4.2 о монотонности системы по экстремальному индексу. Получены достаточные условия монотонности систем со смешанным обслуживанием по экстремальному индексу (Теорема 4.5). В разделе 4.10. приведены примеры результатов регенеративной оценки экстремального индекса стационарного времени ожидания в системах со смешанным входным процессом.

В Главе 5 найдены нормализующая последовательность и предельное распределение максимума стационарного времени ожидания в системах вида $GI/Pareto/1$ с распределением Парето времени обслуживания (Теорема 5.1, раздел 5.1.). Получены нормализующая последовательность и предельное распределение максимума стационарного времени ожидания в системах $GI/We/1$

с распределением Вейбулла времени обслуживания для случая $\beta \in (0, 1)$ (Теорема 5.2). В разделе 5.2. доказана Теорема 5.3, согласно которой, если в системе $GI/G/1$ время обслуживания задано конечной смесью распределений, r -доминирующая компонента равновесного распределения которого является субъэкспоненциальной, то хвост предельного распределения стационарного времени ожидания эквивалентен хвосту этого распределения с точностью до константы. Более того, предельное распределение максимума стационарного времени ожидания относится к области притяжения распределения экстремальных значений того же типа, что и доминирующая компонента, а экстремальный индекс стационарного времени ожидания равен 0. В разделе 5.3. получены нормализующая последовательность и предельное распределение максимума стационарного времени ожидания в системах $GI/G/1$ с показательным-Парето распределением времени обслуживания (Теорема 5.4). В разделе 5.4. для систем с повторными вызовами с m классами орбит получены асимптотические распределения незавершенной работы в минорантной и мажорантной системах при условии, что время обслуживания относится к классу субъэкспоненциальных распределений (Леммы 5.3 и 5.4). Для систем с повторными вызовами и m классами орбит и распределением Парето времени обслуживания найдены предельные распределения максимумов незавершенной работы в минорантной, исходной и мажорантной системах (Теорема 5.5). В разделе 5.5. для систем $M/G/2$, в которых времена обслуживания имеют распределение Парето, получены предельные распределения стационарного времени ожидания (Лемма 5.5 для $\rho < 1$ и Лемма 5.6 для $1 < \rho < 2$).

В Главе 6 приведены основные сведения о классической регенерации. В разделах 6.2. и 6.3. описан новый метод регенеративных огибающих для многосерверных систем $GI/G/N/\infty$ и систем обслуживания со смешанными распределениями. В разделе 6.4. на основе метода расщепления переходного ядра цепи Маркова получены соотношения для частоты (интенсивности) моментов искусственной регенерации в системах вида $Pareto/M/1$ и $H_2/M/1$, а также приведены результаты численных экспериментов, которые подтверждают полученные теоретические соотношения. В разделе 6.5. приведен метод искусственной регенерации на основе экспоненциального расщепления для построения регенераций в системах обслуживания с суперпозицией входных потоков с тяжелым хво-

стом. Доказана Теорема 6.2 о том, что в моменты искусственной регенерации система регенерирует в классическом смысле. В Лемме 6.2 получены достаточные условия экспоненциального расщепления распределения Парето. Для системы обслуживания со смешанным входным процессом, включающим распределения Парето, получены условия применимости метода экспоненциального расщепления. В разделе 6.6. приведен разработанный алгоритм построения моментов искусственной регенерации для процесса стационарного времени ожидания в системах со смешанным входным процессом. В разделе 6.7. приведен метод регенеративного оценивания для оценки характеристик систем обслуживания, а в разделе 6.8. приведена регенеративная оценка экстремального индекса, которая использована при оценке экстремального индекса стационарного времени ожидания. В разделе 6.9. описан комплекс программ для ЭВМ.

Глава 1. Отношения порядка конечных смесей распределений

В данной главе представлены теоретические результаты, касающиеся отношений порядка конечных смесей распределений, а также некоторых классов распределений. Отношения порядка составляют теоретическую основу для исследования свойств монотонности систем, представленных в Главах 2 и 5.

Конечные смеси распределений широко используются при статистическом анализе и моделировании сложных систем с неоднородными данными [30, 117]. Такие модели применяются во многих прикладных областях, например, в астрономии, биологии, экологии, генетике, медицине, психиатрии, финансах, инженерном деле и маркетинге [169, 58]. В этих приложениях модели конечных смесей лежат в основе различных методов анализа данных, включая кластерный анализ и анализ скрытых классов, дискриминантный анализ, анализ изображений, анализ выживаемости. Использование смесей распределений при моделировании систем обслуживания мотивировано тем, что процессы прибытия и обслуживания могут фактически представлять собой потоки неоднородных данных и генерироваться несколькими разными источниками с различными распределениями.

При исследовании систем обслуживания со смешанными распределениями в Главах 2 и 5 используются свойства монотонности, основанные на сохранении отношений порядков для таких базовых характеристик качества обслуживания, как время ожидания и число заявок в системе. В этой связи в работе рассмотрены понятия стохастического порядка, упорядоченности по интенсивности отказов, порядка с вероятностью 1. Их свойства приведены в работах [116, 145, 156].

Поскольку соотношения между параметрами различных распределений, гарантирующие упорядоченность случайных величин (с.в.), для многих распределений удастся получить в явном виде из условия упорядоченности их функций интенсивности отказов, а не из стохастического порядка, то в исследовании особое место отводится сравнению смесей по интенсивности отказов. Отметим, что функция интенсивности отказов играет важную роль в теории надежности технических и информационных систем обслуживания [2, 40, 106]. Отношения

порядков (стохастического и по интенсивности отказов) для конечных смесей распределения рассмотрены в работах [31, 43, 101, 116, 130].

В реально функционирующих системах часто времена обслуживания могут генерироваться распределениями, которые характеризуются либо лёгкими, либо тяжёлыми хвостами. Такая ситуация обусловлена тем, что часть данных, например, показания температуры, давления или освещённости, снимается с датчиков. В то же время в систему нередко поступают очень большие объёмы трафика — к примеру, мультимедийный контент [71]. Следовательно, модель смеси адекватно описывает время обработки информации. Поскольку, как правило, «лёгкие» данные хорошо аппроксимируются экспоненциальным распределением, а «тяжёлые» данные в сетях связи — распределением Парето, то в настоящей главе исследуется смесь этих двух распределений, а именно многокомпонентное экспоненциально-паретовское распределение, а также приводится ряд его свойств.

Для оценки влияния компонент на смесь в диссертационной работе используется расстояние Колмогорова-Смирнова (равномерная метрика) [93, 104, 155].

1.1. Основные термины и определения

Пусть с.в. X задана функцией распределения (ф.р.) $F_X(x) = P(X \leq x)$, $x \in \mathbb{R}$.

Будем называть правой конечной точкой ф.р. F_X величину

$$x_R = \sup\{x : F_X(x) < 1\} \leq \infty.$$

Заметим, что $F_X(x) < 1$ для $x < x_R$ и $F_X(x) = 1$ для $x \geq x_R$.

Будем называть левой конечной точкой ф.р. F_X

$$x_L = \inf\{x : F_X(x) > 0\} \geq -\infty,$$

т. е. $F_X(x) > 0$ для $x > x_L$ и $F_X(x) = 0$ для $x \leq x_L$. Напомним, что $\bar{F}_X(x) = 1 - F_X(x) = P(X > 0)$ — хвост ф.р. и $f_X(x)$ — плотность распределения с.в. X (когда она существует).

Для с.в. $X \geq 0$ с конечным математическим ожиданием, $MX < \infty$, определим

функцию равновесного распределения следующим образом [36, 108]:

$$F_e(x) = \frac{1}{MX} \int_0^x \bar{F}_X(t) dt, \quad x \geq 0. \quad (1.1)$$

Отметим, что равновесное распределение возникает как предельное распределение времени до ближайшего восстановления в теории надежности. Более того, знание равновесного распределения F_e позволяет получить основные свойства исходной ф.р. F_X [89].

Определение 1.1. ([40]) Пусть $X \geq 0$ – абсолютно непрерывная с.в. с плотностью распределения $f_X(x)$, и для каждого $x \in \mathbb{R}$ определена функция

$$r_X(x) = \begin{cases} \frac{f_X(x)}{\bar{F}_X(x)}, & \text{если } \bar{F}_X(x) > 0; \\ \infty, & \text{если } \bar{F}_X(x) = 0. \end{cases} \quad (1.2)$$

$r_X(x)$ называется функцией интенсивности отказов ф.р. F_X или с.в. X .

Заметим, что

$$r_X(x) = \lim_{\Delta x \downarrow 0} \frac{P(x < X \leq x + \Delta x | X > x)}{\Delta x}.$$

Функция $r_X(x)$ играет важную роль в различных приложениях. В теории надежности $r_X(x)\Delta x$ это вероятность того, что устройство с наработкой x (т. е. оно работало безотказно до момента x) откажет в интервале времени $[x, x + \Delta x]$ [1]. Тем самым функция интенсивности отказов является мерой, выражающей устойчивость устройства к отказам в зависимости от времени. В актуарной математике $r_X(x)$ используют при расчете коэффициентов смертности [56], в теории очередей $r_X(x)$ – это, например, условная вероятность того, что время ожидания заявки в очереди равно $x + \Delta x$, если она уже прождала время x [40]. Для абсолютно непрерывной с.в. X функцию интенсивности отказов можно найти из соотношения

$$r_X(x) = -\frac{d \log \bar{F}_X(x)}{dx}. \quad (1.3)$$

Более того, функция интенсивности отказов для распределений, определенных

на носителе $(0, \infty)$, обладает следующими свойствами [116]:

$$\begin{aligned} r_X(x) &\geq 0 \text{ для всех } x > 0; \\ \int_0^x r_X(t)dt &< \infty \text{ для некоторого } x > 0; \\ \int_0^\infty r_X(t)dt &= \infty; \\ \int_0^x r_X(t)dt &= \infty \text{ влечет } r_X(z) = \infty \text{ для всех } z > x. \end{aligned}$$

Пусть F_X есть ф.р. с.в. $X \geq 0$ и $F_X(0) = P(X = 0) = 0$. Для каждого u на событии $X > u$ определим с.в. $X_u = X - u \mid X > u$, называемую величиной превышения уровня u с.в. X .

Определение 1.2. ([68]) *Функцией распределения величины превышения уровня u , называется ф.р., хвост которой задается как*

$$1 - F_{X_u}(x) = \bar{F}_{X_u}(x) = P(X_u > x) = P(X > x + u \mid X > u) = \frac{\bar{F}_X(x + u)}{\bar{F}_X(u)}, \quad x \geq 0, \quad (1.4)$$

и определенная для всех $u < x_R$ таких, что $\bar{F}_X(u) > 0$.

Таким образом, с.в. $X_u = X - u \mid X > u$ задается ф.р. F_{X_u} , а если существует плотность f_X , то X_u имеет плотность

$$f_{X_u}(x) = \frac{f_X(x + u)}{\bar{F}_X(u)}, \quad x \geq 0$$

и функцию интенсивности отказов

$$r_{X_u}(x) = \frac{f_X(x + u)}{\bar{F}_X(x + u)} = r_X(x + u), \quad x \geq 0.$$

Математическое ожидание с.в. X_u равное

$$\begin{aligned} e_X(u) : &= \mathbf{M}X_u = \mathbf{M}(X - u | X > u) = \int_0^\infty \frac{\bar{F}_X(x+u)}{\bar{F}_X(u)} dx = \\ &= \frac{\int_u^\infty \bar{F}_X(x) dx}{\bar{F}_X(u)} = \frac{\int_u^\infty x f_X(x) dx}{\bar{F}_X(u)} - u \end{aligned} \quad (1.5)$$

является *средней величиной превышения уровня u* [88].

Заметим, что функция интенсивности отказов $r_X(u)$ характеризует информацию о с.в. X в интервале малой длины $(u, u + \Delta u)$, в то время как средняя величина превышения уровня u – информацию после достижения уровня u . Более того, взяв производную $e'_X(u)$ в соотношении (1.5), получим связь между этими двумя характеристиками:

$$e'_X(u) = e_X(u)r_X(u) - 1.$$

Распределения с монотонной функцией интенсивности отказов было бы естественно определить через возрастающую (убывающую) функцию r_X , но это требует существования плотности распределения, более того, в случае убывающей функции интенсивности отказов область существования u не может быть расширена до $-\infty$. Поэтому монотонность распределений по функции интенсивности отказов определяется через функцию величины превышения уровня следующим образом.

Определение 1.3. ([2]) Абсолютно непрерывное распределение F_X относится к типу распределений с возрастающей интенсивностью отказов (ВФИ-распределение), если $F_X(0) = 0$ и функция

$$\frac{F_X(x+u) - F_X(u)}{\bar{F}_X(u)}, \quad x > 0$$

является возрастающей функцией по u для всех $u \geq 0$ таких, что $F_X(u) < 1$.

Определение 1.4. ([2]) Абсолютно непрерывное распределение F_X относится к типу распределений с убывающей интенсивностью отказов (УФИ-

распределение), если $F_X(0) = 0$ и

$$\frac{F_X(x+u) - F_X(u)}{\bar{F}_X(u)}, \quad x > 0$$

является убывающей функцией по u для всех $u \geq 0$ таких, что $F_X(u) < 1$.

Следующие два утверждения эквивалентны определению 1.3, но более удобны для проверки соответствующих условий.

Утверждение 1.1. ([116]) *Распределение F_X является ВФИ (УФИ)-распределением тогда и только тогда, когда для всех u таких, что $\bar{F}_X(u) > 0$, функция $\bar{F}_{X_u}$ убывает (возрастает) по u для всех $x \geq 0$.*

Следующее утверждение верно лишь при условии существования плотности распределения на положительной полуоси.

Утверждение 1.2. ([2]) *Пусть с.в. $X \geq 0$ задана ф.р. $F_X, F_X(0) = 0$ и существует $f_X(x) = F'_X(x)$ для $x > 0$. Тогда F_X является ВФИ (УФИ)-распределением тогда и только тогда, когда функция интенсивности отказов $r_X(x)$ есть возрастающая (убывающая) функция.*

Примеров распределений, имеющих монотонные функции интенсивности отказов, достаточно много. Типичные примеры УФИ-распределений: распределение Парето, гиперэкспоненциальное и логнормальное распределения. Равномерное непрерывное распределение относится к классу ВФИ-распределений. Характер монотонности функции интенсивности отказов распределения Вейбулла и гамма-распределения зависит от значения параметра степени. При параметре степени меньше (больше) единицы эти распределения относятся к классу УФИ (ВФИ)-распределений. Тем не менее, помимо распределений с постоянными и монотонными интенсивностями отказов есть и другие классы распределений. Например, функция интенсивности отказов может сначала убывать (возрастать), затем принимать постоянное значение, и далее снова возрастать (убывать). В этом случае ее называют выпуклой вниз функцией интенсивности отказов, ВНФИ. (Для выпуклой вверх функции, соответственно, сокращение ВВФИ.)

Определение 1.5 ([106, 173]). *Распределение F_X имеет выпуклую вниз функцию интенсивности отказов, ВНФИ (выпуклую вверх функцию интенсивности отказов, ВВФИ), если эта функция имеет следующий вид:*

$$r_X(x) = \begin{cases} r_1(x), & r'_1(x) < (>) 0, \quad x \leq a_1; \\ r_2(x), & r'_2(x) = 0, \quad a_1 < x \leq a_2; \\ r_3(x), & r'_3(x) > (<) 0, \quad x > a_2, \end{cases} \quad (1.6)$$

где константы $a_1 \leq a_2$. Если $a_1 = a_2$, то ВНФИ называется U-образной.

1.2. Стохастический порядок и упорядоченность по интенсивности отказов

Рассмотрим две с.в. X и Y с ф.р. $F_X(x)$ и $F_Y(x)$.

Определение 1.6. ([116]) *С.в. X стохастически меньше с.в. Y (меньше по распределению), если*

$$\bar{F}_X(x) \leq \bar{F}_Y(x), \quad x \in \mathbb{R}. \quad (1.7)$$

Стохастический порядок будем обозначать $X \underset{st}{\leq} Y$.

Определение 1.7. ([116]) *С.в. X почти наверное меньше с.в. Y , если*

$$P(X \leq Y) = 1. \quad (1.8)$$

Упорядоченность с вероятностью 1 (почти наверное) будем обозначать

$$X \leq Y \text{ с в. 1.}$$

Определение 1.8. ([145]) *Пусть $r_X(x)$ и $r_Y(x)$ – функции интенсивности отказов с.в. X и Y ($X, Y \geq 0$), соответственно. Будем говорить, что с.в. X меньше с.в. Y по интенсивности отказов, если выполнено следующее соотношение:*

$$r_X(x) \geq r_Y(x), \quad x \in \mathbb{R}. \quad (1.9)$$

Упорядоченность по интенсивности отказов будем обозначать

$$X \underset{r}{\leq} Y.$$

Далее везде будем рассматривать положительные с.в., что естественно для моделирования характеристик систем обслуживания и охватывает все рассматриваемые в работе примеры.

Следующее утверждение показывает, что упорядоченность по интенсивности отказов влечет стохастический порядок, а из последнего вытекает упорядоченность в среднем.

Утверждение 1.3. ([24, 145]) *Если $MX^k, MY^k, 1 \leq k < \infty$, то выполнены следующие соотношения:*

$$X \underset{r}{\leq} Y \text{ влечет } X \underset{st}{\leq} Y \text{ влечет } MX^k \leq MY^k. \quad (1.10)$$

Соотношения (1.10) используются далее при исследовании свойств монотонности систем обслуживания.

Приведем некоторые примеры.

Пример 1.1. С.в. $X \sim \text{Exp}(\lambda)$, т. е.

$$F_X(x) = 1 - e^{-\lambda x}, \quad x \geq 0, \quad (1.11)$$

имеет функцию интенсивности отказов $r_X(x) = \lambda$ и математическое ожидание $MX = 1/\lambda$.

Пример 1.2. Пусть две с.в. имеют показательное распределение с параметрами λ_1 и λ_2 , соответственно. При $\lambda_1 \geq \lambda_2$ функции интенсивности отказов связаны неравенством $r_{X_1} \geq r_{X_2}$, т. е. $X_1 \underset{r}{\leq} X_2$.

Пример 1.3. Пусть с.в. $Y \sim \text{Pareto}(\alpha, x_0)$ с ф.р. [87]

$$F_Y(x) = 1 - \left(\frac{x_0}{x_0 + x} \right)^\alpha, \quad x \geq 0, \quad x_0 > 0, \quad \alpha > 0. \quad (1.12)$$

Функция интенсивности отказов имеет вид

$$r_Y(x) = \frac{\alpha}{x_0 + x}, \quad x \geq 0, \quad (1.13)$$

и убывает при $x \rightarrow \infty$, а математическое ожидание равно

$$MY = \frac{x_0}{\alpha - 1}, \quad \alpha > 1. \quad (1.14)$$

Пример 1.4. Пусть $Y_1 \sim Pareto(\alpha_1, x_0^1)$ и $Y_2 \sim Pareto(\alpha_2, x_0^2)$. Тогда для упорядоченности по интенсивности отказов, $Y_1 \leq_r Y_2$, достаточно выполнения следующих двух условий:

$$\alpha_1 \geq \alpha_2, \quad x_0^1 \leq x_0^2. \quad (1.15)$$

Пример 1.5. Пусть заданы две с.в. X и Y с распределениями (1.11) и (1.12). Если выполнено неравенство

$$\lambda \geq \frac{\alpha}{x_0}, \quad (1.16)$$

то с.в. упорядочены по интенсивности отказов, $X \leq_r Y$.

1.3. Каплинг случайных величин и стохастическая упорядоченность

Каплинг представляет собой совместное распределение с.в., заданных, вообще говоря, на разных вероятностных пространствах так, что маргинальные распределения совпадают с распределениями исходных с.в. [66, 109, 171].

Определение 1.9. ([171]) С.в. X' является копией с.в. X , если она имеет то же самое распределение, $X \stackrel{st}{=} X'$, т. е.

$$F_X(x) = F_{X'}(x), \quad x \geq 0. \quad (1.17)$$

Определение 1.10. ([171]) Каплингом семейства с.в. $(X_i, i \in J)$, где J - некоторое множество индексов, называется набор копий $(X'_i, i \in J)$, удовлетворяющих соотношению (1.17) для всех $i \in J$.

Заметим, что в определении каплинга используется покомпонентное равенство распределений, в то время как совместное распределение каплинга $(X'_i, i \in J)$ может отличаться от совместного распределения $(X_i, i \in J)$.

Определим *квантиль-функцию* для некоторой (исходной) ф.р. F_X как

$$F_X^{-1}(u) = \inf\{x \in R : F_X(x) \geq u\}, \quad u \in [0, 1].$$

Если F_X – непрерывная, строго возрастающая функция, то обратная функция F_X^{-1} существует и тоже строго возрастает. Если с.в. U равномерно распределена на $[0, 1]$, тогда с.в. $X' = F_X^{-1}(U)$ является копией с.в. X . Действительно,

$$P(X' \leq x) = P(F_X^{-1}(U) \leq x) = P(U \leq F_X(x)) = F_X(x), \quad x \in \mathbb{R}.$$

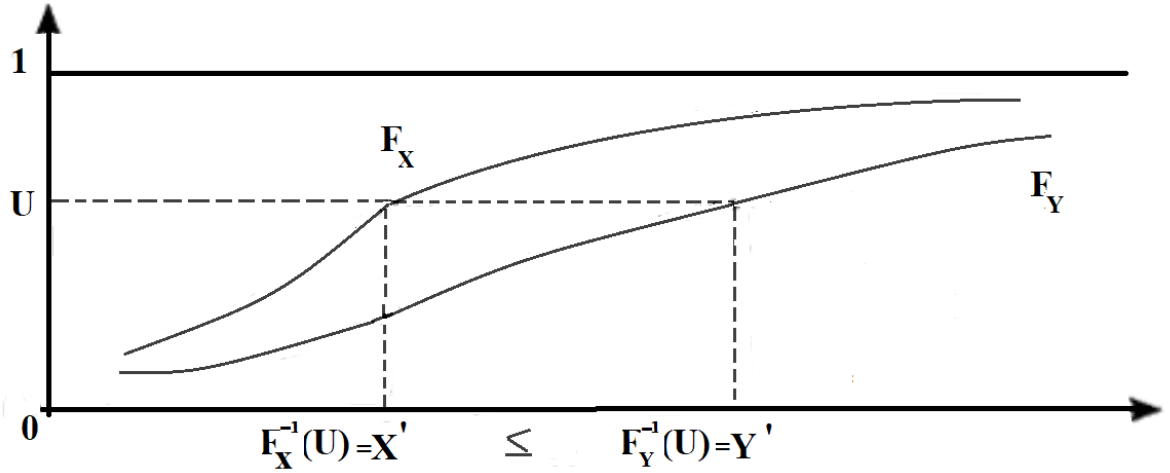


Рисунок 1.1. Стохастический порядок и порядок почти наверное

Если с.в. X стохастически меньше, чем с.в. Y , т. е. $\bar{F}_X(x) \leq \bar{F}_Y(x)$, то из неравенства $F_Y(x) \geq u$ следует, что $F_X(x) \geq u$ и,

$$\{x \in \mathbb{R} : F_Y(x) \geq u\} \subseteq \{x \in \mathbb{R} : F_X(x) \geq u\}.$$

Таким образом, $F_X^{-1}(u) \leq F_Y^{-1}(u)$, что влечет $F_X^{-1}(U) \leq F_Y^{-1}(U)$ (см. Рисунок 1.1). Эти выводы суммированы в следующем Утверждении.

Утверждение 1.4. ([171]) *Две с.в. X и Y стохастически упорядочены,*

$$X \underset{st}{\leq} Y,$$

тогда и только тогда, когда существует каплинг (X', Y') с.в. X и Y такой, что

$$P(X' \leq Y') = 1.$$

Утверждение 1.4 дает связь определений стохастического порядка, порядка почти наверное и каплинга и будет использовано для доказательства Теорем 1.1 и 1.2 ниже.

1.4. Достаточные условия упорядоченности по интенсивности отказов

Следующая Лемма 1.1 дает достаточное условие упорядоченности распределений по интенсивности отказов, которое более удобно для проверки монотонности, чем проверка стохастического порядка. Более того, как показано далее, оно является необходимым для сравнения систем обслуживания с конечным буфером.

Лемма 1.1. Пусть r_X, r_Y – функции интенсивности отказов распределений F_X, F_Y , соответственно. Тогда, если

$$\sup_{x \geq 0} r_X(x) < \infty, \quad \inf_{x \geq 0} r_Y(x) > 0 \quad (1.18)$$

и

$$\sup_{x \geq 0} r_X(x) \leq \inf_{x \geq 0} r_Y(x), \quad (1.19)$$

то

$$X \underset{r}{\geq} Y.$$

Сформулируем условия, при которых соотношения (1.18) выполняются для различных классов распределений.

Лемма 1.2. Условие (1.18) выполнено для с.в. X и Y , если

а) для ф.р. F_X выполнено одно из следующих условий:

(а1) F_X есть УФИ-распределение и

$$\sup_x r_X(x) = \lim_{u \rightarrow x_L} r_X(u) < \infty;$$

(а2) F_X есть ВФИ-распределение и

$$\sup_x r_X(x) = \lim_{u \rightarrow x_R} r_X(u) < \infty;$$

(а3) F_X есть ВВФИ-распределение и

$$\sup_x r_X(x) = \max_{u \in (a_1, a_2]} r_X(u) < \infty,$$

где a_1, a_2 определены соотношением (1.6);

(а4) функция интенсивности отказов есть постоянная величина,

$$r_X(x) = \text{const};$$

б) для распределения F_Y выполнено одно из следующих условий:

(б1) F_Y есть УФИ-распределение и

$$\inf_x r_Y(x) = \lim_{u \rightarrow x_R} r_Y(u) > 0;$$

(б2) F_Y есть ВФИ-распределение и

$$\inf_x r_Y(x) = \lim_{u \rightarrow x_L} r_Y(u) > 0;$$

(б3) F_Y есть ВНФИ-распределение и

$$\inf_x r_Y(x) = \min_{x \in (a_1, a_2]} r_Y(u) > 0,$$

где a_1, a_2 определены соотношением (1.6);

(б4) функция интенсивности отказов есть постоянная величина,

$$r_Y(x) = \text{const}.$$

Приведем примеры, иллюстрирующие применение Лемм 1.1 и 1.2.

Пример 1.6. Пусть с.в. X имеет показательное распределение с параметром $\lambda > 0$ и параметром сдвига $a > 0$, так что плотность имеет вид

$$f_X(x) = \begin{cases} \lambda e^{-\lambda(x-a)}, & x \geq a; \\ 0, & x < a. \end{cases}$$

В данном случае функция интенсивности отказов – постоянная величина $r_X(x) = \lambda, x \geq a$.

Пусть с.в. Y имеет равномерное распределение на отрезке $[a, b]$ с плотно-

СТЬЮ

$$f_Y(x) = \begin{cases} \frac{x-a}{b-a}, & x \in [a, b]; \\ 0, & x \notin [a, b]. \end{cases}$$

Функция интенсивности отказов с.в. Y возрастает и имеет вид

$$r_Y(x) = \frac{1}{b-x}, \quad x \in [a, b),$$

причем,

$$\inf_x r_Y(x) = r_Y(a) = \frac{1}{b-a}.$$

Поэтому условие (1.19) примет вид

$$\inf_x r_Y(x) = 1/(b-a) \geq \sup_x r_X(x) = \lambda.$$

Следовательно, если $\lambda(b-a) \leq 1$, то $X \underset{r}{\geq} Y$ и, согласно условию (1.10), $X \underset{st}{\geq} Y$. Заметим, что соотношение для стохастического порядка в явном виде в данном случае выписать нельзя.

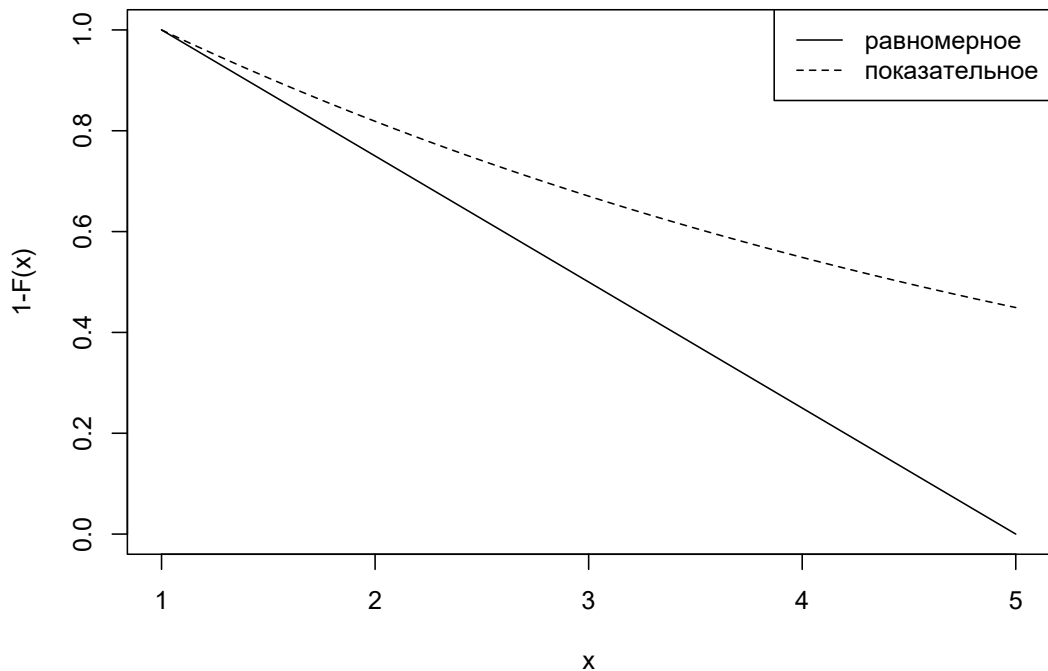


Рисунок 1.2. Хвосты ф.р. показательного и равномерного распределений, $\lambda = 0.2$, $a = 1$, $b = 5$.

Рисунок 1.2 демонстрирует стохастический порядок с.в., имеющих показательное распределение с параметром $\lambda = 0.2$ и равномерное распределение на отрезке $[1, 5]$, соответственно. При таких параметрах условие $\lambda(b-a) \leq 1$

выполнено.

Пример 1.7. Пусть с.в. X имеет распределение Парето $Pareto(\alpha, x_0)$, а с.в. Y показательная с ф.р. $F_Y(x) = 1 - e^{-\lambda x}$, $x \geq 0$. Заметим, что

$$\sup_x r_X(x) = r_X(0) = \frac{\alpha}{x_0} \leq \lambda = \inf_x r_Y(x).$$

Следовательно, если

$$\lambda \geq \alpha/x_0, \quad \text{то} \quad r_X(x) \leq r_Y(x), \quad x \geq 0$$

и из $X \underset{r}{\geq} Y$ следует $X \underset{st}{\geq} Y$. Рисунок 1.3 демонстрирует упорядоченность функций интенсивности отказов при $\lambda = 4.2$; $x_0 = 0.5$; $\alpha = 2.1$.

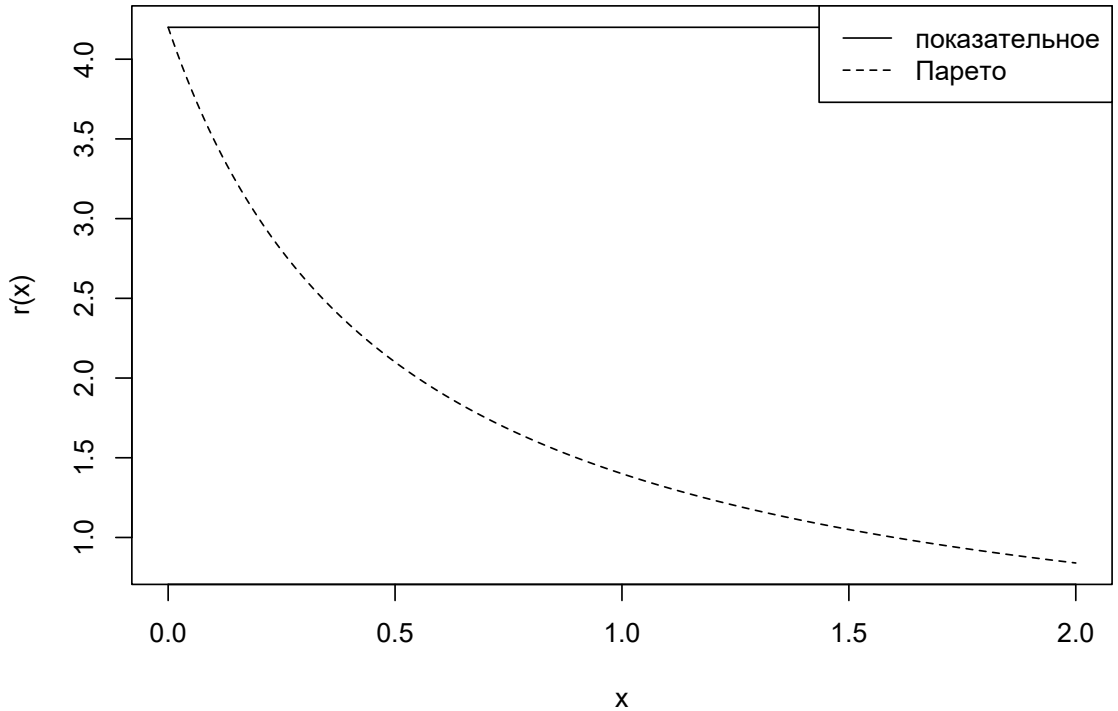


Рисунок 1.3. Функции интенсивности отказов показательного и Парето распределений, $\lambda = 4.2$; $x_0 = 0.5$; $\alpha = 2.1$.

Пример 1.8. Пусть с.в. Y имеет распределение Вейбулла, $Y \sim We(\alpha, \beta)$ с ф.р. вида

$$F_Y(x) = 1 - e^{-(x/\alpha)^\beta}, \quad x \geq 0, \quad \alpha > 0, \quad \beta > 0,$$

и монотонной функцией интенсивностей отказов

$$r_Y(x) = \frac{\beta}{\alpha} \left(\frac{x}{\alpha} \right)^{\beta-1}, \quad x \geq 0. \quad (1.20)$$

Распределение Вейбулла является УФИ-распределением при $\beta < 1$ или ВФИ-распределением при $\beta > 1$ (т.к. r_Y возрастает при $\beta > 1$ и убывает при $\beta < 1$). Для того, чтобы неравенство (1.19) было выполнено необходимо, чтобы были выполнены условия а1) или б1) Леммы 1.2 при $\beta < 1$, либо условия а2) и б2) при $\beta > 1$, что невозможно, поскольку

$$\inf_x r_Y(x) = 0, \quad \sup_x r_Y(x) = \infty, \quad \beta > 0, \beta \neq 1.$$

Следовательно, неравенство (1.19) не может быть выполнено как для случая УФИ-распределения, так и ВФИ-распределения Вейбулла.

1.5. Функция интенсивности отказов распределений с длинными хвостами

Определение 1.11. ([87]) *Ф.р. F_X имеет длинный хвост, если для любого $x > 0$ выполнено соотношение*

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

Это асимптотическое свойство означает, что на событии $X > u$ для больших u , с.в. X превзойдет также пороговое значение $x + u$, для любого $x > 0$.

УФИ-распределения в литературе часто отождествляют с ф.р. с длинными хвостами. Приведем примеры, в которых УФИ-распределения не имеют длинного хвоста.

Распределения с длинными хвостами имеют функции интенсивностей отказов, асимптотически убывающие до нуля (не обязательно монотонно) [87], т.е.

$$\lim_{x \rightarrow \infty} r_X(x) = 0. \tag{1.21}$$

Поэтому функции интенсивностей отказов таких распределений нельзя использовать в правой части неравенства (1.19).

Пример 1.9. Важным примером распределения с длинным хвостом является распределение Парето (1.12), функция интенсивности отказов которого моно-

тонно убывает до нуля,

$$r(x) = \frac{\alpha}{x_0 + x} \rightarrow 0, \text{ при } x \rightarrow \infty. \quad (1.22)$$

Пример 1.10. Рассмотрим обобщение распределения Парето – распределение Бурра [187], обозначаемое $Burr(x_0, \alpha, k)$

$$F_X(x) = 1 - \left(\frac{x_0}{x_0 + x^k} \right)^\alpha, \quad x \geq 0, \quad x_0 > 0, \quad \alpha > 0, \quad k > 0, \quad (1.23)$$

и имеющее функцию интенсивности отказов

$$r_X(x) = \frac{\alpha k x^{k-1}}{x_0 + x^k}, \quad x \geq 0, \quad x_0 > 0, \quad \alpha > 0, \quad k > 0. \quad (1.24)$$

Заметим, что при $k \geq 2$ функции интенсивности отказов сначала возрастает, достигая максимума

$$r_X(x^*) = \frac{\alpha k (x_0(k-1))^{1-1/k}}{x_0 k} \quad (1.25)$$

в точке

$$x^* = (x_0(k-1))^{1/k},$$

а при $x > x^*$ убывает. Таким образом, распределение Бурра относится к классу ВВФИ-распределений для $k \geq 2$. При $k < 1$ функция интенсивности отказов монотонно убывает, но тем не менее, свойство (1.21) выполняется для любого $k > 0$.

Заметим, что свойство (1.21) не обязательно выполняется для УФИ-распределений. Рассмотрим пример УФИ-распределения с функцией интенсивности отказов, сходящейся в правой конечной точке к некоторому положительному числу.

Пример 1.11. Применим Лемму 1.1 для нахождения условия на параметры, при котором гиперэкспоненциальное распределение и распределение Бурра связаны отношением порядка. Пусть с.в. X имеет гиперэкспоненциальное распределение

$$F_X(x) = 1 - \sum_{i=1}^n p_i e^{-\mu_i x}, \quad x \geq 0, \quad (1.26)$$

где

$$p_i \geq 0, \mu_i > 0, i = 1, \dots, n; p_1 + \dots + p_n = 1.$$

Функция интенсивности отказов имеет следующий вид:

$$r_X(x) = \frac{\sum_{i=1}^n p_i \mu_i e^{-\mu_i x}}{\sum_{i=1}^n p_i e^{-\mu_i x}}, \quad x \geq 0. \quad (1.27)$$

Гиперэкспоненциальное распределение относится к классу распределений с легким хвостом, при этом функция интенсивности отказов $r_X(x)$ не сходится к нулю, хотя и монотонно убывает:

$$\lim_{x \rightarrow \infty} r_X(x) = \min(\mu_1, \dots, \mu_n) =: \mu^*. \quad (1.28)$$

Предположим, что с.в. $Y \sim Burr(x_0, \alpha, k)$ с ф.р. вида (1.23) при $k \geq 2$. Пусть с.в. X имеет гиперэкспоненциальное распределение вида (1.26). Зафиксируем параметры $x_0, k, \mu_i, p_i, i = 1, \dots, n$ и найдем значение α^* , гарантирующее порядок

$$X \underset{r}{\leq} Y \quad \text{для} \quad \alpha \leq \alpha^*.$$

Для этого необходимо найти единственное решение уравнения $r_X(x) = r_Y(x)$ для $x > 0$, что возможно лишь, если $n \neq 1$. Заметим, что поскольку из выражения (1.25) следует

$$r_Y(x) \leq r_Y(x^*), \quad x \geq 0,$$

то из неравенства

$$r_Y(x^*) = \frac{\alpha k (x_0(k-1))^{1-1/k}}{x_0 k} \leq \mu^*,$$

следует $r_Y(x) \leq r_X(x), x \geq 0$. Пусть

$$\alpha^* = \frac{x_0 k}{k(x_0(k-1))^{1-1/k}} \mu^*. \quad (1.29)$$

Тогда условие

$$\alpha \leq \alpha^* \quad (1.30)$$

гарантирует выполнение неравенства $r_Y(x) \leq r_X(x), x \geq 0$. На Рисунке 1.4

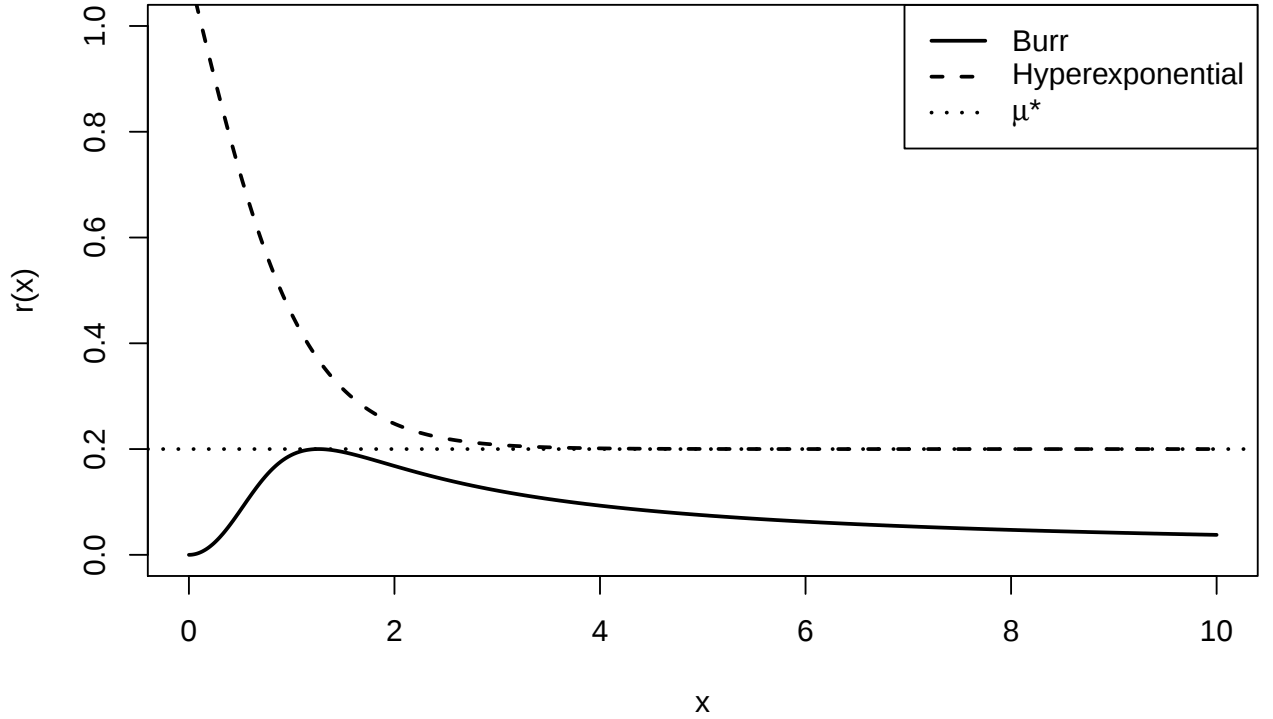


Рисунок 1.4. Сравнение интенсивностей отказов распределения $Burr(x_0, \alpha^*, k)$ и гиперэкспоненциального распределения с параметрами $x_0 = 1, k = 3, n = 2, p_1 = p_2 = 0.5, \mu_1 = 0.2, \mu_2 = 2$ и значением $\alpha^* = 0.12$, удовлетворяющим условию (1.29).

приведено сравнение функций интенсивностей отказов гиперэкспоненциального распределения и распределения Бурра с параметрами $x_0 = 1, k = 3, n = 2, p_1 = p_2 = 0.5, \mu_1 = 0.2, \mu_2 = 2$. Таким образом, свойство (1.19) выполнено, так как $\inf_x r_X(x) = \mu^*$ и

$$\sup_x r_Y(x) = r_Y(x^*),$$

где x^* удовлетворяет соотношению (1.25).

Вышеприведенные примеры показывают, что распределения с длинным хвостом не обязательно имеют монотонно убывающую функцию интенсивности отказов, а также что существуют УФИ-распределения, не обладающие длинным хвостом.

1.6. Функция интенсивности отказов конечной смеси распределений

Рассмотрим семейство с.в. $X_1, \dots, X_n$, заданных ф.р. $F_{X_1}, \dots, F_{X_n}$, соответственно.

Определение 1.12. ([117]) С.в. X называется n -компонентной смесью распределений, если ее ф.р. F_X имеет следующий вид:

$$F_X(x) = p_1 F_{X_1}(x) + \cdots + p_n F_{X_n}(x), \quad (1.31)$$

где вероятность p_i называется i -м коэффициентом пропорции смеси, $\sum_{i=1}^n p_i = 1$, $p_i > 0$, $i = 1, \dots, n$.

Функция интенсивности отказов смеси равна

$$r_X(x) = \frac{p_1 f_{X_1}(x) + \cdots + p_n f_{X_n}(x)}{p_1 \bar{F}_{X_1}(x) + \cdots + p_n \bar{F}_{X_n}(x)}, \quad (1.32)$$

где f_{X_i} – плотность i -й компоненты X_i смеси. В частности, для двухкомпонентной смеси с ф.р.

$$F_X(x) = p F_{X_1}(x) + (1 - p) F_{X_2}(x), \quad p \in (0, 1), \quad (1.33)$$

функция интенсивности отказов (1.32) принимает вид

$$r_X(x) = \frac{p f_{X_1}(x) + (1 - p) f_{X_2}(x)}{p \bar{F}_{X_1}(x) + (1 - p) \bar{F}_{X_2}(x)} = \pi(x) r_{X_1}(x) + (1 - \pi(x)) r_{X_2}(x), \quad (1.34)$$

где

$$\pi(x) := \frac{p \bar{F}_{X_1}(x)}{p \bar{F}_{X_1}(x) + (1 - p) \bar{F}_{X_2}(x)},$$

а математическое ожидание равно

$$MX = pMX_1 + (1 - p)MX_2.$$

Предположим, что с.в. X_1 , X_2 независимы и I – дискретная с.в., независимая от (X_1, X_2) , с распределением

$$I = \begin{cases} 1, & \text{с вероятностью } p; \\ 0, & \text{с вероятностью } 1 - p. \end{cases}$$

Тогда распределение с.в.

$$X = I X_1 + (1 - I) X_2 \quad (1.35)$$

имеет вид (1.33) двухкомпонентной смеси. Известно [116], что для любого $x \geq 0$

$$\min(r_{X_1}(x), r_{X_2}(x)) \leq r_X(x) \leq \max(r_{X_1}(x), r_{X_2}(x)). \quad (1.36)$$

Функция интенсивности отказов смеси $r_X(x)$ в пределе с ростом x соответствует наименьшей из функций интенсивностей отказов компонент [116], т. е.

$$\lim_{x \rightarrow \infty} \frac{r_X(x)}{\min(r_{X_1}(x), r_{X_2}(x))} = 1. \quad (1.37)$$

Приведем примеры конечных смесей распределений важных для задач, рассматриваемых в диссертационной работе.

Пример 1.12. Пусть с.в. X_1, X_2 имеют показательное распределение, $X_i \sim \text{Exp}(\lambda_i)$, $i = 1, 2$. Тогда их смесь имеет гиперэкспоненциальное распределение вида (1.26) с убывающей функцией интенсивности отказов,

$$r_X(x) = \frac{p\lambda_1 e^{-\lambda_1 x} + (1-p)\lambda_2 e^{-\lambda_2 x}}{pe^{-\lambda_1 x} + (1-p)e^{-\lambda_2 x}} \xrightarrow{x \rightarrow \infty} \min(\lambda_1, \lambda_2).$$

Напомним, что гиперэкспоненциальное распределение относится к классу ВФИ-распределений, при этом i -я компонента имеет функцию интенсивности отказов $r_{X_i}(x) = \lambda_i$, $i = 1, 2$.

Пример 1.13. Рассмотрим смесь двух распределений Парето вида

$$Z = IZ_1 + (1 - I)Z_2, \quad (1.38)$$

где $Z_i \sim \text{Pareto}(\alpha_i, x_0)$, так что хвост ф.р. Z имеет вид

$$\bar{F}_Z(x) = p \left(\frac{x_0}{x_0 + x} \right)^{\alpha_1} + (1 - p) \left(\frac{x_0}{x_0 + x} \right)^{\alpha_2}. \quad (1.39)$$

Известно [116], что смесь двух УФИ-распределений является УФИ-распределением. Действительно, для смеси (1.38) функция интенсивности отказов вычисляется по формуле

$$r_Z(x) = \frac{pr_{Z_1}(x)a(x) + (1-p)r_{Z_2}(x)}{pa(x) + (1-p)}, \quad (1.40)$$

где

$$a(x) = x_0^{\alpha_1 - \alpha_2} (x_0 + x)^{\alpha_2 - \alpha_1},$$

и кроме того, $r_Z(x) \rightarrow 0$, $x \rightarrow \infty$. Поскольку производная

$$r'_Z(x) = -(1-p) \frac{pa(x)(r_{Z_1}(x) - r_{Z_2}(x))^2 + r_{Z_2}^2(x)/\alpha(pa(x) + (1-p))}{(pa(x) + (1-p))^2} < 0,$$

то r_Z убывает и F_Z является УФИ-распределением.

1.7. Стохастический порядок и упорядоченность по интенсивности отказов двухкомпонентных смесей

Следующая Теорема 1.1 демонстрирует стохастический порядок и упорядоченность по интенсивности отказов двух компонент и их смеси. Подробный анализ двухкомпонентной смеси существенно облегчает дальнейший анализ смеси с произвольным (конечным) числом компонент, приведенный в заключительных разделах данной главы. Также отметим, что полученные результаты играют важную роль в анализе монотонности систем обслуживания в Главе 2.

Теорема 1.1. *Если с.в. X является смесью с.в. X_1 и X_2 вида (1.35), то*

$$X_1 \underset{st}{\leq} X_2 \Rightarrow X_1 \underset{st}{\leq} X \underset{st}{\leq} X_2; \quad (1.41)$$

$$X_1 \underset{r}{\leq} X_2 \Rightarrow X_1 \underset{r}{\leq} X \underset{r}{\leq} X_2. \quad (1.42)$$

Доказательство. Пусть (X'_1, X'_2, I') – каплинг с.в. X_1, X_2, I такой, что и $I \overset{st}{=} I'$

$$X'_1 \leq X'_2 \text{ с в. 1.}$$

Тогда $X' = I'X'_1 + (1 - I')X_2$ есть стохастическая копия с.в. X , и

$$X'_1 \leq X', \quad X' \leq X_2 \text{ с в. 1.}$$

Тогда из Утверждения 1.4 следует стохастический порядок (1.41)

$$X_1 \underset{st}{\leq} X \underset{st}{\leq} X_2.$$

Теперь докажем упорядоченность смеси по интенсивности отказов. Из стохастического порядка $X_1 \leq_{st} X_2$ следует, что $\bar{F}_{X_1}(x) \leq \bar{F}_{X_2}(x)$, $x \geq 0$. Тогда очевидно,

$$\bar{F}_{X_1}(x) \leq p\bar{F}_{X_1}(x) + (1-p)\bar{F}_{X_2}(x) \leq \bar{F}_{X_2}(x).$$

С другой стороны, упорядоченность по интенсивности отказов $X_1 \leq_r X_2$ означает

$$f_{X_2}(x) \leq \frac{f_{X_1}(x)\bar{F}_{X_2}(x)}{\bar{F}_{X_1}(x)},$$

и, следовательно,

$$r_X(x) \leq \frac{pf_{X_1}(x) + (1-p)f_{X_1}(x)\bar{F}_{X_2}(x)/\bar{F}_{X_1}(x)}{p\bar{F}_{X_1}(x) + (1-p)\bar{F}_{X_2}(x)} = r_{X_1}(x), \quad x \geq 0,$$

откуда следует неравенство $X_1 \leq_r X$. Аналогично, из неравенств

$$f_{X_1}(x) \geq \frac{f_{X_2}(x)\bar{F}_{X_1}(x)}{\bar{F}_{X_2}(x)}$$

и

$$r_X(x) \geq \frac{pf_{X_2}(x)\bar{F}_{X_1}(x)/\bar{F}_{X_2}(x) + (1-p)f_{X_2}(x)}{p\bar{F}_{X_1}(x) + (1-p)\bar{F}_{X_2}(x)} = r_{X_2}(x), \quad x \geq 0,$$

следует, что $X \leq_r X_2$, что доказывает (1.42). $\square$

Заметим, что утверждения Теоремы 1.1 можно также получить из соотношения (1.36). Другое доказательство утверждения (1.42) приведено в Теореме 1.B.22 [156].

Сравним две смеси с одинаковым коэффициентом пропорции p :

$$\begin{aligned} F_X(x) &= pF_{X_1}(x) + (1-p)F_{X_2}(x), \\ F_Y(x) &= pF_{Y_1}(x) + (1-p)F_{Y_2}(x). \end{aligned}$$

Следующая Теорема 1.2 приведена в работе [116, Proposition 3F1] без детального доказательства. В данной работе она доказывается методом каплинга.

Теорема 1.2. *Предположим, что компоненты смесей стохастически упорядочены*

$$X_1 \leq_{st} Y_1, \quad X_2 \leq_{st} Y_2. \quad (1.43)$$

Если с.в. X_1, Y_1, X_2, Y_2 независимы, то их смеси

$$X = IX_1 + (1 - I)X_2 \text{ и } Y = IY_1 + (1 - I)Y_2 \quad (1.44)$$

стохастически упорядочены:

$$X \underset{st}{\leq} Y. \quad (1.45)$$

Доказательство. Пусть (X'_1, Y'_1, I') есть каплинг с.в. (X_1, Y_1, I) , т.е. $X_1 \overset{st}{=} X'_1$, $Y_1 \overset{st}{=} Y'_1$, $I \overset{st}{=} I'$, такой, что

$$P(X'_1 \leq Y'_1) = 1.$$

Построим каплинг (X'_2, Y'_2, I') с.в. (X_2, Y_2, I) , т.е. $X_2 \overset{st}{=} X'_2$, $Y_2 \overset{st}{=} Y'_2$, такой, что

$$P(X'_2 \leq Y'_2) = 1.$$

Предположим, что пары (X'_1, Y'_1) и (X'_2, Y'_2) независимы. Тогда $X' := I'X'_1 + (1 - I')X'_2$ есть стохастическая копия с.в. X , а $Y' := I'Y'_1 + (1 - I')Y'_2$ – стохастическая копия с.в. Y . Таким образом, (X', Y') есть каплинг (X, Y) , для которого верно

$$P(X' \leq Y') = 1.$$

Следовательно, из Утверждения 1.4 следует соотношение (1.45). $\square$

Для того, чтобы сравнить смеси по интенсивности отказов потребуются дополнительные условия. Следующая теорема (модификация Теоремы 2.1 в [31]) позволяет сравнить две смеси по интенсивности отказов.

Теорема 1.3. *Предположим, что для двух смесей X и Y вида (1.44) выполнены условия:*

$$X_1 \underset{r}{\geq} X_2 \text{ или } Y_1 \underset{r}{\geq} Y_2; \quad (1.46)$$

$$\min(X_1, Y_2) \underset{st}{\leq} \min(X_2, Y_1); \quad (1.47)$$

$$X_1 \underset{r}{\leq} Y_1, \quad X_2 \underset{r}{\leq} Y_2. \quad (1.48)$$

Тогда смеси (1.44) упорядочены по интенсивности отказов, т. е.

$$X \underset{r}{\leq} Y. \quad (1.49)$$

Доказательство. Пусть $X_1 \underset{r}{\geq} X_2$. Из неравенства (1.47) следует, что

$$\frac{\bar{F}_{X_2}(x)\bar{F}_{Y_1}(x)}{\bar{F}_{X_1}(x)\bar{F}_{Y_2}(x)} \geq 1, \quad x \geq 0.$$

Из неравенств (1.48) следует, что

$$f_{X_1}(x) \geq f_{Y_1}(x) \frac{\bar{F}_{X_1}(x)}{\bar{F}_{Y_1}(x)}$$

и

$$f_{X_2}(x) \geq f_{Y_2}(x) \frac{\bar{F}_{X_2}(x)}{\bar{F}_{Y_2}(x)}.$$

Следовательно,

$$r_X(x) = \frac{pf_{X_1} + (1-p)f_{X_2}}{p\bar{F}_{X_1} + (1-p)\bar{F}_{X_2}} \geq \frac{\bar{F}_{X_2}}{\bar{F}_{Y_1}} \cdot \frac{pf_{Y_1} + (1-p)f_{Y_2} \frac{\bar{F}_{X_2}\bar{F}_{Y_1}}{\bar{F}_{X_1}\bar{F}_{Y_2}}}{p\bar{F}_{Y_1} + (1-p)\bar{F}_{Y_2}} \geq r_Y(x), \quad x \geq 0,$$

и соотношение (1.49) выполнено. $\square$

Результаты Теорем 1.2 и 1.3 позволяют исследовать свойства монотонности систем обслуживания, где в качестве распределений интервалов между приходами заявок или времен обслуживания выступают различные двухкомпонентные смеси распределений. При этом сравнение распределений по интенсивности отказов очень часто позволяет найти соотношения между параметрами, гарантирующие стохастический порядок, а, следовательно, монотонность исследуемых характеристик систем.

Теорема 1.4. *Пусть*

$$\bar{F}_X(x) = p\bar{F}_{X_1}(x) + (1-p)\bar{F}_{X_2}(x),$$

где F_{X_1} и F_{X_2} являются УФИ распределениями. Пусть величина превышения уровня u для смеси $X_u = X - u | X > u$ задана хвостом ф.р.

$$\bar{F}_{X_u}(x) = \frac{\bar{F}_X(x+u)}{\bar{F}_X(u)}.$$

Тогда для всех u таких, что $\bar{F}_{X_u}(x) > 0$, $\bar{F}_{X_u}(x)$ возрастает по u для всех $x \geq 0$.

Доказательство. Известно, что смесь двух УФИ-распределений является УФИ-распределением [116, Proposition 2.C.14]. Следовательно, F_X относится к классу УФИ распределений. С другой стороны, для любого УФИ распределения ф.р. величины превышения уровня u возрастает по u [156, Theorem 1.A.30], следовательно, $\bar{F}_{X_u}(x)$ возрастает по u . $\square$

Пример 1.14. Легко проверить, что для двухкомпонентной смеси распределений Парето вида (1.39) хвост ф.р. величины превышения уровня u

$$\bar{F}_{Z_u}(x) = \left(\frac{x_0 + u}{x_0 + x + u} \right)^{\alpha_1} \frac{px_0^{\alpha_1 - \alpha_2} + (1 - p)(x_0 + x + u)^{\alpha_1 - \alpha_2}}{px_0^{\alpha_1 - \alpha_2} + (1 - p)(x_0 + u)^{\alpha_1 - \alpha_2}},$$

возрастает по u .

1.8. Показательное-Парето распределение

В системах обслуживания со смешанными распределениями входные потоки или времена обслуживания очень часто оказываются смесью распределений с легкими и тяжелыми хвостами. Для изучения поведения таких систем введено показательное-Парето распределение, которое является смесью показательного распределения и распределения Парето.

Определение 1.13. *Смесь*

$$Z = I \cdot X + (1 - I) \cdot Y \quad (1.50)$$

двух с.в. $X \sim \text{Exp}(\lambda)$ и $Y \sim \text{Pareto}(\alpha, x_0)$ имеет показательное-Парето распределение с параметрами $\lambda > 0$, $\alpha > 0$, $x_0 > 0$, $0 < p < 1$, если ее ф.р. имеет вид

$$F_Z(x) = 1 - pF_X(x) - (1 - p)F_Y(x) = 1 - pe^{-\lambda x} - (1 - p) \left(\frac{x_0}{x_0 + x} \right)^\alpha, \quad x \geq 0. \quad (1.51)$$

Распределение (1.51) будем обозначать через $\text{Exp} - \text{Pareto}(\lambda, \alpha, x_0, p)$. Плотность показательного-Парето распределения имеет вид

$$f_Z(x) = p\lambda e^{-\lambda x} + (1 - p) \frac{\alpha x_0^\alpha}{(x_0 + x)^{\alpha+1}}$$

и пределы

$$\lim_{x \rightarrow \infty} f_Z(x) = 0; \quad \lim_{x \rightarrow 0} f_Z(x) = p\lambda + \frac{(1-p)\alpha}{x_0}.$$

На Рисунке 1.5 изображены графики ф.р. показательного-Парето распреде-

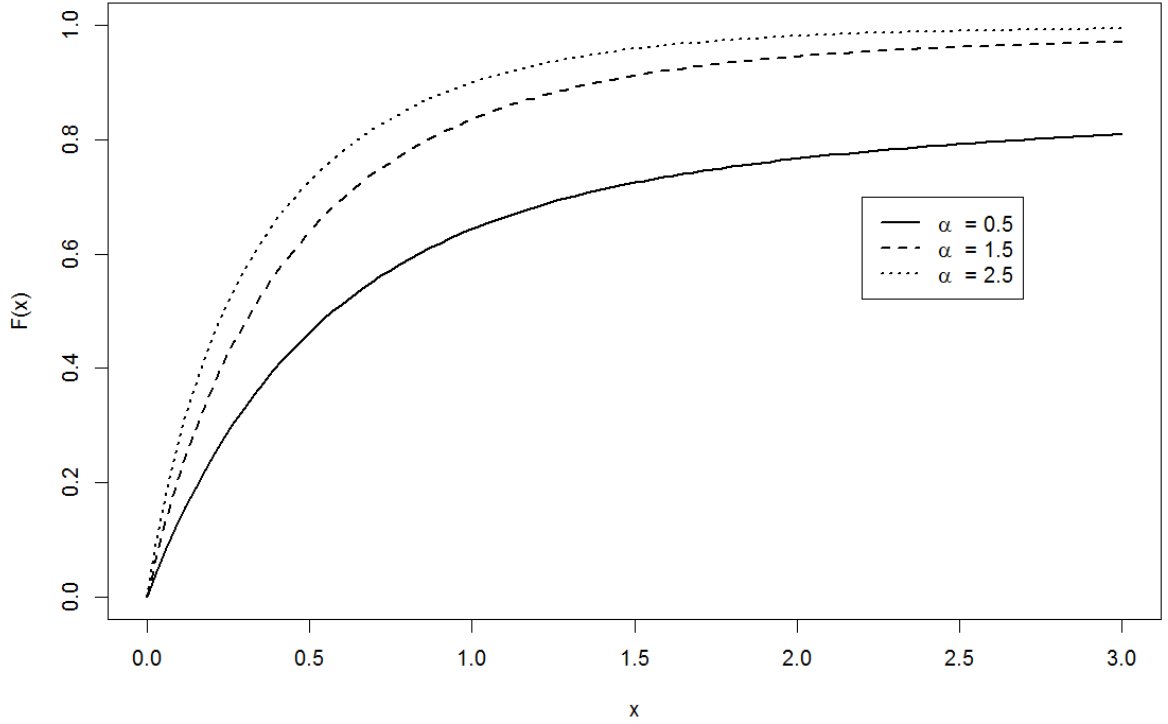


Рисунок 1.5. Ф.р. показательного-Парето распределения.

ления с параметрами $x_0 = 0.5, \lambda = 2, p = 0.5$ при значениях параметра $\alpha = 0.5, 1.5, 2.5$. Для последующего анализа важно выяснить некоторые свойства данного распределения. В частности, имеет место следующий результат.

Лемма 1.3. *С.в. Z с показательным-Парето распределением имеет характеристическую функцию вида*

$$\phi_Z(t) = \frac{\lambda\alpha}{\lambda - ipt} e^{-i(1-p)tx_0} \sum_{k=0}^{\infty} \frac{(i(1-p)tx_0)^k}{k!(\alpha - k)}. \quad (1.52)$$

Доказательство. Вычислим х.ф.

$$\phi_Z(t) = \mathbf{M} e^{it(pX + (1-p)Y)} = \mathbf{M} e^{pitX} \mathbf{M} e^{(1-p)itY} = \phi_X(pt) \phi_Y((1-p)t),$$

где х.ф. показательной компоненты X имеет вид

$$\phi_X(t) = \int_0^{\infty} e^{itx} \lambda e^{-\lambda x} dx = \frac{\lambda}{\lambda - it}.$$

Х.ф. распределения Парето можно вычислить, используя разложение экспоненциальной функции в ряд Тейлора в окрестности нуля

$$\begin{aligned} \phi_Y(t) &= \int_0^{\infty} e^{itx} \alpha x_0^{\alpha} (x_0 + x)^{-\alpha-1} dx \\ &= e^{-itx_0} \int_{x_0}^{\infty} e^{itx} x_0^{\alpha} \alpha x^{-\alpha-1} dx \\ &= e^{-itx_0} x_0^{\alpha} \alpha \int_{x_0}^{\infty} \sum_{k=0}^{\infty} \frac{(itx)^k}{k!} x^{-\alpha-1} dx \\ &= e^{-itx_0} x_0^{\alpha} \alpha \sum_{k=0}^{\infty} \frac{(it)^k}{k!} \int_{x_0}^{\infty} x^{k-\alpha-1} dx \\ &= e^{-itx_0} \alpha \sum_{k=0}^{\infty} \frac{(itx_0)^k}{k!(\alpha - k)}. \end{aligned}$$

Таким образом, х.ф. показательного-Парето распределения имеет вид (1.52). □

Используя свойство х.ф., позволяющее вычислять начальные моменты [17] как

$$i^{-k} \frac{d^k \phi_Z(t)}{dt^k} \Big|_{t=0} = \mathbf{M}Z^k,$$

найдем моменты k -го порядка с.в. Z в предположении, что $\alpha > k$:

$$\mathbf{M}Z^k = \Gamma(1+k) \left(\frac{p}{\lambda}\right)^k + \frac{\Gamma(1+k)x_0^k(1-p)^k}{\Gamma(\alpha)} \sum_{j=0}^{k-1} \left(\frac{p}{x_0\lambda(1-p)}\right)^j \Gamma(\alpha - k + j),$$

где $\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t} dt$ – Гамма-функция. В частности, математическое ожидание равно

$$\mathbf{M}Z = \frac{p}{\lambda} + \frac{(1-p)x_0}{\alpha - 1}, \quad \alpha > 1,$$

второй начальный момент равен

$$MZ^2 = \frac{2p^2}{\lambda^2} + \frac{2p(1-p)x_0}{\lambda(\alpha-1)} + \frac{2(1-p)^2x_0^2}{(\alpha-1)(\alpha-2)}, \quad \alpha > 2,$$

а дисперсия равна

$$DZ = \frac{p^2}{\lambda^2} + \frac{(1-p)^2x_0^2\alpha}{(\alpha-1)^2(\alpha-2)}, \quad \alpha > 2.$$

Подчеркнем, что моменты порядка k существуют только для $k < \alpha$.

Пусть с.в. X_u , Y_u – величины превышения уровня u для с.в. X и Y , соответственно, хвосты ф.р. которых вычисляются по формуле (1.4). Обозначим через Z_u с.в., соответствующую ф.р. величины превышения уровня u с хвостом распределения, т.е.

$$\bar{F}_{Z_u}(x) = \frac{\bar{F}_Z(x+u)}{\bar{F}_Z(u)} = \frac{pe^{-\lambda(x+u)} + (1-p)\left(\frac{x_0}{x_0+x+u}\right)^\alpha}{pe^{-\lambda u} + (1-p)\left(\frac{x_0}{x_0+u}\right)^\alpha}, \quad x \geq 0.$$

Поскольку хвост условного распределения величины превышения уровня u сходится к 1 при $u \rightarrow \infty$

$$\lim_{u \rightarrow \infty} \bar{F}_{Z_u}(x) = 1,$$

то в силу определения 1.11 показательное-Парето распределение относится к классу распределений с длинными хвостами.

Вычислим среднюю величину превышения уровня u смеси Z с ф.р. (1.51) по формуле (1.5):

$$\begin{aligned} e_Z(u) = MZ_u &= \int_0^\infty \frac{\bar{F}_Z(x+u)}{\bar{F}_Z(u)} dx = \frac{\int_u^\infty \bar{F}_Z(x) dx}{\bar{F}_Z(u)} \\ &= \frac{\int_u^\infty (pe^{-\lambda x} + (1-p)x_0^\alpha(x_0+x)^{-\alpha}) dx}{pe^{-\lambda u} + (1-p)x_0^\alpha(x_0+u)^{-\alpha}} \\ &= \frac{pe^{-\lambda u}(\alpha-1)(x_0+u)^\alpha + \lambda(1-p)x_0^\alpha(x_0+u)}{\lambda p(\alpha-1)(x_0+u)^\alpha e^{-\lambda u} + \lambda(1-p)(\alpha-1)x_0^\alpha}, \quad \alpha > 1. \end{aligned}$$

Отметим, что средняя величина $e_Z(u)$ превышения уровня u является возрастающей функцией и неограниченно растет с увеличением порогового значения

u , а именно $e_Z(u) \rightarrow \infty$ при $u \rightarrow \infty$. Функция интенсивности отказов с.в. Z вычисляется по формуле

$$r_Z(x) = \frac{p r_X(x) a(x) + (1-p) r_Y(x)}{p a(x) + (1-p)}, \quad (1.53)$$

где

$$a(x) = e^{-\lambda x} \left(1 + \frac{x}{x_0}\right)^\alpha. \quad (1.54)$$

Заметим, что

$$r_Z(x) \longrightarrow 0 \quad \text{при} \quad x \rightarrow \infty,$$

что характерно для распределений с длинными хвостами, и что

$$r_Z(x) \longrightarrow p\lambda + (1-p)\frac{\alpha}{x_0} \quad \text{при} \quad x \rightarrow 0.$$

При выполнении соотношения $\lambda \geq \alpha/x_0$ между параметрами смеси функция интенсивности отказов ограничена сверху, $r_Z(x) \leq \lambda$.

Установим монотонность функции интенсивности отказов. Действительно, производная

$$r'_Z(x) = -(1-p) \frac{p a(x) (r_X(x) - r_Y(x))^2 + r_Y^2(x) / \alpha (p a(x) + (1-p))}{(p a(x) + (1-p))^2} < 0.$$

Поэтому $r_Z(x)$ монотонно убывает и, следовательно, показательное-Парето распределение относится к классу УФИ-распределений. Функция интенсивности отказов величины Z_u превышения уровня u , вычисляется по формуле:

$$r_{Z_u}(x) = \frac{p r_{X_u}(x) a(x+u) + (1-p) r_{Y_u}(x+u)}{p a(x+u) + (1-p)} = r_Z(x+u), \quad (1.55)$$

где функция $a(x)$ задана формулой (1.54). Вычислим функцию равновесного распределения F_e с.в. Z ($x \geq 0$):

$$F_e(x) = \frac{1}{MZ} \int_0^x \bar{F}_Z(t) dt = 1 - \frac{1}{MZ} \left(\frac{p e^{-\lambda x}}{\lambda} + \frac{(1-p)x_0^\alpha}{(\alpha-1)(x_0+x)^{\alpha-1}} \right). \quad (1.56)$$

Эта функция далее в Главе 5 будет использована для получения предельного распределения максимума систем обслуживания с показательным-Парето вре-

менем обслуживания.

Теорема 1.5 показывает, при каком условии на параметры показательного-Парето распределения, смесь (1.50) ограничена по интенсивности отказов: сверху распределением Парето, а снизу – показательным распределением.

Теорема 1.5. Пусть с.в. Z имеет показательное-Парето распределение (1.51). Тогда, если выполнено неравенство

$$\frac{\alpha}{x_0} \leq \lambda, \quad (1.57)$$

то

$$X \leq \frac{Z}{r} \leq Y, \quad (1.58)$$

$$X_u \leq \frac{Z_u}{r} \leq Y_u, \quad (1.59)$$

$$e_X(u) \leq e_Z(u) \leq e_Y(u). \quad (1.60)$$

Доказательство. Найдем условия, при которых

$$r_Y(x) \leq r_Z(x) \leq r_X(x), \quad x \geq 0. \quad (1.61)$$

Согласно Лемме 1.1, для выполнения $r_Y(x) \leq r_X(x)$ достаточно выполнения соотношения

$$r_Y(x) \leq \sup_{x \geq 0} r_Y(x) = r_Y(0) = \frac{\alpha}{x_0} \leq r_X(x) = \lambda. \quad (1.62)$$

Поскольку неравенство (1.62) соответствует условию (1.57), то по Теореме 1.1 верно соотношение (1.58).

В работе [132] показано, что исходные с.в. упорядочены по интенсивности отказов тогда и только тогда, когда величины превышения уровня u имеют стохастический порядок. Покажем, что величины X_u и Y_u упорядочены по интенсивности отказов для любых $x, u \geq 0$. Действительно,

$$r_{Y_u}(x) = r_Y(x + u) \leq \sup_{x \geq 0} r_Y(x + u) = r_Y(u) = \frac{\alpha}{x_0 + u} \leq \frac{\alpha}{x_0} \leq r_X(x + u) = \lambda.$$

Отсюда следует, что $r_Y(x + u) \leq r_X(x + u)$ для всех $x, u \geq 0$, что, согласно Теореме 1.1, доказывает утверждение (1.59). Вычислим средние $e_X(u)$ и $e_Y(u)$.

Для показательной с.в. X $e_X(u) = 1/\lambda$. Для распределения Парето

$$e_Y(u) = \frac{x_0 + u}{\alpha - 1}, \quad u \geq 0, \quad \alpha > 1.$$

Если условие (1.57) выполнено, то верны неравенства

$$\lambda \geq \frac{\alpha}{x_0} \geq \frac{\alpha - 1}{x_0 + u},$$

и тогда, как несложно проверить, выполнены также неравенства

$$\frac{1}{\lambda} \leq \frac{pe^{-\lambda u}(\alpha - 1)(x_0 + u)^\alpha + \lambda(1 - p)x_0^\alpha(x_0 + u)}{\lambda p(\alpha - 1)(x_0 + u)^\alpha e^{-\lambda u} + \lambda(1 - p)(\alpha - 1)x_0^\alpha} \leq \frac{x_0 + u}{\alpha - 1},$$

что и доказывает (1.60). □

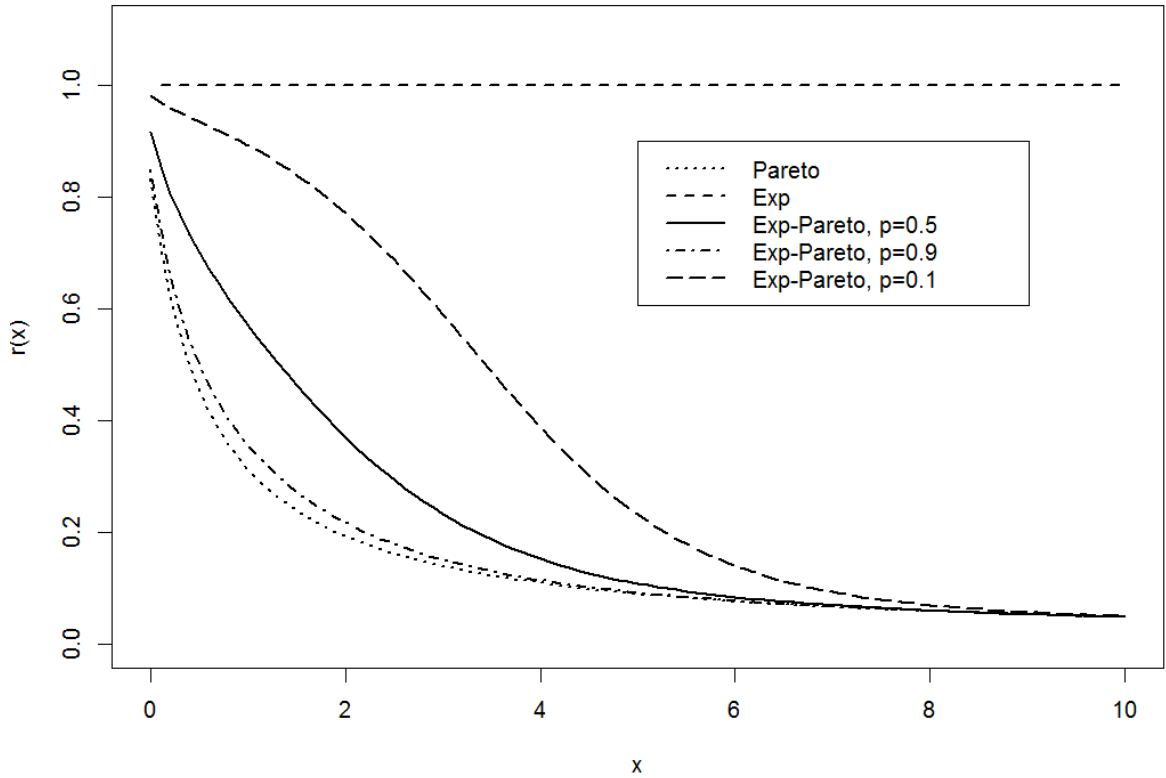


Рисунок 1.6. Функции интенсивностей отказов распределений $Exp(1)$, $Pareto(0.5; 0.6)$ и смеси $Exp - Pareto(1; 0.5; 0.6; p)$ при различных значениях коэффициента смеси p .

Рисунок 1.6 демонстрирует функции интенсивностей отказов распределений с.в. X, Y, Z при $\lambda = 1, \alpha = 0.5, x_0 = 0.6$. Видно, что, при заданных параметрах, функции интенсивности отказов упорядочены в соответствии с со-

отношением (1.61). Следовательно, с.в. упорядочены по интенсивности отказов согласно неравенству (1.60).

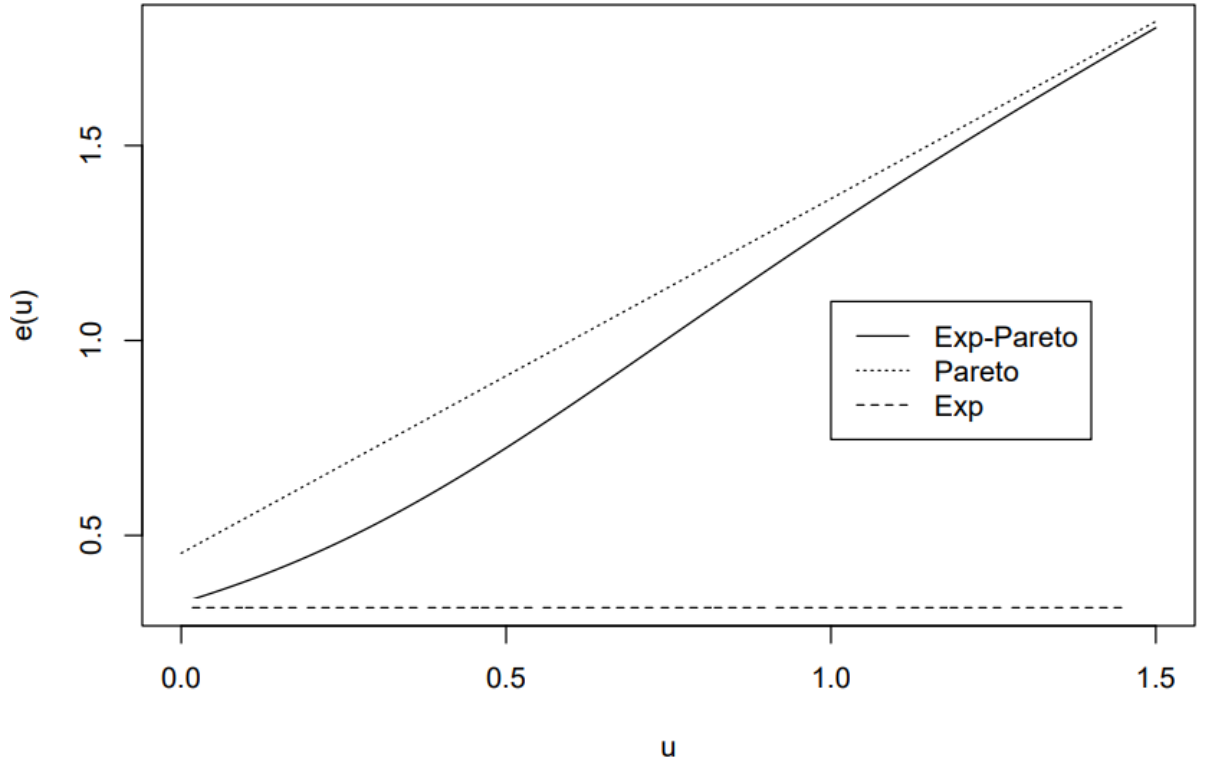


Рисунок 1.7. Средние величины $e(u)$ превышения уровня u для $Exp(5)$, $Pareto(2.1; 0.5)$ и смеси $Exp - Pareto(5; 2.1; 0.5; 0.5)$.

На Рисунке 1.7 видно, что при заданных параметрах ($\lambda = 5, \alpha = 2.1, x_0 = 0.5, p = 0.5$) средние упорядочены в соответствии с соотношением (1.60).

1.9. Сравнение некоторых двухкомпонентных смесей

Предположим, что с.в. X имеет гиперэкспоненциальное распределение вида

$$\bar{F}_X(x) = pe^{-\lambda_1 x} + (1-p)e^{-\lambda_2 x}, \quad x \geq 0, \quad \lambda_1, \lambda_2 > 0. \quad (1.63)$$

С.в. Y имеет показательное-Парето распределение вида

$$\bar{F}_Y(x) = pe^{-\lambda x} + (1-p) \left(\frac{x_0}{x_0 + x} \right)^\alpha, \quad x \geq 0, \quad (1.64)$$

и пусть (хвост) распределения с.в. Z задан смесью распределений Парето

$$\bar{F}_Z(x) = p \left(\frac{x_0}{x_0 + x} \right)^{\alpha_1} + (1-p) \left(\frac{x_0}{x_0 + x} \right)^{\alpha_2}, \quad x \geq 0. \quad (1.65)$$

В Теореме 1.6 получены соотношения между параметрами $\lambda, \lambda_1, \lambda_2, \alpha, \alpha_1, \alpha_2$ распределений (1.63)–(1.65), гарантирующие стохастический порядок смесей $X = IX_1 + (1 - I)X_2, Y = IY_1 + (1 - I)Y_2, Z = IZ_1 + (1 - I)Z_2$, где $P(I = 1) = p, P(I = 0) = 1 - p$.

Теорема 1.6. Пусть ф.р. с.в. X, Y, Z заданы соотношениями (1.63) – (1.65), соответственно, и выполняются следующие соотношения между их параметрами:

$$\lambda_1 \geq \lambda \geq \frac{\alpha_1}{x_0}, \quad (1.66)$$

$$\lambda_2 \geq \frac{\alpha}{x_0} \geq \frac{\alpha_2}{x_0}. \quad (1.67)$$

Тогда смеси имеют стохастический порядок

$$X \underset{st}{\leq} Y \underset{st}{\leq} Z. \quad (1.68)$$

Доказательство. Из соотношений

$$\lambda_1 \geq \lambda, \quad \lambda_2 \geq \frac{\alpha}{x_0} \quad (1.69)$$

следует, что компоненты X и Y упорядочены по интенсивности отказов:

$$X_1 \underset{r}{\leq} Y_1, \quad X_2 \underset{r}{\leq} Y_2. \quad (1.70)$$

С другой стороны, из (1.70) вытекает стохастическая упорядоченность:

$$X_1 \underset{st}{\leq} Y_1, \quad X_2 \underset{st}{\leq} Y_2. \quad (1.71)$$

Следовательно, на основании Теоремы 1.2, смеси X и Y имеют стохастический порядок, $X \underset{st}{\leq} Y$. Если параметры распределений смесей Y и Z связаны соотношениями

$$\lambda \geq \frac{\alpha_1}{x_0}, \quad \alpha \geq \alpha_2, \quad (1.72)$$

то функции интенсивностей отказов компонент удовлетворяют неравенствам

$$r_{Y_i}(x) \geq r_{Z_i}(x), \quad x \geq 0.$$

Следовательно, они упорядочены по интенсивности отказов, и, значит стоха-

стически упорядочены:

$$Y_i \leq_r Z_i, \quad Y_i \leq_{st} Z_i, \quad i = 1, 2. \quad (1.73)$$

Тогда, на основании Теоремы 1.2 (раздел 1.7.), смеси Y и Z также имеют стохастический порядок:

$$Y \leq_{st} Z.$$

Таким образом, при условиях (1.66), (1.67), смеси X , Y , Z стохастически упорядочены в соответствии с (1.68). $\square$

В Теореме 1.7 получены соотношения между параметрами распределений (1.63)-(1.65), гарантирующие их упорядоченность по интенсивности отказов.

Теорема 1.7. *Если с.в. X , Y , Z заданы распределениями (1.63) – (1.65), соответственно, и выполнены следующие соотношения:*

$$\frac{\alpha_1}{x_0} \leq \lambda \leq \frac{\alpha_2}{x_0} = \frac{\alpha}{x_0} \leq \lambda_2 \leq \lambda_1, \quad (1.74)$$

то смеси упорядочены по интенсивности отказов, а именно,

$$X \leq_r Y \leq_r Z. \quad (1.75)$$

Доказательство. Во-первых, заметим, что (1.74) влечет условия (1.66) и (1.67) на параметры. Это в свою очередь влечет стохастическую упорядоченность компонент

$$X_i \leq_{st} Y_i, \quad Y_i \leq_{st} Z_i, \quad i = 1, 2.$$

Поэтому из Теоремы 1.6 следует стохастическая упорядоченность смесей (1.68).

При $\lambda_1 \geq \lambda$, очевидно, что $X_1 \leq_r Y_1$, а при $\lambda_2 \geq \frac{\alpha}{x_0}$ выполнено соотношение $X_2 \leq_r Y_2$. Следовательно, выполнено условие (1.48) Теоремы 1.3 для смесей X и Y . Чтобы проверить условие

$$\min(X_1, Y_2) \leq_{st} \min(X_2, Y_1),$$

заметим, что оно может быть переписано в виде

$$\left(\frac{x_0}{x_0 + x} \right)^\alpha \leq e^{-(\lambda + \lambda_2 - \lambda_1)x}, \quad x \geq 0. \quad (1.76)$$

Чтобы его проверить отметим, что условие $\lambda \leq \frac{\alpha}{x_0}$ гарантирует $Y_1 \underset{r}{\geq} Y_2$, а, следовательно, $Y_1 \underset{st}{\geq} Y_2$. Это можно записать в виде соотношения

$$\bar{F}_{Y_2}(x) = \left(\frac{x_0}{x_0 + x} \right)^\alpha \leq \bar{F}_{Y_1}(x) = e^{-\lambda x}, \quad x \geq 0. \quad (1.77)$$

Тогда при $\lambda_1 \geq \lambda_2$ верно соотношение

$$e^{-\lambda x} \leq e^{-(\lambda + \lambda_2 - \lambda_1)x},$$

а ввиду (1.77) также верно неравенство (1.76). Условие $\lambda \leq \frac{\alpha}{x_0}$ гарантирует, что $Y_1 \underset{r}{\geq} Y_2$, и выполнено соотношение (1.46). Таким образом, при выполнении неравенств

$$\lambda_1 \geq \lambda_2 \geq \frac{\alpha}{x_0} \geq \lambda, \quad (1.78)$$

на основании Теоремы 1.3 (раздел 1.7.) смеси X и Y упорядочены по интенсивности отказов, т. е. $X \underset{r}{\leq} Y$.

Теперь докажем, что смеси Y и Z упорядочены по интенсивности отказов. Пусть $\alpha_1 \leq \alpha_2$, тогда $Z_1 \underset{r}{\geq} Z_2$. При $\lambda \geq \frac{\alpha_1}{x_0}$ с.в. Y_1 и Z_1 упорядочены по интенсивности отказов, $Y_1 \underset{r}{\leq} Z_1$, а при $\alpha = \alpha_2$ функции интенсивности отказов с.в. Y_2 и Z_2 совпадают, $r_{Y_2}(x) = r_{Z_2}(x)$. Таким образом, условия (1.46) и (1.48) Теоремы 1.3 для смесей Y и Z выполнены.

Проверим выполнение условия

$$\min(Y_1, Z_2) \underset{st}{\leq} \min(Y_2, Z_1), \quad (1.79)$$

которое в данном случае эквивалентно условию

$$e^{-\lambda x} \leq \left(\frac{x_0}{x_0 + x} \right)^{\alpha_1 + \alpha - \alpha_2}, \quad x \geq 0. \quad (1.80)$$

При условии $\lambda \geq \frac{\alpha_1}{x_0}$ с.в. Y_1 и Z_1 упорядочены по интенсивности отказов, $Y_1 \underset{r}{\leq} Z_1$, а, следовательно, стохастически упорядочены, $Y_1 \underset{st}{\leq} Z_1$. Поэтому для хвостов ф.р. верно неравенство

$$\bar{F}_{Y_1}(x) = e^{-\lambda x} \leq \bar{F}_{Z_1}(x) = \left(\frac{x_0}{x_0 + x} \right)^{\alpha_1}, \quad x \geq 0,$$

и неравенство (1.79) выполняется при $\alpha = \alpha_2$. Таким образом, при выполнении соотношений между параметрами

$$\alpha = \alpha_2 \geq \alpha_1, \quad \lambda x_0 \geq \alpha_1, \quad (1.81)$$

все условия Теоремы 1.3 для смесей Y и Z выполнены и, значит,

$$Y \underset{r}{\leq} Z.$$

Соотношения (1.78) и (1.81) вместе дают условие (1.74). □

1.10. Равномерная метрика для двухкомпонентных смесей распределения

В этом разделе рассмотрена равномерная метрика между двумя распределениями, которая широко используется в моделировании, в том числе и в данной работе, при анализе чувствительности систем.

Определение 1.14. ([104]) *Величина*

$$\Delta(F_X, F_Y) = \sup_x |F_X(x) - F_Y(x)| \quad (1.82)$$

называется расстоянием Колмогорова-Смирнова (равномерным расстоянием) между двумя распределениями F_X и F_Y .

Пусть с.в. X – смесь с.в. X_1 и X_2 . Равномерное расстояние $\Delta(F_X, F_{X_1})$ между распределением смеси $F_X(x) = pF_{X_1}(x) + (1-p)F_{X_2}(x)$ и распределением F_{X_1} первой компоненты смеси равно

$$\begin{aligned} \Delta(F_X, F_{X_1}) &= \sup_{x \geq 0} |pF_{X_1}(x) + (1-p)F_{X_2}(x) - F_{X_1}(x)| \\ &= (1-p) \sup_{x \geq 0} |F_{X_1}(x) - F_{X_2}(x)|. \end{aligned} \quad (1.83)$$

Если существуют плотности f_{X_i} и точка x^* , в которой достигается супремум выражения (1.83),

$$\Delta(F_X, F_{X_1}) = (1-p)|F_{X_1}(x^*) - F_{X_2}(x^*)|,$$

то x^* удовлетворяет равенству

$$f_{X_1}(x^*) = f_{X_2}(x^*). \quad (1.84)$$

Выразим плотности через функции интенсивности отказов r_{X_i} , тогда равенство (1.84) преобразуется в равенство

$$r_{X_1}(x^*)\bar{F}_{X_1}(x^*) = r_{X_2}(x^*)\bar{F}_{X_2}(x^*).$$

Таким образом, выражение (1.83) может быть записано в более удобной форме как

$$\begin{aligned} \Delta(F_X, F_{X_1}) &= (1-p) \frac{|r_2(x^*) - r_1(x^*)|}{r_2(x^*)} \bar{F}_{X_1}(x^*) \\ &= (1-p) \frac{|r_{X_1}(x^*) - r_{X_2}(x^*)|}{r_{X_1}(x^*)} \bar{F}_{X_2}(x^*). \end{aligned} \quad (1.85)$$

Заметим, что (1.85) позволяет получить верхнюю оценку для равномерного расстояния $\Delta(F_X, F_{X_1})$:

$$\Delta(F_X, F_{X_1}) \leq (1-p) \frac{|r_{X_2}(x^*) - r_{X_1}(x^*)|}{r_{X_1}(x^*)} =: \delta(x^*). \quad (1.86)$$

Пример 1.15. Для гиперэкспоненциального распределения вида (1.63) смеси двух с.в. $X_i \sim \text{Exp}(\lambda_i), i = 1, 2$, точка супремума равна

$$x^* = \frac{\log \lambda_1 - \log \lambda_2}{\lambda_1 - \lambda_2},$$

и в этом случае оценка (1.85) принимает вид

$$\Delta(F_X, F_{X_1}) = (1-p) \frac{|\lambda_2 - \lambda_1|}{\lambda_2} \left(\frac{\lambda_1}{\lambda_2} \right)^{-\frac{\lambda_1}{\lambda_1 - \lambda_2}} \leq (1-p) \frac{|\lambda_2 - \lambda_1|}{\lambda_2}. \quad (1.87)$$

Выражение (1.87) является частным случаем более общего результата, полученного в работе [154] для одномерных *смесей* X_M вида

$$X_M \stackrel{st}{=} \frac{X_1}{Y}, \quad (1.88)$$

где $X \sim Exp(\lambda_1)$, а с.в. Y имеет распределение:

$$Y = \begin{cases} 1, & \text{с вероятностью } p; \\ \lambda_1/\lambda_2 & \text{с вероятностью } 1 - p. \end{cases} \quad (1.89)$$

В общем случае смеси вида (1.88) называются масштабными смесями [15], и они имеют ф.р.

$$\hat{F}_{X_M}(x) = \int_0^\infty F_{X_1}(\theta x) dF_Y(\theta),$$

где F_{X_1} – ф.р. смешиваемой с.в. X_1 , а F_Y – ф.р. смешивающей с.в. $Y > 0$. По сути дела, (1.88) является преобразованием масштаба, и если $Y \in \{y_1, \dots, y_m\}$ – дискретная с.в., то (1.88) преобразуется в конечную смесь

$$X_M = \sum_{i=1}^m \frac{1}{y_i} I(Y = y_i) X_1. \quad (1.90)$$

Для случая $X_1 \sim Exp(\lambda_1)$ известна оценка равномерной метрики [154, 155]

$$\Delta(\hat{F}_{X_M}, F_{X_1}) \leq M|Y - 1|. \quad (1.91)$$

Вычислив математическое ожидание с.в. Y из (1.89) и подставив его в неравенство (1.91), получим правую часть выражения (1.87).

Пример 1.16. Рассмотрим двухкомпонентную смесь

$$Y = IY_1 + (1 - I)Y_2,$$

где $Y_i \sim Pareto(\alpha_i, x_0)$, $i = 1, 2$. Из (1.85) следует, что в этом случае

$$\begin{aligned} \Delta(F_Y, F_{Y_1}) &= (1 - p) \frac{|\alpha_1 - \alpha_2|}{\alpha_2} \left(\frac{\alpha_2}{\alpha_1} \right)^{\frac{\alpha_1}{\alpha_1 - \alpha_2}} \\ &\leq (1 - p) \frac{|\alpha_1 - \alpha_2|}{\alpha_2} =: \delta(\alpha_1, \alpha_2), \end{aligned} \quad (1.92)$$

а точка x^* , удовлетворяющая равенству (1.84), равна

$$x^* = x_0 \left(\left(\frac{\alpha_1}{\alpha_2} \right)^{\frac{1}{\alpha_1 - \alpha_2}} - 1 \right).$$

Заметим, что оценки равномерных расстояний для смеси двух распределений Парето (1.92) и для гиперэкспоненциального распределения (1.87) совпадают при равенстве параметров, $\alpha_i = \lambda_i$, $i = 1, 2$.

Рисунок 1.8 демонстрирует зависимость расстояния Колмогорова-Смирнова $\Delta(F_Y, F_{Y_1})$ и $\delta(\alpha_1, \alpha_2)$ от параметра α_2 при фиксированных параметрах $\alpha_1 = 2$ и $p = 0.9$, в то время как параметр α_2 варьируется в интервале $(1, 5)$.

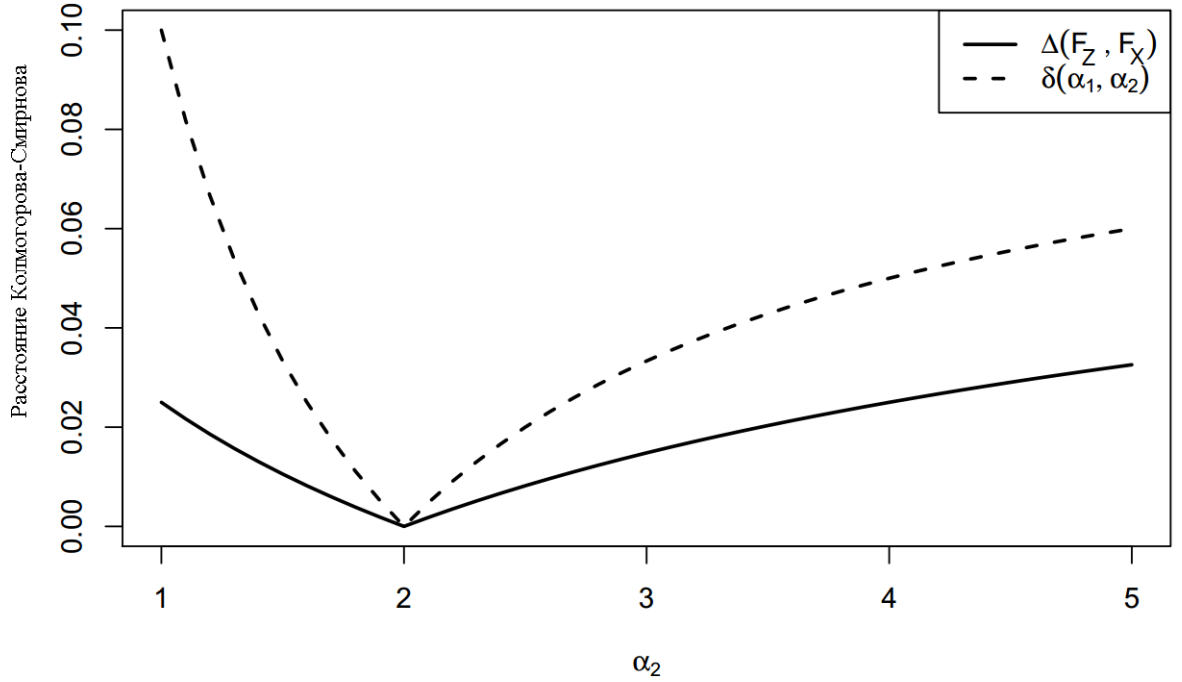


Рисунок 1.8. Зависимость равномерной метрики $\Delta(F_Y, F_{Y_1})$ с параметром смеси $p = 0.9$ и верхней границей $\delta(\alpha_1, \alpha_2)$ от параметра α_2 .

Пусть с.в. Z задана смесью

$$Z = IX + (1 - I)Y$$

и пусть с.в. X_u , Y_u – соответствующие величины превышения (фиксированного) уровня u . Обозначим через Z_u величину превышения уровня u , заданную хвостом ф.р.

$$\bar{F}_{Z_u}(x) = \frac{\bar{F}_Z(x + u)}{\bar{F}_Z(u)}, \quad x \geq 0.$$

Равномерная метрика между $\bar{F}_{Z_u}$ и $\bar{F}_{X_u}$ для фиксированного $u \geq 0$ вычисляется

по формуле

$$\begin{aligned}\Delta(\bar{F}_{Z_u}, \bar{F}_{X_u}) &:= \sup_x |\bar{F}_{Z_u}(x) - \bar{F}_{X_u}(x)| \\ &= (1-p) \sup_x \left| \frac{\bar{F}_Y(x+u)\bar{F}_X(u) - \bar{F}_Y(u)\bar{F}_X(x+u)}{\bar{F}_X(u)(p\bar{F}_X(u) + (1-p)\bar{F}_Y(u))} \right|. \quad (1.93)\end{aligned}$$

Следующая теорема является весьма полезной при анализе чувствительности функции превышения уровня u смеси распределений.

Теорема 1.8. *Если существуют плотности f_X и f_Y и точка x^* , в которой достигается супремум (1.93) для фиксированного $u \geq 0$, то x^* удовлетворяет равенству*

$$\frac{r_Y(x^* + u)}{r_X(x^* + u)} = \frac{\bar{F}_{X_u}(x^*)}{\bar{F}_{Y_u}(x^*)}. \quad (1.94)$$

Доказательство. Взяв производную выражения $|\bar{F}_{Z_u}(x) - \bar{F}_{X_u}(x)|$ по x , получим

$$\frac{f_Y(x+u)}{f_X(x+u)} = \frac{\bar{F}_Y(u)}{\bar{F}_X(u)}.$$

Далее, умножив левую и правую части последнего соотношения на $\bar{F}_X(x+u)/\bar{F}_Y(x+u)$, получим (1.94). $\square$

Например, для гиперэкспоненциального распределения (1.63) решение (1.94) имеет следующий вид:

$$x^* = \frac{\log \lambda_2 - \log \lambda_1}{\lambda_2 - \lambda_1},$$

и совпадает с выражением x^* для равномерной метрики $\Delta(F_X, F_{X_1})$ между гиперэкспоненциальным распределением F_X с параметрами $\lambda_1 > \lambda_2$ и показательным распределением F_{X_1} с параметром λ_1 . Выражение (1.93) в этом случае принимает вид

$$\begin{aligned}\Delta(\bar{F}_{Z_u}, \bar{F}_{X_u}) &= \frac{1-p}{pe^{-(\lambda_1-\lambda_2)u} + (1-p)} \frac{|\lambda_2 - \lambda_1|}{\lambda_2} \left(\frac{\lambda_1}{\lambda_2} \right)^{-\frac{\lambda_1}{\lambda_1-\lambda_2}} \\ &= \frac{1}{pe^{-(\lambda_1-\lambda_2)u} + (1-p)} \Delta(\bar{F}_Z, \bar{F}_X), \quad (1.95)\end{aligned}$$

и более того,

$$\Delta(\bar{F}_{Z_u}, \bar{F}_{X_u}) \rightarrow \frac{\Delta(\bar{F}_Z, \bar{F}_X)}{1-p}, \quad u \rightarrow \infty. \quad (1.96)$$

Для смеси распределений Парето (1.65) из равенства (1.94) для фиксированного u находим

$$x^* = (x_0 + u) \left(\frac{\alpha_2}{\alpha_1} \right)^{1/(\alpha_2 - \alpha_1)} - x_0 - u. \quad (1.97)$$

Далее из (1.96) находим, что

$$\begin{aligned} \Delta(\bar{F}_{Z_u}, \bar{F}_{X_u}) &= \sup_x \left| \frac{p \left(\frac{x_0}{x_0+x+u} \right)^{\alpha_1} + (1-p) \left(\frac{x_0}{x_0+x+u} \right)^{\alpha_2}}{p \left(\frac{x_0}{x_0+u} \right)^{\alpha_1} + (1-p) \left(\frac{x_0}{x_0+u} \right)^{\alpha_2}} - \left(\frac{x_0+u}{x_0+x+u} \right)^{\alpha_1} \right| \\ &= \sup_x \left| \left(\frac{p + (1-p) \left(1 + \frac{x+u}{x_0} \right)^{\alpha_1 - \alpha_2}}{p + (1-p) \left(1 + \frac{u}{x_0} \right)^{\alpha_1 - \alpha_2}} - 1 \right) \left(\frac{x_0+x+u}{x_0+u} \right)^{-\alpha_1} \right| \\ &= \frac{(x_0+u)^{\alpha_1 - \alpha_2}}{px_0^{\alpha_1 - \alpha_2} + (1-p)(x_0+u)^{\alpha_1 - \alpha_2}} \Delta(\bar{F}_Z, \bar{F}_X) \\ &\rightarrow \frac{1}{1-p} \Delta(\bar{F}_Z, \bar{F}_X), \quad u \rightarrow \infty, \end{aligned} \quad (1.98)$$

где, согласно (1.92), равномерная метрика между хвостами распределений с.в. Z и X равна

$$\Delta(\bar{F}_Z, \bar{F}_X) = (1-p) \frac{\alpha_1 - \alpha_2}{\alpha_2} \left(\frac{\alpha_2}{\alpha_1} \right)^{\frac{\alpha_1}{\alpha_1 - \alpha_2}}.$$

Следовательно, в пределе равномерная метрика принимает значение

$$\lim_{u \rightarrow \infty} \Delta(\bar{F}_{Z_u}, \bar{F}_{X_u}) = \frac{\alpha_1 - \alpha_2}{\alpha_2} \left(\frac{\alpha_2}{\alpha_1} \right)^{\frac{\alpha_1}{\alpha_1 - \alpha_2}}.$$

Пример 1.17. Пусть с.в. Z имеет показательное-Парето распределение вида (1.50). Из неравенства Пинскера-Чизара [185] вытекает, что равномерная метрика ограничена сверху следующим образом:

$$\Delta(F_X, F_Y) \leq \sqrt{2D_{KL}(X||Y)}, \quad (1.99)$$

где

$$D_{KL}(X||Y) := \int_0^{\infty} f_X(x) \log \left(\frac{f_X(x)}{f_Y(x)} \right) dx \quad (1.100)$$

— дивергенция Кульбака-Лейблера. В работе [185] показано, что минимальная дивергенция между распределениями $Exp(\lambda)$ и $Pareto(\alpha, x_0)$ достигается при условии $\lambda = \frac{\alpha-1}{x_0}$ и, более того,

$$\min D_{KL}(Y||X) \leq \frac{1}{\alpha(\alpha-1)}. \quad (1.101)$$

Очевидно, что если дивергенция Кульбака-Лейблера близка к нулю, то равно-

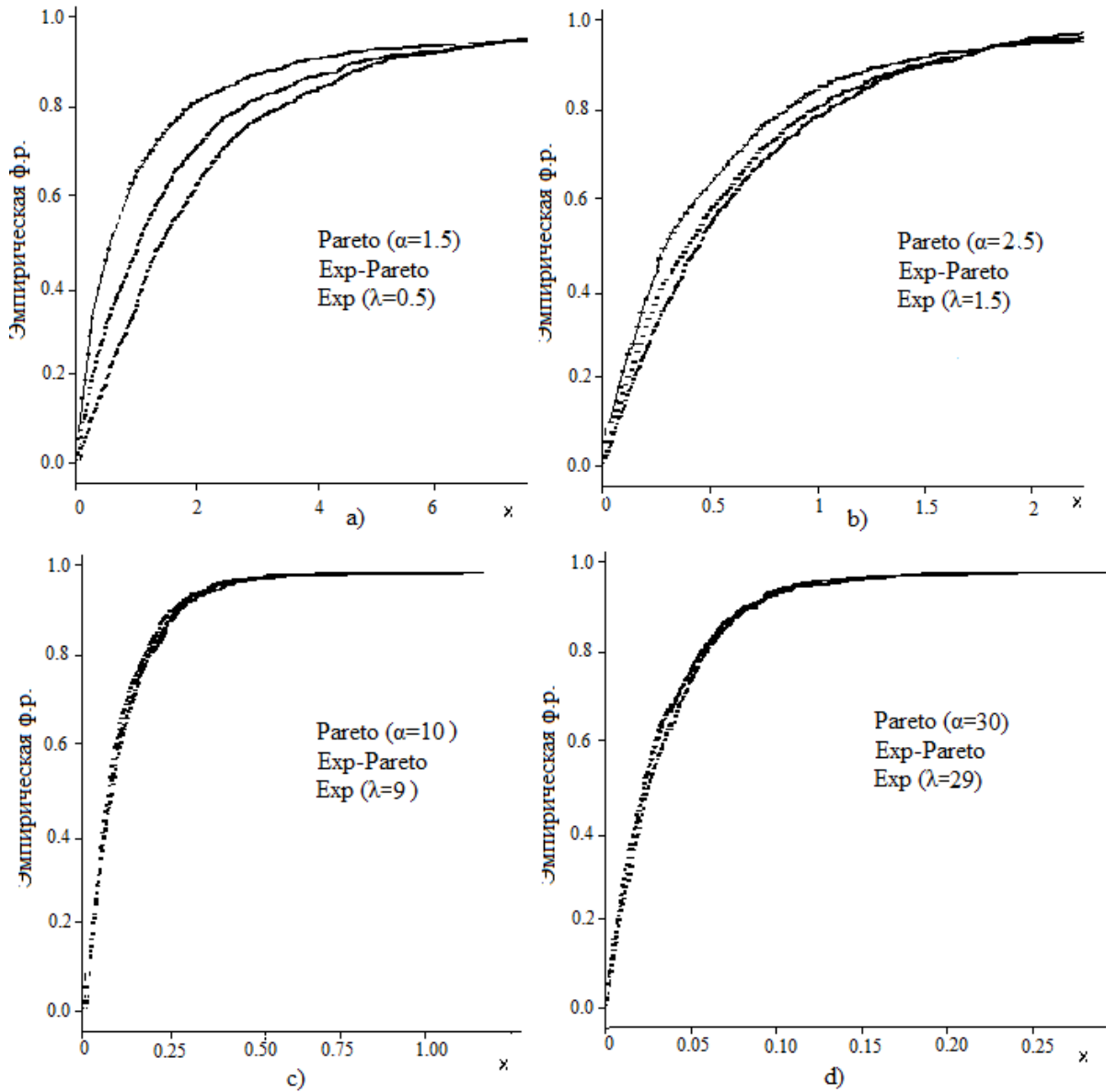


Рисунок 1.9. Эмпирические ф.р. показательного, показательного-Парето и Парето распределений при различных α, λ , и $x_0 = 1, p = 0.5$

мерная метрика ведет себя аналогично, что означает близость распределений показательного и Парето.

Применим (1.99) для оценки равномерной метрики между распределениями показательным-Парето и Парето (для случая $\alpha > 1$). Получаем

$$\begin{aligned}\Delta(F_Z, F_Y) &= \sup_x |pF_X(x) + (1-p)F_Y(x) - F_Y(x)| = \sup_x |pF_X(x) - pF_Y(x)| \\ &= p\Delta(F_X, F_Y) \leq p\sqrt{2D_{KL}(Y||X)} \leq \frac{\sqrt{2p}}{\sqrt{\alpha(\alpha-1)}}.\end{aligned}\quad (1.102)$$

Последнее неравенство демонстрирует скорость сходимости $\Delta(F_Z, F_Y)$ к нулю при $\alpha \rightarrow \infty$. Тот же эффект получается, если аппроксимировать распределение Парето показательным распределением при $\alpha \rightarrow \infty$ [185]. Именно, пусть

$$\lambda = \frac{\alpha + o_\alpha(1)}{x_0}, \quad (1.103)$$

где $o_\alpha(1) \rightarrow 0$ при $\alpha \rightarrow \infty$, тогда

$$\bar{F}_Y(x) = \lim_{\alpha \rightarrow \infty} \left(1 + \frac{\lambda x}{\alpha(1 + o_\alpha(1))}\right)^{-\alpha} = e^{-\lambda x} = \bar{F}_X(x), \quad x \geq 0. \quad (1.104)$$

Рассуждения, приведенные выше, позволяют сформулировать следующую Лемму 1.4 об аппроксимации показательного-Парето распределения показательным распределением при больших значениях параметра α .

Лемма 1.4. Пусть ф.р. $F_Z(x)$ имеет вид (1.51), тогда

$$F_Z(x) \rightarrow 1 - e^{-\lambda x} = F_X(x) \quad \text{при } \alpha \rightarrow \infty, \quad \text{для всех } x \geq 0.$$

На Рисунке 1.9 представлены графики эмпирических ф.р. показательного, показательного-Парето и Парето распределений при $x_0 = 1, p = 0.5$, а) $\alpha = 1.5, \lambda = 2.5$, б) $\alpha = 2.5, \lambda = 3.5$, в) $\alpha = 10, \lambda = 11$, д) $\alpha = 30, \lambda = 31$ для объема выборки $n = 1000$. Как и ожидалось, тест Колмогорова-Смирнова подтверждает, что наборы данных из показательного-Парето и Парето распределений близки только для случая д) $\alpha = 30$. Равномерная метрика $\Delta(F_Z, F_Y)$ равна, соответственно, 0.79, 0.34, 0.074, 0.012, для случаев а)-д). Верхняя граница для равномерной метрики $\frac{\sqrt{2p}}{\sqrt{\alpha(\alpha-1)}}$ (см. соотношение (1.102)) равна 0.81, 0.36, 0.074, 0.023, соответственно.

1.11. Стохастический порядок и упорядоченность по интенсивности отказов многокомпонентных смесей

Проведенный анализ двухкомпонентных смесей позволяет в этом разделе достаточно легко обобщить результаты на случай, когда число компонент m произвольно и конечно.

Рассмотрим семейство неотрицательных независимых с.в. $\{X_i\}$ с ф.р. $F_{X_i}(x)$, $i = 1, \dots, m$. С.в. X является m -компонентной смесью, если ее ф.р. F_X имеет следующий вид [117]:

$$F_X(x) = p_1 F_{X_1}(x) + \dots + p_m F_{X_m}(x), \quad (1.105)$$

где $\sum_{i=1}^m p_i = 1$, $p_i \geq 0$. Заметим, что m -компонентную смесь можно представить в виде

$$X = \sum_{i=1}^m I_i X_i, \quad (1.106)$$

где X_i заданы ф.р. F_{X_i} , и с.в. I_i принимают значение 1 с вероятностью p_i и значение 0, с вероятностью $1 - \sum_{j \neq i}^m p_j$, $i = 1, \dots, m$, $\sum_{i=1}^m I_i = 1$.

Функция интенсивности отказов m -компонентной смеси определяется по формуле

$$r_X(x) = \frac{p_1 f_{X_1}(x) + \dots + p_m f_{X_m}(x)}{p_1 \bar{F}_{X_1}(x) + \dots + p_m \bar{F}_{X_m}(x)},$$

для x , таких, что $\bar{F}_X(x) := P(X > x) > 0$, где f_{X_i} есть плотность с.в. X_i , $i = 1, \dots, m$.

Теорема 1.9 утверждает, что если компоненты смеси упорядочены (стохастически или по интенсивности отказов), то смесь имеет соответствующие верхнюю и нижнюю границы.

Теорема 1.9. *Если ф.р. F_X с.в. X является смесью (1.105), то*

$$X_1 \leq_{st} \dots \leq_{st} X_m \quad \text{влечет} \quad X_1 \leq_{st} X \leq_{st} X_m; \quad (1.107)$$

$$X_1 \leq_r \dots \leq_r X_m \quad \text{влечет} \quad X_1 \leq_r X \leq_r X_m. \quad (1.108)$$

Доказательство. Из стохастического порядка компонент $X_1 \leq_{st} \dots \leq_{st} X_m$ сле-

дует, что $\bar{F}_{X_1}(x) \leq \dots \leq \bar{F}_{X_m}(x)$, и тогда

$$\bar{F}_{X_1}(x) \leq p_1 \bar{F}_{X_1}(x) + \dots + p_m \bar{F}_{X_m}(x) \leq \bar{F}_{X_m}(x).$$

С другой стороны, упорядоченность по интенсивности отказов $X_1 \leq_r X_i$, $i = 1, \dots, m$ означает, что

$$f_{X_i}(x) \leq \frac{f_{X_1}(x) \bar{F}_{X_i}(x)}{\bar{F}_{X_1}(x)}, \quad i = 1, \dots, m,$$

и, следовательно,

$$r_X(x) \leq \frac{p_1 f_{X_1}(x) \frac{\bar{F}_{X_2}(x)}{\bar{F}_{X_1}(x)} + \dots + p_m f_{X_1}(x) \frac{\bar{F}_{X_m}(x)}{\bar{F}_{X_1}(x)}}{p_1 \bar{F}_{X_1}(x) + p_2 \bar{F}_{X_2}(x) + \dots + p_m \bar{F}_{X_m}(x)} = r_{X_1}(x),$$

откуда следует, что $X_1 \leq_r X$. Аналогично, неравенства

$$f_{X_i}(x) \geq \frac{f_{X_m}(x) \bar{F}_{X_i}(x)}{\bar{F}_{X_m}(x)}, \quad i = 1, \dots, m,$$

и

$$r_X(x) \geq \frac{p_1 f_{X_m}(x) \frac{\bar{F}_{X_1}(x)}{\bar{F}_{X_m}(x)} + \dots + p_{m-1} f_{X_m}(x) \frac{\bar{F}_{X_{m-1}}(x)}{\bar{F}_{X_m}(x)} + p_m f_{X_m}(x)}{p_1 \bar{F}_{X_1}(x) + \dots + p_m \bar{F}_{X_m}(x)} = r_{X_m}(x)$$

влекут $X \leq_r X_m$, что завершает доказательство. $\square$

Рассмотрим распределение величины превышения уровня u для m -компонентной смеси. На событии $\{X_i > u\}$ определим с.в. $X_i^u = X_i - u \mid X_i > u$ с ф.р. величины превышения уровня u

$$P(X_i^u \leq x) = P(X_i - u \leq x \mid X_i > u), \quad u \geq 0, \quad x \geq 0.$$

Для m -компонентной смеси X с распределением (1.106) определим с.в. X^u , со-

ответствующую ф.р., хвост которой задан формулой

$$\begin{aligned}\bar{F}_X^u(x) &:= \mathbf{P}(X - u > x | X > u) = \frac{\bar{F}_X(x + u)}{\bar{F}_X(u)} \\ &= \frac{p_1 \bar{F}_{X_1}(x + u) + \cdots + p_m \bar{F}_{X_m}(x + u)}{p_1 \bar{F}_{X_1}(u) + \cdots + p_m \bar{F}_{X_m}(u)}, \quad u \geq 0, \quad x \geq 0.\end{aligned}$$

Теорема 1.10. *Если компоненты смеси (1.106) упорядочены по интенсивности отказов как*

$$X_1 \underset{r}{\leq} \cdots \underset{r}{\leq} X_m, \quad (1.109)$$

то для каждого $u \geq 0$ верно следующее утверждение:

$$X_1^u \underset{r}{\leq} X^u \underset{r}{\leq} X_m^u. \quad (1.110)$$

Доказательство. Известно, что если с.в. упорядочены по интенсивности отказов, то величины превышения уровня u также упорядочены по интенсивности отказов [116, Proposition 2.A.18]. Отсюда следует, что

$$X_1 \underset{r}{\leq} \cdots \underset{r}{\leq} X_m$$

влечет

$$X_i^u \underset{r}{\leq} \cdots \underset{r}{\leq} X_m^u.$$

Из соотношений (1.109), согласно утверждению (1.108), вытекают соотношения (1.110). $\square$

Следующая теорема позволяет стохастически сравнивать m -компонентные смеси с различными коэффициентами смесей.

Теорема 1.11. *Рассмотрим m -компонентные смеси распределений*

$$F_X(x) = p_1 F_{X_1}(x) + \cdots + p_m F_{X_m}(x),$$

$$F_Y(x) = q_1 F_{Y_1}(x) + \cdots + q_m F_{Y_m}(x),$$

с независимыми компонентами X_i, Y_i , $i = 1, \dots, m$. Предположим, что вы-

полнены следующие условия:

$$X_1 \underset{st}{\geq} \cdots \underset{st}{\geq} X_m \quad \text{или} \quad Y_1 \underset{st}{\geq} \cdots \underset{st}{\geq} Y_m; \quad (1.111)$$

$$X_i \underset{st}{\leq} Y_i, \quad i = 1, \dots, m; \quad (1.112)$$

$$\sum_{i=j}^m p_i \geq \sum_{i=j}^m q_i, \quad j = 2, \dots, m. \quad (1.113)$$

Тогда смеси стохастически упорядочены, а именно

$$X \underset{st}{\leq} Y. \quad (1.114)$$

Доказательство. Предположим, что $X_1 \underset{st}{\geq} \cdots \underset{st}{\geq} X_m$. Тогда из условий (1.111), (1.113) и Утверждения 2.1 в работе [130] следует, что

$$p_1 \bar{F}_{X_1}(x) + \cdots + p_m \bar{F}_{X_m}(x) \leq q_1 \bar{F}_{X_1}(x) + \cdots + q_m \bar{F}_{X_m}(x).$$

Поскольку условие (1.112) влечет $\bar{F}_{X_i}(x) \leq \bar{F}_{Y_i}(x)$, $i = 1, \dots, m$, то

$$q_1 \bar{F}_{X_1}(x) + \cdots + q_m \bar{F}_{X_m}(x) \leq q_1 \bar{F}_{Y_1}(x) + \cdots + q_m \bar{F}_{Y_m}(x),$$

откуда вытекает требуемое соотношение $X \underset{st}{\leq} Y$. Для стохастического порядка $Y_1 \underset{st}{\geq} \cdots \underset{st}{\geq} Y_m$ доказательство аналогично. $\square$

Для того, чтобы сравнивать m -компонентные смеси по интенсивности отказов, требуются более ограничительные условия [31], а именно соотношение

$$X \underset{r}{\leq} Y$$

верно, если

$$X_1 \underset{r}{\geq} \cdots \underset{r}{\geq} X_m \quad \text{либо} \quad Y_1 \underset{r}{\geq} \cdots \underset{r}{\geq} Y_m,$$

а также

$$\begin{aligned}\frac{\bar{F}_{Y_{i+1}}(x)}{\bar{F}_{X_{i+1}}(x)} &\leq \frac{\bar{F}_{Y_i}(x)}{\bar{F}_{X_i}(x)}, \quad i = 1, \dots, m-1, \quad x \geq 0, \\ X_i &\leq_r Y_i, \quad i = 1, \dots, m, \\ \frac{p_i}{p_j} &\leq \frac{q_i}{q_j}, \quad 1 \leq i \leq j \leq m.\end{aligned}$$

1.12. Многокомпонентное показательное-Парето распределение

Рассмотрим семейство независимых с.в. $X_1, \dots, X_{k+n}$ с ф.р. $F_i(x)$, $i = 1, \dots, k+n$, $k \geq 1, n \geq 1$. Пусть ф.р. с.в. X имеет вид

$$F_X(x) = p_1 F_{X_1}(x) + \dots + p_k F_{X_k}(x) + p_{k+1} F_{X_{k+1}}(x) + \dots + p_{k+n} F_{X_{k+n}}(x), \quad (1.115)$$

где $\sum_{i=1}^{k+n} p_i = 1$, $p_i \geq 0$.

Выражение (1.115) назовем $(k+n)$ -компонентной смесью распределений.

Определение 1.15. *С.в. X имеет $(k+n)$ -компонентное показательное-Парето распределение, если k ее компонент $X_1, \dots, X_k$ имеют показательное распределение*

$$F_{X_i}(x) = 1 - e^{-\lambda_i x}, \quad x \geq 0, \quad \lambda_i > 0, \quad i = 1, \dots, k,$$

а компоненты $X_{k+1}, \dots, X_{k+n}$ имеют распределение Парето

$$F_{X_j}(x) = 1 - \left(\frac{x_0}{x_0 + x} \right)^{\alpha_j}, \quad x \geq 0, \quad x_0 > 0, \quad \alpha_j > 0, \quad j = k+1, \dots, k+n.$$

Функция интенсивности отказов распределения (1.115) равна

$$r_X(x) = \frac{\sum_{i=1}^k p_i \lambda_i e^{-\lambda_i x} + \sum_{j=k+1}^{k+n} p_j \alpha_j x_0^{\alpha_j} / (x_0 + x)^{\alpha_j+1}}{\sum_{i=1}^k p_i e^{-\lambda_i x} + \sum_{j=k+1}^{k+n} p_j x_0^{\alpha_j} / (x_0 + x)^{\alpha_j}},$$

убывает по x , и, более того,

$$\sup_{x \geq 0} r_X(x) = \lim_{x \rightarrow 0} r_X(x) = \sum_{i=1}^k p_i \lambda_i + \sum_{j=k+1}^{k+n} p_j \alpha_j / x_0.$$

В следующей Теореме 1.12 приведены условия, гарантирующие, что ф.р. (1.115) ограничена сверху и снизу компонентами, имеющими минимальную и максимальную функции интенсивности отказов, соответственно.

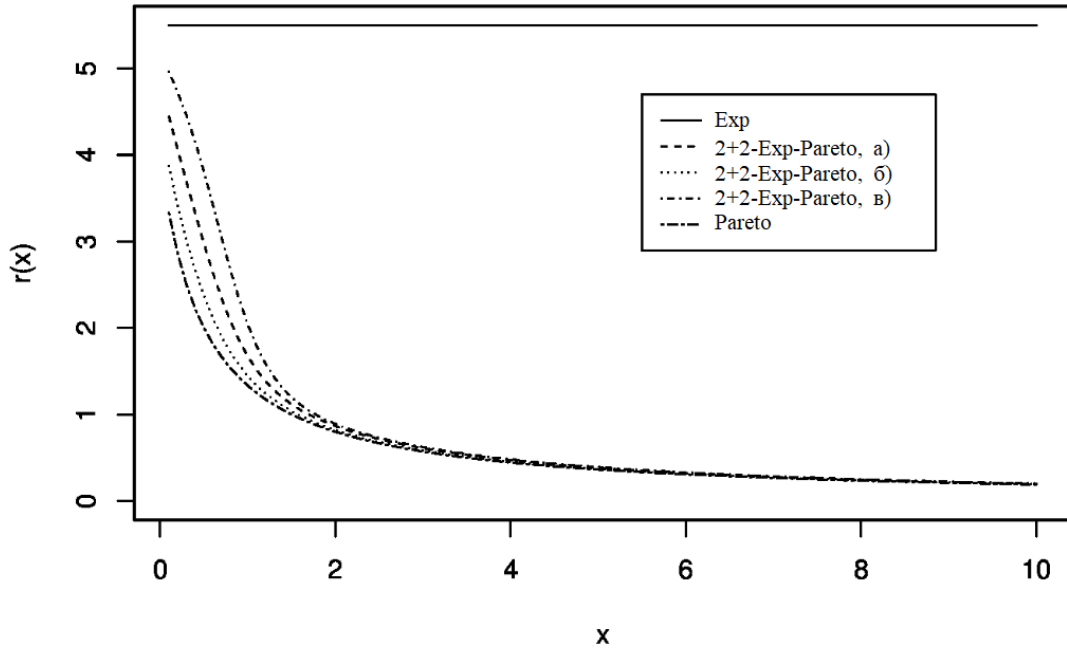


Рисунок 1.10. Функции интенсивности отказов $Exp(5.5)$, $Pareto(2, 0.5)$ и 2+2-компонентного показательного-Парето распределения с параметрами $x_0 = 0.5$, $\lambda_1 = 5.5$, $\lambda_2 = 5.0$, $\alpha_3 = 2.5$, $\alpha_4 = 2.0$ для различных значений пропорции смеси (а): $p_i = 0.25, i = 1, \dots, 4$; б): $p_1 = 0.10, p_2 = 0.15, p_3 = 0.15, p_4 = 0.60$, в): $p_1 = 0.60, p_2 = 0.15, p_3 = 0.15, p_4 = 0.10$.

Теорема 1.12. Если параметры распределения (1.115) удовлетворяют условиям

$$\lambda_1 \geq \dots \geq \lambda_k \geq \frac{\alpha_{k+1}}{x_0 + x} \geq \dots \geq \frac{\alpha_{k+n}}{x_0 + x}, \quad x > 0, \quad (1.116)$$

то

$$X_1 \underset{r}{\leq} X \underset{r}{\leq} X_{k+n}. \quad (1.117)$$

Доказательство. Доказательство следует из соотношения (1.108), поскольку неравенства (1.116) влекут требуемую упорядоченность функций интенсивно-

сти отказов. □

Рисунок 1.10 показывает упорядоченность функций интенсивности отказов распределения Парето, показательного распределения и $(2+2)$ -компонентного показательного-Парето распределения с параметрами $x_0 = 0.5$, $\lambda_1 = 5.5$, $\lambda_2 = 5.0$, $\alpha_3 = 2.5$, $\alpha_4 = 2.0$, и коэффициентами пропорции смеси при параметрах а): $p_i = 0.25, i = 1, \dots, 4$; б): $p_1 = 0.10, p_2 = 0.15, p_3 = 0.15, p_4 = 0.60$, в): $p_1 = 0.60, p_2 = 0.15, p_3 = 0.15, p_4 = 0.10$, соответственно. При таком наборе параметров соотношения (1.116) выполнены и распределения упорядочены по интенсивности отказов в соответствии с (1.117).

1.13. Выводы по результатам главы 1

В Главе 1 получены следующие теоретические результаты.

Доказано достаточное условие упорядоченности по интенсивности отказов, которое будет применяться в Главе 2 для получения свойств монотонности систем с конечным буфером (Лемма 1.1, раздел 1.4.) В Лемме 1.2 для различных классов распределений (различающимися типами функции интенсивности отказов) получены условия, при которых верна Лемма 1.1.

Доказано, что если компоненты конечной смеси распределений упорядочены (стохастически или по интенсивности отказов), то компоненты и смесь упорядочены соответствующим образом (Теорема 1.1, раздел 1.7.)

Методом каплинга доказано, что если соответствующие компоненты двух смесей упорядочены, то две смеси имеют стохастический порядок (Теорема 1.2). В Теореме 1.3 доказаны достаточные условия упорядоченности по интенсивности отказов смесей распределения. Теорема 1.4 утверждает, что если компоненты смеси являются УФИ-распределениями, то хвост распределения величины превышения порога u смеси возрастает с ростом u (раздел 1.7.)

Введено и исследовано новое показательное-Парето распределение. В частности, в Теореме 1.5 получены соотношения между параметрами показательного-Парето распределения, при которых компоненты смеси и сама смесь упорядочены по интенсивности отказов. Такая же упорядоченность сохраняется для распределений величины превышения уровня и их средних (раздел 1.8.)

Получены достаточные условия на параметры двухкомпонентных смесей распределения: двухкомпонентного Парето, двухкомпонентного гиперэкспоненциального и двухкомпонентного показательного-Парето распределений, гарантирующие их стохастический порядок (Теорема 1.6). В Теореме 1.7 получены достаточные условия для упорядоченности по интенсивности отказов указанных двухкомпонентных смесей (раздел 1.9.)

Доказаны Теоремы 1.9 и 1.10, в которых обобщены результаты для стохастического порядка и упорядоченности по интенсивности отказов на случай m -компонентных смесей распределений. В Теореме 1.11 получены достаточные условия, при которых стохастически упорядочены две m -компонентные смеси с различными коэффициентами пропорции смесей (раздел 1.11.)

Получены достаточные условия на параметры $(k + n)$ -компонентного показательного-Парето распределения, при которых гарантируется упорядоченность по интенсивности отказов смеси и ее компонент (Теореме 1.12, раздел 1.12.)

Основные результаты главы 1 опубликованы в работах автора [195, 197, 200, 201, 207, 208, 209].

Глава 2. Монотонность характеристик качества обслуживания

В данной главе исследуются свойства монотонности систем обслуживания, которые лежат в основе сравнения характеристик систем и позволяют вместо исследования характеристик исходных сложных систем (например, со смешанными распределениями) получать их оценки, используя аналитические результаты и методы моделирования более простых систем.

Исследуются свойства монотонности различных систем: с неограниченным буфером, с конечным буфером, со смешанным входным процессом, со смешанным обслуживанием, с повторными вызовами, а также с одновременным обслуживанием случайным числом серверов.

Свойства монотонности систем обслуживания, ввиду их важности, активно изучались в литературе и ранее ([34, 47, 167, 168, 177]). В частности, в работе [177] исследуются свойства монотонности многосерверных систем с ограниченным и неограниченным буфером, а в работе [47] рассмотрена монотонность систем с групповым входным потоком. Тем не менее, насколько нам известно, свойства монотонности систем со смешанными потоками ранее систематически не изучались, а также ранее не рассматривались системы, в которых входные потоки или времена обслуживания представляют собой смесь различных распределений.

Модели систем с постоянной частотой повторных вызовов используются в различных приложениях, например, для моделирования динамики систем беспроводной связи, систем телефонной станции [69]. В работах [49, 61, 62] аналогичная система используется для описания широкого класса протоколов множественного доступа. Более того, системы с постоянной частотой повторных вызовов использовались для описания ТСП-трафика с короткими HTTP-соединениями [41, 42] и оптико-электрической гибридной схемой разрешения конфликтов [178, 182]. Далее в диссертационной работе рассматриваются свойства внешней монотонности таких систем.

Модель высокопроизводительного кластера, т. е. многосерверной системы с одновременным обслуживанием несколькими серверами получила широкое распространение в связи с развитием распределенных систем. Отличительной

особенностью таких систем является то, что приходящему клиенту требуется случайное число серверов (работающих параллельно) на случайное время обслуживания. В частности, если клиенту требуется больше серверов, чем в данный момент свободно, он ожидает в очереди. Таким образом, дисциплина обслуживания является неконсервативной, поскольку незанятые серверы и непустая очередь могут сосуществовать. Отметим, что в статье [147] получен критерий стационарности высокопроизводительного кластера с произвольным числом серверов. Ниже используются свойства модифицированной рекурсии Кифера-Вольфовица [99] для процесса незавершенной работы для установления ее верхней и нижней границ.

Вопросы аппроксимации и корреляционной структуры систем со смешанным входным процессом (суперпозицией входных потоков) интенсивно изучались с 1970-х годов (см., например, [27, 37, 105, 110, 137, 179, 180]). Хорошо известно, что, если компоненты такого входного процесса не являются пуассоновскими, то система не является регенерирующей. С другой стороны, свойство регенерации позволяет точно оценивать стационарные характеристики систем [38, 82, 83, 158]. В Главе 6 регенерация таких систем будет описана подробно, а в данной Главе рассматривается свойство монотонности таких систем.

2.1. Понятие монотонности систем обслуживания

Пусть Σ – стохастическая модель системы обслуживания. В данном разделе частично используется материал и обозначения из книги [24]. Обозначим через c_Σ некоторую системную характеристику данной модели, т. е. характеристику, описывающую поведение системы в целом, которая зависит от времени и определена либо для всех t (например, средняя незавершенная нагрузка), либо лишь для некоторой последовательности (случайных) моментов $t_n, n \geq 1$ (например, время ожидания). Предположим, что характеристика системы c_Σ в момент времени t зависит от «управляющих» с.в. $\xi_1, \xi_2, \dots$ и начального состояния системы Σ_0

$$c_\Sigma(t) = c_\Sigma(t, \Sigma_0, \xi_1, \xi_2, \dots).$$

Пусть $\prec_c$ – некоторое отношение порядка на множестве возможных значений системной характеристики c_Σ . Отношением порядка может выступать сто-

хастический порядок, упорядоченность по интенсивности отказов, упорядоченность в среднем, рассмотренные в Главе 1, а также другие отношения порядка [116].

Стохастические модели могут обладать свойствами как внутренней, так и внешней монотонности.

Определение 2.1. ([24]) *Две стохастические модели Σ_1 и Σ_2 с одинаковой структурой и одинаковым начальным распределением обладают свойством внешней монотонности, если из упорядоченности управляющих с.в. $\xi_{j,1} \prec_j \xi_{j,2}, j = 1, \dots, t$, следует, что для системных характеристик выполняется соотношение $c_{\Sigma_1} \prec_c c_{\Sigma_2}$, где $c_{\Sigma_i} = c_{\Sigma}(\xi_{1,i}, \dots, \xi_{t,i}, \dots)$, $i = 1, 2$.*

Свойство *внешней монотонности* позволяет сравнивать между собой системные характеристики стохастических моделей с одинаковой структурой, но различными распределениями управляющих с.в.

Свойство внутренней монотонности характеризует связь управляющих с.в. и характеристик рассматриваемой системы. Так в односерверных системах обслуживания с неограниченной очередью свойство внутренней монотонности базируется на рекуррентной формуле Линдли для времени ожидания [34]:

$$W_{n+1} = \max(0, W_n + S_n - \tau_n) \quad n \geq 1. \quad (2.1)$$

В частности, известно следующее свойство внутренней монотонности.

Утверждение 2.1. ([24, теорема 5.1.1]) *Если ф.р. F_{W_1} выбрана так, что $W_1 \leq_{st} W_2$, то для всех $n \geq 2$*

$$W_n \leq_{st} W_{n+1},$$

и для стационарного времени ожидания W , если оно существует, выполнено соотношение

$$W_n \leq_{st} W, \quad n \geq 1.$$

Далее в качестве стохастических моделей будем рассматривать модели систем обслуживания и перейдем к изложению их свойств внешней монотонности. Обозначим через $GI/G/N/K$ систему обслуживания, в которой в случайные моменты времени $t_1 = 0 < t_2 < \dots$ приходят заявки на N обслуживающих

устройств. Пусть $\{\tau_n = t_{n+1} - t_n\}$ – н.о.р. интервалы между приходами заявок с ф.р. $A(x) = P(\tau \leq x)$, а $\{S_n\}$ – н.о.р. времена обслуживания с.ф.р. $B(x) = P(S \leq x)$. Также обозначим $\lambda = 1/M\tau$, $\mu = 1/MS$. Заявки поступают в систему в порядке первым пришел – первым обслужен (FIFS). Если все обслуживающие устройства заняты, то заявка помещается в буфер размером K . Если при этом все места в буфере заняты, то заявка покидает систему не обслуженной. К характеристикам, описывающим поведение систем обслуживания, относятся время ожидания W_n заявки с номером n , число заявок в очереди Q_n и число заявок в системе ν_n в момент времени t_n^- . В качестве управляющих последовательностей рассматриваются последовательность интервалов между приходами заявок $\{\tau_n, n \geq 1\}$, заданная ф.р. $A(x)$, а также последовательность времен обслуживания $\{S_n, n \geq 1\}$, заданная ф.р. $B(x)$.

2.2. Монотонность многосерверных систем обслуживания $GI/G/N/\infty$ с неограниченным буфером

Рассмотрим две системы обслуживания $\Sigma^{(1)}$ и $\Sigma^{(2)}$ вида $GI/G/N/\infty$, в каждой из которых N идентичных серверов обслуживания, работающих параллельно, а дисциплина обслуживания – FIFS. Обозначим $S_n^{(i)}$ – время обслуживания заявки с номером n в i -й системе, и $t_n^{(i)}$ – момент прихода в i -ю систему заявки с номером n , $i = 1, 2$. Н.о.р. времена между приходами заявок обозначим

$$\tau_n^{(i)} = t_{n+1}^{(i)} - t_n^{(i)}, \quad n \geq 1, \quad i = 1, 2.$$

Времена обслуживания $\{S_n^{(i)}, n \geq 1\}$ – н.о.р. с.в., и не зависят от $\{\tau_n^{(i)}, n \geq 1\}$. Через $Q_n^{(i)}$ обозначим число заявок в i -й системе и пусть

$$W_n^{(i)} = (W_n^{(i1)}, \dots, W_n^{(iN)}), \quad n \geq 1, \quad i = 1, 2,$$

вектор незавершенной нагрузки в системе i в момент времени $t_n^{(i)}$ – прихода n -й заявки. Другими словами, $W_n^{(ik)}$ – это незавершенная работа, которую должен выполнить сервер k в i -й системе в момент времени $t_n^{(i)-}$. Обозначим также через $\nu_n^{(i)}$ число заявок, а через $Q_n^{(i)}$ – размер очереди в i -й системе в момент

времени $t_n^{(i)-}$. Предположим, что существуют пределы:

$$Q_n^{(i)} \Rightarrow Q^{(i)}, \quad \nu_n^{(i)} \Rightarrow \nu^{(i)}, \quad W_n^{(i)} \Rightarrow W^{(i)}, \quad n \rightarrow \infty, \quad i = 1, 2.$$

Известно, что эти пределы существуют, если выполнено условие отрицательного сноса [34, 127]

$$MS^{(i)} < NM\tau^{(i)}, \quad (2.2)$$

и (при числе серверов $N > 1$) также условие регенерации

$$P(\tau^{(i)} > S^{(i)}) > 0, \quad i = 1, 2. \quad (2.3)$$

(При $N = 1$ последнее условие следует из (2.2).) При условии (2.3) (дискретная) длина цикла регенерации является непериодической с.в. [34, 127].

Следующая Теорема 2.1 является модификацией Теорем 4 и 5 из работы [177], в которых вместо стохастической упорядоченности времен обслуживания предполагается упорядоченность по интенсивности отказов.

Теорема 2.1. Пусть в системах $\Sigma^{(1)}$ и $\Sigma^{(2)}$ выполнены условия (2.2), (2.3). Тогда:

1. Если выполнены соотношения

$$\nu_1^{(1)} = \nu_1^{(2)} = 0, \quad \tau_{st}^{(1)} = \tau_{st}^{(2)}, \quad S_r^{(1)} \leq S_r^{(2)}, \quad (2.4)$$

то

$$\nu_{st}^{(1)} \leq \nu_{st}^{(2)}, \quad Q_{st}^{(1)} \leq Q_{st}^{(2)}. \quad (2.5)$$

2. Если выполнены соотношения

$$W_{st}^{(1)} = W_{st}^{(2)} = 0, \quad \tau_r^{(1)} \geq \tau_r^{(2)}, \quad S_r^{(1)} \leq S_r^{(2)}, \quad (2.6)$$

то

$$W_{st}^{(1)} \leq W_{st}^{(2)}. \quad (2.7)$$

Доказательство Теоремы 2.1 очевидно, поскольку из упорядоченности по интенсивности отказов времен обслуживания вытекает их стохастический порядок [145]. Тогда соотношение (2.5) следует из Теоремы 5 [177], а неравенство

(2.7) – из Теоремы 4 [177].

Упорядоченность по интенсивности отказов является более ограничительным условием по сравнению со стохастическим порядком, но часто оказывается более удобным для практического применения, поскольку для многих распределений позволяет найти явные соотношения между параметрами.

2.3. Монотонность многосерверных систем обслуживания $GI/G/N/K$ с конечным буфером

В этом разделе рассмотрим свойства монотонности многосерверных систем с конечным буфером. Рассмотрим две системы $\Sigma^{(1)}$ и $\Sigma^{(2)}$, в которых $N^{(i)}$ серверов работают параллельно, а размеры буферов равны $K^{(i)}$, $i = 1, 2$. Все остальные обозначения сохраняются как в разделе 2.2.

Следующая Теорема 2.2 является модификацией Теоремы 6 из [177] и позволяет сравнить длины очередей двух систем с конечным буфером с разными распределениями входного потока и времени обслуживания.

Теорема 2.2. *Предположим, что в системах $\Sigma^{(1)}$ и $\Sigma^{(2)}$ выполнены следующие условия:*

$$K^{(1)} \geq K^{(2)}, \quad K^{(1)} + N^{(1)} \leq K^{(2)} + N^{(2)}, \quad (2.8)$$

а также условие упорядоченности по интенсивности отказов для интервалов между приходами заявок

$$\tau^{(1)} \underset{r}{\geq} \tau^{(2)}. \quad (2.9)$$

Пусть также функции интенсивностей отказов времен обслуживания удовлетворяют неравенству

$$\inf_{x \geq 0} r_{S^{(1)}}(x) \geq \sup_{x \geq 0} r_{S^{(2)}}(x). \quad (2.10)$$

Тогда стационарные длины очередей стохастически упорядочены как

$$Q_{st}^{(1)} \leq Q_{st}^{(2)}. \quad (2.11)$$

Доказательство вытекает из Теоремы 6 [177], если заметить, что, согласно Лемме 1.1, из неравенства (2.10) следует

$$S^{(1)} \leq_r S^{(2)}. \quad (2.12)$$

Пример 2.1. Сравним ф.р. стационарных длин очередей систем $M/M/N/K$ (система $\Sigma^{(1)}$) и $M/G/N/K$ (система $\Sigma^{(2)}$ с произвольным распределением времени обслуживания). Для первой системы известно [13] выражение для стационарного распределения длины очереди $Q^{(1)}$:

$$P(Q^{(1)} = i) = \begin{cases} \pi_0 \frac{\rho_1^i}{i!}, & i = 0, \dots, N, \quad N \geq 1, \\ \pi_0 \frac{\rho_1^i}{N^{i-N} N!}, & i = N + 1, \dots, N + K, \end{cases} \quad (2.13)$$

где $\rho_1 = MS^{(1)}/M\tau^{(1)}$ и

$$\pi_0 = \left[\sum_{k=0}^N \frac{\rho_1^k}{k!} + \frac{\rho_1^N}{N!} \sum_{k=N+1}^{N+K} \left(\frac{\rho_1}{N}\right)^{k-N} \right]^{-1}.$$

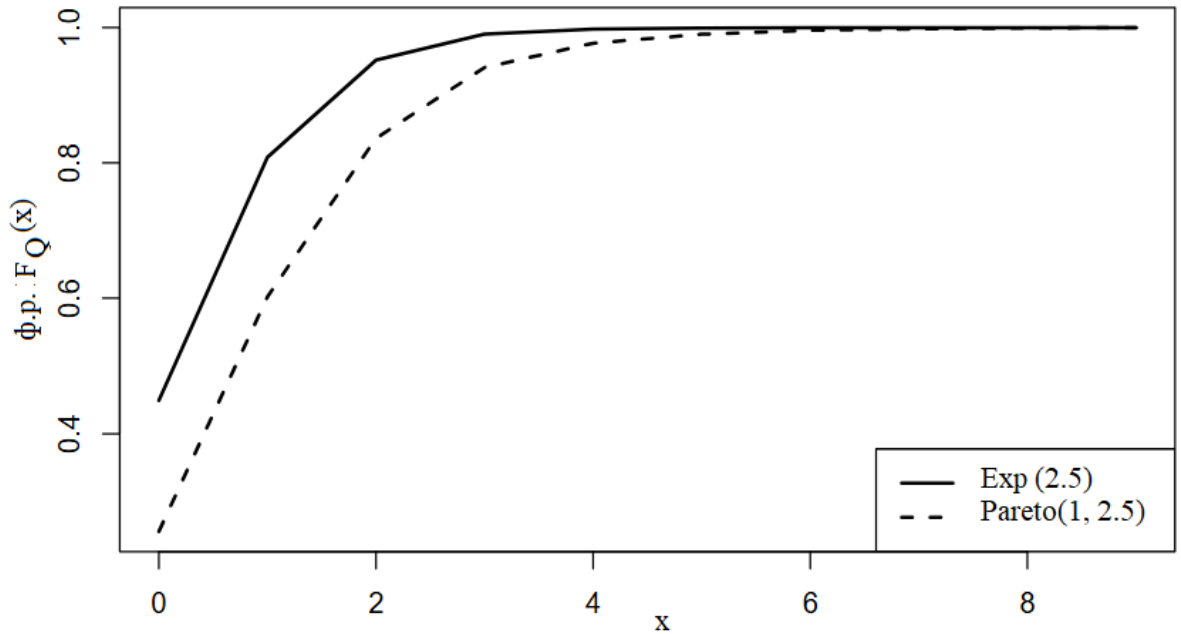


Рисунок 2.1. Сравнение эмпирической ф.р. длины очереди $Q^{(1)}$ в системе $M/G/4/5$ с распределением времени обслуживания $Pareto(1, 2.5)$ с эмпирической ф.р. длины очереди $Q^{(2)}$ в системе $M/M/4/5$ с распределением времени обслуживания $Exp(2.5)$; входной поток - пуассоновский с параметром $\lambda_\tau = 2$ в обеих системах.

В численном эксперименте сравниваются эмпирические ф.р. стационарной длины очереди двух систем при $N = 4, K = 5$. В системе $\Sigma^{(2)}$ время обслуживания имеет распределение Парето с параметрами $\alpha_s = 2.5, x_0 = 1$, а в системе $\Sigma^{(1)}$ – показательное с параметром $\lambda_s = 2.5$. Входной поток в обеих системах – пуассоновский с параметром $\lambda_\tau = 2$. Заметим, что $\inf_{x \geq 0} r_{S^{(1)}}(x) = \lambda_s = 2.5$ и, поскольку распределение Парето относится к классу УФИ-распределений, то

$$\sup_{x \geq 0} r_{S^{(2)}}(x) = r_{S^{(2)}}(0) = \alpha_s/x_0 = 2.5.$$

Следовательно, условие (2.10) Теоремы 2.2 выполнено.

Результаты численных экспериментов на Рисунке 2.1 подтверждают, что длина очереди в системе $\Sigma^{(1)}$ стохастически меньше, чем в системе $\Sigma^{(2)}$, т. е. $Q^{(1)} \leq_{st} Q^{(2)}$.

Пример 2.2. Численное моделирование двух систем $G/G/4/5$. В первой системе $\Sigma^{(1)}$ с.в. $\tau^{(1)}$ распределена по закону Бурра [187], $\tau^{(1)} \sim Burr(x_1, \alpha_1, k)$, с ф.р.

$$F_{\tau^{(1)}}(x) = 1 - \left(\frac{x_1}{x_1 + x^k} \right)^{\alpha_1}, \quad x \geq 0, \quad x_1 > 0, \quad \alpha_1 > 0, \quad k > 0, \quad (2.14)$$

для которого момент m -го порядка существует при $k\alpha_1 > m$. Время обслуживания $S^{(1)}$ в системе $\Sigma^{(1)}$ имеет показательное распределение с параметром λ . В системе $\Sigma^{(2)}$ интервалы между приходами заявок $\tau^{(2)}$ имеют двухкомпонентное гиперэкспоненциальное распределение с параметрами μ_1, μ_2 и p_1, p_2 , а время на обслуживания $S^{(2)}$ – распределение $Pareto(\alpha_2, x_2)$. Заметим, что функция интенсивности отказов с.в. $\tau^{(1)}$ имеет вид

$$r_{\tau^{(1)}}(x) = \frac{\alpha_1 k x^{k-1}}{x_1 + x^k}, \quad x \geq 0. \quad (2.15)$$

При $k \geq 2$ функция $r_{\tau^{(1)}}(x)$ возрастает, достигая максимума

$$r_{\tau^{(1)}}(x^*) = \frac{\alpha_1 k (x_1(k-1))^{1-1/k}}{x_1 k} \quad (2.16)$$

в точке

$$x^* = (x_1(k-1))^{1/k},$$

после которой убывает. При $k < 1$ функция интенсивности отказов монотонно убывает. В этом случае распределение Бурра относится к классу УФИ-распределений.

Функция интенсивности отказов гиперэкспоненциальной с.в. $\tau^{(2)}$ имеет вид

$$r_{\tau^{(2)}}(x) = \frac{p_1\mu_1 e^{-\mu_1 x} + p_2\mu_2 e^{-\mu_2 x}}{p_1 e^{-\mu_1 x} + p_2 e^{-\mu_2 x}} \quad (2.17)$$

и с ростом x монотонно сходится к наименьшему параметру, т. е.

$$\lim_{x \rightarrow \infty} r_{\tau^{(2)}}(x) = \min(\mu_1, \mu_2) =: \mu^*. \quad (2.18)$$

Чтобы получить стохастическую упорядоченность интервалов (2.9) заметим, что поскольку

$$r_{\tau^{(1)}}(x) \leq r_{\tau^{(1)}}(x^*), \quad x \geq 0,$$

то из неравенства

$$r_{\tau^{(1)}}(x^*) = \frac{\alpha_1 k (x_1(k-1))^{1-1/k}}{x_1 k} \leq \mu^*$$

следует $r_{\tau^{(1)}}(x) \leq r_{\tau^{(2)}}(x)$, $x \geq 0$. Следовательно, условие

$$\alpha_1 \leq \frac{x_1 k}{k(x_1(k-1))^{1-1/k}} \mu^* \quad (2.19)$$

влечет упорядоченность по интенсивности отказов

$$\tau^{(1)} \underset{r}{\geq} \tau^{(2)}.$$

Теперь найдем соотношение между параметрами распределений времен обслуживания в двух системах так, чтобы выполнялось условие (2.10). Вычислим

$$\sup_x r_{S^{(2)}}(x) = r_{S^{(2)}}(0) = \frac{\alpha_2}{x_2}$$

и

$$\inf_x r_{S^{(1)}}(x) = \lambda.$$

Следовательно, неравенство

$$\frac{\alpha_2}{x_2} \leq \lambda \quad (2.20)$$

влечет условие (2.10).

Рассмотрим численный пример с параметрами

$$x_1 = 1, k = 3, p_1 = p_2 = 0.5, \mu_1 = 3, \mu_2 = 30, \lambda = 2.5, x_2 = 1, \alpha_2 = 2.5,$$

для которых соотношения (2.19) и (2.20) выполнены. Следовательно, все условия Теоремы 2.2 выполнены, что влечет стохастический порядок стационарных длин очередей, $Q^{(1)} \leq_{st} Q^{(2)}$. В эксперименте параметр $\alpha_2 \in \{1.5, 2, 2.5\}$ так, что условие (2.20) выполнено. Результаты моделирования показывают, что равномерная метрика (расстояние) между эмпирическими ф.р. растет с уменьшением параметра α_2 , см. Рисунок 2.2.

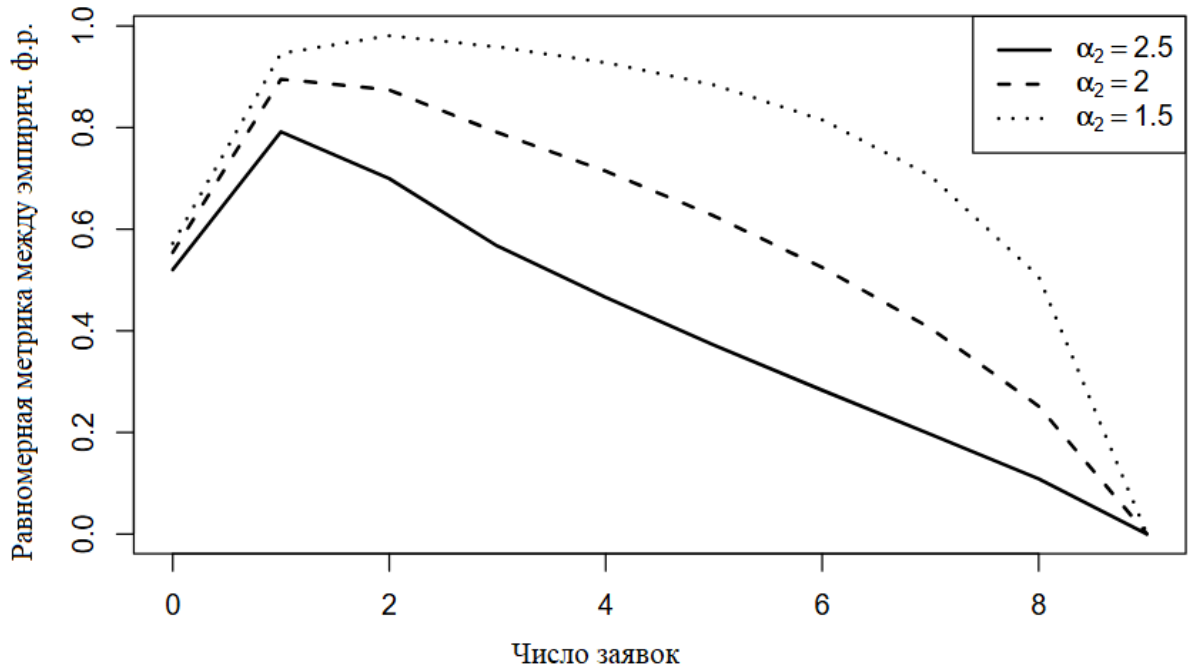


Рисунок 2.2. Равномерная метрика между двумя эмпирическими ф.р. стационарной длины очереди в двух системах $G/G/4/5$: $Burr(x_1, \alpha_1, k)/Exp(\lambda)/4/5$ и $H_2/Pareto(\alpha_2, x_2)/4/5$ (для различных α_2), где $x_1 = 1, k = 3, p_1 = p_2 = 0.5, \mu_1 = 3, \mu_2 = 30, \alpha_1 = 1.8, \lambda = 2.5$ и $x_2 = 1, \alpha_2 \in \{1.5, 2, 2.5\}$.

2.4. Монотонность в системах со смешанным обслуживанием

В данном разделе обобщим утверждение Теоремы 2.1 на систему $GI/G/N/\infty$, в которой времена обслуживания заданы конечной смесью рас-

пределений. Пусть в исходной системе, обозначенной через Σ , ф.р. времени обслуживания является m -компонентной смесью

$$F_S(x) = p_1 F_{S^{(1)}}(x) + \dots + p_m F_{S^{(m)}}(x), \quad x \geq 0, \quad \sum_{k=1}^m p_k = 1, \quad (2.21)$$

где независимые с.в. $S^{(i)}$ заданы ф.р. $F_{S^{(i)}}(x)$. Пусть входной процесс задан последовательностью $\{t_n\}$ моментов прихода заявок в систему и н.о.р интервалами между приходами $\tau_n = t_{n+1} - t_n$, $n \geq 1$. Наконец, предположим, что выполнено условие стационарности

$$MS < NM\tau. \quad (2.22)$$

Построим две новые системы: $\Sigma^{(1)}$ со временем обслуживания $S^{(1)}$ и $\Sigma^{(2)}$ со временем обслуживания $S^{(m)}$. Обозначим Q_n , $Q_n^{(1)}$, $Q_n^{(m)}$ – размер очереди, ν_n , $\nu_n^{(1)}$, $\nu_n^{(m)}$ – число заявок в системе, W_n , $W_n^{(1)}$, $W_n^{(m)}$ – незавершенная работа в момент времени t_n^- в системах Σ , $\Sigma^{(1)}$ и $\Sigma^{(2)}$, соответственно. Пусть существуют стационарные пределы

$$Q_n^{(i)} \Rightarrow Q^{(i)}, \quad W_n^{(i)} \Rightarrow W^{(i)}, \quad n \rightarrow \infty, \quad i = 1, m.$$

Теорема 2.3. *Предположим, что для исходной системы Σ выполнены условия*

$$\begin{aligned} \nu_1^{(1)} &= \nu_1^{(m)} = \nu_1 = 0, \\ \tau_{st}^{(1)} &= \tau_{st}^{(m)} = \tau, \\ S^{(1)} &\underset{r}{\leq} S^{(2)} \underset{r}{\leq} \dots \underset{r}{\leq} S^{(m)}. \end{aligned} \quad (2.23)$$

тогда

$$\begin{aligned} Q^{(1)} &\underset{st}{\leq} Q \underset{st}{\leq} Q^{(m)}, \\ W^{(1)} &\underset{st}{\leq} W \underset{st}{\leq} W^{(m)}. \end{aligned} \quad (2.24)$$

Утверждение теоремы следует из Теоремы 1.9 (раздел 1.11.), согласно которой времена обслуживания упорядочены по интенсивности отказов как

$$S^{(1)} \underset{r}{\leq} S \underset{r}{\leq} S^{(m)},$$

а, следовательно, из Теоремы 2.1 (раздел 2.2.) следует, что для стационарного времени ожидания и стационарного размера очереди справедлив стохастический порядок

$$Q^{(1)} \underset{st}{\leq} Q \underset{st}{\leq} Q^{(m)}, \quad W^{(1)} \underset{st}{\leq} W \underset{st}{\leq} W^{(m)}.$$

Таким образом, система $\Sigma^{(1)}$ является минорантной, а система $\Sigma^{(2)}$ – мажорантной для исходной системы Σ .

Пример 2.3. Пусть в исходной системе Σ время обслуживания имеет показательное-Парето распределение с ф.р. (1.51). Построим минорантную систему $\Sigma^{(1)}$, в которой время обслуживания $S^{(1)}$ – показательное с параметром λ_s , и мажорантную систему $\Sigma^{(2)}$, в которой время обслуживания $S^{(2)}$ имеет распределение $Pareto(\alpha_s, x_0)$. Входной поток во все три системы одинаковый, $\tau^{(1)} = \tau = \tau^{(2)}$, пуассоновский с параметром λ .

Предположим, что выполнено неравенство для параметров исходной системы:

$$\lambda_s \geq \frac{\alpha_s}{x_0}. \quad (2.25)$$

Тогда по Теореме 1.5 (раздел 1.8.), времена обслуживания упорядочены по интенсивности отказов

$$S^{(1)} \underset{r}{\leq} S \underset{r}{\leq} S^{(2)}.$$

Более того, при условии (2.25) коэффициенты загрузки $\rho_i = MS^{(i)}/M\tau$ удовлетворяют неравенствам

$$\rho_1 = \frac{1}{\lambda_s M\tau} \leq \rho \leq \rho_2 = \frac{x_0}{M\tau(\alpha_s - 1)}, \quad \alpha_s > 1, \quad (2.26)$$

где

$$\rho = \lambda_\tau \left(\frac{p}{\lambda_s} + \frac{(1-p)x_0}{\alpha_s - 1} \right).$$

Также дополнительно предположим, что $\rho_2 < N$, тогда все три системы стационарны. Предположим, что начальные состояния систем совпадают, т. е. $\nu_1^{(1)} = \nu_1^{(2)} = \nu_1$. Из Теоремы 2.3 вытекает стохастический порядок длин очереди и незавершенной нагрузки,

$$Q^{(1)} \underset{st}{\leq} Q \underset{st}{\leq} Q^{(2)} \quad \text{и} \quad W^{(1)} \underset{st}{\leq} W \underset{st}{\leq} W^{(3)}. \quad (2.27)$$

Рассмотрим численный пример для многосерверной системы $M/G/4/\infty$ с параметрами

$$\alpha_s = 2.5, \lambda_s = 5, x_0 = 0.5, p = 0.5, \lambda_\tau = 10.$$

Таким образом, время обслуживания в минорантной системе $\Sigma^{(1)}$ имеет распределение $Exp(5)$, а в мажорантной системе – распределение $Pareto(2.5, 0.5)$. Заметим, что условие (2.25) выполнено, и что все три системы стационарны, так как

$$\rho_1 = 2 < \rho \approx 2.7 < \rho_2 \approx 3.3 < 4.$$

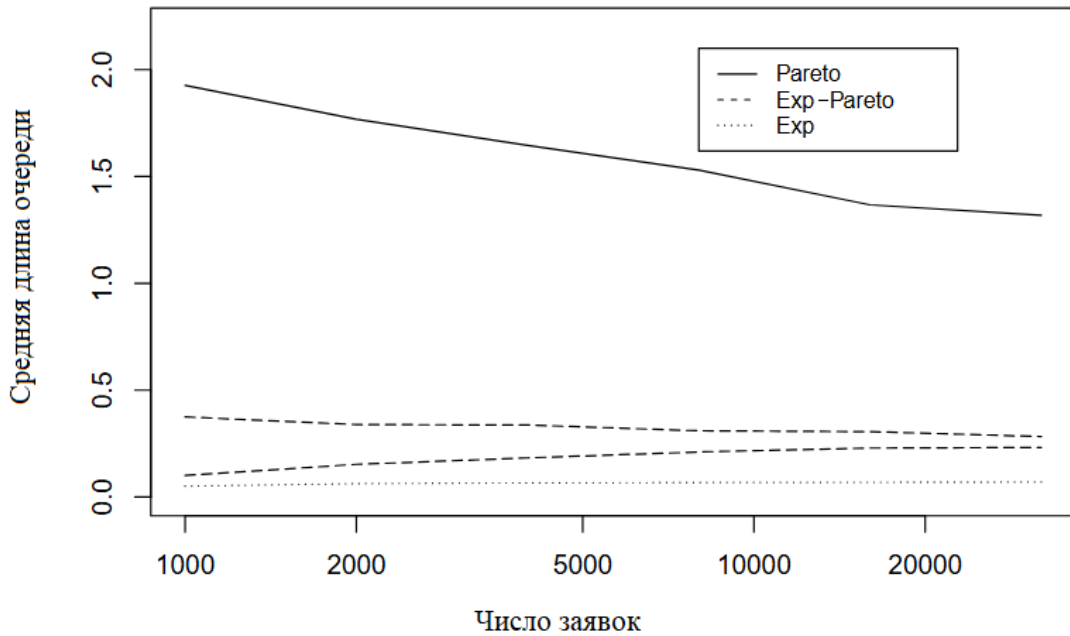


Рисунок 2.3. Нижняя и верхняя границы доверительного интервала для средней длины очереди в 4-х серверной системе с показательным-Парето распределением времени обслуживания.

На Рисунке 2.3 показаны границы доверительного интервала для средней длины стационарной очереди в исходной системе: верхняя линия соответствует правой границе доверительного интервала для мажорантной системы, нижняя граница соответствует левой границе доверительного интервала для минорантной системы. Две пунктирные линии – доверительный интервал для средней стационарной длины очереди в исходной системе. Доверительные интервалы построены методом регенеративного моделирования. Рисунок 2.4 демонстрирует нижнюю и верхнюю границы (построенные для минорантной и мажорантной

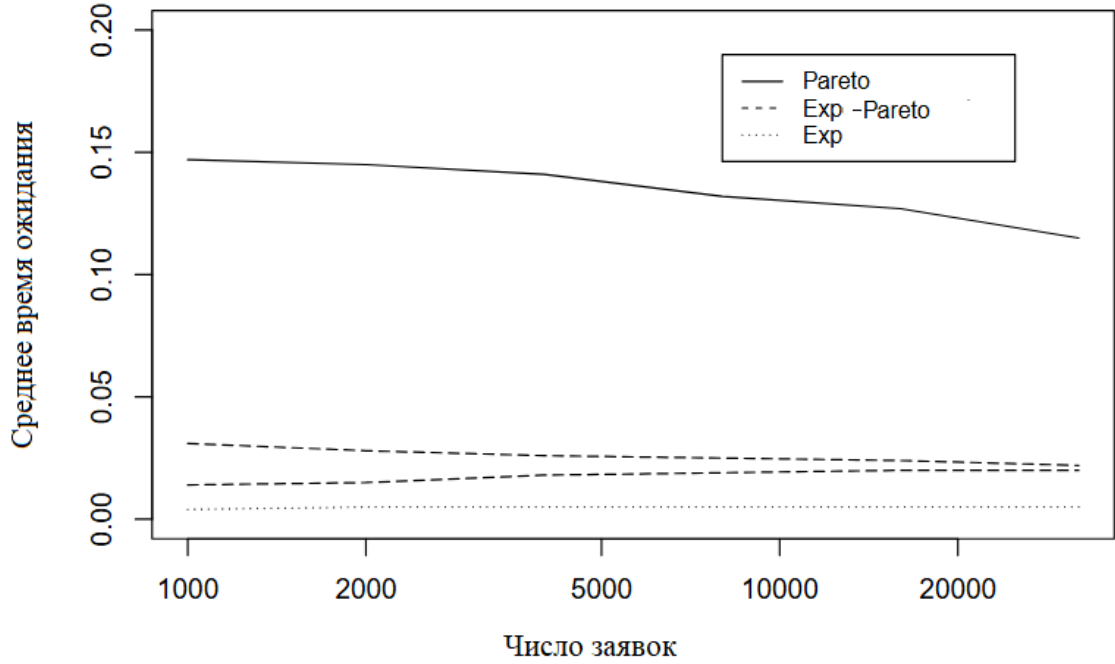


Рисунок 2.4. Нижняя и верхняя границы доверительного интервала для среднего времени ожидания в 4-х серверной системе с показательным-Парето распределением времени обслуживания.

систем, соответственно) доверительного интервала для среднего стационарного времени ожидания в исходной системе с показательным-Парето временем обслуживания.

Таким образом, результаты численных экспериментов иллюстрируют утверждения Теоремы 2.3.

Пример 2.4. Пусть в исходной системе Σ время обслуживания S имеет $(k+n)$ -компонентное показательное-Парето распределение с параметрами $\lambda_i, \alpha_j, x_0, p_i, p_j, i = 1, \dots, k, j = k+1, \dots, k+n$, и ф.р. (1.115). Пусть также выполнено условие (1.116). Тогда из Теоремы 1.12 (раздел 1.12.) следует упорядоченность по интенсивности отказов

$$S^{(1)} \underset{r}{\leq} S \underset{r}{\leq} S^{(k+n)}.$$

Построим две новые системы $\Sigma^{(1)}$ и $\Sigma^{(2)}$ следующим образом. В системе $\Sigma^{(1)}$ время обслуживания $S^{(1)}$ распределено по показательному закону с параметром λ_1 , а в системе $\Sigma^{(2)}$ время обслуживания $S^{(k+n)}$ имеет распределение $Pareto(\alpha_{k+n}, x_0)$. Входной поток во всех трех системах – пуассоновский с пара-

метром λ_τ .

Из Теорем 1.12 и 2.3 следует, что при условии (1.116), стационарные длины очередей и времена ожидания упорядочены как

$$Q^{(1)} \leq_{st} Q \leq_{st} Q^{(k+n)}, \quad W^{(1)} \leq_{st} W \leq_{st} W^{(k+n)}. \quad (2.28)$$

На Рисунке 2.5 представлены границы доверительного интервала для среднего размера очереди в исходной системе Σ с 2+2-компонентным показательным-Парето временем обслуживания с параметрами

$$p_i = 0.25, \quad x_0 = 0.5, \quad \lambda_1 = 5.5, \quad \lambda_2 = 5.0, \quad \alpha_3 = 2.5, \quad \alpha_4 = 2.0. \quad (2.29)$$

Во всех трех системах интенсивность входного потока равна $\lambda_\tau = 10$, число

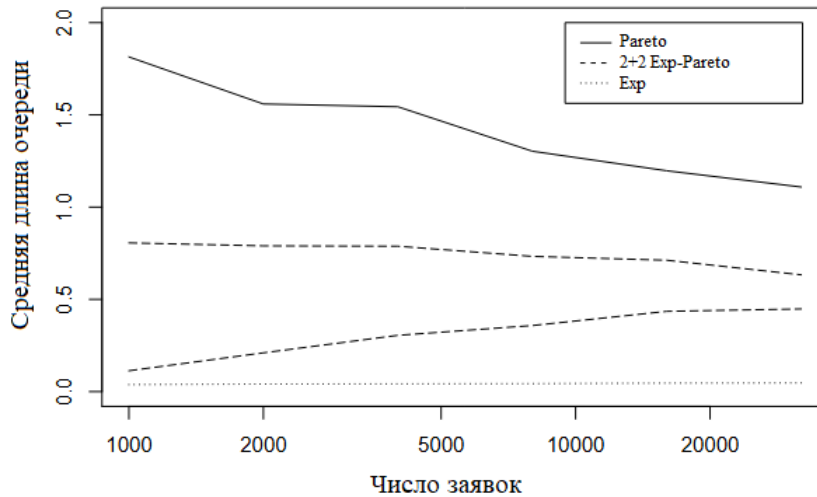


Рисунок 2.5. Доверительные интервалы для средней длины очереди в системах Σ , $\Sigma^{(1)}$, $\Sigma^{(2)}$; $N = N_1 = N_2 = 4$,

серверов $N = 4$. Из соотношений (2.28) следует, что система $\Sigma^{(1)}$ является минорантной, а система $\Sigma^{(2)}$ – мажорантной для системы Σ . Проверим условия стационарности систем. Обозначим через ρ , ρ_1 , ρ_2 коэффициенты загрузки в системах Σ , $\Sigma^{(1)}$, $\Sigma^{(2)}$, соответственно. При заданных параметрах легко проверить, что

$$\rho_1/N \approx 0.45 < \rho/N \approx 0.75 < \rho_2/N \approx 0.83.$$

На Рисунке 2.6 приведены доверительный интервал для среднего времени ожидания в системе Σ , а также верхняя и нижняя границы доверительных интер-

валов для среднего времени ожидания в системах $\Sigma^{(2)}$ и $\Sigma^{(1)}$, соответственно. Доверительные интервалы построены методом регенеративного моделирования с доверительной вероятностью 0.95, а число заявок в эксперименте равнялось 10^4 .

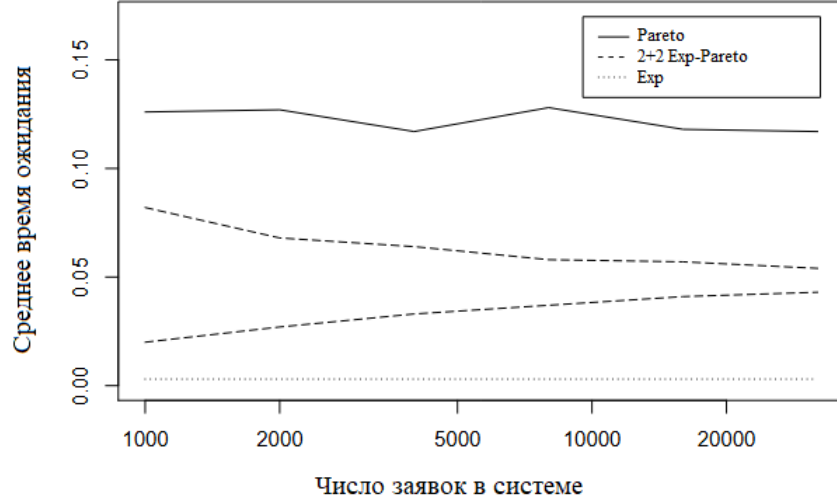


Рисунок 2.6. Доверительные интервалы для среднего времени ожидания в системах Σ , $\Sigma^{(1)}$, $\Sigma^{(2)}$; $N = N_1 = N_2 = 4$.

Рассмотрим две системы, Σ и $\tilde{\Sigma}$, в которых ф.р. времени обслуживания являются смесями вида

$$F_S(x) = p_1 F_{S^{(1)}}(x) + \cdots + p_m F_{S^{(m)}}(x), \quad x \geq 0, \quad \sum_{k=1}^m p_k = 1,$$

$$F_{\tilde{S}}(x) = q_1 F_{\tilde{S}^{(1)}}(x) + \cdots + q_m F_{\tilde{S}^{(m)}}(x), \quad x \geq 0, \quad \sum_{j=1}^m q_j = 1,$$

где независимые с.в. $S^{(i)}$, $\tilde{S}^{(i)}$ заданы ф.р. $F_{S^{(i)}}(x)$, $F_{\tilde{S}^{(i)}}(x)$, соответственно. Следующая теорема устанавливает условия, при которых в таких системах стационарное время ожидания и стационарная длина очереди стохастически упорядочены.

Теорема 2.4. *Предположим, что для стационарных систем Σ и $\tilde{\Sigma}$ выполнены*

условия

$$\begin{aligned} \nu_1 &= \tilde{\nu}_1 = 0, \\ \tau &\stackrel{st}{=} \tilde{\tau}, \\ S^{(1)} &\stackrel{st}{\geq} \dots \stackrel{st}{\geq} S^{(m)} \quad \text{или} \quad \tilde{S}^{(1)} \stackrel{st}{\geq} \dots \stackrel{st}{\geq} \tilde{S}^{(m)}; \end{aligned} \quad (2.30)$$

$$S^{(i)} \stackrel{st}{\leq} \tilde{S}^{(i)}, \quad i = 1, \dots, m; \quad (2.31)$$

$$\sum_{i=j}^m p_i \geq \sum_{i=j}^m q_i, \quad j = 2, \dots, m. \quad (2.32)$$

Тогда

$$\begin{aligned} Q &\stackrel{st}{\leq} \tilde{Q}, \\ W &\stackrel{st}{\leq} \tilde{W}. \end{aligned} \quad (2.33)$$

Для доказательства теоремы заметим, что согласно Теореме 1.11 (раздел 1.11.), при выполнении условий (2.30)-(2.32) времена обслуживания в этих системах имеют стохастический порядок, т.е.

$$S \stackrel{st}{\leq} \tilde{S}.$$

Далее, применяя теоремы 4 и 5 [177], получаем (2.33).

Сформулируем свойства монотонности для системы с конечным буфером $GI/G/N/K$, в которой времена обслуживания заданы m -компонентной смесью распределений (2.21).

Построим две новые системы: минорантную систему $\Sigma^{(1)}$ со временем обслуживания $S^{(1)}$ и мажорантную систему $\Sigma^{(m)}$ со временем обслуживания $S^{(m)}$. Обозначим через Q_n , $Q_n^{(1)}$, $Q_n^{(m)}$ число заявок в очереди в системах Σ , $\Sigma^{(1)}$ и $\Sigma^{(m)}$, соответственно. Пусть существуют стационарные пределы

$$Q_n \Rightarrow Q, \quad Q_n^{(i)} \Rightarrow Q^{(i)}, \quad i = 1, m,$$

а $\tau, \tau^{(1)}, \tau^{(m)}$ – интервалы входных потоков в системах $\Sigma, \Sigma^{(1)}, \Sigma^{(m)}$, соответственно.

Теорема 2.5. Пусть в исходной системе Σ время обслуживания S задано

m -компонентной смесью (2.21). Предположим, что входные потоки в минорантную систему $\Sigma^{(1)}$, мажорантную систему $\Sigma^{(m)}$ и исходную систему Σ удовлетворяют неравенствам:

$$\tau^{(1)} \underset{r}{\geq} \tau \underset{r}{\geq} \tau^{(m)}, \quad (2.34)$$

а времена обслуживания $S^{(1)}$, S и $S^{(m)}$ удовлетворяют неравенствам

$$S^{(1)} \underset{r}{\leq} \dots \underset{r}{\leq} S^{(m)}. \quad (2.35)$$

Пусть выполнены соотношения

$$C^{(1)} \geq C \geq C^{(m)}, \quad C^{(1)} + N^{(1)} \leq C + N \leq C^{(m)} + N^{(m)}. \quad (2.36)$$

Тогда

$$1. \text{ Если } \inf_{x \geq 0} r_{S^{(1)}}(x) \geq \sup_{x \geq 0} r_S(x), \quad (2.37)$$

$$\text{то } Q^{(1)} \underset{st}{\leq} Q. \quad (2.38)$$

$$2. \text{ Если } \inf_{x \geq 0} r_S(x) \geq \sup_{x \geq 0} r_{S^{(m)}}(x), \quad (2.39)$$

$$\text{то } Q \underset{st}{\leq} Q^{(m)}. \quad (2.40)$$

Доказательство следует из Теоремы 1.9 (раздел 1.11.), согласно которой времена обслуживания упорядочены по интенсивности отказов как

$$S^{(1)} \underset{r}{\leq} S \underset{r}{\leq} S^{(m)},$$

и утверждение данной теоремы следует из Теоремы 2.2 (раздел 2.3.)

Пример 2.5. Пусть в исходной системе Σ время обслуживания имеет $k + n$ -компонентное показательное-Парето распределение (1.115) с упорядоченными компонентами, т. е.

$$S^{(1)} \underset{r}{\leq} \dots \underset{r}{\leq} S^{(k+n)}.$$

Очевидно, при условии

$$\inf_x r_{S^{(1)}}(x) = \lambda_1 \geq \sum_{i=1}^k p_i \lambda_i + \sum_{j=k+1}^{k+n} p_j \alpha_j / x_0 = \sup_x r_S(x), \quad (2.41)$$

выполнено соотношение (2.37). Следовательно, стационарные длины очередей стохастически упорядочены и утверждение (2.38) вытекает из Теоремы 2.5.

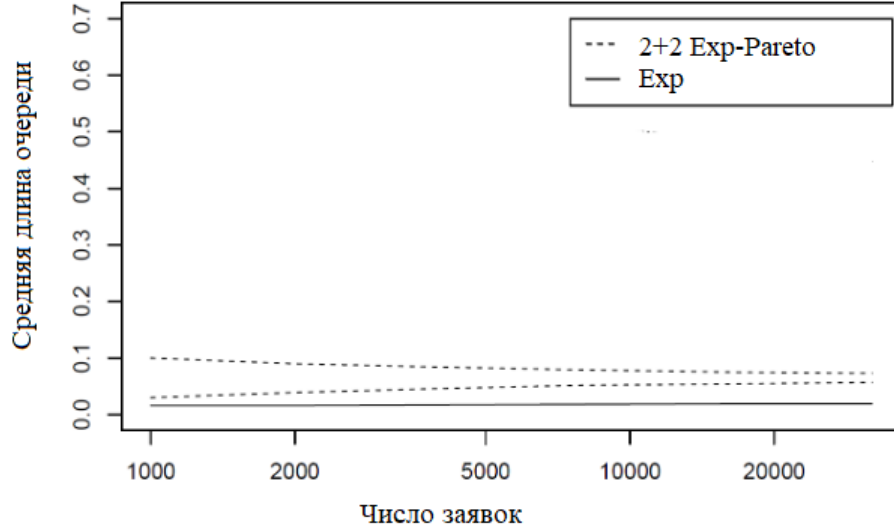


Рисунок 2.7. Доверительный интервал для средней стационарной длины очереди в системах Σ , $\Sigma^{(1)}$, в которых размеры буферов $C = C_1 = 1$, число серверов $N = N_1 = 4$ и $p_1 = 0.1, p_1 = 0.1, p_2 = 0.15, p_3 = 0.15, p_4 = 0.6$.

На Рисунке 2.7 представлены результаты моделирования: построены границы доверительного интервала для средней длины очереди в зависимости от числа заявок в системе Σ с размером буфера $C = 1$, в которой время обслуживания имеет 2+2-компонентное показательное-Парето распределение с параметрами $p_1 = 0.10, p_2 = 0.15, p_3 = 0.15, p_4 = 0.60$, остальные параметры заданы соотношением (2.29), так, что условие (2.41) выполнено. Система $\Sigma^{(1)}$ с показательным обслуживанием с параметром λ_1 является минорантной для системы Σ . На рисунке 2.7 сплошной линией изображена нижняя граница доверительного интервала для средней длины очереди системы $\Sigma^{(1)}$.

Пример 2.6. Пусть в исходной системе Σ время обслуживания имеет показательное-Парето распределение с параметрами $\lambda_s, \alpha_s, x_0, p$. Построим две новые системы: минорантную $\Sigma^{(1)}$, в которой время обслуживания $S^{(1)}$

показательное с параметром λ_s , и мажорантную $\Sigma^{(2)}$, в которой время обслуживания $S^{(2)}$ имеет распределение $Pareto(\alpha_s, x_0)$. В разделе 1.8. было показано, что функция интенсивности отказов для показательного-Парето распределения имеет вид

$$r_S(x) = \frac{p r_{S^{(1)}}(x) a(x) + (1-p) r_{S^{(2)}}(x)}{p a(x) + (1-p)}, \quad x \geq 0,$$

где

$$a(x) = e^{-\lambda_s x} \left(1 + \frac{x}{x_0}\right)^{\alpha_s}. \quad (2.42)$$

Заметим, что

$$\sup_{x \geq 0} r_S(x) = r_S(0) = p\lambda_s + (1-p)\frac{\alpha_s}{x_0}.$$

Отсюда следует, что если $\lambda_s \geq \alpha_s/x_0$, то также $\lambda_s \geq p\lambda_s + (1-p)\alpha_s/x_0$. Следовательно,

$$\inf_{x \geq 0} r_{S^{(1)}}(x) = \lambda_s \geq \sup_{x \geq 0} r_S(x) = r_S(0),$$

и условие (2.37) также выполнено. Наконец, справедливость (2.39) следует из соотношений

$$\begin{aligned} \inf_{x \geq 0} r_S(x) &= \lambda_s, \\ \sup_{x \geq 0} r_{S^{(2)}}(x) &= r_{S^{(2)}}(0) = \frac{\alpha_s}{x_0}. \end{aligned} \quad (2.43)$$

Из Теоремы 2.5 вытекают соотношения (2.38)–(2.40). Следовательно, в трех рассмотренных системах стационарные длины очередей имеют следующий стохастический порядок

$$Q^{(1)} \underset{st}{\leq} Q \underset{st}{\leq} Q^{(2)}.$$

2.5. Монотонность в системах с повторными вызовами

2.5.1. Минорантная и мажорантная системы

Рассмотрим односерверную систему с повторными вызовами Σ , в которой обслуживаются m классов заявок. Заявки k -го класса поступают в соответствии с пуассоновским потоком с параметром λ_k , $k = 1, \dots, m$. Если заявка застаёт сервер пустым, то она немедленно начинает обслуживаться, в противном случае, заявка уходит на k -ю (виртуальную) орбиту бесконечного объема,

образуя очередь в порядке поступления на орбиту, $k = 1, \dots, m$. Первая в очереди на k -й орбите заявка производит независимые попытки присоединиться к обслуживанию на сервере через некоторое случайное время. Интенсивность обслуживания на орбите не зависит от размера орбиты (но может зависеть от класса орбиты). Такие системы называют системами с постоянной скоростью попыток.

Обозначим через t_n момент прихода n -й заявки в общий пуассоновский входной поток (образованный суперпозицией m входных потоков, и пусть $t_1 = 0$), $S_n(k)$ – время обслуживания n -й заявки k -го класса $k = 1, \dots, m, n \geq 1$. Пусть последовательности н.о.р. интервалов между приходами заявок $\{\tau_n := t_{n+1} - t_n, n \geq 1\}$ и времен обслуживания $\{S_n(k), n \geq 1\}$ независимы. Предположим, что интервалы между приходами заявок с (непустой) k -й орбиты распределены показательно с параметром η_k и не зависят от размера орбиты (числа заявок на k -й орбите). Время обслуживания заявок k -го класса $S(k)$ имеет произвольное распределение $F_{S(k)}$, $k = 1, \dots, m$. Обозначим

$$\lambda = \sum_{k=1}^m \lambda_k, \quad \rho_k = \lambda_k M S(k), \quad \rho = \sum_{k=1}^m \rho_k. \quad (2.44)$$

Пусть $W(t)$ есть незавершенная работа в момент времени t^- , и система в начальный момент времени пуста, $W(0) = 0$. Обозначим $W_n = W(t_n)$, $n \geq 1$. Известно [125], что неравенство

$$\rho + \max_{k=1, \dots, m} \frac{\lambda}{\lambda + \eta_k} < 1 \quad (2.45)$$

является достаточным условием стационарности системы. При таком условии существует предел

$$W_n \Rightarrow W \quad \text{при} \quad n \rightarrow \infty.$$

Явный вид ф.р. F_W стационарной незавершенной работы W в общем случае неизвестен. С другой стороны, W является важной характеристикой качества обслуживания системы. Построим верхнюю и нижнюю границы F_W , используя классические системы $M/G/1$ с неограниченной очередью, в которых время обслуживания является конечной смесью.

Рассмотрим две новые системы: $\Sigma^{(1)}$ и $\Sigma^{(2)}$. Входной поток во все три си-

стемы – пуассоновский с параметром λ (это суперпозиция входных потоков, образованных заявками разных классов).

Пусть в системе $\Sigma^{(1)}$ время обслуживания $S^{(1)} = S$ задано смесью распределений вида

$$S = \sum_{k=1}^m I(k)S(k), \quad n \geq 1, \quad (2.46)$$

где индикаторы $I(k)$ таковы, что $\sum I(k) = 1$, среднее $MI(k) = p_k = \lambda_k/\lambda$, а $S(k)$ – время обслуживания заявки k -го класса. Заметим, что система $\Sigma^{(1)}$ является классической односерверной системой вида $M/G/1$ (с дисциплиной обслуживания FIFO), в которой время обслуживания (2.46) является смесью времен обслуживания заявок всех классов исходной системы.

Пусть $\Sigma^{(2)}$ – это система вида $M/G/1$, в которой заявка обслуживается в течение времени S , заданного соотношением (2.46), плюс время ξ_0 , имеющее показательное распределение с параметром

$$\mu_0 = \min_{1 \leq k \leq m} (\lambda + \eta_k),$$

т. е.

$$S^{(2)} = S + \xi_0. \quad (2.47)$$

Таким образом, с.в. ξ_0 соответствует самой «медленной» орбите (с «наибольшими» в среднем интервалами между попытками). Заметим, что система $\Sigma^{(2)}$ имеет коэффициент загрузки

$$\rho^{(2)} = \lambda MS + \frac{\lambda}{\mu_0} = \rho + \frac{\lambda}{\mu_0}, \quad (2.48)$$

и условие стационарности (2.45) для нее принимает вид

$$\rho^{(2)} < 1.$$

В следующей Теореме 2.6 даны верхняя и нижняя границы стационарной незавершенной работы W в исходной системе с повторными вызовами Σ .

Теорема 2.6. Пусть в системах $\Sigma^{(1)}, \Sigma^{(2)}, \Sigma$ в начальный момент времени

$$W_1^{(1)} = W_1 = W_1^{(2)} = 0. \quad (2.49)$$

Тогда при условии (2.45) стационарные времена незавершенной работы имеют следующий стохастический порядок

$$W^{(1)} \underset{st}{\leq} W \underset{st}{\leq} W^{(2)}. \quad (2.50)$$

Доказательство. Как показано в работе [125], условие (2.45) является условием стационарности для системы $\Sigma^{(2)}$, из которого вытекает стационарность исходной системы. Заметим также, что условие $\rho < 1$ – это хорошо известное условие стационарности системы $\Sigma^{(1)}$. Так что эта система стационарна, если выполнено (2.45). Более того, в работе [125, Теорема 3] доказано, что незавершенная нагрузка $W^{(2)}$ стохастически больше незавершенной нагрузки W , что доказывает верхнюю границу неравенства (2.50).

Доказательство нижней границы в неравенстве (2.50) базируется на рекурсии Линдли для процесса незавершенной нагрузки в исходной системе Σ в момент времени t_n :

$$W_{n+1} = \max(0, W_n + S_n - Z_n), \quad n \geq 1, \quad (2.51)$$

где S_n – время обслуживания n -й заявки и Z_n – время занятости сервера в интервале $[t_n, t_{n+1}]$, которое равно проделанной системой работе за этот интервал времени. Теперь применим каплинг для моментов $\{t_n^{(1)}\}$ и н.о.р. времен обслуживания $\{S_n^{(1)}\}$ в системе $\Sigma^{(1)}$ так, что интервалы между приходами заявок и времена обслуживания в системах Σ и $\Sigma^{(1)}$ совпадают с в. 1, т. е., $\tau_n = \tau_n^{(1)}$, $S_n = S_n^{(1)}$, $n \geq 1$. Незавершенная нагрузка в системе $\Sigma^{(1)}$ удовлетворяет классической рекурсии Линдли [34]:

$$W_{n+1}^{(1)} = \max(0, W_n^{(1)} + S_n - \tau_n), \quad n \geq 1. \quad (2.52)$$

Очевидно, что $Z_n \leq \tau_n$ с в. 1 для любого $n \geq 1$. Следовательно,

$$W_2^{(1)} = \max(0, W_1^{(1)} + S_1 - \tau_1) \leq \max(0, W_1 + S_1 - Z_1) = W_2$$

с вероятностью 1, и методом математической индукции по n легко показать, что

$$W_n^{(1)} \leq W_n, \quad n \geq 1. \quad (2.53)$$

Поскольку выполнено условие стационарности (2.45), то соотношение (2.53) выполнено для стационарных версий незавершенной нагрузки, т. е. для W и $W^{(1)}$. $\square$

2.5.2. Распределение стационарного времени ожидания в минорантной и мажорантной системах

В этом разделе рассмотрим стационарное время ожидания во введенных выше системах $\Sigma^{(1)}$ и $\Sigma^{(2)}$ с пуассоновским входным потоком с параметром λ , используя формулу Поллачека-Хинчина.

Рассмотрим вначале систему $\Sigma^{(1)}$ с н.о.р. временами обслуживания $\{S_n, n \geq 1\}$ вида (2.46) с плотностью

$$f_S(x) = \sum_{k=1}^m p_k f_{S(k)}(x), \quad x \geq 0, \quad (2.54)$$

хвостом ф.р.

$$\bar{F}_S(x) = \sum_{k=1}^m p_k \bar{F}_{S(k)}(x), \quad x \geq 0 \quad (2.55)$$

и математическим ожиданием

$$MS = \sum_{k=1}^m p_k MS(k).$$

Обозначая через $\mu = 1/MS$ интенсивность обслуживания, а через $\mu_k = 1/MS(k)$ – интенсивность k -й компоненты, из (2.55) получим соотношение для интенсивности, выраженное через величину

$$\mu = \left[\sum_{k=1}^m \frac{p_k}{\mu_k} \right]^{-1}. \quad (2.56)$$

Введем преобразование Лапласа-Стилтьеса (ПЛС)

$$\psi_{S_e}(z) = \int_0^\infty e^{-zt} dF_{S_e}(t), \quad (2.57)$$

где F_{S_e} есть равновесное распределение с.в. S с плотностью

$$f_{S_e}(x) = \mu \bar{F}_S(x), \quad x \geq 0. \quad (2.58)$$

Распределение F_{S_e} соответствует стационарному перескоку процесса восстановления, образованного последовательностью н.о.р. времен обслуживания $\{S_n\}$ с ф.р. (2.55), см., например, [34].

Выведем ПЛС $\psi_{S_e}(z)$ через ПЛС $\psi_{S_e(k)}(z)$ равновесных распределений компонент смеси. Из соотношений (2.55) – (2.58) следует, что

$$\psi_{S_e}(z) = \mu \int_0^\infty e^{-zt} \bar{F}_S(t) dt = \mu \sum_{k=1}^m p_k \int_0^\infty e^{-zt} \bar{F}_{S(k)}(t) dt = \sum_{k=1}^m \frac{\rho_k}{\rho} \psi_{S_e(k)}(z). \quad (2.59)$$

Заметим, что ПЛС (2.59) есть линейная комбинация ПЛС компонент смеси с коэффициентами, вообще говоря, не равными коэффициентам пропорции смеси $p_1, \dots, p_m$.

Подставив (2.59) в известное для систем вида $M/G/1$ уравнение Поллачека-Хинчина [13] для ПЛС стационарной нагрузки $W^{(1)}$

$$\psi_{W^{(1)}}(z) = \frac{1 - \rho}{z(1 - \rho \psi_{S_e}(z))}, \quad (2.60)$$

получим ПЛС в терминах компонент смеси,

$$\psi_{W^{(1)}}(z) = \frac{1 - \rho}{z \left(1 - \rho \sum_{k=1}^m \frac{\rho_k}{\rho} \psi_{S_e(k)}(z) \right)} = \frac{1 - \rho}{z \left(1 - \sum_{k=1}^m \rho_k \psi_{S_e(k)}(z) \right)}. \quad (2.61)$$

Сформулируем полученный результат в виде следующей леммы.

Лемма 2.1. *В системе $\Sigma^{(1)}$ с распределением времени обслуживания вида (2.46) ПЛС стационарной незавершенной работы $W^{(1)}$ удовлетворяет выражению (2.61).*

Для получения подобного результата в системе $\Sigma^{(2)}$, необходимо найти ПЛС с.в. $S^{(2)}$, определяемой соотношением (2.47), и соответствующее ПЛС для равновесного распределения с.в. $S_e^{(2)}$. Из (2.47) и свойств ПЛС (см., например, [151]), следует, что

$$\psi_{S^{(2)}}(z) = \psi_S(z) \frac{\mu_0}{\mu_0 + z}. \quad (2.62)$$

Вычислим теперь ПЛС $S_e^{(2)}$, используя свойства ПЛС [108]:

$$\begin{aligned}\psi_{S_e^{(2)}}(z) &= \frac{1 - \psi_{S^{(2)}}(z)}{z\mathbf{M}(S + \xi_0)} = \frac{1 - \psi_S(z)\frac{\mu_0}{\mu_0+z}}{z\mathbf{M}(S + \xi_0)} = \\ &= \frac{\mu_0}{\mu_0 + z} \left(\frac{1 - \psi_S(z)}{z\mathbf{M}(S + \xi_0)} + \frac{\mathbf{M}\xi_0}{\mathbf{M}(S + \xi_0)} \right).\end{aligned}\quad (2.63)$$

Из соотношения (2.63), с учетом (2.59), следует

$$\psi_{S_e^{(2)}}(z) = \frac{\mathbf{M}S}{\mathbf{M}(S + \xi_0)} \sum_{k=1}^m \frac{\rho_k}{\rho} \psi_{S_e(k)}(z) \frac{\mu_0}{\mu_0 + z} + \frac{\mathbf{M}\xi_0}{\mathbf{M}(S + \xi_0)} \frac{\mu_0}{\mu_0 + z}.\quad (2.64)$$

Из (2.64) следует, что $S_e^{(2)}$ на самом деле есть смесь экспоненциальной компоненты ξ_0 и сумм вида $S(k) + \xi_0$, $k = 1, \dots, m$. Этот результат может быть интерпретирован в терминах процесса восстановления, где $S + \xi_0$ интервал процесса восстановления, а $S_e^{(2)}$ – стационарное остаточное (равновесное) время восстановления. Подставив (2.64) в соотношение (2.60) (с заменой S_e на $S_e^{(2)}$ и ρ на $\rho^{(2)}$), и обозначив

$$p_0 = \frac{\mathbf{M}S}{\mathbf{M}S + \mathbf{M}\xi_0},\quad (2.65)$$

получим следующее выражение:

$$\psi_{W^{(2)}}(z) = \frac{1 - \rho^{(2)}}{z \left(1 - \rho^{(2)} \left(p_0 \sum_{k=1}^M \frac{\rho_k}{\rho} \psi_{S_e(k)}(z) + 1 - p_0 \right) \frac{\mu_0}{\mu_0+z} \right)}.$$

Полученный результат сформулируем в виде Леммы 2.2, используя (2.48) и (2.65).

Лемма 2.2. *В системе $\Sigma^{(2)}$ с распределением времени обслуживания вида (2.47), ПЛС стационарной незавершенной работы $W^{(2)}$ удовлетворяет выражению*

$$\psi_{W^{(2)}}(z) = \frac{1 - \rho - \frac{\lambda}{\mu_0}}{z \left(1 - \frac{\mu_0}{\mu_0+z} \left(\sum_{k=1}^m \rho_k \psi_{S_e(k)}(z) + \frac{\lambda}{\mu_0} \right) \right)}.\quad (2.66)$$

2.5.3. Система с повторными вызовами и распределением Парето времени обслуживания

Пусть время обслуживания k -го класса заявок в исходной системе с повторными вызовами Σ задано распределением $Pareto(\alpha_k, \beta_k)$

$$F_{S(k)}(x) = 1 - \left(\frac{\beta_k}{\beta_k + x} \right)^{\alpha_k}, \quad x \geq 0, \alpha_k > 0, \beta_k > 0, \quad k = 1, \dots, m. \quad (2.67)$$

Предположим, что α_k не являются целыми числами при $k = 1, \dots, m$. Тогда ПЛС для распределения (2.67) имеет вид

$$\psi_{S(k)}(z) = 1 - (\beta_k z)^{\alpha_k} e^{\beta_k z} \Gamma(1 - \alpha_k, \beta_k z), \quad (2.68)$$

$$\psi_{S_e(k)}(z) = (\alpha_k - 1)(\beta_k z)^{\alpha_k - 1} e^{\beta_k z} \Gamma(1 - \alpha_k, \beta_k z), \quad (2.69)$$

где $\Gamma(a, z) = \int_z^\infty t^{a-1} e^{-t} dt$ – верхняя неполная Гамма-функция.

Для изучения стационарной незавершенной работы в исходной системе Σ , выведем распределение незавершенной работы в явном виде для минорантной системы $\Sigma^{(1)}$, и найдем ПЛС для этого процесса в мажорантной системе $\Sigma^{(2)}$, которые описаны ниже.

Минорантная система $\Sigma^{(1)}$ построена таким образом, что времена обслуживания в ней заданы смесью (2.55), и входной пуассоновский поток совпадает с входным потоком исходной системы Σ .

Следующая теорема является обобщением для смеси распределений результатов, полученных в работе [141], в которой также использованы формулы из работ [170] и [134] для явного представления ПЛС времен обслуживания, заданных распределением Парето.

Теорема 2.7. *Стационарная незавершенная работа $W^{(1)}$ в системе $M/G/1$, в которой время обслуживания задано смесью (2.55), где с.в. $S(k)$ имеют распределение $Pareto(\alpha_k, \beta_k)$, $k = 1, \dots, M$, имеет ф.р.*

$$F_{W^{(1)}}(x) = 1 - \rho(1 - \rho) \int_0^\infty \frac{e^{-xu}}{u} \frac{\tilde{I}(u)}{(1 - \rho \tilde{R}(u))^2 + (\pi \rho \tilde{I}(u))^2} du, \quad x \geq 0, \quad (2.70)$$

где ρ определяется формулой (2.44),

$$\tilde{R}(u) = \sum_{k=1}^m \frac{\rho_k}{\rho} R(\beta_k u, \alpha_k - 1) \quad \text{и} \quad \tilde{I}(u) = \sum_{k=1}^m \frac{\rho_k}{\rho} I(\beta_k u, \alpha_k - 1),$$

коэффициенты ρ_k определяются формулой (2.44), $k = 1, \dots, m$, и

$$I(x, a) = \frac{x^a e^{-x}}{\Gamma(a)}, \quad R(x, a) = {}_1F_1(1, 1 - a, -x) - \pi I(x, a) \cot(\pi a), \quad (2.71)$$

где ${}_1F_1(a, b, z)$ – конфлюэнтная гипергеометрическая функция первого рода (функция Куммера $M(a, b, z)$ [25])

$${}_1F_1(a, b, z) = \sum_{i=1}^{\infty} \frac{(a)_i}{(b)_i i!} z^i, \quad (2.72)$$

и $(a)_i = \prod_{j=0}^{i-1} (a + j)$ – возрастающий факториал, $(a)_0 = 1$ [25].

Доказательство. В работе [141] получено распределение стационарной незавершенной работы $F_{\tilde{W}}$ в системе $M/G/1$ с пуассоновским входным потоком с параметром λ и временем обслуживания с распределением $Pareto(\alpha, \beta)$ (для нецелого $\alpha > 1$) с коэффициентом загрузки

$$\tilde{\rho} = \lambda\beta/(\alpha - 1). \quad (2.73)$$

Это распределение имеет следующий вид:

$$F_{\tilde{W}}(x) = 1 - \int_0^{\infty} \frac{\tilde{\rho}(1 - \tilde{\rho})u^{\alpha-2}}{\Gamma(\alpha - 1)H_1(u, \alpha, \tilde{\rho})} e^{-(1+x/\beta)u} du, \quad (2.74)$$

где функция H_1 определена формулой

$$H_1(x, \alpha, \rho) = [1 - \rho R(x, \alpha - 1)]^2 + [\pi \rho I(x, \alpha - 1)]^2,$$

функция $I(x, a)$ определена в (2.71), а функция $R(x, a)$ определена формулой

$$R(x, a) = 1 + \sum_{r=1}^{\infty} \frac{x^r}{(a - 1) \dots (a - r)} - I(x, a) (\pi a). \quad (2.75)$$

Заметим, что определение функции $R(x, a)$ в формуле (2.75) совпадает с ее представлением в (2.71) с учетом (2.72).

Для обращения ПЛС будем использовать интеграл Бромвича [141]

$$F_{W^{(1)}}(x) = \frac{1}{2\pi i} \int_{s-i\infty}^{s+i\infty} e^{xz} \psi_{W^{(1)}}(z) dz, \quad (2.76)$$

где $s > 0$ достаточно большое. Тогда (2.76) принимает вид

$$F_{W^{(1)}}(x) = 1 - \frac{1}{2\pi i} \int_0^\infty e^{-xu} (\psi_{W^{(1)}}(ue^{i\pi}) - \psi_{W^{(1)}}(ue^{-i\pi})) du. \quad (2.77)$$

Теперь применим распределение (2.77) к смеси (2.55). Ключевой аргумент в [141] для получения (2.77) основан на том, что уравнение $1 - \tilde{\rho}\psi_{\tilde{S}_e}(z) = 0$ не имеет нулей на комплексной плоскости; здесь с.в. $\tilde{S}_e$ имеет распределение $Pareto(\alpha, \beta)$ с параметром $\alpha > 1$, а $\tilde{\rho} \in [0, 1)$ – константа, определенная в (2.73). Повторим этот аргумент для действительной аналитической функции $1 - \tilde{\rho}\psi_{S_e}(z)$, где ПЛС $\psi_{S_e}(z)$ определено формулой (2.59). Пусть $z_1 = x_1 + iy_1$ – корень уравнения $1 - \tilde{\rho}\psi_{S_e}(z)$, тогда $\bar{z}_1 = x_1 - iy_1$ также корень, при котором верно тождество

$$\tilde{\rho}(\psi_{S_e}(z_1) - \psi_{S_e}(\bar{z}_1)) = 0. \quad (2.78)$$

Используя (2.59), перепишем уравнение (2.78) в виде

$$\tilde{\rho} \sum_{k=1}^m \frac{\rho_k}{\rho} (\psi_{S_e(k)}(z_1) - \psi_{S_e(k)}(\bar{z}_1)) = 0. \quad (2.79)$$

В работе [141] показано, что для всех $k = 1, \dots, m$, преобразование (2.69) на множестве $\mathbb{D} = \mathbb{C} \setminus (-\infty, 0]$ (комплексная плоскость на положительной полуоси) эквивалентно преобразованию

$$\psi_{S_e(k)}(z) = \frac{\alpha_k - 1}{\Gamma(\alpha_k)} \int_0^\infty \frac{t^{\alpha_k-1} e^{-t}}{(t + \beta_k z)^2} dt, \quad z \in \mathbb{D}. \quad (2.80)$$

Используя (2.80), уравнение (2.79) перепишем в виде

$$2iy_1 \frac{\tilde{\rho}}{\rho} \sum_{k=1}^m \frac{\lambda_k \beta_k^2}{\Gamma(\alpha_k)} \int_0^\infty \frac{t^{\alpha_k-1} e^{-t}}{(t + \beta_k x_1)^2 + (\beta_k y_1)^2} dt = 0. \quad (2.81)$$

Если $y_1 = 0$, то z_1 является действительным корнем, однако, т.к. для любого $k = 1, \dots, m$ выполнено неравенство $\psi_{S_e(k)}(x) \leq 1$ [141], то из (2.59) следует $\psi_{S_e}(x) \leq 1$ и, следовательно, $1 - \tilde{\rho}\psi_{S_e}(z_1) > 0$ для $\tilde{\rho} \in [0, 1)$.

В противном случае, т. е. если $y_1 \neq 0$ и $z_i \in \mathbb{D}$, то левая часть (2.81) как линейная комбинация вещественных неотрицательных интегралов не равна нулю. Наконец, на отрицательной полуоси $z_1 = xe^{\pm i\pi}$ (для $x > 0$), как следует из работы [141], для каждого $k = 1, \dots, m$ выполнено соотношение

$$\psi_{S_e(k)}(xe^{\pm i\pi}) = R(\beta_k x, \alpha_k - 1) \mp i\pi I(\beta_k x, \alpha_k - 1), \quad (2.82)$$

где функции I и R определены соотношением (2.71). Таким образом, для $x > 0$, из соотношений (2.59) и (2.71) следует, что $\psi_{S_e(k)}(xe^{\pm i\pi})$ имеет ненулевую часть

$$\mp i\pi \sum_{k=1}^m \frac{\rho_k}{\rho} \frac{(\beta_k x)^{\alpha_k - 1} e^{-\beta_k x}}{\Gamma(\alpha_k - 1)}.$$

Из соотношения (2.77), подставляя (2.61) и (2.82), получаем (2.70). $\square$

Таким образом, выражение для $W^{(1)}$ получено в явном виде. Для вывода распределения $W^{(2)}$ необходимо численно обратить ПЛС, используя (2.69) для $W^{(2)}$ в форме (2.66), полученной в Лемме 2.2.

Пример 2.7. Для иллюстрации результатов, полученных в Теоремах 2.6 и 2.7 проведен численный эксперимент. Рассмотрим систему с двумя классами заявок с параметрами:

$$\alpha_1 = 3.3, \beta_1 = 10, \alpha_2 = 4.3, \beta_2 = 2, p_1 = 0.6, \eta_1 = 1, \eta_2 = 2, \rho = 0.7. \quad (2.83)$$

В системе $\Sigma^{(1)}$ выберем коэффициент загрузки ρ так, чтобы $\lambda \approx 0.246$. В системе $\Sigma^{(2)}$ вычислим коэффициент $\mu_0 \approx 1.246$. Следовательно, коэффициент загрузки системы $\Sigma^{(2)}$, $\rho_2 \approx 0.897$, и все три системы стационарны. Используем (2.74) для получения распределения незавершенной работы в системе $\Sigma^{(1)}$. Для получения оценки стационарной незавершенной работы в системе $\Sigma^{(2)}$ применена рекурсия (2.51) и использовано 10^6 событий. Результаты эксперимента показаны на Рисунке 2.8 и демонстрируют стохастический порядок $W^{(1)} \underset{st}{\leq} W \underset{st}{\leq} W^{(2)}$.

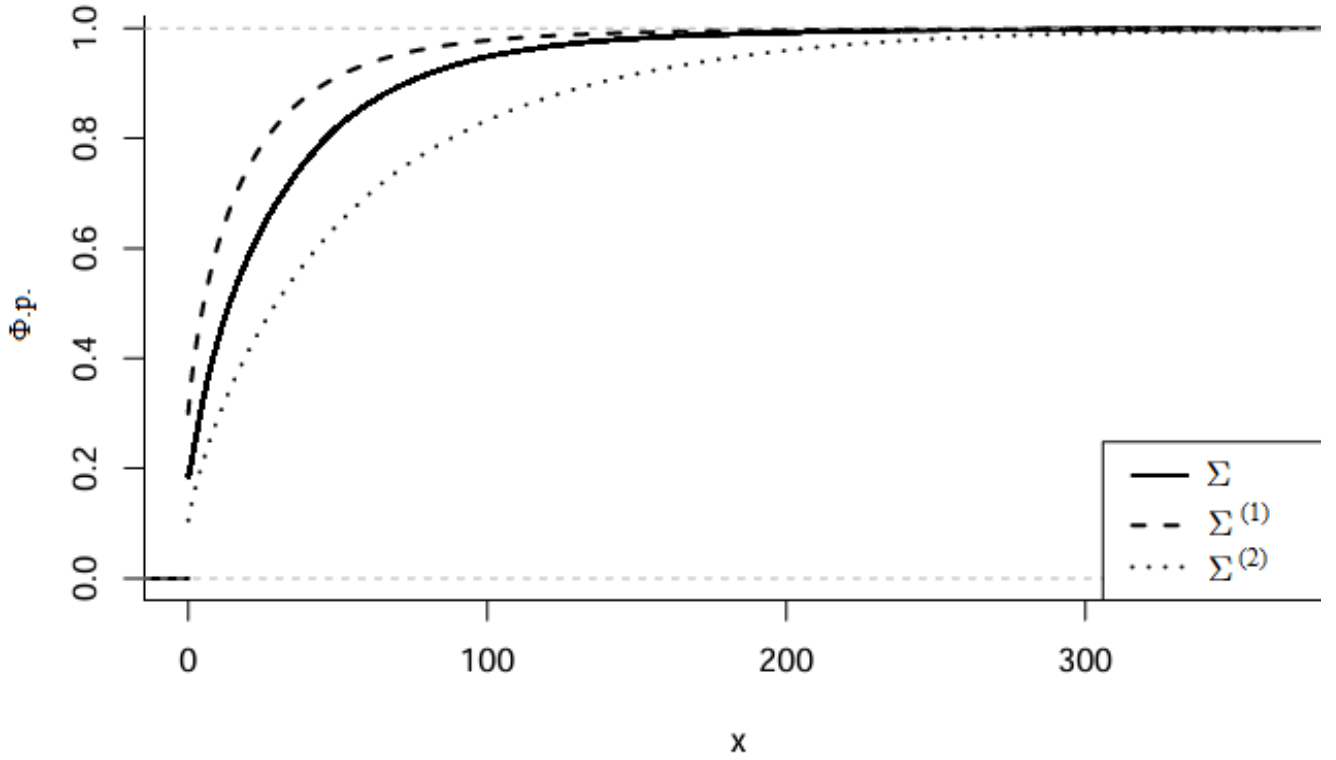


Рисунок 2.8. Теоретическая ф.р. (2.74) для системы $\Sigma^{(1)}$ и эмпирические ф.р. для систем Σ и $\Sigma^{(2)}$ стационарной незавершенной работы в системах вида $M/Pareto/1$ с параметрами (2.83).

2.5.4. Границы незавершенной нагрузки W в системе с показательными временами обслуживания

Пусть в исходной системе с повторными вызовами Σ время обслуживания k -го класса $S(k)$ имеет показательное распределение с ф.р.

$$F_{S(k)}(x) = 1 - e^{-\mu_k x}, \quad x \geq 0, \quad \mu_k > 0, \quad k = 1, \dots, m.$$

В качестве минорантной системы $\Sigma^{(1)}$ рассмотрим систему с неограниченной очередью $M/H_m/1$, в которой время обслуживания $S^{(1)} = S$ имеет гиперэкспоненциальное распределение (смесь m -показательно распределенных с.в. $S(k)$)

$$F_{S^{(1)}}(x) = 1 - \sum_{k=1}^m p_k e^{-\mu_k x}, \quad \mu_k > 0, \quad p_k \geq 0, \quad k = 1, \dots, m, \quad \sum_{k=1}^m p_k = 1. \quad (2.84)$$

Обозначим коэффициент загрузки $\rho^{(1)} = \sum_{k=1}^m \lambda p_k / \mu_k$ в системе $\Sigma^{(1)}$. Поскольку

$$\rho^{(1)} \leq \rho + \frac{\lambda}{\mu_0},$$

то условие (2.45) (предполагается выполненным) влечет стационарность системы $\Sigma^{(1)}$. Рассмотрим ПЛС

$$\psi_{S_e}(z) = \int_0^\infty e^{-zt} dF_{S_e}(t), \quad (2.85)$$

где F_{S_e} есть ф.р. с плотностью

$$f_{S_e}(x) = \mu \bar{F}_S(x), \quad x \geq 0. \quad (2.86)$$

В Лемме 2.1 доказано, что ПЛС стационарной незавершенной работы $W^{(1)}$ выражается через ПЛС компонент смеси времен обслуживания. Подставляя ПЛС для показательного распределения $\psi_{S_e(k)}(z) = \frac{\mu_k}{\mu_k + z}$ в формулу (2.61), получим

$$\psi_{W^{(1)}}(z) = \frac{1 - \sum_{k=1}^m \lambda_k / \mu_k}{z \left(1 - \sum_{k=1}^m \frac{\lambda_k}{\mu_k + z} \right)}.$$

Применяя формулу Поллачика-Хинчина, получим среднюю стационарную незавершенную работу в системе $\Sigma^{(1)}$ в виде

$$MW^{(1)} = \frac{\lambda M(S^{(1)})^2}{2(1 - \rho^{(1)})} = \frac{\sum_{k=1}^m \rho_k^2 + \rho^2}{2\lambda(1 - \rho)}. \quad (2.87)$$

Рассмотрим теперь введенную ранее мажорантную систему $\Sigma^{(2)}$, время обслуживания в которой равно сумме с.в. S с гиперэкспоненциальным распределением (2.84) и с.в. ξ_0 , т. е. $S^{(2)} = S + \xi_0$. Обозначим $\mu_0 = \min_{1 \leq k \leq m} (\lambda + \eta_k)$ и пусть с.в. ξ_0 имеет показательное распределение с параметром μ_0 . Условие стационарности такой системы совпадает с (2.45).

По Лемме 2.2 в системе $\Sigma^{(2)}$ ПЛС стационарной незавершенной работы $W^{(2)}$

имеет следующий вид:

$$\psi_{W^{(2)}}(z) = \frac{1 - \rho - \frac{\lambda}{\mu_0}}{z \left(1 - \frac{\mu_0}{\mu_0 + z} \left(\sum_{k=1}^m \rho_k \psi_{S_e(k)}(z) + \frac{\lambda}{\mu_0} \right) \right)}. \quad (2.88)$$

Подставив $\psi_{S_e(k)}(z) = \frac{\mu_k}{\mu_k + z}$ в (2.88), получим

$$\psi_{W^{(2)}}(z) = \frac{1 - \sum_{k=1}^m \frac{\lambda_k}{\mu_k} - \frac{\lambda}{\mu_0}}{z \left(1 - \frac{\mu_0}{\mu_0 + z} \left(\sum_{k=1}^m \frac{\lambda_k}{\mu_k + z} + \frac{\lambda}{\mu_0} \right) \right)}.$$

Аналогично, можно получить выражение для средней стационарной незавершенной работы

$$\mathbf{MW}^{(2)} = \frac{\lambda \mathbf{M}(S^{(2)})^2}{2(1 - \rho^{(2)})} = \frac{\mu_0^2(\rho^2 + \sum_{k=1}^m \rho_k^2) + 2\lambda(\lambda + \rho\mu_0)}{2\lambda\mu_0(\mu_0 - \rho\mu_0 - \lambda)}. \quad (2.89)$$

Из Теоремы 2.6 (раздел 2.5.1.) следует, что стационарные времена незавершенной работы стохастически упорядочены

$$W^{(1)} \leq_{st} W \leq_{st} W^{(2)},$$

и, поэтому

$$\mathbf{MW}^{(1)} \leq \mathbf{MW} \leq \mathbf{MW}^{(2)}.$$

Действительно, легко проверить, что

$$\mathbf{MW}^{(2)} = \frac{\lambda \mathbf{M}(S^{(2)})^2}{2(1 - \rho^{(2)})} = \frac{\mu_0^2 \lambda \mathbf{M}(S^{(1)})^2 + 2\lambda(\lambda + \rho\mu_0)}{\mu_0^2 2(1 - \rho^{(1)}) - 2\lambda^2 \mu_0} \geq \mathbf{MW}^{(1)}.$$

Пример 2.8. Рассмотрим систему с повторными вызовами Σ с тремя классами заявок ($m = 3$), в которую поступает пуассоновский поток с интенсивностью $\lambda = 10$. Пусть время обслуживания k -го класса $S(k)$, $k = 1, 2, 3$ имеет показательное распределение с параметром $\mu_1 = 10, \mu_2 = 30, \mu_3 = 60$, соответственно. Пусть коэффициенты пропорции гиперэкспоненциального распределения (2.84) принимают значения $p_1 = 1/2, p_2 = 1/3, p_3 = 1/6$, соответственно. Будем полагать, что интенсивности показательно распределенных попыток с орбиты при-

нимают значения

$$\eta_1 = 50, \eta_2 = 100, \eta_3 = 150.$$

В этом случае $\mu_0 = 50$, а коэффициенты загрузки равны

$$\rho^{(1)} = \rho = \sum_{k=1}^3 \lambda p_k / \mu_k = \frac{23}{36}, \quad \rho^{(2)} = \rho + \lambda / \mu_0 = \frac{29}{36} < 1$$

и условие стационарности (2.45) выполнено. Формулы (2.87)-(2.89) дают значения математических ожиданий $MW^{(1)} \approx 0.093$ и $MW^{(2)} \approx 0.106$, соответственно.

Рассмотрим минорантную систему $\Sigma^{(1)}$, в которой время обслуживания имеет гиперэкспоненциальное распределение

$$\bar{F}_S(x) = \sum_{k=1}^3 p_k e^{-\mu_k x}.$$

Для таких классических систем вида $M/G/1$ распределение числа клиентов в системе в стационарном режиме, обозначаемое через $N^{(1)}$, получено в работе [142] в виде смеси

$$\pi_n^* := P(N^{(1)} = n) = \sum_{k=1}^3 \beta_k (r_k)^n,$$

где параметры смеси β_k и геометрические компоненты r_k , $k = 1, 2, 3$ вычислены и равны

$$\beta = (0.040, 0.075, 0.245), \quad r = (0.146, 0.268, 0.711). \quad (2.90)$$

Замечание. В работе [142] с помощью результата [92] было получено стационарное распределение времени *пребывания* клиента в системе, $V^{(1)}$, в виде

$$\bar{F}_{V^{(1)}}(x) = \sum_{k=1}^3 \gamma_k e^{-\theta_k x}, \quad (2.91)$$

где коэффициенты γ_k и параметры θ_k , $k = 1, 2, 3$, для исходных параметров

равны

$$\gamma_1 = 0.047, \gamma_2 = 0.103, \gamma_3 = 0.850, \quad \theta_1 = 58.633, \theta_2 = 27.307, \theta_3 = 4.059.$$

При этом в утверждении Теоремы 3 в работе [142] ошибочно указано, что $\sum_{k=1}^3 \gamma_k = 1 - \rho^{(1)}$. Это было бы верно, если бы распределение F_V соответствовало времени *ожидания*. Как видно, на самом деле, $\sum_{k=1}^3 \gamma_k = 1$. Заметим, что π_{n+1} есть стационарная вероятность наблюдать n клиентов в очереди, т. е.

$$\mathbf{P}(Q^{(1)} = n) = \pi_{n+1}^*,$$

где $Q^{(1)}$ есть число клиентов в *очереди* в стационарном режиме. Вычислив производящую функцию вероятностей $\pi(z)$ для $Q^{(1)}$, получим

$$\pi(z) = \sum_{n=0}^{\infty} z^n \pi_{n+1}^* = \sum_{k=1}^3 \frac{\beta_k}{z} \sum_{n=1}^{\infty} (r_k z)^n = \sum_{k=1}^3 \frac{\beta_k}{z} \left(\frac{1}{1 - r_k z} - 1 \right) = \sum_{k=1}^3 \frac{\beta_k r_k}{1 - r_k z}. \quad (2.92)$$

С другой стороны, выполняя замену переменной $s = \lambda - \lambda z$ в нижеследующем соотношении между производящей функцией стационарной очереди $\pi(z)$ и ПЛС для стационарного времени ожидания $\psi_{W^{(1)}}(z)$

$$\pi(z) = \sum_{n=0}^{\infty} \int_0^{\infty} z^n e^{-\lambda x} \frac{(\lambda x)^n}{n!} dW^{(1)}(x) = \int_0^{\infty} e^{-(\lambda - \lambda z)x} dW^{(1)}(x) = \psi_{W^{(1)}}(\lambda - \lambda z),$$

получим

$$\psi_{W^{(1)}}(s) = \sum_{k=1}^3 \frac{\beta_k r_k}{1 - r_k(1 - s/\lambda)} = \sum_{k=1}^3 \frac{\beta_k r_k}{1 - r_k} \frac{\theta_k}{\theta_k + s},$$

где, как и в работе [142],

$$\theta_k = \frac{\lambda(1 - r_k)}{r_k}.$$

Таким образом, $\psi_{W^{(1)}}$ соответствует обобщенной смеси экспоненциальных распределений (сумма коэффициентов в смеси не равна 1, см. (2.94)). Отметим при этом, что, в отличие от [142], коэффициенты смеси имеют вид

$$\hat{\gamma}_k = \frac{\beta_k r_k}{1 - r_k}.$$

Таким образом, правильное выражение для ф.р. (2.91) имеет вид

$$\bar{F}_{W^{(1)}}(x) = \sum_{k=1}^3 \hat{\gamma}_k e^{-\theta_k x}, \quad (2.93)$$

где $\hat{\gamma} = (0.007, 0.027, 0.604)$. Заметим, что

$$\bar{F}_{W^{(1)}}(0) = \sum_{k=1}^3 \hat{\gamma}_k = 1 - \rho^{(1)} \approx 0.638 < 1. \quad (2.94)$$

Используем выражение (2.93) для построения ф.р. в системе $\Sigma^{(1)}$, а для построения оценок в исходной Σ и мажорантной $\Sigma^{(2)}$ системах – имитационное моделирование. Построим графики ф.р. и эмпирических функций распределения для незавершенной работы в трех системах. Как и следовало ожидать, из Рисунка 2.9 следует, что для стационарных времен ожидания имеет место стохастический порядок (2.50).

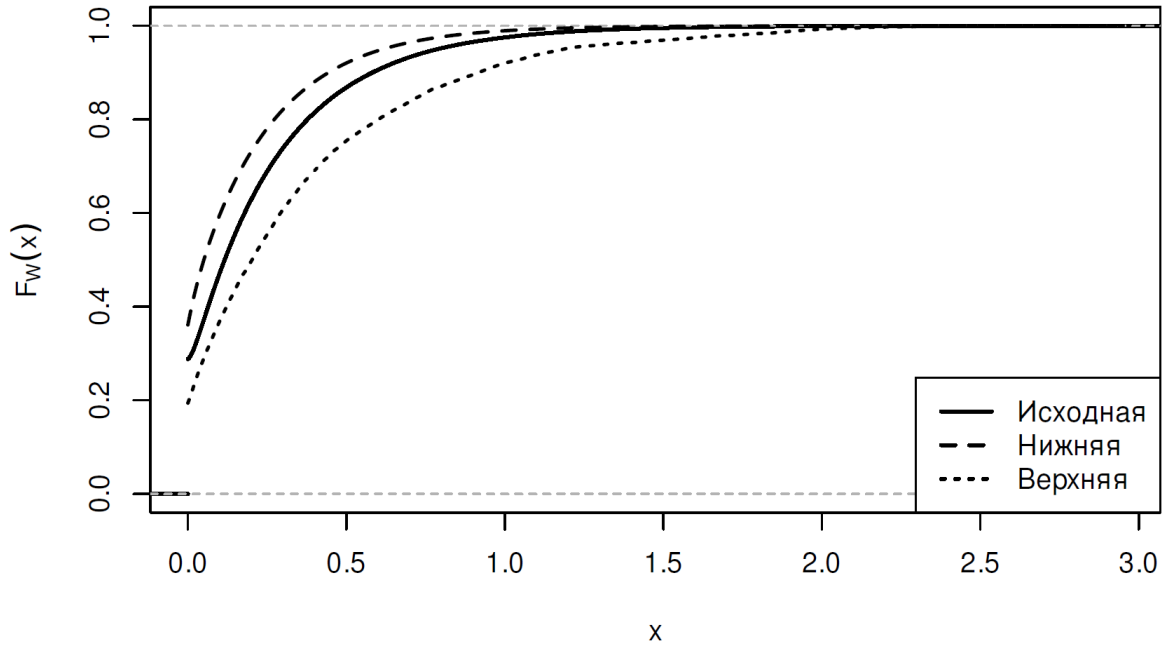


Рисунок 2.9. Ф.р. в минорантной $\Sigma^{(1)}$, исходной Σ и мажорантной $\Sigma^{(2)}$ системах при $\lambda = 10, p_1 = 1/2, p_2 = 1/3, p_3 = 1/6; \mu_1 = 10, \mu_2 = 30, \mu_3 = 60, \eta_1 = 50, \eta_2 = 100, \eta_3 = 150$.

2.6. Монотонность в системах со смешанным входным процессом

Рассмотрим односерверную систему обслуживания, в которую поступает m независимых входных потоков восстановления заявок разных классов. Распределение времени до прибытия первой заявки является стационарным распределением перескока, т. е. удовлетворяет соотношению

$$P(\tau_1^{(i)} \leq x) = \frac{1}{M\tau^{(i)}} \int_0^x (1 - A_i(u)) du, \quad i = 1, \dots, m.$$

Пусть i -й поток задан н.о.р. интервалами между приходами заявок $\{\tau_k^{(i)}, k \geq 2\}$ с ф.р. A_i и средним $M\tau^{(i)} := 1/\lambda_i$. Времена обслуживания заявок i -го входного потока есть н.о.р. с.в. $\{S_k^{(i)}, k \geq 1\}$ с ф.р. B_i , независимые от с.в. $\{\tau_k^{(i)}, i = 1, \dots, m\}$. Таковую систему будем обозначать $\sum_{i=1}^m G_i / \sum_{i=1}^m G_i / 1$. Заметим, что хвост ф.р. времени обслуживания может быть представлен в виде m -компонентной смеси как

$$\bar{B}(x) = \sum_{i=1}^m p_i \bar{B}_i(x), \quad (2.95)$$

где $p_i = \lambda_i / \lambda$, $\lambda = \lambda_1 + \dots + \lambda_m$. Обозначим с.в. с ф.р. B через S . Пусть $\{\tau_k, k \geq 1\}$ – интервалы между приходами заявок в смешанном входном процессе, образованном суперпозицией m входных потоков. Заметим, что в такой системе условие стационарности имеет вид

$$\rho := \lambda \sum_{i=1}^m p_i M S^{(i)} = \sum_{i=1}^m \lambda_i M S^{(i)} < 1. \quad (2.96)$$

В этом случае доказательство стационарности основано на так называемой конструкции Лойнса, см., например, главу 2 в [44]. Для системы $\sum_{i=1}^m G_i / \sum_{i=1}^m G_i / 1$ (обозначим ее Σ) построим минорантную систему $\Sigma^{(1)}$, в которой интервалы между приходами заявок заданы с.в. $\{\tau_k^{(1)}, k \geq 1\}$ из первого входного потока, а обслуживание – с.в. $\{S_k^{(1)}, k \geq 1\}$. В мажорантной системе $\Sigma^{(m)}$ входной поток совпадает с входным потоком исходной системы τ , а время обслуживания задано с.в. $\{S_k^{(m)}, k \geq 1\}$.

Обозначим $W(t)$, $W^{(1)}(t)$, $W^{(m)}(t)$; W_n , $W_n^{(1)}$, $W_n^{(m)}$ – незавершенную ра-

боту в моменты t и t_n^- в системах Σ , $\Sigma^{(1)}$ и $\Sigma^{(m)}$, соответственно. Рассмотрим стационарные пределы

$$W_n^{(i)} \Rightarrow W^{(i)}, \quad W_n \Rightarrow W, \quad W^{(i)}(t) \Rightarrow W_*^{(i)}, \quad W(t) \Rightarrow W_*, \quad n, t \rightarrow \infty, \quad i = 1, m.$$

Эти пределы существуют, если

$$\rho < 1. \quad (2.97)$$

Теорема 2.8. Пусть выполнено условие стационарности (2.97), начальные интервалы между приходами заявок совпадают, $\tau_1 = \tau_1^{(1)}$, и

$$S^{(1)} \leq_r \dots \leq_r S^{(m)}.$$

Тогда

$$W^{(1)} \leq_{st} W \leq_{st} W^{(m)}. \quad (2.98)$$

Доказательство. При условии (2.97) все три системы Σ , $\Sigma^{(1)}$ и $\Sigma^{(m)}$ стационарны. Очевидно, что последовательность моментов $\{t_k^{(1)}, k \geq 1\}$ поступления заявок в систему $\Sigma^{(1)}$ является подпоследовательностью моментов $\{t_k, k \geq 1\}$ поступления заявок в систему Σ , $\{t_k^{(1)}, k \geq 1\} \subseteq \{t_k, k \geq 1\}$, откуда следует, что [177, теорема 2]

$$\tau^{(1)} \geq_{st} \tau.$$

Из Теоремы 1.9 (раздел 1.11.) следует, что

$$S^{(1)} \leq_r S \leq_r S^{(m)},$$

тогда из Теоремы 4 [177] вытекает соотношение (2.98). $\square$

Теорема 2.9. Пусть в системе Σ входной поток задан суперпозицией m входных потоков с моментами прихода в общем потоке $\{t_k, k \geq 1\}$. Времена обслуживания заданы распределением (2.95). В системе $\Sigma^{(1)}$ моменты прихода $\{t_k^{(1)}, k \geq 1\}$ определяются первым входным потоком $\{\tau_k^{(1)}, k \geq 1\}$, а времена обслуживания $\{S_k^{(1)}, k \geq 1\}$ представляют собой первую компоненту смеси (2.95). Предположим, что система Σ стационарна. Пусть $\tau_1 = \tau^{(1)}$, $W_1 = W_1^{(1)}$. Тогда

$$W^{(1)} \leq_{st} W.$$

Доказательство. Прежде всего заметим, что $\{t_k^{(1)}, k \geq 1\} \subseteq \{t_k, k \geq 1\}$. Построим вспомогательную систему $\tilde{\Sigma}$, в которой входной поток такой же, как в системе Σ , а времена обслуживания представляют собой с.в. $\tilde{S}_k = I_k^{(1)} S_k^{(1)}$, где $\sum_{i=1}^m I^{(i)} = 1$, а с.в. $I^{(i)}$ принимает значения 1 с вероятностью $p_i = \lambda_i / \lambda, i = 1, \dots, m$, и 0, иначе. Очевидно, $\tilde{S} \leq_{st} S$. Далее, поскольку

$$\tilde{W}_2 = (\tilde{W}_1 + \tilde{S}_1 - \tau_1)^+ \leq (W_1 + S_1 - \tau_1)^+ = W_2,$$

то индукцией по k можно показать, что

$$\tilde{W}_{k+1} = (\tilde{W}_k + \tilde{S}_k - \tau_k)^+ \leq (W_k + S_k - \tau_k)^+ = W_{k+1}, \quad k \geq 1.$$

Из последовательности $\{t_k, k \geq 1\}$ выделим моменты времени когда приход в систему $\tilde{\Sigma}$ совпадает с приходом в систему $\Sigma^{(1)}$. Пусть $t_k^* = \min\{t_j \geq t_{k-1}^* : I_j^{(1)} = 1, j \geq k-1\}$. Остается заметить, что

$$W^{(1)}(t_k^{(1)}) = W_k^{(1)} = \tilde{W}(t_k^*) \leq W(t_k^*) = W_k, \quad k \geq 1.$$

Следовательно,

$$W_k^{(1)} \leq W_k, \quad k \geq 1.$$

□

Следствие 2.1. Пусть в системе Σ входной поток задан суперпозицией m входных потоков с моментами прихода $\{t_k, k \geq 1\}$. В системе $\tilde{\Sigma}$ моменты прихода $\{\tilde{t}_k, k \geq 1\}$ представляют собой подпоследовательность, $\{\tilde{t}_k, k \geq 1\} \subseteq \{t_k, k \geq 1\}$. Предположим, что времена обслуживания в обеих системах одинаковы и не зависят от класса, $S \stackrel{st}{=} \tilde{S}$. Тогда при условии (2.96) стационарные времена ожидания в этих системах имеют стохастический порядок:

$$\tilde{W} \leq_{st} W.$$

Пример 2.9. Рассмотрим систему обслуживания Σ , в которой входной поток представляет собой суперпозицию двух независимых входных потоков. Пусть

i -й поток задан с.в. $\{\tau_k^{(i)}, k \geq 1\}$ с распределением $Pareto(\alpha_i, x_0)$

$$A_i(x) = 1 - \left(\frac{x_0}{x_0 + x} \right)^{\alpha_i}, x \geq 0, x_0 > 0, \alpha_i > 0.$$

Пусть время обслуживания $S \sim Exp(\mu)$ не зависит от класса обслуживания. Для исходной системы построим минорантную систему $\tilde{\Sigma}$, в которой моменты прихода определяются одним из входных потоков, обозначим его i_0 . Построенная система является системой вида $GI/M/1$ и для нее известен вид распределения стационарного времени ожидания [63]:

$$F_{W_{i_0}}(x) = P(W_{i_0} \leq x) = 1 - \sigma e^{-\mu(1-\sigma)x}, x \geq 0, \quad (2.99)$$

где σ есть корень уравнения

$$z - \psi_{A_{i_0}}(\mu(1-z)) = 0 \quad (2.100)$$

внутри единичного круга и $\psi_{A_{i_0}}(z)$ – ПЛС распределения входного потока A_{i_0} , которое имеет вид

$$\begin{aligned} \psi_{A_{i_0}}(z) &= - \int_0^\infty e^{-zx} d\bar{A}_{i_0}(x) = 1 - z \int_0^\infty e^{-zx} \bar{A}_{i_0}(x) dx \\ &= 1 - z \int_0^\infty e^{-zx} \left(\frac{x_0}{x_0 + x} \right)^{\alpha_{i_0}} dx = 1 - z \int_{x_0 z}^\infty e^{-u+x_0 z} x_0^{\alpha_{i_0}} (u/z)^{-\alpha_{i_0}} du/z \\ &= 1 - e^{x_0 z} (x_0 z)^{\alpha_{i_0}} \Gamma(-\alpha_{i_0} + 1, x_0 z), \end{aligned}$$

где $\Gamma(s, x)$ – верхняя неполная Гамма-функция.

На Рисунке 2.10 приведены графики эмпирической ф.р. стационарного времени ожидания для исходной системы $\Sigma = G_1 + G_2/M/1$ при $\mu = 4, \alpha_1 = 2, \alpha_2 = 3$ (коэффициент загрузки $\rho = 0.75$), а также теоретические ф.р. $F_{W_i}(x)$ для систем $G_i/M/1, i = 1, 2$. В этом случае $\sigma_1 = 0.409, \sigma_2 = 0.64$ – корни уравнения (2.100) для систем $G_1/M/1$ и $G_2/M/1$, соответственно.

На Рисунке 2.11 приведены аналогичные графики для $\mu = 5$ (коэффициент загрузки $\rho = 0.6$). В этом случае $\sigma_1 = 0.39, \sigma_2 = 0.506$ для систем $G_1/M/1$ и

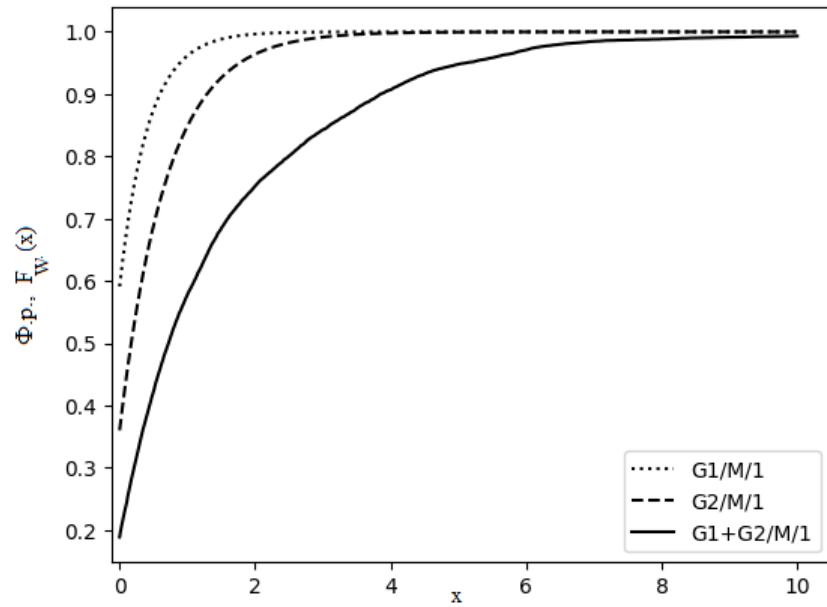


Рисунок 2.10. Эмпирическая ф.р. стационарного времени ожидания в исходной системе $\Sigma = \sum_{i=1}^2 G_i/M/1$, теоретические ф.р. $F_{W_i}, i = 1, 2$ в системах $G_1/M/1$ и $G_2/M/1$ при $\mu = 4, \alpha_1 = 2, \alpha_2 = 3$.

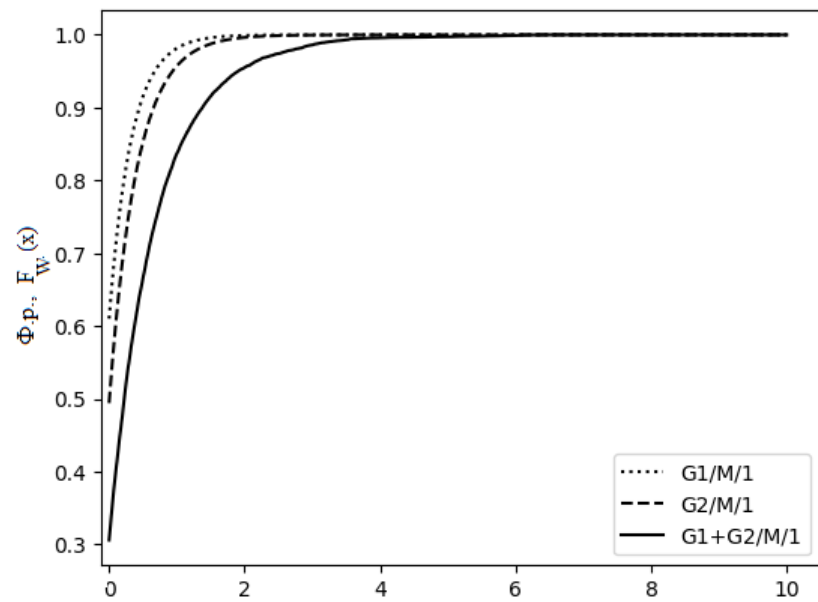


Рисунок 2.11. Эмпирическая ф.р. стационарного времени ожидания в исходной системе $\Sigma = \sum_{i=1}^2 G_i/M/1$, теоретические ф.р. $F_{W_i}, i = 1, 2$ в системах $G_1/M/1$ и $G_2/M/1$ при $\mu = 5, \alpha_1 = 2, \alpha_2 = 3$.

$G_2/M/1$, соответственно.

Пример 2.10. Рассмотрим систему Σ обслуживания, в которой входной поток представляет собой суперпозицию двух независимых входных потоков. Пусть

i -й поток задан с.в. $\{\tau_k^{(i)}, k \geq 1\}$ с распределением $We(\alpha_i, \beta_i)$

$$A_i(x) = 1 - e^{-(x/\alpha_i)^{\beta_i}}, x \geq 0, \beta_i > 0, \alpha_i > 0.$$

Пусть в исходной системе время обслуживания $S \sim Exp(\mu)$ не зависит от класса обслуживания. В минорантной системе $\tilde{\Sigma}$ моменты прихода определяются одним из входных потоков, обозначим его i_0 . Найдем ПЛС распределения Вейбулла для того, чтобы найти корень уравнения (2.100) и ф.р. (2.99).

$$\begin{aligned} \psi_{A_{i_0}}(z) &= - \int_0^\infty e^{-zx} d\bar{A}_{i_0}(x) = 1 - z \int_0^\infty e^{-zx} \bar{A}_{i_0}(x) dx \\ &= 1 - z \int_0^\infty e^{-zx} e^{-(x/\alpha_{i_0})^{\beta_{i_0}}} dx = 1 - z \sum_{k=0}^\infty \int_0^\infty (-1)^k \frac{(zx)^k}{k!} e^{-(x/\alpha_{i_0})^{\beta_{i_0}}} dx = \\ &= 1 - \frac{z\alpha_{i_0}}{\beta_{i_0}} \sum_{k=0}^\infty (-1)^k \frac{(z\alpha_{i_0})^k}{k!} \int_0^\infty u^{(k+1)/\beta_{i_0}} e^{-u} du \\ &= 1 - \frac{z\alpha_{i_0}}{\beta_{i_0}} \sum_{k=0}^\infty \frac{(-1)^k (z\alpha_{i_0})^k}{k!} \Gamma\left(\frac{k+1}{\beta_{i_0}}\right). \end{aligned}$$

2.7. Монотонность модели высокопроизводительного кластера

Рассмотрим многосерверную систему с N работающими параллельно серверами и неограниченным буфером, в которую поступает m независимых входных потоков восстановления. Пусть i -й поток задан н.о.р. интервалами между приходами заявок $\{\tau_k^{(i)}, k \geq 2\}$ с ф.р. A_i и средним $M\tau^{(i)} = 1/\lambda_i$. Распределение времени до прибытия первой заявки является стационарным распределением перескока, т. е. удовлетворяет соотношению

$$P(\tau_1^{(i)} \leq x) = \frac{1}{M\tau^{(i)}} \int_0^x (1 - A_i(u)) du, \quad i = 1, \dots, m.$$

Моменты прихода заявок в смешанный входной процесс $\{t_k, k \geq 1\}$ формируют последовательность н.о.р. интервалов между приходами заявок $\tau_k :=$

$t_{k+1} - t_k, i \geq 1, t_1 := 0$. Времена обслуживания заявок i -го входного потока есть н.о.р с.в. $\{S_k^{(i)}, k \geq 1\}$ с ф.р. B_i , независимые от с.в. $\{\tau_k^{(i)}\}, i = 1, \dots, m$. Время обслуживания может быть представлено в виде m -компонентной смеси распределений B_i , а именно:

$$\bar{B}(x) = \sum_{i=1}^m p_i \bar{B}_i(x), \quad (2.101)$$

где $p_i = \lambda_i / \lambda, \lambda = \lambda_1 + \dots + \lambda_m$. Обозначим через S с.в. с ф.р. B .

В отличие от модели, представленной в разделе 2.6., каждой заявке с номером k требуется для обслуживания C_k серверов одновременно. С.в. $\{C_k\}$ н.о.р. с общим элементов C и ф.р. $\mathbf{C}$. Обозначим $W_{k,j}$ — j -ю наименьшую незавершенную работу (среди всех занятых серверов) в момент времени t_k^- прихода k -й заявки, и пусть $W_k := (W_{k,1}, \dots, W_{k,N})$, таким образом, $W_{k,j} \leq W_{k,j+1}$. Если каждой заявке требуется только один сервер, $C_k = 1$, то система становится классической N -серверной системой вида $GI/G/N/\infty$, в которой векторы $\{W_k\}$ удовлетворяют рекурсии Кифера-Вольфовица

$$W_{k+1} = R(W_k + eS_k - \mathbf{e}\tau_k)^+, \quad (2.102)$$

где оператор R упорядочивает компоненты вектора по возрастанию, $e = (1, 0, \dots, 0)$ и $\mathbf{e}$ — единичный вектор. Для описанной кластерной модели справедлива следующая модификация вектора Кифера-Вольфовица

$$W_{k+1} = R(\overbrace{W_{k,C_k} + S_k - \tau_k, \dots, W_{k,C_k} + S_k - \tau_k}^{C_k \text{ компонент}}, W_{k,C_k+1} - \tau_k, \dots, W_{k,N} - \tau_k)^+. \quad (2.103)$$

Заметим, что последовательность $\{W_k, k \geq 1\}$ принимает значения на множестве

$$E = \{x \in \mathbf{R}^N : 0 \leq x_1 \leq \dots \leq x_N\}.$$

Ключевое отличие рекурсии (2.103) от классической рекурсии (2.102) состоит в том, что в кластерной модели время обслуживания S_k k -й заявки добавляется после момента прихода t_k к оставшейся работе на первых C_k серверах (с наименьшей оставшейся работой). Поскольку все сервера стохастически экви-

валентны, то в моменты прихода стохастическая эквивалентность сохраняется после перенумерации серверов (внутри рекурсии). Заметим, что вектор незавершенной нагрузки в кластерной модели включает в себя времена простоя, поскольку после прихода заявка, вообще говоря, должна ждать пока не освободится нужное число серверов. Условие стационарности таких систем получено в работе [147].

Для исходного N -серверного кластера (обозначим его Σ) построим минорантную и мажорантную системы (кластеры), которые будем обозначать $\Sigma^{(1)}$ и $\Sigma^{(m)}$, соответственно. В минорантной системе моменты прихода $\{t_k^{(1)}, k \geq 1\}$ определяются первым входным потоком и являются подпоследовательностью моментов прихода $\{t_k, k \geq 1\}$ в систему Σ , а время обслуживания – первая компонента смеси S , т. е. $S^{(1)}$. В мажорантной системе входной поток совпадает с потоком исходной системы, а время обслуживания – m -я компонента времени обслуживания, т. е. $S^{(m)}$. (ниже все неравенства между векторами понимаются покомпонентно). Напомним, что все соотношения между исходными с.в. понимаются в стохастическом смысле, однако они могут пониматься как соотношения с вероятностью 1 между с.в., стохастически эквивалентными исходным с.в., и определенными на одном и том же вероятностном пространстве [177]. При таком переопределении за новыми с.в. будем для удобства сохранять исходные обозначения.

Теорема 2.10. *Предположим, что в системах $\Sigma^{(1)}$, Σ и $\Sigma^{(m)}$ начальные моменты прихода совпадают, $t_1 = t_1^{(1)} = t_1^{(m)}$, выполнены условия стационарности и соотношения*

$$S_{st}^{(1)} \leq S_{st}^{(2)} \leq \dots \leq S_{st}^{(m)}, C_{st}^{(1)} \leq C_{st} \leq C_{st}^{(m)}. \quad (2.104)$$

Тогда

$$W_{st}^{(1)} \leq W_{st} \leq W_{st}^{(m)}. \quad (2.105)$$

Доказательство. Построим вспомогательную систему $\tilde{\Sigma}^{(1)}$, в которой входной поток такой же, как в системе Σ , $\tilde{C} = C_{st}^{(1)}$, а времена обслуживания $\tilde{S}_k = I_k^{(1)} S_k^{(1)}$, где $\sum_{i=1}^m I^{(i)} = 1$, где с.в. $I^{(i)}$ принимает значения 1 с вероятностью $p_i = \lambda_i / \lambda, i = 1, \dots, m$, и 0, иначе. Очевидно, $\tilde{W}_1 = W_1$ и $\tilde{S} \leq S_{st}$. По

индукции предположим, что $\tilde{W}_k \leq W_k$ для некоторого $k \geq 1$ и рассмотрим компоненты вектора незавершенной работы $\tilde{W}_{k+1,j}, W_{k+1,j}$ до применения оператора упорядочивания $R(\cdot)^+$. При $1 \leq j \leq \tilde{C}_k$ для всех компонент выполнено неравенство

$$\tilde{W}_{k,\tilde{C}_k} + \tilde{S}_k - \tau_k \leq W_{k,C_k} + S_k - \tau_k.$$

Если $\tilde{C}_k < j \leq C_k$, то

$$\tilde{W}_{k,j} - \tau_k \leq W_{k,j} - \tau_k \leq W_{k,C_k} + S_k - \tau_k.$$

В противном случае, если $C_k < j \leq N$, то

$$\tilde{W}_{k,j} - \tau_k \leq W_{k,j} - \tau_k.$$

Известно [212] и легко проверить, что оператор $R(\cdot)^+$ сохраняет порядок компонент, поэтому $\tilde{W}_{k+1} \leq W_{k+1}, k \geq 1$. Далее, из последовательности $\{t_k, k \geq 1\}$ выделим моменты времени, когда моменты прихода в систему $\tilde{\Sigma}^{(1)}$ совпадают с моментами прихода в систему $\Sigma^{(1)}$, т. е. $\tilde{t}_k = \min\{t_j \geq \tilde{t}_{k-1} : I_j^{(1)} = 1, j \geq k-1\}$. Остается заметить, что

$$W^{(1)}(t_k^{(1)}) = W_k^{(1)} = \tilde{W}(\tilde{t}_k) \leq W(\tilde{t}_k) = W_k, \quad k \geq 1.$$

Для доказательства того, что $\Sigma^{(m)}$ является мажорантной системой для исходной системы Σ построим вспомогательную систему $\hat{\Sigma}^{(m)}$, в которой входной поток такой же, как в системе Σ и $\hat{C} = C^{(m)}$, а времена обслуживания представляют собой смесь

$$\hat{S}_k = I_k^{(m)} S_k^{(m)}, \quad \sum_{i=1}^m I^{(i)} = 1.$$

Очевидно, $\hat{W}_1 = W_1$ и $S \leq_{st} \hat{S}$. Предположим, что $W_k \leq \hat{W}_k$ для некоторого $k \geq 1$ и рассмотрим компоненты вектора незавершенной работы $\hat{W}_{k+1,j}, W_{k+1,j}$ до применения оператора упорядочивания $R(\cdot)^+$. При $1 \leq j \leq C_k$ для всех компонент выполнено неравенство

$$W_{k,C_k} + S_k - \tau_k \leq \hat{W}_{k,\hat{C}_k} + \hat{S}_k - \tau_k.$$

Если $C_k < j \leq \hat{C}_k$, то

$$W_{k,j} - \tau_k \leq \hat{W}_{k,j} - \tau_k \leq \hat{W}_{k,\hat{C}_k} + \hat{S}_k - \tau_k.$$

В противном случае, если $\hat{C}_k < j \leq N$, то

$$W_{k,j} - \tau_k \leq \hat{W}_{k,j} - \tau_k.$$

Таким образом, $W_{k+1} \leq \hat{W}_{k+1}$, $k \geq 1$. □

В следующей Лемме 2.3 получены верхняя и нижняя границы для вектора незавершенной работы в исходной системе Σ через вектор незавершенной работы системы $\tilde{\Sigma}$, в которой время обслуживания меньше, а интервалы между приходами больше на случайное время исходных в системе Σ .

Лемма 2.3. *Предположим, что в системах Σ , $\tilde{\Sigma}$ начальная работа упорядочена $\tilde{W}_1 \leq W_1$ с в. 1, $\tilde{C}_i = C_i$, $i \geq 1$, и выполнены неравенства*

$$\tilde{\tau}_i = \tau_i + \xi_i, \quad \tilde{S}_i = S_i - \eta_i > 0, \quad i \geq 1, \quad (2.106)$$

где с.в. ξ_i , η_i неотрицательны. Тогда с в. 1

$$\tilde{W}_i \leq W_i \leq \tilde{W}_i + \sum_{j=1}^{i-1} (\xi_j + \eta_j) \mathbf{e}, \quad i \geq 1. \quad (2.107)$$

Доказательство. По индукции предположим, что (2.107) выполнено для некоторого $i \geq 1$. Заметим, что для любого $a \in \mathbf{R}$ и $b \geq 0$,

$$(a)^+ \leq (a - b)^+ + b.$$

Для любого j ,

$$(W_{i,j} - \tau_i)^+ \leq (W_{i,j} - \tau_i - \xi_i)^+ + \xi_i = (W_{i,j} - \hat{\tau}_i)^+ + \xi_i,$$

и, аналогично

$$(W_{i,j} + S_i - \tau_i)^+ \leq (W_{i,j} + \hat{S}_i - \hat{\tau}_i)^+ + \xi_i + \eta_i, \quad i \geq 1.$$

Теперь из (2.103) следует

$$\tilde{W}_{i+1} \leq W_{i+1} \leq \tilde{W}_{i+1} + \sum_{j=1}^i (\xi_j + \eta_j) \mathbf{e}, \quad i \geq 1. \quad (2.108)$$

□

Границы (2.107) важны для оценки близости между системами когда элементы управляющей последовательности (время обслуживания или интервалы между приходами заявок) отличаются на случайную величину.

2.8. Выводы по результатам главы 2

В Главе 2 получены следующие результаты.

Представлены новые модели систем обслуживания со смешанными распределениями, в которых времена обслуживания задаются конечной смесью распределений (раздел 2.4.), и модели систем обслуживания со смешанным входным процессом, заданным суперпозицией независимых стационарных входных потоков (раздел 2.6.)

Получены достаточные условия монотонности стационарного времени ожидания, стационарной длины очереди и стационарного числа заявок в многосерверных системах вида $GI/G/N/\infty$ с неограниченным буфером (Теорема 2.1, раздел 2.2.)

Получены достаточные условия монотонности стационарной длины очереди многосерверных систем обслуживания $GI/G/N/K$ с конечным буфером (Теорема 2.2, раздел 2.3.)

Получены достаточные условия монотонности стационарной длины очереди и стационарного времени ожидания в многосерверных системах со смешанными распределениями и неограниченной очередью (Теорема 2.3, раздел 2.4.)

Получены достаточные условия монотонности систем со смешанным обслуживанием (Теорема 2.4, раздел 2.4.)

Доказаны условия, при которых выполняется свойство монотонности для стационарной длины очереди в многосерверной системе с конечным буфером и смешанным обслуживанием (Теорема 2.5, раздел 2.4.)

Для односерверной системы с повторными вызовами m классов получены верхняя и нижняя границы незавершенной работы (Теорема 2.6). Также получены выражения для ПЛС стационарной незавершенной нагрузки в минорантной и мажорантной системах (Леммы 2.1 и 2.2). Получено в явном виде распределение стационарной незавершенной работы в минорантной системе для случая распределения Парето времени обслуживания (Теорема 2.7, раздел 2.5.)

Получены достаточные условия монотонности систем со смешанным входным процессом, заданным суперпозицией независимых стационарных входных потоков (Теоремы 2.8 и 2.9, раздел 2.6.)

Получены достаточные условия монотонности стационарной незавершенной работы высокопроизводительного кластера (Теорема 2.10, раздел 2.7.)

Основные результаты главы 2 диссертационной работы опубликованы в работах автора [194, 196, 197, 200, 201, 206, 207, 208, 209, 214, 220].

Глава 3. Системы обслуживания со смешанным входным процессом

В данной главе рассматриваются односерверные системы обслуживания со смешанным входным процессом, т. е. с суперпозицией m независимых входных потоков $\sum_i^m G_i / \sum_i^m G_i / 1$, свойства монотонности которых были исследованы в разделе 2.6. В таких системах распределение времени обслуживания зависит от класса обслуживания, определяемого входным потоком. В силу зависимости интервалов между приходами заявок (в случае непуассоновских потоков) такие системы обладают сложной корреляционной структурой и не регенерируют в классическом смысле.

Суперпозиции процессов изучались в работах [27, 37, 105, 137, 110, 180, 179]. Аппроксимация некоторых систем с суперпозицией входных процессов с помощью другой классической регенеративной системы вида $GI/G/1$ с единым входным процессом была реализована в ряде работ, например, [27, 179, 180, 181]. В этих работах, в частности, был использован так называемый метод стационарных интервалов (основанный на применении функций Пальма) для аппроксимации распределения времени между приходами заявок с помощью распределения стационарных н.о.р. интервалов. Такая замена является аппроксимацией, поскольку последовательные интервалы в суперпозиции, вообще говоря, зависимы. Проблема зависимости была проанализирована в статье [110], в которой подробно рассматривалась корреляционная структура последовательных периодов между событиями.

3.1. Аппроксимация суперпозиции процессов восстановления

Приведем вкратце теоретические результаты, касающиеся процессов восстановления и точечных процессов, представленные в работах [34, 44].

Рассмотрим н.о.р. неотрицательные с.в. $\{X_k, k \geq 2\}$ и не зависящую от них с.в. $X_1 \geq 0$. Пусть с.в. X – стохастическая копия с.в. $X_k, k \geq 2$. Предположим,

что интенсивность процесса восстановления $\lambda = 1/\mathbf{M}X < \infty$, а

$$G_X(x) = \lambda \int_0^x (1 - F_X(u)) du, \quad x \geq 0, \quad (3.1)$$

ф.р. с.в. X_1 . Известно, что (3.1) представляет собой распределение стационарного остаточного времени восстановления (равновесное распределение [161]). Обозначим через $T_k = \sum_{1 \leq i \leq k} X_i$ моменты восстановления, $k \geq 1$. Процесс

$$N_s(t) := \max\{k > 0 : T_k \leq t\}$$

представляет число моментов восстановления в интервале $[0, t]$, где $N_s(0) = 0$. Процесс $\{N_s(t), t \geq 0\}$ является процессом восстановления со стационарными приращениями [34, 127], т. е. удовлетворяет соотношению $\mathbf{M}N_s(t) = \lambda t$, $t \geq 0$, и распределение приращений $N_s(t+u) - N_s(u)$ не зависит от $u \geq 0$ для каждого фиксированного t .

Рассмотрим m независимых стационарных процессов восстановления. Пусть $\{T_k^{(i)}, k \geq 1\}$ – моменты восстановления для i -го процесса. Через $X^{(i)}$ обозначим обобщенную с.в. (интервалы между восстановлениями с интенсивностью $\lambda_i = 1/\mathbf{M}X^{(i)}$) для i -го процесса, где $i = 1, \dots, m$. Эти процессы формируют смешанный процесс восстановления $\mathcal{T} = \{T_k^{(i)}, k \geq 1; i = 1, \dots, m\}$. Обозначим через $\{T_k\}$ упорядоченную последовательность моментов из множества $\mathcal{T}$, т. е.

$$0 =: T_0 < T_1 < T_2 < \dots. \quad (3.2)$$

В отличие от обычного процесса восстановления, интервалы между моментами восстановления в смешанном процессе, образованном суперпозицией процессов восстановления, вообще говоря, зависимы.

Поскольку стационарные компоненты смешанного процесса имеют независимые приращения, то процесс

$$N_s(t) := N_s^{(1)}(t) + \dots + N_s^{(m)}(t), \quad t \geq 0, \quad (3.3)$$

является (считающим) процессом с независимыми стационарными приращениями (но не является процессом восстановления) [44].

Последовательность $\{T_k\}$ формирует считающий процесс со стационарными интервалами, являющийся стационарным процессом Пальма [44]. Рассмотрим считающий процесс $\{N(t), t \geq 0\}$, задающий число интервалов между событиями в интервале $[0, t]$,

$$N(t) = \max\{k \geq 0 : T_k \leq t\}, \quad t \geq 0, \quad (3.4)$$

где $N(0) = 1$. Обозначим через $X_k := T_k - T_{k-1}$ интервалы между k -м и $(k-1)$ -м моментами, $k \geq 1$. Такие интервалы одинаково распределены с некоторой ф.р. F_X [34, 44, 180]. Далее приведем соотношения, связывающие распределения Пальмовского процесса (3.4) и стационарного во времени (также считающего) процесса $\{N_s(t), t \geq 0\}$. Заметим, что последовательность интервалов $\{X_i\}$ (формирующая Пальмовский процесс) стационарна, если совместное распределение k -мерного вектора $(X_{i_1+h}, \dots, X_{i_k+h})$ не зависит от целого h для всех целых $k > 0$ и неотрицательных $(i_1, \dots, i_k)$. Пальмовский процесс $\{N(t)\}$ и ассоциированная с ним последовательность интервалов $\{X_k\}$ не всегда одновременно стационарны [180]. Тем не менее существует взаимно-однозначное соответствие между считающим стационарным процессом $\{N(t)\}$ и стационарной последовательностью $\{X_k\}$ [44]. Обозначим через $\varphi_k(t) = P(N(t) = k)$ функцию Пальма – вероятность появления k моментов ($k \geq 0$) в интервале $[0, t]$ при условии, что при $t = 0$ произошло событие. Функции Пальма описывают динамику Пальмовского процесса (3.4). Связь между распределениями процесса $\{N_s(t)\}$ и ассоциированного с ним процесса Пальма (3.4) имеет вид [98, 180]:

$$\begin{aligned} P(N_s(t) = k) &= \lambda \int_0^t (\varphi_{k-1}(u) - \varphi_k(u)) du, \quad k \geq 1, \\ P(N_s(t) = 0) &= 1 - \lambda \int_0^t \varphi_0(u) du =: 1 - G_X(t), \end{aligned} \quad (3.5)$$

$$G_X(t) = \lambda \int_0^t (1 - F_X(x)) dx, \quad t \geq 0, \quad (3.6)$$

откуда, в частности, следует

$$\varphi_0(t) = 1 - F_X(t) =: \bar{F}_X(t), \quad t \geq 0, \quad (3.7)$$

где F_X упомянутая выше ф.р. интервалов стационарного процесса Пальма. Процесс $\{N_s(t)\}$ по построению удовлетворяет условиям [180]:

$$MN_s(t) < \infty, \quad \lim_{t \rightarrow \infty} N_s(t)/t =: \lambda < \infty.$$

Процесс $\{N_s^{(i)}(t)\}$ представляет собой i -ю компоненту (стационарного процесса восстановления) $\{N_s(t)\}$. Из соотношения (3.1) следует, что первый интервал i -го процесса задан ф.р.

$$G_i(t) = \lambda_i \int_0^t (1 - F_i(u)) du, \quad (3.8)$$

где $\lambda_i = 1/MX^{(i)}$, $i = 1, \dots, m$. Пусть $\lambda = \lambda_1 + \dots + \lambda_m$. Известно [98, 180, 44], что функции Пальма $\varphi_k(t)$ стационарного процесса Пальма $\{N(t)\}$ могут быть выражены через распределение исходных стационарных во времени процессов $\{N_s^{(i)}(t)\}$ и их стационарных Пальмовских копий $\{N^{(i)}(t)\}$ следующим образом:

$$\begin{aligned} \varphi_k(t) &= \sum_{\sum k_i = k} \sum_{i=1}^m \frac{\lambda_i}{\lambda} \mathbf{P} \left(\sum_{j=1}^{i-1} N_s^{(j)}(t) + N^{(i)}(t) + \sum_{j=i+1}^m N_s^{(j)}(t) = k \right) \\ &= \sum_{\sum k_i = k} \sum_{i=1}^m \frac{\lambda_i}{\lambda} \mathbf{P}(N^{(i)}(t) = k_i) \prod_{j, j \neq i}^m \mathbf{P}(N_s^{(j)}(t) = k_j). \end{aligned} \quad (3.9)$$

Подчеркнем, что суммирование берется по всем различным разбиениям числа $k \geq 0$, таким что $k = k_1 + \dots + k_m$.

Из (3.8) следует, что хвост ф.р. F_X удовлетворяет [180]:

$$\begin{aligned} \bar{F}_X(t) &= \varphi_0(t) = \mathbf{P}(N(t) = 0) = \sum_{i=1}^m \frac{\lambda_i}{\lambda} (1 - F_i(t)) \prod_{j, j \neq i} G_j(t) \\ &= \sum_{i=1}^m \frac{\lambda_i}{\lambda} (1 - F_i(t)) \prod_{j, j \neq i} \lambda_j \int_t^\infty (1 - F_j(u)) du, \quad t \geq 0. \end{aligned} \quad (3.10)$$

Легко проверить, что $\bar{F}_X(0) = \varphi(0) = 1$. Заметим, что процесс Пальма $\{N(t)\}$

со стационарными интервалами, а процесс $\{N_s(t)\}$ имеет независимые приращения. Кроме того, поскольку для однокомпонентного процесса с распределением времени между событиями F и средним λ верно

$$\lambda \int_0^\infty \bar{F}_X(t) dt = 1,$$

то уравнение (3.5) влечет, $P(N_s(t) = 0) = G_X(t)$, удовлетворяющее соотношению (3.6).

Поскольку процесс восстановления является частным случаем точечного процесса, стоит отметить, что F_X представляет собой точное предельное (одномерное) распределение времени между приходами в смешанном процессе Пальма $\{N(t)\}$. Однако это распределение не отражает зависимость между соседними интервалами и поэтому (3.10) дает лишь аппроксимацию для исходного процесса.

3.2. Аппроксимация входного процесса, заданного суперпозицией показательного и Парето распределений

Обозначим исходную систему $\sum_{i=1}^m G_i / \sum_{i=1}^m G_i / 1$ через Σ . Пусть i -я компонента входного процесса задана н.о.р. интервалами между приходами заявок $\{\tau_k^{(i)}, k \geq 2\}$ с ф.р. A_i , а ф.р. времени до момента первого прихода удовлетворяет (3.1) и (3.8)

$$P(\tau_1^{(i)} \leq x) = \frac{1}{M\tau_2^{(i)}} \int_0^x (1 - A_i(u)) du, \quad i = 1, \dots, m.$$

Времена обслуживания i -го класса $\{S_k^{(i)}, k \geq 1\}$ являются н.о.р. и задаются ф.р. B_i , $i = 1, \dots, m$. Предположим, что первые n_1 потоков являются пуассоновскими с ф.р. интервалов $\tau_k^{(i)}$ между заявками

$$A_i(x) = 1 - e^{-\lambda_i x}, \quad x \geq 0, \quad \lambda_i > 0, \quad i = 1, \dots, n_1, \quad (3.11)$$

интервалы в остальных n_2 компонентах заданы распределениями Парето

$$P(\tau_k^{(j)} \leq x) =: A_j(x) = 1 - \left(\frac{x_0}{x_0 + x} \right)^{\alpha_j}, \quad x \geq 0, \quad k \geq 2, \quad x_0 > 0, \quad \alpha_j > 1, \quad (3.12)$$

$$j = 1, \dots, n_2.$$

Пусть $\mu_i = 1/MS_k^{(i)}$, $i = 1, \dots, m$. Далее индексом i будем обозначать компоненты, имеющие распределение (3.11), а индексом j – компоненты с распределением (3.12). Подставив эти распределения в формулу для аппроксимации входного потока (3.10), получим ф.р.

$$A(x) = 1 - pe^{-\lambda x} \left(\frac{x_0}{x_0 + x} \right)^{\alpha - n_2} - (1 - p)e^{-\lambda x} \left(\frac{x_0}{x_0 + x} \right)^{\alpha - n_2 + 1}, \quad (3.13)$$

где

$$\lambda = \lambda_1 + \dots + \lambda_{n_1}, \quad \alpha = \alpha_1 + \dots + \alpha_{n_2},$$

и

$$p = \frac{\lambda x_0}{\lambda x_0 + \alpha - n_2}. \quad (3.14)$$

Заметим, что хвост ф.р. (3.13) представляет собой смесь вида

$$\bar{A}(x) = p\bar{F}_1(x) + (1 - p)\bar{F}_2(x), \quad (3.15)$$

где

$$\begin{aligned} \bar{F}_1(x) &= e^{-\lambda x} \left(\frac{x_0}{x_0 + x} \right)^{\alpha - n_2}, \\ \bar{F}_2(x) &= e^{-\lambda x} \left(\frac{x_0}{x_0 + x} \right)^{\alpha - n_2 + 1}, \end{aligned} \quad (3.16)$$

с коэффициентом смешивания p , определенным в (3.14). Обозначим через τ_A с.в. с ф.р. (3.15). Обозначим через Y с.в. с ф.р. F_1 , а через Z с.в. с ф.р. F_2 . Тогда согласно (3.16)

$$\begin{aligned} Y &= \min(Y_1, Y_2), \quad \text{где } Y_1 \sim \text{Exp}(\lambda), \quad Y_2 \sim \text{Pareto}(\alpha - n_2, x_0), \\ Z &= \min(Y_1, Z_1), \quad \text{где } Z_1 \sim \text{Pareto}(\alpha - n_2 + 1, x_0), \end{aligned}$$

и с.в. τ_A может быть представлена в виде

$$\tau_A = IY + (1 - I)Z, \quad (3.17)$$

где I – индикатор, $P(I = 1) = p$. Если класс обслуживания определяется номером входного потока, k , то хвост ф.р. времени обслуживания в системе Σ может быть также представлен в виде смеси ф.р. $\{B_k\}$ времен обслуживания разных классов, а именно

$$\bar{B}(x) = \sum_{i=1}^{n_1} p_i \bar{B}_i(x) + \sum_{j=1}^{n_2} p_j \bar{B}_j(x), \quad (3.18)$$

где

$$p_i = \frac{\lambda_i x_0}{\lambda x_0 + \alpha - n_2}, \quad i = 1, \dots, n_1; \quad p_j = \frac{\alpha_j - 1}{\lambda x_0 + \alpha - n_2}, \quad j = 1, \dots, n_2.$$

Таким образом, вместо исходной системы Σ рассмотрим классическую систему $\hat{\Sigma}$, в которой интервалы между приходами заявок имеют ф.р. (3.13), а времена обслуживания заданы смесью (3.18), т. е. систему вида $GI/G/1$.

Для конечности средней длины цикла регенерации необходимо выполнение условия стационарности системы $\hat{\Sigma}$. Обозначим через S с.в., соответствующую ф.р. (3.18). Условие стационарности системы $\hat{\Sigma}$ имеет вид [127]:

$$\rho := \frac{MS}{M\tau} < 1, \quad (3.19)$$

где

$$MS = \sum_{i=1}^{n_1} p_i MS^{(i)} + \sum_{j=1}^{n_2} p_j MS^{(j)} = \sum_{i=1}^{n_1} \frac{p_i}{\mu_i} + \sum_{j=1}^{n_2} \frac{p_j}{\mu_j}. \quad (3.20)$$

Вычислим MY :

$$\begin{aligned} MY = M \min(Y_1, Y_2) &= \int_0^\infty x f_{Y_1}(x) \left(\int_x^\infty f_{Y_2}(y) dy \right) dx + \int_0^\infty y f_{Y_2}(y) \left(\int_y^\infty f_{Y_1}(x) dx \right) dy \\ &= \int_0^\infty x \lambda e^{-\lambda x} x_0^{\alpha-n_2} (x_0 + x)^{-\alpha+n_2} dx + \int_0^\infty x e^{-\lambda x} (\alpha - n_2) x_0^{\alpha-n_2} (x_0 + x)^{-\alpha+n_2-1} dx \end{aligned}$$

$$= \frac{1}{\lambda} \left(1 - (\alpha - n_2) e^{\lambda x_0} (\lambda x_0)^{\alpha - n_2} \Gamma(-\alpha + n_2, \lambda x_0) \right).$$

Далее находим

$$MZ = x_0 e^{\lambda x_0} (\lambda x_0)^{\alpha - n_2} \Gamma(-\alpha + n_2, \lambda x_0),$$

что дает

$$M\tau_A = pMY + (1 - p)MZ = \frac{p}{\lambda} = \frac{x_0}{\lambda x_0 + \alpha - n_2}. \quad (3.21)$$

Пусть

$$\rho_i := \frac{MS^{(i)}}{M\tau^{(i)}} = \frac{\lambda_i}{\mu_i}, \quad i = 1, \dots, n_1, \quad \rho_j := \frac{MS^{(j)}}{M\tau^{(j)}} = \frac{\alpha_j - 1}{x_0 \mu_j}, \quad j = 1, \dots, n_2.$$

Тогда легко проверить, используя (3.20) и (3.21), что

$$\rho = \sum_{i=1}^{n_1} \rho_i + \sum_{j=1}^{n_2} \rho_j. \quad (3.22)$$

Будем обозначать через W_n время ожидания n -й заявки в исходной системе Σ со смешанным входным процессом, а через W_n^A – время ожидания n -й заявки в аппроксимирующей системе $\hat{\Sigma}$. Рассмотрим стационарные пределы

$$W_n \Rightarrow W, \quad W_n^A \Rightarrow W^A, \quad n \rightarrow \infty.$$

Эти пределы существуют, если выполнено условие стационарности $\rho < 1$.

3.2.1. Два класса обслуживания

Предположим, что $\lambda_1 = \dots = \lambda_{n_1} =: \lambda$ и $\alpha_1 = \dots = \alpha_{n_2} =: \alpha$. В этом случае в системе только два класса обслуживания, а ф.р. времени обслуживания принимает вид

$$\overline{B}(x) = p\overline{B}_1(x) + (1 - p)\overline{B}_2(x), \quad x \geq 0. \quad (3.23)$$

Тогда хвост ф.р. входного потока в системе $\hat{\Sigma}$ принимает вид

$$\overline{A}(x) = p e^{-\lambda n_1 x} \left(\frac{x_0}{x_0 + x} \right)^{n_2(\alpha-1)} + (1 - p) e^{-\lambda n_1 x} \left(\frac{x_0}{x_0 + x} \right)^{n_2(\alpha-1)+1}, \quad (3.24)$$

где

$$p = \frac{n_1 x_0 \lambda}{n_1 x_0 \lambda + n_2(\alpha - 1)}.$$

3.2.2. Суперпозиция двух распределений Парето

Пусть входной поток задан двумя потоками ($n_1 = 0, n_2 = 2$), $\tau^{(i)} \sim \text{Pareto}(\alpha_i, x_0)$, $\alpha_i > 1, i = 1, 2$. Используя (3.10), найдем, что хвост распределения входного потока в системе $\widehat{\Sigma}$ имеет вид

$$\overline{A}(x) = \left(\frac{x_0}{x_0 + x} \right)^{\alpha_1 + \alpha_2 - 1}, \quad x \geq 0. \quad (3.25)$$

Таким образом, входной поток имеет также распределение Парето с параметрами $\alpha_1 + \alpha_2 - 1, x_0$. Время обслуживания можно представить в виде смеси распределений с коэффициентом пропорции смеси

$$p = \frac{\alpha_1 - 1}{\alpha_1 + \alpha_2 - 2},$$

а условие стационарности (3.22) преобразуется в выражение

$$\rho = \frac{\alpha_1 - 1}{x_0 \mu_1} + \frac{\alpha_2 - 1}{x_0 \mu_2} < 1. \quad (3.26)$$

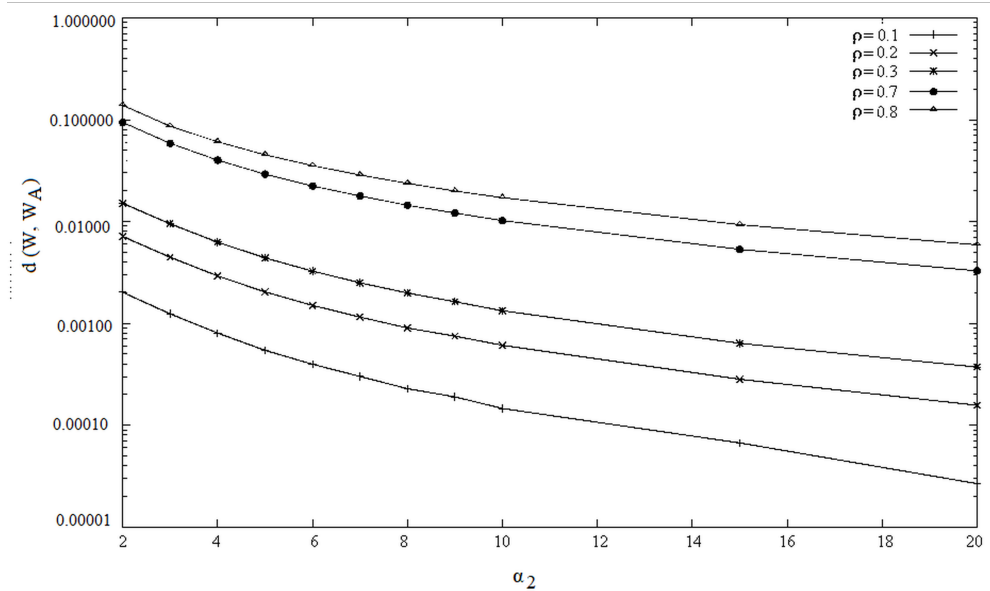


Рисунок 3.1. Расстояние Колмогорова между эмпирическими ф.р. в системах Σ и $\widehat{\Sigma}$ в зависимости от ρ и α_2 .

Рассмотрим систему, в которой входной процесс является суперпозицией двух потоков с распределением Парето с параметрами $\alpha_1 = 2, \alpha_2 \in [2, \dots, 20]$ и $x_0 = 1$. Время обслуживания показательное с параметром μ и не зависит от

класса. На Рисунке 3.1 представлены результаты экспериментов по вычислению расстояния Колмогорова между эмпирической ф.р. времени ожидания W в исходной системе Σ и эмпирической ф.р. в аппроксимирующей системе $\tilde{\Sigma}$ в зависимости от коэффициента загрузки ρ и параметра α_2 . Как видно аппроксимация точнее для больших значений α_2 и малых значениях коэффициента загрузки.

3.2.3. Один класс обслуживания

Рассмотрим систему $\sum_{i=1}^m G_i/M/1$, в которой все заявки имеют один класс обслуживания с показательным распределением с параметром $\mu := \mu_1 = \dots = \mu_n$, $S \sim \text{Exp}(\mu)$. В этом случае аппроксимирующая система – это классическая система $GI/M/1$, для которой известны некоторые аналитические результаты, см., например, [63]. Вначале найдем ПЛС ф.р. (3.13)

$$\begin{aligned} \psi_A(z) &= - \int_0^\infty e^{-zx} d\bar{A}(x) = 1 - z \int_0^\infty e^{-zx} \bar{A}(x) dx \\ &= 1 - pz \int_0^\infty e^{-zx} e^{-\lambda x} \left(\frac{x_0}{x_0 + x} \right)^{\alpha - n_2} dx - (1 - p)z \int_0^\infty e^{-zx} e^{-\lambda x} \left(\frac{x_0}{x_0 + x} \right)^{\alpha - n_2 + 1} dx \\ &= 1 - pze^{x_0(z+\lambda)} x_0^{\alpha - n_2} (z + \lambda)^{\alpha - n_2 - 1} \Gamma(-\alpha + n_2 + 1, x_0(z + \lambda)) \\ &\quad - (1 - p)ze^{x_0(z+\lambda)} x_0^{\alpha - n_2 + 1} (z + \lambda)^{\alpha - n_2} \Gamma(-\alpha + n_2, x_0(z + \lambda)), \end{aligned}$$

где $\Gamma(\xi, x)$ – неполная верхняя Гамма-функция. Далее, применяя свойство Гамма-функции

$$\Gamma(\xi + 1, x) = \xi \Gamma(\xi, x) + x^\xi e^{-x},$$

получим

$$\begin{aligned} \psi_A(z) &= 1 - \frac{pz}{z + \lambda} + \frac{z}{z + \lambda} e^{x_0(z+\lambda)} x_0^{\alpha - n_2} (z + \lambda)^{\alpha - n_2} \\ &\quad \times \Gamma(-\alpha + n_2, x_0(z + \lambda)) (p(\alpha - n_2) - (1 - p)x_0(z + \lambda)). \end{aligned} \quad (3.27)$$

Это ПЛС используется затем для нахождения корня уравнения

$$z - \psi_A(\mu(1 - z)) = 0, \quad (3.28)$$

который входит в известную формулу ф.р. стационарного времени ожидания [34]

$$F_W(x) = 1 - \sigma e^{-\mu(1-\sigma)x}, \quad x > 0. \quad (3.29)$$

Таблица 3.1. Расстояние Колмогорова между эмпирической ф.р. времени ожидания W в исходной системе Σ и эмпирической ф.р. W_A в аппроксимирующей системе $\tilde{\Sigma}$, теоретической ф.р. с.в. W_T . при $\alpha_1 = 2$ и $x_0 = 1$.

α_2	ρ	μ	σ	$d(W, W_A)$	$d(W, W_T)$	$d(W_T, W_A)$
8	0.8	10	0.8196	0.02369	0.02368	1.307×10^{-5}
3	0.75	4	0.8078	0.07127	0.07126	1.306×10^{-5}
2	0.67	3	0.7641	0.08289	0.08288	1.030×10^{-5}
3	0.5	6	0.5766	0.02633	0.02632	6.125×10^{-6}
4	0.5	8	0.5587	0.01757	0.01756	5.919×10^{-6}
5	0.5	10	0.5475	0.01253	0.01252	6.097×10^{-6}
10	0.5	20	0.5244	0.00401	0.00400	5.044×10^{-6}
3	0.3	10	0.3655	0.00949	0.00949	2.680×10^{-6}

Рассмотрим пример из раздела 3.2.2., сравнивая полученные в результате имитационного моделирования эмпирическую ф.р. стационарного времени ожидания W в исходной системе Σ , эмпирическую ф.р. стационарного времени ожидания W_A в аппроксимирующей системе $\tilde{\Sigma}$ с теоретической ф.р. W_T (формула (3.29)). Результаты обобщены в Таблице 3.1, где также приведены значения σ при различных значениях μ . Результаты моделирования подтверждают хорошую точность аппроксимации при $\rho < 0.6$.

Графики ф.р. для «наихудшего» и «наилучшего» случаев приведены на рисунках 3.2 и 3.3, соответственно. В обоих случаях, теоретическая ф.р. с.в. W_T близка к эмпирической ф.р. W_A в аппроксимирующей системе с точностью до 10^{-5} (см. соответствующие значения $d(W_T, W_A)$ в Таблице 3.1).

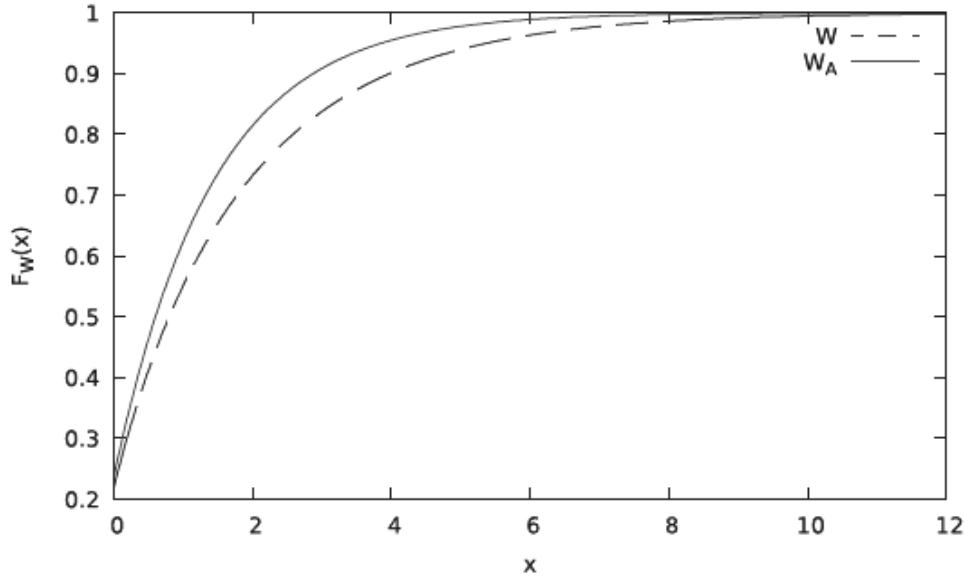


Рисунок 3.2. Графики эмпирических ф.р. с.в. W и W_A при $\alpha_1 = 2$, $\alpha_2 = 2$, $x_0 = 1$, и $\mu = 3$.

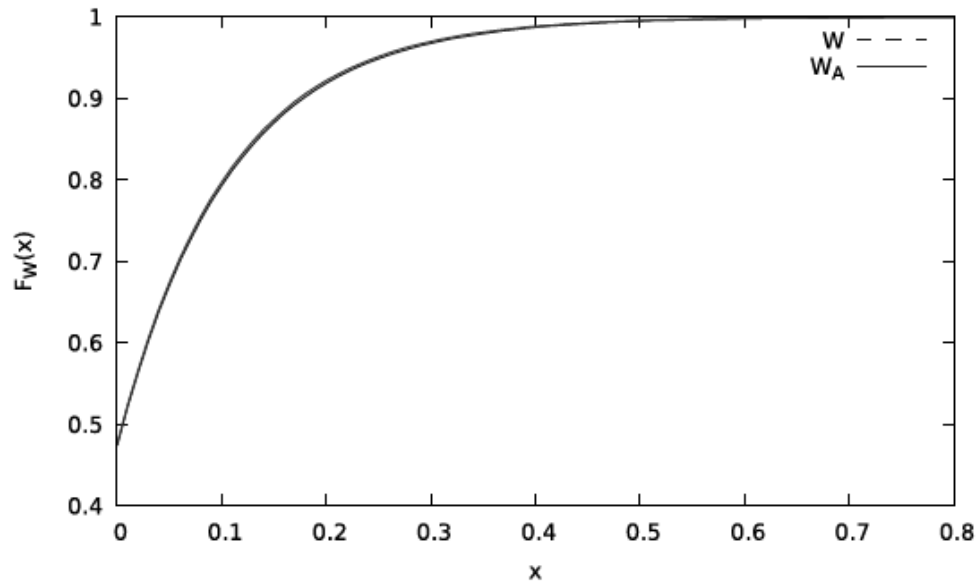


Рисунок 3.3. Графики эмпирических ф.р. с.в. W и W_A при $\alpha_1 = 2$, $\alpha_2 = 10$, $x_0 = 1$, и $\mu = 20$.

3.2.4. Два входных потока: пуассоновский и Парето

Рассмотрим суперпозицию двух входных потоков ($n_1 = 1, n_2 = 1$), заданных ф.р.

$$A_1(x) = 1 - e^{-\lambda x} \quad \text{и} \quad A_2(x) = 1 - \left(\frac{x_0}{x_0 + x} \right)^\alpha. \quad (3.30)$$

Подставим ф.р. (3.30) в (3.10), получим хвост ф.р. с.в. τ_A

$$\bar{A}(x) = pe^{-\lambda x} \left(\frac{x_0}{x_0 + x} \right)^{\alpha-1} + (1-p)e^{-\lambda x} \left(\frac{x_0}{x_0 + x} \right)^{\alpha}, \quad (3.31)$$

где

$$p = \frac{\lambda x_0}{\alpha - 1 + \lambda x_0}. \quad (3.32)$$

Определим ф.р. $\bar{F}_1(x)$ и $\bar{F}_2(x)$ как

$$\bar{F}_1(x) = e^{-\lambda x} \left(\frac{x_0}{x_0 + x} \right)^{\alpha-1} \quad \text{и} \quad \bar{F}_2(x) = e^{-\lambda x} \left(\frac{x_0}{x_0 + x} \right)^{\alpha}.$$

Тогда с.в. τ_A с ф.р. (3.31) может быть представлена в виде двухкомпонентной смеси (3.17) с.в. Y , имеющей ф.р. F_1 , и с.в. Z , имеющей ф.р. F_2 (с коэффициентом пропорции смеси (3.32)), причем

$$Y = \min(Y_1, Y_2) \quad \text{где} \quad Y_1 \sim \text{Exp}(\lambda), \quad Y_2 \sim \text{Pareto}(x_0, \alpha - 1)$$

и

$$Z = \min(Y_1, Z_1) \quad \text{где} \quad Z_1 \sim \text{Pareto}(x_0, \alpha).$$

Отметим, что в этом случае условие стационарности принимает вид

$$\rho = \frac{\lambda}{\mu_1} + \frac{\alpha - 1}{x_0 \mu_2} < 1.$$

Пример 3.1. Пусть в исходной системе входной поток есть суперпозиция потока с показательным распределением с параметром $\lambda = 1$ и потока с распределением Парето с параметрами $\alpha = 2$ и $x_0 = 1$. Время обслуживания не зависит от класса и является показательным с параметром $\mu = 3$. Результаты моделирования представлены на Рисунке 3.4 и в Таблице 3.2.

Пример 3.2. Пусть в системе два класса обслуживания с распределениями $\text{Exp}(\mu_1)$ и $\text{Exp}(\mu_2)$. В Таблице 3.3 представлено расстояние Колмогорова $d(W, W_A)$ при различных значениях параметров μ_1 и μ_2 . В этом случае ф.р. времени обслуживания (3.23) представляет собой гипергеометрическое распределение с параметрами μ_1 , μ_2 и p . Пусть параметры входного потока $\alpha_1 = 2$,

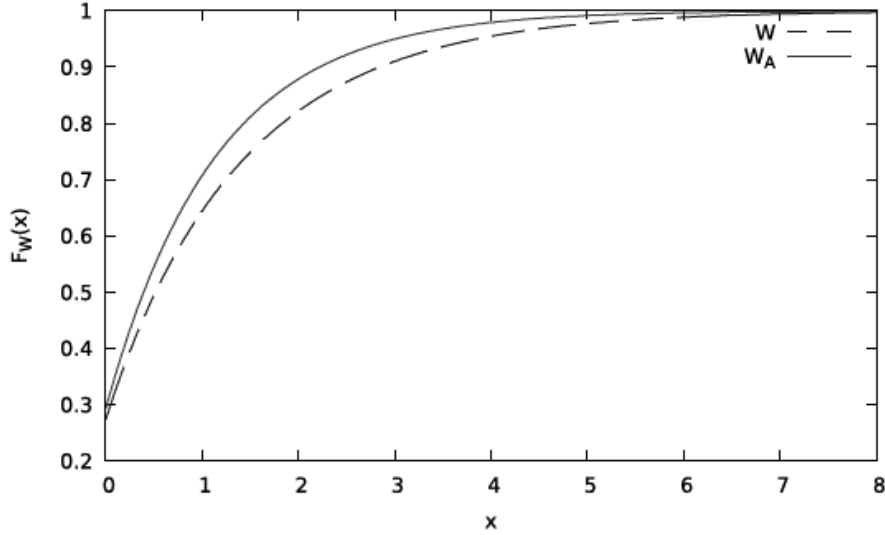


Рисунок 3.4. Графики эмпирических ф.р. стационарного времени ожидания W в исходной системе Σ и стационарного времени ожидания W_A в аппроксимирующей системе $\tilde{\Sigma}$ при $\alpha = 2$, $x_0 = 1$, $\lambda = 1$, и $\mu = 3$.

Таблица 3.2. Расстояние Колмогорова между эмпирической ф.р. времени ожидания W в исходной системе Σ и эмпирической ф.р. W_A в аппроксимирующей системе $\tilde{\Sigma}$, теоретической ф.р. с.в. W_T .

α	x_0	λ	μ	σ	$d(W, W_A)$	$d(W, W_T)$	$d(W_T, W_A)$
2	1	1	3	0.707	0.064	0.064	9.84×10^{-6}
2	1	1	4	0.548	0.032	0.032	5.97×10^{-6}
3	1	0.5	4	0.688	0.023	0.023	7.56×10^{-6}
2	4	1	3	0.425	0.001	0.006	2.73×10^{-6}

$\alpha_2 = 3$ и $x_0 = 1$. Как показывает таблица значения расстояния Колмогорова $d(W, W_A) < 5\%$ при $\rho < 0.6$.

3.3. Аппроксимация входного процесса, заданного суперпозицией распределений Вейбулла

Пусть интервалы между приходами заявок $\{\tau_k^{(i)}, k \geq 1\}$ имеют распределение $We(\alpha_i, \beta_i)$

$$A_i(x) = 1 - e^{-(x/\alpha_i)^{\beta_i}}, \quad x \geq 0, \quad \alpha_i > 0, \quad 0 < \beta_i < 1, \quad i = 1, \dots, m. \quad (3.33)$$

Таблица 3.3. Расстояние Колмогорова между эмпирической ф.р. стационарного времени ожидания W в исходной системе Σ , стационарного времени ожидания W_A в аппроксимирующей системе $\tilde{\Sigma}$ с двумя классами обслуживания.

μ_1	3	4	4	4	4	6	6	6
μ_2	4	5	6	7	8	7	8	9
ρ	0.833	0.650	0.583	0.536	0.472	0.452	0.417	0.389
$d(W, W_A)$	0.118	0.054	0.045	0.040	0.037	0.023	0.020	0.018

Подставляя (3.33) в формулу (3.10), получим хвост распределения смеси

$$\bar{A}(x) = \sum_{i=1}^m p_i \bar{F}_i(x), \quad (3.34)$$

где

$$\bar{F}_i(x) = e^{-(x/\alpha_i)^{\beta_i}} \prod_{j=1, j \neq i}^m \frac{1}{\Gamma(1/\beta_j)} \Gamma(1/\beta_j, (x/\alpha_j)^{\beta_j}), \quad (3.35)$$

с коэффициентами пропорции

$$p_i = \frac{\beta_i \prod_{j=1, j \neq i}^n \alpha_j \Gamma(1/\beta_j)}{\sum_{i=1}^m \beta_i \prod_{j=1, j \neq i}^m \alpha_j \Gamma(1/\beta_j)}, \quad i = 1, \dots, m. \quad (3.36)$$

Используя представление для Гамма-функции

$$\Gamma(\xi, x) = \Gamma(\xi) - \gamma(\xi, x),$$

получим выражение (3.35) в виде

$$\bar{F}_i(x) = e^{-(x/\alpha_i)^{\beta_i}} \prod_{j=1, j \neq i}^m \left(1 - \frac{1}{\Gamma(1/\beta_j)} \gamma(1/\beta_j, (x/\alpha_j)^{\beta_j}) \right). \quad (3.37)$$

Обозначим через τ_A интервал между приходами заявок, соответствующий ф.р. (3.34). Тогда распределение с.в. τ_A можно записать в виде смеси

$$\tau_A = I_1 X_1 + \dots + I_n X_m, \quad (3.38)$$

где индикатор $I_i = 1$ с вероятностью p_i , $\sum_{i=1}^m I_i = 1$ и

$$X_i = \min(Z_i, Y_j, j \neq i, j = 1, \dots, m), \quad (3.39)$$

где с.в.

$$Z_i \sim We(\alpha_i, \beta_i), \quad Y_j \sim SGGa(\alpha_j, 1/\beta_j, \beta_j),$$

а через $SGGa(\alpha_j, a_j, c_j)$ обозначено обобщенное распределение Стэйси-Гамма с ф.р.

$$F_{Y_j}(x) = \frac{1}{\Gamma(a_j)} \gamma(a_j, (x/\alpha_j)^{c_j}), \quad x \geq 0, \quad (3.40)$$

плотностью

$$f_{Y_j}(x) = \frac{c_j}{\alpha_j \Gamma(a_j)} (x/\alpha_j)^{c_j a_j - 1} e^{-(x/\alpha_j)^{c_j}}, \quad x \geq 0$$

и первыми двумя моментами:

$$MY_j = \frac{\alpha_j \Gamma(a_j + 1/c_j)}{\Gamma(a_j)}, \quad MY_j^2 = \frac{\alpha_j^2 \Gamma(a_j + 2/c_j)}{\Gamma(a_j)},$$

где параметры $a_j = 1/\beta_j, c_j = \beta_j, j = 1, \dots, m$.

Пример 3.3. Рассмотрим систему, в которой входной поток представляет собой суперпозицию двух входных потоков с распределением Вейбулла (3.33) $i = 1, 2$. Время обслуживания не зависит от класса и имеет показательное распределение с параметром μ . Условие стационарности такой системы имеет вид

$$\rho = \frac{\beta_1}{\alpha_1 \mu_1 \Gamma(1/\beta_1)} + \frac{\beta_2}{\alpha_2 \mu_2 \Gamma(1/\beta_2)} < 1.$$

Далее приведены результаты численных экспериментов по оценке расстояния Колмогорова между эмпирической ф.р. времени ожидания W в исходной системе Σ и эмпирической ф.р. времени ожидания W_A в аппроксимирующей системе $\tilde{\Sigma}$.

Первый эксперимент проведен при параметрах входного потока $\beta_1 = \beta_2 = 1, \alpha_1 = 1, \alpha_2 = 4$ (суперпозиция двух пуассоновских входных потоков). В этом случае аппроксимирующая система это система $M/M/1$, в которой входной поток – пуассоновский с параметром $1/\alpha_1 + 1/\alpha_2$, что позволило сравнить траектории времени ожидания W в исходной системе и W_A в аппроксимирующей системе. Отметим, что абсолютная ошибка не превышает 8×10^{-5} для любых

Таблица 3.4. Расстояние Колмогорова $d(W, W_A)$, оценки м.о. $Mean(W)$, $Mean(W_A)$ и дисперсий $Var(W)$, $Var(W_A)$ при различных ρ .

$\beta_1 = 5, \beta_2 = 10, \alpha_1 = 1, \alpha_2 = 1$					
ρ	$d(W, W_A)$	$Mean(W)$	$Mean(W_A)$	$Var(W)$	$Var(W_A)$
0.1	0.00277	0.00234	0.00260	0.00021	0.00025
0.2	0.01181	0.00948	0.01189	0.00169	0.00237
0.3	0.02499	0.02295	0.03152	0.006107	0.00983
0.4	0.03849	0.04763	0.06805	0.01745	0.03007
0.5	0.05085	0.09289	0.13311	0.04568	0.07991
0.6	0.06189	0.17661	0.24946	0.11651	0.20209
0.7	0.07168	0.33954	0.46978	0.30896	0.52809
0.8	0.08027	0.70258	0.95107	0.95681	1.61544
0.9	0.08808	1.87013	2.47942	4.89730	8.23029

$\rho < 0.9$.

В Таблице 3.4 приведены результаты численных экспериментов при параметрах $\beta_1 = 5, \beta_2 = 10, \alpha_1 = 1, \alpha_2 = 1$, а в Таблице 3.5 при параметрах входного потока $\beta_1 = 0.1, \beta_2 = 0, 2, \alpha_1 = 1, \alpha_2 = 1$.

По результатам численных экспериментов можно сделать вывод о том, что расстояние Колмогорова $d(W, W_A)$ меньше 5% при любом коэффициенте загрузки ρ для случая распределений Вейбулла с тяжелым хвостом (Таблица 3.5). В случае $\beta_i > 1$ (Таблица 3.4) расстояние Колмогорова не превосходит 5% только при $\rho \leq 0.5$, т. е. при $\rho > 0.5$ аппроксимацию следует использовать с осторожностью.

3.4. Выводы по результатам главы 3

Получено аппроксимирующее распределение входного потока, заданного суперпозицией показательного и Парето распределений (раздел 3.2.) Это распределение имеет вид конечной смеси. Для случая одного класса заявок (когда время обслуживания задано одним и тем же показательным распределением) получено ПЛС стационарного времени ожидания, что позволяет получить теоретическую ф.р. для стационарного времени ожидания в явном виде для аппро-

Таблица 3.5. Расстояние Колмогорова $d(W, W_A)$, оценки м.о. $Mean(W)$, $Mean(W_A)$ и дисперсий $Var(W)$, $Var(W_A)$ при различных ρ .

$\beta_1 = 0.1, \beta_2 = 0, 2, \alpha_1 = 1, \alpha_2 = 1$					
ρ	$d(W, W_A)$	$Mean(W)$	$Mean(W_A)$	$Var(W)$	$Var(W_A)$
0.1	1.909×10^{-5}	95.0969	95.0917	11325.2	11322.9
0.2	8.228×10^{-5}	355.945	355.87	143765	143709
0.3	9.561×10^{-5}	874.066	873.498	827005	825552
0.4	2.952×10^{-5}	1827.71	1827.88	$3.52 \times 10^{+6}$	$3.52 \times 10^{+6}$
0.5	11.72×10^{-5}	3572.09	3571.48	$1.32 \times 10^{+7}$	$1.32 \times 10^{+7}$
0.6	20.18×10^{-5}	6869.13	6865.69	$4.82 \times 10^{+7}$	$4.81 \times 10^{+7}$
0.7	13.88×10^{-5}	13652.6	13652.4	$1.89 \times 10^{+8}$	$1.89 \times 10^{+8}$
0.8	69.37×10^{-5}	30214.2	30144.3	$9.19 \times 10^{+8}$	$9.13 \times 10^{+8}$
0.9	14.63×10^{-4}	90455.3	90122.8	$8.19 \times 10^{+9}$	$8.11 \times 10^{+9}$

ксимирующей системы. Результаты численных экспериментов подтверждают, что аппроксимация является достаточно точной, когда коэффициент загрузки системы не больше 0.6.

Получено аппроксимирующее распределение входного потока, заданного суперпозицией входных потоков с распределением Вейбулла. Такое распределение имеет вид конечной смеси распределений Вейбулла и обобщенных распределений Стейси-Гамма (раздел 3.3.) Результаты численных экспериментов подтверждают, что аппроксимация является достаточно точной при $\beta_i \leq 1$ (т. е. в случае, когда распределение Вейбулла относится к классу распределений с тяжелым хвостом), а в случае $\beta_i > 1$ это верно, когда коэффициент загрузки системы не больше 0.5.

Основные результаты Главы 3 опубликованы в работах автора [191, 193].

Глава 4. Экстремумы стационарных последовательностей и их применение к системам обслуживания

В данной главе введены понятия и представлены некоторые сведения из теории экстремальных значений, на основе которых получены результаты об упорядоченности экстремальных индексов стационарного времени ожидания в системах $M/G/1$ и $GI/M/1$ для случая, когда уравнение $Me^{\gamma(S-\tau)} = 1$ имеет действительный положительный корень γ (такое решение существует для с.в. $S - \tau$, имеющих распределение с легким хвостом). Материал данной главы также использован в Главе 5 при исследовании монотонности экстремального индекса в системах с субъэкспоненциальным обслуживанием.

Моделирование и прогнозирование экстремальных значений при анализе производительности систем массового обслуживания является важной задачей, которую обычно решают с помощью *теории экстремальных значений* [68, 91, 107]. Классические предельные теоремы об экстремальных значениях для н.о.р. с.в. в таких сложных системах не работают, поскольку исследуемые процессы сильно коррелированы [48].

Предельное распределение максимумов строго стационарных последовательностей с.в. выражается через так называемый *экстремальный индекс* θ , который отражает кластерную структуру последовательности или ее локальную структуру зависимости [107, 143, 166, 164, 175]. Например, для последовательности независимых с.в. экстремальный индекс $\theta = 1$, а превышения высокого порогового значения u_n происходят по одиночке. При $\theta < 1$ превышения происходят группами (образуют кластеры), и величина $1/\theta$ аппроксимирует средний размер кластера превышения уровня. Существуют и другие интерпретации экстремального индекса [51, 113]. Например, θ представляет собой долю непустых промежутков между превышениями порогового значения [72]. Таким образом, с практической точки зрения экстремальный индекс полезен для оценки размера кластеров или средней длины промежутков между превышениями. Методы вычисления оценок экстремального индекса обсуждаются в работах [144, 166, 186]. Метод вычисления экстремального индекса положительно возвратной по Хар-

рису цепи Маркова, у которой маргинальное распределение принадлежит области притяжения распределения Гумбеля, приведен в [94]. Метод адаптирован для вычисления экстремального индекса стационарного времени ожидания в системах $GI/G/1$ и $GI/M/c$. Для систем, для которых нельзя найти численно или в явном виде экстремальный индекс в работе строится его регенеративная оценка. В связи с этим в Главе 6 приведен алгоритм регенеративной оценки экстремального индекса стационарного времени ожидания заявки в очереди системы обслуживания.

Проблема исследования экстремальных значений характеристик производительности систем массового обслуживания изучалась в работах [34, 35, 95, 140, 146]. В статье [112] получены распределения кластерных и межкластерных размеров стационарных времен ожидания в системе $GI/G/1$.

4.1. Асимптотическое распределение экстремальных значений

Рассмотрим последовательность н.о.р. с.в. $\{X_n, n \geq 1\}$, заданных общей ф.р. F , и обозначим через M_n максимум первых n величин

$$M_1 = X_1, \quad M_n = \max(X_1, \dots, X_n), \quad n \geq 2. \quad (4.1)$$

Ф.р. максимума M_n вычисляется по формуле:

$$P(M_n \leq x) = P\left(\bigcap_{i=1}^n \{X_i \leq x\}\right) = \prod_{i=1}^n P\{X_i \leq x\} = F^n(x), \quad x \in \mathbb{R}, \quad n \geq 1.$$

Экстремумы случаются «около» правой концевой точки (определение см. в разделе 1.1.) носителя распределения, поэтому, интуитивно, асимптотическое поведение M_n должно быть связано с правым хвостом ф.р. F вблизи ее правой концевой точки x_R [143].

Действительно, для любого $x < x_R$

$$P(M_n \leq x) = F^n(x) \rightarrow 0 \quad \text{при} \quad n \rightarrow \infty.$$

В случае, когда $x_R < \infty$, получим, что для $x \geq x_R$

$$\mathbf{P}(M_n \leq x) = \mathbf{F}^n(x) = 1, \text{ при любом } n \geq 1.$$

Следовательно,

$$M_n \xrightarrow{P} x_R \text{ при } n \rightarrow \infty, \quad (4.2)$$

где $x_R \leq \infty$. Поскольку последовательность $\{M_n, n \geq 1\}$ является неубывающей, то сходимость по вероятности (4.2) влечет сходимость с в.1 [5], а именно

$$\lim_{n \rightarrow \infty} M_n = x_R \text{ с в. 1.} \quad (4.3)$$

Определение 4.1. ([107]) *Распределения $\mathbf{F}_1$ и $\mathbf{F}_2$ относятся к одному типу, если они переводятся друг в друга линейным преобразованием аргумента, т. е. существуют такие константы $a > 0$ и b , что*

$$\mathbf{F}_1(x) = \mathbf{F}_2(ax + b)$$

и, соответственно,

$$\mathbf{F}_2(x) = \mathbf{F}_1\left(\frac{x - b}{a}\right).$$

Пусть с.в. ξ_i заданы ф.р. $\mathbf{F}_i, i = 1, 2$. Тогда с.в.

$$\xi_1 \stackrel{st}{=} \frac{\xi_2 - b}{a} \text{ и } \xi_2 \stackrel{st}{=} a\xi_1 + b$$

имеют ф.р. $\mathbf{F}_1$ и $\mathbf{F}_2$, соответственно.

Одним из наиболее важных результатов теории экстремальных значений является теорема Фишера-Типпетта-Гнеденко или «Теорема о типах экстремальных распределений», которая указывает возможные типы предельного распределения с.в. M_n .

Утверждение 4.1 (О типах экстремальных распределений). ([68]) *Пусть $\{X_n, n \geq 1\}$ – последовательность н.о.р. с.в., $M_n = \max(X_1, \dots, X_n), n \geq 1$. Если для некоторых последовательностей констант $a_n > 0$, b_n выполняется*

$$\lim_{n \rightarrow \infty} \mathbf{P}((M_n - b_n)/a_n \leq x) = G(x), \quad (4.4)$$

и G является невырожденной ф.р., то G принадлежит одному из трёх сле-

дующих классов ф.р. экстремальных значений:

$$\Phi_{\eta}(x) = \begin{cases} 0, & x < 0; \\ \exp(-x^{-\eta}), & x \geq 0, \end{cases} \quad \eta > 0, \quad (4.5)$$

$$\Psi_{\eta}(x) = \begin{cases} \exp(-(-x)^{\eta}), & x < 0; \\ 1, & x \geq 0, \end{cases} \quad \eta > 0, \quad (4.6)$$

$$\Lambda(x) = \exp(-e^{-x}), \quad x \in \mathbb{R}. \quad (4.7)$$

Верно и обратное утверждение: каждая ф.р. G , класс которой является одним из классов ф.р. экстремальных значений, может быть пределом для (4.4) и в действительности является таковым, когда G является ф.р. с.в. X .

Распределение Φ_{η} называется распределением Фреше, распределение Ψ_{η} называется инверсным распределением Вейбулла, распределение Λ называется распределением Гумбеля.

Распределения (4.5)-(4.7) можно записать в более общем виде так называемого обобщенного распределения экстремальных значений [129]:

$$P(X \leq x) = \begin{cases} \exp\left(-\left(1 + \frac{x - \mu}{\eta\sigma}\right)^{-\eta}\right), & \eta \neq 0; \\ \exp\left(-\exp\left(-\frac{x - \mu}{\sigma}\right)\right), & \eta = 0, \end{cases} \quad (4.8)$$

где $1 + \frac{x - \mu}{\eta\sigma} > 0$, параметры μ и σ являются параметрами центрирования и масштаба, соответственно. Параметр формы η называется *индексом экстремальных значений* и играет ключевую роль в поведении правого хвоста. При $\eta > 0$ получим распределение Фреше, при $\eta < 0$ – инверсное распределение Вейбулла, при $\eta = 0$ – распределение Гумбеля. Математическое ожидание и дисперсия обобщенного распределения экстремальных значений равны, соответственно,

$$MX = \mu + (\Gamma(1 - 1/\eta) - 1)\sigma\eta, \quad \eta > 1,$$

$$DX = (\Gamma(1 - 2/\eta) - \Gamma^2(1 - 1/\eta))\sigma^2\eta^2, \quad \eta > 2.$$

При $\eta = 0$

$$MX = \mu + \lambda_E\sigma, \quad DX = \sigma^2\pi^2/6,$$

где $\lambda_E \approx 0.5772..$ – константа Эйлера.

4.2. Области притяжения распределений экстремальных значений

При исследовании асимптотического поведения экстремумов важно знать какой именно из трех типов предельных законов возникает. Далее пойдет речь об условиях, при которых максимум последовательности н.о.р. с.в. с той или иной ф.р. относится, соответственно, к распределениям Гумбеля, Фреше или инверсного Вейбулла.

Определение 4.2. ([68]) *Если существуют последовательности констант $\{a_n > 0, n \geq 1\}, \{b_n, n \geq 1\}$ таких, что для невырожденного распределения G верно соотношение*

$$\lim_{n \rightarrow \infty} P((M_n - b_n)/a_n \leq x) = G(x), \quad (4.9)$$

то говорят, что ф.р. F принадлежит области притяжения максимума ф.р. G , что обозначается $F \in MDA(G)$.

Далее приведены известные результаты о принадлежности к области притяжения максимум-устойчивых распределений для всех трех типов.

Утверждение 4.2 (Область притяжения распределения Гумбеля). ([79]) *Пусть F абсолютно непрерывная ф.р., для которой существует такое z_0 , что F дважды дифференцируема по $x \in (z_0, x_R)$ и $F''(x) < 0$ для всех $z_0 < x < x_R$. Тогда, если*

$$\lim_{x \uparrow x_R} \frac{F''(x)(1 - F(x))}{(F'(x))^2} = -1, \quad (4.10)$$

то $F \in MDA(\Lambda)$. Нормализующие константы могут быть выбраны, исходя из условий

$$\bar{F}(b_n) = (1/n), \quad a_n = f_{\#}(b_n),$$

где $f_{\#}(x) = (1 - F(x))/F'(x)$.

Пример 4.1. Пусть с.в. X имеет распределение Вейбулла с ф.р.

$$F(x) = 1 - e^{-(x/\alpha)^\beta}, \quad \alpha > 0, \beta > 0, \quad x \geq 0. \quad (4.11)$$

Вычислим первую и вторую производные ф.р. $F(x)$

$$\begin{aligned} F'(x) &= \frac{\beta}{\alpha} \left(\frac{x}{\alpha}\right)^{\beta-1} e^{-(x/\alpha)^\beta}, \\ F''(x) &= \frac{\beta(\beta-1)}{\alpha^2} \left(\frac{x}{\alpha}\right)^{\beta-2} e^{-(x/\alpha)^\beta} - \frac{\beta^2}{\alpha^2} \left(\frac{x}{\alpha}\right)^{2\beta-2} e^{-(x/\alpha)^\beta} \\ &= F'(x) \frac{\beta - \beta(x/\alpha)^\beta - 1}{x}. \end{aligned}$$

Заметим, что $F''(x) < 0$ для $x > \alpha \left(\frac{\beta-1}{\beta}\right)^{1/\beta}$ при $\beta > 1$ и $F''(x) < 0$ для $x \geq 0$ при $0 < \beta \leq 1$. Поскольку

$$\lim_{x \rightarrow \infty} \frac{F''(x)(1 - F(x))}{(F'(x))^2} = \lim_{x \rightarrow \infty} \frac{\beta - \beta(x/\alpha)^\beta - 1}{x(\beta/\alpha)(x/\alpha)^{\beta-1}} = -1,$$

то условие (4.10) выполнено. Следовательно, условия Утверждения 4.2 также выполнены, и распределение Вейбулла относится к области притяжения распределения Гумбеля. Поскольку в данном случае

$$f_{\#}(x) = \frac{\alpha(x/\alpha)^{1-\beta}}{\beta},$$

то получим нормализующие константы в виде

$$\begin{aligned} b_n &= \alpha(\ln n)^{1/\beta}, \\ a_n &= f_{\#}(b_n) = \frac{\alpha(\ln n)^{1/\beta-1}}{\beta}, \quad n \geq 1. \end{aligned}$$

Утверждение 4.3. ([79, 86]) (*Область притяжения распределения Фреше*). Пусть правая конечная точка ф.р. F бесконечна, $x_R = \infty$, и существует константа $\eta > 0$ такая, что для каждого $x > 0$ выполнено предельное соотношение

$$\lim_{t \rightarrow \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-\eta}. \quad (4.12)$$

Тогда существует последовательность нормализующих констант $\{a_n > 0\}$ такая, что

$$\lim_{n \rightarrow \infty} P(M_n \leq a_n x) = \Phi_{\eta}(x).$$

При этом $\{a_n\}$ могут быть выбраны, исходя из условия

$$a_n = \inf\{x : 1 - F(x) \leq 1/n\}, \quad n \geq 1.$$

Пример 4.2. Пусть $X \sim \text{Pareto}(\alpha, x_0)$. Очевидно, соотношение (4.12) выполнено, т.к.

$$\lim_{t \rightarrow \infty} \frac{\bar{F}(tx)}{\bar{F}(t)} = \left(\frac{x_0}{x_0 + tx} \right)^\alpha \left(\frac{x_0 + t}{x_0} \right)^\alpha = \left(\frac{1}{x} \right)^\alpha = x^{-\alpha},$$

и нормализующие константы могут быть выбраны в виде

$$a_n = \bar{F}^{-1}(1/n) = \frac{x_0}{(1/n)^{1/\alpha}} - x_0 = x_0 n^{1/\alpha} - x_0, \quad n \geq 1.$$

Утверждение 4.4 (Область притяжения инверсного распределения Вейбулла). ([79]) Пусть правая конечная точка ф.р. F конечна, $x_R < \infty$ и для функции $F^*(x) = F(x_R - 1/x)$, $x > 0$ и некоторой константы $\eta > 0$ выполнено условие

$$\lim_{n \rightarrow \infty} \frac{1 - F^*(tx)}{1 - F^*(t)} = x^{-\eta}.$$

Тогда существуют константы $\{a_n > 0\}$ и $\{b_n\}$ такие, что

$$\lim_{n \rightarrow \infty} P(M_n \leq a_n x + b_n) = \Psi_\eta(x).$$

При этом нормализующие константы могут быть выбраны в виде

$$b_n = x_R, \quad a_n = x_R - \inf\{x : 1 - F(x) \leq 1/n\}, \quad n \geq 1.$$

Определение 4.3. ([68]) Две ф.р. F_X и F_Y имеют эквивалентные хвосты, если они имеют одинаковую правую конечную точку, $x_{R_X} = x_{R_Y} = x_R$, и выполнено соотношение

$$\lim_{x \uparrow x_R} \frac{\bar{F}_X(x)}{\bar{F}_Y(x)} = c, \tag{4.13}$$

где $c \in (0, \infty)$ – некоторая константа.

Согласно следующим трем утверждениям области притяжения максимума замкнуты для ф.р., имеющих эквивалентные хвосты распределений. Иными словами, если для F_X и F_Y выполнено соотношение (4.13), то $F_X \in MDA(G)$ тогда и только тогда, когда $F_Y \in MDA(G)$.

Утверждение 4.5. ([68]) Пусть F_X и F_Y – ф.р. с одинаковыми правыми конечными точками, $x_{R_X} = x_{R_Y} = x_R$, и $F_X \in MDA(\Lambda)$ с последовательностями нормализующих констант $\{a_n > 0\}$ и $\{b_n \in \mathbb{R}\}$, т. е.

$$\lim_{n \rightarrow \infty} F_X^n(a_n x + b_n) = \Lambda(x), \quad x \in \mathbb{R}.$$

Тогда

$$\lim_{n \rightarrow \infty} F_Y^n(a_n x + b_n) = \Lambda(x + c), \quad x \in \mathbb{R},$$

для некоторой константы $c > 0$ тогда и только тогда, когда

$$\lim_{x \uparrow x_R} \frac{\bar{F}_X(x)}{\bar{F}_Y(x)} = e^c. \quad (4.14)$$

Утверждение 4.6. ([68]) Пусть F_X и F_Y – ф.р., и $F_X \in MDA(\Phi_\eta)$ с нормализующими константами $\{a_n\}$, т. е.

$$\lim_{n \rightarrow \infty} F_X^n(a_n x) = \Phi_\eta(x), \quad x > 0.$$

Тогда

$$\lim_{n \rightarrow \infty} F_Y^n(a_n x) = \Phi_\eta(cx), \quad x > 0,$$

для некоторой константы $c > 0$ тогда и только тогда, когда

$$\lim_{x \rightarrow \infty} \frac{\bar{F}_X(x)}{\bar{F}_Y(x)} = c^\eta. \quad (4.15)$$

Утверждение 4.7. ([68]) Пусть F_X и F_Y – ф.р. с одинаковыми конечными правыми конечными точками, $x_{R_X} = x_{R_Y} = x_R < \infty$, и $F_X \in MDA(\Psi_\eta)$ с нормирующими константами a_n , т. е.

$$\lim_{n \rightarrow \infty} F_X^n(a_n x + x_R) = \Psi_\eta(x), \quad x < 0.$$

Тогда

$$\lim_{n \rightarrow \infty} F_Y^n(a_n x + x_R) = \Psi_\eta(cx), \quad x < 0,$$

для некоторой константы $c > 0$ тогда и только тогда, когда

$$\lim_{x \uparrow x_R} \frac{\bar{F}_X(x)}{\bar{F}_Y(x)} = c^{-\eta}. \quad (4.16)$$

Утверждения 4.5, 4.6, 4.7 использованы в Главе 5 для вывода предельного распределения максимума стационарного времени ожидания с учетом эквивалентности хвоста его ф.р. и хвоста ф.р. равновесного времени обслуживания для субэкспоненциальных распределений.

4.3. Линейные преобразования над максимумами

Пусть $v \in (0, \infty)$ и предположим, что существует такая последовательность вещественных чисел $\{u_n, n \geq 1\}$, что

$$n\bar{F}(u_n) \rightarrow v \text{ при } n \rightarrow \infty. \quad (4.17)$$

Тогда [107]

$$P(M_n \leq u_n) \rightarrow e^{-v} \text{ при } n \rightarrow \infty \quad (4.18)$$

и, наоборот, условие (4.18) (для некоторого $0 \leq v \leq \infty$) влечет выполнение (4.17).

В случае выполнения соотношения (4.17), сходимость (4.18) сохраняется для членов семейства линейных нормализующих последовательностей

$$u_n(x) = a_n x + b_n, \quad n \geq 1, \quad x \in \mathbb{R},$$

и (4.18) преобразуется в сходимость

$$P(M_n \leq u_n(x)) \rightarrow v(x),$$

где вид функции $v(x)$ зависит от типа предельного распределения, а именно: для распределения Гумбеля $v(x) = e^{-x}$, для распределения Фреше $v(x) = x^{-\eta}$ и для инверсного распределения Вейбулла $v(x) = (-x)^\eta$.

Приведем примеры линейных нормализующих последовательностей $\{u_n(x)\}$ для некоторых распределений.

Пример 4.3. Показательное распределение $Exp(\lambda)$ с нормализующей последовательностью

$$u_n(x) = \frac{1}{\lambda}(x + \ln n), \quad n \geq 1, \quad (4.19)$$

относится к области притяжения распределения Гумбеля. Действительно,

$$n\bar{F}(u_n(x)) = ne^{-\lambda \frac{x + \ln n}{\lambda}} = n \frac{e^{-x}}{n} = e^{-x}.$$

Пример 4.4. Распределение Парето вида

$$F(x) = 1 - \left(\frac{x_0}{x_0 + x} \right)^\alpha, \quad \alpha > 0, \quad x_0 > 0, \quad x \geq 0, \quad (4.20)$$

относится к области притяжения распределения Фреше, при этом в качестве нормализующей последовательности можно выбрать

$$u_n(x) = x_0 n^{1/\alpha} x - x_0, \quad n \geq 1. \quad (4.21)$$

Действительно,

$$\begin{aligned} n\bar{F}(u_n(x)) &= n \frac{x_0^\alpha}{(x_0 + (x_0 n^{1/\alpha} - x_0)x)^\alpha} \\ &= \frac{n}{n \left(\frac{1}{n^{1/\alpha}} + x - \frac{x}{n^{1/\alpha}} \right)^\alpha} \rightarrow x^{-\alpha} \text{ при } n \rightarrow \infty. \end{aligned}$$

Пример 4.5. Как показано ранее, распределение Вейбулла (4.11) с нормализующей последовательностью

$$u_n(x) = \frac{x \alpha (\ln n)^{1/\beta-1}}{\beta} + \alpha (\ln n)^{1/\beta}, \quad n \geq 1, \quad (4.22)$$

принадлежит области притяжения распределения Гумбеля. Условие (4.17) при выборе (4.22) выполнено, так как

$$\begin{aligned} n\bar{F}(u_n) &= ne^{-\left(\frac{\alpha (\ln n)^{1/\beta-1}}{\beta} x + \alpha (\ln n)^{1/\beta} \right)^\beta / \alpha^\beta} \\ &= ne^{-\ln n \left(1 + \frac{x}{\ln n} + o\left(\frac{1}{\ln n} \right) \right)} \rightarrow e^{-x} \text{ при } n \rightarrow \infty. \end{aligned} \quad (4.23)$$

Пример 4.6. Пусть $X \sim Exp - Pareto(\lambda, \alpha, x_0, p)$ с хвостом ф.р.

$$\bar{F}_X(x) = pe^{-\lambda x} + (1-p) \left(\frac{x_0}{x_0 + x} \right)^\alpha, \quad x \geq 0. \quad (4.24)$$

Чтобы проверить, что (4.24) принадлежит области притяжения распределения

Фреше достаточно показать, что для функции интенсивности отказов r_X выполнено условие $x \cdot r_X(x) \rightarrow \text{const} > 0$ при $x \rightarrow \infty$ [68]. Действительно,

$$x r_X(x) = \frac{p x \lambda a(x) + (1-p) x \alpha / (x_0 + x)}{p a(x) + (1-p)} \rightarrow \alpha \quad \text{при } x \rightarrow \infty, \quad (4.25)$$

где

$$a(x) = e^{-\lambda x} \left(1 + \frac{x}{x_0}\right)^\alpha. \quad (4.26)$$

Проверим условие (4.17) для нормализующей последовательности

$$u_n(x) = a_n x + b_n = x_0 x n^{1/\alpha} - x_0, \quad n \geq 1. \quad (4.27)$$

Подставив (4.27) в (4.17), получим

$$\begin{aligned} n \bar{F}(u_n(x)) &= n \left[p e^{-\lambda u_n(x)} + \frac{(1-p)x_0^\alpha}{(x_0 + u_n(x))^\alpha} \right] \\ &= n \left[p e^{-\lambda u_n(x)} + \frac{(1-p)x_0^\alpha}{(x_0 n^{1/\alpha} x)^\alpha} \right] \\ &= \frac{n(1-p)x^{-\alpha}}{n} \left[\frac{p n e^{-\lambda u_n(x)} x^\alpha}{1-p} + 1 \right] \\ &\rightarrow (1-p)x^{-\alpha}, \end{aligned}$$

при $n \rightarrow \infty$, поскольку

$$u_n(x) \rightarrow \infty, \quad n e^{-\lambda u_n(x)} \rightarrow 0.$$

Следовательно, показательное-Парето распределение принадлежит области притяжения максимума распределения Фреше.

В следующей Теореме 4.1 пример 4.6 обобщен на случай многомерной смеси.

Теорема 4.1. Пусть последовательность н.о.р. с.в. $X_1, \dots, X_n$ имеет $(k+m)$ -показательное-Парето распределение ($m > 1$)

$$F_X(x) = 1 - \sum_{i=1}^k p_i e^{-\lambda_i x} - \sum_{j=k+1}^{k+m} p_j \left(\frac{x_0}{x_0 + x} \right)^{\alpha_j}, \quad x \geq 0, \quad x_0 > 0, \quad \alpha_j > 2, \quad \lambda_i > 0,$$

где $\sum_{i=1}^{k+m} p_i = 1$. Обозначим $M_n = \max(X_1, \dots, X_n)$ и выберем

$$a_n = x_0 n^{1/\alpha^*}, \quad b_n = -x_0, \quad \alpha^* = \min(\alpha_{k+1}, \dots, \alpha_{k+m}), \quad n \geq 2.$$

Тогда

$$F_X \in MDA(\Phi_{\alpha^*}).$$

Доказательство. Подставим $u_n(x) = x_0 n^{1/\alpha^*} x - x_0$ в соотношение (4.17) ($x > 0$), получим

$$\begin{aligned} n\bar{F}(u_n(x)) &= n \left[\sum_{i=1}^k p_i e^{-\lambda_i u_n(x)} + \sum_{j=k+1}^{k+m} p_j \left(\frac{x_0}{x_0 + u_n(x)} \right)^{\alpha_j} \right] \\ &= n(x_0 + u_n(x))^{-\alpha^*} p^* x_0^{\alpha^*} \\ &\quad \times \left[\sum_{i=1}^k \frac{p_i e^{-\lambda_i u_n(x)} (x_0 x n^{1/\alpha^*})^{\alpha^*}}{p^* x_0^{\alpha^*}} + \sum_{j=k+1, j \neq j^*}^{k+m} \frac{p_j x_0^{\alpha_j} (x_0 + u_n(x))^{\alpha^*}}{p^* x_0^{\alpha^*} (x_0 + u_n(x))^{\alpha_j}} + 1 \right] \\ &= \frac{n x^{-\alpha^*} p^*}{n} \left[\sum_{i=1}^k \frac{p_i e^{-\lambda_i u_n(x)} x^{\alpha^*} n}{p^*} + \sum_{j=k+1, j \neq j^*}^{k+m} \frac{p_j x^{\alpha^* - \alpha_j}}{p^* n^{\alpha_j/\alpha^* - 1}} + 1 \right]. \end{aligned}$$

Тогда

$$n\bar{F}(u_n(x)) \rightarrow p^* x^{-\alpha^*} \quad \text{при } n \rightarrow \infty,$$

поскольку

$$n^{\alpha_j/\alpha^* - 1} \rightarrow \infty \quad \text{и} \quad n e^{-\lambda_i u_n(x)} \rightarrow 0.$$

Далее, ввиду (4.18), получаем следующее асимптотическое распределение максимума для $k + m$ -показательного-Парето распределения:

$$P(M_n \leq u_n(x)) \rightarrow e^{-p^* x^{-\alpha^*}} \quad \text{при } n \rightarrow \infty, \quad (4.28)$$

т. е. $(M_n - b_n)/a_n$ сходится к распределению Фреше. □

4.4. Асимптотическое поведение функции распределения величины превышения уровня t

Пусть с.в. X задана ф.р. F_X , а F_{X_t} – ф.р. величины превышения уровня t с.в. $X_t = X - t | X > t$ (см. определение 1.2, раздел 1.1.) Между областями

притяжения нормализованного максимума $(M_n - b_n)/a_n$ и нормализованной величины $(X_t - b(t))/a(t)$ превышения уровня t существует взаимосвязь [45]. Для начала приведем возможные предельные распределения величины X_t превышения уровня t .

Утверждение 4.8. ([45]) Пусть F – ф.р. с.в. X и $F(x) < 1$ для всех x . Предположим, что для каждого $t > 0$ существуют такие функции $a(t) > 0$ и $b(t)$, что при каждом x

$$\lim_{t \rightarrow \infty} F_t(a(t)x + b(t)) = G(x), \quad (4.29)$$

где G – невырожденная ф.р. Тогда G относится к одному из следующих четырех типов распределений:

$$\Gamma_\eta(x) = 1 - (1 + x)^{-\eta}, \quad x \geq 0, \quad (4.30)$$

$$\Pi(x) = 1 - e^{-x}, \quad x \geq 0, \quad (4.31)$$

$$\Gamma_{\gamma, \eta}(x) = \exp(-\gamma[1 + \eta \ln(1 + x)]), \quad x \geq 0, \quad (4.32)$$

$$\Pi_\gamma(x) = 1 - \exp(-\gamma[1 + x]), \quad x \geq 0, \quad (4.33)$$

где $\eta > 0, \gamma > 0$ и $[a]$ – целая часть a .

Будем говорить, что распределение F_t величины превышения уровня t принадлежит области притяжения распределения G , т. е. $F_t \in DA(G)$, если выполнено (4.29). Обозначим также через D_0 множество таких ф.р., что $F(x) < 1$ для всех x .

Утверждение 4.9. ([45])

$$DA(\Pi) = MDA(\Lambda) \cap D_0,$$

$$DA(\Gamma_\eta) = MDA(\Phi_\eta) \quad \text{для всех } \eta > 0.$$

Таким образом, нормализованный максимум $(M_n - b_n)/a_n$ имеет в пределе распределение Фреше Φ_η тогда и только тогда, когда нормализованная величина превышения уровня t имеет предельное распределение (4.30). Аналогично, нормализованный максимум $(M_n - b_n)/a_n$ имеет в пределе распределение Гумбея Λ тогда и только тогда, когда нормализованная величина превышения уровня t имеет предельное распределение (4.31).

Пример 4.7. В примере 4.4 показано, что распределение нормализованного максимума распределения Парето (4.20) сходится к распределению Фреше, Φ_α . Покажем, что нормализованная величина превышения уровня t для этого распределения сходится к Γ_α (т. е. $\eta = \alpha$ в (4.30)). Действительно,

$$F_t(x) = 1 - \frac{(x_0 + t)^\alpha}{(x_0 + t + x)^\alpha}, \quad x \geq 0, t > 0.$$

Пусть $a(t) = t, b(t) = -x_0$. Тогда

$$\lim_{t \rightarrow \infty} F_t(a(t)x + b(t)) = \lim_{t \rightarrow \infty} \left(1 - \frac{(x_0 + t)^\alpha}{(x_0 + t + tx - x_0)^\alpha} \right) = 1 - (1 + x)^{-\alpha} = \Gamma_\alpha(x),$$

что и требовалось.

4.5. Экстремальное поведение строго стационарных последовательностей

Теорема об экстремальных типах распределений применима к *строго стационарной последовательности* зависимых с.в. при определенных условиях на перемешивание [107].

Предположим, что члены последовательности $\{X_n, n \geq 1\}$ зависимы, одинаково распределены и образуют стационарную в узком смысле последовательность, т. е. удовлетворяют тождеству [8]:

$$(X_{k+1}, \dots, X_{k+n}) \stackrel{st}{=} (X_1, \dots, X_n) \text{ для всех } k, n \geq 1.$$

Зависимость членов последовательности определяется следующим условием перемешивания [107]. Будем говорить, что для заданной последовательности вещественных чисел $\{u_n = a_n x + b_n, a_n > 0\}$, выполнено условие перемешивания $D(u_n)$, если для любых целых чисел

$$1 \leq i_1 < \dots < i_k < j_1 < \dots < j_m \leq n,$$

таких, что $j_1 - i_k \geq l$, выполняется неравенство

$$|F_{i_1, \dots, i_k, j_1, \dots, j_m}(u_n) - F_{i_1, \dots, i_k}(u_n)F_{j_1, \dots, j_m}(u_n)| \leq v_{n,l}, \quad (4.34)$$

где $v_{n,l} \rightarrow 0$ при $n \rightarrow \infty$, $l \rightarrow \infty$ и для совместной ф.р.

$$F_{i_1, \dots, i_k}(x_1, \dots, x_k) = P(X_1 \leq x_1, \dots, X_k \leq x_k)$$

использовано обозначение $F_{i_1, \dots, i_k}(u_n) := F_{i_1, \dots, i_k}(u_n, \dots, u_n)$ для каждого n (т. е. все $x_i = u_n$). Заметим, что расстояние l между компонентами зависит от общего числа компонент n , и что скорость убывания $v_{n,l}$ в (4.34), вообще говоря, зависит как от длины серии n так и от расстояния l . В этой связи приведем условие ϕ -перемешивания последовательности $\{X_n, n \geq 1\}$ (условие сильного перемешивания, см. [53]), которое формулируется следующим образом.

Определение 4.4. Пусть $\mathcal{F}_1^n$ – σ -поле, порожденное с.в. $X_1, \dots, X_n$, а $\mathcal{F}_{n+l}^\infty$ – σ -поле, порожденное с.в. $X_{n+l}, X_{n+l+1}, \dots$. Последовательность с.в. $\{X_n, n \geq 1\}$ обладает условием ϕ -перемешивания, если при $l \rightarrow \infty$

$$\phi(l) := \sup_n \sup_{A \in \mathcal{F}_1^n} \sup_{B \in \mathcal{F}_{n+l}^\infty} |P(A \cap B) - P(A)P(B)| \rightarrow 0.$$

Поскольку в последнем условии сходимость равномерна по n , то оно, вообще говоря, сильнее (более ограничительно), чем условие $D(u_n)$. Отметим также, что условия перемешивания интерпретируются как сходимость к нулю ковариаций, поскольку $P(A \cap B) - P(A)P(B) = \text{cov}(I_A, I_B)$.

Утверждение 4.10. ([107]) Предположим, что для некоторого $v > 0$ определена такая последовательность $\{u_n(v)\}$, что $n\bar{F}(u_n(v)) \rightarrow v$ при $n \rightarrow \infty$, и условие $D(u_n(v))$ выполнено для каждого такого v . Тогда

$$P(M_n \leq u_n(v)) \rightarrow e^{-\theta v} \text{ при } n \rightarrow \infty, \quad (4.35)$$

где параметр $\theta \in [0, 1]$.

Заметим, что если $P(M_n \leq u_n(v))$ сходится хотя бы для одного значения $v > 0$, то (4.35) выполнено при всех $v > 0$ для некоторого фиксированного θ , $0 \leq \theta \leq 1$ [107].

Определение 4.5. ([107]) Строго стационарная последовательность $\{X_n, n \geq 1\}$ с общей ф.р. F имеет экстремальный индекс θ ($\theta \in [0, 1]$), если для каждого $v > 0$ существует такая последовательность вещественных чисел $\{u_n = u_n(v)\}$, что выполнены соотношения (4.17) и (4.35).

Из определения 4.5 и утверждения 4.10 следует, что параметр θ для стационарной последовательности $\{X_n, n \geq 1\}$ можно вычислить по формуле [68]:

$$\theta = \lim_{n \rightarrow \infty} \frac{\ln P(M_n \leq u_n)}{n \ln F(u_n)}. \quad (4.36)$$

Экстремальный индекс является важной характеристикой стационарных последовательностей, поскольку выступает мерой кластеризации больших значений элементов последовательности [111]. В частности, чем меньше θ , тем больше размер кластера.

Легко увидеть, что для независимых с.в. экстремальный индекс равен единице.

Известно, что стационарная последовательность и ассоциированная с ней последовательность независимых с.в. имеют предельные распределения максимумов одного и того же типа [107]. А именно, пусть $\{\hat{X}_n, n \geq 1\}$ — последовательность независимых с.в., заданных такой же ф.р. F , что и исходная стационарная последовательность $\{X_n, n \geq 1\}$. Обозначим $\hat{M}_n = \max(\hat{X}_1, \dots, \hat{X}_n)$. Если $P(\hat{M}_n \leq u_n(x)) \rightarrow H(x)$ и $H(x)$ — невырожденная ф.р., то

$$P(M_n \leq u_n(x)) \rightarrow H(ax + b) \text{ при } n \rightarrow \infty, \quad (4.37)$$

где $u_n(x) = a_n x + b_n$ и

$$H(ax + b) = [H(x)]^\theta$$

для некоторых $a > 0$ и b , θ — экстремальный индекс. Более того,

$$P(M_n \leq \hat{u}_n(x)) \rightarrow H(x) \text{ при } n \rightarrow \infty,$$

где $\hat{u}_n(x) = a_n/ax + b_n - ab/a_n$. Следовательно, зная $\lim_{n \rightarrow \infty} P(\hat{M}_n \leq u_n(x))$ и экстремальный индекс θ , можно получить предельное распределение нормированного максимума M_n исходной стационарной последовательности, $\lim_{n \rightarrow \infty} P(M_n \leq \hat{u}_n(x))$. Таким образом, экстремальный индекс θ является ключевым параметром для нахождения асимптотических распределений максимумов стационарных последовательностей.

4.6. Экстремальное поведение в системах $GI/G/1/\infty$

Рассмотрим систему $GI/G/1/\infty$. Обозначим $\{S_i, i \geq 1\}$, $\{\tau_i, i \geq 1\}$ – н.о.р. времена обслуживания и н.о.р. интервалы между приходами заявок, соответственно. Рассмотрим рекурсию Линдли

$$W_{i+1} = (W_i + X_i)^+, \quad i \geq 1,$$

где $(x)^+ = \max(x, 0)$ и $X_i = S_i - \tau_i$ – н.о.р. нерешетчатые с.в. (то есть не существует такого числа $d > 0$, что выполняется равенство $P\left(\sum_{j=1}^{\infty} \{X = a + j \cdot d\}\right) = 1$, где a – любое число) с невырожденной ф.р. F и $MX < 0$. Также предположим, что существует число γ такое, что $M[Xe^{\gamma X}] < \infty$ и

$$Me^{\gamma X} = 1. \quad (4.38)$$

Пусть $Y_n = X_1 + \dots + X_n$, $n \geq 1$ ($Y_0 = 0$). Обозначим через W_n время ожидания начала обслуживания n -й заявки, пришедшей в систему в момент времени t_n . Процесс $\{W_n, n \geq 1\}$ это процесс Линдли, порождаемый случайным блужданием $\{Y_n, n \geq 0\}$ [34]. В частности, $W_n = \max_{0 \leq k \leq n} Y_k$ и, если система стационарна, т. е. $\rho = MS/M\tau < 1$, то существует предел для последовательности $\{W_n, n \geq 1\}$

$$W_n \Rightarrow W, \quad n \rightarrow \infty.$$

Известно [34, 95], что если система стационарна и уравнение (4.38) имеет действительное положительное решение $\gamma \neq 0$ [174], то

1) если $M[Xe^{\gamma X}] < \infty$, то хвост распределения $\sup_{n \geq 0} Y_n$ асимптотически (с точностью до некоторой константы $K > 0$) эквивалентен экспоненциальной функции, а именно

$$\lim_{x \rightarrow \infty} \frac{P(\max_{n \geq 0} Y_n > x)}{e^{-\gamma x}} = K.$$

2) если $M[Xe^{\gamma X}] = \infty$, то $\lim_{x \rightarrow \infty} P(\max_{n \geq 0} Y_n > x) = o(e^{-\gamma x})$, при $x \rightarrow \infty$.

Обозначим через $M_n^W = \max(W_1, \dots, W_n)$ максимальное время ожидания среди $1, \dots, n$ -й заявок. Если $\rho \leq 1$ и существует действительное положитель-

ное решение уравнения (4.38), то [63, 95]

$$\lim_{n \rightarrow \infty} \mathbf{P}(\gamma M_n^W - \ln(b\theta_W n) \leq x) = \Lambda(x), \quad (4.39)$$

где $\Lambda(x) = \exp(-e^{-x})$ – распределение Гумбеля и b – константа. Более того, $M_n^W / \ln n$ сходится по вероятности к $1/\gamma$ [63]:

$$\frac{M_n^W}{\ln n} \xrightarrow{p} 1/\gamma \text{ при } n \rightarrow \infty.$$

Для системы $GI/M/1$ при условии стационарности $\rho < 1$, уравнение (4.38) можно переписать в виде

$$\mathbf{M}e^{-\gamma\tau} = \frac{\mu - \gamma}{\mu}, \quad (4.40)$$

и оно имеет единственное положительное решение γ [34]. Ф.р. стационарного времени ожидания является смесью атома в нуле и показательного распределения с параметром γ и коэффициентами пропорции смеси γ/μ и $1 - \gamma/\mu$, соответственно, [34][Теорема 5.1]:

$$\mathbf{F}_W(x) = 1 - (1 - \gamma/\mu)e^{-\gamma x}, \quad x > 0. \quad (4.41)$$

Более того, из Теоремы 7.5 [63] следует, что экстремальный индекс стационарного времени ожидания для системы $GI/M/1$ вычисляется по формуле

$$\theta_W = \gamma \left(\frac{d}{d\gamma} \psi_\tau(\gamma) + \frac{1}{\mu} \right), \quad (4.42)$$

где $\psi_\tau(\gamma) = \mathbf{M}e^{-\gamma\tau}$ – ПЛС интервалов между приходами заявок, τ , и, кроме того, $b = 1 - \gamma/\mu$ в (4.39). Найдём по ф.р. (4.41) ПЛС для W

$$\psi_W(z) = \frac{\gamma}{\mu} + \int_{0+}^{\infty} e^{-zx} d\mathbf{F}_W(x) = \frac{\gamma}{\mu} + \frac{\gamma(1 - \gamma/\mu)}{\gamma + z}.$$

Поскольку

$$\mathbf{M}W = -\psi'_W(0) = \frac{1 - \gamma/\mu}{\gamma},$$

то легко получить выражение для корня γ через $\mathbf{M}W$:

$$\gamma = \frac{\mu}{\mu \mathbf{M}W + 1}. \quad (4.43)$$

Для стационарной системы $M/G/1$ уравнение (4.38) имеет следующий вид [34]:

$$Me^{\gamma S} = 1 + \frac{\gamma}{\lambda}. \quad (4.44)$$

Более того, если уравнение (4.44) имеет решение γ , то экстремальный индекс θ_W вычисляется по формуле [94]:

$$\theta_W = \frac{\gamma(1 - \rho)}{\gamma + \lambda}. \quad (4.45)$$

Пример 4.8. Можно проверить, что в системе $M/M/1$ соотношение (4.39) выполняется при $b = \rho = \lambda/\mu$, $\gamma = \mu - \lambda$, и что

$$\theta_W = (1 - \rho)^2. \quad (4.46)$$

Пример 4.9. В системе $M/H_m/1$ с гиперэкспоненциальным распределением времени обслуживания с параметрами $p_i, \mu_i, i = 1, \dots, m$ и показательным распределением входного потока $Exp(\lambda)$ уравнение (4.44) для нахождения корней принимает вид

$$\sum_{i=1}^m \frac{p_i \mu_i}{\mu_i - \gamma} = \frac{\lambda + \gamma}{\lambda}. \quad (4.47)$$

Уравнение (4.47) может быть решено численно. Например, при $\lambda = 1, \mu_1 = 4, \mu_2 = 2, p = 0.5, m = 2$, численное решение дает два положительных корня

$$\gamma_1 = 1.38, \gamma_2 = 3.62,$$

которым соответствуют два значения экстремального индекса

$$\theta_{W_1} = 0.36, \theta_{W_2} = 0.49.$$

Согласно [63] нужно выбрать наименьший корень $\gamma_1 = 1.38$, что дает $\theta_{W_1} = 0.36$. Результаты регенеративного оценивания экстремального индекса также подтверждают, что $\theta_W = 0.36$ (см. раздел 4.9.1.)

Пример 4.10. Экспоненциальные моменты $Me^{\gamma S}$ для распределения Вейбулла $We(\alpha, \beta)$ существуют только при $\beta \geq 1$, поэтому для системы $M/We/1$ значе-

ния γ найдем из уравнения

$$\sum_{k=0}^{\infty} \frac{(\alpha\gamma)^k}{k!} \Gamma\left(\frac{k}{\beta} + 1\right) = 1 + \frac{\gamma}{\lambda}, \text{ при } \beta \geq 1. \quad (4.48)$$

Например, при $\lambda = 2, \alpha = 0.25, \beta = 1.5$, численно получаем единственный корень $\gamma = 3.7$, следовательно, $\theta_W = 0.35$. При параметрах $\lambda = 2, \alpha = 0.5, \beta = 1.5$, получаем $\gamma = 0.29, \theta_W = 0.012$, а при $\lambda = 2, \alpha = 0.4, \beta = 1.2$ получаем $\gamma = 0.79$ и $\theta_W = 0.07$.

При $\beta < 1$ оценку экстремального индекса возможно получить с помощью регенеративного моделирования, что будет показано в Главе 5.

Пример 4.11. Для системы $H_m/M/1$ с гиперэкспоненциальным распределением входного потока с параметрами $p_i, \mu_i, i = 1, \dots, m$ и показательным распределением времени обслуживания $Exp(\mu)$ при подстановке параметров в уравнение $Me^{\gamma X} = 1$ получаем следующее уравнение для нахождения корней:

$$\sum_{i=1}^m \frac{p_i \lambda_i}{\lambda_i + \gamma} = \frac{\mu - \gamma}{\mu}. \quad (4.49)$$

Подставив величину

$$\frac{d}{d\gamma} \psi_\tau(\gamma) = - \int_0^\infty x e^{-\gamma x} dA(x) = - \sum_{i=1}^m \frac{p_i \lambda_i}{(\lambda_i + \gamma)^2}$$

в (4.42), получим выражение для экстремального индекса

$$\theta_W = \gamma \left(\frac{1}{\mu} - \sum_{i=1}^m \frac{p_i \lambda_i}{(\lambda_i + \gamma)^2} \right).$$

Например, единственное численное положительное решение при $p = 0.5, \lambda_1 = 1, \lambda_2 = 3, \mu = 4, m = 2$ получим при $\gamma = 2.24$, при котором $b = 1 - \gamma/\mu = 0.44, \theta = 0.33$. Более того, нормализующая последовательность для предельного распределения максимума стационарного времени ожидания в данном примере имеет вид

$$u_n(x) = (x + \ln(b\theta_W n))/\gamma = (x + \ln(0.145n))/2.24, \quad n \geq 1.$$

4.7. Монотонность систем $GI/G/1/\infty$ по экстремальному индексу

Рассмотрим две системы обслуживания $\Sigma^{(1)}$ и $\Sigma^{(2)}$ вида $GI/G/1$. (Будем снабжать индексом i величины, относящиеся к i -й системе.) В обеих системах дисциплина обслуживания FIFS. Обозначим $S_n^{(i)}$ – время обслуживания n -й заявки и $\tau_n^{(i)}$ – интервал между приходами n -й и $n+1$ -й заявок в i -ю систему, $M\tau^{(i)} = 1/\lambda$, $i = 1, 2$. Пусть $\nu_n^{(i)}$ – число заявок в системе, $Q_n^{(i)}$ – размер очереди и $W_n^{(i)}$ – время ожидания в очереди в момент прихода в систему $\Sigma^{(i)}$ заявки с номером n , $i = 1, 2$. Обозначим (в случае их существования) пределы по распределению

$$Q_n^{(i)} \Rightarrow Q^{(i)}, \nu_n^{(i)} \Rightarrow \nu^{(i)}, W_n^{(i)} \Rightarrow W^{(i)}, n \rightarrow \infty, i = 1, 2.$$

Как уже не раз отмечалось, эти пределы существуют, в частности, если времена между приходами заявок $\tau^{(i)}$ являются нерешетчатыми и $\rho_i = \lambda_i M S^{(i)} < 1$, $i = 1, 2$ [34].

Сравним в системах $\Sigma^{(1)}$ и $\Sigma^{(2)}$ экстремальные индексы стационарных длин очередей и времен ожидания. Обозначим максимумы

$$M_n^{Q^{(1)}} = \max(Q_1^{(1)}, \dots, Q_n^{(1)}), M_n^{Q^{(2)}} = \max(Q_1^{(2)}, \dots, Q_n^{(2)}), n \geq 1,$$

$$M_n^{W^{(1)}} = \max(W_1^{(1)}, \dots, W_n^{(1)}), M_n^{W^{(2)}} = \max(W_1^{(2)}, \dots, W_n^{(2)}), n \geq 1.$$

Предположение 4.1. *Предположим, что в системах $\Sigma^{(1)}$ и $\Sigma^{(2)}$ коэффициенты загрузки $\rho_i < 1$, $i = 1, 2$, и существуют предельные распределения для максимумов стационарных времен ожидания $M_n^{W^{(1)}}$, $M_n^{W^{(2)}}$ и стационарных длин очередей $M_n^{Q^{(1)}}$, $M_n^{Q^{(2)}}$ при $n \rightarrow \infty$. Пусть $\theta_{Q^{(i)}}, \theta_{W^{(i)}}$ – экстремальные индексы $Q^{(i)}$, $W^{(i)}$, $i = 1, 2$, соответственно. Тогда*

1) *Если выполнены соотношения*

$$\nu_1^{(1)} = \nu_1^{(2)} = 0, \tau_{st}^{(1)} = \tau_{st}^{(2)}, S_r^{(1)} \leq S_r^{(2)}, \quad (4.50)$$

то

$$\theta_{Q^{(1)}} \geq \theta_{Q^{(2)}}. \quad (4.51)$$

2) Если выполнены соотношения:

$$W_1^{(1)} = W_1^{(2)} = 0, \quad \tau^{(1)} \underset{r}{\geq} \tau^{(2)}, \quad S^{(1)} \underset{r}{\leq} S^{(2)}, \quad (4.52)$$

то

$$\theta_{W^{(1)}} \geq \theta_{W^{(2)}}. \quad (4.53)$$

Следующая теорема утверждает, что Предположение 4.1 выполнено для систем $M/G/1$ при условии существования действительных положительных корней уравнения (4.38).

Теорема 4.2. Пусть для систем $\Sigma^{(1)}$ и $\Sigma^{(2)}$ вида $M/G/1$ выполнены условия стационарности, $\rho_i < 1, i = 1, 2$, $M[S^{(i)}e^{\gamma S^{(i)}}] < \infty, i = 1, 2$, а также выполнены соотношения (4.52). Пусть существуют действительные положительные корни уравнения (4.38). Тогда экстремальные индексы стационарных времен ожидания упорядочены как

$$\theta_{W^{(1)}} \geq \theta_{W^{(2)}}.$$

Доказательство. Прежде всего заметим, что из упорядоченности с.в. по интенсивности отказов следует упорядоченность экспоненциальных моментов [116]. Следовательно, из соотношений (4.52) вытекает, что

$$Me^{\gamma S^{(1)}} \leq Me^{\gamma S^{(2)}}, \quad Me^{-\gamma \tau^{(1)}} \leq Me^{-\gamma \tau^{(2)}}, \quad \text{для любого } \gamma > 0.$$

Обозначим через γ_i положительный действительный корень уравнения (4.38) для системы $\Sigma^{(i)}, i = 1, 2$. Если таких корней несколько, то согласно Теореме 7.2 из [63], выбирается наименьший из них. Тогда равенства

$$Me^{\gamma_1(S^{(1)} - \tau^{(1)})} = Me^{\gamma_2(S^{(2)} - \tau^{(2)})} = 1$$

выполнены, если

$$\gamma_1 \geq \gamma_2.$$

Поскольку интервалы между приходами заявок упорядочены по интенсивности отказов, $\tau^{(1)} \underset{r}{\geq} \tau^{(2)}$, то $\lambda_1 \leq \lambda_2$. Следовательно, $\lambda_1/\gamma_1 \leq \lambda_2/\gamma_2$. Более того, в силу $S^{(1)} \underset{r}{\leq} S^{(2)}$, имеем также $MS^{(1)} \underset{r}{\leq} MS^{(2)}$ и $1 - \rho_1 \geq 1 - \rho_2$. Подставив эти нера-

венства в выражение для экстремального индекса (4.45), получим утверждение теоремы, а именно,

$$\theta_{W^{(1)}} = \frac{1 - \rho_1}{1 + \lambda_1/\gamma_1} \geq \frac{1 - \rho_2}{1 + \lambda_2/\gamma_2} = \theta_{W^{(2)}}.$$

□

Пример 4.12. Для системы $M/M/1$, как уже было сказано выше, экстремальный индекс стационарного времени ожидания имеет вид (4.46). Рассмотрим две системы вида $M/M/1$, в которых времена обслуживания показательные с параметрами μ_1 и μ_2 , соответственно, и $\mu_1 > \mu_2 > 0$. Предположим, что $\lambda_1 < \lambda_2$ и $\rho_i = \lambda_i/\mu_i < 1, i = 1, 2$. В этом случае условия (4.52) выполнены и

$$\gamma_1 = \mu_1 - \lambda_1 > \gamma_2 = \mu_2 - \lambda_1, \quad \theta_{W^{(1)}} = (1 - \rho_1)^2 > (1 - \rho_2)^2 = \theta_{W^{(2)}},$$

т. е. выполнено неравенство (4.53).

Этот пример иллюстрирует утверждение Теоремы 4.2. Следующий пример также иллюстрирует этот вывод.

Пример 4.13. Сравним две системы $\Sigma^{(1)}$ и $\Sigma^{(2)}$ вида $M/H_2/1$. Пусть в системе $\Sigma^{(1)}$ обслуживание имеет гиперэкспоненциальное распределение с параметрами $\mu_1^{(1)} = 2, \mu_2^{(1)} = 4, p^{(1)} = 0.5$, а в системе $\Sigma^{(2)} : \mu_1^{(2)} = 1.5, \mu_2^{(2)} = 2, p^{(2)} = 0.5$. Входной поток в обеих системах одинаковый, $\tau^{(1)} \stackrel{st}{=} \tau^{(2)} \sim \text{Exp}(1.5)$. Отметим, что функции интенсивности отказов времен обслуживания связаны неравенством $r_{S^{(1)}} = \mu_1^{(1)} = 2 > r_{S^{(2)}} = \mu_1^{(2)} = 1.5$, поэтому времена обслуживания упорядочены по интенсивности отказов, $S^{(1)} \leq_r S^{(2)}$. Численное решение уравнения (4.47) дает по два ненулевых действительных корня для каждой системы,

$$\gamma_1^{(1)} = 1, \gamma_2^{(1)} = 3.5; \quad \gamma_1^{(2)} = 0.21, \gamma_2^{(2)} = 1.8.$$

Подставив наименьшие, т. е. $\gamma_1^{(1)}$ и $\gamma_1^{(2)}$, в (4.45), получим упорядоченность экстремальных индексов стационарного времени ожидания

$$\theta_{W^{(1)}} = 0.175 > \theta_{W^{(2)}} = 0.015.$$

Таким образом, условия (4.52) выполнены, $\gamma_1 > \gamma_2$ и $\theta_{W^{(1)}} > \theta_{W^{(2)}}$.

Пример 4.14. Рассмотрим $M/We/1$ систему. Экспоненциальные моменты $Me^{\gamma S}$ для распределения Вейбулла существуют для $\beta \geq 1$ так, что в этом случае (4.44) принимает вид 4.48. Пусть $\lambda_1 = 2, \alpha_1 = 0.25, \beta_1 = 1.5$ в системе $\Sigma^{(1)}$ и $\lambda_2 = 2, \alpha_2 = 0.4, \beta_2 = 1.2$ в системе $\Sigma^{(2)}$. Численное решение уравнения (4.48) дает единственные корни $\gamma_1 = 3.7$ и $\gamma_2 = 0.79$, соответственно, следовательно, $\theta_{W^{(1)}} = 0.35 > \theta_{W^{(2)}} = 0.07$.

Следующая теорема утверждает, что Предположение 4.1 выполнено для систем $GI/M/1$ при условии существования действительных положительных корней уравнения (4.38).

Теорема 4.3. Пусть для двух систем $\Sigma^{(1)}$ и $\Sigma^{(2)}$ вида $GI/M/1$ выполнены условия стационарности, $\rho_i < 1$, $i = 1, 2$, а также соотношение (4.52). Пусть существуют действительные положительные корни уравнения (4.38) для этих систем и

$$M[\tau^{(1)} e^{-\gamma_1 \tau^{(1)}}] \leq M[\tau^{(2)} e^{-\gamma_2 \tau^{(2)}}]. \quad (4.54)$$

Тогда экстремальные индексы стационарных времен ожидания упорядочены как

$$\theta_{W^{(1)}} \geq \theta_{W^{(2)}}.$$

Доказательство. В доказательстве Теоремы 4.2 показано, что корни γ_i уравнения (4.38) для систем $\Sigma^{(i)}$ связаны неравенством

$$\gamma_1 \geq \gamma_2,$$

а в силу (4.52) экспоненциальные моменты также упорядочены. Следовательно,

$$\psi_{\tau^{(1)}}(\gamma_1) := Me^{-\gamma_1 \tau^{(1)}} \leq Me^{-\gamma_2 \tau^{(2)}} =: \psi_{\tau^{(2)}}(\gamma_2).$$

Заметим, что $M\tau e^{-\gamma\tau} < \infty$ для любой с.в. $\tau \geq 0$ и любого $\gamma \geq 0$ и

$$\frac{d}{d\gamma_i} \psi_{\tau^{(i)}}(\gamma_i) = -M(\tau^{(i)} e^{-\gamma_i \tau^{(i)}}), \quad i = 1, 2.$$

Таким образом, из условия (4.54) следует, что

$$\frac{d}{d\gamma_1} \psi_{\tau^{(1)}}(\gamma_1) \geq \frac{d}{d\gamma_2} \psi_{\tau^{(2)}}(\gamma_2).$$

Уравнение (4.38) для рассматриваемой системы эквивалентно выражению

$$\psi_{\tau^{(i)}}(\gamma_i) + \gamma_i/\mu_i = 1, \quad i = 1, 2.$$

Следовательно, выражение (4.42) может быть переписано в виде

$$\theta_{W^{(i)}} = \gamma_i \left(\frac{d}{d\gamma_i} \psi_{\tau^{(i)}}(\gamma_i) + 1/\mu_i \right) = (1 - \psi_{\tau^{(i)}}(\gamma_i)) \left(\mu_i \frac{d}{d\gamma_i} \psi_{\tau^{(i)}}(\gamma_i) + 1 \right), \quad i = 1, 2.$$

Из упорядоченности по интенсивности отказов времен обслуживания следует, что $\mu_1 \geq \mu_2$. Подставляя полученные неравенства в выражение для экстремального индекса (4.42), получим требуемое неравенство

$$\theta_{W^{(1)}} = (1 - \psi_{\tau^{(1)}}(\gamma_1)) \left(\mu_1 \frac{d}{d\gamma_1} \psi_{\tau^{(1)}}(\gamma_1) + 1 \right) \geq (1 - \psi_{\tau^{(2)}}(\gamma_2)) \left(\mu_2 \frac{d}{d\gamma_2} \psi_{\tau^{(2)}}(\gamma_2) + 1 \right) = \theta_{W^{(2)}}.$$

□

Заметим, что условие (4.54) выполняется например, для $M/M/1$, поскольку

$$\frac{d}{d\gamma_i} \psi_{\tau^{(i)}}(\gamma_i) = -\frac{\lambda_i}{(\lambda_i + \gamma_i)^2} = -\frac{\lambda_i}{\mu_i^2},$$

и если $\lambda_1 \leq \lambda_2, \mu_1 \geq \mu_2$, то условие (4.54) выполняется.

Пример 4.15. Рассмотрим две системы $\Sigma^{(1)}$ и $\Sigma^{(2)}$ вида $H_2/M/1$. Пусть в системе $\Sigma^{(1)}$ интервалы между приходами заявок $\tau^{(1)}$ заданы гиперэкспоненциальным распределением с параметрами $p = 0.5, \lambda_1 = 1, \lambda_2 = 3$, а время обслуживания $S^{(1)} \sim \text{Exp}(4)$. Единственное положительное решение уравнения (4.38) есть $\gamma_1 = 2.24$, при котором $\theta_{W^{(1)}} = 0.33$. Пусть в системе $\Sigma^{(2)}$ с.в. $\tau^{(2)}$ заданы гиперэкспоненциальным распределением с параметрами $p = 0.5, \lambda_1 = 2, \lambda_2 = 3$, а $S^{(2)} \sim \text{Exp}(3)$. Тогда единственное положительное решение (4.38) есть $\gamma_2 = 0.58$, при котором $\theta_{W^{(2)}} = 0.0385$. При таких параметрах условие (4.54), а также выполнены условия (4.52), поскольку

$$r_{\tau^{(1)}}(x) = 1 < r_{\tau^{(2)}}(x) = 2; \quad r_{S^{(1)}}(x) = 4 > r_{S^{(2)}}(x) = 3,$$

и

$$\gamma_1 > \gamma_2, \quad \theta_{W^{(1)}} > \theta_{W^{(2)}}.$$

Пример 4.16. Рассмотрим систему $Pareto/M/1$. Вычислим ПЛС

$$\psi_\tau(\gamma) = 1 - e^{\gamma x_0}(\gamma x_0)^\alpha \Gamma(-\alpha + 1, \gamma x_0)$$

и его производную

$$\frac{d}{d\gamma}\psi_\tau(\gamma) = -e^{\gamma x_0}(\gamma x_0)^\alpha \Gamma(-\alpha + 1, \gamma x_0)(x_0 + \alpha/\gamma) + x_0.$$

В силу (4.40) производную ПЛС можно переписать в виде

$$\frac{d}{d\gamma}\psi_\tau(\gamma) = -\gamma/\mu(x_0 + \alpha/\gamma) + x_0.$$

Тогда формула (4.42) для расчета экстремального индекса примет вид

$$\theta_W = \gamma \left(x_0 + \frac{1 - \alpha}{\mu} - \frac{x_0 \gamma}{\mu} \right).$$

Сравним экстремальные индексы в двух системах. Пусть $x_0 = 1, \mu_1 = 2.2, \mu_2 = 1.375, \alpha_1 = \alpha_2 = 2.1$. В этом случае $\gamma_1 = 0.704 > \gamma_2 = 0.12375, \theta_{W(1)} = 0.126 > \theta_{W(2)} = 0.0136$.

4.8. Монотонность систем $GI/M/N$ по экстремальному индексу

Экстремальное поведение многосерверной системы $GI/M/N$ с интенсивностью обслуживания μ соответствует экстремальному поведению системы $GI/M/1$ с интенсивностью обслуживания μN с идентичным входным потоком [94]. Более того, если многосерверная система стационарна, $\rho = (\mu M \tau)^{-1} < N$, то время ожидания имеет предельное распределение смеси с атомом в нуле и показательным распределением с параметром γ , где γ есть единственное решение уравнения (см. Теорему 3.2 в [34])

$$\int_0^\infty e^{-\gamma x} F_\tau(dx) = 1 - \frac{\gamma}{\mu N}.$$

Подставляя μN вместо μ в соотношение (4.42) можно легко получить соотношение для экстремального индекса стационарного времени ожидания в системе

$GI/M/N$, а именно,

$$\theta_W = \gamma \left(\frac{d}{d\gamma} \psi_\tau(\gamma) + \frac{1}{\mu N} \right). \quad (4.55)$$

Последнее соотношение позволяет сравнивать экстремальные индексы в двух системах $GI/M/N_i$, добавляя условие $N_1 \geq N_2$ на число серверов в Теореме 4.3. Сформулируем полученный результат в виде теоремы.

Теорема 4.4. Пусть выполнены условия стационарности, $\rho_i = \lambda_i MS^{(i)} < N_i, i = 1, 2$, и соотношения (4.52) в двух системах $\Sigma^{(1)}$ и $\Sigma^{(2)}$ вида $GI/M/N_i$. Предположим, что существуют положительные корни уравнения (4.38) для этих систем и

$$N_1 \geq N_2.$$

Тогда для экстремальных индексов времен ожидания верно неравенство

$$\theta_{W^{(1)}} \geq \theta_{W^{(2)}}.$$

В частности, для $M/M/N$ легко проверить, что

$$\gamma = \mu N - \lambda, \theta_w = \left(1 - \frac{\lambda}{\mu N} \right)^2.$$

Следовательно, если $\lambda_1 \leq \lambda_2, \mu_1 \geq \mu_2, N_1 \geq N_2$, то

$$\theta_{W^{(1)}} \geq \theta_{W^{(2)}}.$$

4.9. Монотонность систем $GI/G/1$ со смешанным обслуживанием по экстремальному индексу

Рассмотрим односерверную систему Σ вида $GI/G/1$ с дисциплиной обслуживания FIFS. Пусть τ_n – интервалы между приходами n -й и $n+1$ -й заявок, S_n – время обслуживания n -й заявки. Пусть $\{\tau_n, n \geq 1\}, \{S_n, n \geq 1\}$ последовательности н.о.р. с.в. Обозначим $M\tau = 1/\lambda, MS = 1/\mu$. Предположим, что время обслуживания S задано m -компонентной смесью вида

$$B(x) = \sum_{i=1}^m p_i B_i(x), \quad \sum_{i=1}^m p_i = 1, \quad p_i \geq 0, \quad i = 1, \dots, m. \quad (4.56)$$

Предположим, что независимые с.в. $S^{(1)}, \dots, S^{(m)}$ имеют ф.р. $B_i(x)$, $i = 1, \dots, m$. Пусть индикаторы $I_i = 0$ с вероятностью p_i и $I_i = 1$ с вероятностью $1 - p_i$ и не зависят от $\{S^{(i)}\}$, $i = 1, \dots, m$. Тогда с.в.

$$S = I_1 S^{(1)} + \dots + I_m S^{(m)}, \quad I_1 + \dots + I_m = 1 \quad (4.57)$$

является m -компонентной смесью с ф.р. (4.56) и

$$MS = 1/\mu = \sum_{i=1}^m p_i/\mu_i,$$

где $MS^{(i)} = 1/\mu_i$. Обозначим через

$$\rho = \lambda MS = \sum_{i=1}^m \frac{\lambda p_i}{\mu_i} = \sum_{i=1}^m p_i \rho_i$$

коэффициент загрузки, где $\rho_i = \lambda/\mu_i$.

Предположим, что компоненты $S^{(i)}$ времени обслуживания S стохастически упорядочены

$$S^{(1)} \leq_{st} \dots \leq_{st} S^{(m)}.$$

Рассмотрим две системы обслуживания $\Sigma^{(1)}$ и $\Sigma^{(m)}$ с таким же входным потоком, как в исходной системе Σ , $M\tau^{(i)} = 1/\lambda_i = 1/\lambda$, $i = 1, m$. Как обычно будем снабжать индексом i величины, относящиеся к i -й системе. Время обслуживания $S^{(1)}$ в системе $\Sigma^{(1)}$ задано ф.р. B_1 . Время обслуживания $S^{(m)}$ в системе $\Sigma^{(m)}$ задано ф.р. B_m .

Теперь сравним в системах Σ , $\Sigma^{(1)}$ и $\Sigma^{(m)}$ экстремальные индексы стационарных времен ожидания. Пусть ν_n , $\nu_n^{(1)}$, $\nu_n^{(m)}$ – число заявок в момент времени t_n^- , Q_n , $Q_n^{(1)}$, $Q_n^{(m)}$ – размер очереди в момент t_n^- , а W_n , $W_n^{(1)}$, $W_n^{(m)}$ – время ожидания в очереди в момент прихода заявки с номером n в системах Σ , $\Sigma^{(1)}$, $\Sigma^{(m)}$, соответственно. Обозначим (когда существуют) пределы по распределению

$$Q_n^{(i)} \Rightarrow Q^{(i)}, \quad \nu_n^{(i)} \Rightarrow \nu^{(i)}, \quad W_n^{(i)} \Rightarrow W^{(i)}, \quad n \rightarrow \infty, \quad i = 1, m. \quad (4.58)$$

Пусть $\rho < 1$ для исходной системы Σ и существуют пределы

$$Q_n \Rightarrow Q, \quad \nu_n \Rightarrow \nu, \quad W_n \Rightarrow W, \quad n \rightarrow \infty.$$

Пусть

$$\begin{aligned} M_n^W &= \max(W_1, \dots, W_n), \\ M_n^{W^{(1)}} &= \max(W_1^{(1)}, \dots, W_n^{(1)}), \\ M_n^{W^{(m)}} &= \max(W_1^{(m)}, \dots, W_n^{(m)}), \quad n \geq 1. \end{aligned} \quad (4.59)$$

Предположение 4.2. Предположим, что в системах $\Sigma^{(1)}$ и $\Sigma^{(m)}$ коэффициенты загрузки $\rho_i < 1, i = 1, m, \rho < 1$, и

$$W_1^{(1)} = W_1^{(m)} = W_1 = 0. \quad (4.60)$$

Пусть также выполнены следующие стохастические соотношения:

$$\tau^{(1)} \underset{r}{\geq} \tau \underset{r}{\geq} \tau^{(m)} \quad (4.61)$$

и компоненты смеси времени обслуживания упорядочены как

$$S^{(1)} \underset{r}{\leq} S^{(2)} \underset{r}{\leq} \dots \underset{r}{\leq} S^{(m)}. \quad (4.62)$$

Предположим, что существуют предельные распределения максимумов стационарных времен ожидания $M_n^W, M_n^{W^{(1)}}$ и $M_n^{W^{(m)}}$ при $n \rightarrow \infty$. Тогда

$$\theta_{W^{(1)}} \geq \theta_W \geq \theta_{W^{(m)}}, \quad (4.63)$$

где $\theta_W, \theta_{W^{(i)}}$ – экстремальные индексы последовательностей $\{W_n\}, \{W_n^{(i)}\}, i = 1, m$, соответственно.

Построенные системы $\Sigma^{(1)}$ и $\Sigma^{(m)}$ будем называть минорантной и мажорантной системами для исходной системы Σ .

Следующая теорема утверждает, что Предположение 4.2 верно для системы со смешанным обслуживанием вида $M/G/1$, для которой существует действительный положительный корень уравнения (4.38).

Теорема 4.5. Пусть для систем $\Sigma, \Sigma^{(1)}$ и $\Sigma^{(2)}$ вида $M/G/1$ выполнены условия стационарности, $\rho_i < 1, i = 1, m, \rho < 1$, а также верны соотношения (4.60)-(4.62). Пусть существуют действительные положительные корни уравнения

(4.38) для всех трех систем и $M[S^{(i)}e^{\gamma S^{(i)}}] < \infty, i = 1, m, M[Se^{\gamma S}] < \infty$. Тогда

$$\theta_{W^{(1)}} \geq \theta_W \geq \theta_{W^{(m)}}.$$

Доказательство теоремы следует из Теорем 1.1 и 4.2.

4.9.1. Экстремальный индекс времени ожидания в системе $M/H_m/1$

В качестве исходной системы Σ рассмотрим систему $M/H_m/1$ с гиперэкспоненциальным (m -компонентным) распределением времени обслуживания

$$\bar{B}_S(x) = \sum_{i=1}^m p_i e^{-\mu_i x}, \quad \mu_i > 0, p_i \geq 0, i = 1, \dots, m, \quad \sum_{i=1}^m p_i = 1.$$

В качестве минорантной и мажорантной систем рассмотрим две системы $\Sigma^{(1)}$ и $\Sigma^{(m)}$ (вида $M/M/1$), в которых времена обслуживания $S^{(i)}$ имеют показательное распределение, $\bar{B}_{S^{(i)}}(x) = e^{-\mu_i x}, i = 1, m$. Интервалы между приходами заявок во все три системы имеют показательное распределение с параметром λ .

Пусть во всех системах выполнены условия стационарности:

$$\rho_i = \lambda/\mu_i < 1, \quad i = 1, m, \quad \rho = \lambda \sum_{i=1}^m p_i/\mu_i < 1.$$

Функция интенсивности отказов времени обслуживания S в исходной системе Σ равна

$$r_S(x) := \frac{\sum_{i=1}^m p_i \mu_i e^{-\mu_i x}}{\sum_{i=1}^m p_i e^{-\mu_i x}}, \quad x \geq 0.$$

Функции интенсивности отказов времен обслуживания $S^{(1)}$ и $S^{(m)}$ в системах $\Sigma^{(1)}$ и $\Sigma^{(m)}$, соответственно, равны $r_{S^{(1)}}(x) = \mu_1, r_{S^{(m)}}(x) = \mu_m$. Нетрудно проверить, что при условии

$$\mu_1 \geq \mu_2 \geq \dots \geq \mu_m \tag{4.64}$$

функции интенсивности отказов упорядочены следующим образом:

$$r_{S^{(1)}}(x) \geq r_S(x) \geq r_{S^{(m)}}(x), \quad x \geq 0,$$

а, следовательно, времена обслуживания в этих системах упорядочены по интенсивности отказов как

$$S^{(1)} \leq_{\frac{r}{r}} S \leq_{\frac{r}{r}} S^{(m)}.$$

Отсюда, в свою очередь, следует стохастический порядок [145]:

$$S^{(1)} \leq_{\frac{st}{st}} S \leq_{\frac{st}{st}} S^{(m)}.$$

Поскольку системы $\Sigma^{(1)}$ и $\Sigma^{(m)}$ являются системами вида $M/M/1$, то ф.р. стационарного времени ожидания равна

$$F_{W^{(i)}}(x) = 1 - \rho_i e^{-(\mu_i - \lambda)x}, \quad i = 1, m. \quad (4.65)$$

Выбирая нормализующие последовательности для максимума стационарного времени ожидания в виде

$$u_n^i(x) = \frac{x + \ln(\rho_i n)}{\mu_i - \lambda}, \quad i = 1, m,$$

получим предельное распределение Гумбеля [95]:

$$\lim_{n \rightarrow \infty} P(M_n^{W^{(i)}} \leq u_n^i(x)) = \Lambda(x), \quad i = 1, m. \quad (4.66)$$

Кроме того, для систем $M/M/1$ известны выражения для экстремальных индексов стационарных времен ожидания в явном виде [94]:

$$\theta_{W^{(1)}} = (1 - \rho_1)^2, \quad \theta_{W^{(m)}} = (1 - \rho_m)^2. \quad (4.67)$$

Вернемся к исходной системе $M/H_m/1$. Для систем $M/G/1$ экстремальный индекс можно вычислить по формуле (4.45). Более того, из условия для параметров (4.64) и уравнения (4.47) следует, что

$$\frac{\lambda + \gamma}{\lambda} = \sum_{i=1}^m \frac{p_i \mu_i}{\mu_i - \gamma} \geq \sum_{i=1}^m \frac{p_i}{1 - \gamma/\mu_1} = \frac{\mu_1}{\mu_1 - \gamma},$$

а, следовательно,

$$\gamma(\mu_1 - \lambda - \gamma) \geq 0$$

и $\gamma \leq \mu_1 - \lambda = \gamma_1$. Далее, поскольку, $\rho_1 \leq \rho$, то

$$\theta_W = \frac{1 - \rho}{1 + \lambda/\gamma} \leq \frac{1 - \rho_1}{1 + \lambda/\gamma_1} = (1 - \rho_1)^2 = \theta_{W^{(1)}}.$$

Аналогично, можно показать, что

$$\frac{\lambda + \gamma}{\lambda} = \sum_{i=1}^m \frac{p_i \mu_i}{\mu_i - \gamma} \leq \frac{\mu_m}{\mu_m - \gamma},$$

и $\gamma \leq \mu_m - \lambda = \gamma_m$, а, следовательно,

$$\theta_W = \frac{1 - \rho}{1 + \lambda/\gamma} \geq \frac{1 - \rho_m}{1 + \lambda/\gamma_m} = (1 - \rho_m)^2 = \theta_{W^{(m)}}.$$

Полученный результат сформулируем в виде леммы.

Лемма 4.1. *Предположим, что в системах $\Sigma = M/H_m/1$, $\Sigma^{(1)}$ и $\Sigma^{(m)}$ коэффициенты загрузки $\rho_i < 1, i = 1, m$, $\rho < 1$, и выполнены соотношения (4.60) и (4.64). Тогда*

$$\theta_{W^{(1)}} \geq \theta_W \geq \theta_{W^{(m)}}.$$

Рассмотрим систему $M/H_2/1$ с параметрами $\lambda = 1, \mu_1 = 4, \mu_2 = 2, p = 0.5$. В примере 4.9 были численно найдены два положительных корня, среди которых был выбран наименьший $\gamma = 1.38$. Подставив его в (4.45), получим значение экстремального индекса $\theta = 0.36$.

В минорантной и мажорантной системах экстремальные индексы стационарного времени ожидания, соответственно, равны $\theta_{W^{(1)}} = (1 - \rho_1)^2 = 0.5625$ и $\theta_{W^{(2)}} = (1 - \rho_2)^2 = 0.25$.

Полученные значения экстремальных индексов для стационарного времени ожидания согласуются с Предположением 4.2 о том, что если выполнены соотношения (4.61) и (4.62), то экстремальные индексы в этих системах связаны неравенством (4.63), поскольку

$$\theta_{W^{(1)}} = 0.5625 > \theta_W = 0.36 > \theta_{W^{(m)}} = 0.25. \quad (4.68)$$

Кроме того,

$$\gamma_1 = \mu_1 - \lambda = 3 > \gamma = 1.38 > \gamma_2 = \mu_2 - \lambda = 1.$$

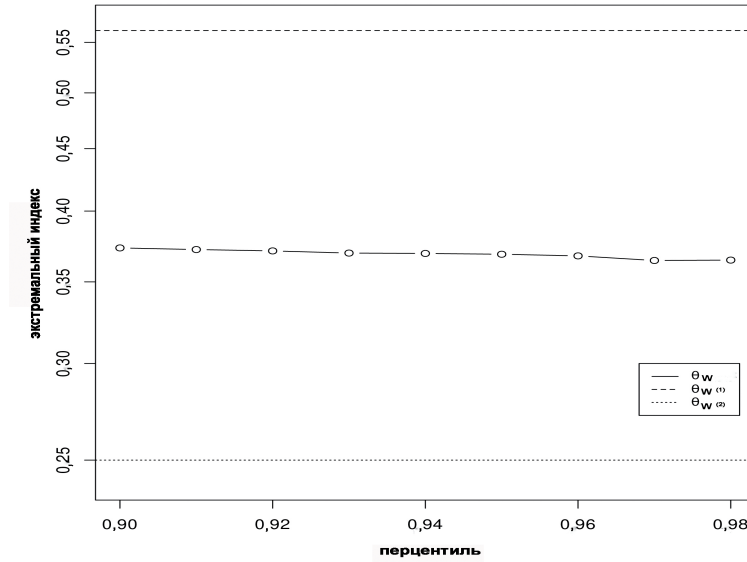


Рисунок 4.1. Экстремальные индексы стационарного времени ожидания в системах $M/M/1(\mu_1 = 4)$, $M/H_2/1(\mu_1 = 4, \mu_2 = 2, p = 0.5)$ и $M/M/1(\mu_2 = 2)$, $\lambda = \lambda_1 = \lambda_2 = 1$ в зависимости от перцентиля времени ожидания.

На Рисунке 4.1 приведены результаты эксперимента по регенеративной оценке экстремального индекса стационарного времени ожидания системы $M/H_2/1$. Экстремальный индекс в исходной системе сравнивался с экстремальными индексами двух систем $M/M/1$ с таким же входным потоком и временем обслуживания, распределенным по показательному закону с параметром $\mu_1 = 4$ (в минорантной системе) и $\mu_2 = 2$ (в мажорантной системе). Результаты моделирования подтверждают справедливость неравенства (4.63).

4.9.2. Экстремальный индекс времени ожидания в системе $H_m/G/1$

Пусть в исходной системе Σ интервалы между приходами заявок заданы гиперэкспоненциальным (m -компонентным) распределением

$$A(x) = 1 - \sum_{i=1}^m p_i e^{-\lambda_i x}, \quad \lambda_i > 0, \quad p_i \geq 0, \quad i = 1, \dots, m, \quad \sum_{i=1}^m p_i = 1.$$

В качестве минорантной и мажорантной систем рассмотрим две системы $\Sigma^{(1)}$ и $\Sigma^{(m)}$ (вида $M/M/1$), в которых интервалы между приходами заявок $\tau^{(i)}$ имеют показательный закон распределения с ф.р. $A_i(x) = 1 - e^{-\lambda_i x}$, $i = 1, m$. Н.о.р. времена обслуживания во всех трех системах имеют показательное распределение с параметром μ .

Пусть выполнены условия стационарности во всех системах

$$\rho_i = \lambda_i / \mu < 1, \quad i = 1, \dots, m, \quad \rho = 1 / \left(\sum_{i=1}^m p_i / \rho_i \right) < 1$$

и

$$\lambda_1 \leq \dots \leq \lambda_m. \quad (4.69)$$

Тогда функции интенсивности отказов упорядочены следующим образом:

$$r_{\tau^{(1)}}(x) \leq r_{\tau}(x) \leq r_{\tau^{(m)}}(x), \quad x \geq 0,$$

а, следовательно, интервалы между приходами заявок в этих системах упорядочены по интенсивности отказов как

$$\tau^{(1)} \underset{r}{\geq} \tau \underset{r}{\geq} \tau^{(m)}.$$

Согласно Предположению 4.2 экстремальные индексы в трех системах удовлетворяют неравенству (4.63). Более того, в силу (4.39) предельное распределение максимума стационарного времени ожидания в исходной системе $H_m/M/1$ есть распределение Гумбеля при нормализующей последовательности

$$u_n(x) = \frac{x + \ln((1 - \gamma/\mu)n)}{\gamma},$$

где γ есть единственное положительное решение уравнения (4.49), а экстремальный индекс рассчитывается по формуле (4.42). Более того, из условий (4.69) следует, что

$$\frac{\mu - \gamma}{\mu} = \sum_{i=1}^m \frac{p_i \lambda_i}{\lambda_i + \gamma} \geq \sum_{i=1}^m \frac{p_i}{1 + \gamma/\lambda_1} = \frac{\lambda_1}{\lambda_1 + \gamma}.$$

Из последнего соотношения следует, что

$$\gamma(\mu - \lambda_1 - \gamma) \geq 0,$$

а, следовательно, $\gamma \leq \mu - \lambda_1 = \gamma_1$. Подставив выражение

$$\frac{d}{d\gamma} \psi_{\tau}(\gamma) = - \sum_{i=1}^m \frac{p_i \lambda_i}{(\lambda_i + \gamma)^2}$$

в (4.42), получим

$$\theta_W = \frac{\gamma}{\mu} \left(1 - \mu \sum_{i=1}^m \frac{p_i \lambda_i}{(\lambda_i + \gamma)^2} \right) \leq \frac{\mu - \lambda_1}{\mu} \left(1 - \frac{\mu \lambda_1}{(\lambda_1 + \gamma)^2} \right) \leq (1 - \rho_1)^2 = \theta_{W^{(1)}}.$$

Аналогично, можно показать, что

$$\frac{\mu - \gamma}{\mu} \leq \sum_{i=1}^m \frac{p_i}{1 + \gamma/\lambda_m} = \frac{\lambda_m}{\lambda_m + \gamma},$$

и

$$\gamma \geq \mu - \lambda_m = \gamma_m,$$

а, следовательно,

$$\theta_W \geq \frac{\mu - \lambda_m}{\mu} \left(1 - \frac{\mu \lambda_m}{(\lambda_m + \gamma)^2} \right) \leq (1 - \rho_m)^2 = \theta_{W^{(m)}}.$$

Полученный результат сформулируем в виде леммы.

Лемма 4.2. *Предположим, что в системах $\Sigma = H_m/M/1$, $\Sigma^{(1)}$ и $\Sigma^{(m)}$ коэффициенты загрузки $\rho_i < 1, i = 1, m$, $\rho < 1$, и выполнены соотношения (4.60) и (4.69). Тогда*

$$\theta_{W^{(1)}} \geq \theta_W \geq \theta_{W^{(m)}}.$$

Пусть, например, $\lambda_1 = 2, \lambda_2 = 10, p = 0,25, \mu = 12$. В этом случае $\rho_1 = 1/6, \rho_2 = 5/6, \rho = 0.417$. Экстремальный индекс в минорантной системе $\theta_{W^{(1)}} = (1 - \rho_1)^2 = 0.694$, в мажорантной системе $\theta_{W^{(2)}} = (1 - \rho_2)^2 = 0.028$. Решение (4.49) дает единственный положительный корень $\gamma = 5.29$, следовательно, экстремальный индекс в исходной системе $\theta_W = 0.33$. Как видим,

$$\gamma_1 = \mu - \lambda_2 = 10 > \gamma = 5.29 > \gamma_2 = \mu - \lambda_1 = 2$$

и

$$\theta_{W^{(1)}} > \theta_W > \theta_{W^{(2)}}.$$

4.10. Экстремальное поведение в системах со смешанным входным процессом

Рассмотрим систему $\sum_i^m G_i / \sum_i^m G_i / 1$, в которую поступает m независимых входных потоков $\{\tau_k^{(i)}, k \geq 1\}$ с распределением A_i и средним $M\tau^{(i)} := 1/\lambda_i$. Эта система была подробно рассмотрена в разделе 2.6. Время обслуживания S является m -компонентной смесью распределений времен обслуживания разных классов

$$\bar{B}(x) = \sum_{i=1}^m p_i \bar{B}_i(x), \quad (4.70)$$

где $p_i = \lambda_i / \lambda$ $i = 1, \dots, m$, $\lambda = \lambda_1 + \dots + \lambda_m$. Пусть $\{\tau_k, k \geq 1\}$ – интервалы между приходами заявок в смешанном входном процессе, образованном суперпозицией m входных потоков и выполнено условие стационарности (2.96). Для исходной системы $\sum_{i=1}^m G_i / \sum_{i=1}^m G_i / 1$ (обозначим ее Σ) построим минорантную систему $\Sigma^{(1)}$, в которой интервалы между приходами заявок заданы с.в. $\tau_k^{(1)}$ из первого входного потока, а время обслуживания есть с.в. $S^{(1)}$. В мажорантной системе $\Sigma^{(m)}$ входной поток совпадает с входным потоком исходной системы τ , а обслуживание задано с.в. $S^{(m)}$.

Обозначим через W_n , $W_n^{(1)}$, $W_n^{(m)}$ – незавершенную работу в момент времени t_n^- в системах Σ , $\Sigma^{(1)}$ и $\Sigma^{(m)}$, соответственно. Введем максимумы стационарных последовательностей незавершенной работы

$$M_n^W = \max(W_1, \dots, W_n),$$

$$M_n^{W^{(i)}} = \max(W_1^{(i)}, \dots, W_n^{(i)}), i = 1, m.$$

Предположение 4.2 можно применить и для систем со смешанным входным процессом. А именно, пусть выполнено условие стационарности (2.97), начальные интервалы между приходами заявок совпадают, $\tau_1 = \tau_1^{(1)} = \tau_1^{(m)}$ и

$$S^{(1)} \leq_r \dots \leq_r S^{(m)}.$$

Тогда согласно Предположению 4.2

$$\theta_{W^{(1)}} \geq \theta_W \geq \theta_{W^{(m)}}, \quad (4.71)$$

где $\theta_W, \theta_{W^{(i)}}$ – экстремальные индексы $W, W^{(i)}, i = 1, m$, соответственно.

Пусть в системе Σ входной поток задан суперпозицией m входных потоков с моментами прихода $\{t_k, k \geq 1\}$. В системе $\tilde{\Sigma}$ моменты прихода $\{\tilde{t}_k, k \geq 1\}$ представляют собой подпоследовательность моментов прихода в систему Σ , т. е. $\{\tilde{t}_k, k \geq 1\} \subseteq \{t_k, k \geq 1\}$. Пусть времена обслуживания в обеих системах одинаковые и не зависят от класса, $S \stackrel{st}{=} \tilde{S}$. Тогда при выполнении условия стационарности (2.96) Теорема 2.8 (раздел 2.6.) верна, и в Предположении 4.2 экстремальные индексы стационарного времени ожидания в этих системах удовлетворяют неравенству

$$\theta_{\tilde{W}} \geq \theta_W.$$

Пример 4.17. Для примера 2.9 (раздел 2.6.) построим оценки экстремальных индексов стационарного времени ожидания в исходной системе, которую обозначим $\sum_{i=1}^2 G_i/M/1$, и системах $G_1/M/1, G_2/M/1$. На Рисунке 4.2 приведены оценки экстремального индекса стационарного времени ожидания в исходной системе Σ при $\mu = 4, \alpha_1 = 2, \alpha_2 = 3$ (коэффициент загрузки $\rho = 0.75$), а также в системах $G_i/M/1, i = 1, 2$.

На Рисунке 4.3 приведены аналогичные графики, но $\mu = 5$ (коэффициент загрузки $\rho = 0.6$).

В обоих случаях (при $\mu = 4$ и $\mu = 5$) экстремальный индекс в системе $G_2/M/1$ является наиболее близкой оценкой для экстремального индекса исходной системы. Заметим также, что на Рисунке 4.2 экстремальный индекс меньше во всех трех системах, чем экстремальный индекс в соответствующих системах на Рисунке 4.3. Это можно объяснить тем, что коэффициенты загрузки в них больше, чем в соответствующих системах, приведенных на Рисунке 4.3.

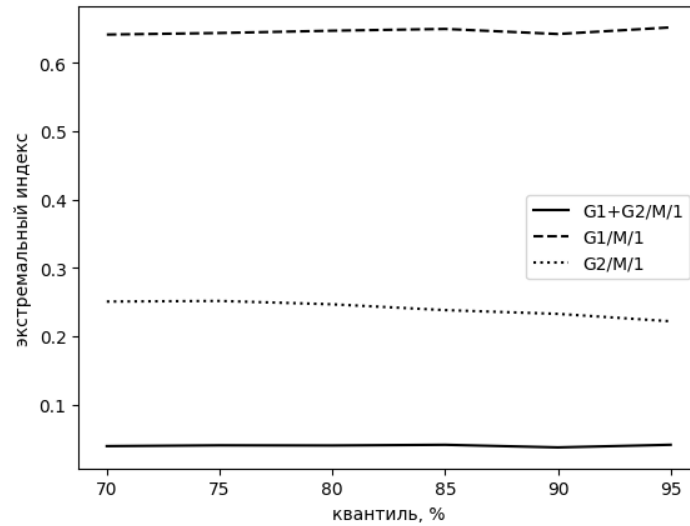


Рисунок 4.2. Оценки экстремальных индексов стационарного времени ожидания в исходной системе $\sum_{i=1}^2 G_i/M/1$ и системах $G_1/M/1$ и $G_2/M/1$, при $\mu = 4, \alpha_1 = 2, \alpha_2 = 3$.

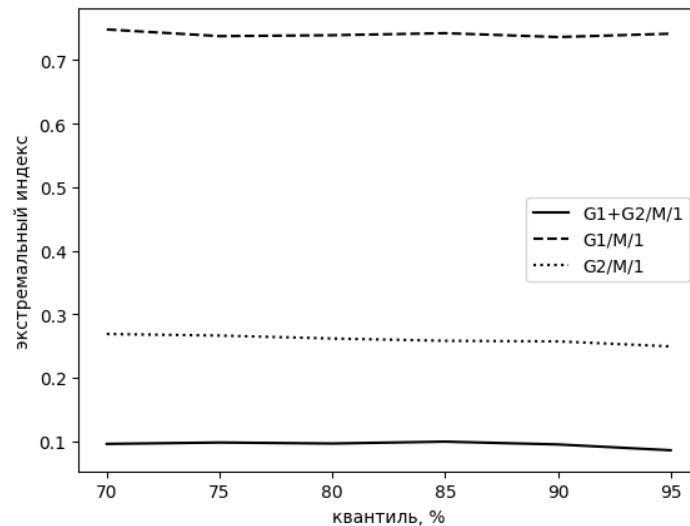


Рисунок 4.3. Оценки экстремального индекса стационарного времени ожидания в исходной системе $\sum_{i=1}^2 G_i/M/1$ и системах $G_1/M/1$ и $G_2/M/1$, при $\mu = 5, \alpha_1 = 2, \alpha_2 = 3$.

4.11. Сравнение экстремальных индексов стохастически упорядоченных строго стационарных последовательностей

В данном разделе обсуждается более широкий контекст Предположений 4.1 и 4.2 на произвольные строго стационарные стохастически упорядоченные

последовательности $\{X_n, n \geq 1\}$ и $\{Y_n, n \geq 1\}$, заданные (маргинальными) ф.р. F_X и F_Y , соответственно. Обозначим максимумы

$$M_n^X = \max(X_1, \dots, X_n), \quad M_n^Y = \max(Y_1, \dots, Y_n), \quad n \geq 1.$$

Пусть $\hat{M}_n^X, \hat{M}_n^Y, n \geq 1$, – максимумы последовательностей независимых с.в., ассоциированных с исходными последовательностями, т. е. заданными (маргинальными) ф.р. F_X и F_Y .

Пусть последовательности $\{X_n, n \geq 1\}$ и $\{Y_n, n \geq 1\}$ взаимно независимы. Предположим, что существуют такие нормализующие последовательности

$$\{u_n(x) = a_n x + b_n\} \text{ и } \{u'_n(x) = a'_n x + b'_n\},$$

что $a_n, a'_n > 0, n \geq 1, u_n(x), u'_n(x) \rightarrow \infty$ для каждого x при $n \rightarrow \infty$. Пусть также существуют такие невырожденные распределения $H_X(x)$ и $H_Y(x)$, что при $n \rightarrow \infty$,

$$\begin{aligned} P((\hat{M}_n^X \leq u_n(x)) &\rightarrow H_X(x), \\ P((\hat{M}_n^Y \leq u'_n(x)) &\rightarrow H_Y(x), \end{aligned} \quad (4.72)$$

$$\begin{aligned} n\bar{F}_X(u_n(x)) &\rightarrow \tau(x), \\ n\bar{F}_Y(u'_n(x)) &\rightarrow \tau'(x), \end{aligned} \quad (4.73)$$

и, кроме того,

$$u_n(x) \geq u'_n(x) \text{ для всех } x, n \geq 1. \quad (4.74)$$

Пусть также выполнены условия перемешивания $D(u_n)$ (см. раздел 4.5.), а последовательности имеют стохастический порядок, т. е.

$$\bar{F}_X(x) \leq \bar{F}_Y(x) \text{ для всех } x \in \mathbf{R}. \quad (4.75)$$

Тогда предположим, что экстремальные индексы θ_X и θ_Y последовательностей $\{X_n, n \geq 1\}$ и $\{Y_n, n \geq 1\}$ упорядочены следующим образом:

$$\theta_X \geq \theta_Y. \quad (4.76)$$

Пример 4.18. Пусть заданы две последовательности с.в.

$$\{X_n = c_1 Z_{n-1} \vee c_2 Z_n, n \geq 1\} \quad \text{и} \quad \{Y_n = d_1 Z_{n-1} \vee d_2 Z_n, n \geq 1\},$$

где $X \vee Y := \max(X, Y)$, константы $c_1, c_2, d_1, d_2 > 0$, ф.р. с.в. Z_n имеет вид

$$F_Z(x) = e^{-1/x}, \quad x > 0.$$

Вычислим ф.р. X и Y :

$$F_X(x) = e^{-\frac{c_1 + c_2}{x}}, \quad F_Y(x) = e^{-\frac{d_1 + d_2}{x}}, \quad x > 0.$$

Известно, что максимум последовательности $\{X_n, n \geq 1\}$ равен [16]:

$$M_n^X = c_1 Z_0 \vee \left((c_1 \vee c_2) \bigvee_{i=1}^{n-1} Z_i \right) \vee c_2 Z_n, \quad n > 1$$

Тогда [16] ф.р. M_n^X :

$$P(M_n^X \leq x) = F_Z(x/c_2) F_Z(x/c_1) F_Z^{n-1}(x/(c_1 \vee c_2)).$$

Экстремальный индекс равен

$$\theta_X = \frac{c_1 \vee c_2}{c_1 + c_2}.$$

Аналогично,

$$\theta_Y = \frac{d_1 \vee d_2}{d_1 + d_2}.$$

Очевидно, если $c_1 + c_2 \leq d_1 + d_2$, то $F_X(x) \geq F_Y(x)$ для всех $x > 0$. Следовательно, $X \underset{st}{\leq} Y$. Поскольку

$$\theta_X = \frac{c_1 \vee c_2}{c_1 + c_2} \geq \theta_Y = \frac{d_1 \vee d_2}{d_1 + d_2},$$

то вывод об упорядоченности экстремальных индексов для заданных последовательностей верен.

Пример 4.19. Пусть в примере 4.18 ф.р. последовательности $\{Z_n, n \geq 0\}$ имеет

распределение Парето:

$$F_Z(x) = 1 - \left(\frac{x_0}{x_0 + x} \right)^\alpha, \quad \alpha > 2, x \geq 0, x_0 > 0.$$

Вычислим ф.р. с.в. X

$$F_X(x) = \left(1 - \left(\frac{x_0 c_1}{x_0 c_1 + x} \right)^\alpha \right) \left(1 - \left(\frac{x_0 c_2}{x_0 c_2 + x} \right)^\alpha \right)$$

и ф.р. с.в. Y

$$F_Y(x) = \left(1 - \left(\frac{x_0 d_1}{x_0 d_1 + x} \right)^\alpha \right) \left(1 - \left(\frac{x_0 d_2}{x_0 d_2 + x} \right)^\alpha \right).$$

Легко проверить, что если $c_1 + c_2 \leq d_1 + d_2$, то $X \underset{st}{\leq} Y$. Получаем неравенство для экстремальных индексов

$$\theta_X = \frac{(c_1 \vee c_2)^\alpha}{c_1^\alpha + c_2^\alpha} \geq \frac{(d_1 \vee d_2)^\alpha}{d_1^\alpha + d_2^\alpha} = \theta_Y.$$

Таким образом, экстремальные индексы для заданных последовательностей также упорядочены в соответствии с (4.76).

4.12. Выводы по результатам главы 4

Построена нормализующая последовательность для максимума последовательности с.в., имеющих $k + m$ -показательное-Парето распределение, а также доказано, что распределение максимума сходится к распределению Фреше (раздел 4.3, Теорема 4.1).

Доказаны Теоремы 4.2, 4.3 и 4.4 о монотонности по экстремальному индексу стационарного времени ожидания систем вида $M/G/1$, $GI/M/1$ и $GI/M/N$, соответственно.

Получены достаточные условия монотонности систем со смешанным обслуживанием по экстремальному индексу (Теорема 4.5).

Для систем $M/H_m/1$ и $H_m/M/1$ получены соотношения для расчета экстремального индекса стационарного времени ожидания, а также доказаны Леммы 4.1 и 4.2 о монотонности системы по экстремальному индексу.

Основные результаты Главы 4 изложены в работах автора [189, 198, 199].

Глава 5. Экстремальное поведение систем с субъэкспоненциальным обслуживанием

В данной главе использованы теоретические результаты, полученные в предыдущих главах и позволяющие получить предельные распределения максимума стационарного времени ожидания (стационарной задержки) в системах с временем обслуживания, имеющим субъэкспоненциальное распределение, в том числе системах со смешанным обслуживанием.

5.1. Экстремальное поведение в системах $GI/G/1/\infty$

Рассмотрим систему $GI/G/1/\infty$. Обозначим через S_i – время обслуживания i -й заявки, τ_i – интервал между приходами i -й и $i + 1$ -й заявок, $i \geq 1$. Пусть последовательности $\{S_i, i \geq 1\}$, $\{\tau_i, i \geq 1\}$ состоят из н.о.р. с.в. Рассмотрим случайное блуждание $Y_n = X_1 + \dots + X_n$, $n \geq 1$ ($Y_0 = 0$) с н.о.р. нерешетчатыми с.в. $X_i = S_i - \tau_i$ и ф.р. F_X .

Обозначим через $B_S(x) = P(S \leq x)$ ф.р. времени обслуживания S .

Определение 5.1. ([100]) *Ф.р. B_S называется субъэкспоненциальной, если*

$$\lim_{x \rightarrow \infty} \frac{\overline{B_S^{*n}}(x)}{n \overline{B_S}(x)} = 1 \quad \text{для всех } n \geq 2,$$

где $\overline{B_S^{*n}}(x)$ – хвост n -кратной свертки распределения $B_S(x)$ с собой, т. е. $\overline{B_S^{*n}}(x) = P(S^1 + \dots + S^n > x)$, где S^i – стохастическая копия S , $i = 1, \dots, n$.

Известно, что если времена обслуживания имеют субъэкспоненциальное распределение, то $Me^{\gamma X} = \infty$ для всех $\gamma > 0$ [34].

Обозначим класс субъэкспоненциальных распределений через $\mathcal{S}$. Также обозначим через S_e стационарное остаточное время обслуживания, заданное плотностью распределения $\overline{B_S}(x)/MS$, и пусть $B_{S_e}(x)$ есть ф.р. с.в. S_e .

Ниже приведено Утверждение 5.1, на которое будут опираться дальнейшие рассуждения относительно экстремального поведения стационарного времени ожидания в системах $GI/G/1$, где время обслуживания имеет субъэкспоненциальное распределение.

Утверждение 5.1. ([174]) *Если система $GI/G/1$ стационарна, т. е. $\rho < 1$, и ф.р. S_e относится к классу субэкспоненциальных распределений ($S_e \in \mathcal{S}$), то ф.р. стационарного времени ожидания W также относится к классу субэкспоненциальных распределений и хвосты этих распределений асимптотически эквивалентны:*

$$\lim_{x \rightarrow \infty} \frac{P(W > x)}{P(S_e > x)} = \frac{\rho}{1 - \rho}. \quad (5.1)$$

Более того, в силу дистрибутивного закона Литтла [36] для систем вида $M/G/1$, Утверждение 5.1 можно сформулировать для стационарной длины очереди, а именно

$$\lim_{k \rightarrow \infty} \frac{P(Q > k)}{P(\lambda S_e > k)} = \frac{\rho}{1 - \rho}. \quad (5.2)$$

В отличие от случая, когда X имеет распределение с легким хвостом, в рассматриваемых системах с субэкспоненциальным обслуживанием предельное распределение максимума стационарного времени ожидания близко к распределению с.в. $X = S - \tau$, т. е. к ф.р. F_X . Пусть $F_W(x)$ – ф.р. стационарного времени ожидания W . Обозначим через $M_n^W = \max(W_1, \dots, W_n)$. Тогда имеет место

Утверждение 5.2. ([36])

$$P(M_W^{\hat{\Delta}} > x) \sim M \hat{\Delta} \bar{F}_X(x) \quad \text{при } x \rightarrow \infty, \quad (5.3)$$

где $M_W^{\hat{\Delta}} = \max(W_n, n = 0, \dots, \hat{\Delta} - 1)$ – максимум значений стационарного регенерирующего процесса $\{W_n, n \geq 0\}$ на цикле регенерации ($\hat{\Delta}$ – число приходов заявок на цикле регенерации, или, иначе, дискретная длина цикла регенерации).

Известно, что $M_W^{\hat{\Delta}} \stackrel{st}{=} M_X^{\hat{\Delta}}$ [36]. Поскольку экстремальный индекс определяется на цикле регенерации как [146]

$$\theta_W = \lim_{x \rightarrow \infty} \frac{P(M_W^{\hat{\Delta}} > x)}{M \hat{\Delta} P(W > x)},$$

то из Утверждений 5.1 и 5.2 и известного соотношения для субэкспоненциаль-

ных распределений (см. [78]),

$$\lim_{x \rightarrow \infty} \frac{\bar{F}_X(x)}{\bar{F}_{X_e}(x)} = 0,$$

вытекает, что экстремальный индекс стационарного времени ожидания для систем с субэкспоненциальным обслуживанием равен нулю [36], т. е.

$$\theta_W = 0. \quad (5.4)$$

5.1.1. Распределение Парето времени обслуживания

Пусть время обслуживания в системе $GI/G/1$ задано распределением Парето

$$B_S(x) = 1 - \left(\frac{x_0}{x_0 + x} \right)^\alpha, \quad \alpha > 1, \quad x_0 > 0, \quad x \geq 0,$$

которое относится к классу распределений с тяжелыми хвостами. Производящая функция моментов равна

$$M_S(\gamma) := \int_0^\infty e^{\gamma x} dB_S(x) = \alpha(-x_0\gamma)^\alpha e^{-\gamma x_0} \Gamma(-\alpha, -\gamma x_0),$$

где $\Gamma(t, x) = \int_x^\infty e^{-y} y^{t-1} dy$ – неполная верхняя Гамма-функция. В этом случае уравнение (4.38) не имеет решения $\gamma \neq 0$, поэтому Теоремы 4.1 и 4.3 не применимы.

В следующей теореме доказано, что нормализованный максимум стационарного времени ожидания M_n^W имеет предельное распределение Фреше с параметром $\alpha - 1$ при $n \rightarrow \infty$.

Теорема 5.1. *Если система $GI/G/1$ стационарна и время обслуживания S задано распределением $Pareto(\alpha, x_0)$, $\alpha > 1$, то при любом $x > 0$*

$$\lim_{n \rightarrow \infty} P(M_n^W \leq u_n(x)) = e^{-\frac{\lambda x_0}{\alpha - 1 - \lambda x_0} x^{-\alpha+1}}, \quad (5.5)$$

где $u_n(x) = x_0 n^{1/(\alpha-1)} x - x_0$, $n \geq 1$.

Доказательство. В стационарной системе коэффициент загрузки $\rho = \lambda x_0 / (\alpha -$

1) < 1 , $\alpha > 1$. Вычислим ф.р. с.в. S_e :

$$B_{S_e}(x) = \mathbf{P}(S_e \leq x) = \frac{\alpha - 1}{x_0} \int_0^x \left(\frac{x_0}{x_0 + y} \right)^\alpha dy = 1 - \left(\frac{x_0}{x_0 + x} \right)^{\alpha-1}.$$

Очевидно, S_e имеет распределение $Pareto(\alpha-1, x_0)$ и, следовательно, относится к классу субъэкспоненциальных распределений [100]. Согласно Утверждению 5.1, стационарное время ожидания W и S_e имеют асимптотически эквивалентные хвосты, и выполнено соотношение (5.1). Пусть $u_n(x) = x_0 n^{1/(\alpha-1)} x - x_0$. Тогда при $n \rightarrow \infty$,

$$n \bar{B}_{S_e}(u_n(x)) = \left(\frac{x_0}{x_0 + x_0 n^{1/(\alpha-1)} x - x_0} \right)^{\alpha-1} \rightarrow x^{-\alpha+1}, \quad (5.6)$$

что влечет $B_{S_e} \in MDA(\Phi_{\alpha-1})$ [107] и гарантирует выполнение условия (4.15) Утверждения 4.6 при $\eta = \alpha - 1$ и

$$c = \left(\frac{1 - \rho}{\rho} \right)^{1/(\alpha-1)} = \left(\frac{\alpha - 1 - \lambda x_0}{\lambda x_0} \right)^{1/(\alpha-1)}.$$

Следовательно,

$$\mathbf{P}(M_{W_n} \leq u_n(x)) \rightarrow \Phi_\eta(cx) \quad \text{при } n \rightarrow \infty.$$

□

На Рисунке 5.1 представлены результаты численного моделирования экстремального индекса для системы, в которой время обслуживания имеет распределение $Pareto(\alpha = 2.1, x_0 = 0.5)$, а входной поток пуассоновский с параметром $\lambda_\tau = 1.87$.

Несложно проверить, что

$$\lim_{x \rightarrow \infty} \frac{\mathbf{P}(S - \tau > x)}{\mathbf{P}(W > x)} = 0.$$

Рисунок 5.1 также демонстрирует, что значения экстремального индекса при любых значениях перцентиля действительно близки к нулю.

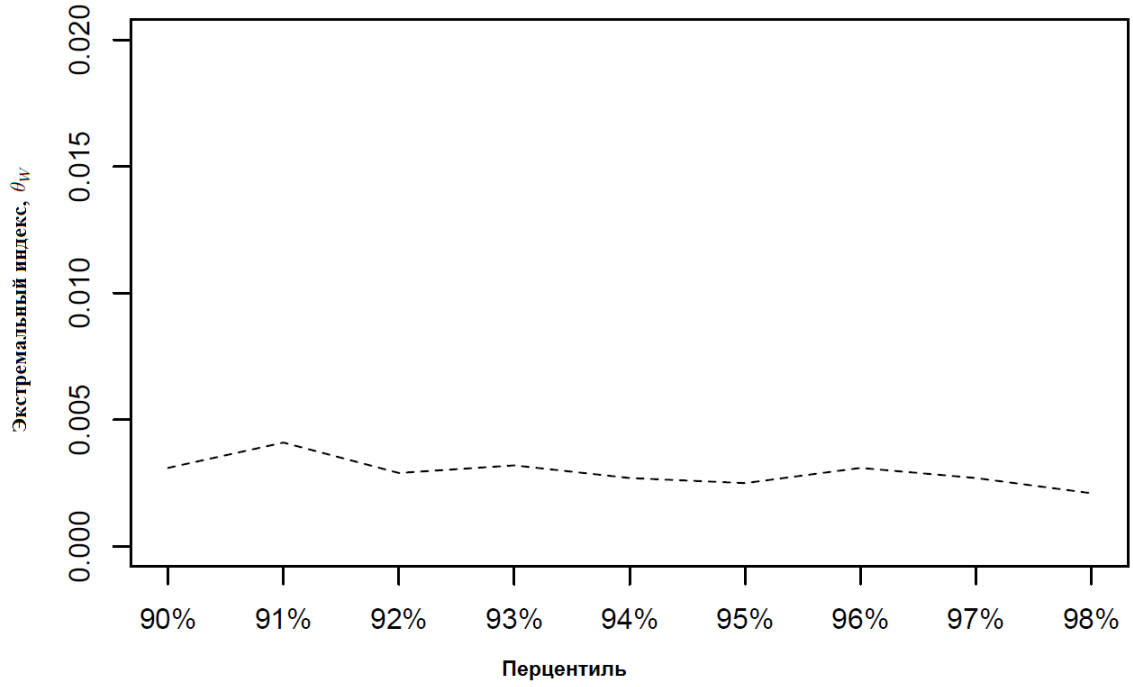


Рисунок 5.1. Зависимость экстремального индекса стационарного времени ожидания от перцентиля порогового значения в системе $M/Pareto/1$ при $\lambda_\tau = 1.87$, $\alpha_2 = 2.1$, $x_0 = 0.5$.

5.1.2. Распределение Вейбулла времени обслуживания

Предположим, что S имеет распределение Вейбулла

$$B_S(x) = 1 - e^{-(x/\alpha)^\beta}, \quad \alpha > 0, \quad \beta < 1, \quad x \geq 0, \quad (5.7)$$

с плотностью

$$f_S(x) = (\beta/\alpha)(x/\alpha)^{\beta-1} e^{-(x/\alpha)^\beta}.$$

Вычислим ф.р. с.в. S_e

$$B_{S_e}(x) = \frac{1}{MS} \int_0^x \overline{B}_S(t) dt = \frac{1}{\Gamma(1/\beta)} \int_0^{(x/\alpha)^\beta} e^{-y} y^{1/\beta-1} dy = \frac{\gamma(1/\beta, (x/\alpha)^\beta)}{\Gamma(1/\beta)}, \quad (5.8)$$

где $\Gamma(t)$ есть Гамма-функция и

$$\gamma(t, x) = \int_0^x e^{-y} y^{t-1} dy$$

– неполная нижняя Гамма-функция. Хвост ф.р. $\overline{B}_{S_e}$ удовлетворяет асимптотическому выражению

$$\overline{B}_{S_e}(x) = \frac{\Gamma(1/\beta, (x/\alpha)^\beta)}{\Gamma(1/\beta)} \sim \frac{(x/\alpha)^{1-\beta} e^{-(x/\alpha)^\beta}}{\Gamma(1/\beta)}, \text{ при } x \rightarrow \infty, \quad (5.9)$$

где

$$\Gamma(t, x) = \int_x^\infty e^{-y} y^{t-1} dy,$$

неполная верхняя Гамма-функция.

Чтобы проверить, что распределения B_S и B_{S_e} (с параметром $\beta \in (0, 1)$) относятся к классу субъэкспоненциальных распределений, достаточно показать, что B_S относится к специальному подклассу субъэкспоненциальных распределений $\mathcal{S}^*$. А именно, распределение $B_S \in \mathcal{S}^*$, если с.в. S имеет конечное среднее $1/\mu < \infty$ и выполнено предельное соотношение [87]

$$\lim_{x \rightarrow \infty} \int_0^x \frac{\overline{B}_S(x-y)}{\overline{B}_S(x)} \overline{B}_S(y) dy = \frac{2}{\mu}.$$

Рассмотрим функцию интенсивности отказов

$$r_S(x) = \frac{f_S(x)}{\overline{B}_S(x)}, \quad x \geq 0,$$

и пусть $R_S(x) = -\ln \overline{B}_S(x)$.

Определение 5.2. ([159]) *Ф.р. B_S имеет правильно меняющийся хвост с параметром $\delta > 0$, если $\overline{B}_S(x) > 0$ для всех x , и для любого фиксированного $t > 0$ выполнено соотношение*

$$\lim_{x \rightarrow \infty} \frac{\overline{B}_S(tx)}{\overline{B}_S(x)} = t^{-\delta}.$$

Будем обозначать через $\mathcal{R}(\delta)$ класс правильно меняющихся функций [87]. Далее используем следующий критерий принадлежности распределения классу

$\mathcal{S}^*$.

Утверждение 5.3. ([87]) *Предположим, что*

$$\lim_{x \rightarrow \infty} r_S(x) = 0, \quad \lim_{x \rightarrow \infty} x r_S(x) = \infty,$$

и выполнено одно из следующих условий:

$$a) \quad \limsup_{x \rightarrow \infty} \frac{x r_S(x)}{R_S(x)} < 1; \quad (5.10)$$

$$b) \quad r_S(x) \in \mathcal{R}(-\delta) \text{ для } \delta \in (0, 1]; \quad (5.11)$$

$$c) \quad R_S(x) \in \mathcal{R}(\delta) \text{ для } \delta \in (0, 1). \quad (5.12)$$

Тогда $B_S \in \mathcal{S}^*$.

Лемма 5.1. *Если B_S является распределением Вейбулла вида (5.7) с параметром $\beta \in (0, 1)$, то $B_S \in \mathcal{S}^*$.*

Доказательство. Проверим условия Леммы 5.3 для распределения Вейбулла. Действительно,

$$\begin{aligned} r_S(x) &= (\beta/\alpha)(x/\alpha)^{\beta-1} \rightarrow 0 \text{ при } x \rightarrow \infty; \\ x r_S(x) &= \beta(x/\alpha)^\beta \rightarrow \infty \text{ при } x \rightarrow \infty; \\ \frac{x r_S(x)}{R_S(x)} &= \beta < 1 \text{ для всех } x. \end{aligned}$$

Следовательно, условие (5.10) выполнено. □

Известно, что если $B_S \in \mathcal{S}^*$, то оба распределения B_S и B_{S_e} относятся к классу субэкспоненциальных, $B_S, B_{S_e} \in \mathcal{S}^*$ (см. [100]). Субэкспоненциальность B_{S_e} следует из соотношений:

$$\begin{aligned} r_{S_e}(x) = \frac{f_{S_e}(x)}{B_{S_e}(x)} &= \frac{(\beta/\alpha)e^{-(x/\alpha)^\beta}}{\Gamma(1/\beta, x^\beta)} \sim (\beta/\alpha)(x/\alpha)^{\beta-1} \rightarrow 0 \text{ при } x \rightarrow \infty; \\ x r_{S_e}(x) &\sim (\beta/\alpha)(x/\alpha)^\beta \rightarrow \infty \text{ при } x \rightarrow \infty; \\ \frac{r_{S_e}(tx)}{r_{S_e}(x)} &\rightarrow t^{\beta-1} = t^{-\delta} \text{ при } x \rightarrow \infty, \delta = 1 - \beta < 1. \end{aligned}$$

Напомним, что $\lambda = 1/M\tau$ – интенсивность входного потока. Для существования стационарного режима коэффициент загрузки для случая, когда время

обслуживания имеет распределение (5.7), должен удовлетворять неравенству

$$\rho = \frac{\lambda\alpha\Gamma(1/\beta)}{\beta} < 1. \quad (5.13)$$

Из соотношений (5.1) и (5.8) следует, что хвост распределения стационарного времени ожидания

$$P(W > x) \sim \frac{\lambda\alpha}{\beta - \lambda\alpha\Gamma(1/\beta)} (x/\alpha)^{1-\beta} e^{-(x/\alpha)^\beta} \text{ при } x \rightarrow \infty. \quad (5.14)$$

Теорема 5.2. Если время обслуживания в стационарной системе $GI/We/1$ имеет распределение Вейбулла вида (5.7) с параметром $\beta \in (0, 1)$, то предельное распределение нормализованного максимума последовательности n стационарных времен ожидания M_n^W при $n \rightarrow \infty$ является распределением типа Гумбеля,

$$P(M_n^W \leq u_n(x)) \rightarrow e^{-\frac{\lambda\alpha}{(\beta - \lambda\alpha\Gamma(1/\beta))\beta} e^{-x}}, \quad (5.15)$$

где нормализующая последовательность имеет вид

$$u_n(x) = \frac{x\alpha(\ln n)^{1/\beta-1}}{\beta} + \alpha(\ln n)^{1/\beta} + \frac{\alpha(1-\beta)}{\beta^2}(\ln n)^{1/\beta-1} \ln(\ln n). \quad (5.16)$$

Доказательство. Подставим хвост ф.р. (5.9) и последовательность (5.16) в выражение (4.17). Получим

$$\begin{aligned} n\bar{B}_e(u_n(x)) &= \frac{n}{\Gamma(1/\beta)} \left(\frac{u_n(x)}{\alpha} \right)^{1-\beta} e^{-(u_n(x)/\alpha)^\beta} = \\ &= \frac{n(\ln n)^{1/\beta-1}}{\Gamma(1/\beta)} \left(1 + \frac{x}{\beta \ln n} + \frac{1-\beta}{\beta^2} \frac{\ln(\ln n)}{\ln n} \right) \times \\ &\times \left(1 + \frac{x}{\ln n} + \frac{(1-\beta) \ln(\ln n)}{\beta \ln n} + o(1/\ln n) \right)^{-1} \times \\ &\times e^{-\ln n \left(1 + \frac{x}{\ln n} + \frac{(1-\beta) \ln(\ln n)}{\beta \ln n} + o(1/\ln n) \right)}. \end{aligned}$$

Поскольку

$$e^{-\ln n \left(1 + \frac{x}{\ln n} + \frac{(1-\beta) \ln(\ln n)}{\beta \ln n} + o(1/\ln n) \right)} \rightarrow$$

$$\rightarrow \frac{1}{n}(\ln n)^{1-1/\beta} e^{-x} \quad \text{при } n \rightarrow \infty,$$

то

$$n\overline{B}_e(u_n(x)) \rightarrow \frac{e^{-x}}{\Gamma(1/\beta)\beta}.$$

Подставив последнее соотношение и (5.1) в Утверждение 4.5 при

$$c = \ln \left(\frac{\beta - \lambda\alpha\Gamma(1/\beta)}{\lambda\alpha\Gamma(1/\beta)} \right),$$

получим, что максимум стационарной последовательности независимых с.в., заданных ф.р. $B_{S_e}(x)$, имеет действительно предельное распределение типа Гумбея. $\square$

На Рисунке 5.2 показаны результаты регенеративного оценивания экстремальных индексов стационарных времен ожидания двух систем. В системе $M/M/1$ интенсивность входного потока $\lambda_1 = 0.4$, интенсивность времени обслуживания $\mu = 0.5$. В системе $M/We/1$ входной поток также пуассоновский с параметром $\lambda_2 = 0.4$, а времена обслуживания имеют распределение $We(1, 0.5)$. В обеих системах коэффициенты загрузки $\rho_1 = \rho_2 = 0.8$.

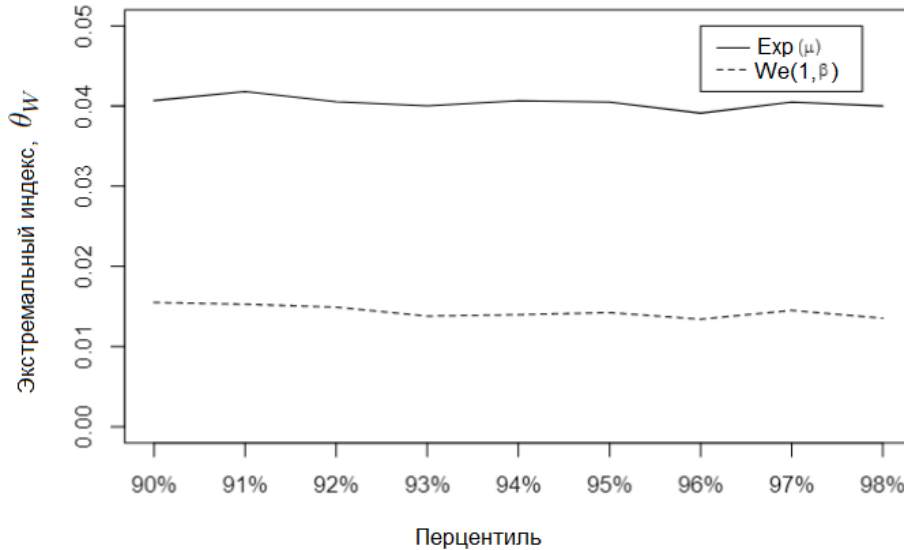


Рисунок 5.2. Экстремальные индексы стационарного времени ожидания в системах: $M/M/1$ с интенсивностью входного потока $\lambda = 0.4$, интенсивностью обслуживания $\mu = 0.5$, в системе $M/We/1$ с интенсивностью входного потока $\lambda = 0.4$ и распределением Вейбулла времени обслуживания с параметрами $\alpha = 1$, $\beta = 0.5$

Результаты оценки экстремального индекса в системе $M/M/1$ дают значения близкие к точному решению, которое задано формулой (4.46) и равно в этом случае $\theta_W = 0.04$. Напомним, что чем меньше экстремальный индекс, тем чаще встречаются большие значения процесса загрузки. Экстремальный индекс во второй системе близок к 0, что соответствует (5.4) (коэффициенты загрузки одинаковые).

5.2. Система $GI/G/1$ со смешанным обслуживанием с субъэкспоненциальными компонентами

Рассмотрим систему Σ вида $GI/G/1$ с дисциплиной обслуживания FIFS. Пусть $M\tau = 1/\lambda$, $MS = 1/\mu$ и время обслуживания S задано m -компонентной смесью распределений с ф.р.

$$B_S(x) = \sum_{i=1}^m p_i B_{S^{(i)}}(x), \quad \sum_{i=1}^m p_i = 1, \quad p_i \geq 0, \quad i = 1, \dots, m. \quad (5.17)$$

Предположим, что с.в. $S^{(1)}, \dots, S^{(m)}$ независимы с ф.р. $B_{S^{(i)}}(x)$, $i = 1, \dots, m$. Пусть индикаторы $I_i = 1$ с вероятностью p_i и $I_i = 0$ с вероятностью $1 - p_i$ и не зависят от $\{S^{(i)}\}$, $i = 1, \dots, m$. Тогда с.в.

$$S = I_1 S^{(1)} + \dots + I_m S^{(m)}, \quad I_1 + \dots + I_m = 1 \quad (5.18)$$

является m -компонентной смесью с ф.р. (5.17) и

$$MS = 1/\mu = \sum_{i=1}^m p_i / \mu_i,$$

где $MS^{(i)} = 1/\mu_i$. Обозначим через

$$\rho = \lambda MS = \sum_{i=1}^m \frac{\lambda p_i}{\mu_i} = \sum_{i=1}^m p_i \rho_i$$

коэффициент загрузки системы Σ , где $\rho_i = \lambda / \mu_i$.

Пусть Q_n – размер очереди, а W_n – время ожидания в момент прихода в t_n^- заявки с номером n в систему Σ , $n \geq 1$. Предположим, что существуют

пределы

$$Q_n \Rightarrow Q, \quad W_n \Rightarrow W, \quad n \rightarrow \infty. \quad (5.19)$$

Пусть $F_W(x)$ – ф.р. стационарного времени ожидания в исходной системе Σ . Обозначим через $M_n^W = \max(W_1, \dots, W_n)$ максимум стационарной последовательности времен ожидания $\{W_n, n \geq 1\}$ и через $\bar{B}_{S_e^{(i)}}(x) = \frac{1}{MS^{(i)}} \int_x^\infty \bar{B}_{S^{(i)}}(y) dy$ – функцию равновесного распределения с.в. $S_e^{(i)}, i = 1, \dots, m$.

Определение 5.3. ([103]) Компонента $F^{(j)}, j \in \{1, \dots, m\}$ является асимптотически r -доминирующей для m -компонентной смеси распределений $F(x) = p_1 F^{(1)}(x) + \dots + p_m F^{(m)}(x), \sum_{i=1}^m p_i = 1$, если

$$\lim_{x \rightarrow x_R} \frac{\bar{F}^{(i)}(x)}{\bar{F}^{(j)}(x)} = r_i, \quad i \in \{1, \dots, m\}, i \neq j, \quad (5.20)$$

где $r = (r_1, \dots, r_m), r_j = 1$, и $0 \leq r_i < 1$, для $i \neq j$, x_R – правая конечная точка $F(x)$.

Следующая теорема утверждает, что если в системе $GI/G/1$ время обслуживания задано конечной смесью (5.17), r -доминирующая компонента равновесного распределения которой относится к классу субэкспоненциальных распределений, $B_{S_e^{(j)}} \in \mathcal{S}$, то хвост предельного распределения стационарного времени ожидания эквивалентен хвосту этого распределения с точностью до константы,

$$\lim_{x \rightarrow \infty} (P(W > x) / \bar{B}_{S_e^{(j)}}(x)) = \gamma.$$

Более того, предельное распределение максимума стационарного времени ожидания относится к области притяжения распределения экстремальных значений того же типа, что и максимум с.в. с ф.р. $B_{S_e^{(j)}}$, а экстремальный индекс стационарного времени ожидания равен 0.

Теорема 5.3. Пусть исходная система Σ стационарна, $\rho < 1$. Пусть время обслуживания задано m -компонентной смесью распределений (5.17) и существует набор чисел $r = (r_1, \dots, r_m), 0 \leq r_i < 1, i \neq j, r_j = 1$, таких, что равновесное распределение j -й компоненты является r -доминирующим в смеси и относится к классу субэкспоненциальных распределений, $B_{S_e^{(j)}} \in \mathcal{S}$. Тогда

1) Равновесное распределение времени обслуживания принадлежит классу субэкспоненциальных распределений, $B_{S_e} \in \mathcal{S}$.

2) Хвосты равновесных распределений B_{S_e} и $B_{S_e^{(j)}}$ эквивалентны с точностью до константы,

$$\lim_{x \rightarrow x_R} \frac{\overline{B}_{S_e}(x)}{\overline{B}_{S_e^{(j)}}(x)} = \sum_{i=1}^m q_i r_i, \quad (5.21)$$

где

$$r_i := \lim_{x \rightarrow x_R} \frac{\overline{B}_{S_e^{(i)}}(x)}{\overline{B}_{S_e^{(j)}}(x)}, \quad 0 \leq r_i < 1, \quad i = 1, \dots, m, i \neq j, r_j = 1;$$

x_R – правая конечная точка ф.р. B_{S_e} и $q_i = \frac{p_i \text{MS}^{(i)}}{\text{MS}}$.

3) Хвосты распределения стационарного времени ожидания, $P(W > x)$, и равновесного распределения $B_{S_e^{(j)}}$ эквивалентны с точностью до константы γ ,

$$\lim_{x \rightarrow \infty} \frac{P(W > x)}{\overline{B}_{S_e^{(j)}}(x)} = \frac{\lambda \sum_{i=1}^m p_i r_i \text{MS}^{(i)}}{1 - \lambda \sum_{i=1}^m p_i \text{MS}^{(i)}} := \gamma. \quad (5.22)$$

4) Если $B_{S_e^{(j)}} \in \text{MDA}(G)$, то ф.р. стационарного времени ожидания в исходной системе $F_W \in \text{MDA}(G^\gamma)$.

5) Экстремальный индекс стационарного времени ожидания в исходной системе $\theta_W = 0$.

Доказательство. 1). Найдем равновесное распределение времени обслуживания в исходной системе

$$\begin{aligned} \overline{B}_{S_e}(x) &= \frac{1}{\text{MS}} \int_x^\infty (p_1 \overline{B}_{S_e^{(1)}}(y) + \dots + p_m \overline{B}_{S_e^{(m)}}(y)) dy = \sum_{i=1}^m \frac{p_i}{\text{MS}} \int_x^\infty \overline{B}_{S_e^{(i)}}(y) dy \\ &= \sum_{i=1}^m \frac{p_i \text{MS}^{(i)}}{\text{MS}} \overline{B}_{S_e^{(i)}}(x) = \sum_{i=1}^m q_i \overline{B}_{S_e^{(i)}}(x). \end{aligned}$$

Очевидно, $\sum_{i=1}^m q_i = 1$. Таким образом, равновесное распределение времени обслуживания в исходной системе представляет собой смесь равновесных распределений компонент с коэффициентами пропорции q_i ,

$$S_e = J_1 S_e^{(1)} + \dots + J_m S_e^{(m)}, \quad J_1 + \dots + J_m = 1$$

где индикатор J_i принимает значение 1 с вероятностью q_i , иначе 0, $i = 1, \dots, m$. Известно, что если хотя бы одна из компонент конечной смеси имеет субъэкспоненциальное распределение, то смесь относится к классу субъэкспоненциальных распределений [78]. Следовательно, п. 1) доказан.

2). Поскольку B_{S_e} – это m -компонентная смесь с коэффициентами пропорции q_i , где j -я компонента является r -доминирующей, то B_{S_e} и $B_{S_e^{(j)}}$ имеют эквивалентные хвосты с точностью до константы $\sum_{i=1}^m q_i r_i$ [103].

3). В силу Утверждения 5.1 и п. 2) настоящей теоремы

$$P(W > x) \sim \frac{\rho}{1-\rho} \bar{B}_{S_e}(x) \sim \frac{\rho}{1-\rho} \sum_{i=1}^m q_i r_i \bar{B}_{S_e^{(j)}}(x) = \gamma \bar{B}_{S_e^{(j)}}(x), \quad \text{при } x \rightarrow \infty,$$

$$\text{где } \gamma = \frac{\rho}{1-\rho} \sum_{i=1}^m q_i r_i = \frac{\lambda \sum_{i=1}^m p_i r_i M S^{(i)}}{1 - \lambda \sum_{i=1}^m p_i M S^{(i)}}.$$

Утверждение п. 4) следует из доказанного п. 3) настоящей теоремы и [103, теорема 8].

5). Поскольку $B_{S_e} \in \mathcal{S}$, то для систем вида $GI/G/1$ с субъэкспоненциальным обслуживанием экстремальный индекс стационарного времени ожидания равен нулю [36]. $\square$

Следствие 5.1. *Предположим, что система Σ стационарна, $\rho < 1$. Пусть время обслуживания задано m -компонентной смесью распределений (5.17) с упорядоченными по интенсивности отказов компонентами*

$$S^{(1)} \underset{r}{\leq} S^{(2)} \underset{r}{\leq} \dots \underset{r}{\leq} S^{(m)}. \quad (5.23)$$

Пусть $S_e^{(m)} \in \mathcal{S}$. Тогда все утверждения Теоремы 5.3 верны.

Доказательство. Для доказательства достаточно показать, что распределение $B_e^{(m)}$ является r -доминирующим для B_{S_e} . Так как из упорядоченности по интенсивности отказов (5.23) следует стохастический порядок с.в. $S_e^{(i)}$,

$$S_e^{(1)} \underset{st}{\leq} S_e^{(2)} \underset{st}{\leq} \dots \underset{st}{\leq} S_e^{(m)},$$

то

$$\frac{\overline{B}_{S_e^{(i)}}(x)}{\overline{B}_{S_e^{(m)}}(x)} \leq 1 \quad \text{для всех } x.$$

Очевидно, что существуют такие r_i , что $\lim_{x \rightarrow x_R} \frac{\overline{B}_{S_e^{(i)}}(x)}{\overline{B}_{S_e^{(m)}}(x)} = r_i$, $0 \leq r_i \leq 1$, $i = 1, \dots, m-1$, $r_m = 1$. В этом случае распределение $B_e^{(m)}$ является асимптотически r -доминирующим для распределения смеси B_{S_e} , $r = (r_1, \dots, r_m)$, $r_m = 1$, и условия Теоремы 5.3 выполнены при $j = m$. $\square$

5.3. Система $M/G/1$ с показательным-Парето временем обслуживания

Пусть входной поток в систему - пуассоновский с параметром λ_τ , а времена обслуживания S заданы *показательным-Парето распределением* с ф.р.

$$B_S(x) = 1 - pe^{-\lambda x} - (1-p) \left(\frac{x_0}{x_0 + x} \right)^\alpha, \quad \lambda > 0, \alpha > 0, x_0 > 0, x \geq 0. \quad (5.24)$$

Функция равновесного распределения имеет вид

$$B_{S_e}(x) = \mu \int_0^x \overline{B}(t) dt = 1 - \mu \left(\frac{pe^{-\lambda x}}{\lambda} + \frac{(1-p)x_0^\alpha}{(\alpha-1)(x_0+x)^{\alpha-1}} \right), \quad (5.25)$$

где $\mu = 1/MS$. Заметим, что

$$\overline{B}_{S_e}(x) = q_1 \overline{B}_e^{(1)}(x) + q_2 \overline{B}_e^{(2)}(x)$$

и

$$q_1 = \mu p / \lambda, \quad q_2 = \mu(1-p)x_0 / (\alpha-1).$$

Напомним, что функция интенсивности отказов имеет вид

$$r_S(x) = \frac{p \lambda a(x) + (1-p) \alpha / (x_0 + x)}{p a(x) + (1-p)}, \quad (5.26)$$

где

$$a(x) = e^{-\lambda x} \left(1 + \frac{x}{x_0}\right)^\alpha.$$

Следующая лемма показывает, что показательное-Парето распределения и его равновесное распределение (1.56) относятся к классу субэкспоненциальных распределений.

Лемма 5.2. *Показательное-Парето распределение (с параметром $\alpha > 1$) принадлежит классу $\mathcal{S}^*$ субэкспоненциальных распределений.*

Доказательство. Применим следующий критерий принадлежности ф.р. F_Z некоторой с.в. Z классу $\mathcal{S}^*$ [87]:

1. Если

$$\limsup_{x \rightarrow \infty} x r_Z(x) < \infty, \quad (5.27)$$

то $F_Z \in \mathcal{S}^*$.

2. Пусть

$$\lim_{x \rightarrow \infty} r_Z(x) = 0. \quad (5.28)$$

Тогда

$$F_Z \in \mathcal{S}^* \text{ тогда и только тогда, когда } \lim_{x \rightarrow \infty} \int_0^x e^{y r_Z(x)} \bar{F}_Z(y) dy = MZ. \quad (5.29)$$

Легко проверить, что для показательного-Парето распределения

$$r_S(x) \longrightarrow 0 \text{ и } x r_S(x) \longrightarrow \alpha \text{ при } x \rightarrow \infty.$$

Более того,

$$\begin{aligned} \int_0^x e^{y r_S(x)} \bar{B}_S(y) dy &= -\frac{p e^{-y(\lambda - r_S(x))}}{\lambda - r_S(x)} \Big|_0^x + (1-p) \frac{(x_0 r_S(x))^\alpha}{r_S(x)} e^{-x_0 r_S(x)} \\ &\times [\gamma(-\alpha + 1, x_0 r_S(x) + x r_S(x)) - \gamma(-\alpha + 1, x_0 r_S(x))] \longrightarrow \\ &\longrightarrow \frac{p}{\lambda} + \frac{(1-p)x_0}{\alpha - 1} = MZ \text{ при } x \rightarrow \infty, \end{aligned}$$

поскольку

$$\gamma(-\alpha + 1, x_0 r_S(x) + x r_S(x)) \rightarrow \gamma(-\alpha, \alpha),$$

и

$$\gamma(-\alpha + 1, x_0 r_S(x)) \sim \frac{(x r_S(x))^{-\alpha+1}}{-\alpha + 1} \quad \text{при } x \rightarrow \infty,$$

где $\gamma(\alpha, x)$ – нижняя неполная Гамма-функция. Следовательно, условия (5.27)–(5.29) выполнены и показательное-Парето распределение принадлежит классу $\mathcal{S}^*$. $\square$

Для рассматриваемой системы коэффициент загрузки имеет вид

$$\rho = \lambda_\tau \left[\frac{p}{\lambda} + \frac{(1-p)x_0}{\alpha - 1} \right]. \quad (5.30)$$

Поскольку, согласно Лемме 5.2, показательное-Парето распределение относится к классу $\mathcal{S}^*$, то для исходной системы верно асимптотическое выражение (5.1). Подставив выражение (5.25) в (5.1), получим, что хвост ф.р. стационарного времени ожидания удовлетворяет асимптотическому соотношению

$$\mathbf{P}(W > x) \sim \frac{\mu \lambda_\tau}{\mu - \lambda_\tau} \left(\frac{p e^{-\lambda x}}{\lambda} + \frac{(1-p)x_0^\alpha}{(\alpha - 1)(x_0 + x)^{\alpha-1}} \right) \quad \text{при } x \rightarrow \infty. \quad (5.31)$$

Теорема 5.4. *Если время обслуживания в стационарной системе $GI/G/1$ (т. е. при условии $\rho < 1$) имеет показательное-Парето распределение (5.24) с параметром $\alpha > 1$, то предельное распределение максимума стационарного времени ожидания M_n^W при $n \rightarrow \infty$ является распределением Фреше вида*

$$\lim_{n \rightarrow \infty} \mathbf{P}(M_n^W \leq u_n(x)) = e^{-\frac{(1-p)\mu\lambda_\tau x_0}{(\alpha - 1)(\mu - \lambda_\tau)} x^{1-\alpha}} \quad (5.32)$$

с нормализующей последовательностью (для $x > 0$)

$$u_n(x) = a_n x + b_n = x_0 x n^{1/(\alpha-1)} - x_0, \quad n \geq 1. \quad (5.33)$$

Доказательство. 1-й способ. Для доказательства достаточно подставить в вы-

ражение (4.17) хвост ф.р. (5.31) и нормализующую последовательность (5.33)

$$\begin{aligned}
n\bar{F}(u_n(x)) &= n \frac{\mu\lambda_\tau}{\mu - \lambda_\tau} \left[\frac{p}{\lambda} e^{-\lambda u_n(x)} + \frac{(1-p)x_0}{\alpha - 1} \frac{x_0^{\alpha-1}}{(x_0 + u_n(x))^{\alpha-1}} \right] \\
&= n \frac{\mu\lambda_\tau}{\mu - \lambda_\tau} \frac{(1-p)x_0}{\alpha - 1} x_0^{\alpha-1} (x_0 + x_0 x n^{1/(\alpha-1)} - x_0)^{-\alpha+1} \\
&\quad \times \left[1 + \frac{p(\alpha-1)(x_0 + u_n(x))^{\alpha-1} e^{-\lambda u_n(x)}}{\lambda(1-p)x_0^\alpha} \right] \\
&= \frac{\mu\lambda_\tau}{\mu - \lambda_\tau} \frac{(1-p)x_0}{\alpha - 1} x^{-\alpha+1} \left[1 + \frac{p(\alpha-1)(x_0 + u_n(x))^{\alpha-1} e^{-\lambda u_n(x)}}{\lambda(1-p)x_0^\alpha} \right] \\
&\longrightarrow \frac{\mu\lambda_\tau}{\mu - \lambda_\tau} \frac{(1-p)x_0}{\alpha - 1} x^{-\alpha+1} \text{ при } n \rightarrow \infty,
\end{aligned}$$

поскольку $u_n(x) \rightarrow \infty$ при $n \rightarrow \infty$.

2-й способ. Для доказательства воспользуемся Теоремой 5.3. В случае показательного-Парето распределения вторая компонента распределения является r -доминирующей при $r_1 = 0, r_2 = 1$. Действительно,

$$\lim_{x \rightarrow \infty} \frac{\bar{B}_{S_e^{(1)}}(x)}{\bar{B}_{S_e^{(2)}}(x)} = \lim_{x \rightarrow \infty} \frac{e^{-\lambda x}}{x_0^{\alpha-1}/(x_0 + x)^{\alpha-1}} = 0 = r_1.$$

Поскольку $B_{S_e^{(2)}} \in MDA(\Phi_{\alpha-1})$, то в силу п. 4 Теоремы 5.3, $F_W \in MDA(\Phi_{\alpha-1}^\gamma)$.

Вычислив

$$\gamma = \frac{\lambda_\tau(pr_1MS^{(1)} + (1-p)r_2MS^{(2)})}{1 - \lambda_\tau(pMS^{(1)} + (1-p)MS^{(2)})} = \frac{\mu\lambda_\tau}{\mu - \lambda_\tau} \frac{(1-p)x_0}{\alpha - 1},$$

получим утверждение теоремы. □

Таким образом, асимптотическое поведение максимума стационарного времени ожидания определяется второй компонентой распределения (распределением Парето).

Пример 5.1. Рассмотрим три системы вида $M/G/1$: систему $\Sigma^{(1)}$ с показательным временем обслуживания с интенсивностью $\lambda = 4.3$, систему $\Sigma^{(2)}$ с распределением Парето времени обслуживания с параметрами $\alpha = 2.1, x_0 = 0.5$ и систему $\Sigma^{(3)}$ с показательным-Парето обслуживанием с параметрами $x_0 = 0.5, \lambda = 4.3, \alpha = 2.1$ и коэффициентом пропорции смеси $p = 0.5$. Интенсивность входного потока во всех системах $\lambda_\tau = 1.87$. Как видно по результатам числен-

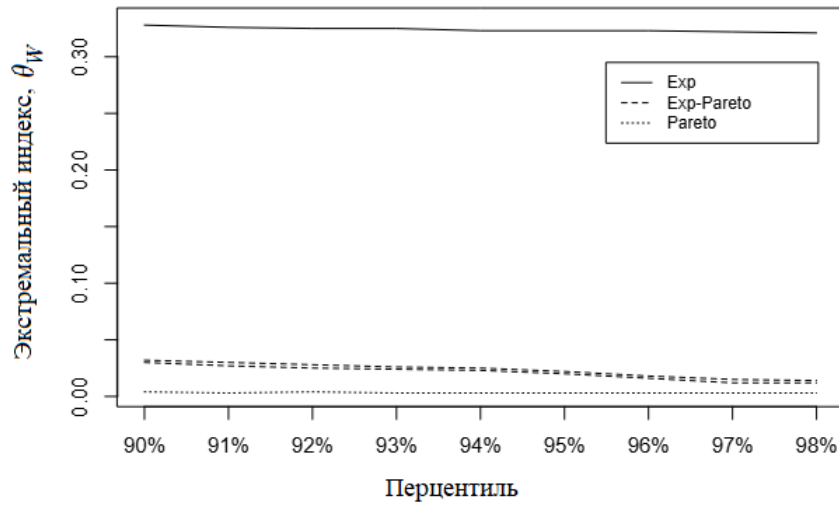


Рисунок 5.3. Экстремальный индекс стационарного времени ожидания при $\lambda_\tau = 1.87$, $p = 0.5$, $x_0 = 0.5$, $\lambda = 4.3$, $\alpha = 2.1$.

ного моделирования на Рисунке 5.3 оценки экстремального индекса упорядочены следующим образом:

$$\theta_{W(1)} > \theta_W = \theta_{W(2)} = 0. \quad (5.34)$$

Соотношение (5.34) соответствует Предположению 4.2 поскольку времена об-

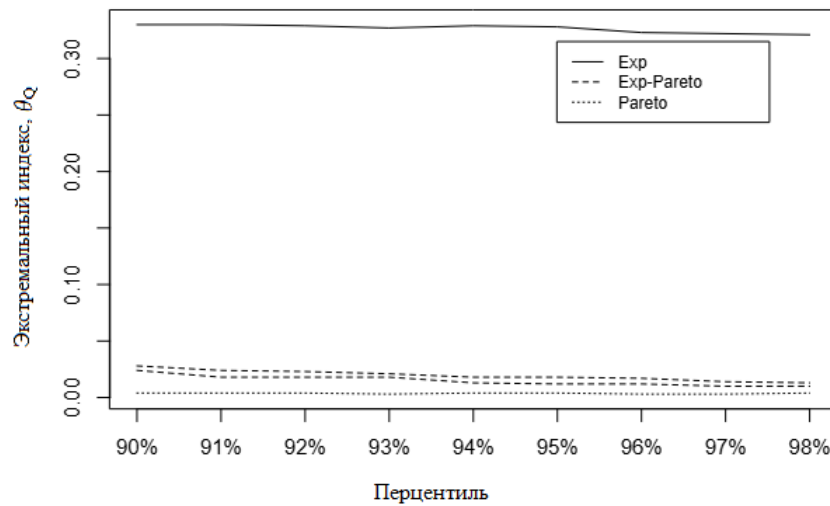


Рисунок 5.4. Экстремальный индекс стационарной длины очереди при $\lambda_\tau = 1.87$, $p = 0.5$, $x_0 = 0.5$, $\lambda = 4.3$, $\alpha = 2.1$.

служивания упорядочены по интенсивности отказов

$$S^{(1)} \underset{r}{\leq} S^{(3)} \underset{r}{\leq} S^{(2)}.$$

Рисунок 5.4 демонстрирует результаты моделирования экстремального индекса для стационарного времени ожидания в описанных выше системах.

5.4. Система с m классами повторных вызовов и временем обслуживания, заданным смесью субэкспоненциальных распределений

Рассмотрим систему с повторными вызовами Σ , а также минорантную $\Sigma^{(1)}$ и мажорантную $\Sigma^{(2)}$ системы, описанные в разделе 2.5.. Пусть выполнено условие стационарности (2.45) и существуют пределы

$$W_n \Rightarrow W, \quad W_n^{(i)} \Rightarrow W^{(i)}, \quad i = 1, 2, \quad n \rightarrow \infty.$$

В случае субэкспоненциальных распределений времен обслуживания для изучения предельного распределения максимума незавершенной нагрузки $\{W_n, n \geq 1\}$ в исходной системе можно применить асимптотику (5.1) для ф.р. незавершенной нагрузки F_W . Напомним, что ф.р. $F_{W^{(1)}}$ стационарной незавершенной нагрузки в минорантной системе $\Sigma^{(1)}$ (которая является классической системой $M/G/1$) получена в явном виде в (2.70), в то время как для незавершенной нагрузки в системе $\Sigma^{(2)}$ известно только ПЛС (2.66).

Применим асимптотику (5.1) для нахождения распределения стационарной незавершенной нагрузки в минорантной системе $\Sigma^{(1)}$, в которой ф.р. времени обслуживания S является смесью вида (2.55). Из (2.59) следует, что ф.р. S_e также есть m -компонентная смесь распределений с коэффициентами пропорции $\rho_1/\rho, \dots, \rho_m/\rho$, где $S_e(k)$ относятся к классу субэкспоненциальных с.в. Подставив (2.55) в (5.1), получим

$$P(W^{(1)} > x) \sim \frac{\rho}{1-\rho} \sum_{k=1}^m \frac{\rho_k}{\rho} \bar{F}_{S_e(k)}(x) = \sum_{k=1}^m \frac{\rho_k}{1-\rho} \bar{F}_{S_e(k)}(x), \quad x \rightarrow \infty. \quad (5.35)$$

Сформулируем полученный результат в виде леммы.

Лемма 5.3. Если $\rho < 1$, время обслуживания S задано ф.р. (2.55) с конечным математическим ожиданием, а ф.р. с.в. $S(k)$ и $S_e(k)$ относятся к классу субэкспоненциальных распределений, $k = 1, \dots, m$, то ф.р. стационарной незавершенной нагрузки $W^{(1)}$ в системе $\Sigma^{(1)}$ удовлетворяет соотношению

$$P(W^{(1)} > x) \sim \sum_{k=1}^m \frac{\rho_k}{1 - \rho} \bar{F}_{S_e(k)}(x), \quad x \rightarrow \infty. \quad (5.36)$$

Из (2.64) следует, что с.в. $S_e^{(2)}$ в системе $\Sigma^{(2)}$ является смесью с.в. $\xi_0 \sim \text{Exp}(\mu_0)$ и суммы с.в. вида $S_e(k) + \xi_0$. Известно, что свертка субэкспоненциальных распределений относится к классу субэкспоненциальных распределений [87, Теорема 5.1]. С.в. $S_e(k) + \xi_0$ и с.в. $S_e(k)$ имеют асимптотически эквивалентные хвосты распределений. Компоненту ξ_0 в смеси (2.64) можно не рассматривать в данном асимптотическом приближении, поскольку хвост субэкспоненциального распределения «тяжелее», чем экспоненциальный. Это следует из свойства (1.4) в [87]. Более того, из формулы (2.66) следует, что поскольку в мажорантной системе $\Sigma^{(2)}$ коэффициент загрузки $\rho^{(2)}$ определяется формулой (2.48), то в качестве коэффициентов пропорции выступает величина $\rho^{(2)}$ (вместо ρ) в формуле (5.36). Полученный результат сформулируем в виде леммы для стационарной незавершенной нагрузки в системе $\Sigma^{(2)}$.

Лемма 5.4. Если $\rho^{(2)} < 1$, время обслуживания S задано ф.р. (2.55) с конечным математическим ожиданием и с.в. $S(k)$ и $S_e(k)$ относятся к классу субэкспоненциальных распределений, $k = 1, \dots, m$, то ф.р. стационарной незавершенной нагрузки $W^{(2)}$ в системе $\Sigma^{(2)}$ удовлетворяет соотношению

$$P(W^{(2)} > x) \sim \sum_{k=1}^m \frac{\rho_k}{1 - \rho^{(2)}} \bar{F}_{S_e(k)}(x), \quad x \rightarrow \infty. \quad (5.37)$$

Заметим, что поскольку $\rho^{(2)} > \rho$, то коэффициенты пропорции в (5.36) и (5.37) удовлетворяют неравенству

$$\frac{\rho_k}{1 - \rho^{(2)}} > \frac{\rho_k}{1 - \rho}, \quad k = 1, \dots, m.$$

5.4.1. Экстремальное поведение в системе с M орбитами и распределениями Парето времен обслуживания

В данном разделе продолжим исследование предельного распределения стационарной незавершенной работы в исходной системе Σ , в которой имеется m орбит и время обслуживания заявок k -го класса $S(k)$ имеет распределение $Pareto(\alpha_k, \beta_k)$ вида (2.67), при условии

$$\alpha_1 \geq \dots \geq \alpha_m \geq 1.$$

Легко проверить, что $S_e(k)$ имеет ф.р.

$$F_{S_e(k)}(x) = \frac{\alpha_k - 1}{\beta_k} \int_0^x \left(\frac{\beta_k}{\beta_k + y} \right)^{\alpha_k} dy = 1 - \left(\frac{\beta_k}{\beta_k + x} \right)^{\alpha_k - 1}, \quad x \geq 0. \quad (5.38)$$

Подставив (5.38) в асимптотическое выражение (5.36), получим

$$\bar{F}_{W^{(1)}}(x) \sim \sum_{k=1}^m \frac{\rho_k}{1 - \rho} \left(\frac{\beta_k}{\beta_k + x} \right)^{\alpha_k - 1} \sim \frac{\rho_m}{1 - \rho} \left(\frac{\beta_m}{\beta_m + x} \right)^{\alpha_m - 1}, \quad x \rightarrow \infty. \quad (5.39)$$

Аналогично, из (5.37) следует, что, при $x \rightarrow \infty$,

$$\bar{F}_{W^{(2)}}(x) \sim \sum_{k=1}^m \frac{\rho_k}{1 - \rho^{(2)}} \left(\frac{\beta_k}{\beta_k + x} \right)^{\alpha_k - 1} \sim \frac{\rho_m}{1 - \rho^{(2)}} \left(\frac{\beta_m}{\beta_m + x} \right)^{\alpha_m - 1}. \quad (5.40)$$

Обозначим через $M_n^W = \max(W_1, \dots, W_n)$ максимум из n значений последовательности $\{W_n, n \geq 1\}$ в исходной системе Σ с орбитами, через $M_n^{W^{(i)}} = \max(W_1^{(i)}, \dots, W_n^{(i)})$ максимум из n членов последовательности незавершенной работы $\{W_n^{(i)}, n \geq 1\}$ в системе $\Sigma^{(i)}$, $i = 1, 2$.

Теорема 5.5. *Предположим, что выполнено условие стационарности (2.45), времена обслуживания имеют распределение (2.67) и*

$$\alpha_1 \geq \dots \geq \alpha_m \geq 1. \quad (5.41)$$

Тогда предельные распределения стационарной незавершенной работы в исходной системе Σ , минорантной и мажорантной системах $\Sigma^{(i)}$, $i = 1, 2$, принадлежат классу распределений Фреше, $F_{M_n^W}, F_{M_n^{W^{(i)}}} \in MDA(\Phi_{\alpha_m - 1})$, $i = 1, 2$,

соответственно.

Доказательство. Из (5.39) следует, что хвост $\bar{F}_{W_n^{(1)}}$ асимптотически эквивалентен хвосту $\bar{F}_{S_e(m)}$. Подставим нормализующую последовательность

$$u_n(x) = \beta_m x n^{1/(\alpha_m-1)} - \beta_m, \quad n \geq 1 \quad (5.42)$$

и $\bar{F}_{S_e(m)}$ ((5.38) при $k = m$) в (4.17), получим

$$\begin{aligned} n\bar{F}_{S_e(m)}(u_n(x)) &= n \left(\frac{\beta_m}{\beta_m + u_n(x)} \right)^{\alpha_m-1} = \\ &= n \left(\frac{\beta_m}{\beta_m x n^{1/(\alpha_m-1)}} \right)^{\alpha_m-1} \rightarrow \\ &\rightarrow x^{-\alpha_m+1}, \quad \text{при } n \rightarrow \infty, \quad x > 0. \end{aligned} \quad (5.43)$$

Таким образом, $F_{S_e(m)} \in MDA(\Phi_{\alpha_m-1})$. Далее, используя свойство сохранения областей притяжения для ф.р. с асимптотически эквивалентными хвостами распределений (Утверждение 4.6, раздел 4.2.), получим, что $F_{W_n^{(1)}} \in MDA(\Phi_{\alpha_m-1})$ при

$$c = \left(\frac{1-\rho}{\rho_m} \right)^{1/(\alpha_m-1)},$$

и, наконец,

$$P(M_n^{W^{(1)}} \leq u_n(x)) \rightarrow e^{-\frac{\rho_m}{1-\rho} x^{-\alpha_m+1}} \quad \text{при } n \rightarrow \infty.$$

Аналогично, используя асимптотическое выражение (5.40), можно показать, что $F_{W_n^{(2)}} \in MDA(\Phi_{\alpha_m-1})$ и, используя (5.42), получаем

$$P(M_n^{W^{(2)}} \leq u_n(x)) \rightarrow e^{-\frac{\rho_m}{1-\rho^{(2)}} x^{-\alpha_m+1}} \quad \text{при } n \rightarrow \infty.$$

Согласно Теореме 2.6 (раздел 2.5.1.), стационарная незавершенная работа в трех системах имеет следующий стохастический порядок, $W^{(1)} \leq W \leq W^{(2)}$, что означает, что $F_{W_n^{(1)}}(x) \geq F_{W_n}(x) \geq F_{W_n^{(2)}}(x)$ для каждого x . Следовательно,

$$P(M_{W_n^{(1)}} \leq u_n(x)) = F_{W_n^{(1)}}^n(u_n(x)) \geq F_{W_n}^n(u_n(x)) \geq$$

$$\geq F_{W_n^{(2)}}^n(u_n(x)) = P(M_{W_n^{(2)}} \leq u_n(x)), \quad x \geq 0.$$

Поскольку существование пределов левой и правой частей неравенство доказано выше, то $\lim_{n \rightarrow \infty} P(M_n^W \leq u_n(x))$ существует и более того, $F_W \in MDA(\Phi_{\alpha_m-1})$. $\square$

5.5. Экстремальное поведение в системе $M/G/2$

Рассмотрим систему $M/G/2$ с двумя идентичными серверами обслуживания и дисциплиной обслуживания FIFS. Заявки поступают в систему в моменты времени t_n , в соответствии с пуассоновским потоком с параметром λ . Н.о.р. времена обслуживания $\{S_n\}$ имеют произвольное распределение с ф.р. $B(x)$. Обозначим через $\rho = \lambda MS$ коэффициент загрузки системы.

Пусть $W_{n,i}$ — незавершенная работа на i -м сервере в момент t_n^- прихода n -й заявки, $i = 1, 2$. Рассмотрим вектор Кифера–Вольфовица $W_n = (W_{n,1}, W_{n,2})$, содержащий упорядоченную по возрастанию незавершенную работу на серверах, включая работу, требуемую для ожидающих в очереди заявок, т. е.

$$W_{n+1} = R((W_{n,1} + S_n - \tau_n)^+, (W_{n,2} - \tau_n)^+), \quad n \geq 1, \quad (5.44)$$

где оператор $R(x) = (R_1(x), R_2(x))$ упорядочивает компоненты вектора по возрастанию, $R_1(x) \leq R_2(x)$. Значение $W_{n,1}$ представляет собой задержку n -й заявки в системе, которая вызвана ее ожиданием в очереди перед обслуживанием. Предположим, что выполнено условие стационарности системы

$$\rho = \lambda MS < 2, \quad (5.45)$$

и заметим, что существует предел

$$W_{n,1} \Rightarrow W \quad \text{при} \quad n \rightarrow \infty.$$

Функция распределения F_W стационарной задержки W для исходной системы неизвестна. С другой стороны, W служит важной характеристикой качества обслуживания системы.

Обозначим через $M_n^W = \max(W_1, \dots, W_n)$ максимум из n с.в. стационарной последовательности времен ожидания, $n \geq 1$. Через $\hat{M}_n^W$ обозначим ассоцииро-

ванную с исходной последовательность независимых с.в. с ф.р. F_W . Рассмотрим асимптотическое поведение вероятности $P(M_n^W > x)$ при $x \rightarrow \infty$, когда времена обслуживания заданы распределением Парето

$$\bar{B}_S(x) = \left(\frac{x_0}{x_0 + x} \right)^\alpha, \quad x_0 \geq 0, \alpha > 1, x > 0. \quad (5.46)$$

Определение 5.4. ([77]) *Распределение B_S относится к классу распределений с периодически правильно меняющимся хвостом, $B_S \in \mathcal{IR}$, если $\bar{B}_S(x) > 0$ для всех x и*

$$\lim_{\epsilon \downarrow 0} \liminf_{x \rightarrow \infty} \frac{\bar{B}_S(x(1 + \epsilon))}{\bar{B}_S(x)} = 1.$$

Напомним, что ранее были рассмотрены классы субэкспоненциальных распределений $\mathcal{S}$ и распределений с правильно меняющимися хвостами $\mathcal{R}(\delta)$. Эти классы распределений имеют следующую взаимосвязь [77]:

$$\mathcal{R}(\delta) \subset \mathcal{IR} \subset \mathcal{S}. \quad (5.47)$$

Ф.р. $B_S(x)$ вида (5.46) относится к распределениям с правильно меняющимся хвостом, $B_S \in \mathcal{R}(\alpha)$. Известно, что если $B_S \in \mathcal{R}(\alpha)$, то $B_{S_e} \in \mathcal{R}(\alpha - 1)$ [188]. Например, для распределения Парето вида (5.46) выражение

$$\bar{B}_{S_e}(x) = \frac{\alpha - 1}{x_0} \int_x^\infty \left(\frac{x_0}{x_0 + y} \right)^\alpha dy = \left(\frac{x_0}{x_0 + x} \right)^{\alpha-1} \quad (5.48)$$

представляет собой хвост распределения Парето с параметром $\alpha - 1$.

5.5.1. Случай максимальной стабильности

Пусть выполнено условие максимальной стабильности в исходной двухсерверной системе, т. е. $\rho < 1$. Предположим, что $B_S \in \mathcal{R}(\alpha)$, $B_{S_e} \in \mathcal{S}$, $\alpha > 1$. Известно [77], что в этом случае

$$\bar{F}_W(x) = \bar{P}(W > x) \sim c_1(\bar{B}_{S_e}(x))^2, \quad x \rightarrow \infty, \quad (5.49)$$

где константа

$$c_1 = \frac{\rho^2}{2-\rho} \left(1 + \frac{\rho}{\alpha-1} \int_0^\infty \frac{dz}{(1+z)^{\alpha-1}(1+z(1-\rho))^\alpha} \right).$$

Лемма 5.5. Пусть в системе $M/G/2$ время обслуживания S имеет распределение Парето (5.46) и $\rho < 1$. Тогда распределение стационарного времени ожидания принадлежит области притяжения распределения Фреше, $F_W \in MDA(\Phi_{2\alpha-2})$, и более того,

$$\lim_{n \rightarrow \infty} P(\hat{M}_n^W \leq u_n(x)) = e^{-c_1 x^{-2\alpha+2}}, \quad (5.50)$$

где

$$u_n(x) = x_0 x n^{1/(2\alpha-2)} - x_0, \quad n \geq 1.$$

Доказательство. Известно [68], что если $G \in MDA(\Phi_{2\alpha-2})$ и $\lim_{n \rightarrow \infty} G^n(u_n(x)) = \Phi_{2\alpha-2}(x)$, $x > 0$, то

$$\lim_{n \rightarrow \infty} F^n(u_n(x)) = \Phi_{2\alpha-2}(\hat{b}x), \quad x > 0, \quad (5.51)$$

для некоторой константы $\hat{b} > 0$ тогда и только тогда, когда ф.р. F и G имеют асимптотически эквивалентные хвосты с точностью до константы,

$$\lim_{x \rightarrow \infty} \frac{\overline{G}(x)}{\overline{F}(x)} = \hat{b}^{2\alpha-2}. \quad (5.52)$$

Пусть

$$\overline{G}(x) := (\overline{B}_{S_e}(x))^2 = \left(\frac{x_0}{x_0 + x} \right)^{2\alpha-2}.$$

Из асимптотики (5.49) следует, что

$$\lim_{x \rightarrow \infty} \frac{\overline{G}(x)}{\overline{F}_W(x)} = 1/c_1,$$

следовательно, F_W и G имеют эквивалентные хвосты с точностью до константы $1/c_1$. Найдем предельное распределение максимума последовательности с.в.,

заданных распределением G . Подставив $\bar{G}$ в (4.17), получим

$$n\bar{G}(u_n(x)) = n \left(\frac{x_0}{x_0 + u_n(x)} \right)^{2\alpha-2} = \frac{nx^{-2\alpha+2}}{n} \xrightarrow{n \rightarrow \infty} x^{-2\alpha+2} =: v(x).$$

Подставляя $v(x)$ и $\hat{b} = c_1^{1/(-2\alpha+2)}$ в (5.51), получим требуемый результат (5.50). $\square$

Заметим, что предельное распределение максимума, вообще говоря зависящих, стационарных времен ожидания M_n^W получим, подставляя $a = \theta^{-1/(2\alpha-2)}$, $b = 0$ в соотношение (4.37), т. е.

$$\lim_{n \rightarrow \infty} P(M_n^W \leq u_n(x)) = e^{-\theta c_1 x^{-2\alpha+2}}.$$

5.5.2. Случай минимальной стабильности

Пусть выполнено условие минимальной в стабильности, т. е. $1 < \rho < 2$. Предположим, что $B_S \in \mathcal{S}$, $B_{S_e} \in \mathcal{IR}$. Известно, что в этом случае [77]

$$\bar{F}_W(x) = P(W > x) \sim \frac{\rho}{2-\rho} \bar{B}_e \left(\frac{\rho}{\rho-1} x \right), \quad x \rightarrow \infty. \quad (5.53)$$

Лемма 5.6. Пусть в системе $M/G/2$ времена обслуживания S заданы распределением Парето вида (5.46) и $1 < \rho < 2$. Тогда распределение стационарного времени ожидания принадлежит области притяжения максимума распределения Фреше, $F_W \in MDA(\Phi_{\alpha-1})$, и более того,

$$\lim_{n \rightarrow \infty} P(\hat{M}_n^W \leq u_n(x)) = e^{-c_2 x^{-\alpha+1}}, \quad (5.54)$$

где

$$c_2 = \frac{\lambda x_0}{2\alpha - 2 - \lambda x_0},$$

а нормализующая последовательность имеет вид

$$u_n(x) = (x n^{1/(\alpha-1)} - 1) \frac{x_0(\rho-1)}{\rho}, \quad n \geq 1.$$

Доказательство. Поскольку для распределения Парето ф.р. $B_S \in \mathcal{S}$ и $B_{S_e} \in \mathcal{IR}$, то можно использовать асимптотическую эквивалентность $\bar{B}_{S_e}$ и $\bar{F}_w$ и применить асимптотику (5.53) для вывода предельного распределения $\hat{M}_n$. Подста-

вив $u_n(x)$ и равновесное распределение (5.48) в (5.53), получим (для каждого x)

$$\begin{aligned} \frac{n\rho}{2-\rho} \overline{B}_{S_e} \left(\frac{\rho}{\rho-1} u_n(x) \right) &= \frac{n\lambda x_0}{2\alpha-2-\lambda x_0} \left(\frac{x_0(\rho-1)/\rho}{x_0(\rho-1)/\rho + u_n(x)} \right)^{\alpha-1} = \\ &= \frac{n\lambda x_0}{2\alpha-2-\lambda x_0} \frac{x^{-\alpha+1}}{n} \rightarrow \\ &\rightarrow \frac{\lambda x_0}{2\alpha-2-\lambda x_0} x^{-\alpha+1} = v(x) \text{ при } n \rightarrow \infty. \end{aligned}$$

Подставляя $v(x)$ в (4.17), получим (5.54). $\square$

Как и в случае максимальной стабильности предельное распределение максимума зависимых стационарных времен ожидания M_n^W следует из соотношения (4.37), но при других значениях параметров распределения, а именно при $a = \theta^{-1/(\alpha-1)}$ и $b = 0$.

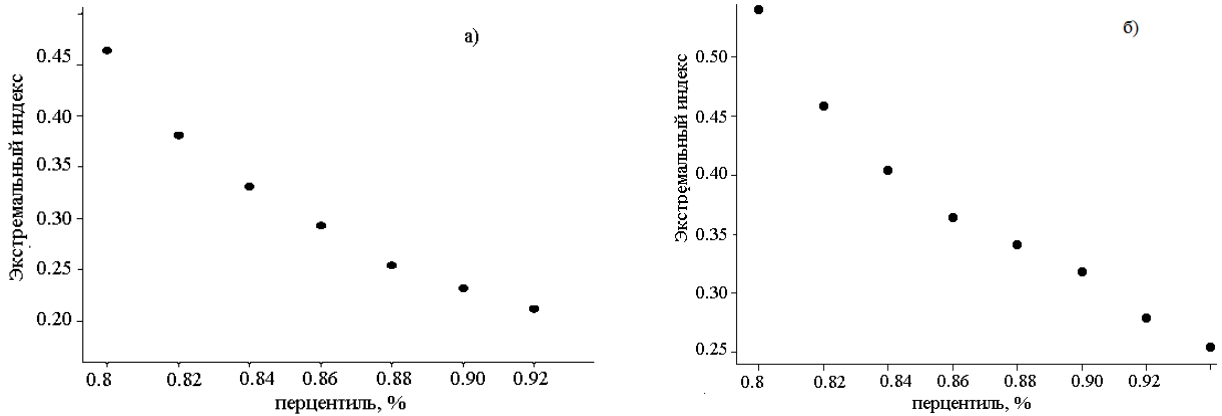


Рисунок 5.5. Зависимость экстремального индекса стационарного времени ожидания от перцентиля в системах $M/G/2$ на рисунке 5.5а: $x_0 = 1, \alpha = 2, \lambda = 0.8$; на рисунке 5.5б: $x_0 = 1, \alpha = 3, \lambda = 1.6$.

Пример 5.2. Пусть входной поток в систему с двумя идентичными серверами – пуассоновский с параметром λ . Время обслуживания имеет распределение Парето (5.46). Коэффициент загрузки системы равен $\rho = \lambda x_0 / (\alpha - 1)$. Ниже приведены результаты численных экспериментов по оценке экстремального индекса методом регенеративного моделирования.

Рисунок 5.5 демонстрирует результаты численного эксперимента при $\alpha = 2, x_0 = 1, \lambda = 0.8$ и $\alpha = 3, x_0 = 1, \lambda = 1.6$. В обоих случаях коэффициент загрузки $\rho = 0.8$ и, следовательно, асимптотическое поведение стационарной

задержки описывается распределением, заданным соотношением (5.49). Распределение максимума стационарной задержки определяется выражением (5.50), где

$$c_1 = \begin{cases} \frac{\rho \ln(1 - \rho)}{\rho - 2}, & \alpha = 2, \\ \frac{3(2\rho(1 - \rho) + \rho(\rho^2 - \rho + 2))}{4\rho(2 - \rho)}, & \alpha = 3. \end{cases}$$

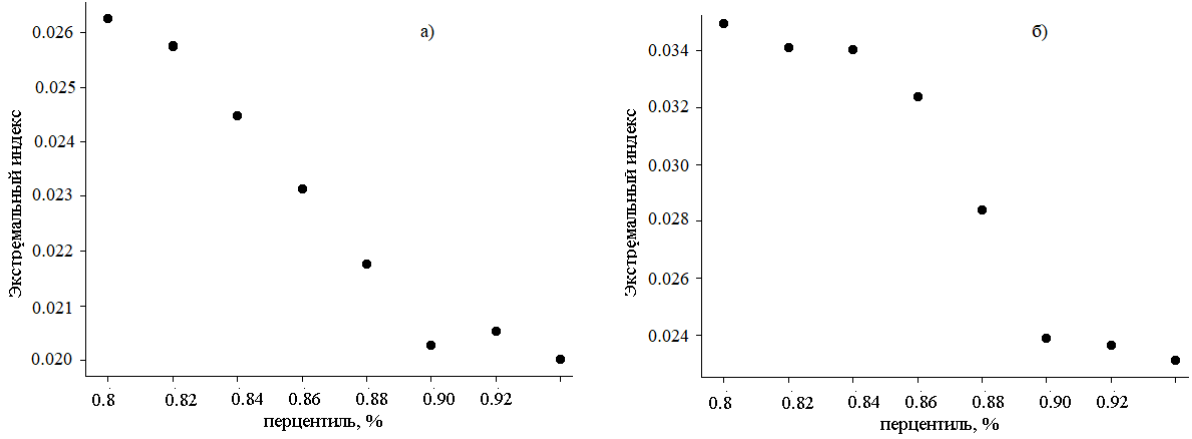


Рисунок 5.6. Зависимость экстремального индекса стационарного времени ожидания от квантиля порогового значения в системах $M/G/2$ на рисунке 5.6а: $x_0 = 1, \alpha = 2, \lambda = 1, 5$; на рисунке 5.6б: $x_0 = 1, \alpha = 3, \lambda = 3$.

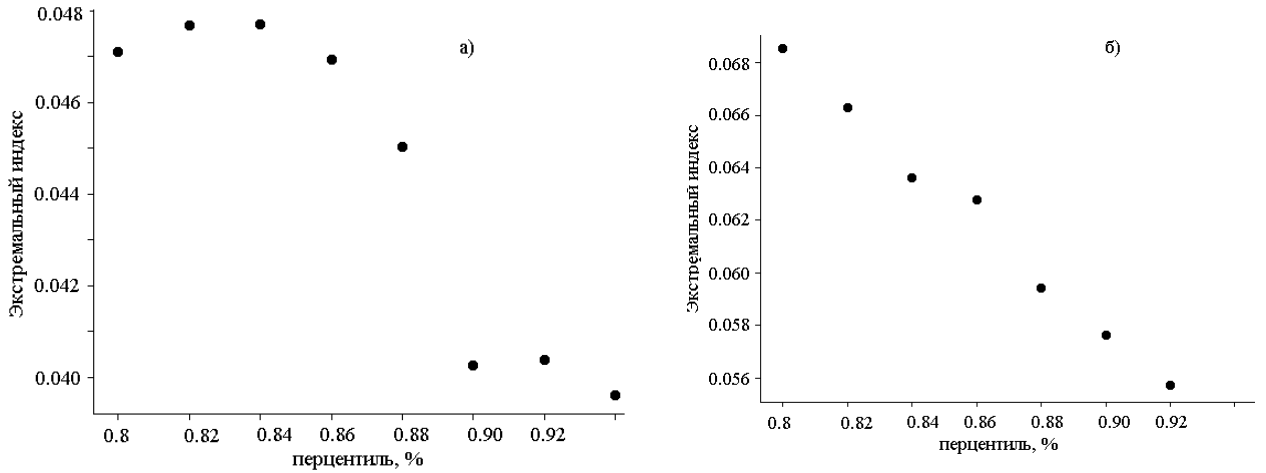


Рисунок 5.7. Зависимость экстремального индекса стационарного времени ожидания от перцентиля в системах $M/G/2$ на рисунке 5.7а: $x_0 = 1, \alpha = 6, \lambda = 7, 5$; на рисунке 5.7б: $x_0 = 1, \alpha = 11, \lambda = 15$

На Рисунке 5.6 отображены результаты численного эксперимента для систем с параметрами $\alpha = 2, x_0 = 1, \lambda = 1, 5$ и $\alpha = 3, x_0 = 1, \lambda = 3$. В обоих

случаях коэффициент загрузки $\rho = 1,5$ и, следовательно, асимптотическое поведение стационарной задержки описывается распределением (5.53), а максимум стационарной задержки — соотношением (5.54), где $c_2 = \frac{\lambda x_0 - \alpha + 1}{\lambda x_0}$. На Рисунке 5.7 коэффициент загрузки по-прежнему $\rho = 1,5$, но $\alpha = 6$ и $\alpha = 11$.

Заметим, что при одних и тех же значениях коэффициента загрузки ρ , чем меньше α (т. е. чем тяжелее хвост распределения времени обслуживания), тем меньше экстремальный индекс θ_W . При одном и том же α , чем меньше ρ (т. е. чем менее нагружена система), тем больше экстремальный индекс θ_W .

5.6. Выводы по результатам главы 5

Получены нормализующая последовательность и предельное распределение максимума стационарного времени ожидания в системах $GI/G/1$ с распределением Парето времени обслуживания (Теорема 5.1, раздел 5.1.) и с распределением Вейбулла времени обслуживания, $\beta \in (0, 1)$ (Теорема 5.2, раздел 5.1.)

Доказана Теорема 5.3 о том, что если в системе $GI/G/1$ время обслуживания задано конечной смесью (5.17), r -доминирующая компонента равновесного распределения которой относится к классу субэкспоненциальных распределений, то хвост предельного распределения стационарного времени ожидания эквивалентен хвосту этого распределения с точностью до константы. Более того, предельное распределение максимума стационарного времени ожидания относится к области притяжения распределения экстремальных значений того же типа, что и максимум с.в., заданных ф.р. доминирующей компоненты.

Доказано Следствие 5.1 о том, что если компоненты времени обслуживания упорядочены по интенсивности отказов и самая «тяжелая» компонента относится к классу субэкспоненциальных, то верна Теорема 5.3 (раздел 5.2.)

Получены нормализующая последовательность и предельное распределение максимума стационарного времени ожидания в системах $GI/G/1$ с распределением показательным-Парето временем обслуживания, (Теорема 5.4, раздел 5.3.)

Для систем с повторными вызовами с орбитами получены асимптотические распределения незавершенной работы в минорантной и мажорантной системах при условии, что время обслуживания относится к классу субэкспоненциальных распределений (Леммы 5.3 и 5.4, раздел 5.4.)

Для систем с повторными вызовами с орбитами и распределением Парето времени обслуживания найдены предельные распределения максимумов незавершенной работы в минорантной, исходной и мажорантной системах (Теорема 5.5, раздел 5.4.)

Для систем $M/G/2$, в которых времена обслуживания имеют распределение Парето, получены предельные распределения стационарного времени ожидания (Леммы 5.5 и 5.6 при условии максимальной, $\rho < 1$, и минимальной, $1 < \rho < 2$, стабильности, соответственно, раздел 5.5.)

Основные результаты главы представлены в работах автора [189, 192, 196, 198, 199, 202, 203, 205, 216].

Глава 6. Методы и алгоритмы регенеративного моделирования систем обслуживания

Поскольку поведение многих реальных коммуникационных систем сложно описать марковскими процессами, то для их анализа широко используется регенеративный метод [164]. Траектория регенерирующего процесса делится на н.о.р. части (случайные элементы), называемые циклами регенерации. Моменты начала цикла называются моментами регенерации, и они образуют вложенный процесс восстановления. Такая регенерация (когда циклы независимы) называется классической [12, 67]. При не ограничительных (приводимых ниже) условиях существует предельное распределение регенерирующего процесса.

Все численные результаты диссертационной работы опираются на тот или иной вариант регенеративного метода имитационного моделирования и оценивания характеристик систем. По этой причине в данной главе представлены основные варианты регенеративного метода, использованные для решения различных задач, рассмотренных в предыдущих главах диссертации.

Чем сложнее система, тем менее вероятно появление моментов классической регенерации, поскольку она, как правило, связана с попаданием основного процесса в определенное состояние, как правило, нулевое (опустошение системы). Вероятность таких событий может быть мала, а следовательно, моменты регенерации редки. Однако частота моментов регенерации является критически важной для надежного оценивания в условиях ограниченного времени.

Для оценки характеристик систем, в которых число моментов классической регенерации недостаточно или они очень редки, что может драматически сказываться на скорости получения оценки с заданной точностью, может быть применен метод рандомизации переходного ядра, предложенный в [29]. Метод позволяет ускорить получение моментов искусственной регенерации, и значит, получать регенеративные оценки быстрее (в смысле затрат времени моделирования), чем на основе классической регенерации. Важно подчеркнуть, что искусственная (рандомизированная) регенерация на самом деле также является классической в смысле независимости получаемых в результате циклов, и различная терминология используется для возможности сравнения частоты регенераций системы, полученных либо как число ее опустошений (классическая

регенерация), либо путем расщепления переходного ядра основного марковского процесса (искусственная регенерация).

В случае, если исследуемая система не регенерирует, может быть использован, например, метод регенеративных огибающих. Суть метода состоит в том, что для исходной системы строятся регенерирующие минорантная и мажорантная системы, а регенеративные оценки характеристик этих систем являются границами для исследуемых характеристик исходной системы.

В системах обслуживания, в которых входной поток или время обслуживания заданы распределением с тяжелым хвостом, для увеличения частоты моментов регенерации вполне естественно использовать метод экспоненциального расщепления, предложенный в работе [28]. Суть метода заключается в том, что вместо исходной с.в. моделируется стохастически эквивалентная с.в., плотность распределения которой задается как смесь плотностей показательного распределения и некоторого другого распределения. Вид последнего распределения определяется условием расщепления.

В системах со смешанным входным процессом проблема частоты регенераций тем более актуальна, когда несколько входных потоков имеют распределения с тяжелым хвостом. Более того, в таких системах интервалы между приходами заявок, вообще говоря, зависимы. В этой связи возникает необходимость адаптации метода искусственной регенерации, основанного на экспоненциальном расщеплении, для систем, в которых входные потоки имеют распределения с тяжелыми хвостами.

Все перечисленные выше методы развиваются в данной главе. Надо отметить, что моменты регенерации строятся искусственным образом так, что система обслуживания в эти моменты оказывается пустой, т. е., как было отмечено выше, вообще говоря, регенерирует в классическом смысле. Также в разделе 6.7. приведен метод регенеративного оценивания, применяемый в Главе 2 для оценки характеристик качества систем обслуживания. В разделе 6.8. приведена регенеративная оценка экстремального индекса, которая использована при оценивании экстремального индекса в Главах 4 и 5.

6.1. Классическая регенерация

Этот раздел опирается в основном на работы [14, 34, 96, 163]. Рассмотрим случайный процесс $Z = \{Z(t)\}_{t \in T}$, где $T = [0, \infty)$ или $T = \{0, 1, \dots\}$. В случае дискретного времени процесс будем обозначать $Z = \{Z_n\}_{n \geq 0}$. Предположим, что процесс Z имеет непрерывные справа траектории (при $T = [0, \infty)$).

Рассмотрим пару (Z, β) , образованную случайным процессом Z и неубывающей последовательностью с.в. $\beta = (\beta_0, \beta_1, \dots)$, $\beta_0 < \beta_1 < \dots$. Обозначим $((Z(\beta_k + t))_{t \geq 0}, (\beta_k - \beta_i)_{i \geq k}) = \Theta_{\beta_k}(Z, \beta)$, $k \geq 0$. Θ – оператор сдвига.

Определение 6.1. ([96]). Случайный процесс Z называется регенерирующим в классическом смысле, если для него выполнены следующие условия:

- (1) Распределение случайной пары $\Theta_{\beta_k}(Z, \beta)$ не зависит от $k \geq 0$;
- (2) Распределение случайной пары $\Theta_{\beta_k}(Z, \beta)$ не зависит от предистории $\{(Z(t))_{t < \beta_k}, \beta_0, \dots, \beta_k\}$, $k \geq 0$.

Процесс β называется процессом восстановления, вложенным в процесс Z , а моменты β_k называются моментами классической регенерации. Заметим, что последовательность моментов регенераций не обязательно единственная. Значения процесса между соседними моментами регенерации (м. р.) называют циклами регенерации, т. е. k -й цикл определяется как

$$G_k = \{Z(t + \beta_k)\}_{0 \leq t < \hat{\Delta}_k}, \quad k \geq 0,$$

а элемент $\{Z(t)\}_{t < \hat{\Delta}_0}$ называют задержкой (если $\hat{\Delta}_0 = \beta_0 = 0$, то говорят о процессе без задержки). Будем считать, что задержка конечна с в. 1, $\hat{\Delta}_0 < \infty$. Длины циклов регенерации $\{\hat{\Delta}_k = \beta_{k+1} - \beta_k, k \geq 0\}$ классически регенерирующих процессов являются н.о.р. с.в., а сами циклы являются н.о.р. случайными элементами. Регенерирующие процессы с конечной средней длиной цикла регенерации $M\hat{\Delta}_1 < \infty$ называются положительно возвратными.

Одно из важных свойств регенерирующих процессов состоит в том, что предельное распределение процесса существует при достаточно слабых условиях. Например, для процессов с непрерывным временем достаточно, чтобы $M\hat{\Delta}_1 < \infty$, с.в. $\hat{\Delta}_1$ была нерешетчатой (то есть не существует такого чис-

ла $d > 0$, что выполняется равенство $P \left(\sum_{j=1}^{\infty} \{ \hat{\Delta}_1 = a + j \cdot d \} \right) = 1$, где число $a > 0$), а выборочные траектории процесса должны быть непрерывны справа [34].

Рассмотрим процесс с нулевой задержкой (т. е. циклы G_0 и G_1 имеют одинаковое распределение). В этом случае будем математическое ожидание обозначать M_0 . Любую измеримую функцию $f(Z) : E \rightarrow \mathbb{R}^+$ назовем характеристикой процесса Z .

Утверждение 6.1. [34] Пусть длина первого цикла $\hat{\Delta}_1$ положительно обратного регенерирующего процесса Z – нерешетчатая. Тогда существует предельное распределение P_e процесса Z при $t \rightarrow \infty$, и для характеристики $f(Z) : E \rightarrow \mathbb{R}^+$ выполнено соотношение

$$M_e f(Z) = \frac{1}{M \hat{\Delta}_1} M_0 \int_0^{\hat{\Delta}_1} f(Z(s)) ds. \quad (6.1)$$

Заметим, что если функция $f(Z(s)) = I_{Z(s) \leq x}$ (x – произвольное фиксированное), то соотношение (6.1) дает предельное распределение процесса Z .

Классическая регенерация процесса $Z = \{Z_n, n \geq 1\}$ с дискретным временем определяется аналогично, а процесс восстановления β определен на множестве натуральных чисел $\mathbb{N}$. Для регенерирующего в классическом смысле процесса с дискретным временем циклы $G_k = \{Z_i\}_{\beta_k \leq i < \beta_{k+1}}$ независимы и одинаково распределены для всех $k \geq 1$. Соотношение (6.1) для процесса с дискретным временем принимает вид

$$M_e(f(Z)) = \frac{M_0 \left(\sum_{i=0}^{\beta_1-1} f(Z_i) \right)}{M \hat{\Delta}_1}. \quad (6.2)$$

В качестве моментов классической регенерации при исследовании систем часто используются моменты прихода заявки в пустую систему. Примером системы, в которой критерий стационарности гарантирует наступление такого опустошения, является односерверная система $GI/G/1$. Напомним, в такой системе в моменты времени $0 = t_0 < t_1 < \dots$ (образующие процесс восстановления) поступают клиенты. При этом клиенту i требуется S_i единиц времени для обслуживания на единственном сервере. Клиент i ожидает в общей неограниченной

очереди время $W_i \geq 0$, и по окончании обслуживания покидает систему. Времена между приходами $\{\tau_i := t_i - t_{i-1}\}$ и времена обслуживания $\{S_i\}$, $i \geq 1$ являются независимыми последовательностями н.о.р. с.в. Время ожидания каждого клиента определяется с помощью рекурсии Линдли:

$$W_{i+1} = (W_i + S_i - \tau_i)^+, \quad i \geq 1. \quad (6.3)$$

Стационарное время ожидания существует тогда и только тогда, когда

$$\mathbf{M}(\tau - S) > 0 \text{ или } \rho = \frac{\lambda}{\mu} < 1, \quad (6.4)$$

где $\mathbf{M}\tau = 1/\lambda$. $\mathbf{M}S = 1/\mu$. Как отмечено выше, всюду при использовании регенеративного метода в качестве моментов регенерации рассматриваются моменты прихода клиентов в пустую систему, именно:

$$\beta_k = \min\{i > \beta_{k-1} : W_i = 0\}, \quad k \geq 1 \quad (\beta_0 := 0).$$

Поскольку (6.4) влечет

$$\mathbf{P}(\hat{\Delta}_1 = 1) = \mathbf{P}(\tau - S > 0) > 0,$$

то с.в. $\hat{\Delta}_1$ – нерешетчатая. Отметим, что цикл регенерации в такой системе состоит из *периода занятости* сервера (в течение которого каждый приходящий клиент вынужден ожидать) и последующего *интервала простоя* сервера. Подчеркнем, что процесс $\{W_i\}_{i \geq 0}$ с моментами регенерации $\beta = \{\beta_k\}$ является положительно возвратным тогда и только тогда, когда выполнен критерий (6.4) [34, 124].

6.2. Метод регенеративных огибающих

Для исходной системы Σ вида $GI/G/N/\infty$ построим мажорантную систему $\bar{\Sigma}$ и минорантную систему $\underline{\Sigma}$. Входной поток во все три системы один и тот же (будем снабжать соответствующие величины в системах $\bar{\Sigma}$ и $\underline{\Sigma}$ верхним и нижним подчеркиванием, соответственно). Пусть t_n – момент прихода n -й

заявки в систему. Пусть

$$W_n = (W_n^{(1)}, \dots, W_n^{(N)}), \quad n \geq 1$$

вектор незавершенной нагрузки в системе Σ в момент времени t_n^- прихода n -й заявки. Обозначим через ν_n число заявок в системе в момент t_n^- , а Q_n – число заявок, ожидающих в очереди в момент времени t_n^- . Предположим, что существуют следующие пределы

$$Q_n \Rightarrow Q, \quad \nu_n \Rightarrow \nu, \quad W_n \Rightarrow W, \quad n \rightarrow \infty.$$

Сначала построим систему $\bar{\Sigma}$, и обозначим через $\bar{z}_n$ момент ухода n -й заявки. Заметим, что в общем случае для многосерверных систем порядок приходов и уходов заявок не совпадают. Пусть $\bar{\mathcal{M}}_n = \{i : t_i \leq t_n < \bar{z}_i\}$ – множество номеров заявок, которые обслуживаются в системе $\bar{\Sigma}$ в момент времени t_n и обозначим через $\bar{S}_i(n)$ остаточное время обслуживания i -й заявки в момент времени t_n^- . Зафиксируем произвольное целое число $\nu_0 \geq 0$, константы $0 \leq a \leq b < \infty$ и определим момент

$$\bar{\beta}_1 = \inf \left\{ k : \bar{\nu}_k = \nu_0, \bar{S}_i(k) \in (a, b), i \in \bar{\mathcal{M}}_k \right\}, \quad (6.5)$$

предполагая, что последовательности времен обслуживания во всех трех системах идентичны до момента $\bar{\beta}_1$. Тогда в момент $k = \bar{\beta}_1$, заменим все ненулевые остаточные времена обслуживания $\bar{S}_i(k), i \in \bar{\mathcal{M}}_k$, на верхнюю границу b . Таким образом, в момент $t_{\bar{\beta}_1}$ времена обслуживания заявок, находящихся на обслуживании, $\{\bar{S}_n\}$, увеличиваются, в то время как в исходной системе Σ они остаются неизменными. Рассмотрим процесс $\bar{\mathbf{X}}_n := \{\bar{\nu}_n, \bar{S}_i(n), i \in \bar{\mathcal{M}}_n\}, n \geq 1$. Этот процесс является регенерирующим. Для $n \geq 1$ определим моменты регенерации системы $\bar{\Sigma}$ следующим образом:

$$\bar{\beta}_{n+1} = \inf \left\{ k > \bar{\beta}_n : \bar{\nu}_k = \nu_0, \bar{S}_i(k) \in (a, b), \bar{\mathcal{M}}_{\bar{\beta}_n} \cap \bar{\mathcal{M}}_k = \emptyset, i \in \bar{\mathcal{M}}_k \right\}. \quad (6.6)$$

В каждый такой момент заменим остаточное время обслуживания на b . Таким образом, в моменты $\bar{\beta}_n$, исходный процесс переходит в фиксированное состояние $(\nu_0, \mathbf{b})$, где N -мерный вектор $\mathbf{b}$ содержит точно $\min(\nu_0, N)$ компонент b

и $(N - \nu_0)^+$ нулей. Условие $\overline{\mathcal{M}}_{\bar{\beta}_n} \cap \overline{\mathcal{M}}_k = \emptyset$ означает, что все заявки, обслуживаемые в момент $\bar{\beta}_n$ покидают систему до момента $\bar{\beta}_{n+1}$. Таким образом, на событии $\{\bar{\beta}_n = k\}$, распределение $\overline{\mathbf{X}}_k = (\nu_0, \mathbf{b})$ (в силу свойства регенерации) не зависит от k и предыстории $\{\overline{\mathbf{X}}_l, l < k\}$. Кроме того, пост-процесс $\{\overline{\mathbf{X}}_l, l \geq k\}$ имеет одинаковое распределение для всех k . Другими словами, процесс $\{\overline{\mathbf{X}}_n\}$ регенерирует в классическом смысле в моменты времени $\{\bar{\beta}_k\}$ и имеет н.о.р. циклы регенерации $\{\overline{\mathbf{X}}_k, \bar{\beta}_n \leq k < \bar{\beta}_{n+1}\}$ с н.о.р. длинами $\bar{\beta}_{n+1} - \bar{\beta}_n, n \geq 1$. Таким образом, системы $\Sigma, \bar{\Sigma}$ имеют одинаковое начальное состояние и, кроме того, $\bar{\tau}_n = \tau_n, n \geq 1, \bar{S} \underset{st}{\geq} S$. Тогда из теорем 4 и 5 [177] следует, что

$$\overline{W} \underset{st}{\geq} W, \quad \bar{\nu} \underset{st}{\geq} \nu, \quad \bar{Q} \underset{st}{\geq} Q. \quad (6.7)$$

Это позволяет использовать классическую регенерацию для построения верхней границы характеристик качества обслуживания системы Σ .

Аналогично можно построить минорантную систему $\underline{\Sigma}$, где в каждый момент $\{\underline{\beta}_n\}$, остаточные времена обслуживания заменяются на нижнюю границу a . Полученный результат сформулирован в следующей теореме.

Теорема 6.1. *Минорантная и мажорантная системы регенерируют в классическом смысле, а стационарное время ожидания $W, \underline{W}, \overline{W}$ и стационарная длина очереди $Q, \underline{Q}, \overline{Q}$ в системах $\Sigma, \underline{\Sigma}, \bar{\Sigma}$, соответственно, удовлетворяют неравенствам*

$$\underline{W} \underset{st}{\leq} W \underset{st}{\leq} \overline{W}, \quad \underline{Q} \underset{st}{\leq} Q \underset{st}{\leq} \overline{Q}. \quad (6.8)$$

Для доказательства отметим, что по построению системы $\Sigma, \bar{\Sigma}, \underline{\Sigma}$ имеют одинаковое начальное состояние, $\underline{\tau}_n = \bar{\tau} = \tau_n, n \geq 1$, а времена обслуживания имеют стохастический порядок $\bar{S} \underset{st}{\geq} S \underset{st}{\geq} \underline{S}$. Поэтому из теорем 4 и 5 [177] следует (6.8). Доказательство того, что построенные минорантная и мажорантная системы обладают свойством регенерации приведено перед формулировкой теоремы.

Сформулируем суть *метода регенеративных огибающих*.

1. Для исходной системы Σ построить минорантную $\underline{\Sigma}$ и мажорантную $\bar{\Sigma}$ системы, регенерирующие в классическом смысле.

2. Для систем $\underline{\Sigma}$ и $\bar{\Sigma}$ методом регенеративного оценивания вычислить доверительные оценки искомых характеристик. Полученные оценки являются со-

ответствующими нижней и верхней границами для искомой характеристики в исходной системе.

Пример 6.1. Рассмотрим систему $M/G/10$, в которой $M\tau = 1$, а времена обслуживания заданы распределением Парето I типа с плотностью

$$f_S(x) = \frac{\alpha k^\alpha}{x^{\alpha+1}}, \quad x \geq k, \quad (6.9)$$

где $\alpha > 1$, $k > 0$ и $MS = k\alpha/(\alpha - 1)$. Пусть коэффициент загрузки $\rho = 9$ и $\alpha = 2.5$. Тогда

$$k = \rho(\alpha - 1)/\alpha = 5.4. \quad (6.10)$$

Применим метод регенеративных огибающих для построения оценки средней длины стационарной очереди MQ . В качестве постоянных a , b (см. (6.6)) возьмем эмпирические квантили для остаточного времени обслуживания. Пусть a – 0.2-квантиль, так что 80% остаточных времен обслуживания будет не больше, чем a , и в качестве b – 0.8-квантиль. Выберем также $\nu_0 = 10$. Рисунок 6.1 показывает, что процессы длины очереди во всех трех системах $(\Sigma, \bar{\Sigma}, \underline{\Sigma})$ очень близки. Отметим, что в эксперименте с 10^4 заявок, ни одна заявка не встречает пустую систему, т. е. классические регенерации отсутствуют.

Пример 6.2. Рассмотрим систему $GI/G/10$, в которой времена обслуживания заданы распределением Парето I типа, где $\alpha = 2.5$ и k определяется формулой (6.10). Пусть τ имеет равномерное распределение на отрезке $[0.3, 0.9]$, что влечет $P(\tau > S) = 0$ и исключает классическую регенерацию. Выберем в качестве a и b 0.2- и 0.8-квантили остаточного времени обслуживания, соответственно. На Рисунке 6.2 приведен 95%-доверительный интервал для средней стационарной длины очереди в исходной системе, полученный как верхняя и нижняя граница в системах $\bar{\Sigma}$, и $\underline{\Sigma}$, соответственно.

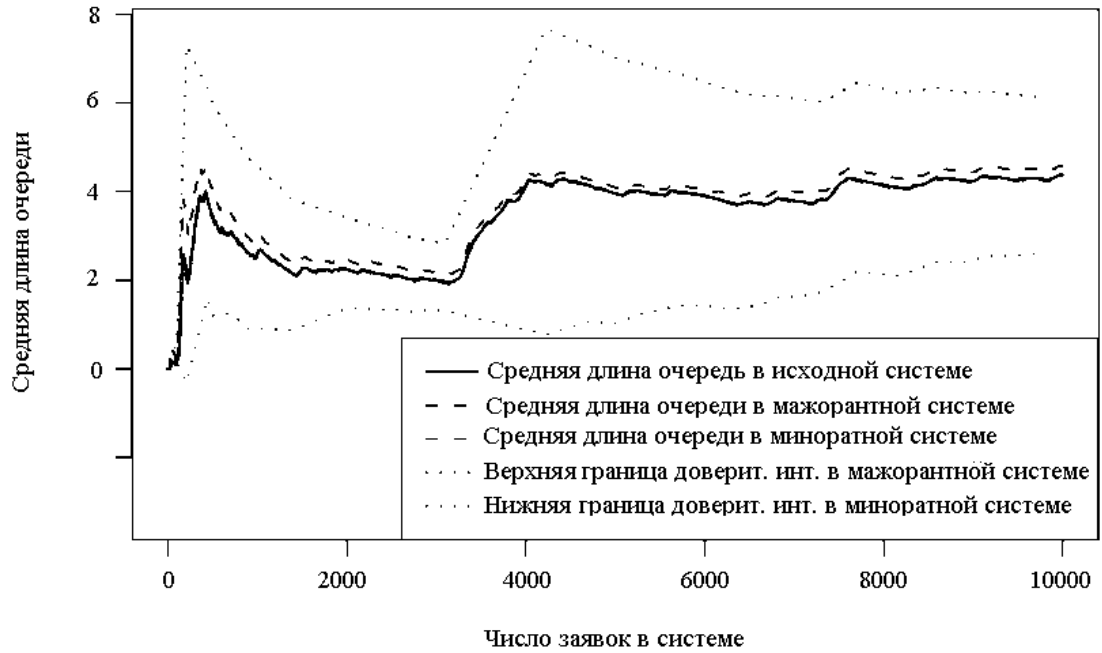


Рисунок 6.1. Система $M/G/10$ ($\rho = 9$) Оценки средних длин очередей и 95% доверительный интервал, построенный на основе доверительных интервалов для средней длины очереди в мажорантной и минорантной системах.

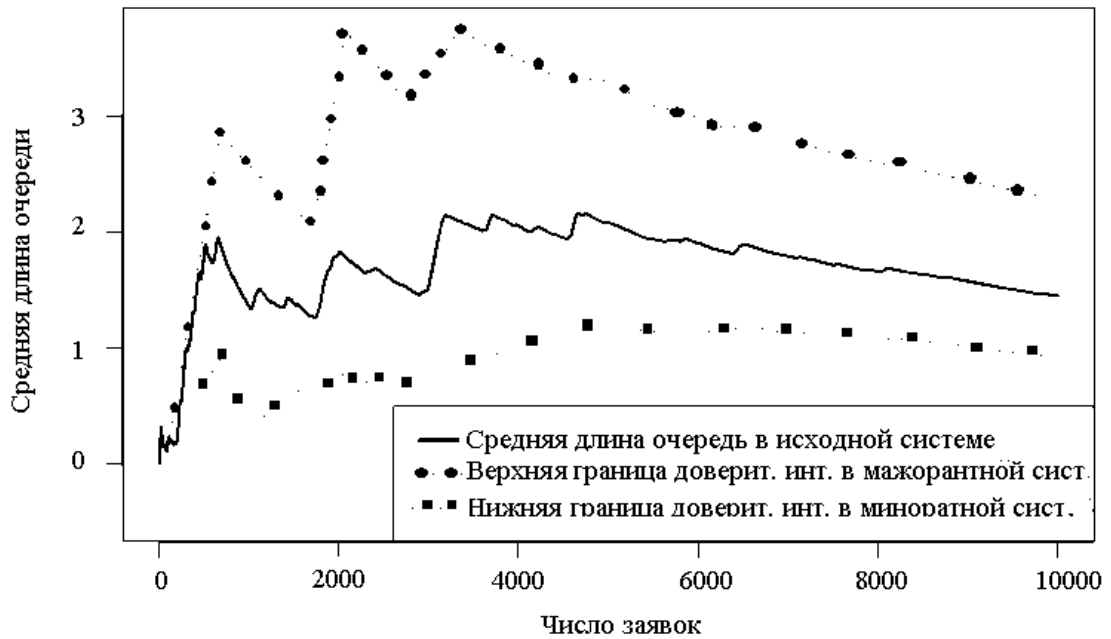


Рисунок 6.2. Система $GI/G/10$ ($\rho = 9$): 95% доверительный интервал для средней стационарной длины очереди.

6.3. Метод построения регенеративных огибающих для систем обслуживания со смешанными распределениями

Рассмотрим систему Σ (типа $GI/G/N/\infty$), в которой время обслуживания задано смесью распределений вида (1.31)

$$F_S(x) = p_1 F_{S^{(1)}}(x) + \dots + p_m F_{S^{(m)}}(x), \quad x \geq 0, \quad \sum_{k=1}^m p_k = 1, \quad (6.11)$$

где независимые с.в. $S^{(i)}$ заданы ф.р. $F_{S^{(i)}}(x)$. Входной процесс задан последовательностью н.о.р. интервалами между приходами $\tau_n = t_{n+1} - t_n$, $n \geq 1$. Наконец, предположим, что выполнено условие стационарности

$$MS < NM\tau. \quad (6.12)$$

Построим две новые системы: минорантную систему $\underline{\Sigma}$ со временем обслуживания $S^{(1)}$ и мажорантную систему $\bar{\Sigma}$ со временем обслуживания $S^{(m)}$. Обозначим через Q_n , $Q_n^{(1)}$, $Q_n^{(m)}$ – размер очереди, ν_n , $\nu_n^{(1)}$, $\nu_n^{(m)}$ – число заявок, W_n , $W_n^{(1)}$, $W_n^{(m)}$ – незавершенную работу в момент времени t_n^- в системах Σ , $\underline{\Sigma}$ и $\bar{\Sigma}$, соответственно. Пусть существуют стационарные пределы

$$Q_n^{(i)} \Rightarrow Q^{(i)}, \quad W_n^{(i)} \Rightarrow W^{(i)}, \quad n \rightarrow \infty, \quad i = 1, m.$$

Из Теоремы 2.3 следует, что, если для исходной системы Σ выполнены условия (2.23), то

$$Q^{(1)} \leq_{st} Q \leq_{st} Q^{(m)}, \quad (6.13)$$

$$W^{(1)} \leq_{st} W \leq_{st} W^{(m)}. \quad (6.14)$$

Следовательно, если системы $\underline{\Sigma}$ и $\bar{\Sigma}$ регенерируют в классическом смысле, то для исходной системы Σ , используя метод регенеративных огибающих, можно построить регенеративные оценки стационарного времени ожидания и стационарной длины очереди на основе оценок для систем $\underline{\Sigma}$ и $\bar{\Sigma}$.

6.4. Искусственная регенерация на основе метода расщепления

Как известно, метод вложенных цепей Маркова можно применить для построения различных типов регенераций в системах $M/G/1$ и $GI/M/1$ [34]. Так, фиксируя попадание процесса очереди в различные (фиксированные) состояния в моменты ухода (прихода) заявок в системах $M/G/1$ и $GI/M/1$ (т. е. в соответствующих вложенных цепях Маркова), можно получить несколько различных последовательностей точек регенерации, а следовательно, и построить несколько доверительных интервалов для оценки характеристики системы.

В данном разделе рассматривается более общая теоретическая конструкция, основанная на методе расщепления, который применяется для ускорения моделирования. Идея метода состоит в фиксации момента перехода процесса времени ожидания из состояния x в состояние y (не зависящее от x) с вероятностью $\delta(x)$ [29].

Рассмотрим обобщенную однородную цепь Маркова $Z = (Z_n, n \geq 0)$, заданную на измеримом пространстве $(\mathcal{E}, \mathcal{B})$ с начальным состоянием $Z_1 = x \in \mathcal{E}$. Предположим, что σ -алгебра $\mathcal{B}$ порождается счётным семейством подмножеств $\mathcal{E}$, а переходное одношаговое ядро определяется как

$$P(x, \cdot) = P(Z_{k+1} \in \cdot | Z_k = x) =: P_x(Z_k \in \cdot), \quad k \geq 0.$$

Предположим, что для некоторого множества $\mathcal{A} \in \mathcal{B}$, и каждого $x \in \mathcal{A}$ и $B \in \mathcal{B}$, существуют $\delta(x) \in (0, 1)$ и вероятностная мера φ на $(\mathcal{E}, \mathcal{B})$ такие, что выполнено следующее условие миноризации:

$$P(x, B) \geq \delta(x)\varphi(B). \tag{6.15}$$

Множество $\mathcal{A}$ и величина $\delta(x)$ называются множеством регенерации и мерой регенерации, соответственно [163].

Ключевое следствие условия (6.15) состоит в том, что после посещения множества $\mathcal{A}$ цепь с вероятностью $\delta(x)$ имеет (на следующем шаге) распределение φ , не зависящее от состояния x на предыдущем шаге. Это событие соответствует рандомизированной (или искусственной) регенерации цепи [29, 28].

Более того, условие (6.15) допускает следующее разложение переходного ядра:

$$P(x, B) = \delta(x)\varphi(B) + (1 - \delta(x))q(x, B), \quad (6.16)$$

где

$$q(x, B) = \frac{P(x, B) - \delta(x)\varphi(B)}{1 - \delta(x)}. \quad (6.17)$$

Согласно (6.16) состояние цепи Маркова на следующем шаге можно получить с помощью следующей процедуры расщепления ядра. Пусть I_x – с.в., принимающая значение 1 с вероятностью $\delta(x)$, и значение 0, иначе. Если цепь находится в состоянии $x \in \mathcal{E}$ и $I_x = 1$, то следующее состояние выбирается из распределения φ ; в противном случае — из распределения $q(x, \cdot)$.

Предположим, что цепь Маркова имеет стационарное распределение π , и обозначим через $K_\varphi(n)$ число моментов искусственной регенерации в интервале $[0, n]$. Известно [29], что при $n \rightarrow \infty$,

$$\frac{K_\varphi(n)}{n} \rightarrow \int_{\mathcal{E}} \delta(x)\pi(dx) := p_\varphi \quad \text{с в. 1}, \quad (6.18)$$

где $\pi(B) = \int_{\mathcal{E}} \pi(dx)P(x, B)$ стационарное распределение цепи Z . Таким образом, p_φ – асимптотическая частота моментов (рандомизированной) искусственной регенерации, полученных расщеплением переходного ядра.

6.4.1. Система $M/M/1$

Пусть интервалы между приходами заявок $\{\tau_k, k \geq 1\}$ заданы показательным распределением с параметром λ , а времена обслуживания $\{S_k, k \geq 1\}$ – показательным распределением с параметром μ . В [29] показано, что для последовательности времён ожидания $\{W_n\}$ в системе $M/M/1$ частота классических регенераций по мере увеличения времени моделирования/наблюдения приближается к $1 - \rho$. С другой стороны, в [29] также показано, что частота искусственных регенераций в системе $M/M/1$ (при $\lambda = 1$), полученная методом расщепления переходного ядра, приближается к $1 - \rho^2$. Таким образом, соотношение между

этими двумя частотами регенерации равно

$$\frac{1 - \rho^2}{1 - \rho} = 1 + \rho.$$

Это равенство указывает на то, что число рандомизированных регенераций может превышать число классических регенераций максимум в 2 раза при $\rho \uparrow 1$.

Далее докажем, что $K(n)/n \rightarrow 1 - \rho^2$ при произвольном $\lambda < \mu$. Используя рекурсию Линдли

$$W_{n+1} = (W_n + S_n - \tau_n)^+, \quad n \geq 0,$$

можем легко вывести, что

$$P(W_{n+1} = 0 | W_n = x) = \frac{\mu}{\mu + \lambda} e^{-\lambda x}, \quad x \geq 0$$

и

$$P(W_{n+1} \leq y | W_n = x) = \begin{cases} \frac{\mu}{\mu + \lambda} e^{-\lambda(x-y)}, & y < x; \\ 1 - \frac{\lambda}{\mu + \lambda} e^{-\mu(y-x)}, & y \geq x. \end{cases}$$

Дифференцируя по y условную ф.р. $P(W_{n+1} \leq y | W_n = x)$, получим

$$f(x, y) = \frac{\mu\lambda}{\mu + \lambda} e^{-\lambda(x-y)^+} e^{-\mu(y-x)^+}.$$

Пусть

$$\phi(y) := P(W_{n+1} \leq y | W_n = 0) = 1 - \frac{\lambda}{\mu + \lambda} e^{-\mu y}, \quad y \geq 0.$$

Известно, что для каждого x , оптимальное значение $\delta^*(x)$ величины $\delta(x)$, при котором частота моментов регенерации максимальна, определяется как [29]

$$\delta^*(x) := \min \left\{ \frac{P(W_{n+1} = 0 | W_n = x)}{\phi(0)}, \inf_{y > 0} \frac{f(x, y)}{f(0, y)} \right\}. \quad (6.19)$$

Найдем оптимальную вероятность регенерации (6.19) для $M/M/1$. Для этого вычислим

$$\frac{P(W_{n+1} = 0 | W_n = x)}{\phi(0)} = e^{-\lambda x}, \quad x \geq 0.$$

Далее заметим, что

$$\inf_{y>0} \frac{f(x, y)}{f(0, y)} = \begin{cases} e^{-\lambda x} e^{\lambda y} \geq e^{-\lambda x}, & y < x; \\ e^{\mu x} \geq e^{-\lambda x}, & y \geq x. \end{cases}$$

Следовательно,

$$\delta^*(x) = e^{-\lambda x}, \quad x \geq 0. \quad (6.20)$$

Заметим, что $\delta^*(0) = 1$. Далее, используя стационарное распределение времени ожидания

$$\pi(x) = \mathbf{P}(W \leq x) = 1 - \rho e^{-\mu(1-\rho)x}, \quad x \geq 0,$$

найдем

$$\pi(0) = 1 - \rho, \quad \pi'(x) = \rho\mu(1 - \rho)e^{-\mu(1-\rho)x}, \quad x > 0.$$

Тогда получаем требуемое

$$p_\phi = \int_0^\infty \pi(dx) \delta^*(x) = 1 - \rho + \int_0^\infty \rho\mu(1 - \rho)e^{-\mu(1-\rho)x} e^{-\lambda x} dx = 1 - \rho^2.$$

Таким образом, при значениях коэффициента загрузки близких к 1, частота моментов искусственной регенерации практически в 2 раза больше, чем классических 0-регенераций. Также легко найти распределение

$$q(x, y) = 1 - \frac{\frac{\lambda}{\mu + \lambda} e^{-\mu y} (e^{\mu x} - e^{-\lambda x})}{1 - e^{-\lambda x}}, \quad y \geq x,$$

$q(x, y) = 0$ для $y < x$.

6.4.2. Система *Pareto/M/1*

Рассмотрим односерверную систему *Pareto/M/1* с дисциплиной обслуживания FIFO. Пусть интервалы между приходами заявок $\{\tau_k, k \geq 1\}$ заданы распределением Парето

$$A(x) = 1 - \left(\frac{x_0}{x_0 + x} \right)^\alpha, \quad x \geq 0, x_0 > 0, \alpha > 1,$$

а времена обслуживания $\{S_k, k \geq 1\} \sim \text{Exp}(\mu)$,

$$B(x) = 1 - e^{-\mu x}, x \geq 0, \mu > 0.$$

Пусть выполнено условие стационарности, а именно

$$\rho = \frac{\alpha - 1}{x_0 \mu} < 1.$$

Для такой системы найдем оптимальную вероятность регенерации $\delta^*(x)$ и частоту моментов искусственной регенерации, получаемых методом расщепления.

Рассмотрим последовательность времен ожидания $\{W_k, k \geq 0\}$. Вычислим вероятность перехода последовательности из состояния x в состояние 0 для $x \geq 0$:

$$\begin{aligned} P(W_{n+1} = 0 | W_n = x) &= P(x + S_n - \tau_n < 0) = \\ &= \int_0^\infty P(\tau > z + x) \mu e^{-\mu z} dz = \\ &= \int_0^\infty \mu e^{-\mu z} \frac{x_0^\alpha}{(x_0 + x + z)^\alpha} dz. \end{aligned}$$

Выполнив замену переменной $u = \mu(x_0 + x + z)$; $du = \mu dz$, и подставив в последний интеграл, получим

$$\begin{aligned} P(W_{n+1} = 0 | W_n = x) &= x_0^\alpha \mu \int_{\mu(x_0+x)}^\infty e^{-u+\mu(x_0+x)} u^{-\alpha} \mu^\alpha du / \mu = \\ &= (x_0 \mu)^\alpha e^{\mu(x_0+x)} \Gamma(-\alpha + 1, \mu(x_0 + x)), \end{aligned}$$

где $\Gamma(s, x)$ – неполная верхняя Гамма-функция. Пусть $y \geq x$, тогда

$$\begin{aligned} P(W_{n+1} > y | W_n = x) &= P(\tau_n < S_n + x - y) = \\ &= \int_{y-x}^\infty P(\tau < z + x - y) \mu e^{-\mu z} dz = \end{aligned}$$

$$\begin{aligned}
&= \int_{y-x}^{\infty} \left(1 - \left(\frac{x_0}{x_0 + z + x - y} \right)^{\alpha} \right) \mu e^{-\mu z} dz = \\
&= e^{-\mu(y-x)} - \int_{y-x}^{\infty} \left(\frac{x_0}{x_0 + z + x - y} \right)^{\alpha} \mu e^{-\mu z} dz.
\end{aligned} \tag{6.21}$$

Сделаем замену переменной: $u = \mu(x_0 + z + x - y)$; $dz = du/\mu$, подставив в последний интеграл, можем записать (6.21) как

$$\begin{aligned}
&e^{-\mu(y-x)} - x_0^{\alpha} e^{\mu(x_0+x-y)} \int_{\mu x_0}^{\infty} e^{-u} (u/\mu)^{-\alpha} du/\mu = \\
&= e^{-\mu(y-x)} - (\mu x_0)^{\alpha} e^{-\mu(y-x-x_0)} \Gamma(-\alpha + 1, \mu x_0).
\end{aligned}$$

Следовательно, для $y \geq x$, получаем

$$P(W_{n+1} \leq y | W_n = x) = 1 - e^{-\mu(y-x)} + (\mu x_0)^{\alpha} e^{-\mu(y-x-x_0)} \Gamma(-\alpha + 1, \mu x_0).$$

Аналогично для $x > y$,

$$\begin{aligned}
P(W_{n+1} > y | W_n = x) &= P(\tau_n < S_n + x - y) = \\
&= \int_0^{\infty} P(\tau < z + x - y) \mu e^{-\mu z} dz = \\
&= \int_0^{\infty} \left(1 - \left(\frac{x_0}{x_0 + z + x - y} \right)^{\alpha} \right) \mu e^{-\mu z} dz.
\end{aligned}$$

Сделаем замену переменной $u = \mu(x_0 + z + x - y)$; $dz = du/\mu$ и, подставив в последний интеграл, получим

$$\begin{aligned}
P(W_{n+1} > y | W_n = x) &= 1 - x_0^{\alpha} \mu \int_{\mu(x+x_0-y)}^{\infty} e^{-u} e^{\mu(x+x_0-y)} (u/\mu)^{-\alpha} du/\mu = \\
&= 1 - (\mu x_0)^{\alpha} e^{\mu(x_0+x-y)} \Gamma(-\alpha + 1, \mu(x_0 + x - y)).
\end{aligned}$$

Следовательно, для $x > y$,

$$P(W_{n+1} \leq y | W_n = x) = (\mu x_0)^{\alpha} e^{\mu(x_0+x-y)} \Gamma(-\alpha + 1, \mu(x_0 + x - y)).$$

Таким образом, для всех $x, y \geq 0$

$$\begin{aligned} \mathbf{P}(W_{n+1} \leq y | W_n = x) &= 1 - e^{-\mu(y-x)^+} + \\ &+ (\mu x_0)^\alpha e^{\mu(x+x_0-y)} \Gamma(-\alpha + 1, \mu(x_0 + (x-y)^+)). \end{aligned} \quad (6.22)$$

Дифференцируя условную ф.р. (6.22) по y , получим

$$f(x, y) = \begin{cases} \mu \left(\frac{x_0}{x_0+x-y} \right)^\alpha - \mu(x_0\mu)^\alpha e^{\mu(x_0+x-y)} \times \\ \times \Gamma(-\alpha + 1, \mu(x_0 + x - y)), & y < x; \\ \mu e^{-\mu(y-x)} - \mu(x_0\mu)^\alpha e^{\mu(x_0+x-y)} \Gamma(-\alpha + 1, \mu x_0), & y \geq x. \end{cases}$$

Пусть $W_n \Rightarrow W$, где W – стационарное время ожидания в системе $GI/M/1$. Известно [63, 153], что распределение стационарного времени ожидания в системе $GI/M/1$ имеет вид

$$\pi(x) = \mathbf{P}(W \leq x) = 1 - \sigma e^{-\mu(1-\sigma)x}, \quad x \geq 0, \quad (6.23)$$

где σ – корень уравнения

$$z - \psi_A(\mu(1-z)) = 0$$

и $\psi_A(z)$ есть ПЛС ф.р. A . Поэтому

$$\pi(dx) = \sigma \mu (1 - \sigma) e^{-\mu(1-\sigma)x} dx, \quad \pi(0) = 1 - \sigma.$$

Теперь найдем оптимальную вероятность регенерации (6.19). Пусть

$$\phi(y) = \mathbf{P}(W_{n+1} \leq y | W_n = 0) = 1 - e^{-\mu y} + (\mu x_0)^\alpha e^{\mu(x_0-y)} \Gamma(-\alpha + 1, \mu x_0).$$

Тогда

$$\frac{\mathbf{P}(W_{n+1} = 0 | W_n = x)}{\phi(0)} = e^{\mu x} \frac{\Gamma(-\alpha + 1, \mu(x_0 + x))}{\Gamma(-\alpha + 1, \mu x_0)} \leq e^{\mu x}, \quad x \geq 0.$$

Далее, для всех $y \geq x$ легко вычислить, что

$$\frac{f(x, y)}{f(0, y)} = e^{\mu x}.$$

Аналогично, для $y < x$ находим

$$\frac{f(x, y)}{f(0, y)} = \frac{(x_0 + x - y)^{-\alpha} - \mu^\alpha e^{\mu(x_0 + x - y)} \Gamma(-\alpha + 1, \mu(x_0 + x - y))}{(x_0 - y)^{-\alpha} - \mu^\alpha e^{\mu(x_0 - y)} \Gamma(-\alpha + 1, \mu(x_0 - y))}.$$

Следовательно,

$$\begin{aligned} \delta^*(x) &= \min \left\{ \frac{\mathbf{P}(W_{n+1} = 0 | W_n = x)}{\varphi(0)}, \inf_{y > 0} \frac{f(x, y)}{f(0, y)} \right\} \\ &= \frac{\mathbf{P}(W_{n+1} = 0 | W_n = x)}{\varphi(0)} = e^{\mu x} \frac{\Gamma(-\alpha + 1, \mu(x_0 + x))}{\Gamma(-\alpha + 1, \mu x_0)}. \end{aligned}$$

Далее, подставив $\delta^*(x)$ и (6.23) в интеграл (6.18), получим

$$\begin{aligned} \int_0^\infty \pi(dx) \delta^*(x) &= 1 - \sigma + \int_0^\infty \sigma \mu (1 - \sigma) e^{-\mu(1-\sigma)x} \times \\ &\times e^{\mu x} \frac{\Gamma(-\alpha + 1, \mu(x_0 + x))}{\Gamma(-\alpha + 1, \mu x_0)} dx. \end{aligned} \quad (6.24)$$

Сделаем замены переменных в (6.24):

$$u = \Gamma(-\alpha + 1, \mu(x_0 + x)), \quad du = -(\mu(x_0 + x))^{-\alpha} e^{-\mu(x_0 + x)} \mu dx,$$

и

$$dv = e^{\mu \sigma x} dx, \quad v = e^{\mu \sigma x} / (\mu \sigma).$$

Теперь (6.24) преобразуется к виду

$$\begin{aligned} \int_0^\infty \pi(dx) \delta^*(x) &= \frac{(1 - \sigma) e^{-\mu \sigma x_0}}{\Gamma(-\alpha + 1, \mu x_0)} \int_{(1-\sigma)\mu x_0}^\infty e^{-u} \left(\frac{u}{1 - \sigma} \right)^{-\alpha} du / (1 - \sigma) \\ &= (1 - \sigma)^\alpha e^{-\mu \sigma x_0} \frac{\Gamma(-\alpha + 1, \mu x_0 - \sigma \mu x_0)}{\Gamma(-\alpha + 1, \mu x_0)}. \end{aligned} \quad (6.25)$$

Таким образом, частота моментов искусственной регенерации удовлетворяет

соотношению

$$\frac{K_\varphi(n)}{n} \rightarrow p_\varphi = (1 - \sigma)^\alpha e^{-\mu\sigma x_0} \frac{\Gamma(-\alpha + 1, \mu x_0 - \sigma\mu x_0)}{\Gamma(-\alpha + 1, \mu x_0)}$$

при $n \rightarrow \infty$ с в. 1.

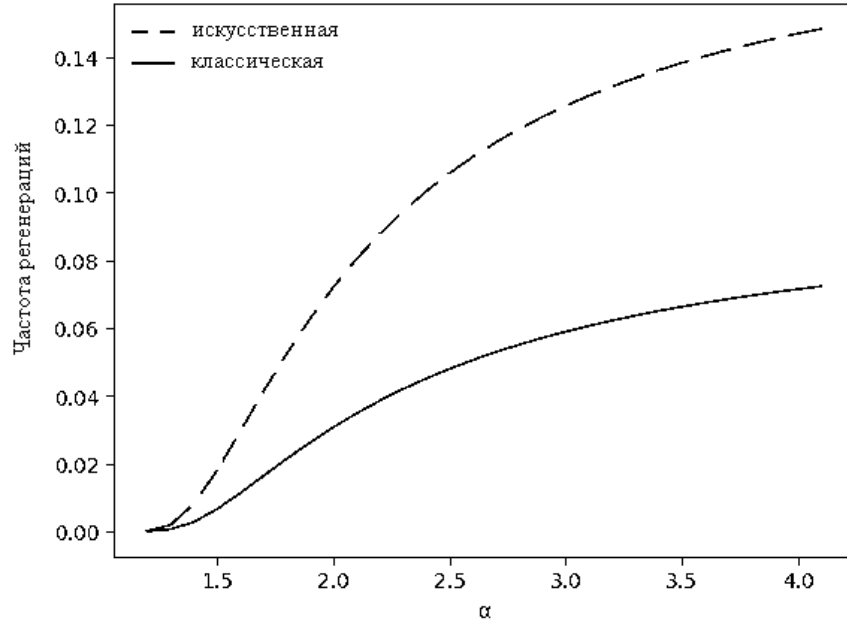


Рисунок 6.3. Сравнение частоты классической регенерации и частоты искусственной регенерации в системе $Pareto/M/1$ в зависимости от α .

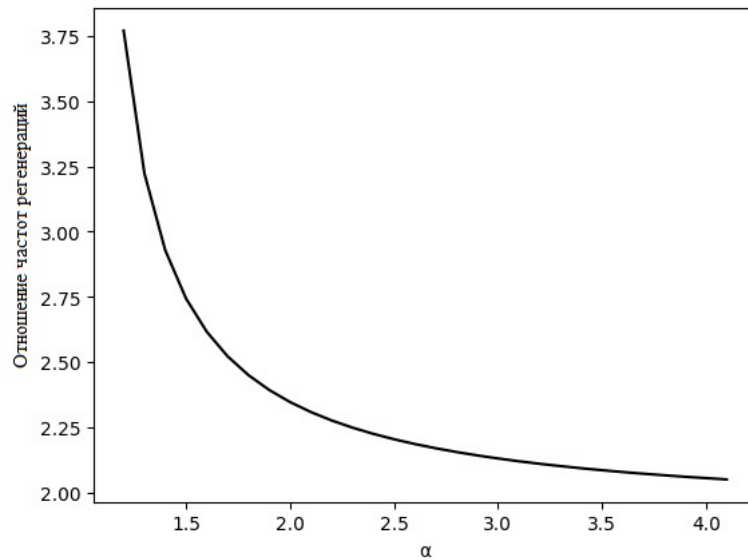


Рисунок 6.4. Отношение частоты искусственной регенерации к частоте классической регенерации в зависимости от α в системе $Pareto/M/1$.

На рисунке 6.3 сравниваются частота искусственной регенерации с частотой классической регенерации в зависимости от α для системы $Pareto/M/1$ с

параметрами $\lambda = 1$, $x_0 = 1$, $\rho = 0.9$.

Рисунок 6.4 иллюстрирует отношение частоты искусственной регенерации к частоте классической регенерации как функцию от параметра α для системы *Pareto/M/1* при $\lambda = 1$, $x_0 = 1$ и $\rho = 0.9$. При всех значениях α , частота искусственной регенерации по крайней мере в два раза больше, чем частота классической регенерации, и отношение растет с уменьшением значения α .

6.4.3. Система $H_2/M/1$

Построим искусственную регенерацию для системы $H_2/M/1$, где входной поток задан гиперэкспоненциальным распределением

$$A(x) = 1 - pe^{-\lambda_1 x} - (1-p)e^{-\lambda_2 x}, \quad \lambda_1, \lambda_2 > 0, \quad x \geq 0,$$

и выполнено условие стационарности

$$\rho = \frac{\lambda_1 \lambda_2}{\mu(p\lambda_2 + (1-p)\lambda_1)} < 1.$$

Как в предыдущем разделе вычислим вероятности

$$P(W_{n+1} = 0 | W_n = x) = \frac{\mu p}{\lambda_1 + \mu} e^{-\lambda_1 x} + \frac{\mu(1-p)}{\lambda_2 + \mu} e^{-\lambda_2 x}, \quad x \geq 0$$

и

$$P(W_{n+1} \leq y | W_n = x) = \begin{cases} \frac{p\mu}{\lambda_1 + \mu} e^{-\lambda_1(x-y)} + \frac{(1-p)\mu}{\lambda_2 + \mu} e^{-\lambda_2(x-y)}, & y < x; \\ 1 - \left(\frac{\lambda_1 p}{\mu + \lambda_1} + \frac{\lambda_2(1-p)}{\mu + \lambda_2} \right) e^{-\mu(y-x)}, & y \geq x. \end{cases}$$

Дифференцируя последнее соотношение по y , получим для $x, y \geq 0$

$$f(x, y) = \left(\frac{p\mu\lambda_1}{\lambda_1 + \mu} e^{-\lambda_1(x-y)^+} + \frac{(1-p)\mu\lambda_2}{\lambda_2 + \mu} e^{-\lambda_2(x-y)^+} \right) e^{-\mu(y-x)^+}.$$

Аналогично, подставляя

$$\phi(y) = P(W_{n+1} \leq y | W_n = 0) = 1 - \left(\frac{\lambda_1 p}{\mu + \lambda_1} + \frac{\lambda_2(1-p)}{\mu + \lambda_2} \right) e^{-\mu y},$$

получим

$$\frac{P(W_{n+1} = 0 | W_n = x)}{\phi(0)} = \frac{p(\lambda_2 + \mu)e^{-\lambda_1 x} + (1-p)(\lambda_1 + \mu)e^{-\lambda_2 x}}{p\lambda_2 + (1-p)\lambda_1 + \mu}, \quad x \geq 0.$$

При $y \geq x$

$$\frac{f(x, y)}{f(0, y)} = e^{\mu x}.$$

При $y < x$

$$g(x, y) := \frac{f(x, y)}{f(0, y)} = \frac{p\lambda_1(\lambda_2 + \mu)e^{-\lambda_1 x} + (1-p)\lambda_2(\lambda_1 + \mu)e^{-\lambda_2 x}e^{(\lambda_2 - \lambda_1)y}}{p\lambda_1(\lambda_2 + \mu) + (1-p)\lambda_2(\lambda_1 + \mu)e^{(\lambda_2 - \lambda_1)y}}.$$

Вычислим производную функции $g(x, y)$ по y , для $x, y > 0$

$$g'_y(x, y) = \frac{p(1-p)\lambda_1\lambda_2(\lambda_1 + \mu)(\lambda_2 + \mu)(\lambda_2 - \lambda_1)(e^{-\lambda_2 x} - e^{-\lambda_1 x})}{(p\lambda_1(\lambda_2 + \mu) + (1-p)\lambda_2(\lambda_1 + \mu)e^{(\lambda_2 - \lambda_1)y})^2}e^{(\lambda_2 - \lambda_1)y} < 0.$$

Следовательно, функция убывает по y и $\inf_y g(x, y) = \lim_{y \rightarrow \infty} g(x, y) = 1$. Вычислим оптимальную вероятность регенерации

$$\begin{aligned} \delta^*(x) &= \min \left\{ \frac{\mathbf{P}(W_{n+1} = 0 | W_n = x)}{\phi(0)}, \inf_{y > 0} \frac{f(x, y)}{f(0, y)} \right\} \\ &= \frac{p(\lambda_2 + \mu)e^{-\lambda_1 x} + (1-p)(\lambda_1 + \mu)e^{-\lambda_2 x}}{p\lambda_2 + (1-p)\lambda_1 + \mu}, \quad x \geq 0 \end{aligned} \quad (6.26)$$

и

$$\begin{aligned} q(x, y) &= 1 - e^{-\mu y} \left(\frac{p\lambda_1}{\mu + \lambda_1} + \frac{(1-p)\lambda_2}{\mu + \lambda_2} \right) \times \\ &\times \frac{e^{\mu x}(p\lambda_2 + (1-p)\lambda_1 + \mu) - (p(\lambda_2 + \mu)e^{-\lambda_1 x} + (1-p)(\lambda_1 + \mu)e^{-\lambda_2 x})}{p\lambda_2 + (1-p)\lambda_1 + \mu - p(\lambda_2 + \mu)e^{-\lambda_1 x} - (1-p)(\lambda_1 + \mu)e^{-\lambda_2 x}}. \end{aligned}$$

Далее находим частоту моментов искусственной регенерации

$$\begin{aligned} p_\phi &= \int_0^\infty \pi(dx) \delta^*(x) = 1 - \sigma + \int_0^\infty \sigma \mu (1 - \sigma) e^{-\mu(1-\sigma)x} \times \\ &\times \frac{p(\lambda_2 + \mu)e^{-\lambda_1 x} + (1-p)(\lambda_1 + \mu)e^{-\lambda_2 x}}{p\lambda_2 + (1-p)\lambda_1 + \mu} dx = \\ &= (1 - \sigma) + \frac{\mu \sigma (1 - \sigma)}{p\lambda_2 + (1-p)\lambda_1 + \mu} \left(\frac{p(\lambda_2 + \mu)}{\mu(1 - \sigma) + \lambda_1} + \frac{(1-p)(\lambda_1 + \mu)}{\mu(1 - \sigma) + \lambda_2} \right). \end{aligned}$$

Лемма 6.1. Частота моментов классической регенерации в системе $H_2/M/1$ меньше, чем частота моментов искусственной регенерации, т. е.

$$p_\phi \geq 1 - \sigma.$$

Доказательство. Пусть $\lambda_1 \leq \lambda_2$. Доказательство очевидно, поскольку

$$p_\phi \geq (1 - \sigma) \left(1 + \frac{(\lambda_1 + \mu)\mu\sigma}{(\lambda_2 + \mu)(\lambda_2 + \mu(1 - \sigma))} \right) \geq 1 - \sigma.$$

□

Заметим, что отношение частот искусственных и классических регенераций определяется как

$$\frac{p_\phi}{1 - \sigma} = 1 + \frac{\mu\sigma}{p\lambda_2 + (1 - p)\lambda_1 + \mu} \left(\frac{p(\lambda_2 + \mu)}{\mu(1 - \sigma) + \lambda_1} + \frac{(1 - p)(\lambda_1 + \mu)}{\mu(1 - \sigma) + \lambda_2} \right) \geq 1 + \frac{\mu}{\mu + \lambda_2}.$$

Пример 6.3. Пусть $p = 0.25$, $\lambda_1 = 2$, $\lambda_2 = 10$, $\mu = 12$. При таких параметрах $\rho = 0.417$, $\sigma = 0.559$. Следовательно, частота моментов классической регенерации в пределе равна $K(n)/n \rightarrow 1 - \sigma = 0.441$, а частота моментов искусственной регенерации равна $K_\phi(n)/n \rightarrow 0.7$. Построим для исходной системы две системы вида $M/M/1$. Пусть в минорантной системе входной поток пуассоновский с параметром λ_1 , а в мажорантной системе – пуассоновский с параметром λ_2 . В этом случае в минорантной системе

$$\rho_1 = 0.167, 1 - \rho_1 = 0.833, 1 - \rho_1^2 = 0.972.$$

В мажорантной системе

$$\rho_2 = 0.833, 1 - \rho_2 = 0.167, 1 - \rho_2^2 = 0.305.$$

Соответственно, частота моментов искусственной регенерации в минорантной системе больше, чем в исходной, а частота моментов искусственной регенерации в исходной системе больше, чем в мажорантной. Для всех систем частота моментов искусственной регенерации больше, чем частота моментов классической регенерации.

6.4.4. Сравнение частоты классических и искусственных регенераций

В этом разделе мы сравниваем число классических регенераций N_{Class} и число искусственных регенераций N_{Artif} в зависимости от коэффициента загрузки.

Обозначим через β_i i -й момент регенерации, тогда $\hat{\Delta}_i = \beta_{i+1} - \beta_i$ длина i -го цикла регенерации, $i = 1, \dots, N_{\text{Artif}}$.

В основе имитационного моделирования методом искусственной регенерации (для n заявок) лежит алгоритм, предложенный в работе [85]:

Шаг 1. Вычислим W_{j+1} , используя рекурсию Линдли:

$$W_{j+1} = (W_j + S_j - \tau_j)^+, \quad W_0 = 0.$$

Шаг 2. Выберем $x := W_j$, $y := W_{j+1}$ и вычислим $\delta(x)$, $\varphi(y)$, и $P(x, y)$.

Шаг 3. Сгенерируем с.в. I_x , которая принимает значение 1 с вероятностью

$$\frac{\delta(x)\varphi(y)}{P(x, y)}.$$

Если $I_x = 1$, то момент прихода $j + 1$ -й заявки это момент регенерации, $\beta_k = t_{j+1}$, и $\beta_0 = 0$ и увеличиваем счетчик $k \leftarrow k + 1$.

Шаг 4. Повторяем шаги 1–3 для $j \leftarrow j + 1$ до тех пор, пока $j + 1 = n$.

В Таблице 6.1 представлены результаты моделирования системы $M/M/1$ при $\lambda = 1$, $n = 10^5$ и различных значениях μ , которые показывают хорошую согласованность с теоретическими значениями. Так, частота классических регенераций, N_{Class}/n , хорошо согласуется с вероятностью классической регенерации, $1 - \rho$, в то время, как частота искусственных регенераций, N_{Artif}/n , близка к теоретическому значению $1 - \rho^2$. Например, при $\rho = 0.5$, частота классической регенерации $N_{\text{Class}}/n = 49.8 \times 10^3/10^5 = 0.498 \approx 1 - \rho = 0.5$, а частота моментов искусственной регенерации $N_{\text{Artif}}/n = 75.6 \times 10^3/10^5 = 0.756$, что близко к теоретическому значению $1 - \rho^2 = 0.75$.

В Таблице 6.2 приведены результаты моделирования системы $H_2/M/1$ при $\lambda_1 = 1$, $\lambda_2 = 3$, $p = 0.5$, $n = 10^5$ и различных значениях μ . Для этой системы также наблюдается хорошая согласованность эмпирической частоты классической регенерации, N_{Class}/n , с вероятностью $1 - \sigma$, и частоты искусственной регенерации, N_{Artif}/n , с вероятностью p_φ . Более того, при $\rho > 0.7$, число искусственных регенераций, N_{Artif} , приблизительно в 2 раза больше числа классических, N_{Class} .

В Таблице 6.3 представлены результаты моделирования для системы $Pareto/M/1$ при $x_0 = 1$, $\mu = 1$, $n = 10^4$ и различных значениях α .

Таблица 6.1. Число регенераций в системе $M/M/1$ ($\lambda = 1$).

ρ	N_{Class}	N_{Artif}	$1 - \rho$	$1 - \rho^2$
0.1	89.0×10^3	99.0×10^3	0.90	0.99
0.2	80.1×10^3	96.1×10^3	0.80	0.96
0.3	70.0×10^3	91.4×10^3	0.70	0.91
0.4	60.0×10^3	84.5×10^3	0.60	0.84
0.5	49.8×10^3	75.6×10^3	0.50	0.75
0.6	39.9×10^3	65.1×10^3	0.40	0.64
0.7	30.2×10^3	52.4×10^3	0.30	0.51
0.8	19.6×10^3	36.6×10^3	0.20	0.36
0.9	10.0×10^3	19.9×10^3	0.10	0.19
0.98	1.8×10^3	3.7×10^3	0.02	0.04

Таблица 6.2. Число регенераций в системе $H_2/M/1$.

ρ	N_{Class}	N_{Artif}	$1 - \sigma$	p_φ
0.1	87.1×10^3	98.4×10^3	0.871	0.984
0.2	75.0×10^3	94.2×10^3	0.751	0.940
0.3	64.0×10^3	88.1×10^3	0.638	0.875
0.4	53.3×10^3	80.0×10^3	0.533	0.793
0.5	43.5×10^3	70.1×10^3	0.434	0.695
0.6	34.0×10^3	59.8×10^3	0.340	0.582
0.7	24.6×10^3	46.3×10^3	0.251	0.456
0.8	16.9×10^3	33.1×10^3	0.164	0.317
0.9	7.9×10^3	17.0×10^3	0.081	0.165
0.98	1.6×10^3	3.6×10^3	0.016	0.034

Таблица 6.3. Число регенераций в системе *Pareto/M/1* ($\mu = 1, x_0 = 1$).

ρ	N_{Class}	N_{Artif}	$1 - \sigma$	p_φ
0.1	8.2×10^3	9.7×10^3	0.82	0.97
0.2	7.3×10^3	9.3×10^3	0.73	0.93
0.3	6.4×10^3	8.8×10^3	0.64	0.87
0.4	5.6×10^3	8.2×10^3	0.55	0.80
0.5	4.6×10^3	7.2×10^3	0.45	0.71
0.6	3.5×10^3	6.1×10^3	0.36	0.60
0.7	2.6×10^3	4.6×10^3	0.27	0.48
0.8	1.8×10^3	3.5×10^3	0.18	0.34
0.9	1.0×10^3	2.0×10^3	0.09	0.18
0.98	1.9×10^2	3.8×10^2	0.02	0.04

В Таблице 6.4 $\alpha = 2.1$, а μ меняется так, чтобы получать различные значения коэффициента загрузки ρ . Для больших значений ρ число регенераций (классических и искусственных) небольшое, поскольку $\alpha = 2.1$ мало, но достаточно для регенеративного моделирования.

Таблица 6.4. Число регенераций в *Pareto/M/1* ($\alpha = 2.1, x_0 = 1$).

ρ	N_{Class}	N_{Artif}	$1 - \sigma$	p_φ
0.1	8.2×10^3	9.7×10^3	0.82	0.97
0.2	6.9×10^3	9.1×10^3	0.67	0.90
0.3	5.3×10^3	8.1×10^3	0.54	0.80
0.4	4.2×10^3	7.0×10^3	0.42	0.69
0.5	3.2×10^3	5.8×10^3	0.32	0.56
0.6	2.3×10^3	4.5×10^3	0.22	0.44
0.7	1.4×10^3	3.1×10^3	0.15	0.31
0.8	9.3×10^2	2.0×10^3	0.09	0.18
0.9	350	860	0.034	0.080
0.98	44	106	0.005	0.012

Таблица 6.5. Длины циклов регенераций в $Pareto/M/1$ ($\rho = 0.9, x_0 = 1$).

α	3.0	2.5	2.1	1.9	1.7	1.5	1.4	1.3
$\hat{\Delta}_{Class}$	15	21	29	38	61	155	364	1714
$\hat{\Delta}_{Artif}$	7	10	12	16	24	56	123	520

Таблица 6.6. Оценки дисперсии в системе $Pareto/M/1$ при $\rho = 0.9$ и $x_0 = 1$.

α	Δ_{Class}	Δ_{Artif}	s_{Class}^2	s_{Artif}^2	$\frac{\hat{\eta}_{Class}^2}{4s_{Class}^2}$	$\frac{\hat{\eta}_{Artif}^2}{4s_{Artif}^2}$
3	15	7	0.7×10^4	0.7×10^4	6.2×10^6	5.9×10^6
2.5	21	10	3.0×10^4	2.7×10^4	5.9×10^6	4.2×10^6
2.1	29	12	2.0×10^5	1.9×10^5	2.3×10^7	2.1×10^7
1.9	38	16	1.9×10^5	1.8×10^5	7.6×10^7	2.3×10^7
1.7	61	24	2.1×10^6	2.0×10^6	7.8×10^8	3.8×10^8
1.5	155	56	2.3×10^7	2.2×10^7	2.9×10^9	2.2×10^9
1.4	364	123	1.2×10^8	1.1×10^8	5.6×10^{10}	2.8×10^{10}
1.3	1714	520	1.5×10^8	1.3×10^8	6.0×10^{10}	4.3×10^{10}

Таблица 6.5 показывает, что при малых α длина цикла искусственной регенерации существенно короче, а частота моментов больше, что важно при имитационном моделировании процессов при ограничении времени. Так, при $\alpha = 1.4$, средняя длина цикла классической регенерации 364, т.е. 29 циклов (в то время как циклов искусственной регенерации около 80) для $n = 10^4$ заявок.

Длина циклов регенерации для обоих типов регенерации, а также соответствующие оценки s^2 и η^2 (описание см. раздел 6.7.) представлены в Таблице 6.6 для системы $Pareto/M/1$ с $\rho = 0.9$ и различными значениями α , используя $n = 10^4$ значений. Мы обозначаем через $\hat{\Delta}_{Class}$ ($\hat{\Delta}_{Artif}$) среднюю длительность цикла при классической (искусственной) регенерации. Аналогично, s_{Class}^2 и $\hat{\eta}_{Class}^2$ (s_{Artif}^2 , $\hat{\eta}_{Artif}^2$) обозначают оценки σ^2 и η^2 , полученные с помощью классической (искусственной) регенерации.

Результаты моделирования в Таблице 6.6 показывают, что s^2 не зависит от типа регенерации, тогда как дисперсия η^2 меньше при использовании искусственной регенерации, что подтверждает теоретический тезис о том, что чем

Таблица 6.7. Оценка $P(W > u_w, Q > u_q)$ для $Pareto/M/1$ при $u_w = 5$, $u_q = 10$.

α	$P_W^{u_w}$	$\hat{P}_{Class}^{u_q}$	$\hat{P}_{Artif}^{u_q}$	$\hat{P}_{Class}^{u_w, u_q}$	$\hat{P}_{Artif}^{u_w, u_q}$
3.5	0.372	0.481	0.481	0.359	0.360
3.0	0.489	0.510	0.511	0.445	0.446
2.5	0.638	0.585	0.585	0.574	0.575
2.1	0.780	0.637	0.638	0.637	0.638
1.5	0.988	0.970	0.971	0.970	0.970

чаще моменты регенерации, тем меньше дисперсия η^2 [29].

Через Q_n обозначим число заявок в очереди в момент t_n^- . Используя технику регенеративного моделирования, найдем оценку совместной вероятности того, что произойдут одновременно два события: время ожидания W превысит заданный порог u_w , а очередь Q превысит порог u_q .

В Таблице 6.7 представлены результаты оценки вероятности $P(W > u_w, Q > u_q)$ для $Pareto/M/1$ при $\rho = 0.9$, $x_0 = 1$, $n = 10^3$, $u_w = 5$, и $u_q = 10$. Аналитических формул для вычисления подобных вероятностей не существует, а регенеративный метод позволяет построить соответствующую оценку с заданной точностью. Обозначим через $P_W^{u_w}$ теоретическое значение $P(W > u_w)$ и через $P_{Class}^{u_w, u_q}$ и $P_{Artif}^{u_w, u_q}$ – соответствующие оценки вероятности $P(W > u_w, Q > u_q)$, полученные на основе классической и искусственной регенерации, соответственно. Аналогично, пусть $\hat{P}_{Class}^{u_q}$ и $\hat{P}_{Artif}^{u_q}$ – оценки $P(Q > u_q)$.

Заметим, что оценки вероятностей $P(W > u_w, Q > u_q)$ и $P(Q > u_q)$ близки и не зависят от выбора типа регенерации, поскольку число регенераций достаточно для получения оценки на основе регенеративного метода..

6.5. Метод экспоненциального расщепления

В данном разделе рассмотрена другая конструкция искусственной регенерации, в которой вместо представления переходного ядра на основе (6.16) расщепляется плотность распределения интервалов между приходами заявок. Этот метод, насколько нам известно, был впервые для моделирования систем обслуживания представлен в работе [28] и основан на замене исходной с.в. стохастически

эквивалентной, представляющей собой смесь показательного распределенной с.в. и некоторой другой, имеющей специальный вид. Свойство потери памяти показательного распределения позволяет построить регенеративную структуру и применить регенеративное моделирование для доверительного оценивания характеристик систем обслуживания.

Предположим, что абсолютно непрерывная с.в. T задана плотностью распределения g_T , которая удовлетворяет неравенству (аналог условия миноризации (6.15) для цепей Маркова)

$$g_T \geq \delta g_0, \quad (6.27)$$

где $\delta \in (0, 1)$ – некоторая константа, g_0 – некоторая плотность распределения. Определим с.в. I_T так, что $P(I_T = 1) = \delta$ и будем называть ее индикатором расщепления. Построим новую с.в. T' в виде смеси двух с.в. T_0 и T_1 , а именно:

$$T' = I_T T_0 + (1 - I_T) T_1, \quad (6.28)$$

где с.в. T_0 задана плотностью g_0 , а с.в. T_1 задана плотностью

$$g_1 = \frac{f - \delta g_0}{1 - \delta}. \quad (6.29)$$

С.в. T' является стохастической копией с.в. T , и в результате представления (6.28) ее плотность удовлетворяет свойству (6.27) с вероятностью не меньше, чем δ .

Будем говорить, что с.в. T экспоненциально расщеплена, если существуют константы $\eta > 0$ и $\delta \in (0, 1)$ такие, что выполнено неравенство

$$g(x) \geq \delta \eta e^{-\eta x}, \quad x \geq 0. \quad (6.30)$$

Таким образом, новая с.в. T' представляет собой смесь с.в. $T_0 \sim \text{Exp}(\eta)$ (с плотностью распределения $g_0(x) = \eta e^{-\eta x}$, $x \geq 0$), и с.в. T_1 с плотностью распределения

$$g_1(x) = \frac{g(x) - \delta \eta e^{-\eta x}}{1 - \delta}, \quad x \geq 0. \quad (6.31)$$

6.5.1. Экспоненциальное расщепление распределения Парето

Лемма 6.2. *Предположим, что с.в. T задана распределением $Pareto(\alpha, x_0)$ с плотностью*

$$g(x) = \alpha x_0^\alpha (x_0 + x)^{-(\alpha+1)}, \quad \alpha > 0, \quad x_0 > 0, \quad x \geq 0. \quad (6.32)$$

Тогда если выполнено условие

$$\alpha + 1 \leq \eta x_0 \leq \alpha / \delta, \quad (6.33)$$

где η и δ определены формулами (6.30), то с.в. T' имеет распределение (6.28), с.в. T_0 имеет показательное распределение с параметром η , и выполнено неравенство (6.30). Кроме того, ф.р. с.в. T_1 имеет вид

$$G_1(x) = 1 - \frac{1}{1 - \delta} \left(\frac{x_0}{x_0 + x} \right)^\alpha + \frac{\delta}{1 - \delta} e^{-\eta x}, \quad x \geq 0, \quad (6.34)$$

с плотностью g_1 , удовлетворяющей соотношению (6.31).

Доказательство. Рассмотрим с.в. $X \sim Exp(\eta)$ и с.в. $Y \sim Pareto(\alpha + 1, x_0)$. Из условия (6.33) следует, что

$$\eta \geq \frac{\alpha + 1}{x_0},$$

а, следовательно, с.в. X и Y упорядочены по интенсивности отказов, $X \leq_r Y$ (см. пример 1.7), что влечет выполнение следующих неравенств для всех $x \geq 0$:

$$r_X(x) = \eta \geq \frac{\alpha + 1}{x_0} \geq \frac{\alpha + 1}{x_0 + x} = r_Y(x), \quad x \geq 0,$$

которые, в свою очередь, приводят к стохастическому порядку, $X \leq Y$ [145]. Это означает, что хвосты ф.р. удовлетворяют неравенству

$$e^{-\eta x} \leq \left(\frac{x_0}{x_0 + x} \right)^{\alpha+1}, \quad x \geq 0.$$

Если $x_0 \delta \eta \leq \alpha$, то выполнено неравенство

$$\frac{x_0 \delta \eta}{\alpha} e^{-\eta x} \leq \left(\frac{x_0}{x_0 + x} \right)^{\alpha+1}.$$

Следовательно,

$$\delta\eta e^{-\eta x} \leq \frac{\alpha}{x_0} \left(\frac{x_0}{x_0 + x} \right)^{\alpha+1} = g(x),$$

и неравенство (6.30) выполнено. $\square$

На Рисунке 6.5, демонстрирующем роль условий (6.33), показаны графики плотности распределения Парето (с параметрами $x_0 = 1$ и $\alpha = 3$) и показательного распределения $\delta\eta e^{-\eta x}$ при различных значениях η и δ . Видно, что плот-

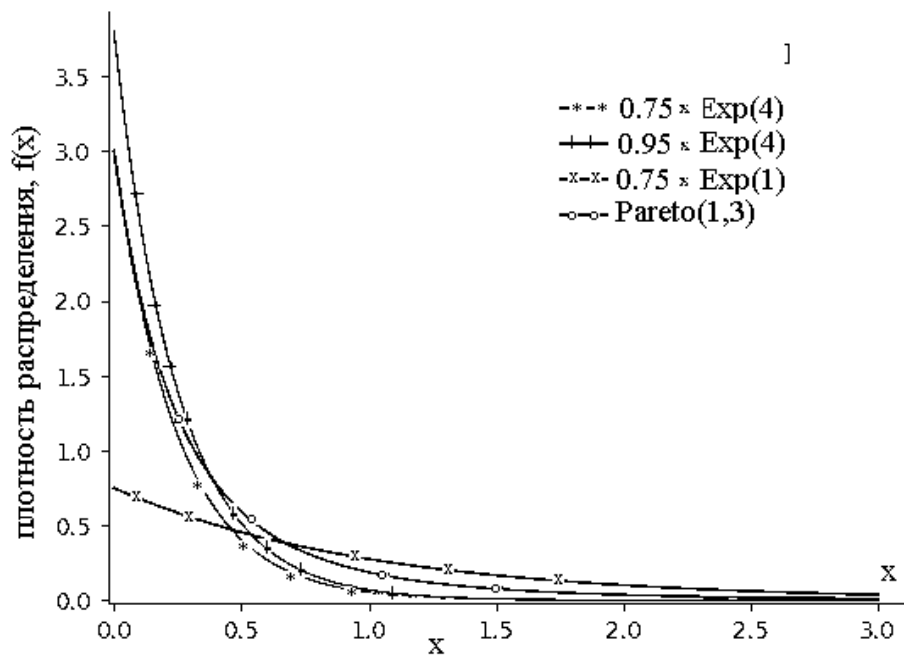


Рисунок 6.5. Графическая интерпретация неравенств (6.33).

ность показательного распределения сверху ограничена плотностью распределения Парето только при выполнении неравенств (6.33) (при $\delta = 0.75, \eta = 4$). Две кривые (при $\delta = 0.75, \eta = 1$ и $\delta = 0.95, \eta = 4$) не удовлетворяют неравенствам (6.33) и, соответственно, для них не выполнено условие расщепления (6.30).

6.6. Метод искусственной регенерации в системах обслуживания с суперпозицией входных потоков

Рассмотрим односерверную систему обслуживания $\sum_{i=1}^m GI_i/G/1$, в которой входной поток представляет собой суперпозицию m независимых стационарных

входных процессов восстановления. Пусть i -й поток задан н.о.р. интервалами между приходами заявок $\{\tau_k^{(i)} = t_{k+1}^{(i)} - t_k^{(i)}, k \geq 0\}$ и пусть $\{t_k, k \geq 1\}$ – моменты прихода в общий входной поток. Для построения моментов регенерации входного процесса $\{\gamma_k\}$, выберем и зафиксируем входной поток с номером i_0 (с моментами прихода $\{t_k^{(i_0)} \equiv t_k^{(0)}, k \geq 1\}$). Будем говорить, что j -й входной поток имеет экспоненциальную фазу расщепления, если с.в. $\tau_k^{(j)}$ представима в виде (6.28) с параметрами расщепления η_j, δ_j , для ее плотности выполнено неравенство (6.30) и $\delta_j = 1$. Рассмотрим последовательность событий:

$$\mathcal{E}_k^{(0)} = \{ j\text{-й входной поток имеет экспоненциальную фазу расщепления} \\ \text{в моменты времени } t_k^{(0)}, \forall j \neq i_0 \}, k \geq 1. \quad (6.35)$$

Это событие означает, что в момент поступления k -й заявки из входного потока с номером i_0 , времена между приходами заявок всех остальных потоков находятся в экспоненциальной фазе (т. е. индикаторы расщепления для них равны 1).

Для каждого целого $k \geq 1$ введем индекс $n_0(k)$ такой, что $t_k = t_{n_0(k)}^{(0)}$. Моменты регенерации $\{\gamma_k\}$ входного потока, образованного суперпозицией входных потоков, определим следующим образом: $\gamma_0 = 0$,

$$\gamma_{k+1} = \min\{i : I(t_{n_0(i)}^{(0)} > \gamma_k) \cdot I(\mathcal{E}_{n_0(i)}^{(0)} = 1)\}, k \geq 0. \quad (6.36)$$

Легко проверить, что разность $\Delta_k := \gamma_{k+1} - \gamma_k$ – число приходов из общего входного процесса (образованного суперпозицией) между событиями $\mathcal{E}_k^{(0)}$ и $\mathcal{E}_{k+1}^{(0)}$. Моменты времени $\{\gamma_k\}$ являются моментами классической регенерации, а $\{\Delta_k\}$ – н.о.р. (дискретными) длинами циклов регенерации. Обозначим W_n – время ожидания n -й заявки в очереди.

Теорема 6.2. *Случайный процесс $W = \{W_n, n \geq 1\}$ регенерирует в классическом смысле, если выполнены два условия:*

- 1) *множество моментов регенерации входного потока вида (6.36) не пусто, $\{\gamma_k, k \geq 1\} \neq \emptyset$;*
- 2) *в моменты регенерации входного потока система свободна, т.е.*

$$\beta_0 = 0, \beta_{k+1} = \min_i \{\gamma_i > \beta_k : W_{\gamma_i} = 0\}, k \geq 0, \quad (6.37)$$

С.в. $\hat{\Delta}_k := \beta_{k+1} - \beta_k, k \geq 1$ являются н.о.р. длинами циклов с моментами регенерации вида (6.37).

Для доказательства теоремы отметим, что по построению в моменты (6.37) система пуста, следовательно, распределение случайной пары $((W_{\beta_k+t})_{t \geq 0}, (\beta_k - \beta_i)_{i \geq k})$ не зависит от $k \geq 0$ и не зависит от предыстории $\{(Z_t)_{t < \beta_k}, \beta_0, \dots, \beta_k\}, k \geq 0$.

Пусть входной поток состоит из суперпозиции $m \geq 2$ потоков $\tau^{(j)}, j = 1, \dots, m$, заданных распределениями $Pareto(\alpha_j, x_0), j = 1, \dots, m$.

Приведем алгоритм, реализующий описанный выше метод расщепления.

Алгоритм построения моментов регенерации методом экспоненциального расщепления в системах обслуживания с суперпозицией входных потоков, имеющих распределения Парето.

Шаг 1. Ввод данных. Ввод числа входных потоков $m \geq 2$, параметров распределений входных потоков $\alpha_j > 2, x_0$, параметров распределения времени обслуживания с ф.р. B_j , числа заявок K .

Шаг 2. Генерация суперпозиции $m \geq 2$ входных потоков.

Генерируем с.в. $\tau_n^{(j)}, j = 1, \dots, m, n \geq 1$, в соответствии с распределением $Pareto(x_0, \alpha_j), j = 1, \dots, m$.

Вычисляем моменты прихода заявок из каждого входного потока $t_n^{(j)}, j = 1, \dots, m, n \geq 1$.

Сортируем моменты прихода в общий входной поток по возрастанию $\{t_k, k \geq 1\} = \text{sort}(t_n^{(j)})\}, j = 1, \dots, m, n \geq 1$ и формируем общий входной поток.

Вычисляем интервалы между приходами заявок в общем входном потоке $\tau_k = t_{k+1} - t_k, k \geq 1$.

Шаг 3. Генерация времени обслуживания. Генерируем с.в. $s_n^{(j)}, j = 1, \dots, M, n \geq 1$, в соответствии с ф.р. B_j .

Шаг 4. Выбираем входной поток $1 \geq i_0 \geq m$.

Шаг 5. Выбираем константы $\delta_j \in (0, 1), \eta_j > 0$ так, чтобы были выполнены для всех входных потоков $j \neq i_0$ условия

$$\eta_j x_0 \geq \alpha_j + 1, \quad \delta_j \eta_j x_0 \leq \alpha_j, \quad j \neq i_0.$$

Шаг 6. Генерируем с.в. $I_n^{(j)}, j \neq i_0$, которые принимают значение 1 с веро-

яностями δ_j и 0, иначе.

Шаг 7. Если $I_n^{(j)} = 1$, то генерируем $\tau_n^{(j)}$ в соответствии с распределением $Exp(\eta_j)$, иначе $\tau_n^{(j)}$ генерируем, используя ф.р.

$$G_1^{(j)}(x) = 1 - \frac{1}{1 - \delta_j} \left(\frac{x_0}{x_0 + x} \right)^{\alpha_j} + \frac{\delta_j}{1 - \delta_j} e^{-\eta_j x}, \quad x \geq 0.$$

Шаг 8. Для каждого целого $k \geq 1$ запоминаем моменты прихода из потока i_0 в общий входной поток τ , т.е. фиксируем номера заявок $n_0(k)$ такие, что $t_k = t_{n_0(k)}^{(0)}$.

Шаг 9. Вычисляем моменты регенерации входного потока $\{\gamma_l\}$ рекурсивно : $\gamma_0 = 0$,

$$\gamma_{l+1} = \min\{i : I(t_{n_0(i)}^{(i_0)} > \gamma_l) \cdot I(\mathcal{E}_{n_0(i)}^{(i_0)} = 1)\}, \quad l \geq 1. \quad (6.38)$$

Шаг 10. Вычисляем времена ожидания по формуле:

$$W_0 = 0, \quad W_{k+1} = \max(W_k + s_k - \tau_{k+1}; 0).$$

Шаг 11. Вычисляем моменты регенерации процесса времени ожидания по формуле:

$$\beta_0 = 0, \quad \beta_{h+1} = \min\{\gamma_l > \beta_h : W_{\gamma_l} = 0\}, \quad h \geq 0,$$

Шаг 12. Вычисляем длины циклов регенерации $\hat{\Delta}_h := \beta_{h+1} - \beta_h, h \geq 1$.

Пример 6.4. Чтобы оценить влияние процесса i_0 , параметров распределения расщепления δ_j и η_j и интенсивности трафика ρ был проведен ряд численных экспериментов по моделированию системы с суперпозицией двух входных процессов, заданных распределением Парето. Пусть

$$\tau^{(1)} \sim Pareto(\alpha_1 = 2, x_0 = 1), \quad \tau^{(2)} \sim Pareto(\alpha_1 = 2, x_0 = 1),$$

а время обслуживания $S \sim Exp(\mu)$.

Выберем первую компоненту суперпозиции входного процесса (с параметром $\alpha_1 = 2$) в качестве фиксированного потока i_0 -процесса и применим метод экспоненциального расщепления ко второму потоку (с параметром $\alpha_2 = 3$).

В следующих таблицах приведены результаты численных экспериментов, где через RegIn обозначено число регенераций входного потока, т. е. число со-

бытий (6.36), а через RegW – число регенераций вида (6.37) процесса W .

Таблица 6.8. Число регенераций RegIn и RegW и полуширина доверительного интервала 99%CI при различных δ и η .

$i_0 = 1$ и $\mu = 4$ ($\rho = 0.75$).				
δ	η	RegIn	RegW	99%CI
.10	4	166.65×10^6	13.5×10^6	4.59×10^{-4}
0.33	4	555.61×10^6	45.1×10^6	4.55×10^{-4}
0.50	4	833.40×10^6	67.5×10^6	4.58×10^{-4}
0.75	4	1250.05×10^6	101.3×10^6	4.58×10^{-4}
0.10	5	133.29×10^6	9.1×10^6	4.34×10^{-4}
0.33	5	444.41×10^6	30.4×10^6	4.56×10^{-4}
0.50	5	666.67×10^6	45.6×10^6	4.55×10^{-4}
0.10	6	111.11×10^6	6.5×10^6	4.46×10^{-4}

В Таблице 6.8 интенсивность обслуживания $\mu = 4$ (т. е. $\rho = 0,75$), а в Таблице 6.9 $\mu = 6$ (т. е. $\rho = 0,5$). В обеих таблицах оценки среднего времени ожидания одинаковы для всех вариантов выбора параметров распределения и совпадают с результатами, полученными с помощью традиционного моделирования (1,358 и 0,2597 для $\mu = 4$ и $\mu = 6$, соответственно). Чтобы оценить качество оценок, в последнем столбце приведены значения половины ширины 99% доверительного интервала для среднего времени ожидания, с использованием моментов регенерации RegW (число наблюдений равно 10^{10}).

Сравнивая Таблицы 6.8 и 6.9, заметим, что при фиксированных значениях δ и η число точек регенерации входного процесса RegIn примерно одинаково. Этот факт объясняется тем, что входной процесс не зависит от скорости обслуживания μ . С другой стороны, число моментов регенерации процесса времени ожидания RegW увеличивается с ростом μ , так как, чем меньше коэффициент загрузки ρ , тем больше вероятность простоя сервера. Это увеличение, вместе с уменьшением среднего времени ожидания, приводит к разнице в ширине доверительного интервала примерно в 10 раз.

Наконец, результаты моделирования в случае фиксации первой компоненты (при $\alpha_1 = 2$) представлены в Таблицах 6.10 и 6.11, которые подтверждают

Таблица 6.9. Число регенераций RegIn и RegW и полуширина доверительного интервала 99%CI при различных δ и η .

$i_0 = 1$ и $\mu = 6$ ($\rho = 0.5$).				
δ	η	RegIn	RegW	99%CI
0.10	4	166.67×10^6	42.7×10^6	4.29×10^{-5}
0.33	4	555.61×10^6	142.4×10^6	4.59×10^{-5}
0.50	4	833.32×10^6	213.6×10^6	4.20×10^{-5}
0.75	4	1250.06×10^6	320.5×10^6	4.57×10^{-5}
0.10	5	133.34×10^6	29.8×10^6	4.47×10^{-5}
0.33	5	444.41×10^6	99.5×10^6	4.27×10^{-5}
0.50	5	666.60×10^6	149.3×10^6	4.55×10^{-5}
0.10	6	111.10×10^6	22.1×10^6	4.58×10^{-5}

предыдущие наблюдения относительно RegIn, RegW и ширины доверительного интервала.

По результатам моделирования можно сделать следующие замечания относительно выбора параметров расщепления. С одной стороны, чем больше параметр δ (для фиксированного значения η), тем чаще встречается экспоненциальная фаза. В результате RegIn также увеличивается. С другой стороны, чем больше параметр η (для фиксированного δ), тем меньше число регенераций RegIn. Более того, результаты моделирования показывают, что наибольшее число регенераций RegW системы обслуживания достигается при выполнении равенства в обоих неравенствах (6.33) для всех входных потоков $i \neq i_0$.

В [38] упоминается, что более частые регенерации обычно сокращают требуемое время моделирования. Это замечание может оказаться полезным, когда время моделирования ограничено, и тогда число полученных регенераций становится критически важным для построения требуемого доверительного интервала.

Таблица 6.10. Число регенераций RegIn и RegW и полуширина доверительного интервала 99%CI при различных δ и η .

$i_0 = 2$ и $\mu = 4$ ($\rho = 0.75$).				
δ	η	RegIn	RegW	99%CI
0.50	3	1111.11×10^6	88.7×10^6	4.56×10^{-4}
0.10	4	166.65×10^6	11.1×10^6	4.51×10^{-4}
0.33	4	555.57×10^6	37.1×10^6	4.61×10^{-4}
0.50	4	833.35×10^6	55.6×10^6	4.55×10^{-4}
0.10	6	111.11×10^6	5.6×10^6	4.54×10^{-4}
0.20	6	222.18×10^6	11.3×10^6	4.60×10^{-4}
0.33	6	370.41×10^6	18.8×10^6	4.56×10^{-4}

Таблица 6.11. Частота регенераций RegIn и RegW и полуширина доверительного интервала 99%CI при различных δ и η .

$i_0 = 2$ и $\mu = 6$ ($\rho = 0.5$).				
δ	η	RegIn	RegW	99%CI
0.50	3	1111.14×10^6	282.5×10^6	4.52×10^{-5}
0.10	4	166.65×10^6	36.7×10^6	4.41×10^{-5}
0.33	4	555.53×10^6	122.4×10^6	4.23×10^{-5}
0.50	4	833.36×10^6	183.6×10^6	4.18×10^{-5}
0.10	6	111.09×10^6	19.5×10^6	3.84×10^{-5}
0.20	6	222.21×10^6	39.0×10^6	4.41×10^{-5}
0.33	6	370.35×10^6	65.0×10^6	4.43×10^{-5}

6.6.1. Суперпозиция входных потоков, заданных распределением Вейбулла

Пусть входной процесс представляет собой суперпозицию независимых входных потоков, заданных н.о.р. интервалами между приходами заявок $\{\tau_k^{(i)}, k \geq 1\}$ с ф.р.

$$A_i(x) = 1 - e^{-(x/\alpha_i)^{\beta_i}}, \quad x \geq 0, \quad \alpha_i > 0, \quad \beta_i > 0, \quad i = 1, \dots, m.$$

Подставим плотность распределения Вейбулла в неравенство (6.30), получим

$$\frac{\beta_j}{\alpha_j}(x/\alpha_j)^{\beta_j-1}e^{-(x/\alpha_j)^{\beta_j}} \geq \delta_j\eta_j e^{-\eta_j x}, \quad (6.39)$$

которое выполняется при условиях:

$$0 < \delta_j \leq \beta_j/\alpha_j, \quad \eta_j = \alpha_j^{-\beta_j}, \quad 0.5 \leq \beta_j < 1, \quad \alpha_j \geq 1. \quad (6.40)$$

Найдем распределение с.в. T_1^j , участвующей в расщеплении (6.27) как

$$G_1^j(x) = 1 - \frac{1}{1 - \delta_j}e^{-(x/\alpha_j)^{\beta_j}} + \frac{\delta_j}{1 - \delta_j}e^{-\eta_j x}, \quad x \geq 0.$$

Пример 6.5. Для оценки влияния выбора процесса i_0 , параметра расщепления δ , параметра η и интенсивности трафика ρ был проведен ряд численных экспериментов для $\alpha_1 = 1.5$, $\beta_1 = 0.5$, $\alpha_2 = 2$ и $\beta_2 = 2/3$ и для 10^9 заявок.

Таблица 6.12. Число регенераций RegW, Среднее время ожидания Mean(W) и полуширина доверительного интервала 99%CI при $i_0 = 2$, $\eta = 0.8164$, $\beta_1 = 0.5$, $\beta_2 = 0.67$, $\alpha_1 = 1.5$, $\alpha_2 = 2$.

ρ	$\delta = 0.33$			$\delta = 0.1$		
	RegW	Mean(W)	99%CI	RegW	Mean(W)	99%CI
0.1	5.2×10^7	0.048	1.8×10^{-5}	1.5×10^7	0.048	1.8×10^{-5}
0.2	3.8×10^7	0.184	7.1×10^{-5}	1.1×10^7	0.184	7.1×10^{-5}
0.3	2.7×10^7	0.439	18.5×10^{-5}	8.3×10^6	0.439	18.5×10^{-5}
0.4	1.9×10^7	0.875	41.7×10^{-5}	5.8×10^6	0.875	41.6×10^{-5}
0.5	1.3×10^7	1.604	89.3×10^{-5}	4.0×10^6	1.604	89.2×10^{-5}
0.6	8.9×10^6	2.852	19.4×10^{-4}	2.7×10^6	2.854	19.4×10^{-4}
0.7	5.6×10^6	5.162	45.6×10^{-4}	1.7×10^6	5.162	4.6×10^{-3}
0.8	3.2×10^6	10.139	13.1×10^{-3}	9.6×10^5	10.142	13.2×10^{-3}
0.9	1.4×10^7	25.905	65.6×10^{-3}	4.1×10^5	25.885	64.7×10^{-3}
0.92	1.0×10^6	33.977	10.7×10^{-2}	3.1×10^5	33.927	10.6×10^{-2}
0.95	6.2×10^5	58.202	28.9×10^{-2}	1.8×10^5	58.190	25.3×10^{-2}

Поскольку регенеративное моделирование позволяет вычислять доверительные интервалы, в следующих таблицах приведено (для различных значе-

ний ρ) не только среднее значение процесса времени ожидания W , $\text{Mean}(W)$, но также полуширина 99% доверительного интервала, 99%CI, и число циклов регенерации, RegW . Отметим, что число регенераций входного процесса RegIn не зависит от ρ , но пропорционально δ .

Выберем $i_0 = 2$, т. е. экспоненциальное расщепление применяется к первому потоку. Из неравенств (6.40) легко видеть, что, выбрав $\eta = 1/\sqrt{1.5}$, экспоненциальное расщепление возможно при $0 < \delta \leq 1/3$. Результаты моделирования приведены только для двух значений δ .

Как отмечено в Таблице 6.12, когда интенсивность трафика ρ увеличивается, RegW уменьшается, в то время как $\text{Mean}(W)$ и 99%CI увеличиваются. Значение δ не оказывает существенного влияния на параметры, за исключением RegW . В частности, для $\delta = 0.33$ $\text{RegIn} \approx 7.2 \times 10^7$, тогда как для $\delta = 0.1$ значение $\text{RegIn} \approx 2.16 \times 10^7$ примерно в ≈ 3.3 раза меньше. Аналогичное наблюдение верно для значений RegW .

Таблица 6.13. Число регенераций RegW , Среднее время ожидания $\text{Mean}(W)$ и полуширина доверительного интервала 99%CI при $i_0 = 1, \eta = 0.63, \beta_1 = 0.5, \beta_2 = 0.67, \alpha_1 = 1.5, \alpha_2 = 2$.

ρ	$\delta = 0.33$			$\delta = 0.1$		
	RegW	Mean(W)	99%CI	RegW	Mean(W)	99%CI
0.1	6.2×10^7	0.048	1.8×10^{-5}	1.85×10^7	0.048	1.8×10^{-5}
0.2	4.6×10^7	0.184	7.1×10^{-5}	1.38×10^7	0.184	7.1×10^{-5}
0.3	3.4×10^7	0.439	18.5×10^{-5}	1.03×10^7	0.439	18.5×10^{-5}
0.4	2.5×10^7	0.875	41.7×10^{-5}	7.63×10^6	0.875	41.7×10^{-5}
0.5	1.8×10^7	1.604	89.2×10^{-5}	5.50×10^6	1.604	89.3×10^{-5}
0.6	1.3×10^7	2.853	19.3×10^{-4}	3.818×10^6	2.855	19.4×10^{-4}
0.7	8.3×10^6	5.161	4.5×10^{-3}	2.49×10^6	5.162	45.6×10^{-4}
0.8	4.8×10^6	10.147	1.3×10^{-2}	1.45×10^6	10.148	13.1×10^{-3}
0.9	2.1×10^6	25.855	6.5×10^{-2}	6.33×10^5	25.904	65.5×10^{-2}
0.92	1.6×10^6	33.956	0.107	4.93×10^5	33.903	10.7×10^{-2}
0.95	9.8×10^5	58.278	0.291	2.96×10^5	58.275	29.4×10^{-2}

Аналогичные выводы можно привести, когда экспоненциальное расщепле-

ние применяется ко второму потоку (см. Таблицу 6.13). В этом случае, согласно неравенствам (6.40), параметр $\eta = 2^{-2/3}$, и экспоненциальное расщепление возможно применить при $0 < \delta \leq 1/3$. В этом случае входной процесс регенерирует чаще, но связь между числом регенераций и значением δ остается почти такой же (а именно, $\text{RegIn} \approx 9.3 \times 10^7$ и $\text{RegIn} \approx 2.8 \times 10^7$ для $\delta = 0.33$ и $\delta = 0.1$, соответственно, что примерно в ≈ 3.3 раза меньше).

6.7. Регенеративное оценивание

Предположим, что необходимо оценить математическое ожидание некоторой функции от исследуемого процесса (в дискретном времени) $\mathbf{M}f(Z) = r$. Например, Z_n – число заявок в системе в момент времени t_n , Z – стационарное число заявок, а $f(Z)$ – задержка заявки в системе или ее энергопотребление. Предположим, что $\{Z_n, n \geq 1\}$ регенерирует в классическом смысле с моментами регенерации $\{\beta_j, j \geq 0\}$ и н.о.р. длинами циклов регенерации $\{\hat{\Delta}_j, j \geq 0\}$.

Пусть $Y_j = \sum_{i=\beta_j}^{\beta_{j+1}-1} f(Z_i)$ – сумма значений характеристики процесса на j -м цикле регенерации. Для классически регенерирующего процесса $\{Y_j\}$ и $\{V_j = Y_j - r\hat{\Delta}_j\}$ образуют последовательности н.о.р. с.в. и $\mathbf{M}V_j = 0$. В силу (6.2), статистическое оценивание характеристики процесса сводится к следующему: при наличии н.о.р. наблюдений $(Y_j, \hat{\Delta}_j)$, $j = 1, \dots, n$, необходимо оценить величину

$$r = \frac{\mathbf{M}Y_1}{\mathbf{M}\hat{\Delta}_1}.$$

Пусть $\{y_i\}_{i=1}^n$, $\{v_i\}_{i=1}^n$, $\{a_i\}_{i=1}^n$ – реализации с.в. Y_1 , V_1 , $\hat{\Delta}_1$, соответственно. Обозначим $\bar{y}_n$, $\bar{a}_n$, $\bar{v}_n$ – средние выборочные для с.в. Y_i , $\hat{\Delta}_i$, V_j соответственно, т.е. $\bar{y}_n = \frac{1}{n} \sum_{j=1}^n y_j$, $\bar{a}_n = \frac{1}{n} \sum_{j=1}^n a_j$, $\bar{v}_n = \frac{1}{n} \sum_{j=1}^n v_j$.

Очевидно, что $\bar{v}_n = \bar{y}_n - \hat{r}\bar{a}_n$, $\hat{r} = \bar{y}_n/\bar{a}_n$.

Пусть $\sigma^2 = \mathbf{D}V_1 = \mathbf{D}(Y_1 - r\hat{\Delta}_1)$. Из ЦПТ следует, что

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\frac{\sqrt{n} \cdot \bar{v}_n}{\sigma} \leq x \right) = \Phi^*(x)$$

или

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\frac{\sqrt{n}(\hat{r} - r)}{\sigma/\bar{a}_n} \leq x \right) = \Phi^*(x),$$

где $\Phi^*(x)$ – ф.р. стандартного нормального закона.

Заметим, что дисперсия σ^2 может быть записана как $\sigma^2 = \mathbf{D}Y_1 - 2rcov(Y_1, \hat{\Delta}_1) + r^2\mathbf{D}\hat{\Delta}_1$.

Пусть s_{11}^2, s_{22}^2 – выборочные дисперсии с.в. Y_j и α_j соответственно; s_{12}^2 – выборочный второй смешанный момент с. в. Y_j и $\hat{\Delta}_j$, т. е.

$$\begin{aligned} s_{11}^2 &= \frac{1}{n-1} \sum_{j=1}^n (y_j - \bar{y}_n)^2 = \frac{1}{n-1} \sum_{j=1}^n y_j^2 - \frac{1}{n(n-1)} \left(\sum_{j=1}^n y_j \right)^2; \\ s_{22}^2 &= \frac{1}{n-1} \sum_{j=1}^n (a_j - \bar{a}_n)^2 = \frac{1}{n-1} \sum_{j=1}^n a_j^2 - \frac{1}{n(n-1)} \left(\sum_{j=1}^n a_j \right)^2; \\ s_{12}^2 &= \frac{1}{n-1} \sum_{j=1}^n (y_j - \bar{y}_n)(a_j - \bar{a}_n) = \frac{1}{n-1} \sum_{j=1}^n y_j a_j - \frac{1}{n(n-1)} \sum_{j=1}^n y_j \sum_{j=1}^n a_j. \end{aligned}$$

Обозначим $s^2 = s_{11}^2 - 2\hat{r}s_{12}^2 + \hat{r}^2 s_{22}^2$. Из усиленного закона больших чисел следует, что $s^2 \rightarrow \sigma^2$ при $n \rightarrow \infty$ с в. 1. Тогда

$$\lim_{n \rightarrow \infty} \mathbf{P} \left(\frac{\sqrt{n}(\hat{r} - r)}{s/\bar{a}_n} \leq x \right) = \Phi^*(x).$$

Пусть

$$z_\gamma = \Phi^{-1}((1 - \gamma)/2), \text{ где } \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-x^2/2} dx.$$

Таким образом, асимптотический $100(1 - \gamma)\%$ доверительный интервал для оценки неизвестного $r = \mathbf{M}(f(Z))$ имеет следующий вид:

$$\left[\hat{r} - \frac{z_\gamma s}{\sqrt{n\bar{a}_n}}; \hat{r} + \frac{z_\gamma s}{\sqrt{n\bar{a}_n}} \right], \quad (6.41)$$

где n – число циклов регенерации, полученных в процессе моделирования.

Заметим, что $\sigma^2 < \infty$, если время обслуживания и интервалы между приходами с.в. имеют конечные вторые моменты [172]. Конечность $\text{cov}(Y_1, \hat{\Delta}_1)$ следует из конечности $\mathbf{M}(Y_1^2)$, $\mathbf{M}(\hat{\Delta}_1^2)$ и неравенства Гельдера.

Из конструкции доверительного интервала, построенного по моментам регенерации (6.41), следует, что именно от числа моментов регенерации N , а не числа наблюдений зависит ширина доверительного интервала, а, следовательно, и точность оценивания. Поэтому частота моментов регенерации играет критическую роль при регенеративном оценивании в условиях ограниченного вре-

мени моделирования. Известно, что качество регенеративной оценки связано с оценкой дисперсии σ^2 , а именно, согласно ЦПТ для σ (см. [84]):

$$n^{1/2}(s - \sigma) \Rightarrow N(0, \frac{\eta^2}{4\sigma^2}), \quad \text{при } n \rightarrow \infty,$$

где

$$\eta^2 = \frac{M(U_1 - cV_1)^2}{M\hat{\Delta}_1},$$

и

$$U_i = V_i^2 - \sigma^2 \hat{\Delta}_i, i = 1, \dots, n, \quad \text{и } c = 2M(V_1 \hat{\Delta}_1)/M\hat{\Delta}_1.$$

При этом асимптотическая дисперсия η^2 зависит от выбора последовательности моментов регенерации. В Теореме 1 [29] доказано, что чем больше число регенераций, тем меньше асимптотическая дисперсия η^2 .

6.8. Регенеративная оценка экстремального индекса

Рассмотрим пару (Z, β) , образованную регенерирующей случайной последовательностью $\{Z_n, n \geq 1\}$ и последовательностью моментов регенерации $\beta = (\beta_0, \beta_1, \dots)$, $\beta_0 < \beta_1 < \dots$.

Обозначим через M_k максимум последовательности на k -м цикле регенерации:

$$M_k = \sup_i \{Z_i, \beta_k \leq i < \beta_{k+1}\}.$$

Пусть $\hat{\Delta}_k = \beta_{k+1} - \beta_k$, $k \geq 0$, и $M\hat{\Delta}$ — математическое ожидание длины цикла регенерации. Стационарная регенерирующая последовательность $\{Z_n, n \geq 0\}$ имеет экстремальный индекс θ тогда и только тогда, когда существует предел [146]

$$\theta_\Delta := \lim_{n \rightarrow \infty} \frac{P(M_1 > u_n)/M\hat{\Delta}}{P(Z_1 > u_n)}, \quad (6.42)$$

для нормализующей последовательности $\{u_n\}$, удовлетворяющей соотношению $n\bar{F}(u_n) \rightarrow v$ при $n \rightarrow \infty$. Из соотношения (6.42) вытекает, что оценка экстремального индекса на основе регенеративного метода (для большого n) имеет

вид

$$\tilde{\theta}_{\Delta}(n) = \frac{1}{\tilde{\Delta}} \frac{\ln \left(1 - \frac{m_{\Delta}(u_n)}{k_{\Delta}} \right)}{\ln \left(1 - \frac{N(u_n)}{n} \right)}, \quad (6.43)$$

где $m_{\Delta}(u_n) = \#(i \leq k_{\Delta} : M_i > u_n)$ — число циклов регенерации, на которых превышено пороговое значение u_n , $N(u_n) = \#(i \leq n : Z_i > u_n)$ — число превышений случайной последовательностью $\{Z_n, n \geq 0\}$ порогового значения u_n , k_{Δ} — число циклов регенерации и $\tilde{\Delta}$ — оценка $\mathbf{M}\hat{\Delta}$.

Заметим, что условие перемешивания $D(u_n)$ (необходимое для существования экстремального индекса) для стационарных регенерирующих процессов с конечной средней длиной цикла выполняется для любых $\{u_n\}$ [146]. Если множества значений $\hat{Z} = (Z_{i_1}, \dots, Z_{i_k})$ и $\tilde{Z} = (Z_{j_1}, \dots, Z_{j_m})$ принадлежат разным циклам регенерации, $\hat{Z} \in C_i$, $\tilde{Z} \in C_j, i \neq j$, то в силу независимости C_i, C_j разность в левой части соотношения (4.34) равна нулю. В противном случае, если часть значений $\hat{Z}$ или $\tilde{Z}$ попадают в один и тот же цикл регенерации, то условие положительной возвратности, $\mathbf{M}\hat{\Delta} < \infty$, и $l \geq \mathbf{M}\hat{\Delta}$ гарантируют выполнение соотношения (4.34).

6.9. Описание комплекса программ

Комплекс программ для ЭВМ написан на языке программирования Python и предназначен для регенеративного моделирования систем обслуживания и расчета регенеративных характеристик качества обслуживания.

Каждая программа для ЭВМ включает в себя следующие блоки:

Блок 1. Генерация входных данных : время обслуживания, входной поток, число серверов, число входных потоков, число мест обслуживания.

Блок 2. Построение моментов регенерации : методом классической регенерации, методом экспоненциального расщепления или методом расщепления переходного ядра.

Алгоритм построения моментов искусственной регенерации методом экспоненциального расщепления представлен в разделе 6.6.

Алгоритм построения моментов искусственной регенерации методом расщепления переходного ядра представлен в разделе 6.4.4.

Блок 3. Расчет регенеративной оценки характеристик качества обслуживания (экстремального индекса, среднего времени ожидания, средней длины очереди).

Формула расчета доверительного интервала характеристики на основе регенеративного метода представлена в разделе 6.7.

Формула расчета оценки экстремального индекса на основе регенеративного метода представлена в разделе 6.8.

Расчет экстремального индекса производится по формулам (4.42) и (4.45), если существует положительное решение уравнения (4.38) и его можно найти в явном виде или численно (см. раздел 4.6.)

Блок 4. Вывод результатов (графики, расчетные значения характеристик).

6.10. Выводы по результатам главы 6

Разработан новый метод регенеративных огибающих для многосерверных систем $GI/G/N/\infty$ и систем обслуживания со смешанными распределениями (разделы 6.2. и 6.3.)

Получены соотношения для частоты (интенсивности) моментов искусственной регенерации в системах $Pareto/M/1$ и $H_2/M/1$ (раздел 6.4). Численные эксперименты подтверждают, что полученные теоретические величины согласуются с эмпирическими оценками, полученными регенеративным методом.

Получены достаточные условия экспоненциального расщепления распределения Парето (Лемма 6.2, раздел 6.4.)

Разработан метод искусственной регенерации на основе экспоненциального расщепления для построения регенераций в системах обслуживания с суперпозицией входных потоков с тяжелым хвостом, доказана теорема о том, что система регенерирует в классическом смысле (Теорема 6.2)

Для системы обслуживания со смешанным входным процессом, включающим распределения Парето, получены условия применимости метода экспоненциального расщепления. Разработан алгоритм построения моментов регенерации для процесса стационарного времени ожидания (раздел 6.5.)

Основные результаты главы опубликованы в работах автора [190, 191, 193, 204, 210, 211, 212, 213, 215].

Заключение

Диссертационная работа посвящена исследованию нового класса систем обслуживания со смешанными распределениями времени обслуживания и смешанным входным процессом. Такие системы адекватно описывают динамику многих современных телекоммуникационных систем и сетей, обрабатывающих большие наборы неоднородных данных. Свойства монотонности систем позволяют строить верхнюю и нижнюю границы исследуемых характеристик.

Впервые исследованы свойства монотонности по экстремальному индексу различных систем обслуживания, в том числе систем со смешанным обслуживанием. Теоретические выводы подтверждены рядом примеров и численных экспериментов, в которых для оценки экстремального индекса использовалось регенеративное моделирование. Для исследования предельного распределения максимума стационарного времени ожидания рассматриваемых систем построены нормализующие последовательности для каждой из исследованных систем.

В системах со смешанным входным процессом интервалы между приходами заявок зависимы, поэтому классическая регенерация для оценивания характеристик не применима. В этой связи разработан метод искусственной регенерации, основанный на экспоненциальном расщеплении, для моделирования систем обслуживания со смешанным входным процессом, содержащим распределения с тяжелым хвостом.

Основными научными достижениями данного диссертационного исследования являются:

1. Для систем со смешанными распределениями доказаны свойства внешней монотонности стационарной длины очереди и стационарного времени ожидания. Полученные результаты позволяют строить верхние и нижние границы основных стационарных характеристик систем. Кроме того, полученные теоретические результаты позволяют сравнивать соответствующие характеристики двух стационарных систем с различными смешанными распределениями.

2. Получена нижняя граница стационарной незавершенной работы в модели высокопроизводительного кластера со смешанным входным процессом, в котором каждая заявка занимает случайное число обслуживающих устройств, работающих параллельно.

3. Для системы с входным потоком, заданным суперпозицией показательного и Парето распределений построено аппроксимирующее распределение входного потока, которое является смесью распределений. Для системы с входным потоком, заданным суперпозицией входных потоков с распределением Вейбулла показано, что аппроксимирующее распределение имеет вид конечной смеси распределений Вейбулла и обобщенных Стейси-Гамма распределений. Результаты численных экспериментов подтверждают, что аппроксимация является достаточно точной, когда коэффициент загрузки системы не больше 0.6.

4. Доказаны теоремы о монотонности по экстремальному индексу стационарного времени ожидания в системах $GI/M/N$ и $M/G/1$. Найдено предельное распределение максимума стационарного времени ожидания для систем со смешанным обслуживанием с доминирующей компонентой, имеющей субэкспоненциальное распределение.

5. Разработан метод искусственной регенерации, основанный на экспоненциальном расщеплении плотности, для моделирования систем обслуживания со смешанным входным процессом, содержащим распределения с тяжелым хвостом. Этот метод позволяет увеличить частоту регенераций и тем самым сокращать время моделирования. Для системы обслуживания, в которой смешанный входной процесс включает потоки, заданные распределением Парето, получены достаточные условия применимости экспоненциального расщепления.

Полученные в диссертации теоретические результаты для моделей со смешанными распределениями существенным образом расширяют класс систем обслуживания, для которых исследованы свойства монотонности и их экстремальное поведение, и, тем самым, возможности анализа стационарных характеристик соответствующих систем с неоднородными данными.

Практическая значимость работы определяется применимостью полученных свойств монотонности и регенеративного метода оценивания для получения надежных оценок стационарных характеристик систем обслуживания со смешанными распределениями. Перспективы развития тематики работы связаны с обобщением полученных теоретических результатов, а также методов регенеративного моделирования на сети обслуживания.

Список сокращений и условных обозначений

с.в.	случайная величина
ф.р.	функция распределения
х.ф.	характеристическая функция
н.о.р.	независимые, одинаково распределенные
$F_X(x) = P(X \leq x)$	ф.р. с.в. X
$\bar{F}_X(x)$	хвост ф.р. $F_X(x)$ с.в. X
$F_e(x)$	функция равновесного распределения с.в. X с ф.р. F
$f_X(x)$	плотность распределения с.в. X
MX	математическое ожидание с.в. X
DX	дисперсия с.в. X
$r_X(x)$	функция интенсивности отказов с.в. X
$\Delta(F_X, F_Y)$	равномерная метрика между ф.р. F_X и F_Y
$D_{KL}(X Y)$	дивергенция Кульбака-Лейблера с.в. X и Y
X_u	величина превышения уровня u
$X \stackrel{st}{=} Y$	эквивалентность с.в. X и Y по распределению
$X \underset{r}{\leq} Y$	с.в. X меньше по интенсивности отказов, чем с.в. Y
$X \underset{st}{\leq} Y$	с.в. X стохастически меньше, чем с.в. Y
$X \leq Y$ с в. 1	с.в. X с вероятностью 1 меньше, чем с.в. Y
$X_n \Rightarrow X$ при $n \rightarrow \infty$	последовательность с.в. $\{X_n, n \geq 1\}$ сходится по распределению к с.в. X

$X_n \xrightarrow{P} X$ при $n \rightarrow \infty$	последовательность с.в. $\{X_n, n \geq 1\}$ сходится по вероятности к с.в. X
$X \sim \text{Exp}(\lambda)$	с.в. X с показательным распределением с параметром λ
$X \sim \text{Pareto}(\alpha, x_0)$	с.в. X с распределением Парето с параметрами α, x_0
$X \sim \text{We}(\alpha, \beta)$	с.в. X имеет распределение Вейбулла с параметрами α, β
$X \sim \text{Burr}(x_0, \alpha, k)$	с.в. X имеет распределение Бурра XII с параметрами x_0, α, k
$X \sim \text{Exp} - \text{Pareto}(\lambda, \alpha, x_0, p)$	с.в. X имеет показательное-Парето распределение с параметрами λ, α, x_0, p
$X \sim H_m$	с.в. X имеет m -мерное гипергеометрическое распределение
$\Gamma(t, z) = \int_z^\infty y^{t-1} e^{-y} dy$	неполная верхняя Гамма-функция
$\gamma(t, z) = \int_0^z y^{t-1} e^{-y} dy$	неполная нижняя Гамма-функция
$\Gamma(z) = \int_0^\infty y^{t-1} e^{-y} dy$	Гамма-функция
$F \in MDA(G)$	ф.р. F принадлежит области притяжения максимума G
ВФИ-распределение	распределение с возрастающей функцией интенсивности отказов
УФИ-распределение	распределение с убывающей функцией интенсивности отказов
ВВФИ-распределение	распределение с выпуклой вверх функцией интенсивности отказов
ВНФИ-распределение	распределение с выпуклой вниз функцией интенсивности отказов
FIFS	дисциплина первым пришел – первым обслужен

ПЛС	преобразование Лапласа-Стилтьеса
$\mathcal{R}(\delta)$	класс регулярно меняющихся функций
$\mathcal{S}$	класс субэкспоненциальных функций
$\mathcal{IR}(\delta)$	класс частично регулярно меняющихся функций
$\mathcal{HT}$	класс функций распределения с тяжелым хвостом
x_R	правая концевая точка распределения
x_L	левая концевая точка распределения
z^+	$\max(0, z)$
I_A	индикатор случайного события A
$\mathbb{R}$	$(-\infty, \infty)$
$:=$	равно по определению
$[a]$	целая часть, не превосходящая a
термины «ф.р.» и «распределение»	используются в работе как синонимы

Список использованной литературы

1. Байхельт Ф. Надежность и техническое обслуживание. Математический подход : пер. с нем. / Ф. Байхельт. П. Франкен. – М. : Радио и связь, 1988. – 392 с.
2. Барлоу Р. Математическая теория надежности / Р. Барлоу, Ф. Прошан. – М. : Сов. радио, 1969. – 488 с.
3. Барлоу Р. Статистическая теория надежности и испытания на безотказность / Р. Барлоу, Ф. Прошан. – М. : Наука, 1984. – 328 с.
4. Борисов А.В. Вероятностный анализ класса марковских скачкообразных процессов / А. В. Борисов, Ю. Н. Куринов, Р. Л. Смелянский // Информатика и ее применения. 2024. Т. 18, вып. 3. С. 30–37.
5. Боровков А. А. Теория вероятностей / А. А. Боровков. – М. : Наука, 1972. – 287 с.
6. Боровков А. А. Вероятностные процессы в теории массового обслуживания / А. А. Боровков. – М. : Наука, 1972.
7. Бочаров П. П. Теория массового обслуживания : учебник / П. П. Бочаров, А. В. Печинкин. – М. : Изд-во РУДН, 1995. – 529 с.
8. Булинский А. В. Теория случайных процессов / А. В. Булинский, А. Н. Ширяев. – М. : ФИЗМАТЛИТ, 2005. – 408 с.
9. Гумбель Э. Статистика экстремальных значений / Э. Гумбель. – М. : Мир, 1965. – 453 с.
10. Ермаков С. М. Математический эксперимент с моделями сложных систем / С. М. Ермаков, Д. С. Мелас. – СПб. : Санкт-Петербургский государственный университет, 1993.
11. Зорин А.В. К определению периода занятости при нелокальном описании потоков / А.В. Зорин // Информатика и ее применения. 2024. Т. 18, вып. 3. С. 45–51.

12. Иглехарт Д. Л. Регенеративное моделирование сетей массового обслуживания / Д. Л. Иглехарт, Д. С. Шедлер. – М. : Радио и связь, 1984.
13. Клейнрок Л. Теория массового обслуживания / Л. Клейнрок. – М. : Машиностроение, 1979. – 432 с.
14. Крэйн М. Введение в регенеративный метод анализа моделей / М. Крэйн, О. Лемуан. – Москва : Наука, 1982.
15. Королев В. Ю. Вероятностные модели : учебное пособие / В. Ю. Королев, О. В. Шестаков ; Московский государственный университет имени М. В. Ломоносова, Факультет вычислительной математики и кибернетики. – М. : МАКС Пресс, 2020. – 266 с.
16. Лебедев А.В. Основы стохастической теории экстремумов / А. В. Лебедев. – М. : Ленанд, 2018. – 104 с.
17. Лукач Е. Характеристические функции / Е. Лукач. – М. : Наука, 1979. – 424 с.
18. Матвеев В. Ф. Системы массового обслуживания / В. Ф. Матвеев, В. Г. Ушаков. – М. : Изд-во МГУ, 1984. – 240 с.
19. Рыков В. В. Основы теории массового обслуживания (Основной курс: марковские модели, методы марковизации) : учебное пособие / В. В. Рыков, Д. В. Козырев. – М. : ИНФРА-М, 2024. – 223 с. – (Высшее образование).
20. Саати Т. Элементы теории массового обслуживания и ее приложения / Т. Саати. – М. : Сов. радио, 1971.
21. Тихонов В. И. Выбросы траекторий случайных процессов / В. И. Тихонов, В. И. Хименко. – М. : Наука, 1987. – 306 с.
22. Федоткин М.А. Стохастические модели процессов адаптивного управления конфликтными потоками неоднородных требований / Федоткин М.А., Зорин А.В. // Теория вероятностей и ее применения. 2020. Т. 65. Вып. 1. С. 163–164.
23. Феллер В. Введение в теорию вероятностей и ее приложения / В. Феллер. Т. 1. – Москва : Мир, 1984. – 528 с.

24. Штойан Д. Качественные свойства и оценки стохастических моделей / Д. Штойан. – М. : Мир, 1979. – 269 с.
25. Abramowitz M. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables / M. Abramowitz, I. A. Stegun. – Washington, D. C. : U. S. Government Printing Office, 1972. – 1046 p.
26. Ancona-Navarrete M. A. A Comparison of Methods for Estimating the Extremal Index / M. A. Ancona-Navarrete, A. Tawn Jonathan // Extremes. – 2000. – Vol. 3, no. 1. – P. 5–38.
27. Albin S. L. On Poisson Approximations for Superposition Arrival Processes in Queues / S. L. Albin // Management Science. – 1982. – Vol. 28. – P. 126–137.
28. Andronov A. Artificial regeneration points for stochastic simulation of complex systems / A. Andronov // Proceedings of the 10th European Simulation Symposium (ESS'98). Simulation Technology: Science and Art. – 1998. – P. 34–40.
29. Andradottir S. Accelerated regeneration for markov chain simulation / S. Andradottir, J. M. Calvin, P. W. Glynn // Probability in the Engineering and Informational Sciences. – 1995. – Vol. 9. – P. 497–523.
30. Al-Hussaini E. K. Reliability and hazard based on finite mixture models / E. K. Al-Hussaini, K. S. Sultan // Advances in Reliability, Handbook of Statistics. – 2001. – Vol. 20. – P. 139–183.
31. Amini-Seresht E., Zhang Y. Stochastic comparisons on two finite mixture models / E. Amini-Seresht, Y. Zhang // Operations Research Letters. – 2017. – Vol. 45. – P. 475–480.
32. Artalejo J. R. Accessible bibliography on retrial queues / J. R. Artalejo // Mathematical and computer modelling. – 1999. – Vol. 30 (3-4). – P. 1–6.
33. Artalejo J. Retrial Queueing Systems: A Computational Approach / J. Artalejo J., A. Gomez-Corral. – Berlin ; Heidelberg : Springer-Verlag, 2008. – 318 p.

34. Asmussen S. Applied Probability and Queues. – 2 nd ed. / S. Asmussen. – New York : Springer-Verlag, 2003.
35. Asmussen S. Extreme Value Theory for Queues Via Cycle Maxima / S. Asmussen // Extremes. – 1998. – Vol. 1, iss. 2. – P. 137–168.
36. Asmussen S. Sampling at subexponential times with queueing applications / S. Asmussen, C. Kluppelberg, K. Sigman // Stochastic Processes and their Applications. – 1999. – Vol. 79, iss. 2. – P. 265–286.
37. Asmussen S. Tail probabilities for non-standard risk and queueing processes with subexponential jumps / S. Asmussen, H. Schmidli, V. Schmidt // Adv. Appl. Probab. – 1999. – Vol. 31. – P. 422–447.
38. Asmussen S. Stochastic Simulation: Algorithms and Analysis / S. Asmussen, P. W. Glynn. – Cham, Switzerland : Springer Nature, 2007.
39. Athreya K. B. A new approach to the limit theory of recurrent Markov Chains / K. B. Athreya, P. Ney // Transactions of the American Mathematical Society. – 1978. – Vol. 245. – P. 493–501.
40. Aven T. Stochastic Models in Reliability / T. Aven, U. Jensen. – New York : Springer-Verlag, 2013.
41. Avrachenkov K. Retrial networks with finite buffers and their application to Internet data traffic / K. Avrachenkov, U. Yechiali // Probability in the Engineering and Informational Sciences. – 2008. – Vol. 22. – P. 519–536.
42. Avrachenkov K. On tandem blocking queues with a common retrial queue / K. Avrachenkov, U. Yechiali // Computers and Operations Research. – 2010. – Vol. 37 (7). – P. 1174–1180.
43. Advances in Distribution Theory, Order Statistics, and Inference / Eds. N. Balakrishnan, J. M. Sarabia, E. Castillo. – Boston, MA : Birkhäuser, 2007. – 484 p.
44. Baccelli F. Elements of Queueing Theory. Palm Martingale Calculus and Stochastic Recurrences / F. Baccelli, P. Bremaud. – 2 nd ed. – Berlin ; Heidelberg, Germany : Springer, 2003.

45. Balkema A. A. Residual life time at great age / A. A. Balkema, L. De Haan // The Annals of Probability. – 1974. – Vol. 2, no. 5. – P. 792–804.
46. Bertail P. Regenerative block-bootstrap confidence intervals for the extremal index / P. Bertail, S. Cléménçon, J. Tressou // Electronic Journal of Statistics. – 2013. – Vol. 7. – P. 1224–1248.
47. Berger A. Comparisons of Multi-Server Queues with Finite Waiting Rooms / A. Berger, W. Whitt // Stochastic Models. – 1992. – Vol. 8, no. 4. – P. 719–732.
48. Berger A. Maximum values in queueing processes / A. Berger, W. Whitt // Probability in the Engineering and Informational Sciences. – 1995. – Vol. 9. – P. 375–409.
49. Bertsekas D. Data Networks / D. Bertsekas, R. Gallager. – Prentice-Hall International, 1992.
50. Borisov A. Partial Diffusion Markov Model of Heterogeneous TCP Link: Optimization with Incomplete Information / Borisov A., Bosov A., Miller G., and Sokolov I. // Mathematics. 2021. – Vol. 9, No. 14? art.no. 1632. PP. 1–31.
51. Beirlant J. Statistics of Extremes: Theory and Applications / J. Beirlant, Y. Goegebeur, J. Teugels, J. Segers. – Chichester, England : John Wiley & Sons, 2004.
52. Bertail P. Extreme values statistics for Markov chains via the (pseudo-) regenerative method / P. Bertail, S. Cléménçon, J. Tressou // Extremes. – 2009. – Vol. 12. – P. 327–360.
53. Billingsley P. Convergence of probability measures / P. Billingsley. – New York : Wiley, 1968.
54. Block H. W. Initial and final behaviour of failure rate functions for mixtures and systems / H. W. Block, Y. Li, T. H. Savits // Journal of Applied Probability. – 2003. – Vol. 40. – P. 721–740.
55. Borovkov A. A. Asymptotic Methods in Queueing Theory / A. A. Borovkov. – New York : Wiley, 1984.

56. Borowiak D. S. Financial and Actuarial Statistics: An Introduction. – 2 nd ed.
/ D. S. Borowiak, A. F. Shapiro. – Ohio, USA : CRC Press, 2013. – 392 p.
57. Boxma O. J. Waiting-time asymptotics for the M/G/2 queue with heterogeneous
servers / O. J. Boxma, Q. Deng, A. P. Zwart // Queueing Systems. – 2002. –
Vol. 40. – P. 5–31.
58. Bishop C. M. Neural Networks for Pattern Recognition / C. M. Bishop. –Oxford :
Oxford University Press, 1995. – 482 p.
59. Bradley R. C. Basic Properties of Strong Mixing Conditions. A Survey and Some
Open Questions // R.C. Bradley // Probability Surveys. – 2005. – Vol. 2(0). –
P. 107–144.
60. Charlot F. Irréductibilité et récurrence au sens de Harris des "Temps
d'attente"des files GI/G/q / F. Charlot, M. Ghidouche, M. Hamami //
Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete. – 1978. –
Vol. 43. – P. 187–203.
61. Choi B. D. Retrial queues with collision arising from unslotted CSMA/CD
protocol / B. D. Choi, Y. W. Shin, W. C. Ahn // Queueing Systems. – 1992. –
Vol. 11. – P. 335–356.
62. Choi B. D. The M/G/1 retrial queue with retrial rate control policy / B. D. Choi,
K. H. Rhee, K. K. Park // Probability in the Engineering and Informational
Sciences. – 1993. – Vol. 7. – P. 29–46.
63. Cohen J. W. The Single Server Queue – 2 nd impression / J. W. Cohen.
–Amsterdam : North-Holland, 1992.
64. Cohen J. W. Extreme value distribution for the M/G/1 and the G/M/1 queueing
system / J. W. Cohen // Ann. Inst. Henri Poincaré. – 1968. – Vol. IV, no. 1. –
P. 83–98.
65. Cohen J. W. The distribution of the maximum number of customers present
simultaneously during a busy period for the system M/G/1 and for the system
G/M/1 / J. W. Cohen // Journ. Appl. Prob. – 1967. – Vol. 4. – P. 162–179.

66. Colonius H. An invitation to coupling and copulas: With applications to multisensory modeling / H. Colonius // Journal of Mathematical Psychology. – 2015. – Vol. 74.
67. Cox D. R. Renewal Theory / D. R. Cox. – London ; UK : Methuen, 1962.
68. Embrechts P. Modelling Extremal Events for Insurance and Finance / P. Embrechts, C. Klüppelberg, T. Mikosch. – Heidelberg ; New York ; Dordrecht ; London : Springer, 1997.
69. Fayolle G. A Simple Telephone Exchange with Delayed Feedbacks / G. Fayolle // Proceedings of the International Seminar on Teletraffic Analysis and Computer Performance Evaluation. – 1986. – P. 245–253.
70. Feldman A. Fitting mixtures of exponentials to long-tail distributions to analyze network performance models / A. Feldman, W. Whitt // Performance Evaluation. – 1998. – Vol. 31, iss. 3-4. – P. 245–279.
71. Feitelson D. G. Workload Modeling for Computer Systems Performance Evaluation / D. G. Feitelson. – Cambridge University Press, 2014. – 597 p.
72. Ferro C. A. T. Inference for Clusters of Extreme Values / C. A. T. Ferro, J. Segers // Journal of the Royal Statistical Society. Series B: Statistical Methodology. – 2003. – Vol. 65, no. 2. – P. 545–556.
73. Finkelstein M. On Mixture Failure Rates Ordering / M. Finkelstein, V. Esaulova // Communications in Statistics – Theory and Methods. – 2006. – Vol. 35, no. 11. – P. 1943–1955.
74. Foss S. G. On the ergodicity conditions for stochastically recursive sequences / S. G. Foss // Queueing Systems. – 1992. – Vol. 12. – P. 287–296.
75. Foss S. Regeneration and renovation in queues / S. Foss, V. Kalashnikov // Queueing Systems. – 1991. – No. 8. – P. 211–224.
76. Foss S. An overview of some stochastic stability methods / S. Foss, T. Konstantopoulos // Journal of the Operations Research Society of Japan. – 2004. – Vol. 47, no. 4. – P. 275–303.

77. Foss S. On Large Delays in Multi-Server Queues with Heavy Tails / S. Foss, D. Korshunov // Mathematics of Operations Research. – 2012. – Vol. 37, no. 2. – P. 201–218.
78. Foss S., Korshunov D., Zachary S. An Introduction to Heavy-Tailed and Subexponential Distributions / S. Foss, D. Korshunov, S. Zachary. – New York : Springer New York, 2013.
79. Galambos J. The Asymptotic Theory of Extreme Order Statistics / J. Galambos. – New York : Wiley, 1978.
80. Glynn P. Wide-sense regeneration for Harris recurrent Markov processes: an open problem / P. Glynn // Queueing Systems. – 2011. – Vol. 68, no. 3-4. – P. 305–311.
81. Glynn P. W. A GSMP formalism for discrete event systems / P. W. Glynn // Proceedings of IEE. – 1989. – Vol. 77. – P. 14–23.
82. Glynn P. W. Some topics in regenerative steady-state simulation / P. W. Glynn // Acta Applicandae Mathematicae. – 1994. – Vol. 34. – P. 225–236.
83. Glynn P. W. Conditions for the applicability of the regenerative method / P. W. Glynn, D. L. Iglehart // Management Science. – 1993. – Vol. 39. – P. 1108–1111.
84. Glynn P.W, Iglehart D.L. A joint central limit theorem for the sample mean and regenerative variance estimator / P. W. Glynn, D. L. Iglehart // Annals of Operations Research. – 1987. – Vol. 8. – P. 41–55.
85. Glynn P. W. and L'Ecuyer P. Likelihood Ratio Gradient Estimation for Stochastic Recursions / P. W. Glynn, P. L'Ecuyer // Advances in Applied Probability. – 1995. – Vol. 27, No. 4. – P. 1019–1053.
86. Gnedenko B. On the limiting distribution of the maximum term in a random series / B. Gnedenko // Breakthroughs in Statistics. – Vol. 1: Foundations and Basic Theory / Eds. S. Kotz, N. L. Johnson. – New York : NY, USA, Springer Science + Business Media, 1993. – P. 185–225.

87. Goldie C. Subexponential distributions / C. Goldie, C. Klüppelberg // A Practical Guide to Heavy Tails: Statistical Techniques for Analyzing Heavy Tails. – Boston, MA : Birkhäuser, 1998. – P. 435–459.
88. Guess F. Mean Residual Life: Theory and Applications / F. Guess, F. Proschan // Handbook of Statistics. – Vol. 7: Quality Control and Reliability / Eds. P. R. Krishnaiah, C. R. Rao. – Amsterdam ; The Netherlands, North-Holland, 1988. – P. 215–224.
89. Gupta R. C. Role of equilibrium distribution in reliability studies / R. C. Gupta // Probability in the Engineering and Informational Sciences. – 2007. – Vol. 21. – P. 315–334.
90. Gut A. Stopped Random Walks: Limit Theorems and Applications. – 2 nd ed. / A. Gut. – Berlin ; Heidelberg, Germany : Springer, 2009.
91. de Haan L. Extreme Value Theory: An Introduction / L. de Haan, A. Ferreira. – New York : Springer, 2006.
92. Haji R. A Relation between Stationary Queue and Waiting Time Distributions / R.Haji, G. F. Newell // Journal of Applied Probability. – 1971. – Vol. 8. – P. 617–620.
93. Hall P. On measures of the distance of a mixture from its parent distribution / P. Hall // Stochastic Processes and their Applications. – Amsterdam : North-Holland Publishing Company, 1979. – Vol. 8, iss. 3. – P. 357–365.
94. Hooghiemstra G. Computing the extremal index of special Markov chains and queues / G. Hooghiemstra, L. E. Meester // Stochastic Processes and their Applications. – 1996. – Vol. 65 (2). – P. 171–185.
95. Iglehart D. L. Extreme values in GI/G/1 queue / D. L. Iglehart // Annals of Mathematical Statistics. – 1972. – Vol. 43 (2). – P. 627–635.
96. Kalashnikov V. Topics on Regenerative Processes / V. Kalashnikov. – Boca Raton, USA : CRC Press, 1994.

97. Kalashnikov V. V. Regenerative queueing processes and their qualitative and quantitative analysis / V. V. Kalashnikov // Queueing Systems. – 1990. – Vol. 6. – P. 113–136.
98. Khintchine A. Y. Mathematical Methods in Queueing Theory / A. Y. Khintchine. – London, UK : Griffin, 1960.
99. Kiefer J. On the theory of queues with many servers / J. Kiefer, J. Wolfowitz // Transactions of the American Mathematical Society. – 1955. – P. 1–18.
100. Klüppelberg C. Subexponential Distributions and Integrated Tails / C. Klüppelberg // Journal of Applied Probability. – 1988. – Vol. 25 (1). – P. 132–141.
101. Kochar S. C. Stochastic Comparisons with Applications in Order Statistics and Spacings / S. C. Kochar. – Cham : Springer, 2022. – 280 p.
102. Batrakova D. A new method for the probabilistic and statistical analysis of information flows in telecommunication networks / D. Batrakova, V. Korolev // Information Applications. – 2007. – Vol. 1. – P. 40–53.
103. Kang S. Maxima of sojourn times in acyclic Jackson queueing networks / Kang S., Serfozo R. // Computerx Ops Rex. – 1997. – Vol. 24, No. II. – P. 1085–1095.
104. Korolev V. Yu. Stability of finite mixtures of generalized Gamma-distributions with respect to disturbance of parameters / V. Yu. Korolev, V. A. Krylov, V. Yu. Kuz'min // Information Applications. – 2011. – Vol. 5 (1). – P. 31–38.
105. Lam T. Superposition of Renewal Processes / T. Lam, J. P. Lehoczky // Advances in Applied Probability. – 1991. – Vol. 23. – P. 64–85.
106. Lai C.-D. Stochastic Ageing and Dependence for Reliability / C.-D. Lai, M. Xie. – New York : Springer, 2006. – 437 p.
107. Leadbetter M. Extremes and Related Properties of Random Sequences and Processes / M. Leadbetter, G. Lindgren, H. Rootzen. – New York : Springer-Verlag, 1983. – 336 p.
108. Lin X. S. Integrated Tail Distribution. Wiley StatsRef: Statistics Reference Online / X. S. Lin. – 2014.

109. Lindvall T. Lectures on the Coupling Method / T. Lindvall. – 1992.
110. Lawrence A. J. Dependency of Intervals Between Events in Superposition Processes / A. J. Lawrence // Journal of the Royal Statistical Society. Series B: Methodological. – 1973. – Vol. 35. – P. 306–315.
111. Markovich N. Clustering and Hitting Times of Threshold Exceedances and Applications / N. Markovich // International Journal of Data Analysis Techniques and Strategies. – 2017. – Vol. 9, iss. 4. – P. 331–347.
112. Markovich N. Cluster Modeling of Lindley Process with Application to Queuing / N. Markovich, R. Razumchik // Distributed Computer and Communication Networks, Lecture Notes in Computer Science. – Cham, Switzerland : Springer, 2019. – Vol. 11965. – P. 330–341.
113. Markovich N. M. Modeling clusters of extreme values / N. M. Markovich // Extremes. – 2014. – Vol. 17 (1). – P. 97–125.
114. Markovich N. M. Erratum to: modeling clusters of extreme values / N. M. Markovich // Extremes. – 2016. – Vol. 19 (1). – P. 139–142.
115. Markovich N. M. Clusters of extremes: modeling and examples / N. M. Markovich // Extremes. – 2017. – Vol. 20. – P. 519–538.
116. Marshall A. Distributions: Structure of Nonparametric, Semiparametric, and Parametric Families / A. Marshall, I. Olkin. – New York : Springer, 2007.
117. Mclachlan G. I. Finite Mixture Models / G. I. Mclachlan, D. Peel. – USA : John Wiley & Sons, 2004.
118. Medhi J. Stochastic models in queueing theory / J. Medhi. – Boston : Academic Press, 1991.
119. Mikosch T. Gumbel and Fréchet convergence of the maxima of independent random walks / T. Mikosch, J. Yslas // Advances in Applied Probability. – 2020. – Vol. 52. – P. 213–236.
120. Moka S.B. Regenerative simulation for queueing networks with exponential or heavier tail arrival distributions / S.B. Moka, S. Juneja // ACM Trans. Model. Comput. Simul. (TOMACS). – 2015. – Vol. 25. – P. 1–22.

121. Morozov E. An extended regenerative structure and queueing network simulation. Preprint No. 1995-08, ISSN 0347-2809. Dept. Math. / E. Morozov. – Gothenburg, Sweden, Chalmers University, 1995.
122. Morozov E. The tightness in the ergodic analysis of regenerative queueing processes / E. Morozov // Queueing Systems. – 1997. – Vol. 27. – P. 179–203.
123. Morozov E. Stochastic boundness of some queueing systems. Preprint No. R-95-2022, ISSN 0908-1216, Dept. Math. and Computer Science / E. Morozov. – Aalborg, Denmark : Aalborg University, 1995.
124. Morozov E. V. Stability analysis of regenerative queueing systems (In Russ.) / E. V. Morozov, R. Delgado // Avtomatika i Telemekhanika. – 2009. – Vol. 70. – P. 42–58.
125. Morozov E. V. Performance analysis and stability of multiclass orbit queue with constant retrial rates and balking / E. V. Morozov, A. S. Rumyantsev, S. Dey, T. G. Deepak // Performance Evaluation. – 2019. – Vol. 134. – P. 1–17.
126. Morozov E. Stability Analysis of a MAP/M/s Cluster Model by Matrix-Analytic Method / E. Morozov, A. Rumyantsev // Lecture Notes in Computer Science: Computer Performance Engineering, 13th European Workshop. –2016. – Vol. 9951. – P. 63–76.
127. Morozov E. Stability Analysis of Regenerative Queueing Models / E. Morozov, B. Steyaert. – Cham, Switzerland : Springer, 2021.
128. Morozov E. On the Positive Recurrence of Finite Regenerative Stochastic Models / E. Morozov, V. Rykov // Mathematics. – 2023. – Vol. 11 (23). – P. 47–54.
129. Naess A. Applied Extreme Value Statistics: With a Special Focus on the ACER Method / A. Naess. – New York : Springer, 2024. – 277 p.
130. Navarro J. Stochastic comparisons of generalized mixtures and coherent systems / J. Navarro // TEST. – 2016. – Vol. 25. – P. 150–169.
131. Navarro J. Hazard rate ordering of order statistics and systems / J. Navarro, M. Shaked // Journal of Applied Probability. – 2006. – Vol. 43. – P. 391–408.

132. Navarro J. Stochastic comparisons of distorted distributions, coherent systems, and mixtures with ordered components / J. Navarro, Y. Águila // *Metrika*. – 2017. – Vol. 80. – P. 627–648.
133. Nadarajah S. On the convolution of Pareto and gamma distributions / S. Nadarajah, S. Kotz // *Computer Networks*. – 2007. – Vol. 51, no. 12. – P. 3650–3654.
134. Nadarajah S. On the Laplace transform of the Pareto distribution / S. Nadarajah // *Queueing Systems*. – 2006. – Vol. 54, no. 4. – P. 243–244.
135. Nazarov A. A survey of recent results in finite-source retrial queues with collisions / A. Nazarov, J. Sztrik, A. Kvach // *Communications in Computer and Information Science*. – 2018. – Vol. 912. – P. 1–15.
136. Nazarov A. Mathematical Model of Call Center in the Form of Multi-Server Queueing System / A. Nazarov, A. Moiseev, S. Moiseeva // *Mathematics*. – 2021. – Vol. 9, iss. 22, art. no. 2877. – 13 p.
137. Newell G. F. Approximations for Superposition Arrival Processes in Queues / G. F. Newell // *Management Science*. – 1984. – Vol. 30. – P. 623–632.
138. Nummelin E. A Splitting Technique for Harris Recurrent Markov Chains / E. Nummelin // *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*. – 1978. – Vol. 43. – P. 309–318.
139. O'Brien G. L. Extreme values for stationary Markov sequences / G. L. O'Brien // *Ann. prob.* – 1987. – Vol. 15. – P. 281–291.
140. Pakes A. G. On the Tails of Waiting-Time Distributions / A. G. Pakes // *Journal of Applied Probability*. – 1975. – Vol. 12, no. 3. – P. 555–564.
141. Ramsay C. M. Exact waiting time and queue size distributions for equilibrium M/G/1 queues with Pareto service / C. M. Ramsay // *Queueing Systems*. – 2007. – Vol. 57, no. 4. – P. 147–155.
142. Rego V. Some explicit formulas for mixed exponential service systems / V. Rego // *Computers and Operations Research*. – 1988. – Vol. 15, iss. 6. – P. 509–520.

143. Resnick S. I. Extreme Values, Regular Variation and Point Processes / S. I. Resnick. – New York : Springer, 1987. – 320 p.
144. Rodionov I. Threshold selection for extremal index estimation / I. Rodionov, N. Markovich // Journal of Nonparametric Statistics. – 2023. – Vol. 36(3). P. 527–546.
145. Ross S. On increasing-failure-rate random variables / S. Ross, J. Shanthikumar, Z. Zhu // Journal of Applied Probability. – 2005. – Vol. 42. – P. 797–809.
146. Rootzen H. Maxima and exceedances of stationary Markov chains / H. Rootzen // Advances in Applied Probability. – 1988. – Vol. 20 (2). – P. 371–390.
147. Rumyantsev A. Stability criterion of a multiserver model with simultaneous service / A. Rumyantsev, E. Morozov // Annals of Operations Research. – 2017 (First Online: 2015). – Vol. 252, no. 1. – P. 29–39.
148. Rykov V. Application of Decomposable Semi-Regenerative Processes to the Study of k-out-of-n Systems / V. Rykov, N. Ivanova, D. Kozyrev // Mathematics. – 1933. – Vol. 9 (16). – P. 2021.
149. Rykov V. Reliability of engineering systems and technological Risk / V. Rykov. – USA : John Wiley & Sons, 2016. – 228 p.
150. Scheller-Wolf A. Delay moments for FIFO $GI/GI/s$ queues / A. Scheller-Wolf, K. Sigman // Queueing Systems. – 1997. – Vol. 25, no. 14. – P. 77–95.
151. Schiff J. L. The Laplace Transform / J. L. Schiff. – New York : Springer, 1999.
152. Senreta E. Regularly varying functions / E. Senreta. – Berlin ; Heidelberg ; New York : Springer-Verlag, 1976.
153. Serfozo R. F. Basics of Applied Stochastic Processes / R. F. Serfozo. – Berlin ; Heidelberg, Germany : Springer, 2009.
154. Shaked M. Mixtures and monotonicity of failure rate functions / M. Shaked, F. Spizzichino // Handbook of Statistics. – 2001. – Vol. 20. – P. 185–198.

155. Shaked M. Bounds on the Distance of a Mixture from Its Parent Distribution / M. Shaked // Journal of Applied Probability. – 1981. – Vol. 18, iss. 4. – P. 853–863.
156. Shaked S. Stochastic Orders / S. Shaked, J. G. Shanthikumar. – New York : Springer Science + Business Media, LLC, 2007. – 486 p.
157. Shedler G. S. Regeneration and Networks of Queues / G. S. Shedler. – New York : Springer-Verlag, 1987.
158. Shedler G. S. Regenerative Stochastic Simulation / G. S. Shedler. – Cambridge : Academic Press, MA, USA, 1993.
159. Sigman K. Appendix: A primer on heavy-tailed distributions / K. Sigman // Queueing Systems. – 1999. – Vol. 33. – P. 261–275.
160. Sigman K. Exact simulation of the stationary distribution of the FIFO M/G/c queue: the general case for $\rho < c$ / K. Sigman // Queueing Systems. – 2012. – Vol. 70. – P. 37–43.
161. Sigman K. One-dependent regenerative processes and queues in continuous time / K. Sigman // Math. Oper. Res. – 1990. – Vol. 15. – P. 175–189.
162. Sigman K. Queues as Harris recurrent Markov chains / K. Sigman // Queueing Systems. – 1988. – Vol. 3. – P. 179–198.
163. Sigman K. A review of regenerative processes. / K. Sigman, R. W. Wolff // SIAM Review. – 1993. – Vol. 35. – P. 269–288.
164. Smith R. L. The extremal index for a Markov chain / R. L. Smith // Journal of Applied Probability. – 1992. – Vol. 29 (1). – P. 37–45.
165. Smith W. L. Regenerative stochastic processes / W. L. Smith // Proceedings of the Royal Society. Series A. – 1955. – Vol. 232. – P. 6–31.
166. Smith L. Estimating the extremal index / L. Smith, I. Weissman // Journal of the Royal Statistical Society. Series B: Methodological. – 1994. – Vol. 56 (3). – P. 515–528.

167. Sonderman D. Comparing multi-server queues with finite waiting rooms, I: Same number of servers / D. Sonderman // Advances in Applied Probability. – 1979. – Vol. 11. – P. 439–447.
168. Sonderman D. Comparing multi-server queues with finite waiting rooms, II: Different number of servers / D. Sonderman // Advances in Applied Probability. – 1979. – Vol. 11 (2). – P. 448–455.
169. Titterington D. M. Statistical Analysis of Finite Mixture Distributions / D. M. Titterington, A. F. M. Smith, U. E. Makov. – USA : John Wiley and Sons Ltd., 1985.
170. Thorin O. On the infinite divisibility of the Pareto distribution / O. Thorin // Scandinavian Actuarial Journal. – 1977. – Vol. 1977, no. 1. – P. 31–40.
171. Thorisson H. Coupling, Stationarity, and Regeneration / H. Thorisson. – New York : Springer, 2000.
172. Thorisson H. The queue $GI/GI/1$: Finite moments of the cycle variables and uniform rates of convergence / H. Thorisson // Stochastic Processes and Their Applications. – 1985. – Vol. 19. – P. 85–99.
173. Varnosafaderani S. Modeling Repairs of Systems with a Bathtub-Shaped Failure Rate Function / S. Varnosafaderani, S. Chukova // Proceedings of 46th Annual Conference of the Operations Research Society of New Zealand (ORSNZ), Dec., 2012, Victoria University of Wellington. – Wellington : New Zealand, 2012. – P. 362–370.
174. Veraverbeke N. Asymptotic behaviour of Wiener-Hopf factors of a random walk / N. Veraverbeke // Stochastic Processes and Their Applications. – 1977. – Vol. 5, no. 1. – P. 27–37.
175. Weissman I. The Extremal Index and Clustering of High Values for Derived Stationary Sequences / I. Weissman, U. Cohen // Journal of Applied Probability. – 1995. – Vol. 32, no. 4. – P. 972–981.
176. Whitt W. Embedded renewal processes in the $GI/G/s$ queue / W. Whitt // Journal of Applied Probability. – 1972. – Vol. 9. – P. 650–658.

177. Whitt W. Comparing counting processes and queues / W. Whitt // Advances in Applied Probability. – 1981. – Vol. 13, no. 1. – P. 207–220.
178. Wong E. W. M. Towards a bufferless optical internet / E. W. M. Wong, L. L. H. Andrew, T. Cui [et al.] // Journal of Lightwave Technology. – 2009. – Vol. 27. – P. 2817–2833.
179. Whitt W. Characterizing superposition arrival processes in packet multiplexers for voice and data / W. Whitt // Journal of Selected Areas in Communications. – 1986. – Vol. 4. – P. 833–846.
180. Whitt W. Approximating a point process by a renewal process, I: Two basic methods / W. Whitt // Operations Research. – 1982. – Vol. 30. – P. 125–147.
181. Whitt W. Queues with superposition arrival processes in heavy traffic / W. Whitt // Stoch. Process. Their Appl. – 1985. – Vol. 21. – P. 81–91.
182. Yao S. Electrical ingress buffering and traffic aggregation for optical packet switching and their effect on TCP-level performance in optical mesh networks / S. Yao, F. Xue, B. Mukherjee [et al.] // IEEE Communications Magazine. – 2002. – Vol. 40 (9). – P. 66–72.
183. Zolotarev V. M. Modern Theory of Summing Independent Random Variables / V. M. Zolotarev. – Moscow : Nauka, 1986. (In Russ.)
184. Zorine A. Controlled M/M/1-RQ System with Randomized Acceptance from the Orbit / A. Zorine // In: Phung-Duc, Tuan; Kasahara, Shoji; Wittevrongel, Sabine (szerk.) Queueing Theory and Network Applications: 14th International Conference, QTNA 2019 : Proceedings Cham, Svájc : Springer Cham. 2019. PP. 64–76.
185. Weinberg G. V. Kullback-Leibler divergence and the Pareto-Exponential approximation / G. V. Weinberg // SpringerPlus. – 2016. – Vol. 5, art. no. 604. – P. 1–9.
186. Weissman I. On blocks and runs estimators of extremal index / I. Weissman, S. Novak // Journal of Statistical Planning and Inference. – 1998. – Vol. 66, no. 2. – P. 281–288.

187. Zimmer W. The Burr XII Distribution in Reliability Analysis / W. Zimmer, J. Keats, F. Wang // Journal of Quality Technology. – 1998. – Vol. 30 (4). – P. 386–394.
188. Zwart A. P. Queueing systems with heavy tails. PhD Thesis, Mathematics and Computer Science / A. P. Zwart. – Eindhoven : Technische Universiteit Eindhoven, 2001.

Публикации автора по теме диссертации

Публикации в рецензируемых научных изданиях, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук, и приравненных к ним научных журналах, входящих в наукометрические базы данных:

189. Peshkova I. The waiting time extremal index in $GI/G/1$ system / I. Peshkova // Discrete and Continuous Models and Applied Computational Science. – 2026. – Vol. 34(1). P. 24–39. (K2, BAK, Scopus)
190. Peshkova I. Splitting-Based Regenerations for Accelerated Simulation of Queues / I. Peshkova, M. Pagano, E. Morozov // Mathematics. – 2025. – Vol. 13, art.no. 2883. – P. 1–18. – URL : <https://www.mdpi.com/2227-7390/13/17/2883> (дата обращения : 01.03.2026). (K1, WoS)
191. Peshkova I. Regeneration and approximation of a queueing system fed by superposed input with Weibull components / Peshkova, M.Pagano, E. Morozov // Reliability: Theory & Applications, Special Issue. – 2025. – Vol. 20, no 8 (85). – P. 108–117. (K1, Scopus)
192. Пешкова И. В. О максимумах стационарной задержки в системах $M/G/2$ / И. В. Пешкова // Вестник Московского университета. Серия 15: Вычислительная математика и кибернетика. – 2024. – Вып. 2. – С. 38–46. (K2, BAK)

193. Peshkova I. Regenerative analysis and approximation of queueing systems with superposed input processes / I. Peshkova, E. Morozov, M. Pagano // Mathematics. – 2024. – Vol. 12, no. 14, art. no. 2202. – P. 1–22. – URL : <https://www.mdpi.com/2227-7390/12/14/2202> (дата обращения : 01.03.2026). (K1, WoS)
194. Пешкова И. В. Границы незавершенной работы в системе с повторными вызовами разных классов и показательным временем обслуживания / И. В. Пешкова // Информатика и её применения. – 2023. – Т. 17, вып. 4. – С. 57–63. (K1, BAK, RSCI)
195. Peshkova I. Exponential-Pareto mixture distribution / I. Peshkova // Reliability: Theory & Applications. – 2023. – Vol. 18, no. 4 (76). – P. 632–645. (K1, Scopus)
196. Morozov E. V. Bounds and Maxima for the Workload in a Multiclass Orbit Queue / E. V. Morozov, I. V. Peshkova, A. S. Rumyantsev // Mathematics. – 2023. – Vol. 11, art. no. 564. – P. 1–15. – URL : <https://www.mdpi.com/2227-7390/11/3/564> (дата обращения : 01.03.2026). (K1, WoS)
197. Peshkova I. On comparison of multiserver systems with multicomponent mixture distributions / I. Peshkova, E. Morozov // Journal of Mathematical Sciences. – 2022. – Vol. 267, no. 2. – P. 260–272. (K1, WoS)
198. Пешкова И. В. Сравнение экстремальных индексов времен ожидания в системах обслуживания M/G/1 / И. В. Пешкова // Информатика и её применения. – 2022. – Т. 16, вып. 1. – С. 61–67. (K1, BAK, RSCI)
199. Пешкова И. В. Границы экстремального индекса времени ожидания в системе M/G/1 с распределением времени обслуживания в виде конечной смеси / И. В. Пешкова // Информатика и её применения. – 2022. – Т. 16, Вып. 4. – С. 26–33. (K1, BAK, RSCI)
200. Morozov E. V. Far-end-tail estimation of queueing system performance / E. V. Morozov, A. S. Rumyantsev, I. V. Peshkova // Journal of Mathematical Sciences. – 2020. – Vol. 248, art. no. 1. – P. 80–91. (K1, Scopus)

201. Morozov E. Sensitivity Analysis and Simulation of a Multiserver Queueing System with Mixed Service Time Distribution / E. Morozov, M. Pagano, I. Peshkova, A. Rumyantsev // Mathematics. – 2020. – Vol. 8, art. no. 1277. – P. 1–16. – URL : <https://www.mdpi.com/2227-7390/8/8/1277> (дата обращения : 01.03.2026). (K1, WoS)
Статьи в сборниках материалов конференций, индексируемых в международных базах данных в Scopus и/или Web of Science
202. Peshkova I. On waiting time maxima in queues with exponential-pareto service times / I. Peshkova, A. Golovin, M. Maltseva // Distributed Computer and Communication Networks. DCCN 2022. Communications in Computer and Information Science. – Cham : Springer, 2023. – Vol. 1748. – P. 147–160.
203. Peshkova I. On comparison of waiting time extremal indexes in queueing systems with Weibull service times / I. Peshkova, E. Morozov, M. Maltseva // ITMM 2021: Information Technologies and Mathematical Modelling: Queueing Theory and Applications, Communications in Computer and Information Science. – Switzerland: Springer Nature, 2022. – Vol. 1605. – P. 80–92.
204. Rumyantsev A. S. Exponential splitting based artificial regeneration in supercomputer queueing model / A. S. Rumyantsev, I. V. Peshkova // DCCN 2022: Distributed Computer and Communication Networks: Control, Computation, Communications. Lecture Notes in Computer Science. – Switzerland : Springer Nature, 2022. – Vol. 13766. – P. 385–396.
205. Peshkova I. On regenerative estimation of extremal index in queueing systems / I. Peshkova, E. Morozov, M. Maltseva // DCCN 2021: Distributed Computer and Communication Networks: Control, Computation, Communications. Lecture Notes in Computer Science / Eds. V. M. Vishnevskiy, K. E. Samouylov, D. V. Kozyrev. – Cham : Springer, 2021. – Vol. 13144. – P. 251–264.
206. Peshkova I. On comparison of multiserver systems with Exponential-Pareto mixture distribution / I. Peshkova, E. Morozov, M. Maltseva // Communications in Computer and Information Science. – 2020. – Vol. 1231. – P. 141–152.
207. Peshkova I. On Comparison of Multiserver Systems with Two-component Mixture Distributions / I. Peshkova, E. Morozov, M. Maltseva //

- Communications in Computer and Information Science. – 2020. – Vol. 1337. – P. 340–352.
208. Peshkova I. On stochastic and failure rate orderings in systems with two-component service time mixture / I. Peshkova, E. Morozov // SMARTY 2020 – Proceedings of the 2nd International Workshop on Stochastic Modeling and Applied Research of Technology. – 2020. – Vol. 2792. – P. 203–212.
 209. Morozov E. On Failure Rate Comparison of Finite Multiserver Systems / E. Morozov, I. Peshkova, A. Rumyantsev // Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). – 2019. – Vol. 11965. – P. 419–431.
 210. Rumyantsev A. Artificial regeneration based regenerative estimation of multiserver system with multiple vacations policy / A. Rumyantsev, I. Peshkova // Information Technologies and Mathematical Modelling: Queueing Theory and Applications. – Cham : Springer, 2019. – P. 38–50.
 211. Peshkova I. Classical regeneration based approaches for output analysis of multiserver systems simulation / I. Peshkova // SMARTY 2018 – Proceedings of the 1st International Workshop on Stochastic Modeling and Applied Research of Technology. – 2018. Vol. 2278. – P. 38–45.
 212. Morozov E. A regeneration-based estimation of high performance multiserver systems / E. Morozov, R. Nekrasova, I. Peshkova, A. Rumyantsev // Communications in Computer and Information Science.– Springer, 2016. – Vol. 608. – P. 271–282.
 213. Morozov E. On Regenerative Envelopes for Cluster Model Simulation / E. Morozov, I. Peshkova, A. Rumyantsev // DCCN 2016 : Distributed Computer and Communication Networks: 19th International Conference / Eds. V. M. Vishnevskiy, K. E. Samouylov, D. V. Kozyrev. – Cham : Springer, Revised Selected Papers, 2016. – Vol. 678. – P. 222–230.
 214. Morozov E. Monotonicity and stochastic bounds for simultaneous service multiserver systems / E. Morozov, A. Rumyantsev, I. Peshkova // International

Congress on Ultra Modern Telecommunications and Control Systems and Workshops. – 2016. – P. 294–297.

Публикации в прочих научных изданиях

215. Пешкова И. В. Методы регенеративного моделирования для анализа многосерверных систем обслуживания / И. В. Пешкова, А. С. Румянцев // Труды Карельского научного центра Российской академии наук. – 2018. – Вып. 7. – С. 68–82.
216. Пешкова И. В. Об экстремальном индексе стационарного времени ожидания в системах M/G/1 с неоднородными входными потоками / И. В. Пешкова // Распределенные компьютерные и телекоммуникационные сети: управление, вычисление, связь (DCCN-2023) : материалы XXVI Международной научной конференции. – М., 2023. – С. 158–162.
217. Пешкова И. В. О предельном распределении максимума стационарного времени ожидания в GI/G/1 с экспоненциальным Парето обслуживанием / И. В. Пешкова // Распределенные компьютерные и телекоммуникационные сети: управление, вычисление, связь (DCCN-2022) : материалы XXV Международной научной конференции / под общ. ред. В. М. Вишневого, К. Е. Самуйлова. – М., 2022. – С. 274–278.
218. Пешкова И. В. Использование экстремального индекса для анализа характеристик качества обслуживания программно-конфигурируемых сетей / И. В. Пешкова, А. В. Жуков, М. А. Мальцева // Цифровые технологии в образовании, науке, обществе : материалы XV Всероссийской научно-практической конференции. – Петрозаводск, 2021. – С. 92–93.
219. Пешкова И. В. Метод регенеративных огибающих для анализа характеристик качества обслуживания программно-конфигурируемых сетей / И. В. Пешкова, А. В. Жуков, М. А. Мальцева // Цифровые технологии в образовании, науке, обществе : материалы XIV Всероссийской научно-практической конференции. – Петрозаводск, 2020. – С. 123–124.
220. Пешкова И. В. Использование функции интенсивности отказов и каплинга при сравнении систем обслуживания / И. В. Пешкова, М. А. Мальцева

- // Цифровые технологии в образовании, науке, обществе : материалы XIV Всероссийской научно-практической конференции. – Петрозаводск, 2020. – С. 124–126.
221. Rumyantsev A. Artificial regeneration based regenerative estimation of multiserver systems performance / A. Rumyantsev, I. Peshkova // Информационные технологии и математическое моделирование (ИТММ-2019) : материалы XVIII Международной конференции имени А. Ф. Терпугова. – Томск, 2019. – Р. 101–103.
222. Morozov E. On artificial regeneration for stochastic simulation of the high performance cluster model / E. Morozov, I. Peshkova, A. Rumyantsev // Distributed computer and communication networks: control, computation, communications (DCCN-2017) : материалы Двадцатой международной научной конференции / под общ. ред. В. М. Вишневого. – М., 2017. – Р. 95–97.
223. Морозов Е. В. Метод регенеративных огибающих для оценивания высокопроизводительных многосерверных систем / Е. В. Морозов, А. С. Румянцев, И. В. Пешкова // Вероятностные методы в дискретной математике. IX Международная Петрозаводская конференция : расширенные тезисы / науч. ред. В. В. Мазалов ; отв. ред. Е. Н. Спектор. – Петрозаводск : Институт прикладных математических исследований Карельского научного центра Российской академии наук, 2016. – С. 63–64.
224. Morozov E. On regenerative envelopes for high performance cluster simulation / E. Morozov, I. Peshkova, A. Rumyantsev // Distributed computer and communication networks: control, computation, communications (DCCN-2016). – Moscow, 2016. – Р. 337–341.
225. Гусев О. В. Способ идентификации перегрузки web-сервера при помощи нейронной сети / О. В. Гусев, А. В. Жуков, И. В. Пешкова // Ученые записки Петрозаводского государственного университета. – 2013. – Вып. 8 (137). – С. 114–117.

Перечень рисунков

1.1	Стохастический порядок и порядок почти наверное	31
1.2	Хвосты ф.р. показательного и равномерного распределений, $\lambda = 0.2$, $a = 1$, $b = 5$	34
1.3	Функции интенсивности отказов показательного и Парето распределений, $\lambda = 4.2$; $x_0 = 0.5$; $\alpha = 2.1$	35
1.4	Сравнение интенсивностей отказов распределения $Burr(x_0, \alpha^*, k)$ и гиперэкспоненциального распределения с параметрами $x_0 = 1$, $k = 3$, $n = 2$, $p_1 = p_2 = 0.5$, $\mu_1 = 0.2$, $\mu_2 = 2$ и значением $\alpha^* = 0.12$, удовлетворяющим условию (1.29).	39
1.5	Ф.р. показательного-Парето распределения.	47
1.6	Функции интенсивностей отказов распределений $Exp(1)$, $Pareto(0.5; 0.6)$ и смеси $Exp-Pareto(1; 0.5; 0.6; p)$ при различных значениях коэффициента смеси p	52
1.7	Средние величины $e(u)$ превышения уровня u для $Exp(5)$, $Pareto(2.1; 0.5)$ и смеси $Exp-Pareto(5; 2.1; 0.5; 0.5)$	53
1.8	Зависимость равномерной метрики $\Delta(F_Y, F_{Y_1})$ с параметром смеси $p = 0.9$ и верхней границей $\delta(\alpha_1, \alpha_2)$ от параметра α_2	60
1.9	Эмпирические ф.р. показательного, показательного-Парето и Парето распределений при различных α , λ , и $x_0 = 1$, $p = 0.5$	63
1.10	Функции интенсивности отказов $Exp(5.5)$, $Pareto(2, 0.5)$ и 2+2-компонентного показательного-Парето распределения с параметрами $x_0 = 0.5$, $\lambda_1 = 5.5$, $\lambda_2 = 5.0$, $\alpha_3 = 2.5$, $\alpha_4 = 2.0$ для различных значений пропорции смеси (а): $p_i = 0.25, i = 1, \dots, 4$; б): $p_1 = 0.10, p_2 = 0.15, p_3 = 0.15, p_4 = 0.60$, в): $p_1 = 0.60, p_2 = 0.15, p_3 = 0.15, p_4 = 0.10$	70
2.1	Сравнение эмпирической ф.р. длины очереди $Q^{(1)}$ в системе $M/G/4/5$ с распределением времени обслуживания $Pareto(1, 2.5)$ с эмпирической ф.р. длины очереди $Q^{(2)}$ в системе $M/M/4/5$ с распределением времени обслуживания $Exp(2.5)$; входной поток - пуассоновский с параметром $\lambda_T = 2$ в обеих системах.	79

- 2.2 Равномерная метрика между двумя эмпирическими ф.р. стационарной длины очереди в двух системах $G/G/4/5$: $Burr(x_1, \alpha_1, k)/Exp(\lambda)/4/5$ и $H_2/Pareto(\alpha_2, x_2)/4/5$ (для различных α_2), где $x_1 = 1, k = 3, p_1 = p_2 = 0.5, \mu_1 = 3, \mu_2 = 30, \alpha_1 = 1.8, \lambda = 2.5$ и $x_2 = 1, \alpha_2 \in \{1.5, 2, 2.5\}$ 82
- 2.3 Нижняя и верхняя границы доверительного интервала для средней длины очереди в 4-х серверной системе с показательным-Парето распределением времени обслуживания. 85
- 2.4 Нижняя и верхняя границы доверительного интервала для среднего времени ожидания в 4-х серверной системе с показательным-Парето распределением времени обслуживания. 86
- 2.5 Доверительные интервалы для средней длины очереди в системах $\Sigma, \Sigma^{(1)}, \Sigma^{(2)}$; $N = N_1 = N_2 = 4$, 87
- 2.6 Доверительные интервалы для среднего времени ожидания в системах $\Sigma, \Sigma^{(1)}, \Sigma^{(2)}$; $N = N_1 = N_2 = 4$ 88
- 2.7 Доверительный интервал для средней стационарной длины очереди в системах $\Sigma, \Sigma^{(1)}$, в которых размеры буферов $C = C_1 = 1$, число серверов $N = N_1 = 4$ и $p_1 = 0.1, p_1 = 0.1, p_2 = 0.15, p_3 = 0.15, p_4 = 0.6$ 91
- 2.8 Теоретическая ф.р. (2.74) для системы $\Sigma^{(1)}$ и эмпирические ф.р. для систем Σ и $\Sigma^{(2)}$ стационарной незавершенной работы в системах вида $M/Pareto/1$ с параметрами (2.83). 103
- 2.9 Ф.р. в минорантной $\Sigma^{(1)}$, исходной Σ и мажорантной $\Sigma^{(2)}$ системах при $\lambda = 10, p_1 = 1/2, p_2 = 1/3, p_3 = 1/6; \mu_1 = 10, \mu_2 = 30, \mu_3 = 60, \eta_1 = 50, \eta_2 = 100, \eta_3 = 150$ 108
- 2.10 Эмпирическая ф.р. стационарного времени ожидания в исходной системе $\Sigma = \sum_{i=1}^2 G_i/M/1$, теоретические ф.р. $F_{W_i}, i = 1, 2$ в системах $G_1/M/1$ и $G_2/M/1$ при $\mu = 4, \alpha_1 = 2, \alpha_2 = 3$ 113
- 2.11 Эмпирическая ф.р. стационарного времени ожидания в исходной системе $\Sigma = \sum_{i=1}^2 G_i/M/1$, теоретические ф.р. $F_{W_i}, i = 1, 2$ в системах $G_1/M/1$ и $G_2/M/1$ при $\mu = 5, \alpha_1 = 2, \alpha_2 = 3$ 113

3.1	Расстояние Колмогорова между эмпирическими ф.р. в системах Σ и $\tilde{\Sigma}$ в зависимости от ρ и α_2	129
3.2	Графики эмпирических ф.р. с.в. W и W_A при $\alpha_1 = 2$, $\alpha_2 = 2$, $x_0 = 1$, и $\mu = 3$	132
3.3	Графики эмпирических ф.р. с.в. W и W_A при $\alpha_1 = 2$, $\alpha_2 = 10$, $x_0 = 1$, и $\mu = 20$	132
3.4	Графики эмпирических ф.р. стационарного времени ожидания W в исходной системе Σ и стационарного времени ожидания W_A в аппроксимирующей системе $\tilde{\Sigma}$ при $\alpha = 2$, $x_0 = 1$, $\lambda = 1$, и $\mu = 3$. .	134
4.1	Экстремальные индексы стационарного времени ожидания в системах $M/M/1(\mu_1 = 4)$, $M/H_2/1(\mu_1 = 4, \mu_2 = 2, p = 0.5)$ и $M/M/1(\mu_2 = 2)$, $\lambda = \lambda_1 = \lambda_2 = 1$ в зависимости от перцентиля времени ожидания.	171
4.2	Оценки экстремальных индексов стационарного времени ожидания в исходной системе $\sum_{i=1}^2 G_i/M/1$ и системах $G_1/M/1$ и $G_2/M/1$, при $\mu = 4$, $\alpha_1 = 2$, $\alpha_2 = 3$	176
4.3	Оценки экстремального индекса стационарного времени ожидания в исходной системе $\sum_{i=1}^2 G_i/M/1$ и системах $G_1/M/1$ и $G_2/M/1$, при $\mu = 5$, $\alpha_1 = 2$, $\alpha_2 = 3$	176
5.1	Зависимость экстремального индекса стационарного времени ожидания от перцентиля порогового значения в системе $M/Pareto/1$ при $\lambda_\tau = 1.87$, $\alpha_2 = 2.1$, $x_0 = 0.5$	184
5.2	Экстремальные индексы стационарного времени ожидания в системах: $M/M/1$ с интенсивностью входного потока $\lambda = 0.4$, интенсивностью обслуживания $\mu = 0.5$, в системе $M/We/1$ с интенсивностью входного потока $\lambda = 0.4$ и распределением Вейбулла времени обслуживания с параметрами $\alpha = 1$, $\beta = 0.5$	188
5.3	Экстремальный индекс стационарного времени ожидания при $\lambda_\tau = 1.87$, $p = 0.5$, $x_0 = 0.5$, $\lambda = 4.3$, $\alpha = 2.1$	197
5.4	Экстремальный индекс стационарной длины очереди при $\lambda_\tau = 1.87$, $p = 0.5$, $x_0 = 0.5$, $\lambda = 4.3$, $\alpha = 2.1$	197

5.5	Зависимость экстремального индекса стационарного времени ожидания от перцентиля в системах $M/G/2$ на рисунке 5.5а: $x_0 = 1, \alpha = 2, \lambda = 0.8$; на рисунке 5.5б: $x_0 = 1, \alpha = 3, \lambda = 1.6$	206
5.6	Зависимость экстремального индекса стационарного времени ожидания от квантиля порогового значения в системах $M/G/2$ на рисунке 5.6а: $x_0 = 1, \alpha = 2, \lambda = 1, 5$; на рисунке 5.6б: $x_0 = 1, \alpha = 3, \lambda = 3$	207
5.7	Зависимость экстремального индекса стационарного времени ожидания от перцентиля в системах $M/G/2$ на рисунке 5.7а: $x_0 = 1, \alpha = 6, \lambda = 7, 5$; на рисунке 5.7б: $x_0 = 1, \alpha = 11, \lambda = 15$. . .	207
6.1	Система $M/G/10$ ($\rho = 9$) Оценки средних длин очередей и 95% доверительный интервал, построенный на основе доверительных интервалов для средней длины очереди в мажорантной и минорантной системах.	218
6.2	Система $GI/G/10$ ($\rho = 9$): 95% доверительный интервал для средней стационарной длины очереди.	218
6.3	Сравнение частоты классической регенерации и частоты искусственной регенерации в системе $Pareto/M/1$ в зависимости от α	228
6.4	Отношение частоты искусственной регенерации к частоте классической регенерации в зависимости от α в системе $Pareto/M/1$	228
6.5	Графическая интерпретация неравенств (6.33).	239

Список таблиц

3.1	Расстояние Колмогорова между эмпирической ф.р. времени ожидания W в исходной системе Σ и эмпирической ф.р. W_A в аппроксимирующей системе $\tilde{\Sigma}$, теоретической ф.р. с.в. W_T . при $\alpha_1 = 2$ и $x_0 = 1$	131
3.2	Расстояние Колмогорова между эмпирической ф.р. времени ожидания W в исходной системе Σ и эмпирической ф.р. W_A в аппроксимирующей системе $\tilde{\Sigma}$, теоретической ф.р. с.в. W_T	134
3.3	Расстояние Колмогорова между эмпирической ф.р. стационарного времени ожидания W в исходной системе Σ , стационарного времени ожидания W_A в аппроксимирующей системе $\tilde{\Sigma}$ с двумя классами обслуживания.	135
3.4	Расстояние Колмогорова $d(W, W_A)$, оценки м.о. $Mean(W)$, $Mean(W_A)$ и дисперсий $Var(W)$, $Var(W_A)$ при различных ρ	137
3.5	Расстояние Колмогорова $d(W, W_A)$, оценки м.о. $Mean(W)$, $Mean(W_A)$ и дисперсий $Var(W)$, $Var(W_A)$ при различных ρ	138
6.1	Число регенераций в системе $M/M/1$ ($\lambda = 1$).	233
6.2	Число регенераций в системе $H_2/M/1$	233
6.3	Число регенераций в системе $Pareto/M/1$ ($\mu = 1, x_0 = 1$).	234
6.4	Число регенераций в $Pareto/M/1$ ($\alpha = 2.1, x_0 = 1$).	234
6.5	Длины циклов регенераций в $Pareto/M/1$ ($\rho = 0.9, x_0 = 1$).	235
6.6	Оценки дисперсии в системе $Pareto/M/1$ при $\rho = 0.9$ и $x_0 = 1$	235
6.7	Оценка $P(W > u_w, Q > u_q)$ для $Pareto/M/1$ при $u_w = 5, u_q = 10$	236
6.8	Число регенераций $RegIn$ и $RegW$ и полуширина доверительного интервала 99%CI при различных δ и η	243
6.9	Число регенераций $RegIn$ и $RegW$ и полуширина доверительного интервала 99%CI при различных δ и η	244
6.10	Число регенераций $RegIn$ и $RegW$ и полуширина доверительного интервала 99%CI при различных δ и η	245
6.11	Частота регенераций $RegIn$ и $RegW$ и полуширина доверительного интервала 99%CI при различных δ и η	245

- 6.12 Число регенераций RegW , Среднее время ожидания $\text{Mean}(W)$ и полуширина доверительного интервала $99\%CI$ при $i_0 = 2, \eta = 0.8164, \beta_1 = 0.5, \beta_2 = 0.67, \alpha_1 = 1.5, \alpha_2 = 2$ 246
- 6.13 Число регенераций RegW , Среднее время ожидания $\text{Mean}(W)$ и полуширина доверительного интервала $99\%CI$ при $i_0 = 1, \eta = 0.63, \beta_1 = 0.5, \beta_2 = 0.67, \alpha_1 = 1.5, \alpha_2 = 2$ 247

Приложение А. Свидетельства о регистрации программы для ЭВМ

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2025611480

ExtremeIndexSuperposedG/M/1

Правообладатель: *Пешкова Ирина Валерьевна (RU)*

Автор(ы): *Пешкова Ирина Валерьевна (RU)*

Заявка № 2024693294

Дата поступления 24 декабря 2024 г.

Дата государственной регистрации

в Реестре программ для ЭВМ 20 января 2025 г.



Руководитель Федеральной службы
по интеллектуальной собственности

Ю.С. Зубов

РОССИЙСКАЯ ФЕДЕРАЦИЯ



СВИДЕТЕЛЬСТВО

о государственной регистрации программы для ЭВМ

№ 2025694782

SPLITTINGSUPERPOSEDG/M/1

Правообладатели: **Пешкова Ирина Валерьевна (RU), Калинин Никита Сергеевич (RU)**

Авторы: **Пешкова Ирина Валерьевна (RU), Калинин Никита Сергеевич (RU)**



Заявка № **2025683004**

Дата поступления **07 августа 2025 г.**

Дата государственной регистрации

в Реестре программ для ЭВМ **08 декабря 2025 г.**

Руководитель Федеральной службы
по интеллектуальной собственности

Ю.С. Зубов