

**В. В. АНИСИМОВ
О. К. ЗАКУСИЛО
В. С. ДОНЧЕНКО**

ЭЛЕМЕНТЫ ТЕОРИИ МАССОВОГО ОБСЛУЖИВАНИЯ И АСИМПТОТИЧЕСКОГО АНАЛИЗА СИСТЕМ

Под общей редакцией
профессора
В. В. Анисимова

*Допущено Министерством
высшего и среднего образования
УССР в качестве учебного
пособия для студентов
вузов, обучающихся по
специальностям
«Математика»,
«Прикладная математика»,
«Экономическая
кибернетика»*

Киев
Головное издательство
издательского объединения
«Вища школа»
1987

22.18я73

A67

УДК 658.64; 658.818

Элементы теории массового обслуживания и асимптотического анализа систем / В. В. Анисимов, О. К. Закусило, В. С. Донченко. — К.: Вища шк. Головное изд-во, 1987. — 248 с.

В учебном пособии описаны основные классические модели систем массового обслуживания, модели некоторых специальных типов систем и систем, работающих в условиях большой нагрузки. Получены новые результаты по асимптотическому анализу потоков редких событий с малой вероятностью потери требования и асимптотическому укрупнению состояний стохастических систем. Приведены методы вероятностного моделирования.

Для студентов вузов, обучающихся по специальностям «Математика», «Прикладная математика», «Экономическая кибернетика».

Ил. 3. Библиогр.: 14 назв.

Рецензенты: д-р физ.-мат. наук *А. Д. Шаташвили* (зав. кафедрой математики и математических методов в экономике Донецкого госуниверситета), д-р физ.-мат. наук *А. Ф. Турбин* (зав. лабораторией прикладной статистики Института математики АН УССР).

Редакция учебной и научной литературы по информатике, вычислительной технике, кибернетике и АСУ
Зав. редакцией *С. М. Рудь*.

А 1502000000-082 111-87
М 211(04)-87

© Издательское объединение
«Вища школа», 1987

ОГЛАВЛЕНИЕ

Предисловие	5
Глава 1. Входящие потоки	7
1.1. Классификация систем	7
1.2. Простейший поток	10
1.3. О сходимости входящих потоков к пуассоновскому	20
1.4. Элементы теории восстановления. Полумарковские процессы	30
1.5. Процессы размножения и гибели	45
Глава 2. Модели систем массового обслуживания	51
2.1. Марковские системы	51
2.2. Система $M G 1$	66
2.3. Система $G M 1$	79
2.4. Система $GI G 1$	83
Глава 3. Некоторые специальные типы систем	85
3.1. Системы с групповым поступлением	85
3.2. Приоритетные системы	87
3.3. Система с периодическим регулярным входным потоком	89
3.4. Система с переменной интенсивностью обслуживания	98
Глава 4. Системы массового обслуживания в условиях большой загрузки	106
4.1. Система $M G 1$ в критическом режиме	106
4.2. Диффузионная аппроксимация систем массового обслуживания	110
4.3. Диффузионная аппроксимация величины очереди в системе $GI G g$	117
Глава 5. Предельные теоремы для потоков редких событий	125
5.1. О наступлении редкого события в схеме восстановления	125
5.2. Асимптотическое поведение времени пребывания полумарковского процесса в связанном подмножестве состояний	134

5.3. Асимптотически связанные подмножества	141
5.4. Системы с быстрым обслуживанием	153
5.5. Редящие потоки	163
5.6. Анализ потоков неоднородных редких событий	168
 Глава 6. Асимптотическое укрупнение сложных систем	187
6.1. Процессы с дискретным вмешательством случая	187
6.2. Асимптотическое укрупнение состояний случайных процессов	200
6.3. Асимптотическое укрупнение СМО	209
6.4. Поведение аддитивных функционалов для асимптотически укрупняемых систем	222
 Глава 7. Вероятностное моделирование	228
7.1. Моделирование основных распределений	228
7.2. Гауссовские и равномерные распределения вероятностей	233
7.3. Моделирование многомерных распределений	235
7.4. Имитация марковских процессов	238
7.5. Имитационное моделирование СМО	239
Предметный указатель	246
Список рекомендованной литературы	247

ПРЕДИСЛОВИЕ

Сложность современного производства и в связи с этим сложность и большая размерность соответствующих математических моделей приводит к необходимости развигать методы асимптотического анализа, которые в ряде случаев позволяют приближенно сводить изучение исходной модели к упрощенной и получать простые приближенные формулы.

В настоящем учебном пособии главы 1—3 являются введением в классическую теорию массового обслуживания. Приводится общая классификация систем, изучаются основные типы входящих потоков, даются необходимые сведения из теории восстановления, марковских и полумарковских процессов, рассматриваются классические модели и некоторые специальные типы систем.

Главы 4—6 посвящены методам асимптотического анализа систем.

В главе 4 дан краткий обзор результатов по диффузионной аппроксимации систем обслуживания и приведен ряд конкретных результатов по асимптотическому анализу ряда систем в условиях критической и надкритической загрузки. Глава 5 посвящена предельным теоремам для потоков редких событий. Аналитическими методами изучаются модели возникновения редких событий на процессах восстановления и полумарковских процессах, аналитическими конструктивными методами — свойства связанных множеств (марковских систем, обладающих свойствами асимптотической связности в течение времени пребывания в подмножестве), вероятностно-конструктивными методами — потоки неоднородных редких событий (с применением теории суммирования слабозависимых слагаемых и теорем о сходимости суперпозиций случайных процессов). Приведен анализ ряда полумарковских систем в условиях асимптотически быстрого восстановления.

Глава 6 посвящена анализу стохастических систем, допускающих асимптотическое укрупнение пространства со-

стояний. В ней развиваются вероятностно-конструктивные методы, основанные на аппарате переключающихся процессов, которые позволяют анализировать также неоднородные по времени модели. Под понятием асимптотического укрупнения системы понимается, что пространство состояний можно разбить на непересекающиеся области, вероятности переходов между которыми асимптотически малы. Оказывается, что при весьма общих предположениях, связанных с условиями перемешивания в каждой области, укрупненный процесс, описывающий переходы системы между областями, сходится к марковскому процессу, а аддитивные функционалы типа суммарного дохода сходятся к процессу с независимыми приращениями, переключаемому предельным укрупненным процессом. Эти результаты проиллюстрированы на моделях систем обслуживания с переменной структурой обслуживания, где возникают схемы естественного укрупнения пространства состояний.

В главе 7 вводятся основные понятия, связанные с вероятностным моделированием систем. Приводятся методы моделирования одномерных и многомерных распределений, марковских процессов, а также методы и алгоритмы имитационного моделирования систем обслуживания. Асимптотические подходы здесь заложены в структуре некоторых алгоритмов моделирования случайных чисел, а также в использовании метода Монте-Карло. П. 1.3, 1.5, главы 4—6 — написаны В. В. Анисимовым; п. 1.2, главы 2 и 3 — О. К. Закусило; п. 1.1, 1.4 совместно В. В. Анисимовым и О. К. Закусило, глава 7 — В. С. Донченко.

Учебное пособие написано на основании курсов, которые читаются на факультете кибернетики Киевского государственного университета и при Республиканском доме экономической и научно-технической пропаганды, и предназначено для студентов вузов, обучающихся по специальностям «Математика», «Прикладная математика», «Экономическая кибернетика», а также аспирантов и научных работников, специализирующихся в области приложений вероятностно-статистических методов к анализу систем.

ГЛАВА 1. ВХОДЯЩИЕ ПОТОКИ

1.1. Классификация систем

Основным объектом изучения в теории массового обслуживания является система массового обслуживания (СМО). Полагают, что СМО задана, если описан характер поступления требований (или заявок) в систему, т. е. входящий поток, указан характер обслуживания, т. е. количество обслуживающих устройств и закон обслуживания, и дисциплина очереди, т. е. принцип, по которому вновь поступившие в систему требования принимаются на обслуживание. Тем самым входящий поток, величина очереди и характер обслуживания являются основными параметрами СМО. Большое количество разнообразных условий, которым могут удовлетворять эти параметры, делает громоздким описание СМО обычным языком. Это привело к созданию нескольких вариантов условных обозначений для наиболее часто встречающихся СМО. Рассмотрим обозначения, предложенные Д. Кендаллом, а затем модифицированные Г. П. Башариным. В соответствии с ними СМО описывается последовательностью из пяти символов, разделенных вертикальными линиями. Первый из символов задает входящий поток, второй — определяет время обслуживания, третий — количество обслуживающих приборов, четвертый — количество мест для ожидания, пятый характеризует дисциплину выбора на обслуживание, если некоторые из требований имеют приоритет перед другими.

Так, буквы M, D, E, GI на месте первого символа означают, что длительности интервалов между моментами поступления требований независимы, одинаково распределены и имеют соответствующий показательное, вырожденное, эрланговское или произвольное распределение.

Те же буквы (M, D, E, G) на месте второго символа аналогичным образом определяют распределение времени обслуживания требований.

Третий и четвертый символы — числа, а пятый имеет вид f_i^j ; $i = 0, 1, 2$; $j = 0, 2$. При $i = 0$ рассматриваемая СМО является бесприоритетной, при $i = 1$ в ней задан относительный приоритет, а при $i = 2$ — абсолютный приоритет. Значение $j = 0$ указывает, что при отсутствии свободных мест для ожидания вновь прибывшее требование теряется, а $j = 2$ — что при тех же условиях оно вытесняет из очереди некоторое требование с самым низким приоритетом.

При относительном приоритете в момент освобождения обслуживающего устройства на обслуживание поступает требование с наивысшим из имеющихся приоритетов, а при абсолютном — на обслуживании может находиться только требование с наивысшим (из имеющихся) приоритетом.

Первый и второй символы могут иметь векторный вид, например, $\vec{M}$, $\vec{E}$, что указывает на наличие нескольких потоков различных типов требований, имеющих различные распределения времени обслуживания.

А. А. Боровков ввел другие способы задания систем массового обслуживания. Один из них использует последовательности четырехмерных векторов $\{\tau_i^e, \nu_i^e, \tau_i^s, \nu_i^s; i \geq 0\}$ (названные им управляющими), компоненты которых имеют следующий смысл: τ_i^e — длительности последовательных интервалов между поступлениями требований, $i \geq 0$; ν_i^e — количества требований, поступающих в моменты $\tau_0^e + \tau_1^e + \dots + \tau_i^e$, $i \geq 0$; τ_i^s — время обслуживания i -й партии требований; ν_i^s — количество требований, принятых на обслуживание в i -й партии (если в системе находится меньшее количество, то партия будет состоять из имеющихся требований).

Второй способ Боровкова определяет СМО с помощью трехмерного случайного процесса $s(t) = (e(t), r(t), s(t); t \geq 0)$, где $e(t)$ — число требований, поступивших в систему; $r(t)$ — число требований, получивших отказ; $s(t)$ — число обслуженных требований к моменту времени t , причем $e(t) \geq r(t) + s(t)$. Последнее неравенство означает, что очередь $Q(t)$ (количество требований в системе) неотрицательна:

$$Q(t) = e(t) - r(t) - s(t) \geq 0.$$

При пользовании как первым, так и вторым способом, кроме задания управляющих последовательностей или процесса $s(t)$, во многих задачах необходимо знание алгоритмов, определяющих дисциплину выбора на обслуживание,

поведение различных требований в очереди и т. д. Эти же алгоритмы нужно указывать и при пользовании символикой Кендалла — Башарина.

Укажем особенности некоторых СМО, которые либо должны быть указаны дополнительно при пользовании символикой Кендалла-Башарина, либо вообще исключают возможность пользоваться этой символикой:

1) длительности интервалов между моментами поступления требований могут быть зависимыми случайными величинами;

2) СМО могут отличаться начальным количеством требований;

3) СМО может быть замкнутой системой, в которой часть (или все) обслуженных требований формируют входной поток;

4) требования могут быть «скоропортящимися» и их время ожидания (или полное время пребывания в системе) ограничено;

5) требование может находиться в очереди к нескольким или всем обслуживающим устройствам;

6) требование может становиться в самую короткую (для быстрого обслуживания) либо в самую длинную очередь (такие ситуации возникают при появлении товаров, пользующихся повышенным спросом, а также перед окончанием работы одной из касс);

7) выбор требований на обслуживание может осуществляться по принципу «первым прибыл — первым обслужен», случайным образом либо в инверсном порядке — «первым прибыл — последним обслужен». С инверсным порядком приходится сталкиваться при изучении процесса продажи товара, когда первые его партии отправляются в дальние помещения склада либо располагаются на задних полках;

8) отдельные (или все) обслуживающие устройства могут быть специализированными, удовлетворять требования разных типов с разной скоростью;

9) обслуживающие устройства могут выходить из строя, т. е. на их ремонт необходимо затрачивать некоторое время;

10) система может быть многофазной, т. е. такой, в которой обслуживание требования происходит последовательно на нескольких (или одном) обслуживающих устройствах.

В пособии в основном используется символика Кендалла (пятый символ опущен). Так, к примеру, СМО типа $GI|G|s|m$ обозначает систему, в которой имеется s обслуживающих приборов и m мест для ожидания. Требования в систему поступают через независимые одинаково распределенные промежутки времени с некоторой произвольной

функцией распределения. Поступившие требования занимают любой из свободных обслуживающих приборов на случайное время (время обслуживания), которое не зависит от остальных случайных величин, связанных с другими требованиями и имеет функцию распределения $G(x)$. После окончания обслуживания требование покидает систему. Если все приборы заняты, требование становится в очередь, если все m мест для ожидания заняты, требование теряется.

В тех случаях, когда $m = 0$ либо ∞ , обычно четвертый символ опускают, а систему называют системой с потерями либо соответственно системой с очередью.

Основными характеристиками, которые изучаются для таких систем, являются: величина очереди $Q(t)$ — количество всех требований, находящихся в системе в момент t (включая требования, находящиеся на обслуживании и в очереди); время ожидания $w(t)$ (иногда говорят виртуальное время ожидания) — промежуток времени, через который требование, поступившее в систему в момент t , начнет обслуживаться; период занятости — промежуток времени, начинающийся с момента поступления требования в пустую систему и заканчивающийся в ближайший момент полного освобождения системы.

ЗАДАНИЯ ДЛЯ САМОКОНТРОЛЯ

1. Дать пример работы системы массового обслуживания.
2. Указать характеристики, представляющие интерес при изучении систем массового обслуживания.

1.2. Простейший поток

В теории массового обслуживания случайным потоком (или просто потоком) называется неубывающая последовательность неотрицательных случайных моментов времени t_i , $i \geq 1$. Момент t_i может интерпретироваться как момент поступления i -го по счету требования, и в этом случае говорят о входящем потоке. Последовательность $\{t_i, i \geq 1\}$ определяет последовательность $\{\tau_i, i \geq 1\}$, $\tau_1 = t_1$, $\tau_i = t_i - t_{i-1}$, $i \geq 2$ длительностей интервалов между соседними точками потока. Очевидно, что случайные величины τ_i однозначно определяют случайный поток.

Определение 1.1. Случайный поток называется простейшим, если величины τ_i , $i \geq 1$ независимы в совокупности, одинаково распределены и имеют показательное распределение: $P\{\tau_i > t\} = e^{-\lambda t}$; $t \geq 0$, $\lambda > 0$. Изучим некоторые свойства простейшего потока.

1. Средняя длина интервала между точками потока равна $1/\lambda$. В самом деле

$$M\tau_i = \int_0^{\infty} t dP\{\tau_i < t\} = \int_0^{\infty} t \lambda e^{-\lambda t} dt = \frac{1}{\lambda}.$$

2. Для простейшего потока случайная величина t_k имеет плотность распределения $g_k(x) = \lambda \frac{(\lambda x)^{k-1}}{(k-1)!} e^{-\lambda x}$, $x > 0$ и функцию распределения

$$G_k(x) = 1 - e^{-\lambda x} \left(1 + \frac{\lambda x}{1!} + \dots + \frac{(\lambda x)^{k-1}}{(k-1)!} \right), \quad x \geq 0.$$

Для доказательства воспользуемся методом математической индукции.

При $k=1$, $g_1(x) = \lambda e^{-\lambda x}$, $x > 0$, что является плотностью показательного распределения.

Предположим, что $g_k(x)$ имеет указанный выше вид. Так как

$$t_{k+1} = \tau_1 + \tau_2 + \dots + \tau_k + \tau_{k+1} = t_k + \tau_{k+1},$$

где случайные величины t_k , τ_{k+1} независимы, то

$$\begin{aligned} g_{k+1}(x) &= \int_0^x g_k(x-y) g_1(y) dy = \\ &= \frac{\lambda^{k+1}}{(k-1)!} e^{-\lambda x} \int_0^x (x-y)^{k-1} dy = \lambda \frac{(\lambda x)^k}{k!} e^{-\lambda x}. \end{aligned}$$

Для проверки второй половины утверждения достаточно продифференцировать $G_k(x)$ и убедиться, что $G'_k(x) = g_k(x)$. Плотность $g_k(x)$ задает так называемое распределение Эрланга k -го порядка.

3. Количество точек $v(t)$ простейшего потока на интервале $[0, t)$ имеет пуассоновское распределение с математическим ожиданием λt . Имеем $P\{v(t) = k\} = P\{t_k < t \leq t_{k+1}\} = P\{t_{k+1} \geq t\} - P\{t_k \geq t\} = 1 - G_{k+1}(t) - (1 - G_k(t)) = e^{-\lambda t} \left(1 + \frac{\lambda t}{1!} + \dots + \frac{(\lambda t)^k}{k!} \right) - e^{-\lambda t} \left(1 + \frac{\lambda t}{1!} + \dots + \frac{(\lambda t)^{k-1}}{(k-1)!} \right) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}.$

4. Количество $v(t+s) - v(s)$ точек простейшего потока на интервале $[s, t+s)$ также имеет пуассоновское распределение с математическим ожиданием λt и не зависит от $v(s)$. Для проверки этого свойства достаточно показать, что $P\{v(s) = i, v(t+s) - v(s) = j\} = P\{v(s) = i\} P\{v(t) = j\}.$

Рассмотрим сначала случай $i \geq 1, j \geq 1$. Используя формулу полной вероятности и независимость τ_k , получим

$$\begin{aligned}
 P\{v(s) = i, v(t+s) - v(s) = j\} &= \int_0^s P\{t_i \in du, \\
 \tau_{i+1} \geq s-u, \tau_{i+1} + \dots + \tau_{i+j} &< t+s-u, \tau_{i+1} + \dots \\
 \dots + \tau_{i+j+1} \geq t+s-u\} &= \int_0^s P\{t_i \in du\} \int_{s-u}^{s+t-u} P\{\tau_i \in dv\} \times \\
 &\times P\{v + \tau_{i+2} + \dots + \tau_{i+j} < t+s-u, \\
 &\quad v + \tau_{i+2} + \dots + \tau_{i+j+1} \geq t+s-u\} = \\
 &= \int_0^s g_i(u) \int_{s-u}^{s+t-u} P\{\tau_1 \in dv\} P\{\tau_1 + \dots + \tau_{j-1} < t+s-v-u; \\
 &\quad \tau_1 + \dots + \tau_j \geq t+s-v-u\} du = \int_0^s g_i(u) \times \\
 &\times \int_{s-u}^{s+t-u} \lambda e^{-\lambda v} P\{v(t+s-v-u) = j-1\} dv du = \\
 &= \int_0^s \lambda \frac{(\lambda u)^{i-1}}{(i-1)!} e^{-\lambda u} \int_{s-u}^{s+t-u} \lambda e^{-\lambda v} e^{-\lambda(t+s-v-u)} \times \\
 &\times \frac{[\lambda(t+s-v-u)]^{j-1}}{(j-1)!} dv du = \int_0^s \lambda \frac{(\lambda u)^{i-1}}{(i-1)!} e^{-\lambda(t+s)} \times \\
 &\times \frac{(\lambda t)^j}{j!} du = e^{-\lambda s} \frac{(\lambda s)^i}{i!} e^{-\lambda t} \frac{(\lambda t)^j}{j!} = P\{v(s) = i\} P\{v(t) = j\}.
 \end{aligned}$$

Рассмотрим случай, когда $i \geq 1, j = 0$:

$$\begin{aligned}
 P\{v(s) = i, v(t+s) - v(s) = 0\} &= \int_0^s P\{t_i \in du, \\
 \tau_{i+1} \geq t+s-u\} du &= \int_0^s P\{t_i \in du\} P\{\tau_{i+1} \geq t+s-u\} = \\
 &= \int_0^s \lambda \frac{(\lambda u)^{i-1}}{(i-1)!} e^{-\lambda u} e^{-\lambda(t+s-u)} du = e^{-\lambda(t+s)} \frac{(\lambda s)^i}{i!} = \\
 &= P\{v(s) = i\} P\{v(t) = 0\}.
 \end{aligned}$$

При $i = 0, j \geq 1$ имеем

$$\begin{aligned}
 P\{v(s) = 0, v(t+s) - v(s) = j\} &= P\{\tau_1 \geq s; \\
 \tau_1 + \dots + \tau_j < t+s, \tau_1 + \dots + \tau_{j+1} \geq t+s\} &= \\
 &= \int_s^{t+s} P\{\tau_1 \in du\} P\{u + \tau_2 + \dots + \tau_j < t+s, u + \tau_2 + \\
 + \dots + \tau_{j+1} \geq t+s\} &= \int_s^{t+s} \lambda e^{-\lambda u} P\{v(t+s-u) = j-1\} du = \\
 &= \int_s^{t+s} \lambda e^{-\lambda(t+s-u)} \frac{[\lambda(t+s-u)]^{j-1}}{(j-1)!} du = e^{-\lambda(t+s)} \frac{(\lambda t)^j}{j!} = \\
 &= P\{v(s) = 0\} P\{v(t) = j\}.
 \end{aligned}$$

И, наконец, при $i = j = 0$ получим

$$\begin{aligned}
 P\{v(s) = 0, v(t+s) - v(s) = 0\} &= \\
 &= P\{v(t+s) = 0\} = e^{-\lambda(t+s)} = \\
 &= P\{v(t) = 0\} P\{v(s) = 0\}.
 \end{aligned}$$

Итак, $P\{v(s) = i, v(t+s) - v(s) = j\} = P\{v(s) = i\} \times P\{v(t) = j\}$, и требуемое представление получено.

Свойство 4 допускает обобщение: _____

5. Если интервалы $[a_k, b_k)$, $k = 1, n$ не пересекаются, то количества $v(b_k) - v(a_k)$ точек простейшего потока на них — независимые в совокупности случайные величины, имеющие пуассоновские распределения со средними $\lambda(b_k - a_k)$.

Из свойства 3 следует, что длительность τ интервала от случайной точки s до очередной точки потока имеет показательное распределение со средним $1/\lambda$:

$$P\{\tau > t\} = P\{v(t+s) - v(s) = 0\} = e^{-\lambda t}.$$

На первый взгляд, этот результат парадоксален — ведь длительность целого интервала между точками t_k и t_{k+1} имеет среднюю длину $1/\lambda$, и казалось бы естественным, что, выбрав точку s внутри интервала между соседними точками потока, мы разбили его на две части, средняя длина каждой из которых по соображениям симметрии равна $1/2\lambda$.

Разрешение этого кажущегося противоречия заключается в том, что интервал, в который попадает точка s , имеет вероятностные характеристики, отличные от характеристик

интервалов потока $[t_k, t_{k+1})$. Функция распределения его длины θ по формуле полной вероятности имеет вид

$$P\{\theta < t\} = \sum_{k=0}^{\infty} P\{t_k \leq s < t_{k+1}\} P\{t_{k+1} - t_k < t/t_k \leq s < t_{k+1}\}$$

и не обязана совпадать с вероятностью $P\{t_{k+1} - t_k < t\} = 1 - e^{-\lambda t}$, $t \geq 0$. Вычисления показывают, что плотность распределения величины θ в самом деле не является показательной и равна

$$v_s(t) = \begin{cases} \lambda^2 t e^{-\lambda t}, & 0 < t \leq s; \\ \lambda(1 + \lambda s) e^{-\lambda t}, & t > s. \end{cases}$$

Свойство 5 является характеристическим для простейшего потока в том смысле, что любой поток, обладающий им, должен быть простейшим. Докажем даже более сильное утверждение. Обозначим, как и ранее, количество точек потока на интервале $[0, t)$ через $v(t)$ и дадим такие определения.

Определение 1.2. Случайный поток $\{t_k, k \geq 1\}$ называется потоком без последствия, если количества его точек на непересекающихся интервалах — независимые в совокупности случайные величины.

Свойство отсутствия последствия называют также марковским свойством.

Определение 1.3. Поток $\{t_k, k \geq 1\}$ называется стационарным, если распределение вектора $(v(s + b_i) - v(s + a_i), i = \overline{1, k})$ не зависит от s при любом конечном k .

Определение 1.4. Поток $\{t_k, k \geq 1\}$ называется ординарным, если вероятность $\psi_s(t)$, того, что на интервале $[s, s + t)$ находится не менее двух его точек, удовлетворяет соотношению $\frac{1}{t} \psi_s(t) \rightarrow 0$ при $t \rightarrow 0$.

Из свойства 5 следует, что простейший поток является стационарным ординарным потоком без последствия. Замечательным является то, что верно и обратное утверждение.

Теорема 1.1. Стационарный ординарный поток без последствия является простейшим.

Доказательство. Найдем вначале вероятность $P_0(t)$ того, что на интервале $[s, s + t)$ длины t нет точек потока. Для этого разобьем этот интервал на два: $[s, s + t_1)$ и $[s + t_1, s + t)$, $0 \leq t_1 < t$. Так как $P_0(t) = P\{\text{на } [s, s + t_1) \text{ нет точек, на } [s + t_1, s + t) \text{ нет точек потока}\}$, то, исполь-

зая свойства отсутствия последствия и стационарности, получим, что $P_0(t) = P_0(t_1)P_0(t - t_1)$, где $0 \leq t_1 \leq t$.

Так как $P_0(t)$ — вероятность, то $0 \leq P_0(t) \leq 1$. Из курса анализа известно, что ограниченным невозрастающим решением последнего функционального уравнения должна быть экспонента:

$$P_0(t) = e^{-\lambda t}; \quad t \geq 0; \quad 0 \leq \lambda < \infty.$$

Рассмотрим крайние возможности для λ :

а) $\lambda = 0$. В этом случае $P_0(t) = 1$, и на любом интервале $[0, t)$ точки потока отсутствуют с вероятностью 1, что не согласуется с нашим определением потока, в соответствии с которым $P\{t_1 < \infty\} = 1$;

б) $\lambda = \infty$. В этом случае при $t > 0$, $P_0(t) = 0$, и на каждом интервале сколь угодно малой длины Δ присутствует с вероятностью 1 хотя бы одна точка потока. Отсюда немедленно следует, что на любом интервале с вероятностью 1 находится бесконечно много точек, и $\psi_s(t) = 1$, что противоречит условию ординарности.

Итак, остается возможным лишь случай $0 < \lambda < \infty$.

Найдем асимптотическое поведение при $t \rightarrow 0$ вероятности $P_1(t)$, того, что на интервале длины t находится одна точка.

Имеем

$$P_1(t) = 1 - P_0(t) - \psi_s(t) = 1 - e^{-\lambda t} - 0(t) = \lambda t + 0(t).$$

Теперь можно найти совместное распределение случайных величин τ_1 и τ_2 . Обозначим через $A_i(s, t)$ событие, заключающееся в том, что на интервале $[s, t)$ находится i точек потока.

Пусть s и t фиксированы, $\Delta < t$. Так как

$$\begin{aligned} P\{A_0(0, s) \cap A_1(s, s + \Delta) \cap A_0(s + \Delta, s + \Delta + t)\} &\leq \\ &\leq P\{P\tau_1 \in [s, s + \Delta), \tau_2 > t\} \leq P\{A_0(0, s) \cap A_1(s, s + \\ &+ \Delta) \cap A_0(s + \Delta, s + t)\}, \end{aligned}$$

то, используя свойство отсутствия последствия и стационарность потока, получим $P_0(s)P_1(\Delta)P_0(t) \leq P\{\tau_1 \in [s, s + \Delta), \tau_2 > t\} \leq P_0(s)P_1(\Delta)P_0(t - \Delta)$.

В соответствии с полученными выражениями для $P_0(t)$ и $P_1(t)$ из последних неравенств следует, что

$$P\{\tau_1 \in [s, s + \Delta), \tau_2 > t\} = \lambda e^{-\lambda s} \Delta e^{-\lambda t} + 0(\Delta).$$

Разделив обе части последнего равенства на Δ , перейдем к пределу при $\Delta \rightarrow 0$ и продифференцируем полученное выражение по t . В результате получим совместную плотность распределения τ_1 и τ_2 , равную произведению $\lambda e^{-\lambda s} \lambda \times$

$\times e^{-\lambda t}$ плотностей $\lambda e^{-\lambda s}$ и $\lambda e^{-\lambda t}$ случайных величин с показательным распределением ($s > 0, t > 0$). Поэтому τ_1 и τ_2 независимы, одинаково распределены и имеют показательное распределение.

Аналогичным образом утверждение теоремы проверяется и для большего количества τ_i .

Доказанная теорема позволяет понять, почему во многих реальных ситуациях входной поток можно считать простейшим или мало отличающимся от него. Так, поток космических частиц, по-видимому, обладает свойствами стационарности, отсутствия последействия и ординарности (последнее свойство может интерпретироваться как невозможность практически одновременного появления двух частиц).

Поток телефонных вызовов на АТС часто может считаться простейшим. Вместе с тем каждое из трех необходимых для этого свойств при определенных обстоятельствах может быть нарушено. Так, рассматривая поток вызовов на АТС на протяжении суток, необходимо учесть, что в ночное время интенсивность ниже, чем днем, и поэтому отказываться от предположения о стационарности.

В случае, когда руководитель организации поручает секретарю пригласить на совещание необходимых ему лиц, один телефонный звонок вызывает появление множества других, и поэтому более естественно отказаться от ординарности, и отсутствия последействия.

Указанные свойства простейшего потока порождены характеристическим свойством показательного распределения, которое имеют величины τ_i :

$$P\{\tau_i > s + t / \tau_i > s\} = P\{\tau_i > t\}, \quad s \geq 0, \quad t \geq 0. \quad (1.1)$$

Это свойство легко проверить:

$$\begin{aligned} P\{\tau_i > s + t / \tau_i > s\} &= \frac{P\{\tau_i > s + t, \tau_i > s\}}{P\{\tau_i > s\}} = \\ &= \frac{P\{\tau_i > s + t\}}{P\{\tau_i > s\}} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = P\{\tau_i > t\}. \end{aligned}$$

Заметим, что верно и обратное утверждение: если распределение случайной величины τ непрерывно и удовлетворяет (1.1), то τ имеет показательное распределение. Действительно, по формуле условной вероятности (1.1) можно записать в виде

$$\begin{aligned} P\{\tau > s + t / \tau > s\} &= \frac{P\{\tau > s + t, \tau > s\}}{P\{\tau > s\}} = \\ &= \frac{P\{\tau > s + t\}}{P\{\tau > s\}} = P\{\tau > t\} \end{aligned}$$

(поскольку $\{\tau > s + t\} \subset \{\tau > s\}$, то $P\{\tau > s + t, \tau > s\} = P\{\tau > s + t\}$). Обозначим $q(t) = P\{\tau > t\}$. Тогда $q(s + t) = q(s)q(t)$. Отсюда $q(t) = e^{-\lambda t}$, $0 < \lambda$, т. е. $P\{\tau < t\} = 1 - e^{-\lambda t}$, что доказывает наше утверждение. Это свойство также называют свойством отсутствия последствия и оно может быть интерпретировано с помощью такого примера.

Пример 1.1. Пусть длительность безотказной работы прибора имеет показательное распределение с параметром λ и к моменту времени s отказ не наступил. Тогда оставшаяся длительность безотказной работы не зависит от s и имеет то же показательное распределение с параметром λ .

Укажем теперь другой вывод формул для распределения $v(t)$, использующий свойства стационарности, ординарности и отсутствия последствия. Для любого $t > 0$, $h > 0$ имеют место соотношения

$$\begin{aligned} P_0(t + h) &= P_0(t) P_0(h); \\ P_k(t + h) &= p_k(t) P_0(h) + P_{k-1}(t) P_1(h) + \\ &+ \sum_{i=2}^k P_{k-i}(t) P_i(h) \quad (k \geq 1). \end{aligned}$$

Учитывая доказанные соотношения:

$$P_0(h) = 1 - \lambda h + o(h); \quad P_1(h) = \lambda h + o(h); \quad P_i(h) = o(h),$$

$i \geq 2$, получим, что при малых h система примет вид

$$\begin{aligned} P_0(t + h) &= P_0(t) (1 - \lambda h) + o(h); \\ P_k(t + h) &= P_k(t) (1 - \lambda h) + P_{k-1}(t) \lambda h + o(h), \quad k \geq 1. \end{aligned}$$

Перенося $P_k(t)$ в левую часть и разделив обе части k -го равенства на h , получим, что при $h \rightarrow 0$ предел правой части существует, следовательно, существуют и производные $P'_k(t)$ и они удовлетворяют системе дифференциальных уравнений

$$\begin{cases} P'_0(t) = -\lambda P_0(t); \\ P'_k(t) = -\lambda P_k(t) + \lambda P_{k-1}(t), \quad k \geq 1. \end{cases} \quad (1.2)$$

При этом будем считать, что $v(0) = 0$ (в начальный момент нет требований), а это означает, что $P_0(0) = 1$, а $P_k(0) = 0$, $k \geq 1$.

Найдем решение в виде $P_k(t) = e^{-\lambda t} v_k(t)$. Тогда после преобразований, сокращая обе части на $e^{-\lambda t}$, получим

$$v'_0(t) = 0; \quad v'_k(t) = \lambda v_{k-1}(t) \quad (1.3)$$

при начальных условиях $v_0(0) = 1$; $v_k(0) = 0$; $k \geq 1$. Из (1.3) по индукции легко получить, что $v_k(t) = \frac{(\lambda t)^k}{k!}$, откуда

$$P_k(t) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}, \quad k \geq 0. \quad (1.4)$$

Приведем еще один способ решения системы (1.2) с помощью производящих функций, так как подобные рассуждения часто применяются при расчете различных характеристик СМО.

Обозначим $A(z, t) = \sum_{k=0}^{\infty} P_k(t) z^k$. Умножая обе части k -го равенства в (2.2) на z^k и суммируя по k , получим для $A(z, t)$ следующее дифференциальное уравнение:

$$A'_t(z, t) = -\lambda A(z, t) + \lambda z A(z, t)$$

с начальным условием $A(z, 0) = 1$. Отсюда

$$A(z, t) = e^{-\lambda t(1-z)}.$$

Раскладывая последнее выражение в ряд по степеням z , получим

$$A(z, t) = \sum_{k=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!} z^k.$$

Отсюда следует (1.4).

Дадим два важных определения. Пусть $\{t_k, k \geq 1\}$ — стационарный поток; $v(t)$ — количество его точек на интервале $[0, t]$, $P_k(t) = P\{v(t) = k\}$.

Определение 1.5. Параметром стационарного потока называется предел

$$\lim_{t \rightarrow 0} \frac{1 - P_0(t)}{t}.$$

Определение 1.6. Интенсивностью стационарного потока называется среднее количество его точек на интервале единичной длины.

Для простейшего потока параметр равен

$$\lim_{t \rightarrow 0} \frac{1 - P_0(t)}{t} = \lim_{t \rightarrow 0} \frac{1 - e^{-\lambda t}}{t} = \lambda$$

и интенсивность μ совпадает с ним:

$$\mu = Mv(1) = \sum_{k=0}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda.$$

Указанное совпадение не является случайным — оно имеет место для любого стационарного потока, что будет показано в следующем параграфе. В общем случае интенсивность и параметр связаны соотношением $\mu \geq \lambda$. В самом деле

$$\mu t = Mv(t) = \sum_{k=1}^{\infty} kP_k(t) \geq \sum_{k=1}^{\infty} P_k(t) = 1 - P_0(t) \\ \mu \geq \frac{1 - P_0(t)}{t}.$$

Переходя к пределу в последнем неравенстве, получим $\mu \geq \lambda$.

Замечание. Определение 5 основывается на доказанной А. Я. Хинчиным теореме о существовании предела $\lim_{t \rightarrow 0} \frac{1 - P_0(t)}{t}$.

Задания для самоконтроля

1. Пусть $\tau_1, \dots, \tau_k$ — независимые в совокупности показательно распределенные величины с параметрами $\lambda_1, \dots, \lambda_k$ соответственно. Доказать, что величина $\tau = \min\{\tau_1, \dots, \tau_k\}$ имеет показательное распределение с параметром $\lambda = \lambda_1 + \dots + \lambda_k$.

2. Пусть $t_1, t_2, \dots$ — последовательные моменты наступления событий в простейшем потоке, $A_n(t)$ — событие, заключающееся в том, что на отрезке $[0, t]$ произошло в точности n событий. Доказать, что распределение моментов $t_1, \dots, t_n$ при условии осуществления события $A_n(t)$ будет равномерным на отрезке $[0, t]$ (т. е. знание количества пуассоновских событий, происшедших на некотором интервале, в сущности, не несет никакой информации о расположении моментов наступления относительно друг друга).

3. Для любого t определим величину недоскока τ_t^- пуассоновского процесса до уровня t как $\tau_t^- = t - t_{v(t)}$ (время между моментом t и последним предшествующим событием или 0). Найти распределение τ_t^- и показать, что оно стремится к показательному при $t \rightarrow \infty$.

4. Доказать, что поток, образованный наложением конечного числа независимых простейших потоков, также будет простейшим. Найти его параметр.

5. Пусть каждая точка простейшего потока независимо от других точек с вероятностью p_i относится к i -му подпотоку. Доказать, что подпотоки являются независимыми простейшими потоками. Найти их параметры.

6. Пусть $\{\tau_i, i \geq 1\}$ — последовательность независимых, одинаково распределенных случайных величин, $P\{\tau_i > t\} = e^{-\lambda t}$. Пусть также N — случайная величина, не зависящая от $\{\tau_i, i \geq 1\}$, $P\{N = k\} = p^{k-1}q, k \geq 1$. Найти распределение суммы $\tau_1 + \tau_2 + \dots + \tau_N$ и указать на связь с предыдущей задачей.

7. Пусть τ_1 и τ_2 — независимые, одинаково показательно распределенные величины. Доказать, что случайные величины $\xi_1 = \tau_1 \wedge \tau_2$ и $\xi_2 = \tau_1 \vee \tau_2 - \tau_1 \wedge \tau_2$ независимы, найти их распределения и указать на связь с задачей 1.

8. Пусть $\{t_i, i \geq 1\}$ — простейший поток. Рассмотрим процедуру его прореживания с помощью интервалов I_k , при которой:

а) точки потока, попавшие внутрь интервала $I_k, k \geq 1$, выбрасываются;

б) длины ξ_k интервалов независимы, одинаково распределены и не зависят от исходного потока $\{t_i, i \geq 1\}$;

в) начало интервала I_1 совпадает с нулем;

г) если $t_r = \min \{t_k, k \geq 1: t_k \text{ лежит справа от } I_s\}$, то начало I_{s+1} совпадает с t_r .

Выяснить, возможна ли ситуация, при которой прореженный поток будет простейшим.

1.3. О сходимости входящих потоков к пуассоновскому

Входные потоки реальных систем массового обслуживания часто представляют собой суммы большого числа независимых потоков малой интенсивности. Так, поток вызовов, поступающий на телефонную станцию, представляет собой сумму потоков, исходящих от отдельных абонентов. Поток судов, прибывающих в какой-либо порт, является суммой потоков судов, отправляющихся из различных других портов. Поток требований на ремонт оборудования в ремонтную бригаду складывается из большого числа потоков, источниками которых являются каждый прибор, каждый отдельный механизм. Оказывается, что при весьма широких условиях относительно исходных потоков суммарные потоки будут близки к пуассоновским, в том числе и к простейшим.

Входящий поток будем задавать количеством требований $v(t)$, поступивших в систему на промежутке времени $[0, t]$.

Пусть $p_k(t) = P\{v(t) = k\}$; $\pi_k(t) = \sum_{i=k}^{\infty} p_i(t)$, $k = 0, 1, 2, \dots$ — вероятность того, что за время t поступит не менее k требований. Напомним, что параметром стационарного потока называется предел

$$\lambda = \lim_{t \rightarrow 0} \frac{\pi_1(t)}{t},$$

а математическое ожидание числа требований, поступающих за единицу времени, называется интенсивностью потока

$$\mu = Mv(1) = \sum_{k=0}^{\infty} k p_k(1).$$

Стационарный ординарный поток без последствия называется простейшим и, как было показано ранее, простей-

ший поток с параметром λ является пуассоновским потоком с параметром λ , т. е.

$$p_k(t) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}, \quad k = 0, 1, 2, \dots$$

Определение 1.7. Случайный процесс $v(t)$ называется пуассоновским процессом с ведущей функцией $\Lambda(t)$ ($\Lambda(t)$ — неотрицательная, непрерывная слева и $\Lambda(t) = 0$ при $t \leq 0$), если:

- а) $v(t)$ — процесс с независимыми приращениями;
- б) для всех $s < t$ и для любого неотрицательного целого k

$$P\{v(t) - v(s) = k\} = \frac{[\Lambda(t) - \Lambda(s)]^k}{k!} e^{-[\Lambda(t) - \Lambda(s)]}.$$

Легко видеть, что простейший (или пуассоновский) поток с параметром λ является частным случаем определенного выше пуассоновского процесса, ведущая функция которого $\Lambda(t) = \lambda t$. Приведем еще один пример входного пуассоновского потока.

Простейший нестационарный поток. Так назовем ординарные потоки без последствия. Обозначим через $p_k(t_0, t)$ вероятность того, что в интервале времени (t_0, t) поступит ровно k требований, и пусть

$$\pi_1(t_0, t) = \sum_{k=1}^{\infty} p_k(t_0, t) = 1 - p_0(t_0, t);$$

$$\pi_2(t_0, t) = \sum_{k=2}^{\infty} p_k(t_0, t) = 1 - p_0(t_0, t) - p_1(t_0, t).$$

Тогда в силу требования ординарности потока

$$\lim_{h \rightarrow 0} \frac{\pi_2(t, t+h)}{h} = 0 \quad (1.5)$$

для любого $t \geq 0$.

Предположим также, что для любого $t \geq 0$ существует предел

$$\lim_{h \rightarrow 0} \frac{\pi_1(t, t+h)}{h} = \lambda(t), \quad (1.6)$$

который назовем мгновенным значением параметра потока.

Аналогично изучению простейшего потока, пользуясь соотношениями (1.5—1.6), а также условием отсутствия

последствия, получим следующую систему уравнений для определения вероятностей $p_k(t_0, t)$:

$$\begin{aligned}\frac{\partial p_0(t_0, t)}{\partial t} &= -\lambda(t) p_0(t_0, t); \\ \frac{\partial p_k(t_0, t)}{\partial t} &= \lambda(t) [p_{k-1}(t_0, t) - p_k(t_0, t)], \quad k \geq 1.\end{aligned}\quad (1.7)$$

Обозначив через $A(t_0, t, z)$ производящую функцию последовательности $p_k(t_0, t)$:

$$A(t_0, t, z) = \sum_{k=0}^{\infty} p_k(t_0, t) z^k, \quad (1.8)$$

из уравнений (1.7)–(1.8) получаем

$$\frac{\partial A(t_0, t, z)}{\partial t} = (z-1) \lambda(t) A(t_0, t, z),$$

или

$$\frac{\partial \ln A(t_0, t, z)}{\partial t} = (z-1) \lambda(t).$$

Отсюда

$$\ln A(t_0, t, z) - \ln A(t_0, t_0, z) = (z-1) \int_{t_0}^t \lambda(u) du.$$

Так как $A(t_0, t_0, z) = p_0(t_0, t_0) = 1$, то

$$\ln A(t_0, t, z) = (z-1) \int_{t_0}^t \lambda(u) du.$$

Обозначим $\Lambda(t) = \int_0^t \lambda(u) du$, тогда

$$\begin{aligned}A(t_0, t, z) &= e^{(z-1)[\Lambda(t) - \Lambda(t_0)]} = \\ &= e^{-[\Lambda(t) - \Lambda(t_0)]} \sum_{k=0}^{\infty} z^k \frac{[\Lambda(t) - \Lambda(t_0)]^k}{k!}.\end{aligned}$$

Учитывая (1.8), получаем

$$\begin{aligned}P\{v(t) - v(t_0) = k\} &= p_k(t_0, t) = \\ &= e^{-[\Lambda(t) - \Lambda(t_0)]} \frac{[\Lambda(t) - \Lambda(t_0)]^k}{k!}, \quad k = 0, 1, 2, \dots\end{aligned}$$

и, таким образом, простейший нестационарный поток является процессом Пуассона в смысле определения 1.7,

ведущая функция которого $\Lambda(t) = \int_0^t \lambda(u) du$.

Сходимость суммарных потоков к пуассоновским. Будем говорить, что последовательность входящих потоков $v_n(t)$ сходится при $n \rightarrow \infty$ к $v(t)$, если совместное распределение вектора

$$\{v_n(t_1), v_n(t_2), \dots, v_n(t_k)\}$$

при любых k и $t_1, \dots, t_k$ слабо сходится к распределению вектора

$$\{v(t_1), v(t_2), \dots, v(t_k)\}.$$

Входящий поток $v_n(t)$ назовем суммарным потоком, если

$$v_n(t) = \sum_{r=1}^{k_n} v_{nr}(t),$$

где $v_{nr}(t)$, $r = \overline{1, k_n}$ — независимые между собой входящие потоки. Приведем теорему, доказанную Б. И. Григелинисом. Введем следующие обозначения:

$$p_{nr}(k, t, s) = P\{v_{nr}(t) - v_{nr}(s) = k\},$$

$$s < t, k = 0, 1, 2, \dots; \quad (1.9)$$

$$\Lambda_n(t, s) = \sum_{r=1}^{k_n} p_{nr}(1, t, s); \quad (1.10)$$

$$B_n(t, s) = \sum_{r=1}^{k_n} (1 - p_{nr}(0, t, s) - p_{nr}(1, t, s)). \quad (1.11)$$

Теорема 1.2. Если выполнены условия:

$$1. \lim_{n \rightarrow \infty} \max_{1 \leq r \leq k_n} [1 - p_{nr}(0, t, 0)] = 0;$$

$$2. \lim_{n \rightarrow \infty} \Lambda_n(t, s) = \Lambda(t) - \Lambda(s);$$

$$3. \lim_{n \rightarrow \infty} B_n(t, 0) = 0, \text{ то последовательность потоков}$$

$v_n(t) = \sum_{r=1}^{k_n} v_{nr}(t)$ сходится при $n \rightarrow \infty$ к пуассоновскому потоку $v(t)$ с ведущей функцией $\Lambda(t)$.

Доказательство. Пусть $0 \leq t_0 < t_1 < \dots < t_m$ — произвольные действительные числа; $l = (l_1, l_2, \dots, l_m)$, где $l_k \geq 0$ — целые числа;

$$\bar{0} = (0, 0, \dots, 0); \quad \bar{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_m);$$

$$\bar{l}_k = (0, \dots, 0, 1, 0, 0, \dots, 0);$$

$$\bar{T} = (t_0, t_1, \dots, t_m);$$

$$v_{nr}(\bar{T}) = (v_{nr}(t_1) - v_{nr}(t_0), \dots, v_{nr}(t_m) - v_{nr}(t_{m-1}));$$

$$v_n(T) = \sum_{r=1}^{k_n} v_{nr}(T); \quad (\bar{\alpha}, \bar{\beta}) = \sum_{k=1}^m \alpha_k \beta_k;$$

$$p_{nr}(\bar{l}, \bar{T}) = P\{v_{nr}(\bar{T}) = \bar{l}\};$$

$$f_{nr}(\bar{\alpha}, \bar{T}) = M \exp\{i(\bar{\alpha}, v_{nr}(\bar{T}))\};$$

$$f_n(\bar{\alpha}, \bar{T}) = M \exp\{i(\bar{\alpha}, v_n(\bar{T}))\}.$$

Для сходимости распределений векторов $\bar{v}_n(\bar{T})$ к соответствующим распределениям пуассоновского процесса достаточно сходимости их характеристических функций. В силу независимости процессов $\tau_{nr}(t)$ имеем

$$f_n(\bar{\alpha}, \bar{T}) = \prod_{r=1}^{k_n} f_{nr}(\bar{\alpha}, \bar{T}).$$

Но $t_{nr}(\bar{\alpha}, \bar{T}) = \sum_{\bar{l}} p_{nr}(\bar{l}, \bar{T}) e^{i(\bar{\alpha}, \bar{l})} = 1 + \sum_{\bar{l} \neq 0} p_{nr}(\bar{l}, \bar{T}) \times$
 $\times (e^{i(\bar{\alpha}, \bar{l})} - 1)$, (здесь символ $\sum_{\bar{l}}$ обозначает суммирование

по всем возможным целочисленным векторам $\bar{l}$ с неотрицательными компонентами).

$$\text{Оценим } \ln f_n(\bar{\alpha}, \bar{T}) = \sum_{r=1}^{k_n} \ln f_{nr}(\bar{\alpha}, \bar{T}).$$

Используя разложения для функции $\ln(1+x)$, нетрудно получить, что при $|x| \leq \frac{1}{2}$, $|\ln(1+x) - x| \leq x^2$. Тогда

$$|\ln f_{nr}(\bar{\alpha}, \bar{T}) - \sum_{\bar{l} \neq 0} p_{nr}(\bar{l}, \bar{T}) (e^{i(\bar{\alpha}, \bar{l})} - 1)| \leq \left(\sum_{\bar{l} \neq 0} p_{nr}(\bar{l}, \bar{T}) \right)^2.$$

Отсюда

$$\begin{aligned} & |\ln f_{nr}(\bar{\alpha}, \bar{T}) - \sum_{k=1}^m p_{nr}(\bar{l}_k, \bar{T}) (e^{i\alpha_k} - 1)| \leq \\ & \leq \sum_{\substack{\bar{l} \neq 0, \bar{l}_k \\ k=1, \dots, m}} p_{nr}(\bar{l}_k, \bar{T}) + \left(\sum_{\bar{l} \neq 0} p_{nr}(\bar{l}, \bar{T}) \right)^2, \end{aligned}$$

Далее,

$$\begin{aligned} \sum_{\bar{l} \neq 0} p_{nr}(\bar{l}, \bar{T}) &= 1 - P\{v_{nr}(t_m) - v_{nr}(t_0) = 0\} \leq \\ &\leq 1 - P\{v_{nr}(t_m) = 0\} = 1 - p_{nr}(0, t_m, 0); \\ \sum_{\substack{\bar{l} \neq 0, \bar{l}_k \\ k=1, 2, \dots, m}} p_{nr}(\bar{l}, \bar{T}) &= P\{v_{nr}(t_m) - v_{nr}(t_0) \geq 2\} \leq \\ &\leq P\{v_{nr}(t_m) \geq 2\}; \end{aligned}$$

$$\sum_{\bar{l} \neq 0} p_{nr}(\bar{l}, \bar{T}) \leq \sum_{k=1}^m p_{nr}(\bar{l}_k, \bar{T}) + P\{v_{nr}(t_m) \geq 2\}.$$

Заметим, что $p_{nr}(1, t_k, t_{k-1}) = p_{nr}(\bar{l}_k, \bar{T})$, поэтому

$$\begin{aligned} |\ln f_{nr}(\bar{\alpha}, \bar{T}) - \sum_{k=1}^m p_{nr}(1, t_k, t_{k-1})(e^{i\alpha_k} - 1)| &\leq \\ &\leq P\{v_{nr}(t_m) \geq 2\} + (1 - p_{nr}(0, t_m, 0)) \times \\ &\times \sum_{k=1}^m p_{nr}(1, t_k, t_{k-1}), \end{aligned}$$

или, учитывая 1.10—1.11,

$$\begin{aligned} |\ln f_n(\bar{\alpha}, \bar{T}) - \sum_{k=1}^m \Lambda_n(t_k, t_{k-1})(e^{i\alpha_k} - 1)| &\leq \\ &\leq B_n(t_m, 0) + \max_{1 \leq r \leq k_n} (1 - p_{nr}(0, t_m, 0)) \Lambda_n(t_m, 0). \quad (1.12) \end{aligned}$$

Учитывая условия теоремы и переходя к пределу в последнем соотношении, получим

$$\lim_{n \rightarrow \infty} f_n(\bar{\alpha}, \bar{T}) = \prod_{k=1}^m \exp\{[\Lambda(t_k) - \Lambda(t_{k-1})](e^{i\alpha_k} - 1)\},$$

что и требовалось доказать.

Сходимость суммарных потоков к простейшим. Рассмотрим случай, когда слагаемые потоки $v_{nr}(t)$ являются стационарными.

Докажем предварительно одну вспомогательную лемму.

Лемма 1.1. Если неотрицательная, неубывающая на отрезке $[0, a]$ функция $f(x)$ удовлетворяет соотношению

$$f(x+y) \leq f(x) + f(y) \quad (1.13)$$

для всех $x, y, x+y \in [0, a]$, то существует конечный или бесконечный предел $\lim_{x \rightarrow 0} \frac{f(x)}{x}$.

Доказательство.

Положим $q = \sup_{x > 0} \frac{f(x)}{x}$. Если $q < \infty$, то для любого $\varepsilon > 0$ существует x_0 такое, что $\frac{f(x_0)}{x_0} > q - \varepsilon$. Для любого $x < x_0$ представим x_0 в виде $x_0 = nx + \delta$ ($0 \leq \delta < x$). Тогда

$$q - \varepsilon \leq \frac{f(x_0)}{x_0} \leq \frac{nf(x) + f(\delta)}{x_0} = \frac{nx}{x_0} \frac{f(x)}{x} + \frac{f(\delta)}{x_0}.$$

Далее отметим, что при $x \rightarrow 0$ $\frac{nx}{x_0} \rightarrow 1$, $f(x) \rightarrow 0$, так как $q < \infty$. Поэтому

$$q - \varepsilon \leq \lim_{x \rightarrow 0} \frac{f(x)}{x} \leq \overline{\lim}_{x \rightarrow 0} \frac{f(x)}{x} \leq q,$$

откуда в силу произвольности ε ,

$$\lim_{x \rightarrow 0} \frac{f(x)}{x} = q.$$

Если $q = \infty$, то вместо $q - \varepsilon$ достаточно взять произвольно большое заданное число L .

Пусть теперь $\nu(t)$ — некоторый стационарный поток требований, а $p_k(t)$, как и раньше, вероятность того, что за время t поступит k требований, $\pi_1(t) = \sum_{k=1}^{\infty} p_k(t)$. Как видно из определения, функция $\pi_1(t)$ — неотрицательная, неубывающая и для любых $t_1 > 0, t_2 > 0$

$$\pi_1(t_1 + t_2) \leq \pi_1(t_1) + \pi_1(t_2).$$

Последнее соотношение следует из того, что если в промежутке $[0, t_1 + t_2)$ поступило хотя бы одно требование, то хотя бы одно требование должно было поступить в одном из промежутков $[0, t_1]$ или $[t_1, t_1 + t_2)$. Тогда в силу леммы 1.1. для стационарного потока $\nu(t)$ всегда существует конечный или бесконечный предел $\lim_{t \rightarrow 0} \frac{\pi_1(t)}{t} = \lambda$, который мы называли параметром потока.

Рассмотрим стационарный ординарный поток с конечным параметром λ . Важную роль при изучении таких процессов играют функции Пальма-Хинчина.

Обозначим через $H_k(\tau, t)$ вероятность следующего события: на промежутке времени $[t_0, t_0 + \tau)$ поступило по меньшей мере одно требование, а на промежутке $[t_0 + \tau, t_0 + \tau + t)$ поступит не более k требований ($t_0 > 0$ — произвольное число). Вероятность $H_k(\tau, t)$, как функция τ , неотрицательна и монотонно не убывает. Очевидно также, что

$$H_k(\tau_1 + \tau_2, t) \leq H_k(\tau_2, t) + H_k(\tau_1, \tau_2 + t),$$

но $H_k(\tau_1, \tau_2 + t) \leq H_k(\tau_1, t)$, поэтому условие (1.13) леммы 1.1 выполнено, и всегда существует предел (возможно и бесконечный)

$$\lim_{\tau \rightarrow 0} \frac{H_k(\tau, t)}{\tau}. \quad (1.14)$$

Но так как всегда $H_k(\tau, t) \leq \pi_1(\tau)$ и по предположению параметр процесса $\lambda = \lim_{\tau \rightarrow 0} \frac{\pi_1(\tau)}{\tau}$ конечный, то конечным будет и предел (1.14). Поэтому всегда существует предел

$$\Phi_k(t) = \lim_{\tau \rightarrow 0} \frac{H_k(\tau, t)}{\pi_1(\tau)},$$

который можно интерпретировать как условную вероятность поступления в промежутке t не более k требований при условии, что в начальный момент этого промежутка требование поступило. Обозначим $h_k(\tau, t) = H_k(\tau, t) - H_{k-1}(\tau, t)$, $k > 0$;

$$h_0(\tau, t) = H_0(\tau, t);$$

$$\varphi_k(t) = \lim_{\tau \rightarrow 0} \frac{h_k(\tau, t)}{\pi_1(t)} = \Phi_k(t) - \Phi_{k-1}(t), \quad k > 0;$$

$$\varphi_0(t) = \Phi_0(t).$$

Функции $\varphi_k(t)$, $k = 0, 1, 2, \dots$ называются функциями Пальма-Хинчина. Функцию $\varphi_k(t)$ можно толковать как вероятность поступления в промежутке $[0, t]$ ровно k требований при условии, что в начальный момент этого промежутка требование поступило. В силу ординарности потока имеем

$$\begin{aligned} p_k(\tau + t) &= P\{v(t_0 + \tau + t) - v(t_0) = k\} = \sum_{r=1}^k P\{v(t_0 + \tau) - \\ &- v(t_0) = r, v(t_0 + \tau + t) - v(t_0 + \tau) = k - r\} = P\{v(t_0 + \\ &+ \tau) - v(t_0) = 0, v(t_0 + \tau + t) - v(t_0 + \tau) = k\} + P\{v(t_0 + \\ &+ \tau) - v(t_0) = 1, v(t_0 + \tau + t) - v(t_0 + \tau) = \\ &= k - 1\} + 0(\tau). \end{aligned}$$

Но так как

$$\begin{aligned} P\{v(t_0 + \tau) - v(t_0) = 0; v(t_0 + \tau + t) - v(t_0 + \tau) = k\} = \\ P\{v(t) - v(0) = k\} - P\{v(\tau) - v(0) > 0; v(t + \tau) - \\ - v(\tau) = k\} = p_k(t) - h_k(\tau, t); \end{aligned}$$

$$\begin{aligned} P\{v(t_0 + \tau) - v(t_0) = 1; v(t_0 + \tau + t) - v(t_0 + \tau) = k - 1\} = \\ P\{v(\tau) - v(0) > 0; v(\tau + t) - v(\tau) = k - 1\} - \\ - P\{v(\tau) - v(0) > 1; v(\tau + t) - v(\tau) = k - 1\} = \\ = h_{k-1}(\tau, t) + 0(\tau), \end{aligned}$$

то

$$p_k(\tau + t) = p_k(t) - h_k(\tau, t) + h_{k-1}(\tau, t) + 0(\tau),$$

или

$$\frac{dp_k(t)}{dt} = \lambda [\varphi_{k-1}(t) - \psi_k(t)], \quad k > 0.$$

Аналогично находим

$$\frac{dp_0(t)}{dt} = -\lambda \varphi_0(t)$$

или, что то же,

$$\rho_0(t) = 1 - \lambda \int_0^t \varphi_0(z) dz; \quad (1.15)$$

$$\rho_k(t) = \lambda \int_0^t [\varphi_{k-1}(t) - \varphi_k(t)] dt. \quad (1.16)$$

Формулы (1.15), (1.16) носят название формул Пальма-Хинчина.

Пусть теперь $v_n(t) = \sum_{r=1}^k v_{nr}(t)$ — суммарный поток, отдельные слагаемые которого — стационарные ординарные потоки. Тогда существуют пределы

$$\lambda_{nr} = \lim_{t \rightarrow 0} \frac{P\{v_{nr}(t) - v_{nr}(0) > 0\}}{t},$$

которые называются параметрами слагаемых потоков.

Обозначим $\varphi_k^{nr}(t)$, $k = 0, 1, 2, \dots$ — функции Пальма — Хинчина потока $v_{nr}(t)$.

Теорема 1.3. Если выполнены условия:

- 1) $\lim_{n \rightarrow \infty} \max_{1 \leq r \leq k_n} \lambda_{nr} = 0$;
- 2) $\lim_{n \rightarrow \infty} \sum_{r=1}^{k_n} \lambda_{nr} \int_0^t \varphi_0^{nr}(u) du = \Lambda t$;
- 3) $\lim_{n \rightarrow \infty} \sum_{r=1}^{k_n} \lambda_{nr} \int_0^t \varphi_1^{nr}(u) du = 0$,

то суммарный поток $v_n(t) = \sum_{r=1}^{k_n} v_{nr}(t)$ сходится при $n \rightarrow \infty$ к простейшему потоку с параметром Λ .

Доказательство. Интенсивность потока по определению $\mu = Mv(1) = \sum_{k=1}^{\infty} k p_k(1)$. Очевидно, что для стационарного потока $\mu t = Mv(t) = \sum_{k=1}^{\infty} k p_k(t) \geq \sum_{k=1}^{\infty} p_k(t) = \pi_1(t)$.

Поэтому $\mu \geq \frac{\pi_1(t)}{t}$ для любого $t > 0$. А значит, $\mu \geq \lim_{t \rightarrow 0} \frac{\pi_1(t)}{t} = \lambda$ — интенсивность стационарного потока всегда не меньше его параметра. Для стационарных орди-

нарных потоков справедливы формулы Пальма — Хинчина (1.15), (1.16), следовательно,

$$\mu = \sum_{k=1}^{\infty} k p_k(1) = \sum_{k=1}^{\infty} \sum_{r=k}^{\infty} p_r(1) = \lambda \sum_{k=0}^{\infty} \int_0^1 \varphi_k(u) du. \quad (1.17)$$

Но $\varphi_k(u) = \lim_{\tau \rightarrow 0} \frac{h_k(\tau, u)}{\pi_1(\tau)}$ при любых $k \geq 0$ и $u > 0$,

$$\sum_{i=0}^k \frac{h_i(\tau, u)}{\pi_1(u)} = \frac{H_k(\tau, u)}{\pi_1(\tau)} \leq 1, \text{ следовательно, } \sum_{i=0}^k \varphi_i(u) \leq 1 \text{ при}$$

любом $k \geq 0$, или $\sum_{i=0}^k \int_0^1 \varphi_i(u) du \leq 1$.

Отсюда

$$\sum_{k=0}^{\infty} \int_0^1 \varphi_k(u) du \leq 1 \text{ и, значит, } \mu = \lambda \sum_{k=0}^{\infty} \int_0^1 \varphi_k(u) du \leq \lambda.$$

Учитывая $\mu \geq \lambda$, получим, что для стационарного ординарного потока всегда $\lambda = \mu$ — интенсивность совпадает с параметром. Поэтому в условиях нашей теоремы $\lambda_{nr} = Mv_{nr}(1)$, и если функции $\Lambda_n(t, s)$, $p_{nr}(k, t, s)$, $B_n(t, s)$ определены соотношениями (3.5)–(3.7), то

$$1 - p_{nr}(0, t, 0) \leq Mv_{nr}(t) = \lambda_{nr}t$$

и условие 1 теоремы 1.3 влечет выполнение условия 1 теоремы 1.2. Пользуясь формулами Пальма — Хинчина (1.15)–(1.16), имеем

$$p_{nr}(0, t, 0) = 1 - \lambda_{nr} \int_0^t \varphi_0^{nr}(u) du;$$

$$p_{nr}(k, t, 0) = \lambda_{nr} \int_0^t [\varphi_{k-1}^{nr}(u) - \varphi_k^{nr}(u)] du, \quad k \geq 1.$$

В силу стационарности процесса

$$\begin{aligned} \Lambda_n(t + s, s) &= \Lambda_n(t, 0) = \sum_{r=1}^{k_n} p_{nr}(1, t, 0) = \\ &= \sum_{r=1}^{k_n} \lambda_{nr} \int_0^t \varphi_0^{nr}(u) du - \sum_{r=1}^{k_n} \lambda_{nr} \int_0^t \varphi_1^{nr}(u) du; \\ B_n(t, 0) &= \sum_{r=1}^{k_n} \lambda_{nr} \int_0^t \varphi_1^{nr}(u) du. \end{aligned}$$

Поскольку выполнены условия 2—3, то

$$\lim_{n \rightarrow \infty} A_n(t, 0) = \lambda t; \quad \lim_{n \rightarrow \infty} B_n(t, 0) = 0,$$

и утверждение теоремы 1.3 следует из теоремы 1.2.

Замечание. В процессе доказательства теоремы 1.3 мы доказали следующее утверждение, представляющее и самостоятельный интерес (это утверждение впервые было доказано В. С. Королуком): параметр стационарного ординарного потока λ всегда совпадает с его интенсивностью μ . В общем же случае параметр стационарного потока не превосходит интенсивности: $\lambda \leq \mu$.

1.4. Элементы теории восстановления. Полумарковские процессы

Определение 1.8. Последовательность $\{t_i, i \geq 1\}$ называется потоком восстановления или рекуррентным потоком, если величины $\tau_1 = t_1$, $\tau_i = t_i - t_{i-1}$, $i \geq 2$ независимы в совокупности и одинаково распределены, причем $P\{\tau_i > 0\} = 1$.

Как видим, простейший поток является частным случаем потока восстановления, когда $P\{\tau_i < x\} = 1 - e^{-\lambda x}$, $x \geq 0$. Название «поток восстановления» связано с возможностью рассматривать последовательность $\{t_i\}$ как последовательность моментов времени, когда производятся замены вышедших из строя приборов. При этом длительности τ_i безотказной работы приборов независимы и одинаково распределены.

Положим $t_0 = 0$ и будем в дальнейшем называть t_i i -м моментом восстановления, а τ_i — длительностями восстановления. Пусть $v(t)$ — количество моментов восстановления на интервале $[0, t)$, включая точку 0. Обозначим $H(t) = Mv(t)$, $F(t) = P\{\tau_i < t\}$. По определению легко видеть, что $P\{v(t) = n\} = P\{t_{n-1} < t \leq t_n\}$, откуда следует тождество

$$P\{v(t) \leq n\} = P\{t_n \geq t\}, \quad n \geq 1,$$

которое носит название тождества Колмогорова — Феллера. Функция $H(t)$ носит название функции восстановления. Как легко видеть,

$$P\{t_n < t\} = P\{\tau_1 + \dots + \tau_n < t\} = F^{n*}(t),$$

где $F(t) = P\{\tau_1 < t\}$.

При этом $F^{0*}(t) = 0$, если $t = 0$, и 1, если $t > 0$. Покажем, что $H(t) < \infty$.

Теорема 1.4. $H(t) = \sum_{n=0}^{\infty} F^{n*}(t) < \infty$.

Доказательство. Имеем

$$\begin{aligned} Mv(t) &= \sum_{n=1}^{\infty} nP\{v(t) = n\} = \sum_{n=1}^{\infty} nP\{t_{n-1} < t \leq t_n\} = \\ &= \sum_{n=1}^{\infty} n[P\{t_{n-1} < t\} - P\{t_n < t\}] = \sum_{n=1}^{\infty} n[F^{(n-1)*}(t) - \\ &\quad - F^{n*}(t)] = \lim_{m \rightarrow \infty} \sum_{n=1}^m n[F^{(n-1)*}(t) - F^{n*}(t)] = \\ &= \lim_{m \rightarrow \infty} \sum_{n=0}^{m-1} F^{n*}(t) - mF^{m*}(t). \end{aligned} \quad (1.18)$$

Далее, $\{t_n < t\} \subset \{\tau_i < t\}$ при $i \leq n$.

Поэтому

$$P\{t_n < t\} = F^{n*}(t) \leq P\{\tau_1 < t, \dots, \tau_n < t\} = F^n(t).$$

Покажем, что для каждого фиксированного t существует целое r такое, что $F^{r*}(t) < 1$. Для этого выберем $a > 0$, так, чтобы $P\{\tau_1 > a\} = p > 0$. Если $r = [t/a] + 1$, то $F^{r*}(t) = P\{\tau_1 + \dots + \tau_r < t\} = 1 - P\{\tau_1 + \dots + \tau_r \geq t\} \leq 1 - P\{\tau_1 > a, \dots, \tau_r > a\} = 1 - p^r < 1$.

Так как $F^{n*}(t) \geq F^{(n+1)*}(t)$, то

$$\begin{aligned} \sum_{n=0}^{\infty} F^{n*}(t) &\leq r \sum_{i=1}^{\infty} F^{ri*}(t) \leq r \sum_{i=0}^{\infty} F^{ri}(t) = \\ &= r \frac{1}{1 - F^r(t)} < \infty. \end{aligned}$$

Кроме того,

$$mF^{m*}(t) \leq mF^{[m/r]*}(t) \leq m(F^{r*}(t))^{[m/r]} \xrightarrow{m \rightarrow \infty} 0.$$

Поэтому, переходя в равенстве (1.18) к пределу, получим

$$Mv(t) = \sum_{n=0}^{\infty} F^{n*}(t) < \infty.$$

Укажем способ вычисления других характеристик $v(t)$. Обозначим

$$\begin{aligned} A(t) &= MS^{v(t)} = \sum_{n=1}^{\infty} s^n P\{v(t) = n\} = \sum_{n=1}^{\infty} s^n P\{t_{n-1} < \\ &< t \leq t_n\} = \sum_{n=1}^{\infty} s^n (F^{(n-1)*}(t) - F^{n*}(t)). \end{aligned} \quad (1.19)$$

Вычислим

$$A^*(\theta) = \int_0^{\infty} e^{-\theta t} A(t) dt.$$

Положим

$$\varphi(\theta) = \int_0^{\infty} e^{-\theta x} dF(x).$$

Тогда, интегрируя по частям, получим

$$\begin{aligned} \int_0^{\infty} e^{-\theta x} F^{n*}(x) dx &= -\frac{e^{-\theta x}}{\theta} F^{n*}(x) \Big|_0^{\infty} + \\ &+ \frac{1}{\theta} \int_0^{\infty} e^{-\theta x} dF^{n*}(x) = \frac{1}{\theta} \varphi^n(\theta), \end{aligned}$$

поскольку $F^{n*}(0) = 0$, а $\int_0^{\infty} e^{-\theta x} dF^{n*}(x) = M \exp \left\{ -\theta \sum_{k=1}^n \tau_k \right\} = \varphi^n(\theta)$.

Поэтому из (1.19)

$$A^*(\theta) = \frac{1}{\theta} \sum_{n=1}^{\infty} s^n (\varphi^{n-1}(\theta) - \varphi^n(\theta)) = \frac{s}{\theta} \frac{1 - \varphi(\theta)}{1 - s\varphi(\theta)}.$$

Используя тождество Колмогорова — Феллера, которое позволяет выражать распределения $v(t)$ через распределения сумм независимых случайных величин, нетрудно исследовать предельное распределение $v(t)$ при соответствующей нормировке, когда $t \rightarrow \infty$.

Теорема 1.5. Если $a = M\tau_i < \infty$ и

$$P \left\{ \frac{1}{n^{1/\alpha}} \left(\sum_{k=1}^n \tau_k - na \right) < x \right\} \xrightarrow[n \rightarrow \infty]{\text{сл}} F_{\alpha}(x), \quad -\infty < x < +\infty,$$

где $F_{\alpha}(x)$ — функция распределения некоторого устойчивого закона с параметром α , $1 < \alpha \leq 2$, то

$$\begin{aligned} P \left\{ \frac{1}{t^{1/\alpha}} \left(v(t) - \frac{t}{a} \right) < x \right\} &\xrightarrow[t \rightarrow \infty]{\text{сл}} 1 - F_{\alpha} \left(-x a^{\frac{\alpha+1}{\alpha}} \right), \\ &-\infty < x < +\infty \end{aligned}$$

(здесь символ $F_n(x) \xrightarrow[n \rightarrow \infty]{\text{сл}} F(x)$ обозначает слабую сходимость функций распределения, т. е. сходимость в точках непрерывности правой части).

Доказательство.

$$\begin{aligned} P\left\{\frac{1}{t^{1/\alpha}}\left(v(t) - \frac{t}{a}\right) < x\right\} &= P\left\{v(t) < \frac{t}{a} + xt^{1/\alpha}\right\} = \\ &= P\left\{\sum_{k=1}^{\lfloor t/a + xt^{1/\alpha} \rfloor} \tau_k \geq t\right\} = P\left\{\frac{1}{n_t^{1/\alpha}}\left(\sum_{k=1}^{n_t} \tau_k - \right. \right. \\ &\quad \left. \left. - n_t a\right) \geq \frac{1}{n_t^{1/\alpha}}(t - n_t a)\right\}, \end{aligned}$$

где $n_t = \lfloor t/a + xt^{1/\alpha} \rfloor$. Так как $n_t \rightarrow \infty$ при $t \rightarrow \infty$ и $(t - n_t a)/n_t^{1/\alpha} \rightarrow x a^{\frac{\alpha+1}{\alpha}}$, утверждение теоремы следует из свойства: если $F_n(x) \xrightarrow[n \rightarrow \infty]{\text{с.л.}} F(x)$ и последовательность x_n такова, что $x_n \rightarrow x$ и x — точка непрерывности $F(x)$, то $F_n(x_n) \rightarrow F(x)$. Действительно, так как при достаточно больших n $x - \varepsilon < x_n < x + \varepsilon$, то для любого $\varepsilon > 0$

$$F(x - \varepsilon) \leq \lim_{n \rightarrow \infty} F_n(x_n) \leq \overline{\lim}_{n \rightarrow \infty} F_n(x_n) \leq F(x + \varepsilon)$$

Поскольку $F(x + \varepsilon) - F(x - \varepsilon) \rightarrow 0$ при $\varepsilon \rightarrow 0$, то теорема доказана.

Отсюда легко получить

Следствие 1.1. Если $a < \infty$, то

$$\frac{v(t)}{t} \xrightarrow[t \rightarrow \infty]{P} \frac{1}{a}.$$

В случае, когда $0 < \alpha < 1$, т. е. $a = \infty$, тем же методом доказывается

Теорема 1.6. Если для некоторого α , $0 < \alpha < 1$

$$P\left\{\frac{1}{n^{1/\alpha}} \sum_{k=1}^n \tau_k < x\right\} \xrightarrow[n \rightarrow \infty]{\text{с.л.}} F_\alpha(x), \quad x > 0,$$

то

$$P\{t^{-\alpha} v(t) < x\} \xrightarrow[t \rightarrow \infty]{\text{с.л.}} 1 - F_\alpha\left(\frac{1}{x^{1/\alpha}}\right), \quad x > 0.$$

Доказательство этой теоремы предоставим читателю.

При исследовании аналитическими методами различных функционалов от процессов восстановления часто возникают уравнения вида

$$Z(t) = z(t) + \int_0^t Z(t-x) dF(x),$$

где $z(t)$ — некоторая известная функция, а $F(x)$ — функция распределения. Уравнение такого типа называется уравнением восстановления.

Лемма 1.2. Если $z(x)$ — ограниченная функция, $z(x) = 0$ при $x < 0$, то уравнение восстановления имеет решение

$$Z(t) = \int_0^t z(t-s) dH(s).$$

Это решение единственно в классе ограниченных на каждом конечном интервале функций, обращающихся в нуль на отрицательной полуоси.

Доказательство. По теореме 1.4

$$H(t) = \sum_{n=0}^{\infty} F^{n*}(t).$$

Свернув обе части последнего равенства с функцией $z(t)$, получим

$$\begin{aligned} H \times z &= \left(\sum_{n=0}^{\infty} F^{n*} \right) \times z = F^{0*} \times z + \sum_{n=1}^{\infty} (F^{n*} \times z) = \\ &= z + F \times \sum_{n=1}^{\infty} F^{(n-1)*} \times z = z + F \times (H \times z). \end{aligned}$$

Таким образом, функция $Z = H \times z$ удовлетворяет уравнению восстановления.

Докажем теперь единственность решения в указанном классе функций. Пусть Z_1 и Z_2 — два ограниченных на каждом конечном интервале решения, $Z_i(x) = 0$ при $x < 0$. Тогда разность $R = Z_1 - Z_2$ удовлетворяет уравнению

$$R = R \times F \quad (1.20)$$

и также ограничена на каждом конечном интервале и $R(x) = 0$ при $x < 0$. Из равенства (1.20) немедленно следует, что

$$R = (R \times F) \times F = R \times F^{2*} = \dots = R \times F^{n*}. \quad (1.21)$$

Покажем, что

$$F^{n*}(x) = P\{\tau_1 + \dots + \tau_n < x\} \xrightarrow{n \rightarrow \infty} 0.$$

$$Me^{-(\tau_1 + \dots + \tau_n)} = (Me^{-\tau_1})^n = \alpha^n \xrightarrow{n \rightarrow \infty} 0 \quad (\alpha < 1).$$

Применяя неравенство Чебышева, получим

$$\begin{aligned} P\{\tau_1 + \dots + \tau_n < x\} &= P\{e^{-(\tau_1 + \dots + \tau_n)} > e^{-x}\} \leq \\ &\leq \frac{Me^{-(\tau_1 + \dots + \tau_n)}}{e^{-x}} \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

Переходя в равенстве (1.21) к пределу при $n \rightarrow \infty$, получим $R = R \times 0 = 0$, следовательно, $Z_1 = Z_2$.

Уравнением восстановления описываются многие задачи. Рассмотрим, например, вопрос нахождения распределения времени ξ появления первого интервала восстановления $[t_{k-1}, t_k]$, длина τ_k которого больше s (если интерпретировать величины t_i как моменты появления автомобилей на перекрестках, а s — как время, необходимое для перехода перекрестка, то величина ξ будет временем ожидания возможности перехода перекрестка).

Обозначая через $Z(t) = P\{\xi < t\}$ функцию распределения ξ , получим по формуле полной вероятности, что $Z(t) = 0$ при $t < s$, и

$$\begin{aligned} Z(t) &= P\{\tau_1 \geq s\} + \int_0^s Z(t-u) P\{\tau_1 \in du\} = \\ &= 1 - F(s) + \int_0^s Z(t-u) dF(u), \quad t \geq s. \end{aligned}$$

Отсюда следует, что функция $Z(t)$ удовлетворяет уравнению восстановления $Z = z + Z \times G$, где $z(t) = \begin{cases} 0, & t < s; \\ 1 - F(s), & t \geq s, \end{cases}$

$$G(t) = \begin{cases} F(t), & t \leq s; \\ F(s), & t > s. \end{cases}$$

В соответствии с леммой 1.2 $Z(t) = \int_0^t z(t-s) dH(s)$,

где функция восстановления $H(t)$ отвечает функции распределения $G(t)$ (вообще говоря, несобственной). Как видим, функция восстановления $H(t)$ играет важную роль в задачах, связанных с потоками восстановления. Для случая, когда точки t_i образуют простейший поток с параметром λ , функция $H(t)$ является линейной. Оказывается, что линейный рост функции восстановления характерен и для других потоков восстановления.

Теорема 1.7. (теорема восстановления). Если распределение времени восстановления неарифметическое, то

$$H(t) - H(t-h) \xrightarrow{t \rightarrow \infty} \frac{h}{M\tau_1}$$

для любого фиксированного h .

Если распределение τ_1 — арифметическое, то $H(t) = H(t-h) \rightarrow \frac{h}{M\tau_1}$ для значений h , кратных шагу распределения.

(Напомним, что распределение величины τ называется арифметическим с шагом h , если возможные значения τ кратны h). Из теоремы восстановления легко следует, что

$$\frac{H(t)}{t} \xrightarrow[t \rightarrow \infty]{} \frac{1}{M\tau_1}.$$

Приведем другое важное следствие теоремы восстановления.

Следствие 1.2. Если функция $z(t)$ равна нулю вне некоторого интервала и непрерывна, то решение $Z(t)$ уравнения восстановления $Z = z + Z \times F$ для неарифметической функции распределения удовлетворяет соотношению

$$Z(t) \xrightarrow[t \rightarrow \infty]{} \frac{1}{\mu} \int_0^\infty z(u) du$$

Замечание. Утверждение следствия 1.2 справедливо в более общей ситуации для непосредственно интегрируемых по Риману функций $z(t)$ (функция $z(t)$ называется непосредственно интегрируемой по Риману, если верхние и нижние интегральные суммы S и s функции $z(t)$, построенные по разбиениям всей полуоси $[0, \infty)$ существуют, и $S - s \rightarrow 0$ при измельчении этих разбиений). Частным случаем функций, непосредственно интегрируемых по Риману, являются монотонные интегрируемые по Риману на $[0, \infty)$ функции. В качестве применения теоремы восстановления найдем предельное при $t \rightarrow \infty$ распределение длины интервала от случайного момента времени t до следующей точки потока восстановления

$$\xi(t) = t_{v(t)+1} - t.$$

Случайная величина $\xi(t)$ носит название оставшегося времени ожидания, или перескока уровня t . Обозначим

$$p(t, x) = P\{\xi(t) < x\}.$$

Теорема 1.8. Если распределение времени восстановления неарифметическое и $M\tau_1 = \mu < \infty$, то

$$p(t, x) \xrightarrow[t \rightarrow \infty]{} \frac{1}{\mu} \int_0^x (1 - F(u)) du.$$

Доказательство. По формуле полной вероятности

$$P\{\xi(t) < x\} = \sum_{i=0}^t P\{t_i \in du, t - u \leq \tau_{i+1} < t - u + x\} =$$

$$\begin{aligned}
&= \int_0^t \sum_{i=0}^{\infty} P\{t_i \in du\} (F(t-u+x) - F(t-u)) = \\
&= \int_0^t \sum_{i=0}^{\infty} F^{i*}(du) (F(t-u+x) - F(t-u)) = \\
&= \int_0^t (F(t-u+x) - F(t-u)) dH(u). \quad (1.22)
\end{aligned}$$

Из соотношения (1.22) и леммы 1.2 следует, что $p(t, x)$ является решением уравнения восстановления

$$p = z + p \times F,$$

где

$$z(t) = F(t+x) - F(t).$$

Далее, функция $z(t)$ непосредственно интегрируема, так как может быть представлена в виде разности

$$z(t) = (1 - F(t)) - (1 - F(t+x))$$

двух непосредственно интегрируемых функций $1 - F(t)$ и $1 - F(t+x)$ (непосредственная интегрируемость последних функций вытекает из их монотонности и существования интеграла

$$\int_0^{\infty} (1 - F(u)) du = \int_0^{\infty} x dF(x) = \mu < \infty,$$

Используя следствие 1.2 из теоремы восстановления, получим

$$\begin{aligned}
p(t, x) &\xrightarrow[t \rightarrow \infty]{\frac{1}{\mu}} \int_0^{\infty} z(u) du = \frac{1}{\mu} \int_0^{\infty} (1 - F(u)) du - \\
&- \int_0^{\infty} (1 - F(u+x)) du = \frac{1}{\mu} \int_0^x (1 - F(u)) du.
\end{aligned}$$

Определение 1.9. Последовательность $\{t_i, i \geq 1\}$ называется потоком восстановления с запаздыванием, если величины $\tau_1 = t_1$, $\tau_i = t_i - t_{i-1}$, $i \geq 2$ независимы в совокупности, а τ_i при $i \geq 2$ одинаково распределены.

Важным примером потока восстановления с запаздыванием является последовательность $t_i = s_{i+\mu(t)} - t$, $i \geq 1$, где $\{s_i, i \geq 1\}$ — поток восстановления, а $\mu(t)$ — количество его моментов восстановления на интервале $[0, t)$.

Обозначим

$F_1(t) = P\{\tau_1 < t\}$, $F(t) = P\{\tau_i < t\}$, $i \geq 2$, $v_1(t)$ — количество точек t_i на интервале $[0, t)$. Пусть $H_1(t) = Mv_1(t)$ — функция восстановления процесса с запаздыванием, а $H(t)$ — функция восстановления для потока с длительностью восстановления, имеющего функцию распределения $F(t)$. Так же, как и при доказательстве леммы 1.2, можно проверить, что

$$H_1(t) < \infty.$$

Лемма 1.3. $H_1(t) = 1 + (F_1 \times H)(t) < \infty$.

Доказательство.

$$\begin{aligned} H_1(t) &= Mv_1(t) = \sum_{n=1}^{\infty} nP\{v_1(t) = n\} = \\ &= \sum_{n=1}^{\infty} nP\{t_{n-1} < t \leq t_n\} = P\{t \leq t_1\} + \\ &+ \sum_{n=2}^{\infty} n[P\{t_{n-1} < t\} - P\{t_n < t\}] = \\ &= 1 - F_1(t) + \sum_{n=2}^{\infty} n[(F_1 \times F^{(n-2)*})(t) - (F_1 \times F^{(n-1)*})(t)] = \\ &= 1 - F_1(t) + F_1 \times (2F^{0*} + F + F^{2*} + \dots + \\ &+ F^{n*} + \dots)(t) = 1 + (F_1 \times \sum_{n=0}^{\infty} F^{n*})(t) = \\ &= 1 + (F_1 \times H)(t) < \infty. \end{aligned}$$

Для потоков восстановления с запаздыванием справедливы предельные теоремы, указанные для обычных потоков восстановления. Так, если распределение времени восстановления τ_2 неарифметическое и $\mu = M\tau_2 < \infty$, то

$$H_1(t) - H_1(t-h) \xrightarrow{t \rightarrow \infty} \mu^{-1}h,$$

а распределение оставшегося времени ожидания $\xi(t)$ в

пределе при $t \rightarrow \infty$ имеет вид $\lim_{t \rightarrow \infty} P\{\xi(t) < x\} = \frac{1}{\mu} \int_0^x (1 -$

$-F(u)) du$. Необходимость изучения потоков восстановления $\{t_i, i \geq 1\}$ с запаздыванием связана с тем, что, как правило, стационарные потоки с независимыми длительностями $\tau_i = t_i - t_{i-1}$ интервалов между точками не являются потоками восстановления. Более того, оказывается,

что в стационарном случае функции распределения $F_1(t)$ и $F(t)$ потока восстановления с запаздыванием не могут быть выбраны произвольно.

Теорема 1.9. Если поток восстановления с запаздыванием $\{t_i, i \geq 1\}$ является стационарным, $\mu = \int_0^\infty x dF(x) < \infty$, то

$$F_1(t) = \frac{1}{\mu} \int_0^t (1 - F(u)) du.$$

Доказательство. Стационарность потока $\{t_i\}$ влечет за собой линейность функции восстановления. В самом деле, пусть $0 \leq s_1 < s_2 < s_3$ — три неслучайных момента времени. В силу стационарности потока, распределение величины $v_1(s+t) - v_1(s)$ совпадает с распределением величины $v_1(t) - v_1(0)$.

Поэтому

$$\begin{aligned} H_1(s_3) - H_1(s_1) &= H_1(s_3) - H_1(s_2) + \\ &+ H_1(s_2) - H_1(s_1) = M(v_1(s_3) - v_1(s_2)) + \\ &+ M(v_1(s_2) - v_1(s_1)) = M(v_1(s_3 - s_2) - v_1(0)) + \\ &+ M(v(s_2 - s_1) - v_1(0)) = H_1(s_3 - s_2) + H_1(s_2 - s_1) - 2. \end{aligned}$$

Поскольку $H_1(t)$ не убывает, то из полученного функционального уравнения следует линейность $H_1(t)$ при $t > 0$. Следовательно, $H_1(t)$ при $t > 0$ имеет вид $H_1(t) = at + 1$.

По лемме 1.3.

$$\begin{aligned} H_1 &= 1 + F_1 \times H = 1 + F_1 + F_1 \times \sum_{n=1}^{\infty} F^{n*} = \\ &= 1 + F_1 + F \times (F_1 \times \sum_{n=0}^{\infty} F^{n*}) = 1 + F_1 + \\ &+ F \times (F_1 \times H) = 1 + F_1 + F \times H_1 - F. \end{aligned}$$

Если $H_1(t) = 1 + at$, то из последнего равенства вытекает, что

$$at = F_1(t) + \int_0^t a(t-s) dF(s).$$

Интегрируя по частям, получим

$$\int_0^t (t-s) dF(s) = (t-s)F(s) \Big|_0^t + \int_0^t F(s) ds = \int_0^t F(s) ds.$$

Отсюда

$$F_1(t) = at - a \int_0^t F(u) du = a \int_0^t (1 - F(u)) du.$$

Так как

$$\lim_{t \rightarrow \infty} F_1(t) = 1,$$

то

$$a = \left(\int_0^{\infty} (1 - F(u)) du \right)^{-1} = \mu^{-1}$$

и

$$F_1(t) = \frac{1}{\mu} \int_0^t (1 - F(u)) du.$$

Теорема доказана.

Замечание. Условие $\mu < \infty$ является необходимым для стационарного потока $\{t_i\}$, $i \geq 1$.

Замечание. Полученное распределение $F_1(t)$ совпадает с предельным для величины перескока, что согласуется со смыслом стационарности. В самом деле, стационарный поток можно рассматривать как поток, находящийся в установившемся режиме, т. е. начавшийся в бесконечно удаленной точке. Поэтому распределение начального интервала $F_1(t)$ должно совпадать с распределением величины перескока точки 0 процессом восстановления, начавшемся в точке, удаленной на $-\infty$, т. е., что то же самое, с предельным распределением величины перескока обычным процессом восстановления.

Замечание. Для потока Пуассона $F_1(t) = F(t)$, и поэтому для его стационарности необходимо, чтобы выполнялось равенство

$$F(t) = \frac{1}{\mu} \int_0^t (1 - F(u)) du$$

или

$$F'(t) = \frac{1}{\mu} (1 - F(t)).$$

Решая это дифференциальное уравнение с начальным условием $F(0) = 0$, получим $F(t) = 1 - e^{-t/\mu}$, и такой поток является простейшим.

Определение 1.10. Процессом восстановления называется процесс

$$\mu(t) = \max\{n \geq 0 : t_n \leq t\},$$

где $\{t_n\}$, $n \geq 1$ — поток восстановления; $t_0 = 0$.

Из определения непосредственно вытекает, что $\mu(t)$ — ступенчатый процесс, скачки которого равны единице, а

длительности τ_i интервалов между скачками независимы и одинаково распределены. Если последовательность $\{t_i\}$, $i \geq 1$ образует простейший поток, то $\mu(t)$ — пуассоновский процесс. Если же распределение τ_i отлично от показательного, то $\mu(t)$ не будет даже цепью Маркова.

Действительно, при исследовании простейшего потока, только для показательного распределенных промежутков имеет место отсутствие последствия, т. е. величина перескока $t_{\mu(t)+1} - t$ не зависит от прошлого и имеет вновь показательное распределение. В противном случае распределение перескока зависит от величины недоскока $t - t_{\mu(t)}$. Следовательно, для того чтобы вычислить распределение следующего момента скачка в общем случае помимо значения $\mu(t)$ нужно знать еще величину $t - t_{\mu(t)}$.

Вместе с тем значения $\mu(t_n) = \mu_n$ процесса $\mu(t)$ в моменты t_n его переходов образуют простейшую цепь Маркова с дискретным временем, начальное состояние которой $\mu_0 = 0$, множество состояний совпадает с множеством натуральных чисел, а вероятности p_{ij} перехода за один шаг имеют вид $p_{ij} = \delta_{i,j-1}$, $\delta_{ij} = 1$ при $i = j$ и 0 при $i \neq j$. Кроме того, времена пребывания процесса $\mu(t)$ в своих состояниях F независимы и имеют общую функцию распределения $F(t)$. Следовательно, процесс $\mu(t)$ определяется цепью Маркова μ_n , которая носит название вложенной цепи Маркова, и функцией распределения $F(t)$ независимых в совокупности времен пребывания процесса $\mu(t)$ в своих состояниях.

Такой способ задания случайных процессов будем называть конструктивным. Он может быть использован для описания однородной цепи Маркова с непрерывным временем, когда эта цепь $\xi(t)$ задается с помощью характеристик своей вложенной цепи Маркова $\xi_n = \xi(t_n)$ (здесь t_n — момент n -го перехода цепи $\xi(t)$) и функций распределения $F_j(t)$ времен пребывания цепи $\xi(t)$ в состоянии j . При этом $F_j(t) = 1 - e^{-q_j t}$, $t \geq 0$, а времена пребывания в последовательно сменяющихся друг друга состояниях условно независимы при условии, что значения состояний определены:

$$\begin{aligned} P\{t_1 - t_0 < x_0, t_2 - t_1 < x_1, \dots, t_k - t_{k-1} < \\ < x_{k-1} / \xi(0) = i_0, \xi(t_1) = i_1, \dots, \xi(t_{k-1}) = \\ = i_{k-1}\} = \prod_{j=0}^{k-1} (1 - e^{-q_{ij} x_i}), \end{aligned}$$

Естественным обобщением как процессов восстановления, так и однородных цепей Маркова с непрерывным или дис-

кретным временем являются полумарковские процессы (ПМП).

Конструктивно ПМП на множестве состояний $\{1, 2, \dots\}$ задается с помощью

а) дискретного распределения $\{p_i, i \geq 1\}$;

б) вероятностей $p_{ij}, i, j \geq 1$, таких, что $\sum_{j=1}^{\infty} p_{ij} = 1$;

в) набора функций распределения $F_{ij}(t); i, j \geq 1$, таких, что $F_{ij}(+0) = 0$.

Характеристики, указанные в пунктах а) и б), задают вложенную цепь Маркова ξ_n , а $F_{ij}(t)$ — времена пребывания в состояниях. Вложенная цепь Маркова задается начальным распределением $\{p_i, i \geq 1\}$, матрицей вероятностей перехода за один шаг $P = \|p_{ij}\|_{i,j \geq 1}$, а ее значение ξ_n равно значению $\xi(t_n)$ ПМП $\xi(t)$ в момент t_n его перехода (для определенности будем считать, что траектории процесса $\xi(t)$ непрерывны справа с вероятностью единица). Если состояния вложенной цепи Маркова известны, $\xi_0 = i_0, \xi_1 = i_1, \dots, \xi_n = i_n, \dots$, то длительности $t_k - t_{k-1}, k \geq 1$ интервалов между моментами переходов процесса $\xi(t)$ условно независимы и имеют условные функции распределения $F_{i_{k-1}i_k}(t)$:

$$P\{t_1 - t_0 < x_0, t_2 - t_1 < x_1, \dots, t_n - t_{n-1} < x_{n-1} | \xi(0) = i_0, \dots, \xi(t_n) = i_n\} = \prod_{k=1}^n F_{i_{k-1}i_k}(x_{k-1}).$$

Укажем, как выражаются основные вероятностные характеристики ПМП через параметры вложенной цепи Маркова и условные функции распределения $F_{ij}(t)$.

Обозначим

$$Q_{ij}(t) = P\{\xi_n = j, t_n - t_{n-1} < t | \xi_{n-1} = i\};$$

$$Q(t) = \|Q_{ij}(t)\|; F_i(t) = P\{t_n - t_{n-1} < t | \xi_{n-1} = i\},$$

где $F_i(t)$ — распределение времени пребывания процесса в состоянии i . Тогда, используя определение условной вероятности, находим

$$Q_{ij}(t) = p_{ij}F_{ij}(t).$$

Отсюда следует, что

$$F_i(t) = \sum_j Q_{ij}(t) = \sum_j p_{ij}F_{ij}(t).$$

Нахождение переходных вероятностей $p_{ij}(t) = P\{\xi(t) = j | \xi(0) = i\}$ можно осуществлять с помощью такой теоремы.

Теорема 1.10. Функции $p_{ij}(t)$ удовлетворяют системе линейных интегральных уравнений

$$p_{ij}(t) = \delta_{ij}(1 - F_i(t)) + \sum_k \int_0^t Q_{ik}(ds) p_{kj}(t-s).$$

Доказательство. Если $\xi(0) = i$, то событие $\{\xi(t) = j\}$ может наступить двумя различными способами:

$$t_1 \geq t, j = i;$$

$$t_1 < t, \xi(t_1) = k, \xi(t) = j,$$

где k — некоторое состояние полумарковского процесса.

Вероятность первого способа равна $\delta_{ij}(1 - F_i(t))$, а второго —

$$\begin{aligned} & \sum_k \int_0^t P\{t_1 \in ds, \xi(s) = k, \xi(t) = j / \xi(0) = i\} = \\ &= \sum_k \int_0^t p_{ik} F_{ik}(ds) P\{\xi(t-s) = j / \xi(0) = k\} = \\ &= \sum_k \int_0^t Q_{ik}(ds) p_{kj}(t-s), \end{aligned}$$

что доказывает теорему.

Обозначив $\Pi(t) = \|p_{ij}(t)\|$, $\Phi(t) = \|\delta_{ij}(1 - F_i(t))\|$, $Q(t) = \|Q_{kj}(t)\|$, получим, что в матричном виде система интегральных уравнений для переходных вероятностей $p_{ij}(t)$ запишется как

$$\Pi(t) = \Phi(t) + \int_0^t Q(ds) \Pi(t-s).$$

Полученное уравнение в свертках называется многомерным уравнением восстановления (а также уравнением марковского восстановления) и может быть исследовано с помощью преобразований Лапласа.

Одним из его следствий служит теорема о предельном значении переходных вероятностей $p_{ij}(t)$ при $t \rightarrow \infty$.

Теорема 1.11. Если выполняются условия

$$\text{а) для всех } i \quad m_i = \int_0^\infty (1 - F_i(t)) dt < \infty;$$

б) распределение времени возвращения в некоторое состояние k не арифметическое;

в) вложенная цепь Маркова имеет стационарное распределение $\{\pi_i\}$;

г) $\sum_i \pi_i m_i < \infty$,

то

$$\lim_{t \rightarrow \infty} p_{ij}(t) = \pi_j m_j \left(\sum_i \pi_i m_i \right)^{-1}.$$

Доказательство. Рассмотрим процесс с двумя состояниями $\{1, 2\}$, который в состоянии $\{1\}$ сидит случайное время τ_1 , затем переходит в состояние $\{2\}$, в котором пребывает время τ_2 , после чего возвращается в $\{1\}$. Будем считать, что моменты возвращения в $\{1\}$ являются моментами восстановления, величины τ_1 и τ_2 , вообще говоря, могут быть зависимыми. Положим

$$F_1(t) = P\{\tau_1 < t\}, \quad F(t) = P\{\tau_1 + \tau_2 < t\}.$$

Тогда аналогично выводу многомерного уравнения получим, что величины $p_{11}(t)$ удовлетворяют уравнению восстановления вида

$$p_{11}(t) = 1 - F_1(t) + \int_0^t p_{11}(t-u) dF(u).$$

Если $F(u)$ — неарифметическое распределение и $\mu = \int_0^\infty x dF(x) < \infty$, то из теоремы восстановления следует, что

$$\lim_{t \rightarrow \infty} p_{11}(t) = \frac{1}{\mu} \int_0^\infty (1 - F_1(u)) du = \frac{M\tau_1}{\mu}.$$

Заметим, что ПМП можно представить в виде процесса с двумя состояниями $\{1, 2\}$, выбирая в качестве $\{1\}$ состояние $\{1\}$, а в качестве $\{2\}$ — подмножество $\{2, 3, \dots\}$. При этом величина $\tau = \tau_1 + \tau_2$ будет временем 1-го возвращения процесса в $\{1\}$. При наших предположениях величина τ — не арифметическая, при этом, если выполнены условия в) и г), то известно, что

$$M\tau = \frac{1}{\pi_1} \sum_i \pi_i m_i.$$

Тем самым получим, для ПМП

$$\lim_{t \rightarrow \infty} p_{11}(t) = \frac{\pi_1 m_1}{\sum_i \pi_i m_i} = \pi_1.$$

Затем покажем, что предел ρ_1 не зависит от начального состояния. Для любого состояния j обозначим через x_{j1} — время первого достижения $\{1\}$, начиная из состояния $\{j\}$ (т. е. $x_{j1} = \inf \{t: t > 0, \xi(t) = 1\}$, если $\xi(0) = j$). Ясно, что x_{j1} — собственная случайная величина. Пусть

$$\Phi_j(x) = P\{x_{j1} < x\}, \quad \text{тогда} \quad p_{11}(t) = \int_0^t p_{11}(t-u) d\Phi_j(u).$$

Отсюда получим для любого $L > 0, t > L$:

$$\begin{aligned} |p_{11}(t) - \rho_1| &= \left| \int_0^t (p_{11}(t-u) - \rho_1) d\Phi_j(u) + \right. \\ &+ \rho_1(1 - \Phi_j(t)) \left| \leq \int_0^L |p_{11}(t-u) - \rho_1| d\Phi_j(u) + \right. \\ &\left. + 2(1 - \Phi_j(L)) + (1 - \Phi_j(t)). \right| \end{aligned}$$

При $t \rightarrow \infty$ правая часть может быть выбором L сделана сколь угодно малой. Тем самым для любого $i \lim_{t \rightarrow \infty} p_{11}(t) = \rho_1$, что доказывает теорему.

ПМП используются как средство описания и изучения достаточно широкого класса СМО и других моделей, которые нельзя описать в виде марковских процессов с дискретным множеством состояний.

Задания для самоконтроля

1. Найти одномерные распределения процесса восстановления $v(t)$ в случае, когда длительность восстановления имеет:
 - а) равномерное на интервале $(0,1)$ распределение;
 - б) геометрическое распределение со средним a ;
 - в) распределение Эрланга с параметрами (n, μ) .
2. Показать, что соответствие между функциями восстановления $H(t)$ и функциями распределения $F(t)$ длительности восстановления взаимно-однозначно.
3. Найти автокорреляционную функцию для простейшего потока.
4. Показать, что $\frac{v(t)}{t} \xrightarrow[t \rightarrow \infty]{P} \frac{1}{M\tau_1}$ для произвольного процесса восстановления.

1.5. Процессы размножения и гибели

Изучим один важный класс процессов, который часто используется при исследовании различных характеристик марковских СМО. Пусть $x(t), t \geq 0$ — случайный процесс с множеством состояний $\{0, 1, 2, \dots\}$, эволюция которого

протекает следующим образом: если в момент t $x(t) = k$, то независимо от значений $x(s)$, $s < t$, на промежутке $[t, t+h]$ с вероятностью $\lambda_k h + o(h)$ $x(t)$ перейдет в состояние $k+1$, с вероятностью $\mu_k h + o(h)$ — в состояние $k-1$, с вероятностью $1 - (\lambda_k + \mu_k)h + o(h)$ останется в состоянии k , а вероятность других изменений есть $o(h)$. Схематически это можно записать в виде

$$k \xrightarrow{h} \begin{cases} k+1, & \text{с вероятностью } \lambda_k h + o(h); \\ k-1, & \text{с вероятностью } \mu_k h + o(h); \\ k & \text{с вероятностью } 1 - (\lambda_k + \mu_k)h + o(h). \end{cases} \quad (1.23)$$

Определение 1.11. Введенный процесс называется процессом размножения и гибели, задаваемый параметрами λ_k ; μ_k ; $k \geq 0$ (при этом $\mu_0 = 0$).

Величины λ_k называются интенсивностями размножения, а μ_k — интенсивностями гибели. Если $\mu_k \equiv 0$; $k \geq 0$, то $x(t)$ называется процессом чистого размножения. Отметим, что по построению очевидно, что $x(t)$ — марковский процесс.

Пример 1.2. Пусть $v(t)$ — число событий простейшего потока с параметром λ , осуществившихся на промежутке $[0, t]$. Тогда легко видеть, что $v(t)$ — процесс чистого размножения, у которого $\lambda_k \equiv \lambda$; $k \geq 0$.

Пример 1.3. Рассмотрим процесс развития колонии бактерий, который характеризуется тем, что каждая бактерия развивается независимо от других и на промежутке $[t, t+\Delta]$ она с вероятностью $\mu\Delta + o(\Delta)$ может погибнуть, а с вероятностью $\lambda\Delta + o(\Delta)$ разделится на две бактерии, каждая из которых будет развиваться по тем же законам. Если $x(t)$ означает число бактерий в момент t , то нетрудно убедиться в том, что $x(t)$ — процесс размножения и гибели с параметрами $\lambda_k = k\lambda$; $\mu_k = k\mu$; $k \geq 0$.

Рассмотрим некоторые свойства процесса $x(t)$.

1. Найдем распределение времени T_k до первого выхода из состояния k при условии, что вначале процесс находится в состоянии k . Обозначим $q_k(t) = P\{T_k \geq t\}$. По построению процесса

$$q_k(t+h) = q_k(t) q_k(h) = q_k(t) (1 - (\lambda_k + \mu_k)h + o(h))$$

и

$$\frac{q_k(t+h) - q_k(t)}{h} = -(\lambda_k + \mu_k) q_k(t) + o(1).$$

Отсюда $q'_k(t) = -(\lambda_k + \mu_k) q_k(t)$.

Так как по условию $q_k(0) = 1$, то

$$q_k(t) = \exp\{-(\lambda_k + \mu_k)t\},$$

т. е. T_k имеет показательное распределение с параметром $\lambda_k + \mu_k$.

2. Вычислим вероятность того, что переход из k -го состояния произойдет в $(k+1)$ -е состояние. В предположении, что $T_k \in [t, t+h]$, имеем

$$\begin{aligned} P\{k \xrightarrow{[t, t+h]} k+1 / T_k \in [t, t+h]\} &= \\ &= \frac{P\{T_k > t, k \xrightarrow{[t, t+h]} k+1\}}{P\{T_k \in [t, t+h]\}} = \\ &= \frac{e^{-(\lambda_k + \mu_k)t} \lambda_k h + o(h)}{(\lambda_k + \mu_k) e^{-(\lambda_k + \mu_k)t} h + o(h)} \xrightarrow{h \rightarrow 0} \frac{\lambda_k}{\lambda_k + \mu_k}. \end{aligned}$$

Поскольку полученная вероятность не зависит от того, в какой момент произойдет выход из состояния k , то она и будет искомой.

Соответственно вероятность перехода из k состояния в $k-1$ будет $1 - \frac{\lambda_k}{\lambda_k + \mu_k} = \frac{\mu_k}{\lambda_k + \mu_k}$.

3. В силу свойств 1 и 2 развитие процесса $x(t)$ происходит следующим образом: в k -м состоянии $x(t)$ проводит случайное время T_k , распределенное по показательному закону с параметром $\lambda_k + \mu_k$, после чего с вероятностью $\frac{\lambda_k}{\lambda_k + \mu_k}$ переходит в состояние $(k+1)$, либо с вероятностью $\frac{\mu_k}{\lambda_k + \mu_k}$ — в $(k-1)$.

4. Выведем систему дифференциальных уравнений для переходных вероятностей процесса.

Обозначим $p_{ij}(t) = P\{x(t) = j / x(0) = i\}$.

Имеет место система так называемых прямых дифференциальных уравнений Колмогорова:

$$\begin{cases} p'_{i0}(t) = -\lambda_0 p_{i0}(t) + \mu_1 p_{i1}(t); \\ p'_{ij}(t) = (-\lambda_j + \mu_j) p_{ij}(t) + \lambda_{j-1} p_{i,j-1}(t) + \\ + \mu_{j+1} p_{i,j+1}(t), \quad j > 0 \end{cases} \quad (1.24)$$

и система обратных уравнений:

$$\begin{cases} P'_{0j}(t) = -\lambda_0 p_{0j}(t) + \mu_1 p_{1j}(t); \\ P'_{ij}(t) = -(\lambda_i + \mu_i) p_{ij}(t) + \lambda_i p_{i+1,j}(t) + \\ + \mu_i p_{i-1,j}(t). \end{cases} \quad (1.25)$$

Укажем вывод системы (1.24). Так как $x(t)$ марковский процесс, то, используя формулу полной вероятности, нетрудно получить следующее соотношение, известное как тождество Колмогорова — Чепмена,

$$p_{ij}(t+h) = \sum_{k=0}^{\infty} p_{ik}(t) p_{kj}(h), \quad h > 0, \quad (1.26)$$

отсюда, используя (1.23), получим

$$\begin{aligned} \frac{p_{ij}(t+h) - p_{ij}(t)}{h} &= p_{ij}(t) \frac{(p_{ij}(h) - 1)}{h} + \\ + \sum_{k \neq j} p_{ik}(t) \frac{p_{kj}(h)}{h} &= (\lambda_j + \mu_j) p_{ij}(t) + \lambda_{j-1} p_{i,j-1}(t) + \\ + \mu_{j+1} p_{i,j+1}(t) &+ o(1), \end{aligned}$$

поскольку из (1.23) вытекает, что $p_{j-1,j}(h) = \lambda_{j-1}h + o(h)$,

$$\begin{aligned} p_{j+1,j}(h) &= \mu_j h + o(h), \quad P_{j,j}(h) = \\ &= 1 - (\lambda_j + \mu_j)h + o(h), \end{aligned}$$

$\frac{p_{kj}(h)}{h} = o(1)$ при $k \neq j-1, k, j+1$ (для того, чтобы предельный переход в приведенных преобразованиях был строго обоснован, достаточно потребовать, чтобы дополнительно при любом $j \geq 0$

$$\overline{\lim}_{h \rightarrow 0} \sup_{k \neq j-1, j+1} \frac{p_{kj}(h)}{h} < \infty).$$

Аналогично выводятся системы (1.25). Системы (1.24) и (1.25) в дальнейшем будут использоваться для вычисления различных переходных характеристик марковских СМО.

5. Исследуем поведение $p_{ij}(t)$ при $t \rightarrow \infty$.

Если $\lambda_k > 0$, $k \geq 0$ и $\mu_k > 0$, $k > 0$, то множество состояний $x(t)$ неприводимо (т. е. все состояния сообщаются) и из общей теории для цепей Маркова следует, что пределы $\lim_{t \rightarrow \infty} p_{ij}(t) = \pi_j$, $j \geq 0$ существуют, не зависят от

начального состояния i и удовлетворяют системе уравнений, которая формально получается из (1.24) заменой $p_{ij}(t)$ на π_j , а $p'_{ij}(t)$ на 0:

$$\begin{cases} \lambda_0 \pi_0 = \mu_1 \pi_1; \\ (\lambda_j + \mu_j) \pi_j = \lambda_{j-1} \pi_{j-1} + \mu_{j+1} \pi_{j+1}, \quad j \geq 1. \end{cases} \quad (1.27)$$

Если система (1.27) имеет положительное решение $\pi_j > 0$, такое, что $\sum_{j \geq 0} \pi_j = 1$, то $\pi_j, j \geq 0$ называется стационарным распределением. Название стационарное объясняется тем, что распределение π_j удовлетворяет равенству

$$\pi_j = \sum_{k=0}^{\infty} \pi_k p_{kj}(u), \quad (1.28)$$

которое получится, если в тождестве (1.26) перейти к пределу при $t \rightarrow \infty$. А соотношение (1.28) свидетельствует о том, что если в качестве начального взять стационарное распределение π_j , то распределение $x(u)$ в любой момент времени u будет совпадать с $\pi_j, j \geq 0$.

6. Вычислим явный вид величины $\pi_j, j \geq 0$. Обозначим $z_j = \mu_{j+1} \pi_{j+1} - \lambda_j \pi_j, j \geq 0$. Тогда из (1.27) получим соотношения

$$z_0 = 0,$$

$$z_{j-1} = z_j, \quad j > 0.$$

Отсюда

$$z_j \equiv 0, \quad j \geq 0, \quad \pi_{j+1} = \frac{\lambda_j}{\mu_{j+1}} \pi_j, \quad j \geq 0.$$

Обозначим

$$a_j = \frac{\lambda_0 \lambda_1 \dots \lambda_{j-1}}{\mu_1 \mu_2 \dots \mu_j}.$$

По индукции легко вывести, что $\pi_j = a_j \pi_0, j > 0$.

Для того чтобы $\sum_{j=0}^{\infty} \pi_j = 1$, необходимо и достаточно,

чтобы $\sum_{j=0}^{\infty} a_j < \infty$. В этом случае

$$\pi_j = a_j \left(\sum_{k=0}^{\infty} a_k \right)^{-1}, \quad j \geq 0. \quad (1.29)$$

Если же $\sum_{j=0}^{\infty} a_j = \infty$, то тогда $\pi_j \equiv 0, j \geq 0$ и стационарного распределения не существует.

Задания для самоконтроля

1. Выписать систему уравнений для вероятностей $P_k(t)$ для процесса гибели и размножения с параметрами

$$\lambda_k = \begin{cases} (N-k)\lambda, & k \leq N; \\ 0 & k > N, \end{cases} \quad \mu_k = \begin{cases} k\mu, & k \leq N; \\ 0 & k > N. \end{cases}$$

2. Пусть параметры процесса гибели и размножения имеют вид

$$\lambda_k = k\mu, \quad \mu_k = k\mu;$$

а) найти дифференциальное уравнение для производящей функции

$$P(z, t) = \sum_{k=0}^{\infty} P_k(t) z^k;$$

б) показать, что при начальном условии $x(0) = 0$ функция

$$P(z, t) = \frac{\mu(1 - e^{(\lambda - \mu)t}) - (\lambda - \mu e^{(\lambda - \mu)t})z}{\mu - \lambda e^{(\lambda - \mu)t} - \lambda(1 - e^{(\lambda - \mu)t})z}$$

является его решением;

в) разлагая $P(z, t)$ в степенной ряд, показать, что

$$P_k(t) = (1 - \alpha(t))(1 - \beta(t))\beta^{k-1}(t), \quad k \geq 1, \\ P_0(t) = \alpha(t)$$

и найти $\alpha(t)$ и $\beta(t)$;

г) найти $Mx(t)$ и $Dx(t)$;

д) найти $\lim_{t \rightarrow \infty} P_0(t)$.

3. Пусть параметры процесса гибели и размножения равны $\lambda_k = k\lambda + \alpha$, $\mu_k = k\mu$;

а) найти уравнения, связывающие вероятности $P_k(t)$, и уравнение для производящей функции $P(z, t) = \sum_{k=0}^{\infty} P_k(t) z^k$;

б) найти дифференциальное уравнение для $M(t) = Mx(t)$ и решить его;

в) найти $\lim_{t \rightarrow \infty} M(t)$.

4. Пусть $\lambda_k = \begin{cases} \lambda_1, & 0 \leq k \leq N; \\ 2\lambda_1, & k > N \end{cases} \quad \mu_k = \mu, \quad k \geq 1$.

Найти стационарные вероятности π_k . При каких условиях стационарное распределение существует?

5. Пусть $\lambda_k = \alpha^k \lambda$, $k \geq 0$, $0 \leq \alpha < 1$, $\mu_k = \mu$, $k \geq 1$.

Найти π_k .

6. Пусть $\lambda_k = (k+2)\lambda$, $k \geq 0$, $\mu_k = k\mu$, $k \geq 1$.

Найти π_k и среднее значение процесса.

ГЛАВА 2. МОДЕЛИ СИСТЕМ МАССОВОГО ОБСЛУЖИВАНИЯ

2.1. Марковские системы

Система $M/M/1/\infty$. В соответствии с классификацией СМО входящий поток рассматриваемой системы является простейшим, обслуживание производится одним устройством, а время обслуживания имеет показательное распределение. Количество мест для ожидания не ограничено. Очередь в системе образуется по принципу «первый пришел — первый обслужен». Указанная СМО часто обозначается как система $M|M|1$ с ожиданием и носит название модели Эрланга. Обозначим через λ параметр входного потока, а через μ — параметр времени обслуживания, который назовем интенсивностью обслуживания.

Изучим вначале распределение количества требований $Q(t)$ в системе в момент времени t . Легко видеть, что для рассматриваемой системы процесс $Q(t)$ является однородной цепью Маркова с непрерывным временем. В самом деле, пусть нам известно, как развивался процесс $Q(t)$ на интервале $[0, T]$, и, в частности, известно значение $Q(T) = i$. Допустим, например, что $i \geq 1$. Это значит, что на обслуживании находится требование, а оставшиеся $i - 1$ ожидают начала обслуживания. Дальнейшие изменения значения $Q(t)$ при $t > T$ связаны с такими факторами:

- а) моментами поступления новых требований;
- б) количеством времени, необходимого для завершения обслуживания обслуживаемого требования;
- в) количеством времени, необходимого для обслуживания следующих требований.

Если $Q(T) = i \geq 1$ известно, то все эти факторы не зависят от развития процесса до рассматриваемого момента времени T из-за свойств отсутствия последействия простейшего потока и показательного распределения.

Если $Q(T) = 0$, то на дальнейшее развитие процесса $Q(t)$ оказывают влияние лишь факторы, указанные в пунктах а) и в).

Итак, $Q(t)$ — цепь Маркова. Очевидно, что ее состояниями являются натуральные числа. Найдем интенсивности q_{ij} вероятностей ее переходов. По определению, при $j \neq i$

$$q_{ij} = \lim_{t \rightarrow 0} \frac{1}{t} P\{Q(t) = j / Q(0) = i\}.$$

Если $j > i + 1$, то событие $Q(t) = j$ при условии $Q(0) = i$ может произойти лишь при поступлении не менее двух требований, и в силу ординарности простейшего потока

$$P\{Q(t) = j / Q(0) = i\} = o(t), \quad t \rightarrow 0.$$

Аналогично, при $j < i - 1$ для наступления события $Q(t) = j$ при условии $Q(0) = i$ необходимо обслуживание хотя бы двух требований. Поэтому, обозначив времена обслуживания этих требований через ξ_1 и ξ_2 , получим

$$\begin{aligned} P\{Q(t) = j / Q(0) = i\} &\leq P\{\xi_1 < t, \xi_2 < t\} = \\ &= P\{\xi_1 < t\} P\{\xi_2 < t\} = (1 - e^{-\mu t})^2 = o(t), \quad t \rightarrow 0. \end{aligned}$$

Пусть теперь $j = i + 1$. Рассуждения, аналогичные приведенным выше, показывают, что главная часть вероятности

$$P\{Q(t) = i + 1 / Q(0) = i\}$$

равна $P\{\text{на } [0, t) \text{ поступило одно требование}\} = \lambda t + o(t)$. Точно так же при $j = i - 1$, $i > 0$ получим $P\{Q(t) = i - 1 / Q(0) = i\} = P\{\text{на } [0, t) \text{ обслужено одно требование}\} + o(t) = P\{\xi_1 < t\} + o(t) = \mu t + o(t)$.

Поэтому при $i \neq j$

$$q_{ij} = \begin{cases} 0, & |i - j| \neq 1; \\ \lambda, & j = i + 1; \\ \mu, & j = i - 1, \quad i > 0. \end{cases}$$

Так как q_{ij} отличны от 0 лишь при $|i - j| = 1$, то, меняя свои состояния, цепь $Q(t)$ может переходить лишь в соседние. Таким образом, мы видим, что $Q(t)$ образует процесс размножения и гибели с параметрами $\lambda_k \equiv \lambda$, $k = 0, 1, \dots$, $\mu_k \equiv \mu$, $k \geq 1$.

Допустим, что начальным состоянием цепи $Q(t)$ является k и обозначим $p_i(t) = P\{Q(t) = i / Q(0) = k\}$. Тогда

функции $p_i(t)$ удовлетворяют прямой системе уравнений Колмогорова:

$$\begin{cases} \dot{p}_0(t) = -\lambda p_0(t) + \mu p_1(t); \\ \dot{p}_i(t) = -(\lambda + \mu) p_i(t) + \lambda p_{i-1}(t) + \\ + \mu p_{i+1}(t), i \geq 1, \end{cases} \quad (2.1)$$

с начальным условием $p_i(0) = \delta_{ik}$, $i \geq 0$.

В теории цепей Маркова доказывается, что эта система уравнений имеет единственное ограниченное решение (при заданных начальных условиях). Найдем его. Для этого введем производящую функцию

$$P(z, t) = \sum_{i=0}^{\infty} p_i(t) z^i.$$

Умножив i -е уравнение системы (2.1) на z^{i+1} и сложив их, получим, что $P(z, t)$ удовлетворяет уравнению

$$z \frac{\partial P(z, t)}{\partial t} = (1 - z) [(\mu - \lambda z) P(z, t) - \mu p_0(t)]. \quad (2.2)$$

Уравнение (2.2) решим, используя преобразования Лапласа,

$$P^*(z, s) = \int_0^{\infty} e^{-st} P(z, t) dt, \quad P_0^*(s) = \int_0^{\infty} e^{-st} p_0(t) dt.$$

Умножив уравнение (2.2) на e^{-st} и проинтегрировав его от 0 до ∞ , получим

$$z \int_0^{\infty} e^{-st} \frac{\partial P(z, t)}{\partial t} dt = (1 - z) [(\mu - \lambda z) P^*(z, s) - \mu P_0^*(s)]. \quad (2.3)$$

Интегрируя по частям, найдем левую часть (2.3):

$$\begin{aligned} z \int_0^{\infty} e^{-st} \frac{\partial P(z, t)}{\partial t} dt &= z (e^{-st} P(z, t)) \Big|_0^{\infty} + \\ + s \int_0^{\infty} e^{-st} P(z, t) dt &= -z^{k+1} + z s P^*(z, s), \end{aligned} \quad (2.4)$$

так как $P(z, 0) = z^k$. Из равенств (2.3) и (2.4) следует, что

$$P^*(z, s) = \frac{z^{k+1} - \mu (1 - z) P_0^*(s)}{zs - (1 - z)(\mu - \lambda z)}. \quad (2.5)$$

Равенство (2.5) содержит неизвестную пока функцию $P_0^*(s)$, которую найдем, используя ограниченность функции $P^*(z, s)$ при $|z| < 1$, $\text{Re } s > 0$.

Для этого заметим, что знаменатель $\lambda z^2 - (\lambda + \mu + s)z + \mu$ в правой части равенства (2.5) имеет два корня

$$z_{1,2} = z_{1,2}(s) = \frac{\lambda + \mu + s \pm \sqrt{(\lambda + \mu + s)^2 - 4\lambda\mu}}{2\lambda},$$

ровно один из которых находится в круге $|z| < 1$. Для проверки последнего утверждения можно воспользоваться известной в теории функций комплексного переменного теоремой Руше, которая формулируется следующим образом.

Если две функции $f(z)$ и $g(z)$ — аналитические в замкнутой области, ограниченной контуром C , и удовлетворяют на C условию $|g(z)| < |f(z)|$, то внутри C функции $f(z)$ и $f(z) + g(z)$ имеют одинаковое количество нулей.

Действительно, положим $f(z) = (\lambda + \mu + s)z$, $g(z) = \mu + \lambda z^2$. Так как при $\text{Re } s > 0$ $|f(z)| > |g(z)|$ на границе единичного круга, а $f(z)$ имеет один корень $z = 0$, то $f(z) + g(z)$ имеет один корень в круге $|z| < 1$. Этим корнем будет $z_2(s)$ (со знаком «минус» перед радикалом).

Так как $P^*(z, s)$ определена при $z = z_2(s)$, то числитель правой части равенства (1.5) должен обратиться в нуль при $z = z_2(s)$, и получаем уравнение для определения $P_0^*(s)$;

$$z_2^{k+1}(s) - \mu(1 - z_2(s))P_0^*(s) = 0,$$

откуда

$$P_0^*(s) = \frac{z_2^{k+1}(s)}{\mu(1 - z_2(s))}. \quad (2.6)$$

Теперь можем записать, что

$$\begin{aligned} P^*(z, s) &= \frac{z^{k+1} - (1 - z) \frac{z_2^{k+1}(s)}{1 - z_2(s)}}{zs - (1 - z)(\mu - \lambda z)} = \\ &= \frac{z^{k+1} - z_2(s) z^{k+1} - z_2^{k+1}(s) + z z_2^{k+1}(s)}{-\lambda(z - z_1(s))(z - z_2(s))(1 - z_2(s))} = \\ &= \frac{(z - z_2(s))(z^k + z^{k-1} z_2(s) + \dots + z_2^k(s)) - z z_2(s)(z - z_2(s)) \times \\ &\quad \times (z^{k-1} + z^{k-2} z_2(s) + \dots + z_2^{k-1}(s))}{-\lambda(z - z_1(s))(z - z_2(s))(1 - z_2(s))} = \\ &= \frac{z^k + z^{k-1} z_2(s) + \dots + z_2^k(s)}{\lambda z_1(s) \left(1 - \frac{z}{z_1(s)}\right)} + \end{aligned}$$

$$+ \frac{z_2^{k+1}(s)}{\lambda z_1(s) \left(1 - \frac{z}{z_1(s)}\right) (1 - z_2(s))} = \frac{1}{\lambda z_1(s)} (z^k + z^{k-1} z_2(s) + \dots + z_2^k(s)) \sum_{i=0}^{\infty} \left(\frac{z}{z_1(s)}\right)^i + \frac{z_2^{k+1}(s)}{\lambda z_1(s) (1 - z_2(s))} \sum_{i=0}^{\infty} \left(\frac{z}{z_1(s)}\right)^i$$

(ряды в последнем выражении сходятся, так как $|z_1(s)| > 1$). Для нахождения $p_n(t)$ можно определить коэффициент при z^n в выражении для $P^*(z, s)$ — он является преобразованием Лапласа — $\int_0^{\infty} e^{-st} p_n(t) dt = P_n^*(s)$. После этого $p_n(t)$ можно определить по формуле

$$p_n(t) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{st} P_n^*(s) ds,$$

где c — достаточно большая действительная постоянная. Ограничимся случаем, когда $k = 0$. Тогда

$$P^*(z, s) = \frac{1}{\lambda z_1(s) (1 - z_2(s))} \sum_{i=0}^{\infty} \left(\frac{1}{z_1(s)}\right)^i;$$

$$P_n^*(s) = \frac{1}{\lambda z_1^{n+1}(s) (1 - z_2(s))}.$$

Последняя функция является преобразованием Лапласа для

$$p_n(t) = e^{-(\lambda+\mu)t} \left[\left(\frac{\lambda}{\mu}\right)^{\frac{n}{2}} I_{n-1}(2\sqrt{\lambda\mu}t) + \left(\frac{\lambda}{\mu}\right)^{\frac{n-1}{2}} I_{n+1}(2\sqrt{\lambda\mu}t) + \left(1 - \frac{\lambda}{\mu}\right) \left(\frac{\lambda}{\mu}\right)^n \sum_{k=n+2}^{\infty} \left(\frac{\lambda}{\mu}\right)^{\frac{k}{2}} \times \times I_k(2\sqrt{\lambda\mu}t) \right],$$

где $I_\nu(z)$ — обобщенная функция Бесселя первого рода

$$I_\nu(z) = \sum_{k=0}^{\infty} \frac{1}{k! \Gamma(\nu + k + 1)} \left(\frac{z}{2}\right)^{\nu+2k}, \quad I_{-1}(z) = I_1(z).$$

Проведенные вычисления показывают, с какими затруднениями приходится сталкиваться при нахождении явных характеристик даже простейших СМО. Вместе с тем

зачастую интерес представляют лишь пределы этих характеристик при $t \rightarrow \infty$. Займемся поэтому изучением стационарного режима системы $M|M|1| \infty$.

Предполагая, что система находится в стационарном режиме, и, следовательно, $p_i(t) \equiv p_i$, из системы (2.1) получим

$$\begin{cases} -\lambda p_0 + \mu p_1 = 0; \\ -(\lambda + \mu) p_i + \lambda p_{i-1} + \mu p_{i+1} = 0, \quad i \geq 1. \end{cases} \quad (2.7)$$

Обозначим $z_i = -\lambda p_{i-1} + \mu p_i$, $i \geq 1$. Тогда система (2.7) запишется в виде

$$\begin{cases} z_1 = 0; \\ -z_i + z_{i+1} = 0, \quad i \geq 1. \end{cases}$$

Следовательно, $z_i \equiv 0$ и $p_i = \frac{\lambda}{\mu} p_{i-1} = \left(\frac{\lambda}{\mu}\right)^i p_0$. Так как должно выполняться условие нормировки $\sum_{i=0}^{\infty} p_i = 1$, то необходимым условием существования стационарного режима будет неравенство $\lambda < \mu$. В этом случае $p_0 = 1 - \frac{\lambda}{\mu}$, а $p_i = \left(1 - \frac{\lambda}{\mu}\right) \left(\frac{\lambda}{\mu}\right)^i$.

Отношение $\rho = \frac{\lambda}{\mu}$ называют загрузкой системы. Можно показать, что неравенство $\rho < 1$ является и достаточным условием существования стационарного режима.

Знание величины p_i позволяет найти среднее значение числа требований в стационарном режиме:

$$MQ(t) = \sum_{i=0}^{\infty} i p_i = \sum_{i=0}^{\infty} i (1 - \rho) \rho^i = \frac{\rho}{1 - \rho}. \quad (2.8)$$

Формула (2.8) показывает, что при значениях ρ , близких к 1, среднее количество требований быстро растет (при $\rho = 1$ стационарный режим вообще отсутствует!). Это обстоятельство характерно не только для СМО типа $M|M|1| \infty$, но и для многих других и должно учитываться при расчете количества касов в магазине, мест в больницах и т. д., т. е. в случаях, когда длительное пребывание требования в очереди нежелательно.

Система $M|M|1|n$. В указанной системе в очереди может находиться не более n требований. Если в некоторый момент времени, когда все места для ожидания заняты, поступает очередное требование, оно теряется, и система развивается так, как если бы это требование

вообще не поступало. Системы с такими свойствами называют системами с потерями. Как и ранее, обозначим через $Q(t)$ количество требований в системе, $Q(0) = k$, $0 \leq k \leq n+1$, $P\{Q(t) = i\} = p_i(t)$. Процесс $Q(t)$ является процессом гибели и размножения, интенсивности перехода которого имеют вид:

$$\begin{aligned}\lambda_i &= \lambda, & 0 \leq i \leq n; \\ \mu_i &= \mu, & 1 \leq i \leq n+1,\end{aligned}$$

где λ — параметр входного потока, а μ — интенсивность обслуживания.

Прямая система уравнений Колмогорова для вероятностей $p_i(t)$ запишется в виде

$$\begin{cases} p'_0(t) = -\lambda p_0(t) + \mu p_1(t); \\ p'_i(t) = -(\lambda + \mu) p_i(t) + \lambda p_{i-1}(t) + \\ + \mu p_{i+1}(t), & 1 \leq i \leq n; \\ p'_{n+1}(t) = -\mu p_{n+1}(t) + \lambda p_n(t), \end{cases} \quad (2.9)$$

а начальные условия — в виде

$$p_i(0) = \delta_{ik}, \quad 0 \leq i \leq n+1. \quad (2.10)$$

Так как система (2.9) линейна с постоянными коэффициентами, то используя теоремы теории дифференциальных уравнений, нетрудно получить, что при заданном начальном условии (2.10) она имеет единственное решение. Не будем его находить, а займемся лишь вычислением эргодического распределения

$$p_i = \lim_{t \rightarrow \infty} p_i(t),$$

которое заведомо существует, так как цепь Маркова $Q(t)$ имеет конечное количество состояний. Вероятности p_i удовлетворяют системе уравнений

$$\begin{cases} 0 = -\lambda p_0 + \mu p_1; \\ 0 = -(\lambda + \mu) p_i + \lambda p_{i-1} + \mu p_{i+1}; \\ 0 = -\mu p_{n+1} + \lambda p_n, \end{cases} \quad (2.11)$$

с условием нормировки

$$\sum_{i=0}^{n+1} p_i = 1. \quad (2.12)$$

Согласно общим формулам для стационарного распределения процессов размножения и гибели, легко получим

$$p_i = \frac{\lambda}{\mu} p_{i-1} = \left(\frac{\lambda}{\mu}\right)^i p_0, \quad 1 \leq i \leq n+1 \quad (2.13)$$

и, используя условие нормировки (2.12), окончательно

$$p_i = \rho^i \frac{1-\rho}{1-\rho^{n+2}}, \quad 0 \leq i \leq n+1, \quad (2.14)$$

где $\rho = \frac{\lambda}{\mu}$. Если $\rho = 1$, то из (2.13) следует, что $p_i \equiv p_0$, $i = \overline{1, n+1}$, и тогда

$$p_i = \frac{1}{n+2}, \quad 0 \leq i \leq n+1. \quad (2.15)$$

Вероятности $p_i(t)$ (p_i), нахождению которых уделялось все внимание, определялись как нестационарные (стационарные) вероятности того, что в момент времени t количество требований в системе равно i . Оказывается, что их можно интерпретировать и другим, весьма важным, способом. Обозначим $A(t, \Delta) = \{\text{на } [t, t+\Delta] \text{ поступило требование}\}$,

$$R_i(t) = \lim_{\Delta \rightarrow 0} P\{Q(t) = i / A(t, \Delta)\},$$

где $R_i(t)$ следует понимать как условную вероятность того, что в момент прихода требования в систему в последней находилось i требований при условии, что требование поступило в момент t .

Теорема 2.1. Для систем с простейшим входным потоком $p_i(t) = R_i(t)$.

Доказательство. Имеем

$$\begin{aligned} P\{Q(t) = i / A(t, \Delta)\} &= \frac{P\{A(t, \Delta) / Q(t) = i\} P\{Q(t) = i\}}{P\{A(t, \Delta)\}} = \\ &= \frac{P\{A(t, \Delta)\}}{P\{A(t, \Delta)\}} P\{Q(t) = i\} = p_i(t), \end{aligned}$$

так как в силу отсутствия последействия простейшего потока событие $A(t, \Delta)$ не зависит от события $\{Q(t) = i\}$. Поэтому $R_i(t) = p_i(t)$.

Особый интерес для систем с потерями представляет вероятность отказа $P_{\text{отк}}$ — вероятность того, что в стационарном режиме требование, поступившее в момент времени t , будет потеряно. Так как в силу доказанной теоремы

эта вероятность для системы $M|M|1|n$ равна вероятности p_{n+1} того, что в неслучайный момент времени t система занята, то

$$P_{\text{отк}} = \begin{cases} \rho^{n+1} \cdot \frac{1-\rho}{1-\rho^{n+2}}, & \rho \neq 1; \\ \frac{1}{n+2}, & \rho = 1. \end{cases}$$

Система $M|M|\infty$. Так как в этой СМО число обслуживающих устройств бесконечно, то вновь прибывшее требование немедленно поступает на обслуживание.

Обозначим, как и ранее, параметр входного потока через λ , интенсивность обслуживания через μ , а количество требований в системе в момент времени t через $Q(t)$.

Процесс $Q(t)$ в системе $M|M|\infty$ также является цепью Маркова. В самом деле, пусть $Q(T) = k$. Тогда в системе находится k требований, и их оставшиеся времена обслуживания не зависят от развития процесса $Q(t)$ при $t < T$ и имеют показательное распределение с параметром λ . Оставшиеся факторы — моменты поступлений очередных требований и их времена обслуживания не зависят от значений процесса $Q(t)$ при $t \leq T$.

Так как переходы $Q(t)$ возможны только в соседнее состояние, то $Q(t)$ — процесс гибели и размножения. Найдем его характеристики. Очевидно, что, как и ранее,

$$\lambda_i = \lambda, i \geq 0. \text{ Далее } \mu_i = \lim_{t \rightarrow 0} \frac{1}{t} P\{Q(t) = i - 1 | Q(0) = i\}.$$

Пусть $\xi_1, \dots, \xi_i$ — независимые в совокупности величины, имеющие одно и то же показательное распределение с параметром μ . Тогда

$$P\{Q(t) = i - 1 | Q(0) = i\} = \sum_{k=1}^i P\{\xi_1 > t, \dots, \xi_{k-1} > t,$$

$$\xi_k < t, \xi_{k+1} > t, \xi_i > t\} + o(t) = i e^{-\mu(t-1)t} (1 - e^{-\mu t}) + o(t) = i \mu t + o(t).$$

Поэтому $\mu_i = i\mu$ (другими словами, интенсивность обслуживания i устройствами в i раз больше, чем одним).

Допустим, что $Q(t) = i$, и обозначим $p_k(t) = P\{Q(t) = k\}$. Тогда вероятности $p_k(t)$ удовлетворяют прямой системе уравнений Колмогорова:

$$\begin{cases} p'_0(t) = -\lambda p_0(t) + \mu p_1(t); \\ p'_k(t) = -(\lambda + k\mu) p_k(t) + \lambda p_{k-1}(t) + (k+1) \mu p_{k+1}(t), & k \geq 1 \end{cases}$$

с начальным условием $p_k(0) = \delta_{ki}$, $k \geq 0$. Введя производящую функцию $P(z, t) = \sum_{k=0}^{\infty} p_k(t) z^k$, получим, что она удовлетворяет уравнению в частных производных:

$$\frac{\partial P(z, t)}{\partial t} = -\lambda(1-z)P(z, t) + \mu(1-z)\frac{\partial P(z, t)}{\partial z}.$$

Решая его по методу Лагранжа, получим, что общее решение имеет вид

$$P(z, t) = e^{\frac{\lambda}{\mu} z} g[(1-z)e^{-\mu t}], \quad (2.16)$$

где функция $g(z)$ находится из начальных условий. Подставляя в равенство (2.16) значение $t = 0$, получим

$$z^i = e^{\frac{\lambda}{\mu} z} g(1-z),$$

откуда

$$g(z) = (1-z)^i e^{-\frac{\lambda}{\mu}(1-z)}. \quad (2.17)$$

Поэтому

$$P(z, t) = [1 - (1-z)e^{-\mu t}]^i \exp \left\{ -\frac{\lambda}{\mu} (1-z)(1 - e^{-\mu t}) \right\}. \quad (2.18)$$

Разлагая правую часть соотношения (2.18) в ряд по степеням z , можно найти $p_k(t)$ как коэффициент при z^k .

В нашем случае легко видеть, что $P(z, t)$ есть произведение двух производящих функций, первая из которых соответствует биномиальному распределению с параметрами $(i, e^{-\mu t})$, а вторая — пуассоновскому распределению с параметром $\frac{\lambda}{\mu}(1 - e^{-\mu t})$.

Следовательно, искомое распределение есть композиция биномиального и пуассоновского, откуда

$$\begin{aligned} p_k(t) &= \sum_{j=0}^{\min(k, i)} C_i^j e^{-j\mu t} (1 - e^{-\mu t})^{i-j} \times \\ &\times \exp \left\{ -\frac{\lambda}{\mu} (1 - e^{-\mu t}) \right\} \left(\frac{\lambda}{\mu} \right)^{k-j} \frac{(1 - e^{-\mu t})^{k-j}}{(k-j)!} = \\ &= \sum_{j=0}^{\min(k, i)} C_i^j \exp \left\{ -\frac{\lambda}{\mu} (1 - e^{-\mu t}) \right\} \times \\ &\times \frac{\left(\frac{\lambda}{\mu} \right)^{k-j}}{(k-j)!} e^{-j\mu t} (1 - e^{-\mu t})^{k+i-2j}. \end{aligned}$$

Найдем стационарное распределение. Из равенства (2.18) следует, что

$$\lim_{t \rightarrow \infty} P(z, t) = e^{-\frac{\lambda}{\mu}(1-z)} = e^{-\frac{\lambda}{\mu}} \sum_{k=0}^{\infty} \left(\frac{\lambda}{\mu}\right)^k \frac{z^k}{k!},$$

и поэтому $p_k = \lim_{t \rightarrow \infty} p_k(t) = e^{-\frac{\lambda}{\mu}} \left(\frac{\lambda}{\mu}\right)^k \cdot \frac{1}{k!}$. Заметим, что стационарные вероятности можно также непосредственно вычислить из соответствующей системы уравнений:

$$\begin{cases} 0 = -\lambda p_0 + \mu p_1; \\ 0 = -(\lambda + k\mu) p_k + \lambda p_{k-1} + (k+1)\mu p_{k+1}, \quad k \geq 1 \end{cases}$$

с условием нормировки $\sum_{k=0}^{\infty} p_k = 1$:

$$p_k = e^{-\frac{\lambda}{\mu}} \frac{\rho^k}{k!}.$$

Итак, в стационарном режиме число требований в системе $M|M|\infty$ имеет пуассоновское распределение со средним $\frac{\lambda}{\mu}$.

Система $M|M|n|0$. Процесс $Q(t)$, равный количеству требований в системе $M|M|n|0$ в момент времени t , также является процессом гибели и размножения с параметрами

$$\lambda_i = \lambda; \quad 0 \leq i \leq n-1; \quad \mu_i = i\mu; \quad 1 \leq i \leq n.$$

Эргодическое распределение $p_k = \lim_{t \rightarrow \infty} P\{Q(t) = k\}$ является единственным решением системы уравнений

$$\begin{cases} -\lambda p_0 + \mu p_1 = 0; \\ -(\lambda + i\mu) p_i + \lambda p_{i-1} + (i+1)\mu p_{i+1} = 0, \quad 1 \leq i \leq n-1; \\ -\mu p_n + \lambda p_{n-1} = 0, \end{cases}$$

удовлетворяющим условию нормировки $\sum_{i=0}^n p_i = 1$. Решая эту систему, получим

$$p_i = \frac{\rho^i}{i!} p_0; \quad \rho = \frac{\lambda}{\mu},$$

а

$$p_0 = \left(\sum_{i=0}^n \frac{\rho^i}{i!} \right)^{-1},$$

откуда

$$p_i = \frac{\rho^i}{i!} \left(\sum_{j=0}^n \frac{\rho^j}{j!} \right)^{-1}.$$

Последние формулы носят название формул Эрланга, впервые рассмотревшего систему $M|M|n|0$.

Распределение длительности интервала занятости системы $M|M|1|\infty$. Изложенный метод нахождения распределения числа требований $Q(t)$ в системе может быть использован и для определения других характеристик, например, распределения длительности интервала занятости.

Интервалом занятости будем называть максимальный интервал времени, в течение которого система занята. Он начинается в момент поступления требования в свободную систему и заканчивается, как только система снова освободится от требований. Дальнейшее развитие СМО для нашей задачи не имеет значения, и можно предположить, что после освобождения системы требования в нее не поступают.

Обозначим через $\xi(t)$ количество требований в системе при сделанном допущении спустя время t после начала интервала занятости. Тогда $\xi(t)$ является процессом гибели и размножения с интенсивностями перехода $\lambda_i = \lambda$; $i \geq 1$; $\lambda_0 = 0$; $\mu_i = \mu$; $i \geq 1$ и начальным состоянием $\xi(0) = 1$.

Распределения длительности интервала занятости θ и процесса $\xi(t)$ связаны соотношением

$$F(t) = P\{\theta < t\} = P\{\xi(t) = 0\},$$

поэтому, найдя распределение процесса $\xi(t)$, найдем и распределение величины θ . Вероятности $p_k(t) = P\{\xi(t) = k\}$ удовлетворяют прямой системе уравнений Колмогорова

$$\begin{cases} p'_0(t) = \mu p_1(t); \\ p'_i(t) = -(\lambda + \mu) p_i(t) + \mu p_{i+1}(t); \\ p'_k(t) = -(\lambda + \mu) p_k(t) + \lambda p_{k-1}(t) + \mu p_{k+1}(t), \quad k \geq 2 \end{cases} \quad (2.19)$$

с начальным условием $p_k(0) = \delta_{k1}$.

Для производящей функции $P(z, t) = \sum_{k=1}^{\infty} p_k(t) z^k$ из системы (2.19) следует, что справедливо равенство

$$\frac{\partial P(z, t)}{\partial t} = P(z, t) \left(-\lambda - \mu + \lambda z + \frac{\mu}{z} \right) - \mu p_1(t). \quad (2.20)$$

Введем преобразования Лапласа

$$P^*(z, s) = \int_0^{\infty} e^{-st} P(z, t) dt \quad \text{и} \quad P_1^*(s) = \int_0^{\infty} e^{-st} P_1(t) dt,$$

тогда из равенства (2.20) с учетом начальных условий получим

$$-z + sP^*(z, s) = P^*(z, s) \left(-\lambda - \mu + \lambda z + \frac{\mu}{z} \right) = \mu P_1^*(s)$$

или

$$P^*(z, s) = \frac{z^2 - \mu z P_1^*(s)}{z(\lambda + \mu + s) - \lambda z^2 - \mu}. \quad (2.21)$$

По теореме Руше знаменатель в правой части (2.21) имеет при $\operatorname{Re} s > 0$ корень

$$z_2 = z_2(s) = \frac{\lambda + \mu + s - \sqrt{(\lambda + \mu + s)^2 - 4\lambda\mu}}{2\lambda}$$

в единичном круге $|z| < 1$. Так как $P^*(z_2(s), s)$ существует, то числитель в правой части (2.21) должен обратиться в нуль при $z = z_2(s)$:

$$z_2^2(s) - \mu z_2(s) P_1^*(s) = 0,$$

откуда

$$P_1^*(s) = \frac{z_2(s)}{\mu}. \quad (2.22)$$

Из неиспользованного еще начального уравнения системы (2.19) для преобразования Лапласа $P_0^*(s) = \int_0^{\infty} e^{-st} P_0(t) dt$ получим $sP_0^*(s) = \mu P_1^*(s)$, и поэтому

$$P_0^*(s) = \frac{z_2(s)}{s}. \quad (2.23)$$

Найдем преобразование Лапласа распределения интервала занятости θ , которое равно

$$\begin{aligned} \varphi(s) &= M e^{-s\theta} = \int_0^{\infty} e^{-st} dP\{\theta < t\} = \int_0^{\infty} e^{-st} dp_0(t) = \\ &= sP_0^*(s) = z_2(s), \end{aligned} \quad (2.24)$$

и тем самым решим поставленную задачу.

Из представления (2.24) следует, что при $\lambda \leq \mu$ распределение интервала занятости является собственным.

В самом деле, для этого достаточно убедиться, что при $\lambda \leq \mu \lim_{s \rightarrow 0} \varphi(s) = 1$, а это установить несложно. Если же $\lambda > \mu$, то $\lim_{s \rightarrow 0} \varphi(s) = \frac{\mu}{\lambda} < 1$, и распределение θ несобственно, а именно, с вероятностью $1 - \frac{\mu}{\lambda}$ интервал занятости будет бесконечным. Как видим, условие существования стационарного режима $\rho = \frac{\lambda}{\mu} < 1$ несколько отличается от условия конечности интервала занятости $\frac{\lambda}{\mu} \leq 1$. Вместе с тем, легко проверить, что в стационарном режиме средняя длина интервала занятости конечна, а при $\rho = 1$ она бесконечна:

$$M\theta = -\varphi'(0) = \begin{cases} \frac{1}{\mu - \lambda} < \infty & \text{при } \rho < 1; \\ \infty & \text{при } \rho = 1. \end{cases}$$

Распределение времени ожидания в системе $M|M|1|\infty$. Обозначим через ζ время, необходимое для обслуживания всех требований, находящихся в системе $M|M|1|\infty$ в момент времени t . В силу теоремы 2.1 величину ζ можно интерпретировать как время ожидания требования, если бы оно поступило в момент времени t , а обслуживание проводилось в соответствии с дисциплиной «первым пришел — первым обслужен».

Величину ζ можно представить в виде $\zeta = \sum_{i=1}^N \tau_i$, где N — количество требований в системе; τ_1 — оставшееся время обслуживания требования, которое уже обслуживается, $\tau_2, \dots, \tau_k$ — полные времена обслуживания требований, находящихся в очереди. В силу отсутствия последовательного показательного распределения величины $N, \tau_1, \dots, \tau_2, \dots$ — независимы. Кроме того, τ_i имеют показательное распределение с параметром μ . Допустим, что система находится в стационарном режиме. Тогда величина N имеет геометрическое распределение:

$$P\{N = k\} = (1 - \rho)\rho^k, \quad k \geq 0.$$

Поэтому преобразование Лапласа $\varphi(s)$ величины ζ может быть найдено с помощью таких преобразований:

$$\begin{aligned} \varphi(s) &= M e^{-s\theta} = M e^{-s \sum_{i=1}^N \tau_i} = M M_N e^{-s \sum_{i=1}^N \tau_i} = \\ &= M (M e^{-s\tau_1})^N = M \left(\int_0^{\infty} e^{-st} \mu e^{-\mu t} dt \right)^N = M \left(\frac{\mu}{s + \mu} \right)^N = \end{aligned}$$

$$= \sum_{k=0}^{\infty} (1-\rho) \rho^k \left(\frac{\mu}{s+\mu} \right) = 1-\rho + \rho \left(\frac{\mu(1-\rho)}{s+\mu(1-\rho)} \right) = \\ = (1-\rho) \varphi_1(s) + \rho \varphi_2(s),$$

где

$$\varphi_1(s) = 1; \quad \varphi_2(s) = \frac{\mu(1-\rho)}{s+\mu(1-\rho)}.$$

Функция $\varphi_1(s)$ является преобразованием Лапласа функции распределения $E_0(x) = \begin{cases} 0, & x \leq 0 \\ 1, & x > 0 \end{cases}$ нуля, а $\varphi_2(s)$ — показательного распределения $F(x) = 1 - e^{-\mu(1-\rho)x}$, $x \geq 0$. Поэтому $\varphi(s)$ соответствует функция распределения

$$G(x) = (1-\rho) E_0(x) + \rho F(x) = \begin{cases} 1 - \rho e^{-\mu(1-\rho)x}, & x > 0; \\ 0, & x \leq 0, \end{cases}$$

имеющая скачок в нуле, величина которого равна $1-\rho$ — вероятности того, что система свободна.

В заключение найдем среднее время ожидания

$$M\theta = -\varphi'(0) = \frac{\rho}{\mu(1-\rho)}.$$

Задания для самоконтроля

1. Парикмахерская начинает работу в момент времени $t=0$. Пребывание посетителей описывается простейшим потоком с параметром λ . Каждая стрижка занимает χ секунд. Найти вероятность того, что посетитель, пришедший вторым, не будет ждать, а также его среднее время ожидания в случаях:

а) $\chi = c = \text{const}$;

б) χ имеет показательное распределение с показателем μ .

2. Рассмотрим систему $M|M|n| \infty$, в которой в момент поступления рассматриваемого требования A находится $m+n$ требований. Обслуживание производится в порядке очередности поступления требований. Найти:

а) среднее время ожидания требования A ;

б) среднее время, через которое система обслужит имевшиеся вначале $m+n+1$ требования;

в) вероятность того, что требование A покинет систему k -м из имеющихся $m+n+1$ требований;

г) вероятность того, что обслуживание требования A закончится раньше, чем обслуживание требования, непосредственно ему предшествующего в очереди;

д) распределение времени ожидания требования A .

3. Найти для системы $M|M|1| \infty$ следующие характеристики:

а) вероятность занятости системы;

б) вероятность того, что есть требования в очереди;

в) среднее количество требований в очереди;

г) среднее время ожидания.

4. Найти среднее количество требований в системах $M|M|\infty$, $M|M|1$ и $M|M|n$, $n > 0$.

5. Рассмотрим систему $M|M|1$ с нетерпеливыми требованиями. При поступлении требования оно присоединяется к очереди с вероятностью $e^{-\alpha\omega}$ и покидает ее с вероятностью $1 - e^{-\alpha\omega}$, где $\omega = \frac{k}{\mu}$, а k — количество требований в системе в момент поступления рассматриваемого требования. Найти:

- а) стационарное распределение количества требований в системе;
- б) условия существования стационарного режима;
- в) предел стационарного распределения и среднего количества требований в системе при $\alpha \rightarrow \infty$.

2.2. Система $M|G|1$

Рассмотрим системы с бесконечным количеством мест ожидания и четвертый символ будем опускать.

Наиболее важными характеристиками таких СМО являются распределение числа требований в системе и распределение времени ожидания. Вначале рассмотрим частные случаи системы $M|G|1$, когда время обслуживания постоянно либо имеет распределение Эрланга.

Система $M|D|1$. Не ограничивая общности, предположим, что время обслуживания равно единице. Параметр входного потока, как и раньше, обозначим через λ , количество требований в системе в момент времени t через $Q(t)$ и через $p_k(t)$ — распределение $Q(t)$.

Процесс $Q(t)$ марковским не является, однако специальный вид времени обслуживания позволяет довольно просто найти его распределение в стационарном режиме, когда $p_k t \equiv p_k$, $k = 0, 1, \dots$

Для этого введем события

$$A_k = \{\text{на } [t, t+1) \text{ поступило } k \text{ требований}\},$$

$k \geq 0$ и рассмотрим значения процесса $Q(t)$ в моменты времени t и $t+1$. По формуле полной вероятности

$$\begin{cases} p_0(t+1) = (p_0(t) + p_1(t)) P\{A_0\}; \\ p_k(t+1) = p_0(t) P\{A_k\} + p_1(t) P\{A_k\} + \\ + p_2(t) P\{A_{k-1}\} + \dots + p_{k+1}(t) P\{A_0\}, \quad k \geq 1, \end{cases} \quad (2.25)$$

так как за единицу времени обслуживается ровно одно из уже имеющихся требований. В стационарном режиме $p_k(t) \equiv p_k$. Кроме того, $P\{A_k\} = \frac{\lambda^k}{k!} e^{-\lambda}$. Поэтому стационар-

нарное распределение удовлетворяет системе соотношений

$$p_0 = (p_0 + p_1) e^{-\lambda};$$

$$p_k = \left[(p_0 + p_1) \frac{\lambda^k}{k!} + p_2 \frac{\lambda^{k-1}}{(k-1)!} + \dots + p_{k+1} \right] e^{-\lambda} \quad (2.26)$$

в условии нормировки $\sum_{i=0}^{\infty} p_i = 1$.

Для производящей функции $P(z) = \sum_{i=0}^{\infty} p_i z^i$ из системы (2.26) следует равенство

$$P(z) = e^{-\lambda} [(p_0 + p_1) e^{\lambda z} + p_2 z e^{\lambda z} + p_3 z^2 e^{\lambda z} + \dots] =$$

$$= e^{-\lambda(1-z)} \left(p_0 + \frac{P(z) - p_0}{z} \right),$$

откуда

$$P(z) = p_0 \frac{1-z}{1 - z e^{\lambda(1-z)}}. \quad (2.27)$$

Так как по условию нормировки $\lim_{z \uparrow 1} P(z) = 1$, то, переходя к пределу по правилу Лопиталья в правой части равенства (2.27), получим $1 = \frac{p_0}{1-\lambda}$, откуда $p_0 = 1 - \lambda$.

Разложив функцию $P(z) = \frac{(1-\lambda)(1-\lambda)}{1 - z e^{\lambda(1-z)}}$ в степенной ряд, получим

$$p_0 = 1 - \lambda, \quad p_1 = (1 - \lambda)(e^{\lambda} - 1),$$

$$p_k = (1 - \lambda) \sum_{i=1}^k (-1)^{k-i} e^{i\lambda} \left[\frac{(i\lambda)^{k-i}}{(k-i)!} + \frac{(i\lambda)^{k-i-1}}{(k-i-1)!} \right], \quad k \geq 2.$$

Заметим, что необходимым условием существования стационарного режима, как и в системе $M|M|1| \infty$, является требование, чтобы средняя скорость поступления требований была меньше средней скорости обслуживания. Можно показать, что оно является и достаточным.

Знание производящей функции $P(z)$ позволяет найти и среднее количество требований в системе в стационарном режиме:

$$MQ(t) = P'(t) = \frac{\lambda}{1-\lambda} \left(1 - \frac{\lambda}{2} \right).$$

Полученное значение меньше, чем в системе $M|M|1| \infty$ при тех же интенсивностях поступления требований и обслуживания $\mu = 1$.

Система $M|E_k|1$. В рассматриваемой СМО время обслуживания имеет распределение Эрланга порядка k . Это значит, что время обслуживания можно представить в виде суммы k независимых одинаково распределенных случайных величин, имеющих показательное распределение с параметром $k\mu$:

$$\tau_{\text{обс}} = \tau_1 + \tau_2 + \dots + \tau_k;$$

$$P\{\tau_i \in dx\} = k\mu e^{-k\mu x} dx, \quad x > 0. \quad (2.28)$$

Такой выбор параметра удобен, так как в этом случае $M\tau_{\text{обс}} = kM\tau_1 = k \cdot \frac{1}{k\mu} = \frac{1}{\mu}$, и интенсивность обслуживания будет равна μ .

Нетрудно указать явный вид плотности распределения Эрланга порядка k :

$$P\{\tau_{\text{обс}} \in dx\} = \frac{(\mu k)^k}{\Gamma(k)} e^{-\mu k x} x^{k-1} dx, \quad x \geq 0. \quad (2.29)$$

В самом деле,

$$M e^{-s\tau_{\text{обс}}} = (M e^{-s\tau_1})^k = \left(\frac{k\mu}{k\mu + s} \right)^k.$$

Однако преобразование Лапласа плотности, задаваемой правой частью равенства (2.29), тоже равно $\left(\frac{k\mu}{k\mu + s} \right)^k$:

$$\begin{aligned} \int_0^{\infty} e^{-sx} \frac{(\mu k)^k}{\Gamma(k)} e^{-\mu k x} x^{k-1} dx &= \frac{(\mu k)^k}{\Gamma(k)} \int_0^{\infty} e^{-(s+\mu k)x} x^{k-1} dx = \\ &= \frac{(\mu k)^k}{\Gamma(k)} \frac{1}{(s + \mu k)^k} \int_0^{\infty} e^{-y} y^{k-1} dy = \left(\frac{\mu k}{s + \mu k} \right)^k, \end{aligned}$$

чем и заканчивается проверка равенства (2.29), так как соответствие между распределениями и их преобразованиями Лапласа взаимнооднозначно.

Система $M|E_k|1$ представляет интерес не только потому, что показательное распределение является частным случаем распределения Эрланга. Более важным является то обстоятельство, что плотность распределения Эрланга при $k \geq 2$ имеет колоколообразный вид, а при больших значениях k хорошо аппроксимирует плотность нормального распределения. Плотность же показательного распределения убывает с возрастанием аргумента и по этой причине зачастую не может служить достаточно хорошим приближением плотности распределения времени обслуживания реальных систем.

В силу того что в системе $M|E_k|1$ время обслуживания допускает представление (2.28), предположим, что каждое требование при обслуживании проходит через k этапов, длительности которых равны τ_i .

В этом случае общее количество этапов $v(t)$, которые необходимо пройти всем требованиям, находящимся в системе в момент времени t , будет цепью Маркова (но не процессом гибели и размножения при $k \geq 2$). Найдем ее интенсивности переходов q_{ij} . Как и в рассмотренных ранее системах типа $M|M|1$, $q_{ii-1} = k\mu$ при $i \geq 1$. Далее, каждое поступающее с интенсивностью λ требование увеличивает число этапов на k , поэтому $q_{ii+k} = \lambda$. Во всех остальных случаях

$$q_{ij} = 0, \quad i \neq j, \quad q_{ii} = -\sum_{j \neq i} q_{ij}.$$

Предполагая, что стационарные вероятности $p_j = P\{v(t) = j\}$ существуют, получим, что справедлива система уравнений

$$\begin{cases} -\lambda p_0 + k\mu p_1 = 0; \\ -(\lambda + k\mu) p_i + \lambda p_{i-k} + k\mu p_{i+1} = 0 \quad (i \geq 1), \end{cases} \quad (2.30)$$

где $p_i = 0$ при $i < 0$.

Как обычно, введем производящую функцию $P(z) = \sum_{i=0}^{\infty} p_i z^i$. Тогда из системы (2.30) следует, что

$$-\lambda P(z) - k\mu (P(z) - p_0) + \frac{k\mu}{z} (P(z) - p_0) + \lambda z^k P(z) = 0,$$

откуда

$$P(z) = \frac{k\mu p_0 \left(1 - \frac{1}{z}\right)}{\lambda + k\mu - k\mu \frac{1}{z} - \lambda z^k}. \quad (2.31)$$

Из условия нормировки $\sum_{i=0}^{\infty} p_i = 1$ вытекает, что $\lim_{z \uparrow 1} P(z) = 1$, поэтому, переходя в равенстве (2.31) к пределу при $z \uparrow 1$, получим $1 = \frac{\mu p_0}{\mu - \lambda}$.

Следовательно, $p_0 = \frac{\mu - \lambda}{\mu} = 1 - \rho$, $\rho = \frac{\lambda}{\mu}$, и

$$P(z) = \frac{k\mu (1 - \rho) (1 - z)}{k\mu + \lambda z^{k+1} - (\lambda + k\mu) z} \quad (2.32)$$

(заметим, что необходимым условием существования стационарного режима будет $\frac{\lambda}{\mu} < 1$).

Для нахождения p_i необходимо разложить $P(z)$ в ряд по степеням z . Для этого заметим, что знаменатель в (2.32) является полиномом степени $k+1$, один из корней которого равен 1, и поэтому представим в виде

$$k\mu + \lambda z^{k+1} - (\lambda + k\mu)z = (1-z)[k\mu - \lambda(z+z^2+\dots+z^k)].$$

Далее, при $\lambda < \mu$ корни $z_1, z_2, \dots, z_k$ сомножителя $k\mu - \lambda(z+z^2+\dots+z^k)$ находятся вне единичного круга, $|z_k| > 1$, и являются простыми. Следовательно, знаменатель пускает представление

$$k\mu(1-z)(1-z/z_1)\dots(1-z/z_k)$$

и

$$P(z) = \frac{1-\rho}{\left(1-\frac{z}{z_1}\right)\dots\left(1-\frac{z}{z_k}\right)} = (1-\rho) \sum_{r=1}^k \frac{a_r}{1-\frac{z}{z_r}},$$

где

$$a_i = \prod_{j \neq i} \frac{1}{1-z_i/z_j}.$$

Поэтому $P(z)$ разлагается в ряд

$$P(z) = (1-\rho) \sum_{i=0}^{\infty} z^i \sum_{r=1}^k \frac{a_r}{z_r^i},$$

и

$$P_i = (1-\rho) \sum_{r=1}^k \frac{a_r}{z_r^i}, \quad i \geq 0.$$

Для нахождения стационарной вероятности p_i того, что в системе находится i требований, теперь достаточно воспользоваться равенством

$$p_i = \sum_{j=(i-1)k+1}^{ik} p_j.$$

Распределение числа требований в системе $M|G|1$. Пусть, как и ранее, $Q(t)$ — количество требований в системе $M|G|1$ в момент времени t . Если распределение времени обслуживания отлично от показательного, то $Q(t)$ не будет цепью Маркова. Вместе с тем можно указать

такие последовательности $\{\theta_n\}$, $n \geq 1$ случайных моментов времени, что значения $v_n = Q(\theta_n)$ будут образовывать цепь Маркова с дискретным временем. В качестве одной из них может выступать последовательность моментов времени, в которые заканчивается обслуживание n -го требования. В самом деле, если значение v_n известно: $v_n = i$, то это значит, что в системе осталось i требований, обслуживание которых еще не началось. Поэтому дальнейшее развитие последовательности v_m при $m > n$ зависит от длительности времени обслуживания имеющихся i требований, моментов поступления новых требований и длительности времени их обслуживания. Так как входящий поток является простейшим, то эти факторы не зависят от значений последовательности v_k при $k < i$. Итак, v_n — цепь Маркова, которую назовем цепью, вложенной в процесс $Q(t)$.

Обозначим $p_{ij}(t) = P\{Q(t) = j / v(0) = i\}$.

Теорема 2.2. Если цепь v_n имеет эргодическое распределение

$$\pi_k = \lim_{n \rightarrow \infty} P\{v = k / v_0 = i\}, \text{ то } \lim_{t \rightarrow 0} p_{ik}(t) = \pi_k.$$

Доказательство. Зафиксируем начальное состояние i и введем обозначения: t_n — момент прихода n -го требования (без учета начальных требований); θ_n — момент обслуживания n -го требования (с учетом начальных требований),

$$v_n = Q(\theta_n), \mu_n = Q(t_n - 0).$$

Покажем сначала, что $\lim_{n \rightarrow \infty} P\{v_n = k\} = \lim_{n \rightarrow \infty} P\{\mu_n = k\}$.

Для этого введем события

$$A_1 = \{v_{n+i} \leq k\};$$

$A_2 = \{\text{на интервале } (0, \theta_{n+i}) \text{ поступило не более } n+k \text{ требований, из которых } n \text{ обслужено}\};$

$$A_3 = \{\mu_{n+k+1} \leq k\};$$

$A_4 = \{\text{на интервале } (0, t_{n+k+1}) \text{ обслужено не менее } i+n \text{ требований}\}.$

Тогда $A_1 \subset A_2 \subset A_3$ и $A_3 \subset A_4 \subset A_1$.

Следовательно, $A_1 = A_3$ и

$$P\{v_{n+i} \leq k\} = P\{\mu_{n+k+1} \leq k\}.$$

При выполнении условий теоремы

$$P\{v_{n+i} \leq k\} \xrightarrow{n \rightarrow \infty} \sum_{j=0}^k \pi_j,$$

поэтому $P\{\mu_{n+k+1} \leq k\} \xrightarrow{n \rightarrow \infty} \sum_{j=0}^k \pi_j$, что эквивалентно сходимости

$$P\{\mu_n = k\} \xrightarrow{n \rightarrow \infty} \pi_k.$$

Далее, по теореме 1.1,

$$P\{Q(t) = k\} = R_k(t) = \lim_{\Delta \rightarrow 0} P\{Q(t) = k/A(t, \Delta)\},$$

где $A(t, \Delta) = \{\text{на интервале } [t, t + \Delta) \text{ поступило требование}\}$. Поэтому для завершения доказательства достаточно убедиться, что

$$\lim_{t \rightarrow \infty} R_k(t) = \lim_{n \rightarrow \infty} P\{\mu_n = k\}.$$

Имеем

$$\begin{aligned} R_k(t) &= \lim_{\Delta \rightarrow 0} P\{Q(t) = k/A(t, \Delta)\} = \lim_{\Delta \rightarrow 0} \sum_{n=1}^{\infty} P\{Q(t) = k, \\ &\quad t_n \in [t, t + \Delta)/A(t, \Delta)\} = \\ &= \lim_{\Delta \rightarrow 0} \frac{\sum_{n=1}^{\infty} P\{Q(t) = k, t_n \in [t, t + \Delta), A(t, \Delta)\}}{P\{A(t, \Delta)\}} = \\ &= \lim_{\Delta \rightarrow 0} \frac{1}{\lambda \Delta} \sum_{n=1}^{\infty} P\{\mu_n = k\} e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} \lambda \Delta = \\ &= e^{-\lambda t} \sum_{n=1}^{\infty} P\{\mu_n = k\} \frac{(\lambda t)^{n-1}}{(n-1)!}. \end{aligned}$$

Так как $P\{\mu_n = k\} \rightarrow \pi_k$, то при $n > N(\varepsilon)$ $|P\{\mu_n = k\} - \pi_k| < \varepsilon$ для любого наперед заданного $\varepsilon > 0$ и

$$\begin{aligned} e^{-\lambda t} (\pi_k - \varepsilon) \sum_{n=N}^{\infty} \frac{(\lambda t)^{n-1}}{(n-1)!} &\leq R_k(t) \leq e^{-\lambda t} \times \\ \times \left(\sum_{n=1}^{N-1} P\{\mu_n = k\} \frac{(\lambda t)^{n-1}}{(n-1)!} + (\pi_k + \varepsilon) \sum_{n=N}^{\infty} \frac{(\lambda t)^{n-1}}{(n-1)!} \right). \end{aligned}$$

Из последнего неравенства следует, что

$$\pi_k - \varepsilon \leq \lim_{t \rightarrow \infty} R_k(t) \leq \overline{\lim}_{t \rightarrow \infty} R_k(t) \leq \pi_k + \varepsilon.$$

Ввиду произвола в выборе $\epsilon > 0$:

$$\lim_{t \rightarrow \infty} R_k(t) = \pi_k,$$

чем и заканчивается доказательство теоремы.

Поскольку цепь v_n образует один класс неперiodических состояний, то вероятности π_k являются единственным решением системы уравнений

$$\pi_k = \sum_i \pi_i p_{ik}, \quad k \geq 0, \quad \sum_i \pi_i = 1, \quad (2.33)$$

где p_{ik} — вероятности перехода за один шаг цепи v_n .

Найдем вероятности p_{ik} .

Рассмотрим сначала случай $i \geq 1$. Тогда $p_{ik} = P$ {за время обслуживания поступило $k - i + 1$ требование} =

$$= \int_0^{\infty} e^{-\lambda t} \frac{(\lambda t)^{k-i+1}}{(k-i+1)!} dG(t), \quad k \geq i-1; \quad p_{ik} = 0 \text{ при } k < i-1.$$

Если $i = 0$, то

$p_{0k} = P$ {после поступления очередного требования за время его обслуживания поступит k требований} =

$$= p_{1k} = \int_0^{\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!} dG(t).$$

Для решения системы (2.33) введем производящую функцию

$$P(z) = \sum_{k=0}^{\infty} \pi_k z^k.$$

Умножив k -е уравнение системы на z^k и просуммировав по k , получим

$$\begin{aligned} P(z) &= \sum_{k=0}^{\infty} z^k \sum_{i=0}^{\infty} \pi_i p_{ik} = \pi_0 \sum_{k=0}^{\infty} z^k p_{0k} + \sum_{i=1}^{\infty} \pi_i \sum_{k=0}^{\infty} z^k p_{ik} = \\ &= \pi_0 \sum_{k=0}^{\infty} z^k \int_0^{\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!} dG(t) + \sum_{i=1}^{\infty} \pi_i \sum_{k=i-1}^{\infty} z^k \int_0^{\infty} e^{-\lambda t} \times \\ &\quad \times \frac{(\lambda t)^{k-i+1}}{(k-i+1)!} dG(t) = \pi_0 \int_0^{\infty} e^{-\lambda t} e^{\lambda t z} dG(t) + \\ &\quad + \sum_{i=1}^{\infty} \pi_i z^{i-1} \int_0^{\infty} e^{-\lambda t} e^{\lambda t z} dG(t) = \int_0^{\infty} e^{-\lambda t(1-z)} dG(t) \times \\ &\quad \times \left(\pi_0 + \sum_{i=1}^{\infty} \pi_i z^{i-1} \right) = G^*(\lambda(1-z)) \left(\frac{P(z) - \pi_0}{z} + \pi_0 \right), \end{aligned}$$

где

$$G^*(\lambda) = \int_0^{\infty} e^{-\lambda t} dG(t).$$

Решая последнее уравнение относительно $P(z)$, найдем

$$P(z) = \frac{G^*(\lambda(1-z))(1-z)\pi_0}{G^*(\lambda(1-z)) - z}. \quad (2.34)$$

Из условий нормировки следует, что $\lim_{z \rightarrow 1-0} P(z) = 1$. Поэтому, переходя в равенстве (2.34) к пределу при $z \uparrow 1$ и раскрывая получающуюся неопределенность по правилу Лопиталя, получим

$$1 = \frac{\pi_0}{1 + \lambda G^{*'}(0)} = \frac{\pi_0}{1 - \int_0^{\infty} t dG(t)} = \frac{\pi_0}{1 - \lambda m},$$

где m — среднее время обслуживания.

Отсюда следует, что

$$\pi_0 = 1 - \lambda m$$

и

$$P(z) = \frac{G^*(\lambda(1-z))(1-z)(1-\lambda m)}{G^*(\lambda(1-z)) - z}.$$

Заметим, что необходимым условием существования эргодического распределения будет $1 - \lambda m > 0$ или $\lambda < \frac{1}{m}$. Можно показать, что оно является и достаточным условием.

Полученная формула для $P(z)$ позволяет, в частности, найти среднее количество $MQ(t)$ требований в системе $M|G|1$, находящейся в стационарном режиме:

$$MQ(t) = P'(z)|_{z=1}.$$

Представим $P'(z)$ в виде

$$P'(z) = [G^*(\lambda(1-z))]' \frac{(1-z)}{G^*(\lambda(1-z)) - z} (1 - \lambda m) + \\ + G^*(\lambda(1-z)) \left[\frac{1-z}{G^*(\lambda(1-z)) - z} \right]' (1 - \lambda m)$$

и воспользуемся такими легко проверяемыми равенствами

$$\lim_{z \rightarrow 1-0} [G^*(\lambda(1-z))]' = \lambda m;$$

$$\lim_{z \rightarrow 1-0} \frac{1-z}{G^*(\lambda(1-z)) - z} = \lim_{z \rightarrow 1-0} \frac{-1}{[G^*(\lambda(1-z))]' - 1} = \frac{1}{1 - \lambda m};$$

$$\begin{aligned} \lim_{z \rightarrow 1-0} \left[\frac{1-z}{G^*(\lambda(1-z)) - z} \right]' &= \lim_{z \rightarrow 1-0} \frac{z - G^*(\lambda(1-z)) - (1-z)}{(G^*(\lambda(1-z)) - z)^2} \times \\ &\times ([G^*(\lambda(1-z))] - 1) = \lim_{z \rightarrow 1-0} \frac{-(1-z) [G^*(\lambda(1-z))]''}{2(G^*(\lambda(1-z)) - z)} \times \\ &\times \frac{1}{([G^*(\lambda(1-z))] - 1)} = \lim_{z \rightarrow 1-0} \frac{\lambda^2 G^{*''}(0)}{2(1-\lambda m)} \times \\ &\times \frac{(1-z)}{G^*(\lambda(1-z)) - z} = \frac{\lambda^2 m_2}{2(1-\lambda m)} \lim_{z \rightarrow 1-0} \frac{1}{1 - [G^*(\lambda(1-z))]'} = \\ &= \frac{\lambda^2 m_2}{2(1-\lambda m)^2}, \end{aligned}$$

где $m_2 = G^{*''}(0) = \int_0^\infty t^2 dG(t)$.

Поэтому

$$MQ(t) = \lim_{z \rightarrow 1-0} P'(z) = \lambda m + \frac{\lambda^2 m_2}{2(1-\lambda m)}.$$

Полученное равенство носит название формулы Поллачека-Хинчина. Иногда формулой Поллачека-Хинчина называют также полученное нами выражение для производящей функции числа требований:

$$P(z) = \frac{G^*(\lambda(1-z))(1-z)(1-\lambda m)}{G^*(\lambda(1-z)) - z}. \quad (2.35)$$

Распределение времени пребывания в системе $M|G|1$. Пусть начальное количество требований в системе равно i . Обозначим через t_n момент поступления n -го требования в систему $M|G|1$; $s_n = \theta_{n+i}$ — момент завершения его обслуживания. Тогда $x_n = s_n - t_n$ — время пребывания n -го требования в системе.

Найдем предел преобразования Лапласа распределения x_n :

$$\varphi(s) = \lim_{n \rightarrow \infty} M e^{-s x_n}.$$

Для этого сделаем такие замечания:

1) в момент s_n ухода рассматриваемого требования из системы оставшееся количество требований, как следует из п. 2.3, имеет в пределе при $n \rightarrow \infty$ производящую функцию $P(z)$;

2) так как обслуживание требований производится в порядке очередности их поступления, то в момент s_n в системе остаются те и только те требования, которые поступили за время пребывания x_n n -го требования.

Из сделанных замечаний следует, что

$$P(z) = \lim_{n \rightarrow \infty} Mz^{\alpha(s_n) - \alpha(t_n)},$$

где $\alpha(t)$ — количество требований, поступивших на интервале $(0, t)$.

Так как $\alpha(t)$ — пуассоновский процесс с параметром λ и не зависит от значений концов интервала (s_n, t_n) , то

$$\begin{aligned} Mz^{\alpha(s_n) - \alpha(t_n)} &= Mz^{\alpha(s_n - t_n)} = Mz^{\alpha(x_n)} = \\ &= MM_{x_n} z^{\alpha(x_n)} = M(Mz^{\alpha(1)})^{x_n} = \\ &= M(e^{-\lambda + \lambda z})^{x_n} = Me^{-(\lambda - \lambda z)x_n} \xrightarrow{n \rightarrow \infty} \varphi(\lambda - \lambda z). \end{aligned}$$

Итак, $\varphi(\lambda - \lambda z) = P(z)$, откуда следует, что

$$\varphi(s) = P\left(1 - \frac{s}{\lambda}\right).$$

В результате получим

$$\varphi(s) = \frac{sG^*(s)(1 - \lambda m)}{\lambda G^*(s) - \lambda + s}. \quad (2.36)$$

Полученную формулу называют формулой Хинчина, Формула Хинчина позволяет найти и предел $\varphi_1(s)$ преобразования Лапласа времени ожидания y_n начала обслуживания n -м требованием при $n \rightarrow \infty$. В самом деле, $x_n = y_n + \tau_n$, где y_n и τ_n независимы,

$$Me^{-sx_n} \rightarrow \varphi(s), \quad Me^{-sy_n} \rightarrow \varphi_1(s), \quad Me^{-s\tau_n} = G^*(s).$$

Поэтому $\varphi(s) = \varphi_1(s)G^*(s)$ и

$$\varphi_1(s) = \frac{s(1 - \lambda m)}{\lambda G^*(s) - \lambda + s}.$$

Распределение периода занятости системы $M|G|1$. Период занятости — интервал, начинающийся с момента поступления некоторого требования в свободную систему и заканчивающийся в ближайший следующий момент времени, когда система снова освободится. Обозначим его длительность через γ , а преобразование Лапласа через $\Gamma(s)$: $\Gamma(s) = Me^{-s\gamma}$.

Теорема 2.3. Функция $\Gamma(s)$ при $s > 0$ является единственным решением функционального уравнения

$$\Gamma(s) = G^*(s + \lambda - \lambda\Gamma(s)),$$

удовлетворяющим условию $0 \leq \Gamma(s) \leq 1$.

Доказательство. Заметим, что длительность γ периода занятости не зависит от порядка обслуживания

требований. Поэтому можно считать, что обслуживание производится по принципу «последним прибыл — первым обслужен».

За время обслуживания τ_1 требования, с момента поступления которого начался период занятости, в систему поступает $\alpha(\tau_1)$ требований. В силу сделанного предположения о порядке выбора на обслуживание первым из них начнет обслуживаться требование с номером $\alpha(\tau_1)$. Требование с номером $\alpha(\tau_1) - 1$ (если такое существует) начнет обслуживаться лишь после того, как обслужатся требования с номером $\alpha(\tau_1)$ и все требования, которые поступят в систему за время его обслуживания. Таким образом, с требованием номером $\alpha(\tau_1)$ оказывается связанным интервал времени, длительность которого γ_1 имеет то же распределение, что и γ . Подобные интервалы можно связать с каждым из $\alpha(\tau_1)$ требований.

Обозначим их длины через $\gamma_1, \gamma_2, \dots, \gamma_{\alpha(\tau_1)}$. Тогда

$$\psi = \tau_1 + \sum_{i=1}^{\alpha(\tau_1)} \gamma_i,$$

причем случайные величины $\tau_1, \gamma_1, \gamma_2, \dots$ независимы,

$$Me^{-s\tau_1} = G^*(s), \quad Me^{-s\gamma_i} = \Gamma(s).$$

Поэтому

$$\begin{aligned} \Gamma(s) &= Me^{-s\gamma} = Me^{-s(\tau_1 + \sum_{i=1}^{\alpha(\tau_1)} \gamma_i)} = MM_{\tau_1, \alpha(\tau_1)} e^{-s(\tau_1 + \sum_{i=1}^{\alpha(\tau_1)} \gamma_i)} = \\ &= Me^{-s\tau_1} (Me^{-s\gamma_i})^{\alpha(\tau_1)} = Me^{-s\tau_1} (\Gamma(s))^{\alpha(\tau_1)} = \\ &= Me^{-s\tau_1} [e^{-\lambda + \lambda\Gamma(s)}]^{\alpha(\tau_1)} = Me^{-\tau_1(s + \lambda - \lambda\Gamma(s))} = G^*(s + \lambda - \lambda\Gamma(s)). \end{aligned} \quad (2.37)$$

Покажем теперь единственность решения полученного функционального уравнения.

Обозначим $x = s + \lambda - \lambda\Gamma(s)$, $f(x) = \frac{-x + s + \lambda}{\lambda}$. Тогда уравнение переписывается в виде $f(x) = G^*(x)$. Поскольку $0 \leq \Gamma(s) \leq 1$, то $s \leq x \leq \lambda + s$. Далее, так как $G^*(s) < 1$ при $s > 0$, $G^*(\lambda + s) > 0$, то $f(s + \lambda) < G^*(s)$, $f(s) < G^*(s)$. Учитывая, что функция $G^*(s)$ является преобразованием Лапласа и поэтому выпукла, а $f(x)$ линейна, получим, что уравнение $f(x) = G^*(x)$ имеет единственное решение на интервале $s \leq x \leq s + \lambda$, чем и заканчивается доказательство теоремы.

Выясним условия, при которых случайная величина γ конечна с вероятностью 1. Известно, что для этого необходимо и достаточно, чтобы $\Gamma(0+) = 1$.

Допустим, что $m = M\tau_1 = -G^*(0) > \frac{1}{\lambda}$. Так как функция $G^*(x)$ выпукла, $G^*(0) = 1$ и $G^*(x) > 0$, то уравнение $G^*(x) = 1 - \frac{x}{\lambda}$ имеет корни $x_0 = 0$ и $x_1 > 0$. Кроме того, корень уравнения $G^*(x) = f(x) = 1 + \frac{s}{\lambda} - \frac{x}{\lambda}$ при $s > 0$ должен удовлетворять условиям $x > x_1$, $x \rightarrow x_1$ при $s \rightarrow 0+$. Поэтому

$$\lim_{s \rightarrow 0+} \Gamma(s) = \lim_{s \rightarrow 0+} \left(1 + \frac{s}{\lambda} - \frac{x}{\lambda} \right) = 1 - \frac{x}{\lambda} < 1.$$

Аналогичные рассуждения показывают, что при $M\tau_1 \leq \frac{1}{\lambda}$ уравнение $G^*(x) = 1 - \frac{x}{\lambda}$ имеет один корень $x_0 = 0$ и $\lim_{s \rightarrow 0+} \Gamma(s) = 1$.

В заключение найдем среднюю длину периода занятости при условии $\lambda m \leq 1$.

Дифференцируя функциональное уравнение для $\Gamma(s)$ при $s > 0$, получим

$$\Gamma'(s) = G^{*'}(s + \lambda - \lambda\Gamma(s))(1 - \lambda\Gamma'(s))$$

или

$$\Gamma'(s) = \frac{G^{*'}(s + \lambda - \lambda\Gamma(s))}{1 + \lambda G^{*'}(s + \lambda - \lambda\Gamma(s))}.$$

Отсюда

$$M\gamma = \lim_{s \rightarrow 0+} \Gamma'(s) = \begin{cases} \frac{m}{1 - \lambda m} & \text{при } \lambda m < 1; \\ \infty & \text{при } \lambda m = 1. \end{cases}$$

Задания для самоконтроля

1. Найти распределение времени пребывания в системе $M|E_2|1$.
2. Рассмотреть систему $M|G|1$, в которой $1 - G(x) = \frac{1}{2} e^{-\rho x}$, $x > 0$.

Найти:

- а) среднее время и дисперсию времени обслуживания;
 - б) среднее время ожидания;
 - в) распределение времени ожидания.
3. Показать, что для системы $M|G|\infty$

$$\lim_{t \rightarrow \infty} p_k(t) = \frac{(\lambda m)^k}{k!} e^{-\lambda m},$$

где m — среднее время обслуживания.

2.3. Система $G|M|1$

Обозначим через $F(t)$ функцию распределения длины интервала между поступлениями требований входящего потока, через μ — параметр показательного распределения времени обслуживания, а через $Q(t)$ — количество требований в системе в момент времени t . Как и в случае системы $M|G|1$, процесс $Q(t)$ марковским не является. Для нахождения некоторых его важных характеристик воспользуемся методом вложенных цепей Маркова.

В отличие от системы $M|G|1$, цепь Маркова v_n , вложенная в процесс $Q(t)$, определится как $v_n = Q(t_n - 0)$, где t_n — момент поступления n -го требования. Таким образом, v_n равно количеству требований, находящихся в системе непосредственно перед приходом n -го требования.

Марковость последовательности v_n обусловлена тем, что факторы, влияющие на развитие процесса $Q(t)$ при $t \geq t_n$, не зависят от значений $Q(t)$ при $t < t_n$ при условии, что $Q(t_n - 0)$ известно. В самом деле, к этим факторам относятся:

а) длительности интервалов между поступлениями следующих требований;

б) оставшееся время обслуживания;

в) времена обслуживания требований, находящихся в системе в момент времени t_n либо поступающих после t_n , и в силу определения системы $G|M|1$ и свойств показательного распределения марковость последовательности v_n проверена.

Займемся нахождением вероятностей перехода за один шаг цепи v_n :

$$p_{ij} = P\{v_{n+1} = j / v_n = i\}.$$

Рассмотрим такие случаи:

а) $j > i + 1$. Тогда $p_{ij} \leq P\{\text{на интервале } [t_n, t_{n+1}] \text{ поступило не менее двух требований}\} = 0$, и, следовательно, $p_{ij} = 0$;

б) $i \geq 1$; $1 \leq j \leq i + 1$. В этом случае $p_{ij} = P\{\text{на } [t_n, t_{n+1}] \text{ поступило одно требование, обслужено } i + 1 - j \text{ требований}\} = P\{\text{на интервале } [t_n, t_{n+1}] \text{ поступило } i + 1 - j \text{ требований простейшего потока с параметром } \mu\}$, так как при указанных условиях обслуживающее устройство все время занято, и длины интервалов между моментами окончания обслуживания требований независимы и имеют показательное распределение с параметром μ .

Поэтому при $i \geq 1, 1 \leq j \leq i + 1$

$$p_{ij} = \int_0^{\infty} e^{-\mu t} \frac{(\mu t)^{i+1-j}}{(i+1-j)!} dF(t);$$

в) $j = 0$. Так же, как и в случае б), имеем $p_{ij} = P\{\text{на } [t_n, t_{n+1}) \text{ поступило одно требование, обслужены все } i+1 \text{ требования} \mid v(t_n - 0) = i\} = P\{i+1\}$ -я точка простейшего потока с параметром μ лежит левее $t_{n+1} - t_n$.

Поэтому

$$p_{i0} = \int_0^{\infty} G_{i+1}(t) dF(t),$$

где $G_i(t) = 1 - e^{-\mu t} \left(1 + \frac{\mu t}{1!} + \dots + \frac{(\mu t)^{i-1}}{(i-1)!} \right)$ — функция распределения i -й точки потока;

г) $i = 0, j = 1$. В этом случае $p_{01} = P\{\text{на } t_n; t_{n+1}) \text{ имеющееся требование не обслужено}\} = \int_0^{\infty} e^{-\mu t} dF(t)$. Цепь v_n непериодична и образует один класс существенных состояний, а ее стационарное распределение $\{\pi_k, k \geq 0\}$ удовлетворяет системе уравнений

$$\begin{cases} \sum_i \pi_i = 1; \\ \pi_k = \sum_i \pi_i p_{ik}, \quad k \geq 0. \end{cases} \quad (2.38)$$

Будем искать ее решение в виде $\pi_k = cx^k$. В этом случае из системы уравнений (2.38) следует, что при $k \geq 1$

$$\begin{aligned} cx^k &= \sum_{i=0}^{\infty} cx^i p_{ik} = c \sum_{i=k-1}^{\infty} x^i \int_0^{\infty} e^{-\mu t} \frac{(\mu t)^{i+1-k}}{(i+1-k)!} dF(t) = \\ &= cx^{k-1} \int_0^{\infty} e^{-\mu t(1-x)} dF(t). \end{aligned}$$

Поэтому x должно являться решением уравнения

$$x = F^*(\mu(1-x)), \quad \text{где } F^*(s) = \int_0^{\infty} e^{-st} dF(t),$$

лежащим в интервале $(0,1)$.

Лемма 2.1. Уравнение $x = F^*(\mu(1-x))$ имеет решение x_1 в интервале $(0,1)$ тогда и только тогда, когда $F^{*'}(0+) = \int_0^\infty t dF(t) = \frac{1}{\lambda} > \frac{1}{\mu}$. Это решение является единственным в интервале $(0,1)$.

Доказательство. Обозначим $z = \mu(1-x)$, $f(z) = \frac{\mu-z}{\mu}$. Тогда $0 < z < \mu$, $f(z) = F^*(z)$.

Так как функция $f(z)$ линейна, $f(0) = 1$, $f(\mu) = 0$, $F^*(z)$ выпукла на интервале $[0, \infty)$, $F^*(0) = 1$, $F^*(\mu) > 0$,

$$F^{*'}(0) = -\frac{1}{\lambda} > f'(0) = -\frac{1}{\mu},$$

то уравнение $f(z) = F^*(z)$ имеет два корня: $z_0 = 0$ и z_1 , $0 < z_1 < \mu$, что эквивалентно утверждению леммы.

Осталось проверить, что начальное уравнение системы $\pi_0 = \sum_{i=0}^\infty \pi_i p_{i0}$ также превращается в тождество при $\pi_i = cx_i^i$.
Имеем

$$\begin{aligned} \sum_{i=0}^\infty cx_i^i p_{i0} &= c \sum_{i=0}^\infty x^i \int_0^\infty G_{i+1}(t) dF(t) = \\ &= c \int_0^\infty \sum_{i=0}^\infty x^i \left(1 - \sum_{k=0}^i e^{-\mu t} \frac{(\mu t)^k}{k!}\right) dF(t) = \\ &= c \left(\frac{1}{1-x} - \int_0^\infty \sum_{i=0}^\infty \sum_{k=0}^i x^i e^{-\mu t} \frac{(\mu t)^k}{k!} dF(t) \right) = \\ &= c \left(\frac{1}{1-x} - \int_0^\infty e^{-\mu t} \sum_{k=0}^\infty \frac{(\mu t)^k}{k!} x^k \sum_{i=k}^\infty x^{i-k} dF(t) \right) = \\ &= c \frac{1}{1-x} \left(1 - \int_0^\infty e^{-\mu t(1-x)} dF(t) \right) = \\ &= c \frac{1}{1-x} (1 - F^*(\mu(1-x))) = c \end{aligned}$$

при $x = x_1$. Константа c может быть легко определена из условия нормировки. Окончательно получим, что $\pi_k = (1-x_1)x_1^k$, где x_1 — единственное в интервале $(0,1)$ решение уравнения $x = F^*(\mu(1-x))$ (при условии $\lambda < \mu$)

Итак, распределение числа требований в системе $G|M|1$ в момент прихода очередного требования в стационарном случае является геометрическим.

Найдем теперь распределение полного времени пребывания требования в системе в стационарном режиме. Введем обозначение: N — количество требований в системе в момент прихода рассматриваемого требования, $\tau_1, \dots, \tau_N$ — времена обслуживания имеющихся требований, τ_{N+1} — время обслуживания вновь поступившего требования, ξ — время пребывания требования в системе.

Тогда $\xi = \sum_{i=1}^{N+1} \tau_i$. Так как величины $N, \tau_1, \tau_2, \dots$ независимы, $Me^{-s\xi} = \mu/(\mu + s)$, то в стационарном режиме

$$\begin{aligned} Me^{-s\xi} &= Me^{-s \sum_{i=1}^{N+1} \tau_i} = M M_N e^{-s \sum_{i=1}^{N+1} \tau_i} = \\ &= M (Me^{-s\tau_1})^{N+1} = M \left(\frac{\mu}{\mu + s} \right)^{N+1} = \sum_{k=0}^{\infty} \pi_k \left(\frac{\mu}{\mu + s} \right)^{k+1} = \\ &= \frac{(1 - x_1) \mu}{s + (1 - x_1) \mu}. \end{aligned}$$

Поэтому $P\{\xi > t\} = e^{-(1-x_1)\mu t}$, $t > 0$, и время пребывания требования в системе в стационарном режиме имеет показательное распределение.

В заключение найдем распределение времени ожидания γ начала обслуживания в стационарном режиме.

Так как $\gamma = \sum_{i=1}^N \tau_i$, то

$$Me^{-s\gamma} = Me^{-s\tau_1} Me^{-s\tau_1} = Me^{-s\tau_1} \frac{\mu}{\mu + s}.$$

Следовательно,

$$\begin{aligned} Me^{-s\gamma} &= \frac{\mu + s}{\mu} \frac{(1 - x_1) \mu}{s + (1 - x_1) \mu} = \\ &= (1 - x_1) + x_1 \frac{(1 - x_1) \mu}{s + (1 - x_1) \mu} = (1 - x_1) \psi_1(s) + x_1 \psi_2(s), \end{aligned}$$

где $\psi_1(s) = 1 = Me^{-s \cdot 0}$, $\psi_2(s) = \frac{(1 - x_1) \mu}{s + (1 - x_1) \mu}$ — преобразование Лапласа показательного распределения с параметром $(1 - x_1) \mu$.

Поэтому функция распределения $H(x) = P\{\gamma < x\}$ равна $H(x) = (1 - x_1) + x_1 (1 - e^{-(1-x_1)\mu x})$, $x > 0$.

2.4. Система GI|G|1

Изучение системы GI|G|1 затрудняется тем, что наиболее интересные случайные процессы, довольно просто характеризующие ее, оказываются немарковскими. Более того, метод вложенных цепей Маркова, как правило, также неприменим.

Выведем уравнение для распределения времени ожидания начала обслуживания в системе GI|G|1.

Пусть $F(t)$ — функция распределения длины интервала между поступлениями требований; $G(t)$ — функция распределения времени обслуживания; t_n — момент поступления n -го требования; τ_n — время его обслуживания; w_n — время его ожидания; s_n — момент его ухода из системы.

Очевидно, что $w_{n+1} = 0$ при $t_{n+1} \geq s_n$ и $w_{n+1} = s_n - t_{n+1}$ при $t_{n+1} < s_n$, или, что то же самое,

$$w_{n+1} = \max(s_n - t_{n+1}, 0).$$

Так как $s_n = t_n + w_n + \tau_n$, то

$$w_{n+1} = \max(w_n + \tau_n - (t_{n+1} - t_n), 0). \quad (2.39)$$

Фигурирующая в последнем равенстве случайная величина $u_n = \tau_n - (t_{n+1} - t_n)$ не зависит от w_n и имеет функцию

распределения $U(t) = P\{u_n < t\} = \int_0^\infty G(s+t) dF(s)$. Кроме

того, величины u_n независимы в совокупности.

Обозначим $P_n(x) = P\{w_n < x\}$ и покажем, что существует $\lim_{n \rightarrow \infty} P_n(x) = P(x)$.

Имеем $P_1(x) = 1$ при $x > 0$, так как первое требование не ожидает и $w_1 = 0$. Далее, при $x > 0$

$$\begin{aligned} P_2(x) &= P\{\max(u_1, 0) < x\} \leq P\{u_1 < x\} \leq P_1(x); \\ P_3(x) &= P\{\max(w_2 + u_2, 0) < x\} \leq P\{w_2 + u_2 < x\} = \\ &= P\{\max(w_1 + u_1, 0) + u_2 < x\} \leq \\ &\leq P\{\max(u_1, 0) + u_2 < x\} = P\{u_1 + u_2 < x, u_2 < x\} = \\ &= P\{u_1 + u_2 < x, u_1 < x\} \leq P_2(x); \\ &\dots \dots \dots \\ P_{n+1}(x) &= P\{\max(w_n + u_n, 0) < x\} \leq P\{w_n + u_n < x\} = \\ &= P\{\max(w_{n-1} + u_{n-1}, 0) + u_n < x\} = \\ &= P\{w_{n-1} + u_{n-1} + u_n < x, u_n < x\} = \dots = \\ &= P\{u_1 + \dots + u_n < x, u_2 + \dots + u_n < x, \dots, u_{n-1} + u_n < x, \\ &u_n < x\} \leq P\{u_2 + \dots + u_n < x, \dots, u_{n-1} + u_n < x, \\ &u_n < x\} = P\{u_1 + \dots + u_{n-1} < x, \dots; \\ &u_{n-2} + u_{n-1} < x, u_{n-1} < x\} = P_n(x). \end{aligned}$$

При $x \leq 0$, $P_n(x) = 0$. Поэтому последовательность $P_n(x)$ не возрастает по n и, поскольку $P_n(x) \geq 0$, имеет предел $\lim_{n \rightarrow \infty} P_n(x) = P(x)$. Очевидно, что $P(x) = 0$ при $x \leq 0$.

Укажем уравнение, которому удовлетворяет функция $P(x)$.

Так как при $x > 0$

$$\begin{aligned} P_{n+1}(x) &= P\{w_n + u_n < x\} = \int_{-\infty}^{\infty} P_n(x-t) dU(t) = \\ &= \int_{-\infty}^{\infty} P_n(x-t) dU(t), \end{aligned} \quad (2.40)$$

то, переходя к пределу при $n \rightarrow \infty$ и используя теорему Лебега, получим

$$P(x) = \int_{-\infty}^x P(x-t) dU(t). \quad (2.41)$$

Это уравнение носит название уравнения Линдли и принадлежит к типу уравнений Винера — Хопфа. Оно может решаться либо методом последовательных приближений в соответствии с равенством (2.40), либо методом факторизации ядра уравнения.

В заключение отметим интересную вероятностную интерпретацию для распределения величины w_n . Обозначим через

$$U_k^m = \sum_{i=k}^m u_i, \quad 1 \leq k \leq m;$$

$$U_k = \sum_{i=1}^k u_i, \quad U_0 = 0.$$

Из полученных соотношений следует, что

$$w_{n+1} = \max\{0, w_n + u_n\} = \dots = \max\{0, U_n^n, 1 \leq k \leq n\}.$$

Поскольку величины U_k , $k \geq 1$ независимы и одинаково распределены, то распределения величин U_k^n совпадают с распределениями величин U_{n-k+1} , $1 \leq k \leq n$. Поэтому в смысле равенства распределений

$$w_{n+1} = \max\{U_k, 0 \leq k \leq n\}. \quad (2.42)$$

Из полученного выражения также видно, что w_n может только возрастать с ростом n . Поэтому существует предельная (возможно бесконечная) случайная величина w и

$$w = \sup\{U_k, k \geq 0\}.$$

Пусть существует конечное математическое ожидание величин U_k . Используя усиленный закон больших чисел, нетрудно получить, что ω будет собственной случайной величиной тогда и только тогда, когда $MU_k < 0$. Смысл этого условия заключается в том, что среднее время между двумя последовательными поступлениями требований больше, чем среднее время обслуживания. При этом предположении функция $P(x)$, удовлетворяющая уравнению (4.3) будет функцией распределения, т. е. $P(+\infty) = 1$.

ГЛАВА 3. НЕКОТОРЫЕ СПЕЦИАЛЬНЫЕ ТИПЫ СИСТЕМ

3.1. Системы с групповым поступлением

Рассмотрим простейшую систему с групповым поступлением, когда длительности интервалов между поступающими группами требований независимы и имеют показательное распределение с параметром λ , количества N_i требований в группах случайны, независимы, одинаково распределены и не зависят от длительностей интервалов между поступлениями, а время обслуживания одного требования имеет показательное распределение со средним $1/\mu$.

Обозначим $r_k = P\{N_i = k\}$; $Q(t)$ — количество требований в системе.

При сделанных предположениях $Q(t)$ будет однородной цепью Маркова с непрерывным временем, интенсивности q_{ij} переходов которой имеют следующий вид:

$$q_{ij} = \begin{cases} \lambda r_{j-i}, & j > i; \\ \mu, & j = i - 1, \end{cases}$$

$$q_{ii} = -\sum_{j \neq i} q_{ij} = \begin{cases} -\mu - \lambda, & i \neq 0; \\ -\lambda, & i = 0. \end{cases}$$

В силу известных свойств цепей Маркова, стационарное распределение $\{\pi_k, k \geq 0\}$ является решением системы уравнений

$$\begin{cases} \sum_i \pi_i = 1; \\ \lambda \pi_0 = \mu \pi_1; \\ (\lambda + \mu) \pi_k = \mu \pi_{k+1} + \sum_{i=0}^{k-1} \pi_i \lambda r_{k-i}, & k \geq 1. \end{cases}$$

Введем производящую функцию

$$P(z) = \sum_{k=0}^{\infty} \pi_k z^k.$$

Тогда из системы уравнений следует, что

$$\begin{aligned} -\mu\pi_0 + (\lambda + \mu)P(z) &= \frac{\mu}{z}(P(z) - \pi_0) + \\ &+ \lambda \sum_{k=1}^{\infty} z^k \sum_{i=0}^{k-1} \pi_i r_{k-i} = \frac{\mu}{z}(P(z) - \pi_0) + \\ &+ \lambda \sum_{i=0}^{\infty} \pi_i z^i \sum_{k=i+1}^{\infty} z^{k-i} r_{k-i} = \frac{\mu}{z}(P(z) - \pi_0) + \lambda R(z)P(z), \end{aligned}$$

где

$$R(z) = \sum_{k=1}^{\infty} z^k r_k.$$

Поэтому

$$P(z) = \frac{\mu\pi_0(z-1)}{z(\lambda + \mu) - \mu - \lambda z R(z)}.$$

Для нахождения π_0 воспользуемся условием нормировки, из которого следует, что $\lim_{z \rightarrow 1-0} P(z) = 1$.

Имеем

$$1 = \lim_{z \rightarrow 1-0} P(z) = \frac{\mu\pi_0}{\lambda + \mu - \lambda(1 + R'(1))} = \frac{\mu\pi_0}{\mu - \lambda n},$$

где $n = R'(1)$ — среднее количество требований в одной группе.

Поэтому

$$\pi_0 = \frac{\mu - \lambda n}{\mu}$$

и

$$P(z) = \frac{(\mu - \lambda n)(1 - z)}{\lambda z R(z) + \mu - (\lambda + \mu)z}.$$

Заметим, что необходимым условием существования стационарного распределения будет $\mu > \lambda n$. Можно показать, что оно будет и достаточным условием.

Задания для самоконтроля

Найти стационарное распределение и производящую функцию числа требований в системе с групповым поступлением при

а) $r_k = (1 - \alpha)\alpha^{k-1}$, $k \geq 1$, $0 < \alpha < 1$;

б) $r_1 = r_2 = \frac{1}{2}$;

в) $r_2 = 1$.

3.2. Приоритетные системы

Рассмотрим простейший случай системы $M|M|1|\infty|f_2$ с двумя типами требований. В рассматриваемой системе моменты поступления требований первого типа образуют простейший поток с параметром λ_1 , а второго типа — с параметром λ_2 . Интенсивности обслуживания равны μ_1 для требований первого типа и μ_2 для второго. Так как в системе количество мест для ожидания неограничено, то верхний индекс пятого символа может быть опущен.

Будем считать, что внутри одного типа требований выбор на обслуживание проводится по принципу «первым прибыл — первым обслужен». (Напомним, что в соответствии со сделанными соглашениями об обозначениях, требования первого типа имеют абсолютный приоритет перед требованиями второго типа).

Обозначим $Q_1(t)$ и $Q_2(t)$ — количество требований первого и второго типа соответственно, находящихся в системе в момент времени t . Тогда процесс $\bar{Q}(t) = (Q_1(t), Q_2(t)) \in R_+^2$ будет однородной цепью Маркова, интенсивности переходов $q((i, j), (k, l))$ которой имеют при $(k, l) \neq (i, j)$ следующий вид:

$$\begin{aligned} q((i, j), (i+1, j)) &= \lambda_1; \\ q((i, j), (i, j+1)) &= \lambda_2; \\ q((i, j), (i-1, j)) &= \mu_1 \text{ при } i \geq 1; \\ q((0, j), (0, j-1)) &= \mu_2 \text{ при } j \geq 1; \\ q((i, j), (k, l)) &= 0 \text{ в остальных случаях.} \end{aligned}$$

В таком случае стационарное распределение $p(i, j)$ цепи $\bar{Q}(t)$ является нормированным решением системы уравнений

$$\begin{cases} (\lambda_1 + \lambda_2 + \mu_1) p(i, j) = \lambda_1 p(i-1, j) + \lambda_2 p(i, j-1) + \mu_1 p(i+1, j), & i \geq 1, j \geq 1; \\ (\lambda_1 + \lambda_2 + \mu_1) p(i, 0) = \lambda_1 p(i-1, 0) + \mu_1 p(i+1, 0), & i \geq 1; \\ (\lambda_1 + \lambda_2 + \mu_2) p(0, j) = \lambda_2 p(0, j-1) + \mu_2 p(0, j+1), & j \geq 1; \\ (\lambda_1 + \lambda_2) p(0, 0) = \mu_1 p(1, 0) + \mu_2 p(0, 1). \end{cases} \quad (3.1)$$

Введем производящие функции

$$P(x, y) = \sum_{i, j=0}^{\infty} P(i, j) x^i y^j \text{ и } P_2(y) = \sum_{j=0}^{\infty} p(0, j) y^j.$$

Тогда из системы уравнений (3.1) следует равенство

$$\begin{aligned}
 (\lambda_1 + \lambda_2 + \mu_1) P(x, y) - \mu_1 P_2(y) + \mu_2 (P_2(y) - p(0, 0)) = \\
 = \lambda_1 x P(x, y) + \lambda_2 y P(x, y) + \frac{\mu_1}{x} (P(x, y) - P_2(y)) + \\
 + \frac{\mu_2}{y} (P_2(y) - P(0, 0)).
 \end{aligned}$$

Поэтому

$$P(x, y) = \frac{P_2(y) \left(\mu_1 - \frac{\mu_1}{x} - \mu_2 + \frac{\mu_2}{y} \right) + p(0, 0) \left(\mu_2 - \frac{\mu_2}{y} \right)}{\lambda_1 + \lambda_2 + \mu_1 - \lambda_1 x - \lambda_2 y - \frac{\mu_1}{x}}.$$

Знаменатель правой части последнего выражения, рассматриваемый как функция x , имеет корни

$$\begin{aligned}
 x_{1,2} &= x_{1,2}(y) = \\
 &= \frac{(\lambda_1 + \lambda_2 + \mu_1 - \lambda_2 y) \pm \sqrt{(\lambda_1 + \lambda_2 + \mu_1 - \lambda_2 y)^2 - 4\lambda_1 \mu_1}}{2}.
 \end{aligned}$$

Один из этих корней, как следует из теоремы Руше, при $|y| < 1$ лежит в единичном круге. В самом деле, если $f(x) = \lambda_1 + \lambda_2 + \mu_1 - \frac{\mu_1}{x}$, $g(x) = -\lambda_1 x - \lambda_2 y$, то $|f(x)|_{|x|=1} \geq \lambda_1 + \lambda_2$, а $|g(x)|_{|x|=1} < \lambda_1 + \lambda_2$ при $|y| < 1$. Далее, $f(x)$ имеет в круге $|x| < 1$ один корень $x_0 = \frac{\mu_1}{\lambda_1 + \lambda_2 + \mu_1}$, поэтому $f(x) + g(x)$ также имеет в этом круге один корень.

Легко видеть, что из корней $x_1(y)$ и $x_2(y)$ внутри единичного круга должно находиться $x_2(y)$. Так как $P(x_2(y), y)$ существует, то числитель выражения, задающего $P(x, y)$, должен обращаться в нуль при $x = x_2(y)$:

$$P_2(y) \left(\mu_1 - \frac{\mu_1}{x_2(y)} - \mu_2 + \frac{\mu_2}{y} \right) + p(0, 0) \left(\mu_2 - \frac{\mu_2}{y} \right) = 0.$$

Отсюда следует, что

$$P_2(y) = p(0, 0) \frac{\mu_2 \left(\frac{1}{y} - 1 \right)}{\mu_1 \left(1 - \frac{1}{x_2(y)} \right) + \mu_2 \left(\frac{1}{y} - 1 \right)}.$$

Поэтому

$$P(x, y) = p(0, 0) \frac{\mu_2 - \frac{\mu_2}{y} + \frac{\left(\frac{\mu_2}{y} - \mu_2 \right) \left(\mu_1 - \frac{\mu_1}{x} - \mu_2 + \frac{\mu_2}{y} \right)}{\mu_1 \left(1 - \frac{1}{x_2(y)} \right) + \mu_2 \left(\frac{1}{y} - 1 \right)}}{\lambda_1 + \lambda_2 + \mu_1 - \lambda_1 x - \lambda_2 y - \frac{\mu_1}{x}}.$$

Постоянная $p(0, 0)$, как обычно, может быть найдена из условия нормировки $\lim_{\substack{x \rightarrow 1-0 \\ y \rightarrow 1-0}} P(x, y) = 1$.

3.3. Система с периодическим регулярным входным потоком

Пусть в одноканальной системе массового обслуживания с ожиданием длительности интервалов между поступлениями требований образуют неслучайную периодическую с периодом n последовательность: $\tau_k = \tau_{k+n}$, $k \geq 1$. Времена обслуживания различных требований будут предполагаться независимыми и имеющими показательное распределение с показателем μ . Обозначим через ξ_k количество требований в системе в момент времени $t_k - 0 = \sum_{i=1}^k \tau_i - 0$, непосредственно предшествующий моменту появления требования с номером k . Последовательность $\{\xi_k, k \geq 1\}$ образует цепь Маркова с множеством состояний $N = \{0, 1, 2, \dots\}$ и дискретным временем. Эта цепь будет неоднородной во времени, если $n > 1$. (Марковость последовательности $\{\xi_k, k \geq 1\}$ обусловлена марковским свойством показательного распределения времени обслуживания).

Для исследования предельных характеристик цепи $\{\xi_k, k \geq 1\}$ введем векторнозначную цепь Маркова

$$\eta_m = (\xi_{mn}, \xi_{mn+1}, \dots, \xi_{mn+m-1}).$$

В отличие от $\{\xi_k, k \geq 1\}$ последовательность $\{\eta_m, m \geq 1\}$ будет однородной цепью Маркова (проверьте это!). Множеством состояний этой цепи является

$$A = \{(r_0, r_1, \dots, r_{n-1}) \in N^{n*} : r_i \leq r_{i-1} + 1, i = 1, \overline{n-1}\}.$$

Все состояния этой цепи сообщаются и образуют один класс. В силу теоремы солидарности состояния цепи η_m имеют общий период. Этот период будет равен единице, если с положительной вероятностью цепь η_m может остаться в некотором состоянии при очередном переходе. Таким состоянием, в частности, является $(0, 0, \dots, 0)$. В самом деле, вероятность перехода из $(0, 0, \dots, 0)$ в $(0, \dots, 0)$ равна вероятности того, что n поочередно поступающих в свободную систему требований обслуживаются до поступления следующего требования, т. е.

$\prod_{k=1}^n (1 - e^{-\mu \tau_k}) > 0$. Итак, цепь $\{\eta_m, m > 1\}$ неперiodична и имеет не более одного стационарного распределения π

(которое обязано совпадать с эргодическим). Распределение π на одноточечных множествах $I = \{(r_0, r_1, \dots, r_{n-1})\}$ задается соотношением

$$\pi(I) = q_0(r_0) \prod_{k=1}^{n-1} P_k(r_{k-1}, r_k), \quad (3.2)$$

где $q_0(r) = \lim_{m \rightarrow \infty} P\{\xi_{mn} = r\}$,

$$P_k(i, j) = \begin{cases} e^{-\mu\tau_k} \frac{(\mu\tau_k)^{i+1-j}}{(i+1-j)!}, & i \geq j \geq 1; \\ e^{-\mu\tau_k}, & j = i+1, i \geq 0; \\ 1 - \sum_{l=1}^{i+1} P_k(i, l), & j = 0, i \geq 0; \\ 0 & \text{в остальных случаях.} \end{cases} \quad (3.3)$$

Вероятности $P_k(i, j)$ — не что иное, как вероятности перехода за один шаг цепи ξ_m : $P_k(i, j) = P\{\xi_{mn+k} = j / \xi_{mn+k-1} = i\}$, чем и обусловлено равенство (3.2). Поскольку $\xi(t_k + 0) = \xi_k + 1$, то $P_k(i, j) = P\{\text{за время } \tau_k \text{ из имеющихся } i+1 \text{ требований будет обслужено } i+1-j\}$. Если $i \geq j \geq 1$ либо $j = i+1, i \geq 0$, то $P_k(i, j)$ равно вероятности того, что при полной загрузке обслуживающее устройство за время τ_k обслужит $i+1-j$ требований, откуда следуют равенства (3.3).

Из равенства (3.2) вытекает, что распределения π и q_0 существуют одновременно и однозначно определяют друг друга.

Распределение q_0 количества требований в системе. Покажем, что распределение q_0 имеет вид

$$q_0(k) = \sum_{i=1}^n c_{0i} \alpha_i^k; \quad k \geq 0; \quad 0 < |\alpha_i| < 1. \quad (3.4)$$

Замечание. Прежде чем переходить к доказательству этого утверждения, заметим, что из (3.4) следует

$$q_j(k) = \lim_{m \rightarrow \infty} P\{\xi_{mn+j} = k\} = \sum_{i=1}^n c_{ji} \alpha_i^k.$$

В самом деле, из (3.4) получаем

$$q_1(k) = \sum_{i=0}^{\infty} q_0(i) P_1(i, k) = \sum_{i=1}^n c_{0i} \sum_{l=0}^{\infty} \alpha_i^l P_1(l, k).$$

Если $k \geq 1$, то

$$\begin{aligned} q_1(k) &= \sum_{i=1}^n c_{0i} \sum_{k=i-1}^{\infty} \alpha_i^k e^{-\mu\tau_i} \frac{(\mu\tau_i)^{l+1-k}}{(l+1-k)!} = \\ &= \sum_{i=1}^n c_{0i} \alpha_i^{k-1} e^{-\mu\tau_i(1-\alpha_i)} = \sum_{i=1}^n c_{1i} \alpha_i^k, \end{aligned}$$

где

$$c_{1i} = c_{0i} \alpha_i^{-1} e^{-\mu\tau_i(1-\alpha_i)}.$$

В общем случае равенство $q_j(k) = \sum_{i=1}^n c_{ji} \alpha_i^k$ проверяется по индукции.

Обозначим $T = \sum_{k=1}^n \tau_k$ и докажем вначале два вспомогательных утверждения.

Лемма 3.1. Если выполнено условие

$$\mu T > n, \quad (3.5)$$

то уравнение

$$\alpha^n = e^{-\mu T(1-\alpha)} \quad (3.6)$$

имеет n корней в круге $|\alpha| < 1$.

Доказательство. Лемма может быть доказана с использованием теоремы Руше для области $|\alpha| < 1 - \varepsilon$, где ε достаточно мало. В самом деле, пусть $f(\alpha) = \alpha^n$, $g(\alpha) = e^{-\mu T(1-\alpha)}$.

Тогда

$$|f(\alpha)|_{|\alpha|=1-\varepsilon} = (1-\varepsilon)^n = 1 - n\varepsilon + o(\varepsilon),$$

$$|g(\alpha)|_{|\alpha|=1-\varepsilon} = 1 - \varepsilon\mu T + o(\varepsilon).$$

Поэтому при выполнении условия (3.5) и достаточно малых $\varepsilon > 0$

$$|f(\alpha)| > |g(\alpha)| \text{ при } |\alpha| = 1 - \varepsilon.$$

Поскольку уравнение

$$f(\alpha) = 0$$

имеет n корней в круге $|\alpha| < 1 - \varepsilon$, то на основании теоремы Руше уравнение (3.6) имеет n корней в круге $|\alpha| < 1$.

Лемма 3.2. При выполнении условия (3.5) все корни уравнения (3.6) в круге $|\alpha| < 1$ простые.

Доказательство. Рассмотрим совокупность уравнений

$$\alpha e^{2\pi i k/n} = e^{-\frac{\mu T}{n}(1-\alpha)}, \quad 1 \leq k \leq n. \quad (3.7)$$

Так же, как и при доказательстве леммы 3.1 с помощью теоремы Руше устанавливается, что каждое из уравнений совокупности (3.7) имеет один корень α_k в круге $|\alpha| < 1$. Очевидно, что эти корни α_k являются также и корнями уравнения (3.6). Допущение о совпадении некоторых корней приводит к противоречию.

В самом деле, если $\alpha_k = \alpha_l$, $k \neq l$, то

$$\begin{aligned}\alpha_k e^{2\pi_i k/n} &= e^{-\mu \frac{T}{n}(1-\alpha_k)} = e^{-\mu \frac{T}{n}(1-\alpha_l)} = \\ &= \alpha_l e^{2\pi_i l/n} = \alpha_k \cdot e^{2\pi_i l/n}.\end{aligned}\quad (3.8)$$

Так как $\alpha = 0$ заведомо не является корнем (3.7), то из (3.8) следует

$$e^{2\pi_i k/n} = e^{2\pi_i l/n},$$

что невозможно при $k \neq l$, $1 \leq k \leq n$, $1 \leq l \leq n$. Лемма доказана.

Теорема 3.1. Если выполнено условие $\mu T > n$, то распределение q_0 имеет вид

$$q_0(k) = \sum_{i=1}^n c_{0i} \alpha_i^k,$$

где α_i — простые корни уравнения $\alpha^n = e^{-\mu T(1-\alpha)}$, а c_{0i} — единственное решение системы уравнений

$$\begin{aligned}\sum_{i=1}^n \frac{c_{0i}}{1-\alpha_i} \frac{1}{\alpha_i^k} e^{-\mu(1-\alpha)} \sum_{j=1}^k \tau_j &= 1, \\ 0 \leq k \leq n-1.\end{aligned}\quad (3.9)$$

Доказательство. Пусть некоторая последовательность $q_0(k)$ задается равенством (3.4), в котором $0 < |\alpha_i| < 1$, $1 \leq i \leq n$. Тогда $\sum_{k=0}^{\infty} |q_0(k)| < \infty$. Из теории однородных цепей Маркова известно, что $q_0(k)$, $k \geq 0$ будет являться стационарным распределением цепи ξ_{mn} тогда и только тогда, когда

$$\begin{cases} \sum_{k=0}^{\infty} q_0(k) = 1; \\ q_0(k) = \sum_{i=0}^{\infty} q_0(i) P(i, k), \quad k \geq 0, \end{cases}\quad (3.10)$$

где $P(i, k)$ — переходная вероятность цепи ξ_{mn} , $m \geq 1$.

Обозначим $P = \| P(i, j) \|$, $P_k = \| P_k(i, j) \|$, $Q_k = (q_k(0), q_k(1), \dots)$. Тогда в силу марковости цепи ξ_m

$$P = P_1 P_2 \dots P_n, \quad Q_k = Q_0 P_1 P_2 \dots P_k = Q_{k-1} P_k,$$

и система уравнений (3.10) эквивалентна системе

$$\sum_{j=0}^{\infty} q_k(j) = 1, \quad \sum_{j=0}^{\infty} |q_k(j)| < \infty, \quad 0 \leq k \leq n-1, \quad (3.11)$$

$$Q_k = Q_{k-1} P_k, \quad 1 \leq k \leq n. \quad (3.12)$$

Система уравнений (3.11) — (3.12) избыточна. Рассмотрим эквивалентную ей систему

$$\begin{cases} \sum_{k=0}^{\infty} q_k(j) = 1; \\ q_k(j) = \sum_{i=j-1}^{\infty} q_{k-1}(i) P_k(i, j), \quad 1 \leq k \leq n-1, \quad j \geq 1; \\ q_0(j) = \sum_{i=j-1}^{\infty} q_n(i) P_n(i, j), \quad j \geq 1. \end{cases} \quad (3.13)$$

Так как в силу замечания, сделанного выше, $q_k(j)$ имеет вид

$$q_k(j) = \sum_{i=1}^n c_{ki} \alpha_i^j,$$

то с учетом формул (3.3) система (3.13) преобразовывается к

$$\begin{cases} \sum_{i=1}^n c_{ki} \alpha_i^j = \sum_{i=1}^n c_{k-1i} \alpha_i^{j-1} e^{-\mu \tau_k (1-\alpha_i)}, \quad 1 \leq k \leq n-1, \quad j \geq 2; \\ \sum_{i=1}^n c_{0i} \alpha_i^j = \sum_{i=1}^n c_{n-1i} \alpha_i^{j-1} e^{-\mu \tau_n (1-\alpha_i)}, \quad j \geq 2; \\ \sum_{i=1}^n c_{ki} \alpha_i = (q_{k-1}(0) - \sum_{i=1}^n c_{k-1i}) e^{-\mu \tau_k} + \\ + \sum_{i=1}^n c_{k-1i} e^{-\mu \tau_k (1-\alpha_i)}, \quad 1 \leq k \leq n-1; \\ \sum_{i=1}^n c_{0i} \alpha_i = (q_{n-1}(0) - \sum_{i=1}^n c_{n-1i}) e^{-\mu \tau_n} + \\ + \sum_{i=1}^n c_{n-1i} e^{-\mu \tau_n (1-\alpha_i)}; \\ q_k(0) = 1 - \sum_{i=1}^n c_{ki} \frac{\alpha_i}{1-\alpha_i}, \quad 0 \leq k \leq n-1. \end{cases} \quad (3.14)$$

Так как в представлении (3.4) все α_i различны и отличны

от нуля, то две первые серии уравнений системы (3.14) эквивалентны равенствам

$$\begin{cases} c_{ki}\alpha_i = c_{k-1}e^{-\mu\tau_k(1-\alpha_i)}, & 1 \leq k \leq n-1, 1 \leq i \leq n; \\ c_{0i}\alpha_i = c_{n-1}e^{-\mu\tau_n(1-\alpha_i)} & 1 \leq i \leq n. \end{cases} \quad (3.15)$$

С учетом (3.15) и последних уравнений системы (3.14) оставшиеся ее уравнения переписутся в виде

$$\sum_{i=1}^n c_{ki} \frac{1}{1-\alpha_i} = 1, \quad 0 \leq k \leq n-1. \quad (3.16)$$

Возможны два случая:

а) все c_{ki} отличны от нуля и

б) существуют значения k и i такие, что $c_{ki} = 0$.

В случае а), перемножив все уравнения системы (3.15) при фиксированных значениях индекса i , получим, что все α_i удовлетворяют уравнению (3.6). В случае б) из системы (3.15) следует, что $c_{ri} = 0$ для всех значений r . Ввиду специфики рассматриваемых уравнений значение α_i может быть произвольным и мы можем положить его равным некоторому корню уравнения (3.6)

Итак, α_i являются корнями уравнения (3.6). Мы выберем их различными, что позволяют леммы 3.1 и 3.2. Очевидно, что $\alpha_i \neq 0$.

Покажем теперь, что при таком выборе значений α_i система (3.15) — (3.16) совместна.

Из (3.15) получаем

$$c_{ki} = c_{0i}\alpha_i^{-k} e^{-\mu(1-\alpha_i) \sum_{j=1}^k \tau_j}, \quad 0 \leq k \leq n-1.$$

Подставив эти значения в (3.16), сведем задачу к проверке совместности системы

$$\sum_{i=1}^n c_{0i} \frac{1}{1-\alpha_i} \alpha_i^{-k} e^{-\mu(1-\alpha_i) \sum_{j=1}^k \tau_j} = 1, \quad (3.17)$$

$$0 \leq k \leq n-1.$$

Определитель D системы (3.17) является аналитической функцией переменных $\tau_1, \tau_2, \dots, \tau_{n-1}$. При $\tau_1 = \tau_2 = \dots = \tau_{n-1} = \frac{T}{n}$ он с точностью до множителя $\prod_{i=1}^n (1-\alpha_i)^{-1} \neq$

$\neq 0$ совпадает с определителем Вандермонда, порожденными величинами

$$x_1 = \alpha_1^{-1} e^{-\mu(1-\alpha_1) \frac{T}{n}}, \quad x_2 = \alpha_2^{-1} e^{-\mu(1-\alpha_2) \frac{T}{n}}, \quad \dots, \\ x_{n-1} = \alpha_{n-1}^{-1} e^{-\mu(1-\alpha_{n-1}) \frac{T}{n}}$$

Из доказательства леммы 3.2 следует, что $x_k = e^{2\pi i k/n}$, $1 \leq k \leq n-1$, и, следовательно, $x_k \neq x_l$ при $k \neq l$. Поэтому $D \neq 0$ при $\tau_i = \frac{T}{n}$. Как аналитическая функция, $D \neq 0$ на некотором открытом множестве G , всюду плотном в R^{n-1} . Как следствие этого, получаем, что система (3.17) совместна для почти всех значений τ_i , $1 \leq i \leq n-1$ из множества $K = \{(\tau_1, \dots, \tau_{n-1}) : \tau_i > 0, \sum_{i=1}^{n-1} \tau_i \leq T\}$ и для этих значений τ , $1 \leq i \leq n-1$ теорема доказана.

Пусть теперь $(\tau_{10}, \tau_{20}, \dots, \tau_{n-1,0}) \in K \cap \partial G$, где ∂G — граница области G . В таком случае существует последовательность $\tau^r = (\tau_{1r}, \tau_{2r}, \dots, \tau_{n-1r}) \in K \cap G$ ($\tau_{1r}, \tau_{2r}, \dots, \tau_{n-1r}$) $\xrightarrow{r \rightarrow \infty} (\tau_{10}, \tau_{20}, \dots, \tau_{n-1,0})$.

Для последовательности τ_r система (3.17) совместна и стационарное распределение $q_0 = q_0^{(k)}$ имеет производящую функцию

$$q^{(r)}(s) = \sum_{i=1}^n c_{0i}^{(r)} \frac{1}{1 - \alpha_i s}. \quad (3.18)$$

При выполнении условий теоремы $q^{(r)}(s) \xrightarrow{r \rightarrow \infty} q(s)$, где $q(s)$ — производящая функция стационарного распределения $q_0^{(0)}$, отвечающего $(\tau_{10}, \tau_{20}, \dots, \tau_{n-1,0})$. Поскольку на основании (3.18) $q^{(r)}(s)$ является дробно-рациональной функцией со знаменателем, не зависящим от τ , (напомним, что α_i — корни уравнения (3.6), в котором T постоянно), то из сходимости $q^{(r)}(s) \xrightarrow{r \rightarrow \infty} q(s)$ следует, что

$$q(s) = \sum_{i=1}^n c_{0i} \frac{1}{1 - \alpha_i s}, \quad c_{0i} = \lim_{r \rightarrow \infty} c_{0i}^{(r)}.$$

Поэтому стационарное распределение $q_0^{(0)}$ также имеет вид (3.4). Тем самым установлена и совместность системы (3.17) при всех значениях $(\tau_1, \tau_2, \dots, \tau_{n-1})$.

Теорема доказана полностью. Заметим, что в ходе доказательства мы фактически установили, что определитель D системы (3.17) отличен от нуля при всех допустимых значениях $\tau_1, \tau_2, \dots, \tau_n$, т. е. при $(\tau_1, \tau_2, \dots, \tau_{n-1}) \in K$.

Частным случаем рассмотренной системы является система с групповым поступлением n требований через интервалы постоянной длины T . Для этого достаточно в исходной системе массового обслуживания положить $\tau_1 = \tau_2 = \dots = \tau_{n-1} = 0, \tau_n = T$. При этом уравнение (3.6) по-прежнему будет задавать необходимые значения α_i , $1 \leq i \leq n$, а система (3.17), определяющая коэффициенты c_{0i} , приобретет вид

$$\sum_{i=1}^n c_{0i} \frac{1}{1-\alpha_i} \alpha_i^{-k} = 1, \quad 0 \leq k \leq 1. \quad (3.19)$$

Система (3.19), конечно, проще системы (3.17), и ее совместность очевидна, так как ее определитель с точностью до множителя $\prod_{i=1}^n (1-\alpha_i)^{-1} \neq 0$ совпадает с определителем Вандермонда, порожденным величинами α_i^{-1} , $1 \leq i \leq n$. Последний заведомо отличен от нуля, так как $\alpha_i \neq \alpha_j$ при $i \neq j$ (отличие определителя от нуля следует из доказательства теоремы 3.1).

Система (3.19) может быть решена по правилу Крамера:

$$c_{0k} = D_k D^{-1}.$$

Несложные вычисления показывают, что

$$D = \prod_{i=1}^n (1-\alpha_i)^{-1} \cdot \prod_{i>j} (\alpha_i^{-1} - \alpha_j^{-1});$$

$$D_k = \prod_{j \neq k} (1-\alpha_j)^{-1} \prod_{i=j} (\alpha_{i1}^{-1} - \alpha_{j1}^{-1}),$$

где $\alpha_{j1} = \alpha_j$ при $j \neq k$ и $\alpha_{k1} = 1$.

Следовательно,

$$c_{0k} = (1-\alpha_k) \alpha_k^{n-1} \prod_{i \neq k} \frac{1-\alpha_i}{\alpha_k - \alpha_i}. \quad (3.20)$$

Теорема 3.2. В системе с групповым поступлением n требований через интервалы длины T и показательно распределенным временем обслуживания распределение q_0 имеет вид

$$q_0(k) = \sum_{i=1}^n c_{0k} \alpha_i^k,$$

где α_i — корни уравнения (3.6), а c_{0k} задаются равенствами (3.20).

Вернемся к исходной системе. Формула $q(s) = \sum_{i=1}^n c_{0i} \times \times \frac{1}{1 - \alpha_i s}$ для производящей функции случайной величины N с распределением q_0 позволяет находить все его моменты. Так, $q^{(k)}(s) = \sum_{i=1}^n c_{0i} \frac{\alpha_i^k k!}{(1 - \alpha_i s)^{k+1}}$, и поэтому в стационарном режиме

$$MN = \sum_{k=1}^{\infty} k q_0(k) = q'(1) = \sum_{i=1}^n c_{0i} \frac{\alpha_i}{(1 - \alpha_i)^2},$$

$$MN^2 = \sum_{k=1}^{\infty} k^2 q_0(k) = q''(1) + q'(1) = \sum_{i=1}^n c_{0i} \frac{\alpha_i (\alpha_i + 1)}{(1 - \alpha_i)^3}.$$

В заключение найдем распределение времени ожидания W_0 требования, поступающего в момент времени $t_{nm} = 0$ в систему, находящуюся в стационарном режиме. W_0 допускает представление $W_0 = \sum_{k=1}^N \theta_k$, где N — количество требований в системе в момент поступления рассматриваемого требования, а θ_k — их оставшиеся времена обслуживания. В силу отсутствия памяти у показательного распределения, θ_k , $1 \leq k \leq N$ независимы, не зависят от N и имеют показательное распределение с показателем μ . По формуле полного условного математического ожидания имеем

$$\begin{aligned} M e^{-s W_0} &= M M_N e^{-s \sum_{k=1}^N \theta_k} = M (M e^{-s \theta_1})^N = M \left(\frac{\mu}{\mu + s} \right)^N = \\ &= q \left(\frac{\mu}{\mu + s} \right) = \sum_{i=1}^n c_{0i} + \sum_{i=1}^n c_{0i} \frac{\mu \alpha_i}{s + \mu - \mu \alpha_i}. \end{aligned}$$

Отсюда

$$P\{W_0 > t\} = \sum_{i=1}^n c_{0i} \frac{\alpha_i}{1 - \alpha_i} e^{-\mu(1 - \alpha_i)t} \quad \text{при } t > 0.$$

Заметим, что функция распределения W_0 имеет скачок в нуле величины $\sum_{i=1}^n c_{0i} = q_0(0)$.

3.4. Система с переменной интенсивностью обслуживания

В рассмотренных ранее системах массового обслуживания с одним обслуживающим устройством интенсивность обслуживания предполагалась постоянной. В системах типа $M|M|\infty$ интенсивность обслуживания системой прямо пропорциональна количеству требований N , находящихся в системе.

Пусть $\xi(t)$ — объем работы, необходимый для обслуживания требований, находящихся в системе $M|M|\infty$ в момент времени t ; $\alpha(t)$ — объем работы, поступивший в систему за время t , ξ_i ; $i \leq N$ — времена обслуживания имеющихся в системе в момент t требований; μ_1 — интенсивность обслуживания одним обслуживающим устройством. Тогда, сравнивая значения $\xi(t)$ и $\xi(t + \Delta)$, получим

$$\xi(t + \Delta) = \xi(t) - N\mu_1\Delta + \alpha(t + \Delta) - \alpha(t) + o(\Delta), \quad (3.21)$$

где $\frac{o(\Delta)}{\Delta} \xrightarrow{\Delta \rightarrow 0} 0$.

Уравнение (3.21) допускает запись в дифференциалах

$$d\xi(t) = -N\mu_1 dt + d\alpha(t). \quad (3.22)$$

В силу закона больших чисел при $M\xi_1 < \infty$

$$\frac{\xi(t)}{N} = \frac{1}{N} \sum_{i=1}^N \xi_i \xrightarrow{N \rightarrow \infty} M\xi_1$$

и при больших значениях N по крайней мере формально уравнение (3.22) может быть заменено на более простое:

$$d\xi(t) = -\mu\xi(t) dt + d\alpha(t), \quad (3.23)$$

где

$$\mu = \frac{\mu_1}{M\xi_1}.$$

Таким образом, есть основания полагать, что уравнение (3.23) пригодно для описания некоторых процессов, связанных с системой $M|M|\infty$ в условиях, когда в последней находится большое количество требований, например, при больших значениях интенсивности входного потока.

Можно принять другую точку зрения. Именно, можно постулировать, что процесс $\xi(t)$, удовлетворяющий уравнению (3.23), равен объему работы, находящейся в системе в момент времени t . Тогда уравнение (3.23) указывает на следующие особенности системы:

а) при $\alpha(s) = \text{const}$ на интервале $s \in [t, t + u]$

$$d\xi(s) = -\mu \xi(s) ds;$$

$$\xi(s) = \xi(t) e^{-\mu(s-t)},$$

б) при $\alpha(s) - \alpha(s-0) > 0$;

$$\xi(s) - \xi(s-0) = \alpha(s) - \alpha(s-0).$$

Заметим, что свойства а) и б) вытекают из уравнения (3.23) и не требовали указания других атрибутов системы (количество обслуживающих устройств, дисциплину выбора на обслуживание и т. д.). Поэтому уравнение (3.23) пригодно и для описания любых процессов, обладающих свойствами а) и б). В частности, оно может описать следующие процессы:

а) величину заряда $\xi(t)$ на пластинах конденсатора, заряжаемого случайным потоком зарядов $\alpha(t)$;

б) объем воды в хранилище, меняющийся как вследствие фильтрации (если скорость фильтрации пропорциональна объему воды), так и вследствие осадков.

Будем рассматривать систему массового обслуживания, объем незавершенной работы $\xi(t) > 0$ которой в момент времени t удовлетворяет уравнению (3.23), в котором $\mu > 0$, а $\alpha(t)$ — неотрицательный обобщенный пуассоновский процесс с интенсивностью скачков λ и функцией распределения скачка $F(x)$.

Конструктивно процесс $\alpha(t)$ задается так:

а) траектории $\alpha(\cdot)$ кусочно-постоянны и непрерывны справа, а моменты t_i , $i \geq 1$ его разрывов образуют простейший поток с параметром λ ;

б) величины скачков $\xi_i = \alpha(t_i) - \alpha(t_i - 0)$, $i \geq 1$, образуют последовательность положительных, независимых, одинаково распределенных случайных величин, не зависящую от последовательности $\{t_i, i \geq 1\}$; функция распределения ξ_i равна $F(x)$.

Как обычно, формальное уравнение (3.23) считается одной из форм записи уравнения

$$\xi(t) - \xi(0) = -\mu \int_0^t \xi(u) du + \alpha(t), \quad (3.24)$$

в котором интеграл определяется как интеграл Римана и существует с вероятностью единица.

Лемма 3.3. Решение уравнения (3.24) с вероятностью единица имеет вид

$$\xi(t) = \xi(0) e^{-\mu t} + \int_0^t e^{-\mu(t-u)} d\alpha(u), \quad (3.25)$$

в котором интеграл существует как интеграл Стильеса с вероятностью единица.

Доказательство. Доказательство леммы 3.3 проводится подстановкой (3.25) в уравнение (3.24):

$$\begin{aligned}\xi(t) - \xi(0) &= -\mu \int_0^t \left[\xi(0) e^{-\mu u} + \int_0^u e^{-\mu(u-v)} d\alpha(v) \right] \times \\ &\times du + \alpha(t) = -\xi(0)(1 - e^{-\mu t}) - \mu \int_0^t e^{\mu v} \int_v^t e^{-\mu u} du d\alpha(v) + \\ &+ \alpha(t) = -\xi(0)(1 - e^{-\mu t}) + \int_0^t (e^{-\mu(t-v)} - 1) d\alpha(v) + \alpha(t) = \\ &= \xi(0) e^{-\mu t} + \int_0^t e^{-\mu(t-v)} d\alpha(v) - \xi(0).\end{aligned}$$

Лемма 3.4. Уравнение (3.24) имеет единственное решение в классе интегрируемых по Риману случайных процессов для произвольного начального значения $\xi(0) = x_0$.

Доказательство. Пусть $\xi_1(t)$, $\xi_2(t)$ — два решения (3.24) с $\xi_1(0) = \xi_2(0) = x_0$. Тогда разность $\xi_0(t) = \xi_1(t) - \xi_2(t)$ удовлетворяет однородному уравнению

$$\xi_0(t) = -\mu \int_0^t \xi_0(u) du. \quad (3.26)$$

Интегрируя равенство (3.26), получаем

$$\begin{aligned}\xi(t) &= -\mu \int_0^t \left(-\mu \int_0^u \xi_0(v) dv \right) du = \mu^2 \int_0^t \int_v^t du \xi_0(v) dv = \\ &= \mu^2 \int_0^t (t-v) \xi_0(v) dv = \dots = (-\mu)^{n+1} \int_0^t \frac{(t-v)^n}{n!} \xi_0(v) dv \rightarrow 0\end{aligned}$$

при $n \rightarrow \infty$.

Следовательно, $\xi_0(t) \equiv 0$, $\xi_1(t) \equiv \xi_2(t)$.

Обозначим $\varphi(s) = M e^{-s\xi_1}$, $\psi(s; t) = M e^{-s\xi(t)}$ и предположим, что $\xi(0) = x_0 \equiv \text{const}$.

Теорема 3.3. Преобразование Лапласа $\psi(s; t)$ имеет вид

$$\psi(s, t) = \exp \left\{ -sx_0 e^{-\mu t} - \int_0^t (1 - \varphi(s e^{-\mu u})) du \right\}. \quad (3.27)$$

Доказательство. Из равенства (3.25) следует

$$\begin{aligned}\ln \psi(s; t) &= \ln M \exp \left\{ -sx_0 e^{-\mu t} - s \int_0^t e^{-\mu(t-u)} d\alpha(u) \right\} = \\ &= -sx_0 e^{-\mu t} + \ln M \exp \left\{ -s \lim_{n \rightarrow \infty} \sum_{k=0}^{n-1} e^{-\mu(t-u_k)} (\alpha(u_{k+1}) - \alpha(u_k)) \right\},\end{aligned}$$

где точки $u_k = u_k(n)$ удовлетворяют условиям

$$0 = u_0 \leq u_1 \leq \dots \leq u_n = t, \quad \max_{\kappa} (u_{k+1} - u_k) \xrightarrow{n \rightarrow \infty} 0,$$

а предел существует с вероятностью единица. Так как $\alpha(t)$ — обобщенный пуассоновский процесс, то приращения $\alpha(u_{k+1}) - \alpha(u_k)$ независимы:

$$\begin{aligned}P\{\alpha(u_{k+1}) - \alpha(u_k) = 0\} &= e^{-\lambda(u_{k+1} - u_k)} = \\ &= 1 - \lambda(u_{k+1} - u_k) + o(u_{k+1} - u_k); \\ P\{\alpha(u_{k+1}) - \alpha(u_k) \geq x\} &= \lambda(u_{k+1} - u_k) \cdot (1 - F(x)) + \\ &+ o(u_{k+1} - u_k).\end{aligned}$$

Поэтому

$$\begin{aligned}\ln \psi(s; t) &= -sx_0 e^{-\mu t} + \\ &+ \ln \lim_{n \rightarrow \infty} \prod_{k=0}^{n-1} M \exp \{ -se^{-\mu(t-u_k)} (\alpha(u_{k+1}) - \alpha(u_k)) \} = \\ &= -sx_0 e^{-\mu t} + \ln \lim_{n \rightarrow \infty} \prod_{k=0}^{n-1} (1 - \lambda(u_{k+1} - u_k) + \lambda(u_{k+1} - u_k) \times \\ &\times M \exp \{ -se^{-\mu(t-u_k)} \xi_1 \}) = -sx_0 e^{-\mu t} + \lim_{n \rightarrow \infty} \sum_{k=0}^{n-1} (-\lambda(u_{k+1} - \\ &- u_k) (1 - \varphi(se^{-\mu(t-u_k)})) = -sx_0 e^{-\mu t} - \\ &- \lambda \int_0^t (1 - \varphi(se^{-\mu(t-u)})) du = -sx_0 e^{-\mu t} - \\ &- \lambda \int_0^t (1 - \varphi(se^{-\mu u})) du.\end{aligned}$$

Теорема 3.4. Собственное предельное при $t \rightarrow \infty$ распределение процесса $\xi(t)$ существует тогда и только тогда, когда

$$M \ln(\xi_1 + 1) < \infty \quad (3.28)$$

и при выполнении условия (3.28) преобразование Лапласа имеет вид

$$\psi(z) = \lim_{t \rightarrow \infty} \psi(s; t) = \exp \left\{ -\frac{\lambda}{\mu} \int_0^s \frac{1 - \varphi(u)}{u} du \right\}. \quad (3.29)$$

Доказательство. Из равенства (3.27) следует, что преобразование Лапласа предельного (возможно, не-собственного) распределения процесса $\xi_1(t)$ равно

$$\begin{aligned}\psi(s) &= \exp \left\{ -\lambda \int_0^{\infty} (1 - \varphi(se^{-\mu u})) du \right\} = \\ &= \exp \left\{ -\frac{\lambda}{\mu} \int_0^s \frac{1 - \psi(u)}{u} du \right\}.\end{aligned}$$

Собственным предельное распределение будет тогда и только тогда, когда $\psi(s) \rightarrow 1$ при $s \downarrow 0$, т. е. когда

$$\int_0^s \frac{1 - \varphi(u)}{u} du \rightarrow 0 \text{ при } s \downarrow 0.$$

Поскольку

$$\begin{aligned}\int_0^s \frac{1 - \varphi(u)}{u} du &= \int_0^{\infty} \int_0^s \frac{1 - e^{-ux}}{u} du dF(x) = \\ &= \int_0^{\varepsilon} \int_0^s \frac{1 - e^{-ux}}{u} du dF(x) + \int_{\varepsilon}^{\infty} \int_0^s \frac{1 - e^{-ux}}{u} du dF(x) = I_1 + I_2,\end{aligned}$$

$$I_1 \leq \int_0^{\varepsilon} \int_0^s \frac{1 - e^{-us}}{u} du dF(x) \leq \int_0^s \frac{1 - e^{-us}}{u} du \xrightarrow{\varepsilon \rightarrow \infty} 0,$$

$$I_2 = \int_0^{\infty} \int_0^{sx} \frac{1 - e^{-v}}{v} dv dF(x),$$

а

$$\int_0^{sx} \frac{1 - e^{-v}}{v} dv = \ln x + o(\ln x) \text{ при } x \rightarrow \infty,$$

то интегрируемость функции $\frac{1 - \varphi(u)}{u}$ в нуле эквивалентна

сходимости интеграла $\int_0^{\infty} \ln x dF(x)$ для любого $\varepsilon > 0$.

Последнее, в свою очередь, эквивалентно условию

$$\int_0^{\infty} \ln(x+1) dF(x) < \infty, \text{ т. е. } M \ln(\xi_1 + 1) < \infty.$$

Теорема доказана.

Из теоремы 3.4 следует, что стационарное распределение процесса $\xi(t)$ зависит от λ и μ только через их отношение $\rho = \lambda/\mu$.

Изучим теперь предельное поведение стационарного распределения при $\rho \rightarrow \infty$. Нам удобно будет ввести случайную величину ζ , имеющую распределение, совпадающее со стационарным для $\xi(t)$. Тогда $Me^{-s\zeta}$ равно правой части равенства (3.29).

Теорема 3.5. Если $M\xi_1 = m < \infty$, то

$$(m\rho)^{-1} \frac{P}{\rho \rightarrow \infty} 1.$$

Доказательство. Нам достаточно проверить, что $\lim_{\rho \rightarrow \infty} Me^{-\frac{s\zeta}{\rho m}} = e^{-s}$. Имеем $\lim_{\rho \rightarrow \infty} Me^{-\frac{s}{m\rho}\zeta} = \lim_{\rho \rightarrow \infty} \exp\left\{-\rho \times \int_0^{s(\rho m)^{-1}} \frac{1 - \varphi(u)}{u} du\right\} = \lim_{\rho \rightarrow \infty} \exp\{-\rho \cdot (-\varphi'(0)) \cdot s(\rho m)^{-1}\} = \exp\{-s\}$.

Теорема 3.6. Если $m_2 = M\xi_1^2 < \infty$, то

$$\sqrt{2} \frac{\zeta - \rho m}{\sqrt{\rho m}} \xrightarrow[\rho \rightarrow \infty]{\text{сл}} N(0, 1).$$

Доказательство. Пусть $\psi_1(s)$ — характеристическая функция ζ ; $\psi_1(s) = Me^{is\zeta}$. Тогда на основании (3.29)

$$\psi_1(s) = \exp\left\{-\rho \int_0^s \frac{1 - \varphi_1(u)}{u} du\right\},$$

где $\varphi_1(s) = Me^{is\xi_1}$ — характеристическая функция ξ_1 . Если $m_2 < \infty$, то $\varphi_1(s) = 1 + ims - \frac{m_2}{2}s^2 + o(s^2)$ при $s \rightarrow 0$. Поэтому при $s \rightarrow 0$

$$\int_0^s \frac{1 - \varphi(u)}{u} du = -ims + \frac{m_2}{4}s^2 + o(s^2)$$

и

$$\begin{aligned} M \exp\left\{is \sqrt{2} \frac{\zeta - \rho m}{\sqrt{\rho m_2}}\right\} &= \psi_1\left(\frac{\sqrt{2}s}{\sqrt{\rho m_2}}\right) \exp\left\{ism \sqrt{\frac{2\rho}{m_2}}\right\} = \\ &= \exp\left\{ism \sqrt{\frac{2\rho}{m_2}} - \frac{s_2}{2} + o(1)\right\} \cdot \exp\left\{-ism \sqrt{\frac{2\rho}{m_2}}\right\} \xrightarrow[\rho \rightarrow \infty]{} \\ &\xrightarrow[\rho \rightarrow \infty]{} \exp\left\{-\frac{s^2}{2}\right\}. \end{aligned}$$

Другими важными характеристиками системы массового обслуживания, которые мы изучим, будут виртуальные времена ожидания начала обслуживания и пребывания в системе. Так как эти величины зависят от дисциплины выбора на обслуживание, то, начиная с этого момента, мы примем, что в системе действует дисциплина «первым прибыл — первым обслужен». Допустим, что требование, поступающее в систему в момент t_0 , нуждается в y единиц работы и $\xi(t_0 - 0) = x$. Обозначим через $w = w(x, y)$ и $\theta = \theta(x, y)$ их времена ожидания и пребывания в системе соответственно. Для нахождения их распределений введем процесс $\eta(t)$ — объем работы в момент времени $t \geq t_0$, необходимой для обслуживания требований, прибывших в систему до момента t_0 . Не ограничивая общности, положим $t_0 = 0$. Тогда из определения процесса $\eta(t)$ следует равенство

$$P\{w > t\} = P\{\eta(t) > 0\}. \quad (3.30)$$

В области $\eta(t) > 0$ справедливо равенство

$$d\eta(t) = -\mu\xi(t)dt. \quad (3.31)$$

Так как $\eta(0) = \xi(0 - 0) = x$; $\xi(0) = x + y$, то из соотношений (3.23), (3.25) и (3.31) следует

$$\begin{aligned} \eta(t) - x &= \xi(t) - \xi(0) - \alpha(t) = (x + y)(\exp\{-\mu t\} - 1) + \\ &+ \int_0^t (\exp\{-\mu(t-u)\} - 1) d\alpha(u); \\ P\{w > t\} &= P\{x + (x + y)(\exp\{-\mu t\} - 1) + \\ &+ \int_0^t (\exp\{-\mu(t-u)\} - 1) d\alpha(u) > 0\} = \\ &= P\{x + (x + y)(\exp\{-\mu t\} - 1) + \int_0^t (\exp\{-\mu u\} - 1) d\alpha(u) > 0\}. \end{aligned} \quad (3.32)$$

Такие же рассуждения для времени θ приводят к равенству

$$\begin{aligned} P\{\theta > t\} &= P\{(x + y)\exp\{-\mu t\} + \\ &+ \int_0^t (\exp\{-\mu u\} - 1) d\alpha(u) > 0\}. \end{aligned} \quad (3.33)$$

Рассмотрим теперь виртуальные времена ожидания W и время пребывания T в системе, находящейся в стационарном режиме. Их распределения задаются выражениями

(3.32) и (3.33), в которых x имеет распределение, совпадающее со стационарным для $\xi(t)$, а функция распределения y равна F . Таким образом,

$$P\{W > t\} = P\left\{y(\exp\{-\mu t\} - 1) + \int_{-\infty}^0 \exp\{-\mu(t-u)\} d\alpha(u) + \int_0^t (\exp\{-\mu u\} - 1) d\alpha(u) > 0\right\}; \quad (3.34)$$

$$P\{T > t\} = P\left\{y \exp\{-\mu t\} + \int_{-\infty}^0 \exp\{\mu u\} d\alpha(u) + \int_0^t (\exp\{-\mu u\} - 1) d\alpha(u) > 0\right\}. \quad (3.35)$$

Формулы (3.34) и (3.35) позволяют доказать следующий результат.

Теорема 3.7. Если $M\xi_1 = m < \infty$, то при

$$\lambda \rightarrow \infty, \mu \equiv \text{const}, \mu W \xrightarrow{\text{сл}} 1 \text{ и } \mu T \xrightarrow{\text{сл}} 1.$$

Доказательство. Ввиду идентичности доказательств, проверим утверждение только для μW . Заметим, что в правой части формулы (3.34) обе части можно умножить на λ^{-1} . Переходя после этого к пределу при $\lambda \rightarrow \infty$, получим

$$\lambda^{-1} y (\exp\{-\mu t\} - 1) \xrightarrow{P} 0,$$

а в силу закона больших чисел для процесса $\alpha(t)$ с параметром λ

$$\begin{aligned} \lambda^{-1} \exp\{-\mu t\} \int_{-\infty}^0 \exp\{\mu u\} d\alpha(u) &\xrightarrow{P} \frac{m}{\mu} \exp\{-\mu t\}; \\ \lambda^{-1} \int_0^t (\exp\{-\mu u\} - 1) d\alpha(u) &\xrightarrow{P} \frac{m}{\mu} (1 - \exp\{-\mu t\}) - mt \end{aligned}$$

и, следовательно, при $t \neq \mu^{-1}$

$$\begin{aligned} P\{W > t\} &\xrightarrow{\lambda \rightarrow \infty} P\left\{\frac{m}{\mu} \exp\{-\mu t\} + \right. \\ &\left. + \frac{m}{\mu} (1 - \exp\{-\mu t\}) - mt > 0\right\} = P\left\{\frac{1}{\mu} > t\right\}. \end{aligned}$$

Поэтому $P\{\mu W > t\} \xrightarrow{\lambda \rightarrow \infty} P\{1 > t\}$, $t \neq 1$ и

$$\mu W \xrightarrow[\lambda \rightarrow \infty]{\text{сл}} 1.$$

Теорема доказана.

ГЛАВА 4. СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ В УСЛОВИЯХ БОЛЬШОЙ ЗАГРУЗКИ

4.1. Система $M|G|1$ в критическом режиме

В предыдущей главе исследовался стационарный режим ряда СМО. При этом необходимым и достаточным условием существования стационарного режима было условие $\rho < 1$, где ρ — загрузка системы. Для системы $M|M|s$ $\rho = \frac{\lambda}{s\mu}$, для $M|G|1$ $\rho = \lambda m$. В общем случае, ρ представляет собой отношение среднего времени между двумя последовательно обслуженными требованиями при полностью занятой системе к среднему времени между двумя последовательными поступлениями требований. Интуитивно ясно, что при $\rho \nearrow 1$ значения таких случайных величин в стационарном режиме как величина очереди и время ожидания, будут неограниченно возрастать по вероятности. Поэтому возникает естественная проблема — указать порядок роста в зависимости от скорости стремления ρ к единице и найти предельное распределение указанных величин при соответствующей нормировке. В данном параграфе эта проблема описывается для системы $M|G|1$.

Обозначим через Q_p и W_p случайные величины, распределения которых совпадают с распределением величины очереди и времени ожидания соответственно для системы $M|G|1$ в стационарном режиме, рассмотренной в п. 2.2 (при этом, естественно, предполагается, что $\rho = \lambda m < 1$). Пусть, как и прежде, $G(x)$ — функция распределения времени обслуживания в системе $M|G|1$, на вход которой поступает простейший поток с параметром λ . Положим

$$G^*(s) = \int_0^{\infty} e^{-sx} dG(x) \quad (s \geq 0) \quad (G^*(s) \text{ — преобразование Лапласа}$$

времени обслуживания. Предположим, что время обслуживания притягивается к некоторому устойчивому закону

с параметром α . Поскольку у нас $m = \int_0^{\infty} x dG(x) < \infty$, то $1 < \alpha \leq 2$ и при этом, как известно [13, с. 514], при $x \rightarrow +\infty$

$$1 - G(x) \sim x^{-\alpha} L(x),$$

где $L(x)$ — медленно меняющаяся функция (т. е. для любого $c > 0$ $L(cx)/L(x) \rightarrow 1$ при $x \rightarrow +\infty$). Отсюда в силу тауберовых теорем [13, с. 508] следует известное представление в окрестности нуля для преобразования Лапласа $G^*(s)$:

$$G^*(s) = 1 - sm + ds^{\alpha} L\left(\frac{1}{s}\right) (1 + o(1)), \quad (4.1)$$

где $d > 0$, а $o(1) \rightarrow 0$ при $s \rightarrow 0$. Для простоты предположим, что $L(x) \equiv 1$.

Имеет место.

Теорема 4.1. Если параметр входящего потока λ меняется таким образом, что $\rho < 1$, $\rho \rightarrow 1$, и выполнено (4.1), то распределение случайной величины $(1 - \rho)^{\frac{1}{\alpha-1}} W_{\rho}$ слабо сходится к распределению в преобразовании Лапласа $\left(1 + \frac{d}{m} s^{\alpha-1}\right)^{-1}$.

Доказательство. Согласно формуле Хинчина (2.12)

$$\begin{aligned} Me^{-sw_{\rho}} &= \frac{s(1 - \lambda m)}{s + \lambda(C^*(s) - 1)} = \frac{1 - \rho}{1 + \lambda \frac{G^*(s) - 1}{s}} = \\ &= \frac{1 - \rho}{1 - \rho + \lambda ds^{\alpha-1}(1 + o(1))}. \end{aligned}$$

Пользуясь (4.1), получим, что при $\rho \nearrow 1$

$$\begin{aligned} M \exp \left\{ -s (1 - \rho)^{\frac{1}{\alpha-1}} W_{\rho} \right\} &= \\ &= \frac{1 - \rho}{1 - \rho + \lambda ds^{\alpha-1} (1 - \rho)(1 + o(1))} \rightarrow \left(1 + \frac{d}{m} s^{\alpha-1}\right)^{-1}, \end{aligned}$$

поскольку $\lambda \rightarrow \frac{1}{m}$. Так как полученная функция непрерывна по s , то по теореме непрерывности следует наше утверждение.

Замечание. Если $L(x) \not\equiv 1$, то в этом случае утверждение теоремы справедливо для величины $\gamma_{\rho} \theta_{\rho}$, где множитель γ_{ρ} выбирается так, чтобы $\gamma_{\rho}^{\alpha-1} L(\gamma_{\rho}^{-1}) \sim 1 - \rho$.

Замечание. Под сходимостью последовательности процессов $\xi_n(t)$, $n \geq 1$ в равномерной топологии к процессу $\xi(t)$ на отрезке $[0, T]$ понимается, что конечномерные распределения $\xi_n(t)$ слабо сходятся к конечномерным распределениям $\xi(t)$ на $[0, T]$ при $n \rightarrow \infty$ и для любого функционала $f(\cdot)$, заданного на пространстве $C_{[0, T]}$ непрерывных на $[0, T]$ функций и непрерывного в равномерной метрике, распределение $f(\xi_n(\cdot))$ слабо сходится к $f(\xi(\cdot))$.

Отметим также, что предельные распределения времени ожидания и времени пребывания в системе в стационарном режиме совпадают, поскольку допредельные величины отличаются друг от друга на случайную величину, равную времени обслуживания, которая при умножении на нормирующий множитель стремится по вероятности к нулю.

Сформулируем теперь аналогичный результат для величины очереди Q_ρ .

Теорема 4.2. Если $\rho < 1$, $\rho \rightarrow 1$ и выполнено (4.1), то распределение случайной величины $(1 - \rho)^{\frac{1}{\alpha-1}} Q_\rho$ слабо сходится к распределению с преобразованием Лапласа

$$\left(1 + \frac{d}{m\alpha} s^{\alpha-1}\right)^{-1}.$$

Доказательство. Если $P(z)$ — производящая функция, а $\varphi(s)$ — преобразование Лапласа некоторой случайной величины, то очевидно, $\varphi(s) = P(e^{-s})$. Поэтому, воспользовавшись формулой Поллачека — Хинчина (2.11), получим

$$M \exp \{-s\theta_\rho\} = \frac{G^*(\lambda(1 - e^{-s}))(1 - e^{-s})(1 - \rho)}{G^*(\lambda(1 - e^{-s})) - e^{-s}}. \quad (4.2)$$

Как и ранее, для простоты положим $L(x) \equiv 1$. Тогда, поскольку $e^{-s} = 1 - s + \frac{s^2}{2} + o(s^2)$, из (4.1) получим

$$G^*(\lambda(1 - e^{-s})) = \begin{cases} 1 - \lambda ms + d\lambda^\alpha s^\alpha (1 + o(1)), & \alpha < 2; \\ 1 - \lambda ms + \left(d\lambda^2 + \frac{\lambda m}{2}\right) s^2, & \alpha = 2. \end{cases}$$

Теперь из (4.2) выходит, что

$$M \exp \{-s\theta_\rho\} = \frac{s(1 - \rho)(1 + o(1))}{s(1 - \rho) + s^\alpha c_\alpha + o(s^\alpha)},$$

где $c_\alpha = d\lambda^\alpha$ при $\alpha < 2$, и $d\lambda^2 - \frac{1}{2}(1 - \rho)$ при $\alpha = 2$.

Отсюда окончательно получим

$$M \exp \left\{ - (1 - \rho)^{\frac{1}{\alpha-1}} Q_\rho \right\} = \\ = \frac{(1 - \rho)(1 + o(1))}{1 - \rho + (1 - \rho) s^{\alpha-1} c_\alpha + o(1 - \rho) \left(1 + \frac{d}{m^\alpha} s^{\alpha-1} \right)^{-1}}.$$

Следствие 4.1. Если $\int_0^\infty x^2 dG(x) = \mu_2 < \infty$, то при $\rho \rightarrow 1$

$$P \{ (1 - \rho) W_\rho < x \} \rightarrow 1 - \exp \left\{ - \frac{2m}{\mu_2} x \right\}, \quad x \geq 0;$$

$$P \{ (1 - \rho) Q_\rho < x \} \rightarrow 1 - \exp \left\{ - \frac{2m^2}{\mu^2} x \right\}, \quad x \geq 0,$$

т. е. предельные распределения величин W_ρ и Q_ρ будут показательными с параметрами $\frac{2m}{\mu_2}$ и $\frac{2m^2}{\mu^2}$ соответственно. Действительно, если $\mu_2 < \infty$, то $\alpha = 2$,

$$G^*(s) = 1 - sm + \frac{1}{2} \mu_2 s^2 + o(s^2),$$

а преобразования Лапласа предельных распределений для величин

$$(1 - \rho) W_\rho \text{ и } (1 - \rho) Q_\rho \text{ будут } \left(1 + \frac{\mu_2}{2m} s \right)^{-1}$$

и $\left(1 + \frac{\mu_2}{2m^2} s \right)^{-1}$ соответственно, т. е. отвечают показательным законам с параметрами $\frac{2m}{\mu_2}$ и $\frac{2m^2}{\mu^2}$.

Заметим, что в п. 2.1 было показано, что для системы $M|M|1$ в стационарном режиме время ожидания имело показательное распределение со скачком в нуле, а величина очереди — геометрическое. Из полученных результатов вытекает, что для системы $M|G|1$ при $\mu_2 < \infty$ и ρ , достаточно близком к единице, время ожидания и величина очереди также будут иметь распределения, близкие к показательному.

Укажем на интересную вероятностную интерпретацию полученных распределений. Пусть $\xi(t)$ — однородный монотонно неубывающий процесс с независимыми приращениями такой, что $\xi(0) = 0$ и $M \exp \{ -s \xi(t) \} = \exp \{ -ds^{\alpha-1}t \}$, $d > 0$, а η — независимая от $\xi(t)$ случайная величина,

имеющая показательное распределение с параметром единица. Рассмотрим величину $\xi(\eta)$, которая получается остановкой $\xi(t)$ в момент η . Легко проверить, что

$$Me^{-s\xi(\eta)} = \int_0^{\infty} e^{-ds^{\alpha-1}t} e^{-t} dt = (1 + ds^{\alpha-1})^{-1}.$$

Поэтому предельные распределения, появившиеся в наших теоремах, принадлежат классу распределений вида $\xi(\eta)$.

В заключение отметим, что в силу результатов п. 2.4 в стационарном режиме время ожидания в системе $G|G|1$ эквивалентно случайной величине $\sup_{k \geq 0} \{S_k\}$,

где $S_0 = 0$, $S_k = \sum_{i=1}^k (v_i - u_i)$, u_i — промежутки между поступлениями требований, а v_i — времена обслуживания. Из общей теории предельных теорем для случайных процессов известно, что в предположении $M(v_i - u_i) = \alpha$,

$D(v_i - u_i) = \sigma^2$, распределения процессов вида $\frac{1}{\sqrt{n}}(S_{[nt]} - n\alpha)$ слабо сходятся к распределению процесса $\sigma W(t)$ на любом промежутке $[0, T]$, где $w(t)$ — стандартный винеровский процесс, при этом сходятся также распределения достаточно хороших функционалов, в частности супремума. Этот подход был использован Ю. В. Прохоровым и А. А. Боровковым для доказательства сходимости нормированного времени ожидания в достаточно общих системах обслуживания в условиях большой загрузки и предположениях типа применимости центральной предельной теоремы к суммам S_n , к распределению случайной величины вида $\sup_{t \geq 0} (\sigma w(t) - t)$. В нашей более частной схеме это соответствует случаю $\alpha = 2$ и показательности предельного распределения.

4.2. Диффузионная аппроксимация систем массового обслуживания

В последнее время появилось много работ по применению метода диффузионной аппроксимации в теории массового обслуживания. С помощью аппроксимативного метода изучаются различные системы массового обслуживания и всевозможные параметры, характеризующие отдельно взятую систему.

Исследования по диффузионной аппроксимации можно разбить на два направления. Работы первого направления основаны на простой замене дискретного процесса, связан-

ного с некоторой системой массового обслуживания, непрерывным диффузионным процессом. Критерием при выборе подходящего диффузионного процесса служит близость локальных характеристик (в терминах приращений) исходного и аппроксимативного процессов, поведение исходного процесса вблизи границ, а также сравнение с результатами, которые получены другим методом (например, методом имитации). Повышенный интерес к методу диффузионной аппроксимации объясняется тем, что использование диффузионных процессов при решении задач массового обслуживания, особенно при исследовании сложных систем, значительно упрощает математические выкладки. Применение диффузионной аппроксимации возможно и в том случае, когда прямые методы либо неприменимы, либо еще не созданы для данного класса задач. Аппроксимативный диффузионный процесс оказывается полезным также тогда, когда прямые методы позволяют определить поведение системы в стационарном режиме, например, стационарное распределение длины очереди, но мала скорость сходимости к этому стационарному распределению.

Второе направление исследований по диффузионной аппроксимации представлено в работах Ю. В. Прохорова и А. А. Боровкова, в которых дается строгое математическое обоснование сходимости нормированного процесса обслуживания к диффузионному процессу. Изучается также сходимость функционалов. Как показали эти исследования системы массового обслуживания, для которых возможна диффузионная аппроксимация, в том или ином смысле близки к критическим. Сюда можно отнести системы с очередью и загрузкой, близкой к 1, системы с интенсивным входным потоком и большим числом каналов обслуживания и целый ряд других.

Изучим методом диффузионной аппроксимации число занятых приборов в системе $M|M|\infty$, а в п. 4.3 величину очереди в системе $GI|G|r$.

Рассмотрим систему $M|M|\infty$, на которую поступает пуассоновский поток заявок с параметром λ , а время обслуживания имеет одно и тоже экспоненциальное распределение с параметром μ ; число обслуживающих приборов бесконечно. Через $Q(t)$ будем обозначать число занятых приборов в системе. В п. 2.1 для производящей функции числа занятых приборов

$$P(z, t) = \sum_{j=0}^{\infty} P\{Q(t) = j / (Q_0) = i\} z^j = \sum_{j=0}^{\infty} P_j(t) z^j \quad (4.3)$$

получено явное выражение

$$P(z, t) = [1 - (1 - z) e^{-\mu t}]^i \exp \left\{ -\frac{\lambda}{\mu} (1 - z)(1 - e^{-\mu t}) \right\}. \quad (4.4)$$

В дальнейшем будем предполагать, что система $M|M|_{\infty}$ зависит от некоторого параметра n (параметра «серии»), т. е.

$$\lambda = \lambda_n, \mu = \mu_n, Q(t) = Q^{(n)}(t).$$

Рассмотрим последовательность процессов

$$\xi_n(t) = \frac{Q^{(n)}(t) - \lambda_n/\mu_n}{\sqrt{n}}, \quad n \geq 1, t \in [0, T];$$

$$Q^{(n)}(0) = \lambda_n/\mu_n + \sqrt{n}a, \quad a \in (-\infty, \infty). \quad (4.5)$$

При $\lambda_n = \lambda + o(1)$, $\mu_n = n^{-1}\mu + o(n^{-1})$ система $M^{(n)} \times \times |M^{(n)}|_{\infty}$ находится в критическом режиме и возможна аппроксимация $\xi_n(t)$, $n \geq 1$ диффузионным процессом.

Под диффузионным процессом $\xi(t)$, $t \in [0, T]$ будем понимать марковский процесс, у которого вероятность перехода $P(t, x, s, A)$ ($0 \leq t < s < T$, $x \in (-\infty, \infty)$, A — борелевское множество) удовлетворяет следующим свойствам:

1) для всякого $\varepsilon > 0$ и $t \in [0, T]$, $x \in (-\infty, \infty)$

$$\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{|x-y|>\varepsilon} P(t, x, t + \Delta t, dy) = 0;$$

2) существуют такие функции $a(t, x)$ и $b(t, x)$, что для всех $\varepsilon > 0$, $t \in [0, T]$, $x \in (-\infty, \infty)$ выполнены соотношения

$$\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{|x-y|<\varepsilon} (y-x) P(t, x, t + \Delta t, dy) = a(t, x);$$

$$\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int_{|x-y|<\varepsilon} (y-x)^2 P(t, x, t + \Delta t, dy) = b(t, x).$$

Коэффициенты $b(t, x)$ и $a(t, x)$ называются соответственно коэффициентами диффузии и переноса.

Замечание. Если при некотором $\delta > 0$ марковский процесс $\xi(t)$ $t \in [0, T]$ удовлетворяет условию

$$\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int |x-y|^{2+\delta} P(t, x, t + \Delta t, dy) = 0,$$

то для того, чтобы процесс $\xi(t)$ был диффузионным, достаточно,

чтобы существовали такие функции $a(t, x)$ и $b(t, x)$, что для всех t, x выполняются соотношения

$$\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int (y - x) P(t, x, t + \Delta t, dy) = a(t, x);$$

$$\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \int (y - x)^2 P(t, x, t + \Delta t, dy) = b(t, x).$$

При исследовании предельного поведения последовательности процессов $\xi_n(t)$, $n \geq 1$ применим локальный подход, который учитывает свойства диффузионных процессов и основывается на близости локальных характеристик (в терминах приращений) допредельного и предельного процессов.

Впервые конкретные случаи сходимости к диффузионным процессам описали С. Н. Бернштейн и А. Я. Хинчин. В дальнейшем общие условия сходимости к диффузионным и более общим марковским процессам были получены И. И. Гихманом и А. В. Скороходом.

Приведем один вариант теоремы И. И. Гихмана и А. В. Скорохода сходимости последовательности цепей Маркова $\xi_n(t)$, $t \in [0, T]$, $n \geq 1$ к диффузионному процессу $\xi(t)$, $t \in [0, T]$.

Теорема 4.3. Пусть $\xi(t)$, $t \in [0, T]$ — диффузионный процесс с коэффициентами переноса и диффузии, равными соответственно $a(t, x)$ и $b(t, x) = \sigma^2(t, x)$. Функции $a(t, x)$ и $\sigma(t, x)$ — непрерывные по совокупности переменных и удовлетворяют следующему условию.

Существует такая постоянная K , что при $t \in [0, T]$, $x, y \in (-\infty, \infty)$ справедливы неравенства

$$|a(t, x) - a(t, y)| + |\sigma(t, x) - \sigma(t, y)| \leq K|x - y|;$$

$$a^2(t, x) + \sigma^2(t, x) \leq K^2(1 + x^2). \quad (4.6)$$

Для того, чтобы последовательность процессов $\xi_n(t)$, $n \geq 1$, сходилась в равномерной топологии к процессу $\xi(t)$, достаточно, чтобы существовала последовательность разбиений отрезка $[0, T]$

$$\{t_{nk}, k = 0, 1, \dots, m_n; n = 1, 2, \dots\},$$

$$0 = t_{n0} < t_{n1} < \dots < t_{nm_n} = T,$$

$$\lim_{n \rightarrow \infty} \max_k \Delta t_{nk} = 0, \quad \Delta t_{nk} = t_{nk+1} - t_{nk}$$

такая, что

$$1) \lim_{n \rightarrow \infty} \sum_{k=0}^{m_n-1} M((a_n(t_{nk}, \xi_{nk}) - a(t_{nk}, \xi_{nk}))^2 +$$

$$+ (\sigma_n(t_{nk}, \xi_{nk}) - \sigma(t_{nk}, \xi_{nk}))^2) \Delta t_{nk} = 0,$$

где

$$a_n(t_{nk}, x) = \frac{1}{\Delta t_{nk}} M(\xi_{n,k+1} - x / \xi_{nk} = x),$$

$$\sigma_{nk}^2(t_{nk}, x) = \frac{1}{\Delta t_{nk}} M((\xi_{n,k+1} - x)^2 / \xi_{nk} = x),$$

$$\xi_{nk} = \xi_n(t_{nk}), k = \overline{0, m_n};$$

2) при некотором $\delta > 0$

$$\lim_{n \rightarrow \infty} \sum_{k=0}^{m_n-1} M |\xi_{n,k+1} - \xi_{nk}|^{2+\delta} = 0$$

и

$$\sup_{k \geq 0} \sum_{i=k}^{m_n-1} M (|\xi_{n,i+1} - \xi_{ni}|^{2+\delta} / \xi_{nk}) \xrightarrow[n \rightarrow \infty]{P} 0;$$

3) функции $\frac{1}{\sigma_n(t_{nk}, x)}$ и $\frac{a_n(t_{nk}, x)}{\sigma_n(t_{nk}, x)}$ равномерно по x

ограничены относительно n ;

4) распределение $\xi_n(0)$ слабо сходится к распределению некоторой случайной величины $\xi(0)$.

Замечание. Условие 3) можно заменить более слабым условием 3') $\frac{1}{\sigma_n(t_{nk}, x)}, \sup_k M |a_n(t_{nk}, \xi_{nk})|^{2+\delta}$ — равномерно по x ограничены относительно n (δ — любое положительное число, для которого выполнено условие 2)).

Предельное поведение последовательности процессов $\xi_n(t)$, $n \geq 1$, изучено в следующей теореме.

Теорема 4.4. Предположим, что в системе $M^{(n)} | M^{(n)} | \infty$ параметры λ_n и μ_n изменяются таким образом, что

$$\lambda_n = \lambda + o(1); \mu_n = n^{-1}\mu + o(n^{-1});$$

$\xi_n(t)$, $n \geq 1$ нормированный процесс числа занятых приборов, определяемый равенством (4.5). Тогда последовательность процессов $\xi_n(t)$, $n \geq 1$, сходится в равномерной топологии к диффузионному процессу $\xi(t)$, $t \in [0, T]$ с коэффициентом переноса $a(t, x)$ и диффузии $b(t, x)$, равными соответственно $-\mu x$ и 2λ .

Доказательство. Условие (4.6) на коэффициенты $a(t, x)$ и $b(t, x)$, очевидно, выполнено. Не нарушая общности, будем считать, что $\lambda_n = \lambda$, $\mu_n = \frac{1}{n}\mu$. Используя

явное выражение для производящей функции числа занятых приборов (4.4), нетрудно подсчитать, что

$$\begin{aligned} M(\xi_{n,k+1} - x / \xi_{nk} = x) &= -x(1 - e^{-\mu \Delta t_{nk}}); \\ M((\xi_{n,k+1} - x)^2 / \xi_{nk} = x) &= \frac{\lambda}{\mu} (1 + e^{-\mu \Delta t_{nk}}) \times \\ &\times (1 - e^{-\mu \Delta t_{nk}}) + n^{-\frac{1}{2}} x e^{-\mu \Delta t_{nk}} (1 - e^{-\mu \Delta t_{nk}}) + \\ &+ x^2 (1 - e^{-\mu \Delta t_{nk}})^2; \\ \sup_{t \in [0, T]} M \xi_n^2(t) &< L_1, \quad \sup_{t \in [0, T]} M \xi_n^4(t) < L_2. \end{aligned}$$

Для того чтобы проверить условие 1), достаточно убедиться, что равномерно по $k = 0, 1, \dots, m_n - 1$:

$$M(a_n(t_{nk}, \xi_{nk}) - a(t_{nk}, \xi_{nk}))^2 \rightarrow 0 \quad (4.7)$$

и

$$M(\sigma_n(t_{nk}, \xi_{nk}) - \sigma(t_{nk}, \xi_{nk}))^2 \rightarrow 0. \quad (4.8)$$

Поскольку

$$\lim_{n \rightarrow \infty} \max_k \Delta t_{nk} = 0,$$

то

$$\begin{aligned} e^{-\mu \Delta t_{nk}} &= 1 - C'_{nk} \Delta t_{nk}; \\ e^{-\mu \Delta t_{nk}} &= 1 - \mu \Delta t_{nk} + C''_{nk} (\Delta t_{nk})^2, \end{aligned} \quad (4.9)$$

где $C'_{nk} \rightarrow \mu$; $C''_{nk} \rightarrow \frac{1}{2} \mu^2$ при $n \rightarrow \infty$ равномерно по $k = 0, 1, \dots, m_n - 1$.

Таким образом,

$$\begin{aligned} M(a_n(t_{nk}, \xi_{nk}) - a(t_{nk}, \xi_{nk}))^2 &= M\left(-\frac{1}{\Delta t_{nk}} \xi_{nk} (1 - \right. \\ &\quad \left. - e^{-\mu \Delta t_{nk}}) + \mu \xi_{nk}\right)^2 = M(\xi_{nk} (\mu - C''_{nk} \Delta t_{nk}) - \\ &\quad - \mu \xi_{nk})^2 = (C''_{nk})^2 (\Delta t_{nk})^2 M \xi_{nk}^2 \leq L_1 (C''_{nk})^2 (\Delta t_{nk})^2; \quad (4.10) \\ \sigma_n^2(t_{nk}, x) &= \frac{\lambda}{\mu} (1 + e^{-\mu \Delta t_{nk}}) \frac{1 - e^{-\mu \Delta t_{nk}}}{\Delta t_{nk}} + \\ &+ \frac{1 - e^{-\mu \Delta t_{nk}}}{\Delta t_{nk}} \left(x n^{-\frac{1}{2}} e^{-\mu \Delta t_{nk}} + x^2 (1 - e^{-\mu \Delta t_{nk}})\right). \end{aligned}$$

Обозначим через

$$f_{nk}(x) = x n^{-\frac{1}{2}} e^{-\mu \Delta t_{nk}} + x^2 (1 - e^{-\mu \Delta t_{nk}}).$$

Используя (4.9), представим $\sigma_n^2(t_{nk}, x)$ в виде

$$\sigma_n^2(t_{nk}, x) = 2\lambda + b_{nk} \Delta t_{nk} + (\mu - C''_{nk} \Delta t_{nk}) f_{nk}(x),$$

где

$$b_{nk} = -2 \frac{\lambda}{\mu} C''_{nk} - \lambda C_{nk} \left(1 - \frac{C''_{nk}}{\lambda} \Delta t_{nk} \right).$$

Нетрудно видеть, что

$$\inf_{x \in (-\infty, \infty)} f_{nk}(x) = - \frac{e^{-2\mu \Delta t_{nk}}}{4n(1 - e^{-\mu \Delta t_{nk}})}.$$

Выбор последовательности разбиений $\{t_{nk}, k = 0, 1, \dots, m_n, n = \overline{1, \infty}\}$ отрезка $[0, T]$ находится в нашем распоряжении. Важно только, чтобы $\lim_{n \rightarrow \infty} \max_k \Delta t_{nk} = 0$. Выберем теперь последовательность разбиений таким образом, чтобы

$$\lim_{n \rightarrow \infty} \min_k n \Delta t_{nk} = \infty.$$

В качестве такой последовательности разбиений можно взять последовательность равномерных разбиений $[0, T]$ на отрезки длиной $[n^\varepsilon]^{-1}$ ($0 < \varepsilon < 1$, $[\cdot]$ — целая часть). При таком выборе $\{t_{nk}, k = 0, 1, \dots, m_n, n = \overline{1, \infty}\}$:

- а) $\inf_{x \in (-\infty, \infty)} f_{nk}(x) \rightarrow 0$ при $n \rightarrow \infty$ равномерно по $k = 0; 1, \dots, m_n - 1$;
- б) условие 3) выполнено при $\delta = 2$;
- в) не нарушая общности, можно считать, что $\sigma_n^2(t_{nk}, x) \geq \lambda$ для всех $x \in (-\infty, \infty)$.

Теперь нетрудно получить необходимую оценку для

$$\begin{aligned} & M(\sigma_n(t_{nk}, \xi_{nk}) - \sigma(t_{nk}, \xi_{nk}))^2. \\ & M(\sigma_n(t_{nk}, \xi_{nk}) - \sigma(t_{nk}, \xi_{nk}))^2 = \\ & = M\left(\frac{\sigma_n^2(t_{nk}, \xi_{nk}) - \sigma^2(t_{nk}, \xi_{nk})}{\sigma_n(t_{nk}, \xi_{nk}) + \sigma(t_{nk}, \xi_{nk})}\right)^2 \leq \\ & \leq (1 + \sqrt{2})^{-2} \lambda^{-1} M(\sigma_n^2(t_{nk}, \xi_{nk}) - \sigma^2(t_{nk}, \xi_{nk}))^2 = \\ & = (1 + \sqrt{2})^{-2} \lambda^{-1} M(b_{nk} \Delta t_{nk} + (\mu - C''_{nk} \Delta t_{nk}) f_{nk}(\xi_{nk}))^2 = \\ & = (1 + \sqrt{2})^{-2} \lambda^{-1} M(b_{nk} \Delta t_{nk} + (\mu - C''_{nk}) n^{-\frac{1}{2}} \times \\ & \times e^{-\mu \Delta t_{nk}} \xi_{nk} + \xi_{nk} (1 - e^{-\mu \Delta t_{nk}})^2)^2. \end{aligned}$$

Отсюда следует, что

$$M(\sigma_n(t_{nk}, \xi_{nk}) - \sigma(t_{nk}, \xi_{nk}))^2 \leq C_1 n^{-1} + C_2 (\Delta t_{nk})^2, \quad (4.11)$$

где C_1, C_2 — некоторые постоянные. Из оценок (4.10),

(4.11) следует (4.7), (4.8). Таким образом, условие 1 выполнено. Условие 2 при $\delta = 2$ проверяется аналогично, так как

$$M(\xi_{n,k+1} - \xi_{nk})^4 \leq C(\Delta t_{nk})^2,$$

где C — некоторая постоянная. Теорема доказана.

Диффузионный процесс $\xi(t)$, $t \in [0, T]$ с коэффициентом переноса и диффузии, равными соответственно $-\mu x$ и 2λ , является процессом Орнштейна — Уленбека. Переходные плотности $u(x, t, y)$ процесса Орнштейна — Уленбека нормальны с математическим ожиданием $xe^{-\mu t}$ и дисперсией $\frac{\lambda}{\mu}(1 - e^{-2\mu t})$.

4.3. Диффузионная аппроксимация величины очереди в системе GI|G|r

Рассмотрим систему GI|G|r с ограниченным числом каналов ($r < \infty$). В случае, когда загрузка системы $\rho \geq 1$, величина очереди $Q(t)$ неограниченно возрастает при $t \rightarrow \infty$ и можно применять теоремы об аппроксимации $Q(t)$ диффузионными процессами. Если параметры системы меняются так, что $\rho > 1$ и $\rho \rightarrow 1$, то вид аппроксимации может зависеть от взаимной скорости стремления ρ к единице, а t к бесконечности. Ряд достаточно тонких результатов в этом направлении получен А. А. Боровковым. При этом метод доказательства основан на асимптотическом исследовании поведения приращений входящих выходящих потоков и идейно близок к результатам И. И. Гихмана, А. В. Скорохода. Здесь ограничимся более простой ситуацией, когда параметры системы не меняются, а $\rho > 1$.

Предположим, что на вход системы поступает рекуррентный ординарный поток требований (поток восстановления) с интервалом восстановления с функцией распределения $F(x)$, а система содержит r различных приборов с функциями распределения времен обслуживания $G_1(x), \dots, G_r(x)$. Поступившее в систему требование занимает любой свободный прибор либо становится в общую очередь, если все приборы заняты.

Пусть $x(t)$, $t \geq 0$ — входящий поток, а $y_i(t)$, $t \geq 0$ — поток восстановления с интервалом восстановления с функцией распределения $G_i(x)$, $i = \overline{1, r}$. Через $y(t) = \sum_{i=1}^r y_i(t)$ обозначим поток обслуживания (выходящий поток требований), который генерирует система, если все приборы все время заняты.

Положим

$$a = \int_0^{\infty} x dF(x); \quad b_i = \int_0^{\infty} x dG_i(x), \quad i = \overline{1, r}.$$

Под загрузкой нашей системы будем понимать величину

$$\rho = \frac{1}{a} \left(\sum_{i=1}^r \frac{1}{b_i} \right)^{-1}.$$

Условие $\rho > 1$ эквивалентно тому, что

$$\frac{1}{a} > \sum_{i=1}^r \frac{1}{b_i}. \quad (4.12)$$

Заметим теперь, что если с момента $t = 0$ все приборы были заняты и ни один не простаивал, то $Q(t) = Q(0) + x(t) - y(t)$. В общем случае, как легко понять, имеет место неравенство $Q(t) \geq Q(0) + x(t) - y(t)$, которое можно понимать в том смысле, что для любого $x > 0$

$$P\{Q(t) \geq x\} \geq P\{Q(0) + x(t) - y(t) \geq x\}. \quad (4.13)$$

Это непосредственно следует из тех соображений, что если $\Omega_i(t)$ — суммарное время простоя i -го прибора на промежутке $[0, t]$, то количество требований, обслуженных i -м прибором $y_i(t)$ на промежутке $[0, t]$, есть величина $y_i(t - \Omega_i(t))$. Тем самым общее количество требований, обслуженных системой, есть величина

$$\tilde{y}(t) = \sum_{i=1}^r y_i(t - \Omega_i(t)) \leq y(t).$$

Так как $Q(t) = Q(0) + x(t) - \tilde{y}(t)$, то отсюда следует (4.13).

Теперь, как следует из предельных теорем для потоков восстановления (см. п. 1.4), при $t \rightarrow \infty$

$$\frac{1}{t} x(t) \rightarrow \frac{1}{a}, \quad \frac{1}{t} y_i(t) \rightarrow \frac{1}{b_i}, \quad i = \overline{1, r}, \quad (4.14)$$

при этом сходимость, как нетрудно доказать, используя усиленный закон больших чисел, имеет место с вероятностью единица.

Поэтому из (4.12) вытекает, что

$$\frac{1}{t} (x(t) - y(t)) \rightarrow \frac{1}{a} - \sum_{i=1}^r \frac{1}{b_i} > 0,$$

т. е. $x(t) - y(t) \rightarrow \infty$ с вероятностью единица. Это означает, что существует некоторый собственный случайный момент времени κ , такой, что $x(t) - y(t) \geq r$ для всех $t \geq \kappa$, т. е. через время κ ни один из приборов нашей системы не будет затем простаивать.

Введем теперь следующие обозначения. Пусть $z(t)$ — простой поток восстановления с функцией распределения интервала восстановления $A(x)$, а $\tilde{z}(t, v)$ — обобщенный поток, у которого функция распределения начального промежутка имеет вид $1 - (1 - A(x + v))(1 - A(v))^{-1}$. Обозначим

$$z(t, u, v) = \tilde{z}(t - u, v), \quad t \geq u \geq v.$$

В сущности поток $z(t, u, v)$ получается при наблюдении потока $z(t)$ на промежутке $[u, t]$ в предположении, что величина недоскока потока $z(t)$ до момента u равна v (величина недоскока — это промежуток времени от последнего перед t момента прихода требования до момента t). Поэтому очевидно, что распределение величины перескока (промежутка времени от момента t до момента следующего поступления требования) при условии, что недоскок равен v , имеет вид $P\{\tau < x + v/\tau > v\} = 1 - (1 - A(x + v))(1 - A(v))^{-1}$. Соответственно определим потоки $x(t, u, v)$ и $y_i(t, u, v)$, $i = \overline{1, r}$. Пусть в момент T величина недоскока потока $x(t)$ равна α_T , а требование на i -м приборе к этому моменту обслуживалось время β_T^i , $i = \overline{1, r}$. Тогда, если $\kappa < T$, то

$$Q(t) = Q(T) + x(t, T, \alpha_T) - \sum_{i=1}^r y_i(t, T, \beta_T^i). \quad (4.15)$$

Соотношение (4.15) является определяющим при исследовании предельного поведения величины $Q(t)$.

Теорема 4.5. Если выполнено (4.12) и существуют конечные выражения

$$\sigma_0^2 = \int_0^\infty (x - a)^2 dF(x), \quad \sigma_i^2 = \int_0^\infty (x - a)^2 dG_i(x), \quad i = \overline{1, r},$$

то при $n \rightarrow \infty$ последовательность процессов

$$\frac{1}{\sqrt{n}} \left(Q(nt) - nt \left(\frac{1}{a} - \sum_{i=1}^r \frac{1}{b_i} \right) \right)$$

при любом начальном распределении сходится в равномерной топологии к процессу вида $\left(\frac{\sigma_0^2}{a^3} + \sum_{i=1}^r \frac{\sigma_i^2}{b_i^3} \right)^{1/2} w(t)$,

$t \in [0, T]$, где $w(t)$ — стандартный винеровский процесс ($Mw(t) = 0$, $Mw^2(t) = t$).

Доказательство. Установим вначале некоторые асимптотические свойства потоков восстановления. Если считать $x(0) = 1$, то имеет место представление

$$t = \sum_{i=1}^{x(t)-1} \tau_i + \tau^-(t), \quad t \geq 0,$$

где $\tau_1, \tau_2, \dots$ — независимые интервалы восстановления, а $\tau^-(t)$ — величина недоскока до момента t (как и ранее, будем полагать $\sum_{i=1}^0 = 0$).

Тогда

$$\begin{aligned} \frac{x(nt) - nta^{-1}}{\sqrt{n}} &= -\frac{1}{a\sqrt{n}} \sum_{i=1}^{x(nt)-1} (\tau_i - a) - \\ &- \frac{1}{\sqrt{n}} - \frac{1}{a\sqrt{n}} \tau^-(nt). \end{aligned} \quad (4.16)$$

Так как τ_i — независимые одинаково распределенные величины с конечной дисперсией σ_0^2 , то при $n \rightarrow \infty$ распределение величины $\frac{1}{\sqrt{n}} \sum_{i=1}^n (\tau_i - a)$ слабо сходится к нормальному с параметрами $(0, \sigma_0^2)$. Тогда из общих предельных теорем для случайных процессов следует, что последовательность процессов $\frac{1}{\sqrt{n}} \sum_{i=1}^{[nt]} (\tau_i - a)$ сходится в равномерной топологии к процессу $\sigma_0 w(t)$, $t \in [0, T]$ для любого $T > 0$. Далее в силу закона больших чисел $x(t) \times t^{-1} \xrightarrow{P} a^{-1}$ при $t \rightarrow \infty$. Отсюда вытекает, что при $n \rightarrow \infty$ для любого $t > 0$

$$\frac{1}{n} x(nt) \xrightarrow{P} ta^{-1},$$

при этом процесс $\frac{1}{n} x(nt)$ монотонно не убывает по t .

Нетрудно показать, что из поточечной сходимости монотонных функций к непрерывной следует равномерная сходимость. Следовательно, последовательность $\frac{1}{n} x(nt)$ сходится в равномерной топологии к ta^{-1} при $n \rightarrow \infty$.

Теперь можно воспользоваться теоремами о сходимости суперпозиций случайных функций. Из результатов [2] непосредственно следует, что последовательность процессов

$$\frac{1}{a\sqrt{n}} \sum_{i=1}^{x(nt)-1} (\tau_i - a)$$

довательности $\frac{1}{a\sqrt{n}} \sum_{i=1}^{[nta^{-1}]} (\tau_i - a)$ и сходится в равномерной топологии к процессу $\frac{\sigma_0}{a} \omega(ta^{-1})$, эквивалентному процессу $\sigma_0 a^{-3/2} \omega(t)$.

Покажем теперь, что

$$\sup_{t \leq T} \frac{1}{\sqrt{n}} \tau^-(nt) \xrightarrow{P} 0.$$

По построению

$$\sup_{t \leq T} \tau^-(nt) \leq \max_{k \leq x(nT)} \tau_k.$$

Для любого $c > 0$

$$\begin{aligned} P\{\max_{k \leq nc} \tau_k < \sqrt{nx}\} &= P\{\tau_k < \sqrt{nx}\}^{[nc]} = \\ &= (1 - (1 - F(\sqrt{nx}))^{nc}) \rightarrow 1, \end{aligned}$$

поскольку из конечности второго момента следует, что $n^2(1 - F(n)) \rightarrow 0$ и, следовательно, для любых $c > 0$, $x > 0$, $nc(1 - F(\sqrt{nx})) \rightarrow 0$. Теперь, поскольку $n^{-1}x(nt) \xrightarrow{P} ta^{-1}$, то для любых $\varepsilon > 0$, $\delta > 0$,

$$\begin{aligned} P\left\{\frac{1}{\sqrt{n}} \max_{k \leq x(nT)} \tau_k > \varepsilon\right\} &\leq P\left\{\frac{1}{\sqrt{n}} \max_{k \leq n(Ta^{-1} + \delta)} \tau_k > \varepsilon\right\} + \\ &+ P\{x(nT) > n(Ta^{-1} + \delta)\} \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

Отсюда, в силу (4.16), окончательно следует сходимость последовательности $\frac{1}{\sqrt{n}}(x(nt) - nta^{-1})$ в равномерной топологии к процессу $\sigma_0 a^{-3/2} \omega(t)$. Заметим также, что это утверждение справедливо и для процесса $x(nt, u, v)$ при любых фиксированных u и v . Если теперь u и v — некоторые собственные случайные величины, не зависящие от течения процесса $x(t)$ при $t > u$, то по теореме Лебега получим вновь, что последовательность $\frac{1}{\sqrt{n}}(x(nt, u, v) - nta^{-1})$ сходится в равномерной топологии к $\sigma_0 \times \times a^{-3/2} \omega(t)$ (это вытекает из того факта, что если $|f_n(x)| \leq 1$ и $f_n(x) \rightarrow f_0(x)$ при $n \rightarrow \infty$ для любого $x \geq 0$, а $F(x)$ — функция распределения, то и

$$\int_0^\infty f_n(x) dF(x) \rightarrow \int_0^\infty f_0(x) dF(x).$$

Воспользовавшись (4.15), получим при любом $T > 0$;
 $t > T$

$$Q(t) = (Q(T) + x(t, T, \alpha_T) - \sum_{i=1}^r y_i(t, T, \beta_T^i)) \times \\ \times \chi(\kappa < T) + Q(t) \chi(\kappa \geq T).$$

Отсюда

$$\begin{aligned} & \left| M \exp \left\{ i\lambda \frac{1}{V_n} \left(Q(nt) - nt \left(\frac{1}{a} - \sum_{i=1}^r \frac{1}{b_i} \right) \right) \right\} - \right. \\ & \quad - M \exp \left\{ i\lambda \frac{1}{V_n} (x(nt, T, \alpha_T) - nta^{-1}) - \right. \\ & \quad \left. \left. - i\lambda \sum_{i=1}^r \frac{1}{V_n} (y_i(nt, T, \beta_T^i) - ntb_i^{-1}) \right\} \right| \leq \\ & \leq \left| M \exp \left\{ i\lambda \frac{1}{V_n} (x(nt, T, \alpha_T) - nta^{-1}) - \right. \right. \\ & \quad \left. \left. - i\lambda \sum_{i=1}^r \frac{1}{V_n} (y_i(nt, T, \beta_T^i) - ntb_i^{-1}) \right\} \times \right. \\ & \quad \times \left(\exp \left\{ i\lambda \frac{1}{V_n} Q(T) \right\} - 1 \right) \chi(\kappa \leq T) \left| + 2M\chi(\kappa > T) \leq \right. \\ & \leq M \left| \exp \left\{ i\lambda \frac{1}{V_n} Q(T) \right\} - 1 \right| + 2P\{\kappa > T\}. \end{aligned}$$

Так как κ — собственная случайная величина, то, задавшись произвольным $\varepsilon > 0$, можно выбрать T таким, что $P\{\kappa > T\} < \varepsilon$. Далее, при каждом фиксированном T при $n \rightarrow \infty$ $\frac{1}{V_n} Q(T) \xrightarrow{P} 0$, откуда в силу ограниченности и непрерывности функции $|e^{i\lambda x} - 1|$ следует, что при $n \rightarrow \infty$

$$M \left| \exp \left\{ i\lambda \frac{1}{V_n} Q(T) \right\} - 1 \right| \rightarrow 0.$$

Далее при любых фиксированных значениях $\alpha_T, \beta_T^i, i = \overline{1, r}$, процессы $\frac{1}{V_n} (x(t, T, \alpha_T) - nta^{-1})$ и $\frac{1}{V_n} (y_i(t, T, \beta_T^i) - ntb_i^{-1}), i = \overline{1, r}$, независимы в совокупности и сходятся в равномерной топологии к процессам $\sigma_0 a^{-3/2} \omega(t)$ и $\sigma_i b_i^{-3/2} \omega_i(t), i = \overline{1, r}$, соответственно, где $\omega(t)$ и $\omega_i(t), i = \overline{1, r}$ — независимые стандартные винеровские процессы. Поэтому линейная их комбинация также сходится в равномерной топологии к линейной комбинации вида $\sigma_0 a^{-3/2} \times \times \omega(t) + \sum_{i=1}^r \sigma_i b_i^{-3/2} \omega_i(t)$, что эквивалентно процессу $(\sigma_0^2 a^{-3} + \sum_{i=1}^r \sigma_i^2 b_i^{-3})^{\frac{1}{2}} \omega(t)$. Усредняя по любому собствен-

ному распределению величин $\alpha_r, \beta_r^i, i = \overline{1, r}$, получим утверждение теоремы 4.5 в одномерном случае. Совершенно аналогично можно расписать совместные распределения процесса $\frac{1}{\sqrt{n}} \left(Q(nt) - nt \left(\frac{1}{a} - \sum_{i=1}^r \frac{1}{b_i} \right) \right)$, а также

выражение для модуля непрерывности процесса в равномерной топологии, что окончательно доказывает теорему.

Из этих результатов следует вывод о том, что для широкого класса систем величина очереди $Q(nt)$ в случае, когда загрузка больше единицы, растет как процесс $\alpha nt + \beta \sqrt{n} w(t)$, ($\alpha > 0$).

Исследуем теперь асимптотическое поведение времени пребывания в системе при дисциплине обслуживания «первый пришел — первый обслужился». Обозначим здесь эту величину $\hat{w}(t)$, чтобы не путать с обозначением винеровского процесса.

Нетрудно понять, что при условии полной загрузки системы, имеет место соотношение

$$\hat{w}(t) = \inf \left\{ u : \sum_{i=1}^r (y_i(t+u) - y_i(t)) \geq Q(t) \right\}, \quad (4.17)$$

т. е. за промежуток $[t, t + \hat{w}(t)]$ суммарное количество требований, обслуженных системой, должно совпадать с величиной $Q(t)$. Из соотношения (4.17) нетрудно вывести закон больших чисел для $\hat{w}(t)$. В самом деле, при фиксированном $t > 0$ и $n \rightarrow \infty$, как следует из предыдущих рассуждений

$$\frac{1}{n} (y_i(n(t+u)) - y_i(nt)) \xrightarrow{P} \frac{u}{b_i}, \quad i = \overline{1, r},$$

$$\frac{1}{n} Q(nt) \xrightarrow{P} \left(\frac{1}{a} - \sum_{i=1}^r \frac{1}{b_i} \right) t,$$

при этом имеет место также сходимость в равномерной топологии.

Поэтому

$$\begin{aligned} \frac{1}{n} \hat{w}(nt) &= \inf \left\{ u : \frac{1}{n} \sum_{i=1}^r y_i(nu) \geq \frac{1}{n} Q(nt) \right\} \xrightarrow{n \rightarrow \infty} \\ &\rightarrow \inf \left\{ u : u \sum_{i=1}^r \frac{1}{b_i} \geq \left(\frac{1}{a} - \sum_{i=1}^r \frac{1}{b_i} \right) t \right\} = \\ &= t \left(\frac{1}{a} - \sum_{i=1}^r \frac{1}{b_i} \right) \left(\sum_{i=1}^r \frac{1}{b_i} \right)^{-1}. \end{aligned}$$

Изучим теперь поведение величины $\frac{1}{\sqrt{n}}(w(nt) - ntq)$, где $q = \left(\frac{1}{a} - \sum_{i=1}^r \frac{1}{b_i}\right) \left(\sum_{i=1}^r \frac{1}{b_i}\right)^{-1}$.

Теорема 4.6. В условиях теоремы 4.5 последовательность процессов $\frac{1}{\sqrt{n}}(\hat{w}(nt) - ntq)$ сходится в равномерной топологии к процессу $cw(t)$, где

$$c^2 = \frac{1}{a} \left(\frac{\sigma_0^2}{a^2} + \sum_{i=1}^r \frac{\sigma_i^2}{b_i^3} \left(\sum_{j=1}^r \frac{1}{b_j} \right)^{-1} \right) \left(\sum_{i=1}^r \frac{1}{b_i} \right)^{-2}.$$

Доказательство. Воспользуемся следующими конструктивными построениями, основанными на методе одного вероятностного пространства. А. В. Скороходом было доказано, что если последовательность процессов $\xi_n(t)$ сходится в равномерной топологии к процессу $\xi_0(t)$, $t \in [QT]$, то можно построить эквивалентные процессы, заданные на одном вероятностном пространстве, которые по некоторой подпоследовательности $n_k \rightarrow \infty$ будут сходиться к предельным с вероятностью единица в равномерной топологии. Не ограничивая общности, будем считать, что n_k совпадает с n и что исходные и предельные процессы уже заданы на одном вероятностном пространстве и имеет место соответствующая сходимости с вероятностью единица. Поэтому в силу метода одного вероятностного пространства и предыдущих утверждений справедливы следующие представления:

$$\begin{aligned} x(nt) &= nt/a + \sqrt{n} (w_0(t) \sigma_0 a^{-3/2} + \rho_n^{(0)}(t)); \\ y_i(nt) &= nt/b_i + \sqrt{n} (w_i(t) \sigma_i b_i^{-3/2} + \rho_n^{(i)}(t)); \\ Q(nt) &= nt \left(\frac{1}{a} - \sum_{i=1}^r \frac{1}{b_i} \right) + \sqrt{n} (w_0(t) \sigma_0 a^{-3/2} - \\ &\quad - \sum_{i=1}^r w_i(t) \sigma_i b_i^{-3/2} + \rho_n^{(r+1)}(t)), \quad t \in [0, T], \end{aligned} \quad (4.18)$$

где $w_i(t)$, $i = \overline{0, r}$ — независимые в совокупности стандартные винеровские процессы, а $\sup_{t \in [0, T]} |\rho_n^{(i)}(t)| \rightarrow 0$ с вероятностью единица, $i = \overline{1, r+1}$.

Взяв в (4.17) вместо t величину nt , а вместо u величину $ntq + \sqrt{n}v$, получим в силу разложений (4.18)

$$\begin{aligned}
w(nt) &= ntq + \sqrt{n} \inf \{v : v > -\infty, \sum_{i=1}^r (y_i(nt + ntq + \\
&+ \sqrt{n} v) - y_i(nt)) \geq Q(nt)\} = ntq + \sqrt{n} \inf \left\{v : nt \left(\frac{1}{a} - \right. \right. \\
&- \sum_{i=1}^r \frac{1}{b_i} \left. \right) + \sqrt{n} v \sum_{i=1}^r \frac{1}{b_i} + \sqrt{n} \sum_{i=1}^r (\omega_i(t(1+q) + \\
&+ \frac{v}{\sqrt{n}}) - \omega_i(t)) \sigma_i b_i^{-3/2} \geq nt \left(\frac{1}{a} \sum_{i=1}^r \frac{1}{b_i} \right) + \\
&+ \sqrt{n} (\omega_0(t) \sigma_0 a^{-3/2} - \sum_{i=1}^r \omega_i(t) \sigma_i b_i^{-3/2}) + \sqrt{n} \rho_n(t) \left. \right\} = \\
&= ntq + \sqrt{n} \inf \left\{v : v \geq \left(\sum_{i=1}^r \frac{1}{b_i} \right)^{-1} \omega_0(t) \sigma_0 a^{-3/2} - \right. \\
&- \sum_{i=1}^r \sigma_i b_i^{-3/2} \omega_i \left(t(1+q) + \frac{v}{\sqrt{n}} \right) + \rho_n(t) \left. \right\}.
\end{aligned}$$

Поскольку траектории $\omega_i(t)$ непрерывны, а $\sup_{t \in [0, T]} |\rho_n(t)| \rightarrow 0$, с вероятностью единица, то отсюда следует, что

$$\begin{aligned}
\frac{1}{\sqrt{n}} (w(nt) - ntq) &\rightarrow \left(\sum_{i=1}^r \frac{1}{b_i} \right)^{-1} (\omega_0(t) \sigma_0 a^{-3/2} - \\
&- \sum_{i=1}^r \sigma_i b_i^{-3/2} \omega_i(t(1+q))).
\end{aligned}$$

Поскольку процесс $\alpha_1 \omega_1(\beta_1 t) + \alpha_2 \omega_2(\beta_2 t)$ эквивалентен процессу $(\alpha_1^2 \beta_1 + \alpha_2^2 \beta_2)^{1/2} w(t)$, то отсюда вытекает утверждение теоремы 4.6.

Полученные результаты можно распространить на системы с полумарковским входящим потоком и более сложной дисциплиной обслуживания.

ГЛАВА 5. ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ ДЛЯ ПОТОКОВ РЕДКИХ СОБЫТИЙ

5.1. О наступлении редкого события в схеме восстановления

В предыдущей главе использовалось поведение систем в условиях большой загрузки ($\rho > 1$ либо $\rho \rightarrow \infty$). Качественно иная картина будет в случае, когда $\rho \rightarrow 0$. Подобные модели возникают при исследовании потоков потерянных требований систем с асимптотически быстрым обслуживанием, потоков отказов высоконадежных систем,

изучении редящих потоков и т. п. То общее, что роднит эти классы задач, связано с рассмотрением потоков редких событий (вероятность наступления события в некотором смысле асимптотически мала). При этом возникает естественная проблема — исследовать при соответствующей нормировке асимптотическое поведение момента первого наступления редкого события (который может интерпретироваться как момент отказа системы), а также изучить при соответствующем растяжении временной оси поведение потока «редких» событий.

Ряд задач, связанных с асимптотическим исследованием моментов отказов восстанавливаемых систем, приводят к рассмотрению следующей схемы, впервые изученной А. Д. Соловьевым.

Пусть $v(t)$ — простой поток восстановления, $\tau_1, \tau_2, \dots$ — последовательные промежутки между поступлениями требований. $t_k = \sum_{i=1}^k \tau_i$, $k \geq 1$ ($t_0 = 0$). Предположим, что на каждом периоде восстановления $[t_k, t_{k+1})$ независимо от других промежутков может произойти некоторое событие A в некоторой случайной точке $t_k \leq \kappa_k < t_{k+1}$. Обозначим через ζ момент первого наступления события A . Очевидно, ζ представим в виде

$$\zeta = \tau_1 + \dots + \tau_{v-1} + \tau'_v, \quad (5.1)$$

где v — целочисленная случайная величина, равная первому по счету номеру периода восстановления, на котором произошло A , а $\tau'_v = \zeta - t_{v-1}$. Пусть $q = P(A)$ — вероятность осуществления A на одном периоде восстановления (так как $v(t)$ — простой поток восстановления, $q = P\{v = 1\}$). Если $q \rightarrow 0$, то A — редкое событие. В этом случае, очевидно, $v \xrightarrow{P} \infty$. Будем исследовать при этих предположениях предельное поведение ζ . Вычислим вначале явное выражение для преобразования Лапласа величины ζ . Обозначим

$$\begin{aligned} \varphi(\theta) &= M e^{-\theta \zeta}, \quad \chi_1 = \chi\{v = 1\}; \\ \alpha(\theta) &= M e^{-\theta \tau_1}. \end{aligned}$$

Лемма 5.1.

$$\varphi(\theta) = \frac{M \exp\{-\theta \zeta\} \chi_1}{1 - M \exp\{-\theta \tau_1\} (1 - \chi_1)}. \quad (5.2)$$

Действительно,

$$\begin{aligned} \varphi(\theta) &= M e^{-\theta \zeta} \chi_1 + M e^{-\theta \zeta} (1 - \chi_1) = M e^{-\theta \zeta} \chi_1 + M e^{-\theta(\tau_1 + \zeta')} (1 - \chi_1) \\ &= M e^{-\theta \zeta} \chi_1 + M e^{-\theta \tau_1} (1 - \chi_1) M e^{-\theta \zeta'}, \end{aligned}$$

где ζ' — случайная величина, равная интервалу времени до момента наступления A , начиная с момента t_1 , при условии, что $\nu > 1$. Тогда по построению ζ' не зависит от τ_1 и χ_1 и $Me^{-\theta\zeta'} = \varphi(\theta)$. Отсюда получим (5.2).

Лемма 5.2. Если параметры потока меняются таким образом, что для некоторого нормирующего множителя β

$$\begin{aligned} \frac{1}{q} (1 - Me^{-\theta\beta\tau_1}) &\rightarrow a(\theta); \\ \frac{1}{q} Me^{-\theta\beta\tau_1}\chi_1 &\rightarrow \psi_1(\theta), \quad \frac{1}{q} Me^{-\theta\beta\zeta}\chi_1 \rightarrow \psi_2(\theta), \end{aligned} \quad (5.3)$$

где $a(\theta)$ — логарифм преобразования Лапласа некоторого безграничного закона, а $\psi_1(\theta)$ и $\psi_2(\theta)$ — преобразования Лапласа некоторых случайных величин, то

$$\varphi(\beta\theta) \rightarrow \frac{\psi_2(\theta)}{a(\theta) + \psi_1(\theta)}. \quad (5.4)$$

Действительно, представим (5.2) в виде

$$\varphi(\theta) = \frac{\frac{1}{q} Me^{-\theta\zeta}\chi_1}{\frac{1}{q} (1 - Me^{-\theta\tau_1}) + \frac{1}{q} Me^{-\theta\tau_1}\chi_1}, \quad (5.5)$$

откуда непосредственно вытекает (5.4).

Основную сложность в конкретных задачах составляет проверка условий (5.3). Заметим, что в большинстве конкретных схем, связанных с анализом СМО, $\psi_1(\theta) \equiv 1$, откуда следует, что и $\psi_2(\theta) \equiv 1$ (поскольку $\psi_1(\theta) \leq \psi_2(\theta) \leq 1$). Действительно, если $\psi_1(\theta) \equiv 1$, то это означает, что $P\{\beta\tau_1 > \delta/A\} \rightarrow 0$ для любого $\delta > 0$, т. е. при осуществлении события A не происходит больших выбросов периода восстановления, сравнимых с величиной β^{-1} . Поэтому в большинстве реальных задач класса предельных законов для $\varphi(\beta\theta)$ имеет вид

$$(1 + a(\theta))^{-1}. \quad (5.6)$$

Укажем теперь достаточные условия сходимости ζ к показательному закону.

Теорема 5.1. Пусть параметры потока меняются таким образом, что $\alpha = qM\tau_1^2(M\tau_1)^{-2} > 0$ и $\alpha \rightarrow 0$. Тогда

$$P\left\{\frac{q\zeta}{M\tau_1} > x\right\} \xrightarrow{\alpha \rightarrow 0} e^{-x}. \quad (5.7)$$

Доказательство. Достаточно показать, что

$$\varphi\left(\frac{q}{M\tau_1}\theta\right) \xrightarrow{\alpha \rightarrow 0} \frac{1}{1+\theta}.$$

Воспользовавшись очевидным неравенством $0 \leq 1 - e^{-x} \leq x$ при $x \geq 0$, будем иметь для числителя (5.5)

$$\begin{aligned} 0 &\leq 1 - \frac{1}{q} M \exp\left\{-\frac{q}{M\tau_1}\theta\zeta\right\} \chi_1 = \\ &= \frac{1}{q} M \left(1 - \exp\left\{-\frac{q}{M\tau_1}\theta\zeta\right\}\right) \chi_1 \leq \frac{1}{q} M \left(\frac{q}{M\tau_1}\theta\zeta\chi_1\right) = \\ &= \frac{\theta}{M\tau_1} M\zeta\chi_1 \leq \frac{\theta}{M\tau_1} M\tau_1\chi_1. \end{aligned}$$

Далее, используя неравенство Шварца

$$M\xi_1\xi_2 \leq \sqrt{M\xi_1^2 M\xi_2^2},$$

получим

$$\frac{\theta}{M\tau_1} M\tau_1\chi_1 \leq \frac{\theta}{M\tau_1} \sqrt{M\tau_1^2 M\chi_1^2} = \theta \sqrt{\frac{M\tau_1^2}{(M\tau_1)^2} q} \rightarrow 0.$$

Аналогично показывается, что

$$\frac{1}{q} M \exp\left\{-\frac{q}{M\tau_1}\theta\tau_1\right\} \chi_1 \rightarrow 1.$$

Теперь осталось показать, что

$$\frac{1}{q} \left(1 - M \exp\left\{-\frac{q}{M\tau_1}\theta\tau_1\right\}\right) \rightarrow \theta. \quad (5.8)$$

Воспользуемся неравенством $0 \leq e^{-x} - 1 + x \leq \frac{x^2}{2}$, которое справедливо при $x \geq 0$ (так как для функции $f(x) = e^{-x} - 1 + x - \frac{x^2}{2}$, $f(0) = 0$, а $f'(x) = -e^{-x} + 1 - x \leq 0$ при $x \geq 0$, то $f(x) \leq 0$ при $x \geq 0$). Тогда

$$\begin{aligned} 0 &\leq \theta - \frac{1}{q} \left(1 - M \exp\left\{-\frac{q}{M\tau_1}\theta\tau_1\right\}\right) = \\ &= \frac{1}{q} M \left(\exp\left\{-\frac{q}{M\tau_1}\theta\tau_1\right\}\right) - 1 + q\theta \frac{\tau_1}{M\tau_1} \leq \\ &\leq \frac{1}{2q} \frac{\theta^2 q^2}{(M\tau_1)^2} M\tau_1^2 = \alpha \frac{\theta^2}{2} \rightarrow 0, \end{aligned}$$

что доказывает теорему.

Приведем одно важное следствие теоремы 5.1.

Следствие 5.1. Если параметры потока меняются так, что распределение τ_1 фиксированно и $M\tau_1 = m < \infty$, а $q > 0$ и $q \rightarrow 0$, то имеет место (5.7).

В самом деле, как следует из доказательства теоремы 5.1, достаточно показать, что $M\tau_1\chi_1 \rightarrow 0$ и выполнено (5.8). Напишем

$$M\tau_1\chi_1 = \int \tau_1(\omega) \chi_1^{(q)}(\omega) P(d\omega),$$

где $q = M\chi_1^{(q)} \rightarrow 0$. Поскольку $\chi_1^{(q)}$ — индикатор, то $\chi_1^{(q)}$ стремится к нулю по вероятности при $q \rightarrow 0$. Так как $m < \infty$, то, используя теорему Лебега, нетрудно получить, что и $M\tau_1\chi_1 \rightarrow 0$ (действительно, если по некоторой последовательности $q_n \rightarrow 0$ $M\tau_1\chi_1^{(q_n)} \rightarrow \delta > 0$, то выбирая из q_n подпоследовательность q_{n_k} , по которой $\chi_1^{(q_{n_k})} \rightarrow 0$ а вероятностью единица, получим $M\tau_1\chi_1^{(q_{n_k})} \rightarrow 0$. Полученное противоречие доказывает наше утверждение). Теперь, из соотношения $M \exp \{-\theta\tau_1\} = 1 - \theta m + o(\theta)$ непосредственно следует (5.8).

Подобные схемы естественно возникают при исследовании моментов достижения так называемого удаляющегося уровня. Так, к примеру, пусть задан некоторый случайный процесс $\xi(t)$, который принадлежит к классу регенерирующих процессов, т. е. существует последовательность моментов $0 \leq \mu_1 < \mu_2 < \dots$, которые являются точками восстановления (регенерации) для $\xi(t)$ (это означает, что распределения величин $\xi(t + \mu_k) - \xi(\mu_k)$, $t \geq 0$ не зависят от значений процесса $\xi(u)$, $u \leq \mu_k$ и совпадают с распределением величин $\xi(t + \mu_1) - \xi(\mu_1)$ для любого $k \geq 1$). Пусть также задана последовательность областей D_N , $N \geq 1$. Вероятность попадания $\xi(t)$ в D_N на одном периоде восстановления обозначим $q(D_N)$, т. е.

$$q(D_N) = 1 - P\{\xi(t) \notin D_N, \mu_1 \leq t \leq \mu_2\}.$$

Через $\zeta(D_N)$ обозначим момент первого попадания процесса $\xi(t)$ в область D_N .

Тогда если $m = M(\mu_2 - \mu_1) < \infty$ и при $N \rightarrow \infty$ последовательность D_N меняется так, что $q(D_N) > 0$ и $q(D_N) \rightarrow 0$, то, в силу следствия 5.1,

$$P\{m^{-1}q(D_N)\zeta(D_N) > x\} \xrightarrow{N \rightarrow \infty} e^{-x}.$$

В частности, пусть $\xi(t)$ — ПМП с множеством состояний $\{1, 2, \dots\}$, который задается вложенной цепью Маркова с вероятностями переходов $\|p_{ij}\|$, $i, j = \overline{1, \infty}$, и на-

бором времен переходов τ_{ij} , $i, j = \overline{1, \infty}$. Предположим, что все состояния сообщаются и, кроме того, вложенная цепь имеет стационарное распределение $\{\pi_k, k = \overline{1, \infty}\}$. Как известно, ПМП является многомерным процессом восстановления, при этом последовательные моменты попадания в любое фиксированное состояние $\{j\}$ являются точками восстановления. Пусть для простоты $j = 1$. Обозначим

$$\xi_N = \inf \{t: t > 0, \xi(t) \geq N\}.$$

Очевидно в нашем случае $D_N = \{N, N+1, \dots\}$ и при $N \rightarrow \infty$ $q(D_N) \rightarrow 0$.

При этом, как известно из теорем ПМП,

$$m = \frac{1}{\pi_1} \sum_{k, j=1}^{\infty} \pi_k p_{kj} M\tau_{kj}.$$

Для вычисления величины $q(D_N)$ можно выписать систему линейных алгебраических уравнений. Отметим, что при достаточно широких предположениях, в случае, когда $p_{ij} \rightarrow 0$ при $i, j \in D_N$, $N \rightarrow \infty$, следует ожидать, что величина $q(D_N)$ пропорциональна $\sum_{j=N}^{\infty} \pi_j$.

Укажем еще один достаточный критерий проверки условий (5.3), который часто может быть использован в реальных схемах. Положим

$$q(x) = P\{v = 1/\tau_1 = x\}.$$

Лемма 5.3. Если параметры потока меняются так, что

$$\frac{1}{q} (1 - Me^{-\theta \beta \tau_1}) \rightarrow a(\theta), \quad (5.9)$$

а функция $q(x)$ при фиксированных параметрах монотонно не возрастает, то $\psi_1(\theta) \equiv \psi_2(\theta) \equiv 1$.

Практически это условие означает, что чем больше интервал τ_1 , тем меньше вероятность осуществления на нем события A .

Для доказательства укажем одно полезное неравенство Чебышева.

Если $f(x)$ и $g(x)$ — две монотонные неотрицательные функции, $F(x)$ — функция распределения, то

$$\int_{-\infty}^{\infty} f(x) g(x) dF(x) \leq \int_{-\infty}^{\infty} f(x) dF(x) \int_{-\infty}^{\infty} g(y) dF(y), \quad (5.10)$$

если $f(x)$ не убывает, а $g(x)$ не возрастает, и

$$\int_{-\infty}^{\infty} f(x) g(x) dF(x) \geq \int_{-\infty}^{\infty} f(x) dF(x) \int_{-\infty}^{\infty} g(y) dF(y), \quad (5.11)$$

если одновременно $f(x)$ и $g(x)$ не убывают либо не возрастают (в предположении существования соответствующих интегралов).

Докажем первое неравенство (второе доказывается аналогично). По условию в силу монотонности для любых x_1, x_2

$$(f(x_1) - f(x_2))(g(x_1) - g(x_2)) \leq 0.$$

Тогда

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (f(x_1) - f(x_2))(g(x_1) - g(x_2)) dF(x_1) dF(x_2) \leq 0.$$

Поскольку $\int_{-\infty}^{\infty} dF(x) = 1$, а

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1) g(x_2) dF(x_1) dF(x_2) = \int_{-\infty}^{\infty} f(x_1) dF(x_1) \int_{-\infty}^{\infty} g(x_2) dF(x_2),$$

то после преобразований получим

$$2 \int_{-\infty}^{\infty} f(x_1) g(x_1) dF(x_1) - 2 \int_{-\infty}^{\infty} f(x_1) dF(x_1) \int_{-\infty}^{\infty} g(x_2) dF(x_2) \leq 0,$$

что эквивалентно (5.10).

Теперь для доказательства леммы покажем, что $\psi_1(\theta) \equiv 1$. В самом деле

$$\begin{aligned} 0 &\leq 1 - \frac{1}{q} M e^{-\theta \beta \tau_1} \chi_1 = \frac{1}{q} M (1 - e^{-\theta \beta \tau_1}) \chi_1 = \\ &= \frac{1}{q} M (1 - e^{-\theta \beta \tau_1}) q(\tau_1) \leq \frac{1}{q} M (1 - e^{-\theta \beta \tau_1}) M q(\tau_1) = \\ &= M (1 - e^{-\theta \beta \tau_1}) \rightarrow 0, \end{aligned}$$

в силу (5.9). Здесь применено неравенство (5.10), так как функция $1 - e^{-\theta \beta x}$ не возрастает, а $q(x)$ не убывает. Лемма 5.3 доказана.

Применим теперь указанный подход к задаче анализа момента первой потери требования ξ в системе GI | G | r | m — г в условиях, когда требования обслуживаются асимптотически быстро.

Пусть $t_1 < t_2 < \dots$ — последовательные моменты поступления требований в систему, $\xi_i = t_{i+1} - t_i$, $i \geq 1$, а η_i ,

$i \geq 1$ — последовательные времена обслуживания (по определению системы $\xi_i, i \geq 1$ и $\eta_i, i \geq 1$ — последовательности независимых одинаково распределенных случайных величин).

Через $t_1^* < t_2^* < \dots$ обозначим моменты поступления требований в пустую систему. Тогда последовательность $t_1^*, t_2^*, \dots$ образует поток восстановления. Если величины ξ_1 и η_1 меняются таким образом, что вероятность q потери требования (поступления требования в систему, у которой заняты все r приборов и $m-r$ мест для ожидания) на одном периоде восстановления (периоде занятости) стремится к нулю, то рассматривая потерю требования как осуществление редкого события, мы придем к рассмотренной выше модели. Значительные сложности вызывает проверка условий (5.3). А. Д. Соловьевым было показано, что при простых естественных предположениях относительно величин ξ_1 и η_1 предельное распределение $\beta\xi$ будет иметь вид $(1+a(\theta))^{-1}$. Приведем здесь без доказательства теорему, вытекающую из этих результатов.

Теорема 5.2. Если величины ξ_1 и η_1 меняются таким образом, что $M\xi_1 \rightarrow a > 0$, $M\xi_1^p < \infty$ для некоторого $p > 1$,

$$P\{\eta_1 \geq \xi_1 + \xi_2 + \dots + \xi_m\} > 0;$$

$$\sum_{k=1}^{\infty} P\{\eta_1 \geq \xi_1 + \dots + \xi_k\} \rightarrow 0,$$

и для некоторых $\varepsilon_0 > 0$, $p_0 > 0$ $P\{\eta_1 < a - \varepsilon_0\} \geq p_0$, то

$$P\left\{\frac{q\xi}{a} > x\right\} \rightarrow e^{-x}, \quad x \geq 0.$$

Укажем в заключение вероятностную интерпретацию полученных результатов о моменте первого наступления редкого события. В исходной схеме положим

$$\tau_k^- = \{\tau_k / \bar{A}_k\}, \quad \tau_k^+ = \{\tau_k / A_k\}, \quad k \geq 1,$$

где введенные обозначения показывают, что распределение τ_k^- совпадает с условным распределением τ_k при условии, что на k -м периоде восстановления событие A не произошло (соответственно τ_k^+ , если A произошло). Тогда в силу (5.1)

$$\xi = \sum_{k=1}^{v-1} \tau_k^- + \tau_v^+, \quad (5.12)$$

где τ_k^- , $k \geq 1$ — последовательность независимых одинаково распределенных величин, и τ_v^+ не зависит от τ_k^- , $k < v$,

а v по построению имеет геометрическое распределение:

$$P\{v = k\} = (1 - q)^{k-1}q, \quad k \geq 1$$

и не зависит от совокупности $\{\tau_k^+, \tau_k^-, k \geq 1\}$.

Обозначим $\varphi^-(\theta) = Me^{-\theta\tau_1}$, $\varphi^+(\theta) = Me^{-\theta\tau_1^+}$. Тогда из (5.12)

$$Me^{-\theta\tau} = \varphi^+(\theta) \sum_{k=1}^{\infty} \varphi^-(\theta)^{k-1} (1 - q)^{k-1} q = \frac{q \varphi^+(\theta)}{1 - (1 - q) \varphi^-(\theta)}. \quad (5.13)$$

Учитывая, что по построению

$$\varphi^+(\theta) = M[e^{-\theta\tau_1/v} = 1] = \frac{1}{q} Me^{-\theta\tau_1} \chi_1,$$

а

$$\varphi^-(\theta) = M[e^{-\theta\tau_1/v} > 1] = \frac{1}{1 - q} Me^{-\theta\tau_1} (1 - \chi_1),$$

из соотношения (5.13) легко следует (5.2).

Тем самым представление (5.13) дает другой способ вывода формулы (5.2). При этом оно может быть использовано также для доказательства предельных теорем. Действительно, если $q \rightarrow 0$, то легко убедиться в том, что

$$P\{qv > x\} \rightarrow e^{-x}, \quad x \geq 0. \quad (5.14)$$

В самом деле,

$$\begin{aligned} Me^{-q\theta v} &= \sum_{k=1}^{\infty} e^{-q\theta k} (1 - q)^{k-1} q = \\ &= \frac{qe^{-q\theta}}{1 - (1 - q)e^{-q\theta}} = \frac{q + 0(q)}{1 - (1 - q)(1 - q\theta + 0(q))} \rightarrow \frac{1}{1 + \theta}, \end{aligned}$$

что соответствует (5.14).

Предположим теперь, что параметры потока меняются таким образом, что для некоторого нормирующего множителя β

$$\tau^-(t) = \beta \sum_{k=1}^{[tq^{-1}]} \tau_k^- \xrightarrow{c.l.} \xi(t), \quad t \geq 0, \quad (5.15)$$

а

$$\beta \tau_1^+ \xrightarrow{c.l.} \kappa, \quad (5.16)$$

где $\xi(t)$ — некоторый однородный случайный процесс с независимыми приращениями вида $M \exp\{-\theta \xi(t)\} = \exp\{-\bar{a}(\theta)t\}$, при этом (5.15) означает слабую сходимость конечномерных распределений. Заметим, что так как τ_k^- независимы и одинаково распределены, (5.15) эквивалентно условию

$q^{-1}(1 - \varphi^-(\beta\theta)) \rightarrow a^-(\theta)$. Отсюда нетрудно получить, что

$$Me^{-\theta\tau^-(t)} \rightarrow e^{-a^-(\theta)t} \quad (5.17)$$

равномерно по t на любом конечном промежутке на $[0, \infty)$. Теперь, пользуясь теоремой Хелли и соотношениями (5.14), (5.15), получим

$$\begin{aligned} Me^{-\theta\tau} &= Me^{-\theta\tau_1^+} \int_0^\infty Me^{-\theta\tau^-(t)} d_t P\{q\nu < t + q\} \rightarrow \\ &\rightarrow Me^{-\theta x} \int_0^\infty e^{-a^-(\theta)t} e^{-t} dt, \end{aligned} \quad (5.18)$$

что эквивалентно следующему соотношению:

$$\xi \stackrel{\text{сл}}{\Rightarrow} \xi(\eta) + \kappa, \quad (5.19)$$

где η и κ — независимы и не зависят от $\xi(t)$, $t \geq 0$ и η показательно распределена с параметром 1.

Заметим, что, вычисляя интеграл в правой части (5.18), получим выражение $Me^{-\theta x}(1 + a^-(\theta))^{-1}$, что в других обозначениях эквивалентно (5.4).

Таким образом, указанных конструктивный подход приводит к схеме изучения геометрически распределенного (с малым параметром) числа независимых одинаково распределенных слагаемых. Это обстоятельство наталкивает на мысль о тесной связи асимптотических задач ТМО с теорией суммирования независимых случайных величин.

Заметим также, что указанный подход (представление различных функционалов типа моментов осуществления редких событий в виде суммы случайного числа слагаемых) получил существенное развитие в работах ряда авторов применительно к значительно более общим схемам случайной остановки различных классов процессов и их приложений к задачам ТМО, теории надежности и редящим потокам [3, 4, 6, 9].

5.2. Асимптотическое поведение времени пребывания полумарковского процесса в связанном подмножестве состояний

Полумарковские процессы (ПМП) (см. п. 1.4) являются естественным обобщением цепей Маркова с непрерывным временем и в рамках ПМП удается описывать достаточно широкие классы систем, являющихся моделями реальных производственных объектов. При этом зачастую возникают

задачи, связанные с нахождением времени пребывания системы в определенном подмножестве пространства состояний (момент выхода из подмножества обычно интерпретируется как отказ системы либо момент изменения режима функционирования). В тех случаях, когда мы имеем дело с высоконадежными системами, а выход из подмножества сигнализирует о появлении какой-либо неисправности, естественно ожидать, что вероятности выхода достаточно малы. В связи с этим возникает актуальная проблема — исследовать асимптотическое поведение времени пребывания ПМП в подмножестве состояний в том случае, когда вероятности выходов стремятся к нулю. В какой-то мере эта модель обобщает схему, изученную в п. 5.1 о моменте 1-го наступления редкого события на процессе восстановления.

Здесь решим данную проблему для случая конечного подмножества состояний. Изучим вначале неасимптотическую ситуацию.

Пусть $\xi(t)$, $t \geq 0$ ПМП с конечным множеством состояний $\{0, 1, \dots, r\}$, который задается вложенной цепью Маркова x_n , $n \geq 0$, с матрицей вероятностей переходов $\|p_{ij}\|$, $i, j = \overline{0, r}$, и набором независимых в совокупности случайных величин $\tau(i, j)$ — времен переходов из i в j , $i, j = \overline{0, r}$. Через $\Omega(k)$ обозначим время пребывания $\xi(t)$ в подмножестве $I = \{1, \dots, r\}$, начиная из состояния $\{k\}$, т. е.

$$\Omega(k) = \inf \{t : t > 0, \xi(t) = 0\}, \text{ где } \xi(0) = k, k \neq 0.$$

Для величин $\Omega(k)$ легко получить систему стохастических соотношений:

$$\Omega(k) = \begin{cases} \tau(k, j) + \Omega(j), & j = \overline{1, r}, \text{ с вер. } p(k, j); \\ \tau(k, 0), & \text{с вер. } p(k, 0) \end{cases} \quad (5.20)$$

(здесь соотношение

$$\xi = \begin{cases} \xi_k, & \text{с вер. } \alpha_k; \\ k = \overline{1, r}, \end{cases}$$

формально означает, что $P\{\xi < x\} = \sum_{k=1}^r \alpha_k P\{\xi_k < x\}$. Обозначим $\varphi(u, k, j) = Me^{iu\tau(k, j)}$, $k, j = \overline{0, r}$, $\psi(u, k) = Me^{iu\Omega(k)}$. Переходя в (5.20) к характеристическим функциям, получим для величин $\psi(u, k)$ систему линейных алгебраических уравнений:

$$\begin{cases} \psi(u, k) = \sum_{j=1}^r p(k, j) \varphi(u, k, j) \psi(u, j) + p(k, 0) \varphi(u, k, 0); \\ k = \overline{1, r}. \end{cases} \quad (5.21)$$

Обозначим через $\psi(u)$ и $\varphi(u)$ — вектор-столбцы с компонентами $\psi(u, k)$ и $p(k, 0)\varphi(u, k, 0)$ соответственно, а

$$\Phi(u) = \|p(k, j)\varphi(u, k, j)\|, \quad k, j = \overline{1, r}.$$

Тогда в матричном виде (5.21) переписывается

$$\psi(u) = \Phi(u)\psi(u) + \varphi(u). \quad (5.22)$$

Предположим, что множество I обладает тем свойством, что для любого $i \in I$ существует цепочка переходов с положительными вероятностями, ведущая в $\{0\}$, т. е. I не замкнуто и не содержит замкнутых подмножеств. В этом случае все характеристические корни матрицы $P = \|p(k, j)\|$, $k, j = \overline{1, r}$, по модулю меньше единицы и по теореме Фробениуса характеристические корни матрицы $\Phi(u)$ при любом u также по модулю меньше единицы (так как $|\varphi(u, k, j)| \leq 1$). Следовательно, существует матрица $(E - \Phi(u))^{-1}$, где E — единичная матрица и

$$\psi(u) = (E - \Phi(u))^{-1}\varphi(u). \quad (5.23)$$

Из соотношения (5.21) легко получить для вектора с компонентами $M\Omega(k)$, $k = \overline{1, r}$:

$$\bar{M} = (E - P)^{-1}\bar{m},$$

где $\bar{m}$ — вектор-столбец с компонентами

$$m_k = \sum_{j=0}^r p(k, j)M\tau(k, j), \quad k = \overline{1, r}.$$

Предположим теперь, что параметры процесса (цепь $x_n = x_\varepsilon(n)$, величины $p(k, j) = p_\varepsilon(k, j)$ и $\tau(k, j) = \tau_\varepsilon(k, j)$) зависят от некоторого параметра ε , $\varepsilon \rightarrow 0$, так, что $p_\varepsilon(k, 0) \rightarrow 0$, $k = \overline{1, r}$. В этом случае очевидно количество скачков процесса до момента выхода из I стремится по вероятности к бесконечности и имеет смысл исследовать предельное распределение величин $\Omega(j) = \Omega_\varepsilon(j)$.

Предположим, что выполнены условия:

1. $p_\varepsilon(k, j) \rightarrow p_0(k, j)$, $k, j = \overline{1, r}$, и матрице $P_0 = \|p_0(k, j)\|$, $k, j = \overline{1, r}$, соответствует один существенный класс.

2. $\varepsilon^{-1}p_\varepsilon(k, 0) \rightarrow b_k < \infty$, $k = \overline{1, r}$, и $\sum_{k=1}^r b_k \neq 0$.

3. Существует нормирующий множитель β_ε , такой, что

$$a) \varphi_\varepsilon(\beta_\varepsilon u, k, j) = 1 + \varepsilon a_{kj}(u) + o(\varepsilon), \quad k, j = \overline{1, r};$$

$$б) \varphi_\varepsilon(\beta_\varepsilon u, k, 0) \rightarrow \rho_k(u), \quad k = \overline{1, r}.$$

При этих предположениях очевидно

$$\Phi_\varepsilon(\beta_\varepsilon u) \rightarrow P_0,$$

где P_0 — замкнута и, следовательно, $E - \Phi_\varepsilon(\beta_\varepsilon u)$ в пределе необратима. В этом основная сложность, возникающая при асимптотическом исследовании (5.23). В работах [10, 11] разработан аналитический аппарат асимптотического анализа операторных уравнений типа (5.23), основанный на обращении возмущенных на спектре операторов. Рассмотрим конечномерный случай.

Теорема 5.3. В условиях 1—3 распределение величины $\beta_\varepsilon \Omega_\varepsilon(j)$ независимо от начального состояния $\bar{j} = \overline{1, r}$, слабо сходится к распределению с характеристической функцией

$$\sum_{k=1}^r \pi_k b_k \rho_k(u) \left(\sum_{l=1}^r \pi_l b_l - \sum_{i,j=1}^r \pi_i p(i, j) a_{ij}(u) \right)^{-1}. \quad (5.24)$$

Доказательство. Обозначим $\psi_\varepsilon(u, k) = M e^{iu \beta_\varepsilon \Omega(k)}$, $k = \overline{1, r}$, и будем вычислять предел $\psi_\varepsilon(u, k)$ при $\varepsilon \rightarrow 0$. Пусть $\Phi_\varepsilon(u) = \| p_\varepsilon(k, j) \varphi_\varepsilon(k, j, u) \|$, $k, j = \overline{1, r}$, а $\bar{\psi}_\varepsilon(u)$ и $\bar{\varphi}_\varepsilon(u)$ — векторы-столбцы с компонентами $\psi_\varepsilon(u, k)$ и $p_\varepsilon(k, 0) \varphi_\varepsilon(u, k, 0)$. Тогда при наших предположениях

$$\Phi_\varepsilon(\beta_\varepsilon u) = P_\varepsilon + \varepsilon A(u) + \| o(\varepsilon) \|,$$

а

$$\bar{\varphi}_\varepsilon(\beta_\varepsilon u) = \varepsilon \bar{\varphi}_0(u) + o(\varepsilon),$$

где $P_\varepsilon = \| p_\varepsilon(k, j) \|$, $k, j = \overline{1, r}$, $A(u) = \| p_0(k, j) a_{kj}(u) \|$, $k, j = \overline{1, r}$, $\bar{\varphi}_0(u)$ — вектор-столбец с компонентами $b_k \rho_k(u)$, $k = \overline{1, r}$, а элементы матрицы $\| o(\varepsilon) \|$ и вектора $\bar{o}(\varepsilon)$ имеют порядки $o(\varepsilon)$. Тогда (5.23) можно записать в виде

$$\begin{aligned} \bar{\psi}_\varepsilon(u) &= (E - P_\varepsilon - \varepsilon A(u) + \| o(\varepsilon) \|)^{-1} (\varepsilon \varphi_0(u) + \bar{o}(\varepsilon)) = \\ &= (E - G_\varepsilon(\varepsilon A(u) - \| o(\varepsilon) \|))^{-1} G_\varepsilon(\varepsilon \bar{\varphi}_0(u) + \bar{o}(\varepsilon)), \end{aligned} \quad (5.25)$$

где $G_\varepsilon = (E - P_\varepsilon)^{-1}$. Исследуем поведение матрицы G_ε при $\varepsilon \rightarrow 0$. Так как матрица P_0 неприводима, т. е. единица — характеристический корень кратности единица, а все остальные корни по модулю не превосходят единицу, то, разлагая элементы матрицы $(\mu E - P_0)^{-1}$ ($\mu > 1$) на простые дроби, получим

$$(\mu E - P_0)^{-1} = \frac{\pi}{\mu - 1} + Q(\mu),$$

где $Q(\mu) = \sum_{k=2}^r \frac{Q_k}{\mu - \mu_k}$ в случае разных корней $\mu_2, \mu_3, \dots$,

μ_r ($\mu_i \neq 1$, $i = 2, r$) и $Q(\mu) = \sum_{k=x}^s \sum_{i=1}^{\alpha_k} Q_{ki} / (\mu - \mu_k)^i$ в случае, если корень μ_k имеет кратность α_k , $k = 2, s$, а Q_k и Q_{ki} — некоторые матричные коэффициенты. При этом матрица

π имеет одинаковые строки вида $\pi_1, \pi_2, \dots, \pi_r$, где $\{\pi_k\}$, $k = \overline{1, r}$ — стационарное распределение для матрицы P_0 . Отметим, что при $\mu \rightarrow 1$ элементы матрицы $Q(\mu)$ имеют конечные пределы.

Введем стохастическую матрицу $\tilde{P}_\varepsilon = \|\tilde{p}_\varepsilon(k, j)\|$, $\overline{k, j} = \overline{1, r}$, где $\tilde{p}_\varepsilon(k, j) = p_\varepsilon(k, j)(1 - p_\varepsilon(k, 0))^{-1}$. Очевидно, что $\tilde{P}_\varepsilon \rightarrow P_0$, и отсюда несложно получить, что для любого $g_\varepsilon \rightarrow 0$

$$(E - (1 - g_\varepsilon) \tilde{P}_\varepsilon)^{-1} = (1 - g_\varepsilon)^{-1} ((1 - g_\varepsilon)^{-1} E - \tilde{P}_\varepsilon)^{-1} = \\ = \frac{1}{g_\varepsilon} \pi_\varepsilon + \|O(1)\|,$$

где π_ε — матрица с одинаковыми строчками $\{\pi_\varepsilon(1), \pi_\varepsilon(2), \dots, \pi_\varepsilon(r)\}$ — стационарное распределение для матрицы P_ε , а элементы матрицы $\|O(1)\|$ имеют конечные пределы при $\varepsilon \rightarrow 0$. Обозначим

$$g_\varepsilon = \sum_{k=1}^r \pi_\varepsilon(k) p_\varepsilon(k, 0).$$

Тогда

$$(E - P_\varepsilon)^{-1} = (E - (1 - g_\varepsilon) \tilde{P}_\varepsilon + (1 - g_\varepsilon) \tilde{P}_\varepsilon - P_\varepsilon)^{-1} = \\ = \left(E + \frac{1}{g_\varepsilon} \pi_\varepsilon ((1 - g_\varepsilon) \tilde{P}_\varepsilon - P_\varepsilon) + \|O(\varepsilon)\| \right)^{-1} \left(\frac{1}{g_\varepsilon} \pi_\varepsilon + \|O(1)\| \right).$$

Пусть B — произвольная матрица, имеющая одинаковые строчки. Через q обозначим сумму строки B . Вычислим матрицу $(E - kB)^{-1}$. Ответ будем искать в виде $E + xB$. Тогда

$$(E + xB)(E - kB) = (E - kB)(E + xB) = E, \\ E + xB - kB - xkqB = E,$$

что выполнено, если $x - k - xkq = 0$, или $x = \frac{k}{1 - kq}$ при $kq \neq 1$. Отсюда

$$(E - kB)^{-1} = E + \frac{k}{1 - kq} B. \quad (5.26)$$

Так как

$$\sum_{j=1}^r (\sum_{k=1}^r \pi_\varepsilon(k) \tilde{p}_\varepsilon(k, j) (1 - g_\varepsilon) - \sum_{k=1}^r \pi_\varepsilon(k) p_\varepsilon(k, j)) = \\ = 1 - g_\varepsilon - \sum_{k=1}^r \pi_\varepsilon(k) (1 - p_\varepsilon(k, 0)) = 0,$$

то

$$\left(E + \frac{1}{g_\varepsilon} \pi_\varepsilon ((1 - g_\varepsilon) \tilde{P}_\varepsilon - P_\varepsilon) \right)^{-1} = E - \frac{1}{g_\varepsilon} \pi_\varepsilon ((1 - g_\varepsilon) \tilde{P}_\varepsilon - P_\varepsilon).$$

Учитывая, что элементы матрицы $\frac{1}{g_\varepsilon} \pi_\varepsilon ((1 - g_\varepsilon) \tilde{P}_\varepsilon - P_\varepsilon)$

имеют конечные пределы, а сумма строки равна нулю,[†] получим

$$(E - P_\varepsilon)^{-1} = (E + \|O(\varepsilon)\|)^{-1} \left(E - \frac{1}{g_\varepsilon} \pi_\varepsilon ((1 - g_\varepsilon) \bar{P}_\varepsilon - P_\varepsilon) \right) \left(\frac{1}{g_\varepsilon} \pi_\varepsilon + \|O(1)\| \right) = \frac{1}{g_\varepsilon} \pi_\varepsilon + \|O(1)\|.$$

Тем самым получим лемму, имеющую самостоятельный интерес.

Лемма 5.4. Если $\max_{k=\overline{1, r}} p_\varepsilon(k, 0) \rightarrow 0$, а матрица P_0 неприводима, то

$$(E - P_\varepsilon)^{-1} = \frac{1}{g_\varepsilon} \pi_\varepsilon + \|O(1)\|. \quad (5.27)$$

Теперь из (5.25) легко вывести, что

$$\begin{aligned} \bar{\Psi}_\varepsilon(u) &= \left(E - \frac{\varepsilon}{g_\varepsilon} \pi_\varepsilon A(u) - \|O(1)\| \right)^{-1} \left(\frac{\varepsilon}{g_\varepsilon} \pi_\varepsilon \bar{\Phi}_0(u) + o(1) \right) \rightarrow \\ &\rightarrow \left(E - \frac{1}{b} \pi A(u) \right)^{-1} \frac{1}{b} \pi \bar{\Phi}_0(u) = \bar{\Psi}_0(u), \end{aligned}$$

где $b = \sum_{k=1}^r \pi_k b_k$. Через $a(u) = \sum_{i,j=1}^r \pi_i p(i, j) a_{ij}(u)$ обозначим сумму строчки матрицы $\pi A(u)$. Тогда, пользуясь (5.7), получим

$$\begin{aligned} \bar{\Psi}_0(u) &= \left(E - \frac{1}{b-a(u)} \pi A(u) \right) \frac{1}{b} \pi \bar{\Phi}_0(u) = \\ &= \frac{1}{b} \left(1 + \frac{a(u)}{b-a(u)} \right) \pi \bar{\Phi}_0(u) = \frac{\rho(u)}{b-a(u)} \bar{1}, \end{aligned}$$

где $\rho(u) = \sum_{k=1}^r \pi_k b_k \rho_k(u)$, $\bar{1}$ — единичный вектор, что соответствует (5.24). Теорема доказана.

Рассмотрим некоторые важные следствия из теоремы 5.3.

Следствие 5.2. Если времена переходов процесса не зависят от параметра ε (т. е. $\tau_\varepsilon(k, j) = \tau_0(k, j)$, $k, j = \overline{0, r}$) и $M\tau_0(k, j) = m_{kj} < \infty$, а также выполнены условия 1, 2, то для любого $j = \overline{1, r}$

$$P\{\varepsilon \Omega_\varepsilon(j) < x\} \rightarrow 1 - \exp\left\{-\frac{b}{m} x\right\}, \quad x \geq 0, \quad (5.28)$$

где $m = \sum_{k=1}^r \pi_k p_0(k, j) m_{kj}$, т. е. распределение будет показательным с параметром $\frac{b}{m}$.

Действительно, в этом случае

$$M \exp\{i u \varepsilon \tau_0(k, j)\} = 1 + \varepsilon i u m_{kj} + o(\varepsilon), \quad k, j = \overline{0, r}$$

и, следовательно, $a_{kj}(u) = ium_{kj}$, $a(u) = ium$, $\rho_k(u) \equiv 1$ и $\Psi_0(u) = b(b - ium)^{-1} \bar{1}$, что соответствует (5.28).

Пусть $v_e(j) = \inf \{n : n \geq 0, x_e(n+1) = 0\}$, если $x_e(0) = j$, т. е. $v_e(j)$ — количество шагов, проделанных процессом в подмножестве $\{1, \dots, r\}$ до первого попадания в $\{0\}$, начиная из состояния $\{j\}$.

Следствие 5.3. Если выполнены условия 1, 2, то

$$P\{\varepsilon v_e(j) < x\} \rightarrow 1 - e^{-bx}, \quad x \geq 0. \quad (5.29)$$

В самом деле, нам достаточно положить $\tau_e(k, j) \equiv 1$, $k, j = \overline{0, r}$. Тогда $\Omega_e(i) = v_e(i) + 1$ и (5.29) следует из (5.28).

Можно показать, что в этом случае также для любого $m \in M(\varepsilon v_e(j))^m \rightarrow M\eta^m$, где η — показатель с параметром b .

В прикладных задачах представляет интерес вопрос — из какого состояния произошел переход в $\{0\}$, что может соответствовать определенному типу неисправности. Номер этого состояния определяется величиной $x_e(v_e(j))$, если $x_e(0) = j$.

Следствие 5.4. Если выполнены условия 1, 2, то

$$P\{x_e(v_e(j)) = k\} \rightarrow \pi_k b_k b^{-1}, \quad k = \overline{1, r}. \quad (5.30)$$

Действительно, если положить $\tau_e(l, n) = 0$, $l, n = \overline{1, r}$, $\tau_e(k, 0) = 0$, $\tau_e(m, 0) = +\infty$, $m = \overline{1, r}$, $m \neq k$, и вместо характеристических функций рассмотреть преобразование Лапласа, то при любом $s > 0$

$$M \exp\{-s\Omega_e(j)\} = P\{x_e(v_e(j)) = k\},$$

и так как в нашем случае $\varphi_e(u, l, n) \equiv 1$, $k, j = \overline{1, r}$, $\rho_m(-is) \equiv 0$, $s > 0$, $m \neq k$, а $\rho_k(-is) \equiv 1$, то из (5.24) вытекает (5.30).

Укажем конструктивную интерпретацию класса законов вида (5.24).

Пусть $\xi(t)$, $t \geq 0$ процесс с независимыми приращениями такой, что $M \exp\{iu\xi(t)\} = \exp\{ta(u)\}$, ξ — случайная величина с характеристической функцией $\rho(u)$, а η — имеет показательное распределение с параметром $b(\xi(t))$, $t \geq 0$, ξ и η независимы в совокупности). Тогда класс предельных законов можно представить в виде

$$\xi(\eta) + \zeta \quad (5.31)$$

(нетрудно вычислить, что характеристическая функция (5.31) равна (5.24).

Интуитивно это можно объяснить следующим образом. Поскольку вероятности выхода из состояний стремятся к нулю, число скачков в подмножестве стремится к бесконечности и, грубо говоря, прежде чем выйти, процесс успеет войти в стационарный режим. Тогда g_ε — есть стационарная вероятность выхода за 1 шаг, при этом в стационарном режиме характеристическая функция времени перехода будет

$$\sum_{kj} \pi_\varepsilon(k) p_\varepsilon(k, j) \varphi_\varepsilon(u, k, j)$$

и при нормировке времени β_ε имеет разложение $1 + \varepsilon a(u) + o(\varepsilon)$, а $\rho(u)$ есть характеристическая функция предельного распределения времени перехода за 1 шаг из подмножества в $\{0\}$. Тем самым наша схема асимптотически близка к схеме суммирования геометрического числа (с параметром g_ε) слагаемых, и в силу результатов п. 5.1 естественно ожидать класс законов вида (5.19).

5.3. Асимптотически связанные подмножества

В теореме 5.3 исследована ситуация, в которой множество $\{1, \dots, r\}$ в пределе замкнуто и связно. Вместе с тем в реальных задачах, при исследовании СМО в условиях быстрого обслуживания (см. п. 5.4) возникают модели, в которых подмножеству $\{1, \dots, r\}$ в пределе соответствует несколько классов существенных состояний и, возможно, несущественные состояния. В этом случае ситуация существенно будет зависеть от скорости стремления к нулю вероятностей переходов между этими классами. Введем важное понятие множества состояний, связанного в асимптотическом смысле — s -множества.

Пусть $x_\varepsilon(k)$, $k \geq 0$ — цепь Маркова с множеством состояний $\{0, 1, \dots, r\}$ и матрицей переходных вероятностей, $\|p_\varepsilon(i, j)\|$, $i, j = \overline{0, r}$, $\langle \alpha \rangle$ — некоторое подмножество состояний из $\{1, 2, \dots, r\}$. Положим

$$v_\varepsilon(i, \langle \alpha \rangle) = \min \{k : k > 0, x_\varepsilon(k) \notin \langle \alpha \rangle\},$$

где

$$x_\varepsilon(0) = i \in \langle \alpha \rangle,$$

$q_\varepsilon(i, j, \langle \alpha \rangle) = P\{x_\varepsilon(k) = j \text{ хотя бы при одном } k, k < v_\varepsilon(i, \langle \alpha \rangle)/x_\varepsilon(0) = i\}$, $i, j \in \langle \alpha \rangle$, т. е. $q_\varepsilon(i, j, \langle \alpha \rangle)$ — вероятность того, что, начиная из i , цепь попадет в j прежде чем выйдет из $\langle \alpha \rangle$.

Определение 5.1. Подмножество состояний $\langle \alpha \rangle$ образует s -множество, если для любых $i, j \in \langle \alpha \rangle$

$$q_\varepsilon(i, j, \langle \alpha \rangle) \rightarrow 1$$

при $\varepsilon \rightarrow 0$.

Это фактически означает, что, начиная из любого состояния, цепь посетит каждое состояние из $\langle \alpha \rangle$ асимптотически бесконечное число раз, прежде чем выйдет из $\langle \alpha \rangle$. Простейшим примером s -множества может служить такое множество, которое в пределе образует один существенный класс. Оказывается, что асимптотические свойства s -множеств аналогичны свойствам множеств, образующих в пределе один класс.

Обозначим

$$\beta_\varepsilon(i, \langle \bar{\alpha} \rangle) = \sum_{l \in \langle \bar{\alpha} \rangle} p_\varepsilon(i, l),$$

$$g_\varepsilon(\langle \alpha \rangle) = \sum_{i \in \langle \alpha \rangle} \pi_\varepsilon(i, \langle \alpha \rangle) \beta_\varepsilon(i, \langle \bar{\alpha} \rangle), \quad (5.32)$$

где $\pi_\varepsilon(i, \langle \alpha \rangle)$, $i \in \langle \alpha \rangle$ — стационарное распределение для цепи с матрицей

$$\bar{P}_\varepsilon(\langle \alpha \rangle) = \|\bar{p}_\varepsilon(i, j, \langle \alpha \rangle)\|, \quad i, j \in \langle \alpha \rangle,$$

а

$$\bar{p}_\varepsilon(i, j, \langle \alpha \rangle) = p_\varepsilon(i, j) (1 - \beta_\varepsilon(i, \langle \bar{\alpha} \rangle))^{-1}.$$

Положим $u_\varepsilon(i, \langle \alpha \rangle, k) = P\{x_\varepsilon(v_\varepsilon(i, \langle \alpha \rangle)) = k / x_\varepsilon(0) = i\}$, $i \in \langle \alpha \rangle$, $k \in \langle \alpha \rangle$, т. е. $u_\varepsilon(i, \langle \alpha \rangle, k)$ — вероятность, начиная из состояния $i \in \langle \alpha \rangle$, в момент выхода из $\langle \alpha \rangle$ попасть в состояние k .

Лемма 5.5. Если $\langle \alpha \rangle$ представляет собой s -множество, то для всех $i \in \langle \alpha \rangle$, $k \in \langle \alpha \rangle$

$$u_\varepsilon(i, \langle \alpha \rangle, k) = u_\varepsilon(\langle \alpha \rangle, k) (1 + o(1)), \quad (5.33)$$

где

$$u_\varepsilon(\langle \alpha \rangle, k) = g_\varepsilon(\langle \alpha \rangle)^{-1} \sum_{j \in \langle \alpha \rangle} \pi_\varepsilon(j, \langle \alpha \rangle) p_\varepsilon(j, k).$$

Доказательство. Рассмотрим вначале случай, когда матрице $P_\varepsilon(\langle \alpha \rangle) = \|p_\varepsilon(i, j)\|$, $i, j \in \langle \alpha \rangle$ в пределе соответствует один существенный класс. По формуле полной вероятности несложно получить, что

$$u_\varepsilon(i, \langle \alpha \rangle, k) = \sum_{j \in \langle \alpha \rangle} p_\varepsilon(i, j) u_\varepsilon(j, \langle \alpha \rangle, k) + p_\varepsilon(i, k), \quad i \in \langle \alpha \rangle,$$

отсюда

$$\bar{u}_\varepsilon(\langle \alpha \rangle, k) = (E - P_\varepsilon(\langle \alpha \rangle))^{-1} \bar{p}_\varepsilon(k),$$

где $u_e(\langle \alpha \rangle, k)$ и $\bar{p}_e(k)$ — векторы-столбцы с компонентами соответственно $u_e(i, \langle \alpha \rangle, k)$ и $p_e(i, k)$, $i \in \langle \alpha \rangle$, и, пользуясь представлением (3.27), легко получим (5.32).

При этом заметим, что $\langle \alpha \rangle$ обладает следующим свойством, которое назовем свойством В: если величины $p_e(i, j)$, $i, j \in \langle \alpha \rangle$ изменить на величины порядка $O(\beta_e(i, \bar{\alpha}))$, то при этом величины $p_e(i, \langle \alpha \rangle)$, $i \in \langle \alpha \rangle$ изменятся на $o(1) p_e(i, \langle \alpha \rangle)$, если же $p_e(i, j)$ изменить на величины порядка $o(\beta_e(i, \bar{\alpha}))$, то при этом и величины $g_e(\langle \alpha \rangle)$ и $u_e(\langle \alpha \rangle, k)$ изменятся на $o(1) g_e(\langle \alpha \rangle)$ и $o(1) u_e(\langle \alpha \rangle, k)$ соответственно. Будем одновременно с доказательством леммы доказывать свойство В для любого s -множества $\langle \alpha \rangle$.

Пусть теперь множеству $\langle \alpha \rangle$ в пределе соответствует несколько существенных классов и, возможно, несущественные состояния. Укажем алгоритм, который за конечное число шагов, не более числа состояний в $\langle \alpha \rangle$, позволяет проверить, является ли $\langle \alpha \rangle$ s -множеством.

Алгоритм. Будем последовательно строить на $\langle \alpha \rangle$ связанные компоненты графа переходов в соответствии с различными порядками вероятностей переходов. На первом шаге проведем все переходы ($i \rightarrow j$) такие, что $\lim_{\varepsilon \rightarrow 0} p_e(i, j) > 0$. Получим некоторый ориентированный граф.

Если все $\langle \alpha \rangle$ будет связано и замкнуто, то $\langle \alpha \rangle$ в пределе образует один класс и является s -множеством. В противном случае заменим каждую связную замкнутую область $\langle d \rangle$ одним (укрупненным) состоянием $\{d\}$ с усредненными вероятностями переходов

$$\begin{aligned} p_e(\langle d \rangle, j) &= u_e(\langle d \rangle, j); \\ p_e(j, \langle d \rangle) &= \sum_{i \in \langle d \rangle} p_e(j, i), \quad j \in \langle d \rangle, \end{aligned} \quad (5.34)$$

Получим новый граф с меньшим числом состояний. Далее назовем переход из $\{i\}$ в $\{j\}$ «правильным», если $\lim_{\varepsilon \rightarrow 0} p_e(i, j) (1 - p_e(i, i))^{-1} > 0$. Очевидно, что с вероятностью, стремящейся к единице, при выходе из $\{i\}$ могут осуществляться только «правильные» переходы.

В соответствии с этим на 2-м шаге для нового графа проведем все «правильные» переходы. Вновь получим некоторый ориентированный граф, в котором снова каждую замкнутую связную область заменяем одним состоянием по формулам (5.34) и т. д. Тогда на некотором шаге n (n не превосходит числа состояний в $\langle \alpha \rangle$) придем к одной из двух ситуаций:

1. Все $\langle \alpha \rangle$ приведено к одному состоянию.
2. Нашлось некоторое состояние (возможно укрупненное), из которого «правильный» переход ведет во внешность $\langle \alpha \rangle$.

Докажем теперь, что $\langle \alpha \rangle$ образует s -множество в том и только в том случае, когда имеет место первая ситуация. В этом случае назовем n -порядком $\langle \alpha \rangle$.

Для доказательства нашего утверждения проведем индукцию по n . При $n = 1$ лемма доказана. Пусть $n > 1$. Тогда на $(n-1)$ -м шаге $\langle \alpha \rangle$ будет разбито на множество состояний $\{\langle d_1 \rangle\}, \dots, \{\langle d_n \rangle\}$, матрица «правильных» переходов на котором образует один связный замкнутый класс, а каждое $\{\langle d_j \rangle\}$ — либо отдельное состояние, либо s -множество порядка не более $n-1$. По предположению индукции

$$u_e(i, \langle d_j \rangle, k) = u_e(\langle d_j \rangle, k) (1 + o(1)), \quad i \in \langle d_j \rangle, \\ k \notin \langle d_j \rangle$$

и каждое $\langle d_j \rangle$ удовлетворяет свойству B . Теперь по исходной цепи $x_e(k)$ построим вложенную цепь Маркова $\tilde{x}_e(m)$ следующим образом: пусть $0 < \tau_e(0) < \tau_e(1) < \dots$ — последовательные моменты переходов между областями $\langle d_i \rangle$, т. е. если $x_e(\tau_e(k)) \in \langle d_i \rangle$, то

$$\tau_e(k+1) = \min \{l : l > \tau_e(k), x_e(l) \notin \langle d_i \rangle\}, \quad k \geq 0.$$

Тогда положим $\tilde{x}_e(m) = x_e(\tau_e(m))$, $m \geq 0$. Вероятности переходов цепи $\tilde{x}_e(m)$ по построению имеют вид $\tilde{p}_e(i, j) = P\{\tilde{x}_e(1) = j | \tilde{x}_e(0) = i\} = u_e(i, \langle \alpha \rangle, j) = u_e(\langle d_i \rangle, j) (1 + o(1))$, если $i \in \langle d_i \rangle$, $j \notin \langle d_i \rangle$ (поскольку нас интересует асимптотика вероятности перехода из $\langle \alpha \rangle$, нам достаточно определить вероятности переходов только внутри подмножества $\langle \alpha \rangle$ и считать, что $x_e(0) \in \langle \alpha \rangle$).

Рассмотрим теперь новую цепь $\tilde{\tilde{x}}_e(m)$, вероятности переходов которой имеют вид $\tilde{\tilde{p}}_e(i, j) = u_e(\langle d_i \rangle, j)$, если $i \in \langle d_i \rangle$, $j \notin \langle d_i \rangle$. Так как эти вероятности отличаются от вероятностей переходов цепи $x_e(m)$ на величины $o(1) \times \times (\langle d_i \rangle, j)$ соответственно и цепи $\tilde{x}_e(m)$ в пределе соответствует один класс, то в силу сделанного выше замечания о свойстве B величины $u_e(i, \langle \alpha \rangle, k)$, $k \notin \langle \alpha \rangle$, также изменяется на $o(1) u_e(\langle \alpha \rangle, k)$. Однако у вновь полученной цепи $\tilde{\tilde{x}}_e(m)$ вероятности переходов из всех состояний множества $\langle d_i \rangle$ одинаковы ($u_e(\langle d_i \rangle, j)$) и так для любого $\langle d_i \rangle$. Тем самым цепь $\tilde{\tilde{x}}_e(m)$ можно укрупнить, объединяя

все состояния из $\langle d_l \rangle$ в одно состояние. При этом получим новую цепь $\hat{x}_e(m)$ с множеством состояний $\{\langle d_1 \rangle\}, \dots, \{\langle d_{r_n} \rangle\}$ и вероятностями переходов из $\{\langle d_l \rangle\}$ в $\{\langle d_s \rangle\}$:

$$u_e(\langle d_l \rangle, \langle d_s \rangle) = \sum_{j \in \langle d_s \rangle} u_e(\langle d_l \rangle, j),$$

которой в пределе соответствует один класс состояний. Тем самым легко видеть, что так как по «правильным» переходам из состояний $\{\langle d_l \rangle\}$ новая цепь связна и замкнута, то исходное множество $\langle \alpha \rangle$ образует s -множество.

Нетрудно понять, что в случае второй ситуации, если из $\{\langle d_l \rangle\}$ правильный переход ведет вне $\langle \alpha \rangle$, то найдется состояние $j \notin \langle d_l \rangle$ такое, что для любых $i \in \langle d_l \rangle$

$$\lim_{\varepsilon \rightarrow 0} q_\varepsilon(i, j, \langle \alpha \rangle) < 1,$$

т. е. $\langle \alpha \rangle$ не образует s -множество.

Итак пусть $\langle \alpha \rangle$ s -множество порядка n . Докажем представление (5.33). В силу наших построений согласно предположению индукции вероятности выхода для исходной цепи и цепи $\hat{x}_e(m)$ будут отличаться сомножителем $1 + o(1)$. Теперь согласно предыдущим формулам, так как цепи $\hat{x}_e(m)$ в пределе соответствует один класс, то

$$\hat{u}_e(\langle \alpha \rangle, k) = \hat{g}_e(\langle \alpha \rangle)^{-1} \sum_{l=1}^{r_n} \hat{\pi}_e(\langle d_l \rangle) \hat{u}_e(\langle d_l \rangle, k), \quad (5.35)$$

где $\hat{\pi}_e(\langle d_l \rangle)$; $l = \overline{1, r_n}$ — стационарное распределение цепи $\hat{x}_e(m)$,

$$\hat{g}_e(\langle \alpha \rangle) = \sum_{l=1}^{r_n} \hat{\pi}_e(\langle d_l \rangle) \hat{u}_e(\langle d_l \rangle, \langle \bar{\alpha} \rangle),$$

а

$$\hat{u}_e(\langle d_l \rangle, k) = g_e(\langle d_l \rangle)^{-1} \sum_{j \in \langle d_l \rangle} \pi_e(j, \langle d_l \rangle) p_e(j, k),$$

$$\hat{u}_e(\langle d_l \rangle, \langle \bar{\alpha} \rangle) = \sum_{k \in \langle \alpha \rangle} \hat{u}_e(\langle d_l \rangle, k), \text{ а величины } g_e(\langle d_l \rangle) \text{ и}$$

$\pi_e(j, \langle d_l \rangle)$ определяются согласно формулам (5.32).

Введем теперь цепь Маркова $\hat{x}_e(m)$ с множеством состояний $\{\langle d_1 \rangle\}, \dots, \{\langle d_{r_n} \rangle\}$ и матрицей вероятностей переходов

$$p_e(\langle d_l \rangle, \langle d_j \rangle) = \sum_{i \in \langle d_l \rangle} \pi_e(i, \langle d_l \rangle) \sum_{s \in \langle d_j \rangle} p_e(i, s), \quad l, j = \overline{1, r_n}.$$

Легко видеть, что $p_e(\langle d_l \rangle, \langle d_j \rangle) = g_e(\langle d_l \rangle) u_e(\langle d_l \rangle, \langle d_j \rangle)$ при $l \neq j$. Обозначим теперь через $\hat{\pi}_e(\langle d_l \rangle)$, $l = \overline{1, r_n}$, ста-

ционарное распределение цепи $\hat{x}_e(m)$. Непосредственно не-
сложно проверить, что

$$\hat{\pi}_e(\langle d_l \rangle) = \left(\sum_{j=1}^n \hat{\pi}_e(\langle d_j \rangle) g_e(\langle d_j \rangle)^{-1} \right)^{-1} \pi_e(\langle d_l \rangle) g_e(\langle d_l \rangle)^{-1},$$

$$l = \overline{1, r_n}. \quad (5.36)$$

Поэтому формулу (5.35) можно переписать в виде

$$\hat{u}_e(\langle \alpha \rangle, k) = \hat{g}_e(\langle \alpha \rangle)^{-1} \sum_{l=1}^n \hat{\pi}_e(\langle d_l \rangle) \sum_{j \in \langle d_l \rangle} \pi_e(j, \langle d_l \rangle) \times$$

$$\times p_e(j, k),$$

где

$$\hat{g}_e(\langle \alpha \rangle) = \sum_{l=1}^n \hat{\pi}_e(\langle d_l \rangle) \sum_{j \in \langle d_l \rangle} \pi_e(j, \langle d_l \rangle) \beta_e(j, \langle \bar{\alpha} \rangle).$$

Докажем теперь в виде леммы следующее утверждение,
из которого непосредственно вытекает (5.33).

Лемма 5.6. Если $\langle \alpha \rangle$ представляет собой s -множество
и при нашем построении $i \in \langle d_l \rangle$, то

$$\pi_e(i, \langle \alpha \rangle) = \hat{\pi}_e(\langle d_l \rangle) \pi_e(i, \langle d_l \rangle) (1 + o(1)). \quad (5.37)$$

Для доказательства введем некоторые предварительные
построения. Пусть у нас задана цепь Маркова $x(l)$, $l \geq 0$
с конечным множеством состояний I и матрицей вероят-
ностей переходов $P = \|p_{ij}\|$, $i, j \in I$, при этом цепь не-
приводима и $I = \bigcup_k I_k$, где $I_k \cap I_s = \emptyset$, $k \neq s$. Выберем
конкретную область I_k и пусть $\tau_1 < \tau_2 < \dots$ — последо-
вательные моменты попадания цепи $x(l)$ в состояния I_k ,
т. е.

$$\tau_1 = \min \{l : l \geq 0, x(l) \in I_k\};$$

$$\tau_m = \min \{l : l > \tau_{m-1}, x(l) \in I_k\}, m \geq 1.$$

Рассмотрим последовательность $x^{(k)}(m) = x(\tau_m)$, $m \geq 1$.
Покажем, что $x^{(k)}(m)$ — однородная цепь Маркова со ста-
ционарным распределением

$$\pi_i^{(k)} = \pi_i \left(\sum_{j \in I_k} \pi_j \right)^{-1}, \quad i \in I_k, \quad (5.38)$$

где π_i , $i \in I$ — стационарное распределение исходной цепи
 $x(l)$.

Действительно, так как τ_m , $m \geq 1$ — есть последова-
тельность марковских моментов, то $x^{(k)}(m)$ — цепь Мар-
кова. Обозначим через $p_{ij}^{(k)}$, $i, j \in I_k$ ее переходные веро-

ятности. По формуле полной вероятности нетрудно получить

$$p_{ij}^{(k)} = p_{ij} + \sum_{s \in I_k} p_{is} f_{sj}, \quad (5.39)$$

где f_{sj} — вероятность того, что цепь $x(l)$, начиная из состояния $s \in I_k$, когда либо попадет в множество I_k и j — состояние из I_k в момент попадания. Тогда легко видеть, что величины f_{sj} удовлетворяют системе соотношений

$$f_{sj} = p_{sj} + \sum_{l \in I_k} p_{sl} f_{lj}, \quad s \in I_k, j \in I_k. \quad (5.40)$$

Как известно из теории марковских процессов стационарное распределение $\pi_i^{(k)}$ удовлетворяет системе

$$\pi_j^{(k)} = \sum_{i \in I_k} \pi_i^{(k)} p_{ij}^{(k)}, \quad (5.41)$$

решение которой существует и единственно (в классе вероятностных распределений). Поэтому достаточно проверить, что π_i , $i \in I_k$ — решение (5.41), откуда будет следовать (5.38). Подставляя π_i , $i \in I_k$ в правую часть (5.41) и пользуясь (5.39), (5.40) и соотношением

$$\pi_s = \sum_{i \in I_k} \pi_i p_{is} + \sum_{m \in I_k} \pi_m p_{ms},$$

получим

$$\begin{aligned} \sum_{i \in I_k} \pi_i p_{ij}^{(k)} &= \sum_{i \in I_k} \pi_i (p_{ij} + \sum_{s \in I_k} p_{is} f_{sj}) = \sum_{i \in I_k} \pi_i p_{ij} + \\ &+ \sum_{s \in I_k} \sum_{i \in I_k} \pi_i p_{is} f_{sj} = \sum_{i \in I_k} \pi_i p_{ij} + \\ &+ \sum_{s \in I_k} (\pi_s - \sum_{m \in I_k} \pi_m p_{ms}) (p_{sj} + \sum_{l \in I_k} p_{sl} f_{lj}) = \\ &= \pi_j + \sum_{s \in I_k} \pi_s \sum_{l \in I_k} p_{sl} f_{lj} - \sum_{m \in I_k} \pi_m \sum_{s \in I_k} p_{ms} f_{sj} = \pi_j. \end{aligned}$$

Введем теперь на множествах $I_1, I_2, \dots$ укрупненную цепь Маркова $x(l)$, $l \geq 0$ с множеством состояний $\{I_1\}, \dots, \{I_2\}, \dots$ и матрицей вероятностей переходов

$$p(I_k, I_j) = \sum_{i \in I_k} \pi_i^{(k)} \sum_{l \in I_j} p_{il}, \quad k, j = 1, 2, \dots \quad (5.42)$$

Пусть $\pi(I_k)$, $k = 1, 2, \dots$ — стационарное распределение цепи $\hat{x}(l)$. Докажем, что имеет место соотношение

$$\pi_i = \pi(I_k) \pi_i^{(k)}, \quad i \in I_k, \quad k = 1, 2, \dots \quad (5.43)$$

Действительно, по определению

$$\pi(I_j) = \sum_k \pi(I_k) p(I_j, I_k).$$

Используя (5.42) и (5.38), отсюда нетрудно получить, что

$$\pi(I_k) = \sum_{i \in I_k} \pi_i, \quad k = 1, 2, \dots \quad (5.44)$$

откуда непосредственно следует (5.43).

Вернемся теперь к доказательству соотношения (5.37). Рассмотрим на множестве $\langle d_l \rangle$ цепь Маркова $x_e^{(l)}(m)$, $m \geq 1$, построенную по моментам попадания $x_e(k)$ в состояния $\langle d_l \rangle$. Легко видеть в силу формул (5.32), что для переходных вероятностей $p_e^{(l)}(i, j)$ цепи $x_e^{(l)}(m)$ справедливо соотношение

$$p_e^{(l)}(i, j) = \tilde{p}_e(i, j, \langle d_l \rangle) + O(\beta_e(i, \langle \bar{d}_l \rangle)).$$

Теперь $\langle d_l \rangle$ — s -множество порядка не более $n - 1$, и по предположению индукции в силу свойства B стационарное распределение $\pi_e^{(l)}(i)$, $i \in \langle d_l \rangle$ цепи $x_e^{(l)}(m)$ удовлетворяет соотношению $\pi_e^{(l)}(i) = \pi_e(i, \langle d_l \rangle)(1 + o(1))$. Отсюда, используя равенство (5.43) и (5.44), получим в наших обозначениях (5.37). Как уже отмечалось, (5.37) непосредственно влечет (5.33). Для доказательства свойства B на n -м шаге заметим, что при соответствующем изменении вероятностей $p_e(i, j)$ на величины $O(\beta_e(i, \langle \bar{\alpha} \rangle))$, вероятности переходов $\tilde{p}_e(i, \langle d_l \rangle)$, $i \in \langle d_l \rangle$ изменяется на величины $O(\beta_e(i, \langle d_l \rangle))$, следовательно, $\pi_e(i, \langle d_l \rangle)$ изменяется на $o(1)\pi_e(i, \langle d_l \rangle)$, при этом вероятности переходов цепи $\hat{x}_e(m)$ и $\hat{\pi}_e(\langle d_l \rangle, \langle d_s \rangle)$ изменяется на $o(1)$, т. е. величины $\hat{\pi}_e(\langle d_l \rangle)$ изменяется на $o(1)$. Отсюда, согласно формулам (5.36) и (5.37), величины $\pi_e(i, \langle \alpha \rangle)$ также изменяется на $o(1)\pi_e(i, \langle \alpha \rangle)$, что доказывает свойство B .

Заметим, что формулы (5.35) — (5.37) дают алгоритм последовательного вычисления стационарного распределения и переходных вероятностей с точностью до величин более высокого порядка малости.

Полученные результаты дают возможность также распространить результаты п. 5.2 на тот случай, когда подмножество $\{1, 2, \dots, r\}$ образует s -множество.

Пусть задан ПМП $\xi_e(t)$ с множеством состояний $\{0, 1, \dots, r\}$, матрицей вероятностей переходов вложенной цепи $p_e(k, j)$, $k, j = \overline{0, r}$, и набором времен перехо-

дов $\pi_\varepsilon(k, j)$, $k, j = \overline{0, r}$. Предположим, что подмножество $\{1, \dots, r\}$ образует s -множество. Положим

$$g_\varepsilon = \sum_{k=1}^r \pi_\varepsilon(k) p_\varepsilon(k, 0),$$

где $\pi_\varepsilon(k)$, $k = \overline{1, r}$ — стационарное распределение для цепи с матрицей $\|p_\varepsilon(i, j)(1 - p_\varepsilon(i, 0))^{-1}\|$, $i, j = \overline{1, r}$. Предположим далее, что

$$g_\varepsilon^{-1} \pi_\varepsilon(k) p_\varepsilon(k, 0) \rightarrow b_k, \quad k = \overline{1, r},$$

и существует нормирующий множитель β_ε такой, что

$$\begin{aligned} \varphi_\varepsilon(\beta_\varepsilon u, k, j) &= 1 + g_\varepsilon a_{kj}(u) + o(g_\varepsilon), \quad k, j = \overline{1, r}; \\ \varphi_\varepsilon(\beta_\varepsilon u, k, 0) &\rightarrow \rho_k(u), \quad k = \overline{1, r} \end{aligned} \quad (5.45)$$

(здесь как и ранее в п. 5.2 $\varphi_\varepsilon(u, k, j) = M \exp \{iu \pi_\varepsilon(k, j)\}$). Тогда при указанных предположениях справедлива

Теорема 5.4. Распределение величины $\beta_\varepsilon \Omega_\varepsilon(j)$ независимо от начального состояния $j = \overline{1, r}$ слабо сходится к распределению с характеристической функцией

$$\sum_{k=1}^r b_k \rho_k(u) (1 - \sum_{k,j=1}^r \pi_0(k) p_0(k, j) a_{kj}(u))^{-1},$$

где

$$\pi_0(k) = \lim_{\varepsilon \rightarrow 0} \pi_\varepsilon(k); \quad p_0(k, j) = \lim_{\varepsilon \rightarrow 0} p_\varepsilon(k, j); \quad k, j = \overline{1, r}.$$

Доказательство вновь проводится последовательно индукцией по порядку s -множества $\{1, \dots, r\}$. При $n = 1$ утверждение было доказано в п. 5.2. Укажем здесь одно обобщение теоремы 5.3, которое будет использоваться при доказательстве. Пусть выделено некоторое подмножество $\langle \alpha \rangle$. Обозначим

$$\Omega_\varepsilon(k) = \inf \{t : t > 0, \xi_\varepsilon(t) \in \langle \alpha \rangle\},$$

где $\xi_\varepsilon(0) = k \in \langle \alpha \rangle$, а

$$\chi_\varepsilon(\langle \beta \rangle) = \chi \{\xi_\varepsilon(\Omega_\varepsilon(k)) \in \langle \beta \rangle\},$$

где $\langle \beta \rangle \cap \langle \alpha \rangle = \emptyset$, т. е. $\chi_\varepsilon(\langle \beta \rangle)$ — индикатор попадания в подмножество $\langle \beta \rangle$ в момент выхода из $\langle \alpha \rangle$. Обозначим

$$\psi_\varepsilon(u, k, \langle \alpha \rangle, \langle \beta \rangle) = M \exp \{iu \Omega_\varepsilon(k)\} \chi_\varepsilon(\langle \beta \rangle).$$

Подобно соотношениям (5.20) и (5.21), нетрудно получить, что

$$\begin{aligned} \psi_\varepsilon(u, k, \langle \alpha \rangle, \langle \beta \rangle) &= \sum_{j \in \langle \alpha \rangle} p_\varepsilon(k, j) \varphi_\varepsilon(u, k, j) \times \\ &\times \psi_\varepsilon(u, j, \langle \alpha \rangle, \langle \beta \rangle) + \sum_{l \in \langle \beta \rangle} p_\varepsilon(k, l) \varphi_\varepsilon(u, k, l). \end{aligned} \quad (5.46)$$

Из (5.46), пользуясь представлением (5.27), подобно теореме 5.3 можно доказать утверждение:

Лемма 5.7. Если $\langle \alpha \rangle$ в пределе замкнуто и неприводимо и существует нормирующий множитель β_ε такой, что

$$\varphi_\varepsilon(\beta_\varepsilon u, k, j) = 1 + g_\varepsilon(\langle \alpha \rangle) \alpha_\varepsilon a_{kj}(u) + o(g_\varepsilon(\langle \alpha \rangle) \alpha_\varepsilon),$$

$$k, j \in \langle \alpha \rangle,$$

$$\varphi_\varepsilon(\beta_\varepsilon u, k, l) = 1 + \alpha_\varepsilon a_{kl}(u) + o(\varepsilon), \quad k \in \langle \alpha \rangle, \quad l \in \langle \beta \rangle,$$

где $\alpha_\varepsilon \rightarrow 0$, то для любого $k \in \langle \alpha \rangle$

$$\begin{aligned} \psi_\varepsilon(\beta_\varepsilon u, k, \langle \alpha \rangle, \langle \beta \rangle) &= u_\varepsilon(\langle \alpha \rangle, \langle \beta \rangle) \times \\ &\times (1 + \alpha_\varepsilon \sum_{i, j \in \langle \alpha \rangle} \pi_\varepsilon(i, \langle \alpha \rangle) p_\varepsilon(i, j) a_{ij}(u) + \\ &+ u_\varepsilon(\langle \alpha \rangle, \langle \beta \rangle)^{-1} \sum_{i \in \langle \alpha \rangle} \sum_{l \in \langle \beta \rangle} \pi_\varepsilon(i, \langle \alpha \rangle) p_\varepsilon(i, l) \times \\ &\times a_{il}(u)) + o(\alpha_\varepsilon). \end{aligned} \quad (5.47)$$

Из представления (5.47) видно, что главная часть разложения характеристической функции времени перехода из $\langle \alpha \rangle$ в $\langle \beta \rangle$ не зависит от начального состояния $k \in \langle \alpha \rangle$. Далее, проводя индукцию и последовательно укрупняя подмножества состояний в $\langle \alpha \rangle$ аналогично тому, как это делалось при доказательстве леммы 5.5 и пересчитывая стационарные характеристики по формулам (5.36), (5.37), придем к тому, что на $n-1$ -м шаге подмножество $\{1, \dots, r\}$ будет состоять из r_n укрупненных состояний $\{\langle d_1 \rangle\}, \dots, \{\langle d_{r_n} \rangle\}$, характеристические функции переходов между которыми будут иметь разложения типа (5.47), при этом укрупненной цепи $\hat{x}_\varepsilon(m)$ будет соответствовать один класс. После этого можно будет воспользоваться теоремой 5.3 и формулами (5.37), что доказывает нашу теорему.

Положим $\varphi_\varepsilon(u, k, j) \equiv 1$, $k, j = \overline{1, r}$, а $\bar{\varphi}_\varepsilon(u)$ и $\bar{\varphi}_\varepsilon(u)$ — вектора-столбцы с компонентами $M \exp\{iu\Omega_\varepsilon(j)\}$, $j = \overline{1, r}$, и $p_\varepsilon(j, 0) \varphi_\varepsilon(u, j, 0)$, $j = \overline{1, r}$, соответственно. Тогда из результатов теоремы 5.3, согласно формулам (5.21), следует

$$\begin{aligned} \bar{\psi}_\varepsilon(u) &= (E - P_\varepsilon)^{-1} \bar{\varphi}_\varepsilon(u) = \\ &= \frac{1}{g_\varepsilon} \sum_{k=1}^r \pi_\varepsilon(k) p_\varepsilon(k, 0) \varphi_\varepsilon(u, k, 0) \bar{1} (1 + o(1)) \end{aligned}$$

и так как $\varphi_\varepsilon(u, k, 0)$ — произвольные характеристические функции, отсюда следует, что элементы матрицы $(E - P_\varepsilon)^{-1}$ эквиваленты величинам $g_\varepsilon^{-1} \pi_\varepsilon(k)$, $k = \overline{1, r}$. Тем самым получим:

Следствие 5.5. Если множество $\{1, \dots, r\}$ является $\mathcal{S}$ -множеством, то

$$(E - P_\varepsilon)^{-1} = \frac{1}{g_\varepsilon} \pi_\varepsilon (1 + o(1)),$$

где π_ε — матрица с одинаковыми строчками вида $\{\pi_\varepsilon(1), \dots, \pi_\varepsilon(r)\}$, а запись $\pi_\varepsilon(1 + o(1))$ означает, что соответствующие элементы матрицы $\pi_\varepsilon(1 + o(1))$ и π_ε отличаются сомножителями порядка $1 + o(1)$.

Замечание. Непосредственно из доказательства теорем 5.3 и 5.4 вытекает, что в преобразованиях везде участвуют величины $\pi_\varepsilon(k) \times \times p_\varepsilon(k, j) a_{kj}(u)$. Поэтому условие (5.45) на величины $\varphi_\varepsilon(\beta_\varepsilon u, k, j)$, $k, j = \overline{1, r}$, можно ослабить, заменив следующим:

$$g_\varepsilon^{-1} \pi_\varepsilon(k) p_\varepsilon(k, j) (1 - \varphi_\varepsilon(\beta_\varepsilon u, k, j)) \rightarrow \hat{a}_{kj}(u), \quad k, j = \overline{1, r}.$$

При этом характеристическая функция предельного закона будет иметь вид

$$\sum_{k=1}^r b_k \rho_k(u) (1 + \sum_{k,j=1}^r \hat{a}_{kj}(u))^{-1}.$$

Рассмотрим один пример расчета характеристик конкретной цепи.

Пример 5.1. Пусть цепь Маркова $x_\varepsilon(k)$, $k \geq 0$, принимает значения в множестве (i, q) , $i = \overline{1, r}$, $q = \overline{0, m+1}$, причем вероятности переходов $p_\varepsilon((i, q), (j, z))$ ведут себя следующим образом:

1) $p_\varepsilon((i, 0), (j, 0)) \rightarrow p_{ij}$, $i, j = \overline{1, r}$, и матрица $\|p_{ij}\|$, $i, j = \overline{1, r}$, неприводима;

2) $p_\varepsilon((i, q), (j, q+1)) = \varepsilon \alpha_{ij}^{(q)} + o(\varepsilon)$, $q = \overline{0, m}$; $i, j = \overline{1, r}$;

3) $p_\varepsilon((i, q), (f, q)) \rightarrow 0$, $i, f = \overline{1, r}$, $q \geq 1$;

4) $p_\varepsilon((i, q), (j, z)) = 0$ при $z - q \geq 2$, $i, j = \overline{1, r}$.

Будем говорить, что множество состояний $\{(i, q), i = \overline{1, r}\}$ представляет собой q -й уровень, $q = \overline{0, m+1}$. Тогда условия 1—4 означают, что нулевой уровень в пределе образует один существенный класс, далее вероятности перехода с q -го на $(q+1)$ -й уровень порядка ε , на q -м уровне вероятности переходов стремятся к нулю, а суммарная вероятность за 1 шаг вернуться с уровня $q \geq 1$ на нижние уровни стремится к единице. Выделим подмножество состояний $\langle \alpha \rangle = \{(i, q), i = \overline{1, r}, q \leq m\}$. Вычислим главные части стационарных вероятностей $\pi_\varepsilon((i, q), \langle \alpha \rangle) = \pi_\varepsilon(i, q)$, а также величину $g_\varepsilon \langle \alpha \rangle$.

Обозначим $P = \| p_{ij} \|$, $i, j = \overline{1, r}$, $A^{(q)} = \| \alpha_{ij}^{(q)} \|$, $i, j = \overline{1, r}$, $q = \overline{0, m}$, $\{\pi_k, k = \overline{1, r}\}$ — стационарное распределение для цепи с матрицей P , $\bar{\pi} = \{\pi_1, \dots, \pi_r\}$ и $\pi_\varepsilon^{(q)} = \{\pi_\varepsilon(i, q), i = \overline{1, r}\}$ — вектора-строки.

Лемма 5.8. В предположениях 1—4

$$\pi_\varepsilon^{(q)} = \varepsilon^q \bar{\pi} A^{(0)} \dots A^{(q-1)} + o(\varepsilon^q), \quad q = \overline{1, m}, \quad (5.48)$$

а

$$g_\varepsilon(\langle \alpha \rangle) = \varepsilon^{m+1} \bar{\pi} A^{(0)} \dots A^{(m)} \bar{1} + o(\varepsilon^{m+1}). \quad (5.49)$$

Доказательство. Нетрудно видеть, что $\langle \alpha \rangle$ образует s -множество порядка $m+1$. На 1-м шаге нулевой уровень укрупнится в состояние $\{\langle 0 \rangle\}$ со стационарной вероятностью перехода за 1 шаг

$$p_\varepsilon(\{\langle 0 \rangle\}, (i, 1)) = \varepsilon \sum_{k=1}^r \pi_k \alpha_{ki}^{(0)} (1 + o(1)). \quad (5.50)$$

На 2-м шаге укрупняются нулевой и первый уровни в состояние $\{\langle 1 \rangle\}$. При этом несложно вычислить стационарное распределение на множестве $\{\langle 0 \rangle\}, \{(i, 1), i = \overline{1, r}\}$. Так как $p_\varepsilon(\{(i, 1), \langle 0 \rangle\}) \rightarrow 1$, то, согласно (5.50),

$$\pi_\varepsilon(i, 1) = \pi_\varepsilon(\{\langle 0 \rangle\}) p_\varepsilon(\{\langle 0 \rangle\}, (i, 1)) (1 + o(1)).$$

И так как $\pi_\varepsilon(\{\langle 0 \rangle\}) + \sum_{i=1}^r p_\varepsilon(\{\langle 0 \rangle\}, (i, 1)) = 1$, то

$$\pi_\varepsilon(i, 1) = \varepsilon \sum_{k=1}^r \pi_k \alpha_{ki}^{(0)} (1 + o(1)), \quad i = \overline{1, r}.$$

Следовательно, после 2-го шага

$$p_\varepsilon(\{\langle 1 \rangle\}, (i, 2)) = \sum_{\substack{k=\overline{1, r} \\ s=0, 1}} \pi_\varepsilon(k, s) p_\varepsilon((k, s), (i, 2)) (1 + o(1)) = \varepsilon^2 \sum_{k=1}^r \pi_k \sum_{j=1}^r \alpha_{kj}^{(0)} \alpha_{ji}^{(1)} (1 + o(1)),$$

а стационарное распределение аналогично имеет вид

$$\begin{aligned} \pi_\varepsilon(i, 2) &= \pi_\varepsilon(\{\langle 1 \rangle\}) p_\varepsilon(\{\langle 1 \rangle\}, (i, 2)) (1 + o(1)) = \\ &= p_\varepsilon(\{\langle 1 \rangle\}, (i, 2)) (1 + o(1)). \end{aligned}$$

Проводя дальнейшие построения по индукции, легко получить формулы (5.48), (5.49).

Заметим, что эти формулы развивают применительно к подобным схемам принцип монотонности траекторий, предложенный в (9), суть которого заключается в том, что главные части стационарных вероятностей и вероятностей выхода достаточно считать только на монотонных траекториях, т. е. траекториях, которые в данном случае имеют вид $(i_0, 0) \rightarrow (i_1, 1) \rightarrow (i_2, 2) \rightarrow \dots \rightarrow (i_{m+1}, m+1)$.

Как следствие леммы 5.8 и теоремы 5.4 приведем один результат, который в дальнейшем будет использован при асимптотическом анализе систем типа $M|M|n$ с быстрым обслуживанием.

Следствие 5.6. Пусть ПМП $\xi_\varepsilon(t)$ задан вложенной цепью $x_\varepsilon(k)$, $k \geq 0$, удовлетворяющей условиям 1—4, и временами переходов из (l, s) в (j, z) $\tau_\varepsilon((l, s), (j, z))$, удовлетворяющими условию

$$M \exp \{i\theta \beta_\varepsilon \tau_\varepsilon((l, s), (j, z))\} = 1 + a_{lj}(s, z, \theta) \varepsilon^{m+1} + o(\varepsilon^{m+1}),$$

где β_ε — некоторый нормирующий множитель. Через $\Omega_\varepsilon(j, s)$ обозначим момент первого достижения уровня $m+1$, если $\xi_\varepsilon(0) = (j, s)$, $(s \leq m)$. Тогда

$$\lim_{\varepsilon \rightarrow 0} M \exp \{i\theta \beta_\varepsilon \Omega_\varepsilon(j, s)\} = (1 - A(\theta))^{-1},$$

где

$$A(\theta) = \sum_{k,j=1}^r \pi_k p_{kj} a_{kj}(0, 0, \theta) (\bar{\pi} A^{(0)} \dots A^{(m)} \bar{1})^{-1}.$$

В частности, если $a_{lj}(s, z, \theta) = i\theta m_{lj}(s, z)$ ($i = +\sqrt{-1}$), то предельное распределение будет показательным с параметром

$$\bar{\pi} A^{(0)} \dots A^{(m)} \bar{1} \left(\sum_{k,j=1}^r \pi_k p_{kj} m_{kj}(0, 0) \right)^{-1}.$$

5.4. Системы с быстрым обслуживанием

В настоящем параграфе будет показано, как можно с использованием результатов о времени пребывания ПМП в s -множестве исследовать асимптотическое поведение момента первой потери требования в СМО полумарковского типа с быстрым обслуживанием, т. е. в СМО, поведение которых можно описать некоторым вспомогательным ПМП, а времена обслуживания стремятся по вероятности к нулю.

П. 1. Рассмотрим систему $M|M|n$. Входящий поток представляет собой пуассоновский поток, заданный на конечной цепи Маркова и строится следующим образом: задана управляющая цепь Маркова $x(t)$, $t \geq 0$ с непрерывным временем и неприводимым конечным числом состояний $\{1, 2, \dots, r\}$, которая задается плотностями вероятностей переходов $\{a_{ij}, i, j = \overline{1, r}, i \neq j\}$ ($a_{ii} = \sum_{j \neq i} a_{ij}$). Кроме того, задана совокупность положительных параметров $\{\lambda_k, k = \overline{1, r}\}$. Тогда входящий поток $v(t)$ на фиксированной траектории $x(t)$ определяется как неоднородный

пуассоновский процесс с мгновенным значением параметра $\lambda_{x(t)}$. Это означает, что если на промежутке $[t, u]$, $x(s) = k$, то на этом промежутке $v(t)$ получает приращения пуассоновского потока с параметром λ_k . В п. 6.1 двумерный процесс $\{x(t), v(t)\}$, $t \geq 0$, есть марковский процесс, однородный по второй компоненте.

Система имеет n различных обслуживающих устройств, стоящих по порядку, при этом i -е устройство обслуживает по показательному закону с параметром μ_i . Кроме того, система имеет m мест для ожидания. Требования, приходящие в систему, поступают на обслуживание на прибор, имеющий наименьший номер среди свободных, если таковых нет, то становятся в очередь. Если заняты все приборы и места для ожидания, требование теряется.

Предположим, что обслуживание происходит асимптотически быстро, т. е. $\mu_i = \mu_\varepsilon(i) \rightarrow \infty$ при $\varepsilon \rightarrow 0$. Пусть для наглядности

$$\mu_\varepsilon(i) = \varepsilon^{-1} b_i, \quad i = \overline{1, r}.$$

Через $\Omega_\varepsilon(k, q)$ обозначим момент 1-й потери требования, если в начальный момент $x(0) = k$ и в системе было q требований ($q \leq n + s$). Будем исследовать предельное поведение $\Omega_\varepsilon(k, q)$ при $\varepsilon \rightarrow 0$.

Введем следующие вспомогательные характеристики. Пусть $Q_\varepsilon(t)$ — количество требований в момент t в системе $M|n|M|\infty$ (т. е. в предположении, что требования не теряются). Рассмотрим двумерный процесс $z_\varepsilon(t) = \{x(t), Q_\varepsilon(t)\}$, $t \geq 0$. Легко видеть, что $z_\varepsilon(t)$ — цепь Маркова с непрерывным временем со значениями в множестве состояний $\{(i, s), i = \overline{1, r}, s = \overline{0, \infty}\}$. При этом $\Omega_\varepsilon(k, q)$ в таких терминах есть момент первого выхода из подмножества $\langle \alpha \rangle = \{(i, s), i = \overline{1, r}, s = \overline{0, n+m}\}$, если $z_\varepsilon(0) = (k, q)$.

Тем самым задача исследования момента первой потери требования сведена к задаче исследования времени пребывания цепи Маркова в подмножестве состояний.

Вычислим характеристики $z_\varepsilon(t)$. Нетрудно вычислить по построению, что на малом промежутке времени $[t, t+h]$ возможны следующие переходы $z_\varepsilon(t)$:

$$(i, s) \xrightarrow{h} \begin{cases} (j, s) & \text{с вероятностью } a_{ij}(h) + o(h), \quad j \neq i \\ (i, s+1) & \text{с вероятностью } \lambda_i h + o(h), \\ (i, s-1) & \text{с вероятностью } \varepsilon^{-1} \sum_{k=1}^{s_n} b_k h + o(h), \\ (i, s) & \text{с вероятностью } 1 - h(\lambda_i + a_{ii} + \varepsilon^{-1} \times \\ & \times \sum_{k=1}^s b_k) + o(h). \end{cases}$$

Здесь $s_n = \min(s, n)$, а при $s = 0$ считаем $\sum_1^0 = 0$.

При этом в состоянии (i, s) цепь сидит показательное время $\tau_\varepsilon(i, s)$ с параметром $\lambda_i + a_{ii} + \varepsilon^{-1} \sum_{k=1}^{s_n} b_k$, а вероятности переходов вложенной для $z_\varepsilon(t)$ цепи $z_\varepsilon(t)$, $t \geq 0$, имеют вид (подобно тому, как считались соответствующие характеристики для процессов размножения и гибели в п. 1.5)

$$\begin{aligned} p_\varepsilon((i, s), (j, s)) &= a_{ij} (\lambda_i + a_{ii} + \varepsilon^{-1} \sum_{k=1}^{s_n} b_k)^{-1}; \\ p_\varepsilon((i, s), (i, s+1)) &= \lambda_i (\lambda_i + a_{ii} + \varepsilon^{-1} \sum_{k=1}^{s_n} b_k)^{-1}; \\ p_\varepsilon((i, s), (i, s-1)) &= \varepsilon^{-1} \sum_{k=1}^{s_n} b_k (\lambda_i + a_{ii} + \varepsilon^{-1} \sum_{k=1}^{s_n} b_k)^{-1}. \end{aligned}$$

Из полученных формул следует, что при $\varepsilon \rightarrow 0$

$$\begin{aligned} p_\varepsilon((i, s), (j, s)) &= O(\varepsilon), \quad s \geq 1, \\ p_\varepsilon((i, s), (i, s+1)) &= \varepsilon \lambda_i (\sum_{k=1}^{s_n} b_k)^{-1} (1 + o(1)), \quad s \geq 1, \\ p_\varepsilon((i, s), (i, s-1)) &\rightarrow 1, \quad s \geq 1. \quad \text{При } s = 0 \\ p_\varepsilon((i, 0), (j, 0)) &= \frac{a_{ij}}{\lambda_i + a_{ii}}, \quad p_\varepsilon((i, 0), (i, 1)) = \frac{\lambda_i}{\lambda_i + a_{ii}}, \\ i &= \overline{1, r}. \end{aligned}$$

Тем самым попадем в условия примера 3.1 п. 5.3, при этом роль нулевого уровня играет множество $\{(i, 0) \mid i = \overline{1, r}\}$, а роль q -го уровня — $\{(i, q+1) \mid i = \overline{1, r}\}$.

Обозначим через $\pi_\varepsilon(i, q)$ стационарную вероятность состояния (i, q) , $\pi_0(i, q) = \lim_{\varepsilon \rightarrow 0} \pi_\varepsilon(i, q)$. Так как нулевой уровень в пределе образует один существенный класс, то величины $\pi_0(i, q)$, $i = \overline{1, r}$, $q = 0, 1$ удовлетворяют системе уравнений, которая получается в теории марковских цепей из системы прямых дифференциальных уравнений Колмогорова подобно тому, как это делалось для процессов размножения и гибели в п. 1.5:

$$\begin{aligned} \pi_0(j, 0) &= \sum_{i \neq j} \pi_0(i, 0) \frac{a_{ij}}{\lambda_i + a_{ii}} + \pi_0(j, 1); \\ \pi_0(j, 1) &= \pi_0(j, 0) \frac{\lambda_j}{\lambda_j + a_{jj}}, \quad j = \overline{1, r}. \end{aligned}$$

Подставляя 2-е уравнение в 1-е, получим

$$\pi_0(j, 0) \frac{a_{jj}}{\lambda_j + a_{jj}} = \sum_{i \neq j} \pi_0(i, 0) \frac{a_{ij}}{\lambda_i + a_{ii}}, \quad j = \overline{1, r}. \quad (5.51)$$

Обозначим через π_k , $k = \overline{1, r}$, стационарное распределение управляющей цепи $x(t)$. Поскольку

$$\pi_i a_{ij} = \sum_{i \neq j} \pi_i a_{ij}, \quad j = \overline{1, r},$$

то из (5.51) следует, что $\pi_0(i, 0) = B\pi_i(\lambda_i + a_{ii})$, $i = \overline{1, r}$. Далее $\pi_0(i, 1) = B\pi_i\lambda_i$ и, учитывая нормирующее условие $\sum_{i=0}^r (\pi_0(i, 0) + \pi_0(i, 1)) = 1$, легко получим:

$$\pi_0(i, 0) = B\pi_i(\lambda_i + a_{ii});$$

$$\pi_0(i, 1) = B\pi_i\lambda_i, \quad i = \overline{1, r},$$

где $B = (\sum_{k=1}^r \pi_k (2\lambda_k + a_{kk}))^{-1}$.

Теперь по формулам (5.48), (5.49) нетрудно получить, что

$$\pi_e(i, q) = e^{q-1} B \pi_i \lambda_i^q \prod_{s=1}^{q-1} (\sum_{k=1}^s b_k)^{-1} (1 + o(1)), \quad q \geq 1,$$

а

$$g_e(\langle \alpha \rangle) = e^{n+m} B \left(\sum_{k=1}^n b_k \right)^{-m-1} \prod_{s=1}^{n-1} \left(\sum_{j=1}^s b_j \right)^{-1} \times \\ \times \sum_{i=1}^r \pi_i \lambda_i^{n+m+1} (1 + o(1)),$$

где $\prod_1^0 = 1$.

Положим теперь $\beta_e = e^{n+m}$. Тогда, учитывая показательность $\tau_e(j, s)$, легко получим при фиксированном u

$$M \exp \{ i e^{n+m} u \tau_e(j, 0) \} = 1 + e^{n+m} \frac{i u}{\lambda_j + a_{jj}} (1 + o(1));$$

$$M \exp \{ i e^{n+m} u \tau_e(j, s) \} = 1 + o(e^{n+m}), \quad s > 0.$$

Тогда, используя следствие 5.6, непосредственно получим следующее утверждение: (5.1) для системы $M_u | M | n | m$ при указанных предположениях для любых $k = \overline{1, r}$, $q \leq n + m$ распределение величины $e^{n+m} \Omega_e(k, q)$ будет слабо сходиться к показательному с параметром

$$\Lambda = \left(\sum_{k=1}^n b_k \right)^{-m-1} \prod_{s=1}^{n-1} \left(\sum_{j=1}^s b_j \right)^{-1} \sum_{i=1}^r \pi_i \lambda_i^{n+m+1}.$$

Отметим, в частности, что для системы $M | M | n | m$ с пуассоновским входящим потоком с параметром λ и

одинаковыми приборами с интенсивностью обслуживания $\epsilon^{-1}b$, получим предельное значение Λ в виде

$$\Lambda = \frac{1}{nmn!} \left(\frac{\lambda}{b} \right)^{n+m+1}.$$

П. 2. Рассмотрим систему $SM|M|n|m$, которая является в некотором смысле более общей, чем рассмотренная в предыдущем разделе. Входящий поток управляется некоторым конечным ПМП и строится следующим образом: пусть $\xi(t)$ — ПМП в неприводимом множестве состояний $\{1, \dots, r\}$, который задается набором времен сидения в состояниях $\{\tau_k, k = \overline{1, r}\}$ и вложенной цепью Маркова $\xi_l, l \geq 0$ с матрицей вероятностей переходов $P = \|p_{ij}\|, i = \overline{1, r}$. Через $0 = \Theta_0 < \Theta_1 < \dots$ обозначим последовательные моменты скачков $\xi(t)$. Тогда требования в систему поступают в моменты $\Theta_l, l \geq 0$, т. е. входящий поток $v(t)$ — есть число скачков $\xi(t)$ на промежутке $[0, t]$. Система имеет n одинаковых обслуживающих устройств, при этом время обслуживания распределено по показательному закону с параметром $\mu_e = \mu/\epsilon$, и есть m мест для ожидания.

Будем исследовать предельное поведение $\Omega_\epsilon(k, q)$ при $\epsilon \rightarrow 0$. Вновь через $\Theta_\epsilon(t)$ обозначим количество требований в системе в момент t в предположении, что $m = \infty$, т. е. требования не теряются. Введем двумерный процесс $\eta_\epsilon(t) = \{\xi(t), \Theta_\epsilon(t)\}, t \geq 0$. Подобно тому, как исследовалась система $G|m|1$ (п. 2.3), нетрудно получить, что моменты $\theta_1, \theta_2, \dots$ — есть марковские моменты для $\eta_\epsilon(t)$, т. е. последовательность $\zeta_\epsilon(k) = \eta_\epsilon(\theta_k + 0), k \geq 1$, есть цепь Маркова. Построим вспомогательный ПМП $\bar{\eta}_\epsilon(t) = \{\xi(t), \overline{Q_\epsilon(t)}\}, t \geq 0$, у которого на каждом промежутке $[\theta_k, \theta_{k+1})$ $\overline{Q_\epsilon(t)} = Q_\epsilon(\theta_k + 0), k \geq 1$. Поскольку в момент θ_k возможно увеличение очереди не более чем на одно требование, то очевидно, что момент потери требования совпадает с моментом выхода ПМП $\bar{\eta}_\epsilon(t)$ из подмножества $\langle \alpha \rangle = \{(i, s), i = \overline{1, r}, s = 0, n + m\}$. Вычислим теперь переходные характеристики $\bar{\eta}_\epsilon(t)$. Обозначим $q_\epsilon(k) = Q_\epsilon(\theta_k + 0)$. Тогда последовательность $\{\xi_k, q_\epsilon(k)\}, k \geq 1$, определяет вложенную цепь $\zeta_\epsilon(k)$ для ПМП $\bar{\eta}_\epsilon(t)$ (отметим, что при нашем построении $q_0(k) \geq 1$). Обозначим вероятности переходов вложенной цепи из (i, s) в (i, q) через $p_\epsilon((i, s), (i, q))$, а время перехода — $\tau_\epsilon((i, s), (j, q))$. Пусть $G_k(x) = p\{\tau_k < x\}$. Будем предполагать, что

$$(G_k(0+) = 0, k = \overline{1, r}. \quad (5.52)$$

Введем величины

$$v_{\varepsilon}^{(k)}(s) = \int_0^{\infty} e^{-\mu s_n x/\varepsilon} dG_k(x), \quad k = \overline{1, r}, \quad s = 1, 2, \dots,$$

где $s_n = \min(s, n)$. Легко видеть, что

$$p_{\varepsilon}((k, s), (j, s+1)) = p_{kj} v_{\varepsilon}^{(k)}(s), \quad j, k = \overline{1, r}, \quad s \geq 1,$$

причем при $\varepsilon \rightarrow 0$ $v_{\varepsilon}^{(k)}(s) \rightarrow 0$, $s \geq 1$ (в силу (5.52)). Остальные переходные вероятности, как легко видеть, удовлетворяют следующим соотношениям:

$$p_{\varepsilon}((k, s), (j, s)) \rightarrow 0, \quad k, j = \overline{1, r}, \quad s > 1,$$

и

$$p_{\varepsilon}((k, s), (j, s+2)) \equiv 0, \quad k, j = \overline{1, r}, \quad s \geq 0.$$

При этом $p_{\varepsilon}((k, 0), (j, 0)) = p_{kj}(1 - v_{\varepsilon}^{(k)}(1))$. Тем самым вновь попадаем в условия примера 3.1 п. 5.3. Здесь s -му уровню соответствует подмножество $\{(i, s+1), i = \overline{1, r}\}$, $s = 0, 1, \dots$. Предположим, что величины $v_{\varepsilon}^{(k)}(s)$ сравнимы между собой, т. е. существуют константы $b_k^{(s)}$ и сомножитель $\gamma_{\varepsilon} \rightarrow 0$ такие, что

$$v_{\varepsilon}^{(k)}(s) = b_k^{(s)} \gamma_{\varepsilon} (1 + o(1)), \quad k = \overline{1, r}, \quad s \geq 1. \quad (5.53)$$

Тогда по формулам (5.48) и (5.49) нетрудно вычислить, что

$$\pi_{\varepsilon}(i, 1) = \pi_i(1 + o(1)), \quad i = \overline{1, r},$$

где π_i , $i = \overline{1, r}$ — стационарное распределение для вложенной цепи ξ_l , $l \geq 0$. Далее обозначим $A^{(s)} = \|\pi_{ij} b_i^{(s)}\|$, $i, j = \overline{1, r}$. Отсюда, если $\bar{\pi}_{\varepsilon}^{(s)} = \{\pi_{\varepsilon}(i, s), i = \overline{1, r}\}$ и $\bar{\pi} = \{\pi_1, \dots, \pi_r\}$ — векторы-строки, то

$$\bar{\pi}_{\varepsilon}^{(s)} = \gamma_{\varepsilon}^{s-1} \bar{\pi} A^{(1)} \dots A^{(s-1)} (1 + o(1)), \quad s \geq 1,$$

а

$$g_{\varepsilon}(\langle \alpha \rangle) = \gamma_{\varepsilon}^{n+m} \bar{\pi} A^{(1)} \dots A^{(n-1)} (A^{(n)})^{m+1} \bar{1} (1 + o(1)).$$

Теперь осталось оценить разложения в окрестности 0 характеристических функций времен переходов. Предположим, что величины τ_k притягиваются к некоторым устойчивым законам, т. е.

$$M \exp \{-\theta \tau_k\} = 1 - c_k \theta^{\alpha} + o(\theta^{\alpha}), \quad k = \overline{1, r}, \quad (5.54) \\ (0 < \alpha \leq 1)$$

(в данном случае удобнее пользоваться преобразованиями Лапласа случайных величин для проведения соответствующих оценок, здесь $\theta \geq 0$). Тогда положим $\beta_e = \gamma_e^{\frac{m+n}{\alpha}}$. Из (5.54) следует, что

$$M \exp \{-\beta_e \theta \tau_k\} = 1 - \gamma_e^{n+m} c_k \theta^\alpha + o(\gamma_e^{n+m}). \quad (5.55)$$

Получим подобные разложения для времен переходов $\tau_e((i, s), (j, q))$. По определению, если $\xi_1 = i$, $q_e(1) = s$, то

$$\begin{aligned} P\{\tau_e((i, s), (i, q)) < x\} &= P\{\tau_i < x/\xi_2 = j, \\ A_e(s, q-1)\} &= P\{\tau_i < x/A_e(s, q-1)\}. \end{aligned}$$

поскольку в нашей конструкции времена пребывания ПМП в состояниях $\xi(t)$ не зависят от последующего перехода (это ограничение несущественно и просто упрощает форму записи). Здесь $A_e(s, q-1)$ — событие, которое означает, что за время τ_i количество требований в системе стало равным $q-1$.

Рассмотрим вначале случай $s=1$. Через η_e обозначим показательно распределенную величину с параметром μ/e . Тогда легко видеть, что если $\eta_e < \tau_i$, то поступившее требование успеет обслужиться за время τ_i , в противном случае в момент поступления очередного требования в системе будет находиться два требования. Поэтому

$$\begin{aligned} P\{\tau_e((i, 1), (j, 1)) < x\} &= P\{\tau_i < x/\eta_e < \tau_i\}; \\ P\{\tau_e((i, 1), (j, 2)) < x\} &= P\{\tau_i < x/\eta_e \geq \tau_i\}. \end{aligned}$$

Покажем теперь, что

$$|1 - M \exp \{-\beta_e \theta \tau_e((i, 1), (j, 2))\}| = O(\gamma_e^{n+m}). \quad (5.56)$$

Действительно,

$$\begin{aligned} 0 \leq 1 - M[e^{-\theta \tau_i/\eta_e} \geq \tau_i] &= P\{\eta_e \geq \tau_i\}^{-1} (P\{\eta_e \geq \tau_i\} - \\ &- M e^{-\theta \tau_i} \chi(\eta_e \geq \tau_i)) = P\{\eta_e \geq \tau_i\}^{-1} (M(1 - \\ &- e^{-\theta \tau_i}) \chi(\eta_e \geq \tau_i)) = P\{\eta_e \geq \tau_i\}^{-1} \int_0^\infty (1 - e^{-\theta x}) \times \\ &\times P\{\eta_e \geq x\} dG_i(x) \leq P\{\eta_e \geq \tau_i\}^{-1} \int_0^\infty (1 - e^{-\theta x}) \times \\ &\times dG_i(x) \int_0^\infty P\{\eta_e \geq x\} dG_i(x) = 1 - M e^{-\theta \tau_i}. \end{aligned}$$

Здесь мы воспользовались неравенством Чебышева (5.10), поскольку $1 - e^{-\theta x}$ монотонно возрастает на $[0, \infty)$, а $P\{\eta_e \geq x\}$ монотонно убывает на $[0, \infty)$.

Таким образом.

$$0 \leq 1 - M[e^{-\theta \tau_i} / \eta_e \geq \tau_i] \leq 1 - M e^{-\theta \tau_i}.$$

Подставляя вместо θ величину $\beta_e \theta$ и учитывая (5.55), легко получим (5.56).

Теперь по формуле полной вероятности

$$M e^{-\theta \tau_i} = M[e^{-\theta \tau_i} / \eta_e < \tau_i] P\{\eta_e < \tau_i\} + \\ + M[e^{-\theta \tau_i} / \eta_e \geq \tau_i] P\{\eta_e \geq \tau_i\}.$$

Поскольку $P\{\eta_e \geq \tau_i\} = v_e^{(i)}(1) \rightarrow 0$, и выполнено (5.55), (5.56), то отсюда нетрудно получить, что

$$M[e^{-\beta_e \theta \tau_i} / \eta_e < \tau_i] = 1 - \gamma_e^{n+m} c_i \theta^\alpha + o(\gamma_e^{m+n}).$$

Тем самым

$$M \exp\{-\beta_e \theta \tau_e((i, 1), (j, 1))\} = 1 - \gamma_e^{n+m} c_i \theta^\alpha + \\ + o(\gamma_e^{m+n}), \quad i, j = \overline{1, r}. \quad (5.57)$$

Теперь из соотношения

$$M e^{-\theta \tau_i} = \sum_{\substack{j=\overline{1, r} \\ q=\overline{0, s+1}}} M \exp\{-\theta \tau_e((i, s), \\ (j, q))\} p_e((i, s), (j, q)).$$

подобным образом можно получить

$$p_e((i, s), (j, q)) (1 - M \exp\{-\beta_e \theta \tau_e((i, s), \\ (j, q))\}) = O(\gamma_e^{n+m}),$$

для всех $i, j = \overline{1, r}$, $0 \leq q \leq s+1$. Поскольку $p_e(i, s) \rightarrow 0$ при $s > 1$, то, воспользовавшись замечанием к теореме 5.4 и следствием 5.6, получим следующее утверждение: (5.2). Для системы $SM|M|n|m$ в предположениях (5.52) — (5.54) для любых $k = \overline{1, r}$, $q \leq n+m$

$$M \exp\{-\theta \gamma_e^{\frac{n+m}{\alpha}} \Omega_e(k, q)\} \rightarrow (1 + M \theta^\alpha)^{-1},$$

где

$$M = \left(\sum_{i=1}^r \pi_i c_i (\bar{\pi} A^{(1)} \dots A^{(n-1)} (A^{(n)})^{m+1} \bar{1})^{-1} \right).$$

В частности, для системы $GI|M|1|m$ с рекурентным входящим потоком и одним прибором $M = cb^{-(m+1)}$.

П. 3. Исследуем поведение потока потерянных требований. Пусть $\hat{v}_e(t)$ количество требований, получивших отказ на промежутке $[0, t]$. Поскольку для обеих рассмотренных здесь систем суммарная вероятность (для вложенной цепи) перейти с уровня s на уровень $q < s$ при $s \geq 1$ (для системы $M_n|M|n|m|$ и при $s \geq 2$ (для системы $G_n|M|n|m|$) стремится к единице, то отсюда следует, что с вероятностью, близкой к единице, требование, следующее за получившим отказ, поступит в пустую систему. После этого следующее требование получит отказ через время $\Omega_e(i, 0)$ (i — состояние управляющей цепи в момент поступления 1-го требования). Так как предельное распределение не зависит от i , то отсюда следует асимптотическая независимость нормированных промежутков времени между последовательными моментами отказов требований (строгий формальный вывод этого утверждения предоставляем читателю).

Отметим теперь следующий общий результат, который удобно использовать при исследовании асимптотического поведения редких потоков.

5.3. Пусть задан поток требований $v_e(t)$, у которого $0 \leq \mu_e(0) \leq \mu_e(1) \leq \dots$ — последовательные моменты поступления требований, и существует нормирующий множитель β_e такой, что случайные величины $\beta_e(\mu_e(k+1) - \mu_e(k))$, $k \geq 0$, асимптотически независимы и для любого k

$$P\{\beta_e(\mu_e(k+1) - \mu_e(k)) < x\} \xrightarrow{\text{сл}} F(x).$$

Тогда конечномерные распределения потока $v_e(\beta_e^{-1}t)$ слабо сходятся на любом промежутке времени $[0, T]$ к потоку восстановления $v_0(t)$ с распределением промежутка восстановления $F(x)$, где T — точка непрерывности $F^{(m)*}(x)$ при любом $m \geq 1$.

Действительно, из наших условий непосредственно следует, что для любого $k \geq 1$

$$\beta_e \mu_e(k) = \beta_e \sum_{i=1}^k (\mu_e(i) - \mu_e(i-1)) \xrightarrow{\text{сл}} \sum_{i=1}^k \tau_i,$$

где $\tau_1, \tau_2, \dots$ — последовательность независимых одинаково распределенных величин с функцией распределе-

ния $F(x)$. Следовательно, в силу тождества Колмогорова—Феллера

$$\begin{aligned} P\{\nu_e(\beta_e^{-1}t) \leq k\} &= P\{\mu_e(k) \geq \beta_e^{-1}t\} = \\ &= P\{\beta_e \mu_e(k) \geq t\} \xrightarrow{\text{сл}} P\left\{\sum_{i=1}^k \tau_i \geq t\right\} = P\{\nu_0(t) \leq k\}. \end{aligned}$$

Аналогично можно показать сходимость совместных распределений.

Заметим, что если $F(+0) = 0$, то предельный поток будет ординарным.

Из этого утверждения применительно к рассмотренным системам при указанных выше предположениях вытекают следующие результаты:

1) поток отказов системы $M_u | M | n | m$ $\nu_e(e^{-(n+m)}t)$ сходится к пуассоновскому потоку с параметром Λ .

2) поток отказов системы $SM | M | n | m$ $\nu_e(\gamma_e^{-(n+m)}t)$ сходится к потоку восстановления с преобразованием Лапласа промежутка времени между восстановлениями

$$(1 + M\theta^\alpha)^{-1}. \quad (5.58)$$

В заключение отметим, что рассмотренные надежность характеристики системы $SM | M | n | m$ не инвариантны относительно средних значений характеристик входящего потока и обслуживания, т. е., рассматривая системы с одинаковыми средними времен управляющего ПМП $\xi(t)$, мы можем получить совершенно различные (на одинаковых промежутках) потоки отказов (в отличие от того, как, к примеру, стационарное распределение ПМП при заданной вложенной цепи инвариантно относительно средних времен сидения в состояниях).

Действительно, поток отказов определяется коэффициентом нормировки времени γ_e и параметром M (5.58). Величина M определяется поведением функций $G_k(x)$ на бесконечности и в случае существования средних $G_k(x)$ инвариантна относительно их. Порядок же γ_e определяется поведением $G_k(x)$ в окрестности нуля. К примеру, если $G_k(x) \sim x^\alpha$ при $x \rightarrow 0$, то по формуле (5.53) нетрудно вычислить, что $\gamma_e \sim c e^\alpha$. Тем самым легко построить примеры трех систем с одинаковыми средними $G_k(x)$, таких, что на промежутке $[0, t/\epsilon]$ поток отказов одной системы сходится к пуассоновскому с конечным параметром $0 < \Lambda < \infty$, а для двух других систем величины $\nu_e(t/\epsilon)$ стремятся по вероятности к нулю и бесконечности соответственно.

5.5. Редующие потоки

В п. 1.3 было показано, что при определенных условиях суммы большого числа независимых равномерно малых потоков восстановления сходятся к пуассоновскому потоку. Вместе с тем во многих прикладных задачах возникают модели потоков, которые получаются из исходного путем многократного прореживания требований. Точнее, пусть задан поток восстановления $v(t)$, который поступает на некоторое устройство. Это устройство сохраняет каждое требование в потоке с вероятностью p либо выбрасывает с вероятностью $q = 1 - p$. На выходе получается прореженный (или редующий) поток $v^{(p)}(t)$. Пусть исходный поток задается промежутками длительностей между требованиями $\tau_1, \tau_2, \dots$ с функцией распределения $F(x)$. Положим

$$\varphi(\theta) = \int_0^{\infty} e^{-\theta x} dF(x), \quad (\theta \geq 0).$$

Пусть $\tau_1^{(p)}, \tau_2^{(p)}, \dots$ — промежутки длительностей между требованиями потока $v^{(p)}(t)$. Нетрудно видеть, что $\tau_1^{(p)}, \tau_2^{(p)}, \dots$ — независимые одинаково распределенные величины, при этом $\tau_1^{(p)} = \sum_{i=1}^k \tau_i$, т. е.

где k — величина, распределенная по геометрическому закону с параметром p (т. е. $P\{k = k\} = q^{k-1} p, k \geq 1$). Отсюда следует, что

$$\varphi^{(p)}(\theta) = Me^{-\theta \tau_1^{(p)}} = \sum_{k=1}^{\infty} \varphi(\theta)^k \cdot q^{k-1} p = \frac{p \varphi(\theta)}{1 - q \varphi(\theta)}. \quad (5.59)$$

С помощью несложных выкладок можно самостоятельно убедиться в том, что прореживание исходного потока с помощью последовательно соединенных устройств с вероятностями $p_1, p_2, \dots, p_n$ эквивалентно прореживанию одним устройством с вероятностью $p_1 \dots p_n$.

Подобные модели многократного прореживания могут возникать, к примеру, в задачах последовательного контроля изделий, когда отбраковываются неисправные изделия, в задачах последовательного выявления ошибок в текстах, неточностей в записях программ и т. п.

В случае, когда $n \rightarrow \infty$ и $p_1 \dots p_n \rightarrow 0$ возникает естественная проблема исследования класса предельных потоков для $v^{p_1 \dots p_n}(t)$ при соответствующем изменении масштаба времени.

Данная проблема в сущности эквивалентна следующей: пусть требование сохраняется в потоке с вероятностью

$p_e \rightarrow 0$. Исследовать класс предельных потоков для потоков вида $v^{(p_e)}(\beta_e^{-1}t)$, где $\beta_e \rightarrow 0$.

Подобные схемы возникают также при исследовании предельного поведения потока потерянных требований в п. 5.4, поскольку такие потоки также можно рассматривать как прореживание (с помощью СМО) исходного потока.

Исследуем вначале схему прореживания потока восстановления. Пусть у нас есть последовательность простых потоков восстановления $v_e(t)$. Поток $v_e(t)$ задается промежутком времени между двумя последовательными поступлениями τ_e . Имеет место

Теорема 5.5. Если в схеме прореживания потока $v_e(t)$ каждое требование остается с вероятностью $p_e \rightarrow 0$ и существует нормирующий множитель β_e такой, что

$$\varphi_e(\beta_e \theta) = M \exp \{-\beta_e \theta \tau_e\} = 1 - p_e a(\theta) + o(p_e), \quad (5.60)$$

то поток $v_e^{(p_e)}(\beta_e^{-1}t)$ сходится к потоку восстановления с преобразованием Лапласа промежутка между поступлениями $(1 + a(\theta))^{-1}$.

Следствие 5.7 (теорема Реньи). Если $v(t)$ — простой поток восстановления, для которого $\lambda^{-1} = \int_0^\infty x dF(x)$ и $0 < \lambda < \infty$, то поток $v^{(p)}(p^{-1}t)$ сходится при $p \rightarrow 0$ к простейшему потоку с параметром λ .

Доказательство. Так как $v_e^{(p_e)}(\beta_e^{-1}t)$ — поток восстановления, то в силу утверждения 5.3 в п. 5.4, нам достаточно показать слабую сходимость распределений промежутков между требованиями, что соответствует сходимости преобразований Лапласа:

$$M \exp \{-\beta_e \theta \tau_e^{(p_e)}\} \rightarrow (1 + a(\theta))^{-1}. \quad (5.61)$$

Воспользовавшись (5.59) и (5.60), получим

$$\begin{aligned} M \exp \{-\beta_e \theta \tau_e^{(p_e)}\} &= \frac{p_e \varphi_e(\beta_e \theta)}{1 - (1 - p_e) \varphi_e(\beta_e \theta)} = \\ &= \frac{p_e (1 - p_e a(\theta) + o(p_e))}{1 - (1 - p_e) (1 - p_e a(\theta) + o(p_e))} \rightarrow (1 + a(\theta))^{-1}. \end{aligned}$$

В условиях следствия, если $0 < \lambda < \infty$, то $\varphi(\theta) = 1 - \frac{1}{\lambda} \theta + o(\theta)$, откуда следует, что $a(\theta) = \frac{1}{\lambda} \theta$, а выражение $\frac{1}{1 + \frac{1}{\lambda} \theta} = \frac{\lambda}{\lambda + \theta}$ соответствует преобразованию Лапласа

показательного распределения с параметром λ . Заметим, что результат теоремы 5.5 сохраняется также для потоков с запаздыванием. В этом случае надо потребовать дополнительно, чтобы $\beta_e \tau_e^{(1)} \xrightarrow{P} 0$, где $\tau_e^{(1)}$ — время запаздывания.

Доказательство этого несложного факта предоставляем читателю.

Рассмотрим теперь более общие схемы прореживания потоков, которые, как мы видели, могут возникать при изучении потоков, потерянных требований в системах, задачах анализа надежности сложных стохастических систем и т. п.

Пусть задан полумарковский входящий поток $v(t)$, т. е. есть ПМП $\xi(t)$ с конечным множеством состояний $\{1, \dots, \dots, r\}$, который задается вложенной цепью с матрицей переходных вероятностей $P = \|p_{ij}\|$, $i, j = \overline{1, r}$ и набором времен переходов $\{\tau_{ij}, i, j = \overline{1, r}\}$, а $v(t)$ эквивалентно числу скачков $\xi(t)$ на промежутке $[0, t]$, $t > 0$. Задано семейство вероятностей $\bar{p}_e = \{p_e(1), \dots, p_e(r)\}$, $0 \leq p_e \times (i) \leq 1$, которые определяют следующую схему прореживания: пусть $0 = t_0 \leq t_1 \leq t_2 \leq \dots$ — последовательные моменты скачков $\xi(t)$ $x_k = \xi(t_k)$ — вложенная цепь. Тогда если $x_k = i$, то требование, поступившее в момент t_{k+1} , сохраняется с вероятностью $p_e(i)$ и выбрасывается с вероятностью $1 - p_e(i)$. Смысл этого заключается в следующем: положим $F_i(x) = \sum_{j=1}^r p_{ij} P\{\tau_{ij} < x\}$. Тогда если требование поступает через промежуток времени с функцией распределения $F_i(x)$, то оно сохраняется с вероятностью $p_e(i)$, т. е. вероятности прореживания зависят от характеристик временного интервала, через который поступает требование, что естественно может возникать в прикладных задачах.

Теорема 5.6. Если вложенная цепь неприводима, $p_e(k) = eb_k$,

$$0 \leq b_k < \infty, \quad k = \overline{1, r}, \quad \max_{k=1, r} b_k > 0$$

и

$$M \exp\{-\theta \tau_{ij}\} = 1 - a_{ij} \theta^\alpha + o(\theta^\alpha), \quad i, j = \overline{1, r}, \quad (5.62) \\ (0 < \alpha \leq 1),$$

то редущий поток $v^{(\bar{p}_e)}(e^{-\frac{1}{\alpha}t})$ сходится к потоку восстановления с преобразованием Лапласа промежутка между поступлениями $b(b + A\theta^\alpha)^{-1}$, где $A = \sum_{k=1}^r \pi_k p_{kj} a_{kj}$; $b = \sum_{k=1}^r \pi_k b_k$, а π_k , $k = \overline{1, r}$ — стационарное распределение для вложенной цепи.

В частности если $\alpha = 1$, поток будет простейшим с параметром $b \cdot A^{-1}$.

Замечание. В общем случае, если

$$M \exp \{-\theta \tau_{ij}\} = 1 - a_{ij} \theta^{\alpha} h(\theta) + o(\theta^{\alpha} h(\theta)),$$

где $h(\theta)$ — медленно меняющаяся функция, нормирующий множитель β_{ε} следует выбрать так, чтобы $\beta_{\varepsilon}^{\alpha} h(\beta_{\varepsilon}) \sim \varepsilon$.

Доказательство. Построим вспомогательный ПМП $\kappa_{\varepsilon}(t)$ с множеством состояний $\{(i, s), i = \overline{1, r}, s = 0, 1\}$. При этом распределения времен переходов $\tau((i, s), (j, k))$ задаются следующим образом: $P\{\tau((i, s), (j, k)) < x\} = F_{ij}(x)$, $i, j = \overline{1, r}, s, k = 0, 1$, а вероятности переходов вложенной цепи имеют вид:

$$\begin{aligned} p_{\varepsilon}((i, s), (j, 0)) &= (1 - p_{\varepsilon}(i)) p_{ij}; \\ p_{\varepsilon}((i, s), (j, 1)) &= p_{\varepsilon}(i) p_{ij}, \quad i, j = \overline{1, r}, s = 0, 1. \end{aligned}$$

При этом легко понять, что моменты попадания $\kappa_{\varepsilon}(t)$ в подмножество $\{(i, 1), i = \overline{1, r}\}$ представляют собой моменты появления требований в редеемшем потоке $v^{(\bar{p}_{\varepsilon})}(t)$. Поскольку $p_{\varepsilon}(i) = \varepsilon b_i$ и $p_{\varepsilon}((i, 0), (j, 0)) \rightarrow p_{ij}$, $i, j = \overline{1, r}$, то мы находимся в условиях теоремы 5.3. Пусть $\kappa_{\varepsilon}(0) = (i, 0)$. Обозначим $\Omega_{\varepsilon}(i) = \inf\{t : \kappa_{\varepsilon}(t) \in \{(j, 1), j = \overline{1, r}\}\}$. Тогда при наших предположениях из теоремы 5.3 прямо следует

$$\lim_{\varepsilon \rightarrow 0} M \exp \left\{ -\varepsilon^{\frac{1}{\alpha}} \theta \Omega_{\varepsilon}(i) \right\} = (1 + A \theta^{\alpha})^{-1}, \quad i = \overline{1, r}.$$

Теперь при попадании в подмножество $\{(j, 1), j = \overline{1, r}\}$ процесс с вероятностью, стремящейся к единице, в момент следующего скачка вернется в подмножество $\{(i, 0), i = \overline{1, r}\}$. Так как предельное распределение $\Omega_{\varepsilon}(i)$ не зависит от i , то, используя утверждение 5.3 в п. 5.4, нетрудно получить утверждение теоремы.

Воспользовавшись следствием 5.4, можно получить соответствующие результаты для схем редения потоков требований разных типов. К примеру в нашей модели будем считать, что если $x_k = i$, то требование, поступившее в момент t_{k+1} , имеет тип « i ». Тогда при наших предположениях редееющий поток требований сходится к указанному потоку восстановления, у которого каждое появляющееся требование независимо от других с вероятностью $\pi_i b_i b^{-1}$ будет иметь тип « i ».

В частности, если предельный поток пуассоновский ($\alpha = 1$), это означает, что его можно представить в виде суммы r независимых пуассоновских потоков, при этом i -й поток соответствует потоку требований типа « i » и имеет параметр $\pi_i b_i A^{-1}$.

Применяя результаты п. 5.3, несложно получить соответствующие результаты в том случае, когда исходный поток $v(t) = v_\varepsilon(t)$ задан на ПМП $\xi_\varepsilon(t)$, зависящем от параметра ε , при этом необходимо потребовать, чтобы подмножество $\{(i, 0), i = \overline{1, r}\}$ с вероятностями переходов $p_\varepsilon((i, 0), (j, 0))$ образовывало s -множество.

Данные результаты показывают, что потоки восстановления и, в частности, пуассоновские потоки при достаточно общих предположениях могут выступать в качестве хорошего приближения для потоков редких событий, которые строятся на процессах, обладающих точками восстановления.

Заметим, что полученные результаты сохраняются и в том случае, когда исходный поток управляется некоторым случайным процессом, числа попаданий в состояния которого удовлетворяют закону больших чисел. В противном случае (если не выполняется закон больших чисел) классы предельных потоков будут значительно шире и уже не являются потоками восстановления.

Задания для самоконтроля

1. Пусть $\Pi(t)$, $t \geq 0$ — пуассоновский поток с параметром λ , $\Pi^{(p)}(t)$ поток, полученный прореживанием $\Pi(t)$ (каждое требование независимо от других сохраняется с вероятностью p). Показать, что $\Pi^{(p)}(t)$ — пуассоновский поток с параметром $p\lambda$.

2. Пусть $v(t)$, $t \geq 0$ — простой поток восстановления с интенсивностью $\mu = \left(\int_0^\infty x dF(x) \right)^{-1}$. Показать, что поток $v^{(p)}(p^{-1}t)$ также имеет интенсивность μ .

3. Пусть $\Pi(t)$, $t \geq 0$ — пуассоновский поток с параметром λ , при этом каждое требование с вероятностью α_i будет i -го типа, $\sum \alpha_i = 1$. Показать, что $\Pi(t)$ можно представить в виде $\Pi(t) = \sum_i \Pi_i(t)$, где $\Pi_i(t)$ — независимые пуассоновские потоки с параметрами $\alpha_i \lambda$ ($\Pi_i(t)$ — поток требований i -го типа).

Для потоков восстановления в общем случае это неверно.

5.6. Анализ потоков неоднородных редких событий

В предыдущих параграфах данной главы описывались модели, связанные с моментами наступления редких событий на простых либо многомерных (типа ПМП) процессах восстановления. При этом общий метод исследования заключался в построении вспомогательного ПМП и исследовании момента первого выхода из подмножества. В том случае, когда параметры потока либо системы меняются со временем, т. е. схема неоднородна, указанные методы не годятся, поскольку в случае произвольной зависимости параметров от времени соответствующие процессы не обладают моментами восстановления.

Ряд результатов по асимптотическому анализу неоднородных потоков, порожденных системами, у которых есть одно восстанавливающее состояние, а времена сидения в других состояниях пренебрежимо малы по сравнению с временем сидения в этом состоянии, получено в (9). Некоторые результаты по исследованию потоков, порождаемых неоднородными по времени системами с многими точками восстановления, а также без них, приведены в (3). В данном параграфе в 1-й части исследуется модель, которая является распространением схемы первого осуществления редкого события п. 5.1 на неоднородный случай. Во 2-й части исследуется поведение потока редких событий, порождаемого неоднородной системой с произвольным пространством состоянием. Рассматриваются примеры асимптотического анализа неоднородных СМО. Здесь предлагается единый подход к данной проблематике, связанный с исследованием потоков редких индикаторов, заданных на некоторой случайной последовательности (появление единицы трактуется как наступление редкого события). Этот подход развивался в работах (2, 3).

п. 1. Докажем вначале теорему, которая является обобщением теоремы Пуассона (п. 1.3) на случай слабо зависящих редких событий. Пусть задана некоторая дискретная последовательность $x_e(k)$, $k \geq 1$ со значениями в произвольном измеримом пространстве $(X, \mathcal{B}_X)$ и независимые в совокупности семейства случайных индикаторов $\{\chi_e(k, x), x \in X\}$, $k \geq 1$, т. е. при фиксированных $x_1, x_2, \dots$ величины $\chi_e(1, x_1), \chi_e(2, x_2), \dots$ независимы в совокупности (величины $\chi_e(k, x)$ могут принимать только одно из двух значений 0 или 1). Наблюдается последовательность $\chi_e(k) = \chi_e(k, x_e(k))$, $k \geq 1$. Обозначим

$$y_e(m) = \sum_{k=1}^m \chi_e(k), \quad m \geq 1.$$

$y_\varepsilon(m)$ представляет дискретный поток событий, управляемых последовательностью $x_\varepsilon(k)$. Положим

$$p_\varepsilon(k, x) = P\{\chi_\varepsilon(k, x) = 1\}, \quad x \in X,$$

и

$$\rho_\varepsilon(k) = p_\varepsilon(k, x_\varepsilon(k)), \quad k \geq 1.$$

Теорема 5.7. Если при $\varepsilon \rightarrow 0$

$$\gamma_\varepsilon(t) = \sum_{k=1}^{[\varepsilon^{-1}t]} \rho_\varepsilon(k) \xrightarrow{\text{ол}} \gamma_0(t), \quad t \in [0, T], \quad (5.63)$$

где $\gamma_0(t)$ — некоторый неубывающий процесс, и

$$\sum_{k=1}^{[\varepsilon^{-1}t]} \rho_\varepsilon^2(k) \xrightarrow{P} 0, \quad (5.64)$$

то

$$y_\varepsilon([\varepsilon^{-1}t]) \xrightarrow{\text{ол}} \prod (\gamma_0(t)), \quad t \in [0, T],$$

где $\prod(x)$ — пуассоновский поток с параметром единица и $\prod(\cdot)$ и $\gamma_0(\cdot)$ — независимы, тем самым

$$P\{y_\varepsilon([\varepsilon^{-1}t]) = k\} \rightarrow M e^{-\gamma_0(t)} \frac{\gamma_0(t)^k}{k!}, \quad k \geq 0.$$

Доказательство. Запишем следующее представление:

$$\begin{aligned} M \exp\left\{-\Theta \sum_{k=1}^{[\varepsilon^{-1}t]} \chi_\varepsilon(k)\right\} &= M \prod_{k=1}^{[\varepsilon^{-1}t]} M[e^{-\Theta \chi_\varepsilon(k)} | x_\varepsilon(k)] = \\ &= M \prod_{k=1}^{[\varepsilon^{-1}t]} (1 - \rho_\varepsilon(k) (1 - e^{-\Theta})). \end{aligned}$$

Заметим, что для любых чисел α_k, β_k таких, что $|\alpha_k| \leq 1, |\beta_k| \leq 1, k = \overline{1, m}$ имеет место неравенство

$$\left| \prod_{k=1}^m \alpha_k - \prod_{k=1}^m \beta_k \right| \leq \sum_{k=1}^m |\alpha_k - \beta_k|.$$

В самом деле

$$\begin{aligned} \left| \prod_{k=1}^m \alpha_k - \prod_{k=1}^m \beta_k \right| &\leq \sum_{k=1}^m \left| \prod_{k=1}^j \prod_{i=j+1}^m \alpha_k \beta_k - \prod_{k=1}^{j-1} \prod_{i=j}^m \alpha_k \beta_i \right| \leq \\ &\leq \sum_{j=1}^m \prod_{k=1}^{j-1} \prod_{i=j+1}^m |\alpha_k| |\beta_i| |\alpha_j - \beta_j| \leq \sum_{j=1}^m |\alpha_j - \beta_j|. \end{aligned}$$

Пользуясь этим неравенством и неравенством $0 \leq e^{-x} - 1 + x \leq \frac{1}{2} x^2$ при $x \geq 0$, получим для любого $L > 0$

$$\begin{aligned} & |M \exp \{-\theta y_{\varepsilon} ([\varepsilon^{-1} t])\} - M \exp \{-\gamma_{\varepsilon}(t) (1 - e^{-\theta})\}| \leq \\ & \leq 2P \left\{ \sum_{k=1}^{[\varepsilon^{-1} t]} \rho_{\varepsilon}^2(k) > L \right\} + \sum_{k=1}^{[\varepsilon^{-1} t]} M |1 - \rho_{\varepsilon}(k) (1 - e^{-\theta}) - \\ & - e^{-\rho_{\varepsilon}(k)(1 - e^{-\theta})}| \chi \left(\sum_{k=1}^{[\varepsilon^{-1} t]} \rho_{\varepsilon}^2(k) \leq \alpha \right) \leq \\ & \leq 2P \left\{ \sum_{k=1}^{[\varepsilon^{-1} t]} \rho_{\varepsilon}^2(k) > L \right\} + \frac{1}{2} (1 - e^{-\theta})^2 \times \\ & \times M \sum_{k=1}^{[\varepsilon^{-1} t]} \rho_{\varepsilon}^2(k) \chi \left(\sum_{k=1}^{[\varepsilon^{-1} t]} \rho_{\varepsilon}^2(k) \leq L \right) \rightarrow 0 \quad (5.65) \end{aligned}$$

по теореме Хелли в силу (5.64). Теперь, поскольку функция $\{-x(1 - e^{-\theta})\}$ ограничена и непрерывна по x в области $x \geq 0$, то в силу (5.63)

$$M \exp \{-\gamma_{\varepsilon}(t) (1 - e^{-\theta})\} \rightarrow M \exp \{-\gamma_0(t) (1 - e^{-\theta})\}$$

и окончательно

$$M \exp \{-\theta y_{\varepsilon} ([\varepsilon^{-1} t])\} \rightarrow M \exp \{-\gamma_0(t) (1 - e^{-\theta})\},$$

что соответствует слабой сходимости одномерных распределений $y_{\varepsilon}(t)$ к $\prod (\gamma_0(t))$.

Тем же методом нетрудно показать, что для любых $\theta_1, \dots, \theta_m, t_1, \dots, t_m$.

$$\begin{aligned} & M \exp \left\{ - \sum_{i=1}^m \theta_i y_{\varepsilon} ([\varepsilon^{-1} t_i]) \right\} = M \exp \left\{ - y_{\varepsilon} ([\varepsilon^{-1} t_1]) \times \right. \\ & \times \sum_{i=1}^m \theta_i - \sum_{j=1}^{m-1} (y_{\varepsilon} ([\varepsilon^{-1} t_{j+1}]) - y_{\varepsilon} ([\varepsilon^{-1} t_j])) \cdot \sum_{i=j+1}^m \theta_i \left. \right\} \rightarrow \\ & \rightarrow M \exp \left\{ - \gamma_0(t_1) (1 - \exp \left\{ - \sum_{i=1}^m \theta_i \right\}) - \sum_{j=1}^{m-1} (\gamma_0(t_{j+1}) - \right. \\ & \left. - \gamma_0(t_j)) \cdot (1 - \exp \left\{ - \sum_{i=j+1}^m \theta_i \right\}) \right\}, \end{aligned}$$

что соответствует сходимости конечномерных распределений $y_{\varepsilon} ([\varepsilon^{-1} t])$ к $\prod (\gamma_0(t))$. Теорема доказана.

Оценки типа (3.65) позволяют, используя неравенство Эссеена, получать явные оценки скорости сходимости распределений $y_{\varepsilon} ([\varepsilon^{-1} t])$ к распределениям процесса пуассоновского типа $\prod (\gamma_{\varepsilon}(t))$. Действительно, воспользовав-

шись вариантом неравенства Эссена, если функции распределения F и G ступенчатые со скачками в целых точках, а

$$\varphi(u) = \int_{-\infty}^{\infty} e^{iux} dF(x), \quad g(u) = \int_{-\infty}^{\infty} e^{iux} dG(x),$$

то

$$\sup_x |F(x) - G(x)| \leq \frac{1}{4} \left| \frac{\varphi(u) - g(u)}{u} \right| du,$$

читатель может без труда получить следующее:

Следствие 5.8. Если выполнено условие (5.63) и $\sum_{k=1}^{[e^{-1}t]} M\rho_s^2(k) \rightarrow 0$, то $\sup_{k \geq 0} \left| P\{y_s([e^{-1}t]) = k\} - Me^{-\gamma_s(t)} \times \right.$
 $\left. \times \frac{\gamma_s(t)^k}{k!} \right| \leq \frac{1}{4} (1 + 2 \ln \pi) \cdot \sum_{k=1}^{[e^{-1}t]} M\rho_s^2(k) \rightarrow 0$.

Заметим, в частности, что если $\chi_s(k)$ — независимы, то $\rho_s(k) = M\chi_s(k) = P\{\chi_s(k) = 1\}$, и в этом случае если

$$\sum_{k=1}^{[e^{-1}t]} P\{\chi_s(k) = 1\} \rightarrow \Lambda(t)$$

и $\max_{n \leq e^{-1}t} P\{\chi_s(k) = 1\} \rightarrow 0$, то выполнены условия (5.63),

(5.64) и $y_s([e^{-1}t]) \xrightarrow{cn} \prod (\Lambda(t))$, $t \leq 0$, что соответствует сходимости к пуассоновскому потоку с ведущей функцией $\Lambda(t)$.

Рассмотрим теперь схему построения потока событий с помощью сумм индикаторов редких событий. Пусть в дополнение к последовательности $x_s(k)$ и семейству индикаторов $\{\chi_s(k, x), x \in X\}$, $k \geq 1$, заданы независимые от них и в совокупности семейства неотрицательных случайных величин $\{\tau_s(k, x), x \in X\}$, $k \geq 1$. Рассмотрим сложный поток $v_s(t)$, моменты скачков которого есть величины

$$\mu_s(m) = \sum_{k=1}^m \tau_s(k, x_s(k)), \quad m \geq 1 \quad (5.66)$$

(поток $v_s(t)$ можно рассматривать как неоднородный поток восстановления, который управляется последовательностью $x_s(k)$). Заметим, в частности, что если $x_s(k)$ — вырождена ($x_s(k) = x$ с вероятностью единица для всех $k \geq 1$, а распределения $\tau_s(k, x)$ не зависят от k , то $v_s(t)$ — обычный поток восстановления. Если $x_s(k)$ — дискретная цепь Маркова, $v_s(t)$ — поток полумарковского типа. Если теперь

$\chi_\varepsilon(m, x_\varepsilon(m)) = 1$, то будем говорить что в момент $\mu_\varepsilon(m)$ произошло редкое событие (предполагается, что

$$P\{\chi_\varepsilon(m, x_\varepsilon(m)) = 1\} \rightarrow 0.$$

Будем исследовать поток редких событий $\hat{v}_\varepsilon(t)$ при соответствующем растяжении оси времени в предположении, что осуществление или неосуществление события не влияет на поведение последовательности $x_\varepsilon(k)$.

Положим

$$y_\varepsilon(m) = \sum_{k=1}^m \chi_\varepsilon(k, x_\varepsilon(k)), \quad m \geq 1$$

и определим последовательности

$$\begin{aligned} \kappa_\varepsilon(l) &= \min\{m: y_\varepsilon(m) = l\}; \\ \zeta_\varepsilon(l) &= \mu_\varepsilon(\kappa_\varepsilon(l)), \quad l \geq 1. \end{aligned}$$

Легко видеть, что по построению $\zeta_\varepsilon(l)$ — момент l -го осуществления редкого события и нам достаточно установить сходимости, при соответствующей нормировке, величин $\zeta_\varepsilon(l)$.

Обозначим

$$p_\varepsilon(k, x) = P\{\chi_\varepsilon(k, x) = 1\}, \quad k \geq 1, x \in X;$$

$$\Lambda_\varepsilon(m) = \sum_{k=1}^m p_\varepsilon(k, x_\varepsilon(k)), \quad m \geq 1.$$

Теорема 5.8. Если при $\varepsilon \rightarrow 0$

$$\beta_\varepsilon \mu_\varepsilon([\varepsilon^{-1} t]) \xrightarrow{\text{сл}} \tau_0(t), \quad t \geq 0, \quad (5.67)$$

где $\tau_0(t)$ — некоторый стохастически непрерывный процесс

$$\Lambda_\varepsilon([\varepsilon^{-1} t]) \xrightarrow{P} \Lambda(t), \quad t \geq 0, \quad (5.68)$$

где $\Lambda(t)$ — непрерывная неслучайная функция, и

$$\sum_{k=1}^{[\varepsilon^{-1} t]} p_\varepsilon^2(k, x_\varepsilon(k)) \xrightarrow{P} 0, \quad t > 0, \quad (5.69)$$

то поток $\hat{v}_\varepsilon(\beta_\varepsilon^{-1} t)$ слабо сходится к потоку $\hat{v}_0(t)$, последовательные моменты скачков которого есть величины $\tau_0(\kappa_l)$, $l \geq 1$, где κ_l — момент l -го скачка, независимого от $\tau_0(u)$ пуассоновского потока $\prod(\Lambda(t))$ (потока с ведущей функцией $\Lambda(t)$).

Замечание. Если $\tau_0(t)$ — однородный процесс с независимыми приращениями вида $M e^{-\theta \tau_0(t)} = e^{-a(\theta)}$, а $\Lambda(t) = \lambda t$, $t \geq 0$, то нетрудно вычислить, что $\hat{v}_0(t)$ есть простой поток восстановления с преобразованием Лапласа промежутка восстановления $\lambda(\lambda + a(\theta))^{-1}$.

Доказательство. Используя теорему 5.7, получим, что в силу условий (5.68), (5.69)

$$y_e([e^{-1}t]) \xRightarrow{\text{сл}} \prod (\Lambda(t)), \quad t \geq 0, \quad (5.70)$$

при этом из непрерывности $\Lambda(t)$ вытекает ординарность и стохастическая непрерывность потока $\prod (\Lambda(t))$ (т. е. моменты скачков κ_l , $l \geq 1$ имеют непрерывное распределение). Тогда из (5.70) и тождества Колмогорова — Феллера получим

$$\begin{aligned} P\{\varepsilon \mu_e(l) \geq x\} &= P\{y_e([e^{-1}x]) \leq l\} \rightarrow \\ &\rightarrow P\{\prod (\Lambda(x)) \leq l\} = P\{\kappa_l \geq x\}. \end{aligned}$$

Покажем, что $\varepsilon \mu_e(l)$ асимптотически не зависит от $\beta_e \mu_e([e^{-1}t])$. Для этого покажем, что в условиях теоремы $\{\beta_e \mu_e([e^{-1}t]), y_e([e^{-1}t])\} \xRightarrow{\text{сл}} \{\tau_0(t), \prod (\Lambda(t))\}$, $t \geq 0$, (5.71)

где $\tau_0(t)$ и $\prod (\Lambda(t))$ независимы. Для этого достаточно, чтобы для любых $s > 0$, $\theta > 0$

$$\begin{aligned} |M e^{-s \beta_e \mu_e([e^{-1}t]) - \theta y_e([e^{-1}t])} - M e^{-s \beta_e \mu_e([e^{-1}t])} \times \\ \times M e^{-\theta y_e([e^{-1}t])} | \delta_e \rightarrow 0. \end{aligned}$$

Справедливость этого вытекает из следующих оценок

$$\begin{aligned} \delta_e &= |M[M_{x_e(\cdot)} \exp\{-s \beta_e \mu_e([e^{-1}t])\} \cdot M_{x_e(\cdot)} \exp\{-\theta \times \\ &\times y_e([e^{-1}t])\} - M[M_{x_e(\cdot)} \exp\{-s \beta_e \mu_e([e^{-1}t])\}] \times \\ &\times M[M_{x_e(\cdot)} \exp\{-\theta y_e([e^{-1}t])\}]]| \leq M[M_{x_e(\cdot)} \exp\{-\theta y_e \times \\ &\times ([e^{-1}t])\} - M[M_{x_e(\cdot)} \exp\{-\theta y_e([e^{-1}t])\}]], \end{aligned}$$

где $M_{x_e(\cdot)}$ обозначает математическое ожидание при фиксированных значениях $x_e(1)$, $x_e(2)$, ...

При доказательстве теоремы 5.7 было показано, что

$$\begin{aligned} |M_{x_e(\cdot)} \exp\{-\theta y_e([e^{-1}t])\} - \exp\{-(1 - e^{-\theta}) \times \\ \times \sum_{k=1}^{[e^{-1}t]} p_e(k, x_e(k))\}| \leq \frac{1}{2} \sum_{k=1}^{[e^{-1}t]} p_e^2(k, x_e(k)) \times \\ \times \chi\left(\sum_{k=1}^{[e^{-1}t]} p_e^2(k, X_e(k)) \leq L\right) + \\ + 2\chi\left(\sum_{k=1}^{[e^{-1}t]} p_e^2(k, x_e(k)) > L\right) = \alpha_e(L). \end{aligned}$$

Поэтому окончательно

$$\delta_\varepsilon \leq M \left| \exp \left\{ - (1 - e^{-\theta}) \Lambda_\varepsilon ([\varepsilon^{-1}t]) \right\} - M \exp \left\{ - (1 - e^{-\theta}) \times \right. \right. \\ \left. \left. \times \Lambda_\varepsilon ([\varepsilon^{-1}t]) \right\} \right| + M \alpha_\varepsilon(L).$$

Теперь, так как функция $\exp \{ -x(1 - e^{-\theta}) \}$ по x непрерывна и ограничена в области $x \geq 0$ и выполнено (5.68), то

$$M \left| \exp \left\{ - (1 - e^{-\theta}) \Lambda_\varepsilon ([\varepsilon^{-1}t]) \right\} - M \exp \left\{ - (1 - e^{-\theta}) \right\} \times \right. \\ \left. \times \Lambda_\varepsilon ([\varepsilon^{-1}t]) \right\} \big| \rightarrow M \left| \exp \left\{ - (1 - e^{-\theta}) \Lambda(t) \right\} - \right. \\ \left. - \exp \left\{ - (1 - e^{-\theta}) \Lambda(t) \right\} \right| = 0,$$

т. е. $\delta_\varepsilon \rightarrow 0$, так как $M \alpha_\varepsilon(L) \rightarrow 0$ в силу (5.69).

Так как $\tau_0(t)$ и $\prod (\Lambda(t))$ независимы, то соответственно и моменты скачков κ_l , $l \geq 1$ не зависят от $\tau_0(t)$, и при этом из (5.71) следует, что

$$\{\beta_\varepsilon \mu_\varepsilon ([\varepsilon^{-1}t]), \varepsilon \kappa_\varepsilon(l)\} \xrightarrow{\text{сл}} \{\tau_0(t), \kappa_l\}, \quad t \geq 0, \quad l \geq 1.$$

Отметим теперь, что процесс $\mu_\varepsilon ([\varepsilon^{-1}t])$ монотонно не убывает с вероятностью единица, а $\tau_0(t)$ — стохастически непрерывный процесс, поэтому любой не зависящий от $\tau_0(t)$ момент κ_l будет с вероятностью единица точкой непрерывности $\tau_0(t)$. Воспользовавшись теперь теоремой о сходимости значений случайных процессов в случайной точке, являющейся с вероятностью единица точкой непрерывности (2) и представлением $\beta_\varepsilon \zeta_\varepsilon(l) = \beta_\varepsilon \mu_\varepsilon ([\varepsilon^{-1} \varepsilon \kappa_\varepsilon(l)])$, $l \geq 1$, получим, что

$$\beta_\varepsilon \zeta_\varepsilon(l) \xrightarrow{\text{сл}} \tau_0(\kappa_l), \quad l \geq 1,$$

что доказывает теорему.

Аналогично можно доказать более общее утверждение.

Теорема 5.9. Если при $\varepsilon \rightarrow 0$

$$\{\beta_\varepsilon \mu_\varepsilon ([\varepsilon^{-1}t]), \Lambda_\varepsilon ([\varepsilon^{-1}t])\} \xrightarrow{\text{сл}} \{\tau_0(t), \Lambda(t)\}, \quad t \geq 0,$$

где $\tau_0(t)$ — стохастически непрерывен, а $\Lambda(t)$ — монотонно неубывающий непрерывный процесс, и выполнено (5.69), то поток $\hat{\nu}_\varepsilon(\beta_\varepsilon^{-1}t)$ слабо сходится к потоку $\hat{\nu}_0(t)$, который строится так, как указывалось выше, где, однако, $\tau_0(t)$ и $\Lambda(t)$ могут быть зависимы, а процесс $\prod(\cdot)$ не зависит от $(\tau_0(\cdot), \Lambda(\cdot))$.

Доказательство. Прежде всего заметим, что поток $\hat{\nu}_0(t)$ полностью задается совместными распределениями

процесса $\{\tau_0(t), \prod(\Lambda(t))\}$, $t \geq 0$, при этом в случае, если $\prod(\cdot)$ не зависит от $(\tau_0(\cdot), \Lambda(\cdot))$, непосредственно убеждаемся, что

$$Me^{-s\tau_0(t) - \theta \prod(\Lambda(t))} = Me^{-s\tau_0(t) - (1 - e^{-\theta})\Lambda(t)}.$$

Согласно доказательству теоремы 5.8 нам достаточно доказать соотношение (5.71), в котором $\prod(\cdot)$ не зависит от пары $(\tau_0(\cdot), \Lambda(\cdot))$. Ввиду неравенства (5.65)

$$M_{x_\varepsilon(\cdot)} \exp\{-\theta y_\varepsilon([\varepsilon^{-1}t])\} = \exp\{-(1 - e^{-\theta})\Lambda_\varepsilon([\varepsilon^{-1}t], x_\varepsilon(\cdot))\} + \alpha_\varepsilon(x_\varepsilon(\cdot)),$$

где $M\alpha_\varepsilon(x_\varepsilon(\cdot)) \rightarrow 0$.

Поэтому, так как на фиксированной траектории $x_\varepsilon(\cdot)$ величины $\tau_\varepsilon(k, x_\varepsilon(k))$ и $\chi_\varepsilon(k, x_\varepsilon(k))$ независимы, то

$$\begin{aligned} M \exp\{-s\beta_\varepsilon\mu_\varepsilon([\varepsilon^{-1}t]) - \theta y_\varepsilon([\varepsilon^{-1}t])\} &= \\ = M[M_{x_\varepsilon(\cdot)} \exp\{-s\beta_\varepsilon\mu_\varepsilon([\varepsilon^{-1}t])\} \cdot M_{x_\varepsilon(\cdot)} \exp\{-\theta \times \\ \times y_\varepsilon([\varepsilon^{-1}t])\}] &= M[M_{x_\varepsilon(\cdot)} \exp\{-s\beta_\varepsilon\mu_\varepsilon([\varepsilon^{-1}t])\} \times \\ \times (\exp\{-(1 - e^{-\theta})\Lambda_\varepsilon([\varepsilon^{-1}t], x_\varepsilon(\cdot))\} + \alpha_\varepsilon(x_\varepsilon(\cdot))) &= \\ = M \exp\{-s\beta_\varepsilon\mu_\varepsilon([\varepsilon^{-1}t]) - (1 - e^{-\theta})\Lambda_\varepsilon([\varepsilon^{-1}t])\} + o(1) \rightarrow \\ \rightarrow M \exp\{-s\tau_0(t) - (1 - e^{-\theta})\Lambda(t)\} \end{aligned}$$

в силу ограниченности функции $\exp\{-sx - (1 - e^{-\theta})y\}$ в области $x \geq 0$, $y \geq 0$ и теоремы Хелли, что доказывает наше утверждение.

Основную сложность в реальных схемах составляет проверка условий (5.67), (5.68). Предельные теоремы для подобных схем исследовались многими авторами, в основном, когда $x_\varepsilon(k)$ — однородная марковская или стационарная последовательность, а распределения $\tau_\varepsilon(k, x)$ и $\chi_\varepsilon(k, x)$ не зависят от k . Приведем здесь одну теорему для достаточно общего класса последовательностей, удовлетворяющих условию равномерно сильного перемешивания.

Введем следующие обозначения. Пусть $\mathcal{F}_1^k$ и $\mathcal{F}_j$ — σ -алгебры, порожденные величинами $\{x_\varepsilon(i), 1 \leq i \leq k\}$ и $\{x_\varepsilon(l), l \geq j\}$ соответственно. Введем коэффициент равномерно сильного перемешивания

$$\varphi_\varepsilon(j) \sup_{k \geq 1} \sup_{A \in \mathcal{F}_1^k}, B \in \mathcal{F}_{k+j} |P(B/A) - P(B)|, j \geq 1,$$

а $\varphi_\varepsilon(0) = 1$.

Далее положим

$$a_e(s, k, x) = -\ln M \exp \{-s\beta_e \tau_e(k, x)\}, \quad x \in X, \quad k \geq 1, \quad s \geq 0,$$

$$\eta_e(s, k) = a_e(s, k, x_e(k)), \quad k \geq 1;$$

$$a_e(s, k) = M\eta_e(s, k), \quad \beta_e^2(s, k) = D\eta_e(s, k), \quad k \geq 1;$$

$$A_e(s, m) = \sum_{k=1}^m a_e(s, k), \quad m \geq 1.$$

Теорема 5.10. Если для любого фиксированного $s \geq 0$

$$\sum_{k=1}^{[\varepsilon^{-1}t]} \beta_e^2(s, k) \rightarrow 0, \quad (5.72)$$

и

$$q_e = \sum_{1 \leq k < j \leq nT} \sqrt{\varphi_e(j-k)} \beta_e(s, k) \beta_e(s, j) \rightarrow 0, \quad (5.73)$$

а

$$A_e(s, [\varepsilon^{-1}t]) \rightarrow A_0(s, t), \quad t \in [0, T], \quad (5.74)$$

где $A_0(0+, t) \equiv 0$,

то

$$\beta_e \mu_e([\varepsilon^{-1}t]) \xrightarrow{\text{ел}} \tau_0(t), \quad t \in [0, T],$$

где $\tau_0(t)$ — процесс с независимыми приращениями с преобразованием Лапласа вида

$$M \exp \{-s\tau_0(t)\} = \exp \{-A_0(s, t)\}, \quad t \in [0, T].$$

Замечание. Нетрудно проверить, что $q_e \rightarrow 0$, и выполнено (5.72), если

$$\lim_{j \rightarrow \infty} \overline{\lim}_{\varepsilon \rightarrow 0} \varphi_e(j) = 0; \quad (5.75)$$

$$\overline{\lim}_{\varepsilon \rightarrow 0} \sum_{k=1}^{[\varepsilon^{-1}T]} \beta_e(s, k) \leq C_s < \infty; \quad (5.76)$$

$$\beta_e(s) = \max_{k \leq \varepsilon^{-1}T} \beta_e(s, k) \rightarrow 0. \quad (5.77)$$

Доказательство. По построению

$$\begin{aligned} M \exp \{-s\beta_e \mu_e([\varepsilon^{-1}t])\} &= M [M_{x_{e(t)}} \exp \{-s\beta_e \mu_e([\varepsilon^{-1}t])\}] = \\ &= M \exp \left\{ - \sum_{k=1}^{[\varepsilon^{-1}t]} \eta_e(s, k) \right\}. \end{aligned}$$

Покажем теперь, что при любом $s \geq 0$

$$\gamma_e = M \left(\sum_{k=1}^{[\varepsilon^{-1}t]} \eta_e(s, k) - A_e(s, [\varepsilon^{-1}t]) \right)^2 \rightarrow 0.$$

Тогда из неравенства $|e^{-a} - e^{-b}| \leq |a - b|$, справедливого при $a \leq 0, b \leq 0$, которое непосредственно вытекает из формулы Лагранжа для функции e^{-x} на отрезке $[a, b]$, получим

$$|M \exp \left\{ - \sum_{k=1}^{[\varepsilon^{-1}t]} \eta_\varepsilon(s, k) \right\} - \exp \{ -A_0(s, t) \}| \leq \\ \leq \sqrt{M\gamma_\varepsilon} + |A_\varepsilon(s, [\varepsilon^{-1}t]) - A_0(s, t)|,$$

что, согласно (5.74) влечет слабую сходимость одномерных распределений (аналогично доказывается сходимость конечномерных распределений).

В самом деле

$$\gamma_\varepsilon = \sum_{k=1}^{[\varepsilon^{-1}t]} \beta_\varepsilon^2(s, k) + 2 \sum_{1 \leq k < j \leq \varepsilon^{-1}t} M(\eta_\varepsilon(s, k) - a_\varepsilon(s, k)) \times \\ \times (\eta_\varepsilon(s, j) - a_\varepsilon(s, j)).$$

Воспользуемся следующим известным неравенством, если ξ и η — соответственно $\mathcal{F}_0^k$ и $\mathcal{F}_j$ измеримы, $k < j$, то

$$|M \xi \cdot \eta - M \xi M \eta| \leq 2 \sqrt{\varphi_\varepsilon(j-k)} \sqrt{M|\xi|^2} \sqrt{M|\eta|^2}.$$

Отсюда следует оценка

$$\gamma_\varepsilon \leq \sum_{k=1}^{[\varepsilon^{-1}t]} \beta_\varepsilon^2(s, k) + 4q_\varepsilon \rightarrow 0,$$

что доказывает теорему.

Этот результат легко переносится на схемы суммирования индикаторов.

Следствие 5.9. Если выполнено условие (5.75),

$$\lim_{\varepsilon \rightarrow 0} \sum_{k=1}^{[\varepsilon^{-1}T]} \sqrt{M p_\varepsilon^2(k, x_\varepsilon(k))} \leq C < \infty; \quad (5.78)$$

$$\max_{k \leq \varepsilon^{-1}T} P\{\chi_\varepsilon(k, x_\varepsilon(k)) = 1\} \rightarrow 0; \quad (5.79)$$

$$\sum_{k=1}^{[\varepsilon^{-1}T]} M p_\varepsilon(k, x_\varepsilon(k)) \rightarrow \Lambda(t), \quad t \in [0, T], \quad (5.80)$$

то

$$y_\varepsilon([\varepsilon^{-1}t]) \xrightarrow{\text{сл}} \Pi(\Lambda(t)), \quad t \in [0, T].$$

Заметим, что в том случае, когда величины $\tau_\varepsilon(k, x)$ и $\chi_\varepsilon(k, x)$ равномерно по x малы в смысле

$$\lim_{\varepsilon \rightarrow 0} \varepsilon^{-1} \sup_{k \geq 1} \sup_{x \in X} a_\varepsilon(s, k, x) \leq C_s < \infty$$

и

$$\overline{\lim}_{\varepsilon \rightarrow 0} \varepsilon^{-1} \sup_{k \geq 1} \sup_{x \in X} p_\varepsilon(k, x) \leq C < \infty,$$

условия (5.76)–(5.79) автоматически выполняются.

Условия (5.74) и (5.80) обычно можно проверить в тех случаях, когда последовательность удовлетворяет некоторым свойствам эргодичности. Предположим, что существует последовательность вероятностных мер $\pi_\varepsilon(B)$ на (X, B_X) (в частности может быть $\pi_\varepsilon(B) \equiv \pi(B)$, $\varepsilon \geq 1$) таких, что

$$\lim_{\varepsilon \rightarrow 0} \lim_{j \rightarrow \infty} \Psi_\varepsilon(j) = 0, \quad (5.81)$$

где

$$\Psi_\varepsilon(j) = \sup_{k \geq 1} \sup_{A \in \mathcal{F}_1^k, B \in \mathcal{F}_{k+j}} |P(B/A) - \pi_\varepsilon(B)|.$$

В этом случае соответствующие характеристики можно выразить через интегралы по мере $\pi_\varepsilon(B)$ (аналог эргодической меры). Заметим, что если $\pi_\varepsilon(B) \equiv \pi(B)$, условие (5.81) означает равномерную эргодичность последовательности $x_\varepsilon(k)$.

Обозначим

$$b_\varepsilon(s, k) = \sup_{x \in X} a_\varepsilon(s, k, x), \quad k \geq 1,$$

$$\tilde{a}_\varepsilon(s, k) = \int_x a_\varepsilon(s, k, x) \pi_\varepsilon(dx), \quad \tilde{A}_\varepsilon(s, m) = \sum_{k=1}^m \tilde{a}_\varepsilon(s, k), \quad m \geq 1,$$

$$p_\varepsilon(k) = \sup_x p_\varepsilon(k, x), \quad \tilde{\lambda}_\varepsilon(k) = \int_x p_\varepsilon(k, x) \pi_\varepsilon(dx), \quad k \geq 1,$$

$$\Lambda_\varepsilon(m) = \sum_{k=1}^m \tilde{\lambda}_\varepsilon(k), \quad m \geq 1.$$

Тогда из теорем 5.8 и 5.10 как следствие вытекает:

Теорема 5.11. Пусть выполнено условие (5.81) и при любом фиксированном $s \geq 0$, $T > 0$

$$\sup_{k \leq \varepsilon^{-1}T} b_\varepsilon(s, k) \rightarrow 0, \quad \sup_{k \leq \varepsilon^{-1}T} p_\varepsilon(k) \rightarrow 0,$$

$$\overline{\lim}_{\varepsilon \rightarrow 0} \sum_{k=0}^{[\varepsilon^{-1}T]} b_\varepsilon(s, k) \leq C_s < \infty;$$

$$\overline{\lim}_{\varepsilon \rightarrow 0} \sum_{k=0}^{[\varepsilon^{-1}T]} p_\varepsilon(k) \leq C < \infty;$$

$$\tilde{A}_\varepsilon(s, [\varepsilon^{-1}t]) \rightarrow A_0(s, t), \quad t \geq 0;$$

$$\tilde{\lambda}_\varepsilon([\varepsilon^{-1}t]) \rightarrow \Lambda(t), \quad t \geq 0,$$

где $A_0(s, t)$ и $\Lambda(t)$ непрерывны по t и $A_0(0+, t) \equiv 0$.

Тогда

$$\{\beta_e \mu_e([e^{-1}t]), y_e([e^{-1}t])\} \xrightarrow{\text{сл}} \{\tau_0(t), \Pi(\Lambda(t))\}, t \geq 0,$$

где $\tau_0(t)$ и $\Pi(\Lambda(t))$ — независимы,

$$M \exp\{-s\tau_0(t)\} = \exp\{-A_0(s, t)\}, t \geq 0,$$

и при этом

$$\beta_e \zeta_e(l) \xrightarrow{\text{сл}} \tau_0(\chi_l), l \geq 1.$$

Доказательство. Легко видеть, что $\varphi_e(j) \leq 2 \Psi_e(j)$, $j \geq 1$, далее в наших обозначениях $\beta_e(s, k) \leq b_e(s, k)$, $k \geq 1$, $s \geq 0$, $M p_e^2(k, x_n(k)) \leq p_e^2(k)$, $k \geq 1$. Поэтому при наших предположениях выполнены условия (5.72), (5.73) и (5.78), (5.79) теоремы 5.10 и следствия 5.9. Осталось проверить (5.74) и (5.80). Для этого покажем, к примеру, что для любого $t > 0$

$$A_e(s, [e^{-1}t]) - \tilde{A}_e(s, [e^{-1}t]) \rightarrow 0, t > 0.$$

Воспользуемся неравенством:

если

$$\sup_{x \in X} |f(x)| \leq C, \sup_{A \in \mathcal{B}_X} |P(A) - Q(A)| \leq \alpha,$$

то

$$\left| \int_X f(x) P(dx) - \int_X f(x) Q(dx) \right| \leq 2C\alpha.$$

Тогда

$$|a_e(s, k) - \tilde{a}_e(s, k)| \leq 2b_e(s, k) \Psi_e(k)$$

и

$$|A_e(s, [e^{-1}t]) - \tilde{A}_e(s, [e^{-1}t])| \leq 2 \sum_{k=1}^{[e^{-1}t]} b_e(s, k) \Psi_e(k) \rightarrow 0,$$

что доказывает теорему.

П.2. В данной части рассмотрим потоки редких событий, которые генерируются неоднородными системами с непрерывным временем.

Пусть поведение системы описывается случайным процессом $\chi_e(t)$, $t \geq 0$ со значениями в измеримом пространстве $(X, \mathcal{B}_X)$. Система может генерировать некоторые импульсы (их можно интерпретировать как заявки либо отказы некоторых элементов, которые мгновенно ремонтируются), которые не влияют на поведение системы и поток которых $y_e(t)$ описывается следующим образом: задано семейство неотрицательных функций $\lambda_e(t, x)$, $t \geq 0$, $x \in X$, $\mathcal{B} \times \mathcal{B}_X$ — измеримых ($\mathcal{B}$ — σ -алгебра борелевских множеств на $[0, \infty)$),

и если траектория $x_\varepsilon(t) = x_\varepsilon(t, \omega)$ зафиксирована, то $y_\varepsilon(t, \omega)$ есть неоднородный пуассоновский поток с ведущей функцией

$$\gamma_\varepsilon(t, \omega) = \int_0^t \lambda_\varepsilon(u, x_\varepsilon(u, \omega)) du.$$

Если $x_\varepsilon(t)$ — ступенчатый процесс, то $y_\varepsilon(t)$ интерпретируется следующим образом: если $x_\varepsilon(u) = x$ при $t \leq u \leq t+h$, то вероятность появления одного импульса на $[t, t+h]$ есть $\lambda_\varepsilon(t, x)h + o(h)$, а двух и более — $o(h)$.

Тогда по определению можно записать

$$M \exp \{-\theta y_\varepsilon(t)\} = M \exp \{-\gamma_\varepsilon(t)(1 - e^{-\theta})\},$$

где $\gamma_\varepsilon(t) = \int_0^t \lambda_\varepsilon(u, x_\varepsilon(u)) du$. Из указанного представления

непосредственно вытекает утверждение: если

$$\gamma_\varepsilon(\varepsilon^{-1}t) \xrightarrow{\text{сл}} \gamma_0(t), \quad t \geq 0, \quad (5.82)$$

где $\gamma_0(t)$ — некоторый неубывающий процесс, то

$$y_\varepsilon(\varepsilon^{-1}t) \xrightarrow{\text{сл}} \Pi(\gamma_0(t)), \quad t \geq 0.$$

В этом случае, когда процесс $x_\varepsilon(t)$ удовлетворяет условиям перемешивания, можно указать достаточные условия для выполнения (5.82). Обозначим через $\mathcal{F}_0^t$ и $\mathcal{F}_u$ σ -алгебры, порожденные величинами $\{x_\varepsilon(s), s \leq t\}$ и $\{x_\varepsilon(s), s \geq u\}$ соответственно, и пусть

$$\varphi_\varepsilon(u) = \sup_{t \geq 0} \sup_{A \in \mathcal{F}_0^t, B \in \mathcal{F}_{t+u}} |P(B/A) - P(B)|, \quad u \geq 0.$$

Положим

$$\lambda_\varepsilon(u) = M \lambda_\varepsilon(u, x_\varepsilon(u)), \quad \beta_\varepsilon^2(u) = D \lambda_\varepsilon(u, x_\varepsilon(u)),$$

$$\Lambda_\varepsilon(t) = \int_0^t \lambda_\varepsilon(u) du.$$

Аналогично теореме 5.10 и замечанию к ней доказывается утверждение: если

$$q_\varepsilon = \int_0^{\varepsilon^{-1}t} \int_0^{\varepsilon^{-1}t-u} \beta_\varepsilon(u) \beta_\varepsilon(u+v) \sqrt{\varphi_\varepsilon(v)} dudv \rightarrow 0 \quad (5.83)$$

и $\Lambda_\varepsilon(\varepsilon^{-1}t) \rightarrow \Lambda(t)$, $t \geq 0$, то выполнено (5.82) с $\gamma_0(t) \equiv \Lambda(t)$ и

$$y_\varepsilon(\varepsilon^{-1}t) \xrightarrow{\text{сл}} \Pi(\Lambda(t)), \quad t \geq 0. \quad (5.84)$$

В частности, (5.83) выполнено, если:

$$\lim_{u \rightarrow \infty} \overline{\lim}_{\varepsilon \rightarrow 0} \varphi_{\varepsilon}(u) = 0;$$

$$\overline{\lim}_{\varepsilon \rightarrow 0} \int_0^{t/\varepsilon} \beta_{\varepsilon}(u) du \leq C_t < \infty, \quad t > 0,$$

и для некоторого $T > 0$

$$\lim_{\varepsilon \rightarrow 0} \sup_{\varepsilon u < t} \int_0^T \beta_{\varepsilon}(u + z) dz = 0, \quad t > 0.$$

Соответственно теореме 5.11 можно сформулировать аналогичное утверждение для эргодических процессов:

пусть существует последовательность вероятностных мер $\pi_{\varepsilon}(B)$ на $(X, \mathcal{B}_X)$ таких, что

$$\lim_{u \rightarrow \infty} \overline{\lim}_{\varepsilon \rightarrow 0} \sup_{t > 0} \sup_{A \in \mathcal{F}_0^t, B \in \mathcal{F}_{t+u}^t} |P(B/A) - \pi_{\varepsilon}(B)| = 0,$$

$$\overline{\lim}_{\varepsilon \rightarrow 0} \int_0^{t/\varepsilon} b_{\varepsilon}(u) du \leq C_t < \infty, \quad t > 0,$$

и для некоторого $T > 0$

$$\lim_{\varepsilon \rightarrow 0} \sup_{\varepsilon u < t} \int_0^T b_{\varepsilon}(u + z) dz = 0,$$

где $b_{\varepsilon}(u) = \sup_{x \in X} \lambda_{\varepsilon}(u, x)$, $u \geq 0$,

$$\int_0^{\varepsilon^{-1}t} \left(\int_X \lambda_{\varepsilon}(u, x) \pi_{\varepsilon}(dx) \right) du \rightarrow \Lambda(t), \quad t \geq 0.$$

Тогда выполнено (5.84).

В заключение отметим, что аналогично можно ввести неубывающие процессы типа времени наработки $\tau_{\varepsilon}(t)$ на траекториях $x_{\varepsilon}(t)$ так, что при фиксированной $x_{\varepsilon}(t, \omega)$

$$M \exp \{-\tau_{\varepsilon}(t)\} = M \exp \left\{ - \int_0^t a_{\varepsilon}(s, u, x_{\varepsilon}(u, \omega)) du \right\},$$

(т. е. в состоянии x на промежутке $[t, t+h]$ развивается процесс с независимыми приращениями с кумулянтной

$$\int_t^{t+h} a_{\varepsilon}(s, u, x) du.$$

Для процессов $\tau_e(t)$ справедливы указанные выше утверждения, связанные со сходимостью конечномерных распределений (выраженные в терминах характеристик $a_e(s, u, x)$), и имеет место утверждение, аналогичное теореме 5.8.

Рассмотрим теперь применение полученных результатов к анализу конкретных СМО.

1. Пусть дана система типа GI/G/1, входящий в поток в которую задается моментами поступления k -го требования

$Q_e(k) = \sum_{i=1}^k \xi_e(i)$, $k \geq 0$ ($S_e(0) = 0$), а время обслуживания k -го требования — $\eta_e(k+1)$, при этом величины $\xi_e(k)$, $k \geq 0$, и $\eta_e(k)$, $k \geq 1$ — независимы в совокупности. Если требование поступит в занятую систему, оно теряется. Пусть $v_e(t)$, $t \geq 0$ — поток потерянных требований. Справедливо утверждение:

если

$$\sup_{k \geq 1} (P\{\xi_e(k) \leq \eta_e(k)\} + P\{\xi_e(k) + \xi_e(k+1) \leq \eta_e(k)/\xi_e(k) \leq \eta_e(k)\}) \rightarrow 0 \quad (5.85)$$

и существуют нормирующие множители q_e и β_e такие, что

$$\sum_{k=1}^{[q_e^{-1}t]} P\{\xi_e(k) \leq \eta_e(k)\} \rightarrow \Lambda(t), \quad t \geq 0; \quad (5.86)$$

$$\sum_{k=1}^{[q_e^{-1}t]} (1 - M \exp\{-\theta \beta_e \xi_e(k)\}) \rightarrow A(\theta, t), \quad t \geq 0 \quad (5.87)$$

и

$$\sup_{k \geq 1} (1 - M \exp\{-\theta \beta_e \xi_e(k)\}) \rightarrow 0, \quad (5.88)$$

где $A(0^+, t) \equiv 0$, $t \geq 0$, и $\Lambda(t)$ и $A(\theta, t)$ непрерывны по t , то поток $v_e(\beta_e^{-1}t)$ слабо сходится к потоку, который задается последовательными моментами скачков ζ_l , $l \geq 1$, где $\zeta_l = \tau(\chi_l)$, $\tau(t)$ — неубывающий процесс с независимымиращениями вида $M \exp\{-\theta \tau(t)\} = \exp\{-A(\theta, t)\}$, а χ_l — момент l -го скачка процесса $\Pi(\Lambda(t))$, где $\Pi(\cdot)$ и $\tau(\cdot)$ — независимы.

Замечание. В условии (5.85) малость 1-го слагаемого означает, что с вероятностью, стремящейся к единице, каждое требование поступает в пустую систему, а малость 2-го слагаемого означает, что требования с близкой к единице вероятностью теряются по одному. В противном случае, когда, к примеру, $\eta_e(k)$ может с малой вероятностью принимать большие значения, на таких промежутках будет теряться много требований входящего потока и предельный поток отказов будет неординарным.

Доказательство. Положим $\chi_e(k) = 1$, если $\xi_e(k) \leq \eta_e(k)$ и 0 в противном случае. Далее введем семейства независимых в совокупности случайных величин $\xi_e^+(k)$ и $\xi_e^-(k)$, $k \geq 1$, с распределениями вида:

$$P\{\xi_e^+(k) < x\} = P\{\xi_e(k) < x/\xi_e(k) > \eta_e(k)\};$$

$$P\{\xi_e^-(k) < x\} = P\{\xi_e(k) < x/\xi_e(k) \leq \eta_e(k)\}, \quad k \geq 1.$$

Пусть

$$y_e(m) = \sum_{k=1}^m \chi_e(k), \quad m \geq 1.$$

Определим рекуррентным образом последовательности случайных величин:

$$v_e(1) = \min\{m : m \geq 1, y_e(m) = 1\};$$

$$\zeta_e(1) = \sum_{k=1}^{v_e(1)-1} \xi_e^+(k) + \xi_e^-(v_e(1));$$

$$\gamma_e(l) = \max\{s : s \geq 0, \sum_{k=v_e(l)}^{v_e(l)+s} \xi_e(k) \leq \eta_e(v_e(l))\xi_e(v_e(l)) \leq \leq \eta_e(v_e(l))\};$$

$$\hat{\zeta}_e(l) = \sum_{k=v_e(l)+1}^{v_e(l)+\gamma_e(l)+1} \xi_e(k), \quad l \geq 1;$$

$$v_e(l+1) = \min\{m : m > v_e(l) + \gamma_e(l) + 1, \chi_e(m) = 1\};$$

$$\zeta_e(l+1) = \zeta_e(l) + \hat{\zeta}_e(l) + \sum_{k=v_e(l)+\gamma_e(l)+2}^{v_e(l+1)-1} \xi_e^+(k) + \xi_e^-(v_e(l+1))$$

при этом величины $\xi_e^+(k)$, $\xi_e^-(k)$ не зависят от $v_e(l)$, $l \geq 1$.

Из построения следует, что введенные величины эквивалентны следующим функционалам для исходной системы: $\zeta_e(1)$ — моменту первой потери требования, $\gamma_e(1)+1$ — количеству требований, которые поступят на промежутке $[\sum_{k=1}^{v_e(1)-1} \xi_e^+(k), \sum_{k=1}^{v_e(1)-1} \xi_e^+(k) + \eta_e(v_e(1))]$ и, следовательно, потеряются. При этом из условия (5.85) следует, что $\gamma_e(l) \xrightarrow{P} 0$.

Далее, $\hat{\zeta}_e(l)$ — интервал времени, через который очередное требование впервые поступит в пустую систему после момента $\zeta_e(1)$, после чего мы ждем до следующего появления события $\chi_e(k) = 1$, что произойдет при $k = v_e(2)$ и т. д. Из условий (5.85), (5.88) следует, что $\gamma_e(l) \xrightarrow{P} 0$, $\beta_e \hat{\zeta}_e(l) \xrightarrow{P} 0$, $l \geq 1$, и, следовательно, с вероятностью, стремящейся к единице, $v_e(l)$ — номер l -го по порядку потерянного требования, а $\zeta_e(l)$ — момент его потери.

Так же, как это делалось в п. 5.1, используя неравенство Чебышева, получим следующие оценки:

$$0 \leq 1 - M e^{-\theta \beta_e \xi_e^-(k)} = 1 - \frac{M e^{-\theta \beta_e \xi_e^-(k)} \chi(\xi_e^-(k) \leq \eta_e(k))}{q_e(k)} = \\ = \frac{1}{q_e(k)} M (1 - e^{-\theta \beta_e \xi_e^-(k)}) \chi(\xi_e^-(k) \leq \eta_e(k)) \leq 1 - M e^{-\theta \beta_e \xi_e^-(k)},$$

так как функция $1 - e^{-\theta \beta_e x}$ не убывает, а функция $M \chi(x \leq \eta_e(k)) = P\{x \leq \eta_e(k)\}$ не возрастает по x (здесь $q_e(k) = P\{\xi_e^-(k) \leq \eta_e(k)\}$).

Отсюда, в частности, следует

$$\sup_{k \geq 1} (1 - M e^{-\theta \beta_e \xi_e^-(k)}) \rightarrow 0 \quad (5.89)$$

и неравенство

$$M e^{-\theta \beta_e \xi_e^-(k)} \leq M e^{-\theta \beta_e \xi_e^-(k)} \leq 1. \quad (5.90)$$

Теперь

$$M e^{-\theta \beta_e \xi_e^+(k)} = \frac{1}{1 - q_e(k)} M e^{-\theta \beta_e \xi_e^+(k)} \chi(\xi_e^+(k) > \eta_e(k)) = \\ = \frac{1}{1 - q_e(k)} (M e^{-\theta \beta_e \xi_e^+(k)} - q_e(k) M e^{-\theta \beta_e \xi_e^-(k)}).$$

Отсюда, пользуясь (5.90), получим неравенство

$$1 - M e^{-\theta \beta_e \xi_e^-(k)} \leq 1 - M e^{-\theta \beta_e \xi_e^+(k)} \leq \frac{1}{1 - q_e(k)} \times \\ \times (1 - M e^{-\theta \beta_e \xi_e^-(k)}),$$

откуда вытекает с использованием (5.85), (5.87), (5.88), что

$$\sup_{k \geq 1} (1 - M e^{-\theta \beta_e \xi_e^-(k)}) \rightarrow 0 \quad (5.91)$$

и

$$\sum_{k=1}^{[q_e^{-1}t]} (1 - M e^{-\theta \beta_e \xi_e^+(k)}) \rightarrow A(\theta, t), \quad t \geq 0. \quad (5.92)$$

Теперь из условий (5.85), (5.86) следует, что

$$y_e([q_e^{-1}t]) \xrightarrow{\text{сл}} \Pi(\Lambda(t)), \quad t \geq 0,$$

т. е.

$$q_e v_e(l) \xrightarrow{\text{сл}} \kappa_l, \quad l \geq 1.$$

Далее обозначим

$$\mu_e^+(m) = \sum_{k=1}^m \xi_e^+(k), \quad m \geq 1.$$

Из условий (5.91), (5.92) вытекает, что

$$\begin{aligned} M e^{-\theta \beta_e \mu_e^+ ([q_e^{-1} t])} &= \exp \left\{ \sum_{k=1}^{[q_e^{-1} t]} \ln M e^{-\theta \beta_e \xi_e^+(k)} \right\} = \\ &= \exp \left\{ - \sum_{k=1}^{[q_e^{-1} t]} (1 - M e^{-\theta \beta_e \xi_e^+(k)}) + \theta(1) \right\} \rightarrow \exp \{ -A(\theta, t) \}, \end{aligned}$$

что эквивалентно соотношению

$$\mu_e^+ ([q_e^{-1} t]) \xrightarrow{сл} \tau(t), \quad t \geq 0,$$

при этом, по построению, процессы $\tau(t)$ и $\Pi(\Lambda(u))$ независимы. Кроме того, заметим, что из (5.89) следует: $\beta_e \xi_e^-(v_e(l)) \rightarrow 0$ для любого $l \geq 1$.

Используя указанное замечание и теорему 5.8, получим наше утверждение.

Рассмотрим систему, состоящую из r различных приборов с функциями распределения времени работы $F_1(x), \dots, F_r(x)$ соответственно, которая функционирует следующим образом: задана стохастическая матрица $P = \|p_{ij}\|$, $i, j = \overline{1, r}$. Начальный прибор выбирается в соответствии с некоторым начальным распределением, далее после окончания работы i -го прибора с вероятностью p_{ij} мгновенно начинает работать j -й прибор. На промежутке $[t, t+h]$ i -й работающий прибор может отказать с вероятностью $\lambda_e(t, i)h + o(h)$, после отказа система простаивает время $\tau_e(i)$, после чего включается в соответствии с некоторым распределением и работает описанным выше образом. Пусть $v_e(t)$, $t \geq 0$ — поток отказов системы. Имеет место утверждение:

Пусть матрица P неприводима со стационарным распределением Π_k , $\mathcal{F}_k(x)$ — нерешетчатые,

$$\int_0^\infty x dF_k(x) = m_k < \infty, \quad k = \overline{1, r},$$

и существует множитель $\beta_e \rightarrow 0$ такой, что

$$\lim_{e \rightarrow 0} \int_0^{\beta_e^{-1} t} \lambda_e(u, i) du = \Lambda_i(t), \quad i = \overline{1, r}, \quad t \geq 0,$$

где $\Lambda_i(t)$ — непрерывны, $i = \overline{1, r}$, для некоторого $T > 0$

$$\lim_{\varepsilon \rightarrow 0} \sup_{\beta_\varepsilon u < t} \int_0^T \lambda_\varepsilon(u + z, i) dz = 0, \quad i = \overline{1, r}, \quad t > 0$$

и $\beta_\varepsilon \rho_\varepsilon(i) \rightarrow 0, \quad i = \overline{1, r}$.

Тогда поток $v_\varepsilon(\beta_\varepsilon^{-1}t)$ слабо сходится к $\Pi(\Lambda(t)), \quad t \geq 0$, где

$$\Lambda(t) = \left(\sum_{k=1}^r \pi_k m_k \right)^{-1} \cdot \sum_{i=1}^r \pi_i m_i \Lambda_i(t).$$

В самом деле, если $P\{\tau_\varepsilon(i) = 0\} = 1, \quad i = \overline{1, r}$, то поток отказов строится так, как указывалось в п. 2 на процессе $x(t)$, который в нашей схеме представляет ПМП с функциями распределения времен сидения $F_k(x)$ и вложенной цепью Маркова с матрицей вероятностей переходов P . Тогда при наших предположениях в силу эргодической теоремы для ПМП [10]:

$$\sup_{t > 0} \left| P\{x(t+u) = i/x(t) = j\} - \frac{\pi_i m_i}{\sum_{k=1}^r \pi_k m_k} \right| \xrightarrow{u \rightarrow \infty} 0$$

и наше утверждение прямо следует из приведенных выше в п. 2 результатов. Отметим, что так как $\Pi(\Lambda(t))$ на каждом конечном промежутке имеет конечное с вероятностью единица число скачков, а $\beta_\varepsilon \tau_\varepsilon(i) \rightarrow 0, \quad i = \overline{1, r}$, то интервалы простоя асимптотически не влияют на поведение потока отказов, как и начальное распределение в момент включения.

Заметим, что, используя результаты п.5.3, указанные результаты можно распространить также на тот случай, когда ПМП $x(t) = x_\varepsilon(t)$ образует s -множество, а функции $F_k(x)$ также зависят от параметра $\varepsilon > 0$.

Отметим, что в том случае, когда функции $F_k(x)$ не имеют математических ожиданий и $\lambda_\varepsilon(t, i) \equiv \varepsilon \lambda_i, \quad i = \overline{1, r}$, класс предельных потоков для $v_\varepsilon(\varepsilon^{-1}t)$ допускает представление вида

$$\sum_{k=1}^r \Pi_k(\lambda_k \Omega_k(t)), \quad t \geq 0,$$

где $\Pi_k(x), \quad k = \overline{1, r}$ — независимые в совокупности и от $\Omega_k(t), \quad k = \overline{1, r}$, пуассоновские процессы, а

$$\begin{aligned} \Omega_k(t) &= \tau_k(\mu(t) - 0) + (t - \tau(t)) \chi(\tau_k(\mu(t) + 0) \neq \\ &\neq \tau_k(\mu(t) - 0)), \end{aligned}$$

где $\tau_k(u)$, $k = \overline{1, r}$ — неотрицательные неубывающие и независимые процессы с независимыми приращениями,

$$\tau(u) = \sum_{k=1}^r \tau_k(u), \quad \mu(t) = \inf \{u : u \geq 0, \tau(u) \geq t\}.$$

Эти результаты вытекают из общих предельных теорем для суперпозиции случайных процессов, разработанных в [2] и их обсуждение выходит за рамки настоящей книги.

Задания для самоконтроля

1. Рассмотрим систему, которая имеет r состояний $\{1, \dots, r\}$ и в моменты вида kh , $k \geq 1$ ($h > 0$), может изменять свое состояние (из i в j перейти с вероятностью p_{ij} , где матрице $\|p_{ij}\|$, $i, j = \overline{1, r}$, соответствует неприводимая цепь Маркова). Пусть также, если система в момент kh находится в состоянии i , то с вероятностью $\lambda_i h$ она может выдать импульс, не влияющий на ее поведение. Через $y_h(t)$ обозначим количество импульсов на промежутке $[0, t]$. Описать предельное поведение потока $y_h(t)$ при $h \rightarrow 0$.

2. Пусть поведение системы описывается однородной неприводимой цепью Маркова $x(t)$, $t \geq 0$, с непрерывным временем и конечным множеством состояний. Если система находится в состоянии i на промежутке времени $[t, t+s]$, то с вероятностью $\varepsilon_i s + o(s)$ система может выдать импульс, не влияющий на ее поведение. Пусть $y_\varepsilon(t)$ — количество импульсов на промежутке $[0, t]$. Описать предельное поведение потока $y_\varepsilon(\varepsilon^{-1}t)$ при $\varepsilon \rightarrow 0$.

ГЛАВА 6. АСИМПТОТИЧЕСКОЕ УКРУПНЕНИЕ СЛОЖНЫХ СИСТЕМ

6.1. Процессы с дискретным вмешательством случая

При анализе поведения сложных стохастических систем значительный интерес вызывает вопрос о возможности приближения при определенных условиях процесса, описывающего поведение исходной системы, некоторым другим процессом с меньшим числом состояний и имеющим более простую структуру, например марковским или полумарковским процессом. Связано это с тем, что встречающиеся на практике системы, как правило, описываются процессами с большим числом состояний и достаточно сложной структурой. Поэтому проблема укрупнения состояний и тем самым понижения размерности системы играет важную

роль в практических задачах. Основная сложность, которая здесь возникает, связана с возможной потерей информации которая может произойти при переходе к меньшему числу состояний. Так даже при укрупнении дискретных однородных цепей Маркова полученный процесс будет вновь цепью Маркова лишь при весьма жестких линейных связях на переходные вероятности исходной цепи. А именно, пусть x_k , $k \geq 0$ — цепь Маркова с множеством состояний $I = \{1, 2, \dots\}$ и матрицей вероятностей переходов $P = \|p_{ij}\|$, $i, j \geq 1$. Пусть задано разбиение множества I на области $I_1, I_2, \dots (\bigcup_{k \geq 1} I_k = I, I_k \cap I_l = \emptyset, k \neq j)$.

Укрупненный процесс строим следующим образом:

для любого $k \geq 0$ $\hat{x}_k = j$, если $x_k \in I_j$, $j \geq 1$.

Тогда, как хорошо известно, $\hat{x}_k$ будет вновь цепью Маркова при любом начальном распределении тогда и только тогда, когда для любых множеств I_i и I_j величина $p(s, I_j) = \sum_{i \in I_j} p_{si}$ будет одинаковой для всех $s \in I_i$

(в этом случае величины $\hat{p}(I_i, I_j) = p(s, I_j)$, $s \in I_i$ будут представлять собой вероятности переходов из I_i в I_j для укрупненной цепи). Естественно, трудно ожидать, чтобы подобное условие часто выполнялось для большинства систем. Поэтому укрупненный процесс реально может быть лишь близок в той или иной форме к марковскому процессу. Вместе с тем при исследовании так называемых высоконадежных систем, некоторых классов СМО в условиях малой вероятности потери требования возникает проблема асимптотического укрупнения, которая заключается в следующем. Пусть задана система, характеристики которой зависят от некоторого малого параметра $\varepsilon > 0$ так, что множество состояний системы можно разбить на области и при $\varepsilon \rightarrow 0$ вероятности переходов между областями в некотором смысле стремятся к нулю. Проблема состоит в выяснении условий, при которых укрупненный процесс, полученный склеиванием состояний каждой области в одно состояние, будет сходиться при $\varepsilon \rightarrow 0$ к марковскому или полумарковскому процессу, а также, когда будут сходиться аддитивные функционалы типа полного дохода системы к соответствующему функционалу от предельного процесса.

На практике это означает возможность замены при достаточно малых значениях параметра ε исходного процесса предельным и исследовании характеристик последнего, которые с достаточной степенью точности будут приближать соответствующие характеристики исходной системы.

В настоящее время решению проблемы асимптотического укрупнения посвящено значительное количество работ, которые можно разделить на три направления. В работах З. С. Королюка и А. Ф. Турбина [10, 11] развивается асимптотический подход, связанный с обращением возмущенных на спектре операторов. Этим методом для одно- и двумерных марковских и полумарковских процессов с дискретным и произвольным пространством состояний получены условия сходимости укрупненных процессов к марковским и построены также асимптотические разложения по степеням малого параметра. И. Н. Коваленко [9] прямыми вероятностными методами получены оценки близости укрупненных, близких к марковским, процессов к цепи Маркова.

В данной главе развивается подход, изложенный в работах [2, 3]. Данный подход обладает теми преимуществами, что позволяет исследовать схемы асимптотического укрупнения, для которых в настоящее время не разработан соответствующий аналитический аппарат, например неоднородные по времени марковские процессы.

В основе данной методики лежат прямые вероятностные представления искомым характеристикам в терминах исходных вероятностных объектов, при этом существенно используется новый класс процессов, введенный в [2] — переключающиеся процессы.

В связи с этим рассмотрим вначале ряд важных классов случайных процессов, которые относятся к категории процессов с дискретным вмешательством случая.

Полумарковские процессы (ПМП), которые в свое время были введены как естественное обобщение марковских цепей, представляют собой достаточно удобный математический аппарат для описания различного рода прикладных моделей. Основное характеризующее свойство ПМП — наличие вложенной марковской цепи, т. е. таких моментов времени, значения процесса в которые образуют цепь Маркова. Это означает, что при фиксированном значении процесса в этот момент события из «будущего» не зависят от «прошлого» поведения процесса до данного момента. При этом ПМП меняет свое состояние скачком только в эти моменты.

Вместе с тем в естественных ситуациях возникают процессы, у которых можно выделить указанные марковские моменты времени, однако в промежутках между ними процесс меняется по некоторому закону. Подобные классы процессов носят название процессов с дискретным вмешательством случая, либо процессов с дискретной ком-

понентой, подразумевая под этим, что значение компоненты может меняться только скачком в некоторые случайные моменты времени.

Рассмотрим марковские процессы с полумарковским вмешательством случая. Конструктивно их можно описать следующим образом. Пусть задан ПМП $x(t)$ со значениями в X , а в пространстве Y задано параметрическое семейство марковских процессов $y(t, x)$ ($x \in X$ — значение параметра).

Тогда суперпозиция $y(t, x(t))$ и будет процессом с полумарковским вмешательством случая, а двухкомпонентный процесс $(x(t), y(t, x(t)))$ будет процессом с дискретной компонентой полумарковского типа.

В общем случае процесс $\xi(t), t \geq 0$, будет процессом с дискретным вмешательством случая, если существует последовательность случайных моментов времени $\tau_0 \leq \tau_1 \leq \tau_2 \leq \dots$ таких, что для любого k поведение процесса на промежутке (τ_k, ∞) полностью определяется значением $\xi(\tau_k)$ и не зависит от событий, связанных с моментами $t < \tau_k$. При этом двухкомпонентный процесс $(x(t), \xi(t)), t \geq 0$, где $x(t) = \xi(\tau_k)$ при $\tau_k \leq t < \tau_{k+1}, k \geq 0$ будет процессом с дискретной компонентой.

Заметим, что в том случае, когда при каждом $xy(t, x)$ — процесс с независимыми приращениями (ПНП), а $x(t)$ — марковский процесс, процесс $y(t, x(t))$ естественно назвать ПНП, управляемым марковским процессом, а процесс $(x(t), y(t, x(t)))$ называется марковским процессом, однородным по 2-й компоненте (МП ОВК).

Рассмотрим этот класс более подробно, поскольку указанные типы процессов возникают в качестве предельных при исследовании асимптотического поведения интегральных функционалов на процессах, допускающих асимптотическое укрупнение состояний.

Пусть $x(t), t \geq 0$ — марковский процесс с дискретным множеством состояний $\{1, 2, \dots\}$, $\xi(t), t \geq 0, k = 1, 2, \dots$ и $\{\xi_{kj}(t), t \geq 0, k, j = 1, 2, \dots\}$ — независимые в совокупности и от $x(\cdot)$ семейства процессов с независимыми приращениями и случайных величин, независимых при разных t , соответственно со значениями в R . Зададим конструктивно процесс $\{x(t), y(t)\}, t \geq 0$, следующим образом: если $x(0) = i_0$, то $y(t) = \xi_{i_0}(t)$ до тех пор, пока $x(t)$ находится в состоянии i_0 . В момент τ_1 перехода $x(t)$ в другое состояние i_1 , $y(t)$ получает скачок на величину $\xi_{i_0 i_1}(\tau_1)$, после чего $y(t)$ получает приращения, совпадающие с приращениями процесса $\xi_{i_1}(t)$, до момента τ_2 перехода $x(t)$ в следующее состояние i_2 и т. д. Построенный таким образом процесс $x(t), y(t)$ будет МП ОВК. Заметим,

что марковость процесса очевидна по построению, а однородность по второй компоненте означает, что для любых

$$\begin{aligned} i, j = 1, 2, \dots, y \in R, A \in \mathfrak{B}_R, s < t \\ P\{x(t) = j, y(t) \in A / x(s) = i, y(s) = y\} = \\ = P\{x(t) = j, y(t) \in A - y / x(s) = i, y(s) = 0\} \end{aligned}$$

и также следует из построения (здесь $\mathfrak{B}_R$ — σ -алгебра борелевских множеств в R , а множество $A - y = \{a : a + y \in A\}$). В том случае, когда $\xi_k(t)$ однородные по времени процессы, а распределения величин $\xi_{kj}(t)$ не зависят от t , $(x(t), y(t))$ также будет однородным по времени. Как известно, если $\xi_k(0) = 0$, то тогда в нашем случае $M \exp\{i\theta \xi_k(t)\} = \exp\{t A_k(\theta)\}$, $t \geq 0$, где функция $A_k(\theta)$ называется кумулянтной процесса.

Обозначим $\psi_{kj}(\theta) = M \exp\{i\theta \xi_{kj}(t)\}$, $k, j = 1, 2, \dots$

Положим

$$\varphi_{kj}(t, \theta) = M [\exp\{i\theta y(t)\} \chi(x(t) = j) / x(0) = k, y(0) = 0]. \quad (6.1)$$

Если множество состояний $x(t)$ конечно ($\{1, 2, \dots, m\}$), то матрица $\Phi_t(\theta) = \|\varphi_{kj}(t, \theta)\|_{k,j=1, \overline{m}}$ имеет вид $\Phi_t(\theta) = \exp\{tQ(\theta)\}$, где

$$Q(\theta) = \begin{vmatrix} -a_1 + A_1(\theta) a_{12} \psi_{12}(\theta) & \dots & a_{1m} \psi_{1m}(\theta) \\ a_{21} \psi_{21}(\theta) - a_2 + A_2(\theta) & \dots & a_{2m} \psi_{2m}(\theta) \\ \dots & \dots & \dots \\ a_{m1} \psi_{m1}(\theta) a_{m2} \psi_{m2}(\theta) & \dots & -a_m + A_m(\theta) \end{vmatrix},$$

где a_{kj} , $k \neq j$ — интенсивности вероятностей перехода $x(t)$, а $a_k = \sum_{j \neq k} a_{kj}$, $k = \overline{1, m}$.

К примеру отметим, что в п. 5.4 рассматривался пуассоновский поток $y(t)$, управляемый цепью Маркова $x(t)$. Поток $y(t)$ можно интерпретировать следующим образом. Пусть $\Pi_\lambda(t)$ — пуассоновский поток с параметром λ и задана совокупность параметров λ_k , $k = \overline{1, m}$. Тогда $y(t) = \Pi_{x(t)}(t)$, либо $y(t)$ есть вторая компонента МП ОВК $(x(t), y(t))$, где $y(t)$ строится по процессам $\Pi_{\lambda_k}(t)$, а величины скачков $\xi_{kj}(t) \equiv 0$, $t \geq 0$, $k, j = \overline{1, m}$.

Далее заметим, что в общей ситуации, когда вторая компонента аналогично управляется первой компонентой и строится по независимым семействам процессов с независимыми приращениями и величин скачков, но процесс $x(t)$ будет полумарковским, возникает конструкция процесса с независимыми приращениями, управляемого полумарковским процессом (ПНП ПМП). Пусть в этом случае ПМП

$x(t)$ задается вложенной цепью Маркова с матрицей вероятностей переходов $P = \|p_{ij}\|$, $i, j = \overline{1, m}$, и набором времен переходов $\{\tau_{ij}, i \neq j\}$. Обозначим $F_{ij}(z) = P\{\tau_{ij} < z\}$,

$$F_i(z) = \sum_{j \neq i} p_{ij} F_{ij}(z), \quad i = \overline{1, m}.$$

Предположим, что $\xi_k(t)$ — однородные ПНП, которые задаются кумулянтами $A_k(\theta)$, $k = \overline{1, m}$, а распределения величин $\xi_{kj}(t)$ также не зависят от t . Укажем метод нахождения распределений процесса $(x(t), y(t))$. С помощью формулы полной вероятности для величин $\varphi_{kj}(t, \theta)$ (см. (6.1)). Нетрудно получить следующую систему интегральных соотношений:

$$\begin{aligned} \varphi_{kj}(t, \theta) &= \delta_{kj} e^{t A_k(\theta)} (1 - F_k(t)) + \\ &+ \sum_{l \neq k} \int_0^t e^{u A_k(\theta)} p_{kl} \psi_{kl}(\theta) \varphi_{lj}(t-u, \theta) dF_{kl}(u), \\ &k, j = \overline{1, m}, \end{aligned} \quad (6.2)$$

где $\delta_{kj} = 1$, если $k = j$, и 0, если $k \neq j$. Обозначим

$$\rho_{kj}(s, \theta) = \int_0^\infty e^{-st} \varphi_{kj}(t, \theta) dt, \quad \operatorname{Re} s > 0;$$

$$f_{kj}(s) = \int_0^\infty e^{-sx} dF_{kj}(x), \quad f_k(s) = \int_0^\infty e^{-sx} dF_k(x), \quad k, j = \overline{1, m}.$$

Тогда, переходя в (6.2) к преобразованиям Лапласа, после несложных преобразований получим

$$\begin{aligned} \rho_{kj}(s, \theta) &= \delta_{kj} \cdot \frac{1 - f_k(s - A_k(\theta))}{s - A_k(\theta)} + \\ &+ \sum_{l \neq k} p_{kl} \psi_{kl}(\theta) f_{kl}(s - A_k(\theta)) \cdot \rho_{lj}(s, \theta), \quad k, j = \overline{1, m}. \end{aligned} \quad (6.3)$$

Пусть $\bar{\rho}_j$ и $\bar{f}_j$ — векторы-столбцы с компонентами соответственно $\rho_{kj}(s, \theta)$, $k = \overline{1, m}$ и $\delta_{kj} \frac{1 - f_k(s - A_k(\theta))}{s - A_k(\theta)}$, $k = \overline{1, m}$, а Φ — матрица с элементами $p_{kl} \psi_{kl}(\theta) \bar{f}_{kl}(s - A_k(\theta))$. Тогда соотношение (6.3) переписется в виде $\bar{\rho}_j = \bar{f}_j + \Phi \bar{\rho}_j$.

Предположим, что множество состояний ПМП $x(t)$ связно и не содержит мгновенных и поглощающих состояний (общая ситуация была исследована в [1]). При этом предположении при $\operatorname{Re} s > 0$, так как $\operatorname{Re} A_k(\theta) \leq 0$, $|f_{kl}(s - A_k(\theta))| < 1$, $k \neq l$, поэтому у матрицы Φ сумма модулей элементов любой строки меньше единицы. Следовательно, матрица $E - \Phi$ обратима, и тогда $\bar{\rho}_j = (E - \Phi)^{-1} \bar{f}_j$.

Развивая дальше рассмотренную конструкцию процессов и отказываясь от предположения, что на каждом участке постоянства компоненты $x(t)$ вторая компонента развивается как процесс с независимыми приращениями, получим более общий класс — переключающиеся процессы (ПП) [2]. Рассмотрим здесь более подробно конструкцию ПП, поскольку далее этот класс будет выступать в качестве эффективного аппарата исследования схем асимптотического укрупнения.

Пусть D_∞' — пространство функций, заданных на $[0, \infty)$, не имеющих разрывов второго рода, непрерывных справа и принимающих значения в R^r , $\mathcal{U}$ — минимальная σ -алгебра цилиндрических множеств в D_∞' , $\mathfrak{B}$ — σ -алгебра борелевских множеств в R , $(X, \mathfrak{B}_X)$ — некоторое измеримое пространство. Предположим, что заданы:

1) параметрическое семейство случайных процессов $\mathcal{T} = \{\xi^{(l)}(t, x, a), t \geq 0, x \in X, a \in R^r, l \geq 0\}$, распределения которых при разных l независимы;

2) семейство $\mathcal{U}$ — измеримых функционалов $\tau = \tau(x, z(\cdot))$, $x \in X$, $z(\cdot) \in D_\infty'$, заданных на D_∞' , причем для удобства будем считать, что $\tau(x, z(\cdot))$ — марковский момент для функции $z(\cdot)$, т. е. событие $\{\tau(x, z(\cdot)) \leq t\}$ полностью определяется отрезком траектории $z(u)$ на промежутке $[0, t]$;

3) параметрическое семейство случайных величин $T = \{\gamma^{(l)}(x, y, a), x, y \in X, a \in R^r, l \geq 0\}$, распределения которых при разных l независимы;

4) параметрическое семейство вероятностных мер на X

$$\mathcal{P} = \{P(x, A, a), x \in X, a \in R^r, A \in \mathfrak{B}_X\};$$

5) начальное распределение $p = \{p(A), A \in \mathfrak{B}_X\}$.

Построим по семействам 1) — 5) ПП $\{x(t), \xi(t)\}$, $t \geq 0$, со значениями в (X, R^m) следующим образом: определим рекуррентным образом последовательности $x_l, \tau_l, a_l, l \geq 0$:

$$a_0 = 0, \tau_0 = 0, P\{x_0 \in A\} = p(A), A \in \mathfrak{B}_X;$$

$$\tau_{l+1} = \tau_l + \tau(x_l, \xi^{(l)}(\cdot, x_l, a_l)), a_{l+1} = \xi^{(l)}(\tau_{l+1} - \tau_l,$$

$$x_l, a_l), P\{x_{l+1} \in A/x_l = x, a_{l+1} = a\} = P(x, A, a);$$

$l \geq 0$.

Тогда τ_l — последовательные моменты скачков $x(t)$, при этом $x(\tau_l) = x_l, l \geq 0$, а $\xi(0) = 0$ и на промежутке $[\tau_l, \tau_{l+1})$ $\xi(t) = \xi_l + \xi^{(l)}(t - \tau_l, x_l, a_l)$, где $\xi_{l+1} = \xi_l + \xi^{(l)}(\tau_{l+1} - \tau_l, x_l, a_l) + \gamma^{(l)}(x_l, x_{l+1}, a_{l+1}), l \geq 0, \xi_0 = \xi(0) = 0$.

Отметим, что для корректности построения необходимо предполагать выполненными естественные условия измеримости соответствующих характеристик введенных семейств по паре (x, a) относительно σ -алгебры $\mathfrak{B}_X \times \mathfrak{B}$ с тем, чтобы построенные величины x_l, τ_l, a_l были случайными величинами.

Для ПП $\{\kappa(t), \xi(t)\}, t \geq 0$, будем предполагать также выполненными условия регулярности: на каждом конечном промежутке $\kappa(t)$ имеет с вероятностью единица конечное число скачков.

Конструктивно процесс $\{\kappa(t), \xi(t)\}$ развивается следующим образом: $\xi(0) = 0$, а начальное значение $x_0 = \kappa(0)$ выбирается с распределением $p(A), A \in \mathfrak{B}_X$. Далее $\xi(t)$ получает приращения процесса $\xi^{(0)}(t, x_0, 0)$ до первого момента переключения $\tau_1 = \tau(x_0, \xi^{(0)}(\cdot, x_0, 0))$. В момент τ_1 выбирается следующее значение $\kappa(\cdot) x_1$ с распределением $P(x_0, A, a_1)$, где $a_1 = \xi^{(0)}(\tau_1, x_0, 0)$, процесс $\xi(\cdot)$ получает скачок на величину $\gamma^{(0)}(x_0, x_1, a_1)$ и далее получает приращения процесса $\xi^{(1)}(t - \tau_1, x_1, a_1)$ до следующего момента переключения $\tau_2 = \tau_1 + \tau(x_1, \xi^{(1)}(\cdot, x_1, a_1))$ и так далее.

Отметим, что если семейства $\mathcal{A}, \mathcal{T}, \mathcal{P}$ не зависят от параметров a и l , то процесс $\kappa(t)$ будет ПМП с временем пребывания в состоянии $x \tau(x, \xi^{(0)}(\cdot, x))$ и вложенной цепью Маркова, которая определяется вероятностями переходов $p(x, A)$. Если при этом процессы $\xi(t, x)$ являются при каждом x процессами с независимыми приращениями, а моменты $\tau(x, \xi(\cdot, x))$ не зависят от процессов $\xi(\cdot, x)$, то $\{\kappa(t), \xi(t)\}$ представляет собой ПНП ПМП. В частности, если все $\tau(x, \xi(\cdot, x))$ имеют показательное распределение, получим МП ОВК.

Заметим, что если моменты переключения являются моментами достижения некоторой области для процессов $\xi(t, x, a)$, которые при этом являются непрерывными марковскими либо линейными процессами, то $\xi(t)$ описывает эволюцию кусочно-непрерывного либо линейного агрегата, введенных в [9].

В качестве примеров моментов переключения можно рассмотреть следующие:

$$\begin{aligned} \tau(x, z(\cdot)) &= \inf \{t : t \geq 0, \varphi_x(t, z(t)) \in G\}; \\ \tau(x, z(\cdot)) &= \inf \{t : t \geq 0, \varphi_x(t, z(t)) - \varphi_x(t, z(t-0)) \in G\} \\ \tau(x, z(\cdot)) &= \inf \left\{ t : t \geq 0, \int_0^t \varphi_x(u, z(u)) du \in G \right\}, \end{aligned}$$

где $\varphi_x(t, z)$ — при каждом $x \in X$ непрерывная на $[0, \infty) \cdot R^r$

функция со значениями в R^p , а G — некоторая область в R^p . Обычно в реальных задачах возникают моменты подобного рода, связанные с первым попаданием процесса в некоторую область, моментом наступления скачка определенной величины, достижения интегральным функционалом типа дохода системы некоторого уровня и т. п.

Так, момент 1-й потери требования в системе GI/GI/m/r есть момент 1-го достижения процессом $Q(t)$ (величина очереди) уровня $m + r + 1$.

Рассмотрим здесь некоторую специальную схему построения ПП, которая укладывается в общую модель, однако более удобна при рассмотрении схем асимптотического укрупнения. Эта специальная схема связана со специальной конструкцией моментов переключения как моментов скачков некоторых вспомогательных процессов, а параметр a определяется как пара — момент скачка и значение этого процесса в этот момент. Формальная схема следующая. Пусть заданы:

1) параметрическое семейство случайных процессов $\mathcal{X}^{(l)} = \{(\xi^{(l)}(t, u, x), \chi^{(l)}(t, u, x)), t \geq u, x \in X, l \geq 0\}$, распределения которых при разных l независимы. Здесь процессы $\xi(\cdot)$ принимают значения в R^r , а $\chi(\cdot)$ — ступенчатый процесс, величины скачков которого принимают значения в X и $\chi^{(l)}(u, u, x) = 0$ для любых $u \geq 0, x \in X, l \geq 0$;

2) начальное распределение $p = \{p(A), A \in \mathfrak{X}_x\}$.

Обозначим через $\tau(u, x, l)$ — момент первого скачка процесса $\chi^{(l)}(t, u, x)$, а через $\beta(u, x, l)$ его величину (по построению $\tau(u, x, l) \geq u$). Определим вновь рекуррентно последовательности $x_l, \tau_l, l \geq 0$:

$$\tau_0 = 0, P\{x_0 \in A\} = p(A), A \in \mathfrak{X}_x;$$

$$\tau_{l+1} = \tau(\tau_l, x_l, l), x_{l+1} = \beta(\tau_l, x_l, l); \quad (6.4)$$

$$\xi_0 \doteq 0, \xi_{l+1} = \xi_l + \xi^{(l)}(\tau_{l+1}, \tau_l, x_l), l \geq 0,$$

и введем аналогично ПП $\{\kappa(t), \zeta(t)\}$: $\kappa(t) = x_l, \zeta(t) = \xi_l + \xi^{(l)}(t, \tau_l, x_l)$ при $\tau_l \leq t < \tau_{l+1}, l \geq 0$.

Как следует из построения, здесь моменты переключения определяются моментами первых скачков вспомогательных процессов $\chi^{(l)}(t, u, x)$, а величина скачка определяет следующее значение компоненты $\kappa(\cdot)$.

Рассмотрим примеры, иллюстрирующие возможности приложений описанных моделей к схемам неасимптотического укрупнения пространства состояний марковских процессов.

Пример 6.1. Укажем здесь интерпретацию момента первого выхода цепи Маркова из подмножества как момента первого скачка некоторого пуассоновского процесса, заданного на вспомогательной цепи.

Пусть $x(t)$, $t \geq 0$ — неоднородная цепь Маркова с множеством состояний $\{0, 1, \dots, r\}$, которая задается мгновенными интенсивностями вероятностей переходов

$$\lambda_{ij}(t), i, j = \overline{0, r} \quad (\lambda_{ii}(t) = \lambda_i(t) = \sum_{j \neq i} \lambda_{ij}(t)).$$

Если $x(0) = j \neq 0$, положим $\mu_j = \inf\{t : t \geq 0, x(t) = 0\}$.

Далее через $\tilde{x}(t)$, $t \geq 0$, обозначим цепь Маркова с множеством состояний $\{1, \dots, r\}$, которая задается мгновенными интенсивностями вероятностей перехода $\lambda_{ij}(t)$, $i, j = \overline{1, r}$, $i \neq j$, $(\lambda_{ii}(t) = \lambda_i(t) = \sum_{j \neq i, j \neq 0} \lambda_{ij}(t))$ и введем МП ОВК $(x(t), y(t))$, где $y(t)$ — пуассоновский поток, управляемый цепью $x(t)$, который задается совокупностью значений мгновенных параметров $\{\lambda_{i0}(t), i = \overline{1, r}\}$ (т. е. если $\tilde{x}(t) = i$, то мгновенная интенсивность вероятности скачка $y(t)$ $\lambda_{i0}(t)$).

Если $\tilde{x}(0) = j$, положим

$$\tau_j = \inf\{t : t > 0, y(t) > 0\},$$

т. е. τ_j — момент первого скачка $y(t)$. Легко видеть по построению, что распределения μ_j и τ_j совпадают (это непосредственно вытекает из того, что функции распределения величин τ_j и μ_j удовлетворяют одной и той же системе дифференциальных уравнений Колмогорова).

Пример 6.2. Докажем аналогичное утверждение для цепей с дискретным временем. Пусть $x(k)$, $k \geq 0$ — однородная цепь Маркова с множеством состояний $\{0, 1, \dots, r\}$ и матрицей вероятностей переходов $P \| p_{ij} \|$, $i, j = \overline{0, r}$. Если $x(0) = j$, положим

$$\mu_j = \min\{k : k \geq 1, x(k) = 0\}.$$

Через $\tilde{x}(k)$, $k \geq 0$ обозначим цепь Маркова с множеством состояний $\{1, \dots, r\}$ и матрицей вероятностей переходов $\tilde{P} = \| \tilde{p}_{ij} \|$, $i, j = \overline{1, r}$, где $\tilde{p}_{ij} = p_{ij} (1 - p_{i0})^{-1}$, $i, j = \overline{1, r}$. Зададим независимое от $\tilde{x}(k)$ семейство независимых в совокупности индикаторов $\{\chi(k, l), k = \overline{1, r}, l \geq 0\}$, где $P\{\chi(k, l) = 1\} = p_{k0}$, $P\{\chi(k, l) = 0\} = 1 - p_{k0}$, $k = \overline{1, r}, l \geq 0$. Положим

$$y(m) = \sum_{l=0}^{m-1} \chi(\tilde{x}(l), l)$$

и

$$\tau_j = \min \{m : m \geq 1, y(m) > 0\}, (\tilde{x}(0) = j).$$

Тогда распределения μ_j и τ_j совпадают. Докажем это. По определению

$$P\{\mu_j = k\} = \sum_{\langle i_1, \dots, i_{k-1} \rangle} p_{ji_1} p_{i_1 i_2} \dots p_{i_{k-2} i_{k-1}} p_{i_{k-1} 0},$$

для любого $k \geq 1$, где сумма берется по всем возможным перестановкам с повторениями $\langle i_1, \dots, i_{k-1} \rangle$ из множества $\{1, \dots, r\}$. Соответственно легко вычислить, что

$$P\{\tau_j = k\} = \sum_{\langle i_1, \dots, i_{k-1} \rangle} \tilde{p}_{ji_1} \tilde{p}_{i_1 i_2} \dots \tilde{p}_{i_{k-2} i_{k-1}} \times \\ \times (1 - p_j)(1 - p_{i_1 0}) \dots (1 - p_{i_{k-2} 0}) p_{i_{k-1} 0}$$

и по построению $\tilde{p}_{ij}$ нетрудно видеть, что $P\{\mu_j = k\} = P\{\tau_j = k\}$ для любого $k \geq 1$.

Этот результат справедлив для неоднородных цепей, а также для процессов с дискретным или непрерывным временем и произвольным пространством состояний.

Пример 6.3. Изучим схемы укрупнения по первой компоненте МП ОВК в терминах ПП. Пусть $(x(t), \zeta(t))$ — МП ОВК, который задается неоднородным марковским процессом с множеством состояний X и мгновенными интенсивностями вероятностей переходов $\lambda(x, A, t)$, $x \in X$, $A \in \mathfrak{B}_X$, $x \notin A$, $t \geq 0$ (т. е. $P\{x(t+h) \in A / x(t) = x\} = \lambda(x, A, t)h + o(h)$, при этом предполагаем, что $x(t)$ — регулярен и вероятности переходов могут быть найдены как единственное решение второй системы дифференциальных уравнений Колмогорова), а процесс $\zeta(t)$ задается совокупностью процессов с независимыми приращениями $\{\xi(t, x)$, $t \geq 0$, $x \in X\}$ (т. е. когда $x(t) = x$, $\zeta(t)$ получает приращения процесса $\xi(t, x)$). Для простоты считаем, что в моменты изменения состояний $x(t)$ дополнительных скачков $\zeta(t)$ не получает.

Предположим, что задано другое измеримое пространство $(\tilde{X}, \mathfrak{B}_{\tilde{X}})$ и $\mathfrak{B}_X$ — измеримое отображение $X \xrightarrow{f} \tilde{X}$. Тогда процесс $\hat{x}(t) = f(x(t))$ будет представлять собой укрупнение процесса $x(t)$, поскольку все состояния $x \in X$ такие, что $f(x) = \hat{x}$, склеиваются в одно состояние $\hat{x}$ процесса $\hat{x}(t)$. Опишем процесс $(\hat{x}(t), \xi(t))$ в терминах ПП. Для любого $\hat{x} \in \tilde{X}$ положим $I_{\hat{x}} = f^{-1}(\hat{x})$, т. е. $I_{\hat{x}} = \{x : x \in X, f(x) = \hat{x}\}$. Построим вспомогательные МП ОВК $\{x(t, u, \hat{x}, x)$, $(\xi(t, u, \hat{x}, x), \chi(t, u, \hat{x}, x))\}$, $t \geq u$, $x \in I_{\hat{x}}$, $\hat{x} \in \tilde{X}$ следую-

щим образом: $x(t, u, \hat{x}, x)$ — марковский процесс с множеством состояний $I_{\hat{x}}$, начальным состоянием x (т. е. $x(u, u, \hat{x}, x) = x$) и мгновенными интенсивностями вероятностей переходов $\lambda(x, A, t)$, $A \in I_{\hat{x}} \cap \mathfrak{B}_x$, $\xi(t, u, \hat{x}, x)$ строится на $(t, u, \hat{x}, x)$ с помощью процессов $\xi(t, z)$, а $\chi(t, u, \hat{x}, x)$ строится на $x(t, u, \hat{x}, x)$ как обобщенный пуассоновский процесс, управляемый $x(\cdot)$, при этом $\chi(u, u, \hat{x}, x) \equiv 0$, а

$$P\{\chi(t+h, u, \hat{x}, x) \in B | \chi(t, u, \hat{x}, x) = 0, x(t, u, x, x) = z\} = \lambda(z, B, t)h + o(h), \text{ если } B \subset X \setminus I_{\hat{x}} \text{ и } 0,$$

если $B \subset I_{\hat{x}}$. Отсюда непосредственно вытекает, что величина скачка процесса $\chi(t, u, \hat{x}, x)$ принимает значения в множестве $X \setminus I_{\hat{x}}$ и при условии, что в момент скачка t $x(t, u, \hat{x}, x) = z$, имеет распределение $\lambda(z, B, t) \cdot \lambda(z, X \setminus I_{\hat{x}}, t)^{-1}$, $B \subset X \setminus I_{\hat{x}}$.

Заметим, что поскольку нам достаточно знать траектории $\chi(\cdot)$ лишь до момента первого скачка, мы можем не определять, что означают суммы элементов из X .

Построим ПП $(\kappa(t), \xi(t))$ по введенным семействам $(\xi(t, u, \hat{x}, x), \chi(t, u, \hat{x}, x))$, так, как указывалось выше в конструкции $\{\mathcal{F}^{(1)}, p\}$ (см. (6.4)).

Имеет место.

Лемма 6.1. Если начальные распределения процессов совпадают, т. е. $p(A) = P\{x(0) \in A\}$, $A \in \mathfrak{B}_x$, то в этом случае и конечномерные распределения $(\kappa(t), \xi(t))$ совпадают с распределениями $(\hat{x}(t), \xi(t))$, $t \geq 0$.

Доказательство. Введем для процесса $x(t)$ следующие величины, связанные с отображением $f(\cdot)$. Положим $\mu(u, x) = \inf\{t : t \geq u, x(t) \notin I_{f(x)}\}$, где $x(u) = x$ и $\beta(u, x) = x(\mu(u, x))$ (напомним, что рассматриваемые нами процессы непрерывны справа). Далее рекуррентно определим: $\tilde{x}_0 = x(0)$, $\tilde{\tau}_0 = 0$, $\tilde{\tau}_{l+1} = \mu(\tau_l, x_l)$, $\tilde{x}_{l+1} = \beta(\tau_l, x_l)$, $l \geq 0$.

Заметим далее, что распределения $\mu(u, x)$ и $\tau(u, x)$ (момента первого скачка $\chi(t, u, \hat{x}, x)$) совпадают, при этом по построению

$$P\{\beta(u, x) \in B | \mu(u, x) = t, x(\mu(u, x) - 0) = z\} = \frac{\lambda(z, B, t)}{\lambda(z, X \setminus I_{f(x)}, t)}.$$

Отсюда непосредственно вытекает, что если $P\{x(0) \in A\} = p(A)$, то распределения величин $(\tilde{\tau}_1, \tilde{x}_1)$ совпадают

с распределениями величин (τ_1, x_1) , участвующих в построении ПП (см. (1.4)), и далее по индукции легко следует, что конечномерные распределения последовательностей $\{\tilde{x}_l, \tilde{\tau}_l, l \geq 0\}$ и $\{x_l, \tau_l, l \leq 0\}$ совпадают. Так как при этом компонента $\xi(t)$ в обоих случаях строится по одним и тем же ПНП $\xi(t, x)$, то отсюда сразу следует наше утверждение.

Пример 6.4. Сформулируем аналогичное утверждение для процессов с дискретным временем. В этом случае аналогами ПНП выступают суммы независимых случайных величин.

Пусть $x(k)$, $k \geq 0$ — марковский процесс с множеством состояний X , который задается вероятностями переходов на k -м шаге $p(x, A, k)$, т. е.

$$P\{x(k+1) \in A / x(k) = x\} = p(x, A, k), \quad x \in X,$$

$$A \in \mathfrak{B}_x, \quad k \geq 0.$$

Пусть также задано параметрическое семейство случайных величин $\{\xi(x, k), x \in X, k \geq 0\}$ таких, что при фиксированных $x_1, x_2, \dots$ величины $\xi(x_k, k)$, $k \geq 0$, независимы в совокупности, при этом выражения $M \exp\{i\lambda \xi(x, k)\} \mathfrak{B}_x$ — измеримы по x при любом $k \geq 0$. Положим $\zeta(k) = \sum_{l=0}^{k-1} \xi(x(l), l)$ и рассмотрим процесс $(x(k), \zeta(k))$, $k \geq 0$ — аналог МП ОВК с дискретным временем. Пусть снова $f(\cdot) — \mathfrak{B}_x$ — измеримое отображение X на $\hat{X}$. Положим $\hat{x}(k) = f(x(k))$ и рассмотрим процесс $(\hat{x}(k), \zeta(k))$, $k \geq 0$. Опишем его в терминах ПП. Аналогично построим вспомогательные МП ОВК $\{x(k, u, \hat{x}, x), \xi(k, u, \hat{x}, x), \chi(k, u, \hat{x}, x)\}$, $k \geq u$, $x \in I_{\hat{x}}, \hat{x} \in \hat{X}$, следующим образом: $x(k, u, \hat{x}, x)$ — марковский процесс с множеством состояний $I_{\hat{x}}$, $x(k, k, \hat{x}, x) = x$ и

$$\begin{aligned} P\{x(k+1, u, \hat{x}, x) \in A / x(k, u, \hat{x}, x) = z\} = \\ = \frac{p(z, A, k)}{p(z, I_{\hat{x}}, k)}, \quad z \in I_{\hat{x}}, \quad A \subset I_{\hat{x}}, \quad k \geq u. \end{aligned}$$

Далее,

$$\xi(k, u, \hat{x}, x) = \sum_{l=u}^{k-1} \xi(x(l, u, \hat{x}, x), l), \quad k > u,$$

а

$$\chi(k, u, \hat{x}, x) = \sum_{l=u}^{k-1} \chi(x(l, u, \hat{x}, x), l), \quad k > u,$$

где $\{\chi(x, l), x \in X, l \geq 0\}$ — параметрическое семейство

независимых в совокупности случайных величин таких, что

$$P\{\chi(x, l) = 0\} = p(x, I_{l(x)}, l);$$

$$P\{\chi(x, l) \in B\} = p(x, B, l), \quad x \in X, \quad B \subset X \setminus I_{l(x)}, \quad l \geq 0$$

Пусть вновь $\tau(u, x)$ — момент первого скачка $\chi(k, u, \hat{x}, x)$, т. е. $\tau(u, x) = \min\{k: k > u, \chi(k, u, \hat{x}, x) \neq 0\}$, а $\beta(u, x) = \chi(\tau(u, x), u, \hat{x}, x)$ — величина скачка. Определим по введенным семействам ПП $(\kappa(k), \zeta(k))$, $k \geq 0$, согласно формулам (6.4). Аналогично тому, как это делалось в примере 6.3, доказывается утверждение:

Если $p(A) = P\{x(0) \in A\}$, $A \in \mathfrak{B}_x$, то конечномерные распределения ПП $(\kappa(k), \zeta(k))$ совпадают с распределениями $(\hat{\kappa}(k), \hat{\zeta}(k))$, $k \geq 0$.

Несмотря на кажущуюся громоздкость конструкции заметим, что ПП строится по процессам, заданным на областях вида I_x , $x \in X$. Такие процессы проще моделировать, чем исходный, поскольку размерности отдельных компонент определяются размерностями областей и могут быть существенно меньше размерности всей системы. Кроме того, как будет показано далее, этот подход является весьма плодотворным применительно к задачам асимптотического укрупнения.

6.2. Асимптотическое укрупнение состояний случайных процессов

Исследуем общие условия сходимости укрупненных процессов к марковским и полумарковским процессам. Изложение в основном следует работам [1] и [2], п. 6.8.

Пусть $x_\varepsilon(t)$, $t \geq 0$ — некоторый непрерывный справа случайный процесс с дискретным множеством состояний I (характеристики процесса зависят от параметра $\varepsilon > 0$). Предположим, что задано разбиение множества I на под-

множества $I_1, I_2, \dots, I_r$ ($r \leq \infty$) (т. е. $\bigcup_{k=1}^r I_k = I$, $I_k \cap I_j = \emptyset$, $k \neq j$). Введем укрупненный процесс $\hat{x}_\varepsilon(t) = k$, если $x_\varepsilon(t) \in I_k$, $k = \overline{1, r}$, $t \geq 0$.

С процессом $\hat{x}_\varepsilon(t)$ связан вложенный процесс изменения состояний $y_\varepsilon(l)$, $l = 0, 1, \dots$, где $y_\varepsilon(0) = x_\varepsilon(0)$, а $y_\varepsilon(l) = \hat{x}(\mu_\varepsilon(l))$, $l \geq 1$, где $\mu_\varepsilon(l)$ — последовательные моменты изменения состояний $\hat{x}(t)$, т. е. $\mu_\varepsilon(0) = 0$,

а

$$\mu_\varepsilon(l+1) = \inf\{t: t > \mu_\varepsilon(l), \hat{x}_\varepsilon(t) \neq \hat{x}_\varepsilon(\mu_\varepsilon(l))\}, \quad l \geq 0.$$

Изучим условия слабой сходимости конечномерных распределений процессов $y_\varepsilon(k)$ и $\hat{x}_\varepsilon(\beta_\varepsilon^{-1}t)$ (β_ε — нормирующий множитель) к марковским и полумарковским процессам соответственно.

Пусть для каждого $k = \overline{1, r}$ задана последовательность конечных областей $D_N(k)$, $N \geq 1$, такая, что

$$D_N(k) \subset D_{N+1}(k), \quad N \geq 1, \quad \text{и} \quad \bigcup_{N \geq 1} D_N(k) = I_k.$$

Предположим, что начальное распределение процесса $x_\varepsilon(t)$ удовлетворяет условиям: для любого $k = \overline{1, r}$

$$\lim_{\varepsilon \rightarrow 0} P\{x_\varepsilon(0) \in I_k\} = \pi_k; \quad (6.5)$$

$$\begin{aligned} \lim_{N \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} [P\{x_\varepsilon(0) \in I_k \setminus D_N(k)\} + \\ + P\{x_\varepsilon(0) \in \bigcup_{m > N} I_m\}] = 0. \end{aligned} \quad (6.6)$$

Обозначим $\tau_\varepsilon(k) = \mu_\varepsilon(k+1) - \mu_\varepsilon(k)$, $k \geq 0$. Пусть $U_\varepsilon(k)$ — σ -алгебра событий, порожденная траекториями $x_\varepsilon(t)$ до момента $\mu_\varepsilon(k)$, а событие $A_\varepsilon(i, k, l)$ означает, что $x_\varepsilon(\mu_\varepsilon(l)) = i$, $i \in I_k$.

Теорема 6.1. Если выполнены условия (6.5), (6.6) для любых $k, j = \overline{1, r}$, $k \neq j$, $i \in I_k$, $l \geq 0$, и для любого $B_\varepsilon \times \times(k) \in U_\varepsilon(k)$ такого, что

$$\lim_{\varepsilon \rightarrow 0} P\{B_\varepsilon(k)\} > 0,$$

$$\lim_{\varepsilon \rightarrow 0} P\{y_\varepsilon(l+1) = j / A_\varepsilon(i, k, l), B_\varepsilon(k)\} = p_{kj}; \quad (6.7)$$

$$\begin{aligned} \lim_{N \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} P\{x_\varepsilon(\mu_\varepsilon(l+1)) \in I_j \setminus D_N(j) / A_\varepsilon(i, \\ k, l), B_\varepsilon(k)\} = 0; \end{aligned} \quad (6.8)$$

$$\begin{aligned} \lim_{N \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} P\{x_\varepsilon(\mu_\varepsilon(l+1)) \in \bigcup_{m > N} I_m / A_\varepsilon(i, \\ k, l), B_\varepsilon(k)\} = 0, \end{aligned} \quad (6.9)$$

то конечномерные распределения последовательности $y_\varepsilon(l)$, $l \geq 0$, слабо сходятся к распределениям цепи Маркова $y_0(l)$, $l \geq 0$, с множеством состояний $\{1, \dots, r\}$, начальным распределением π_k , $k = \overline{1, r}$, матрицей переходных вероятностей

$$P = \|p_{kj}\|, \quad k, j = \overline{1, r} \quad (p_{kk} = 0).$$

Отметим, что если множество I конечно, то условия (6.6), (6.8), (6.9) автоматически выполняются.

Потребуем теперь дополнительно, чтобы существовал нормирующий множитель β_ε такой, что

$$\lim_{\varepsilon \rightarrow 0} P \{ \beta_\varepsilon \tau_\varepsilon(l) < z, y_\varepsilon(l+1) = j/A_\varepsilon(i, k, l), B_\varepsilon(k) \} = \\ = p_{kj} F_{kj}(z), k, j = \overline{1, r}, k \neq j, i \in I_k, l \geq 0. \quad (6.10)$$

Введем ПМП $\hat{x}_0(t)$, который задается начальным распределением π_k , $k = \overline{1, r}$, вложенной цепью Маркова с матрицей переходных вероятностей $P = \| p_{kj} \|$, $k, j = \overline{1, r}$, и набором времен переходов $\{\tau_{kj}\}$ с функциями распределения $F_{kj}(z)$, $k, j = \overline{1, r}, k \neq j$.

Теорема 6.2. Если выполнены условия (6.5) — (6.10) и процесс $\hat{x}_0(t)$ регулярен, то конечномерные распределения $x_\varepsilon(\beta_\varepsilon^{-1}t)$ слабо сходятся к соответствующим распределениям ПМП $\hat{x}_0(t)$ во всех точках стохастической непрерывности $\hat{x}_0(t)$.

Доказательство. Введем для ПМП $x_0(t)$ следующие величины:

$$\mu_0(0) = 0, y_0(0) = \hat{x}_0(0), \mu_0(l+1) = \inf \{t: t > \\ > \mu_0(l), \hat{x}_0(t) \neq \hat{x}_0(\mu_0(l))\}, y_0(l) = \hat{x}_0(\mu_0(l)), \\ \tau_0(l) = \mu_0(l+1) - \mu_0(l), l \geq 0,$$

т. е. $\mu_0(l)$ — момент l -го скачка, а $y_0(l)$ — вложенная цепь. Заметим, что из условий (6.6) и (6.8) непосредственно вытекает, что

$$\sum_{k=1}^r \pi_k = 1, \sum_{j=1}^r p_{kj} = 1, k = \overline{1, r}.$$

Действительно, для величин π_k , к примеру,

$$1 = \lim_{\varepsilon \rightarrow 0} \sum_{k=1}^{\infty} P \{ x_\varepsilon(0) \in I_k \} \leq \lim_{N \rightarrow \infty} \overline{\lim}_{\varepsilon \rightarrow 0} \left[\sum_{k=1}^N P \{ x_\varepsilon(0) \in I_k \} + \right. \\ \left. + \sum_{k > N} P \{ x_\varepsilon(0) \in I_k \} \right] = \lim_{N \rightarrow \infty} \sum_{k=1}^N \pi_k = \sum_{k=1}^{\infty} \pi_k.$$

Далее предварительно докажем лемму.

Лемма 6.2. Если выполнены условия (6.5) — (6.10), то для любого $m \geq 1$ многомерное распределение вектора $\{y_\varepsilon(l), \beta_\varepsilon \tau_\varepsilon(l), l = \overline{0, m-1}, y_\varepsilon(m)\}$ слабо сходится к распределению вектора $\{y_0(l), \tau_0(l), l = \overline{0, m-1}, y_0(m)\}$.

Доказательство проведем по индукции. При $m=1$ утверждение непосредственно вытекает из наших условий.

Пусть теперь $m > 1$. Введем события

$$C_\varepsilon(m) = \{\beta_\varepsilon \tau_\varepsilon(l) < z_l, y_\varepsilon(l) = i_l,$$

$$l = \overline{0, m-1}, y_\varepsilon(m) = i_m\};$$

$$C_0(m) = \{\tau_0(l) < z_l, y_0(l) = i_l, l = \overline{0, m-1},$$

$$y_0(m) = i_m\}, m \geq 1.$$

По предположению индукции для некоторого $m > 1$ $P\{C_\varepsilon(m)\} \rightarrow P\{C_0(m)\}$ для любых $i_l \in \{1, \dots, r\}$, $l = \overline{0, m}$ и z_l , $l = \overline{0, m-1}$, являющихся точками непрерывности величины $P\{c_0(m)\}$. По определению ПМП $P\{c_0(m+1)\} = P\{c_0(m)\} \cdot p_{i_m, m+1} F_{i_m, m+1}(z_m)$. Отметим теперь следующее неравенство, справедливое для любых чисел $\alpha_j \geq 0$, $\beta_j \geq 0$,

$$\beta_j \leq 1, \sum_{j \geq 1} \alpha_j \leq 1;$$

$$\left| \sum_{j \geq 1} \alpha_j \beta_j - \alpha_0 \beta_0 \right| \leq \left| \sum_{j \geq 1} \alpha_j \beta_j - \beta_0 \sum_{j \geq 1} \alpha_j \right| + \beta_0 \left| \sum_{j \geq 1} \alpha_j - \alpha_0 \right| \leq \left| \sum_{j \geq 1} \alpha_j - \alpha_0 \right| + \max_{j = \overline{1, N}} |\beta_j - \beta_0| + \sum_{j > N} \alpha_j.$$

Используя это неравенство и формулу полной вероятности, расписывая $P\{C_\varepsilon(m+1)\}$ по системе событий $A_\varepsilon(j, i_m, m)$, $j \in I_{i_m}$, нетрудно получить оценку

$$\begin{aligned} & |P\{C_\varepsilon(m+1)\} - P\{C_0(m+1)\}| \leq |P\{C_\varepsilon(m)\} - \\ & - P\{C_0(m)\}| + \max_{j \in D_N(i_m)} |P\{\beta_\varepsilon \tau_\varepsilon(m) < z_m, \\ & y_\varepsilon(m+1) = i_{m+1}/C_\varepsilon(m), A_\varepsilon(j, i_m, m)\}| + \\ & + P\{C_\varepsilon(m), \bigcup_{j \in I_{i_m} \setminus D_N(i_m)} A_\varepsilon(j, i_m, m)\}. \end{aligned}$$

Отсюда очевидно следует, что при $\varepsilon \rightarrow 0$ правая часть выбором N может быть сделана, в силу (6.10), сколь угодно малой (для третьего члена это несложно доказывается по индукции с использованием (6.6) и (6.8)). Лемма доказана.

Из леммы, в частности, следует, что в условиях (6.5) — (6.9) конечномерные распределения последовательности $\{y_\varepsilon(l), l \geq 0\}$, сходятся к распределениям $\{y_0(l), l \geq 0\}$, что доказывает теорему 6.1.

Далее положим

$$v_\varepsilon(t) = \min \{l: l \geq 0, \beta_\varepsilon \mu_\varepsilon(l+1) \geq t\};$$

$$v_0(t) = \min \{l: l \geq 0, \mu_0(l+1) \geq t\}, t \geq 0.$$

Легко видеть, что

$$|P\{\hat{x}_\varepsilon(\beta_\varepsilon^{-1}t) = i\} - P\{\hat{x}_0(t) = i\}| \leq \sum_{k=0}^N |P\{\hat{x}_\varepsilon(\beta_\varepsilon^{-1}t) = i, \\ v_\varepsilon(t) = k\} - P\{\hat{x}_0(t) = i, v_0(t) = k\}| + \\ + P\{v_\varepsilon(t) > N\} + P\{v_0(t) > N\}. \quad (6.11)$$

В силу леммы 6.2

$$P\{\hat{x}_\varepsilon(\beta_\varepsilon^{-1}t) = i, v_\varepsilon(t) = k\} = P\{y_\varepsilon(k) = i, \\ \beta_\varepsilon \mu_\varepsilon(k) < t \leq \beta_\varepsilon \mu_\varepsilon(k+1)\} \xrightarrow{\varepsilon \rightarrow 0} P\{y_0(k) = i, \\ \mu_0(k) < t \leq \mu_0(k+1)\} = P\{\hat{x}_0(t) = i, v_0(t) = k\}$$

в точках непрерывности правой части по t . При этом также $v_\varepsilon(t) \xrightarrow{\text{сл}} v_0(t)$ в точках стохастической непрерывности $v_0(t)$. Так как по условию $\hat{x}_0(t)$ регулярен, т. е. $v_0(t)$ — собственная случайная величина, то $\lim_{N \rightarrow \infty} P\{v_0(t) > N\} = 0$ и

$$\lim_{N \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} P\{v_\varepsilon(t) > N\} \leq \lim_{N \rightarrow \infty} P\{v_0(t) > N\} = 0.$$

Отсюда следует, что правая часть (6.11) при $\varepsilon \rightarrow 0$ выбором N может быть сделана сколь угодно малой, откуда следует сходимость одномерных распределений $\hat{x}_\varepsilon(\beta_\varepsilon^{-1}t)$ к $\hat{x}_0(t)$. Аналогично доказывается сходимость конечномерных распределений. Теорема 6.2 доказана.

Заметим, что для простоты результаты сформулированы для дискретного множества состояний. В том случае, если каждая область I_k может быть произвольным множеством, необходимо потребовать существования некоторой монотонной последовательности $D_N(k) (\bigcup_{k \geq 1} D_N(k) = I_k)$, такой, что условия (6.7) — (6.10) выполнены равномерно по $i \in D_m(k)$ для любого $m \geq 1$.

Из полученных результатов вытекает, что марковский или полумарковский процесс допускает асимптотическое укрупнение состояний в том случае, когда время пребывания в каждой области I_k и вероятности перехода в следующие области в пределе не зависят от начального состояния I_k . Поэтому можно применить результаты об асимптотическом поведении времени пребывания ПМП в подмножестве состояний (пп. 5.2, 5.3) к схемам асимптотического укрупнения ПМП.

Рассмотрим следующие примеры.

Пример 6.5. Исследуем здесь конечные цепи Маркова и ПМП. Пусть $x_\varepsilon(i)$, $i \geq 0$ — однородная цепь Маркова

с конечным множеством состояний I и матрицей вероятностей переходов $P_\varepsilon = \|p_\varepsilon(i, j)\|$, $i, j \in I$. Предположим, что цепь удовлетворяет условию А: задано разбиение $I = \bigcup_{k=1}^r I_k$, при этом каждой матрице $p^{(k)} = \|p_0(i, j)\|$, $i, j \in I_k$ ($p_0(i, j) = \lim_{\varepsilon \rightarrow 0} p_\varepsilon(i, j)$) соответствует один существенный класс и существуют конечные пределы

$$b_{ij} = \lim_{\varepsilon \rightarrow 0} \varepsilon^{-1} p_\varepsilon(i, j), \quad i \in I_k, \quad j \in I_s, \quad k \neq s.$$

Через $q^{(k)}(i)$, $i \in I_k$ обозначим стационарное распределение для матрицы $P^{(k)}$. Положим

$$\lambda_{km} = \sum_{i \in I_k} q^{(k)}(i) \sum_{j \in I_m} b_{ij}, \quad \lambda_k = \sum_{m \neq k} \lambda_{km},$$

$$k, m = \overline{1, r}, \quad k \neq m \quad (6.12)$$

и

$$\hat{x}_\varepsilon(l) = k, \quad (6.13)$$

если $x_\varepsilon(l) \in I_k$, $k = \overline{1, r}$, $l \geq 0$.

Следствие 6.1. Если выполнено условие А и условие (6.5), то конечномерные распределения процесса $\hat{x}_\varepsilon(\lfloor \varepsilon^{-1}t \rfloor)$, $t \geq 0$, слабо сходятся к распределениям цепи Маркова с непрерывным временем $\hat{x}_0(t)$ с множеством состояний $\{1, \dots, r\}$, начальным распределением π_k , $k = \overline{1, r}$, и интенсивностями вероятностей переходов λ_{kj} , $k \neq j$.

Пусть теперь $\kappa_\varepsilon(t)$ — ПМП, который задан вложенной цепью $x_\varepsilon(l)$, $l \geq 0$, и набором времен переходов $\{\tau_\varepsilon(k, j)$, $k \neq j\}$. Введем процесс

$$\hat{x}_\varepsilon(t) = k, \quad (6.14)$$

если $\kappa_\varepsilon(t) \in I_k$; $k = \overline{1, r}$, $t \geq 0$.

Следствие 6.2. Если цепь $x_\varepsilon(l)$ удовлетворяет условиям А и (6.5) и существует нормирующий множитель β_ε такой, что

$$\lim_{\varepsilon \rightarrow 0} \varepsilon^{-1} (M \exp \{i\theta \beta_\varepsilon \tau_\varepsilon(l, j)\} - 1) = a_{lj}(\theta),$$

$$l, j \in I_k, \quad k = \overline{1, r}; \quad (6.15)$$

$$\lim_{\varepsilon \rightarrow 0} M \exp \{i\theta \beta_\varepsilon \tau_\varepsilon(l, j)\} = \rho_{lj}(\theta), \quad l \in I_k, \quad j \in I_m, \quad k \neq m, \quad (6.16)$$

то конечномерные распределения процесса $\hat{x}_\varepsilon(\beta_\varepsilon^{-1}t)$ слабо сходятся к распределениям ПМП $\hat{x}_0(t)$ с множеством состояний $\{1, 2, \dots, r\}$ в точках его стохастической непрерывности (предполагается, что $\hat{x}_0(t)$ — регулярен).

Здесь $x_0(t)$ задается начальным распределением π_k , $k = \overline{1, r}$, вложенной цепью Маркова с вероятностями переходов $\hat{p}_{km} = \lambda_{km} \lambda_k^{-1}$, $k, m = \overline{1, r}$, $k \neq m$, и набором времен переходов $\{\tau_{km}, k, m = \overline{1, r}, k \neq m\}$, где

$$\begin{aligned} M \exp \{i\theta \tau_{km}\} &= \psi_{km}(\theta) (\lambda_k - A_k(\theta))^{-1}, \\ A_k(\theta) &= \sum_{l, j \in I_k} q^{(k)}(l) \cdot p_0(l, j) a_{lj}(\theta); \\ \psi_{km}(\theta) &= \frac{\lambda_k}{\lambda_{km}} \sum_{l \in I_k} q^{(k)}(l) \sum_{j \in I_m} b_{lj} \rho_{lj}(\theta), \quad k \neq m \end{aligned}$$

(для определенности будем считать $0/0 = 1$).

Действительно, как было показано в п. 5.2, предельные распределения времен сидения и вероятностей переходов из каждой области I_k в нашем случае не зависят от начального состояния, следовательно выполнены условия (6.7) и (6.10). Теперь из следствия 5.4 вытекает, что вероятность покинуть область I_k через состояние $s \in I_k$ есть в пределе $\lambda_k^{-1} \cdot q^{(k)}(s) \sum_{l \notin I_k} b_{sl}$, а вероятность, выйдя из I_k , перейти в $l \notin I_k$, есть в пределе $(\sum_{l \notin I_k} b_{sl})^{-1} b_{sl}$.

Поэтому вероятность того, что цепь выйдет из области I_k через состояние $s \in I_k$ и перейдет после этого в состояние $l \notin I_k$, стремится к величине $\lambda_k^{-1} q^{(k)}(s) b_{sl}$. Отсюда непосредственно следуют выражения для величин $\hat{p}_{km}$. Далее в случае следствия 6.1 предельное распределение времени сидения в I_k , как следует снова же из следствия 5.3, будет показательным с параметром λ_k . Это соответствует цепи $x_0(t)$. Аналогично из формулы (5.24), следует, что в наших обозначениях для характеристической функции предельного распределения времени сидения в I_k имеет место формула

$$(\lambda_k - A_k(\theta))^{-1} \sum_{m \neq k} \frac{\lambda_{km}}{\lambda_k} \psi_{km}(\theta),$$

откуда непосредственно вытекают выражения для τ_{km} , поскольку для ПМП, если τ_k — время сидения в состоянии k , а τ_{km} — время перехода из k в m , то

$$M \exp \{i\theta \tau_k\} = \sum_{m \neq k} \hat{p}_{km} M \exp \{i\theta \tau_{km}\}.$$

Замечание. Если в условии (6.16)

$$\rho_{lj}(\theta) \equiv 1, \quad l \in I_k, \quad j \in I_m, \quad k \neq m, \quad -\infty < \theta < +\infty,$$

то распределения времен переходов τ_{km} будут иметь характеристические функции вида $\lambda_k (\lambda_k - A_k(\theta))^{-1}$ и не зависят от индекса m . При этом если $A_k(\theta) \equiv 0$, $-\infty < \theta < +\infty$, то состояние $\{k\}$ у процесса $x_0(t)$ будет мгновенным, если же $|A_k(\theta)| = \infty$, то $\{k\}$ будет поглощающим состоянием.

Используя результаты п. 5.3, можно распространить наши утверждения на более общую ситуацию, когда каждая область I_k может представлять собой s -множество.

Пример 6.6. Пусть снова $x_s(t)$, $t \geq 0$ — конечная однородная цепь Маркова, которая удовлетворяет условию В: задано разбиение $I = \bigcup_{k=1}^r I_k$, при этом каждая область I_k^s образует s -множество.

Согласно п. 5.3 введем следующие величины. Пусть

$$\tilde{P}_s^{(k)} = \|\tilde{p}_s^{(k)}(i, j)\|, \quad i, j \in I_k,$$

где

$$\tilde{p}_s^{(k)}(i, j) = p_s(i, j) (1 - \sum_{s \notin I_k} p_s(i, s))^{-1},$$

$q_s^{(k)}(i)$, $i \in I_k$ — стационарное распределение для матрицы $\tilde{P}_s^{(k)}$, $k = \overline{1, r}$, а

$$q_s(k) = \sum_{i \in I_k} q_s^{(k)}(i) \sum_{i \notin I_k} p_s(i, j);$$

$$\hat{p}_s(k, m) = \sum_{i \in I_k} q_s^{(k)}(i) \sum_{i \in I_m} p_s(i, j), \quad k, m = \overline{1, r}, \quad k \neq m$$

(заметим, что, в частности, некоторые I_k могут представлять собой одно состояние). Пусть также существуют пределы

$$\lim_{\varepsilon \rightarrow 0} g_s(k)^{-1} \hat{p}_s(k, m) = \hat{p}_{km}, \quad k, m = \overline{1, r}, \quad k \neq m, \quad (6.17)$$

и (конечные или бесконечные) пределы

$$\lim_{\varepsilon \rightarrow 0} \varepsilon^{-1} q_s(k) = \lambda_k, \quad k = \overline{1, r}. \quad (6.18)$$

Через $\hat{x}_0(t)$, $t \geq 0$, обозначим цепь Маркова с непрерывным временем, множеством состояний $\{1, \dots, r\}$ и начальным распределением π_k , $k = \overline{1, r}$, которая задается параметрами λ_k , $k = \overline{1, r}$, времен сидения в состояниях и вероятностями переходов вложенной цепи $\hat{p}_{km}$, $k, m = \overline{1, r}$, $k \neq m$ (случай $\lambda_k = \infty$ соответствует мгновенному состоянию $\{k\}$, а $\lambda_k = 0$ — поглощающему). Имеет место

Следствие 6.3. Если цепь $\hat{x}_0(t)$ регулярна и выполнены условия В, (6.5), (6.17) и (6.18), то конечномерные распределения процесса $\hat{x}_\varepsilon(\lfloor \varepsilon^{-1}t \rfloor)$, $t \geq 0$ (6.13), слабо сходятся к распределениям цепи $\hat{x}_0(t)$.

Пусть теперь $\kappa_\varepsilon(t)$ — ПМП с вложенной цепью $x_\varepsilon(l)$, $l \geq 0$, и набором времен переходов $\{\tau_\varepsilon(k, j), k \neq j\}$. Предположим, что существует нормирующий множитель β_ε такой, что выполнено (6.16) и

$$\lim_{\varepsilon \rightarrow 0} g_\varepsilon(k)^{-1} q_\varepsilon^{(k)}(l) p_\varepsilon(l, j) (\text{Mexp} \{i\theta \beta_\varepsilon \tau_\varepsilon(l, j)\} - 1) = a_{ij}(\theta), \times \\ l, j \in I_k, k = \overline{1, r}. \quad (6.19)$$

Обозначим здесь

$$A_k(\theta) = \sum_{l, j \in I_k} a_{lj}(\theta); \\ \psi_{km}(\theta) = \lim_{\varepsilon \rightarrow 0} \hat{p}_\varepsilon(k, m)^{-1} \sum_{l \in I_k} q_\varepsilon^{(k)}(l) \times \\ \times \sum_{j \in I_m} p_\varepsilon(l, j) \rho_{lj}(\theta), k, m = \overline{1, r}, k \neq m.$$

Введем ПМП $\hat{x}_0(t)$ с множеством состояний $\{1, \dots, r\}$, начальным распределением π_k , $k = \overline{1, r}$, вложенной цепью Маркова с вероятностями переходов $\hat{p}_{km}$, $k, m = \overline{1, r}$, $k \neq m$, и набором времен переходов $\{\tau_{km}, k \neq m\}$, где

$$\text{Mexp} \{i\theta \tau_{km}\} = \psi_{km}(\theta) (1 - A_k(\theta))^{-1}.$$

Следствие 6.4. Если ПМП $\hat{x}_0(t)$ регулярен, выполнены условия В, (6.5), (6.16), (6.17), (6.18), то конечномерные распределения процесса $\hat{x}_\varepsilon(\beta_\varepsilon^{-1}t)$ (6.14) слабо сходятся к распределениям $\hat{x}_0(t)$ в точках его стохастической непрерывности.

Доказательство приведенных утверждений непосредственно вытекает из теорем 5.3, 5.4 и результатов п. 5.3.

Далее отметим, что поскольку конечномерные распределения вложенного процесса и распределения времен пребывания в областях слабо сходятся к соответствующим предельным, то и распределение времени пребывания в некотором подмножестве областей сходится к распределению времени пребывания предельного ПМП в соответствующем подмножестве состояний, если оно не замкнуто. А именно, пусть $\kappa_\varepsilon(t)$ — введенный выше ПМП, а $G = \bigcup_{k \in A} I_k$, где

$A \subset \{1, 2, \dots, r\}$ (т. е. G — некоторое подмножество областей I_k).

Следствие 6.5. Если выполнены условия В, (6.16), (6.17), (6.19), у ПМП $\hat{x}_0(t)$ подмножество A не замкнуто и не содержит замкнутых подмножеств, и выполнено (6.5), где $\sum_{k \in A} \pi_k = 1$, то распределение величины $\beta_\varepsilon \Omega_\varepsilon$, где Ω_ε — время пребывания $x_\varepsilon(t)$ в подмножестве G , слабо сходится к распределению времени пребывания ПМП $\hat{x}_0(t)$ в подмножестве A .

6.3. Асимптотическое укрупнение СМО

Выше рассматривались общие условия сходимости укрупненных случайных процессов к марковским и полумарковским процессам, а также условия асимптотического укрупнения цепей Маркова и ПМП. Эти теоремы можно эффективно применять к схемам асимптотического укрупнения СМО. Рассмотрим некоторые достаточно общие классы систем, у которых могут возникать подобные эффекты (под эффектами типа укрупнения понимается ситуация, когда некоторые характеристики системы асимптотически эквивалентны соответствующим характеристикам процесса, состояния которого есть некоторые укрупненные подмножества состояний процесса, описывающего исходную систему). Пусть характеристики системы зависят от параметра ε , $\varepsilon > 0$ и $\varepsilon \rightarrow 0$, при этом система обладает свойством асимптотически быстрого обслуживания, т. е. вероятности обслуживания требования за промежуток времени между двумя последовательными поступлениями требований стремятся в определенном смысле к единице при $\varepsilon \rightarrow 0$. Будем рассматривать системы с ограниченным числом мест для ожидания, считать, что требование, поступившее в переполненную систему, теряется и изучать асимптотическое поведение момента первой потери требования и потока потерянных требований. Для большей конкретности будем предполагать, что на вход системы поступает поток полумарковского типа, а время обслуживания зависит только от типа обслуживаемого требования.

Первый класс систем характеризуется тем, что ПМП, управляющий входящим потоком, допускает асимптотическое укрупнение множества состояний, дисциплина очереди — «первый пришел — первый обслужился», а приборы в системе зафиксированы на своих местах и каждое поступающее требование по определенному фиксированному закону занимает один из свободных приборов либо становится в очередь. В этом случае естественно ожидать, что при достаточно широких предположениях предельное распределение нормированного момента первой потери требования

совпадает с распределением времени пребывания некоторого ПМП в подмножестве состояний, количество которых равно числу состояний, до которых асимптотически укрупняется ПМП, управляющий входящим потоком. Здесь эти результаты будут проиллюстрированы для системы типа $SM|M|n|m$.

Второй класс систем может быть охарактеризован фиксированным характером входящего потока, но переменной структурой системы, которая в зависимости от величины очереди и типов занятых приборов может некоторым образом менять внутреннюю свою структуру (например, порядок приборов). В этом случае при определенных предположениях типа «асимптотически быстрого» обслуживания могут также возникать эффекты асимптотического укрупнения состояний процесса, описывающего поведение системы. Эти результаты будут проиллюстрированы для случая системы $M|M|n$ с переменным порядком приборов.

Заметим что интересные эффекты типа укрупнения могут возникать при рассмотрении систем смешанного типа, а также при более общих зависимостях характера обслуживания и типов поступающих заявок. Для некоторых классов систем, поведение которых можно описать в терминах вспомогательного марковского или полумарковского процесса с конечным числом состояний, можно, используя результаты пп. 5.2—5.4 и п. 6.1, 6.2, непосредственно вычислить соответствующие характеристики (эти системы типов $SM|M|n|m$ либо $M_n|G|1|m$). Вместе с тем широкие классы систем не могут быть описаны в терминах процессов с конечным множеством состояний (например, $GI|G|n$), и тем не менее такие системы также при определенных условиях должны асимптотически укрупняться до процесса с конечным числом состояний (к примеру, если входящий поток управляется асимптотически укрупняемым ПМП). В связи с этим можно указать на широкие перспективы применения метода асимптотического укрупнения к анализу немарковских СМО. Рассмотрим далее некоторые конкретные схемы СМО.

Система типа $SM|M|n|m$. Пусть входящий поток управляется некоторым ПМП $\xi_e(t)$ (т. е. моменты поступления требований совпадают с моментами скачков $\xi_e(t)$), в системе имеется n идентичных приборов, время обслуживания распределено по показательному закону с параметром μ/e , есть m мест для ожидания. Дисциплина очереди — «первый пришел — первый обслужился». Требование, поступившее в занятую систему, теряется. Будем изучать

предельное поведение момента первой потери требования, который обозначим Ω_e .

Пусть $\xi_e(t)$ — ПМП с конечным множеством состояний $\{1, \dots, N\}$, который задается набором времен сидения в состоянии $\{\tau_e(k), k = \overline{1, N}\}$ и вложенной цепью Маркова $x_e(l), l \geq 0$, с матрицей вероятностей переходов $P_e = \|p_e \times \times (i, j)\|, i, j = \overline{1, N}$. Предположим, что матрице P_e в пределе соответствует r классов существенных состояний $I_1, I_2, \dots, I_r (r > 1)$. Напомним, что тот случай, когда $r = 1$, был описан в п. 5.4. Аналогично тому, вновь через $Q_e(t)$ обозначим количество требований в системе в момент t в предположении, что $m = \infty$, т. е. требования не теряются. Введем двумерный процесс $\eta_e(t) = \{\xi_e(t), Q_e(t)\}, t \geq 0$. Пусть $0 = \theta_e(0) \leq \theta_e(1) \leq \theta_e(2) \leq \dots$ — последовательные моменты скачков $\xi_e(t)$. Тогда они есть марковские моменты для $\eta_e(t)$, т. е. последовательность $\xi_e(k) = \eta_e \times \times (\theta_e(k) + 0), k \geq 1$, образует цепь Маркова. Построим вспомогательный ПМП $\bar{\eta}_e(t) = \{\xi_e(t), \bar{Q}_e(t)\}, t \geq 0$, у которого на каждом промежутке $[\theta_e(k), \theta_e(k+1)) \bar{Q}_e(t) = Q_e(\theta_e(k) + 0)$. Тогда момент первой потери требования совпадает с моментом первого выхода $\bar{\eta}_e(t)$ из подмножества состояний $\langle \alpha \rangle = \{(i, s), i = \overline{1, N}, s = \overline{0, n+m}\}$. Разобьем $\langle \alpha \rangle$ на подмножества: $\langle \alpha \rangle = \bigcup_{k=1}^r \langle \alpha \rangle_k$, где $\langle \alpha \rangle_k = \{(i, s), i \in I_k, s = \overline{0, n+m}\}$. Далее положим $G_e(k, x) = P\{\tau_e(k) < x\}, k = \overline{1, N}$. Введем величины

$$v_e^{(k)}(s) = \int_0^\infty e^{-\mu s_n x/e} dG_e(k, x), k = \overline{1, N}, s = 1, 2, \dots,$$

где $s_n = \min(s, n)$. Если

$$\lim_{h \rightarrow +0} \lim_{\varepsilon \rightarrow 0} G_e(k, h) = 0, k = \overline{1, N}, \quad (6.20)$$

то тогда из оценки

$$v_e^{(k)}(s) \leq G_e(k, h) + e^{-\mu s_n h/e},$$

верной для любого $h > 0$, следует, что $v_e^{(k)}(s) \rightarrow 0$ при $\varepsilon \rightarrow 0$ для любых $k = \overline{1, N}, s = 1, 2, \dots$. Вновь предположим, что величины $v_e^{(k)}(s)$ сравнимы, т. е. существуют константы $b_k^{(s)}$ и нормирующий множитель $\gamma_e \rightarrow 0$ такие, что

$$v_e^{(k)}(s) = \gamma_e b_k^{(s)} (1 + o(1)), k = \overline{1, N}, s = 1, 2, \dots \quad (6.21)$$

Пусть $\pi_i^{(k)}, i \in I_k$ — стационарное распределение на классе I_k в пределе для вложенной цепи $x_\varepsilon(l)$. Введем матрицы

$$A_k^{(s)} = \|p_{ij} b_i^{(s)}\|, i, j \in I_k, \text{ где } p_{ij} = \lim_{\varepsilon \rightarrow 0} p_\varepsilon(i, j).$$

Тогда, как показано в п. 5.4, если бы между классами не было переходов, то для случая, когда $\xi_\varepsilon(t)$ вначале находился в классе I_k , для стационарной вероятности выхода из подмножества $\langle \alpha \rangle_k$ имеет место формула

$$q_\varepsilon(\langle \alpha \rangle_k) = \gamma_\varepsilon^{m+n} \bar{\pi}_k A_k^{(1)} \dots A_k^{(n-1)} (A_k^{(n)})^{m+1} 1 (1 + 0(1)), \quad (6.22)$$

где $\bar{\pi}_k$ — вектор-строка с компонентами $\pi_i^{(k)}, i \in I_k$.

Предположим, что вложенная цепь $x_\varepsilon(l)$ удовлетворяет условию:

$$\lim_{\varepsilon \rightarrow 0} \gamma_\varepsilon^{-(n+m)} \sum_{i \in I_k} \pi_i^{(k)} \sum_{j \in I_m} p_\varepsilon(i, j) = u_{km} < \infty, \\ k, m = \overline{1, r}, k \neq m. \quad (6.23)$$

Это означает, что вероятности переходов между классами I_k процесса $\xi_\varepsilon(t)$ имеют порядок γ_ε^{n+m} , т. е. порядка вероятности потери требования.

Рассмотрим подмножество состояний $\langle \alpha \rangle_k$ для процесса $\bar{\eta}_\varepsilon(t)$. Как следует из п. 5.4 (п. 2) и условия 6.23, $\langle \alpha \rangle_k$ образует s -многообразие порядка $m+1$, поскольку вероятности переходов с s -го на $(s+1)$ -й уровень порядка γ_ε для любого $s = \overline{0, n+m}$, а вероятности переходов из $\langle \alpha \rangle_k$ в другие подмножества порядка γ_ε^{n+m} . Поэтому для стационарного распределения на s -многообразии $\langle \alpha \rangle_k$, которое определяется по пересчитанной матрице вероятностей переходов согласно лемме 5.8, сохраняются формулы типа (5.48). Напомним, что если $\pi_\varepsilon(i, l)$ — стационарная вероятность состояния (i, l) подмножества $\langle \alpha \rangle_k$, то $\pi_\varepsilon(i, 0) = \pi_i^{(k)} (1 + 0(1)), i \in I_k$. Обозначим $\langle \alpha \rangle_0 = \{(i, l), i = \overline{1, N}, l \geq n+m+1\}$, т. е. попадание в $\langle \alpha \rangle_0$ означает потерю требования. Определим на классах $\langle \alpha \rangle_k, k = \overline{0, r}$, укрупненный процесс $\kappa_\varepsilon(t) = k$, если $\bar{\eta}_\varepsilon(t) \in \langle \alpha \rangle_k, k = \overline{0, r}$. Так как нас интересует поведение момента первого попадания в $\langle \alpha \rangle_0$, то можно считать 0 поглощающим состоянием для $\kappa_\varepsilon(t)$.

Определим стационарные укрупненные вероятности переходов между классами $\langle \alpha \rangle_k$. Согласно общим формулам,

приведенным в лемме 5.8, и формулам для главных частей стационарных вероятностей в нашем случае легко вычислить, что

$$\hat{p}_e(k, m) = \gamma_e^{n+m} u_{km} (1 + 0(1)), \quad k, m = \overline{1, r}, \quad k \neq m,$$

а

$$\hat{p}_e(k, 0) = \gamma_e^{n+m} u_{k0} (1 + 0(1)), \quad k = \overline{1, r},$$

где

$$u_{k0} = \bar{\pi}_k A_k^{(1)} \dots A_k^{(n-1)} A_k^{(n)m+1} 1.$$

Далее предположим, что величины $\tau_e(j)$ притягиваются к некоторым безгранично делимым законам, т. е. существует нормирующий множитель β_e такой, что

$$M \exp \{-\beta_e \theta \tau_e(j)\} = 1 - \gamma_e^{n+m} a_j(\theta) + 0(\gamma_e^{n+m}), \quad j = \overline{1, N}. \quad (6.24)$$

Тогда следуя рассуждениям п. 2 п. 5.4, получим, что предельное распределение времени пребывания $\Omega_e(\langle \alpha \rangle_k)$ процесса $\bar{\eta}_e(t)$ в подмножестве $\langle \alpha \rangle_k$ не зависит от начального состояния и также последующего перехода и при этом из формул примера 6.6

$$M \exp \{-\theta \beta_e \Omega_e(\langle \alpha \rangle_k)\} \rightarrow c_k (c_k + A_k(\theta))^{-1}, \quad (6.25)$$

где

$$c_k = \sum_{l=0}^r u_{kl};$$

$$A_k(\theta) = \sum_{l \in I_k} \pi_l^{(k)} a_l(\theta), \quad k = \overline{1, r}.$$

Обозначим $\hat{p}_{km}^* = u_{km} c_k^{-1}$, $k = \overline{1, r}$, $m = \overline{0, r}$, $k \neq m$. Пусть $\bar{\kappa}_0(t)$ — ПМП с множеством состояний $\{0, 1, \dots, r\}$, начальным распределением ρ_k , $k = \overline{1, r}$ (0 — поглощающее состояние), вложенной цепью Маркова с вероятностями переходов $\hat{p}_{km}$, $k = \overline{1, r}$, $m = \overline{0, r}$, $k \neq m$, и набором времен сидения $\{\tau_k, k = \overline{1, r}\}$, где

$$M \exp \{-\theta \tau_k\} = c_k (c_k + A_k(\theta))^{-1}. \quad (6.26)$$

Из следствия 6.5 и предыдущих рассуждений вытекает следующая теорема.

Теорема 6.3. Если выполнены условия (6.20), (6.21), (6.23), (6.24) и $P\{\xi_e(0) \in I_k\} \rightarrow \rho_k$, $k = \overline{1, r}$, а $Q_e(0) \leq n + m$, то предельное распределение величины $\beta_e \Omega_e$ (момента первой потери требования) слабо сходится к распределению времени пребывания ПМП $\bar{\kappa}_0(t)$ в подмножестве состояний $\{1, \dots, r\}$.

Отметим, в частности, что если величины $\tau_e(j)$, $j = \overline{1, N}$ не зависят от параметра e и имеют конечные математические ожидания m_j , $j = \overline{1, N}$, то в этом случае следует положить $\beta_e = \gamma_e^{n+m}$ и тогда $a_j(\theta) = \theta m_j$, $j = \overline{1, N}$ и $\tilde{\kappa}_0(t)$ будет представлять собой цепь Маркова с интенсивностями вероятностей переходов

$$\hat{\lambda}_{km} = u_{km} M_k^{-1}, \quad k = \overline{1, r}, \quad m = \overline{0, r},$$

где

$$M_k = \sum_{i \in I_k} \pi_j^{(k)} m_j.$$

Используя результаты, описанные в п. 5.4 (п. 3.), можно также исследовать поток потерянных требований. Нетрудно видеть, что при исходных предположениях, если процесс $\tilde{\eta}_e(t)$ попал на $m+n+1$ — уровень из множества $\langle \alpha \rangle_k$, то с вероятностью, стремящейся к единице, процесс в момент поступления следующего требования вернется в подмножество $\langle \alpha \rangle_k$, при этом требование поступит в пустую систему. Тем самым в терминах укрупненного процесса в пределе нормированные моменты отказа — это моменты попадания в состояние $\{0\}$, при этом, если переход в $\{0\}$ был из состояния $\{k\}$ ($k \neq 0$), то процесс мгновенно возвращается в $\{k\}$. Поэтому можно задать ПМП $\tilde{\kappa}_0(t)$ с множеством состояний $\{1, \dots, r, 0_1, \dots, 0_r\}$, у которого $0_1, \dots, 0_r$ — мгновенны, времена сидения в состояниях $1, \dots, r$ определяются формулой (6.26), а вероятности переходов вложенной цепи имеют вид

$$p(k, m) = \hat{p}_{km}, \quad k, m = \overline{1, r}, \quad k \neq m;$$

$$p(k, 0_k) = \hat{p}_{k0}, \quad p(0_k, k) = 1, \quad k = \overline{1, r},$$

а вероятности других переходов равны 0. Тогда получим следующее утверждение.

Следствие 6.6. В условиях теоремы 6.3 конечномерные распределения потока потерянных требований $v_e(\beta_e^{-1}t)$ слабо сходятся в точках множества $Q \subset [0, \infty)$ к распределениям потока $v_0(t)$, моменты поступления требований которого есть последовательные моменты попадания процесса $\tilde{\kappa}_0(t)$ в подмножество состояний $\{0_1, \dots, 0_r\}$ (здесь Q — множество точек непрерывности функций распределения величин τ_k и их всевозможных сверток).

Доказательство следствия непосредственно вытекает из предыдущих рассуждений и теоремы о сходимости укрупненных ПМП к предельным, которая охватывает и ту

ситуацию, когда у предельного процесса могут быть мгновенные состояния.

Заметим, что читатель в качестве самостоятельного утверждения может распространить приведенные выкладки на случай, когда приборы имеют различные интенсивности обслуживания порядка $O(1/\varepsilon)$ и фиксированный порядок, при этом требование поступает на прибор, имеющий наименьший номер из свободных приборов.

Система $M|M|n$ с переменным порядком обслуживания. Рассмотрим систему с потерями типа $M|M|n$, на которую поступает простейший поток требований с параметром λ , а время обслуживания i -м прибором имеет показательное распределение с параметром c_i/ε , ($c_i > 0$), $i = \overline{1, n}$. Будем считать, что система имеет n позиций ($1, 2, \dots, n$), на которых размещаются обслуживающие приборы. Порядок расположения приборов на позициях будет задаваться перестановками вида $(i_1, \dots, i_n)$, т. е. прибор типа i_k находится на k -й позиции, $k = \overline{1, n}$. Поступающее требование занимает тот свободный прибор, который имеет позицию с наименьшим номером, либо теряется, если все приборы заняты. В момент окончания обслуживания данный прибор мгновенно перемещается на n -ю позицию, а приборы, расположенные на позициях с большими номерами, мгновенно сдвигаются на одну позицию вниз, т. е. если при порядке приборов $(i_1, \dots, i_n)$ прибор i_k закончил обслуживание, то порядок мгновенно изменяется на $(i_1, \dots, i_{k-1}, i_{k+1}, \dots, i_n, i_k)$.

Подобную схему можно проинтерпретировать следующим образом. Некоторое устройство порождает простейший поток поломок. n -ремонтных бригад обслуживают систему так, что первая в очереди бригада устраняет поломку за показательное время со своим параметром, а затем становится в общую очередь из свободных бригад.

Будем снова исследовать предельное поведение момента Ω_ε первой потери требования. Пусть $\eta_\varepsilon^{(i)}(t)$ — обозначает номер прибора, занимающего i -ю позицию в момент t , а $\delta_\varepsilon^{(i)}(t) = 0$, если этот прибор свободен и 1, если он занят. Тогда поведение нашей системы описывается многомерным процессом

$$(\bar{\eta}_\varepsilon(t))_{s_\varepsilon(t)} = (\eta_\varepsilon^{(1)}(t), \dots, \eta_\varepsilon^{(n)}(t))_{s_\varepsilon(t)}, \quad t \geq 0,$$

где $s_\varepsilon(t) = \sum_{i=1}^n \delta_\varepsilon^{(i)}(t)$ — количество занятых приборов (которые по построению всегда расположены на первых пози-

циях, т. е. если $s_s(t) = k$, то $\delta_s^{(1)}(t) = \dots = \delta_s^{(k)}(t) = 1$, $\delta_s^{(k+1)} = \dots = \delta_s^{(n)}(t) = 0$. Введенный процесс будет марковским со следующими инфинитезимальными характеристиками:

$$(i_1, \dots, i_n)_s \xrightarrow{\Delta} \begin{cases} (i_1, \dots, i_n)_{s+1} \text{ с вероятностью } \lambda \Delta + o(\Delta); \\ (i_1, \dots, i_{k-1}, i_{k+1}, \dots, i_n, i_k)_{s-1} \text{ при } k \leq s \\ \quad \text{с вероятностью } \frac{1}{\varepsilon} c_{ik} \Delta + o(\Delta); \\ (i_1, \dots, i_n)_s \text{ с вероятностью } (1 - \\ \quad - (\lambda + \frac{1}{\varepsilon} \sum_{j=1}^s c_{ij})) \Delta + o(\Delta). \end{cases} \quad (6.27)$$

При этом, если $s = n$, то будем считать, что из состояния $(i_1, \dots, i_n)_n$ с интенсивностью λ возможен переход в состояние, которое обозначим $\{n+1\}$. Тем самым момент первого попадания $(\bar{\eta}_s(t))_{s_s(t)}$ в $\{n+1\}$ представляет собой по построению момент первой потери требования (так как нас интересует поведение данного момента, то можно не определять поведение процесса после попадания в $\{n+1\}$).

Сложность прямого расчета данной системы связана с тем, что количество всех возможных состояний процесса $(\bar{\eta}_s(t))_{s_s(t)}$ есть число $(n+1)! + 1$. Поэтому естественно воспользоваться методами асимптотического укрупнения пространства состояний. В дальнейшем величину $s_s(t)$ будем называть уровнем системы. Пусть $\bar{\xi}_s(k)$, $k \geq 0$ — вложенная цепь Маркова для процесса $(\bar{\eta}_s(t))_{s_s(t)}$. Тогда из соотношений (6.27) непосредственно вытекает, что в состояниях 0-го уровня процесс проводит показательное время с параметром λ , а в других состояниях с параметром $\lambda + \frac{1}{\varepsilon} \sum_{k=1}^s c_{ik} \rightarrow \infty$ при $s > 1$. При этом вероятности переходов с уровня s на уровень $s+1$ есть величины порядка $O(\varepsilon)$ при $s \geq 1$, а суммарная вероятность перехода с s -го на $(s-1)$ -й уровень стремится к единице.

Пусть начальным было некоторое состояние $(i_1, \dots, i_n)_0$. Тогда легко видеть, что с вероятностью порядка $1 - O(\varepsilon)$ цепь $\bar{\xi}_s(k)$ последовательно перемещается по цепочке из $2n$ состояний, получающейся циклическим сдвигом вида

$$\begin{aligned} (i_1, \dots, i_n)_0 &\rightarrow (i_1, \dots, i_n)_1 \rightarrow (i_2, \dots, i_n, i_1)_0 \rightarrow \\ &\rightarrow (i_2, \dots, i_n, i_1)_1 \rightarrow (i_3, \dots, i_n, i_1, i_2)_0 \rightarrow \dots \rightarrow \\ &\rightarrow (i_n, i_1, \dots, i_{n-1})_1 \rightarrow (i_1, \dots, i_n)_0. \end{aligned}$$

Обозначим множество указанных состояний через $K(i_1, \dots, i_n)$. Поскольку два множества $K(i_1, \dots, i_n)$ и $K(i'_1, \dots, i'_n)$ совпадают, если вектора $(i_1, \dots, i_n)$ и $(i'_1, \dots, i'_n)$ отличаются только циклической перестановкой компонент, то для определенности будем рассматривать множества вида $K(1, i_2, \dots, i_n)$, где $(i_2, \dots, i_n)$ — всевозможные перестановки чисел $(2, \dots, n)$. Понятно, что множества $K(1, i_2, \dots, i_n)$ будут в пределе представлять собой классы существенных состояний для цепи $\bar{\xi}_0(k)$ (предельной для $\bar{\xi}_\varepsilon(k)$), а остальные состояния будут несущественными. Так как на каждом классе цепь $\bar{\xi}_0(k)$ с вероятностью единица циклически обходит все состояния, т. е. она периодична с периодом $2n$, то стационарное распределение на каждом классе будет равновероятным, т. е. стационарные вероятности каждого состояния равны $1/2n$.

Исследуем поведение стационарного распределения цепи $\bar{\xi}_\varepsilon(k)$ на состояниях из классов K . Из соотношений (6.27) следует, что цепь $\bar{\xi}_\varepsilon(k)$ из состояния $(i_1, \dots, i_n)_1$ с вероятностью $\varepsilon \frac{\lambda}{c_{i_1}} + o(\varepsilon)$ перейдет в состояние $(i_1, \dots, i_n)_2$, а с вероятностью $1 - \varepsilon \frac{\lambda}{c_{i_1}} + o(\varepsilon)$ — в состояние $(i_2, \dots, i_n, i_1)_0$. Далее из состояния $(i_1, \dots, i_n)_2$ с вероятностью $\varepsilon \frac{\lambda}{c_{i_1} + c_{i_2}} + o(\varepsilon)$ возможен переход в $(i_1, \dots, i_n)_3$, а с вероятностями $\frac{c_{i_1}}{c_{i_1} + c_{i_2}} + o(1)$ и $\frac{c_{i_2}}{c_{i_1} + c_{i_2}} + o(1)$ — в состояния $(i_2, \dots, i_n, i_1)_1$ и $(i_1, i_3, \dots, i_n, i_2)_1$ соответственно. При этом заметим, что состояние $(i_2, \dots, i_n, i_1)_1$ принадлежит классу $K(i_1, i_2, \dots, i_n)$, а $(i_1, i_3, \dots, i_n, i_2)_1$ — классу $K(i_1, i_3, \dots, i_n, i_2)$, которые различны при $n \geq 3$. Так как $c_i > 0$, $i = 1, n$, получим, что через 2-й уровень цепь может переходить из класса и при этом множество состояний вида $\{(i_1, \dots, i_n)_k, k = 0, 1, 2\}$, где $(i_1, \dots, i_n)$ — возможные перестановки чисел $(1, \dots, n)$, образует α -множество $\langle \alpha \rangle_2$ порядка 2. Согласно алгоритму, приведенному в п. 5.3, на 1-м шаге укрупним каждый класс $K(1, i_2, \dots, i_n)$ до одного состояния $\{1, i_2, \dots, i_n\}$ со стационарными вероятностями переходов в состояния $(j_1, \dots, j_n)_2$ 2-го уровня

$$P_\varepsilon(\langle 1, i_2, \dots, i_n \rangle, \langle j_1, \dots, j_n \rangle_2) = \varepsilon \frac{\lambda}{2nc_{i_1}} + o(\varepsilon),$$

где $(j_1, \dots, j_n)$ — циклическая перестановка чисел $(1, i_2, \dots, i_n)$, а

$$g_e(\langle 1, i_2, \dots, i_n \rangle) = \varepsilon \lambda / 2n d,$$

где $d = \sum_{k=1}^n 1/c_k$.

Теперь на 2-м шаге построим укрупненную вложенную цепь $\xi_e(k)$ вероятности переходов которой есть вероятности «правильных» переходов и, согласно формулам для величин $u_e(\langle d_l \rangle, \langle d_s \rangle)$ в п. 5.3, в нашем случае имеют вид:

$$u_e(\langle 1, i_2, \dots, i_n \rangle, (j_1, \dots, j_n)_2) = \frac{1}{c_{j_1} d} + o(1), \quad (6.28)$$

$$a \quad u_e((j_1, \dots, j_n)_2, \{\langle 1, i_2, \dots, i_n \rangle\}) = \frac{c_{j_1}}{c_{j_1} + c_{j_2}} + o(1);$$

$$u_e((j_1, \dots, j_n)_2, \{\langle j_1, j_2, \dots, j_n, j_2 \rangle\}) = \frac{c_{j_2}}{c_{j_1} + c_{j_2}} + o(1),$$

где $(j_1, \dots, j_n)$ — циклическая перестановка чисел $(1, i_2, \dots, i_n)$.

Предположим, что все $(n-1)!$ перестановок $(1, i_2, \dots, i_n)$ занумерованы каким-либо образом и K_l — l -й класс, соответствующий l -й перестановке. Для упрощения вычислений стационарного распределения $\xi_e(k)$ воспользуемся следующим приемом. Пусть $\tilde{\xi}_e^*(k)$ — новая цепь Маркова с множеством состояний $\{K_l\}$, $j = \overline{1, (n-1)!}$, значения которой есть значения цепи $\xi_e(l)$ в последовательные моменты попадания в состояния типа $\{K_j\}$. Из формул (5.36), (5.37) непосредственно вытекает, что стационарная вероятность любого из $2n$ состояний класса K_l внутри нашего s -множества $\langle \alpha \rangle_2$ есть величина

$$\frac{1}{2n} \hat{\pi}_e(K_l) g_e(K_l)^{-1} \left(\sum_{l=1}^{(n-1)!} \hat{\pi}_e(K_l) g_e(K_l)^{-1} \right) + \sum_{(\bar{i})_2} \hat{\pi}_e((\bar{i})_2)^{-1}, \quad (6.29)$$

где $(\bar{i})_2$ — состояния 2-го уровня, при этом для них очевидно $g_e((\bar{i})_2) = 1$, а $\hat{\pi}_e(K_l)$ и $\hat{\pi}_e((\bar{i})_2)$ — стационарные вероятности состояний $\{K_l\}$ и $(\bar{i})_2$ соответственно для цепи $\tilde{\xi}_e(k)$. Пусть теперь $\hat{\pi}_e^*(K_l)$, $l = \overline{1, (n-1)!}$ — стационарное распределение для цепи $\tilde{\xi}_e^*(k)$. В силу результатов леммы 5.6

$$\hat{\pi}_e^*(K_l) = \hat{\pi}_e(K_l) q_e^{-1}, \quad l = \overline{1, (n-1)!},$$

где

$$q_e = \sum_{l=1}^{(n-1)!} \hat{\pi}_e(K_l).$$

Отсюда и из формулы (6.29), поскольку $g_e(K_l) = O(e)$, получим для стационарной вероятности любого из $2n$ состояний класса K_j выражение

$$\frac{1}{2n} \hat{\pi}_e^*(K_j) g_e(K_j)^{-1} \left(\sum_{l=1}^{(n-1)!} \hat{\pi}_e^*(K_l) g_e(K_l)^{-1} \right)^{-1} + o(1). \quad (6.30)$$

Тем самым достаточно вычислить стационарное распределение цепи $\hat{\xi}_e^*(k)$. Пусть $K_j = K(i_1, \dots, i_n)$. Тогда вероятности переходов цепи $\hat{\xi}_e^*(k)$ вычисляются по формуле

$$u_e(K_j, K_l) = \sum_{(\bar{i})_1} u_e(K_j, (\bar{i})_2) u_e(\bar{i})_2, K_l),$$

и, согласно (6.28), имеют вид

$$u_e(K_j, K_j) = \frac{1}{d} \sum_{m=1}^n \frac{1}{c_{i_m} + c_{i_{m+1}}} + o(1), \quad (6.31)$$

а

$$u_e(K_j, K_l) = \frac{1}{d} \frac{c_{i_{r+1}}}{c_{i_r}} \cdot \frac{1}{c_{i_r} + c_{i_{r+1}}} + o(1),$$

если класс $K_l = K(i_1, \dots, i_{r-1}, i_{r+1}, i_r, \dots, i_n)$, где полагаем $i_{n+1} = i_1$. При этом возможны переходы только в такие классы, которые отличаются от K_j перестановкой каких-либо двух соседних приборов. Таким образом, количество классов, из которых можно попасть для цепи $\hat{\xi}_e^*(k)$ в K_j за один шаг равно n , и они имеют вид $K(i_1, \dots, i_{r-1}, i_{r+1}, i_r, \dots, i_n)$, $r = \overline{1, n}$. Из формулы (6.31) вероятность перехода из такого класса в K_j есть величина

$$\frac{1}{d} \frac{c_{i_r}}{c_{i_{r+1}}} \cdot \frac{1}{c_{i_r} + c_{i_{r+1}}} + o(1).$$

Запишем общую систему уравнений для стационарного распределения предельной для $\hat{\xi}_e^*(k)$ цепи $\hat{\xi}_0^*(k)$

$$\begin{aligned} \hat{\pi}_0^*(K_j) &= \hat{\pi}_0^*(K_j) \frac{1}{d} \sum_{m=1}^n \frac{1}{c_{i_m} + c_{i_{m+1}}} + \\ &+ \sum_{r=1}^n \hat{\pi}_0^*(K(i_1, \dots, i_{r-1}, i_{r+1}, i_r, \dots, i_n)) \frac{1}{d} \frac{c_{i_r}}{c_{i_{r+1}}} \times \\ &\times \frac{1}{c_{i_r} + c_{i_{r+1}}}. \end{aligned} \quad (6.32)$$

Покажем, что $\hat{\pi}_0^*(K_j) \equiv \pi$, $j = \overline{1, (n-1)!}$ является решением системы. Подставляя в правую часть (6.32) вместо величин $\hat{\pi}_0^*(K_j)$ π и учитывая, что $d = \sum_{k=1}^n \frac{1}{c_k}$, получим

$$\frac{\pi}{d} \left(\sum_{m=1}^n \frac{1}{c_{t_m} + c_{t_{m+1}}} + \sum_{r=1}^n \frac{c_{t_r}}{c_{t_r} + c_{t_{r+1}}} \frac{1}{c_{t_r} + c_{t_{r+1}}} \right) = \frac{\pi}{d} \sum_{r=1}^n \frac{1}{c_{t_{r+1}}} = \pi.$$

Используя единственность стационарного распределения и условие нормировки $\sum_{j=1}^{(n-1)!} \hat{\pi}_0^*(K_j) = 1$, окончательно получим, что

$$\hat{\pi}_0^*(K_j) \equiv 1/(n-1)!, \quad j = \overline{1, (n-1)!}$$

Так как все $g_\varepsilon(K_j)$ одинаковы, то из формулы (6.30) следует, что распределение на всех состояниях 0-го и 1-го уровней s -множества $\langle \alpha \rangle_s$ асимптотически равновероятное и стремится к $1/2n!$.

Теперь можно последовательно вычислять главные части стационарных вероятностей уровней $s \geq 2$ в точности подобно тому, как это делалось в примере п.5.3. Поскольку для исходной цепи $\tilde{\xi}_\varepsilon(k)$ вероятности переходов с s -го на $(s+1)$ -й уровень, как следует из соотношений (6.27), имеют вид

$$\begin{aligned} p_\varepsilon((i_1, \dots, i_n), (i_1, \dots, i_n)_{s+1}) = \\ = \varepsilon \lambda \left(\sum_{j=1}^s c_{t_j} \right)^{-1} + o(\varepsilon) \end{aligned}$$

при $1 \leq s \leq n$, где $(i_1, \dots, i_n)_{n+1} = \{n+1\}$, то из формулы (5.48) и сходимости стационарных вероятностей состояний 0-го и 1-го уровней к $1/2n!$ для цепи $\tilde{\xi}_\varepsilon(k)$ последовательно получим, что

$$\pi_\varepsilon((i_1, \dots, i_n)_{s+1}) = \frac{\varepsilon^s \lambda^s}{2n!} \prod_{k=1}^s \left(\sum_{j=1}^k c_{t_j} \right)^{-1}.$$

При этом множество всех состояний $\{(i_1, \dots, i_n)_s, s = \overline{0, n}\}$, где $(i_1, \dots, i_n)$ — все возможные перестановки чисел $(1, \dots, n)$, образует s -множество $\langle \alpha \rangle_n$ порядка n , и для стационарной вероятности перехода $g_\varepsilon(\langle \alpha \rangle_n)$ в отказовое состояние $\{n+1\}$ имеет место, согласно (5.49), формула

$$g_\varepsilon(\langle \alpha \rangle_n) = \frac{\varepsilon^n \lambda^n}{n!} \sum_{(i_1, \dots, i_n)} \prod_{k=1}^n \left(\sum_{j=1}^k c_{t_j} \right)^{-1},$$

где сумма берется по всем возможным перестановкам.

Докажем по индукции, что

$$\sum_{(i_1, \dots, i_n)} \prod_{k=1}^n (\sum_{j=1}^k c_{i_j})^{-1} = (\prod_{k=1}^n c_k)^{-1}.$$

Для $n = 1$ утверждение очевидно. Далее

$$\begin{aligned} & \sum_{(i_1, \dots, i_{n+1})} \prod_{k=1}^{n+1} (\sum_{j=1}^k c_{i_j})^{-1} = \\ &= (\sum_{j=1}^{n+1} c_j)^{-1} \sum_{(i_1, \dots, i_n) \in (1, \dots, n+1) \setminus \{i_{n+1}\}} \prod_{k=1}^n (\sum_{j=1}^k c_{i_j})^{-1} = \\ &= (\sum_{j=1}^{n+1} c_j)^{-1} \sum_{i_{n+1}=1}^{n+1} (\prod_{j=1}^n c_{i_j})^{-1} = \\ &= (\prod_{j=1}^{n+1} c_{i_j})^{-1} (\sum_{j=1}^{n+1} c_j)^{-1} \sum_{i_{n+1}=1}^{n+1} c_{i_{n+1}} = (\prod_{k=1}^{n+1} c_k)^{-1}. \end{aligned}$$

Окончательно получим

$$g_\varepsilon(\langle \alpha \rangle) = \varepsilon^n \frac{\lambda^n}{2n!} \left(\prod_{k=1}^n c \right)^{-1}. \quad (6.33)$$

Пусть $\tau_\varepsilon((i_1, \dots, i_n)_s)$ — обозначает время сидения процесса $(\bar{\eta}_\varepsilon(t))_{s_\varepsilon(t)}$ в состоянии $(i_1, \dots, i_n)_s$. Тогда из соотношений (3.8) вытекает, что это время имеет показательное распределение с параметром $\lambda_\varepsilon((i_1, \dots, i_n)_s) =$

$= \lambda + \frac{1}{\varepsilon} \sum_{j=1}^n c_{i_j}$, который стремится к ∞ при $s \geq 1$. Отсюда вытекает следующее соотношение:

$$M \exp \{ -\theta \varepsilon^n \tau_\varepsilon((i_1, \dots, i_n)_s) \} = 1 - \frac{\varepsilon^n \theta}{\lambda} \kappa_0(s) + o(\varepsilon^n), \quad (6.34)$$

где $\kappa_0(s) = 1$, $s = 0$, и 0 при $s > 0$.

Учитывая, что состояний 0-го уровня $n!$ и их суммарная стационарная вероятность $n!$. $\frac{1}{2n!} = \frac{1}{2}$ и используя следствие 5.6, окончательно получим следующую теорему.

Теорема 6.4. Для системы $M|M|n$ с переменным порядком обслуживания при наших обозначениях предельное распределение величины $\varepsilon^n \Omega_\varepsilon$ (момента первой потери требования) при $\varepsilon \rightarrow 0$ будет показательным с параметром

$$\frac{\lambda^{n+1}}{n!} \left(\prod_{k=1}^n c_k \right)^{-1}.$$

Следуя результатам п.5.4 (п.3), получим, что при наших предположениях поток потерянных требований $v_e(e^{-nt})$ сходится к простейшему потоку с вышеуказанным параметром.

Заметим, что аналогичное утверждение имеет место для системы типа $M_n | M | n$ с переменным порядком обслуживания, на вход которой поступает пуассоновский поток, параметр которого управляется конечной в пределе неприводимой цепью Маркова. Это можно доказать самостоятельно и получить формулу для соответствующего параметра предельного закона.

6.4. Поведение аддитивных функционалов для асимптотически укрупняемых систем

В пп. 6.2, 6.3 изучалась сходимость укрупненных процессов к марковским и полумарковским процессам и рассматривались применения к анализу конкретных СМО с быстрым обслуживанием. Вместе с тем во многих реальных задачах возникают вопросы, связанные с поведением аддитивных функционалов от таких систем. Общая модель следующая. Пусть задан двумерный процесс $(x_e(t), \zeta_e(t))$, $t \geq 0$, со значениями в некотором множестве состояний (X, R^m) , где $\zeta_e(t)$ строится некоторым образом по траекториям $x_e(t)$ (к примеру $\zeta_e(t) = \int_0^t \varphi_e(u, x_e(u)) du$, где $\varphi_e(u, x)$ — некоторая возможно случайная функция). Рассмотрим процесс $(\hat{x}_e(t), \hat{\zeta}_e(t))$, $t \geq 0$, где $\hat{x}_e(t)$ — укрупнение процесса $x_e(t)$ по некоторому разбиению множества X . Если процесс $\hat{x}_e(t)$ асимптотически сходится к марковскому процессу, а приращения $\hat{\zeta}_e(t)$ слабо зависят от предыстории $x_e(t)$, то при весьма общих предположениях в качестве предельных для $(\hat{x}_e(t), \hat{\zeta}_e(t))$ выступает класс МП ОВК, описанный в п.6.1.

Общий подход к получению подобных предельных теорем был предложен в [2, 3] и использует аппарат переключающихся процессов (введенных в п.6.1), а также предельные теоремы для процессов накопления на слабозависимых процессах. Заметим, что данный подход позволяет также решать задачи асимптотического укрупнения неоднородных по времени марковских процессов и ПМП и в этом направлении продвинуться далее по сравнению с работами [10, 11], в которых изучались модели асимптотического укрупнения только однородных по времени процессов.

Рассмотрим наиболее простые схемы, иллюстрирующие общие результаты, поскольку исследование общих моделей достаточно сложно и связано с привлечением предельных теорем для функционалов от случайных процессов и метода одного вероятностного пространства. Общие схемы с подробными доказательствами приведены в [2, 3].

п.1. Пусть задана неоднородная цепь Маркова $x_e(t)$, $t \geq 0$, с конечным множеством состояний I и мгновенными интенсивностями вероятностей переходов $\lambda_e(i, j, t)$, $i, j \in I$, $i \neq j$, $t \geq 0$. Предположим, что задано разбиение I на непересекающиеся подмножества $I_1, \dots, I_r$, $r > 1$, и

$$\lambda_e(i, j, t) = \begin{cases} \lambda_{ij}^{(k)} + o(1) & \text{при } i, j \in I_k, i \neq j, \\ \alpha_e(i, j, t) & \text{при } i \in I_k, j \in \bar{I}_k, k = \overline{1, r}, \end{cases} \quad (6.35)$$

где $\sup_{e \leq T} |o(1)| \rightarrow 0$ и $\sup_{e \leq T} \alpha_e(i, j, t) \rightarrow 0$ для любого $T > 0$.

Условие (6.35) означает, что на каждом подмножестве I_k интенсивности вероятностей переходов равномерно сходятся к интенсивностям вероятностей переходов некоторой однородной цепи, а интенсивности вероятностей переходов между подмножествами равномерно малы. Обозначим через $x^{(k)}(t)$ однородную цепь Маркова с множеством состояний I_k и интенсивностями вероятностей переходов $\lambda_{ij}^{(k)}$, $i, j \in I_k$, $i \neq j$. Предположим, что каждая цепь $x^{(k)}(t)$ неприводима и ее стационарное распределение обозначим $\pi_i^{(k)}$, $i \in I_k$.

Далее пусть задана совокупность неслучайных интегрируемых функций $\varphi_e(i, t)$, $i \in I$, $t \geq 0$. Введем процесс

$$\xi_e(t) = \int_0^t \varphi_e(x_e(u), u) du. \quad (6.36)$$

Заметим, что $\varphi_e(i, t)$ можно интерпретировать как локальный доход системы в i -м состоянии в момент t , а $\xi_e(t)$ — суммарный доход, полученный на промежутке $[0, t]$.

Введем следующие условия:

$$\lim_{e \rightarrow 0} \int_0^{e^{-1}t} \alpha_e(t, j, u) du = \beta_{ij}(t), \quad (6.37)$$

где $\beta_{ij}(t)$ — непрерывны, $i \in I_k$, $j \in \bar{I}_k$, $k = \overline{1, r}$, и

$$\lim_{e \rightarrow 0} \int_0^{e^{-1}t} \varphi_e(i, u) du = f_i(t), \quad i \in I, t > 0, \quad (6.38)$$

где $\sup_{eu \leq t} |\varphi_e(i, u)| \rightarrow 0$ для любого $t > 0$.

Обозначим

$$\begin{aligned}\Lambda_{km}(t) &= \sum_{i \in I_k} \sum_{j \in I_m} \pi_i^{(k)} \beta_{ij}(t), \quad k \neq m, \\ f^{(k)}(t) &= \sum_{i \in I_k} \pi_i^{(k)} f_i(t), \quad k = \overline{1, r}.\end{aligned}\quad (6.39)$$

Через $\hat{x}_0(t)$ обозначим цепь Маркова с множеством состояний $\{1, \dots, r\}$, у которой вероятности перехода из k в m на промежутке $(t, t + \Delta t)$ есть величины $(\Lambda_{km}(t + \Delta) - \Lambda_{km}(t))(1 + o(1))$, $k \neq m$ (если, к примеру, $\Lambda_{km}(t)$ дифференцируемы и $\hat{\lambda}_{km}(t) = \Lambda_{km}(t)$, то тогда это цепь с мгновенными интенсивностями вероятностей переходов $\hat{\lambda}_{km}(t)$). Введем укрупненный процесс

$$\hat{x}_e(t) = k, \text{ если } x_e(t) \in I_k, \quad k = \overline{1, r}, \quad t \geq 0.$$

Теорема 6.5. Если выполнены условия (6.35), (6.37)

$$\lim_{e \rightarrow 0} P\{x_e(0) \in I_k\} = \rho_k, \quad k = \overline{1, r}, \quad (6.40)$$

то конечномерные распределения процесса $\hat{x}_e(e^{-1}t)$ слабо сходятся к распределениям цепи $\hat{x}_0(t)$ с начальным распределением ρ_k , $k = \overline{1, r}$.

Если дополнительно выполнено (6.38), то распределения процесса $(\hat{x}_e(e^{-1}t), \xi_e(e^{-1}t))$ слабо сходятся к МП ОВК $(x_0(t), \xi_0(t))$.

Здесь $\xi_0(t)$ задается функциями $f^{(k)}(t)$ и $\xi_0(0) = 0$, а далее, если $\hat{x}_0(t) = k$, $t \in [u, s]$, то

$$\xi_0(s) - \xi_0(u) = f^{(k)}(s) - f^{(k)}(u).$$

Доказательство. Воспользуемся вначале схемой, приведенной в примере 6.3, и построим некоторый ПП, описывающий поведение $(\hat{x}_e(t), \xi_e(t))$. Пусть

$$\{x_e(t, u, k, i), (\xi_e(t, u, k, i), \chi_e(t, u, k, i))\},$$

$i \in I_k$, $k = \overline{1, r}$, $t \geq u \geq 0$ — вспомогательные МП ОВК, которые строятся следующим образом: $x_e(t, u, k, i)$, $t \geq u$ — цепь Маркова с множеством состояний I_k , начальным состоянием $x_e(u, u, k, i) = i \in I_k$ и мгновенными интенсивностями вероятностей переходов $\lambda_e(i, j, t)$, $i, j \in I_k$, $i \neq j$. Здесь

$$\lambda_e(i, i, t) = \sum_{j \in I_k, j \neq i} \lambda_e(i, j, t);$$

$\chi_e(t, u, k, i)$ строится на $x_e(t, u, k, i)$ как обобщенный пуассоновский процесс, управляемый $x_e(\cdot)$, при

этом $\chi_e(t, u, k, i)$ принимает значения во множестве $\{0\} \cup I \setminus I_k$ (для определенности полагаем, что $\{0\} \notin I$) и мгновенная интенсивность вероятности того, что величина скачка $\chi_e(t, u, k, i)$ примет значение $j \in I_k$, если $x_e(t, u, k, i) = l \in I_k$, есть $\alpha_e(l, j, t)$; процесс $\xi_e(t, u, k, i)$ строится по функциям $\varphi_e(i, t)$ так, что $\xi_e(u, u, k, i) = 0$ и далее если $x_e(t, u, k, i) = j \in I_k$ на промежутке $v \leq t \leq s$, то

$$\xi_e(s, u, k, i) - \xi_e(v, u, k, i) = \int_v^s \varphi_e(j, y) dy.$$

Тогда, как показано в примере 6.3, построенный по семействам $(x_e(\cdot), (\xi_e(\cdot), \chi_e(\cdot)))$, где моменты переключения определяются моментами скачков $\chi_e(\cdot)$, ПП будет эквивалентен по распределению процессу $(\hat{x}_e(t), \xi_e(t))$.

Далее заметим, что при наших предположениях для любого $k = \overline{1, r}$, $i \in I_k$, $u > 0$, цепь $x_e(t, u, k, i)$, $t \geq u$ — в пределе эргодическая, поскольку ее интенсивности равномерно сходятся к интенсивностям однородной неприводимой цепи. Тем самым цепь удовлетворяет условию равномерно сильного перемешивания и для любых $i, j \in I_k$

$$\lim_{t \rightarrow \infty} \overline{\lim}_{\varepsilon \rightarrow 0} \sup_{u > 0} |P\{x_e(t+u, u, k, i) = j\} - \pi_j^{(k)}| = 0. \quad (6.41)$$

Теперь по построению

$$\xi_e(t, u, k, i) = \int_u^t \varphi_e(x_e(v, u, k, i), v) dv.$$

Воспользовавшись условиями сходимости интегральных функционалов от процессов, удовлетворяющих условию равномерно сильного перемешивания, приведенными в п.5.6, получим, что при выполнении условий (6.38), (6.41) для любых $i \in I_k$, $k = \overline{1, r}$, $t > u$

$$\xi_e(t/\varepsilon, u/\varepsilon, k, i) \xrightarrow{P} f^{(k)}(t) - f^{(k)}(u). \quad (6.42)$$

Далее по аналогии с построениями, приведенными в п. 5.6, получим, что распределения процесса $x_e(t/\varepsilon, u/\varepsilon, k, i)$ независимо от $i \in I_k$ сходятся к распределениям обобщенного пуассоновского процесса, у которого вероятность того, что величина скачка на промежутке $[s, s + \Delta]$ примет значение в множестве I_m , есть $(\Lambda_{km}(s + \Delta) - \Lambda_{km}(s)) \times (1 + o(1))$.

Таким образом, если $x_e(0) = i \in I_k$, то распределение момента первого скачка $\mu_e(1)$ процесса $\hat{x}_e(e^{-1}t)$ совпадает с моментом скачка $\chi_e(t/e, 0, k, i)$ и сходится к моменту первого скачка μ_1 цепи $\hat{x}_0(t)$, у которой $\hat{x}_0(0) = k$. При этом так как $\Lambda_{km}(t)$ непрерывны, то с вероятностью единица момент τ_1 попадает в множество точек непрерывности функции $f^{(k)}(t)$, и тогда из (6.42) следует, что доход, накопленный на промежутке $[0, \mu_e(1)]$, сходится (по распределению) к величине $f^{(k)}(\tau_1) - f^{(k)}(0) = f^{(k)}(\tau_1)$ (строгое обоснование этого факта следует из результатов о сходимости суперпозиций случайных функций в точке непрерывности [2]). Далее распределения значений $\hat{x}_e(t)$ после скачка сходятся к распределениям $\hat{x}_0(\tau_1 + 0)$ и затем, если попадание $\hat{x}_e(\cdot)$ произошло в m -е состояние, прошлое асимптотически забывается и процесса $\hat{x}_e(\cdot)$ развивается совместно с процессом накопления дохода до момента следующего скачка и т. д. Поэтому по индукции легко следует сходимость совместных распределений вектора значений $\hat{x}_e(\cdot)$ в моменты скачков и самих моментов к распределениям соответствующего вектора цепи $\hat{x}_0(\cdot)$. Доказательство этого может быть проведено аналогично доказательству леммы 6.2. А отсюда, подобно доказательству теоремы 6.2, можно получить наше утверждение. Заметим лишь здесь, что регулярность цепи $\hat{x}_0(t)$ следует из того, что множество состояний конечно, а величины $\Lambda_{km}(t)$ ограничены при каждом t . Поэтому если обозначим $v_0(t)$ — число скачков $\hat{x}_0(\cdot)$ на промежутке $[0, t]$, то легко получим оценку

$$P\{v_0(t) \geq k\} \leq P\{\tilde{\Pi}(t) \geq k\},$$

где $\tilde{\Pi}(t)$ — величина, имеющая пуассоновское распределение с параметром $\max_{\substack{k, m=1, r \\ k \neq m}} \Lambda_{km}(t)$, откуда следует, что $v_0(t)$

конечна с вероятностью единица при любом $t \geq 0$.

п.2. Сформулируем аналогичный результат для процессов с дискретным временем. Поскольку доказательство использует конструкцию ПП, приведенную в примере 6.4, и проводится в основном подобно доказательству теоремы 6.5, оно будет опущено.

Пусть $x_e(k)$, $k = 0, 1, \dots$ — неоднородная цепь Маркова с конечным множеством состояний и вероятностями переходов на k -м шаге $p_e(i, j, k)$, $i, j \in I$, $k \geq 0$. Предположим, что задано разбиение I на непересекающиеся подмножества $I_1, \dots, I_r$, $r > 1$, и

$$p_e(i, j, k) = \begin{cases} p_{ij}^{(k)} + o(1) & \text{при } i, j \in I_k; \\ \alpha_e(i, j, k) & \text{при } i \in I_k, j \in \overline{I_k}, \\ & k = \overline{1, r} \end{cases} \quad (6.43)$$

где $\max_{e, k < T} |o(1)| \rightarrow 0$; $\max_{e, k < T} \alpha_e(i, j, k) \rightarrow 0$ для любого $T > 0$.

Далее предположим, что каждой матрице $\|p_{ij}^{(k)}\|$, $i, j \in I_k$ соответствует неприводимая непериодическая цепь, стационарное распределение которой обозначим $\pi_i^{(k)}$, $i \in I_k$.

Пусть задана совокупность неслучайных констант $\varphi_e(i, k)$ $i \in I$, $k \geq 0$. Введем процесс типа доходов системы

$$\xi_e(m) = \sum_{k=0}^m \varphi_e(x_e(k), k).$$

Предположим, что

$$\lim_{e \rightarrow 0} \sum_{k=0}^{[e^{-1}t]} \alpha_e(i, j, k) = \beta_{ij}(t). \quad (6.44)$$

где $\beta_{ij}(t)$ — непрерывны, $i \in I_k$, $j \in \overline{I_k}$, $k = \overline{1, r}$ и

$$\lim_{e \rightarrow 0} \sum_{k=0}^{[e^{-1}t]} \varphi_e(i, k) = f_i(t), i \in I, \quad (6.45)$$

где $\max_{e, k < t} |\varphi_e(i, k)| \rightarrow 0$ для любого $t > 0$.

Построим укрупненный процесс

$$\hat{x}_e(l) = k, \text{ если } x_e(l) \in I_k, k = \overline{1, r}, l \geq 0.$$

Теорема 6.6. Если выполнены условия (6.40), (6.43), (6.44), то в обозначениях (6.39) конечномерные распределения процесса $\hat{x}_e([e^{-1}t])$ слабо сходятся к распределениям цепи $\hat{x}_0(t)$. Если дополнительно выполнено (6.45), то и распределения $(\hat{x}_e([e^{-1}t]), \xi_e([e^{-1}t]))$ слабо сходятся к МП ОВК $(\hat{x}_0(t), \xi_0(t))$.

Здесь цепь $\hat{x}_0(t)$ и процесс $\xi_0(t)$ строятся так, как указывалось в теореме 6.5.

Подобные модели могут возникать при исследовании асимптотически укрупняемых СМО, подобных рассмотренным в п. 6.3, с доходами (т. е. к примеру за единицу времени обслуживания i -м прибором j -го требования взимается плата c_{ij} , за изменение порядка приборов может взиматься своя плата и т. п.). Ряд результатов для таких систем читатель может получить самостоятельно, используя результаты п. 6.3.

В заключение заметим, что подобные схемы возникают в значительно более общих ситуациях при исследовании стохастических систем с доходами, множество состояний

которых можно разбить на непересекающиеся области так, что вероятности переходов между областями могут быть некоторыми функционалами от траектории процесса и в определенном смысле «асимптотически» малы. Для реальных систем переходы между областями могут быть интерпретированы как изменения режима работы, вызванные, к примеру, какой-либо неисправностью. При этом, поскольку у высоконадежных систем соответствующие вероятности достаточно малы, естественно исследовать асимптотическое поведение таких систем. Тогда, если поведение системы в каждой области описывается процессом, обладающим «асимптотическим забыванием прошлого», а доход имеет вид интегрального функционала от траектории системы, то при весьма широких условиях в качестве предельных для $(\tilde{x}_\varepsilon(t), \tilde{z}_\varepsilon(t))$ выступает класс МП ОВК.

При этом возникает ряд важных для теории и практики проблем, одна из которых связана с получением явных оценок скорости сходимости к предельным процессам, а вторая — с построением при фиксированном достаточно малом $\varepsilon > 0$ оптимального разбиения с заданным числом классов, которое в каком-то смысле обладало бы наименьшей потерей информации об исходной системе, т. е. предельный укрупненный по данному разбиению процесс максимально был бы близок к исходному процессу. Эти проблемы до настоящего времени еще не получили окончательного решения и ждут пытливых исследователей.

ГЛАВА 7. ВЕРОЯТНОСТНОЕ МОДЕЛИРОВАНИЕ

7.1. Моделирование основных распределений

При имитационном моделировании основной задачей является получение последовательностей действительных чисел с заданными статистическими характеристиками. Чаще всего такие последовательности должны быть реализациями последовательностей независимых одинаково распределенных случайных величин. Например, при моделировании простейшего пуассоновского потока необходимо получать длины промежутков времени между поступлениями соседних заявок. Длины этих промежутков являются независимыми между собой одинаково распределенными по показательному закону случайными величинами. Значит, при имитации входного потока необходимо иметь устройство, которое бы выдавало положительные числа, которые

можно было бы интерпретировать как значения независимых показательно распределенных случайных величин. Другими, наиболее часто встречающимися в связи с системами массового обслуживания распределениями вероятностей являются: пуассоновское распределение, которое задается вероятностями

$$p_k(\lambda) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad k \geq 0,$$

и распределение Эрланга с параметрами (n, λ) , задаваемое плотностью

$$f(x) = \begin{cases} \frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x}, & x \geq 0; \\ 0, & x < 0. \end{cases} \quad (7.1)$$

Первое из них интерпретируется как распределение числа требований, поступивших за единицу времени при показательном распределении (с параметром λ) промежутка времени между требованиями, второе — как распределение времени поступления n -го требования при тех же предположениях на входной поток.

Как уже отмечалось, при имитационном моделировании на ЭВМ необходимо иметь устройства, которые бы выдавали последовательности чисел с данными статистическими характеристиками, в частности значения независимых одинаково распределенных случайных величин с заданным законом распределения. Такие программно реализуемые устройства называют датчиками случайных чисел. Обычно реализуют датчик равномерного распределения, т. е. распределение с плотностью $f(x) = \begin{cases} 0, & x \notin [0, 1] \\ 1, & x \in [0, 1] \end{cases}$. Как пра-

вило, этот датчик хорошо отработан и его статистические характеристики достаточно хорошо отвечают равномерному распределению. Датчики остальных распределений стараются получить с помощью датчика равномерного распределения. Для дискретных распределений эта возможность вытекает из следующей несложной леммы.

Лемма 7.1. Пусть $(\Omega, \mathfrak{B}_2, P)$ — распределение вероятностей на

$$\Omega = \{x_1, \dots, x_n, \dots\} \quad (P(x_i) = p_i, \sum p_i = 1), \quad (7.2)$$

ξ — равномерно распределенная на $[0, 1]$ случайная величина. Тогда, если последовательность чисел $0 = a_0 < a_1 < \dots \leq 1$, такова, что $a_2 - a_{i-1} = p_i, i \geq 1$, то

случайная величина $\xi = \sum_{i=1}^n x_i \chi_{[a_{i-1}, a_i]}(\xi)$ имеет распределение, совпадающее с [7.2]

Доказательство. Действительно, поскольку $\chi_{[a_{i-1}, a_i]}(x)$ — индикаторная функция и интервалы $[a_i, a_{i-1})$ не пересекаются, то в сумме $\sum_{i=1}^n x_i \chi_{[a_{i-1}, a_i]}(\xi)$ при фиксированном ξ лишь одно слагаемое отлично от нуля, поэтому

$$P\{\xi \in \{x_1, \dots, x_n, \dots\}\} = 1$$

и

$$\begin{aligned} P\{\xi = x_n\} &= P\{\chi_{[a_{n-1}, a_n]}(\xi) = 1\} = \\ &= P\{\xi \in [a_{n-1}, a_n)\} = P\{a_{n-1} \leq \xi < a_n\} = \\ &= a_n - a_{n-1} = p_n. \end{aligned}$$

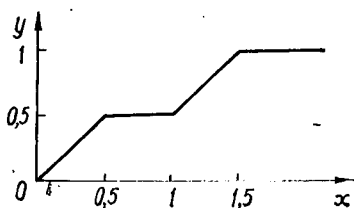


Рис. 7.1.

Из этой леммы вытекает, что для получения очередного значения датчика дискретного распределения с требуемыми характеристиками нужно проверить очередное значение датчика равномерного распределения на принадлежность одному из интервалов $[a_{i-1},$

$a_i)$ и принимать в качестве значения датчика дискретного распределения число x_i с номером того интервала, которому принадлежит значение равномерного датчика.

Другой класс распределений, которые можно получить с помощью равномерного, это те, у которых в области определения функций распределения можно выделить интервалы строгой монотонности со значениями функции, дающими в совокупности отрезок $[0, 1]$. В сущности это класс распределений с непрерывными функциями распределения. Пример такой функции распределения дан на рис. 7.1.

$$F(x) = \begin{cases} 0, & x \leq 0 \\ x, & x \in (0, 0,5]; \\ 0,5, & x \in (0,5, 1]; \\ x - 0,5, & x \in (1, 1,5]; \\ 1, & x \geq 1,5 \end{cases}$$

В области определения F два интервала монотонности $(0, 0,5]$ и $(1, 1,5]$. Нетрудно проверить, что если рассмотреть F на объединении участков строгой монотонности, то такое сужение функции будет строго монотонной функ-

цией между областью определения и областью значений, а поскольку последняя совпадает с $[0, 1]$, то на $[0, 1]$ определена обратная функция, устанавливающая взаимно-однозначное соответствие между отрезком $[0, 1]$ и объединением участков строгой монотонности. Для примера такой обратной будет $\tilde{F}^{-1}$:

$$\tilde{F}^{-1}(x) = \begin{cases} x, & x \in (0, 0.5]; \\ x + 0.5, & x \in (0.5, 1]. \end{cases}$$

Обозначим такой класс распределений через M , а обратные для суженных на интервале строгой монотонности функций $F \in M$ будем обозначать $\tilde{F}^{-1}$. Из определения $\tilde{F}^{-1}$ вытекает, что $F(\tilde{F}^{-1}(x)) = x$, $x \in [0, 1]$.

Справедливо следующее утверждение:

Лемма 7.2. Пусть ξ — случайная величина с $F_\xi \in M$. Тогда $\xi = F_\xi(\zeta)$ есть равномерно распределенная на $[0, 1]$ случайная величина.

Доказательство. Действительно

$$F_\xi(x) = P\{\xi \in (-\infty, x)\} = \begin{cases} 1, & x \geq 1; \\ 0, & x \leq 0. \end{cases}$$

Далее при $x \in (0, 1)$

$$\begin{aligned} F_\xi(x) &= P\{\xi < x\} = P\{F_\zeta(\zeta) < x\} = \\ &= P\{\zeta < F_\zeta^{-1}(x)\} = F_\zeta(F_\zeta^{-1}(x)) = x, \end{aligned}$$

что доказывает наше утверждение. Из этого утверждения непосредственно следует, что случайная величина ζ со значениями в множестве участков строгой монотонности функции распределения $F_\zeta(x)$ и на этих участках имеющая вид $\zeta = \tilde{F}_\zeta^{-1}(\xi)$ (ξ — равномерное на $[0, 1]$ распределение) будет иметь функцию распределения $F_\xi(x)$.

Действительно,

$$P\{\zeta < x\} = P\{\tilde{F}_\zeta^{-1}(\xi) < x\} = P\{\xi < F_\xi(x)\} = F_\xi(x).$$

Как и лемма 7.1 в дискретном случае, лемма 7.2 позволяет моделировать значение независимых случайных величин с распределением $F_\zeta(x)$ по значениям датчика равномерного распределения. Действительно, в обозначениях леммы 7.2 $\zeta = \tilde{F}_\zeta^{-1}(\xi)$. Таким образом, для получения очередного значения случайной величины ζ нужно взять $\tilde{F}_\zeta^{-1}$ от очередного значения случайной величины ξ , которая

равномерно распределена на $[0, 1]$. Так, для показательного распределения с параметром λ

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0; \\ 0, & x < 0, \end{cases}$$

промежуток строгой монотонности есть $[0, \infty)$. Значение обратной функции $\tilde{F}^{-1}(x)$ определяется как единственное значение решения уравнения $1 - e^{-\lambda y} = x$, т. е.

$$\tilde{F}^{-1}(x) = -\frac{\ln(1-x)}{\lambda}, \quad x \in (0, 1). \quad (7.3)$$

Таким образом, случайная величина $\zeta = -\frac{\ln(1-\xi)}{\lambda}$ имеет показательное распределение. Остается только заметить, что случайная величина $1 - \xi$ при равномерном на $[0, 1]$ распределении ξ также имеет равномерное на $[0, 1]$ распределение, т. е. выборку из показательного распределения можно получить из равномерного как с помощью (7.3), так и с помощью (7.4):

$$\zeta = -\frac{\ln \xi}{\lambda}. \quad (7.4)$$

Перейдем к оставшемуся из трех упомянутых распределений — распределению Эрланга, плотность которого задается формулой (7.1). Его можно моделировать, используя лемму 7.2. Однако в отличие от случая показательного распределения, в данном случае это сопряжено с определенными трудностями, связанными с построением функции, обратной к функции распределения. Поэтому выборки из распределения Эрланга получают, используя его интерпретацию с помощью показательного распределения: если $\xi_1, \dots, \xi_n$ — независимые одинаково распределенные по показательному закону с параметром λ случайные величины, то

$$\zeta = \sum_{i=1}^n \xi_i$$

имеет распределение Эрланга с параметрами (n, λ) . Отсюда, если $\xi_1, \dots, \xi_n, \xi_{n+1}, \xi_{n+2}, \dots, \xi_{2n}$ — выборки из показательного распределения с параметром λ , то

$$\zeta_k = \sum_{i=1}^n \xi_{(k-1)n+i}, \quad k \geq 1, \quad (7.5)$$

будет выборкой из требуемого распределения Эрланга. Если ξ_i — интерпретируется как длительность интервала между приходами $(i-1)$ -го и i -го требования, то ζ_k —

момент завершения прихода k -й группы из n -требований и можно говорить о прореживании входного потока $\xi_i, i \geq 1$, т. е. о регистрации моментов прихода каждого n -го требования.

7.2. Гауссовские и равномерные распределения вероятностей

Свойства нормального и гауссовского распределения. Напомним, что равномерным на (a, b) , $a, b \in R^1$, $a < b$ называется распределение вероятностей, задаваемое плотностью

$$f_r(x) = \begin{cases} \frac{1}{b-a} & x \in (a, b); \\ 0 & x \notin (a, b). \end{cases}$$

Гауссовским же или нормальным с параметрами $a, \sigma^2: a \in R^1, \sigma > 0$ — гауссовским $N(a, \sigma^2)$ — называется распределение вероятностей с плотностью $f_v(x)$:

$$f_v(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-a)^2}{2\sigma^2}}.$$

И семейство равномерных и семейство гауссовских распределений вероятностей — двухпараметрическое: для однозначного описания распределения необходимо задать два параметра. Оба семейства замкнуты относительно линейных операций. Важным свойством равномерного и гауссовского распределений является то, что они могут быть получены с помощью линейных операций из выделенных — стандартных распределений. Точнее это утверждение описывается следующими леммами.

Лемма 7.3. Пусть $\xi \simeq [a, b]$ равномерно распределена. Тогда $\xi_0 = (\xi - a)/(b - a) \simeq [0, 1]$ равномерно распределена.

Пусть $\gamma \simeq N(a, \sigma^2)$ распределена, тогда $\gamma_0 = \frac{\gamma - a}{\sigma} \simeq N(0, 1)$ — распределение.

Доказательство. Для доказательства необходимо выписать функцию распределения рассматриваемой случайной величины и продифференцировать. Так, для нормального случая

$$\begin{aligned} F_{\gamma_0}(x) &= P\{\gamma_0 < x\} = P\left\{\frac{\gamma - a}{\sigma} < x\right\} = \\ &= P\{\gamma < a + \sigma x\} = F_{\gamma}(a + \sigma x), \text{ т. е.} \end{aligned}$$

$$f_{v_0}(x) = \frac{d}{dx} F_{v_0}(a + \sigma x) = f_v(a + \sigma x) \cdot \sigma = \\ = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Лемма 7.4. Пусть $\xi \simeq [a, b]$ равномерно распределена. Тогда $\zeta = c\xi + d \simeq [d + ca, d + cb]$ равномерна. Пусть $\xi \simeq N(a, \sigma^2)$ распределена, тогда $\theta = c\xi + d \simeq N(a + d, c^2\sigma^2)$ — гауссовская.

Доказательство проводится, например, как в лемме 7.3. В дальнейшем через $\xi_0(\gamma_0)$ будем обозначать случайные величины с равномерным $[0, 1]$ (соответственно $N[0, 1]$) распределением. Теперь нетрудно установить, что для $\xi \simeq (a, b)$ равномерно распределенной

$$\xi = \frac{\xi - a}{b - a}(b - a) + a = a + \xi_0(b - a)$$

$\xi_0 = \frac{\xi - a}{b - a} \in [0, 1]$ равномерно распределена, а для

$$\gamma = \frac{\gamma - a}{\sigma} \sigma + a = \sigma \gamma_0 + a, \quad \gamma_0 = \frac{\gamma - a}{\sigma} \simeq N(0, 1).$$

Таким образом, любое равномерное или гауссовское распределение может быть получено с помощью некоторого стандартного в своем классе распределения. В принципе, в качестве такового может выступать любое распределение в своем классе. Обычно в качестве стандартных выбирают $[0, 1]$ и $N(0, 1)$ — гауссовское.

Для имитации это означает, что датчик любого равномерного или гауссовского распределения может быть получен с помощью линейного преобразования выхода датчика стандартного распределения.

Имитация стандартного равномерного и гауссовского распределений. Для построения датчика $[0, 1]$ равномерного распределения обычно используется метод вычетов. Он заключается в следующем. Очередное значение выхода датчиков r_{n+1} получается из предыдущего по формуле

$$r_{n+1} = \{Mr_n\}, \quad n \geq 0, \quad r_0 = 2^{-m},$$

где M — достаточно большое целое; m — число двоичных разрядов в мантиссе ячейки ЭВМ; $\{\cdot\}$ обозначает дробную часть. Часто [14] в качестве M берут $M = 5^{2p+1}$, p — целое. Подробное описание такого датчика можно найти, например, в [14].

По рассмотренному принципу построен и датчик RANDU.

При моделировании стандартного гауссовского распределения можно использовать различные алгоритмы. Наиболее простой, пожалуй, тот, который использует центральную предельную теорему применительно к независимым $[0, 1]$ равномерно распределенным случайным величинам

$$r_i: \gamma_0^{(n)} = \sqrt{\frac{12}{n}} \sum_{i=1}^n \left(r_i - \frac{1}{2} \right) \text{ асимптотически } N(0, 1) \text{ нор-}$$

мальна. Уже при $n = 12$, которое удобно из-за множителя, получается достаточно хорошее совпадение. Такой принцип нормирования реализован, например, в датчике GAUSS.

7.3. Моделирование многомерных распределений

Моделирование (a, B) — векторов. Имитация наборов случайных чисел с фиксированным характером статистических связей между компонентами приводит к необходимости моделирования случайных векторов. Взаимосвязь компонент векторов может фиксироваться, например, заданием распределения. Другим типом связей могут быть ограничения на моменты. Отметим, что важнейшим классом многомерных распределений, полностью определяемых своими первыми двумя моментами, является гауссовский.

Напомним, что гауссовским n -мерным вектором ξ с параметрами $a, B: a \in R^n, B — (n \times n)$ — симметричная неотрицательно определенная матрица — называется случайный вектор с характеристической функцией $\varphi_\xi(t) = \exp \{ i t' a - \frac{1}{2} t' B t \}$, $t \in R^n$. Здесь и в дальнейшем все векторы считаются векторами столбцами.

Используя связь между значениями производных характеристической функции и моментами, нетрудно показать, что a — вектор математических ожиданий, а B — матрица ковариаций гауссовского ξ .

Гауссовский вектор будем называть гауссовским $N(a, B)$ -вектором.

В случае, когда B положительно определена, плотность $f_\xi(x)$ гауссовского $N(a, B)$ вектора имеет вид

$$f_\xi(x) = (2\pi)^{-\frac{n}{2}} |B|^{-1/2} \exp \{ (x - a)' B^{-1} (x - a) \},$$

где $x \in R^n$; $|B|$ — определитель матрицы B .

Для моделирования гауссовских векторов и векторов, задаваемых ограничениями на первые и вторые моменты, важную роль играет следующее утверждение.

Лемма 7.5. Симметричная, неотрицательно определенная матрица B однозначно представима в виде $B = UU'$ для нижнетреугольной матрицы U , определяемой по B однозначно (разложение, фигурирующее в лемме 7.5, называют разложением Холецкого).

Доказательство. Элементы матрицы U получают приравниванием элементов матрицы левой и правой части 3.1:

$$\begin{aligned} u_{11} &= b_{11}^{1/2}; \\ u_{i1} &= \frac{b_{21}}{u_{11}}; \\ u_{ii} &= (b_{ii} - \sum_{k=1}^{i-1} u_{ik}^2)^{1/2}; \\ u_{ij} &= \frac{b_{ij} - \sum_{k=1}^{i-1} u_{ik} u_{jk}}{u_{ii}} \end{aligned} \quad (7.6)$$

Нам понадобятся также некоторые свойства ковариаций векторов.

Напомним, что по определению ковариация векторов ξ, η — $\text{cov}(\xi, \eta)$ — определяется соотношением

$$\text{cov}(\xi, \eta) = M(\xi - M\xi)(\eta - M\eta)'. \quad (7.7)$$

Свойства ковариации:

- 1) $\text{cov}(\xi, \eta) = (\text{cov}(\eta, \xi))'$;
- 2) $\text{cov}(\xi_1 + \xi_2, \eta) = \text{cov}(\xi_1, \eta) + \text{cov}(\xi_2, \eta)$;
- 3) $\text{cov}(l\xi, \eta) = l \text{cov}(\xi, \eta)$, $l \in R'$;
- 4) $\text{cov}(A\xi, \eta) = A \text{cov}(\xi, \eta)$, A — $(m \times n)$ — матрица;
- 5) $\text{cov}(a + \xi, \eta) = \text{cov}(\xi, \eta)$, a — неслучайный вектор.

Для примера, рассмотрим свойство 4):

$$\begin{aligned} \text{cov}(A\xi, \eta) &= \text{cov} M(A\xi - MA\xi)(\eta - M\eta)' = \\ &= M[A(\xi - M\xi)(\eta - M\eta)'] = AM(\xi - M\xi)(\eta - M\eta)'. \end{aligned}$$

Здесь мы воспользовались определением математических ожиданий от векторов и матриц из случайных элементов. По определению

$$M \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \begin{pmatrix} M\xi_1 \\ \vdots \\ M\xi_n \end{pmatrix}; \quad M(\xi_{ij})_{i,j=1}^n = (M\xi_{ij})_{i,j=1}^n.$$

Из этого определения следует:

$MA\xi = AM\xi$, A — $(m \times n)$ — неслучайная матрица;

$MA\psi = AM\psi$, ψ — $(n \times k)$ — случайная матрица.

Если $M\xi = a$, $\text{cov}(\xi, \xi) = B$, будем говорить, что $\xi \simeq (a, B)$ распределена, или, что ξ — (a, B) — вектор.

Пусть $a \in R^n$, $B = (n \times n)$ — положительно определенная матрица $B = UU'$, U — нижнетреугольная, определенная однозначно по B в силу леммы 7.5.

Лемма 7.6. Пусть $\gamma \simeq (0, I)$ распределена, I — единичная матрица $n \times n$. Тогда $\xi = a + U\gamma \simeq (a, B)$ — вектор.

Доказательство. $M\xi = a$,

$$\text{cov}(\xi, \xi) = \text{cov}(a + U\gamma, a + U\gamma) = \text{cov}(U\gamma, U\gamma) = U \text{cov}(\gamma, \gamma) U' = UU' = B.$$

Отметим, что $(0, I)$ -вектор может быть получен из n независимых между собой случайных величин с нулевыми математическими ожиданиями и единичными дисперсиями.

Уже разобранные датчики равномерного на $(0, 1)$ распределения — RANDU — и гауссовского $N(0, 1)$ распределения — GAUSS — выдают значения независимых между собой случайных величин. Таким образом, если из n последовательных значений датчика GAUSS сформировать вектор γ_0 и рассмотреть вектор γ_0 , то получится случайный $(0, I)$ -вектор. Такой же вектор $(0, I)$ получится, если сформировать вектор γ_0 из последовательных n значений ρ_i датчика RANDU, предварительно преобразованных в соответствии со следующим выражением:

$$\gamma_i^0 = \left(\rho_i - \frac{1}{2} \right) \cdot \sqrt{12}.$$

В обоих случаях будут получены $(0, I)$ -векторы, различающиеся своими распределениями. В первом случае будет получен гауссовский вектор. Это утверждение вытекает из следующей леммы.

Лемма 7.7. Пусть $\xi = (\xi_1, \dots, \xi_n)' \simeq N(a, B)$, $\eta = A\xi$, где A — $(m \times n)$ матрица, тогда $\eta \simeq N(Aa, ABA')$ — вектор.

Доказательство вытекает из представления характеристической функции вектора η и использовании вида характеристической функции вектора $\xi = (\xi_1 \dots \xi_n)'$: $\varphi_\xi(t) = \exp \left\{ it'a - \frac{1}{2} t' B t \right\}$, $t \in R^n$. Действительно:

$$\begin{aligned} \varphi_\eta(u) &= M \exp \{ iu'\eta \} = M \exp \{ iu' A \xi \} = \\ &= \varphi_\xi(A'u) = \exp \left\{ iu' A a - \frac{1}{2} u' A B A' u \right\}. \end{aligned}$$

Теорема 7.1. (Алгоритм моделирования (a, B) -случайных векторов).

Пусть $\bar{\gamma}_k = (\gamma_{nk+1} \dots \gamma_{(k+1)n})'$ — последовательность векторов, сформированных из последовательных значений датчика случайных величин с нулевыми математическими

ожиданиями и единичными дисперсиями. Тогда $\bar{\xi}_k = U\bar{\gamma}_k + a$ образует последовательность значений независимых (a, B) случайных векторов, где U определяется соотношениями (7.6).

Следствие 7.1. (Алгоритм моделирования гауссовских $N(a, B)$ векторов). Пусть $\gamma_1 \dots \gamma_l$ — последовательность значений с выхода датчика GAUSS. Тогда $\bar{\xi}_1 \dots \bar{\xi}_k \dots$, где

$$\bar{\xi}_k = U (\gamma_{kn+1} \dots \gamma_{(k+1)n})' + a$$

— последовательность значений независимых $N(a, b)$ векторов.

Пример 7.1. Пусть $a = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix} \Rightarrow U = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix}$;
 $\xi = \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix} - (a, B) - \text{вектор.}$

7.4. Имитация марковских процессов

При изучении различного рода процессов, в том числе и систем массового обслуживания, возникает необходимость имитации марковских процессов.

Имитация марковских цепей. Пусть множество состояний описывается отрезком неотрицательного ряда $\{1 \dots n\}$. Как известно, неоднородная цепь Маркова $\xi_0 \dots \xi_m \dots$ задается начальным распределением $(p_1 \dots p_n)$ и последовательностью P_m , $m \geq 1$, $n \times n$ стохастических матриц. Процесс имитации состоит в получении последовательности целых чисел $i_0, \dots, i_m \dots$, являющихся значениями случайных величин $\xi_0, \dots, \xi_m \dots$ соответственно. Имитация начинается с получения i_0 . Значение i_0 — это значение дискретной случайной величины с распределением, задаваемым набором чисел $P_1 \dots P_n$ ($P_i \geq 0$, $\sum P_i = 1$). Для получения соответствующего значения необходимо воспользоваться датчиком дискретного распределения (п. 7.1).

Вероятности $P\{\xi_1 = j / \xi_0 = i_0\}$ описываются i -й строкой матрицы переходных вероятностей на первом шагу P_1 . Таким образом, для получения значения i_1 нужно воспользоваться датчиком дискретного распределения с вероятностями $(p_{i_0 1}^{(1)}, \dots, p_{i_0 n}^{(1)})$, $P_1 = (p_{ij}^{(1)})$. Вообще, если значение i_N уже получено, то i_{N+1} получается с датчика дискретного распределения с вероятностями $(p_{i_N 1}^{(N+1)}, \dots, p_{i_N n}^{(N+1)})$.

Пример 7.2. Пусть необходимо смоделировать однородную цепь Маркова с двумя состояниями $\{1, 2\}$, началь-

ным распределением, задаваемым набором $\left(\frac{1}{3}, \frac{2}{3}\right)$, и матрицей переходных вероятностей $\begin{pmatrix} \frac{1}{8} & \frac{7}{8} \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix}$. Получение i_0 : рассматриваем значение ρ_0 элемента с датчика RANDU.

Полагаем $\xi_0 = \begin{cases} 1 & \rho_0 < \frac{1}{3} \\ 2 & \rho_0 > \frac{1}{3} \end{cases}$. Пусть, например, $\rho_0 < \frac{1}{3} \Rightarrow$

$\Rightarrow \xi_0 = 1$. Распределение ξ_1 в этом случае описывается первой строкой матрицы P . Таким образом, если ρ_1 — значение следующего числа с датчика RANDU, то

$$\xi_1 = \begin{cases} 1 & \rho_1 < 1/8; \\ 2 & \rho_1 > 1/8. \end{cases}$$

Далее процесс повторяется.

7.5. Имитационное моделирование СМО

Классический способ применения математики для описания объекта или изучаемого явления заключаются в таком выборе средств описания, чтобы поведение интересующей нас характеристики выражалось некоторой функцией. Поиск способов аналитического описания этой функции обычно составляет предмет математического моделирования. В простейших случаях требуемую функцию можно выписать аналитически. В других — написать дифференциальное уравнение, которому она удовлетворяет. Однако даже в тех случаях, когда можно написать дифференциальное уравнение, для него не всегда можно написать решение в общем виде, т. е. не всегда можно вывести достаточно простое аналитическое выражение, дающее зависимость изучаемых характеристик от тех, которые на эту изучаемую функцию влияют. Возмозности, представляемые в такой ситуации ЭВМ, трудно переоценить. ЭВМ часто дает возможность получить конкретное числовое значение исследуемой характеристики — в зависимости от конкретных, например, числовых значений исходных характеристик. Математическое описание или модель объекта в этом случае выступает как алгоритм (метод), позволяющий по одним конкретным наборам исходных значений получать другие (исследуемые наборы значений изучаемых характеристик). Математическое описание объекта, алгоритм его функционирования, реализованный на ЭВМ с целью получения

конкретных значений исследуемых характеристик в зависимости от конкретных значений исходных, является имитационной моделью объекта в том смысле, что ЭВМ имитирует поведение изучаемого объекта, выдавая вместо него значение исследуемых характеристик.

У термина «имитационная модель» есть еще один аспект. Имитация подразумевает неотличимость имитируемых характеристик от тех, которые получены естественным путем, т. е. с исследуемого объекта. Регистрация же характеристик реальных явлений так или иначе связана с воздействием случайных факторов. Таким образом, характеристики, вызываемые моделью, должны включать в себя и воздействие случая, т. е. имитировать случайность. Имитация случайности — второе важное достоинство имитационных моделей.

Имитационное моделирование систем массового обслуживания предполагает имитацию входных потоков, а также обслуживающих приборов для изучения различных показателей, характеризующих работу системы на интервале моделирования. Такими показателями для СМО применительно к интервалу моделирования могут быть доля времени занятости приборов, число обслуживаний, средняя продолжительность одного обслуживания, максимальная длина очереди, средняя длина очереди, общее число входов в очередь, общее число входов в очередь без последующего ожидания, среднее время, проведенное в очереди — а также некоторые др.

Для обеспечения возможности получения соответствующих показателей общая логика построения имитационной модели должна обеспечивать в случае необходимости вычисление требуемых характеристик.

Среди общих подходов к моделированию СМО отметим два. Один из них связан с разбиением временного промежутка, на котором моделируется система, на достаточно малые дискреты. После чего на каждом из получившихся дискретов с помощью соответствующих датчиков бернуллиевских случайных величин «разыгрывается» одно из событий: «пришла заявка», «окончилось обслуживание на одном из приборов», а также проверяется, не является ли соответствующий дискрет последним на интервале моделирования.

При другом способе моделирования — по наступлению «критических событий» — просматриваются только моменты прихода заявок и окончания обслуживания, а также проверяется, не истекло ли время моделирования. Каждый из способов имеет свои достоинства, однако второй в боль-

шинстве случаев предпочтительнее. В дальнейшем мы подробно остановимся на методе моделирования по критическим событиям. Имитацию СМО по моментам наступления критических событий (МКС) проиллюстрируем на примере простейшей СМО с одним входным потоком, характеризующимся функцией распределения $F(x)$, и одним прибором, время обслуживания которого характеризуется функцией распределения $G(x)$. Датчики этих распределений в дальнейшем будем называть соответственно F — датчик, G — датчик, а их значения обозначать соответствующими малыми f, g , снабжая, если это необходимо, их подходящими индексами. Общий интервал имитации обозначим T .

При имитации СМО с помощью МКС на интервале имитации T последовательно моделируются моменты критических событий. Это моделирование происходит в процессе выполнения некоторых стандартных шагов — тактов. Главными величинами, определяющими очередное значение МКС, являются: значение ближайшего для данного такта момента окончания обслуживания заявки прибором обозначим через TF , а также значение ближайшего момента прихода заявки будем обозначать TA . Эти величины также изменяются от такта к такту, а их изменение в очередном такте вычисления МКС составляет подготовительную работу для проведения очередного такта вычислений МКС. Согласованное изменение величин МКС, TF , TA , составляющее суть имитации с помощью МКС, происходит следующим образом.

В начальный момент TA полагается равным нулю. Что касается начального значения TF , то к нему вернемся несколько позже.

При изменении величин от такту к такту будем употреблять для них одни и те же обозначения, иногда снабжая переменные индексами «нов», «стар», чтобы отличить новое значение величины, построенной в такте, от того, которое она имела в начале такта.

В начальный момент работы СМО прибор не занят, и первый МКС отвечает приходу заявки, если этот приход вообще уложился в интервал моделирования T . Момент прихода заявки имитируется значением f F -датчика. Значение переменной TA — момента ближайшего прихода заявки — изменяется: $TA_{\text{нов}} = TA_{\text{стар}} + f$, напомним, что $TA_{\text{стар}} = 0$. Все готово для проведения такта вычисления МКС: $МКС = TA$. Точнее было бы написать $МКС = \min(T, TA)$, чтобы учесть возможность непопадания заявки в интервал моделирования. Действительно, если на интервале моделирования нет ни одной заявки, критиче-

ским событием является конец моделирования, что и отражено в формуле. Такт вычисления МКС сопровождается подготовкой к вычислению следующего значения МКС. В рассматриваемой ситуации эта подготовка заключается в вычислении с одной стороны нового значения ТА — момента ближайшего прихода заявки. Новое значение ТА вычисляется обращением к очередному значению f F -датчика и выполнению операции $ТА_{нов} = ТА_{стар} + f$. С другой — вычисление момента ТF ближайшего окончания обслуживания, с учетом того, что прибор запущен на обслуживание заявки. Новое значение ТF вычисляется обращением к G -датчику и присваиванию: $ТF = МКС + g$, поскольку запуск на обслуживание произошел в момент МКС. Это завершает подготовительную работу для вычисления очередного МКС с учетом того, что уже вычисленный МКС — это момент прихода заявки. Будем называть такую систему подготовительных мероприятий отработкой прихода заявки.

Теперь все готово для вычисления очередного МКС — проведения очередного такта вычислений. МКС является ближайшим из трех моментов: ТА, ТF, Т.

$$МКС = \min(ТА, ТF, Т). \quad (7.9)$$

Формула (7.9) является основной рабочей формулой для вычисления МКС в каждом из тактов вычислений. Для работы в очередном такте необходим пересчет значений ТА, ТF. Этот пересчет происходит по-разному в зависимости от того, что отвечает вычисленному значению МКС — минимуму в (7.9) — ТА или ТF. На рис. 7.2 это блоки 5, 7 соответственно. Собственно, с тем, как пересчитываются значения, когда МКС — момент прихода заявки, было сказано выше: имитируется интервал до прихода следующей заявки обращением к F -датчику, изменяется значение ТА: $ТА = ТА_{стар} + f$; если прибор был не занят, происходит обращение к G -датчику и изменяется значение ТF: $ТF_{нов} = ТF_{стар} + g$. Если прибор занят, обращение к датчику не происходит и вся подготовительная работа заключается в изменении ТА. Отработка прихода заявки представлена блоком 7.

Разберем подготовительные мероприятия, если вычисленным значением МКС оказался момент окончания обслуживания, этот набор операций назовем отработкой конца обслуживания. Если в системе нет очереди, то ближайшим МКС может быть только приход заявки, момент прихода которой зафиксирован значением ТА. Для сохранения единообразия формы вычислений по формуле (7.9) в этом случае ТF полагают заведомо большим ТА и Т, например,

$T + 1$. В этом и заключается обработка окончания обслуживания, если в СМО нет очереди. Если же очередь есть, то происходит обращение к очередному значению G -датчика, вычисляется ближайший момент окончания обслуживания: $TF_{\text{нов}} = TF_{\text{стар}} + g$, и можно переходить к очередному такту вычислений МКС.

Теперь можно вернуться к обсуждению задания начальных значений TF . Поскольку в начале работы СМО ближайшим МКС может быть только приход заявки, то,

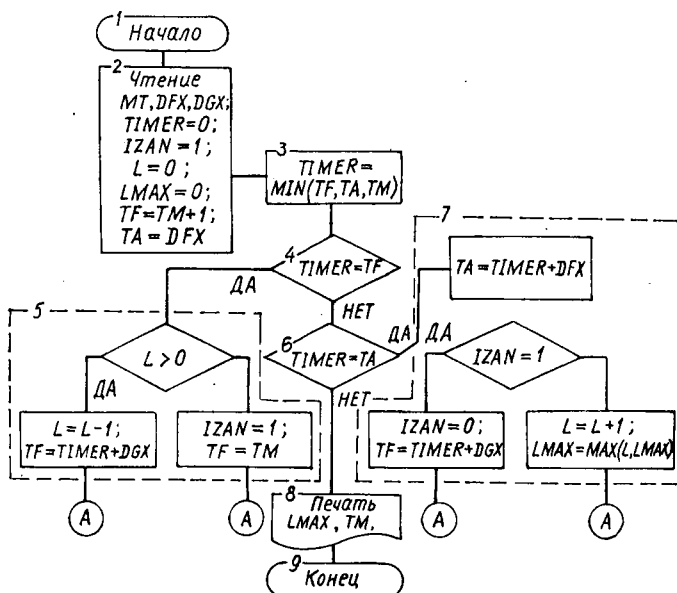


Рис. 7.2.

как и при обработке окончания обслуживания, для обеспечения единообразия вычислений по формуле (7.9) полагают $TF = T + 1$, и тогда и на начальном этапе работы СМО вычисление очередного МКС можно проводить в соответствии с (7.9).

Окончание моделирования происходит, когда очередное вычисленное значение МКС станет равным T .

Выше были рассмотрены вопросы, связанные с просмотром МКС на интервале моделирования. Что касается вычисления исследуемых характеристик СМО, например максимальной длины очереди, то они пересчитываются на каждом шагу вычисления МКС. Таким образом, нет нужды хранить вычисляемые значения МКС до конца моделирования.

На рис. 7.2 приведена блок-схема программы, имитирующая СМО, описанную выше, с неограниченной очередью. При этом исследуемой характеристикой является максимальная длина очереди. Дисциплина обслуживания — «первым пришел — первым обслужился».

Переменный МКС в программе обозначен как TIMER, для ТА, TF оставлены прежние обозначения. Интервал моделирования обозначен ТМ, максимальная длина очереди — L. Значения F, G-датчиков обозначены соответственно DFX, DGX.

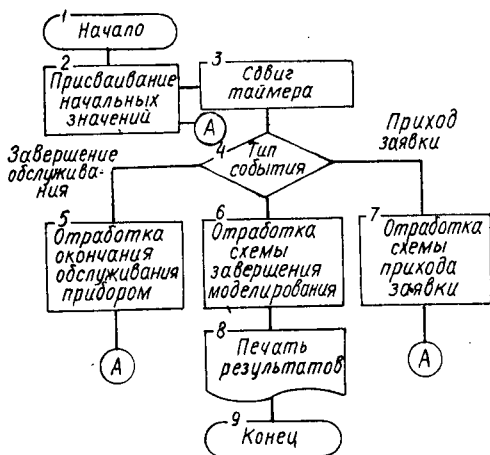


Рис. 7.3.

Кроме уже отмеченных переменных TIMER, TF, ТМ, ТА, датчиков, а также выходных переменных, отвечающих значениям исследуемых характеристик в модели, должны быть предусмотрены внутренние переменные, описывающие занятость прибора. Обозначим их IZAN1, ..., IZANK по числу приборов. Переменные принимают два значения: 0 — прибор занят; 1 — прибор свободен. Общая схема работы модели представлена на рис. 7.3. Поскольку способ изменения переменных и подготовка к новому сдвигу зависят от типа обслуживания, объединим соответствующие операции в блоки соответственно: «обработка логической схемы прихода заявки», «обработка логической схемы завершения обслуживания прибором», «обработка логической схемы окончания моделирования».

Номера блоков рис. 7.2 соответствуют номерам блоков рис. 7.3 иногда укрупненных. В соответствии с обсуждавшейся выше общей схемой переменные, используемые в программе: TIMER — переменная моделирования момен-

гов «критических событий»; TA — переменная моделирования ближайшего момента прихода заявки; TF — переменная моделирования ближайшего момента времени окончания обслуживания заявки; $IZAN$ — переменная, характеризующая занятость прибора: значение 0 соответствует свободному, 1 — занятому прибору; TM — время моделирования; L — длина очереди; $LMAX$ — выходная переменная — максимальная длина очереди; DFX — переменная, отвечающая выходу датчика случайных чисел с распределением, описываемым $F(x)$; DGX — переменная, отвечающая выходу датчика случайных чисел с распределением, описываемым $G(x)$.

Переменные DFX , DGX , TM — входные; $TIMER$, TF , TA , $TZAN$, L — внутренние; $LMAX$ — выходная.

При моделировании СМО с несколькими входными потоками и несколькими приборами логика работы по имитации, отраженная на рис. 7.2, полностью сохраняется. Изменения в имитационной программе носят скорее количественный характер: увеличивается число датчиков случайных чисел: по одному на каждый входной поток и на каждый обслуживающий прибор. Соответственно изменяется число переменных типа TA : их теперь $TA_1, \dots, TA_k$ — по числу k каналов поступления заявок, и число переменных типа TF , $IZAN$: их теперь M по числу обслуживающих приборов. Значение очередного МКС — переменная $TIMER$ — находится по формуле типа (7.9)

$$TIMER = \min(TA_1, TA_2, \dots, TA_k, TF_1, \dots, TF_m).$$

Дальнейшие вычисления: выход по TA или по TF превращается в выход по TA_i , TF_i . Соответственно вычисления происходят для того индекса, по которому произошел выход из блока 4.

Задания для самоконтроля

1. Написать блок-схему алгоритма работы G_1 , G_2 модели для двух приборов, характеризующимися функциями распределения с общим входным потоком, дисциплиной обслуживания «первым пришел — первым обслужен» первым освобожденным прибором. Если приборы заняты, требование становится в очередь. Исследуемые параметры: максимальная длина очереди и число обслуживаний каждым прибором.

2. В условиях предыдущей задачи написать блок-схему модели исследования средней длины очереди.

3. Для системы, описанной в первой задаче, выдать в качестве характеристики среднее время, проведенное требованиями в очереди, если считать только те требования, которые стояли в очереди.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- Аппроксимация диффузионная 110
 Время пребывания ПМП в подмножестве 134
 Датчик случайных чисел распределения вероятностей 229
 — гауссовского 235
 — дискретного 229
 — многомерного 237
 — равномерного 234
 Загрузка системы 56
 Значение параметра мгновенное 21
 Имитация распределения вероятностей 229
 — марковского процесса 238
 — системы массового обслуживания 240
 — , язык GPSS 245
 Интенсивность гибели 46
 — потока 24
 — размножения 46
 Коэффициент равномерно сильного перемешивания 176
 Метод одного вероятностного пространства 124
 S — множество 141
 Модель Эрланга 51
 Неравенство Чебышева 130
 Параметр потока 18
 Переход правильный 143
 Поток последствий 14
 — восстановления 30
 — входящий 7
 — ординарный 14
 — потерянных требований (отказов) 161, 214
 — простейший 10
 — нестационарный 21
 — редкий 163
 — редких событий 168
 — рекуррентный 30
 — с запаздыванием 37
 — случайный 10
 — стационарный 14
 Принцип монотонности траекторий 168
 Приоритет абсолютный 8, 87
 — относительный 8
 Процесс восстановления 40
 — диффузионный 112
 — марковский, однородный по 2-й компоненте 190
 — марковский с полумарковским вмешательством случая 190
 — Орштейна-Уленбека 117
 — переключающийся 193
 — полумарковский 42
 — пуассоновский 21
 — размножения и гибели 46
 — чистого 46
 — чистой гибели 46
 — с дискретной компонентой 190
 — с дискретным вмешательством, случай 189
 — с независимыми приращениями, управляемый марковским процессом 190
 — — полумарковским процессом 191
 — укрупненный 188
 Распределение стационарное 49
 Система с быстрым обслуживанием 153
 — с групповым поступлением 85
 — $GI[G]1$ 83
 — $G[M]1$ 79
 — $M[G]1$ 70
 — марковская 51
 — с потерями 57
 — $M[M]1$ 62
 — $M[D]1$ 79
 — $M[E_R]1$ 83
 — с переменной интенсивностью обслуживания 98
 — с рекуррентным входящим потоком 161
 — с переменным порядком обслуживания 215
 — с периодическим регулярным входным потоком 89
 Событие критическое в СМО 241
 Событие редкое 126, 172
 Теорема восстановления 35
 — Григелиониса 23
 — Королюка 30
 — Ренья 164
 Тождество Колмогорова-Феллера 30
 Уравнение восстановления 34
 — — многомерное 43
 — марковского восстановления 43
 — Линдли 84
 Уравнения Колмогорова дифференциальные 47, 48
 Формула Поллачека-Хинчина 75
 — Хинчина 76
 — Эрланга 62
 Функция ведущая 21
 — восстановления 30
 — Пальма-Хинчина 28
 Цепь Маркова вложенная 41

СПИСОК РЕКОМЕНДУЕМОЙ ЛИТЕРАТУРЫ

1. *Анисимов В. В.* Асимптотическое укрупнение состояний случайных процессов //Кибернетика.— 1973.—№ 3. — С. 109—117.
2. *Анисимов В. В.* Предельные теоремы для случайных процессов и их применение к дискретным схемам суммирования.—К.: Вища шк. Головное изд-во, 1976.—80 с.
3. *Анисимов В. В.* Асимптотические методы анализа стохастических систем.— К.; Тбилиси: Мешниереба, 1984.—180 с.
4. *Гнеденко Б. Н., Коваленко И. Н.* Введение в теорию массового обслуживания.—М.: Наука, 1966.—431 с.
5. *Ермаков С. М., Михайлов Г. А.* Статистическое моделирование.— 2-е изд.—М.: Наука, 1982.—296 с.
6. *Ивченко Г. И.* Теория массового обслуживания.—М.: Высш. шк., 1982.—244 с.
7. *Клейнрок Л.* Теория массового обслуживания.— М.: Машиностроение, 1979.—432 с.
8. *Климов Г. П.* Стохастические системы обслуживания.—М.: Наука, 1966.—243 с.
9. *Коваленко И. Н.* Исследования по анализу надежности сложных систем.—К.: Наук. думка, 1975.—210 с.
10. *Королюк В. С., Турбин А. Ф.* Полумарковские процессы и их приложения.— К.: Наук. думка, 1976.—184 с.
11. *Королюк В. С., Турбин А. Ф.* Математические основы фазового укрупнения сложных систем.— К.: Наук. думка, 1978.—218 с.
12. *Саати Т. Л.* Элементы теории массового обслуживания и ее приложения.—М.: Сов. радио, 1971.—520 с.
13. *Феллер В.* Введение в теорию вероятностей и ее приложения. Т. 2.—М.: Мир, 1967.—752 с.
14. *Шрайбер Т. Дж.* Моделирование на GPSS.—М.: Машиностроение, 1980.—592 с.

Учебное пособие

Владимир Владиславович Анисимов

Олег Каленикович Закусило

Владимир Степанович Донченко

Элементы теории массового обслуживания и асимптотического анализа систем

Редактор *Ж. Г. Давиденко*

Художественный редактор *С. П. Духленко*

Технический редактор *Г. Б. Верник*

Корректор *А. П. Долгая*

Информ. бланк № 9785

Сдано в набор 07.08.85. Подп. в печать 13.11.86. БФ 02709. Формат 84×108/32. Бумага типогр. № 3 Лит. гарн. Выс. печать. Усл. печ. л. 13,02. Усл. кр.-отт. 13,23. Уч.-изд. л. 13,33. Тираж 2000 экз. Изд. № 7024. Зак. 6-246. Цена 65 к.

Головное издательство издательского объединения «Вища школа», 252054, Киев-54, ул. Гоголевская, 7

Отпечатано с матриц книжной фабрики им. М. В. Фрунзе в Харьковской городской типографии № 16. 310003, Харьков-3, ул. Университетская, 16. Зак. 2006.