

А. В. Дигрис

ДИСКРЕТНО-СОБЫТИЙНОЕ МОДЕЛИРОВАНИЕ

Курс лекций

МИНСК
БГУ
2011

УДК 519.8(075.8)
ББК 22.18я7
Д44

*Печатается по решению
Редакционно-издательского совета
Белорусского государственного университета*

Р е ц е н з е н т ы :
доктор физико-математических наук,
профессор *В. М. Борздов*;
доктор физико-математических наук,
профессор *М. А. Маталыцкий*

Дигрис, А. В.
Д44 Дискретно-событийное моделирование [Электронный ресурс] :
курс лекций / А. В. Дигрис. – Минск : БГУ, 2011. – Режим доступа :
<http://www.elib.bsu.by>, ограниченный.
ISBN 978-985-518-511-7.

Рассматриваются базовые принципы построения математических моделей, их классификация и виды моделирования. Излагаются методы моделирования случайных величин, случайных процессов и различных типов потоков случайных событий. Описываются преобразования случайных потоков, обобщенный подход к анализу статистических характеристик потоков случайных событий в физическом эксперименте, а также наиболее часто используемые при этом математические алгоритмы.

Предназначено для студентов факультета радиофизики и компьютерных технологий БГУ.

**УДК 519.8(075.8)
ББК 22.18я7**

ISBN 978-985-518-511-7

© Дигрис А. В., 2011
© БГУ, 2011

1. ВВЕДЕНИЕ

Л е к ц и я 1

МОДЕЛИРОВАНИЕ КАК МЕТОД НАУЧНОГО ПОЗНАНИЯ

В настоящее время, с развитием многочисленных научных направлений в различных областях знаний, все более актуальной и важной становится проблема разработки методов исследования, которые позволяют наиболее эффективно и быстро решать задачи, возникающие в той или иной научной сфере. Одним из наиболее распространенных методов изучения процессов и явлений в современной науке является метод моделирования.

Моделированием называется замещение одного объекта (оригинала) другим (моделью) и фиксация или изучение свойств оригинала путем исследования свойств модели.

Благодаря интенсивному развитию вычислительной техники и кибернетики моделирование приобретает общенаучный характер и применяется в исследованиях живой и неживой природы, в науках о человеке и обществе. Моделирование используется в тех случаях, когда исследование самого объекта-оригинала невозможно или невыгодно по каким-либо причинам (высокая стоимость создания оригинала, его сложное строение, малые пространственно-временные размеры и т. д.). В процессе моделирования объектом-оригиналом может служить любая реальная или воображаемая система. Каждый моделируемый объект имеет ряд составляющих его элементов и связей между ними, а следовательно, обладает структурой. Состояние объекта определяется в любой момент времени пространственным расположением его элементов, которое с течением времени меняется в результате взаимодействия частей объекта как между собой, так и с окружающей средой. Структурность оригинала и его поведение могут быть отражены с помощью множества параметров рассматриваемой системы, а также ряда ее характеристик, зависящих от параметров, причем каждая характеристика определяется некоторым их подмножеством. Модель в свою очередь также обладает рядом характеристик и параметров, часть из которых совпадает с параметрами объекта-оригинала.

При построении модели необходимо учитывать, что она правомерно может заменить рассматриваемый объект лишь в случае, если соответствующие исследуемые характеристики модели и исходного объекта опре-

деляются однотипными подмножествами параметров и связаны с этими параметрами одинаковыми зависимостями. С другой стороны, моделирование будет целесообразным только тогда, когда модель не обладает теми признаками объекта-оригинала, которые препятствуют его эффективному исследованию, в то время как другие параметры модели могут способствовать фиксации и изучению исследуемых свойств.

Классификация моделей. В силу многообразия критериев, по которым все модели можно разбить на классы или группы, задача построения единой классификации моделей не имеет однозначного решения. Группирование моделей может проводиться по характеру моделей, по характеру моделируемых объектов, по сферам приложения и его уровням. Таким образом, любая классификация моделей, а следовательно, и методов моделирования не может претендовать на полноту и единственность.

Далее будет рассмотрена одна из наиболее распространенных классификаций моделей, в основу которой положена степень абстрагирования модели от оригинала. Согласно выбранному критерию все множество моделей, прежде всего, может быть разбито на два класса: физические (материальные) и математические (абстрактные).

Физическая модель представляет собой некоторую систему, которая эквивалентна или подобна оригиналу, или имеет сходный с ним процесс функционирования. Различают физические модели следующих видов:

1. *Натурные модели* обычно представляют собой реальные изучаемые системы, примерами которых могут служить макеты и опытные образцы. Такие модели полностью или практически адекватны объекту-оригиналу и обеспечивают достаточно высокую точность и достоверность результатов моделирования. В то же время создание и эффективное исследование таких моделей возможно лишь для сравнительно узкого класса систем ввиду дороговизны данного подхода и возможности присутствия у реальной системы свойств, затрудняющих анализ требуемых характеристик.

2. *Квазинатурные модели* – это модели, которые могут быть представлены в виде совокупности натурных и математических моделей. Необходимость их использования обусловлена наличием ситуаций, когда для одной из частей рассматриваемой системы математическая модель не приемлема (модель человека-оператора), либо когда часть моделируемой системы еще не существует на практике, или ее натурное моделирование затруднительно. Такие модели часто используются при разработке новых систем, а также при создании и тестировании их программного обеспечения.

3. *Масштабные модели* – системы аналогичной оригиналу физической природы, но отличающиеся по масштабам. В качестве методологической основы масштабного моделирования используется теория подобия, которая используется для построения соотношений между параметрами и характеристиками модели и объекта-оригинала. Примером использования масштабной модели может служить процесс продувки моделей самолетов в аэродинамических трубах.

4. *Аналоговыми моделями* являются системы, функционирование которых аналогично исходному объекту, а физическое строение может в той или иной степени отличаться. Такие модели требуют наличия математического описания исследуемой системы, так как для их использования необходима тождественность безразмерных математических отображений исследуемых процессов для моделируемого объекта и его модели. Наиболее широко известны аналоговые модели, где напряжение или сила тока заменяют некоторый механический процесс.

Моделирование с участием описанных выше видов моделей, в ходе которого модель воспроизводит основные геометрические, физические, динамические и функциональные характеристики оригинала, называется *предметным моделированием*. Если модель и моделируемый объект имеют одну и ту же физическую природу, то говорят о *физическом моделировании*.

Преимущество физического моделирования перед натурным экспериментом заключается в том, что условия создания и исследования модели могут значительно отличаться от аналогичных условий, свойственных оригиналу, и выбираются, исходя из удобства и простоты предполагаемого исследования. Поскольку при моделировании нет необходимости сохранять размеры систем-оригиналов, и в точности соблюдать все условия их функционирования, имеется возможность получить существенный выигрыш во времени и стоимости исследования. Однако условия моделирования выбираются не абсолютно произвольно. Между оригиналом и моделью должны быть сохранены некоторые соотношения подобия, вытекающие из закономерностей физической природы явлений и гарантирующие возможность использования сведений, полученных путем моделирования, для оценки свойств оригинала.

Математическая модель представляет собой формализованное описание системы с помощью абстрактного языка, в частности с помощью математических соотношений, отражающих процесс функционирования системы.

Базовым аппаратом, на котором основано математическое моделирование, является по существу вся математика. Здесь применяется диф-

ференциальное, интегральное, алгебраическое исчисление, теории множеств, различные математические алгоритмы и т. д. Стоит отметить, что зачастую одни и те же математические модели могут использоваться для изучения реальных систем различной природы, поскольку математическое описание исходных систем имеет одну и ту же форму. Дальнейшее разделение математических моделей на классы, так же как и для моделей в целом, зависит от выбора критерия классификации.

Так, например, *в соответствии с характером соотношений*, которые определяют зависимости *между характеристиками и параметрами математических моделей*, вся их совокупность может быть разделена на два основных типа:

а) *детерминированные модели* – модели, для которых в любой заданный момент времени устанавливается однозначное соответствие между параметрами и характеристиками рассматриваемой системы;

б) *вероятностные (стохастические) модели* – модели, для которых в любой заданный момент времени могут быть зафиксированы лишь распределения вероятностей для характеристик системы.

По принципу построения и характеру функционирования выделяют следующие типы моделей:

а) *аналитические модели* – модели, при построении которых объект-оригинал описывается в виде совокупности математических конструкций (чаще всего различных уравнений и их систем), которые были получены на основе представлений о функционировании моделируемой системы. В результате при использовании такой модели процесс функционирования исходной системы, как правило, в явном виде не воспроизводится;

б) *имитационные модели* – модели, при построении и использовании которых моделирующий алгоритм с той или иной степенью точности воспроизводит функционирование исходной системы. При этом зачастую осуществляется имитация элементарных явлений, протекающих в системе, без нарушения взаимосвязи между ними и их логической структуры.

Имитационные модели могут быть созданы и успешно использоваться для более широкого класса систем, чем аналитические, так как они позволяют легко выделить необходимые исследователю характеристики системы и тем самым избежать процессов, затрудняющих изучение исходного объекта.

Для изучения любой реальной системы с применением математического моделирования требуется в первую очередь построить ее математическое описание, которое и будет являться математической моделью. При построении математических моделей, независимо от их принадлеж-

ности к тому или иному классу в соответствии с выбранным принципом классификации, используют единый набор правил. Общие правила построения математических моделей могут быть представлены в виде ряда этапов, которые будут рассмотрены ниже.

Этапы построения математической модели. Процесс создания математической модели состоит из нескольких этапов, начиная с изучения моделируемой системы и условий ее функционирования и заканчивая проверкой (тестированием) полученной модели.

Первый этап состоит в построении *содержательного описания* исследуемой системы, которое в дальнейшем будет являться исходным материалом для последующих этапов построения математической модели. Содержательное описание в словесном выражении концентрирует сведения о физической природе и количественных характеристиках элементарных явлений, протекающих в исследуемой системе, о степени и характере взаимодействия между ними, о месте и значении каждого элементарного явления в общем процессе функционирования рассматриваемой реальной системы. Также формируется представление о структуре системы, свойствах ее элементов и причинно-следственных связях, присущих исследуемой системе и существенных для достижения цели моделирования. Содержательное описание исследуемой системы составляется, как правило, специалистами соответствующей научной или технической области, практически без участия математиков. Поэтому оно может и не иметь строгой математической формулировки.

Важным моментом при построении содержательного описания является четкое изложение идеи предполагаемого исследования и формулировка цели моделирования. Здесь необходимо отметить, что в общем случае возможно построение множества различных моделей для одной и той же реальной исследуемой системы, которые могут различаться по степени детализации, могут быть ориентированы на выделение определенных функций и характеристик, являющихся наиболее важными для исследователя. Поэтому формулировка цели моделирования важна уже с точки зрения выбора наиболее подходящей модели. Правильное определение целей моделирования способствует повышению в дальнейшем эффективности созданной модели.

После построения содержательного описания исследуемой системы выполняется создание *формализованной схемы*. Формализованная схема является промежуточным звеном между содержательным описанием и математической моделью.

Для построения формализованной схемы необходимо выбрать оптимальную систему параметров, определяющих поведение исследуемой

системы, определить характеристики, по которым можно будет оценивать характер поведения системы, а также определить все зависимости между выделенными характеристиками и параметрами. На этапе построения формализованной схемы должна быть дана точная математическая формулировка задачи исследования с указанием перечня искомых величин и оцениваемых зависимостей.

При построении формализованной схемы, как правило, выполняется ряд последовательных шагов, позволяющих создать основу будущей математической модели.

Первым шагом является выполнение *детализации модели*, т. е. выделение отдельных подсистем. Количество выделяемых подсистем определяется установленным оптимальным количеством варьируемых параметров, а также целями моделирования (например, если цель моделирования состоит в изучении поведения и характеристик некоторых элементов исходной системы, то целесообразно каждый такой элемент представить в виде отдельной подсистемы). Уровень детализации, т. е. глубина разбиения на составляющие, для различных частей системы может быть различной.

Следующим шагом на данном этапе является *локализация модели*. Она осуществляется путем представления всех источников внешних воздействий на систему, если таковые существуют, в виде специальных элементов – генераторов внешних воздействий, которые также включаются в создаваемую модель.

Завершающим шагом является *структуризация и выделение процессов или состояний*. Процесс структуризации включает в себя указание количества и характера связей между элементами в системе. В достаточно сложных системах, которые содержат многофункциональные элементы, часто обладающие множеством входных и выходных связей, выделяют специальные управляющие элементы, которые направляют процессы передачи данных между элементами системы, определяют алгоритмы функционирования самих элементов и т. д. Если рассмотренные ранее действия определяли статику моделируемой системы, то выделение процессов определяет динамику ее функционирования. Если исследуемая система является системой со структурным принципом управления, рассматривают не процессы, протекающие в системе, а состояния, в которых могут оказаться элементы системы в течение времени моделирования. При таком подходе каждому элементу системы ставится в соответствие один или несколько параметров, изменение значений которых будет определять состояние этого элемента в процессе функционирования системы.

Отдельной, не менее важной, задачей при построении формализованной схемы является систематизация и уточнение совокупности всех исходных данных, значений известных параметров исследуемой системы и начальных условий. О важности данной задачи говорит уже то, что достоверность результатов моделирования во многом зависит от точности и полноты исходных данных. В некоторых случаях задача определения начальных значений параметров может быть осложнена по следующим причинам:

а) значения параметров могут часто быть не только детерминированными, но и носить случайный характер;

б) не все параметры моделируемой системы могут быть стационарными;

в) часто конструируются модели систем, еще не существующих на практике, а следовательно, значения их параметров неизвестны.

Указанные выше случаи требуют поиска путей преодоления возникающих трудностей. Наиболее эффективным способом, позволяющим решить проблему представления случайных величин, является задание их с помощью законов распределения. Это будет способствовать в дальнейшем как созданию аналитической модели рассматриваемой системы, так и проведению имитационного моделирования ввиду наличия разработанных методов моделирования случайных величин с заданными распределениями. Особую важность использование законов распределения для случайных параметров приобретает в случае, если такой параметр зависит от времени, поскольку при необходимости рассмотрения поведения системы на определенном временном отрезке, задание совокупности значений этого параметра, например с помощью таблиц, затруднительно. В случае если для некоторых параметров нет вообще никаких данных о возможных значениях, экспертами в данной области, в соответствии с их личным опытом, осуществляется выдвижение гипотез об этих значениях. В этом случае корректность и точность исходных данных полностью зависит от квалификации специалистов.

Кроме определения начальных значений параметров системы также устанавливаются границы изменения для непрерывных и области принимаемых значений для дискретных параметров.

При отсутствии явной зависимости между входными и выходными параметрами и наличии лишь совокупности их значений возникает проблема аппроксимации этой зависимости определенным математическим выражением, для чего могут применяться различные методы интерполяции. После того как для набора входных и выходных параметров в соответствии с их значениями была определена некоторая приближенная за-

висимость, устанавливающая связь между ними, далее она проверяется и при необходимости уточняется.

Чаще всего содержательное описание включает сведения, достаточные для построения форматизированной схемы. Если этих сведений недостаточно, то требуются дополнительные эксперименты или наблюдения, уточняющие представления об исследуемой системе. Следует подчеркнуть, что формализованная схема полностью подводит итог изучению и экспериментальному обследованию системы. Все сведения об исследуемой системе, полученные из эксперимента и технической документации, должны быть использованы на этапе построения формализованной схемы.

Следующим этапом при построении математической модели является *преобразование формализованной схемы в математическую модель*, которое выполняется при помощи математических методов. Важно отметить, что при выполнении данного преобразования опираются только на сведения, полученные после разработки формализованной схемы. В процессе создания непосредственно математической модели записывают в аналитической форме все соотношения, которые еще не были записаны, выражают логические условия в виде системы неравенств и по возможности придают аналитическую форму всем другим сведениям, содержащимся в формализованной схеме. В результате получают математическую модель системы, которая представляет собой совокупность соотношений, связывающих выходные характеристики и характеристики состояния системы с параметрами системы, входными сигналами и характеристиками состояний в предыдущие моменты времени.

При построении математической модели всегда стремятся к представлению информации, содержащейся в формализованной схеме, в виде аналитических зависимостей. Это значительно упрощает последующее использование построенной модели как с точки зрения полноты получаемых результатов, так, нередко, и с точки зрения требуемых вычислительных затрат. В случае невозможности построения аналитической модели изучаемой системы в целом (например, из-за сложности изучаемой системы), оно осуществляется в определенной степени для ее отдельных элементов с последующим применением, например, имитационного подхода.

Последним этапом при создании математической модели является *проверка ее адекватности исследуемой системе и оценка эффективности построенной модели*. В процессе создания модели адекватность может быть искажена по причине чрезмерного ее упрощения при проведении детализации и локализации в результате пренебрежения какими-

либо внешними факторами, оказывающими существенное влияние на работу системы, или игнорирования влияния некоторых важных параметров, либо неправильной аппроксимации закона их поведения. Для оценки адекватности созданной модели существует ряд критериев, использование которых зависит, прежде всего, от специфики самой системы. При обнаружении неадекватности созданной модели проводят ее корректировку, внося в построенную модель на определенном этапе изменения.

Оценка эффективности построенной модели производится методами, которые могут использовать как один, так и несколько критериев. В качестве критериев оценки эффективности модели обычно используются такие, как стоимость, надежность модели, ее производительность и др. При оценке эффективности важно учитывать внешние воздействия на систему, при их наличии, так как они могут оказать существенное влияние на вышеуказанные критерии.

Использование построенной математической модели в процессе моделирования зависит в первую очередь от характера самой модели, а также математических соотношений, которые ее составляют. В зависимости от того как может использоваться построенная модель выделяют различные виды моделирования.

Виды моделирования. Выделяют *три основных вида моделирования: аналитическое, численное и имитационное.*

Аналитическое моделирование. При разработке математической модели в процессе ее построения на основе формализованной схемы обычно стремятся, насколько это возможно, представить содержимое модели в виде аналитических зависимостей. В идеальной ситуации стремятся построить математическую модель в виде такой системы соотношений между характеристиками, параметрами системы и внешними воздействиями, которая допускает выполнение аналитического моделирования, под результатами которого понимают построение явных формул для искомых величин, либо проведение исследования уравнений качественными методами. Получение результатов такого характера является настолько полным решением задачи, что к аналитическому моделированию стремятся в первую очередь. Однако воспользоваться аналитическими исследованиями удается сравнительно редко, т. к. построение математической модели в виде системы уравнений, допускающих простое аналитическое решение, является трудной задачей, а для сложных систем непреодолимой. Тем не менее при решении многих прикладных задач зачастую используют упрощения ради возможности получить хотя бы

приближенное аналитическое решение задачи, что продемонстрировано на следующем примере.

Пример 1.1. Типичным примером упрощения реальной картины явлений может служить предположение о пропорциональности силы $F = \psi\xi(t)$, действующей на маятник со стороны пружины, растяжению пружины $\xi(t)$. При таком предположении дифференциальное уравнение, описывающее гармонические колебания маятника, имеет вид

$$m \frac{d^2 \xi(t)}{dt^2} = -\psi \xi(t), \quad (1.1)$$

где $\xi(t)$ – отклонение центра масс пружинного маятника от положения равновесия в момент времени t , m – его масса, а $\psi\xi(t)$ – сила, действующая на маятник со стороны пружины, ψ – жесткость пружины. Как показывает эксперимент, на самом деле, зависимость силы от растяжения пружины $F = F(\xi)$ оказывается более сложной. Но получить с учетом этой зависимости уравнение, допускающее аналитическое решение, не представляется возможным. Полученное без учета силы внутреннего трения, а также сопротивления воздуха, решение уравнения (1.1) плохо согласуется с реальными колебаниями пружинного маятника в экспериментах. В частности, колебания никогда не бывают свободными, они всегда сопровождаются заметным затуханием. Однако это решение позволяет получить многие полезные выводы и рекомендации, и поэтому широко используются в теоретических и прикладных задачах.

В тех случаях, когда не удастся преобразовать математическую модель в подходящую систему уравнений, либо не удастся получить решение полученной системы уравнений аналитическими методами, а упрощения приводят к недопустимо грубым результатам, от аналитического моделирования отказываются и переходят к другим способам использования математической модели.

Численное моделирование. Более широкую сферу применения имеет исследование систем при помощи численных методов, особенно с учетом бурного развития вычислительной техники. Содержание работы при численном исследовании остается в основном таким же, как и при использовании аналитических методов. Разница заключается в том, что после выполнения наиболее трудной части работы – представления математической модели в виде систем уравнений, оказывается, что решение такой системы с применением аналитических методов либо в целом, либо в частности, в явном виде невозможно. В данной ситуации прибегают к реализации и использованию соответствующих численных методов, при помощи которых и выполняются необходимые вычисления. При исследовании систем численными методами результатами являются таблицы значений искомых величин для конечного набора значений парамет-

ров системы, начальных и граничных условий. Использование численных методов в настоящее время является достаточно эффективным ввиду хорошо разработанного специального программного обеспечения и наличия вычислительных систем высокой производительности. Все это способствует тому, что класс уравнений, которые могут быть решены численно значительно шире, чем класс уравнений, доступных аналитическому решению. Однако численные решения являются менее полными по сравнению с результатами аналитического моделирования.

Использование численного моделирования, хотя и не требует получения аналитических выражений, связывающих искомые величины с входными параметрами исследуемой системы, предполагает представление математической модели в виде некоторой системы уравнений, устанавливающих соотношение между входными параметрами и искомыми величинами. При построении моделей сложных или недостаточно изученных систем, часто построение такой системы уравнений оказывается невозможным. В таких случаях прибегают к использованию имитационного моделирования.

Имитационное моделирование. Имитационное моделирование является в настоящее время наиболее универсальным методом изучения систем и количественной оценки их функционирования. При использовании имитационного моделирования исследования производятся на имитационной модели, где процессы, протекающие в реальной системе, заменяются имитируемыми с сохранением некоторого соотношения длительностей и последовательности во времени отдельных явлений. В процессе проведения имитационного моделирования, как и при экспериментах с самой реальной системой, по фиксируемым значениям параметров определяют требуемые характеристики системы. Преимуществом имитационного моделирования является то, что оно может использоваться в отличие от аналитического и численного моделирования в тех случаях, когда не все элементы системы представлены в виде строгих математических зависимостей. Кроме того, при наличии достаточно мощной вычислительной системы имитационное моделирование может быть выполнено практически при любой степени детализации созданной модели. По этой причине в настоящее время с интенсивным развитием информационных технологий имитационные методы получают все более широкое применение.

Пример 1.2. В качестве примера использования имитационного моделирования можно рассмотреть моделирование стратегии технического обслуживания автобуса. Пусть E – основное состояние автобуса (он исправен и осуществляет N рейсов за смену), A – состояние, когда автобус нуждается в мелком профилактическом ремон-

те в продолжение времени одного рейса, B – состояние, когда автобус нуждается в немедленном ремонте длительностью в одну смену. Предположим, что p_A – вероятность перехода автобуса из состояния E в состояние A , p_B – это вероятность перехода автобуса из состояния A в состояние B . Необходимо выбрать одну из двух стратегий обслуживания. Стратегия α : как только автобус перешел в состояние A , он ремонтируется. Стратегия β : автобус работает до тех пор, пока не перейдет в состояние B . Задача состоит в выборе лучшей из приведенных выше стратегий, которая дает наибольшее количество рейсов в день. Автобус выходит на линию в состоянии E , то есть при любой стратегии автобус, заканчивающий N рейсов в состоянии A или B , ночью ремонтируется.

Решение приведенной выше задачи может быть выполнено с использованием имитационного моделирования. При моделировании используется генератор случайной величины, равномерно распределенной на отрезке $[0,1]$, реализации которой вместе с вероятностями p_A и p_B перехода автобуса между состояниями позволяют определить состояние автобуса в ходе каждого рейса. Вначале имитируется поведение по стратегии α и подсчитывается количество рейсов, затем то же самое выполняется для стратегии β . После этого, сравнив число успешных рейсов в каждом случае, можно сделать вывод о наиболее приемлемой стратегии.

Решение многих практических задач приводит к необходимости анализировать системы, которые подвергаются воздействию случайных факторов или носят стохастический характер. В отличие от других методов имитационное моделирование оказывается весьма удобным аппаратом для исследования стохастических систем. При исследовании систем с учетом случайных факторов возникают дополнительные трудности. Для имитационного моделирования эти трудности обычно сравнительно легко преодолимы, благодаря наличию достаточно хорошо разработанных алгоритмов моделирования случайных величин, процессов, потоков и т. д. Результаты моделирования, полученные при однократном воспроизведении процесса функционирования системы, не могут объективно характеризовать изучаемый объект. Поэтому при моделировании систем с учетом случайных факторов приходится воспроизводить большое количество реализаций и определять искомые величины как средние значения.

Имитационное моделирование может оказаться наиболее приемлемым в следующих случаях:

- 1) не существует законченной математической постановки задачи;
- 2) аналитические модели имеются, но математические процедуры столь сложны и трудоемки, что имитационное моделирование дает более простой или единственно возможный способ решения задачи;

3) кроме оценки выходных параметров, желательно выполнить наблюдение за поведением системы в течение определенного периода времени;

4) имитационное моделирование может оказаться единственно возможным вследствие невозможности постановки эксперимента и наблюдений в реальных условиях (например, изучение поведения космических кораблей в межпланетном пространстве);

5) при моделировании необходимо изменение временной шкалы (замедление либо ускорение изучаемых процессов);

6) необходимо выполнить предварительную проверку новых стратегий и правил принятия решений перед проведением экспериментов на реальной системе;

7) в случае если для стохастической модели данных только о моментах распределений недостаточно и особое значение имеет последовательность событий в исследуемой системе.

Стоит отметить, что имитационным методом решаются не только задачи со случайными воздействиями, а также чисто детерминированные. Причем при моделировании детерминированных систем необходимо предварительно построить стохастическую систему, выходные характеристики которой позволяют оценить искомые характеристики исследуемой системы. Для иллюстрации сказанного выше можно привести следующий пример.

Пример 1.3. Пусть необходимо методом имитационного моделирования найти оценку площади фигуры, ограниченной осями координат, ординатой $x = 1$ и кривой $y = f(x)$. При этом для определенности предполагается, что $0 \leq f(x) \leq 1$ для всех $x \in [0, 1]$.

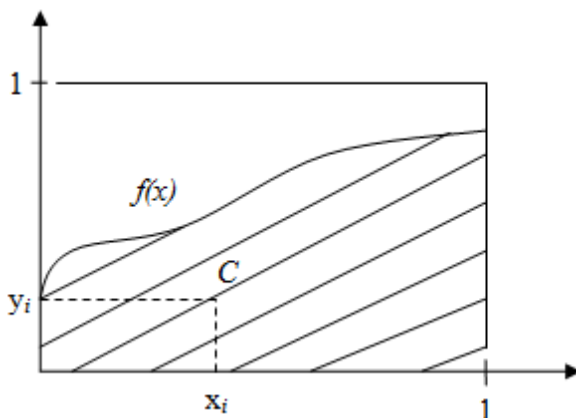


Рис. 1.1. Вычисление площади под кривой $f(x)$ с применением имитационного моделирования

Данная задача определения площади фигуры сводится к вычислению определенного интеграла:

$$S_{\Phi} = \int_0^1 f(x) dx.$$

Для решения задачи с применением имитационной модели выполняются следующие шаги (рис. 1.1):

1. Генерируется пара независимых случайных величин $x_i, y_i; i = 1, 2, \dots$, равномерно распределенных на $[0; 1]$.

2. Определяется координата (x_i, y_i) точки C и проверяется, попала ли она в площадь фигуры, т. е. выполняется ли

неравенство $y_i \leq f(x_i)$. Если да, то исход испытания считается положительным и к счетчику h , который перед началом моделирования был равен 0, добавляется единица.

3. Проведя N указанных выше испытаний, вычисляем $\frac{h}{N}$. Это и есть искомый интеграл.

Принципы изменения модельного времени. Наиболее важным параметром имитационной модели является модельное время, изменение которого отражает процесс функционирования реальной системы. В зависимости от характера изменения данного параметра различают *имитационное моделирование с приращением временного интервала* и *имитационное моделирование с продвижением времени до особых состояний*.

Процесс функционирования моделируемой системы можно рассматривать как последовательную схему ее состояний, описываемую характеристиками $z_1(t), z_2(t), \dots, z_n(t)$ в n -мерном пространстве. Задача моделирования процесса функционирования системы состоит в построении функций $z_1(t), z_2(t), \dots, z_n(t)$, а также в вычислении некоторых величин зависящих от этих функций. В распоряжении имеются соотношения математической модели, связывающие характеристики состояний системы $z_1(t), z_2(t), \dots, z_n(t)$, с ее параметрами и временем, а также начальные условия $z_1^0, z_2^0, \dots, z_n^0$ для начального момента времени t_0 .

Принцип приращения временного интервала. Преобразуем математическую модель к такому виду, чтобы было удобно вычислить $z_1(t + \Delta t), z_2(t + \Delta t), \dots, z_n(t + \Delta t)$, по значениям $z_i(t)$. Возьмем t_0 – начальный момент времени и прибавим некоторый шаг Δt . Вычислим $z_i(t_0 + \Delta t)$ и затем перейдем дальше к вычислению $z_i(t_0 + \Delta t + \Delta t)$. Если шаг достаточно мал, то таким путем можно получить приближенный вид $z_i(t)$. Приведенный выше алгоритм справедлив для детерминированных систем. Для систем со случайными воздействиями алгоритм в целом остается прежним, лишь с некоторыми отличиями. В начальный момент t_0 разыгрывается одно из возможных состояний z_i^0 в соответствии с заданным законом распределения вероятностей. Затем вычисляется условное распределение z_i в момент времени $t_0 + \Delta t$ при условии, что в начальный момент времени были получены некоторые уже известные z_i^0 . Далее разыгрывается состояние в момент времени $t_0 + \Delta t$, и уже

при известных $z_i(t_0 + \Delta t)$ аналогичным образом получаются значения z_i после очередного приращения временного аргумента. Описанная выше процедура повторяется до тех пор, пока не будет построена одна из возможных реализаций случайного многомерного процесса $z(t)$ в заданном интервале времени (t_0, T) . Таким образом, приведенный принцип построения алгоритма имитационной модели позволяет определить последовательные состояния моделируемой системы через некоторые интервалы времени. Данный подход является наиболее универсальным, охватывающим весьма широкий класс реальных систем, однако вместе с тем он не экономичен с точки зрения расхода вычислительного времени.

Принцип продвижения времени до особых состояний. При рассмотрении некоторых систем можно обнаружить существенную неоднородность, неравноправность состояний системы в заданном интервале времени. Удастся выделить два типа состояний: особые и не особые (обычно состояния, в которых система находится постоянно). Особое состояние – это состояние, характерное для системы в некоторые изолированные моменты времени, совпадающие, например, с моментами поступления в систему входных сигналов от внешней среды. Часто для таких состояний характерно то, что координаты $z_i(t)$ в эти моменты времени изменяются скачком, а между особыми состояниями изменение $z_i(t)$ происходит плавно и непрерывно. При использовании принципа продвижения времени до особых состояний изменение модельного времени происходит путем перехода из текущего особого состояния системы в последующее, которое является наиболее ранним из всех прогнозируемых состояний такого рода. При использовании данного подхода имитационная модель включает в себя процедуру определения момента времени, соответствующего следующему особому состоянию по известным характеристикам данного и предыдущих состояний. Примером особого состояния системы, в частности при исследовании систем массового обслуживания (СМО), может служить момент смены состояния обслуживающего прибора при начале или окончании обслуживания заявки. Данный принцип изменения временного параметра является более предпочтительным, так как экономит время вычислений, поскольку предполагает ровно столько промежуточных моментов времени, сколько возникает особых состояний.

На практике обычно не выдерживается строго определенный принцип изменения модельного времени, а используется их совокупность или в целом, или на отдельных этапах выполнения задачи.

Наряду со всеми преимуществами имитационных моделей, приведенными выше, они также имеют и ряд существенных недостатков. Разработка хорошей модели часто требует много времени. Весьма существенно то, что при использовании имитационной модели исследователь может получить интересующие его ответы, касающиеся поведения исследуемой системы, только после очередного имитационного эксперимента и возможности прогнозирования значительно меньше, чем при использовании аналитического подхода. Тем не менее высокий уровень детализации систем любого уровня сложности, возможность исследования динамики развития процессов обеспечивают имитационному методу большую перспективу широкого практического применения.

Выбор того или иного метода моделирования во многом определяется спецификой математической модели системы и возможностью ее представления в виде разрешимых математических конструкций. Отметим также, что часто бывает полезно использовать аналитическое и имитационное моделирование совместно, если они могут дополнить друг друга.

Таким образом, рассмотрев основные этапы построения модели и подготовки процесса моделирования, можно заключить, что моделирование сегодня имеет основательно разработанную методологическую базу, что дает возможность результативно и целенаправленно проводить с его помощью исследование систем различной сложности во многих областях науки и техники. Метод моделирования может эффективно взаимодействовать с другими общенаучными методами, как используя полученные с их помощью результаты, так и давая информационную базу для проведения дальнейших исследований.

2. МЕТОДЫ МОДЕЛИРОВАНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Л е к ц и я 2

МОДЕЛИРОВАНИЕ БАЗОВОЙ СЛУЧАЙНОЙ ВЕЛИЧИНЫ

Основной целью стохастического моделирования является создание стохастических моделей сложных систем, при построении которых необходимо использовать алгоритмы моделирования случайных величин, случайных процессов и потоков случайных событий. Несмотря на то что указанные выше алгоритмы отличаются друг от друга по своей структуре, любой из них в конечном итоге является некоторым преобразованием одной или нескольких реализаций *базовой случайной величины*. Именно по этой причине умение качественно моделировать базовую случайную величину приобретает особенно важное значение при создании и реализации моделей сложных стохастических систем.

Моделирование базовой случайной величины. *Базовой случайной величиной* называется непрерывная случайная величина α , равномерно распределенная на отрезке $[0, 1]$.

Интегральная функция распределения базовой случайной величины имеет вид

$$F_{\alpha}(x) = P\{\alpha \leq x\} = \begin{cases} 0, & x < 0 \\ x, & 0 \leq x \leq 1, \\ 1, & x > 1 \end{cases} \quad (2.1)$$

а ее плотность распределения соответственно

$$f_{\alpha}(x) = \begin{cases} 1, & x \in [0, 1] \\ 0, & x \notin [0, 1] \end{cases}. \quad (2.2)$$

Математическое ожидание M_{α} и дисперсия D_{α} базовой случайной величины равны

$$M_{\alpha} = \int_0^1 x f_{\alpha}(x) dx = \frac{1}{2}; \quad D_{\alpha} = \int_0^1 (x - M_{\alpha})^2 f_{\alpha}(x) dx = \frac{1}{12}. \quad (2.3)$$

Базовая случайная величина также обладает следующими свойствами.

Свойство 1. Случайные величины α и $1-\alpha$ эквивалентны.

Обозначим через ε дискретную случайную величину, принимающую два возможных значения 0 и 1 с одинаковой вероятностью:

$$P\{\varepsilon = 0\} = P\{\varepsilon = 1\} = \frac{1}{2}. \quad (2.4)$$

Представим случайную величину α в виде бесконечной двоичной дроби:

$$\alpha = 0.\varepsilon_1\varepsilon_2\dots = \sum_{k=1}^{\infty} \varepsilon_k 2^{-k}, \quad (2.5)$$

где ε_k – случайные величины, принимающие два возможных значения: 0 или 1, $k = 1, 2, \dots$.

Свойство 2. Если случайная величина α равномерно распределена на отрезке $[0, 1]$, то случайные величины $\varepsilon_1, \varepsilon_2, \dots$ в представлении (2.5) независимы и подчиняются распределению (2.4).

Свойство 3. Если случайные величины $\varepsilon_1, \varepsilon_2, \dots$ независимы и распределены по закону (2.4), то формула (2.5) дает случайную величину, равномерно распределенную на $[0, 1]$.

Среди способов моделирования базовой случайной величины наиболее известны следующие три: *табличный, физический, численный (арифметический)*.

Табличный способ предполагает использование специальных таблиц, содержащих реализации либо самой случайной величины α , либо случайных величин ε , при помощи которых затем можно получать реализации α согласно приведенному выше выражению (2.5). Сами таблицы составляются при помощи специально разработанных устройств. Для использования такого способа генерации на практике данные из таблицы считываются в память компьютера и далее используются для моделирования. Однако использование данного способа не получило широкого распространения в основном по двум причинам:

- запас чисел в таблице как правило ограничен;
- для хранения информации о таблице необходимы значительные затраты памяти ЭВМ.

Физический способ предполагает использование специального устройства, оформленного в виде приставки к ЭВМ. Характер работы устройства может быть самым разнообразным, начиная с механических устройств и заканчивая устройствами, использующими шумы радиоэлектронных ламп и явление радиоактивного распада. Данный способ генерации базовой случайной величины также обладает рядом существенных недостатков, среди которых можно выделить следующие:

- невозможность точного повторного воспроизведения последовательности сгенерированных реализаций;
- необходимость многократных тестовых проверок генератора;
- необходимость создания специального устройства для генерации случайных чисел.

Численный или арифметический способ появился во многом благодаря недостаткам первых двух способов, использование которых было затруднено по указанным выше причинам. Численные методы моделирования случайных чисел сводятся, как правило, к использованию некоторого детерминированного математического алгоритма, который, будучи реализованным в виде программы, позволяет получать каждое новое число на основе одного или нескольких сгенерированных ранее чисел. Достоинством арифметических генераторов являются:

- возможность повторного воспроизведения ранее сгенерированной выборки;
- необходимость лишь однократной проверки вместо периодического тестирования, необходимого в случае физического способа генерирования;
- высокая скорость получения реализаций базовой случайной величины при минимуме требуемого для генератора объема оперативной памяти.

Числа, вырабатываемые арифметическими генераторами, называются «псевдослучайными». Действительно, если известна некоторая реализация базовой случайной величины, то следующая реализация, получаемая с помощью детерминированного алгоритма, полностью определена и всегда предсказуема. Поэтому получаемые таким образом числа нельзя назвать случайными в полном смысле этого слова. Однако для целей моделирования более важными являются другие свойства генерируемых чисел, а именно:

- 1) равномерность распределения на отрезке $[0,1]$;
- 2) возможность получения независимых выборок базовой случайной величины, что должно подтверждаться прохождением ряда статистических проверок, о которых будет сказано ниже.

Генераторы, соответствующие двум приведенным выше требованиям, могут считаться хорошими арифметическими генераторами.

Одним из наиболее широко используемых классов арифметических генераторов являются *конгруэнтные генераторы*. Алгоритм, лежащий в их основе, представлен следующим выражением:

$$z_i = g(z_{i-1}, z_{i-2}, \dots) \bmod m, \quad (2.6)$$

где g – постоянная детерминированная функция предыдущих значений $z_j; j = i-1, i-2, \dots$, m – модуль, z и m являются неотрицательными целыми числами. Операция ‘mod’ возвращает остаток от целочисленного деления, т. е.

$$y = x \bmod m = x - m[x/m], \quad (2.7)$$

где $[]$ – операция выделения целой части числа. Генерируемые значения для такого генератора лежат в диапазоне $0 \dots m-1$, а реализации базовой случайной величины в данном случае могут быть получены по формуле

$$\alpha_i = z_i / m. \quad (2.8)$$

Важной характеристикой арифметических генераторов является их период. *Периодом арифметического генератора* называется максимальное количество реализаций, после которого начинается повторение ранее сгенерированного набора. Поскольку при моделировании обычно требуется огромное количество случайных чисел, то период разрабатываемых генераторов стремятся сделать как можно большим.

В зависимости от вида функции g , можно выделить следующие типы генераторов:

Многократные рекурсивные генераторы. Для них функция g является линейной и использует q ранее сгенерированных либо заданных в качестве начальных значений:

$$g(z_{i-1}, z_{i-2}, \dots) = a_1 z_{i-1} + a_2 z_{i-2} + \dots + a_q z_{i-q}, \quad (2.9)$$

где $a_1, a_2 \dots a_q$ – константы. В ряде публикаций показано, что при соответствующем подборе этих констант периоды таких генераторов могут достигать значений $m^q - 1$.

Квадратичный конгруэнтный генератор задается следующей функцией:

$$g(z_{i-1}, z_{i-2}, \dots) = a_1 z_{i-1}^2 + a_2 z_{i-1} + a_3.$$

Поскольку для данного генератора z_i зависит только от z_{i-1} , а также учитывая что $0 \leq z_i \leq m-1$, максимальный период таких генераторов равен m .

Линейные конгруэнтные генераторы (ЛКГ) задаются следующей функцией:

$$g(z_{i-1}, z_{i-2}, \dots) = a z_{i-1} + c, \quad (2.10)$$

где a (множитель) и c (приращение) являются неотрицательными целыми числами. ЛКГ были созданы Д. Х. Лемером в 1951 г. и являются од-

ними из первых арифметических генераторов. Если $c > 0$, то ЛКГ называют *смешанным*, а при $c = 0$ – *мультипликативным*. Исторически и по причине своей простоты данные генераторы получили достаточно широкое распространение. Далее рассмотрим более подробно именно этот тип арифметических генераторов.

По тем же причинам, что и для квадратичного конгруэнтного генератора, для ЛКГ максимально возможный период не превышает m . Если период ЛКГ равен m , то говорят, что ЛКГ имеет *полный период*. Для ЛКГ с полным периодом длина генерируемой цепочки различных значений зависит только от выбора m . Если же период не является полным, то длина вышеуказанной цепочки значений будет определяться начальным значением z_0 и в этом случае говорят о *периоде начального значения*.

Наиболее важной задачей при создании ЛКГ является задача выбора его коэффициентов (a, c, m) . Кроме приведенных выше критериев, обеспечивающих качество генерируемых последовательностей случайных чисел, выбор коэффициентов ЛКГ выполняется также и исходя из ряда посылок, связанных непосредственно с алгоритмом генерации, лежащим в его основе.

Во-первых, исходя из максимально возможного периода ЛКГ и диапазона генерируемых им значений, очевидно, что при использовании ЛКГ можно получить лишь значения α равные $0, 1/m \dots (m-1)/m$ и нет возможности получать остальные значения из диапазона $[0, 1]$. По этой причине при создании ЛКГ модуль m стараются выбрать как можно больше с целью увеличения плотности заполнения диапазона $[0, 1]$ генерируемыми точками. Поскольку в большинстве используемых в настоящее время ЭВМ для хранения целых чисел используется 32 бита, а крайний левый бит является знаковым, то часто выбирают $m = 2^b$, где $b = 31$. Такой выбор также позволяет избежать достаточно медленного явного деления на m и использовать вместо этого операцию переполнения целых чисел. Наибольшее целое число, которое может быть представлено в b -разрядной ЭВМ, равно $2^b - 1$. Любая попытка сохранить большее целое число W завершится потерей всех, расположенных слева, двоичных знаков, порядковый номер которых превосходит b . Оставшиеся b двоичных знаков составляют ровно $W \bmod 2^b$.

Во-вторых, для практических целей предпочтительно использовать ЛКГ с полным периодом, поскольку в данном случае гарантируется получение каждого значения из набора $0 \dots m-1$ один раз в течение цикла генерации. По этой причине полезно знать, как выбирать коэффициенты

a , c , m так, чтобы у ЛКГ был полный период. Такое правило выбора коэффициентов дается в теореме, доказанной в 1962 г. Халлом и Доббелом.

Теорема 2.1. ЛКГ, определенный по формуле (2.10), имеет полный период тогда и только тогда, когда выполняются следующие три условия:

- 1) единственное положительное целое число, на которое без остатка делятся как m , так и c , равно 1;
- 2) если некоторое число q является простым числом (делится только само на себя и на 1), на которое делится m , то $a-1$ делится на q ;
- 3) если m делится на 4, то $a-1$ тоже делится на 4.

При условии, что $m = 2^b$, согласно теореме 2.1, ЛКГ будет обладать полным периодом, если c будет нечетным, а $a-1$ будет делиться на 4. В ранних работах по ЛКГ акцентировалось внимание на следующем представлении a :

$$a = 2^l + 1, l > 0, \quad (2.11)$$

что приводит к возможности получения $az_{i-1} = 2^l z_{i-1} + z_{i-1}$ и замене умножения на операции сдвига влево на l бит и последующего прибавления z_{i-1} . Однако, как показали последующие исследования, ЛКГ с множителями, представленными выражением (2.11), обладают плохими статистическими характеристиками.

В ходе исследований также выяснилось, что более простые мультипликативные ЛКГ (МЛКГ) зарекомендовали себя не хуже смешанных. По этой причине большинство ЛКГ, применяемых сегодня, являются мультипликативными, поскольку факт превосходства смешанных генераторов в эффективности пока не доказан.

При использовании МЛКГ вначале пытались выбирать модуль $m = 2^b$. Однако было доказано, что максимальный период при этом составляет лишь 2^{b-2} (т. е. только $m/4$) при условии, что z_0 является нечетным, а параметр a задается выражениями вида $8k + 3$ или $8k + 5$, где $k = 0, 1, \dots$. По этим правилам, например, построен известный генератор RANDU, для которого $m = 2^{31}$, $a = 2^{16} + 3 = 65\,539$. Как оказалось, при дальнейших исследованиях приведенные выше правила построения МЛКГ имеют ряд недостатков:

- поскольку максимальный период составляет лишь $m/4$, то не все числа из диапазона $0 \dots m-1$ могут быть сгенерированы. Более того, неизвестно куда попадут сгенерированные числа, что может

привести к появлению недопустимо больших пробелов между соседними по величине реализациями базовой случайной величины;

- при выборе $a = 2^l + j$, где $j > 0$, можно получить неудовлетворительные статистические свойства сгенерированной последовательности.

Из-за вышеуказанных недостатков было предложено отказаться от выбора $m = 2^b$ в МЛКГ. Вместо этого были предложены *МЛКГ с простым модулем*, для которых m и a выбираются по следующим правилам:

- 1) m должно быть наибольшим простым числом меньше 2^b . Например, при $b = 31$, $m = 2^b - 1$;
- 2) a – первообразный элемент по модулю m , то есть наименьшее целое число l , для которого $a^l - 1$ делится на m , составляет $m - 1$.

Если m и a выбраны по этим правилам, то период МЛКГ будет равен $m - 1$.

Задача определения a как первообразного элемента по модулю m оказалась достаточно сложной с вычислительной точки зрения. С целью выбора хороших значений a было проведено множество исследований. Для $m = 2^{31} - 1$ широко используются $a = 7^5 = 16\,807$ и $a = 630\,360\,016$. Первый из этих множителей применяется в известном генераторе Парка – Миллера.

Для проверки качества получаемых с помощью арифметических генераторов случайных чисел используются специально разработанные тесты. Некоторые из них будут рассмотрены ниже.

Тестирование генераторов базовой случайной величины. Критерий «хи-квадрат» разработан, чтобы установить, являются ли получаемые при помощи тестируемого генератора значения $\alpha_i, i = 1, 2, \dots$ равномерно распределенными между 0 и 1. Интервал $[0, 1]$ делится на k подынтервалов одинаковой длины, и затем генерируются реализации базовой случайной величины $\alpha_1, \alpha_2, \dots, \alpha_n$ (как правило, значение n должно равняться, по меньшей мере, 100, а величина n/k должна составлять хотя бы 5). Предположим, что $j = 1, 2, \dots, k$, а f_j будет числом значений α_i , попадающих в j -й подынтервал, а также

$$\chi_n^2 = \frac{k}{n} \sum_{j=1}^k \left(f_j - \frac{n}{k} \right)^2. \quad (2.12)$$

Для больших значений n величина χ_n^2 будет иметь распределение «хи-квадрат» с $k-1$ степенями свободы. При проверке качества сгенерированных чисел фиксируется доверительная вероятность γ (обычно 0.9, 0.95, 0.99 или 0.999) и определяется верхняя критическая точка $\chi^2(k-1, 1-\gamma)$ уровня доверительной вероятности $1-\gamma$, которая является корнем уравнения

$$\int_{\chi^2(k-1, 1-\gamma)}^{\infty} \varphi_{k-1}(x) dx = 1-\gamma, \quad (2.13)$$

где $\varphi_{k-1}(x)$ – плотность распределения χ^2 . На практике решение уравнения (2.13) находят из специальных таблиц. Если вычисленное по формуле (2.12) значение χ_n^2 удовлетворяет неравенству $\chi_n^2 < \chi^2(k-1, 1-\gamma)$, то гипотеза о равномерности распределения смоделированных значений на отрезке $[0, 1]$ принимается.

Для больших значений k существует приближенная формула для расчета верхней критической точки

$$\chi_{k-1, 1-\gamma}^2 \approx (k-1) \left\{ 1 - \frac{2}{9(k-1)} + z_{1-\gamma} \sqrt{\frac{2}{9(k-1)}} \right\}^3, \quad (2.14)$$

где $z_{1-\gamma}$ – верхняя критическая точка уровня доверительной вероятности $1-\gamma$ нормального распределения $N(0,1)$.

Критерий серий является обобщением проверки по критерию «хи-квадрат» до высших измерений. Если значения $\alpha_i, i=1, 2, \dots$ являются независимыми реализациями базовой случайной величины, то непересекающиеся d -мерные последовательности $A_1 = (\alpha_1, \alpha_2, \dots, \alpha_d)$, $A_2 = (\alpha_{d+1}, \alpha_{d+2}, \dots, \alpha_{2d})$, ... должны представлять собой независимые случайные векторы, равномерно распределенные в d -мерном единичном гиперкубе $[0,1]^d$. Для получения численной оценки по критерию серий разделим интервал $[0, 1]$ на k одинаковых подынтервалов и сгенерируем случайные вектора $A_1, A_2, \dots, A_n$ (для которых требуется $n \cdot d$ значений α_i). Пусть $f_{j_1 j_2 \dots j_d}$ будет числом $A_i, i=1, 2, \dots, n$, содержащим первый компонент в подынтервале j_1 , второй компонент в подынтервале j_2 и т. д. Если построить величину

$$\chi_n^2(d) = \frac{k^d}{n} \sum_{j_1=1}^k \sum_{j_2=1}^k \dots \sum_{j_d=1}^k \left(f_{j_1 j_2 \dots j_d} - \frac{n}{k^d} \right)^2, \quad (2.15)$$

то она будет иметь приближенное распределение «хи-квадрат» с $k^d - 1$ степенями свободы (желательно, чтобы $n/k^d \geq 5$). Проверка d -мерной равномерности распределения выполняется точно так же, как и проверка для обычного критерия «хи-квадрат».

Критерий серий позволяет косвенно проверить независимость отдельных значений α_i , поскольку, если отдельные значения α_i коррелированы, то распределение d -мерных векторов будет отклоняться от равномерности. Например, если смежные значения α_i стремятся к положительной корреляции, то пары (α_i, α_{i+1}) имеют тенденцию собираться вдоль нисходящей диагонали единичного квадрата, что отражается на значении $\chi_n^2(2)$.

Критерий восходящих серий является более прямой проверкой допущения о независимости сгенерированных значений базовой случайной величины. Равномерность распределения по данному критерию не проверяется. Для построения численного значения данного критерия рассмотрим некоторую последовательность сгенерированных значений $\alpha_i, i = 1, 2, \dots$. Данную последовательность всегда можно разбить на подпоследовательности, в пределах которых идущие друг за другом значения α_i возрастают. Такая подпоследовательность называется *восходящей серией*. После выделения восходящих серий вычисляется их длина, которая равна числу принадлежащих серии значений α_i . Далее подсчитывается число восходящих серий с длиной 1, 2, 3, 4, 5 и ≥ 6 и вычисляется величина

$$r_i = \begin{cases} \text{число восходящих серий длиной } i & \text{для } i = 1, \dots, 5; \\ \text{число восходящих серий длиной } \geq 6 & \text{для } i = 6 \end{cases}.$$

Тогда численное значение критерия восходящих серий представляет собой величину

$$R = \frac{1}{n} \sum_{i=1}^6 \sum_{j=1}^6 a_{ij} (r_i - nb_i)(r_j - nb_j),$$

где a_{ij} и b_i являются соответственно матрицей и вектором констант, значения которых приводятся в литературе. Для больших значений n ($n \geq 4000$) статистический критерий R будет иметь приближенное распределение «хи-квадрат» с 6-ю степенями свободы, и, соответственно, его проверка может быть выполнена при помощи таблиц для распределения «хи-квадрат».

Критерий Колмогорова – Смирнова позволяет сравнить функцию эмпирического распределения с функцией предполагаемого распределе-

ния F . В случае если известны все параметры рассматриваемого распределения, преимуществом данного критерия является его точная достоверность для любого объема выборки n . Для определения статистики, лежащей в основе данного критерия, будем использовать ступенчатую эмпирическую функцию распределения:

$$F_n(x_i) = \frac{i}{n}, i = 1, 2, \dots, n. \quad (2.16)$$

Статистикой распределения Колмогорова – Смирнова D_n будет являться наибольшее (вертикальное) расстояние между F и F_n (рис. 2.1)

$$D_n = \max\{D_n^+, D_n^-\}, \quad (2.17)$$

где $D_n^+ = \max_{1 \leq i \leq n} \left\{ \frac{i}{n} - F(x_i) \right\}$, $D_n^- = \max_{1 \leq i \leq n} \left\{ F(x_i) - \frac{i-1}{n} \right\}$, x_i – отсортирован-

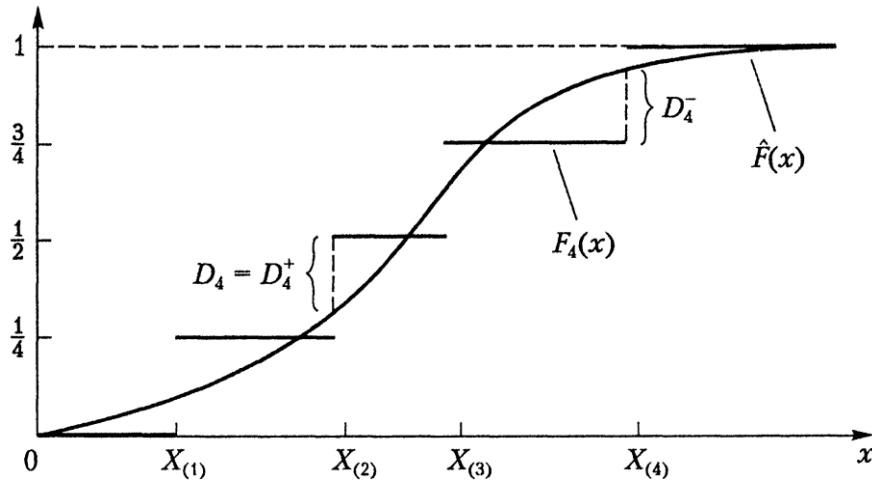


Рис. 2.1. Графическое представление вычисления статистики Колмогорова – Смирнова для D_n при $n = 4$

ная по возрастанию последовательность сгенерированных значений базовой случайной величины.

Большое значение D_n указывает на неудовлетворительное согласие, так что форма критерия должна опровергать гипотезу о равномерности распределения полученных реализаций базовой случайной величины, если D_n превышает некоторую константу $d_{n,1-\gamma}$, где γ – указанный уровень критерия. Численное значение критической точки $d_{n,1-\gamma}$ зависит от того, как было определено предполагаемое распределение F .

Если ни один из параметров F не оценивался каким-либо образом по данным, и в предположении, что функция F является непрерывной,

распределение D_n не зависит от F . Этот факт означает, что при всех формах непрерывных распределений достаточно одной таблицы значений для $d_{n,1-\gamma}$. Для облегчения проверки данного критерия Стефенс разработал точное приближение, согласно которому вместо сравнения D_n и $d_{n,1-\gamma}$ проверяют выполнение следующего неравенства:

$$\left(\sqrt{n} + 0,12 + \frac{0,11}{\sqrt{n}} \right) D_n \leq c_{1-\gamma}. \quad (2.18)$$

Значения $c_{1-\gamma}$ зависят только от γ и могут быть найдены в специальных таблицах.

Л е к ц и я 3

СТАНДАРТНЫЕ МЕТОДЫ МОДЕЛИРОВАНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Метод обратных функций. Предположим, что необходимо сгенерировать реализацию некоторой непрерывной случайной величины ξ . Пусть непрерывная и строго возрастающая функция F является функцией распределения случайной величины ξ . Тогда если известна функция F^{-1} , которая является обратной для функции F , то согласно *методу обратных функций* реализация X случайной величины ξ может быть сгенерирована по следующим правилам:

- 1) генерируем реализацию базовой случайной величины α ;
- 2) реализация случайной величины ξ будет равна: $X = F^{-1}(\alpha)$.

Необходимо отметить, что значение X будет всегда определено и однозначно, поскольку функция F определена в любой точке области $[0, 1]$ и монотонно возрастает. Для того чтобы доказать, что случайная величина, реализации которой сгенерированы по методу обратных функций, действительно имеет функцию распределения F , приведем следующую теорему:

Теорема 3.1. Случайная величина ξ , удовлетворяющая уравнению $\xi = F^{-1}(\alpha)$, где α базовая случайная величина, имеет функцию распределения $F(x)$.

Доказательство. Теорема будет доказана, если для любого вещественного числа x справедливо выражение $P(\xi \leq x) = F(x)$ (по определению функции распределения). Учитывая обратимость $F(x)$,

$$P(\xi \leq x) = P(F^{-1}(\alpha) \leq x) = P(\alpha \leq F(x)) = F(x).$$

Справедливость последнего равенства следует из определения функции распределения для базовой случайной величины в области $[0, 1]$ (формула (2.1)).

Теорема доказана.

Если обратная функция F^{-1} может быть получена аналитически, то реализации случайной величины ξ получают прямо из уравнения, приведенного выше. В противном случае возможно либо численное решение уравнения $F(X) = \alpha$, либо применение для генерации ξ альтернативных методов.

Рассмотрим применение метода обратных функций на примерах.

Пример 3.1. Пусть случайная величина ξ подчиняется экспоненциальному закону распределения

$$F(x) = 1 - e^{-\lambda(x-a)}, \quad x \geq a.$$

Тогда в соответствии с методом обратных функций реализации случайной величины ξ могут быть получены из уравнения $1 - e^{-\lambda(x-a)} = \alpha$, откуда $x = a - \ln(1 - \alpha) / \lambda = a - \ln(\alpha) / \lambda$. Замена $1 - \alpha$ на α возможна благодаря тому, что эти случайные величины распределены одинаково.

Пример 3.2. Пусть случайная величина ξ подчиняется нормальному закону распределения со средним 0 и дисперсией 1. Функция плотности распределения для такой случайной величины равна

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}},$$

а интегральная функция распределения соответственно имеет вид

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

В данном случае требуемое уравнение

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du = \alpha$$

в явном виде неразрешимо. По этой причине для моделирования данной случайной величины следует применять другие методы.

Метод суперпозиции. Метод суперпозиции является методом, позволяющим повысить эффективность моделирования распределений за счет представления их в виде некоторой комбинации более простых распределений и использования многоэтапности. Предположим, что функция распределения случайной величины ξ представлена в виде

$$F(x) = \sum_{j=1}^{\infty} p_j F_j(x), \quad (3.1)$$

где $F_j(x)$ также являются функциями распределения некоторых других случайных величин, $p_j \geq 0$ и соответственно $\sum_{j=1}^{\infty} p_j = 1$. Введем дискретную случайную величину η , которая принимает значения $1, 2, \dots$ с вероятностями $p_1, p_2, \dots$. Тогда справедлива следующая теорема.

Теорема 3.2. Пусть α_1 и α_2 независимые базовые случайные величины. Если с помощью α_1 разыгрывать значение j дискретной величины η , а затем из уравнения $F_j(x) = \alpha_2$ определять реализации случайной величины ξ , то функция распределения ξ будет задана выражением (3.1).

Доказательство. Для доказательства воспользуемся теоремой о полной вероятности и вычислим функцию распределения для случайной величины ξ :

$$P(\xi \leq x) = \sum_{j=1}^{\infty} P(\xi \leq x | \eta = j) P(\eta = j) = \sum_{j=1}^{\infty} F_j(x) p_j = F(x).$$

Теорема доказана.

Рассмотрим применение метода суперпозиции на примере.

Пример 3.3. Пусть случайная величина ξ имеет следующую функцию распределения:

$$F(x) = \frac{1}{12} [5x + (x-1)^5 + 1].$$

Представим эту функцию распределения в виде суперпозиции функций распределения двух других случайных величин ξ_1 и ξ_2

$$F_1(x) = \frac{1}{2}x \text{ и } F_2(x) = \frac{1}{2}[(x-1)^5 + 1],$$

тогда $F(x) = \frac{5}{6}F_1(x) + \frac{1}{6}F_2(x)$. Согласно методу суперпозиции введем дискретную случайную величину η , которая может принимать значение 1 с вероятностью $5/6$ и значение 2 с вероятностью $1/6$. Будем моделировать данную случайную величину следующим образом: моделируем базовую случайную величину α_1 ; если $\alpha_1 < 5/6$, то определяем реализацию случайной величины ξ_1 методом обратных функций $X_1 = 2\alpha_2$ и используем ее в качестве реализации ξ . Если $\alpha_1 \geq 5/6$, то определяем реализацию случайной величины ξ_2 методом обратных функций $X_2 = 1 + \sqrt[5]{2\alpha_2 - 1}$ и используем ее в качестве реализации ξ .

Кроме базовой формулировки метода суперпозиции существует также ряд его модификаций. Первой из них является *модифицированный метод суперпозиции*, представленный теоремой 3.3.

Теорема 3.3. Пусть α базовая случайная величина. Если с ее помощью разыгрывать значение j дискретной величины η , а затем по разыгранному значению j определить реализацию случайной величины ξ из уравнения $F_j(x) = \Theta$, где $\Theta = \left(\alpha - \sum_{i=1}^{j-1} p_i \right) p_j^{-1}$, то функция распределения ξ будет задана выражением (3.1).

Справедливость последней теоремы основана на том, что величина Θ , как показано в литературе, является базовой случайной величиной, независимой по отношению к α . Если сказанное выше справедливо, то справедливость последней теоремы напрямую следует из теоремы 3.2. Условие, сформулированное в теореме 3.3, позволяет построить более экономичный алгоритм моделирования по методу суперпозиции, в котором для моделирования каждой реализации случайной величины ξ требуется сгенерировать лишь одну реализацию базовой случайной величины вместо двух. Стоит также отметить, что при малых вероятностях p_i могут возникать отклонения распределения случайной величины Θ от равномерного, что делает использование модифицированного метода суперпозиции невозможным. Рассмотрим использование данного метода на примере.

Пример 3.4. Используем модифицированный метод суперпозиции для построения алгоритма моделирования случайной величины ξ , приведенной в примере 3.3. Для дискретной случайной величины η величина Θ , определенная в теореме 3.3, будет равна

$$\Theta = \begin{cases} \frac{6}{5}\alpha & \text{при } \eta = 1 \\ 6\alpha - 5 & \text{при } \eta = 2 \end{cases}.$$

Тогда если реализация базовой случайной величины $\alpha < 5/6$, то определяем реализацию случайной величины ξ как $X = \frac{12}{5}\alpha$. Если $\alpha \geq 5/6$, то определяем реализацию случайной величины ξ как $X = 1 + \sqrt[5]{12\alpha - 11}$.

Второй модификацией метода суперпозиции является *метод интегральной суперпозиции*. Рассмотрим непрерывную случайную величину Θ в декартовых координатах на плоскости (x, y) . Обозначим плотность

распределения Θ как $f_{\Theta}(x, y)$ и введем случайную величину ξ с плотностью распределения равной

$$f_{\xi}(x) = \int_{-\infty}^{\infty} f_{\Theta}(x, y) dy.$$

Задача состоит в построении алгоритма моделирования для случайной величины ξ .

Моделирование координат точки Θ можно выполнять в любом порядке. Воспользуемся следующим представлением:

$$f_{\Theta}(x, y) = f_{\eta}(y) f_{\xi}(x | y), \quad (3.2)$$

где $f_{\eta}(y)$ плотность распределения некоторой вспомогательной случайной величины η . Если возможно представление (3.2), то моделирование выполняется в следующем порядке:

- 1) моделируется случайная величина η (например, как $F_{\eta}(y) = \alpha_1$);
- 2) моделируется искомая случайная величина ξ (например, как $F_{\xi}(x | y) = \alpha_2$).

Чаще всего данный метод применяется в случае, если плотность распределения $f_{\xi}(x)$ задана в виде интеграла по параметру.

Пример 3.5. Плотность распределения случайной величины ξ задана в виде

$$f(x) = n \int_1^{\infty} y^{-n} e^{-xy} dy, \quad n > 0, \quad 0 < x < \infty.$$

В данном случае $f_{\Theta}(x, y) = ny^{-n} e^{-xy}$. Тогда

$$f_{\eta}(y) = \int_0^{\infty} f(x, y) dx = \int_0^{\infty} ny^{-n} e^{-xy} dx = ny^{-n-1}, \quad 1 < y < \infty.$$

Согласно формуле полной вероятности $f_{\xi}(x | y) = \frac{f_{\Theta}(x, y)}{f_{\eta}(y)} = ye^{-xy}$.

Соответствующие интегральные функции распределения

$$F_{\eta}(y) = \int_1^y nu^{-n-1} du = n \frac{u^{-n}}{-n} \Big|_1^y = 1 - y^{-n}; \quad F_{\xi}(x | y) = \int_0^x ye^{-ay} da = 1 - e^{-xy}.$$

Тогда по методу обратных функций сначала моделируется η как $y = \alpha_1^{-1/n}$, а затем ξ как $x = -\frac{1}{y} \ln(\alpha_2)$.

Метод суперпозиции часто применяют для моделирования случайных величин, для которых поведение функций плотности распределения

резко отличается в различных подобластях области определения (рис. 3.1).

В данном случае вся область определения разбивается на отдельные подобласти Δ_j $j=1, \dots, m$ и формируются вспомогательные случайные величины с плотностями:

$$f_j(x) = \begin{cases} f(x)/p_j & x \in \Delta_j \\ 0 & x \notin \Delta_j \end{cases},$$

где $p_j = \int_{\Delta_j} f(x) dx$. Очевидно, что $\sum_{j=1}^m p_j = 1$.

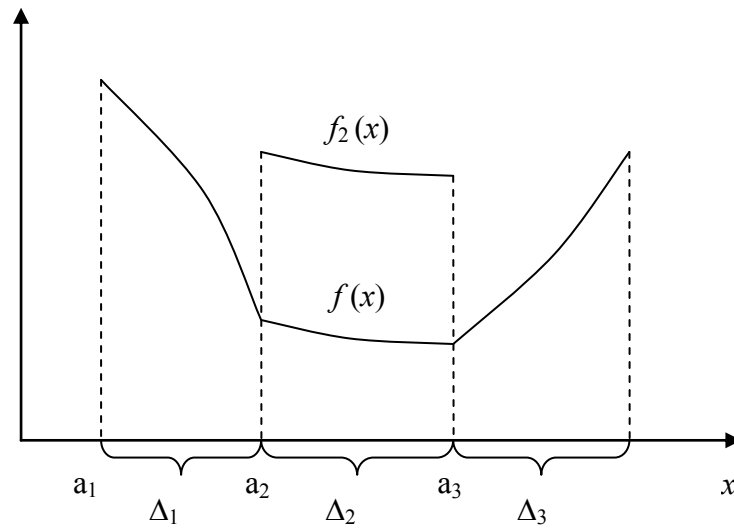


Рис. 3.1. Применение метода суперпозиции для моделирования случайных величин с резко изменяющейся плотностью распределения

Тогда плотность распределения исходной случайной величины представима в виде: $f(x) = \sum_{j=1}^m p_j f_j(x)$ и согласно теореме 3.2 может быть смоделирована по методу суперпозиции. В этом случае сначала будет разыгрываться номер подобласти, а затем моделироваться искомая случайная величина, например, с использованием уравнения

$$\int_{a_j}^x f(u) du = p_j \alpha,$$

где a_j — левый конец Δ_j .

Метод кусочной аппроксимации. Данный метод является приближенным методом моделирования, поскольку предполагает использова-

ние некоторой аппроксимации плотности распределения моделируемой случайной величины.

Пусть необходимо получить реализации X некоторой случайной величины ξ , имеющей плотность распределения $f(x)$ (рис. 3.2). Разобьем область определения $f(x)$ на n интервалов $[a_k, a_{k+1}]$, $k = 0, \dots, n-1$. Тогда вероятность попадания реализации случайной величины ξ в каждый такой интервал будет определена следующим образом:

$$p_k = \int_{a_k}^{a_{k+1}} f(x) dx. \quad (3.3)$$

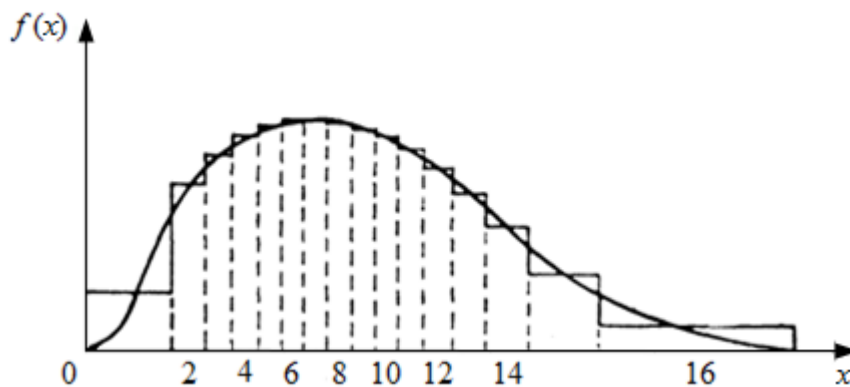


Рис. 3.2. Метод кусочной аппроксимации

Для выполнения процесса моделирования метод кусочной аппроксимации требует нахождения номера интервала, в который попадет реализация случайной величины ξ . Наиболее удобным для практического использования является случай, когда интервалы разбиения сформированы таким образом, что все p_k равны между собой. Если при этом область задания случайной величины ξ разбита на n интервалов, то

$p_k = \frac{1}{n}$. Тогда необходимо разбить интервал $[0,1]$ на n равных частей,

сгенерировать реализацию базовой случайной величины α_1 и определить номер k подынтервала на отрезке $[0, 1]$, в который попала данная реализация. Полученное значение k и будет определять номер интервала в пределах области распределения ξ , куда попадет ее реализация.

Введем для каждого из выделенных интервалов вспомогательные случайные величины η_k , значения которых лежат в пределах соответствующего k -го интервала. Обозначим плотности распределения случай-

ных величин η_k как $\varphi_k(x)$. В общем случае функции $\varphi_k(x)$ могут быть различны. Для того чтобы сгенерированные реализации случайных величин η_k можно было использовать как реализации случайной величины ξ , их плотности распределения $\varphi_k(x)$ в пределах соответствующих интервалов должны быть связаны с плотностью $f(x)$ выражением

$$\varphi_k(x) = \frac{f(x)}{p_k}, \quad (3.4)$$

что ведет к использованию алгоритма метода суперпозиции, изложенного ранее. Однако использование напрямую выражения (3.4) на практике часто оказывается затруднительным, что связано с видом функции $f(x)$. В такой ситуации применяется рассматриваемый метод кусочной аппроксимации, согласно которому $\varphi_k(x)$ заменяются некоторыми аппроксимациями $\varphi_k^*(x)$, от вида которых зависит дальнейший ход моделирования. В остальном ход моделирования будет совпадать с алгоритмом метода суперпозиции.

Рассмотрим ряд аппроксимаций, которые с определенной степенью точности могут использоваться при любом виде плотности распределения $f(x)$.

Наиболее простым вариантом является использование $\varphi_k^* = \text{const}$, которые задаются выражением

$$\varphi_k^*(x) = \frac{1}{a_{k+1} - a_k}.$$

Учитывая, что в данном случае η_k – равномерно распределенная на отрезке $[a_k, a_{k+1}]$ случайная величина, для реализации ξ получаем следующее выражение

$$X = a_k + (a_{k+1} - a_k)\alpha_2,$$

где α_2 – независимая реализация базовой случайной величины.

Если необходимо достичь большей точности при применении данного приближенного метода могут использоваться другие варианты функции $\varphi_k^*(x)$. При линейной аппроксимации $\varphi_k(x)$ на отрезке $[a_k, a_{k+1}]$, $\varphi_k^*(x)$ будет иметь вид

$$\varphi_k^*(x) = s_k + g_k x,$$

где $g_k = \frac{f_{k+1} - f_k}{a_{k+1} - a_k}$ – угловой коэффициент прямой, $f_k = f(a_k)$, s_k может

быть определено из условия нормировки функции $\varphi_k^*(x)$:

$$\int_{a_k}^{a_{k+1}} \varphi_k^*(x) dx = 1.$$

Описанный выше подход к разбиению области определения случайной величины ξ на интервалы, вероятности попадания в которые одинаковы, в ряде случаев может оказаться неприемлемым. Это связано, прежде всего, с тем, что при таком выборе системы интервалов точность аппроксимации функции $f(x)$ неодинакова по всей области определения и зависит от значений $f(x)$ (рис. 3.2). При малых $f(x)$ точность аппроксимации хуже. По этой причине при выборе системы интервалов для разбиения области определения случайной величины ξ следует учитывать необходимость обеспечения заданной точности на интервалах с наименьшими значениями $f(x)$.

Метод Неймана. Данный метод генерации случайных величин был впервые предложен фон Нейманом в 1951 г. Суть метода изложена в условии теоремы 3.4.

Теорема 3.4. Пусть ξ – непрерывная случайная величина, имеющая плотность распределения $f(x)$. Пусть существует некоторая функция $g(x)$, такая, что для любого x выполняется неравенство $f(x) \leq g(x)$, и кроме того, площадь под кривой $g(x)$ конечна, т. е. $S = \int_{-\infty}^{\infty} g(x) dx < \infty$.

Определим вспомогательные случайные величины η , которая имеет плотность распределения $r(x) = g(x)/S$ и ψ , которая равномерно распределена на отрезке $[0, g(x)]$ и соответственно имеет плотность распределения $l(y) = 1/g(x)$. Тогда реализации случайной величины ξ могут быть получены следующим образом:

- 1) получаем реализацию X' случайной величины η ;
- 2) получаем независимо от X' реализацию Y случайной величины ψ ;
- 3) если $Y \leq f(X')$, то реализация X случайной величины ξ равна $X = X'$, иначе переходим к пункту 1.

Доказательство. Покажем, что случайная величина, полученная по алгоритму, сформулированному в условии теоремы, действительно имеет плотность распределения $f(x)$, т. е. $P(X \leq x) = \int_{-\infty}^x f(u) du$.

Поскольку реализация X случайной величины ξ равна реализации X' случайной величины η при условии, что $Y \leq f(X')$, то

$$P(X \leq x) = P(X' \leq x | Y \leq f(X')) = \frac{P(X' \leq x, Y \leq f(X'))}{P(Y \leq f(X'))}. \quad (3.5)$$

Определим вероятности, стоящие в правой части последнего выражения. Учитывая то, что случайные величины η и ψ генерируются независимо, имеем

$$P(X' \leq x, Y \leq f(X')) = \int_{-\infty}^x \frac{g(u)}{S} \int_0^{f(u)} \frac{1}{g(u)} dv du = \frac{1}{S} \int_{-\infty}^x f(u) du,$$

$$P(Y \leq f(X')) = \int_{-\infty}^{\infty} \frac{g(u)}{S} \int_0^{f(u)} \frac{1}{g(u)} dv du = \frac{1}{S} \int_{-\infty}^{\infty} f(u) du = \frac{1}{S}.$$

Тогда, подставляя последние выражения в (3.5), получаем

$$P(X \leq x) = \int_{-\infty}^x f(u) du.$$

Теорема доказана.

Как видно из алгоритма метода Неймана, некоторые сгенерированные пары точек (X', Y) могут остаться неиспользованными, поскольку не удовлетворяют условию $Y \leq f(X')$, однако получение этих точек требует вычислительных затрат. В связи с этим вводится понятие эффективности метода. *Эффективностью метода Неймана* называется вероятность κ того, что точка X' станет реализацией X случайной величины ξ , а не будет отброшена. Эффективность может быть определена как отношение площади под графиком $f(x)$ к площади под графиком $g(x)$ на заданной области определения. Учитывая, что $f(x)$ является плотностью распределения и площадь под данной кривой равна 1 согласно условию нормировки, эффективность можно представить следующим образом:

$$\kappa = \frac{1}{S},$$

где S — площадь под кривой $g(x)$.

Функцию $g(x)$ стараются выбирать таким образом, чтобы эффективность κ была как можно ближе к 1, что увеличивает долю полезных генераций X' .

Пример 3.6. Случайная величина ξ определена на отрезке $[-R, R]$ и имеет плотность распределения

$$f(x) = \frac{2}{\pi R^2} \sqrt{R^2 - x^2}.$$

Максимальное значение функции $f(x)$ на $[-R, R]$ равно

$$B = \max_{-R \leq x \leq R} f(x) = 2/(\pi R).$$

Тогда плотность распределения вспомогательной случайной величины η представим в виде

$$r(x) = \frac{2}{\pi R S} = \left[S = 2R \cdot \frac{2}{\pi R} \right] = \frac{1}{2R}.$$

Данную случайную величину можно моделировать методом обратных функций:

$$\int_{-R}^x r(u) du = \frac{x + R}{2R} = \alpha_1 \Rightarrow X' = R(2\alpha_1 - 1),$$

где α_1 – реализация базовой случайной величины. Для моделирования второй вспомогательной случайной величины ψ также воспользуемся методом обратных функций:

$$\int_0^y l(v) dv = \int_0^y \frac{\pi R}{2} dv = \frac{\pi R y}{2} = \alpha_2 \Rightarrow Y = \frac{2\alpha_2}{\pi R},$$

где α_2 – отдельная независимая реализация базовой случайной величины. Поскольку согласно теореме 3.4 для получения реализации X случайной величины ξ требуется выполнение условия $Y \leq f(X')$, получаем

$$\frac{2\alpha_2}{\pi R} \leq \frac{2}{\pi R^2} \sqrt{R^2 - [R(2\alpha_1 - 1)]^2} \Rightarrow \alpha_2 \leq \sqrt{1 - (2\alpha_1 - 1)^2}.$$

Таким образом, моделирование случайной величины ξ заключается в следующем. Разыгрываем две независимые реализации α_1, α_2 базовой случайной величины. Если выполняется последнее неравенство, считаем, что $X = X' = R(2\alpha_1 - 1)$. В противном случае разыгрываем следующую пару случайных чисел. Эффективность метода в данном случае равна

$$\kappa = \frac{\pi R}{4R} = \frac{\pi}{4}.$$

Л е к ц и я 4

СПЕЦИАЛЬНЫЕ МЕТОДЫ МОДЕЛИРОВАНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН

Несмотря на то что большинство случайных величин может быть смоделировано с использованием одного из стандартных методов, изложенных ранее, известны также и методы, которые основываются на каком-нибудь специальном свойстве моделируемой случайной величины. Часто таким свойством является возможность представления моделируемой случайной величины в виде комбинации других случайных величин, которые могут быть смоделированы гораздо проще. По причине применения только для моделирования определенных случайных величин специальные методы не имеют какой-либо общей формы. Далее рассмотрим несколько таких методов.

Метод моделирования распределения Эрланга. Задача состоит в моделировании случайной величины ξ , которая подчиняется распределению Эрланга порядка n . Функция плотности распределения такой случайной величины имеет вид

$$f_{n,\lambda}(x) = \begin{cases} 0, & x \leq 0 \\ \frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x}, & x > 0. \end{cases} \quad (4.1)$$

Теорема 4.1. Если случайная величина $\xi = \sum_{i=1}^n \eta_i$, где η_i – независимые случайные величины, имеющие одинаковое экспоненциальное распределение с параметром λ , и функции плотности распределения которых равны

$$\varphi_i(x) = \lambda e^{-\lambda x}, x > 0, i = 1, 2, \dots, n, \quad (4.2)$$

то функция плотности распределения случайной величины ξ будет задаваться выражением (4.1).

Доказательство. Доказательство данной теоремы выполняется методом математической индукции.

При $n=1$ формула (4.1) переходит в (4.2) и, соответственно, утверждение теоремы выполняется.

Предположим, что оно выполняется и при $n=k$, т. е. случайная величина $\xi' = \sum_{i=1}^k \eta_i$ имеет плотность распределения вида

$$f'(x) = \frac{\lambda^k}{(k-1)!} x^{k-1} e^{-\lambda x}, x \geq 0.$$

Найдем плотность распределения случайной величины $\zeta = \xi' + \eta_{k+1}$. Поскольку случайные величины ξ' и η_{k+1} независимы, плотность распределения случайной величины ζ представима в виде свертки

$$f_{\zeta}(x) = \int_{-\infty}^{\infty} \varphi_{k+1}(x-u) f'(u) du = \int_0^x \varphi_{k+1}(x-u) f'(u) du, x \geq 0,$$

где учтено, что $f'(u) = 0$ при $u < 0$ и $\varphi_{k+1}(x-u) = 0$ при $x-u < 0$. Подставляя в последнее выражение значения $\varphi_{k+1}(x-u)$ и $f'(u)$, получим

$$f_{\zeta}(x) = \frac{\lambda^{k+1}}{(k-1)!} e^{-\lambda x} \int_0^x u^{k-1} du = \frac{\lambda^{k+1}}{k!} x^k e^{-\lambda x}, x \geq 0.$$

Аналогично можно продолжить доказательство до включения в сумму η_n .

Теорема доказана.

Случайные величины η_i могут быть легко смоделированы методом обратных функций. Их реализации: $Y_i = -\frac{1}{\lambda} \ln(\alpha_i)$, где α_i – независимые реализации базовой случайной величины. Тогда алгоритм моделирования ξ может быть значительно упрощен. Как следует из выше приведенной теоремы, получив n независимых реализаций $\alpha_1, \dots, \alpha_n$, можно разыграть выборочное значение X случайной величины ξ по формуле:

$$X = -\frac{1}{\lambda} \ln \alpha_1 - \frac{1}{\lambda} \ln \alpha_2 - \dots - \frac{1}{\lambda} \ln \alpha_n = -\frac{1}{\lambda} \ln \prod_{i=1}^n \alpha_i.$$

Последнее выражение для генерации ξ позволяет избежать необходимости вычислять натуральный логарифм n раз. Однако данное выражение должно использоваться с осторожностью при больших n , поскольку число, стоящее под логарифмом, может стать близким к 0, что в свою очередь может привести к проблемам при вычислении логарифма. В этом случае можно, например, разбить выражение на несколько логарифмов.

Методы моделирования нормального распределения. Задача состоит в моделировании случайной величины ζ , распределенной по нормальному закону со средним m и дисперсией σ^2 . Ее плотность распределения

$$f_{\zeta}(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}. \quad (4.3)$$

Для моделирования обычно используют представление

$$\zeta = \sigma\xi + m, \quad (4.4)$$

где ξ представляет собой случайную величину, распределенную по нормальному закону со средним 0 и дисперсией 1. Справедливость (4.4) может быть доказана следующим образом:

$$\begin{aligned} P(\zeta \leq x) &= P\left(\xi \leq \frac{x-m}{\sigma}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(x-m)/\sigma} e^{-\frac{t^2}{2}} dt = \left[t = \frac{u-m}{\sigma}\right] = \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^x e^{-\frac{(u-m)^2}{2\sigma^2}} du. \end{aligned}$$

Таким образом, задача моделирования сводится к моделированию случайной величины ξ , имеющей плотность распределения

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \quad (4.5)$$

Метод, основанный на переходе в полярную систему координат. Одним из способов моделирования такой случайной величины является использование перехода в полярную систему координат. Выберем независимую от ξ случайную величину η , также распределенную по нормальному закону со средним 0 и дисперсией 1. Рассмотрим на плоскости XU случайную точку Q с координатами ξ и η . Плотность распределения такой точки

$$f_Q(x, y) = f_{\xi}(x)f_{\eta}(y) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2}} = \frac{1}{2\pi} e^{-\frac{r^2}{2}} = c(r). \quad (4.6)$$

Представим функцию $f_Q(x, y)$ в полярных координатах. Связь между декартовыми и полярными координатами

$$x = r \cos \varphi$$

$$y = r \sin \varphi$$

и Якобиан перехода к полярным координатам:

$$\begin{vmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi} \end{vmatrix} = r.$$

Тогда

$$f_Q(r, \varphi) = rc(r) = \frac{r}{2\pi} e^{-\frac{r^2}{2}}. \quad (4.7)$$

Как видно из выражения (4.7), $f_Q(r, \varphi)$ не зависит от φ . Область изменения полярных координат для Q : $0 \leq r \leq \infty$, $0 \leq \varphi \leq 2\pi$. Найдем частные плотности распределения случайных величин по отдельным полярным координатам ρ и Θ :

$$f_\rho(r) = \int_0^{2\pi} f_Q(r, \varphi) d\varphi = \int_0^{2\pi} rc(r) d\varphi = 2\pi rc(r).$$

Из условия нормировки

$$\int_0^\infty \int_0^{2\pi} rc(r) d\varphi dr = \int_0^{2\pi} d\varphi \int_0^\infty rc(r) dr = 2\pi \int_0^\infty rc(r) dr = 1 \Rightarrow f_\Theta(\varphi) = \int_0^\infty rc(r) dr = \frac{1}{2\pi}.$$

Исходя из полученных выражений для $f_\rho(r)$ и $f_\Theta(\varphi)$, можно сделать вывод, что $f_Q(r, \varphi) = f_\rho(r)f_\Theta(\varphi)$, а следовательно, случайные величины ρ и Θ независимы, и поэтому их можно моделировать отдельно, например, методом обратных функций:

$$\begin{aligned} \int_0^a f_\rho(r) dr &= \int_0^a 2\pi rc(r) dr = \int_0^a re^{-\frac{r^2}{2}} dr = -\int_0^a e^{-\frac{r^2}{2}} d\left(-\frac{r^2}{2}\right) = 1 - e^{-\frac{a^2}{2}} = \\ &= \alpha_1 \Rightarrow a = \sqrt{-2\ln(\alpha_1)}; \end{aligned}$$

$$\int_0^h f_\Theta(\varphi) d\varphi = \frac{1}{2\pi} \int_0^h d\varphi = \frac{h}{2\pi} = \alpha_2 \Rightarrow h = 2\pi\alpha_2,$$

где a и h – реализации случайных величин ρ и Θ соответственно. Переходя обратно к декартовым координатам, получаем следующие выражения для реализаций X и Y случайных величин ξ и η :

$$\begin{aligned} X &= a \cos(\varphi) = \sqrt{-2\ln(\alpha_1)} \cos(2\pi\alpha_2); \\ Y &= a \sin(\varphi) = \sqrt{-2\ln(\alpha_1)} \sin(2\pi\alpha_2). \end{aligned} \quad (4.8)$$

Формулы (4.8) позволяют найти сразу две реализации случайной величины, распределенной по нормальному закону с параметрами $(0,1)$. Однако данные формулы содержат операции $\ln$, $\sin$, $\cos$, которые выполняются достаточно долго. Поэтому в ряде случаев применение данного метода может оказаться неприемлемым с точки зрения вычислительных затрат.

Метод, основанный на применении центральной предельной теоремы теории вероятностей. Для построения алгоритма моделирования случайной величины ξ , подчиняющейся нормальному распределению с параметрами $(0,1)$, введем понятие нормированной случайной величины. Для случайной величины η нормированная случайная величина определяется как

$$\eta_0 = (\eta - M_\eta) / \sigma_\eta,$$

где M_η и σ_η – соответственно математическое ожидание и среднеквадратическое отклонение случайной величины η . Пусть случайная величина $\xi^{(n)}$ определяется как сумма n независимых нормированных базовых случайных величин:

$$\xi^{(n)} = \sum_{i=1}^n \frac{\alpha_i - 1/2}{\sqrt{1/12}} = \sqrt{3} \sum_{i=1}^n (2\alpha_i - 1). \quad (4.9)$$

Математическое ожидание такой случайной величины равно 0, а ее дисперсия равна n . Нормируя $\xi^{(n)}$, получим

$$\xi_0^{(n)} = \sqrt{\frac{3}{n}} \sum_{i=1}^n (2\alpha_i - 1). \quad (4.10)$$

Пусть $F_0^{(n)}(x)$ функция распределения случайной величины $\xi_0^{(n)}$. Тогда согласно центральной предельной теореме

$$\lim_{n \rightarrow \infty} F_0^{(n)}(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{u^2}{2}} du.$$

Как следствие при достаточно большом n для моделирования можно применять выражение (4.10). На практике вполне достаточно ограничиться значением $n=12$. Тогда реализации X случайной величины ξ можно получать по формуле

$$X = \sum_{i=1}^{12} \alpha_i - 6.$$

Моделирование случайных векторов. Все методы моделирования, представленные ранее, могут применяться для моделирования отдельных случайных величин с некоторым заданным распределением. Однако при построении некоторых имитационных моделей возникает необходимость моделировать случайные вектора вида $X = (X_1, X_2, \dots, X_n)$ из заданного совместного (или многомерного) распределения, где отдельные координаты такого вектора могут быть зависимыми либо независимыми случайными величинами. Предположим, что случайные величины $\xi_1, \xi_2, \dots, \xi_n$ имеют совместную плотность распределения $f(x_1, x_2, \dots, x_n)$. Для моделирования вышеуказанных случайных величин можно использовать условные плотности распределения, которые могут быть найдены следующим образом:

$$f_1(x_1) = \int \dots \int_{\substack{-\infty \\ n-1}}^{\infty} f(x_1, x_2, \dots, x_n) dx_2 \dots dx_n;$$

$$f_2(x_2 | x_1) = \int \dots \int_{\substack{-\infty \\ n-2}}^{\infty} f(x_1, x_2, \dots, x_n) dx_3 \dots dx_n / f_1(x_1); \quad (4.11)$$

...

$$f_n(x_n | x_1, \dots, x_{n-1}) = f(x_1, x_2, \dots, x_n) / [f_1(x_1) \dots f_{n-1}(x_{n-1} | x_1, \dots, x_{n-2})].$$

При необходимости использования для моделирования отдельных случайных величин ξ_i методов, оперирующих с функциями распределения (например, метод обратных функций), можно ввести условные функции распределения:

$$F_i(x_i | x_1, \dots, x_{i-1}) = \int_{-\infty}^{x_i} f_i(u | x_1, \dots, x_{i-1}) du. \quad (4.12)$$

Моделирование случайного вектора в данном случае сводится к последовательному моделированию его координат, т. е. по $f_1(x_1)$ (или по $F_1(x_1)$) моделируется реализация ξ_1 , затем по $f_2(x_2 | x_1)$ определяется реализация ξ_2 и так далее до тех пор, пока не будет получена реализация последней случайной величины ξ_n . Стоит отметить, что поскольку переменные функции $f(x_1, x_2, \dots, x_n)$ могут быть занумерованы $n!$ способами, то существует столько же способов представления условных плотностей (4.11). Часто порядок нумерации определяет простоту последующего моделирования отдельных случайных величин ξ_i , а следовательно, и простоту моделирования случайного вектора в целом.

Если случайные величины, определяющие координаты случайного вектора, являются независимыми, то процесс моделирования упрощается, поскольку в этом случае $F(x_1, x_2, \dots, x_n) = \prod_{i=1}^n F_i(x_i)$ и отдельные координаты случайного вектора могут быть смоделированы независимо друг от друга.

Рассмотрим процесс моделирования случайных векторов на нескольких примерах.

Пример 4.1. Случайный вектор Q с декартовыми координатами $\xi_1, \xi_2, \dots, \xi_n$ распределен равномерно в n -мерном параллелепипеде $\Pi\{a_i < x_i < b_i; i = 1, 2, \dots, n\}$. Тогда совместная плотность распределения для такого случайного вектора будет равна

$$f_Q(x_1, x_2, \dots, x_n) = \begin{cases} c, & (x_1, x_2, \dots, x_n) \in \Pi \\ 0, & (x_1, x_2, \dots, x_n) \notin \Pi \end{cases}$$

где $\frac{1}{c} = \prod_{i=1}^n (b_i - a_i)$ – объем параллелепипеда.

Проинтегрировав $f_Q(x_1, x_2, \dots, x_n)$ по всем переменным, кроме x_i , получим функции распределения по отдельным координатам

$$f_i(x_i) = \begin{cases} \frac{1}{b_i - a_i}, & x_i \in [a_i, b_i] \\ 0, & x_i \notin [a_i, b_i] \end{cases}$$

Тогда отдельные координаты вектора можно смоделировать методом обратных функций:

$$F_i(x_i) = \frac{x_i - a_i}{b_i - a_i} = \alpha_i \Rightarrow x_i = a_i + \alpha_i(b_i - a_i).$$

Пример 4.2. Случайный вектор $Q(\xi, \eta)$ может принимать значения в треугольнике (рис. 4.1): $x + y < 1$, $x > 0$, $y > 0$.

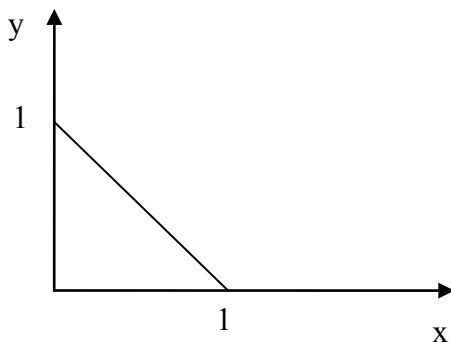


Рис. 4.1. Область значений случайной величины $Q(\xi, \eta)$

Совместная плотность распределения случайных величин ξ и η задана функцией $f(x, y) = 6x$. Рассмотрим два возможных случая нумерации переменных:

Пусть первой переменной будет ξ . Тогда

$$f_\xi(x) = \int_0^{1-x} f(x, y) dy = 6x(1-x)$$

при $0 < x < 1$;

$$f_\eta(y|x) = \frac{f(x, y)}{f_\xi(x)} = \frac{1}{1-x}$$

при $0 < y < 1-x$.

Соответствующие интегральные функции распределения случайных величин ξ и η будут равны

$$F_{\xi}(x) = \int_0^x f_{\xi}(u) du = 3x^2 - 2x^3 \text{ при } 0 < x < 1,$$

$$F_{\eta}(y|x) = \int_0^y f_{\eta}(v|x) dv = \frac{y}{1-x} \text{ при } 0 < y < 1-x.$$

После применения метода обратных функций для получения реализации X случайной величины ξ можно использовать выражение $3X^2 - 2X^3 = \alpha_1$, и затем для получения реализации Y случайной величины η : $Y = (1-X)\alpha_2$.

Теперь рассмотрим альтернативный вариант построения алгоритма моделирования случайного вектора и в качестве первой переменной выберем η . В этом случае плотности распределения для случайных величин ξ и η будут иметь вид

$$f_{\eta}(y) = \int_0^{1-y} f(x, y) dx = 3(1-y)^2 \text{ при } 0 < y < 1,$$

$$f_{\xi}(x|y) = \frac{f(x, y)}{f_{\eta}(y)} = \frac{2x}{(1-y)^2} \text{ при } 0 < x < 1-y.$$

Соответствующие интегральные функции распределения:

$$F_{\eta}(y) = \int_0^y f_{\eta}(v) dv = 1 - (1-y)^3 \text{ при } 0 < y < 1;$$

$$F_{\xi}(x|y) = \int_0^x f_{\xi}(u|y) du = \frac{x^2}{(1-y)^2} \text{ при } 0 < x < 1-y.$$

Учитывая свойство базовой случайной величины о том, что α и $1-\alpha$ эквивалентны и применяя метод обратных функций, получаем следующие выражения для моделирования случайных величин η и ξ :

$$Y = 1 - \sqrt[3]{\alpha_1}; \quad X = (1-Y)\sqrt{\alpha_2}.$$

Как видно из результирующих выражений для моделирования, полученных при различном порядке следования переменных случайного вектора $Q(\xi, \eta)$, во втором случае можно напрямую использовать метод обратных функций для обеих переменных, в то время как для первого варианта это не так. Поэтому с точки зрения вычислительных затрат второй порядок следования переменных выглядит предпочтительнее.

3. МОДЕЛИРОВАНИЕ СЛУЧАЙНЫХ ПРОЦЕССОВ

Л е к ц и я 5

ОПИСАНИЕ СЛУЧАЙНЫХ ПРОЦЕССОВ И ОБЩИЕ МЕТОДЫ ИХ МОДЕЛИРОВАНИЯ

При моделировании различных систем в ряде случаев возникает необходимость получения реализаций случайной величины с учетом ее изменения во времени. Например, может понадобиться генерировать последовательность величин, характеризующих размер поступающих сообщений в системе телекоммуникаций. Если предполагается, что реализации такой случайной величины в разные моменты времени будут независимыми и одинаково распределенными, то задача моделирования может быть сведена к получению реализаций из соответствующего одномерного распределения при помощи методов, которые были рассмотрены ранее. Однако если между наблюдениями в различные моменты времени существует некоторая зависимость, то применение ранее рассмотренных методов будет некорректным, поскольку не будет учитываться взаимосвязь между реализациями с течением времени. Далее будут рассмотрены методы моделирования случайных процессов, которые позволяют учитывать такие связи.

Под *случайным процессом* будем понимать функцию $\xi(t)$, значения которой при фиксированном t являются случайными величинами, $t \in \Omega = [T_0, T]$.

В то время как случайная величина определяется множеством ее значений и распределением вероятностей на этом множестве, случайный процесс характеризуется множеством функций времени и вероятностной мерой, заданной на этом множестве функций. Каждая такая функция времени называется *реализацией случайного процесса* $\xi(t)$. Множество значений параметра t называют *областью определения случайного процесса*, а множество значений, которые может принимать функция $\xi(t)$ — *пространством значений случайного процесса*.

Для описания случайных процессов используются характеристики, аналогичные тем, которые использовались для описания случайных величин. Далее кратко рассмотрим основные способы описания случайных процессов, которые применяются на практике.

По аналогии с описанием непрерывных случайных величин при помощи интегральной функции распределения и функции плотности распределения случайный процесс может быть описан аналогичными характеристиками. Поскольку в отличие от случайных величин случайный процесс является функцией переменной t , то в общем случае он описывается многомерными функциями распределения. Рассмотрим n точек параметра t $t_1, t_2, \dots, t_n$. Тогда n -мерная интегральная функция распределения случайного процесса будет иметь вид

$$F_n(x_1, t_1; \dots; x_n, t_n) = P\{\xi(t_1) \leq x_1; \dots; \xi(t_n) \leq x_n\}. \quad (5.1)$$

n -мерная функция плотности вероятности $f_n(x_1, t_1; \dots; x_n, t_n)$ представляет собой вероятность того, что значения случайного процесса $\xi(t)$ в n точках $t_1, t_2, \dots, t_n$ заключены в соответствующих малых интервалах $[x_1, x_1 + dx_1), [x_2, x_2 + dx_2), \dots, [x_n, x_n + dx_n)$. Данная вероятность равна $f_n(x_1, t_1; \dots; x_n, t_n) dx_1 \dots dx_n$. Функция $f_n(x_1, t_1; \dots; x_n, t_n)$ позволяет судить о связи между вероятными значениями случайного процесса в n произвольных точках аргумента t . Функция плотности распределения $f_n(x_1, t_1; \dots; x_n, t_n)$ может быть получена как частная производная степени n от n -мерной интегральной функции распределения рассматриваемого случайного процесса:

$$f_n(x_1, t_1; \dots; x_n, t_n) = \frac{\partial^n}{\partial x_1 \dots \partial x_n} F_n(x_1, t_1; \dots; x_n, t_n). \quad (5.2)$$

Необходимо отметить, что при описании случайного процесса с помощью многомерных функций распределения, приведенных выше, полная информация о процессе может быть получена, если эти многомерные функции распределения известны для любого n произвольно выбранных значений $t_1, t_2, \dots, t_n$.

Рассмотренный выше способ описания случайных процессов позволяет теоретически полностью описать рассматриваемый случайный процесс, однако для практических целей такой способ описания часто оказывается слишком громоздким и неудобным для использования. Кроме того, для случайных процессов, которые встречаются на практике, построение многомерных функций распределения является в большинстве случаев достаточно сложной задачей, а зачастую и вообще не представляется возможным. По этой причине для описания случайных процессов часто используют моментные функции. По аналогии с моментами случайных величин различают начальные и центральные моментные функции.

Под начальными моментными функциями случайного процесса $\xi(t)$, заданного на некотором интервале, понимают математические ожидания соответствующих произведений:

$$\alpha_k(t) = M\{\xi^k(t)\} = \int_{-\infty}^{\infty} x^k f(x, t) dx; \quad (5.3)$$

$$\alpha_{k,m}(t_1, t_2) = M\{\xi^k(t_1) \xi^m(t_2)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1^k x_2^m f(x_1, t_1; x_2, t_2) dx_1 dx_2,$$

где $f(x, t)$ и $f(x_1, t_1; x_2, t_2)$ соответственно одномерная и двумерная плотности распределения случайного процесса $\xi(t)$. Для дальнейшего использования наиболее важную роль будут играть первый начальный момент или *математическое ожидание случайного процесса*:

$$m_\xi(t) = \alpha_1(t) = M\{\xi(t)\} = \int_{-\infty}^{\infty} x f(x, t) dx \quad (5.4)$$

и начальный момент $\alpha_{1,1}(t_1, t_2)$, который называют *корреляционной функцией случайного процесса*:

$$R_\xi(t_1, t_2) = \alpha_{1,1}(t_1, t_2) = M\{\xi(t_1) \xi(t_2)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f(x_1, t_1; x_2, t_2) dx_1 dx_2. \quad (5.5)$$

Кроме начальных моментов можно рассматривать также центральные моменты, которые определяются следующими выражениями:

$$\mu_k(t) = M\{[\xi(t) - m_\xi(t)]^k\} = \int_{-\infty}^{\infty} [x - m_\xi(t)]^k f(x, t) dx;$$

$$\mu_{k,m}(t_1, t_2) = M\{[\xi(t_1) - m_\xi(t_1)]^k [\xi(t_2) - m_\xi(t_2)]^m\} =$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [x_1 - m_\xi(t_1)]^k [x_2 - m_\xi(t_2)]^m f(x_1, t_1; x_2, t_2) dx_1 dx_2. \quad (5.6)$$

Из центральных моментов особо стоит выделить $D_\xi(t) = \mu_2(t)$, называемый *дисперсией случайного процесса*, а также $\mu_{1,1}(t_1, t_2)$, называемый его *ковариационной функцией*. Как следует из приведенных выше выражений, ковариационная и корреляционная функции случайного процесса связаны между собой следующим соотношением:

$$\mu_{1,1}(t_1, t_2) = R_\xi(t_1, t_2) - m_\xi(t_1) m_\xi(t_2). \quad (5.7)$$

Стоит также отметить, что в отличие от $\xi(t)$, представленные выше моментные функции являются детерминированными функциями времени.

Прежде чем непосредственно перейти к рассмотрению алгоритмов моделирования случайных процессов, выделим некоторые их типы, понятия которых потребуются в дальнейшем.

Случайный процесс называется *стационарным в узком смысле*, если все его конечные функции распределения любого порядка инвариантны относительно сдвига параметра t на любое вещественное число, т. е. для любых n и t_0 справедливы равенства:

$$\begin{aligned} F_n(x_1, t_1; \dots; x_n, t_n) &= F_n(x_1, t_1 - t_0; \dots; x_n, t_n - t_0); \\ f_n(x_1, t_1; \dots; x_n, t_n) &= f_n(x_1, t_1 - t_0; \dots; x_n, t_n - t_0). \end{aligned} \quad (5.8)$$

Случайный процесс называется *стационарным в широком смысле*, если его математическое ожидание $m_\xi(t)$ и корреляционная (а соответственно и ковариационная) функция $R(t_1, t_2)$ не зависят от сдвига параметра t на любое вещественное число. Таким образом, для стационарного процесса $m_\xi(t) = \text{const}$, а $R(t_1, t_2) = R(t_2 - t_1) = R(\tau)$.

Случайный процесс называется *гауссовским (или нормальным)*, если все его конечномерные распределения являются многомерными нормальными.

Задача моделирования случайного процесса состоит в получении n его реализаций в различные дискретные моменты времени $t_1, t_2, \dots$ в пределах области определения Ω . Для построения алгоритмов моделирования случайных процессов существует ряд общих методов.

Метод линейного преобразования. Предположим, что необходимо получить реализации y_i стационарного случайного процесса $\xi(t)$ с заданной корреляционной функцией $R(\tau)$ и математическим ожиданием m_ξ . Реализации такого случайного процесса могут быть сгенерированы путем линейного преобразования реализаций x_i независимых случайных величин η_i , для которых математическое ожидание $m_\eta = 0$, а дисперсия $\sigma_\eta^2 = 1$.

Для начала рассмотрим случай, когда $m_\xi = 0$, а затем адаптируем полученный метод для генерации случайных процессов с любым другим m_ξ . Линейное преобразование для получения реализаций случайного

процесса $\xi(t)$ в матричной форме может быть представлено следующим выражением:

$$\|Y\| = \|A\| \|X\|, \quad (5.9)$$

где $\|Y\|$ – вектор реализаций искомого случайного процесса $\xi(t)$; $\|X\|$ – вектор реализаций случайных величин η_i ; $\|A\|$ – некоторая пока неизвестная матрица линейного преобразования. Выберем $\|A\|$ в виде треугольной матрицы, тогда уравнение (5.9) будет иметь вид

$$\begin{aligned} y_1 &= a_{11}x_1 \\ y_2 &= a_{21}x_1 + a_{22}x_2 \\ &\dots \\ y_n &= a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n \end{aligned} \quad (5.10)$$

Коэффициенты матрицы $\|A\|$ могут быть найдены при помощи известной корреляционной функции случайного процесса. На основании данной функции построим матрицу корреляционных коэффициентов R_{nm} размерности $n \times n$. Для этого используем разности между моментами времени, в которые необходимо сгенерировать реализации искомого случайного процесса:

$$\|R\| = \begin{bmatrix} R(t_1 - t_1) & R(t_2 - t_1) & \dots & R(t_n - t_1) \\ & R(t_2 - t_2) & \dots & R(t_n - t_2) \\ & & \dots & \\ & & & R(t_n - t_n) \end{bmatrix}. \quad (5.11)$$

По определению корреляционные коэффициенты задаются следующим образом:

$$R_{ij} = M\{y_i y_j\} = M\{y_j y_i\} = R_{ji}.$$

Принимая во внимание, что дисперсия случайной величины η равна 1 и, следовательно,

$$M\{x_i x_j\} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad (5.12)$$

с учетом (5.10) вычислим коэффициенты R_{ij} :

$$\begin{aligned}
R_{11} &= M\{(a_{11}x_1)^2\} = a_{11}^2 M\{x_1^2\} = a_{11}^2 \\
R_{12} &= M\{(a_{11}x_1)(a_{21}x_1 + a_{22}x_2)\} = a_{11}a_{21} \\
R_{22} &= M\{(a_{21}x_1 + a_{22}x_2)^2\} = a_{21}^2 + a_{22}^2 \\
R_{13} &= M\{(a_{11}x_1)(a_{31}x_1 + a_{32}x_2 + a_{33}x_3)\} = a_{11}a_{31} \\
R_{23} &= a_{21}a_{31} + a_{32}a_{22} \\
R_{33} &= a_{31}^2 + a_{32}^2 + a_{33}^2 \\
&\dots
\end{aligned}$$

Из полученных уравнений выражаем элементы матрицы $\|A\|$:

$$\begin{aligned}
a_{11} &= \sqrt{R_{11}}; \\
a_{i1} &= R_{i1}/a_{11}, i = 2, \dots, n; \\
a_{il} &= \frac{R_{li} - \sum_{j=1}^{l-1} a_{ij}a_{lj}}{a_{ll}}, l = 2, \dots, i-1,
\end{aligned} \tag{5.13}$$

где $a_{ll} = \sqrt{R_{ll} - \sum_{j=1}^{l-1} a_{lj}^2}$.

Таким образом, в соответствии с известным выражением для корреляционной функции можно вычислить элементы матрицы линейного преобразования, используя выражения (5.13), а затем получить реализации моделируемого случайного процесса при помощи выражения (5.9), предварительно сгенерировав реализации случайных величин η_i . Случайный процесс, сгенерированный таким образом, будет иметь нулевое математическое ожидание. Если необходимо сгенерировать реализации случайного процесса с некоторым m_ξ , отличным от 0, то это может быть выполнено путем прибавления к элементам вектора $\|Y\|$ соответствующей величины m_ξ .

Представленный алгоритм моделирования позволяет получать лишь необходимые корреляционные связи между реализациями случайного процесса в различные моменты времени. Законы же распределения в различные моменты времени нигде не принимаются во внимание, а следовательно, могут быть любыми. Единственным требованием является выполнение выражения (5.12) для случайных величин η .

Стоит отметить, что при увеличении n количество вычислений и оперативной памяти ЭВМ, требуемые для данного метода, увеличивают-

ся из-за роста размеров используемых матриц. Данный факт является недостатком приведенного выше метода.

Метод канонического разложения. Согласно данному методу моделируемый случайный процесс $\xi(t)$ представляется в виде канонического разложения вида

$$\xi(t) = \sum_{k=1}^{\infty} V_k \varphi_k(t), \quad (5.14)$$

где V_k – некоррелированные случайные величины с математическим ожиданием 0 и дисперсией σ_k^2 , $\varphi_k(t)$ – детерминированные функции канонического разложения.

Выражение (5.14) используется непосредственно при моделировании для получения реализаций искомого случайного процесса. Учитывая тот факт, что случайные величины V_k некоррелированы между собой и их математическое ожидание равно 0, корреляционная функция искомого случайного процесса будет выглядеть следующим образом:

$$R(t_1, t_2) = M\{\xi(t_1)\xi(t_2)\} = \sum_{k=1}^{\infty} \sigma_k^2 \varphi_k(t_1) \varphi_k(t_2).$$

В общем случае сумма в выражении (5.14) содержит бесконечное число слагаемых, однако на практике ее заменяют некоторой конечной суммой, где переменная суммирования k изменяется от 1 до N .

Основной задачей при использовании метода канонического разложения является определение дисперсий σ_k^2 случайных величин V_k и вида функций $\varphi_k(t)$, которые в общем случае могут быть заданы различными способами.

Часто в качестве координатных функций выбирают систему ортонормированных функций, которые удовлетворяют следующему условию:

$$\int_{-\infty}^{\infty} \varphi_n(t) \varphi_m(t) dt = \delta_{nm} = \begin{cases} 1, & n = m \\ 0, & n \neq m \end{cases}.$$

Теоретически доказано (теорема Карунена – Лоева), что разложение случайного процесса в ряд с некоррелированными коэффициентами по ортонормированной системе функций может быть произведено всегда. При этом дисперсии σ_k^2 находятся как собственные значения, а функции $\varphi_k(t)$ как собственные функции следующего интегрального уравнения:

$$\int_G R(t, s) \varphi(s) p(s) ds = \lambda \varphi(t), \quad (5.15)$$

где G – интервал разложения (в том числе и бесконечный), $R(t, s)$ – корреляционная функция случайного процесса $\xi(t)$, $p(s)$ – некоторые произвольно выбираемые неотрицательные весовые функции, которые могут использоваться для получения более удобного для вычислений результата.

Прямое использование уравнения (5.15) для практических целей в большинстве случаев является неудобным по причине вычислительных сложностей, которые возникают при его решении. По этой причине часто пользуются приближенными способами определения случайных коэффициентов и функций канонического разложения.

Предположим, что необходимо на конечном числе временных точек $t_1, \dots, t_n$ получить реализации случайного процесса $\xi(t)$, для которого задана корреляционная функция $R(t, t')$. Случайный процесс $\xi(t)$ может быть аппроксимирован другим случайным процессом $\xi^*(t)$, представимым в виде канонического разложения (5.14) и таким, что его корреляционная функция $R^*(t, t')$ совпадает в выделенных точках с $R(t, t')$. Показано, что для этого можно использовать каноническое разложение с конечным числом слагаемых, равным выбранному числу точек по времени:

$$\xi(t) = \sum_{k=1}^n V_k \varphi_k(t),$$

причем σ_k и $\varphi_k(t)$ определяются по следующим рекуррентным формулам:

$$\begin{aligned} \sigma_1^2 &= R(t_1, t_1); \varphi_1(t) = \frac{1}{\sigma_1} R(t, t_1); \\ \sigma_k^2 &= R(t_k, t_k) - \sum_{i=1}^{k-1} \sigma_i^2 \varphi_i(t_k); \varphi_k(t) = \\ &= \frac{1}{\sigma_k} \left[R(t, t_k) - \sum_{i=1}^{k-1} \sigma_i^2 \varphi_i(t) \varphi_i(t_k) \right], k = 2, \dots, N. \end{aligned} \quad (5.16)$$

В результате получаем следующий алгоритм для моделирования:

- на основании заданной корреляционной функции моделируемого случайного процесса $R(t, t')$ получаем дисперсии σ_k^2 и функции канонического разложения по формулам (5.16);
- получаем $x_1, \dots, x_n$ – реализации некоррелированных случайных величин с математическим ожиданием 0 и дисперсией 1;

- получаем реализации y_i случайного процесса $\xi(t)$ по формуле

$$y_i = \sum_{k=1}^n x_k \sigma_k \varphi_k(t_i).$$

Строго говоря, корреляционная функция смоделированного таким образом случайного процесса совпадает с корреляционной функцией исходного процесса только в выбранных точках. Но если точки подобраны так, что значения в соседних точках сильно коррелированы, то $\xi^*(t)$ может использоваться в этом случае для моделирования и в других промежуточных точках. Если положить закон распределения случайных коэффициентов V_k нормальным, то получим алгоритм для точного моделирования нормального случайного процесса (не в рамках корреляционных представлений, а в рамках многомерных распределений).

Преимущество приведенного выше приближенного алгоритма состоит в том, что он позволяет получить каноническое разложение с помощью обычных алгебраических операций, не прибегая к решению интегральных уравнений. Данный алгоритм особенно удобен при небольшом числе дискретных точек по времени.

Метод разложения в ряд Фурье. Наиболее простым частным случаем ортогонального разложения на конечном интервале $[0, T]$ для стационарных случайных процессов является разложение, при котором в качестве собственных функций выбирают синусы и косинусы. Тогда случайный процесс $\xi(t)$ может быть представлен в следующем виде:

$$\xi(t) = \sum_{k=0}^{\infty} V_k \cos(\omega_k t) + U_k \sin(\omega_k t), \quad 0 \leq t \leq T, \quad (5.17)$$

где V_k и U_k – некоррелированные случайные амплитуды гармоник; $\omega_k = \omega_1 k$ – частоты, кратные основной частоте. Таким образом, как следует из (5.17), при действительном t $\xi(t)$ является случайной периодической функцией с периодом $T_1 = 2\pi/\omega_1$. В общем случае T_1 не совпадает с T .

Построим алгоритм моделирования, основанный на представлении (5.17). Поскольку V_k и U_k некоррелированы, то корреляционная функция $R(t, t')$ для $\xi(t)$ может быть представлена в следующем виде:

$$R(t, t') = \sum_{k=0}^{\infty} D(V_k) \cos(\omega_k t) \cos(\omega_k t') + D(U_k) \sin(\omega_k t) \sin(\omega_k t'), \quad (5.18)$$

где $D(V_k)$ и $D(U_k)$ – дисперсии случайных коэффициентов V_k и U_k . Если принять $D(V_k) = D(U_k) = \sigma_k^2$, то формула (5.18) будет иметь вид

$$\begin{aligned} R(t, t') &= \sum_{k=0}^{\infty} \sigma_k^2 (\cos(\omega_k t) \cos(\omega_k t') + \sin(\omega_k t) \sin(\omega_k t')) = \\ &= \sum_{k=0}^{\infty} \sigma_k^2 \cos \omega_k (t - t') = \sum_{k=0}^{\infty} \sigma_k^2 \cos \omega_k \tau = R(\tau). \end{aligned} \quad (5.19)$$

Как видно из выражения (5.19), корреляционная функция в данном случае зависит лишь от разности аргументов. При этом $R(\tau)$ является периодической функцией с периодом T_1 равным периоду $\xi(t)$, а дисперсии σ_k^2 являются коэффициентами разложения $R(\tau)$ в ряд Фурье по косинусам.

Зная величину интервала разложения T , находим коэффициенты разложения корреляционной функции $R(\tau)$ заданного процесса в ряд Фурье по косинусам на удвоенном интервале $(-T, T)$:

$$\begin{aligned} \sigma_0^2 &= \frac{1}{T} \int_0^T R(\tau) d\tau; \\ \sigma_k^2 &= \frac{2}{T} \int_0^T R(\tau) \cos(k\omega_1 \tau) d\tau, \end{aligned} \quad (5.20)$$

где $\omega_1 = \frac{\pi}{T}$; $k = 1, 2, \dots$.

Величины σ_0^2, σ_k^2 в формулах (5.20) представляют собой дисперсии случайных коэффициентов V_k и U_k в искомом разложении. Если величина интервала T такова, что при $\tau \geq T$ значения корреляционной функции пренебрежимо малы, то верхний предел интегрирования в выражениях (5.20) можно заменить на бесконечность. Тогда выражения (5.20) будут иметь вид

$$\begin{aligned} \sigma_0^2 &= \frac{1}{T} \int_0^{\infty} R(\tau) d\tau = \frac{G(0)}{2T}; \\ \sigma_k^2 &= \frac{2}{T} \int_0^{\infty} R(\tau) \cos(k\omega_1 \tau) d\tau = \frac{G(k\omega_1)}{T}, k = 1, 2, \dots, \end{aligned} \quad (5.21)$$

где $G(\omega) = \int_{-\infty}^{\infty} R(\tau) e^{-j\omega\tau} d\tau = 2 \int_0^{\infty} R(\tau) \cos(\omega\tau) d\tau$ – энергетический спектр моделируемого случайного процесса. Дисперсии σ_k^2 с точностью до мно-

жителя совпадают со значениями спектральной плотности $G(\omega)$ в точках $\omega_k = k\omega_1 = k\pi/T$, $k = 0, 1, \dots$. Данный факт делает процесс вычисления дисперсий весьма простым при известной функции спектральной плотности.

Если выбрать некоторый постоянный шаг дискретизации по времени $\Delta t = T/N$, то моменты времени будут представлены в виде $t_n = n\Delta t$. Тогда реализации случайного процесса в такие моменты времени можно получить по формуле

$$\xi(t_n) = \sum_{k=0}^m V_k \cos \frac{k\pi n}{N} + U_k \sin \frac{k\pi n}{N}, \quad n = 1, \dots, N, \quad (5.22)$$

где учтено, что $\omega_k t = \frac{k\pi}{T} n \frac{T}{N} = \frac{k\pi n}{N}$. На практике число слагаемых в формуле (5.22) можно выбрать исходя из следующего условия:

$$1 - \frac{1}{R(0)} \sum_{k=0}^m \sigma_k^2 < \varepsilon,$$

где ε – некоторая достаточно малая величина. Это неравенство выражает тот факт, что сумма дисперсий σ_k^2 должна быть равна дисперсии моделируемого случайного процесса.

При моделировании нормального случайного процесса V_k и U_k будут являться нормальными случайными величинами. В этом случае алгоритм моделирования может быть представлен в виде

$$\xi(t_n) = \sum_{k=0}^m E_k \cos \left(\frac{k\pi n}{N} + \beta_k \right), \quad (5.23)$$

где E_k – случайные коэффициенты, подчиняющиеся распределению Релея, функция плотности распределения которого задается формулой

$$f_E(x) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}, \quad x \geq 0,$$

где параметр $\sigma = \sigma_k$, β_k – случайные фазы гармоник, независимые от E_k и равномерно распределенные в интервале $[0, 2\pi]$.

Алгоритмы (5.22) и (5.23) для получения реализаций моделируемого случайного процесса требуют вычисления значений тригонометрических функций. Алгоритмы вычисления тригонометрических функций обычно реализуются в виде стандартных подпрограмм и требуют значительных вычислительных затрат, что снижает эффективность моделирования. При постоянном шаге дискретизации объем вычислений можно значи-

тельно сократить, если использовать рекуррентные алгоритмы для вычисления тригонометрических функций от аргументов вида $\frac{k\pi n}{N}$, $n = 1, \dots, N$. Для построения таких рекуррентных алгоритмов можно использовать следующие равенства:

$$\begin{aligned}\sin\left(\frac{k\pi}{N}n\right) &= \sin\left(\frac{k\pi}{N}(n-1) + \frac{k\pi}{N}\right) = \\ &= \sin\left(\frac{k\pi}{N}(n-1)\right)\cos\left(\frac{k\pi}{N}\right) + \cos\left(\frac{k\pi}{N}(n-1)\right)\sin\left(\frac{k\pi}{N}\right); \\ \cos\left(\frac{k\pi}{N}n\right) &= \cos\left(\frac{k\pi}{N}(n-1) + \frac{k\pi}{N}\right) = \\ &= \cos\left(\frac{k\pi}{N}(n-1)\right)\cos\left(\frac{k\pi}{N}\right) - \sin\left(\frac{k\pi}{N}(n-1)\right)\sin\left(\frac{k\pi}{N}\right).\end{aligned}\tag{5.24}$$

Как видно из выражений (5.24), достаточно вычислить значения тригонометрических функций при $n = 1$, а их значения при других n могут быть получены рекуррентным образом. Использование рекуррентных формул (5.24) вместо прямого вычисления тригонометрических функций на каждом шаге вычислений сокращает число используемых элементарных операций на порядок и более, в зависимости от того, какое число элементарных операций задействовано при вычислении тригонометрических функций. Данный факт позволяет существенно сократить время моделирования.

Лекция 6

МЕТОД СКОЛЬЗЯЩЕГО СУММИРОВАНИЯ

Ранее был рассмотрен ряд методов, позволяющих выполнять моделирование случайных процессов, заданных на конечном интервале времени, в рамках корреляционной теории. Как было отмечено, основным недостатком этих методов является значительное увеличение вычислительных затрат на моделирование при формировании реализаций большой длины. Данный факт часто затрудняет использование описанных ранее методов моделирования на практике.

В общем случае при моделировании больших реализаций случайных процессов не существует единых эффективных методов моделирования. Однако на практике приходится чаще всего сталкиваться с необходимостью моделирования случайных процессов, обладающих определенными свойствами и относящихся к некоторому достаточно узкому

классу. Примерами таких процессов могут служить, прежде всего, стационарные нормальные случайные процессы и стационарные случайные процессы, являющиеся порождением нормальных в нелинейных системах. Для этих классов случайных процессов существуют достаточно эффективные алгоритмы моделирования.

Моделирование стационарного нормального случайного процесса $\xi(t)$ в некоторых точках области определения $t_n, n = 1, \dots, M$ основано на линейном преобразовании стационарной последовательности $x[n]$ независимых нормальных случайных величин (дискретного белого шума) в последовательность $\xi[n]$, которая коррелирована по заданному закону. Особенностью методов, которые будут рассмотрены далее, является то, что с их помощью можно получать реализации нормального случайного процесса сколь угодно большой длины.

Метод скользящего суммирования. Согласно данному методу реализации моделируемого нормального случайного процесса могут быть получены с использованием следующего линейного преобразования:

$$\xi[n] = \sum_{k=1}^N c_k x[n-k], \quad (6.1)$$

где $x[n]$ – последовательность независимых нормальных случайных величин с $m_x = 0$ и $\sigma_x^2 = 1$, c_k – пока неизвестные коэффициенты линейного разложения. Заметим, что для получения первой реализации случайного процесса требуется N реализаций $x[n]$. Для вычисления каждой последующей реализации $\xi[n]$ требуется получить одно дополнительное случайное число последовательности $x[n]$. При этом для вычисления очередного значения $\xi[n]$ используется N последних реализаций $x[n]$. Таким образом, при вычислении значений случайного процесса операция суммирования в выражении (6.1) смещается (скользит) вдоль последовательности $x[n]$ в сторону значений, полученных последними. Отсюда происходит и название рассматриваемого метода.

Зависимость (коррелированность) между получаемыми значениями случайного процесса $\xi[n]$ и $\xi[n+m]$ обеспечивается за счет того, что в их вычислении участвует $N-m$ общих случайных величин исходной последовательности $x[n]$. При $m \geq N$ значения случайного процесса $\xi[n]$ и $\xi[n+m]$ становятся некоррелированными. Из выражения (6.1) очевидно, что характер корреляционных связей между $\xi[n]$ и $\xi[n+m]$ может устанавливаться только за счет выбора коэффициентов разложения c_k .

Обозначим известную корреляционную функцию моделируемого по формуле (6.1) стационарного нормального случайного процесса $\xi(t)$ как $R(\tau)$. Тогда значения данной корреляционной функции могут быть использованы для нахождения неизвестных до сих пор весовых коэффициентов c_k , которые входят в выражение (6.1). Для решения данной задачи существует несколько подходов, которые будут рассмотрены далее.

Получение коэффициентов c_k путем решения нелинейной алгебраической системы уравнений.

С учетом определения корреляционной функции (выражение (5.5)) можно установить связь в точках t_n между значениями $R(\tau)$ и коэффициентами c_k в разложении (6.1), которая выглядит следующим образом:

$$\begin{aligned} R[0] &= c_1^2 + \dots + c_N^2, \\ R[1] &= c_1 c_2 + \dots + c_{N-1} c_N, \\ &\dots \\ R[N-1] &= c_1 c_N, \end{aligned} \quad (6.2)$$

где учтено, что для независимых случайных величин $x[n]$ справедливо равенство

$$M\{x[k]x[k+m]\} = \begin{cases} 1, & m = 0 \\ 0, & m \neq 0 \end{cases}.$$

Система нелинейных относительно весовых коэффициентов c_k уравнений (6.2) может быть представлена в матричной форме следующим образом:

$$\begin{pmatrix} R[0] \\ R[1] \\ \vdots \\ R[N-2] \\ R[N-1] \end{pmatrix} = \begin{pmatrix} c_1 & c_2 & \dots & c_{N-1} & c_N \\ c_2 & c_3 & \dots & c_N & 0 \\ \dots & \dots & \dots & \dots & \dots \\ c_{N-1} & c_N & \dots & 0 & 0 \\ c_N & 0 & \dots & 0 & 0 \end{pmatrix} \times \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_{N-1} \\ c_N \end{pmatrix}$$

Для решения нелинейной системы уравнений (6.2) можно в общем случае воспользоваться численными методами, однако при увеличении размерности системы вычисления будут становиться более трудоемкими. Кроме того, если точки $t_n = n\Delta t$ и система (6.2) была решена численно для некоторого шага дискретизации, то при изменении Δt требуется повторно решать нелинейную систему уравнений, что обычно связано со значительными вычислительными затратами, либо применять алгоритмы

интерполяции по ранее полученным значениям случайного процесса со старым шагом дискретизации, что сопряжено с потерей точности вычислений.

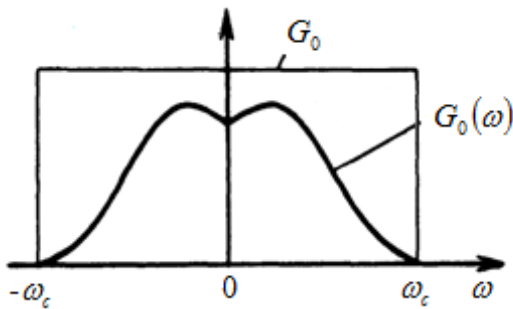
Получение коэффициентов c_k путем разложения корня квадратного из функции спектральной плотности моделируемого процесса в ряд Фурье.

Данный подход к получению коэффициентов c_k разложения (6.1) позволяет представить их в виде аналитических выражений. Получить данные аналитические выражения можно представив формирование случайного процесса $\xi(t)$ как результат прохождения белого шума через некоторую линейную систему.

Пусть необходимо смоделировать непрерывный стационарный нормальный случайный процесс $\xi(t)$ с нулевым математическим ожиданием и заданной корреляционной функцией $R(\tau)$. Энергетический спектр данного случайного процесса вычисляется как преобразование Фурье от его корреляционной функции

$$G(\omega) = \int_{-\infty}^{\infty} R(\tau) e^{-j\omega\tau} d\tau. \quad (6.3)$$

Обычно спектральная плотность $G(\omega)$ при достаточно больших ω убывает и после некоторой частоты ω_c становится пренебрежимо малой (рис. 6.1). Тогда при моделировании случайный процесс $\xi(t)$ можно с достаточной степенью точности заменить случайным процессом $\xi_0(t)$, который имеет ограниченную спектральную плотность вида



$$G_0(\omega) = \begin{cases} G(\omega), & |\omega| \leq \omega_c \\ 0, & |\omega| > \omega_c \end{cases}$$

Рис. 6.1. Спектральные плотности входного и выходного сигналов

Пусть случайный процесс $\xi_0(t)$ является результатом воздействия на некоторую непрерывную линейную систему непрерывного нормально-го белого шума $x(t)$ со спектром G_0 , ограниченным частотой ω_c . Тогда передаточная функция такой системы может быть получена из следующего соотношения:

$$G_0(\omega) = |K(j\omega)|^2 G_0. \quad (6.4)$$

Соотношение (6.4) выражает известный в теории случайных процессов факт: энергетический спектр шума на выходе линейной системы равен произведению энергетического спектра входного шума на квадрат модуля передаточной функции (комплексной частотной характеристики) системы. В общем случае выражению (6.4) удовлетворяет множество линейных систем, которые могут отличаться друг от друга своими фазочастотными характеристиками, которые являются аргументами комплексной функции $K(j\omega)$. Для дальнейшего рассмотрения будем использовать одну из таких систем с нулевой фазочастотной характеристикой. Тогда передаточная функция K_0 такой системы будет вещественной и, как следует из выражения (6.4), будет равна

$$K_0(j\omega) = \left(\frac{1}{G_0} G_0(\omega) \right)^{1/2}. \quad (6.5)$$

Пусть рассматриваемая нами линейная система обладает некоторой импульсной переходной характеристикой $h(t)$, которая определяется как реакция системы в некоторый момент времени на δ – импульс, поступивший на вход системы на τ единиц времени раньше. Тогда связь входного и выходного сигналов будет иметь вид

$$\xi_0(t) = \int_{-\infty}^{\infty} h(\tau) x(t - \tau) d\tau. \quad (6.6)$$

Импульсная переходная характеристика системы связана с ее передаточной функцией преобразованием Фурье:

$$h(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} K_0(j\omega) e^{j\omega t} d\omega. \quad (6.7)$$

Тогда с учетом формулы (6.5) выражение для переходной характеристики $h(t)$ будет иметь вид

$$h(t) = \frac{1}{\pi} \int_0^{\omega_c} \left(\frac{1}{G_0} G_0(\omega) \right)^{1/2} \cos(\omega t) d\omega, \quad (6.8)$$

где учтен тот факт, что $K_0(j\omega)$ является действительной четной функцией.

Покажем, что значения $\xi_0[n]$ моделируемого случайного процесса $\xi_0(t)$ на выходе рассматриваемой нами линейной системы в моменты времени $t_n = n\Delta t$, $n = 1, \dots, M$ можно выразить через дискретные значе-

ния $x[n]$ входного процесса $x(t)$ и дискретные значения $h[n]$ импульсной переходной характеристики системы $h(t)$.

Случайный процесс $x(t)$ на входе системы может быть представлен в виде ряда Котельникова:

$$x(t) = \sum_{n=-\infty}^{\infty} x[n] f_n(t), \quad (6.9)$$

где $f_n(t) = \frac{\sin[\omega_c(t - n\Delta t)]}{\omega_c(t - n\Delta t)}$, $\Delta t = \frac{\pi}{\omega_c}$ – шаг дискретизации, $x[n] = x(n\Delta t)$ –

последовательность независимых нормально распределенных случайных

величин с параметрами $(0, \sigma^2)$, $\sigma^2 = \frac{1}{\pi} \int_0^{\omega_c} G_0 d\omega = G_0 \frac{\omega_c}{\pi}$. Выражение (6.9)

показывает связь между непрерывным белым шумом $x(t)$ со спектром, ограниченным частотой ω_c , и дискретным белым шумом $x[n]$, значения которого совпадают с $x(t)$ в точках $n\Delta t$. Поскольку дисперсия $x[n]$ при использовании в выражении (6.1) равна 1, то

$$G_0 \frac{\omega_c}{\pi} = 1 \Rightarrow G_0 = \frac{\pi}{\omega_c} = \Delta t. \quad (6.10)$$

Последнее выражение определяет связь между спектральной плотностью входного белого шума, а также шагом дискретизации по времени и частотой ω_c , ограничивающей энергетический спектр моделируемого случайного процесса.

Значения моделируемого случайного процесса в выбранных нами временных точках с учетом соотношения (6.6) могут быть получены при помощи следующего выражения:

$$\xi_0[n] = \int_{-\infty}^{\infty} h(\tau) x(n\Delta t - \tau) d\tau. \quad (6.11)$$

Для нахождения $\xi_0[n]$ разложим функцию $h(t)$, обладающую в силу (6.8) ограниченным спектром, в ряд Котельникова:

$$h(t) = \sum_{n=-\infty}^{\infty} h[n] f_n(t), \quad (6.12)$$

где

$$h[n] = h(n\Delta t) = \frac{1}{\pi} \int_0^{\omega_c} K_0(j\omega) \cos\left(\frac{n\pi\omega}{\omega_c}\right) d\omega.$$

Подставив в (6.11) выражения (6.9) и (6.12), получим

$$\xi_0[n] = \sum_{k=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} h[k]x[m] \int_{-\infty}^{\infty} f_k(\tau)f_m(n\Delta t - \tau)d\tau. \quad (6.13)$$

Учитывая то, что $f_m(n\Delta t - \tau) = f_{n-m}(\tau)$, а также то, что функции $f_n(t)$ являются ортогональными, т. е.

$$\int_{-\infty}^{\infty} f_k(t)f_{n-m}(t)dt = \begin{cases} \frac{\pi}{\omega_c}, & k = n - m \\ 0, & k \neq n - m \end{cases},$$

окончательно получаем

$$\xi_0[n] = \sum_{k=-\infty}^{\infty} c_0[k]x[n-k], \quad (6.14)$$

где с учетом (6.8)

$$c_0[k] = \frac{1}{\omega_c} \int_0^{\omega_c} \left[\frac{\omega_c}{\pi} G(\omega) \right]^{1/2} \cos \frac{k\pi\omega}{\omega_c} d\omega. \quad (6.15)$$

Как следует из (6.15), коэффициенты $c_0[k]$ являются коэффициентами Фурье в разложении функции $K_0(j\omega)$ на интервале $(-\omega_c, \omega_c)$. Как правило, коэффициенты $c_0[k]$ достаточно быстро убывают. Если, например, функция $K_0(j\omega)$ непрерывна, то для $c_0[k]$ существует следующая оценка:

$$c_0[k] \leq \frac{Q}{k^2},$$

где Q – некоторое положительное число.

По этой причине на практике достаточно ограничиться небольшим числом членов в (6.14). На основании этого данное выражение можно переписать в следующем виде:

$$\xi_0[n] \approx \xi^*[n] = \sum_{k=-p}^p c_0[k]x[n-k]. \quad (6.16)$$

Поскольку исходная последовательность $x[n]$ стационарна, то статистические свойства последовательности $\xi^*[n]$ не изменятся, если формулу (6.16) переписать в следующем виде:

$$\xi^*[n] = \sum_{k=1}^N c_k x[n-k], \quad (6.17)$$

где $N = 2p + 1$, а $c_k = c_o[k - p - 1]$. Выражение (6.17) совпадает с общей формулой для метода скользящего суммирования (6.1), где коэффициенты c_k могут быть найдены с использованием заданной корреляционной функции моделируемого случайного процесса при помощи (6.3) и (6.15).

На практике количество слагаемых в разложении (6.1), а соответственно и коэффициентов c_k , можно найти исходя из следующего условия:

$$\left| 1 - \frac{1}{\sigma^2} \sum_{k=1}^N c_k^2 \right| < \varepsilon,$$

где σ^2 – дисперсия моделируемого случайного процесса, ε – некоторая малая константа. Данное неравенство основано на том, что сумма квадратов весовых коэффициентов должна быть равна дисперсии моделируемого случайного процесса.

Л е к ц и я 7

МЕТОД РЕКУРРЕНТНЫХ РАЗНОСТНЫХ УРАВНЕНИЙ

При использовании метода рекуррентных разностных уравнений реализации моделируемого случайного процесса $\xi[n]$ в M точках области его определения могут быть получены с использованием следующего линейного преобразования:

$$\xi[n] = \sum_{k=0}^l a_k x[n-k] - \sum_{k=1}^m b_k \xi[n-k], \quad (7.1)$$

где $x[n]$ – последовательность независимых нормальных случайных величин с $m_x = 0$ и $\sigma_x^2 = 1$, a_k и b_k – пока неизвестные коэффициенты. Как видно из выражения (7.1), для получения очередной реализации моделируемого случайного процесса $\xi[n]$ используются как реализации независимых случайных величин $x[n]$, так и реализации самого моделируемого случайного процесса, которые были получены в ряде предыдущих точек. По этой причине формула, используемая для выполнения моделирования, представляет собой рекуррентное выражение. Алгоритм, представленный соотношением (7.1), требует задания начальных условий, т. е. некоторых предыдущих значений $\xi[n-k]$, которые будут использоваться для вычисления первого значения этой последовательности. Эти значения можно выбрать нулевыми. В этом случае в начале сгенерированной последовательности $\xi[n]$ будет наблюдаться некоторый переходной процесс и на начальном участке полученная последовательность будет

искаженной. Однако после окончания переходного процесса $\xi[n]$ становится стационарной, и получаемые далее значения могут использоваться для построения реализаций $\xi(t)$.

Необходимо отметить, что рекуррентный алгоритм вида (7.1) применим только для моделирования ограниченного класса стационарных нормальных процессов, корреляционные функции которых обладают специальными свойствами. Преобразование Фурье от таких корреляционных функций, которое позволяет получить функцию спектральной плотности случайного процесса, представляет собой дробно-рациональную функцию вида

$$G(\omega) = \int_{-\infty}^{\infty} R(\tau) e^{-j\omega\tau} d\tau = \frac{G_1(\omega)}{G_2(\omega)}, \quad (7.2)$$

где $G_1(\omega)$ и $G_2(\omega)$ – полиномы степени l' и $m' > l'$ соответственно.

Корреляционные функции, соответствующие спектральной плотности (7.2), в общем виде задаются следующим выражением:

$$R(\tau) = \sum_k C_k e^{\lambda_k \tau}, \quad \lambda_k = \alpha_k + j\beta_k; \beta_k > 0, \quad (7.3)$$

где C_k – полиномы относительно τ . Для случая вещественных процессов корреляционная функция (7.3) может быть представлена в следующем виде:

$$R(\tau) = \sum_k [A_k(\tau) \cos \alpha_k \tau + B_k(\tau) \sin \alpha_k |\tau|] e^{-\beta_k |\tau|}, \quad (7.4)$$

где множители $A_k(\tau)$ и $B_k(\tau)$ представляют собой полиномы относительно τ .

Основной задачей при построении алгоритма моделирования с использованием метода рекуррентных разностных уравнений является нахождение коэффициентов a_k и b_k , выражения для которых будут определяться корреляционной функцией моделируемого случайного процесса. Получение данных параметров рекуррентного алгоритма возможно методами приведенными далее.

Метод факторизации спектральной плотности случайных процессов.

Пусть требуется смоделировать стационарный случайный процесс $\xi(t)$, корреляционная функция которого $R(\tau)$ задается выражением (7.4). Поскольку при моделировании задача состоит в нахождении реализаций $\xi[n] = \xi(n\Delta t)$ случайного процесса $\xi(t)$ на множестве дискретных

точек, то корреляционная функция моделируемого случайного процесса для данного случая может быть переписана в дискретной форме

$$R[n] = \sum_k [A_k[n] \cos \bar{\alpha}_k n + B_k[n] \sin \bar{\alpha}_k n] e^{-\bar{\beta}_k |n|}, \quad (7.5)$$

где $A_k[n]$ и $B_k[n]$ представляют собой дискретные многочлены, $\bar{\alpha}_k = \alpha_k \Delta t$ и $\bar{\beta}_k = \beta_k \Delta t$ – безразмерные параметры.

По аналогии с непрерывным случайным процессом $\xi(t)$ для связанного с ним дискретного случайного процесса $\xi[n]$ вводится понятие спектральной плотности, которая определяется следующим выражением:

$$\Phi(\bar{\omega}) = \sum_{n=-\infty}^{\infty} R[n] e^{-j\bar{\omega}n}, \quad (7.6)$$

где $\bar{\omega} = \omega \Delta t$ – безразмерная частота. Обозначив $z = e^{-j\bar{\omega}}$, можно переписать выражение (7.6) в следующем виде:

$$\Phi(\bar{\omega}) = F(z) = \sum_{n=-\infty}^{\infty} R[n] z^n. \quad (7.7)$$

Как и для случая непрерывных функций $R(\tau)$ и $G(\omega)$, спектральная плотность $F(z)$ случайного процесса с корреляционной функцией (7.5) является дробно-рациональной функцией относительно $z = e^{-j\bar{\omega}}$ и может быть представлена в виде

$$F(z) = \frac{A'(z)}{B'(z)}, \quad (7.8)$$

где $A'(z) = A'_0 + A'_1 z + \dots + A'_{l'} z^{l'}$, $B'(z) = B'_0 + B'_1 z + \dots + B'_{m'} z^{m'}$, $l' < m'$. Для вещественных процессов коэффициенты в данных полиномах являются вещественными числами.

Из теории случайных процессов известно, что энергетический спектр шума на выходе некоторой линейной системы равен произведению энергетического спектра входного шума на квадрат модуля передаточной функции системы. Если на вход системы подается дискретный нормированный белый шум $x[n]$, спектральная плотность которого $G_x = 1$, а передаточная функция рассматриваемой системы задана выражением

$$K(z) = \frac{A(z)}{B(z)}, \quad (7.9)$$

то спектральная плотность шума на выходе системы в соответствии с вышесказанным будет равна

$$F(z) = |K(z)|^2 = \left| \frac{A(z)}{B(z)} \right|^2 = \frac{A(z)}{B(z)} \frac{A(z^{-1})}{B(z^{-1})}. \quad (7.10)$$

Если в (7.10) $A(z)$ и $B(z)$ являются полиномами относительно z , то спектральная плотность в (7.10) является дробно-рациональной. В этом случае, сравнивая (7.8) и (7.10), приходим к выводу, что дискретный случайный процесс $\xi[n]$, который соответствует моделируемому непрерывному случайному процессу $\xi(t)$, может быть получен путем пропускания дискретного нормированного белого шума через дискретный линейный фильтр с рациональной передаточной функцией, удовлетворяющей следующему уравнению:

$$|K(z)|^2 = \left| \frac{A(z)}{B(z)} \right|^2 = \frac{A'(z)}{B'(z)}. \quad (7.11)$$

Выражение (7.11) позволяет определить вид передаточной функции $K(z)$ по известной спектральной плотности $F(z)$. Поскольку уравнение (7.1) фактически описывает поведение линейной системы с передаточной функцией, которая задается формулой (7.9), то коэффициенты полиномов $A(z)$ и $B(z)$, входящих в (7.9) и будут определять коэффициенты a_k и b_k в рекуррентном уравнении рассматриваемого метода.

Таким образом, можно выделить следующие шаги при построении алгоритма моделирования.

Шаг 1. Нахождение спектральной плотности $F(z)$ моделируемого случайного процесса $\xi[n]$ по заданной корреляционной функции $R(\tau)$ с использованием выражений (7.5) – (7.7). На данном этапе часто перед нахождением спектральной плотности $F(z)$ корреляционную функцию моделируемого случайного процесса удобнее представить в комплексном виде (7.3). Кроме того, для вычисления суммы ряда (7.7) удобно представить его в следующем виде:

$$F(z) = F^+(z) + F^+(z^{-1}) - R[0], \quad (7.12)$$

где

$$F^+(z) = \sum_{n=0}^{\infty} R[n] z^n. \quad (7.13)$$

В данном случае сначала согласно (7.13) вычисляется $F^+(z)$, а затем $F(z)$ на основании (7.12).

Шаг 2. Факторизация $F(z)$, т. е. разбиение ее на два сомножителя:

$$F(z) = \frac{A'(z)}{B'(z)} = \frac{A(z)}{B(z)} \frac{A(z^{-1})}{B(z^{-1})} = |K(z)|^2$$

с целью нахождения выражения для дробно-рациональной передаточной функции $K(z) = \frac{A(z)}{B(z)}$.

Выполнение факторизации для дробно-рациональной спектральной плотности $F(z)$ возможно на основании приводимой в литературе теоремы о разложении дробно-рациональных неотрицательных функций.

Теорема 7.1. Всякая неотрицательная рациональная относительно z функция:

$$F(z) = \frac{A'(z)}{B'(z)} = \frac{\sum_{k=0}^{l'} A'_k z^k}{\sum_{k=0}^{m'} B'_k z^k} = \frac{C' \prod_{k=1}^{l'} (z - v'_k)}{\prod_{k=1}^{m'} (z - w'_k)}$$

может быть представлена в виде

$$\begin{aligned} F(z) &= \left| \frac{\sum_{k=0}^l A_k z^k}{\sum_{k=0}^m B_k z^k} \right|^2 = C \left| \frac{\prod_{k=1}^l (z - v_k)}{\prod_{k=1}^m (z - w_k)} \right|^2 = \\ &= C \frac{\prod_{k=1}^l (z - v_k) \prod_{k=1}^l (z^{-1} - v_k^*)}{\prod_{k=1}^m (z - w_k) \prod_{k=1}^m (z^{-1} - w_k^*)} = \frac{A(z)}{B(z)} \frac{A(z^{-1})}{B(z^{-1})}, \end{aligned} \quad (7.14)$$

где C' и C – некоторые константы, v'_k и w'_k соответственно корни полиномов $A'(z)$ и $B'(z)$, v_k – те из корней v'_k , которые по модулю больше 1 и половина корней равных 1, w_k – те из корней w'_k , которые по модулю больше 1. Отметим, что в общем случае приведенные выше корни полиномов являются комплексными числами.

Тогда, как следует из теоремы, передаточная функция $K(z)$ будет иметь вид

$$K(z) = \sqrt{C} \frac{\prod_{k=1}^l (z - v_k)}{\prod_{k=1}^m (z - w_k)} = \frac{\sum_{k=0}^l A_k z^k}{\sum_{k=0}^m B_k z^k}. \quad (7.15)$$

Множитель C находится из условия:

$$C \frac{\left| \prod_{k=1}^l (z - v_k) \right|^2}{\left| \prod_{k=1}^m (z - w_k) \right|^2} = F(z).$$

Шаг 3. Преобразование передаточной функции $K(z)$ к виду

$$K(z) = \frac{a_0 + a_1 z + \dots + a_l z^l}{1 + b_1 z + \dots + b_m z^m}.$$

Данное преобразование состоит в делении числителя и знаменателя в (7.15) на B_0 .

Метод дискретизации непрерывных формирующих фильтров.

Пусть необходимо смоделировать случайный процесс $\xi(t)$, который является реакцией линейной системы на воздействие нормированного белого шума $x(t)$ с единичной спектральной плотностью. Выходной сигнал линейной системы задается следующим выражением:

$$\xi(t) = \int_{-\infty}^{\infty} x(\tau) h(t - \tau) d\tau, \quad (7.16)$$

где $h(t)$ – импульсная переходная характеристика рассматриваемой линейной системы. Для физически реализуемых систем $h(t) = 0$, $x(t) = 0$, $t < 0$. Тогда выражение (7.16) приобретает вид

$$\xi(t) = \int_{-\infty}^{\infty} x(\tau) h(t - \tau) d\tau. \quad (7.17)$$

Пусть известна импульсная переходная характеристика линейной системы, которая задана следующим образом:

$$h(t) = C e^{-\alpha t}, \quad t \geq 0. \quad (7.18)$$

Тогда с учетом (7.16) моделируемый случайный процесс может быть получен с использованием следующего выражения:

$$\xi(t) = \int_0^t x(\tau) h(t - \tau) d\tau = C \int_0^t x(\tau) e^{-\alpha(t-\tau)} d\tau. \quad (7.19)$$

С учетом выражения (7.19) значение $\xi(t)$ в точке $t + \Delta t$ будет равно

$$\xi(t + \Delta t) = C \int_0^{t+\Delta t} x(\tau) e^{-\alpha(t+\Delta t-\tau)} d\tau. \quad (7.20)$$

Разобьем интервал интегрирования в выражении (7.20) на два подынтервала и перепишем (7.20) в виде

$$\begin{aligned} \xi(t + \Delta t) &= e^{-\alpha\Delta t} C \int_0^t x(\tau) e^{-\alpha(t-\tau)} d\tau + C \int_t^{t+\Delta t} x(\tau) e^{-\alpha(t+\Delta t-\tau)} d\tau = \\ &= e^{-\alpha\Delta t} \xi(t) + \Delta\xi. \end{aligned} \quad (7.21)$$

$\xi(t)$ и $\Delta\xi$ – независимые между собой интегралы от белого шума по неперекрывающимся временным интервалам. Учитывая, что значения $x(t)$ в различные моменты времени независимы, получим выражение для дисперсии $\Delta\xi$:

$$\sigma_{\Delta}^2 = \int_t^{t+\Delta t} h^2(t + \Delta t - \tau) d\tau = \int_0^{\Delta t} h^2(t) dt = C^2 \int_0^{\Delta t} e^{-2\alpha t} dt = \frac{C^2}{2\alpha} (1 - e^{-2\alpha\Delta t}). \quad (7.22)$$

С учетом выражений (7.21) и (7.22) получаем рекуррентный алгоритм для моделирования случайного процесса $\xi(t)$:

$$\xi[n] = e^{-\alpha\Delta t} \xi[n-1] + \alpha x[n] = \rho \xi[n-1] + \sqrt{(C^2 / 2\alpha)(1 - \rho^2)} x[n],$$

где $\rho = e^{-\alpha\Delta t}$, $x[n]$ – последовательность независимых нормальных случайных величин с параметрами (0,1).

Рассмотренные методы *скользящего суммирования* и *рекуррентных разностных уравнений* предназначены для моделирования стационарных нормальных случайных процессов. Оба метода основаны на выполнении линейного преобразования последовательности независимых нормальных случайных величин в последовательность реализаций случайного процесса с заданной корреляционной функцией. Для каждого из рассмотренных методов моделирования существуют различные подходы к получению коэффициентов соответствующего линейного преобразования.

Метод скользящего суммирования является универсальным методом, пригодным для моделирования стационарных нормальных случайных процессов независимо от конкретного вида корреляционной функции или спектральной плотности моделируемого процесса. Если *моделируемый процесс задан своей корреляционной функцией, а его спектральная плотность неизвестна* и ее вычисление путем взятия преобразования Фурье от корреляционной функции затруднительно, то получение весовых коэффициентов в методе скользящего суммирования может

быть выполнено путем аналитического или численного решения *нелинейной алгебраической системы уравнений*. Если же *спектральная плотность* рассматриваемого процесса задана или может быть легко получена путем преобразования Фурье от известной корреляционной функции, то весовые коэффициенты метода скользящего суммирования могут быть получены путем *разложения корня квадратного из функции спектральной плотности моделируемого процесса в ряд Фурье*. Стоит также отметить, что метод скользящего суммирования является, вообще говоря, приближенным методом, поскольку при моделировании используется конечное число слагаемых. Точность может быть увеличена за счет увеличения числа слагаемых, однако при этом будут увеличиваться требуемые вычислительные затраты.

Метод рекуррентных разностных уравнений предпочтительно использовать в случае, если спектральная плотность моделируемого стационарного нормального случайного процесса является дробно-рациональной функцией. В этом случае данный метод будет более экономичен с точки зрения вычислительных затрат.

Л е к ц и я 8

МОДЕЛИРОВАНИЕ ТИПОВЫХ НОРМАЛЬНЫХ И НЕГАУССОВСКИХ СЛУЧАЙНЫХ ПРОЦЕССОВ

В изложенном ранее материале было приведено описание методов моделирования стационарных нормальных случайных процессов, а также были сформулированы критерии, определяющие как выбрать тот или иной метод для построения алгоритма моделирования при решении конкретной задачи. Далее будут приведены примеры построения алгоритмов моделирования для некоторых типовых нормальных случайных процессов с применением различных подходов, изложенных ранее, а также рассмотрены некоторые аспекты моделирования негауссовских случайных процессов.

Моделирование стационарного нормального случайного процесса с корреляционной функцией:

$$R(\tau) = \frac{\sigma^2}{1 + \omega_*^2 \tau^2}. \quad (8.1)$$

Как следует из сравнения формулы (8.1), для корреляционной функции рассматриваемого случайного процесса с обобщенным выражением для корреляционной функции случайного процесса с дробно-рациональным спектром (формула (7.4)), данный случайный процесс не

относится к случайным процессам с дробно-рациональным спектром. По этой причине для моделирования данного случайного процесса будем использовать метод скользящего суммирования. Найдем выражение для спектральной плотности рассматриваемого случайного процесса, используя преобразование Фурье от его корреляционной функции (формула (6.3)):

$$G(\omega) = \int_{-\infty}^{\infty} \frac{\sigma^2}{1 + \omega_*^2 \tau^2} e^{-j\omega\tau} d\tau = \left\{ \frac{1}{1 + \tau^2} \rightarrow \pi e^{-|\omega|}; R(a\tau) \rightarrow \frac{1}{a} G\left(\frac{\omega}{a}\right) \right\} =$$

$$= \frac{\pi \sigma^2}{\omega_*} e^{-\left|\frac{\omega}{\omega_*}\right|}. \quad (8.2)$$

В подтверждение выше сказанного из формулы (8.2) видно, что полученная функция спектральной плотности не является дробно-рациональной функцией (отношением полиномов), а следовательно, использование рекуррентных методов невозможно.

Графические зависимости корреляционной функции и спектральной плотности рассматриваемого случайного процесса приведены на рис. 8.1.

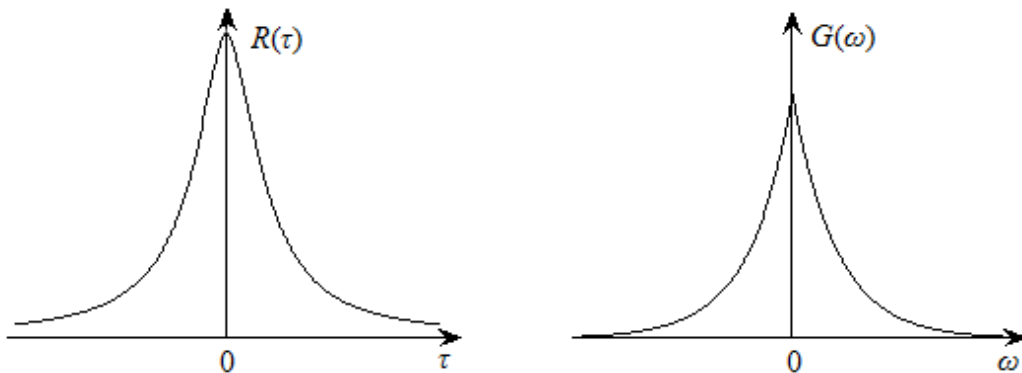


Рис. 8.1. Графическое представление корреляционной функции случайного процесса (формула (8.1)) и его спектральной плотности (формула (8.2))

Для построения алгоритма моделирования найдем коэффициенты c_k в формуле (6.1) метода скользящего суммирования. Для этого воспользуемся подходом, основанным на разложении квадратного корня спектральной плотности в ряд Фурье. В соответствии с формулой (6.15) получим

$$\begin{aligned}
c_0[k] &= \frac{1}{\omega_c} \int_0^{\omega_c} \left[\frac{\omega_c}{\pi} \frac{\pi \sigma^2}{\omega_*} e^{-\left| \frac{\omega}{\omega_*} \right|} \right]^{\frac{1}{2}} \cos \frac{k \pi \omega}{\omega_c} d\omega = \\
&= \frac{\sigma}{\omega_c} \int_0^{\omega_c} \left[\frac{\omega_c}{\omega_*} e^{-\left| \frac{\omega}{\omega_*} \right|} \right]^{\frac{1}{2}} \cos \frac{k \pi \omega}{\omega_c} d\omega.
\end{aligned} \tag{8.3}$$

Введем обозначение $\gamma = \frac{\omega_c}{\omega_*}$ и сделаем замену переменных $x = \frac{\omega}{\omega_c}$.

Тогда выражение (8.3) примет следующий вид:

$$c_0[k] = \sigma \sqrt{\gamma} \int_0^1 e^{-\gamma x / 2} \cos(k \pi x) dx.$$

Интеграл в последнем выражении является табличным:

$$\int e^{ax} \cos bx = \frac{e^{ax}}{a^2 + b^2} (a \cos bx + b \sin bx).$$

Учитывая, что в нашем случае $a = -\gamma/2$, $b = k\pi$ и, подставляя пределы интегрирования, получим

$$\begin{aligned}
c_0[k] &= \frac{\sigma \sqrt{\gamma}}{\gamma^2/4 + k^2 \pi^2} e^{-\frac{\gamma}{2} x} \left(k \pi \sin(k \pi x) - \frac{\gamma}{2} \cos(k \pi x) \right) \Big|_0^1 = \\
&= \frac{4 \sigma \sqrt{\gamma}}{\gamma^2 + 4 k^2 \pi^2} \left(-(-1)^{|k|} \frac{\gamma}{2} e^{-\frac{\gamma}{2}} + \frac{\gamma}{2} \right).
\end{aligned}$$

После элементарных преобразований последнего выражения окончательно получаем

$$c_0[k] = 2 \sigma \gamma^{3/2} \frac{1 - (-1)^{|k|} e^{-\gamma/2}}{\gamma^2 + 4 \pi^2 k^2}. \tag{8.4}$$

При достаточно больших ω_c $\gamma \gg 1$ и последнее выражение с учетом того, что экспонента будет достаточно быстро спадать, может быть представлено в упрощенном виде

$$c_0[k] = 2 \sigma \gamma^{3/2} \frac{1}{\gamma^2 + 4 \pi^2 k^2}. \tag{8.5}$$

Вычислив искомые коэффициенты по формуле (8.4) или (8.5), можно сформировать методом скользящего суммирования (по формуле (6.1)) последовательность $\xi[n]$ с требуемой корреляционной функцией (8.1).

Моделирование стационарного нормального случайного процесса с корреляционной функцией

$$R(\tau) = \sigma^2 e^{-\omega_* |\tau|} \cos(\omega_0 \tau). \quad (8.6)$$

Как следует из формы выражения (8.6), рассматриваемый случайный процесс является случайным процессом с дробно-рациональным спектром (формула (7.4)). Найдем функцию спектральной плотности данного случайного процесса:

$$\begin{aligned} G(\omega) &= \int_{-\infty}^{\infty} \sigma^2 e^{-\omega_* |\tau|} \cos(\omega_0 \tau) e^{-j\omega \tau} d\tau = \left\{ e^{-|\tau|} \rightarrow \frac{2}{1 + \omega^2}; R(a\tau) \cos(b\tau) \rightarrow \right. \\ &\rightarrow \left. \frac{1}{2a} \left(G\left(\frac{\omega - b}{a}\right) + G\left(\frac{\omega + b}{a}\right) \right) \right\} = \frac{\sigma^2}{2\omega_*} \left[\frac{2}{1 + \left(\frac{\omega - \omega_0}{\omega_*}\right)^2} + \frac{2}{1 + \left(\frac{\omega + \omega_0}{\omega_*}\right)^2} \right]. \end{aligned}$$

Выполняя приведение к общему знаменателю и необходимые сокращения, а также вводя обозначения $x = \frac{\omega}{\omega_*}$ и $x_0 = \frac{\omega_0}{\omega_*}$, получаем окончательное выражение для функции спектральной плотности

$$G(\omega) = \frac{2\sigma^2}{\omega_*} \frac{1 + x_0^2 + x^2}{(1 + (x + x_0)^2)(1 + (x - x_0)^2)}. \quad (8.7)$$

Корреляционная функция (8.6) и спектральная плотность (8.7) рассматриваемого случайного процесса представлены графически на рис. 8.2.

Поскольку рассматриваемый стационарный нормальный случайный процесс является случайным процессом с дробно-рациональной спектральной плотностью, то для построения алгоритма его моделирования выберем метод рекуррентных разностных уравнений. Для получения неизвестных коэффициентов метода (формула (7.1)) применим подход, основанный на факторизации спектральной плотности случайного процесса.

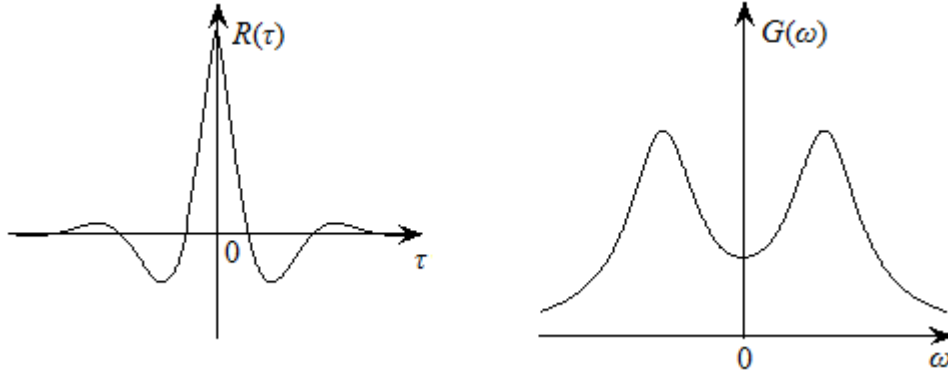


Рис. 8.2. Графическое представление корреляционной функции случайного процесса (формула (8.6)) и его спектральной плотности (формула (8.7))

С учетом (8.6) корреляционная функция дискретного процесса $\xi[n]$, соответствующего моделируемому случайному процессу, будет иметь вид

$$R[n] = \sigma^2 e^{-\gamma_* |n|} \cos(\gamma_0 n),$$

где $\gamma_* = \omega_* \Delta t$, $\gamma_0 = \omega_0 \Delta t$, Δt – шаг дискретизации.

Запишем функцию $R[n]$ в комплексном представлении (7.3):

$$R[n] = \frac{\sigma^2}{2} e^{\lambda_1 n} + \frac{\sigma^2}{2} e^{\lambda_2 n},$$

где $\lambda_{1,2} = -\gamma_* \pm j\gamma_0$.

Найдем спектральную плотность случайного процесса $\xi[n]$. Подставляя выражение для $R[n]$ в (7.13), получим

$$\begin{aligned} F^+(z) &= \sigma^2 \sum_{n=0}^{\infty} \left(\frac{1}{2} e^{\lambda_1 n} + \frac{1}{2} e^{\lambda_2 n} \right) z^n = \frac{\sigma^2}{2} \left(\sum_{n=0}^{\infty} (e^{\lambda_1} z)^n + \sum_{n=0}^{\infty} (e^{\lambda_2} z)^n \right) = \\ &= \left\{ \sum_{n=0}^{\infty} x^n = \frac{1}{1-x} \right\} = \frac{\sigma^2}{2} \left(\frac{1}{1 - e^{\lambda_1} z} + \frac{1}{1 - e^{\lambda_2} z} \right) = \sigma^2 \frac{1 - \rho \cos(\gamma_0) z}{1 - 2\rho \cos(\gamma_0) z + \rho^2 z^2}, \end{aligned} \quad (8.8)$$

где $\rho = e^{-\gamma_*}$ и учтено, что бесконечные суммы являются суммами бесконечных геометрических прогрессий. Подставляя полученное выражение в (7.12) и учитывая, что $R[0] = \sigma^2$, получим формулу для спектральной плотности $F(z)$:

$$F(z) = \sigma^2 \left[\frac{1 - \rho \cos(\gamma_0)z}{1 - 2\rho \cos(\gamma_0)z + \rho^2 z^2} + \frac{1 - \rho \cos(\gamma_0)z^{-1}}{1 - 2\rho \cos(\gamma_0)z^{-1} + \rho^2 z^{-2}} - 1 \right] =$$

$$= \sigma^2 \frac{-z^{-1}(A_0 + A_1 z + A_0 z^2)}{|1 + B_1 z + B_2 z^2|^2}$$

где выполнено приведение к общему знаменателю и $A_0 = (1 - \rho^2)\rho \cos \gamma_0$, $A_1 = -(1 - \rho^4)$, $B_1 = -2\rho \cos \gamma_0$, $B_2 = \rho^2$.

Следующим этапом является выполнение факторизации полученной спектральной плотности $F(z)$. Как видно из последнего полученного выражения, знаменатель уже является произведением двух комплексно-сопряженных многочленов и в факторизации не нуждается. Для представления в аналогичной форме числителя найдем корни соответствующего многочлена

$$v'_{1,2} = \frac{-A_1 \pm \sqrt{A_1^2 - 4A_0^2}}{2A_0} = -\frac{A_1}{2A_0} \pm \sqrt{\frac{A_1^2}{4A_0^2} - 1} = v_0 \pm \sqrt{v_0^2 - 1},$$

$$\text{где } v_0 = \frac{1 + \rho^2}{2\rho \cos \gamma_0} = \frac{ch \gamma_*}{ch \gamma_0}.$$

В рассматриваемом случае оба полученных корня $v'_{1,2}$ обращают в тождество следующее уравнение:

$$-\sigma^2 z^{-1}(A_0 + A_1 z + A_0 z^2) = C(z - v_1)(z^{-1} - v_1),$$

где $v_1 = v'_{1,2}$, а $C = \sigma^2 \frac{A_0}{v_1}$. По этой причине анализ модуля этих корней

можно не выполнять, а в качестве v_1 может быть выбран любой из них. Таким образом, в результате проведенной факторизации получаем следующее представление спектральной плотности $F(z)$:

$$F(z) = \frac{C(z - v_1)(z^{-1} - v_1)}{(1 + B_1 z + B_2 z^2)(1 + B_1 z^{-1} + B_2 z^{-2})}.$$

В соответствии с соотношением (7.15) получаем выражение для передаточной функции линейной системы

$$K(z) = \frac{\sqrt{C}(z - v_1)}{1 + B_1 z + B_2 z^2} = \frac{a_0 + a_1 z}{1 + b_1 z + b_2 z^2},$$

где $a_0 = -\sqrt{C}v_1 = -\sigma\sqrt{A_0 v_1}$, $a_1 = \sqrt{C} = \sigma\sqrt{A_0/v_1}$, $b_1 = B_1$, $b_2 = B_2$.

Соответственно рекуррентный алгоритм для моделирования рассматриваемого случайного процесса будет выглядеть следующим образом:

$$\xi[n] = a_0 x[n] + a_1 x[n-1] - b_1 \xi[n-1] - b_2 \xi[n-2].$$

Моделирование стационарного нормального случайного процесса с корреляционной функцией:

$$R(\tau) = \sigma^2 e^{-\omega_* |\tau|}. \quad (8.9)$$

Как видно из сопоставления выражений (8.9) и (8.6), данный случайный процесс является частным случаем случайного процесса с корреляционной функцией (8.6), где $\omega_0 = 0$. Тогда с учетом (8.7) функция спектральной плотности рассматриваемого процесса будет равна:

$$G(\omega) = \frac{2\sigma^2}{\omega_*} \left(\frac{1}{1+x^2} \right); \quad x = \frac{\omega}{\omega_*}. \quad (8.10)$$

Графики корреляционной функции и спектральной плотности приведены на рис. 8.3.

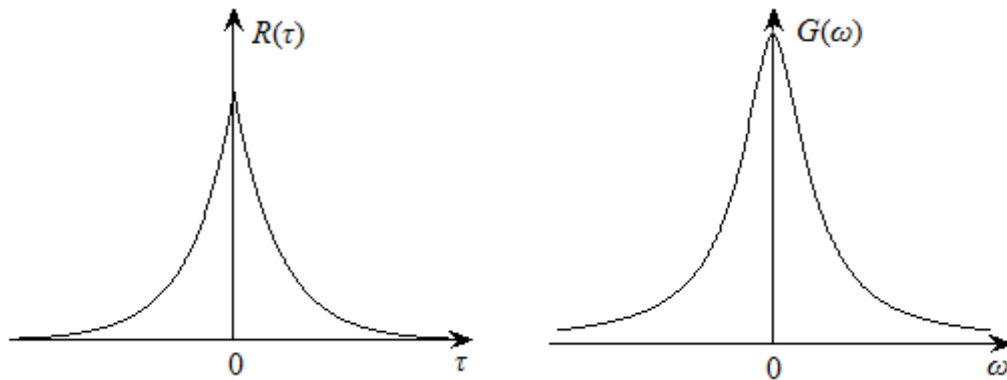


Рис. 8.3. Графическое представление корреляционной функции случайного процесса (формула (8.9)) и его спектральной плотности (формула (8.10))

Для построения алгоритма моделирования воспользуемся результатами, полученными в предыдущем примере. На основании выражения (8.8) с учетом того, что для рассматриваемого случайного процесса $\gamma_0 = 0$, получим

$$F^+(z) = \sigma^2 \frac{1 - \rho z}{1 - 2\rho z + \rho^2 z^2} = \frac{1}{1 - \rho z}, \quad \rho = e^{-\gamma_*}, \quad \gamma_* = \omega_* \Delta t.$$

Из этого выражения с учетом (7.12) получаем формулу для функции спектральной плотности $F(z)$:

$$F(z) = \sigma^2 \left[\frac{1}{1 - \rho z} + \frac{1}{1 - \rho z^{-1}} - 1 \right] = \sigma^2 \frac{1 - \rho^2}{(1 - \rho z)(1 - \rho z^{-1})}. \quad (8.11)$$

Как видно из последнего выражения, факторизацию числителя и знаменателя в данном случае проводить не требуется. Передаточная функция соответствующей линейной системы задается выражением:

$$K(z) = \sigma \frac{\sqrt{1 - \rho^2}}{1 - \rho z}.$$

Тогда рекуррентный алгоритм моделирования принимает вид

$$\xi[n] = a_0 x[n] - b_1 \xi[n-1] = \sigma \sqrt{1 - \rho^2} x[n] + \rho \xi[n-1]. \quad (8.12)$$

Моделирование негауссовских случайных процессов. Часто информация о негауссовском случайном процессе $\xi(t)$, который необходимо смоделировать, задается совокупностью его одномерной функции плотности вероятности $f_\xi(y, t)$ и корреляционной функции $R(t_1, t_2)$. В таком виде задача моделирования является неопределенной, поскольку не задано совместное многомерное распределение значений случайного процесса в различные моменты времени. В общем случае существует целый класс процессов с различными многомерными распределениями и с одинаковыми $f_\xi(y, t)$ и $R_\xi(t_1, t_2)$. По этой причине из этого класса можно выбрать какой-либо один из случайных процессов, построение которого будет наиболее приемлемым с точки зрения разрабатываемого алгоритма моделирования.

Простейшее решение в данной ситуации состоит в приведении моделируемого негауссова случайного процесса к совокупности гауссовых процессов посредством некоторого преобразования вида

$$\xi(t) = g[v_1(t), v_2(t), \dots, v_k(t), t], \quad (8.13)$$

где $v_i(t)$ – независимые гауссовы случайные процессы с одномерными плотностями распределения:

$$f_v(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

В стационарном случае случайные процессы $v_i(t)$ могут быть смоделированы одним из методов, представленных ранее. Преобразование (8.13) строится исходя из следующей формулы:

$$\xi(t) = f(\alpha_1, \dots, \alpha_k, t), \quad (8.14)$$

выражающей моделируемый процесс $\xi(t)$ через независимые базовые случайные величины α_i , $i=1,2,\dots,k$. Полагая $\alpha_i = F_{v_i}[v_i(t)]$ (поскольку функция распределения переводит значение случайной величины в значение базовой случайной величины), можно с учетом (8.14) представить формулу (8.13) в виде

$$\xi(t) = f[F_{v_1}\{v_1(t)\}, \dots, F_{v_k}\{v_k(t)\}, t]. \quad (8.15)$$

Корреляционная функция моделируемого случайного процесса $R_\xi(t_1, t_2)$ также может быть представлена в виде некоторой функции от корреляционных функций процессов $v_i(t)$:

$$R_\xi(t_1, t_2) = G\{R_{v_1}(t_1, t_2), \dots, R_{v_k}(t_1, t_2)\}. \quad (8.16)$$

Если удастся подобрать корреляционные функции гауссовых процессов $R_{v_i}(t_1, t_2)$, которые удовлетворяют равенству (8.16) и являются удобными для моделирования $v_i(t)$, то задача моделирования $\xi(t)$ оказывается решенной. Далее рассмотрим пример моделирования негауссова случайного процесса:

Пример 8.1. Моделирование стационарного процесса с экспоненциальным распределением.

Задача состоит в моделировании случайного процесса $\xi(t)$ с одномерной плотностью вероятностей

$$f_\xi(y) = \lambda e^{-\lambda y}, \quad y \geq 0$$

и корреляционной функцией $R_\xi(\tau)$. Для данного случайного процесса формула (8.13) может быть представлена в виде

$$\xi(t) = v_1^2(t) + v_2^2(t), \quad (8.17)$$

где $v_1(t)$ и $v_2(t)$ являются независимыми гауссовыми случайными процессами с нулевыми средними и дисперсиями $\sigma_{v_1}^2 = \sigma_{v_2}^2 = 1/(2\lambda)$. Как видно из (8.17), $M\{\xi(t)\} = M\{v_1^2(t)\} + M\{v_2^2(t)\} = 1/\lambda$. Кроме того, можно показать, что корреляционная функция $\xi(t)$ будет выражаться через корреляционные функции случайных процессов $v_1(t)$ и $v_2(t)$ следующим образом:

$$R_\xi(\tau) = \frac{1}{\lambda^2} + 2(R_{v_1}^2(\tau) + R_{v_2}^2(\tau)). \quad (8.18)$$

При подборе $R_{v_1}(\tau)$ и $R_{v_2}(\tau)$, удовлетворяющих последнему выражению, возможны различные варианты. Например, можно положить:

$$R_{v_1}(\tau) = R_{v_2}(\tau) = \frac{1}{2} \sqrt{R_\xi(\tau) - \frac{1}{\lambda^2}}. \quad (8.19)$$

Тогда моделирование $\xi(t)$ сводится к получению реализаций двух независимых одинаково распределенных нормальных процессов с корреляционными функциями (8.19) и к последующему получению реализации $\xi(t)$ в соответствии с (8.17). Стоит отметить, что полученные в (8.19) корреляционные функции для нормальных случайных процессов должны в точке 0 быть равными $1/(2\lambda)$, поскольку в данной точке при нулевых средних и при условии стационарности рассматриваемого случайного процесса корреляционная функция равна дисперсии. В противном случае выражение (8.17) не будет давать возможность получать искомый случайный процесс $\xi(t)$ со средним $1/\lambda$.

Если, например, корреляционная функция стационарного случайного процесса с экспоненциальным распределением задается следующим образом:

$$R_{\xi}(\tau) = \frac{1}{\lambda^2} (1 + e^{-a|\tau|}),$$

то с учетом выражения (8.19) корреляционные функции стационарных нормальных процессов будут иметь вид

$$R_{v_1}(\tau) = R_{v_2}(\tau) = \frac{1}{2\lambda} e^{-a|\tau|/2}.$$

Тогда с учетом алгоритма моделирования стационарных случайных процессов с нормальным распределением (8.12) и зависимости (8.17) получаем следующий рекуррентный алгоритм для моделирования стационарного случайного процесса с экспоненциальным распределением:

$$\xi[n] = \left(\sigma \sqrt{1 - \rho^2} x_1[n] + \rho \xi_1[n-1] \right)^2 + \left(\sigma \sqrt{1 - \rho^2} x_2[n] + \rho \xi_2[n-1] \right)^2,$$

где $\sigma = 1/\sqrt{2\lambda}$, $\rho = e^{-\gamma_*}$, $\gamma_* = \omega_* \Delta t = a \Delta t / 2$.

4. МОДЕЛИРОВАНИЕ СЛУЧАЙНЫХ ПОТОКОВ

Л е к ц и я 9

ОПИСАНИЕ ПОТОКОВ СЛУЧАЙНЫХ СОБЫТИЙ. ОБЩИЕ ПОДХОДЫ К МОДЕЛИРОВАНИЮ СЛУЧАЙНЫХ ПОТОКОВ

Случайный поток это последовательность случайных событий, наступающих в отдельные моменты времени на интервале моделирования $\Omega[T_0; T]$ и характеризующихся только моментами своего появления.

Графически поток случайных событий представлен на рис. 9.1 в виде следующих одна за другой точек на временной оси. Моменты наступления событий потока обозначены как ξ_i $i = 1, 2, \dots$

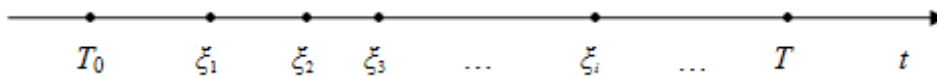


Рис. 9.1. Поток случайных событий

Далее будем предполагать, что моменты наступления событий рассматриваемого потока упорядочены во времени по возрастанию, т. е. $T_0 \leq \xi_1 \leq \xi_2 \leq \dots \leq \xi_i \leq \dots \leq T$.

Можно сопоставить понятия случайного потока и случайного вектора. В обоих случаях говорят о наборе реализаций в общем случае зависимых случайных компонент, однако основное отличие состоит в том, что число компонент случайного вектора фиксировано, а число компонент случайного потока является случайным, и, более того, может быть бесконечным в случае, если область определения потока имеет бесконечную протяженность. Иначе говоря, случайный поток может рассматриваться как случайный вектор со случайным числом компонент.

Процесс моделирования случайного потока состоит в последовательной генерации упорядоченных по возрастанию моментов наступления его событий.

Для задания случайных потоков часто используют *параметрический способ*. Поток задан параметрически, если известны совместные плотности распределения $\pi_i(t_1, \dots, t_i; \Omega)$, $i = 1, 2, \dots$ и вероятности $p_k(\Omega)$, $k = 0, 1, 2, \dots$, где $\pi_i(t_1, \dots, t_i; \Omega) dt_1 \dots dt_i$ – вероятность совместного попадания событий потока в интервалы $(t_j - dt_j, t_j]$, $j = 1, 2, \dots, i$, при условии, что на Ω выпало i событий, а $p_k(\Omega)$ – вероятность, что на Ω

происходит ровно k событий потока, $\sum_{k=0}^{\infty} p_k(\Omega) = 1$. Для получения реализации потока случайных событий, заданного вышеуказанным способом, необходимо сначала смоделировать случайную величину η , с дискретным распределением $P_k, P\{\eta=0\} = p_0(\Omega), P\{\eta=1\} = p_1(\Omega), \dots$, получая таким образом число событий на интервале Ω . Пусть в процессе моделирования дискретной случайной величины η была получена ее реализация i . Тогда, если $i=0$, то это означает, что на рассматриваемой области моделирования случайного потока не произошло ни одного события. Если $i \neq 0$, то в случайные моменты $t_1, \dots, t_i$ наступают события, совместная плотность вероятности которых равна $\pi_i(t_1, \dots, t_i; \Omega)$. Таким образом, оставшаяся часть моделирования случайного потока сводится к задаче моделирования случайного вектора с известной совместной плотностью распределения, для решения которой можно воспользоваться методом обратных функций и получить систему уравнений относительно искомым $t_1, \dots, t_i$:

$$F_j^i(t_j / t_1, \dots, t_{j-1}) = \alpha_j, j = 1, 2, \dots, i, \quad (9.1)$$

где α_j – независимые реализации базовой случайной величины и условная функция распределения:

$$F_j^i(t_j / t_1, \dots, t_{j-1}) = \int_{t_{j-1}}^{t_j} \pi_j^i(t / t_1, \dots, t_{j-1}) dt. \quad (9.2)$$

В свою очередь, условная плотность распределения, в подынтегральном выражении определяется следующим образом:

$$\pi_j^i(t_j / t_1, \dots, t_{j-1}) = \frac{\int_{t_j \leq t_{j+1} \leq \dots \leq t_i \leq T} \pi_i(t_1, \dots, t_i; \Omega) dt_{j+1} \dots dt_i}{\pi_{j-1}^i(t_{j-1} / t_1, \dots, t_{j-2}; \Omega) \dots \pi_1^i(t_1; \Omega)}, \quad (9.3)$$

где $t_0 = T_0$.

Представленный выше метод моделирования случайных потоков непригоден для случая, когда $T = \infty$ и с вероятностью равной 1 на интервале наблюдения наступает бесконечное число событий, поскольку в этом случае все $p_k(\Omega)$ и $\pi_i(t_1, \dots, t_i; \Omega)$ с конечным индексом равны 0.

В связи с этим рассмотрим еще один способ параметрического задания потоков случайных событий. Поток будет задан, если известны ненормированные плотности $a_i(t_1, \dots, t_i), i = 1, 2, \dots$, для которых $a_i(t_1, \dots, t_i) dt_1 \dots dt_i$ есть вероятность наступления на Ω не менее i собы-

тий, первые i из которых попали в интервалы $(t_j - dt_j, t_j]$, $j = 1, 2, \dots, i$. Предположим, что первые $i-1$ событие наступили в моменты времени $t_1, \dots, t_{i-1}$. Для моделирования момента наступления следующего события найдем условную плотность вероятности

$$q_i(t_i / t_1, \dots, t_{i-1}) = \frac{a_i(t_1, \dots, t_i)}{a_{i-1}(t_1, \dots, t_{i-1})}; a_0 \equiv 1 \quad (9.4)$$

и вероятность d_i наступления i -го события на Ω при условии, что $i-1$ событие наступило в моменты времени $t_1, \dots, t_{i-1}$:

$$d_i = \int_{t_{i-1}}^T q_i(t / t_1, \dots, t_{i-1}) dt, t_0 = T_0, \quad (9.5)$$

где $q_i(t_i / t_{i-1}, \dots, t_1) dt$ – вероятность наступления i -го события в интервале $(t_i - dt_i, t_i]$ при условии, что $i-1$ предыдущих событий наступили в моменты времени $t_1, \dots, t_{i-1}$.

Для каждого d_i определяем дискретную случайную величину η_i , принимающую значения 0 или 1 с вероятностями $P\{\eta = 0\} = 1 - d_i, P\{\eta = 1\} = d_i$ и получаем ее реализацию. Если полученное значение η_i равно 0, то это означает, что на интервале наблюдения наступило $i-1$ событие и моделирование прекращается. Если $\eta_i = 1$, то на Ω наступит по крайней мере i событий. В этом случае момент наступления i -го события t_i может быть найден с использованием условной функции распределения с применением, например, метода обратных функций. Тогда уравнения для нахождения t_i будет выглядеть следующим образом:

$$F_i(t_i / t_1, \dots, t_{i-1}) = \alpha_i, \quad (9.6)$$

где

$$F_i(t_i / t_1, \dots, t_{i-1}) = \int_{t_{i-1}}^{t_i} \frac{q_i(t / t_1, \dots, t_{i-1})}{d_i} dt, \quad (9.7)$$

а α_i – независимые реализации базовой случайной величины. Процесс моделирования случайного потока продолжается до тех пор, пока значение реализации дискретной случайной величины η_i не станет равно 0.

Часто при моделировании случайных потоков с использованием совместных плотностей $a_i(t_1, \dots, t_i), i = 1, 2, \dots$ можно существенно упростить алгоритм моделирования, если расширить интервал наблюдения

$\Omega = [T_0; T]$ в сторону бесконечности и заменить его интервалом $\Omega' = [T_0; \infty)$. При использовании расширенной области Ω' , процесс моделирования продолжается до тех пор, пока очередной полученный момент наступления события t_i не станет больше T , а реализацией потока будут события, попавшие в интервал $[T_0; T]$. Использование расширенной области моделирования Ω' особенно удобно в случае, если для моделируемого случайного потока для любого конечного k $p_k(\Omega) = 0$. Для таких потоков из алгоритма моделирования, приведенного выше, можно исключить генерирование дискретной случайной величины η_i , поскольку для них все $d_i \equiv 1$ (обязательно наступает любое конечное число событий).

Между представленными выше плотностями распределения $\pi_i(t_1, \dots, t_i; \Omega)$ и $a_i(t_1, \dots, t_i)$ существует связь, которая определяется следующими математическими выражениями:

$$\begin{aligned} \pi_i(t_1, \dots, t_i; \Omega) &= a_i(t_1, \dots, t_i) - \int_{t_i}^T a_{i+1}(t_1, \dots, t_i, t_{i+1}) dt_{i+1} \\ a_i(t_1, \dots, t_i) &= \pi_i(t_1, \dots, t_i; \Omega) + \\ &+ \sum_{j=i+1}^{\infty} \int_{t_i}^T \int_{t_{i+1}}^T \dots \int_{t_{j-1}}^T \pi_j(t_1, \dots, t_i, t_{i+1}, \dots, t_j; \Omega) dt_{i+1} \dots dt_j, a_0 \equiv 1. \end{aligned} \quad (9.8)$$

Справедливость выражений (9.8) следует из вероятностного смысла плотностей $\pi_i(t_1, \dots, t_i; \Omega)$ и $a_i(t_1, \dots, t_i)$. Действительно, дифференциал $a_i(t_1, \dots, t_i) dt_1 \dots dt_i$ есть вероятность появления на Ω либо ровно i событий рассматриваемого потока в малой окрестности точек $t_1, \dots, t_i$ (что соответствует $\pi_i(t_1, \dots, t_i; \Omega)$), либо любого другого числа событий $v > i$, первые i из которых попали в малые окрестности точек $t_1, \dots, t_i$, а оставшиеся $v - i$ могут располагаться справа друг от друга в любом месте правее точки t_i (что определяется суммой во втором выражении (9.8)). Фактически интегрирование приведенных выше плотностей по некоторой переменной означает, что соответствующее событие потока может располагаться где угодно в пределах интервала интегрирования. Аналогичным образом можно дать интерпретацию первому из выражений (9.8). Согласно определению $a_{i+1}(t_1, \dots, t_i, t_{i+1}) dt_1 \dots dt_{i+1}$ есть вероятность возникновения на Ω $i + 1$ или более событий, причем первые $i + 1$ событие попадают в малые окрестности точек $t_1, \dots, t_{i+1}$. Проинтегрировав

данную плотность по переменной t_{i+1} , получим вероятность возникновения на Ω $i+1$ или более событий, при условии, что первые i событий попали в малые окрестности точек $t_1, \dots, t_i$, а оставшиеся находятся правее точки t_i . В то же время $a_i(t_1, \dots, t_i)dt_1 \dots dt_i$ есть вероятность появления на Ω i или более событий, первые i из которых попали в малые окрестности точек $t_1, \dots, t_i$. Тогда разность, представленная в первом уравнении (9.8), умноженная на $dt_1 \dots dt_i$, будет соответствовать вероятности возникновения на Ω ровно i событий потока в малых окрестностях точек $t_1, \dots, t_i$. Как следует из выше сказанного и в соответствии с определением плотностей $\pi_i(t_1, \dots, t_i, \Omega)$, разность в правой части первого выражения (9.8) есть ничто иное, как $\pi_i(t_1, \dots, t_i, \Omega)$.

Важно отметить, что плотности $a_i(t_1, \dots, t_i)$ могут быть представлены как плотности $\pi_i(t_1, \dots, t_i, [T_0, t_i])$. Для подтверждения этого рассмотрим, как изменяются плотности $\pi_i(t_1, \dots, t_i; \Omega)$ при переходе от области $\Omega = [T_0, T]$ к области $[T_0, t] \subset \Omega$. Из вероятностного смысла плотностей $\pi_i(t_1, \dots, t_i; [T_0, t])$ следует, что дифференциал $\pi_i(t_1, \dots, t_i; [T_0, t])dt_1 \dots dt_i$ представляет собой вероятность выпадения на $[T_0, t]$ ровно i точек потока в интервалах $(t_j - dt_j, t_j]$, $j = 1, 2, \dots, i$ независимо от факта появления в области $[t, T]$ точек рассматриваемого потока. Тогда

$$\begin{aligned} \pi_i(t_1, \dots, t_i, [T_0, t]) &= \pi_i(t_1, \dots, t_i, \Omega) + \\ &+ \sum_{j=i+1}^{\infty} \int_{t_{i+1}}^T \int_{t_{i+1}}^T \dots \int_{t_{j-1}}^T \pi_j(t_1, \dots, t_i, t_{i+1}, \dots, t_j; \Omega) dt_{i+1} \dots dt_j. \end{aligned} \quad (9.9)$$

Сравнивая последнее выражение со вторым выражением из (9.8), получаем

$$a_i(t_1, \dots, t_i) = \pi_i(t_1, \dots, t_i, [T_0, t_i]). \quad (9.10)$$

В качестве альтернативы приведенным выше способам задания случайных потоков при помощи совместных распределений моментов наступления событий $\xi_1, \dots, \xi_i$, $i = 1, 2, \dots$ можно рассмотреть способы, использующие совместные распределения интервалов между соседними событиями случайного потока $\Delta_1, \dots, \Delta_i$, $i = 1, 2, \dots$, где

$$\Delta_j = t_j - t_{j-1}, \quad j = 1, 2, \dots, i. \quad (9.11)$$

Определим для интервалов между событиями случайного потока системы совместных плотностей, аналогичных совместным плотностям для моментов наступления событий, которые были определены выше.

Пусть $\pi_i^\Delta(\tau_1, \dots, \tau_i, \Omega), i=1, 2, \dots$ – совместная плотность распределения, для которой $\pi_i^\Delta(\tau_1, \dots, \tau_i, \Omega) d\tau_1 \dots d\tau_i$ – совместная вероятность попадания временных интервалов между событиями потока в малые области $(\tau_j - d\tau_j, \tau_j], j=1, 2, \dots, i$, при условии, что на интервале наблюдения было зафиксировано ровно i интервалов между событиями случайного потока. Соответственно $a_i^\Delta(\tau_1, \dots, \tau_i)$ – совместные плотности распределения, для которых $a_i^\Delta(\tau_1, \dots, \tau_i) d\tau_1 \dots d\tau_i$ есть вероятность того, что на Ω было зафиксировано не менее i интервалов между событиями случайного потока, первые i из которых попали в малые области $(\tau_j - d\tau_j, \tau_j], j=1, 2, \dots, i$.

Плотности для интервалов между событиями случайного потока связаны с плотностями для моментов наступления его событий следующим образом:

$$\pi_i^\Delta(\tau_1, \dots, \tau_i, \Omega) = \pi_i(T_0 + \tau_1, T_0 + \tau_1 + \tau_2, \dots, T_0 + \tau_1 + \tau_2 + \dots + \tau_i, \Omega);$$

$$\pi_i(t_1, t_2, \dots, t_i, \Omega) = \pi_i^\Delta(t_1 - T_0, t_2 - t_1, \dots, t_i - t_{i-1}, \Omega);$$

$$a_i^\Delta(\tau_1, \dots, \tau_i) = a_i(T_0 + \tau_1, T_0 + \tau_1 + \tau_2, \dots, T_0 + \tau_1 + \tau_2 + \dots + \tau_i);$$

$$a_i(t_1, t_2, \dots, t_i) = a_i^\Delta(t_1 - T_0, t_2 - t_1, \dots, t_i - t_{i-1}),$$

где $i=1, 2, \dots$. Процесс моделирования случайных интервалов между событиями аналогичен приведенному ранее процессу моделирования моментов появления событий потока. После получения интервалов между событиями случайного потока моменты возникновения событий могут быть получены с использованием (9.11):

$$t_i = \tau_i + t_{i-1}.$$

При моделировании случайных потоков важным моментом является построение единого подхода к их математическому описанию, который позволял бы одинаковым образом получать как описанные выше совместные плотности вероятностей, используемые в процессе моделирования, так и основные характеристики моделируемого потока, позволяющие контролировать правильность работы моделирующего алгоритма. Для построения такого математического описания случайных потоков используем введенные ранее совместные плотности $\pi_i(t_1, \dots, t_i; \Omega)$. При помощи данных плотностей можно получать выражения для вероятностей $p_i(\Omega)$ выпадения на Ω ровно i событий:

$$p_i(\Omega) = \int_{T_0}^T \int_{t_1}^T \dots \int_{t_{i-1}}^T \pi_i(t_1, \dots, t_i; \Omega) dt_1 \dots dt_i. \quad (9.12)$$

Для удобства использования в дальнейшем доопределим функцию $\pi_i(t_1, \dots, t_i; \Omega)$ на всю i -мерную область изменения ее аргументов так, чтобы она была симметричной при произвольной перестановке аргументов. Для этого используем операцию симметризации, в соответствии с которой доопределенная функция будет средним арифметическим исходных функций со всевозможными перестановками ее аргументов. Полученная плотность распределения однозначно описывает рассматриваемый поток и допускает расширение интегрирования в выражении (9.12) на всю область Ω^i . С учетом того, что возможно $i!$ различных перестановок аргументов, выражение (9.12) может быть представлено в следующем виде:

$$\begin{aligned} p_i(\Omega) &= \frac{1}{i!} \int_{T_0}^T \int_{T_0}^T \dots \int_{T_0}^T \pi_i(t_1, \dots, t_i; \Omega) dt_1 \dots dt_i = \\ &= \frac{1}{i!} \int_{\Omega^i} \pi_i(t_1, \dots, t_i; \Omega) dt_1 \dots dt_i, \quad i = 0, 1, \dots \end{aligned} \quad (9.13)$$

Для вероятностей $p_i(\Omega)$, представленных в (9.13), выполняется условие нормировки

$$\sum_{i=0}^{\infty} p_i(\Omega) = 1. \quad (9.14)$$

Поскольку плотности $\pi_i(t_1, \dots, t_i; \Omega)$ определяют как число событий потока, возникающих на интервале наблюдения, так и их местоположение на временной оси, то их удобно использовать для выполнения операции усреднения. Пусть $q(\xi_1, \xi_2, \dots, \xi_v)$ – некоторая функция, которая зависит как от числа точек потока v , так и от моментов наступления событий $\xi_1, \xi_2, \dots, \xi_v$ в реализации потока. Тогда среднее от такой функции по числу событий и их местоположению на временной оси задается следующим выражением:

$$\langle q \rangle_{\xi, v} = \sum_{i=0}^{\infty} \frac{1}{i!} \int_{\Omega^i} q_i(t_1, \dots, t_i) \pi_i(t_1, \dots, t_i; \Omega) dt_1 \dots dt_i. \quad (9.15)$$

Используя операцию усреднения, представленную в выражении (9.15), введем понятие *производящего функционала потока случайных событий (ПФЛ)*:

$$L[u, \Omega] = \left\langle \prod_{j=1}^v [u(\xi_j) + 1] \right\rangle_{\xi, v} = \sum_{i=0}^{\infty} \frac{1}{i!} \int_{\Omega^i} \pi_i(t_1, \dots, t_i; \Omega) \prod_{j=1}^i [u(t_j) + 1] dt_1 \dots dt_i, \quad (9.16)$$

где $u(\cdot)$ – некоторая пробная функция одного аргумента.

ПФЛ случайного потока, представленный выражением (9.16), является единым и достаточно универсальным математическим инструментом для описания случайных потоков событий. Важная роль ПФЛ при описании случайных потоков состоит в том, что, зная его выражение для рассматриваемого потока, можно достаточно просто получать для этого потока совместные плотности распределения $\pi_i(t_1, \dots, t_i; \Omega)$ и $a_i(t_1, \dots, t_i)$, а также ряд характеристик потока, таких, например, как его интенсивность, корреляционная функция и распределение числа событий на интервале наблюдения.

Приведем некоторые свойства, которыми обладает ПФЛ.

Свойство 1. Как следует из выражений (9.16), (9.14) и (9.13) $L[0, \Omega] = 1$.

Свойство 2. $L[-1, \Omega] = p_0(\Omega) = \pi_0(\Omega)$.

Свойство 3. При наложении m независимых случайных потоков с ПФЛ $L^{(i)}[u, \Omega]$, $i = 1, \dots, m$, ПФЛ суммарного потока равен произведению ПФЛ отдельных потоков, т. е.

$$L[u; \Omega] = \prod_{i=1}^m L^{(i)}[u; \Omega].$$

Свойство 4. При использовании ПФЛ для различных целей часто применяется операция функционального дифференцирования. Приведем ее упрощенную интерпретацию, которая используется при решении прикладных задач. Пусть $B[u]$ – некоторый функционал, аргумент $u(t)$ которого является функцией от переменной t . Дадим аргументу $u(t)$ некоторое приращение $\delta u(t)$ и выделим линейную часть приращения функционала

$$B[u + \delta u] - B[u] = \int_{\Omega} E[u; t] \delta u(t) dt + o(\delta u),$$

где $E[u; t]$ зависит от t как обычная функция, является функционалом по $u(t)$ и представляет собой функциональную производную $B[u]$ по u в точке t :

$$E[u; t] = \frac{\delta B[u]}{\delta u(t)}.$$

Вторая производная является производной от первой и т. д. Связь между производными соседних порядков выглядит следующим образом:

$$\frac{\delta^i B[u]}{\delta u(t_1) \dots \delta u(t_{i-1}) \delta u(t_i)} = \frac{\delta}{\delta u(t_i)} \frac{\delta^{i-1} B[u]}{\delta u(t_1) \dots \delta u(t_{i-1})}.$$

Рассмотрим ряд наиболее типичных примеров вычисления функциональной производной от различных функционалов:

- 1) пусть $B[u] = a(t)$, причем $a(t)$ не зависит от $u(t)$. Тогда $\delta B[u] / \delta u(t) \equiv 0$;
- 2) пусть $B[u] = \int_{\Omega} a(t) u(t) dt$. В этом случае $\delta B[u] / \delta u(t) = a(t)$;
- 3) пусть $B[u] = a(t) u(t)$. Тогда $\delta B[u] / \delta u(t_1) = a(t_1) \delta(t_1 - t)$, где $\delta(t)$ – функция Дирака;
- 4) пусть $B[u] = \int_{\Omega^2} F(t_1, t_2) u(t_1) u(t_2) dt_1 dt_2$. Тогда

$$\frac{\delta B[u]}{\delta u(t)} = \int_{\Omega} F(t, t_2) u(t_2) dt_2 + \int_{\Omega} F(t_1, t) u(t_1) dt_1 = 2 \int_{\Omega} \{F(t, \tau)\}_{sim(2)} u(\tau) d\tau,$$

где операция $\{\dots\}_{sim(i)}$ обозначает операцию симметризации функции, стоящей в фигурных скобках, относительно аргументов, число которых указано в круглых скобках после знака *sim*. Если в рассматриваемом примере функция $F(t_1, t_2)$ симметрична, т. е. $F(t_1, t_2) = F(t_2, t_1)$, то

$$\frac{\delta B[u]}{\delta u(t)} = 2 \int_{\Omega} F(t, \tau) u(\tau) d\tau;$$

- 5) пусть $B[u] = f[v[u, \tau]]$, где $f[v]$ – функционал от функции $v[u, \tau]$ аргумента τ , которая в свою очередь одновременно является функционалом от u . В данном случае формула функционального дифференцирования аналогична взятию производной от сложной функции

$$\frac{\delta B[u]}{\delta u(t)} = \int \frac{\delta f[v]}{\delta v[u; \tau]} \frac{\delta v[u; \tau]}{\delta u(t)} d\tau.$$

При помощи ПФЛ случайного потока можно получать совместные плотности распределения событий потока. Для этой цели применим определенную выше операцию функционального дифференцирования. Функциональная производная i -го порядка при $u(t) = -1$ от ПФЛ, представленного выражением (9.16), дает плотность $\pi_i(t_1, \dots, t_i; \Omega)$

$$\pi_i(t_1, \dots, t_i; \Omega) = \frac{\delta^i L[u, \Omega]}{\delta u(t_1) \dots \delta u(t_i)} \Big|_{u(\cdot)=-1}, i=1,2,\dots \quad (9.17)$$

Для нахождения плотностей $a_i(t_1, \dots, t_i)$ воспользуемся их связью с плотностями $\pi_i(t_1, \dots, t_i; \Omega)$, которая представлена в выражении (9.10). С учетом данной связи очевидно, что плотности $a_i(t_1, \dots, t_i)$ могут быть получены путем i -кратного дифференцирования ПФЛ (9.16) при $u(t) = -1$ и при условии, что область наблюдения $\Omega = [T_0, T]$ заменена на интервал $[T_0, t_i]$:

$$a_i(t_1, \dots, t_i) = \frac{\delta^i L[u, [T_0, t_i]]}{\delta u(t_1) \dots \delta u(t_i)} \Big|_{u(\cdot)=-1}, i=1,2,\dots \quad (9.18)$$

Кроме нахождения совместных плотностей распределения случайного потока ПФЛ может быть использован также для нахождения различных статистических характеристик рассматриваемого потока событий. Так, в результате i -кратного дифференцирования ПФЛ при $u(t) = 0$ могут быть получены симметрические *моментные функции* $f_i(t_1, \dots, t_i), i=1,2,\dots$.

$$f_i(t_1, \dots, t_i) = \frac{\delta^i L[u, \Omega]}{\delta u(t_1) \dots \delta u(t_i)} \Big|_{u(\cdot)=0}, i=1,2,\dots \quad (9.19)$$

Учитывая формулы (9.16) и (9.19), можно получить выражение, связывающее моментные функции $f_i(t_1, \dots, t_i)$ и совместные плотности распределения $\pi_i(t_1, \dots, t_i; \Omega)$:

$$f_i(t_1, \dots, t_i) = \sum_{k=0}^{\infty} \frac{1}{k!} \int_{\Omega^k} \pi_{i+k}(t_1, \dots, t_i, t_{i+1}, \dots, t_{i+k}; \Omega) dt_{i+1} \dots dt_{i+k} \quad (9.20)$$

Последнее соотношение позволяет дать вероятностную интерпретацию моментных функций $f_i(t_1, \dots, t_i)$. Дифференциал $f_i(t_1, \dots, t_i) dt_1 \dots dt_i$ соответствует вероятности попадания событий в малые интервалы $[t_j, t_j + dt_j), j=1, \dots, i$ безотносительно к появлению точек вне данных интервалов. Особую роль среди моментных функций играют моментные функции первого и второго порядков. Моментная функция первого порядка $f_1(t)$ представляет собой *интенсивность случайного потока*. Моментная функция второго порядка $f_2(t_1, t_2)$ представляет собой *корреляционную функцию случайного потока*.

При описании случайных потоков, кроме ПФЛ, важную роль также играет производящая функция (ПФ) числа точек случайного потока на интервале наблюдения Ω , определяемая следующим образом:

$$\Theta(z; \Omega) = \langle z^v \rangle_v, \quad (9.21)$$

где z – некоторая вспомогательная переменная, v – число точек потока на интервале наблюдения. С учетом выражения (9.15) ПФ (9.21) может быть представлена следующим образом:

$$\Theta(z; \Omega) = \sum_{i=0}^{\infty} z^i p_i(\Omega), \quad (9.22)$$

где $p_i(\Omega)$ – вероятность выпадения на Ω ровно i событий потока, определяемая выражением (9.13). Тогда из формул (9.22) и (9.16) следует, что ПФ случайного потока и его ПФЛ связаны следующим соотношением:

$$\Theta(z; \Omega) = L[z-1; \Omega]. \quad (9.23)$$

Таким образом, как следует из выражения (9.23), ПФ случайного потока может быть также получена из его ПФЛ путем подстановки в выражение для ПФЛ вместо $u(t)$ величины $z-1$. В результате, зная ПФЛ случайного потока, и получив в соответствии с (9.23) его ПФ, можно найти выражение для еще одной важной характеристики потока – *распределения числа его событий на интервале наблюдения* путем i -кратного дифференцирования ПФ по вспомогательной переменной z

$$p_i(\Omega) = \frac{d^i \Theta[z; \Omega]}{i! dz^i} \Big|_{z=0}. \quad (9.24)$$

При решении многих прикладных задач возникает необходимость моделирования потоков, обладающих определенными свойствами, причем их моделирование зачастую оказывается более простым и удобным. Далее рассмотрим ряд таких свойств.

Поток однородных событий называется *стационарным*, если $p_k(t, t_0)$ вероятность появления k событий на интервале времени $(t_0, t_0 + t)$ не зависит от начального момента времени t_0 , а зависит только от t и k . Для стационарных потоков их интенсивность $f_1(t) = \text{const}$, а корреляционная функция $f_2(t_1, t_2) = f_2(t_2 - t_1) = f_2(\tau)$ зависит только от разности аргументов.

Случайный поток событий обладает свойством *ординарности*, если в нем практически невозможны наступления двух или более событий в один и тот же момент времени. Это означает, что для ординарного пото-

ка вероятность $\psi(t_0, t)$ появления двух и более событий в интервале $(t_0, t_0 + t)$ при любом t_0 характеризуется следующим свойством:

$$\lim_{t \rightarrow 0} \frac{\psi(t_0, t)}{t} = 0.$$

Случайный поток называется потоком с *ограниченным последствием*, если случайные величины Δ_i (интервалы между временами наступления событий потока) независимы. Совместная функция плотности распределения интервалов между событиями:

$$\varphi(\tau_1, \tau_2, \dots, \tau_i) = \varphi_1(\tau_1)\varphi_2(\tau_2) \dots \varphi_i(\tau_i),$$

где $\tau_i = t_i - t_{i-1}$, $i = 2, 3, \dots$, а $\tau_1 = t_1 - T_0$ – расстояние до первого события на интервале наблюдения Ω .

Случайный поток называется *потоком без последствия*, если для любой системы попарно непересекающихся подынтервалов $\Omega_i, i = 1, \dots, n$ число событий, попадающих в каждый из этих подынтервалов, представляет собой случайную величину, не зависящую от числа событий в других интервалах.

Важную роль на практике играет ординарный поток без последствия, который еще часто называют потоком Пуассона. Его прикладное значение состоит в том, что суперпозиция большого числа независимых потоков произвольного типа в условиях физического эксперимента, приближается к потоку Пуассона. Далее будут рассмотрены различные типы случайных потоков и приведены алгоритмы их численного моделирования.

Л е к ц и я 10

МОДЕЛИРОВАНИЕ ТИПОВЫХ СЛУЧАЙНЫХ ПОТОКОВ (Часть 1)

Одноточечный поток. *Одноточечный поток* – это поток, состоящий из не более чем одного события, момент наступления которого имеет заданную плотность распределения $s(t)$.

Теоретическое описание. В общем случае, если плотность одноточечного потока не нормирована к 1 на интервале наблюдения $\Omega = [T_0, T]$ (т. е. $\int_{\Omega} s(t) dt < 1$), то выражения для определенных выше совместных плотностей $\pi_i(t_1, \dots, t_i; \Omega)$ будут выглядеть следующим образом:

$$\pi_i(t_1, \dots, t_i; \Omega) = \begin{cases} 1 - \int_{\Omega} s(t) dt, & i = 0 \\ s(t_1) & i = 1 \\ 0 & i > 1 \end{cases} \quad (10.1)$$

В случае если $\int_{\Omega} s(t) dt = 1$, то последнее выражение будет иметь более простой вид

$$\pi_i(t_1, \dots, t_i; \Omega) = \begin{cases} s(t_1) & i = 1 \\ 0 & i \neq 1 \end{cases} \quad (10.2)$$

Подставляя выражение (10.1) в общую формулу ПФЛ случайного потока (9.16), получаем выражение для ПФЛ одноточечного потока:

$$\begin{aligned} L[u; \Omega] &= \pi_0 + \int_{\Omega} \pi_1(t; \Omega)(u(t) + 1) dt = \\ &= 1 - \int_{\Omega} s(t) dt + \int_{\Omega} s(t) dt + \int_{\Omega} s(t) u(t) dt = 1 + \int_{\Omega} s(t) u(t) dt. \end{aligned} \quad (10.3)$$

Производящая функция числа событий одноточечного потока на интервале наблюдения в соответствии с (9.23)

$$\Theta(z; \Omega) = 1 + (z - 1) \int_{\Omega} s(t) dt.$$

Тогда число событий потока на Ω :

$$p_i(\Omega) = \begin{cases} \Theta(z; \Omega)|_{z=0} = 1 - \int_{\Omega} s(t) dt & i = 0 \\ \left. \frac{d\Theta(z; \Omega)}{dz} \right|_{z=0} = \int_{\Omega} s(t) dt & i = 1 \end{cases} \quad (10.4)$$

Поток Бернулли. *Потоком Бернулли* называется случайный поток, порождаемый фиксированным числом k независимых источников, каждый из которых генерирует не более 1 события на интервале наблюдения Ω в случайный момент времени ξ с парциальной плотностью распределения $s_i(t), i = 1, \dots, k$. В общем случае парциальные плотности распределения могут быть не нормированы на Ω , т. е.

$$\int_{\Omega} s_i(t) dt \leq 1.$$

Теоретическое описание. Построим ПФЛ потока Бернулли и найдем в соответствии с приведенной ранее общей теорией выражения для со-

вместных плотностей, необходимых при построении алгоритмов моделирования. Кроме того, найдем основные характеристики данного потока.

Производящий функционал потока Бернулли в соответствии со свойствами ПФЛ и с учетом выражения (10.3) имеет следующий вид:

$$L[u; \Omega] = \prod_{i=1}^k \left[1 + \int_{\Omega} s_i(t) u(t) dt \right]. \quad (10.5)$$

Выполняя функциональное дифференцирование последнего выражения и с учетом формулы (9.17), получаем выражения для плотностей $\pi_i(t_1, \dots, t_i; \Omega)$ потока Бернулли:

$$\left. \begin{aligned} \pi_0(\Omega) &= \prod_{i=1}^k (1 - \varphi_i) \\ \pi_1(t_1; \Omega) &= \pi_0(\Omega) \sum_{i=1}^k \frac{s_i(t_1)}{1 - \varphi_i} \\ \pi_2(t_1, t_2; \Omega) &= \pi_0(\Omega) 2! \left\{ \sum_{i < j} \frac{s_i(t_1) s_j(t_2)}{(1 - \varphi_i)(1 - \varphi_j)} \right\}_{sim(2)} \\ &\dots \\ \pi_k(t_1, \dots, t_k; \Omega) &= k! \left\{ \prod_{i=1}^k s_i(t_i) \right\}_{sim(k)} \end{aligned} \right\}, \quad (10.6)$$

где φ_i – вероятность того, что i -й одноточечный источник сгенерирует событие на интервале наблюдения, которая в соответствии с выражением (10.4) равна

$$\varphi_i = \int_{\Omega} s_i(t) dt. \quad (10.7)$$

Найдем моментные функции первого и второго порядков для потока Бернулли, определяющие соответственно его интенсивность и корреляционную функцию. В соответствии с выражениями (9.19) и (10.5) получаем

$$\begin{aligned} f_1(t) &= \sum_{i=1}^k s_i(t), \\ f_2(t_1, t_2) &= 2! \left\{ \sum_{j < i} s_i(t_1) s_j(t_2) \right\}_{sim(2)}. \end{aligned} \quad (10.8)$$

Производящая функция числа событий на интервале наблюдения для потока Бернулли в соответствии с выражениями (9.23) и (10.5) имеет вид

$$\Theta(z; \Omega) = \prod_{i=1}^k [1 + (z-1)\varphi_i]. \quad (10.9)$$

Тогда в соответствии с формулой (9.24) получаем выражения для вероятностей выпадения определенного числа событий потока Бернулли на интервале наблюдения

$$\left. \begin{aligned} p_0(\Omega) &= \prod_{i=1}^k (1 - \varphi_i) \\ p_1(\Omega) &= \sum_{i=1}^k \varphi_i \prod_{j=1, j \neq i}^k (1 - \varphi_j) = p_0(\Omega) \sum_{i=1}^k \frac{\varphi_i}{1 - \varphi_i} \\ p_2(\Omega) &= p_0(\Omega) \sum_{i < j} \frac{\varphi_i \varphi_j}{(1 - \varphi_i)(1 - \varphi_j)} \\ &\dots \\ p_k(\Omega) &= \prod_{i=1}^k \varphi_i \end{aligned} \right\}. \quad (10.10)$$

Выражения (10.10) могут быть легко получены с использованием теории вероятностей путем представления вероятностей $p_k(\Omega)$ в виде суммы вероятностей возникновения событий на Ω от определенного числа одноточечных источников с определенными порядковыми номерами, при условии, что события от остальных источников не были сгенерированы. В случае если плотности распределения, а следовательно, и вероятности φ_i для всех одноточечных потоков одинаковы, т. е. $\varphi_1 = \varphi_2 = \dots = \varphi_k = \varphi$, то выражения (10.10) могут быть записаны в едином виде

$$p_i(\Omega) = C_k^i \varphi^i (1 - \varphi)^{k-i}, \quad (10.11)$$

где $C_k^i = \frac{k!}{i!(k-i)!}$ — число комбинаций из k по i .

Алгоритмы моделирования. Исходя из определения, можно представить поток Бернулли как поток, образованный путем наложения k независимых одноточечных потоков. На основании данного механизма формирования потока Бернулли может быть построен один из методов его моделирования. Данный метод состоит из следующих шагов.

Шаг 1. Для всех k источников одноточечных потоков определим, были ли сгенерированы их события. С этой целью в соответствии с из-

ложенным выше общим подходом к моделированию, для каждого одноточечного потока введем вспомогательную дискретную случайную величину $\eta_i, i = 1, \dots, k$, принимающую два возможных значения: 0 с вероятностью $P\{\eta_i = 0\} = 1 - \varphi_i$ и 1 с вероятностью $P\{\eta_i = 1\} = \varphi_i$, где φ_i определено выражением (10.7). Получив реализации η_i , определим, были ли сгенерированы события соответствующих одноточечных потоков на интервале наблюдения. Индикатором того, что единственное событие i -го одноточечного потока будет сгенерировано, является получение реализации соответствующей случайной величины η_i , равной 1.

Шаг 2. Для тех одноточечных потоков, для которых на предыдущем шаге $\eta_i = 1$, генерируем местоположение их событий на временной оси, используя следующие выражения для плотностей распределения $\rho_i(t) = s_i(t)/\varphi_i$. На данном шаге может применяться наиболее удобный алгоритм для моделирования непрерывных случайных величин.

Шаг 3. События всех одноточечных потоков, полученные на предыдущих этапах, объединяются и упорядочиваются по времени путем сортировки по возрастанию. В результате полученная совокупность событий и будет реализацией потока Бернулли.

Преимуществом представленного выше алгоритма моделирования является его универсальность, поскольку он может использоваться для моделирования любого потока Бернулли. Однако вычислительное время, необходимое для моделирования, значительно возрастает при увеличении числа накладываемых одноточечных потоков по причине наличия операции сортировки. В связи с этим для потоков Бернулли, обладающих определенными свойствами, были предложены альтернативные методы моделирования.

Рассмотрим частный случай – *однородный поток Бернулли*. Поток Бернулли называют однородным, если все парциальные плотности одинаковы и нормированы на интервале наблюдения, т. е. $s_i(t) = s(t), i = 1, 2, \dots, k$; $\int_{\Omega} s(t) dt = 1$.

В данном случае можно построить алгоритм моделирования, основанный на приведенном ранее общем методе, использующем выражения для плотностей $\pi_i(t_1, \dots, t_i; \Omega)$ и вероятностей $p_i(\Omega)$ появления точно i событий на интервале наблюдения.

Найдем выражения для плотностей $\pi_i(t_1, \dots, t_i; \Omega)$ и вероятностей $p_i(\Omega)$ однородного потока Бернулли. С учетом выражений (10.6) и (10.11) получим

$$\begin{cases} \pi_i(t_1, \dots, t_i; \Omega) = 0, & p_i = 0, \quad i \neq k \\ \pi_k(t_1, \dots, t_k; \Omega) = k! \prod_{i=1}^k s(t_i), & p_k = 1 \end{cases} \quad (10.12)$$

Таким образом, задача моделирования такого потока сводится к задаче моделирования случайного вектора с совместной плотностью распределения $\pi_k(t_1, \dots, t_k; \Omega)$. При этом можно воспользоваться методом, предложенным ранее в разделе, посвященном моделированию случайных векторов. Отдельные координаты случайного вектора можно получить последовательно с использованием условных функций распределения и метода обратных функций в соответствии с (9.1), где условные функции распределения могут быть получены в соответствии с (9.2), (9.3) и имеют вид

$$F_i^k(t_i | t_1, \dots, t_{i-1}) = 1 - \left[\frac{\int_{t_i}^T s(x) dx}{\int_{t_{i-1}}^T s(x) dx} \right]^{k-i+1}, \quad i = 1, \dots, k; t_0 = T_0. \quad (10.13)$$

Пример 10.1. Применим описанный выше метод моделирования для генерации однородного потока Бернулли, для которого парциальные плотности определены как $s(t) = 1$ на интервале наблюдения $\Omega = [0; 1]$. Для такого потока в соответствии с формулой (10.13) получаем

$$F_i^k(t_i / t_{i-1}, \dots, t_1) = 1 - \left(\frac{1 - t_i}{1 - t_{i-1}} \right)^{k-i+1}.$$

Далее, используя метод обратных функций в соответствии с (9.1), получаем выражение для вычисления моментов наступления событий генерируемого потока:

$$t_i = 1 - (1 - t_{i-1}) \alpha_i^{1/(k-i+1)}; \quad i = 1, \dots, k, t_0 = 0,$$

где α_j – независимые реализации базовой случайной величины.

Для однородного потока Бернулли с $s(t) = 1$ на интервале наблюдения $\Omega = [0; 1]$ существует метод моделирования, основанный на получении событий потока Бернулли на основе сгенерированных событий стационарного потока Пуассона с интенсивностью $\lambda = 1$ (алгоритмы моделирования потока Пуассона будут рассмотрены ниже). В соответствии с данным методом алгоритм моделирования строится следующим образом.

Шаг 1. Последовательно получают $k + 1$ момент наступления событий стационарного потока Пуассона с интенсивностью $\lambda = 1$ на интервале $[0, \infty]$. Обозначим эти моменты времени как $\tau_1, \tau_2, \dots, \tau_{k+1}$.

Шаг 2. Вычисляем моменты наступления событий потока Бернулли в соответствии со следующими выражениями:

$$t_1 = \frac{\tau_1}{\tau_{k+1}}; t_2 = \frac{\tau_2}{\tau_{k+1}}; \dots; t_k = \frac{\tau_k}{\tau_{k+1}}.$$

Зависимость временных затрат на моделирование от числа источников в схеме Бернулли качественно представлена на рис. 10.1.

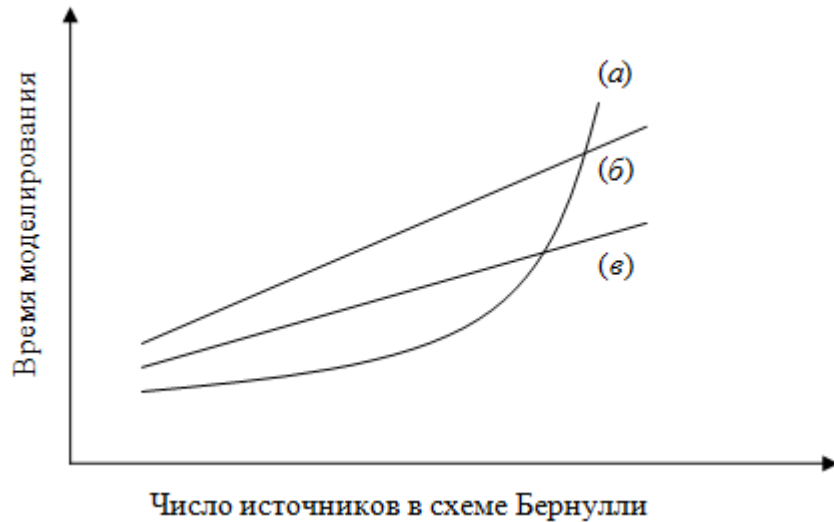


Рис. 10.1. Зависимость временных затрат на моделирование от числа источников в схеме Бернулли: (а) – алгоритм, основанный на наложении односточечных потоков; (б) – алгоритм на основе выражения (10.13); (в) – алгоритм, основанный на генерации стационарного потока Пуассона

Рекуррентный поток. Рекуррентным называется случайный поток, у которого протяженности интервалов между соседними точками независимы и имеют одинаковые распределения.

Теоретическое описание. Пусть $b(\tau)$ – функция плотности распределения интервалов между событиями рекуррентного потока на интервале наблюдения $\Omega = [T_0, T]$. Обозначим через $b_1(\tau)$ функцию плотности распределения расстояния от начала интервала наблюдения до первого события потока. Тогда выражение для совместных плотностей распределения расстояний между событиями рекуррентного потока $a_i^\Delta(\tau_1, \dots, \tau_i)$ можно представить в следующем виде:

$$a_i^\Delta(\tau_1, \dots, \tau_i) = b_1(\tau_1) \prod_{j=2}^i b(\tau_j), i = 1, 2, \dots \quad (10.14)$$

Алгоритм моделирования. Выражение (10.14) может быть использовано для построения алгоритма моделирования рекуррентного потока. Перейдем к расширенной области наблюдения $\Omega' = [T_0, \infty]$. Тогда, под-

ставляя (10.14) в выражения для условных плотностей вероятности аналогичных (9.4) и вероятностей (9.5), получим

$$\begin{aligned} q_1^\Delta(\tau_1) &= b_1(\tau_1); \\ q_i^\Delta(\tau_i | \tau_1, \dots, \tau_{i-1}) &= b(\tau_i), \quad i = 2, 3, \dots, \end{aligned}$$

где $d_i = 1$, $i = 1, 2, \dots$.

Из последних выражений следует, что расстояния от начала интервала наблюдения до первого события потока τ_1 и между событиями рекуррентного потока τ_i могут быть найдены путем решения следующих уравнений:

$$\begin{cases} \int_0^{\tau_1} b_1(x) dx = \alpha_1 \\ \int_0^{\tau_i} b(x) dx = \alpha_i, \quad i = 2, 3, \dots \end{cases}, \quad (10.15)$$

где $\alpha_1, \alpha_2, \dots$ – независимые реализации базовой случайной величины. Получив реализации интервалов между событиями потока можно определить местоположение событий потока согласно (9.11)

$$t_i = T_0 + \sum_{j=1}^i \tau_j, \quad i = 1, 2, \dots \quad (10.16)$$

Моделирование потока продолжается до тех пор, пока очередное сгенерированное событие не окажется за пределами области Ω .

Пример 10.2. Задача состоит в моделировании рекуррентного потока на интервале наблюдения $\Omega = [0, c]$, для которого заданы следующие плотности распределения первого и последующих интервалов между событиями:

$$\begin{aligned} b(\tau_i) &= b(\tau) = \frac{1}{c}, \quad i = 2, 3, \dots, \\ b_1(\tau_1) &= \frac{1 - \int_0^{\tau_1} b(\tau) d\tau}{\int_0^c \tau b(\tau) d\tau} = \frac{2}{c^2} (c - \tau_1). \end{aligned}$$

Интегрируя $b_1(\tau)$ и $b(\tau)$ и решая уравнения (10.15), получаем

$$\begin{cases} \int_0^{\tau_1} \frac{2}{c^2} (c - x) dx = 1 - \left(1 - \frac{\tau_1}{c}\right)^2 = \alpha_1 \Rightarrow \tau_1 = c(1 - \sqrt{\alpha_1}) \\ \int_0^{\tau_i} \frac{1}{c} dx = \frac{\tau_i}{c} = \alpha_i \Rightarrow \tau_i = c\alpha_i, \end{cases} \quad i = 2, 3, \dots$$

Определив таким образом реализации τ_i , далее находим реализации событий потока по формуле (10.16).

Поток Пуассона. Наиболее простое определение потока Пуассона может быть дано при помощи потока Бернулли. *Потоком Пуассона* называется однородный поток Бернулли, для которого число k независимых одноточечных источников случайно и подчиняется дискретному закону Пуассона с параметром Λ :

$$p_k = \frac{\Lambda^k}{k!} e^{-\Lambda}, \Lambda > 0. \quad (10.17)$$

Теоретическое описание. Найдем ПФЛ потока Пуассона, используя представленную в его определении связь данного потока с потоком Бернулли. ПФЛ потока Пуассона может быть получен путем усреднения ПФЛ потока Бернулли по числу ν одноточечных источников. Используя выражение (10.5) и учитывая, что для однородного потока Бернулли все парциальные плотности равны между собой, получаем

$$\begin{aligned} L[u, \Omega] &= \left\langle \prod_{i=1}^{\nu} \left[1 + \int_{\Omega} s(t) u(t) dt \right] \right\rangle_{\nu} = \left\langle \left[1 + \int_{\Omega} s(t) u(t) dt \right]^{\nu} \right\rangle_{\nu} = \\ &= \sum_{k=0}^{\infty} p_k \left[1 + \int_{\Omega} s(t) u(t) dt \right]^k. \end{aligned}$$

Подставляя в последнее выражение значения вероятностей p_k , определенные в (10.17), получаем

$$L[u; \Omega] = \sum_{k=0}^{\infty} \frac{\Lambda^k}{k!} e^{-\Lambda} \left[1 + \int_{\Omega} s(t) u(t) dt \right]^k. \quad (10.18)$$

Последнее выражение представляет собой не что иное, как разложение экспоненты в ряд Тейлора

$$e^{\Lambda(x-1)} = \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{d^k}{dx^k} (e^{\Lambda(x-1)}) \right)_{x=0} x^k = \sum_{k=0}^{\infty} \frac{\Lambda^k}{k!} e^{-\Lambda} x^k. \quad (10.19)$$

Сравнивая выражения (10.18) и (10.19), получаем окончательное выражение для ПФЛ потока Пуассона

$$L[u; \Omega] = \exp \left\{ \int_{\Omega} \lambda(t) u(t) dt \right\}, \quad (10.20)$$

где функция $\lambda(t) = \Lambda s(t)$, как будет показано далее, представляет собой *интенсивность потока Пуассона*. Как следует из выражения (10.20), ПФЛ потока Пуассона, а следовательно, и все его основные характеристики, полностью определяются интенсивностью $\lambda(t)$. Интенсивность потока Пуассона $\lambda(t)$ следует рассматривать как мгновенную интенсивность потока в некоторый момент времени t . Кроме того, вводится также понятие средней интенсивности потока на некотором интервале времени $[t_0, t_0 + \Delta t]$, которая может быть определена следующим образом:

$$\frac{\int_{t_0}^{t_0+\Delta t} \lambda(t) dt}{\Delta t} = \frac{\Lambda(t_0, \Delta t)}{\Delta t}, \quad (10.21)$$

где $\Lambda(t_0, \Delta t)$ – среднее число событий на интервале $[t_0, t_0 + \Delta t]$. Аналогично можно определить смысл параметра дискретного распределения Пуассона Λ , входящего в формулу (10.17). Его значение может быть получено путем интегрирования интенсивности потока Пуассона $\lambda(t)$ по всему интервалу наблюдения

$$\int_{\Omega} \lambda(t) dt = \int_{\Omega} \Lambda s(t) dt = \Lambda \int_{\Omega} s(t) dt = \Lambda. \quad (10.22)$$

Из последнего выражения следует, что Λ представляет собой среднее число событий потока Пуассона на интервале наблюдения Ω .

Выполняя функциональное дифференцирование выражения (10.20) с учетом формулы (9.18), получаем выражения для плотностей $a_i(t_1, \dots, t_i)$ потока Пуассона

$$a_i(t_1, \dots, t_i) = \exp\left\{-\int_{T_0}^{t_i} \lambda(t) dt\right\} \prod_{j=1}^i \lambda(t_j), \quad i = 1, 2, \dots \quad (10.23)$$

Найдем моментные функции первого и второго порядков для потока Пуассона, определяющие соответственно его интенсивность и корреляционную функцию. В соответствии с выражениями (9.19) и (10.20) получаем

$$\begin{aligned} f_1(t) &= \lambda(t), \\ f_2(t_1, t_2) &= \lambda(t_1)\lambda(t_2). \end{aligned} \quad (10.24)$$

Производящая функция числа событий на интервале наблюдения для потока Пуассона в соответствии с выражениями (9.23) и (10.20) имеет вид

$$\Theta(z; \Omega) = \exp \left\{ (z-1) \int_{\Omega} \lambda(t) dt \right\} = \exp \{ (z-1) \Lambda \}. \quad (10.25)$$

Тогда в соответствии с формулой (9.24) можно получить выражение для вероятностей выпадения определенного числа событий потока Пуассона на интервале наблюдения. Данное выражение, как и следовало ожидать, совпадает с дискретным распределением Пуассона, представленным ранее формулой (10.17).

Алгоритмы моделирования. Для моделирования потока Пуассона будем использовать его параметрическое задание на интервале наблюдения $\Omega = [T_0, T]$ при помощи системы совместных плотностей $a_i(t_1, \dots, t_i)$, $i = 1, 2, \dots$. Подставляя (10.23) в выражения (9.4) и (9.5), получим

$$q_i(t_i | t_1, \dots, t_{i-1}) = \lambda(t_i) \exp \left[- \int_{t_{i-1}}^{t_i} \lambda(t) dt \right], \quad t_0 = T_0, \quad (10.26)$$

$$d_i = \int_{t_{i-1}}^T \lambda(x) e^{-\int_{t_{i-1}}^x \lambda(t) dt} dx = - e^{-\int_{t_{i-1}}^x \lambda(t) dt} \Big|_{t_{i-1}}^T = 1 - \exp \left[- \int_{t_{i-1}}^T \lambda(t) dt \right].$$

Подставляя полученные выражения (10.26) в выражение для условной функции распределения (9.7), получаем для случая потока Пуассона

$$F_i(t_i | t_1, \dots, t_{i-1}) = \left\{ 1 - \exp \left[- \int_{t_{i-1}}^{t_i} \lambda(t) dt \right] \right\} / d_i. \quad (10.27)$$

Согласно общему подходу к моделированию случайных потоков с использованием плотностей $a_i(t_1, \dots, t_i)$, который был изложен выше, для моделирования потока Пуассона необходимо выполнить следующие шаги.

Шаг 1. Для моделирования i -го ($i = 1, 2, \dots$) события потока Пуассона необходимо получить независимую реализацию базовой случайной величины α'_i . Данная реализация используется для проверки того, появится ли очередное (i -е) событие потока на интервале наблюдения. Для этого необходимо сравнить α'_i и d_i . Если $\alpha'_i < d_i$, то это означает, что событие должно быть сгенерировано. В этом случае необходимо перейти к шагу 2. В противном случае моделирование реализации потока Пуассона завершается.

Шаг 2. Если на шаге 1 было установлено, что i -е событие потока имеет место на интервале наблюдения, то для определения его временного положения необходимо сгенерировать дополнительную независимую реализацию базовой случайной величины α_i'' . С ее использованием местоположение i -го события определяется согласно (9.6) и (10.27) из уравнения

$$\left\{ 1 - \exp \left[- \int_{t_{i-1}}^{t_i} \lambda(t) dt \right] \right\} / d_i = \alpha_i'' \Rightarrow \int_{t_{i-1}}^{t_i} \lambda(t) dt = -\ln(1 - d_i \alpha_i''). \quad (10.28)$$

Выражение (10.28) приобретает более простой вид в случае, если задача состоит в моделировании стационарного потока Пуассона с интенсивностью $\lambda(t) = \lambda = \text{const}$. В данном случае удобно перейти к расширенной области моделирования $\Omega' = [T_0, \infty]$. Тогда согласно выражению (10.26) все вероятности $d_i \equiv 1$. Учитывая это выражение (10.28) может быть переписано в следующем упрощенном виде:

$$t_i = t_{i-1} - \frac{\ln \alpha_i''}{\lambda}. \quad (10.29)$$

При этом процесс моделирования заканчивается при условии, что момент появления очередного события моделируемого потока Пуассона оказался за пределами интервала моделирования $\Omega = [T_0, T]$.

Использование приведенного выше метода моделирования потока Пуассона может быть затруднено в случае, если не удастся легко найти аналитическое решение уравнений (10.26) (для d_i) и (10.28) (вычисление интеграла). В таком случае представленный метод моделирования может быть несколько модифицирован и будет содержать следующие шаги.

Шаг 1. Моделируется вспомогательный поток Пуассона с интенсивностью $\beta(t) \geq \lambda(t)$ и такой, что моделирование вспомогательного потока является достаточно простым. Например, может использоваться $\beta(t) = \text{const} = \max_{\Omega}(\lambda(t))$. Обозначим моменты появления событий вспомогательного потока как $t_1, t_2, \dots$.

Шаг 2. Для каждого события вспомогательного потока $t_i, i = 1, 2, \dots$ моделируется случайная величина, равномерно распределенная на интервале $[0, \beta(t_i)]$. В случае если ее реализация $y_i \leq \lambda(t_i)$, то в момент времени t_i регистрируется событие исходного моделируемого потока с интенсивностью $\lambda(t)$. В противном случае t_i отбрасывается.

Фактически моделирование с использованием приведенного метода основано на прореживании вспомогательного потока Пуассона с большей интенсивностью и в некотором роде напоминает процедуру прореживания, использовавшуюся при моделировании непрерывных случайных величин методом Неймана. По аналогии с методом Неймана для приведенного метода моделирования потока Пуассона вводится понятие *эффективности* метода. Значение эффективности может быть вычислено как отношение среднего числа событий искомого потока к среднему числу событий вспомогательного потока

$$\kappa = \frac{\int_{\Omega} \lambda(t) dt}{\int_{\Omega} \beta(t) dt}.$$

Для повышения эффективности приведенного выше метода прореживания можно использовать следующую модификацию данного метода. После получения каждого события t_i искомого потока интенсивность вспомогательного потока меняется так, чтобы выполнялось более мягкое условие, чем в исходной формулировке метода: $\beta_i(t) \geq \lambda(t)$ только в пределах интервала $[t_i, T]$. Например, при генерации потока Пуассона с монотонно убывающей интенсивностью $\lambda(t)$ в качестве интенсивности вспомогательного потока для получения i -го события можно выбирать значения $\beta = \lambda(t_{i-1})$.

Пример 10.3. Задача состоит в моделировании потока Пуассона с интенсивностью:

$$\lambda(t) = a(1 + \sin t), \quad a > 0$$

на интервале $\Omega = [0, T]$.

Используем метод прореживания вспомогательного потока. В качестве интенсивности вспомогательного потока выберем $\beta(t) = \max_{\Omega}(\lambda(t)) = 2a$. Тогда с учетом (10.29) события вспомогательного потока могут быть получены по следующей формуле:

$$t_i = t_{i-1} - \frac{\ln \alpha'_i}{2a}.$$

После получения каждого t_i проверяем условие окончания моделирования $t_i > T$. Если это условие еще не выполняется, то проверяем, будет ли в точке t_i зарегистрировано событие исходного потока. Для этого вычисляем реализацию $y_i = 2a\alpha'_i$ случайной величины равномерно распределенной на интервале $[0, 2a]$ (α'_i —

независимая реализация базовой случайной величины). Если $y_i \leq a(1 + \sin t_i)$, то в момент времени t_i регистрируется событие исходного потока.

Величина эффективности метода для данного случая:

$$\kappa = \left(\int_0^T a(1 + \sin t) dt \right) / \int_0^T 2a dt = \frac{1}{2} + \frac{1 - \cos T}{2T}.$$

Кроме перечисленных выше методов существует также алгоритм моделирования потока Пуассона, основанный на его определении, как однородного потока Бернулли со случайным числом источников. Данный алгоритм при его использовании для моделирования потока Пуассона с интенсивностью $\lambda(t)$ на интервале Ω содержит следующие шаги.

Шаг 1. По формуле (10.22) находят параметр Λ соответствующего дискретного распределения Пуассона. Далее моделируют реализацию k дискретной случайной величины, вероятности которой заданы выражением (10.17), получая, таким образом, число одноточечных источников в схеме Бернулли.

Шаг 2. Зная число источников k , моделируют однородный поток Бернулли с k источниками и парциальными плотностями $s(t) = \lambda(t) / \Lambda$ при помощи наиболее подходящего метода. События такого потока Бернулли и будут событиями исходного потока Пуассона с интенсивностью $\lambda(t)$.

Л е к ц и я 11

МОДЕЛИРОВАНИЕ ТИПОВЫХ СЛУЧАЙНЫХ ПОТОКОВ (Часть 2)

Отрицательно-биномиальный поток. *Отрицательно-биномиальный поток* – это поток Пуассона, интенсивность которого является случайным процессом вида

$$\lambda(t) = \gamma \bar{\lambda}(t), \quad (11.1)$$

где $\bar{\lambda}(t)$ – детерминированная функция, а γ – неотрицательная случайная величина, распределенная по экспоненциальному закону. Функция распределения случайной величины γ имеет вид

$$F_\gamma(y) = 1 - e^{-y}, \quad y \geq 0. \quad (11.2)$$

Теоретическое описание. Найдем ПФЛ отрицательно-биномиального потока, используя представленную в его определении связь данного потока с потоком Пуассона. ПФЛ отрицательно-

биномиального потока может быть получен путем усреднения ПФЛ потока Пуассона по γ . Используя выражение (10.20) с учетом (11.1), получаем

$$L[u, \Omega] = \left\langle \exp \left\{ \gamma \int_{\Omega} \bar{\lambda}(t) u(t) dt \right\} \right\rangle_{\gamma} = \int_0^{\infty} e^{x \int_{\Omega} \bar{\lambda}(t) u(t) dt} e^{-x} dx = \int_0^{\infty} e^{-x \left\{ 1 - \int_{\Omega} \bar{\lambda}(t) u(t) dt \right\}} dx.$$

Вычисляя интеграл в последнем выражении, получаем окончательную формулу для ПФЛ отрицательно-биномиального потока

$$L[u; \Omega] = \frac{1}{1 - \int_{\Omega} \bar{\lambda}(t) u(t) dt}. \quad (11.3)$$

Выполняя функциональное дифференцирование выражения (11.3) с учетом формулы (9.18), получаем выражения для плотностей $a_i(t_1, \dots, t_i)$ отрицательно-биномиального потока

$$a_i(t_1, \dots, t_i) = \frac{i! \prod_{j=1}^i \bar{\lambda}(t_j)}{\left[1 + \int_{T_0}^{t_i} \bar{\lambda}(t) dt \right]^{i+1}}, \quad i = 1, 2, \dots \quad (11.4)$$

Найдем выражения для моментных функций отрицательно-биномиального потока. Для этого выполним функциональное дифференцирование ПФЛ (11.3) при $u(t) = 0$ в соответствии с (9.19). В результате получим

$$f_i(t_1, \dots, t_i) = i! \prod_{j=1}^i \bar{\lambda}(t_j), \quad i = 1, 2, \dots \quad (11.5)$$

Соответственно интенсивность и корреляционная функция отрицательно-биномиального потока будут выглядеть следующим образом:

$$\begin{aligned} f_1(t) &= \bar{\lambda}(t); \\ f_2(t_1, t_2) &= 2\bar{\lambda}(t_1)\bar{\lambda}(t_2). \end{aligned} \quad (11.6)$$

Производящая функция числа событий на интервале наблюдения для отрицательно-биномиального потока в соответствии с выражениями (9.23) и (11.3) имеет вид

$$\Theta(z; \Omega) = \frac{1}{1 - (z-1) \int_{\Omega} \bar{\lambda}(t) dt}. \quad (11.7)$$

Тогда, дифференцируя (11.7) при $z=0$ в соответствии с формулой (9.24), получаем выражения для вероятностей выпадения определенного числа событий отрицательно-биномиального потока на интервале наблюдения Ω

$$p_i(\Omega) = \frac{1}{i!} \frac{\left(\int_{\Omega} \bar{\lambda}(t) dt \right)^i}{\left(1 - (z-1) \int_{\Omega} \bar{\lambda}(t) dt \right)^{i+1}} \Bigg|_{z=0} = \frac{\left(\int_{\Omega} \bar{\lambda}(t) dt \right)^i}{\left(1 + \int_{\Omega} \bar{\lambda}(t) dt \right)^{i+1}}. \quad (11.8)$$

Далее перейдем к рассмотрению алгоритмов моделирования отрицательно-биномиального потока.

Алгоритмы моделирования. Поскольку определение отрицательно-биномиального потока основано на использовании понятия потока Пуассона, то наиболее очевидный алгоритм его моделирования может быть построен на использовании одного из представленных ранее алгоритмов моделирования потока Пуассона. Для того чтобы сделать возможным использование алгоритмов моделирования потока Пуассона в рассматриваемом случае необходимо определить интенсивность потока Пуассона, для чего в свою очередь необходимо получить реализацию y_i случайной величины γ . Тогда алгоритм моделирования отрицательно-биномиального потока будет выглядеть следующим образом.

Шаг 1. Для моделирования j -й реализации отрицательно-биномиального потока получаем реализацию y_j случайной величины γ . Для этого удобно воспользоваться методом обратных функций. С учетом выражения (11.2) получим

$$F_{\gamma}(y) = 1 - e^{-y_j} = \alpha_j \Rightarrow y_j = -\ln \alpha_j,$$

где α_j – независимая реализация базовой случайной величины.

Шаг 2. Получаем события $t_i, i=1,2,\dots$ потока Пуассона с интенсивностью $\lambda_j(t) = y_j \bar{\lambda}(t)$ наиболее удобным алгоритмом из тех, что были приведены ранее в соответствующем разделе. Полученные события будут представлять собой реализацию моделируемого отрицательно-биномиального потока.

Использование приведенного выше алгоритма позволяет получить реализации отрицательно-биномиального потока как совокупность отдельных реализаций потоков Пуассона с интенсивностями, отличающимися некоторым множителем, каждый раз определяемым соотношением

соседних реализаций случайной величины γ . В то же время рассматриваемый алгоритм требует дополнительных вычислительных затрат, связанных с генерированием случайной величины, распределенной по экспоненциальному закону.

Альтернативой приведенному выше методу может служить подход, основанный на непосредственном использовании плотностей $a_i(t_1, \dots, t_i)$, $i = 1, 2, \dots$, полученных для отрицательно-биномиального потока. Учитывая выражение (11.4), как и ранее для потока Пуассона, найдем выражения для условных плотностей $q_i(t_i | t_1, \dots, t_{i-1})$ и вероятностей d_i . Подставляя (11.4) в (9.4) и (9.5), получим

$$\begin{aligned}
 q_i(t_i | t_1, \dots, t_{i-1}) &= i\bar{\lambda}(t_i) \frac{\left(1 + \int_{T_0}^{t_{i-1}} \bar{\lambda}(t) dt\right)^i}{\left(1 + \int_{T_0}^{t_i} \bar{\lambda}(t) dt\right)^{i+1}}, \quad t_0 = T_0, \\
 d_i &= \int_{t_{i-1}}^T i\bar{\lambda}(x) \frac{\left(1 + \int_{T_0}^{t_{i-1}} \bar{\lambda}(t) dt\right)^i}{\left(1 + \int_{T_0}^x \bar{\lambda}(t) dt\right)^{i+1}} dx = - \frac{\left(1 + \int_{T_0}^{t_{i-1}} \bar{\lambda}(t) dt\right)^i}{\left(1 + \int_{T_0}^x \bar{\lambda}(t) dt\right)^i} \Bigg|_{t_{i-1}}^T = \\
 &= 1 - \frac{\left(1 + \int_{T_0}^{t_{i-1}} \bar{\lambda}(t) dt\right)^i}{\left(1 + \int_{T_0}^T \bar{\lambda}(t) dt\right)^i}. \tag{11.9}
 \end{aligned}$$

Подставляя (11.9) в выражение для условной функции распределения (9.7), получаем для случая отрицательно-биномиального потока

$$F_i(t_i | t_1, \dots, t_{i-1}) = \frac{1}{d_i} \left[1 - \frac{\left(1 + \int_{T_0}^{t_{i-1}} \bar{\lambda}(t) dt\right)^i}{\left(1 + \int_{T_0}^{t_i} \bar{\lambda}(t) dt\right)^i} \right]. \tag{11.10}$$

Согласно алгоритму моделирования случайных потоков с использованием плотностей $a_i(t_1, \dots, t_i)$ для моделирования отрицательно-биномиального потока необходимо выполнить следующие шаги.

Шаг 1. Для моделирования i -го ($i=1, 2, \dots$) события отрицательно-биномиального потока необходимо получить независимую реализацию базовой случайной величины α'_i . Данная реализация используется для проверки того, появится ли очередное (i -е) событие потока на интервале наблюдения. Для этого необходимо сравнить α'_i и d_i , определенное выражением (11.9). Если $\alpha'_i < d_i$, то это означает, что событие должно быть сгенерировано. В этом случае необходимо перейти к шагу 2. В противном случае моделирование реализации отрицательно-биномиального потока завершается.

Шаг 2. Если на шаге 1 было установлено, что i -е событие потока имеет место на интервале наблюдения, то для определения его временного положения необходимо сгенерировать дополнительную независимую реализацию базовой случайной величины α''_i . С ее использованием местоположение i -го события на оси времени определяется согласно (9.6) и (11.10) из уравнения

$$\begin{aligned} \frac{1}{d_i} \left[1 - \frac{1 + \int_{T_0}^{t_{i-1}} \bar{\lambda}(t) dt}{1 + \int_{T_0}^{t_i} \bar{\lambda}(t) dt} \right]^i &= \alpha''_i \Rightarrow \frac{1 + \int_{T_0}^{t_{i-1}} \bar{\lambda}(t) dt}{1 + \int_{T_0}^{t_i} \bar{\lambda}(t) dt} = 1 - d_i \alpha''_i \Rightarrow \\ \Rightarrow \int_{T_0}^{t_i} \bar{\lambda}(t) dt &= \frac{1 + \int_{T_0}^{t_{i-1}} \bar{\lambda}(t) dt}{(1 - d_i \alpha''_i)^{1/i}} - 1. \end{aligned} \quad (11.11)$$

Как и для случая потока Пуассона для отрицательно-биномиального потока выражение для получения момента возникновения очередного события потока приобретает более простой вид при рассмотрении стационарного случая, когда $\bar{\lambda}(t) = \bar{\lambda} = \text{const}$. При переходе к расширенной области моделирования $\Omega' = [T_0, \infty]$ для стационарного отрицательно-биномиального потока выражение (11.11) приобретает вид

$$t_i = T_0 + \frac{\frac{1}{\bar{\lambda}} + t_{i-1} - T_0}{\alpha''_i^{1/i}} - \frac{1}{\bar{\lambda}}.$$

При этом, как следует из выражения (11.9) вероятности $d_i \equiv 1$, что также упрощает алгоритм моделирования потока.

Поток Кокса. *Потоком Кокса* называется поток Пуассона, у которого интенсивность представляет собой некоторый случайный процесс $\xi(t)$.

Теоретическое описание. Как и для рассмотренного ранее отрицательно-биномиального потока ПФЛ потока Кокса может быть получен путем усреднения ПФЛ потока Пуассона по реализациям случайного процесса $\xi(t)$:

$$L[u, \Omega] = \left\langle \exp \left\{ \int_{\Omega} \xi(t) u(t) dt \right\} \right\rangle_{\xi(t)}. \quad (11.12)$$

Найдем выражения для моментных функций потока Кокса. Для этого выполним функциональное дифференцирование ПФЛ (11.12) при $u(t) = 0$ в соответствии с (9.19). В результате получим

$$f_i(t_1, \dots, t_i) = \langle \xi(t_1) \xi(t_2) \dots \xi(t_i) \rangle_{\xi(t)}, i = 1, 2, \dots \quad (11.13)$$

Таким образом, как видно из выражения (11.13), моментные функции потока Кокса совпадают с моментными функциями случайного процесса $\xi(t)$.

Производящая функция числа событий на интервале наблюдения для потока Кокса в соответствии с выражениями (9.23) и (11.12) имеет вид

$$\Theta(z; \Omega) = \left\langle \exp \left\{ (z - 1) \int_{\Omega} \xi(t) dt \right\} \right\rangle_{\xi(t)}. \quad (11.14)$$

Тогда, дифференцируя (11.14) при $z = 0$, в соответствии с формулой (9.24), получаем выражения для вероятностей выпадения определенного числа событий потока Кокса на интервале наблюдения Ω :

$$p_i(\Omega) = \int_0^{\infty} \frac{W^i e^{-W}}{i!} \rho(W) dW, \quad (11.15)$$

где $W = \int_{\Omega} \xi(t) dt$, а $\rho(W)$ – плотность распределения величины W . В ли-

тературе формула (11.15) часто называется формулой Манделя.

В ряде случаев при известной плотности распределения $\rho(W)$ выражение (11.15) может принимать более простой вид. Например, рассмотрим поток Кокса, имеющий в качестве интенсивности случайный

процесс с экспоненциальной плотностью распределения $\rho_\xi(x) = \frac{1}{m_\xi} e^{-\frac{x}{m_\xi}}$,

где m_ξ – математическое ожидание случайной величины с плотностью распределения $\rho_\xi(x)$. Для достаточно малых интервалов Ω можно считать, что $W = A\xi(t)$, где A – длина интервала Ω . Тогда плотность распределения $\rho(W)$ случайной величины W можно рассчитать по следующей формуле:

$$\rho(W) = \rho_\xi(\xi) \left| \frac{d\xi}{dW} \right| = \rho_\xi\left(\frac{W}{A}\right) \frac{1}{A} = \frac{1}{m_\xi A} e^{-\frac{W}{m_\xi A}} = \frac{1}{m} e^{-\frac{W}{m}}.$$

Подставляя полученное выражение в (11.15), получим, что в данном случае распределение числа событий i потока Кокса на интервале Ω будет распределением Бозе – Эйнштейна:

$$\begin{aligned} p_i(\Omega) &= \frac{1}{i!m} \int_0^\infty W^i e^{-W} e^{-\frac{W}{m}} dW = \frac{1}{i!m} \int_0^\infty W^i e^{-W\left(\frac{m+1}{m}\right)} dW = \\ &= \frac{1}{i!m} \frac{m}{m+1} \left[-W^i e^{-W\left(\frac{m+1}{m}\right)} \right]_0^\infty + i \int_0^\infty W^{i-1} e^{-W\left(\frac{m+1}{m}\right)} dW = \\ &= \frac{1}{i!m} \frac{m}{m+1} i \int_0^\infty W^{i-1} e^{-W\left(\frac{m+1}{m}\right)} dW = \dots \\ &= \frac{1}{i!m} \left(\frac{m}{m+1} \right)^i i! \int_0^\infty e^{-W\left(\frac{m+1}{m}\right)} dW = \frac{m^i}{(m+1)^{i+1}}. \end{aligned} \quad (11.16)$$

Алгоритм моделирования. Алгоритм моделирования потока Кокса основан на его определении, как специальной разновидности потока Пуассона. Для моделирования данного потока можно воспользоваться алгоритмом прореживания вспомогательного потока Пуассона более высокой интенсивности. Данный алгоритм включает следующие шаги.

Шаг 1. Получаем моменты времени $\tau_1, \dots, \tau_k \in \Omega$, представляющие собой реализацию вспомогательного потока Пуассона с детерминированной интенсивностью $\lambda(t) \gg \xi(t)$, где k – число событий в конкретной реализации вспомогательного потока. Вид функции $\lambda(t)$ выбирается из следующего условия:

$$P\{\xi(t) > \lambda(t)\} = \varepsilon \ll 1, \quad (11.17)$$

т. е. вероятность превышения случайным процессом значения интенсивности $\lambda(t)$ должна быть незначительной.

Шаг 2. Генерируем значения $y(\tau_1), \dots, y(\tau_k)$ случайного процесса $\xi(t)$ в полученные на шаге 1 моменты времени $\tau_1, \dots, \tau_k \in \Omega$.

Шаг 3. Для каждой точки $\tau_i, i=1, \dots, k$ генерируем реализацию φ_i случайной величины, равномерно распределенной на интервале $[0, \lambda(\tau_i)]$. Если выполняется неравенство $\varphi_i < y(\tau_i)$, то точка τ_i включается в реализацию моделируемого потока Кокса. В противном случае τ_i отбрасывается. Оставшиеся после прореживания точки и будут представлять собой реализацию потока Кокса.

Как видно из представленного выше алгоритма моделирования потока Кокса, для получения реализации потока требуется моделировать случайный процесс $\xi(t)$. В общем случае, если $\xi(t)$ не является гауссовским случайным процессом, его моделирование может оказаться достаточно сложной задачей. Однако в некоторых случаях удастся представить реализации негауссовского случайного процесса $\xi(t)$ в виде некоторой комбинации реализаций гауссовских случайных процессов, для которых существуют достаточно эффективные алгоритмы моделирования. Так, например, приведенный выше экспоненциальный случайный процесс может быть представлен в виде (8.17). При этом его корреляционная функция задается соотношением (8.18). Алгоритм моделирования экспоненциального случайного процесса был приведен ранее в разделе, посвященном моделированию случайных процессов.

Парнокоррелированный поток. Одним из наиболее известных частных случаев потока Кокса является парнокоррелированный поток. *Парнокоррелированный поток* – это поток Кокса, интенсивность которого $\xi(t)$ представляет собой нормальный случайный процесс. Отметим, что парнокоррелированный поток полностью определяется своими интенсивностью $f_1(t)$ и корреляционной функцией $f_2(t_1, t_2)$.

Теоретическое описание. Пусть $\xi(t)$ – нормальный случайный процесс с математическим ожиданием $g_1(t)$ и ковариационной функцией $g_2(t_1, t_2)$. Тогда с учетом выражения (11.12) для ПФЛ потока Кокса, ПФЛ парнокоррелированного потока будет иметь вид:

$$L[u, \Omega] = \exp \left\{ \int_{\Omega} g_1(t) u(t) dt + \frac{1}{2} \int_{\Omega^2} g_2(t_1, t_2) u(t_1) u(t_2) dt_1 dt_2 \right\}, \quad (11.18)$$

где $g_1(t) - \int_{\Omega} g_2(t, \tau) d\tau > 0$.

Найдем выражения для моментных функций первого и второго порядка парнокоррелированного потока. Для этого выполним функциональное дифференцирование ПФЛ (11.18) при $u(t) = 0$ в соответствии с (9.19). В результате получим

$$\begin{aligned} f_1(t) &= g_1(t), \\ f_2(t_1, t_2) &= g_2(t_1, t_2) + g_1(t_1)g_1(t_2). \end{aligned} \quad (11.19)$$

Алгоритмы моделирования. Поскольку исходя из определения парнокоррелированный поток является частным случаем потока Кокса, то для его моделирования можно использовать общий метод моделирования потока Кокса, основанный на операции прореживания. При использовании метода прореживания необходимо учитывать тот факт, что реализации нормального процесса $\xi(t)$ могут принимать отрицательные значения на заданном интервале моделирования, что недопустимо для интенсивности случайного потока. По этой причине необходимо модифицировать алгоритм моделирования $\xi(t)$ таким образом, чтобы исключить использование отрицательных значений его реализаций. Например, можно заменять получаемые отрицательные значения нулевыми. Использование таких замен в общем случае приводит к искажению функций $f_1(t)$ и $f_2(t_1, t_2)$, характеризующих парнокоррелированный поток. Однако при определенных условиях использование алгоритма моделирования с приведенными выше модификациями все же возможно. Так, при моделировании методом прореживания парнокоррелированных потоков, у которых среднеквадратическое отклонение случайного процесса $\xi(t)$ значительно меньше его математического ожидания, вносимые искажения практически не сказываются на получаемых характеристиках потока.

В связи с недостатками применения метода прореживания, приведенными выше, был разработан альтернативный метод моделирования парнокоррелированного потока. Данный метод основан на представлении парнокоррелированного потока в виде совокупности событий связанных между собой вспомогательных потоков. Первый из указанных выше потоков называют базовым (или потоком базовых точек), а остальные (вторичные) потоки образуются путем сдвига событий базового потока на случайную величину. Распределение случайной величины, определяющей сдвиг событий вторичных потоков относительно базового, выбирается таким образом, чтобы события вторичных потоков наступали позже во времени относительно событий базового потока. Таким обра-

зом, алгоритм моделирования строится в соответствии с механизмом образования парнокоррелированного потока.

Пусть на интервале наблюдения $\Omega = [T_0, T]$ необходимо смоделировать последовательность событий парнокоррелированного потока с интенсивностью $f_1(t)$ и ковариационной функцией вида $g_2(t_1, t_2) = f_2(t_1, t_2) - f_1(t_1)f_1(t_2)$. Алгоритм моделирования включает следующие шаги.

Шаг 1. Выберем в качестве базового потока поток Пуассона с интенсивностью $f_1^A(t)$, которая определяется следующим выражением:

$$f_1^A(t) = f_1(t) - \int_{T_0}^t g_2(x, t) dx \quad (11.20)$$

и получим его реализацию.

Шаг 2. Для каждого события базового потока генерируем смещенный поток, состоящий из не более чем одного события. Условная плотность вероятности появления смещенного события вторичного потока в малом интервале $(t - dt, t]$ при условии, что соответствующее событие базового потока появилось в момент времени τ , имеет следующий вид:

$$\psi(t | \tau) = \begin{cases} 0, & t \leq \tau \\ g_2(\tau, t) / f_1^A(\tau), & t > \tau \end{cases} \quad (11.21)$$

Процесс моделирования смещенного события можно разделить на два этапа. Вначале определим, будет ли иметь место событие вторичного потока. Для этого вычислим вероятность $p(\tau) = \int_{\tau}^T \psi(x | \tau) dx$ появления события вторичного потока на интервале $(\tau, T]$. Затем получаем реализацию базовой случайной величины α' и проверяем, выполняется ли условие $\alpha' \leq p(\tau)$. Если нет, то точка вторичного потока не порождается. В противном случае событие вторичного потока имеет место и далее определяется его местоположение. С учетом того что плотность $\psi(t | \tau)$ не нормирована на интервале $(\tau, T]$, уравнение для определения местоположения события вторичного потока имеет вид:

$$\frac{\int_{\tau}^t \psi(x | \tau) dx}{p(\tau)} = \alpha'', \quad (11.22)$$

где α'' – независимая реализация базовой случайной величины.

Шаг 3. Выполняем наложение реализаций базового и всех вторичных потоков, после чего, выполнив сортировку по возрастанию, получаем реализацию парнокоррелированного потока.

Пример 11.1. На интервале наблюдения $\Omega = [T_0, T]$ смоделировать парнокоррелированный поток со следующими характеристиками:

$$\begin{aligned} f_1(t) &= f = \text{const} \\ g_2(t_1, t_2) &= ae^{-a(t_2-t_1)}. \end{aligned}$$

Воспользуемся методом, основанным на моделировании базового и смещенных потоков. Интенсивность базового потока Пуассона

$$f_1^A(t) = f - a \int_{T_0}^t e^{-a(t-x)} dx = f - 1 + e^{-a(t-T_0)}.$$

Условная плотность вероятности появления события вторичного потока в соответствии с выражением (11.21)

$$\psi(t | \tau) = \frac{ae^{-a(t-\tau)}}{f - 1 + e^{-a(\tau-T_0)}}.$$

Тогда вероятность появления события вторичного потока будет иметь вид

$$p(\tau) = \int_{\tau}^T \frac{ae^{-a(x-\tau)}}{f - 1 + e^{-a(\tau-T_0)}} dx = - \frac{e^{-a(x-\tau)}}{f - 1 + e^{-a(\tau-T_0)}} \Big|_{\tau}^T = \frac{1 - e^{-a(T-\tau)}}{f - 1 + e^{-a(\tau-T_0)}}.$$

С целью нахождения выражения для определения местоположения событий вторичных потоков воспользуемся выражением (11.22)

$$\frac{\int_{\tau}^t ae^{-a(x-\tau)} dx}{1 - e^{-a(T-\tau)}} = \frac{1 - e^{-a(t-\tau)}}{1 - e^{-a(T-\tau)}} = \alpha'' \Rightarrow t = \tau - \frac{1}{a} \ln[1 - \alpha''(1 - e^{-a(T-\tau)})].$$

Л е к ц и я 12

ПОТОКИ ГРУППИРОВАННЫХ ТОЧЕК. СТАТИСТИЧЕСКИЙ АНАЛИЗ ГЕНЕТИЧЕСКИ СВЯЗАННЫХ ПОТОКОВ

Часто потоки событий, с которыми приходится сталкиваться при проведении физических экспериментов, можно рассматривать как последовательности групп событий определенной природы. При этом интерпретация понятия группы может, вообще говоря, быть различной и зависит от конкретной рассматриваемой ситуации. Так, для одиночных потоков в качестве группы может рассматриваться отдельная реализация потока. В то же время при изучении генетически связанных потоков, т. е. таких потоков, для которых события одного из них (первичного) могут порождать последовательности событий других (вторичных), в качестве

групп рассматриваются реализации вторичных потоков. Однако вне зависимости от интерпретации понятия группы для всех потоков, которые могут рассматриваться как потоки группированных точек, может использоваться единый подход к изучению их статистических характеристик.

Для описания потоков группированных точек введем понятия *внутригрупповых и общегрупповых параметров*. *Внутригрупповые параметры* представляют собой моменты наступления событий, принадлежащих отдельно выделенной группе. В дальнейшем при рассмотрении генетически связанных потоков, в качестве таких отдельных групп будем рассматривать отдельные вторичные потоки. В свою очередь *общегрупповые параметры* характеризуют всю группу в целом (например, определяют начало группы или число событий в группе).

Вначале предположим, что рассматриваемый случайный поток имеет один случайный общегрупповой параметр γ , который задан на некоторой области G при помощи функции плотности вероятности $g(x)$, причем $\int_G g(x)dx = 1$. В данном случае внутригрупповые параметры пред-

ставляют собой моменты возникновения случайных событий на некоторой области наблюдения $\Omega = [T_0, T]$. Для описания такой одиночной группы событий можно использовать совместные плотности распределения $\pi_{1i}(\tau; G; t_1, \dots, t_i; \Omega)$, $i = 0, 1, \dots$, которые определяют совместную вероятность принятия общегрупповым параметром γ значения из малой окрестности точки τ и появления на Ω ровно i событий, наступающих в окрестностях точек $t_1, \dots, t_i$:

$$\begin{aligned} & \pi_{1i}(\tau; G; t_1, \dots, t_i; \Omega) d\tau dt_1 \dots dt_i = \\ & = P\{\gamma \in [\tau, \tau + d\tau); \xi_1 \in [t_1, t_1 + dt_1), \dots, \xi_i \in [t_i, t_i + dt_i); v = i\}. \end{aligned}$$

В случае если значение общегруппового параметра зафиксировано, то рассматриваемый поток превращается в обычный поток событий, который задается условными плотностями

$$\pi_{i|1}(t_1, \dots, t_i; \Omega | \tau; G) = \frac{\pi_{1i}(\tau; G; t_1, \dots, t_i; \Omega)}{g(\tau)}, i = 0, 1, \dots \quad (12.1)$$

Плотности (12.1) характеризуют совместное распределение моментов наступления событий рассматриваемого потока при условии, что $\gamma = \tau$. Данные плотности подчиняются обычной для случайных потоков нормировке:

$$\sum_{i=0}^{\infty} \frac{1}{i!} \int_{\Omega^i} \pi_{i|1}(t_1, \dots, t_i; \Omega | \tau; G) dt_1 \dots dt_i = 1$$

и позволяют ввести для описания рассматриваемого потока условный ПФЛ, который имеет следующий вид:

$$L[u; \Omega | \tau; G] = \sum_{i=0}^{\infty} \frac{1}{i!} \int_{\Omega^i} \pi_{i|1}(t_1, \dots, t_i; \Omega | \tau; G) \prod_{j=1}^i [1 + u(t_j)] dt_1 \dots dt_i. \quad (12.2)$$

Безусловные характеристики рассматриваемого потока могут быть найдены путем усреднения условного ПФЛ (12.2) по общегрупповому параметру γ :

$$\begin{aligned} L[u; \Omega] &= \int_G L[u; \Omega | \tau; G] g(\tau) d\tau = \\ &= \sum_{i=0}^{\infty} \frac{1}{i!} \int_{\Omega^i} \int_G \pi_{i|1}(t_1, \dots, t_i; \Omega | \tau; G) g(\tau) d\tau \prod_{j=1}^i [1 + u(t_j)] dt_1 \dots dt_i. \end{aligned} \quad (12.3)$$

Учитывая, что по определению безусловный ПФЛ рассматриваемого потока имеет вид

$$L[u; \Omega] = \sum_{i=0}^{\infty} \frac{1}{i!} \int_{\Omega^i} \pi_i(t_1, \dots, t_i; \Omega) \prod_{j=1}^i [1 + u(t_j)] dt_1 \dots dt_i$$

и сравнивая последнее выражение с (12.3), получаем выражения для безусловных совместных плотностей рассматриваемого потока

$$\begin{aligned} \pi_i(t_1, \dots, t_i; \Omega) &= \int_G \pi_{i|1}(t_1, \dots, t_i; \Omega | \tau; G) g(\tau) d\tau, \\ f_i(t_1, \dots, t_i) &= \int_G f_{i|1}(t_1, \dots, t_i | \tau; G) g(\tau) d\tau. \end{aligned}$$

Примером потока с одним общегрупповым параметром может служить отрицательно-биномиальный поток, у которого общегрупповым параметром является случайная величина γ , входящая в выражение для интенсивности такого потока, а в качестве группы рассматривается поток Пуассона, интенсивность которого равна $\tau \bar{\lambda}(t)$, где τ является реализацией случайной величины γ .

При рассмотрении потока группированных точек число групп в потоке может быть больше 1. Предположим, что в потоке можно выделить k групп с общегрупповыми параметрами $\gamma_1, \dots, \gamma_k$ и соответствующими внутригрупповыми параметрами $t_1^1, \dots, t_{i_1}^1, \dots, t_1^k, \dots, t_{i_k}^k$. В общем случае внутригрупповые параметры разных групп и общегрупповые параметры могут иметь различную природу. Описание такого потока может быть

выполнено при помощи системы совместных плотностей обще- и внутригрупповых параметров:

$$\pi_{ki_1 \dots i_k}(\tau_1, \dots, \tau_k; G; t_1^1, \dots, t_{i_1}^1, \dots, t_1^k, \dots, t_{i_k}^k; \Omega), i_j = 0, 1, \dots; j = 1, \dots, k; \tau \in G; t \in \Omega.$$

В дальнейшем для рассмотрения генетически связанных потоков будем полагать, что точки различных групп неразличимы, число групп является случайной величиной, а общегрупповые параметры образуют случайный поток, заданный на области G , которая в общем случае не совпадает с Ω . Поток общегрупповых параметров будем называть *первичным потоком*. Для задания первичного потока будем использовать системы функций $\pi_m(\tau_1, \dots, \tau_m; G)$ и симметрических моментных функций $f_m(\tau_1, \dots, \tau_m)$, на основе которых может быть построен ПФЛ первичного потока

$$L[v; G] = \sum_{m=0}^{\infty} \frac{1}{m!} \int_{G^m} \pi_m(\tau_1, \dots, \tau_m; G) \prod_{j=1}^m [1 + v(\tau_j)] d\tau_1 \dots d\tau_m,$$

или

$$L[v; G] = 1 + \sum_{m=1}^{\infty} \frac{1}{m!} \int_{G^m} f_m(\tau_1, \dots, \tau_m) \prod_{j=1}^m v(\tau_j) d\tau_1 \dots d\tau_m,$$

где $v(\tau)$ – пробная функция ПФЛ первичного потока.

Поскольку в рассматриваемом случае точки, принадлежащие различным группам, являются неразличимыми, то можно считать, что все внутригрупповые параметры образуют единый поток без разбиения на группы. Такой единый поток будем называть в дальнейшем *суммарным*.

Общегрупповые параметры $\gamma_1, \dots, \gamma_\mu$ и точки суммарного потока $\xi_1, \dots, \xi_\nu$ могут быть описаны системой совместных плотностей $\pi_{mk}(\tau_1, \dots, \tau_m; G; t_1, \dots, t_k; \Omega)$, $m = 0, 1, \dots$, $k = 0, 1, \dots$, такой, что

$$\begin{aligned} & \pi_{mk}(\tau_1, \dots, \tau_m; G; t_1, \dots, t_k; \Omega) d\tau_1 \dots d\tau_m dt_1 \dots dt_k = \\ & P\{\gamma_1 \in [\tau_1, \tau_1 + d\tau_1), \dots, \gamma_m \in [\tau_m, \tau_m + d\tau_m); \mu = m; \\ & \xi_1 \in [t_1, t_1 + dt_1), \dots, \xi_k \in [t_k, t_k + dt_k); \nu = k\}, \end{aligned}$$

где μ – число событий потока общегрупповых параметров, а ν – число событий суммарного потока.

Совместный ПФЛ потока группированных точек определяется следующим образом:

$$L[v; G; u; \Omega] = \left\langle \prod_{i=1}^{\mu} [1 + v(\tau_i)] \prod_{j=1}^{\nu} [1 + u(t_j)] \right\rangle_{\gamma, \mu, \xi, \nu}. \quad (12.4)$$

Для изучения свойств суммарного потока при фиксированных значениях общегрупповых параметров используются условные плотности:

$$\pi_{k|m}(t_1, \dots, t_k; \Omega | \tau_1, \dots, \tau_m; G) = \frac{\pi_{mk}(\tau_1, \dots, \tau_m; G; t_1, \dots, t_k; \Omega)}{\pi_m(\tau_1, \dots, \tau_m; G)} \quad (12.5)$$

и соответствующий условный ПФЛ:

$$\begin{aligned} L[u; \Omega | \tau_1, \dots, \tau_m; G] &= \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} \int_{\Omega^k} \pi_{k|m}(t_1, \dots, t_k; \Omega | \tau_1, \dots, \tau_m; G) \prod_{j=1}^k [1 + u(t_j)] dt_1 \dots dt_k. \end{aligned} \quad (12.6)$$

С учетом (12.5) найдем связь между совместным ПФЛ потока группированных точек (12.4) и условным ПФЛ суммарного потока (12.6):

$$\begin{aligned} L[v; G; u; \Omega] &= \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \frac{1}{m!} \frac{1}{k!} \int_{G^m} \int_{\Omega^k} \pi_{mk}(\tau_1, \dots, \tau_m; G; t_1, \dots, t_k; \Omega) \times \\ &\quad \times \prod_{i=1}^m [1 + v(\tau_i)] \prod_{j=1}^k [1 + u(t_j)] dt_1 \dots dt_k d\tau_1 \dots d\tau_m = \\ &= \sum_{m=0}^{\infty} \frac{1}{m!} \int_{G^m} \pi_m(\tau_1, \dots, \tau_m; G) \sum_{k=0}^{\infty} \frac{1}{k!} \int_{\Omega^k} \pi_{k|m}(t_1, \dots, t_k; \Omega | \tau_1, \dots, \tau_m; G) \times \\ &\quad \times \prod_{j=1}^k [1 + u(t_j)] dt_1 \dots dt_k \prod_{i=1}^m [1 + v(\tau_i)] d\tau_1 \dots d\tau_m = \\ &= \sum_{m=0}^{\infty} \frac{1}{m!} \int_{G^m} \pi_m(\tau_1, \dots, \tau_m; G) \prod_{i=1}^m [1 + v(\tau_i)] L[u; \Omega | \tau_1, \dots, \tau_m; G] d\tau_1 \dots d\tau_m. \end{aligned} \quad (12.7)$$

Если связь между группами выражается только через соответствующий общегрупповой параметр, то условный ПФЛ (12.6) может быть выражен через условные ПФЛ одиночных групп $L^B[u; \Omega | \tau_i; G]$ вида (12.2) следующим образом:

$$L[u; \Omega | \tau_1, \dots, \tau_m; G] = \prod_{i=1}^m L^B[u; \Omega | \tau_i; G], \quad (12.8)$$

где $L^B[u; \Omega | \tau_i; G]$ – условный ПФЛ одиночной группы с общегрупповым параметром $\gamma_i = \tau_i$. Подставляя (12.8) в (12.7), получаем

$$L[v; G; u; \Omega] = L^A[(1 + v(\cdot))L^B[u; \Omega | \cdot; G] - 1; G].$$

Полагая в последнем выражении пробную функцию $v = 0$, получаем формулу, связывающую ПФЛ суммарного потока, условные ПФЛ одиночных групп и ПФЛ потока общегрупповых параметров:

$$L^C[u; \Omega] = L^A[L^B[u; \Omega | \cdot; G] - 1; G]. \quad (12.9)$$

Используя выражение (12.9), можно получить связь между плотностями суммарного потока, условными плотностями одиночных групп и безусловными плотностями первичного потока. Выполняя функциональное дифференцирование (12.9) при $u(t) = 0$, можно получить аналогичную связь между моментными функциями. Выражение для первой моментной функции (интенсивности) суммарного потока будет иметь вид

$$f_1^C(t) = \int_G f_1^A(\tau) f_{1|1}^B(t | \tau; G) d\tau, \quad (12.10)$$

где обозначения A, B, C соответствуют первичному, индивидуальному вторичному и суммарному потокам.

Выражение (12.10) может быть использовано для восстановления интенсивности генетически связанных потоков при их статистическом анализе. При обработке экспериментальных данных, которые представляют собой реализации генетически связанных потоков случайных событий, несущих информацию об исследуемых явлениях, часто используется модель, согласно которой каждое событие первичного потока, воздействуя на исследуемый объект, порождает вторичный случайный поток. Вторичные потоки образуют суммарный поток, который и регистрируется при проведении эксперимента. Целью проводимого анализа является определение тех или иных характеристик (например, интенсивности) отдельных вторичных потоков. В дальнейшем задача анализа генетически связанных потоков будет рассмотрена на примере восстановления интенсивности отдельного вторичного потока.

Пусть на интервале $G = [T_1, T_2]$ задан первичный поток A . Каждое его событие независимо от других событий первичного потока порождает индивидуальный вторичный поток B_τ , где τ — момент наступления порождающего события. Будем полагать, что потоки B_τ идентичны и отличаются друг от друга лишь моментами возникновения. Тогда, исходя из сделанных предположений, условные моментные функции индивидуальных вторичных потоков будут иметь вид

$$f_{i|1}(t_1, \dots, t_i | \tau; G) = \begin{cases} 0, & \min_{1 \leq j \leq i} t_j < \tau \\ f_i(t_1 - \tau, \dots, t_i - \tau) & \min_{1 \leq j \leq i} t_j \geq \tau \end{cases}. \quad (12.11)$$

Суперпозиция вторичных потоков образует суммарный поток C , исследуемый на заданном интервале Ω . Статистический анализ генетически связанных потоков предполагает, как правило, нахождение харак-

теристик потока B_τ по известным характеристикам потока A и полученным из эксперимента характеристикам потока C . При этом искомые характеристики индивидуального вторичного потока могут быть определены по следующей схеме:

- регистрируем требуемые характеристики потока C на интервале Ω ;
- находим из ПФЛ потока C аналитические выражения характеристик, полученных из эксперимента;
- полученные аналитические выражения в дальнейшем используются в качестве уравнений для нахождения неизвестных характеристик вторичного потока аналитическими либо численными методами.

В процессе решения поставленной выше задачи существенное значение играет способ выбора интервала наблюдения Ω , от которого в конечном счете зависят получаемые уравнения для нахождения характеристик индивидуальных вторичных потоков.

Пусть интервал наблюдения Ω , на котором регистрируется суммарный поток C , выбирается вне зависимости от того, в какие моменты времени наступали события первичного потока A . Тогда, учитывая выражения (12.10) и (12.11), уравнение, определяющее связь интенсивности суммарного потока C с интенсивностями первичного и индивидуального вторичного потоков запишется в виде

$$f_1^C(t) = \int_{T_1}^{T_2} f_1^A(\tau) f_1^B(t - \tau) d\tau, \quad (12.12)$$

где $f_1^B(x) \equiv 0$, если $x < 0$; $G = [T_1, T_2]$, $t \in \Omega = [T_0, T]$. Решение интегрального уравнения (12.12) позволяет получить интенсивность индивидуального вторичного потока.

При решении интегрального уравнения (12.12) существенным является учет расположения области наблюдения Ω . Пусть $\Omega = G$, $T_1 \neq -\infty$. Тогда, положив начало отсчета в точке T_1 из (12.12), получим

$$f_1^C(t) = \int_0^t f_1^A(t - \tau) f_1^B(\tau) d\tau. \quad (12.13)$$

Интегральное уравнение (12.13) представляет собой уравнение типа свертки. Для решения такого уравнения может использоваться преобразование Лапласа. Преобразование Лапласа функции $f(t)$ дает ее изображение $\Phi(p)$, которое определяется следующей формулой:

$$\Phi(p) = \int_0^{\infty} f(t) e^{-pt} dt.$$

Применив преобразование Лапласа к обеим частям выражения (12.13) и учитывая, что изображение по Лапласу свертки двух функций представляет собой произведение их изображений, получим

$$\Phi_1^C(p) = \Phi_1^A(p) \Phi_1^B(p), \quad (12.14)$$

где $\Phi_1^A(p)$, $\Phi_1^B(p)$ и $\Phi_1^C(p)$ – изображения по Лапласу соответственно функций $f_1^A(t)$, $f_1^B(t)$ и $f_1^C(t)$. Из уравнения (12.14) легко находится изображение искомой интенсивности индивидуального вторичного потока

$$\Phi_1^B(p) = \frac{\Phi_1^C(p)}{\Phi_1^A(p)}. \quad (12.15)$$

Вычислив изображение $\Phi_1^B(p)$ и применив обратное преобразование Лапласа, можно получить выражение для искомой интенсивности индивидуального вторичного потока $f_1^B(t)$.

Пример 12.1. Пусть интенсивность первичного потока имеет вид

$$f_1^A(t) = a e^{bt}, \quad a > 0.$$

Тогда ее изображение по Лапласу в соответствии с таблицами данного преобразования имеет вид

$$\Phi_1^A(p) = \frac{a}{p+b}.$$

Подставляя последнее выражение в (12.15), получим изображение интенсивности индивидуального вторичного потока:

$$\Phi_1^B(p) = \frac{1}{a} [p \Phi_1^C(p) + b \Phi_1^C(p)].$$

В соответствии с таблицами преобразования Лапласа получим выражение для интенсивности индивидуального вторичного потока. Учитывая, что умножение изображения на p эквивалентно дифференцированию соответствующей ему функции во временной области, находим

$$f_1^B(t) = \frac{1}{a} \left(\frac{d}{dt} f_1^C(t) + b f_1^C(t) \right).$$

В случае если первичный поток является стационарным, т. е. $f_1^A(t) = a$, то последнее выражение упрощается. Для случая стационарного потока необходимо положить $b = 0$. Тогда получим

$$f_1^B(t) = \frac{1}{a} \frac{d}{dt} f_1^C(t).$$

Стоит отметить, что решение интегральных уравнений типа свертки относится к разряду плохо обусловленных задач. Поэтому восстановление искомой интенсивности $f_1^B(t)$ при помощи преобразования Лапласа требует определения интенсивности суммарного потока $f_1^C(t)$ с высокой точностью. По вышеуказанным причинам в ряде случаев имеет смысл применять альтернативные подходы к определению $f_1^B(t)$, для которых базовые уравнения отличаются от уравнения (12.13).

В качестве примера рассмотрим ситуацию, когда первичный поток является регулярным, т. е. его интенсивность задается выражением

$$f_1^A(\tau) = \sum_{i=0}^m \delta(\tau - T_1 - i\Delta t), \quad (12.16)$$

где $m = [(T_2 - T_1)/\Delta t]$, $\Delta t > 0$, $[x]$ – целая часть действительного числа x . Подставляя (12.16) в (12.12), получим

$$f_1^C(t) = \int_{T_1}^{T_2} \sum_{i=0}^m \delta(\tau - T_1 - i\Delta t) f_1^B(t - \tau) d\tau = \sum_{i=0}^m f_1^B(t - T_1 - i\Delta t). \quad (12.17)$$

Выражение (12.17) дает возможность получить рекуррентное соотношение для определения интенсивности индивидуального вторичного потока:

$$f_1^B(t) = \begin{cases} f_1^C(t + T_1) - \sum_{i=1}^l f_1^B(t - i\Delta t), & l < m \\ f_1^C(t + T_1) - \sum_{i=1}^m f_1^B(t - i\Delta t), & l \geq m \end{cases},$$

где $l = t/\Delta t$; $\sum_{i=1}^0 f_1^B(t - i\Delta t) = 0$.

Отдельный интерес представляет ситуация, когда первичный поток является стационарным, т. е. $f_1^A(t) = a$, причем $G = [-\infty; +\infty]$ и $-\infty < T_0, T < +\infty$. Тогда уравнение (12.12) преобразуется к виду

$$f_1^C(t) = a \int_0^{\infty} f_1^B(x) dx \equiv ab, \quad (12.18)$$

где $b = \int_0^{\infty} f_1^B(x) dx$. Как видно из выражения (12.18), в данном случае невозможно получить однозначно искомую интенсивность индивидуально-

го вторичного потока, так как соотношению (12.18) удовлетворяет любая функция $f_1^B(x)$, для которой выполняется условие

$$\int_0^{\infty} f_1^B(x) dx = b < +\infty,$$

т. е. среднее число событий задаваемого ею индивидуального вторичного потока B_{τ} конечно. В таких ситуациях для получения $f_1^B(t)$ следует применять методы корреляционного анализа или методы, базирующиеся на использовании статистических данных, полученных с учетом моментов наступления событий первичного потока.

В заключение отметим, что как следует из выражения (12.9), суммарный поток, получаемый после наложения индивидуальных вторичных потоков, в общем случае имеет сложную структуру, даже если первичный поток является элементарным, например, потоком Пуассона, отрицательно-биномиальным потоком и т. д. В то же время в некоторых ситуациях при наложении определенных ограничений на отдельные вторичные потоки, тип суммарного потока S может быть простым при простом типе первичного потока.

С учетом выражения (12.11) условный ПФЛ индивидуальных вторичных потоков может быть записан через их моментные функции следующим образом:

$$L[u; \Omega | \tau; G] = 1 + \sum_{m=1}^{\infty} \frac{1}{m!} \int_{\Omega^m} f_m^B(t_1 - \tau, \dots, t_m - \tau) \prod_{j=1}^m u(t_j) dt_1 \dots dt_m. \quad (12.19)$$

Пренебрегая в (12.19) всеми слагаемыми начиная с третьего и подставляя приближенную формулу для $L[u; \Omega | \tau; G]$ в (12.9), получаем следующее выражение для ПФЛ суммарного потока

$$L[u; \Omega] = L \left[1 + \int_{\Omega} f_1^B(t - \tau) u(t) dt - 1; G \right] = L \left[\int_{\Omega} f_1^B(t - \tau) u(t) dt; G \right]. \quad (12.20)$$

Далее предположим, что первичный поток является потоком Пуассона. Тогда, используя выражение (10.20) для ПФЛ потока Пуассона и выражение (12.20), получим

$$\begin{aligned} L[u; \Omega] &= \exp \left\{ \int_G f_1^A(\tau) \left(\int_{\Omega} f_1^B(t - \tau) u(t) dt \right) d\tau \right\} = \\ &= \exp \left\{ \int_{\Omega} \left(\int_G f_1^A(\tau) f_1^B(t - \tau) d\tau \right) u(t) dt \right\}, \end{aligned} \quad (12.21)$$

где выполнено изменение порядка интегрирования. С учетом выражения (12.12), (12.21) может быть переписано в следующем виде:

$$L[u; \Omega] = \exp \left\{ \int_{\Omega} f_1^C(t) u(t) dt \right\}. \quad (12.22)$$

Как видно из полученного выражения для ПФЛ суммарного потока (12.22), в рамках ограничений, позволяющих выполнить приведенные выше преобразования, суммарный поток в данном случае также является потоком Пуассона. Смысл основного из ограничений, наложенных на вторичные потоки, состоит в том, что число событий в их реализациях должно быть достаточно малым. При выполнении данного условия можно показать, что тип суммарного потока совпадает с типом первичного потока также и в случаях, когда первичный поток является отрицательно-биномиальным потоком или потоком Бернулли.

Л е к ц и я 13

ПРЕОБРАЗОВАНИЕ СЛУЧАЙНЫХ ПОТОКОВ. СМО КАК ПРЕОБРАЗОВАТЕЛЬ СЛУЧАЙНЫХ ПОТОКОВ

При проведении ряда экспериментов, в качестве экспериментальных данных регистрируются реализации потоков случайных событий, которые несут информацию об исследуемом объекте. Такие потоки случайных событий являются результатом ряда преобразований некоторого исходного случайного потока при его прохождении через исследуемую систему. Такие преобразования исходного потока в общем случае носят случайный характер и связаны преимущественно с изменением положения точек в реализации исходного потока, а также с изменением их количества. Причинами изменения местоположения точек могут быть, например, неточности, возникающие при фиксации момента появления регистрируемого события, а также особенности работы самого детектора. Причинами изменения количества событий в реализациях потока являются, как правило, потери событий в детекторах, например, из-за отличной от единицы квантовой эффективности и наличия конечного мертвого времени. Вместе с тем возможны и случаи, когда число событий в регистрируемой реализации потока возрастает по сравнению с реализацией исходного потока. Такая ситуация обычно имеет место в случае, если в изучаемой системе присутствуют различного рода шумы (например, при использовании детекторов с электронными умножителями), а также послеимпульсы.

Для формализованного описания процесса трансформации исходного потока в поток, который получается после некоторого преобразования, будем использовать способ совместного задания моментов наступления событий этих потоков.

Пусть исходный поток событий задан на интервале $G = [T_1, T_2]$ совокупностью моментов наступления его событий $\gamma_1, \gamma_2, \dots$. Поток после преобразования наблюдается на интервале $\Omega = [T_0, T]$, где регистрируется совокупность событий преобразованного потока $\xi_1, \xi_2, \dots$. Для совместного описания совокупности случайных величин $\gamma_1, \gamma_2, \dots$ и $\xi_1, \xi_2, \dots$ по аналогии с тем, как это было сделано для отдельного случайного потока, введем систему совместных плотностей

$$\{\pi_{mk}(\tau_1, \dots, \tau_m; G; t_1, \dots, t_k; \Omega), m = 0, 1, \dots, k = 0, 1, \dots\}.$$

Данная система плотностей определяет совместную вероятность наступления точно m событий исходного потока на области G , лежащих в окрестностях точек $\tau_1, \dots, \tau_m$ и наступления точно k событий преобразованного потока на области Ω , которые попали в окрестности точек $t_1, \dots, t_k$. Представленные выше совместные плотности позволяют записать совместный ПФЛ двух систем точек для рассматриваемого преобразования:

$$\begin{aligned} L[v; G; u; \Omega] = & \sum_{m=0}^{\infty} \sum_{k=0}^{\infty} \frac{1}{m!} \frac{1}{k!} \int_{G^m} \int_{\Omega^k} \pi_{mk}(\tau_1, \dots, \tau_m; G; t_1, \dots, t_k; \Omega) \times \\ & \times \prod_{i=1}^m [1 + v(\tau_i)] \prod_{j=1}^k [1 + u(t_j)] dt_1 \dots dt_k d\tau_1 \dots d\tau_m. \end{aligned} \quad (13.1)$$

Приведенный выше совместный ПФЛ может быть переписан с учетом выражения для условных плотностей и соответственно определяемого через них условного ПФЛ преобразованного потока

$$\begin{aligned} \pi_{k|m}(t_1, \dots, t_k; \Omega | \tau_1, \dots, \tau_m; G) &= \frac{\pi_{mk}(\tau_1, \dots, \tau_m; G; t_1, \dots, t_k; \Omega)}{\pi_m(\tau_1, \dots, \tau_m; G)}, \\ L[u; \Omega | \tau_1, \dots, \tau_m; G] &= \\ &= \sum_{k=0}^{\infty} \frac{1}{k!} \int_{\Omega^k} \pi_{k|m}(t_1, \dots, t_k; \Omega | \tau_1, \dots, \tau_m; G) \prod_{j=1}^k [1 + u(t_j)] dt_1 \dots dt_k \end{aligned} \quad (13.2)$$

в следующем виде:

$$L[v; G; u; \Omega] = \sum_{m=0}^{\infty} \frac{1}{m!} \int_{G^m} \pi_m(\tau_1, \dots, \tau_m; G) \prod_{i=1}^m [1 + v(\tau_i)] L[u; \Omega | \tau_1, \dots, \tau_m; G] d\tau_1 \dots d\tau_m, \quad (13.3)$$

где $\pi_m(\tau_1, \dots, \tau_m; G)$, $m = 0, 1, \dots$ – плотности исходного потока, $L[u; \Omega | \tau_1, \dots, \tau_m; G]$ – ПФЛ преобразованного потока при условии, что события исходного потока произошли в моменты времени $\tau_1, \dots, \tau_m$. Условные плотности (13.2) полностью отражают процесс преобразования исходного потока.

Зная выражение для совместного производящего функционала $L[v; G; u; \Omega]$, нетрудно восстановить ПФЛ входного потока

$$L^I[v; G] = L[v; G; u; \Omega]_{u(t)=0} \quad (13.4)$$

и найти ПФЛ выходного потока

$$L^{II}[u; \Omega] = L[v; G; u; \Omega]_{v(t)=0}. \quad (13.5)$$

Далее будут рассмотрены некоторые простейшие виды преобразований исходного потока и получены соответствующие выражения для ПФЛ выходного потока.

Тождественное преобразование. Рассмотрим случай, когда выполняемое преобразование является тождественным, т. е. входной поток не подвергается никаким изменениям. Пусть $\Omega = G$, тогда очевидно:

$$\pi_{k|m}(t_1, \dots, t_k; \Omega | \tau_1, \dots, \tau_m; G) = \begin{cases} m! \prod_{i=1}^m \delta(t_i - \tau_i), & k = m \\ 0, & k \neq m \end{cases}. \quad (13.6)$$

Подставив последнее выражение в выражение для условного ПФЛ преобразованного потока (13.2), получим

$$L[u; \Omega | \tau_1, \dots, \tau_m; G] = \prod_{i=1}^m [1 + u(\tau_i)], \quad m = 0, 1, \dots \quad (13.7)$$

Заметим, что как видно из выражения (13.7), условный ПФЛ в рассматриваемом случае может быть представлен в следующем виде:

$$L[u; \Omega | \tau_1, \dots, \tau_m; G] = \prod_{i=1}^m L[u; \Omega | \tau_i], \quad (13.8)$$

где

$$L[u; \Omega | \tau_i] = \sum_{k=0}^{\infty} \frac{1}{k!} \int_{\Omega^k} \pi_{k|1}(t_1, \dots, t_k; \Omega | \tau_i; G) \prod_{j=1}^k [1 + u(t_j)] dt_1 \dots dt_k. \quad (13.9)$$

Представление (13.8) характерно для преобразований, действующих независимо по каждому событию исходного потока. Подставляя выражение (13.7) для условного ПФЛ в формулу (13.3), получаем следующее представление совместного ПФЛ исходного и преобразованного потоков:

$$L[v; G; u; \Omega] = \sum_{m=0}^{\infty} \frac{1}{m!} \int_{G^m} \pi_m(\tau_1, \dots, \tau_m; G) \prod_{i=1}^m [1 + v(\tau_i)] [1 + u(\tau_i)] d\tau_1 \dots d\tau_m,$$

откуда окончательно получаем

$$L[v; G; u; \Omega] = L^I [[1 + v(\cdot)] [1 + u(\cdot)] - 1; G]. \quad (13.10)$$

Потери событий. Перейдем к рассмотрению преобразования типа потерь событий. Предположим, что всякое событие исходного потока, которое наступило в момент времени τ , независимо от других событий остается в преобразованном потоке с вероятностью $q(\tau)$, либо, соответственно, теряется с вероятностью $1 - q(\tau)$. Функция $q(\tau)$, значения которой принадлежат отрезку $[0, 1]$, называется *функцией прорезживания* исходного потока. Так как рассматриваемое преобразование выполняется независимо по каждому из отдельных событий исходного потока, то для него справедливо соотношение (13.8). Тогда для получения условного ПФЛ преобразованного потока $L[u; \Omega | \tau_1, \dots, \tau_m; G]$ остается найти выражения для условных ПФЛ $L[u; \Omega | \tau]$. Очевидно, что плотности, на основе которых строятся эти ПФЛ, в рассматриваемой ситуации имеют вид

$$\pi_{k|1}(t_1, \dots, t_k; \Omega | \tau; G) = \begin{cases} 1 - q(\tau) & k = 0; \\ q(\tau) \delta(t_1 - \tau), & k = 1; \\ 0, & k > 1. \end{cases} \quad (13.11)$$

Подставляя выражение (13.11) в формулу (13.9), получаем условный ПФЛ $L[u; \Omega | \tau]$:

$$\begin{aligned} L[u; \Omega | \tau] &= 1 - q(\tau) + q(\tau) \int_{\Omega} \delta(t_1 - \tau) [1 + u(t_1)] dt_1 = \\ &= 1 - q(\tau) + q(\tau) [1 + u(\tau)] = 1 + q(\tau) u(\tau). \end{aligned}$$

Далее с учетом выражения (13.8) получаем формулу для условного ПФЛ преобразованного потока событий

$$L[u; \Omega | \tau_1, \dots, \tau_m; G] = \prod_{i=1}^m [1 + q(\tau_i)u(\tau_i)].$$

Используя полученное выражение для условного ПФЛ и учитывая (13.3), получаем формулу для совместного ПФЛ исходного и преобразованного потоков при преобразовании типа потерь событий:

$$L[v; G; u; \Omega] = \sum_{m=0}^{\infty} \frac{1}{m!} \int_{G^m} \pi_m(\tau_1, \dots, \tau_m; G) \prod_{i=1}^m [1 + v(\tau_i)] [1 + q(\tau_i)u(\tau_i)] d\tau_1 \dots d\tau_m,$$

или, учитывая общий вид ПФЛ исходного потока и переписывая последнее выражение в более компактном виде

$$L[v; G; u; \Omega] = L^I [[1 + v(\cdot)] [1 + q(\cdot)u(\cdot)] - 1; G]. \quad (13.12)$$

Последнее выражение позволяет найти формулу для ПФЛ прореженного потока с функцией прореживания $q(\tau)$, которое с учетом (13.5) имеет вид

$$L^II [u; \Omega] = L^I [q(\cdot)u(\cdot); G]. \quad (13.13)$$

Для преобразования типа потерь событий, в частности, можно рассмотреть случай, когда $\Omega \subset G$, а все события исходного потока, которые попадают в область $\Omega_1 = G \setminus \Omega$, теряются. Такое преобразование является частным случаем предыдущего при условии, что функция прореживания имеет следующий вид:

$$q(\tau) = \begin{cases} 1, & \tau \in \Omega \\ 0, & \tau \in \Omega_1 \end{cases}.$$

Подставляя последнее выражение в (13.13), получаем ПФЛ преобразованного потока на подобласти Ω :

$$L^II [u; \Omega] = L^I [u(\cdot); G]_{t \in \Omega}, \quad (13.14)$$

что, как и ожидалось, отражает равенство в пределах Ω ПФЛ преобразованного и исходного потоков.

Изменение местоположения событий. Рассмотрим линейное независимое преобразование точек исходного потока, имеющее вид

$$\xi_i = \gamma_i + \beta_i, \quad i = 1, 2, \dots,$$

где β_i – одинаково распределенные случайные величины, характеризующие смещение точек исходного потока. Обозначим через $B(t)$ функцию распределения случайных величин β_i и найдем выражение для условного ПФЛ $L[u; \Omega | \tau]$ на расширенной области $\Omega' = [-\infty; +\infty]$. Условные плотности, соответствующие данному ПФЛ, имеют вид

$$\pi_{k|1}(t_1, \dots, t_k; \Omega' | \tau; G) = \begin{cases} \frac{d}{dt} B(t_1 - \tau), & k = 1 \\ 0, & k \neq 1 \end{cases}.$$

Тогда выражение для самого условного ПФЛ $L[u; \Omega' | \tau]$ будет выглядеть следующим образом:

$$\begin{aligned} L[u; \Omega' | \tau] &= \int_{\Omega'} [1 + u(t_1)] \frac{d}{dt} B(t_1 - \tau) dt_1 = \\ &= \int_{\Omega'} dB(t_1 - \tau) + \int_{\Omega'} u(t_1) dB(t_1 - \tau) = 1 + \int_{\Omega'} u(t) dB(t - \tau). \end{aligned} \quad (13.15)$$

Проведем обратное сужение области Ω' к Ω . Данную операцию можно рассматривать как операцию по удалению точек. На основании этого, воспользовавшись выражением (13.14), получим выражение для $L[u; \Omega | \tau]$:

$$L[u; \Omega | \tau] = 1 + \int_{\Omega} u(t) dB(t - \tau). \quad (13.16)$$

Далее с учетом выражений (13.8) и (13.16) получаем формулу для условного ПФЛ преобразованного потока событий:

$$L[u; \Omega | \tau_1, \dots, \tau_m; G] = \prod_{i=1}^m \left[1 + \int_{\Omega} u(t) dB(t - \tau_i) \right]. \quad (13.17)$$

Выражение для совместного ПФЛ исходного и преобразованного потоков может быть получено для преобразования, изменяющего местоположения событий, путем подстановки выражения (13.17) в формулу (13.3). Выполняя данную подстановку и записывая полученное выражение в компактном виде, получаем

$$L[v; G; u; \Omega] = L^I \left[\left[1 + v(\cdot) \right] \left[1 + \int_{\Omega} u(t) dB(t - \cdot) \right] - 1; G \right]. \quad (13.18)$$

Последнее выражение позволяет получить ПФЛ преобразованного потока, который с учетом (13.5) имеет вид

$$L^{II}[u; \Omega] = L^I \left[\int_{\Omega} u(t) dB(t - \cdot); G \right]. \quad (13.19)$$

Используя формулы, связывающие ПФЛ для входного и выходного потоков для определенного вида преобразования, можно получать выражения для характеристик выходного потока. Так, например, используя полученную формулу (13.19), можно найти выражение для интенсивности преобразованного потока. Для этого достаточно выполнить функ-

циональное дифференцирование выражения (13.19) при условии, что пробная функция $u(t) = 0$. В результате получим выражение, связывающее интенсивность $f_1^{II}(t)$ преобразованного потока с интенсивностью $f_1^I(\tau)$ исходного потока и функцией распределения случайного смещения:

$$f_1^{II}(t) = \int_G f_1^I(\tau) \frac{dB(t-\tau)}{dt} d\tau.$$

Рассмотрим более общий случай, когда точки исходного потока независимо друг от друга подвергаются некоторому нелинейному преобразованию вида

$$\xi_i = D(\gamma_i), \quad i = 1, 2, \dots \quad (13.20)$$

Обозначим через D^{-1} обратное к (13.20) преобразование и будем считать, что оно является однозначным, т. е. $\gamma_i = D^{-1}(\xi_i)$. Тогда выражение для условных плотностей $\pi_{k|1}(t_1, \dots, t_k; \Omega | \tau; G)$ будет иметь вид

$$\pi_{k|1}(t_1, \dots, t_k; \Omega | \tau; G) = \begin{cases} \delta(t_1 - D(\tau)), & k = 1 \\ 0, & k \neq 1 \end{cases}.$$

С учетом этого получаем выражение для $L[u; \Omega | \tau]$:

$$L[u; \Omega | \tau] = 1 + u(D(\tau)).$$

Отсюда, находя условный ПФЛ преобразованного потока согласно (13.8) и далее подставляя полученное выражение в формулу (13.3), получаем выражение для совместного ПФЛ исходного и преобразованного потоков:

$$L[v; G; u; \Omega] = L^I [[1 + v(\cdot)] [1 + u(D(\cdot))] - 1; G].$$

Как следствие ПФЛ выходного потока для рассматриваемого преобразования имеет вид

$$L^{II}[u; \Omega] = L^I [u(D(\cdot)); G]. \quad (13.21)$$

Выполняя функциональное дифференцирование выражения (13.21) при условии $u(t) = 0$, получаем связь между интенсивностями исходного и преобразованного потоков:

$$f_1^{II}(t) = f_1^I(D^{-1}(t)) \left| \frac{dD^{-1}(t)}{dt} \right|.$$

Стоит отметить, что при исследовании реальных систем обработки информации часто обнаруживается, что в процессе их функционирования входные данные и результаты их обработки системой представимы в

виде последовательности событий, имеющих определенное положение во времени. Для построения математических моделей таких систем используется понятие системы массового обслуживания (СМО).

СМО как преобразователь случайных потоков. Обобщенная схема системы массового обслуживания представлена на рис. 13.1.

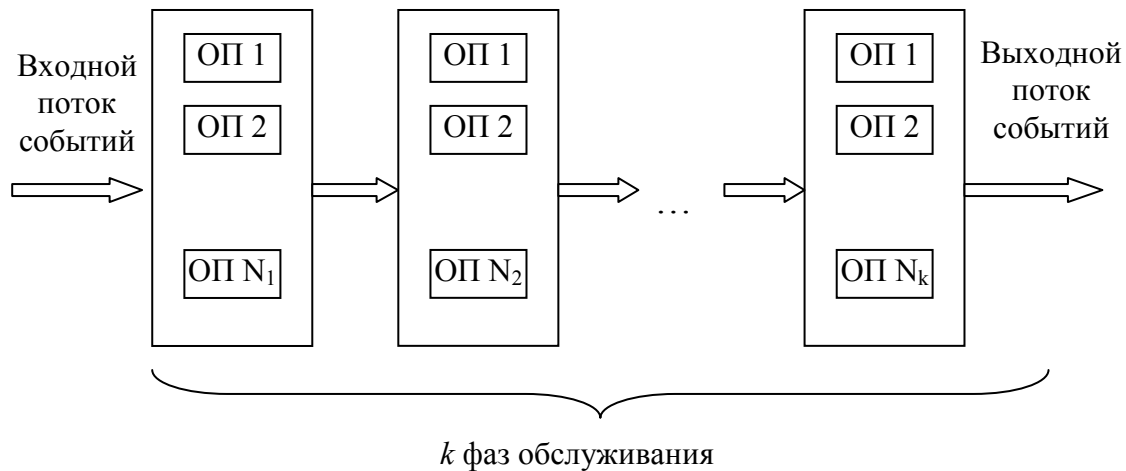


Рис. 13.1. Схематичное представление СМО

Рассмотрим основные структурные элементы, составляющие систему массового обслуживания.

Потоки событий. Поток событий представляет собой совокупность событий, каждое из которых характеризуется только определенным положением во времени. *Входной поток* – это совокупность событий, поступающих на вход системы и требующих обработки. *Выходной поток* – это совокупность событий, сформированных в результате работы системы. В общем случае события различных потоков в пределах СМО могут различаться по типу и приоритету обслуживания.

Обслуживающие приборы. Обслуживающие приборы являются элементами СМО, которые выполняют обработку событий, поступающих на их вход, и формируют выходной поток событий. Смысл процесса обслуживания события входного потока заключается обычно в изменении его местоположения на временной оси в соответствии с *функцией распределения времени обслуживания прибора* $B(t, \tau)$. Функция распределения времени обслуживания зависит от момента времени τ , в который начинается обслуживание очередного события, и определяет вероятность того, что поступившее в момент времени τ на обслуживающий прибор событие обслужится к моменту времени t :

$$B(t, \tau) = P\{\xi < t\},$$

где ξ – случайная величина, характеризующая время обслуживания.

Очереди событий. Очереди событий используются в СМО для размещения событий, поступивших в систему и ожидающих момента времени, когда соответствующий обслуживающий прибор освободится и сможет начать их обработку. При наличии очередей возможны различные схемы, задающие порядок обслуживания ожидающих в очереди событий, которые называются *дисциплинами обслуживания*. Среди наиболее известных можно выделить такие дисциплины обслуживания ожидающих событий, как:

- *FIFO (First In First Out).* Данная дисциплина определяет прямой порядок обслуживания событий, при котором события обслуживаются в порядке поступления.
- *LIFO (Last In First Out).* Данная дисциплина определяет обратный или стековый порядок обслуживания, когда на обслуживание выбирается событие, поступившее в систему последним.
- *Случайный порядок обслуживания,* при котором очередное событие для обслуживания выбирается случайно с равной вероятностью.

При построении СМО возможны две основные схемы их поведения в случае, если в момент поступления какого-либо входного события все доступные обслуживающие приборы оказываются занятыми. Согласно первой схеме (*системы с потерями*) такое событие просто получает отказ (или, как говорят, «теряется») и все дальнейшее течение процесса обслуживания идет так, как если бы этого события вообще не поступало. При другом устройстве СМО (*системы с ожиданием*) событие, заставшее все приборы занятыми, сохраняется в очереди событий как претендент на будущее обслуживание и в дальнейшем занимает один из освободившихся приборов. Кроме того, возможен и некоторый промежуточный вариант поведения СМО (*системы с очередью ограниченной длины*), когда события, поступающие для обслуживания, остаются в очереди, если число ранее прибывших и ожидающих обслуживания событий не превышает заданного числа k , в противном случае событие теряется.

В некоторых СМО пребывание событий в очереди может ограничиваться по времени. В данном случае возможен вариант, когда каждое требование остается в обслуживающей системе не более чем на время Δt , даже если началось его обслуживание. Возможен также и вариант, когда время ожидания ограничено величиной Δt , но если до истечения этого срока обслуживание началось, то оно доводится до конца.

В описанных выше вариантах поведения СМО предполагалось, что обслуживающие приборы обладают абсолютной надежностью, и сами

никогда не выходят из рабочего состояния. Это несколько идеализирует реальные системы. По этой причине при построении ряда СМО учитывается возможность выхода обслуживающих приборов из строя.

Еще одной особенностью построения СМО, которая встречается достаточно часто, является *обслуживание с преимуществом*. Эта особенность состоит в следующем. В систему массового обслуживания поступает не один, а несколько потоков требований. Требования потоков с меньшими номерами пользуются преимущественным правом обслуживания и становятся при появлении в очередь впереди всех ранее поступивших требований с более высокими номерами породивших их потоков. Здесь в свою очередь возможны два варианта функционирования СМО: требование высокого ранга не прерывает обслуживание требований менее высоких рангов; требование высокого ранга прерывает уже проводившееся обслуживание более низкого ранга. Во втором случае также различают две возможности: при возвращении вытесненного требования в обслуживающий прибор предыдущая работа забывается и обслуживание начинается сначала, либо ранее произведенная работа не забывается.

При всем многообразии представленных выше схем для построения СМО процесс функционирования любой из возможных систем сводится к преобразованию входного потока событий в выходной поток, при котором события входного потока могут изменять свое местоположение во времени, а также может изменяться число событий в выходном потоке по сравнению с входным. Как было показано выше, такие типы преобразования могут быть описаны математически с применением аппарата ПФЛ.

Рассмотрим входной и выходной потоки событий, которые присутствуют в структуре СМО, как исходный и преобразованный потоки событий, которые рассматривались при описании преобразования случайных потоков. В таком представлении становится очевидным, что теория, приведенная ранее и основанная на использовании ПФЛ, становится справедливой и для описания СМО. Для совместного описания входного и выходного потоков заявок могут использоваться системы совместных плотностей $\{\pi_{mk}(\tau_1, \dots, \tau_m; G; t_1, \dots, t_k; \Omega), m = 0, 1, \dots, k = 0, 1, \dots\}$ и выражения для совместного ПФЛ входного и выходного потоков (13.1) и (13.3), а также для ПФЛ входного (13.4) и выходного (13.5) потоков. При построении совместного ПФЛ, его конечное выражение будет отражать структуру конкретной рассматриваемой системы, а также будет зависеть от функций распределения времени обслуживания приборов, входящих в СМО.

Кроме характеристик входного и выходного потоков, которые могут быть получены с использованием выражений (13.4) и (13.5) для их ПФЛ соответственно, при помощи совместного ПФЛ можно найти и некоторые дополнительные характеристики рассматриваемой СМО. Так, например, одной из важнейших характеристик любой СМО является распределение числа событий $\eta(t)$, находящихся в системе в произвольный момент времени t . Обозначим производящую функцию $\eta(t)$ как $\varphi_1(z, t)$. Для нахождения $\varphi_1(z, t)$ прервем поступление заявок на вход системы в момент времени t . Тогда число заявок на выходе системы на интервале $[t, \infty]$ в точности совпадает с $\eta(t)$. Отсюда, учитывая связь производящей функции и ПФЛ $\Theta(z, D) = L[z - 1; D]$, получим выражения для $\varphi_1(z, t)$:

$$\varphi_1(z, t) = L[v; [T_1, t]; z - 1; [t, \infty]]_{v(\tau)=0}, \quad (13.22)$$

где T_1 – начало интервала, на котором определен входной поток.

Также легко, при известном ПФЛ выходного потока, определяется производящая функция числа заявок, обслуженных за произвольный интервал времени Ω :

$$\varphi_2(z, t) = L''[z - 1; \Omega]. \quad (13.23)$$

Сами распределения, для которых выше были представлены производящие функции, могут быть получены путем многократного дифференцирования по переменной z так же, как это было показано ранее для распределения числа событий обычного случайного потока (формула (9.24)).

Продemonстрируем использование аппарата ПФЛ для описания СМО на примере СМО с бесконечным числом обслуживающих приборов.

Пример 13.1. Рассмотрим СМО с бесконечным числом обслуживающих приборов. Пусть на вход такой системы поступает произвольный поток событий, задаваемый ПФЛ $L^I[v; G]$. Пусть функции распределения времени обслуживания $B(\tau, t)$ одинаковы для всех обслуживающих приборов в рассматриваемой системе. Так как число приборов в системе бесконечно, то моменты окончания обслуживания событий входного потока при фиксированных моментах их поступления в систему независимы. По этой причине условный ПФЛ выходного потока представим в виде

$$L[u; \Omega | \tau_1, \dots, \tau_m; G] = \prod_{i=1}^m L[u; \Omega | \tau_i],$$

где $L[u; \Omega | \tau_i]$ – ПФЛ потока состоящего из одного события, отстоящего от точки τ на некоторое случайное расстояние, задаваемое распределением $B(\tau, t)$. Согласно

рассмотренному ранее преобразованию смещения точек потока событий ПФЛ этого одноточечного потока по аналогии с (13.16) имеет вид

$$L[u; \Omega | \tau] = 1 + \int_{\Omega} u(t) dB(\tau, t).$$

Тогда, с целью получения совместного ПФЛ входного и выходного потоков для рассматриваемой СМО, необходимо проделать преобразования аналогичные приведенным выше, при рассмотрении преобразования событий случайного потока типа смещения во времени. На этом основании выражение для совместного ПФЛ входного и выходного потоков в данном случае будет выглядеть следующим образом:

$$L[v; G; u; \Omega] = L' \left[\left[1 + v(\cdot) \right] \left[1 + \int_{\Omega} u(t) dB(\cdot, t) \right] - 1; G \right].$$

Откуда ПФЛ выходного потока:

$$L''[u; \Omega] = L' \left[\int_{\Omega} u(t) dB(\cdot, t); G \right].$$

На основании (13.22) получаем ПФ числа заявок в системе на момент времени t

$$\varphi_1(z, t) = L'[(z-1)(1-B(\cdot, t)); [T_1, t]].$$

В частности, если входной поток является потоком Пуассона с интенсивностью $\lambda(t)$, то, учитывая выражение (10.20) для ПФЛ такого потока, получим следующее выражение для $\varphi_1(z, t)$:

$$\varphi_1(z, t) = \exp \left\{ (z-1) \int_{T_1}^t (1-B(\tau, t)) \lambda(\tau) d\tau \right\}.$$

5. АНАЛИЗ СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК ПОТОКОВ СЛУЧАЙНЫХ СОБЫТИЙ

Л е к ц и я 14

ОБОБЩЕННАЯ СХЕМА АНАЛИЗА СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК ПОТОКОВ СЛУЧАЙНЫХ СОБЫТИЙ

При выполнении практических исследований в различных областях науки данные, получаемые в процессе проведения экспериментов, могут быть представлены в виде последовательности событий, которые образуют поток. Более того, часто приходится иметь дело с экспериментами, при проведении которых потоки случайных событий различной природы претерпевают ряд преобразований. При этом под воздействием исследуемой системы на исходный случайный поток изменяется местоположение событий и их количество. Изменения, которые отражаются на характере регистрируемого потока, несут информацию об исследуемой системе. Для извлечения данной информации зарегистрированная реализация преобразованного случайного потока подвергается предварительной обработке с целью получения различного рода статистических характеристик, которые далее используются в процессе анализа для получения значений параметров, характеризующих те или иные свойства исследуемой системы.

Одними из наиболее ярких примеров экспериментов, результаты которых представимы в дискретно-событийном виде, могут служить эксперименты во флуоресцентной спектроскопии. В основе таких экспериментов лежит способность молекул ряда веществ поглощать падающие на них фотоны определенной длины волны, и затем через некоторое время испускать фотоны на другой длине волны, свойственной данному виду молекул. Время между поглощением фотона и его повторным испусканием является случайной величиной и несет информацию об исследуемом образце, а также о процессах, которые протекают в рассматриваемой молекулярной системе.

Далее будет дано краткое описание ряда флуоресцентных экспериментов, а также дано представление о характере получаемых в результате их проведения данных.

Импульсные измерения кривых затухания флуоресценции. В экспериментах по импульсному измерению кривых затухания флуоресценции (рис. 14.1) на исследуемый образец направляется поток фотонов.

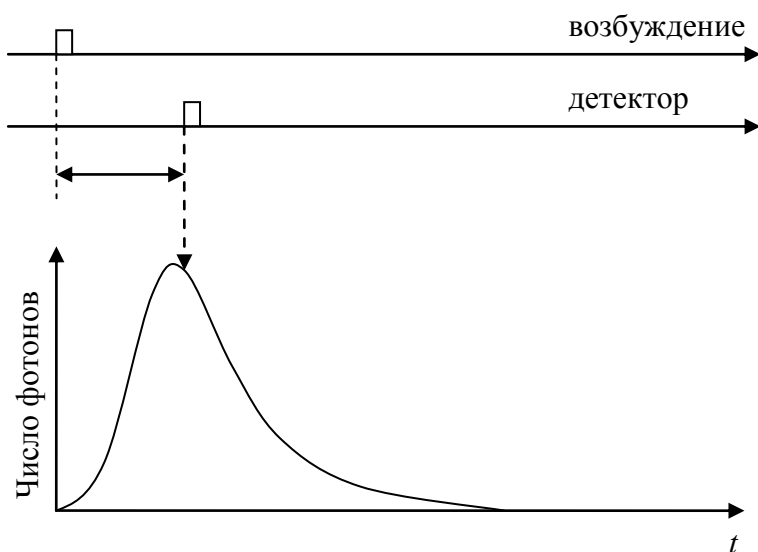


Рис. 14.1. Формирование кривой затухания флуоресценции

После некоторой временной задержки, зависящей от оптических и фотофизических процессов в веществе, фотоны, поглощенные образцом, переизлучаются и регистрируются на детекторе. Детектирующее устройство позволяет зарегистрировать поток случайных событий, представленный временами прихода фотонов. При этом нулевой момент времени соответствует приходу импульса возбуждения. Для регистрации кривых затухания флуоресценции наиболее часто применяется метод счета единичных фотонов. Согласно данному методу исследуемый образец возбуждается последовательностью импульсов настолько малой интенсивности, чтобы на детекторе можно было зарегистрировать время прихода единичного фотона, причем этот фотон является единственным приблизительно на 20 импульсов возбуждения. Поскольку испускание флуоресценции является случайным явлением, то вероятность испускания единичного фотона с определенной временной задержкой в различные моменты времени будет той же, что и вероятность испускания фотонов с той же временной задержкой для большого числа флуорофоров.

В качестве анализируемой статистической характеристики регистрируемого потока фотонов используется *интенсивность затухания флуоресценции*, которая представляет собой гистограмму, построенную по интервалам времени между регистрацией переизлученных фотонов на детекторе и временем прихода оптического импульса возбуждения.

Флуоресцентная спектроскопия отдельных неподвижных молекул. Данный вид флуоресцентных экспериментов является достаточно чувствительным и эффективным методом для исследования свойств отдельных молекул. При проведении типичного эксперимента такого типа можно выделить три потока событий, которые представлены на рис. 14.2.

Первый поток событий представляет собой фотоны возбуждения, временное местоположение которых полностью определяется функцией

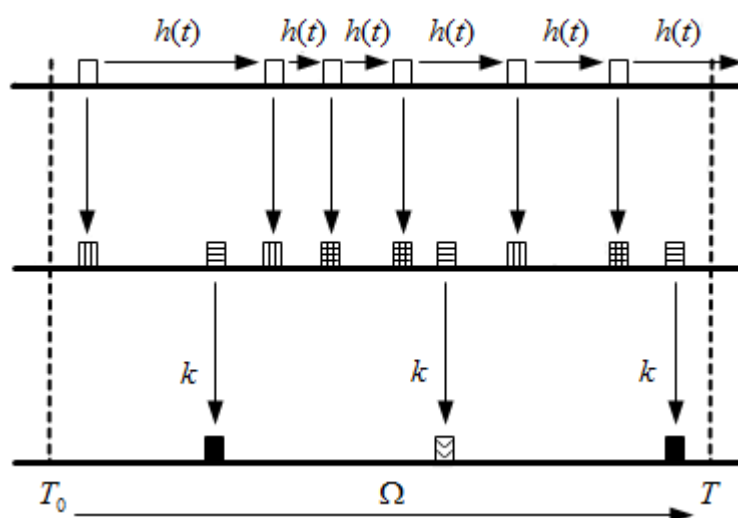


Рис. 14.2. Схема регистрации фотонов при исследовании отдельных неподвижных молекул методом флуоресцентной спектроскопии

плотности распределения расстояний между фотонами $h(t)$. В случае если возбуждение выполняется при помощи хорошо стабилизированного лазера и является непрерывным, то временные промежутки между соседними фотонами распределены по экспоненциальному закону. В случае импульсного возбуждения каждый возбуждающий импульс состоит из множества фотонов, в то время как отдельной молекулой поглощается только один из них. В этом случае расстояние между соседними фотонами можно считать одинаковым, а $h(t)$ будет представлять собой дельта-функцию.

Отдельная молекула в невозбужденном состоянии может поглощать фотон (вертикально заштрихованные символы), переходя при этом в возбужденное состояние. В случае если другие фотоны попадают на возбужденную молекулу, то они теряются (вертикально и горизонтально заштрихованные символы). Спустя некоторое время молекула переходит обратно в невозбужденное состояние либо путем испускания фотона (горизонтально заштрихованные символы), либо безизлучательно. Испу-

щенные фотоны (закрашенные символы) регистрируются детектором в пределах интервала наблюдения $\Omega = [T_0; T]$ с вероятностью k , которая определяет эффективность детектирующей системы. В противном случае они теряются (символы, заштрихованные волнистой линией). Испущенные фотоны представляют собой второй поток событий. Фотоны, зафиксированные детектором, образуют третий поток.

Статистические характеристики зарегистрированного детектором потока фотонов содержат информацию о фотонах возбуждения, а также о свойствах исследуемой молекулы. Среди таких статистических характеристик можно выделить корреляционную функцию, распределение интервалов между соседними фотонами, распределение числа фотонов на интервале наблюдения, распределения первого и единственного события на интервале наблюдения. При анализе перечисленных выше статистических характеристик с использованием соответствующих моделей можно получить значения параметров, характеризующих определенные свойства исследуемой молекулы.

Флуоресцентная флуктуационная спектроскопия. При проведении такого рода экспериментов фиксируются, а затем анализируются статистические флуктуации интенсивности флуоресценции, полученной путем засветки малого объема, через который могут проходить молекулы исследуемого образца. Молекулы образца заключены в пределах кюветы некоторого объема V , в пределах которого они могут свободно диффундировать. Луч хорошо стабилизированного лазерного излучения фокусируется на центре кюветы, формируя в ее пределах область с неоднородной интенсивностью засветки (максимальной в фокусе и быстро спадающей к периферии). Каждый раз, когда флуоресцентная молекула входит в освещенный объем, регистрируется всплеск интенсивности флуоресценции. Изменение интенсивности флуоресценции регистрируется даже тогда, когда молекула остается в объеме, но изменяет свое местоположение вследствие сильной неоднородности засветки (максимальная в фокусе и быстро спадающая к периферии). Флуктуации интенсивности могут быть также вызваны процессами, влияющими на квантовый выход флуоресценции вещества (например, переход в триплетное состояние). Из теории детектирования известно, что при постоянной интенсивности излучения, падающего на детектор, на выходе детектора будет наблюдаться поток событий, который является потоком Пуассона. В случае если в падающей интенсивности наблюдаются флуктуации случайного характера, то на выходе детектора будет наблюдаться уже дважды стохастический поток событий. Таким образом, флуктуации интенсивности флуоресценции, порождаемые различными процессами, происходящими

в изучаемой молекулярной системе, определяют отклонение характера потока на выходе детектора от потока Пуассона. Как следствие описанные выше флуктуации будут отражаться и на форме статистических характеристик, построенных на базе случайного потока, регистрируемого на выходе детектора. В результате при анализе построенных статистических характеристик, таких как корреляционная функция, распределение числа фотонов на интервале наблюдения и т. д., с использованием соответствующих моделей можно получать значения ряда параметров, отражающих те или иные свойства исследуемой молекулярной системы (коэффициент диффузии, время нахождения в триплетном состоянии и т. д.).

Как следует из приведенных выше примеров различного рода экспериментов, в результате предварительной обработки зарегистрированных в ходе их проведения случайных потоков получают некоторый набор статистических характеристик, которые несут информацию об определенных свойствах исследуемой системы. Стоит отметить, что получаемые из экспериментальных случайных потоков статистические характеристики не являются абсолютно точными, а содержат искажения, обусловленные наличием статистического шума, характер которого, вообще говоря, зависит как от вида самой статистической характеристики, так и от типа проводимого эксперимента. По этой причине для получения количественных оценок параметров, характеризующих интересующие исследователя свойства, построенные статистические характеристики должны быть проанализированы с применением ряда специально предназначенных для этого алгоритмов.

Вне зависимости от конкретного эксперимента, при проведении которого были получены статистические характеристики случайного потока событий, их обработка может проводиться по единой схеме, которая позволяет определить четкую структуру выполняемого анализа. Это в свою очередь дает возможность выработать методику проведения вычислений и использовать ряд методов, для которых уже существует достаточно мощный и хорошо разработанный аналитический аппарат. Обобщенная схема проведения анализа статистических характеристик, случайного потока представлена на рис. 14.3.

Как видно из приведенной схемы, выполнение анализа обеспечивается рядом взаимодействующих друг с другом объектов, каждый из которых выполняет достаточно четко определенную отдельную задачу. Процесс анализа представляет собой организованную по определенным правилам последовательность взаимодействий между объектами, включенными в схему.



Рис. 14.3. Обобщенная схема анализа статистических характеристик потока событий

Рассмотрим отдельные этапы подготовки и выполнения анализа статистических характеристик случайного потока событий.

Первым предварительным этапом проведения анализа является *подготовка исходных данных* в виде набора статистических характеристик. Предназначенные для анализа статистические характеристики могут быть либо получены непосредственно в ходе измерений, либо вычисляться на основании измеренного потока случайных событий. Местоположение отдельных событий потока на оси времени несет информацию об исследуемом образце, породившем случайный поток. Часть данной информации будет содержаться в определенной статистической характеристике, построенной на основании анализируемого случайного потока. Способ, применяемый для вычисления требуемой статистической характеристики исследуемого потока случайных событий, зависит от вида выбранной характеристики.

Поскольку каждая статистическая характеристика может содержать различные части информации из исходного потока случайных событий, то построение и анализ нескольких статистических характеристик дают более полную картину об изучаемом образце.

После получения исходных статистических характеристик исследуемой системы необходимо выбрать *набор моделей*, которые могли бы быть поставлены им в соответствие. Модель представляет собой математически формализованное описание некоторой статистической характеристики потока, которое отражает процессы, протекающие в исследуемом образце. В то время как исходная статистическая характеристика потока представляет собой данные, которые должны быть проанализированы, модель позволяет построить соответствующую теоретическую зависимость (теоретическую характеристику), которая сравнивается в процессе анализа с исходной. Форма теоретической характеристики, сгенерированной моделью, зависит от значений одной или более переменных, называемых параметрами модели. Математически теоретическая характеристика может быть представлена в следующей обобщенной форме:

$$F^{th} = f(\Omega, \vec{a}), \quad (14.1)$$

где Ω представляет собой область значений аргумента, на которой определена выбранная статистическая характеристика потока, $\vec{a} = \{a_1, a_2 \dots a_m\}$ – вектор параметров, f обозначает алгоритм, позволяющий вычислить теоретическую характеристику на Ω в соответствии со значениями $\vec{a}$.

Каждый параметр $a_i, i = 1 \dots m$ в пределах выбранной модели имеет достаточно определенный смысл и характеризует систему, описываемую

моделью в целом. Например, в широко известной много-экспоненциальной модели, применяемой для анализа кривых затухания флуоресценции, параметр, называемый временем жизни, обозначает среднее время, в течение которого флуорофор остается в возбужденном состоянии.

В зависимости от формы алгоритма f все модели могут быть разделены на две большие группы: *аналитические* и *имитационные*. В случае использования аналитических моделей теоретическая характеристика вычисляется в соответствии с некоторой математической формулой. Вследствие этого, теоретические характеристики, сгенерированные при помощи аналитических моделей, не содержат статистического шума. Имитационные модели используют для построения теоретической характеристики алгоритмы статистического моделирования. Примерами имитационных моделей могут служить достаточно часто упоминаемые в литературе модели, построенные при помощи метода Монте-Карло. Алгоритм генерации теоретической характеристики в имитационных моделях обычно очень детально отражает природу процессов, протекающих в исследуемом образце. Основным недостатком имитационных моделей является то, что, по сравнению с аналитическими моделями, они требуют значительно большего времени вычислений для получения теоретической характеристики. Поскольку алгоритмы, заложенные в основу имитационных моделей, используют генераторы случайных величин, теоретические характеристики, получаемые с помощью таких моделей, содержат статистический шум.

Следующим элементом, играющим важную роль в процессе анализа статистических характеристик потока случайных событий, является *алгоритм генерации начальных приближений* для параметров выбранной модели. Задачей данного алгоритма является получение такого набора значений параметров модели, при котором исходная статистическая характеристика потока и теоретическая характеристика, получаемая при помощи модели, будут максимально близки друг к другу. Для получения соответствующих значений параметров модели алгоритм генерации начальных приближений использует информацию об исходной статистической характеристике, построенной по имеющемуся потоку случайных событий. Значения параметров, получаемые при помощи алгоритма генерации начальных приближений, обычно несколько отличаются от оптимальных значений. Однако использование этих значений на следующем этапе итерационного анализа приводит к значительному ускорению процедуры анализа и получению более точных результатов.

После вычисления исходной статистической характеристики, выбора модели и генерации начальных приближений для ее параметров, может быть выполнена основная часть анализа, представляющая собой итерационную процедуру подгонки теоретической характеристики к исходной. Данная процедура позволяет получить такой набор параметров моделей, при котором достигается наилучшее совпадение исходной и теоретической характеристик. *Процедура итерационного анализа* строится на базе двух компонент: целевого критерия и метода оптимизации.

Целевой критерий предназначен для сравнения исходной и теоретической характеристик, а также для вычисления значения, которое показывает степень их совпадения. Большинство целевых критериев, представленных в литературе, являются следствием теории максимального правдоподобия. В соответствии с этой теорией целью анализа является максимизация функции правдоподобия, задаваемой следующим образом:

$$L(\vec{a}) = f(x_1, x_2, \dots, x_k, \vec{a}), \quad (14.2)$$

где x_i – значение исходной статистической характеристики в i -й точке ($i = 1, \dots, k$) области значений аргумента Ω , $\vec{a} = \{a_1, a_2 \dots a_m\}$ – вектор из m параметров.

Функция $f(x_1, x_2, \dots, x_k, \vec{a})$ представляет собой совместную плотность вероятности того, что именно значения $x_1, x_2, \dots, x_k$ будут значениями исходной статистической характеристики в $1, 2, \dots, k$ -й точках пространства аргументов. Для выполнения анализа необходимо найти такую комбинацию параметров $a_1, a_2 \dots a_m$, которая дает максимум функции правдоподобия $L(\vec{a})$. Более подробно некоторые целевые критерии будут рассмотрены позднее.

Методы оптимизации предназначены для поиска оптимального набора значений анализируемых параметров, при котором достигается максимум соответствия между исходной и теоретической характеристиками. Каждый метод оптимизации использует собственный, отличный от других математический алгоритм, предназначенный для изменения значений параметров в процессе проведения анализа. Все множество методов оптимизации может быть разделено на две большие группы: методы, не требующие вычисления значений производных целевого критерия по параметрам, и градиентные методы, алгоритмы которых требуют вычисления вышеуказанных производных.

Среди методов, не использующих значения производных от целевого критерия, можно выделить несколько классов, упоминаемых в литературе. *Методы прямого поиска* выполняют одномерный поиск по каждо-

му анализируемому параметру (сеточный метод, метод Хука-Дживса, метод Пауела, метод Розенброка). *Симплекс-методы* используют построение многоугольника (симплекса) в пространстве анализируемых параметров и перемещают центр данного многоугольника в точку в пространстве параметров, где значение целевого критерия является наилучшим. *Стохастические методы* используют процедуру поиска, основанную на случайном определении направления в пространстве параметров (метод случайного поиска, метод Монте-Карло, «simulated annealing» метод).

Применение *градиентных методов оптимизации* обеспечивает более быструю сходимость к оптимальному значению целевого критерия по сравнению с методами, не использующими производных. Однако в то же время данные методы требуют вычисления первых, а в некоторых случаях и вторых производных от целевого критерия, что не всегда является простой задачей. Наиболее часто в литературе упоминают следующие градиентные методы: метод наискорейшего спуска, метод Гаусса – Ньютона, метод Марквардта.

После нахождения оптимального набора параметров модели при помощи итерационной процедуры анализа, описанной выше, необходимо оценить *качество полученных результатов*. Кроме конечного значения целевого критерия для оценки качества полученных результатов, можно использовать ряд предназначенных для этой цели критериев. Наиболее часто для оценки качества проведенного анализа используются графические зависимости взвешенных остатков и их автокорреляционной функции от значений аргумента в области Ω , а также построение доверительных интервалов для полученных значений параметров, для чего может, например, использоваться метод исчерпывающего поиска. В случае если критериев, указанных выше, недостаточно для оценки качества проведенного анализа, могут быть использованы дополнительные критерии, такие как стандартное отклонение целевого критерия (Z_{χ^2}), параметр Дурбина – Ватсона, тест серий и т. д.

Отдельно стоит выделить такую часть обобщенной схемы анализа, как *связывание параметров*. Для того чтобы понять, что такое связывание параметров и с какой целью оно используется в процессе анализа, рассмотрим следующий пример.

Пусть для получения параметров, характеризующих свойства исследуемой системы, было построено несколько статистических характеристик, которые соответствуют нескольким проведенным экспериментам. Пусть в процессе проведения вышеуказанных экспериментов изменялись

условия, влияющие на состояние исследуемой системы, что в свою очередь привело к изменению значений части параметров, характеризующих состояние системы, в то время как значения оставшейся части параметров остались неизменными. Пусть модели построенных статистических характеристик содержат как параметры, значения которых отличались в различных экспериментах, так и параметры, значения которых должны были остаться неизменными.

В такой ситуации весь набор построенных статистических характеристик анализируется совместно, а сама процедура получения оптимального набора параметров с использованием совокупности статистических характеристик называется *глобальным анализом*. Такой тип анализа имеет смысл лишь в том случае, когда часть параметров моделей, соответствующих различным статистическим характеристикам, имеет одинаковые значения. Именно для таких параметров и применяется процедура связывания, что делает их значения одинаковыми в процессе анализа. Связанные параметры рассматриваются как один параметр, от которого зависят теоретические кривые всех моделей, параметры которых были связаны. Как следствие, возможность связывания параметров уменьшает общее число параметров подлежащих оптимизации. Это делает проводимый анализ более устойчивым к попаданию в локальные экстремумы целевого критерия и увеличивает вероятность определения набора параметров, который является наиболее оптимальным. Кроме того, при проведении совместного анализа нескольких статистических характеристик увеличивается количество экспериментальной информации, доступной для анализирующих алгоритмов, что делает анализ более устойчивым.

Таким образом, методика анализа статистических характеристик потока случайных событий, приведенная выше, содержит полный набор алгоритмов, позволяющий получать оценки значений параметров, характеризующих исследуемую систему, а также оценивать качество полученных результатов. Далее некоторые наиболее важные компоненты данной схемы будут рассмотрены более детально.

Л е к ц и я 15

АЛГОРИТМЫ ДЛЯ АНАЛИЗА ДАННЫХ (Часть 1)

Целевые критерии. Одним из наиболее важных элементов, входящих в схему анализа статистических характеристик случайных потоков, является целевой критерий. Данный элемент предназначен, прежде всего, для численной оценки степени подобия анализируемой статистиче-

ской характеристики изучаемой системы, полученной из эксперимента, и соответствующей ей теоретической кривой, полученной при помощи используемой для анализа модели при некоторой комбинации значений оцениваемых параметров. Необходимость в разработке и последующем использовании в процессе анализа специальных целевых критериев возникает по причине наличия случайных искажений в точках оцениваемых экспериментальных кривых. При использовании для нахождения оценок неизвестных параметров итерационных процедур оптимизации, значения целевого критерия в различных точках пространства параметров позволяют алгоритмам оптимизации вести поиск экстремального значения целевого критерия (т. е. в зависимости от конкретной ситуации его минимума или максимума). Поиск экстремума целевого критерия, который соответствует оптимальному набору значений оцениваемых параметров, обычно и является целью проведения оптимизации. Отсюда происходит и название самого «целевого» критерия.

Конкретный аналитический вид целевого критерия, который используется методами оптимизации для нахождения оценок неизвестных параметров, определяется характером статистического шума в анализируемой экспериментальной статистической характеристике. Для получения выражений целевых критериев обычно используется единый подход, который известен в литературе как метод максимального правдоподобия.

В основе метода максимального правдоподобия лежит предположение о том, что при определенном наборе значений параметров, характеризующих систему, вероятность получить статистическую характеристику, близкую к экспериментальной, будет максимальна.

Пусть в результате проведения эксперимента была получена некоторая статистическая характеристика, состоящая из k точек, принадлежащих области значений аргумента Ω , в которых ее значения равны $x_1, x_2, \dots, x_k$. Здесь важно отметить, что в простейшем случае рассматриваемая статистическая характеристика может быть функцией одного аргумента. Однако в общем случае это функция, заданная на многомерном пространстве аргументов, и тогда указанные выше значения $x_1, x_2, \dots, x_k$ представляют общую совокупность точек измеренной статистической характеристики. Пусть экспериментальная статистическая характеристика анализируется при помощи разработанной для ее описания теоретической модели, которая позволяет получать теоретическую кривую зависящую от m -мерного вектора параметров $\vec{a} = \{a_1, a_2 \dots a_m\}$. Тогда базовое выражение для функции правдоподобия в самом общем виде будет выглядеть следующим образом:

$$L(\vec{x}, \vec{a}) = f(x_1, x_2, \dots, x_k, \vec{a}), \quad (15.1)$$

где $f(x_1, x_2, \dots, x_k, \vec{a})$ представляет собой совместную плотность вероятности того, что именно значения $x_1, x_2, \dots, x_k$ будут значениями исходной статистической характеристики в $1, 2, \dots, k$ -й точках аргумента.

Для нахождения оценок неизвестных параметров $a_1, a_2 \dots a_m$ необходимо найти такую комбинацию их значений, которая дает максимум функции правдоподобия $L(\vec{a})$, представленной выражением (15.1). Компоненты вектора параметров $\vec{a}$, от которого зависит функция правдоподобия $L(\vec{a})$, определяются моделью, поставленной в соответствие анализируемой исходной статистической характеристике. По этой причине функция правдоподобия в выражении (15.1) будет зависеть от значений теоретической кривой, генерируемой моделью, которая в свою очередь является функцией искомого вектора параметров $\vec{a}$:

$$L(\vec{x}, F^{th}(\vec{a})) = f(x_1, x_2, \dots, x_k, F^{th}(\vec{a})). \quad (15.2)$$

Для получения окончательных выражений целевых критериев, используемых при проведении анализа, функция правдоподобия (15.2) нормируется на значение функции правдоподобия $L(\vec{x}, \vec{x})$, которая гарантирует абсолютно точное совпадение с исходной статистической характеристикой:

$$L_N(\vec{x}, F^{th}(\vec{a})) = \frac{L(\vec{x}, F^{th}(\vec{a}))}{L(\vec{x}, \vec{x})}. \quad (15.3)$$

Ввиду того, что функция логарифма является монотонно возрастающей, поиск максимума (15.3) эквивалентен поиску минимума логарифма нормированной функции правдоподобия, взятого со знаком '-'. При получении целевых критериев, применяемых на практике, используют следующее исходное выражение:

$$\chi^2(\vec{a}) = -2 \ln L_N(\vec{x}, F^{th}(\vec{a})) = -2 \ln \frac{L(\vec{x}, F^{th}(\vec{a}))}{L(\vec{x}, \vec{x})}. \quad (15.4)$$

Возможность дальнейших преобразований выражения (15.4) зависит от наличия или отсутствия связей между случайными величинами, реализации которых определяют значения в отдельных точках исходной статистической характеристики. В случае если значения в различных точках исходной статистической характеристики независимы, то совместная плотность вероятности в выражении (15.1) может быть представлена в следующем виде:

$$f(x_1, x_2, \dots, x_k, F^{th}(\vec{a})) = \prod_{i=1}^k f(x_i, F_i^{th}(\vec{a})), \quad (15.5)$$

где $f(x_i, F_i^{th}(\vec{a}))$ представляет собой плотность вероятности того, что значение x_i будет получено в i -й точке исходной статистической характеристики, $F_i^{th}(\vec{a})$ – значение теоретической характеристики в i -й точке пространства аргументов.

С учетом (15.5) выражение (15.4) может быть представлено в виде

$$\begin{aligned} \chi^2(\vec{a}) &= -2 \ln \frac{\prod_{i=1}^k f(x_i, F_i^{th}(\vec{a}))}{\prod_{i=1}^k f(x_i, x_i)} = -2 \ln \prod_{i=1}^k \frac{f(x_i, F_i^{th}(\vec{a}))}{f(x_i, x_i)} = \\ &= -2 \sum_{i=1}^k \ln \frac{f(x_i, F_i^{th}(\vec{a}))}{f(x_i, x_i)}. \end{aligned} \quad (15.6)$$

Формулы (15.4) и (15.6) обычно используются на практике для получения окончательных выражений целевых критериев. Для проведения дальнейших преобразований этих формул необходимо знать статистику шума, присутствующего в исходной статистической характеристике, что определяет конкретный вид функции $f(x_1, x_2, \dots, x_k, F^{th}(\vec{a}))$ в выражении (15.2) или $f(x_i, F_i^{th}(\vec{a}))$ в выражении (15.6). Далее рассмотрим несколько наиболее известных распределений шума, которые встречаются на практике (в частности, при получении экспериментальных кривых в ходе проведения флуоресцентных экспериментов), и получим соответствующие им целевые критерии.

Независимые случайные величины с Гауссовским распределением. Одним из наиболее используемых на практике является случай, когда значения, получаемые в отдельных точках исходной статистической характеристики, являются независимыми случайными величинами, каждая из которых подчиняется распределению Гаусса. Примером статистической характеристики, которая в точности соответствует вышеуказанному случаю, является кривая затухания флуоресценции, измеренная методом однофотонного счета. В данном примере при большом числе зарегистрированных фотонов шум в отдельных точках кривой затухания флуоресценции с достаточной степенью точности может считаться Гауссовским. Для нормальных случайных величин функция $f(x_i, F_i^{th}(\vec{a}))$ имеет вид

$$f(x_i, F_i^{th}(\vec{a})) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{(x_i - F_i^{th}(\vec{a}))^2}{2\sigma_i^2}\right), \quad (15.7)$$

где σ_i – среднеквадратическое отклонение шума в i -й точке пространства аргументов.

Подставляя выражение (15.7) в выражение (15.6), получим

$$\begin{aligned} \chi_G^2(\vec{a}) &= -2 \sum_{i=1}^k \ln \left[\frac{\frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{(x_i - F_i^{th}(\vec{a}))^2}{2\sigma_i^2}\right)}{\frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{(x_i - x_i)^2}{2\sigma_i^2}\right)} \right] = \\ &= 2 \sum_{i=1}^k \frac{(x_i - F_i^{th}(\vec{a}))^2}{2\sigma_i^2} = \sum_{i=1}^k \frac{(x_i - F_i^{th}(\vec{a}))^2}{\sigma_i^2}. \end{aligned} \quad (15.8)$$

Выражение (15.8) в точности соответствует целевому критерию хи-квадрат, использование которого лежит в основе хорошо известного метода наименьших квадратов. На практике обычно используется нормированное значение данного критерия, получаемое путем деления выражения (15.8) на число степеней свободы:

$$\chi_G^2(\vec{a}) = \frac{1}{k - m - 1} \sum_{i=1}^k \frac{(x_i - F_i^{th}(\vec{a}))^2}{\sigma_i^2}. \quad (15.9)$$

Значение целевого критерия, задаваемого выражением (15.9), полученное после выполнения анализа, может использоваться в качестве критерия оценки качества полученных результатов. Анализ считается удачным, если значение целевого критерия χ_G^2 близко к 1.

Независимые случайные величины с распределением Пуассона. В случае если число регистрируемых фотонов при получении кривой затухания флуоресценции невелико, то распределение Пуассона для отдельных точек кривой уже не может аппроксимироваться с достаточной степенью точности распределением Гаусса. Это имеет место в частности при проведении экспериментов в одномолекулярной флуоресцентной спектроскопии, когда число испущенных исследуемой системой фотонов достаточно мало. Вместе с тем случайные величины в различных точках исходной статистической характеристики по-прежнему остаются независимыми. В данном случае функция $f(x_i, F_i^{th}(\vec{a}))$ имеет вид

$$f(x_i, F_i^{th}(\vec{a})) = \frac{[F_i^{th}(\vec{a})]^{x_i}}{x_i!} \exp(-F_i^{th}(\vec{a})). \quad (15.10)$$

Подставляя выражение (15.10) в выражение (15.6), получим

$$\begin{aligned} \chi_P^2(\vec{a}) &= -2 \sum_{i=1}^k \ln \left[\frac{\frac{[F_i^{th}(\vec{a})]^{x_i}}{x_i!} \exp(-F_i^{th}(\vec{a}))}{\frac{[x_i]^{x_i}}{x_i!} \exp(-x_i)} \right] = \\ &= -2 \sum_{i=1}^k \ln \left[\left(\frac{F_i^{th}(\vec{a})}{x_i} \right)^{x_i} \exp(x_i - F_i^{th}(\vec{a})) \right] = \\ &= 2 \sum_{i=1}^k x_i \ln \left[\frac{x_i}{F_i^{th}(\vec{a})} \right] - x_i + F_i^{th}(\vec{a}). \end{aligned} \quad (15.11)$$

Как и в случае Гауссовской статистики, описанной ранее, на практике при проведении анализа используют значение $\chi_P^2(\vec{a})$, нормированное на число степеней свободы $\nu = k - m - 1$. При достаточно хорошем совпадении исходной и теоретической статистических характеристик значение нормированного целевого критерия $\chi_P^2(\vec{a})$ будет близким к 1.

Мультиномиальное распределение. При использовании данной статистики отдельные точки экспериментальной кривой не могут считаться независимыми. Вследствие этого мультиномиальное распределение является многомерным, а следовательно, для получения выражения для целевого критерия в данном случае следует использовать более общее базовое выражение (15.4). Многомерная функция плотности вероятности для мультиномиального распределения имеет вид

$$f(x_1, x_2, \dots, x_k, F_i^{th}(\vec{a})) = N! N^N \prod_{i=1}^k \frac{[F_i^{th}(\vec{a})]^{x_i}}{x_i!}, \quad (15.12)$$

где $N = \sum_{i=1}^k x_i$ (в случае если в качестве исходной кривой рассматривается кривая затухания флуоресценции, то N представляет собой общее число зарегистрированных фотонов). Подставляя выражение (15.12) в формулу (15.4), в итоге получаем

$$\begin{aligned}\chi_M^2(\vec{a}) &= -2 \ln \left[\frac{N! N^N \prod_{i=1}^k \frac{[F_i^{th}(\vec{a})]^{x_i}}{x_i!}}{N! N^N \prod_{i=1}^k \frac{[x_i]^{x_i}}{x_i!}} \right] = -2 \ln \prod_{i=1}^k \left[\frac{F_i^{th}(\vec{a})}{x_i} \right]^{x_i} = \\ &= -2 \sum_{i=1}^k \ln \left[\frac{F_i^{th}(\vec{a})}{x_i} \right]^{x_i} = 2 \sum_{i=1}^k x_i \ln \left[\frac{x_i}{F_i^{th}(\vec{a})} \right].\end{aligned}\quad (15.13)$$

Величина $\chi_M^2(\vec{a})$, представленная выражением (15.13), может использоваться в качестве целевого критерия для процедур оптимизации. На практике, как и в двух предыдущих случаях, рассмотренных ранее, используется нормированное значение целевого критерия (15.13) $\chi_M^2(\vec{a})/(k - m - 1)$, которое при хорошем совпадении исходной и теоретической кривых должно стремиться к 1.

При использовании на практике целевого критерия (15.13) в процессе оптимизации требуется проведение дополнительной нормировки вычисляемой теоретической кривой. Данная нормировка должна обеспечивать выполнение следующего условия:

$$N = \sum_{i=1}^k F_i^{th}(\vec{a}). \quad (15.14)$$

Смысл нормировки (15.14) состоит в том, что площади под графиками исходной и соответствующей ей теоретической статистических характеристик должны быть всегда одинаковыми в процессе проведения оптимизации.

При анализе статистических характеристик с соответствующими распределениями значений в отдельных точках пространства аргументов, полученные выше целевые критерии могут использоваться совместно с различными методами оптимизации для получения оценок неизвестных параметров.

Методы оптимизации. Методы оптимизации играют достаточно важную роль в процессе проведения анализа статистических характеристик исследуемого потока событий. Используя значения целевого критерия в различных точках пространства оцениваемых параметров и стремясь к нахождению его экстремума, метод оптимизации, таким образом, стремится обеспечить максимальное сходство между исходной и соответствующей ей теоретической статистическими характеристиками. Для поиска экстремума целевого критерия каждый метод оптимизации использует уникальный математический алгоритм для изменения значений

оцениваемых параметров. В зависимости от характера данного алгоритма все методы оптимизации могут быть разделены на две большие группы:

- неградиентные методы, алгоритмы оптимизации которых не требуют вычисления производных от целевого критерия;
- градиентные методы, которые основаны на алгоритмах, требующих вычисления производных от функции целевого критерия.

Для каждой из представленных выше групп рассмотрим по одному методу оптимизации, наиболее широко используемому на практике.

Симплекс-метод. Данный метод является одним из наиболее популярных неградиентных методов, который может использоваться для анализа статистических характеристик случайных потоков событий. В основе симплекс-метода лежит алгоритм оптимизации, в соответствии с которым выполняется эмпирический поиск экстремума целевого критерия в пространстве параметров. Поисковый алгоритм данного метода может адаптироваться определенным образом к изменениям поведения функции целевого критерия в различных областях пространства параметров. Метод достаточно прост в понимании, не требует вычисления производных от функции целевого критерия и может достаточно эффективно применяться для получения оценок небольшого числа неизвестных параметров.

С момента, когда симплекс-метод был представлен впервые в работе Спэндли, было предложено множество модификаций, призванных улучшить его эффективность. Наиболее важным из предложенных усовершенствований стало введение Нелдером и Мидом симплекса, способного изменять свой размер в процессе проведения оптимизации. Это позволило значительно сократить число шагов, которое требуется симплекс-методу для поиска экстремума целевого критерия. Ниже будет рассмотрена версия метода, базирующаяся на алгоритме Нелдера – Мида.

Пусть задача состоит в поиске минимума одного из представленных ранее целевых критериев $\chi^2(a_1, a_2, \dots, a_m)$, где $a_1, a_2, \dots, a_m$ представляют собой m параметров, значения которых могут изменяться для достижения минимума χ^2 . Поисковый алгоритм, который используется симплекс-методом, состоит из повторяющихся шагов поиска, включающих построение в пространстве параметров многоугольника, называемого симплексом. В m -мерном пространстве симплекс представляет собой фигуру, состоящую из $m + 1$ вершины. На плоскости это треугольник, а в трехмерном пространстве тетраэдр (рис. 15.1).

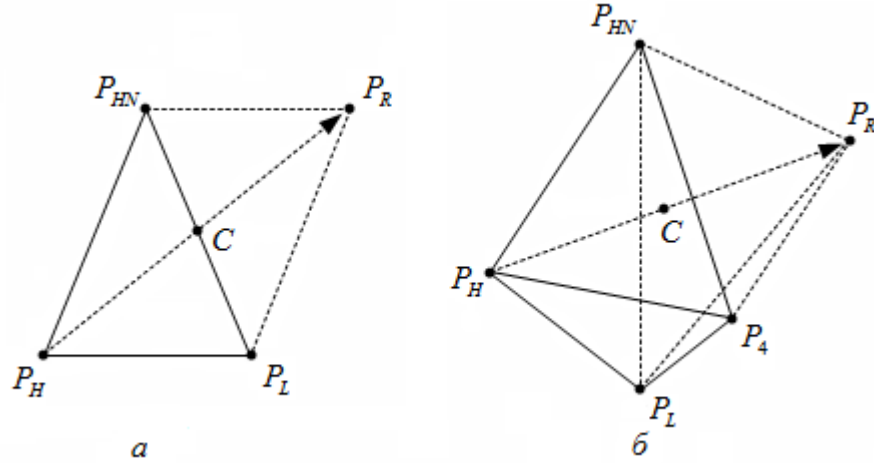


Рис. 15.1. Форма симплекса в 2-мерном (а) и 3-мерном (б) пространстве

Блок-схема симплекс-метода представлена на рис. 15.2. Точки, задающие $m+1$ вершину симплекса в m -мерном пространстве, определяются m -мерными векторами, компоненты которых строятся на базе выбранного набора начальных приближений для оцениваемых параметров $a_1, a_2, \dots, a_m$ следующим образом:

$$\begin{aligned}
 P_1 &= P(a_1, a_2, \dots, a_m); \\
 P_2 &= P(Ka_1, a_2, \dots, a_m); \\
 P_3 &= P(a_1, Ka_2, \dots, a_m); \\
 &\vdots \\
 P_{m+1} &= P(a_1, a_2, \dots, Ka_m).
 \end{aligned}
 \tag{15.15}$$

Коэффициент K , который фигурирует в выражении (15.15), позволяет разнести различные вершины симплекса в пространстве параметров на определенное расстояние. Поскольку вершины симплекса изначально могут быть выбраны относительно далеко от области глобального минимума целевого критерия, то для предотвращения попадания в локальный минимум предпочтительно строить исходный симплекс с вершинами, разнесенными на достаточно большое расстояние друг от друга. Конкретный выбор значения коэффициента K будет зависеть от конкретной решаемой задачи.

После определения координат вершин симплекса для каждой из них вычисляется значение целевого критерия χ^2 :

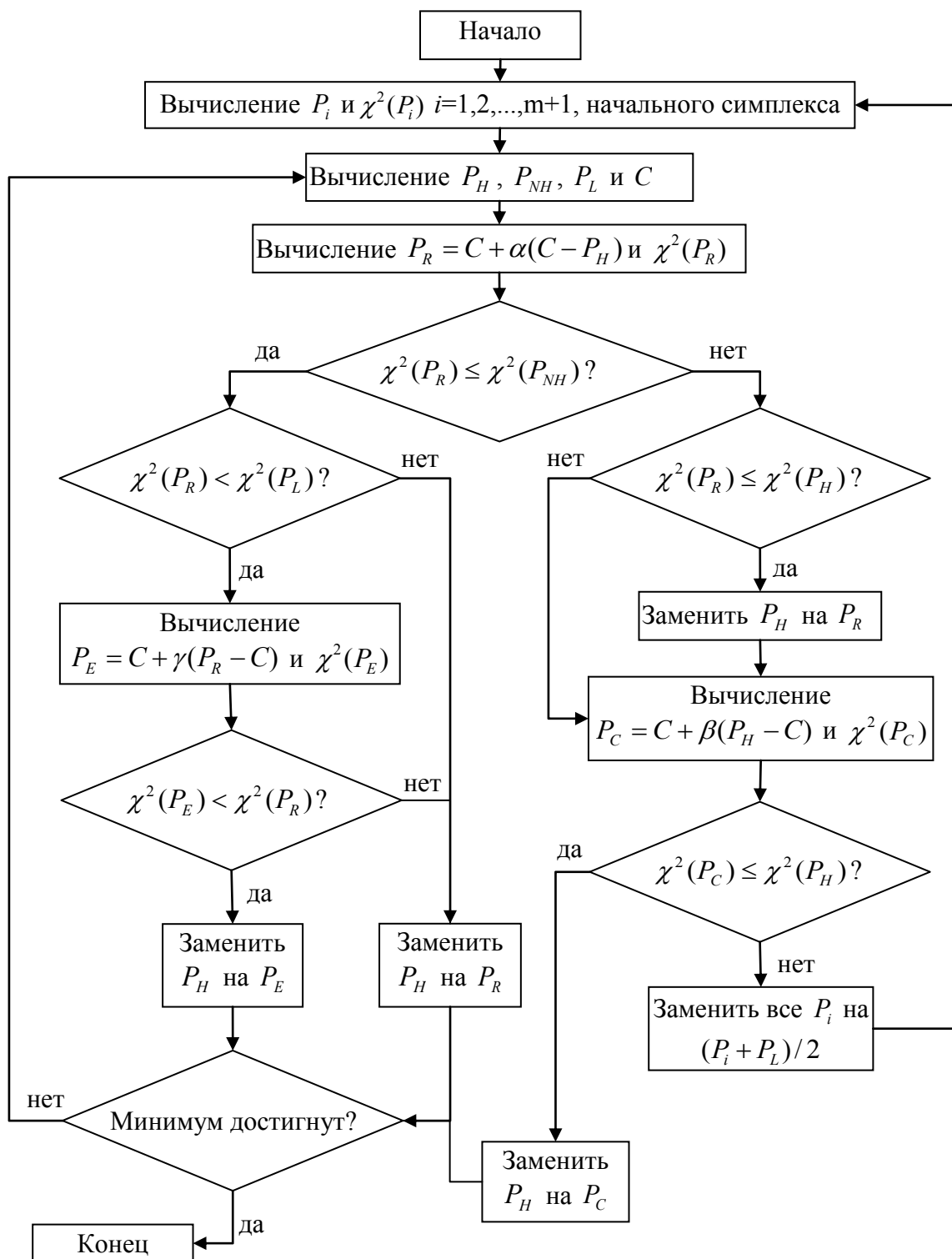


Рис. 15.2. Блок-схема симплекс-метода

$$\chi^2(P_i), i = 1, \dots, m + 1. \quad (15.16)$$

В дальнейшем будем обозначать вершину симплекса, которой соответствует наибольшее значение χ^2 , как P_H , вершину с наибольшим значением χ^2 без учета P_H , как P_{HN} и вершину с наименьшим значением χ^2 , как P_L .

После построения начального симплекса и определения вершин P_H , P_{HN} и P_L выполняется перемещение симплекса в пространстве параметров и одновременное изменение его размера с целью поиска минимума целевого критерия χ^2 . Стратегия данного перемещения основана на выполнении ряда операций, изложенных далее.

Логично предположить, что минимум целевого критерия будет находиться в области пространства параметров, противоположной точке P_H относительно всех других точек симплекса. По этой причине для определения предполагаемого направления на минимум χ^2 строится *центр* C для всех точек симплекса за исключением P_H , по следующей формуле:

$$C = \frac{1}{m} \sum_{i=1, i \neq H}^{m+1} P_i. \quad (15.17)$$

На плоскости, когда симплекс представлен треугольником, C является серединой отрезка, соединяющего P_{HN} и P_L (рис. 15.1). После определения C точка P_H отражается относительно C , в результате чего определяется *отраженная точка* P_R в соответствии со следующей формулой:

$$P_R = C + \alpha(C - P_H), \quad (15.18)$$

где α – коэффициент отражения. Нелдер и Мид предлагают использовать $\alpha = 1$, что ведет к расположению P_R и P_H на одинаковом расстоянии от C . Однако, в определенной ситуации, это может приводить к возникновению осцилляций, когда вышеуказанные точки будут переходить друг в друга. Для исключения таких ситуаций коэффициент отражения можно принимать несколько меньшим, например, 0,9985.

После получения отраженной точки вычисляется соответствующее ей значение целевого критерия. При этом полученное значение $\chi^2(P_R)$ может по-разному соотноситься со значениями целевого критерия в других точках симплекса, что определяет дальнейший ход алгоритма. Можно выделить три основных случая, которые будут рассмотрены далее.

Во-первых, *отражение может быть умеренно успешным*. В данном случае выполняется следующее условие:

$$\chi^2(P_L) \leq \chi^2(P_R) \leq \chi^2(P_{NH}). \quad (15.19)$$

Выполнение (15.19) означает, что величина целевого критерия в точке P_R лучше, чем в P_{HN} , но хуже, чем в P_L . В данном случае формируется новый симплекс путем замены P_H на P_R . Затем процесс построения и тестирования отраженной точки повторяется с использованием вновь полученного симплекса.

Отраженная точка считается очень успешной, если выполняется условие

$$\chi^2(P_R) < \chi^2(P_L). \quad (15.20)$$

В данном случае считается, что симплекс перемещается в наиболее перспективном направлении и поиск в выбранном направлении необходимо расширить. По этой причине формируется точка расширения P_E в соответствии со следующей формулой:

$$P_E = C + \gamma(P_R - C), \quad (15.21)$$

где γ представляет собой коэффициент расширения. Нелдер и Мид предлагают использовать значение γ , которое размещает P_E по линии $P_H - C$, на расстоянии в два раза дальше от C , чем расположена точка P_R . Для обеспечения стабильности работы алгоритма γ выбирается немного отличным от 2 (например, 1,95). Операция расширения считается успешной, если выполняется условие

$$\chi^2(P_E) < \chi^2(P_R). \quad (15.22)$$

Тогда P_H заменяется точкой расширения P_E . В противном случае P_H заменяется точкой P_R . После выполнения одной из указанных выше замен строится новый симплекс и вновь ищется отраженная точка.

В случае если значение целевого критерия для отраженной точки превосходит значение критерия в точке P_{HN} , т. е. выполняется неравенство вида

$$\chi^2(P_R) > \chi^2(P_{HN}), \quad (15.23)$$

то это означает, что симплекс расположен в области экстремума целевого критерия. В данном случае необходимо построить новый симплекс, который будет покрывать область экстремума целевого критерия. Для этого необходимо сравнить значения целевого критерия в точках P_H и P_R .

В случае если операция отражения оказывается неудачной, т. е.

$$\chi^2(P_R) > \chi^2(P_H), \quad (15.24)$$

то экстремум целевого критерия, скорее всего, расположен с той же стороны от центроида, что и точка P_H . В противном случае искомый экстремум, вероятнее всего, расположен в области, охватываемой симплексом, у которого точка P_H заменена на P_R .

Вне зависимости от того, была ли в данной ситуации наихудшая точка заменена отраженной или нет, полученный в итоге симплекс с высокой степенью вероятности покрывает область экстремума целевого критерия. По этой причине делается попытка сократить размеры симплекса. Для этого рассчитывается точка сжатия P_C в соответствии со следующим выражением:

$$P_C = C + \beta(P_H - C), \quad (15.25)$$

где β – коэффициент сжатия. Для $\beta = 0,5$, P_C будет расположена ровно на середине отрезка, соединяющего точки P_H и C . Так же как и ранее, для исключения неустойчивости в работе алгоритма, обычно коэффициент β выбирают несколько отличным от 0,5 (например, 0,4985). В случае если сжатие оказалось успешным, т. е.

$$\chi^2(P_C) \leq \chi^2(P_H), \quad (15.26)$$

то P_H заменяется на P_C , и далее процедура поиска повторяется нахождением новой отраженной точки. В противном случае каждая точка симплекса переопределяется в соответствии со следующим выражением:

$$P_i = (P_i + P_L) / 2, \quad i = 1, \dots, m + 1. \quad (15.27)$$

В данном случае, как следует из (15.27), изменяются все точки, кроме P_L . После построения нового симплекса переходят к очередной итерации алгоритма.

Таким образом, рассмотренный метод после построения исходного симплекса определяет направление в пространстве оцениваемых параметров, которое является наиболее перспективным для достижения экстремума целевого критерия, и перемещается вдоль него. При обнаружении вдоль выбранного направления области с худшими значениями целевого критерия, симплекс-метод либо меняет направление движения, либо, если обнаружена область экстремума, сжимается.

Критерием для окончания оптимизации может служить следующее правило:

$$\sqrt{\frac{1}{m} \sum_{i=1; i \neq H}^{m+1} (\chi^2(P_i) - \chi^2(C))^2} < \varepsilon, \quad (15.28)$$

что определяет некоторую величину, имеющую смысл стандартной ошибки в значениях целевого критерия. Стоит отметить, что поскольку выражение (15.28) использует только информацию о значениях целевого критерия в точках симплекса, то при достаточно пологом поведении критерия в области экстремума значения отдельных параметров при выполнении (15.28) могут значительно отличаться от оптимальных. Однако в данном случае большего уточнения и не требуется, поскольку описанное выше поведение целевого критерия предполагает слабую зависимость от соответствующего параметра, которая лишь указывает на плохую определенность последнего.

Кроме выражения (15.28) для завершения работы алгоритма симплекс-метода также могут применяться критерии, основанные на использовании информации о значениях координат отдельных точек симплекса. Такие критерии могут учитывать расстояние между наихудшей и наилучшей точками или относительные расстояния между вершинами симплекса и центроидом.

На практике симплекс-метод является достаточно эффективным при небольшом числе оцениваемых параметров. При росте числа оцениваемых параметров и как следствие усложнении поверхности целевого критерия в пространстве параметров возрастает вероятность попадания симплекс-метода в локальный минимум.

Л е к ц и я 16

АЛГОРИТМЫ ДЛЯ АНАЛИЗА ДАННЫХ (Часть 2)

Метод Марквардта. Данный метод является одним из наиболее широко используемых при решении задач оптимизации градиентных методов. Алгоритм метода Марквардта основан на широко известном из математического анализа факте, что в точке экстремума функции многих переменных все ее частные производные первого порядка равны 0. Для случая поиска минимума целевого критерия $\chi^2(a_1, a_2, \dots, a_m)$ данное правило приобретает следующий вид:

$$\frac{\partial \chi^2}{\partial a_j} = 0, \quad j = 1 \dots m, \quad (16.1)$$

то есть частные производные по искомым параметрам должны быть равны 0. Разложим функции (16.1) в ряд Тейлора по параметрам $a_1, a_2, \dots, a_m$:

$$\frac{\partial \chi^2}{\partial a_j} = \frac{\partial \chi^2}{\partial a_j} \Big|_{a_l=a_l^0} + \sum_{l=1}^m \frac{\partial^2 \chi^2}{\partial a_l \partial a_j} \Big|_{a_l=a_l^0} \Delta a_l + \dots, \quad (16.2)$$

где $a_1^0, a_2^0, \dots, a_m^0$ определяют точку, которая является очередным приближением в пространстве параметров, вокруг которой выполняется разложение. Ограничим ряд в правой части (16.2) первыми двумя слагаемыми и подставим при этом условия (16.2) в (16.1). В результате получим

$$\frac{\partial \chi^2}{\partial a_j} \Big|_{a_l=a_l^0} = - \sum_{l=1}^m \frac{\partial^2 \chi^2}{\partial a_l \partial a_j} \Big|_{a_l=a_l^0} \Delta a_l, \quad (16.3)$$

или в матричном представлении

$$AX = B, \quad (16.4)$$

где

$$A_{lj} = - \frac{\partial^2 \chi^2}{\partial a_l \partial a_j} \Big|_{a_l=a_l^0}; \quad X_l = \Delta a_l; \quad B_j = \frac{\partial \chi^2}{\partial a_j} \Big|_{a_l=a_l^0}.$$

Выражение (16.3) является базовым для построения алгоритма метода Марквардта (рис. 16.1), предназначенного для поиска минимума целевого критерия. Данное выражение представляет собой систему из m линейных уравнений относительно приращений Δa_l . Если задать набор начальных приближений для оцениваемых параметров, то после вычисления производных в начальной точке система уравнений (16.3) может быть решена, например, численными методами. Полученные в результате значения приращений параметров используются для вычисления следующего набора приближений по формуле

$$a_j^{n+1} = a_j^n + \Delta a_j^n, \quad (16.5)$$

где n – номер итерации. Фактически, поскольку выражение (16.3) получено исходя из условия минимизации целевого критерия, то, если бы (16.2) было точным, получаемые приращения определяли бы смещения параметров, необходимые для достижения минимума целевого критерия.

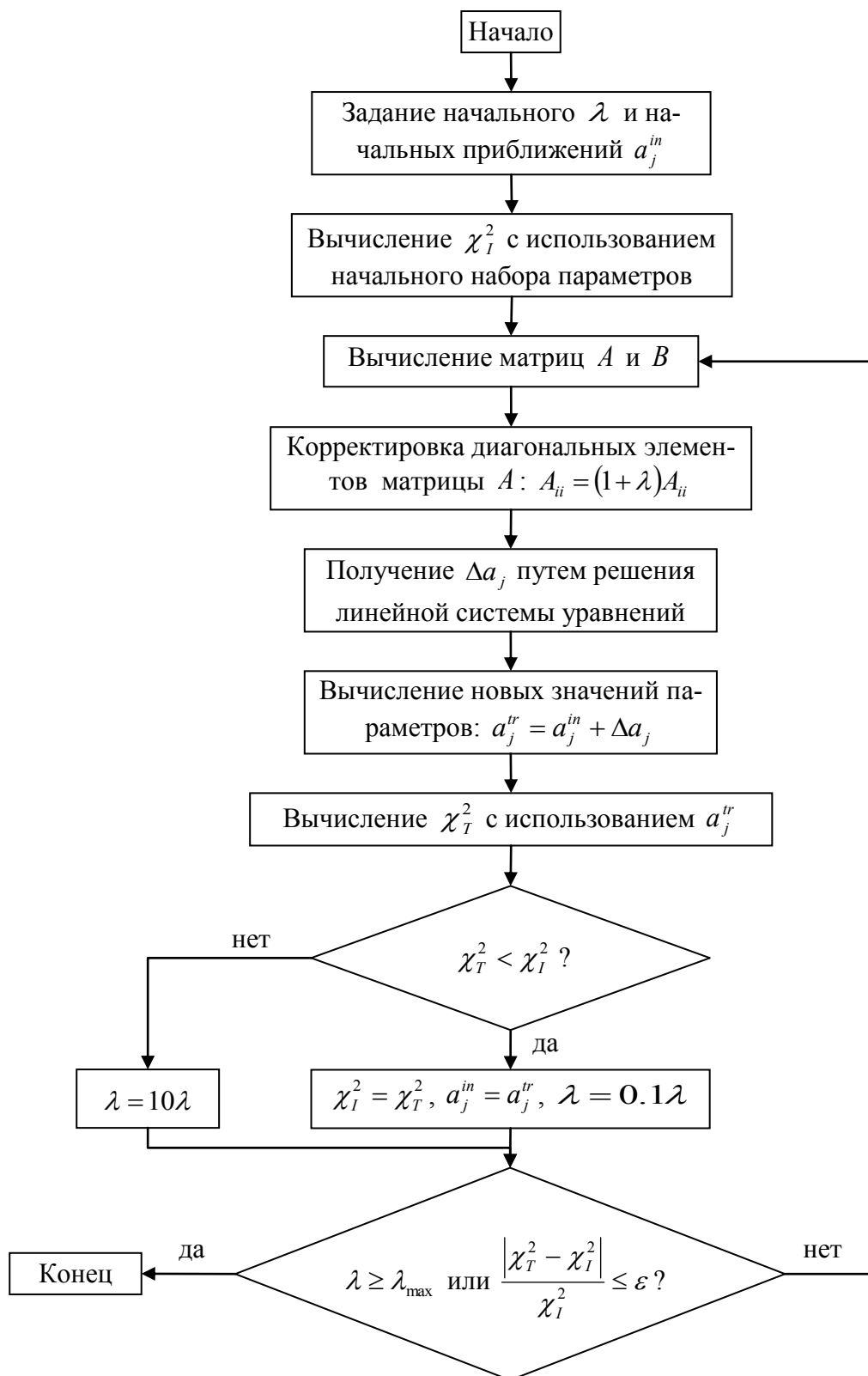


Рис. 16.1. Блок-схема метода Марквардта

Метод Марквардта использует для нахождения приращений параметров выражение, которое базируется на (16.3), однако имеет некоторое отличие в правой части. Данное выражение имеет вид

$$\left. \frac{\partial \chi^2}{\partial a_j} \right|_{a_l = a_l^0} = - \sum_{l=1}^m (1 + \lambda_{lj}) \left. \frac{\partial^2 \chi^2}{\partial a_l \partial a_j} \right|_{a_l = a_l^0} \Delta a_l; \quad \lambda_{lj} = \begin{cases} \lambda, & l = j \\ 0, & l \neq j \end{cases} \quad (16.6)$$

где величина λ является параметром, который определяет поведение метода оптимизации. Если значение данного параметра относительно велико, что обычно имеет место вначале процесса оптимизации, то недиагональные элементы будут менее значимыми по сравнению с диагональными. В этом случае поведение рассматриваемого метода будет аналогично поведению градиентных методов, которые основаны на том факте, что направление на минимум целевого критерия прямо противоположно градиенту целевого критерия в пространстве параметров. Это позволяет достаточно быстро достигать области минимума целевого критерия в случае, если начальные приближения были выбраны достаточно далеко от оптимальной точки. По мере приближения к минимуму значение параметра λ уменьшается и при достаточно малых его значениях уравнение (16.6) становится близким к уравнению (16.3), которое позволяет получать достаточно хорошие значения приращений параметров в некоторой малой области около оптимального их набора. Причиной этому является то, что ошибка, внесенная при ограничении числа слагаемых в ряде Тейлора, в области минимума χ^2 значительно уменьшается и выражение (16.3) является достаточно точным. Таким образом, рассматриваемый метод регулирует параметр λ с целью обеспечения уменьшения целевого критерия на каждой итерации. Если значения χ^2 уменьшаются, то λ также уменьшается. Если же на некотором шаге значение χ^2 в новой точке, полученной в соответствии с (16.5) увеличилось, то метод наращивает значение λ до тех пор, пока не будет достигнуто снижение χ^2 на очередном шаге.

В итоге алгоритм метода включает следующие шаги.

Шаг 1. Устанавливается начальное значение λ (например, $\lambda = 0,001$) и задаются начальные приближения для оцениваемых параметров $a_j = a_j^{in}$, $j = 1, \dots, m$.

Шаг 2. Для заданных начальных приближений вычисляется значение целевого критерия χ_I^2 .

Шаг 3. Рассчитываются матрицы производных A и B (выражения (16.3) и (16.4)).

Шаг 4. Модифицируются диагональные элементы матрицы A в соответствии с (16.6).

Шаг 5. Решается система уравнений (16.4), откуда находят значения приращений параметров Δa_j .

Шаг 6. Вычисляется очередной пробный набор параметров $a_j^{tr}, j=1, \dots, m$ в соответствии с (16.5), а для полученного набора параметров вычисляется значение целевого критерия χ_T^2 .

Шаг 7. Если $\chi_T^2 < \chi_I^2$, то пробный набор параметров принимается в качестве исходного, т. е. $a_j^{in} = a_j^{tr}$ и $\chi_I^2 = \chi_T^2$, уменьшается значение параметра $\lambda = 0,1\lambda$ и далее, если критерий выхода не удовлетворен, выполняется переход к шагу 3. В противном случае, если $\chi_T^2 \geq \chi_I^2$ увеличивается значение параметра $\lambda = 10\lambda$ и, если критерий выхода не удовлетворен, выполняется переход к шагу 4.

В качестве критериев выхода из процесса оптимизации могут использоваться следующие соотношения:

1. В случае если метод не в состоянии найти направление на минимум целевого критерия и $\lambda \geq \lambda_{\max}$, где $\lambda_{\max}$ достаточно велико (например, $\sim 10^{10}$).
2. Если область минимума целевого критерия была достигнута, то для окончания процесса оптимизации можно использовать следующее условие:

$$\frac{|\chi_T^2 - \chi_I^2|}{\chi_I^2} \leq \varepsilon,$$

где ε – некоторая заранее заданная малая величина, которая сравнивается с относительным изменением целевого критерия на двух последовательных итерациях.

Алгоритм (16.6) метода Марквардта приобретает окончательный вид в случае, если известно выражение для целевого критерия.

Пример 16.1. Пусть необходимо решить задачу минимизации целевого критерия, полученного при гауссовском распределении шума в отдельных точках анализируемой статистической характеристики. Тогда, как было показано ранее, целевой критерий имеет вид

$$\chi_G^2(a_1, a_2, \dots, a_m) = \frac{1}{k-m-1} \sum_{i=1}^k \frac{(x_i - F_i^{th}(a_1, a_2, \dots, a_m))^2}{\sigma_i^2}. \quad (16.7)$$

Для выражения (16.7) найдем производные, входящие в (16.3):

$$\begin{aligned} \frac{\partial \chi^2}{\partial a_j} &= -\frac{2}{k-m-1} \sum_{i=1}^k \frac{(x_i - F_i^{th})}{\sigma_i^2} \frac{\partial F_i^{th}}{\partial a_j}, \\ \frac{\partial^2 \chi^2}{\partial a_l \partial a_j} &= \frac{2}{k-m-1} \left[\sum_{i=1}^k \frac{1}{\sigma_i^2} \frac{\partial F_i^{th}}{\partial a_l} \frac{\partial F_i^{th}}{\partial a_j} - \sum_{i=1}^k \frac{(x_i - F_i^{th})}{\sigma_i^2} \frac{\partial^2 F_i^{th}}{\partial a_l \partial a_j} \right]. \end{aligned}$$

Второй элемент, расположенный в квадратных скобках, в выражении для второй производной при увеличении k стремится к 0 гораздо быстрее, чем первый. По этой причине им можно пренебречь. С учетом этого подставим полученные выражения для производных в (16.6). Тогда, сокращая одинаковые константы, получаем

$$\sum_{l=1}^m (1 + \lambda_{lj}) \left[\sum_{i=1}^k \frac{1}{\sigma_i^2} \frac{\partial F_i^{th}}{\partial a_l} \frac{\partial F_i^{th}}{\partial a_j} \right] \Delta a_l = \sum_{i=1}^k \frac{(x_i - F_i^{th})}{\sigma_i^2} \frac{\partial F_i^{th}}{\partial a_j}.$$

Критерии оценки качества проведенного анализа. После проведения оптимизации и получения параметров модели, поставленной в соответствие исследуемой статистической характеристике, необходимо дать оценку качества проведенного анализа. На первый взгляд самым простым способом для оценки качества проведенного анализа может быть визуальное сравнение исходной статистической характеристики и соответствующей ей теоретической кривой. Однако такое сравнение позволяет обнаружить некачественно проведенный анализ лишь в тех случаях, когда две кривые, упомянутые выше, отличаются очень значительно. Если же кривые достаточно близки друг к другу и с учетом того, что исходная статистическая характеристика содержит статистический шум, незначительные систематические отклонения могут быть не обнаружены приведенным выше простым способом. В этом случае необходимо использовать ряд критериев оценки качества проведенного анализа, некоторые из которых будут рассмотрены ниже. Стоит также отметить, что судить о качестве полученных результатов и о том, являются ли полученные оценки параметров приемлемыми, лучше по результатам применения некоторой совокупности таких критериев.

В качестве первого критерия оценки качества проведенного анализа может выступать *значение целевого критерия*, использовавшегося при оптимизации, которое было получено после ее проведения. Значение целевого критерия, как упоминалось ранее, отражает степень подобия исходной и теоретической статистических характеристик. При адекватной нормировке полученное после хорошо проведенного анализа значение целевого критерия должно быть близким к 1.

Для оценки качества полученных результатов могут использоваться *дополнительные графические критерии*, которые оцениваются визуально, и дают более адекватную информацию о степени подобия исходной и теоретической статистических характеристик, чем их непосредственное визуальное сравнение.

Первым из таких критериев является *график взвешенных остатков*, построенный на области определения статистической характеристики. Значения взвешенных остатков в каждой точке показывают отличие исходной статистической характеристики от теоретической. Непосредственное выражение для расчета значений взвешенных остатков в различных точках пространства аргументов зависит напрямую от того, каков характер шума в различных точках исходной статистической характеристики. Например, в случае гауссовского распределения шума выражение для расчета взвешенных остатков имеет вид

$$R_i = \frac{x_i - F_i^{th}(a_1, a_2, \dots, a_m)}{\sigma_i}, i = 1, \dots, k, \quad (16.8)$$

где k – число точек в статистической характеристике, x_i , F_i^{th} и σ_i – соответственно значения исходной и теоретической статистических характеристик, а также среднеквадратическое отклонение шума в точке с порядковым номером i . Для успешно проведенного анализа значения взвешенных остатков распределены случайным образом около нулевой линии. Визуальная оценка графика взвешенных остатков эффективна при наличии достаточно выраженных неслучайных отклонений теоретической кривой от исходной.

В случае если полученные в результате неудачно проведенного анализа систематические отклонения между исходной и теоретической статистическими характеристиками достаточно малы, то часто они могут выражаться в неизбежных корреляциях взвешенных остатков. При этом обнаружение таких отклонений при помощи визуальной оценки графика взвешенных остатков может быть затруднительно. Для обнаружения вышеуказанных корреляций используется *график автокорреляционной функции взвешенных остатков*, который также строится в пределах области определения исходной статистической характеристики и оценивается визуально. При этом для расчета автокорреляционной функции используется следующая формула:

$$C_j = \frac{\frac{1}{k-j+1} \sum_{i=1}^{k-j+1} R_i R_{i+j-1}}{\frac{1}{k} \sum_{i=1}^k R_i^2}. \quad (16.9)$$

Верхняя граница для j обычно выбирается так, чтобы число слагаемых в числителе было достаточным для корректного усреднения (обычно $1 \leq j \leq k/2$). Как следует из выражения (16.9) $C_1 = 1$, что означает, что каждое значение взвешенных остатков идеально коррелирует с самим собой. Если значения исходной статистической характеристики случайно разбросаны вдоль кривой теоретической статистической характеристики, то при больших k C_j , $j \neq 1$ будут стремиться к 0. Поскольку на практике число точек в кривых конечно, то для хорошо проведенного анализа автокорреляционная функция должна демонстрировать частые осцилляции малой амплитуды около нулевой отметки.

Еще одним важным компонентом в системе оценки качества проведенного анализа является поиск *доверительных интервалов* для оцениваемых параметров, которые позволяют определить, с какой точностью были получены их значения. Введение доверительных интервалов обусловлено тем фактом, что получаемое в результате анализа значение некоторого параметра зависит от конкретной реализации шума в анализируемой статистической характеристике. Для различных реализаций шума получаемые значения оцениваемых параметров могут, вообще говоря, быть различными. По этой причине возникает вопрос об определении некоторой области, в пределах которой должно находиться искомое значение параметра.

Доверительный интервал представляет собой некоторую область вокруг минимума целевого критерия, в пределах которой значение целевого критерия не увеличивается значительно по сравнению с минимальным значением.

Для каждого из оцениваемых параметров в данной области существует некоторый интервал значений, в пределах которого с заданной доверительной вероятностью будет находиться полученное значение данного параметра.

Для поиска доверительных интервалов оцениваемых параметров могут использоваться различные методы, наиболее известными среди которых являются *метод стандартных ошибок*, *метод исчерпывающего поиска*, *метод Монте-Карло*. В качестве примера рассмотрим более подробно метод исчерпывающего поиска.

Метод исчерпывающего поиска. Данный метод основан на том факте, что значение целевого критерия, рассчитанного с использованием значения оцениваемого параметра, лежащего вне границ доверительного интервала, будет превосходить некоторый максимальный уровень.

Пусть есть некоторая статистическая характеристика, которая содержит нормальный шум. Пусть после анализа данной характеристики при помощи выбранной модели был получен набор значений параметров, которые соответствуют минимуму целевого критерия. В соответствии с рассматриваемым подходом, доверительные интервалы для параметров модели получаются последовательно для каждого параметра в отдельности.

Пусть из совокупности оцениваемых параметров выбран тот, для которого требуется найти доверительный интервал. Для нахождения границы доверительного интервала (например, правой) значение выбранного параметра меняется с некоторым шагом от оптимального значения в сторону искомой границы (например, вправо). При каждом получаемом при таком изменении значении, рассматриваемый параметр фиксируется (т. е. процедура оптимизации не может изменять его значение) и выполняется оптимизация остальных параметров, для которых в качестве начальных приближений выбираются найденные оптимальные значения. Полученное после выполнения такой оптимизации значение целевого критерия χ^2 будет определяться следующим выражением:

$$\chi^2 = \chi_{\min}^2 + \chi_{par}^2, \quad (16.10)$$

где $\chi_{\min}^2$ – значение целевого критерия при оптимальных значениях всех оцениваемых параметров, χ_{par}^2 – некоторое приращение, возникающее по причине наличия внесенной в один из параметров ошибки.

Для определения доверительного интервала рассмотрим отношение χ^2 и $\chi_{\min}^2$. В этом случае выражение (16.10) перепишется в виде

$$\frac{\chi^2}{\chi_{\min}^2} = 1 + \frac{\chi_{par}^2}{\chi_{\min}^2}. \quad (16.11)$$

Входящее в последнее выражение отношение $\chi_{par}^2 / \chi_{\min}^2$ может быть рассчитано с использованием F-статистики. С учетом этого, а также формулы (16.11), может быть получено выражение для предельного значения целевого критерия $\chi_{\lim}^2$, при превышении которого изменение целевого критерия может с определенной доверительной вероятностью считаться значимым:

$$\chi_{\text{lim}}^2 = \chi_{\text{min}}^2 \left(1 + \frac{m}{k - m - 1} F(m, k - m - 1, 1 - p) \right), \quad (16.12)$$

где m – число оцениваемых параметров, k – число точек в анализируемой статистической характеристике, p – доверительная вероятность (обычно 0,67, 0,95, 0,99 и т. д.), с которой определяется доверительный интервал, $F(m, k - m - 1, 1 - p)$ – F-статистика.

Таким образом, значение выбранного параметра изменяется до тех пор, пока получаемое при его фиксированном значении в процессе оптимизации значение χ^2 не станет равным χ_{lim}^2 , рассчитанному по формуле (16.12). Если значение параметра, удовлетворяющее вышеуказанному условию, найдено, то оно и будет являться границей доверительного интервала.

ЛИТЕРАТУРА

- Альянах, И. Н.* Моделирование вычислительных систем / И. Н. Альянах. – Ленинград, 1988. – 224 с.
- Апанасович, В. В.* Статистический анализ потоков в физическом эксперименте / В. В. Апанасович, А. А. Коляда, А. Ф. Чернявский. – Минск: Университетское, 1988. – 256 с.
- Апанасович, В. В.* Цифровое моделирование стохастических систем / В. В. Апанасович, О. М. Тихоненко. – Минск: Университетское, 1986. – 127 с.
- Бусленко, В. В.* Моделирование сложных систем / В. В. Бусленко. – М.: Наука, 1978. – 400 с.
- Быков, В. В.* Цифровое моделирование в статистической радиотехнике / В. В. Быков. – М.: Советское радио, 1971. – 328 с.
- Губанов, В. А.* Введение в системный анализ / В. А. Губанов, В. В. Захаров, А. Н. Коваленко. – Ленинград, 1988. – 232 с.
- Лоу, А. М.* Имитационное моделирование / А. М. Лоу, В. Д. Кельтон. – 3-е изд. – СПб.: Питер, 2004. – 848 с.
- Максимей, И. В.* Математическое моделирование больших систем / И. В. Максимей. – Минск, 1985. – 120 с.
- An experimental comparison of the Maximum Likelihood Estimation and Nonlinear Least-Squares Fluorescence Lifetime Analysis of Single Molecules / M. Maus [et al.] // *Anal. Chem.* – 2001. – Vol. 73. – P. 2078–2086.
- Demas, J. N.* Excited State Lifetime Measurements / J. N. Demas. – New York: Academic Press, 1983. – 273 p.
- Integrated data analysis in time-resolved fluorescence and fluorescence correlation spectroscopy / M. M. Yatskou [et al.] // *Recent Research Developments in Physical Chemistry.* – 2004. – Vol. 7, Part I. – P. 165–183.
- Johnson, M. L.* Parameter Estimation by Least-Squares Methods / M. L. Johnson, L. M. Faunt // *Methods in enzymology.* – 1992. – Vol. 210. – P. 1–37.

ОГЛАВЛЕНИЕ

1. ВВЕДЕНИЕ	3
Лекция 1. Моделирование как метод научного познания	3
2. МЕТОДЫ МОДЕЛИРОВАНИЯ СЛУЧАЙНЫХ ВЕЛИЧИН	19
Лекция 2. Моделирование базовой случайной величины	19
Лекция 3. Стандартные методы моделирования случайных величин	29
Лекция 4. Специальные методы моделирования случайных величин	40
3. МОДЕЛИРОВАНИЕ СЛУЧАЙНЫХ ПРОЦЕССОВ	48
Лекция 5. Описание случайных процессов и общие методы их моделирования	48
Лекция 6. Метод скользящего суммирования	59
Лекция 7. Метод рекуррентных разностных уравнений	66
Лекция 8. Моделирование типовых нормальных и негауссовских случайных процессов	73
4. МОДЕЛИРОВАНИЕ СЛУЧАЙНЫХ ПОТОКОВ	83
Лекция 9. Описание потоков случайных событий. Общие подходы к моделированию случайных потоков	83
Лекция 10. Моделирование типовых случайных потоков (Часть 1)	94
Лекция 11. Моделирование типовых случайных потоков (Часть 2)	107
Лекция 12. Потоки группированных точек. Статистический анализ генетически связанных потоков	117
Лекция 13. Преобразование случайных потоков. СМО как преобразователь случайных потоков	127
5. АНАЛИЗ СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК ПОТОКОВ СЛУЧАЙНЫХ СОБЫТИЙ	139
Лекция 14. Обобщенная схема анализа статистических характеристик потоков случайных событий	139
Лекция 15. Алгоритмы для анализа данных (Часть 1)	149
Лекция 16. Алгоритмы для анализа данных (Часть 2)	162
ЛИТЕРАТУРА	172

Учебное издание

Дигрис Анатолий Владимирович
ДИСКРЕТНО-СОБЫТИЙНОЕ
МОДЕЛИРОВАНИЕ
Курс лекций

В авторской редакции

Художник обложки *Т. Ю. Таран*
Технический редактор *Г. М. Романчук*
Корректор *Н. П. Ракицкая*

Ответственный за выпуск *Т. М. Турчиняк*

Электронный ресурс 5,78 Мб
Режим доступа: www.elib.bsu.by, ограниченный
Дата доступа:

Белорусский государственный университет.
ЛИ № 02330/0494425 от 08.04.2009.
Пр. Независимости, 4, 220030, Минск.