ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО РЫБОЛОВСТВУ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Калининградский государственный технический университет»

Балтийская государственная академия рыбопромыслового флота

И.П. Корнева

СЛУЧАЙНЫЕ ПРОЦЕССЫ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Учебное пособие
для студентов, обучающихся по направлению
26.03.01 «Управление водным транспортом
и гидрографическое обеспечение судоходства»,
профиль «Управление водными
и мультимодальными перевозками»
всех форм обучения

Калининград
Издательство БГАРФ
2021

УДК 519.21
К 67

Корнева, И.П. Случайные процессы и математическая статистика: учебное пособие для студентов, обучающихся по направлению 26.03.01«Управление водным транспортом и гидрографическое обеспечение судоходства», профиль «Управление водными и мультимодальными перевозками» всех форм обучения / И.П. Корнева; БГАРФ ФГБОУ ВО «КГТУ». – Калининград: Издательство БГАРФ, 2021. – 72 с. – Библиогр.: с. 72. – ISBN 978-5-7481-0479-1. – Текст: непосредственный

ISBN 978-5-7481-0479-1

В учебном пособии даны основные понятия случайных процессов, их классификация, законы распределения и основные характеристики, рассмотрены основные типы случайных процессов и их практические приложения.

Содержание теоретического материала соответствует государственному образовательному стандарту по направлению 26.03.01 «Управление водным транспортом и гидрографическое обеспечение судоходства», профиль «Управление водными и мультимодальными перевозками». Задачи приведены с подробными решениями и демонстрационными примерами.

Пособие предназначено для студентов всех форм обучения.

Библиогр. – 10 назв.

Печатается по решению редакционно-издательского совета Балтийской государственной академии рыбопромыслового флота.

Рецензенты: **Авдеева Н.Н.,** кандидат педагогических наук,
доцент секции прикладной математики
кафедры высшей математики КГТУ;
Алексеева Е.Е., кандидат педагогических наук,
доцент Института образования БФУ им. И. Канта

ISBN 978-5-7481-0479-1 УДК 519.21
К 67

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ

Теория случайных процессов – математическая наука, изучающая случайные явления в динамике их развития. Математическими моделями для описания случайных явлений, развивающихся во времени, являются случайные процессы. При этом предполагается, что состояние в текущий момент времени есть случайная величина.

Подобно тому, как понятие случайной величины есть обобщение понятия случайного события, понятие случайной функции (случайного процесса) является обобщением понятия случайной величины, так как значение случайной функции при определенном значении ее аргумента является случайной величиной.

Продолжительное время все практические задачи, возникающие при исследовании случайных явлений, полностью укладывались в схему случайных величин. Но после открытия броуновского движения в 1827 году для исследования аналогичных процессов потребовался математический аппарат для качественного анализа не случайных величин, а случайных функций или случайных процессов. Изучение таких взаимосвязей осложняется тем, что они не являются строгими и в зависимости от исхода опыта принимают различный функциональный вид, заранее неизвестно какой. Однако есть закономерности, присущие множеству значений, принимаемых случайной функцией как закономерности массового явления. Если для случайной функции в качестве аргумента рассматривать время, то получаем случайный процесс.

Математическая теория случайных процессов (а также более общих случайных функций произвольного аргумента) является важной главой вероятностей теории. Первые шаги по созданию теории случайных процессов относились к ситуациям, когда время t изменялось дискретно, а система могла иметь лишь конечное число разных состояний, то есть к схемам последовательности зависимых испытаний (Марков А.А. и др.).

Развитие теории случайных процессов, зависящих от непрерывно меняющегося времени, является заслугой советских математиков Е.Е. Слуцкого, А.Н. Колмогорова и А.Я. Хинчина, американских математиков Н. Винера, В. Феллера и Дж. Дуба, французского математика П. Леей, шведского математика Х. Крамера и др. Наиболее детально разработана теория некоторых специальных классов случайных процессов, в первую очередь – марковских процессов и стационарных случайных процессов, а также ряда подклассов и обобщений, указанных двух классов случайных процессов (цепи Маркова, ветвящиеся процессы, процессы с независимыми приращениями, мартингалы, процессы со стационарными приращениями и др.).

Особенно возросло число задач, требующих для решения привлечение аппарата теории случайных функций, в связи с развитием систем автоматического управления и регулирования.

При подготовке студентов, обучающихся по направлению 26.03.01 «Управление водным транспортом и гидрографическое обеспечение судоходства» по профилю «Управление водными и мультимодальными перевозками», большое внимание уделено методам построения и анализа стохастических моделей.

Дисциплина «Случайные процессы и математическая статистика» формирует математическую базу, позволяющую строить стохастические модели случайных процессов, протекающих в различных практических ситуациях, обеспечивает целенаправленную ориентацию в сфере будущей профессиональной деятельности будущего инженера водного транспорта. На базе изучения дисциплины осуществляется формирование навыков программирования, моделирования, освоения современных компьютерных технологий, их применение для разрешения социально-экономических проблем.

Цель данного учебного пособия – помочь студентам, овладевшим основами теории вероятностей, познакомиться с основными понятиями теории случайных процессов и математической статистики, овладеть методами решения задач, связанных с дискретными цепями Маркова, корреляционной теорией случайных процессов.

Для успешного усвоения дисциплины «Случайные процессы и математическая статистика» студент обязан знать основные разделы дисциплины «Математика» по теории вероятностей, математическому анализу, теории функций действительного и комплексного переменного, линейной и векторной алгебре. Модели и методы теории случайных процессов используются в следующих дисциплинах: «Моделирование транспортных процессов», «Эконометрика», «Основы системного анализа», «Проектирование и управление мультимодальными перевозками».

Излагаемый материал может служить основой для анализа управляемых систем массового обслуживания, которые описывают процессы функционирования широкого класса реальных экономических и технических систем – систем связи, систем снабжения, экономических систем (система офисов туристической фирмы, магазины, транспортные кассы и т. п.), транспортные системы и т. д. Кроме этого излагаемая теория может быть применена к анализу процесса управления состоянием технических систем на периоде их эксплуатации вне зависимости от конкретных задач, решаемых рассматриваемой системой.

Таким образом, теорию случайных процессов можно применить для построения большинства моделей материальных, финансовых и других потоков.

ГЛАВА 1. Основные понятия теории случайных процессов

1.1. Классификация процессов

На практике решение вопроса о случайности или детерминированности процесса основывается на способности воспроизвести процесс в ходе проводимого эксперимента. Если это приводит к одним и тем же результатам, то процесс считается детерминированным.

Рис. 1. Классификация процессов

Определение 1.1. *Детерминированные процессы* – это процессы, описываемые математическими формулами. Например, движение спутника на околоземной орбите, измерение температуры воды при нагревании и т. д. Такие процессы, как высота волн, напряжение в электросети, изменение численности жителей какого-либо города с течением времени не являются детерминированными, для их описания требуются *вероятностные понятия и статистические характеристики. Это процессы случайные*. К числу случайных процессов могут быть причислены и многие производственные процессы, сопровождающиеся случайными флуктуациями, а также ряд процессов, встречающихся в геофизике (например, вариации земного магнитного поля), физиологии (например, изменение биоэлектрических потенциалов мозга, регистрируемых на электроэнцефалограмме), экономике и логистике.

Следующий вид процессов – *хаотические.* С одной стороны, они детерминированные, но с очень сильной зависимостью от начальных условий, которые в реальности точно повторить нельзя, и поведение системы через некоторое время становится непредсказуемым. На выходе такие системы имеют случайные траектории, для изучения которых требуются вероятностные подходы.

1.2. Основные определения и примеры

Для возможности применения математических методов к изучению случайных процессов требуется, чтобы мгновенное состояние системы можно было схематически представить в виде точки некоторого фазового пространства (пространства состояний) R, при этом случайный процесс будет представляться функцией X (t) времени t со значениями из R.

Пусть имеются результаты наблюдений некоторого явления на бесконечном интервале времени. Эти наблюдения описать конкретной математической зависимостью нельзя, поскольку каждое наблюдение какого-либо процесса дает невоспроизводимый результат и можно указать только вероятность получения данного результата. Поэтому говорят, что результаты образуют случайную реализацию $\mathrm{x}(\mathrm{t})$, которая является лишь одной из бесконечного множества потенциально возможных. Совокупность всех реализаций образует случайный процесс $X(t) = \{x_k(t)\}$, где индекс $\boldsymbol{k}$ меняется от нуля до бесконечности и обозначает номер реализации (эксперимента).

Определение 1.2. *Случайный процесс (случайная функция)* – это семейство случайных величин $X(t), t \in T,$ заданных на одном и том же вероятностном пространстве и зависящих от параметра t, принадлежащего некоторому множеству *T*. Обычно t интерпретируется как время. Другими словами, это процесс, значение которого при любом фиксированном $t = t_i$ является случайной величиной $X(t_i)$.

Определение 1.3. *Реализацией случайного процесса* $X(t)$ называется неслучайная функция $x(t)$, в которую превращается случайный процесс $X(t)$ в результате опыта (рис. 2).

Случайный процесс можно рассматривать как совокупность всех возможных его реализаций. Бесконечное множество реализаций $\{x_k(t)\}$ называется *ансамблем.* Именно для ансамбля реализаций, регистрируемых на бесконечном интервале времени, вводятся вероятностные характеристики, определяющие свойства случайного процесса.

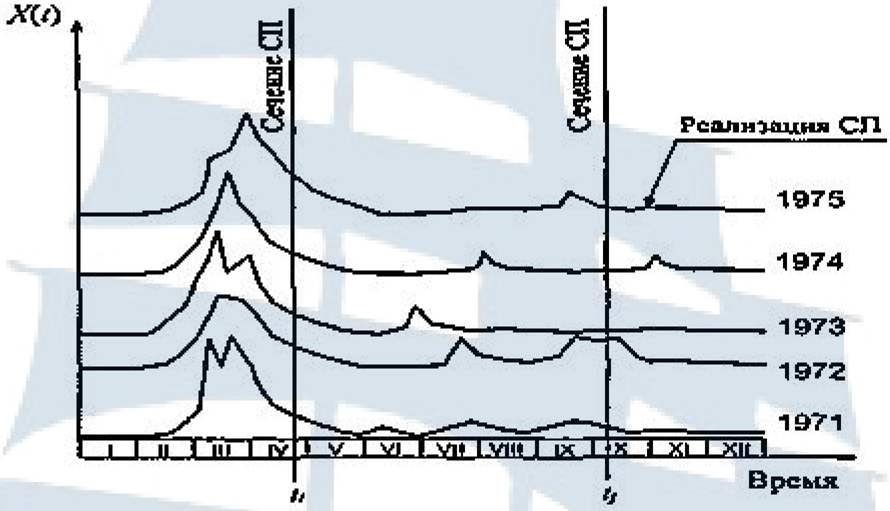


Рис. 2. Сечения и реализации случайного процесса

Это означает, что случайный процесс может быть описан с помощью количественных характеристик только при наличии большого числа идентичных испытаний.

Определение 1.4. *Сечение случайного процесса (случайной функции)* – это случайная величина $X(t_i)$ при $t = t_i$.

Результат исхода испытания $x_k(t_i)$ для любой реализации и в любой фиксированный момент времени t_i представляет собой значение случайной величины $X(t_i)$ (рис. 2).

Пример 1.1. (*точечные процессы*). На телефонную станцию поступают заявки на междугородние разговоры, и при этом фиксируется время поступления заявки. В таких процессах с появлением определенных событий в случайные моменты времени обычно полагают *x(t)*, равной числу заявок, поступивших за промежуток времени $[0,t]$. Эти процессы служат хорошими математическими моделями при проектировании систем обслуживания, при анализе транспортных потоков на магистралях; они используются в ядерной физике и т. д. Множество Т в данном случае – отрезок R^+ вида [0,Т] с возможным бесконечным значением Т. Пространство $X(t)$ значений случайного процесса при любом $t \in T$ совпадает с множеством неотрицательных целых чисел. Траектория имеет вид ступенчатой функции, возрастающей скачками в случайные моменты времени, и величина каждого скачка равна единице.

Пример 1.2. (*ветвящиеся процессы*). Наблюдается некоторая биологическая популяция, состоящая из особей, способных размножаться и гибнуть. Такие данные, как число потомков в определенном колене отдельной особи, численность популяции к фиксированному моменту времени *t*, количество погибших и новорожденных особей и т. п., составляют особый интерес для популяционной генетики, и трудно переоценить роль вероятностных моделей в изучении динамики развития биологической популяции. Аналогичные модели используются в физике элементарных частиц, особенно при изучении ядерных реакций. Пространства T и $X(t)$ те же, что и в первом примере 1.1, траектории также имеют вид ступенчатых функций, но величины скачков – произвольные целые числа.

Пример 1.3. (Броуновское движение в капилляре). Длинный тонкий капилляр наполняется жидкостью, и в середину капилляра помещается частица, диаметр которой ненамного меньше диаметра капилляра. Под действием молекул жидкости частица совершает хаотические движения, и для наблюдения за ними вводится система координат: капилляр рассматривается как действительная ось R с нулем в середине капилляра. В каждый момент времени *t* (непрерывно) регистрируется расстояние *x(t)* частицы от середины капилляра (очевидно, x (0) = 0) с учетом знака (минус – слева от середины, плюс – справа). Если изобразить теперь траекторию движения частицы на плоскости в координатах *(t, x(t))*, то мы получим то, что физики называют *траекторией одномерного броуновского движения*. Вероятностные модели, определяющие распределения таких процессов, были предложены Винером, Эйнштейном и Смолуховским. В этом примере: T – отрезок временной оси; $X(t)$ = R.

Пример 1.4. Рассмотрим случайный процесс, заданный аналитически формулой: $X(t) = Y \cdot sint,$ где $t \geq 0, Y \sim R[0,1]$ – случайная величина, имеющая равномерное распределение.

Решение. Если случайная величина Y = 1 в первом опыте, Y = 2 – во втором, Y = 1/2 – в третьем, то получим три реализации случайного процесса: $x_1(t) = sint;\ x_2(t) = 2sint;\ x_3(t) = \frac{1}{2}sint -$ неслучайные функции. Если зафиксировать момент времени, например $t = \frac{\pi}{6}$, то получим сечение процесса: $X(\pi/_6)$, то есть $X(\pi/_6) = \frac{1}{2}$Y – случайная величина.

Реализации случайного процесса и ансамбль реализаций являются множеством случайных величин. Поэтому вероятностное описание случайных величин служит одновременно вероятностным описанием случайного процесса, а статистические характеристики случайной величины совпадают с характеристиками случайного процесса.

Как правило, в распоряжении исследователя имеется одна реализация случайного процесса конечной длины $x(t)$, которая называется *выборочной функцией,* или просто *выборкой.* Если значения выборки берутся в дискретные моменты времени, то мы имеем дело с *временным* рядом. Вычисляемые по временному ряду (выборке) или по конечному набору временных рядов характеристики случайного процесса называются *выборочными* и служат лишь оценкой истинных характеристик случайного процесса.

1.3. Законы распределения случайных процессов

Значения случайных величин нельзя предугадать даже при полностью известных условиях эксперимента, в котором они измеряются. Возможно указать лишь вероятности того, что случайная величина примет то или иное значение или попадет в заданный интервал. Однако, зная *распределение вероятностей* случайных величин, возможно делать выводы о свойствах реализаций случайного процесса и об их характерных особенностях.

Сечение $x(t)$ случайного процесса $X(t)$ при любом фиксированном значении аргумента *t* представляет собой случайную величину, которая имеет закон распределения:

$$F_t(x) = P(X(t) < x). \tag{1.1}$$

Это *одномерный закон распределения случайного процесса* $X(t)$, который не является исчерпывающей характеристикой случайного процесса, так как характеризует свойства отдельно взятого сечения и не дает представления о совместном распределении двух и более сечений. На рис.3 показаны два случайных процесса с разными вероятностными структурами, но примерно одинаковыми распределениями случайных величин в каждом сечении.

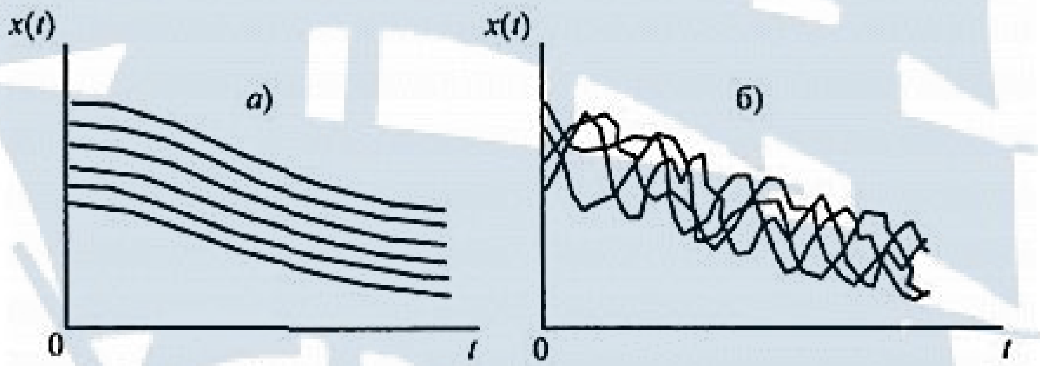


Рис. 3. Случайные процессы с разными вероятностными структурами

Производная $p(x)$ функции$F_t(x)$ называется *одномерной плотностью вероятности* (или просто плотностью вероятности).

Предположим, например, что большое число одинаковых генераторов шума было включено в некоторый момент времени в прошлом и с тех пор они работают до настоящего времени.

С выходом всей совокупности генераторов в данный момент времени связывают плотность вероятности $p(x)$, которая характеризует вероятность того, что в определенный момент времени выход *k*-го генератора сигналов x_k лежит в интервале между значениями *a* и *b*.

Плотность вероятности $p(x)$ процесса $X(t)$ можно также построить, рассматривая относительное время пребывания бесконечной реализации $x(t)$ одного генератора в интервале между значениями a и b. В интегральной форме связь вероятности с функцией $p(x)$ записывается следующим образом:

$$F_t(a < x_k \leq b) = \int_a^b \mathrm{p}(\mathrm{x})dx. \tag{1.2}$$

Плотность вероятности $p(x)$ задает скорость изменения вероятности в зависимости от значения реализации *x(t).* Функция $p(x)$ обычно оценивается путем вычисления вероятности того, что мгновенное значение отдельной реализации заключено в узком интервале, центр которого пробегает область значений процесса, с последующим делением на ширину интервала. Общая площадь, ограниченная графиком плотности вероятности, равна единице, что указывает на достоверность события, заключающегося в том, что значения реализации попадают в интервал от -∞ до ∞.

Одномерная плотность вероятности, помимо описания вероятностной структуры процесса, обычно применяется с целью проверки нормальности процесса, выявления нелинейностей и анализа экстремальных значений. Однако она не может полностью охарактеризовать свойства случайного процесса, поскольку не учитывает статистическую взаимосвязь (корреляцию) между значениями случайного процесса в различные моменты времени.

Для характеристики степени статистической зависимости вводится понятие *совместной плотности вероятности*, которая определяет вероятность того, что одновременно выполняются несколько условий. Например, для анализа статистической взаимосвязи между значениями случайного процесса, сдвинутыми на интервал времени τ, необходимо рассмотреть двумерную плотность вероятности процессов $X(t)$ и $Y(t) = X(t + \tau)$.

Эта функция определяет одновременное выполнение двух условий: случайная величина $\boldsymbol{X}$ принимает значение из интервала $x_1 < X \leq x_2$ и случайная величина $\boldsymbol{Y}$ лежит в интервале $y_1 < Y \leq y_2$. В интегральной форме это условие выглядит следующим образом:

$$\mathrm{F_t}(\mathrm{x_1} < X \leq \mathrm{x_2}, \mathrm{y_1} < Y \leq \mathrm{y_2}) = \int_{\mathrm{x_1}}^{\mathrm{x_2}} \int_{\mathrm{y_1}}^{\mathrm{y_2}} \mathrm{p}(\mathrm{x}, \mathrm{y}, \tau)\mathrm{dxdy}\,. \tag{1.3}$$

Аналогично строятся совместные плотности вероятности большей размерности – *многомерные плотности вероятности*, характеризующие статистическую взаимосвязь значений случайного процесса $X(t + \tau_i)$, сдвинутых на различные временные интервалы τ_i, где $i = 0, \dots, n$; n определяет размерность плотности вероятности; n-мерная плотность вероятности является наиболее полной характеристикой случайного процесса, которая учитывает корреляции значений случайного процесса в различные моменты времени. Чем больше n, тем лучше плотность вероятности описывает свойства случайного процесса. Поскольку комбинации временных сдвигов являются произвольными, то можно построить бесконечное число совместных плотностей вероятности различной размерности. Совокупность возможных

n-мерных плотностей вероятности образует иерархию, которая полностью определяет случайный процесс.

Однако многомерные законы распределения случайных процессов являются сравнительно громоздкими характеристиками и с ними крайне трудно оперировать на практике. Поэтому при изучении случайных процессов часто ограничиваются случаями, когда для описания случайного процесса достаточно знать только его одномерный или двумерный закон распределения.

Пример 1.5. Случайный процесс, который полностью характеризуется одномерной плотностью вероятности, так называемый *чистый случайный процесс*, или *белый шум*. Значения $X(t)$ в этом процессе, взятые в разные моменты времени t, совершенно независимы друг от друга, как бы близко ни были выбраны эти моменты времени. Это означает, что кривая белого шума содержит всплески, затухающие за бесконечно малые промежутки времени. Поскольку значения $X(t)$, например, в моменты времени t_1 и t_2 независимы, то вероятность совпадения событий, заключающихся в нахождении $X(t)$ между x_1 и $x_1 + dx_1$ в момент времени t_1 и между x_2 и $x_2 + dx_2$ в момент t_2, равна произведению вероятностей каждого из этих событий, поэтому

$$p_2(x_1, t_1; x_2, t_2) = p_1(x_1, t_1)p_1(x_2, t_2)$$

вообще для белого шума:

$$p_n(x_1, t_1; x_2, t_2; \dots; x_n, t_n) = p_1(x_1, t_1)p_1(x_2, t_2) \dots p_1(x_n, t_n),$$

то есть все плотности вероятности белого шума определяются из одномерной плотности вероятности.

Однако существует особый класс случайных процессов, впервые исследованных известным русским математиком А.А. Марковым (1856–1922) и называемых *марковскими случайными процессами*, для которых знание значения процесса в момент t_k уже содержит в себе всю информацию о будущем ходе процесса, какую только можно извлечь из поведения процесса до этого момента.

В случае марковского случайного процесса для определения вероятностных характеристик процесса в момент времени t_k достаточно знать вероятностные характеристики для любого одного предшествующего момента времени, то есть t_{k-1}.

Таким образом, марковские случайные процессы полностью характеризуются двумерной плотностью вероятности.

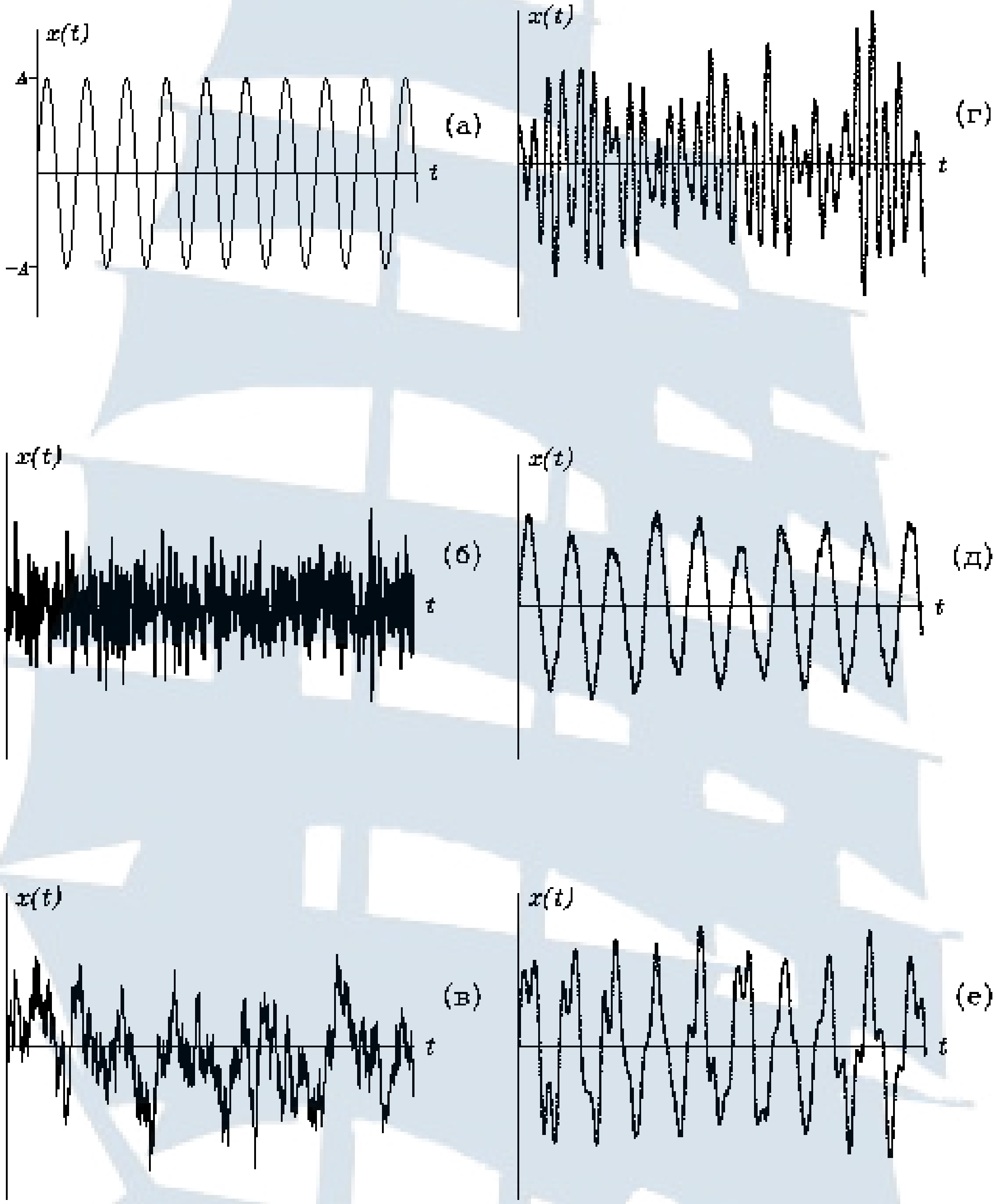


Рис. 4. Реализации случайных процессов

На рис. 4 представлены примеры реализаций случайных процессов: гармонический сигнал со случайной начальной фазой (а); белый шум (б); цветной шум (в); узкополосный шум (г); гармонический сигнал со случайной начальной фазой плюс цветной шум (д); гармонический сигнал со случайной начальной фазой плюс узкополосный шум (е).

Очевидно, что на основе выборок случайного процесса построить иерархию плотностей вероятности невозможно. Однако в некоторых случаях для того, чтобы охарактеризовать процесс, достаточно знать конечное число совместных плотностей вероятности. Это справедливо, например, для случайных процессов, состоящих из статистически независимых случайных величин, и для квазидетерминированных случайных процессов. Поэтому важно определить принадлежность случайного процесса к классу подобных случайных процессов. Среди вероятностных распределений существуют такие, свойства которых очень хорошо изучены и которые наиболее часто используются на практике. К ним относится, прежде всего, нормальное (гауссово) распределение, распределения Стьюдента и хи-квадрат, F-распределение Фишера, показательное, биномиальное и пуассоновское распределения.

Пример 1.6. На рис. 4 показаны типичные реализации некоторых случайных процессов, а также их комбинаций. На рис. 5 приведены плотности вероятности случайных процессов, реализации которых представлены на рис. 4, (а, б, в, г).

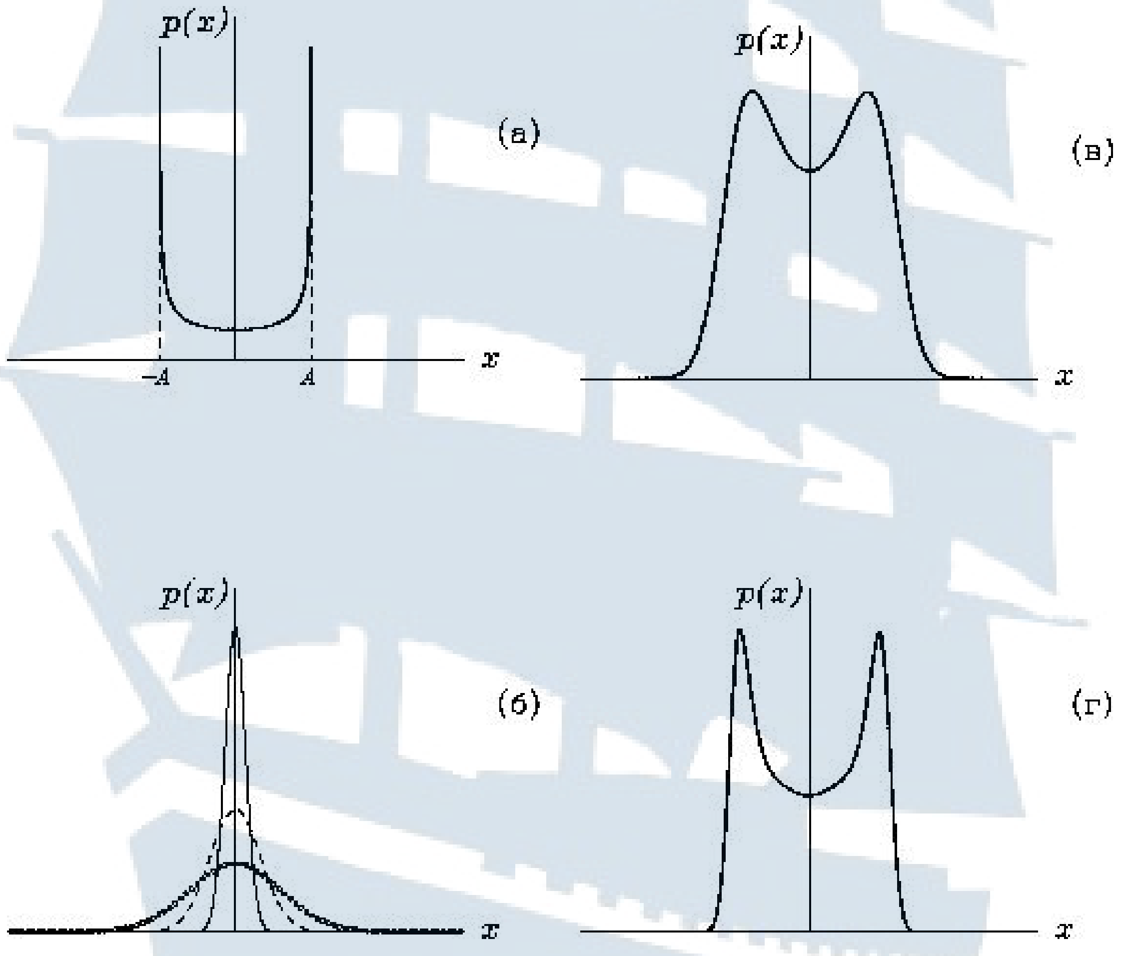


Рис. 5. Плотности вероятности гармонического сигнала: со случайной начальной фазой (а); белого шума (маркер «о» на рис. б); цветного шума (сплошная линия на рис. б); узкополосного шума (пунктирная линия на рис. б); суммы гармонического сигнала со случайной начальной фазой и цветного шума (в); суммы гармонического сигнала со случайной начальной фазой и узкополосного шума (г)

Таким образом, учитывая, что случайный процесс представляет собой совокупность всех сечений при различных значениях t, то для его полного описания надо рассматривать совместную функцию распределения сечений процесса:

$$F_{t_1,t_2,\dots,t_n}(x_1, x_2, \dots, x_n) = P\{X(t_1) < x_1,\ X(t_2) < x_2, \dots ,\ X(t_n) < x_n\}, \quad (1.4)$$

так называемый *конечномерный закон распределения* случайного процесса в моменты $t_1, t_2, \dots, t_n$, то есть рассматривать многомерную случайную величину $(X(t_1), X(t_2), \dots, X(t_n))$.

Если функция (1.4) достаточное число раз дифференцируема, то n-мерная совместная плотность вероятностей случайного процесса $X(t)$ имеет вид:

$$f_{t_1,t_2,\dots,t_n}(x_1, x_2, \dots, x_n) = \frac{\partial^n F_{t_1,t_2,\dots,t_n}(x_1,x_2,\dots,x_n)}{\partial x_1 \partial x_2 \cdots \partial x_n}. \quad (1.5)$$

Чем больше *n*, тем полнее описывает случайный процесс функция распределения (1.4) или плотность вероятностей (1.5). Эти функции учитывают связь хотя и между любыми, но лишь фиксированными сечениями этого процесса. Случайный процесс считается *заданным*, если задано множество всех его *n*-мерных законов распределения или *n*-мерных плотностей вероятностей для любых *n*.

Таким образом, понятие случайного процесса является обобщением понятия системы случайных величин, когда этих величин – бесконечное множество.

Понятие о функции распределения и плотности вероятности случайного процесса обычно используют при теоретических построениях и определениях. На практике вместо многомерных законов распределения используют более простые, хотя и менее полные числовые характеристики случайного процесса, аналогичные числовым характеристикам случайных величин, такие как математическое ожидание, дисперсия, начальные и центральные моменты, но только для случайного процесса они будут не числами, а функциями.

1.4. Основные характеристики случайных процессов

1.4.1. Математическое ожидание случайного процесса

Определение 1.5. Математическим ожиданием случайного процесса $X(t)$ или его средним значением называется неслучайная функция $m_X(t)$, которая при любом фиксированном значении аргумента t равна математическому ожиданию соответствующего сечения случайного процесса:

$$m_X(t) = MX(t). \quad (1.6)$$

Функция $m_X(t)$ характеризует поведение случайного процесса в среднем, представляет собой некоторую «среднюю» функцию, вокруг которой происходит разброс случайного процесса, то есть около которой группируются реализации этого процесса.

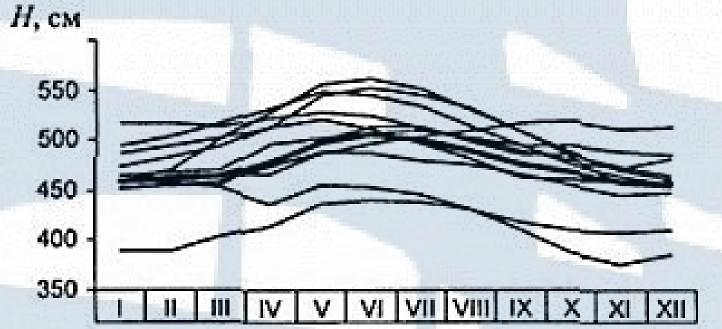


Рис. 6

Математическое ожидание случайного процесса в каждый фиксированный момент времени t_k равно математическому ожиданию соответствующего сечения случайного процесса $X(t_k)$ и представляет собой операцию вероятностного усреднения случайной величины $X(t_k)$. Математическое ожидание называют *средним значением случайного процесса* по множеству (средним по ансамблю, статистическим средним), поскольку оно представляет собой вероятностно усредненное значение бесконечного множества реализации случайного процесса.

Используя свойства математического ожидания случайной величины и учитывая, что $X(t)$ – случайный процесс, а $f(t)$ – неслучайная функция, получаем *свойства математического ожидания случайного процесса:*

1. Математическое ожидание неслучайной функции равно самой функции, то есть

$$M\big(f(t)\big) = f(t).$$

2. Неслучайный множитель можно выносить за знак математического ожидания случайного процесса, то есть

$$M(f(t) \cdot X(t)) = f(t) \cdot m_X(t).$$

3. Математическое ожидание суммы (разности) двух случайных процессов равно сумме (разности) математических ожиданий слагаемых, то есть

$$M(X(t) \pm Y(t)) = m_X(t) \pm m_Y(t).$$

Заметим, что если зафиксировать значение аргумента t и перейти от случайного процесса к его сечению, то можно найти математическое ожидание этого процесса. Например, если сечение случайного процесса при данном t – непрерывная случайная величина, то с плотностью $f_t(x)$ его математическое ожидание можно найти по формуле

$$MX(t) = \int_{-\infty}^{+\infty} x \cdot f_t(x)dx. \qquad (1.7)$$

Пример.1.7. Случайный процесс задан формулой

$$Y(t) = X \cdot e^{-t}, t > 0, X \sim N(3; 1),$$

то есть X – случайная величина, распределенная по нормальному закону с $a = 3, \sigma = 1$. Найти математическое ожидание случайного процесса *Y(t).*

Решение. Согласно свойству 2

$$m_Y(t) = M(X \cdot e^{-t}) = e^{-t} \cdot MX.$$

Так как $X \sim N(3; 1)$, то MX = 3, то $m_Y(t) = 3e^{-t}$.

1.4.2. Дисперсия случайного процесса

Для того чтобы учесть степень разбросанности реализации случайного процесса относительно его среднего значения, вводят понятие дисперсии случайного процесса.

Определение 1.6. *Дисперсией случайного процесса* $X(t)$ называется неслучайная функция $D_X(t)$, которая при каждом значении t равна дисперсии соответствующего сечения:

$$DX(t) = M(X(t) - m_x(t))^2. \qquad (1.8)$$

Или $D_X(t)$ можно определить по формуле

$$D_X(t) = DX(t) = MX^2(t) - m_X^2(t). \qquad (1.9)$$

Дисперсия $D_X(t)$ характеризует разброс (рассеяние) возможных значений (реализаций) случайного процесса относительно его математического ожидания.

Например, если сечение случайного процесса при данном t – непрерывная случайная величина с плотностью $f_t(x)$, то его дисперсию можно найти по формуле

$$DX(t) = \int_{-\infty}^{+\infty} (x - m_x(t))^2 \cdot f_t(x)dx. \qquad (1.10)$$

Определение 1.7. *Средним квадратическим отклонением* случайного процесса $X(t)$ называется неслучайная функция $\sigma_X(t)$, которая равна квадратному корню из дисперсии случайного процесса:

$$\sigma_X(t) = \sqrt{D_X(t)}. \qquad (1.11)$$

Свойства дисперсии случайного процесса:

1. Дисперсия неслучайной функции равна нулю, то есть $D\big(f(t)\big) = 0$.

2. Дисперсия случайного процесса неотрицательна, то есть $D_X(t) = \sigma_X^2(t) \geq 0$.

Если дисперсия случайного процесса равна 0, то все его реализации совпадают с математическим ожиданием $m_x(t)$, а сам случайный процесс превращается в обычную функцию действительной переменной t.

3. Дисперсия произведения неслучайной функции на случайную функцию равна произведению квадрата неслучайной функции на дисперсию случайной функции, то есть

$$D\big(f(t) \cdot X(t)\big) = f^2(t) \cdot D_X X(t).$$

4. Дисперсия суммы случайного процесса и неслучайной функции равна дисперсии случайного процесса, то есть

$$D(X(t) \pm f(t)) = D_X(t).$$

Пример 1.8. Используя условие примера 1.7, найти дисперсию и среднее квадратическое отклонение случайного процесса.

Решение. Находим дисперсию, используя ее свойство 3:

$$D_Y(t) = D(X \cdot e^{-t}) = (e^{-t})^2 \cdot DX = e^{-2t} \cdot DX.$$

Поскольку $X \sim N(3; 1)$, то $DX = \sigma_X^2(t) = 1$.

Поэтому $D_Y(t) = e^{-2t} \cdot DX = e^{-2t}$.

Тогда $\sigma_X(t) = \sqrt{D_X(t)} = \sqrt{e^{-2t}} = e^{-t}$.

Таким образом, математическое ожидание и дисперсия случайного процесса $X(t)$ зависят от одномерной плотности вероятностей и являются неслучайными функциями времени t.

1.4.3. Корреляционная функция случайного процесса

Математическое ожидание и дисперсия не являются исчерпывающими характеристиками случайного процесса, но они не дают достаточного представления о том, какой характер будут иметь отдельные реализации случайного процесса. Зная их, ничего нельзя сказать о зависимости двух и более сечений случайного процесса. Чтобы в какой-то мере охарактеризовать внутреннюю структуру случайного процесса, то есть учесть связь между значениями случайного процесса в различные моменты времени или, иными словами, учесть степень изменчивости случайного процесса, необходимо к перечисленным характеристикам добавить корреляционную функцию случайного процесса.

Определение 1.8. *Корреляционной функцией* случайного процесса $X(t)$ называется неслучайная функция $K_X(t_1, t_2)$, которая при каждой паре значений t_1 и t_2 равна корреляционному моменту соответствующих сечений $X(t_1)$ и $X(t_2)$:

$$K_X(t_1, t_2) = M((X(t_1) - m(t_1)) \cdot (X(t_2) - m(t_2)) \quad (1.12)$$

или

$$K_X(t_1, t_2) = M\big(X(t_1) \cdot X(t_2)\big) - m_X(t_1) \cdot m_X(t_2). \quad (1.13)$$

Функция $K_X(t_1, t_2)$ характеризует *степень связи* между ординатами случайного процесса $X(t)$ для двух моментов времени t_1 и t_2. При этом чем больше корреляционная функция, тем более плавными являются реализации случайного процесса $X(t)$ и наоборот.

Основные свойства корреляционной функции $K_X(t_1, t_2)$ случайного процесса:

1. Корреляционная функция при одинаковых значениях аргументов равна дисперсии случайного процесса, то есть $K_X(t, t) = D_X(t)$.

2. Симметричность: корреляционная функция не меняется при перестановке аргументов местами, то есть $K_X(t_1, t_2) = K_X(t_2, t_1)$.

Это свойство непосредственно вытекает из определения корреляционной функции.

3. Если к случайному процессу прибавить неслучайную функцию, то корреляционная функция не изменится, то есть если $Y(t) = X(t) + f(t)$, то $K_Y(t_1, t_2) = K_X(t_1, t_2)$.

Действительно, так как $m_Y(t) = m_X(t) + f(t)$, то

$$Y(t) - m_Y(t) = X(t) + f(t) - m_X(t) - f(t) = X(t) - m_X.$$

Следовательно, $K_Y(t_1, t_2) = K_X(t_2, t_1)$.

4. Модуль корреляционной функции не превосходит произведения среднеквадратических отклонений, то есть

$$|K_X(t_1, t_2)| \le \sqrt{D_X(t_1) \cdot D_X(t_2)} = \sigma_X(t_1) \cdot \sigma_X(t_2).$$

5. При умножении случайного процесса $X(t)$ на неслучайный множитель $f(t)$ его корреляционная функция умножится на произведение $f(t_1) \cdot f(t_2)$, то есть если $Y(t) = X(t) \cdot f(t)$, то

$$K_Y(t_1, t_2) = f(t_1) \cdot f(t_2) \cdot K_X(t_1, t_2).$$

Часто наряду с корреляционной функцией случайного процесса рассматривается нормированная корреляционная функция $r_X(t_2, t_1)$, определяемая равенством:

$$r_X(t_1, t_2) = \frac{K_X(t_1, t_2)}{\sigma_X(t_1) \cdot \sigma_X(t_2)}. \quad (1.14)$$

Свойства нормированной корреляционной функции аналогичны свойствам коэффициента корреляции для случайных величин:

1. $|r_X(t_1, t_2)| \le 1$.

2. $r_X(t,t) = 1$.
3. $r_X(t_1,t_2) = r_X(t_2,t_1)$.

Пример 1.9. Используя условие примера 1.7, найти корреляционную и нормированную корреляционную функции случайного процесса $Y(t)$.

Решение. $K_Y(t_1,t_2) = M\big((Xe^{-t_1} - 3e^{-t_1})\cdot(Xe^{-t_2} - 3e^{-t_2})\big) =$

$$= M(X^2e^{-t_1}\cdot e^{-t_2} - 6Xe^{-t_1}\cdot e^{-t_2} + 9e^{-t_1}\cdot e^{-t_2}) =$$

$$= M(e^{-t_1}\cdot e^{-t_2}(X-3)^2) = e^{-t_1}\cdot e^{-t_2}\cdot DX = e^{-t_1}\cdot e^{-t_2}\cdot 1 = e^{-(t_1+t_2)}.$$

Тогда $r_X(t_1,t_2) = \frac{e^{-(t_1+t_2)}}{e^{-t_1}\cdot e^{-t_2}} = 1$.

Пример 1.10. Найти математическое ожидание и корреляционную функцию случайного процесса: $X(t) = \omega\cdot sint$, где $t \geq 0$; ω – непрерывная случайная величина, равномерно распределенная на отрезке $[-1,1]$ с плотностью вероятностей:

$$f(x) = \begin{cases} 0{,}5, & x \in [-1,1] \\ 0, x < -1, x > 1. \end{cases}$$

Решение. При фиксированном ω реализацией случайного процесса $X(t)$ является синусоида с амплитудой ω.

По формулам (1.7) и (1.13):

$$MX(t) = \int_{-\infty}^{+\infty} x\cdot f_t(x)dx = \int_{-1}^{1} 0{,}5xdx = 0\,.$$

$$K_X(t_1,t_2) = M\big(X(t_1)\cdot X(t_2)\big) - m_X(t_1)\cdot m_X(t_2) = M\big(X(t_1)\cdot X(t_2)\big) =$$

$$= M\big(\omega\cdot \sin(t_1)\cdot\omega\cdot sin(t_2)\big) = \sin(t_1)\cdot sin(t_2)\cdot M(\omega^2) =$$

$$= \sin(t_1)\cdot sin(t_2)\cdot\int_{-1}^{1} 0{,}5x^2\,dx = \frac{1}{3}\cdot\sin(t_1)\cdot sin(t_2).$$

1.4.4. Взаимная корреляционная функция случайного процесса

Для определения степени зависимости сечений двух случайных процессов используют корреляционную функцию связи или взаимную корреляционную функцию.

Определение 1.9. *Взаимной корреляционной функцией* случайных процессов $X(t)$ и $Y(t)$ называется неслучайная функция $R_{XY}(t_1,t_2)$ двух независимых аргументов t_1 и t_2, которая при каждой паре значений t_1 и t_2 равна корреляционному моменту двух сечений $X(t_1)$ и $X(t_2)$:

$$R_{XY}(t_1,t_2) = M\Big(\big(X(t_1) - m_X(t_1)\big)\cdot\big(Y(t_2) - m_Y(t_2)\big)\Big). \quad (1.15)$$

Определение 1.10. Два случайных процесса $X(t)$ и $Y(t)$ называются *некоррелированными*, если их взаимная корреляционная функция тождественно равна нулю, то есть $R_{XY}(t_1, t_2) = 0$ для любых значений t_1 и t_2. В противном случае, случайные процессы $X(t)$ и $Y(t)$ называются *коррелированными* (или *связанными*).

Свойства взаимной корреляционной функции непосредственно вытекают из ее определения и свойств корреляционного момента:

1. При одновременной перестановке индексов и аргументов взаимная корреляционная функция не меняется, то есть $R_{XY}(t_1, t_2) = R_{YX}(t_2, t_1)$.

2. Модуль взаимной корреляционной функции двух случайных процессов не превышает произведения их средних квадратических отклонений, то есть $|R_{XY}(t_1, t_2)| \le \sigma_X(t_1) \cdot \sigma_Y(t_2)$.

3. Корреляционная функция не изменится, если к случайным процессам $X(t)$ и $Y(t)$ прибавить неслучайные функции $f(t)$ и $\varphi(t)$ соответственно, то есть $R_{X_1Y_1}(t_1, t_2) = R_{XY}(t_1, t_2)$, где $X_1(t) = X(t) + f(t), Y_1(t) = Y(t) + \varphi(t)$.

4. Неслучайные множители можно вынести за знак корреляции, то есть если

$$X_1(t) = X(t) \cdot f(t), Y_1(t) = Y(t) \cdot \varphi(t), \text{ то}$$

$$R_{X_1Y_1}(t_1, t_2) = f(t_1) \cdot \varphi(t_2) \cdot R_{XY}(t_1, t_2).$$

5. Если $X(t) = X_1(t) + X_2(t)$, то

$$K_X(t_1, t_2) = K_{X_1}(t_1, t_2) + K_{X_2}(t_1, t_2) + R_{X_1X_2}(t_1, t_2) + R_{X_2X_1}(t_1, t_2).$$

6. Если случайные процессы $X_1(t)$ и $X_2(t)$ некоррелированы, то корреляционная функция их суммы равна сумме их корреляционных функций:

$$K_X(t_1, t_2) = K_{X_1}(t_1, t_2) + K_{X_2}(t_1, t_2).$$

Для оценки степени зависимости сечений двух случайных процессов используют *нормированную взаимную корреляционную функцию* $r_{XY}(t_1, t_2)$, определяемую равенством:

$$r_{XY}(t_1, t_2) = \frac{R_{XY}(t_1, t_2)}{\sqrt{D_X(t_1) \cdot D_Y(t_2)}}. \quad (1.16)$$

Пример 1.11. Найти взаимную корреляционную функцию двух случайных процессов $X(t) = t \cdot V$ и $Y(t) = (t + 2) \cdot V$, где V – случайная величина, причем $DV=3$.

Решение. $m_X(t) = M(t \cdot V) = t \cdot MV = t \cdot m_V$,

$$m_Y(t) = M((t + 2) \cdot V) = (t + 2) \cdot m_V, \text{ тогда}$$

$$R_{XY}(t_1, t_2) = M(t_1V - t_1m_V)\big((t_2 + 2)V - m_V(t_2 + 2)\big) =$$

$$= M\big(t_1(V - m_V)\big)\big((t_2 + 2)(V - m_V)\big) = t_1\big((t_2 + 2)DV\big) = 3t_1(t_2 + 2).$$

Пример 1.12. На рис. 7 показаны теоретически ковариационные функции процессов, с которыми часто приходится иметь дело на практике. Показаны ковариационные функции гармонического сигнала со случайной начальной фазой (а), белого шума (б), цветного шума (в), узкополосного шума (г), суммы гармонического сигнала со случайной начальной фазой и цветного шума (д), суммы гармонического сигнала со случайной начальной фазой и узкополосного шума (е).

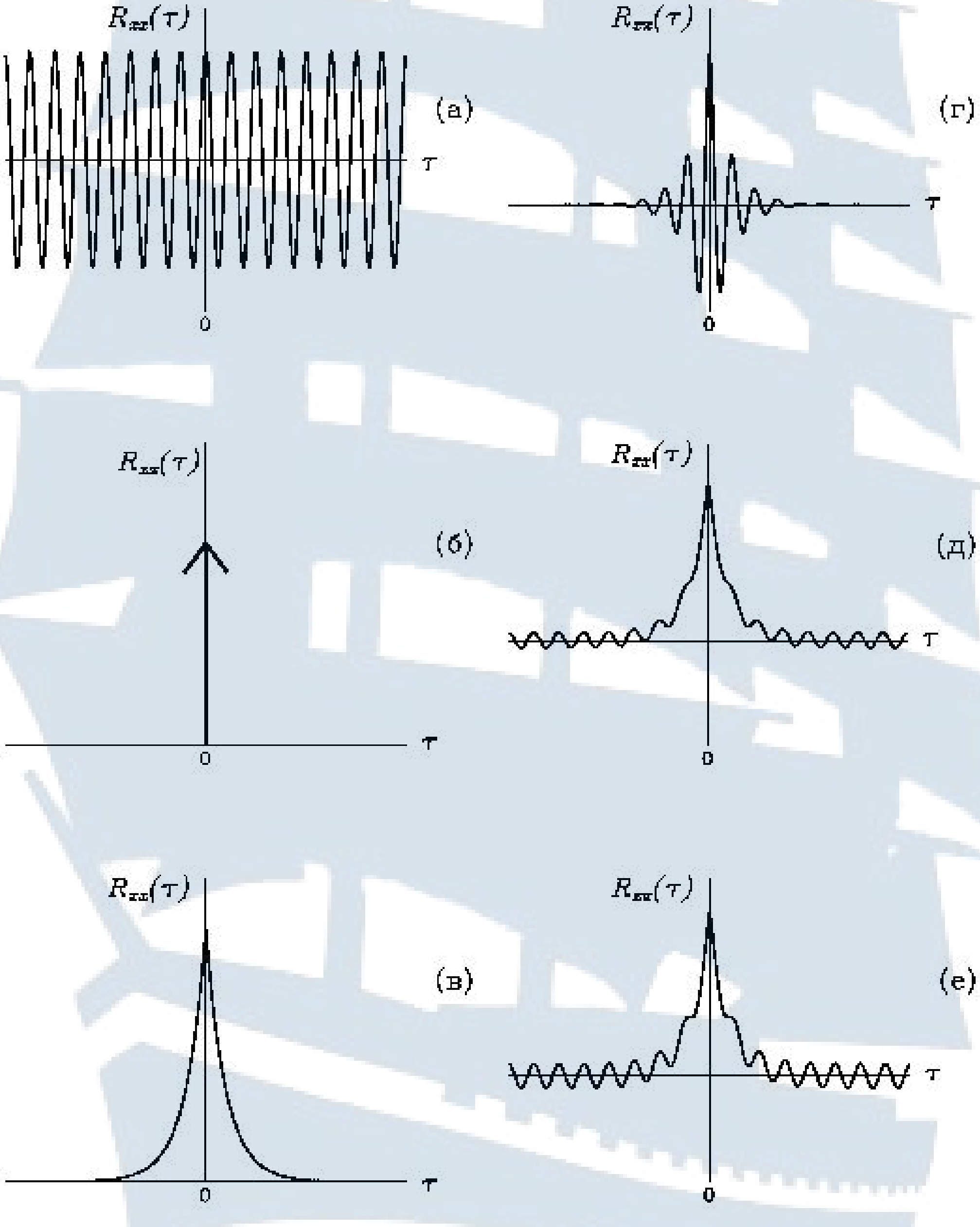


Рис. 7. Ковариационные функции процессов

ГЛАВА 2. Основные случайные процессы

2.1. Классификация случайных процессов

Случайный процесс, протекающий в любой физической системе, представляет собой случайные переходы системы из одного состояния в другое. Случайный процесс находится в некотором *состоянии*, если он полностью описывается значениями переменных, которые задают это состояние. Процесс совершает переход из одного состояния в другое, если описывающие его переменные изменяются от значений, задающих одно состояние на значения, определяющие другое состояние. Случайный процесс состоит в том, что с течением времени процесс переходит из одного состояния в другое, *заранее неизвестно какое.*

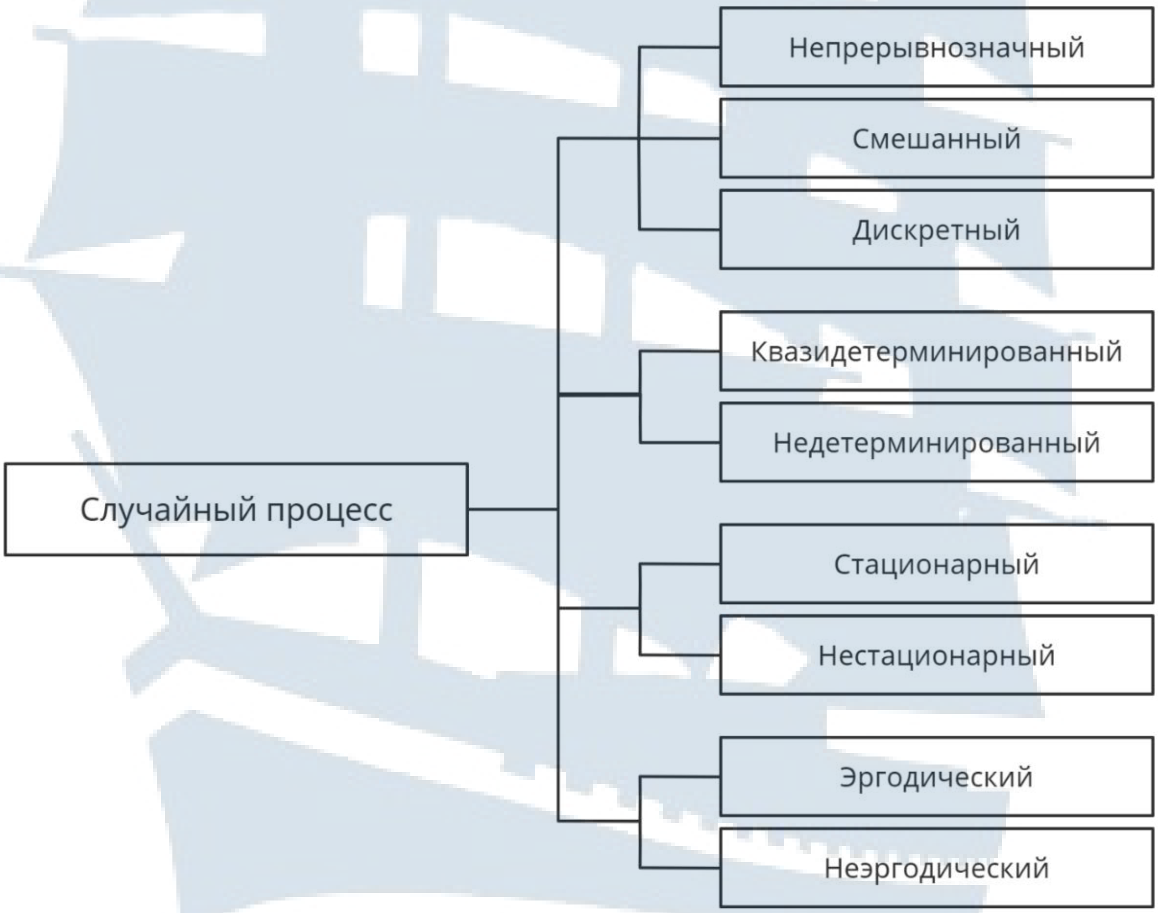


Рис. 8. Классификация случайных процессов

В зависимости от множества этих состояний W, от множества T значений аргумента t все случайные процессы делят на следующие классы:

1) дискретный процесс (дискретное состояние) с дискретным временем;

2) дискретный процесс с непрерывным временем;

3) непрерывный процесс (непрерывное состояние) с дискретным временем;

4) непрерывный процесс с непрерывным временем.

В случаях 1) и 3) множество T дискретно, то есть система, в которой протекает процесс, может менять свои состояния только в моменты времени $t_1, t_2, t_3, \ldots, t_n$, число которых конечно или счетно. При этом в случае 1) в любой момент времени t множество состояний W конечно или счетно, то есть в любой момент времени t сечение случайного процесса характеризуется дискретной случайной величиной, а в случае 3) – непрерывной случайной величиной. В случаях 2) и 4) множество T непрерывно, то есть переходы системы из состояния в состояние могут происходить в любой момент времени t наблюдаемого периода. При этом в случае 2) множество состояний системы W конечно или счетно, а в 4) W – несчетно.

Определение 2.1. *Непрерывнозначный* случайный процесс – это процесс, для которого случайные величины $\mathrm{X}(\mathrm{t_i})$ могут принимать любые значения в пределах заданной области возможных значений, например, тепловые шумы в проводниках, дробовые шумы в транзисторах, скорость ветра. Типичная реализация такого случайного процесса имеет вид, показанный на рис. 9, а.

Определение 2.2. *Дискретный* случайный процесс – это процесс, для которого случайные величины могут принимать только определенные значения, например, напряжение, которое случайным образом принимает два значения – нулевое и ненулевое в зависимости от того, открыт или заперт коммутатор (рис. 9, б).

Существуют также смешанные случайные процессы, которые имеют как непрерывнозначную, так и дискретную составляющие (рис. 9, в).

Большинство *гидрологических процессов* являются процессами с непрерывным состоянием и непрерывным временем. Но при вводе шага дискретности по времени они превращаются из непрерывного процесса с непрерывным временем в непрерывный процесс с дискретным временем.

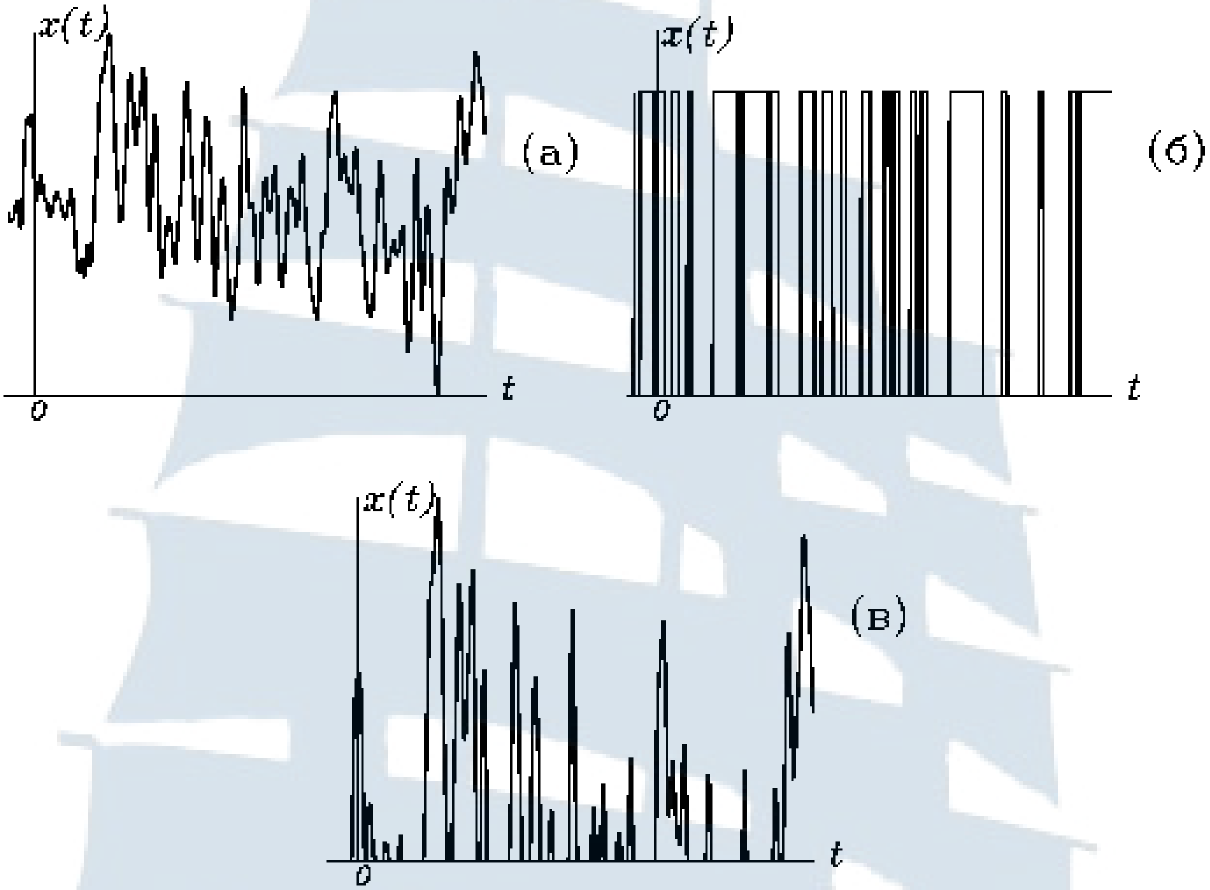


Рис. 9. Примеры реализаций непрерывнозначного (а),
Дискретного (б) и смешанного (в) случайных процессов

Для случайных процессов с непрерывным временем время между переходами из одного состояния в другое случайно. Это означает, что вероятность перехода из одного состояния в другое не может быть задана, поскольку вероятность такого перехода точно в произвольный момент времени равна нулю.

Интенсивность перехода g_{ij} из состояния E_i в состояние E_j определяется как предел отношения вероятности перехода $P_{ij}(\Delta\tau)$ системы за промежуток времени $\Delta\tau$ из E_i в E_j к длине этого промежутка:

$$g_{ij} = \lim_{\Delta\tau\to\infty} \frac{P_{ij}(\Delta\tau)}{\Delta\tau} (i, j = 1, \dots n, i \neq j).$$

Отсюда следует, что вероятность перехода за бесконечно малый промежуток времени $\Delta\tau$ равна: $g_{ij}\Delta\tau$ $(i \neq j)$. Вероятность двух и более переходов за время $\Delta\tau$ имеет порядок $(\Delta\tau)^2$ и выше и предполагается бесконечно малой величиной.

Интенсивности переходов задаются в виде квадратной матрицы: $G = [g_{ij}], i,j = 1, \dots, n$, называемой *матрицей интенсивностей переходов,* диагональные элементы которой определяются из условия:

$$\sum_{j=1}^{n} g_{ij} = 0 (i = 1, \dots n), \text{ откуда } g_{ii} = -\sum_{j=1, j\neq i}^{n} g_{ij} (i,j = 1, \dots n).$$

Определение 2.3. Если реализация случайного процесса является случайной функцией времени и ее будущие значения могут быть точно предсказаны на основе зарегистрированных раньше значений, то такой случайный процесс называется *квазидетерминированным*. Если предсказание будущих значений невозможно, то процесс называется *недетерминированным*.

Определение 2.4. *Стационарными* называются такие процессы, вероятностные свойства которых не зависят от начала отсчета времени.

Некоторые стационарные случайные процессы обладают свойством, заключающимся в том, что каждый член ансамбля реализаций ведет себя в статистическом смысле так же, как и весь ансамбль. В этом случае все характеристики случайных процессов можно проанализировать путем исследования свойств только одной реализации.

Определение 2.5. Случайные процессы называются *эргодическими в строгом смысле,* если об этом процессе в целом можно судить по одной его достаточно продолжительной реализации. Такие случайные процессы всегда являются стационарными. Все нестационарные случайные процессы неэргодичны, однако неэргодическими могут быть и стационарные случайные процессы.

Пример 2.1. Рассмотрим квазидетерминированный случайный процесс вида

$$X(t) = Y \cdot \cos(\omega t + \theta),$$

где ω – постоянная; $\boldsymbol{Y}$ – случайная относительно ансамбля реализаций величина; Θ – случайная величина, равномерно распределенная в интервале $[0,2\pi]$, причем Θ и $\boldsymbol{Y}$ статистически *независимы*.

Поскольку $Y = Y_k = const$ для отдельной k-й реализации, но принимает другие (постоянные) значения для других реализаций, то этот процесс не является эргодическим, однако можно показать, что он является стационарным.

В общем случае трудно доказать, что *эргодичность* – обоснованное допущение для какого-либо физического случайного процесса, так как на практике обычно возможно измерить лишь одну реализацию этого процесса. Однако часто оказывается полезным предположить эргодичность случайного процесса, если только отсутствуют веские доводы физического характера, препятствующие этому. В предположении эргодичности анализ временных рядов значительно упрощается. Поэтому далее будем рассматривать только эргодические и, следовательно, стационарные случайные процессы.

2.2. Стационарные случайные процессы

Важным классом случайных процессов являются стационарные случайные процессы.

Определение 2.6. *Стационарные случайные процессы* – это процессы, не изменяющие свои вероятностные характеристики с течением времени, то есть $m_X = const;\ D_X = const$.

Они имеют вид непрерывных случайных колебаний вокруг некоторого среднего значения. Например, давление газа в газопроводе, колебания самолета при «автополете», колебания напряжения в электрической сети и т. д. Отличие стационарных и нестационарных случайных процессов показано на рис. 10.

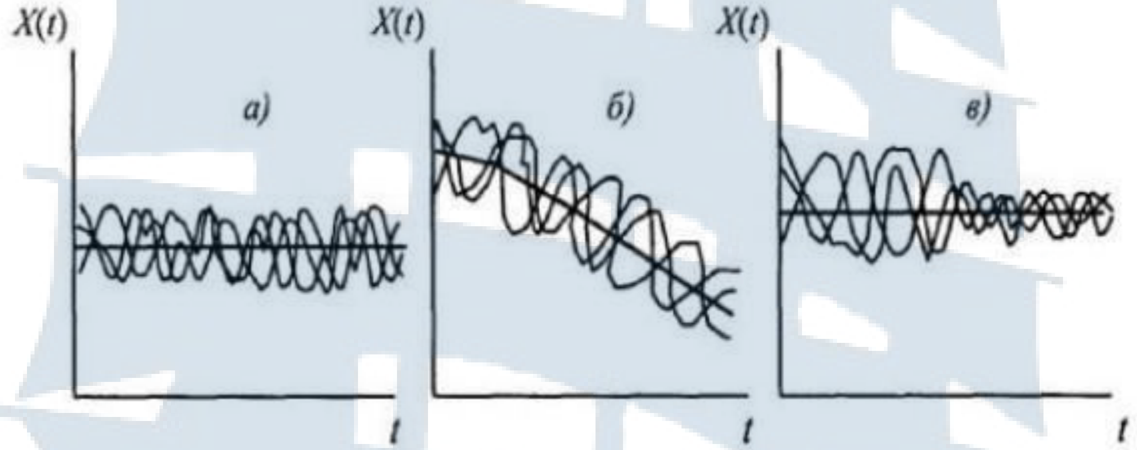


Рис. 10. а) стационарный случайный процесс;
б) нестационарный случайный процесс по математическому ожиданию;
в) нестационарный случайный процесс по дисперсии

Определение 2.7. Случайный процесс $X(t)$ называется *стационарным в широком смысле,* если его математическое ожидание постоянно, а корреляционная функция зависит только от разности аргументов, то есть

$$m_X(t) = m = const;\ \ K_X(t_1, t_2) = K_X(t_2 - t_1).$$

Из этого определения следует, что корреляционная функция стационарного процесса есть функция одного аргумента:

$$K_X(t_1, t_2) = K_X(r), \text{где } r = t_2 - t_1.$$

Это обстоятельство упрощает операции над стационарными случайными процессами.

Определение 2.8. Случайный процесс $X(t)$ называется *стационарным в узком смысле,* если все его характеристики зависят не от значений аргументов, а лишь от их взаимного расположения. Это означает, что совместное распределение случайных величин:

$\{X(t_1+h), X(t_2+h), \dots, X(t_n+h)\}$ и $\{X(t_1), X(t_2), \dots, X(t_n)\}, t_i + h \in T$ одинаково и не зависит от h, то есть

$$F_{t_1+h, t_2+h, \dots, t_n+h}(x_1, x_2, \dots, x_n) = F_{t_1, t_2, \dots, t_n}(x_1, x_2, \dots, x_n),$$

при любых $h > 0, n \geq 1, t_1, t_2, \dots, t_n \in T$.

Отсюда для плотности вероятностей f имеет место соотношение

$$f_{t_1+h, t_2+h, \dots t_n+h}(x_1, x_2, \dots, x_n) = f_{t_1, t_2, \dots, t_n}(x_1, x_2, \dots, x_n).$$

Для стационарного случайного процесса $X(t)$, учитывая, что

$f_{t+h}(x) = f_t(x)$ при $h = -t$, имеем: $f_t(x) = f_0(x) = f(x)$,

где $f_t(x)$ – одномерная плотность вероятностей. Отсюда

$$m_X(t) = \int_{-\infty}^{+\infty} x \cdot f_t(x) dx = \int_{-\infty}^{+\infty} x \cdot f(x) dx = m_X.$$

Аналогично для двумерной плотности вероятностей из равенства

$f_{t_1+h, t_2+h}(x_1, x_2) = f_{t_1, t_2}(x_1, x_2)$ при $h = -t_1$, получим

$$f_{t_1, t_2}(x_1, x_2) = f_{t_1, t_2}(x_1, x_2, 0, t_2 - t_1).$$

Отсюда для корреляционной функции

$$K_X(t_1, t_2) = M((X(t_1) - m(t_1)) \cdot (X(t_2) - m(t_2)) =$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (X(t_1) - m_X)(X(t_2) - m_X) f_{t_1, t_2}(x_1, x_2, t_2 - t_1) dx_1 dx_2 = K_X(t_2 - t_1).$$

Таким образом, для процессов, стационарных в узком смысле, математическое ожидание является постоянной величиной, а корреляционная функция зависит только от разности аргументов, то есть

$$K_X(t_1, t_2) = K_X(t_2 - t_1) = K_X(t_1 - t_2),$$

так как корреляционная функция симметрична.

Очевидно, что стационарный в узком смысле случайный процесс является стационарным и в широком смысле. Обратное в общем случае неверно.

Теория стационарных процессов разработана наиболее полно и позволяет сравнительно просто производить расчеты для многих практических случаев. Поэтому допущение о стационарности иногда целесообразно делать также и для тех случаев, когда случайный процесс хотя и не стационарен, но на рассматриваемом отрезке времени работы системы статистические характеристики сигналов не успевают сколь-нибудь существенно измениться. В дальнейшем, если не будет оговорено особо, будут рассматриваться случайные процессы, стационарные в широком смысле.

Пример 2.2. Пусть $X(t), t \geq 0,$ – случайный процесс, описываемый дифференциальным уравнением: $\dot{x} + ax = \omega,$ где a – положительная постоянная, а ω – случайная величина с нулевым математическим ожиданием и постоянной дисперсией σ^2, не зависящая от начального условия $x(0) = 0$. Найти характеристики случайного процесса $X(t)$.

Решение. Решив данную задачу Коши, получим траекторию данного процесса: $x(t) = \frac{\omega}{a}(1 - e^{-at}), t \geq 0$. Так как $M(\omega) = 0$, находим:

$$m_X(t) = M\left(\frac{\omega}{a}(1 - e^{-at})\right) = \frac{1}{a}(1 - e^{-at})M(\omega) = 0.$$

Далее,

$$D_X(t) = MX^2(t) - m_X^2(t) = \frac{1}{a^2}(1 - e^{-at})^2 M(\omega^2) = \frac{\sigma^2}{a^2}(1 - e^{-at})^2;$$

$$K_X(t_1, t_2) = M(X(t_1)) \cdot X(t_2)) = M(\frac{\omega^2}{a^2}(1 - e^{-at_1})(1 - e^{-at_2})) =$$

$$= \frac{\sigma^2}{a^2}(1 - e^{-at_1})(1 - e^{-at_2}) \neq K_X(t_2 - t_1),$$

то есть рассматриваемый процесс не является стационарным в широком смысле (тем более в узком смысле). Заметим, что при t_1 и t_2, стремящихся к бесконечности, корреляционная функция $K_X(t_1, t_2)$ стремится к $\frac{\sigma^2}{a^2}$.

Таким образом, в пределе при $t \to \infty$ процесс становится стационарным в широком смысле.

Пример 2.3. Случайный процесс, задающий гармонические колебания:

$$X(t) = A \cdot \cos(\omega t + \varphi).$$

Можно показать, что $m_X(t) = 0$, а

$$K_X(t_1, t_2) = \frac{1}{2}M(A^2) \cdot \cos\omega(t_1 - t_2) = \sigma_X^2 \cdot cos\omega\tau.$$

Основные свойства корреляционной функции стационарного случайного процесса (ССП):

1. Дисперсия ССП постоянна и равна корреляционной функции в нуле, то есть $D_X(t) = K_X(0) = const$.

Действительно, $D_X(t) = K_X(t, t) = K_X(t - t) = K_X(0)$.

2. Корреляционная функция ССП является четной, то есть

$$K_X(\tau) = K_X(-\tau).$$

Для доказательства достаточно положить $t_2 - t_1 = t$ и воспользоваться свойством симметричности корреляционной функции.

3. Модуль корреляционной функции ССП не превышает ее значения при $\tau = 0$, то есть $|K_X(\tau)| \leq K_X(0)$.

В самом деле, из неравенства Коши-Буняковского имеем:

$$K_X^2(t_2 - t_1) \leq K_X(t_1 - t_1)K_X(t_2 - t_2) = K_X^2(0).$$

Определение 2.9. Случайные процессы $X(t)$ и $Y(t)$ называются *стационарно связанными,* если их взаимно корреляционная функция $K_{XY}(t_1, t_2)$ зависит только от $\tau = t_2 - t_1$.

2.3. Эргодические случайные процессы

Большинство ССП обладают важным для практики *эргодическим свойством*, сущность которого состоит в том, что по одной достаточно продолжительной реализации можно судить о всех свойствах процесса так же, как и по любому количеству реализаций, то есть отдельные характеристики ССП $\{m_X, K_X(\tau)\}$ могут быть определены как соответствующие средние по времени для одной реализации достаточно большой продолжительности.

В теории случайных процессов пользуются *двумя понятиями средних значений.* Первое понятие о среднем значении – это *среднее значение по множеству (или математическое ожидание)*, которое определяется на основе наблюдения над множеством реализации случайного процесса в один и тот же момент времени. В общем случае среднее значение по множеству является функцией времени. Другое понятие о среднем значении – это *среднее значение по времени*, которое определяется на основе наблюдения за отдельной реализацией случайного процесса $x(t)$ на протяжении достаточно длительного времени Т.

В качестве оценок характеристик эргодических ССП принимают усредненное по времени значение:

$$\widetilde{m}_X(t) = \frac{1}{T}\int_0^T x(t)dt; \tag{2.1}$$

$$\widetilde{K}_X(\tau) = \frac{1}{T}\int_0^{T-\tau} \big(x(t) - \widetilde{m}_X(t)\big) \cdot \big(x(t+\tau) - \widetilde{m}_X(t)\big)dt. \tag{2.2}$$

Интегралы в правых частях равенств вычисляются приближенно.

Среднее значение по времени в общем случае различно для отдельных реализаций множества, определяющих случайный процесс.

Вообще, для одного и того же случайного процесса среднее по множеству и среднее по времени различны, однако для эргодических стационарных случайных процессов среднее значение по множеству совпадает со средним значением по времени. Равенство средних по множеству и по времени вытекает из *эргодической теоремы*, в которой для некоторых стационарных случайных процессов доказано, что любая статистическая характеристика, полученная усреднением по множеству, с вероятностью, сколь угодно близкой к единице, совпадает с характеристикой, усредненной по времени.

Эргодическая теорема доказана не для всех стационарных процессов, поэтому в тех случаях, где она еще не доказана, говорят об эргодической гипотезе. Следует заметить, что не всякий стационарный процесс является эргодическим.

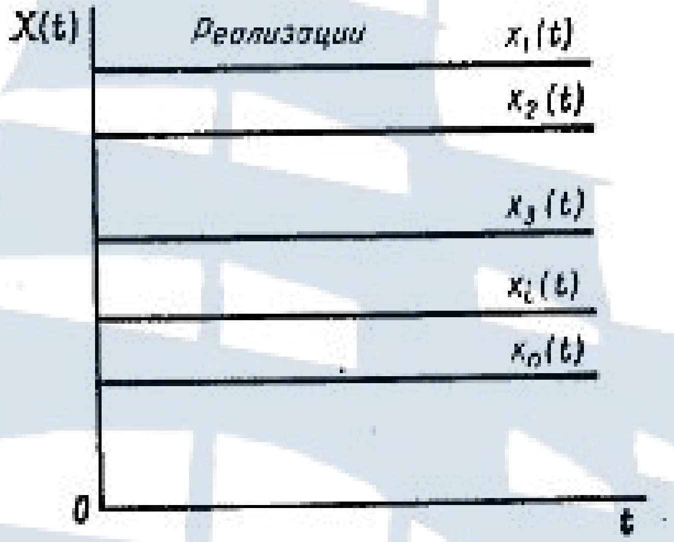


Рис. 11. График стационарного неэргодического процесса

На рис. 11 изображен график стационарного неэргодического процесса, для которого равенство средних по множеству и по времени не выполняется. Один и тот же случайный процесс в общем случае может быть эргодическим по отношению к одним статистическим характеристикам и не эргодическим по отношению к другим. В дальнейшем будем считать, что условия эргодичности для математического ожидания и корреляционной функции выполняются.

Физический смысл эргодической теоремы (или гипотезы) глубок и имеет большое практическое значение. Для определения статистических свойств эргодических стационарных процессов, если трудно осуществить одновременное наблюдение за множеством подобных систем в произвольно выбранный момент времени, например при наличии одного опытного образца, его можно заменить длительным наблюдением за одной системой. Этот факт лежит в основе экспериментального определения корреляционной функции стационарного случайного процесса по одной реализации. Наоборот, при наличии большой партии изделий массовой продукции для аналогичных исследований можно провести одновременное наблюдение за всеми образцами партии или их достаточно представительной выборкой.

Определение 2.10. Говорят, что случайный процесс обладает *эргодическим свойством,* если по истечении достаточно большого промежутка времени вероятности состояний стремятся к предельным значениям $p_1, p_2, \dots, p_n$, не зависящим от начальных вероятностей$p_1(0), p_2(0), \dots, p_n(0)$ и от текущего момента времени.

Таким образом, для случайных процессов, обладающих эргодическим свойством, имеет место: $\lim_{t\to\infty} P(t) = P(\infty) = P$, где $P = (p_1, p_2, \dots, p_n)$ – вектор вероятностей состояний системы, называемых стационарными вероятностями.

Случайный процесс с дискретным временем обладает эргодическим свойством, если матрица вероятностей переходов не является *периодической* или *разложимой.*

Определение 2.11. *Разложимой* называется матрица, которая может быть приведена к одному из следующих видов:

$$1)\begin{bmatrix} A & 0 \\ 0 & D \end{bmatrix};\ \ 2)\begin{bmatrix} A & 0 \\ C & D \end{bmatrix};\ \ 3)\begin{bmatrix} A & B \\ 0 & D \end{bmatrix},$$

где A, B, C, D – ненулевые квадратные подматрицы; 0 – нулевая квадратная подматрица.

В первом случае состояния, соответствующие подмножествам A и D, называются *замкнутыми,* так как система, находясь в каком-то состоянии одного из этих подмножеств, никогда не сможет перейти в какое-либо состояние другого подмножества. Состояния, соответствующие подмножеству D во втором случае и подмножеству A в третьем случае, называются *невозвратными,* поскольку после того, как процесс покинет эти состояния, невозможен обратный переход в эти состояния из состояний, соответствующих другим подмножествам.

Определение 2.12. Матрица называется *периодической*, если она может быть приведена к виду: $\begin{bmatrix} 0 & B \\ C & 0 \end{bmatrix}$.

Случайный процесс в этом случае по очереди будет переходить из состояний, соответствующих B, в состояния, соответствующие C.

Достаточным условием эргодического случайного процесса $X(t)$ относительно математического ожидания и корреляционной функции является стремление к нулю его корреляционной функции при $\tau \to \infty$ то есть $\lim\limits_{\tau \to \infty} K_X(\tau) = 0$. Это означает, что при увеличении сдвига между сечениями корреляционная функция затухает (рис. 12).

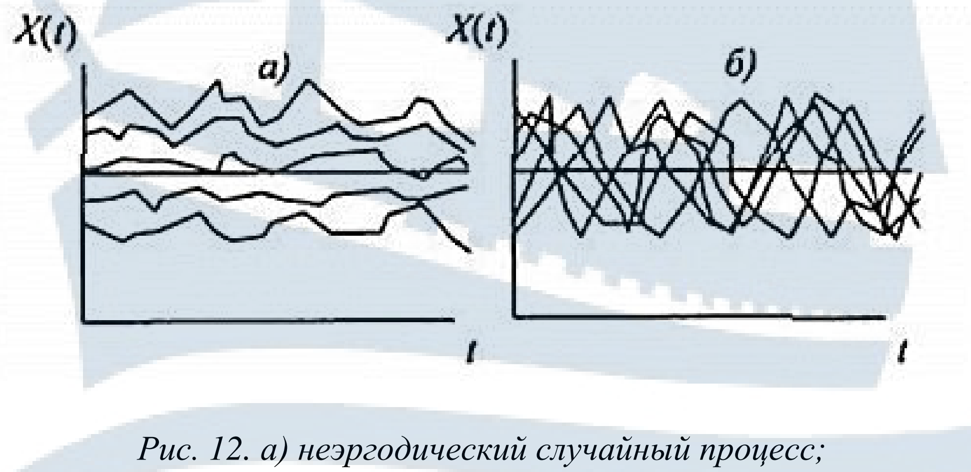


Рис. 12. а) неэргодический случайный процесс;
б) эргодический случайный процесс

2.4. Марковский процесс

Среди случайных процессов особое место занимают марковские случайные процессы (А.А. Марков (1856–1922) – русский математик).

Рассмотрим некоторую физическую систему S, в которой происходит случайный процесс. С течением времени под влиянием случайных факторов она может переходить из одного состояния в другое.

Определение 2.13. Дискретный случайный процесс называется *марковским* или *процессом без последствия,* если условная вероятность каждого из состояний системы S в будущем зависит только от ее состояния в настоящем и не зависит от того, когда и как система пришла в это состояние.

Таким образом, марковский процесс допускает зависимость исхода любого испытания только от исхода предыдущего испытания, то есть будущее развитие его зависит только от настоящего состояния и не зависит от «предыстории» процесса.

Пример 2.4. Пусть система S представляет собой техническое устройство, которое уже проработало некоторое время, «износилось» и пришло в некоторое состояние, характеризующееся определенной степенью изношенности S. Как будет работать система в будущем? Очевидно, что характеристики работы системы в будущем (частота отказов, потребность в ремонте) зависят от состояния устройства в настоящий момент и не зависят от того, когда и как устройство достигло своего настоящего состояния.

На практике часто встречаются случайные процессы, которые, с той или иной степенью приближения, можно считать марковскими.

Теория марковских случайных процессов является в настоящее время очень обширным разделом теории вероятностей с широким спектром различных приложений – от описания физических явлений типа диффузии или перемешивания шихты во время плавки в доменной печи до процессов образования очередей или распространения мутаций генов в биологической популяции. Нас будут интересовать, главным образом, применения теории марковских случайных процессов к построению математических моделей операций, ход и исход которых существенно зависит от случайных факторов.

Марковские случайные процессы делятся на классы по некоторым признакам, в зависимости от того, как и в какие моменты времени система S может менять свои состояния.

Определение 2.14. Случайный процесс называется *процессом с дискретными состояниями*, если возможные состояния системы:

$$S_1, S_2, S_3, \ldots$$

можно перечислить (перенумеровать) одно за другим, а сам процесс состоит в том, что время от времени система S скачком (мгновенно) перескакивает из одного состояния в другое.

Кроме процессов с дискретными состояниями существуют случайные процессы с *непрерывными состояниями*: для этих процессов характерен постепенный, плавный переход из состояния в состояние. Например, процесс изменения напряжения в осветительной сети представляет собой случайный процесс с непрерывными состояниями.

Определение 2.15. Случайный процесс $X(t)$ называют марковским, если для любых n моментов времени $t_1 < t_2 < \cdots < t_{n-1} < t_n$ из T условная функция распределения вероятностей случайной величины $X(t_n)$ при известных $X(t_1), X(t_2), \dots, X(t_{n-1})$ есть

$$P(X(t_n) < x_n | X(t_{n-1}) = x_{n-1}, \dots, X(t_2) = x_2, X(t_1) = x_1) = \\ = P(X(t_n) < x_n) | X(t_{n-1}) = x_{n-1}) \qquad (2.3)$$

для произвольных чисел $x_1, x_2, \dots, x_n$.

Равенство (2.3) в терминах условной плотности вероятностей случайной величины $X(t_n)$ означает, что

$$f(X(t_n) < x_n | X(t_{n-1}) = x_{n-1}, \dots, X(t_2) = x_2, X(t_1) = x_1) = \\ = f(X(t_n) < x_n | X(t_{n-1}) = x_{n-1}), \qquad (2.4)$$

или

$$f(x_n | x_{n-1}, \dots, x_1) = f(x_n | x_{n-1}). \qquad (2.5)$$

Функция (2.4) или (2.5) называется *условной переходной плотностью вероятностей.*

Пусть $X(t), t \in T,$ – марковский процесс с условной переходной плотностью вероятностей $f\big(x(t)\big|x(\tau)\big)$, определенной для любых значений t и τ из $T, t > \tau$. Пусть $f(x_n, x_{n-1}, \dots, x_1)$ – совместная плотность распределения случайных величин $X(t_{n-1}), \dots, X(t_2), X(t_1)$ для любых n и $t_1 < t_2 < \cdots < t_{n-1} < t_n$ из *T*. Согласно формуле умножения вероятностей,

$$f(x_n, x_{n-1}, \dots, x_1) = f(x_n | x_{n-1}, \dots, x_1) \cdot f(x_{n-1}, \dots, x_1). \qquad (2.6)$$

Но в силу того, что процесс марковский, имеет место равенство (2.5). Из него и из равенства (2.6) получим:

$$f(x_n, x_{n-1}, \dots, x_1) = f(x_n | x_{n-1}) \cdot f(x_{n-1}, \dots, x_1). \qquad (2.7)$$

Повторив предыдущие рассуждения, найдем:

$$f(x_{n-1}, \dots, x_1) = f(x_{n-1} | x_{n-2}) \cdot f(x_{n-2}, \dots, x_1). \qquad (2.8)$$

Отсюда из равенства (2.7) получаем:

$$f(x_n, x_{n-1}, \dots, x_1) = f(x_n | x_{n-1}) \cdot f(x_{n-1} | x_{n-2}) \cdot f(x_{n-2}, \dots, x_1).$$

Повторив этот процесс (*n*-1) раз, окончательно придем к равенству:

$$f(x_n, x_{n-1}, \dots, x_1) = f(x_n | x_{n-1}) \cdot f(x_{n-1} | x_{n-2}) \cdot \dots \cdot f(x_2 | x_1) \cdot f(x_1). \qquad (2.9)$$

Для марковских процессов условные переходные плотности вероятностей $f(x_n|x_{n-1}),\ k = 2,3,\dots,n,$ известны. Поэтому из равенства (2.9) следует, что при известной плотности распределения вероятностей $f(x_1)$ можно найти совместную плотность распределения вероятностей марковского процесса.

Для того, чтобы случайный процесс с непрерывным временем был *марковским*, необходимо, чтобы интервалы времени между соседними переходами из состояния в состояние были распределены по экспоненциальному закону, то есть функция распределения имела вид:

$$F_{ij}(\tau) = 1 - e^{-\alpha_{ij}\tau}. \tag{2.10}$$

2.4.1. Дискретный марковский процесс с дискретным временем. Марковские цепи. Переходная матрица, ее свойства. Вероятности многошаговых переходов

Одним из простейших случайных процессов является так называемый конечный марковский процесс (конечная марковская цепь).

Определение 2.16. Случайный процесс называется *процессом с дискретными состояниями*, если возможные состояния системы:

$$S_1, S_2, S_3, \dots$$

можно перечислить (перенумеровать) одно за другим, а сам процесс состоит в том, что время от времени система S скачком (мгновенно) перескакивает из одного состояния в другое.

При анализе случайных процессов с дискретными состояниями очень удобно пользоваться геометрической схемой – так называемым *графом состояний*. Граф состояний геометрически изображает возможные состояния системы и ее возможные переходы из состояния в состояние. Будем изображать каждое состояние прямоугольником, а возможные переходы («перескоки») из состояния в состояние – стрелками, соединяющими эти прямоугольники.

Пример 2.5. Техническое устройство S состоит из двух узлов I и II, каждый из которых может в ходе работы устройства отказать (выйти из строя). Построить граф состояний. Предполагается, что ремонт узлов в ходе процесса не производится.

Решение. Процесс, протекающий в системе, состоит в том, что она случайным образом, в какие-то моменты времени, переходит (перескакивает) из состояния в состояние. Перед нами – процесс с дискретными состояниями.

Возможны следующие состояния системы: S_1 – оба узла работают; S_2 – первый узел отказал, второй работает; S_3 – второй узел отказал, первый работает; S_4 – оба узла отказали.

Отметим, что на графе (рис. 13) не показан возможный переход из состояния S_1 непосредственно в S_4, который осуществится, если строго одновременно выйдут из строя оба узла. Возможностью такого события мы пренебрегаем.

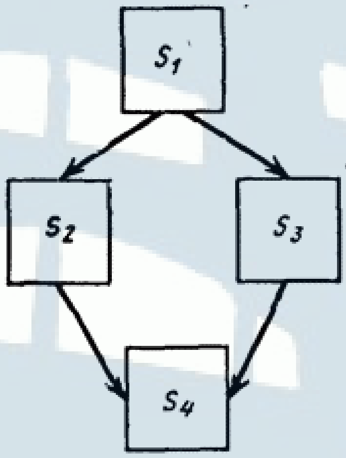


Рис. 13. Граф состояний

Пример 2.6. Система S – автомашина, которая может находиться в одном из пяти возможных состояний: S_1 – исправна, работает; S_2 – неисправна, ожидает осмотра; S_3 – осматривается; S_4 – ремонтируется; S_5 – списана (рис. 14).

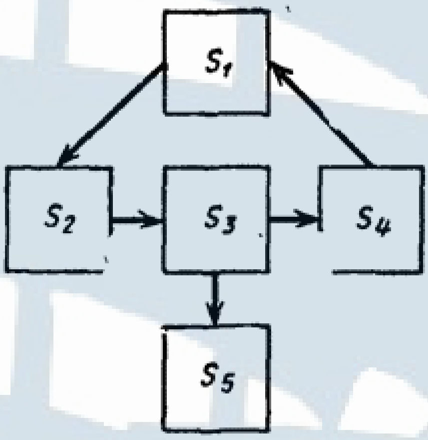


Рис. 14. Граф состояний системы

Пример 2.7. Система S представляет собой техническое устройство, состоящее из двух узлов I и II, каждый из них может в какой-то момент времени отказать. Отказавший узел немедленно начинает восстанавливаться (рис. 15).

Решение. Возможные состояния системы: S_1– оба узла работают;

S_2 – первый узел восстанавливается, второй работает;

S_3– первый узел работает, второй восстанавливается;

S_4 – оба узла восстанавливаются.

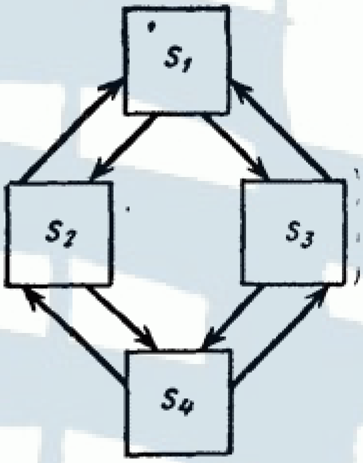


Рис. 15. Граф состояний системы

Пусть в некоторой системе S происходит случайный процесс с дискретными состояниями $s_1, s_2, \dots, s_n$ и дискретным временем, то есть переход системы из одного состояния в другое происходит только в определенные моменты времени $t_0, t_1, t_2, \dots$. Эти моменты называют *шагами* процесса. Поэтому процесс можно рассматривать как последовательность (цепь) событий S(0), S(1), S(2),…(S(0) – начальное состояние системы, то есть перед первым шагом; S(1) – состояние системы после первого шага; S(2) – состояние системы после второго шага и т. д.), то есть событий вида

$$\{S(k) = s_i\}, \text{где } i = 1,2,3, \dots, n; k = 0,1,2, \dots. \qquad (2.11)$$

Определение 2.17. Марковский случайный процесс с дискретными состояниями и дискретным временем называют *марковской цепью* или *цепью Маркова.*

Определение 2.18. Цепь, в которой условные вероятности состояний системы в будущем зависят только от состояния системы на последнем шаге, называют *простой цепью Маркова.*

Определение 2.19. Множество всех возможных состояний, которые может принимать процесс, называется *пространством состояний.*

Дискретный марковский случайный процесс удобно иллюстрировать с помощью так называемого *графа состояний.* В нем состояния $S_1, S_2, \dots$ системы S изображаются прямоугольниками (или кружками), а возможные непосредственные переходы из состояния в состояние – стрелками (или ориентированными дугами), соединяющими состояния.

Для описания дискретных процессов пользуются вероятностями состояний системы S, то есть значениями $p_1(k), p_2(k), \dots, p_n(k), i = 1,2,3, \dots, n.$

Рассмотрим простую цепь Маркова и вычислим вероятности состояний системы:

$$p_i(k) = P\{S(k) = s_i\}, i = 1,2,3, \dots, n; k = 0,1,2, \dots$$

Здесь $P\{S(k) = s_i\}$ – безусловная вероятность того, что на k-м шаге система будет находиться в состоянии S_i.

Для нахождения безусловной вероятности нужно знать начальное распределение вероятностей $p_1(0), p_2(0), \dots, p_n(0)$, то есть вероятности состояний $p_i(0)$ в момент времени $t_0 = 0$ (начало процесса) и так называемые переходные вероятности $p_{ij}(k)$ на k-м шаге.

Определение 2.20. *Переходной вероятностью* $p_{ij}(k)$ называют условную вероятность перехода системы S на k-м шаге в состояние S_j, если известно, что на предыдущем шаге она была в состоянии S_i, то есть

$$p_{ij}(k) = P\{S(k) = s_j | S(k-1) = s_i\}, i, j = 1,2,3, \dots, n. \qquad (2.12)$$

Определение 2.21. Цепь Маркова называется *однородной,* если $p_{ij}(k) = p_{ij}$, то есть условные вероятности $p_{ij}(k)$ не зависят от номера испытаний (рис. 16).

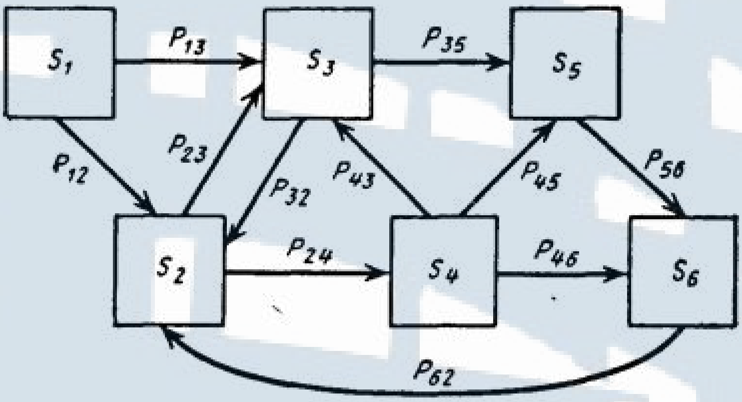


Рис. 16. Пример однородной цепи

Определение 2.22. *Переходной матрицей* марковской цепи называется квадратная матрица порядка n вида $P = (p_{ij})$, где p_{ij} – вероятность того, что при отдельном испытании (за один шаг) система переходит из состояния S_i в состояние S_j.

Элементы строки матрицы перехода P определяют вероятности того, что рассматриваемая система находится в определенном состоянии. Так как строка переходной матрицы учитывает все возможные состояния системы, определяя таким образом вероятности полной группы событий, то сумма вероятностей, записанных в строке, равна единице.

Свойства элементов переходной матрицы:

1) $p_{ij} \geq 0$;

2) $\sum_{j=1}^{n} p_{ij} = 1$ $(i = 1,2,3, \dots, n)$, то есть сумма вероятностей каждой строки матрицы перехода равна 1.

Пример 2.8. Фишка находится на средней клетке в ряду из 5 клеток: *A,B,C,D,E,* считая слева направо, и перемещается на соседнюю клетку по правилу: если она не на крайней клетке, то перемещается вправо, когда выпадет герб (Г) при подбрасывании монеты, а влево, когда выпадет решка (Р); если фишка на крайней клетке, то остается на ней независимо от выпадения Г или Р. Построить матрицу перехода марковской цепи.

Решение. При первом подбрасывании монеты в зависимости от выпадения Г или Р фишка переместится на одну клетку вправо или влево соответственно. При следующем подбрасывании монеты новое положение фишки будет зависеть только от предыдущего ее местоположения, то есть фишка может оказаться либо с краю, либо опять в центре. Здесь имеется 5 состояний: *A,B,C,D,E.* По условию задачи переход из состояний *B,C,D* в любое соседнее состояние A,B,C,D,E осуществляется с вероятностью 0,5. Если же фишка находится в крайних клетках A или E, то вероятность перехода из A в B и из E в D равна нулю в силу условий, а вероятность перехода из A в A и из E в E равна единице. В результате искомая матрица перехода имеет вид:

$$P = \begin{bmatrix} 10\ 0\ 00 \\ 0{,}5\ 0\ 0{,}5\ 0\ 0 \\ 0\ 0{,}5\ 00{,}5\ 0 \\ 00\ 0{,}5\ 0\ 0{,}5 \\ 00\ 0\ 0\ 1 \end{bmatrix}.$$

Пример 2.9. В любой данный день погода может быть хорошей (Х), посредственной (П) и плохой (Пл). Предположим, что если сегодня погода хорошая, то завтра она будет хорошей с вероятностью 0,6, посредственной с вероятностью 0,3 и плохой с вероятностью 0,1. Если сегодня погода посредственная, то завтра она будет хорошей, посредственной или плохой с вероятностями 0,2; 0,5 и 0,3 соответственно. Если сегодня плохая погода, то завтра она будет хорошей, посредственной или плохой с вероятностями 0,2; 0,3 и 0,5 соответственно. Этот процесс можно описать в виде марковской цепи с тремя исходами s_1, s_2, s_3, соответствующими хорошей, посредственной и плохой погоде в любой данный день. Переходная матрица этой марковской цепи имеет вид

$$P = \begin{bmatrix} 0{,}6 & 0{,}3 & 0{,}1 \\ 0{,}2 & 0{,}5 & 0{,}3 \\ 0{,}2 & 0{,}3 & 0{,}5 \end{bmatrix}. \qquad (2.13)$$

Вероятности многошаговых переходов

Пусть $P = (p_{ij})$ – переходная матрица. Здесь p_{ij}– условная вероятность того, что из состояния s_i, в котором система оказалась в результате некоторого испытания, она перейдет при следующем испытании в состояние s_j. Эта условная вероятность называется *одношаговой вероятностью* *перехода* из состояния s_i в состояние s_j. Двухшаговая вероятность $p_{ij}^{(2)}$ определяется как вероятность того, что система, находясь в состоянии s_i переходит в состояние s_j в результате двух испытаний. Переходя из состояния s_i в состояние s_j за два шага, система может пройти через некоторое промежуточное состояние $s_k, k = 1,2, \ldots$ Вероятность такого перехода есть $p_{ik}p_{kj}$, поскольку последовательные испытания независимы. По формуле полной вероятности получаем, что двухшаговая вероятность перехода

$$p_{ij}^{(2)} = p_{i1}p_{1j} + p_{i2}p_{2j} + \cdots + p_{in}p_{nj} = \sum_{k=1}^{n} p_{ik}p_{kj} \,. \qquad (2.14)$$

Это выражение представляет собой ij-й элемент матрицы P^2, где P– одношаговая переходная матрица этой марковской цепи.

Матрица P^2 называется *двухшаговой переходной матрицей марковской цепи.*

Аналогично, m – шаговая переходная матрица P^m определяется как m-я степень одношаговой переходной матрицы P.

Пример 2.10. Для одношаговой матрицы перехода (2.13) найти матрицу перехода из состояния s_i в состояние s_j за два шага.
Решение. Искомой матрицей является матрица P^2.

$$P^2 = \begin{bmatrix} 0,6 & 0,3 & 0,1 \\ 0,2 & 0,5 & 0,3 \\ 0,2 & 0,3 & 0,5 \end{bmatrix} \cdot \begin{bmatrix} 0,6 & 0,3 & 0,1 \\ 0,2 & 0,5 & 0,3 \\ 0,2 & 0,3 & 0,5 \end{bmatrix} = \begin{bmatrix} 0,44 & 0,36 & 0,20 \\ 0,28 & 0,40 & 0,32 \\ 0,28 & 0,36 & 0,36 \end{bmatrix}.$$

Полученный результат говорит о том, что если погода сегодня хорошая, то вероятность хорошей погоды двумя днями позже есть $p_{11}^{(2)} = 0{,}44$, а вероятность того, что если сегодня посредственная погода, то послезавтра она будет плохой, равна $p_{23}^{(2)} = 0{,}32$.

Замечание. Цепь Маркова представляет собой дальнейшее обобщение схемы Бернулли на случай зависимых испытаний; независимые испытания являются частным случаем марковской цепи. Под «событием» понимается состояние системы, а под «испытанием» – изменение ее состояния.

2.4.2. Дискретный марковский процесс с непрерывным временем. Плотности вероятностей перехода. Размеченный граф состояния

Если в некоторой системе S происходит марковский случайный процесс с дискретными состояниями $s_1, s_2, \ldots, s_n$ и переходы системы из состояния в состояние происходят в случайные моменты времени (что чаще встречается на практике), то такой процесс называется *марковским дискретным процессом с непрерывным временем.* Например, выход из строя (отказ) любого элемента аппаратуры может произойти в любой момент времени; окончание ремонта (восстановление) этого элемента также может произойти в заранее незафиксированный момент времени. Для описания таких процессов может быть применена схема марковского случайного процесса с дискретными состояниями и непрерывным временем, кратко называемой *непрерывной цепью Маркова.*

Пусть имеется ряд дискретных состояний: $S_1, S_2, \dots, S_n$ и переход (перескок) системы S из одного состояния в другое может осуществляться в любой момент времени.

В случае процесса с непрерывным временем не задаются переходные вероятности, так как вероятность перехода (перескока) системы из состояния s_i в состояние s_j точно в момент t будет равна нулю (как вероятность любого отдельно взятого значения непрерывной случайной величины). Вместо переходных вероятностей (2.12) вводят в рассмотрение плотности вероятностей перехода λ_{ij}.

Пусть система S в момент t находится в состоянии s_i. Рассмотрим элементарный промежуток времени Δt, примыкающий к моменту t (рис.17).

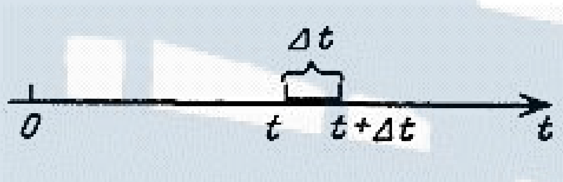


Рис. 17. Промежуток времени Δt состояния s_i.

Определение 2.23. *Плотностью вероятности перехода* λ_{ij} называется предел отношения вероятности перехода системы за время Δt из состояния s_i в состояние s_j к длине промежутка Δt:

$$\lambda_{ij} = \lim_{\Delta t \to 0} \frac{p_{ij}(\Delta t)}{\Delta t}, \tag{2.15}$$

где $p_{ij}(\Delta t)$ – вероятность того, что система, находившаяся в момент t в состоянии s_i, за время Δt перейдет из него в состояние $s_j, i \neq j$.

Из формулы (2.15) следует, что при малом Δt вероятность перехода $p_{ij}(\Delta t)$(с точностью до бесконечно малых высших порядков) равна $\lambda_{ij}\Delta t$:

$$p_{ij}(\Delta t) \approx \lambda_{ij}\Delta t. \tag{2.16}$$

Определение 2.24. Если все плотности вероятностей перехода λ_{ij} не зависят от t, то марковский процесс называется *однородным*; в противном случае – *неоднородным,* то есть если эти плотности представляют собой какие-то функции времени $\lambda_{ij}(t)$.

Предположим, что нам известны плотности вероятностей перехода $p_{ij}(\Delta t)$ для всех пар состояний (s_i, s_j).

Построим граф состояний системы S (рис. 18), где против каждой стрелки проставим соответствующую плотность вероятности перехода.

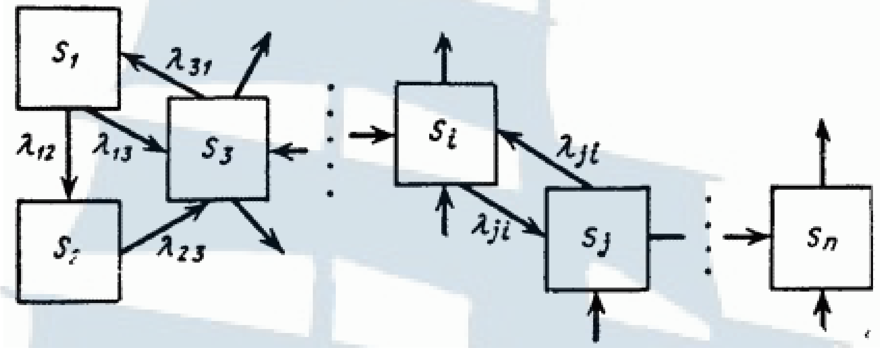


Рис. 18. Граф состояний

Такой граф называют *размеченным графом состояний*, зная который можно определить вероятности состояний как функции времени:

$$p_1(t), p_2(t), \dots, p_n(t). \qquad (2.17)$$

Эти вероятности удовлетворяют дифференциальным уравнениям, называемыми *уравнениями Колмогорова.* Решая эти уравнения, мы получим вероятности (2.17).

Пример 2.11. Пусть система S имеет четыре возможных состояния: S_1, S_2, S_3, S_4.

Размеченный граф состояний системы показан на рис.19.

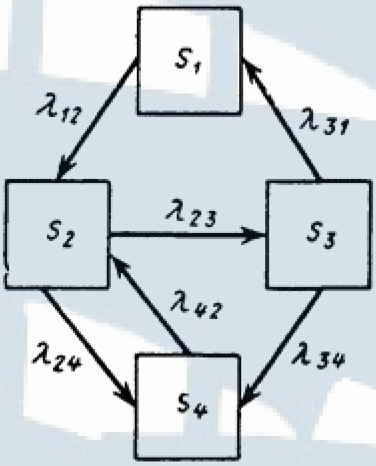


Рис. 19. Граф состояний

Найдем одну из вероятностей состояний, например, $p_1(t)$. Это есть вероятность того, что в момент t система будет находиться в состоянии S_1.

Придадим t малое приращение Δt и найдем вероятность того, что в момент система будет находиться в состоянии S_1. Это может произойти двумя способами: – в момент t система уже была в состоянии S_1 а за время Δt не вышла из этого состояния или в момент t система была в состоянии S_3, а за время Δt перешла из него в S_1. Вероятность первого варианта найдем как произведение вероятности $p_1(t)$ того, что в момент t система была в состоянии S_1 на условную вероятность того, что, будучи в состоянии S_1 система за время Δt не перейдет из него в S_2. Эта условная вероятность (с точностью до бесконечно малых высших порядков) равна $1 - \lambda_{12}\Delta t$.

Аналогично, вероятность второго варианта равна вероятности того, что в момент t система была в состоянии S_3, умноженной на условную вероятность перехода за время Δt в состояние S_1, то есть $p_3(t)\lambda_{31}\Delta t$.

Применяя правило сложения вероятностей, получим:

$$p_1(t + \Delta t) = p_1(t)(1 - \lambda_{12}\Delta t) + p_3(t)\lambda_{31}\Delta t.$$

Раскроем скобки в правой части, перенесем $p_1(t)$ в левую и разделим обе части равенства на Δt, получим:

$$\frac{p_1(t + \Delta t) - p_1(t)}{\Delta t} = -\lambda_{12}p_1(t) + \lambda_{31}p_3(t).$$

Теперь устремим Δt к нулю и перейдем к пределу:

$$\lim_{\Delta t \to 0} \frac{p_1(t + \Delta t) - p_1(t)}{\Delta t} = -\lambda_{12}p_1(t) + \lambda_{31}p_3(t).$$

Левая часть есть не что иное, как производная функции $p_1(t)$:

$$\frac{dp_1(t)}{dt} = -\lambda_{12}p_1(t) + \lambda_{31}p_3(t). \qquad (2.18)$$

Таким образом, дифференциальному уравнению (2.18) должна удовлетворять функция $p_1(t)$. Аналогичные дифференциальные уравнения могут быть выведены и для остальных вероятностей состояния: $p_2(t)$, $p_3(t)$, $p_4(t)$.

Рассмотрим второе состояние S_2 и найдем $p_2(t + \Delta t)$ – вероятность того, что в момент $t + \Delta t$ система S будет находиться в состоянии S_2. Это событие может произойти следующими способами:

– в момент t система уже была в состоянии S_2 а за время Δt не перешла из него ни в S_3, ни в S_4;

– в момент t система была в состоянии S_1 а за время Δt перешла из него в S_2;

– в момент t система была в состоянии S_4, а за время Δt перешла из него в S_2.

Вероятность первого варианта вычисляется так: $p_2(t)$ умножается на условную вероятность того, что система за Δt не перейдет ни S_3, ни в S_4.

Так как события, состоящие в переходе за время Δt из S_2 в S_3 и из S_2 в S_4 несовместны, то вероятность того, что осуществится один из этих переходов, равна сумме их вероятностей, то есть $\lambda_{23}\Delta t + \lambda_{24}\Delta t$ (с точностью до бесконечно малых высших порядков). Вероятность того, что не осуществится ни один их этих переходов, равна $1-\lambda_{23}\Delta t+\lambda_{24}\Delta t$. Отсюда вероятность первого варианта:

$$p_2(t)(1-\lambda_{23}\Delta t+\lambda_{24}\Delta t).$$

Прибавляя сюда вероятности второго и третьего вариантов, получим:

$$p_2(t+\Delta t)=p_2(t)(1-\lambda_{23}\Delta t+\lambda_{24}\Delta t)+p_1(t)\lambda_{12}\Delta t+p_4(t)\lambda_{42}\Delta t.$$

Перенося $p_2(t)$ в левую часть, деля на Δt и переходя к пределу, получим дифференциальное уравнение для $p_2(t)$:

$$\frac{dp_2(t)}{dt}=-\lambda_{23}p_2(t)-\lambda_{24}p_2(t)+\lambda_{12}p_1(t)+\lambda_{42}p_4(t). \qquad (2.19)$$

Рассуждая аналогично для состояний S_3 и S_4, получим в результате систему дифференциальных уравнений, составленных по типу (2.18) и (2.19). Отбросим в них для краткости аргумент t у функций p_1, p_2, p_3, p_4 и перепишем эту систему в виде:

$$\begin{cases} \frac{dp_1}{dt}=-\lambda_{12}p_1+\lambda_{31}p_3, \\ \frac{dp_2}{dt}=-\lambda_{23}p_2-\lambda_{24}p_2+\lambda_{12}p_1+\lambda_{42}p_4, \\ \frac{dp_3}{dt}=-\lambda_{31}p_3-\lambda_{34}p_3+\lambda_{23}p_2, \\ \frac{dp_4}{dt}=-\lambda_{42}p_4+\lambda_{24}p_2+\lambda_{34}p_3. \end{cases} \qquad (2.20)$$

Уравнения системы (2.20) называются *уравнениями Колмогорова*.

Интегрируя систему (2.20), получим искомые вероятности состояний как функции времени. Начальные условия берутся в зависимости от того, каково было начальное состояние системы S. Например, если в начальный момент времени (при $t=0$)система S находилась в состоянии S_1, то надо принять начальные условия:

При $t=0 \quad p_1=1, \quad p_2=p_3=p_4=0.$

Заметим, что все четыре уравнения для (2.20) можно было бы и не записывать, так как любую из вероятностей системы можно выразить через три остальные и подставить в другие уравнения.

Правило построения системы уравнений Колмогорова

В левой части каждого уравнения стоит производная вероятности состояния, а правая часть содержит столько членов, сколько стрелок связано с данным состоянием. Если стрелка направлена из состояния, соответствующий член имеет знак «минус»; если в состояние – знак плюс». Каждый член равен произведению плотности вероятности перехода, соответствующей данной стрелке, умноженной на вероятность того состояния, из которого исходит стрелка.

Это правило составления дифференциальных уравнений для вероятностей состояний является общим и справедливо для любой непрерывной марковской цепи; с его помощью можно механически записывать дифференциальные уравнения для вероятностей состояний непосредственно по размеченному графу состояний.

Пример 2.12. Размеченный граф состояний системы S имеет вид, показанный на рис.20. Написать систему дифференциальных уравнений Колмогорова и начальные условия для решения этой системы, если известно, что в начальный момент система находится в состоянии S_1.

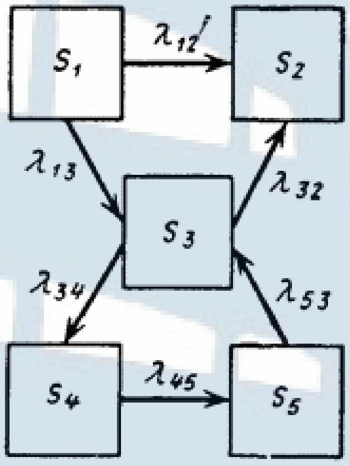


Рис. 20. Размеченный граф состояний

Решение. Система уравнений Колмогорова имеет вид:

$$\frac{dp_1}{dt} = -(\lambda_{12} + \lambda_{13})p_1;$$

$$\frac{dp_2}{dt} = \lambda_{12}p_1 + \lambda_{32}p_3;$$

$$\frac{dp_3}{dt} = -(\lambda_{32} + \lambda_{34})p_3 + \lambda_{13}p_1 + \lambda_{53}p_5;$$

$$\frac{dp_4}{dt} = -\lambda_{45}p_4 + \lambda_{34}p_3;$$

$$\frac{dp_5}{dt} = -\lambda_{53}p_5 + \lambda_{45}p_4.$$

Начальные условия: при $t = 0$ $p_1 = 1;$ $p_2 = p_3 = p_4 = p_5 = 0.$

Процессы указанного типа используются, в частности, для исследования реальных систем массового обслуживания (СМО); процессы в них протекают в непрерывном времени.

Марковский процесс служит моделью для многих процессов в биологии (распределение эпидемий, рост популяции), в физике (распад радиоактивного вещества), в теории массового обслуживания. В системах массового обслуживания множество состояний системы определяется числом каналов, то есть линий связи и т. д.; переходы между состояниями системы S происходят под воздействием потока событий (то есть потока заявок, требований, отказов и т. д.), которые являются простейшими, пуассоновскими.

2.4.3. Потоки событий: регулярные, ординарные, стационарные, без последействия. Интенсивность потока. Пуассоновский процесс

Наиболее распространенным марковским процессом является *пуассоновский процесс*. Этим процессом описываются такие величины, как число фотонов, испускаемых веществом при радиоактивном распаде, число телефонных вызовов из данного района, число происшествий на данном перекрестке, число заявок на обслуживание и др.

При рассмотрении случайных процессов, протекающих в системах с дискретными состояниями и непрерывным временем, часто приходится встречаться с так называемыми «потоками событий».

Определение 2.25. *Потоком событий* называется последовательность однородных событий, следующих одно за другим в какие-то случайные моменты времени.

Примеры: поток вызовов на телефонной станции; поток включений приборов в бытовой электросети; поток грузовых составов, поступающих на железнодорожную станцию; поток неисправностей (сбоев) вычислительной машины; поток выстрелов, направляемых на цель, и т. д.

Будем изображать поток событий последовательностью точек на оси времени t (рис.21), учитывая,что положение каждой точки на оси абсцисс случайно.

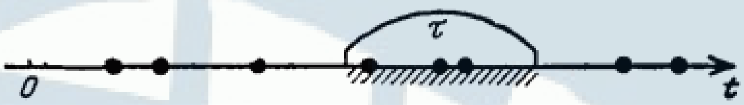


Рис. 21. Изображение потока событий

Определение 2.26. Поток событий называется *регулярным*, если собы-тия следуют одно за другим через строго определенные промежутки времени.

Такой поток сравнительно редко встречается на практике, чаще приходится встречаться с потоками событий, для которых и моменты наступления событий, и промежутки времени между ними случайны.

Определение 2.27. Поток событий называется *стационарным,* если вероятность попадания того или иного числа событий на участок времени длиной τ зависит только от длины участка и не зависит от того, где именно на оси $0t$ расположен этот участок.

Определение 2.28. Поток событий называется *потоком без последействия*, если для любых непересекающихся участков времени число событий, попадающих на один из них, не зависит от того, сколько событий попало на другой (или другие, если рассматривается больше двух участков).

Определение 2.29. Поток событий называется *ординарным,* если вероятность попадания на элементарный участок двух или более событий пренебрежимо мала по сравнению с вероятностью попадания одного события.

Стационарность потока означает его однородность по времени: вероятностные характеристики такого потока не должны меняться в зависимости от времени. В частности, интенсивность потока событий (среднее число событий в единицу времени) для стационарного потока должна оставаться постоянной. Это не значит, что фактическое число событий, появляющихся в единицу времени, постоянно, поток может иметь местные сгущения и разрежения. Важно, что для стационарного потока эти сгущения и разрежения не носят закономерного характера, а среднее число событий, попадающих на единичный участок времени, остается постоянным для всего рассматриваемого периода.

На практике часто встречаются потоки событий, которые на ограниченном участке времени могут рассматриваться как стационарные. Например, поток вызовов, поступающих на телефонную станцию на интервале от 13 до 14 часов, может считаться стационарным, хотя этот поток в течение целых суток уже не будет стационарным (ночью интенсивность потока вызовов гораздо меньше, чем днем). Заметим, что большинство физических процессов, которые мы называем «стационарными», в действительности стационарны только на ограниченном участке времени.

Отсутствие последействия в потоке означает, что события, образующие поток, появляются в последовательные моменты времени независимо друг от друга. Например, поток пассажиров, входящих на станцию метро, можно считать потоком без последействия, потому что причины, обусловившие приход отдельного пассажира именно в данный момент, а не в другой, как правило, не связаны с аналогичными причинами для других пассажиров. Если такая зависимость появляется, условие отсутствия последействия оказывается нарушенным.

Рассмотрим, например, поток грузовых поездов, идущих по железнодорожной ветке. Если, по условиям безопасности, они не могут следовать один за другим чаще, чем через интервал времени τ_0 то между событиями в потоке имеется зависимость, и условие отсутствия последействия нарушается. Если интервал τ_0 мал по сравнению со средним интервалом τ между поездами, такое нарушение несущественно, но если интервал τ сравним с τ_0, его приходится учитывать.

Ординарность потока означает, что события в потоке приходят поодиночке, а не парами, тройками и т. д. Например, поток клиентов, направляющихся в парикмахерскую, поток атак истребителей по бомбардировщику, если они атакуют цель поодиночке. Если в неординарном потоке события происходят только парами, только тройками и т. д., то можно его рассматривать как ординарный «поток пар», «поток троек» и т. д.

Пример неординарного потока событий – поток товарных вагонов, прибывающих на сортировочную станцию случайным образом.

Определение 2.30. Поток событий, не имеющий последействия, который ординарен, но не стационарен, называется *нестационарным пуассоновским потоком.* В таком потоке интенсивность λ (среднее число событий в единицу времени) зависит от времени:

$$\lambda = \lambda(t),$$

тогда как для простейшего потока

$$\lambda = const.$$

Пуассоновский поток событий (как стационарный, так и нестационарный) тесно связан с *распределением Пуассона.* А именно, число событий потока, попадающих на любой участок, распределено по закону Пуассона.

Рис. 22. Участок времени длины, τ

Рассмотрим на оси 0t, где наблюдается поток событий, некоторый участок времени длины τ (рис. 22), начинающийся в момент t_0 и заканчивающийся в момент $t_0 + \tau$. Из курса теории вероятностей известно, что вероятность попадания на этот участок ровно m событий выражается формулой

$$P_m = \frac{a^m}{m!} e^{-a} \; (m = 0,1, \dots),$$

где *a* – среднее число событий, приходящееся на участок τ.

Для стационарного (простейшего) пуассоновского потока величина a равна интенсивности потока, умноженной на длину интервала: $a = \lambda\tau$, то есть не зависит от того, где на оси 0t взят участок τ.

Для нестационарного пуассоновского потока величина *a* выражается формулой

$$a = \int_{t_0}^{t_0+\tau} \lambda(t) dt, \tag{2.21}$$

следовательно, зависит от того, в какой точке t_0 начинается участок τ.

Рассмотрим на оси 0t простейший поток событий с интенсивностью λ (рис. 23). Нас будет интересовать интервал времени T между соседними событиями в этом потоке. Очевидно, Т есть величина случайная; найдем ее закон распределения. Сначала найдем функцию распределения

$$F(t) = P(T < t).$$

Отложим от начала интервала Т (точки t_0) отрезок t и найдем вероятность того, что интервал Т будет меньше *t*. Для этого нужно, чтобы на участок длины *t*, примыкающий к точке t_0 попало *хотя бы одно* событие потока. Через вероятность противоположного события (на участок t не попадет ни одного события потока) получим:

$$F(t) = 1 - P_0, \quad \text{где} \quad P_0 = \frac{a^0}{0!}e^{-a} = e^{-a} = e^{-\lambda t},$$

откуда функция распределения величины Т будет:

$$F(t) = 1 - e^{-\lambda t}(t > 0). \qquad (2.22)$$

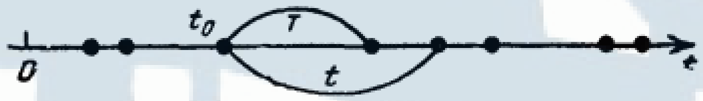


Рис. 23. Простейший поток событий с интенсивностью λ

Чтобы найти плотность распределения *f(t)* случайной величины Т, продифференцируем выражение (2.22) по *t:*

$$f(t) = \lambda e^{-\lambda t} \ (t > 0). \qquad (2.23)$$

Закон распределения с плотностью (2.23) называется *показательным* или *экспоненциальным с параметром* λ , где λ интенсивность потока (см. стр. 72–75 [4]).

Таким образом, промежуток времени Т между соседними событиями в *простейшем потоке распределен по показательному закону.* Для нестационарного пуассоновского потока закон распределения промежутка Т уже не будет показательным. Вид этого закона будет зависеть, во-первых, от того, где на оси 0*t* расположено первое из событий, и, во-вторых, от вида зависимости $\lambda(t)$, характеризующей интенсивность потока. Если $\lambda(t)$ меняется сравнительно медленно, и его изменение за время между двумя событиями невелико, то закон распределения промежутка времени между событиями можно приближенно считать показательным.

Рассмотрим на оси 0*t* простейший поток событий с интенсивностью λ и элементарный участок Δt, прилежащий в точке t (рис. 23).

Найдем вероятность того, что на участке Δt появится какое-то событие потока, то есть участок не будет «пуст». Так как поток ординарен, вероятностью появления на участке Δt более чем одного события можно пренебречь. Обозначим $P_0(\Delta t)$ вероятность того, что на участке Δt не будет события, $P_1(\Delta t)$ – вероятность того, что на нем появится одно событие.

В силу ординарности потока

$$P_1(\Delta t) \approx 1 - P_0(\Delta t),$$

где вероятность

$$P_0(\Delta t) = \frac{a^0}{0!} e^{-a} = e^{-a} = e^{-\lambda \Delta t},$$

откуда

$$P_1(\Delta t) \approx 1 - e^{-\lambda \Delta t}.$$

Разлагая $e^{-\lambda \Delta t}$ в ряд по степеням $\lambda \Delta t$ и пренебрегая величинами высшего порядка малости, получим

$$P_1(\Delta t) \approx 1 - (1 - \lambda \Delta t).$$

Отсюда

$$P_1(\Delta t) \approx \lambda \Delta t, \tag{2.24}$$

то есть вероятность появления на элементарном участке времени Δt какого-то события потока приближенно равна $\lambda \Delta t$, гдеλ – интенсивность потока. Эту вероятность мы будем называть «элементом вероятности появления события».

Очевидно, такая же формула будет справедлива и для нестационарного пуассоновского потока, с той разницей, что величину λ нужно брать равной ее значению в той точке t, к которой примыкает участок Δt: $P_1(\Delta t) \approx \lambda(t) \Delta t$.

2.4.4. Пуассоновские потоки событий и непрерывные марковские цепи

Рассмотрим некоторую физическую систему S с дискретными состояниями $S_1, S_2, \dots, S_n$, которая переходит из состояния в состояние под влиянием случайных событий, представляющих собой какие-то потоки событий (потоки вызовов, потоки отказов, потоки выстрелов и т. д.).

Пусть система S с графом состояний, показанным на рис. 24, в момент t находится в состоянии S_i и может перейти из него в состояние S_j под влиянием какого-то пуассоновского потока событий с интенсивностью λ_{ij}.

Вероятность этого перехода за элементарный промежуток времени Δt равна $\lambda_{ij} \Delta t$.

Таким образом, плотность вероятности перехода λ_{ij} в непрерывной цепи Маркова представляет собой интенсивность потока событий, переводящего систему по соответствующей стрелке.

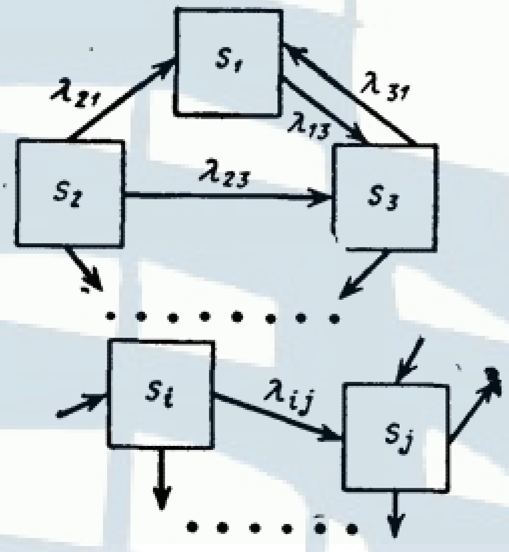


Рис. 24. Граф состояний

Если все потоки событий, переводящие систему S из состояния в состояние, *пуассоновские,* то процесс, протекающий в системе, будет *марковским*. Действительно, пуассоновский поток обладает отсутствием последействия, поэтому, при заданном состоянии системы в данный момент, ее переходы в другие состояния в будущем обусловлены только появлением каких-то событий в пуассоновских потоках, а вероятности появления этих событий не зависят от «предыстории» процесса.

В дальнейшем, рассматривая марковские процессы в системах с дискретными состояниями и непрерывным временем (непрерывные марковские цепи), удобно рассматривать переходы системы из состояния в состояние как происходящие под влиянием каких-то потоков событий, хотя бы в действительности эти события были единичными. Например, работающее техническое устройство мы будем рассматривать как находящееся под действием потока отказов, хотя фактически оно может отказать только один раз. Действительно, если устройство отказывает в тот момент, когда приходит первое событие потока, то совершенно все равно – продолжается после этого поток отказов или же прекращается.

Итак, рассматривается система S, в которой переходы из состояния в состояние происходят под действием пуассоновских потоков событий с определенными интенсивностями. Проставим эти интенсивности (плотности вероятностей переходов) на графе состояний системы у соответствующих стрелок. Получим размеченный граф состояний (рис. 24), по которому, пользуясь правилом, сформулированным в п. 2.3.2, можно сразу записать дифференциальные уравнения Колмогорова для вероятностей состояний.

Пример 2.13. Техническая система S состоит из двух узлов: I и II; каждый из них независимо от другого может отказывать (выходить из строя). Поток отказов первого узла – пуассоновский, с интенсивностью – λ_{I} второго – также пуассоновский, с интенсивностью λ_{II}. Каждый узел сразу после отказа начинает ремонтироваться (восстанавливаться). Поток восстановлений для обоих узлов – пуассоновский с интенсивностью λ.

Составить уравнения Колмогорова и определить при каких начальных условиях нужно решать эти уравнения, если в начальный момент ($t = 0$)система работает исправно.

Решение. Возможные состояния системы: S_{11} – оба узла исправны; S_{21} – первый узел ремонтируется, второй исправен; S_{12} – первый узел исправен, второй ремонтируется; S_{22} – оба узла ремонтируются.
Размеченный граф состояний системы показан на рис.25.

Интенсивности потоков событий на рис. 25 проставлены из следующих соображений. Если система S находится в состоянии S_{11}, то на нее действуют два потока событий: поток неисправностей узла I с интенсивностью λ_{I}, переводящий ее в состояние S_{21} и поток неисправностей узла II с интенсивностью λ_{II} переводящий ее в S_{12}.

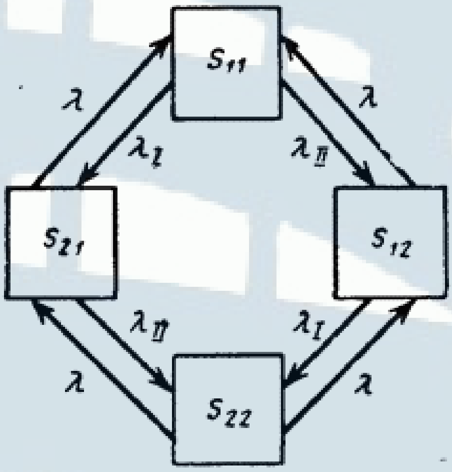


Рис. 25. Интенсивности потоков событий

Пусть теперь система находится в состоянии S_{21} (узел I ремонтируется, узел II – исправен). Из этого состояния система может вернуться в S_{11} (это происходит под действием потока восстановлений с интенсивностью λ) или перейти в состояние S_{22} (когда ремонт узла I еще не закончен, а узел II тем временем вышел из строя; этот переход происходит под действием потока отказов узла II с интенсивностью λ_{II}.

Обозначая вероятности состояний $p_{11}, p_{12}, p_{21}, p_{22,}$ запишем уравнения Колмогорова для вероятностей состояний:

$$\frac{dp_{11}}{dt} = -(\lambda_{\mathrm{I}} + \lambda_{\mathrm{II}})p_{11} + \lambda p_{21} + \lambda p_{12};$$

$$\frac{dp_{21}}{dt} = -(\lambda + \lambda_{\mathrm{II}})p_{21} + \lambda_{\mathrm{I}} p_{11} + \lambda p_{22};$$

$$\frac{dp_{12}}{dt} = -(\lambda + \lambda_{\mathrm{I}})p_{12} + \lambda_{\mathrm{II}} p_{11} + \lambda p_{22};$$

$$\frac{dp_{22}}{dt} = -2\lambda p_{22} + \lambda_{\mathrm{II}} p_{21} + \lambda_{\mathrm{I}} p_{12}.$$

Начальные условия, при которых нужно решать эту систему:
при $t = 0 \quad p_{11} = 1; \quad p_{12} = p_{21} = p_{22} = 0$.

Заметим, что, пользуясь условием: $p_{11} + p_{12} + p_{21} + p_{22} = 1$, можно было бы уменьшить число уравнений на одно, так как любую из этих вероятностей можно выразить через остальные и подставить в уравнения системы.

Пример 2.14. Группа в составе пяти самолетов в строю «колонна» (рис. 26) совершает налет на территорию противника. Передний самолет (ведущий) является постановщиком помех; до тех пор, пока он не сбит, идущие за ним самолеты не могут быть обнаружены и атакованы средствами ПВО противника. Атакам подвергается только постановщик помех. Поток атак – пуассоновский, с интенсивностью λ (атак/час). В результате атаки постановщик помех поражается с вероятностью р.

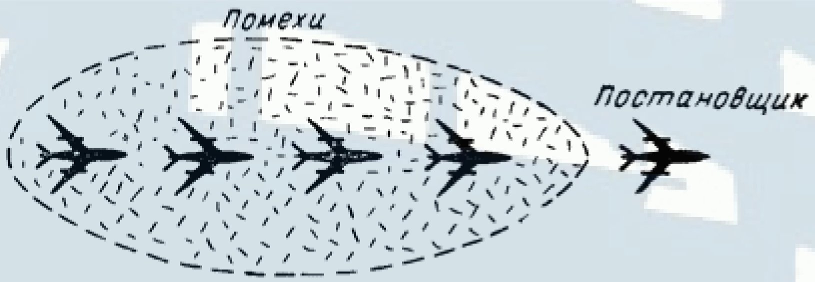


Рис. 26. Колонна из пяти самолетов

Если постановщик помех поражен (сбит), то следующие за ним самолеты обнаруживаются и подвергаются атакам ПВО; на каждый из них направляется пуассоновский поток атак с интенсивностью λ; каждой атакой самолет поражается с вероятностью р. Когда самолет поражен, атаки по нему прекращаются, но на другие самолеты не переносятся.

Написать уравнения Колмогорова для вероятностей состояний системы и указать начальные условия.

Решение. Введем обозначения для состояний системы: S_5 – все самолеты целы; S_4 – постановщик помех сбит, остальные самолеты целы; S_3 – постановщик помех и один бомбардировщик сбиты, остальные самолеты целы; S_2 – постановщик помех и два бомбардировщика сбиты, остальные самолеты целы; S_1 – постановщик помех и три бомбардировщика сбиты, один самолет цел; S_0– все самолеты сбиты.

Граф состояний системы показан на рис.27. Чтобы разметить этот граф, определим интенсивности потоков событий, переводящих систему из состояния в состояние

Из состояния S_5 систему переводит поток поражающих (или «успешных») атак, то есть тех атак, которые приводят к поражению постановщика (разумеется, если он раньше не был поражен).

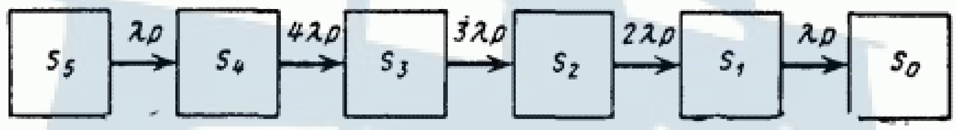


Рис. 27. Граф состояний системы

Интенсивность потока атак равна λ, но не все они – поражающие: каждая из них оказывается поражающей только с вероятностью p. Очевидно, интенсивность потока поражающих атак равна λp; эта интенсивность и проставлена в качестве λ_{54} у первой слева стрелки на графе (рис.27).

Найдем интенсивность λ_{43}. Система находится в состоянии S_4, то есть целы и могут быть атакованы четыре самолета. Она перейдет в состояние S_3 за время Δt, если за это время какой-нибудь из самолетов будет сбит. Найдем вероятность противоположного события – за время Δt ни один самолет не будет сбит:

$$(1-\lambda p\Delta t)(1-\lambda p\Delta t)(1-\lambda p\Delta t)(1-\lambda p\Delta t) =$$
$$= (1-\lambda p\Delta t)^4 \approx 1-4\lambda p. \qquad (2.25)$$

Здесь отброшены члены высшего порядка малости относительно Δt. Вычитая эту вероятность из единицы, получим вероятность перехода из S_4 за время Δt: $4\lambda p\Delta t$, откуда получим: $\lambda_{43} = 4\lambda p$.

Заметим, что интенсивность этого потока событий просто равна сумме интенсивностей потоков поражающих атак, направленных на отдельные самолеты. Уравнения Колмогорова вероятностей состояний имеют вид:

$$\frac{dp_5}{dt} = -\lambda p p_5;$$

$$\frac{dp_4}{dt} = -4\lambda p p_4 + \lambda p p_5;$$

$$\frac{dp_3}{dt} = -3\lambda p p_3 + 4\lambda p p_4;$$

$$\frac{dp_2}{dt} = -2\lambda p p_2 + 3\lambda p p_3;$$

$$\frac{dp_1}{dt} = -\lambda p p_1 + 2\lambda p p_2;$$

$$\frac{dp_0}{dt} = \lambda p p_1.$$

Так как в начальный момент (при $t = 0$) все самолеты целы, начальные условия будут таковы:

$$\text{при } t = 0 \quad p_5 = 1; \quad p_0 = p_1 = p_2 = p_3 = p_4 = 0.$$

2.4.5. Предельные вероятности состояний

Пусть имеется физическая система S с дискретными состояниями: $S_1, S_2, \ldots, S_n$, являющаяся цепью Маркова (рис. 28).

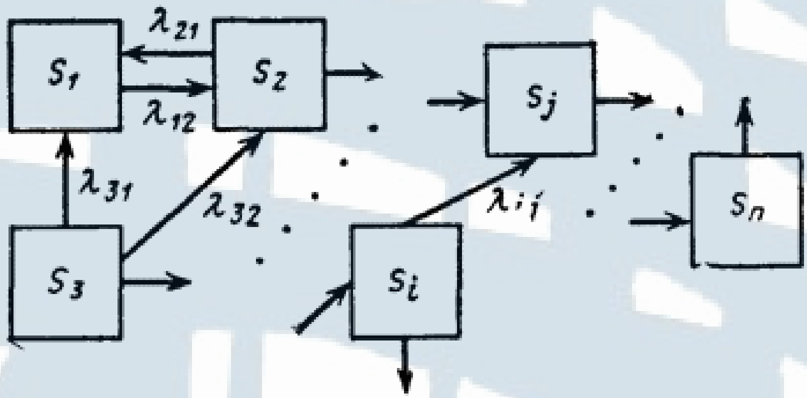


Рис. 28. Граф состояний системы

Предположим, что все интенсивности потоков событий, переводящих систему из состояния в состояние, постоянны:

$$\lambda_{ij} = const,$$

то есть все потоки событий – простейшие (стационарные пуассоновские) потоки.

Записав систему дифференциальных уравнений Колмогорова для вероятностей состояний и проинтегрировав эти уравнения при заданных начальных условиях, мы получим вероятности состояний как функции времени:

$$p_1(t), p_2(t), \ldots, p_n(t) \quad \text{и} \quad \sum_{i=1}^{n} p_i(t) = 1.$$

Определение 2.31. Если существуют пределы:

$$\lim_{t \to \infty} p_i(t) = p_i \ (i = 1,2, \ldots n), \qquad (2.26)$$

то они называются *предельными вероятностями* состояний.

Теорема (без доказательства): если число состояний системы S конечно и из каждого состояния можно перейти (за то или иное число шагов) в каждое другое, то предельные вероятности состояний существуют и не зависят от начального состояния системы.

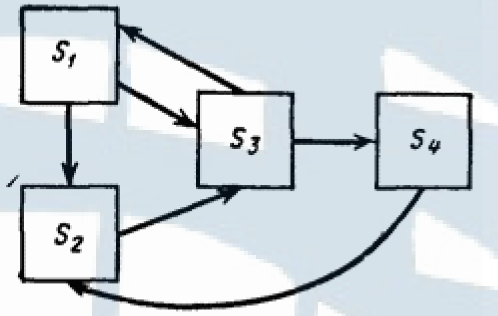


Рис. 29. Граф состояний, удовлетворяющий условию теоремы

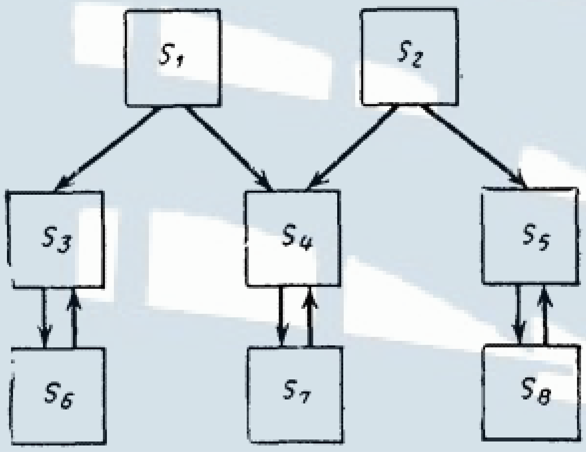


Рис. 30. Граф состояний, где условие теоремы не выполнено

Предположим, что предельные вероятности существуют:

$$\lim_{t\to\infty} p_i(t) = p_i \ (i = 1,2, \dots n), \text{где } \sum_{i=1}^{n} p_i(t) = 1.$$

Таким образом, при $t \to \infty$ в системе S устанавливается некоторый предельный стационарный режим, состоящий в том, что система случайным образом меняет свои состояния, но вероятность каждого из них уже не зависит от времени. Каждое из состояний осуществляется с некоторой постоянной вероятностью, которая представляет собой среднее относительное время пребывания системы в данном состоянии. Пусть у системы S три возможных состояния: S_1, S_2, S_3 и их предельные вероятности равны 0,2, 0,3 и 0,5. Это означает, что после перехода к установившемуся режиму система S в среднем две десятых времени будет находиться в состоянии S_1; три десятых – в состоянии S_2 и половину времени в состоянии S_3.

Так как в предельном (установившемся) режиме все вероятности состояний постоянны, то их производные равны нулю. Поэтому для нахождения предельных вероятностей нужно положить все левые части уравнений Колмогорова (производные) равными нулю. Тогда система дифференциальных уравнений превратится в систему линейных алгебраических уравнений. Совместно с условием

$$\sum_{i=1}^{n} p_i = 1$$

(так называемым «нормировочным условием») эти уравнения дают возможность вычислить все предельные вероятности: $p_1, p_2, \dots, p_3$.

Пример 2.15. Физическая система 5 имеет возможных состояний: S_1, S_2, S_3, S_4, S_5; размеченный граф которых дан на рис. 31. Вычислить предельные вероятности состояний.

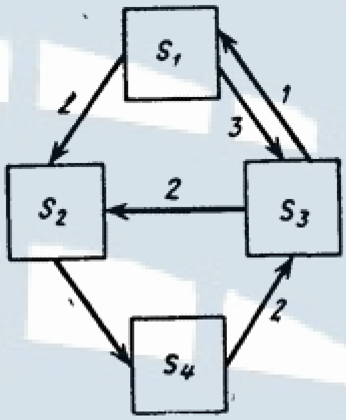


Рис. 31. Размеченный граф

Решение. Пишем уравнения Колмогорова для вероятностей состояний (2.27):

$$\begin{cases} \frac{dp_1}{dt} = -5p_1 + p_3, \\ \frac{dp_2}{dt} = 2p_1 - p_2 + 2p_3, \\ \frac{dp_3}{dt} = 3p_1 - 3p_3 + 2p_4, \\ \frac{dp_4}{dt} = -2p_4 + p_2. \end{cases} \quad (2.27)$$

Полагая левые части равными нулю, получим систему алгебраических уравнений для предельных вероятностей состояний:

$$\left.\begin{aligned} 0 &= -5p_1 + p_3, \\ 0 &= -p_2 + 2p_1 + 2p_3, \\ 0 &= -3p_3 + 3p_1 + 2p_4, \\ 0 &= -2p_4 + p_2. \end{aligned}\right\} \qquad (2.28)$$

Уравнения (2.28) – однородные уравнения, из которых величины p_1, p_2, p_3, p_4 определяются только с точностью до постоянного множителя. Нормировочное условие:

$$p_1 + p_2 + p_3 + p_4 = 1,$$

совместно с уравнениями (2.28), дает возможность найти все неизвестные вероятности.

Действительно, выразим из (2.28) все неизвестные вероятности через одну из них, например, через p_1. Из первого уравнения:

$$p_3 = 5p_1.$$

Подставляя во второе уравнение, получим:

$$p_2 = 2p_1 + 2p_3 = 2p_1 + 10p_1 = 12p_1.$$

Четвертое уравнение дает:

$$p_4 = {}^1/_2\, p_2 = 6p_1.$$

Подставляя все эти выражения вместо $\mathrm{p}_2, \mathrm{p}_3, \mathrm{p}_4$ в нормировочное условие, получим:

$$p_1 + 12p_1 + 5p_1 + 6p_1 = 1.$$

Отсюда

$$24p_1 = 1;\ \ p_1 = {}^1/_{24};\ \ p_2 = {}^1/_2\,;\ \ p_3 = {}^5/_{24}\,,;\ \ p_4 = {}^1/_4.$$

Это значит, что в предельном, установившемся режиме система S будет проводить в состоянии S_1 в среднем одну двадцать четвертую часть времени, в состоянии S_2 – половину времени, в состоянии S_3 – пять двадцать четвертых и в состоянии S_4 – одну четверть времени.

Предельные вероятности можно найти, записав алгебраические уравнения для предельных вероятностей непосредственно, не проходя через этап дифференциальных. Проиллюстрируем это на примере.

Пример 2.16. Граф состояний системы показан на рис. 32. Написать алгебраические уравнения для предельных вероятностей состояний.

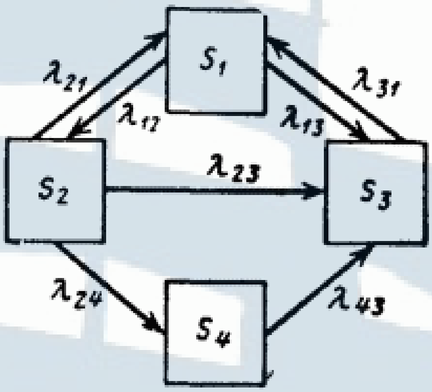


Рис. 32. Граф состояний

Решение. Запишем соответствующие правые части, приравниваем их к нулю и перенося их в другую часть, получим:

$$\begin{cases} \lambda_{21}p_2 + \lambda_{31}p_3 = (\lambda_{12} + \lambda_{13})p_1, \\ \lambda_{12}p_1 = (\lambda_{23} + \lambda_{24})p_2, \\ \lambda_{13}p_1 + \lambda_{23}p_2 + \lambda_{43}p_4 = \lambda_{31}p_3, \\ \lambda_{24}p_2 = \lambda_{43}p_4. \end{cases} \tag{2.29}$$

Правило: для каждого состояния сумма членов, соответствующих входящим стрелкам, равна сумме членов, соответствующих выходящим стрелкам. Каждый член равен интенсивности потока событий, переводящего систему по данной стрелке, умноженной на вероятность того состояния, из которого выходит стрелка.

Пример 2.17. Написать алгебраические уравнения для предельных вероятностей состояний системы S, граф состояний которой дан на рис. 33. Решить эти уравнения.

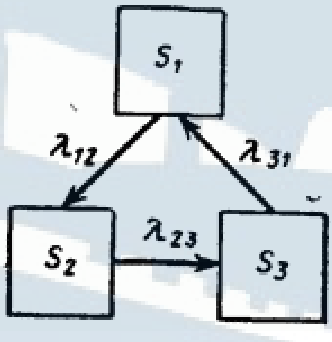


Рис. 33. Граф состояний

Решение. Запишем алгебраические уравнения для предельных вероятностей состояний:

$$\lambda_{12}p_1 = \lambda_{23}p_2; \;\; \lambda_{31}p_3 = \lambda_{12}p_1; \;\; \lambda_{31}p_3 = \lambda_{23}p_2. \tag{2.30}$$

Нормировочное условие:

$$p_1 + p_2 + p_3 = 1. \qquad (2.31)$$

Выразим с помощью первых двух уравнений p_3, p_2 через p_1:

$$p_3 = \frac{\lambda_{12}}{\lambda_{31}} p_1; \;\; p_2 = \frac{\lambda_{12}}{\lambda_{23}} p_1. \qquad (2.32)$$

Подставим их в нормировочное условие

$$p_1 + \frac{\lambda_{12}}{\lambda_{31}} p_1 + \frac{\lambda_{12}}{\lambda_{23}} p_1 = 1,$$

откуда

$$p_1 = \frac{1}{1 + \frac{\lambda_{12}}{\lambda_{23}} + \frac{\lambda_{12}}{\lambda_{31}}} = 1.$$

Далее, из (2.32) получим:

$$p_2 = \frac{\frac{\lambda_{12}}{\lambda_{23}}}{1 + \frac{\lambda_{12}}{\lambda_{23}} + \frac{\lambda_{12}}{\lambda_{31}}}; \quad p_3 = \frac{\frac{\lambda_{12}}{\lambda_{31}}}{1 + \frac{\lambda_{12}}{\lambda_{23}} + \frac{\lambda_{12}}{\lambda_{31}}}.$$

2.4.6. Процессы «гибели и размножения»

В предыдущем параграфе мы убедились, что зная размеченный граф состояний системы, можно сразу написать алгебраические уравнения для предельных вероятностей состояний.

Таким образом, если две непрерывные цепи Маркова имеют одинаковые графы состояний и различаются только значениями интенсивностей λ_{ij}, то не надо находить предельные вероятности состояний для каждого из графов в отдельности: достаточно составить и решить в буквенном виде уравнения для одного из них, а затем подставить вместо λ_{ij} соответствующие значения для другого графа.

Определение 2.32. Марковская непрерывная цепь называется «*процессом гибели и размножения»*, если все ее состояния можно вытя-нуть в одну цепочку, в которой каждое из средних состояний $(S_2, \ldots, S_{n-1})$ связано прямой и обратной связью с каждым из соседних состояний, а крайние состояния (S_1, S_n) – только с одним соседним состоянием (рис. 34).

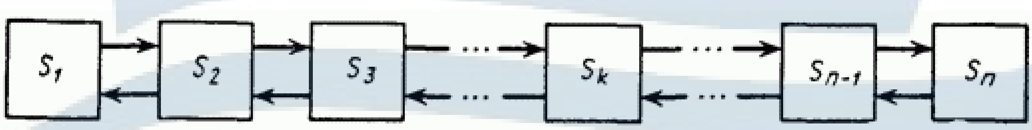


Рис. 34. Граф состояний

Пример 2.18. Техническое устройство состоит из трех одинаковых узлов; каждый из них может выходить из строя (отказывать); отказавший узел немедленно начинает восстанавливаться.

Состояния системы нумеруем по числу неисправных узлов: S_0 – все три узла исправны; S_1 – один узел отказал (восстанавливается), два исправны; S_2 – два узла восстанавливаются, один исправен; S_3 – все три узла восстанавливаются.

Из графа состояний (рис. 35) видно, что процесс, протекающий в системе, представляет собой процесс «гибели и размножения».

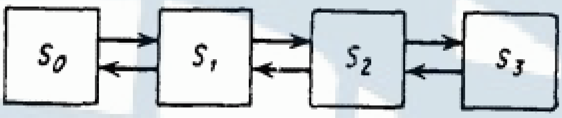


Рис. 35. Граф состояний

Схема гибели и размножения очень часто встречается в самых разнообразных практических задачах, поэтому рассмотрим эту схему в общем виде и решим соответствующую систему алгебраических урав-нений с тем, чтобы в дальнейшем, встречаясь с конкретными процессами, протекающими по такой схеме, пользоваться уже готовым решением.

Итак, рассмотрим случайный процесс гибели и размножения с графом состояний (рис. 36).

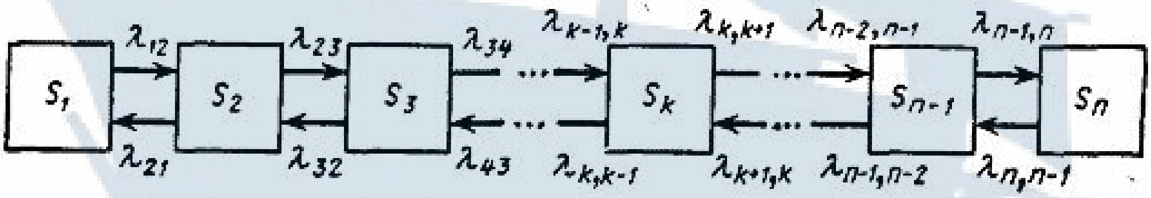


Рис. 36. Граф состояний

Напишем алгебраические уравнения для вероятностей состояний. Для первого состояния S_1 имеем:

$$\lambda_{12} p_1 = \lambda_{21} p_2. \qquad (2.33)$$

Для второго состояния S_2 суммы членов, соответствующих входящим и выходящим стрелкам, равны:

$$\lambda_{32} p_3 + \lambda_{12} p_1 = \lambda_{21} p_2 + \lambda_{23} p_2.$$

В силу (2.33) имеем:

$$\lambda_{32} p_3 = \lambda_{23} p_2.$$

Далее аналогично получим:

$$\lambda_{k-1,k}p_{k-1} = \lambda_{k,k-1}p_k, \quad (2.34)$$

где k принимает все значения от 2 до n.

Таким образом, для схемы гибели и размножения члены, соответствующие стоящим друг над другом стрелкам, равны между собой, а предельные вероятности состояний $S_1, \dots, S_n$ в любой схеме гибели и размножения удовлетворяют уравнениям:

$$\lambda_{12}p_1 = \lambda_{21}p_2;\ \lambda_{32}p_3 = \lambda_{23}p_2;\ \lambda_{32}p_3 = \lambda_{23}p_2; \dots; \lambda_{k-1,k}p_{k-1} = \lambda_{k,k-1}p_k; \dots;$$
$$\lambda_{n-1,n}p_{n-1} = \lambda_{n,n-1} \quad (2.35)$$

и нормировочному условию:

$$p_1 + p_2 + p_3 + \dots + p_n = 1.$$

Будем решать эту систему следующим образом: из первого уравнения системы (2.29) выразим p_2:

$$p_2 = \frac{\lambda_{12}}{\lambda_{21}}p_1, \quad (2.36)$$

из второго, с учетом (2.36), получим:

$$p_3 = \frac{\lambda_{23}}{\lambda_{32}}p_2, p_3 = \frac{\lambda_{23}}{\lambda_{32}}\frac{\lambda_{12}}{\lambda_{21}}p_1, \quad (2.37)$$

из третьего, с учетом (2.37):

$$p_4 = \frac{\lambda_{34}}{\lambda_{43}}\frac{\lambda_{23}}{\lambda_{32}}\frac{\lambda_{12}}{\lambda_{21}}p_1, \quad (2.38)$$

$$p_k = \frac{\lambda_{k-1,k}}{\lambda_{k,k-1}}\frac{\lambda_{k-2,k-1}\dots}{\lambda_{k-1,k-2}\dots}\frac{\lambda_{12}}{\lambda_{21}}p_1. \quad (2.39)$$

Формула (2.39) справедлива для любого k от 2 до n. Ее структура: в числителе стоит произведение всех плотностей вероятности перехода (интенсивностей) λ_{ij}, стоящих у стрелок, направленных *слева направо*, с начала и до состояния S_k; в знаменателе – произведение всех интенсивностей λ_{ij}, стоящих у стрелок, идущих *справа налево*, с начала и до стрелки, исходящей из состояния S_k. При $n = k$ в числителе будет стоять произведение интенсивностей λ_{ij}, стоящих *у всех* стрелок, идущих *слева направо*, а в знаменателе – *у всех* стрелок, идущих *справа налево*.

Итак, все вероятности $p_1, p_2, \dots, p_n$ выражены через одну из них – p_1. Подставим эти выражения в нормировочное условие:

$$p_1 + p_2 + p_3 + \dots + p_n = 1.$$

Получим:

$$p_1 + \frac{\lambda_{12}}{\lambda_{21}} p_1 + \frac{\lambda_{23}}{\lambda_{32}} \frac{\lambda_{12}}{\lambda_{21}} p_1 + \cdots + \frac{\lambda_{k-1,k}}{\lambda_{k,k-1}} \frac{\lambda_{k-2,k-1} \cdots \lambda_{12}}{\lambda_{k-1,k-2} \cdots \lambda_{21}} p_1 + \cdots +$$

$$+ \frac{\lambda_{n-1,n}}{\lambda_{n,n-1}} \frac{\lambda_{n-2,n-1} \cdots \lambda_{12}}{\lambda_{n-1,n-2} \cdots \lambda_{21}} p_1 = 1,$$

откуда

$$p_1 = \frac{1}{1 + \frac{\lambda_{12}}{\lambda_{21}} + \frac{\lambda_{23}}{\lambda_{32}} \frac{\lambda_{12}}{\lambda_{21}} + \cdots + \frac{\lambda_{n-1,n}}{\lambda_{n,n-1}} \frac{\lambda_{n-2,n-1} \cdots \lambda_{12}}{\lambda_{n-1,n-2} \cdots \lambda_{21}}}.$$

Остальные вероятности выражаются через p_1:

$$p_2 = \frac{\lambda_{12}}{\lambda_{21}} p_1;$$

$$p_3 = \frac{\lambda_{23}}{\lambda_{32}} \frac{\lambda_{12}}{\lambda_{21}} p_1;$$

$$p_4 = \frac{\lambda_{34}}{\lambda_{43}} \frac{\lambda_{23}}{\lambda_{32}} \frac{\lambda_{12}}{\lambda_{21}} p_1; \ldots;$$

$$p_k = \frac{\lambda_{k-1,k}}{\lambda_{k,k-1}} \frac{\lambda_{k-2,k-1} \cdots \lambda_{12}}{\lambda_{k-1,k-2} \cdots \lambda_{21}} p_1; \ldots;$$

$$p_n = \frac{\lambda_{n-1,n}}{\lambda_{n,n-1}} \frac{\lambda_{n-2,n-1} \cdots \lambda_{12}}{\lambda_{n-1,n-2} \cdots \lambda_{21}} p_1.$$

Таким образом, задача «гибели и размножения» в общем виде решена: предельные вероятности состояний найдены.

Пример 2.19. Найти предельные вероятности состояний для процесса гибели и размножения (рис. 37).

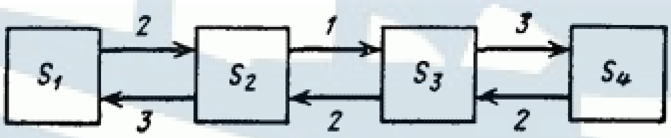


Рис. 37. Граф состояний

Решение. По формулам (2.39) имеем:

$$p_1 = \frac{1}{1 + \frac{2}{3} + \frac{2 \cdot 1}{3 \cdot 2} + \cdots + \frac{2 \cdot 1 \cdot 3}{3 \cdot 2 \cdot 2}} = \frac{2}{5};$$

$$p_2 = \frac{2}{3} \cdot \frac{2}{5} = \frac{4}{15}, p_3 = \frac{1}{3} \cdot \frac{2}{5} = \frac{2}{15}, p_4 = \frac{1}{2} \cdot \frac{2}{5} = \frac{1}{15}.$$

ГЛАВА 3. Задачи теории массового обслуживания. Классификация систем массового обслуживания. Логистические потоки

На практике очень часто приходится сталкиваться с анализом работы своеобразных систем, называемых *системами массового обслуживания* (СМО). Примерами таких систем могут служить: телефонные станции, ремонтные мастерские, билетные кассы, справочные бюро, магазины, парикмахерские и т. п.

Каждая СМО состоит из какого-то числа обслуживающих единиц, которые мы будем называть *каналами обслуживания.* Ими могут быть линии связи, рабочие точки, приборы, железнодорожные пути, лифты, автомашины и т. д.

Системы массового обслуживания могут быть *одноканальными* или *многоканальными.* Каждая СМО предназначена для обслуживания какого-то потока заявок, поступающих на СМО в случайные моменты времени. Обслуживание поступившей заявки продолжается некоторое время, после чего канал освобождается и готов к принятию следующей заявки. Случайный характер потока заявок приводит к тому, что в какие-то промежутки времени на входе СМО скапливается излишне большое число заявок (они либо образуют очередь, либо покидают СМО необслуженными); в другие же периоды СМО будет работать с недогрузкой или вообще простаивать.

Каждая система массового обслуживания, в зависимости от числа каналов и их производительности, а также от характера потока заявок, обладает какой-то *пропускной способностью*, позволяющей ей более или менее успешно справляться с потоком заявок. Предмет теории массового обслуживания – установление зависимости между характером потока заявок, числом каналов, их производительностью, правилами работы СМО и успешностью (эффективностью) обслуживания.

В качестве *характеристик эффективности обслуживания* (в зависимости от условий задачи и целей исследования) применяют различные величины и функции:

– среднее количество заявок, которое может обслужить СМО в единицу времени;

– средний процент принятых заявок или покидающих СМО необслуженными;

– среднее время ожидания в очереди;

– закон распределения времени ожидания;

– среднее количество заявок, находящихся в очереди;

– закон распределения числа заявок в очереди;

– средний доход, приносимый СМО в единицу времени, и т. д.

Случайный характер потока заявок приводит к тому, что в системе массового обслуживания будет происходить случайный процесс. Для рациональной его организации необходимо изучить случайный процесс, протекающий в системе, и описать его математически. Этим и занимается теория массового обслуживания.

Заметим, что за последние годы область применения математических методов теории массового обслуживания непрерывно расширяется. Как своеобразные системы массового обслуживания могут рассматриваться: системы сбора и обработки информации; автоматизированные производственные цехи, поточные линии; транспортные системы; системы противовоздушной обороны и т. п. Близкими к задачам теории массового обслуживания являются многие задачи, возникающие при анализе надежности технических устройств.

Математический анализ работы СМО значительно облегчается, если случайный процесс, протекающий в системе, является марковским. Тогда удается сравнительно просто описать работу СМО с помощью аппарата обыкновенных дифференциальных (в предельном случае – линейных алгебраических) уравнений и выразить в явном виде основные характеристики эффективности обслуживания через параметры СМО и потока заявок.

Для того, чтобы процесс, протекающий в системе, был марковским, нужно, чтобы все потоки событий, переводящие систему из состояния в состояние, были *пуассоновскими* (потоками без последействия). Если эти потоки не являются пуассоновскими, то математическое описание процессов, происходящих в СМО, становится сложным и требует более громоздкого аппарата. В этих случаях довести описание до явных, аналитических формул удается только в редких, простейших случаях. Однако, если процесс, протекающий в СМО, отличен от марковского, то с его помощью характеристики эффективности СМО могут быть оценены приближенно. Чем сложнее СМО, чем больше в ней каналов обслуживания, тем точнее оказываются приближенные формулы, полученные с помощью марковской теории. Следует также заметить, что в ряде случаев для принятия обоснованных решений по управлению работой СМО вовсе и не требуется точного знания всех ее характеристик – зачастую достаточно и приближенного.

Системы массового обслуживания вообще могут быть двух типов:

1. *Системы с отказами*. В таких системах заявка, поступившая в момент, когда все каналы заняты, получает «отказ», покидает СМО и в дальнейшем процессе обслуживания не участвует.

2. *Системы с ожиданием (с очередью).* В таких системах заявка, поступившая в момент, когда все каналы заняты, становится в очередь и ожидает, пока не освободится один из каналов. Как только освободится канал, принимается к обслуживанию одна из заявок, стоящих в очереди.

Обслуживание в системе с ожиданием может быть «упорядоченным» (заявки обслуживаются в порядке поступления) и «неупорядоченным» (заявки обслуживаются в случайном порядке).

Системы с очередью делятся на *системы с неограниченным ожиданием* и *системы с ограниченным ожиданием.*

В системах с неограниченным ожиданием каждая заявка, поступившая в момент, когда нет свободных каналов, становится в очередь и ждет освобождения канала, который примет ее к обслуживанию. Любая заявка, поступившая в СМО, будет обслужена.

В системах с ограниченным ожиданием на пребывание заявки в очереди накладываются те или другие ограничения. Эти ограничения могут касаться длины очереди (числа заявок, одновременно находящихся в очереди), времени пребывания заявки в очереди (после какого-то срока пребывания в очереди заявка покидает очередь и уходит), общего времени пребывания заявки в СМО и т. д.

В зависимости от типа СМО, при оценке ее эффективности применяются так называемые показатели эффективности. Например, для СМО с отказами одной из важнейших характеристик ее продуктивности является так называемая *абсолютная пропускная способность* – среднее число заявок, которое может обслужить система за единицу времени. Наряду с абсолютной, часто рассматривается *относительная пропускная способность* СМО – средняя доля поступивших заявок, обслуживаемая системой (отношение среднего числа заявок, обслуживаемых системой в единицу времени, к среднему числу поступающих за это время заявок).

Помимо абсолютной и относительной пропускной способностей, при анализе СМО с отказами в зависимости от задачи исследования интересуют и другие характеристики: – среднее число занятых каналов; – среднее относительное время простоя системы в целом и отдельного канала и т. д.

Очевидно, для *СМО с неограниченным ожиданием* как абсолютная, так и относительная пропускная способность теряют смысл, так как каждая поступившая заявка рано или поздно будет обслужена. Для такой СМО весьма важными характеристиками являются: – среднее число заявок в очереди; – среднее число заявок в системе (в очереди и под обслуживанием); – среднее время ожидания заявки в очереди; – среднее время пребывания заявки в системе (в очереди и под обслуживанием) и другие характеристики ожидания.

Для СМО с ограниченным ожиданием интерес представляют обе группы характеристик: как абсолютная и относительная пропускная способности, так и характеристики ожидания.

Для анализа процесса, протекающего в СМО, существенно знать основные параметры системы: число каналов, интенсивность потока заявок, производительность каждого канала (среднее число заявок обслуживаемое каналом в единицу времени), условия образования очереди (ограничения, если они есть).

В зависимости от этих параметров выражают характеристики эффективности работы СМО. Будем считать все потоки событий, переводящие СМО из состояния в состояние, пуассоновскими. Напомним, что в случае, когда пуассоновский поток стационарен (простейший поток), интервал времени Т между событиями в этом потоке есть случайная величина, распределенная по показательному закону:

$$f(t) = \lambda e^{-\lambda t}, \tag{3.1}$$

где λ – интенсивность потока событий.

В случае, когда из какого-то состояния систему выводят сразу несколько простейших потоков, величина T – время пребывания системы (подряд) в данном состоянии есть случайная величина, распределенная по закону (3.1), где λ – суммарная интенсивность всех потоков событий, выводящих систему из данного состояния.

При использовании *логистического подхода* объектом управления также становится *поток* – множество, совокупность объектов, воспринимаемая как единое целое.

Поток является не только главной категорией логистики, но и в сочетании «*управление потоком»* становится основным логистическим инструментом. Однако, для описания потоков даётся лишь самая простая их классификация: внешний – внутренний поток; непрерывный – дискретный; детерминированный – стохастический; стабильный – нестабильный; стационарный – нестационарный; равномерный – неравномерный; периодический – непериодический; ритмичный – неритмичный; простой – сложный; управляемый – неуправляемый.

Существующая классификация, по сути, неприменима для составления моделей, характеризующих потоки, и, следовательно, не касается дальнейшей работы с ними.

Научное управление потоками требует упорядочения и формализации всего многообразия потоков.

Основными параметрами, характеризующими поток, выступают: начальный пункт потока (a), конечный пункт потока (b), траектория пути (F), длина пути (D), скорость перемещения объектов потока (V), время (t), промежуточные пункты (q), интенсивность потока (λ).

Пусть ρ – поток, S – логистическая система. Тогда состояние потока может быть охарактеризовано с помощью каких-либо численных переменных (его параметров) как функции от времени (3.2):

$$\rho\ (\mathrm{a, b, F, D, q, V}, \lambda\ ) = \mathrm{f}\ (\mathrm{t}). \tag{3.2}$$

Основным параметром, характеризующим поток, является его *плотность*, представляющая собой количество перемещаемых объектов в единицу времени.

Существующую *условную классификацию потоков* можно представить в виде *двух укрупнённых видов потоков: детерминированных и стохастических.* Основные модели, представляющие данную классификацию, приведены в таблице (см. ниже).

Детерминированные потоки – это потоки, значения параметров которых являются определёнными на любой момент времени (рис. 38). Интервалы между событиями являются строго одинаковыми и равными определённой неслучайной величине. Поэтому детерминированные потоки также называют *регулярными.*

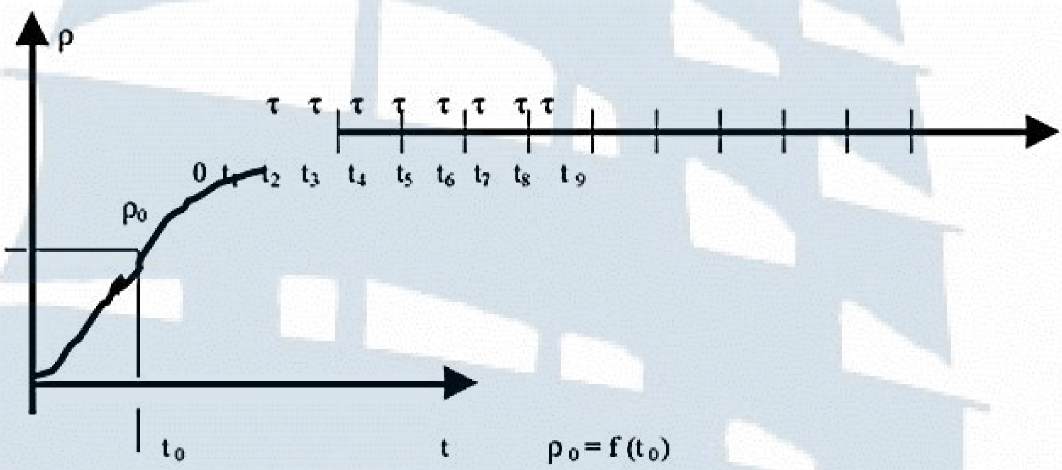


Рис. 38. Детерминированный поток

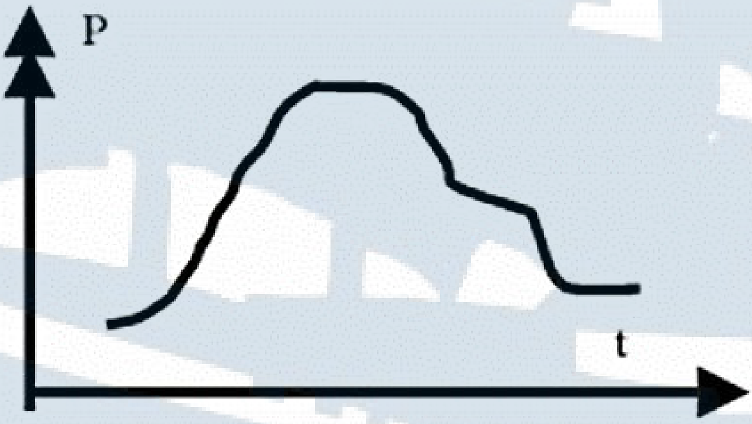
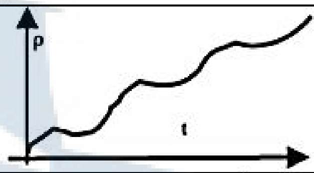
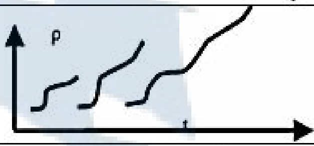
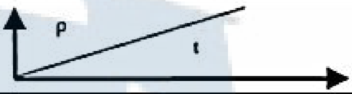
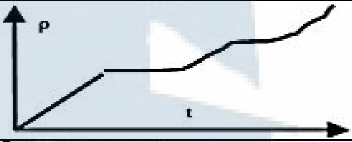
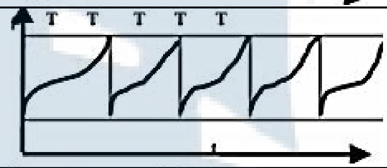
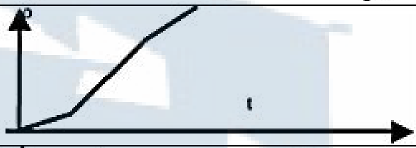


Рис. 39. Стохастический поток

Стохастические потоки – потоки, значения параметров которых являются случайными величинами (рис. 39). В том или ином состоянии система находится с некоторой вероятностью P.

Таблица

Вид потока	Модель	Схема модели
Детерминированные потоки		
1. Стабильные потоки	$P = f(t) = const;$ $t = 0, \infty$	
Нестабильные потоки	$P = f(t) \neq const$	
2. Равномерные потоки	$V = const;$ $S = V*t$	
Неравномерные потоки	$V = const;$ $S \neq V*t$	
3. Периодические потоки	$P = f(T)$	
Непериодические потоки	$P = f(t_i) = T$	
4. Ритмичные потоки	$P = V1*t1 + V2*t2 + \ldots + Vn*tn$	
Неритмичные потоки	$P = f(V,t)$	
Стохастические потоки		
1. Стационарные потоки	$\lambda = const$	
Нестационарные потоки	$\lambda \neq const$	
2. Непрерывные потоки	$P = \int_{t1} f(t)dt$	
Дискретные потоки	$P = \sum_{i=1} pi$	

Детерминированный поток довольно редко встречается на практике; он представляет определённый интерес как предельный случай для других потоков. Вообще, в природе не существует совершенно не случайных процессов, но есть процессы, на ход которых случайные факторы влияют так слабо, что при описании состояния системы ими можно пренебречь. Однако существуют и такие процессы, где случайность играет основную роль. Между двумя крайними случаями лежит целый спектр процессов, в которых случайность играет большую или меньшую роль. Учитывать или не учитывать случайность потока зависит, в основном, от того, какая практическая задача решается.

Многие прикладные задачи можно решить с помощью сравнительно простых методов расчёта, если отказаться от рассмотрения случайных процессов самого общего вида и ограничиться только процессами, обладающими некоторыми специальными свойствами. Такими, в частности, являются уже рассмотренные выше марковские процессы. В отличие от случайных процессов общего вида, исчерпывающей характеристикой которых выступает многомерный закон распределения, для полной характеристики марковских процессов достаточно знать двумерные законы распределения.

Рассмотрим ряд процессов, происходящих на предприятиях автомобильного транспорта. Например, поток заявок на перевозку грузов, обслуживание населения автомобилями – такси, текущий и капитальный ремонт автомобилей, ТО-2, капитальный ремонт агрегатов, организация технической помощи на линии. Во всех этих потоках имеют место случайные элементы. Так возникновение требований на текущий ремонт автомобиля всегда является случайным. В один момент времени их может быть больше, а в другой – меньше, хотя в среднем за определённое время их число, возможно, будет постоянным. Случайным будет и время, затрачиваемое на текущий ремонт каждого автомобиля, так как оно зависит от того, какая неисправность явилась причиной появления требования на текущий ремонт, от квалификации рабочего и от других факторов. Это свидетельствует о том, что большинство входящих потоков на автомобильном транспорте зависят от ряда случайных факторов, что также относится и ко времени обслуживания. Поэтому эти величины обычно описываются с помощью вероятностных характеристик.

Целью изучения стохастических потоков является обеспечение эффективной работы, которая в каждом случае имеет свой конкретный смысл. Она должна определяться не качественно, а количественно, то есть определённым числом, что требует математического представления каждого процесса массового обслуживания. Стохастические потоки широко применяются в теории восстановления – разделе теории надёжности технических устройств.

В теории массового обслуживания, в основном, рассматривают простейший поток требований, то есть обладающий свойствами стационарности, ординарности и отсутствия последействия. С помощью методов теории массового обслуживания устанавливается зависимость между характером потоков и параметрами обслуживания, определяются распределения интервалов поступления заявок и времени обслуживания, вероятности состояния системы, очереди заявок, вычисляется пропускная способность системы. Применение математического аппарата теории массового обслуживания может дать много ценного для методического обеспечения логистической деятельности на предприятиях.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Вентцель А.Д. Курс теории случайных процессов. – М.: Наука, 1996. – 320 с.

2. Вентцель Е.С., Овчаров Л.А. Теория случайных процессов и ее инженерные приложения. – М.: Высшая школа, 2001. – 382 с.

3. Вентцель Е. С., Овчаров Л. А. Задачи и упражнения по теории вероятностей: учебное пособие для втузов. – 5-е изд., испр. – М.: Издательский центр «Академия», 2003.

4. Корнева, И.П. Специальные главы математики. Теория вероятностей и математическая статистика: учебное пособие. – Калининград: Изд-во БГАРФ, 2019. –143 с.

5. Миллер Б.М., Панков А.Р. Теория случайных процессов в примерах и задачах. – М.: ФИЗМАТЛИТ, 2002. – 320 с.

6. Письменный, Д.Т. Конспект лекций по теории вероятностей, математической статистике и случайным процессам. – М.: Айрис, 2008.

7. Прохоров С.А. Моделирование и анализ случайных процессов: лабораторный практикум. – 2-е изд., перераб. и дополненное. – СНЦ РАН, 2002. – 277 с. [Электронный ресурс].

8. Свешников А.А. Прикладные методы теории Марковских процессов: Учебное пособие. – Санкт-Петербург: Лань, 2007.

9. Свешников А.А. Сборник задач по теории вероятностей, математической статистике и теории случайных функций: учебное пособие, 4-е изд. – Санкт-Петербург: Лань, 2008.

10. Хрущева И.В., Щербаков В.И., Леванова Д.С. Основы математической статистики и теории случайных процессов: учебное пособие. – 2-е изд., испр. – Санкт-Петербург: Лань, 2009.

Корнева Ирина Петровна

СЛУЧАЙНЫЕ ПРОЦЕССЫ И МАТЕМАТИЧЕСКАЯ СТАТИСТИКА

Учебное пособие для студентов,
обучающихся по направлению 26.03.01
«Управление водным транспортом
и гидрографическое обеспечение судоходства»,
профиль «Управление водными
и мультимодальными перевозками»
всех форм обучения

Ведущий редактор М.Б. Априянц

Младший редактор Г.В. Деркач

Компьютерное редактирование
И.В. Леонова

Подписано в печать 25.11.2021 г.
Усл. печ. л. 5,5. Уч.-изд. л. 4,5.

Лицензия № 021350 от 28.06.99.

Печать офсетная.

Формат 60х90/16.

Заказ № 1724. Тираж 50 экз.

Доступ к архиву публикации и условия доступа к нему:
https://bgarf.ru/akademia/#biblioteka

БГАРФ ФГБОУ ВО «КГТУ»

Издательство БГАРФ,
член Издательско-полиграфической ассоциации высших учебных заведений
236029, Калининград, ул. Молодежная, 6.