Лекция № 1. Вводная

Математическое моделирование рабочих процессов на автомобильном транспорте. Курс предусматривает 18 часов лекций, 18 часов лабораторных работ, 36 часов СРС, 36 часов контроль, курсовой проект, экзамен.

Лекционный курс включает:

1. Изучение методов обработки опытных данных (регрессионный и корреляционный анализ).
2. Вероятностные законы и их инженерное приложение.
3. Случайные функции и случайные процессы и их моделирование.
4. Характеристика функционирования систем массового обслуживания.

Лабораторные работы:

1. Основы вычислений в Microsoft Excel.
2. Использование математических и статистических функций в Microsoft Excel.
3. Работа с логическими функциями в Microsoft Excel.
4. Подбор формул по данным опыта методом наименьших квадратов в Microsoft Excel.
5. Законы распределения дискретной случайной величины.
6. Методы обработки опытных данных. Построение линейных и нелинейных детерминированных моделей.
7. Обработка экспериментальных данных в Microsoft Excel.
8. Проверка гипотезы о нормальном распределении генеральной совокупности по критерию Пирсона.
9. Обработка экспериментальных данных по закону Вейбулла.
10. Моделирование случайных величин методом Монте-Карло.
11. Моделирование процессов ТО и ремонта автомобилей.
12. Моделирование функционирования систем массового обслуживания.
13. Расчет показателей сетевых графиков.

Литература:

1. Тарасик, В.П. Математическое моделирование технических систем [Электронный ресурс]: учебник. – Электрон. дан. – Минск: Новое знание, 2013. – 584 с. – Режим доступа: http://e.lanbook.com/books/element.php?pl1_id=4324 – Загл. с экрана.
2. Введение в математическое моделирование транспортных потоков [Электронный ресурс]: учебное пособие. – Электрон. дан. – М.: МЦНМО (Московский центр непрерывного математического образования), 2013. – 426 с. – Режим доступа: http://e.lanbook.com/books/element.php?pl1_id=56419 – Загл. с экрана.
3. Голубева Н.В. Математическое моделирование систем и процессов [Электронный ресурс]: учебное пособие. – Электрон. дан. – СПб.: Лань, 2013. – 192 с. – Режим доступа: http://e.lanbook.com/books/element.php?pl1_id=4862 – Загл. с экрана.

4. Гмурман В.Е. Теория вероятностей и математическая статистика: учеб. пособие. – 12-е изд. – М.: Высшее образование, 2008. – 479 с. ISBN 978-5-9692-0192-7.

5. Гмурман, В.Е. Руководство к решению задач по теории вероятностей и математической статистике: учебное пособие для вузов: [для подготовки бакалавров и специалистов] / В.Е. Гмурман. – Москва: Высшее образование, 2009. – 404 с. – ISBN 978-5-9692-0384-6.

6. Баженов, М.Ю. Моделирование производственных процессов: методические указания к лабораторным работам / М.Ю. Баженов; Владимирский государственный университет (ВлГУ), 2013. – 99 с.

7. Дьяконов, В.П. VisSim + Mathcad + MATLAB. Визуальное математическое моделирование [Электронный ресурс]. – Электрон. дан. – М.: СОЛОН-Пресс, 2008. – 384 с. – Режим доступа: http://e.lanbook.com/books/element.php?pl1_id=13679 – Загл. с экрана.

8. Аверченков, В.И. Основы математического моделирования технических систем [Электронный ресурс]: учебное пособие / В.И. Аверченков, В.П. Федоров, М.Л. Хейфец. – Электрон. дан. – М.: ФЛИНТА, 2011. – 271 с. – Режим доступа: http://e.lanbook.com/books/element.php?pl1_id=44652 – Загл. с экрана.

ВВЕДЕНИЕ

ЭВМ прочно вошла в нашу жизнь, и практически нет такой области человеческой деятельности, где не применялась бы ЭВМ. ЭВМ сейчас широко используется в процессе создания и исследования новых машин, новых технологических процессов и поиске их оптимальных вариантов; при решении экономических задач, при решении задач планирования и управления производством на различных уровнях. Создание же крупных объектов в ракетотехнике, авиастроении, судостроении, а также проектирование плотин, мостов, и др. вообще невозможно без применения ЭВМ.

Для использования ЭВМ при решении прикладных задач, прежде всего прикладная задача должна быть "переведена" на формальный математический язык, т.е. для реального объекта, процесса или системы должна быть построена его математическая модель.

Слово "Модель" происходит от латинского modus (копия, образ, очертание). Моделирование – это замещение некоторого объекта А другим объектом Б. Замещаемый объект А называется оригиналом или объектом моделирования, а замещающий Б – моделью. Другими словами, модель – это объект-заменитель объекта-оригинала, обеспечивающий изучение некоторых свойств оригинала.

Целью моделирования являются получение, обработка, представление и использование информации об объектах, которые взаимодействуют между собой и внешней средой; а модель здесь выступает как средство познания свойств и закономерности поведения объекта.

Моделирование широко используются в различных сферах человеческой деятельности, особенно в сферах проектирования и управления, где особенными являются процессы принятия эффективных решений на основе получаемой информации.

Теорией моделирования является раздел науки, изучающий способы исследования свойств объектов-оригиналов, на основе замещения их другими объектами-моделями. В основе теории моделирования лежит теория подобия. При моделировании абсолютное подобие не имеет места и лишь стремится к тому, чтобы модель достаточно хорошо отображала исследуемую сторону функционирования объекта. Абсолютное подобие может иметь место лишь при замене одного объекта другим точно таким же.

Классификация моделей

Все модели можно разделить на два класса:

- вещественные,
- идеальные.

В свою очередь вещественные модели можно разделить на:

- натурные,
- физические,
- математические.

Идеальные модели можно разделить на:

- наглядные,
- знаковые,
- математические.

Вещественные натурные модели – это реальные объекты, процессы и системы, над которыми выполняются эксперименты научные, технические и производственные.

Вещественные физические модели – это макеты, муляжи, воспроизводящие физические свойства оригиналов (кинематические, динамические, гидравлические, тепловые, электрические, световые модели).

Вещественные математические – это аналоговые, структурные, геометрические, графические, цифровые и кибернетические модели.

Идеальные наглядные модели – это схемы, карты, чертежи, графики, графы, аналоги, структурные и геометрические модели.

Идеальные знаковые модели – это символы, алфавит, языки программирования, упорядоченная запись, топологическая запись, сетевое представление.

Идеальные математические модели – это аналитические, функциональные, имитационные, комбинированные модели.

В приведенной классификации некоторые модели имеют двойное толкование (например – аналоговые). Все модели, кроме натурных, можно объединить в один класс мысленных моделей, т.к. они являются продуктом абстрактного мышления человека.

Математическое моделирование

Остановимся на одном из наиболее универсальных видов моделирования – математическом, ставящим в соответствие моделируемому физическому процессу систему математических соотношений, решение которой позволяет получить ответ на вопрос о поведении объекта без создания физической модели, часто оказывающейся дорогостоящей и неэффективной.

Математическое моделирование – это средство изучения реального объекта, процесса или системы путем их замены математической моделью, более удобной для экспериментального исследования с помощью ЭВМ.

Математическая модель является приближенным представлением реальных объектов, процессов или систем, выраженным в математических терминах и сохраняющим существенные черты оригинала. Математические модели в количественной форме, с помощью логико-математических конструкций, описывают основные свойства объекта, процесса или системы, его параметры, внутренние и внешние связи.

Построение математической модели заключается в определении связей между теми или иными процессами и явлениями, создании математического аппарата, позволяющего выразить количественно и качественно связь между теми или иными процессами и явлениями, между интересующими специалиста физическими величинами, и факторами, влияющими на конечный результат.

Обычно их оказывается настолько много, что ввести в модель всю их совокупность не удается. При построении математической модели перед исследованием возникает задача выявить и исключить из рассмотрения факторы, несущественно влияющие на конечный результат (математическая модель обычно включает значительно меньшее число факторов, чем в реальной действительности). На основе данных эксперимента выдвигаются гипотезы о связи между величинами, выражающими конечный результат, и факторами, введенными в математическую модель. Такая связь зачастую выражается системами дифференциальных уравнений в частных производных (например, в задачах механики твердого тела, жидкости и газа, теории фильтрации, теплопроводности, теории электростатического и электродинамического полей).

Конечной целью этого этапа является формулирование математической задачи, решение которой с необходимой точностью выражает результаты, интересующие специалиста.

Для автомобильного транспорта (АТ) расчетные методы моделирования позволяют определить необходимое число постов обслуживания, запасных частей и других ресурсов для осуществления транспортного процесса.

Основным производственным процессом АТ является транспортный. На осуществление этого процесса работает ряд служб АТП – техническая, коммерческая (эксплуатации), служба снабжения, служба главного механика и т.д.

Каждая служба ведет свое <u>производство</u>, которое определяет свои конкретные <u>цели</u>, для достижения которых требуется решение определенных <u>задач</u>, каждая из которых требует свои <u>методы решения</u>.

Всю эту последовательность действий можно представить в виде цепочки

Пр – Ц – З – Р.

При изучении производственного процесса, последний заменяют моделью с целью экономии затрат (денежных, временных, трудовых).

Математическими моделями называют комплекс математических зависимостей и логических выражений, отображающих существенные характеристики изучаемого явления (процесса).

Преимущества математических моделей (математического моделирования) перед другими заключаются:

а) в низкой стоимости их создания;

б) в быстром получении результатов исследования;

в) в возможности проведения расчетных экспериментов и проверки правильности построения модели.

К ***недостаткам*** математических моделей следует отнести то, что они абстрактны.

Примерами математических моделей могут служить следующие зависимости:

а) $y = \int_b^a f(x)dx$ – определенный интеграл для вычисления площади, работы, скорости и т.д.;

б) $y = a + bx$ – линейная зависимость;

в) математические модели вероятностных законов, например, $f(x) = \lambda e^{-\lambda x}$ – показательный закон;

г) $y = \sum_{i=1}^{n} c_i x_i$, где c_i - параметр; x_i - переменная; y - искомая величина.

Классификация математических моделей

По ***принципам построения*** математические модели разделяют на аналитические и имитационные.

В аналитических моделях процессы функционирования реальных объектов, процессов или систем записываются в виде явных функциональных зависимостей. Аналитическая модель разделяется на типы в зависимости от математической проблемы:

– уравнения (алгебраические, трансцендентные, дифференциальные, интегральные),

– аппроксимационные задачи (интерполяция, экстраполяция, численное интегрирование и дифференцирование),

– задачи оптимизации,

– стохастические проблемы.

Рис. 1. Классификация математических моделей

Однако по мере усложнения объекта моделирования построение аналитической модели превращается в трудноразрешимую проблему. Тогда исследователь вынужден использовать имитационное моделирование.

В имитационном моделировании функционирование объектов, процессов или систем описывается набором алгоритмов. Алгоритмы имитируют реальные элементарные явления, составляющие процесс или систему с сохранением их логической структуры и последовательности протекания во времени. Имитационное моделирование позволяет по исходным данным получить сведения о состояниях процесса или системы в определенные моменты времени, однако прогнозирование поведения объектов, процессов или систем здесь затруднительно. Можно сказать, что имитационные модели - это проводимые на ЭВМ вычислительные эксперименты с математическими моделями, имитирующими поведение реальных объектов, процессов или систем.

В зависимости от ***характера исследуемых реальных процессов*** и систем математические модели могут быть детерминированные и стохастические.

В детерминированных моделях предполагается отсутствие всяких случайных воздействий, элементы модели (переменные, математические связи) достаточно точно установленные, поведение системы можно точно определить. При построении детерминированных моделей чаще всего используются алгебраические уравнения, интегральные уравнения, матричная алгебра.

Стохастическая модель учитывает случайный характер процессов в исследуемых объектах и системах, который описывается методами теории вероятности и математической статистики.

По ***виду входной информации*** модели разделяются на непрерывные и дискретные.

Если информация и параметры являются непрерывными, а математические связи устойчивы, то модель - непрерывная. И наоборот, если информация и параметры - дискретны, а связи неустойчивы, то и математическая модель - дискретная.

По ***поведению моделей во времени*** они разделяются на статические и динамические.

Статические модели описывают поведение объекта, процесса или системы в какой-либо момент времени. Динамические модели отражают поведение объекта, процесса или системы во времени.

По степени соответствия между математической моделью и реальным объектом, процессом или системой математические модели разделяют на изоморфные (одинаковые по форме) и гомоморфные (разные по форме).

Модель называется изоморфной, если между нею и реальным объектом, процессом или системой существует полное поэлементное соответствие. Гомоморфной – если существует соответствие лишь между наиболее значительными составными частями объекта и модели.

Последовательность решения инженерной задачи на ПК

Решение задачи на ПК представляет собой трудоемкий процесс, который требует знания из различных областей науки:

– математики;

– вычислительной техники;

– программирования;

– специальных дисциплин (если решаются задачи по выбранной специальности).

Если в качестве обработки данных выступает ПК, то решение этих задач можно разбить на следующие основные этапы:

Первый этап «Постановка задачи» заключается в изучении исследуемого явления, формулировке задачи и целей ее решения.

Определяются исходные данные и ожидаемые результаты (их содержание, объем, достоверность описываемого явления).

Этот этап заканчивается словесным описанием задачи.

Первый этап непосредственно связан со *вторым* «Математическая формулировка задачи». В итоге второго этапа должна быть получена математическая модель.

На третьем этапе «Разработка метода решения» производится выбор целесообразного (оптимального) математического метода решения поставленной задачи.

Результат – математическое описание выбранного метода решения.

Четвертый этап – «Составление алгоритма решения задачи». Здесь производится детальный анализ выбранного метода решения. Составляется с необходимой степенью детализации с помощью блок-схемы.

Пятый этап «Составление и отладка программы» состоит в записи разработанного алгоритма на одном из языков программирования.

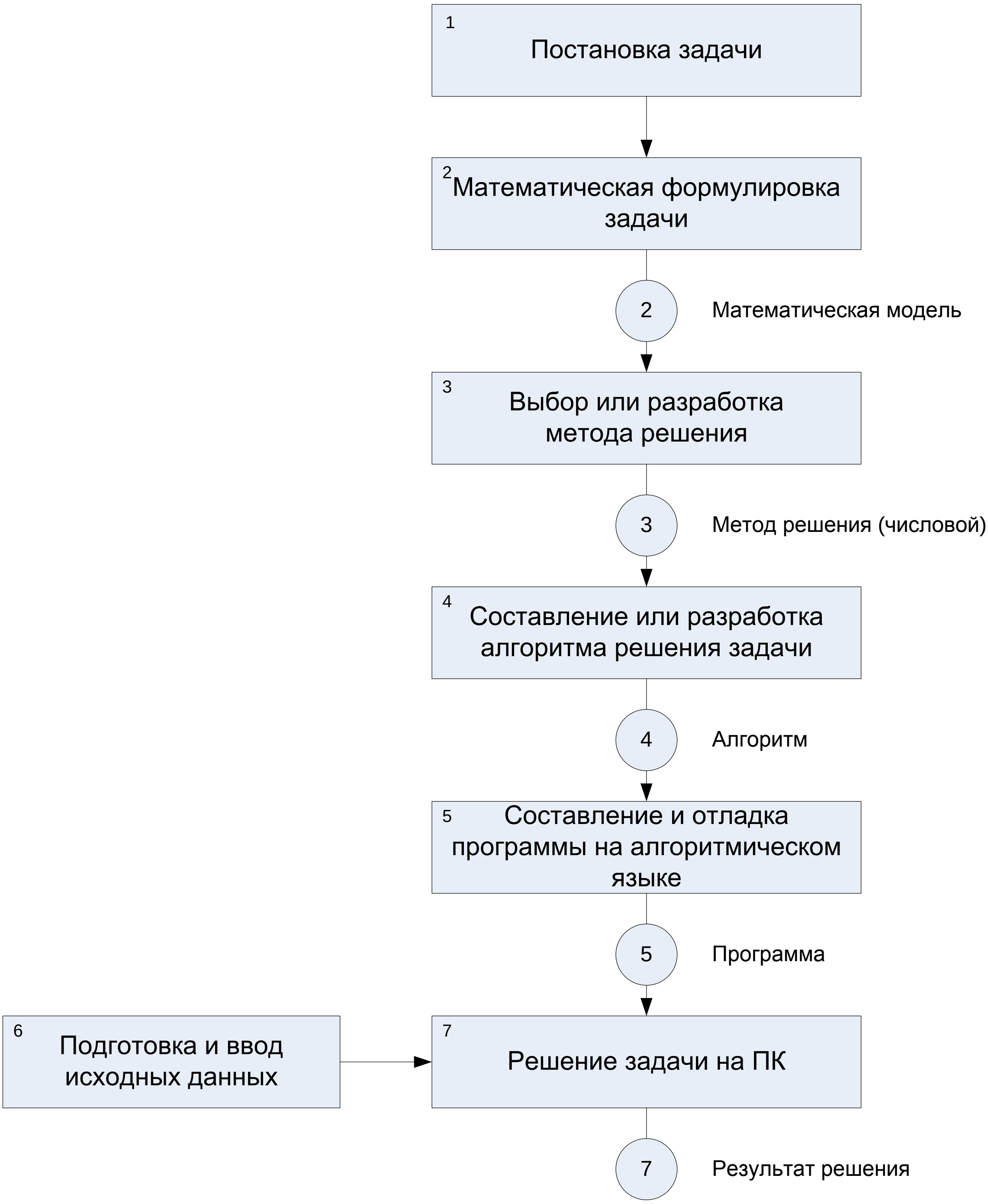


Рис. 2. Последовательность подготовки и решение задач на ПК

Этапы 5, 6, 7 соответствуют этапу отладки программы на ПК. Программа вводится в ПК и обрабатывается специальной программой – транслятором, которая переводит программу с языка программирования в машинный код, состоящий из двоичных данных и инструкций процессора. Транслятор диагностирует все ошибки, связанные с записью программы на языке программирования. Ошибки исправляются и вновь отдаются транслятору на обработку.

Когда будут исправлены все ошибки, транслятор строит машинную программу. На данном этапе производится счет по специально подготовленным контрольным примерам, поиск и исправление ошибок.

После получения правильных результатов по всем контрольным примерам программа считается отлаженной и готовой к эксплуатации.

Этап «Решение задачи на ПК» заключается в получении результатов для различных вариантов исходных данных. Этот этап является регулярным счетом по программе. Для каждой готовой программы должна быть написана соответствующая инструкция по эксплуатации.

Лекция № 2. Алгоритмы решения инженерных задач

После построения математической модели и установления её решения составляется алгоритм решения задачи.

Алгоритм – конечная последовательность строго определённых правил приводящих на основе исходных данных к решению задачи. В практике встречаются *словесные* и *графические* описания алгоритмов.

Словесное описание – описание алгоритма на естественном языке.

Пример: Вычислить значение функции

$$y = arctg(\sqrt{a \cdot x^3 + b} + \ln(a \cdot x^3 + b))$$

при x = ?; a = ?; b = ?

Процесс решения заданный на естественном языке запишется так:

1. Ввести значения x, a, b.
2. Вычислить $t = a \cdot x^3 + b$; $z = \sqrt{t}$.
3. Вычислить $s = \ln(t)$.
4. Вычислить $y = \mathrm{arctg}(z + s)$.
5. Печатать y.

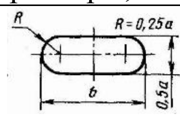
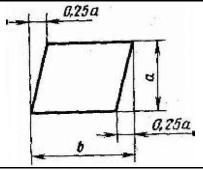
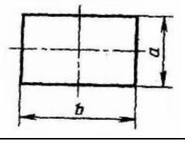
Словесным описанием алгоритма пользуются редко, в основном алгоритм решения инженерных задач описывают *графическим* методом, с помощью блок-схемы.

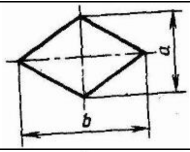
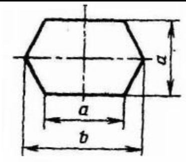
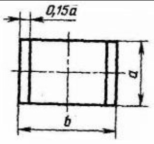
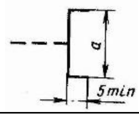
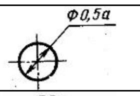
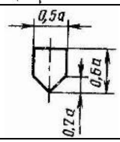
Блок-схема – наглядное графическое изображение алгоритма, когда отдельные действия (этапы) алгоритма изображаются при помощи различных геометрических фигур (блоков).

Обозначение блоков и их содержимого определены ГОСТами: ГОСТ 19002-80, ГОСТ 19003-80, ГОСТ 19005-85, ГОСТ 19701-90.

Наиболее часто встречающиеся операции при разработке алгоритма представлены таблице.

Таблица.

Наименование	Обозначение и размеры, мм	Функция
Пуск - останов		Начало, конец, прерывание процесса обработки данных или выполнения программы Примечание: Размер *a* должен выбираться из ряда *10, 15, 20* мм. Допускается увеличивать размер *a* на число, кратное 5. *R=0,25a; b=1,5a*
Ввод - вывод		Преобразование данных в форму, пригодную для обработки (ввод) или отображения результатов обработки (вывод)
Процесс вычисления		Выполнение операции или группы операций, в результате которых изменяется значение, форма представления или расположение данных

Принятие решения		Выбор направления выполнения алгоритма или программы в зависимости от некоторых переменных условий
Модификация		Организация циклического процесса (выполнение операций, меняющих команды или группы команд, изменяющих программу)
Предопределенный процесс		Использование ранее созданных и отдельно описанных алгоритмов или программ
Комментарий		Связь между элементом схемы и пояснением
Линия потока		Указание последовательности связей между символами
Соединитель		Указание связи между прерванными линиями потока, связывающими символы
Межстраничный соединитель		Указание связи между разъединенными частями схем алгоритмов и программ, расположенных на разных листах

Блоки соединяются линиями потока информации. Внутри блоков записываются выполняемые действия. Линии потока определяют направление вычислений, причем если линии идут сверху вниз и слева направо, то стрелки не ставятся. Если необходимо отразить другое направление (снизу вверх или справа налево), то на линиях ставятся стрелки. Блоки на схеме нумеруются цифрами, которые ставятся в разрыве верхней линии слева.

Практически любой сложный алгоритм представляет собой комбинацию трех типов структур алгоритмов:

- линейного;
- разветвляющегося;
- циклического.

Линейные алгоритмы

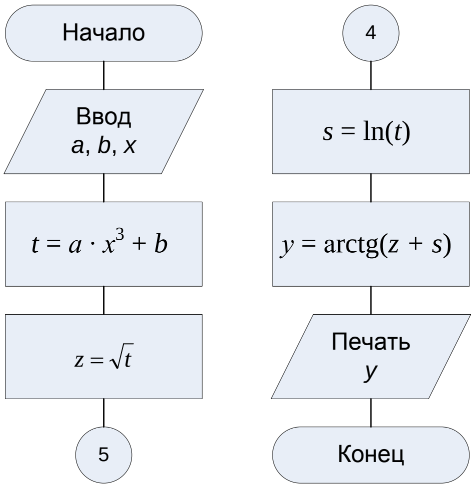
Линейный алгоритм состоит из последовательности операций, выполняющихся только один раз в порядке их следования (см. рис. 1).

Как правило, все «действия», этапы являются чисто арифметическими.

На практике линейные алгоритмы встречаются редко. Как правило, все этапы являются чисто арифметическими.

Пример 1. Линейный алгоритм.

Вычислить: $y = arctg(\sqrt{a \cdot x^3 + b} + \ln(a \cdot x^3 + b))$,

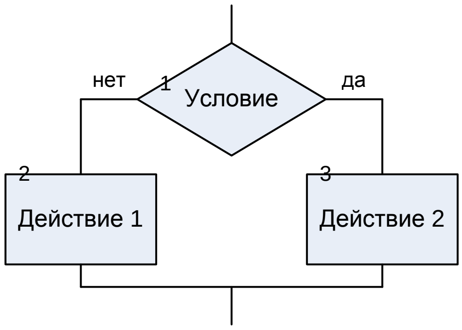


при a = 2, b = 7, x = 1,2,3,4 и т.д.

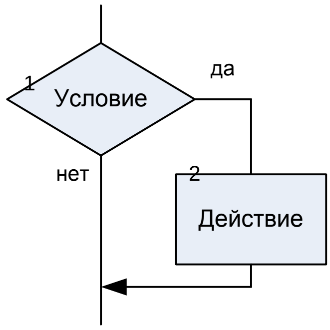
В процессе решения задачи всегда имеем некоторые другие вспомогательные этапы (ввод данных, печать результатов счета и т.д.)

Разветвляющиеся алгоритмы

Разветвляющийся алгоритм – содержит блок проверки некоторого условия, и в зависимости от результата проверки выполняется та или иная последовательность операций, называемая ветвью. При этом форма разветвлений может быть как полной, так и сокращенной.



Полная форма разветвления



Сокращенная форма разветвления

Пример 2. Разветвляющийся алгоритм

Вычислить квадрат наибольшего из трёх чисел: ***a***, ***b***, ***c***.

Вначале сравнивают два числа: ***a*** и ***b***. Большее из них принимают за максимальное. Затем производят сравнение полученного результата с третьим числом ***c***. Если значение ***c*** оказывается больше, то его принимают за максимальное и возводят в квадрат. В противном случае наибольшим считается результат сравнения ***a*** и ***b***.

Алгоритм на естественном языке:

1. Сравнить ***a*** и ***b***. Если ***a*** > ***b***, то принять ***y*** = ***a***. В противном случае принять ***y*** = ***b***.

2. Сравнить ***c*** и ***y***. Если ***c*** > ***y***, то заменить ***y*** = ***c***. В противном случае оставить без изменений.

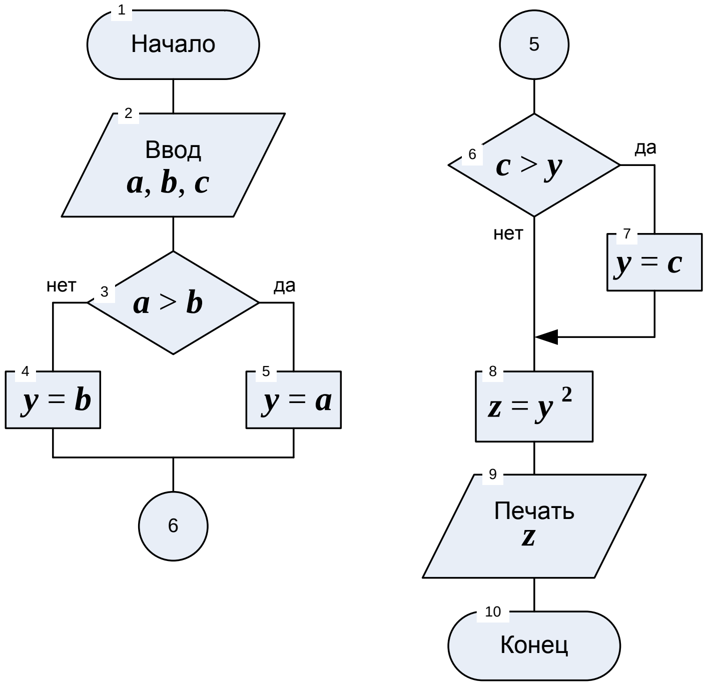
3. Вычислить $\boldsymbol{z} = \boldsymbol{y}^2$.

Алгоритм в виде блок-схемы:

Схема алгоритма содержит два разветвления:

– полное (блоки 3,4,5);

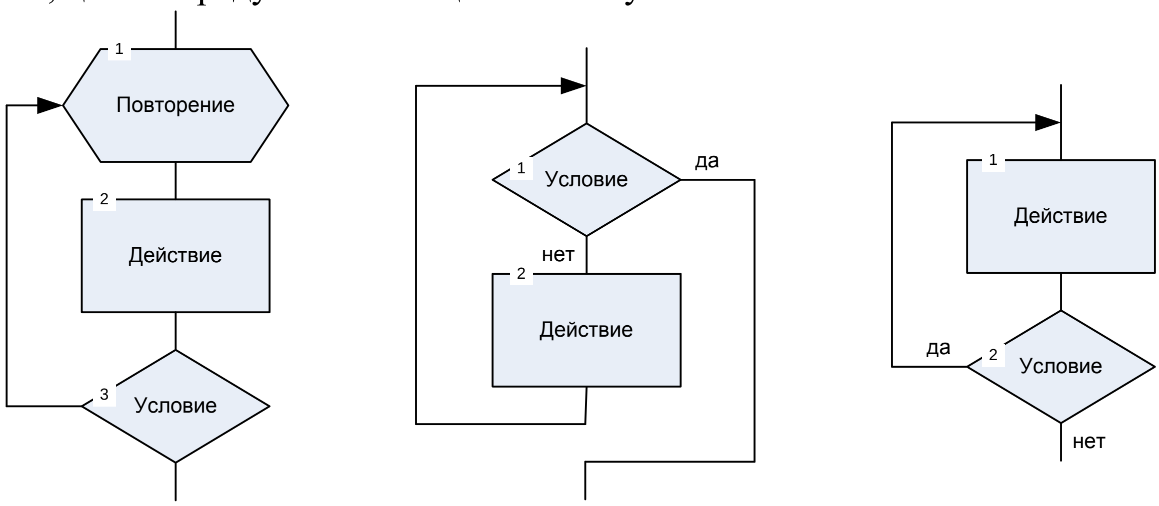
– сокращенное (блоки 6 и 7).



Циклические алгоритмы

Циклический алгоритм – последовательность операций, выполняющихся многократно.

Различают следующие структуры циклических алгоритмов: цикл с повторением, цикл с предусловием и цикл с постусловием.



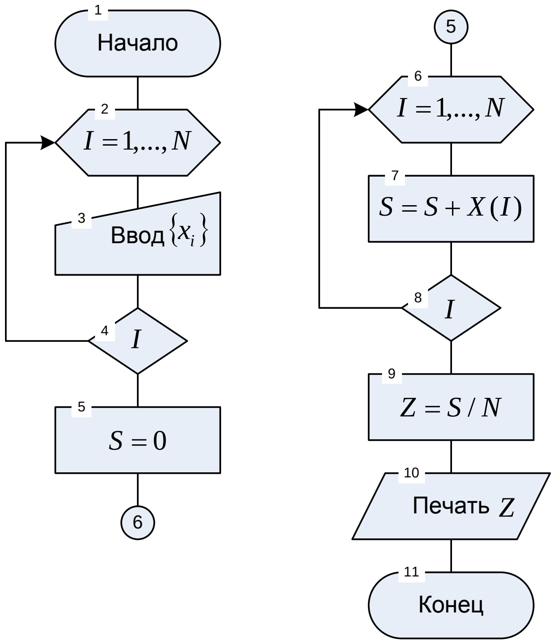
Цикл с повторением *Цикл с предусловием* *Цикл с постусловием*

Любой циклический алгоритм содержит несколько типов блоков. Основной блок, называемый *телом цикла*, производит требуемые вычисления.

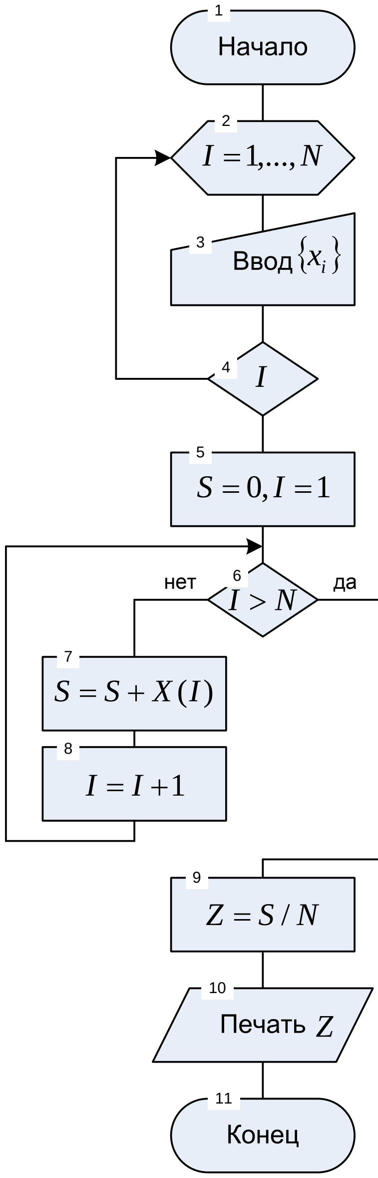
Остальные блоки имеют вспомогательное значение. Они организуют циклический процесс: устанавливают начальные и новые значения данных, проверяют условие окончания или продолжения цикла циклического процесса.

Пример 3. Вычислить среднее арифметическое чисел

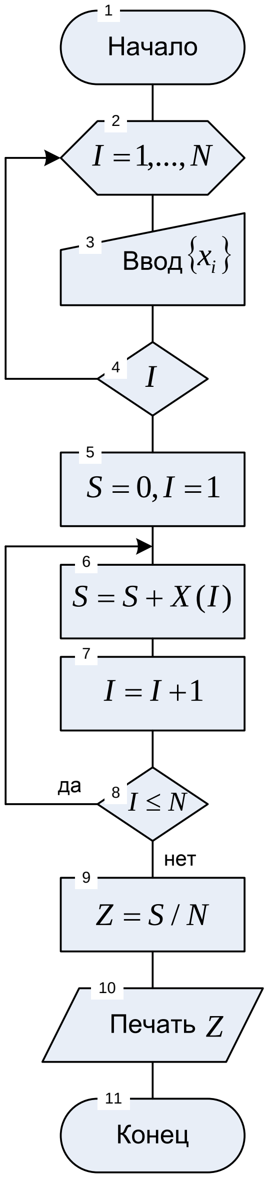
$$\{x_i\} = x_1, x_2, x_3, ..., x_n$$



Циклический алгоритм с повторением



Циклический алгоритм с предусловием



Циклический алгоритм с постусловием

Лекция № 3. Корреляционно-регрессионный анализ

При обработке опытных данных и построении математических моделей используют теорию корреляционно-регрессионного анализа. Рассмотрим содержание каждого из них. Начнем с регрессионного анализа.

Регрессионный анализ устанавливает математическую модель, связывающую зависимую переменную $\boldsymbol{y}$ с исследуемой переменной $\boldsymbol{x}$, т.е. позволяет получить зависимость вида $\boldsymbol{y = f(x)}$, которую называют регрессией.

Различают ***парные*** и ***множественные*** регрессии. Приведем примеры парной регрессии:

$y = b + ax$ – линейная;

$$\left.\begin{aligned} &y = b + \frac{a}{x} \text{ – гиперболическая;} \\ &y = bx^a \text{ – степенная;} \\ &y = ba^x \text{ – показательная;} \\ &y = b + a\ln x \text{ – логарифмическая и т.д.} \end{aligned}\right\} \text{Нелинейные регрессии}$$

Когда значение переменной $\boldsymbol{y}$ зависит сразу от нескольких переменных $\boldsymbol{x_1}$, $\boldsymbol{x_2}$, $\boldsymbol{x_3}$ и т.д., то имеем множественную регрессию, например:

$y = a + bx_1 + cx_2$,

$$y = b_0x_0 + b_1x_1 + ... + b_nx_n = \sum_{i=1}^{n} b_ix_i \ .$$

При практических исследованиях, когда имеем статистические данные, регрессионный анализ позволяет получить числовые значения параметров ***“a”*** и ***“b”***, т.е. получить эмпирические формулы, где зависимую переменную $\boldsymbol{y}$ называют результативным или функциональным признаком, а соответственные независимые переменные $\boldsymbol{x_1}$, $\boldsymbol{x_2}$ называют факторными признаками.

Корреляционный анализ устанавливает количественную оценку тесноты связи между изучаемыми признаками (факторами).

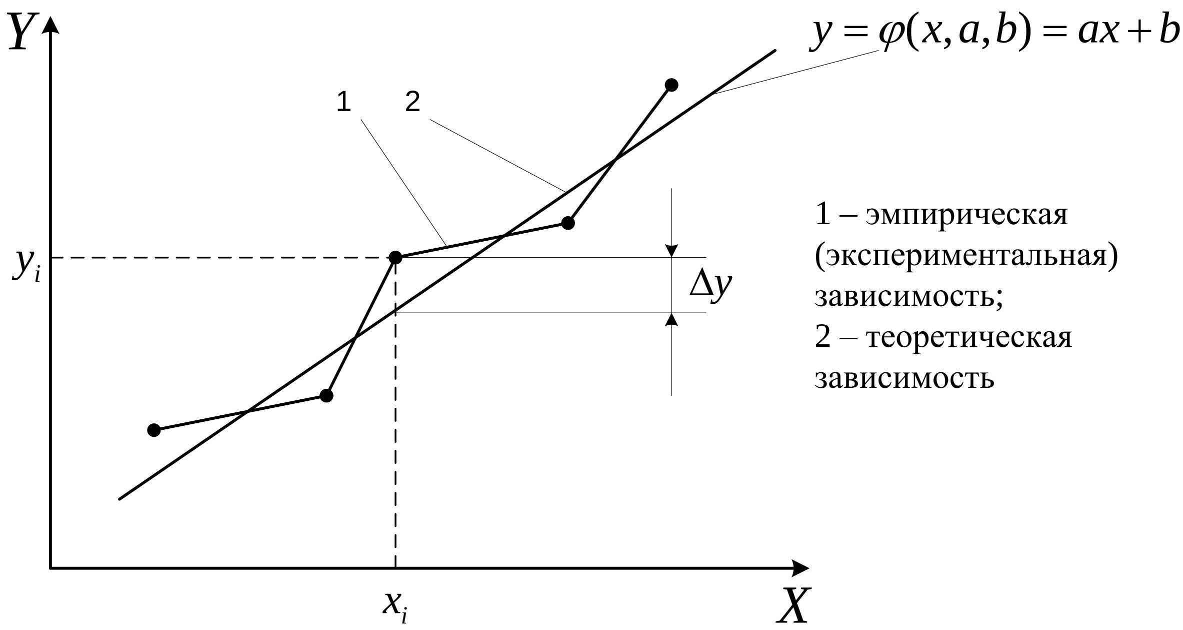
Парная регрессия

Пусть в результате опытов найдены некоторые значения $\boldsymbol{x_i}$ и соответствующие им $\boldsymbol{y_i}$, которые запишем в виде таблицы

x	x_1	x_2	…	x_i	…	x_n
y	y_1	y_2	…	y_i	…	y_n

Требуется найти зависимость вида $y = f(x)$, т.е. парную регрессию, удовлетворяющую этим опытным данным $y = \varphi(x, a, b)$. Одним из наиболее распространенных методов получения таких зависимостей является метод наименьших квадратов.

Метод наименьших квадратов позволяет подобрать более точные значения параметров ***“a”*** и ***“b”.*** Предварительно необходимо установить общий вид статистической функции, который можно выявить по опытным данным, если их нанести на плоскость с координатами ***X0Y***.



Зависимость *Y* от *X*, изображаемая аналитической функцией $y = f(x)$, не может совпадать с экспериментальными значениями y_i во всех ***n*** точках. Это означает, что для всех или некоторых точек имеем разность $\Delta y_i = y_i - \varphi(x_i, a, b)$, отличную от нуля.

Метод наименьших квадратов заключается в том, что подбираются параметры ***a*** и ***b*** таким образом, чтобы сумма квадратов разностей была наименьшей, т.е.

$$z = \sum_{i=1}^{n} \Delta y_i^2 = \sum \left[y_i - \varphi(x_i, a, b)\right]^2 \to \min$$

Если вид функции *y* = *f*(*x*) установлен, то ее можно представить в виде

$$y = f(x) = \varphi(x, a, b),$$

где *a* и *b* – искомые параметры, тогда

$$z = \sum_{i=1}^{n} \left[y_i - \varphi(x_i, a, b)\right]^2 \to \min . \qquad (1)$$

Для нахождения минимума выражения (1) вычисляем частные производные по аргументам *a* и *b* и приравниваем эти производные к нулю, получим:

$$\begin{cases} \dfrac{\partial z}{\partial a} = 2\sum\limits_{i=1}^{n} \left[y_i - \varphi(x_i, a, b)\right]\varphi'_a(x_i, a, b) = 0, \\ \dfrac{\partial z}{\partial b} = 2\sum\limits_{i=1}^{n} \left[y_i - \varphi(x_i, a, b)\right]\varphi'_b(x_i, a, b) = 0. \end{cases} \qquad (2)$$

Система (2) содержит два уравнения с двумя неизвестными *a* и *b*. Решив систему (2), найдем значения параметров *a* и *b*. При найденных значениях параметров величина z будет наименьшей, т.е. аналитическая зависимость будет наилучшим образом описывать экспериментальные данные.

Пример:

Линейная регрессия

Пусть эмпирические данные необходимо описать зависимостью $y = ax + b$, т.е. $y = \varphi(x, a, b) = ax + b$.

Тогда, согласно методу наименьших квадратов, запишем

$$z = \sum_{i=1}^{n} (y_i - ax_i - b)^2 \to \min \tag{3}$$

Выбираем числа a и b так, чтобы величина z была наименьшей, для чего вычисляем частные производные выражения (3) по a и b, получим

$$\begin{cases} \dfrac{\partial z}{\partial a} = 2\sum_{i=1}^{n} (y_i - ax_i - b)(-x_i) = 0, \\ \dfrac{\partial z}{\partial b} = 2\sum_{i=1}^{n} (y_i - ax_i - b)(-1) = 0. \end{cases} \tag{4}$$

Эти два условия дают следующую систему уравнений

$$\begin{cases} \sum_{i=1}^{n} x_i y_i - a\sum_{i=1}^{n} x_i^2 - b\sum_{i=1}^{n} x_i = 0, \\ \sum_{i=1}^{n} y_i - a\sum_{i=1}^{n} x_i - nb = 0. \end{cases} \tag{5}$$

Откуда

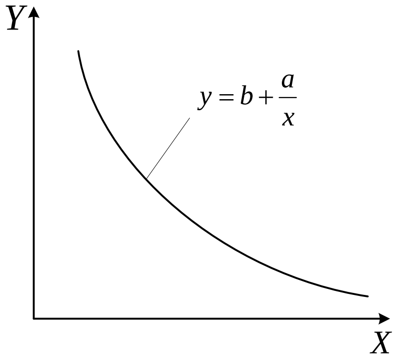
$$\begin{cases} a = \dfrac{\sum x_i y_i - \dfrac{\sum x_i \sum y_i}{n}}{\sum x_i^2 - \dfrac{\left(\sum x_i\right)^2}{n}}; \\ b = \dfrac{\sum y_i}{n} - \dfrac{a\sum x_i}{n}. \end{cases} \tag{6}$$

Для определения численной величины параметров a и b составляем расчетную таблицу или программу для расчета на ПК.

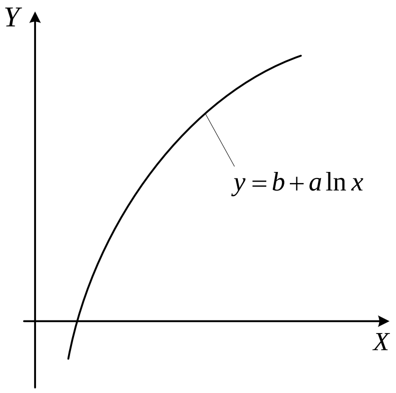
i	x_i	y_i	$x_i y_i$	x_i^2
1	x_1	y_1	$x_1 y_1$	x_1^2
2	x_2	y_2	$x_2 y_2$	x_2^2
…	…	…	…	…
n	x_n	y_n	$x_n y_n$	x_n^2
	$\sum_{i=1}^{n} x_i$	$\sum_{i=1}^{n} y_i$	$\sum_{i=1}^{n} x_i y_i$	$\sum_{i=1}^{n} x_i^2$

Используя метод наименьших квадратов аналогично можно установить значения коэффициентов ***a*** и ***b*** для других парных регрессий.

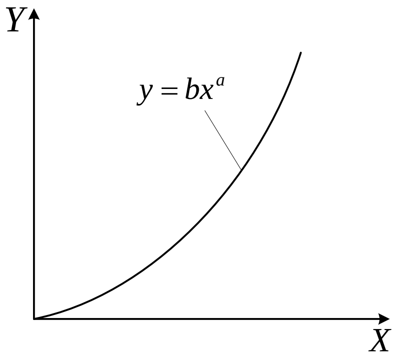
Однако этого можно и не делать, если есть возможность перейти от нелинейной зависимости к линейной (так называемая *линеаризация* аналитических зависимостей):



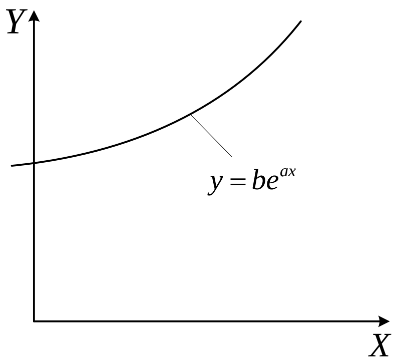
а) пусть $y = \frac{a}{x} + b$ – гиперболическая зависимость; заменим $\frac{1}{x} = x'$, получим линейную зависимость $y = ax' + b$;



б) $y = b + a\ln x$ – логарифмическая зависимость; заменим $\ln x = x'$, получим $y = ax' + b$;



в) $y = bx^a$ – степенная зависимость; логарифмируя, получим $\ln y = \ln b + a\ln x$. Заменим $\ln y = y'$; $\ln b = b'$; $\ln x = x'$. Имеем $y' = ax' + b'$.



г) $y = be^{ax}$ – показательная зависимость; логарифмируя, получим $\ln y = \ln b + ax$. Полагая $\ln y = y'$; $\ln b = b'$, имеем $y' = b' + ax$.

Для линейных зависимостей коэффициенты *a, a´*, *b*, *b´* находим из (6).

Лекция № 4. Корреляционно-регрессионный анализ (продолжение)

Множественная регрессия

Ряд задач автомобильного транспорта требуют построения множественных регрессий вида $Y = f(x_1, x_2, \ldots, x_m)$.

Так, если будем рассматривать деятельность АТП, то одним из основных показателей его работы является коэффициент выпуска автомобилей на линию – $\alpha_в$, на формирование которого оказывает влияние большое число факторов. Основными из них будут:

$\boldsymbol{x_1}$ – обеспеченность водителями;

$\boldsymbol{x_2}$ – обеспеченность ремонтными рабочими;

$\boldsymbol{x_3}$ – обеспеченность инженерными кадрами;

$\boldsymbol{x_4}$ – обеспеченность запасными частями и т.д.

Для рассматриваемого случая

$$\alpha_в = y = f(x_1, x_2, x_3, x_4),$$

т.е. имеем дело с построением множественной регрессии.

В более общем виде линейная множественная регрессия записывается так:

$$y = b_0 x_0 + b_1 x_1 + b_2 x_2 + \ldots + b_m x_m = \sum_{i=0}^{m} b_i x_i \ ,$$

где y – теоретическое значение результативного признака;

x_i – аргументы (факторы);

m – число изучаемых факторов;

b_i – частные коэффициенты регрессии, показывающие степень влияния каждого из факторов на функцию;

b_0 – остаточный член, характеризующий среднее значение функции.

Для получения таких регрессий необходимо располагать статистическими (экспериментальными) данными, которые в общем случае можно представить таблицей

		J = 1, 2, … , n – (столбец)						
		1	2	…	J	…	n	∑
I = 1, 2, …, m – (строка)	X_1	X_{11}	X_{12}	…	X_{1j}	…	X_{1n}	$\sum X_{1j}$
	X_2	X_{21}	X_{22}	…	X_{2j}	…	X_{2n}	$\sum X_{2j}$
	…	…	…	…	…	…	…	…
	X_i	X_{i1}	X_{i2}	…	X_{ij}	…	X_{in}	$\sum X_{ij}$
	…	…	…	…	…	…	…	…
	X_m	X_{m1}	X_{m2}	…	X_{mj}	…	X_{mn}	$\sum X_{mj}$
	y	y_1	y_2	…	y_j	…	y_n	$\sum y_j$

Процесс построения множественной регрессии поясним на примере.

<u>Пример.</u> Для случая $y = f(x_1, x_2)$ линейная регрессия имеет вид $y = a + bx_1 + cx_2$. При вычислении коэффициентов *a, b, c* воспользуемся методом наименьших квадратов. Тогда будем иметь

$$z = \sum_{j=1}^{n} \left(y_j - a - bx_{1j} - cx_{2j}\right)^2 \Rightarrow \min \tag{7}$$

Вычислим частные производные выражения (7) по параметрам *a, b* и *c*, получим:

$$\left\{\begin{aligned} \frac{\partial z}{\partial a} &= 2\sum_{i=1}^{n}(y_i - a - bx_{1i} - cx_{2i})(-1) = 0 \\ \frac{\partial z}{\partial b} &= 2\sum_{i=1}^{n}(y_i - a - bx_{1i} - cx_{2i})(-x_{1i}) = 0 \\ \frac{\partial z}{\partial c} &= 2\sum_{i=1}^{n}(y_i - a - bx_{1i} - cx_{2i})(-x_{2i}) = 0 \end{aligned}\right\} \tag{8}$$

Преобразуя систему (8), получим

$$\left\{\begin{aligned} & na + b\sum_{i=1}^{n} x_{1i} + c\sum_{i=1}^{n} x_{2i} = \sum_{i=1}^{n} y_i \\ & a\sum_{i=1}^{n} x_{1i} + b\sum_{i=1}^{n} x_{1i}^2 + c\sum_{i=1}^{n} x_{1i}x_{2i} = \sum_{i=1}^{n} x_{1i}y_i \\ & a\sum_{i=1}^{n} x_{2i} + b\sum_{i=1}^{n} x_{1i}x_{2i} + c\sum_{i=1}^{n} x_{2i}^2 = \sum_{i=1}^{n} x_{2i}y_i \end{aligned}\right\} \tag{9}$$

Система (9) представляет систему линейных уравнений, решая которую одним из известных методов, найдем значения неизвестных параметров *a, b, c*.

Методом Крамера параметры *a, b* и *c* определяются из выражений:

$$a = \frac{Д_a}{Д}; b = \frac{Д_b}{Д}; c = \frac{Д_c}{Д}$$

где *Д* – главный определитель системы линейных уравнений (9);

$Д_a$ – определитель системы уравнений, в котором столбец коэффициентов при *a* заменен столбцом свободных членов;

$Д_b$ – определитель системы, в которой столбец коэффициентов при *b* заменен столбцом свободных членов;

$Д_c$ – определитель системы, в которых столбец коэффициентов при *c* заменен столбцом свободных членов.

Если требуется построить математическую модель вида $y = \sum_{i=0}^{m} b_i x_i$, то для вычисления коэффициентов b_i (i = 0, 1, 2, …, m) опять воспользуемся методом наименьших квадратов, получаем систему нормальных уравнений вида

$$\begin{cases} nb_0 + b_1\sum x_{1j} + ... + b_m \sum x_{mj} = \sum y_j; \\ b_0 \sum x_{1j} + b_1 \sum x_{1j}x_{2j} + ... + b_m \sum x_{1j}x_{mj} = \sum x_{1j} y_j; \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ b_0 \sum x_{mj} + b_1 \sum x_{1j}x_{mj} + ... + b_m \sum x_{mj}x_{mj} = \sum x_{mj} y_j. \end{cases} \qquad (10)$$

Решая систему (10) известным методом, найдем значения параметров b_i уравнения регрессии. Наиболее распространенным методом решения системы (10) является метод обращения матриц, запрограммированный в ряде стандартных программ аппроксимации.

В матричной форме система (10) запишется

$$XB = Y, \qquad (11)$$

где X – исходная матрица;

B – матрица исходных значений параметров;

Y – матрица свободных членов.

$$X = \begin{pmatrix} n & \sum x_{1j} & .... & \sum x_{mj} \\ \sum x_{1j} & \sum x_{1j}x_{2j} & .... & \sum x_{1j}x_{mj} \\ .... & .... & .... & .... \\ \sum x_{mj} & \sum x_{1j}x_{mj} & .... & \sum x_{mj}x_{mj} \end{pmatrix}; \quad B = \begin{pmatrix} B_0 \\ B_1 \\ .... \\ B_m \end{pmatrix}; \quad Y = \begin{pmatrix} \sum y_j \\ \sum x_{1j} y_j \\ ..... \\ \sum x_{mj} y_j \end{pmatrix}.$$

При решении системы линейных уравнений матричным способом необходимо вычислить обратную матрицу (X^{-1}) и умножить её на правую и левую части уравнения (11), получим

$$X^{-1} \cdot X \cdot B = X^{-1} \cdot Y.$$

Так как $X^{-1} \cdot X = E$ и $E \cdot B = B$, то окончательное выражение для определения матричным способом параметров B_i примет вид

$$B = X^{-1} \cdot Y.$$

Корреляционный анализ

Различают функциональные и корреляционные зависимости. *Функциональной* называют такую зависимость, когда одному значению независимого фактора соответствует только одно значение зависимого фактора. *Корреляционная* зависимость – это такая зависимость, при которой одному значению независимой переменной могут соответствовать несколько значений зависимой переменной. Поэтому корреляционные зависимости могут быть установлены только при обработке большого количества наблюдений. При обработке таких данных пользуются корреляционным анализом.

Корреляционный анализ представляет совокупность методов и приемов выявления количественной оценки, тесноты связи. Тесноту связи (наличие корреляции) между двумя величинами можно определить визуально по полю корреляции.

Корреляционным полем называют нанесенные на график в определенном масштабе точки, соответствующие одновременным значениям двух величин.

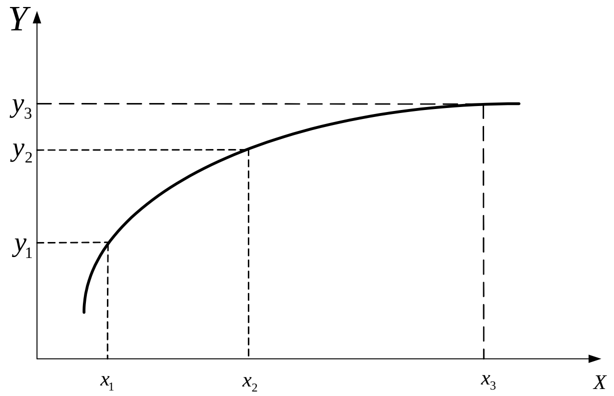


График функциональной зависимости

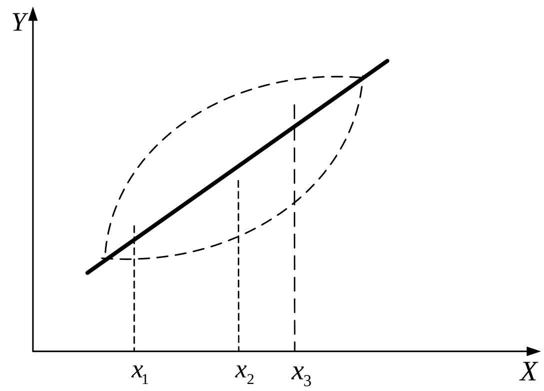


График корреляционной зависимости

В нашем случае теснота связи между параметрами ***x*** и ***y*** определяется визуально по соотношению короткой и продольной осями эллипса рассеяния наблюдений, нанесенных на поле корреляции. Чем больше отношение продольной оси к короткой, тем связь теснее.

Более точно теснота связи оценивается коэффициентом корреляции ***r***.

Коэффициент корреляции лежит в пределах $0 \le |r| \le 1$. При $r = 0$ связи нет. Если $|r| = 1$, то между двумя величинами существует функциональная связь.

Итак, по величине коэффициента корреляции можем сделать следующее заключение:

$0 \le |r| < 0{,}2$ – связи практически нет;
$0{,}2 \le |r| < 0{,}5$ – связь слабая;
$0{,}5 \le |r| < 0{,}75$ – связь средняя;
$0{,}75 \le |r| < 0{,}95$ – связь сильная;
$0{,}95 \le |r| \le 1$ – практически функциональная связь.

При положительных r наблюдается прямая связь, т.е. с увеличением независимой переменной увеличивается и зависимая. При отрицательном коэффициенте корреляции существует обратная связь – с увеличением независимой переменной зависимая переменная уменьшается.

Способы вычисления коэффициента корреляции

Коэффициент парной корреляции

Пусть имеем исходные данные для расчета:

x	x_1	x_2	…	x_i	…	x_n
y	y_1	y_2	…	y_i	…	y_n

Требуется установить зависимость $y = \varphi(x, a, b)$. Можно ли это сделать? Да, если связь существует, нет – если связи нет и результаты дают ложную информацию. Поэтому предварительно рассчитаем коэффициент корреляции. Существует ряд формул для расчета коэффициента корреляции:

1) $$r_{xy} = \frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sqrt{\sum\left(x_i - \bar{x}\right)^2 \sum\left(y_i - \bar{y}\right)^2}},$$

где r_{xy} – коэффициент корреляции;

x_i, y_i – текущие значения наблюдаемых величин;

$\bar{x}$, $\bar{y}$ – средние значения этих величин.

2) $$r_{xy} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sigma_x \sigma_y},$$

где $\overline{xy}$ – среднее значение произведения двух корреляционных величин;

σ_x, σ_y – средние квадратичные отклонения соответствующих величин, которые определяются так:

$$\sigma_x = \sqrt{\bar{x}^2 - (\bar{x})^2}, \sigma_y = \sqrt{\bar{y}^2 - (\bar{y})^2}.$$

Для линейной регрессии коэффициент корреляции r является не только критерием тесноты связи, но и критерием точности аппроксимации (подбора формулы, выражающий зависимость).

Оценка точности аппроксимации криволинейной зависимостью производится при помощи корреляционного отношения

$$\eta = \sqrt{1 - \frac{\sum(y_i - \tilde{y}_i)^2}{\sum(y_i - \bar{y}_i)^2}},$$

где y_i – текущие значения зависимой переменной;

$\tilde{y}_i$ – теоретические значения;

$\bar{y}$ – средние значения.

Корреляционное отношение принимает значения $0 \leq \eta \leq 1$, оно всегда положительно. Если $\eta > r$, то кривая точнее аппроксимирует зависимость, чем прямая; для прямой $r = \eta$.

Дополнительной оценкой точности аппроксимации часто применяют при оценке нелинейной регрессии, является средняя относительная ошибка аппроксимации $\bar{\varepsilon}$, которая определяется по формуле

$$\bar{\varepsilon} = \frac{1}{N} \Sigma \left| \frac{y_i - \tilde{y}_i}{y_i} \right| 100.$$

<u>Коэффициент множественной корреляции</u>

На автомобильном транспорте встречаются задачи, в которых на результативный признак влияет не один, а несколько факторов. Поэтому изучение парных зависимостей обычно малоэффективно. Исходя из этого, в моделях необходимо учитывать совокупное влияние нескольких факторов. Это совокупное влияние факторов определяется методом множественной корреляции.

При оценке взаимного влияния трех и более переменных используют коэффициент множественной корреляции *R*, который для трех переменных определяется по формуле:

$$R = \sqrt{\frac{r_{yx_1}^2 + r_{yx_2}^2 - 2r_{yx_1} r_{yx_2} r_{x_1x_2}}{1 - r_{x_1x_2}^2}}\ .$$

При расчете совокупного коэффициента корреляции необходимо предварительно определить парные коэффициенты корреляции $r_{yx_1}, r_{yx_2}, r_{x_1x_2}$. После того как все они определены, их записывают в квадратную симметричную матрицу:

$$\begin{bmatrix} 1 & r_{yx_1} & r_{yx_2} & r_{yx_3} \\ r_{yx_1} & 1 & r_{x_1x_2} & r_{x_1x_3} \\ r_{yx_2} & r_{x_1x_2} & 1 & r_{x_2x_3} \\ r_{yx_3} & r_{x_1x_3} & r_{x_2x_3} & 1 \end{bmatrix}\ .$$

Тогда множественный коэффициент корреляции определяется формулой

$$R = \sqrt{1 - \frac{D}{D_{11}}}\ ,$$

где D – определитель матрицы парных коэффициентов корреляции;

D_{11} – определитель той же матрицы с вычеркнутыми первой строчкой и первым столбцом, т.е. определитель матрицы парных коэффициентов корреляции между факторами независимыми переменными.

Лекция № 5. Решение задач автомобильного транспорта методами теории вероятностей и математической статистики

На предыдущих лекциях мы рассмотрели методику построения детерминированных математических моделей для опытных данных, представленных в виде двумерных массивов. Для чего использовали метод наименьших квадратов или, более грамотно, "корреляционно-регрессионный анализ". При этом качество модели оценивается количественно с помощью коэффициента корреляции.

На практике в области автомобильного транспорта часто приходится обрабатывать одномерные массивы вида $\{x_i\}$.

Примеры:

– пробег шин до выхода из строя: 10,2; 15,6; 18,8; 9,5; 25,2 и т.д.;

– ежесуточный сход автомобилей с линии: 2; 1; 3; 4; 5; 7; 2; 0 и т.д.

Здесь имеем дело со случайными величинами.

Случайной величиной называется такая величина, которая в результате опыта может принять то или иное значение, которое с точностью нельзя предсказать до опыта.

Между частными значениями случайной величины и вероятностями их появления существует определенная зависимость. Указанная зависимость называется законом распределения случайной величины. Закон распределения случайной величины можно задать в виде таблицы, графика или формулы.

Все случайные величины делятся на ***дискретные*** и ***непрерывные***. Дискретная случайная величина принимает фиксированные значения на отрезке $[a, b]$. Непрерывная случайная величина принимает на отрезке $[a, b]$ любое значение.

Основными вероятностными законами распределения дискретной случайной величины являются биноминальный закон и закон Пуассона, а непрерывной случайной величины – нормальный, показательный, Вейбулла, Релея и др.

Генеральная и выборочная совокупности и их характеристики

При решении задач автомобильного транспорта приходится сталкиваться с понятием ***генеральной*** и ***выборочной*** совокупностей.

Генеральной совокупностью называют совокупность всех объектов (элементов), подлежащих изучению. Очевидно, что подвергать исследованию всю генеральную совокупность затруднительно или нецелесообразно. В связи с этим из генеральной совокупности извлекают лишь некоторую ее часть, называемую выборочной совокупностью или просто выборкой.

Например, изучить автомобиль ВАЗ-2110 на устойчивость или эффективность торможения; или электролампочки на надежность и продолжительность работы и т.п. Определяют характеристики выборочной совокупности и по определенным правилам переносят их на генеральную совокупность.

Генеральная и выборочная совокупности имеют свои характеристики.

Итак, пусть требуется исследовать некоторую генеральную совокупность «Г.с.» (рис. 1), которая характеризуется параметрами:

$M(x)$ – математическое ожидание[1];

$Д(x)$ – дисперсия[2];

$\sigma(x)$ – среднее квадратическое отклонение[3];

v – коэффициент вариации;

$F(x)$ – функция распределения[4];

$f(x)$ – плотность распределения[5].

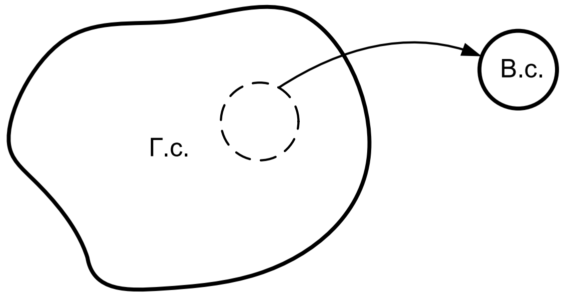


Рис. 1. Схема процесса выборки

Непосредственно вычислить их невозможно. Однако можно оценить (принять) по данным выборочной совокупности. Для чего из генеральной совокупности извлечем выборку «В.с.», для которой методами математической статистики можем вычислить:

$\overline{X}$ – среднее арифметическое;

$Д^*$(x) – статистическая дисперсия;

$\sigma^*(x)$ – среднее квадратическое отклонение;

v – коэффициент вариации;

$W(x)$ – относительная частота;

$F^*(x)$ – статистическая (экспериментальная) функция распределения.

Математическая статистика – это прикладная наука, занимающаяся разработкой методов сбора, описания и обработки результатов наблюдений (опытов), имеющихся в массовых случайных явлениях (процессах).

---

[1] Математическим ожиданием дискретной случайной величины называют сумму произведений всех её возможных значений на их вероятности, т.е. $M(X) = \sum_{i=1}^{\infty} x_i p_i$

[2] Дисперсией (рассеянием) дискретной случайной величины называют математическое ожидание квадрата отклонения случайной величины от её математического ожидания, т.е. $D(X) = M[X - M(X)]^2$

[3] Средним квадратическим отклонением случайной величины X называют квадратный корень из дисперсии

[4] Функцией распределения называют функцию $F(x)$, определяющую вероятность того, что случайная величина X в результате испытания примет значение, меньшее x, т.е. $F(x) = P(X < x)$, где x – действительное число.

[5] Плотностью распределения вероятностей непрерывной случайной величины X называют функцию $f(x)$ – первую производную от функции распределения $F(x)$: $f(x) = F'(x)$.

Теория вероятностей изучает и устанавливает закономерности массовых случайных явлений.

Математическая статистика при решении своих задач использует законы теории вероятностей. Поэтому теория вероятностей и математическая статистика являются взаимосвязанными науками.

Итак, задачи, возникающие при изучении процессов автомобильного транспорта, требуют знаний основных положений теории вероятностей и математической статистики. Если математическая статистика занимается разработкой методов сбора и обработки результатов наблюдений случайных процессов, то теория вероятностей изучает их закономерности.

Используя методы математической статистики, возможно определить числовые характеристики выборочной совокупности. И перенеся их по определенным правилам на генеральную совокупность, оценить числовые характеристики последней.

Для количественной оценки случайной величины используют числовые характеристики.

Среднее арифметическое случайной величины служит характеристикой математического ожидания распределения случайной величины и вычисляется по формуле

$$\overline{X} = (X_1 + X_2 + ... + X_n)/n = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

где X_1, X_2, ..., X_n – значения элементов ряда;

n – число элементов ряда.

Статистическая дисперсия характеризует разброс случайной величины относительно ее среднего значения. Она вычисляется по формуле

$$Д^* = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2.$$

Среднее квадратическое отклонение служит мерой рассеивания случайной величины относительно ее среднего значения и вычисляется по формуле

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2}$$

Коэффициент вариации ряда определяется отношением $v = \frac{\sigma}{\overline{X}}$.

По коэффициенту вариации приближенно определяется закон распределения случайной величины, так при $V \leq 0{,}3$ распределение подчиняется нормальному закону, при $V = 0{,}52$ - закону распределения Релея, а при $V = 1{,}0$ - экспоненциальному (показательному) закону распределения.

Статистическая оценка ***коэффициента асимметрии*** дает дополнительную информацию о форме распределения случайной величины (рис. 2). Асимметрия или скошенность вычисляется по формуле

$$\overline{A} = \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^3}{\sigma^3}.$$

Статистическая оценка ***коэффициента эксцесса*** дает дополнительную информацию о форме распределения случайной величины (рис. 3).

Эксцесс, или островершинность вычисляется по формуле

$$\overline{Э} = \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^4}{\sigma^4} - 3.$$

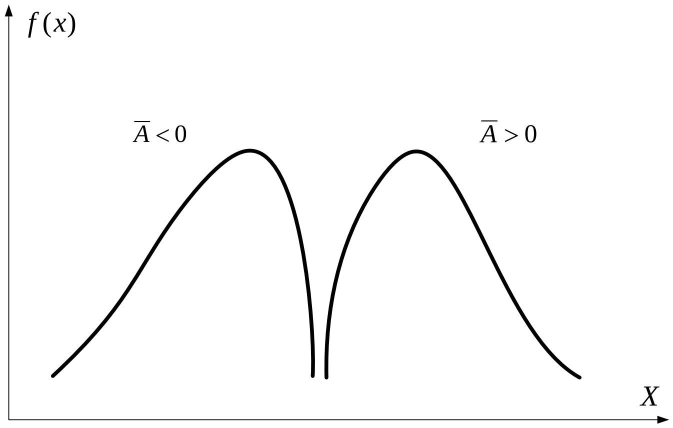


Рис. 2. Положительная и отрицательная симметрия

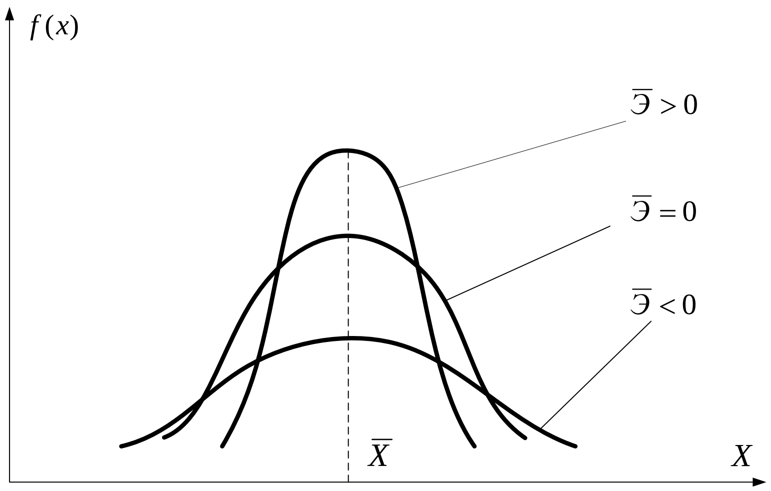


Рис. 3. Положительный и отрицательный эксцессы

Найдя интересующие нас числовые характеристики выборочной совокупности, можем перенести их при определенных условиях на всю генеральную совокупность, т.е. принять:

$$M(x) = \overline{X};\ Д(x) = Д^*(x);\ \sigma(x) = \sqrt{Д^*(x)}.$$

Лекция № 6. Законы распределения случайной величины

Между частными значениями случайной величины и вероятностями их появления существует определенная зависимость. Указанная зависимость называется *законом распределения*.

Закон распределения случайной величины можно задать в виде таблицы, графика или формулы. Различают законы распределения *дискретной случайной величины* и законы распределения *непрерывной случайной величины*.

При табличном задании закона распределения дискретной случайной величины первая строка таблицы содержит возможные значения, а вторая – их вероятности:

X	x_1	x_2	…	x_n
P	p_1	p_2	…	p_n

Приняв во внимание, что в одном испытании случайная величина принимает одно и только одно возможное значение, заключаем, что события $X = x_1$, $X = x_2$, … , $X = x_n$ образуют полную группу; следовательно, сумма вероятностей этих событий, т.е. сумма вероятностей второй строки таблицы, равна единице:

$$p_1 + p_2 + \ldots + p_n = 1.$$

Для наглядности закон распределения дискретной случайной величины можно изобразить и графически, для чего в прямоугольной системе координат строят точки (x_i, p_i), а затем соединяют их отрезками прямых. Полученную фигуру называют *многоугольником распределения*. Ниже мы рассмотрим построение такого многоугольника.

Законы распределения дискретной случайной величины

Основными законами распределения дискретной случайной величины являются биномиальный закон и закон Пуассона.

Биномиальное распределение

Биномиальное распределение возникает при выполнении следующих условий:

– в результате одного испытания может появиться одно из двух противоположных событий A или $\overline{A} = B$;

– вероятности указанных событий от опыта к опыту не меняются и составляют $P(A) = p$ и $P(B) = q$;

– проводится n независимых испытаний.

При выполнении указанных условий возникают различные комбинации таких событий, вероятность появления которых определяется по формуле Бернулли, называемой биномиальным законом распределения

$$P_{(m,n)} = \frac{n!}{m!\left(n-m\right)!} q^{n-m} p^m = C_n^m q^{n-m} p^m \,, \qquad (1)$$

где n – число независимых испытаний;

m – число появления события A; $m = 0, 1, 2, \ldots, n$

$P_{(m,n)}$ – вероятность того, что при n испытаниях событие A появится ровно m раз;

p и q – вероятности появления соответственно событий A и B, где $q = 1 - p$;

$C^m{}_n$ – число сочетаний из n элементов по m.

Напишем биномиальный закон в виде таблицы:

X	n	n-1	…	m	…	0
P	p^n	nqp^{n-1}	…	$C_n^m q^{n-m} p^m$	…	q^n

Для биномиального распределения числовых характеристик: математическое ожидание $M(m)$ и дисперсия $Д(m)$ выражаются с помощью формул

$$M[m] = np \text{ и } Д[m] = npq.$$

Пример: Монета брошена 3 раза. Написать в виде таблицы закон распределения случайной величины X – числа выпадений «герба».

Решение. Вероятность появления «герба» в каждом бросании монеты p = ½, следовательно, вероятность непоявления «герба» $q = 1 - ½ = ½$.

При двух бросаниях монеты «герб» может появиться либо 3 раза, либо 2 раза, либо 1 раз, либо совсем не появиться. Таким образом, возможные значения X таковы: $x_1 = 3$, $x_2 = 2$, $x_3 = 1$, $x_4 = 0$. Найдем вероятности этих возможных значений по формуле Бернулли:

$$P_{(3,3)} = \frac{n!}{m!(n-m)!} q^{n-m} p^m = \frac{3!}{3!(3-3)!} \cdot 0{,}5^0 \cdot 0{,}5^3 = 1 \cdot 1 \cdot 0{,}125 = 0{,}125$$

$$P_{(2,3)} = \frac{n!}{m!(n-m)!} q^{n-m} p^m = \frac{3!}{2!(3-2)!} \cdot 0{,}5^1 \cdot 0{,}5^2 = 3 \cdot 0{,}5 \cdot 0{,}25 = 0{,}375$$

$$P_{(1,3)} = \frac{n!}{m!(n-m)!} q^{n-m} p^m = \frac{3!}{1!(3-1)!} \cdot 0{,}5^2 \cdot 0{,}5^1 = 3 \cdot 0{,}25 \cdot 0{,}5 = 0{,}375$$

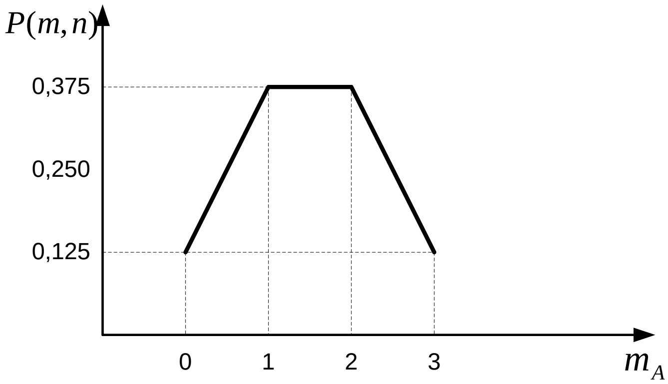
$$P_{(0,3)} = \frac{n!}{m!(n-m)!} q^{n-m} p^m = \frac{3!}{0!(3-0)!} \cdot 0{,}5^3 \cdot 0{,}5^0 = 1 \cdot 0{,}125 \cdot 1 = 0{,}125$$

Напишем искомый закон распределения с помощью таблицы, которую называют рядом распределения вероятности:

X	3	2	1	0
p	0,125	0,375	0,375	0,125

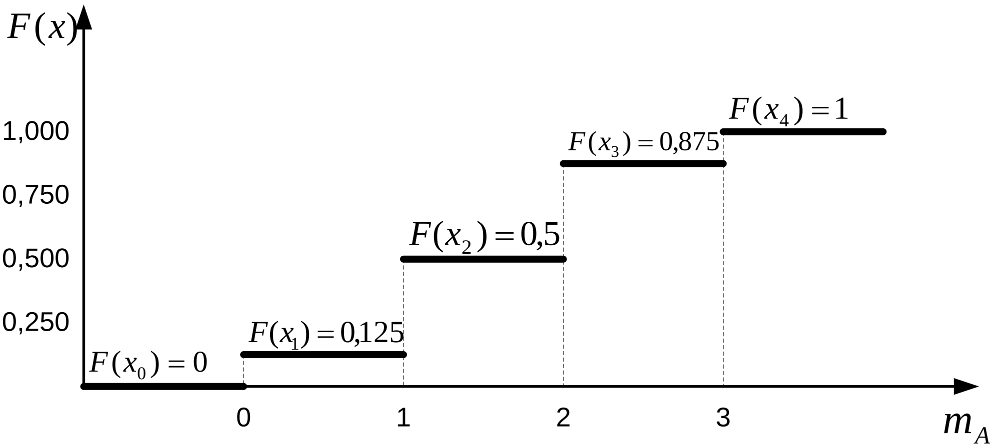
Контроль: 0,125 + 0,375 + 0,375 + 0,125 = 1.

Эти данные можно представить графически, с помощью многоугольника распределения вероятности.



На основании ряда можно построить график функции распределения. *Функцией распределения* называют функцию $F(x)$, определяющую вероятность того, что случайная величина X в результате испытания примет значение, меньшее x, т.е. $F(x) = P(X < x)$. Иногда функцию распределения называют интегральной функцией.

$$F(x) = \begin{cases} P(X \le 0) = 0 \\ P(X \le 1) = 0{,}125 \\ P(X \le 2) = 0{,}5 \\ P(X \le 3) = 0{,}875 \\ P(X > 3) = 1 \end{cases}$$



Распределение Пуассона

Распределение Пуассона представляет собой предельный случай биномиального распределения для условий, когда $p \to 0$; $n \to \infty$ и $np = a$.

Преобразуя выражение биномиального закона при приведенных выше условиях, получим формулу распределения Пуассона:

$$P_{(m,n)} = \frac{a^m}{m!} e^{-a}, \qquad (2)$$

где n – число испытаний;

m – число появлений события A ($m = 0, 1, 2, \dots, n$);

$P_{(m,\,n)}$ – вероятность того, что при n испытаниях событие A появится ровно m раз;

a – параметр закона ($a = np$);

p – вероятность появления события A в одном испытании.

В связи с тем, что вероятность появления отдельных событий в распределении Пуассона характеризуется малой вероятностью ($p \to 0$), закон Пуассона называют законом редких явлений.

Математическое ожидание $M(m)$ и дисперсия $Д(m)$ для распределения Пуассона равны и определяются по выражению

$$M[m] = Д[m] = np = a.$$

Пример: Завод отправил на базу 5000 доброкачественных изделий. Вероятность того, что в пути изделие повредится, равно 0,0002. Найти вероятность того, что на базу прибудут 3 негодных изделия.

Решение. По условию задачи $n = 5000$, $p = 0{,}0002$, $m = 3$. Найдем a:

$$a = np = 5000 \cdot 0{,}0002 = 1.$$

По формуле Пуассона искомая вероятность приближенно равна

$$P_{(m,n)} = \frac{a^m}{m!} e^{-a} = \frac{1^3}{3!} e^{-1} \approx 0{,}06\,.$$

Закон Пуассона описывает:

– поток требований в зону ремонта и ТО;

– поток заявок на запасные части, узлы, агрегаты;

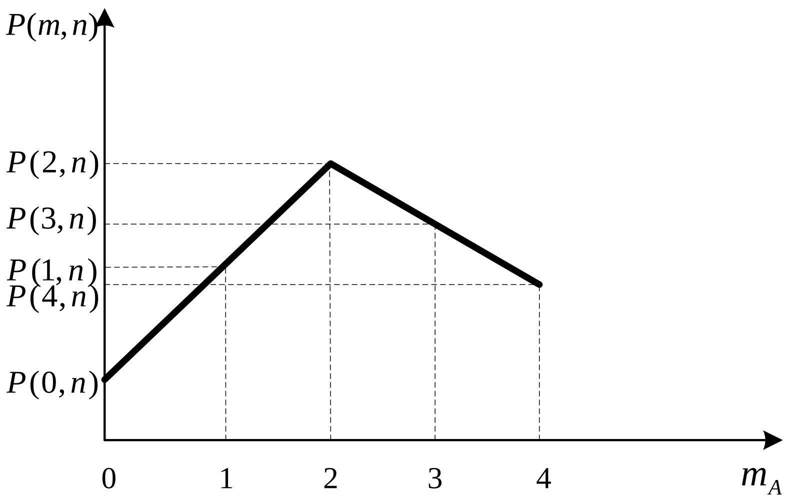
– случайное число отказов в течение фиксированной наработки.

Дискретная случайная величина, кроме формул 1 и 2, также может быть задана:

а) рядом распределения вероятности

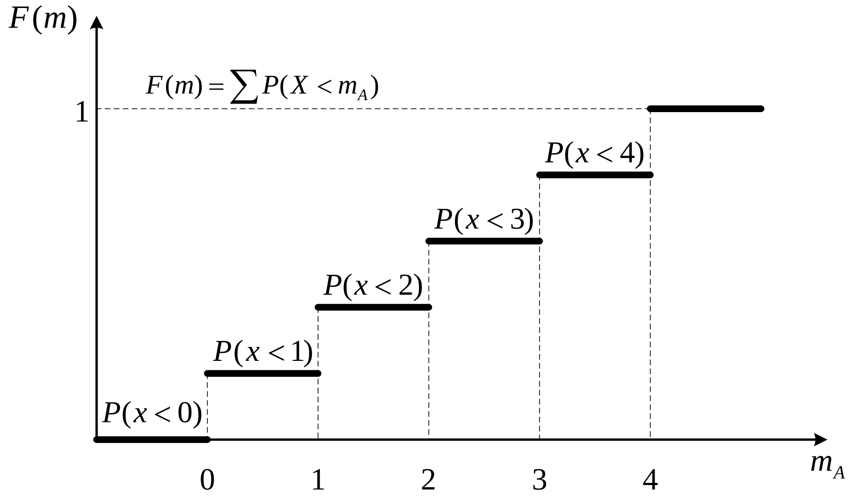
Частота значения события A	m_A	0	1	2	3	…	m
Вероятности, отвечающие частным значениям появления случайной величины	$P(m, n)$	$P(0, n)$	$P(1, n)$	$P(2, n)$	$P(3, n)$	…	$P(m, n)$

б) многоугольником распределения вероятности



в) графиком функции распределения вероятности.

На основании ряда распределения или многоугольника распределения может быть построен график функции распределения.



Основные числовые характеристики дискретной случайной величины в этом случае определяются по формулам

$M[m]=\sum_{i=0}^{m} m_i p_i$ – математическое ожидание;

$Д[m]=\sum_{i=0}^{m} (m_i - M[m])^2 p_i$ – дисперсия.

Лекция № 7. Законы распределения непрерывной случайной величины

Непрерывной называют случайную величину, которая может принимать все значения из некоторого конечного или бесконечного промежутка. Очевидно, число возможных значений непрерывной случайной величины бесконечно.

При описании непрерывных случайных величин производственных процессов автомобильного транспорта широко используются следующие *вероятностные законы*:

– закон равномерного распределения;
– нормальный закон распределения;
– показательный закон распределения;
– законы Релея, Вейбулла и др.

Закон равномерного распределения вероятностей

Распределение вероятностей называют *равномерным*, если на интервале, которому принадлежат все возможные значения случайной величины, плотность распределения сохраняет постоянное значение.

Найдём плотность равномерного распределения $f(x)$, считая, что все возможные значения случайной величины заключены в интервале (a, b), на котором функция $f(x)$ сохраняет постоянные значения. X не принимает значений вне интервала (a, b), поэтому $f(x) = 0$ при $x < a$ и $x > b$.

Найдём постоянную C. Так как все возможные значения случайной величины принадлежат интервалу (a, b), то должно выполняться соотношение

$$\int_a^b f(x)dx = 1, \text{ или } \int_a^b Cdx = 1.$$

Отсюда

$$C = \frac{1}{\int_a^b dx} = \frac{1}{(b-a)}.$$

Итак, искомая плотность вероятности равномерного распределения

$$f(x) = \begin{cases} 0 & \text{при} \quad x \le a, \\ 1/(b-a) & \text{при} \quad a < x \le b, \\ 0 & \text{при} \quad x > b. \end{cases} \tag{1}$$

График плотности равномерного распределения изображен на рис. 1, а график функции распределения – на рис. 2.

Числовые характеристики закона определяются по следующим формулам:

$$M(x) = \int_a^b xf(x)dx = \int_a^b x\frac{1}{b-a}dx = \frac{1}{b-a}\left.\frac{x^2}{2}\right|_a^b = \frac{a+b}{2};$$

$$D(x)=\int_a^b\left(x-\frac{a+b}{2}\right)^2\frac{1}{b-a}dx=\frac{(b-a)^2}{12};$$

$$\sigma(x)=\sqrt{D(x)}=\sqrt{\frac{(b-a)^2}{12}}.$$

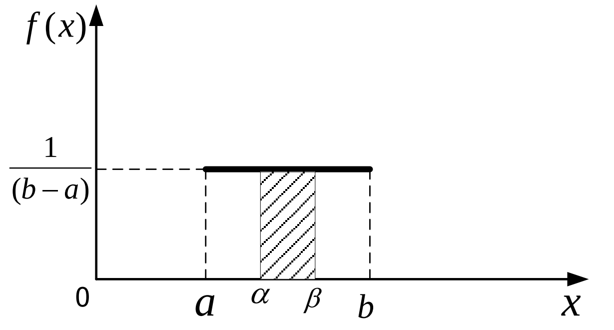


Рис. 1. График плотности равномерного распределения

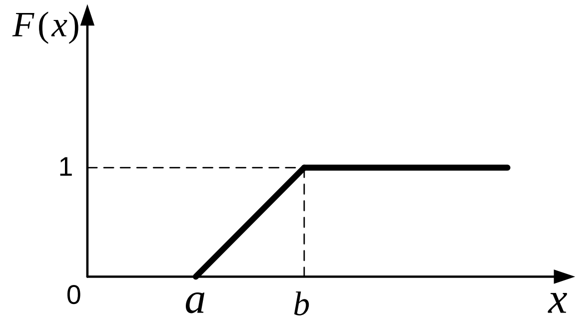


Рис. 2. График функции равномерного распределения

Вероятность попадания случайной величины X в отрезок [α, β], где [α, β] $\in [a, b]$ (см. рис. 1) определяются по формуле

$$P(X\in[\alpha,\beta])=F(\beta)-F(\alpha)=\frac{\beta-\alpha}{b-a}.$$

Равномерный закон распределения описывает процессы, связанные с работой светофоров, используется в задачах массового обслуживания, при статистическом моделировании процессов автомобильного транспорта.

Нормальный закон распределения вероятностей

Нормальным называют распределение вероятностей непрерывной случайной величины, которое описывается плотностью

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

где $f(x)$ – плотность вероятности распределения непрерывной случайной величины;

x_i – текущие значения случайной величины;

$\bar{x}$ – математическое ожидание случайной величины (среднее значение);

σ – среднее квадратическое отклонение случайной величины;

π – число Пи, равное 3,14159265358979;

e – экспонента, равная 2,71828182845904.

Из формулы (2) видим, что нормальное распределение определяется двумя параметрами: $\bar{x}$ и σ. Достаточно знать эти два параметра, чтобы задать нормальное распределение.

График плотности нормального распределения называют *нормальной кривой* или *кривой Гаусса* (рис. 3).

Выясним, как влияют на форму и расположение нормальной кривой значения параметров $\bar{x}$ и σ. Изменение величины параметра $\bar{x}$ (математического ожидания) не изменяет формы нормальной кривой, а приводит лишь к ее сдвигу вдоль оси *x*: вправо, если $\bar{x}$ возрастает, и влево, если $\bar{x}$ убывает. С возрастанием σ максимальная ордината нормальной кривой убывает, а сама кривая становится более пологой, т.е. сжимается к оси *x*; при убывании σ нормальная кривая становится более островершинной и растягивается в положительном направлении вдоль оси *y*.

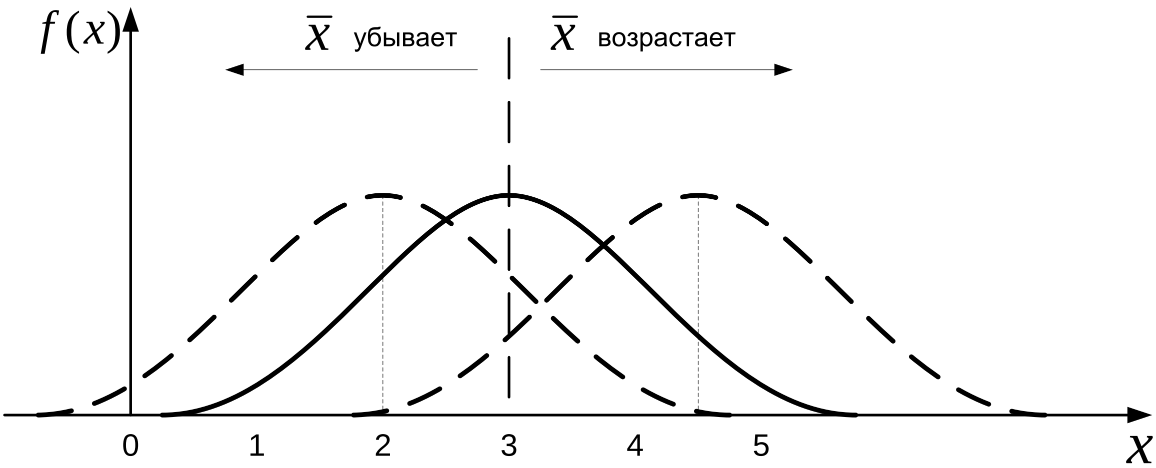


Рис. 3 а). Влияние математического ожидания нормального распределения на форму нормальной кривой

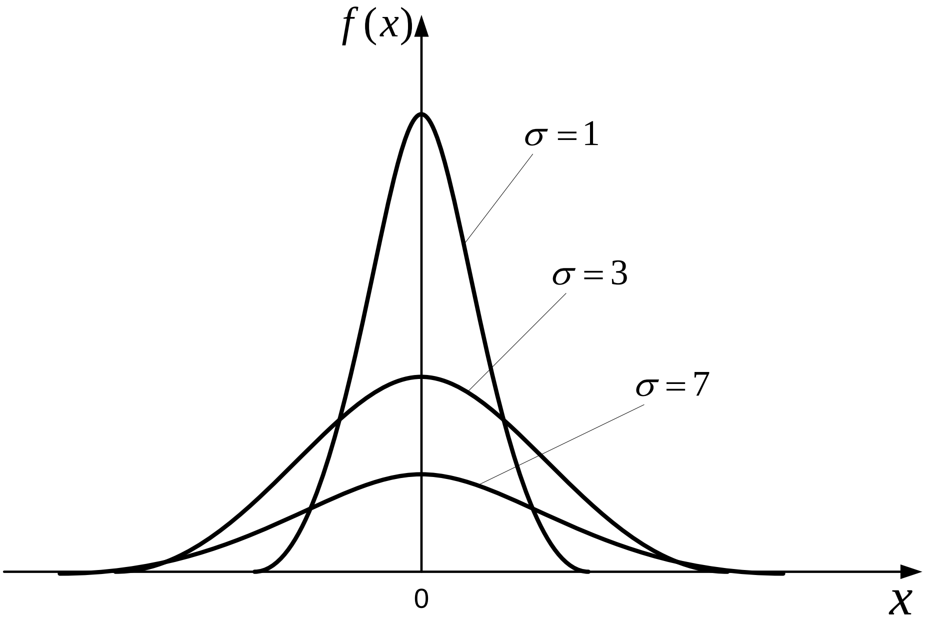


Рис. 3 б). Влияние среднеквадратического отклонения нормального распределения на форму нормальной кривой

Заметим, что при $\bar{x}$ = 0 и σ = 1 нормальную кривую $f(x) = \dfrac{1}{\sqrt{2\pi}} e^{-x^2/2}$ называют *нормированной*.

Числовые характеристики нормального закона:

– математическое ожидание непрерывной случайной величины *X*:

$$M(x) = \int_a^b xf(x)dx;$$

– дисперсия непрерывной случайной величины:

$$D(x) = \int_a^b [x - M(x)]^2 f(x)dx;$$

– среднее квадратическое отклонение непрерывной случайной величины:

$$\sigma(x) = \sqrt{D(x)}.$$

Для нормального закона распределения случайной величины верно правило *трех сигм*: если случайная величина распределена нормально, то абсолютная величина ее отклонения от математического ожидания не превосходит утроенного среднего квадратического отклонения, т.е. $\bar{x} \pm 3\sigma$

Для нормального закона распределения имеет место коэффициент вариации $v = \dfrac{\sigma}{\bar{x}} \le 0{,}33$.

В отрасли автомобильного транспорта нормальный закон описывает:

– пробег до капитального ремонта агрегатов и узлов автомобиля;

– суточные пробеги автомобилей;

– время на операции ТО и их трудоемкости;

– наработка деталей с постепенным характером отказов;

– время на капитальный ремонт агрегатов и т.д.

Показательный закон распределения вероятностей

Показательным (*экспоненциальным*) называют распределение вероятностей непрерывной случайной величины *X*, которое описывается плотностью

$$f(x) = \begin{cases} 0 & \text{при } x < 0 \\ \lambda e^{-\lambda x} & \text{при } x >= 0 \end{cases}$$

где λ – параметр закона (постоянная положительная величина).

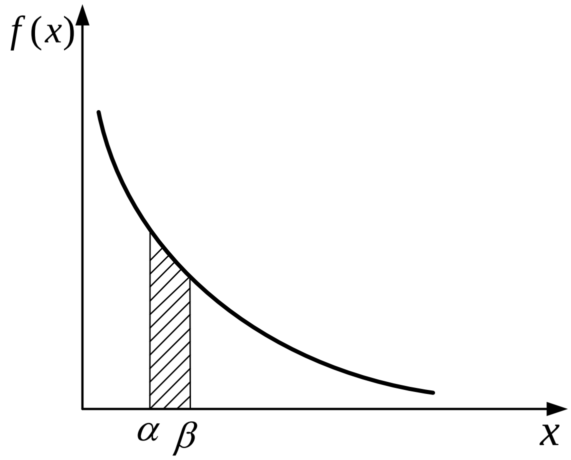


Рис. 4 а). График плотности показательного распределения

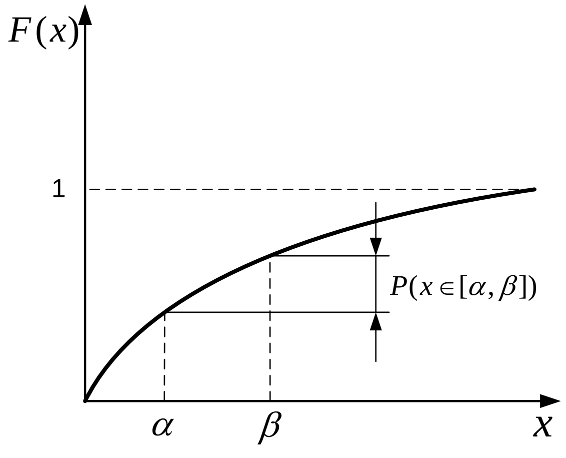


Рис. 4 б). График функции показательного распределения

Лекция № 8. Законы распределения непрерывной случайной величины (продолжение)

Мы видим, что показательное распределение определяется одним параметром λ. Эта особенность показательного распределения указывает на его преимущество по сравнению с распределениями, зависящими от большего числа параметров. Обычно параметры неизвестны и приходится находить их оценки (приближенные значения). Разумеется, проще оценить один параметр, чем два, три и т.д.

Найдём функцию распределения показательного закона (рис. 4 б)

$$F(x) = \int_{-\infty}^{x} f(x)dx = \int_{-\infty}^{0} 0dx + \lambda\int_{0}^{x} e^{-\lambda x}dx = 1 - e^{-\lambda x}.$$

Итак,

$$F(x) = \begin{cases} 0 & \text{при } x < 0, \\ 1 - e^{-\lambda x} & \text{при } x \geq 0. \end{cases}$$

В теории надежности $F(x)$ характеризует вероятность распределения отказов, а $R = 1 - F(x)$ – характеризует вероятность безотказной работы автомобиля (узла, агрегата).

Числовые характеристики показательного закона вычисляются по формулам

$$M(x) = \int_{0}^{\infty} x\lambda e^{-x\lambda}dx = \frac{1}{\lambda};$$

$$D(x) = \int_{0}^{\infty}\left(x - \frac{1}{\lambda}\right)^2 \lambda e^{-\lambda x} = \frac{1}{\lambda^2};$$

$$\sigma(x) = \frac{1}{\lambda}.$$

Сравнивая полученные выражения, заключаем, что $M(x) = \sigma(x) = 1/\lambda$, т.е. математическое ожидание и среднее квадратическое отклонение показательного распределения равны между собой.

Вероятность попадания случайной величины в интервал $[\alpha, \beta]$ (см. рис. 4а) и 4б) определяется выражением

$$P(x \in [\alpha, \beta]) = \int_{\alpha}^{\beta} \lambda e^{-\lambda x}dx = e^{-\lambda\alpha} - e^{-\lambda\beta}.$$

Пример. Время обслуживания автомобиля на СТОА распределено по показательному закону с параметром $\lambda = 3$ авт./час. Определить сколько автомобилей будут обслужены за время от $t = 0{,}13$ до $t = 0{,}7$.

Решение. $P(0{,}3 < t < 0{,}7) = e^{-3 \cdot 0{,}13} - e^{-3 \cdot 0{,}7} = 0{,}553$.

Показательный закон встречается в задачах надежности и массового обслуживания. Ему подчиняются:

– наработка деталей с внезапным характером отказов;

– промежутки времени между поступлениями автомобилей в зону ремонта;

– время восстановления автомобилей при текущем ремонте.

Закон Вейбулла

Плотность распределения вероятности закона Вейбулла имеет вид

$$f(t)=\begin{cases} n\mu^n t^{n-1}e^{-(\mu\cdot t)^n} & \text{при} \quad t\ge 0, n\ge 0, \mu\ge 0, \\ 0 & \text{при} \quad t<0, n<0, \mu<0, \end{cases}$$

где t – случайная величина (время, пробег и т. д.);

n – параметр формы (при n = 1 закон Вейбулла преобразуется в показательный закон, при n = 2 – в закон Релея, при n = 3,25 – в нормальный закон);

μ – параметр масштаба.

Итак, плотность распределения Вейбулла задается двумя параметрами n и μ, что обусловливает широкий диапазон его применения на практике.

Распределение Вейбулла находит широкое применение при исследовании функционирования автотранспортных средств. Хорошо описывает постепенные отказы изделий.

В некоторых случаях вместо μ применяют величину, обработанную по параметру масштаба α = 1/μ, тогда плотность вероятности записывается так:

$$f(t)=\frac{n}{\alpha}\left(\frac{t}{\alpha}\right)^{n-1}e^{-\left(\frac{t}{\alpha}\right)^n}.$$

График плотности распределения Вейбулла приведен на рис. 5.

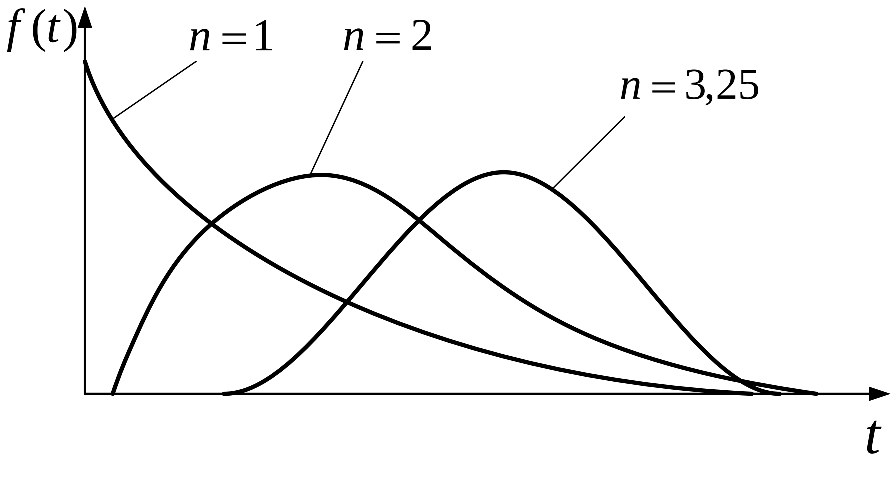


Рис. 5. Графики плотности распределения

Функция распределения закона Вейбулла имеет вид

$$F(t)=\int_0^t n\mu^n t^{n-1}e^{-(\mu t)^n}=1-e^{-\left(\frac{t}{\alpha}\right)^n}.$$

В теории надежности кривая функции распределения $F(t)$ характеризует вероятность отказа изделия, а функция

$$R(t)=1-F(t)=e^{-\left(\frac{t}{\alpha}\right)^n}$$

характеризует вероятность исправного состояния изделия и называется *кривой ресурса*.

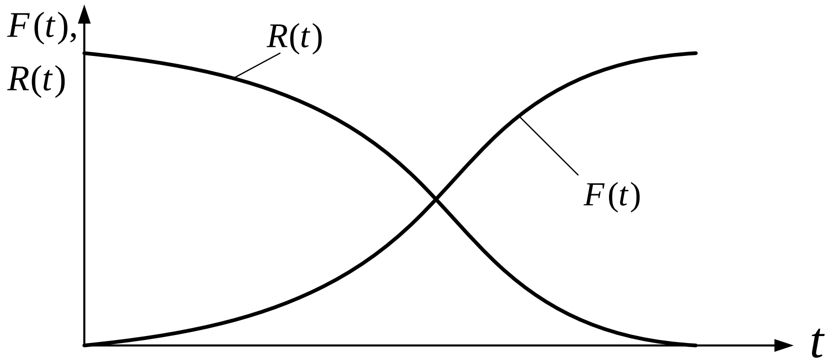


Рис. 6. Функции распределения вероятности отказов и безотказной работы

При решении задач надежности приходится вычислять интенсивность отказов изделий, которая в общем случае равна отношению плотности распределения к вероятности безотказной работы изделия

$$\lambda(t) = f(t)/R(t).$$

Очевидно, что если по мере течения времени вероятность исправной работы изделия уменьшается, то и значение интенсивности отказа изделия изменяется (возрастает).

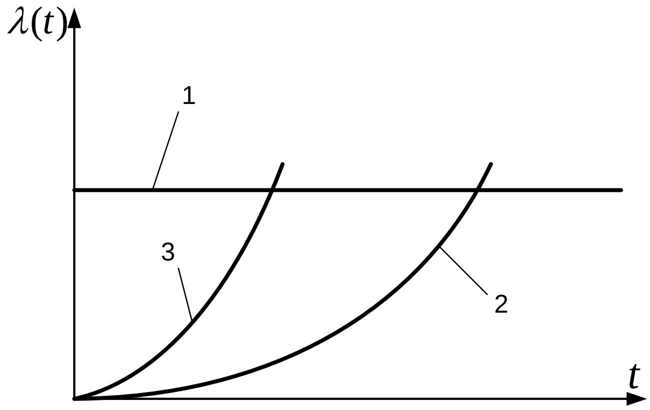


Рис. 7. Кривые интенсивности отказов:
1 – показательного закона (внезапные отказы); 2 – закона Вейбулла (постепенные отказы); 3 – нормального закона (постепенные отказы)

Формулы математического ожидания и дисперсии закона Вейбулла имеют вид:

$$M(t) = \int_0^\infty tn\mu^n t^{n-1} e^{-(\mu t)^n} dt = \int_0^\infty t e^{-(\mu t)^n} d\left(\mu^n t^n\right), \qquad (*)$$

$$D(t) = \int_0^\infty t^2 e^{-(\mu t)^n} d\left(\mu^n t^n\right) - \left[M(t)\right]^2. \qquad (**)$$

Указанные интегралы вычисляются с помощью гамма-функции

$$\Gamma(\alpha)=\int_0^\infty x^{\alpha-1}e^x dx.$$

Значения гамма-функции Эйлера в зависимости от параметра α подсчитаны и приведены в специальной справочной литературе.

Преобразуя выражения (*) и (**) к виду, удобному для применения гамма-функции Эйлера, получим:

$$M(t)=\frac{1}{\mu}\Gamma\left(1+\frac{1}{n}\right),$$

$$D(t)=\frac{1}{\mu^2}\Gamma\left(1+\frac{2}{n}\right)-\left[\frac{\Gamma\left(1+\frac{1}{n}\right)}{\mu}\right]^2.$$

Формула для вычисления коэффициента вариации в этом случае принимает вид:

$$V=\frac{\sigma(t)}{M(t)}=\frac{\sqrt{\Gamma\left(1+\frac{2}{n}\right)-\Gamma^2\left(1+\frac{1}{n}\right)}}{\Gamma\left(1+\frac{1}{n}\right)}=\varphi(n).$$

Как видим, коэффициент вариации является функцией параметра формы закона n. В свою очередь, параметр формы закона n является функцией коэффициента вариации V:

$$n=\psi(V)=\psi\left[\frac{\sigma(t)}{M(t)}\right].$$

Следовательно, если известны $M(t)$ и $\sigma(t)$ закона Вейбулла, то можем определить значения параметра формы n и на основании этого определить параметр масштаба μ.

Для удобства вычисления параметра формы заранее составлены таблицы, приведенные в справочной литературе.

Лекция № 9. Статистическая обработка опытных данных

Пусть исследуется некоторый непрерывный признак, принимающий значения в интервале [a, b]. Например, наработка какого-либо узла или агрегата до отказа:

$$\{x_i\} = 20{,}5; 300; 215; 105{,}5 \text{ тыс. км и т.д.}$$

При обработке и анализе таких статистических данных прибегают к построению интервального вариационного ряда и гистограммы.

Построение интервального вариационного ряда и гистограммы

Вариационным рядом называют представление статистических данных, позволяющих судить о частоте распределяемого признака.

Например:

3, 4, 5, 2, 2, 3, 3, 4, 5, 3 – статистический ряд

2, 2, 3, 3, 3, 3, 4, 4, 5, 5 – ранжированный статистический ряд

2, 4, 2, 2 – вариационный ряд

Пусть имеем доброкачественный объём выборки (статистический ряд). Порядок обработки его следующий:

1) зарегистрированные значения рассматриваемого признака X_i расположить в возрастающем порядке;

2) найти наибольшее $X_{\max}$ и наименьшее $X_{\min}$ значения параметра;

3) определить размах измерения значений параметра $R = X_{max} - X_{min}$;

4) вычислить число интервалов K в зависимости от объёма выборки n

$$K = 1 + 3{,}32\, lg\, n;$$

5) определить ширину частичного интервала h:

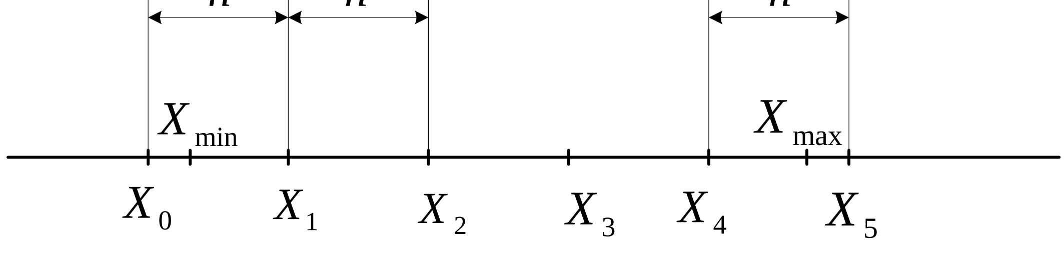
$$h = \frac{R}{K} = \frac{X_{\max} - X_{\min}}{1 + 3{,}32 \lg n};$$

6) определить границы интервалов, для чего установить нулевое (крайнее) значение интервала X_0:

$$X_0 = X_{min} - h/2.$$

Следующие границы интервалов определяются последовательным прибавлением ширины интервала h к предыдущему значению границы: $X_1 = X_0 + h$, $X_2 = X_1 + h$ и т.д. до тех пор, пока X_i не будет больше X_{max};

Примечание: весь диапазон изменения признака от $X_{\min}$ до $X_{\max}$ разбивается на определенное число (K) равных отрезков (интервалов) шириной h



7) определить число элементов значений признаков, попавших в i-й интервал (эту величину называют опытной частотой m_i^* данного интервала);

8) определить относительную частоту (частость) i-го интервала W_i:

$$W_i = m_i/n;$$

9) определить накопленную частость $W_i^н$.

$$W_i^н = \sum_{i=1}^{i} W_i$$

Накопление частости $W^н$ получается путём последовательного прибавления частости W_i очередного интервала: $W_1^н = W_1$, $W_2^н = W_1^н + W_2$, $W_3^н = W_2^н + W_3$, и т.д., для последнего интервала $W_n^н = 1$;

10) результаты расчёта свести в таблицу, которую называют интервальным вариационным рядом:

Номер интервала	Границы интервала $X_{i-1} \div X_i$	Середина интервала $\bar{X}_i$	Частота m_i^*	Частость W_i	Накопленная частость $W_i^н$
1	$X_0 \div X_1$	$\bar{X}_1$	m_1^*	W_1	$W_1^н$
2	$X_1 \div X_2$	$\bar{X}_2$	m_2^*	W_2	$W_2^н$
…	…	…	…	…	…
k	$X_{k-1} \div X_k$	$\bar{X}_k$	m_k^*	W_k	$W_k^н$

Для проверки вычислений: $\sum m_i = n$, $\sum W_i = 1$.

Основные числовые характеристики для интервального вариационного ряда вычисляются по следующим формулам:

– среднее арифметическое:

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} \bar{X}_i m_i^* = \sum_{i=1}^{n} \bar{X}_i W_i$$

– статистическая дисперсия:

$$Д^*(X) = \frac{1}{n}\sum_{i=1}^{n} (\bar{X}_i - \bar{X})^2 m_i^* = \sum_{i=1}^{n} (\bar{X}_i - \bar{X})^2 W_i$$

– среднее квадратическое отклонение:

$$\sigma = \sqrt{Д^*(X)}$$

Графическое выражение закона распределения можно представить в виде гистограммы и накопленной (кумулятивной) кривой.

Гистограмма представляет собой набор прямоугольников, основанием каждого является длина частичного интервала, а высота равна m_i^* или W_i.

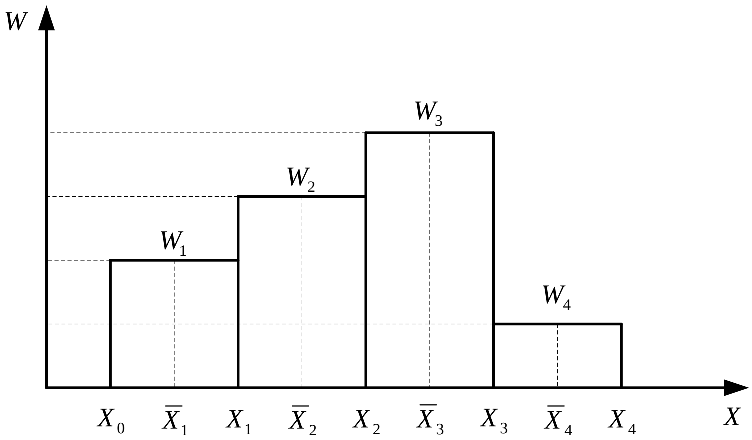


Рис. 1. Гистограмма распределения признака

По построенной гистограмме назначается теоретический закон распределения случайной величины $f(x)$, для которого по формуле

$$F_{\text{т}}(x) = \int_{-\infty}^{x} f(x)dt$$

определяется теоретическая функция распределения $F_{\text{Т}}(x)$.

Кумулятивная кривая строится по накопленным частостям $W_i^{\text{н}}$, она соответствует опытной функции распределения признака $W_i^{\text{н}} = F_{\text{оп}}^*(x)$ (рис. 2). Соответствие опытной $F_{\text{оп}}^*(x)$ и теоретической $F_{\text{т}}(x)$ функции распределения может быть оценено с помощью критерия согласия.

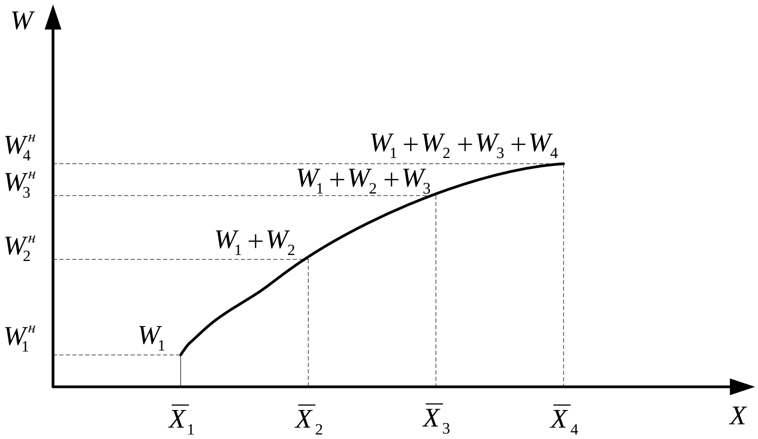


Рис. 2. Кумулятивная кривая

Лекция № 10. Статистическая оценка гипотез. Критерии согласия

Статистической называют гипотезу о виде неизвестного распределения, или о параметрах известных распределений. Например, статистическими являются гипотезы:

1) генеральная совокупность распределена по закону Пуассона;

2) дисперсии двух нормальных совокупностей равны между собой.

В первой гипотезе сделано предположение о виде неизвестного распределения, во второй – о параметрах двух известных распределений.

Наряду с выдвинутой гипотезой рассматривают и противоречащую ей гипотезу. Если выдвинутая гипотеза будет отвергнута, то имеет место противоречащая гипотеза.

Нулевой (*основной*) называют выдвинутую гипотезу H_0.

Конкурирующей (*альтернативной*) называют гипотезу H_1, которая противоречит нулевой.

При выдвижении и принятии указанных гипотез возникают следующие четыре случая:

1. Гипотеза H_0 верна и принимается.

2. Гипотеза H_0 верна, но ошибочно отвергается. Возникающую при этом ошибку называют *ошибкой первого рода*, а вероятность ее появления называют *уровнем значимости* и обозначают α.

$$\alpha = 1 - P_{д},$$

где $P_{д}$ – доверительная вероятность.

3. Гипотеза H_0 не верна и отвергается.

4. Гипотеза H_0 не верна, но ошибочно принимается. Возникающую при этом ошибку называют *ошибкой второго рода*, а вероятность её появления обозначают β.

Для решения отмеченной задачи предложены соответствующие критерии и заранее, при заданном уровне значимости α, подсчитаны и составлены таблицы, в которых помещены критические (табличные) значения указанных критериев.

Статистическим критерием (или просто *критерием*) называют случайную величину K, которая служит для проверки нулевой гипотезы.

Для проверки гипотезы по данным выборок вычисляют частные значения входящих в критерий величин и таким образом получают опытное (наблюдаемое) значения критерия.

При практической проверке рассматриваемых гипотез происходит сопоставление опытных значений критерия $K_{оп}$ с табличным значением критерия $K_{теор}(K_{табл})$ и далее в зависимости от соотношения $K_{опытн} \geq< K_{теор}$ принимают или отвергают выдвинутую гипотезу.

Рассмотрим порядок статистической проверки правдоподобия гипотезы о принадлежности опытных данных к заданному виду вероятностного закона. Решение этой задачи производится в два этапа:

– по виду гистограммы или, исходя из физической сущности рассматриваемого явления, делают предварительное суждение, т.е. выдвигается гипотеза о принадлежности опытных данных к конкретному вероятностному закону;

– применяя метод моментов, производят проверку правдоподобия выдвинутой гипотезы.

Сущность метода моментов состоит в том, что параметры сглаживающего закона должны сохранить основные черты статистического распределения, т.е. чтобы было равенство математического ожидания и дисперсии статистического и сглаживающего распределений.

Проверка правдоподобия гипотезы о принадлежности опытных данных к заданному виду вероятностного закона может производиться с помощью критериев: Пирсона, Колмогорова, Романовского, Фишера, Кохрена, Стьюдента и др.

1. *Критерий χ^2 Пирсона*

Критерий χ^2 Пирсона записывается в виде следующего альтернативного условия:

$$P_{\text{опытн}}(\chi^2, r) = \begin{cases} \geq \alpha & \text{– гипотеза о принадлежности опытных данных к рассматриваемому закону не отвергается;} \\ < \alpha & \text{– гипотеза о принадлежности опытных данных к рассматриваемому вероятностному закону отвергается} \end{cases}$$

где $\alpha = 0,05$,

χ^2 вычисляется по формуле

$$\chi^2 = \sum_{i=1}^{k} \frac{\left(m_i^* - m_i\right)^2}{m_i}$$

r - число степеней свободы, $r = k - 1 - s$;

k - число интервалов гистограммы;

s - число параметров предполагаемого распределения.

В частности, если предполагаемое распределение – нормальное, то оценивают два параметра (математическое ожидание и среднее квадратическое отклонение), поэтому $s = 2$ и число степеней свободы $r = k - 1 - 2 = k - 3$.

Если, например, предполагают, что генеральная совокупность распределена по закону Пуассона, то оценивают один параметр λ, поэтому s=1 и r=k–2.

Значения вероятностей $P_{\text{опытн}}(\chi^2, r)$, вычисленные в зависимости от числа степеней свободы r и опытного значения χ^2 приведены в специальной таблице.

2. *Критерий Романовского*

$$K_p = \frac{\chi^2 - K}{\sqrt{2K}} = \begin{cases} \le 3 & \text{– гипотеза о принадлежности опытных данных к рассматриваемому закону не отвергается;} \\ > 3 & \text{– гипотеза о принадлежности опытных данных к рассматриваемому вероятностному закону отвергается} \end{cases}$$

где χ^2 – критерий Пирсона; K – число интервалов.

3. *Критерий Колмогорова*

Критерий Колмогорова записывается в виде альтернативного условия:

$$P(\lambda) = P\left\{\max\left|F^*_{\text{опытн}}(x) - F_{\text{теор}}(x)\right|\sqrt{n}\right\} = \begin{cases} \ge P_{\text{Д}} & \text{– гипотеза о принадлежности опытных данных к рассматриваемому вероятностному закону не отвергается;} \\ < P_{\text{Д}} & \text{– гипотеза о принадлежности опытных данных к рассматриваемому вероятностному закону отвергается,} \end{cases}$$

где n – объем выборки (число всех испытаний);

$F^*(x)_{\text{опытн}}$ – опытное значение функции распределения;

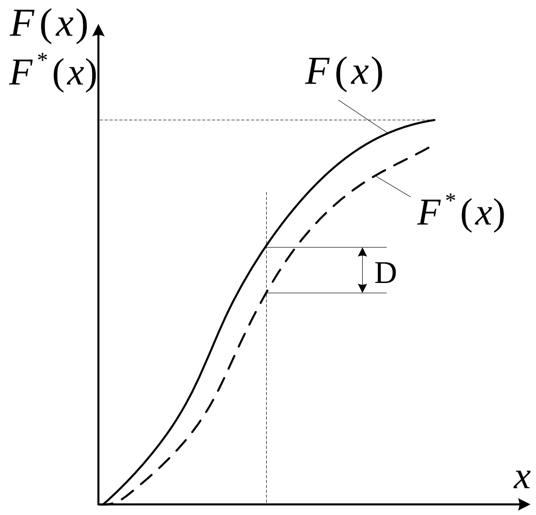
$F(x)_{\text{теор}}$ – теоретическое значение функции распределения;

$P_{\text{Д}}$ – доверительная вероятность (для автомобильного транспорта $P_{\text{Д}}$ = 0,95).

Для критерия Колмогорова также имеется заранее составленная таблица.

Критерий Колмогорова считается наиболее простым критерием проверки гипотезы о виде закона распределения. Суть его состоит в следующем.

1). По результатам “n” произведённых измерений (опытов) строится эмпирическая функция распределения $F^*(x)$.



2). На этом же графике строится предполагаемая теоретическая функция распределения $F(x)$, рассчитанная для принятого закона распределения.

3). Определяется максимальная величина разности их ординат, которая в дальнейшем берётся по модулю:

$$\text{D} = \max\left|F^*(x) - F(x)\right|,$$

т.е. определяется максимальное значение абсолютной величины разности между эмпирической функцией распределения

$F^*(x)$ и теоретической функцией распределения $F(x)$.

4). Вычисляется величина $\lambda = D\sqrt{n}$.

5). Затем по таблице находим величину $P(\lambda)$. Если вероятность $P(\lambda)$ мала (менее 0,6), то гипотеза отвергается. При сравнительно большой вероятности $P(\lambda)$ (более 0,6) гипотеза принимается.

Лекция № 11. Теория случайных процессов и её инженерное приложение

Теория случайных процессов описывает процессы, протекающие в системах массового обслуживания (СМО). К СМО относятся АЗС, СТОА, ремонтные мастерские и т.п.

Случайные процессы и их классификация

Определение: Случайными процессами называют процессы, которые в результате опыта могут принять заранее неизвестный вид.

Чтобы придать этому вопросу прикладной характер мы будем изучать случайные процессы не вообще, а случайные процессы, протекающие в системах ***S***. При этом под системой ***S*** будем понимать техническое устройство (автомобиль, ремонтная мастерская, СТОА, АЗС, и т.д.).

Если состояние системы ***S*** меняется случайным образом, то будем говорить, что в системе протекает случайный процесс.

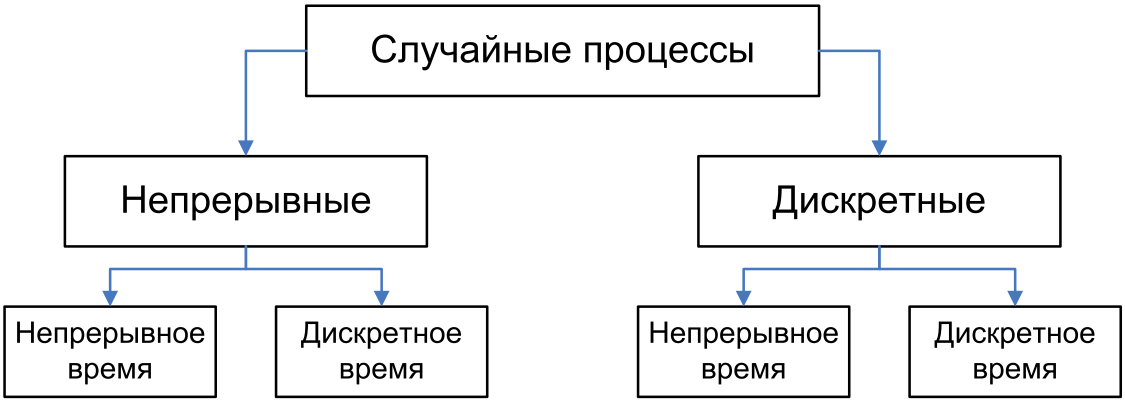


Рис. 1. Классификация случайных процессов

Случайные процессы классифицируются «по времени» и «по состояниям» системы.

Случайный процесс называется процессом с ***непрерывным состоянием***, если его течение в любой момент времени представляет собой непрерывную случайную величину и множество её значений несчетно. Проще говоря, этот процесс характеризуется плавным, постепенным переходом из состояния в состояние. Например, процесс изменения напряжения в осветительной сети, температура двигателя и т.д.

Случайный процесс называется процессом ***с дискретным состоянием***, если в любой момент времени множество его состояний конечное или счетное. Этот процесс характеризуется тем, что система S скачкообразно (мгновенно) время от времени перескакивает (переходит) из одного состояния в другое. Для такого процесса возможные состояния можно пронумеровать, например, S_1, S_2, S_3, S_4 и т.д. и составить геометрическую схему, которую называют ***графом состояний***. В этом случае каждое состояние изображают в виде прямоугольника или кружка, а возможные переходы из состояния в состояние – стрелками, соединяющие эти прямоугольники (кружки).

Пример: Система S – автомобиль, может находиться в одном из пяти возможных состояний:

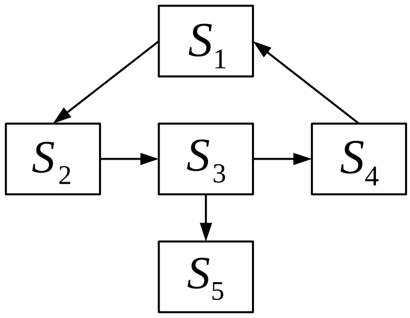
S_1 – исправен, работает;

S_2 – неисправен, ожидает ремонта;

S_3 – диагностируется;

S_4 – ремонтируется;

S_5 – списан.



Случайный процесс называется процессом ***с дискретным временем***, если система, в которой он протекает, может менять свое состояние только в моменты t_1, t_2, …, t_n, число которых конечно. В промежутках времени между этими моментами система S сохраняет свое состояние.

Случайный процесс называется процессом ***с непрерывным временем***, если переходы системы из одного состояния в другое могут происходить в любой момент времени наблюдаемого периода.

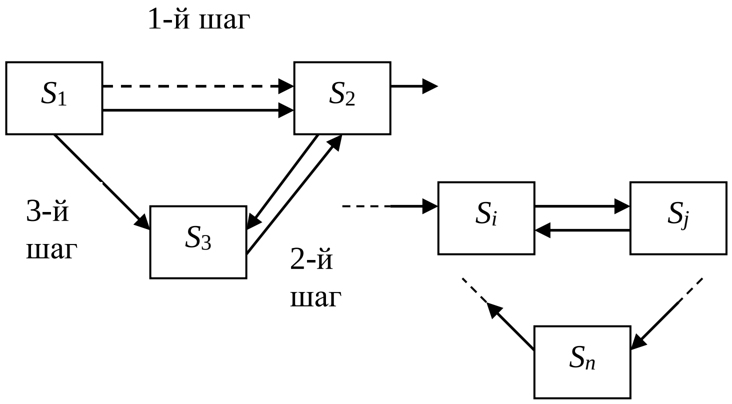
Все эти случайные процессы относятся к так называемым марковским случайным процессам.

Случайный процесс с дискретными состояниями и дискретным временем

Пусть имеется система S с дискретными состояниями S_1, S_2, …, S_i, …, S_n. Случайные переходы (перескоки) системы из одного состояния в другое могут происходить только в определенные (фиксированные) моменты времени t_1, t_2, t_3 … Эти моменты принято называть шагами процесса; $t_0 = 0$ – его начало. Сам процесс представляет собой случайное блуждание системы S по состояниям. После первого шага система может оказаться в одном (и только в одном) из своих возможных состояний $S_1^{(1)}, S_2^{(1)}, ..., S_i^{(1)}, ..., S_n^{(1)}$, на втором шаге – $S_1^{(2)}, S_2^{(2)}, ..., S_i^{(2)}, ...S_n^{(2)}$ и т.д.

Предположим, что граф состояний системы S имеет вид, представленный на рис. 3.

Процесс блуждания системы S по состояниям можно представить как последовательность или «цепь» событий, состоящих в том, что в начальный момент времени $t_0 = 0$ система находится в одном из состояний (например, в состоянии $S_2^{(0)}$; в момент первого шага перешла из него скачком в состояние $S_2^{(1)}$, из которого на втором шаге перешла в $S_2^{(2)}$, на третьем шаге перешла из $S_3^{(2)}$ в $S_3^{(3)}$, что означает задержку системы в состоянии S_3, и т.д.



Переход системы из состояния S_i в него же означает задержку системы в состоянии S_i. Процесс задержки возможен для любого состояния, однако он стрелками на графе не отмечается, так как все расчеты можно вести и без них.

При описании таких случайных процессов вводят две основных характеристики: вероятности состояний – $P_i(k)$ и вероятности перехода – P_{ij}.

Вероятность состояния $P_i(k)$ означает вероятность того, что система S после k-го шага находится в S_i -ом состоянии:

$P_i(k) = P\{S(k) = S_i\}$ $(i = 1, 2, \ldots, n;\ k = 0, 1, \ldots)$.

Вероятности перехода P_{ij} означают вероятность перехода системы S из состояния S_i в состояние S_j за один шаг. Вероятность P_{ii} означает вероятность задержки системы в состоянии S_i за один шаг.

Если известны вероятности перехода из одного состояния в другое за один шаг, то их можно записать в виде квадратичной таблицы (матрицы) размерности mn:

$$\left\| P_{ij} \right\| = \begin{Vmatrix} P_{11} & P_{12} & P_{1j} & P_{1n} \\ P_{21} & P_{22} & P_{2j} & P_{2n} \\ P_{i1} & P_{i2} & P_{ij} & P_{in} \\ P_{n1} & P_{n2} & P_{nj} & P_{nn} \end{Vmatrix}$$

По главной диагонали матрицы стоят вероятности задержки системы в данном состоянии S_j $(j = 1, 2, \ldots, n)$.

Так как на каждом шаге системы S может находиться только в одном из возможных состояний, то для любой i-й строки матрицы сумма всех стоящих в ней вероятностей P_{ij} равна единице:

$$\sum_{i=1}^{n} P_{ij} = 1.$$

В силу последнего условия вероятность задержки для любого состояния можно получить из выражения

$$P_{ij} = 1 - \sum_{j \neq i} P_{ij}.$$

Чтобы найти вероятность состояний $P_i(k)$ необходимо также знать начальное распределение вероятностей. Так, если известно, что в начальный момент система S находится во вполне определенном состоянии, например S_1, то вероятность $P_1(0)$ равна единице, а все остальные – нулю:

$$P_1(0) = 1, \quad P_2(0) = P_3(0) = ... = P_n(0) = 0.$$

При нахождении вероятностей состояний на k-ом шаге $P_i(k)$ (k = 1, 2, …) удобно пользоваться так называемым размеченным графом состояний системы S, где возле каждой стрелки, ведущей из состояния S_i в состояние S_j, проставлена переходная вероятность P_{ij}; вероятности задержки на размеченном графе не проставляются, а просто получаются дополнением до единицы сумм вероятностей, стоящих у всех стрелок, ведущих из данного состояния S_i.

Рассмотрим размеченный граф состояний (рис. 4.). Для этого графа состояний вероятности задержки равны:

$$P_{11} = 1 - (P_{12} + P_{13});\ P_{22} = 1 - P_{23};\ P_{23} = 1 - P_{32}.$$

Теперь покажем, как найти вероятность нахождения системы S на k-м шаге в состоянии S_i (i = 1, 2, 3), если задана матрица переходных вероятностей $\|P_{ij}\|$ или, что равнозначно, размеченный граф состояний.

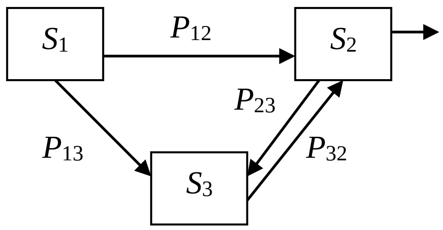


Рис. 4

Предположим, что в начальный момент (перед первым шагом) система находится в состоянии S_1, тогда для начального момента(0)

$$P_1(0) = 1;\ P_2(0) = 0;\ P_3(0) = 0.$$

Очевидно, вероятности состояний после первого шага будут равны:

$$P_1(1) = P_{11};\ P_2(1) = P_{12};\ P_3(1) = P_{13},$$

так как за первый шаг система может перейти в состояние S_1, S_2, S_3 с вероятностями P_{11}, P_{12}, P_{13}.

Вероятности состояний после второго шага вычисляются по формуле полной вероятности с гипотезами (предположениями), что после первого шага система была в одном из состояний S_1, S_2, S_3. Получим:

$$\begin{aligned} P_1(2) &= P_1(1)P_{11} + P_2(1)P_{21} + P_3(1)P_{31} \\ P_2(2) &= P_1(1)P_{12} + P_2(1)P_{22} + P_3(1)P_{32} \\ P_3(2) &= P_1(1)P_{13} + P_2(1)P_{23} + P_3(1)P_{33} \end{aligned} \qquad (1)$$

В общем виде уравнения (1) можно записать так:

$$P_i(2) = \sum_{j=1}^{n} P_j(1)P_{ji}\ .$$

Таким образом, мы выразим распределение вероятностей на втором шаге через распределение вероятностей на первом шаге и матрицу $\|P_{ij}\|$. Переходя таким же способом от k = 2 к k = 3 и т.д., получим рекуррентную формулу

$$P_i(k) = \sum_{j=1}^{n} P_j(k-1)P_{ji} \quad (k = 1,2,...,\ i = 1,2,...,n),$$

позволяющую шаг за шагом вычислять вероятности состояний системы S.

Лекция № 12. Теория случайных процессов и её инженерное приложение (продолжение)

Случайный процесс с дискретными состояниями и непрерывным временем

Рассмотрим систему S, у которой переход из состояния S_i в состояние S_j происходит в произвольный момент времени t. Поставим задачу найти вероятности состояний такого процесса, т.е.

$$P_1(t), P_2(t),\ldots, P_i(t),\ldots, P_n(t),$$

где $P_i(t)$ - вероятность того, что в момент времени t система будет находиться в состоянии S_i ($i = 1, 2, \ldots, n$).

Отметим, что система вероятностей состояний для любого момента времени равна 1, так как события состоящие в том, что в момент времени t система находится в состояниях S_1, S_2, …, S_{n} несовместимы и образуют полную группу событий

$$\sum_{i=1}^{n} P_i(t) = 1.$$

Для решения поставленной задачи введем новую характеристику случайного процесса λ_{ij} – плотность вероятности перехода (интенсивность потока событий).

Плотностью вероятности перехода λ_{ij} называют предел отношения вероятности перехода системы за время Δt из состояния S_i в состояние S_j, к длине промежутка Δt, т.е.

$$\lambda_{ij} = \lim_{\Delta t \to 0} \frac{P_{ij}(\Delta t)}{\Delta t}, \tag{1}$$

где $P_{ij}(\Delta t)$ – вероятность того, что система, находящаяся в момент времени t в состоянии S_i, за время Δt переходит из него в состояние S_j.

Из выражения (1)

$$P_{ij}(\Delta t) = \lambda_{ij}\Delta t.$$

Если известны плотности вероятностей перехода λ_{ij} для всех пар состояний S_i, S_j, то представляется возможность определить вероятности состояний $P_1(t)$, $P_2(t)$, …, $P_n(t)$ как функцию времени. Для чего проставим λ_{ij} на граф состояний, получим размеченный граф состояний. Запишем систему дифференциальных уравнений Колмогорова, решая которую, находим вероятности состояний $P_1(t)$, $P_2(t)$, …, $P_n(t)$.

Правила составления дифференциальных уравнений Колмогорова рассмотрим на конкретном примере.

Пусть имеем размеченный граф состояний дискретного случайного процесса с непрерывным временем (рис. 1), где S_1, S_2, S_3 – возможные состояния системы.

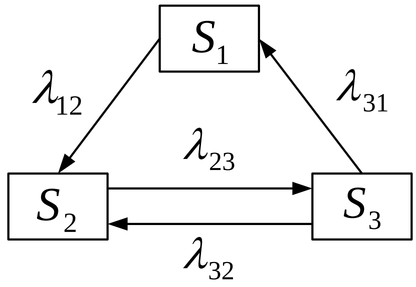


Рис. 1. Размеченный граф состояний

Отметим, что в моменты времени t система S может находиться в любом из состояний S_1, S_2, S_3.

Рассмотрим состояние S_1 и найдем по формуле полной вероятности $P_1(t+\Delta t)$ – вероятность того, что в момент времени $(t+\Delta t)$ система S будет находиться в состоянии S_1, которая определится выражением:

$$P_1(t+\Delta t) = P_1(t)(1-\lambda_{12}\Delta t) + P_2(t)\lambda_{21}\Delta t, \qquad (2)$$

где $P_1(t)(1-\lambda_{12}\Delta t)$ – вероятность того, что в момент времени t система была в состоянии S_1 и за время Δt не перейдет в другое состояние (в нашем случае S_2);

$P_3(t)\lambda_{31}\Delta t$ – вероятность того, что в момент времени t система была в состоянии S_2 и за время Δt перешла из него в состояние S_1.

Преобразуем выражение (2): перенесем $P_1(t)$ в левую часть, разделим на Δt и перейдем к пределу, получим

$$\frac{dP_1(t)}{dt} = \lim_{\Delta t\to 0}\frac{P_1(t+\Delta t)-P_1(t)}{\Delta t} = -\lambda_{12}P_1(t)+\lambda_{31}P_3(t).$$

Рассуждая аналогично для состояний S_2 и S_3, получим систему дифференциальных уравнений, называемую уравнениями Колмогорова:

$$\left.\begin{aligned}\frac{dP_1(t)}{dt} &= -\lambda_{12}P_1(t)+\lambda_{31}P_3(t)\\ \frac{dP_2(t)}{dt} &= -\lambda_{23}P_2(t)+\lambda_{12}P_1(t)+\lambda_{32}P_3(t)\\ \frac{dP_3(t)}{dt} &= -(\lambda_{31}+\lambda_{32})P_3(t)+\lambda_{23}P_2(t)\end{aligned}\right\} \qquad (3)$$

Для решения системы уравнений (3) необходимо выбрать начальные условия. Например, если при $t = 0$ система находится в состоянии S_1, то для начального условия (t = 0) будут справедливы значения:

$$P_1(0) = 1;\ P_2(0) = P_3(0) = 0.$$

Одно из уравнений системы (3) можно заменить нормировочным условием $\sum_{i=1}^{n} P_i(t) = 1$, которое справедливо для любого момента времени t.

Примечание. Структура записи уравнений системы (3) подчинена определенным правилам:

а) в левой части каждого уравнения записывается производная вероятности рассматриваемого состояния;

б) правая часть содержит столько членов (слагаемых), сколько стрелок связано с данным состоянием;

в) каждое слагаемое равно произведению плотности вероятности, соответствующей данной стрелке, умноженной на вероятность того состояния, из которого исходит стрелка;

г) если стрелка направлена из состояния, то соответствующее слагаемое имеет знак «минус», а если в состояние – знак «плюс».

Отметим, что при $t \to \infty$ в системе устанавливается стационарный режим, для которого вероятности P_1, P_2, P_3 уже не зависят от времени, так как

$$\lim_{t\to\infty} P_i(t) = P_i \,.$$

Вероятности P_1, P_2, P_3 и т.д. называют предельными вероятностями состояний, для вычислений которых необходимо в системе (3) положить, что все производные равны нулю, т.е.

$$\frac{dP_i}{dt} = 0 \,.$$

Действительно, в предельном (установившемся) режиме все вероятности состояний постоянны, значит и их производные равны 0.

В этом случае система дифференциальных уравнений (3) превращается в систему обычных линейных алгебраических уравнений (4), которое совместно с нормировочным условием $\sum P_i = 1$ дает возможность вычислить все вероятности состояний

$$\left.\begin{aligned} \lambda_{12}P_1 - \lambda_{31}P_3 &= 0 \\ \lambda_{23}P_2 - \lambda_{12}P_1 - \lambda_{32}P_3 &= 0 \\ (\lambda_{31} + \lambda_{32})P_3 - \lambda_{23}P_2 &= 0 \end{aligned}\right\} \qquad (4)$$

Процесс «гибели» и «размножения»

Случайный процесс, который имеет размеченный граф в виде цепочки (рис. 2), в которой каждое состояние связано прямой и обратной связью с соседними состояниями, при этом в любой момент времени он может увеличиться на единицу, либо уменьшиться на единицу, либо остаться неизменным, называют процессом гибели и размножения.

Интенсивность потока событий, ведущих к увеличению функции $X(t)$ («размножению»), обозначены λ_{ij} , где индекс i соответствует индексу того состояния, из которого выходит стрелка, а индекс j – тому состоянию, в которое входит стрелка. Интенсивности потока событий, ведущих к уменьшению функции («гибели»), обозначены λ_{ji}.

Очевидно, что одномерный закон распределения процесса гибели и размножения $X(t)$ можно описать с помощью системы уравнений Колмогорова для размеченного графа, изображенного на рис. 2.

$$\left.\begin{aligned}
&\frac{dP_0(t)}{dt} = -\lambda_{01}P_0(t) + \lambda_{10}P_1(t)\\
&\frac{dP_1(t)}{dt} = -(\lambda_{12} + \lambda_{10})P_1(t) + \lambda_{01}P_0(t) + \lambda_{21}P_2(t)\\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\\
&\frac{dP_n(t)}{dt} = -\lambda_{n,n-1}P_n(t) + \lambda_{n-1,n}P_{n-1}(t)
\end{aligned}\quad,\right\} \qquad (5)$$

где $P_i(t) = P\{S(t) = S_i\} = P\{X(t) = i\}$ $(i = 0, 1, 2, \dots, n)$.

Систему уравнений (5) нужно решать при начальных условиях $P_i(0) \ge 0$ $(i = 0, 1, 2, \dots, n)$, при этом $\sum_{i=0}^{n} P_i(0) = 1$.

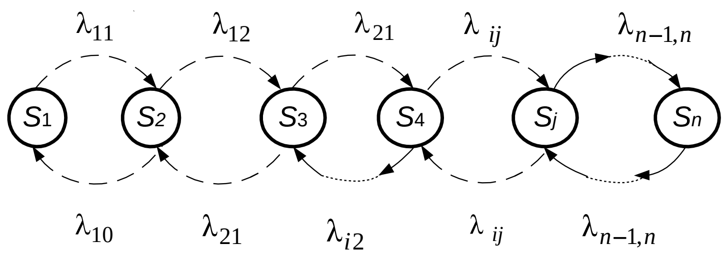


Рис. 2

Предельные вероятности состояний для простейшего процесса гибели и размножения, находящегося в стационарном режиме, могут быть получены из системы уравнений (5), в которой нужно все интенсивности потоков взять постоянными, а все производные вероятностей состояний положить равными нулю:

$$\left.\begin{aligned}
&-\lambda_{01}P_0 + \lambda_{10}P_1 = 0\\
&-\lambda_{12}P_1 - \lambda_{10}P_1 + \lambda_{01}P_0 + \lambda_{21}P_2 = 0\\
&-\lambda_{23}P_2 - \lambda_{21}P_2 + \lambda_{12}P_1 + \lambda_{32}P_3 = 0\\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\\
&-\lambda_{n,n-1}P_0 + \lambda_{n-1,n}P_{n-1} = 0
\end{aligned}\right\} \qquad (6)$$

Решим систему (6) при условии, что

$$P_0 + P_1 + P_2 + \dots + P_n = 1.$$

Выразим из первого уравнения системы (6) вероятность P_1 через P_0, получим

$$P_1 = \frac{\lambda_{01}}{\lambda_{10}}P_0.$$

Так как $\lambda_{01}P_0 = \lambda_{10}P_1$, то из второго уравнения системы (6) имеем:

$$P_2 = \frac{\lambda_{12}}{\lambda_{21}}P_1 = \frac{\lambda_{01}\lambda_{12}}{\lambda_{10}\lambda_{21}}P_0.$$

Далее аналогично имеем:

$$\left.\begin{aligned}
&P_3 = \frac{\lambda_{23}}{\lambda_{32}} P_2 = \frac{\lambda_{01}\lambda_{12}\lambda_{23}}{\lambda_{10}\lambda_{21}\lambda_{32}} P_0 \\
&P_j = \frac{\lambda_{ij}}{\lambda_{ji}} P_i \\
&P_n = \frac{\lambda_{n-1,n}}{\lambda_{n,n-1}} P_{n-1} = \frac{\lambda_{01}\lambda_{12}\ldots\lambda_{n-1,n}}{\lambda_{10}\lambda_{21}\ldots\lambda_{n,n-1}} P_0
\end{aligned}\right\} \qquad (7)$$

Таким образом, любую предельную вероятность P_i ($i = 1, 2, \ldots, n$) можно выразить через предельную вероятность P_0. Вероятность P_0 можно найти из нормировочного условия

$$P_0 + P_0 \sum_{i=1}^{n} \prod_{k=1}^{i} \frac{\lambda_{k-1,k}}{\lambda_{k,k-1}} = 1.$$

Откуда

$$P_0 = \frac{1}{1 + \sum_{i=1}^{n} \prod_{k=1}^{i} \frac{\lambda_{k-1,k}}{\lambda_{k,k-1}}} \qquad (8)$$

Формулы (7) и (8) дают возможность вычислить предельные вероятности состояний простейшего процесса гибели и размножения, находящегося в стационарном режиме при конечном числе состояний.

Лекция № 13. Моделирование методами теории массового обслуживания

Теория массового обслуживания описывает процессы, протекающие в системах массового обслуживания (СМО) (рис. 1). К СМО относятся ремонтные мастерские, станции технического обслуживания, автозаправочные станции и т.д.

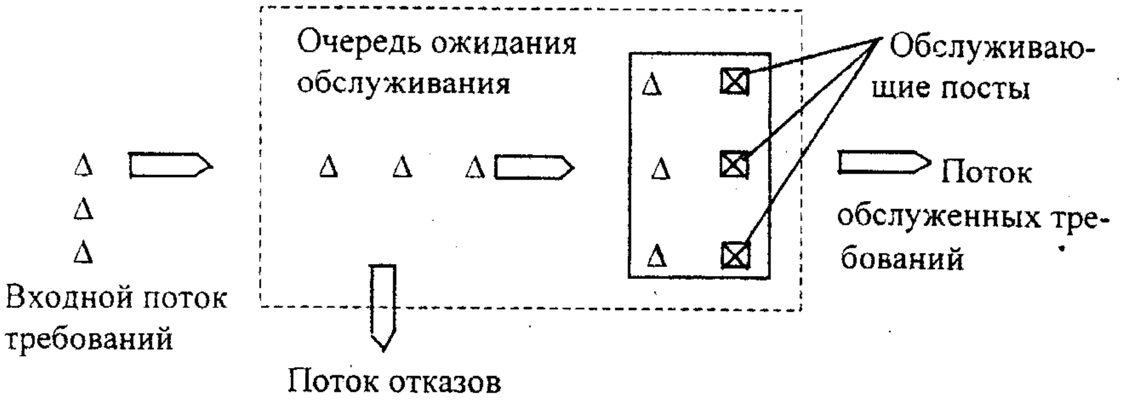


Рис. 1

Случайный характер потока заявок приводит к тому, что в СМО происходит какой-то случайный процесс. Если случайный процесс Марковский, то функционирование СМО можно описать системой дифференциальных уравнений, а в предельном случае – системой линейных алгебраических уравнений, решением которых определяются характеристики работы СМО.

Методы теории массового обслуживания позволяют решать следующие задачи автомобильного транспорта:

- определить количество линий или постов ТО и Р автомобилей;
- определить рациональное количество оборотных агрегатов;
- производить расчет количества постов погрузки и разгрузки автомобилей, а также многие другие задачи.

При анализе работы СМО необходимо знать её основные исходные параметры:

- интенсивность потока заявок - λ;
- трудоёмкость обслуживания одной заявки - Тр;
- число каналов обслуживания - n;
- число мест ожидания - m;
- количество операторов на каждом канале - d;
- условия, накладываемые на образование очереди.

При описании режима работы СМО используются промежуточные параметры:

- время обслуживания одной заявки - t_0 = Тр/d;

- производительность каждого канала обслуживания - $\mu = 1/t_0$; (среднее число заявок, обслуживаемое каналом в единицу времени);

- приведенная интенсивность обслуживания $\alpha = \lambda/\mu$;

- коэффициент загрузки $\beta = \alpha/n$.

Классификация и показатели работы систем массового обслуживания

Системы массового обслуживания подразделяются по ряду признаков:

1. По времени ожидания (T_x):

- без ожидания или с потерями требований ($T_x = 0$);
- с ожиданием или без потерь ($T_x = 0 - \infty$);
- с ограниченным временем ожидания ($T_x = 0 – \tau$).

Системы с неограниченным ожиданием начала технического воздействия является наиболее реальным. Это обусловлено те, что в условиях Российской Федерации перегоны (нулевые пробеги) до поста обслуживания весьма велики. Существует дефицит запчастей и материалов, отсутствует оперативная связь между региональными СТОА, не развиты фирменные системы сервиса и т.п. Поэтому клиент вынужден ждать начала обслуживания иногда и вне зоны ожидания.

2. По числу требований в сутки ($N_с$):

- с ограниченным входящим потоком, или замкнутые ($N_с = 0…k$);
- с неограниченным потоком, или открытые ($N_с = 0…\infty$).

Число требований на СТОА не может быть заранее ограничено, системы являются открытыми (разомкнутыми).

3. По количеству каналов (обслуживающих аппаратов):

- с ограниченным числом обслуживающих аппаратов ($x = 0…n$);
- с неограниченным числом обслуживающих аппаратов ($x = 0…\infty$).

4. По числу фаз (r):

- однофазные ($r = 1$);
- многофазные ($r > 1$).

Многофазные СМО используются при поточном виде обслуживания (например, диагностика - ТО).

5. По уровню организации:

- упорядоченные (энтропия Э = 0);
- неупорядоченные (Э > 0).

Порядок обслуживания устанавливается в результате анализа СМО.

При дальнейшем рассмотрении вопросов функционирования СМО мы остановимся на двух основных видах:

- СМО с отказами;
- СМО с ожиданием обслуживания.

В системах с отказами обслуживаются только те требования, которые поступают в момент времени, когда хотя бы один из каналов обслуживания был свободен. Если все каналы заняты, то заявка покидает СМО необслуженной.

В СМО с ожиданием заявка, поступившая в момент, когда все каналы заняты, ждет освобождения канала обслуживания, то есть встает в очередь.

Различают одноканальные и многоканальные СМО с ожиданием, при этом на длину очереди могут быть наложены ограничения, определяющие максимальную длину очереди.

Все показатели (характеристики) функционирования СМО можно разделить на четыре группы:

- вероятностные;
- количественные;
- временные;
- качественные.

Вероятностные показатели

Будем говорить, что требование находится в СМО, если оно ожидает обслуживания или находится на обслуживании. Обозначим J число требований, находящихся в СМО. Так как с течением времени это число меняется, то J является случайной величиной. Предположим, что произведено очень большое количество наблюдений над СМО, в результате которых установлена доля случаев, когда в системе наблюдалось ровно J требований. Эта величина называется вероятностью p_J того, что в СМО имеется J требований.

Все вероятности состояний должны удовлетворять условию нормировки, согласно которому сумма всех вероятностей должна равняться единице, т. е.

$$\sum_{j=0}^{m+s} p_j = 1.$$

Требование получает отказ в том и только в том случае, если в момент своего поступления в СМО застает занятыми как все каналы, так и все места накопления. Если в системе имеется n аппаратов и m мест накопителя, то вероятность данного события равна

$$P_{\text{отк}} = P_{m+n}.$$

Эта вероятность называется вероятностью отказа.

Вероятность нахождения в СМО не более i требований

$$P_{<i} = \sum_{j=0}^{i} p_j .$$

Вероятность отсутствия очереди

$$P_{\text{от.оч}} = P_{<n} = \sum_{j=0}^{n} p_j ;$$

Вероятность того, что требованию не придется ожидать начала обслуживания

$$P_{(\text{Ам} < n)} = 1 - P_{(\text{Ам} - n)} = P_{\text{от.оч}} - P_n;$$

Вероятность наличия очереди

$$P_{\text{н.оч}} = \sum_{j=m+1}^{m+m} p_j \ ;$$

Вероятность того, что все обслуживающие аппараты заняты

$$P_{(\text{Ам} + n)} = \sum p_j = P_{\text{н.оч}} + P_n;$$

Вероятность обслуживания

$$P_{\text{обс}} = 1 - P_{\text{отк}} \qquad .$$

Количественные показатели

Среднее число свободных обслуживающих аппаратов

$$A_c = \sum_{i=0}^{n} (n-i)p_i .$$

Среднее число требований занятых обслуживающих аппаратов

$$A_м = \sum_{i=0}^{n} ip_i + n\sum_{i=n+1}^{m+n} p_i = n - A_c .$$

Среднее число требований в накопителе

$$A_{\text{н}} = \sum_{i=s+1}^{m+n} (i-n)p_i .$$

Среднее число требований в обслуживающей системе

$$A = \sum_{i=0}^{n+m} ip_i = A_{\text{н}} + A_{\text{м}} .$$

Среднее число требований, получающих отказ за единицу времени

$$A_{\text{отк}} = \lambda P_{\text{отк}}.$$

Временные показатели

Важным показателем является среднее время ожидания начала обслуживания $T_{\text{ож}}$. Сумма среднего времени ожидания $T_{\text{ож}}$ и обслуживания $T_{\text{обсл}}$ равна среднему времени пребывания требования в системе $T_{\text{сист}}$

$$T_{\text{сист}} = T_{\text{ож}} + T_{\text{обсл}}.$$

Установим полезную связь между средним числом требований, находящихся в системе, и средним временем пребывания требования в системе. Так как за единицу времени в систему поступает λ требований, а средняя продолжительность пребывания одного требования в системе есть $T_{сист}$, то суммарная продолжительность нахождения всех требований в системе за единицу времени равна ($\lambda T_{сист}$). Но так как в системе находится в среднем A требований, то эта величина равна произведению единицы времени на A, т. е.

$$A = \lambda T_{сист}.$$

Аналогично можно вывести формулу

$$A_{н} = \lambda T_{ож}.$$

Качественные показатели

Приведенные показатели характеризуют степень использования обслуживающих аппаратов и затраты времени на пребывание требований в системе (в очереди и на обслуживание). Обычно бывает необходимым соизмерить эти величины, чтобы найти экономически оптимальное решение. В связи с этим предположим, что за единицу времени пребывания требования в системе обслуживания имеют место убытки, равные C. Кроме того, эксплуатация одного аппарата приводит за единицу времени к расходам, равным K. Отметим, что в величину K обычно входят как эксплуатационные расходы, так и удельные капитальные затраты на один аппарат, связанные с его приобретением и приходящиеся на единицу времени. Пусть далее отказ в обслуживании одного требования влечет убыток, равный $C_{о}$.

Тогда приведенные средние затраты, связанные с эксплуатацией n аппаратов, поступлением в систему в единицу времени λ требований и убытками от отказов требованиям, в единицу времени в в среднем составляют

$$Э = C_{о}\lambda P_{отк} + CT_{сист}\lambda + Kn.$$

В этом выражении $T_{сист}$ (среднее время пребывания одного требования в системе), а вместе с ним и слагаемое ($CT_{сист}\lambda$) уменьшаются с ростом числа обслуживающих аппаратов n. В противоположность этому слагаемое Kn с ростом числа аппаратов увеличивается. Задача состоит, следовательно, в выборе такого значения n, при котором критерий (Э) будет минимален. Такой выбор n обеспечивает минимизацию приведенных затрат, связанных с приобретением и эксплуатацией аппаратов и непроизводительным пребыванием требований в системе обслуживания.

Отметим также, что за счет дополнительных затрат часто имеется возможность уменьшить среднюю длительность обслуживания одного требования $T_{обс}$. При этом будет уменьшаться и величина $T_{сист}$ - среднее время пребывания требования в системе. В этом случае для нахождения оптимального решения в критерий Э следует ввести также указанные дополнительные затраты.

Следующим показателем качества функционирования СМО являются:

- коэффициент загрузки поста

$$K_з = \lambda\mu/n.$$

- коэффициент использования аппаратов

$$K_{исп} = A_м/n.$$

Лекция № 14. Характеристики систем массового обслуживания

1. Системы массового обслуживания с отказами

Пусть имеем n-канальную СМО с отказами, в которую поступает поток требований (автомобилей) на обслуживание с интенсивностью λ, интенсивность обслуживания одного канала равна μ. Определим показатели эффективности работы такой системы:

Построим размеченный граф состояний системы (рис. 1), где состояния системы пронумерованы по числу занятых каналов, т. е.

$S0$ - все каналы свободны;

$S1$ - занят один канал, остальные свободны;

$S2$ - заняты два канала, остальные свободны;

…………………………………………….

Sn - заняты все n каналов.

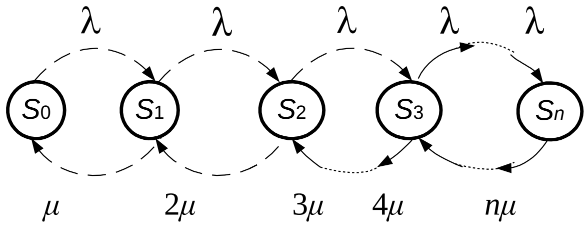


Рис. 1. Размеченный граф состояний

По стрелке слева направо систему переводит поток заявок с интенсивностью λ. По стрелкам справа налево систему переводит поток обслуживаний интенсивностью $k\mu$, где k - число занятых каналов.

Для вычислений предельных вероятностей состояний системы (P_i) запишем систему линейных уравнений. Согласно рис. 1 будем иметь:

Таблица 1

k	Цепочка уравнений	Вероятность состояния системы
1	$\lambda P_0 = \mu P_1$	$P_1 = \frac{\alpha}{1!} P_0$
2	$\lambda P_1 = 2\mu P_2$	$P_2 = \frac{\alpha}{2} P_1 = \frac{\alpha^2}{2!} P_0$
3	$\lambda P_2 = 3\ \mu\ P_3$	$P_3 = \frac{\alpha}{3} P_2 = \frac{\alpha^3}{3!} P_0$
4	$\lambda\ P_3 = 4\mu P_4$	$P_4 = \frac{\alpha}{4} P_4 = \frac{\alpha^4}{4!} P_0$
….	……………………	…………………………………
k	$\lambda P_{k-1} = k\mu P_k$	$P_k = \frac{\alpha}{k} P_{k-1} = \frac{\alpha^k}{\kappa!} P_0$

…	…………………….	………………………………….
n	$\lambda P_{n-1} = n\mu P_n$	$P_n = \frac{\alpha}{n} P_{n-1} = \frac{\alpha^n}{n!} P_0$

Рекуррентное выражение для определения вероятности состояния будет равно

$$P_k / P_{k-1} = \frac{\alpha}{k},$$

где $\alpha = \lambda/\mu$.

Запишем нормировочное условие:

$$P_0 + P_1 + P_2 + ... + P_n = 1.$$

Решая его совместно с системой уравнений табл. 1, получим:

$$P_0 = \frac{1}{1 + \frac{\alpha}{1!} + \frac{\alpha^2}{2!} + \frac{\alpha^3}{3!} + ... + \frac{\alpha^n}{n!}}.$$

Остальные вероятности состояний найдём из выражений табл. 1.

Зная предельные вероятности состояний P_0, P_1, P_2, …, P_n, вычислим характеристики работы СМО.

Вероятностные характеристики:

- вероятность того, что все каналы заняты и отказ в обслуживании

$$P_{\text{отк}} = P_n;$$

- относительная пропускная способность

$$q = 1 - P_n.$$

Количественные характеристики:

- среднее число занятых постов

$$M_n = 0P_0 + 1P_1 + 2P_2 + \ldots + nP_n = \alpha(1 - P_{\text{отк}});$$

- среднее число каналов, свободных от обслуживания

$$N_0 = n - M_n;$$

- абсолютная пропускная способность

$$A = \lambda(1 - P_n).$$

Временная характеристика:

- среднее время обслуживания

$$t_{\text{обс}} = 1/\mu.$$

Качественные характеристики:

- коэффициент занятости каналов

$$K_{\text{з}} = M_n/n;$$

- коэффициент простоя каналов

$$K_{пр} = 1 - K_з.$$

2. Системы массового обслуживания с ожиданием

Одноканальная СМО с ожиданием

Пусть СМО имеет один канал, на которую поступает поток заявок с интенсивностью λ. Интенсивность обслуживания канала равна μ. Заявка, поступившая в момент, когда канал занят, становится в очередь и ожидает обслуживания. Предположим, что число мест в очереди равно m, то есть, если заявка, пришедшая в момент, когда в очереди стоит m заявок, покидает СМО необслуженной.

Будем нумеровать состояния СМО по числу заявок, находящихся в системе (как обслуживаемых, так и ожидающих обслуживания):

S_0 - канал свободен;

S_1 - канал занят, очереди нет;

S_2 - канал занят, одна заявка в очереди;

………………………………………

S_{m+1} - канал занят, m заявок в очереди.

Составим размеченный граф состояний системы (рис. 2.)

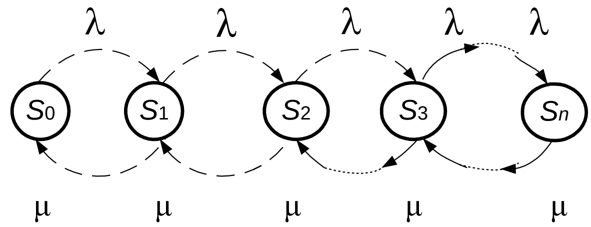


Рис. 2. Размеченный граф состояния

Система дифференциальных уравнений Колмогорова для данного процесса имеет вид:

$$\left.\begin{aligned}
&\frac{dP_0(t)}{dt} = -\lambda_0 P_0(t) + \mu P_1(t);\\
&\frac{dP_1(t)}{dt} = -(\lambda + \mu)P_1(t) + \lambda P_0(t) + \mu P_2(t);\\
&\frac{dP_2(t)}{dt} = -(\lambda + \mu)P_2(t) + \lambda P_1(t) + \mu P_3(t);\\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\\
&\frac{\mathrm{dP}_{\mathrm{m+1}}}{\mathrm{dt}} = -\mu P_{m+1}(t) + \lambda P_m(t).
\end{aligned}\right\} \qquad (1)$$

Добавим к этой системе начальные условия. Например, если при $t = 0$ система находится в состоянии S_0, то начальные условия примут вид:

$$P_0(0) = 1,\ P_1(0) = P_2(0) = \ldots = P_{m+1}(0) = 0.$$

Проинтегрировав систему (1) при принятых начальных условиях, получим все вероятности состояния как функции времени

$$P_0(t);\ P_1(t);\ P_2(t);\ldots\ P_{m+1}(t),$$

которые в любой момент времени t удовлетворяют условию

$$\sum_{i=0}^{m+1} P_i(t) = 1 \qquad (2)$$

Если в системе дифференциальных уравнений (1) положить все производные равные нулю (при $t \to \infty$ зависимость от времени пропадает), то она превращается в систему обычных линейных алгебраических уравнений, которая совместно с нормированным условием (2) дает возможность вычислить все предельные вероятности состояний.

При $\dfrac{dP_i}{dt} = 0$ имеем

$$\left.\begin{aligned} &\lambda P_0 = \mu P_1, \\ &\lambda P_1 + \mu P_1 = \lambda P_0 + \mu P_2, \\ &\lambda P_2 + \mu P_2 = \lambda P_1 + \lambda P_3, \\ &\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ &\mu P_{m+1} = \lambda P_m. \end{aligned}\right\} \qquad (3)$$

После преобразований система (3) перепишется в виде цепочки уравнений (табл. 2).

Таблица 2

1	$P_1 = \alpha P_0$
2	$P_2 = \alpha P_1 = \alpha^2 P_0$
3	$P_3 = \alpha P_2 = \alpha^3 P_0$
…	……………………
$m+1$	$P_{m+1} = \alpha P_m = \alpha^{m+1} P_0$

Окончательное выражение для P_0 имеет вид

$$P_0 = \frac{1}{1 + \alpha + \alpha^2 + \ldots. + \alpha^{m+1}}.$$

Остальные вероятности состояний можно найти по формулам табл. 2.

Многоканальная СМО с ожиданиями

Пусть имеем n-канальную СМО с ожиданием, на которую поступает поток заявок с интенсивностью λ, интенсивность обслуживания одного канала равна μ, число мест в очереди ограничено заданным числом m. Вычислить основные характеристики СМО.

Состояния системы будем нумеровать по числу заявок, связанных с системой:

$S0$ - все каналы свободны;

$S1$ - занят только один канал;

$S2$ - заняты только два канала;

S_n - заняты все n каналов.

Когда СМО находится в любом из этих состояний, очереди еще нет. После того, как будут заняты все каналы обслуживания, а заявки продолжают поступать, образуется очередь. Тогда состояния системы будут:

S_{n+1} - заняты все n каналов и одна заявка в очереди;

S_{n+2} - заняты все n каналов и две заявки в очереди;

…………………………………………………

S_{n+m} - заняты все n каналов и все m мест в очереди.

Граф состояний системы представлен на рис. 3.

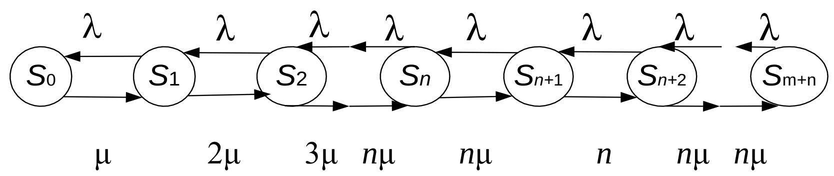


Рис. 3. Размеченный граф состояния

Действительно, переход системы в состояние с большими номерами (слева направо) вызывается только потоком заявок с интенсивностью λ. По стрелкам справа налево систему переводит поток обслуживания, интенсивность которого равна μ, умноженная на число занятых каналов. Действительно, полная интенсивность потока обслуживания возрастает с подключением новых каналов вплоть до такого состояния S_μ, когда все n каналов скажутся занятыми. С появлением очереди интенсивность обслуживания больше не увеличивается, так как она уже достигла максимума, равного ($n\mu$).

Не повторяя соответствующих рассуждений, запишем сразу в окончательном виде основные формулы, отражающие работу СМО с ожиданием, введя для упрощения записи обозначение

$\lambda/\mu = \alpha$ - приведенная интенсивность; $\alpha/n = \beta$ - нагрузка.

Таблица 3

До возникновения очереди

k	Цепочка уравнений	Вероятность состояния системы
1	$\lambda P_0 = \mu P_1$	$P_1 = \frac{\alpha}{1!} P_0$
2	$\lambda P_1 = 2\mu P_2$	$P_2 = \frac{\alpha}{2} P_1 = \frac{\alpha^2}{2!} P_0$
3	$\lambda P_2 = 3\mu P_3$	$P_3 = \frac{\alpha}{3} P_2 = \frac{\alpha^3}{3!} P_0$
4	$\lambda P_3 = 4\mu P_4$	$P_4 = \frac{\alpha}{4} P_4 = \frac{\alpha^4}{4!} P_0$
…	……………	…………………………………
k	$\lambda P_{k-1} = k\mu P_k$	$P_k = \frac{\alpha}{k} P_{k-1} = \frac{\alpha^k}{k!} P_0$
…	……………………..	………………………………..
n	$\lambda P_{n-1} = n\mu P_n$	$P_n = \frac{\alpha}{n} P_{n-1} = \frac{\alpha^n}{n!} P_0$

После возникновения очереди

k	Цепочка уравнений	Вероятность состояния системы
1	$\lambda P_n = n\mu P_{n+1}$	$P_{n+1} = \beta P_n = \beta^1 P_n = \beta \frac{\alpha^n}{n!} P_0$
2	$\lambda P_{n+1} = n\mu P_{n+2}$	$P_{n+2} = \beta P_{n+1} = \beta^2 P_n = \beta^2 \frac{\alpha^n}{n!} P_0$
3	$\lambda P_{n+2} = n\mu P_{n+3}$	$P_{n+3} = \beta P_{n+2} = \beta^3 P_n = \beta^3 \frac{\alpha^n}{n!} P_0$
4	$\lambda P_{n+3} = n\mu P_{n+4}$	$P_{n+4} = \beta P_{n+3} = \beta^4 P_n = \beta^4 \frac{\alpha^n}{n!} P_0$
…	………………	………………………………..
l	$\lambda P_{n+l-1} = n\mu P_{n+l}$	$P_{n+l} = \beta P_{n+l-1} = \beta^l P_n = \beta^l \frac{\alpha^n}{n!} P_0$
…	……………….....	………………………………..
m	$\lambda P_{n+m-1} = n\mu P_{n+m}$	$P_{n+m} = \beta P_{n+m-1} = \beta^m P_n = \beta^m \frac{\alpha^n}{n!} P_0$

Вычислим основные характеристики СМО, для чего запишем нормировочное условие

$$P_0 + P_1 + P_2 + ... + P_{n+m} = 1$$

и подставим в него значения вероятностей P_i выраженные через вероятность P_0, тогда вероятность того, что все каналы свободны - определяется по выражению

$$P_0 = \frac{1}{\sum_{k=0}^{n} \frac{\alpha^k}{k!} + \frac{\alpha^n}{n!} \sum_{l=1}^{m} \beta^l} .$$

Другие вероятности состояний определяются по данным табл. 3.

Вероятность того, что все каналы и места ожидания заняты и заявка получает отказ в обслуживании

$$P_{\text{отк}} = P_{n+m} = \beta^m \frac{\alpha^n}{n!} P_0.$$

Относительная пропускная способность

$$q = 1 - P_{\text{отк}}.$$

Среднее число каналов, занятых обслуживанием

$$M_{з} = \sum_{k=1}^{n} kP_k + n\sum_{k=1}^{m} P_{n+k} = aq.$$

Среднее число каналов, свободных от обслуживания

$$N_0 = \sum_{k=0}^{n} (n-k)P_k = n - M_{з}.$$

Среднее число заявок в накопителе

$$A_{н} = \sum_{k=n+1}^{n+m} (k-n)P_k.$$

Абсолютная пропускная способность

$$A = \lambda q.$$

Среднее время ожидания обслуживания заявки в очереди

$$\overline{t_o} = \frac{A_{н}}{\lambda}.$$

Среднее время пребывания заявки в системе

$$\overline{t_c} = \overline{t_o} + \frac{1}{\mu}.$$

Коэффициент занятости каналов

$$K_{з} = M_{з}/n.$$

Коэффициент простоя каналов

$$K_{пр} = 1 - K_{з}.$$