

О г л а в л е н и е

| | |
|---|-----------|
| Введение | 5 |
| Глава 1. Понятие модели и моделирования..... | 8 |
| 1.1. Общее определение модели | 8 |
| 1.2. Классификация моделей и моделирования | 13 |
| 1.2.1. Классификация моделей и моделирования по признаку «характер моделируемой стороны объекта»..... | 14 |
| 1.2.2. Классификация моделей и моделирования по признаку «характер процессов, протекающих в объекте» | 14 |
| 1.2.3. Классификация моделей и моделирования по признаку «способ реализации модели» | 16 |
| 1.3. Этапы моделирования | 19 |
| 1.4. Адекватность модели | 23 |
| 1.5. Требования, предъявляемые к моделям..... | 26 |
| 1.6. Компьютерное моделирование..... | 27 |
| Вопросы для самоконтроля..... | 29 |
| Глава 2. Концепция дискретных систем для имитационного моделирования..... | 32 |
| 1. Основные понятия | 32 |
| 2. Классификация моделей массового обслуживания | 44 |
| 3. Параметры и характеристики систем массового обслуживания | 50 |
| Вопросы для самоконтроля..... | 55 |
| Глава 3. Имитационное статистическое моделирование | 56 |
| 3.1. Сущность имитационного моделирования | 57 |
| 3.2. Общая характеристика метода имитационного моделирования | 64 |
| 3.3. Статистическое моделирование при решении детерминированных задач..... | 68 |
| 3.4. Моделирование равномерно распределенной случайной величины..... | 71 |
| 3.5. Моделирование случайной величины с произвольным законом распределения | 79 |
| 3.6. Моделирование единичного события | 86 |
| 3.7. Моделирование полной группы несовместных событий | 87 |
| 3.8. Моделирование совместных независимых событий..... | 94 |
| 3.8.1. Определение совместных исходов по жребию | 94 |
| 3.8.2. Последовательная проверка исходов..... | 95 |
| 3.9. Моделирование совместных зависимых событий | 96 |
| 3.10. Классификация случайных процессов | 99 |
| 3.11. Способы продвижения модельного времени..... | 103 |
| 3.12. Модель противоборства двух сторон..... | 112 |
| 3.13. Модель противоборства как процесс блуждания по решётке..... | 120 |
| 3.14. Типовая схема имитационной модели с продвижением времени по событиям..... | 124 |
| 3.15. Имитационная модель системы массового обслуживания | 128 |
| Вопросы для самоконтроля..... | 138 |

| | |
|--|------------|
| Глава 4. Планирование экспериментов..... | 140 |
| 4.1. Сущность и цели планирования эксперимента | 140 |
| 4.2. Элементы стратегического планирования экспериментов..... | 142 |
| 4.3. Стандартные планы | 146 |
| 4.4. Формальный подход к сокращению общего числа прогонов | 151 |
| 4.5. Элементы тактического планирования | 152 |
| 4.6. Точность и количество реализаций модели при определении средних значений параметров..... | 154 |
| 4.6.1. Определение оценки математического ожидания | 154 |
| 4.6.2. Определение оценки дисперсии..... | 158 |
| 4.7. Точность и количество реализаций модели при определении вероятностей исходов | 159 |
| 4.8. Точность и количество реализаций модели при зависимом ряде данных | 169 |
| 4.9. Проблема начальных условий | 172 |
| Вопросы для самоконтроля..... | 173 |
| Глава 5. Обработка результатов имитационного эксперимента | 175 |
| 5.1. Характеристики случайных величин и процессов..... | 175 |
| 5.2. Требования к оценкам характеристик..... | 177 |
| 5.3. Оценка характеристик случайных величин и процессов | 178 |
| 5.4. Гистограмма | 182 |
| 5.5. Элементы дисперсионного анализа. Критерий Фишера | 184 |
| 5.6. Критерий Вилкоксона | 188 |
| 5.7. Однофакторный дисперсионный анализ | 192 |
| 5.8. Выявление несущественных факторов | 197 |
| 5.9. Сущность корреляционного анализа..... | 201 |
| 5.10. Обработка результатов эксперимента на основе регрессии..... | 208 |
| Вопросы для самоконтроля..... | 224 |
| Глава 6. Современные теории имитационного моделирования | 226 |
| 6.1. Распределённое имитационное моделирование | 226 |
| 6.1.1. Понятие о распределённой информационной системе | 226 |
| 6.1.2. Необходимость перехода к параллельному и распределённому имитационному моделированию | 229 |
| 6.1.3. Подходы к сокращению времени имитационного эксперимента .. | 233 |
| 6.1.4. Вариант построения распределённой имитационной модели | 234 |
| 6.2. Агентное моделирование | 238 |
| 6.2.1. Мультиагентные системы и агенты | 238 |
| 6.2.2. Подходы к построению мультиагентной модели | 241 |
| 6.2.3. Общая схема системы агентного моделирования..... | 249 |
| Вопросы для самоконтроля..... | 253 |
| Заключение | 254 |
| Список литературы..... | 259 |
| Приложение 1 | 262 |
| Приложение 2 | 264 |
| Глоссарий | 265 |

Введение

Данное учебное пособие «Моделирование и проектирование систем. Часть I. Курс лекций» рекомендуется курсантам при изучении курса «Моделирование и проектирование систем». Пособие может быть также полезно адъюнктам при изучении дисциплины «Моделирование». Оно окажет также существенную помощь при курсовом и дипломном проектировании, поскольку разработка имитационной модели системы или процесса может составлять предмет не только специальной части, но и курсовых, и выпускных квалификационных работ в целом. Предполагается, что читатель знает основы теории вероятностей, математической статистики и программирования на ЭВМ, а также соответствующую предметную область.

Имитационное моделирование (ИМ) – один из самых мощных инструментов анализа, которыми располагают специалисты, занимающиеся исследованием, разработкой и функционированием сложных процессов и систем. Идея имитационного моделирования проста и в то же время интуитивно привлекательна. Каждый современный исследователь должен уметь пользоваться этим методом моделирования.

Имитационное моделирование как необходимая часть инженерного образования сложилось в середине двадцатого века. Воспринятое поначалу как своеобразный численный метод решения сложных задач, как «младший брат» аналитического моделирования, оно постепенно стало основным, подчас единственным методом при анализе и синтезе сложных систем и процессов.

Общеизвестно, что правильно поставленный натурный эксперимент, то есть исследование свойств объекта на самом объекте, максимально информативен. Оказывается, что эксперимент с компьютерной имитационной моделью вполне конкурентоспособен с натурным. Не говоря о том, что натурный эксперимент в ряде случаев вообще невозможен или нецелесообразен, эксперимент с имитационной моделью может быть приемлемо информативным, и выполнен значительно быстрее и дешевле натурального. Это и предопределило стремительное и повсеместное внедрение моделирования.

В популяризации имитационного моделирования заметную роль сыграли работы Р. Шеннона [51] и Т. Д. Шрайбера [52]. В свое время эти работы были широко известны в среде научных работников и инженеров. Большую положительную роль в распространении компьютерного имитационного моделирования у нас в стране сыграли работы

по моделированию сложных систем на ЭВМ члена-корреспондента АН СССР Н. П. Бусленко (1922-1977) и выдающегося математика академика АН СССР А. А. Самарского (1919-2008). Их работы в области математического моделирования и вычислительного эксперимента широко используются на практике.

Постепенно популярность методов и идей ИМ возрастала во всем мире, что вызывало потребность в общении пользователей, обмене знаниями и опытом, теоретическом осмыслении этого нового метода познания. В результате во многих странах мира были созданы национальные и даже наднациональные (международные) общества имитационного моделирования. Таких обществ сейчас в мире много [33].

Несмотря на то, что ИМ известно и применяется в нашей стране давно, профессиональных сообществ в этой области, как в других странах, до сих пор не создавалось. ***В феврале 2011 года зарегистрировано Некоммерческое партнерство «Национальное общество имитационного моделирования» (НП НОИМ), Санкт-Петербург.*** Можно ожидать, что в ближайшее время ИМ России будет развиваться еще более интенсивно, и мы действительно войдем в международный мир ИМ полноправными участниками. Именно такие цели и задачи ставит перед собой Национальное общество ИМ НОИМ.

НОИМ необходимо много сделать в деле популяризации и продвижения идей ИМ в различных областях науки, технологий и промышленности, государственных и отраслевых органах управления. В Интернете создан сайт НОИМ www.simulation.su, на котором размещена различная информация по имитационному моделированию.

Курс имитационного моделирования является обязательным для изучения в учебных планах технических высших и средних учебных заведений, в том числе и военных.

Первая глава носит вводный характер. Разъясняются понятия моделирования и основных терминов: модель, моделирование, классификация, этапы, адекватность, компьютерное моделирование. Сформулированы и пояснены требования к моделям.

Во второй главе излагается концепция дискретных систем – систем массового обслуживания, реализованная во многих системах имитационного моделирования, в том числе в GPSS World и AnyLogic, приемы построения моделей инструментальными средствами которых и проведение исследований будут рассматриваться на практических занятиях. Многие задачи, решаемые при совершенствовании и разра-

ботке реальных военно-технических систем, описываются именно моделями систем массового обслуживания.

В третьей главе излагается сущность имитационного статистического моделирования, рассмотрены обоснования и методы имитационного моделирования случайных дискретных событий – единичного, совместных и несовместных, независимых и зависимых, и случайных процессов. Представлены с подробным описанием примеры, демонстрирующие способы построения алгоритмов имитационных статистических моделей.

В четвёртой главе излагаются основы планирования компьютерного эксперимента: цели, задачи, стратегическое и тактическое планирование, типовые планы. Подчёркивается необходимость знания этих основ, несмотря на наличие встроенных типовых планов в математические программные пакеты, в том числе имитационного моделирования.

Пятая глава посвящена обзору некоторых наиболее употребительных приёмов статистической обработки результатов компьютерного эксперимента: дисперсионный анализ, корреляционный анализ, регрессионный анализ. Обзор подкреплён примерами.

В шестой главе даётся краткий обзор современных теорий имитационного моделирования, в том числе распределённого имитационного и имитационного мультиагентного моделирования.

Каждая глава завершается вопросами для самоконтроля.

В заключении излагаются основные принципы моделирования как систематизированная последовательность этапов, методов и средств по организации процесса моделирования.

В пособии для лучшего усвоения материала фрагменты, представляющие наибольший интерес, выделяются разными шрифтами, что позволяет акцентировать внимание читателя на аспектах, которые, по мнению автора, являются важными для понимания моделей и методов.

При работе над пособием авторы опирались на свой опыт моделирования и преподавания, а также на многие издания по теме моделирования. В большей степени были учтены работы, указанные в списке литературы.

В учебном пособии «Моделирование и проектирование систем. Часть II. Пособие для практических занятий» изложены методики разработки моделей в системе моделирования AnyLogic.

Авторы благодарны Д. В. Боеву, помощь которого в подготовке и оформлении электронной версии рукописи была существенна.

Глава 1. Понятие модели и моделирования

Первая тема имеет вводный, в основном, терминологический характер. Подробно раскрываются понятия модели и моделирования, их назначение как основного, а подчас, и единственного метода анализа и синтеза сложных систем и процессов. Даётся вариант классификации моделей и моделирования, в некоторой мере упрощенный, но достаточный для полного уяснения сущности моделирования вообще, и математического в частности. Сам по себе процесс моделирования в полной мере не формализован, большая роль в этом принадлежит опыту разработчика. Но, тем не менее, рассматриваемый в теме процесс создания модели может стать основой для начинающих и по мере накопления опыта может быть индивидуализирован.

Математическая модель, являясь абстрактным образом моделируемого объекта или процесса, не может быть его полным аналогом. Достаточно сходства в тех элементах, которые определяют цель исследования. Для качественной оценки сходства вводится понятие адекватности модели объекту и, в связи с этим, раскрываются понятия изоморфизма и изофункционализма. Формальных приёмов, позволяющих автоматически, «бездумно», создавать адекватные математические модели, нет. Окончательное суждение об адекватности модели даёт практика, то есть сопоставление модели с действующим объектом. Тем не менее, усвоение всех тем пособия позволит исследователю справляться с проблемой обеспечения адекватности моделей.

Завершается тема изложением требований к моделям, которые были сформулированы Р. Шенноном на заре компьютерного моделирования более тридцати лет назад в книге «Имитационное моделирование систем – искусство и наука». Актуальность этих требований сохраняется и в настоящее время.

1.1. Общее определение модели

Практика свидетельствует: самое лучшее средство для определения свойств объекта – *натурный эксперимент*, т. е. исследование свойств и поведения самого объекта в нужных условиях. Дело в том, что при проектировании невозможно учесть многие факторы, расчёт ведётся по усредненным справочным данным, используются новые, недостаточно проверенные элементы (прогресс нетерпелив!), меняются условия внешней среды и многое другое. Поэтому натурный эксперимент –

необходимое звено исследования. Неточность расчётов компенсируется увеличением объёма натуральных экспериментов, созданием ряда опытных образцов и «доводкой» изделия до нужного состояния. Так поступали и поступают при создании, например, любого нового образца техники и вооружения.

Однако во многих случаях натуральный эксперимент невозможен.

Моделирование предоставляет возможность исследования таких объектов, прямой эксперимент с которыми:

- опасен;
- экономически невыгоден;
- долговременен;
- кратковременен;
- протяжён в пространстве;
- невозможен;
- неповторим;
- ненагляден.

Остановимся на трёх последних причинах необходимости моделирования.

Невозможен. Часто человек имеет дело с ситуацией, когда объекта нет, он ещё только проектируется. При проектировании важно не только представить себе будущий объект, но и испытать его виртуальный аналог до того, как дефекты проектирования проявятся в оригинале.

Важно: моделирование теснейшим образом связано с проектированием. Обычно сначала проектируют систему, потом её испытывают, потом снова корректируют проект и снова испытывают, и так до тех пор, пока проект не станет удовлетворять предъявляемым к нему требованиям. Процесс «проектирование-моделирование» цикличен. При этом цикл имеет вид спирали – с каждым повтором проект становится все лучше, так как модель становится все более детальной, а уровень описания точнее.

Неповторим. Это достаточно редкий случай, когда эксперимент повторить нельзя; в такой ситуации модель – единственный способ изучения таких явлений. Пример – исторические процессы, – ведь повернуть историю вспять невозможно.

Ненагляден. Модель позволяет заглянуть в детали процесса, его стадии. При построении модели исследователь вынужден описать причинно-следственные связи, позволяющие понять все в единстве, системе. Построение модели дисциплинирует мышление.

Важно: модель играет системообразующую и смыслообразующую роли в научном познании, позволяет *понять* явление, структуру изучаемого объекта. Не построив модель, вряд ли удастся понять логику действия системы. Это означает, что модель позволяет разложить систему на элементы, связи, механизмы, требует объяснить действие системы, определить причины явлений, характер взаимодействия составляющих.

Время подготовки натурального эксперимента и проведение мероприятий по обеспечению безопасности часто значительно превосходят время самого эксперимента. Многие испытания, близкие к граничным условиям, могут протекать настолько бурно, что возможны аварии и разрушения части или всего объекта.

Из сказанного следует, что натуральный эксперимент необходим, но в то же время невозможен либо нецелесообразен.

Выход из этого противоречия, как видим, есть и называется он «моделирование».

Моделирование – это замещение одного *исходного объекта-оригинала* другим объектом с целью получения информации о свойствах объекта-оригинала.

Отсюда следует.

Моделирование – это, во-первых, процесс создания или отыскания в природе объекта, который в интересующем исследователя смысле может заменить исследуемый объект. Этот промежуточный объект называется **моделью** (рис. 1.1). Модель может быть материальным объектом той же или иной природы по отношению к изучаемому объекту (оригиналу). Модель может быть мысленным объектом, воспроизводящим оригинал логическими построениями или математическими формулами и компьютерными программами.

Моделирование, во-вторых, это испытание, исследование модели. То есть, моделирование связано с экспериментом, отличающимся от натурального тем, что в процесс познания включается «промежуточное звено» – модель. Следовательно, **модель** является одновременно *средством эксперимента* и *объектом эксперимента*, заменяющим изучаемый объект.

Моделирование, в-третьих, это перенос полученных на модели сведений на оригинал или, иначе, приписывание свойств модели оригиналу. Чтобы такой перенос был оправдан, между моделью и оригиналом должно быть сходство, *подобие*.

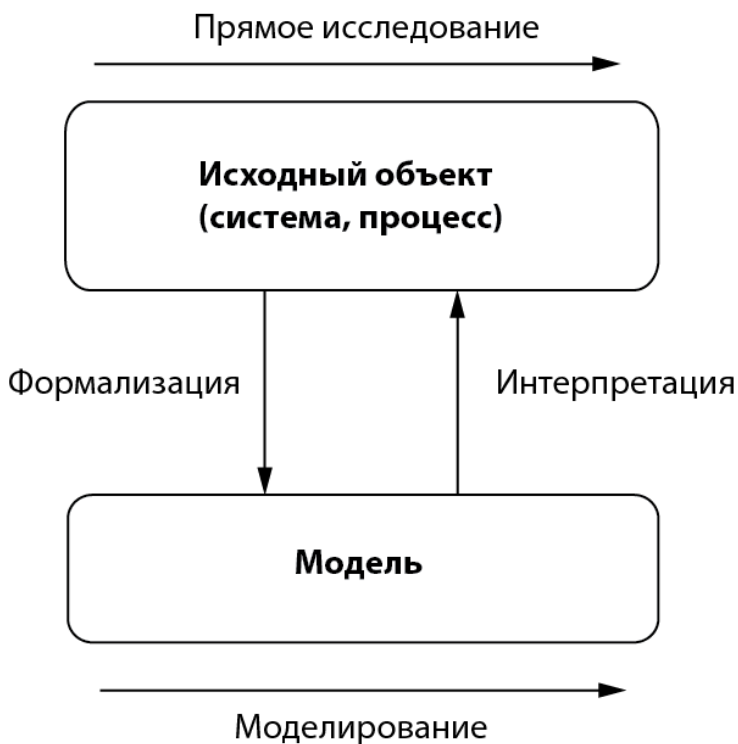


Рис. 1.1. К понятию «моделирование»

Подобие может быть физическим, геометрическим, структурным, функциональным и т. д. **Степень подобия** может быть разной – от тождества во всех аспектах до сходства только в главном. Очевидно, модели не должны воспроизводить полностью все стороны изучаемых объектов. Достижение абсолютной одинаковости сводит моделирование к натурному эксперименту, о возможности или целесообразности которого было уже сказано.

Остановимся на основных **целях моделирования**.

Прогноз – оценка поведения системы при некотором сочетании её управляемых и неуправляемых параметров. Прогноз – главная цель моделирования.

Объяснение и лучшее понимание объектов. Здесь чаще других встречаются задачи оптимизации и анализа чувствительности. *Оптимизация* – это точное определение такого сочетания факторов и их величин, при котором обеспечиваются наилучший показатель каче-

ства системы, наилучшее по какому-либо критерию достижение цели моделируемой системой. *Анализ чувствительности* – выявление из большого числа факторов тех, которые в наибольшей степени влияют на функционирование моделируемой системы. Исходными данными при этом являются результаты экспериментов с моделью.

Часто модель создается для применения в качестве *средства обучения*: модели-тренажеры, стенды, учения, деловые игры и т. п.

Моделирование как метод познания применялось человечеством – осознанно или интуитивно – всегда. На стенах древних храмов предков южно-американских индейцев обнаружены графические модели мироздания. Учение о моделировании возникло в средние века. Выдающаяся роль в этом принадлежит Леонардо да Винчи (1452-1519).

Гениальный полководец А. В. Суворов (1730-1800) перед атакой крепости Измаил тренировал солдат на модели измаильской крепостной стены, построенной специально в тылу.

Наш знаменитый механик-самоучка И. П. Кулибин (1735-1818) создал модель одноарочного деревянного моста через р. Неву, а также ряд металлических моделей мостов. Они были полностью технически обоснованы и получили высокую оценку российскими академиками Л. Эйлером (1707-1783) и Д. Бернулли (1700-1782). К сожалению, ни один из этих мостов не был построен.

Огромный вклад в укрепление обороноспособности нашей страны внесли работы по моделированию взрыва – выпускник Михайловского артиллерийского училища 1869 г. генерал-инженер Н. Л. Кирпичев (1850-1927), моделированию в авиастроении – М. В. Келдыш (1911-1978), С. В. Ильющин (1894-1977), А. Н. Туполев (1888-1972) и др., моделированию ядерного взрыва – И. В. Курчатов (1903-1960), А. Д. Сахаров (1921-1989), Ю. Б. Харитон (1904-1996) и др.

Широко известны работы Н. Н. Моисеева (1917-2000) по моделированию систем управления. В частности, для проверки одного нового метода математического моделирования была создана математическая модель Синопского сражения – последнего сражения эпохи парусного флота. В 1833 году адмирал П. С. Нахимов (1802-1855) разгромил главные силы турецкого флота. Моделирование на вычислительной машине показало, что Нахимов действовал практически безошибочно. Он настолько верно расставил свои корабли и нанес первый удар, что единственное спасение турок было отступление. Иного выхода у них не было. Они не отступили и были разгромлены.

Сложность и громоздкость технических объектов, которые могут изучаться методами моделирования, практически неограниченны. В последние годы все крупные сооружения исследовались на моделях – плотины, каналы, Братская и Красноярская ГЭС, системы дальних электропередач, образцы военных систем и др. объекты.

Поучительный пример недооценки моделирования – гибель английского броненосца «Кэптен» в 1870 году. В стремлении еще больше увеличить своё тогдашнее морское могущество и подкрепить империалистические устремления в Англии был разработан суперброненосец «Кэптен». В него было вложено все, что нужно для «верховой власти» на море: тяжелая артиллерия во вращающихся башнях, мощная бортовая броня, усиленное парусное оснащение и очень низкими бортами – для меньшей уязвимости от снарядов противника. Консультант инженер Рид построил математическую модель остойчивости «Кэптена» и показал, что даже при незначительном ветре и волнении ему грозит опрокидывание. Но лорды Адмиралтейства настояли на строительстве корабля. На первом же учении после спуска на воду налетевший шквал перевернул броненосец. Погибли 523 моряка. В Лондоне на стене одного из соборов прикреплена бронзовая плита, напоминающая об этом событии и, добавим мы, о тупоумии самоуверенных лордов Британского Адмиралтейства, пренебрегших результатами моделирования.

Моделирование является технологией решения задач. Это замечание – очень важное. Так как технология есть способ достижения результата с известным заранее качеством и гарантированными затратами и сроками, то моделирование, как дисциплина:

изучает способы решения задач, то есть является инженерной наукой;

является универсальным инструментом, гарантирующим решение любых задач, независимо от предметной области.

1.2. Классификация моделей и моделирования

Каждая модель создается для конкретной цели и, следовательно, уникальна. Однако наличие общих черт позволяет сгруппировать все их многообразие в отдельные классы, что облегчает их разработку и изучение. В теории рассматривается много признаков классификации и их количество не установилось. Тем не менее, наиболее актуальны следующие **признаки классификации**:

характер моделируемой стороны объекта;
характер процессов, протекающих в объекте;
способ реализации модели.

1.2.1. Классификация моделей и моделирования по признаку «характер моделируемой стороны объекта»

В соответствии с этим признаком модели могут быть:
функциональными (кибернетическими);
структурными;
информационными.

Функциональные модели отображают только поведение, функцию моделируемого объекта. В этом случае моделируемый объект рассматривается как «черный ящик», имеющий входы и выходы. Физическая сущность объекта, природа протекающих процессов, структура объекта остаются вне внимания исследователя, хотя бы потому, что они неизвестны. При функциональном моделировании эксперимент состоит в наблюдении за выходом моделируемого объекта при изменении входных воздействий. По этим данным и строится модель поведения в виде некоторой математической функции.

Компьютерная шахматная программа – функциональная модель работы человеческого мозга при игре в шахматы.

Структурное моделирование – это создание и исследование модели, структура которой (элементы, связи) подобна структуре моделируемого объекта. Как выяснили ранее, подобие устанавливается не вообще, а относительно цели исследования. Поэтому она может быть описана на разных уровнях детализации. Наиболее общее описание структуры – это топологическое описание с помощью теории графов.

Учение войск – структурная модель вида боевых действий.

Информационная модель – совокупность данных и отношений между ними, описывающая различные свойства реального изделия, интересующие разработчика модели и потенциального или реального пользователя.

1.2.2. Классификация моделей и моделирования по признаку «характер процессов, протекающих в объекте»

По этому признаку модели могут быть детерминированными или стохастическими, статическими или динамическими, дискретными или непрерывными или дискретно-непрерывными.

Детерминированные модели отображают процессы, в которых отсутствуют случайные воздействия.

Стохастические модели отображают вероятностные процессы и события.

Статические модели служат для описания состояния объекта в какой-либо момент времени.

Динамические модели отображают поведение объекта во времени.

Дискретные модели отображают поведение систем с дискретными состояниями.

Непрерывные модели представляют системы с непрерывными процессами.

Дискретно-непрерывные модели строятся тогда, когда исследователя интересуют оба эти типа процессов.

Очевидно, конкретная модель может быть стохастической, статической, дискретной или какой-либо другой (рис. 1.2).

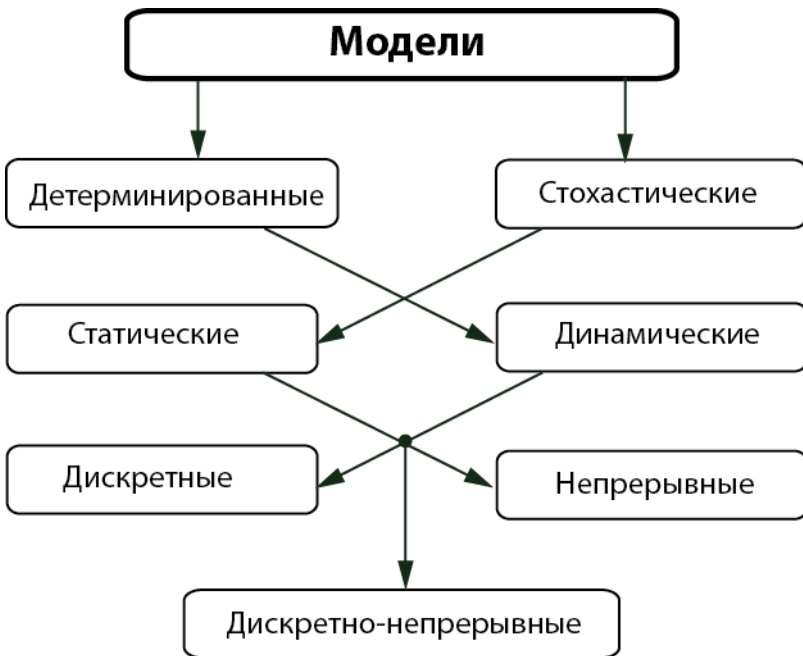


Рис. 1.2. Классификация моделей по признаку «характер процессов, протекающих в объекте»

1.2.3. Классификация моделей и моделирования по признаку «способ реализации модели»

Согласно этому признаку модели делятся на два обширных класса (рис. 1.3):

- абстрактные (мысленные) модели;
- материальные модели.

Нередко в практике моделирования присутствуют смешанные, абстрактно-материальные модели.

Абстрактные модели представляют собой определенные конструкции из общепринятых знаков на бумаге, другом материальном носителе или в виде компьютерной программы.

Абстрактные модели, не вдаваясь в излишнюю детализацию, можно разделить на:

- символьные;
- математические.

Символьная модель – это логический объект, замещающий реальный процесс и выражающий основные свойства его отношений с помощью определенной системы знаков или символов. Это либо слова естественного языка, либо слова соответствующего тезауруса, графики, диаграммы и т. п.

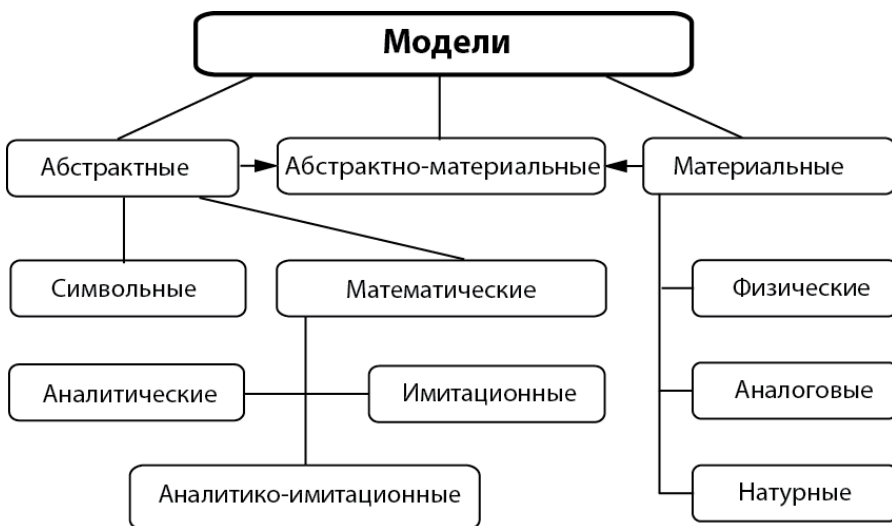


Рис. 1.3. Классификация по способу реализации модели

Символическая модель может иметь самостоятельное значение, но, как правило, ее построение является начальным этапом любого другого моделирования.

Математическое моделирование – это процесс установления соответствия моделируемому объекту некоторой математической конструкции, называемой математической моделью, и исследование этой модели, позволяющее получить характеристики моделируемого объекта.

Математическое моделирование – главная цель и основное содержание изучаемой дисциплины.

Математические модели могут быть:

аналитическими;

имитационными;

смешанными (аналитико-имитационными).

Аналитические модели – это функциональные соотношения: системы алгебраических, дифференциальных, интегро - дифференциальных уравнений, логических условий. Уравнения Максвелла – аналитическая модель электромагнитного поля. Закон Ома – модель электрической цепи. Также аналитическими моделями являются первый и второй законы Ньютона.

Преобразование математических моделей по известным законам и правилам можно рассматривать как эксперименты. Решение на основе аналитических моделей может быть получено в результате однократного просчета безотносительно к конкретным значениям характеристик («в общем виде»). Это наглядно и удобно для выявления закономерностей. Однако для сложных систем построить аналитическую модель, достаточно полно отражающую реальный процесс, удаётся не всегда. Тем не менее, есть процессы, например, марковские, актуальность моделирования которых аналитическими моделями доказана практикой.

Имитационное моделирование. Создание вычислительных машин обусловило развитие нового подкласса математических моделей – имитационных.

Имитационное моделирование предполагает представление модели в виде некоторого алгоритма – компьютерной программы, – выполнение которого имитирует последовательность смены состояний в системе и таким образом представляет собой поведение моделируемой системы. Имитационная модель реализует временную диаграмму функционирования моделируемой системы.

Процесс создания и испытания таких моделей называется имитационным моделированием, а сам алгоритм – имитационной моделью.

Имитационная модель – универсальное средство исследования сложных систем, представляющее собой логико-алгоритмическое описание поведения отдельных элементов системы и правил их взаимодействия, отображающих последовательность событий, возникающих в моделируемой системе.

В чем заключается принципиальное отличие имитационных и аналитических моделей?

В случае аналитического моделирования компьютер является мощным калькулятором, арифмометром. Аналитическая модель *решается* на компьютере.

В случае же имитационного моделирования имитационная модель – программа – *реализуется* на компьютере.

Имитационные модели достаточно просто учитывают влияние случайных факторов. Для аналитических моделей это серьезная проблема. При наличии случайных факторов необходимые характеристики моделируемых процессов получают многократными прогонами (реализациями) имитационной модели и дальнейшей статистической обработкой накопленной информации. Поэтому часто имитационное моделирование процессов со случайными факторами называют **имитационным статистическим моделированием**.

Если исследование объекта затруднено использованием только аналитического или имитационного моделирования, то применяют смешанное (*комбинированное*), аналитико-имитационное моделирование. При построении таких моделей процессы функционирования объекта декомпозируют на подпроцессы, для которых используют аналитические модели, а для остальных подпроцессов строят имитационные модели.

Материальное моделирование основано на применении моделей, представляющих собой реальные технические конструкции. Это может быть сам объект или его элементы (*натурное моделирование*). Это может быть специальное устройство – модель, имеющая либо физическое, либо геометрическое подобие оригиналу. Это может быть устройство иной физической природы, чем оригинал, но процессы, в котором описываются аналогичными математическими соотношениями. Это так называемое *аналоговое моделирование*. Такая аналогия наблюдается, например, между колебаниями антенны спутниковой

связи под ветровой нагрузкой и колебанием электрического тока в специально подобранной электрической цепи.

Нередко создаются **материально-абстрактные модели**. Та часть операции, которая не поддается математическому описанию, моделируется материально, остальная – абстрактно. Таковы, например, командно-штабные учения, когда работа штабов представляет собой натуральный эксперимент, а действия войск отображаются в документах.

Приведённая выше классификация является *идеальной*. Модели сложных систем обычно имеют комплексный вид, используют в своём составе сразу несколько представлений. Если удаётся свести модель к одному типу, для которого уже есть математический аппарат, то исследование модели, решение задач на ней существенно упрощается, становится типовым. Для этого модель должна быть различными способами приведена к каноническому виду, то есть к виду, для которого уже есть методы решения.

1.3. Этапы моделирования

Математическое моделирование как, впрочем, и любое другое, считается искусством и наукой. Известный специалист в области имитационного моделирования Роберт Шеннон так назвал свою широко известную в научном и инженерном мире книгу: «Имитационное моделирование – искусство и наука». Поэтому в инженерной практике нет формализованной инструкции, как создавать модели. И, тем не менее, анализ приёмов, которые используют разработчики моделей, позволяет выделить этапы моделирования.

Первый этап: уяснение целей моделирования. Вообще-то это главный этап любой деятельности. Цель существенным образом определяет содержание остальных этапов моделирования. Заметим, что различие между простой системой и сложной порождается не столько их сущностью, но и целями, которые ставит исследователь.

Обычно целями моделирования являются:

прогноз поведения объекта при новых режимах, сочетаниях факторов и т. п.;

подбор сочетания и значений факторов, обеспечивающих оптимальное значение показателей эффективности процесса;

анализ чувствительности системы на изменение факторов;

проверка различного рода гипотез о характеристиках случайных параметров исследуемого процесса;

определение функциональных связей между поведением («реакцией») системы и влияющими факторами, что может способствовать прогнозу поведения или анализу чувствительности;

уяснение сущности, лучшее понимание объекта исследования, а также формирование первых навыков для эксплуатации моделируемой или действующей системы.

Второй этап: построение концептуальной модели. Концептуальная модель (от лат. *conceptio*) – модель на уровне определяющего замысла, который формируется при изучении моделируемого объекта. На этом этапе исследуется объект, устанавливаются необходимые упрощения и аппроксимации. Выявляются существенные аспекты, исключаются второстепенные. Устанавливаются единицы измерения и диапазоны изменения переменных модели. Если возможно, то концептуальная модель представляется в виде известных и хорошо разработанных систем: массового обслуживания, управления, авторегулирования, разного рода автоматов и т. д. Концептуальная модель полностью подводит итог изучению проектной документации или экспериментальному обследованию моделируемого объекта.

Результатом второго этапа является обобщенная схема модели, полностью подготовленная для математического описания – построения математической модели.

Третий этап: выбор языка программирования или моделирования, разработка алгоритма и программы модели. Модель может быть аналитической или имитационной, или их сочетанием. В случае аналитической модели исследователь должен владеть методами решения.

В истории математики (а это, впрочем, и есть история математического моделирования) есть много примеров тому, когда необходимость моделирования разного рода процессов приводила к новым открытиям. Например, необходимость моделирования неравномерного движения привела к открытию и разработке дифференциального исчисления (Лейбниц (1646-1716) и Ньютон (1643-1727)) и соответствующих методов решения. Проблемы аналитического моделирования остойчивости кораблей привели академика Крылова А. Н. (1863-1945) к созданию теории приближённых вычислений и аналоговой вычислительной машины.

Результатом третьего этапа моделирования является программа, составленная на наиболее удобном для моделирования и исследования языке – универсальном или специальном.

Четвертый этап: планирование эксперимента. Математическая модель является объектом эксперимента. Эксперимент должен быть в максимально возможной степени информативным, удовлетворять ограничениям, обеспечивать получение данных с необходимой точностью и достоверностью. Существует теория планирования эксперимента, нужные нам элементы этой теории мы изучим в соответствующем месте дисциплины.

Результат четвёртого этапа – план эксперимента.

Пятый этап: выполнение эксперимента с моделью. Если модель аналитическая, то эксперимент сводится к выполнению расчётов при варьируемых исходных данных. При имитационном моделировании модель реализуется на компьютере с фиксацией и последующей обработкой получаемых данных. Эксперименты проводятся в соответствии с планом, который может быть включен в алгоритм модели. В современных системах моделирования такая возможность есть.

Шестой этап: обработка, анализ и интерпретация данных эксперимента. В соответствии с целью моделирования применяются разнообразные методы обработки: определение разного рода характеристик случайных величин и процессов, выполнение анализов – дисперсионного, регрессионного, корреляционного и др. Многие из этих методов входят в системы моделирования (GPSS World, AnyLogic и др.) и могут применяться автоматически. Не исключено, что в ходе анализа полученных результатов модель может быть уточнена, дополнена или даже полностью пересмотрена.

После анализа результатов моделирования осуществляется их интерпретация, то есть перевод результатов в термины предметной области. Это необходимо, так как обычно специалист предметной области (тот, кому нужны результаты исследований) не обладает терминологией математики и моделирования и может выполнять свои задачи, оперируя лишь хорошо знакомыми ему понятиями.

На этом рассмотрении последовательности моделирования закончим. Приведём ещё вариант этапов моделирования (рис. 1.4) и сделаем весьма важный вывод о необходимости документирования результатов каждого этапа. Это необходимо в силу следующих причин.

Во-первых, моделирование процесс итерационный, то есть с каждого этапа может осуществляться возврат на любой из предыдущих этапов (рис. 1.4) для уточнения информации, необходимой на этом этапе, а документация может сохранить результаты, полученные на предыдущих итерациях.



Рис. 1.4. Вариант этапов моделирования

Во-вторых, в случае исследования сложной системы в нём участвуют большие коллективы разработчиков, причём различные этапы

выполняются различными коллективами. Поэтому результаты, полученные на каждом этапе, должны быть переносимы на последующие этапы, то есть иметь унифицированную форму представления и понятное другим заинтересованным специалистам содержание.

В-третьих, результат каждого из этапов должен являться самоценным продуктом. Например, концептуальная модель может и не использоваться для дальнейшего преобразования в математическую модель, а являться описанием, хранящим информацию о системе, которое может использоваться как архив, в качестве средства обучения.

1.4. Адекватность модели

Итак, мы установили: модель предназначена для замены оригинала при исследованиях, которым подвергать оригинал нельзя или нецелесообразно. Но замена оригинала моделью возможна, если они в достаточной степени похожи или адекватны.

Адекватность означает, достаточно ли хорошо с точки зрения целей исследования результаты, полученные в ходе моделирования, отражают истинное положение дел. Термин происходит от латинского *adaequatus* – приравненный.

Говорят, что модель адекватна оригиналу, если при ее интерпретации возникает «портрет», в высокой степени сходный с оригиналом.

До тех пор, пока не решён вопрос, правильно ли отображает модель исследуемую систему (то есть адекватна ли она), ценность модели нулевая!

Термин «адекватность» как видно носит весьма расплывчатый смысл. Понятно, что результативность моделирования значительно возрастёт, если при построении модели и переносе результатов с модели на систему – оригинал может воспользоваться некоторой теорией, уточняющей идею подобия, связанную с используемой процедурой моделирования.

К сожалению теории, позволяющей оценить адекватность математической модели и моделируемой системы, нет, в отличие от хорошо разработанной теории подобия явлений одной и той же физической природы.

Проверку адекватности проводят на всех этапах построения модели, начиная с самого первого этапа – концептуального анализа. Если описание системы будет составлено не адекватно реальной системе, то и модель, как бы точно она не отображала описание системы, не

будет адекватной оригиналу. Здесь сказано «как бы точно», так как имеется в виду, что вообще не существуют математические модели, абсолютно точно отображающие процессы, существующие в реальности.

Если изучение системы проведено качественно и концептуальная модель достаточно точно отражает реальное положение дел, то далее перед разработчиками стоит лишь проблема эквивалентного преобразования одного описания в другое.

Итак, можно говорить об адекватности модели в любой её форме и оригинала, если:

описание поведения, созданное на каком-либо этапе, достаточно точно совпадает с поведением моделируемой системы в одинаковых ситуациях;

описание убедительно представительно относительно свойств системы, которые должны прогнозироваться с помощью модели.

Предварительно исходный вариант математической модели подвергается следующим проверкам:

все ли существенные параметры включены в модель;

нет ли в модели несущественных параметров;

правильно ли отражены функциональные связи между параметрами;

правильно ли определены ограничения на значения параметров;

не даёт ли модель абсурдные ответы, если её параметры принимают предельные значения.

Такая предварительная оценка адекватности модели позволяет выявить в ней наиболее грубые ошибки.

Но все эти рекомендации носят неформальный, рекомендательный характер. Формальных методов оценки адекватности не существует! Поэтому, в основном, качество модели (и в первую очередь степень ее адекватности системе) зависит от опыта, интуиции, эрудиции разработчика модели и других субъективных факторов.

Окончательное суждение об адекватности модели может дать лишь практика (хотя для оценки адекватности используются и экспертные методы), то есть сравнение модели с оригиналом на основе экспериментов с объектом и моделью. Модель и объект подвергаются одинаковым воздействиям и сравниваются их реакции. Если реакции одинаковы (в пределах допустимой точности), то делается вывод, что модель адекватна оригиналу. Однако надо иметь в виду следующее:

воздействия на объект носят ограниченный характер из-за возможного разрушения объекта, недоступности к элементам системы и т. д.;

воздействия на объект имеют физическую природу (изменение питающих токов и напряжений, температуры, скорости вращения валов и т. д.), а на математическую модель – это числовые аналоги физических воздействий.

Для оценки степени подобия структур объектов (физических или математических) существует понятие *изоморфизма* (изо – одинаковый, равный, морфе – форма, греч.).

Две системы считаются изоморфными, если существует взаимно однозначное соответствие между элементами и отношениями (связями) этих систем.

Изоморфны, например, множество действительных положительных чисел и множество их логарифмов. Каждому элементу одного множества – числу соответствует значение его логарифма в другом, умножению двух чисел в первом множестве – сложение их логарифмов в другом. С точки зрения пассажира план метрополитена, находящийся в каждом вагоне поезда метро, изоморфен реальному географическому расположению рельсовых путей и станций, хотя для рабочего, ремонтирующего рельсовые пути, этот план естественно не является изоморфным. Фотография является изоморфным отображением реального лица для полицейского, родителей, знакомых, но не является таковым для художника.

При моделировании сложных систем достигнуть такого полного соответствия трудно, да и нецелесообразно. При моделировании абсолютное подобие не имеет места. Стремятся лишь к тому, чтобы модель достаточно хорошо отражала исследуемую сторону функционирования объекта. Модель по сложности может стать аналогичной исследуемой системе и никакого упрощения исследования не будет.

Для оценки подобия в поведении (функционировании) систем существует понятие *изофункционализма*.

Две системы произвольной, а подчас неизвестной структуры изофункциональны, если при одинаковых воздействиях они проявляют одинаковые реакции. Такое моделирование называется функциональным или кибернетическим и в последние годы получает все большее распространение, например, при моделировании человеческого интеллекта (игра в шахматы, доказательство теорем, распознавание образов и т. д.). Функциональные модели не копируют структуры. Но копируя поведение, исследователи последовательно «подбираются» к познанию структур объектов (человеческого мозга, Солнца, и др.).

1.5. Требования, предъявляемые к моделям

Итак, общие требования к моделям.

1. Модель должна быть *актуальной*. Это значит, что модель должна быть нацелена на важные для лиц, принимающих решения, проблемы.

2. Модель должна быть *результативной*. Это значит, что полученные результаты моделирования могут найти успешное применение. Данное требование может быть реализовано только в случае правильной формулировки требуемого результата.

3. Модель должна быть *достоверной*. Это значит, что результаты моделирования не вызовут сомнения. Данное требование тесно связано с понятием адекватности, то есть, если модель неадекватна, то она не может давать достоверных результатов.

4. Модель должна быть *экономичной*. Это значит, что эффект от использования результатов моделирования превышает расходы ресурсов на её создание и исследование.

Эти требования (обычно их называют внешними) выполнимы при условии обладания моделью внутренними свойствами.

Модель должна быть:

1. *Существенной*, т. е. позволяющей вскрыть сущность поведения системы, вскрыть неочевидные, нетривиальные детали.

2. *Мощной*, т. е. позволяющей получить широкий набор существенных сведений.

3. *Простой* в изучении и использовании, легко рассчитываемой на компьютере.

4. *Открытой*, т.е. позволяющей её модификацию.

В заключение темы сделаем несколько замечаний.

Трудно ограничить область применения математического моделирования. При изучении и создании промышленных и военных систем практически всегда можно определить цели, ограничения и предусмотреть, чтобы конструкция или процесс подчинялись естественным, техническим и (или) экономическим законам.

Круг аналогий, которые можно использовать в качестве моделей, также практически неограничен. Следовательно, надо постоянно расширять свое образование в конкретной области, но, в первую очередь, в математике.

В последние десятилетия появились проблемы с неясными и противоречивыми целями, диктуемыми политическими и социальными

факторами. Математическое моделирование в этой области пока еще проблематично. Что это за проблемы? Защита от загрязнения окружающей среды; предсказаний извержений вулканов, землетрясений, цунами; рост городов; руководство боевыми действиями и ряд других. Но, тем не менее, «процесс пошёл», прогресс не остановим, и проблемы моделирования таких сверхсложных систем постоянно находят своё разрешение. Здесь следует отметить лидирующую роль отечественных учёных и, в первую очередь, академика Н. Н. Моисеева, его учеников и последователей.

1.6. Компьютерное моделирование

В данном пособии будем рассматривать только один из видов компьютерного моделирования (КМ), а именно имитационное моделирование (ИМ) СМО, и частично вопросы планирования экспериментов и статистической обработки результатов ИМ.

Тем не менее, приведём краткое описание математического моделирования (ММ), которое позволяет благодаря абстрактным математическим формулам точно, однозначно и количественно оценить исследуемый объект. Усложнение исследуемых систем привело к резкому усложнению их математического описания, что, в свою очередь, приводило к необходимости делать всевозможные упрощающие допущения. При этом возникла опасность ухода от реального представления о системе. Выходом из этого положения являлся либо прогресс самих математических методов, либо изыскание иных методов описания. Появление мощных современных компьютеров и возникновение информационных технологий привело ко второму рождению ММ. Оно стало вторгаться практически во все сферы человеческой деятельности. В ряде областей ММ стало вытеснять физическое моделирование, так произошло, в частности, в авиационной промышленности, где начался демонтаж аэродинамических труб. Дальнейший прогресс ММ идёт за счёт:

разработки новых численных методов решения задач моделирования, возможных только при условии использования информационных технологий (ИТ);

стремительного увеличения объёмов памяти и производительности ЭВМ, что позволяет на много порядков увеличить размерности решаемых задач и перейти благодаря этому к качественно новым задачам моделирования.

Так, прогресс ММ позволил: исследовать эффекты синергизма (комбинированное действие, когда выходной эффект системы превышает действие, оказываемое компонентами по отдельности); оценивать бифуркационные состояния (вероятностное разветвление процесса функционирования системы); прогнозировать развитие диссипативных структур (переход в качественно новое состояние, характеризующееся более высоким уровнем самоорганизации); создавать более совершенные модели Вселенной (Большой взрыв, горячая Вселенная, теория струн), теории искусственного интеллекта, исследований в разных областях социологии, экономики и т. д.

В то же время возникла мощная оппозиция применению ММ в плохо структурируемых, не формализуемых областях, т. е. в таких областях, где человек (оператор, ЛОР, ЛПР) является основным элементом. По определению академика РАН РФ А. А. Самарского, процесс ММ базируется на триаде «модель—алгоритм—программа». До появления ЭВМ основную роль играла модель в виде математических уравнений, алгоритм представлял собой схему ручных расчётов для приближенного решения уравнений, а программа отсутствовала вообще. В начале использования ЭВМ первого поколения программе отводилась второстепенная роль – представление алгоритма в машинных кодах. Развитие ИТ привело к тому, что ЭВМ стали использовать для моделирования процессов функционирования системы, причём в этом случае имелись алгоритм и программа, а математическая модель в её классическом виде практически отсутствовала или молчаливо предполагалось, что математической моделью является одно из аналитических представлений. Это направление получило название имитационного моделирования и представлено в работах Н. П. Бусленко, Н. Н. Моисеева, Р. Шеннона и многих других. Таким образом, в ММ началось опережающее развитие третьей компоненты триады – программы или программного обеспечения процесса моделирования.

Необходимо четко понимать разницу между программированием и моделированием. Программирование в настоящее время располагает большим арсеналом языков программирования, средств автоматизации управления вычислительными ресурсами, создания программ, автоматизации работ с большими массивами данных, так называемых систем управления базами данных (СУБД).

Однако весь этот арсенал направлен на создание программ, но модель не должна превращаться только в программу, описывающую аб-

структные алгоритмы или логические отношения. Компьютерная модель должна оставаться, прежде всего, моделью реального объекта не зависимо от того, чем описывается его поведение: набором формул или правил, графиком или прогнозными оценками экспертов. Поэтому модель должна допускать исследование всех интересных возможностей: анализ чувствительности, изменение выходных характеристик, определение областей устойчивости и степень робастности, оптимизацию параметров, оценку вариантов построения и т. д.

В связи со сказанным все чаще в литературе появляется термин *компьютерное моделирование*. КМ объединяет достижения математического моделирования, системного программирования и информационных технологий.

Компьютерное моделирование обладает:

способностью понимать, интерпретировать и использовать формализованную и не формализованную информацию (математические формулы, логические правила, вербальные описания и т. п.);

различными формами представления данных и знаний, заполняя пространство между ММ с его аналитическими формами описания и искусственным интеллектом с его формами и правилами представления знаний;

способностью участвовать в процессе не только автоматизации научных исследований за счёт использования самой ЭВМ для модификации различных режимов применения КМ, но и интеграции всех этапов жизненного цикла системы путём использования быстро развивающихся методов ИТ (широко распространенные во всем мире CALS-технологии, CASE-технологии, технологии ICAM и IDEF);

возможностью расширения круга пользователей, от узкого круга специалистов-математиков и профессиональных программистов до большого класса исследователей, не обладающих профессиональными знаниями в областях математики и программирования, но хорошо знающих предметную область и умеющих обращаться с пакетами прикладных программ.

Вопросы для самоконтроля

1.1. Что такое модель? Раскройте смысл фразы: «модель есть объект и средство эксперимента».

1.2. Обоснуйте необходимость моделирования.

1.3. Назовите общие классификационные признаки моделей.

- 1.4. Нужно ли стремиться к абсолютному подобию модели и оригинала?
- 1.5. Назовите и поясните три аспекта процесса моделирования.
- 1.6. Что значит структурная модель?
- 1.7. Что такое функциональная модель?
- 1.8. Дайте характеристику символьной модели.
- 1.9. В каком случае можно утверждать, что имитационная модель системы создана?
- 1.10. Классификация моделей по характеру процессов, протекающих в моделируемых объектах.
- 1.11. Классификация моделей по способам их реализации.
- 1.12. Сущность математического моделирования и его основных классов: аналитического и имитационного.
- 1.13. Назовите этапы моделирования и дайте им краткую характеристику.
- 1.14. Чем вызвана итерационность процесса моделирования?
- 1.15. Дайте характеристику концептуальной модели.
- 1.16. На какой триаде базировалось математическое моделирование? Дайте пояснения.
- 1.17. На какой триаде базируется компьютерное моделирование? Дайте пояснения.
- 1.18. Что такое адекватность модели?
- 1.19. Дайте понятия изоморфизма. Приведите примеры.
- 1.20. Дайте понятия изофункционализма. Приведите примеры.
- 1.21. Общие требования (внешние) к моделям.
- 1.22. Почему говорят, что имитационное моделирование это «наука и искусство»?
- 1.23. Внутренние свойства модели.
- 1.24. Приведите примеры объектов и возможных их моделей в своей предметной области.
- 1.25. В качестве примера модели поведения можно назвать:
 - а) правила техники безопасности в компьютерном классе;
 - б) список курсантов взвода;
 - в) план оповещения при приведении в соответствующую степень боевой готовности;
 - г) план военного городка.
- 1.26. Понятие модели имеет смысл при наличии:
 - а) моделирующего субъекта и моделируемого объекта;

б) цели моделирования и моделируемого объекта;
в) моделирующего субъекта, цели моделирования и моделируемого объекта;

г) цели моделирования и двух различных объектов.

1.27. Сколько моделей можно создать при описании боя:

а) 1;

б) 5;

в) 15;

г) множество;

д) более 15.

1.28. Утверждение истинно:

а) «Модель обладает всеми признаками объекта-оригинала»;

б) «Можно создавать и использовать разные модели объекта»;

в) «Можно создавать и использовать только натурные модели объекта»;

г) «Можно создавать и использовать только единственную модель объекта».

1.29. Процесс построения модели объекта, как правило, предполагает описание:

а) всех свойств исследуемого объекта;

б) свойств безотносительно к цели моделирования;

в) всех возможных пространственно-временных характеристик;

г) наиболее существенных с точки зрения цели моделирования свойств объекта.

1.30. Поясните отличие имитационной модели от программы.

Глава 2. Концепция дискретных систем для имитационного моделирования

1. Основные понятия

Для описания одного и того же понятия многочисленные литературные источники по моделям и методам теории массового обслуживания зачастую используют разные термины. Сама «теория массового обслуживания» часто называется «теорией очередей» (в англоязычной литературе Queue Theorie). Наряду с термином «обслуживающий прибор» используются термины «устройство», «канал», «линия» и т. д. Обычно это связано с прикладной областью, в которой применяются модели массового обслуживания. Например, термины «вызов» и «линия» используются в телефонии (откуда собственно и пошла теория массового обслуживания). В связи с этим, желательно иметь однозначные термины и понятия.

Рассматривая модели массового обслуживания как абстрактные математические модели, мы будем использовать в данной главе термины безотносительно прикладной области применения этих моделей. Примеры применения будут приведены в последующих главах.

Система массового обслуживания (СМО) – математический (абстрактный) объект, элементами которого являются (рис. 2.1):

- входной (входящий) поток требований (заявок) на обслуживание;
- приборы (каналы) обслуживания;
- очередь заявок, ожидающих обслуживания;
- выходной (выходящий) поток обслуженных заявок;
- поток не обслуженных заявок.

Заявка (требование, запрос, вызов, клиент, сообщение, пакет) – объект, поступающий в СМО и требующий обслуживания в обслуживающем приборе. Совокупность заявок, распределенных во времени, образуют *входной поток заявок*.

Обслуживающий прибор или просто **прибор (устройство, канал, линия, орудие)** – элемент СМО, функцией которого является обслуживание заявок.

В каждый момент времени в приборе на обслуживании может находиться только одна заявка.

Обслуживание – задержка заявки на некоторое время в обслуживающем приборе.

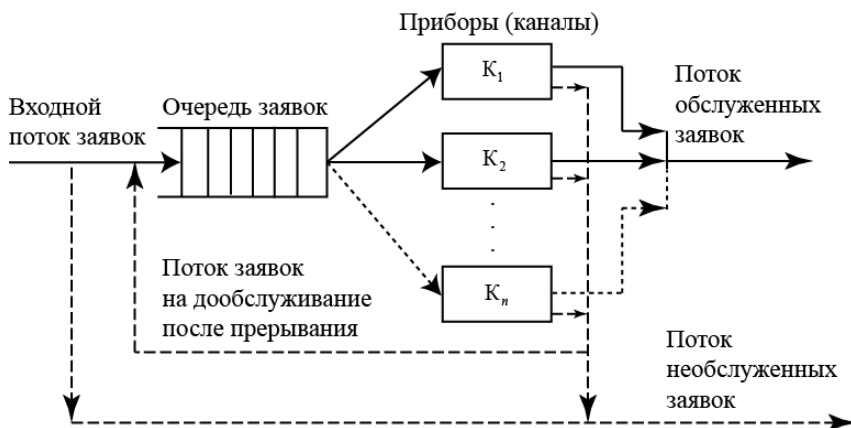


Рис. 2.1. Система массового обслуживания

Длительность обслуживания – время задержки (обслуживания) заявки в приборе.

Накопитель (буфер) – совокупность мест для ожидания заявок перед обслуживающим прибором. Количество мест для ожидания определяет **ёмкость накопителя**.

Заявка, поступившая на вход СМО, может находиться в двух состояниях:

в состоянии *обслуживания* (в приборе);

в состоянии *ожидания* (в накопителе), если все приборы заняты обслуживанием других заявок.

Заявки, находящиеся в накопителе и ожидающие обслуживания, образуют **очередь** заявок. Количество заявок, ожидающих обслуживания в накопителе, определяет **длину очереди**.

Дисциплина буферизации (дисциплина постановки в очередь) – правило занесения поступающих заявок в накопитель (буфер).

Дисциплина обслуживания – правило выбора заявок из очереди для обслуживания в приборе.

Приоритет – преимущественное право на занесение (в накопитель) или выбор из очереди (для обслуживания в приборе) заявок одного класса по отношению к заявкам других классов.

Существует большое многообразие систем массового обслуживания, отличающихся структурной и функциональной организацией. В то же время, разработка аналитических методов расчёта характеристик функционирования СМО во многих случаях предполагает нали-

чие ряда ограничений и допущений, сужающих множество исследуемых СМО.

Далее при рассмотрении СМО, если не указано другое, будем использовать следующие **предположения**:

поступившая в систему заявка *мгновенно* попадает на обслуживание, если прибор свободен;

в приборе на обслуживании в каждый момент времени может находиться только *одна* заявка;

после завершения обслуживания какой-либо заявки в приборе очередная заявка выбирается на обслуживание из очереди мгновенно, то есть, другими словами, прибор *не простаивает*, если в очереди есть хотя бы одна заявка;

поступление заявок в СМО и длительности их обслуживания не зависят от того, сколько заявок уже находится в системе, или от каких-либо других факторов;

длительность обслуживания заявок не зависит от скорости (интенсивности) поступления заявок в систему.

Сеть массового обслуживания (СеМО) – совокупность взаимосвязанных СМО, в среде которых циркулируют заявки (рис. 2.2а).

Основными элементами СеМО являются узлы ($У$) и источники (генераторы) заявок ($Г$).

Узел сети представляет собой систему массового обслуживания.

Источник – генератор заявок, поступающих в сеть и требующих определенных этапов обслуживания в узлах сети.

Для упрощенного изображения СеМО используется граф СеМО.

Граф СеМО – ориентированный граф, вершины которого соответствуют узлам СеМО, а дуги отображают переходы заявок между узлами (рис. 2.2б).

Переходы заявок между узлами СеМО, в общем случае, могут быть заданы в виде вероятностей передач. Путь движения заявок в СеМО называется **маршрутом**.

Остановимся на некоторых элементах СМО более подробно.

Входной (входящий) поток заявок.

Совокупность событий, распределенных во времени, называется **поток**ом. Если событие заключается в появлении заявок, имеем **поток заявок**.

Для описания потока заявок, в общем случае, необходимо задать интервалы времени $\tau = t_k - t_{k-1}$ между соседними моментами t_{k-1} и

t_k поступления заявок с порядковыми номерами $k-1$ и k соответственно ($k = \overline{1, 2, \dots}$, $t_0 = 0$ – начальный момент времени).

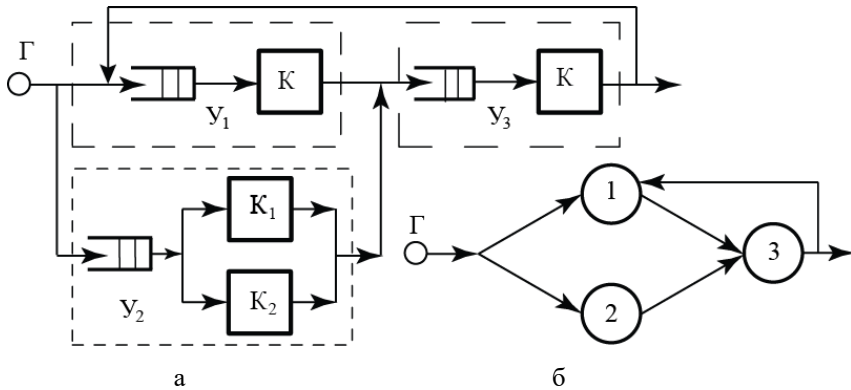


Рис. 2.2. Сеть массового обслуживания

Основной характеристикой потока заявок является его **интенсивность** λ – среднее число заявок, проходящих через некоторую границу за единицу времени. Величина $\tau = 1/\lambda$ определяет **средний интервал времени между двумя последовательными заявками**.

Поток, в котором интервалы времени τ_k между соседними заявками принимают определенные заранее известные значения, называется **детерминированным**. Если при этом интервалы одинаковы ($\tau_k = \tau$ для всех $k = \overline{1, 2, \dots}$), то поток называется **регулярным**. Для полного описания регулярного потока заявок достаточно задать интенсивность потока λ или значение интервала $\tau = 1/\lambda$.

Поток, в котором интервалы времени τ_k между соседними заявками представляют собой случайные величины, называется **случайным**. Для полного описания случайного потока заявок, в общем случае, необходимо задать законы распределений $F_k(\tau_k)$ всех интервалов τ_k , $k = \overline{1, 2, \dots}$.

Случайный поток, в котором все интервалы τ_1, τ_2, \dots между заявками независимы в совокупности и описываются функциями распределений $F_1(\tau_1), F_2(\tau_2), \dots$ называется потоком с **ограниченным последствием**.

Случайный поток, в котором все интервалы τ_1, τ_2, \dots распределены по одному и тому же закону $F(t)$, называется **рекуррентным**.

Поток заявок называется **стационарным**, если интенсивность λ и закон распределения $F(t)$ интервалов между последовательными заявками не меняются со временем. В противном случае поток заявок является **нестационарным**.

Поток заявок называется **ординарным**, если в каждый момент времени t_k может появиться только одна заявка. Если в какой-либо момент времени может появиться более одной заявки, то имеем **неординарный** или **групповой** поток заявок.

Поток заявок называется потоком **без последствий**, если заявки поступают *независимо* друг от друга, то есть момент поступления очередной заявки не зависит от того, когда и сколько заявок поступило до этого момента.

Стационарный ординарный поток без последствий называется **простейшим**.

Интервалы времени τ между заявками в простейшем потоке распределены по **экспоненциальному закону** с функцией распределения $F(t) = 1 - e^{-\lambda t}$, плотностью распределения $f(t) = \lambda e^{-\lambda t}$, где $\lambda > 0$ – параметр распределения, представляющий собой интенсивность потока заявок.

Простейший поток часто называют **пуассоновским**, поскольку число заявок k , поступающих за некоторый заданный промежуток времени Δt , распределено по **закону Пуассона**:

$$P_k(\Delta t) = \frac{(\lambda \cdot \Delta t)^k}{k!} e^{-\lambda \cdot \Delta t},$$

где $P_k(\Delta t)$ – вероятность поступления ровно k заявок за некоторый фиксированный интервал времени Δt .

Следует отметить, что пуассоновский поток, в отличие от простейшего, может быть:

стационарным, если интенсивность λ не меняется со временем;

нестационарным, если интенсивность потока зависит от времени:

$$\lambda = \lambda(t).$$

В то же время, простейший поток, по определению, всегда является стационарным.

Аналитические исследования моделей массового обслуживания часто проводятся в предположении о простейшем потоке заявок, что обусловлено рядом присущих ему замечательных особенностей.

1. Суммирование (объединение) потоков. Сумма N независимых стационарных ординарных потоков с интенсивностями $\lambda_1, \lambda_2, \dots, \lambda_N$ образует простейший поток с интенсивностью $\lambda = \sum_{i=1}^N \lambda_i$

при условии, что складываемые потоки оказывают более или менее одинаково малое влияние на суммарный поток. На практике суммарный поток близок к простейшему при $N \geq 5$. Значит, *при суммировании независимых простейших потоков суммарный поток будет простейшим* при любом значении N . Простейший поток в теории СМО аналогичен нормальному закону распределения в теории вероятностей: к простейшему потоку приводит предельный переход для потока, являющегося суммой потоков с произвольными характеристиками при бесконечном увеличении числа слагаемых и уменьшении их интенсивности.

2. Вероятностное разрежение потока. *Вероятностное* (но не *детерминированное*) разрежение *простейшего* потока заявок, при котором любая заявка случайным образом с некоторой вероятностью p исключается из потока независимо от того, исключены другие заявки или нет, приводит к образованию *простейшего* потока с интенсивностью $\lambda^* = p\lambda$, где λ – интенсивность исходного потока. Поток исключенных заявок с интенсивностью $\lambda^{**} = (1-p)\lambda$ – тоже *простейший* поток.

3. Если обслуживающие каналы (приборы) рассчитаны на простейший поток заявок с интенсивностью λ , то обслуживание других типов потоков (с той же интенсивностью) будет обеспечено с не меньшей эффективностью.

4. Простота. Предположение о простейшем потоке заявок позволяет для многих математических моделей получить в явном виде зависимости характеристик СМО от параметров. Наибольшее число аналитических результатов получено для простейшего потока заявок.

Анализ моделей с потоками заявок, отличными от простейших, обычно усложняет математические выкладки и не всегда позволяет получить аналитическое решение в явном виде. Своё название «*простейший*» поток получил именно благодаря этой особенности.

Часто встречаются системы, у которых поток входных заявок зависит от количества заявок, находящихся в обслуживании. Такие СМО называют *замкнутыми* (иначе – *разомкнутыми*). Например, работа ремонтного подразделения части может быть представлена моделью замкнутой СМО. Пусть это подразделение предназначено для обслуживания радиостанций, которых в объединении m . Каждая из них имеет интенсивность отказов λ . Входной поток отказавшей аппаратуры будет иметь интенсивность λ_p :

$$\lambda_p = \lambda(m - n),$$

где n – количество радиостанций, уже находящихся в мастерской на ремонте.

Заявки могут иметь разные права на начало обслуживания. В этом случае говорят, что заявки *неоднородные*. Преимущества одних потоков заявок перед другими задаются приоритетами.

Важной характеристикой входного потока является *коэффициент вариации*:

$$v = \frac{\sigma}{\tau_{\text{инт}}},$$

где $\tau_{\text{инт}}$ – математическое ожидание длины интервала;

σ – среднее квадратическое отклонение длины интервала $\tau_{\text{инт}}$.

Для простейшего потока $\left(\sigma = \frac{1}{\lambda}, \tau = \frac{1}{\lambda}\right)$: $v = 1$. Для большинства реальных потоков $0 \leq v \leq 1$. При $v = 0$ поток регулярный, детерминированный. Коэффициент вариации – характеристика, отражающая степень неравномерности поступления заявок.

Каналы (приборы) обслуживания.

В СМО могут быть один или несколько обслуживающих приборов (каналов). Согласно с этим СМО называют одноканальными или многоканальными.

Многоканальные СМО могут состоять из однотипных или разнотипных приборов.

Основная характеристика канала – длительность обслуживания.

Длительность обслуживания – время нахождения заявки в приборе – в общем случае величина случайная. В случае неоднородной нагрузки СМО длительности обслуживания заявок разных классов могут различаться законами распределений или только средними значениями.

При этом обычно предполагается независимость длительностей обслуживания заявок каждого класса.

Часто длительность обслуживания заявок предполагается распределенной по *экспоненциальному закону*, что существенно упрощает аналитические выкладки. Это обусловлено тем, что процессы, протекающие в системах с экспоненциальным распределением интервалов времени, являются *марковскими* процессами:

$$F(t) = 1 - e^{-\mu t}, \quad f(t) = \mu e^{-\mu t},$$

где μ – *интенсивность обслуживания*, здесь $\mu = \frac{1}{\tau_{\text{обсл}}}$; $\tau_{\text{обсл}}$ – математическое ожидание времени обслуживания.

При исследовании СМО выпадает из рассмотрения сущность обслуживания, качество обслуживания.

Каналы могут быть *абсолютно надежными*, то есть не выходить из строя. Вернее, так может быть принято при исследовании. Каналы могут обладать *конечной надежностью*. В этом случае модель СМО значительно сложнее.

Стратегия управления потоками заявок в моделях массового обслуживания задается в виде:

- дисциплины буферизации;
- дисциплины обслуживания.

Дисциплины буферизации и дисциплины обслуживания могут быть классифицированы по следующим признакам:

- наличие приоритетов между заявками разных классов;
- способ (режим) вытеснения заявок из очереди (для дисциплин буферизации) и назначения заявок на обслуживание (для дисциплин обслуживания);
- правило вытеснения или выбора заявок на обслуживание;
- возможность изменения приоритетов.

Одна из возможных **классификаций дисциплин буферизации (постановки в очередь)** в соответствии с перечисленными признаками представлена на рис. 2.3.

В зависимости от *наличия или отсутствия приоритетов* между заявками разных классов все дисциплины буферизации могут быть разбиты на две группы:

- бесприоритетные;
- приоритетные.



Рис. 2.3. Вариант классификации дисциплин буферизации

По способу вытеснения заявок из накопителя можно выделить следующие классы дисциплин буферизации:

- без вытеснения заявок – заявки, поступившие в систему и заставшие накопитель заполненным до конца, теряются;
- с вытеснением заявки данного класса, то есть такого же класса, что и поступившая;
- с вытеснением заявки самого низкоприоритетного класса;
- с вытеснением заявки, принадлежащей группе низкоприоритетных классов.

Два первых класса относятся к беспriorитетным дисциплинам буферизации, а остальные – к приоритетным.

Дисциплины буферизации могут использовать следующие *правила вытеснения заявок из накопителя*:

- вытеснение случайное;
- вытеснение последней заявки, то есть поступившей в систему позже всех;
- вытеснение «долгой» заявки, то есть находящейся в накопителе дольше всех.

Иногда ёмкость накопителя в моделях полагают неограниченной, хотя в реальной системе ёмкость ограничена. Такое допущение оправдано, когда вероятность потери заявки в реальной системе из-за переполнения ёмкости накопителя меньше 10^{-3} . В этом случае дисциплина практически не влияет на характеристики обслуживания заявок.

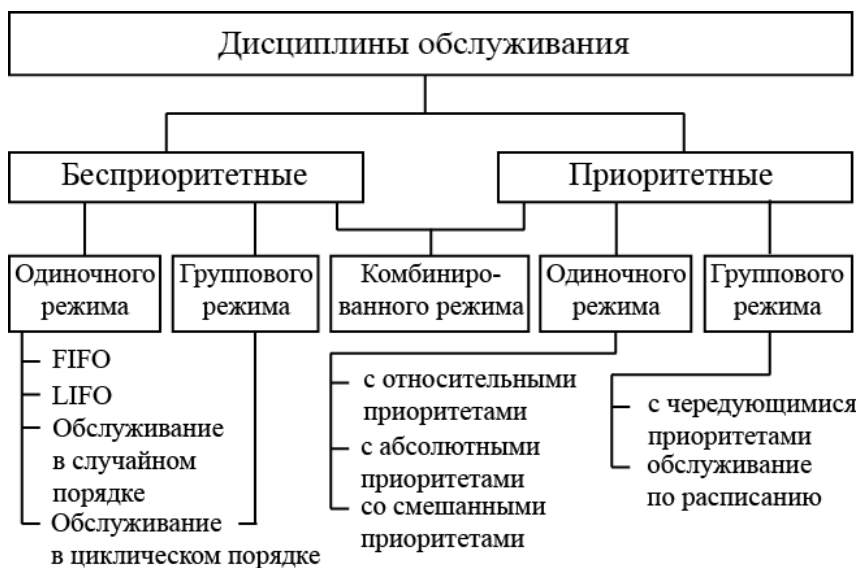


Рис. 2.4. Вариант классификации дисциплин обслуживания

На рис. 2.4 представлена **классификация дисциплин обслуживания** заявок в соответствии с теми же признаками, что и для дисциплин буферизации.

В зависимости от *наличия или отсутствия приоритетов* между заявками разных классов все дисциплины обслуживания, как и дисциплины буферизации, могут быть разбиты на две группы:

- бесприоритетные;
- приоритетные.

По *способу назначения заявок на обслуживание* дисциплины обслуживания могут быть разделены на дисциплины:

- одиночного режима;
- группового режима;
- комбинированного режима.

В дисциплинах обслуживания **одиночного режима** всякий раз на обслуживание *назначается только одна заявка* (просмотр очередей с целью назначения на обслуживание в приборе очередной заявки выполняется после обслуживания каждой заявки).

В дисциплинах обслуживания **группового режима** всякий раз на обслуживание *назначается группа заявок* одной очереди (просмотр очередей с целью очередного назначения на обслуживание выполня-

ется только после обслуживания всех заявок ранее назначенной группы). В предельном случае назначаемая на обслуживание группа заявок может включать в себя все заявки данной очереди. Заявки назначенной на обслуживание группы *последовательно выбираются из очереди* и обслуживаются прибором, после чего на обслуживание назначается следующая группа заявок другой очереди в соответствии с заданной дисциплиной обслуживания.

Комбинированный режим – комбинация одиночного и группового режимов, когда часть очередей заявок обрабатывается в одиночном режиме, а другая часть – в групповом.

Дисциплины обслуживания могут использовать следующие *правила выбора заявок на обслуживание*:

бесприоритетные

обслуживание в порядке поступления (FIFO – First In First Out), когда на обслуживание выбирается заявка, поступившая в систему раньше других;

обслуживание в обратном порядке (LIFO – Last In First Out), когда на обслуживание выбирается заявка, поступившая в систему позже других;

обслуживание в случайном порядке, когда на обслуживание заявка выбирается случайным образом;

обслуживание в циклическом порядке, когда на обслуживание заявки выбираются в процессе циклического опроса накопителей в последовательности 1, 2, ..., N (N – количество накопителей), после чего указанная последовательность повторяется;

приоритетные

с **относительными приоритетами** – приоритеты учитываются только в моменты завершения обслуживания заявок при выборе новой заявки на обслуживание и не влияют на процесс обслуживания низкоприоритетной заявки в приборе; другими словами, поступление заявки в СМО с более высоким приоритетом по сравнению с обслуживаемой в приборе не приводит к прерыванию процесса обслуживания этой заявки;

с **абсолютными приоритетами**, означающими, что, в отличие от относительных приоритетов, при поступлении высокоприоритетной заявки обслуживание заявки с низким приоритетом прерывается и на обслуживание принимается поступившая высокоприоритетная заявка; при этом прерванная заявка может быть возвращена в накопитель или

удалена из системы; если заявка возвращена в накопитель, то её дальнейшее обслуживание может быть продолжено с прерванного места или начато заново, то есть с самого начала;

со **смешанными приоритетами** – любой комбинацией без приоритетного обслуживания, относительных и абсолютных приоритетов;

с **чередующимися приоритетами** – аналогом относительных приоритетов, проявляется только в моменты завершения обслуживания группы заявок одной очереди и назначения новой группы;

обслуживание по расписанию, когда заявки разных классов (находящиеся в разных накопителях) выбираются на обслуживание в соответствии с некоторым расписанием (планом), задающим последовательность опроса очередей заявок, например, в случае трёх классов заявок (накопителей) расписание может иметь вид: {2, 1, 3, 3, 1, 2}.

Дисциплины FIFO, LIFO, относительных, абсолютных и смешанных приоритетов относятся к дисциплинам *одиночного режима*. Очевидно, что дисциплины *группового режима* обслуживания в циклическом порядке, с чередующимися приоритетами и обслуживание по расписанию, в частном случае могут быть реализованы как дисциплин обслуживания одиночного режима, если размер назначаемой на обслуживание группы равен 1, при этом дисциплина обслуживания с чередующимися приоритетами вырождается в дисциплину обслуживания с относительными приоритетами.

Среди представленных дисциплин обслуживания особое место занимают дисциплины со смешанными приоритетами, обладающие общностью по отношению к перечисленным дисциплинам обслуживания одиночного режима.

Для математического описания дисциплин обслуживания со смешанными приоритетами используется **матрица приоритетов**, представляющая собой квадратную матрицу: $Q = [q_{ij}]$, $i, j = \overline{1, N}$, N – число классов заявок, поступающих в систему.

Элемент q_{ij} матрицы задает приоритет заявок класса i по отношению к заявкам класса j и может принимать следующие значения:

- 0 – нет приоритета;
- 1 – приоритет относительный;
- 2 – приоритет абсолютный.

Элементы матрицы приоритетов должны удовлетворять следующим *требованиям*:

$q_{ii} = 0$, так как между заявками одного и того же класса не могут быть установлены приоритеты;

если $q_{ij} = 1$ или 2, то $q_{ji} = 0$, так как если заявки класса i имеют приоритет к заявкам класса j , то последние не могут иметь приоритет к заявкам класса i ($i, j = 1 \dots H$).

В зависимости от *возможности изменения приоритетов* в процессе функционирования системы приоритетные дисциплины буферизации и обслуживания делятся на два класса:

со статическими приоритетами, которые не изменяются со временем;

с динамическими приоритетами, которые могут изменяться в процессе функционирования системы в зависимости от разных факторов, например, при достижении некоторого критического значения длины очереди заявок какого-либо класса, обладающего низким приоритетом, ему может быть предоставлен более высокий приоритет.

2. Классификация моделей массового обслуживания

При моделировании реальных систем с дискретным характером функционирования широкое применение находят базовые модели в виде СМО, которые могут быть классифицированы (рис.2.5):

по числу мест в накопителе;

по числу обслуживающих приборов;

по количеству классов заявок, поступающих в СМО.

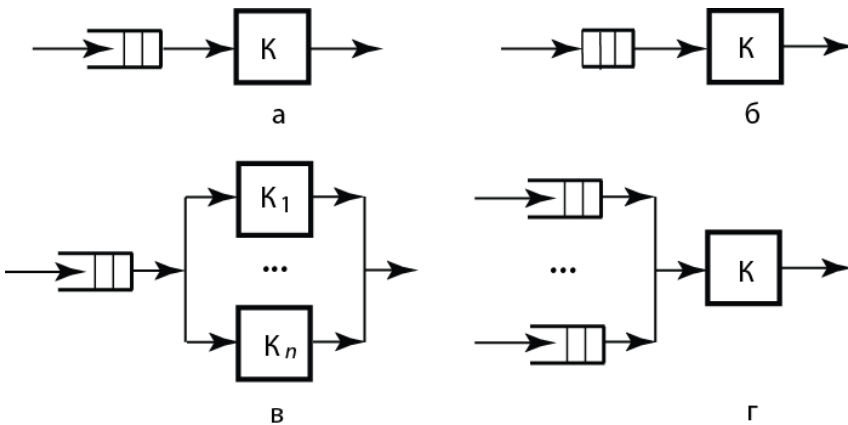


Рис. 2.5. Базовые модели СМО

1. По числу мест в накопителе СМО делятся на системы:

без накопителя (СМО с отказами), в которых заявка, поступившая в систему и заставшая все обслуживающие приборы занятыми обслуживанием более высокоприоритетных заявок, получает отказ и теряется;

с накопителем ограниченной ёмкости (СМО с потерями) (рис. 2.5б), в которых поступившая заявка теряется, если она застаёт накопитель заполненным полностью;

системы с накопителем неограниченной ёмкости (СМО без потерь) (рис. 2.5а, в, г), в которых для любой поступившей заявки всегда найдется место в накопителе для ожидания.

Как отмечалось ранее, предположение о неограниченной ёмкости накопителя может использоваться для моделирования реальных систем, в которых вероятность потери заявки из-за переполнения накопителя ограниченной ёмкости меньше 10^{-3} .

2. По количеству обслуживающих приборов СМО делятся на:

одноканальные (рис. 2.5а, б, г), содержащие один прибор (канал) K ;

многоканальные (рис. 2.5в), содержащие n обслуживающих приборов K_1, \dots, K_n , $n > 1$.

В многоканальных СМО обычно предполагается, что все приборы идентичны и равнодоступны для любой заявки, то есть при нескольких свободных приборах поступившая заявка с равной вероятностью может попасть в любой из них на обслуживание.

3. По количеству классов (типов) заявок, поступающих в СМО, различают системы:

с однородным потоком заявок (рис. 2.5а, б, в);

с неоднородным потоком заявок (рис. 2.5г).

Однородный поток заявок образуют заявки одного класса, а неоднородный поток – это поток заявок нескольких классов.

В СМО, представляющей собой абстрактную математическую модель, заявки относятся к разным классам в том случае, если они в моделируемой реальной системе различаются хотя бы одним из следующих факторов:

длительностью обслуживания;

приоритетами.

Если заявки не различаются длительностью обслуживания и приоритетами, они могут быть представлены как заявки одного класса.



Рис. 2.6. Вариант классификации моделей сетей массового обслуживания

В зависимости от структуры и свойств исследуемых систем их моделями могут служить СеМО различных классов. Одна из возможных классификаций сетевых моделей приведена на рис. 2.6.

1. В зависимости от *характера процессов поступления и обслуживания заявок* в сети СеМО делятся на:

стохастические, в которых процессы поступления и/или обслуживания заявок носят случайный характер, то есть интервалы времени между поступающими заявками и/или длительности их обслуживания в узлах представляют собой случайные величины, описываемые соответствующими законами распределений;

детерминированные, в которых интервалы времени между поступающими заявками и длительности их обслуживания в узлах являются детерминированными величинами.

2. По *виду зависимостей, связывающих интенсивности потоков заявок в разных узлах*, СеМО делятся на:

линейные, если эти зависимости линейные;

нелинейные, если эти зависимости являются нелинейными.

В *линейных* СеМО, как это следует из определения, интенсивность потока заявок в узел j связана с интенсивностью потока заявок в узел i линейной зависимостью: $\lambda_j = \alpha_{ij} \cdot \lambda_i$, где α_{ij} – коэффициент пропорциональности, показывающий, во сколько раз отличаются интенсивности потоков заявок в узел j и в узел i , $i, j = \overline{1, n}$.

Поскольку указанная зависимость справедлива для любой пары узлов, это выражение можно записать в несколько ином виде и выразить интенсивность поступления заявок во все узлы $j = \overline{1, n}$ через одну и ту же интенсивность, например, через интенсивность λ_0 потока заявок, поступающих в СеМО из источника заявок: $\lambda_j = \alpha_j \lambda_0$.

В последнем выражении коэффициент пропорциональности $\alpha_j \geq 0$ показывает, во сколько раз интенсивность потока заявок в узел j ($i, j = \overline{1, n}$) отличается от интенсивности источника заявок, и называется **коэффициентом передачи**. Коэффициент передачи может принимать любое положительное значение.

Коэффициент передачи играет важную роль при разработке математических зависимостей и расчёте характеристик функционирования сетевых моделей. Это обусловлено тем физическим смыслом, который несет в себе коэффициент передачи.

Коэффициент передачи можно трактовать как *среднее число попаданий заявки в данный узел за время её нахождения в сети*. Например, если коэффициент передачи узла СеМО равен 3, то это означает, что любая заявка за время нахождения в сети *в среднем* 3 раза побывает на обслуживании в данном узле. Значение коэффициента передачи, равное 0,25, означает, что *в среднем* только одна заявка из четырёх попадёт на обслуживание в данный узел, а три другие обойдут данный узел.

В *нелинейных* СеМО интенсивности потоков заявок в узлах связаны более сложными нелинейными зависимостями, что значительно усложняет их исследование.

Нелинейность СеМО может быть обусловлена:

потерей заявок в сети, например из-за ограниченной ёмкости накопителей в узлах;

размножением заявок в сети, заключающимся, например, в формировании нескольких новых заявок после завершения обслуживания некоторой заявки в одном из узлов сети.

Таким образом, СеМО является линейной, если в ней заявки не размножаются и не теряются. Далее рассматриваются, в основном, линейные СеМО.

3. По числу циркулирующих в сети заявок различают СеМО:
 разомкнутые;
 замкнутые;
 замкнуто-разомкнутые.

Разомкнутая (открытая) СеМО содержит один или несколько *внешних независимых источников* заявок, генерирующих заявки в сеть независимо от числа заявок, находящихся в сети (рис.2.7а). В разомкнутой СеМО одновременно может находиться *любое число заявок*, в том числе, и сколь угодно большое, то есть от 0 до бесконечности. С разомкнутой СеМО связана внешняя среда, из которой поступают заявки в сеть и в которую они возвращаются после обслуживания в сети. Внешняя среда в разомкнутой СеМО обозначается обычно как нулевой узел «0», и разомкнутая СеМО, в этом случае, изображается в виде рис. 2.7б.

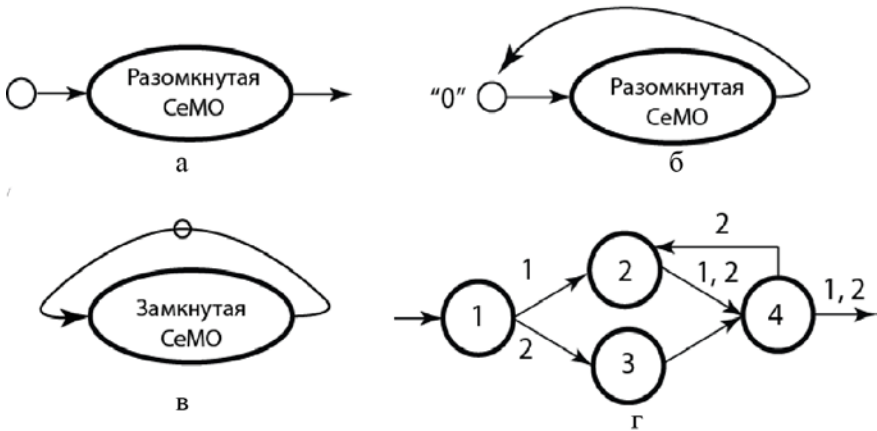


Рис. 2.7. Виды сетей массового обслуживания

Замкнутая (закрывающаяся) СеМО не содержит *независимых внешних источников* заявок и характеризуется тем, что в ней циркулирует *постоянное число заявок M* (рис. 2.7в). На графе замкнутой СеМО из физических соображений, связанных с конкретным представлением процесса функционирования исследуемой реальной системы, обычно выделяется особая дуга, отображающая процесс завершения обслуживания заявок в сети и мгновенного формирования новой заявки с та-

кими же параметрами обслуживания, что и завершившая обслуживание. Такая трактовка позволяет рассматривать завершившую обслуживание заявку как новую заявку, поступившую в сеть из *зависимого источника* заявок.

По аналогии с разомкнутой СеМО на выделенной дуге замкнутой СеМО отмечается условная точка «0», рассматриваемая как нулевой узел и трактуемая иногда как фиктивная СМО с нулевой длительностью обслуживания. Или как зависимый источник заявок, генерирующий заявки только в момент поступления некоторой заявки на его вход. Выделение нулевого узла в замкнутой СеМО преследует двоякую цель: во-первых, достигается однозначность в представлении и математическом описании разомкнутой СеМО и замкнутой СеМО; во-вторых, обеспечивается возможность определения временных характеристик замкнутой СеМО относительно выделенного узла «0». В частности, *время пребывания заявок в замкнутой СеМО* рассматривается как промежуток времени между двумя соседними моментами прохождения заявки через нулевой узел.

В качестве замкнутой СеМО можно рассматривать систему восстановления вооружения и техники объединения при условии отсутствия отправок для восстановления в другие системы.

Замкнуто-разомкнутая СеМО (комбинированная) представляет собой комбинацию замкнутой СеМО и разомкнутой СеМО, в которую, кроме постоянно циркулирующих в сети M^* заявок, из внешнего независимого источника поступают заявки такого же или другого класса, при этом суммарное число заявок в сети $M \geq M^*$.

Систему восстановления вооружения и техники можно считать замкнуто-разомкнутой при наличии отправок в другие системы.

4. По *типу циркулирующих заявок* различают СеМО:

однородные, в которых циркулирует один класс заявок (однородный поток заявок);

неоднородные, в которых циркулирует несколько классов заявок (неоднородный поток заявок), различающихся хотя бы одним из следующих факторов:

длительностями обслуживания в узлах;

приоритетами;

маршрутами.

Маршруты заявок разных классов задаются путем указания номеров классов заявок на соответствующих дугах сети (рис. 2.7г).

3. Параметры и характеристики систем массового обслуживания

Для описания СМО используются три группы параметров:

структурные;

нагрузочные;

функциональные параметры (параметры управления).

К *структурным параметрам* относятся:

количество обслуживающих приборов n , равное 1 для одноканальной СМО и $n > 1$ для многоканальной СМО;

количество m и ёмкости накопителей E_j ($j = \overline{1, m}$);

способ взаимосвязи накопителей с приборами (в случае многоканальных СМО), например, в виде матрицы связей.

Нагрузочные параметры СМО включают в себя:

количество поступающих в систему классов заявок H , которое равно 1 для СМО с однородным потоком заявок и $H > 1$ для СМО с неоднородным потоком;

закон распределения $F_i(t)$ интервалов времени между поступающими в систему заявками класса $i = \overline{1, H}$ или, по крайней мере, первые два момента распределения, задаваемые, например, в виде интенсивности поступления заявок λ_i и коэффициента вариации ν_i интервалов;

закон распределения $F_i(t)$ длительности обслуживания заявок класса $i = \overline{1, H}$ или, как минимум, первые два момента распределения, в качестве которых обычно используются средняя длительность $t_{i\text{обсл}}$ или интенсивность $\mu_i = 1/t_{i\text{обсл}}$ обслуживания и коэффициент вариации ν_i .

Задание двух первых моментов нагрузочных параметров зачастую оказывается достаточным для оценки характеристик обслуживания заявок на уровне средних значений. Отметим, что в имитационной модели для описания простейшего потока достаточно задать только интенсивность поступления заявок в систему или среднее время поступления заявок.

Функциональные параметры задаются в виде конкретных стратегий управления потоками заявок в СМО, определяющих правило занесения заявок разных классов в накопители ограниченной ёмкости

(дисциплина буферизации) и правило выбора их из очереди на обслуживание (дисциплина обслуживания).

Для компактного описания систем массового обслуживания часто используются обозначения, предложенные Д. Кендаллом, в виде:

$$A/B/N/L,$$

где **A** и **B** – задают законы распределений соответственно интервалов времени между моментами поступления заявок в систему и длительности обслуживания заявок в приборе;

N – число обслуживающих приборов в системе ($n = 1, 2, \dots$);

L – число мест в накопителе, которое может принимать значения $0, 1, 2, \dots$ (отсутствие **L** означает, что накопитель имеет неограниченную ёмкость).

Для задания законов распределений **A** и **B** используются следующие обозначения:

G (General) – произвольное распределение общего вида;

M (Markovian) – экспоненциальное (показательное) распределение;

D (Deterministik) – детерминированное распределение;

U (Uniform) – равномерное распределение;

Ek (Erlangian) – распределение Эрланга k -го порядка (с k последовательными одинаковыми экспоненциальными фазами);

hk (hipoexponential) – гипоэкспоненциальное распределение k -го порядка (с k последовательными разными экспоненциальными фазами);

hr (Hiperexponential) – гиперэкспоненциальное распределение порядка r (с r параллельными экспоненциальными фазами);

g (gamma) – гамма-распределение;

P (Pareto) – распределение Парето и т.д.

Примеры:

M/M/1 – одноканальная СМО с накопителем неограниченной ёмкости, в которую поступает однородный поток заявок с экспоненциальным распределением интервалов времени между последовательными заявками (простейший поток) и экспоненциальной длительностью обслуживания заявок в приборе.

M/G/3/10 – трёхканальная СМО с накопителем ограниченной ёмкости, равной 10, в которую поступает однородный поток заявок с экспоненциальным распределением интервалов времени между последовательными заявками (простейший поток) и длительностью обслуживания заявок, распределённой по закону общего вида.

D/E2/7/0 – семиканальная СМО без накопителя (ёмкость накопителя равна 0), в которую поступает однородный поток заявок с детерминированными интервалами времени между последовательными заявками (детерминированный поток) и длительностью обслуживания заявок в приборе, распределённой по закону Эрланга 2-го порядка.

Для обозначения более сложных СМО дополнительно могут использоваться обозначения, описывающие неоднородный поток заявок и приоритеты между заявками разных классов.

СМО может работать в следующих режимах:

установившемся или **стационарном**, когда вероятностные характеристики системы не изменяются со временем;

неустановившемся, когда характеристики системы изменяются со временем, что может быть обусловлено:

началом работы системы, когда значения характеристик функционирования, меняясь со временем, стремятся в пределе к стационарным значениям (**переходной режим**);

нестационарным характером потока заявок и обслуживания в приборе (**нестационарный режим**).

Кроме этого, в некоторых системах, например в *СМО с накопителем неограниченной ёмкости*, неустановившийся режим функционирования может быть обусловлен *перегрузкой системы*, когда интенсивность поступления заявок превышает интенсивность обслуживания, и система не справляется с возлагаемой на нее нагрузкой (**режим перегрузки**). При этом характеристики функционирования СМО с течением времени растут неограниченно. В частности, длина очереди перед прибором с течением времени становится всё больше и в пределе стремится к бесконечности.

Обычно исследование СМО с накопителем неограниченной ёмкости проводится в предположении о существовании установившегося режима, непременным условием которого является требование отсутствия перегрузок, для чего нужно, чтобы интенсивность поступления заявок была меньше, чем интенсивность обслуживания. Это требование записывается для одноканальных СМО в виде условия: $\lambda < \mu$.

Для многоканальных СМО аналогичное условие имеет вид: $\lambda < t\mu$, где t – число обслуживающих приборов, а значение $t\lambda$ представляет собой суммарную интенсивность обслуживания заявок в t -канальной СМО.

В СМО с накопителем ограниченной ёмкости превышение интенсивности поступления заявок над суммарной интенсивностью обслуживания не приводит к неограниченному росту длины очереди, что обусловлено потерей заявок. Значит, в СМО с накопителем ограниченной ёмкости перегрузки не приводят к работе системы в неустановившемся режиме, а приводят лишь к росту числа потерянных заявок. При этом потеря части поступающих заявок при наличии накопителя ограниченной ёмкости может рассматриваться как один из механизмов борьбы с перегрузками.

Характеристики СМО с однородным потоком заявок.

Характеристики систем со стохастическим характером функционирования являются случайными величинами и полностью описываются соответствующими законами распределений. На практике при моделировании часто ограничиваются определением только средних значений (математических ожиданий), реже – определением двух первых моментов этих характеристик.

В качестве основных характеристик СМО с однородным потоком заявок используются следующие величины:

нагрузка системы;

коэффициент простоя системы;

вероятность потери заявок;

вероятность обслуживания заявки, то есть вероятность того, что поступившая в систему заявка будет обслужена;

производительность системы, представляющая собой интенсивность потока обслуженных заявок, выходящих из системы;

интенсивность потока потерянных (не обслуженных) заявок из-за ограниченной ёмкости накопителя;

среднее время ожидания заявок в очереди;

среднее время пребывания заявок в системе, складывающееся из времени ожидания и времени обслуживания;

средняя длина очереди заявок;

среднее число заявок в системе (в очереди и на обслуживании в приборе).

Характеристики СМО с неоднородным потоком заявок.

Для СМО с неоднородным потоком заявок, в которую поступают N классов заявок с интенсивностями μ_1, \dots, μ_N и средними длительностями обслуживания $t_{1\text{обсл}}, t_{2\text{обсл}}, \dots, t_{N\text{обсл}}$, определяются две груп-

пы характеристик обслуживания заявок:

характеристики по каждому классу (поток) заявок;

характеристики объединённого (суммарного) потока заявок;

Характеристики по каждому классу заявок $i = \overline{1, N}$ идентичны характеристикам СМО с однородным потоком.

Характеристики объединённого (суммарного) потока заявок позволяют определить усредненные по всем классам заявок показатели эффективности функционирования СМО.

Параметры и характеристики СеМО.

Для описания *линейных разомкнутых и замкнутых однородных экспоненциальных СеМО* используется следующая совокупность параметров:

число узлов в сети;

число обслуживающих приборов в узлах сети;

матрица вероятностей передач, где p_{ij} – вероятность передачи за-

явки из узла i в узел j ;

интенсивность источника заявок, поступающих в разомкнутую СеМО, или число заявок M , циркулирующих в замкнутой СеМО;

средние длительности обслуживания заявок в узлах сети.

Заметим, что состав параметров разомкнутых и замкнутых СеМО различается только одним параметром, а именно: для замкнутых СеМО, в отличие от разомкнутых СеМО, вместо интенсивности поступления заявок в сеть необходимо задать число постоянно циркулирующих в сети заявок M .

В случае *не экспоненциальных разомкнутых СеМО* дополнительно необходимо задать законы распределения или, по крайней мере, вторые моменты интервалов времени между поступающими в разомкнутую сеть заявками и длительностей обслуживания заявок в узлах сети.

В случае *неоднородных СеМО* необходимо дополнительно задать количество классов заявок H в сети. Для каждого класса задать матрицы вероятностей передач $\mathbf{P}(h)$, интенсивности $\lambda(h)$ или число заявок $M(h)$. А также средние длительности обслуживания $\mu_i(h)$ заявок класса $h = \overline{1, H}$ в узле $i = \overline{1, n}$. При необходимости могут быть заданы законы распределений интервалов между поступающими в распределённую СеМО заявками и законы распределений длительностей обслуживания заявок разных классов в узлах сети.

Вопросы для самоконтроля

- 2.1. Что понимают под системой массового обслуживания?
- 2.2. Назовите элементы СМО.
- 2.3. Приведите вариант классификации СМО.
- 2.4. Показатели СМО с отказами.
- 2.5. Показатели СМО с ожиданием.
- 2.6. Насколько предположение о простейшем характере потока заявок соответствует реальности?
- 2.7. Когда оправдано использование предположения о простейшем характере потока заявок?
- 2.8. Какой поток заявок называется однородным? В каких случаях поток заявок в СМО является неоднородным?
- 2.9. Почему в СМО с накопителем неограниченной ёмкости, работающей без перегрузок, возникают очереди? В каких случаях они не возникают?
- 2.10. В каких случаях заявки в СМО относятся к разным классам?
- 2.11. Что в реальной системе может служить основанием для того, чтобы в соответствующей математической модели заявки были разделены на разные классы?
- 2.12. Как называется стационарный ординарный поток без последствия?
- 2.13. Когда поток заявок является стационарным? Привести примеры нестационарного потока заявок.
- 2.14. Какой поток заявок называется ординарным? Привести примеры неординарного потока заявок.
- 2.15. Что понимается под обслуживанием заявок в СМО? Что такое интенсивность обслуживания заявок в СМО, и какова её размерность?
- 2.16. Перечислить возможные дисциплины буферизации. В каких СМО не используются дисциплины буферизации?
- 2.17. Краткая характеристика приоритетных дисциплин обслуживания заявок.
- 2.18. Какими особенностями обладает простейший поток заявок?
- 2.19. Характеристики СМО с однородным потоком заявок.
- 2.20. Характеристики СМО с неоднородным потоком заявок.
- 2.21. Характеристики СеМО.
- 2.22. Что понимается под замкнутой СМО? Приведите примеры.
- 2.23. Приведите примеры краткого описания СМО.

Глава 3. Имитационное статистическое моделирование

Имитационная модель представляет собой программу, реализованную на компьютере, описывающую (моделирующую) функционирование элементов моделируемой системы, их связь между собой и внешней средой. Поэтому имитационную модель следует называть компьютерной моделью и имитационное моделирование – компьютерным моделированием, тем более, что название «имитационное моделирование» несёт в себе тавтологию: имитация и моделирование – очень сходные понятия. Тем не менее, термин «имитационное моделирование» прижился и широко используется в отечественной технической и научной литературе.

Имитационная модель является функциональной, так как она создается для получения характеристик моделируемого процесса, а не структуры. Однако, моделирующий алгоритм, как правило, имеет модульную структуру, аналогичную размещению и связям элементов в моделируемом объекте.

Имитационная модель даёт численное решение задачи, что не позволяет непосредственно усматривать функциональные связи между параметрами процесса, как это демонстрируют аналитические модели. Однако, выполнив серию экспериментов с моделью, направленно изменяя значения исследуемого фактора, и, выполнив обработку результатов, можно построить искомую связь между показателем эффективности системы и исследуемым фактором.

Имитационную модель, в отличие от аналитической модели, можно разработать с любой детализацией процесса или явления.

Как правило, имитационные модели создают для исследования процессов, на течение которых влияют различного рода случайности: отказы и сбои технических устройств, неточности измерений, рассеивание попаданий относительно точек прицеливания, и многое другое. Следовательно, результат такого процесса случаен. В имитационной модели случайные факторы моделируются при помощи специально подобранных генераторов случайных величин, которые входят в современные системы имитационного моделирования. Для получения характеристик таких вероятностных операций имитационная модель многократно реализуется на компьютере. Полученный при этом ряд значений исследуемого параметра подвергается статистической обра-

ботке, в результате которой и определяются характеристики случайных показателей процесса – математическое ожидание, дисперсия, закон распределения и т. п. Такие имитационные модели называют статистическими имитационными.

В этой теме излагаются сущность и основные аспекты имитационного моделирования: моделирования случайных величин, событий и процессов. Рассмотрены имитационные модели систем массового обслуживания, варианты модели противоборства двух сторон.

3.1. Сущность имитационного моделирования

Сущность имитационного моделирования рассмотрим на примере.

Пример 3.1. По объекту наносится одиночный ракетный удар. Радиус поражения R .

Попадание ракеты в цель характеризуется рассеиванием, распределенным по нормальному закону со среднеквадратическими отклонениями:

по дальности σ_x ;

по направлению σ_y .

Цель будет уничтожена, если расстояние r от нее (то есть от точки прицеливания) до центра взрыва ракеты будет меньше или равно R , то есть $r \leq R$. Так как $R \gg$ размеров объекта, то цель можно считать точечной.

Наличие рассеивания исключает однозначный ответ: «цель поражена – цель не поражена». Задача носит вероятностный характер, поэтому в результате моделирования может быть получен ответ: цель будет поражена с вероятностью P .

Цель моделирования: определить вероятность $P = P(r \leq R)$ поражения объекта одиночным ракетным ударом.

Решение

Построим декартову систему координат так, чтобы точечный объект находился в начале координат, а направление пуска ракеты совпадало с осью x (рис. 3.1).

Возьмем две последовательности нормально распределенных случайных чисел:

$$x_1, x_2, \dots, x_i, \dots, x_N;$$

$$y_1, y_2, \dots, y_i, \dots, y_N.$$

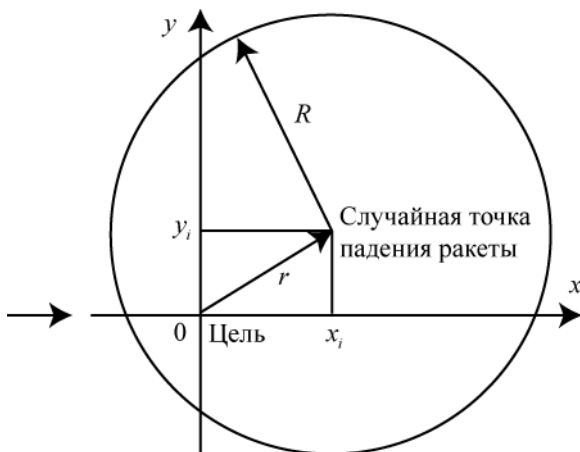


Рис. 3.1. Иллюстрация к нанесению удара

Первая последовательность соответствует распределению $M[x]=0, \sigma_x = a$, вторая – $M[y]=0, \sigma_y = b$. Математические ожидания $M[x]$, $M[y]$ взяты равными нулю, так как объект поражения (точка прицеливания) находится в начале координат, то есть имеет координаты $x=0$ и $y=0$.

Закон и характеристики случайных чисел x и y соответствуют закону рассеивания пуска ракет.

Моделирование

1. Имитируем удар, то есть мысленно нанесем удар по объекту путем определения координат взрыва. В силу идентичности закона рассеивания и его характеристик с законами распределения случайных чисел такими координатами могут быть x_1 и y_1 , взятые из последовательностей случайных чисел.

2. Вычислим расстояние r_1 от места взрыва ракеты до цели:

$$r_1 = \sqrt{x_1^2 + y_1^2}.$$

3. Оценим результаты имитации удара, то есть, установим факт поражения или непоражения объекта:

если $r_1 \leq R$, то объект поражен;

если $r_1 > R$, то объект не поражен.

4. Если объект поражен, запомним этот факт увеличением M на единицу, то есть $M = M + 1$ (в начале $M = 0$).

5. Для нахождения вероятности поражения объекта повторим имитацию нанесения удара N раз.

6. Оценим вероятность через частоту поражения объекта:

$$P(r \leq R) = \frac{M}{N}.$$

Возможность оценки вероятности частотой свершения события доказывается теоремой Я. Бернулли: при неограниченном числе однородных независимых опытов с практической достоверностью можно утверждать, что частота события будет сколь угодно мало отличаться от его вероятности в отдельном опыте (Бернулли Якоб 1 – самый старший из восьми представителей этой швейцарской семьи – выдающихся ученых).

Чем больше число N (число реализаций, число испытаний, число прогонов модели), тем точнее будет оценка вероятности P .

В рассмотренном примере 3.1 при $R = 2$ км, $\sigma_x = 1,5$ км, $\sigma_y = 0,8$ км оценки вероятностей P поражения цели при различном числе реализаций модели показаны в табл. 3.1.

Таблица 3.1

Оценки вероятностей поражения цели

| | | | | |
|-----|-----|------|--------|--------|
| N | 20 | 200 | 2000 | 10 000 |
| P | 0,8 | 0,75 | 0,7615 | 0,7644 |

При $N = 10\,000$, тех же характеристиках рассеивания и других радиусах поражения R получим:

$$R = 1 \rightarrow P = 0,33; \quad R = 1,5 \rightarrow P = 0,576; \quad R = 2,5 \rightarrow P = 0,8823.$$

В одной из последующих тем мы установим количественную связь между числом реализаций модели N , требуемой точностью и достоверной вероятностью результата моделирования, в данном случае оценки вероятности P .

Данный пример иллюстрирует *сущность метода имитационного моделирования*, который заключается в следующем.

1. Создается модель, поведение которой подчиняется тем же вероятностным законам, что и интересующий нас процесс.

2. По известным законам распределения для отдельных характеристик процесса выбираются их случайные значения.

3. Вычисляются параметры исхода процесса при случайных значениях характеристик, полученных на этапе 2, и запоминаются. Этапы 2 и 3 соответствуют одному статистическому испытанию.

4. В результате N статистических испытаний (повторений этапов 2 и 3) получают N значений параметров исхода процесса. Вероятностные характеристики параметров исхода процесса получают в результате статистической обработки полученных случайных величин.

Статистическая обработка и оценка точности результатов моделирования основываются на предельных теоремах теории вероятностей: теореме Чебышева и теореме Бернулли.

Рассмотрим ещё один пример.

Пример 3.2. Транспорт 1 с грузом отправился из пункта A в пункт C через пункт B . Одновременно из пункта D в пункт E через пункт B отправился транспорт 2. Скорости движения транспортных распределены по нормальному закону с математическими ожиданиями V_1 и V_2 и стандартными отклонениями σ_1 и σ_2 .

Построить алгоритм имитационной модели (ИМ) с целью определения вероятности встречи транспортных 1 и 2 в пункте B . Расстояние от пункта A до пункта B S_1 , а от пункта D до пункта B – S_2 . Событие встречи считать состоявшимся, если их времена прибытия в пункт B либо равны, либо отличаются на величину, не превышающую Δt .

Решение

Построим схему движения транспортных 1 и 2 (рис. 3.2).

Возьмем две последовательности нормально распределенных случайных чисел:

$$V_{11}, V_{12}, \dots, V_{1i}, \dots, V_{1N};$$

$$V_{21}, V_{22}, \dots, V_{2i}, \dots, V_{2N},$$

характеристики которых соответствуют математическим ожиданиям и стандартным отклонениям скоростей движения транспортных 1 и 2.

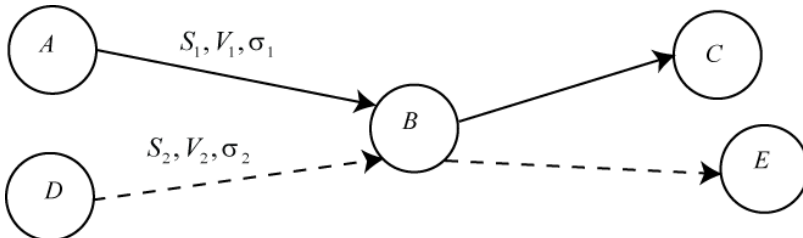


Рис. 3.2. Схема движения транспортных

1. Имитируем движение транспортов 1 и 2 до пункта B со скоростями V_{11} и V_{21} соответственно, взятыми из последовательностей нормально распределенных случайных чисел.

2. Вычислим время t_1 и t_2 прибытия в пункт B транспортов:

$$t_1 = \frac{S_1}{V_{11}}, \quad t_2 = \frac{S_2}{V_{21}}.$$

3. Оценим результат имитации движения транспортов 1 и 2, т. е. установим факт наличия или отсутствия их встречи:

если $|t_1 - t_2| \leq \Delta t$, встреча состоялась;

если $|t_1 - t_2| > \Delta t$, встреча не состоялась.

4. Если встреча состоялась, зафиксируем этот факт увеличением значения M на 1, т. е. $M = M + 1$ (вначале $M = 0$).

5. Для нахождения вероятности встречи транспортов 1 и 2 повторим имитацию их движения и встречи N раз.

Рассчитаем вероятность встречи:

$$P(\Delta t \leq |t_1 - t_2|) = \frac{M}{N}.$$

Результаты моделирования при $V_1 = V_2 = 25$ км/ч, $\sigma_1 = \sigma_2 = 3$ км/ч, $S_1 = S_2 = 50$ км/ч и $N = 10000$:

$$\Delta t = 0,2 \text{ ч} \rightarrow P = 0,0908;$$

$$\Delta t = 0,3 \text{ ч} \rightarrow P = 0,1548;$$

$$\Delta t = 0,4 \text{ ч} \rightarrow P = 0,2543;$$

$$\Delta t = 0,5 \text{ ч} \rightarrow P = 0,3779;$$

$$\Delta t = 0,6 \text{ ч} \rightarrow P = 0,5571.$$

Очевидно, изложенный процесс имитации легко может быть реализован на компьютере. Представим алгоритмы моделей примеров 3.1 и 3.2 схемами рис. 3.3 и 3.4 соответственно.

В рассмотренных примерах исследуются различные процессы. Но алгоритмы моделей этих процессов имеют общую, практически идентичную часть (блоки 1, 5...9, на рис. 3.3 и 3.4 они выделены) и часть, которая непосредственно имитирует исследуемый процесс (блоки 2...4).

Подобное сходство и различие ещё раз подтверждают сформулированную нами ранее сущность имитационного моделирования.

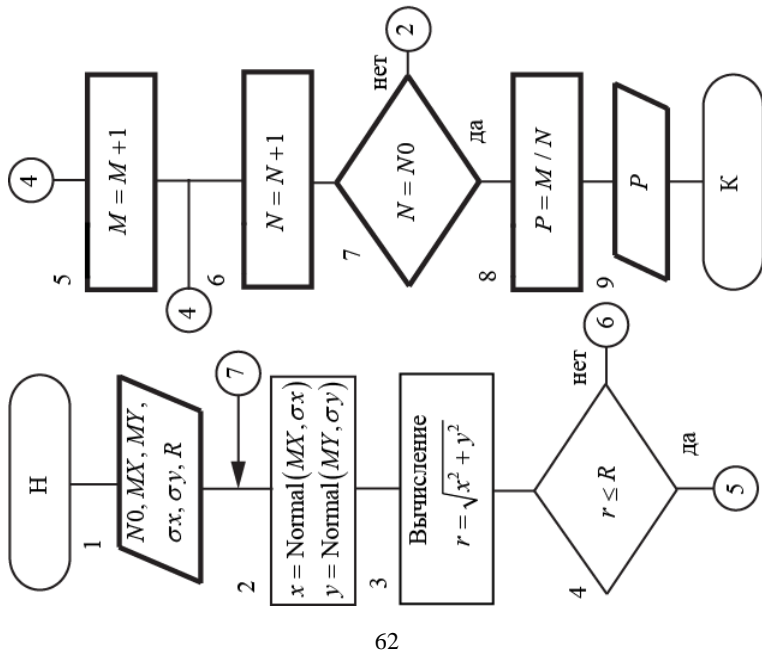


Рис. 3.3

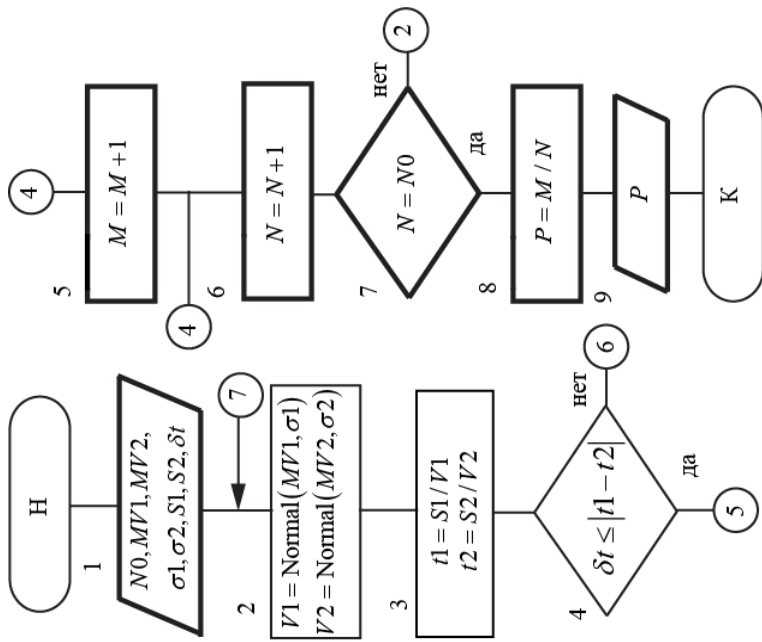


Рис. 3.4

Пример 3.3. По объекту наносится не одиночный удар, а три последовательных ракетных удара. При поражении объекта любой из трёх ракет пуски прекращаются. Остальные условия те же, что и в примере 3.1.

Алгоритм ИМ приведен на рис. 3.5. На нем выделены блоки 1, 8...11, выполняющие те же функции, что блоки 1, 5...9 в алгоритмах ИМ на рис. 3.3 и 3.4. Блоки 2...7 непосредственно имитируют нанесение удара по объекту, т. е. выполняют одну реализацию (один прогон модели).

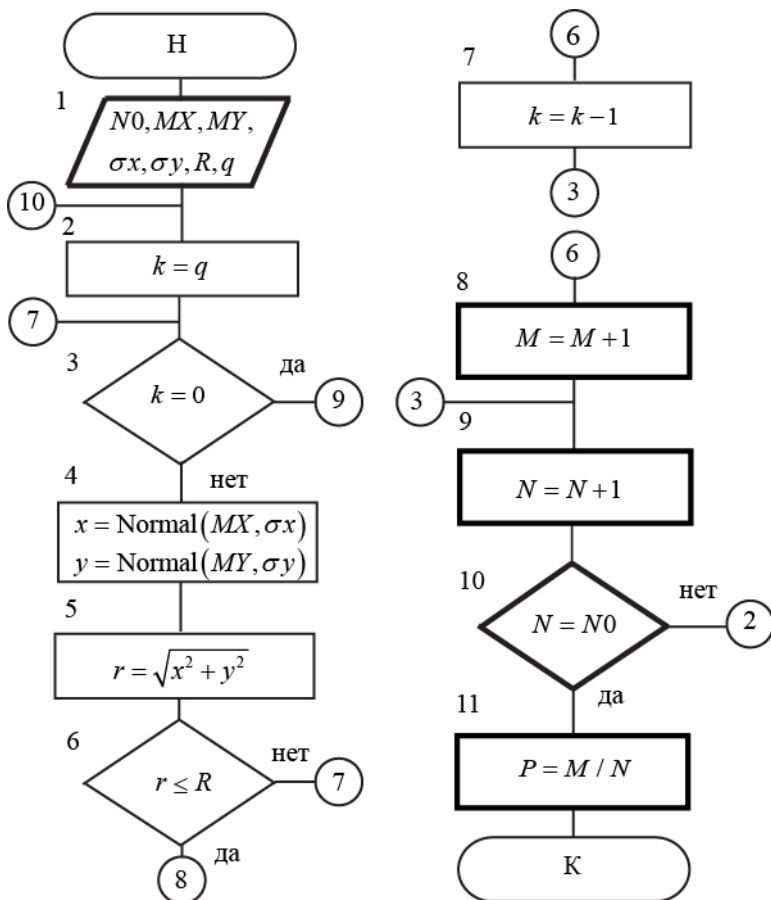


Рис. 3.5. Алгоритм модели нанесения удара тремя ракетами

В блоке 2 переменной k присваивается начальное число пусков ракет. Далее эта переменная используется для организации внутреннего цикла по числу пусков.

После каждого пуска значение k уменьшается на 1 (блок 7). При $k = 0$ (блок 3) реализация завершается. Завершается она также и при поражении объекта (блок 6). Но при этом предварительно значение переменной M увеличивается на 1. По завершении N_0 реализаций рассчитывается оценка математического ожидания вероятности поражения объекта тремя последовательными пусками ракет.

3.2. Общая характеристика метода имитационного моделирования

Когда явления в системе настолько сложны и многообразны, что аналитическая модель становится слишком грубым приближением к действительности, то исследователь вынужден использовать имитационное моделирование. По Р. Шеннону, имитация – это «процесс конструирования реальной системы и постановки эксперимента на ней».

ИМ – это метод исследования, заключающийся в имитации на ЭВМ (с помощью комплекса программ) процесса функционирования системы или отдельных ее частей и элементов. Сущность метода ИМ заключается в разработке таких алгоритмов и программ, которые имитируют поведение системы, её свойства и характеристики в необходимом для исследования системы составе, объёме и области изменения её параметров. При ИМ реализующий модель алгоритм воспроизводит процесс функционирования системы во времени, причём имитируются явления, составляющие процесс, с сохранением их логической структуры и последовательности протекания во времени (хронологическом порядке), что позволяет по исходным данным получить сведения о состояниях процесса в определенные моменты времени, дающие возможность оценить характеристики системы.

Имитационное моделирование позволяет осуществлять многократные испытания модели с нужными входными данными, чтобы определить их влияние на выходные критерии оценки работы системы. При таком моделировании компьютер используется для численной оценки модели, а с помощью полученных данных рассчитываются её реальные характеристики.

Имитационная модель отображает стохастический процесс смены дискретных состояний системы. При реализации модели на компью-

тере производится накопление статистических данных по показателям модели, которые являются предметом исследований.

По окончании моделирования накопленная статистика обрабатывается, и результаты моделирования получаются в виде выборочных распределений исследуемых величин. Таким образом, математическая статистика и теория вероятностей являются математическими основами имитационного моделирования.

Имитационные модели могут быть реализованы средствами универсальных языков программирования (Паскаль, Си++ и др.). Они предоставляют практически неограниченные возможности в разработке и отладке программ моделей. Однако, модель в виде программы на универсальном языке программирования часто непонятна исследователю. Ведь совершенно необязательно исследователь, специалист в конкретной предметной области должен знать тонкости программирования на каком-либо языке. Поэтому были созданы специализированные языки моделирования, которые существенно упрощают создание моделей и обработку результатов моделирования (Simula, Simgscript, Арена, семейство языков GPSS, AnyLogic и др.).

Имитационная модель может представить объект практически любой сложности. Ограничениями могут служить лишь недостаточная квалификация исполнителя, а также требование адекватности модели и достижения очень большой точности результата. А это связано с получением статистических выборок большого объема, что ведёт к необходимости получения большого числа реализаций модели и, следовательно, высокопроизводительных компьютеров.

Если сложность аналитической модели с усложнением моделируемого объекта возрастает с ускорением, как показано на рис. 3.6, то сложность имитационной модели, начиная с некоторого уровня S_0 , растёт незначительно.

Недостатком имитационного моделирования является то, что решение, результат является численным, частным, справедливым только для конкретных значений исходных данных. Чтобы получить функциональные зависимости между параметрами исследуемого процесса (системы) нужно сделать большое количество вариантов решений. Аналитическая же модель даёт, как правило, функциональные зависимости. Если сложность задачи, требуемая точность решения, возможности математики и способности исследователя позволяют построить аналитическую модель, то следует использовать её.

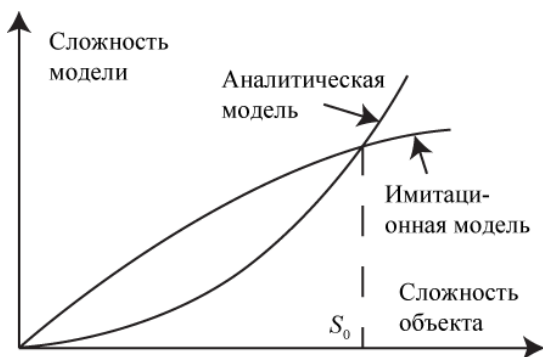


Рис. 3.6. Иллюстрация роста сложности моделей

Одной из основных целей имитационного моделирования является определение показателей эффективности различных операций, выполняемых при функционировании исследуемой системой.

Показатели эффективности могут выступать в виде оценок характеристик случайных величин или процессов или вероятностей исхода операций. В первом случае – это время, расход ресурсов, численности противоборствующих сторон, расстояния и т. п. Во втором случае показатель эффективности выступает в качестве вероятности, например, достижения цели операции в заданный срок, исправного состояния техники и т. д.

Имитационное моделирование может применяться в самых различных сферах деятельности. Ниже приведен перечень задач, при решении которых моделирование особенно эффективно:

- проектирование и анализ производственных систем;
- оценка различных систем вооружений и требований к их материально-техническому обеспечению;
- определение требований к сетям связи;
- определение требований к оборудованию и программному обеспечению различных компьютерных систем;
- проектирование и анализ работы транспортных систем, например: аэропортов, автомагистралей, портов и метрополитена;
- оценка проектов создания центров обработки заказов, заведений быстрого питания, больниц, отделений связи;
- определение политики в системах управления запасами;
- при подготовке специалистов и освоении новой техники на имитаторах (тренажерах).

Имитационные модели позволяют достаточно просто учитывать такие факторы, как наличие дискретных и непрерывных элементов, нелинейные характеристики элементов системы, многочисленные случайные воздействия и другие, которые часто создают трудности при аналитических исследованиях. В настоящее время имитационное моделирование – наиболее эффективный метод исследования больших систем, а часто и единственный практически доступный метод получения информации о поведении системы, особенно на этапе её проектирования. Несмотря на это, существуют обстоятельства, из-за которых могут получиться неадекватные результаты моделирования:

- нечёткая постановка задачи и цели моделирования;
- недостаточность или неполнота исходных данных;
- неверное определение источников и распределений случайных величин в реальных системах;
- недостаточный уровень проработки модели;
- недостаточные знания методологии моделирования;
- неподходящее программное обеспечение для моделирования;
- неправильное использование анимации;
- анализ выходных данных, полученных в результате недостаточного количества прогонов модели.

В настоящее время имитационное моделирование широко применяется в мире для исследования сложных систем. Этому способствуют преимущества, присущие этому методу, а именно:

1. Большинство сложных реальных систем с вероятностными параметрами нельзя точно описать математическими моделями.
2. Путём моделирования можно разработать ряд альтернативных вариантов моделей системы и затем определить, какой из них наиболее соответствует исходным требованиям.
3. Имитационное моделирование менее затратное, чем проведение экспериментов с реальными системами, тем более что иногда эксперименты на реальных системах в принципе невозможны.
4. Моделирование позволяет изучить длительный интервал функционирования системы в сжатые сроки или, наоборот, изучить более подробно работу системы в развернутый интервал времени.
5. При имитационном моделировании можно получать любое количество оценок вероятностной модели, проводя её прогоны.
6. Моделирование позволяет оценить некоторые эксплуатационные показатели системы при различных условиях эксплуатации.

Понятия «статистическое» и «имитационное моделирование» часто рассматривают как синонимы. Однако следует иметь в виду, что статистическое моделирование не обязательно является имитационным. Например, решение детерминированной задачи – вычисление определённого интеграла методом Монте-Карло путём определения подынтегральной площади на основе множества статистических испытаний, относится к статистическому моделированию, но не может ИМ.

Рассмотрим примеры вычисления определённых (собственных) интегралов.

3.3. Статистическое моделирование при решении детерминированных задач

Метод статистических испытаний может быть использован как численный метод решения и математических задач. Именно в таком качестве он был применен в США в 1944 г. Джоном фон Нейманом при расчётах по созданию ядерного реактора.

Применение метода рассмотрим на примере вычисления некоторого интеграла.

Пример 3.4. Пусть $0 \leq f(x) \leq 1, 0 \leq x \leq 1$. Полагаем, что функция $f(x)$ такова, что интеграл относится к «не берущимся».

Требуется вычислить $S = \int_0^1 f(x) dx$.

Решение

Представим функцию в координатах $f(x)$ и x как показано на рис. 3.7. Как известно, численное значение интеграла данного вида (определённого интеграла) равно площади S фигуры, ограниченной кривой, описываемой подынтегральной функцией $f(x)$. Площадь S состоит из множества элементарных площадок. Количество элементарных площадок (например, точек) в этой площади и будет численным значением искомого интеграла.

Имитируем координаты каждой точки значениями x_i и y_i , принадлежащими равномерному распределению на участке $[0, 1]$:

$$x_1, x_2, \dots, x_i, \dots, x_N;$$

$$y_1, y_2, \dots, y_i, \dots, y_N$$

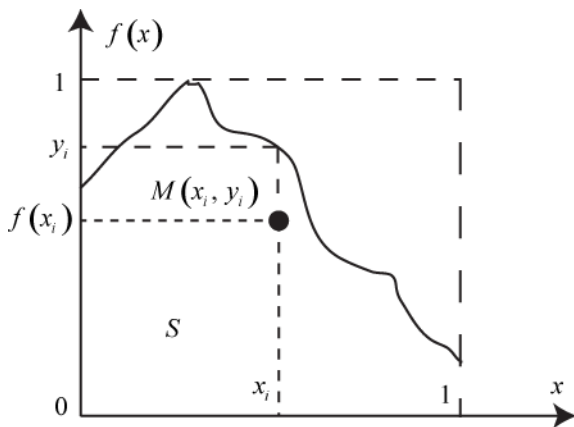


Рис. 3.7. Вычисление интеграла

Рассмотрим пару чисел x_i, y_i . Вычислим $f(x_i)$ и сравним с y_i . Если $f(x_i) \leq y_i$, то это означает, что точка $M(x_i, y_i)$ принадлежит площади S . Если $f(x_i) > y_i$, то это означает, что точка $M(x_i, y_i)$ не принадлежит площади S .

Введем:

$$z_i = \begin{cases} 1, & f(x_i) \leq y_i, \\ 0, & f(x_i) > y_i. \end{cases}$$

Число точек, попавших в границы S равно $\sum_{i=1}^N z_i$, где N – общее число точек, попавших в единичную площадь существования функции и аргумента. Отсюда следует:

$$\frac{\sum_{i=1}^N z_i}{N} \approx S = \int_0^1 f(x) dx.$$

Чем больше будет элементарных площадей – точек, тем точнее будет вычислен интеграл. Приведенное решение примера справедливо для единичных областей существования функции и аргумента. Однако это несущественно, так как произвольные границы существования $a \leq x \leq b, 0 \leq f(x) \leq A$ заменой переменных можно свести к единичным границам.

Известны статистические алгоритмы численного решения многократных интегралов.

Пример 3.5. Найти оценку I^* интеграла $I = \int_0^1 dx \int_x^1 (x + y) dy$.

Решение

Область интегрирования ограничена линиями $y = x, y = 1, x = 0$, т. е. принадлежит единичному квадрату (рис. 3.8).

Площадь области интегрирования (прямоугольного треугольника) $S = (1 \cdot 1) / 2 = 0,5$.

Используем формулу

$$I^* = S \cdot \frac{\sum_{i=1}^M f(x_i, y_i)}{M} = 0,5 \cdot \frac{\sum_{i=1}^M f(x_i, y_i)}{M},$$

в которой M – число случайных точек (x_i, y_i) , принадлежащих области интегрирования. У этих точек $y_i \geq x_i$. Если данное условие выполняется, то вычисляется

$$\sum_{i=1}^M f(x_i, y_i) = \sum_{i=1}^M f(x_i, y_i) + (x_i + y_i),$$

а число случайных точек M увеличивается на 1: $M = M + 1$.

Результаты моделирования приведены в табл. 3.2.

Из данных табл. 3.2 (верхние пять строк) видно, что с увеличением числа реализаций N ошибка $\Delta S = S - I^*$ в определении оценки интеграла I^* уменьшается и при $N = 1000000$ становится равной нулю.

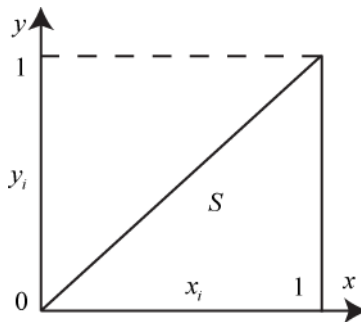


Рис. 3.8. Иллюстрация к примеру 3.5

Таблица 3.2

Результаты моделирования примера 3.5

| | | | | | |
|----------------------------|-------|---------|----------|-----------|------------|
| N | 10 | 1000 | 10000 | 100000 | 1000000 |
| M | 5 | 500 | 5038 | 49658 | 500364 |
| $\sum_{i=1}^M f(x_i, y_i)$ | 4,773 | 487,695 | 5006,152 | 49533,242 | 500191,650 |
| I^* | 0,477 | 0,488 | 0,497 | 0,499 | 0,500 |
| $\Delta S = S - I^*$ | 0,023 | 0,012 | 0,003 | 0,001 | 0 |
| M | 6 | 503 | 4935 | 49833 | |
| $\sum_{i=1}^M f(x_i, y_i)$ | 5,025 | 494,593 | 4917,236 | 49802,019 | |
| I^* | 0,419 | 0,492 | 0,498 | 0,500 | |
| $\Delta S = S - I^*$ | 0,081 | 0,008 | 0,002 | 0 | |

В четырех нижних строках табл. 3.2 приведены результаты моделирования с другими начальными числами генераторов равномерно распределенных случайных чисел. Видно, ошибка в оценке интеграла равна нулю уже при $N = 100000$ реализаций модели.

В заключение ещё раз отметим, что имитационное моделирование целесообразно применять в случаях:

- когда нет законченной математической постановки задачи;
- когда нет аналитических методов решения поставленной задачи;
- когда аналитические методы есть, но они не удовлетворяют требованиям точности и достоверности;
- когда аналитические методы есть, но их вычислительные процедуры сложны даже для компьютера;
- когда реализация известных процедур сталкивается с недостаточной математической подготовкой исследователя;
- когда исследователю нужно знать не только оценки искомых характеристик, но и динамику всего случайного процесса.

3.4. Моделирование равномерно распределенной случайной величины

Особое значение в статистическом моделировании имеет непрерывная равномерно распределенная случайная величина. Особая значимость этой случайной величины объясняется тем, что, во-первых, она сама по себе необходима для моделирования случайных процес-

сов и величин и, во-вторых, случайные величины с другими законами распределения формируются на её основе.

Определение. Непрерывная случайная величина γ имеет равномерное распределение в интервале $[a, b]$, если её плотность вероятности $f(x)$ определяется так (рис. 3.9).

Значения характеристик равномерного закона распределения:

математическое ожидание $M[\gamma] = \frac{a+b}{2}$;

дисперсия $D[\gamma] = \frac{(b-a)^2}{12}$.

При моделировании часто используются случайные числа из интервала $[0, 1]$. Непрерывная случайная величина γ равномерно распределена в интервале $[0, 1]$, если:

$$f(x) = \begin{cases} 1, & 0 \leq x \leq 1, \\ 0, & x < 0, x > 1. \end{cases}$$

В этом случае $M[\gamma] = \frac{1}{2}$, $D[\gamma] = \frac{1}{12}$.

Случайное число x_i из интервала $[0, 1]$ легко преобразуется в случайное число x'_i для интервала $[a, b]$:

$$x'_i = (b-a) \cdot x_i + a.$$

Применительно к двоичным дробям случайное число из интервала $[0, 1]$ представляет собой бесконечную дробь:

$$\gamma = z_1 \cdot 2^{-1} + z_2 \cdot 2^{-2} + \dots + z_i \cdot 2^{-i} + \dots; \quad i = \overline{1, \infty};$$

$$z_i = \begin{cases} 1, & p = 0,5, \\ 0, & p = 0,5. \end{cases}$$

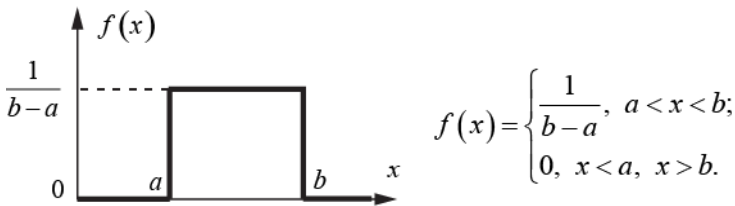


Рис. 3.9. Плотность вероятности равномерного распределения

Очевидно, реализовать такую дробь в компьютере невозможно, так как разрядная сетка компьютера ограничена. В компьютере можно формировать дискретные последовательности случайных чисел, которые не могут отличаться друг от друга только на величину меньше 2^{-n} (n – число разрядов в сетке компьютера). То есть непрерывного, «теоретического» распределения на компьютерах получить нельзя. Если эти числа равновероятны, то такое распределение случайных чисел называют *квазиравномерным*.

Заметим, что непрерывные случайные величины существуют только в теории. На практике все случайные величины дискретны и шаг дискретности равен наименьшей единице измерения.

Случайная величина ξ , имеющая квазиравномерное распределение в интервале $[0, 1]$, принимает значения

$$x_i = \frac{i}{2^n - 1}, \quad i = 0, 1, 2, \dots, 2^n - 1$$

с вероятностями $p_i = 0,5^n$.

Можно показать, что случайная величина ξ имеет характеристики:

$$M[\xi] = \frac{1}{2}, \quad D[\xi] = \frac{1 \cdot (2^n + 1)}{12 \cdot (2^n - 1)}.$$

Современные компьютеры имеют разрядность не менее 32. Следовательно, $M[\xi] = M[\gamma]$, а дисперсии тоже практически совпадают.

Для формирования случайных чисел в компьютере может использоваться один из двух основных способов:

- аппаратный (физический);
- алгоритмический (программный).

Аппаратный способ. При этом способе случайные числа формируются специальным устройством, использующим надёжные источники энтропии. Например, тепловой шум, шумы в электронных приборах, фотоэлектрический эффект, квантовые явления и т. д. Эти процессы в теории абсолютно непредсказуемы, на практике же получаемые из них случайные числа проверяются с помощью специальных статистических тестов.

- Преимущества такого способа:
- количество случайных чисел неограниченно;
- не требует затрат оперативной памяти;

требует малые вычислительные ресурсы компьютера.

Однако, такой датчик (генератор) случайных чисел имеет существенные недостатки, которые в настоящее время исключили его из инженерной практики:

трудность настройки;

необходимость периодической проверки формируемой последовательности на соответствие закону распределения;

обеспечение стабильности условий работы устройства – питания, влажности, температуры, старения приборов и элементов;

при необходимости невозможно повторить эксперимент при одной и той же последовательности случайных чисел.

Алгоритмический способ. При этом способе случайные числа формируются с помощью специальных программ.

Достоинства способа:

в настоящее время предлагается достаточно датчиков, генерирующих случайные числа, проверенных практикой и, следовательно, не нуждающихся в особых проверках;

можно многократно воспроизвести одну и ту же последовательность;

в памяти компьютера хранится только программа датчика (генератора), занимающая, как правило, малый объем;

алгоритмический датчик может быть реализован и аппаратно, за счёт чего существенно сокращается время формирования случайного числа и в целом время моделирования.

Недостатки:

на формирование случайного числа при программной реализации датчика требуются затраты машинного времени;

любой алгоритмический датчик может сгенерировать ограниченное количество неповторяющихся чисел.

В настоящее время практически везде применяются алгоритмические датчики случайных чисел. Создание высокопроизводительных компьютеров существенно снижает роль первого недостатка (затраты машинного времени). Второй недостаток устраняется использованием в одной модели нескольких датчиков случайных чисел (ДСЧ).

Алгоритмические датчики не обеспечивают получение теоретически «чистой» случайности чисел, так как их формирование идёт по формулам. Вследствие этого, рано или поздно последовательность случайных чисел станет повторяться или выродится. Последнее озна-

чает, что, начиная с некоторого числа, все последующие числа будут равны нулю.

Поэтому алгоритмические датчики называют датчиками *псевдослучайных* чисел, обладающими:

статистическими свойствами случайных чисел, определяемых путём их проверки специальными тестами;

периодичностью, то есть повторяемостью через определённые промежутки времени.

Качество алгоритмического датчика оценивается тем, насколько полно он удовлетворяет следующим требованиям:

закон распределения формируемых чисел должен быть равномерным (квазиравномерным);

числа должны быть статистически независимыми;

числа в последовательности не должны повторяться;

формирование чисел должно занимать минимальное машинное время и минимальный объем памяти.

В дальнейшем алгоритмический ДСЧ, выдающий детерминированную, псевдослучайную последовательность квазиравномерно распределённых случайных чисел, будем называть датчиком равномерно распределённых случайных чисел и обозначать $\gamma \sim \text{Rav}[0, 1]$.

Для формирования равномерно распределённых случайных чисел в интервале $[0, 1]$ могут использоваться следующие методы:

метод квадратов;

метод произведений;

мультипликативный конгруэнтный метод;

методы, представляющие модификации перечисленных методов.

Исторически первым является датчик, в котором был реализован «метод срединных квадратов». Он служит хорошей иллюстрацией принципа алгоритмического формирования равномерно распределённых случайных чисел. Сущность **метода квадратов** заключается в следующем:

подбирается начальное целое число, например, четырёхразрядное число 2152;

вычисляется квадрат: $2152^2 = 04 \mathbf{6311} 04$;

выделяется четыре средних разряда, которые рассматриваются как дробная часть первого случайного числа 0,6311;

вычисляется квадрат $6311^2 = 39 \mathbf{8287} 21$;

выделяется четыре средних разряда, которые рассматриваются как дробная часть второго случайного числа 0,8287;

вычисляется квадрат $8287^2 = 68\ 6743\ 69$;

третье случайное число 0,6743 и т. д.

Очевидно, что максимальная длина периода генератора, то есть максимальное количество неповторяющихся случайных чисел определяется количеством разрядов в дробной части. В нашем примере максимально возможная длина периода равна 9999 (от 0,0001 до 0,9999). Однако в действительности длина периода меньше максимально возможной и зависит от исходного целого числа. Неудачно выбранное значение исходного числа может привести к двум неприятностям: маленькой длине периода или даже к вырождению генератора, когда значения случайной величины начинают повторяться.

Метод произведений аналогичен методу квадратов. Отличие состоит в том, что перемножаются два n -разрядных целых числа, одно из которых, называемое *ядром* или *множителем*, не меняется, а второе, называемое *множимым*, формируется из n последних (правых) разрядов полученного $2n$ -разрядного числа – произведения ядра и множимого. Естественно, что вначале, как и в методе квадратов, необходимо грамотно выбрать исходные значения ядра и множителя.

В настоящее время очень широкое распространение в практике моделирования получили **конгруэнтные методы** формирования псевдослучайной последовательности.

Два целых числа a и b называются **конгруэнтными** (сравнимыми) по модулю M , где m – целое число, если разность $(a - b)$ делится на M без остатка, а числа a и b дают одинаковые остатки от деления на M .

Например, 2568 и 148 (по модулю 10), 1746 и 511 (по модулю 5), 6493 и 2221 (по модулю 2) и т. д.

Конгруэнтные методы описываются в виде рекуррентного соотношения следующего вида:

$$X_{i+1} = (\lambda X_i + \mu) \pmod{M}, \quad i = 0, 1, 2, \dots$$

где X_i, λ, μ, M – неотрицательные целые числа; X_0 – произвольное неотрицательное нечётное число; λ – множитель; μ – аддитивная константа; m – модуль.

Каждое новое значение X_{i+1} псевдослучайной последовательности представляет собой целочисленный остаток от деления на модуль M суммы произведения предыдущего значения X_i на множитель λ и

аддитивной константы μ . Далее псевдослучайное число x_{i+1} в интервале $[0, 1]$ получается путем деления целочисленного значения X_{i+1} на модуль M : $x_{i+1} = X_{i+1} / M$.

Описанный метод генерирования псевдослучайных чисел получил название **смешанного конгруэнтного метода**.

В некоторых случаях используется более простой метод генерирования псевдослучайных чисел, представляющий частный случай смешанного метода, когда $\mu = 0$, получивший название **мультипликативного конгруэнтного метода**. В этом случае рекуррентное соотношение имеет вид:

$$X_{i+1} = \lambda X_i \pmod{M}, i = 0, 1, \dots,$$

где X_0 – произвольное нечётное число, неотрицательное;

$\lambda = 8t \pm 3$, t – любое целое положительное число;

M – значение модуля. Для реализации на компьютере удобно $M = p^q$, где p – основание системы счисления (2 или 10), q – число разрядов в случайном числе.

В этом случае взятие числа по модулю сводится к выделению q младших разрядов произведения λX_i .

Алгоритм мультипликативного конгруэнтного метода:

1. Выбрать X_0 , например, $X_0 = 234567$.
2. Вычислить коэффициент λ . Пусть $t = 1$, тогда $\lambda = 8 \cdot 1 - 3 = 5$.
3. Выбрать модуль M . Пусть система счисления десятичная ($p = 10$), разрядность случайных чисел $q = 6$. Тогда $M = 10^6$.
4. Вычислить произведение λX_0 : $\lambda X_0 = 5 \cdot 234567 = 1172835$.
5. Найти остаток от деления по модулю M :
$$X_1 = \lambda X_0 \pmod{10^6} = 1172835 \pmod{10^6} = 172835.$$
6. Найти число x_1 последовательности случайных чисел из интервала $[0, 1]$: $x_1 = X_1 / 10^6 = 0,172835$.
7. Присвоить $X_0 = X_1$ и перейти к п. 4.

Достоверность и точность результатов имитационного статистического моделирования в значительной степени определяется качеством используемых в моделях алгоритмических датчиков псевдослучайных последовательностей.

Проверка алгоритмических датчиков равномерно распределенных псевдослучайных чисел предполагает формирование большой совокупности или, как говорят, представительной выборки случайных чисел и выполнение множества проверочных тестов.

Датчики равномерно распределенных псевдослучайных чисел проверяют на: периодичность, равномерность, случайность.

Проверка на периодичность требует обязательного определения длины периода. Чем больше длина периода, тем генератор более качественный.

Проверка на равномерность распределения. Генератор должен выдавать близкие значения к следующим значениям статистических параметров равномерного распределения:

математическое ожидание $M[\gamma] \approx 0,5$;

дисперсия $D[\gamma] \approx 0,0833$;

среднеквадратическое отклонение $\sigma \approx 0,2887$.

Частотный тест позволяет выявить: сколько чисел попало в интервал $(M[\gamma] - \sigma; M[\gamma] + \sigma)$, то есть $(0,5 - 0,2887; 0,5 + 0,2887)$ или, в конечном итоге, $(0,2113; 0,7887)$. Так как $0,7887 - 0,2113 = 0,5774$, заключаем, что в хорошем генераторе в этот интервал должно попадать около 57.7% из всех выпавших случайных чисел.

При **проверке на случайность** можно использовать совокупность тестов проверки: частот, пар, комбинаций, серий, корреляции.

Тест проверки частот предполагает разбиение диапазона распределения на несколько интервалов и подсчет количества (частот или вероятностей) попаданий случайных чисел в выделенные интервалы.

Тест проверки пар заключается в подсчете количества 1 для каждого разряда всей совокупности выработанных генератором двоичных случайных чисел. Число 1 во всех разрядах должно составлять примерно 50% от количества выработанных генератором чисел.

Тест проверки комбинаций сводится к подсчету 1 в случайных числах, количество которых в среднем должно составлять половину от количества разрядов.

Тест проверки серий заключается в подсчете количества различных длин последовательностей одинаковых значений (1 или 0).

Тест проверки корреляции заключается в определении коэффициента корреляции между последовательностями случайных чисел, вырабатываемых двумя разными генераторами.

3.5. Моделирование случайной величины с произвольным законом распределения

Методы формирования псевдослучайных чисел с заданным законом распределения основаны на использовании генераторов равномерно распределённых случайных величин. При этом наибольшее распространение получили следующие методы:

аналитический (метод обратной функции);

табличный;

метод композиций, основанный на функциональных особенностях генерируемых распределений.

Аналитический метод заключается в построении математической зависимости, связывающей значения случайной величины с заданным законом распределения со значениями случайной величины, распределённой равномерно в интервале $[0, 1]$.

В основе аналитического метода лежит, как правило, метод обратной функции, который позволяет при моделировании случайных величин учесть все их статистические свойства.

Метод основан на следующей теореме.

Теорема. Если непрерывная случайная величина Y имеет плотность распределения вероятностей $f(y)$, то распределение случайной величины

$$F(y) = \int_{-\infty}^y f(y) dy \quad (3.1)$$

равномерно в интервале $[0, 1]$, т. е.

$$F(y) = \gamma \sim \text{Rav}[0, 1].$$

По определению, $F(y)$ является функцией распределения случайной величины Y .

Теорема может быть проиллюстрирована графиками, представленными на рис. 3.10.

Обозначим: $x_i - i$ -е число из $\gamma \sim \text{Rav}[0, 1]$ равномерного распределения, $y_i - i$ -е случайное число из произвольного распределения.

Из (3.1) следует:

$$x_i = \int_{-\infty}^{y_i} f(y) dy .$$

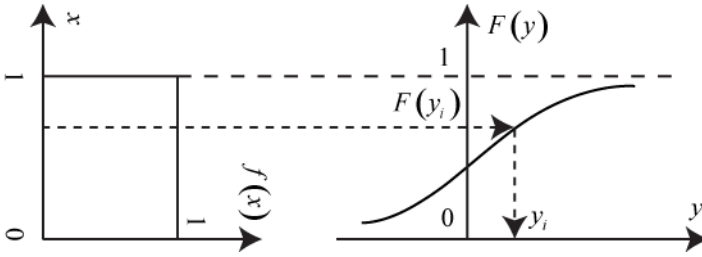


Рис. 3.10. Иллюстрация к методу обратной функции

Моделировать равномерно распределенное случайное число x_i мы уже умеем. Нужно найти неизвестное y_i , находящееся в верхнем пределе интегрирования.

Относительно y_i выражение принимает вид:

$$y_i = F^{-1}(x_i).$$

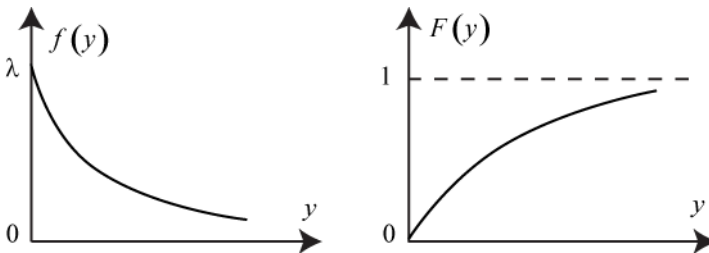
Отсюда и название – «метод обратной функции».

Пример 3.6. Получить формулу для моделирования случайных чисел, распределенных по экспоненциальному закону, с параметром λ (математическим ожиданием $1/\lambda$). Плотность $f(y)$ и функция $F(y)$ этого распределения имеют вид (рис. 3.11):

Решение

$$x_i = \int_{-\infty}^{y_i} f(y) dy = \int_{-\infty}^0 f(y) dy + \int_0^{y_i} f(y) dy = \int_0^{y_i} f(y) dy = 1 - e^{-\lambda y_i};$$

$$x_i \in \text{Rav}[0, 1]; \quad e^{-\lambda y_i} = 1 - x_i; \quad -\lambda y_i = \ln(1 - x_i); \quad y_i = -\frac{1}{\lambda}(1 - x_i).$$



$$f(y) = \begin{cases} 0 & \text{при } y \leq 0, \\ \lambda e^{-\lambda y} & \text{при } y > 0. \end{cases}$$

$$F(y) = \begin{cases} 0 & \text{при } y \leq 0, \\ 1 - e^{-\lambda y} & \text{при } y > 0. \end{cases}$$

Рис. 3.11. Плотность и функция экспоненциального распределения

Поскольку случайная величина $(1 - x_i)$ имеет равномерное распределение в интервале $[0, 1]$, как и x_i , то справедливо:

$$y_i = -\frac{1}{\lambda} \ln x_i.$$

Достоинства аналитического метода:

высокая точность метода;

не требуется составления и хранения в памяти таблиц, как в табличном методе.

Недостатки аналитического метода:

метод распространяется только на функции, которые позволяют вычислить интеграл от функции плотности аналитически;

использование численных методов вычисления не берущихся интегралов приводит к погрешностям и большим затратам машинного времени;

выражение, используемое для вычислений, содержит в себе функции вычисления логарифмов, возведения в степень, вычисления радикалов, что требует значительных затрат машинного времени.

Табличный метод. В современных системах моделирования применяется приближенный метод обратной функции, основанный на кусочно-линейной аппроксимации функции распределения моделируемой случайной величины.

Суть метода заключается в следующем.

Требуемый закон распределения случайной величины размещается в памяти компьютера в виде таблицы, содержащей координаты функции распределения. Каждая координата состоит из случайного числа y_i и соответствующего значения функции распределения $F(y_i)$:

$$y_1, F(y_1)/y_2, F(y_2)/\dots/y_k, F(y_k)/\dots/y_n, F(y_n).$$

Чем больше координат, тем точнее будет моделирование. Приемлемая точность обеспечивается заданием 20...30 координат.

При обращении за очередным случайным числом нужного закона распределения сначала генерируется случайное число из $\text{Rav}[0, 1]$.

Это число сравнивается со значениями $F(y_k)$, $k = \overline{1, n}$. При совпадении выдаётся соответствующее случайное число y_k .

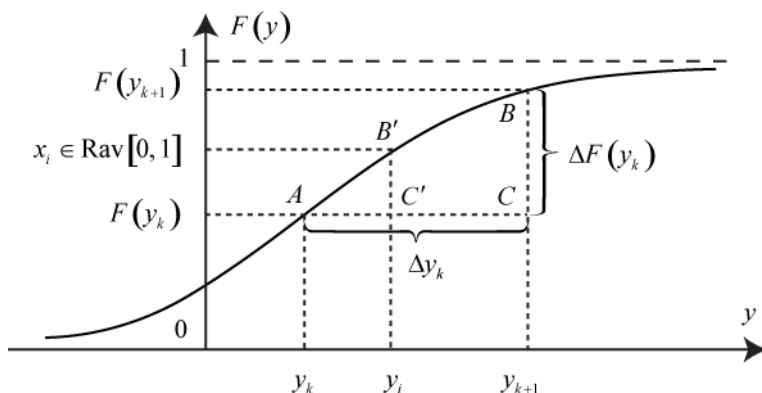


Рис. 3.12. Иллюстрация к методу кусочно-линейной аппроксимации

Если нет совпадения, то случайное число y_k вычисляется из подобия треугольников, как показано на рис. 3.12.

Из подобия треугольников ABC и $AB'C'$ следует:

$$\frac{\Delta F(y_k)}{\Delta y_k} = \frac{x_i - F(y_k)}{y_i - y_k}.$$

Отсюда по $x_i \in \text{Rav}[0, 1]$ находится значение y_i .

В ранних версиях GPSS использовались табличные генераторы для формирования случайных чисел, распределённых по экспоненциальному закону и нормальному закону, описываемые в виде функций, например, XpDis и SNorm соответственно.

```
XpDis    FUNCTION    RN247,C24
0,0/.1,.104/.2,.222/.3,.355/.4,.509/.5,.69/.6,.915/.7,1.2/.75,1.38/
.8,1.6/.84,1.83/.88,2.12/.9,2.3/.92,2.52/.94,2.81/.95,2.99/.96,3.2/
.97,3.5/.98,3.9/.99,4.6/.995,5.3/.998,6.2/.999,7/.9998,8
```

```
SNorm    FUNCTION    RN123,C25
0,-5/.00003,-4/.00135,-3/.00621,-2.5/.02275,-2/.06681,-1.5/
.11507,-1.2/.15866,-1/.21186,-.8/.27245,-.6/.34458,-.4/
.42074,-.2/.5,0/.57926,.2/.65542,.4/.72575,.6/.78814,.8/
.84134,1/.88493,1.2/.93319,1.5/.97725,2/.99379,2.5/
.99865,3/.99997,4/1,5
```

Представленные в XpDis значения функции $F(x)$ и x (24 координаты C24) соответствуют экспоненциальному распределению с мате-

математическим ожиданием, равным единице. Если математическое ожидание экспоненциально распределённой случайной величины отличается от 1, то полученное значение случайной величины умножается на значение математического ожидания.

Значения функции распределения $F(x)$ изменяются от 0 до 0,9998, а соответствующие им значения случайной величины x от 0 до 8. Отсюда видно, что получаемое от генератора случайное число может быть больше среднего значения (матожидания) в восемь раз.

В генераторе нормально распределённых случайных чисел значения случайной величины могут меняться только в интервале от -5 до $+5$, хотя теоретически могут лежать в пределах от $-\infty$ до $+\infty$ (25 координат – С25). Отсюда следует, что среднеквадратическое отклонение генерируемого нормального распределения должно быть, как минимум, в пять раз меньше математического ожидания. В противном случае от датчика может быть получено отрицательное число, что не имеет физического смысла.

Достоинства табличного метода:

существует принципиальная возможность построения таблицы для формирования случайных последовательностей с любым законом распределения, в том числе полученного экспериментальным путём;

можно обеспечить любую заданную точность генерирования случайных чисел за счет увеличения количества интервалов табуляции (уменьшения шага табуляции);

для генерирования случайных величин с заданным законом распределения вероятностей требуется только генератор равномерно распределённых случайных чисел и выполнение несложных операций, занимающих мало времени.

Недостатки табличного метода:

значительные затраты памяти для хранения большого числа таблиц с разными законами распределений;

наличие методической погрешности, обусловленной применением линейной интерполяции для определения значений случайных чисел, находящихся между узлами табуляции;

для уменьшения методической погрешности формирования случайных последовательностей при использовании линейной интерполяции следует увеличивать количество точек табуляции, что приводит к увеличению размера таблиц и, как следствие, к дополнительным затратам памяти и времени;

в связи с неодинаковой скоростью изменения функции распределения для обеспечения высокой точности формирования случайных последовательностей табулирование должно выполняться с переменным шагом, выбор которого связан с определёнными проблемами.

Значительную роль в моделировании играет случайная величина, имеющая нормальное распределение. Метод обратной функции в аналитическом виде здесь неприемлем, так как интеграл (3.1) не берущийся, а его численное решение громоздко.

Для генерации случайных чисел, подчинённых нормальному распределению, применяется метод обратной функции с кусочно-линейной аппроксимацией, как было показано выше, а также метод, основанный на центральной предельной теореме (ЦПТ) теории вероятностей.

Как известно, ЦПТ даёт теоретическое объяснение подтвержденному практикой наблюдению: если исход случайного события определяется большим числом случайных факторов, и влияние каждого фактора мало (ни один из факторов не имеет превалирующего значения), то такой случайный исход хорошо аппроксимируется нормальным распределением. Эта теорема имеет много формулировок. Одна из наиболее практичных для целей моделирования случайных последовательностей – теорема Леви-Линдберга.

Теорема. Случайная величина

$$\eta = \frac{\sum_{i=1}^N x_i - NM[x]}{\sqrt{ND[x]}}$$

где $\sum_{i=1}^N x_i$ – сумма N случайных чисел одного и того же распределения с математическим ожиданием $M[x]$ и дисперсией $D[x]$ при $N \rightarrow \infty$ асимптотически стремится к нормальному распределению с $M[\eta] = 0$ и дисперсией $D[\eta] = 1$.

Теорема не накладывает условий на вид распределения, из которого берутся случайные числа x_i . Поэтому удобно случайные числа x_i брать из рассмотренного датчика равномерно распределённых случайных чисел $\gamma \sim \text{Rav}[0, 1]$. В этом случае $M[\gamma] = \frac{1}{2}$, $D[\gamma] = \frac{1}{12}$.

Хорошее приближение к нормальному распределению получается уже при числе $N = 6$. Каждое случайное число при $N = 6$ генерируется так:

$$y_i = \frac{\sum_{i=1}^N x_i - 3}{\sqrt{0,5}} = \sqrt{2} \left(\sum_{i=1}^6 x_i - 3 \right).$$

Недостаток способа состоит в том, что он не экономичен, так как для генерирования одного случайного числа y_i требуется шесть случайных чисел из распределения $\gamma \sim \text{Rav}[0, 1]$.

В ряде случаев применяют ещё более неэкономичные датчики с числом $N = 12$. Тогда

$$y_i = \frac{\sum_{i=1}^{12} x_i - 6}{\sqrt{1}} = \sum_{i=1}^{12} x_i - 6.$$

Если датчик случайных чисел нормального распределения формирует стандартную последовательность чисел с $M = 0$, $\sigma = 1$, то последовательность нормального распределения с произвольными значениями его характеристик получается по формуле:

$$y'_i = y_i \sigma + m,$$

где m – требуемое значение математического ожидания;

σ – требуемое значение среднего квадратического отклонения;

y'_i – случайное число из нормального распределения с математическим ожиданием m и средним квадратическим отклонением σ .

При этом, ещё раз обратим ваше внимание на то, среднее квадратическое отклонение должно быть как минимум в пять раз меньше среднего значения.

В современных системах моделирования, как уже отмечалось, имеются встроенные датчики, позволяющие непосредственно получать нужную случайную величину с требуемыми значениями характеристик. Однако если исследователя эти возможности не удовлетворяют (например, по точности представления функции распределения вероятностей), то он может задать требуемый закон распределения самостоятельно.

3.6. Моделирование единичного события

Под *единичным событием* мы будем понимать смену состояний одного элемента (системы), причем состояний всего два: оборудование исправно – неисправно, канал СМО свободен – занят, цель поражена – не поражена и т. п.

Переход из одного состояния в другое – случайный. В любой момент времени система находится в одном состоянии с вероятностью P , в другом – с вероятностью $1 - P$.

Цель моделирования: имитировать состояние такого элемента.

Теорема. Пусть некоторое событие A свершается с вероятностью $p(A)$. Это может быть отказ техники, поступление сообщения, уничтожение цели, обнаружение цели и т. п.

Моделью свершения такого единичного события A является попадание значения x_i случайной величины γ , равномерно распределенной в интервале $[0, 1]$, в числовой интервал $[0, P(A)]$.

Доказательство

Как известно

$$P(a \leq x \leq b) = \int_a^b f(x) dx.$$

Для $x_i \in \gamma \sim \text{Rav}[0, 1]$, $a = 0$, $b = P(A)$

$$P(0 \leq x_i \leq P(A)) = \int_0^{P(A)} f(x) dx = \int_0^{P(A)} 1 \cdot dx = P(A),$$

так как $f(x) = 1$ для $\gamma \sim \text{Rav}[0, 1]$ на интервале $[0, P(A)]$.

Пример 3.7. Пусть вероятность состояния элемента $P(A) = 0,9$. В i -ой реализации случайное число $x_i \in \gamma \sim \text{Rav}[0, 1]$ после розыгрыша равно 0,955. Это означает, что в i -ой реализации модели событие A не свершилось (рис. 3.13). Естественно, одна реализация ни о чём не говорит. Реальная ситуация будет отображена на множестве реализаций и чем их больше, тем точнее.

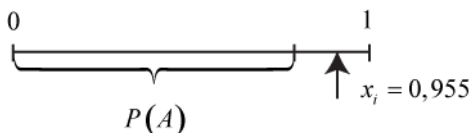


Рис. 3.13. Событие A не произошло

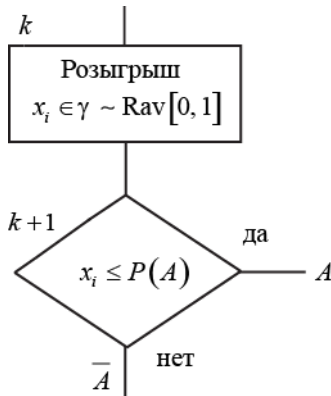


Рис. 3.14. Фрагмент алгоритма имитации единичного события

Фрагмент алгоритма имитации в модели единичного случайного события приведен на рис. 3.14.

3.7. Моделирование полной группы несовместных событий

Элемент системы (или система в целом) может находиться во многих (больше двух) несовместных состояниях. Известны вероятности нахождения системы в этих состояниях. Например, оружие может находиться:

в боеготовом состоянии с вероятностью P_1 ;

в неисправном состоянии и ремонтироваться силами своего расчета, вероятность этого состояния P_2 ;

ремонтироваться в мастерской части – P_3 ;

ремонтироваться на заводе – P_4 ;

Очевидно, что $P_1 + P_2 + P_3 + P_4 = 1$.

Такие и аналогичные события называются *полной группой несовместных событий*.

Алгоритм моделирования основан на следующей теореме.

Теорема. В полной группе несовместных событий моделью свершения события A_m , происходящего с вероятностью P_m , является попадание значения $x_i \in \gamma \sim \text{Rav}[0, 1]$ в отрезок, равный P_m , числовой

шкалы $\sum_{m=1}^n P_m = 1$, где n – число несовместных событий (рис. 3.15):

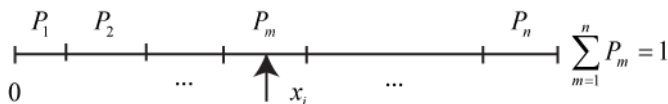


Рис. 3.15. Событие A_m произошло

Доказательство.

Введём численные обозначения концов отрезков P_m по нарастающей:

$$l_r = \sum_{m=1}^r P_m .$$

В этом случае, согласно теореме, условием свершения события A_m является попадание случайного числа x_i в интервал:

$$l_{m-1} < x_i \leq l_m .$$

Следовательно

$$\begin{aligned} P(l_{m-1} < x_i \leq l_m) &= \int_{l_{m-1}}^{l_m} \gamma(x) dx = \int_{l_{m-1}}^{l_m} 1 \cdot dx = \\ &= \left(l_m = \sum_1^m P_m \right) - \left(l_{m-1} = \sum_1^{m-1} P_m \right) = P_m . \end{aligned}$$

Такой способ моделирования полной группы несовместных событий обычно называют *определением исходов по жребию*. То есть моделью полной группы несовместных событий является эксперимент с m исходами.

Отметим, что последовательность точек l_1, l_2, \dots, l_r в данном случае есть не что иное, как последовательность значений выборочной функции распределения.

Алгоритм, реализующий способ определения исходов по жребию, может быть построен тремя вариантами, представленными на рис. 3.16.

Первый вариант (рис. 3.16а) применяется тогда, когда число возможных исходов невелико и не равно степени по основанию два.

На рис. 3.16б алгоритм построен по способу половинных сечений для четырех исходов.

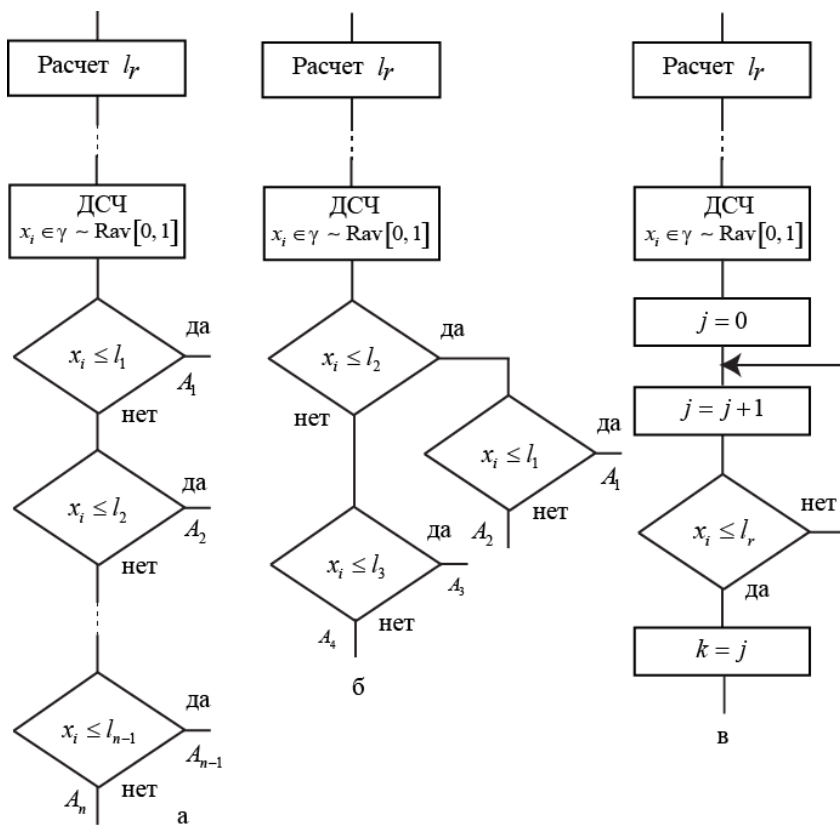


Рис. 3.16. Варианты алгоритма определения исходов по жребью

Третий вариант алгоритма (рис. 3.16в) в цикле определяет исход (событие), номер которого присваивается переменной k . Далее этот номер используется для организации нужной работы алгоритма. Этот вариант наиболее часто используется при построении алгоритмов имитационных моделей.

Пример 3.8. Канал передачи данных может находиться в одном из четырёх несовместных состояний:

A_1 – исправен и свободен, $P_1 = 0,15$;

A_2 – исправен и занят, $P_2 = 0,4$;

A_3 – неисправен, $P_3 = 0,25$;

A_4 – подавлен помехами, $P_4 = 0,2$.

Решение

Представим необходимые для определения исходов по жребью данные табл. 3.3.

Таблица 3.3

Данные для определения исходов по жребью

| Вероятности | Событие | | | |
|---------------------------------|---------|-------|-------|-------|
| | A_1 | A_2 | A_3 | A_4 |
| Вероятности событий | 0,15 | 0,4 | 0,25 | 0,2 |
| Суммарные вероятности (l_r) | 0,15 | 0,55 | 0,8 | 1,0 |
| Номера интервалов (r) | 1 | 2 | 3 | 4 |

Предположим, что при выполнении i -ой реализации датчик равномерно распределенных случайных чисел $\gamma \sim \text{Rav}[0, 1]$ сгенерировал $x = 0,525$. Путем последовательных сравнений определяется, что $l_1 < 0,525 \leq l_2$. Значит в данной реализации канал находится в состоянии A_2 – исправен и занят.

Пример 3.9. По каналу передачи данных (КПД) передаются сообщения трёх видов S_1, S_2, S_3 . Вероятности поступления сообщений соответствующих видов показаны в табл. 3.4.

Таблица 3.4

| S_i | S_1 | S_2 | S_3 |
|-------|-------|-------|-------|
| P_i | 0,24 | 0,36 | 0,4 |

КПД может находиться в одном из четырёх A_1, A_2, A_3, A_4 несовместных состояний (A_1 – КПД исправен и свободен, A_2 – КПД исправен и занят, A_3 – КПД неисправен, A_4 – КПД подавлен помехами) с вероятностями, показанными в табл. 3.5.

Таблица 3.5

| A_j | A_1 | A_2 | A_3 | A_4 |
|-------|-------|-------|-------|-------|
| P_j | 0,4 | 0,15 | 0,25 | 0,2 |

Построить алгоритм ИМ оценки вероятности передачи сообщений третьего вида S_3 при поступлении на КПД S_0 сообщений всех видов.

Решение.

На КПД поступают сообщения трёх видов. Поступление сообщения – это событие. Данная группа сообщений это группа несовместных событий. Полная группа несовместных событий моделируется определением исходов по жребию.

Предварительно нужно по данным табл. 3.4 построить числовую шкалу l_k вероятностей $\sum_{i=1}^3 P_i = 1$ (табл. 3.6). Свершением события – поступлением сообщения S_i будет попадание числа, полученного от датчика $x_i \in \gamma \sim \text{Rav}[0, 1]$ равномерно распределённых случайных чисел в отрезок числовой шкалы, равной P_i .

Таблица 3.6

| | | | |
|-------|-------|-------|-------|
| S_i | S_1 | S_2 | S_3 |
| P_i | 0,24 | 0,36 | 0,4 |
| l_k | 0,24 | 0,60 | 1,00 |
| k | 1 | 2 | 3 |

Состояния КПД по условиям постановки задачи – полная группа несовместных событий. Для моделирования состояний КПД предварительно также нужно построить числовую шкалу q_m вероятностей (табл. 3.7) по данным табл. 3.5.

Таблица 3.7

| | | | | |
|-------|-------|-------|-------|-------|
| A_j | A_1 | A_2 | A_3 | A_5 |
| P_j | 0,40 | 0,15 | 0,25 | 0,20 |
| q_m | 0,40 | 0,55 | 0,80 | 1,00 |
| m | 1 | 2 | 3 | 4 |

Сообщения по КПД не будут передаваться при состояниях $S_2 \dots S_3$.

Для работы ИМ необходимо ввести:

данные числовой шкалы l_i массивом $L = \{0, 24; 0, 6; 1, 0\}$;

данные числовой шкалы q_j массивом $Q = \{0, 4; 0, 55; 0, 8; 1, 0\}$;

$N0$ – количество прогонов модели;

$S0$ – количество всего поступающих сообщений всех трёх видов.

Для проведения промежуточных расчётов и вывода результатов моделирования потребуются переменные:

N – счётчик текущего числа прогонов модели;

S – счётчик текущего числа поступивших сообщений всех видов;

x – число – результат обращения к ДРРСЧ при определении вида поступившего сообщения;

y – число – результат обращения к ДРРСЧ при определении состояния КПД;

z – номер вида поступившего на КПД сообщения;

h – номер состояния КПД;

$S3$ – количество переданных сообщений третьего вида во всех прогонах модели;

$P3$ – вероятность передачи сообщений третьего вида.

Вероятность передачи сообщений третьего вида $P3$ можно определять двумя способами:

1) количество переданных сообщений третьего вида $S3$ во всех прогонах модели разделить на число поступивших на канал передачи данных сообщений третьего вида во всех прогонах модели;

2) число переданных сообщений третьего вида $S3$ во всех прогонах модели разделить на число поступивших на канал передачи данных сообщений всех трёх видов во всех прогонах: $P3 = S3 / (S0 * N0)$.

В первом случае мы определим вероятность передачи сообщений третьего вида из числа их поступивших. Во втором случае – вероятность передачи сообщений третьего вида из числа поступивших сообщений всех видов. Здесь не нужно будет считать количество поступивших на канал передачи данных сообщений третьего вида во всех прогонах модели. Воспользуемся вторым способом.

Условием завершения одного прогона модели является выполнение равенства $S = S0$, то есть равенства текущего числа поступивших на КПД сообщений заданному их числу.

Алгоритм ИМ приведен на рис. 3.17.

Блок 1 предназначен для ввода исходных данных.

Блоки 2...18 предназначены для моделирования непосредственно передачи сообщений в зависимости от их вида и состояния КПД.

Блоки 4...8 предназначены для определения (розыгрыша) вида поступившего сообщения. Розыгрыш производится с использованием данных числовой шкалы l_i . Код вида сообщения заносится в переменную z (блок 8).

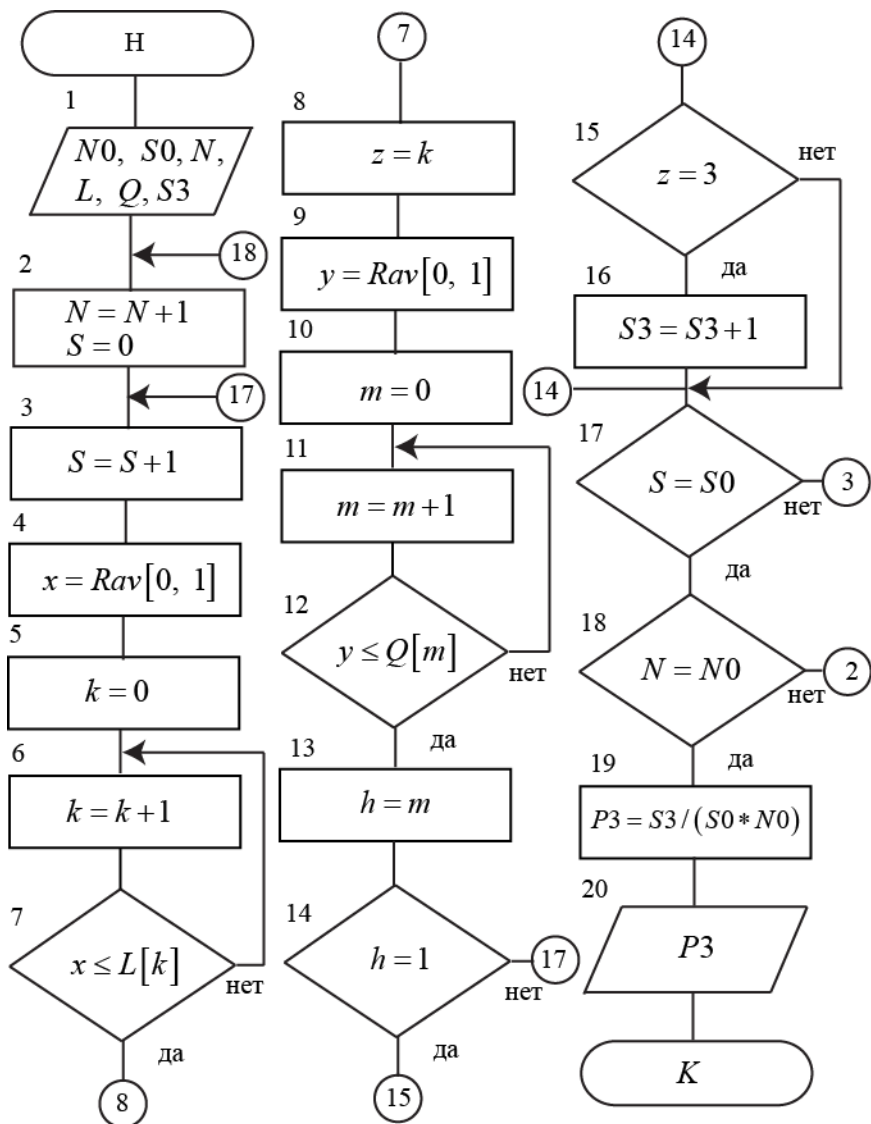


Рис. 3.17. Алгоритм имитационной модели передачи сообщений

Блоки 9...13 предназначены для определения (розыгрыша) состояния КПД. Розыгрыш производится с использованием данных числовой шкалы q_j . Код состояния заносится в переменную h (блок 13).

В блоке 14 проверяется состояние КПД. Если он исправен и не занят, т.е. код $h = 1$, то в следующем блоке 15 проверяется вид сообщения. Если $z = 3$, то сообщение третьего вида передаётся по КПД, что фиксируется в блоке 16, т.е. $S3$ увеличивается на 1. Сообщения других видов не распознаются и не считаются, хотя при необходимости реализовать распознавание и счёт можно аналогично.

В блоке 17 проверяется: все ли $S0$ сообщений поступили? Если нет, то условие $S = S0$ не выполняется и имитация передачи очередного сообщения, начиная с блока 3, повторяется.

При выполнении условия $S = S0$ в блоке 18 проверяется условие $N = N0$. Если оно не выполняется, то, начиная с блока 2, выполняется очередной прогон модели.

Если условие $N = N0$ выполняется, значит, выполнилось заданное $N0$ число прогонов модели.

В блоке 19 рассчитывается искомая вероятность, а в блоке 20 она выводится на носитель.

3.8. Моделирование совместных независимых событий

Рассмотрим моделирование совместных независимых событий.

Способ моделирования состоит в том, что совместные независимые события сводятся к одному сложному событию.

Для лучшего понимания и обозримости способа рассмотрим моделирование двух событий A и B . Увеличение числа событий ничего принципиально нового в моделирование не вносит.

Пусть независимые события A и B происходят с вероятностями $P(A)$ и $P(B)$ соответственно. Например, это могут быть попадания в цель двух независимо ведущих по ней огонь орудий.

Моделирование такой ситуации может быть выполнено двумя способами:

- определение совместных исходов выбором по жребью;
- последовательная проверка исходов.

3.8.1. Определение совместных исходов по жребью

Прежде всего, по вероятностям $P(A)$ и $P(B)$ нужно определить вероятности возможных исходов, т. е. появления совместных независимых событий. Возможные исходы совместного события Q_i и соответствующие вероятности P_i представлены в табл. 3.8.

Возможные исходы совместного события

| Q_i | AB | \overline{AB} | $A\overline{B}$ | $\overline{A}\overline{B}$ |
|-------|------------------|----------------------------|----------------------------|----------------------------|
| P_i | $P(A)P(B)$ | $[1-P(A)]P(B)$ | $P(A)[1-P(B)]$ | $1-P_i$ |
| l_r | $l_1 = P(A)P(B)$ | $l_2 = l_1 + [1-P(A)]P(B)$ | $l_3 = l_2 + P(A)[1-P(B)]$ | 1 |

Совместное событие в i -ой реализации определяется выбором исхода по жребию.

Если случайное число $x_i \in \gamma \sim \text{Rav}[0, 1]$ при очередной реализации окажется, например, на участке $l_1 < x_i \leq l_2$, то в данной реализации фиксируется свершение сложного события \overline{AB} . Если же окажется $x_i > l_3$, то фиксируется событие $\overline{A}\overline{B}$. Алгоритм может быть построен по одному из приведенных на рис. 3.16 вариантов.

3.8.2. Последовательная проверка исходов

Алгоритм способа последовательной проверки исходов приведен на рис. 3.18.

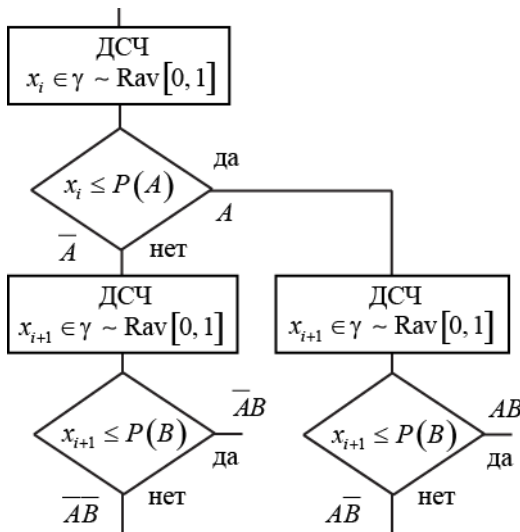


Рис. 3.18. Алгоритм последовательной проверки исходов

Проверку свершения каждого из совместных событий надо осуществлять разными случайными числами, так как события независимые. При первом способе достаточно одного случайного числа x_i , но сравнений может быть больше. Кроме того, нужно предварительно рассчитывать вероятности возможных исходов.

3.9. Моделирование совместных зависимых событий

Пусть события A и B имеют вероятности свершения $P(A)$ и $P(B)$ соответственно. Условная вероятность $P(B/A)$ известна.

Покажем способ моделирования совместных зависимых событий на примере.

Пример 3.10. При испытании нового автомата определены вероятности горизонтального и вертикального отклонений пробойн от точки прицеливания $P(A) = P(r_1 \leq 10 \text{ см})$, $P(B) = P(r_2 \leq 10 \text{ см})$.

Вероятность отклонения пробойн по высоте относительно тех, которые уложились в пределы допустимого бокового отклонения, равна:

$$P(B/\bar{A}) = \frac{P(B) - P(A)P(B/A)}{1 - P(A)}.$$

Соответствующий фрагмент модели приведен на рис. 3.19.

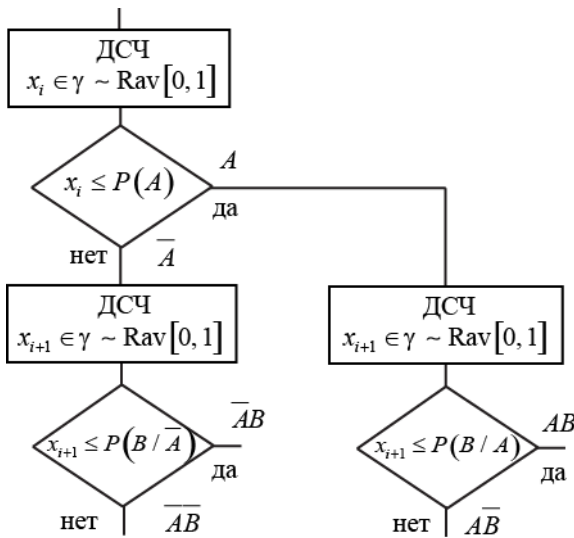


Рис. 3.19. Алгоритм моделирования совместных зависимых событий

Пример 3.11. В ремонтное подразделение поступают вышедшие из строя средства связи (СС). В каждом СС могут быть неисправными в любом сочетании блоки A, B, C . Вероятности выхода из строя блоков P_A, P_B, P_C соответственно. Ремонт производится путем замены неисправных блоков исправными блоками. В момент поступления неисправного СС вероятности наличия исправных блоков P_{HA}, P_{HB}, P_{HC} соответственно. При отсутствии хотя бы одного из исправных блоков A, B, C ремонт неисправного СС не производится.

Построить алгоритм имитационной модели с целью определения абсолютного и относительного количества отремонтированных СС с неисправными блоками A, B, C и A, B из общего количества R поступивших в ремонт СС.

Решение

Для имитации неисправных блоков СС и имитации наличия исправных блоков в ремонтном подразделении воспользуемся способом определения по жребью. Для этого рассчитаем вероятности исходов и сведём их в табл. 3.9 и 3.10 соответственно.

Таблица 3.9

Вероятности появления неисправных блоков

| Q_i | ABC | $ABC\bar{C}$ | С другими блоками |
|-------|---------------------|---------------------------------|-------------------|
| P_i | $P_A P_B P_C$ | $P_A P_B (1 - P_C)$ | |
| l_r | $l_1 = P_A P_B P_C$ | $l_2 = l_1 + P_A P_B (1 - P_C)$ | 1 |
| k | 1 | 2 | 3 |

Таблица 3.10

Вероятности наличия исправных блоков

| Q_{Hi} | ABC | $ABC\bar{C}$ | С другими блоками |
|----------|------------------------------|--|-------------------|
| P_{Hi} | $P_{HA} P_{HB} P_{HC}$ | $P_{HA} P_{HB} (1 - P_{HC})$ | |
| l_{Hr} | $l_1 = P_{HA} P_{HB} P_{HC}$ | $l_2 = l_1 + P_{HA} P_{HB} (1 - P_{HC})$ | 1 |

Так как нужно определить абсолютное и относительное (вероятность ремонта) количества отремонтированных СС, поступивших с неисправными блоками A, B и A, B, C то нет необходимости рассчитывать вероятности для других сочетаний блоков.

Так как, согласно постановке задачи, ремонт производится только тогда, когда имеются соответствующие исправные блоки, то нужно смоделировать событие с неисправными блоками и событие с наличием исправных блоков. Если порядковые номера согласно шкалам вероятностей совпадут, то событие – ремонт произойдёт и его нужно зафиксировать.

Для работы ИМ нужно ввести количество прогонов модели $N0$, количество всего поступающих в ремонт CC , данные числовой шкалы l_r массивом L , данные числовой шкалы l_{nr} массивом LH .

Нужны также переменные для:

N – счёта текущего числа прогонов модели;

A – записи и последующего вывода абсолютного количества отремонтированных CC ;

D – записи и последующего вывода относительного количества отремонтированных CC ;

i – счёта текущего числа событий;

j – счёта текущего числа поступивших в ремонт неисправных CC ;

M – счёта текущего числа отремонтированных CC во всех прогонах модели;

x – записи числа – результата обращения к датчику равномерно распределённых случайных чисел.

Алгоритм имитационной модели приведен на рис. 3.20.

Согласно постановке задачи в блоках 3...7 по данным табл. 3.9 разыгрывается, с какими неисправными блоками поступает CC в ремонт. В результате розыгрыша определяется номер интервала числовой шкалы вероятностей (столбца табл. 3.9) и запоминается как переменная k .

Аналогично в блоках 8...11 разыгрывается по данным табл. 3.10 наличие необходимых для замены исправных блоков. Если такие блоки имеются, т. е. выполняется условие $k = i$ в блоке 12, в счетчик M (блок 13) добавляется единица и осуществляется переход к блоку 14 для проверки количества поступивших CC в ремонт.

В ИМ организованы два цикла:

внутренний цикл по количеству поступивших всего в ремонт неисправных CC (блок 14);

внешний цикл по числу реализаций модели (блок 16).

Если условие $j < R$ в блоке 14 выполняется, в блоке 15 число прогонов увеличивается на 1.

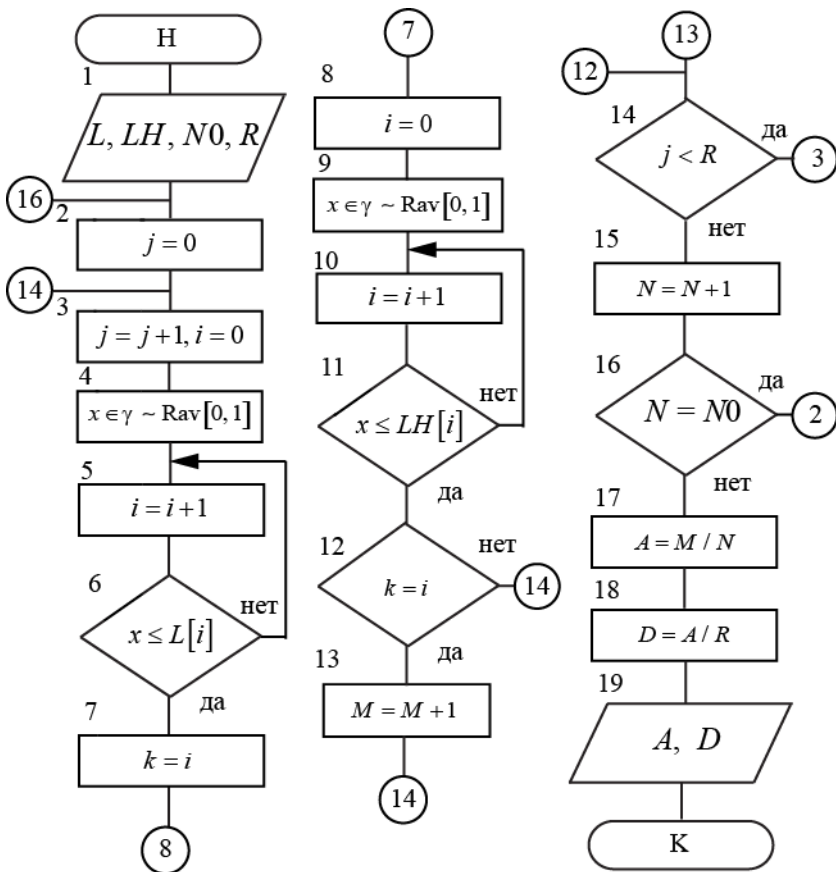


Рис. 3.20. Алгоритм имитационной модели ремонта СС

И далее, если в блоке 16 выполняется условие $N = N0$, то есть выполнено заданное число прогонов модели, в блоках 17 и 18 вычисляются абсолютное количество отремонтированных СС и вероятность ремонта СС соответственно.

В блоке 19 результаты моделирования выводятся на носитель.

3.10. Классификация случайных процессов

Случайная величина $X(t)$, зависящая от одного неслучайного вещественного аргумента t , называется случайным процессом. $X(t)$ является случайной величиной при каждом фиксированном значении

аргумента. Обычно (во всяком случае, для процессов, протекающих в технических системах) в качестве вещественного аргумента выступает время, поэтому случайный процесс будем обозначать $X(t)$.

Определим два понятия, присущие случайным процессам: сечение и реализация (рис. 3.21).

Сечением случайного процесса $X(t)$ называется случайная величина $x(t_j)$, являющаяся значением случайного процесса в фиксированный момент времени t_j .

Реализацией случайного процесса $X(t)$ называется функция времени $x_i(t)$, описывающая его течение в некотором i -м опыте.

Случайный процесс $X(t)$ и аргумент t могут быть дискретными или непрерывными.

Очевидно, вследствие особенностей представления информации в компьютере моделью случайного процесса будет модель дискретной последовательности дискретного случайного процесса. Следовательно, чтобы смоделировать реальный случайный процесс, необходимо выполнить следующее:

разбить интервал исследования на M временных точек t_j , которых должно быть столько, чтобы обеспечить необходимую точность воспроизведения исследуемого процесса;

выполнить одну реализацию случайного процесса, то есть для каждого момента времени t_j определить сечение, разыграв случайное число, имеющее характеристики случайного процесса;

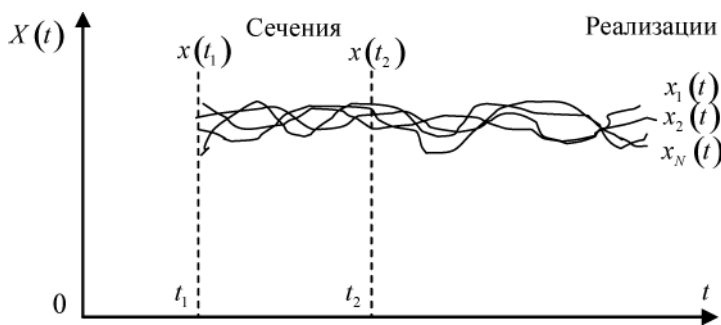


Рис. 3.21. Реализации и сечения случайного процесса

определить аналогичные сечения для каждой из N реализаций случайного процесса (число N выбирается таким, чтобы обеспечить необходимые точность и достоверность результатов).

Случайные процессы являются удобной математической моделью функций времени, значения которых случайные величины. Например: число звонков, поступающих в единицу времени на телефонную станцию, являясь случайной величиной, зависит от времени суток; расход электроэнергии в единицу времени – тоже функция времени со случайными значениями. То есть можно сказать, что случайный процесс – это однопараметрическое семейство случайных величин, зависящих от значений параметра, имеющего физический смысл времени.

Случайные процессы могут быть:

стационарные;

нестационарные.

На практике часто встречаются случайные процессы, у которых все реализации однородны в вероятностном смысле. То есть значения всех сечений представляют собой случайные числа, одинаково распределенные с одинаковыми математическими ожиданиями и дисперсиями:

$$M[x(t)] = M[x]; \quad D[x(t)] = D[x].$$

Такие процессы называют *стационарными*.

Что касается автокорреляционной функции $K_k(t_k, t_3)$, то её значение в стационарном процессе зависит только от разности $t_k - t_3$ и не зависит от того, в каком месте временной оси находятся точки t_k и t_3 .

Для стационарного процесса нет необходимости определять искомые характеристики для всех M сечений, а достаточно только для одного сечения N реализаций случайного процесса. То есть вместо $M \times N$ измерений выполнить только N измерений. По данным этих измерений рассчитываются оценки $\bar{x}(t_j)$ и $S_{x(t_j)}^2$, которые в силу стационарности и являются оценками характеристик всего случайного процесса $M[x]$ и $D[x]$.

Если сечения случайного процесса неоднородны в вероятностном смысле, то такой процесс называется *нестационарным*.

В работе модели стационарного процесса может присутствовать и нестационарный период. Это разного рода переходные процессы.

Например, характеристики начального периода работы модели нестационарные из-за того, что начальные установки характеристик процесса были не равны характеристикам, значения которых установятся в дальнейшем. Естественно, речь идёт о средних значениях характеристик.

Важнейшим свойством случайного процесса является свойство эргодичности.

Свойство *эргодичности* заключается в том, что все реализации случайного процесса имеют одинаковые статистические характеристики. Отсюда следует, что одна реализация случайного процесса характеризует весь случайный процесс $X(t)$, следовательно, для определения статистических характеристик процесса достаточно выполнить одну реализацию.

Обычно рассматривают свойство эргодичности по отношению к одной какой-либо характеристике случайного процесса. Относительно оценки математического ожидания свойство эргодичности формально выглядит так:

$$\frac{1}{N} \sum_{i=1}^N x_i(t_j) = \frac{1}{M} \sum_{j=1}^M x_i(t_j).$$

Свойством эргодичности обладают многие случайные процессы и, в том числе, все стационарные.

Таким образом, можно сформулировать определение эргодического процесса.

Случайный процесс $X(t)$ называется *эргодическим*, если его основные характеристики $M[x]$ и $D[x]$ могут быть получены не только усреднением по множеству реализаций, но и усреднением по времени одной реализации.

Например, при изучении флуктуационного шума радиоприемников, представляющего собой стационарный случайный процесс, достаточно ограничиться измерением сечений в течение заданного времени T в одном конкретном образце. Результаты, полученные при обработке данных измерений, могут быть распространены на все идентичные радиоприёмники. Аналогичные рассуждения могут быть проведены и относительно шума в канале связи.

3.11. Способы продвижения модельного времени

При реализации имитационной модели используются обычно три представления времени:

реальное время системы, функционирование которой имитируется;
модельное время, по которому организуется синхронизация событий в модели;

машинное время имитации, отражающее затраты ресурса времени компьютера.

Время в компьютерной модели принципиально не может протекать непрерывно. В компьютере в каждый момент времени выполняется одна команда. Но даже отдельные события реального процесса, протекающие, скажем, одновременно и мгновенно, в имитационной модели представляются цепочкой команд, на выполнение которых тратится машинное время.

Следовательно, время в модели, то есть модельное время (МВ), продвигается дискретно, скачками.

Продвижение времени в модели может быть организовано двумя способами:

продвижение модельного времени с фиксированным переменным шагом Δt ;

продвижение модельного времени до очередного события (по принципу Δx).

Сущность первого способа поясним временными диаграммами, показанными на рис. 3.22.

На диаграммах *a...z* показаны моменты смены дискретных состояний элементами 1...4 системы. На диаграмме *d* – временная последовательность смены состояний системой. На диаграмме *e* – точки модельного времени, то есть время смены состояний системы, показанных на диаграмме *d*.

Так как моменты модельного времени на диаграмме *e* не связаны с моментами появления событий *a...z*, то имитационная модель с фиксированным шагом продвижения времени искажает действительные процессы в системе: разновременные события представляются одновременными, моменты свершения событий фиксируются, как правило, с опозданием. Уменьшая величину Δt , можно уменьшить искажение действительного процесса. Однако это приводит к увеличению затрат машинного времени, особенно, если интервалы между сменами состояний в среднем больше, чем Δt .

На диаграмме ж рис. 3.22 демонстрируется сущность второго способа. Она заключается в том, что модельное время сдвигается вперед не на фиксированную величину Δt , а точно до времени наступления самого раннего из очередных событий – на Δx .

Видно, что недостатки, присущие первому способу, здесь исключены: события рассматриваются и моделируются в моменты их свершения, и одновременно (события a_1, z_1), если у них одинаковое время появления. Промежутки времени, когда в модели «ничего не происходит», пропускаются без особых затрат машинного времени. Эти пропуски все равно учитываются в модельном времени.

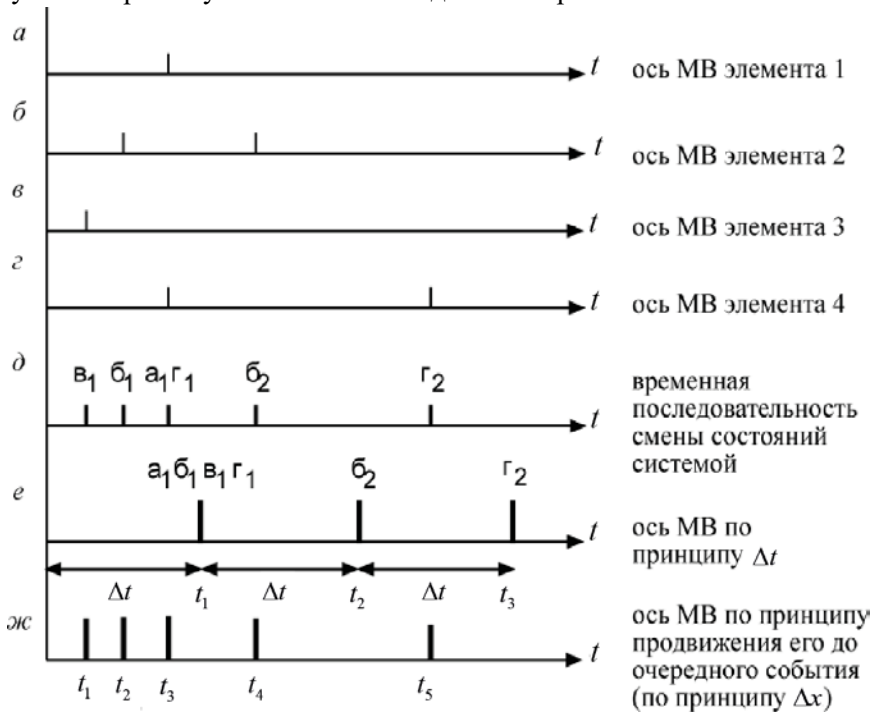


Рис. 3.22. Временная диаграмма работы модели

Однозначных рекомендаций по выбору того или иного способа продвижения модельного времени нет. Из общих рассуждений можно установить, что, если смена состояний в моделируемой системе происходит регулярно и часто, нет ограничений на расход машинного времени, то продвижение модельного времени фиксированными шагами Δt вполне приемлемо.

Если же смена состояний происходит редко и нерегулярно, кроме того, предъявляются повышенные требования к точности моделирования, то целесообразней второй способ продвижения модельного времени – скачками до ближайшего по времени события.

Мы рассмотрели способы продвижения модельного времени при так называемом последовательном (квазипараллельном) имитационном моделировании, характерным признаком которого является наличие централизованного списка событий и глобальных часов модельного времени. Обычно в таких имитационных моделях исследуемого процесса устанавливаются (или определяются) реальные затраты времени моделируемой системы (процесса) в масштабе, который устанавливает сам исследователь. Как правило, эти затраты безотносительны к естественному движению времени, которое обычно называют «реальным», хотя правильнее называть его естественным или натуральным временем.

Пример 3.12. На узел связи поступают заявки на передачу сообщений. Интервалы времени поступления заявок подчинены показательному закону с математическим ожиданием T_1 . На узле связи имеются два канала передачи данных. При поступлении очередной заявки в интервале времени $[0..T_2]$ вероятности того, что каналы A и B будут свободны, соответственно равны P_{1A} и P_{1B} .

При поступлении заявок после времени T_2 вероятности того, что каналы A и B будут свободны, соответственно равны P_{2A} и P_{2B} . Сообщение передаётся по любому свободному каналу. Если оба канала заняты, заявка теряется.

Построить алгоритм имитационной модели с целью определения абсолютного и относительного числа обслуженных заявок (вероятности обслуживания) из их общего количества, поступивших на узел связи за время моделирования T_3 .

Решение

В ранее рассмотренных постановках задач отсутствовала динамика или фактор времени. Например, просто указывалось, что поступает такое-то количество заявок, а за какое время и с какими интервалами – неизвестно. Настоящая постановка задачи отличается тем, что в неё введена динамика – определены интервалы времени поступления заявок – сообщений, подчинённые показательному закону. Поэтому число поступивших сообщений не указывается, так как оно неизвестно, случайно.

В модели должно быть модельное время. Оно предназначено для синхронизации работы элементов имитационной модели. Модельное время масштабируется по отношению к реальному времени работы моделируемой системы или процесса. Например, в данном случае время моделирования работы узла T_3 , пусть 2 часа, а одна единица модельного времени пусть равна одной секунде. Тогда время моделирования равно 7200 единицам модельного времени. Также масштабируются и остальные временные характеристики модели.

Согласно постановке задачи имеются два канала и известны вероятности, когда они свободны. Канал свободен – это единичное событие. Для его моделирования нужно обратиться к датчику равномерно распределённых случайных чисел и полученное случайное число сравнить с вероятностью. Если оно меньше или равно заданной вероятности, то канал свободен. Если нет, то канал занят.

Поскольку каналов два, то нужно обращаться к датчику равномерно распределённых случайных чисел в случае, если первый канал оказался занятым и нужно определить, свободен или занят второй канал, так как это разные независимые события.

Для моделирования интервалов времени между соседними запросами нужно обращаться к датчику экспоненциально распределённых случайных чисел, указав среднее значение T_1 .

Один из вариантов построения алгоритма модели может быть таким. Имитируется поступление заявок на передачу сообщений через интервалы времени, подчинённые экспоненциальному закону. При очередном поступлении заявки текущее модельное время увеличивается на величину интервала. Это текущее модельное время сравнивается со временем T_2 . Если оно меньше, то разыгрывается передача сообщений по каналам, свободным с вероятностями P_{1A} и P_{1B} . Если же текущее модельное время больше T_2 , но меньше T_3 , то разыгрывается вероятность передачи сообщений по тем же каналам, но свободным с вероятностями P_{2A} и P_{2B} .

Для работы алгоритма необходимо ввести:

N_0 – количество прогонов модели;

T_1 – средний интервал поступления запросов;

T_2 – время, после которого изменяются вероятности занятости каналов;

$P_{1A}, P_{1B}, P_{2A}, P_{2B}$ – вероятности нахождения каналов в свободных состояниях;

$T3$ – время моделирования.

Потребуется также переменные для проведения промежуточных расчётов и вывода результатов моделирования.

N – счётчик текущего числа прогонов модели;

M – счётчик числа обслуженных заявок (переданных сообщений) за $N0$ прогонов модели;

S – счётчик числа поступивших заявок (сообщений) во всех прогонах модели;

t – текущее модельное время одного прогона модели;

x – число – результат обращения к ДРРСЧ при определении состояния канала;

A – абсолютное число обслуженных заявок;

P – относительное число обслуженных заявок (вероятность обслуживания заявок).

Алгоритм имитационной модели приведен на рис. 3.23.

В связи с введением фактора времени следует обратить внимание на то, что алгоритм модели содержит два цикла: первый, внутренний определяется временем моделирования $T3$ (блок 14), а второй цикл, внешний – количеством реализаций модели N (блок 16).

Блок 3 предназначен для определения (розыгрыша) интервала времени поступления очередной заявки и счёта количества поступивших заявок за все прогоны модели.

В блоке 4 проверяется текущее модельное время: если условие выполняется, то обслуживание имитируется блоками 5...8, в противном случае – блоками 9...12.

В блоке 13 фиксируется выполнение очередной заявки, если это имело место.

В блоке 14 проверяется: не завершилось ли время моделирования? Если нет, то условие $t \geq T3$ не выполняется, и имитация поступления очередной заявки, начиная с блока 3, повторяется.

При выполнении условия $t \geq T3$ в блоке 16 проверяется условие $N = N0$. Если оно не выполняется, то, начиная с блока 3, выполняется очередной прогон модели.

Если условие $N = N0$ выполняется, значит, выполнилось заданное $N0$ число прогонов модели. В блоке 18 определяется абсолютное число обслуженных заявок, а в блоке 19 рассчитывается искомая вероятность и в блоке 20 результаты моделирования выводятся на носитель.

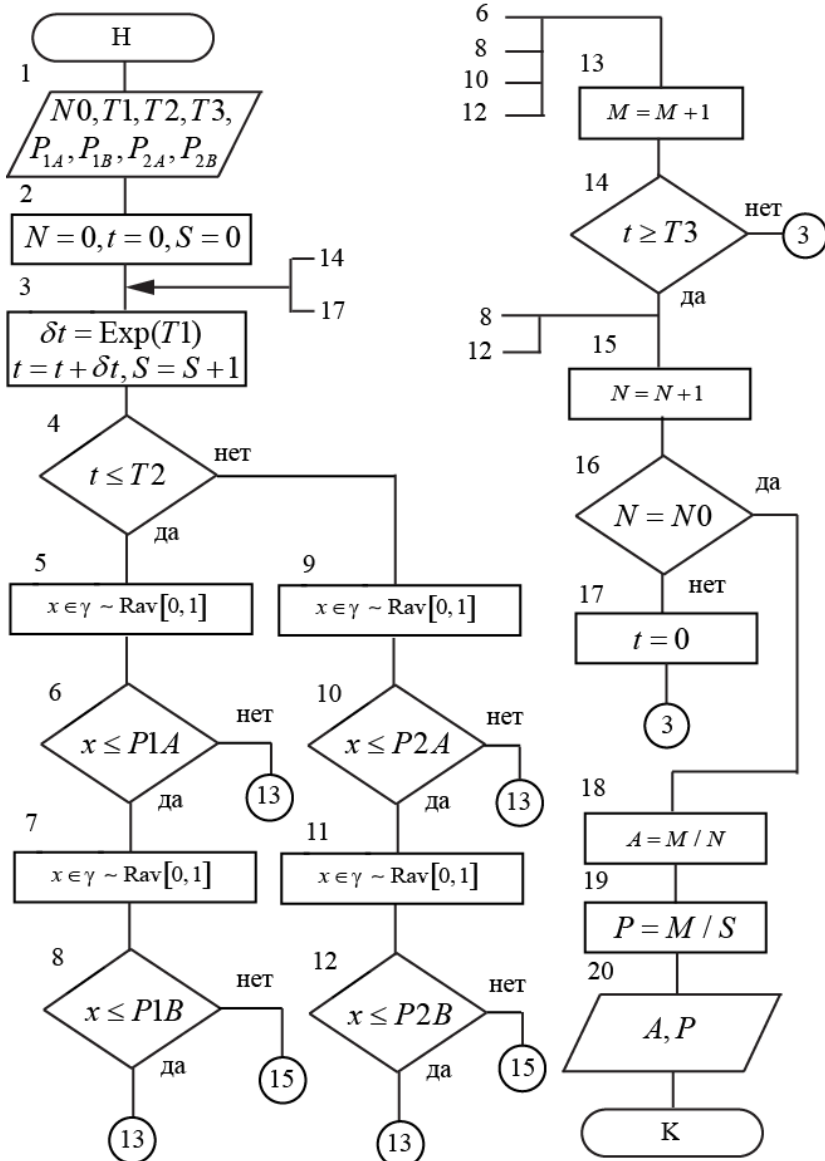


Рис. 3.23. Схема алгоритма имитационной модели передачи сообщений

В примере 3.11 мы построили алгоритм ИМ ремонта СС без учёта фактора времени. Построим этот же алгоритм в динамике.

Пример 3.13. В ремонтное подразделение поступают вышедшие из строя средства связи (СС) с интервалами времени T_1 , распределёнными по экспоненциальному закону. В каждом СС могут быть неисправными в любом сочетании блоки A, B, C . Вероятности выхода из строя блоков P_A, P_B, P_C соответственно. Ремонт производится путем замены неисправных блоков исправными блоками. В момент поступления неисправного СС вероятности наличия исправных блоков P_{HA}, P_{HB}, P_{HC} соответственно. При отсутствии хотя бы одного из исправных блоков A, B, C ремонт неисправного СС не производится. При наличии всех исправных блоков A, B, C средство связи ремонтируется. Время ремонта подчиняется экспоненциальному закону со средним значением T_3 . СС также не ремонтируется, если имеются исправные блоки A, B, C , но ремонтное подразделение занято ремонтом предыдущего СС.

Построить алгоритм имитационной модели с целью определения абсолютного и относительного (вероятности ремонта) количества отремонтированных СС с неисправными блоками A, B, C и A, B из общего количества поступивших в ремонт СС за время T_2 .

Решение.

Постановка данной задачи от рассмотренной в примере 3.11 отличается тем, что в ней присутствует динамика – интервалы времени поступления неисправных СС и время моделирования – работы подразделения.

Расчёты вероятностных шкал появления неисправных блоков (табл. 3.9) и наличия исправных блоков (табл. 3.10) остаются прежними. Прежней остаётся и идея построения алгоритма: так как ремонт производится только тогда, когда имеются соответствующие исправные блоки, то нужно смоделировать событие с неисправными блоками и событие с наличием исправных блоков. Если порядковые номера согласно шкалам вероятностей совпадут, то событие – ремонт произойдёт и его нужно зафиксировать.

Для работы алгоритма модели нужно ввести следующие исходные данные:

T_1 – средний интервал времени поступления неисправных СС;

T_2 – время моделирования;

T_3 – среднее время ремонта неисправного СС;

N_0 – количество прогонов модели;

данные числовой шкалы l_r массивом L ;
 данные числовой шкалы l_{nr} массивом LH .

Потребуется также переменные для:

N – счёта текущего числа прогонов модели;

t – текущего модельного времени;

A – записи и вывода абсолютного количества отремонтированных СС – одного из результатов моделирования;

D – записи и вывода относительного количества отремонтированных СС – второго результата моделирования;

i – счёта текущего числа событий;

j – счёта текущего числа поступивших в ремонт неисправных СС за все прогоны модели;

M – счёта текущего числа отремонтированных СС за все прогоны модели;

x – числа – результата обращения к датчику равномерно распределённых случайных чисел;

δt – интервала поступления очередного неисправного СС;

δtr – времени ремонта одного неисправного СС;

tr – текущего модельного времени ремонта неисправного СС.

Внесём в алгоритм (рис. 3.20) изменения, необходимые для учёта времени ремонта неисправных СС. Рассмотрим вначале процесс ремонта по временной диаграмме (рис. 3.24).

На диаграмме показаны две оси модельного времени (МВ): ось МВ поступления неисправных СС и ось МВ работы ремонтного подразделения (РП).

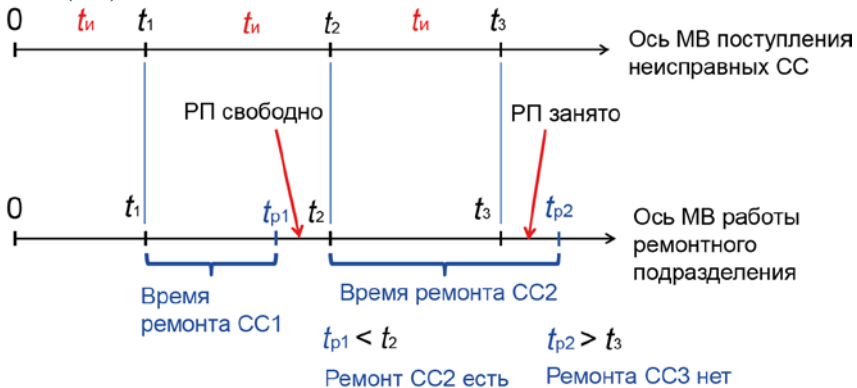


Рис. 3.24. Временная диаграмма работы ремонтного подразделения

На рис. 3.24: t_1, t_2, t_3 – моменты МВ поступления неисправных СС через интервалы t_n , распределённые по экспоненциальному закону.

В t_1 поступления СС1 ремонтное подразделение свободно и приступает к ремонту СС1. Время ремонта δtr также распределено по экспоненциальному закону.

Время окончания ремонта СС1 $t_{p1} = t_1 + \delta tr$. В период от t_{p1} до поступления неисправного СС2 ремонтное подразделение свободно, то есть $t_{p1} \leq t_2$, и сразу начинается ремонт СС2, время окончания – $t_{p2} = t_2 + \delta tr$.

Это время больше, чем время поступления неисправного СС3, то есть $t_3 < t_{p2}$, поэтому ремонта СС3 не будет.

Посмотрим, как это реализовано в алгоритме. Алгоритм имитационной модели приведен на рис. 3.25.

В блоке 14 проверяется, свободно ли ремонтное подразделение. Если свободно, то в блоке 15 фиксируется, что ремонт будет ($M = M + 1$). Разыгрывается время ремонта неисправного СС $\delta tr = \text{Exp}(T3)$ и оно добавляется к текущему модельному времени – получаем время окончания ремонта одного неисправного СС. Далее алгоритм работает так же, как и предыдущий алгоритм.

Если ремонтное подразделение не свободно (блок 14), то есть текущее модельное время работы ремонтного подразделения больше текущего модельного времени – времени поступления очередного неисправного СС, ремонт не фиксируется и осуществляется переход к блоку 16. Далее алгоритм работает так, как и предыдущий алгоритм.

В алгоритм также добавлены: обнуление tr – текущего модельного времени работы ремонтного подразделения в блоках 2 и 19.

Алгоритм модели относительно прост. Но если увеличить количество элементов и различных процессов системы, ввести их динамику, например, обслуживания заявок, то алгоритм модели усложнится. Для построения имитационных моделей функционирования сложных систем «лобовой» подход неприемлем.

Далее мы рассмотрим способы, позволяющие в некотором роде унифицировать построение моделей сложных систем с продвижением в них модельного времени по принципу Δt , а также до ближайшего события по принципу Δx .

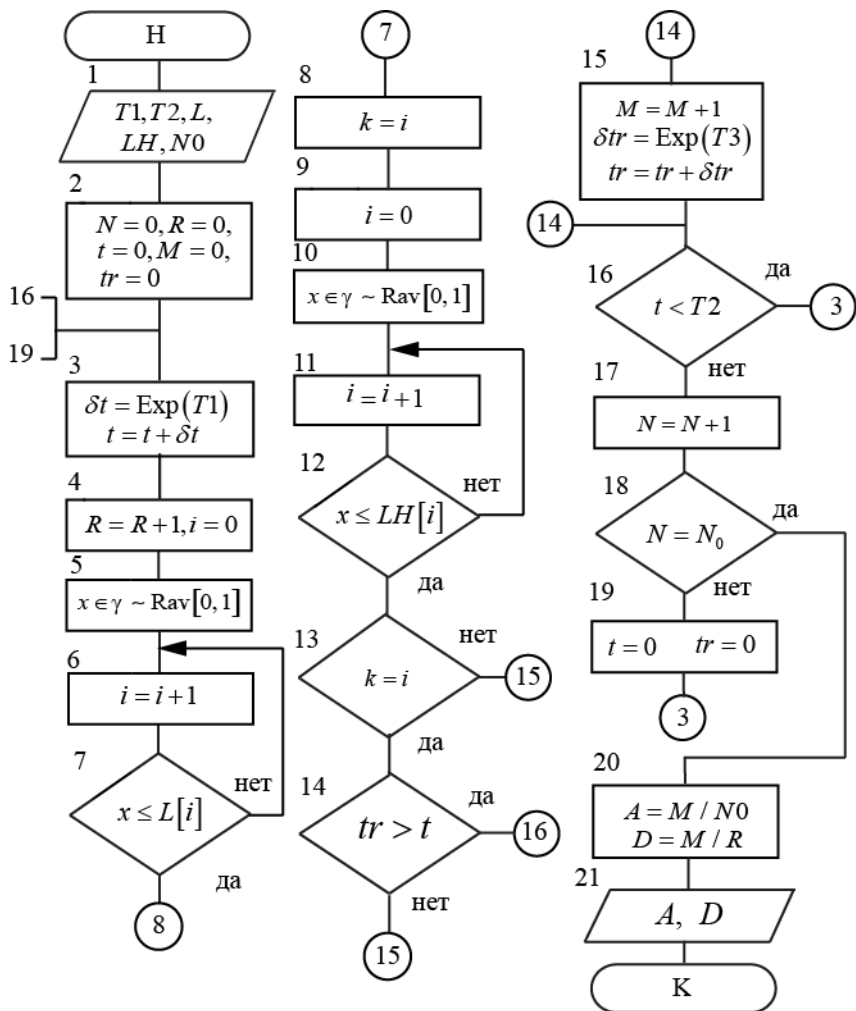


Рис. 3.25. Схема алгоритма имитационной модели ремонта СС

3.12. Модель противоборства двух сторон

Приемлемая по точности математическая модель такой сложной системы как бой невозможна из-за наличия неопределённых и не формализуемых факторов и уникальных ситуаций. Однако, приближительные частные модели возможны и целесообразны для количественного обоснования некоторых решений, оценки обстановки, прогнозирования результатов решений и др.

Имитационную модель, в которой продвижение модельного времени реализовано с фиксированным шагом, рассмотрим на примере огневого противоборства группировок A и B – модели высокоорганизованного боя.

Постановка задачи

Две группировки A и B ведут бой. В составе группировок A и B N_1 и N_2 боевых единиц со скорострельностями λ_1 и λ_2 и вероятностями поражения цели при одном выстреле P_1 и P_2 соответственно. Каждая группировка однородна, но не обязательно группировки однородны между собой. Например, бой танков с танками, танков с противотанковыми средствами и т.п.

Высокоорганизованным боем называют бой с полной информацией, а именно:

любая боевая единица одной стороны, пока она не поражена, может вести огонь по любой непораженной боевой единице другой стороны;

разведка, связь и управление организованы идеально, то есть перенос огня каждого средства на новую цель происходит мгновенно после поражения предыдущей цели;

пораженная боевая единица в дальнейших действиях не участвует, то есть за время боя не восстанавливается, пополнения сторон нет;

временем полёта носителя заряда пренебрегаем;

перенос огня не влияет на скорострельность боевых единиц и вероятность поражения;

количество боеприпасов неограниченно.

Цель моделирования. Прогнозирование средних количеств поражённых и непоражённых боевых единиц каждой группировки на любой момент времени.

Для начала, чтобы не загромождать алгоритм модели, введём два ограничения:

вероятность поражения цели одним выстрелом $\leq 0,1$;

вероятность одновременного поражения двумя средствами друг друга пренебрежимо мала.

Впоследствии мы убедимся, что в имитационной модели рассматриваемого боя эти ограничения могут быть сняты.

Учебная задача состоит в том, чтобы изучить структуру конкретного алгоритма, в котором продвижение модельного времени реализовано с фиксированным шагом Δt .

Введём обозначения:

A, B – первоначальные численности группировок;

a, b , – текущие значения численностей группировок A и B соответственно;

a_k, b_k – число оставшихся средств каждой из сторон в конце k -го интервала моделирования (после k -ой реализации (прогона) модели);

n – текущее число реализаций модели;

N – заданное число реализаций модели случайного процесса;

t – текущее время;

T – длительность интервала моделирования;

$\bar{a}(T), \bar{b}(T)$ – средние за N реализаций модели численности оставшихся средств (боевых единиц) сторон в конце каждого интервала моделирования:

$$\bar{a}(T) = \frac{\sum_{k=1}^N a_k}{N}, \quad \bar{b}(T) = \frac{\sum_{k=1}^N b_k}{N},$$

i, j – переменные счета средств сторон A и B соответственно.

Идея построения имитационной модели состоит в реализации модели противоборства N раз и фиксации остатков сторон после каждой реализации. По выполнении N реализаций будут определены оценки численностей сторон $\bar{a}(T), \bar{b}(T)$.

Определим величину Δt . На выбор величины Δt влияют два противоречивых требования:

на протяжении отрезка времени Δt не должно происходить много событий, так как они будут зафиксированы как одновременные, что исказит исследуемый реальный процесс;

на протяжении отрезка времени Δt должно произойти хотя бы одно событие, иначе будет много «пустых» прогонов модели, что увеличит машинное время.

С учетом приведенных выше требований разобьем интервал моделирования T на равные отрезки Δt такие, чтобы каждое огневое средство любой из сторон могло выстрелить не более одного раза.

Например, скорострельность средств поражения стороны A $\lambda_A = 2$ выстр./мин, а средств стороны B – $\lambda_B = 3$ выстр./мин. Так как $\lambda_B > \lambda_A$, то Δt следует выбрать исходя из условия $\Delta t < 1 / \lambda_B$.

Алгоритм модели противоборства двух сторон состоит из четырёх модулей:

М1 – установка начальных условий;

М2 – продвижение модельного времени;

М3 – формирование результата моделирования с заданными точностью и достоверностью;

М4 – имитация противоборства двух сторон.

Схема алгоритма модели противоборства двух сторон показана на рис. 3.26.

Модуль установки начальных условий (М1). Он состоит из трёх блоков 1...3.

Блок 1 – установка начальных условий на весь процесс моделирования: число реализаций N модели, интервал исследования T (время моделирования), величина временного шага Δt , установка в нуль ячеек $\sum a_k, \sum b_k$.

Блок 2 – установка начальных условий на очередную реализацию модели: восстановление численностей сторон $a = A, b = B$ и исходного времени $t = 0$.

Блок 3 – установка начальных условий на очередной отрезок Δt модельного времени: подготовка перебора ($i = 0, j = 0$) средств поражения каждой стороны.

Модуль продвижения модельного времени (М2). Состоит из двух блоков 4..5.

Блок 4 – продвижение модельного времени на очередной временной отрезок Δt ($t = t + \Delta t$).

Блок 5 – проверка условия окончания очередной реализации модели ($t \geq T$).

Модуль формирования результата и обеспечения заданной точности (М3). Состоит из блоков 6...9.

Блок 6 – накопление суммы остатков средств каждой стороны за текущее количество интервалов моделирования (реализаций модели).

Блок 7 – счётчик числа реализаций модели ($n = n + 1$).

Блок 8 – осуществляет контроль над выполнением заданного числа реализаций модели. Число реализаций модели определяется, исходя из заданных исследователем точности и достоверности результатов моделирования.

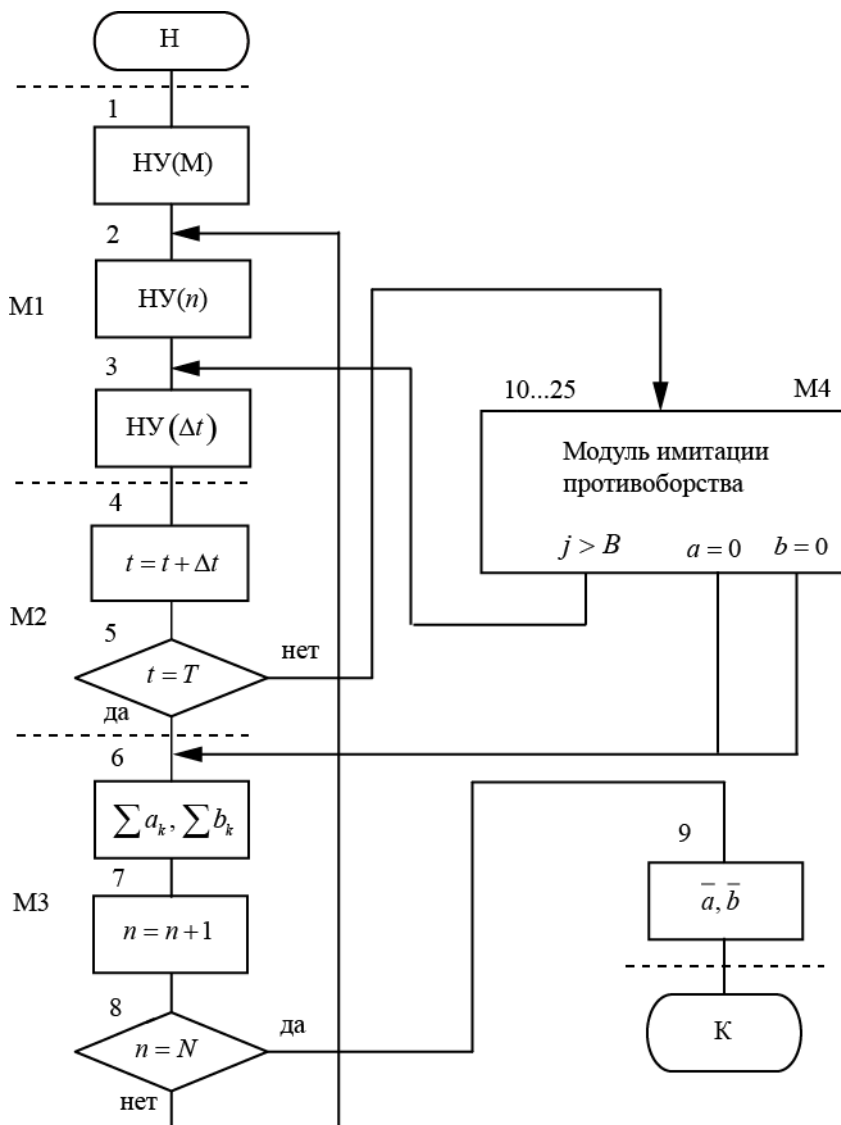


Рис. 3.26. Схема алгоритма модели противоборства двух сторон

Блок 9 – формирование конечного результата моделирования.

В данном случае – \bar{a} и \bar{b} .

При необходимости определяются и другие статистические характеристики.

Модуль имитации противоборства сторон (М4). Этот модуль – основной. Если структура предыдущих модулей, в общем-то, стандартна, то реализация данного модуля носит функционально-индивидуальный характер. Модуль включает блоки 10...25. Схема алгоритма модуля представлена на рис. 3.27.

Каждое средство противоборствующих сторон идентифицируется его номером. Номера средств стороны $A: i = \overline{1, A}$, средств стороны $B: j = \overline{1, B}$.

Блок 10 – выбор очередного средства стороны $A: i = i + 1$.

Блок 11 – проверка: все ли средства стороны A получили право на выстрел? Если $i > A$, то управление передается блоку 18 для имитации выстрелов средствами стороны B . В противном случае управление передается блоку 12.

Блок 12 – проверка: боеспособно ли выбранное средство? Состояние средств сторон A и B определяют переменные α_i и β_j :

$$\left\{ \begin{array}{l} \alpha_i = 0, \text{ если } i\text{-ое средство стороны } A \text{ уничтожено;} \\ \alpha_i = 1, \text{ если } i\text{-ое средство стороны } A \text{ боеспособно;} \\ \beta_j = 0, \text{ если } j\text{-ое средство стороны } B \text{ уничтожено;} \\ \beta_j = 1, \text{ если } j\text{-ое средство стороны } B \text{ боеспособно.} \end{array} \right.$$

Если окажется $\alpha_i = 0$, то управление передается блоку 10 для выбора очередного средства стороны A . Иначе – переход к блоку 13.

Блок 13 – выбор цели из средств стороны B . Выбор цели может быть организован либо случайной, либо детерминированной процедурами. Самый простой способ: последовательная проверка средств стороны B с выбором первого непораженного средства.

Блок 14 – проверка выбранной цели: не уничтожена ли она была на предыдущих этапах данной реализации модели? Если $\beta_j = 0$, то переход к блоку 13 для выбора непораженной цели. Иначе – переход к блоку 15 для имитации выстрела.

Выстрел – одиночное событие со случайным исходом. Моделью такого события является известная нам конструкция, состоящая из двух блоков 15 и 15.1.

Блок 15 – обращение к ДСЧ за равномерно распределенным случайным числом $x_i \sim \text{Rav}[0, 1]$.

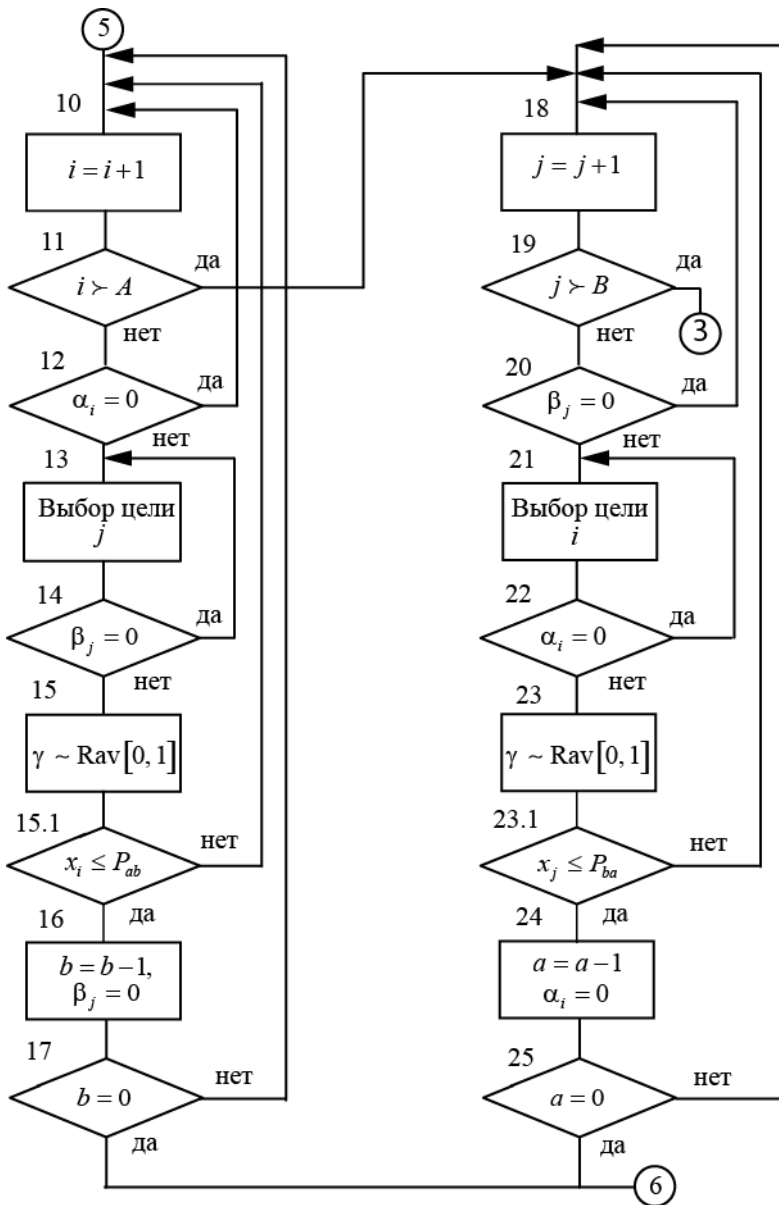


Рис. 3.27. Схема алгоритма модуля противоборства двух сторон

Блок 15.1 – проверка результата выстрела. Если $x_i \leq P_{ab}$, цель поражена и управление передается блоку 16 для фиксации этого факта. Если $x_i > P_{ab}$, то промах и управление передается блоку 10 для выбора очередного стреляющего средства стороны А.

Блок 16 – уменьшение числа средств стороны В: $b = b - 1$ и установка признака состояния пораженного средства: $\beta_i = 0$.

Блок 17 – не уничтожена ли вся группировка В? Если да, то данная реализация модели заканчивается и управление передается блоку 6 для фиксации оставшихся боеспособных средств стороны А. Если нет, то управление передается блоку 10 для выбора очередного средства стороны А и т. д.

После предоставления права на выстрел всем средствам стороны А соответствующее право дается средствам стороны В – переход из блока 11 в блок 18.

Функции блоков 18...25 попарно одинаковы с функциями блоков 10...17, изменены только обозначения – вместо А указано В, вместо $\alpha_i \rightarrow \beta_j$, вместо $P_{ab} \rightarrow P_{ba}$ и т. д.

По окончании перебора всех средств стороны В и, если не зафиксировано полное уничтожение средств стороны А ($a = 0$), управление передается блоку 3 для моделирования очередного скачка времени на величину Δt .

Примерная диаграмма изменения численностей сторон А и В на интервале $0 \dots T$ в k -й реализации показана на рис. 3.28.

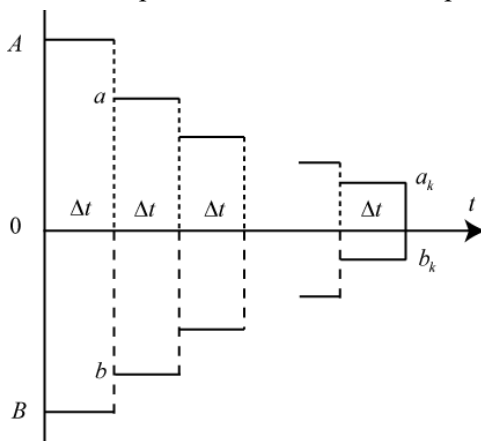


Рис. 3.28. Диаграмма изменения численности сторон

3.13. Модель противоборства как процесс блуждания по решётке

Итак, чтобы построить имитационную модель какого либо процесса, надо скопировать и описать в виде программы элементарные действия процесса и соединить их в нужной временной хронологической последовательности. Такое прямолинейное моделирование, в общем, понятно и естественно.

Однако, иногда можно увидеть совсем неожиданную аналогию (структурную или функциональную) между исследуемым процессом и некоторым другим, на первый взгляд ничего не имеющим общего с оригиналом. Имитационная модель этого нового процесса может оказаться проще и, тем не менее, позволит определить искомые характеристики.

В частности, для исследования процесса противоборства двух сторон таким «дублером» может быть процесс «блуждания частицы по решётке».

Рассмотрим, как перемещение некоторого объекта по решётке становится аналогом, моделью противоборства двух сторон, если целью моделирования является определение оценок численностей сторон в конце операции – интервала исследования.

Противоборствуют две стороны: A и B .

Первоначальные численности сторон: A и B .

Текущие численности сторон: a и b ; $a = \overline{0, A}$, $b = \overline{0, B}$.

Время «жизни» каждого средства случайно, имеет экспоненциальное распределение с параметрами λ_a и μ_b сторон A и B .

Обе стороны образуют единую систему $S_{a,b}$. В ходе боя численности сторон изменяются (уменьшаются) или остаются неизменными. Мы это трактуем как переход системы из одного состояния в другое: $S_{a,b} \rightarrow S_{a-1,b}$ или $S_{a,b} \rightarrow S_{a,b-1}$ и т. д.

Граф состояний такой системы имеет вид двумерной целочисленной решётки с однонаправленными связями. Переходу системы из одного состояния в другое можно поставить в соответствие случайное блуждание некоторой частицы по решётчатой структуре. Отсюда и название моделирующей абстракции.

Граф состояний приведен на рис. 3.29.

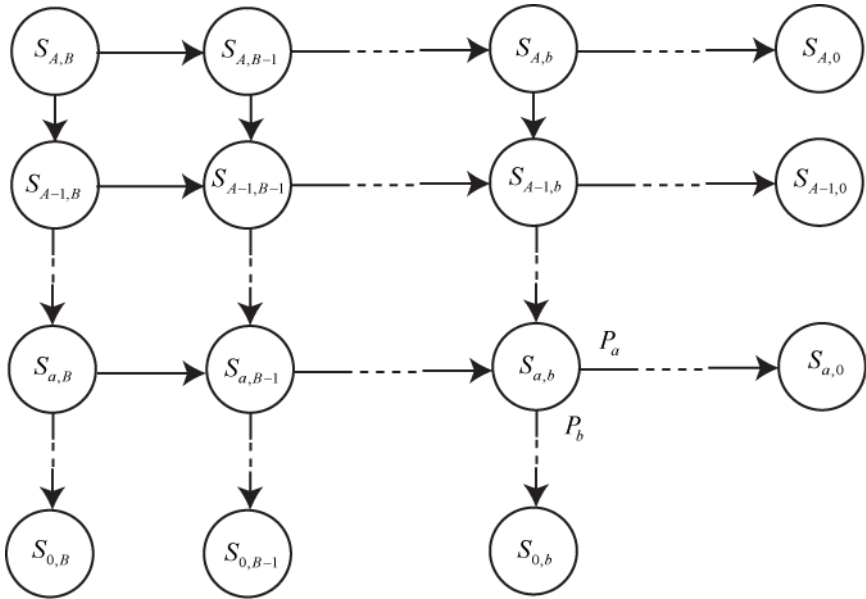


Рис. 3.29. Граф состояний моделируемой системы противоборства

Исходное состояние системы $S_{a,b}$.

Стороны ведут взаимоуничтожающий огонь. В предположении ординарности потока событий, переводящего систему из одного состояния в другое, система из состояния $S_{a,b}$ может перейти только лишь в состояние $S_{a-1,b}$ или $S_{a,b-1}$ (или остаться в прежнем состоянии $S_{a,b}$). Очевидно, направления переходов в графе-решётке только вправо или вниз.

Процесс противоборства заканчивается при полном уничтожении средств одной из сторон, т. е. при переходе в состояние $S_{0,b}$ или $S_{a,0}$, а также по окончании временного интервала исследования T . Так заканчивается одна, k -я реализация процесса противоборства. Зафиксировав остатки средств сторон a_k и b_k и усреднив их за N реализаций, найдем решение поставленной задачи – определение оценок численностей каждой из противоборствующих сторон на момент окончания операции.

Продвижение времени в модели – фиксированными шагами Δt . Величина Δt такова, чтобы на этом временном участке происходило не более одного события.

Суммарные интенсивности огня сторон A и B зависят от численностей их боеспособных средств и равны $a\lambda_a$ и $b\mu_b$.

Следовательно, вероятности уничтожения одного из средств (перехода системы в очередное состояние) равны:

$$P(S_{a,b} \rightarrow S_{a,b-1}) \cong a\lambda_a\Delta t = P_a,$$

$$P(S_{a,b} \rightarrow S_{a-1,b}) \cong b\mu_b\Delta t = P_b.$$

Эти приближенные равенства тем точнее, чем меньше Δt .

Схема алгоритма ИМ противоборства двух сторон, построенная способом «блуждания частицы по решётке» представлена на рис. 3.30.

Блоки 1 и 2 – установка начальных условий на весь процесс моделирования и на каждую очередную реализацию.

Блок 3 – расчет вероятностей P_a и P_b .

Блоки 4...10 – определение исхода противоборства на очередном временном отрезке Δt . Если $x_i \leq P_a$, то свершилось событие $S_{a,b-1}$, иначе – проверка условия $x_i \leq P_b$? Если это условие выполняется, то свершилось событие $S_{a-1,b}$, если не выполняется, то пораженных средств не оказалось. Вероятность этого события $S_{a,b} \rightarrow S_{a,b}$ равна $1 - P_a - P_b$.

Блоки 11...16 аналогичны соответствующим блокам предыдущей имитационной модели противоборства двух сторон (см. рис. 3.26).

Сравним обе рассмотренные модели противоборства двух сторон.

Последняя заметно проще. Упрощение связано с тем, что в ней не моделируется поведение каждого отдельного средства, а рассматривается поведение системы в целом.

Однако, первую модель несложно развить, например, учесть неоднородность средств в каждой группировке, указав соответствующие значения P_{ab} и P_{ba} . Можно учесть и другие, важные с точки зрения цели моделирования факторы, условия и ситуации.

Последняя модель таких возможностей не имеет.

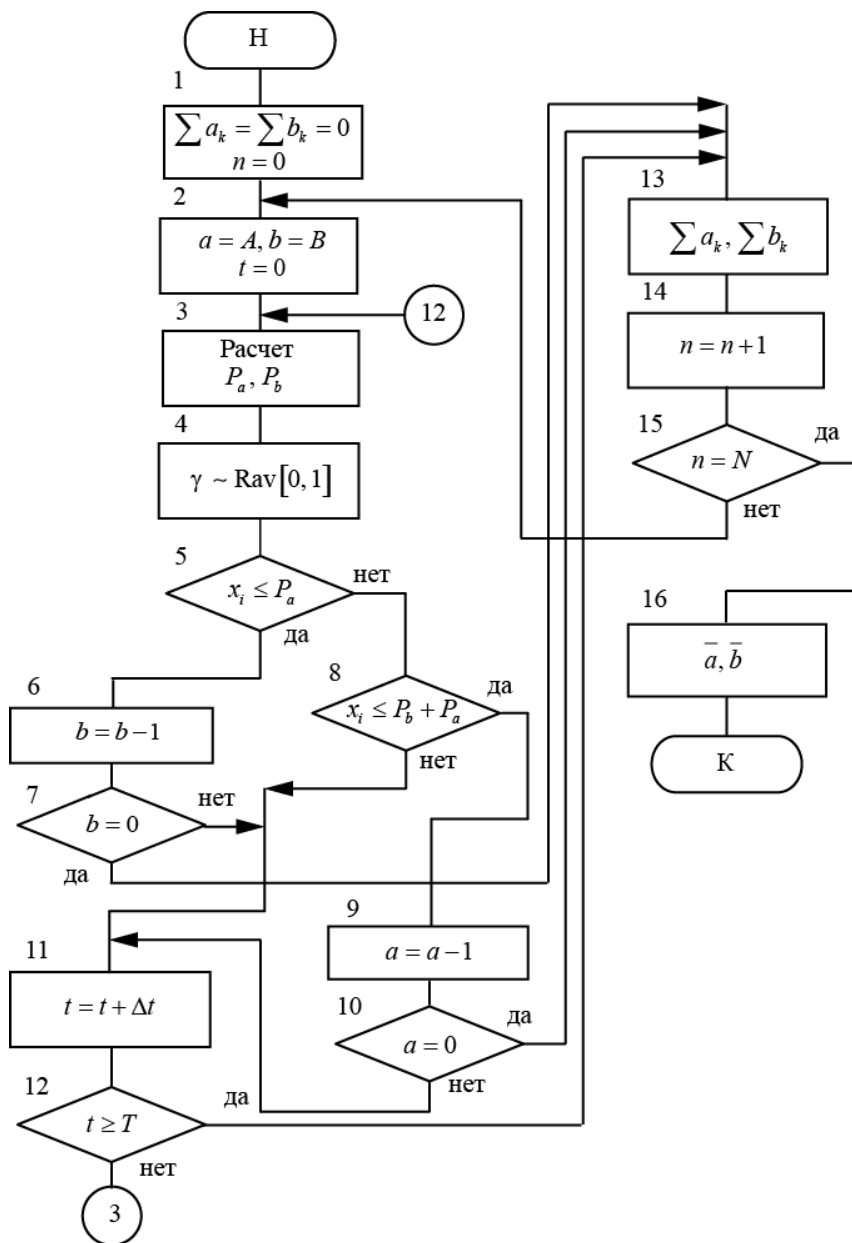


Рис. 3.30. Схема модели «блуждание частицы по решётке»

3.14. Типовая схема имитационной модели с продвижением времени по событиям

Уточним понятие события.

Под *событием* будем понимать смену состояния системы.

Виды событий:

- выход из строя технических средств;
- восстановление работоспособности техники;
- поступление сообщения на узел связи;
- начало передачи сообщения;
- конец передачи сообщения;
- уничтожение цели и т. п.

Вид события связан с типом элемента, сменяющего свои состояния. Смена состояний системой есть объединение смен состояний её элементов.

Будем считать, что событие совершается в конкретный момент времени мгновенно.

Вообще-то событие – смена состояния – совершается за некоторый отрезок времени. Но в интересующих нас процессах длительность смены состояний много меньше времени нахождения системы в каком-либо состоянии, поэтому длительностью события можно пренебречь.

Каждому событию соответствует пространственно-временная точка $t^{j,k}$, где t – момент свершения события; k – тип элемента системы, сменившего состояние; j – вид или номер события этого элемента.

Различают события активные и пассивные.

Активное событие – это смена состояния элемента под воздействием присущих ему внутренних причин.

Например, отказ техники – активное событие, так как оно определяется износом или скрытыми дефектами. Конец передачи сообщения по каналу связи – активное событие, так как при прочих равных условиях время передачи определяется быстроедействием передающей аппаратуры связи.

Пассивное событие – событие, возникшее под воздействием активного события.

Например:

начало передачи сообщения по каналу связи, так как начало зависит от другого активного состояния – поступления сообщения на узел связи;

изменение числа устройств, ожидающих ремонта. Увеличение их численности зависит от поступления неисправных устройств в мастерскую; уменьшение – от освобождения мастерской; и др.

В одном и том же элементе системы могут происходить активные и пассивные события.

Например:

начало решения задачи на компьютере – событие пассивное, так как оно происходит от другого события: поступления запроса на решение;

завершение решения – событие активное, так как оно зависит от внутренних свойств компьютера (производительности);

поступление заявок на обслуживание в очередь и покидание очереди – события пассивные, но уход из очереди «нетерпеливых» заявок – события активные.

Обычно элемент, в котором происходит активное событие, называют активным, пассивное событие – пассивным.

Основу имитационной модели системы с продвижением модельного времени по событиям составляют модели активных событий. Пассивные события моделируются просто, если выявлены их причинно-следственные отношения с активными событиями.

Множество точек $t^{j,k}$ представляет собой поведение системы во времени. Множество точек $t^{j,k}$ активных событий называется *списком событий*.

Список событий может быть сформирован либо перед началом моделирования, либо формироваться в ходе моделирования. На рис. 3.22 список событий представлен диаграммой *ж*.

Имитационная модель с продвижением времени по событиям любой системы состоит из следующих функциональных модулей:

модуль установки начальных условий;

модуль продвижения модельного времени;

модули реакции;

модуль обеспечения заданной точности и достоверности;

модуль формирования результата моделирования.

Модуль установки начальных условий. Модуль устанавливает значения разного рода констант, например, заданное число N реализаций модели, критерий завершения интервала моделирования T и др., а также элементы списка событий $t^{j,k}$ – все или только начальные для каждого вида события.

Если в предыдущем способе продвижения модельного времени на фиксированные промежутки Δt модуль состоял из трёх блоков установки начальных условий, то здесь достаточно двух: установка начальных условий на весь процесс моделирования (НУ(М)) и установка начальных условий на каждую реализацию (НУ(n)).

Модуль продвижения модельного времени. Его функции: выбор в списке событий момента времени свершения очередного, ближайшего события $\min t^{j,k}$ и фиксация очередного момента модельного времени:

$$t = t_0 + \min t^{j,k}.$$

Модули реакции. Представляют собой программы, имитирующие действия соответствующих активных событий и связанных с ними пассивных событий. Модулей реакции столько, сколько в системе активных событий. Подключение того или иного модуля реакции происходит согласно значениям j и k .

Помимо имитации событий в модулях реакции могут выполняться следующие действия:

фиксация текущих значений параметров, интересующих исследователя, так называемых числовых атрибутов, и, возможно, их текущая статистическая обработка;

подсчёт числа свершившихся событий данного вида;

розыгрыш времени свершения $t^{j,k}$ очередного события данного вида и занесение этого времени в список событий (если список формируется в ходе моделирования). Розыгрыш осуществляется с помощью соответствующего датчика случайных чисел, адекватно имитирующего временные интервалы между событиями.

Модуль обеспечения заданной точности и достоверности ведёт подсчёт числа реализаций модели $n = n + 1$. При достижении $n > N$ моделирование заканчивается.

Модуль формирования результата. Выполняет окончательную статистическую обработку данных моделирования и обеспечивает отображение результата в данном виде.

Взаимодействие модулей показано на рис. 3.31.

Переход на соответствующий модуль реакции A , B и т. д. происходит по значению индекса k текущего значения $t^{j,k}$.

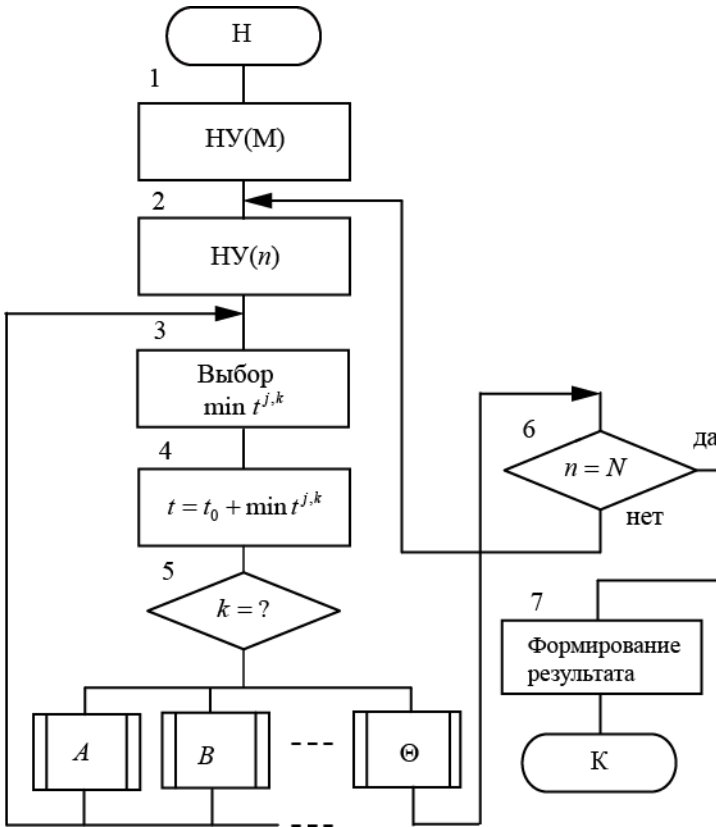


Рис. 3.31. Схема модели с продвижением времени по событиям

Безусловный выход из любого модуля реакции – в блок выбора очередного значения $\min t^{j,k}$ модуля продвижения времени.

В общий цикл предлагаемой структуры моделирующего алгоритма удобно вписывается определение конца интервала исследования T и момента конца моделирования ($n = N$). Для этого вся система интерпретируется как один обобщённый активный элемент, имеющий два состояния: включен и выключен.

В списке событий этому обобщённому элементу соответствует одна пространственно-временная точка $t^{j,k} = t_0 + T$. Здесь t_0 – начальная точка отсчета интервала исследования, обычно $t_0 = 0$. Когда в процес-

се работы моделирующего алгоритма окажется, что $\min t^{j,k} = t_0 + T$, управление будет передано специальному модулю реакции Θ , который зафиксирует конец текущей реализации ($n = n + 1$).

Этот модуль также в случае $n < N$ передаст управление блоку (НУ(N)) на выполнение очередной реализации, а в случае выполнения $n = N$ – модулю формирования результатов.

Ранее было отмечено, что некоторые элементы в зависимости от состояния могут быть активными либо пассивными.

В случае перехода такого элемента в пассивное состояние надо исключить возможность выбора его блоком «Выбор $\min t^{j,k}$ ». Для этого в список событий записывается время $t^{j,k} > T$ его переключения на свой модуль реакции. В этом случае модуль продвижения времени никогда не найдет этой временной точки, так как интервал исследования будет заканчиваться раньше.

Но когда этот элемент вновь станет активным, время его активного события $t^{j,k}$ должно быть занесено в список событий.

3.15. Имитационная модель системы массового обслуживания

Для иллюстрации способа продвижения модельного времени до ближайшего события рассмотрим имитационную модель СМО следующего вида:

многофазная;

многоканальная;

с несколькими неоднородными потоками заявок на обслуживание;

разомкнутая;

абсолютно надежная;

с очередями неограниченной ёмкости на всех фазах обслуживания.

СМО такого вида показана на рис. 3.32.

На рис. 3.32:

$\Gamma_1 \dots \Gamma_G$ – источники потоков заявок на обслуживание;

$K_{1,Q_1} \dots K_{S,Q_S}$ – каналы обслуживания;

S – число фаз;

$Q_1 \dots Q_S$ – число каналов в каждой фазе;

$L_1 \dots L_S$ – очереди заявок на входах соответствующих фаз.

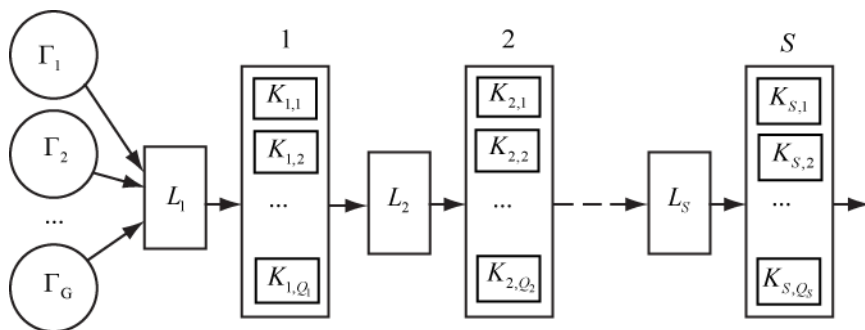


Рис. 3.32. Многофазная система массового обслуживания

Такой СМО может быть, например, сборочный цех предприятия. Из нескольких цехов ($\Gamma_1 \dots \Gamma_G$) поступают готовые блоки для сборки изделия. Сначала они поступают на диспетчерские посты, на которых происходит оформление соответствующей документации. После этого блоки направляются на сборочные участки.

Диспетчеризация – первая фаза обслуживания.

Сборка изделий – вторая фаза обслуживания.

Затем идёт стендовый контроль и регулировка – это третья фаза обслуживания.

Четвёртая фаза – приёмка готового изделия (например, военная приёмка).

Далее может быть упаковка изделия – пятая фаза.

Упакованные изделия поступают в транспортное подразделение для отправки – это шестая фаза.

На каждом из этих шести участков работают инженеры, техники, наладчики, контролёры, упаковщики. В силу различного рода случайностей на входе каждой фазы могут образовываться очереди комплектующих блоков, непроверенных изделий, готовых изделий. А также могут быть простои работников, что в терминах СМО трактуется как очереди свободных каналов.

Вернёмся к общей постановке задачи.

Целью моделирования является определение показателей исхода операции массового обслуживания на временном интервале $0..T$. Например, оценка относительной пропускной способности СМО, среднее количество заявок в очередях, коэффициент использования каналов в каждой фазе и др.

Проанализируем элементы СМО.

1. *Входные потоки заявок на обслуживание*

Поступление заявок на вход первой фазы – событие активное. Следовательно, в модели должны быть предусмотрены модули реакции на это событие. Каждый поток обслуживает свой модуль реакции, значит, модулей реакции этого типа должно быть столько, сколько имеется потоков заявок, то есть G . Моменты поступления заявок имитируются соответствующими генераторами (датчиками) случайных интервалов между заявками потоков $\Gamma_1 \dots \Gamma_G$.

2. *Каналы обслуживания*

Каждый канал имеет два состояния: свободен или занят. Переход в состояние «занят» – событие пассивное, так как определяется поступлением заявки. Переход в состояние «свободен» – событие активное, оно определяется внутренними свойствами канала, например, производительностью. Следовательно, в модели должны быть модули реакции на освобождение каналов. И этих модулей должно быть столько, сколько имеется в нашей СМО каналов:

$$Q_1 + Q_2 + \dots + Q_S.$$

3. *Очереди заявок на обслуживание на входах фаз*

Эти элементы СМО – пассивные, число состояний неограниченно.

4. *Очереди свободных каналов в каждой фазе обслуживания*

Эти элементы СМО также пассивные, число состояний равно числу каналов в каждой фазе.

5. *Обобщённый элемент «СМО в целом»*

Характеризуется одним активным событием – завершение интервала исследования T . Следовательно, для обработки этого события в модели предусматривается один специальный модуль реакции.

Изменения состояний считаем мгновенными.

Общее число модулей реакции в рассматриваемой модели СМО должно быть:

$$R = G + \sum_{i=1}^S Q_i + 1.$$

При большом числе входных потоков и каналов обслуживания количество модулей реакции будет значительным, и модель будет труднообозримой, громоздкой. Однако структуру модели можно существенно упростить, назначив однородным активным событиям по одному модулю реакции. Или: освобождение любого канала вызывает

обращение к другому общему модулю реакции. Естественно, на входе таких универсальных модулей реакции должен присутствовать блок настройки на конкретное активное состояние объединенной группы.

Строить же общий модуль реакции на освобождение всех каналов СМО нецелесообразно, так как освобождение каналов последней фазы СМО требует действий, заметно отличающихся от тех, которые необходимо выполнить при освобождении каналов не последней фазы.

Таким образом, имитационная модель СМО будет иметь всего четыре модуля реакции:

модуль реакции на поступление заявки (А);

модуль реакции на освобождение канала не последней фазы (В);

модуль реакции на освобождение канала последней фазы (С);

модуль реакции на завершение интервала исследования T (Θ) –

реакция на изменение состояния обобщенным элементом «СМО в целом».

Блок-схема имитационной модели СМО состоит из общей части и модулей реакции. Общая часть имеет стандартный вид и была рассмотрена ранее (см. рис. 3.31). Структура общей части определяется способом продвижения модельного времени по событиям. Она показана на рис. 3.33.

Блок 1 – установка начальных условий на весь процесс моделирования. Вводятся количество прогонов модели N , модельное время моделирования T (реальное время исследования в масштабе), указываются число источников заявок, число каналов. Счётчик числа реализаций n устанавливается в нуль (или в N – в зависимости от того, как организован счёт реализаций модели). При необходимости указываются данные планирования эксперимента с моделью, описываются необходимые переменные для счёта, хранения результатов моделирования, последующей их статистической обработки и вывода на соответствующие носители и др.

Блок 2 – установка начальных условий на очередную реализацию модели. Текущее время, счетчики числа поступивших и обслуженных заявок устанавливаются в нуль. Каналы обслуживания и очереди к ним устанавливаются в исходные состояния (установки исходных состояний определяются требованиями к исследованию) и др.

Наборы исходных данных $HU(M)$ и $HU(n)$ определяются конкретным назначением моделируемой СМО и характером последующих экспериментов.

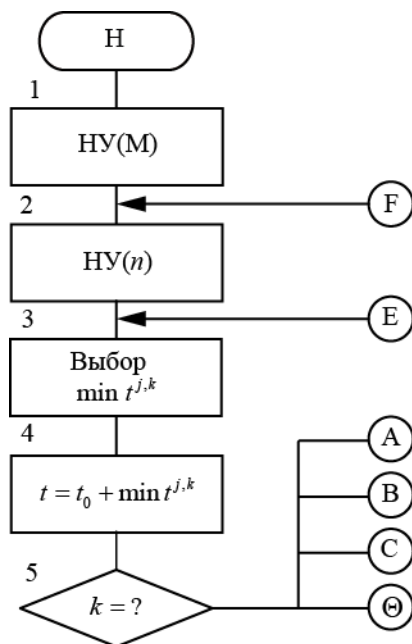


Рис. 3.33. Схема имитационной модели СМО

Блоки 3, 4 реализуют продвижение времени по событиям.

Блок 5 – выбор модуля реакции. В зависимости от индекса k пространственной временной точки $\min t^{j,k}$ передает управление соответствующему модулю реакции – А, В, С или Θ .

В рассматриваемой модели элементы СМО имеют только два состояния, одно из которых является начальным. Поэтому индекс j не имеет смысла.

Рассмотрим состав и функционирование модулей реакции.

Допустим, что в рассматриваемый момент времени свершилось активное событие – поступление заявки от одного из источников. Блок 5 передает управление модулю реакции А.

Схема алгоритма модуля реакции А приведена рис. 3.34.

Блок А5 – блок настройки. Обеспечивает доступ к ДСЧ, ячейкам и т. д. модуля, предназначенным для обслуживания конкретного потока заявок, которому принадлежит поступившая заявка.

Блок А6 – проверка наличия свободных каналов первой фазы. Количество свободных каналов в очереди обозначено $l_{ск}$.

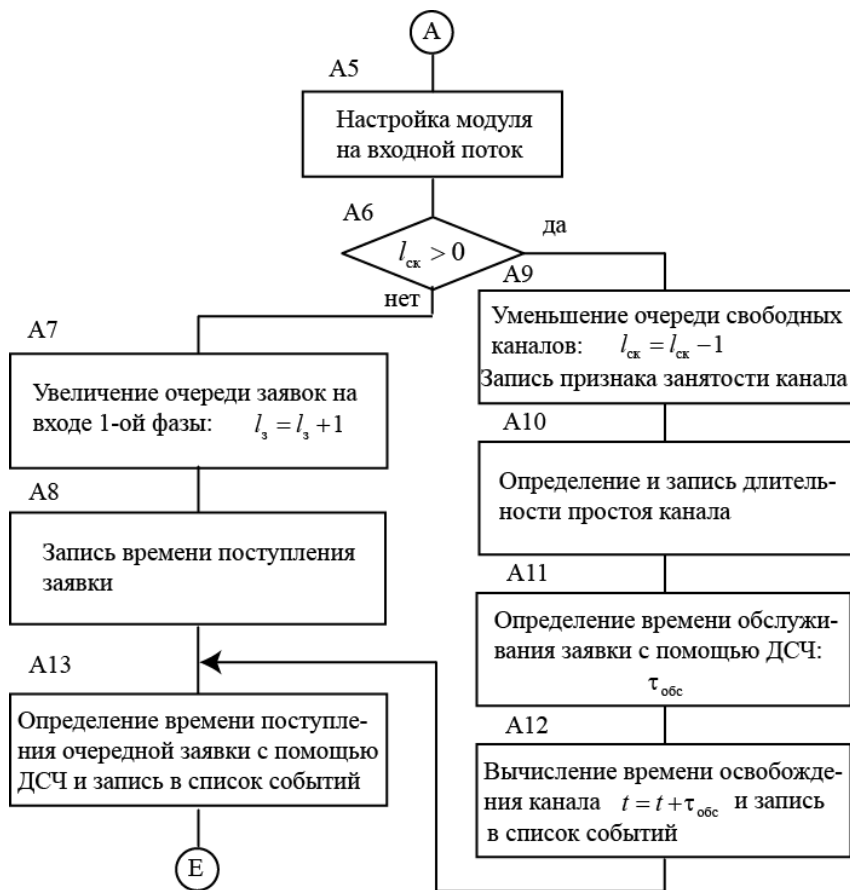


Рис. 3.34. Схема алгоритма модуля реакции А

Если все каналы заняты ($l_{\text{ск}} = 0$), управление передается блоку А7. Если свободные каналы есть $l_{\text{ск}} > 0$, управление передается блоку А9.

Блок А7 – размещение поступившей заявки в соответствующей очереди первой фазы ($l_3 + 1$). Текущее содержимое этой очереди обозначено как l_3 .

Блок А8 – запоминание времени постановки заявки в очередь и передача управления блоку А13.

Правая ветвь алгоритма (см. рис. 3.34) выполняет действия, связанные с постановкой заявки на обслуживание.

Блок А9 – имитация загрузки одного из свободных каналов поступившей заявкой. Очередь свободных каналов уменьшается на 1 ($l_{\text{ск}} - 1$). Выбранному каналу присваивается признак занятости – 1. У свободных каналов этот признак равен 0.

Блок А10 – определение и запоминание длительности простоя канала. Накопление времени простоя необходимо, например, для определения коэффициента занятости канала на всем интервале исследования.

Блок А11 – имитация времени обслуживания заявки.

Случайное время обслуживания $\tau_{\text{обс}}$ формируется обращением к соответствующему датчику случайных чисел.

Блок А12 – прогноз времени окончания обслуживания $t + \tau_{\text{обс}}$ и занесение новой пространственно-временной точки в список событий.

Блок А13 – прогноз момента времени поступления очередной заявки. Соответствующий датчик случайных чисел выдает длительность случайного временного интервала между заявками данного типа. Момент поступления очередной заявки вычисляется как $t + \tau_{\text{инт}}$. Полученная таким образом пространственно-временная точка заносится в список событий.

Управление передается в точку Е – блоку 3 для определения очередного ближайшего события.

Если очередным ближайшим событием окажется освобождение канала не последней фазы, то блок 5 передает управление модулю реакции В, имитирующему действия в СМО при свершении этого активного события.

Схема алгоритма модуля реакции В представлена на рис. 3.35.

Модуль В имеет две функции. Во-первых, попытаться загрузить освободившийся канал заявкой из очереди на входе фазы, в которой находится только что освободившийся канал. Во-вторых, имитировать продвижение заявки, освободившей этот канал, в следующую фазу обслуживания.

Блок В5 – настройка модуля на освободившийся канал конкретной фазы.

Каждая фаза обслуживания имеет свою очередь заявок на входе, свою очередь свободных каналов и, возможно, свои датчики случайных чисел. Поэтому не будем вводить для них новые обозначения, оставим те, которые были использованы при описании модуля А.

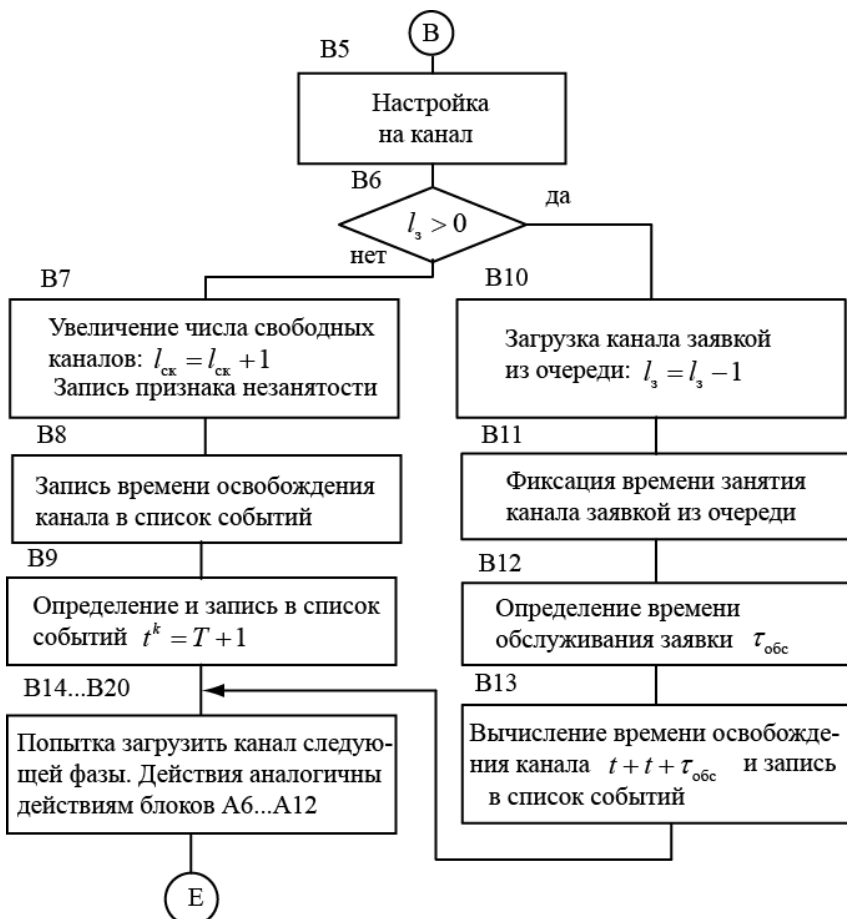


Рис. 3.35. Схема алгоритма модуля реакции В

Блок В6. Проверка: есть ли в очереди на входе фазы заявки, ожидающие обслуживания ($l_3 > 0$)? Если есть, то освободившийся канал должен быть немедленно загружен (блоки 10...13). Если нет, то канал должен быть переведен в режим ожидания (блоки В7...В9).

Блок В7. Очередь свободных каналов увеличивается на 1 ($l_{ск} + 1$). Признак занятости канала устанавливается в нуль.

Блок В8. Запоминание момента освобождения канала.

Блок В9. Перевод освободившегося канала в пассивное состояние. Для этого пространственно-временной точке t^k , относящейся к осво-

бодившемуся каналу, присвоить значение времени, превышающего интервал исследования T , то есть $t^k = T + 1$. Это значение t^k заносится в список событий. Управление передается блоку В14.

Блоки **В14...В20** моделируют размещение заявки, освободившей канал, в следующей фазе обслуживания. Функции и взаимные связи этих блоков аналогичны блокам А6 ... А12, рассмотренным ранее:

В14 → А6; В15 → А7; В16 → А8; В17 → А9; В18 → А10; В19 → А11; В20 → А12.

Блок 20 передает управление блоку 3 для идентификации очередного активного события. Этим событием может быть поступление заявки от какого-либо источника на вход первой фазы или окончание обслуживания заявки каким-либо каналом на первой или других фазах. В этих случаях управление будет снова передано модулям реакции А или В.

Если же очередным активным событием окажется освобождение канала последней фазы, то управление передается модулю реакции С (рис. 3.36).

Блок С5 – настройка на соответствующий канал последней фазы. Аналогичен блоку В5.

Блоки С6...С13 – загрузка освободившегося канала ожидающей заявкой, если таковая есть на входе последней фазы. Функции и взаимосвязи блоков аналогичны блокам В6...В13. Блок С6 соответствует блоку В6, блок С7 – блоку В7 и т. д. по одинаковым номерам.

Блок С14 формирует числовые атрибуты заявок, покидающих СМО. Такими атрибутами могут быть: суммарное время обслуживания заявки на всех фазах, суммарное время ожидания в очередях, число обслуживаний без ожидания и др. Атрибуты запоминаются и управление передается блоку 2.

Если очередное значение $\min t^k = T$, что означает окончание интервала исследования работы СМО, управление передается модулю реакции Θ (рис. 3.37).

Построить имитационную модель с продвижением времени можно и по-другому. Однако рассмотренная структура модели обладает важными достоинствами:

1. Моделирующий алгоритм нагляден и прост, что существенно упрощает программирование и, особенно, отладку модели, а также изучение готовых моделей.

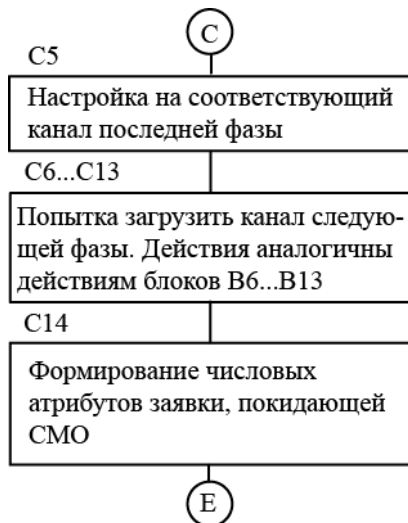


Рис. 3.36. Схема алгоритма модуля реакции С

2. Рассмотренная структура алгоритма позволяет параллельно разрабатывать модель по отдельным модулям и, следовательно, сокращать время на ее создание.



Рис. 3.37. Схема алгоритма модуля реакции Θ

3. Облегчены корректировка и развитие модели. Например, в рассмотренной модели СМО легко учесть выход из строя и восстановление каналов обслуживания. Для этого надо добавить два модуля реакции: один – на возникновение отказа, другой – на восстановление работоспособности. Оба эти события рассматриваются как активные, если неисправность возникает из-за внутренних дефектов оборудования, а восстановление является результатом усилий обслуживающего персонала, входящего в состав СМО.

Вопросы для самоконтроля

3.1. Что такое имитационная статистическая модель? Сравните её с аналитической моделью.

3.2. Назначение датчиков случайных чисел (генераторов) в имитационном моделировании.

3.3. Способы формирования случайных чисел в алгоритмических датчиках случайных чисел.

3.4. Почему случайные числа, формируемые в компьютере, являются псевдослучайными квазиравномерными?

3.5. Способы проверки датчиков равномерно распределённых случайных чисел.

3.6. Формирование случайных величин с произвольными законами распределения вероятностей методом обратной функции.

3.7. Формирование равномерно распределённых случайных чисел на произвольном отрезке $[a, b]$.

3.8. Формирование нормально распределённых случайных чисел с произвольными значениями m и σ .

3.9. Создать программу мультипликативного датчика равномерно распределённых случайных чисел для 64-разрядной сетки:

$$M = 2^{63} = 9\ 223\ 372\ 036\ 854\ 775\ 808; \lambda = 2^{32} + 3 = 4\ 294\ 967\ 299;$$

$$X_i = \lambda = 4\ 294\ 967\ 299.$$

Определить период повторения случайных чисел.

3.10. Усовершенствуйте алгоритм имитационной модели (рис. 3.5) так, чтобы можно было определять не только оценку математического ожидания вероятности поражения объекта, но и оценку математического ожидания расхода ракет.

3.11. Можно ли на компьютере формировать непрерывные случайные последовательности? Если нет, то почему?

- 3.12. Моделирование единичных событий.
- 3.13. Моделирование полной группы несовместных событий.
- 3.14. Алгоритмы моделирования полной группы несовместных событий. Их достоинства и недостатки.
- 3.15. Моделирование совместных независимых событий.
- 3.16. Алгоритмы моделирования совместных независимых событий. Их достоинства и недостатки.
- 3.17. Моделирование совместных зависимых событий.
- 3.18. Алгоритм моделирования совместных зависимых событий.
- 3.19. Что понимают под случайным процессом?
- 3.20. Моделирование случайных процессов. Сечение. Реализация.
- 3.21. Возможно ли на компьютере моделирование непрерывного случайного процесса? Если нет, то почему?
- 3.22. Стационарный и нестационарный случайные процессы.
- 3.23. Эргодический и неэргодический случайные процессы.
- 3.24. Способы продвижения модельного времени в имитационной модели. Их достоинства и недостатки.
- 3.25. Какие виды времени различают при имитационном моделировании?
- 3.26. В чём заключается масштабирование при имитационном моделировании?
- 3.27. Что значит квазипараллельное моделирование?
- 3.28. Постановка задачи на разработку имитационной модели высокоорганизованного противоборства двух сторон.
- 3.29. Модель противоборства двух сторон, назначение блоков. Общая идея построения модели.
- 3.30. Модуль имитации противоборства двух сторон, назначение блоков, работа модуля.
- 3.31. Поясните приём моделирования противоборства двух сторон методом «блуждания по решётке».
- 3.32. Понятие активного и пассивного элемента в модели СМО.
- 3.33. Схема имитационной модели СМО с продвижением времени до ближайшего события.
- 3.34. Что понимается в имитационной модели под списком событий и как он формируется?
- 3.35. Модули реакции в модели СМО. Приведите примеры.
- 3.36. Достоинства структуры модели СМО с продвижением времени до ближайшего события.

Глава 4. Планирование экспериментов

Создание модели – акт необходимый при анализе и синтезе сложных систем, но далеко не конечный. Модель – не цель исследователя, а только инструмент для проведения исследований, инструмент эксперимента. В первых темах мы достаточно полно раскрыли афоризм: «Модель есть объект и средство эксперимента».

Эксперимент должен быть информативен, то есть давать всю нужную информацию, которой следует быть полной, точной, достоверной. Но она должна быть получена приемлемым способом. Это означает, что способ должен удовлетворять экономическим, временным и, возможно, другим ограничениям. Такое противоречие разрешается с помощью рационального (оптимального) планирования эксперимента.

Теория планирования эксперимента сложилась в шестидесятые годы двадцатого века благодаря работам выдающегося английского математика, биолога, статистика Рональда Айлмера Фишера (1890–1962 гг.). Одно из первых отечественных изданий: Федоров В. В. Теория оптимального эксперимента, 1971 г. Несколько позже сложилась теория и практика планирования имитационных экспериментов, элементы которых рассматриваются в данной теме. Раскрывается сущность планирования компьютерного эксперимента, его необходимость и основные пути достижения приемлемых информативности и экономичности, точности и достоверности результатов моделирования.

4.1. Сущность и цели планирования эксперимента

Итак, как мы уже знаем, модель создаётся для проведения на ней экспериментов. Будем считать, что эксперимент состоит из *наблюдений*, а каждое наблюдение – из *прогнозов (реализаций) модели*.

Для организации экспериментов наиболее важно следующее.

1. Простота повторений условий эксперимента.
2. Возможность управления экспериментом, включая его прерывание и возобновление.
3. Легкость изменения условий проведения эксперимента (воздействий внешней среды).
4. Исключение корреляции между последовательностями данных, снимаемых в процессе эксперимента с моделью.
5. Определением интервала времени исследования модели.

Компьютерный эксперимент с имитационной моделью обладает преимуществами перед натурным экспериментом.

Что же такое компьютерный (машинный) эксперимент?

Компьютерный эксперимент представляет собой процесс использования модели с целью получения и анализа интересующей исследователя информации о свойствах моделируемой системы.

Эксперимент требует затрат труда и времени и, следовательно, финансовых затрат. Чем больше мы хотим получить информации от эксперимента, тем он дороже.

Средством достижения приемлемого компромисса между максимумом информации и минимумом затрат ресурсов является план эксперимента.

План эксперимента определяет:

объем вычислений на компьютере;

порядок проведения вычислений на компьютере;

способы накопления и статистической обработки результатов моделирования.

Планирование экспериментов имеет следующие цели:

сокращение общего времени моделирования при соблюдении требований к точности и достоверности результатов;

увеличение информативности каждого наблюдения;

создание структурной основы процесса исследования.

Итак, план эксперимента на компьютере представляет собой метод получения с помощью эксперимента необходимой информации.

Конечно, можно проводить исследования и по такому плану: исследовать модель во всех возможных режимах, при всех возможных сочетаниях внешних и внутренних параметров, повторять каждый эксперимент десятки тысяч раз – чем больше, тем точнее!

Очевидно, что пользы от такой организации эксперимента мало, так как полученные данные трудно обзреть и проанализировать. Кроме того, существенными будут затраты ресурсов, а они всегда ограничены.

Весь комплекс действий по планированию эксперимента разделяют на две самостоятельные функциональные части:

- стратегическое планирование;
- тактическое планирование.

Стратегическое планирование позволяет ответить на вопрос, при каком сочетании уровней внешних и внутренних факторов может

быть получена наиболее полная и достоверная информация о поведении системы.

Тактическое планирование обеспечивает достижение заданных точности и достоверности результатов.

4.2. Элементы стратегического планирования экспериментов

Формирование стратегического плана выполняется в факторном пространстве. *Факторное пространство* – это множество внешних и внутренних параметров, значения которых исследователь может контролировать в ходе подготовки и проведения эксперимента. На рис. 4.1 изображена плоскость, являющаяся факторным пространством двух факторов.

Объектами стратегического планирования являются:

выходные переменные (отклики, реакции, экзогенные переменные);

входные переменные (факторы, эндогенные переменные);

уровни факторов.

Математические методы планирования экспериментов основаны на так называемом кибернетическом представлении процесса проведения эксперимента (рис. 4.2).

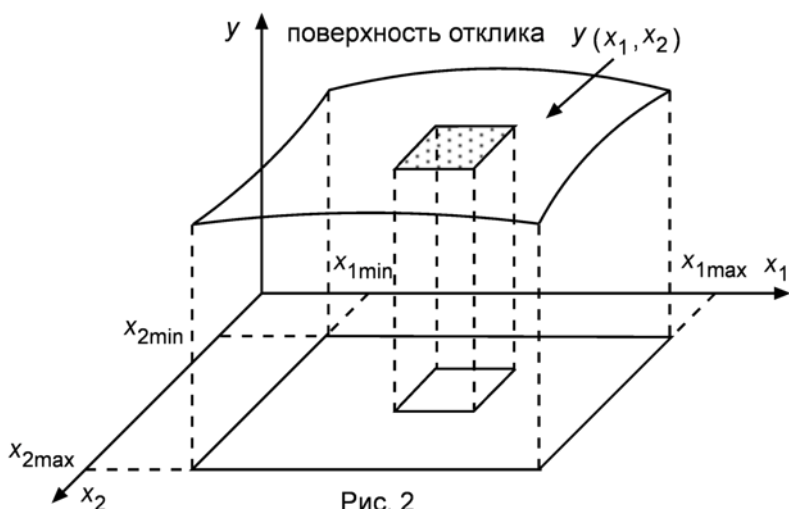


Рис. 4.1. Плоскость факторного пространства двух факторов

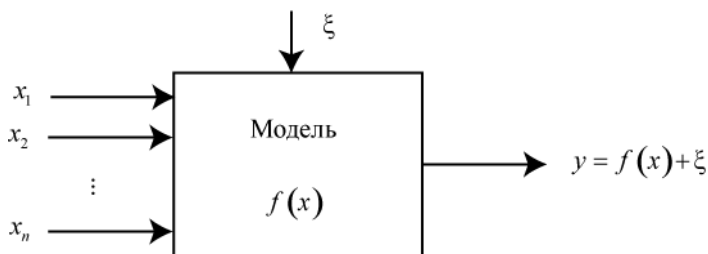


Рис. 4.2. Кибернетическое представление эксперимента

На рис. 4.2:

$x_i, i = \overline{1, n}$ – входные переменные, факторы;

$y = f(x) + \xi$ – выходная переменная (реакция, отклик);

ξ – ошибка, помеха, вызываемая наличием случайных факторов;

$f(x)$ – оператор, моделирующий действие реальной системы, определяющий зависимость выходной переменной y от факторов x_i .

Иначе: $f(x)$ – модель процесса, протекающего в системе.

Первой проблемой, решаемой при стратегическом планировании, является выбор отклика (реакции), то есть определение, какие величины нужно измерять во время эксперимента, чтобы получить искомые ответы. Естественно, выбор отклика зависит от цели исследования.

Например, при моделировании информационной системы пункта управления может интересовать исследователя время ответа системы на запрос. Но может интересовать и такой показатель как максимальное число обслуженных запросов за интервал времени. А может, то и другое. Измеряемых откликов может быть достаточно много: y_1, y_2, \dots . В дальнейшем будем говорить об одном отклике y .

Второй проблемой стратегического планирования является выбор (определение) существенных факторов и их сочетаний, влияющих на работу моделируемого объекта. Факторами могут быть питающие напряжения, температура, влажность, ритмичность поставок комплектующих и многое другое. Обычно число факторов велико и чем меньше мы знакомы с моделируемой системой, тем большее, нам кажется, число их влияет на работу системы. В теории систем приводится так называемый принцип Парето:

20% факторов определяют 80% свойств системы;

80% факторов определяют 20% свойств системы.

Следовательно, надо уметь выделять существенные факторы. А это достигается достаточно глубоким изучением моделируемого объекта и протекающих в нём процессов.

Факторы могут быть количественными и (или) качественными.

Количественные факторы – это те факторы, значения которых числа. Например, интенсивности входных потоков и потоков обслуживания моделируемой системы (объекта), ёмкость буфера (накопителя), число каналов в СМО, вероятность попадания в цель, вероятность обнаружения цели и др.

Качественные факторы – дисциплины обслуживания (LIFO, FIFO и др.) в СМО, квалификация личного состава, его отношение к выполнению поставленных задач и т. п.

Фактор должен быть управляемым.

Управляемость фактора – это возможность установки и поддержания значения фактора постоянным или изменяющимся в соответствии с планом эксперимента. Возможны и неуправляемые факторы, например, влияние внешней среды.

К совокупности воздействующих факторов предъявляются два основных требования:

совместимость;

независимость.

Совместимость факторов означает, что все комбинации значений факторов осуществимы.

Независимость факторов определяет возможность установления значения фактора на любом уровне независимо от уровней других факторов.

В стратегических планах факторы обозначают латинской буквой x_i , где индекс i указывает номер (тип) фактора. Встречаются и такие обозначения факторов: A, B, C, \dots и т. д.

Третьей проблемой стратегического планирования является выбор значений каждого фактора, называемых *уровнями фактора*.

Число уровней может быть два, три и более. Например, если в качестве одного из факторов выступает температура, то уровнями могут быть: $80^\circ\text{C}, 100^\circ\text{C}, 120^\circ\text{C}$.

Для удобства и, следовательно, удешевления эксперимента количество уровней нужно выбирать поменьше. Но достаточное для обеспечения требуемых точности и достоверности эксперимента. Минимальное число уровней – два.

С точки зрения удобства планирования эксперимента целесообразно устанавливать одинаковое число уровней у всех факторов. Такое планирование называют *симметричным*.

Анализ данных эксперимента существенно упрощается, если назначить уровни факторов, равноотстоящие друг от друга. Такой план называется *ортогональным*. Ортогональность плана обычно достигают так: две крайние точки области изменения фактора выбирают как два уровня, а остальные уровни располагают так, чтобы они делили полученный отрезок на две части.

Например, диапазон питающего напряжения 30...50 В на пять уровней будет разбит так: 30 В, 35 В, 40 В, 45 В, 50 В.

Эксперимент, в котором реализуются все сочетания уровней всех факторов, называется *полным факторным экспериментом* (ПФЭ).

План ПФЭ предельно информативен, но он может потребовать неприемлемых затрат ресурсов.

Если отвлечься от компьютерной реализации плана эксперимента, то число измерений откликов (реакций) модели N_C при ПФЭ равно:

$$N_C = q_1 \cdot q_2 \cdot \dots \cdot q_k$$

где q_i – число уровней i -го фактора, $i = \overline{1, k}$;

k – число факторов эксперимента.

Величина N_C определяет структуру стратегического плана, то есть количество наблюдений (информационных точек).

При машинной реализации ПФЭ в каждом наблюдении (информационной точке) нужно выполнить определённое число прогонов (реализаций) модели, чтобы обеспечить заданную точность и достоверность значений откликов. Определение числа прогонов модели является предметом тактического планирования.

Обозначим число прогонов в каждом наблюдении p . Тогда для симметричного ПФЭ число N необходимых прогонов модели равно:

$$N = pq^k.$$

Пример 4.1. Планируется провести компьютерный эксперимент, в котором на отклик модели влияют три фактора. Для каждого фактора установлены три уровня. Требования по точности и достоверности требуют 6000 прогонов модели на каждом уровне (для каждого наблюдения). Время одного прогона модели равно 2 с.

Оценить затраты времени на проведение эксперимента.

Решение.

Исходные данные: $k = 5, q = 2, p = 6000, t_p = \frac{2}{60} = \frac{1}{30}$ мин.

Число прогонов модели: $N = pq^k = 6000 \cdot 2^5 = 6000 \cdot 32 = 192000$.

Затраты времени: $T = \frac{N \cdot t_p}{60} = \frac{192000 \cdot 1}{60 \cdot 30} \approx 106$ ч.

Итак, при стратегическом планировании выбираются:

отклики (реакции), то есть, какие характеристики (показатели) нужно измерять во время эксперимента, чтобы получить искомые ответы;

существенные факторы и их сочетания, влияющие на функционирование моделируемого объекта;

значения каждого фактора, называемые *уровнями фактора*.

Ранее на рис. 4.2 был показан оператор преобразования – математическая модель процесса $f(x)$. В некоторых исследованиях, если $f(x)$ – имитационная модель процесса, требуется вместо неё получить так называемую «вторичную модель» в виде аналитической зависимости. В дальнейшем «вторичная модель» может быть использована на практике или в других исследованиях. В таких случаях математическая модель формируется по данным эксперимента методом регрессионного анализа, что будет предметом нашего рассмотрения в следующей теме.

4.3. Стандартные планы

Многие планы экспериментов в настоящее время стандартизованы. Они имеются в справочниках, математических пакетах программ и системах моделирования. Однако исследователь должен быть готов при необходимости к модификации имеющихся стандартных планов, то есть к приспособлению их к специфическим условиям своих конкретных задач.

С полным факторным экспериментом мы уже знакомы. Это, как отмечалось ранее, самый информативный план, понятный по структуре, но и самый неэкономичный. Поэтому ПФЭ применяют тогда, когда число факторов невелико. В приведенном примере 4.1 при $k = 5, q = 2, p = 6000, t_p = 2 / 60 = 1 / 30$ мин затраты времени на проведение компьютерного эксперимента ожидаются в 106 ч.

Поэтому актуальной становится проблема более или менее обоснованного сокращения плана эксперимента (числа наблюдений). Способов сокращения плана и, следовательно, уменьшения затрат времени на проведение экспериментов, много, но все они, в конечном счёте, основаны на пренебрежении эффектами парных, тройных и более взаимодействий факторов. Естественно, это снижает точность моделирования, но во многих случаях допустимо.

Рассмотрим несколько примеров.

Пример 4.2. Необходимо провести эксперимент с моделью, имеющей три двухуровневых фактора, с целью построения математической модели («вторичной модели») процесса в виде:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 + b_{123}x_1x_2x_3.$$

Уравнение имеет восемь коэффициентов, следовательно, достаточно провести восемь наблюдений. Это уравнение соответствует ПФЭ типа $N = 2^3$.

Полный факторный эксперимент дает возможность определить не только коэффициенты b_1, b_2, b_3 , соответствующие так называемым *линейным эффектам* (их также называют *главными*), но и коэффициенты $b_{12}, b_{13}, b_{23}, b_{123}$, соответствующие всем эффектам взаимодействия факторов, а также свободный член b_0 .

Эффекты взаимодействия двух и более факторов проявляются, если влияние каждого из них на отклик зависит от уровней, на которых установлены другие факторы.

Теперь допустим, что число наблюдений в эксперименте, равное восьми, неприемлемо и план надо сократить.

Вполне естественно предположить, что эффекты взаимодействия оказывают на реакцию системы существенно меньшее влияние, чем линейные, или даже отсутствуют вовсе, если факторы обладают свойством независимости.

Исключим их и тогда модель процесса (уравнение отклика, уравнение реакции, «вторичная модель») принимает вид:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3.$$

Теперь число неизвестных коэффициентов b_i сократилось вдвое и число необходимых наблюдений для их определения стало равно четырём.

Что это за наблюдения?

Четыре наблюдения достаточны для проведения ПФЭ при двух факторной модели $N = 2^2$. Этими факторами, например, могут быть x_1, x_2 или другая двухфакторная комбинация из трёх факторов.

Уровни третьего фактора x_3 получают из первых двух с помощью, так называемого *генерирующего соотношения*:

$$x_3 = x_1 \cdot x_2.$$

Поскольку факторы двухуровневые, то в общем виде уровни путём перехода к безразмерной системе координат принято обозначать:

верхний уровень: +1;

нижний уровень: -1.

Новый, сокращенный план эксперимента называют *полуреplikой* и обозначают 2^{3-1} . План приведен в табл. 4.1.

Таблица 4.1

План ПФЭ 2^2

| № | x_0 | План ПФЭ 2^2 | | $x_3 = x_1 \cdot x_2$ | Отклик y_i |
|---|-------|----------------|-------|-----------------------|--------------|
| | | x_1 | x_2 | | |
| 1 | 1 | -1 | -1 | +1 | y_1 |
| 2 | 1 | +1 | -1 | -1 | y_2 |
| 3 | 1 | -1 | +1 | -1 | y_3 |
| 4 | 1 | +1 | +1 | +1 | y_4 |

Единичный столбец x_0 обеспечивает вычисление свободного члена b_0 в модели объекта (процесса).

Таким же образом можно проводить дальнейшее сокращение планов типа 2^{k-1} , получая четверть реплики 2^{k-2} и более мелкие реплики.

Естественно, такое сокращение числа экспериментов приводит к «огрублению» коэффициентов b_i . Следовательно, полученную модель процесса $y = f(x)$ нужно проверять на адекватность.

Рассмотренное планирование является основой и составной частью для разработки более сложных – несимметричных многоуровневых планов.

Не менее часто целью экспериментов является проверка разного рода гипотез о природе сравниваемых объектов. Например, однородны ли выходы двух систем в смысле законов распределения, характеристик этих законов. Поскольку обработка данных эксперимента ведется методами дисперсионного анализа, то и планы в данном случае называются планами дисперсионного анализа. Сущность дисперсионного анализа мы рассмотрим в следующей теме.

Планы дисперсионного анализа могут быть полные, если используются все возможные сочетания условий (аналогично ПФЭ), и неполные, которые применяются тогда, когда полные планы оказываются громоздкими и неэкономичными. Сокращение дисперсионных планов происходит, как и ранее, за счёт исключения некоторых сочетаний факторов (взаимодействий) и уровней случайным или традиционным образом.

Наиболее популярным из неполных планов является «латинский квадрат». Этот план целесообразно применять, когда из всех существенных факторов можно выделить один доминирующий (самый существенный) фактор. Все факторы должны варьироваться на одинаковом числе уровней. Латинские квадраты можно накладывать друг на друга, образуя греко-латинские квадраты.

В планах дисперсионного анализа часто факторы обозначают латинскими буквами A, B, C, \dots , а уровни – индексами при соответствующих факторах: A_1, A_2, B_1 .

Пример 4.3. Построить план «латинский квадрат» симметричного трёхфакторного четырёхуровневого эксперимента. Доминирующий фактор A .

Решение

Исходные данные: $k = 3, q = 4$.

Введём обозначения факторов и уровней:

A_1, A_2, A_3, A_4 – уровни доминирующего фактора A ;

B_1, B_2, B_3, B_4 – уровни фактора B ;

C_1, C_2, C_3, C_4 – уровни фактора C .

План приведен в табл. 4.2.

В этом плане число наблюдений $N = 4 \cdot 4 = 16$. В полном плане их было бы $N = 4^3 = 64$. Сокращение произошло за счёт исключения некоторых комбинаций: $A_1B_2C_1, A_1B_2C_2$ и др. План позволяет получить несмещённые оценки главных эффектов.

План «латинский квадрат»

| Уровни B | Уровни C | | | |
|------------|------------|-------|-------|-------|
| | C_1 | C_2 | C_3 | C_4 |
| B_1 | A_1 | A_2 | A_3 | A_4 |
| B_2 | A_2 | A_3 | A_4 | A_1 |
| B_3 | A_3 | A_4 | A_1 | A_2 |
| B_4 | A_4 | A_1 | A_2 | A_3 |

Планы латинских (греко-латинских) квадратов используются в тех случаях, когда требуется оценить влияние факторов, варьируемых более чем на двух уровнях и заранее известно, что между факторами нет взаимодействий или ими можно пренебречь. Имеются таблицы различных размеров, но не существует греко-латинского квадрата для шести уровней.

В практике планирования экспериментов встречаются и такие неполные планы: один из факторов меняет свои значения при фиксированных значениях других. То есть исследуется поочередно влияние каждого фактора в отдельности.

Иногда применяются и так называемые *рандомизированные* планы. В таких планах сочетания факторов и уровней для каждого прогона модели выбираются случайно. Вид случайности и объём выборки определяется исследователем.

Когда используется выборка меньшая, чем требует полный факторный эксперимент, плата за это осуществляется смешиванием эффектов. Исследователь, измеряя один эффект, не может сказать главный это эффект или эффект взаимодействия:

$$\begin{aligned}
 I = BCE = ADE = ABCD, & \quad A = DE = BCD = ABCE, \\
 B = CE = ACD = ABDE, & \quad C = BE = ABD = ACDE, \\
 D = AE = ABC = BCDE, & \quad E = BC = AD = ABCDE, \\
 AB = CD = ACE = BDE, & \quad AC = BD = ABE = CDE.
 \end{aligned}$$

В приведенном примере вследствие сокращения плана эксперимента невозможно отличить, например, главный это эффект от фактора A или эффект взаимодействий $DE = BCD = ABCE$. Как видно, то же самое происходит и с эффектами взаимодействий, например, эффект взаимодействия AC или взаимодействий $BD = ABE = CDE$.

4.4. Формальный подход к сокращению общего числа прогонов

Рассмотренные способы сокращения общего числа прогонов носят эвристический (субъективный) характер. Они осуществлялись за счёт исключения каких-то комбинаций уровней факторов.

Однако во многих случаях исследователь имеет свободу действий в выборе числа факторов k , числа уровней q и числа прогонов p модели в одном наблюдении. Каждый из этих аргументов в конкретной ситуации по-разному влияет на общее число прогонов модели N .

Исследуем эти влияния.

Как нам уже известно, общее число прогонов (реализаций) модели равно:

$$N = pq^k.$$

Рассмотрим относительное влияние аргументов p, q, k на число реализаций N .

Сначала нужно получить выражения для вычисления скоростей изменения функции N при изменении одного аргумента и неизменных остальных аргументах. Для этого последовательно найдем частные производные первого порядка от функции N по этим аргументам:

$$1. \frac{\partial N}{\partial k} = pq^k \ln q; \quad 2. \frac{\partial N}{\partial q} = pkq^{k-1}; \quad 3. \frac{\partial N}{\partial p} = q^k.$$

Теперь сравним попарно полученные производные:

$$1. \frac{\partial N}{\partial k} / \frac{\partial N}{\partial q} = \frac{q \ln q}{k}; \quad 2. \frac{\partial N}{\partial p} / \frac{\partial N}{\partial q} = \frac{q}{kp}; \quad 3. \frac{\partial N}{\partial p} / \frac{\partial N}{\partial k} = \frac{1}{p \ln q}.$$

Из соотношений 1 и 2 следует: если $(kp > q)$ и $(k < q \ln q)$, то наибольшее влияние на N оказывает изменение числа уровней q .

Из соотношений 3 и 1 следует: если $(p \ln q > 1)$ и $(k < q \ln q)$, то наибольшее влияние на N оказывает изменение числа факторов k .

Из соотношений 2 и 3 следует: если $(q > kp)$ и $(1 - p \ln q)$, то наибольшее влияние на число N оказывает изменение числа реализаций модели на каждом уровне факторов (на каждом наблюдении).

Рассмотренный формальный подход к сокращению числа реализаций не совсем корректен, так как функция общего числа прогонов N носит не непрерывный, а дискретный характер. Тем не менее, такой

подход применяется с последующим округлением результатов до целых чисел.

Покажем применение формального подхода сокращения реализаций на примере.

Пример 4.4. На вход модели объекта действуют четыре трёхуровневых фактора ($k = 4, q = 3$). В каждом наблюдении предполагаются восемь прогонов модели ($p = 8$). Полный факторный эксперимент потребует $N = pq^k = 8 \cdot 3^4 = 648$ прогонов или 81 наблюдение. Такие затраты ресурсов неприемлемы.

Требуется определить, какой из аргументов (k, q, p) следует уменьшить, чтобы достичь наиболее существенного уменьшения числа реализаций N .

Решение

Подготовим данные для сравнений:

$$q \ln q = 3 \ln 3 = 3,3; \quad kp = 4 \cdot 8 = 32; \quad p \ln q = 8 \cdot 1,1 = 8,8.$$

Соблюдается условие:

$$(kp > q) \wedge (k > q \ln q), \text{ так как } (32 > 3) \wedge (4 > 3,3).$$

Следовательно, наибольшее влияние на изменение N оказывает изменение числа уровней q .

Уменьшим q на единицу: $q = 3 - 1 = 2$. В этом случае при ПФЭ потребуется выполнить $N = 8 \cdot 2^4 = 128$ прогонов или 16 наблюдений, то есть в пять раз меньше.

Варьирование факторов на двух уровнях встречается часто и решение $q = 2$ будет приемлемо, если нет обстоятельств, не устраивающих это решение.

4.5. Элементы тактического планирования

Задачей тактического планирования является обеспечение результатам компьютерного эксперимента требуемых точности и достоверности.

Рассмотрим случай, когда имитационная модель строилась для определения характеристик некоторых случайных величин.

Таковыми случайными величинами могут быть:

время обслуживания заявки в СМО;

численности противоборствующих сторон;

расход боеприпасов;

число обнаруженных средствами разведки объектов противника;
время наработки на отказ технического устройства и др.

Из характеристик случайных величин, как правило, интересуют среднее значение (матожидание), дисперсия и характеристика связи случайных величин – коэффициент корреляции.

Характеристику случайной величины будем обозначать греческой буквой Θ . С помощью имитационного моделирования точное значение Θ определить нельзя, так как число N реализаций модели конечно. При конечном числе реализаций модели определяется приближенное значение характеристики. Обозначим это приближение $\bar{\Theta}$.

Приближенное значение $\bar{\Theta}$ называют *оценкой* соответствующей характеристики: оценкой математического ожидания, оценкой дисперсии, оценкой коэффициента корреляции.

Точностью характеристики $\bar{\Theta}$ называют величину ε в отношении

$$\left| \bar{\Theta} - M[\Theta] \right| < \varepsilon ,$$

где $M[\Theta]$ – математическое ожидание случайной величины.

Величина ε представляет собой абсолютное значение ошибки в определении значения искомой характеристики.

Достоверность оценки характеристики $\bar{\Theta}$ называют вероятностью α того, что заданная точность достигается:

$$P\left(\left|\bar{\Theta} - M[\Theta]\right| < \varepsilon\right) = \alpha. \quad (4.1)$$

Достоверность характеризует повторяемость, устойчивость эксперимента и трактуется так: если для оценки $M[\Theta]$ использовать величину $\bar{\Theta}$, то в среднем на каждые 1000 применений этого правила в $1000 \cdot \alpha$ случаев величина $\bar{\Theta}$ будет отличаться от $M[\Theta]$ на величину меньше ε .

В ряде случаев целесообразно пользоваться понятием относительной точности $d = \frac{\varepsilon}{M[\Theta]}$. Тогда достоверность оценки имеет вид:

$$P\left(\left(\frac{\left|\bar{\Theta} - M[\Theta]\right|}{M[\Theta]}\right) < d\right) = \alpha .$$

4.6. Точность и количество реализаций модели при определении средних значений параметров

Найдем функциональную связь точности ε и достоверности α с количеством реализаций модели, когда в качестве показателей эффективности выступают математическое ожидание и дисперсия некоторой случайной величины (времени, расстояния и т. п.).

4.6.1. Определение оценки математического ожидания

Найдем искомую связь для случая, когда целью эксперимента является определение оценки математического ожидания некоторой случайной величины.

В N прогонах модели получены независимые значения интересующего нас показателя эффективности:

$$a_1, a_2, \dots, a_i, \dots, a_N.$$

В качестве оценки математического ожидания возьмем выборочное среднее (среднее арифметическое):

$$\bar{a} = \left(\sum_{i=1}^N a_i \right) / N.$$

В последующей теме мы покажем, что оценка такого вида является наилучшей.

Согласно центральной предельной теореме, если значения a_i независимы и имеют конечные дисперсии одного порядка, то при большом числе слагаемых N случайная величина \bar{a} имеет практически нормальное распределение с математическим ожиданием и дисперсией соответственно:

$$M[\bar{a}] = M[a], \quad \sigma_a^2 = \frac{\sigma_a^2}{N}, \quad \sigma_{\bar{a}} = \frac{\sigma_a}{\sqrt{N}},$$

где σ_a^2 – дисперсия искомой случайной величины a .

Следовательно, справедливо

$$P\left(\left|\bar{a} - M[a]\right| < t_\alpha \sigma_{\bar{a}}\right) = \Phi^*(t_\alpha),$$

где $\Phi^*(t_\alpha) = \frac{2}{\sqrt{2\pi}} \int_0^{t_\alpha} e^{-\frac{z^2}{2}} dz$ – интеграл вероятности.

В некоторых изданиях под интегралом вероятности понимают несколько иное выражение, поэтому целесообразно пользоваться интегралом Лапласа, который связан с интегралом вероятности так: $\Phi^*(t_\alpha) = 2\Phi(t_\alpha)$. $\Phi(t_\alpha)$ – интеграл Лапласа. Из приведенного выражения следует:

$$P\left(\left|\bar{a} - M[a]\right| < t_\alpha \sigma_a\right) = 2\Phi(t_\alpha).$$

Сравнивая это выражение с выражением (4.1), имеем:

$$\varepsilon = t_\alpha \sigma_a = t_\alpha \frac{\sigma_a}{\sqrt{N}}, \quad 2\Phi(t_\alpha) = \alpha, \quad t_\alpha = \Phi^{-1}\left(\frac{\alpha}{2}\right).$$

Интеграл Лапласа табулирован, следовательно, задавая значением достоверности α , определяется аргумент t_α .

Итак, искомая связь между точностью ε , достоверностью α и числом реализаций модели получена:

$$\varepsilon = t_\alpha \frac{\sigma_a}{\sqrt{N}}, \quad N = t_\alpha^2 \frac{\sigma_a^2}{\varepsilon^2}. \tag{4.2}$$

Из выражений (4.2) следует:

увеличение точности на один порядок (уменьшение ошибки на один порядок) потребует увеличения числа реализаций модели на два порядка;

число необходимых реализаций модели N не зависит от величины искомого параметра a , а зависит от дисперсии σ_a^2 .

Достоверность результата α указана значением аргумента функции Лапласа t_α . Связь значения t_α с достоверностью α находится из таблицы значений функции (интеграла) Лапласа. Наиболее употребительные соответствия t_α и α приведены в табл. 4.3.

Таблица 4.3

Фрагмент таблицы функции (интеграла) Лапласа

| | | | | | | | |
|------------|------|------|------|------|------|-------|-------|
| α | 0,80 | 0,85 | 0,90 | 0,95 | 0,99 | 0,995 | 0,999 |
| t_α | 1,28 | 1,44 | 1,65 | 1,96 | 2,58 | 2,81 | 3,3 |

Чтобы пользоваться формулами (4.2), нужно знать дисперсию σ_a^2 . Очень редки случаи, когда значение дисперсии известно до эксперимента, поэтому возможны два способа предварительного определения дисперсии.

Первый способ. Иногда заранее известен размах значений искомой случайной величины:

$$R = \max a_i - \min a_i .$$

В предположении нормального распределения случайной величины a , можно с использованием «правила трёх сигм» получить приближенную оценку σ_a :

$$\frac{R}{2} \approx 3\sigma_a, \quad \sigma_a \approx \frac{R}{6}, \quad \sigma_a^2 \approx \frac{R^2}{36}.$$

Второй способ. Надо воспользоваться оценкой дисперсии. Для этого необходимо выполнить предварительный прогон модели в количестве $N^* = 1000$ реализаций. С использованием полученного в результате этих прогонов ряда значений показателя $a_i, i = \overline{1, N^*}$, найдём оценку дисперсии:

$$S_a^2 = \frac{1}{N^* - 1} \sum_{i=1}^{N^*} (a_i - \bar{a})^2, \quad S = \sqrt{\frac{1}{N^* - 1} \sum_{i=1}^{N^*} (a_i - \bar{a})^2}.$$

Здесь \bar{a} – среднеарифметическое значение по N^* измерениям. И в этом случае формулы (4.2) имеют вид:

$$N = t_\alpha^2 \frac{S_a^2}{\varepsilon^2}, \quad \varepsilon = t_\alpha \frac{S_a}{\sqrt{N}}. \quad (4.3)$$

Вычисленную дисперсию S_a^2 подставим в формулу для определения числа прогонов модели N . Если в результате расчёта окажется, что выполняется неравенство $N > N^*$, то моделирование должно быть продолжено до выполнения N реализаций. Если же $N \leq N^*$, то моделирование заканчивается. Необходимая точность ε оценки случайной величины a (искомого показателя эффективности) при заданной достоверности α достигнута.

Если в технических условиях задана относительная точность $d = \frac{\varepsilon}{a}$, то формулы (4.3) принимают вид:

$$N = t_\alpha^2 \frac{S_a^2}{d^2 a^2}, \quad d = t_\alpha \frac{S_a}{a\sqrt{N}}.$$

Значение \bar{a} определяется на основании $N^* = 1000$ прогонов модели. Все дальнейшие расчеты аналогичны только что рассмотренным аналитическим выражениям.

Вышеприведенные рассуждения и выражения были справедливы в предположении нормального закона распределения случайной величины a . Если в этом есть сомнение, то для определения связи ε , α и N можно воспользоваться известным неравенством П. Л. Чебышева:

$$P\left(|\bar{a} - M[a]| \geq \varepsilon\right) \leq \frac{\sigma_a^2}{N\sigma^2}.$$

С учетом направления знаков неравенств получим:

$$\frac{\sigma_a^2}{N\sigma^2} = 1 - \alpha \Rightarrow N = \frac{\sigma_a^2}{\varepsilon^2(1 - \alpha)} \Rightarrow \varepsilon = \sqrt{\frac{\sigma_a^2}{N(1 - \alpha)}}.$$

Также как и в предыдущих случаях вместо неизвестной дисперсии σ_a^2 следует использовать ее оценку S_a^2 , вычисленную по данным N^* прогонов модели.

И еще: обратим внимание, что в данном случае достоверность α участвует в формулах в явном виде.

Итак, в выражениях (4.3) мы вместо неизвестной дисперсии σ_a^2 используем её оценку S_a^2 . В этом случае вместо аргумента функции Лапласа t_α надо использовать параметр распределения Стьюдента t_α^* , значения которого зависят не только от уровня достоверности α , но и от числа так называемых степеней свободы $k = N - 1$. Здесь, как и прежде, N – число прогонов модели. Вообще-то, при $N \rightarrow \infty$ распределение Стьюдента стремится к нормальному распределению, но при малом числе прогонов модели t_α^* заметно отличается от t_α .

Для практических целей значения t_α^* можно взять из табл. 4.4.

Таблица значений t_α^*

| k | α | | | | |
|-----|----------|------|------|------|-------|
| | 0,8 | 0,9 | 0,95 | 0,99 | 0,999 |
| 10 | 1,37 | 1,81 | 2,23 | 3,17 | 4,6 |
| 20 | 1,33 | 1,73 | 2,1 | 2,85 | 3,73 |
| 30 | 1,31 | 1,7 | 2,04 | 2,75 | 3,65 |
| 40 | 1,3 | 1,68 | 2,02 | 2,7 | 3,55 |
| 60 | 1,3 | 1,67 | 2,0 | 2,67 | 3,41 |
| 120 | 1,29 | 1,66 | 1,98 | 2,62 | 3,37 |

Из табл. 4.4 видно, что при $k = N - 1 > 120$ значения t_α^* и t_α практически совпадают. Но при меньших значениях N следует пользоваться величиной t_α^* .

4.6.2. Определение оценки дисперсии

Мы научились находить оценку математического ожидания \bar{a} некоторой случайной величины a с заданными точностью и достоверностью. Теперь рассмотрим задачу определения оценки дисперсии S^2 случайной величины a также с заданными точностью и достоверностью.

Опустим вывод и приведём окончательный вид формул для расчёта значений N и ε :

$$N = t_\alpha^2 \frac{(\mu_4 - \sigma^4)}{\varepsilon^2}; \quad \varepsilon = t_\alpha \sqrt{\frac{(\mu_4 - \sigma^4)}{N}},$$

где μ_4 – эмпирический центральный момент четвертого порядка:

$$\mu_4 = \frac{1}{N} \sum_{i=1}^N (a_i - \bar{a})^4.$$

Неизвестное значение σ заменяется оценкой S , как было рассмотрено ранее.

Если определяемая случайная величина имеет нормальное распределение, то $\mu_4 = 3\sigma^4 \approx 3S^4$ и выражения для N и ε принимают вид:

$$N = t_\alpha^2 \frac{2S^4}{\varepsilon^2}; \quad \varepsilon = t_\alpha \frac{S^2 \sqrt{2}}{\sqrt{N}} \quad (4.4)$$

Как и ранее при малых значениях N ($N < 120$) следует использовать параметр распределения Стьюдента t_{α}^* .

Из сопоставления выражений (4.3) и (4.4) следует, что одно и то же количество реализаций модели обеспечит разное значение ошибки ε при оценке математического ожидания случайной величины a и её дисперсии – при одинаковой достоверности. И иначе: одинаковую точность определения оценок математического ожидания и дисперсии случайного параметра при одинаковой достоверности обеспечит разное число реализаций модели.

Пример 4.5. В результате предварительных прогонов модели $N^* = 1000$ определена оценка дисперсии $S^2 = 10 \varepsilon \delta^2$.

Определить число реализаций модели N_1 и N_2 для определения оценок математического ожидания и дисперсии случайной величины a соответственно с точностью $\varepsilon = 0,1$ и достоверностью $\alpha = 0,9$.

Решение.

$$N_1 = t_{\alpha}^2 \frac{S^2}{\varepsilon^2} = 1,65^2 \frac{10}{0,1^2} \approx 2720.$$

$$N_2 = t_{\alpha}^2 \frac{2S^4}{\varepsilon^2} = 1,65^2 \frac{200}{0,1^2} \approx 54450.$$

4.7. Точность и количество реализаций модели при определении вероятностей исходов

Мы рассматриваем случай, когда в качестве показателя эффективности выступает вероятность свершения (или не свершения) какого-либо события, например, поражения цели, выхода из строя техники, завершения комплекса работ в заданное время и др.

В качестве оценки вероятности P события a выступает частота его свершения:

$$\bar{P} = \frac{m}{N},$$

где N – число реализаций модели;

m – число свершений данного события.

Использование частоты \bar{P} в качестве оценки искомой вероятности P основано на теореме Я. Бернулли, которую в данном случае можно в формализованном виде записать так:

$$\lim_{N \rightarrow \infty} \frac{m}{N} = P.$$

Точность и достоверность этой оценки связаны с уже известным определением достоверности:

$$P\left(\left|\bar{P} - P\right| < \varepsilon\right) = \alpha.$$

Задача сводится к нахождению такого количества реализаций N , чтобы оценка \bar{P} отличалась от искомого значения P менее, чем на ε с заданной достоверностью. Здесь, как и ранее, ε – абсолютное значение, характеризующее точность оценки.

Для нахождения функциональной связи между точностью, достоверностью и числом реализаций модели введем переменную x_i – результат исхода i -й реализации модели:

$$x_i = \begin{cases} 1, & \text{событие свершилось, вероятность } P, \\ 0, & \text{событие не свершилось, вероятность } 1 - P. \end{cases}$$

Тогда частота свершения события (оценка искомой вероятности) будет определяться следующим выражением:

$$\bar{P} = \frac{\sum_{i=1}^N x_i}{N}.$$

Величина $\sum_{i=1}^N x_i$ – случайная и дискретная. Она при таком задании x_i имеет биномиальное распределение (распределение Бернулли) с характеристиками:

$$\text{математическое ожидание } M\left[\sum_{i=1}^N x_i\right] = NP;$$

$$\text{дисперсия } D\left[\sum_{i=1}^N x_i\right] = NP(1 - P).$$

Из этого следует:

$$M\left[\bar{P}\right] = M\left[\frac{\sum_{i=1}^N x_i}{N}\right] = \frac{1}{N} NP = P,$$

$$D[\bar{P}] = D\left[\frac{\sum_{i=1}^N x_i}{N}\right] = \frac{1}{N^2} NP(1-P) = \frac{P(1-P)}{N},$$

$$\sigma_{\bar{P}} = \sqrt{D[\bar{P}]} = \frac{P(1-P)}{N}.$$

В теории вероятностей есть теорема Лапласа (частный случай центральной предельной теоремы), сущность которой состоит в том, что при больших значениях числа реализаций N биномиальное распределение достаточно хорошо согласуется с нормальным распределением.

Следовательно, можно записать:

$$P\left(|\bar{P} - P| < t_\alpha \sigma_{\bar{P}}\right) = 2\Phi(t_\alpha),$$

$$P\left(|\bar{P} - P| < t_\alpha \sqrt{\frac{P(1-P)}{N}}\right) = 2\Phi(t_\alpha).$$

Следуя рассуждениям, приведенным ранее, получим искомые формулы:

$$\varepsilon = t_\alpha \sqrt{\frac{P(1-P)}{N}}, \quad N = t_\alpha^2 \frac{P(1-P)}{\varepsilon^2}. \quad (4.5)$$

Как и ранее, t_α – аргумент функции Лапласа, $t_\alpha = \Phi^{-1}\left(\frac{\alpha}{2}\right)$.

Если априорные сведения хотя бы о порядке искомой вероятности P неизвестны, то использование значения абсолютной ошибки ε может не иметь смысла. Например, может быть так, что исследователь задал значение абсолютной ошибки $\varepsilon = 1$, а искомое значение вероятности оказалось $\bar{P} = 0,01$. Очевидно, явное несоответствие. Поэтому целесообразно оперировать относительной погрешностью:

$$d = \frac{\varepsilon}{P}.$$

В этом случае формулы (4.5) принимают вид:

$$d = t_\alpha \sqrt{\frac{1-P}{PN}}, \quad N = t_\alpha^2 \frac{1-P}{Pd^2}. \quad (4.6)$$

Из формул (4.6) следует, что при определении оценок малых вероятностей с приемлемо высокой точностью необходимо выполнить очень большое число реализаций модели. При отсутствии высокопро-

изводительного компьютера применение статистического моделирования становится проблематичным.

Пример 4.6. Вероятность $P = 0,1$.

Определить число реализаций модели и затраты машинного времени для оценки данной вероятности с относительной точностью $d = 1\%$ и достоверностью $\alpha = 0,9$. На выполнение одной реализации модели требуется 5 с.

Решение.

Из табл. 4.3 находим $t_\alpha = 1,65$. Относительная точность $d = 0,01$.

$$N = t_\alpha^2 \frac{1-P}{Pd^2} = 1,65^2 \frac{1-0,1}{0,1 \cdot 0,01^2} = 2,7252 \frac{0,9}{0,1 \cdot 0,01^2} = \frac{2,45}{10^{-5}} = 245000.$$

Если на выполнение одной реализации требуется 5 с, то затраты машинного времени составят $\frac{245\,000 \cdot 5}{3600} \approx 340$ ч.

В формулах (4.5) и (4.6) для вычисления N или $\varepsilon(d)$ присутствует та же неопределенность, которую мы обсуждали ранее: вычисления требуют знания вероятности P , а она до эксперимента неизвестна. Эта неопределенность снимается так.

Выполняется предварительно N^* прогонов модели. Обычно принимают $N^* = 1000$. По данным этих прогонов вычисляют ориентировочное значение оценки вероятности \bar{P}^* , которую и подставляют в формулу вместо вероятности P .

Если окажется $N > N^*$, моделирование следует продолжить до выполнения N реализаций.

Если окажется $N \leq N^*$, то моделирование заканчивается. При этом если $N < N^*$, то следует определить действительную точность ε или d для $N = 1000$ реализаций. Очевидно, в этом случае достигнутая точность будет выше заданной (ошибка меньше заданной).

Но более удобно рассчитывать число реализаций на так называемый наихудший случай.

Вернемся к формуле (4.5)

$$N = t_\alpha^2 \frac{P(1-P)}{\varepsilon^2}.$$

Анализ формулы показывает, что число реализаций модели в зависимости от вероятности P изменяется от 0 (при $P = 0$) до 0 (при $P = 1$), проходя через максимум. Максимальное значение N принимает при вероятности $P = 0,5$:

$$\frac{\partial N}{\partial P} = \frac{t_\alpha^2}{\varepsilon^2} (1 - 2P) = 0, \quad 1 - 2P = 0, \quad P = 0,5.$$

То есть наибольшее число реализаций модели будет тогда, когда искомая вероятность равна 0,5.

В этом случае число реализаций определяется так:

$$N_m = t_\alpha^2 \frac{P(1-P)}{\varepsilon^2} = t_\alpha^2 \frac{0,5(1-0,5)}{\varepsilon^2} = \frac{t_\alpha^2}{4\varepsilon^2}. \quad (4.7)$$

Если в результате моделирования окажется, что искомая вероятность значительно отличается от 0,5 (в любую сторону), то точность моделирования будет выше заданной (ошибка ε меньше). Для определения этой точности следует воспользоваться уже известной формулой (4.5), но при $N = N_m$:

$$\varepsilon = t_\alpha \sqrt{\frac{P(1-P)}{N_m}}. \quad (4.8)$$

Пример 4.7. Сервер обрабатывает запросы, поступающие с автоматизированных рабочих мест (АРМ) с интервалами, распределенными по экспоненциальному закону со средним значением $T_1 = 2$ мин. Вычислительная сложность запросов распределена по нормальному закону с математическим ожиданием $S_1 = 6 \cdot 10^7$ оп и среднеквадратическим отклонением $S_2 = 2 \cdot 10^5$ оп. Производительность сервера по обработке запросов $Q = 5 \cdot 10^5$ оп/с.

Построить алгоритм имитационной модели с целью определения вероятности обработки запросов за время $T = 1$ час. Исследовать зависимость вероятности обработки запросов от интервалов их поступления, вычислительной сложности и производительности сервера.

Решение

Для построения алгоритма имитационной модели введем следующие идентификаторы:

$t1$ – текущее модельное время поступления запроса;

$t2$ – интервал поступления запросов;

$t3$ – текущее модельное время окончания обработки запроса;
 $t4$ – модельное время обработки запроса;
 k – счётчик количества прогонов модели (реализаций);
 p – вероятность обработки запросов;
 M – счётчик количества обработанных запросов;
 N – заданное количество прогонов модели (реализаций);
 R – количество поступивших запросов за N прогонов модели;
 T – время моделирования.

Алгоритм модели приведен на рис. 4.3.

Рассмотрим работу алгоритма.

1 (номер блока, символа). Обнуляются переменные:

$M = 0$ – счётчик количества обработанных запросов за одно наблюдение;

$k = 0$ – счётчик текущего числа прогонов модели в одном наблюдении;

$R = 0$ – счётчик количества запросов за N прогонов модели (в одном наблюдении).

Устанавливаются:

$N = 1000$ – количество прогонов модели (реализаций);

$T = 3600$ – время моделирования (модельное).

2. Обнуляются переменные:

$t1 = 0$ – текущее модельное время поступления запроса;

$t3 = 0$ – текущее модельное время окончания обработки запроса.

Увеличивается на 1:

$k = k + 1$ – значение счётчика количества прогонов модели.

3. Увеличивается на 1:

$R = R + 1$ – значение счётчика поступивших запросов за одно наблюдение (N прогонов модели).

4. Разыгрывается интервал времени поступления запроса:

$$t2 = \exp(T1).$$

Увеличивается текущее модельное время:

$$t1 = t1 + t2.$$

5. Осуществляется проверка: если выполняется первый прогон модели, то есть $k = 1$, то управление передаётся блоку 6.

6. Разыгрывается время обработки запроса:

$$t4 = [Normal(S1, S2,)] / Q.$$

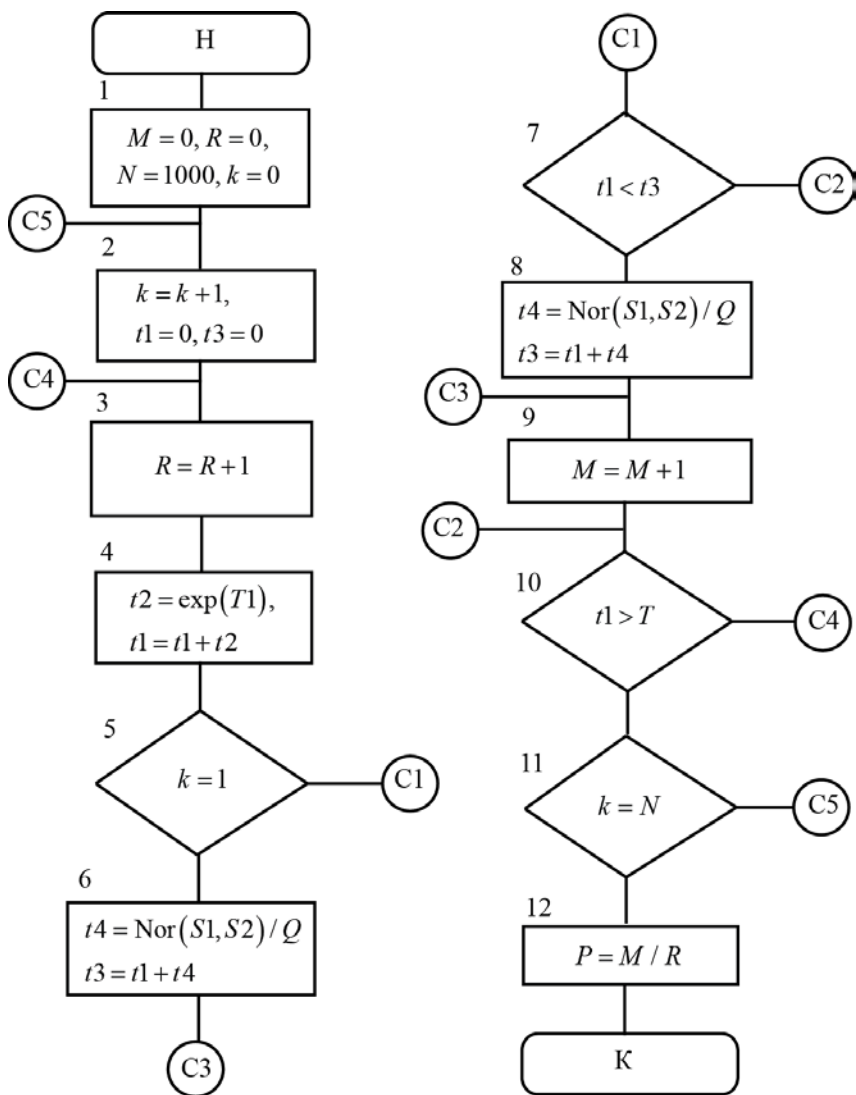


Рис. 4.3. Алгоритм модели обработки запросов сервером

Определяется модельное время окончания обработки запроса – освобождения сервера:

$$t3 = t3 + t4.$$

Управление передаётся блоку 9.

9. Увеличивается на 1 значение счётчика количества обработанных запросов:

$$M = M + 1.$$

10. Проверяется: текущее модельное время больше заданного времени моделирования:

$$t1 > T?$$

Если условие выполняется, а оно не выполнится, так как обрабатывается только первый запрос, управление передаётся блоку 3.

3. Увеличивается на 1:

$$R = R + 1$$

значение счётчика поступивших запросов за одно наблюдение (N прогонов).

4. Разыгрывается интервал времени поступления запроса:

$$t2 = \exp(T1).$$

Увеличивается текущее модельное время;

$$t1 = t1 + t2.$$

Оно становится равным текущему времени ввода в модель очередного запроса.

5. Проверка: если выполняется первый прогон модели, то есть $k = 1$, то управление передаётся блоку 6. Так как выполняется не первый прогон, управление передаётся блоку 7.

Проверяется: текущее модельное время меньше времени окончания обработки запроса, то есть $t1 < t3$? Если меньше, значит сервер занят, запрос теряется. Управление передаётся блоку 10. Если не меньше, то есть сервер свободен, то управление передаётся блоку 8.

8. Разыгрывается время обработки запроса:

$$t4 = [\text{Normal}(S1, S2)] / Q.$$

Определяется модельное время окончания обработки запроса – освобождения сервера:

$$t3 = t1 + t4.$$

Управление передаётся блоку 9.

9. Увеличивается на 1 значение счётчика количества обработанных запросов:

$$M = M + 1.$$

10. Проверяется: текущее модельное время больше заданного времени моделирования:

$$t1 > T?$$

Если условие выполняется, то управление передаётся блоку 11. В противном случае – блоку 3.

11. Проверяется: выполнено ли заданное количество прогонов модели, то есть $k = N$? Если условие не выполняется, то управление передаётся блоку 2. Моделирование продолжается.

В противном случае управление передаётся блоку 12.

12. Расчёт оценки вероятности обработки запросов сервером:

$$P = M / R.$$

Работа алгоритма заканчивается.

Перейдём к стратегическому планированию эксперимента. Выберем интервалы варьирования уровней факторов.

$T_1 = 120 \pm 60$ с – средний интервал поступления запросов.

Для изменения математического ожидания и среднеквадратического отклонения целесообразно ввести коэффициент, принимающий два значения, например, $h_1 = 0,5$ и $h_2 = 1,5$. В этом случае значения математического ожидания и стандартного отклонения будут изменяться пропорционально, благодаря чему не будет совершена ошибка, когда значение стандартного отклонения окажется равным более одной пятой значения математического ожидания.

Тогда $S_1 = 2 \cdot 10^7 \pm 3 \cdot 10^7$ оп, $S_2 = 2 \cdot 10^5 \pm 1 \cdot 10^5$ оп.

$Q = 5 \cdot 10^5 \pm 2 \cdot 10^5$ оп/с – производительность сервера.

В соответствии с интервалами варьирования представим уровни факторов (табл. 4.5), где индексы н и в – нижний и верхний уровни факторов соответственно.

Таблица 4.5

Уровни факторов

| T_1 | | S_1 / S_2 | | Q | |
|----------|----------|-------------------------------|-------------------------------|----------------|----------------|
| $T_{1Н}$ | $T_{1В}$ | $(S_1 / S_2)_{Н}$ | $(S_1 / S_2)_{В}$ | $Q_{Н}$ | $Q_{В}$ |
| 60 | 180 | $3 \cdot 10^7 / 1 \cdot 10^5$ | $9 \cdot 10^7 / 3 \cdot 10^5$ | $3 \cdot 10^5$ | $7 \cdot 10^5$ |
| -1 | +1 | -1 | +1 | -1 | +1 |

Составим план факторного эксперимента (табл. 4.6).

План полного факторного эксперимента

| № | T_1 | S_1 / S_2 | Q | Y | |
|---|-------|-------------|-----|--------------|------------|
| | | | | $N^* = 1000$ | $N = 9604$ |
| 1 | 2 | 3 | 4 | 5 | 6 |
| 1 | -1 | -1 | -1 | 0,375 | 0,375 |
| 2 | -1 | -1 | +1 | 0,584 | 0,583 |
| 3 | -1 | +1 | -1 | 0,166 | 0,167 |
| 4 | -1 | +1 | +1 | 0,320 | 0,319 |
| 5 | +1 | -1 | -1 | 0,640 | 0,642 |
| 6 | +1 | -1 | +1 | 0,809 | 0,808 |
| 7 | +1 | +1 | -1 | 0,376 | 0,375 |
| 8 | +1 | +1 | +1 | 0,584 | 0,583 |

Естественно, план полного факторного эксперимента будет представлен в столбцах 1...4.

Приступим к тактическому планированию эксперимента. Так как до эксперимента искомая оценка вероятности обработки запросов нам не известна, то выполним первое наблюдение при $N^* = 1000$ прогонов модели. Получим вероятность обработки запросов $P = 0,375$. Занесём её в табл. 4.6 (строка 1, столбец 5).

Зададимся точностью $\varepsilon = 0,01$ и доверительной вероятностью $\alpha = 0,95$.

По таблице значений функции Лапласа найдём её аргумент $t_\alpha = 1,96$ (см. табл. 4.3).

По формуле (4.5) рассчитаем требуемое количество прогонов модели при $\varepsilon = 0,01$ и $\alpha = 0,95$:

$$N = t_\alpha^2 \frac{P_0(1-P_0)}{\varepsilon^2} = 1,96^2 \cdot \frac{0,375(1-0,375)}{0,01^2} = 3,8416 \cdot \frac{0,375 \cdot 0,625}{0,01^2} = 9003.$$

При расчёте по формуле (4.7) числа прогонов для «худшего случая» (а такой вариант возможен, так как при выполнении других наблюдений эксперимента, результаты которых занесены в столбец 1 табл.4.6, мы видим, что есть $P \approx 0,5$) получим:

$$N = t_\alpha^2 \frac{0,5 \cdot 0,5}{\varepsilon^2} = 1,96^2 \cdot \frac{0,25}{0,01^2} = 3,8416 \cdot \frac{0,25}{0,01^2} = 9604.$$

4.8. Точность и количество реализаций модели при зависимом ряде данных

До сих пор мы предполагали, что выходные данные модели образуют ряд значений a_i , статистически независимых и принадлежащих одному закону распределения. Однако это не всегда так.

Пусть, например, целью статистического моделирования будет определение математического ожидания времени пребывания заявки в очереди $\bar{t}_{ож}$ одноканальной системы массового обслуживания.

В результате эксперимента с моделью будет получен ряд значений $\bar{t}_{ож i}$, $i = 1, N$, которые заведомо статистически зависимы: при большом времени ожидания k -й заявки значение $\bar{t}_{ож k+1}$, не может быть малым, если обе заявки находились одновременно в очереди. Связь точности оценки ε , среднего времени ожидания $\bar{t}_{ож}$ с количеством реализаций N в этом случае выглядит иначе, чем было рассмотрено ранее. Мы рассмотрим метод определения точности и количества реализаций для статистически зависимых последовательностей – откликов модели, в основе которого лежит так называемый регенеративный метод анализа СМО.

Допустим, что в результате эксперимента с имитационной моделью получен ряд значений $\bar{t}_{ож i}$, приведенный в табл. 4.7.

Здесь i – порядковый номер поступающих заявок.

Таблица 4.7

Результаты эксперимента – время ожидания заявки в очереди

| | | | | | | | | | | | | |
|------------------|---|---|---|---|---|---|---|---|----|----|----|-----|
| i | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | ... |
| $\bar{t}_{ож i}$ | 0 | 5 | 7 | 0 | 3 | 0 | 3 | 9 | 11 | 2 | 0 | ... |

Обратим внимание на то, что заявка 1 застаёт канал обслуживания свободным: её время ожидания в очереди равно нулю. Такая же ситуация возникла для заявок 4, 6 и 11. Период занятости и простоя канала обслуживания образуют цикл его работы. В табл. 4.7 можно выделить три таких цикла, в которые входят следующие наборы обслуженных заявок:

- цикл 1 – заявки 1, 2, 3;
- цикл 2 – заявки 4, 5;
- цикл 3 – заявки 6, 7, 8, 9, 10.

Заявка 11 является началом нового цикла 4 и т. д.

Начала каждого цикла неотличимы друг от друга – заявка поступает на обслуживание без ожидания. Говорят так: система восстанавливается (регенерируется) к началу каждого цикла, следовательно, поведение системы в очередном цикле не зависит от её поведения в предыдущих циклах. Введём обозначения:

Θ_k – сумма времен ожидания k -го цикла, $k = 1, n$;

q_k – количество заявок, образующих k -й цикл.

Для данных, приведенных в табл. 4.5:

$$\Theta_1 = 0 + 5 + 7 = 12, \quad \Theta_2 = 0 + 3 = 3, \quad \Theta_3 = 0 + 3 + 9 + 11 + 2 = 25, \dots;$$

$$q_1 = 3, \quad q_2 = 2, \quad q_3 = 5, \dots$$

Таким образом, мы получили пары чисел – независимых и одинаково распределенных:

$$(\Theta_1, q_1), (\Theta_2, q_2), (\Theta_3, q_3), \dots$$

Заметим, что числа Θ_k и q_k между собой зависимы.

Дальнейшей целью является определение оценки математического ожидания времени пребывания заявки в очереди $\bar{t}_{ож}$, отличающееся от $M[t_{ож}]$ на величину не более ε при заданной достоверности α .

Так как

$$\sum_{i=1}^N t_{ож i} = \sum_{k=1}^n \Theta_k, \quad N = \sum_{k=1}^n q_k,$$

где n – число циклов, то оценка математического ожидания времени пребывания заявки в очереди определяется так:

$$\bar{t}_{ож} = \frac{\sum_{i=1}^N t_{ож i}}{N} = \frac{\sum_{k=1}^n \Theta_k}{\sum_{k=1}^n q_k}.$$

Разделим числитель и знаменатель на число циклов n и получим:

$$\bar{t}_{ож} = \frac{\frac{\sum_{k=1}^n \Theta_k}{n}}{\frac{\sum_{k=1}^n q_k}{n}} = \frac{\bar{\Theta}}{q}.$$

В соответствии с центральной предельной теоремой оценка длительности цикла $\bar{\Theta}$ при числе циклов $n \rightarrow \infty$ есть случайная величина, распределённая по нормальному закону с математическим ожиданием и дисперсией соответственно:

$$M[\bar{\Theta}] = M[\Theta], \quad \sigma_{\bar{\Theta}}^2 = \frac{\sigma^2}{n},$$

где σ^2 – дисперсия, представляющая собой сумму дисперсий зависящих между собой случайных величин Θ и q .

Следовательно, имеет место уже знакомое нам выражение

$$P\left(\left|\bar{\Theta} - M[\Theta]\right| < t_{\alpha} \frac{\sigma}{\sqrt{n}}\right) = 2\Phi(t_{\alpha}).$$

Если ε – граничное значение ошибки для оценки $\bar{t}_{\text{ож}}$, то очевидно граничное значение ошибки для оценки $\bar{\Theta}$ равно $\varepsilon \bar{q}$.

Тогда $t_{\alpha} \frac{\sigma}{\sqrt{n}} = \varepsilon \bar{q}$. Из этого следует:

$$\varepsilon = t_{\alpha} \frac{\sigma}{q\sqrt{n}}, \quad n = t_{\alpha}^2 = \frac{\sigma^2}{\varepsilon^2 \bar{q}}.$$

Коэффициент t_{α} , как и ранее, характеризует достоверность оценки $\bar{t}_{\text{ож}}$ и является аргументом функции Лапласа: $t_{\alpha} = \Phi^{-1}\left(\frac{\alpha}{2}\right)$.

Значения σ^2 и \bar{q} до эксперимента неизвестны. Их ориентировочные значения должны быть определены по данным предварительных прогонов модели в количестве n^* реализаций циклов. Обычно $n^* = 50 \dots 100$.

Оценку дисперсии σ^2 обозначим S^2 . Она вычисляется так:

$$S^2 = S_{\Theta}^2 - 2\bar{t}_{\text{ож}} r_{\Theta, q} + t_{\text{ож}}^2 S_q^2.$$

Здесь:

$$S_{\Theta}^2 = \frac{1}{n^* - 1} \sum_{k=1}^{n^*} (\Theta_k - \bar{\Theta})^2 \quad \text{– оценка дисперсии } \Theta;$$

$$S_q^2 = \frac{1}{n^* - 1} \sum_{k=1}^{n^*} (q_k - \bar{q})^2 \quad \text{– оценка дисперсии } q;$$

$r_{\Theta, q} = \frac{1}{n^* - 1} \sum_{k=1}^{n^*} (\Theta_k - \bar{\Theta})(q_k - \bar{q})$ – корреляционный момент случайных величин Θ и q ;

$$\bar{\Theta} = \frac{\sum_{k=1}^{n^*} \Theta_k}{n^*}, \quad \bar{q} = \frac{\sum_{k=1}^{n^*} q_k}{n^*}, \quad t_{ож} = \frac{\bar{\Theta}}{\bar{q}}.$$

И, наконец, необходимое число циклов будет определено:

$$n = t_{\alpha}^2 \frac{S^2}{\varepsilon^2 q}.$$

Если окажется $n > n^*$, то моделирование продолжается до достижения n циклов. Если же окажется $n \leq n^*$, то моделирование заканчивается и, если необходимо, даётся оценка достигнутой точности.

Признак конца моделирования: $n = n_{зад}$ или количество обслуженных СМО заявок $n = N_{зад} = n_{зад} \cdot \bar{q}$.

Для моделей, предназначенных для приблизительных расчетов, удовлетворительной считается точность 10-15 %, а для моделей, предназначенных для использования в управляющих и контролирующих системах, – 1-2 % [44].

4.9. Проблема начальных условий

К тактическому планированию эксперимента относится и решение так называемой проблемы начальных условий.

В отличие от реальной системы модель работает прогонами – для накопления нужной статистики. Поэтому при каждом новом прогоне модели требуется какое-то время, чтобы установился стационарный режим, характеристики которого интересуют исследователя.

То есть начальные условия искажают характеристики стационарного режима.

Например, моделируется функционирование направления связи. В установившемся режиме входной буфер направления имеет среднее заполнение поступившими, но не обработанными пока сообщениями. Но перед каждым очередным прогоном в модели устанавливаются нулевые начальные условия.

Или ещё: вероятность обслуживания заявки в СМО имеет какое-то стационарное значение. Но вначале эта вероятность равна нулю.

Следовательно, начальные установки регистрируемого параметра (показателя эффективности и др.) искажают результат.

Для устранения ошибок, вызываемых не соответствующей установкой начальных условий, возможно применение следующих мер:

1. Ставить начальные условия, близкие значениям стационарного режима, то есть модель разрабатывается так, что условия функционирования системы типичны с самого начала.

2. Увеличить интервал исследования так, чтобы он стал значительно больше предполагаемого времени установления стационарного режима.

3. Отбросить информацию, снимаемую в промежутке времени от пуска до установившихся стационарных значений, и продолжить моделирование, собирая статистику, на которую уже не влияют нетипичные ситуации.

Первый подход требует от разработчика знания типичных условий работы и умения внести в модель эти условия. В моделях сложных систем это вряд ли выполнимо.

При втором подходе требуется слишком долгое моделирование до наступления состояния, когда исчезает влияние собранных неверных данных. Стоимость такого моделирования для сложных систем может оказаться слишком высокой, что делает этот подход нежелательным.

Третий подход оказывается наиболее удобным. Нужно на определенной стадии моделирования отбросить статистику с последующим продолжением моделирования без каких-либо модификаций модели. Такой подход используется в ряде систем моделирования. Заметим, однако, что время установки стационарных значений в модели трудно определить до эксперимента.

Все эти приёмы могут способствовать уменьшению влияния переходных процессов в модели на результаты эксперимента, однако свети его к нулю они не могут.

Вопросы для самоконтроля

- 4.1. Что понимается под компьютерным экспериментом?
- 4.2. Почему имитационную стохастическую модель нужно прогонять не один, а несколько раз?
- 4.3. Каковы цели планирования экспериментов?

- 4.4. Что такое стратегическое планирование эксперимента?
- 4.5. Что такое тактическое планирование эксперимента?
- 4.6. Что понимается под кибернетическим представлением эксперимента?
- 4.7. Что такое реакция или отклик системы?
- 4.8. Какой величиной является реакция или отклик системы?
- 4.9. Что такое факторы и уровни факторов?
- 4.10. Приведите вариант классификации факторов.
- 4.11. Симметричный факторный эксперимент.
- 4.12. Полный факторный эксперимент (ПФЭ).
- 4.13. Ортогональный факторный эксперимент.
- 4.14. Как определяется количество информационных точек в ПФЭ? В симметричном ПФЭ?
- 4.15. Пути сокращения затрат на проведение эксперимента.
- 4.16. Дайте определение точности оценки характеристики случайной величины.
- 4.17. Дайте определение достоверности оценки характеристики случайной величины.
- 4.18. Как получено выражение, связывающее точность и достоверность оценки с числом реализаций модели?
- 4.19. Способы априорного определения оценки дисперсии.
- 4.20. Как получено выражение $N = t_{\alpha}^2 \frac{P(1-P)}{\varepsilon^2}$?
- 4.21. Аргументы выражения $N = t_{\alpha}^2 \frac{\sigma^2}{\varepsilon^2}$.
- 4.22. Способы априорного определения оценки вероятности какого-либо события, определяемого моделированием.
- 4.23. Что понимается под «худшим случаем» определения числа прогонов модели?
- 4.24. В результате прогонов имитационной модели ожидается получить три показателя со следующими характеристиками:
- $$\sigma = 0,6; \varepsilon = 0,1; \alpha = 0,95;$$
- $$\sigma = 1,2; \varepsilon = 0,2; \alpha = 0,95;$$
- $$\sigma = 1,8; \varepsilon = 0,01; \alpha = 0,9.$$
- Определить требуемое количество N реализаций модели для достижения требуемой точности и достоверности.
- 4.25. Каковы пути разрешения проблемы начальных условий?

Глава 5. Обработка результатов имитационного эксперимента

Современные системы имитационного моделирования предоставляют возможность выполнять автоматически стандартную обработку результатов моделирования:

- определение характеристик случайных параметров, главным образом, их матожиданий и дисперсий;

- фиксация минимальных и максимальных значений исследуемых величин;

- частотное распределение результатов измерений (построение гистограмм);

- расчёт коэффициентов использования объектов модели и др.

Часто приходится выполнять более сложную обработку:

- определение функциональных или статистических зависимостей между исследуемыми величинами;

- выявление существенных или несущественных факторов, участвующих в эксперименте;

- сравнение случайных параметров процессов с целью определения значимости расхождения или совпадения их характеристик и др.

В наиболее развитых системах моделирования предусмотрены средства, обеспечивающие выполнение этих обработок. Но в любом случае исследователь должен понимать сущность обработки, уметь правильно готовить исходные данные, грамотно интерпретировать результаты обработки. При наличии альтернатив обоснованно выбрать метод обработки и, при необходимости, разрабатывать или совершенствовать соответствующие процедуры. В настоящей главе рассматриваются наиболее актуальные для практики понятия и математические методы обработки данных, полученных в соответствии с целью исследования с помощью имитационной модели.

5.1. Характеристики случайных величин и процессов

В результате эксперимента с имитационной статистической моделью, состоящего из N наблюдений, мы получаем N значений исследуемой случайной величины a :

$$a_1, a_2, \dots, a_i, \dots, a_N .$$

По этим данным нужно дать всестороннее описание случайной величины a .

Описать случайную величину – это значит определить её характеристики. В общем случае:

$$\bar{\Theta} = \bar{\Theta}(a_1, a_2, \dots, a_i, \dots, a_N),$$

где $\bar{\Theta}$ – оценка характеристики случайной величины.

Под характеристикой понимают следующее.

Во-первых, это характеристика *величины*:

Математическое ожидание (среднее арифметическое);

медиана (срединное значение);

мода (наиболее вероятное значение);

среднее геометрическое и др.

В рамках задач, характерных для нашей профессии, наиболее актуальным является математическое ожидание. Как известно, математическое ожидание определяет центр рассеивания случайной величины, наиболее полно отмечающее ее положение на числовой оси. Будем обозначать математическое ожидание случайной величины a так: $M[a]$.

Во-вторых, это характеристики *рассеивания*:

дисперсия (математическое ожидание квадрата отклонения случайной величины a);

среднее квадратическое отклонение (квадратный корень из дисперсии, стандартное отклонение), иногда целесообразно пользоваться этой характеристикой, так как она имеет размерность самой случайной величины, а размерность дисперсии – квадрат размерности случайной величины;

размах ($\max a_i - \min a_i$).

В-третьих, это характеристика *связи* между случайными величинами (корреляция); степень связи определяется величиной коэффициента корреляции r . В случайном процессе связь между значениями случайной функции в моменты времени t_k, t_s определяет коэффициент автокорреляции $k(t_k, t_s)$.

В-четвертых, это характеристика *закона распределения вероятностей* случайной величины в виде плотности или функции распределе-

ния: $f(a)$ или $F(a) = \int_{-\infty}^a f(a) da$.

5.2. Требования к оценкам характеристик

Ограниченное число реализаций модели не позволяет точно определить значения этих характеристик, а только приближенно, то есть так называемые *оценки характеристик* $\bar{\Theta}$. Степень приближения оценок $\bar{\Theta}$ зависит от методов их вычислений (формул). Поскольку $\bar{\Theta} = \bar{\Theta}(a_1, a_2, \dots, a_i, \dots, a_N)$, где a_i – случайные значения искомого параметра, то величина $\bar{\Theta}$ – случайная со своими значениями математического ожидания, дисперсии и т. п.

Как правило, математическая статистика может предложить разные формулы для вычисления оценки одной и той же характеристики. Следовательно, оценки могут быть более или менее точными или даже вовсе непригодными при имитационном моделировании.

Чтобы оценка наилучшим образом представляла искомую характеристику, нужно, чтобы она обладала следующими свойствами:

- несмещённостью;
- состоятельностью;
- эффективностью.

Несмещённость. Это свойство означает, что оценка не содержит систематической ошибки. То есть, математическое ожидание оценки совпадает с действительным значением характеристики $\bar{\Theta}$:

$$M[\bar{\Theta}] = M[\Theta].$$

Состоятельность. Это свойство означает, что оценка $\bar{\Theta}$ приближается сколь угодно близко к истинному значению характеристики $\Theta(M[\Theta])$ по мере увеличения объёма выборки, т. е. увеличения числа реализаций модели. Формально это свойство записывают так:

$$P(|\bar{\Theta} - M[\Theta]| < \varepsilon) \rightarrow 1$$

при $N \rightarrow \infty$ и любом $\varepsilon > 0$.

Именно это свойство являлось определяющим при нахождении количественной связи между точностью, достоверностью оценок и числом реализаций модели.

Эффективность. Это свойство означает, что из всех несмещенных и состоятельных оценок следует предпочесть ту, у которой разброс значений меньше. Иначе: эффективной оценкой характеристики слу-

чайной величины называют ту оценку, которая имеет наименьшую дисперсию:

$$D[\bar{\Theta}] = \min_k \bar{\Theta}_k,$$

k – число возможных оценок.

В исследовании свойств оценок большая заслуга принадлежит англичанину Рональду Эйлмеру Фишеру (1890-1962). Основные результаты он получил в 1912 г., когда ему было 22 года.

5.3. Оценка характеристик случайных величин и процессов

Наиболее используемые оценки характеристик приведены в табл. 5.1.

Таблица 5.1

Характеристики случайных величин и их оценки

| Характеристика | Оценка | Среднее квадратическое отклонение оценки |
|--|--|--|
| Математическое ожидание $M[x] = \int_{-\infty}^{+\infty} x f(x) dx$ | $\bar{x} = \frac{\sum_{i=1}^N x_i}{N}$ | $\sigma_{\bar{x}} = \frac{S}{\sqrt{N}}$ |
| Дисперсия $D[x] = M[x^2] - (M[x])^2$ | $S^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2$ | $\sigma_{S^2} = \sqrt{\frac{2}{N}} S^2$ |
| Среднее квадратическое отклонение $\sigma_x = \sqrt{D[x]}$ | $S = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$ | $\sigma_S = \frac{S}{\sqrt{2N}}$ |
| Вероятность события P | $\bar{P} = \frac{m}{N}$ | $\sigma_{\bar{P}} = \sqrt{\frac{P(1-P)}{N}}$ |
| Коэффициент корреляции $\rho_{x,y} = \frac{\text{Cov}(x,y)}{\sigma_x \sigma_y}$ | $r_{xy} = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{S_x S_y}$ | $\sigma_r = \frac{1 - r_{xy}^2}{\sqrt{N}}$ |

Все оценки несмещённые, состоятельные, эффективные.

Интересный факт. Проблемами оценок занимался и Абрахам Вальд (1902-1950), американский математик австрийского происхождения

дения. Во время Второй мировой войны командование американских и британских ВВС поручило Абрахаму Вальду выяснить, какие части фюзеляжа самолета нужно защитить дополнительной броней. Вальд изучал самолеты, возвращавшиеся с боевых вылетов, отмечая места попаданий. В результате он рекомендовал установить дополнительную защиту на те участки (центральную и заднюю части фюзеляжа), где количество пробоин было минимальным. Рекомендация была основана на выводе, что защищать нужно от тех попаданий, которых Вальд не видел, самолеты, которые их получили, не возвращались.

Приведем для иллюстрации оценок два примера.

Пример 5.1. Оценка математического ожидания случайной величины a – среднее арифметическое

$$\bar{a} = \frac{1}{N} \sum_{i=1}^N a_i$$

является несмещённой, состоятельной и эффективной.

Оценка в виде медианы не является эффективной, так как дисперсия в этом случае

$$D[\bar{a}_m] = \frac{\pi \sigma_a^2}{2N}$$

в $\frac{\pi}{2}$ раз больше дисперсии $D[\bar{a}]$, равной, как известно, $D[\bar{a}] = \frac{\sigma_a^2}{N}$.

Пример 5.2. Выборочная дисперсия случайной величины a

$$S^2 = \frac{1}{N} \sum_{i=1}^N (a_i - \bar{a})^2$$

состоятельна, эффективна, но смещена. Смещение образовалось из-за того, что вместо неизвестного $M[a]$ в формуле стоит оценка \bar{a} .

Несмещенная оценка имеет вид:

$$S^2 = \frac{1}{N-1} \sum_{i=1}^N (a_i - \bar{a})^2.$$

Иногда формулы для вычисления оценок математического ожидания и дисперсии используют в рекуррентной форме:

$$M[\bar{a}]_i = M[\bar{a}]_{i-1} \frac{i-1}{i} + \frac{a_i}{i},$$

$$S_i^2 = S_{i-1}^2 \frac{i-1}{i} + \frac{(a_i - M[\bar{a}]_i)^2}{i},$$

где $M \left[\bar{a} \right]_i$, $M \left[\bar{a} \right]_{i-1}$, S_i^2 , S_{i-1}^2 – оценки матожидания и дисперсии, вычисленные по данным i и $(i-1)$ реализаций имитационной модели.

Приведенные в табл. 5.1 формулы соответствуют нормальному закону распределения вероятностей исследуемой величины.

При исследовании случайного процесса $X(t)$ весь временной интервал $(0, T)$ представляется последовательностью из M временных точек t_j , $j = \overline{1, M}$, в каждой из которых измеряется значение сечения $x_i(t_j)$. Индекс i – номер реализации случайного процесса, $i = \overline{1, N}$.

Полученные данные образуют матрицу сечений размером $(M \times N)$, что и является моделью исследуемого процесса (табл. 5.2).

Таблица 5.2

Результаты исследования случайного процесса

| Реализации | Временные точки | | | | | |
|------------|-----------------|------------|-----|------------|-----|------------|
| | t_1 | t_2 | ... | t_j | ... | t_M |
| 1 | $x_1(t_1)$ | $x_1(t_2)$ | ... | $x_1(t_j)$ | ... | $x_1(t_M)$ |
| 2 | $x_2(t_1)$ | $x_2(t_2)$ | ... | $x_2(t_j)$ | ... | $x_2(t_M)$ |
| ... | ... | ... | ... | ... | ... | ... |
| i | $x_i(t_1)$ | $x_i(t_2)$ | ... | $x_i(t_j)$ | ... | $x_i(t_M)$ |
| ... | ... | ... | ... | ... | ... | ... |
| N | $x_N(t_1)$ | $x_N(t_2)$ | ... | $x_N(t_j)$ | ... | $x_N(t_M)$ |

Совокупность сечений в каждой временной точке t_j (столбец матрицы), представляет собой случайные числа некоторой случайной величины в общем случае со своими законами распределения, математическими ожиданиями, дисперсиями:

$$\bar{x}(t_j) = \frac{1}{N} \sum_{i=1}^N x_i(t_j), \quad S_{x(t_j)}^2 = \frac{1}{N} \sum_{i=1}^N \left[x_i(t_j) - \bar{x}(t_j) \right]^2.$$

При решении практических задач последовательности этих оценок матожиданий и дисперсий, определенных в точках $t_1 \dots t_M$, достаточно полно представляют моделируемый случайный процесс. Оценки матема-

тических ожиданий $\bar{x}(t_j)$ и дисперсий $S_{x(t_j)}^2$ можно аппроксимировать подходящими кривыми в предположении непрерывности процесса.

Иногда исследователя интересует связь сечений случайного процесса между собой. Степень зависимости между сечениями определяет автокорреляционная функция. Оценка её имеет вид:

$$\bar{K}_X(t_k, t_s) = \frac{1}{N} \sum_{i=1}^N [x_i(t_k) - \bar{x}(t_k)] \cdot [x_i(t_s) - \bar{x}(t_s)],$$

где $x_i(t_k)$ и $x_i(t_s)$ – значения сечений в точках t_k и t_s соответственно i -й реализации;

$\bar{x}(t_k)$ и $\bar{x}(t_s)$ – оценки математических ожиданий совокупности сечений в точках t_k и t_s соответственно.

Данные расчёта значений автокорреляционной функции $\bar{K}_X(t_k, t_s)$, $k = \overline{1, M}$, $s = \overline{1, M}$ помещают в таблицу, которая и является табличным определением её. В случае необходимости данные таблицы могут быть представлены подходящей аппроксимирующей кривой.

Пример таблицы значений $\bar{K}_X(t_k, t_s)$ для случайного процесса, определенного пятью сечениями ($M = 5$), показан в табл. 5.3.

Таблица 5.3

Значения автокорреляционной функции

| Временные точки | | | | | |
|-----------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| | t_1 | t_2 | t_3 | t_4 | t_5 |
| t_1 | $\bar{K}_X(t_1, t_1)$ | $\bar{K}_X(t_1, t_2)$ | $\bar{K}_X(t_1, t_3)$ | $\bar{K}_X(t_1, t_4)$ | $\bar{K}_X(t_1, t_5)$ |
| t_2 | $\bar{K}_X(t_2, t_1)$ | $\bar{K}_X(t_2, t_2)$ | $\bar{K}_X(t_2, t_3)$ | $\bar{K}_X(t_2, t_4)$ | $\bar{K}_X(t_2, t_5)$ |
| t_3 | $\bar{K}_X(t_3, t_1)$ | $\bar{K}_X(t_3, t_2)$ | $\bar{K}_X(t_3, t_3)$ | $\bar{K}_X(t_3, t_4)$ | $\bar{K}_X(t_3, t_5)$ |
| t_4 | $\bar{K}_X(t_4, t_1)$ | $\bar{K}_X(t_4, t_2)$ | $\bar{K}_X(t_4, t_3)$ | $\bar{K}_X(t_4, t_4)$ | $\bar{K}_X(t_4, t_5)$ |
| t_5 | $\bar{K}_X(t_5, t_1)$ | $\bar{K}_X(t_5, t_2)$ | $\bar{K}_X(t_5, t_3)$ | $\bar{K}_X(t_5, t_4)$ | $\bar{K}_X(t_5, t_5)$ |

Очевидно, что рассчитывать все значения $\bar{K}_X(t_k, t_s)$ для заполнения таблицы (в данном примере их 25) не надо, так как значения K_X при $t_k = t_s$ («северо-западная диагональ») представляют собой значе-

ния соответствующих дисперсий. И $\overline{K}_X(t_k, t_s) = \overline{K}_X(t_s, t_k)$, что исключает необходимость расчёта половины оставшихся значений коэффициентов автокорреляционной функции, расположенных выше или ниже упомянутой диагонали.

5.4. Гистограмма

Одной из задач моделирования может быть определение закона распределения вероятностей исследуемой случайной величины и количественных значений его характеристик.

Аналогом, моделью плотности распределения вероятности случайной величины является гистограмма, которую можно построить (аналитически или графически) по данным имитационного статистического моделирования.

Гистограмма (рис. 5.1) строится так.

В результате N реализаций модели получен ряд случайных значений исследуемого параметра a : a_1, a_2, \dots, a_N . Весь диапазон значений a разбивается на l интервалов (разрядов). Числовой диапазон каждого интервала обозначим $\Delta_j, j = \overline{1, l}$. Обычно все числовые диапазоны одинаковые: $\Delta_j = \Delta$.

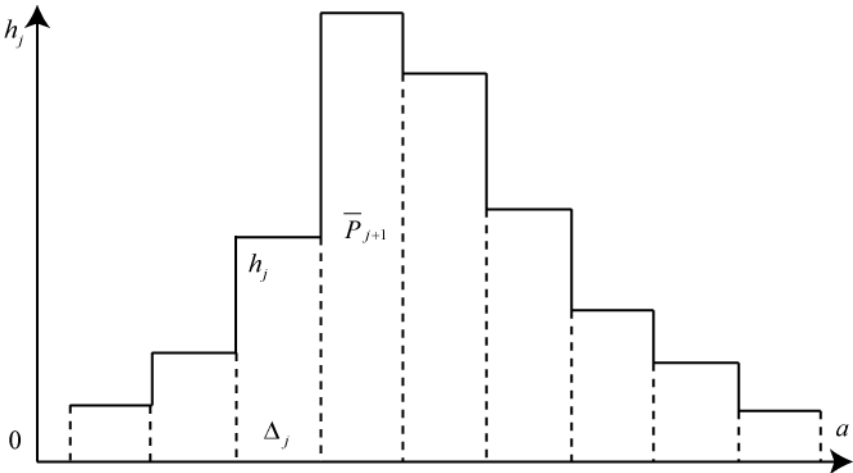


Рис. 5.1. Гистограмма

Для каждого интервала подсчитываем число значений a_i , попавших в него – m_j .

На каждом интервале строят прямоугольник с высотой h_j :

$$h_j = \frac{m_j}{N \cdot \Delta}.$$

Площадь каждого прямоугольника гистограммы равна относительной частоте \bar{P}_j :

$$\bar{P}_j = \frac{m_j}{N \cdot \Delta} \cdot \Delta = \frac{m_j}{N}.$$

По выбору числа интервалов l существуют разные эмпирические рекомендации, например:

$$20 \leq l \leq 30, \quad l = 5 \lg N.$$

Чем больше N и l , а меньше Δ , тем ближе гистограмма совпадает с некоторым теоретическим распределением. Доказал это Валерий Иванович Гливенко – известный отечественный математик.

На основе очертания гистограммы делается предположение (выдвигается гипотеза) о совпадении полученного эмпирического распределения вероятностей с тем или иным теоретическим – нормальным, экспоненциальным, Вейбулла и т. д. Затем выполняется проверка этой гипотезы с помощью *критериев согласия*. В курсе высшей математики рассматриваются некоторые критерии (Колмогорова, Смирнова и др.), наиболее популярным считают критерий хи-квадрат – критерий Пирсона, предложенный им в 1903 г.

Оценки математического ожидания и дисперсии можно получить по данным гистограммы:

$$\bar{a} = \sum_{j=1}^l \bar{a}_j \bar{P}_j, \quad S^2 = \sum_{j=1}^l \bar{a}_j^2 \bar{P}_j - \frac{\Delta}{12},$$

где \bar{a}_j – среднее значение каждого интервала;

\bar{P}_j – оценка по каждому интервалу;

$\frac{\Delta}{12}$ – поправка В. Шеппарда. Обычно поправку применяют при

высокой точности расчётов или при большом числе реализаций модели ($N \geq 500$). При меньшем числе реализаций модели ($N < 500$) поправка не применяется.

5.5. Элементы дисперсионного анализа. Критерий Фишера

Приведем понятия, которые используем в дальнейшем. В математической статистике (а это основной математический аппарат обработки результатов моделирования) используется понятие гипотезы.

Гипотезой называется предположение о:

законах распределения вероятностей случайных величин;

значениях характеристик случайных величин;

совпадении законов распределения двух и более случайных величин и др.

Обычно исходную гипотезу называют нулевой и обозначают H_0 . Противоположное утверждение называют конкурирующей гипотезой и обозначают H_1 .

Гипотеза подвергается проверке. Смысл этой проверки в том, чтобы принять или отклонить её с допустимым минимальным риском. При этом возможны ошибки:

забраковать проверяемую гипотезу, *если она верна*, что соответствует так называемой *ошибке первого рода*;

принять проверяемую гипотезу, когда она *не верна*, значит совершить *ошибку второго рода*.

Правило, по которому принимается суждение об истинности или ложности основной гипотезы H_0 называют *критерием проверки* или *критерием согласия*.

В практике моделирования и обработки экспериментальных данных очень часто необходимо решать проблему подтверждения или опровержения гипотезы о принадлежности двух или более выборок одной генеральной совокупности.

К такой проблеме приводят такие задачи:

сравнительная оценка различных технологических процессов по их производительности, точности, экономичности;

сравнение конструктивных особенностей приборов, машин, средств вооружения и др.

Признаки, по которым проводится сравнительная оценка, часто не являются детерминированными, обладают рассеиванием. Например, точность никогда не может быть абсолютной, так как измерительные приборы всегда несут в себе ошибку.

Наиболее общим и часто применяемым на практике методом сравнения качеств объектов является дисперсионный анализ.

Сущность *дисперсионного анализа* состоит в проверке гипотезы о тождественности выборочных дисперсий одной и той же генеральной дисперсии.

Почему исследователей интересует сравнение именно дисперсий, а не каких-либо других характеристик? Заметим, что есть методики сравнения, например, математических ожиданий и др., но они не обладают такой общностью, как дисперсионный анализ.

А дело в том, что дисперсия характеризует важные конструкторские и технологические показатели такие, как:

точность приборов;

рассеивание точек попадания при стрельбе и др.

И ещё дисперсионный анализ одновременно решает проблему проверки гипотезы о равенстве средних значений выборок.

Задача сравнения дисперсий сводится к проверке исходной гипотезы (нулевой гипотезы H_0) о принадлежности двух выборок одной и той же генеральной совокупности.

Для проверки гипотезы о равенстве дисперсий нужно иметь независимую функцию, вычисляемую по данным эксперимента.

Такой функцией является функция Фишера (распределение Фишера, F -распределение), определяемая так:

$$F = \frac{\frac{U}{k_1}}{\frac{V}{k_2}},$$

где U и V – случайные величины, имеющие распределение χ^2 ;

k_1 и k_2 – соответствующие степени свободы случайных величин U и V соответственно, $k_1 = N_1 - 1$, $k_2 = N_2 - 1$;

N_1 и N_2 – количество испытаний (объемы выборок).

Почему χ^2 является мерой сравнения дисперсий? Потому, что дисперсии, являясь суммой квадратов ошибок, имеют распределение χ^2 .

Распределение *хи-квадрат* определяется следующим образом:

$$f(x) = \left\{ 1 / \left[2^{\nu/2} * \Gamma(\nu/2) \right] \right\} * \left[x^{(\nu/2)-1} * e^{-x/2} \right], \nu = 1, 2, \dots, 0 < x,$$

где ν – число степеней свободы, e – число Эйлера (2,71...), Γ – гамма-функция.

График плотности F -распределения показан на рис. 5.2.

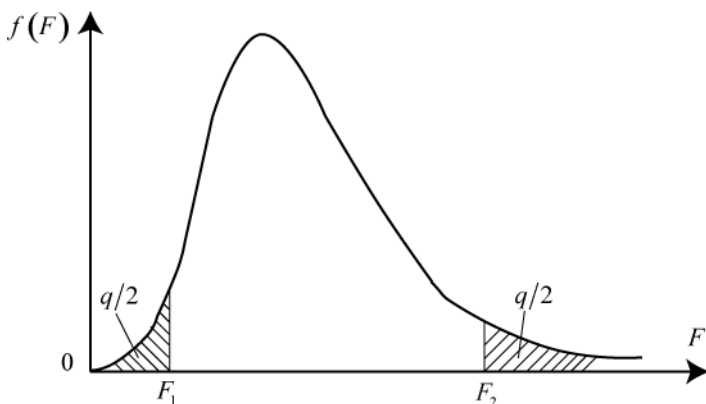


Рис. 5.2. График плотности F -распределения

Итак, случайная величина

$$F = \frac{S_1^2}{S_2^2},$$

где S_1^2 и S_2^2 – несмещённые оценки дисперсий, полученных из независимых выборок, взятых из нормальных совокупностей, имеет распределение Фишера (F -распределение).

Величина F – случайна, судить однозначно по её величине о подтверждении или опровержении гипотезы об однородности исследуемых выборок нельзя.

Поэтому вводится $q\%$ уровень значимости, численно равный вероятности *неприемлемых* отклонений от принятой гипотезы. Области неприемлемых значений F показаны на рис. 5.2 штриховкой. Граничные точки допустимых значений F определяются точками F_1 и F_2 , соответствующих вероятностям $q/2$.

Если вычисленное по данным эксперимента значение F попадает в область между точками F_1 и F_2 :

$$F_1 \leq F \leq F_2,$$

то принятая гипотеза не опровергается.

Заметим, что случайная величина

$$F^* = \frac{1}{F} = \frac{S_2^2}{S_1^2}$$

также имеет F -распределение со степенями свободы $k_1 = N_1 - 1$ и $k_2 = N_2 - 1$ соответственно. Следовательно, вероятность попадания числа F в левую критическую область равна:

$$P(F > F_1) = P\left(\frac{1}{F} > \frac{1}{F_1}\right) = P\left(F^* > \frac{1}{F_1}\right).$$

Отсюда следует, что *левая* критическая точка F -распределения соответствует *правой* критической точке F^* -распределения. Т. е. правые точки распределений F и F^* определяют левую и правую точки F_1 и F_2 . Поэтому в таблицах представлены только правые F_2 критические точки F -распределения.

В таблицах значения F_2 приведены в зависимости от $q/2$, числа степеней свободы $k_1 = N_1 - 1$ и $k_2 = N_2 - 1$.

Обычно при вычислении F в числитель отношения $\frac{S_1^2}{S_2^2}$ ставят значение большей дисперсии.

Итак, при $F < F_2$ принятая гипотеза не опровергается, при $F > F_2$ — не подтверждается.

Пример 5.3. В часть поступили две буссоли. Первая из них при измерении пять раз одного и того же угла показала дисперсию $S_1^2 = 0,1 \text{ град}^2$. По результатам семи измерений второй буссолью того же угла получена дисперсия $S_2^2 = 0,15 \text{ град}^2$.

Однотипны ли буссоли? Одинаковы ли они по точности измерения углов? Выдвинем и проверим гипотезу об их однотипности для уровня значимости $q = 10\%$.

Решение.

$$F = \frac{S_2^2}{S_1^2} = \frac{0,15}{0,1} = 1,5; \quad k_1 = 5 - 1, \quad k_2 = 7 - 1 = 6.$$

По таблице F -распределения для степеней свобод $k_2 = 6$ соответствующей большей дисперсии, и $k_1 = 4$, соответствующей меньшей дисперсии, и уровню значимости $q/2$, находим $F_2 = 6,16$.

Так как $F = 1,5 < F_2 = 6,16$, то для уровня значимости $q = 10\%$ гипотеза об одинаковости буссолей не опровергается.

Итак: чем меньше уровень значимости $q\%$, тем меньше вероятность забраковать проверяемую гипотезу, когда она верна, то есть совершить ошибку первого рода.

Но с уменьшением уровня значимости (увеличением F_2) расширяется область допустимых ошибок, что приводит к увеличению вероятности принятия неверного решения, то есть совершения ошибки второго рода.

В заключение изложенного отметим, что как бы ни был велик объём статистического материала (N_1, N_2) критерий Фишера (впрочем, как и любой другой) не может дать абсолютно достоверный ответ о справедливости или несправедливости проверяемой гипотезы, так как мы оперируем случайными числами.

То есть, опровержение гипотезы при $F > F_2$ не является категорическим доказательством её несправедливости. Равно как и подтверждение гипотезы при $F \leq F_2$ не означает категорического доказательства её справедливости. Не исключено, что в том и в другом случаях решение может оказаться ошибочным.

Суждение о подтверждении или отклонении выдвинутой гипотезы высказывается с определенной степенью достоверности.

Среди инженеров бытует шутливое изречение: статистика, как фонарный столб на улице: света даёт мало, но при случае на него можно опереться.

Но свет-то даёт! И другой альтернативы нет.

5.6. Критерий Вилкоксона

Как и в предыдущем случае решается следующая задача. Имеются две серии независимых наблюдений однородных случайных величин X и Y , причем значения x_i и y_i дают различные значения средних ($\bar{x} \neq \bar{y}$) или (и) различные рассеивания. Возникает вопрос: можно ли считать эти расхождения существенными или расхождения зависят от случайных выборок?

Простой в употреблении и вполне приемлемый по точности критерий для проверки гипотезы о тождественности функций распределения $F(x)$ и $F(y)$ предложил в середине прошлого века Вилкоксон (Уилкоксон). Критерий назван его именем.

Рассматривается нулевая гипотеза: $F(x) = F(y)$. Конкурирующая гипотеза: $F(x) < F(y)$.

Критерий основан на подсчёте числа инверсий. Инверсии определяются так.

Измеренные значения x_i и y_i , $i = \overline{1, N}$ располагаются в общую последовательность в порядке возрастания их значений. Пусть это будет, например, так:

$$y_1, x_1, x_2, y_2, y_3, y_4, x_3, y_5, y_6, x_4,$$

где $x_1 \dots x_4$ – члены, принадлежащие первой выборке;

$y_1 \dots y_6$ – члены второй выборки.

Данная последовательность – не убывающая, содержащая $m+n$ чисел, m – количество чисел последовательности x , n – последовательности y .

Если гипотеза $F(x) = F(y)$ верна, то достаточно очевидно, что числа из обеих последовательностей хорошо перемешиваются. Степень перемешивания определяется числом инверсий членов первой последовательности относительно второй. Если в упорядоченной общей последовательности некоторому x предшествует одно значение y , это означает, что имеет место одна инверсия. Если некоторому x предшествуют k значений y , то это значение x имеет k инверсий.

Для нашего примера член x_1 имеет одну инверсию с y_1 ; член x_2 – тоже одну с y_1 ; член x_3 имеет четыре инверсии (с y_1, y_2, y_3, y_4); член x_4 имеет шесть инверсий (с $y_1, y_2, y_3, y_4, y_5, y_6$).

Таким образом, общее число инверсий:

$$\bar{u} = 1 + 1 + 4 + 6 = 12.$$

Показано, что случайная величина u уже при $m+n > 20$ и $\min(m, n) \geq 3$ дает хорошее приближение к нормальному распределению с математическим ожиданием и дисперсией:

$$M[u] = \frac{m \cdot n}{2}, \quad D[u] = \frac{m \cdot n}{12} (m + n + 1), \quad \sigma_u = \sqrt{D[u]}.$$

При уровне значимости $q\%$ и нормальности распределения u вероятность попадания значения u в критическую область (что означает не подтверждение нулевой гипотезы) равна:

$$P\left\{\left|M[u]-\bar{u}\right| > t_{\alpha} \sigma_u\right\} = q = 1 - 2\Phi(t_{\alpha}),$$

$$P\left\{\left|M[u]-\bar{u}\right| - t_{\alpha} \sigma_u > \bar{u} > M[u] + t_{\alpha} \sigma_u\right\} = q.$$

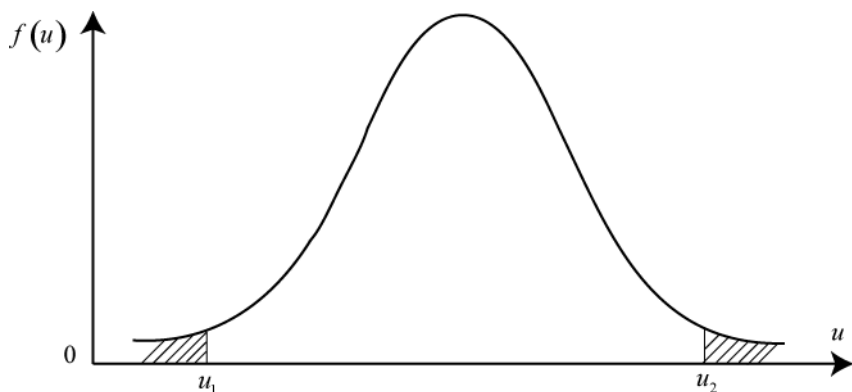


Рис. 5.3. Левая и правая критические границы функции $f(u)$

Отсюда следует, что левая критическая граница и правая критическая граница (рис. 5.3) равны соответственно:

$$u_1 = M[u] - t_{\alpha} \sigma_u, \quad u_2 = M[u] + t_{\alpha} \sigma_u.$$

Значение t_{α} определяется так:

$$1 - 2\Phi(t_{\alpha}) = q; \quad 2\Phi(t_{\alpha}) = 1 - q; \quad \Phi(t_{\alpha}) = \frac{1 - q}{2}; \quad t_{\alpha} = \Phi^{-1}\left(\frac{1 - q}{2}\right).$$

$\Phi(t_{\alpha})$ – функция Лапласа, с которой мы встречались ранее, она табулирована. Наиболее актуальные соответствия уровней значимости q и аргументов функции Лапласа t_{α} указаны в табл. 5.4.

Таблица 5.4

Уровни значимости и аргументы функции Лапласа

| | | | |
|--------------|------|------|------|
| $q\%$ | 2 | 5 | 10 |
| t_{α} | 2,33 | 1,96 | 1,65 |

Пример 5.4. С целью проверки адекватности модели центра коммутации сообщений измерено время задержки передачи сообщений на модели центра и непосредственно на самом центре. Результаты измерений сведены в табл. 5.5.

Таблица 5.5

Результаты измерений задержки сообщений

| | | | | | | | | | | |
|--------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| x, c | 0,8 | 1,9 | 3,0 | 3,5 | 3,8 | 2,5 | 1,7 | 0,9 | 1,0 | 2,3 |
| y, c | 1,4 | 2,1 | 3,1 | 3,6 | 2,7 | 1,8 | 1,1 | 0,2 | 1,6 | 2,8 |

Последовательность x – отклики модели, y – данные, измеренные на центре. Проверка адекватности модели состоит в проверке нулевой гипотезы, то есть в том, что данные измерений идентичны в статистическом смысле.

Решение.

Составим в порядке возрастания общую последовательность времён задержек x и y (табл. 5.6).

Таблица 5.6

Общая последовательность времен задержек сообщений

| | | | | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|----------|----------|
| y_1 | x_1 | x_2 | x_3 | y_2 | y_3 | y_4 | x_4 | y_5 | x_5 |
| 0,2 | 0,8 | 0,9 | 1,0 | 1,1 | 1,4 | 1,6 | 1,7 | 1,8 | 1,9 |
| y_6 | x_6 | x_7 | y_7 | y_8 | x_8 | y_9 | x_9 | y_{10} | x_{10} |
| 2,1 | 2,3 | 2,5 | 2,7 | 2,8 | 3,0 | 3,1 | 3,5 | 3,6 | 3,8 |

Расчёт числа инверсий для x :

$$\bar{u} = 1+1+1+4+5+6+6+8+9+10 = 51.$$

Расчёт характеристик:

$$M[u] = \frac{10 \cdot 10}{2} = 50 \text{ с}, \quad D[u] = \frac{10 \cdot 10}{12} (10+10+1) = 175 \text{ с}^2,$$

$$\sigma_u = \sqrt{175} = 13,23 \text{ с}.$$

Примем уровень значимости $q = 5\%$. Тогда $t_\alpha = 1,96$.

Правая критическая точка:

$$u_2 = M[u] + t_\alpha \sigma_u = 50 + 1,96 \cdot 13,23 \approx 74.$$

Левая критическая точка:

$$u_1 = M[u] - t_\alpha \sigma_u = 50 - 1,96 \cdot 13,23 \approx 24.$$

Проверка гипотезы H_0 :

$$24 < 52 < 74.$$

Гипотеза об идентичности распределений времён ожиданий в модели и в объекте не опровергается.

В заключение отметим, что при малых m и n ($m + n < 20$) для критерия Вилкоксона составлены таблицы критических точек u_1 и u_2 для различных уровней значимости q . Эти таблицы приводятся в широко известных изданиях, например, Б. Л. ван дер Варден «Математическая статистика», Б. В. Гнеденко и др. «Математические методы в теории надежности».

5.7. Однофакторный дисперсионный анализ

В современной жизни – военной и гражданской часто возникают проблемы, решение которых требует научного обоснования.

Идентичны ли патроны для конкретного образца стрелкового вооружения, выпускаемые на разных заводах?

Идентичны ли автоматы, выпускаемые на разных заводах?

Идентичны ли снаряды для конкретного образца артиллерийского вооружения, выпускаемые на разных заводах?

Здесь в качестве исследуемого фактора выступают заводы. Разные, но однотипные по назначению – варианты фактора, которые можно трактовать как уровни факторов.

Аналогичная задача возникает при сравнении однотипных изделий, вырабатываемых с применением различных технологий. Здесь подлежит анализу фактор – технология производства.

Эти и подобные задачи являются задачами однофакторного дисперсионного анализа (ОДА).

Иногда возможны задачи одновременного исследования влияния двух и более факторов. Например, чем объяснить рассеивание попаданий в цель: конструктивными особенностями стрелкового вооружения, выпущенного на разных заводах, или различиями в подготовке стрелков?

Исследованием влияния факторов и занимается *факторный дисперсионный анализ*.

Мы рассмотрим ОДА, наиболее актуальный анализ на практике. Теория рассматривает и многофакторный дисперсионный анализ. В нём процедуры подобны тем, которые мы рассмотрим в ОДА. Усложняются только расчёты и при необходимости, зная ОДА, овладеть методикой многофакторного дисперсионного анализа не составит труда.

Эксперимент для выполнения ОДА состоит в накоплении результатов измерений контролируемого параметра (угла, расстояния, наработки на отказ некоторого изделия и т. д.) при каждом варианте исследуемого фактора.

Введём обозначения:

m – число вариантов фактора;

n – число измерений при каждом варианте;

a_{ij} – результат каждого измерения;

$i = \overline{1, m}$ – номер варианта фактора;

$j = \overline{1, n}$ – номер измерения.

Схема эксперимента заключается в следующем.

Производится n измерений контролируемого параметра при m вариантах фактора.

В принципе, число измерений может быть разным для каждого варианта фактора. Ход дальнейших рассуждений от этого не меняется.

Результаты эксперимента сводятся в таблицу (табл. 5.7).

Таблица 5.7

Результаты эксперимента

| № варианта | Номер измерения | | | | | | Средние значения |
|------------|-----------------|----------|-----|----------|-----|----------|---|
| | 1 | 2 | ... | j | ... | n | |
| 1 | a_{11} | a_{12} | ... | a_{1j} | ... | a_{1n} | $\bar{a}_1 = \frac{1}{n} \sum_{j=1}^n a_{1j}$ |
| 2 | a_{21} | a_{22} | ... | a_{2j} | ... | a_{2n} | $\bar{a}_2 = \frac{1}{n} \sum_{j=1}^n a_{2j}$ |
| ... | | | | | | | |
| i | a_{i1} | a_{i2} | ... | a_{ij} | ... | a_{in} | $\bar{a}_i = \frac{1}{n} \sum_{j=1}^n a_{ij}$ |
| ... | | | | | | | |
| m | a_{m1} | a_{m2} | ... | a_{mj} | ... | a_{mn} | $\bar{a}_m = \frac{1}{n} \sum_{j=1}^n a_{mj}$ |

Вопрос: влияют ли варианты фактора на точность измерений? Или, говоря языком математической статистики, являются результаты $m \cdot n$ измерений выборкой одной генеральной совокупности, или нет? Если да, то варианты фактора несущественны, если нет, то существенны.

Будем исходить из следующей нулевой гипотезы:

наблюдения каждого варианта независимы;

наблюдения каждого варианта имеют нормальное распределение;

имеют одинаковую дисперсию σ^2 ;

имеют одинаковые центры рассеивания.

Очевидно, если систематические ошибки вариантов не одинаковы, следует ожидать повышенного рассеивания выборочных средних \bar{a}_i .

Для подтверждения или отрицания выдвинутой нулевой гипотезы об идентичности вариантов фактора проведём дисперсионный анализ.

Общее среднее арифметическое по всем $m \cdot n$ измерениям:

$$\bar{a} = \frac{\sum_{i=1}^m \sum_{j=1}^n a_{ij}}{m \cdot n}.$$

Сумма квадратов отклонений по всем $m \cdot n$ измерений, то есть по данным всего эксперимента:

$$Q = \sum_{i=1}^m \sum_{j=1}^n (a_{ij} - \bar{a})^2.$$

Эту сумму квадратов отклонений можно разложить на два независимых слагаемых:

$$Q = n \sum_{i=1}^m (\bar{a}_i - \bar{a})^2 + \sum_{i=1}^m \sum_{j=1}^n (a_{ij} - \bar{a}_i)^2.$$

Обозначим:

$$Q_1 = n \sum_{i=1}^m (\bar{a}_i - \bar{a})^2, \quad Q_2 = \sum_{i=1}^m \sum_{j=1}^n (a_{ij} - \bar{a}_i)^2.$$

Что такое Q_1 и Q_2 ?

Q_1 – сумма квадратов отклонений между вариантами фактора, так как \bar{a}_i – среднее значение измеренного параметра i -го варианта фактора;

Q_2 – характеризует отклонения внутри каждого варианта.

Если принятая гипотеза о равенстве центров рассеивания a_i и дисперсий σ^2 верна, тогда все $m \cdot n$ наблюдений значений a_{ij} можно рассматривать как выборку из одной и той же нормальной совокупности с очевидной несмещённой оценкой дисперсии:

$$S^2 = \frac{Q}{m \cdot n - 1}.$$

Можно показать, что величина

$$\frac{1}{m-1} Q_1,$$

имеющая распределение χ^2 со степенями свободы $m-1$, является оценкой дисперсии S^2 .

И величина

$$\frac{1}{m(n-1)} Q_2,$$

имеющая распределение χ^2 со степенями свободы $m(n-1)$, также является оценкой дисперсии S^2 .

Из сказанного следует, что критерий

$$F = \frac{\frac{1}{m-1} Q_1}{\frac{1}{m(n-1)} Q_2}$$

при нашей гипотезе и независимости Q_1 и Q_2 (это можно доказать) имеет F -распределение с $m-1$ и $m(n-1)$ степенями свободы.

А дальше мы уже знаем, как поступить:

выбираем уровень значимости q ;

вычисляем число F ;

из таблицы по величине $\frac{q}{2}$ находим F_2 .

Если окажется

$$F > F_2,$$

то есть мы попали в область маловероятных значений F , то выдвинутая гипотеза *не подтверждается*. А это значит, что варианты фактора не однотипны.

Но если

$$F \leq F_2,$$

то гипотеза об однородности вариантов фактора подтверждается, конечно, в рамках допустимого риска.

Пример 5.5. Необходимо проверить однотипность патронов к автомату Калашникова, изготовленных на трёх заводах.

Для получения необходимых для дисперсионного анализа данных автомат закрепили в специальном станке и сделали из него по 50 выстрелов патронами каждого завода. По результатам стрельбы измерялись радиальные отклонения пробоев от точки прицеливания.

Результаты измерений приведены в табл. 5.8.

Таблица 5.8

Результаты стрельбы

| Заводы | Эксперименты и отклонения, см | | | | | | | | |
|--------|-------------------------------|---|---|-----|----|----|-----|----|----|
| | 1 | 2 | 3 | ... | 26 | 27 | ... | 49 | 50 |
| № 1 | 3 | 2 | 1 | ... | 4 | 3 | ... | 1 | 3 |
| № 2 | 2 | 0 | 4 | ... | 3 | 2 | ... | 2 | 3 |
| № 3 | 2 | 3 | 3 | ... | 1 | 0 | ... | 1 | 5 |

Решение

Проверяем исходную гипотезу: патроны, выпускаемые на трех разных заводах, баллистически однотипны.

При выборе уровня значимости q исходим из того, что более опасна ошибка второго рода – подтвердить ошибочный выбор. Примем $q = 10\%$.

Число вариантов фактора: $m = 3$.

Число измерений: $n = 50$.

При вычислениях опустим очевидные элементарные арифметические детали.

Средние отклонения пробоев при стрельбе патронами заводов № 1, № 2, № 3 равны соответственно:

$$\bar{a}_1 = \frac{3+2+1+\dots+4+3+\dots+1+3}{50} = 2,5 \text{ см},$$

$$\bar{a}_2 = \frac{2+0+4+\dots+3+2+\dots+2+3}{50} = 2,8 \text{ см},$$

$$\bar{a}_3 = \frac{2+3+3+\dots+1+0+\dots+1+5}{50} = 3,1 \text{ см}.$$

Среднее отклонение по 150 выстрелам:

$$\bar{a} = \frac{2,5+2,8+3,1}{3} = 2,8 \text{ см}.$$

Средний квадрат расхождений между вариантами факторов:

$$\begin{aligned} \frac{1}{m-1} Q_1 &= \frac{1}{m-1} n \sum_{i=1}^m (\bar{a}_i - \bar{a})^2 = \\ &= \frac{50}{2} [(2,5 - 2,8)^2 + (2,8 - 2,8)^2 + (3,1 - 2,8)^2] = 4,5 \text{ см}^2. \end{aligned}$$

Число степеней свободы: $m - 1 = 3 - 1 = 2$.

Средний квадрат расхождений внутри вариантов:

$$\begin{aligned} \frac{1}{m(n-1)} Q_2 &= \frac{1}{m(n-1)} \sum_{i=1}^m \sum_{j=1}^n (a_{ij} - \bar{a}_i)^2 = \frac{1}{3 \cdot 49} [(3-2,5)^2 + (2-2,5)^2 + \\ &+ (1-2,5)^2 + \dots + (4-2,5)^2 + (3-2,5)^2 + \dots + (1-2,5)^2 + (3-2,5)^2 + \\ &+ (2-2,8)^2 + \dots + (5-3,1)^2] = 2,2 \text{ см}^2. \end{aligned}$$

Число степеней свободы:

$$m(n-1) = 3 \cdot 49 = 147.$$

Расчёт F -критерия:

$$F = \frac{4,5}{2,2} \approx 2,1.$$

По табл. (приложение 2) при $\frac{q}{2} = 5\%$ верхних пределов отклонения величины F и имеющихся степенях свободы 2 и 147 находим $F_2 = 3,06$. Величина F_2 определена при $R_2 = 150$, так как табличные значения при $k_2 = 147$ не определены. Нетрудно убедиться, что такое приближение вполне допустимо.

Поскольку $2,1 < 3,06$, делаем вывод о том, что выдвинутая гипотеза об однотипности партий патронов к автомату Калашникова, выпускаемых тремя разными заводами, не опровергается (в пределах принятого уровня значимости).

5.8. Выявление несущественных факторов

Большое количество факторов усложняет и снижает эффективность эксперимента. Среди этого множества могут быть и несущественные факторы. Исключение их упростило бы эксперимент, не снижая его информативности.

Несущественный фактор выявляется так.

Выполняются первый эксперимент из N наблюдений с учётом проверяемого фактора и второй эксперимент также из N наблюдений – без него. В обоих случаях фиксируются отклики y . Делается предположение, что обе выборки принадлежат одной генеральной совокупности, то есть, что проверяемый фактор – несущественный (это нулевая гипотеза). Дальнейшие рассуждения должны либо *не опровергнуть* эту гипотезу, либо посчитать её *недостаточно обоснованной*.

Итак, получены две последовательности откликов, в которой y'_i и y''_i – значения откликов в i -м наблюдении при наличии и отсутствии проверяемого фактора соответственно:

$$y'_1, y'_2, y'_3, \dots, y'_i, \dots, y'_N ;$$

$$y''_1, y''_2, y''_3, \dots, y''_i, \dots, y''_N .$$

Согласно принятой гипотезе эти последовательности имеют одинаковые математические ожидания $M[y]$ и дисперсии σ_y^2 .

Рассмотрим случайную величину Z , реализациями которой является последовательность случайных чисел

$$z_i = y'_i - y''_i, \quad i = \overline{1, N} .$$

При независимости z_i и достаточно большом числе наблюдений N согласно центральной предельной теореме:

$$M[\bar{z}] = M[z] = 0, \quad \sigma_z = \frac{\sigma_z}{\sqrt{N}} .$$

Очевидно:

$$\sigma_z^2 = 2\sigma_y^2, \quad \sigma_z = \sigma_y \sqrt{2} .$$

Как отделить случайные отклонения \bar{z} от нуля от тех, которые мы будем считать не подтверждающими принятую гипотезу?

Такое разделение осуществляется по следующему правилу: если вычисленная величина \bar{z} окажется маловероятной, в рамках нормального распределения и данном среднеквадратическом отклонении σ_z , то такое отклонение \bar{z} от нуля считается не соответствующим принятой гипотезе.

Эту малую вероятность называют уровнем значимости и обозначают q . Обычно $q = 2\%, 5\%, 10\%$ – в зависимости от степени опасности совершения ошибки первого или второго рода.

На графике плотности распределения $f(\bar{z})$ уровень значимости q показан заштрихованным участком (рис. 5.4).

Для нормального закона распределения случайной величины \bar{z} вероятность превышения \bar{z} некоторого значения определяется известным выражением:

$$P\left(\left|\bar{z} - 0\right| > t_\alpha \frac{\sigma_z}{\sqrt{N}}\right) = 1 - 2\Phi(t_\alpha),$$

$$P\left(\left|\bar{z}\right| > t_\alpha \frac{\sigma_z}{\sqrt{N}}\right) = 1 - 2\Phi(t_\alpha),$$

где $\Phi(t_\alpha)$ – функция Лапласа.

Следовательно:

$$\text{границное значение } \bar{z}_q = t_\alpha \frac{\sigma_z}{\sqrt{N}};$$

$$\left[1 - 2\Phi(t_\alpha)\right] = q.$$

Аргумент функции Лапласа t_α находим из соответствующего справочника согласно $t_\alpha = \Phi^{-1}\left(\frac{1-q}{2}\right)$ и, как было указано ранее,

$$\sigma_z = \sigma_y \sqrt{2}.$$

Из изложенного следует:

если $\bar{z} > t_\alpha \frac{\sigma_y \sqrt{2}}{\sqrt{N}}$, принятая гипотеза о несущественности проверяемого фактора не подтверждается;

если $\bar{z} \leq t_\alpha \frac{\sigma_y \sqrt{2}}{\sqrt{N}}$, принятая гипотеза не опровергается (в рамках принятого уровня значимости q).

Обычно величина σ_y неизвестна, поэтому следует использовать её оценку S_y :

$$S_y = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (\bar{y} - y_i)^2}.$$

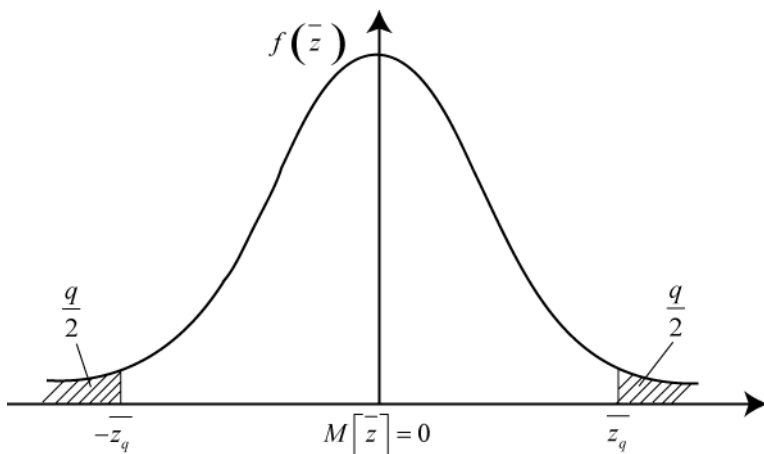


Рис. 5.4. Плотность распределения функции $f(\bar{z})$

Оценку \bar{y} и ряд значений y_i можно получить из данных первого эксперимента (y_i') или второго (y_i''), так как в силу рассматриваемой гипотезы они идентичны. Однако следует помнить, что если $N < 100$, то вместо аргумента функции Лапласа t_α надо брать t'_α – аргумент функции Стьюдента.

Пример 5.6. Исследуется зависимость времени пребывания заявки в системе массового обслуживания от дисциплины выборки заявок из очереди: LIFO или FIFO. Проведены два эксперимента. Первый эксперимент из $N = 100$ наблюдений с дисциплиной FIFO и второй эксперимент также из $N = 100$ наблюдений с дисциплиной LIFO.

Решение

Выдвигается гипотеза о несущественности влияния на время пребывания заявки в системе массового обслуживания изменения дисциплины FIFO на LIFO.

Результаты измерений и вычислений:

$$\bar{z} = 1,8 \text{ мин}, S_y = 2 \text{ мин};$$

для уровня значимости $q = 5\%$, $t_\alpha = 1,96$

$$\bar{z}_{q} \leq t_\alpha \frac{S_y \sqrt{2}}{\sqrt{N}} = 1,96 \frac{2 \cdot 1,41}{10} = 0,55.$$

Так как $\bar{z} = 1,8 > \bar{z}_q = 0,55$, то гипотеза не подтверждается. Для времени пребывания заявки в очереди в системе массового обслуживания не безразлично, какая дисциплина выборки заявок из очереди применена.

В заключение отметим, что рассмотренную проблему можно решать и методом однофакторного дисперсионного анализа. Однако если $m = 2$ (при сравнении двух выборок) целесообразно использовать изложенный метод.

5.9. Сущность корреляционного анализа

Часто при исследовании объекта или его модели необходимо наблюдать за характеристиками двух и более случайных величин. Например, за двумя откликами одного эксперимента. При этом может возникнуть вопрос: есть ли связь между этими случайными величинами? Существенна или несущественна эта связь, если она есть?

Корреляционный анализ – это совокупность методов обнаружения зависимости (корреляции) между двумя или более случайными признаками или процессами.

Под *корреляцией* будем понимать статистическую зависимость между двумя случайными величинами, не имеющую, вообще говоря, строго функционального характера.

Заметим, что корреляционный анализ не позволяет определить вид функциональной связи между случайными величинами, а только наличие или отсутствие предполагаемой связи, например, линейной, параболической, экспоненциальной и т. д. В рамках этого учебного пособия мы ограничимся рассмотрением гипотезы о наличии линейной корреляции.

Определение вида функциональной связи между величинами рассматривается в регрессионном анализе, элементы которого и практическое использование будут рассмотрены в следующем п. 5.10.

Название «корреляционный анализ» происходит от латинского слова *correlatio* – согласование, связь, соотношение, взаимосвязь. Термин впервые введен Гальтоном (1822-1911) в 1888 г.

Обычно исследуют парную корреляцию, то есть зависимость между двумя случайными величинами (процессами), хотя возможны и более сложные ситуации, когда необходимо обнаружить наличие или отсутствие связей между тремя или более случайными величинами.

Мы ограничимся исследованием парной корреляции.

Как известно, связь между двумя случайными величинами можно описать с помощью двумерной функции распределения. Однако такое описание часто очень сложно, а для практических целей можно удовлетвориться определением зависимостей средних значений.

Итак, целью имитационного эксперимента является определение характеристик двух случайных величин a и b . Например:

| a | b |
|---|---|
| Средний балл успеваемости учебной группы по математике. | Средний балл выполнения упражнения по стрельбе. |

| | |
|---|--|
| Рассеивание точки падения снаряда по дальности. | Рассеивание точки падения заряда по боковому отклонению. |
|---|--|

| | |
|----------------|--|
| Вес курсантов. | Успеваемость по физической подготовке. |
|----------------|--|

Необходимо проверить: есть ли связь между величинами a и b ?

Проверка наличия (или отсутствия) связи – корреляции – между случайными величинами выполняется так.

Проводятся два эксперимента, каждый – с соответствующей моделью. В каждом эксперименте – N наблюдений (напоминаем, что компьютерный эксперимент состоит из наблюдений, а наблюдение – из реализаций (прогонов) модели, число которых рассчитывается с учетом требуемой точности и достоверности получаемых результатов моделирования). В результате экспериментов получают два множества значений измеряемых параметров a и b : a_i и b_i , $i = \overline{1, N}$.

Из этих множеств формируются пары:

$$(a_1, b_1), (a_2, b_2), \dots, (a_i, b_i), \dots, (a_N, b_N).$$

Каждая пара интерпретируется как координаты случайной точки в системе координат a, b .

Первичное исследование можно провести графически. Возможны следующие варианты размещения точек на графиках (рис. 5.5).

Корреляция – важное понятие. Научитесь визуально определять по расположению данных, насколько тесно они коррелированы.

Говорят, что две переменные *положительно* коррелированы, если при увеличении значений одной переменной увеличиваются значения другой переменной (рис. 5.5а, б).

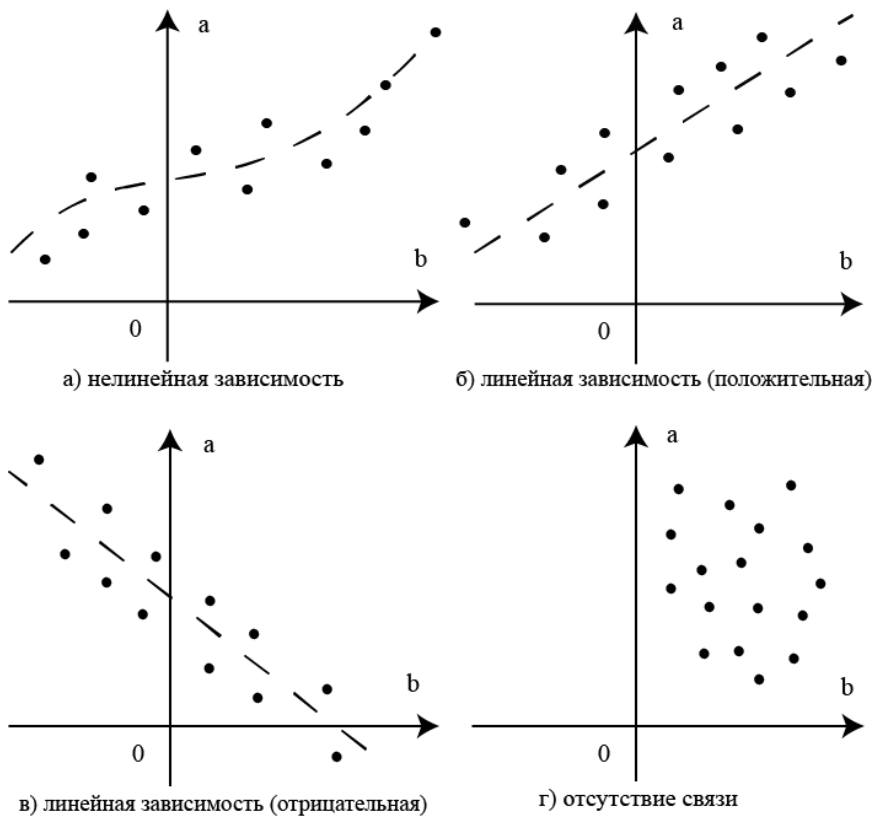


Рис. 5.5. Графическое исследование корреляции

Две переменные *отрицательно* коррелированы, если при увеличении одной переменной другая переменная уменьшается (рис. 5.5в).

Отсутствие корреляции – совместного поведения переменных – обнаруживается хаотическим нагромождением точек, исключающим проведение какой-либо аппроксимирующей линии (см. рис. 5.5г).

Но такое качественное исследование недостаточно. Необходимо иметь количественную оценку степени корреляции между величинами a и b .

Если совместное распределение вероятностей случайных величин a и b нормальное, то количественной характеристикой степени линейной связи между ними является коэффициент корреляции r (введен Пирсоном (Pearson), 1896 г.): $-1 \leq r \leq 1$.

Если $r = 0$, то между a и b линейная независимость.

Равенство $r = \pm 1$ свидетельствует о наличии однозначной функциональной связи между a и b , то есть $b = f(a)$.

При $-1 < r < 1$ между a и b существует стохастическая связь, причём, чем ближе коэффициент корреляции $|r|$ к единице, тем эта связь сильнее. Стохастическая связь означает, что при изменении a имеется лишь тенденция к изменению b .

Коэффициент корреляции r определяется по данным эксперимента, следовательно, можно определить только его оценку \bar{r} . В качестве оценки \bar{r} принят выборочный коэффициент корреляции:

$$\bar{r} = \frac{\frac{1}{N} \sum_{i=1}^N (a_i - \bar{a})(b_i - \bar{b})}{S_a \cdot S_b}, \quad (5.1)$$

где $\bar{a} = \frac{1}{N} \sum_{i=1}^N a_i$, $\bar{b} = \frac{1}{N} \sum_{i=1}^N b_i$ – оценки математических ожиданий $M[a]$ и $M[b]$;

$$S_a = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (a_i - \bar{a})^2}, S_b = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (b_i - \bar{b})^2} \text{ – оценки}$$

средне-кватратических отклонений σ_a и σ_b .

Выборочный коэффициент корреляции \bar{r} , так же как и теоретический, принимает значения: $-1 \leq \bar{r} \leq 1$.

Если $\bar{r} > 0$, то наблюдается положительная корреляция (см. рис. 5.5б). Если $\bar{r} < 0$ – отрицательная корреляция (см. рис. 5.5в). Если $\bar{r} = 0$ – линейная корреляция отсутствует (но не исключена нелинейная). Если $\bar{r} = 1$, то между случайными величинами существует жесткая функциональная связь.

Заметим, что рассматриваемый коэффициент корреляции определяет степень *линейной* связи между случайными величинами a и b . Эта корреляция наиболее популярна, поэтому часто, когда говорят о корреляции, имеют в виду именно корреляцию Пирсона.

Однако этот линейный коэффициент корреляции не является пригодным для оценки нелинейной связи, если таковая присутствует. При

нелинейной зависимости степень связи между случайными величинами устанавливается более сложными характеристиками, например, *корреляционным отношением* (К. Пирсон).

Числитель выражения (5.1) иногда называют *ковариацией* – $\text{cov}(a, b)$.

Если случайные величины a и b независимы, они и не коррелированы ($\bar{r} = 0$). Но некоррелированность a и b не всегда свидетельствует об их независимости. Но если a и b имеют нормальное распределение, то условие $\bar{r} = 0$ является необходимым и достаточным условием независимости этих величин.

И ещё. Наличие корреляции между случайными величинами a и b не всегда свидетельствует об их взаимосвязи. Дело в том, что при независимости a и b каждая из них в отдельности зависит от некоторого случайного фактора ξ , но эта зависимость нами не замечена.

Поэтому хорошим тоном после вычисления корреляций является построение диаграмм рассеяния, которые позволяют понять, действительно ли между двумя исследуемыми переменными имеется связь.

Оценка коэффициента корреляции должна быть определена с требуемой точностью и достоверностью, которые зависят от числа реализаций модели. Найдём эту связь.

В предположении нормальности распределения \bar{r} можно написать:

$$P\left(\left|\bar{r} - r\right| < t_\alpha \sigma_{\bar{r}}\right) = 2\Phi(t_\alpha) \quad (5.2)$$

С выражением (5.2) мы уже знакомы. Здесь:

r – точное значение коэффициента корреляции;

$\sigma_{\bar{r}}$ – среднеквадратическое отклонение случайной величины \bar{r} ;

t_α – аргумент функции Лапласа $\Phi(t_\alpha)$.

Обычно среднеквадратическое отклонение $\sigma_{\bar{r}}$ неизвестно, поэтому нужно брать её оценку.

При больших выборках N оценка среднеквадратического отклонения S_r :

$$\sigma_{\bar{r}} \approx S_r = \frac{1 - \bar{r}^2}{\sqrt{N}}.$$

Из (5.2) следует:

$$\varepsilon = t_\alpha \frac{1 - \bar{r}^2}{\sqrt{N}}, \quad N = t_\alpha^2 \frac{(1 - \bar{r}^2)^2}{\varepsilon^2}, \quad t_\alpha = \Phi^{-1}\left(\frac{\alpha}{2}\right),$$

ε – абсолютная величина ошибки.

Предварительное определение \bar{r} осуществляется по данным пробного эксперимента в количестве $N^* = 500 \dots 1000$ реализаций модели.

На основании изложенного и в силу случайного характера исследуемых величин a и b мы можем утверждать лишь следующее: истинное значение коэффициента корреляции r лежит в пределах

$$\bar{r} - \varepsilon \leq r \leq \bar{r} + \varepsilon$$

с заданной достоверностью α .

В заключение отметим, что если совместное распределение случайных величин a и b не является нормальным, то оценка \bar{r} коэффициента корреляции может выступать в качестве ориентировочной оценки степени тесноты связи a и b .

Пример 5.7. Для оценки конструкции нового крупнокалиберного пулемёта было произведено 96 выстрелов по щиту, отстоявшему на расстоянии 300 метров.

Результаты отклонений попаданий от точки прицеливания (боковые x_i , по высоте y_i) объединены в десятисантиметровые диапазоны и сведены в таблицу (табл. 5.9).

Для оценки конструктивных особенностей пулемёта необходимо узнать: есть ли какая-то связь между боковыми отклонениями и отклонениями по высоте?

Решение

Ответ на поставленный вопрос о наличии или отсутствии связи может дать коэффициент корреляции.

Предварительно заметим, что группировка измерений в десятисантиметровые диапазоны вносит некоторую ошибку в дальнейшие расчёты, однако можно показать, что при данной группировке ошибка несущественна.

В табл.5.9 указаны не реальные отклонения, а центры диапазонов (-25...-15, -15...-5, -5...5 и т. д.).

Таблица 5.9

Отклонения от точки прицеливания

| y_j | Боковые отклонения x_i | | | | | | | Всего n_{y_j} |
|-----------------|--------------------------|-----|----|----|----|----|----|-----------------|
| | -20 | -10 | 0 | 10 | 20 | 30 | 40 | |
| -50 | 0 | 0 | 1 | 0 | 2 | 0 | 0 | 3 |
| -40 | 0 | 1 | 1 | 1 | 2 | 0 | 0 | 5 |
| -30 | 1 | 1 | 3 | 5 | 2 | 1 | 0 | 13 |
| -20 | 1 | 3 | 7 | 3 | 2 | 2 | 0 | 18 |
| -10 | 0 | 2 | 6 | 10 | 3 | 0 | 0 | 21 |
| 0 | 0 | 1 | 6 | 6 | 6 | 1 | 1 | 21 |
| 10 | 0 | 0 | 3 | 3 | 3 | 1 | 0 | 10 |
| 20 | 0 | 1 | 1 | 2 | 1 | 0 | 0 | 5 |
| Всего n_{x_i} | 2 | 9 | 28 | 30 | 21 | 5 | 1 | 96 |

Для определения коэффициента корреляции понадобятся следующие характеристики:

$$\bar{x}, S_x, \bar{y}, S_y, \text{ ковариация } \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}).$$

Все эти характеристики вычисляются по данным измеренных отклонений боковых x и по высоте y .

Для примера, расчет \bar{x} :

$$\bar{x} = \frac{(-20 \cdot 2) + (-10 \cdot 9) + (0 \cdot 28) + \dots + (40 \cdot 1)}{96} = \frac{780}{96} \approx 8,1 \text{ см.}$$

Результаты расчета остальных характеристик:

$$S_x = 11,6 \text{ см, } \bar{y} = -11,6 \text{ см, } S_y = 16,8 \text{ см, } \text{cov}(x, y) = 1518 \text{ см}^2.$$

Теперь оценка коэффициента корреляции:

$$\bar{r} = \frac{1518}{96 \cdot 11,6 \cdot 16,8} = \frac{1518}{18525} = 0,082.$$

Среднеквадратическое отклонение этой оценки:

$$S_r = \frac{1 - \bar{r}^2}{\sqrt{N}} = \frac{1 - 0,082^2}{\sqrt{96}} \approx 0,1 \text{ см.}$$

Из-за малого количества выстрелов оценка \bar{r} определена с ошибкой, которая в предположении о нормальном распределении

случайной величины \bar{r} и достоверности, например, $\alpha = 0,8$ ($t_\alpha = 1,28$) равна:

$$\varepsilon = t_\alpha \sigma_r \approx t_\alpha S_r = 1,28 \cdot 0,1 = 0,128 \text{ см.}$$

Отсюда следует, что истинное значение коэффициента корреляции r лежит в пределах:

$$0,082 - 0,128 \leq r \leq 0,082 + 0,128, \quad -0,046 \leq r < 0,21.$$

Обнаружена небольшая линейная зависимость отклонений боковых и отклонений по высоте. Баллистики, отвергая непосредственную корреляцию между отклонениями x и y , объясняют значение $\bar{r} > 0$ влиянием конструктивных особенностей пулемета. Обнаружена также систематическая ошибка в прицеле: $\bar{x} \neq 0, \bar{y} \neq 0$.

5.10. Обработка результатов эксперимента на основе регрессии

Часто целью исследования является определение функциональной связи между факторами и откликом (реакцией модели) по данным, полученным при экспериментах с моделью объекта или непосредственно с объектом. Такая цель достигается регрессионным анализом значений факторов x и отклика y .

Под *регрессией* в теории вероятностей и математической статистике понимают зависимость среднего значения какой-либо величины от некоторой другой (других) величины. *Регрессионный анализ* – это совокупность методов построения и исследования регрессионной зависимости между величинами (в нашем случае между факторами и откликом) по статистическим данным. Статистические данные определяются и накапливаются при проведении эксперимента.

Формальная схема эксперимента выглядит так (рис. 5.6).

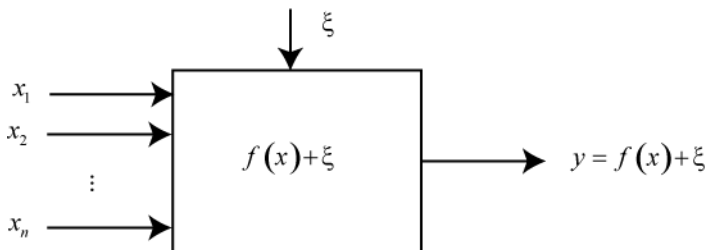


Рис. 5.6. Формальная схема эксперимента

Прямоугольник представляет исследуемый объект или его математическую модель. Обозначения на рис. 5.6:

x_i – значения факторов, $i = \overline{1, n}$;

ξ – случайный фактор, помеха. Будем считать, что эта случайная величина имеет нормальное распределение с матожиданием $M[\xi] = 0$. Влияние помехи на отклик аддитивное, то есть её случайные значения прибавляются к значениям отклика;

$f(x)$ – искомая функциональная зависимость между факторами и откликом.

Отклик y – величина случайная. $f(x)$ представляет собой среднее значение отклика (так как $M[\xi] = 0$): $\bar{y} = f(x)$.

Исследуемый объект представляется как «чёрный ящик», никаких предположений о виде функции $f(x)$ нет. Поэтому представим её в виде аппроксимирующего полинома:

$$\bar{y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n = \sum_{i=0}^n \beta_i x_i, \quad x_0 = 1.$$

Этот полином получил название уравнения регрессии, а коэффициенты β_i – коэффициенты регрессии. От точности подбора коэффициентов регрессии зависит точность представления $f(x)$.

Коэффициенты β_i определяются путем обработки полученных в ходе эксперимента варьируемых значений факторов и откликов.

Однако из-за ограниченного числа наблюдений точные значения β_i получить нельзя, будут найдены их оценки b_i :

$$\bar{\beta}_i = b_i.$$

Поэтому уравнение регрессии принимает вид:

$$\bar{y} = b_0 + b_1 x_1 + b_2 x_2 + \dots + b_i x_i + \dots + b_n x_n = \sum_{i=0}^n b_i x_i, \quad x_0 = 1.$$

Вообще-то метку над \bar{y} теперь надо бы изменить, так как вместо β_i в уравнении теперь стоят b_i , но мы этого делать не будем, чтобы не загромождать изложение новыми знаками.

В уравнении регрессии могут участвовать и так называемые «совместные эффекты» (x_1x_2 , $x_1x_2x_3$ и т. п.) или степени значений факторов (x_1^2 , x_2^3 и т. п.). Совместные эффекты и степени факторов можно обозначать обобщённым фактором.

Например, уравнение регрессии

$$\bar{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_1x_2 + b_4x_2^2$$

можно представить так:

$$\bar{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4, \quad x_3 = x_1x_2, \quad x_4 = x_2^2$$

Итак, для определения выражения $f(x)$ надо:

выбрать степень аппроксимирующего полинома – уравнения регрессии;

определить коэффициенты регрессии.

Выбор уравнения регрессии обычно начинают с линейной модели. Например, для двухфакторного эксперимента её вид:

$$\bar{y} = b_0 + b_1x_1 + b_2x_2.$$

Если окажется, что такая аппроксимация даёт неприемлемые отклонения при сравнении с экспериментальными точками отклика y , то модель усложняется, например, так:

$$\bar{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_1x_2 \quad \text{или}$$

$$\bar{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_1^2 + b_4x_2^2, \quad \text{и т. д.}$$

Коэффициенты регрессии b_i для выбранного уравнения определяются из условия минимума суммы квадратов ошибок, вычисленных по всем экспериментальным точкам.

Это делается так.

Введём обозначения:

x_{il} – значение i -го фактора в наблюдении номер l ;

y_l – значение отклика в l -м наблюдении;

\bar{y} – значение отклика, вычисленное по принятому уравнению регрессии и данным x_{il} .

Очевидно, сумма квадратов ошибок между экспериментальными значениями y_l и вычисленными по уравнению регрессии \bar{y}_l для всех N наблюдений равна:

$$\delta = \sum_{l=1}^N (y_l - \bar{y}_l)^2 = \sum_{l=1}^N \left(y_l - \sum_{i=0}^n b_i x_{li} \right)^2.$$

Для определения минимума ошибки δ возьмём частные производные от δ по всем неизвестным коэффициентам регрессии b_j , $j = \overline{1, n}$ и приравняем их нулю:

$$\frac{\partial \delta}{\partial b_j} = -2 \sum_{l=1}^N \left(y_l - \sum_{i=0}^n b_i x_{li} \right) x_{jl} = 0.$$

Нетрудно убедиться, что это условие минимума, а не максимума. Очевидно:

$$\begin{aligned} \sum_{l=1}^N \left(y_l - \sum_{i=0}^n b_i x_{li} \right) x_{jl} &= 0, \\ \sum_{l=1}^N y_l x_{jl} &= \sum_{l=1}^N \sum_{i=0}^n b_i x_{li} x_{jl}. \end{aligned}$$

Для лучшей наглядности выделим неизвестные коэффициенты регрессии и получим:

$$\sum_{l=1}^N y_l x_{jl} = \sum_{l=1}^N b_i \sum_{i=0}^n x_{li} x_{jl}. \quad (5.3)$$

Выражение (5.3) представляет собой систему из $n+1$ уравнений для нахождения $n+1$ неизвестных коэффициентов регрессии b_i , которые окончательно определяют выбранное уравнение регрессии.

Нахождение коэффициентов регрессии справедливо при следующих допущениях:

1. Случайный фактор ξ имеет нормальное распределение с математическим ожиданием $M[\xi] = 0$.

2. Результаты наблюдений y_l – независимые нормально распределенные случайные величины. Если это не соблюдается, то следует измерять другой отклик, удовлетворяющий этому условию, но функционально связанный с исследуемым откликом y .

3. Точность наблюдений (количество реализаций модели) не меняется от наблюдения к наблюдению.

4. Точность наблюдения x_{il} должна быть выше точности y_l .

Пример 5.8. На модели объекта проведен однофакторный эксперимент из пяти наблюдений, результаты которого сведены в таблицу (табл. 5.10).

Найти функциональную связь фактора с откликом $\bar{y} = f(x)$.

Таблица 5.10

Результаты эксперимента

| Фактор и отклики | Наблюдение | | | | | $\sum_{l=1}^N$ |
|---------------------|------------|-----|-----|-----|-----|----------------|
| | 1 | 2 | 3 | 4 | 5 | |
| x_{il} | 0 | 0,5 | 1,0 | 1,5 | 2,0 | 5 |
| y_l | 7,0 | 4,8 | 2,8 | 1,4 | 0 | 16 |
| $y_l x_{il}$ | 0 | 2,4 | 2,8 | 2,1 | 0 | 7,3 |

Решение

Примем, что кроме управляемого фактора x_{il} при проведении эксперимента на объект воздействует случайный фактор, распределенный по нормальному закону с математическим ожиданием $M[\xi] = 0$. Также предположим, что эта связь – линейная, следовательно, уравнение регрессии нужно определять в виде:

$$\bar{y} = b_0 + b_1 x_1.$$

Неизвестных коэффициентов два: b_0 и b_1 . Запишем (5.3) в виде двух уравнений для $j=0$, $j=1$ и в каждом из них разложим суммы по индексу i :

$$\begin{cases} \sum_{l=1}^N y_l x_{0l} = b_0 \sum_{l=1}^N x_{0l} x_{0l} + b_1 \sum_{l=1}^N x_{1l} x_{0l}, \\ \sum_{l=1}^N y_l x_{1l} = b_0 \sum_{l=1}^N x_{0l} x_{1l} + b_1 \sum_{l=1}^N x_{1l} x_{1l}. \end{cases}$$

Так как $x_{0l} = 1$, получим:

$$\begin{cases} \sum_{l=1}^N y_l = N b_0 + b_1 \sum_{l=1}^N x_{1l}, \\ \sum_{l=1}^N y_l x_{1l} = b_0 \sum_{l=1}^N x_{1l} + b_1 \sum_{l=1}^N x_{1l}^2. \end{cases} \quad (5.4)$$

Подставим данные эксперимента из табл. 5.10 в систему (5.4):

$$\begin{cases} 16 = 5b_0 + 5b_1, \\ 7,3 = 5b_0 + 7,5b_1. \end{cases}$$

Решим систему из двух уравнений и получим: $b_0 = 6,68$, $b_1 = -3,48$.
Следовательно, искомое уравнение регрессии:

$$\bar{y} = 6,68 - 3,48x_1.$$

Доверительные границы для истинных значений β_0 и β_1 примера 5.8 определяются как обычно:

$$b_0 - t_\alpha^* \sigma_{b_0} \leq \beta_0 \leq b_0 + t_\alpha^* S_{b_0}; \quad b_1 - t_\alpha^* \sigma_{b_1} \leq \beta_1 \leq b_1 + t_\alpha^* S_{b_1},$$

где t_α^* – аргумент распределения Стьюдента;

S_{b_0} , S_{b_1} – среднеквадратические отклонения величин b_0 и b_1 соответственно.

Значения t_α^* определяются из таблицы распределения Стьюдента для $N - 2 = 3$ степеней свободы и достоверности α . Пусть $\alpha = 0,9$, тогда $t_\alpha^* \approx 2,35$.

Значения S_{b_0} , S_{b_1} находятся по формулам:

$$S_{b_0} = \sqrt{\frac{\sum_{l=1}^N d_{y_l}^2}{N(N-2)}}, \quad S_{b_1} = \sqrt{\frac{1}{N-2} \frac{\sum_{l=1}^N d_{y_l}^2}{\sum_{l=1}^N d_{x_l}^2}}.$$

Данные для вычисления S_{b_0} , S_{b_1} представлены в табл. 5.11.

Таблица 5.11

Данные для вычисления S_{b_0} , S_{b_1}

| l | x_l | $d_{x_l} = \bar{x} - x_l$ | $d_{x_l}^2$ | y_l | \bar{y}_l | $d_{y_l} = \bar{y}_l - y_l$ | $d_{y_l}^2$ |
|-----|-------|---------------------------|-------------|-------|-------------|-----------------------------|-------------|
| 1 | 0 | 1,0 | 1,00 | 7,0 | 6,68 | -0,32 | 0,1024 |
| 2 | 0,5 | 0,5 | 0,25 | 4,8 | 4,94 | 0,14 | 0,0196 |
| 3 | 1,0 | 0 | 0 | 2,8 | 3,20 | 0,40 | 0,1600 |
| 4 | 1,5 | -0,5 | 0,25 | 1,4 | 1,46 | 0,06 | 0,0036 |
| 5 | 2,0 | -1,0 | 1,00 | 0 | 0,28 | 0,28 | 0,0784 |

$$\bar{x} = \frac{\sum_{l=1}^N x_l}{N} = 1, \quad \sum_{l=1}^N d_{x_l}^2 = 2,5, \quad \sum_{l=1}^N d_{y_l}^2 = 0,364.$$

$$S_{b_0} = \sqrt{\frac{0,364}{5(5-2)}} = 0,156, \quad S_{b_1} = \sqrt{\frac{0,364}{3 \cdot 2,5}} = 0,22.$$

С уровнем достоверности $\alpha = 0,9$ ($t_\alpha^* = 2,35$)

$$6,68 - 2,35 \cdot 0,156 \leq \beta_0 \leq 6,68 + 2,35 \cdot 0,156, \quad 6,31 \leq \beta_0 \leq 7,05;$$

$$-3,48 - 2,35 \cdot 0,22 \leq \beta_1 \leq -3,48 + 2,35 \cdot 0,22, \quad -4,0 \leq \beta_1 \leq -2,96.$$

Большой размах доверительных границ объясняется малым числом наблюдений в данном эксперименте.

Доверительные границы для \bar{y} принимают разные значения в зависимости от значений факторов.

На практике часто ограничиваются обобщёнными оценками адекватности построенной модели: величиной среднего абсолютного отклонения

$$\varepsilon = \frac{\sum_{i=1}^N (|y_i - \bar{y}_i|)}{N},$$

или (и) величиной среднеквадратической ошибки на единицу веса

$$\varepsilon_0 = \sqrt{\frac{\sum_{i=1}^N (|y_i - \bar{y}_i|)}{N - (n+1)}}, \quad N \gg (n+1).$$

Весом или *степенью свободы* эксперимента называют разность $N - (n+1)$ между числом наблюдений N и числом коэффициентов регрессии $(n+1)$.

Предположим, что линейная модель $\bar{y} = 6,68 - 3,48x_1$ недостаточно точно отображает связь между фактором x_1 и откликом y .

Введем в рассмотрение более сложную нелинейную модель:

$$\bar{y} = b_0 + b_1x_1 + b_2x_1^2.$$

Для определения коэффициентов регрессии b_0, b_1, b_2 обозначим $x_1^2 = x_2$ и получим двухфакторную линейную модель:

$$\bar{y} = b_0 + b_1x_1 + b_2x_2.$$

В этом случае уравнение (5.3) раскрывается так:

$$\begin{cases} \sum y_l x_{0l} = b_0 \sum x_{0l} x_{0l} + b_1 \sum x_{1l} x_{0l} + b_2 \sum x_{2l} x_{0l}, \\ \sum y_l x_{1l} = b_0 \sum x_{0l} x_{1l} + b_1 \sum x_{1l} x_{1l} + b_2 \sum x_{2l} x_{1l}, \\ \sum y_l x_{2l} = b_0 \sum x_{0l} x_{2l} + b_1 \sum x_{1l} x_{2l} + b_2 \sum x_{2l} x_{2l}. \end{cases}$$

В уравнениях принято: $\sum = \sum_{l=1}^N$.

Так как $x_{0l} = 1$, $x_{2l} = x_{1l}^2$, то система принимает вид:

$$\begin{cases} \sum y_l = b_0 N + b_1 \sum x_{1l} + b_2 \sum x_{1l}^2 \\ \sum y_l x_{1l} = b_0 \sum x_{1l} + b_1 \sum x_{1l}^2 + b_2 \sum x_{1l}^3, \\ \sum y_l x_{1l}^2 = b_0 \sum x_{1l}^2 + b_1 \sum x_{1l}^3 + b_2 \sum x_{1l}^4. \end{cases}$$

Подставим значения фактора и отклика из табл. 5.10:

$$\begin{cases} 16 = 5b_0 + 5b_1 + 7,5b_2, \\ 7,3 = 5b_0 + 7,5b_1 + 12,5b_2, \\ 7,15 = 7,5b_0 + 12,5b_1 + 22,125b_2. \end{cases}$$

Решим систему из трёх уравнений с тремя неизвестными и получим: $b_0 = 7$, $b_1 = -4,74$, $b_2 = 0,63$.

Таким образом, получено новое уравнение регрессии:

$$\bar{y} = 7 - 4,74x_1 + 0,63x_1^2.$$

По значениям ε и ε_0 нетрудно убедиться в том, что нелинейная модель более точно отображает моделируемый процесс (см. табл. 5.10), чем линейная.

В рассмотренном примере ошибка модели определялась по тем же данным, по которым и была определена сама модель. Однако при сокращенных планах экспериментов (см. п. 4.3) можно выполнить все или часть «сэкономленных» наблюдений для получения так называемых проверочных данных, которые и использовать для вычисления ошибки ε или ε_0 . В этом случае оценка адекватности модели будет более объективна, хотя число наблюдений в эксперименте увеличивается, и экономии их не будет.

По уравнению регрессии можно сделать ориентировочную оценку чувствительности отклика к изменению того или иного фактора.

Например, в уравнении $\bar{y} = 50,5x_1 + 0,74x_2 - 28,3x_3$ влияние фактора x_2 на отклик незначительно по сравнению с другими, так как коэффициент $b_2 = 0,74$ намного меньше остальных коэффициентов.

В пакете MS Excel есть функция «Регрессия», которая выполняет регрессионный анализ данных компьютерного эксперимента.

Пример 5.9. В ремонтное подразделение поступают вышедшие из строя средства связи (СС) с интервалами времени, подчиненными показательному закону с математическим ожиданием t . В каждом СС могут быть неисправными в любом сочетании блоки A, B, C с вероятностями P_{1A}, P_{1B}, P_{1C} соответственно. Ремонтное подразделение ремонтирует СС путем замены неисправных блоков исправными блоками. В момент поступления неисправного СС в ремонтное подразделение вероятности наличия в нем исправных блоков A, B, C, D соответственно $P_{2A}, P_{2B}, P_{2C}, P_{2D}$. Наличие и замена блока D обязательно при любом сочетании неисправных блоков.

Построить имитационную модель «Система ремонта» с целью определения вероятности ремонта СС с неисправными блоками C, D и A, B, D за время T_0 . По результатам эксперимента получить уравнение регрессии, связывающее вероятность ремонта СС с вероятностями $P_{1B}, P_{1C}, P_{2B}, P_{2C}, P_{2D}$.

Решение.

Постановка примера 5.9 аналогична постановке примера 3.11. При разработке модели использовался алгоритм примера 3.11 (см. рис. 3.20).

Для построения уравнения регрессии введём обозначения:

y – отклик модели, вероятность ремонта СС с неисправными блоками C, D и A, B, D за время T_0 ;

x_1 – фактор, представляющий вероятность P_{1B} ;

x_2 – фактор, представляющий вероятность P_{1C} ;

x_3 – фактор, представляющий вероятность P_{2B} ;

x_4 – фактор, представляющий вероятность P_{2C} ;

x_5 – фактор, представляющий вероятность P_{2D} .

Исходные данные и результаты эксперимента – 32 наблюдения приведены в табл. 5.12, результаты анализа – в табл. 5.13.

Таблица 5.12

Результаты эксперимента с моделью «Система ремонта»

| № отклика | Y | P1B | P1C | P2B | P2C | P2D |
|-----------|-------|-----|------|-----|-----|------|
| 1 | 0,088 | 0,3 | 0,55 | 0,5 | 0,2 | 0,65 |
| 2 | 0,127 | 0,3 | 0,55 | 0,5 | 0,2 | 0,95 |
| 3 | 0,303 | 0,3 | 0,55 | 0,5 | 0,8 | 0,65 |
| 4 | 0,442 | 0,3 | 0,55 | 0,5 | 0,8 | 0,95 |
| 5 | 0,099 | 0,3 | 0,55 | 0,9 | 0,2 | 0,65 |
| 6 | 0,146 | 0,3 | 0,55 | 0,9 | 0,2 | 0,95 |
| 7 | 0,317 | 0,3 | 0,55 | 0,9 | 0,8 | 0,65 |
| 8 | 0,46 | 0,3 | 0,55 | 0,9 | 0,8 | 0,95 |
| 9 | 0,116 | 0,3 | 0,85 | 0,5 | 0,2 | 0,65 |
| 10 | 0,167 | 0,3 | 0,85 | 0,5 | 0,2 | 0,95 |
| 11 | 0,445 | 0,3 | 0,85 | 0,5 | 0,8 | 0,65 |
| 12 | 0,653 | 0,3 | 0,85 | 0,5 | 0,8 | 0,95 |
| 13 | 0,12 | 0,3 | 0,85 | 0,9 | 0,2 | 0,65 |
| 14 | 0,175 | 0,3 | 0,85 | 0,9 | 0,2 | 0,95 |
| 15 | 0,452 | 0,3 | 0,85 | 0,9 | 0,8 | 0,65 |
| 16 | 0,66 | 0,3 | 0,85 | 0,9 | 0,8 | 0,95 |
| 17 | 0,118 | 0,3 | 0,55 | 0,5 | 0,2 | 0,65 |
| 18 | 0,173 | 0,9 | 0,55 | 0,5 | 0,2 | 0,95 |
| 19 | 0,336 | 0,9 | 0,55 | 0,5 | 0,8 | 0,65 |
| 20 | 0,486 | 0,9 | 0,55 | 0,5 | 0,8 | 0,95 |
| 21 | 0,158 | 0,9 | 0,55 | 0,9 | 0,2 | 0,65 |
| 22 | 0,228 | 0,9 | 0,55 | 0,9 | 0,2 | 0,95 |
| 23 | 0,373 | 0,9 | 0,55 | 0,9 | 0,8 | 0,65 |
| 24 | 0,544 | 0,9 | 0,55 | 0,9 | 0,8 | 0,95 |
| 25 | 0,127 | 0,9 | 0,85 | 0,5 | 0,2 | 0,65 |
| 26 | 0,184 | 0,9 | 0,85 | 0,5 | 0,2 | 0,95 |
| 27 | 0,457 | 0,9 | 0,85 | 0,5 | 0,8 | 0,65 |
| 28 | 0,67 | 0,9 | 0,85 | 0,5 | 0,8 | 0,95 |
| 29 | 0,137 | 0,9 | 0,85 | 0,9 | 0,2 | 0,65 |
| 30 | 0,201 | 0,9 | 0,85 | 0,9 | 0,2 | 0,95 |
| 31 | 0,471 | 0,9 | 0,85 | 0,9 | 0,8 | 0,65 |
| 32 | 0,689 | 0,9 | 0,85 | 0,9 | 0,8 | 0,95 |

По этим данным сформировано искомое уравнение:

$$\bar{y} = -0,52287 + 0,044568x_1 + 0,270679x_2 + 0,048634x_3 + \\ + 0,559089x_4 + 0,387762x_5.$$

Таблица 5.13

Результаты регрессионного анализа

| | <i>Коэффициенты</i> | <i>Стандартная ошибка</i> | <i>t-статистика</i> |
|-------------------|------------------------|-------------------------------|--------------------------------|
| Y-пересечение | -0,52287 | 0,083821 | -6,23787 |
| Переменная X1 | 0,044568 | 0,034409 | 1,295248 |
| Переменная X2 | 0,270679 | 0,068279 | 3,964334 |
| Переменная X3 | 0,048634 | 0,051209 | 0,949722 |
| Переменная X4 | 0,559089 | 0,034139 | 16,37673 |
| Переменная X5 | 0,387762 | 0,068279 | 5,679123 |
| <i>Наблюдение</i> | <i>Предсказанное Y</i> | <i>Остатки</i> | <i>Стандартные остатки</i> |
| 1 | 0,027558 | 0,060442 | 1,141425 |
| 2 | 0,143886 | -0,01 689 | -0,31889 |
| 3 | 0,363011 | -0,06001 | -1,13329 |
| 4 | 0,47934 | -0,03734 | -0,70515 |
| 5 | 0,047 011 | 0,051989 | 0,981781 |
| 6 | 0,16334 | -0,01734 | -0,32746 |
| 7 | 0,382465 | -0,06547 | -1,23628 |
| 8 | 0,498794 | -0,03879 | -0,7326 |
| 9 | 0,108761 | 0,007239 | 0,136697 |
| 10 | 0,2209 | -0,05809 | -1,09701 |
| 11 | 0,444215 | 0,000785 | 0,014822 |
| 12 | 0,560544 | 0,092456 | 1,745995 |
| 13 | 0,128215 | -0,00822 | -0,15514 |
| 14 | 0,244544 | -0,06954 | -1,31331 |
| 15 | 0,463 669 | -0,01167 | -0,22036 |
| 16 | 0,579 998 | 0,080002 | 1,510813 |
| 17 | 0,027 558 | 0,090442 | 1,707963 |
| 18 | 0,170 627 | 0,002373 | 0,044804 |
| 19 | 0,389 752 | -0,05375 | -1,01509 |
| 20 | 0,506 081 | -0,02008 | -0,37922 |
| 21 | 0,073 752 | 0,084248 | 1,590978 |

| <i>Наблюдение</i> | <i>Предсказанное Y</i> | <i>Остатки</i> | <i>Стандартные остатки</i> |
|-------------------|-------------------------------------|----------------|--------------------------------|
| 22 | 0,190081 | 0,037919 | 0,716081 |
| 23 | 0,409206 | -0,03621 | -0,68374 |
| 24 | 0,525535 | 0,018465 | 0,348706 |
| 25 | 0,135502 | -0,0085 | -0,16057 |
| 26 | 0,251831 | -0,06783 | -1,28096 |
| 27 | 0,470956 | -0,01396 | -0,26356 |
| 28 | 0,587285 | 0,082715 | 1,56204 |
| 29 | 0,154956 | -0,01796 | -0,33909 |
| 30 | 0,271285 | -0,07028 | -1,3273 |
| 31 | 0,49041 | -0,01941 | -0,36655 |
| 32 | 0,606739 | 0,082261 | 1,553472 |

Пример 5.10. На узел связи поступают заявки на передачу сообщений. Интервалы времени поступления заявок подчинены показательному закону с математическим ожиданием T_1 . На узле связи имеются два канала передачи данных. При поступлении очередной заявки в интервале времени $[0..T_2]$ вероятности того, что каналы A и B будут свободны, равны P_{1A} и P_{1B} соответственно.

При поступлении заявок после времени T_2 вероятности того, что каналы A и B будут свободны, равны P_{2A} и P_{2B} соответственно. Сообщение передаётся по любому свободному каналу. Если оба канала заняты, сообщение теряется.

Построить имитационную модель «Обработка запросов на узле связи» с целью определения абсолютного и относительного числа потерянных заявок из их общего количества, поступивших на узел связи за время T_{MOD} , $T_{\text{MOD}} > T_2$.

Получить уравнение регрессии, связывающее относительную долю обслуженных заявок (вероятность обработки) с интервалами их поступления и вероятностями P_{1A} , P_{1B} , P_{2A} , P_{2B} .

Решение.

Имитационная модель построена в соответствии с алгоритмом (см. рис. 3.23). Для построения уравнения регрессии введём обозначения:

y – отклик модели, вероятность ремонта СС с неисправными блоками C, D и A, B, D за время T_0 ;

x_1 – фактор, представляющий вероятность P1A;

x_2 – фактор, представляющий вероятность P1B;

x_3 – фактор, представляющий вероятность P2A;

x_4 – фактор, представляющий вероятность P2B;

x_5 – фактор, представляющий интервалы поступления заявок T .

Исходные данные и результаты эксперимента приведены в табл. 5.14. Для регрессионного анализа также использовалась функция «Регрессия» MS Excel. Получено искомое уравнение:

$$\bar{y} = 0,526135 - 0,23277x_1 + 0,289906x_2 - 0,48314x_3 - 0,25334x_4 + 1,48E-05x_5.$$

Таблица 5.14

Результаты эксперимента с моделью «Обработка запросов на узле связи»

| Номер отклика | Y | P1A | P1B | P2A | P2B | T |
|---------------|--------|-----|-----|-----|-----|---|
| 1 | 0,405 | 0,5 | 0,7 | 0,3 | 0,3 | 1 |
| 2 | 0,406 | 0,5 | 0,7 | 0,3 | 0,3 | 9 |
| 3 | 0,090 | 0,5 | 0,7 | 0,3 | 0,9 | 1 |
| 4 | 0,090 | 0,5 | 0,7 | 0,3 | 0,9 | 9 |
| 5 | 0,195 | 0,5 | 0,7 | 0,7 | 0,3 | 1 |
| 6 | 0,195 | 0,5 | 0,7 | 0,7 | 0,3 | 9 |
| 7 | 0,060 | 0,5 | 0,7 | 0,7 | 0,9 | 1 |
| 8 | 0,060 | 0,5 | 0,7 | 0,7 | 0,9 | 9 |
| 9 | 0,380 | 0,5 | 0,9 | 0,3 | 0,3 | 1 |
| 10 | 0,380 | 0,5 | 0,9 | 0,3 | 0,3 | 9 |
| 11 | 0,650 | 0,5 | 0,9 | 0,3 | 0,9 | 1 |
| 12 | 0,650 | 0,5 | 0,9 | 0,3 | 0,9 | 9 |
| 13 | 0,170 | 0,5 | 0,9 | 0,7 | 0,3 | 1 |
| 14 | 0,170 | 0,5 | 0,9 | 0,7 | 0,3 | 9 |
| 15 | 0,035 | 0,5 | 0,9 | 0,7 | 0,9 | 1 |
| 16 | 0,0348 | 0,5 | 0,9 | 0,7 | 0,9 | 9 |
| 17 | 0,375 | 0,9 | 0,7 | 0,3 | 0,3 | 1 |
| 18 | 0,376 | 0,9 | 0,7 | 0,3 | 0,3 | 9 |
| 19 | 0,060 | 0,9 | 0,7 | 0,3 | 0,9 | 1 |
| 20 | 0,060 | 0,9 | 0,7 | 0,3 | 0,9 | 9 |
| 21 | 0,165 | 0,9 | 0,7 | 0,7 | 0,3 | 1 |
| 22 | 0,165 | 0,9 | 0,7 | 0,7 | 0,3 | 9 |

| Номер отклика | Y | P1A | P1B | P2A | P2B | T |
|---------------|--------|-----|-----|-----|-----|---|
| 23 | 0,030 | 0,9 | 0,7 | 0,7 | 0,9 | 1 |
| 24 | 0,0301 | 0,9 | 0,7 | 0,7 | 0,9 | 9 |
| 25 | 0,370 | 0,9 | 0,9 | 0,3 | 0,3 | 1 |
| 26 | 0,370 | 0,9 | 0,9 | 0,3 | 0,3 | 9 |
| 27 | 0,055 | 0,9 | 0,9 | 0,3 | 0,9 | 1 |
| 28 | 0,055 | 0,9 | 0,9 | 0,3 | 0,9 | 9 |
| 29 | 0,160 | 0,9 | 0,9 | 0,7 | 0,3 | 1 |
| 30 | 0,160 | 0,9 | 0,9 | 0,7 | 0,3 | 9 |
| 31 | 0,025 | 0,9 | 0,9 | 0,7 | 0,9 | 1 |
| 32 | 0,025 | 0,9 | 0,9 | 0,7 | 0,9 | 9 |

Результаты регрессионного анализа, аналогичные рассмотренным результатам в примере 5.9 (табл. 5.13), приведены в табл. 5.15.

Таблица 5.15

Результаты регрессионного анализа

| | <i>Коэффициенты</i> | <i>Стандартная ошибка</i> | <i>t-статистика</i> |
|-------------------|------------------------|---------------------------|----------------------------|
| Y-пересечение | 0,526135 | 0,211881 | 2,483165 |
| Переменная X1 | -0,23277 | 0,112257 | -2,0735 |
| Переменная X2 | 0,289906 | 0,224514 | 1,291259 |
| Переменная X3 | -0,48314 | 0,112257 | -4,30387 |
| Переменная X4 | -0,25334 | 0,074838 | -3,38522 |
| Переменная X5 | 1,48E-05 | 0,005613 | 0,002645 |
| <i>Наблюдение</i> | <i>Предсказанное Y</i> | <i>Остатки</i> | <i>Стандартные остатки</i> |
| 1 | 0,391756 | 0,013244 | 0,113864 |
| 2 | 0,391875 | 0,014125 | 0,12144 |
| 3 | 0,23975 | -0,14975 | -1,28748 |
| 4 | 0,239869 | -0,14987 | -1,2885 |
| 5 | 0,1985 | -0,0035 | -0,03009 |
| 6 | 0,198619 | -0,00362 | -0,03111 |
| 7 | 0,046494 | 0,013506 | 0,116121 |
| 8 | 0,046613 | 0,013387 | 0,1151 |
| 9 | 0,449738 | -0,06974 | -0,59957 |

| <i>Наблюдение</i> | <i>Предсказанное Y</i> | <i>Остатки</i> | <i>Стандартные остатки</i> |
|-------------------|-------------------------------------|----------------|--------------------------------|
| 10 | 0,449856 | -0,06986 | -0,60059 |
| 11 | 0,297731 | 0,35226 | 3,02865 |
| 12 | 0,297850 | 0,35215 | 3,027629 |
| 13 | 0,256481 | -0,08648 | -0,74353 |
| 14 | 0,256600 | -0,0866 | -0,74455 |
| 15 | 0,104475 | -0,06948 | -0,59732 |
| 16 | 0,104594 | -0,06979 | -0,60006 |
| 17 | 0,298650 | 0,07635 | 0,656423 |
| 18 | 0,298769 | 0,077231 | 0,664 |
| 19 | 0,146644 | -0,08664 | -0,74492 |
| 20 | 0,146763 | -0,08676 | -0,74595 |
| 21 | 0,105394 | 0,059606 | 0,512468 |
| 22 | 0,105513 | 0,059487 | 0,511447 |
| 23 | -0,04661 | 0,076612 | 0,65868 |
| 24 | -0,04649 | 0,076594 | 0,658519 |
| 25 | 0,356631 | 0,013369 | 0,114939 |
| 26 | 0,356750 | 0,01325 | 0,113918 |
| 27 | 0,204625 | -0,14963 | -1,28641 |
| 28 | 0,204744 | -0,14974 | -1,28743 |
| 29 | 0,163375 | -0,00338 | -0,02902 |
| 30 | 0,163494 | -0,00349 | -0,03004 |
| 31 | 0,011369 | 0,013631 | 0,117195 |
| 32 | 0,011488 | 0,013512 | 0,116174 |

Кроме оценок коэффициентов регрессии (*Коэффициенты*) функция «Регрессия» выдаёт и другие результаты регрессионного анализа: вычисленные значения откликов y (*Предсказанное Y*);

разность между вычисленными значениями откликов и измеренными в эксперименте в каждом наблюдении $y - \bar{y}$ (*Остатки*);

среднеквадратические ошибки в определении коэффициентов регрессии S_x (*Стандартная ошибка*) и *t*-статистика;

среднеквадратические ошибки в определении коэффициентов откликов при определённых значениях факторов S_y (*Стандартные остатки*) и др.

После вычисления коэффициентов регрессии, представленной в виде линейного полинома, оценивается их значимость для определе-

ния степени влияния различных факторов на выходной параметр. Основной оценки значимости коэффициентов является сопоставление *абсолютного значения*, например коэффициента b_i , и дисперсии ошибки его определения $\sigma_{b_i}^2$. В этом случае с помощью t -критерия проверяется гипотеза о незначимости рассматриваемого коэффициента, т.е. гипотеза о том, что $b_i = 0$ (проверка нуль-гипотезы). При подсчёте экспериментального значения t -статистики в числитель ставится абсолютное значение рассматриваемого коэффициента, а в знаменатель – дисперсия ошибки его определения:

$$t_i = \frac{b_i}{\sqrt{\sigma_{b_i}^2}}.$$

Коэффициент b_i признаётся незначимым, если t_i для числа степеней свободы N меньше $t_{кр}$, т.е. $|t_i| < t_{кр}$, найденного из таблицы распределения Стьюдента (см. табл. 4.4) для заданной достоверности α .

Пример 5.11. По результатам моделирования, полученным в примере 5.10, определить при достоверности $\alpha = 0,95$ незначимые коэффициенты регрессионной зависимости

$$\bar{y} = 0,526135 - 0,23277x_1 + 0,289906x_2 - 0,48314x_3 - 0,25334x_4 + 1,48E-05x_5.$$

Решение.

В примере 5.10 $N = 32$.

Рассчитанные с помощью функции «Регрессия» t -статистики приведены в крайнем правом столбце табл. 5.15. В таблице коэффициентов Стьюдента (см. табл. 4.4) находим $t_{кр} = 2$. t -статистики сравниваем с найденным значением $t_{кр}$ и находим, что незначимыми являются коэффициенты регрессионной зависимости при факторах x_2 и x_3 .

Статистическая незначимость коэффициента b_i может быть вызвана следующими обстоятельствами:

- 1) нулевой уровень данного фактора (или произведения факторов) близок к точке частного экстремума;
- 2) интервал варьирования уровней данного фактора выбран слишком малым;
- 3) данный фактор (взаимодействие факторов) не оказывает влияния на значение выходного отклика (реакции).

Так как применение ортогональных планов даёт возможность оценивать значения всех коэффициентов независимо друг от друга, то, если один или несколько коэффициентов окажутся незначимыми, они могут быть отброшены без пересчёта остальных. Отбросив незначимые коэффициенты, получим уточнённую имитационную модель в виде полинома («вторичную модель»), представляющую зависимость выходного параметра от факторов.

В условиях примера 5.11 получим:

$$\bar{y} = 0,526135 - 0,23277x_1 - 0,48314x_3 - 0,25334x_4.$$

Проверяется адекватность «вторичной модели». Это значит, что в некоторой подобласти, в которую входят и координаты выполненных наблюдений, предсказанное с помощью «вторичной модели» значение отклика не должно отличаться от значения эксперимента более чем на некоторую заранее заданную величину. Для проверки адекватности достаточно оценить отклонение предсказанного значения отклика от результатов эксперимента в соответствующей точке факторного пространства.

При анализе **адекватности** уравнения регрессии («вторичной модели») исследуемому процессу, возможны следующие варианты:

1. Построенная модель на основе F-критерия Фишера в **целом адекватна** и все **коэффициенты** регрессии **значимы**. Такая модель может быть использована для принятия решений и осуществления прогнозов.

2. Модель по F-критерию Фишера **адекватна**, но **часть коэффициентов не значима**. Модель пригодна для принятия некоторых решений, но не для прогнозов.

3. Модель по F-критерию адекватна, но **все коэффициенты регрессии не значимы**. Модель полностью считается неадекватной. На её основе не принимаются решения и не осуществляются прогнозы.

Вопросы для самоконтроля

5.1. Что понимают под характеристиками случайных величин и процессов?

5.2. Что такое несмещённая оценка характеристики случайной величины? Состоятельная? Эффективная?

5.3. Что характеризует гистограмма? Правило построения гистограммы.

- 5.4. В чем состоит сущность дисперсионного анализа?
- 5.5. Что такое ошибки первого рода и второго рода при оценке гипотез?
- 5.6. Что такое F -распределение и почему оно является мерой сравнения дисперсий случайных величин?
- 5.7. Для чего используется критерий Вилкоксона?
- 5.8. В чем состоит методика выявления несущественных (незначимых) факторов?
- 5.9. Назначение корреляционного анализа.
- 5.10. Назначение регрессионного анализа.
- 5.11. Представьте графически виды корреляции между двумя переменными.
- 5.12. Составьте систему уравнений для определения коэффициентов регрессии модели вида:

$$\bar{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_2^2.$$

- 5.13. Для линейной и нелинейной моделей, полученных в п. 5.12, вычислить и сравнить ошибки ε и ε_0 .
- 5.14. С объектом проведено $N = 6$ экспериментов. Данные экспериментов приведены в таблице.

| Измерения | Эксперименты | | | | | | Σ |
|---------------------|--------------|---|---|---|---|---|----------|
| | 1 | 2 | 3 | 4 | 5 | 6 | |
| x_i | 2 | 1 | 4 | 3 | 6 | 5 | |
| y_i | 3 | 2 | 5 | 4 | 8 | 7 | |
| $x_i y_i$ | | | | | | | |
| x_i^2 | | | | | | | |
| \bar{y}_i | | | | | | | |
| $ \bar{y}_i - y_i $ | | | | | | | |

Построить линейную математическую модель функционирования объекта вида: $y = b_0 + b_1x + \varepsilon$. Расчеты провести в таблице. Проверьте адекватность модели при абсолютной точности 0,2.

- 5.15. Как проверяется значимость коэффициентов регрессии?

Глава 6. Современные теории имитационного моделирования

6.1. Распределённое имитационное моделирование

6.1.1. Понятие о распределённой информационной системе

Что такое распределённая система? В литературе упоминаются разные определения понятия распределённой системы (РС), но все они могут быть сведены к следующим определениям:

1. Распределённая система – это набор независимых компьютеров, представляющий их пользователям как единая система.

2. Распределённая информационная система (РИС) – это совокупность взаимодействующих друг с другом программных компонент. Каждая из таких компонент может рассматриваться как программный модуль (приложение), исполняемый в рамках отдельного процесса.

Какие основные задачи распределённых систем? Основная задача – облегчение доступа к удалённым ресурсам и контроль совместного использования этих ресурсов (компьютеров, файлов, данных в базах данных). Для решения этой основной задачи, РС должна удовлетворять следующим требованиям:

- прозрачность;
- открытость;
- гибкость;
- масштабируемость (расширяемость).

Прозрачность. Важная задача РС состоит в том, чтобы скрыть тот факт, что процессы и ресурсы физически распределены по разным компьютерам. РС, которые представляются пользователям и приложениям в виде единой системы называются прозрачными (transparent).

Открытость. Открытая РС (open distributed system) – это система, предлагающая службы, вызов которых требует стандартные синтаксис и семантику. В РС службы (и компоненты) определяются через интерфейсы, которые часто описываются при помощи языка определения интерфейсов (Interface Definition Language, IDL). Описание интерфейса точно отражает синтаксис служб: имена доступных функций, типы параметров, возвращаемых значений и т.п.

Гибкость – легкость конфигурирования системы, состоящей из разных компонентов. То есть РС должна быть расширяемой программно и аппаратно. В построении гибких открытых РС решающим фактором оказывается организация таких систем в виде наборов относительно небольших, легко заменяемых или адаптируемых компонентов. Это предполагает необходимость определения не только интерфейсов верхнего уровня, с которыми работают пользователи и приложения, но и интерфейсов между компонентами системы.

Масштабируемость (возможность расширения). Масштабируемость системы может измеряться по трём разным показателям:

размер – лёгкость подключения новых пользователей и ресурсов;

площадь – пользователи и ресурсы могут быть разнесены в пространстве;

управление – система проста в управлении при работе во многих органах управления различных уровней.

Эта характеристика РС может снижать производительность. Вопросы практической реализации масштабируемости должны рассматриваться вместе с другими требованиями, такими как безопасность и производительность.

Основные технологии масштабирования: сокрытие времени ожидания связи, распределение и репликация.

1. *Сокращение времени ожидания связи* применяется в случае географического масштабирования. Основная идея – постараться избежать ожидания ответа на запрос от удалённого сервера.

2. *Распределение*. Предполагает разбиение компонентов на более мелкие части с последующим их распределением в системе. Пример – система доменных имен Интернета (DNS).

3. *Репликация (дублирование)*. Применение репликации повышает доступность ресурсов, а также помогает выравнять нагрузку компонентов, что ведет к улучшению производительности. Кроме того, в географически распределённых системах репликация (в виде близко лежащей копии) помогает уменьшить проблему ожидания связи. Особую форму репликации представляет собой кеширование (caching). Как и при репликации, результатом кеширования является создание копии ресурса, обычно в непосредственной близости от клиента. В отличие от репликации кеширование – это действие, предпринимаемое со стороны клиента, а не сервера. Поскольку при репликации и кешировании создаются копии ресурса, модификация одной копии

делает её отличной от остальных. Проблема строгой непротиворечивости ресурса состоит в немедленном распространении изменений во все копии. При одновременном изменении двух или более копий необходимо строго соблюдать порядок внесения изменений во все копии. Практически обеспечить это бывает довольно сложно, а иногда – невозможно. Это значит, что репликация может включать и отдельные не масштабируемые решения.

Существуют разные способы организации процессоров в единую РС (вариантов соединения и организации взаимного обмена). Для простоты, все компьютеры в РС мы можем разделить на две группы. Системы, в которых компьютеры используют совместно память, называются *мультипроцессорами*, а работающие каждый со своей памятью – *мультикомпьютерами*. Основная разница между ними – первые имеют единое адресное пространство, которое используется всеми процессорами. В мультикомпьютерах каждая машина использует свою память (пример, обычная сеть компьютеров).

Мультикомпьютерные системы разделяют также на *гомогенные* (homogeneous) и *гетерогенные* (heterogeneous). В первых используется одна компьютерная сеть, использующая единую технологию, однотипные процессоры. *Гетерогенные* системы содержат независимые компьютеры, соединенные разными сетями.

Хотя аппаратные решения важны для РС, наибольшее влияние на них оказывают именно программные решения. Эти решения влияют, в первую очередь, на удобство работы пользователя в РС. С одной стороны, РС работают как менеджеры ресурсов, помогая пользователям совместно использовать такие ресурсы как память, процессоры, периферийное оборудование, сеть и данные. В этом они подобны ОС. С другой – РС скрывают сложность и гетерогенность аппаратуры, предоставляя виртуальную машину для выполнения приложений. Эти функции обычно выполняют ОС.

Для распределённых компьютеров ОС можно разделить на две категории: сильно связанные и слабо связанные.

Сильно связанные ОС обычно называются распределёнными ОС. Они используются для управления мультипроцессорными и гомогенными мультикомпьютерными системами. Основная их цель – скрыть тонкости управления аппаратным обеспечением.

Слабо связанные ОС называются сетевыми ОС (Network Operation System, NOS). Они используются для управления гетерогенными

мультикомпьютерными системами. Для создания РС служб сетевой ОС недостаточно. Необходимо добавить к ним дополнительные компоненты для организации поддержки прозрачности распределения. Эти компоненты образуют средства промежуточного уровня. Таким образом, основную роль в построении РС играют программные средства (службы) промежуточного уровня, между ОС и распределёнными приложениями.

6.1.2. Необходимость перехода к параллельному и распределённому имитационному моделированию

В настоящее время наметилась тенденция перехода от последовательного имитационного моделирования к *параллельному и распределённому* имитационному моделированию.

Параллельное имитационное моделирование предполагает, что компьютерный эксперимент с имитационной моделью выполняется на нескольких процессорах многопроцессорной ЭВМ. Это может быть многопроцессорная ЭВМ с MPP архитектурой или многопроцессорная ЭВМ с общей памятью.

Выполнение вычислительного эксперимента с имитационной моделью на нескольких компьютерах, объединённых в сеть (локальную или глобальную), можно рассматривать как распределённое имитационное моделирование.

Причины перехода к параллельному и распределённому имитационному моделированию обусловлены следующим.

1. Исследователь, который использует ресурсы N процессоров многопроцессорной ЭВМ (или N компьютеров, объединённых в сеть), в конечном счёте, пытается добиться того, чтобы время выполнения имитационного эксперимента сократилось в N раз.

2. Задачи, которые стоят перед имитационным моделированием, становятся всё более сложными и требуют не только больших временных затрат, но и больших ресурсов памяти. Исследователь, использующий ресурсы нескольких процессоров или компьютеров предполагает, что локальная память этих процессоров или компьютеров также будет использована для решения его задачи.

3. Другим важным стимулом использования распределённого моделирования является необходимость объединения нескольких имитационных систем в одну распределённую среду имитационного моделирования. Примером может служить объединение нескольких трена-

жёров, имитирующих, управление самоходным орудием, пусковой установкой, танком, самолётом и других имитаторов (военные приложения). Их объединяют с целью создать распределённую виртуальную среду для обучения человека действиям по возможным сценариям в нестандартных ситуациях. Другим примером может служить необходимость в объединении нескольких изучаемых подсистем (их моделей) в одну с той целью, чтобы можно было проанализировать их взаимодействие. Так, например, моделирование транспортной инфраструктуры города следует рассматривать неразрывно с подсистемой его электроснабжения и с другими коммуникационными инфраструктурами. В данном случае речь идёт о том, что гораздо более экономичным является подход, когда имитационные модели (в военных приложениях или при моделировании инфраструктур) объединяются с помощью дополнительных программных средств (HLA, High Level Architecture). В противном случае возникает необходимость в разработке новой имитационной модели.

4. Ещё одной причиной для перехода к распределённому моделированию является возможность работы нескольких исследователей над одним и тем же имитационным проектом через Интернет. Проект может выполняться на суперкомпьютере или на вычислительной системе с кластерной архитектурой и, таким образом, исследователи используют вычислительные ресурсы суперкомпьютера, наблюдают за ходом имитационного эксперимента и просматривают результаты, находясь в разных городах или странах.

5. Кроме того, используя распределённую систему моделирования с удалённым доступом через Интернет, можно сдать в «аренду» неиспользуемые в данный момент времени вычислительные ресурсы.

6. Ещё такое преимущество использования распределённых систем, как надёжность: под надёжностью будем подразумевать продолжение выполнения имитационного эксперимента, несмотря на то, что какой либо из процессоров (компьютеров) выходит из строя. В этом случае его работа распределяется на другие процессоры.

Распределение является одной из основных тенденций компьютерных технологий настоящего времени (другой тенденцией является интеграция). Это проявляется и в базах данных, и в искусственном интеллекте (например, мультиагентные системы), и в операционных системах. Появление распределённых систем имитационного моделирования – ещё одно проявление этой тенденции.

Существует два термина, которые часто путают: «распределенные вычислительные системы» и «параллельные вычислительные системы». Их необходимо разделить, чтобы не возникало неоднозначности. Термин *параллельные системы*, как правило, применяется к суперкомпьютерам для того, чтобы подчеркнуть использование многопроцессорной архитектуры. Основными классами архитектур современных параллельных компьютеров являются:

SMP – симметричные мультипроцессорные системы.

MPP – массово-параллельные системы.

NUMA – системы с неоднородным доступом к памяти.

PVP – параллельные векторные системы.

Кластеры – используются в качестве дешёвого варианта MPP. В качестве узлов могут выступать серверы, рабочие станции и даже персональные компьютеры. Для связи узлов используется одна из стандартных сетевых технологий. Кластеризация может осуществляться на разных уровнях компьютерной системы, включая аппаратное обеспечение, операционные системы, программы-утилиты, системы управления и приложения.

Термин *распределённая система* обозначает набор независимых компьютеров.

Пользователи представляют их единой объединённой системой. В этом определении подчеркиваются два момента.

Во-первых, все машины автономны.

Во-вторых, распределённая система скрывает сложность и гетерогенную природу аппаратного обеспечения, на базе которого она построена. Организация распределённых систем включает в себя дополнительный уровень программного обеспечения (ПО), находящийся между верхним и нижним уровнем. Верхний уровень представляет собой пользователей и их приложения. Нижний уровень – это операционные системы (ОС).

В некоторых работах под распределённой системой подразумевают *взаимосвязанный набор автономных компьютеров, процессов или процессоров*. Компьютеры, процессы или процессоры упоминаются как узлы распределённой системы. Будучи определёнными как «автономные», узлы должны быть, по крайней мере, оборудованы своим собственным блоком управления. Таким образом, параллельный компьютер с одним потоком управления и несколькими потоками данных (SIMD) не подпадает под определение распределённой системы.

Мы определили узлы «взаимосвязанными». Из этого следует, что узлы должны иметь возможность обмениваться информацией.

Так как процессы могут играть роль узлов системы, определение включает программные системы, построенные как набор взаимодействующих процессов, даже если они выполняются на одной аппаратной платформе.

Таким образом, *распределённой системой* можно считать *локальные и глобальные сети, мультипроцессорные вычислительные системы и взаимодействующие процессы*, которые выполняются на одном компьютере.

Каждая из этих типов распределённых систем обладает своими отличительными характеристиками, которые следует учитывать при разработке программного обеспечения. Так локальные и глобальные сети отличаются параметрами надёжности, временем коммуникации, однородностью вычислительных узлов:

Параметры надёжности. Под параметром надёжности понимают степень надёжности передачи данных в компьютерной сети. Если сеть является глобальной, то вероятность потери переданного сообщения достаточно велика, и алгоритмы должны предусматривать действия, нейтрализующие последствия таких сбоев. Локальные сети более надёжны, и алгоритмы для них могут быть разработаны в предположении абсолютной надёжности коммуникаций.

Время, затрачиваемое на передачу сообщения. Время, затрачиваемое на передачу сообщения в глобальных сетях на порядки больше, чем время передачи в локальных сетях. Временем обработки сообщения в глобальных сетях всегда можно пренебречь при сравнении его со временем, необходимым для передачи.

Однородность. Вычислительные узлы локальных сетей могут отличаться своей производительностью, тем не менее, для обеспечения функционирования сети можно использовать единое программное обеспечение и одни и те же протоколы. В глобальных сетях используется множество различных протоколов и программное обеспечение, соответствующее разным стандартам. Поэтому следует обращать внимание на преобразование информации при переходе от одного протокола к другому и на совместимость стандартов.

Безопасность. При использовании глобальной сети следует уделять большое внимание разработке программных средств защиты вычислительных узлов от внешних атак.

6.1.3. Подходы к сокращению времени имитационного эксперимента

Исторически сложилось так, что термин «распределённая система имитации» относился к системам, выполняемым на вычислительной системе, узлы которой географически могли находиться на весьма отдалённом расстоянии друг от друга. В этих системах время, которое затрачивалось на передачу сообщений между узлами, было сравнительно велико, а производительность систем – низкая. Термин «параллельная система имитации» применялся к системе имитации, функционирующей на высокопроизводительных вычислительных системах. Вычислительные узлы этих систем могут находиться на незначительном расстоянии (например, в одном здании или одной комнате). В этих системах время, затрачиваемое на коммуникацию между узлами, незначительно по сравнению со временем, затрачиваемым на вычисления. Однако в связи с появлением кластеров рабочих станций, корпоративных и GRID-систем, границы между параллельными и распределёнными системами имитации стираются. В дальнейшем все системы имитации, функционирующие на вычислительной системе с несколькими узлами, будем называть *распределёнными*.

Как происходит распределение вычислений между узлами?

За счёт чего можно сократить время проведения имитационного эксперимента?

1. Выполнять параллельно с ходом имитационного эксперимента специализированные функции, а именно: генерацию псевдослучайного числа, управление списком будущих событий, сбор статистических данных (специализированные функции). Следует отметить, что выигрыш во времени в этом случае является небольшим.

2. При декомпозиции иерархической модели происходит декомпозиция события на подсобытия. Эти подсобытия следует выполнить параллельно (иерархическая декомпозиция). Выигрыш во времени при декомпозиции модели зависит от самой модели. Для некоторых моделей (например, для моделей СМО) декомпозиция выполняется легко. Каждая очередь в модели СМО может быть выполнена на отдельном процессоре (компьютере). Для других имитационных моделей декомпозиция не является тривиальной.

3. Несколько имитационных прогонов последовательной имитационной модели следует выполнять на нескольких процессорах (распараллеливание репликаций). В этом случае вычислительные узлы

должны обладать соответствующими вычислительными ресурсами для выполнения имитационного прогона всей модели. В настоящее время мощность компьютеров достаточно велика, но возникает ещё одна проблема: если репликация может быть выполнена только при некотором условии, которое зависит от предыдущей репликации, то они не могут быть выполнены параллельно.

4. Распределить выполнение имитационной модели на несколько процессоров (компьютеров). Это распределение можно назвать ещё распределением на уровне объектов или процессов, в отличие от описанного выше распределения на уровне моделей. («Распределение» на уровне моделей – то есть модель является неделимой и должна полностью выполняться на той машине, на которой она начала выполняться, а разные модели могут выполняться на разных машинах). Это более простой, но неоптимальный вариант. «Распределение» на уровне объектов (процессов) модели означает, что часть объектов одной и той же модели выполняется на одном сервере, а часть на других. Сложность реализации такой системы заключается в том, что если в первом случае все объекты выполняются в контексте одного процесса и коммуникации между ними можно легко реализовать, во втором случае объекты должны взаимодействовать через сеть, что усложняет синхронизацию объектов, отладку моделей и саму реализацию этих объектов.

6.1.4. Вариант построения распределённой имитационной модели

Целью использования распределённого имитационного моделирования **для военных приложений** является интеграция отдельных моделей в единое модельное окружение. Примером могут быть **тренажёры для обучения и имитации сценариев военных действий**.

При распределённом моделировании в отличие от последовательного моделирования первичной единицей является не объект, а так называемый **логический процесс (ЛП)**.

Логический процесс, это последовательная модель, поддерживающая дискретно-событийное моделирование. Она рассмотрена нами ранее в главе 3. Это значит, что каждый логический процесс имеет собственный набор объектов, данные о состоянии объектов, собственную управляющую программу, собственный локальный список событий и собственные часы локального времени (модельного времени – МВ). Логические процессы взаимодействуют исключительно с помощью передачи сообщений.

Распределённую модель, состоящую из N логических процессов, можно определить следующим образом:

$$DM = U_{k=1}^N \cdot l_{p^k},$$

$$l_{p^k} = \{sm^k, T^k, S^k\},$$

$$S^k = \{(Q_1^k, t_1^k), (Q_2^k, t_2^k), \dots, (Q_m^k, t_m^k) \mid t_1^k \leq t_2^k \leq \dots \leq t_m^k\},$$

$$T_k = t_1^k,$$

где DM – распределенная модель, l_p – логический процесс, sm – под-модель ЛП, T – значение локальных часов МВ логического процесса, S – локальный список событий ЛП, Q – событие в списке событий, t – МВ наступления события Q (временная метка события Q).

Глобальных (общих для всей распределённой модели) часов МВ и глобального списка событий явно не существует, так как наличие общей управляющей программы, работающей с этими глобальными структурами, было бы «бутылочным горлышком» для параллельного исполнения.

Текущее модельное время всей модели в каждый момент равно:

$$TIME = \min_{k=1}^N T^k.$$

Каждый ЛП выполняется в собственном модельном времени как автономная последовательная модель. Для этого он снабжен экземпляром управляющей программы. Схема распределённой модели представлена на рис. 6.1.

Логический процесс общается с другими процессами, передавая им сообщения. Сообщения могут быть разных видов. Рассмотрим простейший вид сообщения — асинхронное сообщение, никак не влияющее на выполнение логического процесса — отправителя.

Сообщение M_{ij} , передаваемое от ЛП lp^i логическому процессу lp^j , в простейшем случае имеет следующую структуру:

$$M_{ij} = (Q^j, T^i).$$

Логический процесс lp^j , получив сообщение, которое имеет форму события, вставляет это событие в упорядоченный список своих локальных событий в соответствии со значением времени T^i . Управляющая программа начинает выполнять модифицированный список локальных событий.

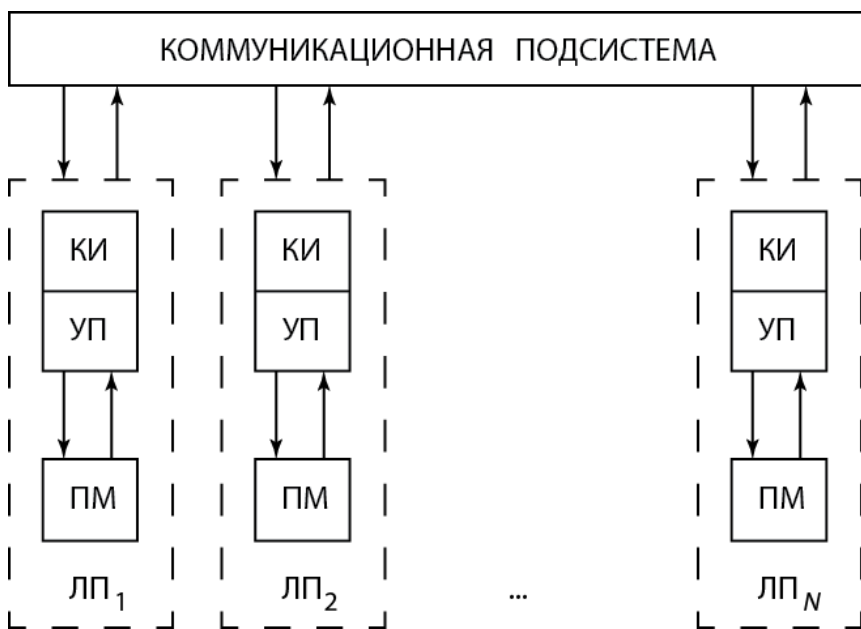


Рис. 6.1. Схема распределённой ИМ: ЛП – логический процесс, ПМ – подмодель, УП – управляющая программа, КИ – коммуникационный интерфейс

Таким образом, полученное сообщение может изменить логику выполнения ЛП – получателя.

Передача сообщений может осуществляться логическими процессами непосредственно с помощью средств операционной системы. Но более перспективна схема, приведенная на рис. 6.1. Она включает коммуникационную подсистему в качестве отдельного компонента. Логические процессы взаимодействуют с коммуникационной подсистемой с помощью определённого интерфейса.

Критерий правильной работы распределённой имитационной модели можно сформулировать так: результат исполнения распределённой модели должен совпадать с результатом исполнения этой модели на однопроцессорном компьютере.

При распределённой реализации модельное время в разных логических процессах «движется» с разными скоростями. Процессорное время не связано с модельным временем явно. Точнее, процессорное время является неубывающей функцией от модельного времени. Но для разных логических процессов эти функции разные. Поэтому через

некоторый интервал естественного времени от начала выполнения модели модельные времена разных логических процессов могут оказаться разными.

Парадокс времени в распределённой модели возникает при получении логическим процессом j от логического процесса i сообщения следующего типа:

$$M_{ij} = (Q^j, T^i) | T^i < T^j.$$

Для борьбы с парадоксами времени в системе распределённого моделирования должна быть реализована программа синхронизации. Эта программа входит в состав коммуникационной подсистемы. Программа синхронизации реализует некоторые абстрактные алгоритмы, которые получили название «алгоритмы синхронизации МВ».

Целью механизма синхронизации МВ является выполнение каждым ЛП событий в порядке не убывания их временных меток. Это требование известно как локальное ограничение причинной связи, так как оно обеспечивает имитацию естественного порядка «от причины – к следствию».

Если не принимать во внимание порядок событий, имеющих одинаковые временные метки (одновременные события), то выполнение этого требования гарантирует, что выполнение программы модели на параллельном компьютере приведет точно к таким же результатам, что и выполнение модели на последовательном компьютере, где все события выполняются строго в порядке возрастания временных меток. Это также гарантирует, что повторное выполнение модели приведет к тем же результатам, если все входные данные совпадают и повторные вычисления, связанные с каждым событием, дают одинаковые результаты.

Таким образом, управление временем в распределённом моделировании не является тривиальным. Оно должно обеспечивать выполнение событий в правильном естественном хронологическом порядке. Более того, на алгоритмы синхронизации возлагается обязанность корректно выполнять повторные имитационные прогоны. При повторном моделировании пользователь должен быть уверен, что он получит те же результаты, что и в первый раз, если входные данные останутся неизменными. Это требование относится только к дискретно-событийному моделированию, но не к тренажёрам.

6.2. Агентное моделирование

6.2.1. Мультиагентные системы и агенты

Одним из самых распространённых способов исследования динамических и сложным образом организованных систем является имитационное моделирование. Данная форма моделирования берёт своё начало в 1960-х годах, и к настоящему моменту она становится всё более актуальной и востребованной. Однако, подобно другим формам, имитационное моделирование имеет ряд трудностей, которые являются следствием попытки представить процессы и явления в реальном мире. В целом, наблюдается тенденция к повышению точности и адекватности создаваемых моделей. Ответом на это требование служит возникновение агентного имитационного моделирования.

Агентное моделирование (АМ) – это новый подход к моделированию систем, содержащих автономных и взаимодействующих агентов.

Агентное моделирование находит всё большее применение. Это объясняется тем, что мы живем в постоянно усложняющемся мире. Прежде всего, системы, которые мы должны анализировать и моделировать, становятся всё более сложными из-за сложных взаимозависимостей их компонентов. Таким образом, традиционные подходы к моделированию (системная динамика, дискретно-событийное моделирование) не могут удовлетворить исследователей.

В настоящее время мультиагентные системы получили широкое применение в таких областях как системы телекоммуникации, поисковые системы в Internet, логистика, компьютерные игры, системы автоматизированного проектирования, системы управления и контроля сложными процессами в медицине и промышленности, программы для электронной коммерции, системы защиты информации. Это и моделирование поведения агентов на фондовых рынках, и моделирование поставок, и предсказание распространения эпидемий, и угрозы биологических войн и т.д.

Выясним более подробно, что такое агенты, их классификацию, рассмотрим архитектуру мультиагентных систем, сферы применения мультиагентных систем, различие между дискретно-событийным и агентным моделированием.

Мультиагентные системы (кросс-платформенные распределённые интеллектуальные системы) представляют собой совокупность интеллектуальных агентов.

Агенты – это автономные объекты, которые могут самостоятельно реагировать на внешние события и выбирать соответствующие действия. В настоящее время в рамках мультиагентных технологий и мультиагентных систем (МАС) разработаны различные типы агентов, которые характеризуются конкретной моделью поведения и свойствами, а также, семейства архитектур и библиотек компонентов, для которых свойственны распределённость и автономность.

Поскольку агенты применяются в самых различных областях, то понятие агента для каждого автора имеет свою смысловую нагрузку: так на производстве в качестве агента может выступать робот, а в системах телекоммуникации понятие «агент» связано с программой, в системах военной связи это могут быть различные объекты.

В зависимости от среды обитания агента наделяют конкретным набором свойств. Существует большое количество типов агентов: автономные агенты, мобильные агенты, персональные ассистенты, интеллектуальные агенты, социальные агенты и т.д.

По способу поведения агенты могут делиться на интеллектуальные (рассуждающие, коммуникативные, ресурсные) и реактивные (не обладающие представлением о внешнем мире).

Существует множество определений агента, данных разными авторами – исследователями МАС. Например: **агент – это аппаратная или программная сущность, способная действовать в интересах достижения целей, поставленных пользователем.**

Для реализации некоторого поведения агент должен иметь специальные устройства (в том числе и программные): *рецепторы* – устройства, непосредственно воспринимающие воздействие внешней среды и *эффекторы* – исполнительные органы, воздействующие на среду, а также процессор – блок переработки информации и память.

Под памятью здесь понимается способность агента хранить информацию о своём состоянии и состоянии среды.

Перечислим свойства агента:

адаптивность (способностью обучаться);

автономность (агент работает как самостоятельная программа, которая ставит себе цели и предпринимает действия для их достижения);

взаимодействие с другими агентами, причём агент может играть разные роли при взаимодействии с одним и тем же агентом;

способность к рассуждениям (агенты могут обладать частичными знаниями или механизмами вывода, например, знаниями, как приводить данные из различных источников к одному виду);

коммуникативность (агенты могут общаться с другими агентами); мобильность (передача кода агента с одного сервера на другой).

Итак, агент должен взаимодействовать *со средой* или *другими агентами*. Агенты *децентрализованы*, а *глобальное состояние* моделируемой системы можно вывести из взаимодействия агентов, зная индивидуальную логику поведения каждого из них.

Программной реализацией агента является объект (в смысле объектно-ориентированного программирования) некоторого класса. Однако существуют два основных отличия агента от традиционного объекта. Во-первых, обмен сообщениями между объектами (рис. 6.2а) сводится в основном к вызову методов, в то время как для агентов это наиболее важный процесс взаимодействия (рис. 6.2б), поэтому они способны к ведению переговоров на основе теории речевых актов для выработки общего решения. Во-вторых, агенты должны обладать психическими характеристиками, сближающими их с живыми существами.

Итак, про агента можно сказать следующее.

Агент идентифицируем, его можно определить как индивидуум, состоящий из частей с определённым набором характеристик и правил, управляющих его поведением и отвечающих за принятие решений.

Агент помещён в определённую среду, живёт в ней и взаимодействует с другими агентами (рис. 6.2б). Агенты имеют протоколы для взаимодействия с другими агентами типа коммуникационного протокола, и способны реагировать на среду. Агенты способны распознавать и различать черты других агентов.

Агент автономный и самонаводящийся. Агент может функционировать в своей среде независимо от совместных действий с другими агентами, по крайней мере, в небольшом числе ситуаций.

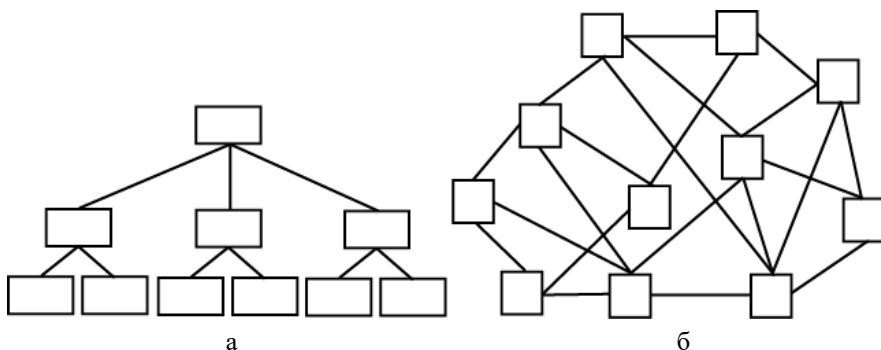


Рис. 6.2. Традиционное и мультиагентное построение модели

Агент гибкий и может обучаться и адаптировать своё поведение по времени, основываясь на опыте. Агент может иметь метаправила, правила, изменяющие другие правила поведения.

Агенты разнообразны, разнородны и динамичны. Агенты активны по отношению к своим атрибутам и правилам поведения. Правила поведения варьируются: они зависят от того, какое количество информации агент учитывает в своих рассуждениях, от внутренней модели внешнего мира, включающей других агентов, от размера памяти для хранения прошлых событий. Все перечисленные факторы агент использует в своих решениях.

Агенты различаются своими атрибутами и накопленными ресурсами. Их природа делает агентное моделирование интересным.

6.2.2. Подходы к построению мультиагентной модели

Предполагается, что модель, создаваемая пользователем, будет инвариантна относительно предметной области. Для этого пользователю предстоит каким-то образом задать агентов, определить их свойства, структуру, поведение и семантику, объединить этих агентов, например, в сеть, определить топологию данной сети, наложить на неё условия моделирования и поместить эту «заготовку» в определённую среду моделирования.

При этом исследователю должны быть даны несколько способов представления агентов. Представление может быть определено внутренним состоянием агента, наличием дополнительных данных (например, собираемой статистикой), поведением, механизмом принятия решений, наличием интеллектуальных средств, механизмов адаптации, средств обучения и пополнения знаний в ходе вывода и т.д.

Для исследования агентного подхода потребуется:

определить сферу применимости агентного подхода;

выявить преимущества и недостатки агентного моделирования;

сравнить агентный подход с традиционными формами имитационного моделирования;

построить математическую модель для разрабатываемой системы;

рассмотреть известные аналоги и результаты их применения.

рассмотреть класс существующих задач, решаемых с помощью агентного подхода.

Напомним, что имитационное моделирование использует и другие подходы (объектные, процессные, дискретно-событийные и т.д.), ко-

торые также нужно учитывать. Более того, для некоторых задач имитационное моделирование неэффективно (вместо него используют, например, аналитическое моделирование). Некоторые задачи, в свою очередь, потребуют комбинированных методов и т.д. Значит, также предстоит выявить особенности и ограничения агентного подхода. Для ответа на все эти перечисленные вопросы обратимся к существующим задачам и способам их решения.

Модель на базе агентного подхода поможет исследователю в изучении сложных, трудноформализуемых, самоорганизующихся процессов и явлений во многих прикладных областях. Однако разработка такой системы весьма не тривиальна, ведь полноценный агент полностью автономен и независим: он совершает целенаправленные действия, вступает в контакт с другими агентами, умеет принимать решения, перемещаться и приспосабливаться к среде и т.п.

Какие задачи могут возникнуть при построении агентной модели? Приведём основные задачи.

Среда моделирования. Следует определить математическую модель среды моделирования, в которой существуют агенты. За основу можно взять, например, списковые структуры, графы, деревья, и т.д.

Методы решения экстремальных задач. Например, в случае графовых моделей агентам потребуется поиск кратчайших путей (циклов, парасочетаний, деревьев, сечений и т.п.).

Способ задания агента. Агент должен обладать памятью, интенциональными характеристиками, механизмами коммуникации, интерфейсами для сбора статистики и т.д.

Способ взаимодействия с окружающей средой. Агент должен знать структуру среды, либо уметь её воспринимать с помощью датчиков/сенсоров.

Способ взаимодействия друг с другом. Необходимо разработать некоторый протокол коммуникации.

Способ задания поведения агента. Каждый агент выполняет свои действия самостоятельно. Это отличает агентный подход к моделированию от всех остальных.

Способ изменения поведения агента. Поведение может измениться, например, в ходе коммуникации с другими агентами или в результате пополнения знаний.

Механизм принятия решений. Агент – автономная сущность, значит, что в случае неопределённости он должен уметь принимать конкретные решения согласно поставленным перед собою целей и задач.

Механизм вывода в условиях неопределённости. В некоторых задачах может возникнуть целый класс проблем, связанный с тем, что стандартный механизм принятия решений не может изменить состояние агента. Агент, по возможности, должен уметь действовать в любых обстоятельствах. В связи с этим требуется реализация вывода на ненадёжных знаниях.

Механизм пополнения знаний. Для адекватного представления модели важно умение агента пополнять свою базу знаний, в том числе во время прогона модели.

Использование метазнаний. Метазнания играют очень важную роль в представлении поведения агента: от управления стратегией вывода до поиска и разрешения противоречий, существующих в текущей базе знаний.

Средства адаптации. Агент должен уметь изменять свою внутреннюю структуру, поведение и семантику при различных воздействиях среды моделирования.

Механизм продвижения времени. Требуется определить, как продвигать время в среде моделирования.

Условия моделирования. На полученную математическую модель пользователем «накладываются» условия моделирования – например, информационные процедуры, которые задают конкретный процесс имитации.

Механизмы сбора статистики. Исследователь должен задать нужные ему параметры моделирования, а специальные программные модули (возможно также агенты) должны уметь собрать эти данные.

Методы отладки моделирования. Отладка поможет исследователю детально изучить конкретный прогон модели. Желательно, чтобы модель умела воссоздать и повторить завершившийся ранее прогон.

Способ задания модели. Это может быть некоторый язык моделирования (графический, текстовый и т.п.), с помощью которого пользователем описывается модель.

Методы распараллеливания имитационного процесса. В случае распределённой системы возникнут проблемы отправки/приёма сообщений, совместного доступа к данным, синхронизации модельного времени и т.д.

Определение общей архитектуры приложения. Следует определить интерфейсы, модули, подпрограммы, компоненты, структуры данных, протоколы, объекты, ресурсы, потоки управления и т.д.

Данный перечень, очевидно, не является полным. Агентное моделирование будет развиваться вместе с развитием других направлений математики и информатики: будут появляться новые численные и статистические методы, направления в развитии искусственного интеллекта, новые алгоритмы адаптации, способы представления нечётких знаний, способы параллельных и распределённых вычислений.

С другой стороны, реализация полного перечня даже этих требований может оказаться весьма трудозатратной. Необходима возможность наращивания возможностей системы. В связи с этим можно прийти к выводу, что ставится задача создания не просто агентной модели, а полноценной платформы, обладающей гибкой, расширяемой и масштабируемой архитектурой для добавления и интеграции новых компонентов.

Дадим краткую характеристику возможных решений, методов и технологий для реализации агентной модели. Рассмотрим, какие подходы уже используются, а какие вопросы остаются нерешёнными.

Нейросетевые технологии. Нейронные сети уже давно используются в моделировании – на их основе базируются множество систем. В этом подходе явно прослеживается проблема адаптации – для интеллектуальной нейронной сети требуется достаточно большая (и желательно репрезентативная) выборка для обучения и тестирования. В ходе прогона модели агент может получить новые данные и на этапе выполнения переобучать сеть.

Методы оптимизации. Некоторые авторы предлагают на базе математической модели построить различные функции (функционалы) и найти условный экстремум. Если мы имеем дело со сложной многокритериальной экстремальной задачей, то стоит задуматься о привлечении методов эвристического поиска (например, на базе генетических алгоритмов). В этом случае также возникает задача адаптации – подбор параметров и анализ пространства решений для ускорения поиска экстремума.

МАС. Теорию мультиагентных систем стоит изучить более детально, т.к. она оперирует общими с агентным моделированием понятиями. Стоит обратить внимание на различные способы представления и коммуникации агентов.

Методы теории игр и исследования операций. В качестве механизмов принятия решений можно применять подходы, основанные на матричных играх, теории полезности и сетевого планирования. Если задача (подзадача) формулируется чётко, то они, безусловно, будут

иметь преимущество перед эвристиками. В одной работе предлагалось сначала математическими методами найти точное решение, а затем его уточнить с помощью МАС.

Нечёткая логика. Предлагаются идеи использования нечёткой логики. Поскольку агент должен уметь принимать решения недетерминированного характера в условиях неопределённости, многозначности, неполноты и не монотонности, то он должен, во-первых, обладать памятью для хранения фактов, а во-вторых, уметь строить рассуждения на их основе. Поскольку факты, как по своей природе, так и в силу коммуникации с другими агентами могут быть разнородны и противоречивы, то требуется реализовать вывод в нечёткой логике.

Вывод на базе знаний. Агент может иметь встроенные средства базы знаний, в т. ч. экспертные системы. Например, предлагают использовать продукционную базу знаний с метаправилами до 3-го уровня включительно. Раз агент в течение продолжительного времени сосуществует с другими агентами в определённой среде, то он может накапливать новые факты. А это вновь потребует реализации некоторых механизмов адаптации – сбора статистики, добавления в ходе прогона новых правил, изменения правил с помощью метаправил, разрешения противоречий в базе знаний и т.д.

Методы децентрализованного принятия решений. Данный аспект прослеживается практически во всех существующих работах. И этого следовало ожидать: агентный подход объясняет процессы в системе через независимое децентрализованное поведение агентов. Более того, принятие решений может быть распределено внутри агента. В этом случае агент действует на базе нескольких источников принятия решений. Тут вновь возникает задача адаптации – если один источник часто даёт неправильные решения, то агент должен полагаться на другие источники.

Voids-моделирование. Хорошо подходит не только для организации поведения животных, но и для более сложных структурных самоорганизующихся систем, например социальных групп. Существуют различные схемы реализации – алгоритмы муравьиных колоний, птичьих стай, сахарных плантаций и т.д.

Теория графов. Очень много работ опираются на применение графов для представления модели. Теория графов достаточно хорошо изучена, имеет методы решения многих задач, как точные, так и эвристические. Поэтому следует учесть опыт исследователей и попробовать применить данную теорию.

Геоинформационные системы (ГИС). Если рассматривать теорию ГИС (без привязки конкретно к картографии), то предлагаемые методы анализа и обработки географически распределённых объектов вполне можно применить к среде, в которой существуют агенты.

Методы теории вероятностей и математической статистики. Предлагают и этот подход, однако авторы утверждают, что его следует использовать с осторожностью. Дело в том, что агентное моделирование представляет взгляд на систему «снизу-вверх», в то время как теория вероятностей больше подходит для типа «сверху-вниз», когда известны глобальные законы распределения, статистические оценки, априорные/апостериорные вероятности тех или иных процессов и т.д. Однако статистика располагает некоторыми полезными приёмами: анализ выборки, построение статистических гипотез, выявление корреляции между факторами, доверительное оценивание, регрессионный анализ и т.д.

Онтологический подход. Агент должен получать, хранить и обрабатывать информацию о текущем состоянии предметной области. Представление агента об окружающем мире, о других агентах и о себе самом может быть специфицировано в форме онтологии. Определение и построение онтологии включает анализ предметной области, выделение базовых онтологических элементов (объектов, их атрибутов, отношений и процессов), проведение операций над этими онтологическими элементами. Если агенты будут оперировать одними и теми же понятиями, то это поможет при решении многих проблем – от методов коммуникации между агентами до способов адаптации к новым условиям. Использование онтологий осложняется тем, что разрабатываемая система моделирования инвариантна относительно предметной области. Поэтому онтология (по крайней мере, предметного уровня) будет изменяться.

Комбинирование различных форм моделирования. Предлагают применение синергетической интеграции различных методов, приёмов, алгоритмов и технологий для получения более правдоподобного результата в моделировании сложных экономических процессов. Разумеется, агентное моделирование должно идти по пути интеграции, т.к. каждая методика обладает своими преимуществами и недостатками. Тут вновь возникает достаточно нетривиальная задача адаптации. Например, агент может с помощью логического вывода или вывода на семантических сетях принимать одни решения, но затем, приспособо-

бившись к среде и пообщавшись с другими агентами, изменить ход вывода и принять другие решения.

Итак, на примере предложенных методов мы убедились, насколько не проста задача агентного моделирования. Нет чётко определённого подхода ни в способах проектирования, ни в выборе методов и алгоритмов, ни в способах представления агента, ни в выборе математической модели, ни в применении средств искусственного интеллекта.

По словам авторов, хорошо себя зарекомендовали теория графов, методы оптимизации и нейронные сети. Однако тот же нейросетевой подход является далеко не единственным и исчерпывающим. Поскольку в системе моделирования требуется решение более широкого класса задач (помимо классификации/кластеризации), то нельзя ограничиваться лишь одним подходом. Аналогично и математическое моделирование – его, безусловно, следует применять, но не в тех случаях, когда задача требует эвристических решений (например, для моделирования поведения человека).

Большинство работ пропагандируют интегрирование различных форм решений проблем агентного моделирования. Действительно, агентный подход явно подразумевает, что моделирование процессов системы суммируется из составных её элементов. Это порождает массу подзадач, для большинства из которых уже существуют решения. Казалось бы, осталось лишь объединить эти решения в одну общую архитектуру. Однако вопрос, каким образом это следует делать, до сих пор остаётся открытым. Более того, решения могут отличаться:

- по степени эффективности (например, генетические алгоритмы могут найти хорошее решение, но не всегда);

- по степени применимости (методы могут работать, но лишь в ограниченных условиях);

- по степени адекватности (разные методы могут давать разные решения – существует проблема выбора с точки зрения задачи).

Из таблицы 6.1 видно, что хорошо проработаны лишь те проблемы, которые характерны для имитационного моделирования вообще: продвижение времени, сбор статистики, отладка. Общие вопросы определения модели, методов коммуникации, механизма принятия решений и т.п. занимают промежуточное положение – имеются готовые решения, однако они не всегда применимы, и при разработке архитектуры системы могут возникнуть трудности. Вопросы адекватного представления поведения агента, методов адаптации, вывода в условиях неопределённости и т.д. можно считать открытыми.

**Применимость существующих решений
к задаче агентного моделирования**

| | | |
|---|--|---|
| Имеются решения, в той или иной мере применимые для агентного моделирования | Имеются решения, но их использование и сопровождение может представить затруднения | Решения, возможно, и существуют, но они могут быть неадекватны, неполны, неэффективны, либо работать лишь в частных случаях |
| Методы решения экстремальных задач | Среда моделирования | Способ задания агента |
| Механизм продвижения времени | Способ взаимодействия с окружающей средой | Способ задания поведения агента |
| Условия моделирования | Способ взаимодействия агентов | Способ изменения поведения агента |
| Механизмы сбора статистики | Механизм принятия решений | Механизм вывода в условиях неопределённости |
| Методы отладки моделирования | Способ задания модели (язык) | Механизм пополнения знаний |
| | Методы распараллеливания имитационного процесса | Использование мета-знаний |
| | | Средства адаптации |
| | | Определение общей архитектуры приложения |

Эти вопросы актуальны в связи с тем, что агенты предпринимают попытку моделировать сложные *естественные* процессы, возникающие в реальном мире.

Попытка учесть всё и всюду вряд ли увенчается успехом, поскольку задача *выбора инструментальных средств для построения системы агентного моделирования* также является сложной и трудно формализуемой. Поэтому остаётся так много открытых вопросов, что создание идеальной системы моделирования, учитывающей все значимые факторы для любой заданной проблемной области, существенно влияющие на результаты моделирования, в настоящее время представляется невозможным.

6.2.3. Общая схема системы агентного моделирования

Рассмотрим общую схему системы агентного моделирования, представленную на рис. 6.3.

Система управления. Управляет процессом имитации. Она должна состоять из нескольких уровней, различающихся по степени абстрактности в управлении данными и процедурами (рис. 6.4):

Нулевой уровень – это часть архитектуры, зависящая от аппаратуры и программного обеспечения. Представляет собой конкретную реализацию функций ядра.

Ядро. Содержит основные функции по управлению и координированию составных частей системы. Функции должны оперировать абстрактными понятиями на уровне интерфейсов, т.е. не должны зависеть от конкретной реализации.

Системные компоненты. Содержат методы, алгоритмы, схемы и структуры данных по управлению ходом имитационного процесса с помощью функций ядра.

API системы. Представляет собой прослойку для доступа внешних компонентов к системным компонентам и/или функциям ядра. Фактически, это – имена процедур, по которым внешние модули могут принять участие в управлении имитационным процессом.

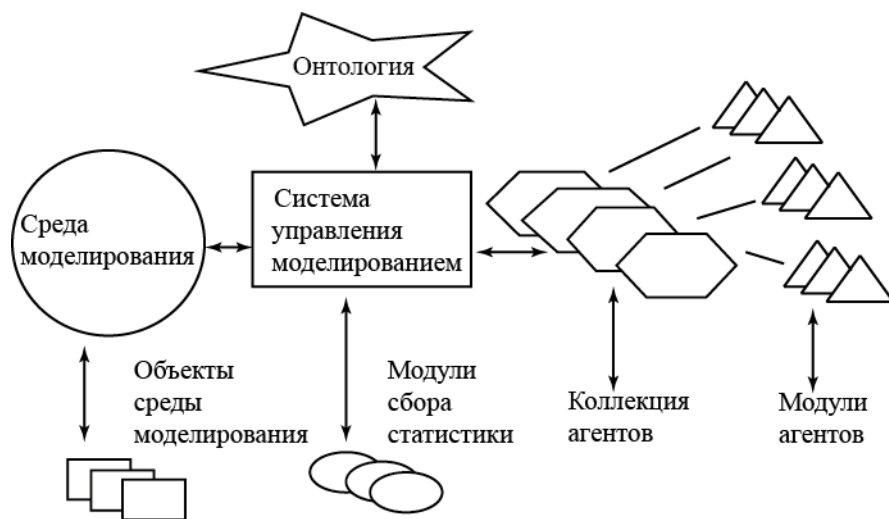


Рис. 6.3. Общая схема системы агентного моделирования

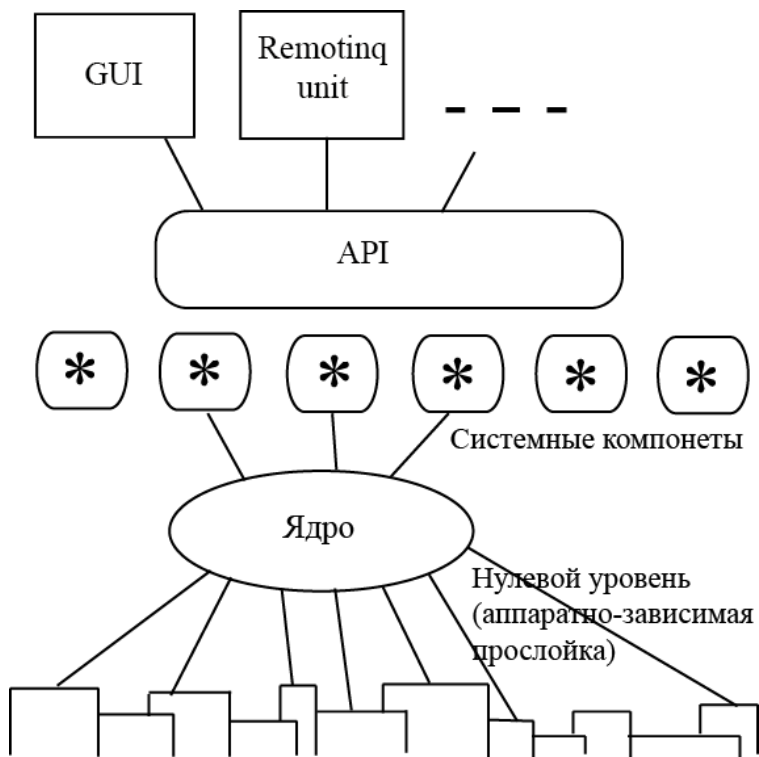


Рис. 6.4. Структура системы управления

Компонента графического интерфейса. Здесь могут использоваться как аппаратно-зависимые интерфейсы (GDI+, X11, Quartz 2D), так и специализированные открытые библиотеки (OpenGL). В общем случае, эту часть будем полагать зависимой от платформы.

Уровень удалённого доступа. Представляет собой дополнительный прокси-уровень, отвечающий за поддержку RPC в системах распределённого моделирования. Это позволит остальным компонентам системы единообразно обращаться к методам локальных и удалённых объектов. Уровень предоставляет общие интерфейсы по маршрутизации и пересылке вызовов объектов, однако конкретная реализация будет зависеть от платформы (XML RPC, RMI, .NET Remoting).

Агенты – это объекты, которые используют интерфейсы, предлагаемые ядром и системными компонентами. Агенты также имеют собственные интерфейсы для контроля и сбора статистики (рис. 6.5).

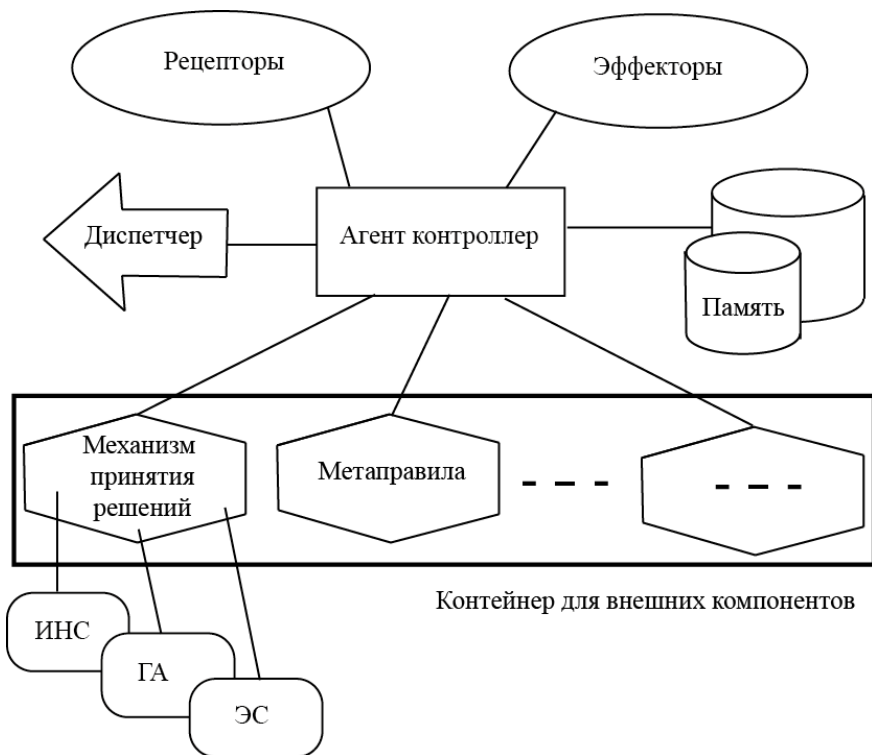


Рис. 6.5. Модульная архитектура агента

Архитектура агента включает центральный *компонент-контроллер*, занимающийся процессом жизнедеятельности агента. Получив управление, он обращается к *рецепторам* с целью получения информации о среде моделирования. Примером может служить зрение человека. Далее контроллер, основываясь на полученных данных и данных из памяти агента, передаёт управление *модулям*.

Под модулями агентов будем понимать внешние присоединяемые компоненты, которые отвечают за структуру, поведение и семантику агента. Это могут быть базы знаний, механизмы принятия решений, механизмы логического вывода, модули сбора информации для настройки весов нейросетей, анализаторы данных, модули управления агентом в условиях неопределённости и т.д. Обычно они представляются в виде функциональных блоков, т.е. содержат вход, выход, подсистему управления и средства реализации.

После отработки всех модулей агент с помощью *эффекторов* предпринимает активные действия (например, движение в некотором направлении). Информация обо всех действиях передаётся системе управления через *диспетчер*. Таким образом, диспетчер служит связующим звеном агента с системой моделирования.

Согласно принципам открытых систем, модули должны иметь доступ к данным и процедурам, доступным на уровне агента (информация о среде моделирования, онтология предметного уровня, функции системы управления). Также они сами должны предоставить доступ к собственным интерфейсам (в частности, для сбора статистики).

В заключение ещё раз отметим, что ведущую роль в изучении сложных динамических процессов получило имитационное моделирование как инструмент выявления закономерностей явлений, имеющих протяжённость во времени.

Новейший инструмент имитационного моделирования – агентный подход – позволяет взглянуть на проблему по индуктивному методу – от более частного к более общему. Это позволит получать информацию об общих глобальных законах исследуемой модели, базируясь на активности её составных элементов. Предложенный способ позволит выявить имплицитированные зависимости и закономерности, которые вряд ли обнаружат другие методы исследования.

Однако создавать агентные модели под каждую конкретную задачу – дело не только дорогостоящее, но и неэффективное. Причина кроется хотя бы в том, что представляемая модель управляется рядом эвристик, которые невозможно не просто доказать формально, но и предположить вплоть до этапа имитационного прогона. Алгоритмы вывода, поведение и методы коммуникации агентов постоянно уточняются по ходу процесса моделирования.

Выходом из этой проблемы служит построение системы имитационного моделирования. Нетрудно заметить, что эта задача является ещё более сложной и дорогостоящей, особенно для агентного моделирования. Тем не менее, разработка агентных моделей, в том числе военного назначения, возможна, так как появляются программные продукты, которые позволяют это сделать, поскольку функции системы агентного моделирования, может быть не всегда нас полностью удовлетворяющие, там имеются. В качестве примера такой системы моделирования, поддерживающей кроме системной динамики, дискретно-событийный и агентный подходы, может служить AnyLogic.

Вопросы для самоконтроля

- 6.1. Что такое распределённая система?
- 6.2. Какие основные задачи распределённых систем?
- 6.3. Приведите требования, предъявляемые к распределённым системам.
- 6.4. Что понимается под прозрачностью распределённой системы? открытостью? гибкостью? масштабируемостью?
- 6.5. Что такое репликация?
- 6.6. Что понимают под гомогенными и гетерогенными мультикомпьютерными системами?
- 6.7. Каковы причины перехода к распределённому имитационному моделированию?
- 6.8. В чём отличие терминов «параллельная система» и «распределённая система»?
- 6.9. Какова цель распределённого имитационного моделирования для военных приложений?
- 6.10. Чем отличается квазипараллельное моделирование от распределённого моделирования?
- 6.11. Что понимают под логическим процессом?
- 6.12. Есть ли глобальные часы модельного времени и глобальный список событий в распределённой модели?
- 6.13. Как можно определить распределённую модель, состоящую из N логических процессов?
- 6.14. Как общаются между собой логические процессы?
- 6.15. Приведите и поясните схему распределённой имитационной модели?
- 6.16. Что является критерием правильной работы распределённой имитационной модели?
- 6.17. Что такое парадокс времени в распределённой имитационной модели?
- 6.18. Каковы причины и цель появления имитационного агентного моделирования?
- 6.19. Что понимают под агентом?
- 6.20. Перечислите основные свойства агента.
- 6.21. Что потребуется для исследования агентного подхода?
- 6.22. Какие задачи возникают при построении агентной модели?
- 6.23. Приведите и поясните общую схему агентной модели.
- 6.24. Приведите и поясните архитектуру агента.

Заключение

В основе исследования сложных систем с использованием математического моделирования лежит системный подход, конечной целью которого является системное проектирование, направленное на построение системы с заданным качеством. В свою очередь системное проектирование базируется на результатах системного анализа, позволяющего выявить причинно-следственные связи между параметрами и показателями исследуемой системы. Системный анализ проводится с использованием математических моделей, предоставляющих возможность прогнозировать эффект, достигаемый при изменении структурно-функциональных параметров системы и параметров нагрузки.

Одним из основных требований, предъявляемых к модели, является её *адекватность* реальной системе, которая достигается за счёт использования моделей с различным уровнем детализации, зависящим от особенностей структурно-функциональной организации системы и целей исследования.

Процессы функционирования реальных систем практически невозможно описать полно и детально, что обусловлено существенной сложностью таких систем. Основная проблема при разработке модели состоит в нахождении компромисса между простотой её описания, что необходимо для её исследования математическими методами, и необходимостью учёта многочисленных особенностей, присущих реальной системе. Попытка построить единую универсальную модель сложной системы, несомненно, обречена на неудачу ввиду её необозримости и невозможности расчета.

Моделирование военно-технических систем в общем случае предполагает выполнение следующих основных этапов:

- формулировка целей моделирования;
- разработка концептуальной модели;
- разработка математической модели;
- параметризация модели;
- выбор методов моделирования;
- выбор средств моделирования;
- проверка адекватности модели (верификация модели);
- проведение экспериментов на модели (расчет характеристик);
- анализ результатов моделирования.

На этапе *определения и формулирования целей моделирования* определяется объект моделирования, формулируются задачи анализа и синтеза, выявляются наиболее важные характеристики, подлежащие исследованию, формулируются требования к качеству функционирования в виде ограничений, налагаемых на характеристики системы, и формулируется критерий эффективности, определяются требования к точности результатов моделирования и форме их представления.

Основное *назначение концептуальной модели* – выявление наиболее существенных сторон структурно-функциональной организации, учёт которых необходим для получения требуемых результатов. В концептуальной модели обычно в словесной форме приводятся сведения о природе и параметрах элементарных явлений исследуемой системы, о степени их взаимодействия, выявляются параметры, оказывающие наиболее существенное влияние на исследуемые характеристики системы. Одна и та же система может представляться различными концептуальными моделями, которые строятся в зависимости от целей исследования, сформулированных на предыдущем этапе. Например, одна концептуальная модель может отображать временные аспекты функционирования системы, другая – надёжностные, третья – структурные аспекты построения системы.

Концептуальная модель служит основой для разработки математической модели в терминах конкретного математического аппарата.

Создание *математической модели* преследует две основные цели:

- 1) дать формализованное описание структуры и процесса функционирования системы для однозначности их понимания;
- 2) попытаться представить процесс функционирования системы в виде, допускающем аналитическое исследование системы с использованием методов, разработанных в рамках данного математического аппарата.

В связи с тем, что состав и номенклатура системных и модельных параметров и показателей, в общем случае, различается, возникает необходимость установления соответствия между значениями системных и модельных параметров и характеристик, которое выполняется на этапе *параметризации модели*.

Выбор метода моделирования зависит от многих факторов, в том числе от целей моделирования, сложности исследуемой системы, требований к номенклатуре исследуемых характеристик, требований к точности получаемых результатов и т.д. При исследовании и проекти-

ровании военно-технических систем наиболее эффективным оказывается использование комбинированного подхода, предполагающего совместное применение аналитических и имитационных методов, что позволяет во многих случаях гарантировать достоверность получаемых результатов. С использованием аналитических методов, применяемых на этапах анализа свойств и синтеза оптимальной системы, решаются задачи, связанные с формированием требований к структурным и функциональным параметрам, обеспечивающим заданное качество функционирования системы, однако получаемые при этом результаты могут иметь значительную погрешность. Имитационные методы, основанные на использовании инструментальных средств систем моделирования, таких как GPSS World, AnyLogic позволяют выполнять исследование систем практически любой сложности с любой степенью детализации и должны применяться на заключительном этапе детального анализа спроектированной системы.

Технические и программные *средства моделирования* выбираются с учётом ряда факторов, к которым относятся достаточность и полнота средств для реализации концептуальной и математической моделей, доступность средств, простота и лёгкость освоения технических и программных средств моделирования, наличие методики применения средств для моделирования систем определённого класса. После выбора средств моделирования разрабатывается программная модель.

Проверка адекватности модели исследуемой системе (верификация модели) заключается в анализе её соответствия исследуемой системе, проявляющегося в близости значений модельных и системных показателей. Отличие модели от исследуемой системы связана с тем, что обычно модель является упрощенным и идеализированным отображением системы. Это обусловлено идеализацией внешних условий и режимов функционирования, не учитывающим в модели несущественных, по мнению исследователя, факторов и параметров, отсутствием точных сведений о внешних воздействиях и о некоторых конкретных особенностях организации системы, введением ряда упрощающих предположений и допущений. На практике верификация модели обычно проводится путем экспертного анализа разумности результатов моделирования. В случае выявления неадекватности модели исследуемой системе необходимо выполнить корректировку модели.

В процессе проверки адекватности модели необходимо определить область применения модели, то есть оценить диапазон изменения па-

раметров, при котором точность результатов моделирования находится в допустимых пределах.

Исследования на моделях заключаются в *проведении экспериментов*, в процессе которых определяются параметры и показатели системы при разных значениях структурно-функциональных параметров и параметров нагрузки. Большой перечень исходных параметров и широкий диапазон их изменения требует предварительного планирования выполняемых на модели экспериментов. Планирование направлено на уменьшение количества и длительности экспериментов при условии обеспечения требуемой точности и достоверности, полноты результатов моделирования. На математическом языке задача планирования эксперимента формулируется следующим образом: нужно выбрать оптимальное, в некотором смысле, расположение точек (координат) в факторном пространстве, чтобы получить некоторое представление о поверхности отклика. Выбор критерия оптимальности в значительной степени произволен.

Особую значимость планирование экспериментов приобретает при использовании методов имитационного моделирования, характеризующихся большими затратами ресурсов ЭВМ.

Одной из основных проблем имитационного моделирования является нахождение компромисса между временем моделирования и затратами памяти ЭВМ, на которой проводится моделирование. Это связано с тем, что имитационное моделирование предъявляет повышенные требования, как к производительности, так и к памяти ЭВМ для проведения имитационных экспериментов. Время, затрачиваемое на проведение одного эксперимента с моделью средней сложности даже на высокопроизводительных ЭВМ, может достигать нескольких десятков минут и, в некоторых случаях, нескольких часов, а потребность в оперативной памяти ЭВМ – десятков и сотен гигабайт. Причём с увеличением числа проводимых имитационных экспериментов соответственно возрастает время моделирования. Все это обуславливает высокую стоимость имитационного моделирования и требует тщательного планирования имитационных экспериментов с целью сокращения затрат на моделирование.

Анализ результатов моделирования направлен на выявление свойств исследуемой системы, и включает в себя следующие этапы:

- 1) обработка результатов для последующего анализа и использования; на этом этапе выделяются наиболее важные, с точки зрения ис-

следователя, результаты, которые представляются в форме, наиболее удобной для изучения свойств исследуемой системы;

2) определение зависимостей показателей от параметров системы путём варьирования исходных параметров нагрузки и структурно-функциональной организации с целью выявления свойств системы;

3) принятие решения о работоспособности исследуемой системы и выработка рекомендаций по наиболее эффективной и рациональной организации проектируемой или модернизируемой системы, которые могут быть использованы в дальнейшем при решении задач синтеза в процессе проектирования.

Синтез оптимальной системы направлен на построение системы, наилучшим образом соответствующей своему назначению. Решение задачи синтеза связано с определением зависимостей показателей функционирования системы от параметров, которые представляются сложными математическими зависимостями. При этом возможность получения приемлемых результатов в процессе решения задач синтеза из-за их сложности и большой трудоёмкости с учётом специфических особенностей реальных систем превосходит возможности математических методов оптимизации, и задача синтеза в общем виде оказывается математически неразрешимой. Для снижения сложности задачи синтеза, процесс проектирования разделяют на этапы. На каждом из них решаются частные задачи синтеза – определяются параметры отдельных сторон структурно-функциональной и информационной организации системы, с использованием различных моделей.

Системы имитационного моделирования GPSS World и AnyLogic предоставляют исследователям возможности проведения оптимизационных экспериментов на имитационных стохастических моделях. Они позволяют получить экстремальное значение целевой функции и необходимые для её достижения оптимальные значения параметров функционирования моделируемой системы.

Синтез оптимальной структуры в данном случае проводится с использованием методов математической статистики, учитывающих стохастический характер результатов. В отношении статистических выводов нужно всегда иметь в виду следующее: положительные результаты проверки статистической гипотезы означают лишь то, что постулируемая гипотеза не противоречит результатам эксперимента. Результаты проверки гипотезы никогда не могут служить доказательством абсолютной справедливости и правильности гипотезы.

Список литературы

1. Алиев Т. И. Основы моделирования дискретных систем: Учеб. пособие.– СПб., СПбГУ ИТМО, 2009.
2. Арлей Н., Бух К. Введение в теорию вероятностей и математическую статистику. – М.: ИЛ, 1951.
3. Боев В. Д. Об адекватности систем имитационного моделирования GPSS World и AnyLogic. – Прикладная информатика, № 6 (30), 2011.
4. Боев В. Д. Об адекватности систем имитационного моделирования GPSS World и AnyLogic. – Прикладная информатика, № 4 (34), 2013.
5. Боев В. Д. Проектирование и моделирование систем: Учеб. пособие. – СПб., ВАС, 2012.
6. Боев В. Д. Концептуальное проектирование систем в AnyLogic 7 и GPSS World: Учеб. пособие. – ИНТУИТ.ru, 2014. ISBN: 978-5-9556-0161-8.
7. Боев В. Д. Имитационное моделирование на GPSS/PC систем обеспечения войск: Учеб. пособие. – СПб.: ВАУ, 1999.
8. Боев В. Д. Моделирование систем. Инструментальные средства GPSS World: Учеб. пособие. – СПб.: БХВ-Петербург, 2004.
9. Боев В. Д. Модель бизнес-процесса и особенности ее реализации в системе моделирования: Статья. – В сб. докладов конференции «Имитационное моделирование. Теория и практика» ИММОД-2005 – СПб.: ФГУП ЦНИИТС, 2005.
10. Боев В. Д. Решение в системе моделирования прямой и обратной задач: Статья. – В сб. докладов конференции «Имитационное моделирование. Теория и практика» ИММОД-2005 – СПб.: ФГУП ЦНИИТС, 2005.
11. Боев В. Д., Таран С. П. Вариант построения имитационной модели функционирования системы восстановления ракетно-артиллерийского вооружения объединения. – Труды Десятой Всероссийской научно-практической конференции. Том 2. – СПб.: НПО специальных материалов, 2007.
12. Боев В. Д., Таран С. П. Прототип автоматизированного рабочего места начальника службы ракетно-артиллерийского вооружения объединения. – Труды Десятой Всероссийской научно-практической конференции. Том 2. – СПб.: НПО специальных материалов, 2007.
13. Боев В. Д., Таран С. П. Модель функционирования системы восстановления ракетно-артиллерийского вооружения объединения в операции. Депонированная рукопись № 15230. – М.: ЦВНИИ МО РФ, 2007.
14. Боев В. Д. Компьютерная методика оценки обеспеченности ракетно-артиллерийским вооружением соединений и частей общевойскового объединения в операции. Депонированная рукопись № 15231. – М.: ЦВНИИ МО РФ, 2007.
15. Боев В. Д., Ушкань А. О. Методика оценки качества обслуживания сети передачи данных: Статья. – В сб. докладов Четвертой Всероссийской конференции «Имитационное моделирование. Теория и практика» ИММОД-2009 – СПб.: СПб.: ФГУП ЦНИИТС, 2009.
16. Боев В. Д. Вторичные модели оценки качества обслуживания сети передачи данных: Статья. – В сб. докладов Четвёртой Всероссийской конференции «Имитационное моделирование. Теория и практика» ИММОД-2009. – СПб.: ФГУП ЦНИИТС, 2009.

17. Боев В. Д. Информационные системы и технологии в экономике и управлении. – Монография. – СПб.: СПб ГПУ, 2010.
18. Боев В. Д., Моисеев Р. А. Некоторые классы типовых объектов сетей связи в AnyLogic. Материалы Всероссийской конференции ИММОД-2013 – Казань, 2013.
19. Боев В. Д., Моисеев Р. А. Имитационная модель самоорганизующейся сети связи. – Инфокоммуникационные технологии в инновациях, медико-биологических и технических науках: сборник научных трудов Пятого международного научного конгресса «Нейробиотелеком-2012». – СПб.: Политехника, 2012.
20. Боев В. Д., Моисеев Р. А. Алгоритм имитационной модели функционирования сети связи. – Сб. трудов № 78, ВАС, 2012.
21. Боев В. Д., Моисеев Р. А. Классы типовых объектов сетей связи в среде имитационного моделирования. – Сб. трудов № 78, ВАС, 2012.
22. Боев В. Д., Бочков А. П., Филюстин А. Е. Датчик равномерно распределенных случайных чисел: Авт. св. №1381499, бюл. №10, 1988.
23. Боев В. Д., Филюстин А. Е. Устройство для определения прогнозных оценок случайного процесса: Авт. св. №1383406, бюл. №11, 1988.
24. Боев В. Д., Филюстин А. Е. Датчик равномерно распределенных случайных чисел: Авт. св. №1594530, бюл. №35, 1990.
25. Боев В. Д., Бочков А. П., Филюстин А. Е. Датчик случайных чисел, распределенных по треугольному закону: Авт. св. №1674116, бюл. №32, 1992.
26. Боев В. Д., Мартыщенко Л. А., Филюстин А. Е. Генератор случайных чисел: Авт. св. №1781681, бюл. №46, 1992.
27. Боев В. Д., Лысенков А. И., Филюстин А. Е. Генератор случайных чисел: Авт. св. №1798780, бюл. №8, 1993.
28. Боев В. Д., Бочков А. П., Филюстин А. Е. Генератор случайных чисел: Авт. св. №1833868, бюл. №30, 1993.
29. Борщёв А. В. Практическое агентное моделирование и его место в арсенале аналитика: Статья. – В сб. докладов конференции ИММОД-2005 «Имитационное моделирование. Теория и практика». – СПб.: ФГУП ЦНИИТС, 2005.
30. Бусленко Н. П. Автоматизация имитационного моделирования сложных систем. – М.: Наука, 1978.
31. Варжапетян А. Г. Имитационное моделирование на GPSS/H: Монография / А. Г. Варжапетян. – Вузовская книга, 2007.
32. Введение в математическое моделирование: Учеб. пособие / Под ред. Г. В. Трусова. – М.: Университетская книга, Логос, 2007.
33. Вентцель Е. С. Исследование операций. – М.: Советское радио, 1972.
34. Вентцель Е. С. Теория вероятностей: Учеб. для вузов / 10-е изд. стер. – М.: Высшая школа, 2006.
35. Вентцель Е. С. Теория случайных процессов и её инженерные приложения: Учеб. для вузов / 4-е изд. стер. – М.: Высшая школа, 2007.
36. Гмурман В. Е. Руководство к решению задач по теории вероятностей и математической статистике: Учеб. пособие. – М.: Высшая школа, 2003.
37. Девятков В. В. Методология и технология имитационных исследований сложных систем: современное состояние и перспективы развития: Монография / В. В. Девятков. М.: Вузовский учебник: ИНФРА-М, 2013.
38. Девятков В. В. Расширенный редактор GPSS World – основные возможности: научное издание / В. В. Девятков. – М.: Издание ООО «Принт-сервис», 2013.

39. Девятков В. В. Мир имитационного моделирования: взгляд из России. – Прикладная информатика, № 4 (34), 2011.
40. Замятина О. М. Моделирование систем: Учеб. пособие. – Томск: Изд-во ТГУ, 2009.
41. Карпов Ю. Г. Имитационное моделирование систем. Введение в моделирование с AnyLogic. – СПб.: БХВ-Петербург, 2005.
42. Кобелев Н. Б. Введение в общую теорию имитационного моделирования. – М.: Принт-Сервис, 2007.
43. Крэйн М., Лемуан О. Введение в регенеративный метод анализа моделей. – М.: Наука, 1982.
44. Лоу А., Кельтон Д. Имитационное моделирование. – СПб.: Питер, БХВ-Петербург, 2004.
45. Митраков А. А. Основы проектирования системы агентного моделирования. – Пермь: ПГНИУ, 2013.
46. Некрасова Р. С. Регенеративный метод и его применение в анализе СМО: монография/Р. С. Некрасова, Е. В. Морозов. – Петрозаводск: Изд-во ПетрГУ, 2013.
47. Окольнішников В. В. Представление времени в имитационном моделировании. – Новосибирск: СО РАН, Вычислительные технологии, том 10, №5, 2005.
48. Смирнов Н. В., Дунин-Барковский И. В. Курс теории вероятностей и математической статистики. – М.: Наука, 1965.
49. Советов Б. Я., Яковлев С. А. Моделирование систем: Практикум. – М.: Высшая школа, 2003.
50. Советов Б. Я., Яковлев С. А. Моделирование систем: Учеб. пособие – М.: Высшая школа, 2010.
51. Шеннон Р. Имитационное моделирование – искусство и наука. – М.: Мир, 1978.
52. Шрайбер Т. Дж. Моделирование на GPSS. – М.: Машиностроение, 1980.

Перечень ресурсов Интернет

53. Боев В. Д. Концептуальное проектирование систем в AnyLogic 7 и GPSS World. Курс лекций. ИНТУИТ.ru, 2014.
54. Боев В. Д., Сыпченко Р. П. Компьютерное моделирование. Курс лекций. ИНТУИТ.ru, 2010.
55. Боев В.Д. Компьютерное моделирование. Пособие для практических занятий, курсового и дипломного проектирования в AnyLogic 7. – www.anylogic.ru, 2014.
56. Сайт национального общества имитационного моделирования России simulation.su.
57. Сайт «Имитационное моделирование систем» www.gpss.ru.
58. Сайт компании разработчика системы имитационного моделирования AnyLogic www.anylogic.ru.

Приложение 1

Таблица значений функции $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-z^2/2} dz$

| x | $\Phi(x)$ | x | $\Phi(x)$ | x | $\Phi(x)$ | x | $\Phi(x)$ |
|------|-----------|------|-----------|------|-----------|------|-----------|
| 0,00 | 0,0000 | 0,35 | 0,1368 | 0,70 | 0,2580 | 1,05 | 0,3531 |
| 0,01 | 0,0040 | 0,36 | 0,1406 | 0,71 | 0,2611 | 1,06 | 0,3554 |
| 0,02 | 0,0080 | 0,37 | 0,1443 | 0,72 | 0,2642 | 1,07 | 0,3577 |
| 0,03 | 0,0120 | 0,38 | 0,1480 | 0,73 | 0,2673 | 1,08 | 0,3599 |
| 0,04 | 0,0160 | 0,39 | 0,1517 | 0,74 | 0,2703 | 1,09 | 0,3621 |
| 0,05 | 0,0199 | 0,40 | 0,1554 | 0,75 | 0,2734 | 1,10 | 0,3643 |
| 0,06 | 0,0239 | 0,41 | 0,1591 | 0,76 | 0,2764 | 1,11 | 0,3665 |
| 0,07 | 0,0279 | 0,42 | 0,1628 | 0,77 | 0,2794 | 1,12 | 0,3686 |
| 0,08 | 0,0319 | 0,43 | 0,1664 | 0,78 | 0,2823 | 1,13 | 0,3708 |
| 0,09 | 0,0359 | 0,44 | 0,1700 | 0,79 | 0,2852 | 1,14 | 0,3729 |
| 0,10 | 0,0398 | 0,45 | 0,1736 | 0,80 | 0,2881 | 1,15 | 0,3739 |
| 0,11 | 0,0438 | 0,46 | 0,1772 | 0,81 | 0,2910 | 1,16 | 0,3770 |
| 0,12 | 0,0478 | 0,47 | 0,1808 | 0,82 | 0,2939 | 1,17 | 0,3790 |
| 0,13 | 0,0517 | 0,48 | 0,1844 | 0,83 | 0,2967 | 1,18 | 0,3810 |
| 0,14 | 0,0557 | 0,49 | 0,1879 | 0,84 | 0,2995 | 1,19 | 0,3830 |
| 0,15 | 0,0596 | 0,50 | 0,1915 | 0,85 | 0,3023 | 1,20 | 0,3849 |
| 0,16 | 0,0636 | 0,51 | 0,1950 | 0,86 | 0,3051 | 1,21 | 0,3869 |
| 0,17 | 0,0675 | 0,52 | 0,1985 | 0,87 | 0,3078 | 1,22 | 0,3883 |
| 0,18 | 0,0714 | 0,53 | 0,2019 | 0,88 | 0,3106 | 1,23 | 0,3907 |
| 0,19 | 0,0753 | 0,54 | 0,2054 | 0,89 | 0,3133 | 1,24 | 0,3925 |
| 0,20 | 0,0793 | 0,55 | 0,2088 | 0,90 | 0,3159 | 1,25 | 0,3944 |
| 0,21 | 0,0832 | 0,56 | 0,2123 | 0,91 | 0,3186 | 1,26 | 0,3962 |
| 0,22 | 0,0871 | 0,57 | 0,2157 | 0,92 | 0,3212 | 1,27 | 0,3980 |
| 0,23 | 0,0910 | 0,58 | 0,2190 | 0,93 | 0,3238 | 1,28 | 0,3997 |
| 0,24 | 0,0948 | 0,59 | 0,2224 | 0,94 | 0,3264 | 1,29 | 0,4015 |
| 0,25 | 0,0987 | 0,60 | 0,2257 | 0,95 | 0,3289 | 1,30 | 0,4032 |
| 0,26 | 0,1026 | 0,61 | 0,2291 | 0,96 | 0,3315 | 1,31 | 0,4049 |
| 0,27 | 0,1064 | 0,62 | 0,2324 | 0,97 | 0,3340 | 1,32 | 0,4066 |
| 0,28 | 0,1103 | 0,63 | 0,2357 | 0,98 | 0,3365 | 1,33 | 0,4082 |
| 0,29 | 0,1141 | 0,64 | 0,2389 | 0,99 | 0,3389 | 1,34 | 0,4099 |
| 0,30 | 0,1179 | 0,65 | 0,2422 | 1,00 | 0,3413 | 1,35 | 0,4115 |
| 0,31 | 0,1217 | 0,66 | 0,2454 | 1,01 | 0,3438 | 1,36 | 0,4131 |
| 0,32 | 0,1255 | 0,67 | 0,2486 | 1,02 | 0,3461 | 1,37 | 0,4147 |
| 0,33 | 0,1293 | 0,68 | 0,2517 | 1,03 | 0,3485 | 1,38 | 0,4162 |
| 0,34 | 0,1331 | 0,69 | 0,2549 | 1,04 | 0,3508 | 1,39 | 0,4177 |

| x | $\Phi(x)$ | x | $\Phi(x)$ | x | $\Phi(x)$ | x | $\Phi(x)$ |
|------|-----------|------|-----------|------|-----------|------|-----------|
| 1,40 | 0,4192 | 1,71 | 0,4564 | 2,04 | 0,4793 | 2,66 | 0,4961 |
| 1,41 | 0,4207 | 1,72 | 0,4573 | 2,06 | 0,4803 | 2,68 | 0,4963 |
| 1,42 | 0,4222 | 1,73 | 0,4582 | 2,08 | 0,4812 | 2,70 | 0,4965 |
| 1,43 | 0,4236 | 1,74 | 0,4591 | 2,10 | 0,4821 | 2,72 | 0,4967 |
| 1,44 | 0,4251 | 1,75 | 0,4599 | 2,12 | 0,4830 | 2,74 | 0,4969 |
| 1,45 | 0,4265 | 1,76 | 0,4608 | 2,14 | 0,4838 | 2,76 | 0,4971 |
| 1,46 | 0,4279 | 1,77 | 0,4616 | 2,16 | 0,4846 | 2,78 | 0,4973 |
| 1,47 | 0,4292 | 1,78 | 0,4625 | 2,18 | 0,4854 | 2,80 | 0,4974 |
| 1,48 | 0,4306 | 1,79 | 0,4633 | 2,20 | 0,4861 | 2,82 | 0,4976 |
| 1,49 | 0,4319 | 1,80 | 0,4641 | 2,22 | 0,4868 | 2,84 | 0,4977 |
| 1,50 | 0,4332 | 1,81 | 0,4649 | 2,24 | 0,4875 | 2,86 | 0,4979 |
| 1,51 | 0,4345 | 1,82 | 0,4656 | 2,26 | 0,4881 | 2,88 | 0,4980 |
| 1,52 | 0,4357 | 1,83 | 0,4664 | 2,28 | 0,4887 | 2,90 | 0,4981 |
| 1,53 | 0,4370 | 1,84 | 0,4671 | 2,30 | 0,4893 | 2,92 | 0,4982 |
| 1,54 | 0,4382 | 1,85 | 0,4678 | 2,32 | 0,4898 | 2,94 | 0,4984 |
| 1,55 | 0,4394 | 1,86 | 0,4686 | 2,34 | 0,4904 | 2,96 | 0,4985 |
| 1,56 | 0,4406 | 1,87 | 0,4693 | 2,36 | 0,4909 | 2,98 | 0,4986 |
| 1,57 | 0,4418 | 1,88 | 0,4699 | 2,38 | 0,4913 | 3,00 | 0,49865 |
| 1,58 | 0,4429 | 1,89 | 0,4706 | 2,40 | 0,4918 | 3,20 | 0,49931 |
| 1,59 | 0,4441 | 1,90 | 0,4713 | 2,42 | 0,4922 | 3,40 | 0,49966 |
| 1,60 | 0,4452 | 1,91 | 0,4719 | 2,44 | 0,4927 | 3,60 | 0,499841 |
| 1,61 | 0,4463 | 1,92 | 0,4726 | 2,46 | 0,4931 | 3,80 | 0,499928 |
| 1,62 | 0,4474 | 1,93 | 0,4732 | 2,48 | 0,4934 | 4,00 | 0,499968 |
| 1,63 | 0,4484 | 1,94 | 0,4738 | 2,50 | 0,4938 | 4,50 | 0,499997 |
| 1,64 | 0,4495 | 1,95 | 0,4744 | 2,52 | 0,4941 | 5,00 | 0,499999 |
| 1,65 | 0,4505 | 1,96 | 0,4750 | 2,54 | 0,4945 | | |
| 1,66 | 0,4515 | 1,97 | 0,4756 | 2,56 | 0,4948 | | |
| 1,67 | 0,4525 | 1,98 | 0,4761 | 2,58 | 0,4951 | | |
| 1,68 | 0,4535 | 1,99 | 0,4767 | 2,60 | 0,4953 | | |
| 1,69 | 0,4545 | 2,00 | 0,4772 | 2,62 | 0,4956 | | |
| 1,70 | 0,4554 | 2,02 | 0,4783 | 2,64 | 0,4959 | | |

Приложение 2

Критические точки распределения F Фишера

($k_1 = N_1 - 1$ – число степеней свободы большей дисперсии,

$k_2 = N_2 - 1$ – число степеней свободы меньшей дисперсии)

| Уровень значимости $q/2 = 0,05$ | | | | | | | | | |
|---------------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| k_2/k_1 | 2 | 4 | 6 | 8 | 10 | 12 | 20 | 24 | 150 |
| 2 | 19,00 | 19,25 | 19,33 | 19,37 | 19,39 | 19,41 | 19,44 | 19,45 | 19,49 |
| 4 | 6,94 | 6,39 | 6,16 | 6,04 | 5,96 | 5,91 | 5,80 | 5,77 | 5,65 |
| 6 | 5,14 | 4,53 | 4,28 | 4,15 | 4,06 | 4,00 | 3,87 | 3,84 | 2,97 |
| 8 | 4,46 | 3,84 | 3,58 | 3,44 | 3,34 | 3,28 | 3,15 | 3,12 | 2,58 |
| 10 | 4,1 | 3,48 | 3,22 | 3,07 | 2,97 | 2,91 | 2,77 | 2,74 | 2,34 |
| 12 | 3,88 | 3,26 | 3,00 | 2,80 | 2,76 | 2,69 | 2,54 | 2,50 | 2,18 |
| 14 | 3,74 | 3,11 | 2,85 | 2,70 | 2,60 | 2,53 | 2,39 | 2,35 | 2,06 |
| 16 | 3,63 | 3,01 | 2,74 | 2,59 | 2,49 | 2,42 | 2,28 | 2,24 | 1,97 |
| 18 | 3,55 | 2,93 | 2,66 | 2,51 | 2,41 | 2,34 | 2,19 | 2,15 | 1,89 |
| 20 | 3,49 | 2,87 | 2,60 | 2,45 | 2,35 | 2,28 | 2,12 | 2,08 | 1,82 |
| 22 | 3,44 | 2,82 | 2,55 | 2,4 | 2,30 | 2,23 | 2,07 | 2,03 | 1,78 |
| 24 | 3,40 | 2,78 | 2,51 | 2,36 | 2,26 | 2,18 | 2,02 | 1,98 | 1,74 |
| 26 | 3,37 | 2,74 | 2,47 | 2,32 | 2,22 | 2,15 | 1,99 | 1,95 | 1,71 |
| 28 | 3,34 | 2,71 | 2,44 | 2,29 | 2,19 | 2,12 | 1,96 | 1,91 | 1,68 |
| 150 | 3,06 | 2,43 | 2,16 | 2,00 | 1,89 | 1,82 | 1,64 | 1,59 | 1,33 |
| Уровень значимости $q/2 = 0,01$ | | | | | | | | | |
| k_2/k_1 | 2 | 4 | 6 | 8 | 10 | 12 | 20 | 24 | 120 |
| 2 | 99,01 | 99,25 | 99,33 | 99,36 | 99,40 | 99,42 | 99,45 | 99,46 | 99,49 |
| 4 | 18,00 | 15,98 | 15,21 | 14,80 | 14,54 | 14,37 | 14,02 | 13,63 | 13,56 |
| 6 | 10,92 | 9,15 | 8,47 | 8,10 | 7,87 | 7,72 | 7,40 | 7,31 | 6,97 |
| 8 | 8,65 | 7,01 | 6,37 | 6,03 | 5,82 | 5,67 | 5,36 | 5,28 | 4,95 |
| 10 | 7,56 | 5,99 | 5,39 | 5,06 | 4,85 | 4,71 | 4,41 | 4,33 | 4,00 |
| 12 | 6,93 | 5,41 | 4,82 | 4,50 | 4,30 | 4,16 | 3,86 | 3,78 | 3,45 |
| 14 | 6,51 | 5,03 | 4,46 | 4,14 | 3,94 | 3,80 | 3,51 | 3,43 | 3,09 |
| 16 | 6,23 | 4,77 | 4,20 | 3,89 | 3,69 | 3,55 | 3,25 | 3,18 | 2,85 |
| 18 | 6,01 | 4,58 | 4,01 | 3,71 | 3,51 | 3,37 | 3,07 | 3,00 | 2,66 |
| 20 | 5,85 | 4,43 | 3,87 | 3,56 | 3,37 | 3,23 | 2,94 | 2,26 | 2,52 |
| 22 | 5,72 | 4,31 | 3,76 | 3,45 | 3,26 | 3,12 | 2,83 | 2,75 | 2,40 |
| 24 | 5,61 | 4,22 | 3,67 | 3,36 | 3,17 | 3,03 | 2,74 | 2,66 | 2,31 |
| 26 | 5,53 | 4,14 | 3,6 | 3,29 | 3,09 | 2,96 | 2,66 | 2,59 | 2,23 |
| 28 | 5,45 | 4,07 | 3,53 | 3,23 | 3,03 | 2,90 | 2,60 | 2,52 | 2,17 |
| 120 | 4,79 | 3,48 | 2,96 | 2,66 | 2,47 | 2,34 | 2,04 | 1,95 | 1,53 |

Глоссарий

Адекватность модели – качественная или количественная оценка соответствия модели моделируемому объекту.

Активное событие – смена состояния элемента под воздействием присутствующих ему внутренних свойств.

Активный элемент – элемент системы, в котором свершаются активные события.

Алгоритмический генератор (датчик) случайной величины – программный модуль, формирующий случайные числа по специально подобранной математической формуле.

Анализ (от греч. *análysis* – разложение, расчленение) – процесс определения свойств, присутствующих в системе.

Генератор (датчик) случайных чисел – аппаратное, табличное или программное (алгоритмическое) включение в компьютер, формирующее случайные числа.

Дискретный марковский процесс – процесс, при котором смена состояний происходит в определённые моменты времени.

Достоверность оценки – вероятность точности оценки.

Имитационная модель – универсальное средство исследования сложных систем, представляющее собой логико-алгоритмическое описание поведения отдельных элементов системы и правил их взаимодействия, отображающих последовательность событий, возникающих в моделируемой системе.

Имитационное моделирование – один из классов математического моделирования; метод исследования, позволяющий строить модели, описывающие процессы так, как они происходили бы в действительности.

Имитационное моделирование распределённое – выполнение единой программы имитационной модели на мультипроцессорной или мультимикрокомпьютерной системе.

Квазиравномерная случайная величина – равномерно распределённая дискретная случайная величина, формируемая в ЭВМ.

Компьютерная модель – компьютерная программа, реализующая математическую модель исследуемых процессов на компьютере.

Компьютерное моделирование – расширение возможностей математического моделирования за счёт использования последних достижений информационных технологий, внедряемых во все сферы человеческой деятельности.

Критерий эффективности – мера эффективности системы, обобщающая все свойства системы в одной оценке – значении критерия эффективности.

Марковский процесс – случайный процесс, в котором вероятность перехода системы в новое состояние зависит только от состояния системы в настоящий момент и не зависит от того, когда и каким образом система перешла в это состояние.

Машинное (компьютерное) время – временные затраты компьютера при моделировании.

Моделирование – создание или отыскание в природе объекта (модели), заменяющего в нужных аспектах реальную систему с целью проведения на нём экспериментов.

Модель – физический или абстрактный объект, адекватно отображающий исследуемую систему.

Модельное время – время, по которому организуются события в модели.

Модуль реакции – совокупность программ, имитирующих функционально самостоятельный фрагмент поведения моделируемого объекта.

Непрерывный марковский процесс – марковский процесс, при котором смена состояний происходит в случайные моменты времени.

Оптимальная система – система, которой соответствует максимальное (минимальное) значение критерия эффективности из всех возможных вариантов построения системы, удовлетворяющих заданным требованиям.

Организация системы – способ достижения поставленной цели за счёт выбора определённой структуры и функции системы. В соответствии с этим различают структурную и функциональную организацию системы.

Отклик – реакция системы (экзогенная переменная), определяемая в ходе компьютерного эксперимента.

Оценка параметра – приближённое значение определяемого параметра, сколь угодно близко приближающаяся к его истинному значению при увеличении числа реализаций модели.

Параметры – величины, описывающие первичные свойства системы и являющиеся исходными данными при решении задач анализа.

Пассивное событие – событие (смена состояния элемента), возникающее под воздействием активного события.

Пассивный элемент – элемент системы, в котором свершаются пассивные события.

План компьютерного эксперимента – оформленный результат применения специальных методов для достижения максимально возможной информативности компьютерного эксперимента с учётом ограничений на ресурсы.

Полный факторный эксперимент (ПФЭ) – компьютерный эксперимент, в котором исследуется поведение системы (процесса) при всех сочетаниях факторов и их уровней.

Показатель эффективности (качества) – мера одного свойства системы. Показатель эффективности всегда имеет *количественный* смысл.

Поток событий – совокупность событий распределённых во времени. Если событие заключается в появлении заявок, то имеем поток заявок.

Поток простейший – стационарный ординарный поток без последствия.

Поток рекуррентный – случайный поток, в котором все интервалы между событиями распределены по одному и тому же закону.

Псевдослучайная величина – случайная величина, формируемая алгоритмическим генератором (датчиком).

Процесс – последовательная смена состояний системы во времени.

Реальное (естественное) время системы – время, в котором функционирует моделируемая система.

Реализация модели – один прогон имитационной модели.

Синтез (от греч. synthesis - соединение, сочетание, составление) – процесс порождения функций и структур, удовлетворяющих требованиям, предъявляемым к эффективности системы.

Система – совокупность взаимосвязанных элементов, объединённых в одно целое для достижения некоторой цели, определяемой назначением системы.

Система моделирования – комплекс программных средств, обеспечивающих процесс создания компьютерных имитационных моделей и проведения с ними экспериментов.

Случайное число – число, порождённое процессом, исход которого непредсказуем и который потом не может быть воспроизведён надёжно.

Случайный процесс – случайная величина, значение которой зависит от неслучайного вещественного аргумента – времени.

Событие – причина, вызывающая переход из состояния в состояние.

Состояние системы – совокупность значений переменных, описывающих это состояние. Система находится в некотором состоянии, если она полностью описывается значениями переменных, которые задают это состояние.

Статистическое моделирование – метод исследования сложных систем, основанный на описании процессов функционирования отдельных элементов в их взаимосвязи с целью получения множества частных результатов, подлежащих обработке методами математической статистики для получения конечных результатов.

Стратегический план эксперимента – план, определяющий организацию компьютерного эксперимента, обеспечивающий приемлемую информативность эксперимента с учётом ограничений на ресурсы.

Структура системы – перечень элементов, входящих в состав системы, и связей между ними.

Тактический план эксперимента – план, обеспечивающий достижение заданной точности характеристик исследуемого процесса.

Точность оценки – степень приближения (абсолютная или относительная) оценки искомой характеристики к её истинному значению.

Уровень фактора – численное значение фактора.

Факторы – входные переменные (эндогенные переменные), влияющие на функционирование моделируемой системы.

Функция системы – правило достижения поставленной цели, описывающее поведение системы и направленное на получение результатов, предписанных назначением системы.

Показатели – величины, описывающие вторичные свойства системы и определяемые в процессе решения задач анализа как функция параметров, то есть эти величины являются вторичными по отношению к параметрам.

Эффективность – степень соответствия системы своему назначению. Эффективность систем обычно оценивается набором показателей эффективности.

GPSS World – система моделирования – программное средство имитационного моделирования дискретных и непрерывных процессов.

AnyLogic – система моделирования – программное средство, поддерживающая три метода построения модели: системной динамики, дискретно-событийный, агентный.