

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«ОРЕНБУРГСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ»**

**МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ ДЛЯ ОБУЧАЮЩИХСЯ
ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ**

**Б1.В.ДВ.04.02 МАТЕМАТИЧЕСКАЯ ТЕОРИЯ ПЛАНИРОВАНИЯ
ЭКСПЕРИМЕНТОВ**

Направление подготовки (специальность):

27.03.04 Управление в технических системах

Профиль образовательной программы:

Интеллектуальные системы обработки информации и управления

Форма обучения: очная

СОДЕРЖАНИЕ

1. Конспект лекций	3
1.1 Лекция № 1 Детерминистический и эмпирический принципы изучения действительности.	
Выбор предмета исследования. Основные этапы экспериментального исследования.	
Классификации методов исследования.....	3
1.2 Лекция № 2 Основные понятия теории подобия. Использование аналогий в практике математического и физического моделирования. Изоморфизм и гомоморфизм.....	11
1.3 Лекция № 3 Основы планирования эксперимента. Пассивный и активный эксперимент.	
Факторы, общая характеристика, функция отклика.....	16
1.4 Лекция № 4 Планы факторного эксперимента. Полный факторный эксперимент.	
Дробный факторный эксперимент.....	21
1.5 Лекция № 5 Теоретические основы обработки экспериментальных данных.....	30
1.6 Лекция № 6 Методы стохастического описания и анализа особенностей процессов управления в технических системах.....	50
1.7 Лекция № 7 Техника экспериментальных измерений. Основные положения теории погрешностей. Адекватность моделей.....	58
1.8 Лекция № 8 Методы оптимизации параметров процессов управления в технических системах.....	65
2. Методические указания по проведению практических занятий	72
2.1 Практическое занятие № ПЗ-1 Проблемы современной фундаментальной науки.	
Основные этапы экспериментального исследования. Классификации методов исследования.....	72
2.2 Практическое занятие № ПЗ-2 Математическое моделирование в инженерных исследованиях.....	72
2.3 Практическое занятие № ПЗ-3 Факторы, методы отбора, общая характеристика, функция отклика.....	73
2.4 Практическое занятие № ПЗ-4 Планы факторного эксперимента	74
2.5 Практическое занятие № ПЗ-5 Теоретические основы обработки экспериментальных данных.....	74
2.6 Практическое занятие № ПЗ-6 Методы стохастического описания и анализа особенностей процессов управления в технических системах.....	79
2.7 Практическое занятие № ПЗ-7 Техника экспериментальных измерений. Основные положения теории погрешностей.....	88

1. КОНСПЕКТ ЛЕКЦИЙ

1. 1 Лекция №1 (2 часа).

Тема: «Детерминистический и эмпирический принципы изучения действительности. Выбор предмета исследования. Основные этапы экспериментального исследования. Классификации методов исследования»

1.1.1 Вопросы лекции:

1. Основные тенденции в развитии современной науки. Детерминистический и эмпирический принципы изучения действительности
2. Организация научно-исследовательской работы
3. Проблемы современной фундаментальной науки

1.1.2 Краткое содержание вопросов:

1. Основные тенденции в развитии современной науки. Детерминистический и эмпирический принципы изучения действительности.

Наука – это непрерывно развивающаяся система знаний объективных законов природы, общества и мышления, получаемых и превращаемых в непосредственную производительную силу общества в результате социально-экономической деятельности.

Это синтез организованной особым образом познавательной деятельности и ее результатов. Под **особым образом познавательной деятельности** понимается методологические и мировоззренческие принципы, обеспечивающие научный подход к выбору, постановке и реализации исследования. Термин наука применяется также и для обозначения отдельной области знаний.

Основная цель науки – познание объективного мира (теоретическое отражение действительности) и воздействие на окружающую среду с целью получения полезных обществу результатов.

Наука поддерживается и развивается в результате исследовательской деятельности общества.

Научное исследование – это форма существования и развития науки. Структуру организации научных исследований целесообразно представить в виде четырех компонентов (рис.1.):

- первый – общие вопросы научных исследований (теория, методология и методы);
- второй – процессы научных исследований (формы, методы и средства познания);
- третий – методика научных исследований (выбор конкретных форм, методов и средств, эффективных для соответствующей области науки



- или отрасли профессиональной деятельности);
- четвертый – технология научных исследований

Детерминизм в современной науке определяется как учение о всеобщей, закономерной связи явлений и процессов окружающего мира. Наличие таких связей является доказательством материального единства мира и существования в мире общих закономерностей. Очень часто детерминизм отождествляется с причинностью, но такой взгляд нельзя считать правильным хотя бы потому, что причинность выступает как одна из форм проявления детерминизма.

Законы, с которыми имеет дело классическая механика, имеют универсальный характер, т. е. они относятся ко всем без исключения изучаемым объектам природы. Отличительная особенность такого рода законов состоит в том, что предсказания, полученные на их основе, имеют достоверный и однозначный характер. Наиболее ярко они проявились после того как на основе закона всемирного тяготения, изложенного И. Ньютона в 1671 г. в "Математических началах натуральной философии" и законов механики возникла небесная механика. На основе законов небесной механики были вычислены отклонения в движении Урана, вызванные возмущающим влиянием неизвестной тогда планеты. Определив величину возмущения, независимо друг от друга по законам механики положение неизвестной планеты рассчитали Д. Адамс и У. Леверье. Всего на угловом расстоянии в 1° от рассчитанного ими положения И. Галле обнаружил планету Нептун. Открытие Нептуна, сделанное на кончике пера, как отметил Ф. Энгельс, блестяще подтвердило справедливость законов небесной механики и наличие в природе однозначных причинных связей. Это позволило французскому механику П. Лапласу сказать: дайте мне начальные условия, и я, с помощью законов механики, предскажу дальнейшее развитие событий. Это вошло в историю как лапласовский, или механистический детерминизм, который допускает однозначные причинные связи в явлениях природы.

Наряду с ними в науке с середины XIX в. стали все шире применяться законы другого типа. Их предсказания не являются однозначными, а являются только вероятностными. Вероятностными они называются потому, что заключения, основанные на них, не следует логически из имеющейся информации, а потому не являются достоверными и однозначными. Информация при этом носит статистический характер, законы, выражающие эти процессы, называют статистическими законами, и этот термин получил в науке большое распространение.

В классической науке статистические законы не признавали подлинными законами, так как ученые в прошлом предполагали, что за ними должны стоять такие же универсальные законы, как закон всемирного тяготения Ньютона. Он считался образцом детерминистического закона, поскольку обеспечивает точные и достоверные предсказания приливов и отливов, солнечных и лунных затмений и других явлений природы. Статистические же законы признавались в качестве удобных вспомогательных средств исследования, дающих возможность представить в компактной и удобной форме всю имеющуюся информацию о каком-либо предмете исследования. Подлинными законами считались именно детерминистические законы, обеспечивающие точные и достоверные предсказания. Эта терминология сохранилась до настоящего времени, когда статистические, или вероятностные, законы квалифицируются как индетерминистические, с чем вряд ли можно согласиться.

Отношение к статистическим законам принципиально изменилось после открытия законов квантовой механики, предсказания которых имеют существенно вероятностный характер.

Таким образом, исторически детерминизм выступает в двух следующих формах:

- ❖ лапласовский, или механистический, детерминизм, в основе которого лежат универсальные законы классической физики;
- ❖ вероятностный детерминизм, опирающийся на статистические законы и законы квантовой физики.

В динамических теориях явления природы подчиняются однозначным (динамическим) закономерностям, а статистические теории основаны на объяснении процессов вероятностными (статистическими) закономерностями. Таким образом, XIX столетие получается столетием динамических теорий; XX столетие — столетием статистических теорий. Значит, динамические теории соответствовали первому этапу в процессе познания природы человеком, тогда как на следующем этапе главную роль стали играть статистические теории.

В современной концепции детерминизма органически сочетаются необходимость и случайность. Признание самостоятельности статистических, или вероятностных, законов, отображающих существование случайных событий в мире, дополняет прежнюю картину строго детерминистического мира. В результате в новой современной картине мира необходимость и случайность выступают как взаимосвязанные и дополняющие друг друга аспекты объяснения окружающего мира.

Рассматривая проблему соотношения между динамическими и статистическими закономерностями, современная наука исходит из концепции примата статистических закономерностей. Не только динамические, но и статистические законы выражают объективные причинно-следственные связи. Более того, именно статистические закономерности являются фундаментальными, более глубокими по сравнению с динамическими закономерностями, они ярче выражают указанные связи.

Современную концепцию детерминизма можно сформулировать следующим образом: динамические законы представляют собой первый, низший этап в процессе познания окружающего мира; статистические же законы более совершенно отображают объективные связи в природе: они являются следующим, более высоким этапом познания. В качестве примера динамических законов можно назвать закон Ома, выражающий зависимость сопротивления от его состава, площади поперечного сечения и длины. Этот закон охватывает множество различных проводников и действует в каждом отдельном проводнике, входящем в это множество.

Статистический характер имеет, например, взаимосвязь изменений давления газа и его объема при постоянной температуре, выявленная Бойлем и Мариоттом. Данная закономерность имеет место лишь в массе хаотически перемещающихся молекул, составляющих тот или иной объем газа. Статистическими являются законы квантовой механики, касающиеся движения микрочастиц; они не в состоянии определить движение каждой отдельной частицы, но определяют движение группы, того или иного множества.

В отличие от динамических законов, статистические законы не позволяют точно предсказать наступление или ненаступление того или иного конкретного явления, направление и характер изменения тех или иных его характеристик. На основе статистических закономерностей можно определить лишь степень вероятности

возникновения или изменения соответствующего явления. Динамические теории не противостоят статистическим теориям, а включаются в рамки последних как предельный случай. Таким образом, согласно современной научной концепции, можно говорить о всеобщности, универсальности вероятностного подхода. Это означает, в частности, что деление фундаментальных теорий на динамические и статистические является, строго говоря, условным. Фактически все фундаментальные теории должны рассматриваться как статистические. Например, классическую механику с полным основанием следует считать статистической теорией, так как лежащий в ее основе принцип наименьшего действия имеет вероятностную природу, потому что, согласно принципу минимума энергии, состояние с наименьшей энергией оказывается наиболее вероятным.

Наука рассматривает два основных типа причинно-следственных связей и соответственно два типа закономерностей — динамические и статистические. Изучение истории возникновения фундаментальных теорий позволяет сделать вывод, что динамические теории соответствовали первому этапу в процессе познания природы человеком, тогда как на следующем этапе главную роль стали играть статистические теории.

2. Организация научно-исследовательской работы

Научная теория — это высшая форма организации теоретического знания, представляющая собой совокупность объединенных в единую систему основных элементов теории (подтвержденных гипотез, понятий, суждений) в соответствующей отрасли (в данном случае в информатике). Критерием истинности теории является ее практическое подтверждение.

Основой любой науки и, в частности, науковедения является **методология**, которая представляет собой учение о структуре, логической организации, методах и средствах деятельности.

В научной литературе под **методологией** обычно понимается, прежде всего, система научного познания, т.е. учение о принципах построения, формах и способах научно-познавательной деятельности.

Методология может быть **специально-научная и философская**.

Специально-научная методология разделяется на несколько уровней: общенаучные методологические концепции и направления, методология отдельных специальных наук, методика и технология исследований. Философская методология определяет систему философских знаний. Частным способом реализации методологии на практике является метод, как система действий в различных видах человеческой деятельности направленных на достижение поставленной задачи.

Научный метод — это система правил и предписаний, направляющих человеческую деятельность (производственную, политическую, культурную, научную, образовательную и т.д.) к достижению поставленной цели.

Если методология — это стратегия научных исследований, обеспечивающих достижение цели, сформулированной в гипотезе предполагаемых научных результатов (генеральный путь познания), то метод — это тактика, показывающая как лучше всего идти этим путем.

Метод (гр. *methodos*) — 1) способ познания, исследования явлений природы и общественной жизни; 2) прием, способ и образ действий.

Метод — путь исследования, способ достижения какой-либо цели, решения конкретных задач. Это совокупность подходов, приемов, операций практического или теоретического освоения действительности.

Из определения метода вытекает, что существуют **две большие группы методов**: познания (исследования) и практического действия (преобразовательные методы) (рис.2).

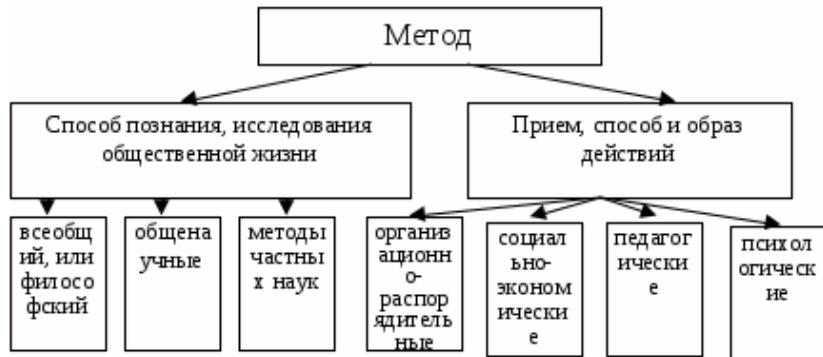


Рис. 2 –Группы научных методов

1) Методы исследования — приемы, процедуры и операции эмпирического и теоретического познания и изучения явлений действительности. С помощью этой группы методов получают достоверные сведения, используемые для построения научных теорий и выработки практических рекомендаций. Система методов исследования определяется исходной концепцией исследователя: его представлениями о сущности и структуре изучаемого, общей методологической ориентации, целей и задач конкретного исследования. Методы подразделяются на следующие:

- всеобщий, или философский, общенаучные и методы частных наук;
- констатирующие и преобразующие;
- эмпирические и теоретические;
- качественные и количественные;
- содержательные и формальные;
- методы сбора эмпирических данных, проверки и опровержения гипотез и теории;
- описания, объяснения и прогноза;
- обработки результатов исследования.

Всеобщий, или философский метод — всеобщий метод материалистической диалектики.

К **общенаучным методам** относятся:

- Наблюдение – это способ познания объективного мира, основанный на непосредственном восприятии предметов и явлений при помощи органов чувств без вмешательства в процесс со стороны исследователя.
- Сравнение - это установление различия между объектами материального мира или нахождение в них общего; осуществляется как при помощи органов чувств, так и при помощи специальных устройств.
- Счет – это нахождение числа, определяющего количественное соотношение однотипных объектов или их параметров, характеризующих те или иные свойства.
- Измерение – это физический процесс определения численного значения некоторой величины путем сравнения ее с эталоном.

- Эксперимент – одна из сфер человеческого практики, в которой подвергается проверке истинность выдвигаемых гипотез или выявляются закономерности объективного мира.
- Обобщение – определение общего понятия, в котором находит отражение главное, основное, характеризующее объекты данного класса.
- Абстрагирование – это мысленное отвлечение от несущественных свойств, связей, отношений предметов и выделение нескольких сторон, интересующих исследователя.
- Формализация – отображение объекта или явления в знаковой форме какого-либо искусственного языка (математики, химии и т.д.).
- Аксиоматический метод – способ построения научной теории, при котором некоторые утверждения принимаются без доказательств.
- Анализ – метод познания при помощи расчленения или разложения предметов исследования на составные части.
- Синтез – соединение отдельных сторон предмета в единое целое.
- Индукция – умозаключение от фактов к некоторой гипотезе (общему утверждению).
- Дедукция – умозаключение, в котором вывод о некотором элементе множества делается на основании знания общих свойств всего множества.
- Аналогия – метод, посредством которого достигается знание о предметах и явлениях на основании того, что они имеют сходство с другими.
- Гипотетический метод познания предполагает разработку научной гипотезы на основе изучения физической, химической и т.п., сущности исследуемого явления, формулирование гипотезы, составление расчетной схемы алгоритма (модели), ее изучение, анализ, разработка теоретических положений.
- Исторический метод познания предполагает исследование возникновения, формирования и развития объектов в хронологической последовательности.
- Идеализация - это мысленное конструирование объектов, которые практически неосуществимы.
- Системные методы: исследование операций, теория массового обслуживания, теория управления, теория множеств и др.

Методы частных наук — специфические способы познания и преобразования отдельных областей реального мира, присущие той или иной конкретной системе знаний (социология — социометрия; психология — психодиагностика).

2) Методы как прием, способ и образ действий (методы практической деятельности) включают в себя способы воздействия, совокупность приемов, операций и процедур подготовки и принятия решения, организации его выполнения.

Для выбора методов на каждом этапе необходимо знать общие и конкретные возможности каждого метода, его место в системе исследовательских процедур. Задача исследователя состоит в том, чтобы для каждого этапа исследования определить оптимальный комплекс методов.

Разнообразные **методы** научного познания условно подразделяются на ряд **уровней**: **эмпирический, экспериментально-теоретический, теоретический и метатеоретический**.

Методы эмпирического уровня: *наблюдение, сравнение, счет, измерение, анкетный опрос, собеседование, тесты, метод проб и ошибок и другие.*

Методы экспериментально-теоретического уровня: *эксперимент, анализ и синтез, индукция и дедукция, моделирование, гипотетический, исторический и логический методы.*

Методы теоретического уровня: абстрагирование, идеализация, формализация, анализ и синтез, индукция и дедукция, аксиоматика, обобщение и другие.

К методам метатеоретического уровня относятся диалектический и метод системного анализа.

3. Проблемы современной фундаментальной науки

Российская академия наук является независимой некоммерческой организацией, имеющей государственный статус. Главным образом РАН занимается проведением фундаментальных исследований в различных областях знаний. При этом при РАН существуют фонды, содействующие реализации наиболее перспективных научных разработок. Это Российский фонд фундаментальных исследований (РФФИ), Российский гуманитарный научный фонд (РГНФ), Фонд содействия развитию малых форм предприятий в научно-технической сфере. В условиях необходимости сохранения целостности государства и стабилизации экономики в первой половине 90-х годов XX века создание этих фондов явилось единственной мерой, предпринятой для поддержки проводимых научных исследований и для содействия внедрению их результатов.

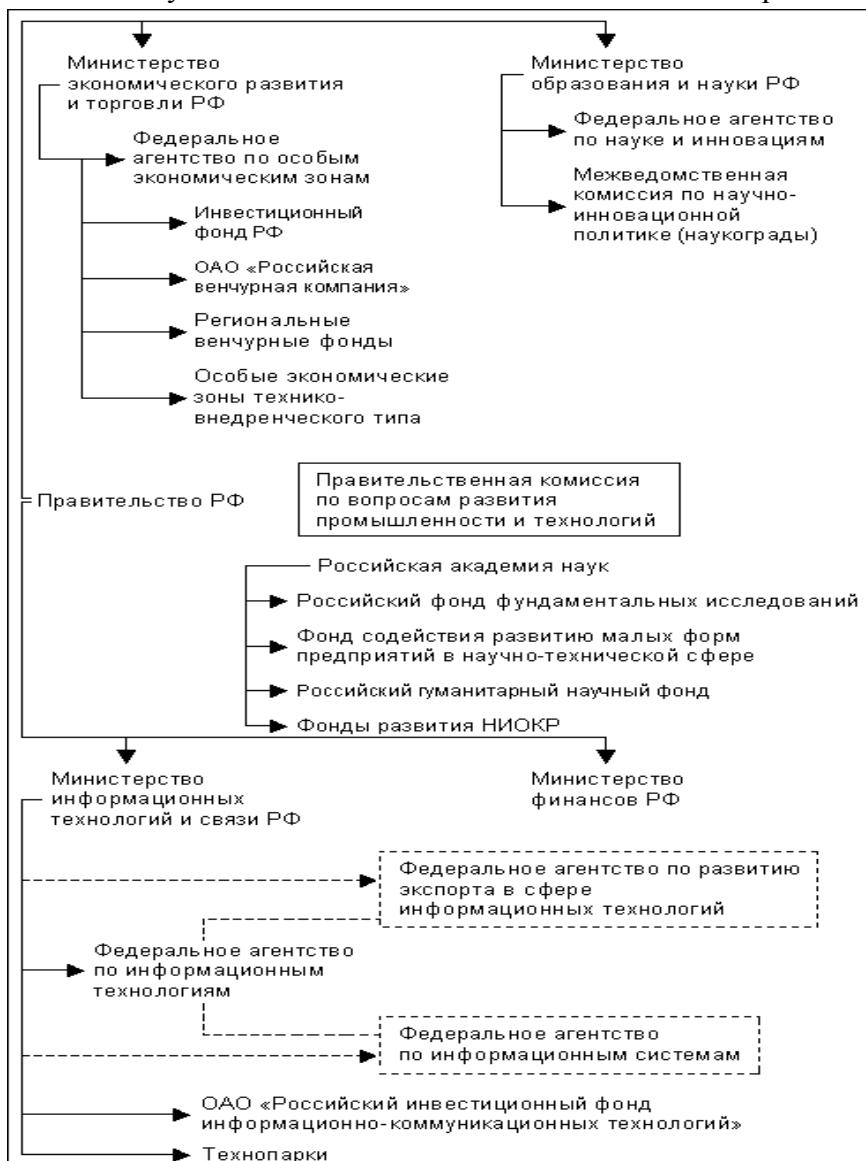


Рис. 3 – Организационная структура науки в России

РФФИ был образован Указом Президента РФ от 27 апреля 1992 года № 426 «О неотложных мерах по сохранению научно-технического потенциала РФ». Фонд

«финансируется из государственного бюджета и поддерживает ученых на безвозвратной основе». Одним из важных направлений в работе РГНФ является создание баз данных по научным разработкам и предоставление информации о них заинтересованным сторонам. РГНФ выделился из состава РФФИ в 1994 году. Главные задачи фонда — «поддержка гуманитарных научных исследований и распространение гуманитарных научных знаний об обществе». Финансируется РГНФ за счет ассигнований в размере 0,5% от средств из федерального бюджета, направляемых на развитие науки. Фонд содействия развитию малых форм предприятий в научно-технической сфере образован 3 февраля 1994 года. Начиная с 2001 года, его размер финансирования вырос с 0,5 до 1,5% средств, направляемых на науку из федерального бюджета. Фонд оказывает финансовую поддержку высокоэффективным научным проектам, разрабатываемым малыми предприятиями. Финансирование проектов осуществляется на паритетной основе с малыми инновационными предприятиями. Отбор проектов, поддерживаемых фондами РАН, проводится на конкурсной основе.

Другим не менее важным органом сферы науки и инноваций ввиду последних изменений является Министерство экономического развития и торговли (МЭРТ), которое сосредотачивает внимание на этапе внедрения разработок, осуществляя инвестирование в инновационные проекты. В рамках МЭРТ недавно образовано Федеральное агентство по управлению особыми экономическими зонами, которое также занимается Инвестиционным фондом РФ. Среди уже созданных и создаваемых типов особых экономических зон (ОЭЗ) в рамках рассматриваемой нами темы важно выделить технико-внедренческие ОЭЗ. К настоящему моменту созданы четыре таких зоны в различных субъектах РФ, имеющие свою специализацию: в Дубне — исследования в области ядерных технологий; в Зеленограде — микроэлектроника; в Санкт-Петербурге — информационные технологии; в Томске — новые материалы.

Целью создания ОЭЗ технико-внедренческого типа является государственная поддержка инновационных предприятий путем предоставления резидентам ОЭЗ налоговых льгот и упрощения таможенного режима. При этом государство берет на себя обязательство по строительству инфраструктуры ОЭЗ. Порядок финансирования создания ОЭЗ устанавливается Соглашением между Правительством РФ в лице МЭРТ, субъектом РФ и администрацией города, на территории которого создана ОЭЗ. Необходимо отметить, что срок действия ОЭЗ составляет 20 лет. Основное требование, которое предъявляется к компаниям, которые желают стать резидентами технико-внедренческой ОЭЗ, — технико-внедренческий характер их деятельности на территории такой ОЭЗ.

Одна из отраслей, на которую Правительство делает ставку, создавая «новую» экономику, — отрасль информационных технологий. Еще одним шагом государства для реализации разработок ИТ-компаний стала одобренная Правительством государственная программа «Создание в Российской Федерации технопарков в сфере высоких технологий». Действующие до настоящего времени технопарки были созданы в разных отраслях экономики благодаря частным инициативам. Важнейшим условием реализации эффективной государственной научно-технической политики является концентрация научного потенциала, финансовых и материально-технических ресурсов на приоритетных направлениях развития науки и техники.

Под приоритетными направлениями развития науки и техники понимаются основные области исследований и разработок, реализация которых должна обеспечить значительный вклад в социально-экономическое и научно-техническое развитие страны и в достижение за счет этого национальных социально-экономических целей.

В каждом из приоритетных направлений развития науки и техники можно выделить некоторую совокупность критических технологий. Под критическими технологиями понимаются такие технологии, которые носят межотраслевой характер, создают существенные предпосылки для развития многих технологических областей или направлений исследований и разработок и дают в совокупности главный вклад в решение ключевых проблем реализации приоритетных направлений развития науки и техники.

1. 2 Лекция №2 (2 часа).

Тема: «Основные понятия теории подобия. Использование аналогий в практике математического и физического моделирования. Изоморфизм и гомоморфизм»

1.2.1 Вопросы лекции:

1. Научная теория как высшая форма организации знания
2. Основные понятия теории подобия. Использование аналогий в практике математического и физического моделирования
3. Изоморфизм и гомоморфизм

1.2.2 Краткое содержание вопросов:

1. Научная теория как высшая форма организации знания

Научная теория – это высшая форма организации теоретического знания, представляющая собой совокупность объединенных в единую систему основных элементов теории (подтвержденных гипотез, понятий, суждений) в соответствующей отрасли (в данном случае в информатике). Критерием истинности теории является ее практическое подтверждение.

Основой любой науки и, в частности, науковедения является **методология**, которая представляет собой учение о структуре, логической организации, методах и средствах деятельности.

В научной литературе под **методологией** обычно понимается, прежде всего, система научного познания, т.е. учение о принципах построения, формах и способах научно-познавательной деятельности.

Методология может быть **специально-научная и философская**.

Специально-научная методология разделяется на несколько уровней: общенаучные методологические концепции и направления, методология отдельных специальных наук, методика и технология исследований. Философская методология определяет систему философских знаний. Частным способом реализации методологии на практике является метод, как система действий в различных видах человеческой деятельности направленных на достижение поставленной задачи.

Научный метод – это система правил и предписаний, направляющих человеческую деятельность (производственную, политическую, культурную, научную, образовательную и т.д.) к достижению поставленной цели.

Если методология – это стратегия научных исследований, обеспечивающих достижение цели, сформулированной в гипотезе предполагаемых научных результатов (генеральный путь познания), то метод – это тактика, показывающая как лучше всего идти этим путем.

Творчество – мышление в его высшей форме, выходящее за пределы известного, а также деятельность, порождающая нечто качественно новое.

В частности, *научное творчество* связано с познанием окружающего мира. *Научно-техническое творчество* имеет прикладные цели и направлено на удовлетворение практических потребностей человека.

Одной из проблем творчества является его мотивационная структура. **Мотивации** (побуждения) связаны с потребностями, которые делятся на три группы: *биологические, социальные и идеальные (подсознательные)*.

Наиболее важным для творчества видом мышления является воображение.

Творческая личность обладает рядом особенностей и прежде всего умением сосредоточить внимание и долго удерживать его на каком-либо вопросе или проблеме.

Общая схема решения научно-технических задач:

- анализ систем задач и выбор конкретной задачи;
- анализ технической системы и разработка ее модели;
- анализ и формулировка условий технической задачи;
- анализ и формулировка условий изобретательской задачи;
- поиск идей решения (принципа действия);
- синтез нового технического решения.

Цель научного исследования – всестороннее, достоверное изучение объекта, процесса или явления; их структуры, связей и отношений на основе разработанных в науке принципов и методов познания, а также получение и внедрение в производство (практику) полезных для человека результатов.

Любое научное исследование имеет свой *объект и предмет*. **Объектом** научного исследования является материальная или идеальная система. **Предмет** – это структура системы, закономерности взаимодействия элементов внутри системы, закономерности развития, различные свойства, качества и т.д.

Научные исследования классифицируются по видам связи с производством и степени важности для него; целевому назначению; источникам финансирования и длительности ведения.

Каждую НИР можно отнести к определенному направлению. **Под научным направлением** понимается наука или комплекс наук, в области которых ведутся исследования (например, техническое, социальное и др.).

Структурными единицами научного направления являются *комплексные проблемы, темы и научные вопросы*.

Проблема – это совокупность сложных теоретических и практических задач, решения которых назрели в обществе (противоречие между знанием и незнанием). Она

возникает тогда, когда человеческая практика встречает затруднения или даже наталкивается на «невозможность» достижения цели.

Тема научного исследования является составной частью проблемы. В результате исследований по теме получают ответы на определённый круг научных вопросов, охватывающих часть проблемы. **Под научными вопросами** понимается мелкие научные задачи, относящиеся к конкретной теме научного исследования.

Выбор направления, проблемы, темы научного исследования и постановка научных вопросов является чрезвычайно ответственной задачей.

При выборе проблемы и темы научного исследования вначале на основе анализа противоречий исследуемого направления формулируется сама проблема и определяются в общих чертах ожидаемые результаты, затем разрабатывается структура проблемы, выделяются темы, вопросы, исполнители, устанавливается их актуальность.

Выбору темы должно предшествовать тщательное ознакомление с отечественными и зарубежными литературными источниками данной и смежной специальностей.

К **процессам научных исследований** относят формы, средства и методы познания, совокупность которых составляет методику исследований конкретной научной области знаний, представляющий собой один из уровней специальной научной методологии.

Научные исследования начинаются с постановки проблемы на основе обнаружения имеющихся противоречий между потребностью научных знаний об объекте и фактическими знаниями об объекте (процессе, явлении) которыми располагает наука на данный период ее развития.

Постановка проблемы определяет выбор темы исследования, уточняет ее название и обеспечивает обоснование актуальности разработки.

Для уточнения задач исследования осуществляется информационный поиск и также проводится научный поиск, обеспечивающий получение научных результатов.

Решающее значение для научных исследований имеют интеллектуальные способности исследователя, его научное мировоззрение, широта научных знаний, системное мышление, ассоциативное восприятие, информационная культура, творческая активность, толерантность. Научные работники должны хорошо владеть психологией научной работы и грамотной организацией научных исследований.

Таким образом, что процесс научных исследований состоит из четырех последовательных и взаимосвязанных этапов (подпроцессов).

2. Основные понятия теории подобия. Использование аналогий в практике математического и физического моделирования

Выявление общего, существенного, присущего всем системам определенного рода производится наиболее общим приемом — математическим моделированием. При математическом моделировании систем наиболее ярко проявляется эффективность единства качественных и количественных методов исследования, характеризующая магистральный путь развития современного научного познания.

Всякая сложная система, модель которой мы создаем, при своем функционировании подчиняется определенным законам — физическим, химическим, биологическим и др. Рассматриваются такие системы, для которых знание законов предполагает известные количественные соотношения, связывающие те или иные характеристики моделируемой системы. Модель создается для ответа на множество

вопросов о моделируемом объекте. Интересуясь некоторыми аспектами функционирующей системы, изучают ее с определенных точек зрения. Направления изучения системы в значительной степени и определяет выбор модели. Опишем процесс построения математической модели сложной системы. Его можно представить состоящим из следующих этапов:

Формулируются основные вопросы о поведении системы, ответы на которые мы хотим получить с помощью модели.

Из множества законов, управляющих поведением системы, учитываются те, влияние которых существенно при поиске ответов на поставленные вопросы.

В дополнение к этим законам, если необходимо, для системы в целом или отдельных ее частей формулируются определенные гипотезы о функционировании. Как правило, эти гипотезы правдоподобны в том смысле, что могут быть приведены некоторые теоретические доводы в пользу их принятия.

Гипотезы, так же как и законы, выражаются в форме определенных математических соотношений, которые объединяются в некоторое формальное описание модели.

На этом заканчивается процесс построения математической модели. Далее следует процесс исследования этих соотношений с помощью аналитических и вычислительных методов, приводящий в конечном итоге к отысканию ответов на предъявляемые модели вопросы. Разрабатывается или используется созданный ранее алгоритм для анализа этой модели. Если модель и алгоритм не слишком сложны, то может оказаться возможным аналитическое исследование модели. В противном случае составляется программа, реализующая этот алгоритм на ЭВМ. После выполнения расчетов по модели на ЭВМ их результаты обязательно сравниваются с фактической информацией из соответствующей предметной области. Это сравнение необходимо для того, чтобы убедиться в адекватности модели, в том, что модельным расчетам можно верить, их можно использовать.

Если модель хороша, то ответы, найденные с ее помощью, как правило, бывают весьма близки к ответам на те же вопросы о моделируемой системе. Более того, в этом случае зачастую с помощью модели удается ответить и на некоторые ранее не ставившиеся вопросы, расширить круг представлений о реальной системе. Если же модель плоха, т. е. недостаточно адекватно описывает систему с точки зрения задаваемых ей вопросов, то она подлежит дальнейшему улучшению или замене. Возможны также ошибки в алгоритме, в программе для ЭВМ. Такие повторные просмотры продолжаются до тех пор, пока результаты расчетов не удовлетворят исследователя. Теперь модель готова к использованию. Критерием адекватности модели служит практика, которая и определяет, когда может закончиться процесс улучшения модели. Итак, ни ЭВМ, ни математическая модель, ни алгоритм для ее исследования порознь не могут решить достаточно сложную исходную задачу. Но вместе они представляют ту силу, которая позволяет познавать окружающий мир, управлять им в интересах человечества.

Достоинствами метода математического моделирования является то, что модель представляет собой формализованную запись тех или иных законов природы, управляющих функционированием системы. Однако определенные трудности возникают при попытке построения математической модели очень сложной системы.

Существуют различные модели, используемые для описания сложных систем,

такие как:

- ❖ дескриптивные (описательные), описывающие происходящие в системе процессы;
- ❖ оптимизационные, управляющие процессом, т. е. принимающие те или иные решения;
- ❖ многокритериальные, рассматривающие систему по многим критериям;
- ❖ игровые, пригодные для исследования и рассматривающие конфликтные ситуации;
- ❖ имитационные, максимально использующие имеющуюся информацию о поведении системы.

Теория подобия — метод математического моделирования, основанный на переходе от обычных физических величин, влияющих на моделируемую систему, к обобщённым величинам комплексного типа, составленным из исходных физических величин, но в определённых сочетаниях, зависящих от конкретной природы исследуемого процесса. Комплексный характер этих величин имеет глубокий физический смысл отражения взаимодействия различных влияний. Теория подобия изучает методы построения и применения этих переменных и применяется в тех случаях математического моделирования, когда аналитическое решение математических задач моделирования невозможно из-за сложности и требований к точности. Теория подобия применяется в этих случаях для синтеза соотношений, получаемых на основе физического механизма изучаемого процесса и данных численного решения или эксперимента.

3. Изоморфизм и гомоморфизм

ИЗОМОРФИЗМ и ГОМОМОРФИЗМ (греч. *isos* — одинаковый, *homoios* — подобный и *morphe* — форма) — понятия, характеризующие соответствие между структурами объектов. Две системы, рассматриваемые отвлеченно от природы составляющих их элементов, являются изоморфными друг другу, если каждому элементу первой системы соответствует лишь один элемент второй и каждой связи в одной системе соответствует связь в другой и обратно. Такое взаимооднозначное соответствие называется ИЗОМОРФИЗМ. Полный ИЗОМОРФИЗМ может быть лишь между абстрактными, идеализированными объектами, напр., соответствие между геометрической фигурой и ее аналитическим выражением в виде формулы. ИЗОМОРФИЗМ связан не со всеми, а лишь с некоторыми фиксированными в познавательном акте свойствами и отношениями сравниваемых объектов, которые в других своих отношениях могут отличаться.

ГОМОМОРФИЗМ отличается от ИЗОМОРФИЗМА тем, что соответствие объектов (систем) однозначно лишь в одну сторону. Поэтому ГОМОФОРМНЫЙ образ есть неполное, приближенное отображение структуры оригинала. Таково, напр., отношение между картиной и местностью, между грамзаписью и ее оригиналом — звуковыми колебаниями воздушной среды. Понятия ИЗОМОРФИЗМ и ГОМОМОРФИЗМ широко применяются в математической логике и кибернетике.

Математическая практика показывает, что даже такое «неглубокое» сходство систем, как изоморфизм, т.е. одинаковость структуры, может оказаться достаточным, чтобы выявить и перенести на другие системы весьма глубокие системные свойства. Здесь уже возникает изоморфизм в системах знаний об изучаемых системах: изоморфизм понятий, утверждений, теорий. При этом «одинаковые», соответствующие

друг другу элементы имеют совершенно различный смысл и исходных системах. В таких случаях иногда образно говорят о более или менее «глубоком» изоморфизме систем. При этом изоморфизм свойств систем, знаний о них является теоретическим следствием изоморфизма систем в классическом узком смысле.

Использование понятия изоморфизма в более широком, размытом смысле, при сопоставлении сложных объектов, процессов и явлений и установлении их аналогий часто не опирается на выявление структуры объектов, их элементов и системообразующих отношений. Иными словами, суть аналогии или одинаковости структуры *не выясняется, и предполагается интуитивно понятной*. Идеальным примером изоморфных систем в широком, размытом смысле считается негатив и позитив фотоснимка, или, например, речь и ее запись на магнитофонной пленке или компактном диске. Техническое устройство и его чертеж на бумаге также находятся в изоморфном соответствии.

Условное понятие «степень изоморфизма» можно наглядно продемонстрировать на примере глубокой аналогии между различными видами колебаний - механическими и акустическими, что и явилось основой создания общей теории колебаний. Говоря о более или менее глубоком изоморфизме, «степени изоморфизма», имеют в виду большее или меньшее число аналогичных свойств у сопоставляемых систем.

В тех случаях, когда понятие изоморфизма используется в широком смысле, без прояснения того, в чем именно состоит аналогии, то по существу нельзя найти обоснование, которое обеспечивало бы перенос свойств известных систем на новые, менее изученные системы. Перенос на новые системы в большей степени играет роль предвидения, чем обоснования. «Степень изоморфизма» между системами условно определяется количеством «совпадающих элементов» и в разных случаях может быть весьма различной.

1. 3 Лекция №3 (2 часа).

Тема: «Основы планирования эксперимента. Пассивный и активный эксперимент. Факторы, общая характеристика, функция отклика»

1.3.1 Вопросы лекции:

1. Основы планирования эксперимента. Пассивный и активный эксперимент
2. Факторы, общая характеристика, функция отклика

1.3.2 Краткое содержание вопросов:

1. Основы планирования эксперимента. Пассивный и активный эксперимент

Эксперимент – одна из сфер человеческого практики, в которой подвергается проверке истинность выдвигаемых гипотез или выявляются закономерности объективного мира.

Цель планирования эксперимента – получение максимума информации о свойствах исследуемого объекта при минимуме опытов.

Такой подход обусловлен высокой стоимостью экспериментов, как физических, так и вычислительных, и вместе с тем необходимостью построения адекватной модели. Планируют как **активный**, так и **пассивный** эксперимент.

К основным преимуществам активного эксперимента можно отнести следующие:

- планирование эксперимента дает четкую последовательную логическую схему построения всего процесса исследования, т. е. известно что, когда и как надо делать;
- внедрение активного планирования позволяет повысить эффективность исследований, извлечь наибольшее количество сведений об изучаемых процессах при ограниченных затратах, сократить объём экспериментальных исследований, повысить надежность и чёткость интерпретации полученных результатов;
- обработка результатов эксперимента осуществляется стандартными приёмами, позволяющими формализовать процесс построения модели и сопоставить материалы различных исследований.

Таким образом, при оптимизации исследуемых процессов активный эксперимент наиболее эффективен при исследовании в лабораторных условиях, т. е. на этапе оптимального проектирования.

В то же время, в процессе производства технологический процесс постоянно подвергается воздействию случайных неконтролируемых возмущений, что приводит к смещению найденного в лабораторных условиях оптимума относительно технологических факторов. Чтобы иметь возможность оценить это смещение и вести процесс при наиболее благоприятных условиях, необходимо после проведения лабораторных исследований продолжить изучение технологического процесса в реальных производственных условиях. Наилучшие результаты при исследовании технологического процесса в производственных условиях, т. е. на этапе оптимального управления, дает пассивный эксперимент.

Пассивный эксперимент сводится к сбору и обработке данных, полученных в результате пассивного наблюдения за технологическим процессом в производственных условиях. Для анализа и обработки этих данных в настоящее время применяется достаточно большое число методов. К ним относится, в первую очередь, регрессионный и корреляционный анализ, а также факторный анализ, метод главных компонент, временные ряды (дрейф параметров во времени) и др.

В результате проведения регрессионного и корреляционного анализа исследуемого процесса в производственных условиях, можно определить уравнение регрессии и найти с помощью коэффициента корреляции степень взаимосвязи изучаемых переменных величин. Однако сами по себе уравнения регрессии и коэффициент корреляции мало что говорят о возможной причинной связи между рассматриваемыми переменными. Для установления этой связи можно использовать факторный анализ, который является довольно гибким количественным методом статистического анализа. Он в большей мере, чем другие методы, может применяться для проверки сложных гипотез и позволяет получить информацию о числе факторов в исследуемой системе, их природе и зависимости, а также степени этой зависимости. Так, по наблюдениям за вариациями 30...40 различных переменных можно с помощью факторного анализа получить конкретную информацию о том, что только пять факторов коррелируют между собой и каждый из них в той или иной степени влияет на изменения соответствующих исходных

переменных. Этим путем можно проверить гипотезу, выдвинутую по результатам наблюдений, полученным при анализе другими методами.

Наиболее существенным недостатком факторного анализа является отсутствие однозначного математического решения проблемы факторных нагрузок, т. е. вклада отдельных факторов в изменения значений функции отклика.

Преимущество пассивного эксперимента состоит в том, что при его применении нет необходимости тратить время и средства на постановку опытов. Полученные результаты можно затем использовать для управления процессом. Однако пассивный эксперимент имеет существенные недостатки, ограничивающие его применение для оптимизации технологических процессов.

Грамотно организованный пассивный эксперимент и анализ его результатов могут дать богатую информацию о реальном процессе и позволяют не только скорректировать результаты предварительно проведенного активного эксперимента, но в ряде случаев даже определить модель исследуемого процесса. Однако это требует глубокого познания механизма явлений изучаемого процесса, и чем он сложнее, тем очевиднее необходимость высокого уровня предварительных теоретических знаний экспериментатора об исследуемом процессе.

Планируемый активный эксперимент при прочих равных условиях точнее и информативнее, а иногда и дешевле пассивного. Это следует учитывать при выборе вида эксперимента. В вычислительном эксперименте, в отличие от физического, нет никаких ограничений на выбор управляемых факторов и характер их изменения. Поэтому вычислительные эксперименты обычно всегда реализуются как активные. В дальнейшем будут рассматриваться в основном вопросы, связанные с планированием активных экспериментов.

При планировании активных экспериментов используются следующие принципы:

- отказ от полного перебора всех возможных состояний объекта;
- постепенное усложнение структуры математической модели;
- сопоставление результатов эксперимента с величиной случайных помех;
- рандомизация опытов;
- оптимальное планирование эксперимента.

2. Факторы, общая характеристика, функция отклика

В качестве целевой функции принимают тот параметр, который наиболее универсально характеризует объект и является обобщенной характеристикой цели исследования. Одним из основных требований, предъявляемых к параметру оптимизации, является универсальность и полнота характеристики, им определяемой. Кроме того, параметр оптимизации должен выражаться в численном виде и быть статистически однозначным, т.е. определенной совокупности значений факторов соответствует единственный параметр оптимизации. Обратное утверждение не имеет смысла, поскольку одному значению параметра оптимизации может соответствовать множество сочетаний значений факторов. Если параметр оптимизации невозможно измерить по причине отсутствия технических средств, то используют принцип ранжирования, в соответствии с которым целевой функции присваивают ранги (баллы), используя специальную шкалу баллов.

Желательно, чтобы параметр оптимизации имел строгий физический смысл и был прост в определении.

Факторами называют параметры, с помощью которых можно воздействовать на состояние объекта. Они обладают следующими основными свойствами:

- факторы должны быть измеряемыми и управляемыми, т.е. должна иметься возможность установить фактор на определенном уровне и измерить его с достаточной степенью точности;
- факторы должны иметь непосредственное однозначное воздействие на объект.

Кроме указанных, существуют требования, предъявляемые к совокупности факторов, определяющих объект исследования:

- факторы должны быть совместимы, т.е. должна иметь место техническая возможность устанавливать любые значения каждого из факторов вне зависимости от других (кроме того, при этом должно соблюдаться условие безопасности эксперимента);
- факторы должны быть независимы, т.е. изменение величины одного фактора не должно быть причиной изменения другого.

При изменении значений факторов изменяется величина целевой функции. Уравнение, связывающее целевую функцию с факторами, называется функцией отклика $y=\varphi(x_1, x_2, \dots, x_n)$.

Функция отклика, описывающая поведение объекта и взаимосвязь между целевой функцией и факторами в некотором диапазоне их изменения, называется интерполяционной моделью.

Геометрическая интерпретация функции отклика называется поверхностью отклика. Такая поверхность имеет место в случае двухфакторной задачи. Во всех остальных случаях под поверхностью отклика понимается некая гиперповерхность, т.е. гипотетическая многомерная поверхность, дать графическую иллюстрацию которой не представляется возможным, вследствие отсутствия соответствующих навыков.

Для описания поверхности отклика может быть применена та или иная математическая зависимость. Такая зависимость, используемая для решения задач по методу планирования эксперимента (т.е. задач интерполяции и оптимизации), должна отвечать следующим требованиям:

- достаточно точно описывать объект (поверхность отклика);
- должна указывать направление движения к оптимуму

Решить экспериментальную задачу можно несколькими методами: путем перебора в экспериментах всех возможных состояний (совокупности факторов); путем случайного перебора возможных сочетаний значений всех факторов; можно использовать некую программу поиска решения задачи.

Детальное представление о свойствах поверхности отклика может быть получено лишь при условии использования густой дискретной сетки значений факторов, покрывающей все факторное пространство. В узлах этой многомерной сетки находятся точки плана, в которых проводятся опыты. В этом случае в принципе можно получить факторную модель, которая будет практически почти полностью соответствовать исходной теоретической модели. Однако в большинстве случаев при решении практических задач, для которых используется факторная модель, такого детального описания не требуется. Выбор структуры факторной модели основан на постулировании определенной степени гладкости поверхности отклика.

Поэтому с целью уменьшения количества опытов принимают небольшое число точек плана, для которых осуществляется реализация эксперимента. В отсутствие априорной информации о свойствах функции отклика нет смысла сразу строить сложную математическую модель объекта.

Если проверка этой модели на адекватность не дает удовлетворительного результата, ее постепенно усложняют путем изменения структуры (например, повышая степень полинома, принятого в качестве факторной модели, или вводя в модель дополнительные факторы и т.п.).

При этом используются результаты опытов, выполненных при построении простой модели, и проводится некоторое количество дополнительных опытов. При большом уровне случайной помехи получается большой разброс значений функции отклика Y в опытах, проведенных в одной и той же точке плана.

В этом случае оказывается, что чем выше уровень помехи, тем с большей вероятностью простая модель окажется работоспособной, чем меньше уровень помехи, тем точнее должна быть факторная модель. Кроме случайной помехи при проведении эксперимента может иметь место систематическая помеха.

Наличие этой помехи практически никак не обнаруживается и результат ее воздействия на функцию не поддается контролю. Однако если путем соответствующей организации проведения опытов искусственно создать случайную ситуацию, то систематическую помеху можно перевести в разряд случайных. Такой принцип организации эксперимента называют *рандомизацией* систематически действующих помех. Наличие помех приводит к ошибкам эксперимента.

Ошибки подразделяют на *систематические* и *случайные*, соответственно наименованиям вызывающих их факторов – помех. В вычислительных активных экспериментах ошибки характерны только для определяемых значений функций отклика. Если исходить из целей построения факторных моделей на основе теоретических моделей, полагая, что теоретические модели дают точное описание физических свойств технического объекта, а регрессионная модель является ее аппроксимацией, то значения функций отклика будут содержать только случайную ошибку.

В этом случае необходимости в рандомизации опытов не возникает. Рандомизацию опытов осуществляют только в физических экспериментах. Следует отметить, что в этих экспериментах систематическую ошибку может порождать, наряду с отмеченными в предыдущем параграфе факторами, также неточное задание значений управляемых факторов, обусловленное некачественной калибровкой приборов для их измерения (инструментальная ошибка), конструктивными или технологическими факторами.

К факторам в активном эксперименте предъявляются определенные требования. Они должны быть:

- 1) **управляемыми** (установка заданных значений и поддержание постоянными в процессе опыта);
- 2) **совместными** (их взаимное влияние не должно нарушать процесс функционирования объекта);
- 3) **независимыми** (уровень любого фактора должен устанавливаться независимо от уровней остальных);

- 4) **однозначными** (одни факторы не должны быть функцией других);
- 5) **непосредственно влияющими на выходные параметры.**

В вычислительном эксперименте реализация трех первых требований не создает никаких затруднений, а в физическом эксперименте могут возникнуть сложности и далее невозможность их осуществления, что приведет к необходимости замены активного эксперимента пассивным.

Функции отклика должны быть:

- 1) **численно измеряемыми;**
- 2) **иметь четкий физический смысл;**
- 3) **однозначными** (характеризовать только одно свойство объекта);
- 4) **информационными** (полностью характеризовать определенное свойство объекта);
- 5) **статистически эффективными** (измеряться с достаточной точностью с целью сокращения дублирования опытов).

1. 4 Лекция №4 (2 часа).

Тема: «Планы факторного эксперимента. Полный факторный эксперимент. Дробный факторный эксперимент»

1.4.1 Вопросы лекции:

1. Оптимизация как цель математического моделирования. Планы факторного эксперимента
2. Полный факторный эксперимент
3. Дробный факторный эксперимент

1.4.2 Краткое содержание вопросов:

1. Оптимизация как цель математического моделирования. Планы факторного эксперимента

Многие задачи, с которыми приходится иметь дело в повседневной практике, являются многовариантными. Среди множества возможных вариантов приходится отыскивать «наилучшие» в некотором смысле, то есть при ограничениях, налагаемых на природные, экономические и технологические возможности. В связи с этим возникла необходимость применять для анализа и синтеза систем математические методы и современную вычислительную технику. Такие методы объединяются под общим названием — математическое программирование.

Математическое программирование — область математики, разрабатывающая теорию и численные методы решения многомерных экстремальных задач с ограничениями, т. е. задач на экстремум функции многих переменных с ограничениями на область изменения этих переменных.

Функцию, экстремальное значение которой нужно найти в условиях экономических возможностей, называют *целевой, показателем эффективности или критерием оптимальности*. Экономические возможности формализуются в виде *системы ограничений*. Все это составляет математическую модель. *Математическая модель* задачи — это отражение оригинала в виде функций, уравнений, неравенств, цифр и т. д. Модель задачи математического программирования включает:

1) совокупность неизвестных величин, действуя на которые, систему можно совершенствовать. Их называют *планом задачи* (вектором управления, решением, управлением, стратегией, поведением и др.);

2) целевую функцию (функцию цели, показатель эффективности, критерий оптимальности, функционал задачи и др.). Целевая функция позволяет выбирать наилучший вариант - из множества возможных. Наилучший вариант доставляет целевой функции экстремальное значение. Это может быть прибыль, объем выпуска или реализации, затраты производства, издержки обращения, уровень обслуживания или дефицитности, число комплектов, отходы и т. д.;

Эти условия следуют из ограниченности ресурсов, которыми располагает общество в любой момент времени, из необходимости удовлетворения насущных потребностей, из условий производственных и технологических процессов. Ограничеными являются не только материальные, финансовые и трудовые ресурсы.

Таковыми могут быть возможности технического, технологического и вообще научного потенциала. Нередко потребности превышают возможности их удовлетворения. Математически ограничения выражаются в виде уравнений и неравенств. Их совокупность образует *область допустимых решений*.

План, удовлетворяющий системе ограничений задачи, называется *допустимым*. Допустимый план, доставляющий функции цели экстремальное значение, называется *оптимальным*. Оптимальное решение, вообще говоря, не обязательно единственно, возможны случаи, когда оно не существует, имеется конечное или бесчисленное множество оптимальных решений.

Математическая теория планирования эксперимента позволяет повысить эффективность экспериментальных исследований. Основы этой теории заложил английский статистик Р. Фишер.

Он впервые показал целесообразность одновременного варьирования многими переменными в противовес широко распространённому однофакторному эксперименту. Планирование эксперимента — это процедура выбора числа и условий проведения опытов, необходимых и достаточных для решения поставленной задачи.

При этом существенно следующее:

- 1) одновременное варьирование всеми переменными, определяющими изучаемый процесс, по специальным правилам, алгоритмам;
- 2) стремление к минимизации общего числа опытов.

В данном случае используется кибернетическое представление о модели объекта в виде «черного ящика». Пусть в процессе исследования какого-либо объекта («черного ящика») некоторое качество его или целевая функция Y зависит от нескольких величин $Y = F\{x_1, x_2, \dots, x_n\}$ (1)

Переменные x_i представляющие варьируемые переменные, называют факторами, а Y — выход, функция отклика, целевая функция. Выходов Y может быть несколько.

В планировании могут участвовать только независимые факторы, которые можно устанавливать и поддерживать на фиксированных уровнях в течение опыта.

При этом каждый фактор x_i может принимать в опыте одно из нескольких допустимых значений. Эти значения называют уровнями, фиксированный набор уровней факторов определяет одно из возможных состояний объекта и один опыт в эксперименте.

Все возможные наборы уровней факторов определяет общее число возможных различных опытов.

В качестве модели рассматривается представление (1) в виде степенного ряда

$$Y = B_0 + B_1 X_1 + \dots + B_n X_n + B_{12} X_1 X_2 + \dots + B_{n-1,n} X_{n-1} X_n + \\ + B_{11} X_1^2 + \dots + R_{nn} X_n^2 + \dots \quad (2)$$

На практике ограничиваются конечным числом членов разложения, т.е. неизвестная функция аппроксимируется усеченным полиномом некоторой степени. Каждый фактор имеет свой допустимый диапазон изменения, т.е. известны граничные значения X_{\min} , X_{\max} . Каждому фактору соответствует своя координатная ось, а образованное таким образом пространство называют факторным пространством. Назначая граничные значения, задаем область определения функции Y в факторном пространстве. На рисунке представлена область экспериментирования функции двух переменных.

Поскольку факторы в общем случае размерные величины, их кодируют, чтобы иметь дело с безразмерными факторами. Операция кодирования представляет собой линейное преобразование факторного пространства, что ведет к переносу начала координат факторного пространства в точку с координатами

$$X_{icp} = \frac{X_{i \min} + X_{i \max}}{2}.$$

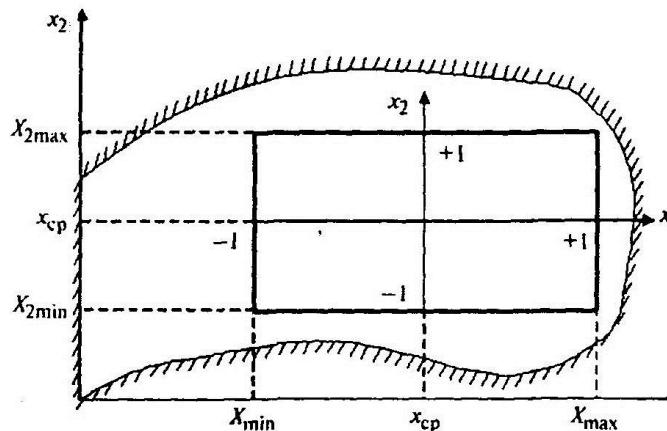


Рис. Область экспериментирования функции двух переменных

Тогда для каждого фактора X_{\min} будет соответствовать -1 , а $X_{\max} +1$. Кодированные значения факторов x_i определяются следующим соотношением:

$$x_i = \frac{x_i - x_{icp}}{x_{icp} - x_{i \min}} = \frac{x_i - x_{icp}}{x_{i \max} - x_{icp}}. \quad (2)$$

Теперь уравнение (2) можно переписать через кодированные факторы в виде

$$Y = b_0 + b_1 x_1 + \dots + b_n x_n + b_{12} x_1 x_2 + \dots + b_{n-1,n} x_{n-1} x_n + \\ + b_{11} x_1^2 + \dots + b_{nn} x_n^2 + \dots \quad (3)$$

$$Y = b_0 + \sum_{i=1}^n b_i x_i + \sum_{\substack{i=1 \\ j \neq i}}^n b_{ij} x_i x_j + \sum_{i=1}^n b_{ii} x_i^2 + \dots$$

или

(4)

Уравнение (4) можно переписать в еще более компактной форме, если ввести фиктивный фактор x_0 , тождественно равный единице, и обозначить все двойные, тройные взаимодействия, а также квадраты факторов символом X_i , а соответствующие коэффициенты символом b_i

$$Y = \sum_{i=0}^m b_i x_i .$$

(5)

Пример. Пусть имеем два фактора X_1 и X_2 . Тогда в соответствии с уравнением (4)

$$Y = b_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2 + b_{11} x_1^2 + b_{22} x_2^2$$

или

в

соответствии

(5.5)

$$Y = b_0 x_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4 + b_5 x_5 ,$$

и

т.д.

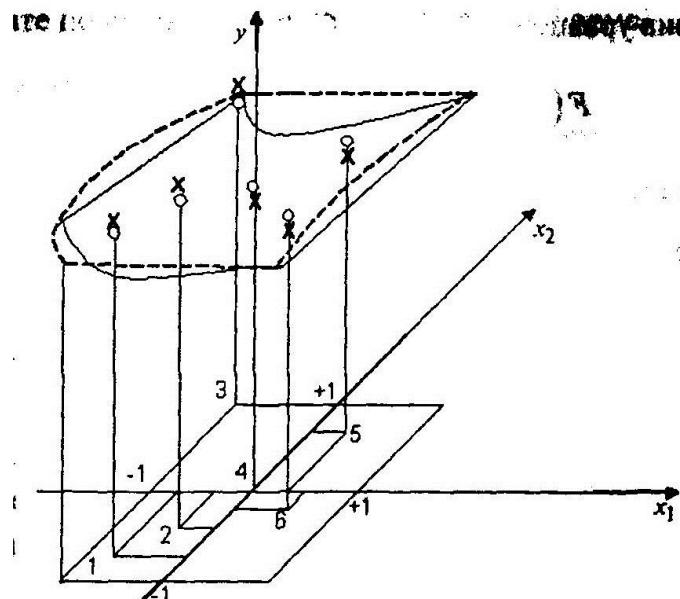


Рис. 2. Поверхность отклика в двухфакторном пространстве

В соответствии с (5) задача нахождения Y заключается в том, чтобы на основе эксперимента при вариации x_i определить неизвестные коэффициенты b_i .

Выражения (1) и (5) называют функцией отклика, представляющей в факторном пространстве в общем случае поверхность отклика. На рис. 2 показана поверхность

отклика в двухфакторном пространстве. Близость аппроксимирующей поверхности к истинной поверхности можно оценить тем или иным критерием. В качестве такой меры близости удобно выбрать квадратичную форму вида

$$\sum_u^N (Y_u - Y_u^*)^2 = \min, \quad (6)$$

где N — число экспериментальных значений функции отклика, число опытов в эксперименте; u — номер опыта; Y — значения функций отклика, предсказанная аппроксимирующим выражением; Y_u — значения истинной поверхности отклика.

Рассмотрим пример для случая одного фактора:

$$Y^* = b_0 + b_1 x. \quad (7)$$

Подставив уравнение (5.7) в (5.6), получим:

$$F(b_0, b_1) = \sum_u^N (Y_u - b_0 - b_1 x_u)^2 = \min.$$

Чтобы найти экстремум, приравняем нулю частные производные:

$$\frac{\partial F(b_0, b_1)}{\partial b_0} = -2 \sum_{u=1}^N (Y_u - b_0 - b_1 x_u) = 0;$$

$$\frac{\partial F(b_0, b_1)}{\partial b_1} = -2 \sum_{u=1}^N (Y_u - b_0 - b_1 x_u) x_u = 0$$

или

$$Nb_0 + b_1 \sum_{u=1}^N x_u = \sum_{u=1}^N Y_u;$$

$$b_0 \sum_{u=1}^N x_u + b_1 \sum_{u=1}^N x_u^2 = \sum_{u=1}^N x_u Y_u.$$

Решая эту систему, получаем искомые значения b_0 и b_1 . Выражения для вычисления b_0 и b_1 , получаются громоздкими, а при большем числе факторов задача вычисления коэффициентов еще больше усложняется.

В связи с этим используют матричную форму записи уравнений и решений относительно коэффициентов b . Если было проведено N опытов, в каждом из которых задавалось определенное сочетание факторов (в n -м опыте — $x_{0u}, x_{1u}, \dots, x_{mu}$), то все возможные сочетания факторов можно представить матрицей X , все результаты — матрицей Y , а все искомые коэффициенты — матрицей B :

$$X = \begin{bmatrix} x_{01} & x_{11} & \cdots & x_{m1} \\ x_{02} & x_{12} & \cdots & x_{m2} \\ \vdots & \vdots & & \vdots \\ x_{0u} & x_{1u} & \cdots & x_{mu} \\ \vdots & \vdots & & \vdots \\ x_{0N} & x_{1N} & \cdots & x_{mN} \end{bmatrix}; \quad Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_u \\ \vdots \\ Y_N \end{bmatrix}; \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_u \\ \vdots \\ b_N \end{bmatrix},$$

В результате получим уравнение в матричной форме

$$Y = XB \quad (1)$$

решение можно получить в виде $B = X^{-1}Y$.

В такой форме надо обращать матрицу, что непросто.

Чтобы упростить решение, запишем (1) в эквивалентной форме $X^T Y = X^T X B$,

где X^T — транспонированная матрица X . Тогда получаем $C B = X^T Y$, (2)

где матрица $C = X^T X$ — квадратная и имеет $(1 + m)$ строк и столбцов.

Для определения коэффициентов A , умножим на матрицу C^{-1} слева обе части уравнения (2). Поскольку $C^{-1}C$ есть единичная матрица, то получаем: $B = C^{-1}X^T Y$ (3)

Анализ матрицы C показывает, что она симметричная: на главной диагонали элементы

имеют вид $\sum_{u=1}^N x_{iu}^2$, а элементы, симметрично расположенные сверху и снизу от главной диагонали, равны между собой, т.е.

$$\sum_{u=1}^N x_{iu} x_{ju} = \sum_{u=1}^N x_{ju} x_{iu}, \text{ где } i = 0, 1, \dots, m, j \neq i.$$

Для получения обратной матрицы C^{-1} матрица C должна быть невырожденной, т.е. её определитель не должен равняться нулю

$-\Delta_C \neq 0$. Если точности аппроксимации не достигается, то надо изменить полином.

Любое изменение полинома приведет к изменению всех коэффициентов b_i . Стало быть снова надо обращать матрицу C , что приводит к существенному росту объема вычислений. Другими словами, при изложенной процедуре все коэффициенты оказываются зависимыми друг от друга. Конечно, хотелось бы избежать этих трудностей и чтобы коэффициенты, вычисленные при исходной, начальной модели, не надо было бы пересчитывать при усложнении модели для достижения большей точности аппроксимации. Оказывается, преодолеть указанные трудности можно с помощью выбора сочетания факторов в каждом опыте по специальному алгоритму.

$$C = \begin{bmatrix} c_{00} & 0 & \cdots & 0 & \cdots & 0 \\ 0 & c_{11} & \cdots & 0 & \cdots & 0 \\ \vdots & \cdots & \cdots & & & \\ 0 & 0 & \cdots & c_{ii} & \cdots & 0 \\ \vdots & \cdots & \cdots & \vdots & & \\ 0 & 0 & \cdots & 0 & \cdots & c_{mm} \end{bmatrix}, \text{ а } C^{-1} = \begin{bmatrix} \frac{1}{c_{00}} & 0 & \cdots & 0 & \cdots & 0 \\ 0 & \frac{1}{c_{11}} & \cdots & 0 & \cdots & 0 \\ \vdots & \cdots & \cdots & & & \\ 0 & 0 & \cdots & \frac{1}{c_{ii}} & \cdots & 0 \\ \vdots & \cdots & \cdots & \vdots & & \\ 0 & 0 & \cdots & 0 & \cdots & \frac{1}{c_{mm}} \end{bmatrix}$$

Вычислительные трудности существенно уменьшаются, если матрица C является диагональной. Тогда обратная матрица находится очень просто. В этом случае получим $1 + m$ независимых уравнений в соответствии с уравнением (5.10)

$$b_i = \frac{1}{c_{ii}} \sum_{u=1}^N x_{iu} Y_u \quad c_{ii} = \sum_{u=1}^N x_{iu}^2,$$

И с учетом того, что

$$b_i = \frac{\sum_{u=1}^N x_{iu} Y_u}{\sum_{u=1}^N x_{iu}^2}.$$

Имеем в итоге:

Итак, необходимо добиться, чтобы матрица С удовлетворяла условию $c_{ij} = c_{ji} = 0$ (элементы над и под главной диагональю) или что тоже

$$\sum_{u=1}^N x_{iu} x_{ju} = \sum_{u=1}^N x_{ju} x_{iu} = 0. \quad (4)$$

Следовательно, условие (4) означает равенство нулю произведения любых двух столбцов матрицы X. Но столбец матрицы можно рассматривать как вектор. Если скалярное произведение двух векторов равно нулю, то векторы ортогональны. Поэтому условие (4) называют условием ортогональности матрицы X, а соответствующий выбор плана эксперимента — ортогональным планированием.

2. Полный факторный эксперимент

Эксперимент, в котором реализуются все возможные сочетания уровней факторов, называется полным факторным экспериментом (ПФЭ). Число возможных сочетаний уровней (или число опытов) определяется по формуле $N = rp$, где r — число уровней; p — число факторов. Рассмотрим свойства и особенности факторного планирования на простых примерах.

Пример 1. Полный фактор эксперимента при $n = 2$ и $p = 2$. Рассмотрим описание модели второго порядка $Y = b_0x_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5$.

Число строк (опытов) в матрице планирования $N = 2^2 = 4$, а максимальное число столбцов — 6. Собственно план эксперимента определяется числом факторов — их два X_1 и X_2 . Остальные столбцы матрицы планирования являются производными столбцов x_1 и x_2 . План эксперимента формируется так, чтобы выполнить условие ортогональности — сумма элементов каждого столбца плана эксперимента должна равняться нулю, т.е.

$$\sum_{u=1}^N x_{iu} x_{ju} = \sum_{u=1}^N x_{ju} x_{iu} = 0.$$

Это достигается чередованием знаков в столбце x_1 , а в каждом последующем столбце плана эксперимента частота чередования вдвое уменьшается по сравнению с предыдущим столбцом. Отметим особенность матрицы планирования. Исходная модель требует определения шести коэффициентов (b_0 — b_5). Следовательно, должно быть шесть уравнений (или шесть столбцов). Но различных столбцов всего четыре, что соответствует числу строк.

u	x_0	x_1	x_2	$x_3 = x_1 x_2$	$x_4 = x_1^2$	$x_5 = x_2^2$
1	+	-		+	+	+
2	+	+		-	+	+
3	+	-	+	-	+	+
4	+	+	+	+	+	+

Однако число различных столбцов равно числу членов аппроксимирующего полинома (*), включая все возможные взаимодействия (в нашем примере одно взаимодействие $b_{12}x_1x_2$). Вывод ПФЭ 2^n позволяет определить все коэффициенты полинома при линейных членах и всех возможных взаимодействиях факторов (включая в общем случае взаимодействие максимально высокого порядка x_1, x_2, \dots, x_n).

То, что столбцы 4 и 5 одинаковы и совпадают с нулевым столбцом свидетельствует о том, что рассматриваемый ПФЭ (и все ПФЭ 2^n) не позволяет определить все коэффициенты, в частности коэффициенты при X_i^2 , точнее невозможно независимо определить коэффициенты b_0, b_{11} и b_{22} , т.е. получается смешанная оценка коэффициентов. Такие планы ПФЭ называют планами первого порядка, так как они позволяют определить коэффициенты лишь при факторах в первой степени.

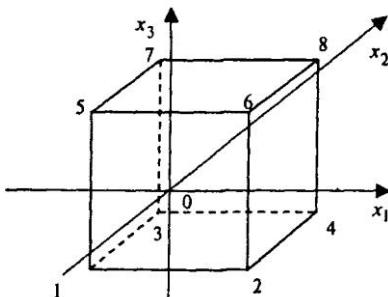
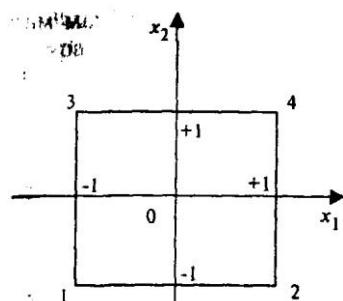
Отсюда вытекает практическая рекомендация — в планах ПФЭ 2^3 и больших степеней нет смысла формировать столбцы для квадратов факторов x_i^2 .

Пример 2. Полный фактор эксперимента при $n = 3$ и $p = 2$. Модель принимает вид $Y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_1x_2 + b_5x_1x_3 + b_6x_2x_3 + b_7x_1x_2x_3$

В этом случае $N=23=8$, т.е. будет восемь строк в матрице планирования.

Из матриц планирования следует характерное для ПФЭ 2^n соотношение

u	x_0	x_1	x_2	x_3	$x_4 = x_1x_2$	$x_5 = x_1x_3$	$x_6 = x_2x_3$	$x_7 = x_1x_2x_3$
1	+	-	-	-	+	+	+	-
2	+	+	-	-	-	-	+	+
3	+	-	+	-	-	+	-	+
4	+	+	+	-	+	-	-	-
5	+	-	-	+	+	-	-	+
6	+	+	-	+	-	+	-	-
7	+	-	+	+	-	-	+	-
8	+	+	+	+	+	+	+	+



т.е. сумма квадратов значений факторов в каждом столбце равна числу опытов N . С учетом этого формула (5.11) для вычисления коэффициентов упрощается и принимает вид

$$b_i = \frac{\sum_{u=1}^N x_{iu} Y_u}{N}.$$

Точки планов ПФЭ 2^n располагаются в вершинах квадрата, куба или в общем случае гиперкуба. Анализируя ПФЭ 2^n , легко заметить, что с ростом числа уровней факторов p и, особенно, с ростом числа факторов n число опытов, предусмотренных планом, будет существенно превышать число искомых коэффициентов, т.е. ПФЭ дает избыточную информацию. Можно ли более экономно организовать эксперимент? Это можно сделать посредством рассмотрения дробных реплик ПФЭ.

3. Дробный факторный эксперимент

Рассмотренные в полном факторном эксперименте примеры свидетельствуют о том, что ПФЭ полностью определяют коэффициенты для линейного уравнения $Y = \sum_{i=0}^n b_i x_i$, при $n < m$.

В такой модели необходимо определить $n + 1$ искомых коэффициентов. Как построить экономно эксперимент? В этом случае строят план эксперимента, который представляет часть плана ПФЭ. Такие планы называют дробными репликами ПФЭ или дробным факторным экспериментом (ДФЭ). Но при этом все требования к столбцам матрицы планирования должны соблюдаться, т.е.

$$\sum_{i,j=1}^N x_i x_j = 0, \quad \sum_{i=1}^N x_i^2 = N, \quad \sum_{i=1}^N x_i = 0.$$

План дробного факторного эксперимента строится очень просто. Так, для линейной модели из трех факторов требуется ДФЭ, содержащий четыре опыта. Таким образом, строится план эксперимента для меньшего числа факторов ($3-1 = 2$), а оставшиеся факторы варьируются как произведения каких-либо факторов, включенных в первую группу. Но, если есть уверенность, что в выбранном интервале варьирования факторов объект может быть оценен линейной моделью, то смешанные коэффициенты b_{12}, b_{13}, \dots стремятся к нулю.

В этом случае факторами x_1 и x_2 варьируют как при ПФЭ, а вместо произведения $x_1 x_2$ ($b_{12} = 0$) вводят дополнительный фактор x_3 . Другими словами укороченную матрицу планирования строим для трех факторов и дополнительный фактор x_3 будет меняться как произведение $x_1 x_2$.

Матрица планирования ДФЭ приведена на рисунке ниже. В матрице планирования четыре пары одинаковых столбцов, а это значит, что соответствующие коэффициенты неразличимы и можно судить лишь об их совместной величине (смешанные оценки коэффициентов) — линейные эффекты смешиваются с эффектами взаимодействия. Но, если объект линейный, b_{ij} и b_{ijk} равны нулю. Тогда из четырех опытов найдем истинные значения b_0 и b_i (b_1, b_2, b_3)

x_0	x_1	x_2	x_3	$x_4 = x_1 x_2$	$x_5 = x_1 x_3$	$x_6 = x_2 x_3$	$x_7 = x_1 x_2 x_3$
1	+	-	-	+	+	-	-
2	+	+	-	-	-	-	+
3	+	-	+	-	-	+	-
4	+	+	+	+	+	+	+

Если объект не является линейным и поверхность отклика, найденная по уравнению, отличается сильно от действительной, то надо ставить ПФЭ.

Дробные реплики особенно удобны при большом числе факторов ($n > 5$), так как при этом коэффициенты при факторах $b_1 — b_n$ смешиваются при тройных и более высоких взаимодействиях, влияние которых существенно слабее, чем при двойных взаимодействиях. Поэтому влиянием этих взаимодействий можно пренебречь.

Общий вывод: планы ПФЭ 2^n или ДФЭ 2^{n-k} позволяют получить решение для линейных и неполных квадратичных полиномов (без квадратичных членов). Невозможность использования этих планов можно считать, например, расхождение между истинным значением Y^n и Y , предсказанным аппроксимирующим выражением в нулевой точке плана ($x_1=x_2=\dots=0$), превышающее допустимое. Более точную аппроксимацию можно получить разными способами. Распространение получил переход к планам второго порядка, позволяющим найти коэффициенты при квадратичных членах аппроксимирующего полинома (x^2).

1. 5 Лекция №5 (2 часа).

Тема: «Теоретические основы обработки экспериментальных данных»

1.5.1 Вопросы лекции:

1. Многомерные СВ, законы их распределения, условные числовые характеристики
2. Функция регрессии, коэффициент детерминации, корреляции, ковариация
3. Виды регрессий, статистическая значимость их параметров. Автокорреляция

1.5.2 Краткое содержание вопросов:

1. Многомерные СВ, законы их распределения, условные числовые характеристики

Совместное рассмотрение двух или нескольких случайных величин приводит к понятию системы случайных величин. Условимся систему нескольких случайных величин X, Y, \dots, W обозначать (X, Y, \dots, W) . Такая система называется также **многомерной случайной величиной**. При изучении системы случайных величин недостаточно изучить отдельно случайные величины, составляющие систему, а необходимо учитывать связи или зависимости между этими величинами.

При рассмотрении системы случайных величин удобно пользоваться геометрической интерпретацией системы. Например, систему двух случайных величин (X, Y) можно рассматривать как случайную точку на плоскости XOY с координатами X и Y или как случайный вектор на плоскости со случайными составляющими X и Y . По аналогии, систему n случайных величин можно рассматривать как случайную точку в n -мерном пространстве или как n -мерный случайный вектор¹.

В дальнейшем, при изучении системы случайных величин ограничимся подробным рассмотрением системы двух случайных величин.

Закон распределения вероятностей системы случайных величин

Законом распределения вероятностей системы случайных величин называется соответствие, устанавливающее связь между областями возможных значений данной системы случайных величин и вероятностями появления системы в этих областях.

Так же, как и для одной случайной величины, закон распределения системы случайных величин может быть задан в различных формах. Рассмотрим таблицу распределения вероятностей системы двух дискретных случайных величин. Пусть X и Y – дискретные случайные величины, возможные значения которых (x_i, y_j) , где $i = \overline{1, n}; j = \overline{1, m}$. Тогда распределение системы таких случайных величин может быть охарактеризовано указанием вероятностей $p_{i,j} = P\{X = x_i, Y = y_j\}$ того, что случайная

величина примет значение \mathbf{x}_i и одновременно с этим случайная величина примет значение \mathbf{y}_j . Вероятности $p_{i,j}$ фиксируются в таблице

\mathbf{y}_i	\mathbf{y}_1	\mathbf{y}_2	\dots	\mathbf{y}_m
\mathbf{x}_j				
\mathbf{x}_1	p_{11}	p_{12}	\dots	p_{1m}
\mathbf{x}_2	p_{21}	p_{22}	\dots	p_{2m}
\vdots	\vdots	\vdots	\vdots	\vdots
\mathbf{x}_n	p_{n1}	p_{n2}	\dots	p_{nm}

Такая таблица называется таблицей распределения вероятностей системы двух дискретных случайных величин с конечным числом возможных значений. Все возможные события $\{X = \mathbf{x}_i, Y = \mathbf{y}_j\}$ при составляют полную группу несовместных событий, поэтому

$$\sum_{i,j} p_{i,j} = \sum_{i,j} P\{X = \mathbf{x}_i, Y = \mathbf{y}_j\} = 1$$

При этом:

$$\sum_j p_{i,j} = \sum_j P\{X = \mathbf{x}_i, Y = \mathbf{y}_j\} = P\{X = \mathbf{x}_i\};$$

$$\sum_i p_{i,j} = \sum_i P\{X = \mathbf{x}_i, Y = \mathbf{y}_j\} = P\{Y = \mathbf{y}_j\}.$$

Функция распределения

Функцией распределения вероятностей системы двух случайных величин называется функция $F(\mathbf{x}; \mathbf{y})$ двух аргументов, равная вероятности совместного выполнения двух неравенств $X < x$ и $Y < y$, то есть $F(\mathbf{x}; \mathbf{y}) = P\{X < x, Y < y\}$.

Геометрически функцию распределения системы двух случайных величин можно интерпретировать как вероятность попадания случайной точки (\mathbf{X}, \mathbf{Y}) в левый нижний бесконечный квадрант с вершиной в точке (x, y) плоскости (см. рис.).

Сформулируем основные свойства функции распределения вероятностей системы двух случайных величин (без доказательства):

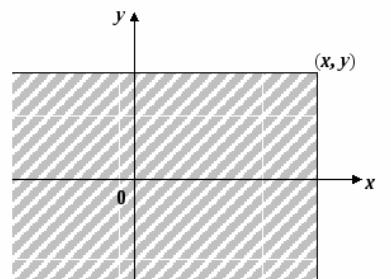
$$1. 0 \leq F(\mathbf{x}, \mathbf{y}) \leq 1, \quad \forall \mathbf{x}, \mathbf{y}$$

$$2. \lim_{y \rightarrow +\infty} F(\mathbf{x}, \mathbf{y}) = F_1(\mathbf{x}), \quad \lim_{x \rightarrow +\infty} F(\mathbf{x}, \mathbf{y}) = F_2(\mathbf{y}) \quad (\text{или } F(\mathbf{x}, +\infty) = F_1(\mathbf{x}); F(+\infty, \mathbf{y}) = F_2(\mathbf{y})).$$

$$3. \lim_{x \rightarrow +\infty} F(\mathbf{x}, \mathbf{y}) = 1 \quad (\text{или } F(+\infty, +\infty) = 1).$$

$$\lim_{x \rightarrow -\infty} F(\mathbf{x}, \mathbf{y}) = \lim_{y \rightarrow -\infty} F(\mathbf{x}, \mathbf{y}) = \lim_{x \rightarrow -\infty} F(\mathbf{x}, \mathbf{y}) = 0$$

$$4. F(\mathbf{x}, -\infty) = F(-\infty, \mathbf{y}) = F(-\infty, -\infty) = 0.$$



5. Функция распределения является неубывающей функцией по каждому из своих аргументов, то есть:

$$F(x_2, y) \geq F(x_1, y), \text{ если } x_2 > x_1;$$

$$F(x, y_2) \geq F(x, y_1), \text{ если } y_2 > y_1.$$

6. Вероятность попадания случайной точки (X, Y) в произвольный прямоугольник со сторонами, параллельными координатным осям (см. рис.) вычисляется по формуле:

$$P(x_1 < X < x_2, y_1 < Y < y_2) = F(x_1, y_1) + F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1),$$

где $x_1 < x_2, y_1 < y_2$.

Плотность распределения вероятностей системы двух случайных величин¹

Предположим, что функция распределения $F(x, y)$ всюду непрерывна и дважды дифференцируема² (за исключением, быть может, конечного числа кривых). Тогда, смешанная частная производная функции

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y} = F''_{xy}(x, y).$$

Функция $f(x, y)$ называется плотностью распределения (или, дифференциальной функцией распределения) системы непрерывных случайных величин.

Геометрически эту функцию можно истолковать как поверхность, которую называют поверхностью распределения.

Зная плотность распределения, можно определить вероятность попадания случайной точки в произвольную область D :

$$P\{(X, Y) \subset D\} = \iint_D f(x, y) dx dy$$

Используя последнюю формулу, выразим интегральную функцию распределения вероятностей системы двух непрерывных случайных величин через плотность распределения:

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(x, y) dx dy$$

Рассмотрим некоторые свойства плотности распределения системы двух непрерывных случайных величин:

$$1. f(x, y) \geq 0, \quad \forall (-\infty < x, y < +\infty);$$

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1$$

2. , если случайная величина распределена на всей координатной плоскости (если же распределена в некоторой плоской области G , то $\iint_G f(x, y) dx dy = 1$)

Условные законы распределения

Пусть известна плотность распределения системы двух случайных величин. Используя свойства функций распределения, можно вывести формулы для нахождения плотности распределения одной величины, входящей в систему:

$$f_1(x) = F_1'(x) = \int_{-\infty}^{+\infty} f(x, y) dy; \quad f_2(y) = F_2'(y) = \int_{-\infty}^{+\infty} f(x, y) dx \quad (*)$$

Перейдём теперь к решению обратной задачи: по известным законам распределения отдельных случайных величин, входящих в систему, найти закон распределения системы. Легко увидеть, что в общем случае эта задача неразрешима. Действительно, с одной стороны, законы распределения отдельных случайных величин, входящих в систему, характеризуют каждую из случайных величин в отдельности, но ничего не говорят о том,

как они взаимосвязаны. С другой стороны, искомый закон распределения системы должен содержать все сведения о случайных величинах системы, в том числе и о характере связей между ними.

Таким образом, если случайные величины \mathbf{X}, \mathbf{Y} взаимозависимы, то закон распределения системы не может быть выражен через законы распределения отдельных случайных величин, входящих в систему. Это приводит к необходимости введения условных законов распределения.

Распределение одной случайной величины, входящей в систему, найденное при условии, что другая случайная величина, входящая в систему, приняла определённое значение, называется условным законом распределения.

Для дискретных случайных величин условным распределением составляющей при условии, что $\mathbf{Y} = y_j$ называется совокупность условных вероятностей $p(x_1 | y_j), p(x_2 | y_j), \dots, p(x_n | y_j)$, вычисленных в предположении, что случайная величина уже приняла значение y_j . Для нахождения $p(x_i | y_j)$ пользуются формулой

$$p(x_i | y_j) = \frac{p(x_i, y_j)}{p(y_j)}, \quad i = \overline{1, n}$$

$$\sum_{i=1}^n p(x_i | y_j) = 1$$

Заметим, что

$$p(y_j | x_i) = \frac{p(x_i, y_j)}{p(x_i)}, \quad j = \overline{1, m}$$

Аналогично находим

Условный закон распределения можно задавать как функцией распределения, так и плотностью распределения. Условная функция распределения обозначается $F(x | y)$; условная плотность распределения обозначается $f(x | y)$.

Плотностью распределения для случайной величины при условии, что случайная величина приняла определённое значение (*условной плотностью распределения*), назовём величину

$$f(x | y) = \frac{f(x, y)}{f_2(y)}$$

Аналогично, плотностью распределения для случайной величины при условии, что случайная величина приняла определённое значение, назовём величину

$$f(y | x) = \frac{f(x, y)}{f_1(x)}$$

Отсюда получаем: $f(x, y) = f_1(x) f(y | x) = f_2(y) f(x | y)$

$$f(x | y) = \frac{f(x, y)}{\int_{-\infty}^{+\infty} f(x, y) dx}; \quad f(y | x) = \frac{f(x, y)}{\int_{-\infty}^{+\infty} f(x, y) dy}$$

или, с учётом формул (*)

Условная плотность распределения обладает всеми свойствами безусловной плотности распределения. В частности,

$$\int_{-\infty}^{+\infty} f(x | y) dx = 1; \quad \int_{-\infty}^{+\infty} f(y | x) dy = 1$$

Числовые характеристики условных законов распределения

Для описания условных законов распределения можно использовать различные характеристики подобно тому, как для одномерных распределений.

Наиболее важной характеристикой является условное математическое ожидание.

Условным математическим ожиданием дискретной случайной величины при $\mathbf{Y} = \mathbf{y}$ (\mathbf{y} – определённое возможное значение случайной величины) называется сумма произведений возможных значений на их условные вероятности:

$$\mathbf{M}(X|Y) = \sum_{i=1}^n x_i P(x_i|y)$$

Для непрерывных случайных величин:

$$\mathbf{M}(X|Y) = \int_{-\infty}^{+\infty} xf(x|y)dx$$

где – условная плотность распределения случайной величины при .

Аналогично, условным математическим ожиданием дискретной случайной величины при $X = x$ (x – определённое возможное значение случайной величины) называется сумма произведений возможных значений на их условные вероятности:

$$\mathbf{M}(Y|X) = \sum_{j=1}^m y_j P(y_j|x)$$

$$\mathbf{M}(Y|X) = \int_{-\infty}^{+\infty} yf(y|x)dy$$

Для непрерывных случайных величин:

где $f(y|x)$ – условная плотность распределения случайной величины при .

Аналогично вводятся условные дисперсии и условные моменты более высоких порядков.

Числовые характеристики системы двух случайных величин

Две случайные величины называются независимыми, если закон распределения одной из них не зависит от того, какие возможные значения приняла другая величина. Из этого определения следует, что условные распределения независимых случайных величин равны их безусловным распределениям. Укажем необходимые и достаточные условия независимости случайных величин.

ТЕОРЕМА 1: Для того чтобы случайные величины X и Y были независимыми, необходимо и достаточно, чтобы функция распределения системы (X, Y) была равна произведению функций распределения составляющих:

$$F(x, y) = F_1(x) \cdot F_2(y)$$

ТЕОРЕМА 2: Для того чтобы случайные величины и были независимыми, необходимо и достаточно, чтобы плотность вероятности системы была равна произведению плотностей вероятностей составляющих:

$$f(x, y) = f_1(x) \cdot f_2(y)$$

Для описания системы двух случайных величин кроме математических ожиданий и дисперсий составляющих используют и другие характеристики, к которым относятся корреляционный момент и коэффициент корреляции.

Корреляционным моментом μ_{xy} случайных величин и называют математическое ожидание произведения отклонений этих величин:

$$\mu_{xy} = M((X - M(X)) \cdot (Y - M(Y)))$$

Для вычисления корреляционного момента дискретных величин используют

$$\mu_{xy} = \sum_{i=1}^n \sum_{j=1}^m (x_i - M(X))(y_j - M(Y))P(x_i, y_j)$$

формулу:

$$\mu_{xy} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - M(X))(y - M(Y))f(x, y)dxdy$$

а для непрерывных величин:

Корреляционный момент служит для характеристики связи между величинами и .

ТЕОРЕМА 3: Корреляционный момент двух независимых случайных величин и равен нулю.

Замечание: из теоремы 3 следует, что если корреляционный момент двух случайных величин и не равен нулю, то и – зависимые случайные величины.

Коэффициентом корреляции r_{xy} случайных величин и называют отношение корреляционного момента к произведению средних квадратических отклонений этих величин:

$$r_{xy} = \frac{\mu_{xy}}{\sigma_x \sigma_y}.$$

Очевидно, коэффициент корреляции двух независимых случайных величин равен нулю (так как $\mu_{xy} = 0$).

2. Функция регрессии, коэффициент детерминации, корреляции, ковариация

Условимся обозначать через X независимую переменную, а через Y – зависимую переменную.

Зависимость величины Y от X называется **функциональной**, если каждому значению величины X соответствует единственное значение величины Y . С функциональной зависимостью мы встречаемся, например, в математике, при изучении физических законов. Обратим внимание на то, что если X – детерминированная величина (т.е. принимающая вполне определённые значения), то и функционально зависящая от неё величина Y тоже является детерминированной; если же X – случайная величина, то и Y также случайная величина.

Однако гораздо чаще в окружающем нас мире имеет место не функциональная, а **стохастическая, или вероятностная, зависимость**, когда каждому фиксированному значению независимой переменной X соответствует не одно, а множество значений переменной Y , причём сказать заранее, какое именно значение примет величина Y , нельзя. Более частое появление такой зависимости объясняется действием на результирующую переменную не только контролируемого или контролируемых факторов (в данном случае таким контролируемым фактором является переменная X), а и многочисленных неконтролируемых случайных факторов. В этой ситуации переменная Y является случайной величиной. Переменная же X может быть как детерминированной, так и случайной величиной. Следует заметить, что со стохастической зависимостью мы уже сталкивались в дисперсионном анализе.

Допустим, что существует стохастическая зависимость случайной переменной Y от X . Зафиксируем некоторое значение x переменной X . При $X = x$ переменная Y в силу её стохастической зависимости от X может принять любое значение из некоторого множества, причём какое именно – заранее неизвестно. Среднее этого множества называют **групповым генеральным средним** переменной Y (при $X = x$) или **математическим ожиданием случайной величины Y , вычисленным при условии, что $X = x$** ; это **условное математическое ожидание обозначают так: $M(Y/X = x)$** . Если существует стохастическая зависимость Y от X , то прежде всего стараются выяснить, изменяются или нет при изменении x условные математические ожидания $M(Y/X=x)$. Если при изменении x условные математические ожидания $M(Y/X=x)$ изменяются, то говорят, что имеет место **корреляционная зависимость** величины Y от X ; если же условные математические ожидания остаются неизменными, то говорят, что корреляционная зависимость величины Y от X отсутствует.

Функция $\phi(x)=M(Y/X=x)$, описывающая изменение условного математического ожидания случайной переменной Y при изменении значений x переменной X , называется **функцией регрессии**.

Выясним, почему именно при наличии стохастической зависимости интересуются поведением условного математического ожидания.

Рассмотрим пример. Пусть X – уровень квалификации рабочего, Y – его выработка за смену. Зависимость Y от X является не функциональной, а стохастической, ведь на выработку помимо квалификации влияет множество других факторов. Зафиксируем значение x уровня квалификации: ему соответствует некоторое множество значений выработки Y . Тогда $M(Y/X = x)$ – средняя выработка рабочего при условии, что его уровень квалификации равен x , или, иначе говоря, $M(Y/X = x)$ – это норматив выработки при уровне квалификации, равном x . Зная зависимость этого норматива от уровня квалификации, можно для любого уровня квалификации рассчитать норматив выработки и, сравнив его с реальной выработкой, оценить работу рабочего.

Обратим внимание на то, что введённые понятия стохастической и корреляционной зависимости относились к генеральной совокупности. Поясним эти понятия числовым примером.

Пример. Допустим, что одновременно изучаются две случайные величины X и Y , или, иначе говоря, двумерная случайная величина (X, Y) , которая задана табл. 1.

Таблица 1.

j	i	1	2	3
y_i	x_i	$x_1 = 2$	$x_2 = 5$	$x_3 = 8$
	1	$y_1 = 0,4$	0,15	0,12
2	$y_2 = 0,8$	0,05	0,30	0,35

Табл. 1 называют **таблицей распределения двумерной величины (X, Y)** ; её следует понимать так. Случайная величина X может принять одно из следующих значений: 2, 5 и 8. Случайная величина Y – значения 0,4 и 0,8. Число 0,15 – это вероятность того, что $X = 2$ и одновременно $Y = 0,4$, или, иначе говоря, вероятность произведения двух событий; события, состоящего в том, что $X = 2$, и события, состоящего в том, что $Y = 0,4$, т.е. $P((X=2)(Y=0,4)) = 0,15$. Аналогично, вероятность $P((X=2)(Y=0,8)) = 0,05$ и т.д. Обратим внимание на следующее: поскольку в табл. 1 указаны все возможные значения величин X и Y , сумма вероятностей, стоящих в таблице, должна быть равна единице: $0,15 + 0,05 + 0,12 + 0,30 + 0,03 + 0,35 = 1$.

Прежде чем выяснить тип зависимости величины Y от X , найдём:

а) Закон распределения величины X . Он представлен табл. 2.

Таблица 2.

x	$x_1 = 2$	$x_2 = 5$	$x_3 = 8$	
$P(X=x)$	$0,15 + 0,05 = 0,2$	$0,12 + 0,30 = 0,42$	$0,35 + 0,03 = 0,38$	$\Sigma = 1$

$$M(X) = 5,54, D(X) = 4,9284$$

Действительно, например, величина X примет значение, равное 2, только в том случае, когда одновременно с этим величина Y примет значение 0,4 или 0,8, т.е.

$$P(X = 2) = P((X = 2)(Y = 0,4)) + P((X = 2)(Y = 0,8)) = 0,15 + 0,05 = 0,2.$$

Справа от ряда распределения величины X находятся её математическое ожидание и дисперсия.

б) Закон распределения величины Y . Он имеет вид табл. 3.

Таблица 3.

y	$y_1 = 0,4$	$y_2 = 0,8$	
$P(Y=y)$	$0,15 + 0,12 + 0,03 = 0,30$	$0,05 + 0,30 + 0,35 = 0,7$	$\Sigma = 1$

$$M(Y) = 0,68, \quad D(Y) = 0,0336$$

в) Условные законы распределения величины Y , а именно закон распределения величины Y сначала при условии, что $X = 2$, затем при условии, что $X=5$, и наконец, при условии, что $X = 8$.

Итак, пусть $X = 2$. Тогда условная вероятность

$$P(Y = 0,4/X = 2) = \frac{P((Y = 0,4)(X = 2))}{P(X = 2)} = \frac{0,15}{0,2} = 0,75,$$

а условная вероятность

$$P(Y = 0,8/X = 2) = \frac{P((Y = 0,8)(X = 2))}{P(X = 2)} = \frac{0,05}{0,2} = 0,25.$$

Таким образом, закон распределения величины Y при условии, что $X = 2$, задан табл. 4.

Таблица 4.

y	$y_1 = 0,4$	$y_2 = 0,8$	
$P(Y = y/X = 2)$	0,75	0,25	$\Sigma = 1$

$$M(Y/X = 2) = 0,4*0,75 + 0,8*0,25 = 0,5, \quad D(Y/X = 2) = 0,03$$

Справа помещено условное математическое ожидание и значение условной дисперсии. Покажем, как вычисляется условная дисперсия. Общая формула условной дисперсии имеет вид

$$D(Y/X = x) = M[(Y/X = x) - M(Y/X = x)]^2. \quad (23)$$

Для табл. 4 получаем

$$D(Y/X = 2) = M [(Y/X = 2) - M(Y/X = 2)]^2 = M [(Y/X = 2) - 0,5]^2 = \\ \sum_{i=1}^2 (y_i - 0,5)^2 \cdot P(Y = y_i/X = 2) = (0,4 - 0,5)^2 \cdot 0,75 + (0,8 - 0,5)^2 \cdot 0,25 = 0,03.$$

$$\text{Пусть } X = 5. \text{ Тогда } P(Y = 0,4/X = 5) = \frac{P((Y = 0,4)(X = 5))}{P(X = 5)} = \frac{0,12}{0,42} = \frac{2}{7}; \quad P(Y = 0,8/X = 5) = \frac{P((Y = 0,8)(X = 5))}{P(X = 5)} = \frac{0,30}{0,42} = \frac{5}{7}.$$

Таким образом, закон распределения величины Y при условии, что $X = 5$, имеет вид табл. 5.

Таблица 5.

y	0,4	0,8	
$P(Y = y/X = 5)$	2/7	5/7	$\Sigma = 1$

$$M(Y/X = 5) = \frac{24}{35} \approx 0,686, \quad D(Y/X = 5) = 0,03265.$$

И наконец, при $X = 8$ ряд распределения задан таблицей 14.

Таблица 6.

y	0,4	0,8	
$P(Y = y/X = 8)$	$\frac{3}{38}$	$\frac{35}{38}$	$\Sigma = 1$

$$M(Y/X = 8) = \frac{73}{95} \approx 0,768, \quad D(Y/X = 8) = 0,01163$$

Из табл. 4–6 видно, что зависимость Y от X стохастическая, поскольку при каждом фиксированном значении величины X величина Y может быть равной либо 0,4, либо 0,8,

причём какому именно из этих чисел она будет равна – сказать заранее нельзя. Ясно прослеживается и корреляционная зависимость величины Y от X , поскольку с изменением значений x величины X меняются и условные математические ожидания $M(Y/X = x)$. Функция регрессии, т.е. зависимость условного математического ожидания $M(Y/X = x)$ от x , задаётся в виде табл. 7.

Таблица 7.

x	2	5	8
$M(Y/X = x)$	0,5	$24/35 \approx 0,686$	$73/95 \approx 0,768$

Выборочный коэффициент корреляции

Выясним, можно ли измерить степень корреляционной и стохастической зависимости величины Y от X . Ответ проиллюстрируем примером 5. Все полученные в примере результаты объединены в табл. 8.

Таблица 8.

x_i	$x_1 = 2$	$x_2 = 5$	$x_3 = 8$
$P(X = x_i)$	0,2	0,42	0,38
$M(Y/X = x_i)$	0,5	0,686	0,768
$D(Y/X = x_i)$	0,03	0,03265	0,01163

$$MY = 0,68 \text{ (см. табл. 3)}, \quad DY = 0,0336 \text{ (см. табл. 3)}$$

Т.к. X – случайная величина, принимающая значения 2, 5 и 8 с вероятностью 0,2; 0,42 и 0,38, то такими же будут вероятности и условных математических ожиданий, и дисперсий. Т.обр., условное математическое ожидание $M(Y/X)$, так же как и условная дисперсия $D(Y/X)$ – случайные величины.

Обратим также внимание на то, что $M(Y)$, найденное в табл. 3, можно вычислить и по табл. 8 следующим образом:

$$M(Y) = M[M(Y/X)] = \sum_{i=1}^3 M(Y/X = x_i)P(X = x_i) = 0,5*0,2 + 0,686*0,42 + 0,768*0,38 = 0,68.$$

Разброс значений величины Y вокруг математического ожидания MY измеряется дисперсией $D(Y)$, или σ_Y^2 :

$$\sigma_Y^2 = D(Y) = M(Y - MY)^2. \quad (24)$$

(По табл. 3: $\sigma_Y^2 = 0,0336$.) Этот разброс может быть вызван:

- зависимостью величины Y от X (эта зависимость может быть обусловлена не только непосредственным влиянием X на Y , но и наличием случайных факторов, действующих на Y через переменную X);
 - зависимостью величины Y от случайных факторов, влияющих только на Y и не влияющих на X ;
- эти факторы называют **остаточными**.

1) Построим показатель разброса значений величины Y , связанного с её зависимостью от фактора X .

Условное математическое ожидание $M(Y/X = x)$ является «представителем греков», которые имеют место при $X = x$. Характеристикой разброса условных математических ожиданий $M(Y/X = x)$ относительно $M(Y)$ является дисперсия $D[M(Y/X)]$, или

$$\sigma_\phi^2 = D[M(Y/X)] = M[M(Y/X) - MY]^2 \quad (25)$$

– эта величина и будет показателем разброса значений величины Y , связанного с её зависимостью от фактора X . По таблице 16 найдём:

$$\sigma_{\varphi}^2 = M[(Y/X) - MY]^2 = (0,5 - 0,68)^2 * 0,2 + (0,686 - 0,68)^2 * 0,42 + (0,768 - 0,68)^2 * 0,38 = 0,0095.$$

2) Теперь построим показатель разброса «игреков», связанного с влиянием остаточных факторов.

Зафиксируем какое-либо значение x величины X . Тогда причиной вариации величины Y при $X = x$ будут остаточные факторы, влияющие только на Y и не влияющие на X . Измерителем этой вариации является условная дисперсия $D(Y/X = x)$. При различных же «каксах» характеристикой разброса «игреков», вызванного влиянием на Y остаточных факторов, будет генеральное среднее из условных дисперсий, или, иначе, математическое ожидание условной дисперсии. Эту величину обозначим σ_0^2 . Имеем

$$\sigma_0^2 = M[D(Y/X)], \quad (26)$$

где при $X = x$ дисперсия $D(Y/X = x)$ вычисляется по формуле (23). (По табл. 16 найдём

$$\sigma_0^2 = M[D(Y/X)] = \sum_{i=1}^3 D(Y/X = x_i) P(X = x_i) = 0,03 * 0,2 + 0,03265 * 0,42 + 0,01163 * 0,38 = 0,0241.)$$

Для вычисленных дисперсий справедливо тождество

$$DY = D[M(Y/X)] + M[D(Y/X)]$$

или

$$\sigma_Y^2 = \sigma_{\varphi}^2 + \sigma_0^2. \quad (27)$$

Степень стохастической зависимости величины Y от X измеряется **генеральным корреляционным отношением**

$$\rho_{Y/X} = + \sqrt{\frac{D[M(Y/X)]}{DY}} = + \sqrt{\frac{\sigma_{\varphi}^2}{\sigma_Y^2}} = + \sqrt{1 - \frac{\sigma_0^2}{\sigma_Y^2}} = + \sqrt{1 - \frac{M[D(Y/X)]}{DY}}. \quad (28)$$

Квадрат корреляционного отношения

$$\rho_{Y/X}^2 = \frac{\sigma_{\varphi}^2}{\sigma_Y^2} = \frac{D[M(Y/X)]}{DY} \quad (26), (25) = \frac{M[M(Y/X)] - MY]^2}{M(Y - MY)^2} \quad (29)$$

называется **генеральным коэффициентом детерминации**; он показывает, какую долю дисперсии величины Y составляет дисперсия условных математических ожиданий, или, иначе говоря, какая доля дисперсии $D(Y)$ объясняется корреляционной зависимостью Y от X . (В примере $\sigma_{\varphi}^2 = 0,0095$, $\sigma_Y^2 = 0,0336$, поэтому $\rho_{Y/X}^2 = \frac{0,0095}{0,0336} = 0,28$, т.е. 28% дисперсии величины Y объясняется её корреляционной зависимостью от X ; $\rho_{Y/X} = + \sqrt{0,28} = 0,53$.)

Свойства генерального корреляционного отношения как измерителя степени корреляционной и стохастической зависимости

$$1. \Theta \leq \rho_{Y/X} \leq 1.$$

Действительно, согласно (29), $\rho_{Y/X} \geq 0$; с другой стороны, из (28) следует, что $\sigma_{\varphi}^2 \leq \sigma_Y^2$, поэтому $\rho_{Y/X} \leq 1$.

2. Условие $\rho_{Y/X} = 0$ является необходимым и достаточным для отсутствия корреляционной зависимости Y от X , т.е. для того, чтобы $M(Y/X) = \text{const}$ при любом значении x величины X .

3. Условие $\rho_{Y/X} = 1$ является необходимым и достаточным для функциональной зависимости величины Y от X .

Следствие. Чем ближе $\rho_{Y/X}$ к единице, тем в силу (28) ближе к нулю $M[D(Y/X)]$, а следовательно, и условные дисперсии $D(Y/X = x)$. Это означает, что при каждом допустимом значении x уменьшается разброс «игреков» относительно $M(Y/X = x)$. Т.о., чем ближе $\rho_{Y/X}$ к единице, тем меньше при каждом x отличие «игреков» от постоянного числа, равного $M(Y/X = x)$, или, иначе говоря, тем выше степень стохастической зависимости Y от X . И, наоборот, чем выше степень стохастической зависимости Y от X , тем ближе $\rho_{Y/X}$ к единице.

В практических задачах наибольший интерес представляют следующие вопросы:

- существует ли корреляционная зависимость Y от X или нет, иначе говоря, отлично ли генеральное корреляционное отношение $\rho_{Y/X}$ от нуля или равно нулю;
- если корреляционная зависимость существует, то какой вид имеет функция регрессии (линейный, параболический или какой-либо другой).

Точно ответить на поставленные вопросы можно лишь в том случае, когда известен закон распределения двумерной величины (X, Y). В примере 5 этот закон задан табл. 9, в которой даны все возможные значения случайных величин X и Y и вероятности совместного появления этих значений. Обычно такими сведениями не располагают; как правило, имеются лишь наблюдавшиеся значения двумерной величины (X, Y). Покажем как, имея наблюдавшиеся значения, ответить на поставленные выше вопросы.

Выборочное корреляционное отношение. Его значимость

Пусть имеется n наблюдений двумерной величины (X, Y). Наблюдавшиеся «иксы» и «игреки» поместим в табл. 17, которая называется **корреляционной таблицей** и строится следующим образом:

- «иксы» группируются в вариационный ряд, число групп которого обозначим v ; если это дискретный ряд, то x_1, x_2, \dots, x_v – различающиеся между собой результаты наблюдений или варианты; если это интервальный ряд, то x_1, x_2, \dots, x_v – центры интервалов;
- «игреки» группируют в вариационный ряд, число групп которого обозначим q : y_1, y_2, \dots, y_q – это либо варианты, если ряд дискретный, либо середины интервалов, если ряд интервальный;
- подсчитывают числа m_{ji} таких наблюдавшихся пар чисел (x, y) , у которых x попадает в группу x_i , а y – в группу y_j , $i = 1, 2, \dots, v$, $j = 1, 2, \dots, q$; например, m_{12} – число пар чисел (x, y) , у которых x попало в группу x_2 , а y – в группу y_1 . Числа $m_{11}, m_{12}, \dots, m_{qv}$ называются **частотами**.

Прежде чем пояснить остальные элементы этой таблицы, сделаем следующее замечание по поводу схемы построения выборочного корреляционного отношения: от табл. 17, содержащей частоты, можно перейти к таблице частостей (табл. 18). Сравним табл. 18 и табл. 9. Их различие состоит в следующем: табл. 9 относилась к генеральной совокупности, поэтому в ней были указаны все мыслимые значения величин X и Y и вероятности комбинаций этих значений; табл. 18 относится к выборочной совокупности, и в ней приведены наблюдаемые значения величин X и Y и частости, или опытные вероятности комбинаций наблюдаемых значений. Поэтому выборочное корреляционное отношение можно строить по той же схеме, что и генеральное корреляционное отношение, если заменить возможные значения величин X и Y на наблюдаемые, вероятности на частости, математические ожидания на средние, дисперсии на выборочные дисперсии.

Однако чаще выборочное корреляционное отношение строят, используя непосредственно табл. 9, а не табл. 1. В табл. 9 кроме сгруппированных наблюдений и частот содержатся следующие данные:

$$- \text{суммы частот по каждой строке} \quad m_1 = \sum_{i=1}^v m_{1i}, m_2 = \sum_{i=1}^v m_{2i}, \dots, m_q = \sum_{i=1}^v m_{qi}, \quad (30)$$

$$- \text{суммы частот по каждому столбцу} \quad n_1 = \sum_{j=1}^q m_{j1}, n_2 = \sum_{j=1}^q m_{j2}, \dots, n_v = \sum_{j=1}^q m_{jv}, \quad (31)$$

Таблица 9.

j	i	1	2	...	v	
	X	x_1	x_2	...	x_v	Σ
1	y_1	m_{11}	m_{12}	...	m_{1v}	m_1
2...	$y_2\dots$	$m_{21\dots}$	$m_{22\dots}$...	$m_{2v\dots}$	$m_{2\dots}$
q	y_q	m_{q1}	m_{q2}		m_{qv}	m_q

Σ	n_1	n_2	...	n_v	$n = \sum_{i=1}^v n_i = \sum_{j=1}^q m_j$
Групповое среднее	$\bar{Y}^{(1)}$	$\bar{Y}^{(2)}$...	$\bar{Y}^{(v)}$	
Групповая выборочная дисперсия	σ_*^2	σ_*^2	...	σ_*^2	

Таблица 10

j	i	1	2	...	v	
	X	x_1	x_2	...	x_v	
1	y_1	p_{11}^*	p_{12}^*	...	p_{1v}^*	
2...	$y_2...$	p_{21}^* ...	p_{22}^*	p_{2v}^* ...	$p_{ij}^* = m_{ij}/n, i = 1, 2, \dots, v$ $j = 1, 2, \dots, q.$
q	y_q	p_{q1}^*	p_{q2}^*		p_{qv}^*	

- групповые средние значения «игреков»

$$\begin{aligned}\bar{Y}^{(1)} &= (y_1 m_{11} + y_2 m_{21} + \dots + y_q m_{q1})/n_1, \\ \bar{Y}^{(2)} &= (y_1 m_{12} + y_2 m_{22} + \dots + y_q m_{q2})/n_2, \\ &\dots \\ \bar{Y}^{(v)} &= (y_1 m_{1v} + y_2 m_{2v} + \dots + y_q m_{qv})/n_v.\end{aligned}\quad (32)$$

Эти средние являются выборочными аналогами соответствующих условных математических ожиданий: $\bar{Y}^{(1)}$ - выборочный аналог математического ожидания величины Y при условии, что $X = x_1$, т.е. аналог величины $M(Y/X=x_1)$; $\bar{Y}^{(2)}$ - аналог $M(Y/X=x_2)$ и т.д.

Групповые выборочные дисперсии:

$$\begin{aligned}\hat{\sigma}_1^2 &= [(y_1 - \bar{Y}^{(1)})^2 m_{11} + (y_2 - \bar{Y}^{(1)})^2 m_{21} + \dots + (y_q - \bar{Y}^{(1)})^2 m_{q1}] / n_1, \\ \hat{\sigma}_2^2 &= [(y_1 - \bar{Y}^{(2)})^2 m_{12} + (y_2 - \bar{Y}^{(2)})^2 m_{22} + \dots + (y_q - \bar{Y}^{(2)})^2 m_{q2}] / n_2, \\ &\dots \\ \hat{\sigma}_v^2 &= [(y_1 - \bar{Y}^{(v)})^2 m_{1v} + (y_2 - \bar{Y}^{(v)})^2 m_{2v} + \dots + (y_q - \bar{Y}^{(v)})^2 m_{qv}].\end{aligned}\quad (33)$$

Эти дисперсии являются выборочными аналогами соответствующих условных дисперсий: $\hat{\sigma}_1^2$ – выборочный аналог условной дисперсии $D(Y/X=x_1)$, $\hat{\sigma}_2^2$ – аналог $D(Y/X=x_2)$ и т.д.

Построим выборочный аналог генерального корреляционного отношения. Выборочным аналогом генеральной дисперсии $\sigma_y^2 = DY$ является величина $\hat{\sigma}_y^2$. Для того чтобы вычислить $\hat{\sigma}_y^2$, найдем сначала среднее \bar{Y} по данным табл. 17. Это можно сделать по одной из следующих тождественных формул:

$$\bar{Y} = \frac{1}{n} (y_1 m_1 + y_2 m_2 + \dots + y_q m_q) = \frac{1}{n} (\bar{Y}^{(1)} n_1 + \bar{Y}^{(2)} n_2 + \dots + \bar{Y}^{(v)} n_v) \quad (34)$$

Напомним, что \bar{Y} – это выборочный аналог $M(Y)$. Теперь найдем

$$S_Y^2 = (y_1 - \bar{Y})^2 m_1 + (y_2 - \bar{Y})^2 m_2 + \dots + (y_q - \bar{Y})^2 m_q = \sum_{j=1}^q (y_j - \bar{Y})^2 m_j \quad (35)$$

Тогда

$$\hat{\sigma}_y^2 = S_Y^2 / n \quad (36)$$

Выборочным аналогом генеральной дисперсии $\sigma_{\phi}^2 = D[M(Y/X)]$ является выборочная дисперсия $\hat{\sigma}_{y(i)}^2$ групповых средних; обозначим ее $\hat{\sigma}_\phi^2$. Имеем

$$\hat{\sigma}_{\phi}^2 = \hat{\sigma}_{y(i)}^2 = \frac{1}{n} \left[(\bar{Y}^{(1)} - \bar{Y})^2 n_1 + (\bar{Y}^{(2)} - \bar{Y})^2 n_2 + \dots + (\bar{Y}^{(v)} - \bar{Y})^2 n_v \right] = S_{\phi}^2/n, \quad (37)$$

где

$$S_{\phi}^2 = (\bar{Y}^{(1)} - \bar{Y})^2 n_1 + (\bar{Y}^{(2)} - \bar{Y})^2 n_2 + \dots + (\bar{Y}^{(v)} - \bar{Y})^2 n_v = \\ = \sum_{i=1}^v (\bar{Y}^{(i)} - \bar{Y})^2 n_i. \quad (38)$$

Выборочным аналогом дисперсии $\sigma_o^2 = D[M(Y/X)]$ является средняя $\hat{\sigma}_i^2$ групповых выборочных дисперсий. Обозначим $\hat{\sigma}_o^2$ эту среднюю

$$\hat{\sigma}_o^2 = \bar{\hat{\sigma}}^2 = \frac{1}{n} (\hat{\sigma}_1^2 n_1 + \hat{\sigma}_2^2 n_2 + \dots + \hat{\sigma}_v^2 n_v) = S_o^2/n, \quad (39)$$

где $S_o^2 = \hat{\sigma}_1^2 n_1 + \hat{\sigma}_2^2 n_2 + \dots + \hat{\sigma}_v^2 n_v$.

Получаем

$$S_o^2 = (y_1 - \bar{Y}^{(1)})^2 m_{11} + (y_2 - \bar{Y}^{(1)})^2 m_{21} + \dots + (y_q - \bar{Y}^{(1)})^2 m_{q1} + (y_1 - \bar{Y}^{(2)})^2 m_{12} + \\ + (y_2 - \bar{Y}^{(2)})^2 m_{22} + \dots + (y_q - \bar{Y}^{(2)})^2 m_{q2} + \dots + (y_1 - \bar{Y}^{(v)})^2 m_{1v} + (y_2 - \bar{Y}^{(v)})^2 m_{2v} + \dots \\ + (y_q - \bar{Y}^{(v)})^2 m_{qv} = \sum_{i=1}^v \sum_{j=1}^q (y_j - \bar{Y}^{(i)})^2 m_{ji}. \quad (40)$$

Для вычисленных дисперсий справедливо тождество

$$\hat{\sigma}_Y^2 = \hat{\sigma}_{Y(i)}^2 + \bar{\hat{\sigma}}^2, \text{ или } \hat{\sigma}_Y^2 = \hat{\sigma}_{\phi}^2 + \hat{\sigma}_o^2, \quad (41)$$

аналогичное тождеству, имеющему место в генеральной совокупности.

Выборочный аналог генерального корреляционного отношения вычисляется следующим образом:

$$\hat{\rho}_{Y/X} = + \sqrt{\frac{\hat{\sigma}_{Y(i)}^2}{\hat{\sigma}_Y^2}} = + \sqrt{\frac{\hat{\sigma}_{\phi}^2}{\hat{\sigma}_Y^2}} = + \sqrt{1 - \frac{\hat{\sigma}_o^2}{\hat{\sigma}_Y^2}} = + \sqrt{1 - \frac{\bar{\hat{\sigma}}^2}{\hat{\sigma}_Y^2}} \quad (42)$$

Величина $\hat{\rho}_{Y/X}$ называется выборочным коэффициентом детерминации. Этот коэффициент показывает, какую долю дисперсии $\hat{\sigma}_Y^2$ составляет выборочная дисперсия групповых средних «игреков» или, иначе говоря, какая доля дисперсии $\hat{\sigma}_Y^2$ объясняется зависимостью Y от X.

Как правило, дисперсии $\hat{\sigma}_Y^2$ и $\hat{\sigma}_{\phi}^2$ находят не по рассмотренным выше формулам, а по следующим, более удобным для вычислений:

$$\hat{\sigma}_Y^2 = \frac{1}{n} \sum_{j=1}^q y_j^2 m_j - (\bar{Y})^2, S_y^2 = \hat{\sigma}_Y^2 n. \quad (43)$$

$$\hat{\sigma}_i^2 = \frac{1}{n_i} \sum_{j=1}^q y_j^2 m_{ji} - (\bar{Y}^{(i)})^2, S_y^2 = \hat{\sigma}_Y^2 n. \quad (44)$$

$$\hat{\sigma}_{\phi}^2 = \frac{1}{n} \sum_{i=1}^v (\bar{Y}^{(i)})^2 n_i - (\bar{Y})^2, S_{\phi}^2 = \hat{\sigma}_{\phi}^2 n. \quad (45)$$

Проверим гипотезу $H_0: P_{Y/X} = 0$ (46)

Предварительно отметим, что в силу свойства 2 корреляционного отношения при выполнении гипотезы имеет место равенство условных, или групповых математических ожиданий величины Y: $M(Y/X=x_1)=M(Y/X=x_2)=\dots=M(Y/X=x_v)$.

Поэтому проверка гипотезы сводится к проверке гипотезы о равенстве групповых математических ожиданий – это задача дисперсионного анализа. Ее можно решить, если выполняются требования, применительно к нашим условиям формулирующиеся следующим образом (*):

- при каждом наблюдаемом значении x_i величины X наблюдения величины Y должны быть независимыми и проводиться в одинаковых условиях; наблюдения должны быть независимы и при различных «иксах»;
- при каждом значении x_i величина Y должна иметь нормальный закон с постоянной для

различных «иксов» генеральной дисперсией (обозначим эту дисперсию σ_0^2 ;

$$\sigma_0^2 = D(Y/X = x_1) = (Y/X = x_2) = \dots = (Y/X = x_v)$$

Допустим, что эти требования выполняются. Тогда для проверки гипотезы (46) следует заполнить табл. 19.

В заключение заметим, что если гипотеза отвергается, то говорят, что выборочное корреляционное отношение статистически значимо. Если гипотеза не отвергается, то говорят, что выборочное корреляционное отношение незначимо.

Обратим внимание на то, что вычислить выборочное корреляционное отношение, а также проверить его значимость можно только в том случае, когда результаты наблюдений сгруппированы в таблицу типа таблицы.

Допустим, что, располагая выборочными данными, мы пришли к выводу, что корреляционная зависимость Y от X существует, т.е. при изменении x изменяются условные математические ожидания $M(Y/X=x)$. Тогда возникает вопрос: каков вид функции регрессии, т.е. функции $\phi(x) = M(Y/X=x)$?

Располагая только выборочными данными, нельзя дать точный ответ на поставленный вопрос, но высказать гипотезу о виде функции $\phi(x)$ можно; также можно провести статистическую проверку этой гипотезы, т.е. выяснить, противоречит или нет эта гипотеза имеющимся выборочным данным.

3. Виды регрессий, статистическая значимость их параметров. Автокорреляция

Исследование начинается с теории, устанавливающей связь между явлениями. Из всего круга факторов, влияющих на результативный признак, выделяются наиболее существенные факторы. После того, как было выявлено наличие взаимосвязи между изучаемыми признаками, определяется точный вид этой зависимости с помощью регрессионного анализа.

Регрессионный анализ заключается в определении аналитического выражения (в определении функции), в котором изменение одной величины (результативного признака) обусловлено влиянием независимой величины (факторного признака). Количественно оценить данную взаимосвязь можно с помощью построения уравнения регрессии или регрессионной функции.

Базисной регрессионной моделью является модель парной (однофакторной) регрессии. Парная регрессия – уравнение связи двух переменных Y и X :

$$Y = f(X)$$

где Y – зависимая переменная (результативный признак);

X – независимая, объясняющая переменная (факторный признак).

В зависимости от характера изменения Y с изменением X различают линейные и нелинейные регрессии.

Линейная регрессия $Y = \beta_0 + \beta_1 X + \varepsilon$

Данная регрессионная функция называется полиномом первой степени и используется для описания равномерно развивающихся во времени процессов.

Наличие случайного члена ε (ошибки регрессии) связано с воздействием на зависимую переменную других неучтенных в уравнении факторов, с возможной нелинейностью модели, ошибками измерения, следовательно, появление случайной ошибки уравнения регрессии может быть обусловлено следующими объективными причинами:

1) нерепрезентативность выборки. В модель парной регрессии включается фактор, не способный полностью объяснить вариацию результативного признака, который может быть подвержен влиянию многих других факторов (пропущенных переменных) в гораздо большей степени. Например, заработная плата может зависеть, кроме квалификации, от уровня образования, стажа работы, пола и пр.;

2) существует вероятность того, что переменные, участвующие в модели, могут быть измерены с ошибкой. Например, данные по расходам семьи на питание составляются на

основании записей участников опросов, которые, как предполагается, тщательно фиксируют свои ежедневные расходы. Разумеется, при этом возможны ошибки.

На основе выборочного наблюдения оценивается выборочное уравнение регрессии (линия регрессии):

$$y_x = a + b x,$$

где a , b – оценки параметров уравнения регрессии (α , β).

Аналитическая форма зависимости между изучаемой парой признаков (регрессионная функция) определяется с помощью следующих методов:

На основе теоретического и логического анализа природы изучаемых явлений, их социально-экономической сущности. Например, если изучается зависимость между доходами населения и размером вкладов населения в банки, то очевидно, что связь прямая.

Графический метод, когда характер связи оценивается визуально.

Эту зависимость можно наглядно увидеть, если построить график, отложив на оси абсцисс значения признака x , а на оси ординат – значения признака y . Нанеся на график точки, соответствующие значениям x и y , получим корреляционное поле:

а) если точки беспорядочно разбросаны по всему полю – это говорит об отсутствии зависимости между этими признаками;

б) если точки концентрируются вокруг оси, идущей от нижнего левого угла в верхний правый – то имеется прямая зависимость между признаками;

в) если точки концентрируются вокруг оси, идущей от верхнего левого угла в нижний правый – то обратная зависимость между признаками.

Если на корреляционном поле соединим точки отрезками прямой, то получим ломаную линию с некоторой тенденцией к росту. Это будет эмпирическая линия связи или эмпирическая линия регрессии. По ее виду можно судить не только о наличии, но и о форме зависимости между изучаемыми признаками.

Построение уравнения парной регрессии

Построение уравнения регрессии сводится к оценке ее параметров. Эти оценки параметров могут быть найдены различными способами. Одним из них является метод наименьших квадратов (МНК). Суть метода состоит в следующем. Каждому значению соответствует эмпирическое (наблюданное) значение y . Построив уравнение регрессии, например уравнение прямой линии, каждому значению будет соответствовать теоретическое (расчетное) значение y_x . Наблюдаемые значения не лежат в точности на линии регрессии, т.е. не совпадают с y_x . Разность между фактическим и расчетным значениями зависимой переменной называется остатком:

$$e = y - y_x$$

МНК позволяет получить такие оценки параметров, при которых сумма квадратов отклонений фактических значений результативного признака от теоретических y_x , т.е. сумма квадратов остатков, минимальна:

$$\Sigma(y - y_x)^2 \rightarrow \min$$

Для линейных уравнений и нелинейных, приводимых к линейным, решается следующая система относительно a и b :

$$a n + b \Sigma x = \Sigma y$$

$$a \Sigma x + b \Sigma x^2 = \Sigma yx$$

где n – численность выборки.

Решив систему уравнений, получим значения a и b , что позволяет записать уравнение регрессии(регрессионное уравнение):

$y_x = a + bx$ где x – объясняющая (независимая) переменная;

y_x –объясняемая (зависимая) переменная;

Линия регрессии проходит через точку (\bar{x}, \bar{y}) и выполняются равенства:
 $e = 0$, $y = \bar{y}_x$

Можно воспользоваться готовыми формулами, которые вытекают из этой системы уравнений:

$$b = \frac{\text{cov}(x, y)}{\sigma_x^2} = \frac{\bar{y} \cdot \bar{x} - \bar{y} \cdot \bar{x}}{\bar{x}^2 - \bar{x}^2};$$

$$a = \bar{y} - b \cdot \bar{x}$$

где – среднее значение зависимого признака;

– среднее значение независимого признака;

$\bar{y} \cdot \bar{x}$ – среднее арифметическое значение произведения зависимого и независимого признаков;

σ_x^2 – дисперсия независимого признака;

$\text{cov}(x, y)$ – ковариация между зависимым и независимым признаками.

Выборочной ковариацией двух переменных x, y называется средняя величина произведения отклонений этих переменных от своих средних

Параметр b при x имеет большое практическое значение и носит название коэффициента регрессии. Коэффициент регрессии показывает, на сколько единиц в среднем изменяется величина y при изменении факторного признака x на 1 единицу своего измерения.

Знак параметра b в уравнении парной регрессии указывает на направление связи:

если $b > 0$, то связь между изучаемыми показателями прямая, т.е. с увеличением факторного признаках увеличивается и результативный признаку, и наоборот;

если $b < 0$, то связь между изучаемыми показателями обратная, т.е. с увеличением факторного признаках результативный признак у уменьшается, и наоборот.

Значение параметра a в уравнении парной регрессии в ряде случаев можно трактовать как начальное значение результативного признака y . Такая трактовка параметра a возможна только в том случае, если значение $x = 0$ имеет смысл.

После построения уравнения регрессии, наблюдаемые значения y можно представить как: $y = \bar{y}_x + e$

Остатки e , как и ошибки, являются случайными величинами, однако они, в отличие от ошибок, наблюдаются. Остаток есть та часть зависимой переменной y , которую невозможно объяснить с помощью уравнения регрессии.

На основании уравнения регрессии могут быть вычислены теоретические значения y для любых значений x .

В экономическом анализе часто используется понятие эластичности функции. Эластичность функции рассчитывается как относительное изменение y к относительному изменению x . Эластичность показывает, на сколько процентов изменяется функция при изменении независимой переменной на 1%.

Поскольку эластичность линейной функции не является постоянной величиной, а зависит от x , то обычно рассчитывается коэффициент эластичности как средний показатель эластичности.

Коэффициент эластичности показывает, на сколько процентов в среднем по совокупности изменится величина результативного признака y при изменении факторного признаках на 1% от своего среднего значения:

$$\mathcal{E}_x = b \frac{\bar{x}}{\bar{y}}$$

где \bar{x}, \bar{y} – средние значения переменных x и y в выборке.

Оценка качества построенной модели регрессии

Качество модели регрессии – адекватность построенной модели исходным (наблюдаемым) данным.

Чтобы измерить тесноту связи, т.е. измерить, насколько она близка к функциональной, нужно определить дисперсию, измеряющую отклонения у от ух и характеризующую остаточную вариацию, обусловленную прочими факторами. Они лежат в основе показателей, характеризующих качество модели регрессии.

При исследовании регрессии устанавливается однофакторная или многофакторная будет строиться модель и вид модели (линейный или нелинейный).

Обоснование вида модели состоит в выборе вида функции (некоторого аналитического выражения), с помощью которого можно будет описать изменение исследуемого показателя под воздействием факторов.

К обоснованию вида функции идут двумя путями: Теоретическим (анализируя экономическую природу x_0 и x_1 , выдвигается гипотеза о характере изменения показателя под действием фактора) И эмпирическим(закон изменения результативного показателя под действием фактора устанавливается путем анализа совокупности фактических данных по полям корреляции).

Наиболее употребительными выражениями при описании связи одного фактора и исследуемого показателя являются:

- Уравнение прямой - $x_0 = a_0 + a_1 x_1$, - Уравнение параболы - $x_0 = a_0 + a_1 x_1 + a_2 x_1^2$, - Уравнение гиперболы - $x_0 = a_0 + \frac{a_1}{x_1}$.

После обоснования парных взаимосвязей переходят к записи многофакторных моделей. В экономических исследованиях чаще всего применяется линейная многофакторная модель - $x_0 = a_0 + a_1 x_1 + \dots + a_n x_n$.

В качестве нелинейных моделей применяются

- Мультипликативная модель - $x_0 = a_0 x_1^{a_1} x_2^{a_2} x_3^{a_3} \dots$ или $x_0 = a_0 a_1^{x_1} a_2^{x_2} a_3^{x_3} \dots$

Для оценки значений параметров регрессионной модели чаще всего используется Метод наименьших квадратов (МНК).Этот метод можно применить как для линейных моделей, так и для нелинейных, допускающих преобразование их к линейному виду путем замены переменных или дифференцированием.

При использовании МНК делаются определенные предпосылки относительно случайной составляющей ε . В модели $y = a + b_1 \cdot x_1 + b_2 \cdot x_2 + \varepsilon$ случайная составляющая ε представляет собой ненаблюданную величину. Поэтому в задачу регрессионного анализа входит не только построение самой модели, но и исследование случайных отклонений ε_i , т. е. остаточных величин.

Остатки представляют собой независимые случайные величины, и их среднее значение равно 0; они имеют одинаковую (постоянную) дисперсию и подчиняются нормальному распределению.

Статистические проверки параметров регрессии, показателей корреляции основаны на непроверяемых предпосылках распределения случайной составляющей ε_i . Связано это с тем, что оценки параметров регрессии должны отвечать определенным критериям: быть Несмешенными, состоятельными и эффективными. Эти свойства оценок, полученных по МНК, имеют чрезвычайно важное практическое значение в использовании результатов регрессии и корреляции.

Коэффициенты регрессии, найденные из системы нормальных уравнений, представляют собой выборочные оценки характеристики силы связи. Их несмешенность является желательным свойством, т. к. только в этом случае они могут иметь практическую значимость.

Несмешенность оценки означает, что математическое ожидание остатков равно нулю. Оценки считаются Эффективными, если они характеризуются наименьшей дисперсией. Поэтому несмешенность оценки должна дополняться минимальной дисперсией. Состоительность оценок характеризует увеличение их точности с увеличением объема выработки.

Указанные критерии оценок (несмешенность, состоятельность, эффективность) обязательно учитываются при разных способах оценивания. Метод наименьших квадратов строит оценки регрессии на основе минимизации суммы квадратов остатков ($y - \hat{y}_x$).

Исследование остатков ε_i предполагают проверку наличия следующих пяти предпосылок МНК:

- случайный характер остатков;
- нулевая средняя величина остатков, не зависящая от x_i ;
- гомоскедастичность – дисперсия каждого отклонение ε_i одинакова для всех значений x ;
- отсутствие автокорреляции остатков, т. е. значения остатков ε_i распределены независимо друг от друга;
- остатки подчиняются нормальному распределению.

С целью проверки случайного характера остатков ε_i строится график зависимости остатков ε_i от теоретических значений результативного признака \hat{y} .

. Если на графике нет направленности в расположении точек ε_i , то остатки ε_i представляют собой случайные величины и МНК оправдан. Также возможны следующие случаи: если ε_i зависит от теоретического значения, то:

Вторая предпосылка МНК относительно нулевой средней величины остатков $\Sigma(y - \hat{y}_x) = 0$ означает, что это выполнимо для линейных моделей и моделей, нелинейных относительно включаемых переменных. Для обеспечения несмешенности оценок коэффициентов регрессии, полученных МНК, необходимо выполнение условий независимости случайных остатков ε_i и переменных x , что исследуется в рамках соблюдения второй предпосылки МНК. С целью проверки выполнение этой предпосылки строится график зависимости случайных остатков ε от факторов, включенных в регрессию x_i . Если расположение остатков на графике не имеет направленности, то они независимы от значений x_i . Если же график показывает наличие зависимости ε_i и x_i , то модель неадекватна.

Предпосылка о нормальном распределении остатков позволяет проводить проверку параметров регрессии и корреляции с помощью критериев t и F. Вместе с тем оценки регрессии, найденные с применением МНК, обладают хорошими свойствами даже при отсутствии нормального распределения остатков, т. е. при нарушении пятой предпосылки метода наименьших квадратов.

В соответствии с третьей предпосылкой МНК требуется, чтобы дисперсия остатков была гомоскедастичной. Это означает, что для каждого значения фактора x_i остатки ε_i имеют одинаковую дисперсию. Если это условие применения МНК не соблюдается, то имеет место гетероскедастичность. Используя трехмерной изображение, рассмотрим отличие гомо- и гетероскедастичности.

Наличие гетероскедастичности будет сказываться на уменьшении эффективности оценок b_i , в частности, становится затруднительным использование формулы

стандартной ошибки коэффициента регрессии, предполагающей единую дисперсию остатков для любых значений фактора.

Наличие гетероскедастичности в остатках регрессии можно проверить с помощью ранговой корреляции Спирмэна. Суть проверки заключается в том, что в случае гетероскедастичности абсолютные остатки ε_i коррелированы со значениями фактора x_i . Этую корреляцию можно измерять с помощью коэффициента ранговой корреляции Спирмэна:

$$\rho = 1 - \frac{6 \sum d^2}{n(n^2 - 1)},$$

где ρ – абсолютная разность между рангами значений x_i и $|\varepsilon_i|$.

Статистическую значимость ρ можно определить с помощью t-критерия:

$$t_\rho = \frac{\rho}{\sqrt{(1 - \rho^2)}} \sqrt{(n - 2)}$$

Принято считать, что если $t_{\text{расч}} > t_{\text{табл}}$, то корреляция

между ε_i и x_i статистически значима, т. е. имеет место гетероскедастичность остатков. В противном случае принимается гипотеза об отсутствии гетероскедастичности остатков.

При построении регрессионных моделей чрезвычайно важно соблюдение четвертой предпосылки МНК – отсутствие автокорреляции остатков, т. е. распределения остатков ε_i и ε_{i-1} независимы. Автокорреляция остатков означает наличие корреляции между остатками текущих и предыдущих (последующих) наблюдений. Находится коэффициент корреляции между ε_i и ε_{i-1} , и если он окажется существенно отличным от нуля, то остатки автокоррелированы и функция плотности вероятности $F(\varepsilon)$ зависит от j -й точки наблюдения и от распределения значений остатков в других точках наблюдения.

Отсутствие автокорреляции остатков обеспечивает состоятельность и эффективность оценок коэффициентов регрессии.

До сих пор в качестве факторов рассматривались экономические переменные, принимающие количественные значения в некотором интервале. Вместе с тем может оказаться необходимым включить в модель фактор, имеющий два или более качественных уровней. Это могут быть разного рода атрибутивные признаки, такие, например, как профессия, пол, образование, климатические условия, принадлежность к определенному региону. Для того, чтобы ввести такие переменные в регрессионную модель, им должны быть присвоены те или иные цифровые метки, т. е. качественные переменные необходимо преобразовать в количественные. Такого вида сконструированные переменные в эконометрике принято называть фиктивными переменными.

Качественные признаки могут приводить к неоднородности исследуемой совокупности, что может быть учтено при моделировании двумя путями:

- регрессия строится для каждой качественно отличной группы единиц совокупности, т. е. для каждой группы в отдельности, чтобы преодолеть неоднородность единиц общей совокупности;
- общая регрессионная модель строится для совокупности в целом, учитывая неоднородность данных. В этом случае в регрессионную модель вводятся фиктивные переменные, т. е. строится регрессионная модель с переменной структурой, отражающей неоднородность данных.

Качественный фактор может иметь только два состояния, которым будут соответствовать 1 и 0. Если же число градаций качественного признака-фактора превышает два, то в модель вводится несколько фиктивных переменных, число которых должно быть меньше числа качественных градаций. Только при соблюдении этого

положения матрица исходных фиктивных переменных не будет линейно зависима и возможна оценка параметров модели.

Коэффициент регрессии при фиктивной переменной интерпретируется как среднее изменение зависимой переменной при переходе от одной категории к другой при неизменных значениях остальных параметров. На основе t-критерия Стьюдента делается вывод о значимости влияния фиктивной переменной, существенности расхождения между категориями.

Такая проверка производится с помощью статистических критериев и на их основе делается вывод о статистической надежности построенного уравнения регрессии, о пригодности модели для анализа и прогнозирования исследуемого показателя.

Для проверки надежности модели в целом используется отношение факторной

$$\frac{s_{\text{факт}}^2}{s_{\text{ост}}^2}$$

дисперсии к остаточной $s_{\text{ост}}^2$. Известно, что отношение этих дисперсий подчиняется распределению Фишера (F-распределение). Расчетное значение F-отношения сравнивается с табличным значением, которое определяется для конкретного уровня значимости α . В экономических исследованиях α принимается равным 0,05 (реже 0,01), число степеней свободы $k_1 = p, k_2 = n - p - 1$. Если $F_{\text{расч}} > F_{\text{табл}}$, то построенная модель считается статистически надежной, а следовательно, отражает закон изменения исследуемого показателя под действием факторов.

Для проверки полноту модели используется $R^2 \cdot 100\%$. Этот показатель показывает, на сколько процентов изменится вариация результативного показателя под влиянием факторов, включенных в модель.

Проверку надежности параметров уравнения регрессии проводят с использованием Т-критерия. Расчетное значение вычисляется по формуле

$$t = \frac{|a_j|}{\sigma_{a_j}}$$

$\sigma_{a_j} = \frac{s_{\text{ост}}}{\sigma_{x_j} \sqrt{n} \sqrt{1 - R^2_{j,1,2,\dots,j-1,j+1,p}}}$. Фактическое значение Т-критерия сравнивается с табличным и если $t_{\text{факт}} > t_{\text{табл}} (t_{\alpha,k}, \alpha = 0,05(0,01), k = n - p - 1)$, то тогда соответствующий коэффициент регрессии значим, т. е. отличен от нуля, а влияние J-го фактора следует считать сильным. Факторы, оказывающие несущественное влияние на исследуемый показатель, из модели исключают.

На этом этапе разрабатываются рекомендации об использовании результатов моделирования. Анализируется уравнение регрессии в натуральном масштабе: коэффициент регрессии a_j показывает, на сколько своих единиц измерения в среднем изменится исследуемый показатель, при увеличении J-го фактора на единицу своего измерения, при условии, что все остальные факторы находятся на постоянном уровне. Свободный член уравнения характеризует изменение результативного показателя за счет изменения факторов, неучтенных в модели.

В связи с тем, что факторы имеют различный физический смысл и различные единицы измерения, коэффициенты регрессии нельзя сравнивать между собой и, следовательно, трудно определить, какой из факторов оказывает наибольшее влияние. Для устранения различий в единицах измерения применяют Частные коэффициенты эластичности $\varepsilon_j = a_j \frac{\bar{x}_j}{\bar{x}_0}$ характеризующие, на сколько % в среднем изменится x_0 при увеличении J-го Фактора на 1% при фиксированном положении других факторов.

При определении степени влияния отдельных факторов необходим показатель, который бы учитывал влияние анализируемых факторов с учетом различий в уровне их вариации. Таким показателем является Коэффициент регрессии в стандартизированном

на масштабе $\beta_j = a_j \frac{\sigma_{x_j}}{\sigma_{x_0}}$. Коэффициент β_j показывает, на какую часть своего среднеквадратического отклонения изменится x_0 при изменении J-го фактора на одно свое среднеквадратическое отклонение при фиксированном значении остальных факторов. Уравнение регрессии в стандартизованном масштабе: $t_{0,1,2,\dots,p} = \sum_{j=1}^p \beta_j t_j$, где $t_j = \frac{x_j - \bar{x}_j}{\sigma_{x_j}}$.

Т. к. в стандартизованном уравнении все факторы и функция измеряются в одинаковых единицах измерения – стандартных отклонениях, то по стандартизованным коэффициентам можно судить о влиянии каждого фактора по сравнению с другими.

1. 6 Лекция №6 (2 часа).

Тема: «Методы стохастического описания и анализа особенностей процессов управления в технических системах»

1.6.1 Вопросы лекции:

1. Простейший поток, его свойства. Классификация потоков.
2. Марковские цепи, их свойства
3. Марковские процессы в инженерной практике

1.6.2 Краткое содержание вопросов:

1. Простейший поток, его свойства. Классификация потоков.

2. Марковские цепи, их свойства

Аппарат теории марковских процессов с дискретными состояниями и цепей Маркова широко используют в теории систем, в исследовании операций и других прикладных дисциплинах. Это обусловлено многими причинами, среди которых отметим следующие:

1) многие реальные технические системы имеют конечные множества возможных состояний, а их поведение в процессе функционирования адекватно моделируется Марковскими процессами,

2) теория марковских процессов с дискретными состояниями и цепей Маркова разработана настолько глубоко, что позволяет решать широкий класс прикладных задач.

Марковские процессы Представление случайных процессов графом состояний

Рассмотрим физическую систему S , в которой протекает случайный процесс с дискретными состояниями: s_1, s_2, \dots, s_i , (1)

число которых конечно (или счетно). Состояния s_1, s_2, \dots могут быть качественными (т. е. описываться словами) или же каждое из них характеризуется случайной величиной (либо случайным вектором).

Прежде всего, рассмотрим множество состояний (1) с точки зрения его структуры - возможности системы S переходить из состояния s_j в данное состояние s_i - непосредственно или через другие состояния. Для этого удобно пользоваться наглядной схемой, так называемым графом состояний. Здесь и далее мы будем отчасти пользоваться терминологией теории графов. Имеются две основные разновидности графов: неориентированные и ориентированные.

Неориентированный граф - совокупность точек (вершин графа) с соединяющими некоторые из них отрезками (ребрами графа).

Ориентированный граф - это совокупность точек (вершин) с соединяющими некоторые из них ориентированными отрезками (стрелками).

При изложении теории случайных процессов с дискретными состояниями мы будем пользоваться только ориентированными графиками. Вершины графа будут соответствовать состояниям системы. Вершину будем изображать прямоугольником, в который вписано обозначение состояния; стрелка, ведущая из вершины s_j в вершину s_i , будет обозначать возможность перехода системы S из состояния s_j в состояние s_i - непосредственно, минуя другие состояния. Стрелки графа могут изображаться не только прямолинейными, но и криволинейными отрезками (рис. 1). Сам график системы S будем обозначать буквой G .

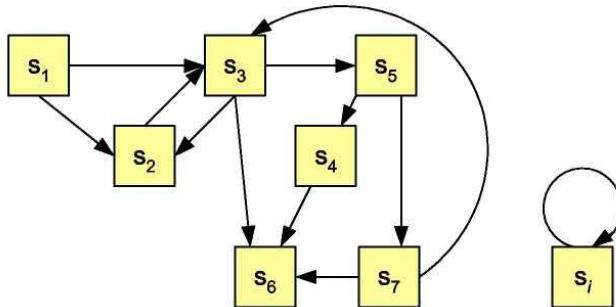


Рисунок 1 – Пример графа состояний

Переход по стрелке, ведущей из состояния s_i в него же, означает задержку системы в состоянии s_i . «Обратные стрелки» можно на графике не проставлять, так как все расчеты можно вести и без них.

Проведем некоторую необходимую для дальнейшего классификацию состояний. Состояние s_i называется источником, если система S может выйти из этого состояния, но попасть в него обратно уже не может, т. е. на графике состояний в состояние s_i не ведет ни одна стрелка. На рисунке 1 состояние s_1 является источником.

Состояние s_i называется концевым (или поглощающим), если система S может попасть в это состояние, но выйти из него уже не может. Для графа состояний это означает, что из состояния s_i не ведет ни одна стрелка (для графа, изображенного на рисунке 1, состояние s_6 поглощающее).

Если система S может непосредственно перейти из состояния s_i в состояние s_j то состояние s_j - называется соседним по отношению к состоянию s_i .

Состояние s_i называется транзитивным, если система S может войти в это состояние и выйти из него, т. е. на графике состояний есть хотя бы одна стрелка, ведущая в s_i и хотя бы одна стрелка, ведущая из s_i . На рисунке 1 все состояния, кроме s_1 и s_6 , являются транзитивными.

Для полноты картины можно рассматривать также и «изолированные» состояния. Состояние s_i называется изолированным, если из него нельзя попасть ни в одно из других состояний и в него нельзя попасть ни из какого другого состояния.

Наряду с отдельными состояниями системы S в ряде задач практически бывает нужно рассматривать подмножества ее состояний.

Обозначим W множество всех состояний системы S (конечное или бесконечное, но счетное) и рассмотрим его подмножество $V \subset W$. Подмножество V называется замкнутым (концевым), если система S , попав в одно (или находясь в одном) из состояний $s_i \in V$, не может выйти из этого подмножества состояний. Концевое подмножество состояний может включать в себя поглощающее состояние, а может и не включать.

Подмножество состояний $V \subset W$ называется связным или эргодическим, если из любого состояния, входящего в него, можно попасть в любое другое состояние, принадлежащее этому подмножеству. Эргодическим может быть и все множество W .

состояний системы S . В эргодическом множестве состояний нет ни источников, ни поглощающих состояний.

Подмножество состояний V называется транзитивным, если система S может войти в это подмножество и выйти из него, т. е. из любого состояния $s_i \subset V$ можно (за то или другое число перескоков) выйти из этого подмножества.

Случайный процесс, протекающий в системе S , можно трактовать как процесс блуждания системы по множеству состояний W . Если подмножество $V \subset W$ является концевым, то, попав в него, система будет продолжать блуждание уже по этому подмножеству состояний V . Если все множество эргодично, то блуждание будет происходить по всем его состояниям.

На практике очень часто встречаются системы, состояния которых образуют цепь (рисунок 2), в которой каждое состояние s_i (кроме двух крайних s_0 и s_n) связано прямой и обратной связью с двумя соседними s_{i-1}, s_{i+1} , а каждое из двух крайних связано прямой и обратной связью только с одним соседним.

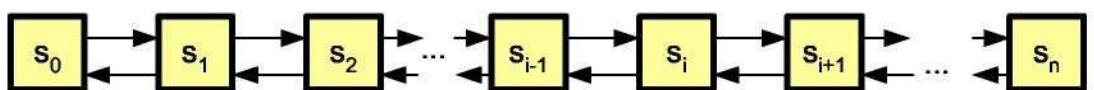


Рисунок 2 - Схема процесса гибели и размножения

Такая схема случайного процесса называется схемой гибели и размножения, а сам процесс — процессом гибели и размножения.

Если на графике состояний системы S стрелки, ведущие справа налево, отсутствуют, то говорят о процессе «чистого размножения», в противоположном случае — о процессе «чистой гибели».

Процесс гибели и размножения может в некоторых случаях иметь не конечное число состояний: $s_1, s_2, \dots, s_i, \dots, s_n$, а бесконечное (счетное): $s_1, s_2, \dots, s_i, \dots$.

При анализе случайных процессов, протекающих в системах с дискретными состояниями, важную роль играют вероятности состояний.

Обозначим $S(t)$ состояние системы S в момент t . Вероятностью i -го состояния в момент t называется вероятность события, состоящего в том, что в момент t система S будет в состоянии s_i . Обозначим ее $p_i(t)$:

$$p_i(t) = P\{S(t) = s_i\}, \quad (2)$$

где $S(t)$ — случайное состояние системы S в момент t . Очевидно, что для системы с дискретными состояниями $s_1, s_2, \dots, s_i, \dots$, в любой момент t сумма вероятностей состояний равна единице:

$$\sum_i p_i(t) = 1, \quad (3)$$

как сумма вероятностей полной группы несовместных событий.

В ряде задач практики нас интересует так называемый установившийся или стационарный режим работы системы, который в ней устанавливается, когда от начала процесса прошло достаточно большое время t . Например, процесс изменения напряжения в сети питания технического устройства, пройдя сразу после включения через ряд колебаний, по прошествии времени, устанавливается. Аналогично этому и в некоторых случайных процессах по прошествии достаточно большого времени t устанавливается стационарный режим, во время которого состояния системы хотя и меняются случайным

образом, но их вероятности $p_i(t)$ ($i = 1, 2, \dots$) остаются постоянными. Обозначим эти постоянные вероятности P_i .

$$P_i = \lim p_i(t) \quad (4)$$

Вероятности $p_i(i = 1, 2, \dots)$, если они существуют, называются финальными (пределными) вероятностями состояний. Финальную вероятность P_i можно истолковать как среднюю долю времени, которую в стационарном режиме проводит система S в состоянии s_i . В дальнейшем будет показано, при каких условиях финальные вероятности существуют и какими они могут быть для разных состояний и подмножеств состояний.

Введем очень важное для дальнейшего понятие марковского случайного процесса.

Случайный процесс, протекающий в системе S , с дискретными состояниями $s_1, s_2, \dots, s_i, \dots$, называется марковским, если для любого момента времени t_0 вероятность каждого из состояний системы в будущем (при $t > t_0$) зависит только от ее состояния в настоящем (при $t = t_0$) и не зависит от того, когда и как она пришла в это состояние; т. е. не зависит от ее поведения в прошлом (при $t < t_0$).

Не следует понимать марковское свойство случайного процесса как полную независимость «будущего» от «прошлого». В общем случае «будущее» зависит от «настоящего», т. е. вероятности $p_i(t)$ при $t > t_0$ зависят от того, в каком состоянии s_i находится система в настоящем (при $t = t_0$); само же это «настоящее» зависит от «прошлого», от того, как вела себя система S при $t < t_0$. Это можно сформулировать следующим образом: для марковского случайного процесса «будущее» зависит от «прошлого» только через «настоящее» (рисунок 3). При фиксированном «настоящем» условные вероятности всех состояний системы в «будущем» не зависят от предыстории процесса, т. е. от того, когда и как система S к моменту t_0 пришла в состояние

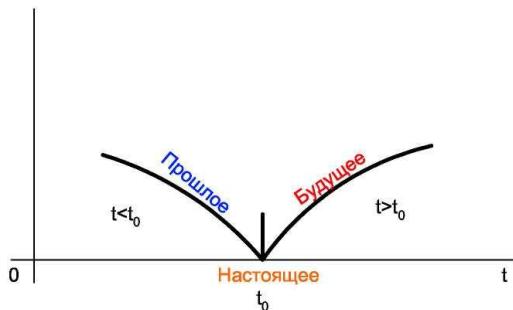


Рисунок 3 – Схема марковского свойства случайного процесса

«Настоящее» может быть задано не одним каким-то состоянием s_i , а целым подмножеством состояний $V \subset W$, где W – множество всех возможных состояний системы.

Подчеркнем также, что «настоящее» может быть задано не только одним состоянием системы S в момент t_0 ; в него при желании можно включить и те элементы из «прошлого», от которых, при заданном «настоящем», зависит будущее. Например, вероятности состояний в «будущем» могут зависеть не только от состояния s_i системы в настоящем, но и от того, из какого состояния s_i система перешла к моменту t_0 в состояние s_i . В этом случае настоящее характеризуется не только состоянием s_i , в которое система перешла к моменту t_0 , но и состоянием s_j , из которого она перешла в s_i . Вводя в состав параметров, характеризующих настоящее состояние системы, те параметры из прошлого, от которых зависит будущее, можно, как говорится, «марковизировать» многие немарковские случайные процессы, но, как правило, это приводит к сильному усложнению математического аппарата.

3. Марковские процессы в инженерной практике

Марковские случайные процессы с дискретными состояниями и дискретным временем

Пусть имеется система S с дискретными состояниями $s_1, s_2, \dots, s_i, \dots, s_n$. Предположим, что случайные переходы («перескоки») системы из состояния в состояние могут происходить только в определенные моменты времени t_0, t_1, t_2, \dots . Эти моменты мы будем называть шагами процесса; $t_0=0$ - его началом. Сам процесс представляет собой случайное блуждание системы S по состояниям. После первого шага система может оказаться в одном (и только в одном) из своих возможных состояний:

$$s_1^{(1)}, s_2^{(1)}, \dots, s_i^{(1)}, \dots, s_n^{(1)}$$

На втором шаге - $s_1^{(2)}, s_2^{(2)}, \dots, s_i^{(2)}, \dots, s_n^{(2)}$, на k -м шаге $s_1^{(k)}, s_2^{(k)}, \dots, s_i^{(k)}, \dots, s_n^{(k)}$ (число состояний в общем случае может быть бесконечным, но счетным).

Здесь же для простоты ограничимся конечным числом n состояний.

Предположим, что граф состояний системы S имеет вид, представленный на рисунке 4. Процесс блуждания системы S по состояниям можно представить как последовательность или «цепь» событий. Эта последовательность состоит в том, что в начальный момент $t_0=0$ система находится в одном из состояний (например, в состоянии $s_1^{(0)}$), в момент первого шага перешла из него скачком в состояние $s_5^{(1)}$, из которого на втором шаге перешла в $s_3^{(2)}$, на третьем шаге перешла в $s_2^{(3)}$ и т. д.

«Траектория» системы, блуждающей по состояниям s_1, s_5, s_3, s_2 показана на рисунке 4 жирными линиями. На каких-то шагах система может задерживаться в том или другом из своих состояний, $s_i^{(k)} = s_i^{(k+1)}$ (это показано «возвратной стрелкой» на рисунке 4) или же вернуться в него после ряда шагов.

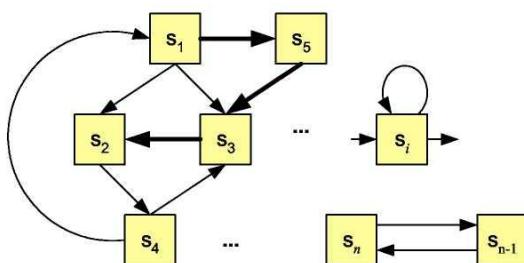


Рисунок 4. Граф состояний системы S

«Траектория» блуждания системы по графу состояний, изображенная на рисунке 4 жирными линиями, представляет собой не что иное, как реализацию случайного процесса, полученную в результате одного опыта. При повторении опыта, естественно, реализации в общем случае не совпадают.

Рассмотрим общий случай. Пусть происходит случайный процесс в системе S с дискретными состояниями $s_1, s_2, \dots, s_i, \dots, s_n$, которые она может принимать в последовательности шагов с номерами $0, 1, 2, \dots, k, \dots$.

Случайный процесс представляет собой последовательность событий вида $\{S(k) = s_i\}$ ($i = 1, 2, \dots, n$; $k = 0, 1, 2, \dots$). Наиболее важной ее характеристикой являются вероятности состояний системы

$$\mathbf{P}\{S(k)=s_i\} \quad (i=1,2,\dots,n; k=0,1,2,\dots), \quad (5)$$

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

Распределение вероятностей (5) представляет собой не что иное, как одномерный закон распределения случайного процесса $S(t)$, протекающего в системе S с «качественными» дискретными состояниями и дискретным временем $t_0, t_1, t_2, \dots, t_k$.

Процесс, протекающий в такой системе S , называется марковским процессом с дискретными состояниями и дискретным временем (или, короче, марковской цепью). При этом выполняется условие: для любого фиксированного момента времени (любого шага k_0) условные вероятности состояний системы в будущем (при $k > k_0$) зависят только от состояния системы в настоящем (при $k = k_0$) и не зависят от того, когда (на каком шаге, при $k < k_0$) и откуда система пришла в это состояние. Марковская цепь представляет собой разновидность марковского процесса, в котором будущее зависит от прошлого только через настоящее.

Цепь, в которой условные вероятности состояний в будущем зависят только от состояния на данном, последнем, шаге и не зависят от предыдущих, иногда называют простой цепью Маркова. В отличие от такой, где будущее зависит от состояний системы не только в настоящем на данном шаге, но и от ее состояний на нескольких предыдущих шагах; такую цепь называют сложной цепью Маркова. Сам А. А. Марков рассматривал сложные цепи, построенные на материале буквенных последовательностей, взятых из текста пушкинского «Евгения Онегина».

Из определения марковской цепи следует, что для нее вероятность перехода системы S в состояние s_i на $(k+1)$ -м шаге зависит только от того, в каком состоянии s_i находилась система на предыдущем k -м шаге и не зависит от того, как она вела себя до этого k -го шага.

Основной задачей исследования марковской цепи является нахождение безусловных вероятностей нахождения системы S на любом k -м шаге в состоянии s_i ; обозначим эту вероятность $p_i(k)$:

$$p_i(k) = P\{S(k) = s_i\} \quad (i = 1, 2, \dots, n; k = 0, 1, 2, \dots). \quad (6)$$

Для нахождения этих вероятностей необходимо знать условные вероятности перехода системы S на k -м шаге в состояние s_i , если известно, что на предыдущем $(k-1)$ -м шаге она была в состоянии s_j . Обозначим эту вероятность

$$p_{ij}(k) = P\{S(k) = s_i | S(k-1) = s_j\} \quad (i, j = 1, 2, \dots, n; k = 0, 1, 2, \dots). \quad (7)$$

Вероятности $p_{ij}(k)$ называются переходными вероятностями марковской цепи на k -м шаге. Вероятность $p_{ij}(k)$ есть вероятность того, что на k -м шаге система задержится (останется) в состоянии s_i .

Переходные вероятности $p_{ij}(k)$ можно записать в виде квадратной таблицы (матрицы) размерности $n \times n$:

$$\|p_{ij}(k)\| = \begin{vmatrix} p_{11}(k) & p_{12}(k) & \dots & p_{1j}(k) & \dots & p_{1n}(k) \\ p_{21}(k) & p_{22}(k) & \dots & p_{2j}(k) & \dots & p_{2n}(k) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ p_{i1}(k) & p_{i2}(k) & \dots & p_{ij}(k) & \dots & p_{in}(k) \\ \dots & \dots & \dots & \dots & \dots & \dots \\ p_{n1}(k) & p_{n2}(k) & \dots & p_{nj}(k) & \dots & p_{nn}(k) \end{vmatrix}. \quad (8)$$

$(k = 0, 1, 2, \dots)$

По главной диагонали матрицы (8) стоят вероятности задержки системы в данном состоянии s_j ($j = 1, \dots, n$) на k -м шаге.

$$p_{11}(k), p_{22}(k), \dots, p_{ii}(k), \dots, p_{nn}(k). \quad (9)$$

Так как на каждом шаге система S может находиться только в одном из взаимно исключающих состояний, то для любой k -й строки матрицы (8) сумма всех стоящих в ней вероятностей равна единице:

$$\sum_{j=1}^n p_{ij}(k) = 1. \quad (10)$$

Матрица, обладающая таким свойством, называется стохастической. Естественно, что все элементы стохастической матрицы отвечают условию $0 \leq p_{ij}(k) \leq 1$. В силу условия (10) можно в матрице (8) не задавать вероятности задержки, а получать их как дополнения до единицы всех остальных членов строки:

$$p_{ii}(k) = 1 - \sum_{j=1}^n p_{ij}(k). \quad (11)$$

Чтобы найти безусловные вероятности $p_i(k)$, недостаточно знать матрицу переходных вероятностей (8); нужно еще знать начальное распределение вероятностей, т. е. вероятности состояний $p_i(0)$, соответствующие началу процесса - моменту $t_0 = 0$:

$$p_1(0), p_2(0), \dots, p_i(0), \dots, p_n(0), \quad (12)$$

в сумме образующие единицу:

$$\sum_{i=1}^n p_i(0) = 1 \quad (13)$$

Если известно, что в начальный момент система S находится во вполне определенном состоянии s_i , то вероятность $p_i(0)$ этого состояния в формуле (13) равна единице, а все остальные - нулю:

$$p_i(0), p_1(0) = p_2(0) = \dots = p_{i-1}(0) = p_{i+1}(0) = \dots = p_n(0) = 0. \quad (14)$$

Цепь Маркова называется однородной, если переходные вероятности $p_{ij}(k)$ не зависят от номера шага k : $p_{ij}(k) = p_{ij}$. Матрица переходных вероятностей для однородной цепи Маркова имеет вид:

$$\|p_{ij}\| = \begin{vmatrix} p_{11} & p_{12} \dots p_{1j} \dots p_{1n} \\ p_{21} & p_{22} \dots p_{2j} \dots p_{2n} \\ \dots & \dots \\ p_{i1} & p_{i2} \dots p_{ij} \dots p_{in} \\ \dots & \dots \\ p_{n1} & p_{n2} \dots p_{nj} \dots p_{nn} \end{vmatrix} \quad (14)$$

При выводе формул для вероятностей состояний, в целях простоты записи, будем рассматривать только однородные цепи Маркова (в случае, когда цепь неоднородна, можно все переходные вероятности в формулах просто положить зависящими от номера шага k).

При нахождении вероятностей состояний марковской цепи на k -м шаге $p_i(k)$ ($k = 1, 2, \dots$) удобно бывает пользоваться так называемым размеченным графом состояний системы S , где возле каждой стрелки, ведущей из состояния s_i в состояние s_j , проставлена переходная вероятность p_{ij} . Вероятности задержки на размеченном графе не проставляются, а просто получаются дополнением до единицы суммы вероятностей, стоящих у всех стрелок, ведущих из данного состояния s_i .

Теперь покажем, как найти для однородной цепи Маркова безусловную вероятность нахождения системы S на k -м шаге в состоянии s_j ($j = 1, 2, \dots, n$)

$$p_j(k) = \mathbf{P}\{S(k)=s_j\}, \quad (15)$$

если задана матрица переходных вероятностей $\|p_{ij}\|$ (или, что равнозначно, размеченный граф состояний) и начальное распределение вероятностей

$$p_i(0) \quad (i = 1, 2, \dots, n).$$

$$\sum_{j=1}^n p_{ij}(k) = 1 \quad . \quad (16)$$

Сделаем гипотезу, состоящую в том, что в начальный момент ($k=0$) система находилась в состоянии s_i . Вероятность этой гипотезы известна и равна $p_i(0)=P\{S(0)=s_i\}$. В предположении, что эта гипотеза имеет место, условная вероятность того, что система S на первом шаге будет в состоянии s_j , равна переходной вероятности $p_{ij}(k)=P\{S(1)=s_j | S(0)=s_i\}$.

По формуле полной вероятности получим:

$$p_j(1) = \sum_{i=1}^n P\{S(1)=s_j | S(0)=s_i\} \cdot P\{S(0)=s_i\} = \sum_{i=1}^n p_{ij} p_i(0), \quad (j = 1, \dots, n) \quad (17)$$

Таким образом, мы нашли распределение вероятностей системы S на первом шаге. Теперь у нас есть все необходимое для того, чтобы найти распределение вероятностей на втором шаге, которое для цепи Маркова зависит только от распределения вероятностей на первом шаге и матрицы переходных вероятностей.

Опять сделаем гипотезу, состоящую в том, что на первом шаге система находится в состоянии s_i . Вероятность этой гипотезы нам уже известна и равна $p_i(1)=P\{S(1)=s_i\}$. При этой гипотезе условная вероятность того, что на втором шаге система S будет в состоянии s_j , равна:

$$p_{ij}(k) = P\{S(2)=s_j | S(1)=s_i\}$$

По формуле полной вероятности находим

$$p_j(2) = \sum_{i=1}^n p_i(1)p_{ij}, \quad (j = 1, 2, \dots, n) \quad (18)$$

Таким образом, мы выразили распределение вероятностей (18) на втором шаге через распределение вероятностей на первом шаге и матрицу $\|p_{ij}\|$. Переходя таким же способом от $k = 2$ к $k = 3$ и т. д., получим рекуррентную формулу:

$$p_j(k) = \sum_{i=1}^n p_i(k-1)p_{ij}, \quad (k = 1, 2, \dots, n; j = 1, 2, \dots, n) \quad (19)$$

При некоторых условиях в цепи Маркова с возрастанием k (номера шага) устанавливается стационарный режим, в котором система S продолжает блуждать по состояниям, но вероятности этих состояний уже от номера шага не зависят. Такие вероятности называются предельными (или финальными) вероятностями цепи Маркова.

Например, если рассматривать ЭВМ в двух состояниях: s_1 - исправна, s_2 - не исправна, то имеет место следующая динамика изменения вероятностей (при начальных условиях):

$$p_1(0) = 1, p_2(0) = 0; p_1(1) = 0,7; p_1(2) = 0,61; p_1(3) = 0,583; p_1(4) = 0,5749.$$

Ниже мы покажем, что в этом случае $p_1 = \lim_{k \rightarrow \infty} p_1(k) = 0,4/(0,4+0,3) = 0,5714$. Таким образом, в рассматриваемой системе стационарный режим наступит практически через четыре шага.

Можно убедиться в том, что в этом примере финальные вероятности не зависят от начальных условий.

Сформулируем условия существования стационарного режима для системы S с конечным числом состояний n , в которой протекает марковский случайный процесс с дискретными состояниями и дискретным временем (цепь Маркова):

1. Множество всех состояний W системы S должно быть эргодическим.

2. Цепь Маркова должна быть однородной:

$$p_{ij}(k) = p_{ij} \quad (20)$$

3. Цепь Маркова должна быть «достаточно хорошо перемешиваемой» (не должна быть «циклической»).

Цепи Маркова, отвечающие этим условиям, будем называть эргодическими цепями Маркова.

1. 7 Лекция №7 (2 часа).

Тема: «Техника экспериментальных измерений. Основные положения теории погрешностей. Адекватность моделей»

1.7.1 Вопросы лекции:

1. Основные положения теории измерений. Техника экспериментальных измерений.
2. Основные положения теории погрешностей.
3. Адекватность моделей

1.7.2 Краткое содержание вопросов:

1. Основные положения теории измерений. Техника экспериментальных измерений.

2. Основные положения теории погрешностей.

Измерение - это процесс нахождения физических величин, параметров, характеристик опытным путем с помощью средства измерения. Найденное значение называют результатом измерения. Измерения посредством измерительного устройства заключается в сравнении измеряемой величины с ее однородной физической величиной, принятой за единицу измерения. Результат выражается числом.

Измерение проводится различными методами, например: методом непосредственной оценки. В этом методе измерения значение измеренной величины определяют непосредственно по отчетному устройству измерительного прибора, предварительного проградуированного по мере. Т.е. при измерении непосредственной оценки меры не происходит, а она передается через предварительно проградуированную оценку. Используют также метод сравнения с мерой. В этом методе сравниваются с однородной величиной воспроизведенной мерой, размер которой известен и который определяет результат измерения.

Технические средства измерения, имеющие нормированные метрологические характеристики и оказывающие определенное влияние на результаты и погрешности измерений, называют средством измерения. В зависимости от назначения средства измерения делятся на следующие виды:

- а) мера – средство измерения, предназначенное для воспроизведения физической величины данного вида;
- б) измерительный прибор – средство измерения,рабатывающее сигнал измерительной информации в форме, доступной для восприятия;
- в) измерительный преобразователь – средство измерения,рабатывающее сигнал измерительной информации в форме, удобной для передачи дальнейшего преобразования и обработки, но не поддающейся непосредственному восприятию. К ним относятся: усилители, входные и выходные делители, измерительные трансформаторы.

Согласно механическим функциям средства измерения подразделяются:

- 1) эталоны - средства измерения, обеспечивающие воспроизведение и хранение единицы измерения и официально утвержденные в качестве эталона;

- 2) образцовые средства измерения – это меры или измерительные приборы, служащие для проверки по ним других средств измерения и утвержденные официально в качестве образцовых;
- 3) рабочие средства измерения.

Различают 2 метода измерения:

Прямое измерение – это измерение, при котором искомое значение величины находят из опытных данных (измерение тока и т.д.). Косвенное измерение – это измерение, при котором измеряется не сама величина, а величина функционально связанная с ней, по значению которой и известной функциональной зависимости определяется измеряемая величина. Например, объем детали, определяемый по результатам измерения геометрических размеров.

Погрешности измерений, погрешности измерительных приборов.

При всяком измерении неизбежны отклонения результатов измерения от истинного значения измеряемой величины, обусловленные различными причинами. Эти отклонения – погрешности измерений. Погрешности классифицируют по причинам возникновения, условиям проведения измерений, характеру появления.

В зависимости от причин возникновения различают погрешности измерения:

- а) погрешность метода измерения – составляющая погрешность измерения, происходящая от несовершенства метода измерения (методическая погрешность);
- б) инструментальная (аппаратурная) погрешность, составляющая погрешность, измерения зависящая от погрешности применяемого средства измерения (от его точности, класса прибора);
- в) субъективная – составляющая общей погрешности измерения, обусловленная несовершенством органов чувств, а также небрежности в процессе измерения и фиксации результата.

По условиям проведения измерения, т.е. зависимости результатов измерения от внешних условий окружающей среды, различают основную и дополнительную погрешности:

- основная погрешность средства измерения, используемая в нормированных климатических условиях. Эта погрешность указывается в паспортных данных или в технических условиях на измерительный прибор.
- дополнительная погрешность – погрешность, вызванная отклонением условий измерения от номинальных. Она может превосходить основную погрешность в несколько раз, для ее учета используют графики, таблицы, формулы, которые даны в документации по эксплуатации прибора.

По характеру появления различают:

- систематические погрешности измерения, являются результатом неправильной градуировки, калибровки прибора;
- случайные погрешности измерения – это составляющая погрешности, появление которой носит случайный характер;
- грубые погрешности.

Многообразные причины появления погрешностей приводят к тому, что, многократно снятые характеристики средств измерений или серии однотипных приборов образуют некоторую область значений. В теории измерений для этого используется понятие полосы неопределенности или полосы погрешностей средства измерения. Средняя линия такой полосы принимается за номинальную характеристику приборов, которая указывается в паспорте и используется для определения результатов измерения. Поэтому погрешность средства измерения есть разница между реальной и номинальной его характеристиками, т.е. является не числом, а функцией измеряемой величины.

$$Y_H = f(x).$$

Разница между реальной и номинальной характеристиками, найденными при заданном значении измеряемой величины, называется абсолютной погрешностью:

$$\Delta_Y = Y_F - Y_N.$$

Знак абсолютной погрешности принимается положительным, если реальная характеристика проходит выше номинальной. Абсолютная погрешность не может служить показателем точности измерений, поскольку точность измерения будет изменяться в зависимости от значения измеряемой величины. Поэтому для характеристики точности результатов измерений используется относительная погрешность, выражаемая в относительных единицах или процентах:

$$\gamma = \frac{\Delta_Y}{Y} ; \gamma = \frac{\Delta_Y}{Y} \cdot 100\%$$

Но эта очень наглядная характеристика точности результата измерения не годится для нормирования погрешности средств измерений. Поэтому для указания и нормирования погрешности средств измерений используется еще одна разновидность погрешности, так называемая *приведенная* погрешность. Она определяется как отношение абсолютной погрешности, выраженной в единицах входной D_X или выходной D_Y величин, к протяженности диапазона изменения соответственно входной X_k или выходной Y_k величины прибора или преобразователя и выражается в относительных единицах или в процентах:

$$\gamma_{\text{пр}} = \frac{\Delta_X}{X_k} = \frac{\Delta_Y}{Y_k}$$

Основное отличие приведенной погрешности от относительной погрешности состоит в том, что D_X или D_Y относится не к переменной текущей величине x или y , а к постоянной величине протяженности диапазона. Приведенная погрешность удобна тем, что для многих систем измерения она имеет одно и то же значение, как для всех точек каждого диапазона, так и для всех его областей, то есть ее удобно использовать для нормирования свойств систем измерения.

Аддитивные и мультипликативные погрешности.

Аддитивные и мультипликативные погрешности используются для описания границ полосы погрешностей средства измерений. При поверке или градуировке средств измерений получают ряд значений входной величины x_i и ряд соответствующих им значений выходной величины y_i . Если эти данные нанести на график с координатами x и y , то полученные точки разместятся в границах некоторой полосы. В том случае, когда эти точки лежат в границах линий, параллельных друг другу, то абсолютная погрешность средства измерений во всем его диапазоне измерений ограничена постоянным, не зависящим от текущего значения входной величины x пределом $\pm D_0$, то такая погрешность называется *аддитивной*, т. е. получаемой путем сложения, или *погрешностью нуля*. Это понятие одинаково применимо как к случайным, так и к систематическим погрешностям.

Примерами систематических аддитивных погрешностей являются погрешности от постороннего груза на чашке весов, от неточной установки прибора на нуль перед измерением, от термо-ЭДС в цепях постоянного тока и т. п. Для устранения таких погрешностей во многих средствах измерений предусмотрено механическое или электрическое устройство для установки нуля (корректор нуля).

Если ширина полосы погрешностей возрастает пропорционально росту входной величины x , а при $x = 0$ также равна нулю, то такая погрешность называется *мультипликативной*. Мультипликативная погрешность, получаемая путем умножения, называют также *погрешностью чувствительности* вне зависимости от того, является ли погрешность случайной или систематической. Причинами возникновения мультипликативных погрешностей могут быть: изменение коэффициента усиления

усилителя, измерение жесткости мембранны датчика манометра или пружинки прибора, изменение спорного напряжения в цифровом вольтметре и т. д.

Погрешность квантования. Это специфическая разновидность погрешности, возникающая в цифровых приборах и дискретных преобразователях. Вследствие того, что измеряемая величина x случайным образом может принимать любые промежуточные значения, погрешность квантования также случайным образом принимает значения в интервале от $+D_0$ до $-D_0$. Поэтому погрешность квантования является инструментальной случайной аддитивной статической погрешностью, так как не зависит ни от текущего значения результата измерения величины x , ни от скорости изменения x во времени.

Методы нормирования погрешностей средств измерений.

Различные средства измерений (СИ) (измерительные приборы и преобразователи, датчики и т.д.) обладают погрешностями, характер проявления которых определяется формой границ полосы погрешности средств измерений (аддитивная, мультиплекционная, или иной другой, более сложной). У каждого конкретного СИ имеется случайная и систематическая составляющие погрешности, причем их соотношение также может быть различным.

Для оценки погрешности, которую внесет данное СИ в конкретный результат, используют *нормированные* значения погрешности. Под нормированным значением понимаются погрешности, являющиеся предельными для данного типа СИ. При этом как систематическая, так и случайная составляющие погрешности отдельных экземпляров СИ одного и того же типа могут различаться, однако в целом для этого типа СИ погрешности не превосходят гарантированного значения. Таким образом, нормируется основная и дополнительная погрешности. Именно эти границы основной погрешности, а также коэффициентов влияния и заносятся в паспорт каждого экземпляра СИ.

Правила, согласно которым назначаются эти границы значений погрешности и форма записи, основываются на системе стандартов, обеспечивающих единство измерений.

Класс точности средств измерений.

Это характеристика, определяющая гарантированные границы значений основных и дополнительных погрешностей, а также другие свойства средств измерений, влияющих на точность. Соответствие погрешности СИ приписанному им классу точности во время эксплуатации проверяется при периодических поверках. Если погрешность оказывается меньше нормированных значений, то СИ продолжает эксплуатироваться, если нет, то подлежит ремонту и регулировке.

Основные способы установления пределов допускаемых погрешностей и обозначения классов точности средств измерений установлены ГОСТ 8.401—80.

Оценка инструментальной статистической погрешности результата измерения по паспортным данным средства измерения.

Результат измерения должен иметь оценку его интервал неопределенности, т. е. степень достоверности. В любой форме представления результатов измерений сообщение о любом результате измерений обязательно должно сопровождаться указанием его погрешности.

Погрешность результата прямого однократного измерения зависит от многих факторов, но, в первую очередь, определяется, естественно, погрешностью используемых средств измерений. Поэтому в первом приближении погрешность результата измерения можно принять равной погрешности, которой в данной точке диапазона измерений характеризуется используемое средство измерений.

Так как погрешности средств измерений изменяются в диапазоне, то вычисление должно производиться по формулам, соответствующие формам границ полосы погрешностей. Вычисляться должна как абсолютная, так и относительная погрешности результата измерения, так как первая из них нужна для округления результата и его

правильной записи, а вторая — для однозначной сравнительной характеристики его точности.

Для разных характеристик нормирования погрешностей СИ эти вычисления производятся по-разному.

1. Класс точности прибора указан в виде одного числа g_s , заключенного в кружок. Тогда относительная погрешность результата (в процентах) $g(x) = g_s$, абсолютная его

$$\Delta(x) = \frac{g_s \cdot x}{100}$$

погрешность как:

2. Класс точности прибора указан одним числом g_0 (без кружка). Тогда абсолютная погрешность результата измерения вычисляется как:

$$\Delta(x) = \frac{g_0 \cdot X_k}{100};$$

где X_k — предел измерений, на котором оно производилось.

3. Класс точности прибора указан двумя числами в виде g_k/g_n . В этом случае удобнее вычислить относительную погрешность результата по формуле:

$$g(x) = g_k + g_n \cdot \left(\frac{X_k}{x} - 1 \right),$$

затем найти абсолютную погрешность как:

$$\Delta(x) = \frac{g(x) \cdot x}{100}$$

При использовании этих формул полезно помнить, что в формулы для определения $g(x)$ значения g_s , g_0 , g_n и g_k подставляются в процентах, поэтому и относительная погрешность результата измерения получается также в процентах.

Пример. На вольтметре класса точности 2.5, с пределом измерений 300 В был получен отсчет измеряемого напряжения $x=267.5$ В. Требуется провести оценку погрешности результата измерения.

Находят абсолютную погрешность:

$$\Delta(x) = \frac{g_s \cdot X_k}{100} = \frac{2.5 \cdot 300}{100} = 7.5 \text{ В};$$

относительная погрешность определяют по уравнению:

$$g(x) = \frac{\Delta(x)}{x} \cdot 100 = \frac{7.5}{267.5} \cdot 100 = 2.81\% \approx 2.8\%$$

Таким образом, окончательно получаем: измерение проведено с относительной погрешностью $g(x)=2.8\%$. Измеренное напряжение $x=(268.5 \pm 7.5)$ В.

3. Адекватность моделей

Вопрос о необходимой и достаточной степени соответствия объекту — оригиналу или адекватности модели относится к числу важнейших в сфере модельной методологии. Под эффективностью понимают практическую полезность. Процесс моделирования неизбежно протекает в условиях диалектического взаимодействия двух противостоящих друг другу тенденций. С одной стороны, исследователь всегда стремится к возможно более полному и точному воспроизведению в модели свойств и характеристик объекта. Неизбежным следствием такого подхода является рост сложности, которая проявляется в числе переменных, числе учитываемых связей и влияний, повышении требования к точности исходных данных и т.д. Именно эта сторона дела — требование полноты соответствия модели объекту — оригиналу рассматривается некоторыми авторами как мера совершенства модели. Однако практика показала неопровергимо: эффективность модели находится в обратной зависимости от её сложности, быстро убывая с ростом последней.

Определить математическим путем наилучшее сочетание полноты-точности создаваемой модели с одной стороны и простоты с другой, практически никогда не

удается из-за неформализуемости и неоднозначности большей части подлежащих учету факторов.

Пара задача-объект в основном определяет номенклатуру подлежащих учету переменных объекта; параметры, входящие в модель, число и характер связей между ними, требования к точности данных и ряд других важнейших характеристик модели. Решающим фактором эффективности сейчас оказывается математический аппарат. Эффективность модели зависит и от такого субъективного момента, как профессиональные качества и уровень подготовки исследователя – исполнителя.

Таким образом, можно сделать заключение: наилучшее в практическом отношении качество или эффективность любой модели достигается как разумный компромисс между близостью модели к оригинал (адекватностью) и простотой, обеспечивающей возможность и удобство использования модели по её прямому назначению; чрезмерная точность модели на практике не менее вредна, чем её неполнота и грубость.

Математическая модель изучаемого процесса или объекта является основой, фундаментом всего исследования. Тем не менее, на сегодняшний день не существует и, по-видимому, не может существовать науки о моделировании реальных процессов и явлений окружающего мира – точно так же, как не существует науки о том, как совершать открытия, изобретения, создавать новые методы научного поиска. Даже математика, одна из наук, которая в большей степени использует дедукцию, своему прогрессу обязана таким “ненаучным” приемам, как интуиция, догадка, фантазия (индуктивному способу мышления).

Моделирование объектов и явлений реальности (на сегодняшний день) в большой степени представляет искусство, а искусству учат на опыте. Человечество обладает таким опытом. Это опыт классиков естествознания, опыт представителей естественных наук, эксплуатирующих для своих целей математический аппарат, и т. д.

В каждом конкретном случае качество модели во многом зависит от способностей исследователя понять существование, физику изучаемого процесса и создать его адекватное математическое описание. Математику привлекают, когда сложен изучаемый или управляемый процесс. Сложность обычно состоит в огромном числе характеристик, его описывающих, и большом числе связей между ними. И задача заключается не только в том, чтобы создать адекватное математическое описание изучаемого процесса, т. е. его модель, но и разработать методику работы с нею. С громоздкими многопараметрическими моделями трудно проводить исследования, поэтому математики вынуждены были при формализации реального процесса отбрасывать многие, на их взгляд менее существенные связи, загрублять математическое описание.

Необходимо обладать незаурядной интуицией для определения, что важно с точки зрения интересующих исследователя вопросов, что – нет. Однако при решении серьезных практических задач невозможно полагаться лишь на интуицию и опыт небольшой группы исследователей, необходима методика, позволяющая с большой степенью достоверности определить адекватность модели и реальности, ею описываемой, область возможного ее применения и круг вопросов, для исследования которых они пригодны. Необходима «система знаний», которая позволила бы, используя накопленный опыт и определенные принципы, выработанные на его основе, а также доказанные или установленные на их базе положения, создавать модели изучаемых процессов, проводить их анализ и определять пути их дальнейшего использования.

Модельное исследование, как любой другой вид осознанной целенаправленной деятельности

1) начинается с возникновения **проблемы** – потребности изменить в лучшую сторону существующее либо ожидаемое положение вещей в той или иной области. Источник проблемы – предшествующее развитие данной области или же внешние факторы.

2) Осмысление или конкретизация проблемы приводит к формулировке **цели или системы целей** как желательного результата будущей деятельности по решению проблемы.

3) Поставленная цель должна быть соотнесена с реальными возможностями ее достижения, т.е. с ресурсами (материальными и другими). Сопоставление целей с ресурсными ограничениями приводит к **формулировке задачи исследования**, которая помимо непротиворечивой системы конкретных целей, учитывающих ресурсные возможности, включает в себя объект моделирования. Задача и объект моделирования должны рассматриваться совместно.

4) Данные о целях исследования, а также исходная информация об объекте моделирования служат для определения **критерия качества создаваемой модели** – количественной меры степени её совершенства. В случае вполне формализованной оптимизационной постановки (например, на основе аппарата линейного программирования) критерий приобретает вид некоторого функционала от переменных и параметров модели, значение которого достигает экстремума при оптимальных ее характеристиках.

5) Следующим шагом в построении модели является основанный на априорных данных **содержательный анализ системы задача-объект и выбор способа формирования модели**.

Если объект не слишком сложен, достаточно изучен и комплекс подлежащих модельному исследованию свойств и характеристик объекта может быть выявлен на основе теоретических представлений и данных (дополняемых необходимым объемом эмпирической информации), целесообразно избрать **аналитический путь построения модели**.

Часто из-за сложности, слабой изученности объекта или отсутствия соответствующих теоретических разработок этот путь не может быть реализован. **Альтернативным является путь идентификации объекта**, т.е. экспериментального определения существенных для решаемой задачи свойств и характеристик объекта, специально ради построения его модели. Эксперимент осуществляется в соответствии со специально разрабатываемым оптимальным планом, данные эксперимента обрабатываются и становятся основой для формализованного описания объекта в виде математической модели вход-выход.

6) Формализованная модель, построенная теоретическим путем или идентифицированная, оценивается в соответствии с выбранным ранее критерием и либо признается удовлетворительной (принимается), либо отвергается как недостаточно совершенная. В последнем случае возникает необходимость в её корректировке и итеративном обращении к ранее выполненным этапам.

7) Решение о принятии модели (в общем случае после i-того итеративного цикла) влечет за собой переход к следующему этапу – **опытной проверке непосредственно в условиях той задачи, для решения которой она построена**. При этом возникают нередко дополнительные требования (например, связанные с удобством использования модели) и необходимость её дополнительной корректировки.

8) Наконец, следует заключительный этап процесса – **использование модели для решения исследовательской или иной задачи**, причем и на этом этапе возможны дальнейшие уточнения и корректировки.

Таким образом, процесс моделирования состоит из трех задач:

1) построение модели (эта задача менее формализуема и конструктивна, в том смысле, что нет алгоритма для построения моделей);

2) исследование модели (эта задача более формализуема, имеются методы исследования различных классов моделей);

3) использование модели (конструктивная и конкретизируемая задача).

1. 8 Лекция №8 (2 часа).

Тема: «Методы оптимизации параметров процессов управления в технических системах»

1.8.1 Вопросы лекции:

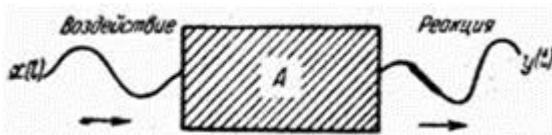
1. Динамическая система. Модель черного ящика. Задачи анализа и синтеза динамических систем.
2. Оператор динамической системы. Классификация операторов.
3. Методы оптимизации параметров процессов управления в динамических системах.

1.8.2 Краткое содержание вопросов:

1. Динамическая система. Модель черного ящика. Задачи анализа и синтеза динамических систем

Имеется некоторая динамическая система A ; под «динамической системой» мы понимаем любой прибор, прицел, счетно-решающий механизм, систему автоматического управления и т. п. Эта система может быть механической, электрической или содержать любые другие элементы. Работу системы будем представлять себе следующим образом: на вход системы непрерывно поступают какие-то входные данные; система перерабатывает их и непрерывно выдает некоторый результат. Условимся называть поступающие на вход системы данные: «воздействием», а выдаваемый результат «реакцией» системы на это воздействие.

Рассмотрим самый простой случай: когда на вход системы A подается только одно воздействие, представляющее собой функцию времени $x(t)$: реакция системы на это воздействие есть другая функция времени $y(t)$. Схема работы системы A условно изображена на рис.

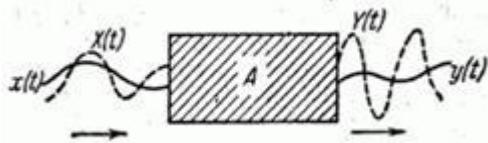


Будем говорить, что система A осуществляет над входным воздействием некоторое преобразование, в результате которого функция $x(t)$ преобразуется в другую функцию $y(t)$.

Преобразование A может быть любого вида и любой сложности. В наиболее простых случаях это, например, умножение на заданный множитель (усилители, множительные механизмы), дифференцирование или интегрирование (дифференцирующие или интегрирующие устройства). Однако на практике системы, осуществляющие в чистом виде такие простейшие преобразования, почти не встречаются; как правило, работа системы описывается дифференциальными уравнениями, и

преобразование A сводится к решению дифференциального уравнения, связывающего воздействие $x(t)$ с реакцией $y(t)$.

При исследовании динамической системы в первую очередь решается основная задача: по заданному воздействию $x(t)$ определить реакцию системы $y(t)$. Однако для полного исследования системы и оценки ее технических качеств такой элементарный подход является недостаточным. В действительности воздействие $x(t)$ никогда не поступает на вход системы в чистом виде: оно всегда искажено некоторыми случайными ошибками (возмущениями), в результате которых на систему фактически воздействует не заданная функция $x(t)$, а случайная функция $X(t)$; соответственно этому система вырабатывает в качестве реакции случайную функцию $Y(t)$, также отличающуюся от теоретической реакции $y(t)$.



Естественно возникает вопрос: насколько велики будут случайные искажения реакции системы при наличии случайных возмущений на ее входе? И далее: как следует выбрать параметры системы для того, чтобы эти искажения были минимальными?

Из двух поставленных выше задач, естественно, более простой является первая - прямая - задача. Сформулируем ее следующим образом.

На вход динамической системы A поступает случайная функция $X(t)$; система подвергает ее известному преобразованию, в результате чего на выходе системы появляется, случайная функция:

$$Y(t) = A(X(t))$$

Известны характеристики случайной функции $X(t)$: математическое ожидание и корреляционная функция. Требуется найти аналогичные характеристики случайной функции $Y(t)$. Короче: по заданным характеристикам случайной функции на входе динамической системы найти характеристики случайной функции на выходе.

Поставленная задача может быть решена совершенно точно в одном частном, но весьма важном для практики случае: когда преобразование A принадлежит к классу так называемых линейных преобразований и соответственно система A принадлежит к классу линейных систем.

2. Оператор динамической системы. Классификация операторов

Понятие оператора является обобщением понятия функции. Когда мы устанавливаем функциональную связь между двумя переменными y и x и пишем:

$$y = f(x)$$

то под символом f мы понимаем правило, по которому заданному значению x приводится в соответствие вполне определенное значение y . Знак f есть символ некоторого преобразования, которому нужно подвергнуть величину x , чтобы получить y . Соответственно виду этого преобразования функции могут быть линейными и нелинейными, алгебраическими, трансцендентными и т. д.

Аналогичные понятия и соответствующая символика применяются в математике и в тех случаях, когда преобразованию подвергаются не величины, а функции. Рассмотрим некоторую функцию $x(t)$ и установим определенное правило A , согласно которому

функция $x(t)$ преобразуется в другую функцию $y(t)$. Запишем это преобразование в следующем виде:

$$y(t) = A\{x(t)\}.$$

Правило A , согласно которому функция $x(t)$ преобразуется в функцию $y(t)$, мы будем называть оператором; например, мы будем говорить: оператор дифференцирования, оператор интегрирования, оператор решения дифференциального уравнения и т. д.

Если динамическая система преобразует поступающую на ее вход функцию $x(t)$ в функцию $y(t)$:

$$y(t) = A\{x(t)\},$$

то оператор A называется оператором динамической системы.

В более общем случае на вход системы поступает не одна, а несколько функций; равным образом на выходе системы могут появляться несколько функций; в этом случае оператор системы преобразует одну совокупность функций в другую. Однако в целях простоты изложения мы рассмотрим здесь лишь наиболее элементарный случай преобразования одной функции в другую.

Преобразования или операторы, применяемые к функциям, могут быть различных типов. Наиболее важным для практики является класс так называемых линейных операторов. Оператор L называется линейным однородным, если он обладает следующими свойствами:

1) к сумме функций оператор может применяться почленно:

$$L(x_1(t) + x_2(t)) = L(x_1(t)) + L(x_2(t));$$

2) постоянную величину c можно выносить за знак оператора:

$$L(cx(t)) = cL(x(t)).$$

Из второго свойства между прочим, следует, что для линейного однородного оператора справедливо свойство

$$L(0) = 0,$$

т. е. при нулевом входном воздействии реакция системы равна нулю.

Примеры линейных однородных операторов:

1) оператор дифференцирования:

$$y(t) = \frac{dx(t)}{dt};$$

2) оператор интегрирования:

$$y(t) = \int_0^t x(\tau) d\tau;$$

3) оператор умножения на определенную функцию $\varphi(t)$:

$$y(t) = \varphi(t)x(t),$$

и т. д.

Кроме линейных однородных операторов, существуют еще линейные неоднородные операторы. Оператор L называется линейным неоднородным, если он состоит из линейного однородного оператора с прибавлением некоторой вполне определенной функции $\varphi(t)$:

$$L(x(t)) = L_0(x(t)) + \varphi(t),$$

где L_0 - линейный однородный оператор.

Примеры линейных неоднородных операторов:

$$1) \quad y(t) = \frac{dx(t)}{dt} + \varphi(t),$$

$$2) \quad y(t) = \int_0^t x(\tau) \varphi(\tau) d\tau + \varphi_1(t),$$

$$3) \quad y(t) = \varphi_1(t)x(t) + \varphi_2(t).$$

где $\varphi(t)$, $\varphi_1(t)$, $\varphi_2(t)$ - вполне определённые функции, а $x(t)$ - преобразуемая оператором функция.

Оператор дифференцирования часто обозначают буквой P : $P = \frac{d}{dt}$, помещаемой в виде множителя перед выражением, подлежащим дифференцированию. При этом запись

$$y(t) = px(t) \quad \text{равносильна записи} \quad y(t) = \frac{dx(t)}{dt}.$$

Двойное дифференцирование обозначается множителем P^2 :

$$P^2 x(t) = \frac{d^2 x(t)}{dt^2} \quad \text{и т. д.}$$

Встречающиеся в технике динамические системы часто описываются линейными дифференциальными уравнениями. В этом случае, как нетрудно убедиться, оператор системы является линейным. Динамическая система, оператор которой является линейным, называется линейной динамической системой.

В противоположность линейным операторам и системам рассматриваются системы и операторы нелинейные. Примерами нелинейных операторов могут служить

$$y(t) = x^2(t), \quad y(t) = \int_0^t x^3(\tau) d\tau, \quad y(t) = \sin x(t),$$

а также решение нелинейного дифференциального уравнения, хотя бы

$$y'(t) + \alpha \cos y(t) = x(t).$$

Динамическая система, оператор которой не является линейным, называется нелинейной системой.

На практике линейные системы встречаются очень часто. В связи с линейностью этих систем к анализу их ошибок может быть с большой эффективностью применен аппарат теории случайных функций. Еще чаще, чем линейные системы, на практике встречаются системы не строго линейные, но в известных пределах допускающие линеаризацию.

3. Методы оптимизации параметров процессов управления в динамических системах

Пусть на вход линейной системы с оператором L воздействует случайная функция $X(t)$, причем известны ее характеристики: математическое ожидание $m_x(t)$ и корреляционная функция $K_x(t, t')$. Реакция системы представляет собой случайную функцию $Y(t) = L(X(t))$.

Требуется найти характеристики случайной функции $Y(t)$ на выходе системы: $m_y(t)$ и $K_y(t, t')$. Короче: по характеристикам случайной функции на входе линейной системы найти характеристики случайной функции на выходе.

Покажем сначала, что можно ограничиться решением этой задачи только для однородного оператора L . Действительно, пусть оператор L неоднороден и выражается формулой $L(X(t)) = L_0(X(t)) + \varphi(t)$,

где L_0 - линейный однородный оператор, $\varphi(t)$ - определенная неслучайная функция.

Тогда $m_y(t) = M[L_0(X(t))] + \varphi(t)$,

т. е. функция $\varphi(t)$ просто прибавляется к математическому ожиданию случайной функции на выходе линейной системы. Что же касается корреляционной функции, то, как известно, она не меняется от прибавления к случайной функции неслучайного слагаемого. Поэтому в дальнейшем изложении под «линейными операторами» будем разуметь только линейные однородные операторы. Решим задачу об определении характеристик на выходе линейной системы сначала для некоторых частных видов линейных операторов.

Дана случайная функция $X(t)$ с математическим ожиданием $m_x(t)$ и корреляционной функцией $K_x(t, t')$. Случайная функция $Y(t)$ связана с $X(t)$ линейным однородным оператором интегрирования:

$$Y(t) = \int_0^t X(\tau) d\tau$$

Требуется найти характеристики случайной функции $Y(t)$, $m_y(t)$ и $K_y(t, t')$.

Представим интеграл как предел суммы:

$$Y(t) = \int_0^t X(\tau) d\tau = \lim_{\Delta\tau \rightarrow 0} \sum_i X(\tau_i) \Delta\tau$$

и применим к равенству операцию математического ожидания. По теореме сложения математических ожиданий имеем:

$$\begin{aligned} m_y(t) &= M[Y(t)] = \lim_{\Delta\tau \rightarrow 0} \sum_i M[X(\tau_i)] \Delta\tau = \\ &= \lim_{\Delta\tau \rightarrow 0} \sum_i m_x(\tau_i) \Delta\tau = \int_0^t m_x(\tau) d\tau \end{aligned}$$

Итак,

$$m_y(t) = \int_0^t m_x(\tau) d\tau$$

т. е. математическое ожидание интеграла от случайной функции равно интегралу от ее математического ожидания. Иными словами: операцию интегрирования и операцию математического ожидания можно менять местами. Это и естественно, так как операция интегрирования по своей природе не отличается от операции суммирования, которую, как мы раньше убедились, можно менять местами с операцией математического ожидания.

Найдём корреляционную функцию $K_y(t, t')$. Для этого перейдём к центрированным случайным функциям:

$$\overset{\circ}{X}(t) = X(t) - m_x(t), \quad \overset{\circ}{Y}(t) = Y(t) - m_y(t)$$

Нетрудно убедиться, что

$$\overset{\circ}{Y}(t) = \int_0^t \overset{\circ}{X}(\tau) d\tau$$

По определению корреляционной функции,

$$K_y(t, t') = M[Y(t)Y(t')]$$

$$Y(t) = \int_0^t X(\tau) d\tau \quad Y(t') = \int_0^{t'} X(\tau') d\tau'$$

Где

Нетрудно убедиться, что произведение двух интегралов в правой части формулы равно двойному интегралу.

Следовательно,

$$Y(t)Y(t') = \int_0^t \int_0^{t'} X(\tau)X(\tau') d\tau d\tau'$$

Применяя к равенству операцию математического ожидания и меняя ее в правой части местами с операцией интегрирования, получим:

$$K_y(t, t') = M[Y(t)Y(t')] = \int_0^t \int_0^{t'} M[X(\tau)X(\tau')] d\tau d\tau'$$

или окончательно:

$$K_y(t, t') = \int_0^t \int_0^{t'} K_x(\tau, \tau') d\tau d\tau'$$

Таким образом, для того чтобы найти корреляционную функцию интеграла от случайной функции, нужно дважды проинтегрировать корреляционную функцию исходной случайной функции: сначала по одному аргументу, затем - по другому.

Дана случайная функция $X(t)$ с математическим ожиданием $m_x(t)$ и корреляционной функцией $K_x(t, t')$. Случайная функция $Y(t)$ связана со случайной функцией $X(t)$ линейным однородным оператором дифференцирования:

$$Y(t) = \frac{dX(t)}{dt}$$

Требуется найти $m_y(t)$ и $K_y(t, t')$.

$$Y(t) = \lim_{\Delta t \rightarrow 0} \frac{X(t + \Delta t) - X(t)}{\Delta t}$$

Представим производную в виде предела:

Применяя к равенству операцию математического ожидания, получим:

$$m_y(t) = M[Y(t)] = \lim_{\Delta t \rightarrow 0} \frac{m_x(t + \Delta t) - m_x(t)}{\Delta t} = \frac{dm_x(t)}{dt}$$

Итак,

$$m_y(t) = \frac{dm_x(t)}{dt}$$

т. е. математическое ожидание производной от случайной функции равно производной от ее математического ожидания.

Следовательно, операцию дифференцирования, как и операцию интегрирования тоже можно менять местами с операцией математического ожидания.

Для определения $K_y(t, t')$ перейдём к центрированным случайным функциям $\overset{\circ}{Y}(t)$ и $\overset{\circ}{X}(t)$; очевидно: $\overset{\circ}{Y}(t) = \frac{d \overset{\circ}{X}(t)}{dt}$.
По определению

$$K_y(t, t') = M[\overset{\circ}{Y}(t)\overset{\circ}{Y}(t')]$$

Подставим вместо $\overset{\circ}{Y}(t)$ и $\overset{\circ}{Y}(t')$ их выражения:

$$K_y(t, t') = M\left[\frac{d \overset{\circ}{X}(t)}{dt} \frac{d \overset{\circ}{X}(t')}{dt'} \right]$$

Мы доказали, что математическое ожидание производной случайной функции равно производной от математического ожидания, т. е. знаки дифференцирования и математического ожидания можно менять местами. Следовательно,

$$K_y(t, t') = M\left[\frac{\partial^2 \overset{\circ}{X}(t) \overset{\circ}{X}(t')}{\partial t \partial t'} \right] = \frac{\partial^2}{\partial t \partial t'} M[\overset{\circ}{X}(t) \overset{\circ}{X}(t')] = \frac{\partial^2}{\partial t \partial t'} K_x(t, t')$$

Таким образом,

$$K_y(t, t') = \frac{\partial^2 K_x(t, t')}{\partial t \partial t'}$$

Итак, чтобы найти корреляционную функцию производной, нужно дважды продифференцировать корреляционную функцию исходной случайной функции: сначала по одному аргументу, затем - по другому.

Можно доказать, что такое правило является общим для всех линейных однородных операторов: если случайная функция $X(t)$ с математическим ожиданием $m_x(t)$ и корреляционной функцией $K_x(t, t')$ преобразуется линейным однородным оператором L в случайную функцию,
 $Y(t) = L(X(t))$,

то для нахождения математического ожидания случайной функции $Y(t)$ нужно применить тот же оператор к математическому ожиданию случайной функции $X(t)$:
 $m_y(t) = L(m_x(t))$,

а для нахождения корреляционной функции нужно дважды применить тот же оператор к корреляционной функции случайной функции $X(t)$, сначала по одному аргументу, затем - по другому:

$$K_y(t, t') = L^{(t)} L^{(t')} (K_x(t, t'))$$

Во многих задачах практики нас, в конечном счёте, интересует не корреляционная функция $K_y(t, t')$ на выходе линейной системы, а дисперсия $D_y(t)$ характеризующая точность работы системы в условиях наличия случайных возмущений.

Дисперсию $D_y(t)$ можно найти, зная корреляционную функцию:
 $D_y(t) = K_y(t, t)$

При этом нужно подчеркнуть, что, как правило, для определения дисперсии на выходе линейной системы недостаточно знать дисперсию на ее входе, а существенно важно знать корреляционную функцию.

Действительно, линейная система может совершенно по-разному реагировать на случайные возмущения, поступающие на ее вход, в зависимости от того, какова внутренняя структура этих случайных возмущений; состоят ли они, например, по преимуществу из высокочастотных или низкочастотных колебаний.

Внутренняя же структура случайного процесса описывается не его дисперсией, а корреляционной функцией.

2. МЕТОДИЧЕСКИЕ УКАЗАНИЯ ПО ПРОВЕДЕНИЮ ПРАКТИЧЕСКИХ ЗАНЯТИЙ

2.1 Практическое занятие № 1 (2 часа).

Тема: «Проблемы современной фундаментальной науки. Основные этапы экспериментального исследования. Классификации методов исследования»

2.1.1 Задание для работы:

1. Исследования в области вещества и энергии
2. Исследования в области новых материалов и технологий
3. Основные этапы экспериментального исследования
4. Классификации методов исследования

2.1.2 Краткое описание проводимого занятия:

- 1. Исследования в области вещества и энергии** – доклады с презентациями по теме, обсуждение
- 2. Исследования в области новых материалов и технологий** – доклады с презентациями по теме, обсуждение
- 3. Основные этапы экспериментального исследования** – доклады с презентациями по теме, обсуждение
- 4. Классификации методов исследования** – доклады с презентациями по теме, обсуждение

2.1.3 Результаты и выводы:

В результате проведенного занятия студенты:

- должны ознакомиться с современными направлениями в исследованиях вещества и энергии, новых материалов и технологий;
- усвоить основные принципы построения глобального информационного пространства;
- выработать навыки по изучению интегративных качеств фундаментальных моделей;
- должны ознакомиться с классификацией НИР, ее особенностями;
- усвоить основные требования к методологии научно-исследовательских работ;
- выработать навыки анализа научной проблемы и построение поэтапного плана ее решения.

2.2 Практическое занятие № 2 (2 часа).

Тема: «Математическое моделирование в инженерных исследованиях»

2.2.1 Задание для работы:

1. Математическая модель: этапы построения, отличительные особенности.
2. Типовые математические модели в инженерных исследованиях.

2.2.2 Краткое описание проводимого занятия:

- 1. Математическая модель: этапы построения, отличительные особенности** -доклады с презентациями, обсуждение
- 2. Типовые математические модели в инженерных исследованиях**- доклады с презентациями, обсуждение

2.2.3 Результаты и выводы:

В результате проведенного занятия студенты:

- рассмотреть типовые математические модели, применяемые в инженерных приложениях и условия их использования;
- усвоить основные понятия, связанные с математическим моделированием;
- выработать навыки анализа этапов построения математических моделей.

2.3 Практическое занятие № 3 (2 часа).

Тема: «Факторы, методы отбора, общая характеристика, функция отклика»

2.3.1 Задание для работы:

1. Факторы, классификация, методы отбора факторов.
2. Общая характеристика факторов.
3. Функция отклика

2.3.2 Краткое описание проводимого занятия:

- 1. Факторы, классификация, методы отбора факторов** – подбор варьируемых факторов и составление матрицы планирования трехфакторного эксперимента (по варианту).
- 2. Общая характеристика факторов** - проведение виртуального эксперимента с назначением выходного параметра.
- 3. Функция отклика** - расчет свободного члена, коэффициентов при линейных факторах и эффектах взаимодействия в уравнении функции отклика.

2.3.3 Результаты и выводы:

В результате проведенного занятия студенты:

- должны ознакомиться с основами теории планирования эксперимента;
- усвоить основные методики на примере планирования трехфакторного эксперимента;
- выработать навыки составления и анализа функции отклика

2.4 Практическое занятие № 4 (2 часа).

Тема: «Планы факторного эксперимента»

2.4.1 Задание для работы:

1. Полный факторный эксперимент типа 2^3 , матрица планирования которого выдается по вариантам.

2.4.2 Краткое описание проводимого занятия:

1. Полный факторный эксперимент типа 2^3 , матрица планирования которого выдается по вариантам – расчет и проверка значимости коэффициентов регрессии, расчет теоретического значения параметра оптимизации, проверка адекватности модели.

2.4.3 Результаты и выводы:

В результате проведенного занятия студенты:

- должны ознакомиться с основами теории планирования эксперимента;
- усвоить основные методики работы с конкретным планом;
- выработать навыки анализа параметра оптимизации и проверки модели на адекватность.

2.5 Практическое занятие № 5 (2 часа).

Тема: «Теоретические основы обработки экспериментальных данных»

2.5.1 Задание для работы:

1. Многомерные СВ, законы их распределения, условные числовые характеристики
2. Функция регрессии, коэффициент детерминации, корреляции, ковариация

2.5.2 Краткое описание проводимого занятия:

1. Многомерные СВ, законы их распределения, условные числовые характеристики

Пример 1. Найти выборочные средние, дисперсии и коэффициент корреляции для выборки, приведенной в таблицей. Построить диаграмму рассеивания.

Решение. Вычисление указанных выборочных характеристик удобно выполнять в следующей последовательности. Сначала вычисляют суммы $\sum x_i$, $\sum y_i$, $\sum x_i^2$, $\sum y_i^2$, $\sum x_i y_i$, $\sum (x_i + y_i)^2$. Для контроля правильности вычислений используется тождество

$$\sum (x_i + y_i)^2 = \sum x_i^2 + 2\sum x_i y_i + \sum y_i^2.$$

Таблица 1

x	y								
8,35	3,50	10,50	6,00	11,35	9,50	12,15	6,00	12,85	9,50
8,74	1,49	10,75	2,50	11,50	6,00	12,25	8,05	13,15	9,02
9,25	6,40	10,76	5,74	11,50	9,00	12,35	5,01	13,25	6,49
9,50	4,50	11,00	8,50	11,62	8,50	12,50	7,03	13,26	10,50
9,75	5,00	11,00	5,26	11,75	10,00	12,76	7,53	13,40	7,51
10,24	7,00	11,25	8,00	12,00	9,00	12,85	6,01	13,50	10,00
13,65	9,50	14,50	10,00	13,75	8,51	14,75	12,00	14,00	11,00
15,25	12,50	14,23	8,40	16,00	11,50	14,26	10,00	16,00	13,00
14,51	9,50	16,25	12,00						

Объем выборки $n = 42$.

Выборочные средние отсюда находятся по формулам

$$\bar{x} = \alpha_{1,0}^* = \frac{1}{n} \sum x_i, \quad \bar{y} = \alpha_{0,1}^* = \frac{1}{n} \sum y_i$$

Затем вычисляются суммы квадратов отклонений от среднего и произведений отклонений

$$Q_x = \sum (x_i - \bar{x})^2 = \sum x_i^2 - \frac{(\sum x_i)^2}{n}, \quad Q_y = \sum (y_i - \bar{y})^2 = \sum y_i^2 - \frac{(\sum y_i)^2}{n},$$

$$Q_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n}$$

$$\text{Отсюда } D_x^* = \frac{1}{n} Q_x, \quad D_y^* = \frac{1}{n} Q_y, \quad r = \frac{\mu_{1,1}^*}{\sqrt{D_x^* D_y^*}} = \frac{Q_{xy}}{\sqrt{Q_x Q_y}}.$$

Предварительно вычислим

$$\sum x_i = 522,23, \quad \sum y_i = 336,41, \quad \sum x_i^2 = 6652,25, \quad \sum y_i^2 = 2987,80, \quad \sum x_i y_i = 4358,626.$$

Тогда найдем $\bar{x} = 12,434$, $\bar{y} = 8,011$.

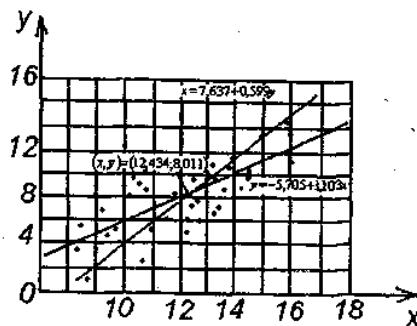
$$\text{Далее находим } Q_x = 6652,25 - \frac{522,23^2}{42} \approx 158,8182, \quad Q_y = 2987,805 - \frac{336,41^2}{42} \approx 292,5958,$$

$$Q_{xy} = 4358,626 - \frac{522,23 \cdot 336,41}{42} \approx 175,1912.$$

Окончательно, получаем

$$D_x^* = \frac{158,8182}{42} \approx 3,7814, \quad D_y^* = \frac{292,5958}{42} \approx 6,9666, \quad r = \frac{175,1912}{\sqrt{158,8182 \cdot 292,5958}} \approx 0,813.$$

Диаграмма рассеивания приведена на рис. 1.



Выборочная линейная регрессия Y на X по выборке (x_i, y_i) , $i=1,2,\dots,n$, определяется

уравнением $y = \beta_0^* + \beta_1^*x = \bar{y} + r \frac{D_Y^*}{D_X^*}(x - \bar{x})$.

Коэффициенты β_0^* и β_1^* называются *выборочными коэффициентами регрессии*. Они

вычисляются по формулам: $\beta_1 = \frac{n \sum x_i y_i - (\sum x_i)(\sum y_i)}{n \sum x_i^2 - (\sum x_i)^2} = \frac{Q_{xy}}{Q_x}$, $\beta_0 = \bar{y} - \beta_1 \bar{x}$.

Аналогично определяется выборочная линейная регрессия X на Y :

$x = \beta_0^{**} + \beta_1^{**}y = \bar{x} + r \frac{D_X^*}{D_Y^*}(y - \bar{y})$ коэффициенты β_0^{**} и β_1^{**} которой находятся по формулам

$$\beta_1^{**} = \frac{n \sum x_i y_i - (\sum x_i)(\sum y_i)}{n \sum y_i^2 - (\sum y_i)^2} = \frac{Q_{xy}}{Q_y}, \quad \beta_0^{**} = \bar{x} - \beta_1^{**} \bar{y}.$$

Для контроля правильности расчетов используют соотношение $\sqrt{\beta_1^* \beta_1^{**}} = |r|$.

Прямые $y = \beta_0^* + \beta_1^*x$, $x = \beta_0^{**} + \beta_1^{**}y$ пересекаются в точке с координатами (\bar{x}, \bar{y}) .

Пример 2. Вычислить выборочные коэффициенты линейной регрессии X на Y и Y на X по выборке из предыдущего примера. Нанести прямые регрессии на диаграмму рассеивания.

Решение. Воспользуемся результатами вычислений в предыдущем примере. По формулам находим

$$\beta_1^* = \frac{175,1912}{158,8182} \approx 1,103, \quad \beta_0^* = 8,011 - 1,103 \cdot 12,434 \approx -5,705.$$

Таким образом, прямая регрессии Y на X имеет уравнение $y = -5,705 + 1,103x$.

$$\text{Аналогично находим } \beta_1^{**} = \frac{175,1912}{292,5958} \approx 0,599, \quad \beta_0^{**} = 12,434 - 0,599 \cdot 8,011 \approx 7,637.$$

Отсюда прямая регрессии X на Y имеет уравнение $x = 7,637 + 0,599y$.

Проверка показывает $\sqrt{1,103 \cdot 0,599} \approx 0,813$, что полученный результат совпадает со значением r , вычисленным в примере (*). Прямые регрессии нанесены на диаграмму рассеивания на рис.1.

Пример 3. Используя группировку выборки, заданной таблицей в примере (*), вычислить выборочные средние, дисперсии, коэффициент корреляции, а также выборочные коэффициенты линейной регрессии X на Y и Y на X .

Решение. Выберем $b_x=1$, $b_y=2$. Прямоугольная сетка, соответствующая этим значениям, нанесена на диаграмму рассеивания (рис. 1). Непосредственно по диаграмме строим корреляционную таблицу (таблица 2). Находим $d_X^* = 11,5$, $d_Y^* = 9$ и вычисляем значения u_i

и v_j по формулам $u_i = \frac{\hat{x}_i - 11,5}{1}$, $i=1,2,\dots,9$, $v_j = \frac{\hat{y}_j - 9}{2}$, $j=1,2,\dots,7$.

Вычисляем следующие суммы:

$\sum n_i u_i = 43$, $\sum n_j v_j = -15$, $\sum n_i u_i^2 = 215$, $\sum n_j v_j^2 = 87$, $\sum_{i=1}^9 \sum_{j=1}^7 n_{ij} u_i v_j = 80$. По формулам

находим $Q_u = 215 - \frac{43^2}{42} \approx 170,976$, $Q_v = 87 - \frac{(-15)^2}{42} \approx 81,643$, $Q_{uv} = 80 - \frac{43 \cdot (-15)}{42} \approx 95,357$.

Далее получаем $\bar{x} = 1 \frac{43}{42} + 11,5 \approx 12,52$, $\bar{y} = 2 \frac{(-15)}{42} + 9 \approx 8,28$, $D_x^* = 1^2 \frac{170,976}{42} \approx 4,071$,

$D_Y^* = 2^2 \frac{81,643}{42} \approx 7,775$, $r = \frac{95,357}{\sqrt{170,976 \cdot 81,643}} \approx 0,807$. Находим выборочные коэффициенты

регрессии: $\beta_1^* = \frac{2}{1} \frac{95,357}{170,976} \approx 1,12$, $\beta_1^{**} = \frac{1}{2} \frac{95,357}{81,643} \approx 0,58$, $\beta_0^* = 8,28 - 1,12 \cdot 12,52 \approx -5,74$,

$$\beta_0^* = 12,52 - 0,58 \cdot 8,28 \approx 7,72.$$

Окончательно получим, что уравнение линейной регрессии Y на X имеет вид $y = -5,74 + 1,12x$, а уравнение линейной регрессии X на Y имеет вид $x = 7,72 + 0,58y$.

Расхождение полученных результатов с результатами выше рассмотренных примеров обусловлено группировкой.

Таблица 2. Корреляционная таблица для диаграммы рассеивания

Границы и средины интервалов для y	v_j	Границы и середины интервалов для x									n_j	$n_j v_j$	$n_j v_j^2$	
		8-9	9-	10-	11-	12-	13-	14-	15-	16-				
		8,5	10	11	12	13	14	15	16	17				
		u_i												
		-3	-2	-1	0	1	2	3	4	5				
0-2 1	-4	1	0	0	0	0	0	0	0	0	1	-4	16	
2-4 3	-3	1	0	1	0	0	0	0	0	0	2	-6	18	
4-6 5	-2	0	2	1	1	1	0	0	0	0	5	-10	20	
6-8 7	-1	0	1	2	1	4	2	0	0	0	10	-10	10	
8-10 9	0	0	0	0	5	3	3	2	0	0	13	0	0	
10-12 11	1	0	0	0	1	0	2	3	0	1	7	7	7	
12-14 13	2	0	0	0	0	0	0	1	1	2	4	8	16	
n_i		2	3	4	8	8	7	6	1	3	$\sum=42$	$\sum=-15$	$\sum=87$	
$n_i v_i$		-6	-6	-4	0	8	14	18	4	15	$\sum=43$			
$n_i v_i^2$		18	12	4	0	8	28	54	16	75	$\sum=215$			

2. Функция регрессии, коэффициент детерминации, корреляции, ковариация

Допустим, что в результате лечения 12 больных с артериальной гипертензией в результате суточного мониторирования sistолического артериального давления (САД) до лечения и после месячного лечения были получены следующие результаты:

№	САД до (x_i)	САД после (y_i)	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$
			-9,6	-7,5	
			-19,6	-17,5	
			-14,6	-12,5	182,5
			-4,6	7,5	-34,5
			0,4	12,5	
			5,4	12,5	67,5
			-9,6	-7,5	
			10,4	7,5	
			15,4	12,5	192,5
			0,4	-7,5	-3

			5,4	-2,5	-13,5
			20,4	2,5	
	$\bar{x} = 169,6$	$\bar{y} = 137,5$			$\sum = 1012,5$
	$\sigma_x = 12,1$	$\sigma_y = 10,6$			

$$r_{x,y} = \frac{\frac{1}{12-1} \cdot 1012,5}{12,1 \cdot 10,6} = 0,718$$

Итак, коэффициент корреляции получился равным 0,718.

Определим, достоверно ли он отличается от нуля. Для этого используем Таблицу 10 приложения. У нас 12 пар измерений, поэтому входим в Таблицу по 12 строке. На пересечении 12 строки и столбца Р=0,05 стоит число 0,576. Полученный коэффициент корреляции (0,718) больше этого числа.

Следовательно, на этом уровне коэффициент корреляции достоверно отличается от нуля, то есть связь есть. На пересечении этой же строки и столбца Р=0,01 стоит число 0,708. Поскольку коэффициент корреляции больше и этого числа, следовательно, мы можем говорить, что связь существует и на этом более значимом уровне. Итак, ответ на первый вопрос таков: существование связи высоко достоверно. Далее, поскольку получено положительное значение коэффициента корреляции, мы заключаем, что связь прямая. Используя Таблицу 2 данного раздела, мы приходим к заключению, что связь сильная.

Найдем коэффициент детерминации:

$$d = r^2 \times 100 = 0,718^2 \times 100 = 0,516 \times 100 = 51,6 \text{ (%)}$$

Таким образом, системическое артериальное давление после лечения на 51,6 % определяется системическим артериальным давлением до лечения, а на 48,4 % другими факторами.

Формы проявления взаимосвязей явлений и процессов весьма разнообразны. Из них в самом общем виде выделяют *функциональную* (полную) и *корреляционную* (неполную) связи.

Математически ковариация представляет собой меру линейной зависимости двух случайных величин.

Коэффициент корреляции - это математическая мера корреляции двух величин. Коэффициенты корреляции могут быть положительными и отрицательными. Иногда показателям тесноты связи можно дать качественную оценку (шкала Чеддока):

Количественная мера тесноты связи	Качественная характеристика силы связи
0,1 - 0,3	Слабая
0,3 - 0,5	Умеренная
0,5 - 0,7	Заметная
0,7 - 0,9	Высокая
0,9 - 0,99	Весьма высокая

2.5.3 Результаты и выводы:

В результате проведенного занятия студенты:

- ознакомиться с основными понятиями многомерного статистического анализа, теории корреляции, классификацией регрессий;

- усвоить алгоритмы нахождения условных законов и числовых характеристик многомерных случайных величин, вычисления коэффициента корреляции, детерминации, ковариации;
- выработать навыки нахождения уравнения регрессии, проверки его параметров на статистическую значимость.

2.6 Практическое занятие № 6 (2 часа).

Тема: «Методы стохастического описания и анализа особенностей процессов управления в технических системах»

2.6.1 Задание для работы:

1. Виды регрессий, статистическая значимость их параметров.
- 2 Автокорреляция

2.6.2 Краткое описание проводимого занятия:

1. Виды регрессий, статистическая значимость их параметров.

Задача:

Имеется связанный выборка из 26 пар значений (x_k, y_k):

k	1	2	3	4	5	6	7	8	9
x_k	25.20000	26.40000	26.00000	25.80000	24.90000	25.70000	25.70000	25.70000	26.10000
y_k	30.80000	29.40000	30.20000	30.50000	31.40000	30.30000	30.40000	30.50000	29.90000

k	11	12	13	14	15	16	17	18	19
x_k	25.90000	26.20000	25.60000	25.40000	26.60000	26.20000	26.00000	22.10000	25.90000
y_k	30.30000	30.50000	30.60000	31.00000	29.60000	30.40000	30.70000	31.60000	30.50000

k	21	22	23	24	25	26
x_k	25.90000	26.30000	26.10000	26.00000	26.40000	25.80000
y_k	30.70000	30.10000	30.60000	30.50000	30.70000	30.80000

Требуется вычислить/построить:

- коэффициент корреляции;
- проверить гипотезу зависимости случайных величин X и Y, при уровне значимости $\alpha = 0.05$;
- коэффициенты уравнения линейной регрессии;
- диаграмму рассеяния (корреляционное поле) и график линии регрессии;

РЕШЕНИЕ:

1. Вычисляем коэффициент корреляции.

Коэффициент корреляции — это показатель взаимного вероятностного влияния двух случайных величин. Коэффициент корреляции R может принимать значения от **-1** до **+1**. Если абсолютное значение находится ближе к **1**, то это свидетельство сильной связи между величинами, а если ближе к **0** — то, это говорит о слабой связи или ее отсутствии. Если абсолютное значение R равно единице, то можно говорить о функциональной связи

между величинами, то есть одну величину можно выразить через другую посредством математической функции.

Вычислить коэффициент корреляции можно по следующим формулам:

$$R_{x,y} = \frac{\text{cov}(X,Y)}{\sigma_x \sigma_y} \quad (1.1), \quad \text{где:}$$

$\text{cov}(X,Y)$ - ковариация случайных величин X и Y

$$\sigma_x^2 = \frac{1}{n} \sum_{k=1}^n (x_k - M_x)^2, \quad \sigma_y^2 = \frac{1}{n} \sum_{k=1}^n (y_k - M_y)^2 \quad (1.2), \quad \text{- оценки дисперсий случайных величин X и Y соответственно.}$$

$$M_x = \frac{1}{n} \sum_{k=1}^n x_k, \quad M_y = \frac{1}{n} \sum_{k=1}^n y_k \quad (1.3), \quad \text{- оценки математического ожидания случайных величин X и Y соответственно.}$$

или по формуле

$$R_{x,y} = \frac{M_{xy} - M_x M_y}{S_x S_y} \quad (1.4), \quad \text{где:}$$

$$M_x = \frac{1}{n} \sum_{k=1}^n x_k, \quad M_y = \frac{1}{n} \sum_{k=1}^n y_k, \quad M_{xy} = \frac{1}{n} \sum_{k=1}^n x_k y_k \quad (1.5)$$

$$S_x^2 = \frac{1}{n} \sum_{k=1}^n x_k^2 - M_x^2, \quad S_y^2 = \frac{1}{n} \sum_{k=1}^n y_k^2 - M_y^2 \quad (1.6)$$

На практике, для вычисления коэффициента корреляции чаще используется формула (1.4) т.к. она требует меньше вычислений. Однако если предварительно была вычислена ковариация $\text{cov}(X,Y)$, то выгоднее использовать формулу (1.1), т.к. кроме собственно значения ковариации можно воспользоваться и результатами промежуточных вычислений.

1.1 Вычислим коэффициент корреляции по формуле (1.4), для этого вычислим значения x_k^2 , y_k^2 и $x_k y_k$ и занесем их в таблицу 1.

Таблица 1

k	x_k	y_k	x_k^2	y_k^2	$x_k y_k$
1	2	3	4	5	6
1	25.2	30.8	635.04000	948.64000	776.16000
2	26.4	29.4	696.96000	864.36000	776.16000
3	26.0	30.2	676.00000	912.04000	785.20000
4	25.8	30.5	665.64000	930.25000	786.90000
5	24.9	31.4	620.01000	985.96000	781.86000
6	25.7	30.3	660.49000	918.09000	778.71000
7	25.7	30.4	660.49000	924.16000	781.28000
8	25.7	30.5	660.49000	930.25000	783.85000
9	26.1	29.9	681.21000	894.01000	780.39000
10	25.8	30.4	665.64000	924.16000	784.32000
11	25.9	30.3	670.81000	918.09000	784.77000

12	26.2	30.5	686.44000	930.25000	799.10000
13	25.6	30.6	655.36000	936.36000	783.36000
14	25.4	31	645.16000	961.00000	787.40000
15	26.6	29.6	707.56000	876.16000	787.36000
16	26.2	30.4	686.44000	924.16000	796.48000
17	26	30.7	676.00000	942.49000	798.20000
18	22.1	31.6	488.41000	998.56000	698.36000
19	25.9	30.5	670.81000	930.25000	789.95000
20	25.8	30.6	665.64000	936.36000	789.48000
21	25.9	30.7	670.81000	942.49000	795.13000
22	26.3	30.1	691.69000	906.01000	791.63000
23	26.1	30.6	681.21000	936.36000	798.66000
24	26	30.5	676.00000	930.25000	793.00000
25	26.4	30.7	696.96000	942.49000	810.48000
26	25.8	30.8	665.64000	948.64000	794.64000

1.2. Вычислим M_x по формуле (1.5).

1.2.1. Сложим последовательно все элементы x_k

$$x_1 + x_2 + \dots + x_{26} = 25.20000 + 26.40000 + \dots + 25.80000 = 669.500000$$

1.2.2. Разделим полученную сумму на число элементов

$$669.50000 / 26 = 25.75000 \quad M_x = \mathbf{25.75000}$$

1.3. Аналогичным образом вычислим M_y .

1.3.1. Сложим последовательно все элементы y_k

$$y_1 + y_2 + \dots + y_{26} = 30.80000 + 29.40000 + \dots + 30.80000 = 793.000000$$

1.3.2. Разделим полученную сумму на число элементов выборки

$$793.00000 / 26 = 30.50000 \quad M_y = \mathbf{30.50000}$$

1.4. Аналогичным образом вычислим M_{xy} .

1.4.1. Сложим последовательно все элементы 6-го столбца таблицы 1

$$776.16000 + 776.16000 + \dots + 794.64000 = 20412.830000$$

1.4.2. Разделим полученную сумму на число элементов

$$20412.83000 / 26 = 785.10885 \quad M_{xy} = \mathbf{785.108846}$$

1.5. Вычислим значение S_x^2 по формуле (1.6.).

1.5.1. Сложим последовательно все элементы 4-го столбца таблицы 1

$$635.04000 + 696.96000 + \dots + 665.64000 = 17256.910000$$

1.5.2. Разделим полученную сумму на число элементов

$$17256.91000 / 26 = 663.72731$$

1.5.3. Вычтем из последнего числа квадрат величины M_x получим значение для S_x^2

$$S_x^2 = 663.72731 - 25.75000^2 = 663.72731 - 663.06250 = \mathbf{0.66481}$$

1.6. Вычислим значение S_y^2 по формуле (1.6.).

1.6.1. Сложим последовательно все элементы 5-го столбца таблицы 1

$$948.64000 + 864.36000 + \dots + 948.64000 = 24191.840000$$

1.6.2. Разделим полученную сумму на число элементов

$$24191.84000 / 26 = 930.45538$$

1.6.3. Вычтем из последнего числа квадрат величины M_y получим значение для S_y^2

$$S_y^2 = 930.45538 - 30.50000^2 = 930.45538 - 930.25000 = \mathbf{0.20538}$$

1.7. Вычислим произведение величин S_x^2 и S_y^2 .

$$S_x^2 S_y^2 = 0.66481 \cdot 0.20538 = 0.136541$$

1.8. Извлечем из последнего числа квадратный корень, получим значение $S_x S_y$.

$$S_x S_y = 0.36951$$

1.9. Вычислим значение коэффициента корреляции по формуле (1.4).

$$R = (785.10885 - 25.75000 \cdot 30.50000) / 0.36951 = (785.10885 - 785.37500) / 0.36951 = -0.72028$$

ОТВЕТ: $R_{x,y} = -0.720279$

2. Проверяем значимость коэффициента корреляции (проверяем гипотезу зависимости). Поскольку оценка коэффициента корреляции вычислена на конечной выборке, и поэтому может отклоняться от своего генерального значения, необходимо проверить значимость коэффициента корреляции. Проверка производится с помощью t-критерия:

$$t = \frac{R_{x,y} \sqrt{n - 2}}{\sqrt{1 - R^2_{x,y}}} \quad (2.1)$$

Случайная величина t следует t-распределению Стьюдента и по таблице t-распределения необходимо найти критическое значение критерия ($t_{kp,\alpha}$) при заданном уровне значимости α . Если вычисленное по формуле (2.1) t по модулю окажется меньше чем $t_{kp,\alpha}$, то зависимости между случайными величинами X и Y нет. В противном случае, экспериментальные данные не противоречат гипотезе о зависимости случайных величин.

2.1. Вычислим значение t-критерия по формуле (2.1) получим:

$$t = \frac{-0.72028 \sqrt{26 - 2}}{\sqrt{1 - (-0.72028)^2}} = -5.08680$$

2.2. Определим по таблице t-распределения критическое значение параметра $t_{kp,\alpha}$. Искомое значение $t_{kp,\alpha}$ располагается на пересечении строки соответствующей числу степеней свободы и столбца соответствующего заданному уровню значимости α .

В нашем случае число степеней свободы есть $n - 2 = 26 - 2 = 24$ и $\alpha = 0.05$, что соответствует критическому значению критерия $t_{kp,\alpha} = 2.064$ (см. табл. 2)

Таблица 2 **t-распределение**

Число степеней свободы ($n - 2$)	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.02$	$\alpha = 0.01$	$\alpha = 0.002$	$\alpha = 0.001$
1	6.314	12.706	31.821	63.657	318.31	636.62
2	2.920	4.303	6.965	9.925	22.327	31.598
3	2.353	3.182	4.541	5.841	10.214	12.924
4	2.132	2.776	3.747	4.604	7.173	8.610
5	2.015	2.571	3.365	4.032	5.893	6.869
6	1.943	2.447	3.143	3.707	5.208	5.959
7	1.895	2.365	2.998	3.499	4.785	5.408
8	1.860	2.306	2.896	3.355	4.501	5.041
9	1.833	2.262	2.821	3.250	4.297	4.781
10	1.812	2.228	2.764	3.169	4.144	4.587
11	1.796	2.201	2.718	3.106	4.025	4.437
12	1.782	2.179	2.681	3.055	3.930	4.318
13	1.771	2.160	2.650	3.012	3.852	4.221
14	1.761	2.145	2.624	2.977	3.787	4.140
15	1.753	2.131	2.602	2.947	3.733	4.073
16	1.746	2.120	2.583	2.921	3.686	4.015

17	1.740	2.110	2.567	2.898	3.646	3.965
18	1.734	2.101	2.552	2.878	3.610	3.922
19	1.729	2.093	2.539	2.861	3.579	3.883
20	1.725	2.086	2.528	2.845	3.552	3.850
21	1.721	2.080	2.518	2.831	3.527	3.819
22	1.717	2.074	2.508	2.819	3.505	3.792
23	1.714	2.069	2.500	2.807	3.485	3.767
24	1.711	2.064	2.492	2.797	3.467	3.745
25	1.708	2.060	2.485	2.787	3.450	3.725
26	1.706	2.056	2.479	2.779	3.435	3.707
27	1.703	2.052	2.473	2.771	3.421	3.690
28	1.701	2.048	2.467	2.763	3.408	3.674
29	1.699	2.045	2.462	2.756	3.396	3.659
30	1.697	2.042	2.457	2.750	3.385	3.646
40	1.684	2.021	2.423	2.704	3.307	3.551
60	1.671	2.000	2.390	2.660	3.232	3.460
120	1.658	1.980	2.358	2.617	3.160	3.373
∞	1.645	1.960	2.326	2.576	3.090	3.291

2.2. Сравним абсолютное значение t-критерия и $t_{\text{кр.}\alpha}$

Абсолютное значение t-критерия не меньше критического $t = 5.08680$, $t_{\text{кр.}\alpha} = 2.064$, следовательно **экспериментальные данные, с вероятностью 0.95 ($1 - \alpha$), не противоречат гипотезе о зависимости случайных величин X и Y.**

3. Вычисляем коэффициенты уравнения линейной регрессии.

Уравнение линейной регрессии представляет собой уравнение прямой, аппроксимирующей (приблизительно описывающей) зависимость между случайными величинами X и Y. Если считать, что величина X свободная, а Y зависимая от X, то уравнение регрессии запишется следующим образом $Y = a + b \cdot X$ (3.1), где:

$$b = R_{x,y} \frac{\sigma_y}{\sigma_x} = R_{x,y} \frac{S_y}{S_x} \quad (3.2),$$

$$a = M_y - b \cdot M_x \quad (3.3)$$

Рассчитанный по формуле (3.2) коэффициент **b** называют коэффициентом линейной регрессии. В некоторых источниках **a** называют постоянным коэффициентом регрессии и **b** соответственно переменным.

Погрешности предсказания Y по заданному значению X вычисляются по формулам :

$$\sigma_{y/x} = \sigma_y \sqrt{1 - R^2_{x,y}} = S_y \sqrt{1 - R^2_{x,y}} \quad (3.4) \quad \text{- абсолютная погрешность,}$$

$$\delta_{y/x} = \frac{\sigma_{y/x}}{M_y} 100\% \quad (3.5) \quad \text{- относительная погрешность}$$

Величину $\sigma_{y/x}$ (формула 3.4) еще называют **остаточным средним квадратическим отклонением**, оно характеризует уход величины Y от линии регрессии, описываемой уравнением (3.1), при фиксированном (заданном) значении X.

3.1. Вычислим отношение $\frac{S_y^2}{S_x^2}$.

$$S_y^2 / S_x^2 = 0.20538 / 0.66481 = 0.30894$$

3.2. Вычислим отношение $\frac{S_y}{S_x}$.

Извлечем из последнего числа квадратный корень - получим:

$$S_y / S_x = 0.55582$$

3.3 Вычислим коэффициент b по формуле (3.2)

$$b = -0.72028 \cdot 0.55582 = -0.40035$$

3.4 Вычислим коэффициент a по формуле (3.3)

$$a = 30.50000 - (-0.40035 \cdot 25.75000) = 40.80894$$

3.5 Оценим погрешности уравнения регрессии.

3.5.1 Извлечем из S_y^2 квадратный корень получим:

$$S_y = \sqrt{0.20538} = 0.45319;$$

3.5.2 Возведем в квадрат $R_{x,y}$ получим: $R_{x,y}^2 = -0.72028^2 = 0.51880$

3.5.3 Вычислим абсолютную погрешность (остаточное среднее квадратическое отклонение) по формуле (3.4)

$$\sigma_{y/x} = 0.45319 \sqrt{1 - 0.51880} = 0.31437$$

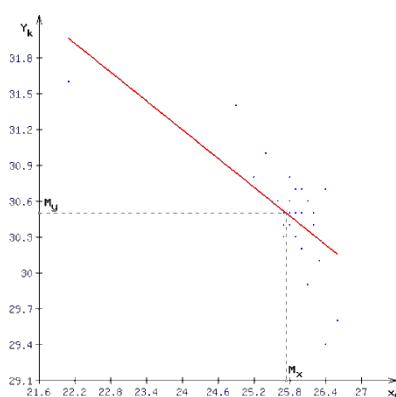
3.5.4 Вычислим относительную погрешность по формуле (3.5)

$$\delta_{y/x} = (0.31437 / 30.50000)100\% = 1.03073\%$$

ОТВЕТ: Уравнение линейной регрессии имеет вид: $Y = 40.80894 -0.40035 X$ (3.6)

Погрешности уравнения: $\sigma_{y/x} = 0.31437$; $\delta_{y/x} = 1.03073\%$

4. Строим диаграмму рассеяния (корреляционное поле) и график линии регрессии.



4.1. Находим минимальный и максимальный элемент выборки X это 18-й и 15-й элементы соответственно, $x_{min} = 22.10000$ и $x_{max} = 26.60000$.

4.2. Находим минимальный и максимальный элемент выборки Y это 2-й и 18-й элементы соответственно, $y_{min} = 29.40000$ и $y_{max} = 31.60000$.

4.3. На оси абсцисс выбираем начальную точку чуть левее точки $x_{18} = 22.10000$, и такой масштаб, чтобы на оси поместились точки $x_{15} = 26.60000$ и отчетливо различались остальные точки.

4.4. На оси ординат выбираем начальную точку чуть левее точки $y_2 = 29.40000$, и такой масштаб, чтобы на оси поместились точки $y_{18} = 31.60000$ и отчетливо различались остальные точки.

4.5. На оси абсцисс размещаем значения x_k , а на оси ординат значения y_k .

4.6. Наносим точки $(x_1, y_1), (x_2, y_2), \dots, (x_{26}, y_{26})$ на координатную плоскость. Получаем диаграмму рассеяния (корреляционное поле), изображенное на рисунке ниже.

4.7. Начертим линию регрессии.

Для этого найдем две различные точки с координатами (x_{r1}, y_{r1}) и (x_{r2}, y_{r2}) удовлетворяющие уравнению (3.6), нанесем их на координатную плоскость и проведем через них прямую. В качестве абсциссы первой точки возьмем значение $x_{min} = 22.10000$. Подставим значение x_{min} в уравнение (3.6), получим ординату первой точки.

Таким образом имеем точку с координатами (22.10000, 31.96127). Аналогичным образом получим координаты второй точки, положив в качестве абсциссы значение $x_{\max} = 26.60000$. Вторая точка будет: (26.60000, 30.15970).

Линия регрессии показана на рисунке ниже красным цветом. Обратите внимание, что линия регрессии всегда проходит через точку средних значений величин X и Y, т.е. с координатами (M_x, M_y).

2. Автокорреляция

Пример 1. Расчет коэффициентов автокорреляции уровней для временного ряда расходов на конечное потребление.

Пусть имеются следующие условные данные о средних расходах на конечное потребление

(y_t , д. е.) за 8 лет (таблица 1).

Таблица 1

Расчет коэффициента автокорреляции первого порядка для временного ряда расходов на конечное потребление, д. е.

t		y_{t-1}	$y_t - \bar{y}_1$	$y_{t-1} - \bar{y}_2$	$(y_t - \bar{y}_1)(y_{t-1} - \bar{y}_2)$	$(y_t - \bar{y}_1)^2$	$(y_{t-1} - \bar{y}_2)^2$
1	7						
2	8	7	-3,29	-3,00	9,86	10,7959	9
3	8	8	-3,29	-2,00	6,57	10,7959	4
4	10	8	-1,29	-2,00	2,57	1,6531	4
5	11	10	-0,29	0,00	0,00	0,0816	0
6	12	11	0,71	1,00	0,71	0,5102	1
7	14	12	2,71	2,00	5,43	7,3673	4
8	16	14	4,71	4,00	18,86	22,2245	16
Итого	86	70	0	0	44	53,4286	38

Разумно предположить, что расходы на конечное потребление в текущем году зависят от расходов на конечное потребление предыдущих лет.

Определим коэффициент корреляции между рядами и измерим тесноту связи между расходами на конечное потребление текущего и предыдущего годов. Добавим в табл. 1 временной ряд .

Одна из рабочих формул для расчета коэффициента корреляции имеет вид:

$$r_{xy} = \frac{\sum (x_j - \bar{x})(y_j - \bar{y})}{\sqrt{\sum (x_j - \bar{x})^2 \cdot \sum (y_j - \bar{y})^2}}.$$

В качестве переменной x мы рассмотрим ряд y_2, y_3, \dots, y_8 , в качестве переменной y – ряд y_1, y_2, \dots, y_7 . Тогда приведенная выше формула примет вид

$$r_1 = \frac{\sum_{t=2}^n (y_t - \bar{y}_1) \cdot (y_{t-1} - \bar{y}_2)}{\sqrt{\sum_{t=2}^n (y_t - \bar{y}_1)^2 \cdot \sum_{t=2}^n (y_{t-1} - \bar{y}_2)^2}}, \text{ где } \bar{y}_1 = \frac{\sum_{t=2}^n y_t}{n-1}, \bar{y}_2 = \frac{\sum_{t=2}^n y_{t-1}}{n-1}$$

Эту величину называют **коэффициентом автокорреляции уровней ряда первого порядка**, так как он измеряет зависимость между соседними уровнями ряда и $t-1$, т. е. при лаге 1.

Для данных примера 1 соотношения (2) составят:

$$\bar{y}_1 = \frac{8+8+10+11+12+14+16}{7} = \frac{79}{7} = 11,29,$$

$$\bar{y}_2 = \frac{7+8+8+10+11+12+14}{7} = \frac{70}{7} = 10.$$

Используя формулу (1), получаем коэффициент автокорреляции первого порядка:

$$r_1 = \frac{44}{\sqrt{53,4286 \cdot 38}} = 0,977.$$

Полученное значение свидетельствует об очень тесной зависимости между расходами на конечное потребление текущего и непосредственно предшествующего годов и, следовательно, о наличии во временном ряде расходов на конечное потребление сильной линейной тенденции.

Аналогично можно определить коэффициенты автокорреляции второго и более высоких порядков. Так, коэффициент автокорреляции второго порядка характеризует тесноту

связи между уровнями и y_{t-2} и определяется по формуле

$$r_2 = \frac{\sum_{t=3}^n (y_t - \bar{y}_3) \cdot (y_{t-2} - \bar{y}_4)}{\sqrt{\sum_{t=3}^n (y_t - \bar{y}_3)^2 \cdot \sum_{t=2}^n (y_{t-2} - \bar{y}_4)^2}}, \text{ где } \bar{y}_3 = \frac{\sum_{t=3}^n y_t}{n-2}, \bar{y}_4 = \frac{\sum_{t=3}^n y_{t-2}}{n-2}$$

Для данных из примера 1 получим:

$$\bar{y}_3 = \frac{8+10+11+12+14+16}{6} = \frac{71}{6} = 11,83,$$

$$\bar{y}_4 = \frac{7+8+8+10+11+12}{6} = \frac{56}{6} = 9,33$$

Построим табл. 2.

Полученные результаты еще раз подтверждают вывод о том, что ряд расходов на конечное потребление содержит линейную тенденцию.

Число периодов, по которым рассчитывается коэффициент автокорреляции, называют **лагом**. С увеличением лага число пар значений, по которым рассчитывается коэффициент автокорреляции, уменьшается. Некоторые авторы считают целесообразным для обеспечения статистической достоверности коэффициентов автокорреляции

использовать правило – максимальный лаг должен быть не больше $\left(\frac{n}{4}\right)$.

Подставив полученные значения в формулу (3), имеем:

$$r_2 = \frac{27,333}{\sqrt{40,8333 \cdot 19,333}} = 0,973$$

Отметим два важных свойства коэффициента автокорреляции:

Во-первых, он строится по аналогии с линейным коэффициентом корреляции и таким образом характеризует тесноту только линейной связи текущего и предыдущего уровней ряда. Поэтому по коэффициенту автокорреляции можно судить о наличии линейной (или близкой к линейной) тенденции. Для некоторых временных рядов, имеющих сильную нелинейную тенденцию (например, параболу второго порядка или экспоненту), коэффициент автокорреляции уровней исходного ряда может приближаться к нулю.

Таблица 2

Расчет коэффициента автокорреляции второго порядка для временного ряда расходов на конечное потребление, д. е.

			$y_t - \bar{y}_3$	$y_{t-2} - \bar{y}_4$	$(y_t - \bar{y}_3)(y_{t-2} - \bar{y}_4)$	$(y_t - \bar{y}_3)^2$	$(y_{t-2} - \bar{y}_4)^2$
1	7						
2	8						
3	8	7	-3,83	-2,33	8,9444	14,6944	5,4444
4	10	8	-1,83	-1,33	2,4444	3,3611	1,7778
5	11	8	-0,83	-1,33	1,1111	0,6944	1,7778
6	12	10	0,17	0,67	0,1111	0,0278	0,4444
7	14	11	2,17	1,67	3,6111	4,6944	2,7778
8	16	12	4,17	2,67	11,1111	17,3611	7,1111
Итого	86	56	0	-4E-15	27,3333	40,8333	19,3333

Во-вторых, по знаку коэффициента автокорреляции нельзя делать вывод о возрастающей или убывающей тенденции в уровнях ряда. Последовательность коэффициентов автокорреляции уровней первого, второго и т.д. порядков называют **автокорреляционной функцией временного ряда**. График зависимости ее значений от величины лага (порядка коэффициента автокорреляции) называется **коррелограммой**.

Анализ автокорреляционной функции и коррелограммы позволяет определить лаг, при котором автокорреляция наиболее высокая, а следовательно, и лаг, при котором связь между текущим и предыдущими уровнями ряда наиболее тесная, т.е. при помощи анализа автокорреляционной функции и коррелограммы можно выявить структуру ряда.

Если наиболее высоким оказался коэффициент автокорреляции первого порядка, исследуемый ряд содержит только тенденцию. Если наиболее высоким оказался коэффициент автокорреляции порядка τ , ряд содержит циклические колебания с периодичностью в моментов времени. Если ни один из коэффициентов автокорреляции не является значимым, можно сделать одно из двух предположений относительно структуры этого ряда: либо ряд не содержит тенденции и циклических, либо ряд содержит сильную нелинейную тенденцию, для выявления которой нужно провести дополнительный анализ. Поэтому коэффициент автокорреляции уровней и автокорреляционную функцию целесообразно использовать для выявления во временном ряде наличия или отсутствия трендовой компоненты (T) и циклической (сезонной) компоненты (S).

2.6.3 Результаты и выводы:

В результате проведенного занятия студенты должны:

- ознакомиться с основными понятиями многомерного статистического анализа, теории корреляции, классификацией регрессий;
- усвоить алгоритмы нахождения условных законов и числовых характеристик многомерных случайных величин, вычисления коэффициента корреляции, детерминации, ковариации;

- выработать навыки нахождения уравнения регрессии, проверки его параметров на статистическую значимость.

2.7 Практическое занятие № 7 (2 часа).

Тема: «Техника экспериментальных измерений. Основные положения теории погрешностей»

2.7.1 Задание для работы:

1. Практика применения основных положений теории погрешностей.

2.7.2 Краткое описание проводимого занятия:

1. Практика применения основных положений теории погрешностей.

Задача. Измерена концентрация активного вещества в шести пробах продукта, получаемого в периодическом химическом процессе. Получены следующие результаты (в г/л) 4,45; 4,40; 4,42; 4,45; 4,38; 4,42. Предполагая, что результаты измерений имеют нормальное распределение, требуется:

- 1) найти точечные несмешанные оценки математического ожидания и среднего квадратического отклонения;
- 2) записать плотность вероятности и функцию распределения СВ X (концентрация вещества);
- 3) найти доверительный интервал, накрывающий математическое ожидание концентрации с заданной доверительной вероятностью $(1-\alpha) = 0,95$, считая s неизвестной;
- 4) найти доверительный интервал, накрывающий неизвестное среднее квадратичное отклонение s с заданной доверительной вероятностью $(1-\alpha) = 0,95$;
- 5) принимая доверительную вероятность $P = 1-\alpha = 0,99$, найти предельную

погрешность, с которой $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ оценивает математическое ожидание a концентрации;

- 6) найти минимальное число проб раствора, концентрации которых надо измерить, чтобы с доверительной вероятностью $(1-\alpha) = 0,95$ можно было бы утверждать, что,

принимая среднее арифметическое \bar{x} за математическое ожидание концентрации, мы совершаляем погрешность, не превышающую $e = 0,5s$, считая $s = S$;

- 7) вычислить $P(4,41 < x < 4,43)$.

Решение:

$$1) \quad \hat{a} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{6} \sum_{i=1}^6 x_i = 4,42 \quad (\text{г/л})$$

$$\text{и} \quad \hat{s} = S = \sqrt{\frac{\sum_{i=1}^{n-1} (x_i - \bar{x})^2}{6-1}} = 0,028 \quad \text{г/л.}$$

- 2) Следовательно, плотность вероятности СВ X (концентрация) имеет вид:

$$f(x) = \frac{1}{0,028\sqrt{2\pi}} \exp\left(-\frac{(x-4,42)^2}{0,016}\right)$$

Функция распределения концентрации имеет вид:

$$F(x) = \int_{-\infty}^x f(x) dx = \frac{1}{0,028\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{(x-4,42)^2}{0,016}\right) dx$$

Используя нормированную функцию Лапласа

$$\Phi(x) = \frac{2}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt$$

можно записать

$$F(x) = \frac{1}{2} + \frac{1}{2} \Phi\left(\frac{x - \bar{x}}{s}\right) = \frac{1}{2} + \frac{1}{2} \Phi\left(\frac{x - 4.42}{0.028}\right)$$

3) Найдем интервальные оценки параметров нормального распределения концентрации. Для нахождения доверительного интервала, накрывающего математическое ожидание, найдем по таблице квантилей распределение Стьюдента по заданной доверительной вероятности $P = 1-a = 0,95$ и числу степеней свободы $n = n-1 = 6-1 = 5$ квантиль

$$t_{\frac{\alpha}{2}, n} = t_{0.025, 5} = 2.571$$

$\frac{t_{\alpha/2, n}}{s}$

Вычислим предельную погрешность интервального оценивания математического ожидания

$$\varepsilon = t_{\frac{\alpha}{2}, n} \frac{s}{\sqrt{n}} = 2.571 \frac{0.028}{\sqrt{6}} = 0.029 \text{ (г/л).}$$

Искомый доверительный интервал, накрывающий математическое ожидание концентрации вещества с заданной доверительной вероятностью $P = 0,95$, равен:

$$\bar{x} - \varepsilon < a < \bar{x} + \varepsilon;$$

$$4,42 - 0,029 < a < 4,42 + 0,029;$$

$$4,391 < a < 4,449.$$

Смысл полученного результата:

если будет произведено достаточно большое число выборок по 6 пробам из бесконечно большой по численности партии химического продукта, то в 95% случаев из них доверительный интервал накроет неизвестное математическое ожидание и только в 5% математическое ожидание может выйти за границы доверительного интервала.

4) Для нахождения доверительного интервала, накрывающего неизвестное среднее квадратическое отклонение s с заданной доверительной вероятностью $(1-a) = 0,95$, найдем по заданной доверительной вероятности 0,95 и числу степеней свободы $n = n-1 = 6-1 = 5$ два числа g_1 и g_2 , т.е. $g_1 = 0,624$ и $g_2 = 2,45$. Искомый доверительный интервал равен:

$$g_1 s < s < g_2 s;$$

$$0.624 * 0.028 < s < 2.45 * 0.028;$$

$$0.017 < s < 0.068.$$

5) Если задать доверительную вероятность $P = 1-a = 0,99$, то предельная погрешность, с

которой среднее арифметическое емкости конденсаторов \bar{x} оценивает неизвестное математическое ожидание, равна:

$$\varepsilon = t_{\frac{\alpha}{2}, n} \frac{s}{\sqrt{n}} = t_{0.005, 5} \frac{0.028}{\sqrt{6}} = 4.032 \frac{0.028}{\sqrt{6}} = 0.046.$$

6) Найдем минимальное число конденсаторов, емкость которых необходимо измерить, чтобы с доверительной вероятностью $P = 1-a = 0,95$ можно было бы утверждать, что,

принимая среднее арифметическое \bar{x} за математическое ожидание концентрации, мы

совершаем погрешность, не превышающую $0,2s = 0,0056$, считая s известным и равны $0,028$.

Искомый объем выборки найдем из соотношения

$$n = \frac{\sigma^2 u_s^2}{\varepsilon^2} = \frac{0.028^2 (1.96)^2}{(0.0056)^2} \geq 96$$

проб.

$$\begin{aligned} P(4.41 < X < 4.43) &= \frac{1}{2} \left[\Phi\left(\frac{4.43 - 4.42}{0.028}\right) - \Phi\left(\frac{4.41 - 4.42}{0.028}\right) \right] = \\ 7) \quad &= \frac{1}{2} [\Phi(0.357) - \Phi(-0.357)] = \Phi(0.357) = 0.279. \end{aligned}$$

2.7.3 Результаты и выводы:

В результате проведенного занятия студенты должны:

- ознакомиться с основными понятиями теории экспериментальных измерений и теории погрешностей;
- усвоить методики определения статистических погрешностей;
- выработать навыки анализа полученных решений.