

**Министерство сельского хозяйства Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования Иркутский государственный аграрный
университет имени А.А. Ежевского**

«Методы научных исследований в информатике»

Учебное пособие

для магистрантов направления подготовки 09.04.03 – Прикладная
информатика

Молодежный, 2020

УДК 001.891

Печатается по решению научно-методического совета Иркутского государственного аграрного университета имени А.А. Ежевского, протокол № 9 от 25 мая 2020 г.

Рецензенты: Галимзянов Т.Р., заместитель начальника отдела сопровождения информационных систем в министерстве финансов Иркутской области; Федурин Н.И., к.т.н, доцент кафедры информатики и математического моделирования.

Бузина Т.С. Учебное пособие «Методы научных исследований в информатике» / Т.С. Бузина- Иркутск: Изд-во Иркутский ГАУ, 2020. – 115 с.

Учебное пособие составлено в помощь магистрантам направления подготовки 09.04.03 – Прикладная информатика для изучения дисциплины «Методы научных исследований в информатике». В работе в комплексе рассмотрены вопросы, связанные с основными понятиями, классификацией и применением методов научных исследований в информатике. В краткой форме изложена теория и приведены примеры выполнения заданий по разделам. Для контроля и систематизации знаний студентов предложены вопросы и задания по каждой теме.

Содержание

1. ОСНОВНЫЕ ПОНЯТИЯ НАУЧНЫХ ИССЛЕДОВАНИЙ В ИНФОРМАТИКЕ	7
1.1. Основные понятия и определения научных исследований	7
1.2. Классификация научных исследований	12
1.3. Методологические основы научных исследований	18
1.4. Классификация методов научных исследований	20
1.5. Контрольные вопросы	25
1.6. Задания для самостоятельной работы	26
2. ИНСТРУМЕНТАЛЬНЫЕ СРЕДСТВА НАУЧНЫХ ИССЛЕДОВАНИЙ	27
2.1. Архитектура информационно- аналитических систем	27
2.2. Применение информационных технологий и систем в научных исследованиях	35
2.3. Статистические пакеты в научных исследованиях	39
2.4. Контрольные вопросы	40
2.5. Задания для самостоятельной работы	41
3. ПРИНЦИПЫ ОРГАНИЗАЦИИ НАУЧНЫХ ИССЛЕДОВАНИЙ В ИНФОРМАТИКЕ	44
3.1. Логические методы в научных исследованиях	44
3.2. Анализ и синтез в научных исследованиях	45
3.3. Метод аналогии в научном исследовании	46
3.4. Сравнение в научных исследованиях	47
3.5. Метод индукции и дедукции	48
3.6. Подходы к организации исследовательской работы	48
3.7. Контрольные вопросы	54
3.8. Задания для самостоятельной работы	55
4. ЭМПИРИЧЕСКИЕ МЕТОДЫ ИССЛЕДОВАНИЙ В ИНФОРМАТИКЕ	57
4.1. Метод наблюдения	57
4.2. Метод экономического эксперимента	60
4.3. Контрольные вопросы	63
4.4. Задания для самостоятельной работы	63
5. ТЕОРЕТИЧЕСКИЕ МЕТОДЫ ИССЛЕДОВАНИЙ В ИНФОРМАТИКЕ	65
5.1. Методы формализации и математизации	65
5.2. Математическое моделирование как метод научных исследований в информатике	66
5.3. Методы оптимизации	68
5.4. Статистические методы в научных исследованиях	71
5.5. Методы прогнозирования в научных исследованиях	85
5.6. Контрольные вопросы	102
5.7. Задания для самостоятельной работы	103
Глоссарий	105
Литература	113

ВВЕДЕНИЕ

Стремительное увеличение объема существующей и циркулирующей в обществе информации ставит современного человека перед проблемой умения работать с ней: находить, отбирать то, что необходимо, хранить, упаковывать и быстро извлекать из хранилища, обрабатывать и преобразовывать. При этом информация все чаще может быть представлена не только в текстовом, наиболее привычном виде, но и в виде видео-и аудиоматериалов, диаграмм и анимационной графики и т.д.

Знание методов, приемов и средств работы с информацией становится одним из основных профессионально важных качеств специалиста любого профиля.

Метод научного исследования — это способ познания объективной действительности, представляющий собой определенную последовательность действий, приемов, операций.

Цель дисциплины - формирование у магистрантов представления о методах научных исследований, используемых на практике и в академической среде, об инструментальных средствах и технологиях для реализации исследований в информатике.

Задачи дисциплины:

- изучение методов научных исследований в области информатики;
- характеристика возможностей современных информационных технологий и систем для реализации научных исследований в информатике;
- формирование у магистрантов логического мышления, необходимого для проведения научных исследований;
- развитие аналитических способностей, и выработка системного видения процессов, происходящих во внешней бизнес-среде и внутри организации;

В процессе освоения дисциплины магистры приобретут следующие компетенции, направленные на формирование:

- навыков использования количественных и качественных методов для проведения научных исследований;
- навыков самостоятельного освоения новых методов исследования, создания научно-производственного профиля своей профессиональной деятельности;
- способности готовить аналитические материалы для управления бизнес-процессами и оценки их эффективности.

Пособие состоит из 5 глав, направленных на изучение методов научных исследований в информатике и их применения в исследовательской работе.

Первая глава содержит основные понятия и определения, лежащие в основе научных исследований. На основании этого выполняется систематизация методологий и методов научных исследований. Изучаются вопросы топики и риторики, являющиеся основой для изложения научных результатов, дискуссий и научных докладов.

Во второй главе описаны возможности информационных технологий и систем как аналитического инструментария для решения задач научных исследований, приведен обзор функциональных возможностей аналитических платформ и приложений, используемых в настоящее время для выполнения исследований в информатике.

В третьей главе приведена характеристика основных логических методов, к числу которых относятся такие методы как методы классификации, обобщения и типологии, анализ и синтез, аналогия сравнительный анализ в информатике и метод дедукции и индукции.

Четвертая глава посвящена рассмотрению эмпирических методов исследований и содержит основные понятия и принципы классификации эмпирических методов. Особое место уделено наблюдению как эмпирическому методу исследований, методологии и организации проведения качественных исследований.

Теоретические методы исследований, практика их применения рассмотрены в пятой главе пособия. Существенное значение в эффективном применении теоретических методов исследований представляет процесс формализации, обеспечивающий построение моделей, на основании которых формируется новое знание об объекте исследования. В учебном пособии уделено особое внимание методам построения математических моделей, проверке их адекватности, инструментальным средствам моделирования. С учетом того, что при выполнении научных исследований в информатике, достаточно часто приходится сталкиваться со случайными событиями, в материал пособия включены общая характеристика статистических методов и детальное описание основных методов статистического анализа, получивших широкое распространение в научных исследованиях.

Кроме того в пятой главе рассмотрены основные методы прогнозирования как наиболее востребованные в решении задач бизнес-аналитики и научных исследований. Поскольку в исследованиях часто приходится иметь дело с данными, представляющими собой историю

изменения объекта исследования во времени, то наиболее подробно в пособии изложены статистические методы прогнозирования временных рядов.

Учебное пособие разработано в соответствии с требованиями ФГОС ВО 3++ по направлению подготовки 09.04.03 – Прикладная информатика, квалификация «Магистр».

1. ОСНОВНЫЕ ПОНЯТИЯ НАУЧНЫХ ИССЛЕДОВАНИЙ В ИНФОРМАТИКЕ

1.1. Основные понятия и определения научных исследований

В условиях цифровизации экономики все большее значение принимает внедрение и развитие новых методов хозяйствования на основе современных информационных технологий. Значительную роль при этом играют научные фундаментальные и прикладные исследования. В Федеральном законе закона РФ от 23 августа 1996 г. «О науке и государственной научно-технической политике» научно-исследовательская деятельность определена как деятельность, направленная на получение и применение новых знаний.

Концептуальную модель процесса исследования можно представить следующим образом (рис. 1.1) [21].

При выполнении научно-исследовательской работы входными данными являются:

- проблемы, выделенные исследователем для решения;
- методы, модели, законы, закономерности и т.д.;
- новейшие результаты, полученные отечественными и зарубежными исследователями в данной области;
- научная информация, на базе которой формируются новые знания.

Управляющими факторами, под действием которых выполняется преобразование входных потоков информации в выходные, являются парадигмы, концепции, методологии, уровень мировых достижений в достижений в решении рассматриваемого круга задач.

Результатом (выходом) научно-исследовательской работы выступают новые знания (закономерности, методологии, принципы, модели, методы и т.д.) и новая научная информация, которая может быть оформлена в виде научных отчетов, диссертаций, статей, монографий, докладов.

Научно-исследовательская деятельность может выполняться коллективами или отдельными физическими лицами с применением необходимого инструментария, технических средств и финансирования.

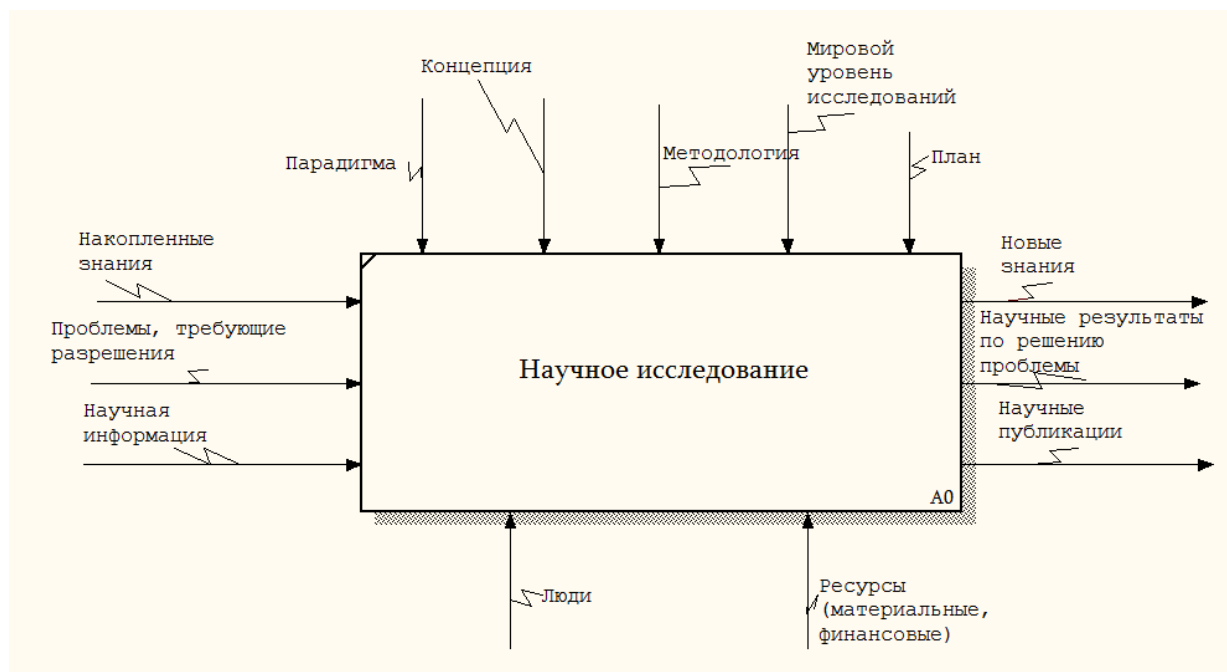


Рисунок 1.1– Концептуальная модель процесса исследования

К основным характеристикам исследования относятся: объект и предмет исследования, тема, цель и задачи исследования, научная идея или концепция, теоретическая и практическая значимость, выводы исследования и сформулированные практические и теоретические рекомендации.

Тема научного исследования отражает определенный аспект проблемы, решение которой будет ставиться в качестве цели исследования, и содержит в себе направленность на принятие решения.

Цель исследования - определение конкретного объекта и всестороннее, достоверное изучение его структуры, характеристик, связей на основе разработанных в науке принципов и методов познания, а также получение полезных для деятельности человека результатов, внедрение в производство с дальнейшим эффектом.

Носитель проблемной ситуации, на которую направлена исследовательская работа представляет собой *объект исследования*.

Предметом исследований является свойство, поведение, функционирование объекта, подлежащего непосредственному изучению. Объект и предмет исследования соотносятся между собой как общее и частное.

Концепция – руководящая идея или совокупность идей, предлагающих новую теоретическую базу для понимания и исследования явлений или процессов, происходящих в обществе, система взглядов на что-либо,

основная мысль, когда определяются цели, задачи исследования и указываются пути его ведения.

Задачи исследования формулируются как конкретные этапы решения представленной проблемы и являются, по сути, содержательной детализацией поставленной цели.

Оценка исследования осуществляется как с точки зрения теоретической ценности (новизна исследования, актуальность, оригинальность), так и с позиций практической значимости.

В конце исследования формируется заключение, содержащее основные выводы и результаты по каждой задаче. Формирование новых знаний основано на использовании различных видов коммуникативного взаимодействия, к которым относятся устная и письменная речь, естественные и искусственные языки[21].

В процессе обмена научными знаниями между людьми в различных областях человеческой деятельности используются различные виды коммуникации. Среди них важную научную роль играют дисциплины по структуре и формированию устной и письменной речи, среди которых можно выделить традиционную риторику (*bene dicendi scientia*– "наука о хорошей речи").

Риторика (греч. *rhetorike* «ораторское искусство»), научная дисциплина, изучающая закономерности порождения, передачи и восприятия хорошей речи и качественного текста.

Риторика, как систематическая дисциплина, развивалась в Древней Греции, когда к навыкам и умению публичного выступления предъявлялись очень высокие требования. Искусство публичного выступления преподавалось в школах софистов, главной задачей которых было научить красноречию чисто практическим путем. Коракс, Исократ, Аристотель, Цицерон и Квинтилиан справедливо считаются величайшими теоретиками риторики. Исторически сложилось два основных направления в определении предмета и целей риторики. Основоположниками первого направления являются Исократ (около 436-388 гг. до н. э.). Он считается основоположником "литературной" риторики, уделявшим первостепенное внимание письменной речи и одним из первых ввел понятие композиции ораторского произведения. Представители этого направления были склонны считать хорошей пышную, построенную по канонам эстетики речь. Убедительность по-прежнему имела значение, но она не была единственным или главным критерием оценки речи.

Второе направление, идущее от Аристотеля (384–322 до н.э.), выделяло достоинства речи по таким признакам как убедительность, способность завоевать признание (согласие, симпатию, сочувствие) слушателей, заставить их действовать определенным образом. Аристотель определял риторику как «способность находить возможные способы убеждения относительно каждого данного предмета». Аристотель явился основоположником превращения риторики в научную дисциплину, поскольку именно им была установлена неразрывная связь между риторикой, логикой и диалектикой. В основных произведениях, посвященных риторике («Риторика», «Топика» и «О софистических опровержениях»), Аристотель указал место риторики в системе наук античности.

Позднее, под влиянием экономических и политических изменений в жизни общества, сформировались новые требования к речевой практике. Возросло влияние "логического" направления, и на смену старой риторике пришла неориторика. Начало этого периода относят ко времени, когда трактат Х был опубликован во Франции. Перельман и Л. Ольбрехт-Титека "новая риторика. Трактат об аргументации" (1958), в котором риторическая система Аристотеля была критически развита на основе современных данных. Авторы трактата рассмотрели взаимосвязь логики и аргументации, понятия аудитории, диалога, двусмысленности, презумпций, топоса, нормативности, ошибок аргументации, категорировали аргументы и подробно проанализировали отдельные категории. Неориторика не отвергала результатов, полученных в "литературном" направлении, что позволяет говорить о взаимообогащении "логического" и "литературного" направлений с доминированием первого.

С точки зрения «логического» направления риторика является наукой о методах, способах и формах убеждения, принципах языкового воздействия на аудиторию, сущности эффективной коммуникации, обеспечивающих достижение желаемой цели.

В "литературном" направлении риторика - это филологическая дисциплина, направленная на изучение способов построения художественной речи, прежде всего прозаической и устной; она тесно связана с поэтикой и стилистикой.

Традиционно различают *общую* и *частную риторику*. Общая риторика представляет собой науку об универсальных принципах и правилах построения хорошей речи, не зависящих от конкретной сферы речевой коммуникации. Частная риторика рассматривает особенности отдельных видов речевой

коммуникации в связи с условиями коммуникации, функциями речи и сферами деятельности человека.

Риторика можно разделить на теоретическую, прикладную и тематическую. Задача теоретической риторики-изучить правила построения качественной речи. Прикладная риторика ориентирована на применение и распространение правил, образцов и закономерностей устной речи, выявленных в ходе теоретических исследований, в практике преподавания литературы. Теоретическая и прикладная риторика идентичны научной и образовательной риторике. Тематическая риторика рассматривает объединение различных видов литературы вокруг соответствующих тем. Примерами могут служить обсуждения важных политических событий (президентские выборы), спортивных событий (Олимпиада) или других вопросов, затрагивающих значительное число людей в стране или мире.

Каноны или правила риторики предполагают наличие устоявшихся частей или этапов риторического развития речи.

Инвенция (*inventio*) – нахождение или создание материала речи или текста–первый этап создания речи, предполагающий разработку проблематики, связанной с предметной областью речи.

Изобретение основано на отборе материала и поиске средств убеждения и требует от исследователя не только определения темы, но и выбора способов ее раскрытия, формулирования аргументов в пользу защищаемой диссертации, построения тематического и содержательного пространства речи. Основными критериями отбора материала являются коммуникативная интенция автора и характер аудитории, к которой он собирается обратиться. Очевидно, что при защите магистерской или докторской диссертации аудитория должна учитывать принятый в научном мире стиль изложения доклада, особенности предметной области, хронометраж выступлений, что предъявляет особые требования к доказательной базе расширенных тезисов и краткости изложения.

Диспозиция (*dispositio*) – расположение или композиция материала. На этом этапе определяется расположение материала, его хронологическая или логическая последовательность. В основе канона «расположение» лежит учение о хрии, или о композиции речи, которые составляют базу таких современных дисциплин как учение о литературной композиции и теория композиции как часть теории текста.

Представленный материал может содержать от трех (введение – основная часть – заключение) до семи (введение – определение темы с ее

подразделениями – изложение – отступление – аргументация или доказательство собственного тезиса – опровержение – заключение) структурных единиц, что определяется объемом материала, его характером и целями изложения.

Элокуция (elocutio) – словесное выражение или дикция.

Изложение или словесное выражение должно отвечать четырем критериям: правильности (соответствовать правилам грамматики, орфографии и произношения), ясности (состоять из общепонятных слов в общих сочетаниях, по возможности не включать абстрактные, заимствованные и другие слова, которые могут быть непонятны аудитории), изящества или декоративности (быть более эстетичным, чем обыденная речь) и актуальности (гармония темы и языковых средств). Эти составляющие канона "словесное выражение" легли в основу современной науки о культуре речи.

Память, запоминание (*memoria*) – раздел риторики, рекомендующий как лучше запомнить необходимую, для выступления, речь. Этот канон предназначен для ораторов, которым необходимо запомнить свои подготовленные речи для последующего публичного воспроизведения. В научных дискуссиях, докладах и презентациях результатов исследований этот канон имеет большое значение, особенно когда необходимо запомнить и аргументированно представить значительный объем информации. В этом разделе риторика предлагает список техник, основанных на сложных визуальных образах. Выступление, выступление (*actio*) предполагает непосредственное выступление оратора перед аудиторией с подготовленным докладом, речью, презентацией, где образ оратора и содержание речи сливаются в единое целое. Цель такой речи – достойно и красиво изложить свои мысли и мнения, сделать свою речь запоминающейся.

Таким образом, риторика является основой для представления научных результатов, дискуссий и научных докладов как видов различных знаний, полученных исследователем. Основные каноны риторики используются также при написании научных публикаций и научных статей, создавая структурную основу для научной работы.

1.2. Классификация научных исследований

Принципы классификации научных исследований основаны на критериях, используемых в типологии научных исследований [21]. В Федеральном законе от 23 августа 1996 г. «О науке и государственной научно-

технической политике» даны понятия *фундаментальных* и *прикладных* научных исследований. *Фундаментальные научные исследования* ориентированы на познание явлений окружающего мира, на выявление новых связей и отношений между объектами и явлениями реальной действительности, открытие новых фундаментальных законов и принципов, создание методологий и технологий [21].

Данный тип исследований связан со значительным риском и неопределённостью, с точки зрения получения конкретного положительного результата, вероятность которого не превышает 10%. Несмотря на это, именно фундаментальные исследования составляют основу развития, как самой науки, так и общественного производства.

Прикладные исследования направлены на получение новых знаний, необходимых для удовлетворения новых практических потребностей общества. Результаты практических исследований используются для совершенствования существующих технологий, материалов и машин или для выявления новых моделей и принципов, разработки новых методов и инструментов, методологий и моделей.

Поисковыми называются научные исследования, направленные на определение перспектив исследований и поиск путей решения научных проблем.

Научные разработки — завершающая часть фундаментальных или прикладных исследований, которая, как правило, включает в себя создание опытных образцов или пилотных проектов, на основе которых формируется решение о дальнейшем развитии исследуемого объекта или процесса.

В таблице 1.1 представлена классификация научных исследований по различным признакам.

Таблица 1.1– Классификация научных исследований

Классификационный признак	Виды научных исследований
Целевое назначение научного исследования	Фундаментальные Прикладные Поисковые Разработки
Источник финансирования	Бюджетные Хоздоговорные Не финансируемые
Длительность научного исследования	Долгосрочные Краткосрочные Экспресс - исследования

Роль в науке	Прорывные Развивающие
Способ инициации научного исследования	Инициативные (спонтанные) Систематизированные Заказные

По критерию "длительности" научные исследования можно разделить на долгосрочные, краткосрочные и экспресс-исследования. Для фундаментальных исследований характерны многолетние исследовательские программы. Прикладные исследования, как правило, чаще всего носят краткосрочный характер, длящийся до 1 года. Примером может служить изучение внешней среды организации, изучение факторов, влияющих на экономику региона или отрасли, изучение явлений и процессов, происходящих в социальной сфере, экономике или управлении под влиянием тех или иных факторов. Экспресс - исследование проводится в основном в информационных целях и выполняется в кратчайшие сроки, по упрощенной схеме, с учетом ограниченных временных ресурсов.

Особое место в классификации научных исследований занимают прорывные исследования, которые изначально или в процессе развития ориентированы на получение принципиально новых знаний, преобразующих фундаментальные представления о природе и обществе. Примером могут служить исследования в области нанотехнологий, фармакологии и биологии.

Развивающие исследования направлены на углубление прорывных исследований с целью дальнейшего внедрения результатов во всех сферах жизнедеятельности человека.

Таким образом, в соответствии с представленной в таблице 1.1 классификацией магистерская работа является научным исследованием, относящимся к виду прикладного или поискового, нефинансируемого, краткосрочного, развивающегося, инициативного исследования.

Исходным моментом научного исследования является выявление проблемы. В общем случае под термином «проблема» понимается противоречие в познании, характеризующееся несоответствием между новыми фактами и данными и старыми способами их объяснения []. В научных исследованиях решение проблемы предполагает получение новых знаний для развития теории и совершенствования практики в экономической деятельности и управлении. Однако в данном случае перед исследователем стоит задача определить, каким путем следует идти к устранению проблемы, то есть выбрать такой подход, который допускает ее полное или частичное

решение. При выявлении проблемы исследователь уже формирует некоторое априорное знание об объекте исследования, которое позволяет ему сформулировать возможные пути решения проблемы. Такое научное предположение, вытекающее из теории или практики, которое еще не подтверждено или опровергнуто и нуждается в научном доказательстве (опровержении), называется гипотезой (от греч. гипотезы - основание, предположение). Гипотеза не является догмой и может быть уточнена, дополнена или изменена в ходе исследования. Первоначальный вариант гипотезы называется рабочей гипотезой. Чтобы превратить его в научную гипотезу, научная гипотеза должна быть проверяемой, т.е. следствия, сформулированные логическим путем на основе дедукции, должны поддаваться проверке и соответствовать фактам, результатам опытов, наблюдениям и т.д.

Научная гипотеза должна соответствовать достаточной общностью и предсказательной силой, т.е. объяснять не только те явления, из рассмотрения которых она возникла, но и все связанные с ними явления и процессы.

Научная гипотеза не должна быть внутренне противоречивой. Различают описательные и объяснительные виды гипотез. Описательные гипотезы □ это предположения о свойствах и характеристиках изучаемых явлений, процессов и объектов, основанные на эмпирических фактах. Объяснительные гипотезы □□ это обоснованные предположения о причинно-следственных зависимостях между изучаемыми процессами. Они выявляют сущность и оценивают силу причинно-следственных связей.

Выдвижение гипотезы - лишь начальный шаг исследования, который является идеей решения проблемы. Выбор подхода к научному исследованию определяет путь решения проблемы. При его выборе значительную роль играют знания, накопленные об объекте исследования, свойства и характеристики объекта, цели и задачи исследования, допущения и ограничения, возможные при проведении исследований.

Для систематизации существующих подходов целесообразно их классифицировать. Классификационными признаками могут быть степень формализации проблемы, Глубина исследования, а также совокупность признаков, заложенных в реализацию исследовательского подхода, которые необходимы для выявления и учета характерных свойств объекта.

В таблице 1.2 представлена классификация подходов к научным исследованиям.

Таблица 1.2 – Классификация подходов к научным исследованиям

Классификационный признак	Виды научных исследований
Степень формализации проблемы	«Жесткий» «Мягкий»
Глубина исследования	Клинический Исследовательский Проектирующий
Возможность учета характерных свойств объекта исследования	Аналоговый Редукционистский (аддитивный) Комплексный Системный Ситуационный Диалектический Логический Прагматический Исторический и др.

Когда цель определена и формализована, а задача предполагает поиск оптимального решения при существующих ограничениях, применяется так называемый "жесткий" подход. Этот подход использует формальный исследовательский аппарат и исключает факторы, не подлежащие формализации. Примерами могут служить проблемы неэффективного использования оборудования, высоких производственных затрат, оптимальных инвестиций и т.д.

Для решения задач с недостаточными знаниями об объекте или его непредсказуемом поведении, со значительной неопределенностью информации и высокими рисками используется "мягкий" подход, основанный на логических методах.

Исследовательский подход ориентирован на применение математических методов и моделей с целью получения оптимального в соответствии с заданными целями решения, но не исключает комбинации с клиническим подходом. Примером может служить исследование, направленное на снижение затрат в хозяйственной отрасли или на предприятии или исследования, направленные на формирование оптимального пакета инвестиционных проектов с минимизацией рисков.

Проектирующий подход в качестве целевой установки ориентирован на изменение свойств объекта исследования или внешнего окружения, вызывающего проблемную ситуацию. Примером может служить реинжиниринг бизнес-процессов в организациях.

С точки зрения учета характерных свойств объекта в процессе исследования выделяют аналоговый, редуccionистский, комплексный, системный, ситуационный, диалектический, логический, прагматический и другие подходы.

Аналоговый подход к исследованиям используется, как правило, в тех случаях, когда проблема или объект исследования мало изучены. В этом случае в основу исследований ложатся известные характеристики и свойства аналога, отобранного исследователем на основе имеющихся знаний и опыта.

Редуccionистский (аддитивный) подход основан на предположении, что свойства объекта сводятся к сумме свойств его элементов. Примером могут служить региональные исследования, базирующиеся на исследованиях субъекта региона.

Достаточно широко находит применение комплексный подход к исследованиям, когда объект рассматривается во всей его полноте с различных точек зрения, а акцент делается на многоаспектности исследования, на рассмотрении объекта, в том числе и элементарного (несистемного), с разных его сторон. Например, исследование с целью повышения эффективности предприятия предполагает применение комплексного подхода, затрагивающего финансовую деятельность предприятия, оценку компании с точки зрения клиентов, исследование внутренних бизнес-процессов и деятельности персонала[21].

Системный подход ориентирован на проведение исследования не столько во всей полноте, сколько во взаимодействии его составных частей. В его основе лежит целостное рассмотрение объекта исследования как совокупности компонентов, их внутренних взаимосвязей и связей с внешней окружающей средой. Примером могут служить исследования, направленные на комплексную автоматизацию хозяйствующих субъектов.

Ситуационный подход предполагает исследование объекта с учетом конкретной ситуации как совокупности внешних и внутренних факторов.

При диалектическом подходе исследователь использует объективно обусловленные, научно разработанные, практически проверенные и наиболее эффективные в конкретных условиях диалектические принципы исследования, например, взаимоотношений общего и частного, качества и количества, целого и части, принципов комплексности, системности, развития и пр.

Исторический подход основан на рассмотрении объекта исследования в некотором временном интервале, рассматривая его свойства и характеристики

в прошлом, настоящем и будущем.

1.3. Методологические основы научных исследований

Основной целью научной деятельности является получение точных и всесторонних научных знаний об окружающем нас мире и его составляющих элементах.

Научное знание – система знаний о законах природы, общества, мышления. Это особый вид познания, который, согласно современным взглядам ученых, характеризуется прежде всего возможностью сопоставления с некоторой объективной реальностью. Потребность в научном знании возникает тогда, когда отсутствует представление об объекте в рамках обыденного мышления и обыденного знания. При этом на основе сопоставления с действительностью конкретные системы научного знания оцениваются как полные, истинные, адекватные, объективные, точные и т. д.

В познавательном процессе выделяют два уровня научных исследований: теоретический и эмпирический. *Теоретические исследования* ориентированы на всестороннее изучение и познание объективной реальности с целью выявления общих для данной предметной области связей и обобщений, составляющих основу развития теории. *Эмпирические исследования* направлены на объект, на конкретное изучение процессов и явлений с целью фактов науки и их обобщений[21].

Любое научное исследование базируется на *методологии*, определяющей последовательность привлечения различных форм знаний, позволяющей осуществить основные принципы научных исследований: объективность, воспроизводимость, доказательность (верификацию) и точность полученных результатов. Дословный перевод с греческого термина «методология» (от греч. *methodos* - путь *logos* - слово, учение) – «учение о методах», однако в литературе под этим термином понимают как учение о научном методе познания, так и совокупность методов, применяемых в какой-либо области науки.

Методология как способ изучения действительности многомерна и многомерна, а также как подходы к научному исследованию могут быть классифицированы с использованием различных признаков. Первой наиболее распространенной классификацией методологий является их деление по принадлежности к той или иной области знания (например, философская

методология, историческая методология, экономическая методология и т. д.). популярным является деление методологий на три уровня своей общности:

- универсальный;
- общий;
- частный.

К *универсальной методологии* относится философия. Именно эта наука дает ученым различных областей знания ориентиры относительно сущности познавательного процесса, его форм, уровней, исходных предпосылок, условий его достоверности и истинности, социально-исторического контекста познания. Философия формулирует общие принципы, используемые учеными, явно или неявно, при осуществлении исследовательского процесса, в том числе и в информатике. Например, философский принцип познаваемости мира относится к познаваемости любых экономических явлений и процессов и означает, что изучаемое экономическое явление познаваемо. Философский принцип целесообразности широко используется в менеджменте: "все, что теоретически возможно и не запрещено законами природы, практически осуществимо." Этот принцип лежит в основе исследований, направленных на управление экономическими процессами и явлениями посредством целенаправленного воздействия на них.

Общая методология ориентирована на применения в нескольких науках. Примером этих методологий могут служить математическая, статистическая, историческая и другие методологии. Очевидно, что эти методы используются не только в математике или истории, но и в экономике, менеджменте, социологии и т. д.

Частные методологии применяются только в одной конкретной науке, например, экономическая методология ориентирована на познавательную деятельность в экономике.

Выбор методологии, как и определение проблемы, является своего рода "искусством", которым необходимо овладеть с первых шагов исследования. Успех в этом направлении основывается на интуиции, сформированной под влиянием:

- знание дисциплины и навыки работы в различных направлениях;
- фундаментальные знания не только в области своей дисциплины, но и в других смежных дисциплинах (для экономистов и менеджеров-математика, информационные технологии, социология, психология и др.);
- анализ современной научной периодики;
- знание круга авторов, работающих над изучаемой проблемой, и

объективная оценка результатов их исследований.

1.4. Классификация методов научных исследований

Успех и достижение исследовательских целей во многом определяются используемыми методами исследования.

Метод (от греч. *methodos* – путь, способ исследования, обучения, действия) — определенная совокупность устойчивых правил, предназначенных для достижения какой-либо цели.

Метод – это прием мышления или практического действия, а также средство или инструмент для исследования какого-либо объекта. «Под методом я разумею достоверные и легкие правила, строго соблюдая которые человек никогда не примет ничего ложного за истинное, и шаг за шагом приумножая знание, придет к истинному познанию всего того, что он будет способен познать» (Р. Декарт).

Выбору адекватных методов исследований способствует их классификация, которая позволит в последующем провести разделение единиц исследования на группы (классы) и облегчит анализ объекта исследования[21]. Рассмотрим варианты классификации, представленные в таблице 1.3.

Таблица 1.3 – Методы научных исследований

Классификационный признак	Виды методов	Примеры
1	2	3
Подходы к исследованию	Научные методы	Методы, основанные на рациональном, доказательном и системном исследовании. Обладают свойствами строгости, однозначности, эффективности, простоты, эвристичности.
	Ненаучные методы	Апелляция к религии, авторитетам, идеологии, обыденных знаниях
Методы научных исследований		
Содержание объекта исследования	Методы естествознания	Методы, применяемые в естествознании
	Методы социально-гуманитарных наук	Методы, применяемые в социально-гуманитарных науках

Отрасль науки	История Философия Экономика Социология и др.	Группы методов, применяемые в конкретной науке
Сфера применения и общность методов	Всеобщие	Философия
	Общие	Математика, статистика, виды эксперимента, виды наблюдений, моделирование и др.
	Частные (специальные)	Экономическое наблюдение Экономический эксперимент и др.
Уровень познания	Эмпирические методы	Наблюдение, сравнение, описание, измерение, анкетный опрос, собеседование, тестирование, эксперимент
	Теоретические методы	Аксиоматический метод. Гипотетический метод. Формализация Абстрагирование. Логические методы (анализ, синтез, индукция, дедукция, аналогия)
Система основания	Формальные	Аналитические методы или методы элементарной математики (функциональный анализ) и методы математического анализа (интегральные, дифференциальные, вариационные исчисления), вероятностно-статистические методы (математическая статистика, теория вероятности), методы исследования операций, методы теории выбора и принятия решений, методы математической логики, математическое и имитационное моделирование и др.
	Эвристические методы	Индукция, дедукция, синтез и др.
Степень применения логики	Логические	Дедукция, индукция, аксиоматика, аналогия, доказательство, логическое обоснование и др.
	Нелогические	Экономическое наблюдение, интуиция и др.
Логические методы исследований		
Система основания	Количественные	Математические методы Статистические методы
	Качественные	Наблюдение, анкетирование, опрос, собеседование и т.д.

В классическом подходе принято различать научные и ненаучные методы исследования. Ненаучные методы-это методы, основанные на ограниченно-рациональном, несистематическом, недоказанном подходе к процессу исследования. Эти методы включают в себя обращение к авторитетам, религии и повседневным экономическим знаниям.

В рамках данного пособия рассматриваются только научные методы, представляющие собой способы освоения действительности, основанные на рациональных, научно обоснованных и систематических исследованиях.

Признаками научных методов являются их *строгость, однозначность, эффективность, простота, эвристичность*.

Под *строгостью* в научном методе понимается его рациональность, доказательность, согласованность всех структурных элементов теоретического построения.

Внутренняя непротиворечивость, соответствие друг другу по смыслу утверждений, присутствующих в теории характеризует *однозначность* научного метода.

Эффективность научного метода определяется возможностью получить решение проблемы за конечное число шагов.

Простота (экономичность) научного метода – это достижение научного результата при минимальном количестве действий.

Эвристичность метода - это его способность приносить новый результат (знание), которое может быть использовано или распространено в новые области данной науки или в других областях знаний.

В зависимости от содержания объектов исследования различают методы естествознания и методы социально-гуманитарных наук. При более детальной классификации выделяют методы по отраслям науки: математические, биологические, правовые, социально- экономические и т.д.

Аналогично классификации методологий научных исследований все методы могут быть разделены на всеобщие, общие и частные. Эти понятия тесно связаны с понятиями всеобщей, общей и частной методологиями.

В зависимости от уровня знаний существуют методы эмпирического и теоретического уровней. Эмпирический уровень включает в себя знания, полученные в результате материальной практики или через некоторый непосредственный контакт с реальностью. Эмпирические методы - это методы познания действительности, действующие на уровне опыта, такие как наблюдение, измерение, эксперимент. Использование этих методов предполагает обращение в той или иной форме к деятельности органов чувств

человека, опирающейся на чувственные формы отражения действительности. Теоретическое знание - это результат обобщения, абстрагирования, введения идеализированных объектов, математизации и т. д. теоретические методы включают использование аналогии, гипотезы, интерпретации, применение математического аппарата, различных моделей, а также историко-эволюционный анализ в экономике и др.

Достаточно широко распространена классификация, при которой научные исследования строятся на двух основополагающих классах методов: формальных и эвристических[21]. Формальные методы опираются на точные абстрактные языки (математические, формальной логики и др.), модели, объекты. К формальным методам относят, например:

- аналитические методы или методы элементарной математики (функциональный анализ) и методы математического анализа (интегральные, дифференциальные, вариационные исчисления);
- вероятностно-статистические методы (математическая статистика, теория вероятности);
- методы исследования операций;
- методы теории выбора и принятия решений;
- методы математической логики;
- математическое и имитационное моделирование.

Слово «эвристика» происходит от греческого слова *heurisco*, оно означает «узнавать новое», «открывать». В этой связи под *эвристическими методами* понимают специальные логические способы решения задач, построенные на методах научного познания, имитирующих процесс человеческого мышления, на использовании специальных правил, приемов, обобщений, различные процедуры, направленные на сокращение перебора вариантов.

К эвристическим методам относят:

- методы как совокупность присущих человеку механизмов, с помощью которых осуществляется решение творческих задач (дедуктивные и индуктивные методы, метод аналогий, анализ и синтез) и относящиеся к общим методам научного познания;
- методы, направленные на сокращение времени решения задач (процедура направленного перебора, матричные методы и др.).
- методы экспертных оценок.

Еще один способ классификации методов научных исследований - это их деление на логические и нелогические методы. Логические методы

основаны на логике: формальной и математической и представляют собой знаково-символьную систему, отражающую инструменты и результаты научного познания. Примерами логических методов могут служить методы индукции, дедукции, формализации и математизации, метод интерпретации в объяснении фактов и явлений, метода аксиоматизации в построении экономических теорий и доказательств, метода аналогий и др.

Нелогические методы являются своего рода антиподами логическим методам, и по своей структуре не связаны с формальной и математической логикой. При этом их значение в научных исследованиях достаточно велико. Например, интуиция как нелогический метод исследования достаточно часто опережает логические методы исследований.

Логические методы исследований в зависимости от того, какие системы обоснования в нем используются, подразделяются на количественные и качественные.

Количественные методы исследования базируются на использовании измеряемых величин, выраженных, как правило, в числах. К числу количественных методов относят математические и статистические методы, однако в прикладном значении эти методы объединяют в один общий метод - эконометрический. Качественные методы исследований базируются на словесных описаниях, интерпретации, истолковании и объяснении свойств и характеристик исследуемого объекта.

При выполнении научных исследований понятие метода следует отличать от понятий техника, процедура, методика научного исследования. На рисунке 1.2 представлена взаимосвязь этих понятий.

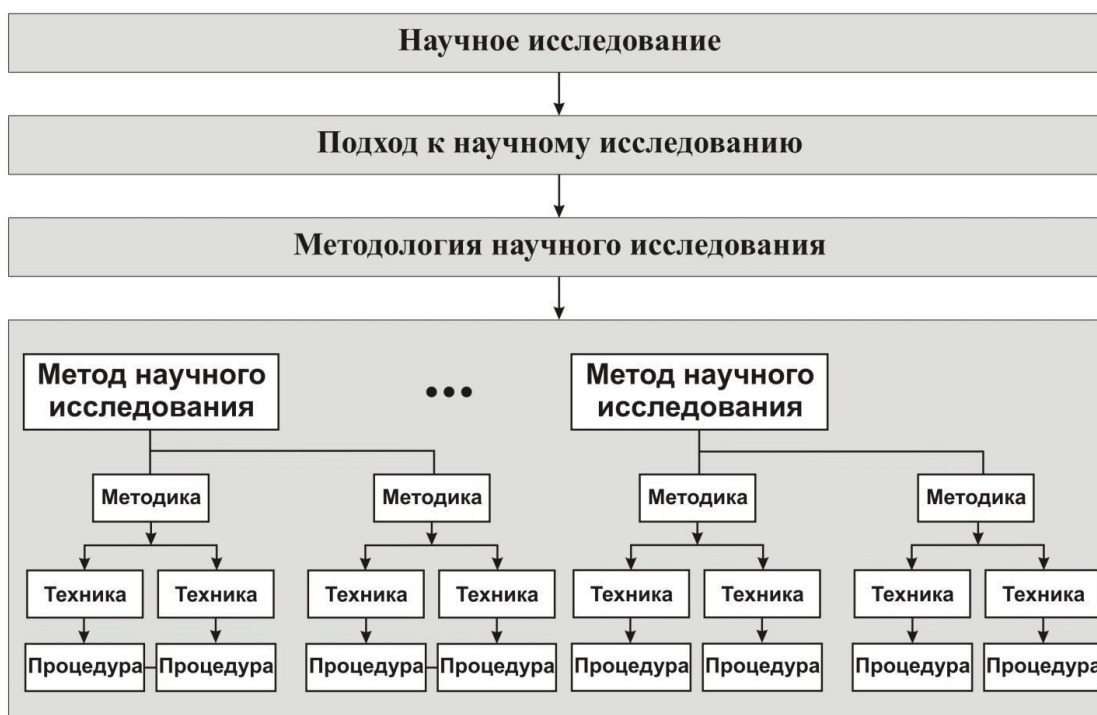


Рисунок 1.2–Взаимосвязь понятий научных исследований

Методология исследования– последовательность привлечения различных форм знаний, которая позволяет осуществить основные принципы научных исследований: объективность, воспроизводимость, доказательность (верификацию) и точность полученных результатов.

Метод– определенная совокупность устойчивых правил, предназначенных для достижения какой-либо цели, способ познания объективной действительности.

Методика исследования– совокупность способов и приемов исследования, порядок их применения и интерпретация полученных с их помощью результатов

Техника исследования– совокупность специальных приемов для использования того или иного метода.

Процедура исследования– последовательность действий, способ организации исследования[21].

1.5. Контрольные вопросы

1. Что понимается под термином «исследование»? Приведите примеры исследований в информатике.
2. По каким критериям классифицируют научные исследования?
3. Какую роль в научном исследовании играет выбор подходов к исследованию? Какие подходы к исследованию вам известны?

4. Дайте определение термину «гипотеза». Охарактеризуйте значение гипотезы в научном исследовании.

5. Дайте определение терминов "метод" и "методология".

6. Какова методология научного исследования.

7. Перечислите общенаучные методы научных исследований.

8. Какие методы причисляют к методам теоретического уровня?

9. Какие методы причисляют к методам эмпирического уровня?

10. Какие методы называются частными?

11. Какие методы называют специальными?

1.6. Задания для самостоятельной работы

1. Автор в процессе написания выпускной квалификационной работы сформулировал тему и цели исследования следующим образом:

Тема исследования: «Разработка и реализация моделей оптимизации производства сельскохозяйственной продукции».

Цель исследования — обосновать с научно-методической точки зрения и разработать модели оптимизации производства сельскохозяйственной продукции в крестьянских фермерских хозяйствах.

Сформулируйте задачи исследования, опираясь на имеющуюся информацию и определения терминов «Тема исследования», «Цель исследования».

2. Автор в процессе работы над выпускной квалификационной работой сформулировал тему исследования следующим образом: «Проектирование информационной системы документооборота центра довузовской подготовки».

Является ли заявленная тема исследования актуальной? Обоснуйте ответ.

3. Определите объект и предмет для следующих исследований:

1) Оптимизация взаимодействия товаропроизводителей различных категорий в агропромышленных объединениях.

2) Оптимизация затрат на производство кормов в сельскохозяйственном предприятии.

3) Оценка эффективности использования складских помещений для хранения овощной продукции.

2. ИНСТРУМЕНТАЛЬНЫЕ СРЕДСТВА НАУЧНЫХ ИССЛЕДОВАНИЙ

2.1. Архитектура информационно-аналитических систем

Инструментами современных научных исследований являются информационные технологии и системы, обеспечивающие обработку и анализ значительных объемов информации. Это требует четкого понимания архитектуры информационных систем, структуры данных, методов и алгоритмов их исследования, а также тенденций развития аналитических решений и инструментов. Знание основ функционирования аналитических информационных систем и методов анализа позволяет ориентироваться в широком спектре прикладных решений в области автоматизации научных исследований и грамотно и эффективно проводить научные исследования.

На рисунке 2.1 представлена общая структура информационно-аналитических систем.

Подсистема хранения данных.

Многомерное хранилище данных может быть организовано в виде одной из следующих структур:

1. физической структуры, называемой MOLAP, в которую с определенной периодичностью загружаются данные из файлов – источников, принадлежащих базам оперативных данных

2. виртуальной структуры, называемой ROLAP, которая динамически используется при запросах. ROLAP – система рассматривается просто как надстройка над реляционными базами данных, обеспечивающая удобный интерфейс пользователя. Типичными инструментальными средствами, поддерживающими ROLAP, является Business Objects.

3. гибридной структуры, называемой HOLAP, которая используется при построении многоуровневых информационных хранилищ, применяемых на разных уровнях управления больших корпорац

Анализ параметров использования ROLAP и MOLAP информационных хранилищ показывает, что внедрение и эксплуатация ROLAP - систем является более простым и дешевым по сравнению с MOLAP – системами, но уступают последним в эффективности оперативного анализа данных.

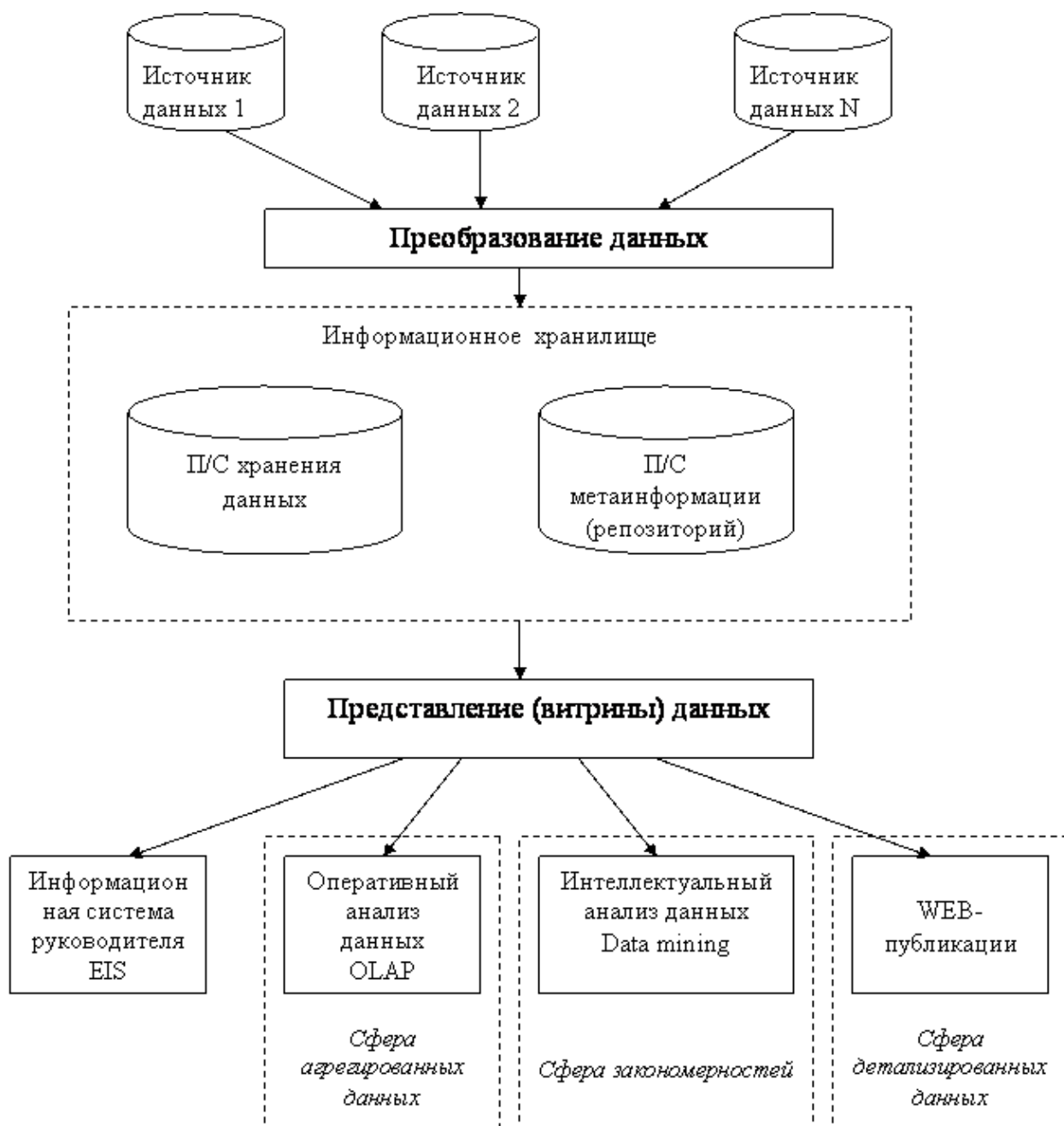


Рисунок 2.1– Структура информационно- аналитических систем

Подсистема метainформации.

Репозиторий представляет собой описание структуры информационного хранилища: состава показателей, иерархии агрегаций измерений, форматов данных, используемых функций, физического размещения на сервере, прав доступа пользователей, частоты обновления.

В репозитории задается схема отображения структуры файлов-источников данных на структуре ИХ, а также схема отображения структуры ИХ на витринах данных. Через репозиторий осуществляется интерпретация запросов к ИХ на проведение оперативного анализа данных.

Подсистема преобразования данных (загрузки хранилища).

Подсистема загрузки ИХ создается только для MOLAP – систем. Для ROLAP – систем в процессе выполнения запросов осуществляется преобразование данных из файлов – источников. В том и другом случаях требуется выполнение следующих основных функций:

- сбор данных,
- очистка данных,
- агрегирование данных.

Сбор данных предполагает передачу данных из источников в ИХ в соответствии со схемой отображения, представленной в репозитории.

В процессе очистки данных осуществляется проверка целостности, исключение дублирования данных, отбраковка случайных данных, восстановление отсутствующих данных, приведение данных к единому формату.

В случае необходимости агрегирования данных осуществляется суммирование итогов по заданным в репозитории признакам.

Подсистема представления данных (организация витрин данных).

Под витриной данных понимается предметно-ориентированное хранилище данных, как правило, агрегированной информации, предназначенное для использования группой пользователей в рамках конкретного вида деятельности предприятия, например маркетинга и т.д.

Как правило, витрины данных являются подмножествами общего хранилища данных компании, которое служит для них источником. Обычно общее ИХ и витрины данных разрабатываются параллельно.

Подсистема оперативного анализа данных.

Подсистема оперативного анализа данных, как правило, используется лицами, подготавливающими информацию для принятия решений, путем выполнения различных статистических группировок исходных данных.

В рамках пользовательского интерфейса для оперативного анализа данных используются следующие базовые операции:

- **Поворот.** Добавление нового признака анализа.
- **Проекция.** Выборка подмножества по задаваемой совокупности измерений. При этом значения, лежащие на оси проекции, суммируются.
- **Раскрытие.** Осуществляется декомпозиция признака агрегации на

компоненты, например, признак года разбивается на кварталы. При этом автоматически детализуются числовые показатели.

- **Свертка.** Операция обратная раскрытию. При этом значения детальных показателей суммируются в агрегируемый показатель.

- **Сечение или срез.** Выделение подмножества данных по конкретным значениям одного или нескольких измерений.

Подсистема интеллектуального анализа данных.

Подсистема интеллектуального анализа данных используется специальной категорией пользователей – аналитиков, которые на основе ИХ обнаруживают закономерности в деятельности предприятия и на рынке, используемые в дальнейшем для обоснования стратегических и тактических решений. Интеллектуальный анализ требует более сложных методов анализа по сравнению со статическими группировками и выполняется путем проведения множества сеансов.

Типичными задачами интеллектуального анализа данных являются:

- Установление корреляций, причинно-следственных связей и временных связей событий, например определение местоположения прибыльных предприятий.

- Классификация ситуаций, позволяющая обобщать конкретные события в классы, например определение типичного профиля покупателя конкретных видов продукции.

- Прогнозирование развития ситуаций, например прогнозирование цен, объемов продаж, производства.

К основным методам интеллектуального анализа данных относятся:

- Методы многомерного статистического анализа,
- Индуктивные методы построения деревьев решений,
- Нейронные сети.

Подсистема «Информационная система руководителя».

Информационная система руководителя предназначена для лиц, непосредственно принимающих решения. Поэтому интерфейс таких систем должен быть в наибольшей степени упрощенным. Обычно в качестве интерфейса руководителям предприятия предлагается набор стандартных отчетов и графиков, настраиваемых на потребности руководителя через систему меню. Часто в качестве интерфейса предлагаются диаграммы Ишикава, представляющие собой саморазворачивающееся дерево показателей, в котором листья ветвей раскрашиваются в разные цвета,

символизирующие характер состояния показателя (нормальный, тревожный, кризисный). Лист любой ветви дерева может быть развернут а таблицу значений показателя или график.

Подсистема WEB – публикации.

Подсистема WEB – публикации предполагает преобразование полученной из ИХ информации в HTML – вид, доступный для ее просмотра удаленными клиентами с помощью браузеров Интернета.

В архитектуре ИАС применяется несколько различных подходов к организации хранилищ данных. Вариант с физическим хранилищем предполагает загрузку данных из OLTP-систем в ХД, где данные приводятся к единому формату, фильтруются, агрегируются, а запросы адресуются непосредственно к ХД.

Такая архитектура предъявляет повышенные требования к безопасности данных, как с точки зрения доступа к ним, так и с точки зрения эффективного хранения больших объемов данных.

При создании виртуального хранилища, данные не копируются в одно хранилище, а извлекаются и интегрируются непосредственно при формировании аналитического запроса. Такой подход, несмотря на возможность работы с реальными детализированными данными, имеет ряд недостатков. К ним относятся увеличение времени обработки запроса, обязательная доступность всех источников информации и возможность обработки только тех исторических данных, которые содержатся в системе OLTP в течение заданного периода времени.

Вариант информационно-аналитической системы с созданием независимых витрин данных отличается простотой их организации, поскольку каждая витрина оперирует с данными одной предметной области (задачи). Недостатками автономных витрин данных является необходимость хранить одинаковые данные в разных витринах данных. Кроме того, в ряде случаев возникает ситуация, при которой пользователям необходимы данные из нескольких витрин одновременно, а отсутствие консолидации данных на уровне независимых витрин не позволяет этого сделать.

В последнее время наибольшее распространение получили информационно-аналитические системы с хранилищем и витринами данных.

На основании детальных данных формируются агрегированные данные, представляющие собой суммы фактических данных по измерениям. Ключевым элементом в ХД являются метаданные - данные о данных. Они содержат всю информацию, необходимую для извлечения, преобразования и загрузки данных из различных источников, а также для последующего использования и интерпретации данных, содержащихся в ХД.

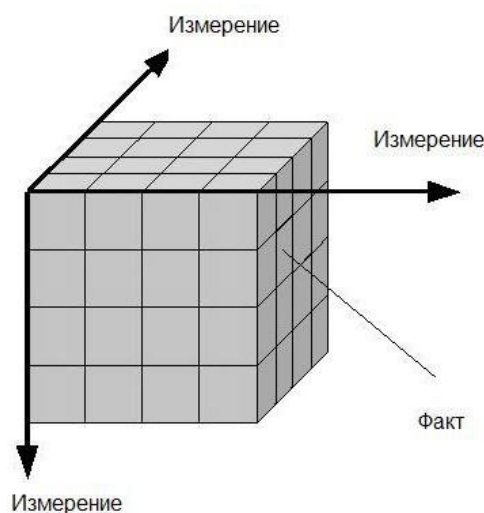


Рисунок 2.2– Представление данных в виде гиперкуба

Одним из критериев эффективности информационно-аналитических систем является скорость выполнения сложных запросов и прозрачность структуры хранения информации. Для обеспечения того, чтобы эти критерии были достижимы, в настоящее время используются два подхода к созданию хранилищ данных. Первый подход основан на многомерной модели данных, в то время как второй подход использует реляционную модель. В обоих случаях данные организованы в виде гиперкуба, который может быть реализован как отдельная многомерная структура, так и в рамках реляционных таблиц. Оси куба-это измерения, которые используются для построения параметров, связанных с анализируемой предметной областью, таких как названия продуктов и названия месяцев года. На пересечении измерительных осей находятся данные, количественно характеризующие анализируемые факты, например, объемы продаж, выраженные в единицах продукции (рис. 2.2).

Дата отгрузки	Наименование товара	Поставщик	Продавец	Цена продажи	Количество
				Сумма	Сумма
06.02.2007	L-Тироксин 100 Берлин-Хеми таб. №100	Поставщик №1	Продавец №1	100,1	3
06.02.2007	Аевит капс. 0.2 x10	Поставщик №1	Продавец №1	12	50
06.02.2007	Азитромицин-АКОС капс. 250мг №6	Поставщик №1	Продавец №1	115,6	4
06.02.2007	Аминазин амп. 2.5% 2мл №10	Поставщик №1	Продавец №1	27,5	5
06.02.2007	Аминостерил КЕ 10% фл. 500мл	Поставщик №1	Продавец №1	359,2	10
06.02.2007	Амиодарон (кордарон) таб. 200мг x30	Поставщик №1	Продавец №1	53	20
06.02.2007	Амлодипин 0.005г №30	Поставщик №1	Продавец №1	91,4	10
06.02.2007	Амлодипин 10мг №30	Поставщик №1	Продавец №1	150	10
06.02.2007	Аммиака р-р 10% фл. 40мл	Поставщик №1	Продавец №4	4,8	74
06.02.2007	Амоксиклав 1200 мг пор. д/в/в x5	Поставщик №1	Продавец №1	813,8	10
06.02.2007	Анальгин 50% амп. 2мл x10	Поставщик №1	Продавец №1	39,2	50
06.02.2007	Ардуан амп. 4мг 2мл x25	Поставщик №1	Продавец №1	1256,8	1
06.02.2007	Аскорутин таб. 0.33 x10	Поставщик №1	Продавец №1	0,8	10
06.02.2007	Аскофен П таб. x10	Поставщик №1	Продавец №1	4,3	50
06.02.2007	Бактисубтил капс. x20	Поставщик №1	Продавец №1	137,7	1
06.02.2007	Бензилпеницилина натр. соль 1млн.Ед.	Поставщик №1	Продавец №1	2,09	15
06.02.2007	Бинг марлевый стер. 5x10	Поставщик №1	Продавец №1	2,4	50
06.02.2007	Бинг марлевый стер. 7x14	Поставщик №1	Продавец №1	4,4	50
06.02.2007	Борная кислота спирт 3% фл. 25мл	Поставщик №1	Продавец №1	3,73	11
06.02.2007	Валерианы экстракт таб. 0.02 x50	Поставщик №1	Продавец №1	5,9	50
06.02.2007	Вага хирургическая нест. упак. 100г	Поставщик №1	Продавец №1	6,8	20
06.02.2007	Верапамил г/х 0,25% 2мл №10	Поставщик №1	Продавец №1	13,5	1
06.02.2007	Верошипактон (спиронолактон) таб. 25мг x20	Поставщик №1	Продавец №1	30,1	50
06.02.2007	Винпоцетин таб.5мг x50	Поставщик №1	Продавец №1	26,2	31
06.02.2007	Вода для инъекций амп. 2мл x10	Поставщик №1	Продавец №1	11,6	5
06.02.2007	Галидор амп. 50мг/2мл x10	Поставщик №1	Продавец №1	266,7	15
06.02.2007	Гепарин р-р д/ин. 5000ЕД/мл фл. 5мл x5	Поставщик №1	Продавец №1	59,1	60
06.02.2007	Гептрал фл.+ раств. 400мг x5	Поставщик №1	Продавец №1	1198,9	1
06.02.2007	Гипотиазид таб. 25мг x20	Поставщик №1	Продавец №1	52,7	15
06.02.2007	Глиатилин р-р 1000мг/4 мл №3 амп.	Поставщик №1	Продавец №1	412,5	8
06.02.2007	Глиатилин р-р 1000мг/4 мл №3 амп.	Поставщик №2	Продавец №2	417,9	18
06.02.2007	Глицин таб. 0.1г №50	Поставщик №1	Продавец №1	9	10
06.02.2007	Диазолин др. 0.1 x10	Поставщик №1	Продавец №1	5,4	50
06.02.2007	Диксин (диоксидин) 1% р-р в/п и наружн. амп. 5мл	Поставщик №1	Продавец №1	18,4	5
06.02.2007	Димедрол 0,05 таб. №10	Поставщик №1	Продавец №1	2,8	100
06.02.2007	Димедрол амп. 1% 1мл x10	Поставщик №1	Продавец №1	67,2	55
06.02.2007	Димексид фл. 100мл	Поставщик №1	Продавец №1	16,4	20

Рисунок 2.3– Представление многомерных данных

В процессе поиска и извлечения из гиперкуба нужной информации над его измерениями производится такие операции как срез, вращение, консолидация и детализация. При выполнении операции срез формируется подмножество гиперкуба, в котором значение одного или более измерений фиксировано (например, значения параметров для фиксированного измерения «Город»). Операция вращения изменяет порядок представления измерений, обеспечивая представление куба в более удобной для восприятия форме. Консолидация – операция перехода от детального представления данных к агрегированному. Например, показатели для отдельных городов могут быть просуммированы с целью получения показателей для всего региона, а показатели для отдельных дней могут быть «свернуты» до показателей за целый год. Операция детализации – это действие над данными, обратное операции консолидации.

При использовании многомерной модели данные хранятся в виде упорядоченных многомерных массивов, а гиперкуб представляется в виде одной плоской таблицы (рисунок 2.3).

Применение многомерных моделей данных в ХД разумно при небольших объемах данных, стабильном наборе информационных

измерений, а также в случаях, когда время ответа системы на нерегламентированные запросы является наиболее критичным параметром или предполагается использование сложных встроенных функций для выполнения кроссмерных вычислений над ячейками гиперкуба, в том числе возможность написания пользовательских функций.

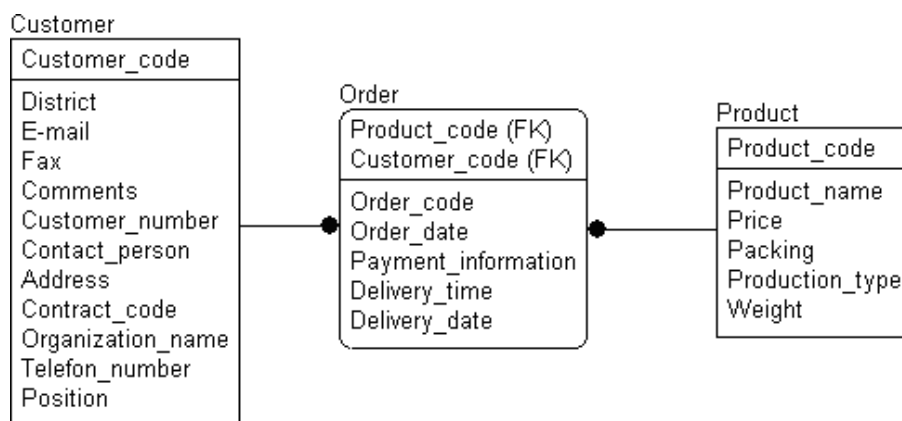


Рисунок 2.4– Модель данных в виде схемы типа «звезда»

При применении реляционной модели данных для построения хранилища данные организуются специальным образом. Чаще всего используется радиальная схема, получившая название «звезда» (рисунок 2.4). В этой схеме, представленной в качестве примера, используется два типа таблиц: таблица фактов (*Order*) и таблицы измерений (*Customer*, *Product*). Если в анализе участвуют связанные измерения, то целесообразно использовать схему «снежинка» (рисунок 2.5).

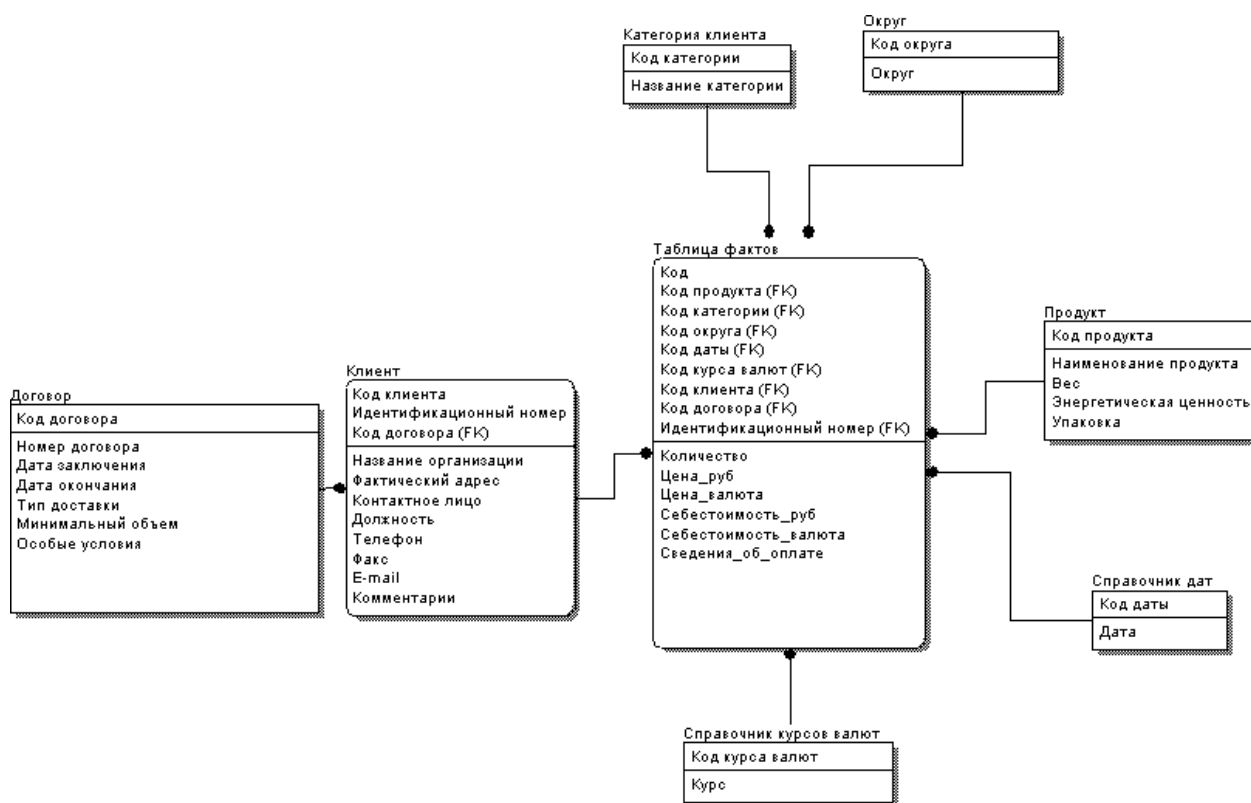


Рисунок 2.5– Модель данных в виде схемы типа «снежинка»

Таким образом, информационно-аналитические системы (ИАС) призваны на основе данных, получаемых в режиме реального времени, помогать в принятии управленческих решений. ИАС - это современный высокоэффективный инструмент поддержки принятия стратегических, тактических и оперативных управленческих решений на основе наглядного и оперативного предоставления всей необходимой совокупности данных пользователям, ответственным за анализ состояния дел и принятие управленческих решений. Комплекс информационно-аналитических систем затрагивает всю управленческую вертикаль: корпоративную отчетность, финансово-экономическое планирование и стратегическое планирование.

2.2. Применение информационных технологий и систем в научных исследованиях

Создание информационного хранилища аналитических данных, основанного на идее сбора разрозненной информации, хранящейся в организациях, в единый массив в форме, пригодной для анализа, имеет большое значение для реализации аналитических процедур в научных

исследованиях и решении практических задач. Если еще в 80 - е годы прошлого века разработка компьютерных программ была неотъемлемой частью самого исследования, когда ученый ставил перед собой задачу, выполнял ее и был в большинстве случаев единственным пользователем созданного приложения, то сейчас существует широкий спектр профессиональных решений, обеспечивающих реализацию научных исследований. Особое место в широком спектре аналитических систем занимают инструменты бизнес-анализа. Это обусловлено, с одной стороны, острой потребностью в исследованиях под влиянием возросшей конкуренции, внешних и внутренних изменений в бизнесе, а также высокими требованиями к эффективности.

Современные технологии, используемые в системах бизнес - анализа, ориентированы на исследование процессов, обеспечивающих принятие обоснованных решений при управлении экономическими объектами. Предварительный анализ обычно выполняется с помощью OLAP-технологий-технологий оперативной аналитической обработки данных, которые просты в использовании, оснащены средствами визуализации данных и возможностью выполнения необходимых расчетов. Для проведения глубоких исследований необходимо использовать более гибкие методы интеллектуального анализа данных. В этом случае предпочтение отдается методам *Data Mining*.

Data Mining – процесс выявления скрытых закономерностей и взаимозависимостей в больших наборах данных для поддержки принятия решений⁸. Основная особенность *Data Mining* – это сочетание широкого математического инструментария (от классического статистического анализа до новых кибернетических методов) и последних достижений в сфере информационных технологий. В технологии *Data Mining* гармонично объединились строго формализованные методы и методы неформального анализа, т.е. количественный и качественный анализ данных.

К методам и алгоритмам *Data Mining* относятся следующие: искусственные нейронные сети, деревья решений, символьные правила, методы ближайшего соседа и k-ближайшего соседа, метод опорных векторов, байесовские сети, линейная регрессия, корреляционно-регрессионный анализ; иерархические методы кластерного анализа, неиерархические методы кластерного анализа, в том числе алгоритмы k-средних и k-медианы; методы поиска ассоциативных правил, в том числе алгоритм Apriori; метод ограниченного перебора, эволюционное программирование и генетические алгоритмы, разнообразные методы визуализации данных и множество других

методов.

Большинство аналитических методов, используемые в технологии Data Mining – это известные математические алгоритмы и методы. Новым является возможность их использования не только в теоретических исследованиях, но и для решения практических задач, что обусловлено возможностями современных информационных систем и технологий.

Выбор инструментальных средств анализа экономической информации определяется видом научного исследования, его целями и задачами, объектом и предметом исследования. Рассмотрим отдельные, наиболее распространенные решения.

В настоящее время корпоративные информационные системы представляют собой сложную аппаратно-программную среду, обеспечивающую автоматизацию процессов организации на операционном, тактическом и стратегическом уровнях (таблица 2.1). На каждом из уровней не только накапливается первичная информация, но осуществляется ее обработка и анализ с целью эффективного управления организацией, выявления особенностей исследуемых процессов и явлений.

Таблица 2.1– Уровни корпоративной информационной системы

Уровень управления	Информационные системы
Уровень 1: Системы обработки транзакций (OLTP-системы)	Системы бухгалтерского учета, управления производством, управления закупками, управления продажами, управления логистикой и др.
Уровень 2: Системы поддержки принятия решений на тактическом уровне	Системы финансового анализа, анализа продаж, закупок, складских запасов, системы управления взаимоотношениями с клиентами, планирование на тактическом уровне управления и др.
Уровень 3: Системы поддержки принятия решений на стратегическом уровне	Системы управления эффективностью бизнеса, системы финансового планирования и бюджетирования, долгосрочного планирования, системы бизнес-анализа (Data Mining, Text Mining и др.)

В зависимости от уровня зрелости компании, ее стратегических целей и задач, возможно использование различных решений, ориентированных на углубленные аналитические исследования. Каждый уровень обеспечивает объем первичных данных, использование которых в научных исследованиях обеспечивает решение практических задач.

Первый уровень автоматизации аналитических задач предусматривает выявление текущей ситуации в компании в прошлые периоды времени путем

создания типовой и регламентированной отчетности. Примером такого подхода являются отчеты о движении денежных средств, остатках товаров и продукции на складах, производственный отчет, включающий номенклатуру и объем произведенной продукции, информацию о затратах запасов на ее производство и т.д., формируемые в транзакционных системах OLTP.

Информационные системы транзакционного уровня имеют достаточно широкие аналитические возможности. Так, системы управления ресурсами предприятия класса ERP (*Enterprise Resource Planning*), содержит следующий функционал:

- *Планирование продаж и производства.* Результатом действия данного функционала является разработка плана производства продукции.

- *Управление спросом.* Основное назначение состоит в формировании прогнозов будущего спроса на продукцию, определении объема заказов клиентов, дистрибьюторов, спроса на сырье и материалы в рамках предприятия и др.

- *Укрупненное планирование мощностей.* Реализует оценку планов производства и степени их выполнимости.

- *Основной план производства.* Формирование плана-графика выпуска продукции со сроками изготовления и количеством.

- *Планирование потребностей в материалах.* Определяются виды материальных ресурсов (сборных узлов, готовых агрегатов, покупных изделий, исходного сырья, полуфабрикатов и др.) и конкретные сроки их поставки для выполнения плана.

- *Спецификация изделий.* Определяет состав конечного изделия, материальные ресурсы, необходимые для его изготовления, и др. Фактически спецификация является связующим звеном между основным планом производства и планом потребностей в материалах.

- *Планирование потребностей в мощностях.* На данном этапе планирования более детально, чем на предыдущих уровнях, определяются производственные мощности.

- *Маршрутизация/рабочие центры.* Формирование маршрутов, в соответствии с которыми выпускаются изделия, с учетом производственных мощностей различного уровня.

- *Управление закупками, запасами, продажами.*

- *Управление финансами* (ведение Главной книги, расчеты с дебиторами и кредиторами, учет основных средств, управление наличными средствами, планирование финансовой деятельности и др.).

- *Управление затратами* (учет всех затрат предприятия и калькуляция себестоимости готовой продукции или услуг).

- Управление проектами/программами.

В соответствии с современными требованиями APICS ERP- система включает следующие модули:

- управления логистическими цепочками – SCM (Supply Chain Management);

- усовершенствованного планирования и составления производственных графиков – APS (Advanced Planning and Scheduling);

- управления взаимоотношениями с клиентами – CRM (Customer Relationship Management);

- электронной коммерции – EC (Electronic Commerce);

- управления данными об изделии – PDM (Product Data Management);

- надстройку Business Intelligence, включающую решения на основе технологии OLAP и DSS (Decision Support Systems);

- автономный модуль, отвечающий за конфигурирование системы – SCE (Stand Alone Configuration Engine);

- окончательного (детализированного) планирования ресурсов – FRP (Finite Resource Planning).

Следующий уровень ориентирован на выявление причин возникновения проблемных ситуаций, для чего в большинстве случаев можно использовать нерегламентированные отчеты, формируемые средствами транзакционных систем или с помощью OLAP- технологий. Оценка текущего состояния и мониторинг производится также средствами OLAP-технологий, аналитические возможности которых позволяют выявить отклонение текущей ситуации плановой в разрезе принятых в компании критериев. OLAP-технологии могут быть внедрены в функционал транзакционной системы или реализовываться с помощью отдельного приложения.

2.3. Статистические пакеты в научных исследованиях

Для принятия решений в условиях неопределенности целесообразно использовать такие методы, как математическая статистика и анализ данных. С середины 80-х годов системы статистического анализа получили широкое применение в научных исследованиях, а в последующие периоды и по настоящее время их пользователями стали коммерческие структуры, медицинские организации, государственные учреждения.

Известные на российском рынке статистические пакеты можно

подразделить на профессиональные и популярные, универсальные и специальные. Профессиональные пакеты (*SAS, BMDP, IMSL*) включают значительное число специальных методов анализа, большинство из которых доступны только математикам-профессионалам. Популярные универсальные пакеты (*StatGraphics, SPSS, SyStat, CSS, Statistica, STADIA*) ориентированы на широкую аудиторию и активно применяются в практике. Специализированные пакеты ориентированы на избранные области анализа данных (Эвриста, Мезозавр, Класс-Мастер, Сани, Сигамнд) и содержат достаточно насыщенный инструментарий статистического анализа. Большой популярностью пользуются статистические средства, включенные табличные процессоры и системы управления базами данных. Наиболее часто в таких пакетах используются средства описательной статистики, регрессионного анализа, анализа временных рядов и прогнозирования.

Устойчивой тенденцией практически во всех новых версиях известных статистических пакетов является включение не только функций, реализующих традиционные методы статистического анализа, но и элементов *Data Mining*. Однако они по – прежнему ориентируются на классические методы – корреляционный, регрессионный, факторный анализ и другие. Недостатками статистических систем, ограничивающими их массовое использование специалистами в области экономики и менеджмента, являются высокие требования к специальной подготовке пользователей, чрезмерная функциональность для использования в финансах и бизнесе, а также высокая стоимость пакетов данного класса.

2.4. Контрольные вопросы

1. Что понимают под термином «инструментальные средства анализа данных»?
2. Дайте общую характеристику функциональным возможностям современных аналитических систем.
3. Проанализируйте и оцените достоинства и недостатки следующих вариантов архитектуры информационно-аналитических систем:
 - Функциональная ИАС.
 - На базе независимых витрин данных.
 - На базе двухуровневого хранилища данных.
 - На базе трехуровневого хранилища данных.
4. В чем состоит различие между виртуальным и физическим Хранилищами данных?

5. «Построение хранилища данных — задача только информационных технологий» - опровергните или подтвердите данный тезис.

6. Дайте описание модели данных, используемой большинством ХД. Приведите примеры схем данных.

2.5. Задания для самостоятельной работы

1. Представьте эссе об OLTP-системах и их применимости для анализа данных.

2. Представьте эссе об OLAP-системах и их применимости для анализа данных.

3. Представьте эссе о ХД. Их классификация по видам и типам. Примеры работающих ХД.

4. С целью повышения эффективности компании руководство приняло решение о проведении исследований, направленных на выявление проблемных ситуаций в головном и региональных офисах компании. Источником информации для решения поставленной задачи могут служить данные информационной системы, представленные в виде справочников «Клиенты», «Товары», «Заказы», «Представительства компании». Для того, чтобы анализ продаж можно было выполнять регулярно, следует воспроизвести в Excel фрагмент ИС, который обеспечит доступ к данным по продажам, а затем сформировать OLAP-отчеты. Составить отчет по проведенным исследованиям и сформулировать предложения, направленные на повышение эффективности компании.

Для выполнения задания реализуйте в Excel модель данных, представленную на рисунке 2.6. Данная модель соответствует работе всех представительств компании.



Рисунок 2.6– Модель данных

Введите исходные данные на лист *Клиенты* в соответствии с рисунком 2.7, а на лист *Товары* в соответствии с рисунком 2.8. Присвойте столбцам (диапазонам ячеек) имена в соответствии с таблицей 2.2. Лист *Заказы* заполните данными и формулами в соответствии с таблицей 2.2 и рисунком 2.9. Анализ хозяйственной деятельности выполните с помощью механизма *Сводные таблицы*, пример которой представлен на рисунке 2.10.

	A	B	C	D	E	F	G	H	I
1	Код клиента	Название фирмы	Контактная персона	Индекс	Город	Улица	Skype	Телефон	Скидка (%)
2	2101	Партия	Иванов Илья	252001	Москва	Донская, 5	TORG1	2345678	10,00%
3	2102	Финиш	Полищук Ирина	220211	Москва	Серпуховская, 14	TORG2	3456789	15,00%
4	2103	Пионер	Резник Павел	234567	Москва	Мытная, 12	TORG3	1234567	10,00%
5	2104	Комплект	Овдеева Ольга	253453	Липецк	Кожуховская, 1	TORG4	3214567	5,00%
6	2105	Москворечье	Ковалев Сергей	335356	Минск	Северянина, 2	TORG5	4216789	20,00%
7	2106	Зенит	Поликарпова Ольга	775634	Киев	Зеленая, 23	TORG6	4126789	0,00%
8	2107	Старт	Абрамов Михаил	233434	Киев	Центральная, 17	TORG7	4567898	0,00%
9	2108	Контакт	Ибрагимов Петр	546734	Липецк	Садовая, 21	TORG8	2346578	0,00%
10	2109	Космос	Мичурин Виктор	453423	Саратов	Мытная, 32	TORG9	9867856	10,00%
11	2110	Авангард	Смирнов Павел	231234	Чита	Строителей, 25	TORG10	5678789	5,00%
12	2111	Золотник	Мсамойлова Татьяна	453456	Коломна	Московская, 23	TORG11	4567879	25,00%

Рисунок 2.7 – Справочник Клиенты

	A	B	C
1	Номер	Наименование товара	Цена
2	101	Intel Celeron 336 (2,8 ГГц)	7 519,00р.
3	102	Intel Celeron 430 (1,80 ГГц)	8 090,00р.
4	103	AMD Sempron LE-1100 (1,9 ГГц)	8 600,00р.
5	104	Intel Celeron E1200 (2x1,60 ГГц)	9 500,00р.
6	201	Принтер Samsung ML-2571N	5 700,00р.
7	202	Принтер Canon LBP-2900B	4 430,00р.
8	203	Принтер HP LaserJet color	13 750,00р.
9	301	Монитор 16 TFT ASUS	4 223,00р.
10	302	Монитор 19 TFT ASUS AS VB191T	6 753,00р.
11	303	Монитор 16 TFT ASUS AS VB191T	6 500,00р.
12	304	Монитор 16 TFT ASUS AS VB191T	6 450,00р.

Рисунок 2.8 – Справочник Товары

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Месяц	Дата	Заказ	Номер товара	Наименование товара	Количество во	Цена за ед. товара	Код клиента	Название фирмы	Сумма заказа	Скидка	Уплачено	Город
2	Январь	14.1.08	10000	101	Intel Celeron 336 (2,8 ГГц)	2	7519	2101	Партия	15038	10%	13534,2	Москва
3	Январь	15.1.08	10001	102	Intel Celeron 430 (1,80 ГГц)	1	8090	2109	Космос	8090	10%	7281	Саратов
4	Январь	15.1.08	10002	103	AMD Sempron LE-1100 (1,9 ГГц)	20	8600	2107	Старт	172000	0%	172000	Киев
5	Январь	16.1.08	10003	201	Принтер Samsung ML-2571N	12	5700	2104	Комплект	68400	5%	64980	Липецк
6	Январь	18.1.08	10004	203	Принтер HP LaserJet color	12	13750	2105	Москворечье	165000	20%	132000	Минск
7	Январь	19.1.08	10005	104	Intel Celeron E1200 (2x1,60 ГГц)	4	9500	2105	Москворечье	38000	20%	30400	Минск
8	Январь	20.1.08	10006	301	Монитор 16 TFT ASUS	45	4223	2110	Авангард	190035	5%	180533,25	Чита
9	Февраль	1.2.08	10007	101	Intel Celeron 336 (2,8 ГГц)	21	7519	2108	Контакт	157899	0%	157899	Липецк
10	Февраль	2.2.08	10008	302	Монитор 19 TFT ASUS AS VB191T	23	6753	2103	Пионер	155319	10%	139787,1	Москва
11	Февраль	3.2.08	10009	203	Принтер HP LaserJet color	34	13750	2102	Финиш	467500	15%	397375	Москва
12	Февраль	15.2.08	10010	201	Принтер Samsung ML-2571N	12	5700	2109	Космос	68400	10%	61560	Саратов
13	Февраль	16.2.08	10011	303	Монитор 16 TFT ASUS AS VB191T	5	6500	2109	Космос	32500	10%	29250	Саратов
14	Февраль	21.2.08	10012	301	Монитор 16 TFT ASUS	26	4223	2106	Зенит	109798	0%	109798	Киев
15	Февраль	23.2.08	10013	201	Принтер Samsung ML-2571N	16	5700	2110	Авангард	91200	5%	86640	Чита
16	Февраль	24.2.08	10014	202	Принтер Canon LBP-2900B	45	4430	2111	Золотник	199350	25%	149512,5	Коломна
17	Февраль	27.2.08	10015	203	Принтер HP LaserJet color	10	13750	2101	Партия	137500	10%	123750	Москва
18	Февраль	28.2.08	10016	101	Intel Celeron 336 (2,8 ГГц)	11	7519	2101	Партия	82709	10%	74438,1	Москва
19	Март	1.3.08	10017	202	Принтер Canon LBP-2900B	3	4430	2102	Финиш	13290	15%	11296,5	Москва
20	Март	2.3.08	10018	302	Монитор 19 TFT ASUS AS VB191T	23	6753	2103	Пионер	155319	10%	139787,1	Москва
21	Март	4.3.08	10019	301	Монитор 16 TFT ASUS	24	4223	2106	Зенит	101352	0%	101352	Киев
22	Март	12.3.08	10020	304	Монитор 16 TFT ASUS AS VB191T	24	6450	2106	Зенит	154800	0%	154800	Киев
23	Март	13.3.08	10021	101	Intel Celeron 336 (2,8 ГГц)	24	7519	2106	Зенит	180456	0%	180456	Киев
24	Март	15.3.08	10022	102	Intel Celeron 430 (1,80 ГГц)	12	8090	2108	Контакт	97080	0%	97080	Липецк

Рисунок 2.9 – Справочник Заказы

Таблица 2.2 – Имена диапазонов ячеек

Имя столбца	Диапазон	Имя
Лист Клиент ы		
Код клиента	=Клиенты!\$A:\$A	Код_клиента
Название фирмы	=Клиенты!\$B:\$B	Название_фирмы
Город	=Клиенты!\$E:\$E	Город
Скидка (%)	=Клиенты!\$I:\$I	Скидка
Лист Товары		
Номер	=Товары!\$A:\$A	Номер
Наименование товара	=Товары!\$B:\$B	Наименование_товара
Цена	=Товары!\$C:\$C	Цена
Лист Заказы		
Дата	=Заказы!\$B:\$B	Дата
Заказ	=Заказы!\$C:\$C	Номер_заказа
Номер товара	=Заказы!\$D:\$D	Номер_товара2
Наименование товара	=Заказы!\$E:\$E	Наименование_товара2
Количество	=Заказы!\$F:\$F	Количество
Цена за ед. товара	=Заказы!\$G:\$G	Цена_за_ед_2
Код клиента	=Заказы!\$H:\$H	Код_клиента2
Название фирмы	=Заказы!\$I:\$I	Название_фирмы2
Сумма заказа	=Заказы!\$J:\$J	Сумма_заказа
Скидка	=Заказы!\$K:\$K	Скидка2
Уплачено	=Заказы!\$L:\$L	Уплачено
Город	=Заказы!\$M:\$M	Город2

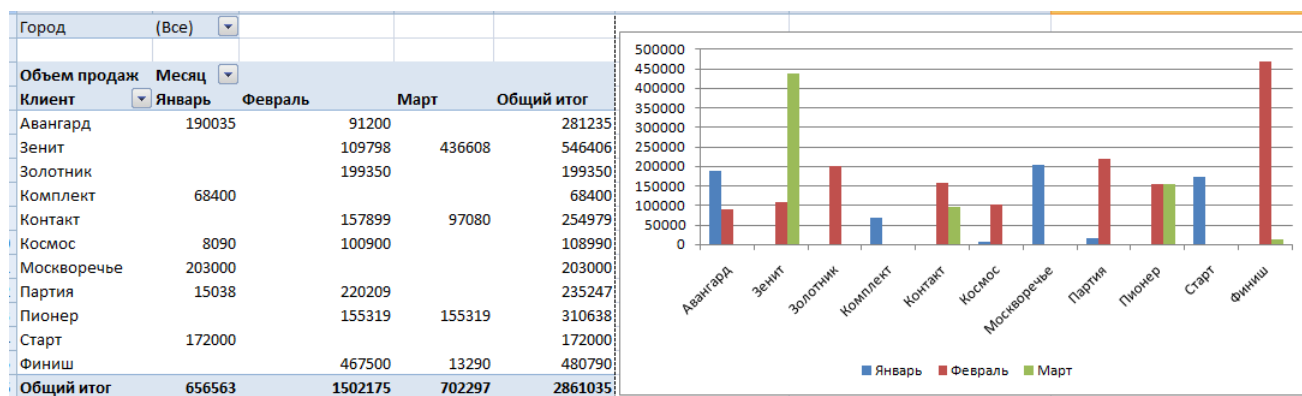


Рисунок 2.10– Пример применения сводных таблиц для анализа данных

3. ПРИНЦИПЫ ОРГАНИЗАЦИИ НАУЧНЫХ ИССЛЕДОВАНИЙ В ИНФОРМАТИКЕ

3.1. Логические методы в научных исследованиях

Группа логических методов состоит из методов, позволяющих находить решения проблем и получать научные результаты, основанные на логических выводах, правилах и опыте. Она включает в себя, прежде всего, такие методы познания, как классификация, обобщение, типология, анализ и синтез, сравнение и сопоставление, дедукция и индукция, метод формализации и математизации, метод логической интерпретации, метод аксиоматизации, метод аналогии, метод доказательства, метод логического обоснования и некоторые другие. В соответствии с принятой в настоящей работе классификацией эти методы носят общий характер и относятся как к эмпирическим, так и к теоретическим методам познания. Одной из характерных черт использования логических методов является субъективизм, но его влияние уменьшается с развитием исследовательских методов и делает эти методы чрезвычайно важными в процессе познания новых истин.

Научные логические методы исследования достаточно часто классифицируют согласно тому, какая система основания (качественная или количественная) используется при рассмотрении проблемы. Такой подход позволяет разделить методы на *количественные* и *качественные*.

Количественные методы исследования основываются на использовании измеряемых величин, что предопределяет выполнение операций (процедур) измерения. *Качественные методы* операции измерения практически не используют, а сосредоточены на словесном описании, интерпретации и объяснении свойств изучаемого экономического объекта, достигая основной цели – понимания сущности исследуемых свойств (качеств) объекта и их интерпретации.

Классификацией называют метод выделения классов из множества объектов, позволяющий отнести каждый из элементов множества к одному из выделенных классов. Метод классификации позволяет выявить разнообразие, свойства, связи и зависимости, общее и специфическое, проникнуть в сущность объектов исследования. С этой целью классификация выполняется по определенному признаку или свойству объекта, который называют классификационным признаком или основанием классификации.

Существует ряд принципов правильной и эффективной классификации,

без которых исследовательская работа не может быть успешной. Прежде всего, должен соблюдаться принцип единства критерия для выделения классов одного порядка, причем критерий не должен меняться в процессе классификации. Например, не следует уточнять, что в исследовании используются эмпирические и логические методы, поскольку первые отражают класс методов по критерию "уровень знаний", а вторые – по критерию "степень применения логики". Принцип пропорциональности-следующий принцип правильного применения метода классификации, означает, что каждый объект из множества других, подлежащих классификации, должен быть включен в один из образованных классов. Нарушение этого правила приводит к искажению представления об объекте исследования. Очевидно, что правильная классификация не позволяет объекту классификации принадлежать сразу к нескольким классам.

Одним из вариантов классификации является *типология*. Отличие типологии от классификации состоит в том, что типология допускает существование таких явлений, которые не соответствуют ни одному из выделенных критериев. Типология превосходит классификацию своей универсальностью и является первоначальной операцией любых систематизаций.

КЛАССИФИКАЦИЯ (от лат. *classis*— разряд, класс и *facio* — делаю, раскладываю).

ТИПОЛОГИЯ (греч. *tipos* - отпечаток, форма, образец и *logos* - слово, учение).

СИСТЕМАТИЗАЦИЯ – от глагола систематизировать, приводит в порядок.

3.2. Анализ и синтез в научных исследованиях

Анализ и *синтез* — эвристические методы, нашедшие широкое применение в научных исследованиях. *Анализ* (от греч. *analysis*— разложение) — метод научного исследования, состоящий в разложении единства на множество, целого — на его части, сложного — на его компоненты, события — на его отдельные ступени, содержания, понятия — на его признаки. *Синтез* — метод, противоположный анализу, означает процесс объединения в одно целое разьединенных ранее частей, компонент, признаков, ступеней в одно целое. Существует большое разнообразие определений, понимания и трактовок понятия «анализ». В математике под анализом понимают разработку приемов

вычислений и их применение к решению различных вопросов о величинах. Достаточно часто под термином «анализ» подразумевают научное исследование. Последняя трактовка скорее относится к обыденной, повседневной, когда анализ рассматривается как процесс изучения, исследования.

Анализ как разложение целого на части очень тесно связан с таким понятием как *абстрагирование* — мысленное отвлечение от второстепенных свойств и качеств объекта и предмета исследования с целью выделить их основные, наиболее важные свойства.

В информатике и управлении анализ применяется с целью выявления сущности, закономерностей, тенденций экономических и социальных процессов, хозяйственной деятельности на всех уровнях (в стране, отрасли, регионе, на предприятии, в частном бизнесе, семье) и служит исходной отправной точкой прогнозирования, планирования, управления экономическими объектами и протекающими в них процессами.

Синтез как метод исследований соединяет в целое раздробленные ранее фрагменты исследовательского материала, ориентирован на формирование новых понятий, принципов и концепций и осуществляет переход к новому экономическому знанию посредством синтетических суждений. Применение этого метода возможно как в рамках одной науки, например, экономической, или на базе междисциплинарного подхода, когда осуществляется соединение в единое целое теорий, методологий и методов различных дисциплин с целью формирования нового знания в рамках одной науки. Междисциплинарный синтез в настоящее время является одним из самых эффективных способов получения нового знания, в том числе и в области экономики и менеджмента.

3.3. Метод аналогии в научном исследовании

Среди различных логических методов, наиболее часто используемых в практической и научной деятельности человека, важное место принадлежит аналогии. *Аналогия* (греч. *analogia* - сходство, соответствие) — метод, согласно которому знание, полученное из рассмотрения какого-либо объекта, переносится на менее изученный, сходный по существенным свойствам и качествам объект. Данный метод является одним из источников научных гипотез, и играет значительную роль на начальных этапах научных исследований. Однако выводы, полученные этим методом, не всегда являются доказательными, поскольку выбранные для сравнения признаки могут оказаться случайными, несущественными или недостаточными для получения

научных выводов.

Метод аналогии нашел свое применение в естественных, технических и точных науках, а также в гуманитарных науках. В экономическом анализе аналогия может быть использована, например, при изучении определенного исторического периода в развитии общества. Учитывая характер развития, смешанное экономическое развитие, культурные особенности и другие характеристики, можно сравнить развитие по сходным характеристикам с развитием в другой стране, последней из таких периодов в ее истории. В данном случае метод аналогии позволяет учесть положительные и отрицательные процессы в развитии общества и избежать существенных ошибок.

В теории управления задачи построения организационной структуры, целевой модели или модели системы функций также решаются с помощью метода аналогии.

3.4. Сравнение в научных исследованиях

Сравнение – это научный метод познания, в процессе которого неизвестное (изучаемое) явление, объект или его характеристики сопоставляются с уже известными, изученными ранее с целью определения общих черт либо различий между ними. Такой подход позволяет выявить специфические особенности исследуемого объекта (явления), определить происходящие изменения и выявить тенденции его развития.

В методах сравнения выделяют:

– горизонтальный сравнительный анализ, который применяется для определения абсолютных и относительных отклонений фактического уровня исследуемых показателей от базового;

– вертикальный сравнительный анализ, используемый для изучения структуры экономических явлений; трендовый анализ, применяемый при изучении относительных темпов роста и прироста показателей за ряд лет к уровню базисного года, т.е. при исследовании рядов динамики.

Чаще всего сравнительный анализ используется для решения таких задач, как оценка достижения плановых показателей при осуществлении производственно-хозяйственной деятельности, сравнительный анализ текущих показателей с показателями предыдущих периодов, сравнение различных управленческих решений с целью выбора оптимального, сравнение динамических рядов с целью установления взаимосвязи исследуемых

показателей и др.

Основными условиями целесообразности проведения сравнительного анализа являются сопоставимость сравниваемых показателей. При этом сравниваемые измеряемые или расчетные показатели должны соответствовать принципам единства объемных, стоимостных, качественных и структурных показателей, относиться к одному и тому же периоду времени и быть сопоставимыми по условиям производства и методам измерения/расчета.

3.5. Метод индукции и дедукции

Методы индукции и дедукции – методы научных исследований, которые применяются практически в каждом исследовании. Под термином «*индукция*» понимается умозаключение, в котором осуществляется переход от частных факторов и положений к общим выводам. Другими словами, индукция позволяет сформулировать общий вывод на основе частного, перейти от единичного к общему. Метод *дедукции* предполагает формирование умозаключений, в которых осуществляется переход от общих фактов и положений к частным, т.е. опираясь на общее знание, исследователь делает вывод частного характера.

Методы дедукции и индукции играют значительную роль в научных исследованиях. Метод индукции широко используется на этапах исследования, когда необходимо обобщить некоторые эмпирические данные, классифицировать их и типологизировать. Его эффективность возрастает при взаимодействии с дедуктивным методом исследования, так как знания, полученные в результате индуктивного рассуждения, расширяются с использованием дедуктивного метода.

3.6. Подходы к организации исследовательской работы

Чаще всего успех исследовательского проекта зависит от умения правильно организовать, спланировать и выполнить этапы исследования в определенной последовательности. Наличие отдельных этапов, а также их содержание в определенной мере зависят от специфики объекта и предмета исследования, а также от вида исследования.

На рисунке 3.1 представлен наиболее типичный цикл исследования проблемы.

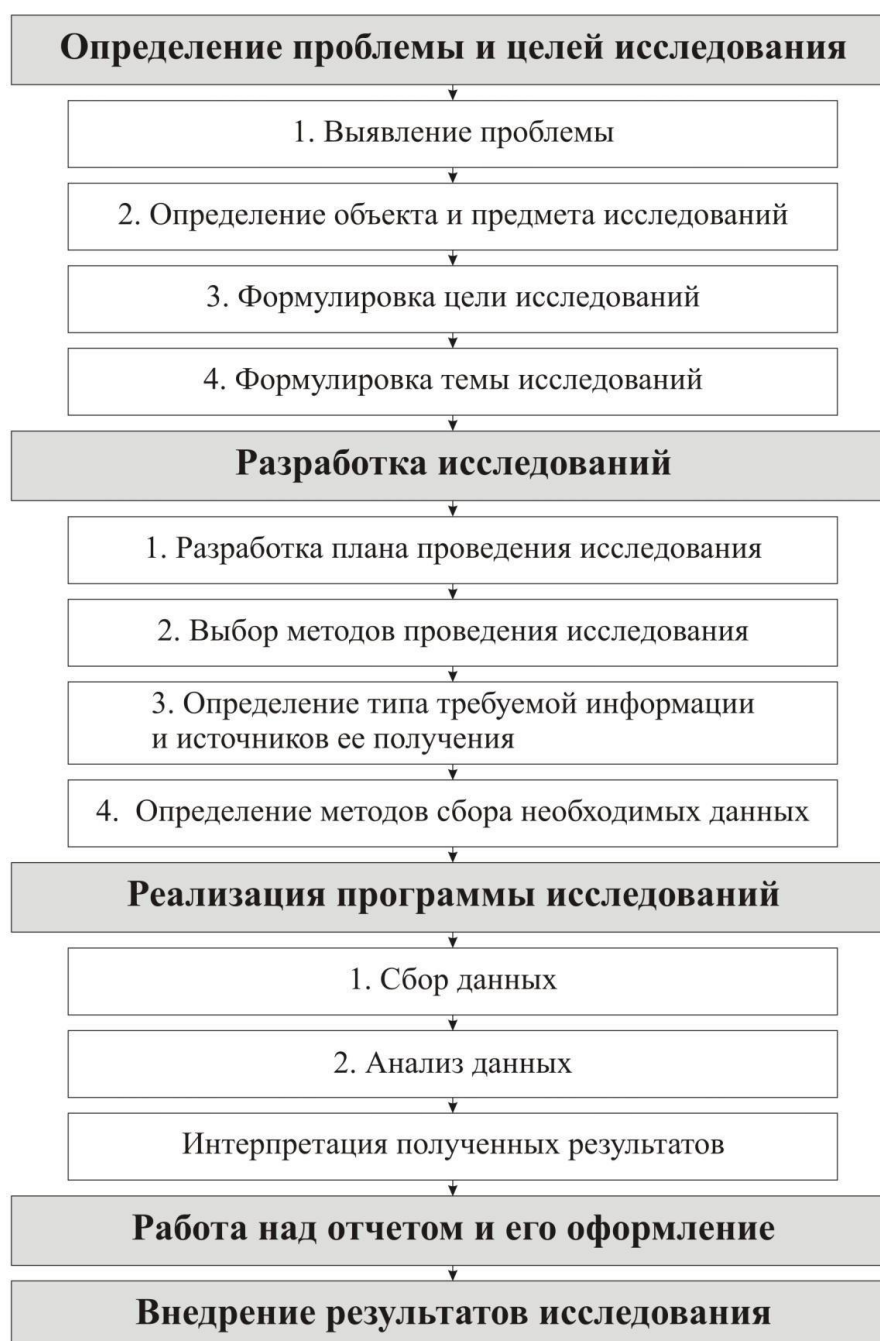


Рисунок 3.1 – Структура проведения научного исследования

На первом подготовительном этапе основное внимание уделяется выявлению проблемы исследования. Исследователю предстоит определить, в чем заключается проблема, решение которой будет способствовать развитию теории и совершенствованию практики в экономической деятельности и управлении. При этом понятие исследования формируется как целостная, логически связанная система взглядов, объединенных общей идеей и направленных на достижение целей исследования.

При формировании концепции используются такие методы как

дивергенция, трансформация, конвергенция. Термин «дивергенция» (*divergence*— расхождение) используется как метод расширения границ в познавательном процессе, направленный на всеобъемлющий поиск идей, подходов, направлений, методов, обеспечивающий достижение целей исследований. К методам дивергенции можно отнести обобщение научной информации, методы визуализации проблемы, обсуждение проблемы, анализ формулировок и формирование понятийного аппарата, накопление информации, инвентаризация точек зрения и подходов к решению проблемы.

В качестве примера использования дивергентного метода при исследовании проблемы повышения инвестиционной привлекательности региона рассматриваются различные аспекты деятельности региона:

- географическое положение;*
- демография;*
- уровень промышленного развития;*
- оценка ресурсов региона;*
- сравнительный анализ деятельности региона с целью выявления направлений инвестиционной привлекательности*
- оценка стратегического потенциала региона;*
- уровень развития инфраструктурных составляющих и т.д.*

Проведенный анализ позволяет выявить наиболее привлекательные направления деятельности в регионе с точки зрения привлечения инвестиций.

Трансформация предполагает изменение восприятия проблемы на тот тип, который наиболее приемлем для исследования. Методами преобразования являются классификация, уточнение структуры проблемы, выбор критериев классификации литературных источников или отдельных элементов проблемы; ранжирование научных взглядов или литературных источников.

Конвергенция (схождение) предопределяет сужение границ исследования, основанное на статистических исследованиях, расчетах, научных обоснованиях, проектировании.

На концепцию исследования оказывает влияние парадигма, преобладающая в данное время в исследуемых научных областях. Научная парадигма (от греч. Paradeigma - пример, образец) представляет собой совокупность теоретических и методологических положений, принятых научным сообществом на известном этапе развития науки, используемых в качестве образца, модели научных исследований, интерпретации, оценки и систематизации научных данных для осмысления гипотез и решения задач в

процессе научного познания. Именно основная научная парадигма определяет направления исследований большинства ученых до тех пор, пока накопленные знания не вступают в противоречие с общепринятыми и не будут сформированы знания, объясняющие возникшие противоречия.

В процессе определения проблемы автор вырабатывает общую картину исследования и в большинстве случаев удается сформировать рабочую гипотезу, которая в ходе исследования должна быть доказана или опровергнута. Кроме того, на этом этапе исследователь начинает идентифицировать объект как носителя проблемной ситуации, а субъект - свойства или характеристики изучаемого предмета исследования.

Исследовательская деятельность направлена на получение и применение нового знания, следовательно, в качестве цели исследования должны выступать достоверное изучение предмета исследования, а также получение практически значимых результатов.

Рассмотрим практический пример формулировки основных частей исследования.

Выявление проблемы

Постоянно меняющиеся условия ведения бизнеса, стремление получить конкурентные преимущества в динамичных условиях рынка заставляют организацию постоянно перестраивать свою деятельность. Появление новых направлений деятельности, вывод на рынок новых продуктов и услуг, установка оборудования с другой технологией, изменения в схемах поставок — все это требует своевременного проведения изменений процессов деятельности организации.

Перечисленные факторы образуют неустойчивую и динамически изменяющуюся среду, подтверждающую, что организации находятся в состоянии постоянного развития. По этой причине процесс изменения непрерывен и является одним из важнейших объектов управления в ходе реализации целей компании. Основное требование к управлению изменениями — обеспечение организационного и технологического соответствия деятельности организации динамичным условиям функционирования внутренней и внешней среды.

Определение научной парадигмы

Развитие теории изменений организации в 60-70-х годах 20-го века способствовало становлению такого нового направления менеджмента как управление изменениями (Change management). В нем получили развитие следующие модели: модель изменений К. Левина; модель управления

изменениями Л. Грейнера; теория изменений, базирующаяся на увеличении прибыли (теория «Е») и теория изменений, ориентированная на развитие организационных способностей (теория «О») М. Бира и Н. Нориа; модель преобразования бизнеса Ф. Гуияра и Дж. Келли; модель «От хорошего к великому» Дж. Коллинза; модель «Кривой перемен» Дж. Дак.

Гипотеза

Анализ существующих моделей в области управления изменениями показывает недостаточность исследования механизмов адаптации процессов организации к условиям функционирования внутренней и внешней среды, нацеленных на ликвидацию разрыва во времени между моментом фиксации принятого решения о необходимости изменений и моментом запуска и актуализации процессов и регламентов организации. Также в существующих моделях не делается попытка формализации принятых решений по построению четкой последовательности действий, необходимых для изменения. В существующих моделях организации процесса управления изменениями не учитываются достижения современного менеджмента, методы управления организацией, основанные на принципах процессного подхода и международных стандартах системы управления качеством.

Подтверждение актуальности исследования

Таким образом, указанные недостатки моделей управления изменениями обуславливают актуальность проведения исследования, в рамках которого необходимым представляется совершенствование моделей и методов управления изменениями бизнес-процессов организации.

Цель исследования

Цель исследования заключается в разработке методик и моделей управления изменениями процессов, направленных на обеспечение своевременной адаптации организации к изменениям внутренней и внешней среды и снижение затрат на организацию и проведение изменений.

Задачи исследования

Для достижения цели исследования поставлены и решены следующие задачи:

- построены классификации изменений бизнес-процессов организации для определения параметров изменений деятельности организации;
- определена концептуальная модель знаний о деятельности для организации хранения и актуализации моделей бизнес-процессов организации;
- разработана система показателей оценки бизнес-процессов с целью выявления проблемных мест для обоснования решений по изменению

процессов организации;

– разработана и формализована методика управления изменениями бизнес- процессов организации, направленной на обеспечение целостности, семантической корректности принятых решений и оперативной адаптации к изменениям внешней и внутренней среды;

– разработана архитектура системы управления изменениями, позволяющая автоматизировать процессы адаптации деятельности организации к изменениям внешней и внутренней среды;

– разработана экономико-математическая модель расчета затрат на выполнение бизнес-процесса в результате его адаптации к изменениям.

Объект и предмет исследования

В качестве объекта исследования выступают процессы изменения деятельности организации.

Предметом исследования являются модели управления изменениями деятельности организации.

Тема такого исследования, очерчивающая границы цели может быть определена как «Разработка методики и моделей управления изменениями бизнес-процессов».

Второй этап исследования - разработка исследовательской программы. Исследовательская программа должна быть ориентирована на выдвинутые гипотезы. Основная его часть включает план исследования и описание основных процедур сбора и анализа эмпирической информации. При составлении плана исследования следует четко придерживаться выбранной темы и не выходить за ее пределы, включать разделы в логическую последовательность, отражающую все аспекты исследования.

В этом же разделе исследований осуществляется выбор методов исследования, показывается связь данных методов с целями, задачами и гипотезами исследований.

Важным этапом научного исследования является реализация исследовательской программы. Эта стадия носит исследовательский характер. Она заключается не только в систематическом изучении литературы по предмету исследования, статистических данных и архивных материалов, но и предполагает осуществление теоретических и эмпирических исследований, обработку, обобщение и анализ полученных результатов, объяснение новых полученных фактов, аргументацию и формулировку положений, выводов, практических рекомендаций и предложений.

Этап «Работа над рукописью и ее оформление» включает определение

композиции, внутренней структуры работы, уточнения темы, названий глав и параграфов, подготовку черновой рукописи, ее редактирование, оформление таких элементов рукописи как список литературы и интернет - источников, графических материалов, приложений. Стандарта по композиции исследовательского труда не существует, и каждый автор выбирает внутреннюю логическую связь между частями рукописи таким образом, чтобы в полной мере раскрыть проблему исследования. Однако сложилась устойчивая тенденция расположения элементов рукописи в следующей последовательности:

1. Титульный лист
2. Оглавление
3. Введение
4. Главы основной части
5. Заключение
6. Список используемых источников
7. Приложения
8. Вспомогательные указатели

Оформление конкретных элементов работы (оглавления, библиографии, индекса) должно соответствовать стандарту. В целом требования к оформлению результатов научно-исследовательских работ содержит ГОСТ 7.32–2001 «Система стандартов по информации, библиотечному и издательскому делу. Отчет о научно-исследовательской работе. Правила оформления», однако в большинстве случаев существуют и корпоративные рекомендации и требования.

Завершающим этапом является внедрение результатов исследования и авторское сопровождение внедряемых разработок.

3.7. Контрольные вопросы

1. Опишите роль логических методов в научных исследованиях?
2. Являются ли логические методы научными?
3. Какие методы входят в группу логических методов?
4. Дайте краткую характеристику методам анализа и синтеза.
5. Какое практическое значение играют анализ и синтез в научных исследованиях?
6. Назовите основные требования к проведению сравнительного анализа.

7. Дайте определение методам индукции и дедукции.
8. Приведите пример использования индукции и дедукции в научных исследованиях.
9. Перечислите основные этапы научного исследования.
10. Что понимается под концепцией исследования.
11. Какие методы используются при разработке концепции?
12. Что вы понимаете под термином «программа исследования»?
13. Является ли процесс формирования научной гипотезы обязательным? Поясните ответ.
14. Назовите основные этапы проведения научного исследования.
15. Поясните содержание методов дивергенции и конвергенции.
16. Какие методы включает в себя трансформация проблемы?
17. Какие отличительные особенности имеет рукопись научной работы от других видов рукописей?

3.8. Задания для самостоятельной работы

1. Перед исследователем стоит задача выявить проблемы занятости населения в городе. Определите характеристики исследования в соответствии со следующим планом.
 - a. Постановка проблемы.
 - b. Формулировка проблемы.
 - c. Цель исследования.
 - d. Задачи исследования.
 - e. Объект исследования и предмет исследования.
 - f. Гипотеза исследования.
 - g. Методология и методы исследования.
2. Перед исследователем стоит задача проведения исследования с целью повышения эффективности работы сельскохозяйственного предприятия. Определите характеристики исследования в соответствии с планом, представленном в п.1.
3. Определите характеристики исследования по плану, представленному в п.1 в соответствии с одной из следующих тем:
 - a. «Совершенствование механизма и особенности деятельности фирмы на международном рынке транспортных услуг».
 - b. «Инвестиционные риски и их расчет методами математического

моделирования».

с. «Стратегическое планирование развития агропромышленных кластеров».

d. «Совершенствование бизнес-процессов в устойчивом развитии фирмы».

e. «Маркетинговые стратегии устойчивого развития фирмы».

f. «Анализ динамики малого и среднего предпринимательства».

g. «Создание потребительской ценности при выходе на международный рынок».

Вариант правильного ответа должен включать классификацию источников информации на вторичные (публикации в открытой печати, обзоры рынка и т.п.) и первичные (результаты опросов потребителей, собственные исследования компании).

4. ЭМПИРИЧЕСКИЕ МЕТОДЫ ИССЛЕДОВАНИЙ В ИНФОРМАТИКЕ

4.1. Метод наблюдения

Отправной точкой научного исследования является первичная информация. Для исследований в области информатики и экономики *первичная информация* – это факты, отражающие тот или иной хозяйственный процесс, явления, события. В эпоху повсеместной компьютеризации значительная часть фактического материала формируется в информационных системах и хранится в базах данных и знаний в виде структурированных или неструктурированных данных.

Значительную роль для проведения научных исследований играют факты, сформированные в первую очередь в ходе эмпирического исследования, которое базируется на непосредственном практическом взаимодействии исследователя с изучаемым объектом. Оно предполагает осуществление *наблюдений* и *экспериментальную деятельность*.

Наблюдение – это целенаправленное, осмысленное, организованное восприятие внешнего мира с целью получения первичной эмпирической информации в форме эмпирических фактов. Наблюдение является методом получения первичных данных от респондентов, которое может быть отнесено как к *качественным* методам исследования, так и к *количественным*. Этот метод находит широкое применение в различных областях и ориентирован не столько на подсчет и статистическое исследование, сколько на содержательный смысл объектов исследования, обобщение, сравнение, кластеризацию и интерпретацию исследуемых процессов и явлений. Наблюдения используются для решения задач различной категории.

Категория 1. Поиск и детализация проблемы исследования:

- детальное определение проблемы дальнейшего исследования;
- выдвижение гипотез для количественного исследования, когда тестируются гипотезы и разрабатываются рекомендации по решению проблем, стоящих перед заказчиком исследования;
- разработка концепций новых продуктов;
- получение предварительной реакции потребителей на концепцию продукта, на варианты рекламных сообщений и т.д.;
- предварительное тестирование анкет до начала массовых опросов.

Категория 2. Ориентирование:

- определение приоритетов потребителей, их лексики для использования ее в составлении рекламных текстов, ориентированных на специфическую целевую аудиторию с определенным сленгом, например, тинейджеров;

- получение информации о незнакомой внешней среде: потребностях, ситуациях использования продукта, проблемах.

Категория 3. Характерные ситуации:

- получение информации, которую невозможно собрать при помощи статистических методов исследования.

Основные виды экономического наблюдения представлены в таблице 4.1.

Таблица 4.1 – Виды наблюдений

Классификационный признак	Виды методов
По степени воздействия на объект исследования	Наблюдение
	Эксперимент
Участие наблюдателя	Включенное наблюдение
	Невключенное наблюдение
Условия проведения	Полевое
	Лабораторное
Руководящая основа	Стандартизированное
	Нестандартизированное
Объект наблюдения	Акторное (наблюдение за действиями и поступками людей)
	Вербальное (наблюдение за словами и мнением респондентов)

Включенное и невключенное наблюдения соответствуют факту (или отсутствию такового) непосредственного участия в наблюдателя в тех или иных наблюдаемых им процессах или явлениях. Способ наблюдения, при котором исследователь становится частью изучаемой группы (фирмы), не ставя при этом в известность остальных членов коллектива, называется *скрытым* включенным наблюдением. *Открытое* включенное наблюдение, когда группа ставится в известность о присутствии наблюдателя, считается неэффективным методом. Непосредственное участие наблюдателя в исследуемых процессах крайне редко, чтобы не повлиять на результаты исследования.

Полевое и лабораторное наблюдения соответствуют условиям проведения исследования. Все наблюдения разделяют на *стандартизированные* и *нестандартизированные*. Это деление обусловлено

необходимостью выполнения некоторых подготовительных работ при применении стандартизированных исследований (формирование опросных листов, тестов, программ мониторинга и т.д.).

Наиболее существенным с точки зрения практического применения наблюдения является его деление на *акторное* и *вербальное*. Акторное наблюдение предполагает наблюдение за действиями и поступками респондентов – акторов, а вербальное – наблюдение за словами и мнениями респондентов.

Рассмотрим более подробно различные виды наблюдений, широко применяемые в маркетинговых исследованиях.

Экспертные интервью (expertinterview) – разновидность глубинных интервью, как метода качественных исследований, которые организуются и проводятся с экспертами. В качестве экспертов могут быть выбраны специалисты из отрасли, менеджеры компаний или владельцы бизнес-процессов. При проведении экспертных интервью исследователи разрабатывают список вопросов, которые будут задавать эксперту. Этот список может носить неструктурированный характер, поскольку эксперт может либо расширить какую-то тему в соответствии со своим знанием отрасли и проблемы, либо, наоборот, игнорировать ответ на один или несколько вопросов. При разработке структурированной анкеты все вопросы должны получить ответ, иначе такие анкеты отбраковываются из массива полученных данных.

Глубинное интервью (in-depth interview) – представляет собой неструктурированное, прямое, личное интервью, в котором одного респондента (потребителя) опрашивает высококвалифицированный интервьюер для определения его основных побуждений, эмоций, отношений и убеждений по определенной теме. Основное отличие глубинного интервью от экспертного в том, что в качестве респондента здесь выступает потребитель продукта или услуги. Цель глубинного интервью - выявление скрытых чувств, мотивов и ценностей респондента.

Индивидуальные глубинные интервью проводятся по заранее составленному плану (сценарию) и имеют продолжительность от 30 до 90 минут в среднем (иногда порядка 45 минут), как правило, проводятся специально оборудованном помещении с односторонним зеркалом, за которым представители заказчика исследования могут наблюдать за ходом проведения интервью. Данный метод находит применение, когда:

- тема исследования очень личная (например, состояние здоровья,

финансовое положение респондента и др.);

- потенциальный участник – компетентное лицо, не имеет права на обсуждение ряда вопросов, имеющих коммерческую тайну в присутствии потенциальных конкурентов;

- труднодостижимость некоторых аудиторий позволяет заключить, что легче провести интервью с представителями этих аудиторий отдельно, чем собирать их в одну фокус-группу;

- товарная категория представляет собой очень сложную область изучения.

Использование метода имеет ряд ограничений, в том числе стоимость исследования, которая выше, чем в других методах, время получения результатов и сложность их интерпретации из-за последовательного получения информации и результатов.

Несмотря на то, что процедуры наблюдения трудно формализовать, тем не менее, существуют требования к регистрации результатов наблюдений. Например, если вы используете очень популярный сегодня метод наблюдения, называемый методом "тайного покупателя", то при планировании исследования необходимо сразу определить параметры для сравнения объектов. Одним из методов экономического наблюдения является *мониторинг* – тип наблюдения, при котором ведется контроль над процессами и явлениями с целью контроля их нахождения в заданном диапазоне изменения. Мониторинг включает в себя систематическое наблюдение, оценку изменений, происходящих с объектом исследования, что позволяет принять оперативные меры в управлении объектом или в последствии осуществить прогноз и выявить тенденции происходящих изменений.

Экономический мониторинг в зависимости от масштаба исследований может рассматриваться как глобальный, национальный, региональный, локальный и точечный. В зависимости от исследуемых процессов выделяют такие типы мониторинга как демографический, аграрный, промышленный, транспортный, экологический и др.

В настоящее время мониторинг находит все большее применение в экономических исследованиях.

4.2. Метод экономического эксперимента

Эксперимент – эмпирический метод исследования, система операций, воздействий и (или) наблюдений, направленных на получение информации об объекте при исследовательских испытаниях.

Одной из основных целей проведения экспериментов является проверка гипотез научных исследований.

Различные способы организации и проведения экспериментов выделить некоторые классификационные признаки и определить несколько видов эксперимента (таблица 4.2).

Таблица 4.2 – Виды экспериментов

Классификационный признак	Виды методов
Природа изучаемого объекта	Физический, экономический, демографический, технологический и др.
Степень воздействия на объект	Пассивный
	Активный
Условия проведения эксперимента	Лабораторный эксперимент
	Натурный эксперимент
Характер взаимодействия с объектом исследования	Материальный
	Вычислительный
	Мыслительный
Число варьируемых факторов	Однофакторный
	Многофакторный
Ожидаемые результаты	Количественные
	Качественные

По степени воздействия на объект исследования выделяют *активные* и *пассивные* эксперименты. При активном эксперименте исследователь на основе выявленных доминирующих факторов варьирует переменные и параметры по разработанной программе, т.е. осуществляется целенаправленный поиск наилучшего решения. Пассивный эксперимент сводится к наблюдению процессов и явлений и измерению параметров и переменных, характеризующих исследуемый объект.

В зависимости от условий проведения эксперимента различают *лабораторный* и *натуральный эксперимент*. Лабораторный эксперимент проводится в специально подготовленных условиях. Для экономических экспериментов это в большинстве случаев – вычислительный эксперимент, состоящий в проведении исследований с помощью вычислительных комплексов или систем с применением математических или имитационных моделей. Особым видом лабораторного эксперимента являются деловые игры, кейсы, базирующиеся на реальных экономических, социальных и других ситуациях. Их реализация также является одним из методов исследований, направленных на поддержку и принятие решений в информатике.

Примером лабораторных исследований могут служить исследования,

проводимые в ситуационных центрах Министерства природных ресурсов РФ, Минатоме и Росэнергоатоме, в МЧС, в некоторых автономных округах и регионах. Ситуационные центры существуют и в крупных промышленных и нефтегазовых компаниях.

Натуральные эксперименты ставят целью изучить процесс или явление в реальных условиях с учетом воздействия внешних и внутренних факторов.

Материальный эксперимент, получивший в литературе название *классического эксперимента*, использует в процессе исследования реальные системы или их физические или аналоговые модели. Этот метод широко распространен в технических исследованиях, однако находит свое развитие и в экспериментальной экономике.

Вычислительный эксперимент – современная технология экспериментирования на компьютере с моделью функционирования объекта на протяжении длительного времени. Такой эксперимент в отличие от материального (классического эксперимента) осуществляется с применением математической или имитационной модели и позволяет исследователю расширить диапазон задач, изучить новые явления и процессы, оценить последствия внешних воздействий и рассмотреть многие ситуации, недоступные в случае классического эксперимента.

Для вычислительного эксперимента этап проектирования процесса эксперимента дополняется разработкой модели, на которой будет проводиться исследование. Такой метод эмпирического исследования позволяет решать широкий круг задач:

- подтвердить или опровергнуть выдвинутые гипотезы,
- оценить внешние (внутренние) факторы, влияющие на объект исследования;
- оптимизировать затраты на экспериментальные исследования, в том числе за счет выявления несущественных факторов и уменьшения размеров модели.

Мыслительный эксперимент представляет собой форму умственной деятельности исследователя, в процессе которого воспроизводится структура реального эксперимента.

Качественный эксперимент ставит перед собой задачу установить наличие или отсутствие предполагаемого теорией явления. *Количественный эксперимент* направлен на оценку свойств и характеристик исследуемого объекта количественными методами.

Однофакторный и *многофакторный* эксперименты отличаются

количеством факторов и независимых переменных, используемых в процессе исследования методом научного эксперимента.

Очевидно, что каждый из рассмотренных видов эксперимента требует от исследователя специальных знаний, больших затрат интеллектуальных и временных ресурсов. При этом следует отметить, что классический и вычислительный эксперименты имеют ряд ограничений. Так, классический эксперимент в силу невозможности выделения фрагмента единого экономического пространства без существенных потерь, не позволяет проверить теоретические гипотезы. Вычислительный эксперимент ограничен невозможностью моделирования значительной области экономических явлений в первую очередь из-за недостатка или незрелости (неформализуемости) содержательного знания. Это в первую очередь относится к тем экономическим процессам, где существенно поведение человека.

4.3. Контрольные вопросы

1. Что вы понимаете под термином «наблюдение»?
2. Какую роль играет наблюдение в научных исследованиях?
3. В чем состоит принципиальное отличие в эмпирических методах «наблюдение» и «эксперимент»?
4. Назовите основные виды наблюдений.
5. Что вы понимаете под термином «эксперимент»?
6. К какому виду экономических исследований относят анкетирование и опрос?
7. Можно отнести к натурному эксперименту создание особых экономических зон? Обоснуйте ответ.
8. Какие этапы включает в себя эксперимент?
9. В чем состоит отличие вычислительного эксперимента от мыслительного? Приведите пример мыслительного эксперимента.

4.4. Задания для самостоятельной работы

1. Выполните проектирование экспериментального исследования для определения зависимости месторасположения продукта в магазине на объем продаж.
2. Определение цены товара может осуществляться с применением

различных методов. Метод, ориентированный на затраты, предполагает расчет покрытия затрат и обеспечения заданного уровня прибыли. Ценообразование, ориентированное на потребителя, использует различные стратегии, основанные на изучении спроса и предпочтений потребителей. Ценообразование, ориентированное на конкурентов, исходит из цен, существующих в настоящее время на рынке. Является ли задача определения цены, исходя из различного уровня планируемой прибыли, вычислительным экспериментом? Зависит ли принадлежность исследования к эксперименту от выбранного метода определения цены?

3. Разработайте опросный лист, позволяющий определить степень лояльности студентов к учебному заведению.

4. Разработайте критерии оценки удовлетворенности персонала медицинского учреждения и анкету для получения возможности оценки критериев.

5. Разработайте анкету, позволяющую оценить факторы, препятствующие развитию малого и среднего бизнеса.

5. ТЕОРЕТИЧЕСКИЕ МЕТОДЫ ИССЛЕДОВАНИЙ В ИНФОРМАТИКЕ

5.1. Методы формализации и математизации

Эмпирический уровень научного познания характеризуется непосредственным исследованием реально существующих, чувственно воспринимаемых объектов. На этом уровне осуществляется процесс накопления информации об исследуемых объектах (путем наблюдения, измерения, экспериментов) здесь происходит первичная систематизация полученных знаний (в виде таблиц, схем, графиков).

Теоретический уровень научного исследования осуществляется на рациональной (логической) ступени познания. На данном уровне происходит выявление наиболее глубоких, существенных сторон, связей, закономерностей, присущих изучаемым объектам и явлениям. Результатом теоретического познания становятся гипотезы, теории, законы.

К теоретическим методам научных исследований относится *формализация*.

Формализация - метод, основанный на выявлении и фиксации формальной структуры исследуемого процесса или явления, приписывающий содержательным элементам процесса (явления) некоторые абстрактные символы и значения. Результатом формализации является построение некоторой модели, позволяющей получить новое знание или информацию об исследуемом объекте.

Процесс формализации может осуществляться различными способами. Первый способ представляет собой формализацию на базе *естественного языка*. Однако этот тип формализации не обеспечивает однозначного использования понятий и, соответственно, приводит к трудностям научной коммуникации. В этой связи полезным является использование *логической формализации*, выполняемой на основе правил формальной и математической логики. Основной частью логической формализации является ее *символизация*, т.е. разработка системы определенных условных обозначений в рамках создания искусственного языка науки.

Наиболее распространенным и эффективным типом формализации является математизация. Математизацию можно определить, как специфический метод формализации, основанный на применении процедур измерения, сравнения и счета.

Выделяют следующие направления математизации в информатике:

– математическое моделирование — построение формально-

количественных, математических моделей исследуемых процессов и явлений;

– эконометрика — решение социально-экономических задач с применением принципов статистического измерения и сбора данных, численного выражения исследуемой реальности с определением меры и границ этого выражения;

– математический эксперимент — метод многократного математического моделирования как эксперимент над моделями.

5.2. Математическое моделирование как метод научных исследований в информатике

Классификация математических моделей

Математическое моделирование является одним из основных инструментов научных исследований, что привело к созданию значительного числа экономико-математических моделей.

Математическая модель является приближенным представлением реальных объектов, процессов или систем, выраженным в математических терминах и сохраняющим существенные черты оригинала. Математические модели в количественной форме, с помощью логико-математических конструкций, описывают основные свойства объекта, процесса или системы, его параметры, внутренние и внешние связи. В таблице 5.1. приведены виды моделей, получивших широкое распространение в научных исследованиях в области информатики.

Таблица 5.1 – Виды математических методов исследования

п/п	Вид модели	Процесс и цель	Методы представления
.	Оптимизация задач с несколькими альтернативами	Найти лучшее решение из небольшого количества альтернатив	Деревья решений. Таблицы решений
.	Оптимизация через алгоритм	Найти лучшее решение из большого или бесконечного числа альтернатив, используя процесс пошагового управления	Линейные и другие модели математического программирования; сетевые модели
.	Имитационные модели	Нахождение приемлемого решения, или лучшего среди проверенных альтернатив, с использованием экспериментов	Различные методы имитационного моделирования

	Эвристика	Найти приемлемое решение, используя правила	Эвристическое программирование; экспертные и интеллектуальные системы
	Другие модели	Нахождение ситуации что-если (what - if), используя правила	Финансовое моделирование
	Предсказательные модели	Предсказание будущего на основе данного сценария	Модели прогнозирования; марковский анализ

Оптимизация с несколькими альтернативами предполагает создание модели, позволяющей найти решение из конечного числа альтернатив, представленных в виде таблицы или графика с их прогнозируемыми вкладами в достижение цели и вероятностями таких вкладов. Модели, описывающие реальные процессы и явления в виде явных функциональных зависимостей (линейных или нелинейных уравнений, дифференциальных или интегральных уравнений, систем этих уравнений), называются аналитическими моделями. Их применение в информатике ограничено сложностью изучаемых процессов и явлений, когда аналитические зависимости не всегда удается построить. В этом случае исследователи прибегают к моделированию.

Имитационное моделирование-это численный метод проведения вычислительных экспериментов с помощью информационных технологий и систем с математическими моделями, моделирующими поведение реальных объектов, процессов и явлений во времени в течение заданного периода. Его реализация базируется на совокупности методов алгоритмизации функционирования объектов исследования, программной реализации алгоритмических описаний, организации, планировании и проведении экспериментов с математическими моделями, моделирующими функционирование исследуемых процессов и явлений в течение заданного периода средствами информационных технологий. Термин "моделирование" означает, что мы имеем дело с такими математическими моделями, которые не могут быть использованы для расчета или прогнозирования поведения системы заранее, а для прогнозирования поведения системы необходим вычислительный эксперимент (моделирование) на математической модели по заданным исходным данным.

Эвристика - специальные методы решения задач (эвристические методы), которые обычно противопоставляются формальным методам решения задач, основанным на точных математических моделях.

Использование эвристических методов (эвристик) сокращает время решения задачи по сравнению с методом полного ненаправленного поиска возможных альтернатив; полученные решения не являются, как правило, наилучшими, а относятся только к набору приемлемых решений.

Предсказательные модели основываются на методах прогнозирования и имеют важное значение в исследовательских проектах, направленных на повышение эффективности экономики.

5.3. Методы оптимизации

В математическом программировании как разделе прикладной математики рассматриваются экстремальные задачи и методы их решения. С помощью построения и решения таких задач возможно улучшение управления и организации производственных процессов. Экстремальные задачи представляют собой задачи условной оптимизации, в которые входят ограничения и целевые функции.

Задачи оптимального или математического программирования широко используются для планирования отраслей сельского хозяйства, рационального использования посевных площадей, моделирования кормовых рационов, оптимизации межхозяйственной кооперации и т.д.

В литературе используется классификация задач математического программирования по типам переменных, их связям, учету времени, свойствам информации, количеству критериев оптимальности, проявлению экстремальных явлений.

В задачах математического программирования встречаются дискретные и непрерывные переменные. В частности, в целочисленных задачах неизвестные могут принимать только целые значения, а при оптимизации посевных площадей переменные соответствуют любым неотрицательным величинам. При этом в ограничениях или целевых функциях связи между переменными имеют линейный или нелинейный вид.

Если переменные задачи оптимального программирования зависят от времени, то их относят к динамическим в отличие от статических задач, где фактор времени не применяется.

Поскольку неизвестные могут принимать случайные или детерминированные значения, то рассматривают определенные и стохастические задачи математического программирования. В ряде случаев, когда неизвестны законы распределения вероятностей случайных величин, говорят о задачах математического программирования с неопределенными

параметрами.

Во многих случаях приходится находить компромиссные решения, если в задачу входят участники с различными целями. В этом случае возникают многокритериальные задачи.

И, наконец, производственные процессы в сельском хозяйстве связаны с внешними факторами. При учете экстремальных природных явлений, наводнений, засух, раннего снега, заморозков и др., причиняющих значительный ущерб сельскохозяйственным предприятиям, интерес вызывают задачи математического программирования с учетом природных событий.

Следует отметить, что многие территории страны относятся к зонам рискованного земледелия, где часто наблюдаются экстремальные природные явления, приносящие значительные ущербы сельскохозяйственным товаропроизводителям. Поэтому модели производственных процессов должны учитывать влияние на сельское хозяйство стихийных явлений.

Задачи линейного программирования

В общем виде задача линейного программирования записывается следующим образом:

$$f = \sum_{j \in J} c_j x_j \rightarrow \max(\min), \quad (5.1)$$

$$\sum_{j \in J} a_{ij} x_j \leq b_i, \quad i \in I, \quad (5.2)$$

$$x_j \geq 0, \quad j \in J, \quad (5.3)$$

где f - целевая функция или критерий эффективности задачи, x_j - переменная, a_{ij} , b_i , c_j - заданные постоянные величины, i, j - индексы.

Задачу (5.1) – (5.2) называют стандартной задачей линейного программирования. Совокупность чисел $X=(x_1, x_2, \dots, x_n)$, удовлетворяющих условиям (5.2)– (5.3), представляет собой допустимое решение или план. План $X^*=(x_1^*, x_2^*, \dots, x_m^*)$, при котором критерий оптимальности (5.1) принимает экстремальное значение, принято называть оптимальным. Очевидно, что, если целевая функция (87) стремится к максимуму, тогда $f(X) \leq f(X^*)$, в противном случае, когда целевая функция достигает минимума, - $f(X) \geq f(X^*)$.

С помощью задачи линейного программирования определяется структура сельскохозяйственного производства предприятия и его отраслей, скотоводства и растениеводства. Кроме того, на основе задач линейного программирования оценивается кормопроизводство и распределение

посевных площадей на различном уровне иерархии, от предприятия до муниципального образования и региона, и структура стада.

Любая стандартная задача линейного программирования может быть приведена к каноническому виду. В этом случае неравенства (5.2) преобразуются в равенства введением дополнительных переменных. Если связь левой и правой частей условия (5.2) определяется символом « \leq », тогда в левую часть добавляют неотрицательную переменную. При другом виде неравенства « \geq » из левой части вычитают дополнительную неотрицательную переменную. Таким образом, ограничение (88) преобразуется в равенство:

$$\sum_{j \in J} a_{ij} x_j = b_i \quad i \in I. \quad (5.4)$$

Кроме преобразования стандартной задачи к каноническому виду, возникают задачи замены минимизации целевой функции ее максимизацией и наоборот. Для этого используют равенство $f_{max} = -f_{min}$ или $f_{min} = -f_{max}$.

Пример. Фирма занимается производством двух видов полок. Дано 100 м³ досок для изготовления полок двух образцов А и В. Для изготовления 1 полки образца А требуется 1 м³ досок, а для производства полок второго образца В – 2 м³ материала. Для изготовления единицы первого образца полка используется 2 ч машинного времени. Одна полка второго образца производится за 3 ч. Общее количество времени для изготовления полок ограничено 180 ч. Согласно договоренности с потребителями фирме необходимо изготавливать не менее 20 изделий первого образца. Стоимость одной полки образца А составляет 5 денежных единиц, а одного изделия вида В – 8 д.е. Построить целевую функцию и ограничения задачи линейного программирования при условии получения максимального дохода.

Решение

Обозначим число полок А переменной x_1 , а количество изделий В – x_2 . Очевидно, что стоимость первой продукции равна $5x_1$, а второй – $8x_2$, поскольку 5 и 8 представляют собой стоимость одной полки разных образцов. Тогда целевая функция примет вид

$$f(X) = 5x_1 + 8x_2 \rightarrow \max.$$

Ограничение по затратам материалов на изготовление полок определяется так

$$1x_1 + 2x_2 \leq 100 \text{ [м}^3\text{]},$$

где $1x_1$ – затраты материала на изготовление полок первого образца, а $2x_2$ – объем досок для производства полок второго образца.

Условие не превышения заданного машинного времени записывается следующим образом

$$2x_1 + 3x_2 \leq 180 \text{ [ч]},$$

где $2x_1$ – затраты времени на производство полок первого образца, а $3x_2$ – количество часов на изготовление полок В.

Ограничение, связанное с необходимостью производства не менее 20 полок образца А, имеет вид

$$x_1 \geq 20.$$

Кроме того, должно соблюдаться условие неотрицательности переменных или $x_2 \geq 0$.

Таким образом, задача включает в себя 2 переменные, состоит из 4 ограничений и целевой функции, которую максимизируют.

5.4. Статистические методы в научных исследованиях

Довольно часто при проведении научных исследований описываются ситуации, в которых рассматриваются детерминированные события функции, когда каждое событие является следствием другого, а физические законы описываются математическими зависимостями одной величины от другой. Человеческая деятельность довольно часто опровергает это правило, и при описании ситуаций фиксируются не только отклонения измеряемых параметров, но и значительная неопределенность некоторых из них. Например, невозможно точно предсказать количество покупателей в магазине или объем продаж в данный момент времени. Значения, точное значение которых неизвестно, называются случайными.

Несмотря на случайный характер большинства параметров, анализируемых в информатике, научные исследования позволяют сформулировать и описать некоторые закономерности, которые широко используются на практике. Эти исследования основаны на статистике-наборе методов, позволяющих принимать решения в условиях неопределенности.

Статистическая наука состоит из нескольких разделов, каждый из которых имеет большое значение для исследований в области информатики. Основой статистики является статистическая теория, отражающая общие принципы статистической науки, ее категории и методы исследования. Ядром является математическая статистика-раздел математики, основанный на численных методах обработки статистических данных, полученных в результате наблюдения и измерения случайных величин. Другой важной

отраслью статистической науки является прикладная статистика. В ней исследуются статистические проблемы, непосредственно связанные с практической деятельностью: измерением, сбором, обработкой и анализом данных в различных областях научных исследований.

Главным условием правильного восприятия и практического использования статистической информации является знание статистической методологии.

Статистическая методология как система приемов, методов и приемов, направленных на изучение количественных закономерностей, проявляющихся во взаимосвязи с социально-экономическими явлениями, включает следующие этапы исследования:

1. Организованная регистрация собранных фактов о массовых социально-экономических, общественных явлениях и процессах.

2. Сводка и группировка статистических данных – это систематизация первичных данных по признакам, объединяющим в качественно однородные группы.

3. Анализ совокупных данных, полученных в результате сводки и группировки, при этом используются обобщенные показатели: абсолютных, относительных и средних величин, показатели вариации, ряды динамики, анализ взаимосвязей и индексы.

Таким образом, основой статистических исследований являются метод статистического наблюдения, группировка и методы статистического анализа.

Метод статистического наблюдения является предварительной стадией статистического исследования, которая представляет собой планомерный, научно организованный учет (сбор) первичных статистических данных о массовых социально-экономических явлениях и процессах. Статистическое наблюдение должно отвечать ряду важнейших требований:

- проводиться непрерывно и систематически;
 - учет массовых данных должен быть таким, чтобы не только обеспечивалась полнота данных, но и учитывалось их постоянное изменение;
 - данные должны быть достоверными и точными;
 - данные должны соответствовать принципу единообразия и сопоставимости.
- исследуемые явления должны иметь не только научную, но и практическую ценность.

При проведении научных исследований не удастся получить всю совокупность значений изучаемой величины (так называемую *генеральную*

совокупность), поэтому в большинстве случаев используется только часть данных (*выборка*).

При проведении исследований необходимо:

- применять однородную, репрезентативную выборку, чтобы результаты исследования могли бы быть распространены на генеральную совокупность;
- выполнять анализ точности полученных результатов;
- обосновывать способы формирования выборки;
- применять для обработки данных способы, соответствующие их особенностям.

Сбор и регистрация статистических фактов предполагает возможность их *измерения* – присвоения чисел исследуемым предметам, событиям, явлениям, процессам. Для регистрации измерений применяются *шкалы*:

- шкала классификации (наименования);
- шкала порядка;
- шкала интервалов;
- шкала отношений.

Шкала классификации (наименований) используется для идентификации и классификации объектов и позволяет выполнить сравнение на уровне «равно» и «не равно».

Шкала порядка (порядковая шкала) – это шкала рангов, позволяющая исследователям определить больше или меньше характеристика одного объекта по сравнению с другим. Однако, используя эту шкалу, мы не можем определить на сколько больше или меньше величина исследуемой характеристики. Например, исследование с использованием порядковой шкалы позволяет выявить отношение респондентов к характеристике объекта в виде «неудовлетворен», «скорее не удовлетворен», «скорее удовлетворен», «удовлетворен», но не позволяет выявить, в какой степени респондент не удовлетворен исследуемым аспектом.

Шкала интервалов позволяет сравнивать величины и определять «насколько больше», «насколько меньше».

Шкалы отношений ориентированы на выяснение вопроса «во сколько раз». В информатике к шкалам отношений относятся такие переменные как доля рынка, объем продаж, количество потребителей.

Пример опроса удовлетворенности персонала включает в себя вопросы с использованием следующих шкал:

1. *Подразделение – шкала наименований.*

2. *Порядковый номер участника соревнований – шкала наименований;*
3. *Стаж работы – шкала отношений;*
4. *Оценка удовлетворенности уровнем заработной платы – шкала интервалов;*
5. *Ранг эффективности использования нового оборудования – порядковая шкала.*

Случайная величина может быть дискретной или непрерывной. Если множество значений случайной величины конечное или счетное, то есть их можно пронумеровать, то случайная величина называется дискретной. Случайная величина называется непрерывной, если она принимает все возможные значения из определенного интервала или по всем числовым осям. Дискретная величина обычно задается рядом распределения, непрерывная величина-функцией или плотностью распределения.

Для обработки данных, собранных в результате наблюдений, используется метод группировки. Группировка-это распределение совокупности единиц исследуемой популяции по группам в соответствии с существенным для данной группы признаком. Несомненным преимуществом данного метода является возможность первичного обобщения данных, сравнительного анализа, обобщения показателей по группам и внутри группы. Появляется возможность сравнивать и анализировать причины различий между группами, изучать взаимосвязи между признаками.

Признак, отличающий группы или типы явлений, называется группировочным или основой группировки. Основа может быть количественной или атрибутивной. Атрибут - это признак, имеющий название (например, профессия: инженер, учитель и т. д.). Если группировка основана на количественном признаке, то возникает вопрос о вычислении интервалов группировки, которые могут быть как равными, так и неравными и, как правило, представляют собой интервал между максимальным и минимальным значениями признака в группе. Устойчивое разграничение объектов представляет собой *классификацию*. Она основывается на устойчивых признаках, (например, классификация отраслей народного хозяйства, классификация основных фондов и т.д.).

Статистический анализ можно рассматривать как количественное исследование в информатике, проведенное в соответствии с требованиями статистической теории и методологии.

Современный статистический анализ располагает широким спектром методов, обеспечивающих эффективную обработку массивов экономической информации. Возможность вычисления статистических функций и применение различных методов анализа определяются свойствами измеряемых величин и различных шкал измерения.

Индексный метод

Индексный метод – сведение количественных оценок в статистические показатели (индексы). *Индекс* представляет собой относительную величину, показывающую во сколько раз уровень изучаемого явления отличается от того же явления в других условиях. Для расчета индексов принято вводить следующие обозначения:

- индивидуальный индекс;
- сводный индекс;

Индивидуальные индексы позволяют осуществить сравнительный анализ отдельных элементов объекта, явления или процесса. Примером могут служить индексы объемов производства, цен на продукцию или товары, индексы себестоимости.

Сводные индексы применяются для оценки сложного, несопоставимого в реальных условиях явления, например изменения объема разноименных товаров (запасов). Сводные индексы могут быть разделены на *групповые* и *общие*. Групповой индекс рассчитывается для группы элементов, входящих в общую совокупность. Примером может служить индекс цен товаров, входящих в какую-либо товарную группу. Общий индекс рассчитывается для всей совокупности элементов изучаемого явления.

Основной формулой для расчета сводного индекса является агрегатная формула, в которой с помощью весов индекса величины приводятся к соизмеримому виду. Примеры расчета индексов представлены на рисунке 5.1.

В зависимости от содержания и характера индексируемой величины различают индексы *качественных* и *количественных показателей*. К индексам количественных (объемных) показателей относятся индексы физического объема производства продукции, стоимости продукции, затрат на выпуск продукции, а также индексы показателей, размеры которых определяются абсолютными величинами.

Качественные показатели определяют уровень исследуемого итогового показателя, формируемого путем деления конечного показателя и определенного количественного показателя (например, средняя заработная

плата определяется соотношением фонда оплаты труда и численности работников). Индексы качественных показателей включают индексы цен, себестоимости, средней заработной платы и производительности труда. Наиболее распространенным индексом в этой группе является индекс цен.

Индивидуальный индекс	
Индивидуальный индекс физического объема продукции	$i_q = \frac{q_1}{q_0}$
Индивидуальный индекс цен	$i_p = \frac{p_1}{p_0}$
Индивидуальный индекс себестоимости	$i_z = \frac{z_1}{z_0}$
Сводные индексы	
Индивидуальный индекс стоимости продукции	$i_{qp} = \frac{q_1 p_1}{q_0 p_0}$
Агрегатный индекс	
Индекс физического объема продукции	$I_q = \frac{\sum q_1 p_0}{\sum q_0 p_0}$
Индекс цен	$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_1}$
Индекс стоимости продукции	$I_p = \frac{\sum p_1 q_1}{\sum p_0 q_0}$

Рисунок 5.1 – Примеры расчета индексов в статистике

Для оценки изменения явления во времени применяют *динамические индексы* (например, изменение цен во времени).

Применение индексного метода позволяет выполнить широкий спектр исследований в информатике:

- проанализировать изменение социально-экономического явления в динамике;
- исследовать соотношения показателей по регионам;
- определить степень влияния одних показателей на другие.

Дескриптивный анализ

Дескриптивный (описательный) анализ базируется на таких статистических показателях, как среднее значение, мода, медиана, стандартное отклонение и др.

Это характеристики переменных, которые могут быть измерены, контролироваться или варьироваться в ходе исследований. Использование описательной статистики позволяет рассмотреть не все значения переменной, а сформировать общее представление о значениях, которые принимает переменная.

Среднее – сумма значений переменной, деленная на n (число значений переменной).

Мода – наиболее часто встречающееся значение переменной.

Медиана – значение, которое разбивает выборку значений переменной на две равные части: одна половина значений переменной лежит ниже медианы, другая – выше. Расчет медианы выполняется следующим образом. Значения переменной упорядочиваются (полученная последовательность называется *вариационным рядом*). При нечетном значении числа наблюдений медиана определяется как среднее число в ряду значений переменной или как среднее арифметическое двух средних чисел.

Максимум и минимум определяются по вариационному ряду.

Дисперсия выборки или *выборочная дисперсия* представляет собой средний квадрат отклонений индивидуальных значений признака от их средней величины. Дисперсия характеризует степень разброса количественных измерений статистической выборки (случайных величин) относительно среднего значения для этой выборки.

Стандартное отклонение (в статистике этот показатель еще называют *среднеквадратическим отклонением*) характеризует степень отклонения данных наблюдений или множеств от среднего значения.

Частота – показатель, показывающий, сколько раз за какой-то период происходило некоторое событие, проявлялось определенное свойство объекта либо наблюдаемый параметр достигал данной величины.

Выбор статистик определится задачами исследования и свойствами выборки. Например, среднее значение, в отличие от медианы, чувствительно к выбросам, что позволяет именно медиану использовать в статистике США в качестве оценки центральной точки доходов населения.

Наибольший интерес в статистических исследованиях представляет оценка вероятности того, что переменная примет данное значение из определенного интервала. Соотношение, устанавливающее связь между случайной величиной и соответствующей вероятностью называется *законом распределения*.

Закон распределения может быть представлен в различной форме: в виде ряда распределения, функции распределения, плотности распределения.

Для дискретной величины ряд распределения задается таблицей, содержащей значения случайной величины и соответствующие им вероятности.

Для непрерывной случайной величины ряд распределения задается

функцией распределения.

Корреляционный анализ

Существенным моментом в статистических исследованиях является выявление взаимосвязи между переменными, характеризующими различные свойства объектов.

Между случайными величинами и может существовать:

1. *Функциональная взаимосвязь* – зависимость, при которой каждому значению переменной соответствует точно определенное значение.
2. *Стохастическая взаимосвязь* – связь, при которой изменение значения одной переменной приводит к изменению закона распределения.
3. *Статистическая взаимосвязь* – зависимость, при которой значение одной переменной изменяется в среднем от того, какие значения принимает другая переменная.

Задачей *корреляционного анализа* является доказательство наличия этой связи и ее силы. Выявление характера (в аналитической форме) этой связи относится к классу задач *регрессионного анализа*.

Если совместное распределение является нормальным, то статистические выводы основываются на выборочном коэффициенте линейной корреляции, в остальных случаях используют коэффициенты ранговой корреляции Кендалла и Спирмена, а для качественных признаков – критерий хи-квадрат.

Независимо от типа, две или более переменных связаны (зависимы), если наблюдаемые значения этих переменных распределены согласованным образом. Другими словами, мы говорим, что переменные зависимы, если их значения систематически согласуются друг с другом в наших наблюдениях.

Корреляция описывает как степень разброса возможных значений случайных величин относительно своих средних значений, так и статистическую связь между ними. Для описания собственно степени статистической связи между случайными величинами используют безразмерную величину, называемую парным коэффициентом корреляции.

Уровень значимости, рассчитанный для каждой корреляции, является основным источником информации о достоверности корреляции. Критерий значимости основан на предположении, что распределение невязок (т. е. отклонений наблюдений от линии регрессии) для зависимой переменной является нормальным (с постоянной дисперсией для всех значений независимой переменной). Исследования методом Монте-Карло показали, что нарушение этих условий не является абсолютно критичным, если размер

выборки не слишком мал и отклонения от нормальности не очень велики.

Отсутствие однородности в выборке также является фактором, смещающим (в ту или иную сторону) выборочную корреляцию. Коэффициент корреляции может быть вычислен по данным, которые поступили из двух или нескольких групп, различающихся по коррелированности признаков.

Регрессионный анализ

Регрессионный анализ является одним из самых распространённых методов обработки результатов наблюдений и в ряде случаев выступает основой для таких методов как планирование эксперимента, многомерный статистический анализ, дисперсионный анализ и др.

В регрессионном анализе различают результирующую (зависимую) переменную и объясняющие (предикторные) переменные.

Функция *регрессии* показывает, каким будет в среднем значение переменной, если переменные примут конкретные значения.

Для решения уравнений регрессии обычно используется метод наименьших квадратов (МНК), позволяющий подбирать такие оценки параметров регрессионной функции, чтобы значения регрессии результирующего показателя как можно меньше отличались от наблюдаемых значений. В тех случаях, когда делается дополнительное предположение о нормальном распределении остатков регрессии, используется метод максимального правдоподобия.

После того, как получены коэффициенты регрессионной зависимости, следует определить величину степени стохастической взаимосвязи результирующей переменной и факторов. Она определяется через общую дисперсию показателя, факторную и остаточную дисперсии.

Для анализа общего качества уравнения линейной многофакторной регрессии используется коэффициент детерминации R^2 . Он определяет долю вариативности результирующего показателя, обусловленную изменением факторных признаков.

Многие важные зависимости в экономике нелинейны. В качестве примера таких регрессионных моделей можно привести зависимость между объемом выпуска продукции и основными факторами производства, функцию спроса.

Если предварительный анализ показывает, что искомая зависимость нелинейна, то выполняются следующие шаги.

Производится попытка так преобразовать анализируемые переменные y , x_1 , x_2 , ..., x_p , чтобы искомая зависимость была представлена в виде линейного

соотношения между преобразованными переменными. Такой процесс называют процедурой линеаризации модели, а сами модели называются внутренне линейными.

Если же линеаризующее преобразование подобрать не удастся, то есть модель внутренне нелинейна, то в этом случае требуется привлечение соответствующих методов нелинейной оптимизации.

Дисперсионный анализ

Результаты экспериментов и испытаний могут зависеть от некоторых факторов, влияющих на изменчивость средних значений случайной величины. Значения факторов называются уровнями факторов, а значение-атрибутом производительности. Например, объем работ, выполняемых на строительной площадке, может зависеть от рабочей команды. В этом случае количество бригад-это уровень фактора, а объем работы за смену-показатель эффективности.

Метод дисперсионного анализа, или *ANOVA* (Analysis of Variance – дисперсионный анализ), предназначается для исследования статистической значимости различия между средними при трех и более выборках (уровнях фактора). Для сравнения средних в двух выборках применяется *t*-критерий.

Процедура сравнения средних называется дисперсионным анализом, так как при исследовании статистической значимости различия между средними нескольких групп наблюдений проводится анализ выборочных дисперсий. Фундаментальная концепция дисперсионного анализа была предложена Фишером.

Суть метода состоит в том, чтобы разделить общую дисперсию на две части, одна из которых обусловлена случайной ошибкой (то есть внутригрупповой изменчивостью), а вторая связана с разницей средних значений. Последний компонент дисперсии затем используется для анализа статистической значимости разницы между средними значениями. Если эта разница существенна, то нулевая гипотеза отвергается и принимается альтернативная гипотеза о существовании разницы между средними значениями.

Переменные, значения которых определяются измерениями в ходе эксперимента (например, экономическая эффективность, урожайность, результат испытаний), называются зависимыми переменными или атрибутами. Переменные, которыми можно управлять в ходе эксперимента (например, уровень управления, тип почвы, методы обучения), называются факторами или независимыми переменными.

При классическом дисперсионном анализе предполагается, что исследуемые величины имеют нормальное распределение с постоянной дисперсией и средними значениями, которые могут отличаться для разных выборочных популяций. В качестве критерия проверки нулевых гипотез используется отношение дисперсии групповых средних и остаточной дисперсии. Однако было показано, что дисперсионный анализ справедлив и для негауссовых случайных величин, а при размере выборки для каждого уровня фактора $n > 4$ погрешность невелика. Если требуется высокая точность вывода и распределение неизвестно, то следует использовать непараметрические критерии, такие как ранговый дисперсионный анализ.

Методы классификации

Под классификацией обычно понимают "деление рассматриваемой совокупности объектов или явлений на однородные в известном смысле группы или отнесение каждого из данной совокупности объектов или явлений к одному из ранее известных классов". Термин классификация используется как для обозначения самого процесса, так и его результата.

Проблемы классификации связаны с проблемами многомерного статистического анализа. Они описывают объекты как векторы в пространстве признаков, характеризующих эти объекты. Задачи классификации делятся на три принципиально различных типа: дискриминантный анализ, кластерный анализ и задачи группировки.

Задача дискриминантного анализа состоит в том, чтобы найти правило, по которому наблюдаемый объект может быть отнесен к одному из указанных классов. Предполагается, что классификация объектов уже известна. Классы могут быть описаны либо непосредственно в математических терминах, либо с использованием обучающих выборок. В обучающих выборках значения объектов и класс, к которому принадлежит данный объект, регистрируются для каждого объекта.

Поскольку классификация объектов известна, методы дискриминантного анализа часто называют классификацией учителя.

Кластерный анализ используется для выделения однородных категорий объектов, когда на основании статистических данных необходимо разделить элементы выборки на группы близких между собой объектов, которые обычно называют кластерами. «Близость» объектов должна обуславливаться совокупностью значений измеренных признаков. Разбиение объектов на кластеры выполняется одновременно с формированием этих кластеров. В связи с тем, что классы и их число, а также сущность классификации заранее

не известны, кластерный анализ называют классификацией без учителя.

В задачах группировки классы заранее не заданы и не требуется, чтобы они были существенно различными. В качестве примера можно привести группировки по уровню дохода, числу детей в семье, номеру группы студента.

Целью как кластеризации, так и группировки является обнаружение и выделение классов. Однако задачей кластерного анализа является поиск достаточно изолированных кластеров объектов (кластеров) на основе имеющихся наблюдений, т. е. выявление "естественного" деления на классы. При группировании объектов они делятся на группы независимо от того, являются ли границы разделения естественными или нет, поэтому соседние группы могут незначительно отличаться друг от друга.

Хотя одни и те же методы могут быть использованы для решения задач кластеризации и группировки, принципиальное различие между ними заключается в том, что на практике кластеризация не всегда может быть успешно выполнена при анализе конкретных данных, если данные достаточно однородны, и в этом случае возможна только группировка.

Если в исследовании имеются как классифицированные данные, так и обучающие выборки на "входе" задачи, то решается задача классификации с обучающими выборками ("классификация с учителем"). Если нет обучающих выборок, то речь идет о проблеме "классификации без учителя".

Выходной результат может быть двух видов:

1. Если число классов и их сущность заранее определены, то каждому из классифицируемых объектов должен быть присвоен номер класса, к которому он принадлежит.

2. Если число классов и их сущность определяются в процессе классификации, то результатом процедуры классификации будет разбиение всей совокупности объектов на определенное число однородных групп (классов).

В первом случае классификация проводится при наличии обучающих выборок и является задачей дискриминантного анализа. Во втором случае имеет место решение задач кластеризации или группировки.

Дискриминантный анализ

Используя дискриминантный анализ, объект может быть отнесен к одной из нескольких predeterminedных групп на основе определенных характеристик.

Линейный дискриминантный анализ базируется на предположении, что данные подчиняются многомерному нормальному закону. В

дискриминантном анализе рассматриваются две задачи:

1) устанавливается правило, в соответствии с которым объект относится к одному из известных классов. Обычно это линейная функция от признаков – функция классификации (или дискриминантная функция):

$$S = b_1x_1 + b_2x_2 + \dots + b_nx_n + a,$$

где x_1, \dots, x_n — значения признаков, соответствующих рассматриваемым объектам; константы b_1, \dots, b_n и a — коэффициенты, которые должны быть оценены с помощью дискриминантного анализа. Целью анализа является определение таких коэффициентов, чтобы по значениям дискриминантной функции можно было с максимальной четкостью провести разделение объектов по группам.

2) по установленному правилу классифицируются новые объекты. Объект относится к классу i , если значение функции классификации S_i наибольшее. Значимость различения объектов оценивается с помощью дисперсионного анализа.

В дискриминантном анализе различие между классами определяется средним значением некоторой переменной или линейной комбинацией переменных, и эта переменная используется для предсказания класса для новых объектов. В связи с тем, что рассматриваются средние значения признаков, при анализе дискриминантной функции используется дисперсионный анализ. Для определения того, какие переменные способствуют различению между группами (классами), вычисляются матрицы общих дисперсий и ковариаций, матрицы внутригрупповых дисперсий и ковариаций. Для определения значимости различий между группами матрицы общих и внутригрупповых дисперсий сравниваются с помощью многомерного F - критерия.

Кластерный анализ

Разберем общую постановку задачи автоматической классификации. В этом случае исследователь не располагает обучающими выборками и имеет только n подлежащих классификации наблюдений, заданных матрицей X .

В общей постановке задача автоматической классификации объектов заключается в разбиении анализируемой совокупности объектов $O = \{O_i\}$ ($i = 1, \dots, n$), статистически представленной в виде матрицы X , на сравнительно небольшое число однородных, в определенном смысле, групп или классов.

Анализируемые объекты могут быть представлены в виде точек в пространстве объектов. Геометрическая близость некоторых точек в этом

пространстве означает близость соответствующих объектов, их однородность. Задача классификации может быть сформулирована как задача нахождения классов объектов, находящихся внутри каждого класса на относительно малых расстояниях друг от друга.

В задаче кластеризации важным шагом является определение понятия близости или однородности объектов. В качестве меры сходства могут быть выбраны коэффициент корреляции, мера расстояния, вероятностная мера сходства. Если задана функция расстояния $d(O_i, O_j)$, то близкие в смысле этой метрики объекты считаются однородными и принадлежащими к одному классу. При этом необходимо сопоставление $d(O_i, O_j)$ с некоторым пороговым значением, определяемым в зависимости от содержательной интерпретации данных.

Методы кластерного анализа имеют эвристический характер. Поэтому важным этапом исследования является выбор наиболее подходящей метрики.

Для переменных, относящихся к интервальной шкале (метрические переменные), основными мерами расстояния являются обычное евклидово расстояние и взвешенное евклидово расстояние.

При реализации кластерных процедур, помимо расчета расстояний между отдельными объектами, вычисляются расстояния между кластерами, сформированными на определенном шаге процедуры.

В зависимости от общего количества классифицируемых объектов все задачи кластерного анализа делятся на два типа. Первый тип включает в себя классификационные задачи для небольшого числа наблюдений, которые включают в себя несколько десятков наблюдений. Ко второму типу относятся задачи классификации больших массивов многомерных наблюдений, включающих сотни и тысячи наблюдений. Для этих двух типов задач используются различные типы кластерных процедур.

Выделяются три основных типа кластер-процедур:

- процедуры иерархические, основанные на пересчете элементов матрицы расстояний $\rho(X_i, X_j)$ и для больших выборок требующие значительных вычислительных мощностей, поэтому они обычно используются в задачах с небольшим числом объектов;
- процедуры параллельные, реализующиеся с помощью итерационных алгоритмов, на каждом шаге которых происходит одновременный расчет всех имеющихся объектов. Обычно используются в задачах с небольшим числом объектов;
- процедуры последовательные, на каждой итерации которых

производится расчет лишь по небольшой части наблюдений. Последовательные процедуры быстро сходятся и могут быть использованы для задач с большим числом объектов.

Иерархические процедуры кластерного анализа

Иерархические процедуры подразделяются на агломеративные и дивизимные.

Основной алгоритм иерархических процедур заключается в последовательном объединении (расщеплении) групп элементов. В агломеративных процедурах предполагается, что первоначально каждый объект является отдельным кластером, и процесс состоит из объединения отдельных объектов, в то время как в разделительных процедурах весь набор объектов изначально является одним кластером, и процесс состоит из разделения объектов.

Иерархические алгоритмы строят полное дерево вложенных кластеров, но число кластеров не определяется самим алгоритмом. Количество кластеров должно определяться исследователем на основе предположений, не связанных с самим алгоритмом: на основе порога слияния кластеров или на основе содержания данных.

Иерархические процедуры, по сравнению с другими кластерными процедурами, обеспечивают более полный и тонкий анализ структуры исследуемых объектов и имеют визуальную интерпретацию в виде дендрограммы. Однако на каждом шаге иерархического алгоритма вычисляется вся матрица расстояний, что требует больших вычислительных мощностей, поэтому алгоритмы используются только для небольшого числа объектов (не более нескольких сотен).

5.5. Методы прогнозирования в научных исследованиях

Прогнозирование в информатике является одной из основных составляющих исследовательского процесса. Правильный выбор методов прогноза, оценки достоверности и интерпретации результатов прогноза требует знания основных положений методологии прогнозирования, возможностей и характерных особенностей методов и способов прогнозирования.

В современном понимании, по мнению Г. Тейла, прогноз-это "определенное суждение о неизвестных, особенно будущих событиях". Сущность прогнозирования заключается в обработке исследователем

имеющейся у него информации о состоянии исследуемого объекта в предыдущий и настоящий моменты и формировании с определенной степенью достоверности системы знаний о будущем состоянии или поведении объекта. Прогнозирование позволяет выявить устойчивые тенденции или выявить существенные изменения социально-экономических процессов, оценить их вероятность на будущий плановый период, выявить возможные альтернативы и накопить научно-эмпирический материал для обоснованного выбора той или иной концепции развития или управленческого решения.

Результатом прогнозирования является прогноз — научно-обоснованное суждение о возможных состояниях объекта в будущем, об альтернативных путях и сроках их осуществления.

В зависимости от целей и задач прогнозов выделяют поисковый и нормативно-целевой прогнозы.

Поисковый прогноз (в литературе также используются термины опрос, исследование, тренд, генетический и др.) - Это определение возможных состояний процесса или явления (объекта исследования) в будущем. При таком прогнозе происходит условное продолжение прошлых тенденций в настоящее и будущее. Поисковое прогнозирование может быть двух видов: традиционное (экстраполяционное) и инновационное (альтернативное). Традиционное прогнозирование основано на предположении, что развитие объекта происходит плавно и непрерывно, т. е. все тенденции, выявленные в прошлом, будут сохранены. Если такой прогноз не учитывает влияние различных факторов на исследуемый объект, а использует зависимость показателей только от времени, то такой прогноз называется "наивным".

Новаторский подход основан на предположении, что исследуемый процесс или явление неоднородны и изменчивы вследствие воздействия различных внешних возмущающих воздействий. Этот подход также называют альтернативным подходом, поскольку он предоставляет множество вариантов развития исследуемого процесса или явления.

Нормативно-целевой прогноз ориентирован на определение путей и сроков достижения цели, за которую принимается возможное состояние объекта или процесса. Если поисковый прогноз основывается на состоянии объекта в прошлом и настоящем, то стандартный целевой прогноз проводится в обратном порядке от заданного состояния в будущем, исходя из наличия материальных, финансовых и иных ресурсов, необходимых для достижения поставленных целей, к существующим тенденциям. Данный вид прогнозирования применяется в тех случаях, когда информации о прошлом и

настоящем объекта исследования недостаточно и невозможно определить тенденции его развития. В этом случае исследователи определяют желаемое конечное состояние объекта, ресурсы, необходимые для перевода объекта из текущего состояния в желаемое, и действия, направленные на осуществление этого перехода.

В зависимости от периода прогнозирования различают *оперативные, краткосрочные, долгосрочные, среднесрочные, дальнесрочные прогнозы*.

При классификации по критерию объекта прогнозирования выделяют *естествоведческие, обществоведческие и научно-технические прогнозы*.

Естественнонаучные прогнозы включают в себя прогнозы явлений или процессов, на которые не может повлиять социальное управление. Примеры включают метеорологические, гидрологические, геологические, биологические, медико - биологические, космологические и физико-химические прогнозы.

Обществоведческие прогнозы осуществляются для явлений, на которые допустимо влияние средств социального управления. К этой группе относятся: социально-медицинские, социально-географические, социально-экологические, социально-космические (перспектив освоения космоса), экономические, социологические, психологические прогнозы (личности, ее поведения, деятельности), демографические, образовательно-педагогические, культурно-эстетические, государственно-правовые, внутривластные, внешнеполитические, военные и другие прогнозы.

Научно-технические прогнозы в узком смысле (технологические, инженерные) относятся к перспективам состояний всех явлений техносферы (материалов, механизмов, приборов и т. п.). В широком смысле они охватывают перспективы развития научно-технического прогресса: науки (ее структуры, кадров, учреждений и т.д.) и техники (в промышленности, строительстве, сельском хозяйстве, связи и т.д.). Одним из наиболее важных классификационных признаков является степень формализации, которая достаточно полно охватывает прогностические методы.

По степени формализации методы экономического прогнозирования можно разделить на экспертные и формализованные. Экспертные (интуитивные) методы прогнозирования применяются в тех случаях, когда невозможно учесть влияние многих факторов из-за значительной сложности объекта прогнозирования. В этом случае используются экспертные оценки. При этом различают индивидуальные и коллективные экспертные оценки.

Индивидуальные экспертные оценки включают в себя: метод

"интервью", предполагающий непосредственный контакт эксперта и специалиста по схеме "вопрос-ответ"; аналитический метод, предусматривающий логический анализ прогнозируемой ситуации, составление аналитических отчетов; метод написания сценария, основанный на определении логики процесса или явления во времени при различных условиях.

Методы коллективных экспертных оценок включают в себя метод «комиссий», «коллективной генерации идей» («мозговая атака»), метод «Дельфи», матричный метод.

Методы экспертной оценки применяются в тех случаях, когда имеется достаточная информация для решения поставленной задачи, но она носит качественный характер или отсутствует достаточная информация для решения поставленной задачи. Данная группа методов основана на том, что при коллективном мышлении, во-первых, точность результата выше, а во-вторых, при обработке индивидуальных независимых оценок, сделанных экспертами, могут возникать продуктивные идеи. Однако методы экспертной оценки также имеют ряд недостатков. Прежде всего, они включают в себя возможность давления на группу экспертов со стороны более активного члена группы экспертов. Кроме того, экспертная оценка требует получения последовательного результата без потери качества прогноза, что не всегда возможно.

В тех случаях, когда удастся формализовать взаимосвязи между основными показателями исследуемых процессов и явлений, применяют формализованные методы прогнозирования. Формализованные методы прогнозирования основаны на статистических методах и методах моделирования объекта исследования с целью выявления устойчивых тенденций его развития.

Статистические методы прогнозирования

Методы статистического прогнозирования охватывают разработку, изучение и применение современных математико-статистических методов прогнозирования на основе объективных данных. Научной основой статистических методов прогнозирования является прикладная статистика и теория принятия решений. Эти методы в настоящее время широко используются для решения плановых, производственных и маркетинговых задач в компаниях различных размеров и форм собственности.

В основе статистических методов прогнозирования лежат данные, представленные в виде временных рядов.

Временной ряд– это последовательность измерений в последовательные моменты времени. Временной ряд определяется как ряд наблюдений $x(t_1), x(t_2), \dots, x(t_N)$ анализируемой случайной величины $f(Y)$, произведенных в последовательные моменты времени t_1, t_2, \dots, t_n .

В большинстве случаев предметом анализа являются временные ряды с равноотстоящими моментами наблюдений, что позволяет представить их в виде:

$$x(1), x(2), \dots, x(N),$$

где $x(t)$ – значение анализируемого показателя, зарегистрированное в t -м такте времени ($t = 1, 2, \dots, N$).

Анализ временных рядов включает широкий спектр разведочных процедур и исследовательских методов, которые решают три основные задачи:

1. Определение природы временного ряда.
2. Предсказание будущих значений временного ряда по настоящим и прошлым значениям (прогнозирование).
3. Выработка стратегии управления и оптимизации анализируемых процессов.

При рассмотрении последовательностей наблюдений значение приобретает их упорядоченность. Если значения некоторого показателя ряда соответствуют некоторому признаку, значения которого расположены в порядке убывания или возрастания, то говорят о динамическом ряде или ряде динамики. Для отсортированного признака в виде времени последовательность показателя называют временным рядом. Примером динамического ряда могут быть средние урожайности по муниципальным образованиям, каждому из которых присваивается упорядоченный номер. Изменение прибыли в хозяйстве за многолетний период является примером временного ряда.

При обработке динамических рядов экономических показателей возникает необходимость выявления:

- 1) тренда развития процесса;
- 2) наличия циклов;
- 3) сезонных изменений;
- 4) случайных колебаний.

Тренд динамического ряда — это основная направленность развития

явления. Для ее выявления используются различные методы: укрупнение интервалов; механическое сглаживание; аналитическое выравнивание; формулы сглаживания Спенсера; избранные точки, гармонический анализ и др.

На предварительном этапе в последовательности значений определяю аномальные величины. Аномальное значение не отвечает потенциальным возможностям ряда, значительно отклоняется от других значений, оказывая существенное влияние на характеристики временного ряда. Аномальные значения связывают с техническими ошибками и объективными обстоятельствами их появления. В первом случае говорят об ошибке первого рода, а во втором – об ошибке второго рода.

Для определения аномальных значений можно использовать метод Ирвина. Согласно этому методу находятся эмпирические значения λ_t по формуле

$$\lambda_t = \frac{|y_t - y_{t-1}|}{\sigma_y}, \quad t = 1, 2, 3, \dots, n, \quad (5.5)$$

где среднеквадратическое отклонение рассчитывается с использованием формул:

$$\sigma_y = \sqrt{\frac{\sum_{t=1}^n (y_t - \bar{y})^2}{n-1}}; \quad (5.6)$$

$$\bar{y} = \frac{\sum_{t=1}^n y_t}{n}. \quad (5.7)$$

Расчетные значения λ_2, λ_3 и т.д. сравниваются с табличными значениями критерия Ирвина λ_α , (таблица 5.2) и если оказываются больше табличных, то соответствующее значение y_t уровня ряда считается аномальным. В таблице 15 приведены значения критерия Ирвина при уровне значимости $\alpha = 5\%$:

После проверки ряда на наличие аномальных значений выясняется их принадлежность к ошибке первого или второго рода. В первом случае (техническая ошибка) значение ряда восстанавливают по уравнению тренда или заменой аномального значения средней арифметической величиной между соседними ординатами ряда. В отличие от ошибок первого рода,

аномальные значения как ошибки второго рода не устранимы.

Таблица 5.2 – Значения критерия Ирвина при уровне значимости $\alpha = 5\%$

n	2	3	10	20	30	50	100
λ	2,	2,	1,	1,	1,	1,	1,0
α	8	3	5	3	2	1	

Проанализировав аномальные явления, переходят к выявлению трендов. В разделе рассмотрены варианты определения тенденций на основе механического и аналитического сглаживания. При механическом сглаживании используют соседние значения рядов. Одним из вариантов механического сглаживания является метод скользящей средней. Во втором случае распространен метод экстраполяции с применением кривых роста.

Пример. Выявить аномальность значений временного ряда урожайности по данным таблицы 5.3.

Таблица 5.3 – Данные для оценки аномального явления

Год	1992	1993	1994	1995	1996	1997	1998	1999
у/га	20	18	33	16	21	15	20	18

Решение

Воспользуемся методом Ирвина (5.5). По формулам (5.6) и (5.7) находим среднее квадратическое отклонение ряда $\sigma_y = 5,59$ и среднее значение $\bar{y} = 20,12$.

Приведем расчетные значения критерия Ирвина:

$\lambda_2 = |18-20| / 5,59 = 0,36$; $\lambda_3 = 2,68$; $\lambda_4 = 3,04$; $\lambda_5 = 0,89$; $\lambda_6 = 1,07$; $\lambda_7 = 0,89$; $\lambda_8 = 0,36$.

Затем сравниваем полученные значения с табличной величиной (таблица 5.2). Поскольку длина ряда $n=8$ методом интерполяции определяется табличное значение $\lambda_\alpha = 1,73$. В двух случаях табличное значение оказалось меньше расчетной величины: $\lambda_3 > \lambda_\alpha$ ($2,68 > 1,73$) и $\lambda_4 > \lambda_\alpha$ ($3,04 > 1,73$). Следовательно, показатель урожайности 1994 года является аномальными уровнем временного ряда.

Скользящая средняя

В статистике различают методы невзвешенной и взвешенной скользящей средней. Метод сглаживания невзвешенной скользящей средней состоит в том, что фактические значения (уровни) последовательности заменяются скользящими средними, рассчитанными из нескольких соседних значений ряда. Период усреднения для динамического ряда должен быть одинаковым, сдвигаясь на единицу от начального уровня. Из периода постоянно исключаются первые уровни и включаются последующие.

При использовании метода невзвешенной скользящей средней период сглаживания может включать нечетное или четное число членов ряда. При нечетном числе уровней на отрезке усреднения, найденные скользящие средние являются центрированными и относятся к середине периода скольжения. Например, при трехлетнем периоде усреднения схема расчета имеет вид

$$\begin{aligned}\bar{y}_2 &= \frac{y_1 + y_2 + y_3}{3}; & \bar{y}_3 &= \frac{y_2 + y_3 + y_4}{3}; \\ \bar{y}_4 &= \frac{y_3 + y_4 + y_5}{3}; & \bar{y}_{n-1} &= \frac{y_{n-2} + y_{n-1} + y_n}{3},\end{aligned}\quad (5.8)$$

где y_i - значение ряда, \bar{y}_{i+1} - усредненные величины, n - длина ряда.

При четном числе членов ряда выбранный интервал усреднения должен соответствовать середине периода. В этом случае вычисляются центрированные скользящие средние. Они рассчитываются путем определения среднего каждой пары значений скользящих средних.

При использовании метода скользящих средних всегда возникает необходимость правильного выбора периода скольжения для погашения случайного отклонения уровней.

Если в развитии явления отмечается определённая цикличность, то интервал скольжения следует брать равным продолжительности цикла. Если же колебания уровней беспорядочные, то целесообразно постепенно укрупнять интервал сглаживания, пока не выявится отчетливая картина тренда.

Метод скользящих средних имеет достоинства и недостатки.

Достоинства метода состоят в том, что расчет скользящих средних по периодам позволяет не только выявить тенденцию, но и отразить ход процесса развития в виде прогностических значений.

Вместе с тем методу скользящих средних присущи недостатки:

- 1) уменьшение информативности ряда;
- 2) невозможность расчета показателей динамики;
- 3) ограниченность применения для динамических рядов со сложными циклическими процессами.

При применении метода взвешенной скользящей средней для определения весов многократно повторяется сглаживание скользящей средней с двухчленным периодом.

Первые скользящие средние находят по формулам:

$$\frac{y_1 + y_2}{2}; \quad \frac{y_2 + y_3}{2}; \quad (5.9)$$

и т.д.

Используя найденные значения, получают вторые скользящие средние для трехчленного периода:

$$\frac{\frac{y_1 + y_2}{2} + \frac{y_2 + y_3}{2}}{2} = \frac{y_1 + 2y_2 + y_3}{4}; \quad \frac{\frac{y_2 + y_3}{2} + \frac{y_3 + y_4}{2}}{2} = \frac{y_2 + 2y_3 + y_4}{4}$$

(5.10) и т.д.

На их основе находят третьи скользящие средние для четырехчленного; периода:

$$\frac{\frac{y_1 + 2y_2 + y_3}{4} + \frac{y_2 + 2y_3 + y_4}{4}}{2} = \frac{y_1 + 2y_2 + 3y_3 + y_4}{8}, \quad (5.11)$$

а далее - четвертые скользящие средние для пятилетнего периода:

$$\frac{y_1 + 4y_2 + 6y_3 + 4y_4 + y_5}{16}. \quad (5.12)$$

Веса уровней при сглаживании образуют коэффициенты бинома Ньютона

		1	2	1		
		1	3	3	1	
	1	4	6	4	1	
1	5	10	10	5	1	

и т.д.

Применение метода скользящей средней позволяет определять тенденции временной изменчивости ряда, выявлять циклы и осуществлять прогнозирование.

Пример. По данным первых двух граф таблица 5.4 рассчитать

скользящие средние за четырехлетний период и центрировать их.

Решение

Используя формулу (5.6) применительно к четырем соседним значениям ряда, нетрудно рассчитать скользящие средние (третья графа). В частности, первое число получено как $(11,4+15,8+14,4+16,7)/4=14,6$, а последнее – $(21,9+18,6+25,9+22,0)/4=22,1$. Центрированные скользящие средние находятся по данным предыдущей графы. Например, $(14,6+17)/2=15,9$, $(17+18,2)/2=17,7$ и т.д. В таблице 5.4 приведены итоги расчетов.

Таблица 5.4 - Динамика объема реализации продукции, млн. руб.

Год	Объем реализации продукции	Скользящие средние (4-летний период)	Центрированные скользящие средние
1989	11,4		
1990	15,8	14,6	15,9
1991	14,4		
1992	16,7	17	17,7
1993	21,7	18,2	18,8
1994	20,2	19,3	20,4
1995	18,5	21,6	21,6
1996	25,8	21,6	21,4
1997	21,9	21,2	22,1
1998	18,6	23,0	22,6
1999	25,9	22,1	
2000	22,0		

Модели роста

Трендовые модели описывают изменения показателя экономического процесса во времени. При этом предполагается, что в последовательности имеет место устойчивая тенденция, которая не изменится в будущем. При построении трендовых моделей могут использоваться кривые роста.

Задача ставится следующим образом. Зная формулу кривой роста и значения временного ряда, на основе некоторого критерия определяют параметры формулы кривой роста.

При использовании линейного тренда

$$\hat{y}_i = at_i + b + \varepsilon \quad (5.13)$$

вычисляются его параметры a и b . В этой формуле a – средняя величина изменения уровней ряда за единицу изменения времени; b – свободный член уравнения; t_i – номера моментов или периодов времени, к которым относятся уровни временного ряда (год, квартал, месяц, дата); ε – остаток ряда.

Часто в качестве критерия применяют метод наименьших квадратов:

$$S = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \rightarrow \min, \quad (5.14)$$

где y_i – значения ряда, \hat{y}_i – уровни, полученные с помощью аналитического выражения.

Применение метода наименьших квадратов для линейной зависимости (41) приводит в конечном итоге к следующим результатам:

$$a = r \frac{\sigma_y}{\sigma_t}, \quad (5.15)$$

$$b = \bar{y} + a\bar{t}, \quad (5.16)$$

$$r = \frac{\sum_{i=1}^n (y_i - \bar{y})(t_i - \bar{t})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2 \sum_{i=1}^n (t_i - \bar{t})^2}}, \quad (5.17)$$

где r – коэффициент корреляции между переменными; \bar{y}, \bar{t} – средние значения рядов y_i и t_i , а σ_y, σ_t – их средние квадратические отклонения.

Помимо линейной функции роста применяют нелинейные выражения. В частности, распространение нашел полином с числом параметров m :

$$y = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + \dots + a_i t^{i1} + a_m t^m, \quad (5.18)$$

где a_0 – свободный член, a_1, a_3, \dots, a_m – коэффициенты формулы, $i = 1, 2, \dots, m$.

Нетрудно показать, что из выражения (5.18) при $m=1$ можно получить линейную зависимость, а в случае $m=2$ – параболу второй степени. Как правило, эти две зависимости наиболее часто используют при выделении тренда.

Кроме полинома, применяют степенные, показательные и логарифмические функции роста. Перечисленные выражения имеют недостаток, который состоит в том, что при возрастании времени в

бесконечность $t \rightarrow \infty$, значение функции увеличивается без ограничения, $y \rightarrow \infty$.

Во многих случаях производственный процесс имеет некоторое предельное значение. В частности, потенциал получения урожайности в той или иной физико-географической зоне, среднегодовой надой молока, среднесуточный привес, настриг шерсти ограничены некоторой верхней границей.

Рассмотренные тенденции производственно-экономических характеристик описывают с помощью логистических кривых или кривых асимптотического роста.

Простейшая логистическая кривая рассматривается в виде следующего дифференциального уравнения

$$\frac{dy}{dt} = K_1 y(A - y), \quad (5.19)$$

где A – уровень насыщения, $K_1 y$ – фактор насыщения, $(A - y)$ – фактор торможения, K_1 – параметр, y – характеристика.

Уравнение изменения во времени y в этом случае можно записать в виде

$$y = \frac{A}{1 + b e^{-AK_1(t-t_0)}}. \quad (5.20)$$

Математический закон асимптотического роста выражают уравнением

$$\frac{dy}{dt} = k(A - y), \quad (5.21)$$

где k – коэффициент роста, A – уровень насыщения некоторой характеристики y за время t .

Уровень изменения y во времени можно записать следующим образом

$$y = A - (A - y_0) e^{-k(t-t_0)}, \quad (5.22)$$

где t_0 и y_0 – начальное и граничное условия.

При составлении перечисленных функций немаловажное значение имеет определение технологических предельных значений производственных процессов.

Предельные значения могут быть получены в виде оценки наибольших (наименьших) реальных предельных границ или на основе определения возможного потенциала системы.

Принимая во внимание то, что сельскохозяйственные предприятия стремятся получать наибольшую прибыль за наименьший промежуток времени, необходимо сокращать периоды достижения предельных значений.

Достигнуть физического или производственного предела возможно за счет варьирования параметра K .

Пример. По данным таблицы 5.3 построить линейный тренд.

Решение

Для определения функциональной зависимости между урожайностью и годами воспользуемся формулами (5.15)-(5.17). В результате получим: $\bar{y} = 20,12$, $\bar{t} = 1995,5$, $\sigma_y = 5,59$, $\sigma_t = 2,45$, $r = -0,28$, $a = -0,63$, $b = 1279$. Таким образом, линейная функция примет вид $y = -0,63t + 1279$.

Модели прогнозирования и оценки устойчивости экономических процессов

Наличие устойчивой тенденции изменения экономического показателя во времени (тренда) позволяет отобразить ее в виде математического выражения. При сохранении выявленного развития можно прогнозировать значения экономического параметра с некоторым упреждением или заблаговременностью. Таким образом, прогнозирование с привлечением трендовой модели – это экстраполяция тенденции в будущее.

В прогнозировании экономических показателей применяют точечные и интервальные оценки. Точечный прогноз представляет собой единственное значение прогнозируемого показателя. В интервальном прогнозе указывается вероятный диапазон изменчивости прогностической величины.

Прогнозирование имеет смысл, если модель качественная. Качество модели определяется адекватностью и точностью. Модель считается адекватной при условии, что остаток ряда (разность между эмпирическими и аналитическими значениями $\varepsilon_i = y_i - \hat{y}_i$) является случайной величиной. Для выполнения этого требования необходимо, чтобы выполнялись следующие условия: колебания значений остатка ряда были случайными; ряд остатка соответствовал нормальному закону распределения; среднее значение последовательности ε_i равнялось бы нулю; значения ряда остатка являлись бы независимыми.

Проверка случайности колебаний уровней остаточной последовательности означает проверку гипотезы о правильности выбора тренда. Характер этих отклонений изучается с помощью ряда непараметрических критериев, одним из которых является критерий серий, основанный на медиане выборки.

Ряд из величин ε_i располагают в порядке возрастания их значений и находят медиану ε_m полученного вариационного ряда. Медиана - срединное значение при нечетном n или среднее арифметическое из двух срединных значений при n четном. Возвращаясь к исходной последовательности ε_i и

сравнивая значения этой последовательности с ε_m , ставится знак «+», если значение ε_t превосходит медиану, и знак «-», если оно меньше медианы. В случае равенства сравниваемых величин соответствующее значение ε_t опускается. Таким образом, получается последовательность, состоящая из плюсов и минусов, общее число которых не превосходит n . Последовательность подряд идущих плюсов или минусов называется *серией*.

Протяженность самой длинной серии обозначается через K_{\max} , а общее число серий – через V . Выборка признается случайной, если выполняются следующие неравенства для уровня значимости 5%:

$$K_{\max} \leq [3,3(\lg n + 1)]; \quad (5.23)$$

$$V \leq \left[\frac{1}{2}n + 1 - 1,96\sqrt{n-1} \right]. \quad (5.24)$$

Вторым требованием к адекватности модели является соответствие случайной компоненты нормальному закону распределения. Поскольку нормальный закон распределения вероятностей является симметричным, то коэффициент асимметрии c_s должен быть близок к нулю или незначимым. Поэтому вначале вычисляют коэффициент асимметрии ряда остатка ε_i , а затем определяют его среднюю квадратическую погрешность:

$$\sigma_{c_s} = \sqrt{\frac{6(n-2)}{(n+1)(n+3)}}, \quad (5.25)$$

где n – длина ряда.

Если выполняется неравенство $|c_s| \leq 1,5\sigma_{c_s}$, тогда считается, что коэффициент асимметрии незначим и принимается гипотеза о соответствии значений остатка нормальному распределению.

Для проверки выполнения третьего условия, соответствия нулю среднего значения остаточной последовательности ε_i , определяется среднее арифметическое ряда остатка и его средняя квадратическая погрешность.

Если их соотношение меньше табличного значения, т.е. $\frac{\bar{\varepsilon}}{\sigma_{\varepsilon}} \leq t_{\alpha}$, тогда

принимается утверждение о соответствии нулю ряда остатка. Табличное значение критерия Стьюдента t_{α} зависит от уровня значимости α и числа степеней свободы $n-1$. Этот критерий определяется с помощью математических справочников или функций, заложенных в различных программных продуктах. В частности, критерий t_{α} можно определить в приложении Excel.

Последним требованием адекватности модели является отсутствие внутрирядных связей между значениями остаточной последовательности. Для решения этой задачи определяется первый коэффициент автокорреляции исследуемого ряда и его средняя квадратическая погрешность. При коэффициенте автокорреляции близком к нулю и преобладании средней квадратической погрешности над ним принимается, что значения ряда остатка независимы.

Только при выполнении перечисленных четырех требований модель считается адекватной. Добавим к этому, что существуют и другие способы оценки адекватности модели, с которыми можно познакомиться в рекомендуемой литературе.

Для оценки точности модели применяют коэффициент детерминации r^2 , среднее квадратическое отклонение аналитических значений от эмпирических величин и среднюю относительную ошибку аппроксимации.

Коэффициент детерминации представляет собой квадрат коэффициента корреляции, рассчитанного по формуле (5.17). Среднее квадратическое отклонение аналитических значений от эмпирических величин находится по формуле (5.21). И, наконец, средняя относительная ошибка аппроксимации имеет вид

$$\delta = \frac{1}{n} \sum_{i=1}^n \left| \frac{y - y_i}{y} \right| \times 100\% , \quad (5.26)$$

где y, y_i – эмпирические и модельные значения

При использовании линейного тренда точечный прогноз определяется по уравнению (5.13). Вместо t_i подставляют значение с учетом упреждения. Если это прогноз, например, на 2009 г. по данным до 2008 г., то в качестве значения используется число 2009.

Для расчета доверительного интервала V можно использовать формулу

$$V = t_{\alpha} \sigma_y \sqrt{1 + \frac{1}{n} + \frac{3(n + 2k - 1)^2}{n(n^2 - 1)}} , \quad (5.27)$$

где t_{α} - значение критерия Стьюдента для уровня значимости α и числа степеней свободы $n-2$; n – длина ряда; k – период заблаговременности. Средняя квадратическая погрешность уравнения регрессии вычисляется по формуле

$$\sigma_y = \sqrt{\frac{\sum_{i=1}^n (y - y_i)^2}{n - m}} , \quad (5.28)$$

где m – число параметров модели.

Используя точечный прогноз и доверительный интервал, интервальное значение прогноза может быть представлено в следующей редакции

$$V_y = y_{t+k} \pm V, \quad (5.29)$$

где Y_{t+k} - точечный прогноз по уравнению (5.13) с заблаговременностью k . Чем меньше полученный интервал, тем точнее считается прогноз.

При рассмотрении будущих ситуаций для проверки модели имеет смысл использовать ретроспективный прогноз. В этом случае из n значений ряда делят на две последовательности, которые имеют длину n_1 и n_2 . По первой последовательности n_1 осуществляется точечный и интервальный прогноз, значения которого сравниваются с фактическими данными второго ряда длиной n_2 . Очевидно, чем ближе прогностические величины к фактическим данным, тем качественнее прогноз.

В заключении обратим внимание на некоторые недостатки прогнозирования с помощью трендовых моделей. Во-первых, хронологическая последовательность может содержать в себе неоднородные данные или несколько трендов. Во-вторых, тренд описывает инерционные процессы (те условия, которые были, не изменятся). И, наконец, они не определяют природу колебаний прогнозируемой характеристики. Тем не менее, тренды позволяют относительно несложно определять будущие ситуации. Следует отметить ограничение периода упреждения или заблаговременности прогноза k . Обычно эта величина не должна превышать третью часть длины последовательности, т.е. $n/3$.

Пример. Оценить случайность колебания уровней остаточной последовательности от тренда по данным таблицы 5.5.

Остаток оценивается как $\varepsilon_i = y_i - \hat{y}_i$ по критерию серий, где \hat{y}_i - значения, снятые с линии тренда, y_i – исходные данные.

Таблица 5.5 – Остаточная последовательность

ε_i	2,3	4,1	5,2	1,7	3,3	1,9	2,5
-----------------	-----	-----	-----	-----	-----	-----	-----

В данном примере, используя формулы правых частей неравенств (5.19) и (5.20), получаем $[3,3(\lg n + 1)] = 6$, $\left[\frac{1}{2}n + 1 - 1,96\sqrt{n-1} \right] = 1$.

Зная медиану $\varepsilon_m = 2,5$, определим серии (таблица 5.6)

Таблица 5.6 – Определение серий

ε_t	2, 3	4, 1	5, 2	1, 7	3, 3	1, 9	2, 5
«+» или «->»	-	+	+	-	+	-	

Поскольку $K_{max} = 2$ меньше, чем 6, а $V=5$ превышает 1, то гипотеза о случайном характере отклонений уровней временного ряда принимается и трендовая модель признается адекватной.

Пример. Привести интервальный прогноз демографического показателя сельского населения старше трудоспособного возраста (тыс. человек) с заблаговременностью 2 года, используя уравнения тренда $Y = 1,42t - 2753,5$, стандартную ошибку уравнения регрессии $\sigma_y = 2,5$, табличное значение критерия Стьюдента $t_\alpha = 1,32$ при $\alpha = 20\%$. Коэффициент детерминации равен 0,92. Продолжительность ряда составила 23 года, 1979-2001 гг.

Решение

Период временного ряда составляет 1979-2001 гг., следовательно, прогноз выполняется на 2002 и 2003 гг.

Подставляя в уравнение модели $Y = 1,42t - 2753,5$ значения t , получим точечные прогнозы демографического показателя на 2002 и 2003 годы:

$$\hat{Y}_{2002} = 1,42 \cdot 2002 - 2753,5 = 89,34;$$

$$\hat{Y}_{2003} = 1,42 \cdot 2003 - 2753,5 = 90,76.$$

Подставляя известные значения в формулу расчета доверительного интервала, получим

$$U_{2003} = 90,76 + 1,32 \cdot 2,5 \cdot \sqrt{1 + \frac{1}{23} + (3 \cdot (23 + 2 \cdot 2 - 1)^2) / (23 \cdot (23^2 - 1))} = 94,75$$

-

верхняя граница;

$$U_{2003} = 90,76 - 1,32 \cdot 2,5 \cdot \sqrt{1 + \frac{1}{23} + (3 \cdot (23 + 2 \cdot 2 - 1)^2) / (23 \cdot (23^2 - 1))} = 86,76$$

-
нижняя граница.

Таким образом, с доверительной вероятностью 80% (уровень значимости – 20%) можно утверждать, что при сохранении сложившихся закономерностей развития прогнозируемая величина демографического показателя в 2003 г. попадет в интервал от 86,76 до 94,75 тыс. человек.

5.6. Контрольные вопросы

1. Назовите основные различия между теоретическим и эмпирическим подходами в научном исследовании.
2. Какую роль играет метод формализации на пути развития экономической науки и науки управления?
3. Что понимается под термином «математическое моделирование»?
4. Дайте краткую характеристику категориям математических моделей.
5. Что вы понимаете под термином «имитационное моделирование»?
6. Дайте определение термину «прогноз».
7. Какие виды прогнозов вам известны?
8. Какие этапы исследования включает статистическая методология?
9. Дайте краткую характеристику методу статистического наблюдения. Каким основным требованиям должны отвечать статистические наблюдения?
10. С какой целью при обработке статистических данных применяется метод группировки?
11. В чем состоит отличие индивидуальных и сводных индексов?
12. Дайте определение и краткую характеристику таким статистическим показателям как «средняя величина», «мода», «медиана».
13. Что такое статистическая связь?
14. Каковы основные характеристики парной статистической связи?
15. Что является целью регрессионного анализа?
16. Как интерпретируется коэффициент детерминации?
17. Каковы свойства коэффициентов множественной регрессии?
18. Каково назначение дисперсионного анализа?

19. Поясните значение термина «классификация».
20. Какие методы классификации используются в кластерном анализе?
21. Какие виды прогнозов вам известны?
22. В каких случаях наиболее целесообразно применение экспертных методов прогнозирования?
23. Какие основные задачи позволяет решить анализ временных рядов?
24. Какие составляющие выделяют в систематической компоненте временного ряда?
25. С какой целью при исследовании временных рядов выполняется спектральный анализ?
26. Что вы понимаете под термином «сглаживание временного ряда»?
27. Какие существуют методы сглаживания временных рядов? Поясните основную суть названных вами методов.

5.7. Задания для самостоятельной работы

1. Определите, в каком количестве необходимо выпускать продукцию четырех типов Прод1, Прод2, Прод3, Прод4, для изготовления которой требуются ресурсы трех видов: трудовые, сырьевые, финансы. Норма расхода, а также прибыль, получаемая от реализации каждого вида продукции, приведены в таблице 5.3. Для решения разработайте математическую модель в соответствии с условиями задачи и выполните исследование средствами MS Excel.

Таблица 5.7 – Исходные данные

Ресурс	Прод1	Прод2	Прод3	Прод4	Знак	Наличие
Прибыль	60	70	120	130	max	
Трудовые	1	1	1	1	<=	16
Сырьевые	6	5	4	3	<=	110
Финансовы е	4	6	10	13	<=	100

2. Выполните научное исследование, направленное на оптимизацию процессов производства с целью получения максимальной прибыли при условии обеспечения продукцией всех заказчиков и отсутствии затоваривания складов. Предприятие производит продукцию трех видов, поставляет их

заказчикам и реализует на рынке. Заказчикам требуется 1000 изделий первого вида, 2000 изделий второго вида и 2500 изделий третьего вида. Условия спроса на рынке ограничивают число изделий первого вида 2000 единицами, второго – 3000 и третьего – 5000 единицами. Для изготовления изделий используется 4 типа ресурсов. Количество ресурсов, потребляемых для производства одного изделия, общее количество ресурсов и прибыль от реализации каждого вида изделия заданы в таблице 5.4.

Таблица 5.8 – Исходные данные

Тип ресурса	Вид изделий			Наличие ресурсов на предприятии
	1	2	3	
1	500	300	1000	25000000
2	1000	200	100	30000000
3	150	300	200	20000000
4	100	200	400	40000000
Прибыль	20	40	50	

Для решения задачи разработайте математическую модель и выполните расчеты средствами *Excel* с помощью надстройки *Поиск решения*. Проиллюстрируйте решение, построив график оптимального выпуска продукции.

3. Выполните сравнительный анализ эффективности инноваций в Иркутском ГАУ, используя следующие характеристики инновационной деятельности:

- Персонал, занятый исследованиями и разработками (человек).
- Внутренние затраты на исследования и разработки (тыс. руб.).
- Затраты на технологические инновации (тыс. руб.).
- Число защищенных кандидатских диссертаций.
- Поступление патентных заявок и выдача охранных документов.
- Объем инновационной продукции по степени новизны (тыс. руб.).
- Удельный вес затрат на технологические инновации в объеме отгруженной продукции инновационно-активных организаций (%).

4. Выполните прогноз численности трудоспособного населения Иркутской обоасти, основываясь на данных Федеральной службы государственной статистики (<http://www.gks.ru>).

Глоссарий

Абстрагирование - отвлечение от второстепенных фактов с целью сосредоточения на важнейших особенностях изучаемого явления.

Автор изобретения - физическое лицо, творческим трудом которого оно создано.

Автор научного открытия - в РФ - физическое лицо, которое путем наблюдения, изучения, эксперимента или рассуждения самостоятельно сделало научное открытие способом, обеспечивающим его установление.

Если открытие сделано группой физических лиц, то любая ссылка на автора научного открытия рассматривается как ссылка на все эти лица.

Автореферат диссертации – научное издание в виде брошюры, содержащее

составленный автором реферат проведенного им исследования, предоставляемого на соискание ученой степени.

Аксиома – исходное положение, которое не может быть доказано, но в то же время и не нуждается в доказательстве.

Аналогия – это способ получения знаний о предметах и явлениях на основании того, что они имеют сходство с другими.

Библиография - информационная инфраструктура, обеспечивающая подготовку, распространение и использование библиографической информации; перечень различных информационных документов с указанием определенных данных.

Внедрение - распространение нововведений; достижение практического использования прогрессивных идей, изобретений, результатов научных исследований (инноваций).

Гипотеза - научное предположение, выдвигаемое для объяснения некоторого явления и требующее верификации.

График – условное изображение соотношения величин в их динамике при помощи геометрических фигур, линий и точек.

Диаграмма – график, построенный с помощью геометрических фигур, таких как прямоугольник, круг.

Данные – сведения, полученные путем измерения, наблюдения, логических или арифметических операций; представленные в форме, пригодной для постоянного хранения, передачи и (автоматизированной) обработки.

Доклад – запись устного сообщения на определенную тему,

предназначаемая для прочтения на семинарском занятии, конференции.

Документ - по законодательству РФ - материальный объект с зафиксированной на нем информацией в виде текста, звукозаписи или изображения, предназначенный для передачи во времени и пространстве в целях хранения и общественного использования. Документ обязательно содержит реквизиты, позволяющие однозначно идентифицировать, содержащуюся в нем информацию.

Задача - координированная и систематизированная серия элементов работы,

используемых для достижения результатов.

Закон – положение, выражающее всеобщий ход вещей в какой-либо области;

высказывание относительно того, каким образом что-либо является необходимым или происходит с необходимостью.

Идея – это: 1) новое интуитивное объяснение события или явления; 2) определяющее стержневое положение в теории.

Издание - документ:

- прошедший редакционно-издательскую обработку;
- полученный печатанием или тиснением;
- полиграфически самостоятельно оформленный;
- имеющий выходные сведения;
- предназначенный для распространения содержащейся в нем информации.

Изобретение - новое и обладающее существенными отличиями техническое решение задачи в любой области экономики, социального развития, культуры, науки, техники, обороны, дающее положительный эффект. Автор изобретения, получивший авторское свидетельство, имеет право дать изобретению свое имя или специальное название. Изобретение является одним из объектов промышленной собственности.

Интеллектуальная собственность - собственность на результаты интеллектуальной деятельности, интеллектуальный продукт, входящий в совокупность объектов авторского и изобретательского права.

Информационное издание - издание, содержащее систематизированные сведения об опубликованных, непубликуемых или неопубликованных документах или результат анализа и обобщения сведений, представленных в первоисточниках.

Информационные ресурсы - в широком смысле - совокупность

данных, организованных для эффективного получения достоверной информации.

Информационные ресурсы - по законодательству РФ - отдельные документы и отдельные массивы документов, документы и массивы документов в информационных системах, библиотеках, архивах, фондах, банках данных, других видах информационных систем.

Источник информации - объект, идентифицирующий происхождение информации; в теории коммуникации - лицо, от которого исходит сообщение; отправитель сообщения; в теории перевода - создатель или автор текста оригинала.

Категория – общее, фундаментальное понятие, отражающее наиболее существенные свойства и отношения предметов и явлений.

Классификация наук - группировка наук на основе определенных принципов.

Конспект - краткое изложение прочитанного.

Концепция – это система теоретических взглядов, объединенных научной идеей (научными идеями).

Лицензия на изобретение - разрешение, выдаваемое одним лицом (лицензиаром) другому лицу (лицензиату) на коммерческое использование изобретения, защищенного патентом в границах строго определенного рынка, в течение определенного срока и за обусловленное вознаграждение.

Логотип - оригинальное начертание, изображение полного или сокращенного

наименования фирмы или товаров фирмы. Логотип специально разрабатывается фирмой с целью привлечения внимания к ней и к ее товарам.

Материалы научной конференции – научный неперIODический сборник, содержащий итоги научной конференции (программы, доклады, рекомендации, решения).

Методика – это совокупность способов и приемов познания.

Методология - 1) совокупность методов, применяемых в какой-либо сфере деятельности (науке, политике и т.д.); 2) учение о научном методе познания.

Моделирование - исследование объектов познания на их моделях. Моделирование предполагает построение и изучение моделей реально существующих предметов, явлений и конструируемых объектов:

- для определения или улучшения их характеристик;
- для рационализации способов их построения;

- для управления и прогнозирования.

Монография - научное или научно-популярное книжное издание:

- содержащее полное и всестороннее исследование одной проблемы или темы;
- принадлежащее одному или нескольким авторам.

Научная деятельность - интеллектуальная деятельность, направленная на получение и применение новых знаний для:

- решения технологических, инженерных, экономических, социальных, гуманитарных и иных проблем;
- обеспечения функционирования науки, техники и производства как единой системы.

Научная информация - логически организованная информация, получаемая в процессе научного познания и отображающая явления и законы природы, общества и мышления.

Научная проблема – это противоречие между знаниями о потребностях общества и незнанием путей и средств их удовлетворения.

Научно-популярное издание - издание, содержащее сведения:

- о теоретических и/или экспериментальных исследованиях в области науки, культуры и техники; изложенное в форме, доступной читателю-неспециалисту.

Научно-техническая информация - документированная информация, возникающая в результате научного и технического развития, а также информация, в которой нуждаются руководители, научные, инженерные и технические работники в процессе своей деятельности, включая специализированную экономическую и нормативно-правовую информацию.

Научное знание - система знаний о законах природы, общества, мышления.

Научное издание - издание, содержащее результаты теоретических и/или

экспериментальных исследований, а также научно подготовленные к публикации памятники культуры и исторические документы.

Научное исследование - процесс изучения, эксперимента, концептуализации и проверки теории, связанный с получением научных знаний. Различают фундаментальные и прикладные научные исследования.

Научное открытие - установление явлений, свойств или законов

материального мира, ранее не установленных и доступных проверке.

Научный вопрос - мелкая научная задача, относящаяся к конкретной области научного исследования.

Научный результат - продукт научной и/или научно-технической деятельности, содержащий новые знания или решения и зафиксированный на любом информационном носителе.

Научный термин – это слово или сочетание слов, обозначающее понятие, применяемое в науке.

Общественные науки - совокупность наук, изучающих различные аспекты жизни человеческого общества.

Объект исследования – это то социальное явление (процесс), которое содержит противоречие и порождает проблемную ситуацию

Объяснение - этап научного исследования, состоящий:

- в раскрытии необходимых и существенных взаимозависимостей явлений или

процессов;

- в построении теории и выявлении закона или совокупности законов, которым

подчиняются эти явления или процессы.

Описание - этап научного исследования, состоящий в фиксировании данных эксперимента или наблюдения посредством определенных систем обозначений, принятых в науке.

Патент - документ: выдаваемый компетентным государственным органом на определенный срок; • удостоверяющий авторство и исключительное право на изобретение; и наделяющий владельца титулом собственника на изобретение.

Патентная информация - информация, публикуемая патентными организациями. Каждая публикация содержит:

- список ключевых слов;

- коды;

- сведения о патентном документе, включающие описание изобретения, фамилии

авторов, дату поступления заявки, дату приоритета, сведения о правовом положении документа.

Первоисточник - источник информации:

- либо являющийся оригинальным документом, содержащим данные исследования;

• либо составленное рукой непосредственного участника описание событий:

дневник, автобиография, письмо, юридический документ, отчет, протокол, деловая

бумага, счет, газета и т.д.

Полезная модель - объект промышленной собственности; конструктивное выполнение средств производства и предметов потребления, а также их составных частей.

Положение – научное утверждение, сформулированная мысль.

Понятие - мысль, отражающая в обобщенной форме предметы и явления

действительности и существенные связи между ними посредством фиксации общих и специфических признаков.

Предмет исследования - существенные свойства или отношения объекта исследования, познание которых важно для решения теоретических или практических проблем. Предмет исследования определяет границы изучения объекта в конкретном исследовании.

Препринт - научное издание, содержащее материалы предварительного характера, опубликованные до выхода в свет издания, в котором они могут быть помещены.

Прикладные научные исследования - исследования, направленные преимущественно на применение новых знаний для достижения практических целей и решения конкретных задач.

Принцип- основное начало, на котором построено что-н. (какая-н. научная система, теория, политика, устройство и т. п.)

Проблема – неразрешенная задача или вопросы, подготовленные к разрешению.

Процедура исследования - последовательность познавательных и организационных действий с целью решения исследовательской задачи.

В общем случае научное исследование предполагает:

1. постановку задачи;
2. предварительный анализ имеющейся информации, условий и методов
3. решения задач данного класса;
4. формулировку исходных гипотез;
5. сбор данных;
6. анализ и обобщение полученных результатов;
7. проверку гипотез;

8. формулирование утверждений.

Публикация - документ, доступный для массового использования.

Рабочая программа – это изложение общей концепции исследования в соответствии с его целями и гипотезами.

Рецензия — это работа, в которой критически оценивают основные положения и результаты научного исследования.

Реферат - краткое изложение содержания отдельного документа, его части или совокупности документов, включающее основные сведения и выводы, а также количественные и качественные данные об объектах описания.

Рубрикация – деление текста на составные части с использованием заголовков, нумерации и т.д.

Сборник научных трудов - сборник, содержащий исследовательские материалы научных учреждений, учебных заведений или обществ.

Способ – это действие или система действий, применяемые при исполнении какой-либо работы, при осуществлении чего-либо.

Сравнение – это сопоставление признаков, присущих двум или нескольким объектам, установление различия между ними или нахождение в них общего.

Суждение – это мысль, в которой утверждается или отрицается что-либо.

Схема – изложение, описание, изображение чего-либо в главных чертах; обычно делается без соблюдения масштаба с помощью условных обозначений.

Счет (количественный метод) - это определение количественных соотношений объектов исследования или параметров, характеризующих их свойства.

Тезисы докладов научной конференции - научный неперIODический сборник, содержащий опубликованные до начала конференции материалы предварительного характера: аннотации, рефераты докладов и/или сообщений.

Тема — это научная задача, охватывающая определенную область научного исследования.

Теория - форма достоверных научных знаний:

- представляющая собой множество логически увязанных между собой допущений и суждений;
- дающая целостное представление о закономерностях и существенных

характеристиках объектов;

- основывающаяся на окружающей реальности.

Учебник - учебное издание, содержащее систематическое изложение учебной дисциплины, ее раздела или части, соответствующее учебной программе и официально утвержденное в качестве учебника.

Учебное издание – это издание, содержащее систематизированные сведения научного или прикладного характера, изложенные в форме, удобной для изучения и преподавания, и рассчитанное на учащихся разного возраста и ступени обучения.

Учебное пособие – это учебное издание, дополняющее или частично заменяющее учебник и официально утвержденное в качестве учебного пособия.

Факт – действительное, вполне реальное событие, явление; нечто сделанное, совершившееся.

Формализация – представление основных положений процессов и явлений в виде формул и специальной символики.

Фундаментальные научные исследования - экспериментальная или теоретическая деятельность, направленная на получение новых знаний об основных закономерностях строения, функционирования и развития человека, общества, окружающей природной среды.

Эксперимент - общенаучный метод получения в контролируемых и управляемых условиях новых знаний о причинно-следственных отношениях между явлениями и процессами.

Эмпирическое обобщение – это система определенных научных фактов, на основании которой можно сделать определенные выводы или выявить недочеты и ошибки.

Литература

1. Российская Федерация. Законы. О науке и государственной научно-технической политике : федер. закон № 127-ФЗ : [от 23.08.1996] // Сборник законов Российской Федерации. Москва : Эксмо, 2010. 1136 с.
2. Об утверждении государственной программы Российской Федерации «Развитие науки и технологий» на 2013–2020 годы (с изменениями и дополнениями) [Электронный ресурс] : постановление Правительства РФ № 301 : [от 15.04.2014]. Доступ из информ.-прав. системы ГАРАНТ. Режим доступа: <http://base.garant.ru/70643478>. Загл. с экрана.
3. О Программе фундаментальных научных исследований в РФ на долгосрочный период (2013–2020 гг.) (с изменениями и дополнениями) [Электронный ресурс] : распоряжение Правительства РФ № 2538-р : [от 27.12.2012]. Доступ из информ.-прав. системы ГАРАНТ. Режим доступа: <http://base.garant.ru/70291160>. Загл. с экрана.
4. Об утверждении Положения о Высшей аттестационной комиссии при Министерстве образования и науки Российской Федерации [Электронный ресурс] : постановление Правительства РФ № 836 : [от 23.09.2013 (ред. от 17.03.2015)]. Доступ из информ.-прав. системы «КонсультантПлюс». Режим доступа: http://www.consultant.ru/document/cons_doc_LAW_152429. Загл. с экрана
5. Российская Федерация. Законы. О высшем и послевузовском профессиональном образовании : федер. закон № 125-ФЗ : [от 22.08.1996] // Сборник законов Российской Федерации. Москва : Эксмо, 2010. 1136 с.
6. Байбородова Л. В., Чернявская А. П. Методология и методы научного исследования. Учебное пособие. — М.: Юрайт. 2018. 222 с.
7. Бузина, Т.С. Оптимизация взаимодействия участников в региональных агропромышленных кластерах /Т.С. Бузина, Я.М. Иванько. — Иркутск: Изд-во Иркутский ГАУ, 2015.— 148 с.
1. Бутырин М.В. Оценка загрязнения почвенного покрова тяжелыми металлами и мышьяком МО г. Свирск Иркутской области /Ш.К. Хуснидинов,

Т.Н. Сосницкая, М.В. Бутырин, Р.В. Замашиков //Вестник БГСХА. – Улан-Удэ – 2014. – №1(34) – С. 45-50.

2. Вашукевич Е.В. Имитационное моделирование в задачах оценки параметров аграрного производства / Е.В. Вашукевич, В.Р. Елохин, Я.М. Иваньо // Тр. XV Байкальской международной школы-семинара «Методы оптимизации и их приложения», посвященной памяти профессора В.П. Булатова. – Иркутск: ИДСТУ СО РАН. – 2011. – С.70-76.

3. Дрецинский В. А. Методология научных исследований. Учебник для бакалавриата и магистратуры. — М.: Юрайт. 2019. 274 с.

4. Иваньо Я.М. и др. Статистика с применением Excel / Я.М. Иваньо, А.Ф. Зверев, Т.Д. Ким, Л.М. Кузнецова / под ред. Я.М. Иваньо, А.Ф. Зверева // Учеб. пособие / Иркутск: ИрГСХА. – 2006.-138 с.

5. Иваньо Я.М. Природные события, техногенные возмущения и оптимизация производственных процессов / Я.М. Иваньо, Н.В. Старкова // Міжнародна науково-практична конференція «Карпатська конференція з проблем охорони довкілля» Мукачєво-Ужгород, Україна, 15-18 травня 2011 р. – Мукачєво-Ужгород. – 2011. – С. 179-180.

6. Иваньо Я.М. Решение задач управления аграрным производством в условиях неполной информации. Монография / Я.М. Иваньо, В.Р. Елохин и др. / Под редакцией Я.М. Иваньо. – Иркутск: Изд-во ИрГСХА. – 2012. - 200 с.

7. Методология и методы научных исследований в экономике и менеджменте [Текст]: пособие для вузов / Завьялова Н.Б., Головина А.Н., Завьялов Д.В., Дьяконова Л.П., Мельников М.С. и др.; под ред. Н. Б. Завьяловой, А.Н. Головиной – Москва- Екатеринбург:, 2014. – 282 с.

8. Никитин А.В. Экономический механизм страхования и преодоления рисков в сельском хозяйстве России при вступлении в ВТО / А.В. Никитин, А.В. Федоренко. – : ФГНУ «Росинформагротех». – 2006. – 220 с.

9. Основы научных исследований : учеб. для техн. ву- зов / В.И. Крутов [и др.] ; под ред. В.И. Крутова, В.В. По- пова. Москва : Высш. шк., 1989. 400 с.

10. Татарова Г. Г. Основы типологического анализа в социологических исследованиях. — М.: Высшее Образование и Наука. 2015. 236 с.

11. Фриман Эдуард Методы изучения истории. — М.: Либроком. 2015. 200 с.

12. Цыпин Г. М. Работа над диссертацией. Навигатор по "трассе" научного исследования. — М.: Юрайн. 2019. 36 с.

Бузина Татьяна Сергеевна

Методы научных исследований в информатике

Учебное пособие

Лицензия на издательскую деятельность

ЛР № 070444 от 11.03.98 г.

Подписано в печать

Тираж

Издательство ФГБОУ ВО Иркутский ГАУ

664038 Иркутская обл., Иркутский район,

пос. Молодежный