

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РФ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ АВТОНОМНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«КАЗАНСКИЙ (ПРИВОЛЖСКИЙ) ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

ИНСТИТУТ УПРАВЛЕНИЯ, ЭКОНОМИКИ И ФИНАНСОВ

Кафедра географии и картографии

**МЕТОДЫ ИССЛЕДОВАНИЯ ТЕРРИТОРИАЛЬНЫХ
СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИСТЕМ**

Учебное пособие

для бакалавров

по направлениям подготовки

05.03.02 «География»

05.03.03 «Картография и геоинформатика»

Казань 2020

Учебное пособие подготовлено:

Доцентом Ю.А. Шакировой, профессором М.В.Панасюком

Рецензенты:

- профессор, д.г.н. В.А.Рубцов, кафедра сервиса и туризма ИУЭФ КФУ

- зав. гидрологической лабораторией Института проблем экологии АН РТ к.г.н.

А.Т. Горшкова

Обсуждено на заседании кафедры географии и картографии (протокол № 5 от 14 января 2020 г.)

Утверждено учебно-методической комиссией Института управления, экономики и финансов КФУ (протокол № 7 от 12 марта 2020 г.)

УДК 303 Методы исследования территориальных социально-экономических систем

Учебное пособие / Ю. А. Шакирова, М.В.Панасюк. – Казань: Казан. федеральный ун-т, 2020. –113 с.

Предлагаемое учебное пособие предназначено для студентов, обучающихся по направлению подготовки бакалавриата 05.03.02 «География», 05.03.03 «Картография и геоинформатика»

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ.....	6
ТЕМА 1. СУЩНОСТЬ ТЕРРИТОРИАЛЬНЫХ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИСТЕМ (ТСЭС)	7
1.1. Понятие о территориальных социально-экономических системах	7
1.2. Основные подсистемы ТСЭС, их взаимосвязи и взаимозависимости..	9
ТЕМА 2. ОСОБЕННОСТИ ИССЛЕДОВАНИЯ ТЕРРИТОРИАЛЬНЫХ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИСТЕМ.....	15
2.1. Основные этапы исследования территориальных социально-экономических систем.....	15
2.2. Источники информации для исследования территориальных социально-экономических систем.....	16
ТЕМА 3. КЛАССИФИКАЦИЯ МЕТОДОВ ИССЛЕДОВАНИЯ ТЕРРИТОРИАЛЬНЫХ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИСТЕМ.....	21
3.1. Методы получения и первичной обработки географической информации	22
3.2. Методы анализа географической информации	26
ТЕМА 4. СТАТИСТИЧЕСКИЕ МЕТОДЫ ИССЛЕДОВАНИЯ ТЕРРИТОРИАЛЬНЫХ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИСТЕМ.....	32
4.1. Статистический анализ	32
4.2. Корреляционный анализ	43
4.3. Регрессионный анализ.....	49
4.4. Кластерный анализ	54
4.5. Анализ динамических рядов	58
ТЕМА 5. МЕТОДЫ МОДЕЛИРОВАНИЯ ТЕРРИТОРИАЛЬНЫХ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИСТЕМ.....	65
5.1. Понятие и виды математико-географических моделей. Трудности моделирования географических объектов.....	65
5.2. Этапы математико-географического моделирования.....	69
5.3. Модели линейного программирования	72

ТЕМА 6. МЕТОДЫ ПРОГНОЗИРОВАНИЯ РАЗВИТИЯ ТЕРРИТОРИАЛЬНЫХ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИСТЕМ.....	78
6.1. Методологические вопросы прогнозирования. Классификация прогнозов	78
6.2. Методы прогнозирования на основе анализа тенденции временных рядов и уравнения регрессии	84
ЗАКЛЮЧЕНИЕ	90
СПИСОК ЛИТЕРАТУРЫ.....	91
ТЕСТОВЫЕ ЗАДАНИЯ	93
ГЛОССАРИЙ.....	105

ВВЕДЕНИЕ

Учебное пособие к курсу «Методы исследования территориальных социально-экономических систем» предназначено для формирования у студентов теоретических знаний и практических навыков проведения комплексных исследований отраслевых, региональных и глобальных географических проблем, прогнозирования и моделирования социально-экономической и хозяйственной деятельности на территории.

В учебном пособии раскрываются особенности методов экономико-географических исследований, возможности их применения для решения практических задач территориальной организации общества. Особое внимание уделено изучению методов математической статистики, которые наиболее часто применяются в прикладных географических исследованиях.

По каждой теоретической теме предлагаются контрольные вопросы и практические упражнения для самостоятельного решения.

Выполнение комплекса заданий обеспечит закрепление и развитие у обучающихся знаний, умений и навыков анализа, моделирования и прогнозирования социально-экономических процессов.

Пособие также содержит тестовые материалы по основным рассмотренным вопросам и глоссарий, содержащий ключевые терминологические понятия курса.

ТЕМА 1. СУЩНОСТЬ ТЕРРИТОРИАЛЬНЫХ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИСТЕМ (ТСЭС)

1.1. Понятие о территориальных социально-экономических системах

Ключевой категорией экономической географии и предметом ее исследования является территориальная социально-экономическая система (ТСЭС), которая рассматривается на разном таксономическом уровне – макро- мезо- и микрорегиональном.

Под **ТСЭС** понимается экономически- и социально-эффективное взаимосвязанное сочетание элементов общественного воспроизводства, функционирующее на определенной территории как звено географического разделения труда.

ТСЭС представляют собой различного ранга формы пространственной организации социально-экономической жизни и деятельности общества в виде: промышленного или сельскохозяйственного предприятия, города, промышленного центра, района, области, страны, региональной интеграционной группировки стран и т.п.

Вне зависимости от значения всем ТСЭС присущи два обязательных свойства – территориальность и комплексность.

ТСЭС являются сложными, открытыми, динамичными системами, имеющими собственный социально-экономический процесс развития, прямые и обратные связи, ограничения, свойства саморегулирования и самовосстановления.

Выявление и выделение ТСЭС осуществляется в результате интегрального районирования на основе следующих принципов:

1. Экономический принцип, так как основной движущей силой развития ТСЭС является материальное производство.
2. Энергетический принцип, заключающийся в более полном использовании энергетических ресурсов и, в первую очередь, энергии

умственного и физического труда населения, что приводит к наибольшей активности и эффективности всей ТСЭС.

3. Принцип социально-экономической целостности, т. е. взаимосвязанности и взаимообусловленности всех элементов ТСЭС, всех структурных образований и подсистем.

4. Принцип открытости выделяемых систем, каждая из которых развивается на основе географического разделения труда.

5. Принцип перспективности развития, выявленного на основе концепций и развития общества, достижений НТР.

6. Принцип экономической и социальной эффективности ТСЭС.

7. Принцип тяготения территории к определенным ядрам – социально-экономическим узлам и центрам с учетом сложившейся системы расселения.

При исследовании ТСЭС следует учитывать, что их формирование начинается и происходит под влиянием географического разделения труда. Каждая из частей территории начинает отбирать для себя те отрасли и производства, для которых внутренне имеется более благоприятное сочетание природных и общественно-исторических условий. Специализация районов на определенных видах производства сопровождается специализацией других районов на иных видах производства. Обмен продукцией завершает географическое разделение труда.

Географическое разделение труда и специализация районов сопровождаются усилением комплексирования и ростом концентрации человеческой деятельности. Это ведет к созданию сначала поселений, а потом к формированию территориальных социально-экономических комплексов.

Вокруг последних образуются поля социально-экономического тяготения. Происходит сочетание разных форм человеческой деятельности – формируются ТСЭС низшего ранга.

Дальнейшее их развитие происходит под влиянием процессов территориальной концентрации и рассеивания, в результате чего ТСЭС усложняется, уплотняется и одновременно растет территориально.

Выделение ТСЭС в процессе районирования происходит не сразу, а в несколько этапов.

На первом этапе осуществляется социально-экономическая дифференциация территории по функциональным признакам, на втором - выявляются ядра социально-экономического тяготения и определяется социально-экономическая целостность путем изучения взаимосвязей между поселениями; на третьем – выделяются границы ТСЭС, на четвертом - проверяется объективность границ, на пятом этапе исследуются процессы развития выделенных систем.

Исследование развития ТСЭС может осуществляться путем выявления внутренних противоречий и механизмов их преодоления, отношений систем с окружающей средой и т. д.

Формирование и развитие ТСЭС осуществляется во времени и пространстве. Процессы, происходящие в их подсистемах (и между ними), отличаются разными материальными носителями, интенсивностью, цикличностью и пр.

1.2. Основные подсистемы ТСЭС, их взаимосвязи и взаимозависимости

ТСЭС состоят из нескольких взаимосвязанных функционально-структурных подсистем: природно-ресурсной, производственной, населения и его расселения, социальной, рекреационной, инфраструктурной.

Применительно к ТСЭС областного или регионального уровней в качестве подсистем выделяются:

- территория;
- природно-ресурсный потенциал
- население и трудовые ресурсы
- промышленность

- сельское хозяйство
- инфраструктура.

Природно-ресурсная подсистема региона определяется совокупностью всех видов природных ресурсов, которые в настоящее время известны и использование которых в обозримом будущем возможно по техническим критериям.

К ним относятся: рудные и нерудные полезные ископаемые, водные, лесные, почвенно-земельные, природно-климатические и др. ресурсы, расположенные на определенной территории.

Как правило, природно-ресурсная подсистема определяет отраслевой профиль региона и является базисом для развития соответствующего территориально-производственного комплекса. Например, в зависимости от почвенно-климатических ресурсов в сельской местности формируются профиль и формы сельскохозяйственных объектов (растениеводство или животноводство).

Наиболее динамичной и доминирующей из подсистем является производственная (промышленная, сельскохозяйственная). В современном мире она оказывает практически определяющее воздействие на все другие подсистемы.

Производственная подсистема региона представляет собой народнохозяйственный комплекс, возникший на базе межрайонного разделения труда и процессов интеграции.

По своему составу данная подсистема отличается сложной структурой и может быть подразделена на более простые структуры родового и видового порядка, например, экономическая (производственная) структура обычно расчленяется на две: производительные силы и производственные отношения (родовой порядок). Производительные силы, в свою очередь, традиционно подразделяются на отрасли (энергопроизводственные циклы, межотраслевые и отраслевые комплексы и др.) – видовой порядок. Отрасли и межотраслевые

образования состоят из более простых индивидуальных структурных образований типа объединений, комбинатов, элементарных комплексов и т. д.

Население как региональная подсистема представляет сложную интегрированную систему, где основное место принадлежит конкретному индивидууму, человеку. Деятельность людей протекает в определенных сферах: производственно-экономической, социально-политической, семейно-бытовой и др.

Каждая из этих сфер представляет собой определенную систему, состоящую в свою очередь из подсистем, так называемых ролевых функций человека. Многообразие деятельности людей требует выделения ведущей формы. В качестве последней можно выделить трудовую деятельность, без которой невозможно существование человека, социально-экономических регионов и всего общества. Главным структурным компонентом регионов является свободный труд на себя и общество. Трудовая деятельность населения выступает важнейшим консолидирующим стержнем, объединяющим все разноуровневые функциональные подсистемы и в первую очередь – природно-ресурсную, производственную, социальную и т. д.

В то же время население каждого региона выступает не только как производитель, но и как потребитель. Формирование потребностей населения происходит под влиянием многоукладной экономики и зависит от интеллектуального и культурного уровня людей, достижений научно-технического прогресса, степени развитости интересов. Интересы людей и формирующиеся на их основе потребности служат мощным стимулом деятельности товаропроизводителей, функционирования рынков и в конечном счете – саморазвития регионов.

Инфраструктура как подсистема складывается из производственной, рыночной, кредитно-финансовой, коммерческой, информационной, институциональной, социальной и экологической инфраструктуры.

Производственная инфраструктура как региональная подсистема представляет комплекс отраслей, обеспечивающих внешние условия

непосредственно процессов производства. В нее включаются грузовой транспорт, оптовая торговля, электро-, газо-, водоснабжение, складское хозяйство, связь, а также сфера деловых услуг, к которой относятся реклама и маркетинговое обслуживание; аренда и лизинг (сфера, объединяющая в себе аренду и кредитование оплаты аренды оборудования), консультационное обслуживание (консультирование по вопросам управления, инвестиционной политики и др.); инжиниринг (сфера инженерно-консультационных и проектных услуг); сфера по обеспечению временных потребностей в рабочей силе и др.

Рыночная инфраструктура представляет совокупность связанных между собой институтов рыночной экономики, выполняющих определенные функции по обеспечению нормального режима функционирования региона. В соответствии с этим, представляется целесообразным выделить следующие составляющие рыночной инфраструктуры и их элементы: кредитно-финансовую, коммерческую, информационную.

Кредитно-финансовая инфраструктура – совокупность институтов (кредитно-банковские учреждения, фондовые биржи, финансовые и страховые компании), действующих в пределах денежного, фондового, инвестиционного рынков и выполняющих функции по обеспечению нормального режима их функционирования и всего регионального комплекса в целом.

Коммерческая инфраструктура – совокупность институтов, обслуживающих функционирование рынка средств производства, товаров и услуг. К ним относятся: товарные биржи, предприятия оптовой и розничной торговли, аукционы, ярмарки, посреднические фирмы.

Информационная инфраструктура – учреждения информационного обслуживания предприятий различных сфер деятельности (получение и обработка информации, производство программного обеспечения, услуги по созданию и использованию баз данных).

Институциональная инфраструктура как региональная функциональная подсистема – это комплекс отраслей и сфер деятельности, осуществляющих макроэкономическое регулирование экономики и поддерживающих наиболее оптимальные макроэкономические пропорции развития национального хозяйства.

Социальная инфраструктура – комплекс отраслей, связанных с воспроизводством рабочей силы: здравоохранением, образованием, розничной торговлей, пассажирским транспортом, жилищно-коммунальным хозяйством, сферой организации досуга, туризмом, общественным питанием, службой быта и пр.

Экологическая инфраструктура рассматривает основные проблемы взаимодействия общества (производство и население) и природы, которые наиболее остро проявляются в процессе природопользования, имеющего региональные особенности. Экологическая инфраструктура включает: использование естественных условий и ресурсов, восстановление и охрана природных богатств, мониторинг за окружающей природной средой, лицензирование природопользования и т. д. В настоящее время роль и значение экологической подсистемы несомненно повысили, поскольку многие экологически опасные явления носят не только региональный, но и планетарный характер. Объемы промышленных и бытовых стоков и ливневого смыва в городской и сельской местности значительно превышают возможности биологической очистки.

Движение, развитие и роль отдельных подсистем различных ТСЭС неоднозначны, так же как неоднозначны роль и развитие самих подсистем. В своем развитии некоторые подсистемы могут опережать другие в силу ряда обстоятельств, отставать или находиться в состоянии стагнации.

В качестве основных системообразующих отношений, объединяющих функционально подсистемы, выделяют:

- а) принадлежность к одному географическому пространству;

b) наличие внутренних вещественных, энергетических, информационных связей;

c) экономическую и социальную эффективность совместного функционирования;

d) невозможность самостоятельного развития отдельных подсистем;

e) управляемость.

В то же время ТСЭС характеризуется разным уровнем подсистем, отличающихся относительной автономностью. Некоторые подсистемы развиваются по своим законам. Так, природно-ресурсный комплекс функционирует по законам природы, правда испытывая влияние законов общественного развития.

Внутреннее содержание ТСЭС определяется их структурой. Под структурой понимается совокупность пространственно-временных отношений между элементами сопряженных функциональных систем. Каждая подсистема имеет свою структуру, которая содержит в себе свойства общей структуры и в то же время выражает собственную специфику.

Структурными образованиями подсистемы производства являются энергопроизводственные циклы; подсистемы населения – функциональные сети поселений; природного комплекса – виды естественных ресурсов; социальной инфраструктуры – виды общественного обслуживания; производственной инфраструктуры – виды транспорта, связи, коммуникаций.

Однако структура целостных систем не сводится к сумме структурных образований подсистем, а представляет собой новое качественное образование, проявляющееся в трех формах: вещественной, территориальной и функциональной.

Вопросы для самоконтроля:

1. Дайте определение понятию «территориальная социально-экономическая система»?

2. На основе каких принципов осуществляется выделение ТСЭС?

3. Назовите основные структурные элементы ТСЭС?

4. Какие функциональные подсистемы можно выделить применительно к ТСЭС областного или регионального уровней?

5. Приведите три-четыре примера географических объектов, представляющих собой сложные территориальные социально-экономические системы. Определите, как проявляются в них отмеченные в данной главе основные признаки системы?

Практическое упражнение:

Составьте графическое изображение структуры ТСЭС федерального округа, города или предприятия.

С помощью блок-схемы отразите все входящие в ТСЭС подсистемы. Стрелками укажите существующие взаимосвязи между отдельными подсистемами ТСЭС.

ТЕМА 2. ОСОБЕННОСТИ ИССЛЕДОВАНИЯ ТЕРРИТОРИАЛЬНЫХ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИСТЕМ

2.1. Основные этапы исследования территориальных социально-экономических систем

Исследование территориальных социально-экономических систем, как правило, включает несколько последовательных этапов:

- инвентаризационный;
- оценочный;
- прогнозный

Инвентаризационный этап включает выявление, описание, систематизацию, картографирование социально-экономических явлений и процессов. Этот этап, как правило, занимает подготовительный и полевой периоды.

Оценочный этап включает картографирование и оценку социально-экономических процессов.

Прогнозный этап направлен на изучение ожидаемых изменений социально-экономических процессов и явлений за определенный период времени.

Любое исследование следует начинать с разработки рабочей программы исследования, состоящей из следующих разделов:

1. формулировка цели исследования (определяется в рамках поставленной темы и направлена на решение проблемы в рамках того научного направления, в котором проводится исследование);
2. определение задач, решение которых позволяет достичь поставленной цели;
3. определение объекта и предмета исследования (в географических исследованиях объект, как правило, имеет территориальную привязку, а предмет определяется темой и целью исследования).

2.2. Источники информации для исследования территориальных социально-экономических систем

Исследования территориальных социально-экономических систем опираются на разнообразные источники сведений, к числу которых относятся:

- письменные документы;
- статистические данные;
- картографические материалы и другие графические изображения;
- аэрофотоснимки и космические снимки.

Местом накопления и хранения этих материалов в основном служат библиотеки, архивы, соответствующие министерства, ведомства и различные учреждения и предприятия.

Экономико-социально-географическая информация может быть первичной, если исследователь собирает ее самостоятельно из документов, фондов, научных отчетов различных организаций, органов статистики, и вторичной, содержащейся в опубликованных материалах, различных справочниках, картографических произведениях (для учебных целей ее можно взять из учебников и учебных пособий).

Важным источником информации по экономико-географическим проблемам является соответствующая научная и учебно-методическая литература. Ориентирами в выявлении этой литературы служат предметные указатели, имеющиеся в разного рода изданиях — энциклопедиях, реферативных журналах, — библиографических указателях, летописях литературных источников библиотечных каталогах.

Наиболее обширный вид документации — **письменные документы**. Они подразделяются на архивные документы, материалы прессы, документы личного характера и косвенную документацию.

Архивные материалы служат в основном для реконструкции географической среды прошлого в исторической географии. Эти материалы имеют различные сроки хранения. Среди текущей документации организаций и предприятий наибольший интерес для географических исследований представляют отчетные документы (годовые, квартальные отчеты промышленных и сельскохозяйственных предприятий); периодическая отчетность министерств, ведомств, организаций; материалы сельских администраций, предприятий, домоуправлений, загсов, органов милиции и т. д. Как правило, эти материалы относятся к мелким территориальным единицам и в силу этого представляют собой лучший источник информации при крупномасштабных исследованиях.

Газетно-журнальная периодика является важным в экономико-географических исследованиях источником информации по изучаемым районам и предметам и отражает не только конкретный факт, но и восстанавливает общий социально-экономический фон, сопутствующий тому или иному событию.

Из *документов личного характера* наибольший интерес для географии представляют дневники полевых исследователей, путешественников.

Косвенная документация представляет собой различного рода печатные издания негеографического содержания, которые могут послужить ценным дополнительным материалом. К таким источникам относятся различные

справочные издания социологического или экономического характера и даже литературно-художественные произведения, отражающие в той или иной мере конкретную географическую среду и географические проблемы.

Статистические данные — важнейший материал, собираемый в ходе эмпирических экономико-географических исследований. Ценным свойством статистических данных является методическая строгость их сбора и обработки и, как следствие этого, большая обоснованность и достоверность выводов, полученных на их основе. В географических исследованиях статистические данные используются для составления карт и изучения территориальных рядов и тенденций территориального распространения. Особую ценность представляют статистические ежегодники «Российский статистический ежегодник», «Россия», «Россия в цифрах», а также отраслевые статистические сборники: «Промышленность в России», «Сельское хозяйство России», «Строительство в России», «Социальное положение и уровень жизни населения России» и др. Статистическая информация применительно к субъектам федерации содержится в двухтомнике «Регионы России». Самым достоверным и подробным источником информации по населению являются переписи, материалы которых публикуются периодически (переписи населения, жилого фонда и др.), а также в статистических сборниках «Демографический ежегодник России», «Население России за 100 лет».

Для исследований в области экономической, социальной и политической географии мира — в разрезе мирового хозяйства и ведущих государств — необходимо пользоваться материалами международной статистики, прежде всего изданиями ООН и ее специализированных организаций (Продовольственной и сельскохозяйственной организации ООН — ФАО, ЮНЕСКО, Всемирного банка и др.), а также публикациями институтов РАН (Института географии, Института мировой экономики и международных отношений, Института США и Канады, Института Африки и т. д.), национальными статистическими изданиями зарубежных стран, справочными изданиями типа «Страны мира».

Различают следующие виды статистических данных:

— демографическая и жилищная статистика – сведения о численности населения, плотности, половозрастной статистике, семейном положении, национальном составе, естественном и механическом движении населения; площадь жилья, в том числе на 1 жителя, сведения о числе комнат, наличии коммунальных удобств и т. д.;

— промышленная статистика и статистика занятости – сведения о числе занятых, производственных мощностях, капитальных затратах за год, физическом и стоимостном объеме произведенной продукции;

— сельскохозяйственная статистика – статистические ежегодники, а также квартальные и подекадные данные статистического учета проведения основных технологических операций (сев и уборка зерновых и т. д.); сведения о площади сельскохозяйственных угодий, их структуре и структуре посевных площадей, урожайности, сумме затрат, объеме производства, поголовье скота, размере доходов и числе занятых, развитию лесного хозяйства, создании защитных лесонасаждений и др.

— транспортная статистика – сведения о наличии и густоте транспортной сети, в том числе по видам, длине и категориям автомобильных дорог; структуре и объемах грузо- и пассажирооборотов.

Картографические материалы часто служат бесценным источником информации, например топографические, отражающие подробности ландшафта, размер и планировку поселений, транспортной сети; карты-планы городов и целых урбанизированных зон, картосхемы плановых и проектных организаций; тематические специальные карты (почвенные, геологические, агроклиматические и др.).

Аэрофотоснимки и космические снимки показывают все видимые детали природных и культурных ландшафтов, в то время как карты отражают только генерализованные характеристики этих ландшафтов. При этом различают перспективный (под углом) и плановый (вниз под прямым углом)

тип аэросъемки. Их обработка позволяет получить двух- и трехмерное изображение объектов дешифрирования.

Важна оценка достоверности документальной информации, которая зависит в первую очередь от первичности ее источника. Как правило, юридические документы, а также материалы, которые подвергаются прямому или косвенному финансовому контролю, содержат более достоверные сведения, чем прочие документы.

Вопросы для самоконтроля:

1. Назовите основные этапы исследования ТСЭС?
2. Перечислите основные виды работ в рамках инвентаризационного этапа географического исследования?
3. Каковы основные разделы программы географического исследования?
4. Назовите основные источники информации для исследования территориальных социально-экономических систем?
5. Перечислите основные источники статистических данных экономико-географических исследований?

Практическое упражнение:

Разработайте рабочую программу экономико-географического исследования особенностей развития сельскохозяйственной отрасли региона (по выбору).

1. Сформулируйте цель и задачи исследования;
2. Используя статистические материалы сайта Федеральной службы государственной статистики <https://www.gks.ru/>, портала официальной статистической информации «ЕМИСС», создайте статистическую базу для исследования.

Статистическая база должны быть представлена совокупностью статистических показателей, сгруппированных по блокам (не менее 5 в каждом блоке).

Рекомендуемые тематические блоки:

- природный блок (температурный режим, режим влажности, количество осадков и др.);
- экономический блок (объем производства основных видов продукции растениеводства и животноводства, стоимость реализованной продукции, площадь земельных угодий, численность поголовья сельскохозяйственных животных, численность работников, занятых в сельскохозяйственном производстве и др.);
- экологический блок (количество внесенных удобрений, использование средств химизации и др.).

ТЕМА 3. КЛАССИФИКАЦИЯ МЕТОДОВ ИССЛЕДОВАНИЯ ТЕРРИТОРИАЛЬНЫХ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИСТЕМ

Исследование ТСЭС предусматривает использование широкого спектра географических методов.

По основной сфере применения методы, применяемые в экономической и социальной географии, делятся на общенаучные (анализ, обобщение, сравнение, и др.); экономические (балансовый, технико-экономические, нормативный); социологические и социально-психологические (опрос населения, анкетирование); информационные (ГИС-технологии, кибернетические подходы); математические и экономико-математические методы; прогностические (метод Дельфи, эвристическое прогнозирование), графические, в том числе картографические.

По своей сущности экономико-географические методы делятся на эмпирические (полевые, камеральные, лабораторные методы) и теоретические (метод аналогий, дедукция и индукция).

В зависимости от последовательности экономико-географических исследований выделяют методы получения исходной информации (наблюдение, метод полевых исследований, дистанционный, анкетирования) и методы анализа и систематизации информации (описание, сравнительно-

географический картографический, типологии и районирования, количественные методы, математическое моделирование и др.).

3.1. Методы получения и первичной обработки географической информации

К методам получения исходной информации относят: наблюдение, метод полевых исследований, дистанционный, анкетирования.

Наблюдение — исторически это один из первых способов получения географической информации. Его сущность заключается в сборе первичной информации о географическом объекте путем непосредственного восприятия и выявления его характеристик в соответствии с целью исследования. Важное значение при этом имеет правильное определение объекта наблюдения, а также его исследуемых признаков.

Главной задачей наблюдения является изучение функционирования территориальных систем в пространстве и времени.

Наблюдение широко применяется при изучении систем расселения, планировки населенных мест, использования земель, выявления пространственных различий в хозяйственной деятельности и быте населения.

Результаты наблюдения фиксируются как при помощи традиционных способов в виде таблиц, записей в полевой дневник, протоколов, карточек, так и при помощи инструментального наблюдения — фотографирования, киносъемки, аэро- и космической съемки.

Особым видом наблюдения является измерение, которое дает информацию о количественных характеристиках объекта.

Метод полевых исследований играет в географии исключительно важную роль. Его сущность заключается в получении первичного фактического материала и непосредственном изучении территории для установления пространственной дифференциации населения и хозяйства, сопоставления взаимообусловленности природных и социально-экономических процессов.

Полевые исследования проводятся по специально разработанным программам и включают сплошное обследование территории, маршрутные наблюдения и изучение типичных (ключевых) участков.

«Ключевые» исследования позволяют более глубоко изучить географические особенности жизни людей, прежде всего их хозяйственной деятельности.

Во время полевых исследований проводятся беседы со специалистами, руководителями производства, которые помогают лучше понять особенности многих явлений территориальной организации производства и населения и выработать конкретные конструктивные предложения.

При использовании **дистанционных методов**, среди которых ведущими являются аэрокосмические, объект изучается на расстоянии, достигающем иногда несколько тысяч километров.

Эти методы очень разнообразны: визуальные, фотографические, телевизионные, электронно-оптические, спектрометрические, радарные, радио-тепло-локационные, лазерные и др.

Основным из аэрометодов является аэрофотосъемка, которая применяется главным образом при топографической съемке и ландшафтных исследованиях. Аэрофотоснимки позволяют получить объемную модель ландшафта, анализировать использование земель, определять границы разных типов земель, структуру ландшафтов.

К космическим методам относятся прямые наблюдения за состоянием атмосферы, земной поверхности, наземных объектов, а также космическая фото- и телесъемка. Полученные космоснимки охватывают большую территорию (до 500х500 км), космическая съемка характеризуется большой скоростью получения и передачи информации, возможностью повторения снимков одних и тех же объектов в разных ракурсах и в различное время, что позволяет проследить динамику географических процессов и явлений.

Методы социологических исследований в экономико-географических исследованиях, как правило, представляют собой анкетный опрос и наблюдение.

Анкетный опрос включает две разновидности: анкетирование, когда анкета заполняется респондентом самостоятельно, и интервьюирование, когда это делает инициатор опроса по ответам респондента.

Наблюдение предполагает непосредственный зрительный контакт исследователя с объектом изучения, в процессе которого устанавливается его состояние или динамика, в соответствии с целью и задачами производимого научного исследования.

Социологическое исследование включает четыре последовательных, сменяющих друг друга организационно-автономных и вместе с тем содержательно взаимосвязанных этапов:

- подготовку исследования;
- сбор первичной социологической информации;
- подготовку собранной информации к обработке и ее компьютерную обработку;
- анализ обработанной информации, подготовку отчета по итогам исследования, формулирование выводов и рекомендаций.

Конкретный вид социологического исследования обусловлен характером поставленной в нем цели, выдвинутых задач.

Именно в соответствии с ними, то есть в зависимости от глубины требуемого анализа предмета исследования, масштабности и сложности решаемых в ходе него задач, и различают три основных вида социологического исследования: разведывательное, описательное и аналитическое.

Разведывательное исследование – наиболее простой вид конкретно-социологического анализа, поскольку решает весьма ограниченные по своему содержанию задачи.

Оно охватывает, как правило, небольшие обследуемые совокупности и основывается на упрощенной программе и сжатом по объему методическом инструментарии.

Разведывательное исследование может использоваться в качестве предварительного этапа глубоких и масштабных исследований, либо сбора «прикидочных» сведений об объекте исследования для общей ориентации.

Потребность в этом возникает особенно в тех случаях, когда предмет социологического исследования относится к числу мало или вообще не изученных проблем.

Описательное исследование – более сложный вид социологического анализа. По своим целям и задачам оно предполагает получение эмпирических сведений, дающих относительно целостное представление об изучаемом явлении, его структурных элементах.

Описательное исследование проводится по полной, достаточно подробно разработанной программе и на базе методически апробированного инструментария. Его надежная методологическая оснащенность делает группировку и классификацию элементов исследуемого объекта по тем характеристикам, которые выделены в качестве существенных в связи с изучаемой проблемой.

Описательное исследование обычно применяется тогда, когда объект анализа – относительно большая общность людей, отличающаяся разнообразными характеристиками.

Аналитическое исследование – самый углубленный вид социологического анализа, ставящего своей целью не только описание структурных элементов изучаемого явления, но и выяснение причин, которые лежат в его основе и обуславливают характер, распространенность, устойчивость или изменчивость и другие свойственные ему черты.

В силу такого предназначения аналитическое исследование имеет особенно большую практическую ценность.

Если в ходе описательного исследования устанавливается, есть ли связь между характеристиками изучаемого социального явления, то в ходе аналитического исследования выясняется, носит ли обнаруженная связь причинный характер.

3.2. Методы анализа географической информации

К методам анализа и систематизации информации относят описание, сравнительно-географический, картографический методы, метод типологии и районирования, количественные методы, математическое моделирование и прогнозирование и др.

Под географическим **описанием** понимается упорядоченная характеристика территории, а также теоретическое обобщение полученного материала, т. е. систематизация, объяснение и построение теории.

Метод описания – самый древний метод географии, появившийся более 2 тыс. лет тому назад.

При географическом описании важное значение имеет выявление специфического, уникального в объектах исследования на основе как интуитивного, так и точного статистического сопоставления.

Сравнительно-географический метод заключается в выявлении черт сходства и различия между исследуемыми объектами. Он используется для сопоставления социально-экономических систем во времени и пространстве, анализе результатов хозяйственной деятельности, развития населения, сферы обслуживания и др.

Сравнение обследуемых объектов производится по их сходству или, наоборот, контрастности. При этом для каждого конкретного исследования должны быть четко определены цель, задачи, принципы и средства сравнения.

Этот метод позволяет раскрыть наиболее общие свойства географической оболочки, служит средством познания ее структуры, характера функционирования, динамики и развития.

Сравнение географических объектов производится по составу, структурным связям, генезису, типу функционирования. Это необходимо для

установления закономерностей территориальной организации общества, особенностей функционирования и развития геосистем. Правильность результатов, полученных в результате сравнения, основана на соблюдении следующих логических правил:

— сравнивать можно только одномасштабные объекты (например, экономические районы одного иерархического ранга);

— сравнивать следует только однопорядковые признаки и характеристики географических объектов, процессов и явлений;

— сравнивать предметы надо сначала по таким признакам, которые имеют существенное значение, а затем по второстепенным;

— при сравнении необходимо анализировать объективные условия функционирования и развития сравниваемых объектов: природные условия, обеспеченность ресурсами, уровень экономического развития региона, вовлеченность в территориальное разделение труда и др.

Под термином «**картографический метод исследования**» понимается метод использования карт для познания отображенных на них явлений и процессов.

Картографический метод широко используется в экономико-географических исследованиях, являясь не только способом выявления пространственных закономерностей, но и часто конечной целью исследования.

Познание явлений реального мира с помощью картографического метода можно свести к следующим этапам:

- наблюдение за окружающей нас действительностью и получение необходимой информации, подлежащей картографированию;
- построение карты на основе обработанной информации;
- изучение карты или серии карт с целью извлечения новой информации;
- создание при необходимости производных карт или других картографических произведений (профилей, графиков, блок-диаграмм и т.п.).

Картографический метод располагает большим числом приемов для изучения объектов, определения их количественных характеристик, размещения в пространстве.

Существуют разные уровни механизации и автоматизации исследований по картам:

- визуальный анализ, т.е. чтение карт, глазомерное сопоставление и зрительная оценка изучаемых объектов;
- инструментальный анализ – применение измерительных приборов и механизмов;
- компьютерный анализ, выполняемый в полностью автоматическом или интерактивном режиме с использованием специальных алгоритмов, программ или геоинформационных систем.
- описание по картам – традиционный и общеизвестный прием анализа карт.

Геоинформационные методы — сравнительно новые методы исследования, основанные на применении ГИС-технологий.

Географические информационные системы (ГИС) представляют собой особую информационную систему, осуществляющую сбор, обработку, хранение, отображение и распространение географических данных, а также данных непространственного характера.

Назначение ГИС — обеспечение принятия решений по оптимальному управлению природными ресурсами, организации промышленности, сельского хозяйства, транспорта, розничной торговли и других пространственных объектов.

Особенность хранения пространственных данных в ГИС – их разделение на слои.

Многослойная организация электронной карты, при наличии гибкого механизма управления слоями, позволяет объединить и отобразить гораздо большее количество информации, чем на обычной карте.

Под **систематизацией** понимают методические приемы, связанные с распределением изучаемых явлений в соответствии с целью исследования и избранным критериям на такие совокупности, которые обладали бы определенной общностью и в то же время отличались один от другого устойчивыми признаками. К таким приемам относят применяемые в географии классификацию, таксономию, типологию и районирование.

Классификация — это группировка изучаемых объектов преимущественно по количественным признакам. К классификации относят количественные градации, которые отражают стадии развития явлений или иерархические уровни территориальных систем.

Таксономия — особый вид систематизации, заключающийся в делении территории на сопоставимые, иерархически соподчиненные территориальные единицы (таксоны), которые связаны той или иной общностью свойств, признаков и благодаря этому могут быть отнесены к определенной таксономической категории.

На базе классификации и таксономии географических объектов строится их типология. Типология — это группировка сложных объектов по совокупностям (типам) в основном по качественным признакам.

Различают два подхода к типологическому изучению экономико-географических объектов.

Первый осуществляется путем обобщения характерных свойств и признаков объектов и предметов, явлений данного множества.

Другой подход основан на детальном изучении одного или нескольких объектов, которые затем выбираются в качестве эталонов по выделенным существенным свойствам. Они являются образцом, а остальные объекты изучаются в сравнении с ними.

Главная методологическая проблема типологии — отбор критериев, которые не могут быть случайными, многочисленными и связанными между собой. В экономической географии при выборе критериев типологии следует иметь в виду, что внешние (морфологические) признаки менее значимы.

Кроме того, экономико-географические объекты характеризуются большей изменчивостью и устойчивостью процессов, определенными взаимоположением и взаимосвязями.

Районирование — это универсальный метод упорядочения и систематизации территориальных систем, широко используемый в географических науках.

Районирование как метод имеет большое значение для решения задач территориального управления и районной группировки, для административного деления и т. п.

Существует много способов районирования, основные из которых картографический, статистический, математический, комплексный и др.

Количественные методы. Экономическая и социальная география оперирует большим количеством исходной информации, что обуславливает необходимость применения математико-статистических методов ее обработки. В экономической и социальной географии они используются для решения следующих задач:

1. количественной характеристики изучаемых явлений и процессов;
2. анализа природных и социально-экономических факторов территориальной дифференциации хозяйства и населения;
3. выявления статистических взаимосвязей между социально-экономическими системами;
4. изучения динамики развития территориальных систем на разных этапах их развития;
5. разработки обобщающих (интегральных) показателей функционирования геосистем;
6. разработки методов автоматизации типологии и районирования как основы для прогнозирования развития территориальных систем населения и хозяйства;
7. выявления пространственно-временных закономерностей, поиска механизмов формирования географических процессов и явлений;

8. научного обоснования вариантов устойчивого развития геосистем и использования результатов в управлении народным хозяйством.

Для выявления и количественной оценки статистических зависимостей применяются различные методы статистического анализа.

Дисперсионный анализ используется для выявления влияния одного или нескольких факторных признаков на результативный признак при небольшом числе наблюдений.

Корреляционный анализ применяется для выяснения формы и степени взаимосвязи между признаками изучаемого объекта.

Регрессионный анализ необходим для определения степени раздельного и совместного влияния факторов на результирующий признак и количественной оценки этого влияния на основе различных критериев.

Ковариационный анализ включает элементы дисперсионного и регрессионного анализа. Он используется для изучения линейной связи двух или более переменных по отдельным группам данных и оценки значимости различий между линиями регрессий внутри этих групп.

Вопросы для самоконтроля:

1. Приведите наиболее распространенные классификации методов географических исследований.

2. Какие признаки положены в основу классификаций методов географических исследований?

3. С какими науками связаны методы географических исследований?

4. Какими причинами вызвано широкое распространение геоинформационных методов? Для решения каких научных и прикладных задач используются эти методы?

5. Для решения каких экономико-географических задач применяют количественные методы исследования?

Практическое упражнение

С помощью картографического метода оцените плотность населения Российской Федерации.

Используя статистические материалы сайта Федеральной службы государственной статистики <https://www.gks.ru/> составьте картограмму плотности населения по регионам РФ.

При построении картограммы используйте следующую шкалу плотности населения: высокая, средняя, низкая. Для определения ширины интервала используйте формулу:

$$h = \frac{x_{max} - x_{min}}{3} \quad (3.1)$$

Картограмму выполните цветовым фоном либо штриховкой.

В разрезе федеральных округов способом столбчатой диаграммы в соответствующем масштабе покажите соотношение городского и сельского населения.

ТЕМА 4. СТАТИСТИЧЕСКИЕ МЕТОДЫ ИССЛЕДОВАНИЯ ТЕРРИТОРИАЛЬНЫХ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИСТЕМ

4.1. Статистический анализ

Использование статистических методов в географии связано с изучением количественной стороны массовых природных и социально-экономических явлений и процессов в целях раскрытия их качественного своеобразия, тенденций и закономерностей их развития в конкретных условиях места и времени.

При географических исследованиях может изучаться как все множество объектов, так и их часть. В первом случае совокупность объектов называется генеральной, во втором – выборочной.

Генеральной совокупностью называют совокупность всех возможных наблюдений, которые могли бы быть проведены в соответствии с целью исследования. Общее число членов генеральной совокупности называют объемом генеральной совокупности. Число членов в генеральной совокупности может быть конечным или бесконечным.

Часто нет необходимости, а иногда и возможности, проанализировать все исходы рассматриваемого явления, поэтому, как правило, для изучения используется только часть генеральной совокупности называемой **выборочной совокупностью**.

Число членов выборочной совокупности называют объемом выборки.

Установлено, чем больше объем выборки, тем лучше она отражает генеральную совокупность.

Обычно оптимальный объем выборки пропорционален степени изменчивости признака.

Если признак изменяется значительно, то количество измерений должно быть достаточно большим, если признак меняется мало, то надежный результат может быть получен и при малом объеме выборки.

Наиболее часто используется случайный способ отбора элементов в выборочную совокупность, при котором все объекты имеют одинаковую возможность попасть в выборку.

В исследованиях используются несколько способов случайного отбора элементов в выборочную совокупность: например, с помощью таблицы случайных чисел.

В результате обработки и систематизации первичных статистических материалов получают ряды цифровых показателей, которые характеризуют отдельные стороны изучаемых явлений либо их изменение во времени.

Упорядоченное распределение единиц совокупности по какому-либо варьирующему признаку называется **рядом распределения**.

В ряду распределения выделяют два структурных элемента:

1. **варианты** — различные значения группировочного признака;
2. **частоты** - числа, показывающие, как часто встречается та или иная варианта в ряду распределения.

Сумма частот ряда распределения равна объему совокупности.

Частоты, выраженные в виде относительных величин, называются **частостями**.

Сумма частотей всегда равна 1, если они выражены в долях единицы, или 100%, если они выражены в процентах.

Накопленные частоты (частости) определяются путем последовательного прибавления к частотам (частостям) первой группы этих показателей последующих групп ряда распределения.

Ряды распределения различаются по виду и характеру вариации признака:

1. По виду признака ряды распределения могут быть атрибутивными и вариационными.

Атрибутивные ряды — это ряды, в которых признак выражен определенным термином, фиксирующим свойство или качество предмета или явления (распределение населения по полу, занятости, национальности, профессии и т.д.).

Вариационные ряды — это ряды, в которых варианты признака выражены цифрами (распределение населения по возрасту, рабочих — по стажу работы, зарплате и т.д.).

2. В зависимости от характера вариации различают дискретные и интервальные вариационные ряды.

Дискретные вариационные ряды — это ряды, в которых признак выражается в виде определенного числа, взятого с заданной степенью точности.

Интервальные вариационные ряды — это ряды, в которых варианты заданы в виде интервалов.

Анализ рядов распределения можно проводить на основе их графического изображения. Для этой цели строят полигон, гистограмму, кумуляту или огиву.

Полигон распределения - ломанная, соединяющая точки, соответствующие значениям варьирующего признака и их частотам.

Для изображения интервальных вариационных рядов применяют **гистограммы**, представляющие собой ступенчатые фигуры, состоящие из

прямоугольников, основания которых равны ширине интервала, а высоты - частотам.

Для графического изображения ряда распределения с помощью **кумуляты** на оси абсцисс наносятся значения вариант (или точки дискретного ряда, или верхние границы интервалов), а по оси ординат – накопленные частоты или частоты.

У **огивы** наоборот, по оси абсцисс – накопленные частоты, а по оси ординат – верхние границы интервалов

Характеристику генеральной совокупности дают по параметрам, полученным на основании данных выборки.

Основные выборочные параметры подразделяют на три группы, которые относятся к числу обобщающих показателей.

Степень соответствия показателей выборки параметрам генеральной совокупности оценивается ошибкой (m).

Первую группу образуют **показатели среднего положения или центральной тенденции ряда**.

К ним относятся: различные виды **средних величин, мода, медиана**. Они выражаются именованными величинами, т.е. сохраняют размерность признака.

Вторую группу образуют **показатели разнообразия признака** (разброса, изменчивости значений выборки относительно средней величины): **амплитуда колебания значений, среднее абсолютное отклонение, среднее квадратическое отклонение, коэффициент вариации**.

Третью группу образуют показатели формы распределения: **коэффициенты асимметрии и эксцесса**.

Одним из основных параметров статистического ряда является **среднее значение величины признака**, или центр, относительно которого распределяются элементы совокупности.

Важнейшее свойство средней заключается в том, что она отражает то общее, что присуще всем единицам исследуемой совокупности.

Средняя арифметическая вычисляется по формуле:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (4.1)$$

где n – число членов выборки,

x_i – конкретное значение ряда.

В тех случаях, когда значения признаков повторяются по несколько или много раз вычисляют среднюю арифметическую взвешенную:

$$\bar{x} = \frac{\sum_{i=1}^n x_i \cdot f_i}{\sum_{i=1}^n f_i} \quad (4.2)$$

где x_i – конкретное значение ряда;

f_i – число единиц, имеющих это значение;

Ошибка среднего арифметического определяется по формуле:

$$m_{\bar{x}} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n(n-1)}} = \sqrt{\frac{\sigma^2}{n}} \quad (4.3)$$

где $m_{\bar{x}}$ – ошибка средней арифметической;

σ_x – среднее квадратическое отклонение;

n – объем выборки.

Структурные средние (мода и медиана) – это наиболее простые представители средних величин, не требующие специальных расчетов.

Медиана (Me) равна значению члена варьирующего ряда, занимающего среднее положение в ряду, в котором величины, образующие его, располагаются в убывающем или возрастающем порядке (т.е. мы имеем простую ранжированную совокупность).

Если число элементов в ранжированном ряду нечетное, то его медианой является значение члена ряда, занимающего центральное положение, если же число членов ряда четное, то за медиану условно принимается среднее значение полученное из двух величин, занимающих центральное положение в ранжированном ряду.

При группировке вариационного ряда в классы медиану определяют по следующей формуле:

$$M_e = x_0 + i \frac{\frac{\sum f_i}{2} - S_{m-1}}{f_m} \quad (4.4)$$

где x_0 – начало класса, в котором находится медиана;

$\sum f_i$ – сумма частот всех интервалов, предшествующих медианному;

f_m – частота медианного интервала.

Мода (Mo) – это наиболее часто встречающийся вариант ряда.

Если вариационный ряд разбит на классы, то мода соответствует максимальной частоте класса, который называется модальным, и определяется по формуле:

$$M_0 = x_m + i \left(\frac{f_2 - f_1}{2f_2 - f_1 - f_3} \right) \quad (4.5)$$

где x_m – меньший предел модального класса; i – классовой интервал;

f_1 – частота класса, предшествующего модальному;

f_2 – частота модального класса;

f_3 – частота класса, следующего за модальным.

Для получения более полного представления о выборочных совокупностях используют показатели рассеяния вариант, или разнообразия признаков: **среднее квадратическое (стандартное) отклонение, средний квадрат отклонений (дисперсия), коэффициент вариации.**

Эти показатели признаков характеризуют различную степень и особенности разброса.

Среднее квадратическое отклонение (σ_x) показывает степень рассеяния значений статистической совокупности около среднего значения.

Среднее квадратическое отклонение σ_x рассчитывается по формуле:

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} \quad (4.6)$$

где x_i – величина признака;

\bar{x} – средняя величина признака;

n – число наблюдений.

Ошибка среднего квадратического отклонения определяется по формуле:

$$m_{\sigma} = \frac{\sigma}{\sqrt{2(n-1)}} \quad (4.7)$$

Квадрат среднего квадратического отклонения называется **дисперсией**:

$$\sigma_x^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} \quad (4.8)$$

Недостаток среднего квадратического отклонения, как и дисперсии, заключается в том, что критерии представляют собой абсолютную именованную величину, поэтому их нельзя использовать при сравнении разнородных рядов, выраженных в различных единицах измерения.

Для этой цели подходит **коэффициент вариации** (C_v) представляющий собой отношение среднего квадратического отклонения к среднему значению ряда:

$$C_v = \frac{\sigma_x}{\bar{x}} * 100\% \quad (4.9)$$

где σ_x – среднее квадратическое отклонение;

\bar{x} – среднее арифметическое значение ряда;

C_v – коэффициент вариации.

Ошибка коэффициента вариации определяется по формуле:

$$m_{C_v} = \frac{C_v}{\sqrt{n}} \cdot \sqrt{\frac{1}{2} + \left(\frac{C_v}{100}\right)^2} \quad (4.10)$$

Коэффициент вариации позволяет оценить вариабельность (разброс) признака в нормированных границах. Если его значение меньше 10 %, то разброс вариант относительно средней арифметической считается слабым, при 10–30 – средним, 30–60 – высоким, 60–100 – высоким, более 100 % – аномальным.

Основная задача анализа вариационных рядов – это выявление подлинной закономерности распределения, которая достигается увеличением объема исследуемой совокупности при одновременном уменьшении интервала ряда.

Если увеличить объем совокупности и уменьшить интервал группировки и изобразить эти данные графически, то полигон (гистограмма) распределения все более приближается к некоторой плавной линии, являющейся для него пределом и носящей название кривой распределения.

Под **кривой распределения** понимается графическое изображение в виде непрерывной линии изменения частот в вариационном ряду, функционально связанного с изменением вариант.

Формы кривых распределения разнообразны.

График **нормального распределения** (Рис.1) представляет собой симметричную одновершинную кривую, напоминающую по форме колокол. Форма нормальной кривой и положение ее на оси абсцисс полностью определяются двумя параметрами – средним арифметическим значением \bar{x} и средним квадратическим отклонением σ_x .

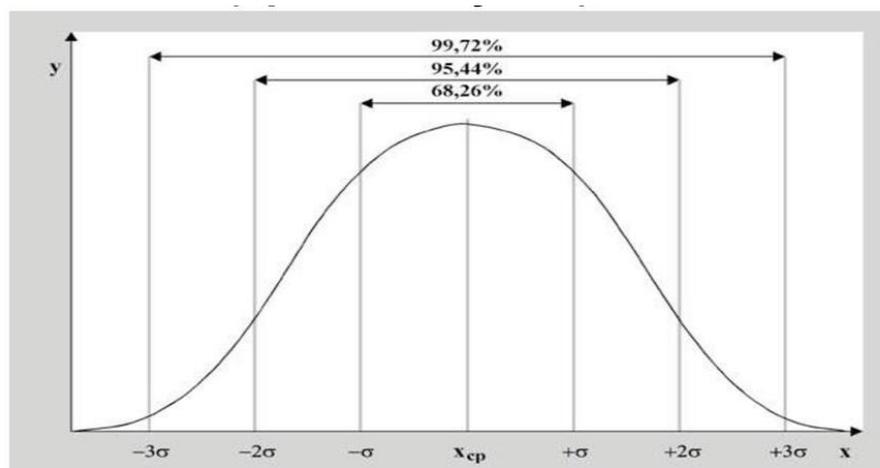


Рис.1. График кривой нормального распределения

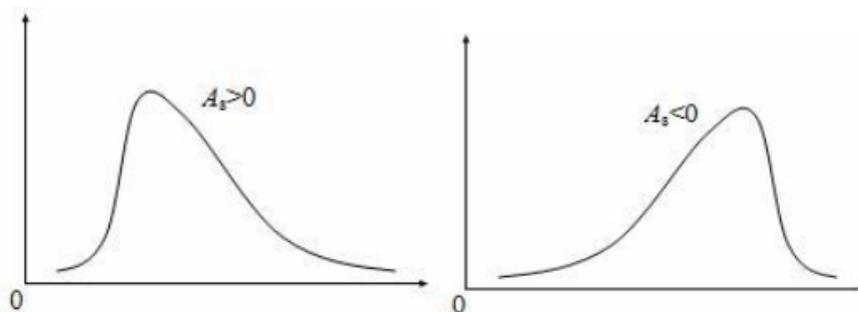
При нормальном распределении в пределах $x \pm \sigma$ находятся 68,3 % всех частот; в пределах $x \pm 2\sigma$ – 95,4% и в пределах $x \pm 3\sigma$ – 99,7 % (правило трех сигм).

Кривые распределения могут быть симметричными и асимметричными.

Для симметричных распределений частоты любых двух вариантов, равноотстоящих в обе стороны от центра, равны между собой.

Если указанные соотношения нарушены, то это свидетельствует о наличии асимметрии распределения.

Различают распределение **скошенное вправо (положительное) или влево (отрицательное)** (Рис.2).



Правосторонняя асимметрия (+) Левосторонняя асимметрия (-)

Рис.2. Положительная и отрицательная асимметрия распределения

При получении асимметричной кривой следует оценить асимметричность распределения. С этой целью вычисляют **коэффициент асимметрии**, его ошибку, затем на основании показателя достоверности устанавливают вид кривой распределения.

Коэффициент асимметрии (A_s) вычисляют по формулам:

$$A_s = \frac{\bar{x} - M_0}{\sigma_x} \quad (4.11)$$

$$A_s = \frac{\bar{x} - Me}{\sigma_x} \quad (4.12)$$

где \bar{x} – средняя величина признака;

M_0 – мода,

Me – медиана,

σ_x – среднее квадратическое отклонение.

Ошибку коэффициента асимметрии (m_{as}) можно рассчитать по следующей формуле:

$$m_{as} = \frac{\sigma}{n+3} \quad (4.13)$$

где n – объем выборки.

Если $A_s > 0$, то асимметрия правосторонняя, а если $A_s < 0$, то асимметрия левосторонняя.

Принято считать, что асимметрия выше 0,5 (независимо от знака) считается значительной; если она меньше 0,25, то незначительной.

Достоверность коэффициента асимметрии определяется по t – критерию Стьюдента:

$$t = \frac{|A_s|}{m_{as}} \quad (4.14)$$

Для оценки степени асимметричности ряда ориентировочно можно использовать следующее соотношение: в случае $t > 3$ асимметрия существенна и распределение признака в генеральной совокупности несимметрично. В противном случае асимметрия несущественна и ее наличие может быть вызвано случайными обстоятельствами.

Эксцесс является мерой «сглаженности» («остро-» или «плосковершинности») распределения.

Если значение эксцесса близко к 0, это означает, что форма распределения близка к нормальному виду.

Отрицательный эксцесс указывает на «плосковершинное» распределение, у которого максимум вероятности выражен не столь ярко, как у нормального (Рис.3).

Положительный эксцесс, напротив, характеризует «островершинное» распределение, график которого более вытянут по вертикальной оси, чем график нормального распределения

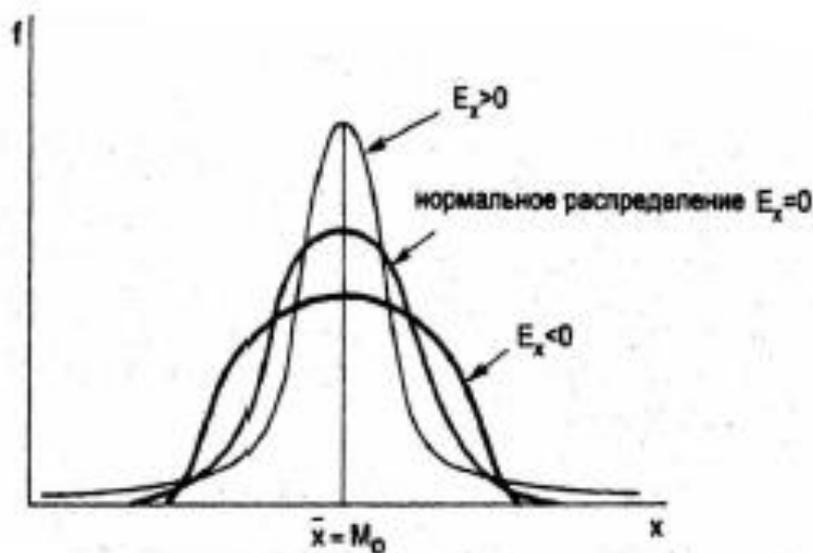


Рис.3. Положительный и отрицательный эксцесс

Считается, что распределение с эксцессом в диапазоне от -1 до +1 примерно соответствует нормальному виду. В большинстве случаев вполне допустимо считать нормальным распределение с эксцессом по модулю не превосходящим 2.

Показатель эксцесса (E) определяется по формуле:

$$E = \frac{\sum_{i=1}^n (x_i - \bar{x})^4}{n \cdot \sigma_x^4} - 3 \quad (4.15)$$

где n – объем выборки;

x_i – значение признака;

\bar{x} – среднее арифметическое значение признака;

σ_x – среднее квадратическое отклонение.

Ошибка коэффициента эксцесса (m_E) вычисляется по следующей формуле:

$$m_E = 2 \sqrt{\frac{6}{n} + 5} \quad (4.16)$$

Хотя показатели асимметрии и эксцесса характеризуют непосредственно лишь форму распределения признака в пределах изучаемой совокупности, но их определение имеет не только описательное значение.

Часто асимметрия и эксцесс дают определенные указания для дальнейшего исследования социально-экономических явлений. Так появление значительного отрицательного эксцесса может указывать на качественную неоднородность исследуемой совокупности.

4.2. Корреляционный анализ

Корреляционным анализом называется математический метод, позволяющий установить направление и тесноту взаимосвязи между определенными явлениями.

Корреляционный анализ широко применяется в прикладных географических исследованиях, однако необходимо иметь в виду, что доказательство математической связи должно опираться на реальную зависимость между явлениями, так как иногда можно установить несуществующие корреляции.

Корреляционный анализ решает следующие задачи:

- установление направления и формы связи,
- оценка тесноты связи,
- оценка репрезентативности статистических оценок взаимосвязи,
- определение величины детерминации (доли взаимовлияния)

коррелируемых факторов.

Для оценки связи используют следующие численные критерии (коэффициенты) корреляционной связи:

- коэффициент корреляции (r) при линейной зависимости,
- корреляционное отношение (η) при нелинейной зависимости,
- коэффициенты множественной регрессии,
- ранговые коэффициенты линейной корреляции.

В отличие от функциональных связей корреляционные связи (корреляция) не являются точной зависимостью одного признака от другого. На графике они выглядят в виде разброса (роя) точек (Рис.4).

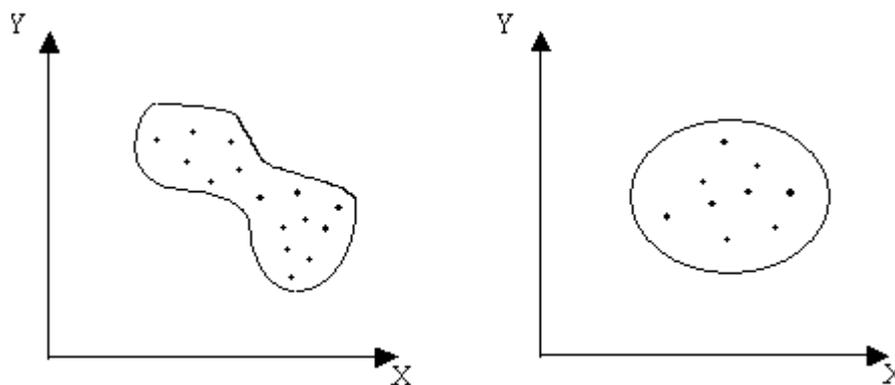


Рис.4. Примеры корреляционной зависимости

Установить наличие связи и получить представление о ее характере можно на основе сопоставления двух или нескольких рядов статистических величин.

По направлению корреляционная связь может быть прямой и обратной; по величине – слабой, средней или сильной; по форме – линейной и нелинейной; по количеству коррелируемых признаков – парной и множественной.

При прямой (положительной) корреляции более высоким значениям одного признака соответствуют более высокие значения другого, а более низким значениям одного признака – низкие значения другого.

При отрицательной (обратной) корреляции соотношения обратные.

Корреляционные связи могут быть однофакторными, когда исследуется связь между одним признаком – фактором и одним признаком – следствием (парная корреляция).

Они могут быть многофакторными, когда исследуется влияние многих взаимодействующих между собой признаков – факторов на признак – следствие (множественная корреляция).

Чаще всего в географических исследованиях используют парный коэффициент корреляции, что, возможно, связано с несложностью его расчетов.

При расчете коэффициента корреляции рекомендуется предварительно построить график и выяснить характер распределения значений параметров.

От формы корреляционной связи зависит дальнейшая обработка экспериментальных или статистических данных (Рис.5).

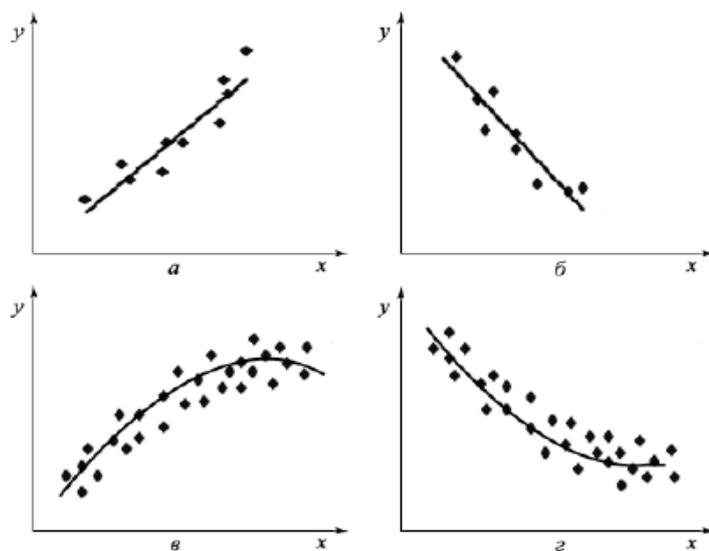


Рис.5. Форма корреляционной связи: а – прямая линейная; б – обратная линейная; в – параболическая; г – гиперболическая

Линейная зависимость предполагает вычисление коэффициента корреляции r , а нелинейная – корреляционного отношения η .

Если зависимость между признаками на графике указывает на линейную корреляцию, рассчитывают коэффициент корреляции Пирсона r .

Величина коэффициента линейной корреляции Пирсона изменяется в пределах от -1 до 1.

Для интерпретации полученной связи важен знак коэффициента корреляции.

Если знак коэффициента линейной корреляции – плюс, то связь между коррелирующими признаками такова, что большей величине (меньшей) одного признака (переменной) соответствует большая (меньшая) величина другого признака (другой переменной).

Иными словами, если один показатель (переменная) увеличивается (уменьшается), то соответственно увеличивается (уменьшается) и другой показатель (переменная).

Если же получен знак минус, то большей величине одного признака соответствует меньшая величина другого. Иначе говоря, при наличии знака минус, увеличению одной переменной (признака, значения) соответствует уменьшение другой переменной.

Если $r=0$, то связи между переменными нет, если он равен +1 или -1, то связь прямолинейная функциональная (точки поля корреляции расположатся строго по линии).

Для оценки тесноты связи по величине коэффициента корреляции используют шкалу Чеддока (Табл.1)

Таблица 1

Шкала Чеддока

Значение коэффициента корреляции	Теснота связи
0,10-0,30	слабая
0,31-0,50	умеренная
0,51-0,70	заметная
0,71-0,90	высокая
0,91-1	весьма высокая

Линейная корреляционная зависимость

При линейной зависимости рассчитывается коэффициент корреляции Пирсона.

$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2 \sum(y_i - \bar{y})^2}} \quad (4.17)$$

где x_i, y_i - значения ряда;

\bar{x} и \bar{y} – средние арифметические значения ряда.

Достоверность вычисленного коэффициента корреляции можно установить через критерий Стьюдента. При использовании критерия Стьюдента для доказательства достоверности r вначале рассчитывают стандартную (квадратическую) ошибку коэффициента корреляции по формуле:

$$m_r = \sqrt{\frac{1-r^2}{n-2}} \quad (4.18)$$

где n – число сопряженных пар данных в сравниваемых выборочных совокупностях.

Затем вычисляют критерий Стьюдента для коэффициента корреляции:

$$t_r = \frac{r}{m_r} \quad (4.19)$$

Вычисленный (фактический) критерий Стьюдента сравнивают с табличным. Если $t_\phi > t_\tau$, то корреляционная связь существенна, при $t_\phi < t_\tau$ – недостоверна.

Нелинейная корреляционная зависимость

При нелинейной зависимости рассчитывается корреляционное отношение. Исходная выборка показателей разбивается на частные группы по значениям y . Их должно быть не менее трех.

Для частных групп рассчитываются средние \overline{y}_Γ и отклонение их от общей средней для выборки \bar{y} , а также отклонения индивидуальных вариантов выборки (y_i) от общей средней (M_y).

Корреляционное отношение вычисляется по формуле:

$$\eta = \sqrt{\frac{n \sum (\overline{y}_\Gamma - \bar{y})^2}{\sum (y_i - \bar{y})^2}} \quad (4.20)$$

где \overline{y}_Γ – среднее арифметическое частных групп; n – число вариантов в частной группе; $(\overline{y}_\Gamma - \bar{y})$ – отклонение общего среднего от средних арифметических частных групп.

Ошибка корреляционного отношения вычисляется следующим образом:

$$m_{\eta} = \sqrt{\frac{(1-\eta)^2}{n-2}} \quad (4.21)$$

Достоверность корреляционного отношения можно установить через критерий Стьюдента:

$$t_{\eta} = \frac{\eta}{m_{\eta}} \quad (4.22)$$

Вычисленный (фактический) критерий Стьюдента сравнивают с табличным. Если $t_{\phi} > t_{\tau}$, то корреляционная связь существенна.

Коэффициент ранговой корреляции Спирмена

Наряду с относительно точными и сложными корреляционными измерениями имеются и более простые способы оценки взаимосвязи между изучаемыми статистическими рядами.

К ним можно отнести коэффициент ранговой корреляции Спирмена.

Благодаря удовлетворительному для многих целей уровню точности данный метод из-за простоты расчетов оказывается исключительно полезным для получения обобщенной оценки корреляции.

Метод Спирмена относится к числу так называемых непараметрических методов, так как он рассматривает не фактические численные значения переменных, а только их место в упорядоченном (ранжированном) списке.

Формула для расчета коэффициента ранговой корреляции Спирмена имеет следующий вид:

$$r_C = 1 - \frac{6\sum D^2}{n(n^2-1)} \quad (4.23)$$

где r_C – коэффициент Спирмена;

D^2 – квадрат разности рангов значений ряда;

n – число сопоставляемых пар рангов.

Достоверность рангового коэффициента корреляции Спирмена можно установить по таблице критических значений коэффициента ранговой корреляции Спирмена.

Ранговый коэффициент корреляции рассчитывается, когда объем выборки не менее 5 и не более 30 элементов.

Значение r_C может меняться в диапазоне от -1 до +1.

Коэффициент ранговой корреляции рассчитывается главным образом тогда, когда нужно выявить приближенную тесноту связи. Однако он представляет весьма значительную ценность для географов в связи с частым использованием данных о многих природных и общественных явлениях, которые выражены в рангах и баллах.

Множественная корреляционная зависимость

Иногда в определенных исследованиях необходимо оценить степень общей взаимосвязи между некоторым показателем x и влияющими на него факторами (y, z, \dots, n).

В этом случае используется коэффициент множественной корреляции R , формула которого для трех коррелируемых величин имеет вид:

$$R = \sqrt{\frac{r_{xy}^2 + r_{xz}^2 - 2r_{xy}r_{xz}r_{yz}}{1 - r_{yz}^2}} \quad (4.24)$$

где r_{xy}, r_{xz}, r_{yz} – коэффициенты парной корреляции.

4.3. Регрессионный анализ

Логическим продолжением корреляционного анализа является регрессионный анализ, который развивает и углубляет представление о корреляционной связи.

Если корреляционный анализ позволяет установить лишь форму и тесноту зависимости между случайными переменными, то регрессионный

анализ математически описывает выявленную зависимость, т. е. дает возможность численно оценить одни параметры через другие.

Суть регрессионного анализа заключается в нахождении наиболее важных факторов, которые влияют на зависимую переменную.

Подобно корреляции регрессия может быть **по числу признаков:** парной (простой) и множественной, **по форме связи:** линейной и нелинейной, **по зависимости:** односторонней (изменяется лишь один признак под влиянием другого) и двусторонней (изменяются оба признака под воздействием друг друга).

Регрессионный анализ возможен при наличии всего лишь нескольких пар сопряженных наблюдений, но при условии сильных связей между признаками ($r \geq 0,7$).

Линейная парная регрессия

Уравнение линейной парной регрессии в общей форме записывается следующим образом:

$$y = ax + b \quad (4.25)$$

где y - зависимая переменная (переменная, описывающая процесс, который мы пытаемся предсказать или понять);

x - независимая переменная (переменная, используемая для моделирования или прогнозирования значений зависимых переменных). В уравнении регрессии они располагаются справа от знака равенства и часто называются объяснительными переменными;

a - коэффициент регрессии, показывающий степень зависимости между переменными;

b - постоянная величина (свободный член уравнения).

В общем случае коэффициент регрессии a показывает, как в среднем изменится результативный признак (y), если факторный признак (x) увеличится на единицу.

Однако в действительности между x и y наблюдается не столь жесткая связь.

Обычно зависимая переменная находится под влиянием целого ряда факторов, в том числе неизвестных исследователю, а также случайных причин. Учитывая возможные отклонения, линейное уравнение будет выглядеть следующим образом:

$$y = ax + b + \xi \quad (4.26)$$

где ξ – случайная переменная.

Для определения параметров регрессионного уравнения a и b наиболее часто используют формулы, полученные на основе метода наименьших квадратов, который минимизирует сумму квадратов отклонения наблюдаемых значений y от модельных значений \hat{y} .

$$a = \frac{\sum(y_i - \bar{y})(x_i - \bar{x})}{\sum(x_i - \bar{x})^2} \quad (4.27)$$

$$b = \bar{y} - a\bar{x} \quad (4.28)$$

Качество модели регрессии связывают с ее адекватностью наблюдаемым (эмпирическим) данным. Проверка адекватности (или соответствия) модели регрессии наблюдаемым данным проводится на основе анализа остатков e_i .

Остаток e_i представляет собой отклонение фактического значения зависимой переменной от ее значения, полученного расчетным путем:

$$e_i = y_i - \hat{y}_i \quad (4.29)$$

Если $e_i = 0$, то результирующий признак y полностью обусловлен влиянием фактора x .

При анализе качества модели регрессии используют коэффициент детерминации R^2 , который показывает долю вариации результативного признака, находящегося под воздействием изучаемых факторов, т. е. определяет какая доля вариации признака y учтена в модели и обусловлена влиянием на него факторов.

$$R^2 = 1 - \frac{\sum_{i=1}^n e_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (4.30)$$

Чем ближе R^2 к единице, тем выше качество модели.

После того, как уравнение регрессии построено, выполняется проверка значимости построенного уравнения в целом и отдельных параметров.

Во-первых, необходимо оценить значимость уравнения регрессии, т.е. установить соответствует ли математическая модель, выражающая зависимость между y и x , фактическим данным и достаточно ли включенных в уравнение объясняющих переменных x для описания зависимой переменной y .

Для проверки значимости уравнения регрессии используется F -критерий Фишера.

Если расчетное значение F с $\nu_1 = k$ и $\nu_2 = n - k - 1$ степенями свободы, где k – количество факторов, включенных в модель, больше табличного, то модель считается значимой.

$$F = \frac{R^2}{1 - R^2} (n - 2) \quad (4.31)$$

Во-вторых, необходим анализ статистической значимости параметров модели регрессии.

Проверка значимости отдельных коэффициентов регрессии связана с определением расчетных значений t -критерия (t -статистики).

$$t_{aP} = \frac{|a|}{s_a} \quad (4.32)$$

$$t_{bP} = \frac{|b|}{s_b} \quad (4.33)$$

$$s_a = \sqrt{\frac{S_e^2}{\sum_{i=1}^n (x_i - \bar{x})^2}} \quad (4.34)$$

$$s_b = \frac{S_e^2 \sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2} \quad (4.35)$$

$$S_e = \sqrt{\frac{1}{n-2} \sum_{i=1}^n e_i^2} \quad (4.36)$$

S_a, S_b – стандартные ошибки коэффициентов уравнения регрессии;

S_e – стандартная ошибка уравнения регрессии.

Расчетные значения t -критерия сравниваются с табличным. Табличное значение критерия определяется при $(n-2)$ степенях свободы (n – число наблюдений) и соответствующем уровне значимости α .

Если расчетное значение t -критерия с $(n-k-1)$ степенями свободы (k – число факторов, включенных в модель) больше его табличного значения при заданном уровне значимости, коэффициент регрессии считается значимым. В противном случае фактор, соответствующий этому коэффициенту должен быть исключен из модели, а оставшиеся в модели параметры пересчитаны.

Нелинейная парная регрессия

При проведении исследований может быть установлена нелинейная зависимость между аргументом и функцией, представляющая собой на графике кривую в виде гиперболы или параболы.

Общее уравнение парной регрессии для гиперболической зависимости имеет вид:

$$y = \frac{a}{x} + b \quad (4.37)$$

Общее уравнение регрессии для параболической зависимости имеет вид:

$$y = ax^2 + bx + c \quad (4.38)$$

где x – аргумент; y – функция; a, b, c – коэффициенты, величину которых следует установить.

Множественная регрессия

Если при установлении зависимости между признаками используется больше одной независимой переменной, то применяют множественный регрессионный анализ.

Проведение такого анализа возможно в следующих условиях: распределение зависимой переменной при различных значениях независимых

должно быть близко к нормальному; дисперсия зависимой переменной при разных значениях признаков x должна считаться одинаковой. С увеличением числа признаков и в случаях нелинейной множественной регрессии необходимо использовать ЭВМ.

Общее уравнение линейной множественной регрессии имеет вид

$$y = a + bx + cz \quad (4.39)$$

где x, z – независимые переменные; y – зависимая переменная; a, b, c – коэффициенты, величину которых следует установить.

4.4. Кластерный анализ

При проведении географических исследований, как правило, возникает проблема объединения по сходству объектов, которые характеризуются множеством признаков, выраженных в разных единицах измерения.

Для этой цели используется кластерный анализ.

Содержательная постановка задачи при кластерном анализе заключается в следующем.

Имеется некоторая совокупность объектов, которые характеризуются рядом признаков. Объекты необходимо разбить на несколько кластеров (классов) таким образом, чтобы объекты из одного класса были сходными по характеризующих их признакам, например, сравнение ландшафтов, выявление сходных тенденций в развитии экономических субъектов.

В качестве характеристики степени сходства между объектами в кластерном анализе используется так называемое «таксономическое расстояние», определяющее степень удаленности данных объектов друг от друга в многомерном математическом пространстве.

Наиболее часто эти расстояния вычисляют по формуле Пифагора, используя Евклидову метрику пространства:

$$d_{ik} = \sqrt{\sum_{j=1}^m (x_{ij} - x_{kj})^2} \quad (4.40)$$

где $i=1 \dots m$

$k=1 \dots n$

m -количество единиц

n -количество признаков.

Решение задач классификации объектов с использованием кластерного анализа проводится в определенной последовательности. Многомерный анализ делится на три этапа:

- составляется таблица исходной информации с указанием объектов и их признаков;
- проводится нормализация исходной информации с использованием среднего квадратического отклонения;
- по нормализованным данным рассчитываются таксономические расстояния, строится дендрограмма и проводится содержательная интерпретация полученных результатов.

Важным условием достоверности результатов, полученных в таксономическом анализе, является «ортогональность» данных, на основе которых проводится исследование, т.е. исходные показатели не должны иметь высоких значений корреляции между собой.

Следующей проблемой, с которой сталкивается исследователь при проведении кластерного анализа, является то, что в большинстве случаев анализируемые показатели, представляют собой несоизмеримые величины (т.е. одни показатели могут быть представлены в кубометрах, в другие в киловаттах на квадратный километр и т.д.). В этом случае возникает необходимость в процедуре стандартизации или нормировки.

Одним из способов подобного преобразования является нормирование с использованием среднего квадратического отклонения, как наиболее часто используемому в таксономическом анализе:

$$\hat{x}_{ij} = \frac{x_{ij} - \bar{x}_j}{\sigma_j} \quad (4.41)$$

где x_{ij} – исходные показатели;

\bar{x}_j - среднее арифметическое для исходного показателя;

\hat{x}_{ij} - нормированные значения исходных показателей;

σ_j - среднее квадратическое отклонение.

Нормированные показатели образуют таблицу нормированных показателей, идентичную таблице исходных показателей.

Все точки измерения показателей представляются в виде точек в многомерном m -мерном пространстве, в котором координатами служат нормированные исходные показатели.

Таксономические расстояния, соединяющие каждую пару точек (d_{ik}), отражают различия по комплексу показателей территориальных единиц.

Поэтому их можно положить в основу дифференциации (т. е. деления на группы) исходной совокупности точек. Они рассчитываются по формуле, координатами в которой являются нормированные показатели:

$$d_{ik} = \sqrt{\sum_{j=1}^m (\hat{x}_{ij} - \hat{x}_{kj})^2} \quad (4.42)$$

где $i=1 \dots m$

$k=1 \dots n$

Все рассчитанные расстояния образуют квадратную симметричную матрицу расстояний.

В прикладных географических исследованиях наиболее часто применяют метод «Вроцлавской таксономии».

Вроцлавский таксономический метод позволяет исследовать внутреннюю структуру совокупности индивидуальных точек.

Осуществляется это путем объединения близлежащих индивидуальных точек в группы; нахождения точек или ареалов, которые являются характерными для данной совокупности индивидуальных точек; деления генеральной совокупности на части.

Характерной особенностью многомерной статистической модели «Вроцлавской таксономии» является то, что она очень проста и удобна для расчетов, при которых не требуется сложная вычислительная техника.

По результатам расчетов по методу «Вроцлавской таксономии» строится дендрит, который затем делится на части, т.е. проводится группировка исходных точек.

Дендрит строится на базе данных матрицы таксономических расстояний по следующему алгоритму: из первого столбца выбирается наименьшее (но не равное 0) расстояние.

Соответствующие номера столбца и строки определяются номерами двух исследуемых точек, которые наносятся на диаграмму и соединяются линией. Длина линии пропорциональна расстоянию между точками. Соединение точек производится до тех пор, пока все точки не будут сведены в одну систему. Необходимо подчеркнуть, что в процессе построения дендрита линия, соединяющая пункты, не должна замыкаться в цепь.

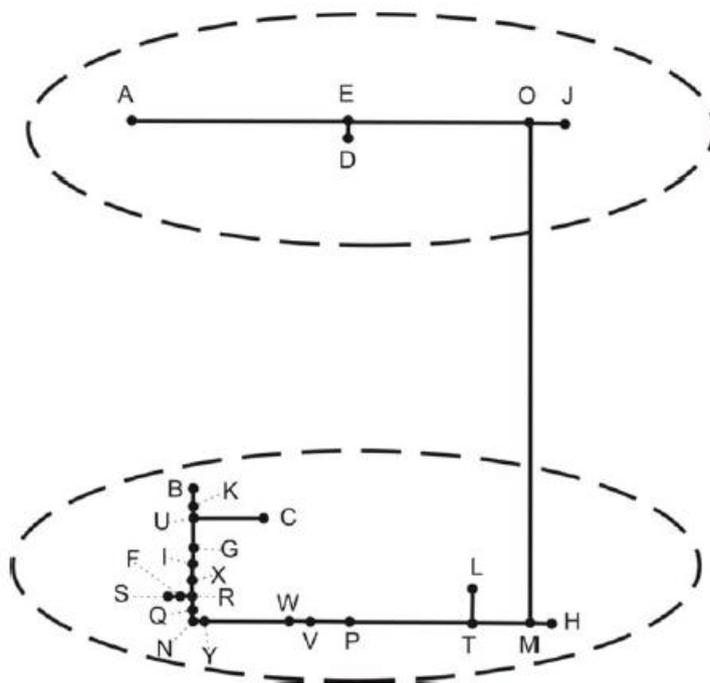


Рис.3. Выделение кластеров методом Вроцлавской таксономии

Наиболее простым способом выделения групп по построенному дендриту, является выделение наиболее длинных отрезков. Однако зачастую полученная конфигурация довольно сложна, поэтому отбирают некоторое условное расстояние, затем все точки отстоящие друг от друга на расстояние, меньшее обозначенного, объединяются в группы.

4.5. Анализ динамических рядов

Ряд расположенных в хронологической последовательности значений статистических показателей представляет собой **временной (динамический) ряд**.

Статистические показатели, характеризующие изучаемый объект, называют **уровнями ряда**. В динамическом ряду они могут быть абсолютными, относительными или средними величинами.

В зависимости от того, выражают уровни ряда состояние явления на определенные моменты времени (на начало месяца, квартала, года и т.п.) или его величину за определенные интервалы времени (например, за сутки, месяц, год и т.п.), различают, соответственно, **моментные и интервальные ряды**.

Ряды динамики могут быть с равноотстоящими (по времени) уровнями и неравноотстоящими (по времени) уровнями.

При построении динамических рядов следует учитывать следующие правила:

1. Все показатели одного динамического ряда должны относиться к равнозначным периодам времени;
2. Показатели динамического ряда должны быть однородны по составу, т. е. иметь одну и ту же полноту охвата объектов наблюдения;
3. Показатели должны быть рассчитаны по единой методологии;
4. При построении ряда динамики должна соблюдаться последовательность и непрерывность ряда.

Одним из важнейших направлений анализа рядов динамики является изучение особенностей развития явления за отдельные периоды времени.

Для выявления специфики развития изучаемых явлений за отдельные периоды времени определяют абсолютные и относительные показатели изменения ряда динамики: абсолютные приросты, темпы роста и прироста.

Рассматривая данные показатели, необходимо правильно выбирать базу сравнения, которая зависит от цели исследования.

При сравнении каждого уровня ряда с предыдущим получают **цепные** показатели: при сравнении каждого уровня с одним и тем же уровнем (базой) получают **базовые** показатели. Также используют средние показатели изменения динамического ряда.

Для выражения абсолютной скорости роста (снижения) уровня ряда динамики вычисляют статистический показатель – **абсолютный прирост** (Δy_t). Его величина определяется как разность двух сравниваемых уровней по формуле:

$$\Delta y_t = y_t - y_{t-1} \quad (4.43)$$

где y_t – уровень t -го года;

y_{t-1} – уровень базисного года.

Интенсивность изменения уровней ряда динамики оценивается отношением текущего уровня к предыдущему или базисному, которое всегда представляет собой положительное число. Этот показатель принято называть **темпом роста** (T_t). Он выражается в процентах и вычисляется по формуле:

$$T_t = \frac{y_t}{y_{t-1}} \times 100\% \quad (4.44)$$

Для выражения изменения величины абсолютного прироста уровней ряда динамики в относительных величинах определяется **темп прироста** (K_t), который рассчитывается по формуле:

$$K_t = T_t - 100\% \quad (4.45)$$

При наличии в динамических рядах значительной случайной ошибки (шума) применяют сглаживание или выравнивание путем укрупнения

интервалов и вычисления групповых средних. Этот метод позволяет повысить наглядность ряда, если большинство «шумовых» составляющих находятся внутри интервалов. Однако, если «шум» не согласуется с периодичностью, распределение уровней показателей становится грубым, что ограничивает возможности детального анализа изменения явления во времени.

Таблица 2

Показатели изменения динамических рядов

Вид показателя	Абсолютный прирост	Темп роста,%	Темп прироста,%
Цепной	$\Delta y_t = y_t - y_{t-1}$	$T_t = \frac{y_t}{y_{t-1}} \times 100\%$	$K_t = T_t - 100\%$
Базисный	$\Delta y_t^\sigma = y_t - y_\sigma$	$T_t^\sigma = \frac{y_t}{y_\sigma} \times 100\%$	$K_t^\sigma = T_t^\sigma - 100\%$
Средний	$\overline{\Delta y} = \frac{y_n - y_1}{n - 1}$	$\overline{T} = \sqrt[n-1]{\frac{y_n}{y_1}} \times 100\%$	$K_t = \overline{T}_t - 100\%$

Более точные характеристики получаются, если используют скользящие средние – широко применяемый способ для сглаживания показателей среднего ряда. Он основан на переходе от начальных значений ряда к средним в определенном интервале времени. В этом случае интервал времени при вычислении каждого последующего показателя как бы скользит по временному ряду.

Применение скользящего среднего полезно при неопределенных тенденциях динамического ряда или при сильном воздействии на показатели циклически повторяющихся выбросов (резко выделяющиеся варианты или интервенция).

Чем больше интервал сглаживания, тем более плавный вид имеет диаграмма скользящих средних. При выборе величины интервала сглаживания необходимо исходить из величины динамического ряда и содержательного смысла отражаемой динамики. Большая величина

динамического ряда с большим числом исходных точек позволяет использовать более крупные временные интервалы сглаживания (5, 7, 10 и т. д.).

Чаще всего величину интервала сглаживания принимают равной 3 или 5.

Вопросы для самоконтроля:

1. Какие группы показателей относятся к основным выборочным параметрам?
2. Объясните цель использования показателей разнообразия признаков?
3. Дайте определение понятию положительная и отрицательная асимметрия и положительный/отрицательный эксцесс?
4. Какие формы корреляционной связи Вы знаете?
5. Что такое ранговая корреляция? В чем ее отличие от обычной линейной парной корреляции?
6. Что такое регрессионный анализ? В чем состоит смысл его применения?
7. Назовите основные виды регрессионного анализа?
8. При решении каких задач наиболее часто используется таксономический анализ?
9. Что принимается за меру различия или сходства показателей в таксономическом анализе?
10. Каковы условия, необходимые для построения и исследования динамических рядов?
11. С помощью какого показателя определяется интенсивность изменения уровней ряда динамики?

Практическое упражнение 1

Используя статистические материалы сайта Федеральной службы государственной статистики <https://www.gks.ru/> создайте выборку показателей

средне-месячной номинальной начисленной заработной платы работников организаций по субъектам РФ.

На основе данных выборки:

1. Постройте интервальный вариационный ряд;

Для определения числа интервалов воспользуйтесь формулой Стерджесса:

$$k = 1 + 3,3 \lg n \quad (4.46)$$

где n - объем выборки

Для определения ширины интервала используйте формулу 3.1.

2. Рассчитайте основные характеристики ряда распределения (частота, частость, накопленная частота, накопленная частость)

3. По показателю частости постройте полигон и гистограмму, по показателю накопленной частости - кумуляту и огиву.

4. Сделайте вывод о форме распределения анализируемого показателя

Практическое упражнение 2

1. Используя интервальный вариационный ряд, построенный в предыдущем задании, определите показатели среднего значения ряда (средняя арифметическая, мода, медиана);

2. Используя исходный дискретный вариационный ряд определите:

- показатели разброса (среднее квадратическое отклонение, коэффициент вариации);

- показатели асимметрии и эксцесса.

3. По каждому из рассчитанных показателей вычислите ошибки и сделайте вывод.

Практическое упражнение 3

Используя статистические материалы сайта Федеральной службы государственной статистики <https://www.gks.ru/> с помощью коэффициента корреляции Спирмена определите зависимость между заболеваемостью

населения регионов одного из федеральных округов РФ и следующими показателями:

- Число больничных коек;
- Численность населения на одну больничную койку;
- Мощность амбулаторно-поликлинических организаций;
- Численность врачей всех специальностей;
- Нагрузка на работников сферы здравоохранения;
- Численность среднего медицинского персонал.

Для расчета коэффициентов корреляции воспользуйтесь следующим алгоритмом:

1. Сопоставьте каждому из признаков их порядковый номер (ранг) по возрастанию или убыванию.
2. Определите разности рангов каждой пары сопоставляемых значений.
3. Возведите в квадрат каждую разность и суммируйте полученные результаты.
4. Вычислите коэффициент корреляции рангов по формуле 4.23.
5. Определите статистическую значимость коэффициента при помощи таблицы критических значений при уровне значимости $\alpha=0,05$ и числе степеней свободы n (n - число элементов ряда).
6. Сравните полученные коэффициенты. Сделайте вывод.

Практическое упражнение 4

Используя статистические данные из предыдущего задания, постройте модель множественной регрессии, описывающую связь между заболеваемостью населения выбранного федерального округа РФ и показателями, продемонстрировавшими наличие корреляционной связи.

Для построения регрессионной модели воспользуйтесь следующим алгоритмом:

- а) Удалите показатели, продемонстрировавшие слишком тесную связь ($r>0,85$);

б) С помощью модуля «Анализ данных» Microsoft Excel рассчитайте уравнение регрессии, если известно, что форма зависимости между показателями линейная;

в) Интерпретируйте полученные результаты (оцените значения параметров коэффициента детерминации, значимость коэффициентов регрессии и уравнения регрессии);

г) Для оценки значимости коэффициентов регрессии и уравнения регрессии используйте данные t-статистики и F-статистики. При определении критических значений показателей воспользуйтесь таблицей критических значений, либо функцией F.ОБР.ПХ и СТЬЮДЕНТ.ОБР.2Х при числе степеней свободы $v_1 = k$, $v_2 = n - k - 1$ и уровне значимости $\alpha = 0,05$.

д) Сделайте вывод.

Практическое упражнение 5

Используя метод кластерного анализа, оцените территориальную дифференциацию уровня жизни населения одного из федеральных округов РФ.

При выполнении работы придерживайтесь следующего алгоритма:

1. На основе статистических материалов сайта Федеральной службы государственной статистики <https://www.gks.ru/> разработайте статистическую базу показателей для оценки уровня жизни региона. Статистическая база должны быть представлена совокупностью показателей, сгруппированных по блокам (не менее 5 в каждом блоке) в разрезе регионов федерального округа.

Рекомендуемые тематические блоки:

- демографический блок;
- экономический блок;
- социальный блок;
- экологический блок.

2. На основе отобранных показателей постройте дендрограмму в программе Statsoft Statistica 6.0

3. Интерпретируйте полученный результат и сделайте вывод.

Практическое упражнение 6

1. Используя статистические материалы сайта Федеральной службы государственной статистики <https://www.gks.ru/>, постройте динамический ряд на основе показателя численности населения одного из федеральных округов за период с 2007-2017 г.

- a) Вычислите значения абсолютного прироста, темпа роста и темпа прироста населения по сравнению с базисным 2007 г.;
- b) Полученные результаты занесите в таблицу;
- c) Постройте график динамики численности населения;
- d) Сделайте вывод.

ТЕМА 5. МЕТОДЫ МОДЕЛИРОВАНИЯ ТЕРРИТОРИАЛЬНЫХ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИСТЕМ

5.1. Понятие и виды математико-географических моделей.

Трудности моделирования географических объектов

Внедрение системной парадигмы в географию в 60-х годах 20 века способствовало широкому распространению методов моделирования при изучении территориальных систем.

Математические модели географических процессов и явлений получили название математико-географических.

Математико-географическое моделирование — это метод формализации географических представлений на основе создания логико-математических конструкций, отражающих количественные отношения реальных географических объектов.

Модели в географии выполняют разнообразные функции:

- a) психологическую (возможность изучения тех объектов и явлений, которые чрезвычайно трудно исследовать иными методами);
- b) собирательную (определение необходимой информации, ее сбор и систематизация);
- c) логическую (выявление и объяснение механизма развития конкретного явления);

d) систематизирующую (рассмотрение действительности как совокупности взаимосвязанных систем);

e) конструктивную (создание теорий и познание законов);

f) познавательную (содействие в распространении научных идей).

Множественность определения моделей и их функций приводит к появлению большого количества подходов к их классификации и типологии.

По целевому назначению математико-географические модели делятся на **теоретико-аналитические**, используемые в исследованиях общих свойств и закономерностей географических процессов, и **прикладные**, применяемые в решении конкретных задач (модели пространственного анализа, прогнозирования, управления).

Математико-географические модели могут предназначаться для исследования разных сторон народного хозяйства (в частности, его производственно-технологической, социальной, территориальной структур) и его отдельных частей.

При классификации моделей по исследуемым процессам и содержательной проблематике можно выделить модели народного хозяйства в целом и его подсистем, отраслей, регионов и т.д., комплексы моделей производства, потребления, формирования и распределения доходов, трудовых ресурсов и т. д.

По форме представления информации модели делятся на материальные и идеальные.

Материальные модели (субстратно-подобные, геометрические, аналоговые, изоморфные) традиционны в географии. Это различные карты и макеты, воспроизводящие природные и социально-экономические объекты.

Идеальные модели в зависимости от степени формализации делятся на неформализованные (концептуальные), частично формализованные и вполне формализованные (математические и информационные).

Концептуальные модели составляют фундамент любой науки. В географии наибольшее значение имеют такие концептуальные модели, как

теория зональности, учение о биосфере В. И. Вернадского, концепция геосистемы В. Б. Сочавы и др.

В частично формализованных моделях формализация информации осуществляется с помощью графических средств, рекомендаций, нормативных актов и т. п.

Полностью формализованные модели отличаются высокой степенью абстракции и использованием богатейшего аппарата прикладной математики.

В зависимости от соотношений, используемых в моделях, они делятся на **детерминистические** и **стохастические**.

Первые позволяют полностью предсказать развитие моделируемой системы во времени и пространстве, основываясь на известных условиях и соотношениях.

Стохастические (вероятностные) модели, напротив, основаны на случайных величинах. Они используются для обобщения событий, которые детерминированы различными факторами, а также для описания событий, на которые влияют случайные условия.

В соответствии с общей классификацией математических моделей они подразделяются на **функциональные** и **структурные**, а также включают промежуточные формы (**структурно-функциональные**).

В исследованиях на макроуровне чаще применяются структурные модели, поскольку для планирования и управления большое значение имеют взаимосвязи подсистем.

Типичными структурными моделями являются модели межотраслевых связей.

Функциональные модели широко применяются в решении задач управления, когда на поведение объекта («выход») воздействуют путем изменения «входа». Один и тот же объект может описываться одновременно и структурной, и функциональной моделью.

Различают модели **дескриптивные** и **нормативные**.

Первые объясняют наблюдаемые факты или дают вероятный прогноз, вторые — предполагают целенаправленную деятельность.

По соотношению экзогенных и эндогенных переменных, включаемых в модель, они могут разделяться на **открытые и закрытые**.

Полностью открытых моделей не существует; модель должна содержать хотя бы одну эндогенную переменную.

подавляющее большинство экономико-математических моделей занимает промежуточное положение и различается по степени открытости (закрытости).

В зависимости от того, включают ли математико-географические модели пространственные факторы и условия или не включают, различают модели **пространственные** (континуальные) и **точечные** (дискретные).

Внедрение методов моделирования в практику географических исследований связано с преодолением определенных трудностей, что главным образом обусловлено природой географических объектов и процессов, спецификой географической науки.

Длительное время главной трудностью практического применения математического моделирования в географии было наполнение разработанных моделей конкретной и качественной информацией.

Точность и полнота первичной информации, реальные возможности ее сбора и обработки во многом определяют выбор типов прикладных моделей. В то же время исследования по моделированию территориальных систем выдвигают новые требования к системе информации.

В зависимости от моделируемых объектов и назначения моделей используемая в них исходная информация имеет существенно различный характер и происхождение. Она может быть разделена на две категории: о прошлом развитии и современном состоянии объектов и о будущем развитии объектов, включая данные об ожидаемых изменениях их внутренних параметров и внешних условий (прогнозы).

Вторая категория информации является результатом самостоятельных исследований, которые также могут выполняться посредством моделирования. Для многих географических процессов выявление закономерностей возможно только на основе достаточно большого количества наблюдений.

Другая проблема порождается динамичностью географических процессов, изменчивостью их параметров и структурных отношений. Вследствие этого они должны постоянно находиться под наблюдением, обеспечивающим устойчивый поток обновляемых данных. Поскольку наблюдения за географическими процессами и обработка эмпирических данных обычно занимают довольно много времени, то при построении математических моделей экономики требуется корректировать исходную информацию с учетом ее запаздывания.

Познание количественных отношений географических процессов и явлений опирается на соответствующие измерения. Точность измерений в значительной степени предопределяет и точность конечных результатов количественного анализа посредством моделирования.

Поэтому необходимым условием эффективного использования математического моделирования является совершенствование системы географических показателей.

Применение математического моделирования заострило проблему измерений и количественных сопоставлений различных аспектов и явлений социально-экономического развития, достоверности и полноты получаемых данных, их защиты от намеренных и технических искажений.

5.2. Этапы математико-географического моделирования

Моделирование географических процессов, как правило, предусматривает несколько последовательных этапов.

На первом подготовительном этапе ставится цель и определяются задачи исследования. Этот этап включает выделение важнейших черт и свойств моделируемого объекта и абстрагирование от второстепенных;

изучение структуры объекта и основных зависимостей, связывающих его элементы; формулирование гипотез (хотя бы предварительных), объясняющих поведение и развитие объекта.

Второй этап включает согласование задач исследования, установление логической последовательности их решения и подготовку исходной информации. Моделирование предъявляет жесткие требования к системе информации. В то же время реальные возможности получения информации ограничивают выбор моделей, предназначенных для практического использования.

При этом принимается во внимание не только принципиальная возможность подготовки информации (за определенные сроки), но и затраты на подготовку соответствующих информационных массивов. Эти затраты не должны превышать эффект от использования дополнительной информации.

В процессе подготовки информации широко используются методы теории вероятностей, теоретической и математической статистики.

На третьем этапе осуществляется построение математической модели. Это этап формализации географической проблемы, выражения ее в виде конкретных математических зависимостей и отношений (функций, уравнений, неравенств и т. д.). Обычно сначала определяется основная конструкция (тип) математической модели, а затем уточняются детали этой конструкции (конкретный перечень переменных и параметров, форма связей).

Построение модели подразделяется, в свою очередь, на несколько стадий. Неправильно полагать, что чем больше фактов учитывает модель, тем она лучше «работает» и дает лучшие результаты. То же можно сказать о таких характеристиках сложности модели, как используемые формы математических зависимостей (линейные и нелинейные), учет факторов случайности и неопределенности и т. д.

Излишняя сложность и громоздкость модели затрудняют процесс исследования. Нужно учитывать не только реальные возможности информационного и математического обеспечения, но и сопоставлять затраты

на моделирование с получаемым эффектом (при возрастании сложности модели прирост затрат может превысить прирост эффекта).

После построения модели выясняются ее общие свойства. Здесь применяются чисто математические приемы исследования. Наиболее важный момент — доказательство существования решений в сформулированной модели (теорема существования).

Если удастся доказать, что математическая задача не имеет решения, то необходимость в последующей работе по первоначальному варианту модели отпадает; следует скорректировать либо постановку экономической задачи, либо способы ее математической формализации.

Решение задачи включает разработку алгоритмов для численного решения задачи, составления программ на ЭВМ и непосредственное проведение расчетов. Трудности этого этапа обусловлены прежде всего большой размерностью географических задач, необходимостью обработки значительных массивов информации.

Обычно расчеты по математико-географической модели носят многовариантный характер. Благодаря высокой эффективности современных ЭВМ удается проводить многочисленные «модельные» эксперименты, изучая поведение модели при различных изменениях некоторых условий.

Исследование, проводимое численными методами, может существенно дополнить результаты аналитического исследования, а для многих моделей оно является единственно осуществимым. Класс географических задач, которые можно решать численными методами, значительно шире, чем класс задач, доступных аналитическому исследованию.

На заключительном этапе (анализ результатов) встает вопрос об адекватности полученных данных реальному объекту и о степени их практической применимости.

Математические методы проверки могут выявлять некорректные построения модели и тем самым сужать класс потенциально правильных моделей. Неформальный анализ теоретических выводов и численных

результатов, получаемых посредством модели, сопоставление их с имеющимися данными также позволяют обнаруживать недостатки постановки экономической задачи, сконструированной математической модели, ее информационного и математического обеспечения.

5.3. Модели линейного программирования

Линейные модели активно используются в экономической географии для решения широкого спектра исследовательских задач.

Линейное программирование – это совокупность методов решения экстремальных задач, в которых цель (критерий оптимальности) и условия (ограничения) заданы уравнениями и неравенствами первой степени.

Задачи, решаемые методами линейного программирования, должны удовлетворять следующим требованиям:

- их решение не должно быть однозначным;
- иметь определенную целевую функцию, для которой ведется поиск максимального и минимального значения;
- иметь условия ограничения, формирующие область допустимых решений задачи.

Все модели линейного программирования состоят из стандартных составных частей:

- совокупность основных переменных;
- линейные ограничения (условия);
- целевая функция, определяющая критерий оптимальности задачи.

Совокупность основных переменных характеризует размеры землепользований, площади, объемы производства, затраты материальных, трудовых, финансовых ресурсов.

Система линейных ограничений (условий) определяет область допустимых значений основных переменных. Каждое отдельное условие

$$Z = \sum_{j=1}^N c_j x_j \rightarrow \max (\min); \sum_{j=1}^N a_{ij} x_j \leq b_i \quad (5.2)$$

$$i=1 \dots M, j=1 \dots N, x_j \geq 0$$

Для решения задач с применением линейного программирования разработан ряд алгоритмов, среди которых наиболее известны симплексный и распределительный методы. Оба метода базируются на последовательном улучшении первоначального плана путем повторения вычислений (итераций). После каждой итерации значение целевой функции улучшается.

Процесс повторяется до получения оптимального плана, а полученный план проверяется на оптимальность простыми критериями.

Распределительный метод среди задач линейного программирования получил распространение из-за упрощения расчетов, точности вычислений и снижения затрат времени на ввод исходной информации. Метод предложили А. Толстой и Л. В. Конторович в 1939 – 40 гг. Первоначально он применялся в задачах, связанных с транспортировкой грузов, их распределением между поставщиком и потребителем, поэтому получил название «транспортная задача». Применяется при решении ряда землеустроительных задач: распределение севооборотов и угодий по участкам, размещение культур на землях различных категорий, перераспределение участков между хозяйствами для экономии транспортных затрат и др.

Суть распределительной задачи следующая. Заданы m источников ресурса (производители продукции, базы с готовой продукцией) и n пунктов его потребления. Запасы ресурса в источниках составляют $A_i, i=1, \dots, m$, потребности – $B_j, j=1, \dots, n$. Стоимость транспортировки единицы ресурса от i -го источника к j -му потребителю C_{ij} . Количество ресурса, транспортируемого от i -го источника к j -му потребителю X_{ij} . Требуется определить такие значения X_{ij} , при которых общие транспортные расходы будут минимальны.

При сбалансированности, когда общий спрос на запас ресурса у поставщиков и общий спрос на него у потребителя равны, задачу называют закрытой:

$$\sum_{i=1}^n A_i = \sum_{j=1}^m B_j \quad (5.3)$$

Если баланс не выдерживается, то транспортная задача является открытой:

$$\sum_{i=1}^n A_i < \sum_{j=1}^m B_j \quad \text{или} \quad \sum_{i=1}^n A_i > \sum_{j=1}^m B_j \quad (5.4)$$

При наличии баланса модель транспортной задачи формулируется следующим образом.

Целевая функция:

$$Z = \sum c_{ij} x_{ij} \rightarrow \max (\min) \quad (5.5)$$

Условия. Ограничения по запасам:

$$\sum_{i=1}^n x_{ij} = A_i, i=1 \dots n \quad (5.6)$$

Ограничения по потребностям:

$$\sum_{j=1}^m x_{ij} = B_j, j=1 \dots m \quad (5.7)$$

Условие баланса:

$$\sum_{i=1}^n A_i = \sum_{j=1}^m B_j \quad (5.8)$$

Условие неотрицательности:

$$X_{ij} \geq 0, i = 1 \dots m, j = 1 \dots n. \quad (5.9)$$

Особенности распределительных транспортных задач следующие:

- условия задачи описываются уравнениями;

- все переменные выражаются в одних и тех же единицах измерения;
- во всех уравнениях коэффициенты при переменных равны единице;
- каждая переменная встречается только в двух уравнениях системы ограничений: в одном по строке (по запасам) и в одном по столбцу (по потребностям).

Целевая функция Z выражает суммарные расходы на транспортировку грузов. Ограничения по запасам и по потребностям означают, что сумма ресурса, забираемого из i -го источника, должна быть равна запасу ресурса в нем, как и сумма ресурса, доставляемого j -му потребителю, должна быть равна его потребности.

Величина C_{ij} может выражать транспортные расходы (минимизация) или прибыль от транспортных операций (максимизация) и другие показатели.

При решении транспортной задачи используются следующие обозначения:

i – индекс поставщика ($i = 1, 2, \dots, m$);

j – индекс потребителя ($j = 1, 2, \dots, n$);

a_i – мощность i -го поставщика;

b_j – спрос j -го потребителя;

C_{ij} – затраты на перевозку продукции от i -го поставщика j -му потребителю;

X_{ij} – количество продукции, которое необходимо перевезти от i -го поставщика j -му потребителю.

Исходная информация для решения транспортной задачи представлена в матрице

Потребители		B_1	B_2	B_n
Поставщики					
A_1					
A_2					
.....					
A_m					

Для решения транспортных задач разработан ряд методов: северо-западного угла, потенциала, дельта-метод, лямбда-задача.

Вопросы для самоконтроля:

1. Каковы основные функции моделей в экономико-географических исследованиях?
2. На какие группы можно разделить географические модели, исходя из их целевого назначения?
3. Приведите пример концептуальных моделей, используемых в географической науке?
4. С какими трудностями связан процесс внедрения методов моделирования в географические исследования?
5. Перечислите основные этапы экономико-географического моделирования?
6. Из каких составных частей состоит модель линейного программирования?

Практическое упражнение

С помощью модели линейного программирования составьте решение транспортной задачи, согласно условиям.

Имеются три пункта отправления A_1, A_2, A_3 однородного груза и пять пунктов B_1, B_2, B_3, B_4, B_5 его назначения. На пунктах A_1, A_2, A_3 груз находится в количестве 90, 70, 110 тонн соответственно. В пункты B_1, B_2, B_3, B_4, B_5 требуется доставить соответственно 50, 60, 50, 40, 70 тонн груза. Расстояние в

сотнях километров между пунктами отправления и назначения приведены в матрице.

Пункты отправления	Пункты назначения				
	B_1	B_2	B_3	B_4	B_5
A_1	9	1	1	5	6
A_2	6	4	6	8	5
A_3	2	9	3	5	3

Необходимо найти такой план перевозок, при котором общие затраты на перевозку грузов будут минимальными.

Указания: 1) считать стоимость перевозок пропорциональной количеству груза и расстоянию, на которое груз перевозится, т.е. для решения задачи достаточно минимизировать общий объем плана, выраженный в тонно-километрах; 2) для решения задачи использовать методы северо-западного угла или метод потенциалов.

ТЕМА 6. МЕТОДЫ ПРОГНОЗИРОВАНИЯ РАЗВИТИЯ ТЕРРИТОРИАЛЬНЫХ СОЦИАЛЬНО-ЭКОНОМИЧЕСКИХ СИСТЕМ

6.1. Методологические вопросы прогнозирования.

Классификация прогнозов

В географических исследованиях особое значение имеют методы прогнозирования, направленные на определение тенденций развития экономических систем и определение будущих параметров их функционирования.

Под **прогнозом** понимается научно обоснованное описание возможных состояний объектов в будущем, а также альтернативных путей и сроков достижения этого состояния.

Процесс разработки прогнозов называется **прогнозированием** (от греч. *prognosis* — предвидение, предсказание).

Прогнозирование должно отвечать на два вопроса:

- Что вероятнее всего ожидать в будущем?
- Каким образом нужно изменить условия, чтобы достичь заданного, конечного состояния прогнозируемого объекта?

Прогнозы, отвечающие на вопросы первого типа, называются **поисковыми**, второго типа — **нормативными**.

Важной характеристикой прогноза является **горизонт прогнозирования** — отрезок времени от момента, для которого имеются последние статистические данные об изучаемом объекте, до момента, к которому относится прогноз.

По горизонту прогнозирования прогнозы делятся на:

- оперативные (до одного месяца);
- краткосрочные (от одного, нескольких месяцев до года);
- среднесрочные (более 1 года, но не превышает 5 лет);
- долгосрочные (более 5 лет).

Наибольший практический интерес, безусловно, представляют краткосрочные и оперативные прогнозы.

Прогнозирование явлений и процессов включает в себя следующие этапы:

1. Постановка задачи и сбор необходимой информации;
2. Первичная обработка исходных данных;
3. Определение круга возможных моделей прогнозирования;
4. Оценка параметров моделей;
5. Исследование качества выбранных моделей, адекватности их реальному процессу и выбор лучшей из моделей;
6. Построение прогноза;
7. Содержательный анализ полученного прогноза.

Всю совокупность методов прогнозирования можно разделить на две большие группы:

1. Качественные (интуитивные) методы.
2. Количественные методы.

Качественные методы подразделяется на: индивидуальные и коллективные

Индивидуальные прогнозы могут делаться как отдельными квалифицированными специалистами, так и едиными группами, работающими как одно целое по общей программе.

В состав индивидуальных методов входят:

- **метод «интервью»**, при котором осуществляется непосредственный контакт эксперта со специалистом по схеме «вопрос – ответ»;
- **аналитический метод**, при котором осуществляется логический анализ какой-либо прогнозируемой ситуации;
- **сценарный метод**, который основан на определении логики процесса или явления во времени при различных условиях;
- **метод аналогов**, построенный на подборе объектов, аналогичных изучаемому, но которые уже прошли тот путь развития, который ещё предстоит изучаемому.
- **морфологический анализ**, состоящий в выделении всех независимых переменных системы, перечислении возможных значений этих переменных и генерировании альтернатив перебором всех возможных сочетаний этих значений.

Коллективные прогнозы даются группой экспертов, которые работают более или менее индивидуально. Как показывает практика коллективные методы эффективнее индивидуальных благодаря тому, что они позволяют предсказать больше вариантов развития, в т.ч. и неожиданных, которые большинству в голову не приходят. В то же время их сложнее организовать и согласовать между собой противоречивые прогнозы разных экспертов. Работа прогнозистов и организаторов процедуры в этом случае при прочих равных условиях более длительная и дорогостоящая.

Эти методы основаны на выявлении и обобщении различных высказываний и мнений экспертов, то есть ученых, специалистов,

потребителей и иных людей, которые обладают достаточными знаниями и квалификацией в изучаемой сфере, а также могут интуитивно чувствовать правильное решение.

Уровень точности, надежности и достоверности итоговых оценок во многом зависит от количества экспертов, степени их компетентности в рассматриваемых вопросах, методов проведения экспертизы, согласованности различных мнений и других факторов.

Наиболее эффективными методы экспертных оценок оказываются тогда, когда исследуемые процессы трудно формализовать и описать количественно, а также в ситуациях, когда временные и финансовые ресурсы сильно ограничены.

Основными из этих методов считаются:

— **Метод «лицом к лицу»** - состоит в опросе экспертов, который проводится в группе численностью от 10 до 30 экспертов.

— **Метод комиссий** отличается от предыдущего тем, что группа экспертов многократно собирается для обсуждения одного и того же вопроса, поэтому число участников не превышает 12. При использовании метода комиссий предварительно разрабатывается программа обсуждения.

— **Метод Дельфи** проводится в заочной форме. Работает группа экспертов до 20 человек. Эксперты на каждом этапе работают независимо друг от друга, предлагая свои оценки для поставленной проблемы. После этого анкеты или ответы собираются, выбирается только информация, непосредственно относящаяся к цели, и статистически усредняется. Среднее значение и обоснования предъявляются всем экспертам. Эта последовательность повторяется примерно три-четыре раза.

— **Метод мозгового штурма**. Начинается как обычно с отбора участников 10-12 человек. Процедура начинается с постановки проблемы. Основная задача - генерация идей по её решению. В отличие от ранее рассмотренных методов здесь не только допускаются, но и приветствуются необычные и даже абсурдные идеи.

Количественные методы пригодны для прогнозирования тех процессов, для которых возможно количественное представление и имеется хоть какая-то теория, которая позволяет моделировать их протекание. В случае субглобальных и более сложных прогнозов эти методы менее затратны.

Вероятность их осуществления зависит от степени изученности моделируемых объектов.

Среди количественных методов можно, в свою очередь, выделить:

- Статистические
- Имитационные
- Натурное моделирование.

Большинство **статистических методов** построены на анализе временных рядов и являются экстраполяцией прошлого в будущее.

Корректное применение экстраполяции требует однородности анализируемого ряда данных, иными словами, отсутствие в них скачков и революционных изменений. Кроме того, ряд должен быть достаточно длинным. В географических прогнозах эта продолжительность зависит от характерного времени процесса, но в любом случае желательно, чтобы ряд состоял хотя бы из 20 наблюдений, соответствующих скорости протекания процесса.

Имитационное моделирование основано на разработке математических моделей поведения сложных систем. Здесь особенно важно знание характера связей и взаимодействий между входными переменными и переменными внутри системы.

Если они отражены адекватно, то результаты получаются хорошие.

Если хорошей теории нет, то используют стохастические модели, основанные на многократном проигрывании ситуации с разной заданной вероятностью реализации тех или иных событий с последующим выявлением наиболее вероятного итогового поведения системы.

Натурное моделирование используется там, где возможно создание уменьшенной копии природного или техногенного объекта.

Прежде чем приступить к процессу построению моделей прогнозирования и получению прогнозных оценок необходимо провести проверку соответствия исходной информации временного ряда.

Одним из важнейших условий, необходимых для правильного отражения временным рядом реального процесса развития, является сопоставимость уровней ряда.

Появление несопоставимых уровней может быть вызвано разными причинами: изменением методики расчета показателя, изменением классификации, терминологии и т. д.

Несопоставимость может возникнуть вследствие территориальных изменений, например, как результат изменения границ области, района, страны.

Другой причиной несопоставимости могут служить структурные изменения. Например, произошло укрупнение нескольких ведомств путем слияния их в единое целое или укрупнение производства за счет слияния нескольких предприятий в одно объединение.

В большинстве случаев удастся устранить несопоставимость, вызванную указанными причинами, путем пересчета более ранних значений показателей с помощью формальных методов.

Для успешного изучения динамики процесса важно, чтобы информация была полной, временной ряд имел достаточную длину (с учетом конкретных целей исследования).

Во временных рядах не должны содержаться пропущенные уровни.

Пропуски могут объясняться как недостатками при сборе информации, так и происходившими изменениями в системе отчетности, в системе фиксирования данных.

В этом случае для использования этого временного ряда в дальнейшем анализе необходимо восстановить пропущенные уровни одним из известных способов восстановления пропусков (выбор метода зависит от специфики конкретного временного ряда).

Если же в систему показателей включен новый признак, учет которого не проводился ранее, то необходимо подождать, пока ряд достигнет требуемой длины или попытаться восстановить прежние значения косвенными методами (через другие показатели), если такой путь представляется возможным.

Одним из способов определения недостающего элемента в ряду называется интерполяция.

Интерполяцией называют приблизительный расчет недостающего уровня, находящегося внутри ранжированного динамического ряда (или внутри однородного периода колеблющегося ряда).

Выявление, исключение аномальных значений, замена их истинными или расчетными является необходимым этапом первичной обработки данных, т.к. применение математических методов к «засоренной» информации приводит к искажению результатов анализа.

6.2. Методы прогнозирования на основе анализа тенденции временных рядов и уравнения регрессии

В основе прогнозирования географических явлений на основе анализа тенденции временных рядов лежит понятие экстраполяции прошлого в будущее.

Под **экстраполяцией** понимается приближенный расчет неизвестных уровней динамического ряда, лежащих за его пределами.

При экстраполяции исходят из предположения, что характер динамики, т.е. та или иная закономерность развития, имевшая место в течение известного периода, сохранится также и в будущем (**перспективная экстраполяция** – прогнозирование) или имела место в прошлом (**ретроспективная экстраполяция**).

Важное значение при экстраполяции имеют вопросы о ее базе расчета и сроках прогнозирования.

База расчета, определяемая на основе теоретического анализа сущности изучаемого явления, должна быть оптимальной по своей длительности.

В качестве такой базы нельзя использовать очень короткий или очень длинный период.

При экстраполяции на практике берут те субпериоды базисного ряда динамики, которые составляют определенный этап в развитии изучаемого явления в конкретных исторических условиях.

Установление сроков прогнозирования зависит от задачи исследования.

Но следует иметь в виду, что чем короче сроки прогноза, тем надежнее результаты экстраполяции.

Обычно рекомендуют, чтобы срок прогноза не превышал одной трети базы расчета.

Применение методов экстраполяции зависит от характера изменений в базисном ряду динамики и предопределяется постановкой задачи исследования.

Экстраполяция может производиться на основе следующих методов:

1) Прогнозирование на основе среднего абсолютного прироста

Прогнозирование методом среднего абсолютного прироста предполагает, что уровни ряда динамики изменяются равномерно (линейно). Применение данного метода прогнозирования возможно при предварительной проверке следующих предпосылок:

1) Абсолютные цепные приросты должны быть приблизительно одинаковыми;

2) Должно выполняться неравенство вида:

$$\sigma_{ост}^2 \leq \rho^2$$

где $\sigma_{ост}^2$ - остаточная дисперсия:

$$\sigma_{ост}^2 = \frac{\sum_{t=1}^n (y_t - \bar{y})^2}{n} \quad (6.1)$$

$$\rho^2 = \frac{1}{2} \cdot \frac{\sum_{t=1}^n \Delta_t^2}{n} \quad (6.2)$$

где Δ_t - цепные абсолютные приросты исходного ряда динамики.

$y_{\bar{d}}$ - теоретические значения уровней ряда, выровненные методом среднего абсолютного прироста.

После проверки и подтверждения выполнения данных предпосылок можно приступить к прогнозированию. Общая модель прогноза имеет вид:

$$y_{t+L} = y_t + \bar{\Delta} \cdot L \quad (6.3)$$

y_t - последний уровень исходного ряда динамики (для перспективного прогноза) или уровень, принятый за базу экстраполяции (во всех остальных случаях)

$\bar{\Delta}$ - средний абсолютный прирост:

$$\bar{\Delta} = \frac{y_t - y_1}{n-1} \quad (6.4)$$

L - горизонт прогнозирования;

y_t - последний уровень исходного ряда динамики;

y_1 - первый уровень исходного ряда динамики;

n - число уровней ряда.

2) Прогнозирование на основе среднего темпа роста

Данный метод применяется, если развитие ряда динамики подчиняется геометрической прогрессии и может быть описан экспоненциальной (показательной) кривой.

Общая модель прогноза имеет вид:

$$y_{t+L} = y_t \cdot (\bar{T})^L \quad (6.5)$$

y_t - последний уровень исходного ряда динамики (для перспективного прогноза) или уровень, принятый за базу экстраполяции (во всех остальных случаях)

\bar{T} - средний темп роста, который определяется как:

$$\bar{T} = \sqrt[n-1]{\frac{y_t}{y_1}} \quad (6.6)$$

y_t - последний уровень исходного ряда динамики;

y_1 - первый уровень исходного ряда динамики;

n - число уровней ряда.

В качестве базового уровня экстраполяции берется последний уровень ряда динамики, так как будущее развитие начинается именно с этого уровня.

Прогнозирование с помощью уравнения регрессии

Регрессионные модели могут быть использованы для прогнозирования возможных ожидаемых значений зависимой переменной.

Прогнозируемое значение переменной y получается при подстановке в уравнение регрессии ожидаемой величины фактора x :

$$y_{np} = ax_{np} + b \quad (6.7)$$

Данный прогноз называется точечным. Значение независимой переменной x_{np} не должно значительно отличаться от значений, входящих в выборку, по которой вычислено уравнение регрессии.

Вероятность реализации точечного прогноза теоретически равна нулю. Поэтому рассчитывается средняя ошибка или доверительный интервал прогноза с достаточно большой надежностью.

Доверительные интервалы зависят от следующих параметров:

- Стандартная ошибка s_e ;
- Удаление x_{np} от своего среднего значения \bar{x} ;
- Количество наблюдений n ;
- Уровень значимости прогноза α .

В частности, для прогноза $y_{np} = ax_{np} + b$ будущие значения y_{np} попадут в доверительный интервал:

$$y_{np} \in \left[y_{np} - s_{et\alpha} \sqrt{1 + \frac{1}{n} + \frac{(x_{np} - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}} \right];$$

$$y_{np} + S_{et_\alpha} \sqrt{1 + \frac{1}{n} + \frac{(x_{np} - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}} \quad (6.8)$$

Значения e_i и S_e рассчитываются по формулам 4.29 и 4.36 соответственно.

Коэффициент Стьюдента t_α рассчитывается по таблице критических значений Стьюдента для числа степеней свободы $\nu = n-2$ при заданном уровне значимости α .

Расположение границ доверительного интервала показывает, что прогноз значений зависимой переменной по уравнению регрессии хорош только в случае, если значение фактора x не выходит за пределы выборки.

Вопросы для самоконтроля:

1. В чем разница между поисковыми и нормативными прогнозами?
2. Каков горизонт прогнозирования у краткосрочных прогнозов?
3. В чем суть качественных (интуитивных) методов прогнозирования?
4. Для каких географических процессов целесообразно использовать количественные методы прогнозирования?
5. Каким условиям должен удовлетворять временной ряд для качественного прогнозирования?
6. С какой целью рассчитывается доверительный интервал при прогнозировании на основании уравнения регрессии?

Практическое упражнение

Составьте прогноз изменения урожайности зерновых и зернобобовых культур в зависимости от количества внесенных удобрений, используя уравнение регрессии.

С помощью данных сайта Федеральной службы государственной статистики <https://www.gks.ru/> постройте однофакторную модель линейной регрессии между показателями «Внесение органических удобрений на один

гектар посева сельскохозяйственных культур, т» и «Урожайность зерновых и зернобобовых культур, ц с 1 га».

Для расчетов используйте модуль «Анализ данных» Microsoft Excel.

Оцените, как изменится урожайность (y_{np}), если количество внесенных органических удобрений (x_{np}) составит 1,5 т, используя полученное уравнение регрессии вида: $y_{np} = ax_{np} + b$.

Рассчитайте доверительный интервал прогноза при уровне значимости $\alpha=0,05$.

ЗАКЛЮЧЕНИЕ

Рассмотренный в пособии материал направлен на ознакомление студентов с подходами и методами, наиболее часто применяемыми в экономико-географических исследованиях, а также практическое применение полученных знаний при решении типовых задач.

Изучение основных теоретических положений, представленных в пособии, позволит сформировать целостное представление о происхождении, возможностях и ограничениях конкретных методов общегеографических и экономико-географических исследований, раскроет современный инструментарий и методики анализа территориальных социально-экономических систем.

Данное пособие призвано оказать помощь студентам при прохождении учебной и производственной практик, подготовке к экзамену или зачёту, написанию курсовых и дипломных проектов.

Решение практических упражнений, представленных в конце каждой теоретической главы, будет способствовать закреплению полученных навыков и формированию профессиональных компетенций, связанных с применением методов в географических исследованиях для обработки, анализа и синтеза географической информации, а также выработки дальнейших рекомендаций по осуществлению эффективной региональной политики.

СПИСОК ЛИТЕРАТУРЫ

1. Бочаров М.К. Методы математической статистики в географии. — М.: Мысль, 1971 — 370 с.
2. Громько Г. Л. Статистические ряды в экономических и экономико-географических исследованиях (теоретические и методические аспекты). — М.: МГУ, 1974. — 264 с.
3. Гетманчук А. В. Экономико-математические методы и модели: Учебное пособие для бакалавров / А. В. Гетманчук, М. М. Ермилов. — М.: Издательско-торговая корпорация «Дашков и К», 2017. — 188 с.
4. Гриценко В. А., Белосевич Е.В., Артищева Е. К. Математические методы в географии: Учебное пособие. — Калининград, 1999. — 75 с.
5. Дружинин А. Г., Угольницкий Г.А. Устойчивое развитие территориальных социально-экономических систем: теория и практика моделирования: монография. — Москва: Вузовская книга, 2013. — 224 с.
6. Дьяконов К.Н., Касимов Н.С., Тикунов В.С. Современные методы географических исследований. — М.: Просвещение, 1996. — 207 с.
7. Екеева Э.В. Методы географических исследований: учебное пособие. — Горно-Алтайск: РИО ГАГУ, 2010. — 48 с.
8. Жуков В. Т., Сербенюк С.Н., Тикунов В.С. Математико-картографическое моделирование в географии. М.: «Мысль», 1980. — 224 с.
9. Звонкова Т.В. Географическое прогнозирование. — М., 1987. — 190 с.
10. Иванова М. Б. Математические методы в социально-экономической географии. — Пермь, 2007. — 315 с.
11. Игонин А. И., Тикунов В. С. Математико-картографическое моделирование и картографирование демографической ситуации в регионах Европы и России // Наука. Инновации. Технологии. — 2019. — № 3. — С. 115–130.
12. Капитанова О.В. Прогнозирование социально-экономических процессов: Учебно-методическое пособие. — Нижний Новгород, 2016. — 74 с.

13. Колеснев В. И., Шафранская И. В. Экономико-математические методы и моделирование в землеустройстве. Практикум: учеб. пособие. Минск, 2007. — 392 с.
14. Лебедева И. М., Федорова А. Ю., Макроэкономическое планирование и прогнозирование. — СПб: Университет ИТМО, 2016. — 54 с.
15. Манак Б.А. Методика экономико-географических исследований. — Минск, 1985. —157 с.
16. Методы экономико-географических исследований /Б. А. Красноярова, Е. П. Крупочкин, Е. В. Мардасова. - Барнаул: Изд-во АлтГУ, 2014. — 189 с.
17. Орлова И. В., Половников В.А. Экономико-математическое методы и модели: компьютерное моделирование: Учебное пособие. — М.: Вуз. учебник: ИНФРА-М, 2010. — 366 с.
18. Пузаченко Ю. Г. Математические методы в экологических и географических исследованиях. — М.: Академия, 2004. —416 с.
19. Статистические методы прогнозирования в экономике: Учебное пособие, практикум, тесты, программа курса / Дуброва Т.А., Архипова М. Ю. — М., 2004. — 136 с.
20. Третьяков А. С. Статистические методы в прикладных географических исследованиях: Учебно-методическое пособие. — Харьков: Шрифт, 2004. — 96 с.
21. Филандышева Л.Б., Сапьян Е.С. Статистические методы в географии: учебно-методическое пособие. — Томск.: Издательский Дом ТГУ, 2015. — 216 с.
22. Червяков В. А. Количественные методы в географии. — Барнаул: Изд-во АГУ, 1998. —259 с.
23. Чертко Н. К. Математические методы в географии: пособие для студентов геогр. фак. / Н. К. Чертко, А. А. Карпиченко. — Минск: БГУ, 2008. —202 с.

ТЕСТОВЫЕ ЗАДАНИЯ

1. Основными свойствами ТСЭС являются:
 - a) территориальность и открытость;
 - b) статичность и детерминированность;
 - c) комплексность и закрытость

2. В основе формирования ТСЭС лежит:
 - a) учение о зональности;
 - b) географическое разделение труда;
 - c) концепция геосистемы В. Б. Сочавы

3. Укажите неверное утверждение:
 - a) Формирование и развитие ТСЭС осуществляется непрерывно;
 - b) Наиболее динамичной и доминирующей подсистемой ТСЭС является производственная подсистема;
 - c) Выделение ТСЭС происходит в несколько этапов;
 - d) Понятие ТСЭС рассматривается только на макроуровне

4. Какой из перечисленных методов используется для сопоставления социально-экономических систем во времени и пространстве?
 - a) сравнительно-географический;
 - b) картографический;
 - c) метод наблюдения;
 - d) метод описания

5. К методам систематизации и анализа географической информации не относится:
 - a) сравнительно-географический;
 - b) геоинформационный;
 - c) дистанционный;

d) картографический

6. Какая группа методов позволяет выявить статистическую взаимосвязь между социально экономическими системами:

- a) картографические методы;
- b) дистанционные методы;
- c) количественные методы;
- d) методы систематизации

7. Множество индивидуально различных объектов, обладающих общими свойствами, представляет собой:

- a) статистическую совокупность;
- b) ряд распределения;
- c) вариационный ряд;
- d) ранжированный список

8. Группировка, выявляющая взаимосвязи между изучаемыми явлениями и их признаками называется:

- a) типологической;
- b) структурной;
- c) аналитической;
- d) комбинационной

9. Группировка населения по полу, возрасту, уровню образования, называется:

- a) типологической
- b) структурной
- c) аналитической
- d) комбинационной

10. Укажите неверное утверждение:

- a) В основание группировки могут быть положены только количественные признаки;
- b) Число групп зависит от задач исследования и вида показателя, положенного в основание группировки, численности совокупности, степени вариации признака;
- c) При небольшом объеме совокупности не следует образовывать большое число групп

11. Укажите неверное утверждение:

- a) Чем меньше объем выборки, тем лучше она отражает генеральную совокупность;
- b) Выборочная совокупность должна полно и правильно отражать свойства генеральной совокупности;
- c) Оптимальный объем выборки пропорционален степени изменчивости признака;
- d) В большинстве случаев достаточно точные результаты получаются при объеме выборки не менее 30-35 вариантов;
- e) Каждая единица генеральной совокупности имеет одинаковую возможность быть отобранной

12. Отбор элементов выборки, при котором единицы, подлежащие изучению, берутся через определенный интервал называется:

- a) серийный отбор;
- b) механический отбор;
- c) случайный отбор

13. К основным характеристикам ряда распределения не относится:

- a) частота;
- b) накопленная частота;

- c) частость;
- d) средняя арифметическая

14. Укажите неверное утверждение:

- a) Частоты всегда являются положительными числами;
- b) Сумма частостей всегда равна 1, если они выражены в долях единицы, или 100%, если они выражены в процентах;
- c) Сумма частот ряда распределения совпадает с объемом совокупности;
- d) Накопленные частоты определяются путем последовательного умножения частот первой группы этих показателей и последующих групп ряда распределения

15. По качественным признакам строятся:

- a) атрибутивные ряды распределения;
- b) вариационные ряды распределения;
- c) дискретные ряды распределения;
- d) интервальные ряды распределения

16. Линия, соединяющая точки, соответствующие значениям варьирующего признака и их частотам называется:

- a) полигон
- b) огива
- c) кумулята
- d) гистограмма

17. Для графического изображения ряда распределения с помощью кумуляты:

- a) на оси абсцисс наносятся значения вариант, а по оси ординат – накопленные частоты;

- b) на оси абсцисс наносятся значения частот, а по оси ординат – значения вариант;
- c) на оси абсцисс наносятся значения накопленных частот, а по оси ординат – значения вариант;
- d) на оси абсцисс наносятся середины интервалов, а по оси ординат – накопленные частоты

18. К показателям среднего положения или центральной тенденции вариационного ряда не относится:

- a) среднее квадратическое отклонение;
- b) мода;
- c) медиана;
- d) средняя

19. Наиболее часто встречающийся вариант ряда – это:

- a) медиана
- b) мода
- c) среднее квадратическое отклонение;
- d) эксцесс

20. Как определить медиану, если число членов ряда четное:

- a) за медиану условно принимается среднее значение, полученное из двух величин, занимающих центральное положение в ранжированном ряду;
- b) медианой является значение члена ряда, занимающего центральное положение;
- c) медианой является разность двух величин, занимающих центральное положение в ранжированном ряду;

21. Среднее квадратическое отклонение (σ_x) показывает:

- a) степень рассеяния значений статистической совокупности около среднего значения;
- b) наиболее часто встречающийся вариант ряда;
- c) степень «сглаженности» («остро-» или «плосковершинности») распределения

22. Какой из показателей разнообразия признака является безразмерной величиной:

- a) среднеквадратическое отклонение;
- b) коэффициент вариации;
- c) амплитуда;
- d) среднее абсолютное отклонение

23. Второе название кривой нормального распределения статистического ряда:

- a) Кривая Гаусса;
- b) Кривая Лейбница;
- c) Кривая Эйлера

24. Если значение эксцесса близко к 0, это означает, что форма распределения:

- a) близка к нормальному виду;
- b) является островершинной;
- c) является плосковершинной

25. Для определения степени «нормальности» распределения используют правило:

- a) 3-х сигм;
- b) 3-х медиан;
- c) 3-х дисперсий

26. Какая характеристика динамического ряда позволяет определить на сколько процентов изменился показатель текущего периода по сравнению с базисным:

- a) абсолютный прирост;
- b) темп роста;
- c) темп прироста

27. Для сглаживания динамического ряда методом скользящей средней, как правило, используют временной промежуток:

- a) 1- год;
- b) 2 года;
- c) 3-5 лет

28. Какие из перечисленных задач не решает корреляционный анализ:

- a) установление направления и формы связи;
- b) оценка тесноты связи;
- c) математическое выражение зависимостей

29. Укажите неверное утверждение:

- a) Чем больше разброс точек в поле корреляции, тем меньше теснота связи;
- b) При прямой корреляции более высоким значениям одного признака соответствуют более высокие значения другого;
- c) По количеству коррелируемых признаков корреляция может быть слабой, средней или сильной

30. Для определения степени тесноты связи между коррелируемыми признаками используют:

- a) шкалу Чеддока;
- b) шкалу Спирмена;

с) шкалу Пирсона

31. При нелинейной зависимости между признаками рассчитывается:

- а) коэффициент корреляции Пирсона;
- б) корреляционное отношение;
- с) коэффициент множественной корреляции

32. Величина коэффициента линейной корреляции Пирсона изменяется в пределах:

- а) от -1 до 1;
- б) от 0 до 1;
- с) от -1 до 0

33. Ранговый коэффициент корреляции рассчитывается, когда объем выборки:

- а) не менее 5 и не более 30 элементов;
- б) не менее 10 и не более 50 элементов;
- с) не менее 20 и не более 100 элементов

34. Укажите неверное утверждение:

- а) Регрессионный анализ возможен при наличии большого количества пар сопряженных наблюдений;
- б) Регрессионный анализ возможен при условии сильных связей между признаками ($r \geq 0,7$);
- с) Если при установлении зависимости между признаками используется больше одной независимой переменной, то применяют множественный регрессионный анализ

35. Для линейного уравнения регрессии $y = ax + b + \xi$ установите соответствие:

a) Независимая переменная	1. y
b) Зависимая переменная	2. a
c) Коэффициент регрессии	3. ξ
d) Случайная компонента	4. x

36. Установите соответствие:

a) Нелинейная множественная регрессия	1. $y = ax + b + \xi$
b) Линейная множественная регрессия	2. $y = a_1x_1 + a_2x_2^2 + b + \xi$
c) Парная линейная регрессия	3. $y = a_1x_1 + a_2x_2 + b + \xi$

37. Дана матрица парных коэффициентов корреляции

	y	x ₁	x ₂	x ₃
y	1			
x ₁	0,72	1		
x ₂	0,48	0,11	1	
x ₃	-0,21	-0,79	-0,51	1

Установите соответствие

a) Связь прямая и сильная	1. x ₁ , x ₃
b) Связь прямая и слабая	2. x ₁ , x ₂
c) Связь обратная и сильная	3. y, x ₃
d) Связь обратная и слабая	4. y, x ₁

38. Факторы, включаемые в модель линейной множественной регрессии должны удовлетворять следующим требованиям (два варианта ответа):

- число факторов должно быть меньше объема совокупности;
- факторы должны представлять временные ряды;

- c) факторы должны иметь одинаковую размерность
- d) между факторами не должно быть слишком высокой корреляции

39. Оцените степень тесноты связи между признаками, если коэффициент корреляции равен 0,33:

- a) слабая;
- b) сильная;
- c) умеренная;

40. Зависимость между урожайностью пшеницы y от количества внесенных удобрений x характеризуется следующим образом: $y=10,6+0,6x+\zeta$, $r_{yx}=0,83$

Дайте интерпретацию коэффициенту регрессии.

- a) 60% вариации урожайности объясняется вариацией внесенных удобрений;
- b) 83% вариации урожайности объясняется вариацией внесенных удобрений;
- c) при увеличении внесенных удобрений на 1 единицу урожайность увеличивается на 0,83;
- d) при увеличении внесенных удобрений на 1 единицу урожайность увеличивается в среднем на 0,6 единиц

41. Какой метод прогнозирования предполагает использование уменьшенной копии природного или техногенного объекта:

- a) имитационное прогнозирование;
- b) натурное моделирование;
- c) экспертное прогнозирование

42. Если коэффициент регрессии является существенным, то фактическое значение статистического критерия:

- a) больше критического (табличного);
- b) меньше критического (табличного);
- c) близко к единице;
- d) близко к нулю

43. Какой из перечисленных методов не относится к методам кластерного анализа:

- a) Метод Вроцлавского дендрита;
- b) Агломеративно-иерархический метод;
- c) Метод цепных подстановок

44. Прогноз, для которого горизонт прогнозирования превышает 5 лет, относится к:

- a) долгосрочным;
- b) краткосрочным;
- c) среднесрочным

45. К индивидуальным прогнозам относят:

- a) Метод «интервью»;
- b) Метод Дельфи;
- c) Метод мозгового штурма;
- d) Натурное моделирование

46. Метод прогнозирования на основе среднего темпа роста применяется в случае, если;

- a) уровни ряда динамики изменяются линейно;
- b) развитие ряда динамики описывается экспоненциальной (показательной) кривой;
- c) уровни ряда динамики изменяются незначительно

47. Уравнение вида $y_{t+L} = y_t + \bar{\Delta} \cdot L$ соответствует модели прогноза на основе:

- a) среднего абсолютного прироста;
- b) среднего темпа роста;
- c) регрессии

48. Какой из перечисленных терминов определяет основную тенденцию развития временного ряда:

- a) тренд;
- b) экстраполяция;
- c) интерполяция

49. Отрезок времени от момента, для которого имеются последние статистические данные об изучаемом объекте, до момента, к которому относится прогноз, называется:

- a) горизонтом прогнозирования;
- b) периодом наблюдения;
- c) ретроспективным участком

50. Если значения цепных абсолютных приростов временного ряда примерно одинаковы, то для вычисления прогнозного значения в следующей точке корректно использовать:

- a) средний абсолютный прирост;
- б) средний темп роста;
- в) средний темп прироста

ГЛОССАРИЙ

Адекватность модели — соответствие модели моделируемому объекту или процессу.

Анализ — метод исследования, характеризующийся выделением и изучением отдельных частей объектов исследования.

Аналитическое исследование – самый углубленный вид социологического анализа, ставящего своей целью не только описание структурных элементов изучаемого явления, но и выяснение причин, которые лежат в его основе и обуславливают характер, распространенность, устойчивость или изменчивость и другие свойственные ему черты.

Аэрофотоснимки — мгновенные изображения всех деталей земной поверхности с большой высоты.

Балансовые методы — группа расчетных методов для анализа, планирования и прогнозирования развития динамических систем с установившимися потоками ресурсов и продукции и с детерминированными зависимостями между приходной и расходной частями.

Временной ряд — ряд наблюдений за значениями некоторого показателя (признака), упорядоченный в хронологической последовательности. Отдельные наблюдения временного ряда называются уровнями этого ряда.

Выборочная совокупность — часть генеральной совокупности, отобранная для выборочного наблюдения.

Генеральная совокупность — совокупность всех возможных наблюдений, которые могли бы быть проведены в соответствии с целью исследования. Общее число членов генеральной совокупности называют объемом генеральной совокупности.

Географический детерминизм — концепция, утверждающая, что процесс общественного развития - это не результат проявления объективных закономерностей развития общества, а следствие влияния природных сил.

Географическое наблюдение — способ получения географической информации, сущность которого заключается в сборе первичной информации о географическом объекте путем непосредственного восприятия и выявления его характеристик.

Географическое описание — упорядоченная характеристика территории, а также теоретическое обобщение полученного материала, т. е. систематизация, объяснение и построение теории.

Географическое прогнозирование — специальное научное исследование конкретных перспектив развития географических явлений.

Географическое разделение труда — разделение труда между отдельными географическими местностями, выражающееся в специализации этих местностей на производстве определенных видов промышленной продукции.

Геоинформационные методы — сравнительно новые методы исследования, основанные на применении ГИС-технологий.

Горизонт прогнозирования — отрезок времени от момента, для которого имеются последние статистические данные об изучаемом объекте, до момента, к которому относится прогноз.

Группировка объектов — объединение объектов, в каком-либо отношении сходных или как-либо зависящих друг от друга.

Дескриптивная модель — модель, предназначенная для описания и объяснения наблюдаемых фактов или прогноза поведения объектов.

Детерминированная модель — аналитическое представление закономерности, операции и т. п., при которых для данной совокупности входных значений на выходе системы может быть получен единственный результат.

Дистанционные методы — изучение объекта на расстоянии, без непосредственного контакта с ним измерительного прибора. Являются одним из перспективнейших направлений изучения природной среды. Служат

важнейшим источником объективной и оперативной информации в различных явлениях, происходящих в географической оболочке Земли.

Дистанционное зондирование Земли — наблюдение поверхности Земли наземными, авиационными и космическими средствами, оснащенными различными видами съемочной аппаратуры.

Интерполяция — приблизительный расчет недостающего уровня, находящегося внутри ранжированного динамического ряда (или внутри однородного периода колеблющегося ряда).

Интуитивные (экспертные) методы прогнозирования — методы прогнозирования, использующие в качестве информации обработанные суждения экспертов, полученные при проведении специальных опросов.

Картографическая информация — любая информация об объектах и явлениях действительности, которая передается с помощью системы специальных условных знаков в графической, цифровой или других читаемых формах на заданной поверхности (носителях).

Картографический метод — один из традиционных методов географии, задача которого заключается в использовании карт для познания изображенных на них явлений.

Картографическое моделирование — создание, анализ, преобразование картографических произведений с целью приобретения нового познания.

Картометрия — измерение по картам расстояний, площадей, координат, высот, глубин, длин, объемов, направлений и т. д.

Классификация — группировка изучаемых объектов преимущественно по количественным признакам. К классификации относят количественные градации, которые отражают стадии развития явлений или иерархические уровни территориальных систем.

Концепция — система взглядов, то или иное понимание явлений, процессов, объектов, руководящая идея для понимания сущности процессов и явлений.

Корреляционный анализ — математический метод, позволяющий установить направление и тесноту взаимосвязи между определенными явлениями.

Космоснимки — данные, получаемые посредством космических аппаратов в различных диапазонах электромагнитного спектра, визуализируемых затем по определённом алгоритму.

Линейное программирование — область математического программирования, посвященная теории и методам решения экстремальных задач, характеризующихся линейной зависимостью между переменными.

Математико-географическое моделирование — метод формализации географических представлений на основе создания логико-математических конструкций, отражающих количественные отношения реальных географических объектов.

Методология науки — учение о текстуре, логической организации, методах и средствах научной деятельности.

Методы — пути исследования, способы изучения объекта; совокупность действий, призванных помочь достижению желаемых результатов; инструменты достижения цели; совокупность процедур и операций практического или теоретического освоения действительности.

Медиана — значение члена варьирующего ряда, занимающего среднее положение в ряду, в котором величины, образующие его, располагаются в убывающем или возрастающем порядке.

Межотраслевой комплекс — интеграционная структура, характеризующая взаимодействие различных отраслей и их элементов, разных стадий производства и распределения продукта.

Метод баллов — цифровая оценка географических объектов и процессов.

Метод Дельфи — метод коллективного экспертного поискового прогнозирования, основанный на выявлении согласованной оценки экспертной группы путем анонимного опроса экспертов в несколько туров,

предусматривающий сообщение экспертам результатов предыдущего тура с целью дополнительного обоснования оценки экспертов в последующем туре.

Метод полевых исследований — метод географического исследования, сущность которого заключается в получении первичного фактического материала и непосредственном изучении территории для установления пространственной дифференциации населения и хозяйства, сопоставления взаимообусловленности природных и социально-экономических процессов.

Мода — значение признака, имеющее наибольшую частоту в статистическом ряду распределения.

Модель — графическое изображение объекта, отражающее структуру и динамические связи, дающее программу дальнейших исследований, упрощенный, неполный аналог изучаемого объекта.

Нормативная модель — модель, предназначенная для нахождения желательного состояния объекта (например, оптимального).

Нормативный прогноз — прогноз, определяющий пути и способы достижения тех или иных альтернатив, рассматривая сами альтернативы как заданные.

Ограничения модели — элемент экономико-математической модели, математические соотношения, отражающие свойства моделируемых объектов во взаимосвязи с внешними (ограничивающими) факторами. Обычно представляя собой систему уравнений и неравенств, они в совокупности определяют область допустимых решений (допустимое множество).

Объект исследования — то, что существует вне нас и независимо от нашего сознания и является предметом познания, практического действия.

Описательное исследование — вид социологического анализа, предполагающее получение эмпирических сведений, дающих относительно целостное представление об изучаемом явлении, его структурных элементах.

Переменная модели — переменная величина, включенная в модель и принимающая различные значения в процессе решения экономико-математической задачи.

Поисковый прогноз — прогноз, определяющий перспективы, возможные состояния, пути развития объекта прогнозирования в будущем и вероятности их достижения.

Предмет исследования — совокупность представлений (теоретических знаний - учений, теорий, гипотез, концепций, парадигм), сложившихся в науке относительно объекта исследования.

Природно-ресурсный потенциал — совокупность природных ресурсов территории, природных условий, явлений и процессов, которые используются или могут быть реально вовлечены в хозяйственную деятельность при данных технических и социально-экономических возможностях общества с условием сохранения среды обитания человека.

Прогноз — научно обоснованное описание возможных состояний объектов в будущем, а также альтернативных путей и сроков достижения этого состояния. Процесс разработки прогнозов называется прогнозированием.

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

Разведывательное исследование — наиболее простой вид конкретно-социологического анализа. Охватывает, как правило, небольшие обследуемые совокупности и основывается на упрощенной программе и сжатом по объему методическом инструментарии.

Районирование — универсальный метод упорядочения и систематизации территориальных систем, широко используемый в географических науках.

Региональный анализ — экономико-географический анализ факторов регионального развития для познания закономерностей и особенностей

развития конкретных районов; для определения рациональных путей развития конкретного района.

Регрессионный анализ — статистический метод исследования зависимости случайной величины от переменных

Ряд распределения — упорядоченное распределение единиц совокупности по какому-либо варьирующему признаку.

Социологическое исследование — система логически последовательных методологических, методических и организационно-технических процедур, связанных между собой единой целью: получить достоверные данные об изучаемом явлении или процессе для их последующего использования в практике социального управления.

Специализация региона — результат территориального разделения труда. Обусловлена возможностью производить на территории региона (за счет использования благоприятных исторических, экономических, природных и др. условий) определенные виды продукции в количестве, значительно превышающем местные потребности, при сравнительно низких затратах труда, то есть возможностью развивать такие отрасли, продукция которых конкурентоспособна на внешнем рынке и преимущественно ориентирована на вывоз.

Сравнительно-географический метод — заключается в выявлении черт сходства и различия между исследуемыми объектами. Используется для сопоставления социально-экономических систем во времени и пространстве, анализе результатов хозяйственной деятельности, развития населения, сферы обслуживания и др.

Среднеквадратическое отклонение — показывает степень рассеяния значений статистической совокупности около среднего значения.

Средняя величина — обобщающий показатель, выражающий типичный уровень (размер) варьирующего признака в расчете на единицу однородной совокупности.

Статистика — отрасль знаний, наука, в которой излагаются общие вопросы сбора, измерения, мониторинга, анализа массовых статистических (количественных или качественных) данных и их сравнение; изучение количественной стороны массовых общественных явлений в числовой форме.

Статистическая информация — совокупность статистических данных, отображающих социально-экономические процессы и используемых в управлении экономикой и общественной жизнью.

Статистические методы прогнозирования — методы прогнозирования, основанные на выявленных в прошлом закономерностях развития объекта и предположениях об инерционном развитии объекта в будущем.

Стохастическая модель — экономико-математическая модель, в которой параметры, условия функционирования и характеристики состояния моделируемого объекта представлены случайными величинами и связаны стохастическими (т.е. случайными, нерегулярными) зависимостями.

Таксономия — особый вид систематизации, заключающийся в делении территории на сопоставимые, иерархически соподчиненные территориальные единицы (таксоны), которые связаны той или иной общностью свойств, признаков и благодаря этому могут быть отнесены к определенной таксономической категории.

Типология — группировка сложных объектов по совокупностям (типам) в основном по качественным признакам.

Территориально-производственный комплекс — совокупность расположенных рядом друг с другом взаимосвязанных производств.

Территориальная социально-экономическая система (ТСЭС) — экономически - и социально-эффективное взаимосвязанное сочетание элементов общественного воспроизводства, функционирующее на определенной территории как звено географического разделения труда.

Теория — форма обобщенного представления действительности, дающая целостное представление о той или иной области знания и подтверждаемая экспериментом или расчетом.

Тренд — изменение, определяющее общее направление развития, основную тенденцию временного ряда. Для определения тренда используются методы выравнивания ряда.

Формализация — построение абстрактных математических моделей; особый подход в научном познании, который заключается в использовании специальной символики, позволяющей отвлечься от изучения реальных объектов, от содержания описывающих их теоретических положений и оперировать вместо этого некоторым множеством символов (знаков).

Целевая функция модели — математическое представление зависимости критерия оптимальности от искомым переменных.

Экстраполяция — приближенный расчет неизвестных уровней динамического ряда, лежащих за его пределами.

Эксцесс — мера остроты пика распределения случайной величины.

Энерго-производственный цикл — устойчивая совокупность производственных процессов, группирующихся вокруг основного процесса, связанного с освоением и использованием конкретного вида сырья или энергии.