

Нижегородский государственный университет им. Н.И. Лобачевского

Национальный исследовательский университет

Учебно-научный и инновационный комплекс

«Социально-гуманитарная сфера и высокие технологии: теория и практика взаимодействия»

Основная образовательная программа

Основная образовательная программа 030600.62 «История», общий профиль

квалификация (степень) бакалавр

Учебно-методический комплекс по дисциплине

«Математические методы в исторических исследованиях»

Негин А.Е., Миронос А.А.

МАТЕМАТИЧЕСКИЕ МЕТОДЫ В ИСТОРИЧЕСКИХ ИССЛЕДОВАНИЯХ

Электронное учебно-методическое пособие

Мероприятие 1.2. Совершенствование образовательных технологий, укрепление материально-технической базы учебного процесса

Нижегород

2012

МАТЕМАТИЧЕСКИЕ МЕТОДЫ В ИСТОРИЧЕСКИХ ИССЛЕДОВАНИЯХ. ., Негин А.Е., Миронос А.А
Электронное учебно-методическое пособие. – Нижний Новгород: Нижегородский
госуниверситет, 2012. – 31 с.

В учебно-методическом пособии рассматриваются вопросы использования в исторических исследованиях методов математической статистики, а также применения средств математического моделирования, для реконструкции исторических событий и процессов. Использование математических методов в исторических исследованиях иллюстрируется конкретными примерами анализа источниковых комплексов, осуществленных при изучении ключевых проблем российской истории. Пособие содержит сведения о структуре курса, список контрольных вопросов и рекомендуемую для самостоятельного изучения литературу.

Электронное учебно-методическое пособие предназначено для студентов ННГУ, обучающихся по направлению подготовки 030600.62 «История», изучающих курс «Математические методы в исторических исследованиях».

ОГЛАВЛЕНИЕ

	стр.
Введение.	4
<u>Раздел 1. Методы математической статистики в исторических исследованиях</u>	5
1.1. Специфика применения математических методов в истории. «Математизация» исторического знания: возможности и ограничения	5
1.2. Выборочный метод	9
1.3. Метод кластерного анализа	12
1.4. Корреляционный, регрессионный и факторный анализ	16
<u>Раздел 2. Моделирование в исторических исследованиях</u>	22
2.1. Виды математических моделей, применяемых в исторических исследованиях	22
2.2. Математические методы в классической и экспериментальной археологии	25
2.3. Проблемы исторического моделирования. Клиодинамика в реконструкции прошлого и прогнозах будущего	28
2.4. Моделирование средствами фрактальной геометрии	30
Структура и содержание дисциплины	34
«Математические методы в исторических исследованиях»	
Вопросы для подготовки к экзамену	38
Рекомендуемая литература	39

Введение.

Развитие исторической науки, как и других областей научного знания, тесно сопряжено освоением новых технологий, расширяющих познавательные возможности. В современных условиях – основные ресурсы сосредоточены в сфере применения компьютерной техники. Именно в этой области сосредоточены многообещающие возможности для совершенствования методологического инструментария исторической науки. Компьютер создает принципиально новые условия работы историка с источником: он делает возможной обработку громадных массивов данных, многомерного анализа и даже моделирования исторических процессов и событий. Современные программные средства предъявляют и новые требования к самому исследователю: освобождая его, зачастую, от необходимости детального знания технологии работы с данными, их «ручной обработки», они заставляют его гораздо пристальнее относиться к формально-логической составляющей исследовательской деятельности. Применение компьютерных технологий в историческом исследовании влечет за собой математизацию исторического знания, обеспечивает базу для более широкого применения междисциплинарных подходов, благодаря которым стало возможным получать более точные данные о прошлом и проверять уже имеющиеся теоретические наработки предыдущих поколений историков. Значение математических методов многогранно, они, одновременно, выступают как мощное средство в исследовательском арсенале, и как «коммуникативный ресурс», обеспечивающий возможность междисциплинарного синтеза.

Введенный в действие образовательный стандарт третьего поколения по направлению подготовки «История» предъявляет повышенные требования к уровню знаний и компетенций будущих выпускников исторических факультетов в сфере использования информационных технологий и математических методов в исторических исследованиях. Современный бакалавр истории, должен уметь использовать в своей профессиональной деятельности *«базовые знания в области основ информатики, элементы естественнонаучного и математического знания»*. В их освоении ведущее место занимает курс «Математические методы в исторических исследованиях». *Необходимой частью учебного процесса в рамках этого курса являются знакомство с имеющимся опытом применения компьютерных технологий и математических методов в конкретных работах современных историков и приобретение практических навыков применения того или иного метода с учетом опыта классических на сегодняшний день исследований в данной области.* Материал, обобщенный в рамках данного учебно-методического пособия, призван помочь студентам освоить наработанный исторической наукой опыт применения математических методов в решении задач исторической реконструкции.

РАЗДЕЛ 1. МЕТОДЫ МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ В ИСТОРИЧЕСКИХ ИССЛЕДОВАНИЯХ

1.1. Специфика применения математических методов в истории. «Математизация» исторического знания: возможности и ограничения

В социальных и гуманитарных науках, изучающих закономерности существования и развития человеческого общества и отдельного человека, традиционными массивами информации, при работе с которыми обычно используются количественные методы, являются т.н. «статистические источники» - данные учета населения, фискальные и кадастровые данные и т.п. Второй группой, в отношении которой также активно используются количественные методы – это «массовые источники» - массивы однотипных по структуре и составу содержащейся в них информации документы (например, периодические издания). Подобная информация легко поддается формализации и, следовательно, приведению к количественному значению с последующей статистической обработкой.

Но не следует, однако, думать, что статистические методы могут использоваться лишь для анализа статистических источников, представляющих собой в исходном виде цифровой материал. Методы статистики пригодны и для работы с неколичественной информацией, ведь они всегда имеют дело с совокупностями, группами, т.е. *массовым материалом*, а не с отдельными случаями, объектами, индивидуумами. Следовательно, и при описании **совокупности данных**, возможен статистический подсчет и, следовательно, применение статистических методов. Таким образом, математизация исторической информации – гораздо более разноплановое и масштабное явление, имеющее не только явное выражение в виде привлечения и обработки данных, содержащих собственно количественную информацию в узком смысле.

Внедрять обработку статистических данных с помощью методов математики в исторических исследованиях и в сопутствующих им вспомогательных исторических дисциплинах стали еще в XIX в. Именно тогда все разрастающаяся источниковая база как письменных, так и археологических источников, потребовала обработки, систематизации и верификации с помощью элементов математического знания.

Своеобразным направлением, позволяющим в конечном счете привести историческую информацию к некому количественному воплощению и, таким образом, обрабатывать ее математическими средствами, является использование экспериментальных методик в истории и археологии. В середине XIX века благодаря усилиям Наполеона III произошло рождение и оформление так называемой военной археологии и реконструкции. Им целенаправленно финансировались раскопки в Алезии, при его поддержке состоялась первая попытка реконструировать античное гребное судно - трирему и средневековую метательную машину – требюше. В этих экспериментах реконструкции древней техники впервые отмечено массовое применение математических методов при изучении развития

античных технологий. На протяжении второй половины XIX – начала XX века последовала целая серия экспериментов на основе математических вычислений, которые ставили своей целью восстановление и испытание действующих моделей греческой и римской осадной техники и метательных машин. Так, спортсмен и филантроп Р. Пейн-Гэллвей реконструировал римскую одноплечевую машину – онагр, довольно туманно описанный Аммианом Марцеллином. Этот большой онагр сумел запустить каменное ядро весом 3,6 кг на расстояние в 450 метров! В начале XX века инициатива перешла к немецким исследователям. Майор Э. Шрамм в сотрудничестве с классическими учеными и при поддержке кайзера Вильгельма II построил двенадцать образцов античных метательных машин. После грандиозной работы проделанной Э. Шраммом в течении следующих шестидесяти лет новых попыток реконструкции не предпринималось, вплоть до появления впоследствии новых археологических находок, уточнивших многие детали.

Касаясь проблем использования статистических методов в исследованиях по античной истории, следует упомянуть, например, расчеты Я. Ле Бозка, приведенные им в его книгах «Третий Августов легион» и «Римская армии эпохи Ранней Империи»¹. Он, например, сравнивал африканские и испанские легионы, у которых соотношение италийцев и местных уроженцев было совершенно различным. Несмотря на это, латинских *cognomina* было преобладающее количество: 96 против 4 для Африки и 94 против 6 для Испании. Он отмечает, что вообще греческие имена у легионеров встречаются крайне редко и их носителей можно подразделить на 3 категории: те, кто действительно происходил с Востока, солдаты из «лагеря» (нет единого мнения по-поводу происхождения термина *origo castris*) и те, кто жил в правление Адриана (как известно - эллинофила).

В Африке, где большую часть времени был размещен только один легион, III Августов, можно проследить изменения этнического состава по документам, особенно многочисленным для II в. и эпохи Северов. В итоге своих подсчетов Я. Ле Бозк пришел к выводу, что I век — это век италийцев и галлов. В начале II в. н.э. в легион начинают вступать африканцы (а некоторые из них сделали это уже в I в.), но их пока еще меньше, чем вифинцев, выходцев с Нижнего Подунавья и особенно сирийцев после парфянских походов того же Траяна. В конце II в. процентное соотношение меняется в обратную сторону — преобладают африканцы, в первую очередь уроженцы Магриба, а затем Нумидии. В начале III в. доля «иностранцев» оставалась стабильной. Легион, распущенный между 238 и 253 гг., был восстановлен, возможно, за счет набора местных жителей; но в середине III в. было уже утрачено обыкновение указывать происхождение новобранца.

Успешное внедрение статистики в изучаемые документы по средневековой и новой истории осуществили историки, работавшие в рамках так называемой школы «Анналов», возникшей на основе одноименного журнала в 1929 г. Представители школы «Анналов» стремились к всестороннему рассмотрению исторического материала, в рамках создания так называемой «тотальной истории» (*histoire totale*). Первая попытка такого воплощения этого идеала всеохватывающей истории приписывается Ф. Броделю, лидеру профессиональных французских историков середины XX в. В его работе

¹ Le Bohec Y. La Troisième Légion Auguste. Paris, 1989; Ле Бозк Я. Римская армия эпохи Ранней Империи/ Пер. с фр. М. Н. Челинцевой. - М., 2001.

«Средиземноморье и средиземноморский мир в эпоху Филиппа II» (1947) были освещены все аспекты этой огромной темы ярко и подробно: физическая география и демография, экономическая и социальная жизнь, политические структуры и политика Филиппа II и его соперников в Средиземноморье. По мнению Броделя, в изучении истории следовало бы как можно шире применять математическое моделирование и разработать подлинную «социальную математику».

Историки школы «Анналов» первыми обратились к локальной истории нового типа. Сила такого подхода «локальной тотальной истории» продемонстрировал другой уже упоминавшийся французский историк Э. Леруа Ладюри в своих работах «Крестьяне Лангедока» (1966) и «Монтайю» (1978). Эти исследования были ограничены масштабами одной деревни на протяжении жизни нескольких поколений.

Близкие к школе «Анналов» методологические разработки применял в своих исследованиях известный российский историк-медиевист Ю. Л. Бессмертный (1923-2000). Так, в своей книге «Жизнь и смерть в Средние века» на материале истории Франции IX–XVIII вв. Ю. Л. Бессмертный проанализировал формы брака и семьи, проследил изменение взглядов на роль женщины в жизни средневекового общества, рассказал об отношении к детству и старости, о «самосохранительном» поведении в разных социальных слоях, воспроизвел средневековые представления о болезни и смерти. Автор исследует изменение важнейших демографических параметров - брачности, рождаемости, смертности, естественного прироста населения.

Уже в конце 50-х гг. возникает и развивается клиометрика (*клиометрия* - англ. Cliometrics) – направление в исторической науке, предполагающее систематическое использование математических методов. Близким, фактически синонимичным понятием, является «квантитативная история» понимаемая как историческое знание, полученное с применением математических методов в исторических исследованиях. Название данного направления производится от имени Клио — музы истории и героической поэзии в греческой мифологии. Клиометрика — это междисциплинарное направление, первоначально связанное с применением эконометрических методов и моделей в исследованиях по экономической истории. Впервые термин клиометрика появился в печати в декабре 1960 года в статье Дж. Хьюгса, Л.Дэвиса и С.Рейтера «Аспекты квантитативного исследования в экономической истории».

Однако бурный всплеск интереса к подобным исследованиям, часто обозначаемый как «клиометрическая революция», связан с 1960-ми гг. Особую роль в развитии данного направления (клиометрические подходы к изучению экономической истории) сыграл американский журнал «Journal of Economic History», редакторами которого в 1960-е гг. стали Дуглас Норт и Уильям Паркер – сторонники клиометрического подхода. В этот же период времени в США стали регулярно проводиться клиометрические конференции. Американские исследователи, опираясь на методы клиометрики, с успехом изучали роль железнодорожного строительства в развитии процессов индустриализации и развития сельского хозяйства США в XIX веке, экономической эффективности рабского труда в американской экономике и т.п.

В 1993 году Роберт Фогель и Дуглас Норт получили Нобелевскую премию по экономике за цикл работ в области клиометрики. В решении Нобелевского комитета отмечается, что

премия присуждена «за развитие новых подходов в исследованиях по экономической истории, основанных на применении экономической теории и количественных методов для объяснения экономических и институциональных изменений».

С 1970-х гг. клиометрический подход начинает активно применяться в исследованиях по экономической истории в Великобритании, скандинавских странах, Испании, Бельгии, Голландии и др. странах. В более широком плане применение количественных методов в исторических исследованиях (квантитативная история) получило распространение в Германии (основную роль здесь играет Центр историко-социальных исследований Кельнского университета) и СССР (России), где «клиометрическая школа» начала складываться в 1970-е гг. прошлого века. Становление квантитативной истории сопровождалась большим количеством научных конференций, публикаций, появлением периодических изданий, таких, например, как "Historical Methods" (с 1967 г., с 1978 г. - "Historical Methods Newsletter") в США, "Computer and the Humanities" (с 1966 г.), "Historische Sozialforschung" (с 1976 г. - "Historical Social Research") в Европе. Данное направление ставило целью качественный переход к пониманию истории как развитой науки (science), систематически применяющей не только методы и модели, но и теории смежных наук. Сильное влияние квантитативных идей испытывали представители "школы Анналов". Известно полемически заостренное высказывание Э. Ле Руа Ладюри: "История, которая не является квантифицируемой, не может претендовать на то, чтобы считаться научной".

В СССР центром исследований по квантитативной истории стал МГУ им. М.В. Ломоносова, где, в 1970-х - 1980-х годах сформировалось сообщество ученых, применяющих математические методы и ЭВМ в исторических исследованиях. Безусловным лидером нового направления стал академик И.Д.Ковальченко. С 1979 г. на базе исторического факультета МГУ действовал и всесоюзный семинар «Количественные методы в исторических исследованиях» (Л. В. Милов, Л. И. Бородкин и др.).

За почти полувековой период активного развития «квантитативной методологии» истории можно говорить о существенной внутренней эволюции как самого научного направления (начинавшегося с клиометрических подходов к изучению экономической истории), так и появления на его базе смежных областей - в частности, активно развивающейся в последние два десятилетия исторической информатики, превратившейся в междисциплинарную область, разрабатывающую теоретические и прикладные проблемы использования информационных технологий в исторических исследованиях и образовании. Однако, все эти междисциплинарные области связаны общностью базового подхода – математизацией исторического знания. Так, Л.И. Бородкин, рассматривая историю возникновения и развития исторической информатики, выделяет на два существенно отличных по своему содержательному наполнению периода: первый - эпоха «больших» ЭВМ (начало 1960-х - конец 1980-х гг.) и второй - «микрокомпьютерная революция» (конец 1980-х - середина 1990-х гг.). К настоящему времени можно говорить о трех последовательных этапах математизации исторической науки: 1) математико-статистическая обработка эмпирических данных и количественная формулировка качественно установленных фактов и обобщений, включающая традиционные математико-статистические методы (дескриптивная статистика, выборочный метод, анализ временных рядов, корреляционный анализ); методы многомерного

статистического анализа; 2) разработка математических моделей явлений и процессов в какой-то области науки; 3) использование математического аппарата для построения и анализа общей научной теории. По мнению Л.И. Бородкина, третий этап в истории пока вообще еще не используется, второй находится в стадии активной разработки.

Уже в конце XX в., как своеобразная реакция на попытки утверждения «сциентизма» в исторических исследованиях, появились и «неоантипозитивистские» концепции, отрицающие возможность научного познания не только прошлого, но и современности. С этой точки зрения отрицается эффективность применения в истории математических методов и предлагается вернуться на позиции художественного, поэтически-метафорического методов ее осмысления и описания, при котором историк кажется все еще больше рассказчиком, нежели исследователем. Очевидными ограничениями, на которые указывают «скептики» в отношении применения количественных методов в исторических исследованиях, связаны с отсутствием прямого наблюдения, субъект-объектной корреляцией, многофакторностью проявлений и соответствующей многоаспектностью изучения, а также со слабой однородностью используемой информации.

Вместе с тем, безусловно, новые методы исторических исследований, основанные на использовании средств математической обработки данных, позволили пересмотреть на ином уровне обобщения ряд уже известных проблем, а также поставить и решить принципиально новые, крупные задачи изучения исторического прошлого.

1.2. Выборочный метод

Зачастую историки имеют в своем распоряжении большой массив источников и данных, которые они не в состоянии полностью обработать. Это касается, в первую очередь, исследований по Новой и Новейшей истории. С другой стороны, чем глубже приходится заглядывать вглубь веков, тем меньшим количеством информации можно оперировать. В обоих этих случаях небесполезно использовать так называемый выборочный метод, суть которого заключается в замене сплошного обследования массовых однородных объектов частичным их исследованием. При этом из генеральной совокупности выделяется часть элементов, именуемая выборкой, и результаты обработки выборочных данных в итоге обобщаются на всю совокупность. Основой для характеристики всей совокупности может служить только репрезентативная выборка, правильно отражающая свойства генеральной совокупности. Это достигается методом случайного отбора элементов генеральной совокупности, при котором у всех ее элементов имеются равные шансы попадания в выборку.

Применение данного метода одинаково подходит и для изучения различных явлений и процессов современности, и для обработки данных проведенных ранее выборочных статистических исследований, таких как переписи. Кроме того, выборочный метод также находит применение при обработке данных естественных выборок, от которых остались лишь фрагментарно сохранившиеся данные. Так, довольно часто, к таковым частично

сохранившимся данным относятся актовые материалы, документы текущего делопроизводства и отчетности.

В зависимости от того, каким образом осуществляется отбор элементов совокупности в выборку, различают несколько видов выборочного обследования, в которых отбор может быть случайным, механическим, типическим и серийным.

Случайным называется отбор, при котором все элементы генеральной совокупности имеют равную возможность быть отобранными, например, с помощью жребия или таблицы случайных чисел.

Способ жеребьевки применяется в том случае, если число элементов всей изучаемой совокупности невелико. При большом объеме данных осуществление случайного отбора при помощи жеребьевки становится сложным. Более пригоден, в случае большого объема обрабатываемых данных, метод использования таблицы случайных чисел.

Способ отбора с помощью таблицы случайных чисел можно рассмотреть на следующем примере. Допустим, что совокупность состоит из 900 элементов, а намеченный объем выборки равен 20 единицам. В таком случае из таблицы случайных чисел следует отбирать числа, не превосходящие 900, до тех пор, пока не будут набраны требуемые 20 чисел. Выписанные числа следует считать порядковыми номерами попавших в выборку элементов генеральной совокупности.

Для очень больших совокупностей лучше применить механический отбор. Так, при формировании 10%-ной выборки из каждых десяти элементов выбирается только один, а вся совокупность условно разбивается на равные части по 10 элементов. Далее из первой десятки наугад выбирается какой-либо элемент (например, при помощи жеребьевки). Остальные элементы выборки определяются указанной пропорцией отбора N номером первого отобранного элемента.

Еще одним видом направленного отбора является типический отбор, когда совокупность разбивается на группы, однородные в качественном отношении. Только после этого уже внутри каждой группы производится случайный отбор. Хотя это более сложный метод, он дает более точные результаты.

Серийный отбор представляет собой вид случайного или механического отбора, осуществляемый для укрупненных элементов исходной совокупности, которая в ходе анализа разбивается на группы (серии).

Изложенные выше способы формирования выборок не исчерпывают собой всех типов отбора, применяемых на практике².

В качестве примера применения выборочного метода в историографии рассмотрим подробнее проведенный отечественными исследователями анализ движения хлебных цен в России в XVIII веке³. Была поставлена задача определить средние цены на хлеб по отдельным губерниям, районам и по России в целом за каждый год XVIII века, а также выявить динамику хлебных цен за столетие. Однако, в ходе исследования стало понятно, что составить таблицы с непрерывным рядом цен не удастся, так как данные в различных архивах сохранились лишь частично. Например, данные за 1708 год имелись только по 36 уездам страны. Только за периоды с 1744 по 1773 и с 1796 по 1801 годы сохранились данные по большинству городов России. В связи с этим было принято решение

² Для наиболее полного ознакомления с различными видами отбора советуем обратиться к книге: *Пейте Ф.* Выборочный метод в переписях и обследованиях. М., 1965.

³ Миронов Б.Н. Хлебные цены в России за два столетия (XVIII–XIX вв.). Л., 1985.

использовать в исследовании аппарат математической статистики, а именно, понятие среднего значения, дисперсии, среднего квадратического отклонения, доверительного интервала. В связи с разрозненностью имеющихся сведений из генеральной совокупности (все сохранившихся и несохранившихся данных о хлебных ценах за 1708 год) была произведена выборка, отражающая совокупность сохранившихся сведений о хлебных ценах. На основании данной выборки была рассчитана средняя цена на хлеб, а также показатель отклонения от среднего значения и был построен доверительный интервал для среднего значения генеральной совокупности с вероятностью 0,95 по следующим данным:

Уезд	Цена (в коп.)	Уезд	Цена (в коп.)	Уезд	Цена (в коп.)
1.	40	2.	43	3.	40
4.	80	5.	74	6.	40
7.	55	8.	42	9.	42
10.	50	11.	40	12.	43
13.	43	14.	35	15.	40
16.	30	17.	36	18.	50
19.	30	20.	29	21.	45
22.	40	23.	42	24.	40
25.	36	26.	50	27.	30
28.	24	29.	25	30.	40
31.	32	32.	30	33.	20
34.	30	35.	25	36.	32

Среднее значение признака, которым является цена на хлеб в 1708 году, был найден по

формуле $\bar{x} = \frac{1}{n} \sum x_i$, где n - объем выборки.

Из имеющихся данных следует, что $\bar{x} = \frac{1}{36} \cdot 1336 = 37,1$ коп

Среднее квадратическое (стандартное) отклонение было найдено по формуле

$$s_x = \sqrt{\left(\frac{1}{n} \sum x_i^2\right) - (\bar{x})^2} = \sqrt{\frac{1}{36} \cdot 58948 - 37,1^2} = 16,16 \text{ коп}$$

Таким образом, согласно данным выборки средняя цена на хлеб в России в 1708 году составляла 37 копеек со стандартным отклонением 16 копеек. Рассчитав коэффициент

вариации, стало ясно, что выборка является неоднородной, вследствие чего возникла необходимость проведения дополнительного анализа цен на хлеб по районам. Нужно было выяснить, насколько средние цены на хлеб, вычисленные по данным выборки, могли отличаться от действительных средних хлебных цен, которые были бы получены, если бы в распоряжении оказались данные за этот год по всем уездам России. Соответственно была определена средняя и предельная ошибки выборки, и построен доверительный интервал. Средняя ошибка для повторной выборки была вычислена по формуле

$\mu_{\bar{x}} = \frac{s_x}{\sqrt{n}}$, где $s_x = \sqrt{\frac{n}{n-1}} \cdot s_x^2$ (исправленное среднее квадратическое отклонение). При подставлении данных, получаем $\mu_{\bar{x}} = \frac{16,38}{\sqrt{36}} = 2,73$ коп.

Предельная ошибка вычисляется по формуле $\Delta = t \cdot \mu$, где значение t зависит от значения вероятности, с которым строится доверительный интервал. При $p = 0,95$ имеем $t = 1,96$. Таким образом, получаем $\Delta = 1,96 \cdot 2,37 = 5,35$ коп.

Далее были проведены вычисления по формуле доверительного интервала: $x_{\theta} - \Delta < x_{ген} < x_{\theta} + \Delta$. Было выяснено с вероятностью 95%, что средняя цена на хлеб в 1708 году по России могла изменяться в пределах от 31,75 копеек до 42,45 копеек.

Таким образом, при помощи аппарата математической статистики, исследователям удалось вычислить средние цены по 10 районам России, а также среднероссийские цены за каждый год XVIII века.

1.3. Метод кластерного анализа

Для типизации в исторических исследованиях наиболее эффективны методы многомерной типологии. Наиболее широко распространен вид типизации по географическому районированию, благодаря которому можно выделить сплошной территориальный комплекс, что важно для раскрытия тех или иных особенностей исторического развития. С другой стороны, территориальное единство изучаемых объектов само по себе не обеспечивает их содержательной однородности. Поэтому, в дополнение к географическому районированию, историки используют социальную типизацию изучаемых объектов, в основе которой лежит не географическое, а социальное пространство. Такой подход уже носит характер многомерной типологии. Наиболее известным методом многомерной типологии является кластерный анализ. Он позволяет выделить кластеры (от англ. Cluster — скопление), группы объектов со сходными свойствами, расположенные в пространстве. Близость этих объектов друг к другу отражает степень их сходства.

Рассмотрим процесс выделения кластеров на примере агломеративно-иерархического метода. Итак, пусть все m признаков будут измерены в количественной шкале. В таком случае каждый n объект будет представлен точкой в m -мерном пространстве признаков. О сходстве объектов можно судить по расстоянию между соответствующими точками. Соответственно, чем ближе объекты находятся друг к другу, тем они более схожи.

Для определения близости пары точек (объектов i и j) в многомерном пространстве используется *евклидово расстояние*, равное корню квадратному из суммы квадратов разностей значений одноименных показателей, взятых для данной пары объектов:

$$d_{ij} = \sqrt{\sum_{k=1}^m (x_{ik} - x_{jk})^2} \quad (i, j = 1, 2, \dots, n),$$

где d_{ij} – евклидово расстояние между i -м и j -м объектами, x_{ik} – значение k -го признака для i -го объекта.

Расстояние между объектами зависит от «масштаба» признаков, который обычно нормализуют, т.е. все признаки приводят к стандартному виду со средним значением, равным нулю, и стандартным отклонением, равным единице.

После нормализации объекты сохраняют свое относительное положение, но «масштаб» измерения признаков уже будет единым.

Обычно близость двух кластеров определяется как среднее значение расстояния между всеми парами объектов, где один объект пары принадлежит к одному кластеру, а другой – к другому:

$$D_{pq}^2 = \sum_{i \in X_p} \sum_{j \in X_q} \frac{d_{ij}^2}{n_p n_q},$$

где D_{pq}^2 – мера близости между p -м и q -м кластерами; X_p – p -й кластер; X_q – q -й кластер; n_p, n_q – число объектов в p -м и q -м кластерах, соответственно.

На первом шаге процедуры агломеративно-иерархического метода кластерного анализа по начальной матрице расстояний между объектами определяется минимальное расстояние. Затем выделяют наиболее близкие объекты, находящиеся друг от друга на этом расстоянии, и объединяют в один кластер. В матрице вычеркивают строку и столбец, соответствующие первому из этих объектов, а расстояния от нового кластера до всех остальных кластеров вычисляют по вышеприведенной формуле. Эти значения вписывают в строку и столбец матрицы расстояний, соответствующие второму объекту из первого кластера.

Второй шаг процедуры предусматривает формирование нового кластера, на основе нового определения минимального расстояния. Этот кластер строят объединением двух объектов, или одного объекта с кластером, построенным на первом шаге. В матрице расстояний вычеркиваются одна строка и один столбец, а одна строка и один столбец пересчитываются и т.д. В конце этой процедуры получится один кластер, объединяющий все n объектов.

С помощью методов кластерного анализа была проведена аграрная типология губерний Европейской России на рубеже XIX—XX вв.⁴ Анализ проводился следующим образом. Для начала были отобраны 19 показателей, характеризующих земельные отношения (размеры крестьянских наделов, удельный вес дворянского землевладения, продажа частновладельческих земель, цена на землю, размеры крестьянской аренды и арендная

⁴ Ковальченко И. Д., Бородкин Л. И. Аграрная типология губерний Европейской России на рубеже XIX—XX веков: (Опыт многомерного количественного анализа) // История СССР. 1979. № 1. С. 59—95.

плата), состояние сельскохозяйственного производства (посевы, сборы и урожайность хлебов, количество рабочего и продуктивного скота, цены на сельскохозяйственную продукцию), глубину и особенности буржуазной аграрной эволюции (применение наемного труда, зарплата сельскохозяйственных рабочих, разложение крестьян). В результате математической обработки данных было выделено 15 взаимосвязанных между собой кластеров с указанием на графике «расстояния», показывающего «близость» губерний, входящих в тот или иной кластер, а, кроме того и самих кластеров. Благодаря такой визуальной подсказке, например, выяснилось, что наиболее сходными по совокупности 19 признаков были губернии VII (Воронежская и Саратовская) и XI (Киевская и Подольская) кластеров. Наименее сходными между собой и в то же время самыми непохожими на все другие были губернии XV кластера (Московская и Петербургская). При этом, однако, кластеры не образовали существенно отличных типов губерний, так как различия между многими из этих кластеров были невелики. Чтобы выделить типы необходимо объединить полученные мини-кластеры в макро-кластеры, после чего уже можно выделить определенные типы. В рассматриваемом примере на основе «расстояний» были выделены следующие типы губерний: I—V кластеры образовали нечерноземный тип аграрного развития, VI—XI кластеры составили среднечерноземный тип, XIII и XIV кластеры обозначили южностепной тип, XV кластер —прибалтийский тип, а XII мини-кластер представлен губерниями столичного типа.

Пример таблицы по кластерам с указанием расстояния, показывающей структуру промышленной типологии губерний Европейской России в начале XX в.:

Кластеры	«Расстояние»							
I	0,13	0,16						
II	0,15		0,21					
III	0,18			0,23				
IV	0,19				0,30			
V	0,27					0,32		
VI	0,28						0,55	
VII	0,50							0,57
VIII	0,46							

В состав указанных в таблице кластеров входят следующие губернии:

I	II	III	IV
Уфимская	Волынская	Киевская	Орловская
Пермская	Минская	Харьковская	Черниговская
Тульская	Витебская	Рязанская	Тамбовская
Астраханская	Ковенская	Новгородская	Смоленская
Виленская	Гродненская	Калужская	Симбирская
	Могилевская	Вологодская	Казанская
			Вятская
V	VI	VII	VIII
Псковская	Донская	Саратовская	Московская
Олонецкая	Херсонская	Оренбургская	Петербургская
Пензенская	Таврическая	Полтавская	Лифляндская
Нижегородская	Курляндская	Бессарабская	Екатерино-
	Самарская		славская
	Курская		Костромская
	Подольская		Тверская
			Ярославская
			Владимирская
			Эстляндская

Наиболее характерные различия между типами устанавливаются путем сопоставления средних значений рассматриваемых признаков в каждом из типов.

Кластерный анализ – это весьма эффективный метод многомерной типологии, хотя и не лишенный недостатков. К таковым относится его ограниченность по части выделения

типов. Кроме того, хотя кластерный анализ и способен показать некое «расстояние» между объектами в мини-кластере и между кластерами, однако эти «расстояния» не способны измерять непосредственно меру сходства и различий между объектами.

Тем не менее, этот метод находит применение и в археологии, так как можно изучать кластерную структуру множества памятников по наличию и частоте встречаемости артефактов. В качестве примера применения метода кластерного анализа в археологии можно привести типологию поселений Алтая VI–II вв. до н.э.⁵ Исследователями был проведен анализ карты расположения известных археологических памятников, на основании которого был сделан вывод о том, что система расположения древних поселков находилась в прямой зависимости от природно-географических условий данной местности, а именно: стационарные поселки древние жители Алтая предпочитали возводить на более высоких террасах и мысах, чем стоянки, а поселения располагались кустами по 8-16 пунктов на крупных реках чаще, чем на их притоках. Для получения скрытой и неярко выраженной информации исследователи выделили 12 видов орнаментов, присутствовавших на фрагментах керамики, обнаруженной на 39 исследованных поселениях. После чего был осуществлен подсчет каждого вида в процентах по каждому поселению. Полученная матрица данных была исследована методом кластерного анализа. В итоге была получена дендрограмма (иерархическая структура), в которой группы объектов могут рассматриваться либо как культурные, либо как территориальные, либо как хронологические.

Однако, кластеры, найденные исследователем, после повторного сбора информации и применения кластерного анализа могут «рассыпаться» из-за случайности выявленной кластерной структуры. Это происходит в том случае, если реальная кластерная структура отсутствует вообще, т. е. исследуемая совокупность является однородной, или когда задано не соответствующее реальности число классов.

Чтобы проверить достоверность наличия кластерной структуры, необходимо привлечение дополнительных фактов и исследование классификации с использованием переменных, как участвующих, так и не участвующих в кластеризации.

1.4. Корреляционный, регрессионный и факторный анализ

Первооткрывателем корреляционного метода является французский естествоиспытатель и натуралист Жорж Кювье (George Cuvier, 1769–1832). Закон Корреляции был выведен им средствами сравнительной анатомии. Кювье понял, что органы одного организма соответствуют друг другу и его общим условиям существования. Так, травоядные имеют зубы, приспособленные для пережевывания растений, а на ногах у них копыта для быстрого бега от хищников. У хищников же выдающиеся клыки, а на ногах когти, и т. д. По его утверждению, ему было под силу восстановить всё животное по одной его части.

Таким образом, корреляция – показатель, отражающий взаимную зависимость двух или более величин. При этом величины должны выбираются случайно, а зависимость может определяться либо совпадением, либо отношениями причинности. Необходимо выяснить, не является ли корреляция ложной, то есть основанной на совпадении. Для этого вводится

⁵ Абдулганеев М.Г., Владимиров В.Н. Типология поселений Алтая VI–II вв. до н.э. Барнаул, 1997. 148 с.

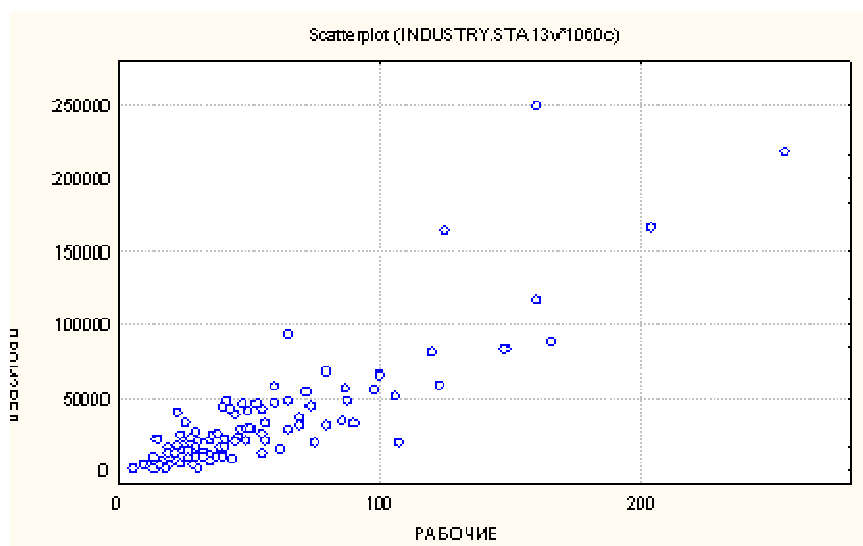
еще одна новая случайная величина. Только при изменении значения одной величины, которое влечет за собой неминуемое систематическое изменение значения другой величины, корреляция считается установленной. Такое изменение может быть выражено в виде коэффициента корреляции, или корреляционного отношения. Коэффициент корреляции показывает, насколько тесно две переменных связаны между собой.

Статистической характеристикой при проверке значимости корреляции служит отношение самого коэффициента к его утроенной ошибке, вычисляемое по формуле:

$$t = r\sqrt{n-2} \frac{1}{\sqrt{1-r^2}},$$

где n – объем выборки. В этой величине известны вероятности всех ее значений. Чем больше значение t , тем меньше его вероятность, т.е. вероятность того, что данная или большая величина корреляции может быть получена в выборке из генеральной совокупности, в которой корреляция равна нулю. В том случае, если эта вероятность окажется меньше выбранного уровня значимости, гипотеза о некоррелированности признаков отклоняется, а связь признается значимой.

Для визуального выявления наличия взаимосвязи между количественными переменными полезно строить диаграммы рассеяния (scatterplot). В этом графике по горизонтальной оси (X) откладывается одна переменная, по вертикальной (Y) другая. При этом каждому объекту на диаграмме соответствует точка, координаты которой равняются значениям пары выбранных для анализа переменных.



Различают два вида зависимостей, которые присущи объективным явлениям природы и общества.

Функциональная зависимость – это взаимосвязь между признаками, в которой каждому значению одного признака соответствует единственное значение другого признака.

Простейшей ее формой является линейная зависимость, характеризующаяся уравнением:

$$y = ax + b .$$

К другими формами функциональной зависимости, относятся: парабола ($y = ax^2 + bx + c$), гипербола ($ax by k+=$), логарифмическая функция ($y = a \lg x$), экспонента ($y = keax$, $k > 0$, $a > 0$).

Функциональная зависимость предполагает изолированность взаимосвязанных признаков от воздействия других факторов. Но такая ситуация в явлениях общественной жизни практически не встречается. В случае, если на связь между признаками влияет множество других факторов, и она проявляется лишь в тенденции, «в среднем», то такая зависимость носит название статистической, или корреляционной.

Для того, чтобы определить тесноту связи между двумя признаками, следует высчитать так называемый *парный линейный коэффициент корреляции*, рассчитывающийся по формуле:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}},$$

где x_i , y_i - значения признаков x и y для i -го объекта; n - число объектов; \bar{x} , \bar{y} - средние арифметические значения признаков x и y . Линейный коэффициент корреляции может принимать значения от -1 до +1, причем чем ближе величина коэффициента корреляции к предельным значениям, тем теснее будет взаимосвязь между признаками. В том случае, если коэффициент равен нулю, линейная связи между признаками будет отсутствовать. Прямая функциональная зависимость будет иметь место, если коэффициент корреляции равен +1 (или -1).

Однако, зачастую необходимо не только оценить тесноту связи между изучаемыми признаками, но и определить ту степень с которой один признак воздействует на другой. В этом случае используется коэффициент детерминации, определяющий процентную долю изменений, происходящих под влиянием факторного признака, в общей изменчивости резульативного признака:

$$D = r^2 100\%,$$

где r - коэффициент корреляции.

В качестве примера подобных вычислений приведем данные из книги Б.Н. Миронова «История в цифрах», где была определена степень корреляционной зависимости между доходом и размерами помещичьего хозяйства в России на рубеже XIX-XX вв. по сведениям о размерах (в десятинах) и доходах (в тыс. руб.) десяти помещичьих имений⁶. Доходность имения зависела от его размера, но, кроме этого на нее влияло и качество земли, и состояние хозяйства, и деловые способности владельца, а также близость рынка и другие факторы. В связи с этим, исследователь поставил задачу узнать, насколько же размер имения влиял на доходность имения.

⁶ Миронов Б.Н. История в цифрах. Л., 1991. С.67.

Исходные данные (x_i - размеры имения в десятинах, y_i - доход имения в тыс. руб.) и промежуточные вычисления были представлены в виде следующей таблицы:

	x_i	y_i	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x})^2$	$(y_i - \bar{y})^2$	$(y_i - \bar{y})(x_i - \bar{x})$
1	240	1.50	-50	-0.10	2500	0.01	5.00
2	255	1.25	-35	-0.35	1225	0.1225	12.25
3	265	1.55	-25	-0.05	625	0.0025	1.25
4	270	1.40	-20	-0.20	400	0.04	4.00
5	285	1.45	-5	-0.15	25	0.0225	0.25
6	295	1.60	5	0	25	0	0
7	310	1.80	20	0.20	400	0.04	4.00
8	320	1.80	30	0.20	900	0.04	6.00
9	325	1.85	35	0.25	1225	0.0625	8.75
10	330	1.90	40	0.30	1600	0.09	12.00

Из которой было выведено, что:

$$\bar{x} = 290; \bar{y} = 1,60; r = \frac{54,0}{\sqrt{8925 \cdot 0,43}} = \frac{54,0}{61,95} \approx 0,87; D = r^2 100\% = 76\% .$$

Следовательно доходность имения на 76% процентов может быть объяснена его размером, и только на 24% другими факторами.

При анализе статистической зависимости важна не только оценка тесноты связи между признаками, но и выявление ее формы. Данная задача решается методами регрессионного анализа.

Регрессионный анализ представляет собой совокупность методов математической статистики, которые позволяют определить форму связи между результативным и факторным признаками, установленной корреляционным анализом. Корреляционная связь описывается уравнением регрессии с помощью с помощью подходящей функции.

Простейшее уравнение линейной регрессии:

$$y = ax + b,$$

где x - факторный признак; y - результативный признак; a и b – параметры уравнения, которые могут быть найдены методом наименьших квадратов по формулам:

$$a = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}, \quad b = \bar{y} - a\bar{x},$$

где x_i, y_i - i -е значение признаков x и y соответственно; \bar{x}, \bar{y} - средние арифметические признаков x и y ; n - число значений признаков x и y .

Регрессионный анализ не используется для определения наличия связи между переменными, ввиду того, что наличие такой связи и есть предпосылка для применения анализа.

Линейная регрессия достаточно хорошо работающим в ряде простых задач. К ее достоинствам относится простота алгоритма и высокое быстродействие. Недостаток только один – неприспособленность к решению существенно нелинейных задач.

Корреляционный анализ выявляет структуру взаимосвязей признаков, характеризующих изучаемое явление или процесс, но не способен объяснить, чем обусловлена именно такая структура связей. Ответить на этот вопрос позволяют методы факторного анализа.

Факторный анализ объединяет методы анализа структуры множества признаков, характеризующих изучаемые явления и процессы, и выявления обобщенных факторов. В его основе лежит положение о том, что корреляционные связи между большим числом наблюдаемых показателей определяются существованием меньшего числа гипотетически наблюдаемых показателей или факторов.

Объяснение множества исходных признаков через небольшое число общих факторов осуществляется сжатием информации, которая содержится в исходных коррелированных признаках.

Основными характеристиками факторного анализа являются факторные нагрузки и факторные веса.

Факторная нагрузка - это значение коэффициентов корреляции каждого из исходных признаков с каждым из выявленных факторов. Чем теснее связь данного признака с рассматриваемым фактором, тем выше значение соответствующей факторной нагрузки. Положительный знак факторной нагрузки указывает на прямую связь данного признака с фактором, а отрицательный знак – на обратную. Если значение факторной нагрузки близко нулю, то это свидетельствует о том, что этот фактор практически не влияет на данный признак.

Факторные нагрузки

№ признаков	№ факторов					
	1	2	...	<i>j</i>	...	<i>k</i>
1	a_{11}	a_{12}	...	a_{1j}	...	a_{1k}
2	a_{21}	a_{22}	...	a_{2j}	...	a_{2k}
·					
<i>i</i>	a_{i1}	a_{i2}	...	a_{ij}	...	a_{ik}
·					
<i>m</i>	a_{m1}	a_{m2}	...	a_{mj}	...	a_{mk}
Вклады факторов	V_1^2	V_2^2	...	V_j^2	...	V_k^2

В приведенной таблице факторных нагрузок содержится *m* строк (по числу признаков) и *k* столбцов (по числу факторов).

Данные о факторных нагрузках позволяют судить о выборе исходных признаков, отражающих тот или иной фактор, и об относительной доле отдельных признаков в структуре каждого фактора.

Факторные веса

№ объектов	№ факторов					
	1	2	...	j	...	k
1	b_{11}	b_{12}	...	b_{1j}	...	b_{1k}
2	b_{21}	b_{22}	...	b_{2j}	...	b_{2k}
·					
i	b_{i1}	b_{i2}	...	b_{ij}	...	b_{ik}
·					
n	b_{n1}	b_{n2}	...	b_{nj}	...	b_{nk}

В таблице факторных весов содержится n строк (что соответствует числу объектов) и k столбцов (по числу факторов).

Факторные веса – это количественные значения выделенных факторов для каждого из n имеющихся объектов. У объектов с большими значениями факторных весов большая степень проявления свойств, присущих данному фактору. Факторы определяются как стандартизированные показатели со средним арифметическим значением 0 и средним квадратическим отклонением 1. Вследствие этого положительные факторные веса соответствуют тем объектам, которые характеризуются степенью проявления свойств больше средней, а отрицательные факторные веса соответствуют тем объектам, в которых степень проявления свойств меньше средней. Объекты ранжируются в пределах каждого фактора исходя из данных о факторных весах. А их, в свою очередь, можно рассматривать как значения индекса, характеризующего уровень развития объектов в рассматриваемом аспекте.

Соответственно, факторные веса могут стать основой классификации исследуемых объектов. Создание многомерной типологии на основе факторного анализа особенно эффективно в том случае, если имеется большое число признаков, характеризующих совокупность объектов, но их содержательный отбор представляет значительные трудности. В такой ситуации необходимо «сжатие» информации, после чего проводится классификация по любому из выделенных факторов.

В качестве примера эффективного использования факторного анализа можно назвать работу И.Д.Ковальченко и Л.И.Бородкина, посвященную изучению аграрной структуры районов Европейской России на рубеже XIX-XX веков⁷. Исследователи не только охарактеризовали основные компоненты аграрной структуры, определив их сравнительные доли, но и получили обобщенные характеристики общего уровня аграрного развития отдельных районов и губерний страны.

⁷ Ковальченко И.Д., Бородкин Л.И. Структура и уровень развития районов Европейской России на рубеже XIX–XX веков (Опыт многомерного анализа) // История СССР. 1981. №1.

РАЗДЕЛ 2. МОДЕЛИРОВАНИЕ В ИСТОРИЧЕСКИХ ИССЛЕДОВАНИЯХ

2.1. Виды математических моделей, применяемых в исторических исследованиях.

Основой вычислительного эксперимента является математическое моделирование. **Математическая модель** – система уравнений (дифференциальных, интегральных и алгебраических), в которой конкретные величины заменяются постоянными и переменными величинами, функциями.

Цель моделирования – замена реального объекта исследования его моделью, которую необходимо исследовать, перенося выводы на объект.

Как и в любом другом эксперименте при математическом моделировании можно выделить ряд общих этапов.

На начальном этапе для исследуемого объекта строится математическая модель. Затем разрабатывается вычислительный алгоритм (в виде совокупности цепочек алгебраических формул и логических условий). На третьем этапе осуществляется разработка компьютерной программы для реализации алгоритма, а далее проводятся собственно расчеты на компьютере. Наконец, на завершающем этапе осуществляется обработка результатов расчетов, которые подвергаются всестороннему анализу.

В литературе называется множество моделей: объясняющие и дескриптивные (описательные), теоретические и эмпирические, алгебраические и качественные, общие и частичные, модели a-priori и a-posteriori, динамические и статические, расширенные и ограниченные, имитационные и экспериментальные, детерминистические и стохастические, семантические и синтаксические.

Применение математических методов в исторических исследованиях обладает определенной спецификой.

Большинство работ, связанных с использованием математических методов в исторических исследованиях, использует статистическую обработку данных исторических источников. Но в 1980-х гг произошло совершенствование методологии исторических исследований, позволившее перейти ко второму этапу – построению математических моделей исторических процессов и явлений.

В работах И.Д. Ковальченко предложена типология моделей исторических процессов и явлений, включающая *отражательно-измерительные* и *имитационные* модели⁸. Исследователь выделяет два этапа моделирования (сущностно-содержательный и

⁸ Ковальченко И.Д. Методы исторического исследования. М., 1987. С. 365–366.

формально-количественный), отмечая, что количественное моделирование состоит в формализованном выражении качественной модели посредством тех или иных математических средств⁹.

Отражательно-измерительные модели представляют изучаемую реальность такой, какой она была в действительности, выявляя и анализируя статистические взаимосвязи в системе показателей, характеризующих изучаемый объект. Целью имитационных моделей является реконструкция отсутствующих данных о динамике изучаемого процесса на некотором интервале времени. Здесь возможен анализ альтернатив исторического развития и теоретическое исследование поведения изучаемого явления (или класса явлений) по построенной математической модели. Выделяют два типа имитационных моделей: *имитационно-контрфактические* и *имитационно-альтернативные* модели исторических процессов.

Обычно контрфактическое моделирование ассоциируется с произвольным перекраиванием исторической реальности, но, с другой стороны, оно может быть эффективным инструментом изучения альтернативных исторических ситуаций. Здесь находят применение аналитические и имитационные модели. Для первых характерна запись процессов функционирования рассматриваемой системы в виде функциональных соотношений (уравнений). *Имитационные* модели воспроизводят сам изучаемый процесс в его функционировании во времени. При этом, имитируются элементарные явления с сохранением их логической структуры и последовательности протекания во времени. С помощью моделирующего алгоритма, по исходным данным о начальном состоянии процесса (входной информации) и его параметрах, можно получить сведения о состояниях процесса на каждом последующем шаге. Преимущество имитационных моделей по сравнению с аналитическими заключается в том, что в них появляется возможность моделирования весьма сложных процессов (с большим числом переменных, нелинейными зависимостями, обратными связями), которые не поддаются аналитическому исследованию. Основным недостатком имитационного моделирования является тот факт, что полученное решение (динамика моделируемого процесса) всегда носит частный характер, отвечая фиксированным значениям параметров системы, входной информации и начальных условий.

Существенное внимание в моделировании привлекают проблемы *верификации* моделей историко-социальных процессов; при этом для многих математических и имитационных моделей параметры фиксируются *a priori*, в то время как в статистических моделях параметры оцениваются из данных, которые верифицируют эту модель.

Решение вопроса о применении математического, статистического или имитационного моделирования для построения теории, зависит от характера и объема имеющихся исходных данных.

⁹ Там же. С. 366.

СРАВНЕНИЕ ТРЕХ ПОДХОДОВ К МОДЕЛИРОВАНИЮ ДИНАМИКИ¹⁰

	<i>Аналитические модели</i>	<i>Статистические модели</i>	<i>Имитационные модели</i>
<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>
<i>Примеры</i>	Дифференциальные уравнения; марковские цепи.	Регрессионные уравнения, фактор-анализ, log-линейные модели.	Системы конечно-разностных уравнений
<i>Ограничения</i>	Одно или несколько уравнений и переменных, простая форма взаимосвязей между ними.	Малое число уравнений, большое число переменных, более сложные связи между ними. Обратные связи трудны для исследования.	Допускается большое число переменных и уравнений. Сложная форма взаимосвязей между ними.
<i>Требования к данным</i>	Модели являются дедуктивными, выводимыми из теории. Данные различного качества необходимы для подтверждения надежности модели.	Модели выводятся из предположений о роли факторов, с привлечением большого количества данных высокого качества.	Модели отчасти выводятся из теории. Возможны данные низкого качества для подтверждения надежности модели.
<i>Значение для построения теории</i>	Ориентированы на анализ динамики. Упрощенное представление о переменных и связях между ними. Результаты моделирования выводятся путем	Весьма ограниченные формы динамических связей. Тенденция к построению сложных измерительных теорий. Дедукции из модели являются тривиальными.	Ориентированы на анализ динамики и допускают нелинейные связи. Тенденция к построению сложных эмпирико-дедуктивных теорий.

	аналитического решения. Предполагаются детерминистические связи между переменными.	Предполагаются стохастические взаимосвязи.	Предполагаются как детерминистические, так и стохастические связи.
<i>Верификация модели</i>	Параметризация проводится либо а priori, либо статметодами. Применение может быть весьма ограниченным. Параметризованные тесты на хорошее соответствие модели возможны только статистическими методами. При несоответствии модели дается некоторая специфическая диагностическая информация.	Параметризация проводится статметодами, из данных. Предположения для оценки могут быть очень сложными для выполнения (например, структура ошибок). Разработаны критерии верификации. Некоторая диагностическая информация возможна в случае соответствия модели данным.	Параметризация проводится либо а priori, либо статметодами. Эмпирически можно проводить сильные тесты модели. Ошибкам измерения особого внимания не уделяется. Отсутствуют параметризованные тесты на соответствие модели. Диагностика в случае несоответствия модели весьма неудовлетворительна.

Математические (аналитические) модели сильны для оценки теории, но при моделировании сложных процессов они становятся очень трудными для решения и понимания, а также могут встретить существенные трудности при верификации на основе ненадежных данных. Имитационные же модели имеют гораздо больше преимуществ, при отображении сложных эмпирических и теоретических взаимосвязей.

2.2. Математические методы в классической и экспериментальной археологии

Задачей археологии является реконструкция целого по неполным данным либо выделение существенных черт целого из большого объема данных. Именно поэтому в археологии широкое применение получили статистические методы. Множество находимых артефактов необходимо упорядочивать, сводить в классы и типы, а это невозможно без их математической обработки. Если при работе с задачами первого типа происходит домысливание, индуктивное расширение информации на основе меньшего количества фактических данных, то второй тип задач характеризуется свертыванием, сжатием информации. При этом большой объем фактических данных подвергается статистическому анализу для того, чтобы выделить его существенную часть, либо

сформировать обобщенные интегрированные показатели. С целью решения этих задач археология одной из первых в числе исторических наук обратилась к математическим методам, а впоследствии и к информационным технологиям. Методы вариационной статистики и геометрии использовались в работах российских археологов уже в 20-х годах XX века.

На Западе статистические методы при изучении палеолитических индустрий были применены Альфредом Киддером в 1936 году, а уже в 40-е годы математические методы стали широко применяться в Америке. Последующие работы Джорджа Брейнерда показали, как можно формулировать и математически решать археологические задачи.

Статистические методы могут быть использованы в том случае, если экспериментальные данные представляют собой значительный объем результатов проведенных «измерений». При этом структура совокупности исходных данных должна содержать в себе определенную неоднородность, выражающую различные соотношения зависимости. В этом случае, проводимый статистический анализ археологических данных позволит выявить скрытые в материале закономерности, для чего существует детально разработанная теория измерений, которая определяет виды признаков и шкал.

На первом этапе археолог составляет более или менее полное описание обнаруженных предметов или следов объектов (построек), и уже на втором этапе производит более или менее обоснованную реконструкцию культурных и исторических реалий, связанных с обнаруженными артефактами.

Вот несколько примеров реконструкционной деятельности, позволяющих восстановить материальную составляющую наших знаний о прошлом.

Ярко выраженным примером проявления возможностей математики в исторической реконструкции является восстановление орнамента на самом широком круге артефактов, ведь орнамент – неотъемлемая часть декорирования многих из них, отражающая общие культурные понятия наших предков. Этот метод основан на исследовании формообразования орнаментальных композиций и методов гармонизации формы при помощи симметричных преобразований, теоретической базой которых послужили работы в области феномена симметрии. Он осуществляется при помощи выборочного статистического наблюдения, результаты которого организуется по принципу случайного отбора. Генеральной совокупностью в данном случае будут все орнаментальные изображения, свойственные той или иной местности в определенный временной отрезок. Далее вычисляются границы доверительных интервалов. После чего строится система соотнесения встречающихся в орнаменте фигур, которые должны быть тщательно изучены, обмерены и классифицированы. В дальнейшем происходит визуальное отнесение тех или иных фигур к конкретным группам образов, которые затем проверяются по каждой части образа. Эти части образа – элементы. Именно они являются объектом статистического наблюдения. Выделенные элементы подразделяются на ряд типов, каждый из которых включает в себя наиболее близкие по графическому контуру конкретные изображения элементов фигур. Рассмотреть вкуче все признаки и элементы с помощью простого наблюдения невозможно. Данную задачу под силу решить лишь только с помощью четкой росписи признаков (качественных и количественных) по определенной системе и последующей обработки этих росписей на счетных машинах.

Метод симметрии используется также при реконструкции конкретных строений и их декора. Так он очень эффективен при восстановлении частично утраченных мозаик,

площадь и границы которых достоверно прослеживаются в ходе археологических раскопок.

Эмпирической базой, обеспечивающей математическое моделирование как отдельных технологических процессов, так и важнейших элементов древних экономик, является экспериментальный метод, который позволяет определить, например, производительность труда в рамках различных производственных процессов, возможности древних транспортных коммуникаций и т.п.

С середины XIX в. экспериментальные методы в археологии постепенно занимают достаточно устойчивые позиции. Так, в 1874 г. во время археологической конференции в Копенгагене была продемонстрирована деревянная постройка, срубленная при помощи каменных орудий. В конце XIX в. Отто Тишлером была экспериментально доказана возможность сверления каменных изделий при помощи деревянного сверла и подсыпаемого под него песка. В 1883 г. была сделана первая экспериментальная попытка доказать возможность плавания в Америку до Христофора Колумба: судно "Викинг", представлявшее собой копию драккара из Гокстада (IX в., Норвегия), за 40 дней благополучно достигло берегов Нового Света.

В 20-х гг. XX в. возникает новая форма археологического эксперимента: исследования приобрели комплексный характер, а участники экспериментов начинают использовать так называемые методики "погружения в историческую эпоху". Так, например, в Швейцарии на берегах Боденского озера были воссозданы поселения каменного и бронзового веков. В Польше подобный центр возник в Бискупине, где с высокой степенью точности было реконструировано городище эпохи железа. Начиная с 1936 г., на территории этого поселения проводились исследования различных архаичных хозяйственных процессов: при помощи древних орудий труда экспериментаторы готовили пищу, реконструировали процессы охоты и обработки земли.

Мощный прорыв в развитии экспериментальных методик в археологии был сделан в 1950-е годы в СССР: С.А. Семеновым была предложена оригинальная методика определения и изучения функций орудий труда по характеру следов работы (трассеологический метод), благодаря которой, например, было установлено, что эффективность рубки леса репликами каменных топоров всего в 3-4 раза ниже, чем при выполнении аналогичной работы современными топорами, сделанными из железа.

Еще одной областью все более широко применяемых методов математических вычислений является т.н. экспериментальная военная археология, стремящаяся к адекватной и научно-обоснованной исторической реконструкции. Исследования, проводимые в данном контексте, включают результаты по формализации и синтезу математических моделей, а также разработку программных средств и получение с их помощью исторических данных, обеспечивающих проведение исторической реконструкции как артефактов, использующих метательный принцип поражения (древних стрел и раннего огнестрельного оружия), так и процессов (боевого функционирования древнего городища или доспеха как объекта, подверженного воздействию метательного оружия; динамической модели сухопутного или же морского боя и т.п.), позволяющих извлечь из исторического источника максимум неявной информации, скрытой при применении традиционных методов исторических исследований.

Подводя итог краткого экскурса, повествующего о применении математических методов в классической и экспериментальной археологии, следует отметить, что к настоящему времени накоплен значительный опыт применения математических методов в археологии,

имеется достаточное количество фундаментальных и основополагающих публикаций по этим вопросам. Однако на сегодняшний день говорить о них как о полностью сформировавшихся научных направлениях преждевременно, поскольку еще не достигнуто определенное согласование предметной области и методов собственно археологии с соответствующими методами математики, компьютерной технологии обработки и анализа информации. Именно это является главной проблемой исследований в области применения информационных технологий в исторической науке в целом.

2.3. Проблемы исторического моделирования. Клиодинамика в реконструкции прошлого и прогнозах будущего

В конце XX века в связи с развитием нелинейной науки (нелинейной динамики и синергетики) произошла научная революция в сфере прогнозирования. Начало ей было положено с открытием в 1963 году явления динамического хаоса, аттракторов и горизонта прогноза. Другим важнейшим понятием в рамках нелинейной науки следует считать бифуркацию (от французского bifurcation – раздвоение, ветвление). Развитие сложных систем сейчас обычно мыслится как прохождение с течением времени последовательности точек бифуркаций, в каждой из которых фактически делается выбор одного из вариантов развития.

Одной из пионерских идей XX века стало представление о самоорганизации или синергетики (дословно с греческого – теории совместного действия)– спонтанном, самопроизвольном возникновении упорядоченности в открытых нелинейных, далеких от равновесия системах.

В России становление синергетики во многом связано с именем специалиста по прикладной математике – С. П. Курдюмова. Его научной школе принадлежит создание и развитие теории режимов с обострением, при которых одна или несколько величин, характеризующих систему, неограниченно возрастают за ограниченное время. Все эти представления нелинейной науки – горизонт прогноза, самоорганизация, параметры порядка, режимы с обострением – используются в стратегическом прогнозировании. Однако, не нужно забывать и законы, выведенные для анализа основополагающих процессов, определяющих алгоритм развития человечества. Это процессы демографического роста, начало изучения которых было положено еще в XVIII веке английским священником Томасом Мальтусом. Его учение получило название у философов – мальтузианство. Т. Мальтус уловил суть основных математических закономерностей развития популяции любых живых существ, которые занимают экологическую нишу с ограниченными ресурсами. Согласно Мальтусу, пока ресурсов много, а численность популяции мала, то скорость ее прироста прямо пропорциональна ее числу. Но когда численность популяции стабилизируется на уровне голодного выживания, минимального душевого потребления, когда голодная смертность компенсирует рождаемость. В конкурентной борьбе за ресурсы нарастает взаимная

агрессия, заканчивающаяся катастрофой. Среди животных это эпидемия, среди людей — еще и войны, революции, которые знаменуют собой конец одного цикла и начало нового.

При этом потенциал роста популяции превышает потенциал роста ее производительных сил. По расчетам Т. Мальтуса население растет в геометрической прогрессии, а производительные силы — только в арифметической. Вышеописанный цикл получил название «мальтузианской ловушки». С современной точки зрения мальтузианская теория имеет много минусов. Так Т. Мальтус не учитывал фактор миграции населения, а также не принимал во внимание механизмы саморегуляции численности человечества, приводящие к демографическому переходу. По мнению Мальтуса ни накопление капитала, ни научно-технический прогресс не компенсируют ограниченность природных ресурсов. Вместе с тем, теория Мальтуса достаточно корректно описывает закономерности экономико-демографической динамики по отношению к доиндустриальному обществу.

Теория демографического цикла Мальтуса в последующее время подверглась переработке, и к сегодняшнему дню уже выявлены десятки демографических циклов в истории Древнего мира, Европы, России, Востока и Китая, которые, как правило, заканчивались демографическими катастрофами, будь то страшные эпидемии или войны.

Российский ученый С. Нефедов построил математическую модель демографического цикла¹¹. Он выделил 40 социально-экономических параметров общества, определяющих цикл. Это позволяет использовать данный комплекс при выделении демографического цикла и без достаточных данных демографической статистики. При умелой постановке исследовательской задачи получаются весьма интересные результаты. Так произошло и при рассмотрении вопроса о причинах возникновения авторитарных режимов.

Авторитарные режимы более характерны для стран с высокой плотностью населения. При наложении на эти данные модели С. А. Нефедова выяснилось, что формирование авторитаризма приходится преимущественно на конец цикла, предшествующий демографической катастрофе. Именно тогда тяжелая и скудная жизнь народа требует деспотизма, безжалостно подавляющего вспышки насилия. Именно революции завершали некоторые из демографических циклов. Пытаясь разгадать загадку революций, американский социолог Дж. Голдстоун обратил внимание на неоднородность человеческих сообществ, выделив на первое место статусную неоднородность по отношению к власти, разделив общество на три страты: власть (людей, наделенных властными полномочиями), элиту и народ. Страты имеют различный доступ к ресурсам. Народ является источником ресурсов для государства и элит, а государство перераспределяет ресурсы через бюджет. Элиты пользуются поддержкой государства, но при наступлении дефицита ресурсов возникает так называемое «перепроизводство элит». Растущим элитам необходимы ресурсы, которые вот-вот иссякнут, и положение не может исправить даже введение жесткой фискальной политики, которая лишь усугубляет недовольство со стороны народа, не имеющего возможности наполнять бюджет. Недовольство растет как в народе, так и в элитах. В находящихся между двух полюсов элитах начинается расслоение, часть их уходит на сторону народа.

¹¹ Нефедов С. А. Концепция демографических циклов. Екатеринбург, 2007.

Структурно-демографический цикл с учетом социального расслоения, описанный Дж. Голдстоуном, обрел математическую модель в 2007 г. благодаря усилиям С. А. Нефедова.

П. В. Турчин – профессор Коннектикутского университета (США), активно разрабатывающий направление клиодинамики (Клио — муза истории), рассматривая причины русской революции, акцентирует внимание на таких ее причинах как быстрый рост социального неравенства и перепроизводство элиты. Кризис биологического существования населения в своих построениях он обходит стороной. Поэтому его построение¹² нельзя отнести ни к неомальтузианству, ни к структурно-демографической теории, согласно которой политический кризис вызывается перепроизводством населения (именно всего населения, а не только элиты) вследствие недостатка ресурсов. П. В. Турчин смещает центр концепции на обеднение и перепроизводство элиты как самостоятельный и решающий фактор исторической динамики, вследствие чего он попадает в серьезные методологические трудности, лишаящие его концепцию точности, строгости, объективности. Его концепция не верифицируется, так как потребности элиты – понятие субъективное и совершенно неопределенное.

Тем не менее, несмотря на все методологические нестыковки и споры, вызванные различными концепциями отдельных исследователей, появление социальной структуры в математической модели демографического цикла — это огромный шаг вперед, благодаря которому исследователь уже на математическом уровне может поставить перед собой, например, такие задачи как раздельное исследование влияния на демографический цикл физического исчерпания ресурсов и их исчерпания вследствие нарастания социальных различий.

Кроме того, следует упомянуть также теорию смены технологических укладов, которые являются движущей силой технологического и социального прогресса, многих кризисов, и, в частности, того, который сейчас переживает современный мир. Это направление экономической науки было развито выдающимся экономистом И. Шумпетером (1883–1950).

По мнению П.В. Турчина XXI век станет веком наук о развитии человеческого общества, но рассмотрен он будет с применением математизации. Математические абстрактные схемы он предлагает накладывать на историческую данность – благо в исторических источниках сохранилось множество количественных параметров, которые можно сравнивать друг с другом, производить различные вычисления с ними и пытаться на основании этих подсчетов выявить модели исторического развития общества.

2.4. Моделирование средствами фрактальной геометрии

В последние годы теория хаоса в виде синергетических представлений активно вторгается в гуманитарные науки. Согласно некоторым исследованиям последних лет, многие социальные процессы, подобно природным явлениям, имеют фрактальную структуру и

¹² Турчин П. В. Историческая динамика: На пути к теоретической истории. М., 2007. С. 176.

развиваются в соответствии с фрактальными закономерностями. При помощи фрактальной геометрии можно создавать модели, подходящие для имитации нелинейности, парадоксальности процессов и структур. Методы фрактальной геометрии широко разрабатываются начиная с 1977 года, после публикации работ Бенуа Мандельброта¹³. Фрактал – это особый тип геометрической фигуры, а «фрактальный» – это характеристика структуры, явления или процесса, обладающих свойствами фрактала. Фрактальные структуры обнаруживаются в природе и изучаются в исследованиях по естественным наукам и связанным с ними прикладных отраслях знания. В последнее время фрактальная теория и методология применяется в социально-экономических и гуманитарных отраслях знания, в том числе и в истории, так как, Согласно некоторым исследованиям последних лет, многие социальные процессы, подобно природным явлениям, имеют фрактальную структуру и развиваются в соответствии с фрактальными закономерностями.

Главным свойством фракталов является, самоподобие, т.е. даже малая часть фрактала содержит: информацию о фрактале в целом. Поэтому, в каком бы масштабе не рассматривался фрактал, всегда будет видно одно и то же. В качестве примера на рисунке представлены несколько масштабов фрактала «Решето Серпински».

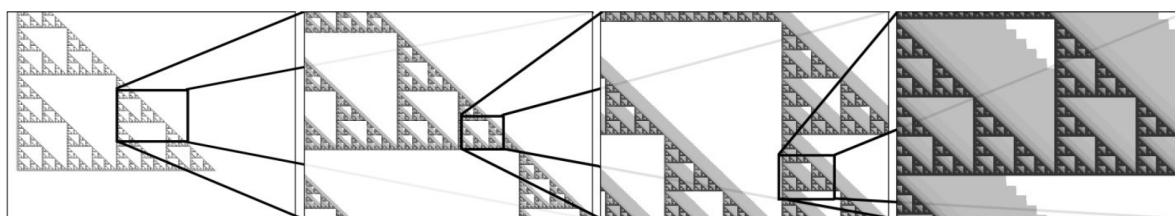


Рисунок Решето Серпински (по Д.С. Жукову и С.К. Лямину, 2011)

Однако, самоподобие бывает частичным или закономерно изменяющимся; на каждом уровне сложности возникает сплав индивидуальных и общих черт всей системы.

По мнению Б. Мандельброта, для природы характерен именно фрактальный способ самоорганизации, благодаря чему можно создавать математические модели социальных и политических явлений и процессов. При помощи специально написанных программ-фракталопостроителей можно проводить компьютерные эксперименты, симулирующие такие явления и процессы. Существует несколько способов такого моделирования.

Во-первых, нелинейную динамическую систему в фазовом пространстве можно рассматривать с помощью итерированной формулы, которая описывает поведение точки (системы в фазовом пространстве). При этом, формула создает череду чисел, значения которых отображают траекторию системы, а совокупность начальных условий системы во многих случаях имеет вид фрактала. Поэтому, при помощи математического описания ряда факторов системы, можно предсказывать ее возможное развитие¹⁴.

¹³ Mandelbrot B.B Fractals: Form, Chance, and Dimension. San Francisco, 1977; Ibid. The Fractal Geometry of Nature. Oxford, 1982.

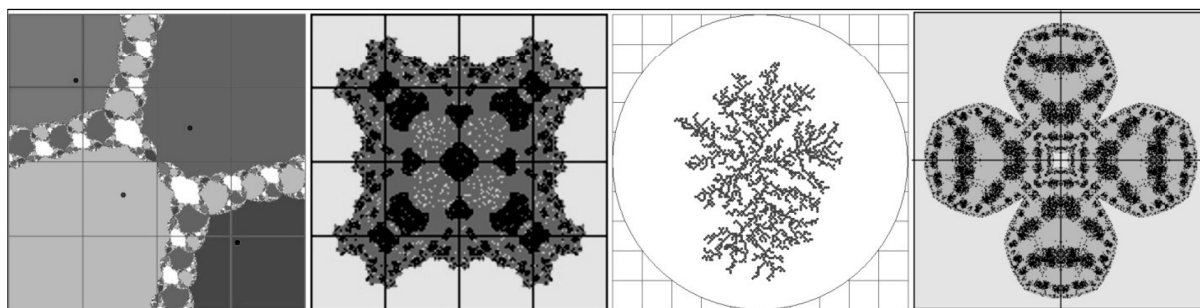
¹⁴ Жуков Д.С., Лямин С.К. Моделирование исторических явлений и процессов средствами фрактальной геометрии // Информационный бюллетень Ассоциации «История и компьютер». 2006. № 34. С. 52

В ходе исследований в Центре фрактального моделирования Д.С.Жуковым и С.К. Ляминим была предложена математическая модель для описания процессов модернизации городской социальной среды и менталитета горожан в пореформенной России («Менталофрактал»), а также разработана модель демографического поведения аграрного населения Центральной России второй половины XIX – начала XX вв. («Демофрактал»). В основу обеих моделей положен сходный математический аппарат, так как они предназначены для имитации схожих процессов форсированной модернизации. В обоих случаях формула аналогична используемой для построения «Фрактала Мандельброта». Вместе с тем, алгоритм генерирования значительно отличается. Математический аппарат «Менталофрактала» и «Демофрактала» содержит итерируемую формулу $Z_{n+1} = Z_n^2 A + C$, (где Z и C – комплексные числа), а также ряд математических условий, позволяющих отождествить геометрический смысл операций над комплексными числами с результатами нуклеарных взаимодействий факторов модели.

Во-вторых, с помощью фракталов появляется возможность имитировать реальные процессы, если в процедуру построения вводятся элементы случайности. Д.С. Жуков и С.К. Лямин применили такой метод моделирования для изучения формирования влияния государства на городские общества во второй половине XIX века¹⁵. С помощью компьютерной программы «Имитация» был сформирован фрактальный кластер, конфигурация которого имитировала результаты взаимодействия следующих факторов: сила модернизационного нажима, инерция (сила сопротивления) традиционного общества, величина объекта модернизационного нажима, количество модернизационных мероприятий. В качестве результатов реального исторического процесса модернизации можно интерпретировать графические результаты работы программы.

Но при разных запусках программы с одними и теми же параметрами вид получившегося фрактала может быть различным, хотя при этом качественные характеристики (величина, «степень разветвлённости» и др.) будут одинаковы, поскольку выражают статистические закономерности взаимодействия вводимых одинаковых параметров.

Некоторые результаты исследований, связанных с построением алгебраических и стохастических фракталов представлены на рисунке ниже.



¹⁵ Жуков Д.С., Лямин С.К. Фрактальное моделирование социально-политических феноменов и процессов // Pro nunc. Современные политические процессы. 2011. Т. 10 №1. С. 161–171.

Рисунок. Некоторые результаты работы фракталопостроителей «Менталофрактал», «Демофрактал», «Имитация» (по Д.С. Жукову и С.К. Лямину, 2011)

Ценность имитационной модели заключается в том, что такая модель позволяет выявить потенциал развития ситуации, если вводить разные значения параметров. В таком случае получаются разные результаты – разные фрактальные кластеры. И самое главное – каждый изолированно рассматриваемый кластер не приносит никаких новых знаний, однако в этом кластере видна взаимосвязь всех исследуемых факторов, вследствие чего череда кластеров позволяет сравнить результаты изменения как одного, так и нескольких факторов.

Метафора фрактала, которому свойственна масштабная инвариантность, позволяет свести всё многообразие фактов, независимо от их масштаба, к определённой закономерности. В этом случае происходит не просто изменение визуального ряда, но происходит смена представлений о существе явлений.

Таким образом, с помощью фрактальной геометрии можно анализировать событийные ряды, и предположить, что фрактальным структурам соответствуют фрактальные процессы их жизнедеятельности. Закручивающимся спиралям в фазовом пространстве в реальной жизни соответствуют затухающие колебательные процессы, характерные для социально-политических процессов на стадии стабилизации. «Раскручивающейся» спирали фазового пространства с аттрактором в бесконечности соответствуют колебательным процессам реального мира, приводящим к дестабилизации и разрушению системы.

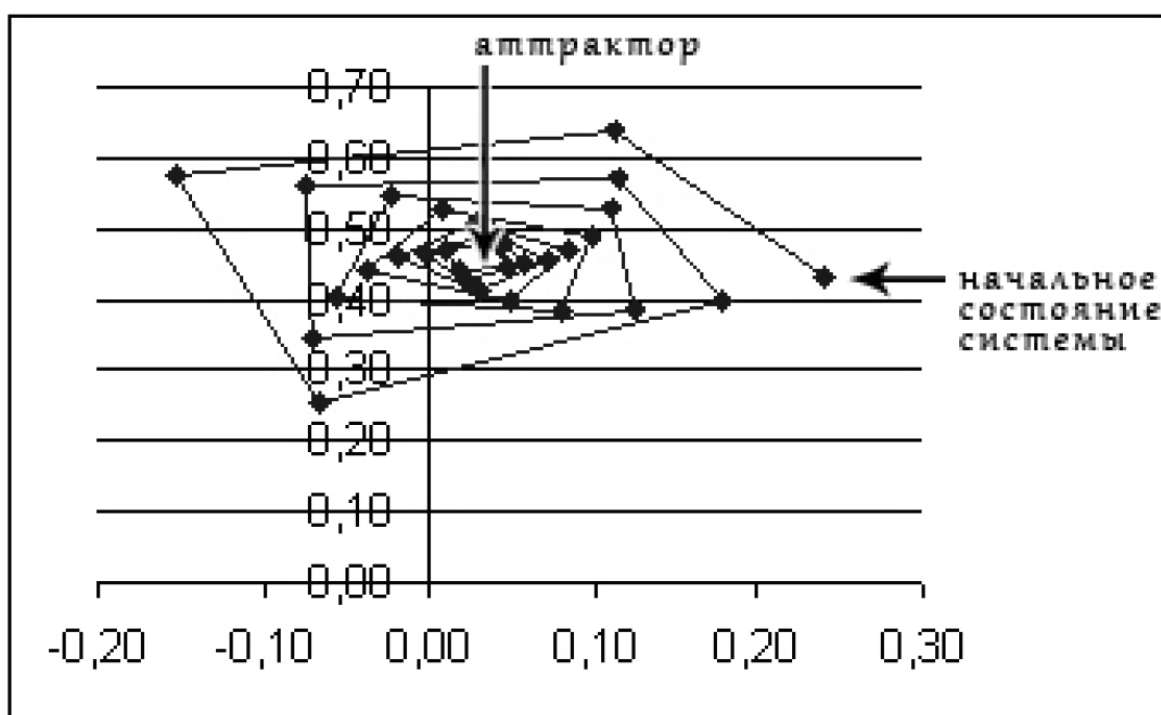


Рисунок 5. Один из результатов интераций формулы Мандельброта (по Д.С. Жукову и С.К. Лямину, 2011)

По мнению разработчиков подобные модели точно не отображают историческую действительность, однако обобщения нескольких факторов, позволяющие произвольно экспериментировать с социальными и политическими явлениями, пригодны для выявления потенциалов и прогнозирования.

Структура и содержание дисциплины «Математические методы в исторических исследованиях»

1. Цели освоения дисциплины

Целями освоения дисциплины «Математические методы в исторических исследованиях» является ознакомление студентов с основными понятиями теории вероятностей, математической статистики, математического моделирования и спецификой их применения в историческом исследовании.

2. Место дисциплины в структуре ООП

Данная дисциплина входит в базовую часть математического и информационно-технологического цикла.

Указанная дисциплина преподается параллельно с информатикой, которая развивает представления о компьютерных методах решения математических задач в истории.

Для освоения данной дисциплины необходимо знание школьного курса математики и информатики. Необходимо, чтобы студент умел производить вычисления над натуральными, целыми, рациональными и действительными числами, мог подсчитать десятичный и натуральный логарифм числа, знал виды функции и способы их задания, смог вычислить площадь с помощью интеграла, смог дифференцировать функцию (найти ее производную).

3 Требования к результатам освоения дисциплины «Математические методы в исторических исследованиях»

Процесс изучения дисциплины направлен на формирование следующих общекультурных компетенций:

1. владение культурой мышления, способностью к обобщению, анализу, восприятию информации, постановке цели и выбору путей ее достижения (ОК-1);
2. способность использовать в познавательной и профессиональной деятельности базовые знания в области основ информатики, элементы естественно-научного и математического знания (ОК-13).

В результате освоения дисциплины обучающийся должен:

Знать разделы математики, которые нашли применение в исторических исследованиях (математическая статистика, методы математического моделирования).

Уметь применять адекватные методы статистического анализа данных исторических источников.

Владеть навыками практического использования программных средств (электронные таблицы Excel и статистические пакеты SPSS или Statistica) для решения задач исторического исследования.

4. Структура и содержание дисциплины «Математические методы в исторических исследованиях»

Общая трудоемкость дисциплины составляет 4 зачетных единиц, что равняется 144 академическим часам. Для оценивания полученных знаний, умений, навыков предусматривается экзамен. Структура и содержание дисциплины отражены в таблице 1.

Таблица 1							
Структура и содержание дисциплины «Математические методы в исторических исследованиях»							
№ п/п	Раздел дисциплины	Семестр	Неделя семестра	Виды учебной работы, включая самостоятельную работу студентов и трудоемкость (в часах)			Формы текущего контроля успеваемости (по неделям семестра) Форма промежуточной аттестации (по семестрам)
				Аудиторная работа студентов	Самостоятельная работа студентов		
1	Введение	3	1	2л			
2	Основы теории множеств и математической логики	3	2-3	4л	2п/з	15	Выполнение практических упражнений в компьютерном классе, расчетных задач дома, тестовых заданий
3	Основы теории вероятностей	3	4	2л	2п/з	15	
4	Математическая статистика	3	5-16	22л	26п/з	40	
5	Методы математического моделирования исторических процессов	3	17-18	4л	2п/з	10	Выполнение работы по поиску информации в сети Интернет по теме лекции
Итого часов:				32	32	80	Экзамен
Примечания: л – лекция, п/з – практические занятия							

Конкретное содержание разделов раскрывается в следующем тематическом плане курса:

Тематический план

Раздел 1. Тема 1. Введение. Процессы информатизации и математизации современных гуманитарных наук. Степени математизации научного знания. Развитие клиометрики и квантитативной истории в зарубежной и отечественной исторической науке. Развитие исторической информатики. Основные направления развития современной количественной истории.

Раздел 2. Тема 2. Основы теории множеств и математической логики. Канторовское понятие множества. Виды множеств. Логические операции над множествами. Математическая логика – теория верных рассуждений. Понятие о высказывании, как основном понятии математической логики. Логические связки.

Раздел 3. Тема 3. Основы теории вероятностей. Случайное событие. Виды случайных событий. Вероятность случайного события. Классическое определение

вероятности. Обратная вероятность. Операции над вероятностями случайных событий. Случайная величина. Виды случайных величин. Основные свойства вероятности. Случайные величины и их распределения. Законы распределения: равномерное; нормальное; биномиальное. Числовые характеристики распределения (математическое ожидание, дисперсия, среднее квадратичное отклонение).

Раздел 4. Тема 4. Введение в математическую статистику. Дескриптивная статистика. Связь теории вероятностей и математической статистики. Виды признаков. Частота признака. Генеральная и выборочная совокупности. вычисление средних, дисперсии, среднеквадратического отклонения, коэффициента вариации. Доверительные интервалы для средних.

Раздел 4. Тема 5. Корреляционный анализ. Линейная корреляция. Парный линейный коэффициент корреляции. Коэффициент детерминации. Проверка значимости коэффициента корреляции. Понятие о нелинейной корреляции. Корреляционное отношение.

Раздел 4. Тема 6. Линейная регрессия (парная и множественная). Уравнение линейной регрессии. Интерпретация коэффициента регрессии. Множественный коэффициент корреляции. Критерии значимости уравнения регрессии. Уравнение множественной регрессии как статистическая объясняющая модель. Ограничения использования регрессионных моделей в исторических исследованиях.

Раздел 4. Тема 7. Факторный анализ. Классификация методов факторного анализа. Факторные нагрузки. Факторные веса. Доля суммарной дисперсии, объясняемой факторами. Ограничения использования факторного анализа в исторических исследованиях.

Раздел 4. Тема 8. Кластер-анализ. Классификация методов кластерного анализа. Иерархический метод, многомерная группировка. Многомерная классификация с использованием нечетких множеств.

Раздел 4. Тема 9. Анализ временных рядов. Этапы анализа временных рядов. Первичный анализ динамики ряда. Цепные и базисные, абсолютные и относительные, средние приросты, темпы роста и темпы прироста. Требования, предъявляемые к построению временного ряда в исторических исследованиях. Вычисление трендов. Сглаживание. Корреляция.

Раздел 4. Тема 10. Анализ качественных признаков. Построение таблиц сопряженности. Вычисление коэффициентов связи качественных признаков.

Раздел 5. Тема 11. Методы математического моделирования исторических процессов. Типы моделей динамики (статистические, имитационные, аналитические). Системы дифференциальных уравнений как основа для построения аналитических динамических моделей. Методы нелинейной динамики в задачах моделирования переходных и неустойчивых процессов. Синергетика в изучении историко-социальных процессов. Конечно-разностные уравнения как аппарат построения имитационных моделей. Понятие о марковских цепях. Возможности и ограничения имитационного моделирования исторических процессов.

5. Образовательные технологии

В проведении лекционных и практических занятий используются следующие образовательные технологии: вводная лекция, лекция-визуализация, лекция-информация, практические занятия, самостоятельная работа.

6. Учебно-методическое обеспечение самостоятельной работы студентов. Оценочные средства для текущего контроля успеваемости, промежуточной аттестации по итогам освоения дисциплины.

Раздел 2. Основы теории множеств и математической логики и Раздел 3. Основы теории вероятностей предусматривает решение задач по теме изучения дома.

Раздел 4. Математическая статистика предусматривает выполнение расчетных задач по каждой из тем дома.

Раздел 5. Методы математического моделирования исторических процессов предусматривает Интернет-поиск по теме изучения. Для этого студентам предлагается список ключевых слов – синергетика и история, клиодинамика, моделирование истории. Также предлагается самостоятельно ознакомиться с информационными бюллетенями Ассоциации «История и компьютер», представленных в электронном виде по адресу www.aik-sng.ru. В этих изданиях необходимо найти статьи, посвященные проблемам моделирования в истории. Результаты поиска должны быть оформлены в виде реферата.

Формами контроля самостоятельной работы по каждому из разделов дисциплины является проверка преподавателем заданий студентов (задач, расчетных задач и реферата) с выставлением оценки «зачтено» или «не зачтено».

ВОПРОСЫ ДЛЯ ПОДГОТОВКИ К ЭКЗАМЕНУ

1. Процессы информатизации и математизации современных гуманитарных наук.
2. Развитие клиометрики и квантитативной истории в зарубежной и отечественной исторической науке.
3. Основные направления развития современной количественной истории.
4. Основные характеристики массовых источников.
5. Альтернатива количественного анализа.
6. Основные характеристики качественного анализа.
7. Формы математизации научного знания.
8. Этапы клиометрического исследования.
9. Правильная постановка исследовательской задачи и ее роль в разработке изучаемой темы.
10. Дескриптивная статистика: основные понятия и характеристика.
11. Характеристики среднего значения признака.
12. Показатели меры вариации признака.
13. Понятие генеральной совокупности.
14. Выборочный метод.
15. Понятие репрезентативности выборки.
16. Виды выборочного изучения.
17. Кластерный анализ как средство типологического анализа.
18. Методика кластерного анализа.
19. Корреляционный анализ в археологии и исторической науке. Линейная корреляция.
20. Проверка значимости коэффициента корреляции.
21. Линейная регрессия (парная и множественная).
22. Понятие коэффициента детерминации.
23. Классификация методов факторного анализа.
24. Метод определения сравнительной достоверности источников.
25. Методы формирования системы репрезентативных данных источника.
26. Понятие репрезентативности «естественных выборок».
27. Суть метода «критерия знаков».
28. Математические методы в исторической науке в XIX веке.
29. Школа «Анналов», «тотальная» и количественная история.
30. Понятие и возможности информационных технологий в построении исторических реконструкций.
31. Проблемы исторического моделирования.
32. История использования математических методов в археологии.
33. Статистические методы обработки археологических данных.

34. Разработка классификации и типологии археологических объектов с помощью средств математической статистики и теории информации.
35. Математическое моделирование в исторической науке и в археологии.
36. Клиодинамика в реконструкции прошлого и прогнозах будущего.
37. Моделирование средствами фрактальной геометрии.

РЕКОМЕНДУЕМАЯ ЛИТЕРАТУРА

- Абрамов В.К. Математические методы в исторических исследованиях. Саранск, 1988
- Белова Е.Б., Бородкин Л.И., Гарскова И.М. Компьютеризованный статистический анализ для историков: учеб. пособие / под ред. Л.И. Бородкина и И.М.Гарсковой. М., 1999.
- Бородкин Л.И. Многомерный статистический анализ в исторических исследованиях. М. : Изд-во МГУ, 1986.
- Бородкин Л.И. Квантитативная история в системе координат модернизма и постмодернизма // Новая и новейшая история. 1998. №5.
- Бородкин Л.И. Клиометрика: pro et contra (виртуальный диалог) // Экономическая история. Обзорение. Вып. 7. М., 2002. С. 114–132.
- Воронов М.В., Мещерякова Г.П. Математика для студентов гуманитарных факультетов. Ростов н/Д: Феникс, 2002.
- Гарскова И.М. Базы и банки данных в исторических исследованиях. М., 1994.
- Грегори П. Поиск истины в исторических данных // Экономическая история: Ежегодник. 1999. М.: РОССПЭН, 1999. С. 471–492.
- Дюран Б., Оделл П. Кластерный анализ. М., 1977.
- Жуков Д.С., Лямин С.К. Моделирование исторических явлений и процессов средствами фрактальной геометрии // Информационный бюллетень Ассоциации «История и компьютер». 2006. № 34. С. 52
- Жуков Д.С., Лямин С.К. Фрактальное моделирование социально-политических феноменов и процессов // Pro nunc. Современные политические процессы. 2011. Т. 10 №1. С. 161–171.
- Информационные технологии в гуманитарных исследованиях: Сборник трудов. Новосибирск: НИИ МИОО НГУ, 1998.
- Информационные технологии для историков: учеб. пособие к практикуму по курсу «Информатика и математика» / Ред. Л.И.Бородкин. М.: Изд-во МГУ, 2006.
- История и Математика: Процессы и модели / Ред. Л.Е.Гринин, А.В.Коротаев, С.Ю.Малков. М., 2009.
- Ковальченко И. Д., Бородкин Л. И. Аграрная типология губерний Европейской России на рубеже XIX—XX веков: (Опыт многомерного количественного анализа) // История СССР. 1979. № 1. С. 59-95.
- Ковальченко И. Д., Бородкин Л. И. Современные методы изучения исторических источников с использованием ЭВМ: учеб. пособие. - М.: Изд-во МГУ, 1987.
- Ковальченко И.Д. Методы исторического исследования. М., 1987.
- Ковальченко И.Д., Бородкин Л.И. Структура и уровень развития районов Европейской России на рубеже XIX-XX веков (Опыт многомерного анализа) // История СССР. 1981. №1.

- Ковальченко И.Д., Милов Л.В. Всероссийский аграрный рынок. XVIII - начало XX в. Опыт количественного анализа. М., 1974.
- Количественные методы в исторических исследованиях: [учеб. пособие для вузов по специальности "История"] /Ред. Л.И. Бородкин, И.М. Гарскова, Т.Ф. Измestьева и др. М.: Высшая школа, 1984.
- Количественные методы в советской и американской историографии: материалы совет.-американ. симпоз. в г. Балтиморе, 1979 г. и г. Таллине, 1981 г. М.: Наука, 1983.
- Круг идей: Модели и технологии исторических реконструкций. Труды XI конференции ассоциации «История и компьютер». Москва-Барнаул-Томск, 2010.
- Математические методы в социально-экономических и археологических исследованиях. М., 1981.
- Математические методы и ЭВМ в историко-типологических исследованиях / Под ред. И.Д. Ковальченко. М., 1989.
- Мазур Л.Н. Методы исторического исследования. Екатеринбург, 2010.
- Методы количественного анализа текстов нарративных источников. М., 1983.
- Мионов Б.Н. История в цифрах. Л., 1991.
- Мионов Б.Н. Хлебные цены в России за два столетия (XVIII–XIX вв.). Л., 1985.
- Мионов Б.Н., Степанов З.В. Историк и математика. М., 1975.
- Нефедов С. А. Концепция демографических циклов. Екатеринбург, 2007.
- Пейте Ф. Выборочный метод в переписях и обследованиях. М., 1965.
- Проблемы математической истории: Математическое моделирование исторических процессов / Ред. Г. Г. Малинецкий, А. В. Коротаев. М.: ЛИБРОКОМ/URSS, 2008.
- Тарасенко В.В. Фрактальная логика. М., 2002.
- Турецкий В.Я. Математика и информатика. - 3-е изд., перераб. и доп. М.:ИНФРА-М, 2007.
- Турчин П. В. Историческая динамика: На пути к теоретической истории. М.: Издательство ЛКИ, 2007.
- Турчин П. В. Историческая динамика: На пути к теоретической истории. М., 2007.
- Турчин П.В.* Перспективы математической истории. Существует ли качественное различие между исторической и естественными науками? // История и Математика: Концептуальное пространство и направления поиска / Ред. П. В. Турчин, Л. Е. Гринин, С. Ю. Малков, А. В. Коротаев. М., 2007. С. 8—18.
- Устинов В.А., Фелингер А.Ф. Историко-социальные исследования, ЭВМ и математика. М., 1973.
- Фёдоров-Давыдов Г.А. О статистическом исследовании взаимовстречаемости признаков и типов предметов в археологических комплексах // Статистико-комбинаторные методы в археологии. М., 1970. С. 123 – 131.
- Фёдоров-Давыдов Г.А. Статистические методы в археологии. М.: Высшая школа, 1987.
- Хвостова К.В. Количественный подход в средневековой социально-экономической истории. М., 1980.
- American Historical Review, February 1998.
- Boldizzoni F. The Poverty of Clio: Resurrecting Economic History. Princeton University Press, 2011.
- Greif A. Analytic Narratives. Oxford University Press, 1998.
- Johnson, E.A. Counting "how it was": Quantitative history in West Germany. In: Historical Methods, 1988. Vol. 21.
- Mandelbrot B.B Fractals: Form, Chance, and Dimension. San Francisco, 1977.
- Mandelbrot B.B The Fractal Geometry of Nature. Oxford, 1982.
- North D., Thomas R. The rise of the Western world: a new economic history, Cambridge University Press, 1973.
- Oberwittler, D. From Coding to Decoding? An Analysis of Historical Social Research in Germany in the 1980s and Early 1990s. In: Historical Methods, Fall 1997, Vol. 30.
- Welling, G. The Prize of Neutrality. A Study in computational history. Amsterdam, 1998.

