

Н.А. ДЕГТЯРЕВА

**ИССЛЕДОВАНИЕ ИНФОРМАТИЗАЦИИ ОБРАЗОВАНИЯ
С ПРИМЕНЕНИЕМ ЭКОНОМЕТРИЧЕСКИХ МОДЕЛЕЙ**

МОНОГРАФИЯ

**Челябинск
2021 г**

УДК 371 : 33
ББК 74.04 : 65в6
Д 26

Дегтярева, Н.А. Исследование информатизации образования с применением эконометрических моделей: монография / Н.А. Дегтярева. – Челябинск: Изд-во ЗАО «Библиотека А.Миллера», 2021. – 160 с.

ISBN 978 – 5 – 91283 - 559-1

В данном исследовании представлены наиболее актуальные проблемы эконометрического моделирования в области информатизации образования, представлена теоретическая и методологическая основа исследования с применением эконометрических методов и моделей, приведены исследования в области применения эконометрических методов и моделей в области информатизации образования для принятия оптимальных управленческих решений.

Рецензенты:

В.Н. Павленко, доктор физико-математических наук, профессор кафедры вычислительной математики ФГБОУ ВО «Челябинский государственный университет».

И.Д. Кутузов, кандидат физико-математических наук, доцент, ФГБОУ ВО «Челябинский государственный университет».

ISBN 978 – 5 – 91283 - 559-1

Дегтярева Н.А, 2021

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ.....	3
1 ЭКОНОМЕТРИЧЕСКОЕ МОДЕЛИРОВАНИЕ.....	9
1.1 Типы переменных и типы данных.....	9
1.2 Основные этапы эконометрического моделирования.....	14
1.3 Классификация моделей.....	20
1.4 Линейная парная регрессия.....	24
2 МОДЕЛЬ МНОЖЕСТВЕННОЙ РЕГРЕССИИ.....	39
2.1 Метод наименьших квадратов оценки параметров.....	39
2.2 Корреляция в КЛММР.....	46
2.3 Оценка качества модели множественной регрессии.....	48
2.4 Мультиколлинеарность и методы ее устранения.....	51
2.5 Линейная модель множественной регрессии с гетероскедастичными остатками.....	58
2.6 Линейная модель множественной регрессии с автокорреляцией остатков.....	66
3 ИСПОЛЬЗОВАНИЕ ЭКОНОМЕТРИЧЕСКИХ МОДЕЛЕЙ В ИССЛЕДОВАНИИ СОЦИАЛЬНЫХ ПРОЦЕССОВ.....	81
3.1 Линейная модель с количественными переменными.....	81
3.2 Линейная смешанная модель (с количественными и качественной переменной).....	103
3.3 Прогнозирование.....	106
3.4 Численное решение задачи с помощью пакета Econometric Views.....	110
ЗАКЛЮЧЕНИЕ.....	123
ГЛОССАРИЙ.....	127
СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ.....	148
ПРИЛОЖЕНИЯ.....	160

ВВЕДЕНИЕ

Информатизация образования – широкомасштабный процесс трансформации содержания, методов и организационных форм учебной работы, обеспечивающий подготовку (эффективную социализацию) школьников к жизни в условиях информационного общества. В западных публикациях такой тип общества сегодня называют «Knowledge Society» — «общество, основанное на знаниях» или «общество знаний». Отечественная традиция использует вместо термина «общество знаний» термин «информационное общество», в ходе движения к которому и разворачивается процесс информатизации школы [25].

Информатизация не ограничивается простым увеличением объемов циркулирующей в мире информации, а также автоматизацией ее сбора, обработки, передачи, хранения, подготовки и представления. Она связана с новым качеством жизни. Подчеркивая это обстоятельство, экономисты стали говорить, что информационная экономика есть экономика, основанная на знаниях. В начале 70-х годов за рубежом отношение рыночной стоимости компаний к их балансовой стоимости составляло менее единицы (около 0,8).

Становление индустриального общества привело к построению школы индустриального общества. Происходящее на наших глазах становление информационного общества также не может не затронуть один из главных механизмов воспроизводства современного человека – общеобразовательную школу [34].

За прошедшие тысячелетия школа как общественный институт пережила не одну трансформацию и смену парадигм [48]. И каждый раз

находились пути и средства, чтобы провести соответствующие преобразования. Сегодня нам нужны способы описания происходящих в школе трансформаций и их оценки. Философия и история педагогики подсказывают общие тенденции развития школы. Мы вынуждены сами искать инструменты для описания этого процесса, разрабатывать модели для управления им. Как показывает опыт подготовки и проведения проектов информатизации системы образования, в нашем распоряжении пока нет адекватных средств, позволяющих описывать направления и оценивать результаты работ по информатизации школы. А это значит, что у нас нет общего языка, на котором мы могли бы обсуждать наши успехи и неудачи, планировать желаемые изменения и оценивать результаты своих усилий.

Поиски решения данной проблемы привели к разработке инструмента, который позволяет обсуждать информатизацию школы не только как глобальный процесс, но и как управляемый поток последовательных изменений в работе отдельных образовательных учреждений и педагогов [20,24].

Использование возможностей современной вычислительной техники, оснащенной пакетами программ машинной обработки статистической информации на ЭВМ, делает практически осуществимым оперативное решение задач изучения взаимосвязи показателей деятельности различных организаций методами корреляционно-регрессионного, факторного и компонентного анализа [31].

При машинной обработке исходной информации на ЭВМ, оснащенных пакетами стандартных программ ведения анализов, вычисление параметров применяемых математических функций

является быстро выполняемой счетной операцией. Результаты выдаются в виде соответствующих распечаток ЭВМ [67].

На сегодняшний день статистические программные продукты (СПП) являются доступными средствами анализа данных. В мировой практике компьютерные системы статистического анализа и обработки данных широко применяются, как в исследовательской работе в области экономики, так и в практической деятельности аналитических, маркетинговых и плановых отделов различных организаций. В последние годы заметно возрос спрос на СПП и в нашей стране [1, 37,74].

СПП позволяют решить широкий спектр задач «разведочного» анализа данных, статистического исследования зависимостей, планирования экспериментов, анализа временных рядов, анализа данных нечисловой природы [19,22,29,33].

Традиционный состав эконометрического инструментария представлен стандартным набором математико-статистических методов [7,43,49]:

- классическая линейная модель множественной регрессии и классический метод наименьших квадратов;
- обобщенная линейная модель множественной регрессии и обобщенный метод наименьших квадратов;
- некоторые специальные модели регрессии (со стохастическими объясняющими переменными, с переменной структурой, с дискретными зависимыми переменными, нелинейные);
- модели и методы статистического анализа временных рядов;
- анализ систем одновременных эконометрических уравнений.
- качественный анализ поставленной проблемы;

- подготовка исходной информации;
- построение математической модели;
- математический анализ модели;
- численное решение с использованием программных продуктов;
- анализ полученных результатов и их интерпретация.

В монографии подробно разобран пример эконометрического исследования с использованием статистических методов, с использованием эконометрического инструментария. Пример основан на официальных статистических данных, собранных в ходе исследования и рассматривается как решение типовой задачи эконометрического анализа социально-экономических данных.

Любое эконометрическое исследование начинается с построения экономико-математической модели, адекватной изучаемому реальному социально – экономическому явлению или процессу и отражающей наиболее существенные с точки зрения исследователя особенности и закономерности развития изучаемого объекта. Полученная модель затем используется для анализа поведения исследуемого явления в различных ситуациях, выявления тенденций в его развитии, расчета его прогнозных количественных и качественных характеристик. Что в свою очередь используется для формулирования рекомендаций и принятия управленческих решений [40,41].

Правильно построенная модель должна давать ответ на вопрос о количественной оценке величины изменения изучаемого явления или процесса в зависимости от изменений внешней среды.

Использование методов экономической теории, математики и статистики (наук, лежащих в основе эконометрики) при моделировании исследуемого процесса осуществлялось в такой последовательности [23,32,43,47].

1) Процесс построения эконометрической модели начинается с качественного исследования проблемы методами экономической теории. На этой основе формулируются цели исследования, выделяются факторы, влияющие на изучаемый показатель, и формулируются предположения о характере предполагаемой зависимости, т. е. создается содержательная модель изучаемого явления или процесса.

2) Сформулированные предположения о характерных особенностях изучаемых зависимостей выражаются в виде математических формул и соотношений. Тем самым формируется математическая модель, отражающая качественную сторону зависимости и содержащая некоторые неопределенные построение модели, параметры.

3) Далее методами статистики осуществляется построение модели, отражающей также и количественную сторону изучаемой зависимости. Для этого производится сбор данных наблюдений о проявлениях изучаемого явления, отбор наиболее значимых факторов, определение наличия и степени тесноты связи между изучаемыми показателями, получение численных значений параметров модели и исследование степени соответствия полученной модели реальной действительности.

1 ЭКОНОМЕТРИЧЕСКОЕ МОДЕЛИРОВАНИЕ

1.1 Типы переменных и типы данных

В любой эконометрической модели, в зависимости от конечных прикладных целей ее использования все участвующие в ней переменные подразделяются на:

-экзогенные переменные, задаваемые как бы извне, автономно, в определенной степени управляемые (планируемые);

-эндогенные переменные, значения которых формируются в процессе и внутри функционирования анализируемой социально-экономической системы под воздействием экзогенных переменных и во взаимодействии друг с другом, являются предметом объяснения в эконометрической модели;

-предопределенные переменные выступают в роли факторов-аргументов или объясняющих переменных;

-лаговые эндогенные переменные входят в уравнения анализируемой эконометрической системы, но измерены в прошлые моменты, а следовательно, являются уже известными, заданными.

Эконометрическая модель служит для объяснения поведения эндогенных переменных в зависимости от значений экзогенных и лаговых эндогенных переменных.

Экономические величины складываются под влиянием множества различных факторов, как количественных, так и качественных по своей природе, поэтому при построении регрессионной модели часто возникает необходимость включить в исследование не только количественные, но и качественные

переменные (например, должность, образование, пол, место проживания, национальность и др.) и исследование их влияния на результирующую переменную.

Регрессионные модели могут содержать одновременно как количественные, так и качественные переменные, и даже только качественные [81,93].

Фиктивной переменной (*dummy variable*) называется атрибутивный или качественный фактор, представленный посредством определённого цифрового кода. Чтобы ввести качественные факторы в регрессионную модель, им должны быть присвоены те или иные цифровые метки, т.е. качественные переменные преобразованы в количественные.

При исследовании влияния качественных признаков в модель вводят фиктивные переменные, которые, как правило, принимают два значения – 0 и 1.

Если включаемый в рассмотрение качественный признак имеет не два, а несколько значений, используют несколько бинарных переменных. При этом нужно следить, чтобы эти переменные оставались линейно независимыми, иначе будет невозможно получить оценки коэффициентов регрессии методом наименьших квадратов. Во избежание линейной зависимости факторов число вводимых бинарных переменных должно быть на единицу меньше числа значений качественного признака.

Фиктивные переменные, несмотря на свою простоту, являются весьма гибким инструментом при исследовании влияния качественных признаков [14]. "Фиктивность" переменных

заключается только в том, что они количественным образом описывают качественные признаки.

Для изучения какого-либо явления необходимо организовать статистическое наблюдение, целью которого является - сбор данные. Сбор данных следует осуществлять в соответствии с определенными переменными.

Статистические данные — это конкретные численные значения статистических показателей, они могут быть пространственными, временными и панельными.

Для моделирования используются данные следующих трех типов.

1. Пространственные данные – данные, полученные в результате статистического наблюдения за несколькими единицами статистической совокупности на один и тот же момент времени или за один и тот же период времени. Такие данные имеют два измерения: признак – объект.

Пусть мы располагаем результатами регистрации значений переменных (X_1, X_2, \dots, X_p) на n статистически обследованных объектах. Так что если i - номер обследованного объекта, то имеющиеся исходные статистические данные состоят из n строк вида $(X_{1i}, X_{2i}, \dots, X_{pi})$, $i = 1, \dots, n$, где X_{ji} - значение j переменной, зарегистрированное на i обследованном объекте. То есть данные могут быть представлены в виде матрицы $n \times p$:

$$X = \begin{pmatrix} X_{11} & \dots & X_{1p} \\ \dots & \dots & \dots \\ X_{n1} & \dots & X_{np} \end{pmatrix}$$

Такой тип данных называется пространственной выборкой или данными поперечного среза (cross-section data).

2. Временные данные (динамические ряды) – данные, полученные в результате статистического наблюдения за одной единицей статистической совокупности за разные моменты или периоды времени. Временные данные имеют два измерения: признак – время.

Пусть данные регистрируются на одном и том же объекте, но в разные периоды времени. Тогда аналогом i будет номер периода времени, к которому привязаны соответствующие данные, а n будет общим числом периодов времени. Такие данные называются временной выборкой, или временными рядами данных (time series data), или данными продольного среза. Для таких данных существенен порядок следования значений переменных.

3. Панельные данные – данные, полученные в результате статистического наблюдения за несколькими единицами статистической совокупности за разные моменты или периоды времени. Панельные данные имеют три измерения: признак – объект – время. Панельные данные позволяют проводить совместный анализ пространственных выборок и анализ временных рядов.

Пусть отслеживается каждый из n объектов в течение T периодов времени. То есть, имеем последовательность матриц вида X , отнесенных к моментам времени $1, 2, \dots, T$:

$$X(t) = \begin{pmatrix} X_{11}(t) & \dots & X_{1p}(t) \\ \dots & \dots & \dots \\ X_{n1}(t) & \dots & X_{np}(t) \end{pmatrix}$$

Такие данные называются панельными, или пространственно-временной выборкой (panel data). Данные сочетают в себе свойства как временных рядов, так и данных поперечного сечения [1].

Подготовка данных заключается в следующем:

1. Каждому индивидууму (экспериментальной единице) присваивается идентификатор.

2. Наблюдения над индивидуумами записываются в матричной форме: строки – индивидуумы, столбцы – наблюдаемые для каждого индивидуума значения признаков.

3. Признаки индивидуума надо упорядочивать в близкие по смыслу группы.

4. Коды названий признаков должны быть такие, чтобы было легко понять, что же это за признак.

5. Отсутствующие значения признаков кодируются, как правило, пробелом или числом, которое является невозможным наблюдением для данного признака (например, рост - 9999).

6. Для измерений в шкале наименований и в порядковой шкале лучше присваивать цифры, а не буквы (например, пол: мужчина – 1, женщина – 0, а не м/ж).

7. Не рекомендуется переводить измерения в интервальной шкале в порядковую (например, возраст целесообразно указывать в годах, а не в шкале «1 – менее 14, 2 – от 14 до 21 и т.д.).

8. Целесообразно объединять две переменные в одну, если это не влечет потери информации. Например, переменная 1 - наличие в семье детей (0 - нет, 1 - есть) и переменная 2 – возраст старшего ребенка могут быть заменены второй (0 – отсутствие детей).

1.2 Основные этапы эконометрического моделирования

Построение эконометрической модели, являющейся основой эконометрического исследования, основывается на предположении о наличии реально существующей зависимости между признаками. От того, насколько хорошо полученная модель описывает изучаемые закономерности между экономическими явлениями или процессами, зависит степень достоверности результатов анализа и их применимости [46,66].

Весь процесс эконометрического моделирования разбивается на следующие основные этапы:

1 этап (постановочный) - определение конечных целей моделирования, набора участвующих в модели факторов и показателей, их роли;

2 этап (априорный) – предмодельный анализ экономической и социальной сущности изучаемого явления, формирование и формализация априорной информации и исходных допущений в частности относящихся к природе исходных данных и случайных остаточных составляющих в виде ряда гипотез;

3 этап (параметризация) - собственно моделирование, т.е. выбор общего вида модели, в том числе состава и формы входящих в неё связей между переменными;

4 этап (информационный) - сбор необходимой статистической информации, т.е. регистрация значений участвующих в модели факторов и показателей;

5 этап (идентификация модели) - статистический анализ модели и статистическое оценивание неизвестных параметров модели.

Непосредственно связан с проблемой идентифицируемости модели. Необходимо предложить и реализовать математически корректную процедуру оценивания неизвестных параметров модели по имеющимся исходным данным;

6 этап (верификация модели) - сопоставление реальных и модельных данных, проверка адекватности модели, оценка точности модельных данных [20].

Подводя итог, можно сделать вывод, что процедура построения эконометрической модели включает три основных этапа:

1. Спецификация модели, т. е. выбор класса моделей, наиболее подходящих для описания изучаемых явлений и процессов. Этот этап предполагает решение двух задач:

а) отбор существенных факторов для их последующего включения в модель;

б) выбор типа модели, т. е. выбор вида аналитической зависимости, связывающей включенные в модель переменные.

2. Оценка параметров модели, т. е. получение численных значений констант модели. При этом используется предварительно полученный массив данных наблюдений.

3. Проверка качества построенной модели и обоснование возможности ее дальнейшего использования.

Важной составляющей процесса построения эконометрической модели является отбор факторов, существенно влияющих на изучаемый показатель и подлежащих включению в разрабатываемую модель. Оптимальный набор факторов определяется на основе качественного и количественного анализа. Прежде всего, на этапе постановки задачи и содержательного экономического анализа

экономической модели отбираются факторы, влияние которых должно быть учтено при построении модели. В ряде случаев набор факторов определяется однозначно или с большой степенью уверенности.

В более сложных случаях целесообразность включения в модель каждого фактора проверяется с помощью формальных статистических методов.

К факторам, включаемым в модель, предъявляются следующие требования:

- 1) факторы не должны коррелировать между собой;
- 2) факторы должны существенно влиять на изменение результативного признака.

При определении «оптимального» набора факторов могут использоваться два подхода: метод включения; метод исключения.

Согласно методу включения сначала строится уравнение регрессии с одним наиболее влияющим фактором. Затем в него последовательно по одному вводятся остальные факторы и определяется пара факторов, совместно наиболее влияющих на результативную переменную. На следующем к первым двум добавляется еще по одному фактору и определяется наилучшая тройка факторов и т. д. В модель включают только значимые (существенные) факторы.

Согласно методу исключения сначала строится уравнение регрессии с полным набором факторов, из числа которых затем последовательно по одному исключаются незначимые факторы. На каждом шаге исключается только один фактор, так как после исключения какого-либо фактора другой фактор, бывший до этого

незначимым, может стать значимым. Процесс заканчивается, когда не остается факторов, которые следует исключить из модели.

Методы включения и исключения не гарантируют определение оптимального набора факторов, но в большинстве случаев дают результаты либо оптимальные, либо близкие к ним.

Не рекомендуется включать в модель очень большое число факторов, так как это может затруднить выявление качественных закономерностей, и возрастает опасность включения в модель несущественных случайных факторов.

Более того, для получения достаточно надежных оценок параметров желательно, чтобы количество наблюдений превышало количество определяемых параметров не менее чем в 6-7 раз.

После выделения совокупности рассматриваемых переменных следующим этапом является определение конкретного вида модели, наилучшим образом соответствующего изучаемому явлению.

По характеру связей факторов с переменной y модели подразделяются на линейные, у которых зависимая переменная является линейной функцией факторов, и нелинейные, к которым относят все остальные модели.

По свойствам своих параметров модели подразделяются на модели с постоянной и переменной структурой. В последнем случае область изменения факторов модели разделена на подобласти, в которых параметры модели имеют различные значения.

Особый вид моделей составляют системы взаимосвязанных эконометрических уравнений, включающие несколько уравнений вид. Каждому уравнению соответствует своя зависимая переменная y_i ,

которая в других уравнениях системы может выступать в качестве фактора.

Если на основе предварительного качественного анализа рассматриваемого явления не удастся однозначно выбрать наиболее подходящий тип модели, то рассматриваются несколько альтернативных моделей, среди которых в процессе исследования выбирается та, которая в наибольшей степени соответствует изучаемому явлению.

При решении проблемы выбора вида аналитической зависимости могут использоваться:

- выводы аналитических исследований о качественном характере зависимости (направление изменения переменных и его особенности),
- описание свойств (характерных особенностей) различных аналитических зависимостей,
- цели построения модели.

Выбор вида эконометрической модели основывается, прежде всего, на результатах предварительного качественного или содержательного анализа, проводимого методами экономической теории. По возможности характер предполагаемой зависимости обосновывается исходя из теоретически предположений о характере закономерности развития изучаемого явления или процесса.

Другой подход основан на анализе массива исходных данных, который позволяет выявить некоторые характеристики предполагаемых зависимостей и на этой основе сформулировать, как правило, несколько предположений о виде аналитической связи. Построенная модель используется для формулирования

предположений о характере закономерности в развитии изучаемого явления, которые проверяются в течение дальнейших исследований.

Рассмотрим проблемы построения регрессионных моделей [68]:

1. Последствия отсутствия в уравнении существенной независимой переменной. Если в уравнение регрессии не включена независимая переменная, оказывающая существенное влияние на результативный признак, то в общем случае это приводит к смещению оценок коэффициентов регрессии. Стандартные ошибки коэффициентов регрессии становятся некорректными, что приводит к неприменимости соответствующих t -тестов. Кроме того, возможно появление автокорреляции и гетероскедастичности остатков. Признаком отсутствия значимой переменной может служить несоответствие знаков коэффициентов теоретическим предположениям. Если нет возможности включить в уравнение регрессии такую переменную, то следует использовать замещающую переменную.

Последствия включения в модель несущественной независимой переменной. Если в уравнение регрессии включена несущественная независимая переменная, то в общем случае это не приводит к смещению оценок коэффициентов регрессии, но значения стандартных ошибок могут возрасти.

Последствия неправильной спецификации формы уравнения регрессии. Использование неверной формы уравнения регрессии приводит к смещенности и несостоятельности оценок параметров и низкому значению коэффициента детерминации R^2 . Возможно также появление автокорреляции и гетероскедастичности остатков.

1.3 Классификация моделей

Наиболее распространенными в эконометрическом моделировании являются следующие образующие четыре группы методы[2,72]:

- классическая линейная модель множественной регрессии (КЛММР) и классический метод наименьших квадратов (МНК);
- обобщенная КЛММР и обобщенный МНК;
- методы статистического анализа временных рядов;
- методы анализа систем одновременных эконометрических уравнений.

Применение этих методов делает возможным построение следующих типов эконометрических моделей:

1. Регрессионные модели с одним уравнением

В таких моделях зависимая (объясняемая) переменная Y представляется в виде функции:

$$Y = f(X, \beta) = f(X_1, \dots, X_k, \beta_1, \dots, \beta_k),$$

где X_1, X_2, \dots, X_k – независимые (объясняющие переменные);

$\beta_1, \beta_2, \dots, \beta_k$ – параметры.

В зависимости от вида функции модели делятся на линейные и нелинейные.

В линейной множественной регрессии параметры при x называются коэффициентами «чистой» регрессии. Они характеризуют среднее изменение результата с изменением соответствующего параметра на единицу при неизменном значении других факторов, закрепленных на среднем уровне.

2. Модели временных рядов

К этому классу относятся модели:

- тренда: $Y(t) = T(t) + \varepsilon_t$,

где t – время;

$T(t)$ - временной тренд заданного параметрического вида
(например, линейный $T(t) = (a + bt)$;

ε - случайная (стохастическая) компонента;

- сезонности: $Y(t) = S(t) + \varepsilon_t$,

где $S(t)$ - периодическая (сезонная) компонента;

ε_t - случайная (стохастическая) компонента;

- тренда и сезонности: $Y(t) = T(t) + S(t) + \varepsilon_t$ (аддитивная) или
 $Y(t) = T(t)S(t) + \varepsilon_t$, (мультипликативная),

где $T(t)$ - временной тренд заданного параметрического вида;

$S(t)$ - периодическая (сезонная) компонента;

ε_t - случайная (стохастическая) компонента.

3. Системы одновременных уравнений

Эти модели описываются системами уравнений. Системы могут состоять из тождеств и регрессионных уравнений, каждое из которых, кроме объясняющих переменных, может включать в себя объясняемые переменные из других уравнений системы. Системы одновременных уравнений требуют сложного математического аппарата и используются для моделей национальной экономики. Поэтому наиболее широко этот подход применяется в макроэкономических исследованиях, а также в исследованиях спроса и предложения.

Различают несколько видов систем уравнений.

1.4 Линейная парная регрессия

Большинство явлений и процессов в экономике находятся в постоянной взаимной и всеохватывающей объективной связи. Исследование зависимостей и взаимосвязей между объективно существующими явлениями и процессами играет большую роль в экономике [35]. Оно дает возможность глубже понять сложный механизм причинно – следственных отношений между явлениями. Для исследования интенсивности, вида и формы зависимостей широко применяется корреляционно – регрессионный анализ, который является методическим инструментарием при решении задач прогнозирования, планирования и анализа хозяйственной деятельности предприятий.

Методы и модели корреляционно - регрессионного анализа занимают центральное место в математическом аппарате эконометрики.

Исследование корреляционных связей называют корреляционным анализом, а исследование односторонних вероятностных зависимостей – регрессионным анализом [36,44].

Задачами корреляционного анализа являются выявление статистической связи между случайными переменными, оценка ее тесноты (силы); отбор факторов, оказывающих наиболее существенное влияние на результативный признак, на основании измерения тесноты связи между явлениями; обнаружение неизвестных причинных связей с помощью логически – профессиональных рассуждений, раскрывающих механизм связей.

Задачами регрессионного анализа являются установление формы зависимости между переменными (линейная или нелинейная, положительная или отрицательная и т.д.), оценка функции регрессии (в виде математического уравнения того или иного типа) и, оценка неизвестных значений (прогноз значений) зависимой переменной.

Рассмотрим некоторый экономический объект (процесс, явление, систему) и выделим только две переменные, характеризующие объект. Обозначим переменные буквами Y и $X=(X_1, X_2, \dots, X_k)$ – одну и набор случайных величин. Будем предполагать, что независимая (объясняющая) переменная X оказывает воздействие на значения переменной Y , которая, таким образом, является зависимой переменной, т.е. имеет место зависимость: $Y = f(X)$

При изучении взаимосвязи между переменными Y и X следует, прежде всего, установить тип зависимости (связи) (природу анализируемых переменных Y и X).

Различают два вида зависимостей между экономическими явлениями и процессами: функциональную и статистическую (стохастическую, вероятностную).

1. Функциональные зависимости имеют место при исследовании связей между неслучайными переменными. Если Y и X являются неслучайными переменными, при этом значения Y строго зависят только от соответствующих значений X и полностью ими определяются. Если каждому значению одной переменной соответствует вполне определенное значение другой, то в этом случае говорят о функциональной зависимости, тогда Y является некоторой функцией от переменной X и верна модель $Y = f(X)$.

2. В большинстве случаев между переменными величинами существуют зависимости, когда каждому значению одной переменной, соответствует не какое – то определенное, а *множество* возможных значений другой переменной, среди которых могут быть факторы, общие для двух случайных величин, что позволяет говорить о вероятностной или статистической зависимости.

Если каждому значению одной переменной соответствует определенное (условное) распределение другой переменной, то говорят о статистической (стохастической, вероятностной) зависимости. Она возникает в следующих случаях: 1) при исследовании связей между случайными переменными Y и X . Величины X и Y подвержены действию множества различных случайных факторов; 2) при исследовании связей между случайными и неслучайными переменными Y и X , измеряемыми с некоторой случайной ошибкой (в этом случае наблюдаются не сами переменные, а искаженные, случайные величины: $Y' = Y + \varepsilon_y$ и $X'_i = X_i + \varepsilon_{x_i}, i = 1, 2, \dots, k$). 3) при анализе влияния на случайную переменную Y неслучайных факторов $X = (X_1, X_2, \dots, X_k)$. Такая связь может быть вызвана двумя причинами: а) ошибками измерения показателя Y , по отношению к которым ошибки измерения факторов X_1, X_2, \dots, X_k пренебрежительно малы; б) если зависимая переменная подвержена влиянию ряда неконтролируемых или неучтенных факторов.

Статистическая зависимость проявляется только в массовом процессе, при большом числе единиц совокупности.

В *частном случае* статистическая зависимость проявляется в том, что при изменении одной из величин изменяется условное

математическое ожидание $M_x(Y)$ или $M_y(X)$ другой. В этом случае говорят о корреляции или корреляционной зависимости.

Если зависимость между двумя переменными такова, что каждому значению одной переменной соответствует определенное условное математическое ожидание (среднее значение) другой, то такая статистическая зависимость называется корреляционной.

Иначе, корреляционной зависимостью между двумя переменными называется функциональная зависимость между значениями одной из них и условным математическим ожиданием другой. Верна модель: $M_x(Y) = \varphi(x)$ или $M_y(X) = \psi(y)$, где $\varphi(x) \neq const$, $\psi(y) \neq const$.

Односторонняя зависимость случайной переменной Y от одной (или нескольких) неслучайной независимой переменной X называется регрессионной. Верна модель: $M_x(Y) = \varphi(x)$, где $\varphi(x) \neq const$. Пример: такая зависимость может возникнуть в случае когда при каждом фиксированном значении X соответствующие значения Y подвержены случайному разбросу за счет действия ряда неконтролируемых факторов. На самом деле, для одного X мы можем наблюдать разные значения Y : а) X – возраст индивидуума, Y – его зарплата; в) X – доход семьи, Y – расходы на питание.

При этом зависимую переменную Y называют также функцией отклика, объясняемой, выходной, результирующей, эндогенной переменной, результативным признаком, а независимую переменную X – объясняющей, входной, предсказывающей, предикторной, экзогенной переменной, фактором, регрессором, факторным признаком.

Уравнение $M_x(Y) = \varphi(x)$, где $\varphi(x) \neq \text{const}$, называется модельным уравнением регрессии Y по X (или просто уравнением регрессии).

Функция $\varphi(x)$ называется модельной функцией регрессии (или просто функцией регрессии Y по X), если она описывает изменение условного среднего значения результирующей переменной Y в зависимости от изменения значений объясняющей переменной X : $\varphi(x) = M(Y|X=x)$.

График функции $\varphi(x)$ называется модельной линией регрессии (или просто линией регрессии).

Для точного описания уравнения регрессии необходимо знать условный закон распределения зависимой переменной Y при условии, что переменная X примет значение x , т.е. $X = x$. В статистической практике такую информацию получить, как правило не удастся, т.к. обычно исследователь располагает лишь выборкой пар значений (x_t, y_t) ограниченного объема n . В этом случае речь может идти об **оценке** (приближенном выражении, аппроксимации) по выборке функции регрессии. Такой оценкой является выборочная линия (кривая) регрессии:

$$\hat{y} = \hat{\varphi}(x, b_0, b_1, \dots, b_p)$$

где \hat{y} - условная (групповая) средняя переменной Y при фиксированном значении $X = x$;

b_0, b_1, \dots, b_p - параметры кривой.

Уравнение $\hat{y} = \hat{\varphi}(x, b_0, b_1, \dots, b_p)$ называется выборочным уравнением регрессии.

Рассмотрим понятие линейной парной регрессии.

Пусть мы располагаем n парами выборочных наблюдений над двумя переменными X и Y : $(x_1, y_1), (x_2, y_2), \dots, (x_t, y_t)$, где $t = \overline{1, n}$.

В силу воздействия неучтенных случайных факторов и причин отдельные наблюдения переменной Y будут в большей или меньшей мере отклоняться от функции регрессии $\varphi(X)$. В этом случае уравнение взаимосвязи двух переменных (парная регрессионная модель) может быть представлена в виде:

$$Y = \varphi(X) + \varepsilon,$$

где ε - случайная переменная (возмущающая переменная), характеризующая отклонение от функции регрессии и называется случайным членом или возмущением, отклонением (или ошибкой).

Присутствие в модели случайной компоненты ε , обусловлено следующими причинами:

1. Ошибки спецификации. Среди них выделяют не включение важных объясняющих переменных, агрегирование (объединение) переменных, неправильную функциональную спецификацию модели.

2. Ошибки измерения. Они связаны со сложностью сбора исходных данных и использованием в модели аппроксимирующих переменных для учета факторов, непосредственное измерение которых невозможно.

3. Ошибки, связанные со случайностью человеческих реакций. Обусловлены тем, что поведение и непосредственное участие человека в ходе сбора и подготовки данных может быть достаточно непредсказуемым, и вносит, таким образом, свой вклад в случайный член.

Таким образом, в регрессионной модели зависимая переменная Y есть некоторая функция $\varphi(X)$ с точностью до случайного возмущения ε .

По расположению эмпирических точек можно предположить наличие линейной корреляционной (регрессионной) зависимости между переменными X и Y . В этом случае линейная парная регрессионная модель будет иметь вид: $y_t = \alpha + \beta x_t + \varepsilon_t, t = 1, \dots, n$.

Уравнение зависимости y_t от x_t вида:

$$y_t = \alpha + \beta x_t + \varepsilon_t \quad t = \overline{1, n}$$

где x_t - неслучайная (детерминированная) величина;

y_t - случайная величина;

ε_t - случайная величина,

называется линейным регрессионным уравнением с двумя переменными (или моделью линейной парной регрессии),

На основе выборочного наблюдения оценивается выборочное линейное уравнение регрессии:

$$\hat{y}_t = a + bx_t,$$

где (a, b) - оценки параметров (α, β) .

В силу того, что случайные отклонения ε_t по выборке не могут быть определены, то они заменяются отклонениями e_t значений y_t переменной Y от оцененной линии регрессии \hat{y}_t , т.е. как разность между фактическим и расчетным значениями зависимой переменной: $e_t = y_t - \hat{y}_t$ - выборочная оценка отклонения ε_t или остаток регрессии. Остатки e_t , как и ошибки ε_t , являются случайными величинами, однако они, в отличие от ошибок ε_t , наблюдаемы.

Воздействие неучтенных случайных факторов и ошибок наблюдений в модели линейной парной регрессии $y_t = \alpha + \beta x_t + \varepsilon_t$ определяется с помощью дисперсии возмущений (ошибок) или остаточной дисперсии σ^2 . Несмещенной оценкой этой дисперсии является выборочная остаточная дисперсия:

$$S^2 = \frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n-2} = \frac{\sum_{t=1}^n e_t^2}{n-2} = \hat{\sigma}^2, \quad ,$$

где $(n-2)$ - число степеней свободы;

\hat{y} - групповая средняя, найденная по уравнению регрессии;

$e_t = y_t - \hat{y}_t$ - выборочная оценка отклонения ε_t или остаток

регрессии.

Величина $S = \sqrt{S^2}$ называется стандартной ошибкой регрессии.

Рассмотрим задачу «наилучшей» аппроксимации набора наблюдений (x_t, y_t) , $t = \overline{1, n}$ линейным уравнением $\hat{y}_t = a + bx_t$.

Величина \hat{y}_t описывается как расчетное значение переменной Y , соответствующее x_t . Наблюдаемые значения y_t , не лежат в точности на линии регрессии, т.е. не совпадают с \hat{y}_t .

Определим остаток e_t , в t -м наблюдении как $e_t = y_t - \hat{y}_t$.

Согласно методу наименьших квадратов неизвестные параметры a и b выбираются таким образом, чтобы сумма квадратов отклонений эмпирических значений y_t от значений \hat{y}_t , найденных по уравнению регрессии $\hat{y}_t = a + bx_t$, была минимальной. Иначе, суть МНК заключается в минимизации суммы квадратов остатков:

$$Q = \sum_{t=1}^n e_t^2 = \sum_{t=1}^n (y_t - \hat{y}_t)^2 = \sum_{t=1}^n (y_t - a - bx_t)^2 \rightarrow \min$$

Здесь (x_t, y_t) – известные значения наблюдения (числа), (a, b) – неизвестные.

На основании необходимого условия существования экстремума функции двух переменных $Q = Q(a, b)$ приравняем к нулю ее частные производные, т.е.

$$\begin{cases} Q'_a = -2 \sum (y_t - a - bx_t) = 0 \\ Q'_b = -2 \sum (y_t - a - bx_t)x_t = 0 \end{cases} \text{ или } \begin{cases} \sum (y_t - a - bx_t) = 0 \\ \sum (y_t - a - bx_t)x_t = 0 \end{cases} \text{ или } \begin{cases} a \cdot n + b \sum x_t = \sum y_t \\ a \sum x_t + b \sum x_t^2 = \sum x_t y_t \end{cases}.$$

После преобразования получим следующую систему нормальных уравнений:

$$\begin{cases} a + b\bar{x} = \bar{y} \\ a\bar{x} + b\overline{x^2} = \overline{xy} \end{cases}$$

где соответствующие средние, определяются по формулам:

$$\bar{x} = \frac{\sum_{t=1}^n x_t}{n}; \quad \bar{y} = \frac{\sum_{t=1}^n y_t}{n}; \quad \overline{xy} = \frac{\sum_{t=1}^n x_t y_t}{n}; \quad \overline{x^2} = \frac{\sum_{t=1}^n x_t^2}{n}.$$

Решение системы:

$$b = \frac{n \sum x_t y_t - \sum x_t \sum y_t}{n \sum x_t^2 - (\sum x_t)^2}; \quad a = \frac{1}{n} \sum y_t - \frac{1}{n} \sum x_t \cdot b \quad \text{или} \quad \begin{cases} b = \frac{\text{cov}(x, y)}{\text{var}(x)} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{x^2 - (\bar{x})^2} \\ a = \bar{y} - b\bar{x} \end{cases}$$

где $\text{var}(x) = S_x^2 = \sigma_x^2$ – выборочная дисперсия переменной X ;

$\text{cov}(x, y)$ – выборочный корреляционный момент или выборочная ковариация.

Подставляя значение $a + b\bar{x} = \bar{y}$ из первого уравнения системы в уравнение регрессии $\hat{y} = a + bx$ получим $\hat{y} - \bar{y} = b(x - \bar{x})$. Из полученного уравнения регрессии следует, что линия регрессии проходит через точку (\bar{x}, \bar{y}) , т.е. $\bar{y} = a + b\bar{x}$.

Коэффициент b называется выборочным коэффициентом регрессии (или угловым коэффициентом регрессии или просто

коэффициентом регрессии) Y по X . Он показывает, на сколько единиц в среднем изменяется переменная Y при увеличении независимой переменной X на одну единицу.

Постоянная a дает прогнозируемое значение зависимой переменной при $x = 0$. Это может иметь смысл в зависимости от того, как далеко находится $x = 0$ от выборочных значений x .

Выборочные дисперсии величин y , \hat{y} , e :

$$\text{var}(y) = \frac{1}{n} \sum (y_t - \bar{y})^2 \text{ — дисперсия наблюдаемых значений } y;$$

$$\text{var}(\hat{y}) = \frac{1}{n} \sum (\hat{y}_t - \bar{y})^2 \text{ — дисперсия расчетных значений } y;$$

$$\text{var}(e) = \frac{1}{n} \sum (y_t - \hat{y}_t)^2 \text{ — дисперсия остатков.}$$

Для того чтобы регрессионный анализ, основанный на МНК, давал наилучшие из всех возможных результаты, должны выполняться определенные условия (условия Гаусса – Маркова) или предпосылки (гипотезы) регрессионного анализа.

Основные гипотезы [60,69]:

1. $y_t = \alpha + \beta x_t + \varepsilon_t \quad t = \overline{1, n}$ - спецификация модели.
2. x_t - неслучайная величина, y_t – случайная величина, ε_t - случайная величина.

Если условие о неслучайности объясняющей переменной не выполняется, то оценки коэффициентов регрессии оказываются смещенными и несостоятельными.

3. Математическое ожидание случайного отклонения ε_t в любом наблюдении равно нулю: $M(\varepsilon_t) = 0$.

(или математическое ожидание зависимой переменной y_t равно линейной функции регрессии: $M(y_t) = a + vx_t$)

Данное условие означает, что случайный член не должен иметь систематического смещения. Если постоянный член включен в уравнение регрессии, то это условие выполняется автоматически.

1. Дисперсия случайного члена ε_t (или зависимой переменной y_t) постоянна для всех наблюдений: $D(\varepsilon_t) = \sigma^2$ (или $D(y_t) = \sigma^2$) – условие гомоскедастичности возмущения (требование постоянства дисперсии регрессионных возмущений) или равноизменчивости возмущения (зависимой переменной).

Данное условие означает, что дисперсия случайного члена в каждом наблюдении имеет только одно значение.

Если условие гомоскедастичности не выполняется, то оценки коэффициентов регрессии будут неэффективными, хотя и несмещенными.

5. Случайные члены ε_i и ε_j (или переменные y_i и y_j) должны быть статистически независимы (некоррелированы) между собой, т.е. $M(\varepsilon_i, \varepsilon_j) = 0, i \neq j$. Иначе, предполагается отсутствие систематической связи между значениями случайного члена в любых двух наблюдениях, которые должны быть абсолютно независимы друг от друга, т.е. остатки ε_i и ε_j некоррелированы.

Данное условие указывает на некоррелированность случайных членов для разных членов. Это условие часто нарушается, когда данные являются временными рядами. В случае, когда это условие не выполняется, говорят об автокорреляции остатков.

Если условие независимости случайных членов не выполняется, то оценки коэффициентов регрессии будут неэффективными, хотя и несмещенными.

б. Случайный член ε_t (или зависимая переменная y_t) есть нормально распределенная величина: $\varepsilon_t \leftrightarrow N(0; \sigma^2)$.

В этом случае модель $y_t = \alpha + \beta x_t + \varepsilon_t$ называется классической нормальной линейной регрессионной моделью.

Предположение о нормальности распределения случайного члена необходимо для проверки значимости параметров регрессии и для их интервального оценивания.

Для получения уравнения регрессии достаточно предпосылок 1-4. Требование выполнения предпосылки 5 (т.е. рассмотрение нормальной регрессии) необходимо для оценки точности уравнения регрессии и его параметров.

Теорема Гаусса – Маркова [82]: если регрессионная модель $y_t = \alpha + \beta x_t + \varepsilon_t$ удовлетворяет предпосылкам 2-5 регрессионного анализа, то оценки a и b , сделанные с помощью МНК, имеют наименьшую дисперсию в классе всех линейных несмещенных оценок (являются наилучшими линейными несмещенными оценками), т.е. обладают следующими свойствами:

1) несмещенности: $E(a) = \alpha$, $E(b) = \beta$.

(Это означает отсутствие систематической ошибки в положении линии регрессии).

2) эффективности: имеют наименьшую дисперсию в классе всех линейных несмещенных оценок, равную:

$$D(a) = \frac{\overline{x^2} \sigma^2}{n \cdot \text{var}(x)}, \quad D(b) = \frac{\sigma^2}{n \cdot \text{var}(x)}; \quad \text{где } \text{var}(x) = \frac{1}{n} \sum_{t=1}^n (x_t - \bar{x})^2$$

3) состоятельности: $\lim_{n \rightarrow \infty} D(a) = 0$, $\lim_{n \rightarrow \infty} D(b) = 0$.

(Это означает, что при достаточно большом n оценки (a, b) близки к (α, β)).

2 МОДЕЛЬ МНОЖЕСТВЕННОЙ РЕГРЕССИИ

Экономические явления, как правило, определяются большим числом одновременно и совокупно действующих факторов. В связи с этим часто возникает задача исследования зависимости одной зависимой переменной Y от нескольких объясняющих переменных x_1, x_2, \dots, x_k . Эта задача решается с помощью множественного регрессионного анализа [90].

Пусть мы располагаем выборочными наблюдениями над k переменными Y_i , и X_{ji} , $j=1, \dots, k$, $i=1, 2, \dots, n$, где n - количество наблюдений (см. таблицу 2.1):

Таблица 2.1 – Таблица выборочных наблюдений

1	2	...	i	...	N
Y_1	Y_2	...	Y_i	...	Y_n
X_{11}	X_{12}	...	X_{1i}	...	X_{1n}
...
X_{k1}	X_{k2}	...	X_{ki}	...	X_{kn}

Множественная регрессия представляет собой регрессию результативного признака с двумя или более числом независимых переменных вида:

$$Y = f(x_1, x_2, \dots, x_k) + \varepsilon.$$

Уравнение зависимости y_t от $x_{t1}, x_{t2}, \dots, x_{tk}$ вида:

$$y_t = \beta_0 + \beta_1 x_{t1} + \dots + \beta_k x_{tk} + \varepsilon_t, t = 1, \dots, n \quad (2.1)$$

ИЛИ

$$y_t = \beta_1 x_{t1} + \beta_2 x_{t2} + \dots + \beta_k x_{tk} + \varepsilon_t, t = 1, \dots, n$$

где x_{tk} - значения регрессора x_k в наблюдении t , а $x_{t0} = 1, t = 1, \dots, n$;

y_t - t -ое наблюдение зависимой переменной y_t ;

β_j - неизвестные коэффициенты (параметры), $j = \overline{0, k}$;

ε_t - случайный член.

называется линейным регрессионным уравнением с k переменными или классической нормальной линейной моделью множественной регрессии (КЛНММР), если ε_t удовлетворяет предпосылкам регрессионного анализа.

Матричное описание регрессии облегчает как теоретический анализ, так и необходимые расчетные процедуры.

Введем обозначения:

$Y = (y_1 y_2 \dots y_n)'$ - матрица-столбец или вектор, значений зависимой переменной размера n ;

$\beta = (\beta_0 \beta_1 \dots \beta_k)'$ - $(k+1) \times 1$ вектор коэффициентов (параметров);

$\varepsilon = (\varepsilon_1 \dots \varepsilon_n)'$ - $n \times 1$ вектор ошибок (возмущений, остатков);

$X = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{pmatrix}$ - $n \times (k+1)$ матрица значений

объясняющих переменных.

Тогда в матричной форме модель (2.1) примет вид:

$$Y = X\beta + \varepsilon$$

Запишем предпосылки множественного регрессионного анализа в матричной форме:

1) $Y = X\beta + \varepsilon$ - спецификация модели.

2) ε - случайный вектор;

X – неслучайная (детерминированная) матрица, столбцы которой должны быть линейно независимые, т.е. ранг матрицы X – максимальный, $r(X) = k + 1$, т.е. ранг матрицы X равен числу ее столбцов.

3) а) $M(\varepsilon) = 0_n, \bar{b}) M(\varepsilon\varepsilon') = \sigma^2 E_n$, где $E_n = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}$ – $n \times n$ единичная

матрица.

с) ε - нормально распределенный случайный вектор, т.е. $\varepsilon \approx N_n(0; \sigma^2 E_n)$.

4) Случайное отклонение должно быть независимо от объясняющих переменных: $\sigma_{\varepsilon x} = 0$.

Модель (2.1), удовлетворяющая приведенным предпосылкам (1-4), кроме 3) о нормальном законе вектора возмущений ε , называется классической линейной моделью множественной регрессии (КЛММР).

На основе n наблюдений оценивается выборочное уравнение множественной регрессии:

$$\hat{y}_t = b_0 + b_1 x_{t1} + b_2 x_{t2} + \dots + b_k x_{tk} + e_t, t = \overline{1, n}$$

или

$$\hat{Y} = Xb + e,$$

где $b = (b_0 b_1 \dots b_k)'$ - оценки параметров $(\beta_0 \dots \beta_k)'$;

$e = (e_1 e_2 \dots e_n)$ - оценки возмущений $(\varepsilon_1 \dots \varepsilon_n)'$.

Воздействие неучтенных случайных факторов и ошибок наблюдений в модели множественной регрессии определяется с помощью дисперсии возмущений (случайных ошибок) или

остаточной дисперсии σ^2 . Несмещенной оценкой этой дисперсии является выборочная остаточная дисперсия:

$$S^2 = \frac{1}{n-k-1} \sum_{i=1}^n e_i^2,$$

где n - число наблюдений;

k - число независимых переменных.

Величина $S = \sqrt{S^2}$ называется стандартной ошибкой регрессии.

Регрессионный анализ является инструментом решения следующих основных задач:

1. Для любых значений объясняющих переменных X_i построить наилучшие по некоторому критерию оценки для неизвестной функции $f(X_i)$.
2. По заданным значениям объясняющих переменных X_i построить наилучший по некоторому критерию прогноз для неизвестного значения результирующей переменной Y .
3. Оценить удельный вес влияния переменных X_i на результирующий показатель Y .

2.1 Метод наименьших квадратов

При оценке параметров модели в качестве исходных данных используется заранее подготовленный массив наблюдений. Так как исходные данные содержат проявления случайных величин, то и полученные оценки являются случайными величинами, зависящими от исходных данных и метода оценивания. Отсюда возникает задача отбора методов оценивания параметров, дающих оценки более высокого качества [16,29].

Согласно теории статистического оценивания качество оценок определяется наличием у них таких свойств как, несмещенность, состоятельность и эффективность.

Оценка параметра называется *несмещенной*, если ее математическое ожидание равно оцениваемому параметру.

Оценка параметра называется *состоятельной*, если она сходится по вероятности к оцениваемому параметру при возрастании количества наблюдений.

Оценка параметра называется *эффективной*, если она имеет наименьшую дисперсию среди возможных несмещенных оценок параметра, вычисленных по выборкам одного и того же объема n .

Наиболее часто для оценки параметров применяются методы наименьших квадратов (МНК) и максимального правдоподобия. При выполнении определенных условий (относительно погрешностей модели ε_i) оценки параметров, полученные с помощью этих методов, обладают свойствами несмещенности, состоятельности и эффективности. Поэтому после получения оценок параметров необходимо проверить выполнение упомянутых условий, чтобы убедиться в качестве полученных оценок. Если эти условия не выполняются, то следует скорректировать модель соответствующим образом.

Причины нарушения условий, налагаемых МНК на погрешности модели ε_i , могут быть следующими: в модели не учтены существенные факторы; неправильно выбран вид модели.

Классический подход к оцениванию параметров уравнения множественной регрессии основан на методе наименьших квадратов

(МНК). В соответствии с МНК минимизируется сумма квадратов остатков:

$$S = \sum_{t=1}^n e_t^2 = \sum (y_t - \hat{y}_t)^2 \rightarrow \min$$

или

$$S = \sum_{t=1}^n e_t^2 = e'e = (Y - Xb)'(Y - Xb) \rightarrow \min ,$$

после раскрытия скобок получим:

$$S = Y'Y - 2b'X'Y + b'X'Xb \rightarrow \min .$$

Необходимым условием ее минимума является равенство нулю

вектора частных производных $\frac{\partial S}{\partial b} = \left(\frac{\partial S}{\partial b_0} \frac{\partial S}{\partial b_1} \dots \frac{\partial S}{\partial b_k} \right) = 0$, т.е.

$\frac{\partial S}{\partial b} = -2X'Y' + 2X'Xb = 0$, откуда получаем систему нормальных уравнений в матричной форме для определения вектора b :
 $X'Xb = X'Y$.

Решением этого уравнения является вектор b :

$$b = (X'X)^{-1} X'Y ,$$

где $(X'X)^{-1}$ - матрица, обратная матрице $(X' X)$ коэффициентов системы;

$X'Y$ - матрица-столбец, или вектор ее свободных членов.

Таким образом, оценки параметров модели множественной регрессии в матричной форме определяются выражением:

$$b = (X'X)^{-1} X'Y ,$$

а несмещенная оценка $S_{b_j}^2$ дисперсии $\sigma_{b_j}^2$ коэффициента регрессии b_j определится по формуле:

$$S_{b_j}^2 = S^2 (X'X)^{-1}_{jj} ,$$

где S^2 – несмещенная оценка параметра σ^2 (остаточная дисперсия);

$(X'X)^{-1}_{jj}$ - диагональный элемент матрицы $(X'X)^{-1}$.

Среднее квадратическое отклонение или стандартная ошибка коэффициента регрессии b_j примет вид: $S_{b_j} = S\sqrt{(X'X)^{-1}_{jj}}$,

Если предпосылки относительно случайного члена ε выполняются (кроме предпосылки о требовании нормальности закона распределения вектора возмущений ε), оценки параметров множественной регрессии являются несмещенными, состоятельными и эффективными.

Согласно теореме Гаусса–Маркова: При выполнении предпосылок множественного регрессионного анализа оценка метода наименьших квадратов $b = (X'X)^{-1}X'Y$ является наиболее эффективной, т.е. обладает наименьшей дисперсией в классе линейных несмещенных оценок.

При нелинейной зависимости признаков, приводимой к линейному виду, параметры множественной регрессии также определяются по МНК с той лишь разницей, что он используется не к исходной информации, а к преобразованным данным.

При этом делаются определенные предпосылки относительно составляющей ε , которая представляет собой ненаблюдаемую величину.

Исследования остатков ε_i - предполагают проверку наличия следующих пяти предпосылок МНК:

1. Случайный характер остатков.

Для проверки случайного характера остатков ε_i , строится график зависимости остатков ε_i от теоретических значений результативного признака.

Если на графике (см. рисунок 2.1) получена горизонтальная полоса, то остатки ε_i , представляют собой случайные величины, теоретические значения Y_x хорошо аппроксимируют фактические значения Y .

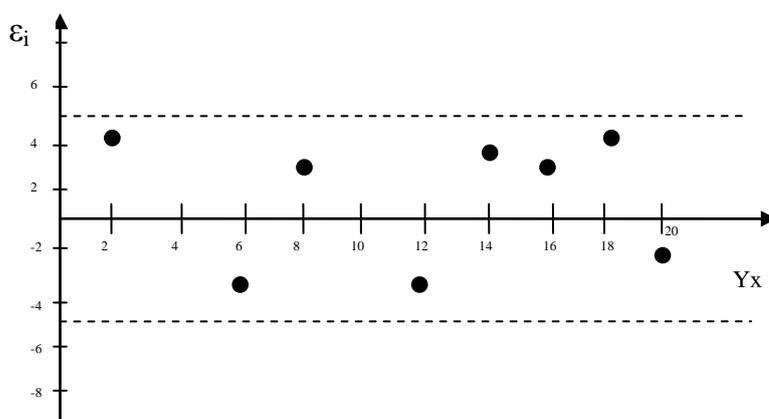


Рисунок 2.1 – Случайный характер остатков ε_i

Возможны следующие случаи: если ε_i зависит от Y_x , то:

- остатки не случайны (рис. 2.2, а);
- остатки носят систематический характер (рис. 2.2, б);
- остатки не имеют постоянной дисперсии (рис. 2.2, в);

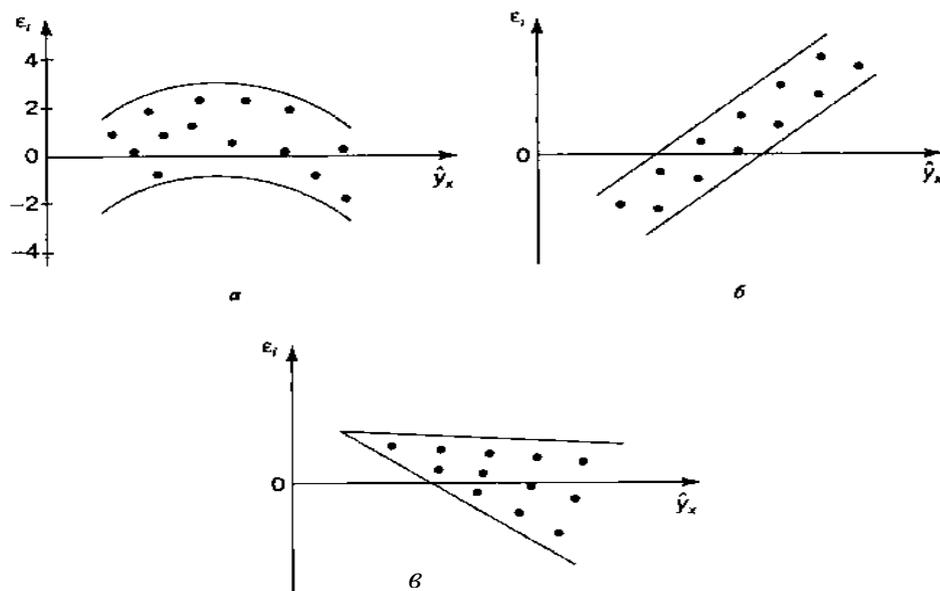


Рисунок 2.2 – Графики зависимости остатков ε_i

В других случаях применяют другую функцию или вводят дополнительную информацию и заново строят уравнение регрессии до тех пор, пока остатки ε_i , не будут случайными величинами.

2. Математическое ожидание случайного члена ε_i в любом отдельном наблюдении равно нулю: $M(\varepsilon_i) = 0$. Или нулевая средняя величина остатков, не зависящая от X_i , то есть $\sum_{i=1}^n (Y_i - Y_{x_i}) = 0$,

Проверка равенства математического ожидания уровней ряда остатков нулю осуществляется с помощью t -критерия Стьюдента. Для этого рассчитывается:

$$t_{расч} = \frac{\frac{|\sum_{i=1}^n \varepsilon_i|}{n}}{\sqrt{\frac{\sum_{i=1}^n (\varepsilon_i - \bar{\varepsilon})^2}{n-k}}}$$

Если $t_{расч} < t_{кр}$, то гипотеза о равенстве математического ожидания уровней ряда остатков принимается.

3. Дисперсия случайного члена ε_i постоянна для всех наблюдений.

Гомоскедастичность – дисперсия каждого отклонения ε_i , одинакова для всех значений X , то есть для каждого значения фактора X_i остатки ε_i , имеют одинаковую дисперсию. Если это условие применения МНК не соблюдается, то имеет место гетероскедастичность.

4. Значения случайного члена в любых наблюдениях ε_i и ε_j не коррелируют между собой: $Cov(\varepsilon_i, \varepsilon_j) = 0$ ($i \neq j$). Отсутствие автокорреляции остатков, т. е. значения остатков ε_i распределены независимо друг от друга. Автокорреляция остатков означает наличие корреляции между остатками текущих и предыдущих (последующих) наблюдений. Отсутствие автокорреляции остаточных величин обеспечивает состоятельность и эффективность оценок коэффициентов регрессии.

5. Значения случайного члена ε_i распределены по нормальному закону распределению.

Соответствие ряда остатков нормальному закону распределения определяется при помощи RS – критерия:

$$RS = \frac{\varepsilon_{\max} - \varepsilon_{\min}}{\sqrt{\frac{\sum_{i=1}^n \varepsilon_i^2}{n - k}}}$$

где ε_{\max} , ε_{\min} – максимальное и минимальное значения остаточного ряда.

Вычисленное значение сравнивают с табличными верхней и нижней границами.

Если значение этого критерия попадает между табулированными границами с заданным уровнем значимости, то гипотеза о нормальном законе распределения ряда остатков принимается.

Предположим, что все предпосылки для множественной регрессии выполняются.

2.2 Корреляция в КЛММР

Для трактовки линейной связи между двумя переменными акцентируют внимание на коэффициенте корреляции [51].

$$r_{yx} = \frac{\sum_{i=1}^n (Y_i - \bar{Y})(X_i - \bar{X})}{\sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2} \sqrt{\sum_{i=1}^n (X_i - \bar{X})^2}} \quad (2.5)$$

Оценку качества построенной модели дает коэффициент детерминации:

$$R^2 = 1 - \frac{\sum_{i=1}^n \varepsilon_i^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2}. \quad (2.6)$$

Согласно формуле (2.6) значение коэффициента детерминации не может быть больше единицы, причем это максимальное значение

будет достигнуто при $\sum_{i=1}^n \varepsilon_i^2 = 0$. Следовательно, значения коэффициента корреляции лежат в числовом промежутке от -1 до +1. Чем лучше регрессия соответствует наблюдениям, тем меньше $\sum_{i=1}^n \varepsilon_i^2$ и тем ближе R^2 к 1, и наоборот, чем "хуже" подгонка линии регрессии к данным, тем ближе значение R^2 к 0.

Проверка гипотезы о значимости коэффициента корреляции равносильна проверке основной гипотезы об отсутствии линейной связи между Y и X . Вычисляя значение t -статистики

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}. \quad (2.7)$$

вывод о значимости r делается при $|t| > t_\alpha$, где t_α - соответствующее табличное значение t -распределения с $(n-2)$ степенями свободы и уровнем значимости α .

Для того чтобы получить адекватную оценку того, насколько хорошо вариация результирующего признака объясняется вариацией нескольких факторных признаков, применяют скорректированный коэффициент детерминации:

$$R_{\text{скорр}}^2 = 1 - (1 - R^2) \frac{n-1}{n-k-1} \quad (2.8)$$

Скорректированный коэффициент детерминации всегда меньше R^2 . Кроме того, в отличие от R^2 , который всегда положителен, $R_{\text{скорр}}$ может принимать и отрицательное значение [19].

2.3 Оценка качества модели множественной регрессии

Проверка общего качества уравнения осуществляется как оценка того, как хорошо эмпирическое уравнение регрессии согласуется со статистическими данными [93,94,30]. Другими словами, насколько широко рассеяны точки наблюдений относительно линии регрессии.

Проверка качества модели множественной регрессии в целом осуществляется с помощью F -критерия Фишера. Для проверки гипотезы о том, что линейная связь между X_1, X_2, \dots, X_k и Y отсутствует:

$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$, пользуются соотношением:

$$F = \frac{R^2}{k} : \frac{1 - R^2}{n - (k + 1)}, \quad (2.9)$$

которое удовлетворяет F - распределению Фишера с $(k, n - (k + 1))$ степенями свободы. Критические значения этой статистики F_α для уровня значимости α затабулированы. Если $F > F_\alpha$, то гипотеза об отсутствии связи между переменными X_1, X_2, \dots, X_k и Y отклоняется, в противном случае гипотеза H_0 принимается и уравнение регрессии не значимо.

Для определения точности модели можно использовать среднюю относительную ошибку аппроксимации:

$$\bar{E}_{\text{ошн}} = \frac{1}{n} \sum_{i=1}^n \frac{|\varepsilon_i|}{Y_i} \cdot 100\%$$

Для модели высокой точности $\bar{E}_{отн} < 10\%$, хорошей – $10\% < \bar{E}_{отн} < 20\%$. удовлетворительной – $20\% < \bar{E}_{отн} < 50\%$, неудовлетворительной – $\bar{E}_{отн} > 50\%$.

Помимо проверки значимости уравнения в целом, можно проверить статистическую значимость каждого из коэффициентов регрессии в отдельности.

Статистическая значимость каждого из коэффициентов регрессии определяется при помощи t -критерия Стьюдента. Решение о том, что коэффициент значим, принимается в случае, когда $|t| > t_\alpha$ иначе принимается альтернативная гипотеза.

Значение t -статистики Стьюдента в случае множественной регрессии определяется по формуле:

$$t_{\beta_i} = \frac{\hat{\beta}_i}{S_{\hat{\beta}_i}} \quad (2.10)$$

где $S_{\hat{\beta}_i}$ - стандартная ошибка коэффициента регрессии β_i , которая определяется по формуле (2.4).

Доверительный интервал для истинного значения коэффициента β_i в этом случае есть

$$[\hat{\beta}_i - t_c S_{\hat{\beta}_i}; \hat{\beta}_i + t_c S_{\hat{\beta}_i}]. \quad (2.11)$$

где t_c - двусторонняя $(1-\alpha)$ -квантиль распределения Стьюдента с $n - k$ степенями свободы,

α - уровень значимости t -статистики Стьюдента.

Коэффициенты регрессионной модели играют важную роль при оценке влияния факторов на результирующий признак. Однако непосредственно с их помощью нельзя сопоставить факторы по степени их влияния на результирующий признак из-за различия единиц измерения и разной степени колеблемости. Для устранения таких различий применяются коэффициенты эластичности и бета - коэффициенты.

На практике часто бывает необходимо сравнение влияния на зависимую переменную различных объясняющих переменных, когда последние выражаются разными единицами измерения. В этом случае используют стандартизированные коэффициенты регрессии (бета - коэффициенты β_j) и средние коэффициенты эластичности $\bar{\epsilon}_j (j = \overline{1, n})$.

Коэффициент β учитывает влияние анализируемых факторов на функцию y с учетом различий в уровне их колеблемости и рассчитывается по формуле:

$$\beta_j = b_j \cdot \frac{S_{x_j}}{S_y},$$

где β_j – коэффициент β j – го фактора;

S_{x_j} - среднее квадратическое отклонение j – го фактора;

$$S^2_{x_j} = \frac{1}{n-1} \sum_{t=1}^n (x_t - \bar{x}_t)^2 ;$$

S_y - среднее квадратическое отклонение функции;

$$S^2_y = \frac{1}{n-1} \sum_{t=1}^n (y_t - \bar{y}_t)^2 ;$$

b_j - коэффициент регрессии при j – м факторе.

Стандартизированный коэффициент регрессии β_j показывает, на сколько величин S_y изменится в среднем зависимая переменная Y

при увеличении только j -й объясняющей переменной x_j на величину своего S_{x_j} .

Из двух факторов x_i и x_j может быть исключен тот фактор, который имеет меньшее значение β – коэффициента.

Допустим, исключению подлежит один из мультиколлинеарных факторов x_i или x_j . Оба фактора управляемы на уровне предприятия, коэффициенты регрессии b_i и b_j статистически значимы. Фактор x_i более тесно связан с y , т.е. $r_{x_i,y} > r_{x_j,y}$, но при этом $\beta_{x_i} < \beta_{x_j}$. В этом случае обычно исключению подлежит фактор x_j .

Средние коэффициенты эластичности рассчитывается по формуле:

$$\bar{\varepsilon}_{yx_j} = b_j \frac{\bar{x}_j}{\bar{y}}$$

Коэффициент эластичности показывает, на сколько процентов (от средней) изменится в среднем функция Y при увеличении только аргумента x_j на 1%.

Указанные коэффициенты позволяют упорядочить факторы по степени влияния факторов на зависимую переменную.

2.4 Мультиколлинеарность и методы ее устранения

Мультиколлинеарность - это линейная взаимосвязь двух или нескольких объясняющих переменных.

Под мультиколлинеарностью понимается высокая взаимная коррелированность объясняющих переменных. Следствием мультиколлинеарности является то, что при отборе переменных-факторов нарушается предположение об отсутствии строгой

линейной зависимости между переменными X_1, X_2, \dots, X_k . В этом случае говорят, что объясняющие переменные $X_j \quad j = 1, 2, \dots, k$ модели характеризуются свойством полной (строгой) мультиколлинеарности. Строгая мультиколлинеарность встречается редко, так как ее несложно избежать на предварительной стадии отбора объясняющих переменных.

Реальная (частичная) мультиколлинеарность возникает в случаях достаточно сильных линейных статистических связей между переменными множественной регрессии [9,91].

Последствия мультиколлинеарности: увеличиваются стандартные ошибки оценок; уменьшаются t-статистики МНК-оценок регрессии; МНК-оценки чувствительны к изменениям данных; возможность неверного знака МНК-оценок; трудность в определении вклада независимых переменных в дисперсию зависимой переменной.

В любом случае мультиколлинеарность затрудняет разделение влияния объясняющих факторов на поведение зависимой переменной и делает оценки коэффициентов регрессии ненадежными. Данный вывод наглядно подтвердим с помощью диаграммы Венна (рисунок 2.3):

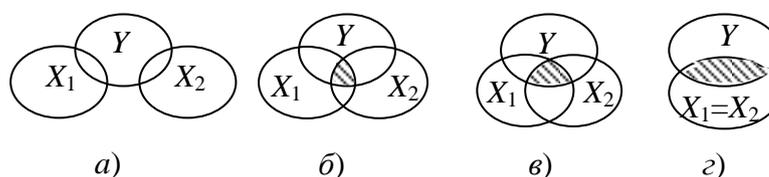


Рисунок 2.3 – Диаграммы Венна

На рисунке 2.3а) коррелированность между объясняющими переменными X_1 и X_2 отсутствует и влияние каждой из них на Y находит отражение в наложении кругов X_1 и X_2 на круг Y . По мере усиления линейной зависимости между X_1 и X_2 соответствующие круги все больше накладываются друг на друга. Заштрихованная область отражает совпадающие части влияния X_1 и X_2 на Y . На рисунке 2.3г) при совершенной мультиколлинеарности невозможно разграничить степени индивидуального влияния объясняющих переменных X_1 и X_2 на зависимую переменную Y .

Точных количественных критериев для проверки наличия мультиколлинеарности не существует, но имеются некоторые практические рекомендации по выявлению мультиколлинеарности:

1. Если среди парных коэффициентов корреляции между объясняющими переменными имеются значения 0,75-0,80 и выше, это свидетельствует о присутствии мультиколлинеарности.

2. Некоторые внешние признаки построенной модели, являющиеся следствиями мультиколлинеарности:

- некоторые из оценок $\beta_j, j = 1, 2, \dots, k$ имеют неправильные с теоретической точки зрения знаки или неоправданно большие по абсолютной величине значения,

- небольшое изменение исходной выборки (добавление или изъятие малой порции данных) приводит к существенному изменению оценок коэффициентов модели вплоть до изменения их знаков,

- большинство оценок коэффициентов регрессии оказываются статистически незначимо отличающимися от нуля, в то

время как в действительности многие из них имеют отличные от нуля значения, а модель в целом является значимой при проверке с помощью F -критерия.

Основными методами устранения мультиколлинеарности являются:

1. Удаление из модели одного или нескольких факторов.
2. Преобразование факторов, при котором уменьшается корреляция между ними.
3. Использование в уравнении регрессии взаимодействия факторов, например, в виде их произведения.
4. Использование так называемой ридж-регрессии (гребневой регрессии).

Это делает получаемые оценки смещенными, но уменьшает средние квадраты ошибок коэффициентов.

5. Использование метода главных компонент.
6. Отбор наиболее существенных объясняющих переменных на основе методов исключения, включения, шаговой регрессии, которые используют для принятия решения F -критерий, и т. д.[18].

Процедура отбора главных факторов обязательно включает следующие этапы:

1. Анализ факторов на мультиколлинеарность и её исключение.

Еще одним из возможных методов устранения или уменьшения мультиколлинеарности является использование пошаговых процедур отбора наиболее информативных переменных. Например, на первом шаге рассматривается лишь одна объясняющая переменная, имеющая с зависимой переменной наибольший коэффициент детерминации; на

втором шаге включается в регрессию новая объясняющая переменная, которая вместе с первоначально отобранной, образует пару объясняющих переменных, имеющую с зависимой переменной наиболее высокий (скорректированный) коэффициент детерминации; на третьем шаге вводится в регрессию еще одна объясняющая переменная, которая вместе с двумя первоначально отобранными образует тройку объясняющих переменных, имеющую с зависимой переменной наибольший (скорректированный) коэффициент детерминации и т.д.

Процедура введения новых переменных продолжается до тех пор, пока будет увеличиваться соответствующий (скорректированный) коэффициент детерминации.

Кроме рассмотренной пошаговой процедуры присоединения объясняющих переменных используется также пошаговые процедуры присоединения – удаления и процедура удаления объясняющих переменных. В большинстве случаев получаемые с помощью пошаговых процедур наборы переменных оказываются оптимальными или близкими к оптимальным.

2. Анализ тесноты взаимосвязи факторов (x) с зависимой переменной (y).

Для анализа тесноты взаимосвязи x и y используются значения коэффициента парной корреляции между фактором и функцией ($r_{x,y}$).

Факторы, для которых $r_{x,y} = 0$, т.е. не связанные с y , подлежат исключению в первую очередь. Факторы, имеющие наименьшее значение $r_{x,y}$, могут быть потенциально исключены из модели.

Вопрос об их окончательном исключении решается в ходе анализа других статистических характеристик.

В модель регрессии включаются те факторы, которые более сильно связаны с зависимой переменной, но слабо связаны с другими факторами.

3. Анализ β -коэффициентов факторов, которые потенциально могут быть исключены.

Коэффициент β учитывает влияние анализируемых факторов на зависимую переменную y с учетом различия в уровне их колеблемости.

Коэффициент β показывает, на сколько σ (СКО) изменяется функция с изменением соответствующего аргумента на одну сигму при фиксированном значении остальных аргументов:

$$\beta_k = \beta_i \frac{\sigma_{x_k}}{\sigma_y}, \quad i = \overline{0, m}, \quad k = \overline{1, m},$$

где β_k - коэффициент β количественного фактора;

σ_{x_k} - СКО количественного фактора;

σ_y - СКО функции;

β_i - коэффициент регрессии при количественном факторе.

Из двух факторов x_i и x_j может быть исключен тот фактор, который имеет меньшее значение β .

Например, исключению подлежит один из мультиколлинеарных факторов x_i или x_j . Оба фактора управляемы на уровне предприятия, коэффициенты регрессии β_i и β_j статистически значимы. Фактор x_i более тесно связан с y , т.е. $r_{x_i y} > r_{x_j y}$, но при этом $\beta_{x_i} < \beta_{x_j}$. В этом случае обычно исключению подлежит фактор x_j .

4. Проверка коэффициентов регрессии на статистическую значимость (по критерию Стьюдента или по критерию Фишера).

5. Анализ факторов на управляемость.

В ходе логического анализа на основе экономических знаний исследователь должен сделать вывод: можно ли разработать организационно – технические мероприятия, направленные на улучшение (изменение) выбранных факторов на уровне предприятия. Если это возможно, то данные факторы управляемы. Неуправляемые факторы на уровне предприятия могут быть исключены из модели.

6. Построение новой регрессионной модели без исключенных факторов.

Для этой модели определяется коэффициент множественной (скорректированный) детерминации \bar{R}^2 .

7. Исследование целесообразности исключения факторов из модели с помощью коэффициента детерминации.

Прежде, чем вынести решение об исключении переменных из анализа в силу их незначимости влияния на зависимую переменную, производят исследования с помощью коэффициента детерминации. Проверяется гипотеза об одновременном равенстве нулю части коэффициентов регрессии.

В первой регрессии содержится m объясняющих переменных, во второй – только часть из них, а именно m_1 объясняющих переменных. При этом $m = m_1 + m_2$, т.е. во вторую регрессию не включили m_2 объясняющих переменных.

Теперь следует проверить, вносят ли совместно эти m_2 переменных существенную долю в объяснение вариации переменной y . Для этого используется статистика:

$$F = \frac{(R_m^2 - R_{m_1}^2) \cdot (n - m - 1)}{(m - m_1) \cdot (1 - R_m^2)},$$

которая имеет F -распределение с $f_1 = m - m_1 = m_2$ и $f_2 = n - m - 1$ степенями свободы.

R_m^2 - коэффициент детерминации регрессии с m объясняющими переменными. $R_{m_1}^2$ - коэффициент детерминации регрессии с m_1 факторами. Разность $(R_m^2 - R_{m_1}^2)$ является мерой дополнительного объяснения вариации переменной y за счет включения m_2 переменных.

Если $F \leq F_{f_1, f_2, \alpha}$, то включение дополнительно объясняющих переменных совместно не оказывает значимого влияния на зависимую переменную y . Если $F > F_{f_1, f_2, \alpha}$, то m_2 объясняющих переменных совместно оказывают существенное влияние на вариацию переменной y , и, следовательно, в этом случае все m_2 переменные нельзя исключать из модели. При реализации первой ситуации ($F \leq F_{f_1, f_2, \alpha}$) факторы окончательно исключаются из модели.

2.5 Линейная модель множественной регрессии с гетероскедастичными остатками

При построении регрессионных моделей достаточно часто возникает ситуация, связанная с неоднородностью выборочных

совокупностей значений статистических признаков (значения объясняющих признаков характеризуют объекты разных масштабов).

Если остатки имеют постоянную дисперсию, они называются гомоскедастичными, но если они непостоянны, то гетероскедастичными (или имеет место гетероскедастичность).

В этой ситуации предположение о постоянстве дисперсии случайной ошибки оказывается не соответствующим действительности. В случаях, когда дисперсия одинакова в каждый момент времени или для каждого значения X , существуют определенные ограничения (в некоторой полосе) для расположения точек на графике X и Y , согласно которым отчетливой тенденции к увеличению или уменьшению дисперсии σ_{ε}^2 по мере роста X не наблюдается.

Ситуации, представленные на рисунке 2.3(а–в) отражают большую вероятность наличия гетероскедастичности случайной ошибки регрессии для рассматриваемых статистических данных.

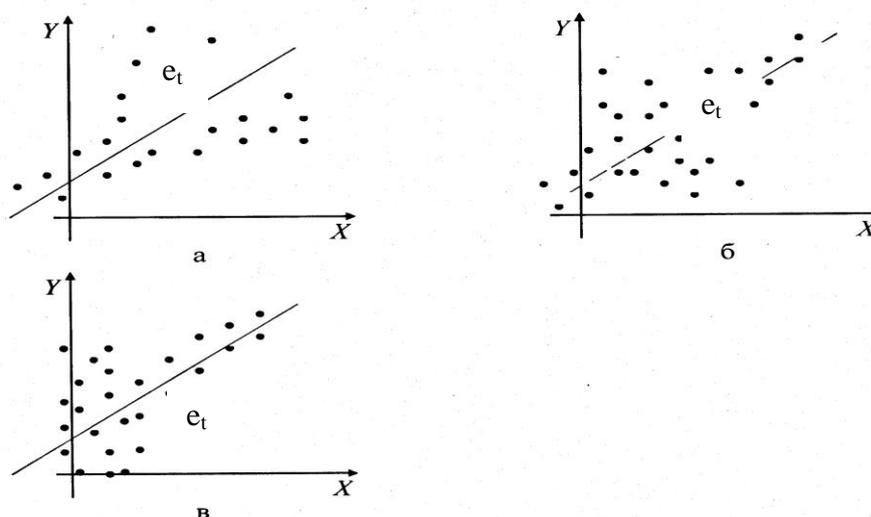


Рисунок 2.3 – Графическое представление отклонений (наличие гетероскедастичности)

На рисунке 2.3 а изображена ситуация, когда значения дисперсии σ_ε^2 растут по мере увеличения значений регрессора X . На рисунке 2.3 б дисперсия ошибки достигает максимальной величины при средних значениях X , уменьшаясь по мере приближения к крайним значениям. На рисунке 2.3 в дисперсия ошибки оказывается наибольшей при малых значениях X , быстро уменьшается и становится однородной по мере увеличения независимой переменной X .

Последствиями гетероскедастичности являются следующие.

Оценки коэффициентов останутся несмещенными и линейными, но не будут эффективными, то есть не будут иметь наименьшую дисперсию по сравнению с другими оценками данного параметра. Увеличение дисперсии оценок снижает вероятность получения максимально точных оценок. Дисперсии оценок будут рассчитываться со смещением. Эта смещенность появляется вследствие того, что не объясненная уравнением регрессии дисперсия, которая используется при вычислении оценок дисперсий всех коэффициентов не является несмещенной. Все выводы, получаемые на основе t - и F - статистик, а также интервальные оценки будут ненадежными.

Следовательно, статистические выводы, получаемые при стандартных проверках качества оценок могут быть ошибочными и приводить к неверным заключениям по построенной модели. Вероятно, что стандартные ошибки коэффициентов будут занижены, а, следовательно, t - статистики будут завышены. Это может привести к признанию статистически значимыми коэффициентов, которые на самом деле ими не являются.

Для обнаружения гетероскедастичности применяются различные методы: тест Гольдфельда-Квандта, тест Глейзера, тест ранговой корреляции Спирмэна, тест Уайта, тест Бреуша-Пагана и др.

Тест Гольдфельда-Квандт применяется в предположении, что средние квадратические отклонения случайного члена σ_i пропорциональны значениям фактора x_i и случайный член распределен по нормальному закону. Процедура применения теста Гольдфельда-Квандта состоит из следующих шагов:

1. Все n наблюдений упорядочиваются по величине X или по убыванию той независимой переменной, относительно которой есть предположение о гетероскедастичности.

2. Вся упорядоченная выборка после этого разбивается на 3 подвыборки размерности k , $n - 2k$, k .

3. Оцениваются отдельные регрессии для первой подвыборки и для третьей подвыборки. Если предположение о пропорциональности дисперсий отклонений значений X верно, то дисперсия регрессии по первой подвыборке (сумма квадратов отклонений $S_1 = \sum_{t=1}^k \varepsilon_t^2$) будет существенно меньше дисперсии

регрессии по третьей подвыборке ($S_3 = \sum_{t=n-k+1}^n \varepsilon_t^2$).

4. Для сравнения соответствующих дисперсий строится следующая F – статистика:

$$F = \frac{S_3 / (n - k - 1)}{S_1 / (n - k - 1)} = \frac{S_3}{S_1}$$

где $(n - k - 1)$ число степеней свободы соответствующих выборочных дисперсий;

k – количество объясняющих переменных в уравнении регрессии.

При сделанных предположениях относительно случайных отклонений построенная F – статистика имеет распределение Фишера с числами степеней свободы $\nu_1 = \nu_2 = n - k - 1$

5. Если $F_{набл} = \frac{S_3}{S_1} > F_{кр} = F_{\alpha, \nu_1, \nu_2}$, то гипотеза об отсутствии

гетероскедастичности отклоняется. Для множественной регрессии этот тест проводится для той объясняющей переменной, которая в наибольшей степени связана с Y при этом число n должно быть больше, чем $(k + 1)$.

Рассмотрим методы смягчения (коррекции) проблемы гетероскедастичности.

Гетероскедастичность приводит к неэффективности оценок, несмотря на их несмещенность. Это может обусловить необоснованные выводы по качеству модели. Поэтому при установлении гетероскедастичности возникает необходимость преобразования модели с целью устранения данного недостатка. Вид преобразования зависит от того, известны или нет дисперсии σ_t^2 отклонений $\varepsilon_t, t = 1, 2, \dots, n$.

1. Метод взвешенных наименьших квадратов - (ВНК).

Этот метод применяется при известных для каждого наблюдения значениях σ_t^2 . В этом случае можно устранить гетероскедастичность, разделив каждое наблюдаемое значение на соответствующее ему значение дисперсии σ_t^2

Опишем метод ВНК на примере парной регрессии: $y_t = \beta_0 + \beta_1 x_t + E_t$

Разделим обе части равенства на известное $\sigma_t = \sqrt{\sigma_t^2}$:

$$\frac{y_t}{\sigma_t} = \frac{1}{\sigma_t} \beta_0 + \frac{x_t}{\sigma_t} \beta_1 + \frac{\varepsilon_t}{\sigma_t}$$

Обозначим:

$$\begin{aligned} \frac{y_t}{\sigma_t} &= y_t^*; & \frac{\varepsilon_t}{\sigma_t} &= V_t \\ \frac{x_t}{\sigma_t} &= x_t^*; & \frac{1}{\sigma_t} &= Z_t \end{aligned}$$

Получим уравнение регрессии без свободного члена, но с дополнительной объясняющей переменной Z и с «преобразованным» отклонением V : $y_t^* = \beta_0 Z_t + \beta_1 x_t^* + V_t$

При этом для V_t выполняются условие гомоскедастичности.

$$\text{Действительно, } \sigma_t^2(V_t) = M(V_t - M(V_t))^2 = M(V_t^2) + M^2(V_t)$$

Так как по предпосылке 1 МНК:

$$M(\varepsilon_t) = 0, \text{ то } M(V_t) = M\left(\frac{\varepsilon_t}{\sigma_t}\right) = \frac{1}{\sigma_t} M(\varepsilon_t) = 0 \text{ и тогда:}$$

$$\sigma_t^2(V_t) = M(V_t^2) = M\left(\frac{\varepsilon_t^2}{\sigma_t^2}\right) = \frac{1}{\sigma_t^2} M(\varepsilon_t^2) = \frac{1}{\sigma_t^2} M(\varepsilon_t - M(\varepsilon_t))^2 = \frac{1}{\sigma_t^2} \sigma_t^2 = 1 - const$$

Следовательно, для преобразованной модели выполняются предпосылки 1-3 МНК. В этом случае оценки полученные по МНК, будут наилучшими линейными несмещенными оценками.

Таким образом ВНК включает следующие этапы:

1. Значения каждой пары наблюдений (x_t, y_t) делят на известную величину σ_t . Тем самым наблюдениям с наименьшими дисперсиями придаются наибольшие «веса», а с максимальными дисперсиями - наименьшие «веса». Действительно, наблюдения с наименьшими дисперсиями отклонений будут более значимыми при оценке

коэффициентов регрессии, чем наблюдения с большими дисперсиями. Учет этого факта увеличивает вероятность получения более точных оценок.

2. По МНК для преобразованных значений $\left(\frac{1}{\sigma_t}, \frac{x_t}{\sigma_t}, \frac{y_t}{\sigma_t}\right)$ строится уравнение регрессии без свободного члена с гарантированными качествами оценок.

1) Дисперсии отклонений неизвестны.

Для применения ВНК необходимо знать фактические значения дисперсий σ_t^2 отклонений. На практике такие значения известны крайне редко. Следовательно, чтобы применить ВНК, необходимо сделать реалистические предположения о значениях σ_t^2 .

Например, может оказаться целесообразным предположить, что дисперсии σ_t^2 отклонений ε_t , пропорциональны значениям x_t или значениям x_t^2 .

а) Дисперсии σ_t^2 пропорциональны x_t , т.е. $\sigma_t^2 = \sigma_t^2 x_t$ (σ_t^2 - коэффициент пропорциональности). Тогда исходное уравнение преобразуется делением его левой и правой частей на $\sqrt{x_t}$:

$$\frac{y_t}{\sqrt{x_t}} = \beta_0 \frac{1}{\sqrt{x_t}} + \beta_1 \frac{x_t}{\sqrt{x_t}} + \frac{e_t}{\sqrt{x_t}} \Rightarrow \frac{y_t}{\sqrt{x_t}} = \beta_0 \frac{1}{\sqrt{x_t}} + \beta_1 \sqrt{x_t} + V_t$$

Можно показать, что для случайных отклонений $V_t = \frac{\varepsilon_t}{\sqrt{x_t}}$ выполняется условие гомоскедастичности. Следовательно, для преобразованной регрессии применим обычный МНК. Действительно,

$$\sigma_t^2 = \sigma^2(\varepsilon_t) = \sigma^2 x_t$$

в силу выполнимости предпосылки $V_t = \frac{\varepsilon_t}{x_t}$ имеем:
 α

$$\sigma_t^2(V_t) = \sigma^2\left(\frac{\varepsilon_t}{\sqrt{x_t}}\right) = \frac{1}{x_t^*} \sigma^2(\varepsilon_t) = \frac{1}{x_t^*} \sigma^2 x_t = \sigma^2 = const$$

Таким образом, оценив для последней модели по МНК-коэффициенты β_0 и, затем возвращаются к исходному уравнению регрессии.

б) Дисперсии σ_t^2 пропорциональны x_t^2 .

Если зависимость σ_t^2 от x_t выражается не линейной функцией а квадратичной то соответствующим преобразованием будет деление уравнения регрессии $y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$, на x_t :

$$\frac{y_t}{x_t} = \beta_0 \frac{1}{x_t} + \beta_1 + \frac{\varepsilon_t}{x_t} \Rightarrow \frac{y_t}{x_t} = \beta_0 \frac{1}{x_t} + \beta_1 + V_t, \text{ где } V_t = \frac{\varepsilon_t}{x_t}$$

Несложно показать, что для отклонений V_t , будет выполняться условие гомоскедастичности. После определения по МНК - оценок коэффициентов β_0 и β_1 для уравнения (4) возвращаются к исходному уравнению $y_t = \beta_0 + \beta_1 x_t + \varepsilon_t$.

Отметим, что для применения описанных выше преобразований весьма значимы знания об исходных значениях дисперсий отклонений σ_t^2 , либо предположения, какими эти дисперсии могут быть. Во многих случаях дисперсии отклонений зависят не от включенных в уравнение регрессий объясняющих переменных, а от тех, которые не включены в модель, но играют существенную роль в исследуемой зависимости. В этом случаи они должны быть включены в модель. В ряде случаев для устранения гетероскедастичности необходимо изменить спецификацию

модели (например, линейную на лог- линейную, мультипликативную на аддитивную и т.п.).

Наличие гетероскедастичности не позволяет получить эффективные оценки, что приводит к необоснованным выводам по их качеству [5,26].

2.6 Линейная модель множественной регрессии с автокорреляцией остатков

Предположим, что нарушается только предпосылка МНК о независимости случайных погрешностей ε_i и ε_j в различных наблюдениях $\text{Cov}(\varepsilon_i, \varepsilon_j) = 0 \quad (i \neq j)$. В этом случае говорят об автокорреляции остатков. Оценки параметров, полученные методом наименьших квадратов, остаются несмещенными, но теряют свою эффективность.

Автокорреляция обычно встречается в регрессионном анализе только при использовании исходных данных в виде временных рядов.

Автокорреляцией остатков модели регрессии называется корреляционная зависимость между настоящими и прошлыми значениями остатков.

Автокорреляция остатков может быть вызвана несколькими причинами, имеющими различную природу:

- 1) наличие ошибок измерения в значениях результативного признака;
- 2) модель может не включать фактор, оказывающий существенное воздействие на результат, влияние которого отражается в остатках, вследствие чего в дальнейшем могут оказаться

автокоррелированными. Очень часто этим фактором является фактор времени t . Кроме того, в качестве таких существенных факторов могут выступать лаговые значения переменных, включенных в модель;

3) модель не учитывает несколько второстепенных факторов, совместное влияние которых на результат существенно ввиду совпадения тенденций их изменения или фаз циклических колебаний;

4) неправильная спецификация функциональной формы модели. В этом случае, следует изменить форму связи факторных и результативного признаков, а не использовать специальные методы расчета параметров уравнения регрессии при наличии автокорреляции остатков.

Применение к модели с автокорреляцией остатков обыкновенного МНК приведет к следующим последствиям:

1. Выборочные дисперсии полученных оценок коэффициентов будут больше по сравнению с дисперсиями по альтернативным методам оценивания, т.е. оценки коэффициентов будут неэффективны.

2. Стандартные ошибки коэффициентов будут оценены неправильно, чаще всего занижены, иногда настолько, что нет возможности воспользоваться для проверки гипотез соответствующими точными критериями - гипотеза о незначимости регрессии, отвергается чаще, чем это следовало бы делать в действительности.

3. Прогнозы по модели получаются неэффективными.

В этом случае возникает проблема тестирования наличия в модели автокорреляции, а также выявления причины автокорреляции

при ее обнаружении: или в модели опущена существенная переменная, или структура ошибок зависит от времени. То есть, исследование остатков позволяет судить о правильности модели и ее пригодности для прогнозирования. Простейшим способом проверки наличия автокорреляции является графическое изображение остатков ε_i . Возможно построение:

- графика временной последовательности, если остатки получены в разные моменты времени;
- графика зависимости остатков от значений Y_x , полученных по регрессии;
- графиков зависимости остатков от объясняющих переменных.

Если изображение остатков представляет собой горизонтальную полосу, это указывает на отсутствие каких-либо проблем, связанных с моделью. В противном случае в зависимости от вида и типа графика можно получить информацию о: неадекватности модели, ошибочности расчетов, необходимости включения в модель линейного или квадратичного члена от времени; наконец о непостоянстве дисперсии.

Ошибки могут коррелировать по-разному, однако без нарушения общности можно рассматривать так называемую сериальную корреляцию (автокорреляцию), когда зависимость между ошибками, отстоящими на некоторое количество шагов s , называемое порядком корреляции (в частности, на один шаг, $s=1$), остается одинаковой.

В исследованиях чаще всего встречается положительная автокорреляция.

Рассмотрим методы определения автокорреляции:

1. Графический метод

На рисунке 2.4 приведены последовательно - временные графики, увязывающие отклонение e_t с моментами t их получения.

По оси абсцисс обычно откладываются либо время получения статистических данных, либо порядковый номер наблюдения, а по оси ординат - отклонения ε_t , (либо оценки отклонений e_t).

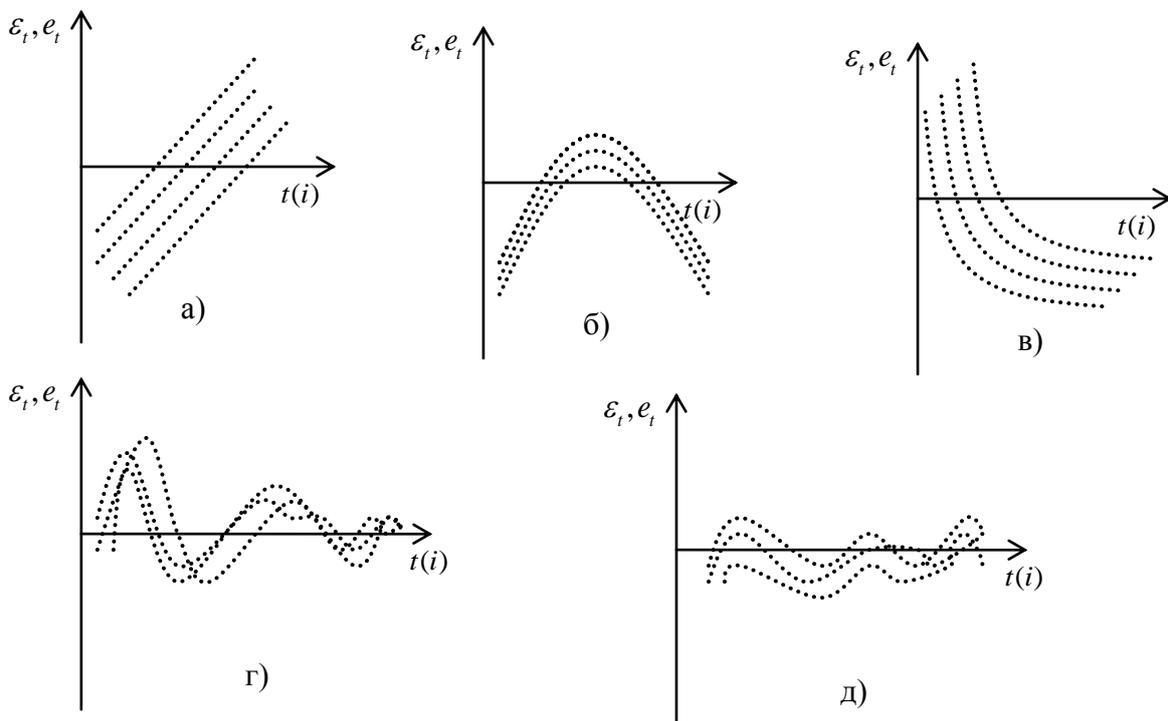


Рисунок 2.4 – Графический метод определения автокорреляции

На рисунке 2.4 а - г) имеются определенные связи между отклонениями, то есть автокорреляция имеет место. Отсутствие зависимости на рисунке д) свидетельствует об отсутствии автокорреляции.

В современных компьютерных прикладных программах для решения задач по эконометрике аналитическое выражение регрессии дополняется графическим представлением результатов. На график

реальных колебаний зависимой переменной накладывается график колебаний переменной по уравнению регрессии. Сопоставив эти два графика, можно выдвинуть гипотезу о наличии автокорреляции остатков. Если эти графики пересекаются редко, то можно предположить наличие положительной автокорреляции остатков.

2. Критерий Дарбина-Уотсона (Durbin - Watson).

Большинство тестов на наличие корреляции по времени в ошибках системы $y = X\beta + \varepsilon$ используют идею: если корреляция есть у ошибок ε , то она присутствует и в остатках e , получаемых после применения к $y = X\beta + \varepsilon$ обычного МНК.

Пусть нулевая гипотеза состоит в отсутствии корреляции, то есть $H_0 : \rho = 0$ (ρ - коэффициент корреляции между соседними ошибками).

В качестве альтернативной может выступать либо просто H_1 : "не H_0 ", либо односторонняя гипотеза: $H_1 : \rho > 0$.

Наиболее широко используется тест Дарбина-Уотсона. Суть его состоит в том, что на основе вычисленной статистики DW (важнейшая характеристика качества регрессионной модели) делается вывод об автокорреляции.

$$DW = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}.$$

Эта статистика тесно связана с выборочным коэффициентом корреляции $r_{e, e_{t-1}}$:

$$DW \approx 2(1 - r_{e, e_{t-1}}),$$

$$\text{где } r \approx \frac{\sum_{t=2}^n e_t e_{t-1}}{\sqrt{\sum_{t=2}^n e_t^2 \sum_{t=1}^{n-1} e_t^2}} \approx \frac{\sum_{t=2}^n e_t e_{t-1}}{\sum_{t=1}^n e_t^2}, \quad 0 \leq DW \leq 4.$$

Содержательный смысл статистики DW : если между e_t и e_{t-1} имеется достаточно высокая положительная корреляция, то e_t и e_{t-1} близки друг к другу и величина статистики DW мала. Действительно, если $r_{e_t e_{t-1}} \approx 1$ (положительная автокорреляция), то $DW = 0$. Отсутствие корреляции означает, что $DW \approx 2$. Если $r_{e_t e_{t-1}} \approx -1$ (отрицательная автокорреляция), то $DW \approx 4$.

Для более точного определения, какое значение DW свидетельствует об отсутствии автокорреляции, а какое - об ее наличии, была построена таблица критических точек распределения Дарбина-Уотсона. По ней для заданного уровня значимости α , числа наблюдений n и количества объясняющих переменных m определяются два значения: d_l - нижняя граница (l - *low* - нижняя), d_u - верхняя граница (u - *upper*) (которые зависят лишь от α, n, m , а, следовательно, могут быть затабулированы) и обладают следующим свойством: если $DW > d_u$, то $DW > d^*$ и, значит, гипотеза H_0 не отвергается; если $DW < d_l$, то $DW < d^*$, и гипотеза H_0 отвергается в пользу H_1 ; если $d_l < DW < d_u$, то ситуация неопределенна, то есть нельзя высказаться в пользу той или иной гипотезы. Если альтернативной является гипотеза об отрицательной корреляции $H_1: \rho < 0$, то соответствующими верхними и нижними границами будут $4 - d_l$ и $4 - d_u$.

Разработаны специальные таблицы критических точек статистики Дарбина-Уотсона, позволяющие при данном числе

наблюдений n , количестве объясняющих переменных m и заданном уровне значимости α определять границы приемлемости (критические точки) наблюдаемой статистики DW .

Для проверки гипотезы об отсутствии автокорреляции остатков используется числовой отрезок, изображенный ниже:



Общая схема критерия Дарбина-Уотсона следующая:

1. По построенному эмпирическому уравнению регрессии:

$$\hat{y}_t = b_0 + b_1 x_{1t} + \dots + b_m x_{mt}$$

определяются значения отклонений $e_t = y_t - \hat{y}_t$ для каждого наблюдения t , $t = 1, 2, \dots, n$.

2. По формуле рассчитывается статистика DW .
3. По таблице критических точек Дарбина-Уотсона определяются два числа d_l и d_u и осуществляются выводы:

Значение статистики DW	Вывод
$4 - d_l < DW < 4$	Гипотеза H_0 отвергается, есть отрицательная корреляция.
$4 - d_u < DW < 4 - d_l$	Неопределенность.
$d_u < DW < 4 - d_u$	Гипотеза H_0 не отвергается.
$d_l < DW < d_u$	Неопределенность.
$0 < DW < d_l$	Гипотеза H_0 отвергается, есть положительная корреляция.

При использовании критерия Дарбина-Уотсона необходимо учитывать следующие ограничения:

1. Критерий DW применяется лишь для тех моделей, которые содержат свободный член.

2. Предполагается, что случайные отклонения ε_t определяются по итерационной схеме: $\varepsilon_t = \rho\varepsilon_{t-1} + v_t$, называемой авторегрессионной схемой первого порядка $AR(1)$; v_t – случайный член.

3. Статистические данные должны иметь одинаковую периодичность (то есть не должно быть пропусков в наблюдениях).

4. Критерий Дарбина-Уотсона не применим для регрессионных моделей, содержащих в составе объясняющих переменных зависимую переменную с временным лагом в один период, то есть для так называемых авторегрессионных моделей вида:

$$y_t = \beta_0 + \beta_1 x_{t1} + \dots + \beta_m x_{tm} + \gamma y_{t-1} + \varepsilon_t.$$

Рассмотрим методы устранения автокорреляции.

Основной причиной наличия случайного члена в модели являются несовершенные знания о причинах и взаимосвязях, определяющих то или иное значение зависимой переменной. Поэтому свойства случайных отклонений, в том числе и автокорреляция, в первую очередь зависят от выбора формулы зависимости и состава объясняющих переменных. Так как автокорреляция чаще всего вызывается неправильной спецификацией модели, то необходимо, прежде всего, скорректировать саму модель. Возможно, автокорреляция вызвана отсутствием в модели некоторой важной объясняющей переменной. Также можно попробовать изменить формулу зависимости (линейную на логарифмическо-линейную).

Однако если все процедуры изменения спецификации исчерпаны, а автокорреляция имеет место, то можно предположить, что она обусловлена какими-то внутренними свойствами ряда $\{e_t\}$. В этом случае можно воспользоваться авторегрессионным преобразованием, в котором

ошибки образуют так называемый авторегрессионный процесс (схему) первого порядка $AR(1)$.

Для изложения $AR(1)$ рассмотрим модель: $y = X\beta + \varepsilon$, где t -я компонента вектора y представляет значение зависимой переменной в момент времени t , $t = \overline{1, n}$. Будем считать, для определенности, что первым регрессором в X является константа.

Запишем уравнение для наблюдения в момент времени t :

$$y_t = \beta_1 + \beta_2 x_{t2} + \dots + \beta_k x_{tk} + \varepsilon_t = x'_t \beta + \varepsilon_t,$$

где $x'_t = (1, x_{t2}, \dots, x_{tk})$ - t -я строка матрицы X .

Один из наиболее простых способов учета коррелированности ошибок (в разные моменты времени) состоит в предположении, что случайная последовательность $\{\varepsilon_t, t = \overline{1, n}\}$ образует авторегрессионный процесс первого порядка. Это означает, что все ошибки удовлетворяют рекуррентному соотношению $\varepsilon_t = \rho \varepsilon_{t-1} + v_t$, где $\{v_t, t = \overline{1, n}\}$ - последовательность независимых нормально распределенных случайных величин с нулевым средним и постоянной дисперсией σ_v^2 , а ρ - некоторый параметр, называемый коэффициентом авторегрессии ($|\rho| < 1$).

Однако на практике, когда значение коэффициента ρ известно, встречается крайне редко; обычно параметр авторегрессии ρ неизвестен и его необходимо оценивать. Существует несколько методов оценивания, как правило, они имеют итеративный характер.

Приведем наиболее употребляемые.

1. Оценивание ρ на основе статистики Дарбина-Уотсона.

Статистика тесно связана с коэффициентом корреляции между соседними отклонениями через соотношение: $DW \approx 2(1 - r_{e_t e_{t-1}})$.

Тогда в качестве оценки коэффициента ρ может быть взят коэффициент $r = r_{e_t e_{t-1}}$, то есть $r \approx 1 - \frac{DW}{2}$.

Этот метод неплох при большом числе наблюдений; оценка r параметра ρ будет достаточно точной.

2. Метод (процедура) Кохрейна-Оркатта (Cochrane-Orcutt)

Опишем итерационный процесс этого метода на примере парной регрессии: $Y = \beta_0 + \beta_1 X + \varepsilon$ и авторегрессионной схемы первого порядка $\varepsilon_t = \rho \varepsilon_{t-1} + v_t$.

1) Оценивается по МНК данная регрессия, и для нее определяются оценки e_t отклонений ε_t , $t = \overline{1, n}$.

2) С использованием схем $AR(1)$ оценивается регрессионная зависимость $e_t = \hat{\rho} e_{t-1} + v_t$, где $\hat{\rho}$ - оценка коэффициента ρ .

3) На основе данной оценки строится уравнение:

$$(y_t - \hat{\rho} y_{t-1}) = \alpha(1 - \hat{\rho}) + \beta(x_t - \hat{\rho} x_{t-1}) + (\varepsilon_t - \hat{\rho} \varepsilon_{t-1})$$

с помощью которого оцениваются коэффициенты α и β (в этом случае значение $\hat{\rho}$ известно).

4) Значения $\beta_0 = \alpha(1 - \hat{\rho})$ и $\beta_1 = \beta$ подставляются в $Y = \beta_0 + \beta_1 X + \varepsilon$.

Вновь вычисляются оценки e_t отклонений и процесс возвращается к этапу 2.

Процесс обычно заканчивается, когда очередное приближение ρ мало отличается от предыдущего. Иногда просто фиксируется

количество итераций. Процедура Кохрейна-Оркатта реализована в большинстве эконометрических компьютерных программ.

3. Процедура Хилдрета-Лу (Hildreth-Lu)

По данному методу регрессия $y_t - \rho y_{t-1} = \beta_0(1 - \rho) + \beta_1(x_t - \rho x_{t-1}) + (\varepsilon_t - \rho \varepsilon_{t-1})$ оценивается для каждого возможного значения ρ из отрезка $[-1; 1]$ с любым шагом (0,1; 0,0001; 0,05 и так далее). Величина $\hat{\rho}$, дающая наименьшую стандартную ошибку регрессии (сумма квадратов отклонений минимальна), принимается в качестве оценки коэффициента ρ . И значения β_0^* и β_1 оцениваются из этого же уравнения регрессии именно с данным значением $\hat{\rho}$. Затем в некоторой окрестности этого значения $\hat{\rho}$ устраивается более мелкая сетка и процесс повторяется. Итерация заканчивается, когда будет достигнута желаемая точность.

Этот итерационный метод широко используется в пакетах прикладных программ.

Таким образом, при установлении автокорреляции необходимо в первую очередь проанализировать правильность спецификации модели. Если после уточнения состава объясняющих переменных, либо изменения формы зависимости автокорреляции по-прежнему имеет место, то, возможно, это связано с внутренними свойствами ряда отклонений $\{\varepsilon_t\}$. В этом случае возможны определенные преобразования, устраняющие автокорреляцию. Среди них выделяется $AR(1)$, которая может быть обобщена в $AR(k)$, $k = 2, 3, \dots$

Для применения схем необходимо оценить коэффициент корреляции между отношениями. Это может быть сделано различными методами: на основе статистики Дарбина-Уотсона,

Кохрейна-Оркатта, Хилдрета-Лу и другие. В случае наличия среди объясняющих переменных лаговой зависимости переменной наличие автокорреляции устанавливается с помощью h статистики Дарбина. А для ее устранения в этом случае предпочтителен метод Хилдрета-Лу.

Последствия автокорреляции:

1. Оценки параметров, оставаясь линейными и несмещенными, перестают быть эффективными. Следовательно, они перестают обладать свойствами наилучших линейных несмещенных оценок (BLUE - оценок).

2. Дисперсии оценок являются смещенными. Часто дисперсии, вычисляемые по стандартным формулам, являются заниженными, что влечет за собой увеличение t -статистик. Это может привести к признанию статистически значимыми объясняющие переменные, которые в действительности таковыми могут и не являться.

3. Оценка дисперсии регрессии $S^2 = \sum_{t=1}^T \frac{e_t^2}{T-m-1}$ (T - объем выборки) является смещенной оценкой истинного значения σ^2 , во многих случаях занижая его.

4. В силу вышесказанного, выводы по t - и F -статистикам, определяющим значимость коэффициентов регрессии и коэффициента детерминации, возможно, будут неверными. Вследствие этого ухудшаются прогнозные качества модели.

Одной из центральных задач эконометрического моделирования является предсказание (прогнозирование) значений зависимой переменной при определенных значениях объясняющих переменных

[94,95]. Здесь возможен двойной подход: либо предсказать условное математическое ожидание зависимой переменной при определенных значениях объясняющих переменных (предсказание среднего значения), либо прогнозировать некоторое конкретное значение зависимой переменной (предсказание конкретного значения).

Одна из важнейших целей моделирования заключается в прогнозировании поведения исследуемого объекта. Термин "прогнозирование" используется в тех ситуациях, когда требуется предсказать состояние системы в будущем.

В основе прогнозирования лежат три взаимодополняющих источника информации о будущем:

- оценка перспектив развития, будущего состояния прогнозируемого явления на основе опыта - при помощи аналогии с достаточно хорошо известными сходными явлениями и процессами;
- условное продолжение в будущее (экстраполяция) тенденций;
- модель будущего состояния явления, построенная сообразно ожидаемым или желательным изменениям ряда условий, перспективы развития которых достаточно хорошо известны.

В соответствии с этим существуют следующие способы разработки прогнозов:

- анкетирование - получение экспертных оценок;
- экстраполирование и интерполирование — построение динамических рядов развития показателей прогнозируемого явления на протяжении периодов основания прогноза в прошлом и упреждения прогноза в будущем;

- моделирование - построение моделей с учетом вероятного или желательного изменения прогнозируемого явления на период упреждения прогноза по имеющимся данным.

Приведенное разделение способов прогнозирования условно, на практике эти способы перекрещиваются и взаимно дополняют друг друга.

Проблема прогнозирования имеет много различных аспектов. Можно различать точечное и интервальное прогнозирование. В первом случае - это конкретное число, во втором - интервал, покрывающий истинное значение переменной с заданным уровнем доверия. Выделяют также условное и безусловное прогнозирование в зависимости от того, известны ли интересующие нас объясняющие переменные точно или приближенно.

Для построения точечного прогноза зависимой переменной в случае безусловного прогнозирования на несколько шагов вперед необходимо знать прогнозные значения всех входящих в модель факторов. Их оценки могут быть получены методами экспертных оценок, или непосредственно заданы исследователем, или могут быть получены на основе экстраполяционных и интерполяционных методов. Прогнозные значения факторов подставляют в модель и получают точечные прогнозные оценки изучаемого показателя [22].

Как определить, как сильно может уклониться модельное среднее значение y^*_0 , рассчитанное по эмпирическому уравнению регрессии, от соответствующего условного математического ожидания. Ответ на этот вопрос дается на основе интервальных оценок, построенных с заданной надежностью.

Доверительный интервал для функции регрессии или для условного математического ожидания зависимой переменной $M_x(Y)$:

$$\hat{y} - t_{1-\alpha;k} \cdot S_{\hat{y}} \leq M_x(Y) \leq \hat{y} + t_{1-\alpha;k} \cdot S_{\hat{y}},$$

где \hat{y} - групповая средняя, определяемая по уравнению регрессии;

$S_{\hat{y}} = S \sqrt{X_0'(X'X)^{-1} X_0}$ - стандартная ошибка уравнения регрессии.

Доверительный интервал y^*_0 для индивидуальных значений зависимой переменной имеет вид:

$$\hat{y}_0 - t_{\alpha;n-k-1} \cdot S_{\hat{y}_0} \leq y^*_0 \leq \hat{y}_0 + t_{\alpha;n-k-1} \cdot S_{\hat{y}_0},$$

где \hat{y}_0 - точечный прогноз зависимой переменной;

$S_{\hat{y}_0} = S \sqrt{1 + X_0'(X'X)^{-1} X_0}$ - стандартная ошибка точечной оценки зависимой переменной.

3 ИСПОЛЬЗОВАНИЕ ЭКОНОМЕТРИЧЕСКИХ МОДЕЛЕЙ В ИССЛЕДОВАНИИ СОЦИАЛЬНЫХ ПРОЦЕССОВ

3.1 Линейная модель с количественными переменными

Область эконометрических исследований в настоящее время значительно расширилась, т.к. в наполнении эмпирическим содержанием нуждаются априорные в разных областях рассуждения. К таким областям относятся исследования и в области медицины, образования, бизнеса, сельского хозяйства и др. [27, 35,39,54].

За прошедшие тысячелетия школа как социальный институт пережила различные изменения. И каждый раз находились пути и средства, чтобы совершенствовать и обновлять школу. Сегодня в условиях происходящих перемен в российской школе и перехода ее к компьютерным технологиям возникают новые задачи. Необходимо детально рассмотреть происходящие в школе изменения и дать им соответствующую оценку. Поиски решения данной проблемы привели к разработке эконометрической модели, которая позволяет рассматривать информатизацию школы как интерактивный процесс в работе муниципальных образовательных учреждений [9].

Информатизация образования – широкомасштабный процесс трансформации содержания, методов и организационных форм учебной работы, обеспечивающий эффективную социализацию школьников к жизни в условиях информационного общества [56].

Решение задач, связанных с оценкой различных форм учебной работы в условиях информатизации общества и школы, эконометрическими методами, методами прикладной статистики,

стало возможным благодаря электронно-вычислительным машинам, оснащенным компьютерными эконометрическими пакетами по реализации математико-статистического анализа.

Сегодня доступными средствами для анализа данных являются различные статистические программные продукты (СПП). Мировая практика исследовательской работы в банковской сфере, в маркетинговых и социологических исследованиях, других социально-экономических областях свидетельствует о широком применении компьютерных систем статистического анализа и обработки данных. За последнее десятилетие вырос спрос на статистические программные продукты и в Российской Федерации. .

В настоящее время существуют множество эконометрических пакетов, в том числе и разработанные в России. Они позволяют наиболее эффективно решать задачи корреляционно-регрессионного анализа и прогнозирования, оказывая существенную помощь исследователю [11,13,53,74].

EconometricViews – специальный пакет для решения эконометрических задач, разработанный компанией QuantitativeMicroSoftware. В работе с пакетом EViews можно рассматривать методы обработки и анализа информации, а также способы табличного и графического представления полученных результатов. Статистические программные продукты позволяют выполнять на компьютере наиболее трудоемкую работу по расчету различных статистик, выявлению наличия статистической зависимости, оценки параметров, построению графиков и прогноза изучаемых показателей [13].

Покажем результаты эконометрического исследования уровня

информатизации общеобразовательных школ Троицкого района Челябинской области на основе имеющихся пространственных данных [34].

Алгоритм решения поставленной проблемы следующий: необходимо определить круг факторов и показателей, участвующих в модели, провести их анализ; осуществить подготовку исходной информации; построить математическую модель и оценить ее параметры; провести численное решение задачи исследования с использованием программных продуктов; провести анализ полученных результатов и их интерпретацию.

Английские ученые разработали Матрицу ВЕСТА для определения состояния информатизации школы [56]:

В Матрице рассматривается шесть уровней информатизации:

1. Процесс информатизации школы еще не организован (Not Yet decided). Использование ИКТ никак не регламентируется.

2. Процесс информатизации связан с отдельными сторонами работы школы (Localized) и в целом не координируется. Учителя используют ИКТ по личной инициативе либо в соответствии с планами работы отдельных методических объединений или кафедр.

3. Процесс информатизации координируется в рамках всей школы (Coordinated). Создан или уже начал действовать механизм поддержки и координации работ по использованию ИКТ.

4. Процесс информатизации развивается (Transformative), для его поддержки и развития систематически прикладываются специальные усилия. Средства ИКТ используются регулярно, хотя, возможно, и не всеми педагогами.

5. Процесс информатизации вышел на новый качественный уровень. Средства ИКТ органично встроены в работу школы (Embedded). Их использование стало нормой, повседневным элементом образовательной работы всех педагогов и учащихся.

6. Использование средств ИКТ преобразует/преобразовало работу школы (Innovative). Они применяются в организованном полном учебно-воспитательном процессе. Учащимся предлагаются новые виды образовательных услуг, цель которых — достижение нового качества образования, решение нетрадиционных образовательных задач.

В соответствии с приведенными уровнями совместно со специалистами МОУ «Информационно-методический центр», «Межшкольный методический центр 74320» была разработана 100-балльная шкала для более качественной оценки уровня информатизации.

1 уровень: 0 – 20 баллов;	4 уровень: 56 – 70 баллов;
2 уровень: 21 – 40 баллов;	5 уровень: 71 – 90 баллов;
3 уровень: 41 – 55 баллов;	6 уровень: 91 – 100 баллов.

В Троицком районе в настоящее время 28 общеобразовательных школ.

В соответствии с матрицей ВЕСТА для определения состояния информатизации школы и разработанной шкалой была проведена детальная оценка уровня информатизации каждой школы (таблица 3.1).

Таблица 3.1 – Оценка уровня информатизации школ

№ п/п	Код школы	Уровень информатизации(бал.)
1	74320_s_001	41
2	74320_s_002	44
3	74320_s_003	42
4	74320_s_004	47
5	74320_s_005	43
6	74320_s_006	38
7	74320_s_007	44
8	74320_s_008	45
9	74320_s_009	39
10	74320_s_010	38
11	74320_s_011	48
12	74320_s_012	43
13	74320_s_013	47
14	74320_s_014	41
15	74320_s_015	35
16	74320_s_016	36
17	74320_s_017	40
18	74320_s_018	30
19	74320_s_019	36
20	74320_s_020	35
21	74320_s_021	33
22	74320_s_022	35
23	74320_s_023	34
24	74320_s_024	33
25	74320_s_025	31
26	74320_s_026	32
27	74320_s_027	32
28	74320_s_028	27

Метод анкетирования был использован в процессе сбора данных [10]. Анкета информатизации была разработана и составлена специалистами Национального Фонда Подготовки Кадров.

Все входящие в анкету показатели делятся на две группы: количественные (объемные), значения которых можно описать с

помощью разностной шкалы (число), и качественные, значения которых можно оценить лишь с помощью шкалы порядка (уровень выполнения некоторого требования).

Количественные показатели описывают, объемы физически доступных ресурсов (например, общее количество компьютеров в школе или количество компьютеров, приходящихся на одного педагога).

Качественные показатели характеризуют обеспеченность тем или иным ресурсом или степень достижения некоторого результата. Так, описывая школьную компьютерную сеть, важно зафиксировать не столько ее технологические характеристики (скорость обмена информацией и пр.), которые определяются уровнем развития ИКТ и не зависят от усилий педагогов, сколько ее распространенность в рамках школы (она может соединять компьютеры внутри одного или нескольких классов, объединять большинство или всего несколько компьютеров, быть доступной с каждого рабочего места, в любом из помещений школы и т.д.). Это позволит оценить, в какой мере компьютерная сеть охватывает школу. Распространенность сети зависит от работников школы и во многом определяет практическую доступность ИКТ-ресурсов.

Вопросы, содержащиеся в Анкете, были условно поделены на 8 групп, в соответствии с которыми определились основные факторы [7]:

Фактор 1 - «доступность аппаратных средств» - описывает состав, количество и распределение аппаратных средств по школе, уровень доступности компьютеров для учителей и учеников. Высокое значение этого фактора наблюдается у тех школ, где не только много компьютеров, видео- и аудиокomплексов, но и сами эти средства находятся в свободном доступе.

Фактор 1 объединил следующую группу вопросов:

- Общая численность учащихся в данном учебном году
- Общее количество учебных помещений школы
- Среднегодовая численность учителей
- Оцените, пожалуйста, общее количество компьютеров, которые реально используются в настоящее время в школе
- Оцените, пожалуйста, общее количество современных компьютеров в школе
- Где постоянно расположены компьютеры, которые используются в вашей школе администраторами, учителями и учениками для учебных целей?
- Оцените, пожалуйста, количество устройств для ввода графической информации, которые используются в школе
- Какова ситуация в школе с цифровыми учебными инструментами?
- Какое из предложенных утверждений лучшим образом характеризует использование принтеров в школе?
- Какое из предложенных утверждений лучшим образом характеризует использование в школе компьютерных проекционных устройств?

Фактор 2 - «уровень организации информационно-технологической среды школы» - описывает (условия для использования ЦОР и ИКТ в учебном процессе).

Высокие значения фактора будут у тех школ, где уделяют внимание организации информационно-технологической среды: налажена работа технической службы, используются базы данных,

выделен специальный персонал и введены соответствующие организационно-распорядительные документы.

Фактор 2 объединил следующую группу вопросов:

- Опишите подключение школьных компьютеров к локальной сети школы.
- Опишите распространенность локальной сети по школе.
- Кто отвечает за информатизацию в школе?
- Как организована в школе техническая поддержка использования ИКТ?
- Как организована в школе методическая поддержка использования ИКТ учителями?
- Какие из предложенных утверждений наилучшим образом характеризуют реально выполняемые регламенты (правила, приказы и т.п.) использования ИКТ в Вашей школе.
- Какие из предложенных утверждений наилучшим образом характеризуют действия администрации, направленные на поддержку использования ИКТ педагогами школы.
- Опишите положение дел в школе с организацией доступа, обмена и хранения учебно-методическими материалами?

Фактор 3 - «использование Интернета в учебном процессе».

Фактор 3 объединил следующую группу вопросов:

- Какое из предложенных утверждений лучшим образом характеризует использование Интернет учащимися школы?
- Какое из предложенных утверждений лучшим образом характеризует использование Интернет в работе педагогов?

- Какое из предложенных утверждений лучшим образом характеризует подключение школы к Интернет?

Фактор 4 - «использование ЦОР в учебном процессе».

Фактор 4 объединил следующую группу вопросов:

- Какое из предложенных утверждений лучшим образом характеризует обеспеченность школы цифровыми учебными материалами для проведения учебных занятий?
- Оцените, пожалуйста, использование педагогами в повседневной практике школы цифровых инструментов и технологии?
- Опишите положение дел в школе в области разработки и использования цифровых учебно-методических материалов?
- Какое из предложенных утверждений лучшим образом характеризует положение дел в школе с использованием цифровых образовательных ресурсов, разработанных в проекте "Информатизация системы образования"

Фактор 5 - «ИКТ-подготовка учителей».

Фактор 5 объединил следующую группу вопросов:

- Оцените, пожалуйста, количество работников школы, которые уверенно и регулярно (не реже одного раза в неделю) используют ИКТ в своей профессиональной деятельности
- Оцените, пожалуйста, количество учителей школы, которые в своей педагогической практике способны реализовать нижеперечисленные профессиональные задачи.
- Оцените, пожалуйста, использование педагогами школы ИКТ в перечисленных ниже формах учебной работы?
- Как в школе организована работа по переподготовке учителей для использования ИКТ в учебном процессе.

Фактор 6 - «использование ИКТ как инструмента учебной работы школьников».

Фактор 6 объединил следующую группу вопросов:

- Опишите положение дел в школе с использованием ИКТ при выполнении общешкольных проектов.
- Какое из предложенных утверждений лучшим образом характеризует положение дел в школе с учебными заданиями, для выполнения которых используются ИКТ?
- Какие учебные задания ожидает получить ученик, когда он идет в школу? (Какого типа задания может получить ученик в вашей школе?)
- Какое из предложенных утверждений лучшим образом характеризует положение дел в школе с использованием ИКТ для выполнения учебных проектов?
- Какое из предложенных утверждений лучшим образом характеризует процедуру установления различных видов связей между работой школьников по информатике и по другим учебным дисциплинам в школе?

Фактор 7 - «использование ИКТ для решения административных задач».

Фактор 7 объединил следующую группу вопросов:

- Укажите, пожалуйста, количество рабочих мест администрации, оснащенных компьютерами
- Какое из предложенных утверждений лучшим образом характеризует положение дел в Вашей школе с использованием ИКТ для ведения школьного делопроизводства?

- Опишите положение дел в школе с использованием ИКТ для планирования учебного процесса?
- Какое из предложенных утверждений лучшим образом характеризует положение дел в школе с использованием ИКТ для контроля над ходом учебного процесса?
- Какое из предложенных утверждений лучшим образом характеризует положение дел в школе с использованием ИКТ для поддержания необходимого уровня материально-технической базы школы?

Фактор 8 - «современность аппаратных средств» - описывает соответствие имеющихся в школе аппаратных средств современным требованиям. Высокое значение данного фактора наблюдается у школ, которые оснащены современной техникой (например, современными компьютерами со встроенными звуковыми платами, станками, лабораториями и пр.

Фактор 8 объединил следующую группу вопросов:

- Укажите, пожалуйста, количество современных компьютеров в школе, которые предназначены для учителей
- Укажите, пожалуйста, количество современных компьютеров, которые находятся в общем пользовании работников школы
- Оцените, пожалуйста, количество современных компьютеров в свободном доступе для учащихся школы

Сбор данных проводился в течение апреля 2007 года с помощью Анкет информатизации школ, которые заполнялись представителями школ (учителями информатики, членами школьных команд, работниками администрации) на специально организованных семинарах. Им рассказывалось о целях работы, и предлагалась заранее

подготовленная инструкция. Анкеты заполнялись в электронном виде (на компьютере), сохранялись в виде отдельного файла.

На некоторые вопросы анкеты следовало дать ответ в виде числа (количество учащихся, компьютеров, проекторов, кабинетов, устройств ввода графической информации и т.п.). Для дальнейшей обработки все абсолютные значения были пересчитаны в относительные (количество учеников на один компьютер, проекторов на один учебный кабинет и т.п.). Каждому вопросу с выбором ответа был присвоен номер, а различным ответам — числовые значения (ранги). Например, вопрос о подключении школы к Интернету предлагал выбор одного из следующих ответов:

- в школе нет доступа к Интернету;
- в школе есть одно рабочее место с доступом к Интернету (модем по телефонной линии);
- в школе есть несколько рабочих мест с доступом к Интернету (свои модемы используют секретарь школы, кабинет информатики, библиотека и т.п.);
- школа подключена к Интернету по выделенной линии.

Кодируя этот вопрос, мы ввели переменную «Интернет», которая принимала целые значения от 0 (в школе нет доступа к Интернету) до 3 (школа подключена к Интернету по выделенной линии). Мы исходили из того, что разница между близлежащими значениями (рангами) незначительна, и не стали использовать весовые коэффициенты. Конечно, разница между рангами 0 (в школе нет доступа к Интернету) и 1 (в школе есть одно рабочее место с доступом к Интернету) не сопоставима с разницей между рангами 2 (в школе есть несколько рабочих мест с доступом к Интернету) и 3 (школа подключена к

Интернету по выделенной линии). Однако в обоих случаях мы оставили ее равной единице. Представляется, что для нашей задачи возникающей погрешностью можно пренебречь.

Введены следующие обозначения:

- Y – уровень информатизации;
- X_1 – первый фактор - «доступность аппаратных средств»;
- X_2 – второй фактор - «уровень организации информационно-технологической среды школы»;
- X_3 – третий фактор - «использование Интернета в учебном процессе»;
- X_4 – четвертый фактор - «использование ЦОР в учебном процессе»;
- X_5 – пятый фактор - «ИКТ-подготовка учителей»;
- X_6 – шестой фактор - «использование ИКТ как инструмента учебной работы школьников»;
- X_7 – седьмой фактор - «использование ИКТ для решения административных задач»;
- X_8 – восьмой фактор - «современность аппаратных средств».

При проведении анализа анкет информатизации были получены следующие значения выбранных факторов (таблица 3.2):

Таблица 3.2 – Значения выбранных факторов

№ п/п	Код школы	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈
1	74320_s_001	0,40	18,00	4,80	11,90	40,30	9,40	9,90	1,03
2	74320_s_002	1,80	18,30	0,00	10,20	18,40	15,90	8,30	0,10
3	74320_s_003	0,76	18,60	3,00	7,80	35,70	9,60	4,80	0,08
4	74320_s_004	0,76	15,40	7,00	11,80	38,50	14,80	6,30	0,06
5	74320_s_005	0,79	11,50	9,80	9,80	18,20	6,40	7,60	1,08
6	74320_s_006	0,78	11,60	0,00	10,30	21,20	14,30	5,30	3,08
7	74320_s_007	0,79	19,00	7,00	11,40	18,40	11,50	7,00	0,07
8	74320_s_008	1,14	19,10	7,00	12,80	18,90	11,70	7,10	0,07
9	74320_s_009	0,85	16,00	0,00	5,60	12,90	8,60	5,00	0,04
10	74320_s_010	0,80	20,00	0,00	11,70	29,40	16,40	8,20	0,09
11	74320_s_011	1,31	12,80	3,00	8,10	25,30	18,30	7,40	1,14
12	74320_s_012	0,81	13,00	4,00	11,90	32,00	16,00	8,50	3,07
13	74320_s_013	1,16	20,60	1,00	9,80	32,60	21,30	8,90	0,10
14	74320_s_014	0,89	20,80	5,00	10,20	25,60	20,60	10,90	0,06
15	74320_s_015	1,83	9,30	0,00	10,40	15,0	11,60	6,10	0,12
16	74320_s_016	1,07	11,10	1,00	9,50	33,30	15,00	7,80	0,15
17	74320_s_017	0,86	14,80	0,00	10,90	29,10	10,60	7,20	2,11
18	74320_s_018	0,38	13,00	0,00	10,40	13,50	13,50	3,60	1,14
19	74320_s_019	0,84	12,50	0,00	11,0	13,40	14,90	4,30	0,13
20	74320_s_020	0,92	16,10	0,00	9,50	16,30	8,60	5,10	0,10
21	74320_s_021	1,23	7,40	0,00	7,20	13,70	8,10	5,70	0,13
22	74320_s_022	0,86	10,20	0,00	8,10	16,50	16,00	6,00	0,13
23	74320_s_023	0,52	13,60	0,00	9,20	22,10	10,80	3,90	0,08
24	74320_s_024	0,51	6,80	0,00	5,40	15,20	14,20	4,30	0,09
25	74320_s_025	1,05	6,60	0,00	7,30	9,00	14,30	6,70	3,12
26	74320_s_026	0,53	14,50	0,00	8,90	19,00	8,70	5,50	0,10
27	74320_s_027	0,66	9,90	0,00	7,00	9,90	10,00	3,50	0,13
28	74320_s_028	1,14	6,00	0,00	10,60	15,30	10,20	2,90	0,19

Таким образом, была получена пространственная выборка, представленная в виде матрицы размерности 28×8 .

В качестве экзогенных (объясняющих) переменных выступают отобранные выше факторы ($X_1 - X_8$), эндогенная (объясняемая) переменная – уровень информатизации (Y) школ Троицкого района.

Проблеме спецификации модели уделено большое внимание. Предварительный анализ данных показал, что использование логарифмической и полулогарифмической моделей затруднено в связи с малыми значениями отдельных факторов (X_1, X_8), поэтому остановились на выборе линейной модели множественной регрессии [6,17] вида:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki} + \varepsilon_i, \quad i=1,2,\dots,n$$

Линейное описание помогает интерпретировать процесс информатизации на разных уровнях, увидеть место, которое занимают те или иные работы по внедрению ИКТ, что нас и интересует; поэтому остановились на выборе линейной модели множественной регрессии. Предварительный анализ данных также показал, что использование логарифмической и полулогарифмической моделей затруднено в связи с малыми значениями отдельных факторов (X_1, X_8).

Рассмотрим линейную модель множественной регрессии

$$\hat{Y} = X\beta, \quad \text{где } X = (1, X_1, \dots, X_8) \quad (3.1)$$

Оценим коэффициенты регрессии с помощью МНК; оценим качество коэффициентов регрессии и модели в целом (см. таблицу 3.3). (Промежуточные вычисления приведены в Приложении 1, таблица 1).

Таблица 3.3 – Оценка качества коэффициентов модели (3.1)

Коэффициенты	Значения	t-статистика	Доверительный интервал	
			Нижняя граница	Верхняя граница
β_0	23,04	8,18	17,14	28,93
β_1	6,97	4,40	3,65	10,29
β_2	0,74	4,76	0,42	1,07
β_3	1,22	5,89	0,78	1,65
β_4	-0,88	-2,90	-1,52	-0,25
β_5	0,21	2,99	0,06	0,36
β_6	0,29	2,02	-0,01	0,60
β_7	-0,72	-1,79	-1,56	0,12
β_8	1,43	2,59	0,27	2,59

Критическое значение распределения Стьюдента при $\alpha=0,05$ и числе степеней свободы $(n-k-1) = (28 - 8 - 1)$ равно $t_{кр} = 2,093$. Из таблицы 3.3 видим:

$$|t_{\beta_6}| = 2,02 < t_{кр} = 2,093,$$

$$|t_{\beta_7}| = 1,79 < t_{кр} = 2,093,$$

кроме того, в доверительный интервал для этих коэффициентов входит ноль, то есть с 95% вероятностью можно утверждать, что коэффициенты при переменных X_6 и X_7 статистически не значимы. По формулам (2.6), (2.8) находим $R^2 = 0,88$, а, следовательно $R_{корр}^2 = 0,83$.

Таким образом скорректированный коэффициент детерминации показывает, что построенная модель достаточно качественно подогнана к наблюдаемым значениям Y_x .

На основе F – статистики Фишера определяем, что модель в целом значима: $F = 17,15 > F_{кр}(0,05;8;19) = 2,48$.

Рассчитаем среднюю относительную ошибку аппроксимации:

$$\bar{E}_{отн} = 4\% , \text{ т.е. модель точная.}$$

Исключим из модели переменные X_6 и X_7 , как статистически незначимые, получим новое уравнение регрессии:

$$\hat{Y} = X\beta, \quad \text{где } X = (1, X_1, \dots, X_5, X_8) \quad (3.2)$$

Снова оценим коэффициенты регрессии с помощью МНК; оценим качество коэффициентов регрессии и модели в целом. (Промежуточные вычисления приведены в Приложении 1, табл. 2).

Результаты вычислений представлены в таблице 3.4.

Таблица 3.4 – Оценка качества коэффициентов модели (3.2)

Коэффициенты	Значения	t-статистика	Доверительный интервал	
			Нижняя граница	Верхняя граница
β_0	24,1	8,18	17,99	30,26
β_1	6,31	4,31	3,27	9,35
β_2	0,67	4,50	0,36	0,98
β_3	1	5,08	0,59	1,40
β_4	-0,8	-2,058	-1,48	-0,13
β_5	0,19	2,74	0,056	0,34
β_8	1,19	2,19	0,06	2,32

Критическое значение распределения Стьюдента при $\alpha = 0,05$ и числе степеней свободы $(n-k-1) = (28 - 6 - 1)$ равно $t_{кр} = 2,080$. Из табл. 3.4 видим: $|t_{\beta_4}| = 2,06 < t_{кр} = 2,080$,

то есть с 95% вероятностью можно утверждать, что коэффициент при переменной X_4 статистически не значим. По формулам (2.6), (2.8) находим $R^2 = 0,84$, а, следовательно $R_{скорр}^2 = 0,80$.

Построенная модель достаточно качественно подогнана к наблюдаемым значениям Y_x .

На основе F – статистики Фишера определяем, что модель в целом значима: $F = 18,88 > F_{кр}(0,05; 6; 21) = 2,57$.

Рассчитаем среднюю относительную ошибку аппроксимации:

$\bar{E}_{отн} = 4,9\%$, т.е. модель точная.

Оценки коэффициентов регрессии показали, что коэффициент при X_4 можно исключить из модели как неинформативный.

Найдем МНК-оценки следующего уравнения (см. табл. 3.5):

$$\hat{Y} = X\beta, \text{ где } X = (1, X_1, \dots, X_3, X_5, X_8) \quad (3.3)$$

(Промежуточные вычисления приведены в Приложении 1, табл. 3).

Таблица 3.5 – Оценка качества коэффициентов модели (3.3)

Коэффициенты	Значения	t-статистика	Доверительный интервал	
			Нижняя граница	Верхняя граница
β_0	19,79	7,50	14,32	25,26
β_1	5,60	3,52	2,28	8,91
β_2	0,56	3,54	0,23	0,88
β_3	0,86	4,11	0,42	1,29
β_5	0,16	2,08	0,00	0,32
β_8	0,87	1,48	-0,35	2,08

Критическое значение распределения Стьюдента при $\alpha = 0,05$ и числе степеней свободы $(n-k-1) = (28-5-1)$ равно $t_{кр} = 2,074$. Из табл. 3.5 видим: $|t_{\beta_8}| = 1,48 < t_{кр} = 2,074$, то есть с 95% вероятностью можно утверждать, что коэффициент при переменной X_8 статистически не значим. По формулам (2.6), (2.7) находим $R^2 = 0,80$, а, следовательно $R_{скорр}^2 = 0,75$.

Построенная модель достаточно качественно подогнана к наблюдаемым значениям Y_x .

На основе F – статистики Фишера определяем, что модель в целом значима: $F = 17,37 > F_{кр}(0,05;5;22) = 2,66$.

Рассчитаем среднюю относительную ошибку аппроксимации:

$\bar{E}_{отн} = 5,4\%$ - прогноз по этой модели будет допустимо точным.

Так как коэффициент при переменной X_8 оказался статистически незначимым по критерию Стьюдента, принимаем решение об исключении его из модели.

Оценим с помощью МНК коэффициенты уравнения (см. таблицу 3.6)

$$\hat{Y} = X\beta, \quad \text{где } X=(1, X_1, \dots, X_3, X_5) \quad (3.4)$$

(Промежуточные вычисления приведены в Приложении 1, табл. 4).

Таблица 3.6 – Оценка качества коэффициентов модели (3.4)

Коэффициенты	Значения	t-статистика	Доверительный интервал	
			Нижняя граница	Верхняя граница
β_0	21,10	8,29	15,84	26,37
β_1	5,38	3,31	2,02	8,75
β_2	0,48	3,15	0,16	0,79
β_3	0,87	4,08	0,43	1,32
β_5	0,18	2,36	0,02	0,34

Критическое значение распределения Стьюдента при $\alpha = 0,05$ и числе степеней свободы $(n-k-1) = (28-4-1)$ равно $t_{кр} = 2,069$. Из таблицы 3.6 видим, что t -статистики всех коэффициентов регрессии больше критического значения, то есть с 95% вероятностью можно утверждать, что все коэффициенты уравнения статистически значимы. По формулам (2.6), (2.7) находим $R^2 = 0,78$, а, следовательно $R^2_{корр} = 0,74$.

Построенная модель достаточно качественно подогнана к наблюдаемым значениям Y_x .

На основе F – статистики Фишера определяем, что модель в целом значима: $F = 20,12 > F_{кр}(0,05;4;23) = 2,80$.

Рассчитаем среднюю относительную ошибку аппроксимации:

$\bar{E}_{отн} = 5,6\%$ - модель допустимо точная.

Итак, получена следующая модель:

$$\hat{Y} = 21,10 + 5,38X_1 + 0,48X_2 + 0,87X_3 + 0,18X_5, \quad (3.5)$$

все коэффициенты которой статистически значимы. Сама модель допустимо точная, коэффициент детерминации $R^2 = 0,78$ показывает,

что вариация зависимой переменной на 78% объясняется вариацией независимых переменных, признаем модель хорошо подогнанной наблюдаемым значениям.

Проверим адекватность модели, выполнение предпосылок МНК.
Случайный характер остатков;

Для проверки случайного характера остатков ε_i , строим график зависимости остатков ε_i от теоретических значений результативного признака.

Из рисунка 3.1 видно, что остатки расположены равномерно (представляют собой горизонтальную полосу), то есть остатки носят случайный характер;

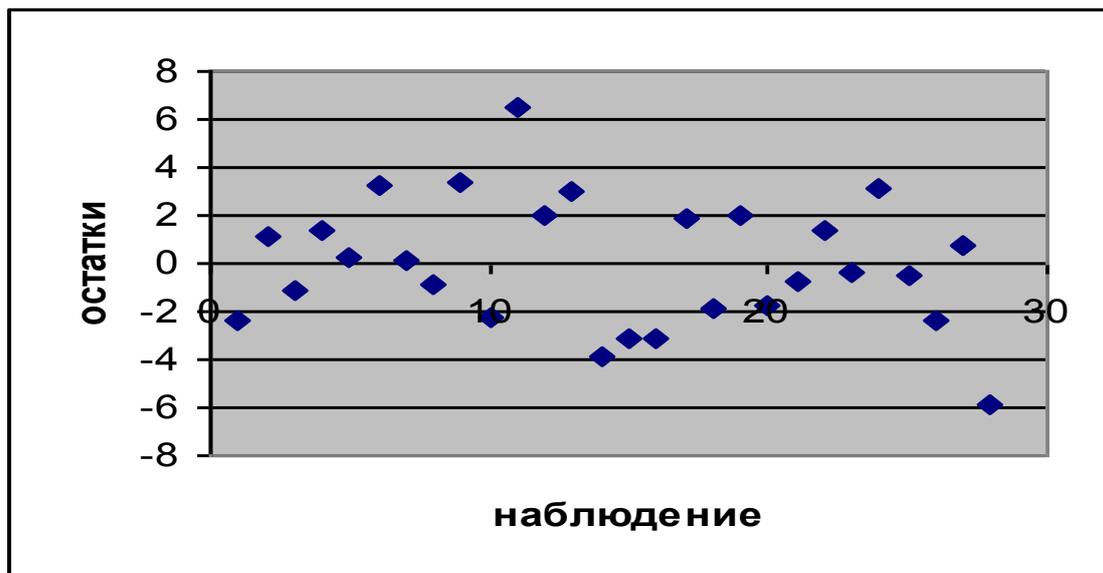


Рисунок 3.1 – График зависимости остатков ε_i от теоретических значений результативного признака

Нулевая средняя величина остатков: в нашем случае $|\varepsilon_i| = 3,55 \cdot 10^{-14}$, поэтому t -критерий выполняется автоматически, математическое ожидание уровней остаточного ряда близко к нулю.

Уровни остаточного ряда независимы, так как автокорреляция отсутствует: значение статистики Дарбина-Уотсона $DW = 2,06$ попадает в интервал $(2; 4-d_U)$ (значение d_U в нашем случае равно 1,747).

Уровни остаточного ряда подчинены нормальному закону распределения, т.к. $RS = 3,86$ попадает в интервал, образованный нижней границей 3,41 и верхней границей 4,80 RS -критерия.

Проверим модель на мультиколлинеарность, построим матрицу парных коэффициентов корреляции (см. таблицу 3.7), проверим статистическую значимость коэффициентов корреляции с помощью t -критерия (см. таблицу 3.8).

Таблица 3.7 – Матрица парных коэффициентов корреляции

	Y	X_1	X_2	X_3	X_5
Y	1				
X_1	0,21	1			
X_2	0,67	-0,11	1		
X_3	0,65	-0,03	0,47	1	
X_5	0,60	-0,16	0,53	0,38	1

Таблица 3.8 – Расчетные значения t -критерия Стьюдента для парных коэффициентов корреляции

	Y	X_1	X_2	X_3
X_1	1,1			
X_2	4,56	-0,33		
X_3	4,39	-0,59	2,03	
X_5	3,79	-0,81	3,16	2,08

Видно, что коэффициенты корреляции между независимыми переменными и зависимой переменной значимы (табличное значение t -статистики $t_{кр} = 2,056$). Корреляционная матрица показывает

умеренную связь между переменными X_2 и X_5 ; X_2 и X_3 , но ее значение не превышает критического уровня, поэтому принимаем решение об отсутствии мультиколлинеарности.

Так как переменная X_2 наиболее тесно связана с зависимой переменной, то проверим модель на гетероскедастичность относительно X_2 , при помощи теста Голдфелда-Куандта:

$$F = \frac{S_3}{S_1} = \frac{25,29}{35,35} = 0,72$$

Табличное значение F -статистики при $\nu_1 = \nu_2 = 5$ степенях свободы и уровне значимости $\alpha = 0,05$ равно 5,05, поэтому принимаем решение об отсутствии гетероскедастичности.

Reset –тест показал, что линейная модель при включении в нее указанных факторов наиболее точно описывает реальный процесс, включение в модель нелинейных членов не целесообразно.

Рассчитаем для нашей модели частные коэффициенты эластичности, которые характеризуют относительное изменение зависимой переменной при изменении объясняющей переменной на 1%, и частные бета-коэффициенты которые показывают, на какую часть величины своего среднего квадратического отклонения изменится в среднем значение результативного признака при изменении факторного признака на величину его среднеквадратического отклонения. Для удобства занесем значения частных коэффициентов эластичности и бета – коэффициентов в таблице 3.9

Таблица 3.9 – Значения частных коэффициентов эластичности и бета – коэффициентов

Переменная	β_i	Частный коэффициент эластичности	Бета-коэффициент
1	21,10	0,55	0,00
X_1	5,38	0,13	0,33
X_2	0,48	0,17	0,37
X_3	0,87	0,04	0,45
X_5	0,18	0,10	0,28

Таким образом, наиболее значимыми факторами, влияющими на уровень информатизации общеобразовательных школ являются:

Фактор 1 - «доступность аппаратных средств»

Фактор 2 - «уровень организации информационно-технологической среды школы»

Фактор 3 - «использование Интернета в учебном процессе».

Фактор 5 - «ИКТ-подготовка учителей».

С помощью этих показателей можно дать оценку уровня информатизации отдельно взятой общеобразовательной школы.

3.2 Линейная смешанная модель (с количественными и качественной переменной)

Содержательная интерпретация фактора X_3 , а также его наблюдаемые значения позволяют сделать предположение о том, что его можно исключить из модели и ввести вместо фактора X_3 качественную переменную D_1 , принимающую следующие значения:

$$D_1 = \begin{cases} 1, & \text{если в школе есть подключение к Internet;} \\ 0, & \text{в противном случае.} \end{cases}$$

Действительно, логично предположить, что при отсутствии Интернета использование ресурсов Сети в образовательной деятельности невозможно, поэтому ответы на такие вопросы Анкеты, как использование Интернета учителями и учащимися автоматически исключаются при отрицательном ответе на вопрос о подключении школы к Сети.

Рассмотрим уравнение:

$$\hat{Y} = X\beta + \alpha_1 D_1 \quad (3.6)$$

Оценим с помощью МНК коэффициенты уравнения (3.6), (см. таблицу 3.10). (Промежуточные вычисления приведены в Приложении 1 таблица 5).

Таблица 3.10 – Оценка качества коэффициентов модели (3.6)

Коэффициенты	Значения	t-статистика	Доверительный интервал	
			Нижняя граница	Верхняя граница
β_0	24,09	6,3	16,08	32,10
β_1	5,28	2,6	1,03	9,53
β_2	0,71	3,45	0,28	1,14
β_3	6,15	3,48	2,45	9,86
β_4	-0,38	-1,00	-1,17	0,42
β_5	0,06	0,60	-0,15	0,28
β_6	0,04	0,24	-0,33	0,42
β_7	-0,31	-0,60	-1,37	0,75
β_8	1,26	1,74	-0,26	2,78

Критическое значение распределения Стьюдента при $\alpha = 0,05$ и числе степеней свободы $(n-k-1) = (28-8-1)$ равно $t_{кр} = 2,093$. Из таблицы

3.10 видно, что большинство коэффициентов уравнения статистически незначимы.

По формулам (2.6), (2.7) находим $R^2 = 0,79$, а, следовательно $R^2_{\text{скорр}} = 0,70$.

Таким образом, можно судить о достаточно точной подгонке построенной модели наблюдаемым данным.

Полученные результаты свидетельствуют о наличии мультиколлинеарности. Проведем анализ переменных на мультиколлинеарность. Для этого рассчитаем матрицу парных коэффициентов корреляции (см. таблицу 3.11):

Таблица 3.11 – Матрица парных коэффициентов корреляции

	Y	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
Y	1								
X_1	0,21	1							
X_2	0,67	-0,06	1						
X_3	0,74	-0,03	0,47	1					
X_4	0,37	0,057	0,45	0,44	1				
X_5	0,6	-0,16	0,53	0,38	0,44	1			
X_6	0,35	0,226	0,25	-0,1	0,17	0,28	1		
X_7	0,64	0,227	0,55	0,46	0,4	0,56	0,46	1	
X_8	0,02	-0,11	-0,26	0	0,12	0,05	0,09	0,2	1

Оценим значимость парных коэффициентов корреляции на основе t -статистики Стьюдента. Расчетные значения приведем в таблице 3.12

Таблица 3.12 – Расчетные значения t -критерия Стьюдента для парных коэффициентов корреляции

	Y	X_1	X_2	X_3	X_4	X_5	X_6	X_7
X_1	1,10							
X_2	4,56	-0,32						
X_3	5,69	-0,12	2,70					
X_4	2,01	0,29	2,58	1,99				
X_5	3,79	-0,80	3,16	4,56	2,52			
X_6	1,89	1,18	1,31	1,30	0,87	1,47		
X_7	4,30	1,19	3,37	3,89	2,25	3,46	2,66	
X_8	0,08	-0,54	-1,36	-0,02	0,61	0,26	0,47	0,79

Видим, что многие коэффициенты корреляции не значимы, так как значения их t -статистик не превышает табличного значения $t_{кр}=2,056$. При этом парные коэффициенты корреляции между X_2 и X_7 , X_2 и X_5 , X_5 и X_7 статистически значимы, то есть из модели в первую очередь необходимо переменные X_2 и X_5 , но практическая значимость переменных X_2 и X_5 не позволяет сделать этого.

Делаем вывод о непригодности использования модели с количественными и качественной переменными для нашего исследования: использование Интернета в учебном процессе является важным аспектом информатизации общеобразовательной школы.

3.3 Прогнозирование

В Троицком районе 17 средних (полных) и 11 основных общеобразовательных школ. Сравним средние уровни информатизации основных и средних (полных) общеобразовательных школ Троицкого района. Для этого вычислим средние значения факторов X_1, X_2, X_3, X_5 для обеих групп школ. Средние значения

показателей информатизации средних (полных) общеобразовательных школ составляют:

\bar{X}_{1cp}	\bar{X}_{2cp}	\bar{X}_{3cp}	\bar{X}_{5cp}
0,99	15,94	3,29	25,98

Средние значения показателей информатизации основных общеобразовательных школ составляют:

$\bar{X}_{1осн}$	$\bar{X}_{2осн}$	$\bar{X}_{3осн}$	$\bar{X}_{5осн}$
0,79	10,60	0,00	14,90

После проведенного анализа была выбрана линейная модель множественной регрессии с количественными переменными:

$$\hat{Y} = 21,10 + 5,38X_1 + 0,48X_2 + 0,87X_3 + 0,18X_5$$

как более соответствующая исходным выборочным данным. Здесь X_1 - «доступность аппаратных средств»; X_2 - «уровень организации информационно-технологической среды школы»; X_3 - «использование Интернета в учебном процессе»; X_5 - «ИКТ-подготовка учителей».

Подставим значения из этих таблиц в модель.

Тогда для средних школ:

$$\hat{Y} = 21,10 + 5,38 \cdot 0,99 + 0,48 \cdot 15,94 + 0,87 \cdot 3,29 + 0,18 \cdot 25,98 = 41,61;$$

для основных школ:

$$\hat{Y} = 21,10 + 5,38 \cdot 0,79 + 0,48 \cdot 10,60 + 0,87 \cdot 0,00 + 0,18 \cdot 14,90 = 33,12.$$

Как видно из полученных оценок, уровни информатизации средних и основных школ существенно различаются. Первые начинают выходить на третий уровень (Coordinated), тогда как вторые уверенно находятся на втором (Localized).

В настоящее время активно ведется подключение школ к сети Интернет, не прекращаются поставки в школы новой компьютерной техники. Однако практически не совсем верно определять возможный уровень информатизации как отдельно взятой школы, так и среднее значение их совокупности, увеличивая (уменьшая) значения только какого-либо фактора.

К концу года все школы района должны иметь свободный доступ в Интернет по выделенной линии или через спутник. Предположим, что ресурсы сети будут активно использоваться во всех школах, как учителями так и учениками, тогда среднее значение фактора X_3 должно быть примерно равно 7.

В настоящее время среднее значение уровня информатизации общеобразовательных школ находится на уровне 38,2.

Оценим уровень информатизации средних и общих общеобразовательных школ Троицкого района в предположении, что среднее значение фактора X_3 равно 7, а средние значения остальных факторов останутся неизменными

Средние значения показателей информатизации средних (полных) общеобразовательных школ при $\bar{X}_3 = 7$ составят:

\bar{X}_{1cp}	\bar{X}_{2cp}	\bar{X}_{3cp}	\bar{X}_{5cp}
0,99	15,94	7	25,98

Средние значения показателей информатизации основных общеобразовательных школ $\bar{X}_3 = 7$ составят:

$\bar{X}_{1осн}$	$\bar{X}_{2осн}$	$\bar{X}_{3осн}$	$\bar{X}_{5осн}$
0,79	10,60	7	14,90

Подставим значения из этих таблиц в модель. Тогда для средних школ:

$$\hat{Y} = 21,10 + 5,38 \cdot 0,99 + 0,48 \cdot 15,94 + 0,87 \cdot 7,00 + 0,18 \cdot 25,98 = 45,38;$$

для основных школ:

$$\hat{Y} = 21,10 + 5,38 \cdot 0,79 + 0,48 \cdot 10,60 + 0,87 \cdot 7,00 + 0,18 \cdot 14,90 = 39,21.$$

Видим, что разница между уровнями информатизации средних и основных школ существенно сократилась.

Оценим уровень информатизации общеобразовательных школ Троицкого района в предположении, что средние значения факторов увеличатся в два раза, а значение фактора X_3 будет равно 7 (рассматривать простое среднее значение фактора X_3 нецелесообразно, так как на время сбора статистических данных ни одна из основных школ не имела доступа к Интернет). Для этого составим в принятых на априорном этапе обозначениях таблицу средних значений показателей информатизации:

\bar{X}_1	\bar{X}_2	\bar{X}_3	\bar{X}_5
1,816497	27,60714	7,00	43,47857

$$\hat{Y} = 21,10 + 5,38 \cdot 1,82 + 0,48 \cdot 27,60 + 0,87 \cdot 7,00 + 0,18 \cdot 43,48 = 58,06.$$

Таким образом, для выхода школ Троицкого района в среднем на 4 уровень информатизации (Transformative) матрицы ВЕСТА необходимо увеличить показатели информатизации в среднем в 2 раза и активно использовать Интернет в образовательной деятельности. К этому показателю планируется прийти к середине следующего года.

3.4 Численное решение задачи с помощью пакета Econometric Views

Econometric Views – специальный пакет для решения эконометрических задач, разработанный компанией Quantitative Micro Software. E Views обеспечивает особо сложный и тонкий инструментарий обработки данных, позволяет выполнять регрессионный анализ, строить прогнозы в Windows-ориентированной компьютерной среде. С помощью этого программного средства можно очень быстро выявить наличие статистической зависимости в анализируемых данных и затем, используя полученные взаимосвязи, сделать прогноз изучаемых показателей. E Views очень легок в освоении благодаря простому построению интерфейса, поэтому именно его эконометристы используют наиболее часто. Этот пакет содержит развитую подсказку, являющуюся справочником по эконометрическим методам.

В исследовании представлено численное решение задачи исследования в Econometric Views 3.1.

Первым этапом эконометрического моделирования является анализ выборочных данных, которые предварительно вводятся в рабочий файл Workfile. Там они хранятся в виде последовательностей чисел, каждая из которых имеет свое имя. Для создания рабочего файла необходимо заполнить диалоговое окно Workfile Range (см. рис. 3.2).

В пакете допускается восемь типов данных. Это могут быть:

- Годовые (Annual);
- Полугодовые (Semi-annual);
- Квартальные (Quarterly);
- Ежемесячные (Monthly);
- Недельные (Weekly) и дневные (5/7 day weeks);
- Недатированные или нерегулярные (Undated or irregular) – допускают работу с данными, не привязанными к определенным временным периодам.

У нас имеется пространственная выборка, поэтому отмечаем пункт «Undated or irregular» в случае необходимо установить двадцать восемь, т.к. наша выборка содержит 28 наблюдаемых объектов.



Рисунок 3.2 – Окно создания рабочего файла Workfile Range

Данные задачи представлены в виде таблицы MS Excel, поэтому ее необходимо импортировать в окно E Views (см. рисунок 3.3):

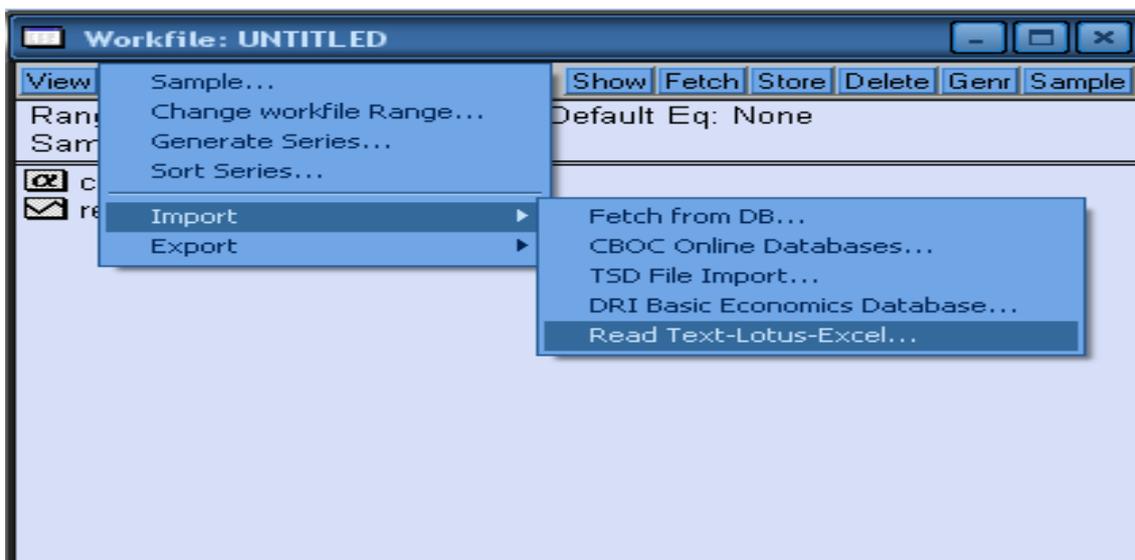


Рисунок 3.3 – Импорт данных

Вводятся соответствующие параметры для импорта данных:

- левая верхняя ячейка диапазона импортируемых данных;
- название страницы рабочей книги, на которой расположена таблица с исходными данными;
- количество переменных, также указывается, каким образом расположены исходные данные – в строках(series in columns) или столбцах(series in rows) (см. рисунок 3.4).

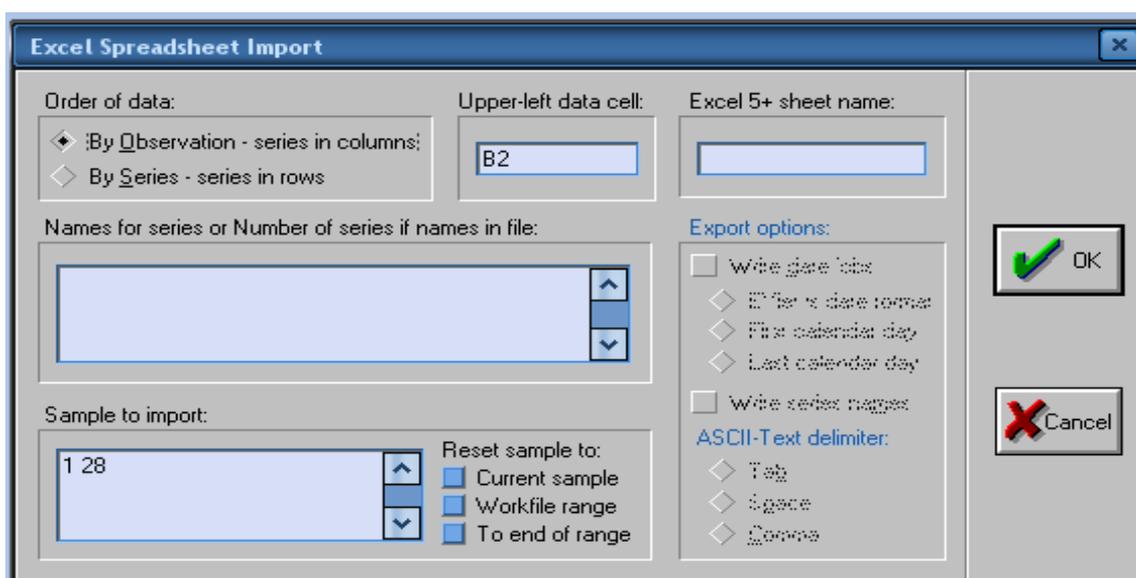


Рисунок 3.4 – Импорт таблицы MS Excel

Для просмотра числовых характеристик (описательных статистик) отмеченных переменных выбирается в рабочем файле View/Descriptive Stats/Individual Samples. В результате появляется окно, представленное на рисунке 3.5. В нем содержатся:

- Mean – Среднее арифметическое значение;
- Median – Медиана;
- Maximum – Максимальное значение;
- Minimum – Минимальное значение;
- Std. Dev. – Стандартное отклонение (среднее квадратическое отклонение);
- Skewness – Коэффициент асимметрии;
- Kurtosis – Эксцесс;
- Observations – Количество наблюдений.

	Y	X1	X2	X3	X4	X5	X6	X7	X8
Mean	38.17857	0.908248	13.80357	1.878571	9.596429	21.73929	12.90357	6.350000	0.634500
Median	38.00000	0.845692	13.30000	0.000000	10.00000	18.65000	12.60000	6.200000	0.124144
Maximum	48.00000	1.828233	20.80000	9.800000	12.80000	40.30000	21.30000	10.90000	3.116279
Minimum	27.00000	0.380952	6.000000	0.000000	5.400000	9.000000	6.400000	2.900000	0.037736
Std. Dev.	5.702849	0.349640	4.457244	2.907635	1.928247	8.898204	3.805501	2.004347	0.997398
Skewness	0.022456	1.013412	-0.090813	1.309856	-0.541704	0.604117	0.406994	0.278818	1.708221
Kurtosis	1.999546	4.163603	1.969342	3.442068	2.546749	2.197526	2.485260	2.465478	4.489613
Jarque-Bera	1.170079	6.372321	1.277784	8.234701	1.609078	2.454427	1.082123	0.696118	16.20620
Probability	0.557084	0.041330	0.527877	0.016288	0.447294	0.293108	0.582130	0.706057	0.000303
Observations	28	28	28	28	28	28	28	28	28

Рисунок 3.5 – Описательные статистики отмеченных переменных

Для оценки параметров множественной регрессии выделяем группу независимых переменных и зависимую переменную, затем открываем окно Equation Specification, где выбираем метод оценивания. В нашем случае это МНК. В верхнем поле отображен список переменных (см. рисунок 3.6).

Оценим параметры модели (см. рисунок 3.7):

$$Y = C(1)*X1 + C(2)*X2 + C(3)*X3 + C(4)*X4 + C(5)*X5 + \\ +C(6)*X6 + C(7)*X7 + C(8)*X8 + C(9) \quad (I)$$

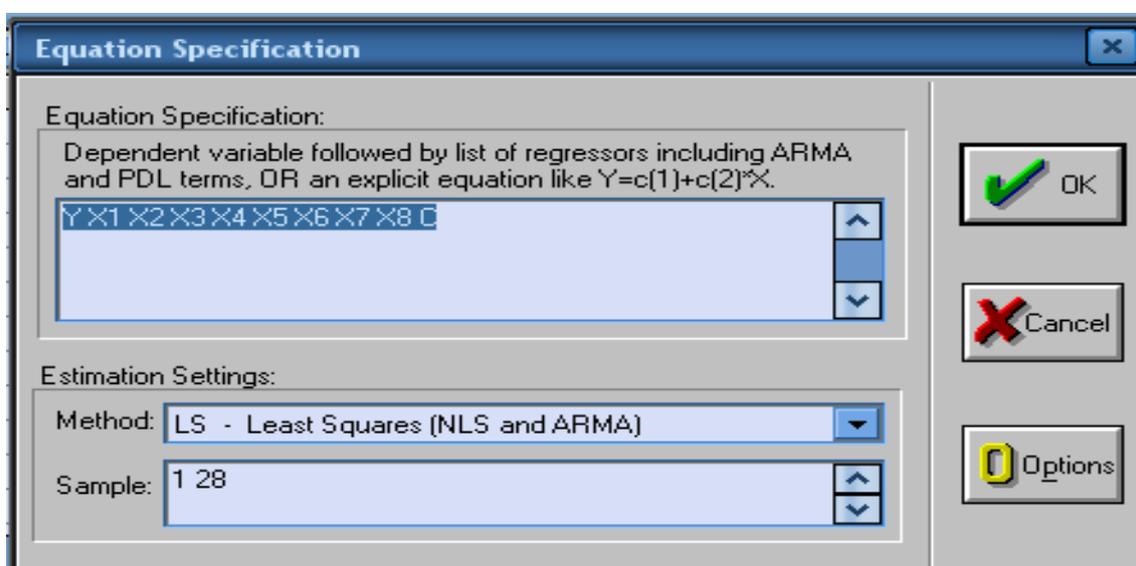


Рисунок 3.6 – Окно Equation Specification. Оценка параметров множественной линейной регрессии (I)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	6.970975	1.584397	4.399765	0.0003
X2	0.742346	0.155990	4.758938	0.0001
X3	1.216707	0.206592	5.889426	0.0000
X4	-0.881086	0.303568	-2.902431	0.0091
X5	0.211422	0.070780	2.987017	0.0076
X6	0.294325	0.145774	2.019054	0.0578
X7	-0.719685	0.401057	-1.794468	0.0887
X8	1.433129	0.553807	2.587775	0.0180
C	23.03645	2.817308	8.176758	0.0000

R-squared	0.878336	Mean dependent var	38.17857
Adjusted R-squared	0.827109	S.D. dependent var	5.702849
S.E. of regression	2.371253	Akaike info criterion	4.819806
Sum squared resid	106.8340	Schwarz criterion	5.248015
Log likelihood	-58.47728	F-statistic	17.14598
Durbin-Watson stat	2.380853	Prob(F-statistic)	0.000000

Рисунок 3.7 – Объект Equation для модели (I)

Так как коэффициенты при X_6 и X_7 статистически незначимы, принимаем решение об их исключении из модели.

Оценим параметры следующей модели (см. рисунки 3.8-3.9):

$$Y = C(1) * X1 + C(2) * X2 + C(3) * X3 + C(4) * X4 + C(5) * X5 + C(8) * X8 + C(9) \quad (II)$$

Equation Specification	
Equation Specification: Dependent variable followed by list of regressors including ARMA and PDL terms, OR an explicit equation like $Y=c(1)+c(2)*X$.	
<input type="text" value="Y X1 X2 X3 X4 X5 X8 C"/>	<input type="button" value="OK"/>
<input type="button" value="Cancel"/>	<input type="button" value="Options"/>
Estimation Settings: Method: <input type="text" value="LS - Least Squares (NLS and ARMA)"/>	
Sample: <input type="text" value="1 28"/>	

Рисунок 3.8 – Окно Equation Specification. Оценка параметров множественной линейной регрессии (II)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	6.310271	1.463068	4.313040	0.0003
X2	0.667460	0.148428	4.496847	0.0002
X3	0.997898	0.196321	5.082998	0.0000
X4	-0.802087	0.323681	-2.058478	0.0617872
X5	0.191881	0.070114	2.736718	0.0124
X8	1.193124	0.544030	2.193121	0.0397
C	24.12809	2.950294	8.178200	0.0000

R-squared	0.843614	Mean dependent var	38.17857
Adjusted R-squared	0.798933	S.D. dependent var	5.702849
S.E. of regression	2.557188	Akaike info criterion	4.928012
Sum squared resid	137.3235	Schwarz criterion	5.261063
Log likelihood	-61.99216	F-statistic	18.88055
Durbin-Watson stat	2.320608	Prob(F-statistic)	0.000000

Рисунок 3.9 – Объект Equation для модели (II)

Из рисунка 3.9 видим, что с 95% вероятностью можно утверждать, что коэффициент при X_4 статистически незначим, принимаем решение об его исключении из модели.

Оценим параметры новой модели (см. рисунки 3.10-3.11):

$$Y = C(1)*X1 + C(2)*X2 + C(3)*X3 + C(5)*X5 + C(8)*X8 + C(9) \quad (III)$$

Equation Specification	
Equation Specification: Dependent variable followed by list of regressors including ARMA and PDL terms. OR an explicit equation like $Y=c(1)+c(2)*X$.	
<input type="text" value="Y X1 X2 X3 X5 X8 C"/>	<input type="button" value="OK"/>
<input type="button" value="Cancel"/>	<input type="button" value="Options"/>
Estimation Settings: Method: <input type="text" value="LS - Least Squares (NLS and ARMA)"/>	
Sample: <input type="text" value="1 28"/>	

Рисунок 3.10 – Окно Equation Specification. Оценка параметров множественной линейной регрессии (III)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	5.602304	1.593751	3.515169	0.0020
X2	0.557024	0.157253	3.542211	0.0018
X3	0.857179	0.208733	4.106586	0.0005
X5	0.158778	0.076449	2.076922	0.0497
X8	0.868351	0.586459	1.480666	0.1529
C	19.78840	2.637134	7.503751	0.0000

R-squared	0.797886	Mean dependent var	38.17857
Adjusted R-squared	0.751951	S.D. dependent var	5.702849
S.E. of regression	2.840278	Akaike info criterion	5.113091
Sum squared resid	177.4780	Schwarz criterion	5.398563
Log likelihood	-65.58327	F-statistic	17.36986
Durbin-Watson stat	2.347614	Prob(F-statistic)	0.000001

Рисунок 3.11 – Объект Equation для модели (III)

Из рисунке 3.11 видим, что с 95% вероятностью можно утверждать, что коэффициент при X_8 статистически незначим, принимаем решение об его исключении из модели.

Оценим параметры новой модели (см. рисунках 3.12-3.13):

$$Y = C(1)*X1 + C(2)*X2 + C(3)*X3 + C(5)*X5 + C(9) \quad (IV)$$

Equation Specification

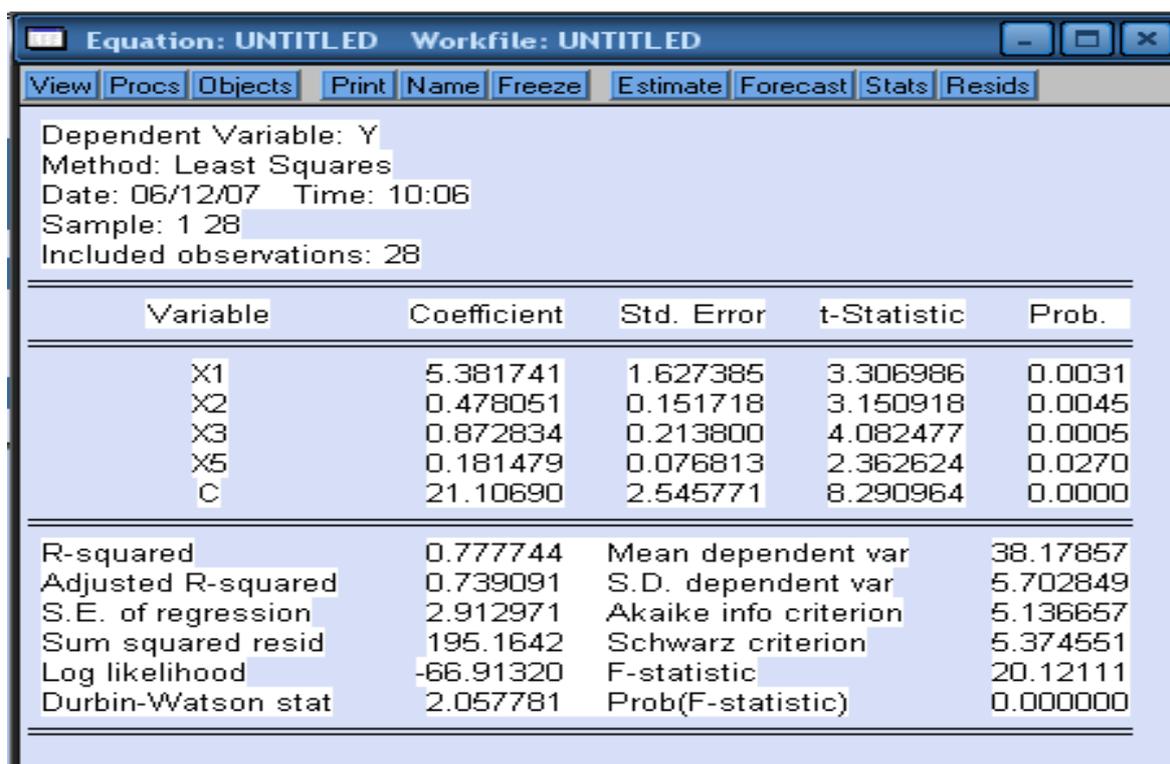
Equation Specification:
 Dependent variable followed by list of regressors including ARMA and PDL terms, OR an explicit equation like $Y=c(1)+c(2)*X$.

Y X1 X2 X3 X5 C

Estimation Settings:
 Method: LS - Least Squares (NLS and ARMA)
 Sample: 1 28

OK
 Cancel
 Options

Рисунок 3.12 – Окно Equation Specification. Оценка параметров множественной линейной регрессии (IV)

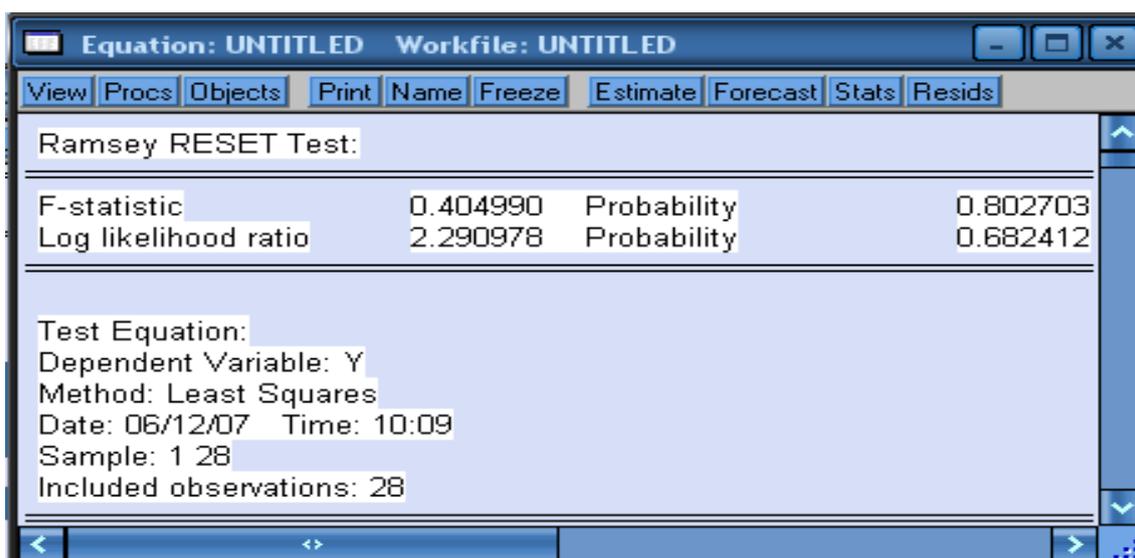


Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	5.381741	1.627385	3.306986	0.0031
X2	0.478051	0.151718	3.150918	0.0045
X3	0.872834	0.213800	4.082477	0.0005
X5	0.181479	0.076813	2.362624	0.0270
C	21.10690	2.545771	8.290964	0.0000

R-squared	0.777744	Mean dependent var	38.17857
Adjusted R-squared	0.739091	S.D. dependent var	5.702849
S.E. of regression	2.912971	Akaike info criterion	5.136657
Sum squared resid	195.1642	Schwarz criterion	5.374551
Log likelihood	-66.91320	F-statistic	20.12111
Durbin-Watson stat	2.057781	Prob(F-statistic)	0.000000

Рисунок 3.13 – Объект Equation для модели (IV)

Как видно из рисунка 3.13, МНК оценки коэффициентов модели (IV) статистически значимы с вероятностью 95%. Проведем Reset – тест модели (IV) на ошибку спецификации (см. рисунок 3.14):



Ramsey RESET Test:			
F-statistic	0.404990	Probability	0.802703
Log likelihood ratio	2.290978	Probability	0.682412

Test Equation:			
Dependent Variable:	Y		
Method:	Least Squares		
Date:	06/12/07	Time:	10:09
Sample:	1 28		
Included observations:	28		

Рисунок 3.14 – Reset – тест модели (IV) на ошибку спецификации

P – значение Reset – теста равно 0,68 (больше 0,05), что указывает на верную спецификацию модели.

Проверим ошибки на гетероскедастичность при помощи теста Уайта (см. рисунок 3.15) .

Так как значение Probability > 0,05 для всех переменных модели (IV), то гипотеза о гомоскедастичности ошибок принимается с 95% уровнем вероятности.

White Heteroskedasticity Test:				
F-statistic	1.457385	Probability	0.237094	
Obs*R-squared	10.64788	Probability	0.222460	
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Date: 06/12/07 Time: 10:16				
Sample: 1 28				
Included observations: 28				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-6.663374	24.37374	-0.273383	0.7875
X1	11.01496	24.46679	0.450200	0.6577
X1^2	-3.323991	10.82226	-0.307144	0.7621
X2	-2.497294	2.862544	-0.872403	0.3939
X2^2	0.052109	0.101680	0.512485	0.6142
X3	5.076433	2.790027	1.819492	0.0846
X3^2	-0.652076	0.330240	-1.974552	0.0630
X5	2.652239	1.421616	1.865651	0.0776
X5^2	-0.053319	0.028316	-1.883016	0.0751
R-squared	0.380282	Mean dependent var	6.970152	
Adjusted R-squared	0.119347	S.D. dependent var	9.909167	
S.E. of regression	9.299069	Akaike info criterion	7.552797	
Sum squared resid	1642.981	Schwarz criterion	7.981006	
Log likelihood	-96.73916	F-statistic	1.457385	
Durbin-Watson stat	1.900823	Prob(F-statistic)	0.237094	

Рисунок 3.15 – Тест Уайта на гетероскедастичность

Рассмотрим модель

$$Y = C(1)*X1 + C(2)*X2 + C(3)*D1 + C(4)*X4 + C(5)*X5 + \\ +C(6)*X6 + C(7)*X7 + C(8)*X8 + C(9), \quad (V)$$

где D1 – фиктивная переменная, принимающая значения 0 или 1.

Оценим параметры модели (V) (см. рисунок 3.16):

Equation: UNTITLED Workfile: UNTITLED

View Procs Objects Print Name Freeze Estimate Forecast Stats Resids

Dependent Variable: Y
Method: Least Squares
Date: 06/12/07 Time: 10:27
Sample: 1 28
Included observations: 28

Variable	Coefficient	Std. Error	t-Statistic	Prob.
X1	5.276299	2.030239	2.598856	0.0176
X2	0.706559	0.204717	3.451398	0.0027
X3	6.149836	1.769058	3.476335	0.0025
X4	-0.375808	0.378922	-0.991782	0.3338
X5	0.062024	0.102931	0.602580	0.5539
X6	0.044057	0.181084	0.243297	0.8104
X7	-0.312210	0.506985	-0.615817	0.5453
X8	1.260625	0.725313	1.738042	0.0984
C	24.08957	3.825886	6.296467	0.0000

R-squared	0.789880	Mean dependent var	38.17857
Adjusted R-squared	0.701408	S.D. dependent var	5.702849
S.E. of regression	3.116241	Akaike info criterion	5.366224
Sum squared resid	184.5082	Schwarz criterion	5.794432
Log likelihood	-66.12713	F-statistic	8.928044
Durbin-Watson stat	2.160330	Prob(F-statistic)	0.000049

Рисунок 3.16 – Объект Equation для модели (V)

Видим – большинство оценок статистически незначимы, что свидетельствует о наличии мультиколлинеарности. С помощью E Views можно получить матрицу парных коэффициентов корреляции для определения мультиколлинеарности между признаками. На рисунке 3.17 показана матрица парных коэффициентов корреляции между переменными модели (V), и результирующим признаком:

Correlation Matrix									
	Y	X1	X2	X3	X4	X5	X6	X7	X8
Y	1.000000	0.210736	0.666434	0.744786	0.366844	0.596667	0.348629	0.644961	0.016439
X1	0.210736	1.000000	-0.063841	-0.025168	0.057482	-0.156088	0.225516	0.227371	-0.105541
X2	0.666434	-0.063841	1.000000	0.468823	0.452177	0.527096	0.250559	0.551771	-0.258370
X3	0.744786	-0.025168	0.468823	1.000000	0.364546	0.666741	0.247755	0.607465	-0.005130
X4	0.366844	0.057482	0.452177	0.364546	1.000000	0.444183	0.168330	0.404737	0.120509
X5	0.596667	-0.156088	0.527096	0.666741	0.444183	1.000000	0.278139	0.561887	0.052748
X6	0.348629	0.225516	0.250559	0.247755	0.168330	0.278139	1.000000	0.462772	0.092004
X7	0.644961	0.227371	0.551771	0.607465	0.404737	0.561887	0.462772	1.000000	0.154570
X8	0.016439	-0.105541	-0.258370	-0.005130	0.120509	0.052748	0.092004	0.154570	1.000000

Рисунок 3.17 – Матрица парных коэффициентов корреляции

Парные коэффициенты корреляции между X_2 и X_7 , X_2 и X_5 , X_5 и X_7 довольно большие, хотя не превышают критического значения, но так как коэффициенты при переменных X_5 и X_7 имеют низкие t – статистики, то их необходимо исключить из модели в первую очередь, но практическая значимость переменной X_5 не позволяет сделать этого.

Поэтому делаем вывод о непригодности использования модели с количественными и качественной переменными для нашего исследования.

Итак, анализ модели, оценка параметров которой представлена на рисунке 3.13, показал, что модель (V) является статистически значимой и имеет статистически значимые коэффициенты, а потому наиболее правильно из всех построенных моделей аппроксимирует реальным данным.

На рисунке 3.18 изображен ряд прогнозных значений модели (V), представлены ошибки прогноза.

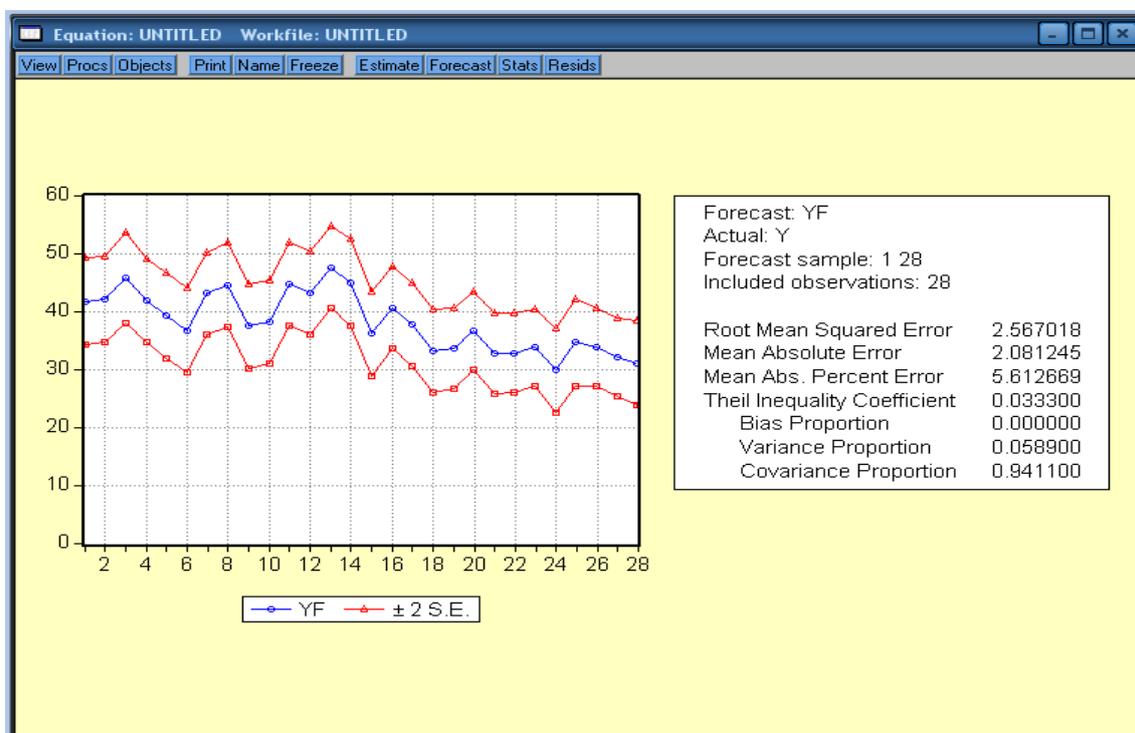


Рисунок 3.18 – Ряд прогнозных значений модели (V)

ЗАКЛЮЧЕНИЕ

В монографии показано использование эконометрических моделей в прогнозировании социальных процессов в обществе. Область эконометрических исследований в настоящее время значительно расширилась, т.к. в наполнении эмпирическим содержанием нуждаются априорные в разных областях рассуждения.

Представлена методология применения многофакторных моделей прогнозирования для анализа и прогнозирования в сфере информатизации муниципальных образовательных учреждений,

Выводы, содержащиеся в исследовании, могут быть использованы руководителями образовательных организаций в принятии эффективных управленческих решений. На основе построенного выборочного регрессионного уравнения они могут делать прогноз на будущее и определять значения факторов, которые можно изменять, имитируя различные сценарии экономического развития анализируемого показателя. Исследование показало, что можно учесть нежелательные тенденции, выявленные при прогнозировании и принять необходимые меры для их предупреждения

В ходе исследования было:

1. В соответствии с шестью уровнями информатизации матрицы ВЕСТА была составлена 100-бальная шкала для оценки состояния информатизации общеобразовательных школ Троицкого района;

2. Для проведения исследования был произведен сбор данных методом анкетирования;

3. Вопросы, содержащиеся в анкете в процессе исследования были условно поделены на восемь групп, в соответствии с которыми были определены основные факторы:

- Фактор 1 - «доступность аппаратных средств»;
- Фактор 2 - «уровень организации информационно-технологической среды школы»;
- Фактор 3 - «использование Интернета в учебном процессе»;
- Фактор 4 - «использование ЦОР в учебном процессе»;
- Фактор 5 - «ИКТ-подготовка учителей»;
- Фактор 6 - «использование ИКТ как инструмента учебной работы школьников»;
- Фактор 7 - «использование ИКТ для решения административных задач»;
- Фактор 8 - «современность аппаратных средств»;

4. Проведен анализ участвующих в модели факторов. Выявлены основные факторы, определена их роль и взаимосвязь.

В ходе исследования были построены линейная модель с количественными переменными, линейная смешанная модель с количественными переменными и качественным признаком. Проведен статистический анализ и проверка качества перечисленных моделей, которые показали, что модель множественной регрессии:

$$Y = 21,10 + 5,38X_1 + 0,48X_2 + 0,87X_3 + 0,18X_5,$$

где X_1 – доступность аппаратных средств;

X_2 – уровень организации информационно-технологической среды школы;

X_3 – использование Интернета в учебном процессе;

X_5 – ИКТ-подготовка учителей;

лучше аппроксимирует эмпирические данные.

Анализ модели показал, что

– Если факторы X_1 , X_2 , X_3 , X_5 принять равными нулю, то уровень информатизации будет равен 21,1 балла, что соответствует первому уровню информатизации школы по матрице ВЕСТА (Not Yet decided);

– При увеличении регрессоров X_1 , X_2 , X_3 , X_5 соответственно на 0,13; 0,17; 0,04; 0,1 процента уровень информатизации увеличится на 0,44 процента.

Построенная модель, отражающая состояние информатизации общеобразовательных школ Троицкого района была использована для прогнозирования уровня информатизации:

– При увеличении основных показателей в два раза и активном использовании Интернет в образовательном процессе;

– При сравнении средних уровней информатизации средних (полных) и основных школ;

– При оценке влияния использования Интернет в образовательном процессе на общий уровень информатизации.

Получены следующие результаты:

– Средний уровень информатизации полных школ составил 41,61 балла, основных – 33,12 балла;

– При активном использовании Интернет уровень информатизации при неизменных показателях составит для средних школ 45,38 балла, для основных школ – 39,21 балла – наблюдается сокращение разрыва между уровнями информатизации средних и основных школ.

– Выход школ Троицкого района в среднем на четвертый уровень информатизации (Transformative) матрицы ВЕСТА к середине следующего года возможен при увеличении к этому времени всех показателей уровня информатизации в среднем в два раза и активном использовании Интернет. Уровень информатизации общеобразовательных школ Троицкого района в этом случае составит 58,06 балла.

Для численного решения задачи использован специальный эконометрический пакет Econometric Views, который имеет большой набор статистических функций: факторный анализ, регрессионный анализ, кластерный анализ, многомерный анализ, критерии согласия и средства для визуальной интерпретации полученных результатов: различные графики, диаграммы, представление данных на географической карте.

Модель может быть применена для принятия эффективного управленческого решения [20,38]. Построенная модель призвана помочь администрации, сотрудникам и учителям рационально использовать имеющиеся ресурсы для качественного последовательного повышения уровня информатизации своей школы.

ГЛОССАРИЙ

Автокорреляция – корреляция между временной переменной и лаговой переменной, составленной от той же переменной.

Авторегрессия – регрессия зависимой временной переменной от лаговой переменной, составленной от той же переменной.

Адаптивные полиномиальные модели – адаптивные модели временного ряда, основанные на представлении уровней ряда в виде полиномов от интервала прогнозирования.

Адекватная модель — модель, для которой ряд возмущений будет удовлетворять основным предпосылкам регрессионного анализа.

Адекватность модели – соответствие построенной модели моделируемому реальному экономическому объекту или процессу.

Априорный этап – начальный этап эконометрического моделирования, на котором проводится анализ сущности изучаемого объекта, формирование и формализация априорной (известной до начала моделирования) информации.

Белый шум – стохастический процесс, представляющий собой последовательность независимых, одинаково распределенных случайных величин.

Временной ряд (динамический ряд, ряд динамики) – последовательность наблюдений некоторого признака (случайной величины) Y в последовательные моменты времени. Отдельные наблюдения называются *уровнями* ряда, которые обозначают y_t .

Временной ряд (ряд динамики, динамический ряд) – упорядоченная во времени последовательность численных

показателей $\{(y_i, t_i), i = 1, 2, \dots, n\}$, характеризующих уровни развития изучаемого явления в последовательные моменты или периоды времени.

Гетероскедастичность – неоднородность относительно дисперсии.

Гомоскедастичность – однородность относительно дисперсии.

Двухшаговый метод наименьших квадратов – применяется для оценки параметров сверхидентифицируемой системы уравнений.

Дискретная случайная величина — множество возможных значений случайной величины, число которых конечно или счетно.

Дисперсионный анализ – статистический метод для оценки влияния различных факторов на результат эксперимента, а также для последующего планирования аналогичных экспериментов.

Дисперсия $D(X)$ случайной величины X — математическое ожидание квадрата ее отклонения от математического ожидания. Дисперсия характеризует отклонение (разброс, рассеяние, вариацию) значений случайной величины относительно среднего значения.

Дисперсия – характеристика рассеяния, разброса, вариации значений случайной величины относительно среднего значения. Дисперсией случайной величины называется математическое ожидание квадрата ее отклонения от математического ожидания.

Доверительная вероятность — достоверность (надежность) определения неизвестного значения параметра с помощью оценки параметра.

Доверительный интервал (при интервальной оценке неизвестного параметра генеральной совокупности) – числовой

интервал, который с заданной доверительной вероятностью накрывает неизвестное значение параметра.

Задача эконометрики – построение экономических моделей, оценка их параметров, проверка гипотез о свойствах экономических показателей, установление видов их взаимосвязей.

Задачи корреляционного анализа: выявление статистической связи между случайными переменными; установление ее направления; количественная оценка ее тесноты (силы); отбор факторов, оказывающих наиболее существенное влияние на результативный признак, на основании измерения тесноты связи между явлениями.

Задачи регрессионного анализа — установление формы зависимости между переменными, оценка функции регрессии, оценка неизвестных значений (прогноз значений) зависимой переменной.

Задачи регрессионного анализа: установление формы зависимости между переменными (линейная или нелинейная, положительная или отрицательная и т.д.); оценка функции регрессии (в виде математического уравнения того или иного типа); оценка качества построенной модели; оценка неизвестных значений (прогноз значений) зависимой переменной.

Закон больших чисел — общий принцип, согласно которому, по формулировке академика А.Н.Колмогорова, совокупное действие большого числа случайных факторов приводит (при некоторых весьма общих условиях) к результату, почти не зависящему от случая. Другими словами, при большом числе случайных величин их средний результат перестает быть случайным и может быть предсказан с большой степенью определенности.

Закон распределения случайной величины — всякое соотношение, устанавливающее связь между возможными значениями случайной величины и соответствующими им вероятностями.

Значимость на уровне - для коэффициента уравнения регрессии это означает, что гипотезу о равенстве его нулю надо отбросить на уровне значимости.

Интервальный прогноз – интервал, в котором с определенной вероятностью находится фактическое значение прогнозируемой переменной.

Классическая нормальная линейная модель множественной регрессии — модель, в которой зависимая переменная возмущения и объясняющие переменные удовлетворяют предпосылкам регрессионного анализа и предпосылке о невырожденности матрицы значений объясняющих переменных.

Ковариация (корреляционный момент) двух случайных величин – математическое ожидание произведения отклонений этих величин от своих математических ожиданий.

Коинтеграция нестационарных временных рядов – возможность получить стационарную линейную комбинацию исходных нестационарных временных рядов.

Концепция — 1) определенный способ понимания, трактовки какого-либо предмета, явления, процесса, основная точка зрения на них, руководящая идея для их систематического освещения; 2) ведущий замысел, конструктивный принцип различных видов деятельности.

Коррелированность – наличие линейной зависимости между двумя случайными величинами. Случайные величины X и Y называются коррелированными, если коэффициент корреляции отличен от нуля и некоррелированными в противном случае. Из независимости случайных величин X и Y следует их некоррелированность (равенство нулю коэффициента корреляции), но из некоррелированности не следует их независимость, т.е. равенство нулю коэффициента корреляции указывает на отсутствие линейной связи между переменными, но не на отсутствие связи между ними вообще.

Коррелограмма — график выборочной автокорреляционной функции.

Корреляции теория — математико-статистическая теория, изучающая зависимости вариации признака от окружающих условий.

Корреляционная связь – зависимость, при которой изменение среднего значения результативного признака обусловлено изменением факторных признаков.

Корреляционная статистическая зависимость — соотношение, соответствие, зависимость между двумя случайными переменными, при этом каждому значению одной случайной переменной соответствует определенное условное математическое ожидание (среднее значение) другой.

Корреляционно-регрессионный анализ - раздел статистической науки, занимающийся исследованием причинных связей между социально-экономическими явлениями и процессами, имеющими количественное выражение.

Корреляционный анализ — раздел математической статистики, изучающий взаимную зависимость случайных величин.

Корреляция (от лат. «*correlatio*» - соотношение, взаимосвязь) - корреляционной зависимостью между двумя переменными величинами называется функциональная зависимость между значениями одной из них и условным математическим ожиданием другой.

Косвенный метод наименьших квадратов – применяется для оценки параметров идентифицируемой системы уравнений.

Косвенный метод наименьших квадратов — применяется, если разрешить систему уравнений относительно Y , так, чтобы в правых частях уравнений оставались только экзогенные переменные X , после чего к полученным уравнениям применяется обычный метод наименьших квадратов для получения оценки некоторых выражений от исходных параметров, из которых затем можно найти оценки и самих параметров.

Коэффициент авторегрессии — коэффициент корреляции между соседними возмущениями или коэффициент автокорреляции.

Коэффициент вариации — показатель относительной колеблемости признака, отношение среднего квадратического отклонения случайной величины к ее математическому ожиданию.

Коэффициент детерминации R^2 – представляет собой квадрат коэффициента корреляции, показывает, какая часть дисперсии результативного признака объяснена уравнением регрессии.

Коэффициент детерминации – одна из наиболее эффективных оценок адекватности регрессионной модели, мера качества уравнения регрессии. Коэффициент детерминации показывает, какая часть

вариации зависимой переменной обусловлена вариацией объясняющей переменной. Чем ближе коэффициент детерминации к единице, тем лучше регрессия аппроксимирует эмпирические данные. Если коэффициент детерминации равен единице, то эмпирические точки лежат на линии регрессии и между переменными X и Y существует линейная функциональная связь. Если коэффициент детерминации равен нулю, то вариация зависимой переменной полностью обусловлена воздействием неучтенных в модели переменных, и линия регрессии параллельна оси абсцисс. В случае парной линейной регрессионной модели коэффициент детерминации равен квадрату коэффициента корреляции.

Коэффициент детерминации – это показатель, который определяет долю разброса зависимой переменной Y , объясняемую регрессией Y на X .

Коэффициент корреляции двух случайных величин — величина, рассчитываемая по наблюдениям над двумя случайными величинами и характеризующая степень их связи; отношение ковариации двух случайных величин к произведению их средних квадратических отклонений.

Коэффициент корреляции двух случайных величин – отношение их ковариации к произведению средних квадратических отклонений этих величин. Коэффициент корреляции определяет тесноту линейной связи двух случайных величин. Чем ближе коэффициент корреляции по модулю к единице, тем теснее линейная связь. В случае равенства нулю коэффициента корреляции линейная корреляционная связь отсутствует и линия регрессии параллельна оси абсцисс.

Коэффициент частной эластичности — предел отношения относительного частного приращения функции к относительному приращению этой переменной.

Коэффициент эластичности - коэффициент, показывающий на сколько процентов изменяется результативный признак Y при изменении факторного признака X на один процент.

Критерий Стьюдента (t-критерий) – применяется для оценки статистической значимости коэффициентов полученного уравнения регрессии.

Критерий Фишера (F-критерий) – применяется для оценки статистической значимости полученного уравнения регрессии в целом.

Критерий Энгеля-Грэнджера – применяется для тестирования временных рядов на коинтеграцию.

Лаг – задержка. Обычно лаг рассматривают между причиной и следствием.

Линии регрессии (кривые регрессии) — графики функций регрессии, или просто регрессии Y по X и X по Y .

Линии регрессии нормально распределенных случайных величин — прямые линии, т.е. нормальные регрессии случайных величин всегда линейны.

Ложная корреляция – высокое значение коэффициентов корреляции при отсутствии истинной зависимости между переменными.

Метод Алмон – применяется для оценки коэффициентов модели с распределенным лагом. Согласно методу Алмон, коэффициенты

представляются в виде полиномов заданной степени k от величины лага j : $b_j = c_0 + c_1j + c_2j^2 + \dots + c_kj^k$.

Метод аналитического выравнивания – заключается в нахождении аналитической функции $\hat{y} = f(t)$, характеризующей основную тенденцию (тренд) изменения уровней временного ряда с течением времени.

Метод включения – применяется при отборе факторов множественной регрессии. Согласно методу включения, сначала строится уравнение регрессии с одним наиболее влияющим фактором (фактор, которому соответствует большее значение парного коэффициента корреляции с результативным признаком r_{yx_i}), затем уравнение регрессии с двумя наиболее влияющими факторами и т. д.

Метод исключения – применяется при отборе факторов множественной регрессии. Согласно методу исключения сначала строится уравнение регрессии с полным набором факторов, из числа которых затем последовательно исключаются наименее значимые факторы.

Метод Койка – применяется для оценки параметров модели с распределенным лагом. Основан на предположении, что коэффициенты при лаговых переменных убывают в геометрической прогрессии.

Метод наименьших квадратов (МНК) – метод оценки параметров уравнения регрессии, основанный на минимизации суммы квадратов отклонений расчетных значений (по уравнению регрессии) зависимой переменной от ее наблюдаемых значений.

Метод наименьших квадратов состоит в минимизации суммы квадратов остатков, т.е. $\sum \varepsilon_i^2 \rightarrow \min$.

Метод скользящих средних — метод выравнивания (сглаживания) временного ряда, т.е. выделения неслучайной составляющей. Основан на переходе от начальных значений членов ряда к их средним значениям на интервале времени, длина которого определена заранее. При этом сам выбранный интервал времени «скользит» вдоль ряда.

Метод сравнения средних – применяется для выявления наличия монотонной тенденции на основе сравнения средних уровней двух отрезков временного ряда.

Метод уменьшения мультиколлинеарности — из двух объясняющих переменных, имеющих высокий коэффициент корреляции (больше 0,8), одну переменную исключают из рассмотрения. При этом, какую переменную оставить, а какую — удалить из анализа, решают в первую очередь на основании экономических соображений. Если с экономической точки зрения ни одной из переменных нельзя отдать предпочтение, то оставляют ту из двух переменных, которая имеет больший коэффициент корреляции с зависимой переменной. Другой метод устранения или уменьшения мультиколлинеарности заключается в переходе от несмещенных оценок, определенных по методу наименьших квадратов, к смещенным оценкам, обладающим, однако, меньшим рассеянием относительно оцениваемого параметра.

Метод Фостера-Стюарта – непараметрический метод, применяемый для выявления наличия монотонной тенденции временного ряда.

Минор элемента a_{ij} — определитель матрицы $(n-1)$ -го порядка, полученной из матрицы A вычеркиванием i -и строки и j -го столбца.

Множественная регрессия – модель зависимости некоторого выходного экономического показателя (объясняемой переменной) от набора входных показателей (объясняющих переменных).

Модель адаптивных ожиданий – динамическая модель, основанная на предположении, что фактическое значение переменной y_t формируется под воздействием ожидаемого значения объясняющей переменной x_{t+1}^e в следующий момент времени.

Модель частичной корректировки – динамическая модель, основанная на предположении, что под воздействием объясняющей переменной x_t формируется не фактическое значение переменной y_t , а ее «желаемый» уровень y_t^* .

Мультиколлинеарность – множественная линейная связь.

Модель множественной регрессии – уравнение вида:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \dots + \beta_k X_{ki} + \varepsilon_i, \text{ или } Y_i = \beta_1 X_{1i} + \dots + \beta_k X_{ki} + \varepsilon_i, \text{ где } i=1, \dots, n$$

Регрессионные модели с одним уравнением – модели, в которых зависимая (объясняемая) переменная y представляется в виде функции $f(x, \beta) = f(x_1, \dots, x_k, \beta_1, \dots, \beta_p)$, где x_1, \dots, x_k – независимые (объясняющие) переменные, а β_1, \dots, β_p – параметры.

Непараметрические тесты стационарности – тесты на стационарность, не выдвигающие каких-либо предположений о законе распределения тестируемого временного ряда, его параметрах.

Оценка β^{\wedge} параметра β называется *несмещенной*, если ее математическое ожидание равно оцениваемому параметру, т.е. $M(\beta^{\wedge}) = \beta$; в противном случае оценка называется *смещенной*. Требование несмещенности гарантирует отсутствие систематических ошибок при оценивании.

Обобщенный метод наименьших квадратов – метод оценки параметров уравнения регрессии, в котором МНК применяется к уравнению регрессии, преобразованному таким образом, чтобы исключить гетероскедастичность остатков.

Параметрические тесты стационарности – тесты на стационарность, основанные на относительно строгих предположениях относительно законов распределения временного ряда, его параметров.

Показательная аналитическая зависимость – зависимость, определяемая соотношением $\hat{y} = a \cdot b^x$.

Поле корреляций – совокупность точек на координатной плоскости, изображающих наблюдения.

Предмет эконометрики – количественная оценка взаимосвязи между случайными событиями, признаками, показателями, факторами переменных экономических объектов.

Приведенная форма модели – система одновременных уравнений, которая отражает зависимость эндогенных переменных только от predetermined переменных.

Распределение(хи-квадрат) с k степенями свободы — распределение суммы квадратов k независимых случайных величин, распределенных по стандартному нормальному закону.

Регрессию уравнение — уравнение линии, вокруг которой группируются точки корреляционного поля; указывает основное направление и тенденцию связи.

Регрессионный анализ – раздел математической статистики, изучающий форму зависимости характеристик стохастического процесса от одного или нескольких факторов.

Регрессионный анализ — 1) раздел математической статистики, изучающий характер связи между случайными переменными; 2) совокупность статистических методов обработки результатов экспериментов, позволяющих в условиях стохастической зависимости выходной переменной от входных параметров определить данную зависимость.

Регрессия – зависимость среднего значения какой-либо величины (y) от некоторой другой величины или от нескольких величин (x_i).

Регрессия (Y по X) – зависимость условного математического ожидания $Mx(Y)$ случайной величины Y (при $X=x$) от x ; аналогично, *регрессия (X по Y)* – зависимость условного математического ожидания $My(Y)$ случайной величины X (при $Y=y$) от y .

Регрессия множественная – зависимость с несколькими зависимыми переменными.

Регрессия парная – зависимость между двумя переменными.

Результативный признак – признак, изменяющийся под действием факторных признаков.

Репрезентативность (представительность) выборки — достаточно полное отражение свойств генеральной совокупности. Должна удовлетворять следующим требованиям: 1) элементы генеральной совокупности выбираются случайным образом; 2) независимость результатов испытаний в выборке; 3) определенный (правильный) подбор объема выборки.

Ряд динамики — числовые значения статистического показателя, представленные во временной последовательности; ряд, расположенный в хронологической последовательности значений статистических показателей.

Случайное блуждание — стохастический процесс, определяемый соотношением $Y_t = Y_{t-1} + \varepsilon_t$.

Стационарный стохастический процесс в широком (слабом) смысле — стохастический процесс, у которого математическое ожидание μ_t и дисперсия σ_t^2 не зависят от времени (одинаковы для всех X_t), а автоковариация $\square_{t_1 t_2}$ зависит только от величины лага $\tau = t_2 - t_1$.

Стационарный стохастический процесс в узком (сильном) смысле — стохастический процесс, для которого совместное распределение вероятностей случайных величин $X_{t_1}, X_{t_2}, \dots, X_{t_n}$ такое же, как у случайных величин

Стохастическая связь — причинная зависимость, которая проявляется не в каждом отдельном случае, а в общем, среднем при большом числе наблюдений.

Стохастический — случайный, вероятностный.

Стохастический процесс — случайная функция $X(t)$ вещественного аргумента t .

Структурная форма модели — исходная система одновременных уравнений, содержащая в качестве факторов зависимые переменные.

Тенденция – изменение, определяющее общее направление развития, основную тенденцию временного ряда.

Теорема Гаусса-Маркова – теорема, определяющая условия, при соблюдении которых оценки параметров линейной регрессии, полученные методом наименьших квадратов, будут несмещенными и эффективными в классе линейных несмещенных оценок.

Тест Глейзера – применяется для проверки остатков регрессии на гетероскедастичность.

Тест Гольдфельда-Квандта – применяется для проверки остатков регрессии на гетероскедастичность. Проверяет равенство дисперсий остатков на двух интервалах изменения фактора.

Тест Дики-Фуллера – применяется для тестирования временного ряда на нестационарность, основан на оценке параметра $\lambda = \alpha_1 - 1$ уравнения $\Delta Y_t = \lambda \cdot Y_{t-1} + \varepsilon_t$. Тест Манна-Уитни – непараметрический тест стационарности применяется для тестирования постоянства математического ожидания и основан на сопоставлении рангов элементов двух совокупностей из рассматриваемого интервала.

Тест на единичный корень (тест Дики-Фуллера) – применяется для тестирования временного ряда на нестационарность, основан на оценке параметра $\lambda = \alpha_1 - 1$ уравнения $\Delta Y_t = \lambda \cdot Y_{t-1} + \varepsilon_t$.

Тест ранговой корреляции Спирмена – применяется для проверки остатков регрессии на гетероскедастичность. Проверяет наличие монотонной зависимости между дисперсией ошибки и величиной фактора.

Тест Сиджела-Тьюки – применяется для тестирования постоянства дисперсии временного ряда и основан на сопоставлении рангов элементов двух совокупностей из рассматриваемого интервала.

Точечный прогноз – среднее прогнозное значение изучаемой переменной экономического объекта.

Тренд – изменение, определяющее общее направление развития, основную тенденцию временного ряда.

Трехшаговый метод наименьших квадратов – применяется для оценки параметров системы уравнений при нарушении предпосылок о независимости и гомоскедастичности остатков системы одновременных уравнений.

Условия Гаусса-Маркова – условия, обеспечивающие состоятельность, несмещенность и эффективность оценок параметров уравнения регрессии с помощью МНК.

Оценка параметра называется *состоятельной*, если ее предел по вероятности при увеличении числа наблюдений стремится к истинному значению параметра, т.е. $\lim_{n \rightarrow \infty} \hat{\beta} = \beta$.

$\sum(Y_i - \bar{Y})^2$ - *TSS* – вся дисперсия;

$\sum(Y_i - \hat{Y}_i)^2$ - *ESS* – необъясненная часть всей дисперсии;

$\sum(Y_i^{\wedge} - Y^-)^2 - RSS$ – объясненная часть всей дисперсии.

Оценка дисперсии ошибок σ^2 . Распределение S^2 .

$$\sigma^2 = S^2 = \frac{e'e}{n-k} = \frac{\sum e_i^2}{n-k}$$

Коэффициент детерминации R^2 (доля объясненной дисперсии):

$$R^2 = 1 - (ESS \div TSS) = RSS \div TSS$$

Статистика Стьюдента

$$t_c = \frac{\beta_{ols,i}^{\wedge} - \beta_1}{S\beta_1^{\wedge}} \approx t(n-k)$$

Статистика F :

$$F = \frac{R^2}{1-R^2} \times \frac{n-k}{k-1} = \frac{RSS}{ESS} \times \frac{n-k}{k-1} = \frac{Y_*^{\wedge} Y_*^{\wedge} / (k-1)}{e'e / (n-k)} \approx F(k-1, n-k)$$

Коэффициент корреляции между двумя переменными, очищенный от влияния других переменных, называется частным коэффициентом корреляции.

Выборочное линейное уравнение парной регрессии: $\hat{y} = a + bx$.

По МНК определяются неизвестные коэффициенты a и b :

$$a = \bar{y} - b\bar{x} \quad \text{или} \quad a = \frac{1}{n} \left(\sum y_i - b \cdot \sum x_i \right)$$

$$b = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{x^2 - (\bar{x})^2} \quad \text{или} \quad b = \frac{n \sum x_i \cdot y_i - \sum x_i \cdot \sum y_i}{n \sum x_i^2 - (\sum x_i)^2}$$

Коэффициент парной корреляции:

$$r_{xy} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\sqrt{x^2 - (\bar{x})^2} \cdot \sqrt{y^2 - (\bar{y})^2}}$$

Если $r_{xy} = \pm 1$, то корреляционная связь представляет функциональную зависимость.

Если $r_{xy} = 0$, то линейная корреляционная связь отсутствует

Коэффициент эластичности (средний): $\bar{\varepsilon} = f'(x) \cdot \frac{\bar{x}}{\bar{y}}$.

Коэффициент β вычисляется по формуле: $\beta = b \cdot \sqrt{\frac{\sum (x_i - \bar{x})^2}{\sum (y_i - \bar{y})^2}}$.

Бетта-коэффициент β - коэффициент показывает: на какую часть величины своего среднего квадратичного отклонения изменится в среднем значение результативного признака y при изменении факторного признака x на величину его среднего квадратичного отклонения при фиксированном на постоянном уровне, значения остальных независимых переменных.

Стандартная ошибка оценки коэффициентов регрессии:

$$S_b = \sqrt{\frac{\sum e_i^2}{n(n-2) \cdot (x^2 - (\bar{x})^2)}} \quad S_a = \sqrt{x^2 \cdot S_b^2}$$

Оценка качества построенной модели с помощью средней ошибки аппроксимации и коэффициента детерминации:

$$\bar{E} = \frac{1}{n} \sum \left| \frac{y_i - \hat{y}_i}{y_i} \right| \cdot 100\%. \quad \text{Если } \bar{E} < 5\% - \text{ модель точная, } \bar{E} < 10\% - \text{ модель}$$

допустимо точная, $\bar{E} > 10\%$ - модель не точная.

Коэффициент детерминации: $R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$ Если $r_{xy}^2 = R^2 -$

высокое качество построенного уравнения.

Фиктивная (искусственная) переменная (D) – переменная, которая отражает 2 противоположных состояния качественного фактора.

ANOVA – модель (модели дисперсионного анализа) - регрессионная модель, содержащая лишь качественные объясняющие переменные.

ANCOVA – модель (модели ковариационного анализа) - модель, в которой объясняющие переменные носят как количественный, так и качественный характер.

Статистика Дарбина-Уотсона (DW):

$$DW = \frac{\sum_{t=2}^n (e_t - e_{t-1})^2}{\sum_{t=1}^n e_t^2}$$

Прогнозирование – построение оценки зависимой переменной.

Точечное прогнозирование: оценка – конкретное число.

Интервальное прогнозирование: оценка – интервал, в котором истинное значение переменной находится с заданным уровнем доверия.

Безусловное – прогнозирование, при котором объясняющие переменные известны точно.

Условное – прогнозирование, при котором объясняющие переменные известны приближенно.

Функция Кобба – Дугласа: $Y = AK^\alpha L^\beta$, где Y – объем выпуска; K, L – затраты капитала, труда; A, α, β – параметры модели.

Функция Энгеля: $Y = AX^\beta$, где A, β – параметры модели.

Двойная логарифмическая модель (модель постоянной эластичности): $\ln Y = \beta_0 + \beta \ln X + \varepsilon$.

Полулогарифмические модели – это модели вида:

$$\ln Y = \beta_0 + \beta X + \varepsilon, \quad Y = \beta_0 + \beta \ln X + \varepsilon.$$

Лог – линейная модель: $\text{Ln}Y_t = \beta_0 + \beta t + \varepsilon_t$.

Линейно – логарифмическая модель: $Y = \beta_0 + \beta \text{Ln}X + \varepsilon$.

Обратная модель: $Y = \beta_0 + \beta_1(1/X) + \varepsilon$.

Степенная модель: $Y = \beta_0 + \beta_1X + \beta_2X^2 + \dots + \beta_nX^n + \varepsilon$.

Показательная модель: $Y = \beta_0 e^{\beta x}$.

Эндогенные переменные - переменные, значения которых определяются внутри модели.

Факторный признак – признак, обуславливающий изменение другого, связанного с ним признака.

Фиктивные переменные – применяются для учета в модели влияния качественного фактора.

Функциональная связь – связь, при которой определенному значению факторного признака соответствует одно и только одно значение результативного признака.

Частные уравнения регрессии – характеризующие изолированное влияние одного из факторов x_i на результативную переменную y при исключении влияния остальных факторов, входящих в общее уравнение регрессии.

Экзогенные – внешние по отношению к модели переменные. Их значения определяются вне модели и поэтому их считают фиксированными.

Экзогенные (внешние, независимые) переменные – переменные, значения которых определяются вне рассматриваемой модели, для которой они являются заданными.

Эконометрика – раздел науки, изучающий конкретные количественные и качественные взаимосвязи экономических

объектов и процессов с помощью математических и статистических методов и моделей.

Экстраполяция – прогноз, получение расчетных значений при условии, что значения аргумента выходят за пределы области определения функции.

Эндогенные (внутренние, зависимые) переменные – переменные, значения которых определяются внутри рассматриваемой модели.

Этапы эконометрического исследования: 1) постановка проблемы, 2) получение данных, анализ их качества, 3) спецификация модели, 4) оценка параметров, 5) интерпретация результатов.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Айвазян, С.А. Программное обеспечение персональных ЭВМ по статистическому анализу данных. // Компьютер и экономика: экономические проблемы компьютеризации общества. М.: Наука, 2019, с.91–107.
2. Айвазян, С.А. Прикладная статистика и основы эконометрики / С.А. Айвазян, В.С. Мхитарян. - М. : ЮНИТИ, 2010. - 656 с.
3. Антохонова, И. В. Методы прогнозирования социально-экономических процессов : учеб. пособие для вузов / И. В. Антохонова. – 2-е изд., испр. и доп. – М. : Издательство Юрайт, 2018. – 213 с.
4. Бабайцев, В. А. Математические методы финансового анализа : учеб. пособие для вузов / В. А. Бабайцев, В. Б. Гисин. — 2-е изд., испр. и доп. — М. : Издательство Юрайт, 2019. — 215 с.
5. Белько, И.В. Эконометрика : практикум: учебное пособие / И.В. Белько, Е.А. Криштапович. - М. : Издательство Гревцова, 2011. - 224 с.
6. Берг, Н.А. Деловая игра как средство организации познавательной деятельности студентов в вузе / Н.А.Берг, Н.А. Дегтярева, // Вестник Уральского института экономики, управления и права. Екатеринбург. Серия: «Экономика, управление и право». - 2016. - № 4. – С. 60-67.
7. Бережная, Е.В. Математические методы моделирования экономических систем: учебное пособие / Е. В. Бережная, В.И. Бережной. - М. : Финансы и статистика, 2006. - 432 с.

8. Бородачев, С. Эконометрика / С.Бородачев. - М. : PalmariumAcademicPublishing, 2012. - 84 с.
9. Бородич, С. А. Эконометрика : практикум : учебное пособие : учебное пособие / С.А. Бородич. - М. : Инфра-М, 2016. - 336 с.
10. Булдык, Г.М. Статистическое моделирование и прогнозирование: учебник / Г.М. Булдык. - Минск : НО ООО «БИП - С», - 2013.- 562 с.
11. Векслер, Л.С. Статистический анализ на персональном компьютере. //«Мир ПК», 1992, №2, – С. 89–97.
12. Водопьян, Г.М. О построении модели процесса информатизации школы / Г.М. Водопьян, А.Ю. Уваров. - М. : РФРДПО, 2006. - 424 с.
13. Вуколов, Э.А. Основы статистического анализа : практикум по статистическим методам и исследованию операций с использованием пакетов Statistica и Excel / Э.А. Вуколов. - М. : ИНФРА - М, 2004. - 464 с.
14. Гармаш, А. Н. Экономико-математические методы и прикладные модели : учебник для бакалавриата и магистратуры / А. Н. Гармаш, И. В. Орлова, В. В. Федосеев ; под ред. В. В. Федосеева. — 4-е изд., перераб. и доп. — М. : Издательство Юрайт, 2017. – 328 с.
15. Гладилин, А.В. Практикум по эконометрике : учебное пособие / А.В. Гладилин, А.Н. Герасимов, Е.И. Громов. - Ростов-на-Дону : Феникс, 2011. - 336 с.
16. Гмурман, В. Е. Теория вероятностей и математическая статистика. Учебник. М.: Юрайт, 2018. – 480 с.

17. Гордеева, Д.С. Экономика образования : учебное пособие для студентов / Д.С. Гордеева, Н.А. Дегтярева. – Челябинск: Цицеро, 2017. – 101 с.

18. Горленко, О. А., Борбаць Н. М. Статистические методы в управлении качеством. Учебник и практикум для спо. М.: Юрайт, 2019. - 270 с.

19. Дегтярева Н.А. Исследование экономических процессов с применением сетевых моделей: монография / Н.А. Дегтярева. - Челябинск: Изд-во ЗАО «Библиотека А.Миллера», 2019.- 160 с.

20. Дегтярева, Н.А. Использование информационных технологий в управлении / Н.А. Дегтярева, И.Д. Колмакова // Вестник факультета управления Челябинского государственного университета. Серия: «Управление». – 2016.- № 1. - С. 99 - 101.

21. Дегтярева, Н.А. Конкурентоспособность регионов и ее оценка / Н.А. Дегтярева, К.О. Латышева // В книге: Конкурентоспособность и развитие социально-экономических систем. Третьей Всероссийской научной конференции памяти А.И. Татаркина. Под общей редакцией В.И. Бархатова, Д.С.Бенц, Челябинск, 2019, С.44 – 45.

22. Дегтярева, Н.А. Анализ социально-экономического развития регионов РФ на основе исследования уровня жизни населения / Н.А. Дегтярева, Д.С. Гордеева, Д.Н. Корнеев, И.И. Плужникова // Азимут научных исследований. Серия: Экономика и управление. – 2018. Т.7, №2 (23). – С.142-146.

23. Дегтярева, Н.А. Введение в экономическую теорию. Практикум по дисциплине «Экономика» : учебное пособие для

студентов вузов / Н.А. Дегтярева. – Челябинск: Изд-во ЗАО «Библиотека А. Миллера», 2020. – 75 с.

24. Дегтярева, Н.А. Информатизация общеобразовательных школ // Функциональная и прикладная наука: сборник научных статей по итогам научно-исследовательской работы за 2014 г./ под науч. ред. М.В. Потаповой.- Челябинск: Изд-во Челяб.гос.пед.ун-т. - 2015. - С. 91-94.

25. Дегтярева, Н.А. Информационные системы в экономике: лабораторный практикум для студентов: учебное пособие / Н.А. Дегтярева. – Челябинск: Цицеро, 2018. - 106 с.

26. Дегтярева, Н.А. Исследование зависимости количества безработных от социально-экономических факторов на основе модели множественной регрессии // Фундаментальная и прикладная наука. – Челябинск: Из-во Челяб.гос.пед.ун-та. - 2016. - № 2. - С. 13-17.

27. Дегтярева, Н.А. Исследование индекса развития человеческого потенциала социально-экономической системы / Н.А. Дегтярева, К.О. Латышева // Ключевые элементы развития человеческого потенциала, экономики и обеспечения экономической безопасности: сборник статей участников Международной научно-практической конференции V Уральского вернисажа науки и бизнеса / под общ. ред. Е.П. Велихова. – Челябинск: Изд-во Челяб. гос. ун-та, 16 марта 2018. - С. 104-108.

28. Дегтярева, Н.А. Макроэкономика. Практикум по дисциплине «Экономика»: учебное пособие для студентов вузов / Н.А. Дегтярева. – Челябинск: Изд-во ЗАО «Библиотека А. Миллера», 2020. – 90 с.

29. Дегтярева, Н.А. Математическая статистика : учебно-практическое пособие / Н.А. Дегтярева. – Челябинск: Изд-во Южно-Урал гос.гуман.- пед. ун-та, 2018. – 122 с.

30. Дегтярева, Н.А. Методические рекомендации по написанию, оформлению и защите курсовой работы по дисциплине «Экономико-математические методы и модели»: учебно-методическое пособие для студентов / Н.А. Дегтярева, Д.С. Гордеева – Челябинск: Изд-во ЗАО «Библиотека А. Миллера», 2021. – 45 с.

31. Дегтярева, Н.А. Методы статистики: учебное пособие для студентов вузов / Н.А. Дегтярева, Д.С. Гордеева – Челябинск: Изд-во ЗАО «Библиотека А. Миллера», 2021. – 63 с.

32. Дегтярева, Н.А. Микроэкономика. Практикум по дисциплине «Экономика» : учебное пособие для студентов вузов / Н.А. Дегтярева. – Челябинск: Изд-во ЗАО «Библиотека А. Миллера», 2020. – 77 с.

33. Дегтярева, Н.А. Модели анализа и прогнозирования на основе временных рядов: монография / Н.А. Дегтярева. - Челябинск: Изд-во ЗАО «Библиотека А.Миллера», 2018.- 160 с.

34. Дегтярева, Н.А. Модели принятия управленческих решений в сфере информатизации муниципальных образовательных учреждений / Н.А. Дегтярева, Е.М. Колмакова, И.Д. Колмакова, // Вестник Челябинского государственного университета. Серия: «Экономические науки ». - № 5 (401) 2017. вып. 57. – С. 15-23.

35. Дегтярева, Н.А. Повышение экономической эффективности функционирования крестьянских (фермерских) хозяйств в условиях рынка: диссертация на соискание ученой степени

кандидата экономических наук / Дегтярева Н.А.; Челябинская государственная агроинженерная академия. - Челябинск, 2000. - 218 с.

36. Дегтярева, Н.А. Практикум по экономико-математическим методам и моделям : учебное пособие для студентов / Н.А. Дегтярева. - Челябинск: Изд-во ЗАО «Библиотека А.Миллера», 2019.– 97 с.

37. Дегтярева, Н.А. Применение статистических методов исследования в сельском хозяйстве / Н.А. Дегтярева, Н.А.Берг // Известия высших учебных заведений. Уральский регион. – 2017. - № 1. – С.42-47.

38. Дегтярева, Н.А. Применение экономико-математических моделей для принятия управленческих решений / Н.А. Дегтярева // В сборнике: Конкурентоспособность и развитие социально-экономических систем Материалы Всероссийской научной конференции памяти академика А. И. Татаркина. Под ред. В. И. Бархатова, Д. А. Плетнёва. 2017. С. 62-64.

39. Дегтярева, Н.А. Применение экономико-математических моделей для принятия управленческих решений по повышению конкурентоспособности предприятия // Региональная конкурентоспособность и образование в контексте глобальных вызовов: сборник материалов международной научно-практической конференции –IV уральский вернисаж науки и бизнеса, 3 марта 2017.

40. Дегтярева, Н.А. Принятие управленческих решений на основе адаптивных моделей / Н.А. Дегтярева, Д.С. Гордеева // Экономика образования.– 2021, № 3(124).– С. 86-91.

41. Дегтярева, Н.А. Принятие эффективных управленческих решений на основе эконометрического прогнозирования. / Н.А. Дегтярева, Н.А. Берг // Вестник Челябинского государственного

университета. Серия: «Экономические науки». - № 4 (414) 2018. вып. 61. – С. 176-183.

42. Дегтярева, Н.А. Профессиональная мобильность в условиях VUCA - мира / Н.А.Дегтярева, Д.С. Гордеева // Национальная безопасность и молодежная политика: киберсоциализация и трансформация ценностей в VUCA-мире: материалы Международной научно-практической конференции. - Челябинск: Издательство Южно-Урал.гос.гуман.-пед.ун-та, 2021. - С.27-31.

43. Дегтярева, Н.А. Сборник задач по статистике : учебное пособие для студентов / Н.А. Дегтярева. – Челябинск: Цицеро, 2017. – 90 с.

44. Дегтярева, Н.А. Сборник задач по экономико-математическим методам и моделям : учебное пособие для студентов / Н.А. Дегтярева. – Челябинск: Цицеро, 2017. – 77 с.

45. Дегтярева, Н.А. Семинарские занятия по дисциплине «Экономика образования» : учебное пособие для студентов / Н.А. Дегтярева. - Челябинск: Изд-во ЗАО «Библиотека А.Миллера», 2019.- 70 с.

46. Дегтярева, Н.А. Эконометрическое моделирование производственной эффективности [Текст] / Н.А. Дегтярева, Д.С. Гордеева, Матвеева П.А., Матвеев В.В., Борисенко Я.М. Федосеев // Азимут научных исследований. Серия: «Экономика и управление». — 2019. — Т. 8, № 2 (27). — С.133–135.

47. Дегтярева, Н.А. Экономико-математические методы и модели. Конспекты лекций : учебное пособие для студентов / Н.А.

Дегтярева. - Челябинск: Изд-во ЗАО «Библиотека А.Миллера», 2019.- 82 с.

48. Дегтярева, Н.А. Экономические аспекты развития системы образования Российской Федерации: учебное пособие / Н.А. Дегтярева, Д.С. Гордеева – Челябинск: Цицеро, 2018. – 145 с.

49. Дегтярева, Н.А. Эконометрические модели анализа и прогнозирования: монография / Н.А. Дегтярева. - Челябинск: Цицеро, 2017. - 170 с.

50. Дубров, А.М. Последовательный анализ в статистической обработке информации. – М.: Статистика, 1976 – 160с.

51. Исследование операций в экономике: учебное пособие/ ред. Н.Ш. Кремер. – Москва: Юрайт, 2017. – 430 с.

52. Ковалев, Е. А. Теория вероятностей и математическая статистика для экономистов : учебник и практикум для бакалавриата, специалитета и магистратуры / Е. А. Ковалев, Г. А. Медведев ; под общ. ред. Г. А. Медведева. — 2-е изд., испр. и доп. — М. : Издательство Юрайт, 2019. — 284 с.

53. Козлов, А.Ю. Статистический анализ данных в MS Excel: Учебное пособие / А.Ю. Козлов, В.С. Мхитарян, В.Ф. Шишов. – М.: Инфра-М, 2018. – 80 с.

54. Колмакова, И.Д. К вопросу о применении эконометрических методов исследования в социальных процессах / И.Д. Колмакова, Н.А. Дегтярева, Е.М. Колмакова // Вестник Челябинского государственного университета. Серия: «Экономика». - 2016. вып. 55. – С. 51-57.

55. Конкурентоспособность и развитие социально-экономических систем: сб. аннот. докл. IV Междунар. науч. конф.

памяти акад. А. И. Татаркина / под ред. В. И. Бархатова, Д. А. Плетнева, О. В. Брижак, Г. П. Журавлевой. Россия, г. Челябинск, 25-26 ноября 2020 года. Челябинск : Изд-во Челяб. гос. ун-та, 2020. – 285 с.

56. Концепция информатизации образования // Информатика и образование. - 1988. - № 6. - С. 3 - 29.

57. Костюнин, В.И. Эконометрика. Учебник и практикум : учебное пособие / В.И. Костюнин. - М. : Юрайт, 2014. - 286 с.

58. Кравченко, А. И. Анализ и обработка социологических данных. Учебник. –М.: КноРус. 2020. – 498 с.

59. Красс, М. С. Математика в экономике. Базовый курс : учебник для бакалавров / М. С. Красс. — 2-е изд., испр. и доп. — М. : Издательство Юрайт, 2019. — 470 с.

60. Кремер, Н.Ш. Эконометрика : учебник для вузов / Н.Ш. Кремер, Б.А. Путко, под ред. Н.Ш.Кремера. - М. : ЮНИТИ - ДАНА, 2012. - 311 с.

61. Кулаичев, А.П. Методы и средства комплексного анализа данных: Учебное пособие / А.П. Кулаичев. – М.: Форум, 2018. – 160 с.

62. Латышева, Е. Лабораторный практикум по математической статистике и эконометрике : учебное пособие / Е. Латышева, С. Строганов. - М. : LAP LambertAcademicPublishing, 2014.- 92 с.

63. Малугин, В. А. Математический анализ для экономистов : учебник и практикум для СПО / В. А. Малугин. — 3-е изд., перераб. и доп. — М. : Издательство Юрайт, 2018. — 557 с.

64. Мардас, А.Н. Эконометрика : учебник и практикум : учебное пособие / А.Н. Мардас. - М. : Юрайт, 2016. -182 с.

65. Митичкин, С.А. Практика программирования в среде 1С:

учебник / С.А. Митичкин. - М. : КомБук, 2004. - 272 с.

66. Мхитарян, В. С. Теория планирования эксперимента и анализ статистических данных. – М.: Юрайт. 2020. 491 с.

67. Наследов, А.Д. SPSS 15: профессиональный статистический анализ данных. - СПб. : Питер, 2008. - 416 с.

68. Ниворожкина, Л.И. Статистические методы анализа данных: Учебник / Л.И. Ниворожкина, С.В. Арженовский, А.А. Рудяга. – М.: Риор, 2018. – 320 с.

69. Орлов, А.И. Эконометрика : учебник для вузов / А.И. Орлов. - 3-е изд., перераб. и доп.- М. : Издательство «Экзамен», 2004. - 576 с.

70. Панкратова, Е.В. Анализ данных в программе SPSS для начинающих социологов / Е.В. Панкратова, И.Н. Смирнова, Н.Н. Мартынова. – М.: Ленанд, 2018. – 200 с.

71. Попов, А. М. Математика для экономистов. В 2 ч. Часть 1 : учебник и практикум для СПО / А. М. Попов, В. Н. Сотников. — 2-е изд., перераб. и доп. — М. : Издательство Юрайт, 2019. — 271 с.

72. Практикум по эконометрике: Учебн. пособие / Под ред. И.И. Елисейевой. – М.: Финансы и статистика, 2013. – 192 с.

73. Ризниченко, Г. Ю. Математическое моделирование биологических процессов. Модели в биофизике и экологии : учеб. пособие для бакалавриата и магистратуры / Г. Ю. Ризниченко. — 2-е изд., перераб. и доп. — М. : Издательство Юрайт, 2019. — 181 с.

74. Салин, В. Н., Чурилова Э. Ю. Статистический анализ данных цифровой экономики в системе "Statistica". Учебно-практическое пособие. – М.: КноРус. 2019. 240 с.

75. Сидняев, Н. И. Теория планирования эксперимента и анализ статистических данных. – М.: Юрайт. 2020. 496 с.

76. Симчера, В.М. Методы многомерного анализа статистических данных / В.М. Симчера. – М.: Финансы и статистика, 2018. – 400 с.

77. Смагин, Б. И. Экономико-математические методы : учебник для академического бакалавриата / Б. И. Смагин. — 2-е изд., испр. и доп. — М. : Издательство Юрайт, 2019. — 272 с.

78. Социально-экономическая статистика : учебник для бакалавров / под ред. М.Р. Ефимовой. 2-е изд., перераб. и доп., - М. : Издательство Юрайт; ИД Юрайт, 2012.

79. Суханова, О.Н. Компьютерная обработка экономической информации. М.:БИБКОМ, 2014.

80. Теория статистики с элементами эконометрики : практикум : учебное пособие. - М. : Юрайт, 2015. - 386 с.

81. Тимофеев, В.С. Эконометрика: учебник для академического бакалавриата / В. С. Тимофеев, А. В. Фаддеенков, В. Ю. Щеколдин. – 2-е изд., перераб. и доп. — М. : Издательство Юрайт, 2018. – 328 с.

82. Тихомиров, И.П. Методы эконометрики и многомерного статистического анализа / И.П. Тихомиров, Т.М. Тихомиров, О.С. Ушмаев. - М. : Экономика, 2011. - 640 с.

83. Тюрин, Ю.Н. Анализ данных на компьютере: Учебное пособие / Ю.Н. Тюрин, А.А. Макаров; Науч. ред. В.Э. Фигурнов. – М.: ИД ФОРУМ, 2017. – 368 с.

84. Фомин, Г. П. Экономико-математические методы и модели в коммерческой деятельности : учебник для бакалавров / Г. П. Фомин. – 4-е изд., перераб. и доп. – М. : Издательство Юрайт, 2019. – 462 с.
85. Халафян, А. А., Боровиков В. П., Калайдина Г. В. Теория вероятностей, математическая статистика и анализ данных. Основы теории и практика на компьютере. Statistica. Excel. Более 150 примеров решения задач. Учебное пособие. – М.: Ленанд. 2017. 320 с.
86. Цибулькинова, В.Е. Методология и методы научного исследования : учебно-методический комплекс дисциплины / Цибулькинова В.Е.. – Москва : Московский педагогический государственный университет, 2016. – 64 с.
87. Чашкин, Ю.Р. Математическая статистика. Анализ и обработка данных: Учебное пособие / Ю.Р. Чашкин; Под ред. С.Н. Смоленский. – Рн/Д: Феникс, 2017. – 236 с.
88. Шипилина, Л.А. Методология профессионально-педагогических исследований : учебное пособие / Шипилина Л.А.. – Омск : Издательство ОмГПУ, 2018. – 282 с.
89. Эконометрика : практикум : учебное пособие / В.А. Валентинов. - М. : Дашков и Ко, 2010. - 436 с.
90. Эконометрика : учебник / В. С. Мхитарян, М. Ю. Архипова, В. А. Балаш, О. С. Балаш, Т. А. Дуброва, В. С. Сиротин. - М. : Проспект, 2010. - 384 с.
91. Эконометрика : учебник для вузов / А.И. Орлов. - М. : Издательство «Экзамен», 2004. - 576 с.
92. Эконометрика : учебно-методическое пособие / Шалабанов А.К., Роганов Д.А. - Казань : ТИСБИ, 2005. - 56 с.

93. Эконометрика: учебник для бакалавриата и магистратуры / ред. И.И. Елисеева. – Москва: Юрайт, 2016.

94. Эконометрическое прогнозирование социально-экономического развития региона / Н.А. Дегтярева, К.О. Латышева // В книге: Конкурентоспособность и развитие социально-экономических систем. Четвертой международной научной конференции памяти А.И. Татаркина. Под общей редакцией В.И. Бархатова, Д.С.Бенц, Челябинск, 2020, С.93 – 94.

95. Экономико-математические методы и прикладные модели : учебное пособие / В.В. Федосеев и др. - М. : ЮНИТИ - ДАНА, 2012. - 304 с. - Режим доступа: <http://www.iprbookshop.ru/15500>.- ЭБС «IPRbooks».

96. Экономическая теория: учебник / Г.П. Журавлева [и др.].– Электрон. текстовые данные.– Москва: Дашков и К, 2019.– 934 с.

97. Юданов, А.Ю. Микроэкономика: учебник. – М.: КНОРУС, 2011.

ПРИЛОЖЕНИЕ 1

Расчетная таблица 1

№	Y	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇	X ₈	Y - \bar{Y}	X ₁ - \bar{X}_1	X ₂ - \bar{X}_2	X ₃ - \bar{X}_3	X ₄ - \bar{X}_4	X ₅ - \bar{X}_5	X ₆ - \bar{X}_6	X ₇ - \bar{X}_7	X ₈ - \bar{X}_8	Y _x	ϵ_i	ϵ_i^2	$(\epsilon_i - \epsilon_{i-1})^2$
1	41,00	0,40	18,00	4,80	11,90	40,30	9,40	9,90	1,03	2,82	-0,51	4,20	2,92	2,30	18,56	-3,50	3,55	0,39	40,16	0,84	0,70	
2	44,00	1,80	18,30	0,00	10,20	18,40	15,90	8,30	0,10	5,82	0,89	4,50	-1,88	0,60	-3,34	3,00	1,95	-0,53	42,93	1,07	1,14	0,05
3	42,00	0,76	18,60	3,00	7,80	35,70	9,60	4,80	0,08	3,82	-0,15	4,80	1,12	-1,80	13,96	-3,30	-1,55	-0,55	45,92	-3,92	15,40	24,91
4	47,00	0,76	15,40	7,00	11,80	38,50	14,80	6,30	0,06	8,82	-0,15	1,60	5,12	2,20	16,76	1,90	-0,05	-0,58	45,94	1,06	1,12	24,85
5	43,00	0,79	11,50	9,80	9,80	18,20	6,40	7,60	1,08	4,82	-0,12	-2,30	7,92	0,20	-3,54	-6,50	1,25	0,45	42,20	0,80	0,64	0,07
6	38,00	0,78	11,60	0,00	10,30	21,20	14,30	5,30	3,08	-0,18	-0,13	-2,20	-1,88	0,70	-0,54	1,40	-1,05	2,45	37,31	0,69	0,48	0,01
7	44,00	0,79	19,00	7,00	11,40	18,40	11,50	7,00	0,07	5,82	-0,12	5,20	5,12	1,80	-3,34	-1,40	0,65	-0,57	43,48	0,52	0,27	0,03
8	45,00	1,14	19,10	7,00	12,80	18,90	11,70	7,10	0,07	6,82	0,23	5,30	5,12	3,20	-2,84	-1,20	0,75	-0,56	44,81	0,19	0,04	0,11
9	39,00	0,85	16,00	0,00	5,60	12,90	8,60	5,00	0,04	0,82	-0,06	2,20	-1,88	-4,00	-8,84	-4,30	-1,35	-0,60	37,61	1,39	1,93	1,44
10	38,00	0,80	20,00	0,00	11,70	29,40	16,40	8,20	0,09	-0,18	-0,11	6,20	-1,88	2,10	7,66	3,50	1,85	-0,54	38,42	-0,42	0,18	3,28
11	48,00	1,31	12,80	3,00	8,10	25,30	18,30	7,40	1,14	9,82	0,40	-1,00	1,12	-1,50	3,56	5,40	1,05	0,51	45,20	2,80	7,86	10,41
12	43,00	0,81	13,00	4,00	11,90	32,00	16,00	8,50	3,07	4,82	-0,10	-0,80	2,12	2,30	10,26	3,10	2,15	2,44	42,49	0,51	0,26	5,28
13	47,00	1,16	20,60	1,00	9,80	32,60	21,30	8,90	0,10	8,82	0,26	6,80	-0,88	0,20	10,86	8,40	2,55	-0,53	45,92	1,08	1,16	0,32
14	41,00	0,89	20,80	5,00	10,20	25,60	20,60	10,90	0,06	2,82	-0,02	7,00	3,12	0,60	3,86	7,70	4,55	-0,57	45,49	-4,49	20,15	30,96
15	35,00	1,83	9,30	0,00	10,40	15,00	11,60	6,10	0,12	-3,18	0,92	-4,50	-1,88	0,80	-6,74	-1,30	-0,25	-0,51	35,89	-0,89	0,80	12,93
16	36,00	1,07	11,10	1,00	9,50	33,30	15,00	7,80	0,15	-2,18	0,16	-2,70	-0,88	-0,10	11,56	2,10	1,45	-0,49	37,60	-1,60	2,57	0,50
17	40,00	0,86	14,80	0,00	10,90	29,10	10,60	7,20	2,11	1,82	-0,05	1,00	-1,88	1,30	7,36	-2,30	0,85	1,47	37,51	2,49	6,22	16,80
18	30,00	0,38	13,00	0,00	10,40	13,50	13,50	3,60	1,14	-8,18	-0,53	-0,80	-1,88	0,80	-8,24	0,60	-2,75	0,51	32,05	-2,05	4,22	20,69
19	36,00	0,84	12,50	0,00	11,00	13,40	14,90	4,30	0,13	-2,18	-0,07	-1,30	-1,88	1,40	-8,34	2,00	-2,05	-0,51	32,80	3,20	10,21	27,56
20	35,00	0,92	16,10	0,00	9,50	16,30	8,60	5,10	0,10	-3,18	0,01	2,30	-1,88	-0,10	-5,44	-4,30	-1,25	-0,54	35,46	-0,46	0,21	13,36
21	33,00	1,23	7,40	0,00	7,20	13,70	8,10	5,70	0,13	-5,18	0,33	-6,40	-1,88	-2,40	-8,04	-4,80	-0,65	-0,51	32,15	0,85	0,73	1,72
22	35,00	0,86	10,20	0,00	8,10	16,50	16,00	6,00	0,13	-3,18	-0,05	-3,60	-1,88	-1,50	-5,24	3,10	-0,35	-0,51	33,54	1,46	2,12	0,36
23	34,00	0,52	13,60	0,00	9,20	22,10	10,80	3,90	0,08	-4,18	-0,39	-0,20	-1,88	-0,40	0,36	-2,10	-2,45	-0,55	33,82	0,18	0,03	1,64
24	33,00	0,51	6,80	0,00	5,40	15,20	14,20	4,30	0,09	-5,18	-0,40	-7,00	-1,88	-4,20	-6,54	1,30	-2,05	-0,54	31,28	1,72	2,97	2,40
25	31,00	1,05	6,60	0,00	7,30	9,00	14,30	6,70	3,12	-7,18	0,14	-7,20	-1,88	-2,30	-12,74	1,40	0,35	2,48	34,58	-3,58	12,79	28,10
26	32,00	0,53	14,50	0,00	8,90	19,00	8,70	5,50	0,10	-6,18	-0,38	0,70	-1,88	-0,70	-2,74	-4,20	-0,85	-0,54	32,41	-0,41	0,17	10,03
27	32,00	0,66	9,90	0,00	7,00	9,90	10,00	3,50	0,13	-6,18	-0,25	-3,90	-1,88	-2,60	-11,84	-2,90	-2,85	-0,50	31,52	0,48	0,23	0,79
28	27,00	1,14	6,00	0,00	10,60	15,30	10,20	2,90	0,18	-11,18	0,23	-7,80	-1,88	1,00	-6,44	-2,70	-3,45	-0,46	30,50	-3,50	12,22	15,78
Σ	1069,00	25,43	386,50	52,60	268,70	608,70	361,30	177,80	17,77	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1069,00	0,00	106,83	254,36
сред.	38,18	0,91	13,80	1,88	9,60	21,74	12,90	6,35	0,63	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	38,18	0,00	3,82	9,42

Продолжение Приложения 1

Расчетная таблица 2

№	Y	X ₁	X ₂	X ₃	X ₄	X ₅	X ₈	Y - \bar{Y}	X ₁ - \bar{X}_1	X ₂ - \bar{X}_2	X ₃ - \bar{X}_3	X ₄ - \bar{X}_4	X ₅ - \bar{X}_5	X ₈ - \bar{X}_8	Y _x	ϵ_i	ϵ_i^2	$(\epsilon_i - \epsilon_{i-1})^2$
1	41,00	0,40	18,00	4,80	11,90	40,30	1,03	2,82	-0,51	4,20	2,92	2,30	18,56	0,39	42,86	-1,86	3,44	
2	44,00	1,80	18,30	0,00	10,20	18,40	0,10	5,82	0,89	4,50	-1,88	0,60	-3,34	-0,53	43,18	0,82	0,67	7,16
3	42,00	0,76	18,60	3,00	7,80	35,70	0,08	3,82	-0,15	4,80	1,12	-1,80	13,96	-0,55	45,00	-3,00	8,97	14,56
4	47,00	0,76	15,40	7,00	11,80	38,50	0,06	8,82	-0,15	1,60	5,12	2,20	16,76	-0,58	44,19	2,81	7,91	33,73
5	43,00	0,79	11,50	9,80	9,80	18,20	1,08	4,82	-0,12	-2,30	7,92	0,20	-3,54	0,45	43,51	-0,51	0,26	11,01
6	38,00	0,78	11,60	0,00	10,30	21,20	3,08	-0,18	-0,13	-2,20	-1,88	0,70	-0,54	2,45	36,28	1,72	2,96	4,95
7	44,00	0,79	19,00	7,00	11,40	18,40	0,07	5,82	-0,12	5,20	5,12	1,80	-3,34	-0,57	43,27	0,73	0,54	0,97
8	45,00	1,14	19,10	7,00	12,80	18,90	0,07	6,82	0,23	5,30	5,12	3,20	-2,84	-0,56	44,48	0,52	0,27	0,05
9	39,00	0,85	16,00	0,00	5,60	12,90	0,04	0,82	-0,06	2,20	-1,88	-4,00	-8,84	-0,60	38,19	0,81	0,65	0,08
10	38,00	0,80	20,00	0,00	11,70	29,40	0,09	-0,18	-0,11	6,20	-1,88	2,10	7,66	-0,54	38,89	-0,89	0,79	2,88
11	48,00	1,31	12,80	3,00	8,10	25,30	1,14	9,82	0,40	-1,00	1,12	-1,50	3,56	0,51	43,62	4,38	19,17	27,76
12	43,00	0,81	13,00	4,00	11,90	32,00	3,07	4,82	-0,10	-0,80	2,12	2,30	10,26	2,44	42,19	0,81	0,66	12,70
13	47,00	1,16	20,60	1,00	9,80	32,60	0,10	8,82	0,26	6,80	-0,88	0,20	10,86	-0,53	44,73	2,27	5,13	2,10
14	41,00	0,89	20,80	5,00	10,20	25,60	0,06	2,82	-0,02	7,00	3,12	0,60	3,86	-0,57	45,41	-4,41	19,48	44,61
15	35,00	1,83	9,30	0,00	10,40	15,00	0,12	-3,18	0,92	-4,50	-1,88	0,80	-6,74	-0,51	36,56	-1,56	2,42	8,17
16	36,00	1,07	11,10	1,00	9,50	33,30	0,15	-2,18	0,16	-2,70	-0,88	-0,10	11,56	-0,49	38,20	-2,20	4,86	0,42
17	40,00	0,86	14,80	0,00	10,90	29,10	2,11	1,82	-0,05	1,00	-1,88	1,30	7,36	1,47	38,77	1,23	1,51	11,79
18	30,00	0,38	13,00	0,00	10,40	13,50	1,14	-8,18	-0,53	-0,80	-1,88	0,80	-8,24	0,51	30,82	-0,82	0,67	4,20
19	36,00	0,84	12,50	0,00	11,00	13,40	0,13	-2,18	-0,07	-1,30	-1,88	1,40	-8,34	-0,51	31,69	4,31	18,59	26,34
20	35,00	0,92	16,10	0,00	9,50	16,30	0,10	-3,18	0,01	2,30	-1,88	-0,10	-5,44	-0,54	36,29	-1,29	1,66	31,34
21	33,00	1,23	7,40	0,00	7,20	13,70	0,13	-5,18	0,33	-6,40	-1,88	-2,40	-8,04	-0,51	33,86	-0,86	0,74	0,18
22	35,00	0,86	10,20	0,00	8,10	16,50	0,13	-3,18	-0,05	-3,60	-1,88	-1,50	-5,24	-0,51	33,20	1,80	3,25	7,08
23	34,00	0,52	13,60	0,00	9,20	22,10	0,08	-4,18	-0,39	-0,20	-1,88	-0,40	0,36	-0,55	33,46	0,54	0,30	1,58
24	33,00	0,51	6,80	0,00	5,40	15,20	0,09	-5,18	-0,40	-7,00	-1,88	-4,20	-6,54	-0,54	30,55	2,45	6,01	3,64
25	31,00	1,05	6,60	0,00	7,30	9,00	3,12	-7,18	0,14	-7,20	-1,88	-2,30	-12,74	2,48	34,75	-3,75	14,04	38,43
26	32,00	0,53	14,50	0,00	8,90	19,00	0,10	-6,18	-0,38	0,70	-1,88	-0,70	-2,74	-0,54	33,77	-1,77	3,14	3,90
27	32,00	0,66	9,90	0,00	7,00	9,90	0,13	-6,18	-0,25	-3,90	-1,88	-2,60	-11,84	-0,50	31,34	0,66	0,43	5,91
28	27,00	1,14	6,00	0,00	10,60	15,30	0,18	-11,18	0,23	-7,80	-1,88	1,00	-6,44	-0,46	29,97	-2,97	8,80	13,14
Σ	1069,00	25,43	386,50	52,60	268,70	608,70	17,77	0,00	0,00	0,00	0,00	0,00	0,00	0,00	1069,00	0,00	137,32	318,67

Расчетная таблица 3

№	Y	X ₁	X ₂	X ₃	X ₅	X ₈	Y - \bar{Y}	X ₁ - \bar{X}_1	X ₂ - \bar{X}_2	X ₃ - \bar{X}_3	X ₅ - \bar{X}_5	X ₈ - \bar{X}_8	Y _x	ε _i	ε _i ²	(ε _i - ε _{i-1}) ²
1	41,00	0,40	18,00	4,80	40,30	1,03	2,82	-0,51	4,20	2,92	18,56	0,39	43,45	-2,45	5,99	
2	44,00	1,80	18,30	0,00	18,40	0,10	5,82	0,89	4,50	-1,88	-3,34	-0,53	43,08	0,92	0,84	11,32
3	42,00	0,76	18,60	3,00	35,70	0,08	3,82	-0,15	4,80	1,12	13,96	-0,55	42,69	-0,69	0,48	2,59
4	47,00	0,76	15,40	7,00	38,50	0,06	8,82	-0,15	1,60	5,12	16,76	-0,58	44,80	2,20	4,86	8,39
5	43,00	0,79	11,50	9,80	18,20	1,08	4,82	-0,12	-2,30	7,92	-3,54	0,45	42,86	0,14	0,02	4,28
6	38,00	0,78	11,60	0,00	21,20	3,08	-0,18	-0,13	-2,20	-1,88	-0,54	2,45	36,66	1,34	1,78	1,44
7	44,00	0,79	19,00	7,00	18,40	0,07	5,82	-0,12	5,20	5,12	-3,34	-0,57	43,79	0,21	0,04	1,28
8	45,00	1,14	19,10	7,00	18,90	0,07	6,82	0,23	5,30	5,12	-2,84	-0,56	45,86	-0,86	0,74	1,13
9	39,00	0,85	16,00	0,00	12,90	0,04	0,82	-0,06	2,20	-1,88	-8,84	-0,60	35,54	3,46	11,99	18,67
10	38,00	0,80	20,00	0,00	29,40	0,09	-0,18	-0,11	6,20	-1,88	7,66	-0,54	40,16	-2,16	4,65	31,58
11	48,00	1,31	12,80	3,00	25,30	1,14	9,82	0,40	-1,00	1,12	3,56	0,51	41,81	6,19	38,30	69,66
12	43,00	0,81	13,00	4,00	32,00	3,07	4,82	-0,10	-0,80	2,12	10,26	2,44	42,76	0,24	0,06	35,39
13	47,00	1,16	20,60	1,00	32,60	0,10	8,82	0,26	6,80	-0,88	10,86	-0,53	43,90	3,10	9,59	8,16
14	41,00	0,89	20,80	5,00	25,60	0,06	2,82	-0,02	7,00	3,12	3,86	-0,57	44,76	-3,76	14,12	46,98
15	35,00	1,83	9,30	0,00	15,00	0,12	-3,18	0,92	-4,50	-1,88	-6,74	-0,51	37,70	-2,70	7,29	1,12
16	36,00	1,07	11,10	1,00	33,30	0,15	-2,18	0,16	-2,70	-0,88	11,56	-0,49	38,21	-2,21	4,90	0,24
17	40,00	0,86	14,80	0,00	29,10	2,11	1,82	-0,05	1,00	-1,88	7,36	1,47	39,29	0,71	0,51	8,58
18	30,00	0,38	13,00	0,00	13,50	1,14	-8,18	-0,53	-0,80	-1,88	-8,24	0,51	32,30	-2,30	5,29	9,09
19	36,00	0,84	12,50	0,00	13,40	0,13	-2,18	-0,07	-1,30	-1,88	-8,34	-0,51	33,71	2,29	5,24	21,07
20	35,00	0,92	16,10	0,00	16,30	0,10	-3,18	0,01	2,30	-1,88	-5,44	-0,54	36,57	-1,57	2,46	14,89
21	33,00	1,23	7,40	0,00	13,70	0,13	-5,18	0,33	-6,40	-1,88	-8,04	-0,51	33,11	-0,11	0,01	2,13
22	35,00	0,86	10,20	0,00	16,50	0,13	-3,18	-0,05	-3,60	-1,88	-5,24	-0,51	33,03	1,97	3,88	4,32
23	34,00	0,52	13,60	0,00	22,10	0,08	-4,18	-0,39	-0,20	-1,88	0,36	-0,55	33,87	0,13	0,02	3,37
24	33,00	0,51	6,80	0,00	15,20	0,09	-5,18	-0,40	-7,00	-1,88	-6,54	-0,54	28,90	4,10	16,83	15,74
25	31,00	1,05	6,60	0,00	9,00	3,12	-7,18	0,14	-7,20	-1,88	-12,74	2,48	33,48	-2,48	6,15	43,32
26	32,00	0,53	14,50	0,00	19,00	0,10	-6,18	-0,38	0,70	-1,88	-2,74	-0,54	33,93	-1,93	3,74	0,30
27	32,00	0,66	9,90	0,00	9,90	0,13	-6,18	-0,25	-3,90	-1,88	-11,84	-0,50	30,69	1,31	1,73	10,55
28	27,00	1,14	6,00	0,00	15,30	0,18	-11,18	0,23	-7,80	-1,88	-6,44	-0,46	32,10	-5,10	25,96	41,08
Σ	1069,00	25,43	386,50	52,60	608,70	17,77	0,00	0,00	0,00	0,00	0,00	0,00	1069,00	0,00	177,48	416,65
сред.	38,18	0,91	13,80	1,88	21,74	0,63	0,00	0,00	0,00	0,00	0,00	0,00	38,18	0,00	6,34	15,43

Расчетная таблица 4

№	Y	X ₁	X ₂	X ₃	X ₅	Y- \bar{Y}	X ₁ - \bar{X}_1	X ₂ - \bar{X}_2	X ₃ - \bar{X}_3	X ₅ - \bar{X}_5	Y _x	ϵ_i	ϵ_i^2	$(\epsilon_i - \epsilon_{i-1})^2$
1	41,00	0,40	18,00	4,80	40,30	2,82	-0,51	4,20	2,92	18,56	43,35	-2,35	5,54	
2	44,00	1,80	18,30	0,00	18,40	5,82	0,89	4,50	-1,88	-3,34	42,89	1,11	1,23	12,01
3	42,00	0,76	18,60	3,00	35,70	3,82	-0,15	4,80	1,12	13,96	43,16	-1,16	1,35	5,17
4	47,00	0,76	15,40	7,00	38,50	8,82	-0,15	1,60	5,12	16,76	45,66	1,34	1,78	6,24
5	43,00	0,79	11,50	9,80	18,20	4,82	-0,12	-2,30	7,92	-3,54	42,73	0,27	0,07	1,13
6	38,00	0,78	11,60	0,00	21,20	-0,18	-0,13	-2,20	-1,88	-0,54	34,70	3,30	10,90	9,18
7	44,00	0,79	19,00	7,00	18,40	5,82	-0,12	5,20	5,12	-3,34	43,90	0,10	0,01	10,28
8	45,00	1,14	19,10	7,00	18,90	6,82	0,23	5,30	5,12	-2,84	45,89	-0,89	0,80	0,97
9	39,00	0,85	16,00	0,00	12,90	0,82	-0,06	2,20	-1,88	-8,84	35,67	3,33	11,12	17,87
10	38,00	0,80	20,00	0,00	29,40	-0,18	-0,11	6,20	-1,88	7,66	40,31	-2,31	5,33	31,85
11	48,00	1,31	12,80	3,00	25,30	9,82	0,40	-1,00	1,12	3,56	41,46	6,54	42,74	78,26
12	43,00	0,81	13,00	4,00	32,00	4,82	-0,10	-0,80	2,12	10,26	41,00	2,00	4,02	20,55
13	47,00	1,16	20,60	1,00	32,60	8,82	0,26	6,80	-0,88	10,86	44,00	3,00	8,97	0,98
14	41,00	0,89	20,80	5,00	25,60	2,82	-0,02	7,00	3,12	3,86	44,84	-3,84	14,77	46,77
15	35,00	1,83	9,30	0,00	15,00	-3,18	0,92	-4,50	-1,88	-6,74	38,11	-3,11	9,70	0,53
16	36,00	1,07	11,10	1,00	33,30	-2,18	0,16	-2,70	-0,88	11,56	39,06	-3,06	9,39	0,00
17	40,00	0,86	14,80	0,00	29,10	1,82	-0,05	1,00	-1,88	7,36	38,08	1,92	3,70	24,87
18	30,00	0,38	13,00	0,00	13,50	-8,18	-0,53	-0,80	-1,88	-8,24	31,82	-1,82	3,32	14,02
19	36,00	0,84	12,50	0,00	13,40	-2,18	-0,07	-1,30	-1,88	-8,34	34,05	1,95	3,81	14,24
20	35,00	0,92	16,10	0,00	16,30	-3,18	0,01	2,30	-1,88	-5,44	36,70	-1,70	2,89	13,33
21	33,00	1,23	7,40	0,00	13,70	-5,18	0,33	-6,40	-1,88	-8,04	33,77	-0,77	0,60	0,86
22	35,00	0,86	10,20	0,00	16,50	-3,18	-0,05	-3,60	-1,88	-5,24	33,62	1,38	1,91	4,64
23	34,00	0,52	13,60	0,00	22,10	-4,18	-0,39	-0,20	-1,88	0,36	34,43	-0,43	0,18	3,26
24	33,00	0,51	6,80	0,00	15,20	-5,18	-0,40	-7,00	-1,88	-6,54	29,83	3,17	10,02	12,90
25	31,00	1,05	6,60	0,00	9,00	-7,18	0,14	-7,20	-1,88	-12,74	31,54	-0,54	0,30	13,76
26	32,00	0,53	14,50	0,00	19,00	-6,18	-0,38	0,70	-1,88	-2,74	34,34	-2,34	5,47	3,22
27	32,00	0,66	9,90	0,00	9,90	-6,18	-0,25	-3,90	-1,88	-11,84	31,19	0,81	0,66	9,92
28	27,00	1,14	6,00	0,00	15,30	-11,18	0,23	-7,80	-1,88	-6,44	32,88	-5,88	34,59	44,79
Σ	1069,00	25,43	386,50	52,60	608,70	0,00	0,00	0,00	0,00	0,00	1069,00	0,00	195,16	401,61
сред.	38,18	0,91	13,80	1,88	21,74	0,00	0,00	0,00	0,00	0,00	38,18	0,00	6,97	14,87

ПРИЛОЖЕНИЕ 2

Информация об общеобразовательных школах Троицкого района

№	Наименование ОУ	Код школы	Ф.И.О. директора	Адрес
1.	Бобровская СОШ	74320_s_001	Воробьева Ксения Павловна	457134, с. Бобровка, IV кв-л, 1а
2.	Белозерская СОШ	74320_s_002	Мудрак Надежда Петровна	457130, п. Белозеры, ул. Центральная, 28
3.	Берлинская СОШ	74320_s_003	Котубаева Зауреш Николаевна	457130, п. Берлин, ул. Школьная, 27
4.	Дробышевская СОШ	74320_s_004	Асоскова Валентина Александровна	457143, с. Дробышево, ул. Школьная, 7а
5.	Каменнореченская СОШ	74320_s_005	Бакайкин Александр Петрович	457138, п. Каменная Речка, ул. Центральная
6.	Карсинская СОШ	74320_s_006	Малявкина Елена Алексеевна	457106, с. Карсы, ул. Восточная, 81
7.	Ключевская СОШ	74320_s_007	Литневский Андрей Леонидович	457136, с. Ключевка, ул. Школьная, 11
8.	Клястицкая СОШ	74320_s_008	Вдовина Елена Александровна	457144, с. Клястицкое, ул. Школьная, 14
9.	Нижнесанарская СОШ	74320_s_009	Дергунова Лариса Геннадьевна	457120, с. Нижняя Санарка, ул. Больничная, 25
10.	Новомирская СОШ	74320_s_010	Астафьева Елена Викторовна	457141, п. Новый Мир, ул. Южная, 19
11.	Песчановская СОШ	74320_s_011	Шумкина Лариса Владимировна	457131, с. Песчаное, ул. Советская, 2
12.	Родниковская СОШ	74320_s_012	Мороз Любовь Андреевна	457123, п. Родники, ул. Просвещения, 2
13.	Скалистская СОШ	74320_s_013	Тележкина Регина Антоновна	457124, п. Скалистый, ул. Школьная, 7
14.	Целинная СОШ	74320_s_014	Смирнова Валентина Ивановна	457122, п. Целинный, ул. Садовая, 19
15.	Чернореченская СОШ	74320_s_015	Давыдов Александр Павлович	457125, п. Черноречье, ул. Мира, 3
16.	Шантаринская СОШ	74320_s_016	Зырянова Валентина Анатольевна	457137, п. Шантарино, ул. Школьная, 1а
17.	Яснополянская СОШ	74320_s_017	Злаказова Ирина Яковлевна	457145, п. Ясные Поляны, ул. Школьная, 2
1.	Белокаменная ООШ	74320_s_018	Гурбанова Софья Искандаровна	457111, п. Белокаменка, ул. Лассалья, 36
2.	Бурханкульская ООШ	74320_s_019	Тараскина Лидия Сергеевна	457133, д. Бурханкуль, ул. Центральная, 40
3.	Кадомцевская ООШ	74320_s_020	Тарасова Анита Валерьевна	457107, п. Кадомцево, ул. Рабочая, 12
4.	Каменносанарская ООШ	74320_s_021	Шишкова Лариса Николаевна	457138, п. Каменная Санарка, ул. Центральная
5.	Каракульская ООШ	74320_s_022	Масленникова Татьяна Александровна	457129, п. Каракулька, ул. Дорожная, 1а
6.	Кварцитная ООШ	74320_s_023	Брюханова Надежда Васильевна	457135, п. Кварцитный
7.	Кособродская ООШ	74320_s_024	Полева Нина Михайловна	457123, п. Кособродка, ул. Красногвардейская
8.	Кумысненская ООШ	74320_s_025	Боржеева Людмила Ивановна	457104, п. Кумысное, ул. Кольцевая, 6а
9.	Морозкинская ООШ	74320_s_026	Васильева Ирина Ивановна	457118, п. Морозкино, ул. Школьная, 47
10.	Подгорненская ООШ	74320_s_027	Койнова Татьяна Васильевна	457124, с. Подгорное, ул. Набережная, 3
11.	Полесская ООШ	74320_s_028	Арискина Марина Васильевна	457113, п. Полесье, ул. Заречная, 3

ПРИЛОЖЕНИЕ 3
Классификация вопросов Анкеты информатизации

Укажите, пожалуйста, количество рабочих мест администрации, оснащенных компьютерами	7 фактор	3	3	1	2	1	1	1	3	2	1	2	3	6	4	2	4	3	0	1	1	2	1	0	0	1	0	0	1
Оцените, пожалуйста, количество устройств для ввода графической информации, которые используются в школе	1 фактор	1	3	1	1	1	4	2	2	2	3	2	3	2	3	1	3	11	0	1	2	1	0	1	0	1	1	1	0
Какова ситуация в школе с цифровыми учебными инструментами?	1 фактор	1	0	0	0	1	3	0	0	0	0	1	3	0	0	0	0	2	1	0	0	0	0	0	0	3	0	0	0
Какое из предложенных утверждений лучшим образом характеризует использование принтеров в школе?	1 фактор	4	4	1	5	3	1	5	3	5	5	6	4	4	3	1	5	6	0	1	3	3	1	1	1	1	3	1	1
Какое из предложенных утверждений лучшим образом характеризует использование в школе компьютерных проекционных устройств?	1 фактор	3	3	1	3	1	3	1	2	3	1	3	1	1	3	2	3	3	1	1	2	2	1	0	0	0	2	0	2
Опишите подключение школьных компьютеров к локальной сети школы.	2 фактор	1	1	1	1	1	1	1	2	0	1	0	0	1	1	1	0	1	0	0	0	0	0	1	0	0	0	0	0
Опишите распространенность локальной сети по школе.	2 фактор	0	1	1	1	0	0	1	2	0	1	0	0	1	1	1	0	1	0	0	0	0	0	1	0	0	0	0	0
Какое из предложенных утверждений лучшим образом характеризует подключение школы к Интернет?	3 фактор	1	0	1	5	5	0	5	5	0	0	1	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
Кто отвечает за информатизацию в школе?	2 фактор	2	3	1	2	2	3	2	2	1	0	1	1	2	3	1	2	2	1	1	3	1	3	2	1	1	2	1	1
Как организована в школе техническая поддержка использования ИКТ?	2 фактор	3	1	1	0	2	1	1	1	1	3	1	2	2	1	0	2	0	1	1	1	1	1	1	1	0	0	1	0
Как организована в школе методическая поддержка использования ИКТ учителями?	2 фактор	5	4	4	2	2	2	5	5	5	5	2	2	5	5	2	2	2	5	3	4	2	2	2	1	1	5	4	1

Продолжение Приложения 3

Какие из предложенных утверждений наилучшим образом характеризуют реально выполняемые регламенты (правила, приказы и т.п.) использования ИКТ в Вашей школе.	2 факт ор	6	6	6	6	2	1	5	4	6	7	5	5	7	7	1	1	6	2	5	5	1	1	4	1	1	5	1	1
Какие из предложенных утверждений наилучшим образом характеризуют действия администрации, направленные на поддержку использования ИКТ педагогами школы.	2 факт ор	0	1	2	2	1	2	2	2	1	1	2	2	1	1	1	2	1	2	0	1	1	2	1	1	1	1	1	1
Какое из предложенных утверждений лучшим образом характеризует обеспеченность школы цифровыми учебными материалами для проведения учебных занятий?	4 факт ор	4	4	4	4	2	3	4	4	2	2	2	4	4	4	1	2	4	1	3	3	3	3	2	0	1	3	1	3
Оцените, пожалуйста, количество работников школы, которые уверенно и регулярно (не реже одного раза в неделю) используют ИКТ в своей профессиональной деятельности	5 факт ор	30	12	24	26	8	13	10	10	6	15	15	21	23	15	5	19	16	4	4	6	9	10	10	10	2	10	5	8
Оцените, пожалуйста, количество учителей школы, которые в своей педагогической практике способны реализовать нижеперечисленные профессиональные задачи.	5 факт ор	7,3	2,8	5	4	1,5	3,3	3,6	4,3	3,3	4,4	2,8	3,2	2,4	2	2,1	4,7	5,1	2,8	3,4	3	1,9	4	5,2	2,9	2,1	2	3	1,9
Оцените, пожалуйста, использование педагогами в повседневной практике школы цифровых инструментов и технологии?	4 факт ор	5,1	3,8	0,2	4,8	4,1	4,1	3,4	6,2	1,2	5,4	3,2	5,4	2,9	3,5	6	4,3	2,8	3,2	3,9	4	2,7	2	3,9	3,6	2,9	5	4	3,9

Научное издание

НИНА АДАМОВНА ДЕГТЯРЕВА

**ИССЛЕДОВАНИЕ ИНФОРМАТИЗАЦИИ ОБРАЗОВАНИЯ
С ПРИМЕНЕНИЕМ ЭКОНОМЕТРИЧЕСКИХ МОДЕЛЕЙ**

Монография

Изд-во ЗАО «Библиотека А.Миллера»,
454080, г. Челябинск, Свердловский пр., 60

Подписано к печати 30.11.2021
Формат 60x84 1/16 Объем 10 уч-изд.л.
Заказ № 537 Тираж 500 экз
Отпечатано на ризографе в типографии ФГБОУ ВО ЮУрГГПУ
454080, г. Челябинск, пр. Ленина, 69