

МИНОБРНАУКИ РОССИИ



Федеральное государственное бюджетное образовательное учреждение
высшего образования

"Российский государственный гуманитарный университет"
(ФГБОУ ВО "РГГУ")

ИНСТИТУТ ИНФОРМАЦИОННЫХ НАУК И ТЕХНОЛОГИЙ БЕЗОПАСНОСТИ
ФАКУЛЬТЕТ ИНФОРМАЦИОННЫХ СИСТЕМ И БЕЗОПАСНОСТИ
КАФЕДРА ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ И СИСТЕМ

Стандартизация в области информационных технологий

ПРАКТИКУМ ПО ДИСЦИПЛИНЕ

Направление подготовки Прикладная информатика 09.03.03

Направленность (профиль): Прикладная информатика в гуманитарной сфере

Уровень квалификации выпускника бакалавр

Форма обучения очная

Москва 2021

Стандартизация в области информационных технологий
Практикум

Авторы-составители:

кандидат технических наук, доцент кафедры информационных технологий и систем *Г.Н. Исаев*,

кандидат технических наук, доцент, заведующий кафедрой информационных технологий и систем *А.А. Роганов*

Ответственный редактор

кандидат технических наук, доцент, заведующий кафедрой информационных технологий и систем *А.А. Роганов*

Утверждено на заседании кафедры информационных технологий и систем
28.06.2021 протокол № 12

1. Цель практических занятий

Целью практических работ (далее Практикум) по дисциплине «Стандартизация в области информационных технологий» является:

1. Закрепление теоретических сведений, полученных в рамках лекционного курса.
2. Получение практических навыков в решении задач стандартизации в области информационных технологий с применением соответствующих методических и программных средств.

2. Задачи Практикума

Задачами Практикума по дисциплине являются:

1. Освоение современных методов решения задач по стандартизации в области информационных технологий (ИТ).
2. Освоение новых программных продуктов в решении задач стандартизации ИТ в области прикладной информатики.
3. Повышение профессионального уровня в решении задач будущего бакалавра – информатика.

3. Методика и средства Практикума

1. Методика каждого практического занятия определяется моделью соответствующей задачи, которая выполняется студентом на занятии по заданию преподавателя.

2. Средствами проведения практического занятия являются:

- Комплект персональных ЭВМ компьютерных классов Института информационных наук и технологий безопасности (ИИНТБ) РГГУ ;
- Комплекс программного обеспечения:
 - операционная система MS Windows версии XP и выше;
 - текстовый редактор MS Word;
 - программный комплекс прикладного статистического анализа данных «Универсальная статистическая диалоговая система «Stadia» (далее Stadia).

Процедурным обеспечением практических работ является:

- программный комплекс «Stadia»,
- сборник заданий студентам по данной дисциплине
- настоящие Методические указания Практического практикума.

Программный комплекс «Stadia». В общем комплексе работ по информационным технологиям значительное место занимают системы автоматизации и ускорения производственных процессов на базе рационализации принятия стратегических и оперативных решений.

Программный продукт «Stadia» является современной технологией моделирования объектов широкого класса и назначения, в том числе и анализа свойств, параметров, показателей информационных технологий (ИТ). На основе аналитических расчетов значений параметров и показателей можно получить достоверные и полные данные для решения задач унификации и стандартизации ИТ и их отдельных компонентов - по аппаратному, программному, информационному, технологическому и организационно-правовому комплексу.

В рамках решения задач анализа и развития ИТ решаются задачи принятия решений стратегического, тактического и оперативного уровней для организаций малого, среднего и большого частного бизнеса и государственных организаций, занятых в той или иной мере в информационной сфере.

Этот программный комплекс является разработкой российских ученых – института Механики МГУ. «Stadia» - это профессиональное компьютерное рабочее место для ежедневной работы специалистов и руководителей фирм, подразделений, отделов планирования, отделов

маркетинга, руководителей научных разработок и проектов, служб сбыта, консультантов по экономическим вопросам и вопросам бизнеса. Главная задача, решаемая с помощью программного продукта «Stadia» - создание вокруг управленческого состава максимально созидательной, информационно-логической среды, удобной для эффективного анализа и последующего быстрого принятия решений на всех уровнях управления. Созданная на основе применения технологии сквозной логики процессов «Stadia» обеспечивает:

- разработку математических и экспериментальных (физических) моделей ИТ в решении задач унификации и стандартизации ИТ;
- проведение анализа ИТ и её компонентов;
- принятие решений по развитию ИТ на основе моделей;
- разработку профессиональных стратегий ИТ;
- принятие комплексных решений по развитию ИТ;
- организацию логики достижения целей, обеспечивая руководителей средством борьбы с проявлениями внутриорганизационного хаоса и инструментом созидания при построении ИТ;
- сокращение времени достижения поставленных целей в несколько раз наряду с улучшением качества этой работы в области изучения информационных объектов;
- максимизацию роста и развития информационного менеджмента организации.

Одно из основных достоинств программы заключается в следующем. Концепция комплекса «Stadia» построена на логических основаниях исследования объектов широкого класса и назначения. Так, например, учитывается дилемма: почему одни информационные системы развиваются и растут, а другие стоят на месте без развития и постепенно умирают. Пользователь может сформулировать проблему, определить возможные альтернативные варианты решения, проанализировать их и выбрать наиболее рациональный. Кроме того, концепция учитывает системные свойства объектов управления, в частности, их декомпозицию. Поэтому программа в соответствии с логическим основанием большой вопрос разбивает на более мелкие, мелкие разбивает на элементарные и элементарные окружает информационной логической средой. И, решая маленькие задачи, пользователь просто решает большие сложные вопросы. Именно так комплекс «Stadia» сложное делает простым.

В программе может работать как один человек, так и группа пользователей организации. Программа построена по такому принципу, что в ней можно решать целые комплексы вопросов, а также и отдельные вопросы, которые важны для пользователя. Например, если программой пользуется вся организация, то работа строится по следующему принципу: сотрудники различных отделов собирают информацию для начальников своих отделов. Затем руководители отделов анализируют собранную информацию и готовят итоговые документы. Далее эта информация концентрируется в блоке отчетности перед высшим руководством, которым и отсылается нажатием одной кнопки. Если же с программой работает один человек, он может решать любые вопросы по любому из блоков программы, которые важны ему в данный момент.

Программа обладает набором и других интересных достоинств, в частности:

- комплексность;
- использование всех возможностей ОС Windows и Microsoft Office;
- использование программы в обучающих целях по дисциплине «Стандартизация в области информационных технологий», а также по другим дисциплинам;
- использование библиотеки с массой примеров, способствующей глубокому рассмотрению вопросов по этим предметам.

В результате - каждое из этих преимуществ позволяет быстро и эффективно реализовать цели исследовательского и проектного секторов и добиться получения максимальных результатов в развитии ИТ. Каким же образом исследования и разработки, построенные на базе программного комплекса «Stadia», обеспечивают реализацию задач развития, являющихся центральными для каждой фирмы? Укажем здесь в качестве примеров несколько блоков задач:

1. Анализ информационной ситуации.

2. Изучение статистических параметров и показателей информационных объектов, в нашем случае - ИТ.
3. Решение математических моделей на основе исходных данных о состоянии информационного объекта.
4. Решение задач корреляционного анализа изучаемых информационных объектов.
5. Регрессионный анализ производительности информационных систем.
6. Получение статистических оценок изучаемых объектов управления, в том числе и ИТ.
7. Устранение неоднородности статистической структуры дефектов обработки данных в технологии информационных систем (ИС) методом кластер-анализа.
8. Изучение эффективности модернизации информационного ресурса.
9. Синтез технологического процесса обработки данных ИС и др.

Это далеко не полный перечень задач по развитию ИТ, которые потенциально может обеспечить комплекс Stadia в направлении стандартизации ИТ.

Программные модули расчетов базируются на соответствующих моделях теории вероятностей и математической статистики. Структура этих моделей соответствует структуре вышеприведенных задач, выполняемых комплексом «Stadia». Институт Механики МГУ постоянно пополняет состав этих моделей, развивая тем самым функциональные возможности программного комплекса.

Значение функциональных возможностей информационных технологии в области менеджмента на базе комплекса «Stadia» существенно увеличивается за счет реализации мультимедийного отображения результатов решения задач. В программе очень полно представлен комплекс гистограмм, диаграмм, таблиц, схем и других средств графического моделирования исходных и результатных данных по решаемым задачам пользователя.

4. Основные этапы работы

1. Постановка задачи практической работы

На первом практическом занятии со студентами проводится общая постановка практических работ. Преподаватель может давать необходимые пояснения по методике предстоящих практических работ. После ознакомления с программным комплексом преподаватель проводит постановку задачи конкретного практического занятия. Здесь разъясняется группе студентов содержание и объем работ, предусмотренных конкретной практической работы. Прежде всего, формулируется цели, задачи, основные этапы работы, последовательность и ход решения задачи практической работы. Определяются содержание и форма представления результатов работы. Необходимо отметить, что каждая практическая работа студента должна быть оформлена в виде индивидуального отчета о практической работе. Поясняется методика составления и оформления отчета о практической работе.

2. Ознакомление студента с содержанием и объемом практической работы

На этом этапе студент должен тщательно изучить содержание и объем предстоящей практической работы. Если постановка задачи недостаточно ясна, он может обратиться к преподавателю за дополнительными разъяснениями. Затем студент приступает к выполнению задания практической работы.

3. Порядок выполнения практической работы

Студент включает компьютер и входит в программно-информационную среду комплекса «Stadia». В соответствии с установленной последовательностью этапов работы выполняет объем работ, предусмотренных заданием практической работы. При выполнении практических работ следует обратить внимание, что комплекс «Stadia» имеет систему справоч.

При условии выполнения полного объема практической работы студент проверяет правильность результатов и предъявляет преподавателю результаты работы, выведенные на экран. В случае замеченных ошибок студент принимает меры к их исправлению и затем снова предъявляет результаты преподавателю для контроля и приема результатов работы. Если в работе ошибок не содержится, то приступает к составлению и оформлению отчета о практической работе.

4. Регистрация результатов и оформление отчета о практической работе

По мере того, как выполняются этапы практической работы, студент регистрирует все результаты своей работы в собственном файле. Этот файл в будущем должен быть оформлен как отчет магистранта о практической работе. Файл должен храниться в папке соответствующего магистранта. Оформление отчета выполнить по следующим правилам. Отчет о практической работе должен содержать следующие обязательные разделы – цель (задачи), методика и средства, основные этапы практической работы, выводы и литература.

Отчет по каждой практической работе составляется по следующей обобщенной структуре:

1. Наименование идентифицирующих признаков: “Отчет о практической работе №__ по теме (наименование темы) студента (указываются фамилия и инициалы, вид обучения, факультет, курс, группа).

2. Цель работы. Формулируется в соответствии с содержанием раздела «Цель работы», соответствующей практической работы.

3. Методика работы. Определяется в соответствии с указанной выше формулировкой и при необходимости уточняется в зависимости от содержания конкретной практической работы.

4. Этапы выполнения работы. Приводятся номера и наименования этапов работы, указанные выше. Последовательно по каждому из этапов приводится характеристика содержания выполненных по этапу работ.

5. Выводы по работе. К этой части работы студент должен быть особенно внимательным. Формулируются выводы теоретического и практического характера о выполненной практической работе. Обычно выводы излагаются последовательно по каждому из этапов работы (отчета) – 1-2 вывода. Указывается место и значение разработанной модели или решенной задачи. Выводы формулируются в сжатой и четкой форме. Вывод должен содержать сжатую мысль о выполненном этапе работы, как результат аналитико-синтетической переработки содержания выполненного этапа. Не следует указывать в выводах содержание и объем выполненных работ.

Текст отчета должен быть изложен лаконично и вместе с тем информативно с соблюдением правил грамматики. Текст практической работы набирается в редакторе Word через 1,5 интервала, 14 кегль, шрифт Courier New Cyr, заголовки разделов и подразделов выделяются жирным шрифтом. В конце отчета может быть указана литература, которую студент применил в практической работе. Библиографические описания литературных источников должны быть оформлены в соответствии с ГОСТ 7.1-84. Правила библиографического описания документации.

5. Заключительная часть практической работы

После окончания составления отчета студент проверяет его правильность и устраняет ошибки. При условии отсутствия ошибок предъявляет экранный отчет преподавателю. Преподаватель читает текст отчета и принимает его. При условии замеченных ошибок преподаватель указывает студенту на эти ошибки. После этого студент исправляет ошибки и повторно предъявляет отчет преподавателю.

После завершения полного объема работ, исправления ошибок по замечаниям преподавателя, сохраняет отчет, выходит из системы «Stadia» и выключает компьютер.

Практическая работа № 1. Решение задачи «Анализ основных параметров стандартизации ИТ»

1.1. Краткая постановка задачи

Основной целью задачи является получение практических навыков в моделировании процесса изучения существенных характеристик функционального состояния исследуемой ИТ, как наиболее распространённого класса информационного ресурса, а также закрепление теоретических сведений, полученных в рамках лекционного курса.

Решение задачи базируется на применении средств системного анализа, теории вероятностей и математической статистики [1,2]. Методика выполнения работы состоит в том, чтобы на основе собранного статистического материала выполнить анализ функционального состояния ИТ программными средствами диалоговой системы STADIA. На основе анализа вырабатывается последующая стратегия и тактика решения задач стандартизации и рационализации функционирования ИТ.

В ходе исследования реально функционирующей ИТ методом прямого наблюдения обнаружены и зафиксированы в «Ведомости дефектов» ошибки по достоверности документной информации, допущенные на различных этапах технологии ИТ. Дефекты по достоверности могут иметь различные модификации, например, замена истинного символа (символов) на другой (ошибочный), транспозиция («наползание») цифр показателя, пропуск реквизита-основания показателя, сбой программы, отказ ЭВМ и др.

В технологии ИТ обнаружение и исправление дефектов, в частности, по достоверности снижает эффективность использования временных, финансовых, трудовых и других ресурсов. Это отрицательно влияет на качество ИТ, в частности, на ее выходные характеристики, например, производительность ИТ, себестоимость обработки документов и др.

В данном примере выборка дефектов по достоверности представлена во временной шкале измерения - в минутах. Каждый дефект как случайное событие представлен количеством минут, затраченных на обнаружение и исправление дефекта по достоверности. Исходные статистические данные идентифицируем как переменную x_1 , состоящую из следующих 30 значений: 3, 1, 5, 2, 4, 6, 1, 4, 2, 5, 6, 4, 6, 45, 2, 4, 3, 7, 5, 4, 6, 4, 3, 5, 4, 5, 3, 5, 4, 3.

1.2. Основные этапы решения задачи

Решение данной задачи состоит из следующих этапов:

1. Ввод статистических экспериментальных данных.
2. Получение описательной статистики основных параметров ИТ по достоверности.
3. Преобразование данных.
4. Визуализация данных.
5. Анализ состояния ИТ по критерию достоверность информации.
6. Формулирование выводов.
7. Составление и оформление отчета о решении задачи.

Ввод статистических экспериментальных данных. В нашем случае применяется ППСА STADIA под управлением ОС WINDOWS. Выполнить загрузку STADIA. Ввод данных производится посредством набора на клавиатуре экспериментальных данных, указанных в разделе 1.1, в форме электронной таблицы или матрицы данных. Данные по каждой переменной записываются в соответствующий вектор-столбец матрицы, в нашем примере в столбец x_1 электронной таблицы (рис.1.1.).

Получение описательной статистики основных параметров ИТ по достоверности. Для расчета значений описательной статистики щелкнуть мышкой

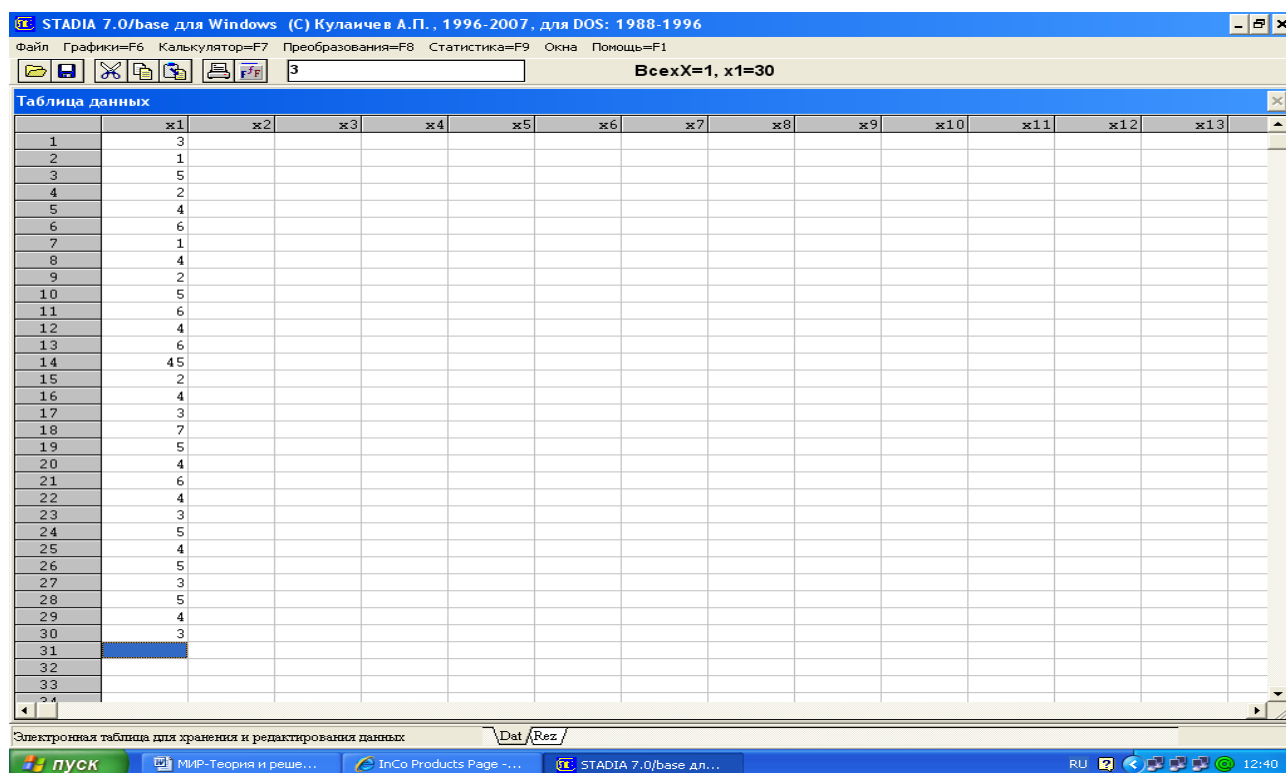


Рис. 1.1. Ввод экспериментальных данных (переменная x1)

по меню “Статистика”. В появившемся подокне «Статистические методы» в подменю “Параметрические тесты” щелкнуть мышкой по опции “Описательная статистика” (рис.1.2).

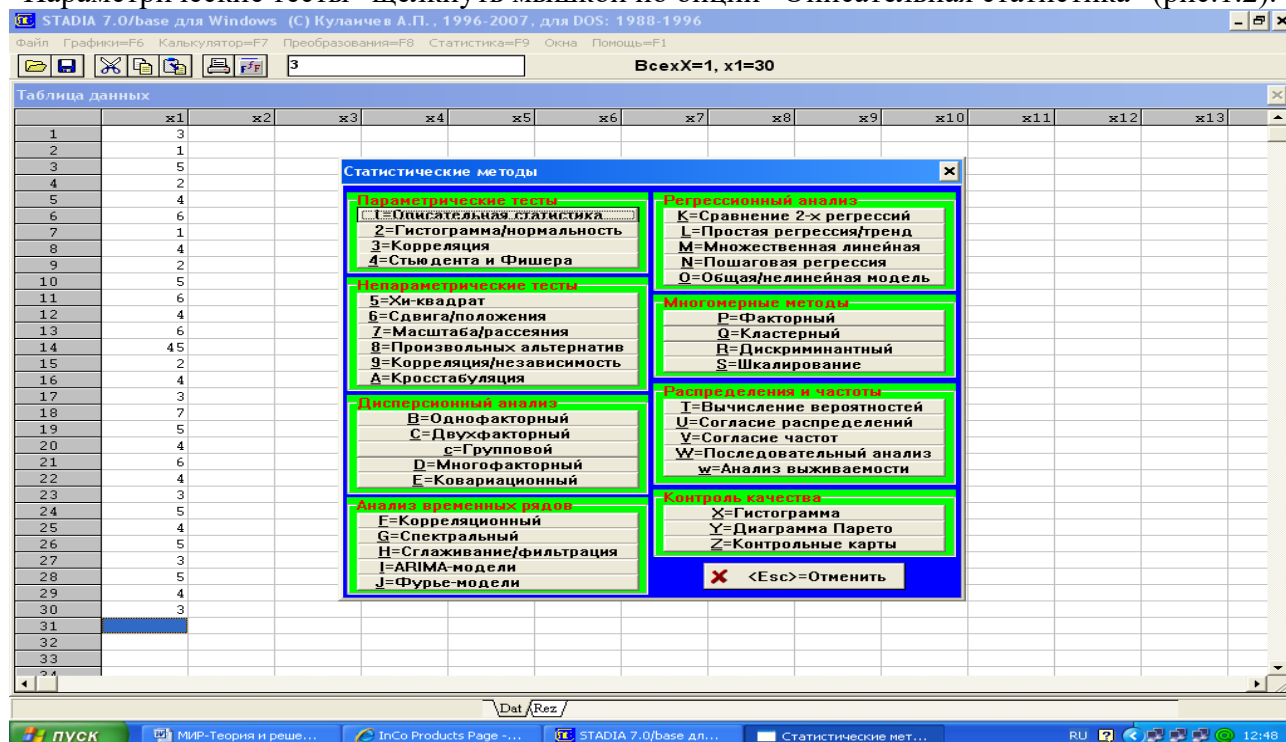


Рис. 1.2. Структура подокна «Статистические методы»

В результате активизации опции «Описательная статистика» появится подокно «Анализ данных» (рис. 1.3).

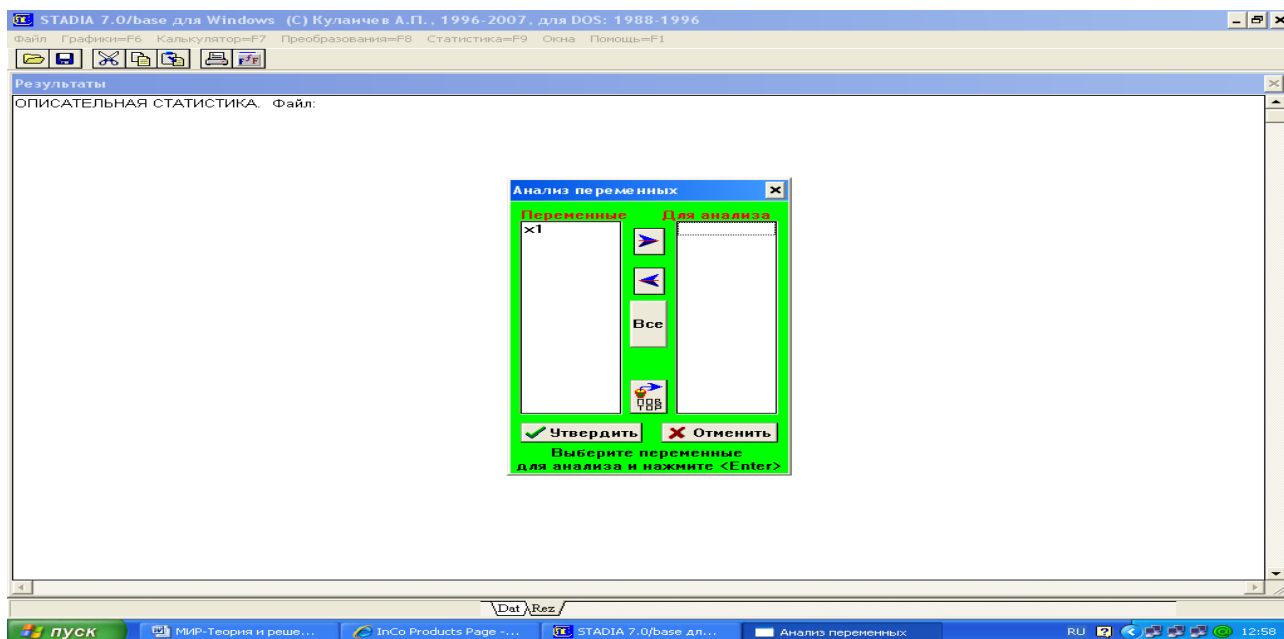
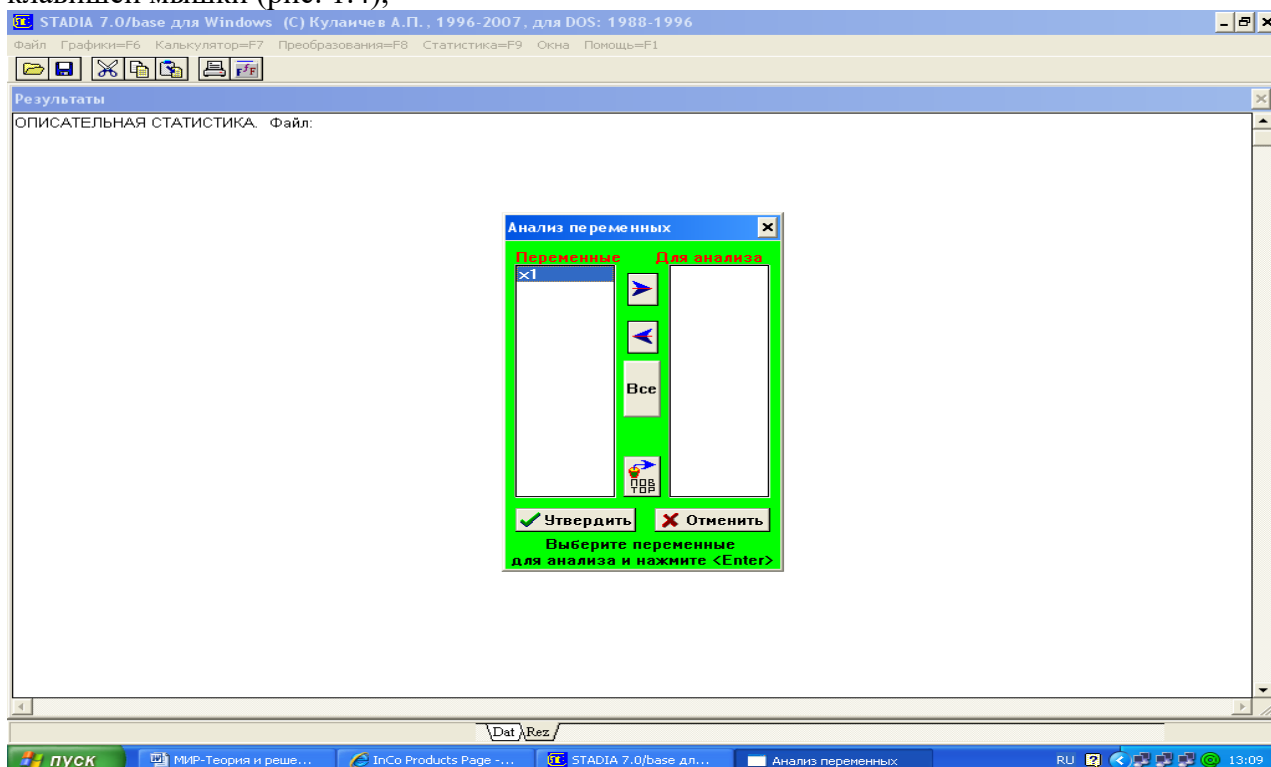


Рис.1.3. Подокно «Анализ переменных»

Затем из субполя «Переменные» переменную x_1 необходимо переслать в субполе «Для анализа». Для этого сначала нужно выделить переменную x_1 , щелкнув по ней однократно левой клавишей мышки (рис. 1.4),

Рис.1.4. Выделение переменной x_1

а затем щелкнуть левой клавишей по кнопке «стрелка вправо» и поместить переменную x_1 в субполе «для анализа» (рис.1.5).

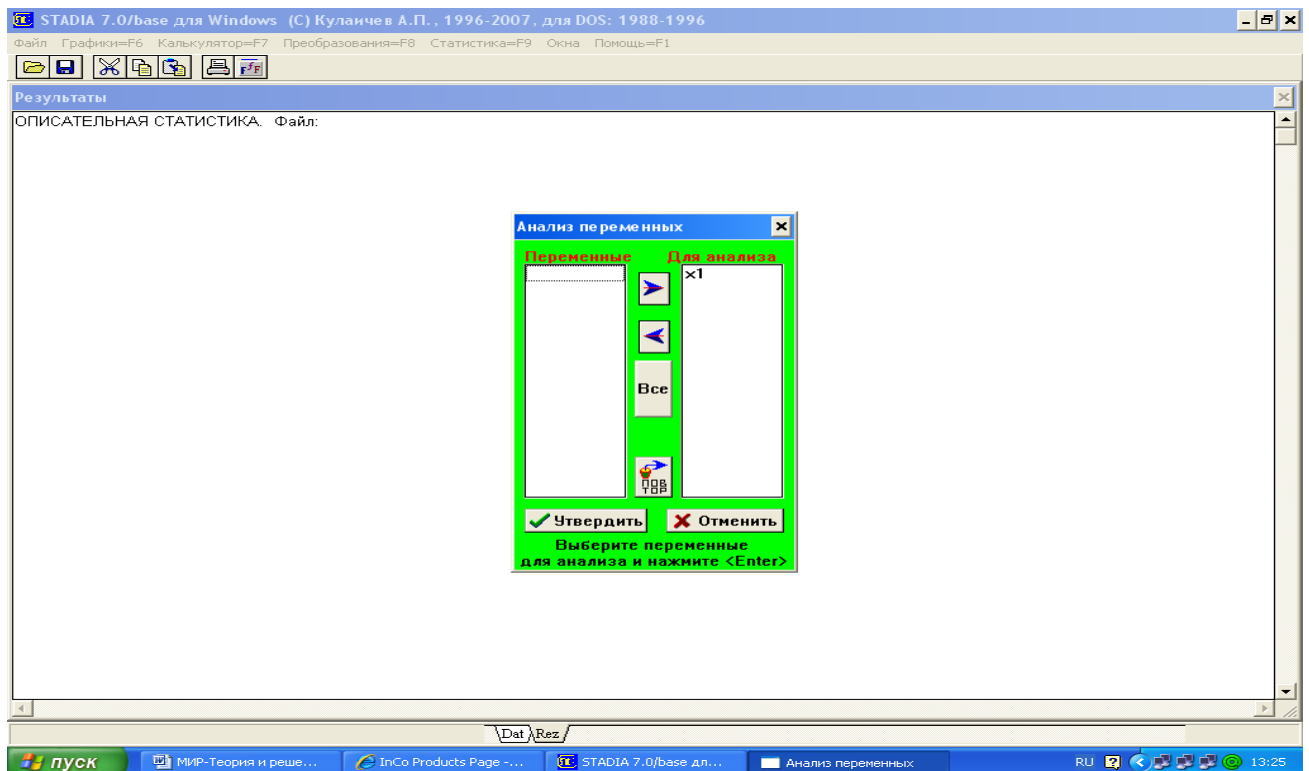


Рис.1.5. Перемещение переменной x_1 в субполе «Для анализа»

С целью устранения необходимости излишних повторов в операционном решении последующих задач, в частности, количества рисунков дадим некоторые пояснения по интерфейсу. В данном подокне кнопки «стрелка левая» и «стрелка правая» служат для пересылки переменных из одного субполя в другое. Нажатие кнопки «Все» реализует команду по перемещению всех переменных в необходимое субполе.

Для вычисления и вывода статистических оценок необходимо кликнуть кнопку «Утвердить», после чего получим результат расчета по переменной x_1 (рис. 1.6). Под рубрикой «ОПИСАТЕЛЬНАЯ СТАТИСТИКА: Файл» приведены краткие данные, которые в нашей задаче следует понимать так: Переменная – x_1 , Размер выборки (количество значений по переменной x_1) – 30, Диапазон значений – от 1 до 45 мин., Среднее выборочное значение – 5.367 мин., Ошибка – 1.394, Дисперсия – 58.31, Стандартное отклонение – 7.636, Сумма всех значений данных переменной x_1 – 161.

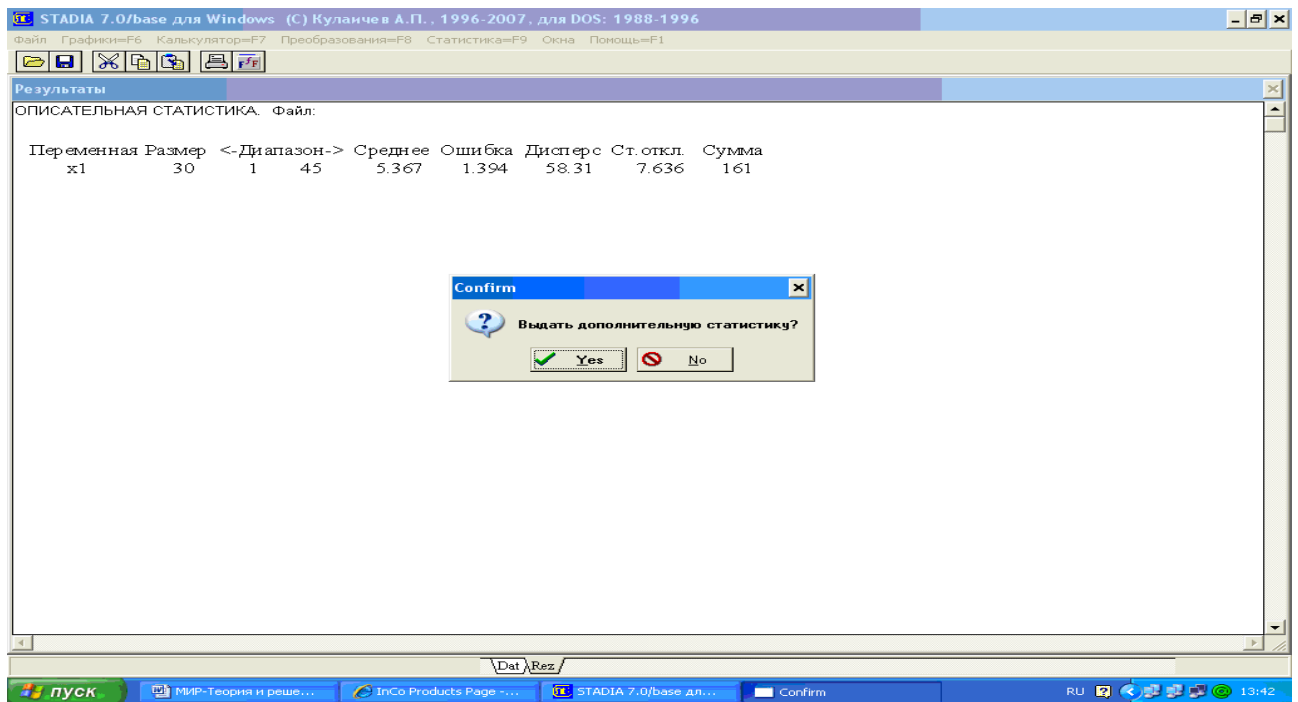


Рис.1.6. Отображение результатов расчета статистических оценок

При необходимости получения дополнительных статистических оценок нужно кликнуть в подокне «Confirm» кнопку «Yes», в противном случае – кнопку «No». В нашем случае представляется целесообразным иметь дополнительную статистику и после активизации кнопки «Yes» получаем результаты с набором дополнительных статистических оценок по переменной x1 (рис.1.7).

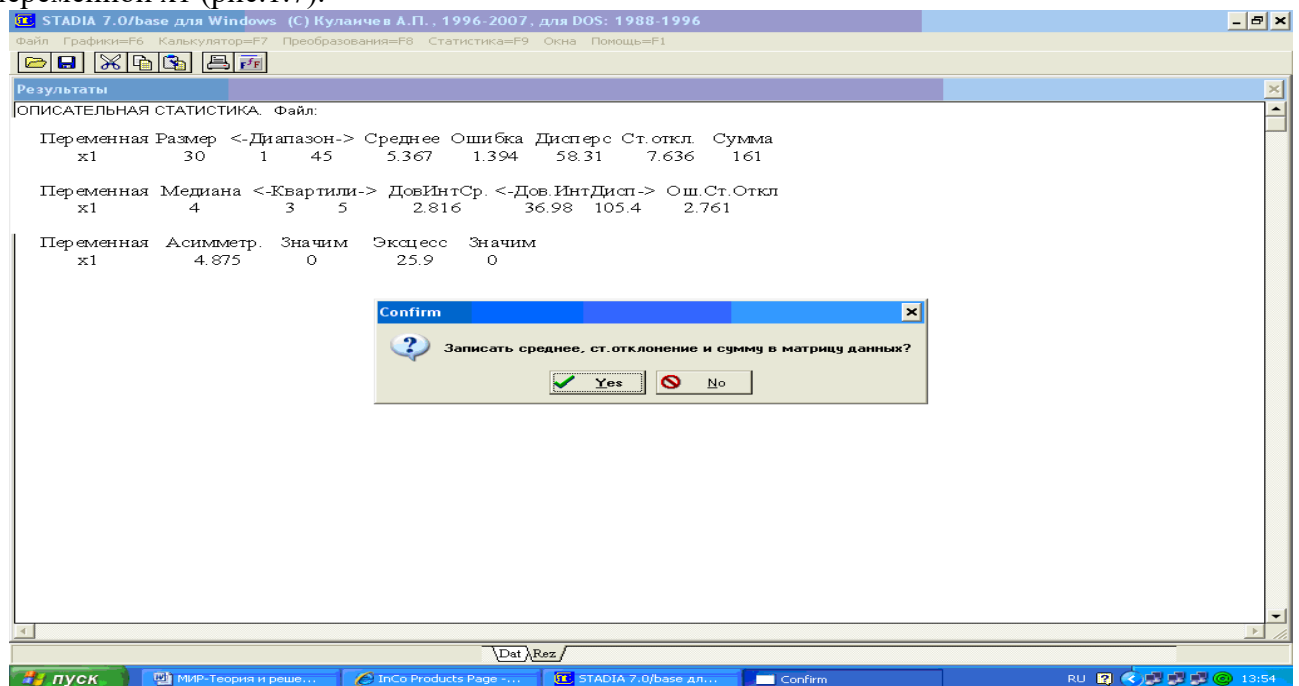


Рис.1.7. Отображение результатов расчета дополнительной статистики

При необходимости записи оценок по параметрам «Среднее», «Стандартное отклонение» и «Сумма» в матрицу данных, то есть в электронную таблицу данных нужно кликнуть в подокне «Confirm» кнопку «Yes», в противном случае – кнопку «No». При условии необходимости сохранения этих оценок, кликнув кнопку «Yes», получим необходимое сохранение этих оценок в столбцах x2, x3, x4 (рис. 1.8).

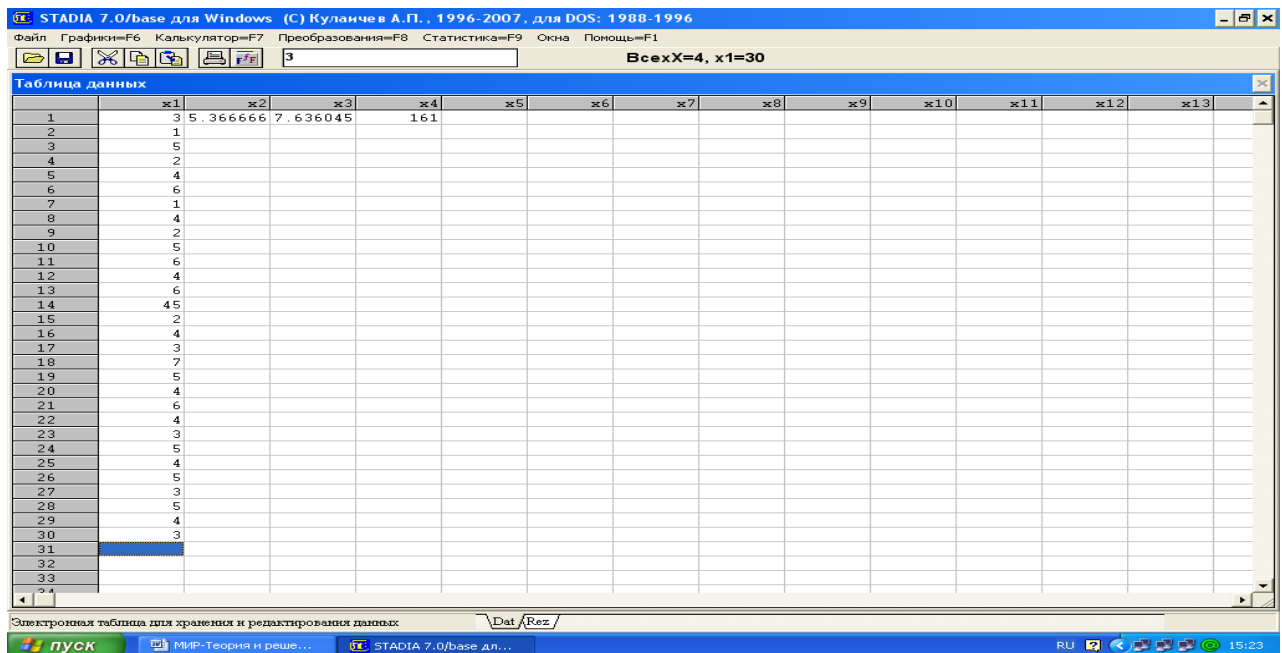


Рис. 1.8. Сохранение статистических оценок в электронной таблице

Описательная статистика представляет в ходе исследования основные параметры, характеризующие собранные данные о дефектах по достоверности, в частности: выборочное среднее значение, разброс значений или стандартное отклонение, характеристики формы распределений значений (мода, медиана, ассиметрия, эксцесс, квартили) и т.д.

Результаты обработки данных представляются в виде соответствующих экранов - исходных данных, результатов вычислений и построенных графиков. Вызов и смена экранов выполняется щелчком мышки на соответствующей надписи («Dat», «Rez») в нижней командной линейке экрана STADIA.

Ниже приводятся краткие определения и пояснения некоторых терминов описательной статистики. Полное описание и порядок вычисления значений оценок в ППСА STADIA изложены в [2].

Выборочное среднее - среднее выборочное значение вычисляется как сумма всех значений, деленная на их число.

Медиана - медиана m представляет собой значение, которое делит выборку пополам: число выборочных значений, меньших m , равно числу выборочных значений, больших m . При симметричном распределении значений переменной выборочное среднее обычно близко к значению медианы.

Квартили - квартили представляют собой два выборочных значения, которые относительно медианы делят выборку на четыре равные части. Здесь наблюдается определенная аналогия с медианой.

Стандартное отклонение или среднеквадратичное отклонение показывает, насколько сильно выборочные значения разбросаны относительно среднего. Для нормально распределенных данных в границах двух стандартных отклонений относительно среднего значения размещается чуть более 95% наблюдений. Эта оценка используется в экспериментальных исследованиях ИП для обозначения практической нормы.

Дисперсия - представляет собой квадрат стандартного отклонения. Вместе с тем, в отличие от последнего дисперсия обладает свойством адитивности, то есть сумма дисперсий нескольких выборок равна дисперсии объединенной выборки.

Ошибка среднего - характеризует точность вычисления среднего значения с учетом величины разброса выборочных значений и размера выборки. Для нормально распределенных данных в границы двух ошибок в вероятность 95% попадают оцениваемое среднее значение генеральной совокупности, что соответствует 95%-му доверительному интервалу среднего.

Доверительный интервал - для некоторой выборочной оценки (среднего, дисперсии, доли и др.) представляет собой диапазон значений наблюдаемой случайной переменной, который с заданной (доверительной) вероятностью «накрывает» (или содержит) неизвестное генеральное значение оцениваемой характеристики этой переменной. Например, среднее выборочное значение оценивается (вычисляется) нами по ограниченной выборке из потенциально бесконечной генеральной совокупности. Отсюда это значение априори содержит неизвестную случайную ошибку – насколько оно отличается от генерального среднего. Ошибка будет статистически уменьшаться при увеличении размера выборки и при уменьшении заданной точности такого предсказания (доверительная вероятность). Так, 95% доверительный интервал относительно вычисленного выборочного среднего говорит о том, что в этом диапазоне с 95% вероятностью содержится оцениваемое генеральное среднее. Доверительные интервалы могут служить эффективным средством количественной проверки статистических гипотез.

Асимметрия и эксцесс - показатели асимметрии и эксцесса характеризуют степень несимметричности выборочного распределения относительно среднего значения и степень выраженности его центрального пика. Для нормального закона распределения эти показатели имеют следующие значения: асимметрия=0 (события распределены симметрично относительно среднего значения выборки), эксцесс=3.

В задачах анализа информационных ресурсов, например, необычные значения переменной и другие отклонения можно лучше увидеть, если воспользоваться визуализацией данных.

Визуализация данных. Для выбора тактики дальнейшего анализа целесообразно применить формы графического представления - визуализации данных посредством отображения значений переменной x_1 на графике. Для этого в главном окне STADIA кликнуть мышкой меню «Графики», после чего появится подокно «Графики данных» (рис. 1.9).

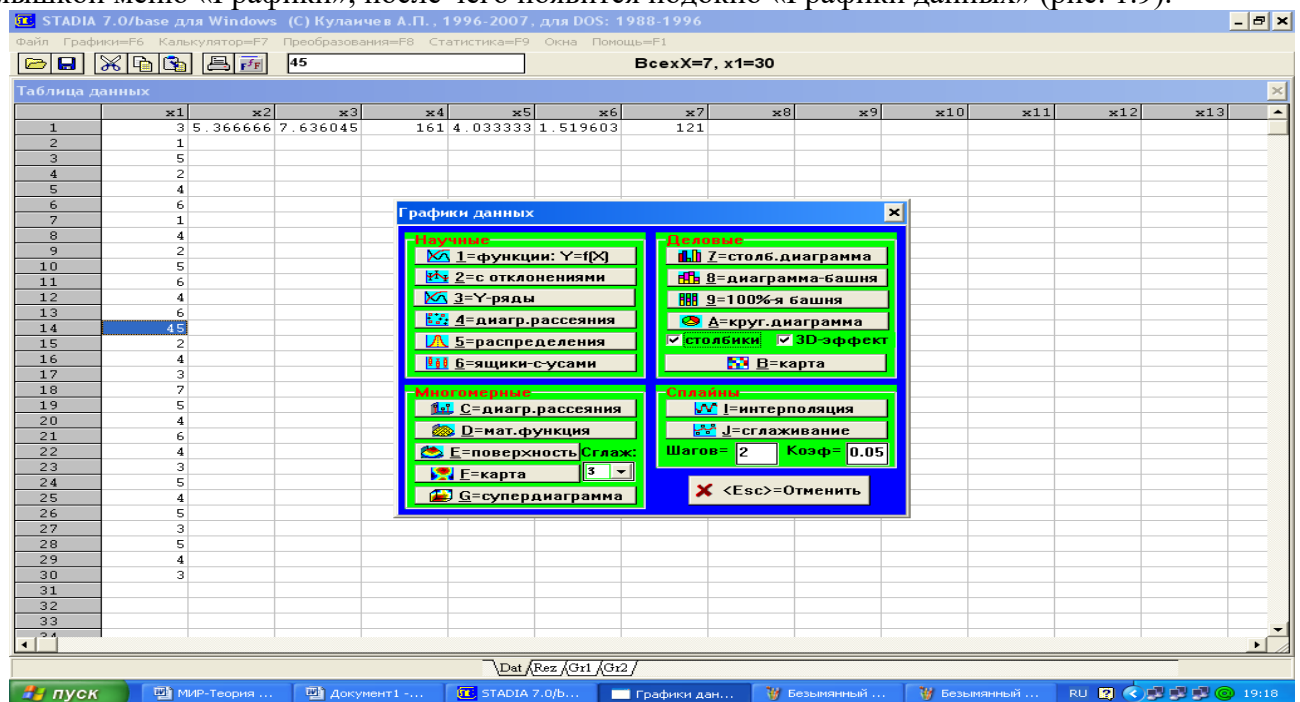


Рис. 1.9. Отображение группы команд «Графики данных»

Затем в группе опций «Научные» кликнуть кнопку «1=функции: $Y=f(X)$ », после чего появится подменю «Графики данных» (рис.1.10).

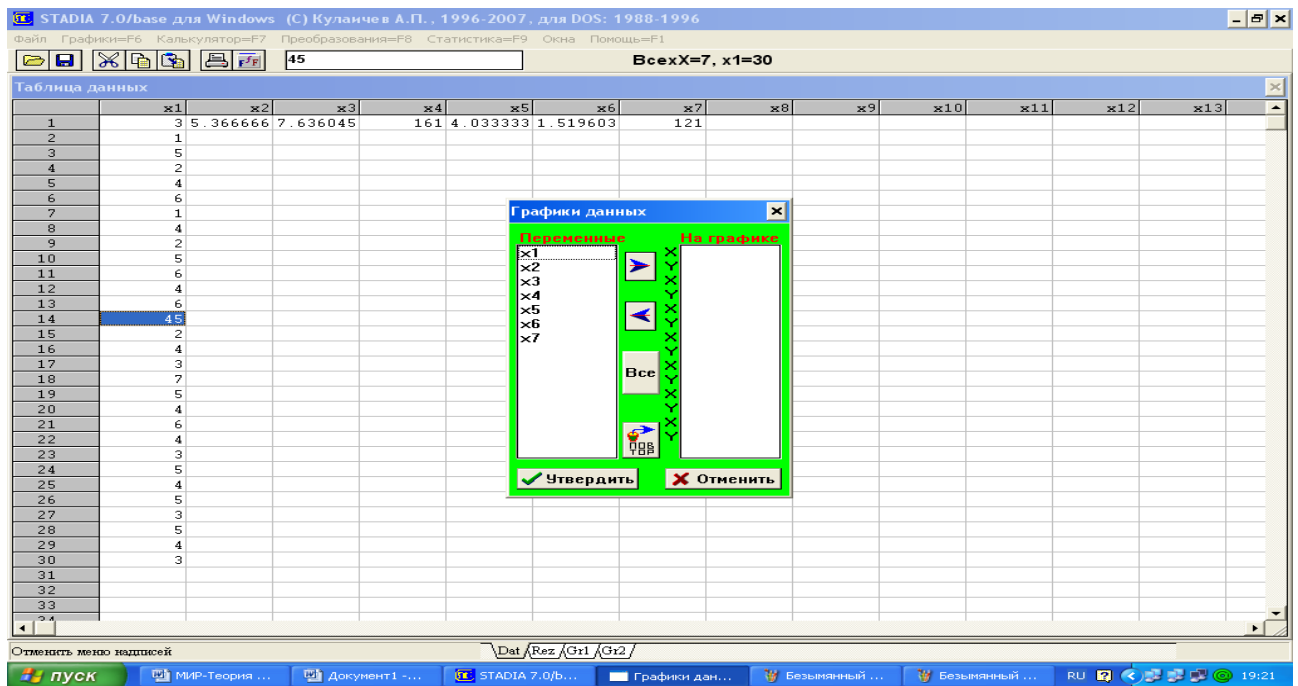


Рис.1.10. Отображение переменных в подокне «Графики данных»

В появившемся подокне «Графики данных» в субполе Переменные выделить мышкой переменную x1 и, щелкнув мышкой по стрелке, направленной вправо, переслать ее в субполе «На графике». Затем щелкнуть мышкой на опции «Утвердить», после чего появится график (рис. 1.11).

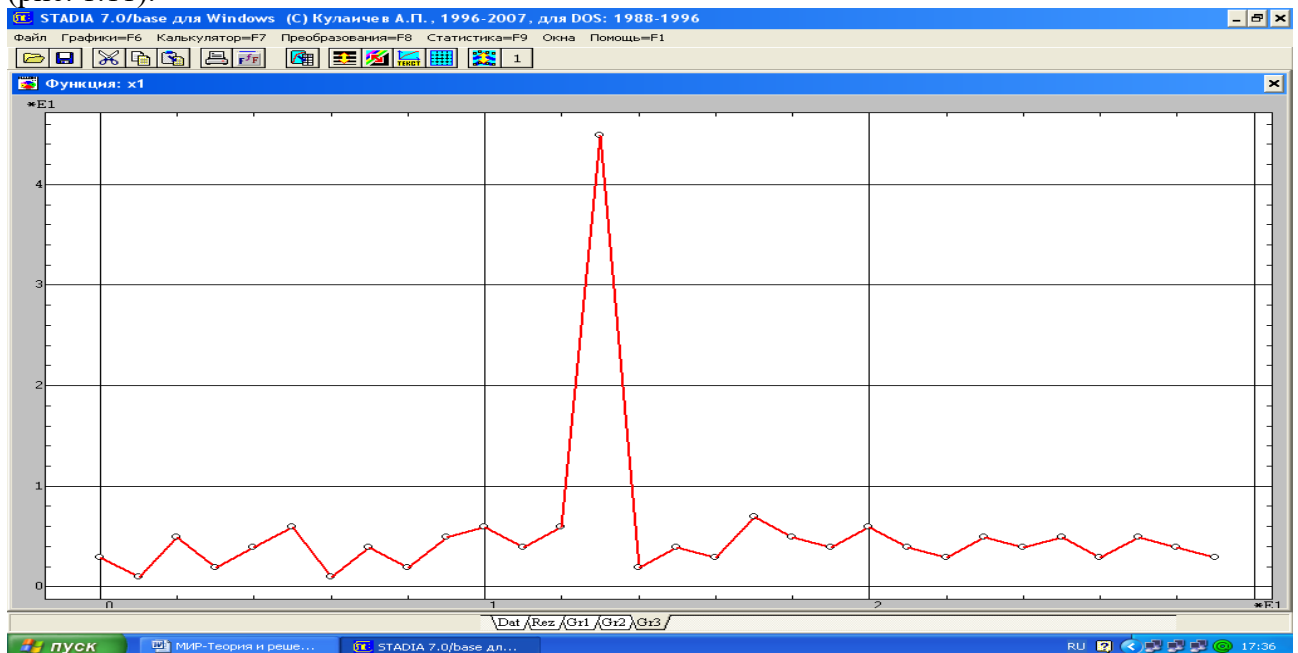


Рис. 1.11. График распределения данных переменной x1

Посредством этого графика можно анализировать свойства переменной x1. На графике видно, что распределение значений, в основном равномерно. Вместе с тем наблюдается одно значение (45 мин.), которое не согласуется с общим характером значений выборки, заходит за границы переменной. Подобное значение обозначается как «выброс», или «флуктуация», которое, как правило, искажает адекватность экспериментальных данных и может снизить уровень качества моделирования. Для устранения такого условия в моделировании можно применить способ «фильтрации» данных, или преобразование данных.

Преобразование данных. Данные в электронной таблице необходимо просмотреть и скорректировать с целью их приведения к адекватному виду решаемой задачи. Нередко

требуется удалить или «сгладить» высокоамплитудные «выбросы», например, в нашем случае 14-й дефект в 45 минут явно не отличается своей репрезентативностью, поэтому его значение в контексте решаемой задачи можно заменить на значение, близкое к значению «средней выборочной», полученному в результате вычислений. Поскольку вычисленное нами «Среднее выборочное» равно 5.367 (см. рис. 1.6), то мы можем заменить в электронной таблице значение 45 на округленное значение 5. По отфильтрованным данным, путем реализации уже описанных выше последовательных команд, получим более адекватные статистические оценки (рис.1.12).

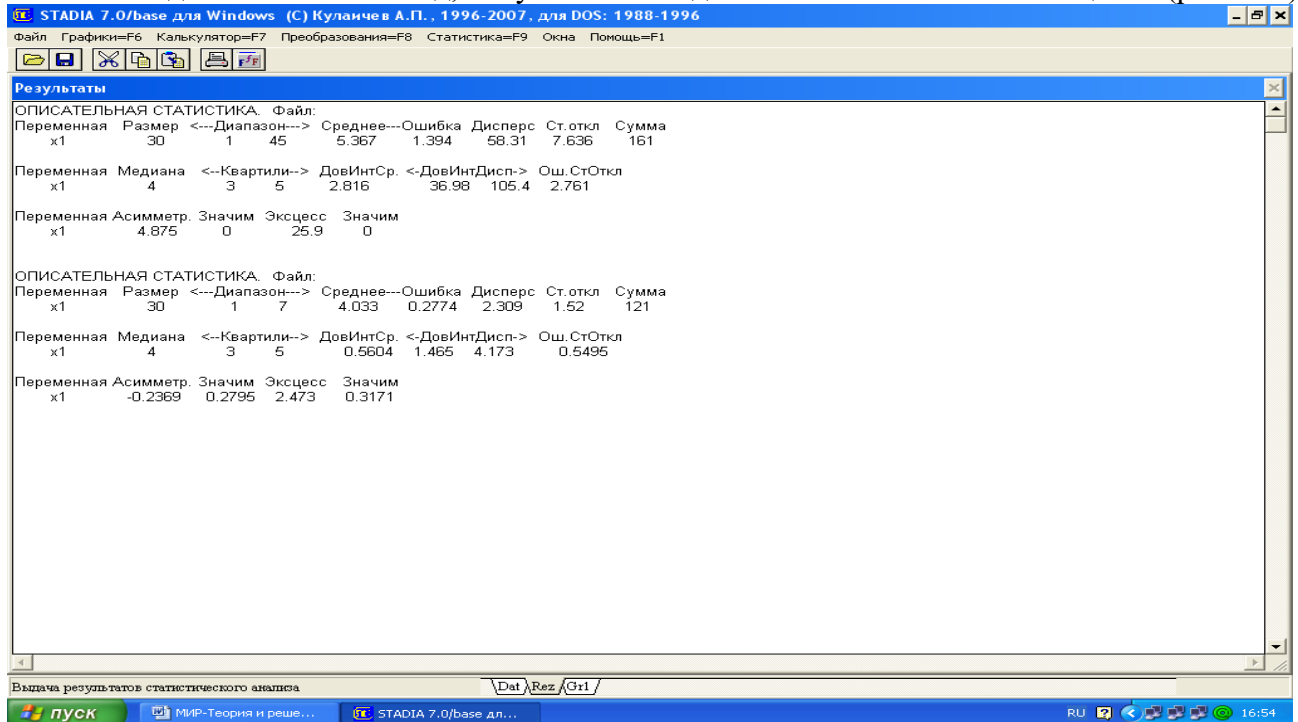


Рис.1.12. Результаты расчета описательной статистики до фильтрации и после фильтрации экспериментальных данных

Для расширения аналитической базы построим график по отфильтрованным экспериментальным данным. В результате реализации последовательности команд, изложенной выше, построим второй график (рис.1.13). На основе описательной статистики и графиков можно выполнить анализ состояния ИТ.

Анализ функционального состояния ИТ по критерию достоверность информации. По результатам анализа значений описательной статистики необходимо интерпретировать функциональное состояние ИТ относительно влияния дефектов по достоверности на уровень качества технологии обработки данных ИТ в контексте решения задачи моделирования ИТ в целом и её компонентов.

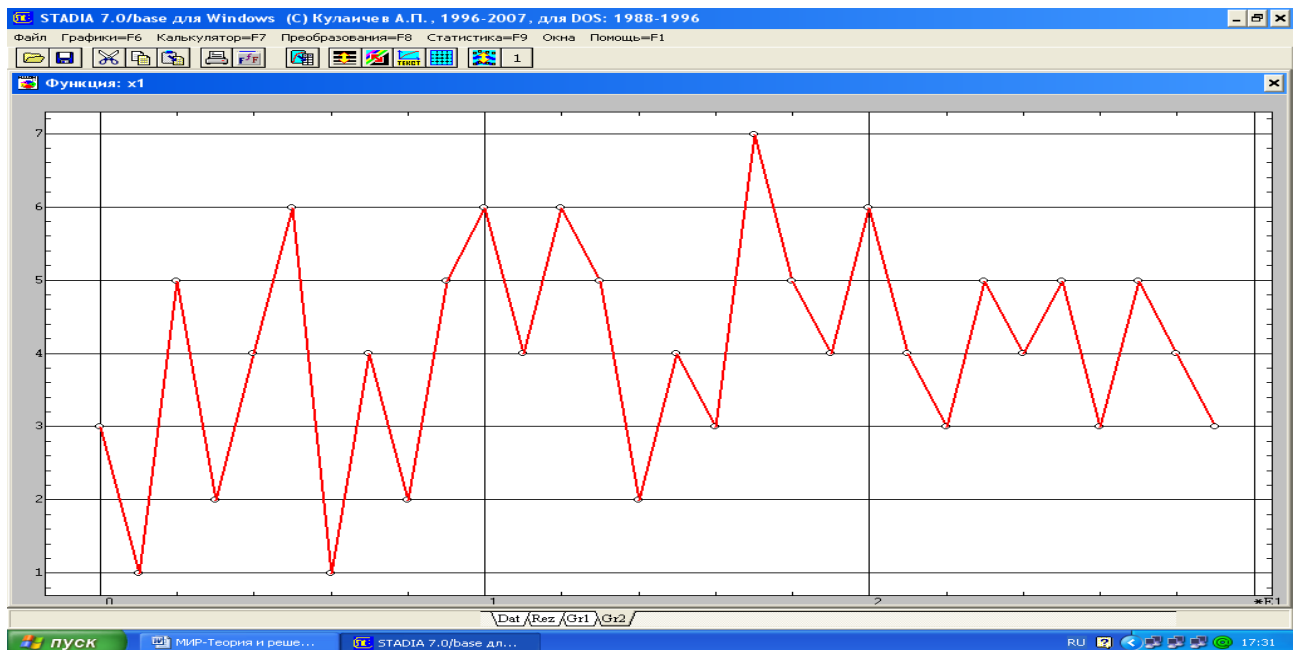


Рис.1.13. График отфильтрованных данных переменной x_1

Априори можно предположить следующее. Ошибки (искажения) истинного значения содержащихся в документах данных являются довольно значимым фактором, отрицательно влияющим на эффективность технологического процесса ИТ. В моделировании функционального состояния ИТ большое значение имеет адекватность привлекаемых к моделированию данных. Сравнительный анализ полученных результатов экспериментальных данных до и после их фильтрации показал следующее.

Значения базовых статистических оценок подтверждает, что фильтрация имеет положительное влияние на уровень достоверности привлекаемых к моделированию экспериментальных данных, а, следовательно и на уровень адекватности модели анализа. Так, например, значение «среднего выборочного» обнаружения и исправления дефекта по достоверности до фильтрации равнялось 5,367 мин., а после фильтрации оно стало равно 4.033 мин. Таким образом, это очень важный показатель в результате уточнения скорректирован на 25%. В практическом отношении это означает то, что в информационно-вычислительной структуре при организации технологического процесса обработки данных трудоёмкость процедур диагностики и корректировки дефектов по достоверности обрабатываемых данных может быть снижена на 25%, а следовательно при определенных условиях и штатная численность работников, занятых в технологии обработки данных ИТ.

Уровень точности расчета значения «Среднего выборочного» подтверждается следующими оценками. Если до фильтрации значение «Ошибки» составляло 1.394, то соответствующее значение после фильтрации стало равно значению 0.2774, явно существенно меньшему. Подобное условие наблюдается и по параметрам «Дисперсия» и «Стандартное отклонение». Если анализировать дополнительные статистические оценки, то они также подтверждают лучшую адекватность отфильтрованных данных. Так, например, если до фильтрации «Значимость» относительно «Ассиметрии» и «Экссесса» имели соответственно нулевые значения, то после фильтрации «Значимость» изменила свои значения на более адекватные, то есть 0.2795 и 0.3171. Это логично, так как значения оценок «Ассиметрии» и «Экссесса» для отфильтрованных данных равны соответственно -0.2369 и 2.473. Указанные значения намного ближе к стандартам математической статистики (нормальное распределение), то есть 0 для «Ассиметрии» и 3 для «Экссесса», чем у данных до фильтрации, соответственно - 4.875 и 25.9. По завершении анализа следует перейти к формулировке выводов.

Формулирование выводов. Необходимо дать лаконичное заключение о решении данной задачи в форме выводов. Целесообразно выводы построить в соответствии с этапами выполнения работы. Необходимо по каждому из этапов сформулировать резюме по наиболее существенному аспекту выполненного этапа. Необходимо учесть, что выводы должны быть связаны между собой, так же как и этапы решения задачи. Содержание выводов должно отражать последовательность решения задачи, с одной стороны, содержание этапа, его место и значение в общем комплексе работ по решению задачи, с другой стороны. Полезно дать собственную оценку эффективности пакета STADIA в решении данной задачи.

По окончании формулирования выводов необходимо ещё раз проверить качество полученных результатов по решению задачи. Затем приступить к составлению и оформлению отчета о решении задачи. Составление и оформление отчета выполнить в соответствии с требованиями, изложенным в разделе 10.1.

Контрольные вопросы и задания

1. Назовите цель решенной Вами задачи?
2. Какие методы и средства применяются для решения данной задачи?
3. Назовите основные этапы решения задачи в их логической последовательности?
4. Для чего необходима фильтрация экспериментальных данных?
5. Как выполняется сравнительный анализ функционального состояния ИТ?
6. Какие статистические оценки в стандартизации привлекаются для анализа функционального состояния ИТ?

Практическая работа № 2. Решение задачи «Определение закона распределения экспериментальных данных при стандартизации технологического процесса обработки данных ИТ»

2.1. Краткая постановка задачи

Основной целью решения этой задачи является получение практических навыков в моделировании определения закона распределения экспериментальных данных о качестве исследуемого технологического процесса ИТ, а также закрепление теоретических сведений, полученных в рамках лекционного курса.

В качестве методов в решении задачи применяются средства теории вероятностей и математической статистики [2]. Методика выполнения работы состоит в том, чтобы посредством последовательных вычислительных процедур установить статистический закон распределения, которому подчиняются экспериментальные данные, полученные в результате сбора данных о дефектах технологического процесса обработки данных ИТ. Необходимые вычисления выполняются программными средствами диалоговой системы STADIA. На основе результатов вычисления определяется закон распределения, в соответствии с которым выполняются последующие вычисления в процессе моделирования и улучшения качества технологического процесса обработки данных.

В практике экспериментального исследования ИТ надежность и достоверность результатов исследования в значительной мере зависит от уровня точности экспериментальных данных случайной выборки, полученной в ходе наблюдения и сбора данных о дефектах, возникающих на различных этапах технологического процесса обработки данных ИТ. Очень часто в практике научного поиска и возникающей при этом необходимости решать задачи моделирования перед исследователями встает вопрос - какова вероятность тех событий, которые отражает статистическая структура дефектов, то есть экспериментальные данные? Каков уровень точности, надежности вычислительных результатов, получаемых на основе обработки экспериментальных данных? Какую методику вычислений следует принять для получения адекватной модели исследуемого технологического процесса обработки данных или какого-либо другого вида информационного ресурса? Обоснованность этих вопросов объясняется многими причинами. Одной из основных причин является то, что автоматизированная информационная система и её компоненты, в частности, технологический процесс обработки данных, в силу своего человеко-машинного характера относятся к классу вероятностных систем. Разумеется, собранные экспериментальные данные и результаты их обработки будут отражать вероятностный характер технологии обработки данных. Ясно, что качество модели ИС будет зависеть от уровня надежности, достоверности результатов расчетов, которые определяются уровнем надежности, или погрешности исходных экспериментальных данных. Одним из способов улучшения надежности результатов вычислений и улучшения качества моделирования является определение характера экспериментальных данных посредством проверки - подчиняются ли каким-либо законам экспериментальные данные о дефектах и, если да, то какому?

Проводимые расчетные работы по моделированию ИТ, как правило, базируются на наборе исходных экспериментальных данных, в частности, на оценках описательной статистики (см. гл. 8, 11). Уровень надежности последних увязывается с определением вида закона распределения вероятностей, которому подчиняются случайные числа выборки.

В качестве исходных экспериментальных данных возьмем выборку по дефектам полноты данных в документах. Такими дефектами являются отсутствие (пропуски) необходимых чисел, слов, показателей и других обязательных элементов документов, текстов, файлов, сообщений, зафиксированных на различных носителях, или передаваемых по каналам связи. В информационно-вычислительных структурах, в технологических процессах обработки

данных ИТ на обнаружение и исправление указанных дефектов расходуются временные, трудовые, аппаратные, программные и финансовые ресурсы. В данной выборке из 30 чисел каждое число представляется дефектом по полноте и измеряется в минутах: 9, 12, 11, 14, 16, 11, 14, 12, 15, 12, 14, 12, 13, 11, 14, 13, 10, 15, 13, 11, 12, 10, 12, 13, 12, 14, 11, 13, 16, 13.

2.2. Основные этапы решения задачи

Решение данной задачи состоит из следующих этапов:

1. Ввод статистических экспериментальных данных.
2. Проверка распределения экспериментальных данных на нормальность.
3. Уточнение вида закона распределения экспериментальных данных.
4. Формулирование выводов.
5. Составление и оформление отчета о решении задачи.

Ввод статистических экспериментальных данных. Ввод в ЭВМ экспериментальных данных в объеме 30 чисел (см. раздел 2.1), выполняется в соответствии с регламентом ввода, указанным в тексте раздела 1.2. В результате процедуры ввода в электронной таблице в столбце x1 размещаются 30 значений, формирующих переменную x1. При условии отсутствия ошибок ввода следует приступить к этапу проверки распределения на нормальность.

Проверка распределения экспериментальных данных на нормальность. Общепринятой формой представления выборочного распределения является гистограмма. Для её вычисления диапазон изменения выборочных значений разбивают на некоторое число равных интервалов и подсчитывают число значений, попадающих в каждый интервал. При графическом представлении гистограммы на каждом интервале строится прямоугольник (столбик), высота которого пропорциональна числу выборочных значений, попавших в интервал.

Для запуска процедуры проверки распределения на нормальность необходимо через меню «Статистика» обратиться к подменю «Статистические методы» и в разделе «Параметрические тесты» кликнуть кнопку «2=Гистограмма/нормальность» (рис. 2.1).

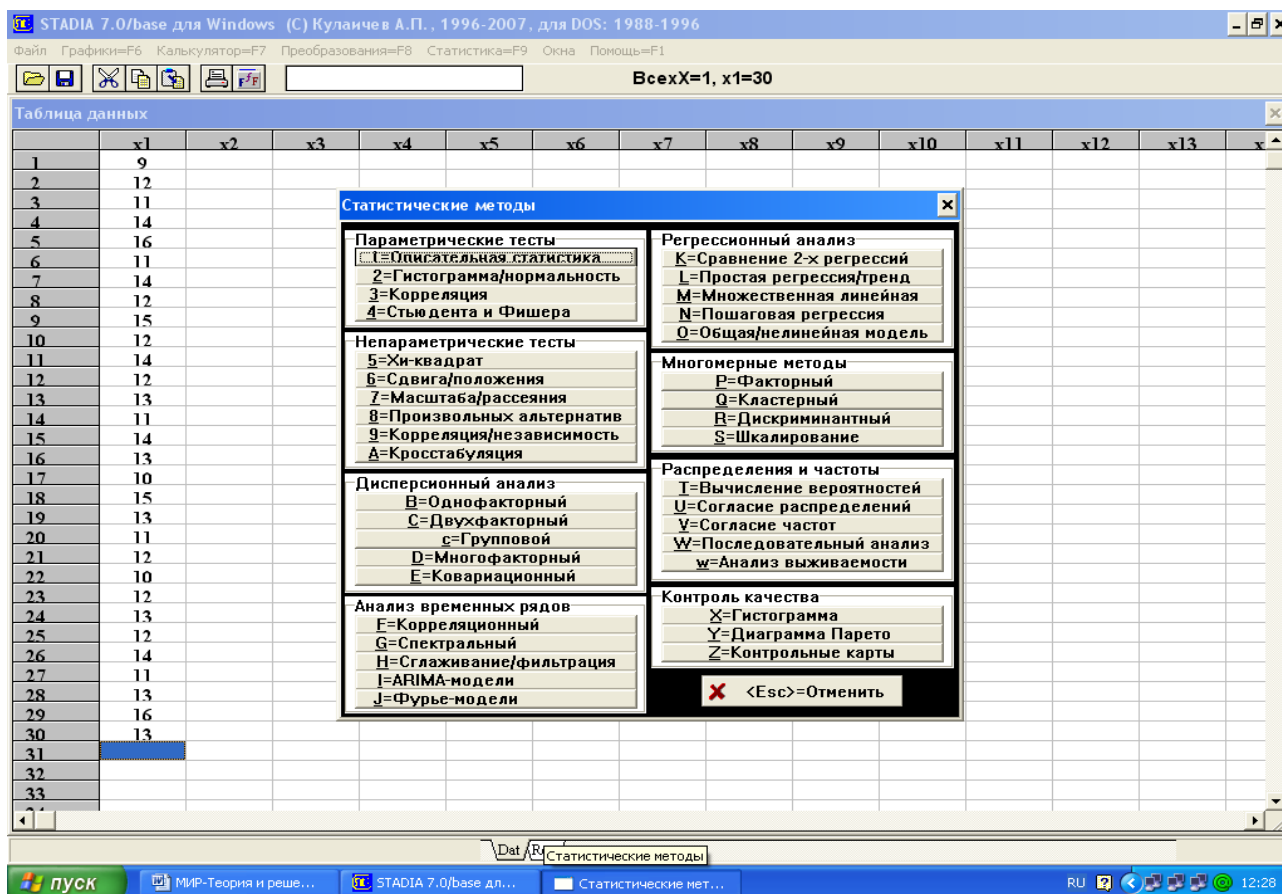


Рис. 2.1. Подокно «Статистические методы»

После чего появится подокно «Гистограмма» (рис. 2.2.).

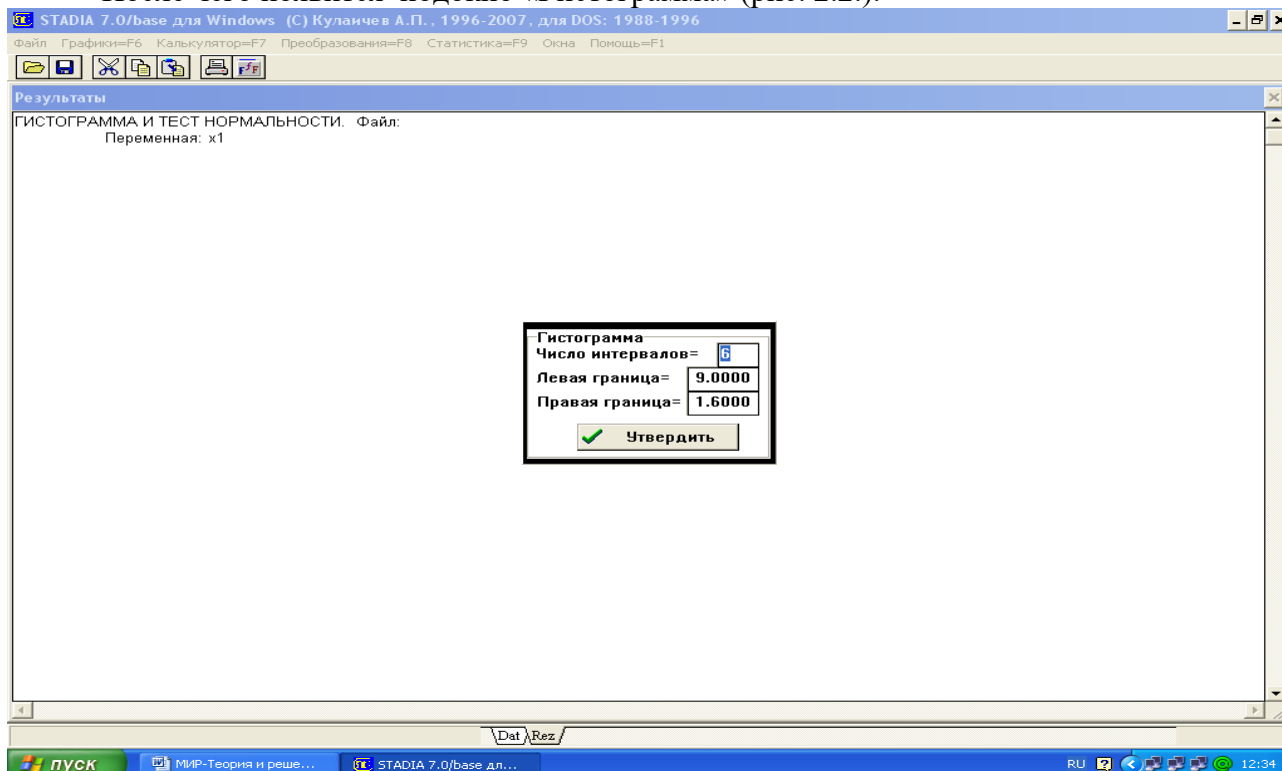


Рис. 2.2. Подокно «Гистограмма»

В этом подокне пользователю программа предлагает количество интервалов гистограммы (в данном случае – 6) с указанием левой и правой границ. Пользователь имеет

возможность изменить количество интервалов по своему усмотрению. Но поскольку программный модуль определяет, как правило, правильно рациональное количество интервалов, то в данном случае можно согласиться с предлагаемым программой количеством интервалов.

Далее следует кликнуть по кнопке «Утвердить», после чего появится окно «Результаты», содержащие данные по гистограмме и тесту нормальности, а также подокно «Confirm» с вопросом «Сохранить гистограмму в матрице данных?» (рис.2.3).

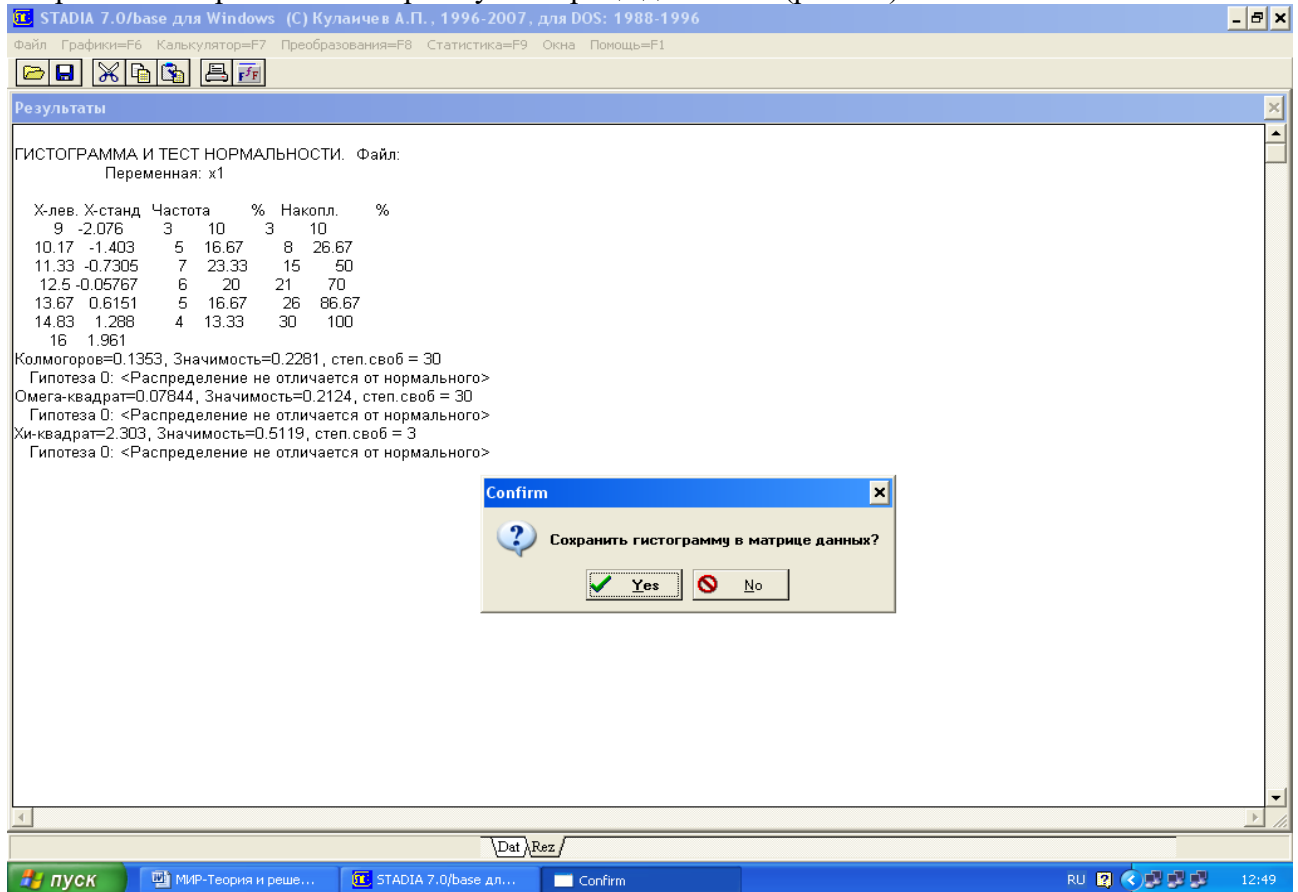


Рис. 2.3. Окно «Результаты», отображающее данные по гистограмме и тесту нормальности

Для получения полного состава данных по проверке экспериментальных данных на нормальность кликнем клавишу «Yes» и получим отображение гистограммы (рис. 2.4).

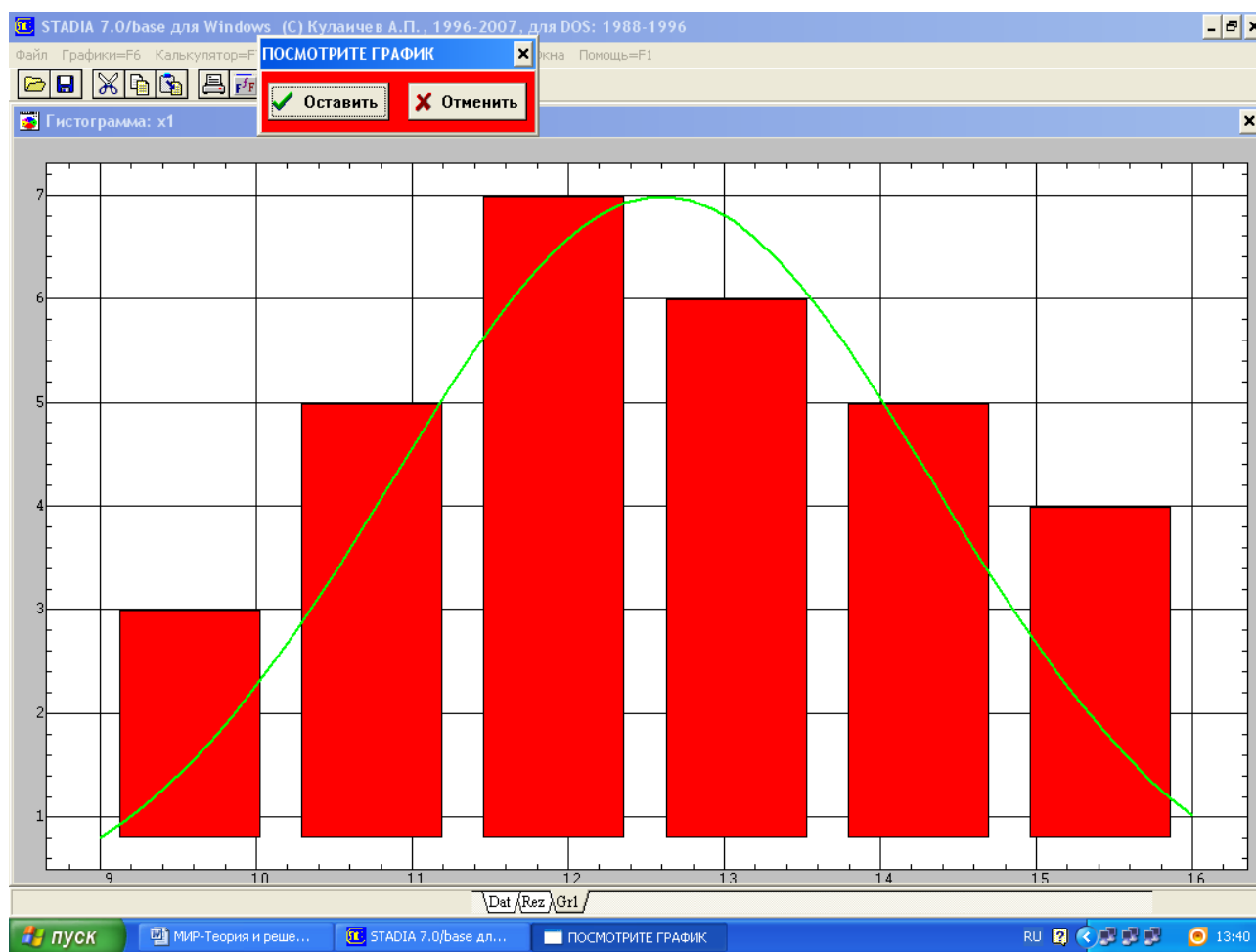


Рис. 2.4. Гистограмма дефектов по показателю «полнота данных»

В подокне «ПОСМОТРИТЕ ГРАФИК» нужно кликнуть клавишу «Оставить», чтобы сохранить гистограмму в окне «Gr1». Затем проведём анализ данных на нормальность распределения экспериментальных данных.

Под записью «Переменная: x1» приводятся следующие данные. Для каждого интервала гистограммы на экран выводятся значения: «X-лев.» - левая граница интервала в исходных единицах, «X-станд.» - значения в единицах стандартного отклонения, «Частота» - число выборочных значений, попавших в интервал, «%» - доля выборочных значений в %, «Накопл.» - накопленное число выборочных значений до текущего интервала включительно, «%» - накопленная доля выборочных значений в % до текущего интервала включительно.

Затем проводится проверка нулевой гипотезы об отсутствии различий между выборочным и нормальным распределением и выдача трех различных статистик (критериев): Колмогорова, Омега-квадрат и Хи-квадрат. По каждому критерию определены 3 параметра – собственно значения критерия, значимость и степень свободы.

Проверка выборочного распределения на нормальность может быть проведена несколькими способами, которые дополняют друг друга:

- визуальный метод в качестве предварительной субъективной оценки может быть осуществлен по рисунку гистограммы выборочного распределения с наложенной кривой плотности вероятности нормального распределения;
- проверка нулевой гипотезы соответствия распределений по коэффициентам асимметрии и эксцесса;
- проверка соответствия формы распределений по критериям Колмогорова, Омега-квадрат и Хи-квадрат.

Пояснения. Критерий Колмогорова реагирует на наибольшую разность между распределениями, которая обычно проявляется вблизи максимума функции плотности

вероятности, поэтому он плохо адаптирован для выявления различий на концах («хвостах») распределений.

В противоположность вышеприведенному критерию критерий Омега-квадрат является более равномерным, учитывая различия между распределениями на всем интервале выборочных значений, однако он сравнительно менее исследован в плане составления таблиц критических значений и предельных аппроксимаций для различного типа распределений.

Критерий Хи-квадрат также достаточно равномерно учитывает различия на всем диапазоне выборочных значений, однако требует большей осторожности при своем применении к непрерывным распределениям, поскольку его результаты существенно зависят от объема выборки и от разбиения выборочного пространства на интервалы.

В математической статистике индикатором для проверки нулевой гипотезы, наряду с другими данными, применяется уровень значимости P . При условии $P > 0.05$ нулевая гипотеза может быть принята. В нашем случае результатные данные показывают, что по критериям Колмогорова, Омега-квадрат и Хи-квадрат уровни значимости P соответственно равны – 0.2281, 0.2124, 0.5119, что означает возможность принятия нулевой гипотезы о соответствии экспериментальных данных нормальному закону распределения.

Уточнение вида закона распределения. В практике научного исследования и моделирования ИР часто требуется более точное определение закона распределения случайных величин выборки. С этой целью прибегают к проверке распределения на соответствие определенному закону. Для установления закона следует щелкнуть мышкой меню «Статистические методы», в подполе «Распределения и частоты» щелкнуть мышкой по кнопке «U-Согласие распределений», после чего на экране появится подокно «Распределения» (рис.2.5).

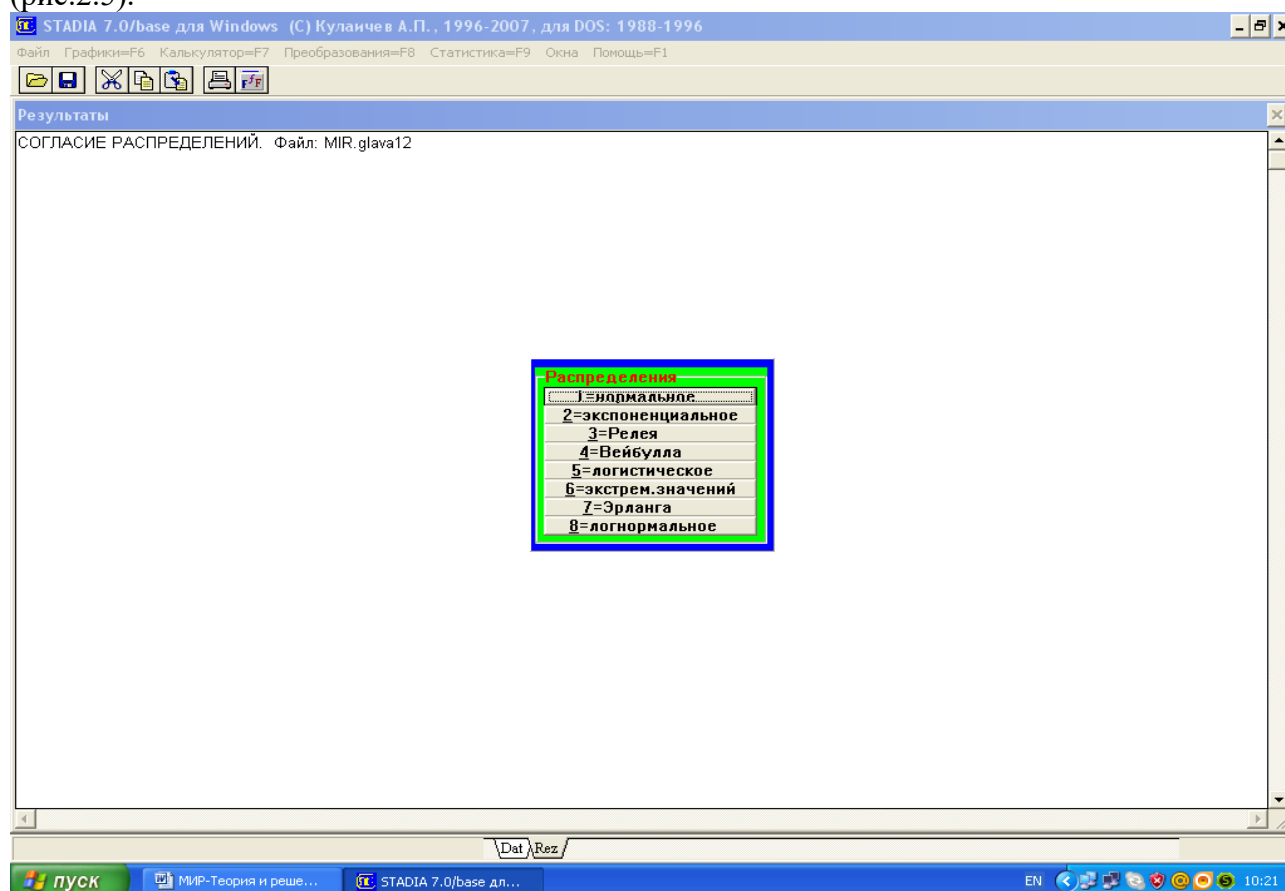


Рис. 2.5. Отображение окна с подокном «Распределения»

Средствами ППСА STADIA можно выполнить проверку следующих законов распределения: нормальное, экспоненциальное, Релея, Вейбулла, логистическое,

экстремальных значений, Эрланга, логнормальное. Это наиболее часто встречающиеся в практике законы распределения, хотя их в математической статистике уже гораздо больше.

В диалоговом окне законов распределения выбрать путем последовательной проверки и анализа распределений наиболее адекватный вид распределения для исследуемой выборки. Подбор закона распределения проводится посредством визуализации и рассмотрения статистических оценок. Визуально подбирается наиболее близкий закон по мере расстояния центра интервальных «сгустков» случайных величин от кривой плотности вероятности распределения. Априори можно предположить, что распределение значений рассматриваемой выборки будет подчиняться закону Пуассона.

Как правило, проверку следует начать с нормального распределения для чего кликнуть мышкой на кнопке «нормальное», после чего получим график (рис.2.6).

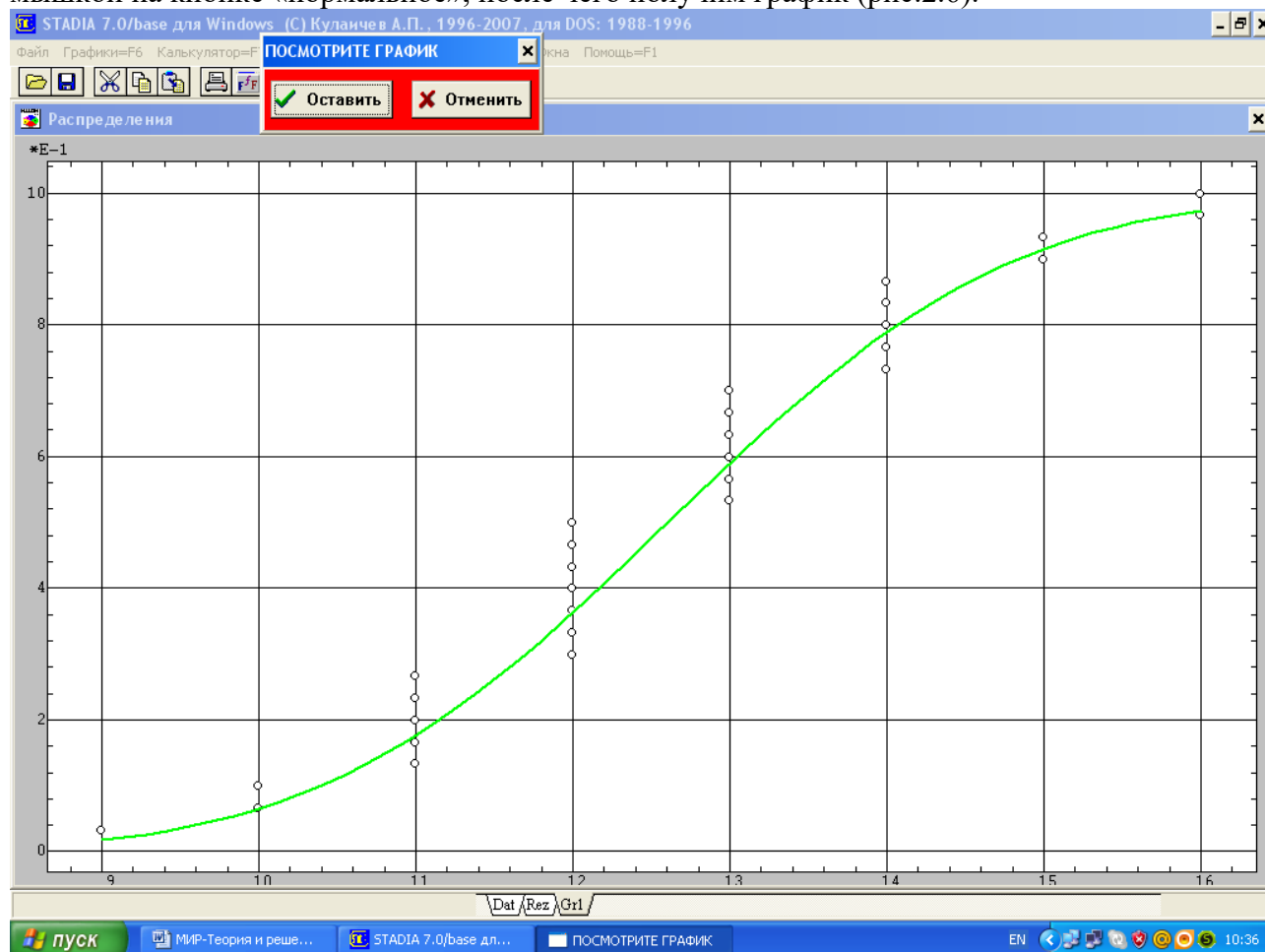


Рис. 2.6. График проверки данных на соответствие нормальному закону

На графике линия отображает теоретические параметры закона нормального распределения. Расположенные вертикально кружочки отображают интервальное распределение исследуемых нами фактических экспериментальных (эмпирических) данных. График довольно чётко показывает совпадение теоретической функции с эмпирической функцией распределения. При этом совпадение имеется как по центральной части гистограммы, так и на её «хвостах», то есть левая и правая части распределения. Более точно заключение дадим ниже на основе результатов расчета статистических параметров распределения.

Для дальнейшего отображения графика на рабочем листе Gr1 электронной таблицы в верхнем окне «ПОСМОТРИТЕ ГРАФИК» кликнуть мышкой кнопку «Оставить», после чего в рабочем листе «Результаты» появится отображение результатов расчетов по проверке соответствия экспериментальных данных и параметров закона нормального распределения (рис.2.7).

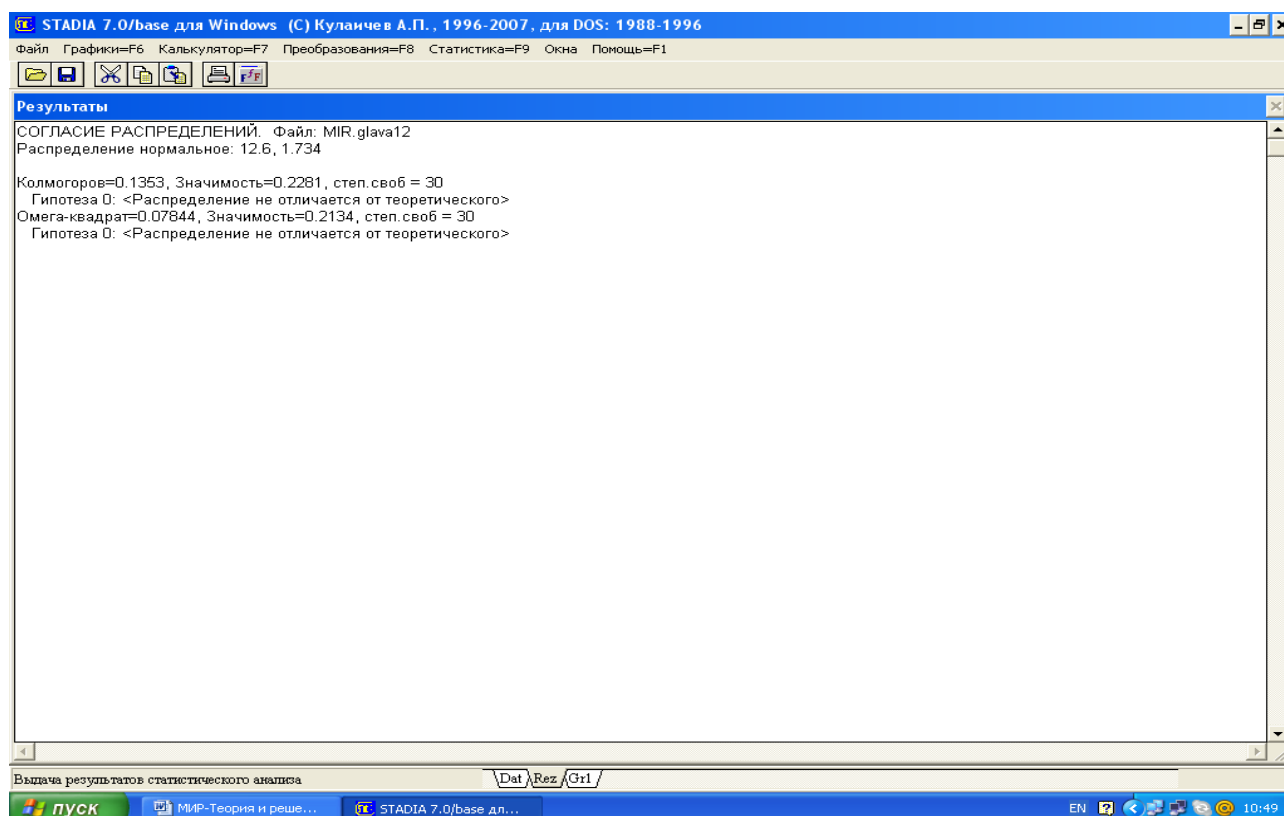


Рис.2.7. Подокно «Результаты» с отображением оценок соответствия данных закону нормального распределения

Под рубрикой «Распределение нормальное» даны два параметра: среднее выборочное значение (в данном случае время обнаружения и исправления дефекта по полноте данных) - 12.6 мин., а также оценка по ошибке среднего выборочного – 1.734, следует сказать - довольно незначительная ошибка.

В соответствии с нулевой гипотезой тот факт, что распределение экспериментальных данных не отличается от теоретического закона нормального распределения данных подтверждается оценками по двум критериям – Колмогорова и Омега-квадрат. Содержание этих критериев прокомментировано выше.

Для проверки распределения наших эмпирических данных на согласие с другим законом вернёмся в окно с подокном «Распределение» (рис. 2.5). Выполним проверку на согласие следующему по списку в подокне «Распределение» закону экспоненциального распределения для чего нужно кликнуть мышкой по кнопке «2=экспоненциальное» и получим соответствующий график (рис. 2.8).

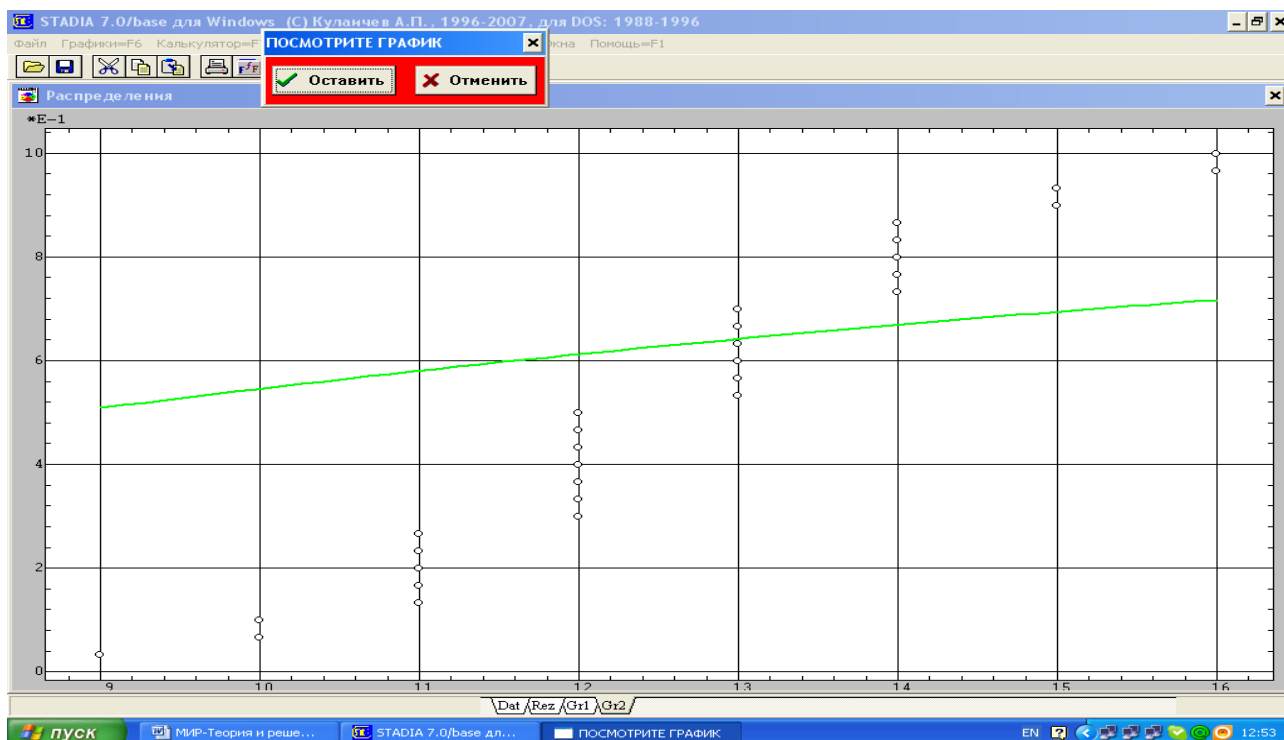


Рис. 2.8. График проверки данных на соответствие экспоненциальному закону

По графику видно, что теоретическая линия закона экспоненциального распределения пересекается только с одним из средних интервалов значений эмпирических данных. Наблюдается резкое расхождение на «хвостах» гистограммы. Априори можно заключить, что согласие в данном случае маловероятно.

Чтобы оставить график кликом мышки кнопку «Оставить» в подокне «ПОСМОТРИТЕ ГРАФИК». Этот график в последующем при необходимости можно отобразить посредством кнопки рабочего листа «Gr2». Одновременно с выдачей графика в рабочем листе электронной таблицы «Rez» выводятся результаты расчетов статистических оценок по закону экспоненциального распределения. Кликнув мышкой по кнопке «Rez» получим отображение указанных результатов (рис. 2.9.)

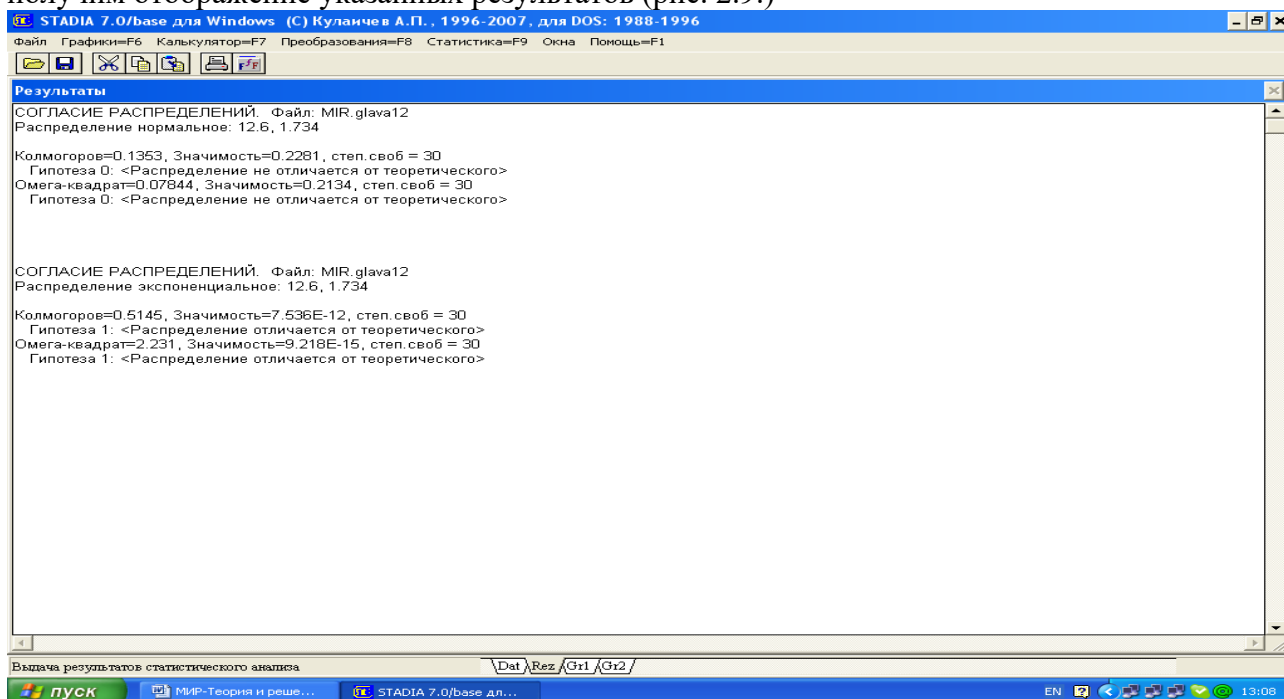


Рис. 2.9. Окно с содержанием оценок по согласию данных с экспоненциальным распределением

При рассмотрении результатов расчета параметров согласия распределения наше предположение, выполненное по анализу графика, об отсутствии согласия подтверждается. В соответствии с проверкой по первой гипотезе программа выдала заключение «Распределение отличается от теоретического», что означает – распределение эмпирических экспериментальных данных не соответствует экспоненциальному закону распределения, что подтверждается также значениями оценок по критериям Колмогорова и Омега-квадрат.

В соответствии с представленным способом проверки согласия следует выполнить проверку по остальным законам распределения. В практическом отношении это расширяет, как правило, представление исследователя о моделируемом информационном ресурсе и принять в дальнейшем правильное решение о стратегии и тактике моделирования.

Выводы. По окончании проверки по всем законам распределения следует перейти к формулированию выводов. В выводах должна содержаться квинтэссенция решения данной задачи.

Оформление отчета о работе. По окончании предусмотренного объема выполненной работы необходимо перейти к составлению и оформлению отчета. По окончании составления отчета необходимо предъявить его преподавателю для контроля качества решения задачи и получения зачета.

Контрольные вопросы и задания

1. Провести проверку согласия распределения экспериментальных данных по полному составу законов распределения, имеющихся в STADIA.
2. Для чего необходима проверка согласия экспериментальных данных в решении задач моделирования?
3. В какой последовательности выполняются этапы проверки согласия?
4. Какие параметры привлекаются для анализа проверки согласия?
5. Каким образом проводится анализ проверки согласия?

Практическая работа № 3. Решение задачи «Оценка эффективности модернизации стандартизации ИТ»

3.1. Краткая постановка задачи

Основной целью решения задачи является получение практических навыков в моделировании процесса проверки эффективности (результативности) мероприятий по улучшению стандартизации ИТ, а также закрепление теоретических сведений, полученных в рамках лекционного курса.

Решение задачи базируется на применении средств системного анализа и математической статистики. Методика выполнения работы состоит в том, чтобы на основе собранных статистических данных выполнить анализ функционального состояния ИТ программными средствами диалоговой системы STADIA. На основе результатов сравнительного анализа вырабатывается последующая стратегия и тактика улучшения качества работы ИТ и, как правило, выбор эталонной (стандартной) модели ИТ.

Допустим, что в результате реализации организационно-технических мероприятий проведена модернизация технологического процесса обработки данных ИТ. Требуется проверить наличие положительного эффекта от модернизации ИТ. Если указанный эффект наблюдается, то необходимо также узнать насколько улучшились функциональные параметры ИТ. Сравнение двух выборок по параметру полноты документной информации ИТ проведём на основе критерия интегральных различий Колмогорова-Смирнова.

По условию решения задачи необходимо в качестве исходных данных взять две выборки по какому-либо определенному параметру, в нашем примере это будет полнота документной информации, обрабатываемой на этапах технологического процесса обработки данных ИТ. В каждой выборке имеется по 30 значений. Первая выборка x_1 проведена до модернизации, а вторая выборка x_2 - после модернизации технологического процесса обработки данных ИТ. Числа первой выборки указаны в тексте раздела 2.1, а числа выборки x_2 следующие: 5, 9, 5, 12, 10, 6, 11, 9, 7, 8, 7, 9, 7, 6, 7, 8, 4, 6, 8, 5, 9, 6, 10, 7, 9, 7, 10, 9, 6, 8.

Задача состоит из основных этапов:

1. Ввод исходных данных.
2. Преобразование данных к рабочему виду.
3. Проверка и анализ интегральных различий двух выборок.
4. Формулирование выводов.
5. Составление и оформление отчета о решении задачи.

3.2. Основные этапы решения задачи

Ввод исходных данных. Вводим данные первой и второй выборки соответственно в столбцы x_1 и x_2 электронной таблицы, формируя тем самым две одноимённые переменные. Ввод осуществляется в соответствии с регламентом ввода данных в электронную таблицу, рассмотренным ранее в разделе 1.2. В результате ввода получим набор чисел по двум переменным x_1 и x_2 (рис.3.1).

Теперь необходимо последовательно получить гистограммы по переменным. Кликнуть переменную x_1 , затем кликнуть правую стрелку, при появлении переменной x_1 в поле «Для анализа» кликнуть кнопку «Утвердить». После этого на экране появится подокно «Гистограмма», идентичное описанному ранее (раздел 2.2.). Необходимо в подокне «Гистограмма» кликнуть кнопку «Утвердить» и появится отображение результатов расчета по x_1 и подокно «Confirm», в котором на вопрос «Сохранить гистограмму в матрице данных?» следует кликнуть кнопку «Утвердить». В результате отобразится гистограмма по переменной x_1 , ранее уже нами полученной (см. рис. 1.4).

Далее для получения данных по переменной x_2 необходимо повторить выполненные ранее процедуры с данными по переменной x_2 . При условии получения данных по обеим переменным можно выполнить визуальную проверку эффективности путем сопоставления двух гистограмм. Если у переменной x_1 минимальное значение дефекта по полноте начинается с 9 минут (левый «хвост» гистограммы), то у переменной x_2 этот параметр меньше – 4 минуты. Подобное условие наблюдается и по правым «хвостам» гистограмм, соответственно у переменной x_1 максимальное значение дефекта равно 16 минутам против 12 минут у переменной x_2 . Отсюда можно сделать предварительное заключение, что в результате модернизации качество технологического процесса обработки данных ИТ улучшилось. Подтверждением нашему заключению могут служить также полученные в результате расчетов значения критериев Колмогорова, Омега-квадрат и Хи-квадрат, - у переменной x_2 указанные значения меньше, чем у переменной x_1 .

Вместе с тем для более достоверного заключения необходимо привлечь к сравнительному анализу и другие более информативные статистические оценки. Для этого необходимо получить оценки описательной статистики по каждой переменной в соответствии с порядком, рассмотренным в разделе 11.2 (см.рис. 1.1-1.8). В результате расчетов получим оценки описательной статистики (рис.3.3). Далее необходимо выполнить проверку и анализ интегральных различий двух выборок.

Проверка и анализ интегральных различий двух выборок. Для получения результатов необходимо через меню «Статистика» в поле «Непараметрические тесты» щелкнуть мышкой по опции «Произвольных альтернатив». В появившемся поле «Анализ переменных» выбрать переменные x_1 и x_2 , для чего переслать их в правое подполе и затем щелкнуть мышкой по опции «Утвердить», в результате чего получим дополнительные результаты расчетов по критерию Колмогорова-Смирнова (рис. 3.4).

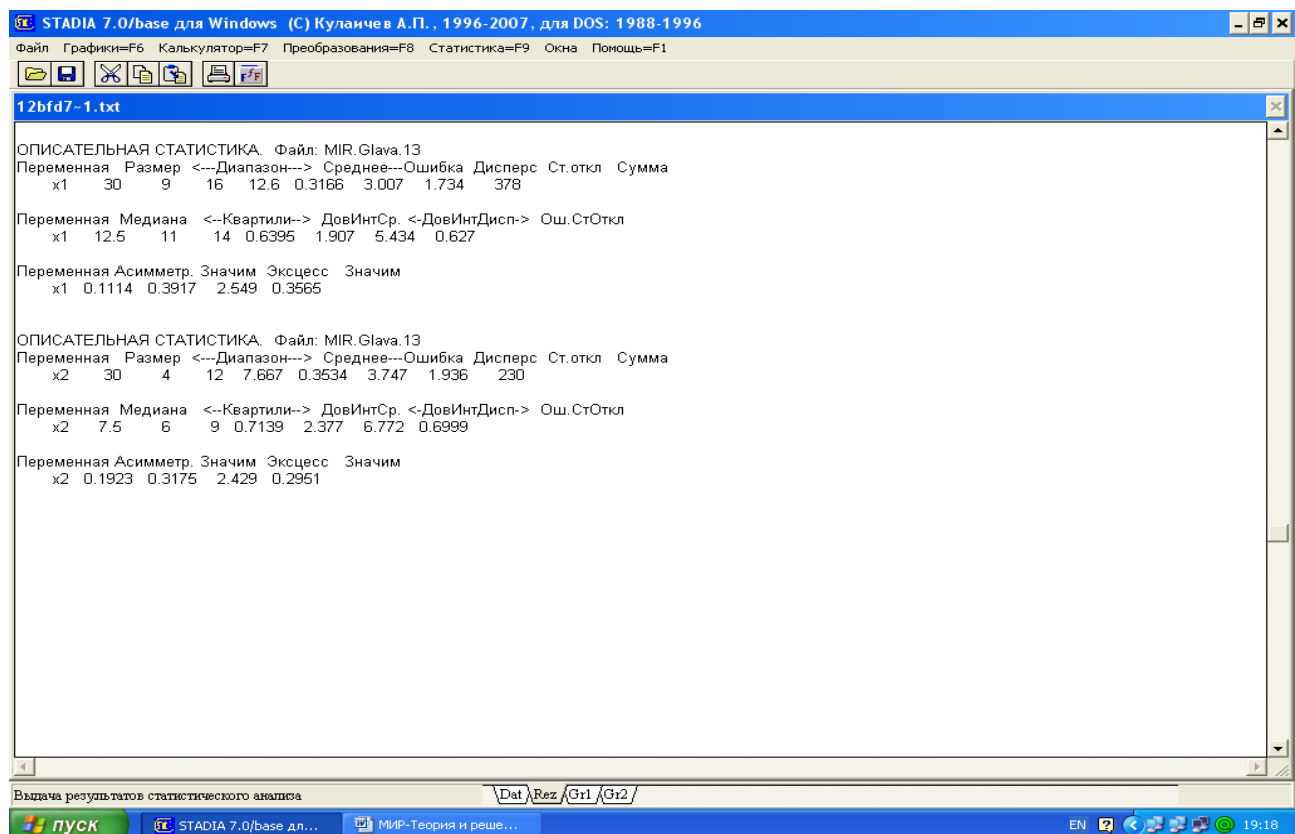


Рис.3.3. Отображение результатов расчета описательной статистики

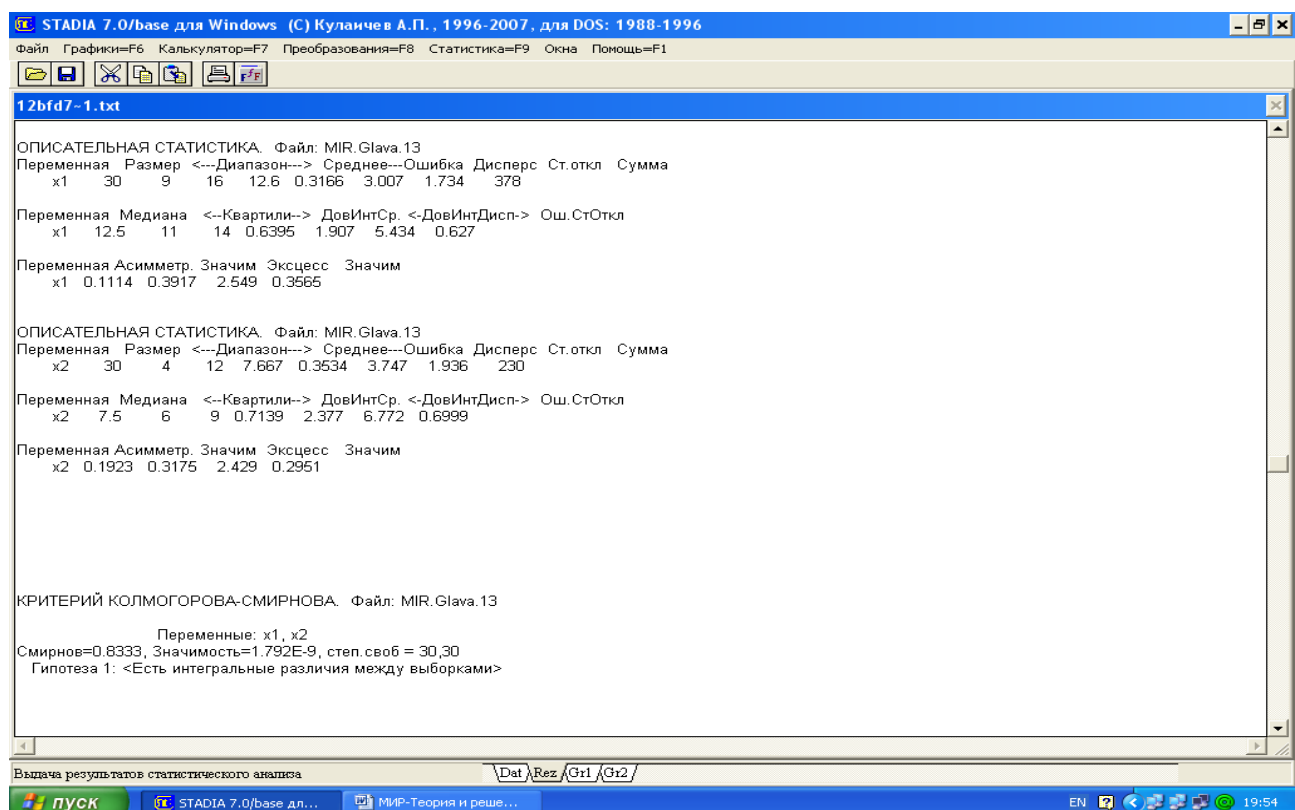


Рис. 3.4. Отображение дополнительных расчетов по критерию Колмогорова-Смирнова

В соответствии с критерием Колмогорова-Смирнова по параметру «гипотеза 1» программа выдала заключение о том, что «Есть интегральные различия между выборками», то есть подтверждается также наше предположение, выдвинутое в результате визуального сравнения гистограмм двух выборок. Вместе с тем необходимо знать эффективность модернизации

технологии обработки данных ИТ в численном измерении. Для этого обратимся к оценкам описательной статистики, в частности, значению «Среднее выборочное», то есть среднее время обнаружения и исправления дефекта по полноте данных в минутах до модернизации (переменная x_1) и после модернизации (переменная x_2). Если у переменной x_1 это значение составляет 12.6 минуты, то у переменной x_2 – 7.667 минут. Сопоставление этих двух значений показывает, что среднее время обнаружения и коррекции дефекта по полноте данных в технологии ИТ существенно сократилось, то есть на 40%.

Формулирование выводов. Необходимо сформулировать выводы в соответствии с основными этапами решения задачи. Особое внимание обратить на последовательность в применении средств сравнительного анализа.

Составление и оформление отчета. После окончания полного объема выполненных работ составить и оформить отчет о решении задачи. Затем провести контроль качества выполненных работ. При условии выявленных дефектов исправить их и предъявить преподавателю для зачета.

Контрольные работы и задания

1. Для чего необходима оценка эффективности модернизации качества технологии ИТ?
2. В какой последовательности выполняются этапы оценки эффективности?
3. Какие средства привлекаются для оценки эффективности?
4. В какой последовательности выполняются этапы оценки эффективности?
5. Какие параметры применяются для сравнительного анализа при оценке эффективности?

Практическая работа № 4. Решение задачи «Распознавание свойств в управлении стандартизацией ИТ»

4.1. Краткая постановка задачи

Основной целью решения задачи является получение практических навыков в выявлении свойств функционирования ИТ в ходе моделирования и экспериментального исследования стандартизации ИТ, а также закрепление теоретических сведений, полученных в рамках лекционного курса.

Методом исследования является кластерный анализ, реализуемый в нашей задаче средствами программы кластер-анализа ППСА STADIA.

В экспериментальных исследованиях ИТ с целью корректности решения задач моделирования перед исследователем всегда возникает задача выявления свойств моделируемого информационного ресурса. Следует отметить, что любое научное исследование, в том числе и процесс моделирования, невозможно выполнить без распознавания свойств изучаемого объекта. На основе познанных свойств строится работа по исследованию, проектированию, построению и эксплуатации как ИТ в целом, так и их отдельных компонентов. Сложность распознавания свойств заключается в том, чтобы в конкретных условиях выбрать эффективный индикатор. Этот индикатор должен обладать способностью отображения наиболее полного состава свойств функционирования ИТ, с одной стороны, и возможности корректного распознавания состава существенных свойств ИТ таким способом, который не требовал бы привлечения сравнительно объемных ресурсов в распознавании. Очевидно то, что подобная противоречивость усложняет решение задачи.

В данной задаче наиболее эффективным индикатором свойств функционирования ИТ является статистическая структура дефектов функционирования ИТ. Сбор и регистрация дефектов может выполняться по различным компонентам ИТ, например, по этапам технологического процесса обработки данных ИТ. Априори подобная выборка дефектов имеет совокупность свойств, то есть по существу является неоднородной. Устранить неоднородность, значит выявить состав свойств, которыми обладает исследуемый технологический процесс.

Одним из методов решения подобных задач является классификация данных выборки и выявление и/или уточнение существенных свойств, характеризующих технологию ИТ. Разделение данных на классы производится по основанию деления, которое может быть представлено мерой близости (удалённости) каждого события, например, дефекта. Основанием деления может быть значение (набор значений) события, отображаемого по временной, стоимостной и другим шкалам измерения.

В нашем случае данные выборки представляют собой набор дефектов обработки документов. Каждый дефект представлен временем и стоимостью его обнаружения и исправления. Задача заключается в том, чтобы по двум переменным - времени и стоимости провести разделение дефектов на классы и выявить на этой основе существенные свойства статистической структуры дефектов обработки документов. Классификация (группировка) данных выборки будет выполнена в решении данной задачи по критерию «мера близости (расстояния)» с применением средств кластерного анализа ППСА STADIA. В решении задач кластеризации имеется несколько «мер близости» [2], в нашей задаче наиболее приемлемым представляется «евклидово расстояние».

Решение задачи выполним с привлечением данных выборки дефектов по двум переменным каждая из которых состоит из 20 чисел. Нижеуказанные переменные x_1 и x_2 отображают соответственно время и стоимость обнаружения и исправления дефектов в минутах и копейках.

x_1 : 3, 11, 2460, 4, 13, 2952, 5, 14, 3444, 3, 12, 2450, 4, 13, 2952, 3, 12, 3440, 2, 15.

x2: 40, 143, 31900, 52, 169, 38100, 72, 185, 44700, 40, 156, 31800, 54, 165, 38100, 39, 156, 44500, 26, 190.

Решение задачи выполняется по следующим основным этапам:

1. Ввод данных статистической структуры дефектов.
2. Выполнение кластерного анализа статистической структуры дефектов.
3. Анализ результатов кластеризации.
4. Формулирование выводов по результатам решения задачи
5. Составление и оформление отчета по решению задачи.

4.2. Основные этапы решения задачи

Ввод данных статистической структуры дефектов. После загрузки ППСА STADIA посредством клавиатуры вводим данные в электронную таблицу. Значения каждой переменной вводятся в соответствующие столбцы x1 и x2 электронной таблицы. Необходимо обеспечить точно такую же последовательность данных, в какой они даны в разделе 4.1. Если будет нарушена последовательность, то это повлечет некорректное решение задачи кластер-анализа. Данные в электронной таблице должны иметь следующий вид (рис. 4.1).

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10	x11	x12	x13
1	3	35											
2	12	290											
3	2460	30115											
4	4	45											
5	13	320											
6	2952	38100											
7	5	70											
8	14	340											
9	3444	42100											
10	3	33											
11	12	290											
12	2450	30110											
13	4	55											
14	13	320											
15	2952	37200											
16	2	26											
17	12	250											
18	3440	42100											
19	2	28											
20	12	280											
21													
22													
23													
24													
25													
26													
27													
28													
29													
30													
31													
32													
33													

Рис. 4.1. Отображение данных переменных x1 и x2

Выполнение кластерного анализа статистической структуры дефектов. Классификация статистической структуры дефектов проводится в следующей последовательности.

Кликнуть меню «Статистика», затем в открывшемся подокне «Статистические методы» в подполе «Многомерные методы» кликнуть кнопку «Q-кластерный». После чего появится подокно «Исходные данные» (рис.4.2).

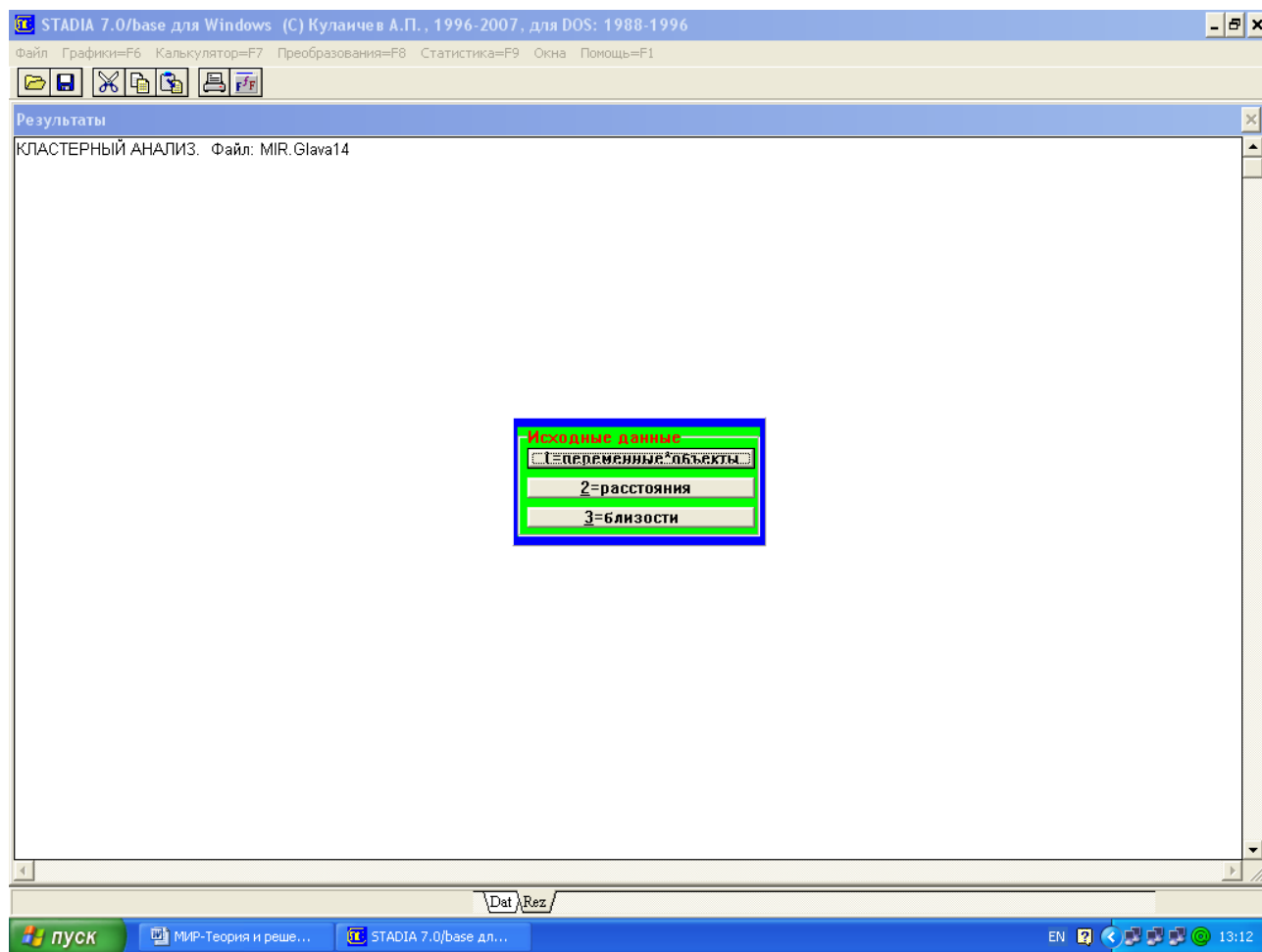


Рис.4.2. Отображение подокна «Исходные данные»

Далее в меню «исходные данные» нужно кликнуть клавишу «1= переменные объекты». В появившемся меню «Метрика вычисления расстояний» (рис.4.3). Нужно кликнуть кнопку «1=Эвклид» после чего появится подокно «Объединяющая» (рис.4.4).

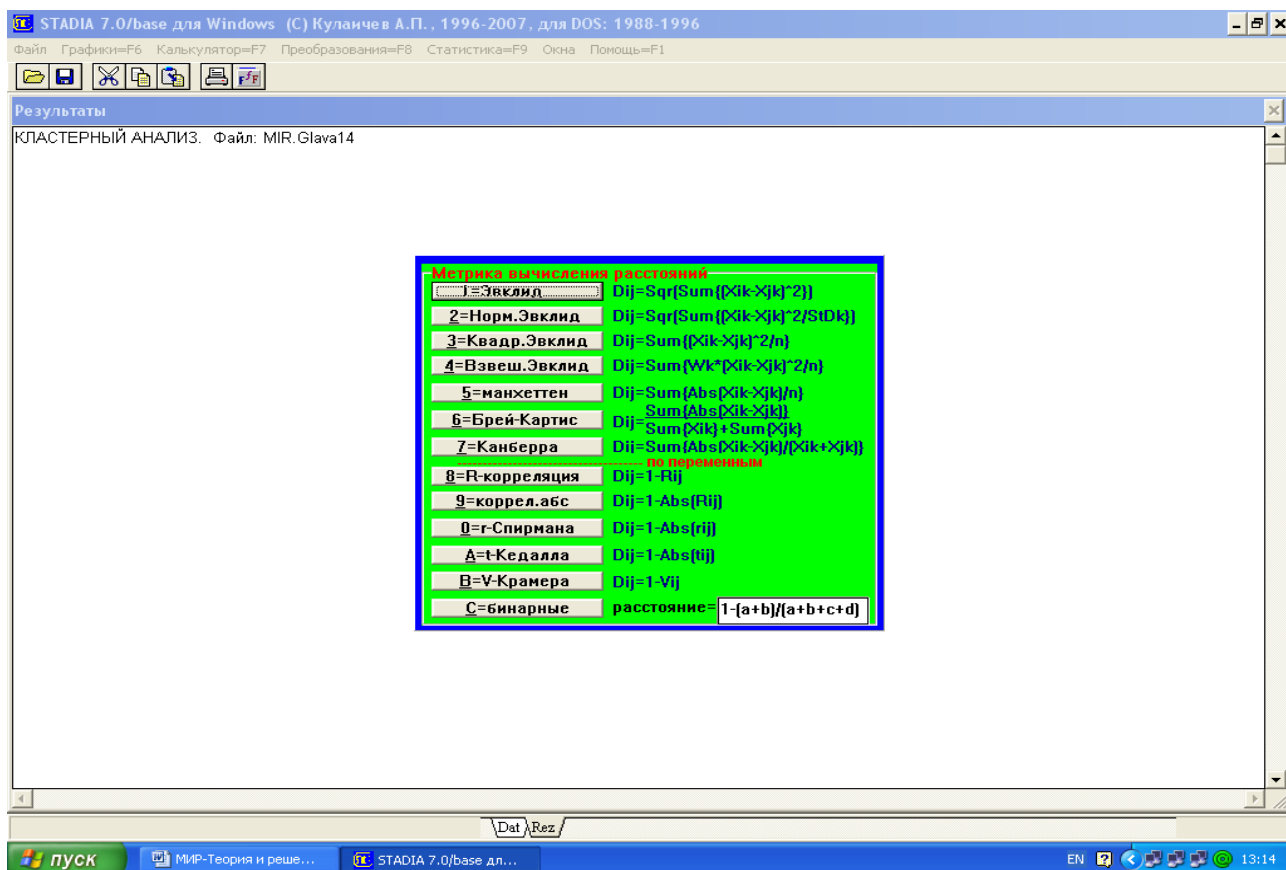


Рис.4.3. Отображение подокна «Метрика вычисления расстояний»

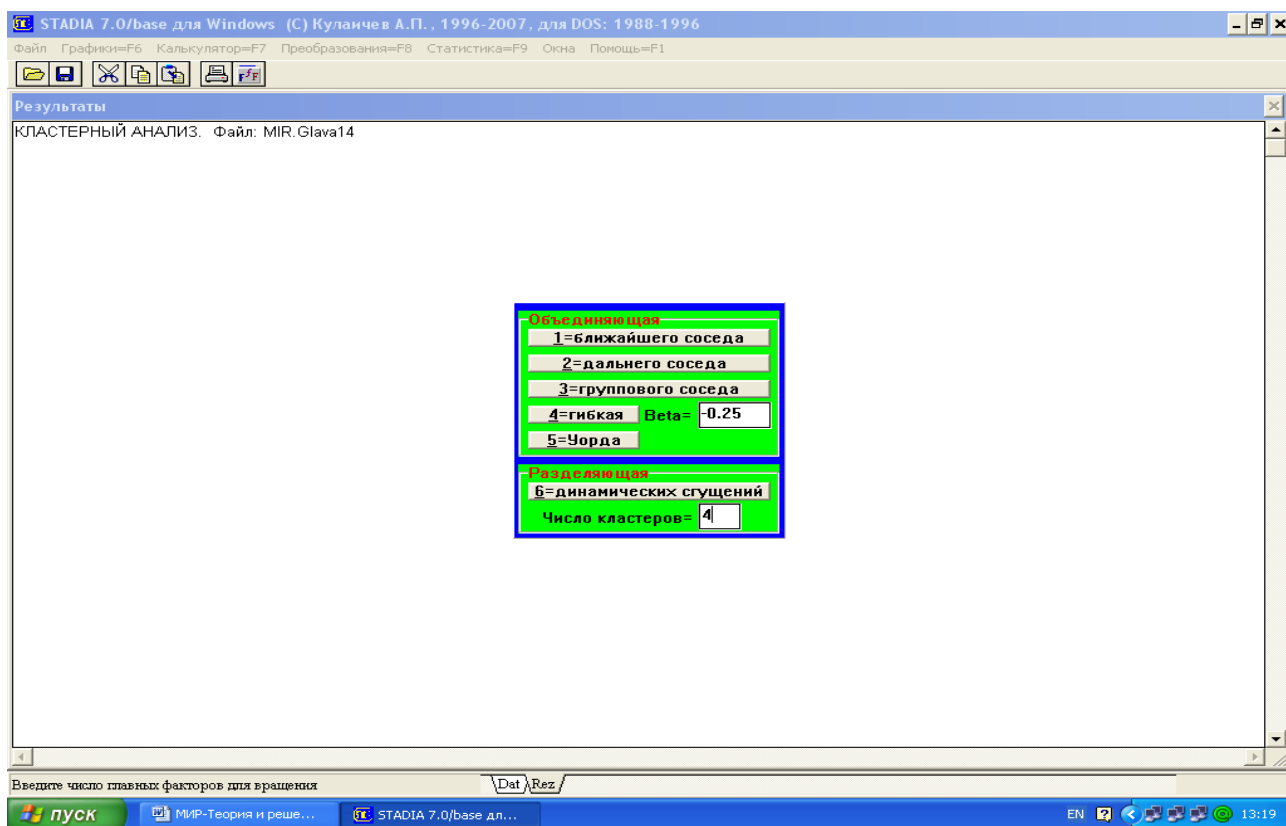


Рис.4.4. Отображение подокна «Объединяющая»

В случае необходимости гибкой стратегии кластеризации в этом же меню необходимо предварительно установить (оставить) коэффициент Beta= - 0.25. Поскольку в нашей задаче

априори неизвестно количество кластеров, задача именно и заключается в их определении, то число кластеров в данном меню не указывается. В субполе «Число кластеров» нужно цифру, если она имеется, заменить на цифру 0. Затем в меню «Объединяющая» нужно кликнуть кнопку «1=ближайшего соседа», после чего на экране появится подполе «Confirm» с вопросом «Выдать таблицу расстояний?» (рис.4.5).

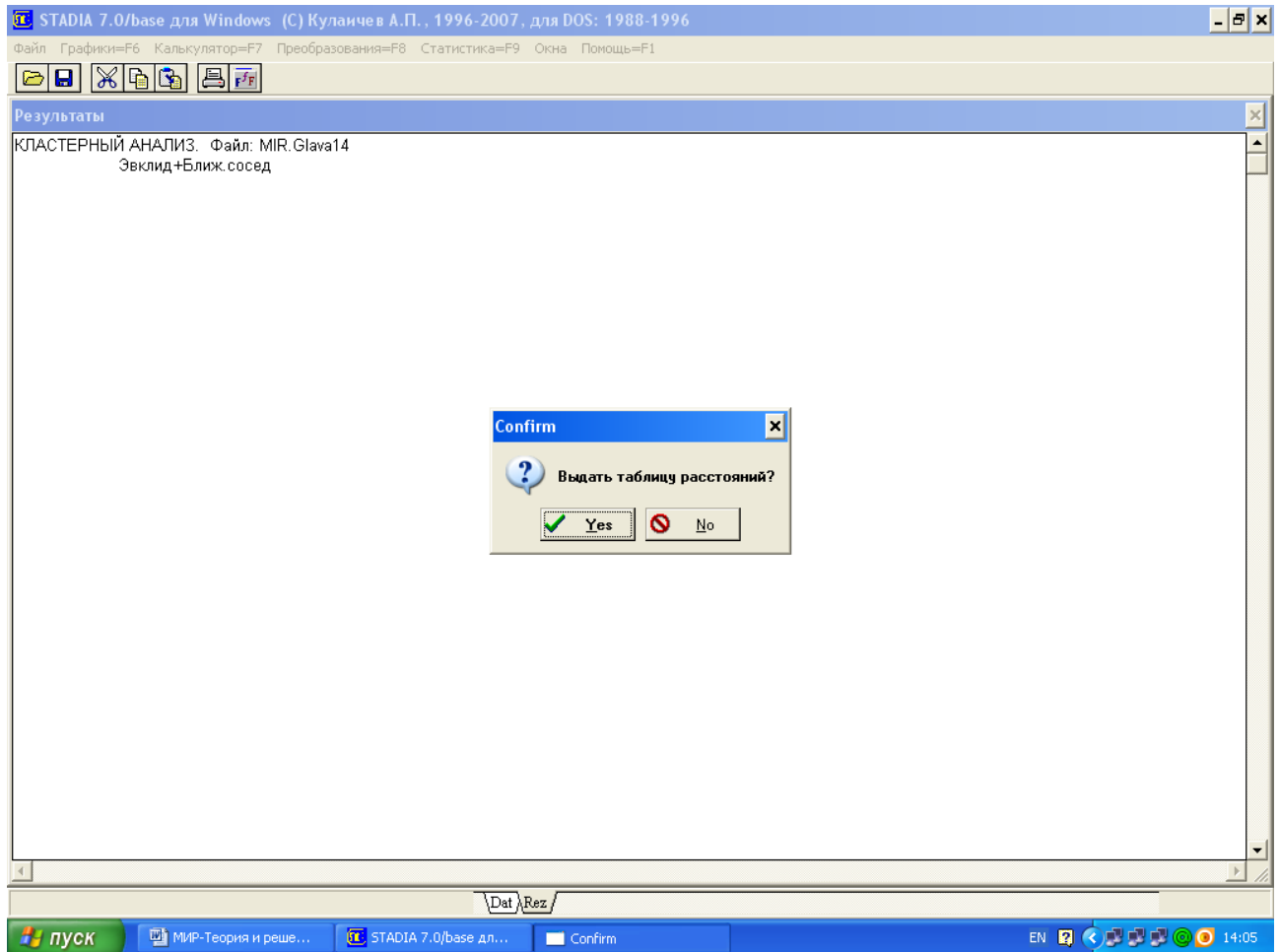
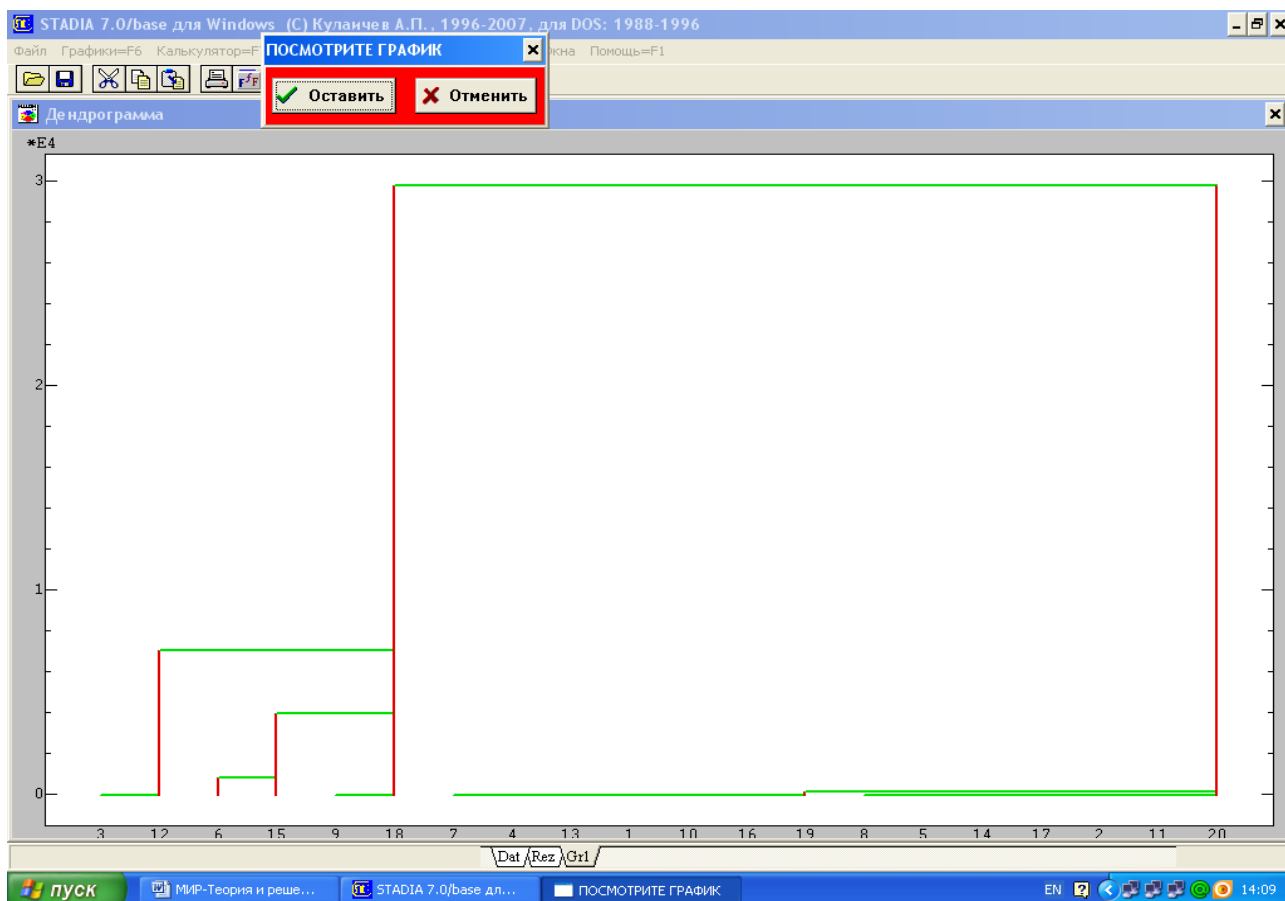


Рис.4.5. Отображение подполя «Confirm»

В подполе «Confirm» необходимо кликнуть кнопку «Yes», после чего будет выдан график, отражающий результаты кластеризации статистической структуры дефектов (рис. 4.6). График в виде дендрограммы отражает объединение кластеров, в нашем случае групп дефектов, обладающих однородными свойствами. По горизонтальной оси указаны порядковые номера дефектов, которые они получили при вводе, - это номера строк электронной таблицы (рис.4.1). По вертикальной оси указаны значения расстояний между дефектами в соответствии со шкалой расстояний. Значения расстояний вычисляются посредством программы кластер-анализа.



4.6. График (дендрограмма) результатов кластеризации дефектов

В подокне «ПОСМОТРИТЕ ГРАФИК» необходимо кликнуть кнопку «Оставить». В состав результатов кластер-анализа входит также и результаты расчетов в виде «Таблицы расстояний» (рис. 4.7).

На этом же экране, ниже «Таблицы расстояний» под рубрикой «Кластеры» указывается список объектов, в нашем случае это номера дефектов (рис.4.8).

STADIA 7.0/base для Windows (C) Кулаичев А.П., 1996-2007, для DOS: 1988-1996

Файл Графики=F6 Калькулятор=F7 Преобразования=F8 Статистика=F9 Окна Помощь=F1

Результаты

КЛАСТЕРНЫЙ АНАЛИЗ. Файл: MIR.Glava14
Эвклид+Ближ.сосед

Таблица расстояний

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
(2) 265.2									
(3) 3.018E4	2.993E4								
(4) 10.05	245.1	3.017E4							
(5) 285.2	30.02	2.99E4	275.1						
(6) 3.818E4	3.792E4	8000	3.817E4	3.789E4					
(7) 35.06	220.1	3.015E4	25.02	250.1	3.814E4				
(8) 305.2	50.04	2.988E4	295.2	20.02	3.787E4	270.1			
(9) 4.221E4	4.195E4	1.203E4	4.22E4	4.192E4	4030	4.217E4	4.19E4		
(10) 2	257.2	3.018E4	12.04	287.2	3.818E4	37.05	307.2	4.221E4	
(11) 255.2	0	2.993E4	245.1	30.02	3.792E4	220.1	50.04	4.195E4	
(12) 3.017E4	2.992E4	11.18	3.016E4	2.989E4	8006	3.014E4	2.987E4	1.203E4	
(13) 20.02	235.1	3.016E4	10	265.2	3.816E4	15.03	285.2	4.219E4	
(14) 285.2	30.02	2.99E4	275.1	0	3.789E4	250.1	20.02	4.192E4	
(15) 3.728E4	3.703E4	7102	3.727E4	3.7E4	900	3.725E4	3.698E4	4925	
(16) 9.055	264.2	3.019E4	19.1	294.2	3.819E4	44.1	314.2	4.221E4	
(17) 215.2	40	2.987E4	205.2	70.01	3.796E4	180.1	90.02	4.199E4	
(18) 4.221E4	4.195E4	1.203E4	4.22E4	4.192E4	4924	4.221E4	4.199E4		
(19) 7.071	262.2	3.019E4	17.12	292.2	3.819E4	42.11	312.2	4.221E4	
(20) 245.2	10	2.994E4	235.1	40.01	3.793E4	210.1	60.03	4.196E4	
	247.2	10	2.993E4	225.1	40.01	3.704E4	254.2	30	4.196E4

ПУСК ММР-Теория и реше... STADIA 7.0/base дл... EN 14:15

Рис.14.7. Отображение «Таблицы расстояний» кластер-анализа дефектов

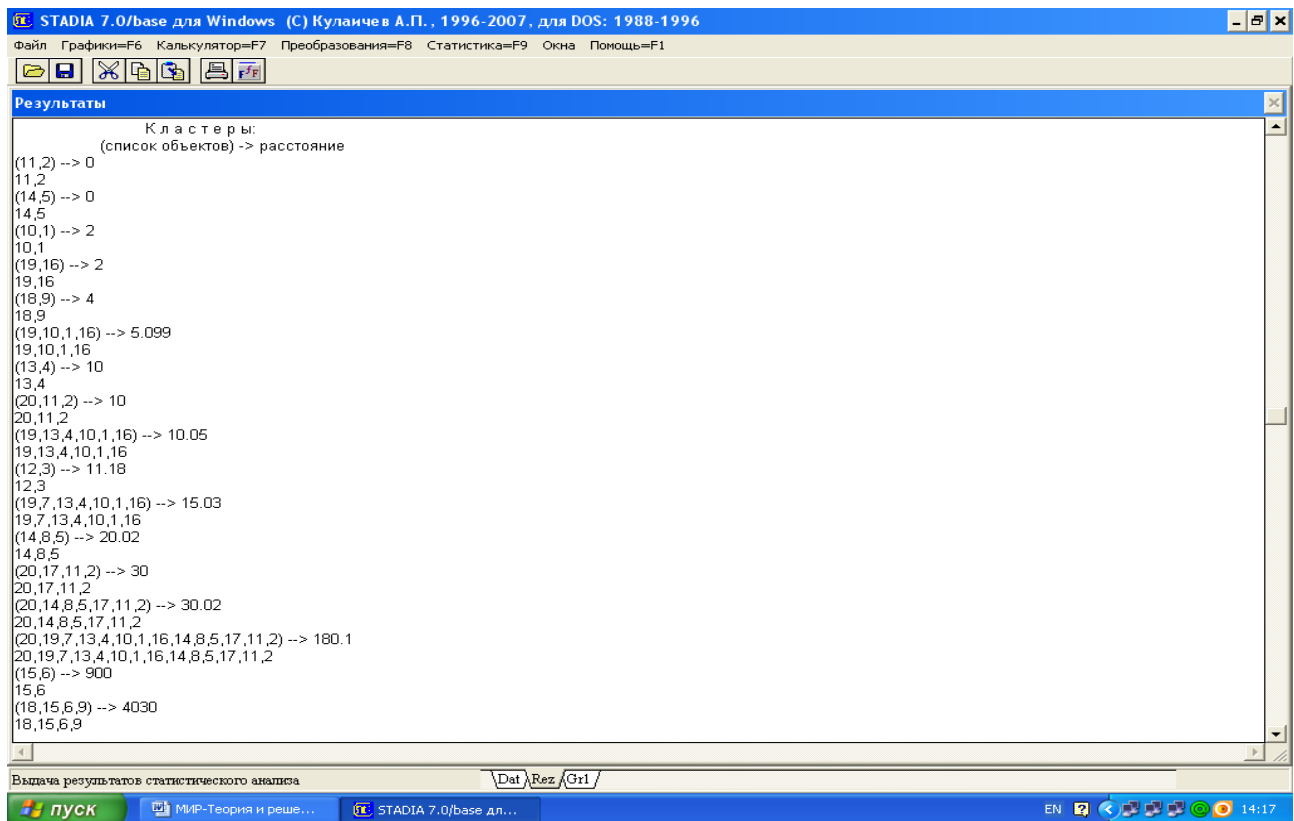


Рис.4.8. Отображение порядка объединения дефектов в кластеры

Здесь указываются данные по величинам среднего внутрикластерного расстояния, по которому можно сравнивать различные варианты кластеризации текущих данных, и найденные кластеры с порядковыми номерами входящих в каждый кластер объектов (дефектов).

Анализ результатов кластеризации. Как видно по графику стратегия «ближайшего соседа» достаточно отчетливо выделяет три группы (класса) дефектов обработки документов. На минимальном расстоянии (примерно 15.03) произошло объединение дефектов 7, 4, 13, 1, 10, 16 и 19. Затем на расстоянии примерно 30.02 произошло объединение дефектов 8, 5, 14, 17, 2, 11 и 20, после чего на расстоянии примерно 4030 произошло объединение дефектов 3, 12, 6, 15, 9 и 18. Зная нумерацию и характер дефектов, которые зарегистрированы в «Ведомости дефектов», можно обозначить эти три группы. В нашем случае дефекты с минимальными значениями времени и стоимости были сформированы прежде всего и составили первый кластер. При идентификации этих дефектов установлено, что они характеризуются наличием искажений символов (цифр, букв) в реквизитах документов, допущенных при выполнении технологических операций с документами. Второй кластер дефектов составили пропуски, или отсутствие элементов документов (реквизитов-оснований), являющихся регламентом структуры того или иного документа. Третий кластер составили дефекты, характеризующиеся опозданием, несвоевременным представлением документов от этапа к этапу технологического процесса обработки данных, то есть нарушение регламентных сроков представления документов. Соответственно эти три сформированных кластера дефектов можно обозначить как дефекты по достоверности, полноте и своевременности. Таким образом, выделяются три наиболее существенных признака (свойства) функционирования изучаемой ИТ в технологии обработки документов. Следует отметить, что при применении кластер-анализа в других ИТ эти кластеры могут иметь и другие свойства, отличные от полученных нами.

Более конкретное рассмотрение может быть проведено в направлении установления меры близости (удаленности) кластеров по шкале расстояний оси ординат графика дендрограммы. Так, например, дефекты достоверности отличаются наименьшим значением разброса, на втором месте дефекты по полноте и далее дефекты по своевременности. Если анализировать меру близости кластеров, то следует отметить, что класс дефектов по полноте имеют меньшую удаленность от класса дефектов по достоверности, чем от класса дефектов

по своевременности. Это объясняется, прежде всего, значениями переменных по времени и стоимости обнаружения и исправления дефектов. Если дефекты по достоверности и полноте в организационно-техническом отношении могут быть сравнительно легко локализованы и устранены, то дефекты по своевременности намного труднее быстро исправить в силу того, что для их исправления необходим более широкий спектр организационно-технических мероприятий.

Формулирование выводов по результатам решения задачи. В выводах необходимо сформулировать место и значение анализа статистической структуры данных выборки на неоднородность. Резюмировать какие пути и инструментальные средства могут быть использованы исследователем для решения этой задачи. Следует указать также - какие критерии классификации данных выборки могут быть применимы и основные этапы кластеризации.

Составление и оформление отчета. После выполнения полного объема работ следует составить отчет по решению задачи и представить его преподавателю для получения зачета.

Контрольные вопросы и задания

1. Выполните решение задачи по распознаванию свойств функционирования ИТ посредством программы кластер-анализа по условию «Уорда» (см. рис. 4.4) с комментированием результатов.
2. Для чего необходимо распознавание свойств в моделировании информационных ресурсов?
3. В чем заключается сущность кластерного анализа?
4. В какой последовательности выполняются основные этапы кластерного анализа?
5. Каким образом комментируются результаты кластер-анализа и определяются свойства моделируемого информационного ресурса?

Практическая работа № 5. Решение задачи «Прогнозирование в управлении стандартизацией ИТ»

5.1. Краткая постановка задачи

Целью решения данной задачи по стандартизации ИТ является получение практических навыков в построении модели прогнозирования функционирования ИТ, а также закрепление теоретических сведений, полученных в рамках лекционного курса. Метод решения задачи базируется на регрессионном анализе. В качестве программного средства регрессионного анализа привлекается соответствующая программа ППСА STADIA.

Во многих практических задачах экспериментального исследования ИТ применяются модели, которые позволяют провести анализ функциональной зависимости между некоторыми переменными. Такие зависимости можно изучить при помощи регрессионных моделей. Методы регрессионного анализа позволяют:

- производить расчет различного вида регрессионных моделей с определением значений характеристик моделей, в частности, расчета коэффициентов при независимых переменных в уравнении регрессионной модели;
- проверить гипотезу адекватности модели имеющимся наблюдениям;
- применить модель для предсказания или прогнозирования значений зависимой переменной при условии появления новых, или до сего времени ненаблюдаемых значений отдельной независимой переменной, или их полного состава;
- определить прогноз потребности в объемах необходимых ресурсов для решения перспективных задач ИТ.

Наиболее применяемыми на практике, простыми для построения и анализа являются однопараметрические и многопараметрические линейные модели.

Допустим, что в результате анализа установлена функциональная зависимость между запаздыванием документов на вход ИТ (независимая переменная x) и производительностью ИТ (зависимая переменная Y). Запаздывание документов измеряется в минутах, а производительность ИТ измеряется в документе-днях, то есть количество документов, которое ИТ способна выдавать абонентам в течение рабочего дня.

Задача состоит в том, что, требуется построить однопараметрическую линейную регрессионную модель, на основе которой можно будет прогнозировать производительность ИТ в задачах управления качеством в зависимости от тех или иных условий запаздывания документов в технологии ИТ.

Решение задачи предусматривает следующие основные этапы:

1. Подготовка исходных данных.
2. Ввод исходных данных.
3. Обработка данных.
4. Анализ полученных результатов.
5. Формулирование выводов.
6. Составление отчета о решении задачи.

5.2. Основные этапы решения задачи

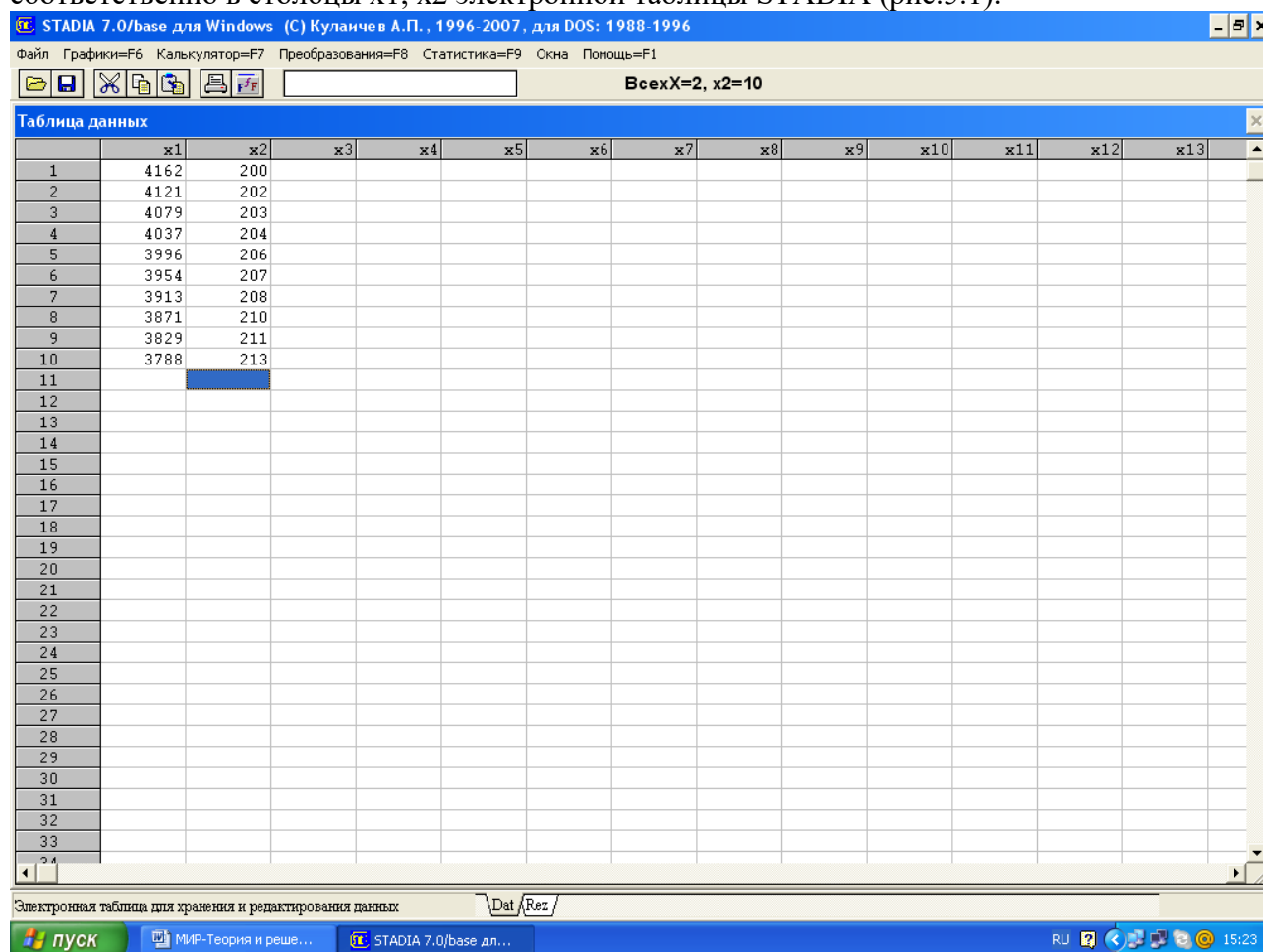
Подготовка исходных данных. Исходные данные представляются в виде матрицы, содержащей одну независимую переменную x и зависимую переменную Y , каждая переменная содержит 10 значений (таблица 5.1). Значений может быть и больше. Каждый исследователь в практической работе применяет данные по своей ИТ, имеющей различные характеристики. Расчет значений независимой (независимых) и зависимой переменных выполняется по

определенной методике. Эта методика независима от специфических характеристик моделируемой ИТ. Для расчета матрицы исходных данных таблица 5.1, содержащей результаты расчета значений двух переменных x и Y , следует ввести данные.

Таблица 5.1.

Пере- менные	Значения переменных									
x	4162	4121	4079	4037	3996	3954	3913	3871	3829	3788
Y	200	202	203	204	206	207	208	210	211	213

Ввод исходных данных. Загрузив ППСА STADIA, вводим данные переменных x и Y соответственно в столбцы $x1$, $x2$ электронной таблицы STADIA (рис.5.1).

Рис.5.1. Отображение исходных данных по переменным x и Y

Обработка данных. Для получения результатов необходимо кликнуть по меню «Статистика», в открывшемся подокне «Статистические методы» в субполе «Регрессионный анализ» кликнуть кнопку «Простая регрессия/тренд» (рис. 5.2).

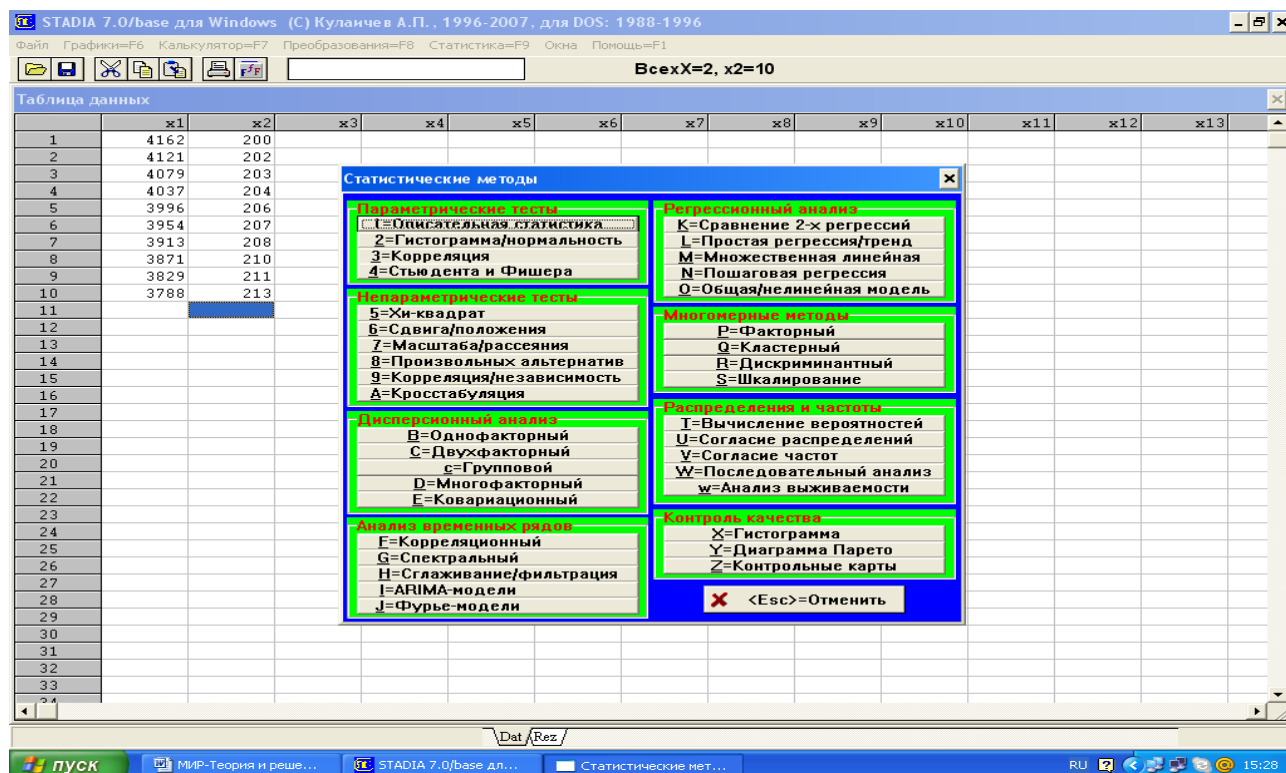


Рис. 5.2. Отображение подокна «Статистические методы»

В появившемся окне «Переменные регрессии» выбрать переменную x1 и поместить в подполе независимой переменной «X-переменные» и переменную x2 в подполе зависимой переменной «Y-переменная», после чего щелкнуть мышкой по опции «Утвердить» (рис.5.3).

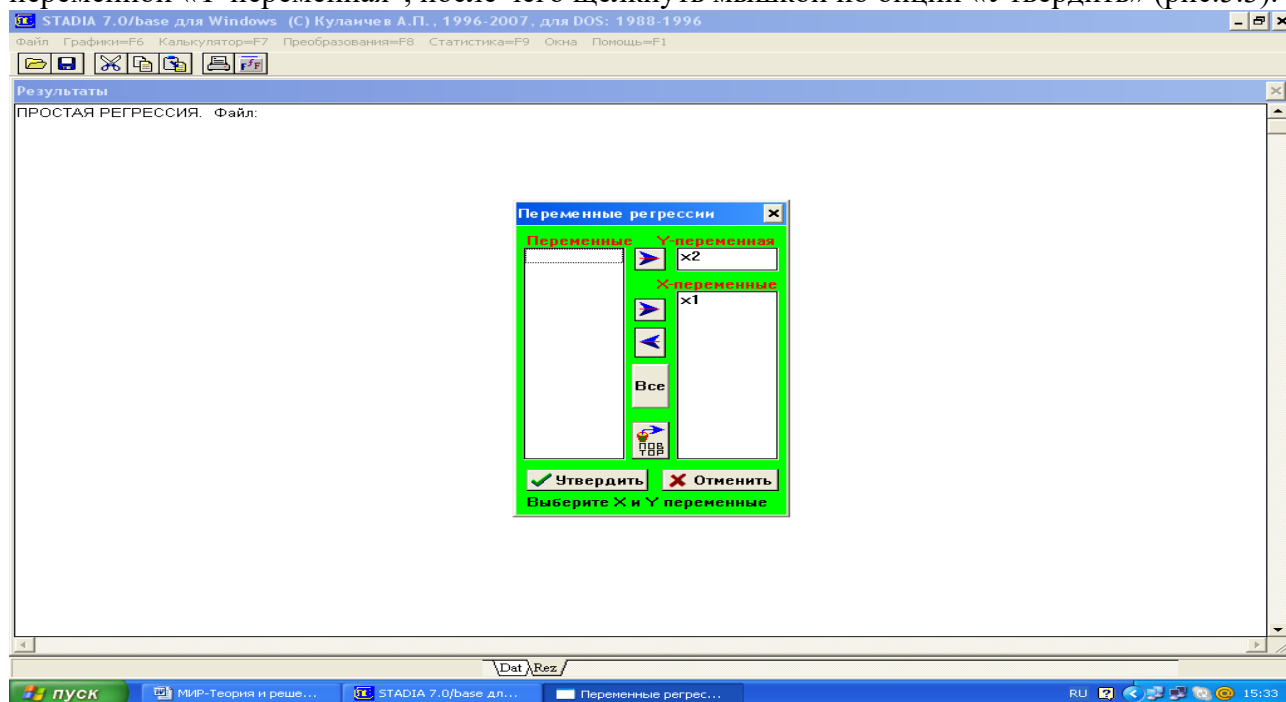


Рис. 5.3. Отображение подокна «Переменные регрессии»

В подокне «Регрессия» (рис. 5.4) надо кликнуть кнопку «1=линейная», выбирая тем самым вид нашей модели для проведения расчетов.

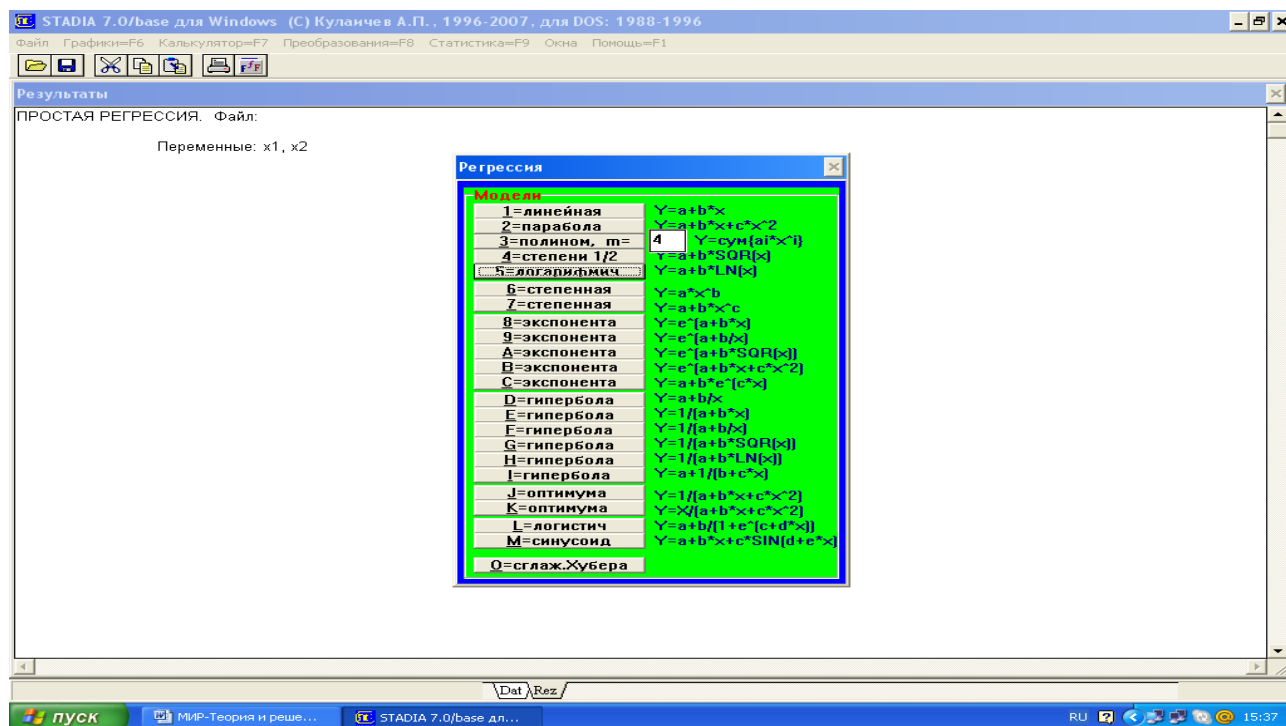


Рис.5.4. Отображение подокна «Регрессия»

Затем появится окно «Результаты» с подокном «Интерполяция» (рис.5.5).

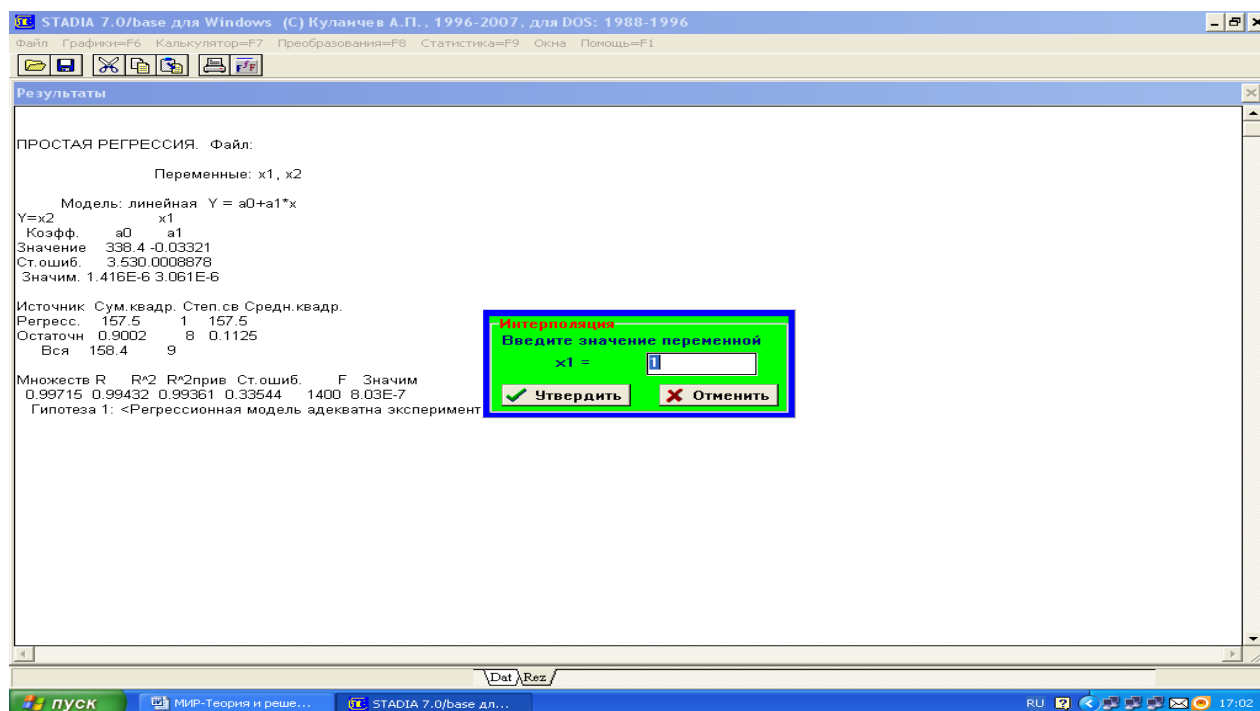


Рис.5.5. Отображение подокна «Интерполяция»

Кликнуть в этом подокне кнопку «Отменить», после чего появится окно «Регрессия» с отображением графика линейной регрессии и подокно «ПОСМОТРИТЕ ГРАФИК» (рис.5.6.).

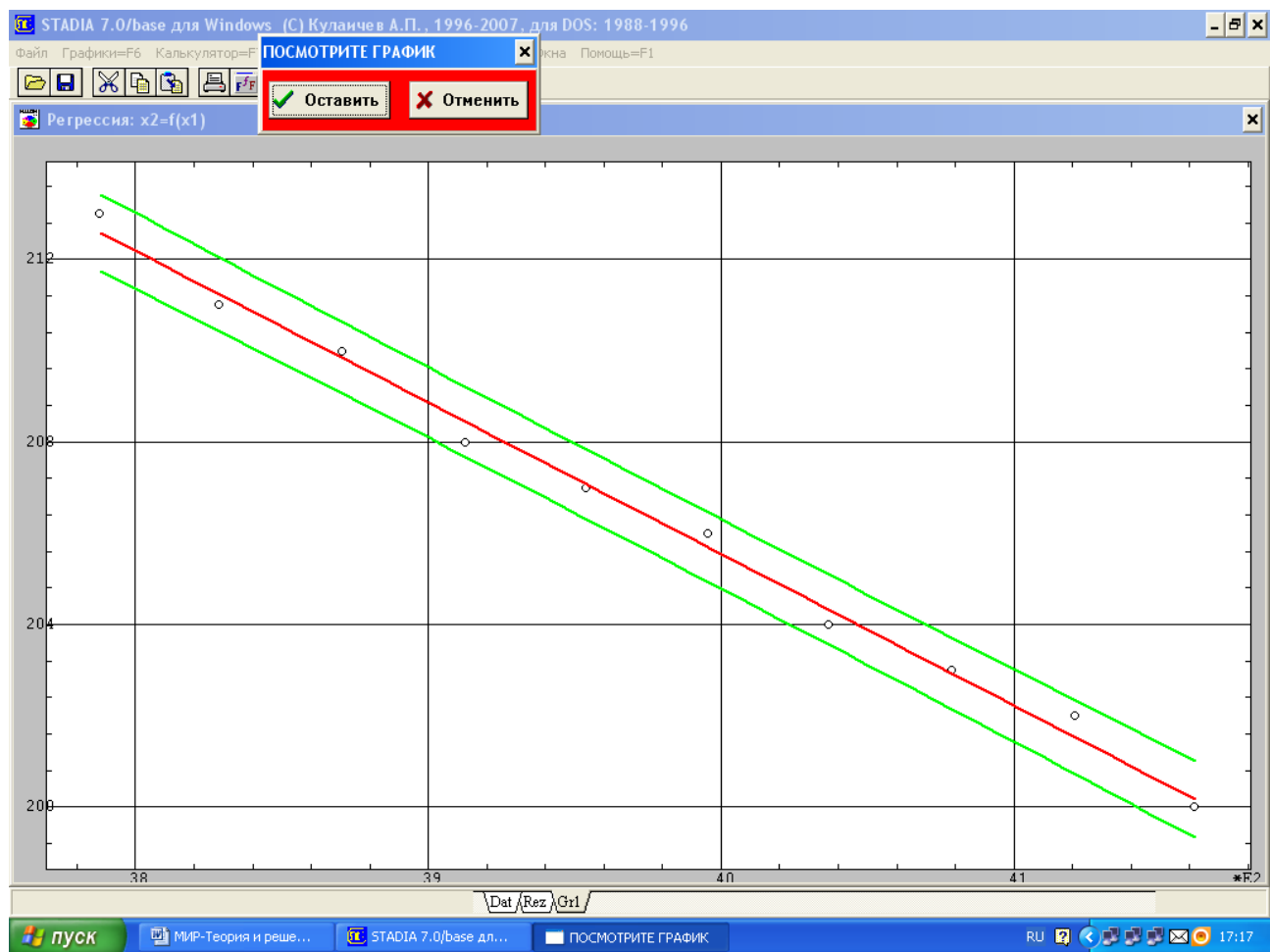


Рис.5.6. Отображение окна «Регрессия» с подокном «ПОСМОТРИТЕ ГРАФИК»

В этом подокне нужно кликнуть кнопку «Оставить», после чего появится отображение подокна «Что дальше?» с полем «Анализ» (рис. 15.7).

Нужно кликнуть кнопку «Анализ остатков» после чего на экране появится подокно «Confirm» (рис.5.8) с вопросом «Выдать график регрессионных остатков от независимой переменной?».

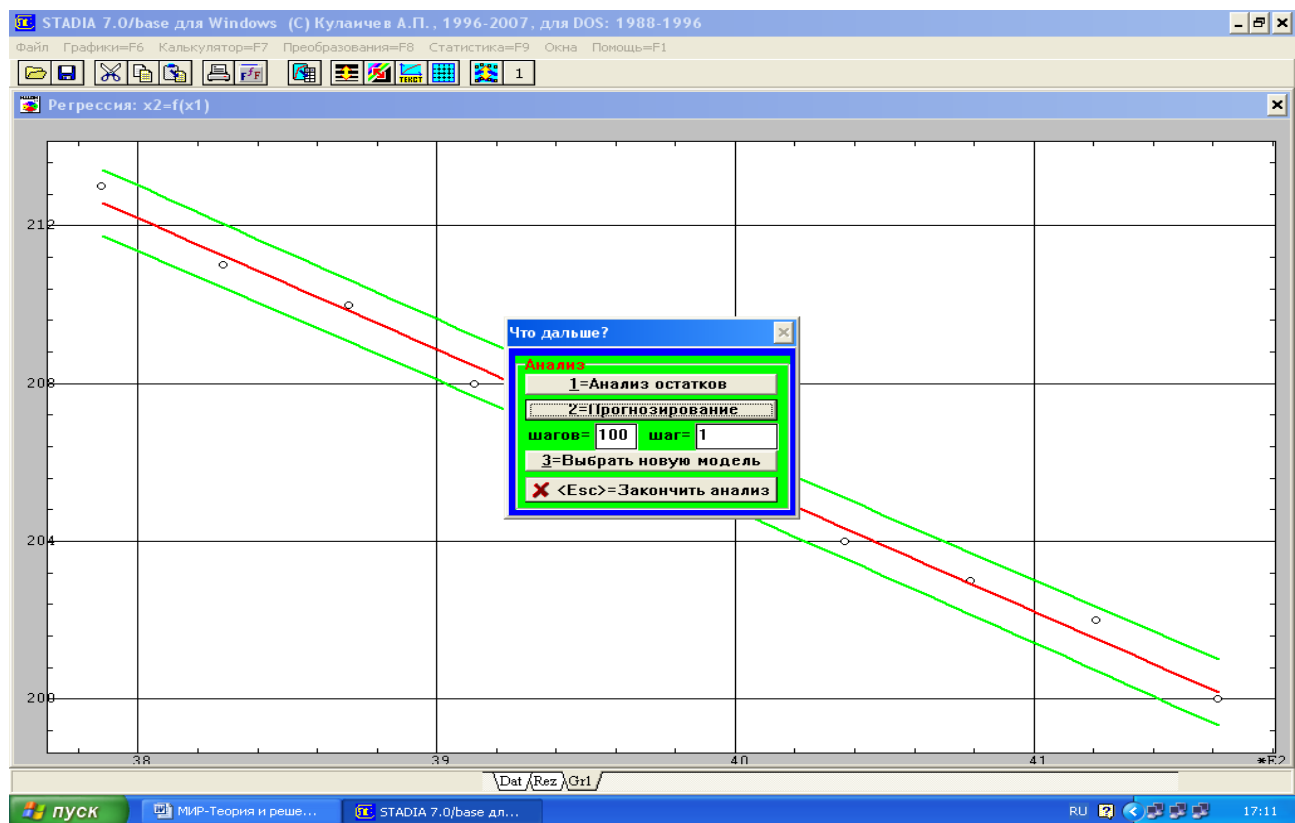


Рис.5.7. Отображение окна «Что дальше?» с субполем «Анализ»

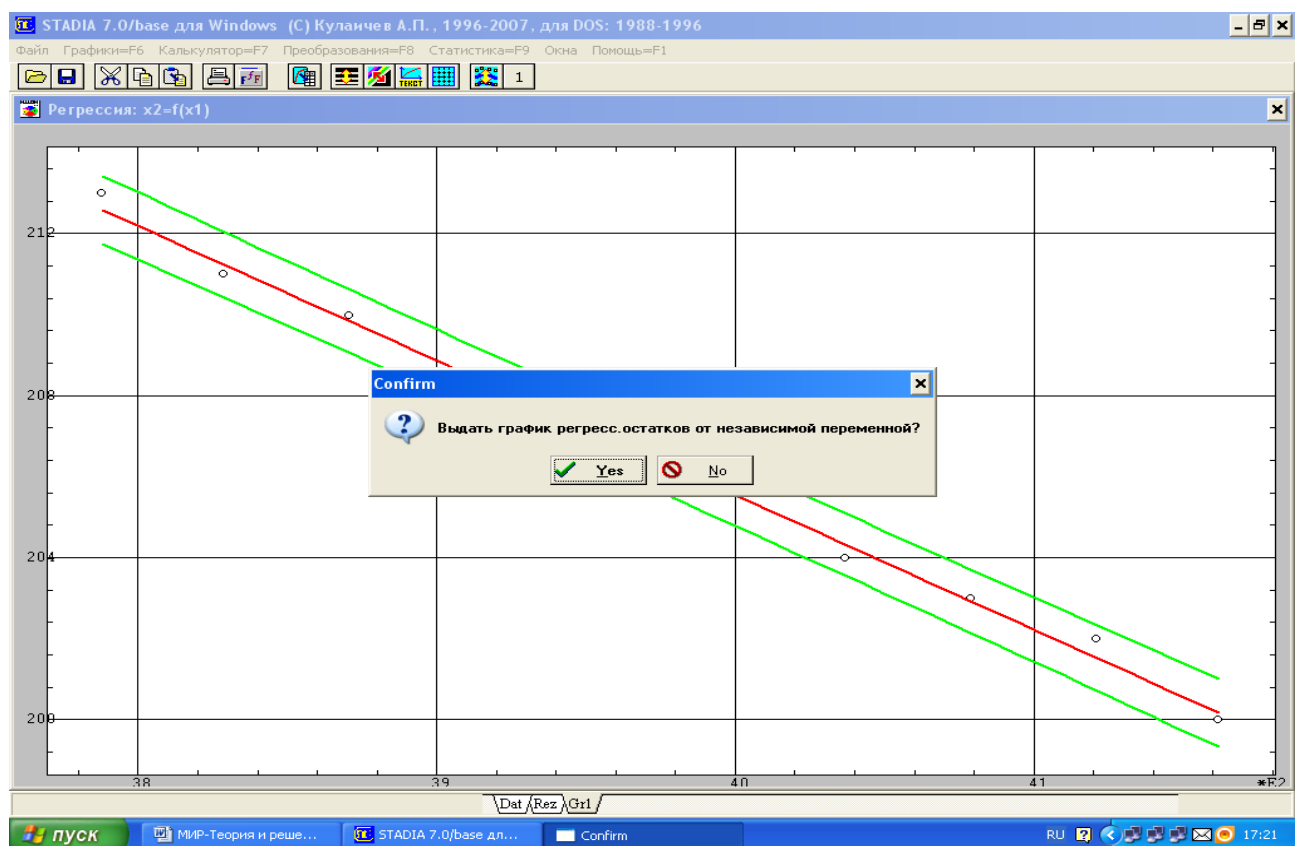


Рис.5.8. Отображение подокна «Confirm»

Нужно кликнуть кнопку «Yes» после чего на экране появится окно «Остатки X» с подокном «ПОСМОТРИТЕ ГРАФИК» (рис.5.9)

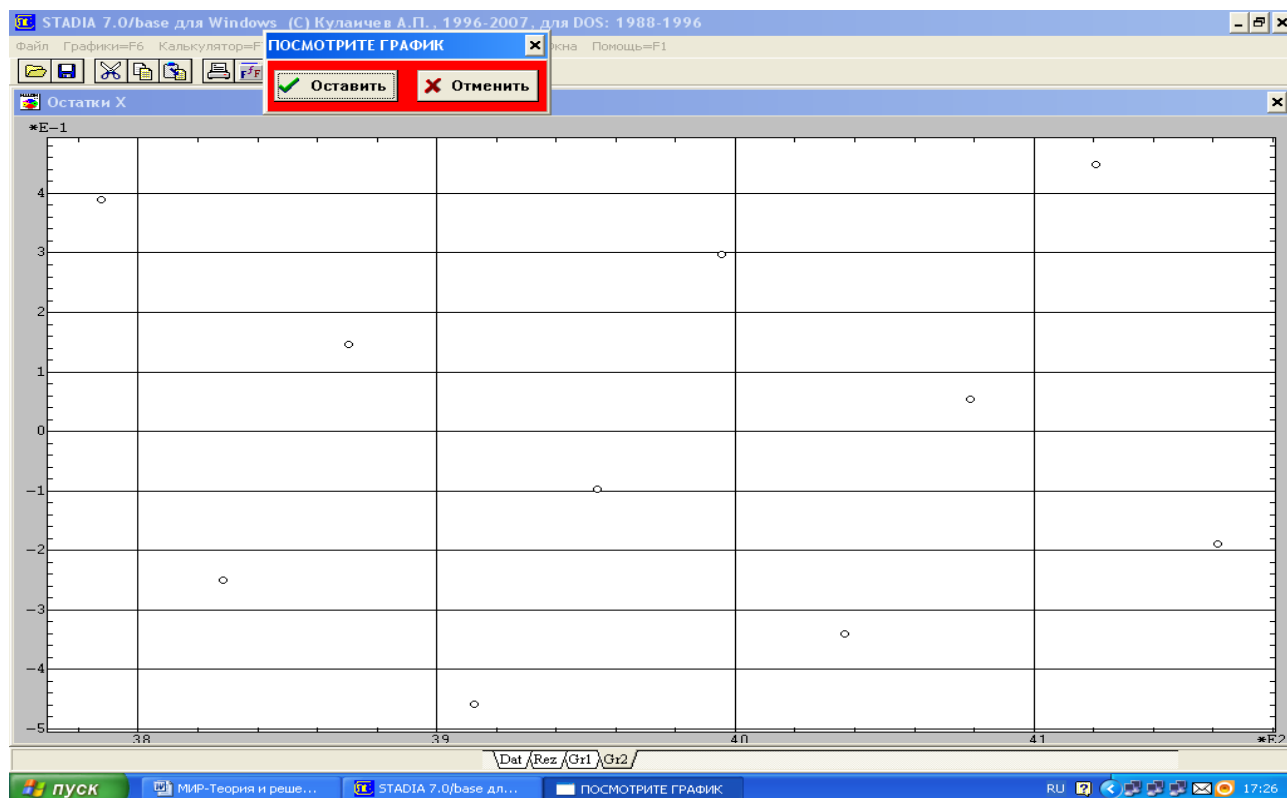


Рис.5.9. Отображение окна «Остатки X» с подокном «ПОСМОТРИТЕ ГРАФИК»

Нужно кликнуть кнопку «Оставить» после чего появится подокно «Confirm» с вопросом «Выдать график регрессионных остатков от регрессионной функции?» (рис.5.10).

Нужно кликнуть кнопку «Yes» после чего появится окно (рис.5.11) с подокном «ПОСМОТРИТЕ ГРАФИК»

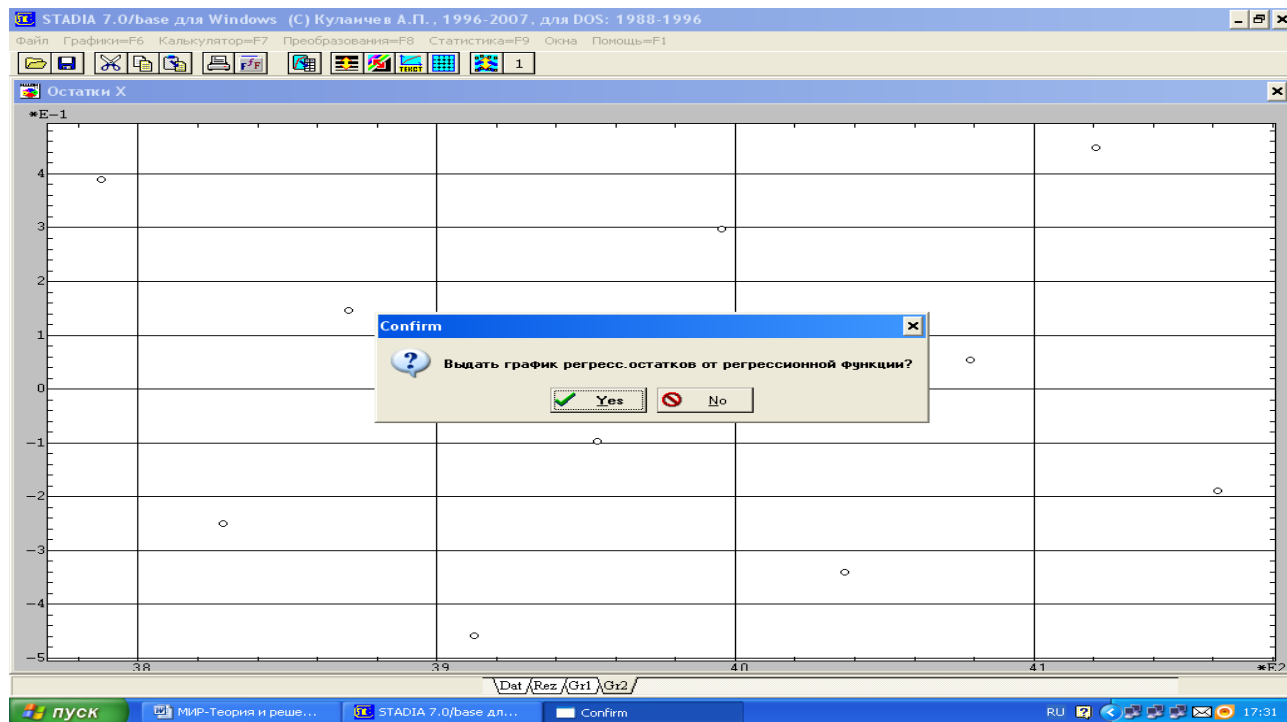


Рис.5.10. Отображение окна «Остатки X» с подокном «Confirm»

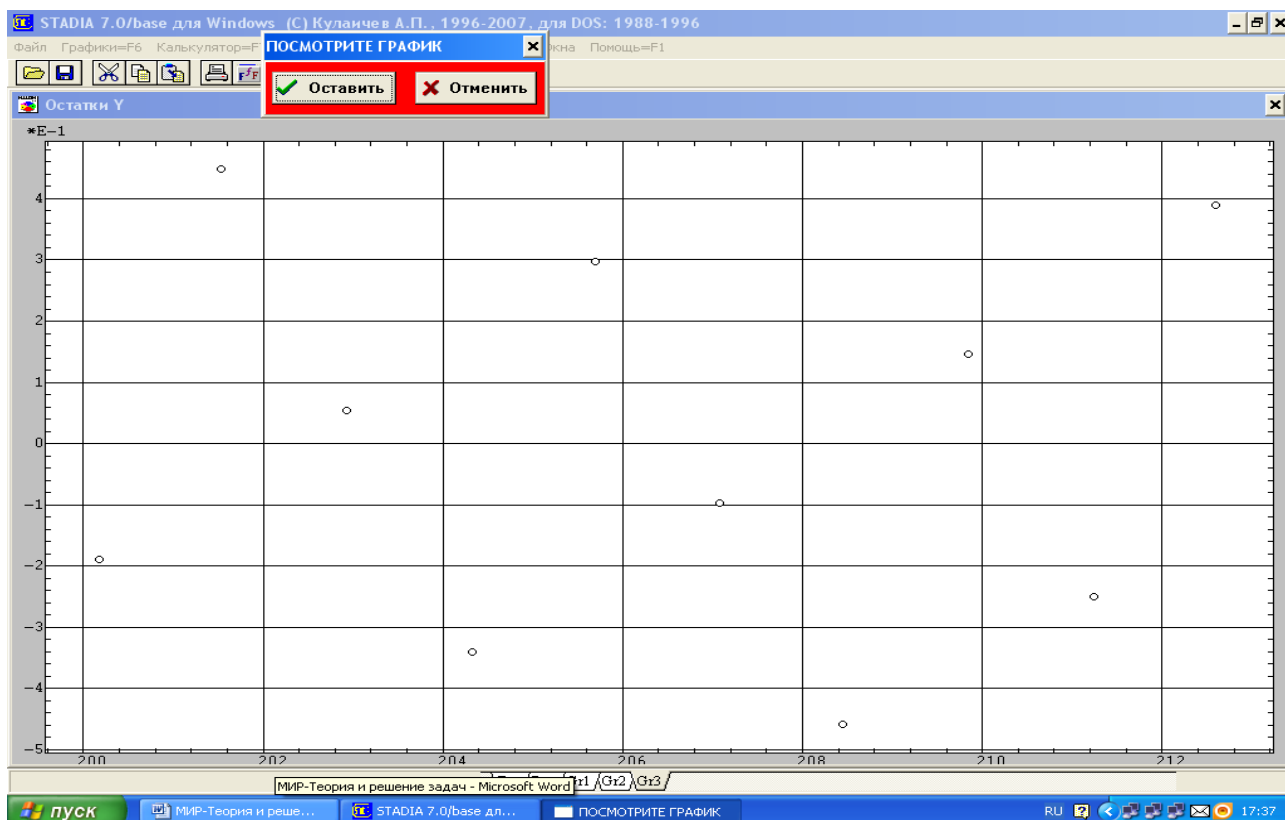


Рис.5.11. Отображение окна «Остатки Y» с подокном «ПОСМОТРИТЕ ГРАФИК»

Нужно кликнуть кнопку «Оставить» после чего опять появится окно с подокном «Что дальше?» (рис.5.12)

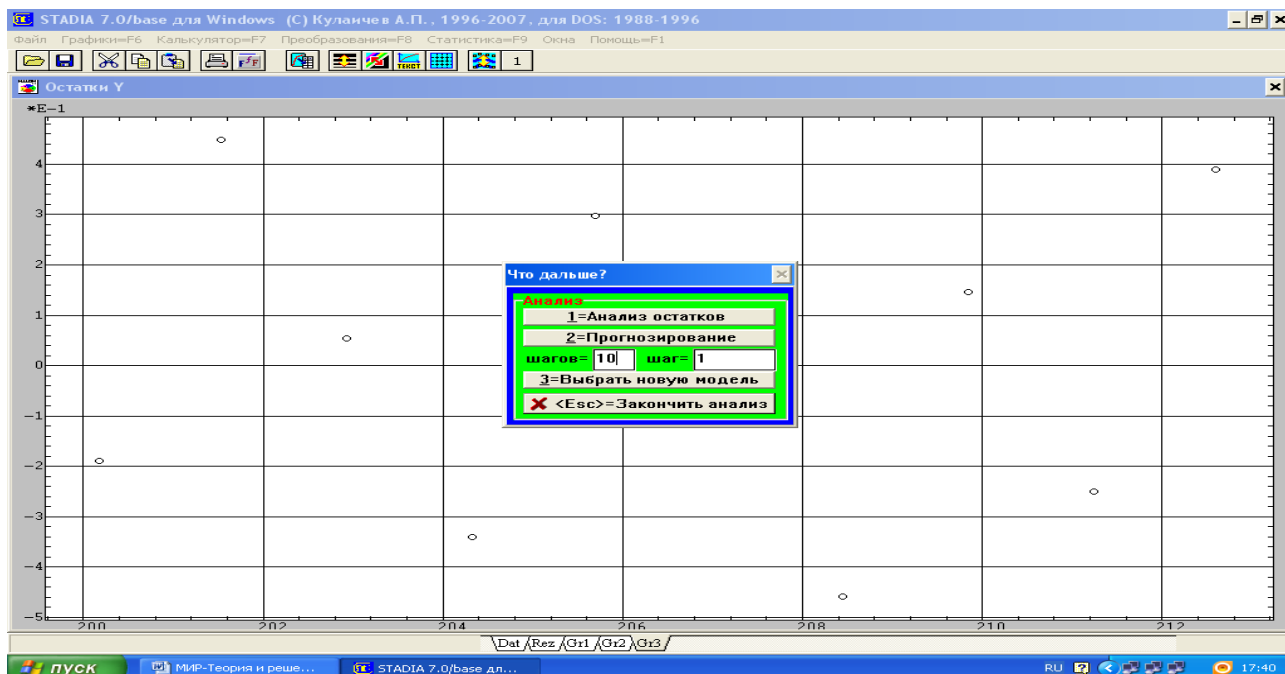


Рис.5.12. Отображение окна «Остатки Y» с подокном «Что дальше?»

Кликнуть кнопку «Прогнозирование» после чего на экране появится (рис.5.13)

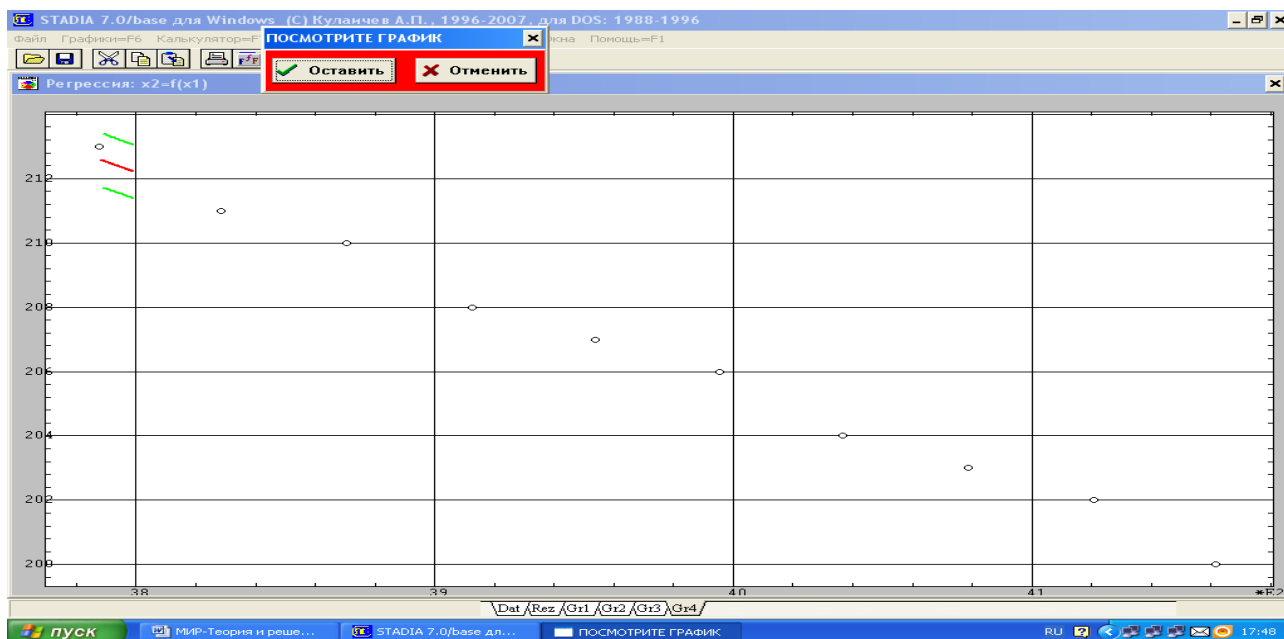


Рис.5.13. Отображение окна «Регрессия» с подокном «ПОСМОТРИТЕ ГРАФИК»

Нужно кликнуть кнопку «Оставить» после чего появится окно «Регрессия» с подокном «Что дальше?» (рис.5.14)

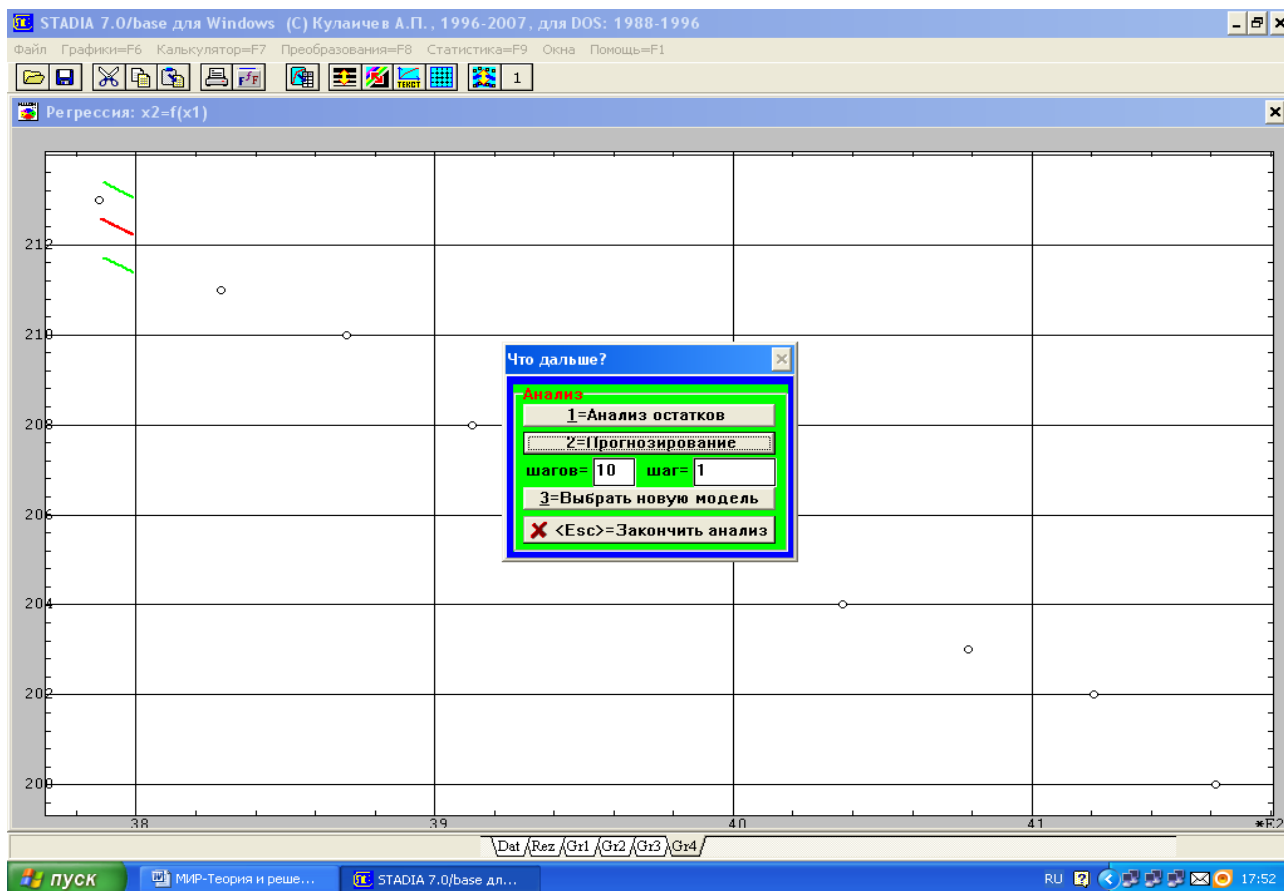


Рис.5.14. Отображение окна «Регрессия» с подокном «Что дальше?»

Нужно кликнуть кнопку «Esc=Закончить анализ». На этой процедуре происходит окончание обработки данных и получение результатов по построению регрессионной модели. Теперь можно переходить к следующему этапу.

Анализ полученных результатов. Прокомментируем графические результаты обработки. На графике линейной регрессии (рис.5.6) по оси ординат обозначена производительность ИТ – документов/день. По оси абсцисс показаны значения дефектов по несвоевременной передаче данных по этапам технологии обработки данных. На графике имеется три линии. Средняя линия обозначает линию регрессии. Нижняя и верхняя линии обозначают доверительный интервал, уровень значимости данных по 10 событиям (шагам) регрессии. На графике видно, что события довольно близко укладываются по линии регрессии. Это свидетельствует о том, что качество прогноза достаточно хорошее. Более конкретное заключение можно сделать на основе анализа остатков, как по переменной x (рис.5.9), так и по переменной Y (рис.5.11). «Разброс» кружочков на графике показывает отсутствие криволинейного тренда. Это свидетельствует о том, что качество линейного прогноза едва ли можно или целесообразно улучшить. Направленность линии регрессии показывает четкую зависимость производительности технологического процесса обработки данных ИТ и значения дефектов по запаздыванию документов. С уменьшением времени запаздывания документов растет производительность ИТ.

Для лучшего комментирования результатов, указанных на рис. 5.5, отобразим их в более укрупненном виде:

ПРОСТАЯ РЕГРЕССИЯ. Файл: Переменные: x1, x2

Модель: линейная $Y = a_0 + a_1 \cdot x$

Коэфф.	a_0	a_1
Значение	338.4	-0.03321
Ст.ошиб.	3,530	0,0008878
Значим.	1.416E-6	3.061E-6

Источник	Сум.квадр.	Степ.св.	Средн.квадр.
Регресс.	157.5	1	157.5
Остаточн.	0,9002	8	0,1125
Вся	158.4	9	

Множеств. R	R^2	R^2 прив.	Ст.ошиб.	F	Значим.
0,99715	0,99432	0,99361	0,33544	1,4E3	8.03E-7

Гипотеза 1: <Регрессионная модель адекватна экспериментальным данным>

X эксп.	Y эксп.	Y регр.	остаток	Ст.остат.	Ст.ошиб.	Довер.инт.
4162	200	200.2	-0.189	-0.5977	0.389	0.8887
4121	202	201.6	0.4492	1.42	0.3749	0.8565
4079	203	202.9	0.05423	0.1715	0.3637	0.8309
4037	204	204.3	-0.3407	-1.077	0.3561	0.8135
3996	206	205.7	0.2975	0.9406	0.3523	0.8048
3954	207	207.1	-0.09749	-0.3083	0.3523	0.8048
3913	208	208.5	-0.4593	-1.452	0.3561	0.8135
3871	210	209.9	0.1458	0.4609	0.3637	0.8309
3829	211	211.2	-0.2492	-0.788	0.3749	0.8565
3788	213	212.6	0.389	1.23	0.389	0.8887

X прогн.	Y прогн.	Ст.ошиб.	Довер.инт.
3789	212.6	0.3886	0.8878
3790	212.5	0.3883	0.887
3791	212.5	0.3879	0.8861
3792	212.5	0.3875	0.8853
3793	212.4	0.3871	0.8844
3794	212.4	0.3868	0.8836
3795	212.4	0.3864	0.8827
3796	212.3	0.386	0.8819
3797	212.3	0.3857	0.8811
3798	212.3	0.3853	0.8802
3799	212.2	0.385	0.8794

Прежде всего, отметим, что полученные статистические оценки в целом подтверждают гипотезу об адекватности модели экспериментальным данным. Получены значения коэффициентов регрессии $a_0 = 338$, $a_1 = -0,03321$, на основе которых можно построить уравнение регрессии для выполнения прогнозов ИТ по производительности от параметров запаздывания документов, то есть от их несвоевременного представления в систему обработки документов. В данном случае регрессионное уравнение будет выглядеть следующим образом: $Y = 338 - 0,03321x$.

Для предсказания производительности ИТ при условии снижения запаздывания документов, например, до 3330 мин. прогнозируемая производительность ИТ в этом случае будет равна $Y = 338 - (0,0332 * 3330) = 218$ документов/день. В уравнение можно подставлять любые значения по запаздыванию документов, которые необходимы для уточнения прогноза. При условии устранения запаздывания документов (значение $x = 0$) ИТ достигнет высшей (идеальной) производительности при данных условиях, то есть 338 документов/день.

С целью более оперативного получения прогнозных данных можно воспользоваться графиком. Если по оси абсцисс установить значение запаздывания и провести от этой точки перпендикулярную прямую до пересечения с функцией регрессии, то можно определить прогнозное значение производительности, проведя от точки пересечения прямую до оси ординат. В точке пересечения на оси ординат будет показано прогнозируемое (приблизительное) значение производительности. На рисунке 5.13 отображена точка начала прогнозного участка линии регрессии, которую нужно продолжить в направлении влево-вверх до значения 338 на оси ординат. Для этого необходимо построить полный график регрессии, состоящий из четырех квадрантов. Тогда на графике можно будет быстро находить прогнозные точки производительности ИТ, что является немаловажным в моделировании стратегии развития ИТ, а также решения задач контроля качества функционирования ИТ.

Анализ количественного отображения полученных значений остатков также подтверждает незначительные отклонения значений экспериментальной производительности ($Y_{\text{эксп.}}$) от значений регрессии, полученных на ЭВМ ($Y_{\text{регр.}}$). При условии завершения всех этапов решения задачи необходимо проверить еще раз правильность выполненных операций, после чего следует приступить к формулированию выводов.

Формулирование выводов. Относительно данной задачи выводы должны содержать оценку значения прогнозирования в моделировании ИР, методы и средства решения задачи. Особое внимание следует уделить последовательности и содержанию выполняемых этапов решения задачи, составу и форме представления результатов, а также методике анализа результатов решения задачи, составлению и оформлению отчета о работе. После завершения выводов следует приступить к составлению и оформлению отчета.

Составление и оформление отчета о решении задачи. Отчет составляется в соответствии с установленными требованиями, а также с учетом этапов и содержания данной задачи. После завершения составления и оформления отчета он представляется преподавателю для контроля и получения зачета по данной работе.

Контрольные вопросы и задания

1. Какую роль играют регрессионные модели в моделировании ИР?
2. Какие взаимосвязи можно исследовать посредством регрессионной модели?
3. Какова последовательность этапов построения регрессионной модели производительности ТПОД?
4. Назовите основные элементы уравнения линейной модели регрессионного анализа?
5. Какие основные параметры применяются для проверки адекватности регрессионной модели?
6. В какой последовательности выполняется проверка адекватности регрессионной модели?

Список литературы Основная литература:

1. Ананьева Т.Н., Новикова Т.Г., Исаев Г.Н. Стандартизация, сертификация и управление качеством программного обеспечения: учебное пособие.-М.: Издательство ИНФРА-М, 2016.-232 с. [ЭБС «znanium.com»]. - Режим доступа: <http://znanium.com/catalog.php?item>
2. Дубовой Н.Д. Основы метрологии, стандартизации и сертификации: Учебное пособие / Н.Д. Дубовой, Е.М. Портнов. - М.: ИД ФОРУМ: НИЦ Инфра-М, 2013.- [ЭБС «znanium.com»]. - Режим доступа: <http://znanium.com/catalog.php?item> - С.165-184. Материалы лекции 4,5.
3. Исаев Г.Н. Моделирование информационных ресурсов: Теория и решение задач: Учебник (гриф УМО).- М.:Альфа-М:ИНФРА-М.-2010.-288 с.
4. Кулаичев А.П. Методы и средства комплексного анализа данных.4-е изд., перераб. и доп.- М.: ФОРУМ:ИНФРА-М, 2006.-512 с.

Дополнительная литература:

5. Исаев Г.Н. Управление качеством информационных систем: Теоретико-методологические основания: монография.-М.: Наука, 2011.-289 с.
6. Исаев Г.Н. Проектирование информационных систем.-М.: Издательство Омега-Л, 2015.-424 с.
7. Исаев Г.Н. Информационные технологии. -М.: Издательство Омега-Л, 2014.-464 с.

Публикуемые отечественные и зарубежные источники по дисциплине в периодических изданиях

8. Научно-техническая информация. Серия 1. Информационные процессы и системы.
9. Научно-техническая информация. Серия 2. Организация и методика информационной работы.
10. Журнал «Информационные технологии».
11. Реферативный журнал «Информатика».

3.2. Перечень программ, тестов, пособий в электронном виде

1. Программный комплекс «STADIA»

