

**ГОУ ВПО
«ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

КОНСПЕКТ ЛЕКЦИЙ
учебной дисциплины вариативной части
профессиональной и практической подготовки
дисциплин самостоятельного выбора ВУЗа
ГОС ВПО по направлению подготовки бакалавра
27.03.02 «Управление качеством»

«Программные статистические комплексы»

Донецк, 2016 г.

**ГОУ ВПО
«ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

КОНСПЕКТ ЛЕКЦИЙ
учебной дисциплины вариативной части
профессиональной и практической подготовки
дисциплин самостоятельного выбора ВУЗа
ГОС ВПО по направлению подготовки бакалавра
27/03.02 «»»Управление качеством

«Программные статистические комплексы»

Рассмотрено на
заседании кафедры
«Управление качеством»
Протокол № от « » 2016

Утверждено на заседании
учебно – издательского
Совета ДонНТУ
Протокол № от « » 2016

Донецк, 2016 г.

УДК 65.012

Конспект лекций по учебной дисциплине «Программные статистические комплексы» (для студентов направления подготовки 27.03.02 «Управление качеством» профиля подготовки «Управление качеством, стандартизация, метрология и сертификация» / Сост.: Н.А.Ченцов, - Донецк: ДонНТУ, 2016. - 23 с.

Описаны методы и подходы использования специализированных программных средств для обработки статистических данных при решении задач управления качеством.

Составители:

НА. Ченцов, проф.

Отв. за выпуск:

Н.Ф. Годына, доц.

Рецензенты:

Е.В. Мирошниченко, доц.

СОДЕРЖАНИЕ

1	ЗНАКОМСТВО С ПАКЕТОМ STATISTICA	5
1.1...	ПРЕДМЕТ СТАТИСТИКИ	5
1.2...	ПРОГРАММНЫЕ ПРОДУКТЫ	5
1.3 --	ИНСТАЛЛЯЦИЯ ПАКЕТА STATISTICA	6
1.4...	ЗАГРУЗКА ПАКЕТА	6
2	ОСНОВНЫЕ ПРАВИЛА РАБОТЫ В ПАКЕТЕ	7
2.1...	МОДУЛИ ПАКЕТА	7
2.2...	ОПРЕДЕЛЕНИЕ СРЕДНЕГО ЗНАЧЕНИЯ ПЕРЕМЕННОЙ	7
2.3...	ХАРАКТЕРИСТИКА ВЫХОДНЫХ ФОРМ	8
2.4...	СЛОЖНОЕ СОХРАНЕНИЕ РЕЗУЛЬТАТОВ АНАЛИЗА	9
3	ОСНОВЫ АНАЛИЗА ОПИСАТЕЛЬНОЙ СТАТИСТИКИ	9
3.1...	ФОРМИРОВАНИЕ ТАБЛИЦЫ ДАННЫХ	9
3.2...	ГРУППИРОВКА ПАРАМЕТРОВ СТАТИСТИКИ	10
3.3...	РУЧНОЕ ВЫЧИСЛЕНИЕ ПАРАМЕТРОВ СТАТИСТИКИ	10
3.4...	АВТОМАТИЧЕСКОЕ ВЫЧИСЛЕНИЕ ПАРАМЕТРОВ СТАТИСТИКИ	11
4	ВИЗУАЛИЗАЦИЯ СТАТИСТИК	12
4.1...	ВИЗУАЛИЗАЦИЯ ОПИСАТЕЛЬНЫХ СТАТИСТИК	12
4.2...	ПОСТРОЕНИЕ ГИСТОГРАММ	12
4.3...	ПОСТРОЕНИЕ КРУГОВЫХ ДИАГРАММ	13
4.4...	ПОСТРОЕНИЕ ДИАГРАММЫ РАССЕИВАНИЯ	13
5	ПРОВЕРКА ГИПОТЕЗЫ	14
5.1...	ПРОВЕРКА ГИПОТЕЗЫ О ВИДЕ РАСПРЕДЕЛЕНИЯ	14
5.2...	ТАБЛИЦА РЕЗУЛЬТАТОВ КРИТЕРИЯ ПИРСОНА	14
5.3...	ПРОВЕРКА ОДНОРОДНОСТИ ВЫБОРОК	15
5.4...	ВЫЯВЛЕНИЕ ПРИЧИН НИЗКОГО УРОВНЯ ЗНАЧИМОСТИ	15
6	СВЯЗИ МЕЖДУ ПЕРЕМЕННЫМИ	16
6.1...	КОРРЕЛЯЦИОННЫЙ АНАЛИЗ	16
6.2...	ПРОСТОЙ (ОДНОФАКТОРНЫЙ) РЕГРЕССИОННЫЙ АНАЛИЗ	17
6.3...	ОЦЕНКА АДЕКВАТНОСТИ МОДЕЛИ ПО ОСТАТКАМ	17
6.4...	СЛОЖНЫЙ РЕГРЕССИОННЫЙ АНАЛИЗ	18
7	ДИАГРАММЫ КАЧЕСТВА	18
7.1...	АНАЛИЗ ПРИЧИННО – СЛЕДСТВЕННЫХ СВЯЗЕЙ	18
7.2...	ПОСТРОЕНИЕ ДИАГРАММЫ ИСИКАВЫ	19
7.3...	ДИАГРАММА ПАРЕТО	19
7.4...	ПОСТРОЕНИЕ ДИАГРАММЫ ПАРЕТО В STATISTICA	20
8	КАРТЫ КОНТРОЛЯ КАЧЕСТВА	20
8.1...	СОДЕРЖАНИЕ КОНТРОЛЬНОЙ КАРТЫ	20
8.2...	КОНТРОЛЬНАЯ КАРТА ИНДИВИДУАЛЬНЫХ ЗНАЧЕНИЙ	21
8.3...	ИНТЕГРАЛЬНЫЕ КОНТРОЛЬНЫЕ КАРТЫ	21
8.4...	ЧТЕНИЕ КОНТРОЛЬНЫХ КАРТ	22

1 Знакомство с пакетом STATISTICA

1.1 Предмет статистики

Статистика — это наука, изучающая количественную характеристику качественной стороны подобных массовых явлений и процессов. Например, успеваемость студентов.

Цель статистики - обнаружение закономерностей, проявляющихся в массовых явлениях или процессах.

Статистическая совокупность — это множество качественно однородных экземпляров (реализаций) объекта. Например, партия деталей, изготовленных по одному чертежу.

Значение числовой характеристики каждой реализации объекта является уникальной и используется для статистического анализа. Например, диаметр втулки в партии деталей.

Статистическая закономерность - форма проявления повторяемости событий, характеристик объекта при неизменности причины, порождающие события. (Сохранение *status quo*.)

Статистические методы определяется объемом области, на которую распространяются выводы, полученные на основе метода.

Описательная статистика используется для простого обобщения данных экспериментальной выборки. Ее выводы распространяются только на объем выборки.

Статистические выводы обеспечивают распространение выводов, полученных из экспериментальной выборки объектов на генеральную их совокупность,

Статистическое исследование включает следующие составные части:

- статистическое наблюдение и фиксация результатов,
- сводка и группировка исходных данных,
- назначение обобщающих показателей характеристики
- их определение из статистического анализа исходных данных.

1.2 Программные продукты

Программные продукты, используемые в рамках статистического исследования можно разделить на три группы, различающиеся по степени специализации.

Универсальное приложение – Microsoft Office. Офисный пакет приложений, созданных корпорацией Microsoft для операционной системы Microsoft Windows. В состав этого пакета входит программное обеспечение для работы с: текстами, электронными таблицами (Excel), базами данных и др.

Microsoft Excel программа для работы с электронными таблицами, которая предоставляет возможности экономико-статистических расчетов, и построения графиков.

Математические пакеты общего назначения имеющие в своем составе средства для статистической обработки данных.

Mathcad -система компьютерной алгебры, ориентированная на подготовку документов с вычислениями, представленными в виде формул и визуальным сопровождением. Реализует подход близкий к работе с бумажным документом.

MatLab пакет прикладных программ для решения задач технических вычислений и одноименный язык программирования, используемый в этом пакете.

Специализированные пакеты разработанные и применяемые для реализации только статистических методов.

STATGRAPHICS, пакет может запускаться в веб-браузере и позволяет импортировать данные из файлов Excel или других форматов. Дополнительно пакет имеет надстройку для реализации метода *Six Sigma*.

STATISTICA, пакет для реализующий функции анализа данных, визуализации данных с привлечением статистических методов. Пакет позволяет выводить информацию в виде различных типов графиков (научные, деловые, трёхмерные и двухмерные графики, специализированные статистические графики — гистограммы, матричные, категорированные графики и др.).

1.3 Инсталляция пакета STATISTICA

Генезис пакета STATISTICA – первая версия была создана в 1986 году, в настоящее время имеется 12я версия. Для изучения будем использовать версию **STATISTICA 6.0**.

Подготовка файловой системы включает операции:

- получить у преподавателя папку *qm16* и переписать ее на ПК студента;
- на своем ПК переименовать папку *StatFIO* заменив *FIO* – на инициалы студента. Например, у Ченцова получим *StatChen*.
- распаковать файл *Statistica_6.0.zip*. *В результате получим инсталляционную папку *Statistica_6.0*.

Обеспечить совместимость STATISTICA 6.0 с Windows 7 и старше:

- в папке *Statistica_6.0* выбрать файл *CDSTART* – установить на него курсор;
- открыть свойства файла – *rClick* / *свойства*;
- перейти на вкладку *Совместимость*, рис.1

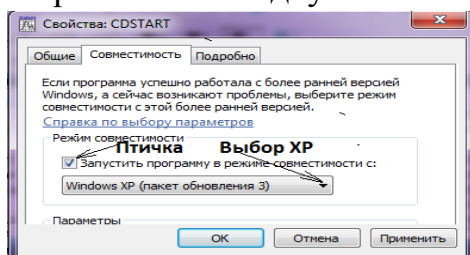


Рис. 1 Совместимость XP

- установить «птичку» возле «Запустить программу в режиме совместимости с:
- Под птичкой выбрать «Windows XP» (пакет обновления 3).
- Нажать «ОК». *Будет установлена совместимость.

Инсталляция пакета из папки *Statistica_6.0*:

- запустить файл *CDSTART* – сделать *wClick* на нем.

Откроется форма инсталляции;

- запустить инсталляцию – нажать *Install STATISTICA*. *Начнется инсталляция.
- Подтверждать все требования до выхода из инсталляции. *Логин и ключи в файле *Content*.



- После окончания инсталляции на столе должна появиться пиктограмма «STATISTICA», а в системе файл "*C:\Program Files (x86)\StatSoft\STATISTICA 6\statist.exe*"

Установка русского языка:

- запустить файл *Ru/stl_asup.exe*. *Будет подключена кириллица;
- запустить файл *RUSSIAN*. Будет введен русский текст.

Запуск пакета – нажать пиктограмму на столе. *Загрузится стартовая форма пакета с пустой таблицей.

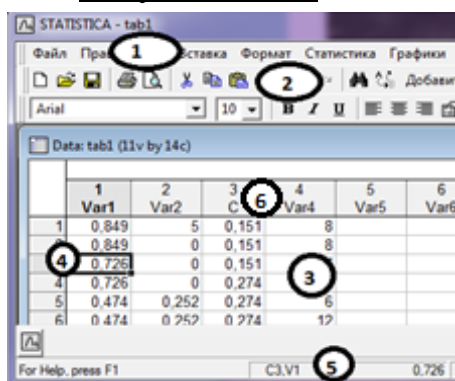


Рис. 1 Стартовая форма

1.4 Загрузка пакета

Стартовая форма пакета включает следующие компоненты (рис. 1):

- 1) меню пакета;
- 2) кнопки управления;
- 3) таблица данных в рабочей зоне;
- 4) активная клетка;
- 5) характеристика активной клетки;

б) имена переменных.

При запуске пакета на стартовую форму загружаются данные последней рабочей таблицы. Ее имя приведено в верхней части таблицы.

Таблица включает строки (регистры), номера которых приведены в левой колонке, и переменные (колонки), номера и имена которых приведены в «шапке» таблицы. *Имена переменных можно корректировать.

Переменные могут иметь типы: *Text*(10 - кол-во символов); *Integer* – целое число; *Double* – десятичное число; *Date* – дата.

Открыть существующую таблицу.

- 1) Выбрать таблицу – *Файл / открыть*. *В файловой системе ПК найти необходимую таблицу *qm16 / StatFIO / ItFriendFIO.sta*.
- 2) Открыть таблицу – сделать на ней *wClick*. *В рабочем окне откроется загрузится выбранная таблица.

2 Основные правила работы в пакете

2.1 Модули пакета

Перечень модулей пакета открывается кнопкой «Статистика» в главном меню. Он выводится на панели, рис. 1. и включает:

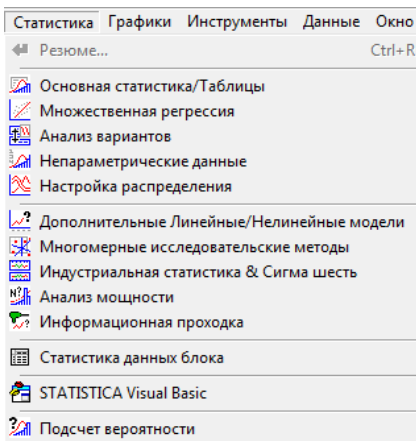


Рис. 1 Модули пакета

- основная статистика/таблицы *Basic statistic* – обеспечивает получение основных показателей описательной статистики;
- множественная регрессия – показывает связь аргумента и отклика;
- непараметрические данные – модуль используется, если данные не соответствуют параметрическим критериям и они не могут быть описаны нормальным законом;
- прочие.

Модуль «Основная статистика/таблицы» (рис.2)

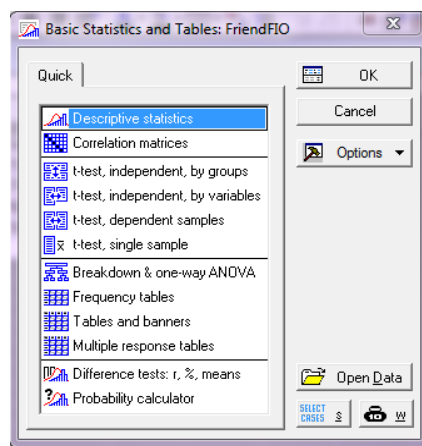


Рис. 2 Задачи «Основная статистика..»

обеспечивает, на основе данных таблицы, решение задач:

- *Descriptive statistic* (описательная статистика);
- *Correlation matrices* (корреляционная матрица);
- *t-test for independent samples* (t-тест для независимых переменных);
- *t-test for dependent samples* (t-тест для зависимых переменных);
- *ANOVA* (дисперсионный анализ);

- *Correlation matrices* (описательная статистика);
- *t-Test for independent variables* (корреляционная матрица);
- *t-Test for dependent samples* (t-тест для независимых переменных);
- *Breakdown & one way analysis* (t-тест для зависимых переменных);
- Прочие.

2.2 Определение среднего значения переменной

Параметры описательной статистики в STATISTICA определяется для выбранной переменной (колонки) загруженной таблицы в последовательности:

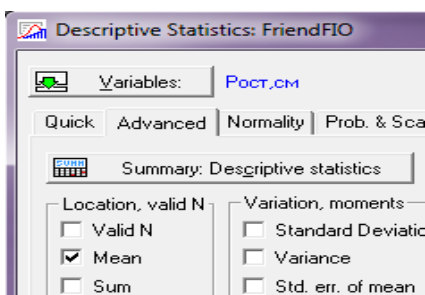


Рис. 3 Настройка «Основная статистика»

1) В меню - *Статистика (Click) / Basic statistic (Click) / Descriptive statistic (wClick)* – откроется панель *Descriptive statistic*, рис.3.

2) Вверху панели нажать *Variables* и выбрать *4-Пост,см*.

3) выбрать вкладку *Advanced*, на которой видны задания на вывод параметров :

- *Valid N* – количества значений переменной;
- *Mean* – среднего значения;
- *Sum* – сумма значений;
- *Standart deviation* стандартное отклонение.
- Др.

Определение среднего роста $n=5$ друзей (параметр описательной статистики).

Вручную

$$H_{CP} = \frac{1}{n} \sum_{i=1}^n H_i = (156 + 212 + 174 + 162 + 156) / 5 = 172 \text{ см.}$$

Автоматически - используя пакет STATISTICA для таблицы с данными и указанной переменной *4-Рост, см.*:

Variable	Mean
Рост, см	172.0000

Рис. 4 Результат обработки статистики – среднее значение переменной *Рост, см*

- ставим птичку возле *Mean* (вид выбранного показателя);
 - нажимаем кнопку *Summary: Descriptive statistic.*
- *Откроется *Date:..* рис. 4..

На ней видны:

- *Mean* - вид показателя;
- *Рост, см* - имя переменной;

172 – среднее значение переменной.

2.3 Характеристика выходных форм

Результаты анализа данных в пакете *STATISTICA* выводятся в три вида форм, которые можно сохранять как отдельные файлы

- 1) **Индивидуальное окно** - имеет вид таблицы, рисунка и др. Определяется содержанием таблицы с данными и видом выполненного анализа.
- 2) **Книга результатов** - представлена отдельными страницами, каждая из которых включает результат одного анализа (сохраненное индивидуальное окно). Например, исходные данные; таблицы с результатами анализа; диаграммы; и т.д. Группировка страниц представлена структурой, выполненной по правилам подобным используемым в файловой системе компьютера.
- 3) **Отчет с результатами анализа** – включает результаты проведенных анализов (индивидуальные окна), представленные в виде обычного текстового файла в формате *rtf*.

Настройка автоматического вывода данных выполняется на панели (*File / OutputManager / Менеджер вывода*), где указываются:

- **направление обязательного** вывода - Индивидуальные окна или книга;
- **направление дополнительного** вывода – Также послать в окно отчета (вывод в текстовый файл).

Вывод в индивидуальное окно является наиболее доступным в настройке. Для его организации на панели *Менеджер вывода* снять все отметки кроме (*Индивидуальные окна / ОК*).

Работа с окном – можно: сохранить и загрузить из файловой системы, вставить в книгу или отчет.

Сохранение окна –

- выбрать его (*Click*);
- запустить сохранение (*Файл / Сохранить как*);
- выбрать папку для сохранения *qm16 / 1aStatFIO*;
- Указать имя файла – в имени таблицы *t* заменить на *a*. Например, *ItFriendFIO* заменить на *1aFriendFIO*.

2.4 Сложное сохранение результатов анализа

Вывод в книгу (новую) включает настройку (*Книга / Одна книга ... / Поместить результаты в книгу автоматически / ОК*). Все прочие отметки снять.

Создание книги – выполняется автоматически после проведения первого анализа. В нее автоматически вводится результат анализа и ей присваивается имя *Workbook№*.

Переименование книги:

- запустить сохранение (*Файл / Сохранить как*);
- выбрать папку для сохранения *qm16 / StatFIO*;
- Указать имя файла – в имени таблицы *t* заменить на *w*. Например, *ItFriendFIO* заменить на *lwFriendFIO*.

Дописывание в книгу – выполняется автоматически после проведения очередного анализа.

Извлечение из книги – в структуре книги выбрать страницу, перетащить в рабочую зону и подтвердить копирование (*Сору*)/

Вывод в отчет

 (новый текстовый файл) включает настройку:

- вывода в индивидуальное окно – *Индивидуальные окна*;
- вывода в отчет - *Также послать в окно отчёта и Один отчет*.
- все прочие отметки снять.

Создание файла отчета – выполняется автоматически после проведения первого анализа. Ему присваивается имя *Report№* и в него вводится результат анализа.

Переименование отчета – выполнить по правилам, принятым для книги. Имя присваивать, используя *r -lrFriendFIO*.

Дописывание в отчет – выполняется автоматически после проведения очередного анализа.

Извлечение из отчета – в структуре выбрать страницу, перетащить в рабочую зону. *Из отчета будет удалена (вернуть - *ctrl+z*).

Прочие варианты вывода:

- автоматически в указанный ранее созданный отчет или книгу;
- индивидуальное окно в указанный ранее созданный отчет или книгу.

3 Основы анализа описательной статистики

3.1 Формирование таблицы данных

Проведение наблюдений и группировка исходных данных составляют 1й этап статистического исследования. Например, в случае исследования друзей результаты наблюдений можно представить в виде таблицы *lFriendFio*.

Таблица - *lFriendFio*

№	ФИО	ДР	Пол	Рост, см	Вес, кг	Балл.
1	Красилин Ю.А.	12.12.1902	М	156	64,66	94
2	Шмелев В.А.	04.05.1995	М	212	105,21	65
3	Фоменко Р.И.	31.01.1998	Ж	174	59,82	78
4	Корина У.Н.	25.08.2001	Ж	162	76,52	87
	Ганжа СС.	17.02.2002	М	156	64,02	94
	<i>Text(20)</i>	<i>Date</i>	<i>Text(1)</i>	<i>Integer</i>	<i>Double</i>	<i>Ineger</i>

Сводка и группировка исходных данных составляют 2й этап. В примере в рамках 2го этапа выявлено значительное отклонение дня рождения у Красилина. Это явилось основанием исключения его из таблицы при формировании сводки.

Новая таблица создается в следующей последовательности.

- 1) Закрыть открытую таблицу – *Файл / закрыть*.

- 2) Создать новую таблицу – *Файл / New*. *Будет предложен размер таблицы по 10 переменных и регистров (строк), которые можно корректировать. *OK*.
- 3) Корректировать имена переменных (колонок) и их типов. *Сделать *wClick* на имени переменной – откроется панель. На панели выполнить корректировку.
- 4) Заполнить таблицу – последовательно в каждую клетку таблицы ввести необходимые данные исключая Красилина.
- 5) Сохранить таблицу – *Файл / сохранить*. *Будет предложено создать имя файлу.

3.2 Группировка параметров статистики

Описательная статистика позволяет определить значительного количества различных показателей.

В **STATISTICA** зявка на их определение указывается на панели *Статистика / (Основная статистика/Таблицы) / Descriptive statistic (wClick)*, где на вкладке *Advanced* приведены три группы.

1.Location, valid N (меры центральной тенденции);

- *Valid N* – количества значений переменной;
- *Mean* – среднего значения;
- *Sum* – сумма значений;
- *Median* – медиана (серединное значение показателя);
- *Mode* – мода (показывает наиболее часто встречающееся значение).
- *Др.*

2.Variation, moments – меры вариативности характеризуют неоднородность выборки, определяемую разницей значений показателей в ней:

- *Standart deviation*- стандартное отклонение;
- *Variance* – дисперсия;
- *Std. err. of mean* – стандартное отклонение среднего;
- *Conf. limits fo means (Interval 95%)* – доверительный интервал в которых с заданной вероятностью (95%) лежит среднее значение величины
- *Др.*

3.Perceniles, ranges - проценты , диапазоны

- *Minimum & maximum* минимальное и максимальное значения величины;
- *Lower & upper quartiles* - Нижняя и верхняя квартили;
- *Persentile boundaries First 10% second 90%* - Границы в процентах Первая 10% Вторая 90%.
- *Др.*

3.3 Ручное вычисление параметров статистики

Для примера рассмотрим доход и образование каждого из 12 человек в группе : 2н, 2с, 3с, 3н, 3т, 4б, 4б, 5т, 6м, 10б, 13м, 19м. Где: н-1 начальное; с-2 среднее; т-3техникум; б-4 бакалавр; м-5 магистр.

Ручной анализ данных позволяет получить значения характеристик описательной статистики:

Центральные характеристики (Location):

- *Valid N = 12* - объем статистики;
- *Mean = (2 + 2 + 3 + 3 + 3 + 4 + 4 + 5 + 6 + 10 + 13 + 19)/12 = 6,17* – среднее арифметическое дохода;

- $Geom\ mean = (2 * 2 * 3 * 3 * 3 * 4 * 4 * 5 * 6 * 10 * 13 * 19) / 12 = 4,74$ – среднее геометрическое геометрическое;
- $Median = (2п + 2в + 3т + 3ч + 3п + 4 + 4 + 5п + 6т + 10т + 13в + 19п) = 4$ – медиана (срединное значение показателя);
- $Mode = (3 + 3 + 3) = 3$ – мода (наиболее часто встречающееся значение);

Характеристика вариации данных (*Variation, moments*):

- *Stdandard Devviation.* = 5,24 – стандартное отклонение (среднеквадратическое). Определяется через дисперсию.

Характеристика разброса данных (*Percentiles, renges*):

- *Lower Quartile* = (2, 2, 3, 3,) = 3 – нижняя квартиль. Значение числа в статистике, которым оканчиваются 25% наименьших значений статистических данных;
- *Upper Quartile для 12 значений* = $(12+1) * 3/4 = 8$ – верхняя квартиль. Значение числа в статистике с которого начинаются 25% наибольших значений статистических данных;
- $Range = (19 - 2) = 17$ – размах значений статистических дапнных;
- $Quartile Range = (8-3) = 5$ –интерквартильная широта разброса, или рассеяния, данных. Равняется разности между верхним и нижним квартилями;

3.4 Автоматическое вычисление параметров статистики

В студии STATISTICA анализ, например, данных о доходах, выполняется в следующей последовательности.

Создание таблицы

- 1) **Создать новую** таблицу *Файл / New -> Крупноформатная таблица*. На вкладке указать:
 - Число переменных = 1;
 - Число регистров (строк) = 12;
 - Размещение = *Как автономное окно*;
- 2) **Изменить переменную** – на имени (*Var1*) сделать *wClick* и изменить:
 - *Name* = *Доход*;
 - *Display format* = *General*;
 - *Type* = *Integer*.
- 3) **Сохранить таблицу** в своей папке с именем *3tDohFIO*
- 4) **Внести в таблицу** данные (доход 12 человек)

Выполнение анализа.

- 1) **Открыть панель** описательной статистики *Статистика / (Основная статистика/Таблицы) / Descriptive statistic (wClick)*;
- 2) **Выбрать переменную** – *Variable = 1- Доход*
- 3) **На вкладке Advanced** указать необходимые характеристики статистики (принятые в ручном анализе);
- 4) **Сформировать результат**- нажать *Summary: Descriptive statistic*. * Результат будет выведен в два новых окна:
 - *Report3tDohFIO1* – содержанием подключенного отчета.
 - *Date:* - индивидуальное окно с результатами анализа, рис.1.

Valid N	Mean	Geometric Mean	Median	Mode	Frequency of Mode	Minimum	Maximum	Lower Quartile	Upper Quartile	Range	Quartile Range	Std.Dev.
12	6,166667	4,738201	4,000000	3,000000	3	2,000000	19,00000	3,000000	8,000000	17,00000	5,000000	5,236816

Рис. 1 Результат обработки статистики

- **Корректировать** имя *Report3tDohFIO1* на *3rDohFio*.

4 Визуализация статистик

4.1 Визуализация описательных статистик

Графики коробок (ящик с усами) используются для визуализации описательных статистик. Такой график позволяет быстро оценить данные по структуре распределения, наличия неправдоподобных результатов, однородности результатов и т.д.

Построение графика средствами *STATISTICA* требует:

- 1) загрузить панель *Статистика / (Основная статистика / Таблицы) / Descriptive statistic*;
- 2) Выбрать вкладку *Options* и на ней указать одно:
 - *Median / Quart. / Range* - Медиана / Квартили / Размах;
 - *Mean / SE / SD* – Среднее / Ошибка среднего / Стандартное отклонение;
 - *Mean / SD / 1,96SD* – Среднее / Стандартное отклонение / Интервал 1,96 стандартного отклонения;
 - *Mean / SE / 1,96SE* – Среднее / Ошибка среднего / Интервал 1,96 ошибки среднего.
- 3) Выбрать вкладку *Quick* и нажать кнопку *Box & Whisker plot for all variables*.

В результате для выбранных опции, например, *Mean / SE / SD* получаем график (рис.1):

- Среднее значение
 $Mean = 6,17$;
- Ошибка среднего
 $\pm SE (4,7. .7,7)$;
- Стандартное отклонение среднего
 $\pm 1,96 * SE = (0,93 \dots 11,4)$.

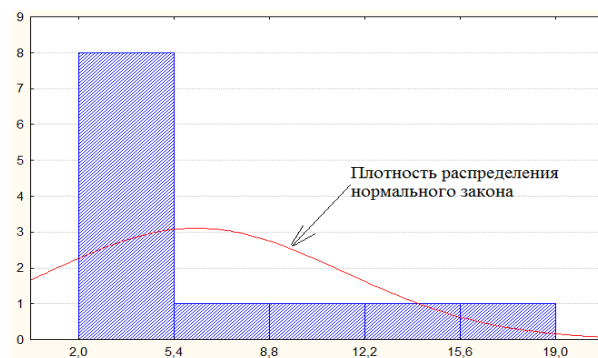


Рис. 2 Гистограмма

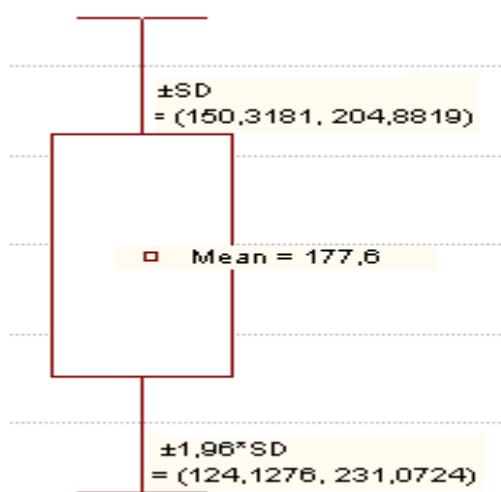


Рис. 1 Ящик с усами

4.2 Построение гистограмм

Графическое представление статистических данных в виде гистограмм позволяет упростить их понимание, увидеть закономерности в хаосе случайных явлений.

STATISTICA обеспечивает построение различных видов графиков. Перечень видов графиков загружается при выборе в главном меню «Графики» и включает три группы:

1) Плоские

- Гистограммы - на основе одной переменной;
- Графики рассеивания – функция одной переменной;
- Поверхность графиков – функции двух переменных.

2) Объемные

- 2d графики Гистограммы - на основе одной переменной;
- 3d последовательные графики;
- 3d XYZ графики

3) Другие.

Гистограмма является простейшей визуализацией статистики одной переменной. Построение выполняется в следующей последовательности:

- 1) Загрузить панель настройки – *Графики / Гистограммы*;
- 2) На вкладке *Быстрый* указать:
 - переменную – *Переменные / 1- Доход*;

- Тип графика – *Regular*
- Кол-во интервалов - Categories =10. *Размер интервала определяется автоматически $d=Range / Categories$;
- Запустить выполнение – *OK*
- *Будет получена гистограмма, см. рис. 2.

4.3 Построение круговых диаграмм

Круговые диаграммы статистических данных весьма показательны когда количество данных не велико. Данные на диаграмме показывают в интервалах и могут характеризоваться в процентах или фактических значениях.

STATISTICA обеспечивает построение одного вида круговой диаграммы, рис 3. Диаграмма разбита на секторы (интервалы). Каждый из них соответствует интервалу значений, например, (3,7 ; 5,4) и включает некоторое количество значений из исходной статистики, например, 3.

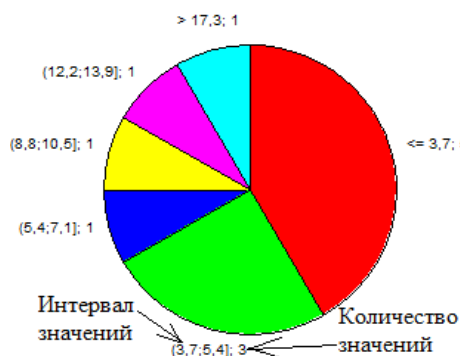


Рис.3 Круговая диаграмма

При построении и визуализации диаграммы могут использоваться различные подходы к определению интервалов и значений показателей в интервалах.

Построение диаграммы:

- 1) Загрузить панель - *Графики / 2DГрафики / Смещанные графики*
- 2) Выбрать вкладку *Быстрый*. *Она используется при быстром построении графика с параметрами по умолчанию.

Основные параметры (на вкладке *Быстрый*) включают:

- Переменные – *выбрать Рост, см;*
- Тип графика: *Pie Chat – Counts*, количество значений в интервале; *Pie Chat – Values*, значение в интервале (количество интервалов равно количеству значений).
- Разворот среза, *номер интервала*. Указанный интервал будет вырван из диаграммы.

Дополнительные параметры (вкладка *Дополнительно*)

- Легенда шрифта *Text and Value*, вывод на график интервала и значения в нем;
- Тип *3D* – построение объемной диаграммы.

4.4 Построение диаграммы рассеивания

Диаграммой рассеивания называется представление элементов выборки как точек на плоскости. Эти точки могут аппроксимированы – представлены некоторой функцией.

STATISTICA обеспечивает использование различных видов функций для аппроксимации, например, полинома 2й степени, рис. 4.

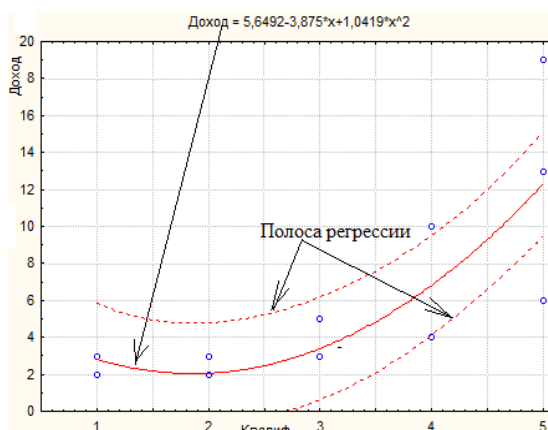


Рис.3 Диаграмма рассеивания

На диаграмме видны:

- Точки выборки;
- Математическое и графическое представление аппроксимирующей функции;
- Полоса регрессии, в которой с заданной вероятностью должны оказаться точки выборки.

Построение диаграммы

- 1) Загрузить панель - *Графики / Графики рассеивания;*
- 2) Выбрать вкладку *Быстрый*.

Основные параметры (на вкладке *Быстрый*)

включают:

- Переменные – указать имя переменной для аргумента (x) и для отклика (y).
- Тип графика: *Regular*, обеспечивает построение одного графика на форме; *Multiple*, обеспечивает построение нескольких графиков на одной форме.
- Полосы регрессии, выбрать *Предсказуемо*. Уровень - 0,5. Не менее 50% точек выборки должны находиться в полосе.

Дополнительные параметры (вкладка *Дополнительно*)

- Подгонка – указать вид функции, используемой для аппроксимации: *Linear* – линейная; *Polynomial* – полином (степень полинома *Опции 2 / Функции подгонки / Полиномиальный порядок*); другие.

5 Проверка гипотезы

5.1 Проверка гипотезы о виде распределения

Задача - имеется экспериментальная статистика для которой определены параметры нормального закона $F(x)$ соответствующие генеральной совокупности (теоретические).

Проверить гипотезу, что данные экспериментальной статистики соответствуют теоретическим получаемым из $F(x)$.

Для проверки гипотезы используем критерий согласия Пирсона *Chi-Square*, у которого уровень значимости p характеризует вероятность ее правильности.

Upper Boundary	Observed Frequency	Cumulative Observed	Percent Observed	Cumul. % Observed	Expected Frequency	Cumulative Expected	Percent Expected	Cumul. % Expected	Observed-Expected
<= 2,00000	0	0	0,00000	0,0000	0,204	0,2045	0,10226	0,1023	-0,2045
3,00000	1	1	0,50000	0,5000	2,040	2,2448	1,02015	1,1224	-1,0403
4,00000	8	9	4,00000	4,5000	11,59462	13,8394	5,79731	6,9197	-3,5946
...
13,00000	0	199	0,00000	99,5000	0,00008	200,0000	0,00004	100,0000	-0,0001
14,00000	0	199	0,00000	99,5000	0,00000	200,0000	0,00000	100,0000	-0,0000

Рис. 2 Таблица результатов Chi-Square

Гипотеза принимается если расчетный p превышает граничный p_z . ($p > p_z$).

В промышленной статистике принимается $p_z = 0,05$. В этом случае при выполнении условия $p > p_z$ ошибка принятой гипотезы возможна менее чем в 5% случаев исследований.

Проверка в среде студии STATISTIKA.

- 1) **Исходные данные** – таблица *5tZakonFIO* включает 200 значений переменной *Var1*.
- 2) **Настроить вывод** в индивидуальные окна (на панели *Менеджер вывода* оставить отметку *Индивидуальные окна / ОК*).

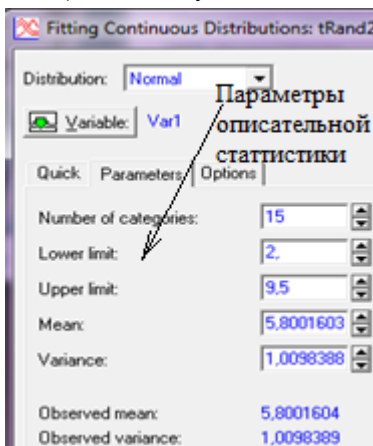


Рис. 1 Настройка проверки

3) **Загрузить панель** настройки проверки - *Статистика / Настройка распределения / Parameters*, (*Fittig Continuous Distributions*) рис. 1.

4) **На панели Distribution Fitting** выбрать нормальный закон – *Normal*.

5) **На панели Fittig Continuous Distributions**, вкладка *Parameters*, рис 1.

- Выбрать *Variable = Var1*.
 - Ознакомиться с параметрами описательной статистики:
- 6) **Запустить выполнение** проверки – *Summary*. *Будет сформирована таблица результатов.

5.2 Таблица результатов критерия Пирсона

Таблица результатов проверки по критерию Пирсона *Chi-Square*, сформированная в разделе 1, включает шапку и строки, рис.2. Количество строк в таблице равно количеству количеству интервалов *Number of categories*.

Данные таблицы сведены в 4ре группы:

- 1) Границы интервалов;
- 2) Экспериментальные данные – *Observed* (4ре переменных);
- 3) Теоретические данные – *Expected* (4ре переменных);

4) Различие теоретических и экспериментальных данных (*Observed – Expected*) for Frequency (1 на переменная).

В шапке таблице приведены характеристики *Chi-Square* :

- *Chi-Square* – значение критерия, например, 14,355;
- *df* – число степеней свободы, например, 3 (*adjusted*);
- *p* – вычисленный уровень значимости, например, 0,00246.

Анализ уровня значимости ($p=0,00246 < p_z=0,05$) в критерии Пирсона говорит о недопустимом различии теоретической и экспериментальной $F(x)$ функций.

Одной из причин может быть ошибка в выборе вида закона распределения. Проверка экспериментальной статистики на соответствие другому закону распределения выполняется в следующей последовательности.

- **Загрузить панель** - *Статистика / Настройка распределения / Parameters, (Fittig Continuous Distributions)* рис. 1.
- **В окне *Distribution*** - выбрать другой закон распределения.
- **Запустить выполнение** проверки – *Summary*. *Будет сформирована таблица результатов.
- **Выполнить анализ** уровня значимости *p*.

5.3 Проверка однородности выборок

Пусть имеются выборки полученные из различных совокупностей. Например, примеси *P1..P4* в одной марке инструментальной стали выплавленной на двух заводах *Z1, Z2*, приведенные в таблице *5tStill*, рис.4.

	1 P1	2 P2	3 P3	4 P4
Z1	82	535	1173	1714
Z2	63	429	995	1307

Рис. 4 Количество примесей в стали

Необходимо проверить гипотезу, что исходные совокупности распределены одинаково.

Проверка гипотезы с использованием критерия Пирсона выполняется в модуле *Статистика / (Дополнительные линейные/нелинейные модели) / Регистрационно - линейный анализ таблиц частоты*.

На открывшейся панели *Log-Linear Analysis*:

- *Input file* – выбрать *Frequencies w/out coding variables (частоты без кодирующих переменных)*;
- *Variables* – выбрать *Select All*;
- *Specify table* – нажать и открыть панель, рис. 4.

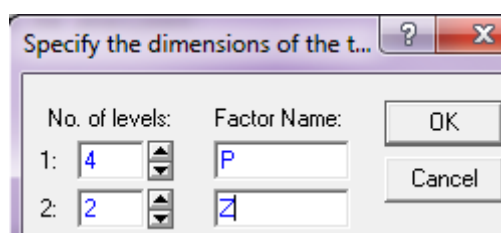


Рис. 4 Количество примесей в стали

В колонку *No. of level:* в 1ю строку вводим кол-во переменных (4); во 2ю – кол-во строк (2).

В колонку *Factor Name* – коды переменных и строк.

Нажать *Ок, Ок* и на вкладке *Advanced* нажать кнопку *Test all marginal....*

Результат анализа будет выведен в таблицу *Result of Fitting....*, где во второй строке приведены показатели критерия:

- *Degrs. of Frydom* = 3 – количество степеней свободы;
- *Max.Lik. Chi-squ.* = 3.59 – значение статистики *Chi-Square*;
- Уровень значимости $p = 0,3088$.

Выполнение условия $p=0.3088 > p_z=0.05$ позволяет принять гипотезу о одинаковом распределении содержания примеси в стали полученной на различных заводах.

5.4 Выявление причин низкого уровня значимости

Нормальный вероятностный график позволяет визуально оценить соответствие экспериментальных данных требованиям нормального закона распределения.

Построение графика средствами *STATISTICA* требует:

- 1) Загрузить панель *Статистика / (Основная статистика / Таблицы) / Descriptive statistic /*);
 - 2) Выбрать переменную *Variable = Var1*;
 - 3) Выбрать вкладку *Prob. & Scatterplots* и на ней нажать кнопку *Normal probability plots*.
- *Будет построен график, рис.5.

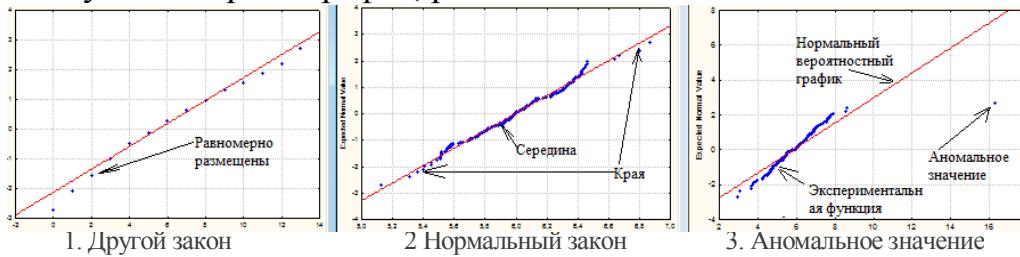


Рис. 5 Использование нормального вероятностного графика

На рисунке приведены характерные случаи поведения экспериментальной функции.

- 1) Особые свойства соответствует другому

- 1) закону распределения.
- 2) Плотные в середине и редкие по краям, симметричные – соответствуют нормальному закону;
- 3) Имеется аномальное значение – говорит о ошибке ввода данных.

Использование гистограммы лежит в основе другого способа анализа экспериментальных данных, рис.6.



Рис. 6 Варианты экспериментальных данных

Внешний вид гистограммы позволяет сделать вывод о соответствии данных нормальному распределению.

6 Связи между пе-

ременными

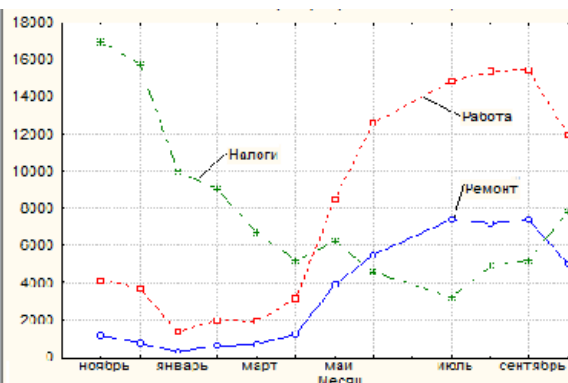
6.1 Корреляционный анализ

Корреляция определяет степень с которой значения двух переменных пропорциональны друг другу.

Коэффициент корреляции r – показатель оценивающий тесноту линейной связи между признаками. Лежит в интервале $(-1..1)$. Знак означает (+) – связь прямая; (-) – обратная. Чем ближе значение к $|1|$ тем теснее линейная связь. Менее 0,3 - связь слабая, более 0,91 – очень тесная.

Исходные данные представлены таблицей бт Ремонт имеющей аргумент (месяц) и отклики (Ремонт, Работа, Налоги) рис.1 . Необходимо оценить связь между этими признаками.

1	2	3	4
Месяц	Ремонт	Работа	Налоги
1 ноябрь	1130,01	4137,63	16952,32
2 декабрь	734,42	3704	15732,32
3 январь	265,4	1329,4	10021,3
4 февраль	586,6	1961,6	9026,65
5 март	666,7	1939,7	6682,3
6 апрель	1232	3116	5214,72
7 май	3888,35	8509,35	6205,71
8 июнь	5465,39	12588,89	4582,14
9 июль	7412,39	14875,5	3240,3
10 август	7168,66	15388,08	4952,44
11 сентябрь	7416,68	15450,67	5215,37
12 октябрь	5001,41	11944,82	7832,32



а. Табличное представление

в. Графическое представление

Рис. 1 Исходные данные к корреляционному анализу

Корреляционный анализ в студии STATISTICA выполняются с использованием модуля *Статистика / (Основная статистика/Таблицы) / Correlation matrices*.

Variable	Месяц	Ремонт	Работа	Налоги
Месяц	1,00	0,88	0,87	-0,74
Ремонт	0,88	1,00	0,99	-0,64
Работа	0,87	0,99	1,00	-0,57
Налоги	-0,74	-0,64	-0,57	1,00

Рис. 2 Коэффициенты корреляции

На панели выполняем: Выбор переменных - *One variable list = Select All*; Запуск анализа – *Summary*.

В результате имеем корреляцию: Очень тесную $r=0.99$ (ремонт/работа); Слабую $r=-0.57$ (работа/налоги).

6.2 Простой (однофакторный) регрессионный анализ

Регрессионный анализ заключается в определении аналитического выражения в котором изменение одной величины (отклика) Y обусловлено влиянием другой величины x (фактор).

Однофакторная (парная) регрессия имеет один фактор и представлена выражением

$$Y=f(x) + \varepsilon,$$

где $f(x)$ - неслучайная составляющая отклика Y зависящая от x ;

ε – остаток или случайная составляющая обусловленная влиянием на отклик множества неучтенных факторов и ошибками измерений.

Парная линейная регрессия представлена выражением

$$Y= b_0 + b_1 x + \varepsilon,$$

где b_0, b_1 – называются коэффициентами регрессии.

Регрессионный анализ в студии STATISTICA выполняют с использованием модуля *Статистика / Множественная регрессия*. *Форма *Multiple Regression Results*.

На вкладке *Advanced* выполняем:

- Variables: *depended = Ремонт* (отклик); *independent = Работа* (фактор) - *Ok*.
- Input file - *Raw Date*;
- Запуск анализа – *Ok*.

В результате имеем форму с результатом анализа в виде текста в ее верхней части и кнопками в нижней.

На вкладке *Advanced* нажать *Summary:Regression results*.

Откроется таблица с результатами анализа в которой:

- колонка *Beta = 0,9935* содержит коэффициент корреляции;
- колонка *B* содержит коэффициенты регрессии $b_0=-588$ (верхняя) ; $b_1=0,506$ (нижняя);
- колонка *p-level =0,0077*-уровень значимости.

При $p\text{-level} < 0,05$ выборка считается однородной и уравнение регрессии принимается к использованию (красный).

6.3 Оценка адекватности модели по остаткам

Остатки это разность между исходными (наблюдаемыми) значениями отклика и предсказанными из модели.

Остатки должны быть нормально распределены, иметь нулевое среднее значение и постоянную дисперсию, независимо от значений аргумента и отклика.

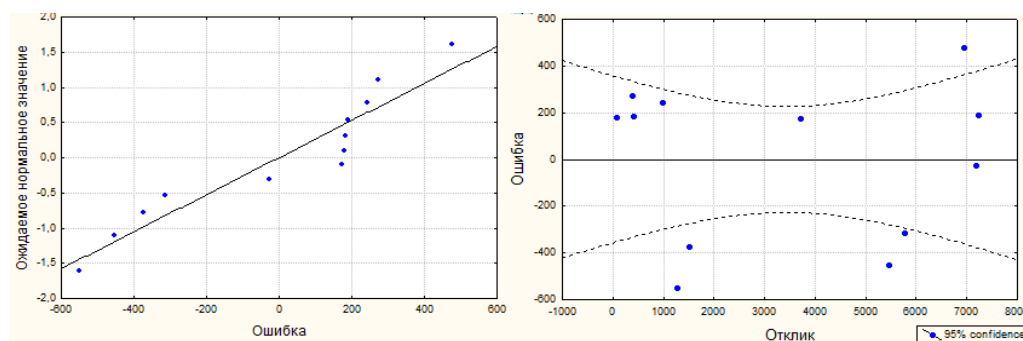
Исследование остатков это продолжение регрессионного анализа на панели *Multiple Regression Results*.

Настройка анализа остатков:

- выбрать вкладку *Residuals / assum....*;
- нажать кнопку *Perform residual analysis*. *Загрузится панель *Residual Analysis*.

Выполнение анализа остатков на панели *Residual Analysis*:

Строим нормальный вероятностный график остатков кладке *Quick* нажать кнопку *Normal plot of residuals*, рис.3а.



а. Нормальный вероятностный график

в. Зависимость остатков от отклика

Рис. 3 Анализ остатков

На рис.3а видно, что что остатки близки к прямой соответствующей нормальному закону.

Строим график зависимости остатков от значения отклика. На вкладке *Scatterplots* нажать кнопку *Predicted*

vs. residuals рис.3в. Как видно остатки хаотично разбросаны относительно прямой и нет резко выделяющихся остатков.

Из поведения остатков на рисунках следует, что предположение о нормальном законе является верным.

6.4 Сложный регрессионный анализ

Сложные регрессионные анализы можно объединить в ряд групп.

Многофакторный регрессионный анализ используется в сложных системах с большим количеством факторов. В рамках анализа могут использоваться факторы; в явном виде – x_1, x_2, x_3 ; обратные к ним $-1/x_1, 1/x_2, 1/x_3$; логарифмы – $\ln(x_1), \ln(x_2), \ln(x_3)$.

Метод пошагового включения предполагает последовательный ввод в модель или исключение факторов в явном или преобразованном виде.

Рациональность включения фактора в модель (степень его независимости) оценивается коэффициентом корреляции R , который должен возрасти и уровнем значимости p -level, который должен быть менее $0,05$.

Гребневая регрессия предполагает получение смещенных оценок (b_0, b_1) модели которые имеют меньшую дисперсию. Такие оценки позволяют получить более точные и приемлемые для практического использования модели.

Показатель λ характеризует смещение оценки модели по отношению к классической, полученной методом наименьших квадратов.

Итерационный (пошаговый) метод определения λ предполагает его увеличение, начиная с $\lambda = 0$, на некоторую величину, например, $0,001$.

На каждом шаге, используя коррелированные значения показателей модели, определяют коэффициент корреляции R , который должен возрасти. Если рост останавливается – дальнейшее изменение λ не целесообразно.

Наилучшая корреляционная модель – понятие субъективное. Однако при ее выборе используют значение: коэффициента корреляции $R > 0,91$; уровень значимости p -level $< 0,05$, остатки должны иметь нормальное распределение, др.

7 Диаграммы качества

7.1 Анализ причинно – следственных связей

Диаграмма причин и результатов (Исикавы) показывает отношение между показателями качества и воздействующими на него факторами.

Объектами исследования с использованием причинно – следственных диаграмм могут быть: рост доли дефектных деталей; увеличение затрат на устранение брака; снижение прибыли от реализации продукции; руководство коллективом..

Диаграмма Исикавы представлена структурой, имеющей несколько уровней, рис.1.

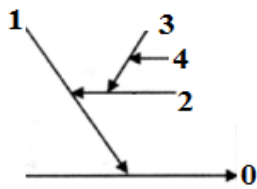


Рис.1. Диаграмма Исикавы

- 0-решаемая проблема;
- 1-главная причина;
- 2-причина вторичного порядка;
- 3-причина третичного порядка;
- прочие уровни причин.

Информацию причин и результатов до второго уровня можно

представить в виде таблицы 7т Сун, рис.2.

Свойства

таблицы

позволяют

представить

данные только 3-х

уровней. На рис.2

показаны:

0-решаемая проблема

(сварить суп):

2 -

главные причины (5-ть шт.);

причины вторичного порядка

(от 2х до 8ми шт.).

Для анализа причин третичного

и выше порядков строят доролнительные

таблицы, где 2й уровень принимается нулевым.

Варка супа					
1	2	3	4	5	
Продукты	Технология варки	Условия варки	Повар	Оборудов. кухни	
1	Вода	Закрытая крышка кастрюли	Время	Опыт работы	Исправная плита
2	Соль	Сначала варить бульон	Температура плиты	Квалификация	Большая кастрюля
3	Мясо	Снимать пену при варке бульона		Знание рецепта	Острый нож
4	Картофель	Нарезать продукты ножом			Доска для нарезки
5	Морковь	Не допускать разваривания			
6	Лук				
7	Лапша				
8	Приправы				

Рис.2. Данные о причинах и результатах

7.2 Построение диаграммы Исикавы

Построение диаграммы Исикавы в студии STATISTICA выполняются с использованием модуля *Статистика / (Индустриальная статистика & Сигма 6) / Анализ процесса / Cause-effect (Ishikawa..)*. *Загрузится панель *Cause – end – Effect*.

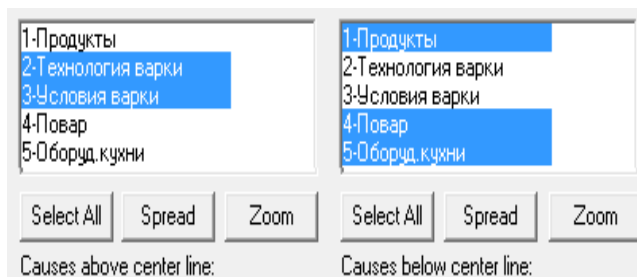


Рис.4. Диаграмма «Суп»

На панели нажать *Variables* и выбрать размещение на диаграмме главных причин относительно центральной линии, (*above, below*) рис. 3.

Запустить **фор-мирование** диаграммы – *Ок, Ок*. *Будет сгенерирована диаграмма, рис.4.

Автоматическая корректировка вида диаграммы. До нажатия *Ок* на панели указать настройки:

- *Arrows* – характеристики стрелок;
- *Font sizes* – размер символов для: заголовка; главной причины; причины вторичного порядка.

Ручная корректировка содержания диаграммы

выполняется на ее сгенерированном рисунке.

Правила редактирования рисунка

- новые объекты (текст, стрелки) создаются путем копирования существующих;
- редактирование существующих объектов выполняется по правилам близким к используемым в графическом редакторе *Word*.

7.3 Диаграмма Парето

Парето обнаружил, что за 80% результата отвечают 20% причин.

Результатом процесса могут быть: благосостояние населения, затраты на ремонты единиц оборудования, трудоемкость изготовления деталей машин и т.д. Обязательным

требованием является наличие большого количества составляющих влияющих на процесс.

Например, процесс (стоимость ремонта оборудования) в зависимости от вида устраняемого дефекта (сбой, износ, погрешность, повреждение цепи и т.д.).

Закон Парето можно использовать для широкого круга процессов:

- 80% покупок делают 20% покупателей;
- 80% работы делают 20% сотрудников;
- 80% потерь на производстве дают 20% видов дефектов, а 80% видов дефектов обуславливают остальные 20% потерь.

Диаграмма Парето – представляет факторы влияющие на процесс в виде причин с последовательно убывающей весомостью, рис.5.

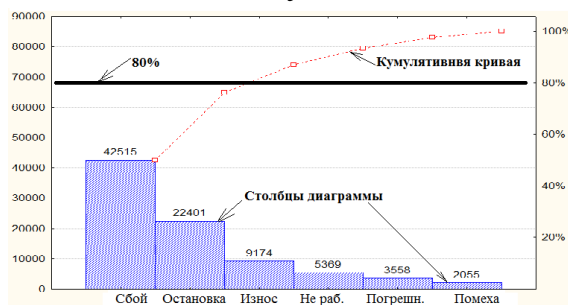


Рис.5. Диаграмма Парето

На диаграмме пока-зывают столбцы, каждый из которых характеризует одну причину, а его размер – соответствует абсолютной величине влияния на процесс.

Кумулятивная кривая показывает рост суммарных потерь при увеличении количества рассматриваемых дефектов. Она позволяет выделить перечень причин оказывающих 80% влияния на процесс и главную причину в наибольшей степени влияющую на процесс.

Например, причина – «Сбой» приводит к потере 42 515 руб.

7.4 Построение диаграммы Парето в STATISTICA

Исходные данные к построению диаграммы представлены в виде таблицы 7е Парето (рис.5.). Например, она включает показатели:

	Станок	Дефект	Потери, руб
1	КТ 4	Сбой	6241
2	Фр 8	Погрешн.	3558
3	КТ4	Сбой	4520
4	Фр 8	Остановк	3850

19	КТ 4	Помеха	2055
20	Св 2	Сбой	1880

Рис.5. Данные к диаграмме Парето

- процесса, который является откликом (потери, руб.);
- факторов влияющих на него (станок, дефект).

Таблица имеет одну переменную характеризующую процесс и несколько факторов характеризующих причины влияющие на процесс.

Построение диаграммы Парето в студии STATISTICA выполняются с использованием модуля

Статистика / (Индустриальная статистика & Сигма б)/ Качество диаграммы управления. *Загрузится панель *Quality a chart specification file*.

На вкладке **Quick** выбрать *Pareto chart analysis*. *Откроется панель *Defining Variables for Pareto Chart*, рис.6. **Панель используется** для указания требований к виду

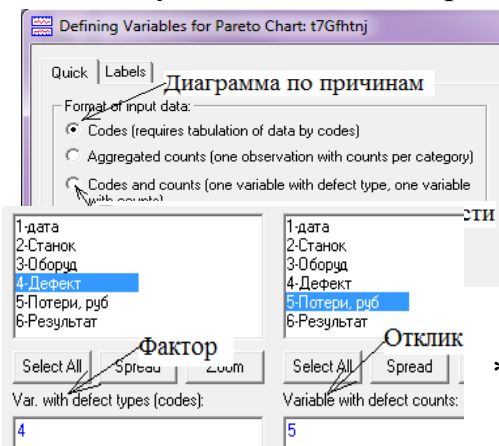


Рис.6. Выбор переменных

диаграммы:

- 1) только по одной причине – одна переменная;
 - 2) по причине (фактор) и стоимости (отклик) – две переменные:
- фактор – Дефект;
 - отклик – Потери, руб.

Запустить формирование диаграммы – **OK**.

*Диаграмма имеет вид показанный на рис.4.

8 Карты контроля качества

8.1 Содержание контрольной карты

Статистически подконтрольный процесс изготовления продукции протекает при его хорошем планировании и правильной

реализации. В этом случае каждый фактор влияющий на качество продукции во время ее изготовления колеблется в некоторых пределах и описывается законом распределения (сохраняется *status quo*).

Однако возможны ситуации когда появляется неслучайная причина приводящая к нарушению процесса изготовления и его выходу из состояния статического контроля.

Для поиска причины в такой ситуации используются специальные методы, например, контрольные карты.

Контрольные карты – предполагают непрерывный контроль во времени за производственным процессом для выявления ситуации когда нарушен *status quo*, рис.1.

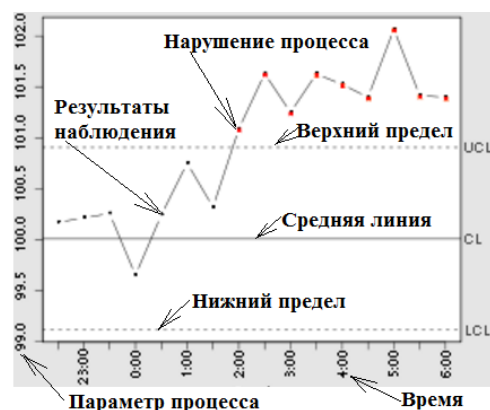


Рис.1. Пример контрольной карты

Ключевые компоненты контрольной карты:

- Оси: аргумент – время; отклик – параметр процесса;
- Контрольные границы: UCL – верхняя, LCL - нижняя; CL – средняя линия.
- Результаты наблюдения;
- Состояние процесса: норма-льное – между UCL и LCL; нарушенное – за пределами UCL и LCL.

Исходные данные к построению контрольной карты представлены таблицей *8t Карта Шухарта* включающей две переменных: аргумент – время; отклик – параметр t^0C процесса, рис. 2.

8.2 Контрольная карта индивидуальных значений

Анализ контрольной карты в студии STATISTICA выполняются с использованием модуля *Статистика / (Индустриальная статистика & Сигма б)/ Качество диаграммы управления*. *Загрузится панель *Quality Control Charts*.

На вкладке *Quick* выбрать *Industryals & moving range*.

На открывшейся вкладке нажать *Variables* и выбрать *Measurements = t⁰C*;

Запустить построение – нажать *Ок*.

Результат анализа представлен двумя группами изображений рис.3

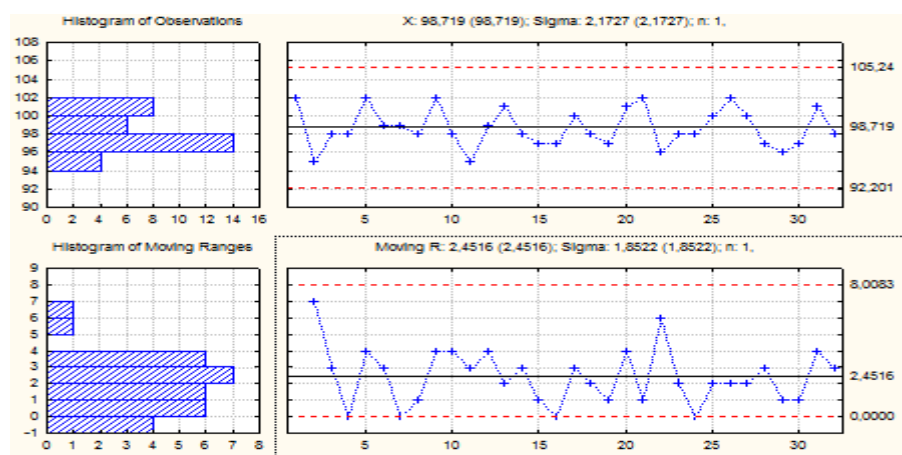


Рис.3. Контрольная карта индивидуальных значений

X- параметр отражает t^0C

исследуемого процесса и включает :

- Гистограмму значений показателей процесса;
- Графическое представление показателей процесса;
- Параметры нормального закона распределения : математическое ожидание $X=98,719$; среднеквадратическое отклонение $Sigma = 2.1727$. Горизонтальными линиями показаны верхняя граница соответствующая $3sigma$ и нижняя.

R – изменение параметра X между двумя результатами.

8.3 Интегральные контрольные карты

Контрольная карта средних значений – применяется при массовом производстве. Она позволяет делать выводы о характеристиках процесса на основе малых выборок из большого числа рассматриваемых единиц продукции.

Порядок построения карты в студии STATISTICA.

- 1) Все единицы продукции делятся на группы (по сменам);
- 2) Из каждой группы делается случайная выборка 5 единиц и строится карта.

Анализ контрольной карты в студии STATISTICA выполняют на панели *Quality Control Charts*, вкладка *Quick*.

- 1) На вкладке выбрать *X-bar & R chat for variablres*.
- 2) На открывшейся вкладке нажать *Variables* и выбрать: *Measurements = процессу; Samplly Idents = номера групп* ;
- 3) Запустить построение – нажать *Ок*. * Будет построена контрольная карта с четырьмя группами.

Контрольная карта накопленных сумм – это карта с памятью. Она более чувствительна к возмущениям и позволяет выявить сдвиг настройки процесса в самом начале.

Анализ контрольной карты в студии STATISTICA выполняют на панели *Quality Control Charts*, вкладка *Variables* .

На вкладке выбрать *CuSum chat for individuals* .

На открывшейся вкладке нажать *Variables* и выбрать: *Measurements = процессу; Part Idents = номера групп*

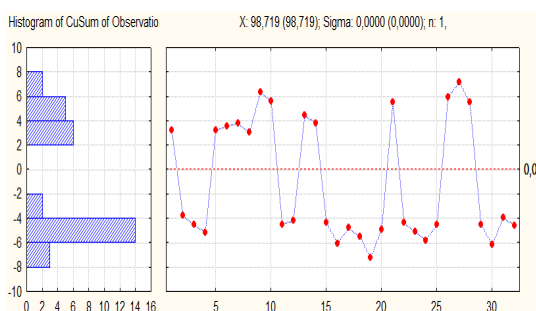


Рис.4. Контрольная карта накопленных сумм

рассматривается изменение рис.5.

Для стабильного процесса контрольные границы не меняются. Выход за границы говорит о причине.

Признаки причин.

- 1) За пределами регулирования находятся X точки и соответствующие им R точки (рис.5, точки 1, 4). Причина – за счет обычных (внутренних) причин увеличилось технологическое рассеивание, т.е. увеличилась сигма.
- 2) За пределами регулирования находятся X точки, а R точки лежат в границах (рис.5, точки 3, 6). Причина - воздействие внешнего фактора.
- 3) Дрейф – 7 и более поднимающихся или ниспадающих точек (рис.5, точки 2, 5). Причина – рост температуры, износ оборудования и др.
- 4) За пределами X и R точки. Причина наличие как обычных так и особых причин нарушающих процесс.

Особые случаи.

Запустить построение – нажать *Ок*. * Будет построена контрольная карта с накопленной суммой.

На карте 6 сигм делятся на 20 равных частей и оценивается количество попаданий в эти части, рис.4.

8.4 Чтение контрольных карт

Анализ контрольных карт выполняется для выявления факта выхода процесса из-под контроля и определения причин выхода.

В рамках анализа параметра X и размаха R , вероятность выхода за 0,01.

о воздействии особой

находятся X точки и

и R точки (рис.5, точки 1, 4). Причина – за счет обычных (внутренних) причин увеличилось технологическое рассеивание, т.е. увеличилась сигма.

находятся X точки, а R точки лежат в границах (рис.5, точки 3, 6). Причина - воздействие внешнего фактора.

7 и более поднимающихся или ниспадающих точек (рис.5, точки 2, 5). Причина – рост температуры, износ оборудования и др.

наличие как обычных так и особых причин нарушающих процесс.

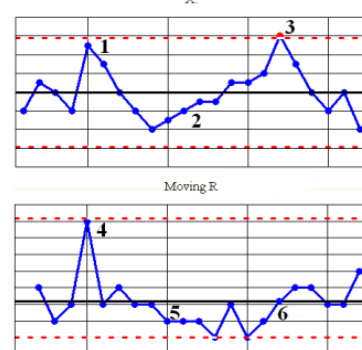


Рис.5. Контрольная карта

- 1) Серия точек – последовательные точки лежат по одну стороу средней линии. Например - более 15ти тточек из 20-ти лежат по одну сторону от средней линии. Причина – внешнее воздействие сдвигающее центр рассеивания от средней линии.
- 2) Периодичность точек – наличие подъемов и спадов с одинаковыми интервалами. Причина – воздействие внешнего периодически изменя.щегося фактора.