**ГОУ ВПО**

**«ДОНЕЦКИЙ НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ»**

**МЕТОДИЧЕСКИЕ УКАЗАНИЯ**

**по выполнению лабораторных работ**

**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»**

**МЕТОДИЧЕСКИЕ УКАЗАНИЯ**

**по выполнению лабораторных работ**

**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 « » 201

**Донецк, 2016 г.**

УДК 65.012

**Методические указания** по выполнению лабораторных работ 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, А.А.Истрати - Донецк: ДонНТУ, 2016. - 27с.

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

Составители: НА. Ченцов, проф.

А.А.Истрати, ассист.

Отв. за выпуск: Н.Ф. Годына, доц.

Рецензенты: Е.В. Мирошниченко, доц.

СОДЕРЖАНИЕ

1 Инсталляция пакета STATISTICA 6

1.1 Оценка уровня студентов 6

1.2 Подготовка файловой системы к инсталляции 6

1.3 Инсталляция пакета STATISTICA 6

2 Знакомство с пакетом STATISTICA 7

2.1 Предмет статистики 7

2.2 Программные продукты 7

2.3 Компоненты пакета STATISTICA 8

3 Модули пакета 9

3.1 Характеристика модулей пакета 9

3.2 Определение среднего значения переменной 9

4 Вывод результатов анализа 10

4.1 Виды вывода результатов анализа 10

4.2 Работа с «Окнами» 10

4.3 Работа с «Книгой» 10

4.4 Работа с «Отчетом» 11

5 Подготовка к анализу описательной статистики 12

5.1 Работа в таблице данных 12

5.2 Формирование таблицы данных 13

5.3 Выбор параметров в STATISTICA 13

6 Вычисление параметров описательной статистики 14

6.1 Ручное вычисление параметров 14

6.2 Автоматизированное вычисление параметров 14

7 Визуализация описательных статистик 16

7.1 Виды гистограмм 16

7.2 Построение гистограмм 16

8 Построение диаграмм 17

8.1 Построение круговых диаграмм 17

8.2 Построение диаграммы рассеивания 18

9 Критерий Пирсона 18

9.1 Проверка гипотезы о виде распределения 18

9.2 Таблица результатов критерия Пирсона 19

10 Сравнение выборок 19

10.1 Проверка однородности выборок 19

10.2 Выявление причин низкого уровня значимости 20

11 Связи между переменными 21

11.1 Корреляционный анализ 21

11.2 Простой (однофакторный) регрессионный анализ 21

12 Оценка модели 22

12.1 Оценка адекватности модели по остаткам 22

12.2 Сложный регрессионный анализ 22

13 Диаграммы качества 23

13.1 Анализ причинно – следственных связей 23

13.2 Построение диаграммы Исикавы 24

14 Диаграмма Парето 24

14.1 Диаграмма Парето 24

14.2 Построение диаграммы Парето в STATISTICA 25

15 Карты контроля качества 25

15.1 Содержание контрольной карты 25

15.2 Контрольная карта индивидуальных значений 26

16 Развитые карты контроля качества 26

16.1 Интегральные контрольные карты 26

16.2 Чтение контрольных карт 27

# Инсталляция пакета STATISTICA

## Оценка уровня студентов

**Каждый студент должен.**

1. **Сообщить о возможности** использования: интернета для общения; ноутбука на занятиях; телефона, планшета для чтения методических указаний.
2. **Прислать Ченцову** на адрес [***prof2007chen@gmail*.com**](mailto:prof2007chen@gmail.com) тестовое письмо с текстом «Сообщаю электронный адрес Иванова Р.П., гр. КСМС 16 для материалов по дисциплине «Программные статистические комплексы».
3. **Придумать себе** идентификатор. Например, у Ченцова идентификатор *Chen*.

**В классе выполнить.**

1. Загрузить ПК.
2. На ПК открыть файл d:/qm/chen/STATISTICA/*3лабор ПрограммныеСтатКомплексы*.doc.

## Подготовка файловой системы к инсталляции

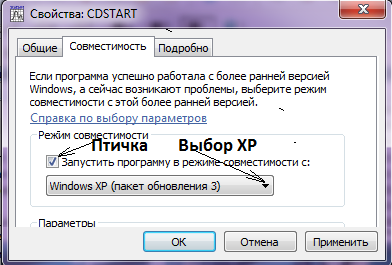
**Подготовка файлов**:

* получить у преподавателя папку *qm16* и переписать ее на ПК студента;
* на своем ПК в папке *qm16* найти папку *StatFIO* и заменить *FIO –* на инициалы студента. Например, у Ченцова получим *StatChen.*
* в папке *StatFIO* создать текстовый файл *z1FIO.doc* для индивидуального задания. \*В первой строке написать «Индивидуальное задание №1 ФИО»
* распаковать *файл Statistica\_6.0.zip* в папку *qm16.* **\***В результате в папке *qm16* получим инсталляционную папку *Statistica\_6.0*.

**Обеспечить совместимость** STATISTICA 6.0c Windows 7 и старше (в случае Windows XP– совместитмость пропустить):

* **в** папке *Statistica\_6.0* **выбрать файл** *CDSTART* –установить на него курсор;
* открыть свойства файла *– rClick / свойства;*
* **перейти на вкладку *Совместимость*, рис.1**
* **установить «птичку» возле «Запустить программу в режиме совместимости с:**

Рис. 1 Совместимость XP



* **Под птичкой выбрать «Windows XP» (пакет обновления 3).**
* **Нажать «ОК». \*Будет установлена совместимость.**

## Инсталляция пакета STATISTICA

**Инсталляция пакета из** папки *Statistica\_6.0 (*логин и ключи в файле *Content*.*txt*)**:**

* **запустить файл** *CDSTART* –сделать *wClick* на нем. \*Откроется форма инсталляции;
* запустить инсталляцию – нажать *Install STATISTICA*. \*Начнется инсталляция.
* Подтверждать все требования до выхода из инсталляции.
* После окончания инсталляции на столе должна появиться пиктограмма «STATISTICA», а в файловой 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 *R****USSIAN*.EXE. \*Будет введен русский текст.

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

**Задание – сделать копию экрана с запущенной** *STATISTICA* **и разместить в файле** *z1….*Файл *z1…* прислать преподавателю на eMail

# Знакомство с пакетом STATISTICA

## 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** характеристикикаждой 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. (Cохранение *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.

## Программные продукты

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

**Универсальное приложение –Microsoft Excel** [программа](https://ru.wikipedia.org/wiki/%D0%9A%D0%BE%D0%BC%D0%BF%D1%8C%D1%8E%D1%82%D0%B5%D1%80%D0%BD%D0%B0%D1%8F_%D0%BF%D1%80%D0%BE%D0%B3%D1%80%D0%B0%D0%BC%D0%BC%D0%B0) для работы с [электронными таблицами,](https://ru.wikipedia.org/wiki/%D0%AD%D0%BB%D0%B5%D0%BA%D1%82%D1%80%D0%BE%D0%BD%D0%BD%D0%B0%D1%8F_%D1%82%D0%B0%D0%B1%D0%BB%D0%B8%D1%86%D0%B0) которая предоставляет возможности статистических расчетов и построения графиков.

**Специализированные пакеты** разработанные и применяемые для реализации только статистических методов.

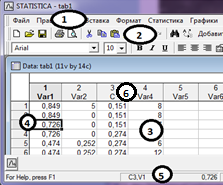
**STATISTICA,** [пакет](https://ru.wikipedia.org/wiki/%D0%9F%D0%B0%D0%BA%D0%B5%D1%82_%D0%BF%D1%80%D0%B8%D0%BA%D0%BB%D0%B0%D0%B4%D0%BD%D1%8B%D1%85_%D0%BF%D1%80%D0%BE%D0%B3%D1%80%D0%B0%D0%BC%D0%BC) для реализующий функции [анализа данных](https://ru.wikipedia.org/wiki/%D0%90%D0%BD%D0%B0%D0%BB%D0%B8%D0%B7_%D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D1%85), [визуализации данных](https://ru.wikipedia.org/wiki/%D0%92%D0%B8%D0%B7%D1%83%D0%B0%D0%BB%D0%B8%D0%B7%D0%B0%D1%86%D0%B8%D1%8F_%D0%B4%D0%B0%D0%BD%D0%BD%D1%8B%D1%85) с привлечением [статистических методов](https://ru.wikipedia.org/wiki/%D0%A1%D1%82%D0%B0%D1%82%D0%B8%D1%81%D1%82%D0%B8%D0%BA%D0%B0). Пакет позволяет выводить информацию в виде различных типов графиков (научные, деловые, трёхмерные и двухмерные графики, специализированные статистические графики — гистограммы, матричные, категорированные графики и др.).

**Задание** - создать файл z2FIO.doc . \*В первой строке написать «Индивидуальное задание №1 ФИО»

## Компоненты пакета STATISTICA

**Студия пакета** включает следующие компоненты (рис. 1):

Рис. 1Стартовая форма



1. меню пакета;
2. кнопки управления;
3. таблица данных в рабочей зоне;
4. активная клетка;
5. характеристика активной клетки;
6. 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 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.

**Загрузить пакет.**

* **Запуск** –файлом *C:\Program Files (x86)\StatSoft\STATISTICA 6\ statist.exe"* или пиктограммой со стола.
* **В рабочую зону** загружается: таблица, открытая при последнем закрытии пакета.

**Открыть существующую таблицу.**

1. Закрыть все таблицы – нажать *Х* в правом верхнем углу.
2. Выбрать таблицу – *Файл / открыть*. \*В файловой системе ПК найти необходимую таблицу *qm16 / StatFIO* / *1tFriendFIO.sta.*
3. Открыть таблицу – сделать на ней *wClick*. \*В рабочем окне загрузится выбранная таблица.

**Задание – в файле** z1FIO.doc указать под именем загруженной таблицы**.**

1. 6й элемент меню пакета;
2. Номер пиктограммы печати;
3. Имя 4й переменной;
4. Значение 4й переменной в 3й строке.

**Корректировка таблицы.**

1. Сохранить таблицу со своим идентификатором. \*У Ченцова -*1tFriendChen.sta* ;
2. Закрыть таблицу и пакет;
3. Открыть пакет и загрузить свою таблицу;
4. В имена переменных добавить свой идентификатор;
5. Добавить одну строку (*rClick* на последней строке, нажать *Add Cases*, подтвердить ввод 1й строки).
6. В новую строку внести данные о своем друге.
7. Создать строки и ввести 2х друзей;

**Создание переменных в открытой таблице.**

1. На имени последней переменной нажать *rClick* и выбрать *Add Variables.*
2. Выбрать характеристики переменной: *Name – размерFIO*; *Display format – General; Type – Integer.*
3. Сохранить – *Ок*.
4. В поле *размерFIO* ввести размер обуви, сохранить.

После ввода данных, закрыть пакет и пригласить преподавателя.

**Задание** - в файл *z1FIO.doc* скопировать экран с модифицированной таблицей.

**Задание** – оформленный файл отчета *z1FIO.rtf* отправить преподавателю на почту.

# Модули пакета

## Характеристика модулей пакета

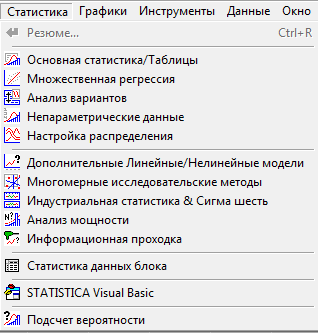
**Создать файл** *z2FIO.doc* (отчет) с первой записью «z2FIO Модули пакета STATISTICA»

**Загрузить пакет** и таблицу*1tFriendFIO.sta*

**Перечень модулей пакета** открывается кнопкой «Статистика» в главном меню. Он выводится на панели, рис.1. и включает 13 компонентов:

* основная статистика/таблицы *Basic statistic* – обеспечивает получение основных показателей описательной статистики;

Рис. 1 Модули пакета

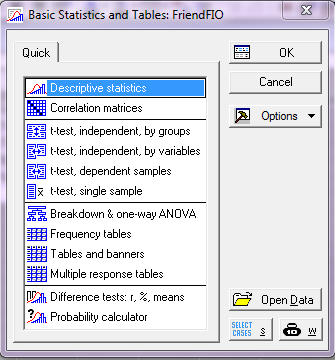


* множественная регрессия – показывает связь аргумента и отклика;
* индустриальная статистика – специализированный модуль для решения задач управления качеством;
* прочие.

**Задание** – студенту, по номеру в журнале загрузить соответствующий модуль. Его копию разместить в отчете, подписать наименование модуля и кол-во активных компонентов в нем.

**Модуль «Основная статистика/таблицы»** (рис.2) обеспечивает, на основе данных таблицы, решение задач.

Рис. 2 Задачи «Основная статистика..



Задание – войти в каждый модуль и записать в отчет наименование блока и кол-во переменных, используемых при решении задач модуля.

## Определение среднего значения переменной

**Вручную рассчитать средний рост друзей из таблицы** *1tFriendFIO.sta.*

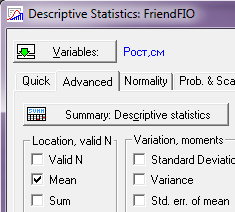
Образец *=(156+212+174+162+156)/5=172cм.*

Вычисления записать в отчет

**Используя STATISTICA определить** средний рост друзей из таблицы:

1. В меню - *Статистика (Click) / Basic statistic (Click) / Descriptive statistic (wClick*) – откроется панель *Descriptive statistic,* рис.3.
2. Вверху панелинажать*Variables* и выбрать *4-Рост,см*.

Рис. 3 Настройка «Основная статистика

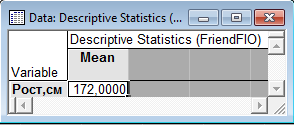


1. выбрать вкладку *Advanced,* на которой видны задания.
2. Выбрать задание на ананализ – *Mean,* cреднее значение;

Для запуска решения нажимаем кнопку *Summary: Descriptive statistic. \**Откроется *Date:..* рис. 4..

**На ней видны**:

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



* *Mean* - вид показателя;
* *Рост ,см* - имя переменной;
* *172* – среднее значение переменной

**Задание** – расчитать среднее значение веса и балла. Все средние значения занести в отчет.

# Вывод результатов анализа

## Виды вывода результатов анализа

**Подготовка** - загрузить пакет и таблицу*1tFriendFIO.sta.*

**Вывод в индивидуальное окно**

**Загрузить менеджер вывода** *Файл / Менеджер вывода / Менеджер вывода.*

**Настроить вывод:**

1. **Поставить метку**- Индивидуальные окна;
2. **Все прочие метки** *снять*;
3. **Нажать *Ок.***

**Выполнить анализ –** определить средний рост друзей**.**

\*Полученную таблицу (индивидуальное окно) сохранить под именем *1twFriendFIO* (По правилам файловой системы).

**Задание -с**делать копию экрана и сохранить в отчете *z2FIO.doc*.

**Вывод в книгу результатов**

**Загрузить менеджер вывода** *Файл / Менеджер вывода / Менеджер вывода.*

**Настроить вывод:**

1. **Поставить метки:** *Книга; Она книга; Поместить результаты в книгу автоматически*;
2. **Все прочие метки** *снять*;
3. **Нажать** *Ок****.***

**Выполнить анализ –**определить средний рост друзей**.**

**Задание.** Полученную книгу сохранить под именем *1tbFriendFIO*

**Выполнить анализ –**определить средние значения всех числовых характеристик друзей**.**

Сделать копию экрана с книгой и сохранить в отчете *z2FIO.doc*.

**Вывод в отчет результатов** (текстовый файл в формате rtf).

**Загрузить менеджер вывода** *Файл / Менеджер вывода / Менеджер вывода.*

**Настроить вывод:**

1. **Поставить метки:** Индивидуальные окна; Также послать окно отчета; Один отчет.
2. **Все прочие метки** *снять*;
3. **Нажать *Ок.***

**Выполнить анализ –**определить средний рост друзей. \*Откроется панель report1.

**Полученный** отчет сохранить под именем *1trFriendFIO*

**Выполнить анализ –**определить средние значения всех числовых характеристик друзей**.**

**Задание.** Сделать копию экрана с книгой и сохранить в отчете *z2FIO.doc*.

## Работа с «Окнами»

**Загрузка сохраненных окон - в**ыполняется по правилам работы в файловой системе.

* **Загрузить ранее** созданное окно.
* **Сохранить его** под иным именем.

## Работа с «Книгой»

**Книга** является структурированным хранилищем результатов анализа.

**Загрузка сохранённой книги - в**ыполняется по правилам работы в файловой системе.

* **Загрузить** ранее созданную книгу.
* **Сохранить ее** под иным именем.
* **Выполнить анализы** друзей с отдельным определением всех числовых характеристик. \*Результат анализа выводить в новую книгу.

**Корректировка книги**.

**В структуре** книги (contents):

* Изменить наименования компонентов книги с учетом их фактического содержания;
* Поменять компоненты местами – для размещения в других местах структуры.

**Задание.** Сделать копию экрана с книгой и сохранить в отчете *z2FIO.doc*.

**Извлечение компоненты из книги**.

* Перетащить в рабочую зону используя *Click*;
* Перетащить в рабочую зону используя *rClick*;

**Ввод окна в книгу**.

* Активировать вводимое окно*- Click* на нем;
* В главном меню выбрать – *Добавить в книгу*.

## Работа с «Отчетом»

**Отчет** является текстовым файлом в формате rtf.

**Загрузка сохранённого отчета - в**ыполняется по правилам работы в файловой системе.

* **Загрузить** ранее созданный отчет.
* **Сохранить его** под иным именем.
* **Выполнить анализы** друзей с отдельным определением всех числовых характеристик. \*Результат анализа выводить в новый отчет.

**Корректировка отчета**.

**В структуре** отчета (contents):

* Изменить наименования компонентов с учетом их фактического содержания;
* Поменять компоненты местами – для размещения в одну строку.

**Задание.** Сделать копию экрана с отчета и сохранить в отчете *z2FIO.doc*.

**В тексте отчета:**

* Изменить наименования компонентов отчета с учетом их фактического содержания;
* Поменять компоненты местами – для размещения в столбик.
* Между компонентами вставит по 3 пустых строки
* Над компонентами написать их наименования.

**Задание.** Сделать копию экрана с отчетом и сохранить в отчете *z2FIO.doc*.

**Извлечение компоненты из отчета**.

* Перетащить в рабочую зону используя *Click*;
* Перетащить в рабочую зону используя *rClick*;

**Ввод окна в отчет**.

* Активировать вводимое окно*- Click* на нем;
* В главном меню выбрать – *Добавить в отчет*.

**Задание.** Разместить в рабочей зоне таблицу друзей, сохраненное окно, созданную книгу и отчет – сделать копию экрана и сохранить в отчете *z2FIO.doc*.

**Задание** – оформленный файл отчета *z2FIO.rtf* отправить преподавателю на почту.

# Подготовка к анализу описательной статистики

## Работа в таблице данных

**Создать файл** *z3FIO.rtf* (отчет) с первой записью «z3FIO Описательная статистика»

**Формализованное** представление исходных данных описательной статистики имеет вид таблицы.

**Создание новой** таблицы.

1. **Закрыть открытую** таблицу – *Файл / закрыть*.
2. **Создать новую** таблицу *– Файл / New*.
3. **Вкладка** – Крупноформатная таблица.
4. **Будет предложен** размер таблицы по 10 переменных и регистров (строк), которые можно корректировать. \*Принять *переменных – 2, регистров -3.*
5. **Размещение** - *Как автономное окно*.
6. **Выполнить создание** - *ОК. \*Откроется таблица с заданными параметрами.*
7. Сохранить таблицу – *Файл / сохранить*. \*Будет предложено создать имя файлу. \*Принять *3tFIO.*
8. **Корректировать имена** переменных (колонок) и их типов (Типы полей см. в разделе 2.3).

* cделать *wClick* на имени переменной *Var1* – откроется панель. \*На ней выполнить:
* Name = *Предмет*.
* Display format = *General;*
* Type = *Text*;
* Length = *22 (символа)*;
* подтвердить корректировку *–Ок*;
* Растянуть колонку для визуализации наименования;
* Выполнить корректировку *Var2: Кол-во пропущенных; Integer*.

1. **Заполнить таблицу**

* В 1ю строку в клетки ввести «ПрограмСтатКомплексы» и кол-во пропущенных занятий.
* Во 2ю строку ввести данные предмета с максимальными пропусками.

1. **Сохранить таблицу** – *Файл / сохранить*.

**Корректировка структуры** таблицы – выполняется для загруженной таблицы.

1. **Добавить переменную**

* На имени последней переменной в таблице нажать *rClick* и не отпуская кнопки перейти на открывшуюся панель.
* На панели выбрать *Add Variables*, ввести характеристики переменной - *Ок*.
* Новая переменная появится в таблице.
* Удаление, перемещение переменной делается через эту-же панель.

1. **Добавить строку**

* В левой колонке нижней строки в таблице нажать *rClick* и не отпуская кнопки перейти на открывшуюся панель.
* На панели выбрать *Add Cases*, ввести характеристики добавляемых строк .
* Новая строка появится в таблице.
* Удаление, перемещение строк делается через эту-же панель.

**Задание в** отчет вставить копию экрана с таблицей *3tFIO.*

## Формирование таблицы данных

**В файле отчета** по образцу создать таблицу *3tFriendFio* с полями образца и дополнительно добавить: пол; Рост, м; Балл (при поступлении в ДонНТУ), № зачетной книжки.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Таблица - *3tFriendFio*   |  |  |  |  | | --- | --- | --- | --- | | № | ФИО | ДР | Вес, кг | | 1 | Красилин Ю.А. | 12.12.1902 | 64 | | 2 | Шмелев В.А. | 04.05.1995 | 105 | | 3 | Фоменко Р.И. | 31.01.1998 | 59 | | 4 | Корина У.Н. | 25.08.2001 | 76 | |  | Ганжа СС. | 17.02.2002 | 64 | |  |  |  |  | |  |  | *Date* |  | | \*Типы полей см. в разделе 2.3.  В таблицу ввести данные о 5ти своих друзьях |

**В пакете STATISTICA** создать таблицу *3tFriendFio* подобную созданной в отчете.

**Задание -**  в отчет вставить копию экрана с таблицей *3tFriendFio.*

## Выбор параметров в STATISTICA

**Описательная статистика** позволяет определить значительного количества различных показателей.

**В *STATISTICA* зявка** на определение параметров описательной статистикиуказывается на панели *Статистика*

**Подготовка**

**Загрузить таблицу** *3tFriendFio* с исходными данными.

**Указать место** вывода результатов анализа - (см.раздел 4.1, Вывод в индивидуальное окно).

**Загрузить панель** *Статистика / (Основная статистика/Таблицы) / Descriptive statistic (wClick);*

**Указать переменную** *variables = Балл*, *Ок*.

**Перейти на вкладку** *Advanced,* где приведены три группы параметров.

**Определить среднее значение**

**В группе *Location, valid N*** **(**м**еры центральной тенденции) выбрать;**

* *Valid N* – количества значений переменной;
* *Mean* – cреднего значения;
* Все прочие отметки во всех группах снять;
* Для выполнения нажать *Summary Descriptive ststistic*

\*Результат будет выведен в Окно – ознакомиться,

**Задание** – Окно скопировать в отчет и закрыть окно.

**Определить все** м**еры центральной тенденции**

* Поставить *Все отметки*;
* Сделать копию экрана.
* Для выполнения нажать *Summary Descriptive ststistic.*
* \*Результат будет выведен в Окно – ознакомиться.

**Задание**

* Окно скопировать в отчет и закрыть окно. Из кармана выложить в отчет копию экрана.

Закрыть STATISTICA. В Word загрузить *z3FIO.rtf* и выполнить его редактирование – сделать отчет удобным для чтения.

# Вычисление параметров описательной статистики

## Ручное вычисление параметров

**Задание** - используя текстовый редактор WORD в ранее созданном файле *z3FIO.rtf* (отчет) добавить запись «Ручное вычисление параметров описательной статистики»

**Формализация исходных данных**

**Вербальное** представление данных.

* известны данные о доходе и образовании каждого из 12 человек в группе : 2н, 2с, 3с, 3н, 3т, 4б, 4б, 5т, 6м, 10б, 13м, 19м. Где образование: н-1 начальное; с-2 среднее; т-3техникум; б-4 бакалавр; м-5 магистр;
* добавить данные о 5-ти своих друзьях с образованием от начального до магистра.

**Задание** - данные ввести в отчет.

**Табличное** представление данных.

**Задание** – в отчете:

* по правилам раздела 5.2 создать таблицу *3tMoneyFio.* Для ввода данных о доходе (для значение дохода создать поле *ДоходFIO* .
* В таблицу ввести даные (для 17 человек).

**Ручной анализ** – выполнить для всех данных таблицы (пример выполнен для 12 человек).

**Задание -** рассчитать и ввести в отчет

**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 арифметитеское 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 (серединное значение показателя);
* *Mode=(*3 + 3 + 3) =*3 –* мода (наиболее часто встречающееся значение);

**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*;*

## Автоматизированное вычисление параметров

**Подготовка к анализу**

* в среде STATISTICA открыть отчет *z3FIO.rtf* и добавить в него запись «Автоматизированное вычисление параметров описательной статистики».
* Направить вывод результатов анализа в индивидуальное окно.

**Создать таблицу** *3tMoneyFio* в пакете STATISTICA*.*

**Ввести в таблицу** данные о доходе и образовании.

**Задание** – копию экрана с таблицей вставить в отчет.

**Определение** центральных характеристик статистики**.**

1. **Открыть панель** описательной статистики *Статистика / (Основная статистика/Таблицы) / Descriptive statistic (wClick);*
2. **Выбрать переменную** *– Variable = 1-* ДоходFIO
3. **На вкладке** *Advanced* указать центральные характеристики статистики (принятые в ручном анализе);
4. **Сформировать** результат- нажать *Summary: Descriptive statistic.\** Результат будет вывелен в *Date:* -индивидуальное окно с результатами анализа, рис.1.

Рис. 1 Результат обработки статистики



**Задание** - используя средства пакета STATISTICA результат анализа добавить в файл отчета.

**Определение** вариации данных**.**

1. **На вкладке** *Advanced* указать паказатели вариации данных;

**Задание** – копию экрана вставить в отчет.

1. **Сформировать** результат- нажать *Summary: Descriptive statistic.*

**Задание** - используя средства пакета STATISTICA результат анализа добавить в файл отчета.

**Определение** характеристик разброса данных**.**

1. **На вкладке** *Advanced* указать характеристик разброса данных;
2. **Сформировать** результат- нажать *Summary: Descriptive statistic.*

**Задание** - используя средства пакета STATISTICA результат анализа добавить в файл отчета.

**Задание** – оформленный файл отчета *z3FIO.rtf* отправить преподавателю на почту.

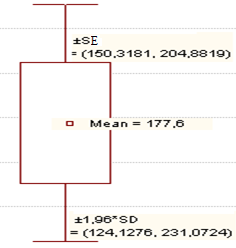
# Визуализация описательных статистик

## Виды гистограмм

**Задание -** создать файл*z4FIO.rtf* (отчет) с первой записью «z4FIO Визуализация описательной статистики»

**Графики коробок** (ящик с усами) используются для визуализации описательных статистик. Рис. 1

Рис. 1 Ящик с усами



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

* Среднее значение *Mean* = 177,6 (из принятой экспериментальной выборки);
* Ошибка среднего ±*SE* (150,3. .204,9) (границы в которых может оказаться среднее значение генеральной совокупности) ;
* Стандартное отклонение среднего ±1,96 \**SD=* (124,1 . . 231,1). (Границы определенны значением 1,96 t-критерия Стюдента в которых с вероятностью 0,95 окажется любое значение из экспериментальной выборки).

**Подготовка к исследлванию.**

**В файловой системе** выполнить копию таблицы *3tMoneyFio* с именем *3tMoneyFio2.*

**В STATISTICA** загрузить таблицу *3tMoneyFio2.*

**Сделать копию** колонки Доход под именем Доход 2

**Построение графика коробок** (ящик с усами) :

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.

1. **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

±*SE* (4,7. .7,7);

* Стандартное отклонение

среднего

±1,96 \**SE=* (0,93 . . 11,4).

## 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.

1. **Объемные**

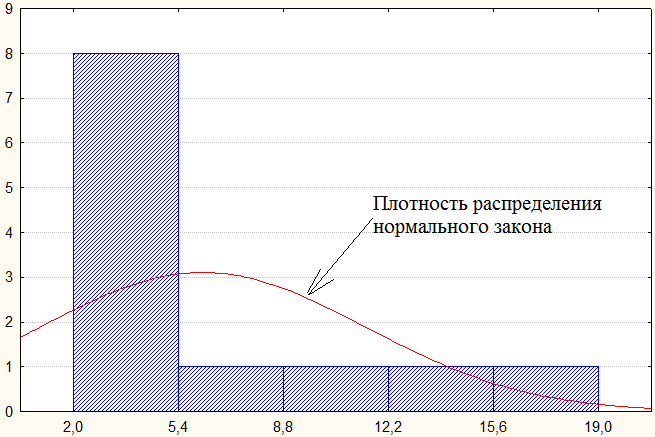
* 2d графики Гистограммы - на основе одной переменной;
* 3d последовательные графики;
* 3d XYZ графики

1. **Другие.**

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

1. Загрузить панель настройки – *Графики / Гистограммы*;
2. На вклаке *Быстрый* указать:

Рис. 2 Гистограмма



* переменную – *Переменные / 1- Доход*;
* Тип графика – *Regular*
* Кол-во интервалов - Categories =10. \*Размер интервала определяется автоматически *d=Range /* Categories;
* Запустить выполне-ние – *ОК*
* \*Будет получена гистограмма, см. рис. 2.

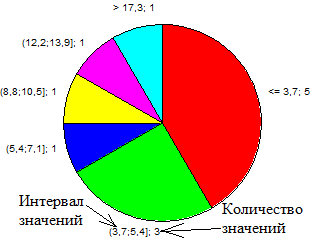
# Построение диаграмм

## 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 фактичестких 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*)

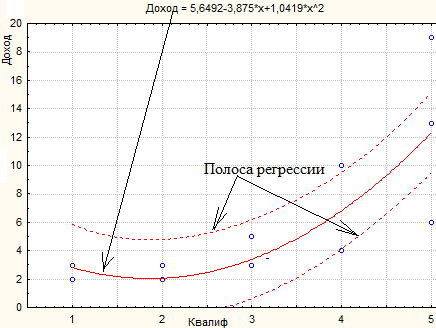
* Легенда шрифта *Text and Value*, вывод на график интервала и значения в нем;
* Тип *3D* – построение объемной диаграммы.

## 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 (*х*) и для отклика (*у*).
* 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.

# Критерий Пирсона

## 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 *Chy-Square*, у которого уровень значимости *р* характеризует вероятность ее правильности.

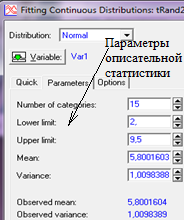
**Гипотеза принимается** если расчетный *p* превышает граничный *рZ*.( *р >* *рZ*).

**В промышленной статистике** принимается *рZ* *= 0,05*. В этом случае при выполнении условия *р >* *рZ* 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*)).
3. **Загрузить панель** настройки проверки - *Статистика / Настройка распределения / Parameters*, *(Fittig Continuous Distributions)* рис. 1.

Рис. 1 Настройка проверки



1. **На панели** *Distribution**Fitting* вбрать нормальный закон – *Normal*.
2. **На панели** *Fittig Continuous Distributions,* вкладка *Parameters,* рис 1.

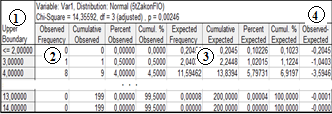
* Выбрать *Variable* = *Var1*.
* Ознакомиться с параметрами описательной статистики:

1. **Запустить выполнение** проверки – *Sammary*. \*Будет сформирована таблица результатов.

## Таблица результатов критерия Пирсона

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

Рис. 2 Таблица результатов Chy-Square



**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 Freguency (*1на переменная*).*

**В шапке таблице** приведены характеристики *Chy-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 *(р=0,00246)<(рZ=0,05)* в критерии Пирсона говорит о недопустимом различии теорети-ческой 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.

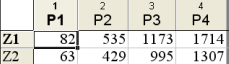
* **Загрузить панель** - *Статистика / Настройка распределения / Parameters, (Fittig Continuous Distributions)* рис. 1.
* **В окне *Distribution*** - выбрать другой закон распределения.
* **Запустить выполнение** проверки – *Sammary*. \*Будет сформирована таблица результатов.
* **Выполнить анализ** уровня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

# Сравнение выборок

## Проверка однородности выборок

**Пусть имеются выборки** полученные из различных совокупностей. Например, примеси *Р1..Р4* в одной марке инструментальной стали выплав-ленной на двух заводах *Z1, Z2*, приведенные в таблице 5*tStill*, рис.4.

Рис. 4 Количество примесей в стали



**Необходимо проверить** гипотезу , что исходные совокупности распределены одинаково.

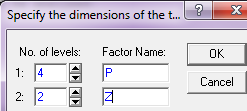
**Проверка гипотезы с** использованием критерия Пирсона выполняется в модуле *Статистика / (Дополнительные линейные/нелинейные модели) / Регистрационно - линейный анализ таблиц частоты.*

**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.

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

Рис. 4 Количество примесей в стали



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

**Нажать *Ок, Ок*** и на вкладке *Advansed* нажать кнопку *Test allmarginal….*

**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 *Chy-Square*;
* Уровень значимости *р = 0,3088*.

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

## Выявление причин низкого уровня значимости

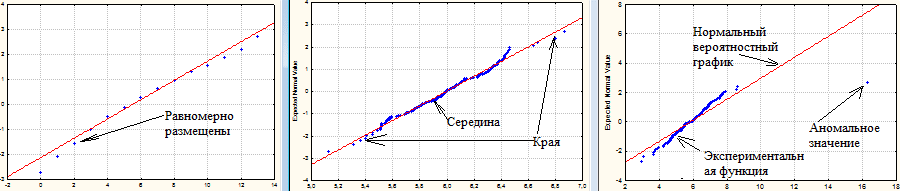
**Нормальный вероятностный** график позволяетвизуально оценить соответствие экспериментальных данных требованиям нормального закона распределения.

**Построение графика** средствами *STATISTICA* требует:

1. З**агрузить 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.

1. Другой закон 2 Нормальный закон 3. Аномальное значение

Рис. 5 Использование нормального еероятностного графика



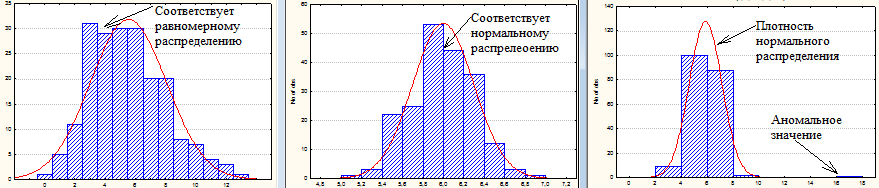
**На рисунке приведены** характерные случаи поведения экспериментальной функции.

1. Особые свойства- соответствует другому 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.

1. Другой закон 2 Нормальный закон 3. Аномальное значение

Рис. 6 Варианты экспериментальных данных



**Внешний вид** гистограммы позволяет сделать выводо соответствии данных нормальному распределению.

# Связи между переменными

## Корреляционный анализ

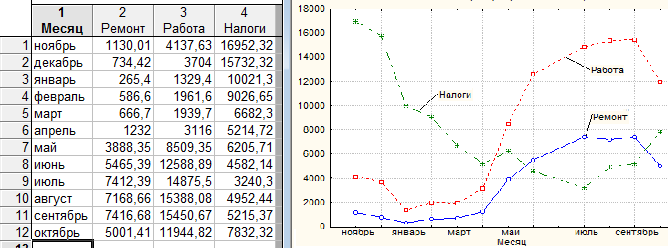
**Корреляция** определяет степень с которой значения двух переменных пропорциональны друг другу.

**Коэффициент корелляции** *r* – показатель оценивающий теснову 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 6t Ремонт имеющей аргумент (месяц) и отклики (Ремонт, Работа, Налоги) рис.1 . Необходимо оценить связь между этими признаками.

а. Табличное представление в. Графическое представление

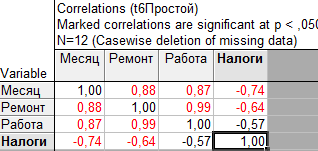
Рис. 1 Исходные данные к корреляционному анализу



**Корелляционный анализ** в студии STATISTICA выполняют с испрользованием модуля *Статистика / (Основная статисттика/Таблицы) / Correlation matrices.*

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

Рис. 2 Коэффициенты корреляции



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

## Простой (однофакторный) регрессионный анализ

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

**Однофакторная (парная) регрессия** имеет один фактор и представлена выражением

*Y=f(x) + ε,*

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

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

**Парная линейная** регрессия представлена выражением

*Y= b0 + b1 x +ε,*

где *b0* , *b1* – называются коэффициентами регрессии.

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

**На вкладке** *Advansed* выполняем:

* Variables: *depended = Ремонт* (отклик); *independent= Работа* (фактор) - *Ок*.
* *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** *Advansed* нажать *Summary:Regression results*.

**Откроется** таблица с результатами анализа в которой:

* колонка *Вeta = 0,9935* содержит коэффициент корелляции;
* колонка *В* содержит коэффициенты регрессии *b0*,*=-588(верхняя) ; b1=0,506(нижняя)*;
* колонка p-level =0,0077-уровеннь значимости.

**При *p-level* < 0,05** выборка считается однородной и уравнение регрессии принимается к использованию (краснаый).

# Оценка модели

## 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:**

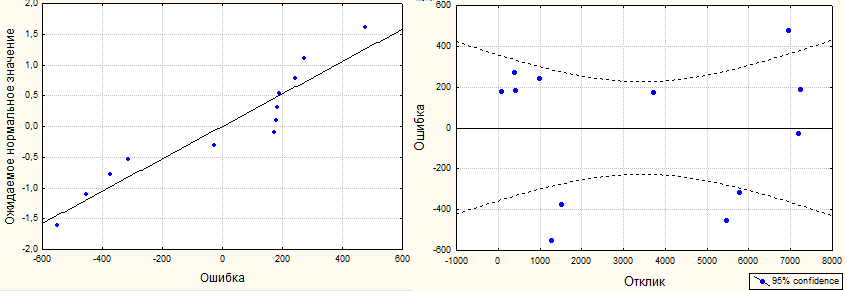
* выбрать вкладку *Residuals / assum*….;
* нажать кнопку *Perform residual analyis. \**Загрузится панель *Residual Analyis.*

**Выполнение анализа** остатков на панели *Residual Analyis***:**

**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 относистельно прямой и нет резко выделяющихся остатков.

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

## 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 – *x1, x2, x3*; обратные к ним -1/ *x1, 1/x2, 1/x3*; логарифмы – *ln(x1), ln(x2), ln(x3).*

**Метод пошагового** включения предполагает последова-тельный ввод в модель или исключение факторов в явном или преобразованном виде.

**Рациональность включения фактора** в модель (степень его независимости) оценивается коэффициентом корелляции *R*, который должен возрасти и уровенем значимость *p-level*, который должен быть менее *0,05*.

**Гребневая регрессия** предполагает получение смещенных оценок (*b0* , *b1* ) 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 *λ = 0*, на некоторую величину, например, 0,001.

**На каждом шаге, и**спользуя кореллированые значения показателей модели, определяют коэффициент коррелляции *R*, который должен возрасти. Если рост останавливается – дальнейшее изменение *λ* не целесообразно.

**Наилучшая корреляционная** модель – понятие субъективное. Однако при ее выборе исползуют значение: коэффицинта корелляции *R>0,91*; уровень значимости *p-level <0.05,* остатки должны иметь нормальное распределение, др.

# Диаграммы качества

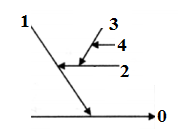
## Анализ причинно – следственных связей

**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** предствалена структурой, имеющей несколько уровней, рис.1.

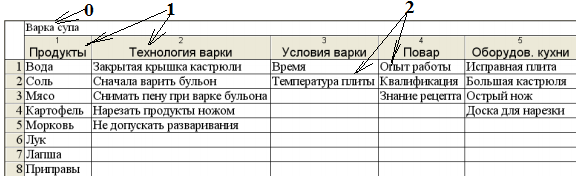
Рис.1. Диаграмма Исикавы



* 0-рещаемая 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 *7t Суп*, рис.2.

Рис.2. Данные о причинах и результатах



**Свойства таблицы** позволяют представвить 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.

## Построение диаграммы Исикавы

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

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

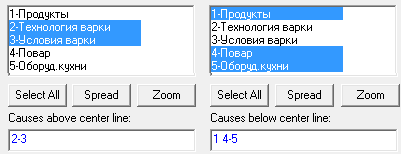


Рис3. Размещение главных причин

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

Рис.4. Диаграмма «Суп»



**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*.

# Диаграмма Парето

## 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.

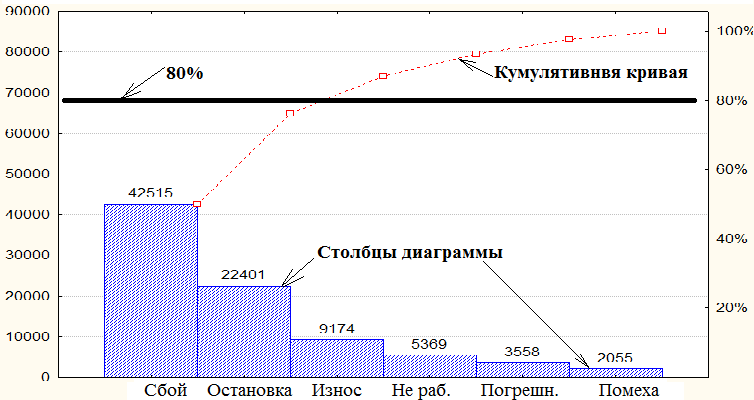
**Например, процесс** (стоимость ремонта оборудования) в зависимости от вида устраняемого дефекта (сбой, износ, погрешность, повреждение цепи и т.д.).

**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 – «Сбои» приводит к потере 42 515 руб.

## Построение диаграммы Парето в STATISTICA

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

Рис.5.Данные к диаграмме Парето



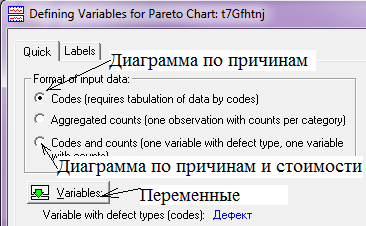
* процесса, который является откликом (потери, руб.);
* факторов влияющих на него (станок, дефект).

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

**Построение диаграммы Парето** в студии STATISTICA выполняют с испрользованием модуля *Статистика / (Индустриальная статистика & Сигма 6)/ Качество диаграммы управления.* \*Загрузится панель *Quflity а chart specification file* .

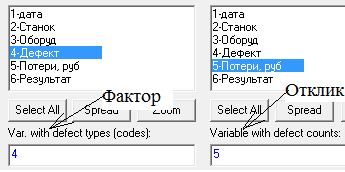
**На вкладке *Quick*** выбрать *Pareto chart analysis. \**Откроется панель *Defining Variables for Pareto Chart,* рис.6. **Панель используется** для указания требований к виду диаграммы:

Рис.6.Требования к диаграмме Парето



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

Рис.6.Выбор переменных



* фактор – Дефект;
* отклик – Потери, руб.

**Запустить** формирование диаграммы – *ОК . \**Диаграммаы имеет вид показанный на рис.4.

# Карты контроля качества

## 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 описываются законом распределения (сохраняется *status quo*).

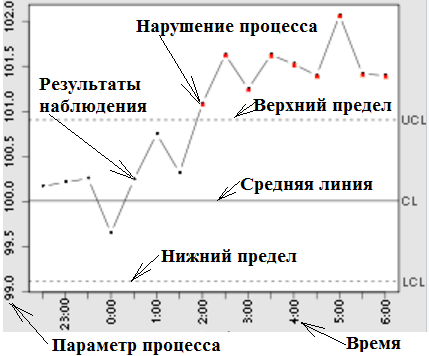
**Однако возможны ситуации** когда появляется неслучайная причина приводящая к нарушению процесса изготовленния и его выходу из состяния статического контроля.

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

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

**Ключевые компоненты** контрольной карты:

Рис.1. Пример 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 *t0C* процесса, рис. 2.

## Контрольная карта индивидуальных значений

**Анализ контрольной карты** в студии STATISTICA выполняют с использованием модуля *Статистика / (Индустриальная статистика & Сигма 6)/ Качество диаграммы управления .* \*Загрузится панель *Quality Conrol Charts*.

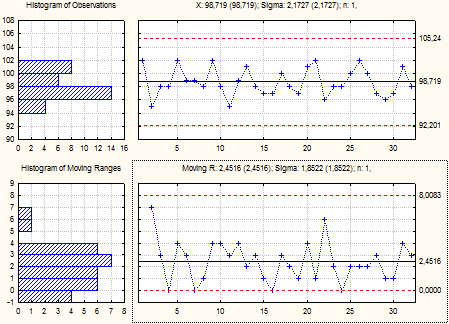
**На вкладке** *Quick*выбрать *Indastryals & moving range*.

**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*- отражает** параметр *t0С* исследуемого процесса и включает :

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

**R – изменение параметра Х** между двуми результатами.

# Развитые карты контроля качества

## 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 студииSTATISTICA.

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 Conrol Charts,* вкладка *Quick*.

1. На вкладке выбрать *X-bar & R chat for variablres.*
2. На открывшейся вкладке нажать *Variables* и выбрать: *Measurements = процессу; Samplу 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 чувчтвительна к возмущениям и позволяет выяить сдвиг настройки процесса в самом начале.

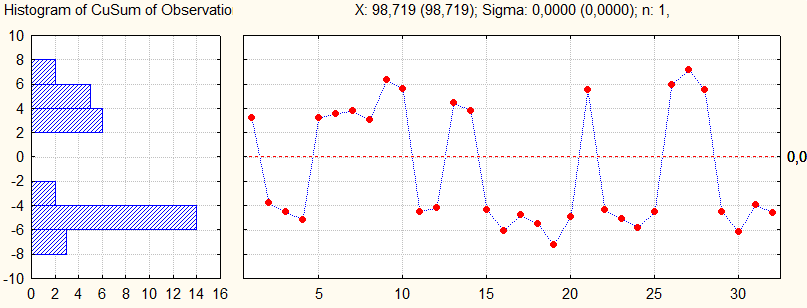
**Анализ контрольной карты** в студии STATISTICA выполняют на панели *Quality Conrol Charts,* вкладка *Variables* .

**На вкладке выбрать** *CuSum*  *chat for individuals* .

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

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

Рис.4. Контрольная карта накопленных сумм

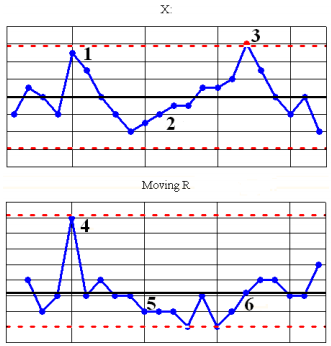


**На карте** 6 сигм деляься на 20 равных частей и оценивается количество попаданий в эти части, рис.4.

## Чтение контрольных карт

**Анализ контрольных карт** выполняется для для выявления факта выхода процесса из-под контроля и определение причин выхода. В рамках анализа рассматривается изменение параметра *Х* и размаха *R* , рис.5.

Рис.5. Контрольная карта



**Для стабильного** процесса вероятность выхода за контрольные границы менене 0,01.

**Выход за границы** говорит о воздействии очобой причины

**Признаки причин**.

1. За границами 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 границами резулирования находятся Х точки, а R точки лежат в границах (рис.5, точки 3, 6). Причина - воздействие внешнего фактора.
3. Дрейф – 7 и более поднимающихся или ниспадающитх точек точек (рис.5, точки 2, 5). Причина – рост температуры, износ оборудования и др.
4. За грааницами X и R точки. Причина наличие как обычных так и особых причин нарушающих процесс.

**Особые случаи.**

1. Серия точек – последовательные точки лежат по одну стороу средней линии. Например - более 15ти тгочек 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.