

доц. Скобова Н.В.

***РУКОВОДСТВО ПО
ИСПОЛЬЗОВАНИЮ
Statistics for windows
версия 8.0***

<http://belsp.by>

1. **Создание нового файла (документа):**
на линейке нажимаем File → New (рис.1).

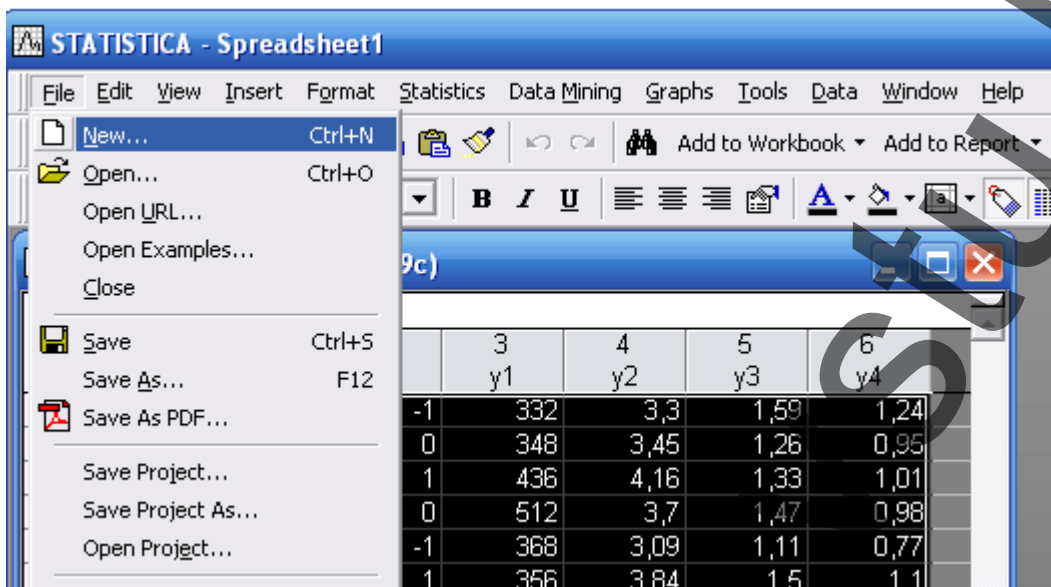


Рисунок 1

В появившемся окне (рисунок 2) в ячейке **Number of variable** - указывают количество столбиков, в ячейке **Number of case** - количество строк; в окне **Display format** выбрать вид переменных - *number* (числовой), после чего станет активной ячейка **Decimal places** - в ней указать количество знаков после запятой.

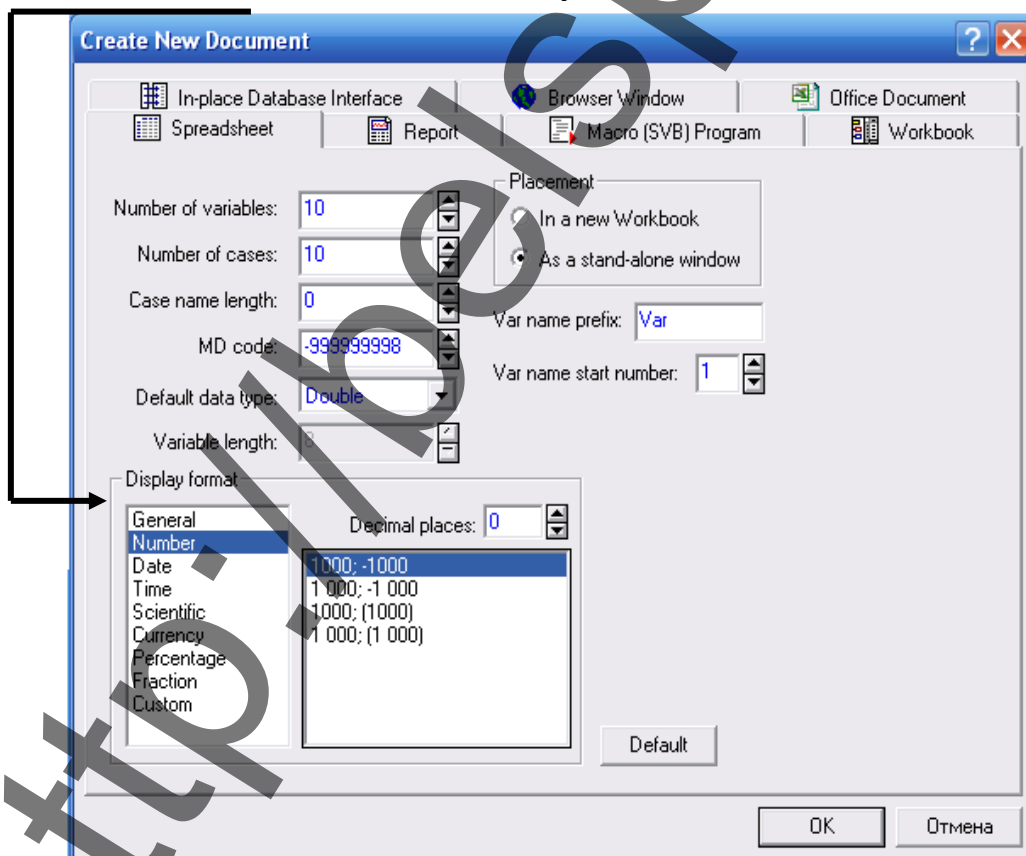


Рисунок 2

1. Расчет среднестатистических характеристик

Для расчета среднего, дисперсии, среднего квадратического отклонения на линейке экрана в меню **Statistics** выбираем строку

Basic Statistics/Tables и далее (рис. 4) строку **Descriptive statistics**.

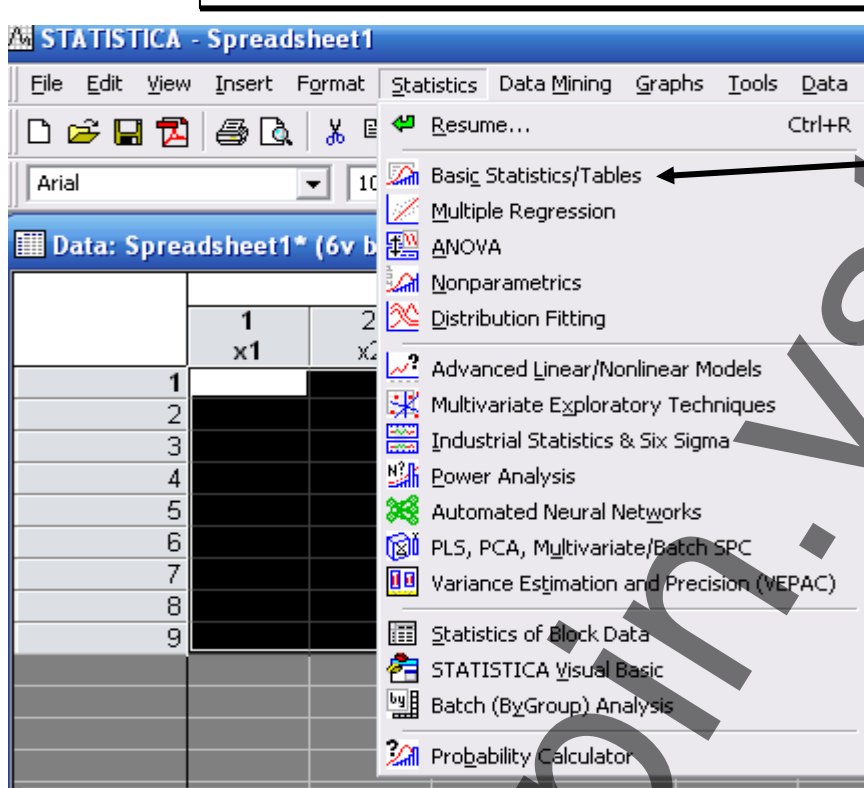


Рисунок 3

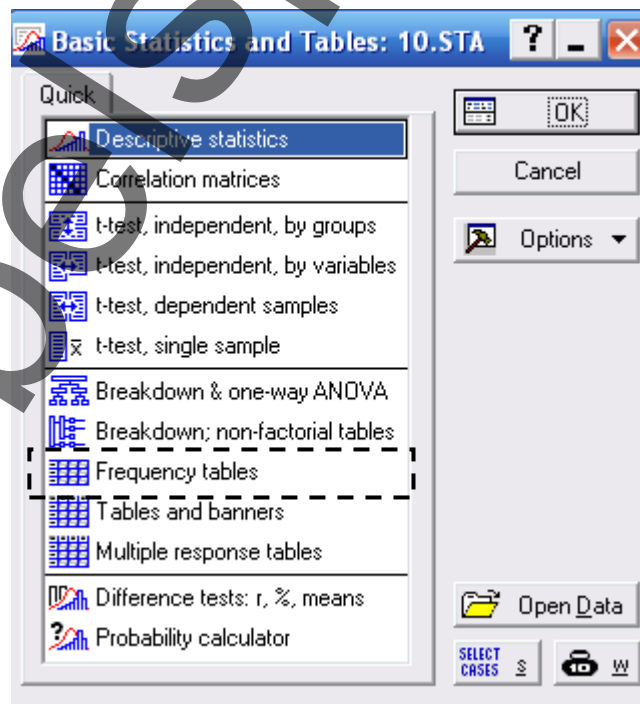


Рисунок 4

В появившемся окне (рисунок 5) в ячейке **Variable** задают переменные, по которым проводят расчет. Далее в закладке **Advanced** проставляют галочки напротив тех параметров, которые требуется рассчитать (таблица 1).

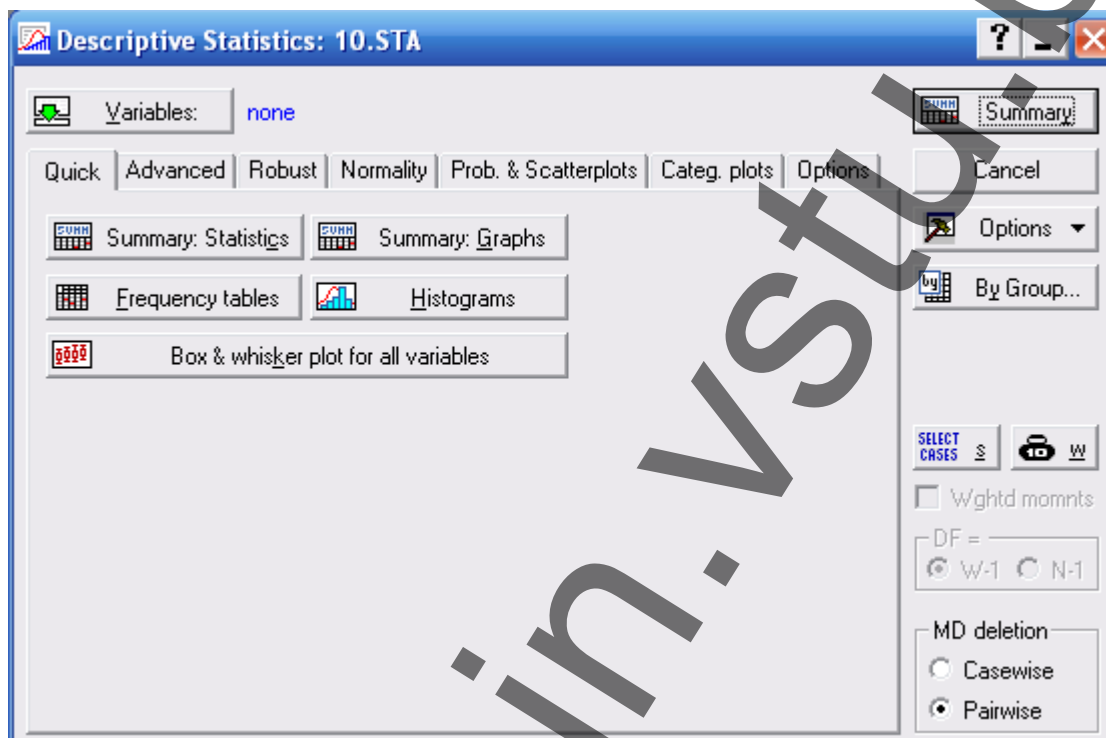


Рисунок 5

Таблица 1 - Среднестатистические характеристики

Valid N	Число элементов совокупности
Mean	Среднее
Sum	Сумма
Median	Медиана
Standart Deviation	Среднее квадратическое отклонение
Variance	Дисперсия
Standart Error of mean	Средняя квадратическая ошибка
95% confidence limits of mean	Уровень доверительной вероятности
Minimum and Maximum	Минимум и максимум совокупности
Lower and upper quartiles	Нижний и верхний квартиль

После выбора параметров для расчета нажимаем кнопку **Summary**

Для построения **гистограммы** в окне **Descriptive Statistics** (рис.5) выбираем переменные для построения, а затем нажимаем кнопку **Histograms**.

Для построения **Box-диаграммы** в окне **Descriptive Statistics** (рис.5) выбираем переменные для построения, а затем нажимаем кнопку **Box & whisker plot all variable**.

3. Построение частотной таблицы

Для разбиения всей выборки на классы (частотный анализ) в меню **Statistics** выбираем «**Basic Statistics**» (рисунок 3) и далее оператор **Frequency Tables** (рис.4 выделено пунктиром). В появившемся окне (рисунок 6) задаемся переменными в ячейке **Variable**, после чего в закладке **Advanced** активируем строку **Step size** и указываем в ней величину интервала, с которым разбивалась выборка на классы. В результате получаем окно (рис.7), в котором в первом столбике указаны границы интервалов, на которые разбита выборка, в столбце **Count** - количество переменных, попавших в указанный класс, **Cumulative Count** - сумма переменных, попавших во все предыдущие классы, **Percent** - процент переменных, попавших в каждый класс.

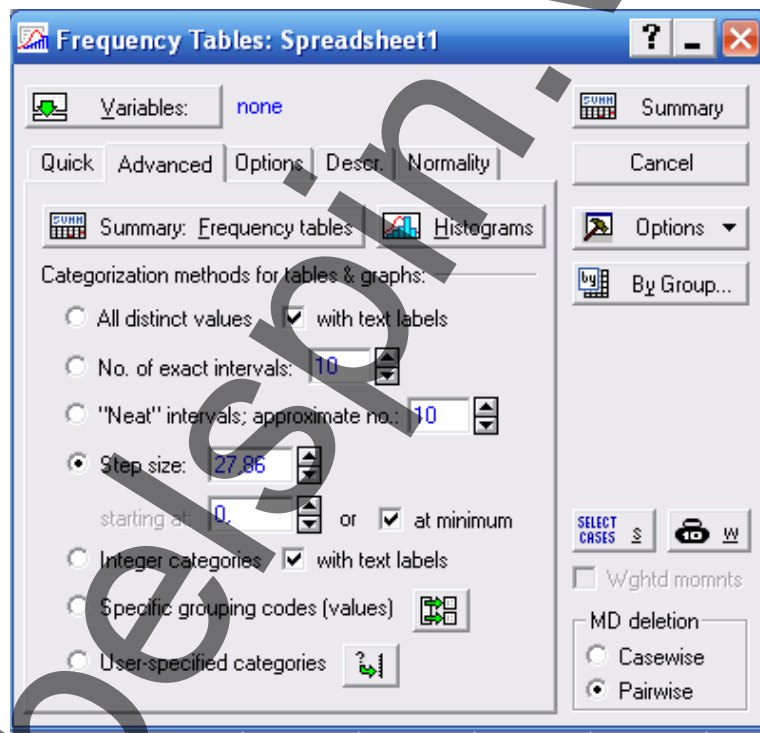


Рисунок 6

		Frequency table: P (Spreadsheet1)			
From	To	Count	Cumulative Count	Percent	Cumulative Percent
12,50000	<=x<13,10000	2	2	22,22222	22,2222
13,10000	<=x<13,70000	4	6	44,44444	66,6667
13,70000	<=x<14,30000	1	7	11,11111	77,7778
14,30000	<=x<14,90000	2	9	22,22222	100,0000
14,90000	<=x<15,50000	0	9	0,00000	100,0000
Missing		0	9	0,00000	100,0000

Рисунок 7

4. Определение закона распределения случайной величины

Для определения закона распределения случайной величины необходимо в меню **Statistics** (рис.3), выбрать раздел «**Distribution fitting**». В появившемся окне (рис.8) в рамке **Continuous Distribution** дважды нажать левой клавишей мышки на строку **Normal**.

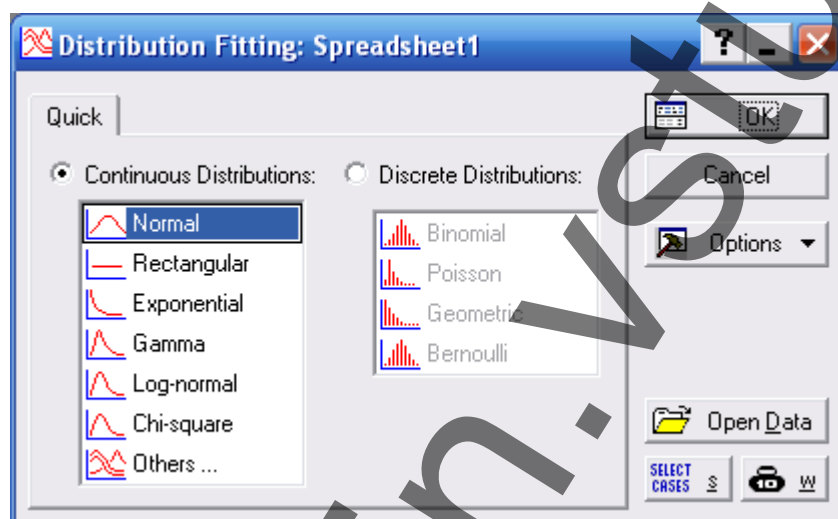


Рисунок 8

В новом окне (рис.9) нажимаем **Variable** и выбираем анализируемые переменные. Далее в закладке **Parameters** в ячейке **Number of categories** указываем количество классов, на которые разбита выборка. Затем необходимо поочередно выбирать в ячейке **Distribution** закон распределения случайной величины (нормальный экспоненциальный, гамма и т.д.) и нажимая кнопку **Summary** получаем расчетное значение критерия Пирсона **Chi-Square** (рис.10 выделено в рамочке), число степеней свободы **df** и уровень значимости критерия $p=$. Наиболее подходящим является тот закон, при котором значение **Chi-Square** не будет превышать критического значения, выбранного по таблице в приложении 2.

Для построения гистограммы выбранного закона распределения в окне **Fitting Continuous** (рисунок 9) в закладке **Quick** выбираем кнопку **Plot of observed and expected distribution** (рисунок 11).

Для расчета критерия Колмагорова-Смирнова в окне **Fitting Continuous** (рис.9) в закладке **Option** активировать строку **Yes (categorized)** в ячейке **Kolmogorov-Smirnov test**.

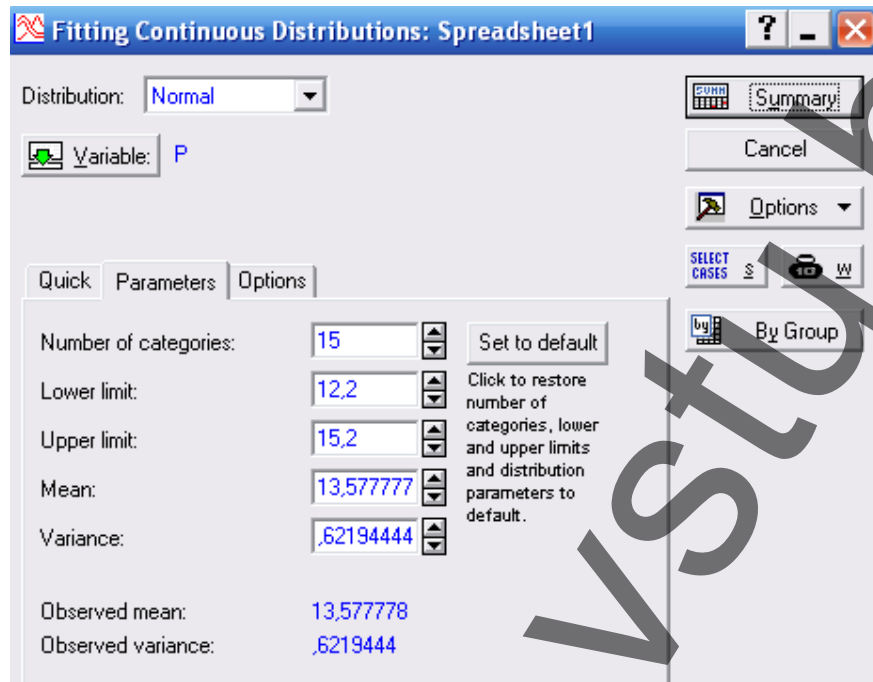


Рисунок 9

Upper Boundary	Observed Frequency	Cumulative Observed	Percent Observed	Cumul. % Observed	Expected Frequency	Cumulative Expected	Percent Expected	Cumul. % Expected	O - E
<= 820,00000	2	2	5,00000	5,0000	0,74459	0,74459	1,86148	1,8615	0,25541
860,00000	5	7	12,50000	17,5000	5,08869	5,83328	12,72173	14,5832	-1,7832
900,00000	14	21	35,00000	52,5000	13,75635	19,58963	34,39086	48,9741	-3,4741
940,00000	12	33	30,00000	82,5000	14,09354	33,68317	35,23384	84,2079	-2,7079
< Infinity	7	40	17,50000	100,0000	6,31683	40,00000	15,79208	100,0000	3,68317

Рисунок 10

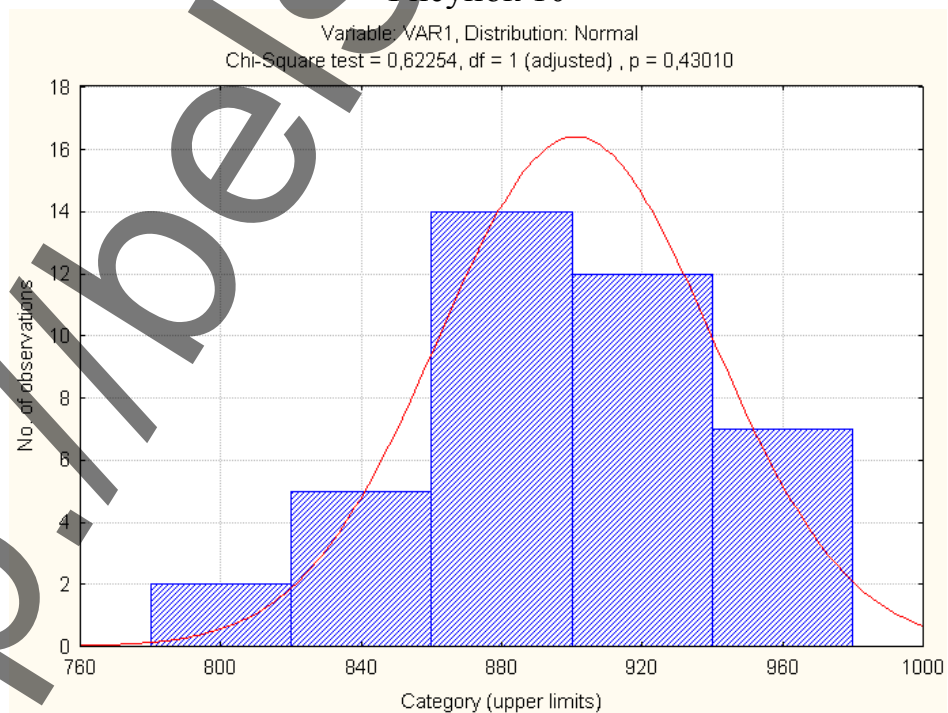


Рисунок 11

5. Корреляционный анализ

Парная корреляция

Для проведения корреляционного анализа необходимо в меню **Statistics** (рис.3), выбрать раздел «**Basic Statistics**». В появившемся окне (рис.4) выбрать раздел **Correlation matrices** и нажать «ОК». В появившемся окне (рис.12) нажимаем кнопку **Two lists (rect. matrix)** и выбираем, какие переменные в корреляционной матрице будут в столбце **First variable list**, а какие - в столбце **Second variable list** и нажимаем «ОК». После этого в закладке **Option** (рис.12) активируем строку **Display r, p-level, and N's** для вывода на экран уровня значимости рассчитанных параметров, после нажимаем **Summary**. Результаты расчета представлены на рисунке 13.

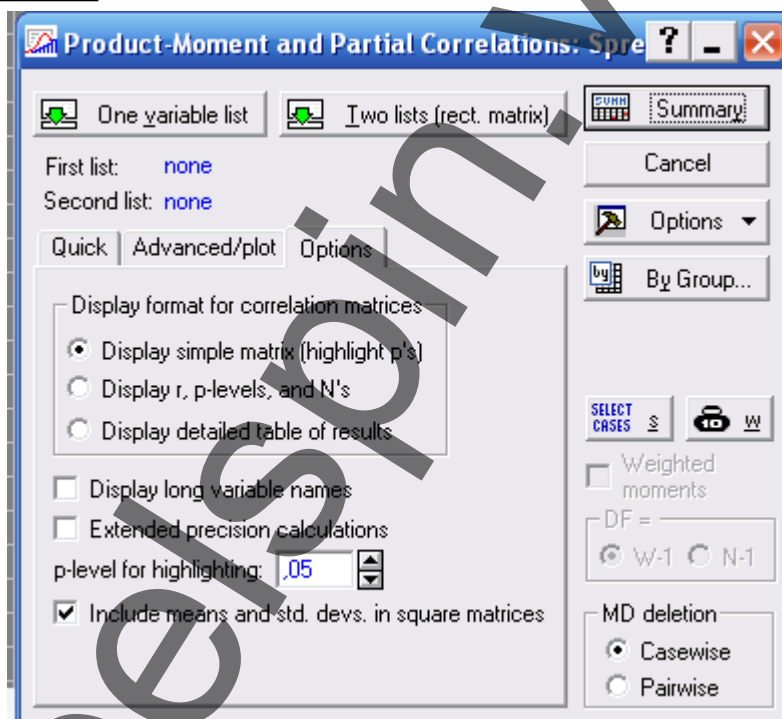


Рисунок 12

Correlations (11111.STA)	
Marked correlations are significant at p < ,05000	
N=4 (Casewise deletion of missing data)	
Variable	VAR2
VAR1	-,6074
	p=,393

Значение коэффициента корреляции

Уровень значимости

Рисунок 13

Для построения графического изображения корреляционной взаимосвязи исследуемых параметров необходимо в стартовом окне (рис.12) в закладке **Advanced/plot** нажать кнопку **2D scatterp**.

Множественная корреляция

Для расчета множественного коэффициента корреляции необходимо в меню **Statistics** (рис.3) выбрать раздел **Multiple Regression**. В появившемся окне нажимаем **Variable** и выбираем переменные, между которыми необходимо рассчитать корреляцию, причем в столбце **Dependent var** указать зависимые переменные (выходные), а в столбце **Independent var** – независимые (входные) переменные. Затем нажать «OK». В появившемся окне (рис.14) в строке Multiple R указан рассчитанный множественный коэффициент корреляции.

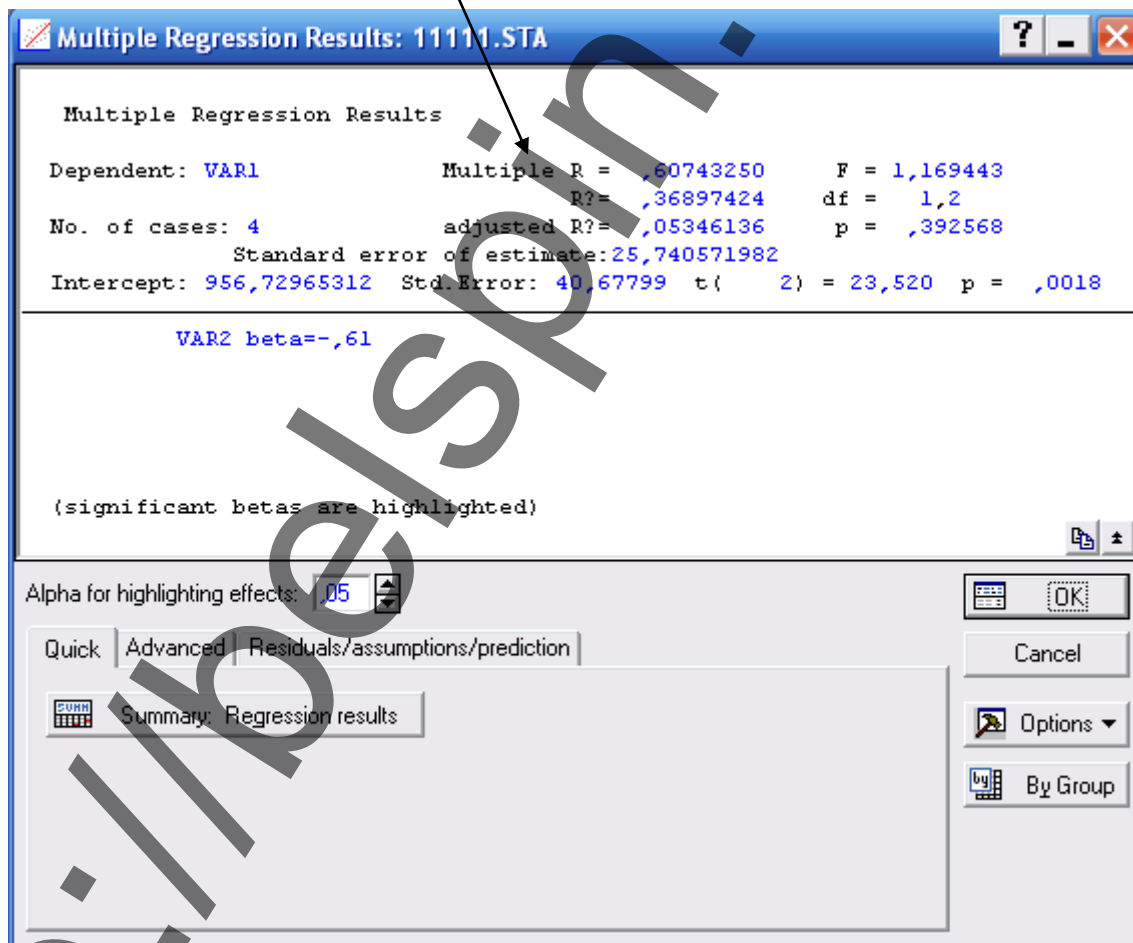


Рисунок 14

6. Регрессионный анализ

Для проведения нелинейного регрессионного анализа необходимо в меню **Statistics** (рис.15) выбрать раздел **Advanced Linear/Nonlinear Model**, а в нем **Nonlinear Estimation**. В появившемся окне выбрать раздел **User-specified regression, custom for loss**. Появится окно (рис.16), в рамке **Estimated function** вводят общий вид регрессионного уравнения, для которого необходимо провести расчет коэффициентов уравнения.

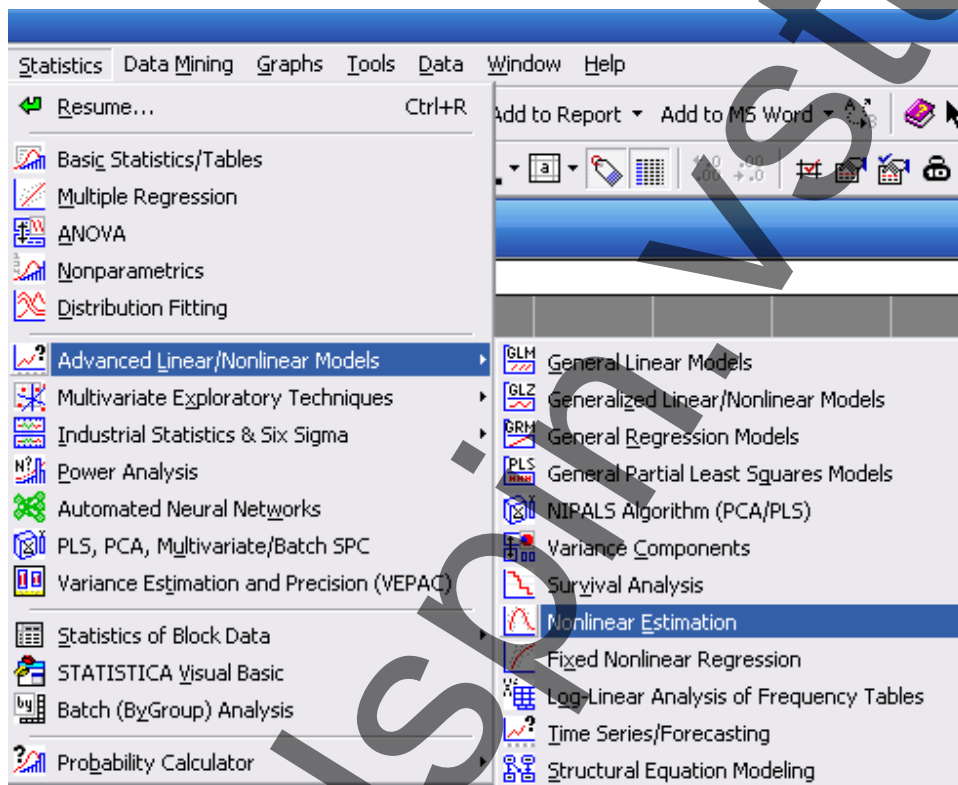


Рисунок 15

Далее 2 раза нажимаем «ОК» и в окне **Model Estimation** в закладке **Advanced** активировать строку (проставить «√») **Asymptotic standart errors** для включения в итоговый отчет расчет ошибок и уровней значимости. Искомые значения коэффициентов модели будут находиться в окне-таблице **Model** (рис.17).

В заголовке таблицы показатель **Variance explained** отражает адекватность модели. В строках таблицы указана следующая информация:

Estimate – численное значение рассчитанных коэффициентов;

Std.Err. -средняя квадратическая ошибка расчета;

t(5) – расчетное значение критерия Стьюдента (цифра в скобках – число степеней свободы);

p-level – уровень значимости рассчитанных коэффициентов.

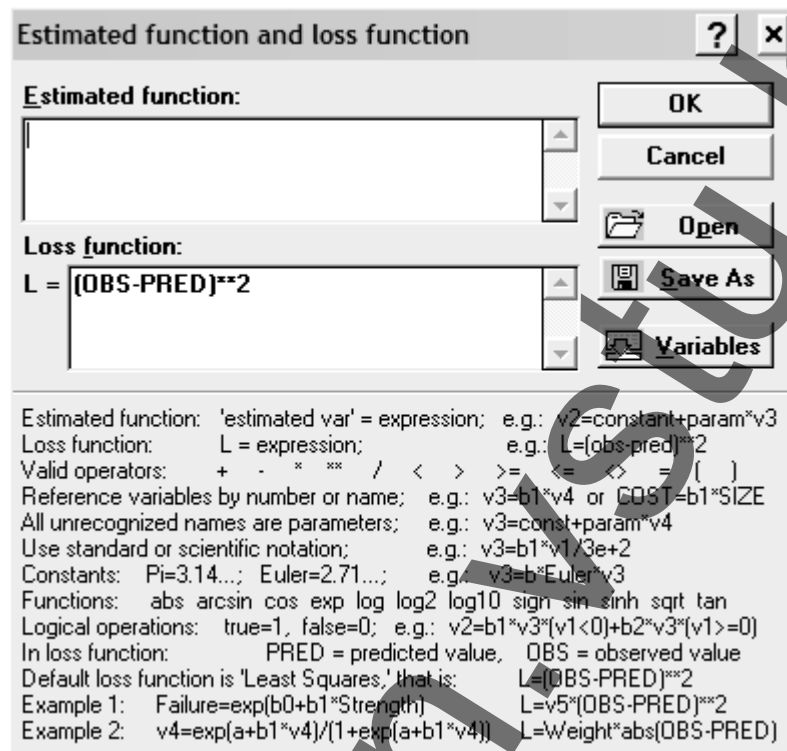


Рисунок 16

Model: H=a0+a1*x1+a2*x2 (Spreadsheet1)			
Dep. var: H Loss: (OBS-PRED)**2			
Final loss: ,002983346 R= ,99547 Variance explained: 99,097%			
N=9	a0	a1	a2
Estimate	5,0933	-0,0950	-0,2134
Std.Err.	0,0074	0,0091	0,0091
t(6)	685,2454	-10,4330	-23,4386
p-level	0,0000	0,0000	0,0000

Рисунок 17

7. Построение графиков

Для построения графика модели используется графический редактор. На линейке экрана выбираем меню **Graphs** (рис.18), в нем раздел **3D XYZ Graphs** и далее **Custom Function Plots**. Появится окно (рис.19), в котором в ячейке *Function* записывают регрессионное уравнение с подстановкой численных значений коэффициентов.

Важно! В уравнении проводят замену: переменную X_1 заменяют на X , X_2 – на Y .

В ячейках X_{min} , X_{max} , Y_{min} , Y_{max} указывают минимальное и максимальное значение, которое могут принимать переменные X_1 (ячейка X)

и X2 (ячейка Y). Далее нажимаем ОК. В результате получаем график (рис.20).

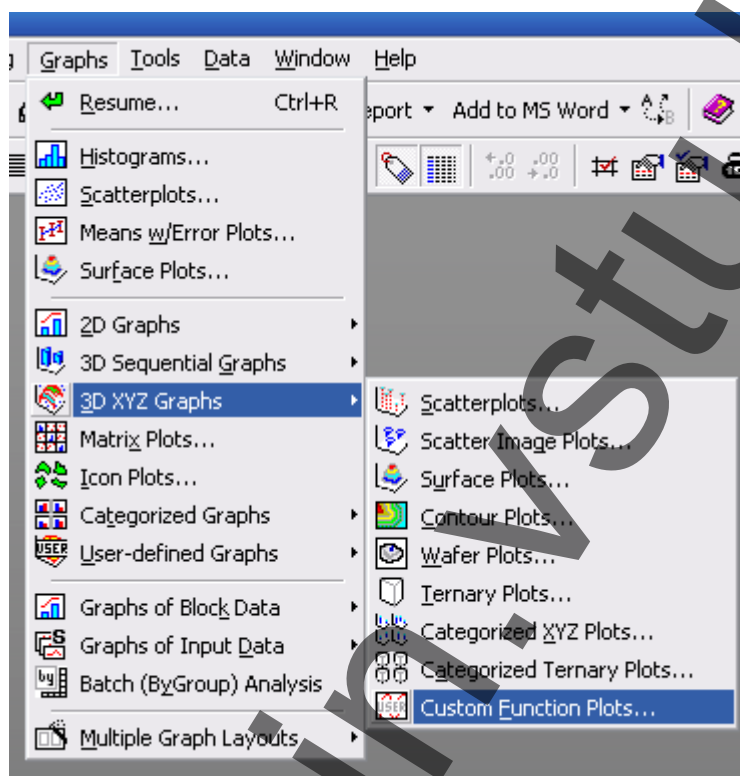


Рисунок 18

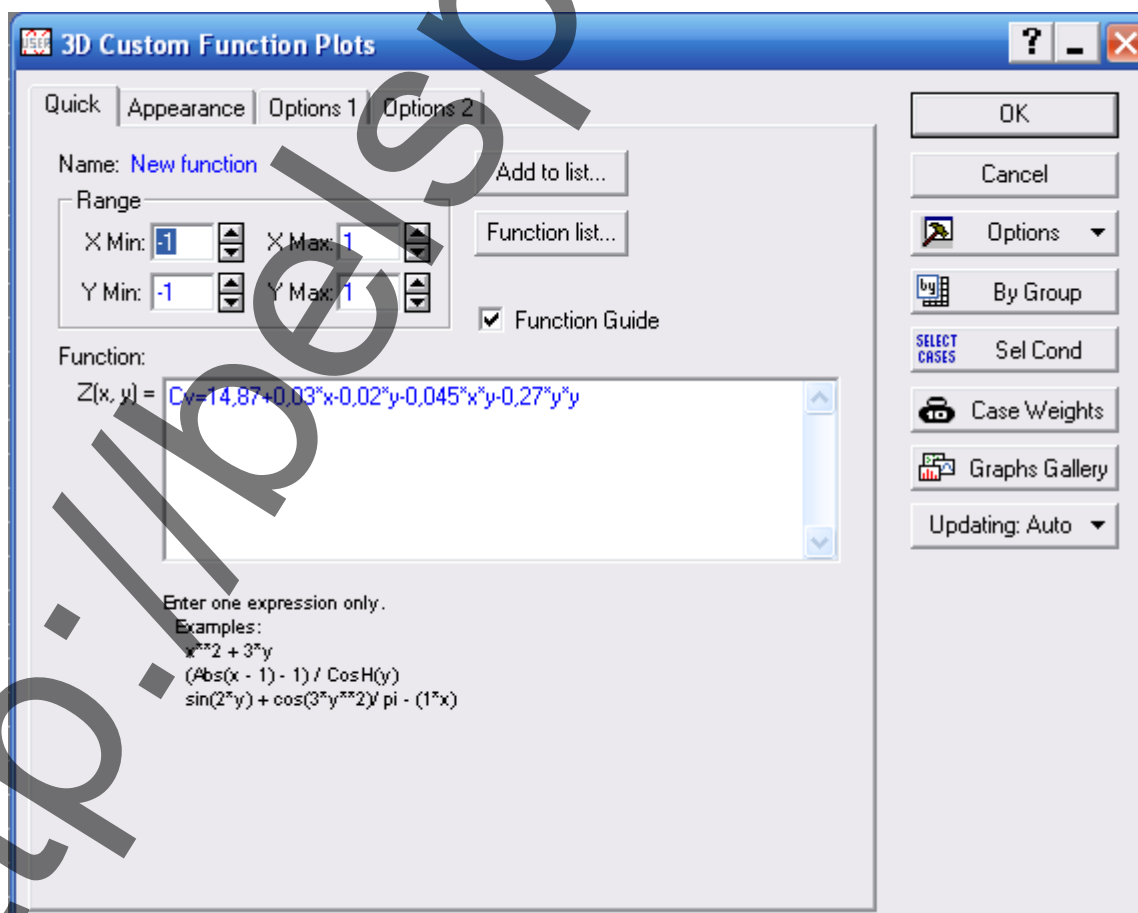


Рисунок 19

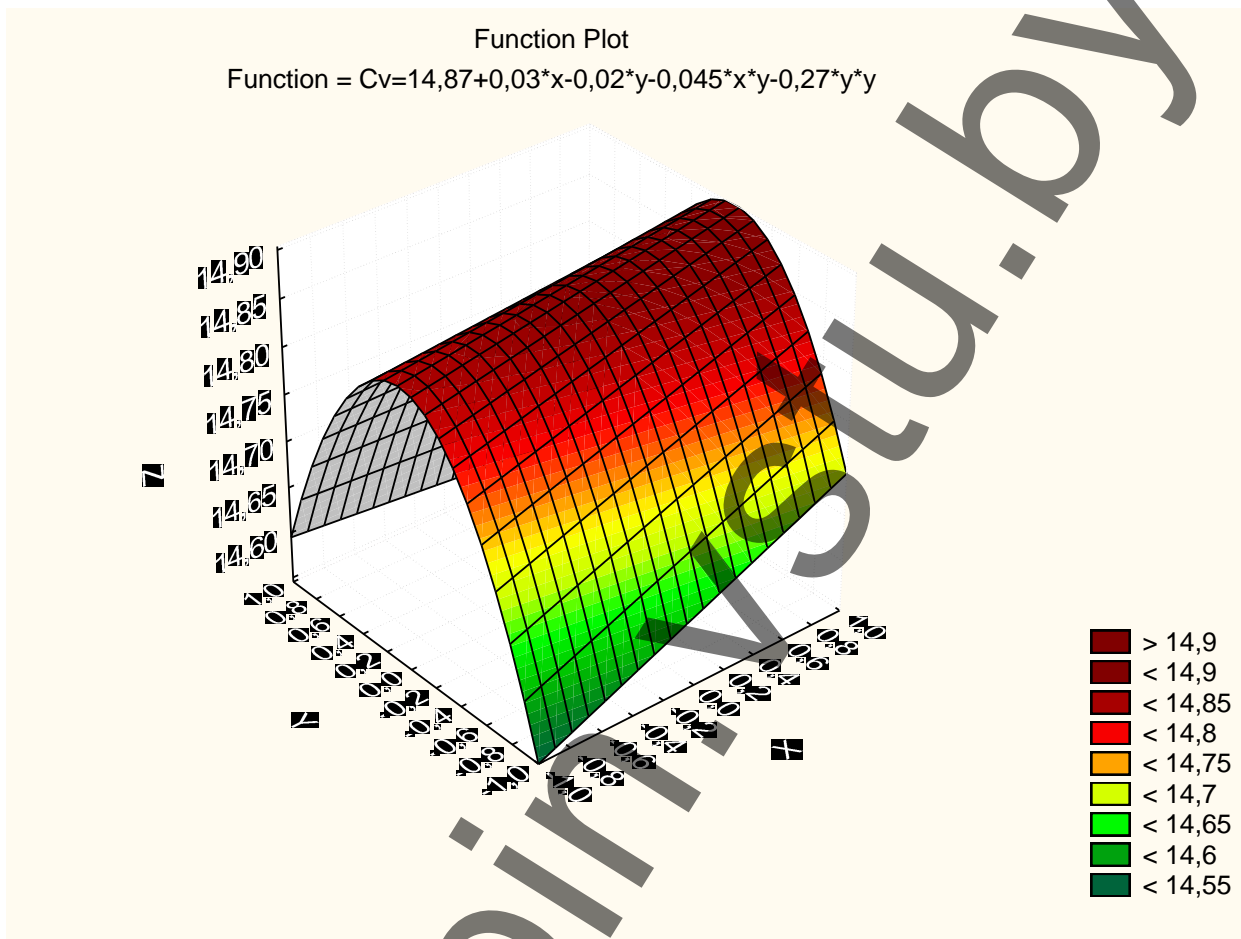


Рисунок 20

Для редактирования внешнего вида графика на поле свободном от графика, но в пределах графика нажимаем правой кнопкой мыши и активируем строку **Graph Properties (All option)**. Появится окно (рис.21) в нем закладки:

Graph Windows внутри **Outside Background** (цвета внутри графика)

Inside Background (цвета снаружи графика)

Graph Layout внутри в ячейке **Graph type** выбирают тип графика:

(Standard – трехмерный,

2D Projection – двухмерный)

Graph Titles/Text внутри в окне записывается название графика

Surface внутри в ячейке **Surface (Projection) contour** поставить «√»

напротив Display (отражение линий на плоскости)

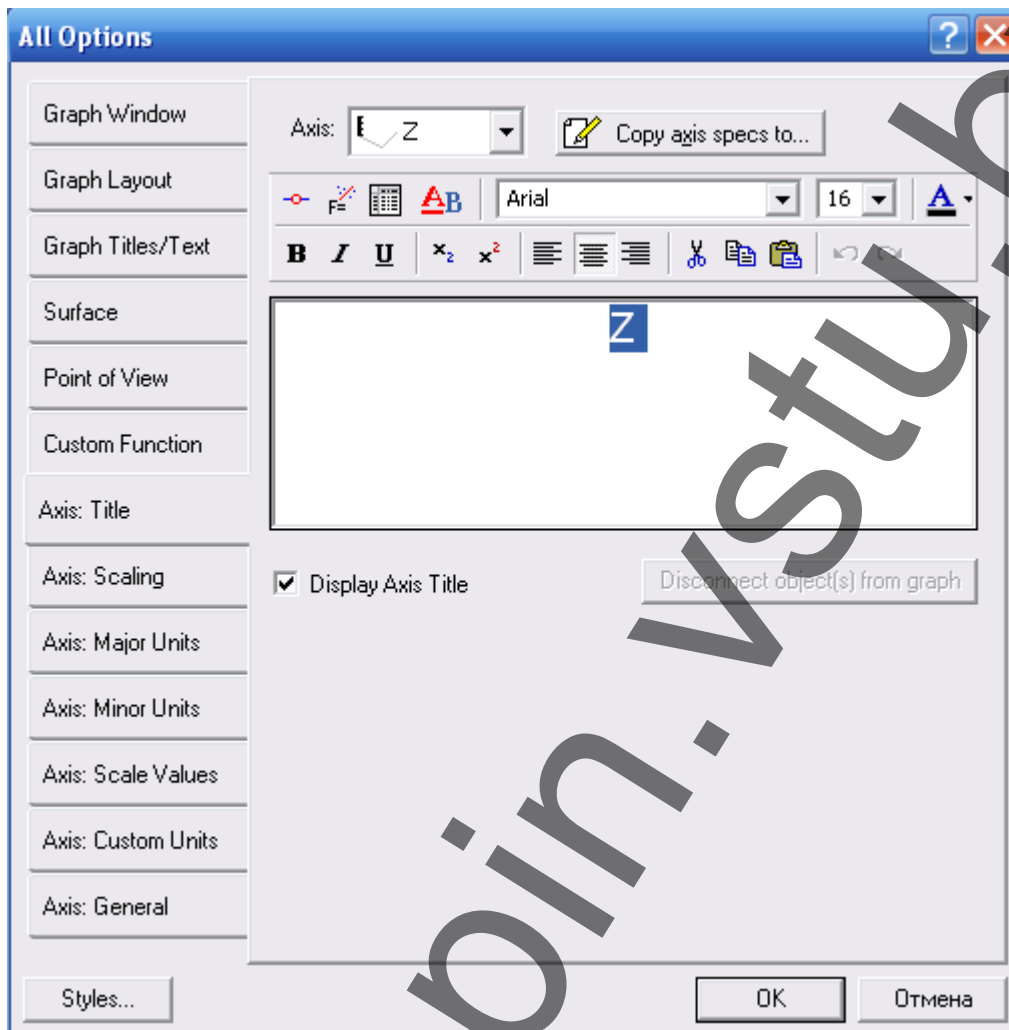


Рисунок 21