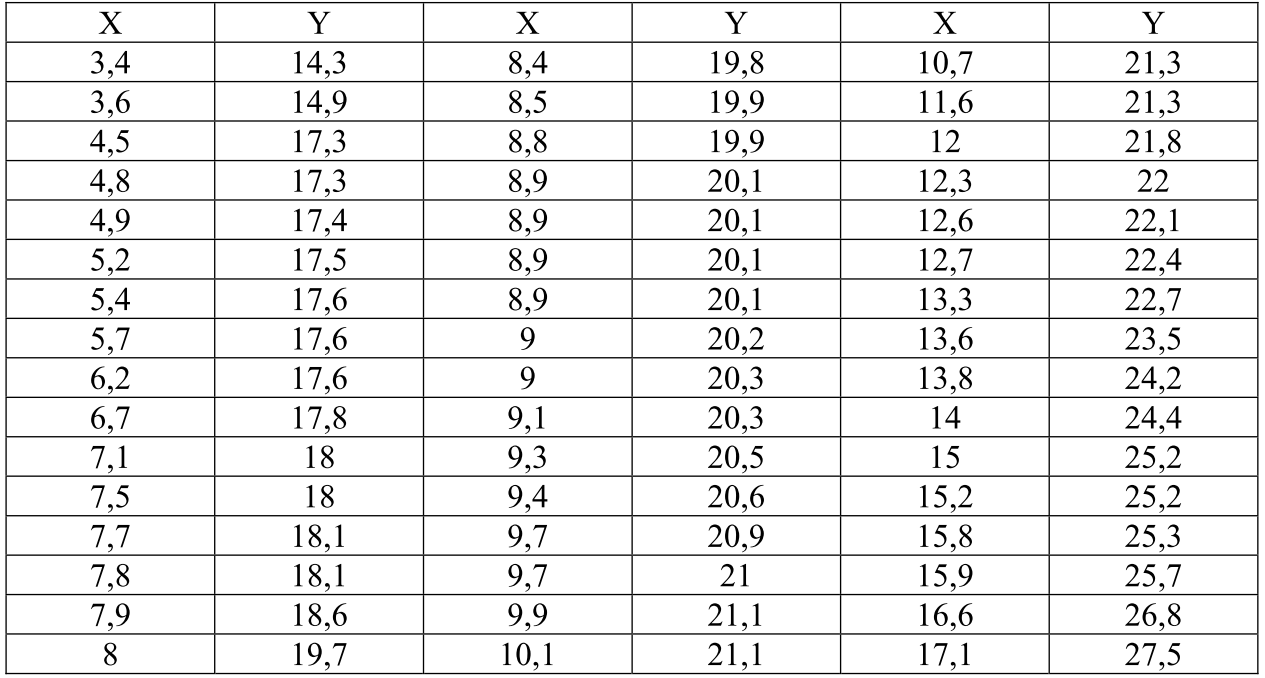
**Лабораторная работа № 4**

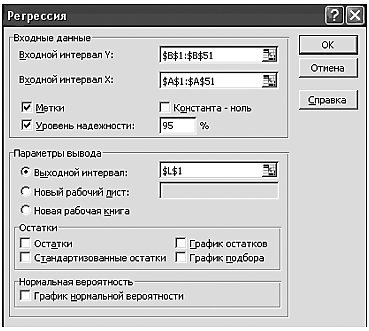
**Регрессионный анализ в среде Excel и STATISTICA**

Регрессионный анализ в среде Excel

Исходные данные представлены в таблице (в табличном редакторе EXCEL данные представлены двумя столбцами):



Открыть модуль **Анализ данных** и выбрать опцию **Регрессия**, после чего щелкнуть мышкой **OK**. В появившемся окне выполнить операции и установки, как показано на рис. 1. Щелкнуть мышкой **OK**. Результат обработки появится в указанном поле (выходной интервал $L$1, табл.1–3).



*Рис.1 – Стартовая панель*

Табл. 1 – Результат обработки



Табл. 2 – Дисперсионный анализ

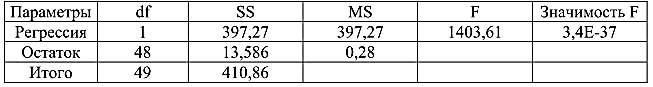
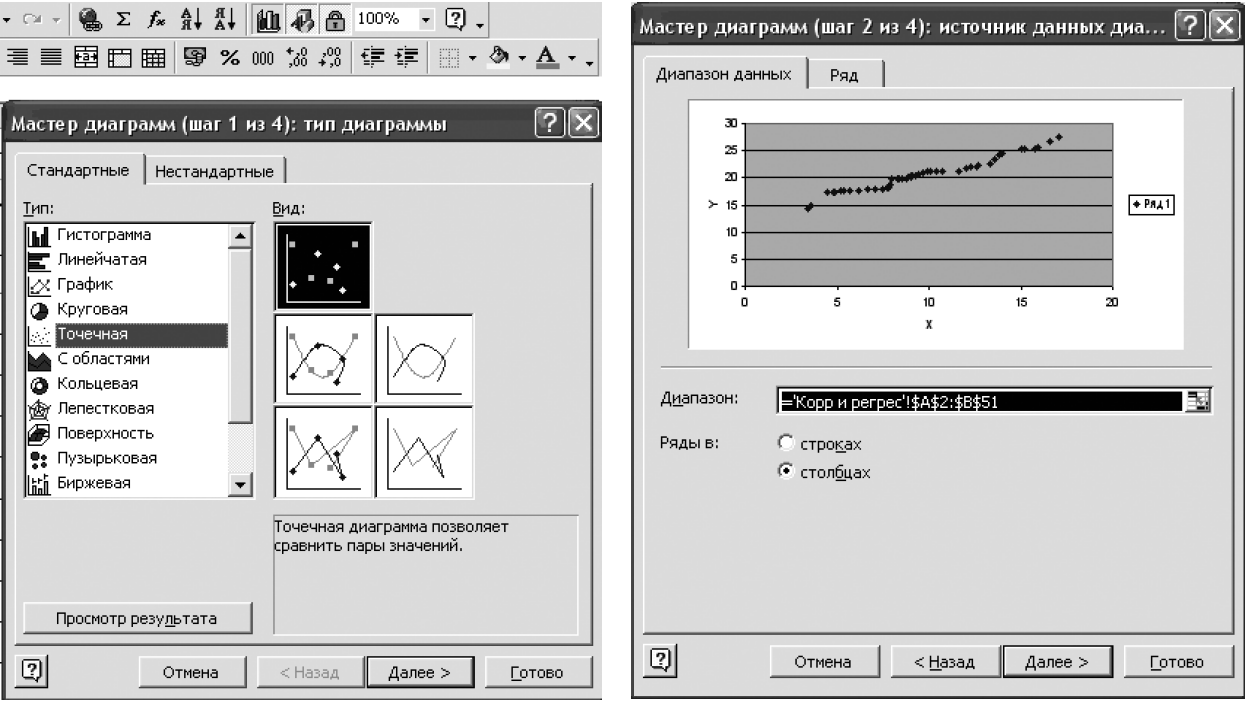


Табл. 3 – Регрессионный анализ



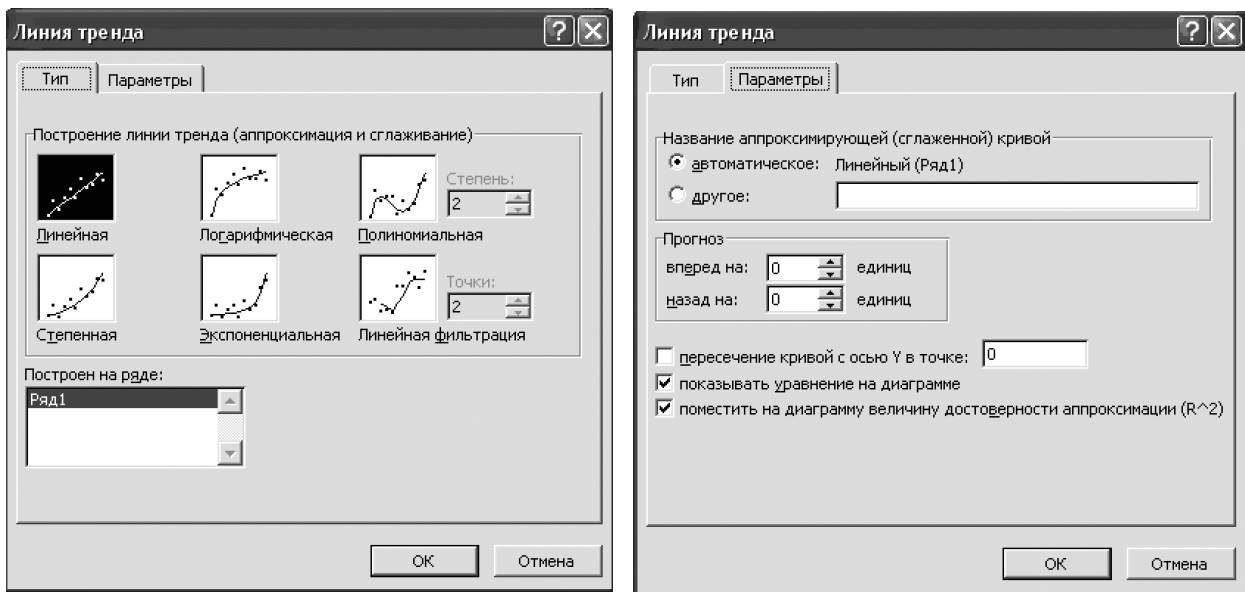
Для данного массива данных получена очень надежная регрессия с высоким коэффициентом корреляции: **Y=12,668+0,8158\*X; r=0,9833276**

Операции регрессии и корреляции можно также выполнить в системе EXCEL, используя модуль **Мастер диаграмм**. Для этого необходимо модуль **Мастер диаграмм** (рис. 2), выбрать **Тип диаграммы Точечная** и щелкнуть по кнопке **Далее**. Выбрать диапазон данных (рис. 3), оформить график и нажать **Готово**:



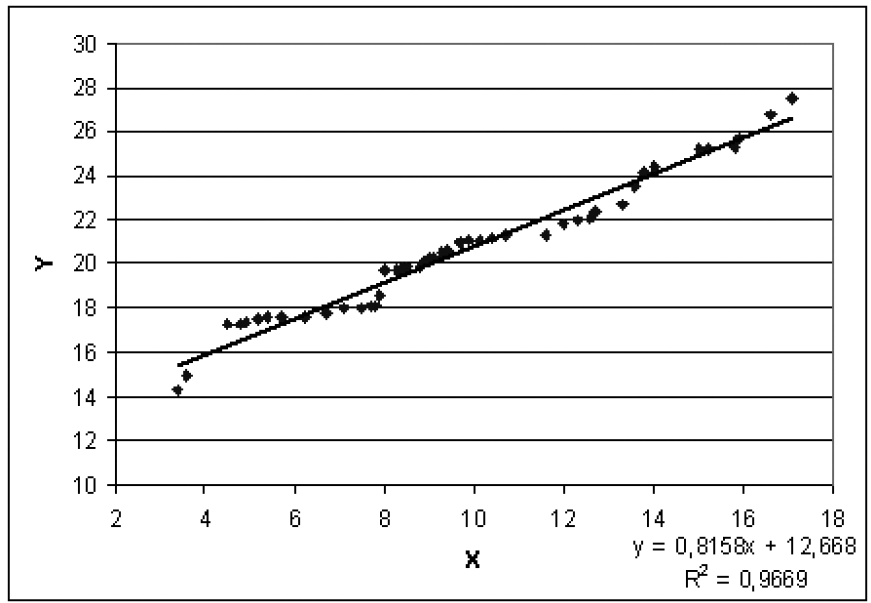
*Рис. 2 – Стартовая панель Рис. 3 – Диапазон данных*

Затем щёлкнуть правой кнопкой по точкам и выбрать опцию **Добавить линию тренда**. Выбрать **Тип – линейная** (рис. 4). В опции **Параметры** выбрать установки, как показано на рис. 5.



*Рис. 4 – Тип аппроксимации Рис. 5 – Параметры кривой*

Отредактированная диаграмма представлена на рис. 6.



*Рис. 6 – Отредактированная диаграмма*

Уравнение регрессии и R2 находятся в правом нижнем углу диаграммы. Как видно, что они такие же, как и при выполнении регрессионного анализа.

**Задание 1.** У 20 взрослых мужчин были измерены высота (длина тела) х (в см) и вес у (в кг):

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| x | 165 | 176 | 175 | 168 | 167 | 172 | 175 | 180 | 179 | 173 |
| у | 56 | 75 | 70 | 61 | 61 | 63 | 72 | 80 | 76 | 68 |
| x | 166 | 178 | 169 | 169 | 170 | 176 | 180 | 169 | 177 | 176 |
| у | 58 | 76 | 60 | 64 | 63 | 71 | 78 | 63 | 75 | 71 |

Определите коэффициент корреляции и регрессии, постройте график и уравнение линейной регрессии.

**Задание 2.** Предполагается, что между количеством настриженной шерсти *у* и живым весом овец *х* имеется зависимость. Для 10 овец были получены следующие данные (в кг):

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| х | 50 | 55 | 60 | 50 | 65 | 60 | 50 | 55 | 50 | 65 |
| у | 4,0 | 4,2 | 4,1 | 4,2 | 4,5 | 4,3 | 4,1 | 4,4 | 4,0 | 4,2 |

Определите коэффициент корреляции и регрессии, постройте график и уравнение линейной регрессии.

**Задание 3.** Были получены следующие данные о длине грудного *х* и брюшного *у* плавника у окуня озера Баторино:

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *x* | 38 | 31 | 36 | 43 | 29 | 33 | 28 | 25 | 36 | 26 | 21 | 30 |
| *у* | 40 | 34 | 38 | 42 | 26 | 33 | 29 | 26 | 36 | 27 | 22 | 32 |
| *x* | 27 | 27 | 28 | 26 | 26 | 25 | 24 | 28 | *28* | 27 | 33 | 27 |
| *у* | 28 | 26 | 32 | 26 | 28 | 27 | 25 | 28 | 30 | 26 | 32 | 27 |
| *x* | 26 | 23 | 22 | 25 | 24 | 29 | 25 | 25 | 30 | 23 | 24 | 32 |
| *у* | 29 | 23 | 24 | 30 | 26 | 30 | 27 | 28 | 32 | 23 | 24 | 32 |
| *x* | 24 | 25 | 30 | 25 | 26 | 30 | 29 | 22 | 29 | 28 | 26 | 28 |
| *у* | 25 | 27 | 33 | 27 | 27 | 32 | 28 | 24 | 31 | 32 | 27 | 30 |
| *x* | 25 | 31 | 25 | 32 | 27 | 31 | 28 | 29 | 26 | 32 | 27 | 31 |
| *у* | 25 | 34 | 26 | 32 | 29 | 30 | 29 | 29 | 26 | 35 | 26 | 33 |
| *x* | 28 | 28 | 26 | 33 | 30 | 27 | 21 | 28 | 26 | 30 | 23 | 27 |
| *у* | 29 | 31 | 29 | 33 | 31 | 31 | 23 | 30 | 27 | 29 | 24 | 28 |

Определите коэффициент корреляции и регрессии, постройте график и уравнение линейной регрессии.

**Задание 4.** Для 10 петушков леггорнов 15-дневного возраста были получены следующие данные о весе их тела *х* (в г) и весе гребня *у*(в мг):

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *х* | 83 | 72 | 69 | 90 | 90 | 95 | 95 | 91 | 75 | 70 |
| *у* | 56 | 42 | 18 | 84 | 56 | 107 | 90 | 68 | 31 | 48 |

Определите коэффициент корреляции и регрессии, постройте график и уравнение линейной регрессии.

**Задание 5.** Путем еженедельного взятия проб с поля было изучено изменение высоты растений сои *у*(в см) с возрастом *х*(в неделях):

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *х* | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| *у* | 5 | 13 | 16 | 23 | 33 | 38 | 40 |

Определите коэффициент корреляции и регрессии, постройте график и уравнение линейной регрессии.

**Задание 6.** Для установления связи между содержанием фосфора в почве *х*и содержанием фосфора в злаковых растениях *у*было проведено 9 анализов со следующими результатами:

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *x* | 1 | 4 | 5 | 9 | 13 | 11 | 23 | 23 | 28 |
| *у* | 64 | 71 | 54 | 81 | 93 | 76 | 77 | 95 | 109 |

Определите коэффициент корреляции и регрессии, постройте график и уравнение линейной регрессии.

**Задание 7.** Было проведено сравнение удоев первой лактации *х* с третьей *у*по 33 коровам холмогорской помеси (в л):

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *x* | 1522 | 239 | 1521 | 2700 | 1789 | 2496 | 1197 | 1105 | 1701 | 2218 | 1790 |
| *у* | 3693 | 4453 | 1446 | 2134 | 2940 | 4353 | 2066 | 2152 | 2396 | 2435 | 3140 |
| *x* | 2964 | 1287 | 1756 | 1406 | 1810 | 1299 | 2609 | 2519 | 1927 | 1655 | 1320 |
| *у* | 4700 | 2113 | 2513 | 3249 | 2553 | 2320 | 4612 | 3201 | 3173 | 3326 | 1639 |
| *x* | 2586 | 1928 | 3884 | 2968 | 2200 | 1753 | 1508 | 1803 | 1811 | 2300 | 1697 |
| *у* | 4562 | 3482 | 4257 | 3465 | 2448 | 3435 | 3747 | 2112 | 3061 | 2985 | 2721 |

Определите коэффициент корреляции и регрессии, постройте график и уравнение линейной регрессии.

**Задание 8.** На белых крысах была показана следующая зависимость между температурой внешней среды *х*(в град.) и количеством поглощенного кислорода *у*(в мл/г веса):

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *x* | 0 | 5 | 10 | 15 | 20 | 25 | 28 | 29 |
| *у* | 3,83 | 3,35 | 2,60 | 2,02 | 1,69 | 1,42 | 1,39 | 1,38 |
| *x* | 30 | 31 | 32 | 33 | 34 | 35 | 40 |  |
| *у* | 1,29 | 1,39 | 1,39 | 1,45 | 1,65 | 1,61 | 2,40 |  |

Определите коэффициент корреляции и регрессии, постройте график и уравнение линейной регрессии.

Регресионный анализ в среде STATISTICA

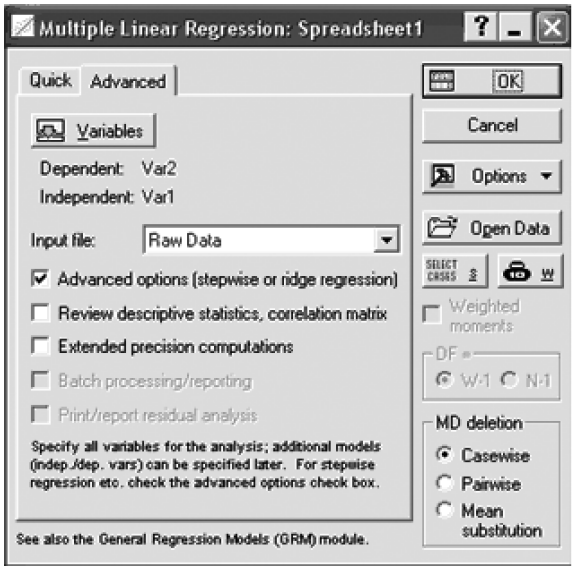
Для начала необходимо ввести исходные данные (рис. 7):

|  |  |  |
| --- | --- | --- |
|  | **X** | **Y** |
| 0,6 | 5,1 |
| 0,8 | 5,1 |
| 0,9 | 5,7 |
| 1,5 | 14,4 |
| 1,9 | 16,2 |
| 3,6 | 16,3 |
| 4,6 | 21,3 |
| 5,1 | 22,5 |
| 6,1 | 28,2 |
| 6,5 | 31,2 |
| 6,7 | 33,5 |
| 7,2 | 34,1 |
| *Рис. 7 – Исходные данные* | 7,4 | 34,9 |

Здесь: **X** – независимая переменная; **Y** – зависимая переменная

**Шаг 1.** Из **Переключателя модулей** STATISTICA откройте модуль **Multiple regression (множественная регрессия)**: высветите название модуля и далее щелкните мышью по названию модуля **Multiple regression**.

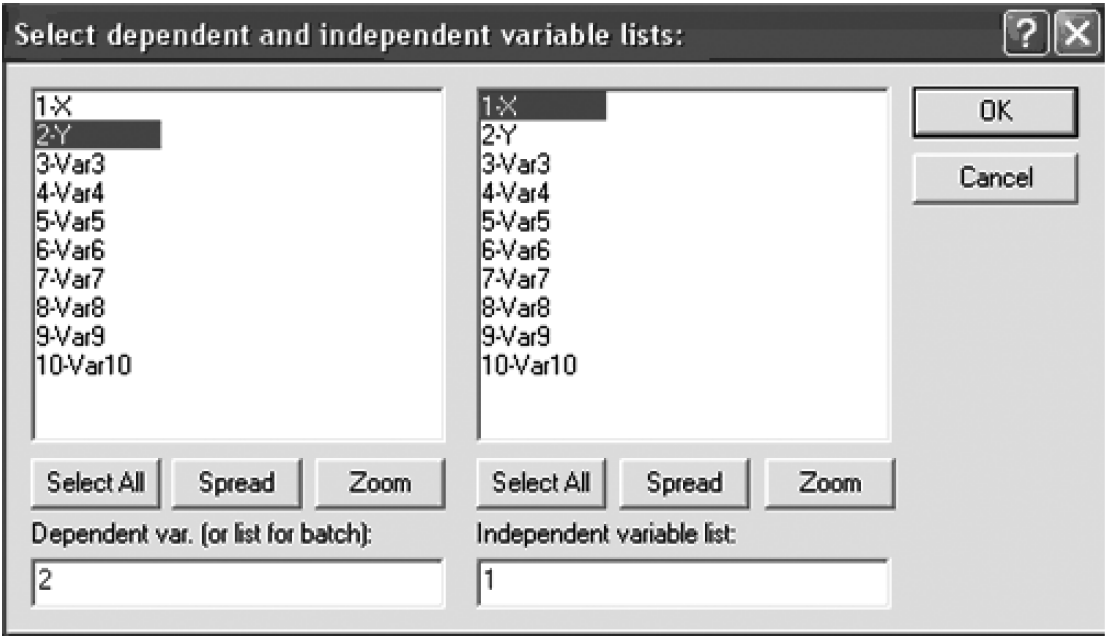
**Шаг 2.** На экране появится стартовая панель модуля (рис. 8):



*Рис. 8 – Стартовая панель модуля Multiple regression*

Выполните установки, как показано на рис. 8. Далее выберите переменные для анализа с помощью кнопки **Variables (переменные)**.

После того как кнопка будет нажата, на вашем экране появится диалоговое окно **Select dependent and independent variable list (выбрать списки зависимых и независимых переменных)** (рис. 9):

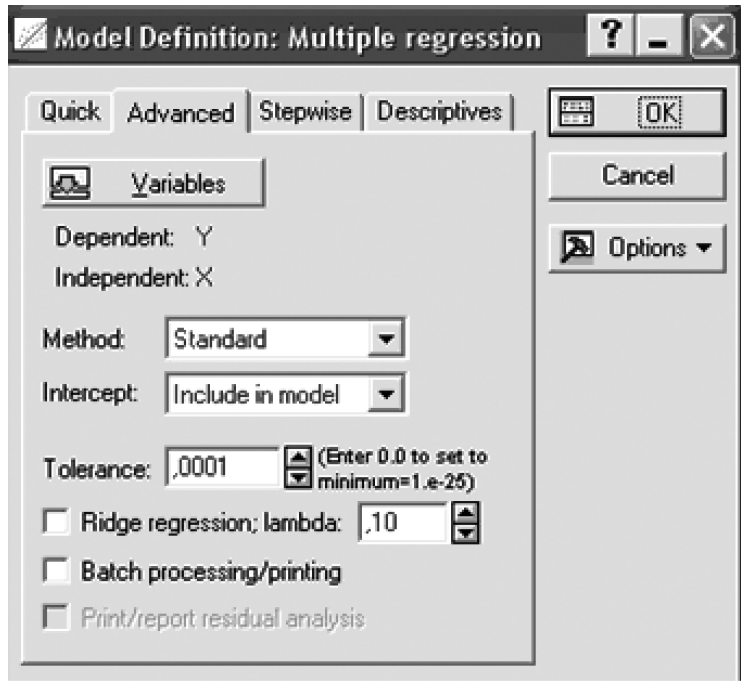


*Рис. 9 – Окно выбора переменных для анализа*

Высветив имя переменной в левой части окна, выберите зависимую переменную. Высветив имя переменной в правой части окна, выберите независимую переменную (напомним, что в данном примере независимой переменной является X, зависимой – Y).

Высветив эти переменные, как показано на рисунке, нажмите кнопку **ОК**. Вы вновь окажетесь в стартовой панели модуля. Нажмите кнопку **ОК** в правом углу стартовой панели.

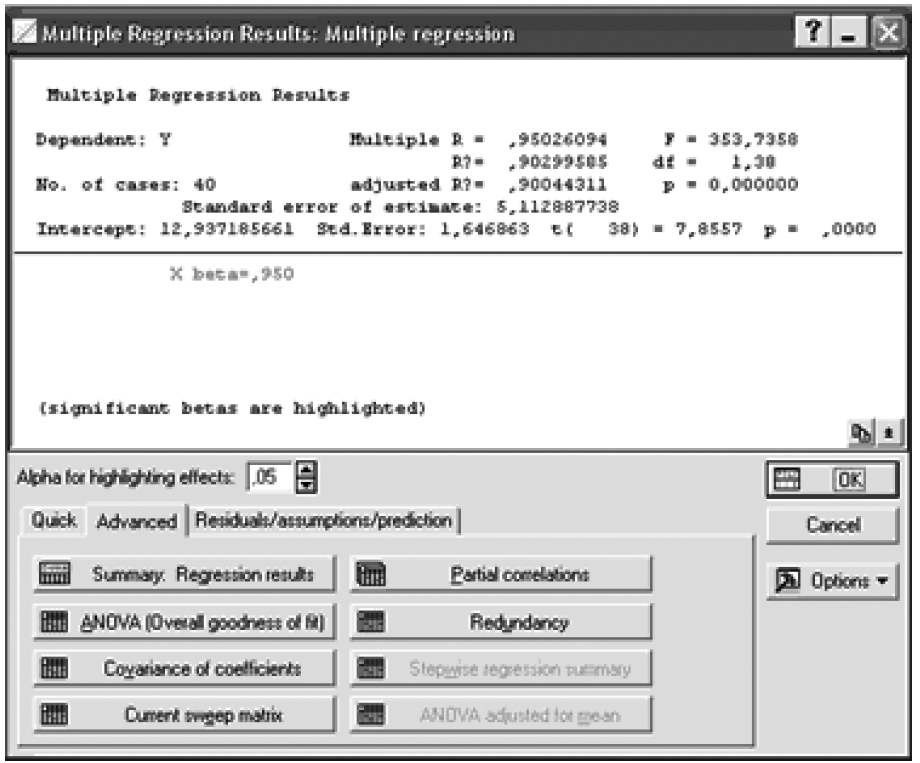
**Шаг 3**. На экране появится диалоговое окно **Model Definition (построение модели)** (рис. 10):



*Рис. 10 – Окно построения модели в модуле* ***Множественная регрессия***

В данном окне выберите стандартный метод оценивания, в опции **Method (метод)**: **Standard (стандартный)**. Далее нажмите кнопку **OK**. Программа произведет оценивание параметров модели и на экране появится диалоговое окно результатов.

**Шаг 4**. В диалоговом окне **Multiple Regression Results (результаты множественной регрессии)** просмотрите результаты оценивания (рис. 11). Результаты можно просмотреть в численном и графическом виде.



*Рис. 11 – Окно оценивания параметров*

- **Dep. Var.** (имя зависимой переменной). В нашем случае – Y.

- **No. of Cases** (число случаев, по которым построена регрессия). В нашем примере это – 40.

- **Multiple R** (коэффициент множественной корреляции).

- **R2** (квадрат коэффициента множественной корреляции), обычно называемый коэффициентом детерминации. Это чрезвычайно важная характеристика, одна из основных статистик в данном окне и показывает долю общего разброса (относительно выборочного среднего зависимой переменной), которая объясняется построенной регрессией.

- **Adjusted R2** (скорректированный коэффициент детерминации).

- **Std. Error of estimate (стандартная ошибка оценки)**. Это статистика является мерой рассеяния наблюдаемых значений относительно регрессионной прямой.

- **Intercept (оценка свободного члена регрессии)**. Значение коэффициента **В0** в уравнении регрессии.

- **Std. Error** (стандартная ошибка оценки свободного члена). Стандартная ошибка коэффициента **В0** в уравнении регрессии.

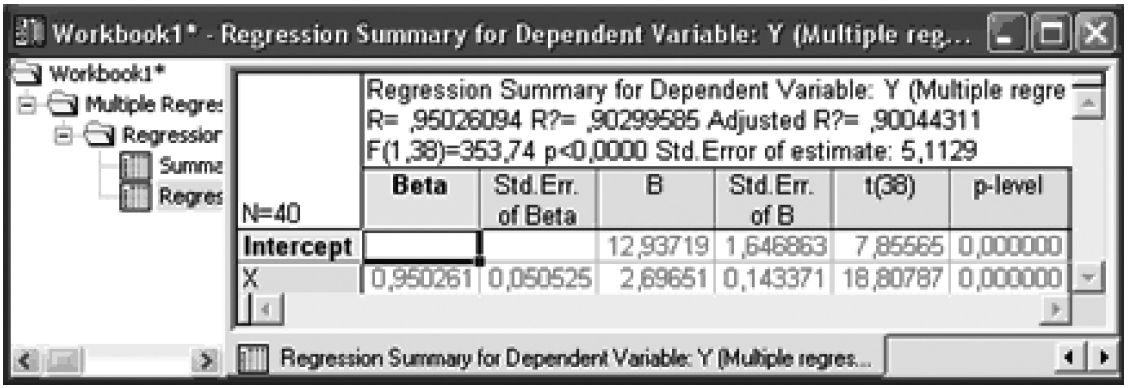
- **t(df) and p-value** (значение t-критерия и уровень р). t-критерий используется для проверки гипотезы о равенстве 0 свободного члена регрессии.

- **F** – значения F-критерия;

- **df** – число степеней свободы F-критерия;

- **р** – уровень значимости.

Нажмите кнопку **Regression summary (итоговый результат регрессии)**. На экране появится электронная таблица вывода, в которой представлены итоговые результаты оценивания регрессионной модели (рис. 12).



*Рис. 12 – Итоговая таблица регрессии*

В первом столбце таблицы даны значения коэффициентов **beta – стандартизованные коэффициенты регрессионного уравнения**, во втором – стандартные ошибки **beta**, в третьем – точечные оценки параметров модели.

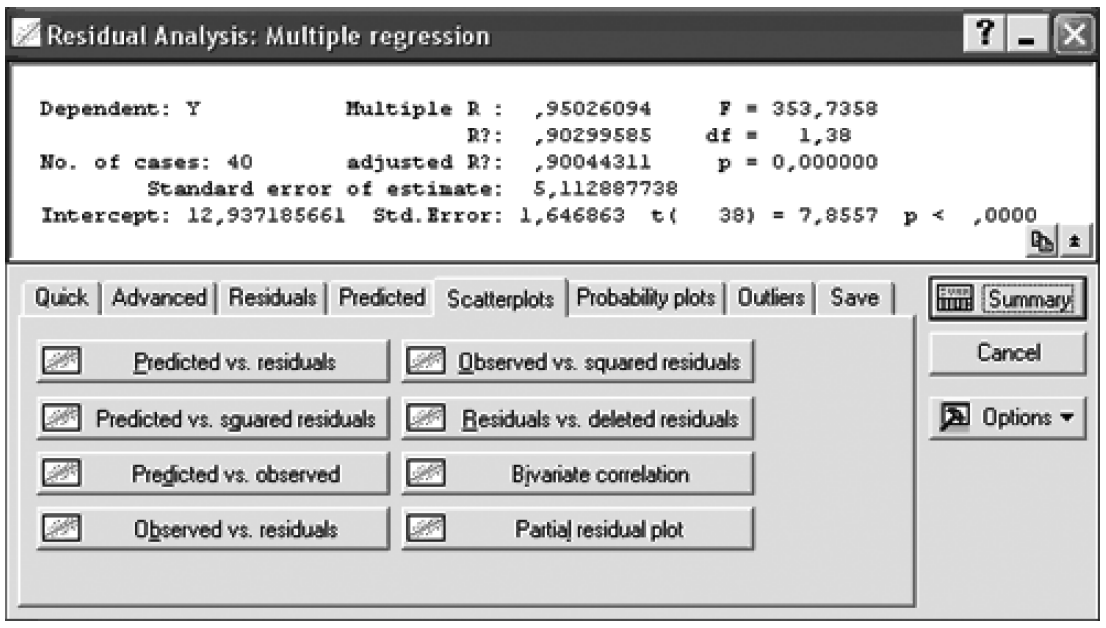
Свободный член **В0 = 12,93719**.

Коэффициент **В1** (при независимой переменной **X**) **= 2,69651**.

Далее, стандартные ошибки для **В0, В1**, значения **t-критерия** и т. д.

Из таблицы видно, что оцененная модель имеет вид:

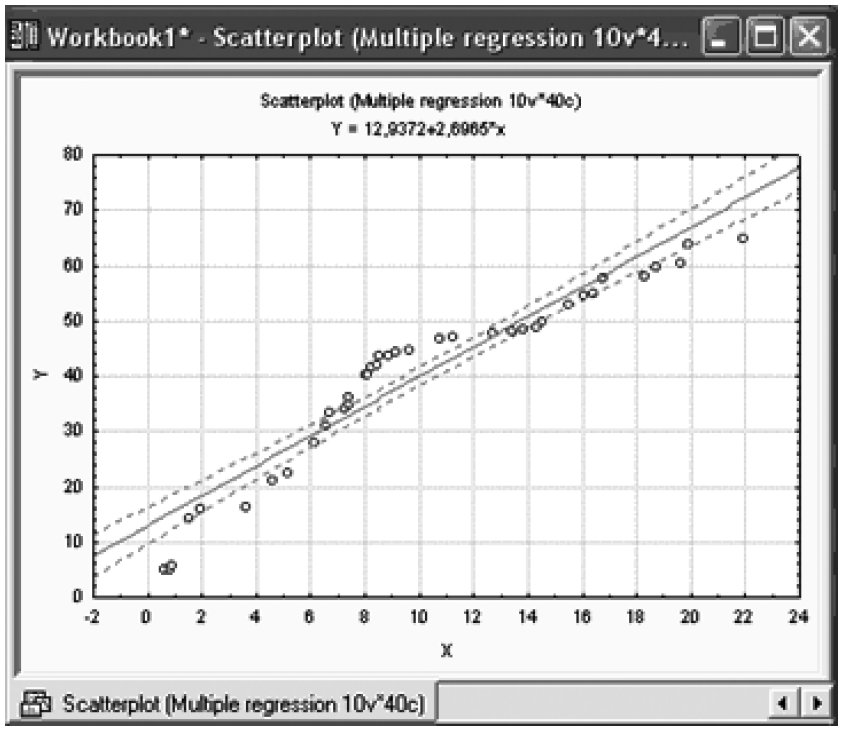
**Y = 2,69651\*X +12,93719**



*Рис. 13 – Диалоговое окно Анализ остатков в модуле*

**Шаг 5**. В окне оценивания параметров (рис. 11) последовательно нажать кнопки **OK** (рис. 13), **Scatterplot**, **Bivariate Correlation**. В появившемся окне выбрать зависимую (**Y variable**) и независимую (**X variable**) переменные. Нажать кнопку **OK**.

На полученном графике данные с подогнанной прямой имеют вид (рис. 14).



*Рис. 14 – Линейная регрессия для данных X и Y*

**Задание 9.** Вычислите коэффициент регрессии по следующему ряду данных (в мм) о длине хвоста *х*и общей длине тела *у*самок королевской змеи *Lampropeltis polyzona*:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *x* | 37 | 49 | 50 | 51 | 53 | 54 | 68 | 86 | 93 | 106 |
| *у* | 284 | 375 | 353 | 366 | 418 | 408 | 510 | 627 | 683 | 820 |
| *x* | 130 | 137 | 142 | 142 | 146 | 149 | 155 | 156 | 187 |  |
| *у* | 1056 | 986 | 1086 | 1086 | 1078 | 1122 | 1254 | 1202 | 1387 |  |

Постройте график и уравнение линейной регрессии.

**Задание 10.** Между возрастом овцематок *х*(в годах) и длительностью плодоношения ягнят *у*(в днях) оказалась следующая зависимость:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| *х* | 2 | 3 | 4 | 5 | 6 | 7 |
| *у* | 149,5 | 149,3 | 150,0 | 150,9 | 150,5 | 151,4 |

Определите коэффициент корреляции и регрессии, постройте график и уравнение линейной регрессии.

**Задание 11.** Фактическая урожайность зерновых культур (в ц/га) в одном совхозе по годам была следующей:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *Годы* | 1953 | 1954 | 1955 | 1956 | 1957 | 1958 | 1959 |
| *Урожайность* | 7,8 | 7,7 | 8,5 | 10,0 | 8,4 | 8,3 | 10,5 |

Определите коэффициент корреляции и регрессии, постройте график и уравнение линейной регрессии.

**Задание 12.** У 10 телят по глубине груди х (в см) и живому весу у (в кг) были получены следующие данные:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *х* | 91 | 86 | 94 | 95 | 104 | 92 | 98 | 84 | 96 | 99 |
| *у* | 62 | 43 | 60 | 73 | 87 | 65 | 79 | 52 | 65 | 68 |

Определите коэффициент корреляции и регрессии, постройте график и уравнение линейной регрессии.

**Задание 13.** Известны данные для 10 бычков о весе при рождении х (в кг) и суточном привесе у (в г):

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *х* | 38,5 | 46,0 | 43,0 | 43,0 | 40,5 | 44,0 | 38,0 | 35,0 | 40,5 | 54,0 |
| *у* | 694 | 901 | 736 | 1005 | 841 | 743 | 896 | 863 | 855 | 830 |

Определите коэффициент корреляции и регрессии, постройте график и уравнение линейной регрессии.

**Задание 14.** Были получены следующие данные о потреблении кислорода *у* пиявками (в г на кг/час) в зависимости от температуры *х* (в град):

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *x* | 5,5 | 5,6 | 6,2 | 8,4 | 9,0 | 10,5 | 16,1 |
| *у* | 16,1 | 14,9 | 18,8 | 32,5 | 32,1 | 37,1 | 88,5 |
| *x* | 16,6 | 17,1 | 18,8 | 19,8 | 20,0 | 20,7 | 26,5 |
| *у* | 91,0 | 94,0 | 122,0 | 162,0 | 167,0 | 187,0 | 436,0 |

Определите коэффициент корреляции и регрессии, постройте график и уравнение линейной регрессии.

**Задание 15.** Под влиянием облучения рентгеновыми лучами наблюдалось следующее замедление размножения вируса мозаики Аукуба *у* (в тыс.) в зависимости от длительности облучения *х* (в мин):

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *х* | 0 | 3 | 7,5 | 15 | 30 | 45 | 60 |
| *у* | 271 | 226 | 209 | 108 | 59 | 29 | 12 |

Определите коэффициент корреляции и регрессии, постройте график и уравнение линейной регрессии.

**Задание 16.** Было учтено среднесуточное количество перевариваемых веществ корма *х* (в кг),съеденного коровой за 12 месяцев лактации:

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *Месяц* | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| х | 14,3 | 12,8 | 13,2 | 13,6 | 13,4 | 13,2 | 12,9 | 12,8 | 12,5 | 12,2 | 11,9 | 11,5 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |

Определите коэффициент корреляции и регрессии, постройте график и уравнение линейной регрессии.