

## Нелинейные регрессионные модели и их графики в STATISTICA

*Кульков Алексей Владимирович*  
*Смоленский государственный университет*  
*магистрант*

### Аннотация

В данной работе предложен способ составления множественных нелинейных регрессионных моделей в системе для статистического анализа Statistica. Также рассмотрен способ построения графика множественных нелинейных регрессионных уравнений в том случае, если имеется две независимых переменных.

**Ключевые слова:** Statistica, регрессионный анализ, множественная нелинейная регрессия, 3D графики.

## Nonlinear regression models and their graphs in STATISTICA

*Kulkov Alexey Vladimirovich*  
*Smolensk State University*  
*postgraduate*

### Abstract

In this paper, we propose a method for constructing multiple nonlinear regression models in a system for statistical analysis of Statistica. A method for plotting the plot of multiple nonlinear regression equations is also considered in the case when there are two independent variables.

**Keywords:** Statistica, regression analysis, multiple nonlinear regression, 3D graphics.

При анализе статистических закономерностей между характеристиками какого-либо процесса или явления и составлении математических моделей этих закономерностей часто используется регрессионный анализ. В большинстве случаев полученные модели имеют нелинейный характер относительно включенных в анализ объясняющих переменных, что вызывает особые трудности в построении данных моделей. Часто эти трудности связаны с выбором средства для проведения регрессионного анализа и с отсутствием знаний о функции регрессии.

В данной работе предложен способ составления множественных нелинейных регрессионных моделей в системе для статистического анализа Statistica, что является актуальным вопросом в области статистического анализа данных.

На первом этапе составления множественной нелинейной регрессионной модели целесообразно построить двумерный график, на

котором по оси  $y$  отметить значения независимой переменной  $Y$ , а на оси  $x$  – одну из зависимых переменных  $X$ . По полученному графику можно предположить о виде зависимости, то есть той функции, которая может её (зависимость) описать. Проводя данную операцию для всех зависимых переменных, можно, приблизительно, судить о функции зависимости для каждой из данных переменных [1].

Рассмотрим конкретный пример. Имеются данные о 30 значений переменной  $Y$  в зависимости от  $X_1$  и  $X_2$ . Все значения занесены в программу Statistica. Для построения графика, на котором отмечены экспериментальные точки  $Y$  и  $X_1$ , необходимо в строке меню программы выбрать пункт «Графика», а далее во вкладке «2М Графики» выбрать «Диаграммы рассеяния». После этого выбираются переменные, которые необходимо отметить на графике. На рисунке 1 «Графики» представлены два графика, на которых отмечены экспериментальные точки для  $Y, X_1$  и  $Y, X_2$ .

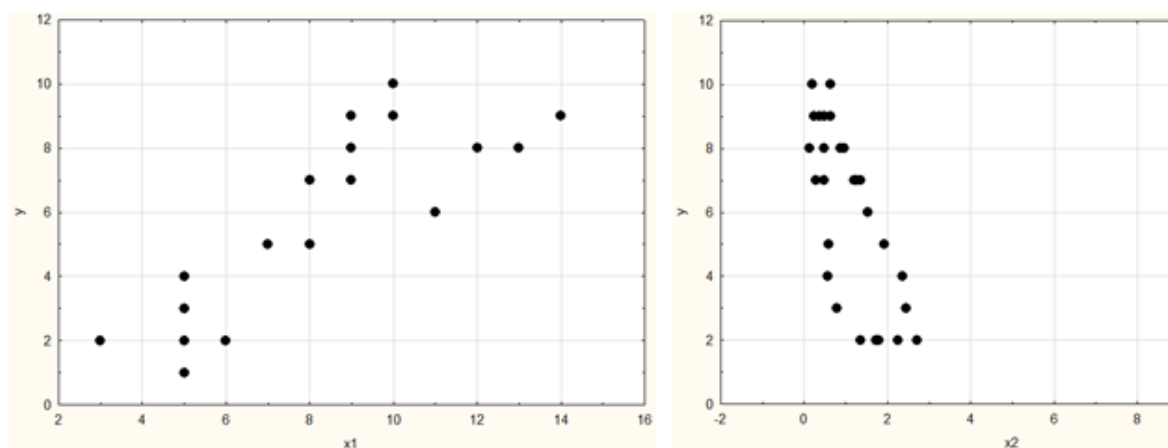


Рисунок 1. Графики

Анализируя рисунок 1, можно предположить, что между  $Y$  и  $X_1$  (рисунок слева) имеет место линейная статистическая зависимость, а между  $Y$  и  $X_2$ . (рисунок справа) – гиперболическая статистическая зависимость. Таким образом, регрессионная модель будет иметь вид:

$$Y = a_0 + a_1 \cdot X_1 + \frac{a_2}{X_2} \quad (1)$$

После решения вопроса о виде зависимостях, можно приступить к построению множественной нелинейной регрессионной модели.

Программа Statistica предоставляет несколько способов для построения нелинейных регрессионных моделей. Первый способ подразумевает выбор функции зависимости автоматически, что является в некоторых случаях недостаточным для построения качественной регрессионной модели.

В пункте «Анализ» во вкладке «Углублённые методы анализа» выбрать «Множественная нелинейная регрессия». В появившемся окне выбрать анализируемые переменные. После этого программа предлагает подключить интересные функции. Исходя из заключения о виде статистической

зависимости, добавляем функцию  $1/x$ . Следуеь отметить, что линейная зависимость включена автоматически. После подключения интересующих функций программа предлагает выбрать зависимые и независимые переменные, что показано на рисунке 2 «Определение модели».

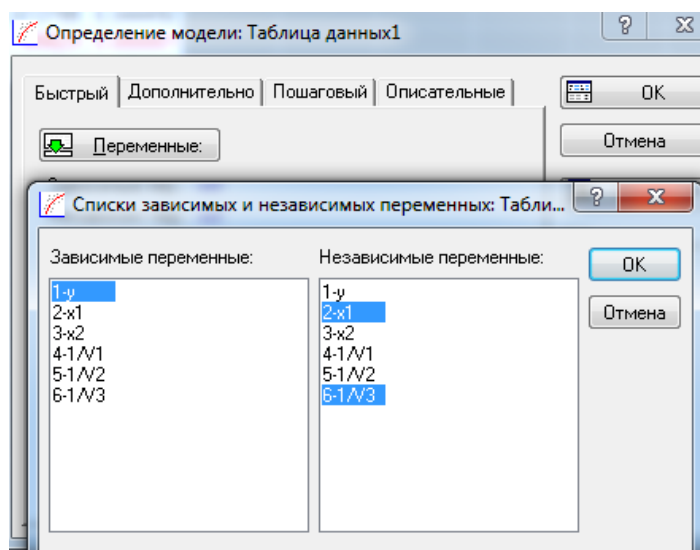


Рисунок 2. Определение модели

В качестве зависимой переменной выбирается  $Y$ , а в качестве независимых  $X_1$  и  $1/V3$ . Символ  $V3$  соответствует той переменной, которая находится в исходной таблице в качестве третьей, то есть в данном случае это  $X_2$ . После задания вида модели появится окно с результатами регрессионного анализа, которые представлены на рисунке 3 «Результаты регрессионного анализа №1».

Итоги регрессии для зависимой переменной: y (Таблица данных1)						
R= ,84927815 R2= ,72127337 Скоррект. R2= ,70062695						
F(2,27)=34,935 p<,00000 Станд. ошибка оценки: 1,5307						
N=30	БЕТА	Ст.Ош. БЕТА	В	Ст.Ош. В	t(27)	p-знач.
Св.член			-0,539833	0,879511	-0,613788	0,544492
x1	0,721655	0,116587	0,720862	0,116459	6,189829	0,000001
1/V3	0,216820	0,116587	0,375185	0,201742	1,859725	0,073853

Рисунок 3. Результаты регрессионного анализа №1

Таким образом, математическая модель данной зависимости имеет вид:

$$Y = -0,54 + 0,72 \cdot X_1 + \frac{0,38}{X_2} \quad (2)$$

Следует отметить, что полученная модель (2) описывает 72% дисперсии зависимой переменной.

Второй способ составления множественной модели является более универсальным и требует от пользователя задание вида уравнения регрессии самостоятельно. Для этого в пункте «Анализ» во вкладке «Углублённые методы анализа» необходимо выбрать «Нелинейное

оценивание». В появившемся окне задания регрессионной функции требуется ввести ту функцию, которую получили при анализе одномерных графиков. В рассмотренном примере данная функция имеет вид  $V1=a0+a1*V2+a2/V3$ . Индекс в символе  $V$  соответствует номеру переменной,  $a0$ ,  $a1$  и  $a2$  – это коэффициенты уравнения. В результате программа выдаст результаты данной регрессионной модели, которые представлены на рисунке 4 «Результаты регрессионного анализа №2».

Модель: $V1=a0+a1*V2+a2/V3$ (Таблица данных1)						
Зав. Пер. : y						
Уров. значимости: 95.0% ( альфа=0.050)						
	Оценка	Стандарт ошиб.	t-знач. сс = 27	p-знач.	Ниж. Дов Предел	Вер. Дов Предел
a0	-0,539833	0,879511	-0,613788	0,544493	-2,34444	1,264774
a1	0,720862	0,116459	6,189829	0,000001	0,48191	0,959816
a2	0,375185	0,201742	1,859725	0,073853	-0,03876	0,789125

Рисунок 4. Результаты регрессионного анализа №2

Анализируя рисунки 3 и 4, можно сделать вывод, что оба способа построения нелинейной множественной регрессионной модели дают одинаковые результаты.

Если в качестве объясняющих факторов (независимых переменных) выступает имеется 2 переменные, то возможно построение трёхмерного графика уравнения множественной регрессии. Для этого необходимо в строке меню программы выбрать пункт «Графика», а далее во вкладке «3М XYZ Графики» выбрать «Графики поверхностей». После этого выбираются переменные, которые необходимо отметить на графике и выполнить построение графика. В результате появится трёхмерная система координат, на которой отмечены экспериментальные точки. При выборе свойств графика, пользователь имеет возможность включить на данный график график функции полученной при регрессионном анализе математической модели. Программа имеет встроенные функции подгонки (то есть функции, которые будут наилучшим образом описывать набор экспериментальных точек), а также предоставляет возможность задание собственной функции подгонки. Для этого в свойствах графика в пункте «Подгонка» необходимо ввести уравнение множественной регрессии. На рисунке 5 «График поверхности» показан график, который получился для составленной регрессионной модели (2).

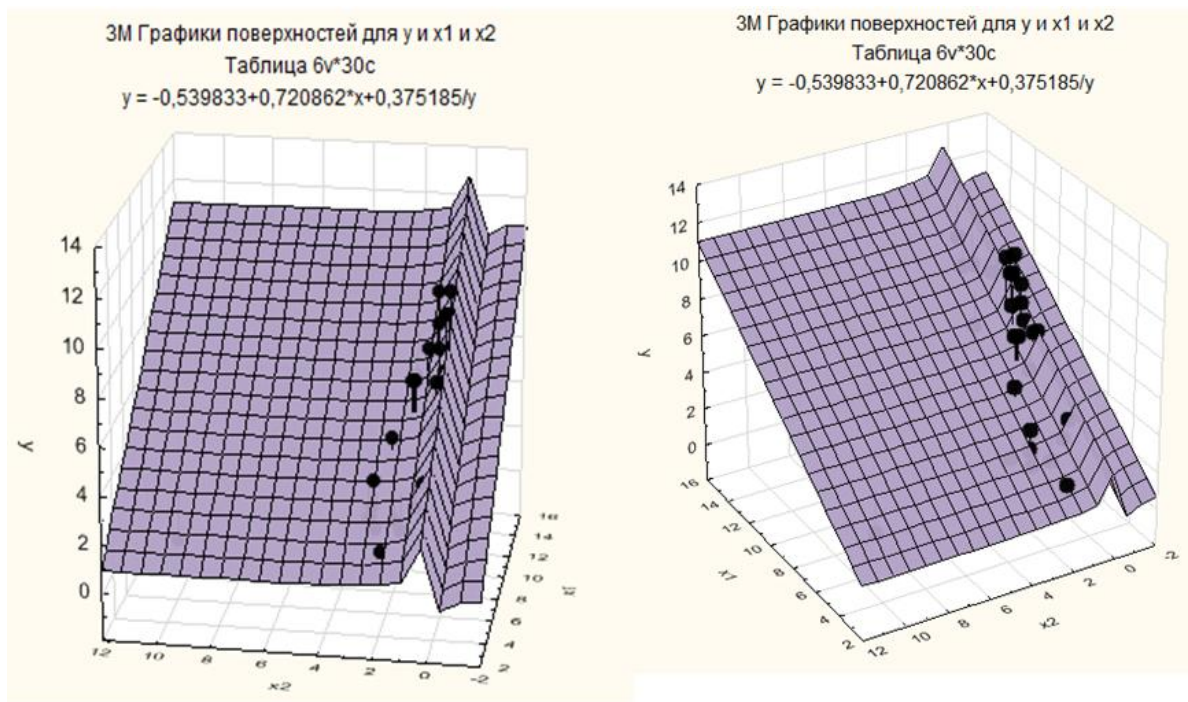


Рисунок 5. График поверхности

На рисунках 6 «Квадратичная подгонка» и 7 «Линейная подгонка» приведены графики квадратической и линейной поверхности для экспериментальных данных, которые построились программой автоматически при выборе соответствующей функции подгонки.

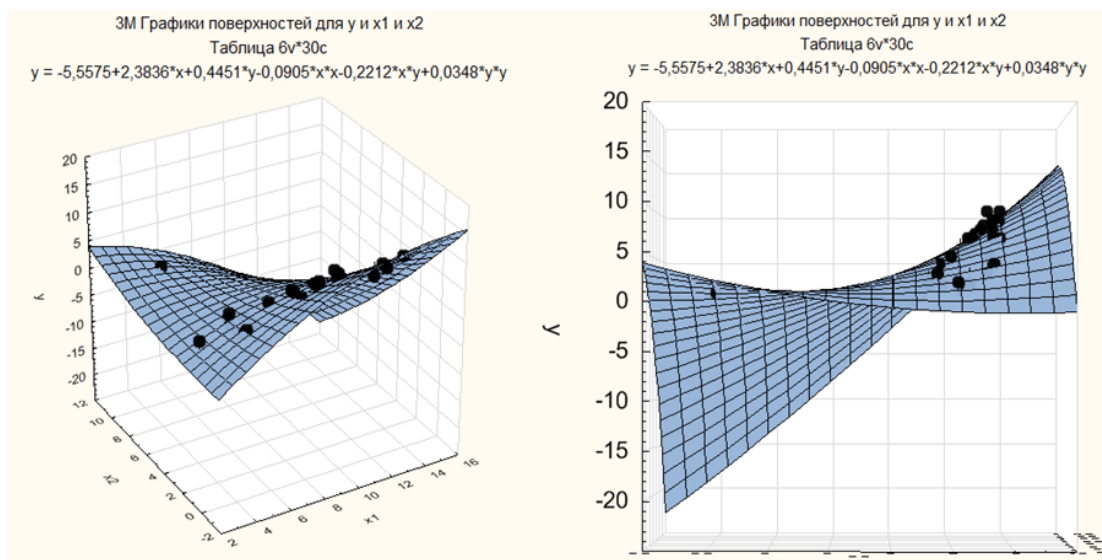


Рисунок 6. Квадратичная подгонка



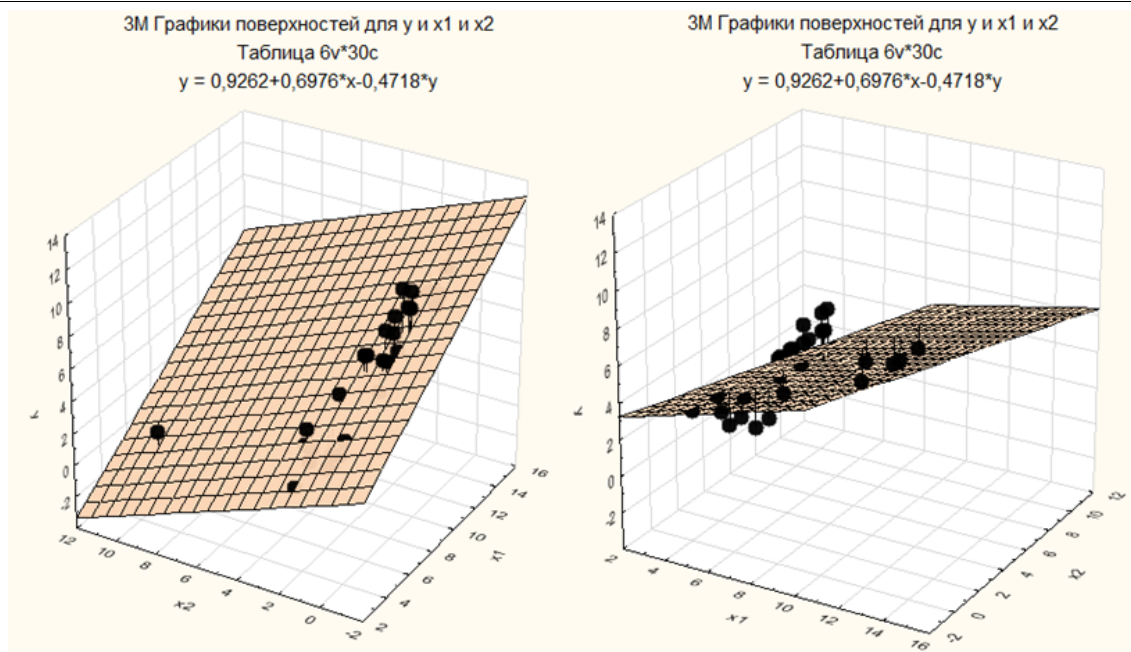


Рисунок 7. Линейная подгонка

Следует отметить, что при выборе линейной или квадратичной функции подгонки программа автоматически рассчитывает и выдаёт уравнение плоскости, которое наилучшим образом описывает набор экспериментальных точек. Таким образом, если в объясняющих (независимых) две, то можно сначала отметить данные точки на графике, добавить одну из двух функций подгонки. Далее посмотреть на аналитический вид получившейся функции и принять её в качестве уравнения множественной регрессии.

Таким образом, нелинейный множественный регрессионный анализ в программе для статистического анализа данных Statistica реализован достаточно хорошо. Помимо построения регрессионной модели и её статистического анализа, Statistica позволяет построить график поверхности для случая, когда имеется две независимых переменных.

### Библиографический список

1. Эконометрика: учебник/ Ю.В. Сажин, И.А. Иванова. Саранск: Мордов. гос.ун-т, 2014. 316 с.