

Лабораторная работа № 3

«КРИВОЛИНЕЙНАЯ КОРРЕЛЯЦИЯ И РЕГРЕССИЯ В RAST»

Если связь между изучаемыми явлениями существенно отклоняется от пропорциональной (это легко установить по графику), то коэффициент линейной корреляции непригоден в качестве меры связи. Он может указать на отсутствие сопряженности там, где налицо сильная криволинейная зависимость. Поэтому необходим новый показатель, который правильно измерял бы степень криволинейной зависимости. Таким показателем является корреляционное отношение, обозначаемое греческой буквой η (эта). Оно измеряет степень корреляции при любой ее форме.

Криволинейная связь между признаками – это такая связь, при которой равномерным изменениям первого признака соответствуют неравномерные изменения второго, причем эта неравномерность имеет определенный закономерный характер.

При графическом изображении криволинейных связей, когда по оси абсцисс откладывают значения первого признака (аргумент, независимая переменная), а по оси ординат – значения второго признака (функция, зависимая переменная) и полученные точки соединяют, получают изогнутые линии. Характер изогнутости зависит от природы коррелируемых признаков.

Корреляционное отношение измеряет степень корреляции при любой ее форме.

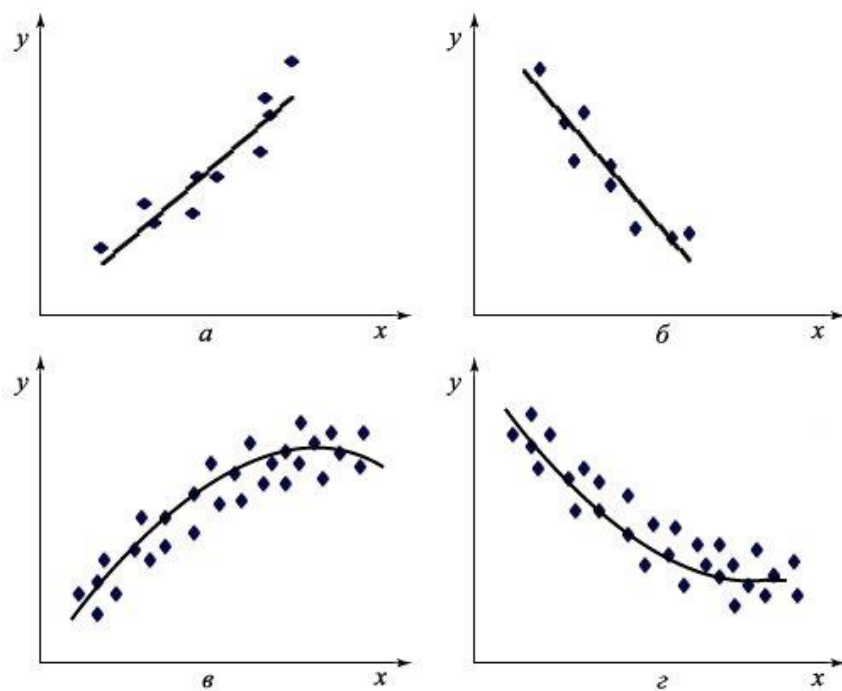


Рисунок 1 – Характер корреляционной связи: а, б – прямолинейная, в, г – криволинейная

В отличие от коэффициента линейной корреляции, который дает одинаковую меру связи признаков (первого со вторым и второго с первым), корреляционное отношение второго признака по первому обычно не бывает равно корреляционному отношению первого признака по второму. Поэтому крайне важно определить какая выборка является аргументом, а какая функцией.

По виду линии на графике можно определить характер связи (прямолинейная или криволинейная), а также тип аппроксимации (рисунок 1).

Задачей исследователя является подобрать вид функции, которая бы наиболее четко ложилась на поле регрессии.

Рассмотрим криволинейную корреляцию и регрессию на следующем примере.

Задание 1. Исследовалось влияние вводимого перорально с кормом сорбента (x) в граммах молочным коровам на снижение содержания ^{137}Cs в молоке (y) в процентах (таблица 1).

Таблица 1 – Дозы вводимого сорбента (x) и снижение содержания ^{137}Cs в молоке (y)

x	14,5	5,1	6,1	6,5	6,7	8,2	7,4	10,7	14,3	8,1	8,1	4,6	8,4
y	49,8	22,5	28,2	31,2	33,5	41,8	34,9	46,9	48,8	40,3	40,3	21,3	42,2
x	9,6	8,8	9,1	9,6	7,4	8,5	11,2	12,7	13,4	13,8	7,2	8,0	
y	44,9	43,9	44,3	44,8	36,2	43,9	47,1	47,9	48,2	48,7	34,1	40,3	

Решение:

Создадим электронную таблицу с данными (рисунок 1)

учебный пример.dat

File Edit Transform Plot Univariate M		
Show		Click mode
☐ Row attributes	☒ Select	
☐ Column attributes	☐ Drag rows/cc	
	X	Y
1	• 14,5	49,8
2	• 5,1	22,5
3	• 6,1	28,2
4	• 6,5	31,2
5	• 6,7	33,5
6	• 8,2	41,8
7	• 7,4	34,9
8	• 10,7	46,9
9	• 14,3	48,8
10	• 8,1	40,3
11	• 8,1	40,3
12	• 4,6	21,3
13	• 8,4	42,2
14	• 9,6	44,9
15	• 8,8	43,9

Рисунок 1 – Электронная таблица с данными

Затем, предварительно выделив данные с переменными выбираем в меню пункт **Model** → **Nonlinear fit** (*Модель* → *Нелинейные зависимости*) – рисунок 2.

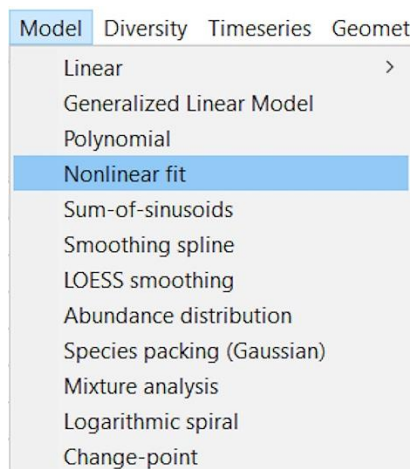


Рисунок 2 – Пункт меню **Nonlinear fit** (*Нелинейные зависимости*)

После выбора этого пункта меню на экране появляется результат в виде линейной регрессии и уравнения ее функции (рисунок 3).

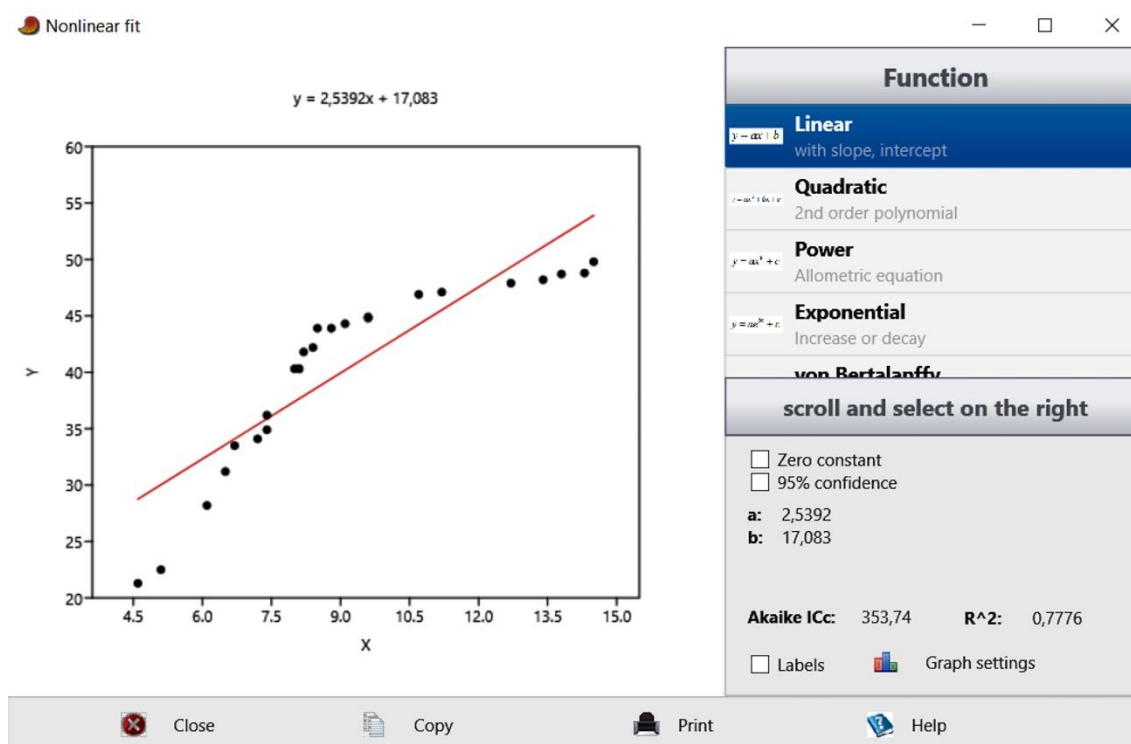


Рисунок 3 – Результаты регрессионного анализа по умолчанию в разделе **Nonlinear fit**

Следует обратить внимание, что по умолчанию программа PAST делает именно линейный регрессионный анализ. С правой стороны окна результатов размещен список функций, которые можно использовать для построе-

ния определенной модели. Для наших целей необходимо выбрать квадратичную функцию, основанную на квадратном уравнении – **Quadratic (2nd order polinomial)** (*Квадратичная, полиномиальная второго порядка*) – рисунок 4.

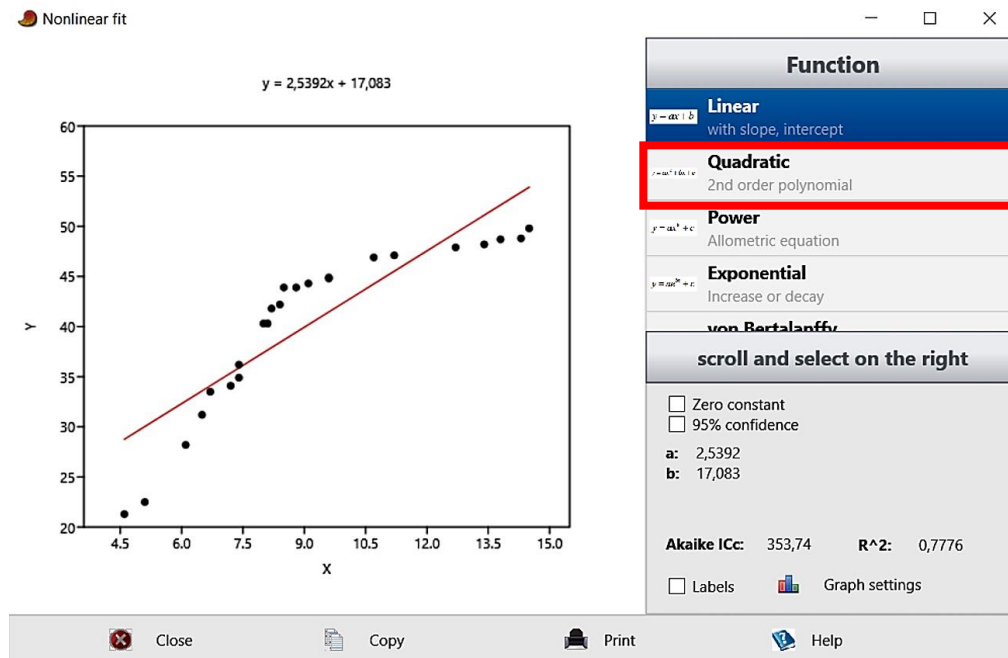


Рисунок 4 – Выбор квадратичной функции **Quadratic**

В итоге в окне результатов будет отображен график и квадратное уравнение полученной криволинейной регрессии (рисунок 5).

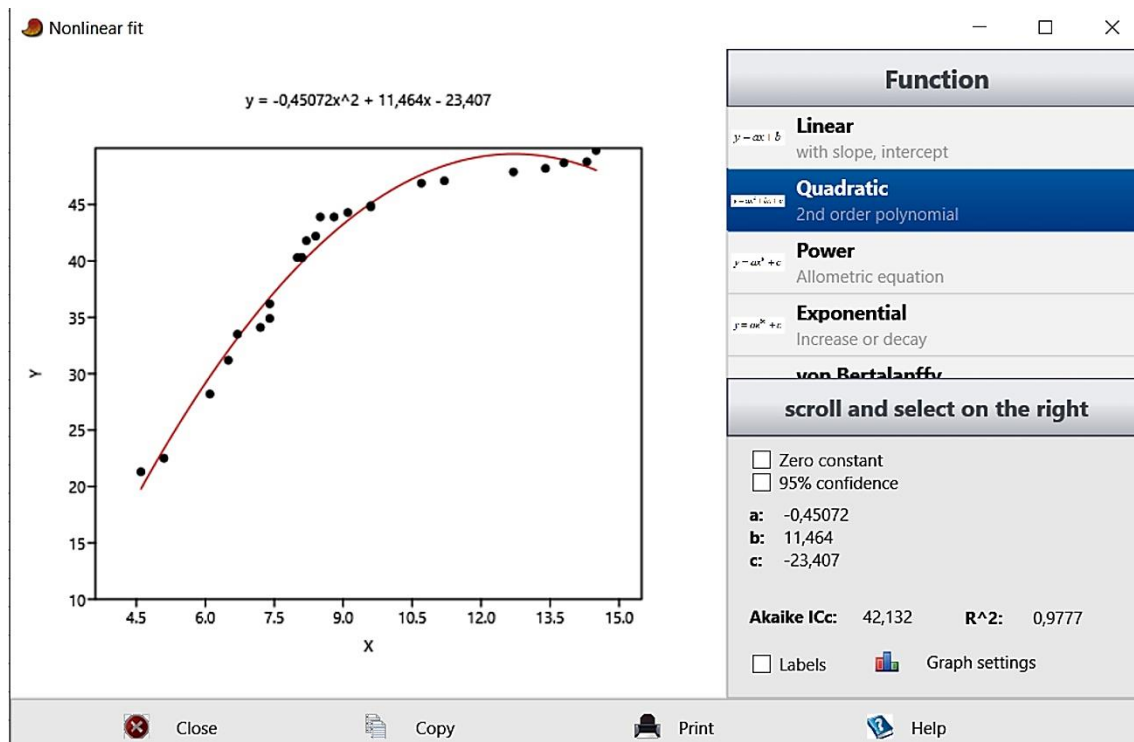


Рисунок 5 – График криволинейной регрессии функции **Quadratic**

Для более полной информации добавим на график диапазон, ограничивающий 95% доверительный интервал – промаркируем бокс **95% confidence** и зайдем в настройки графика (**Graph settings**) подпишем оси и настроим график согласно рисунку 6.

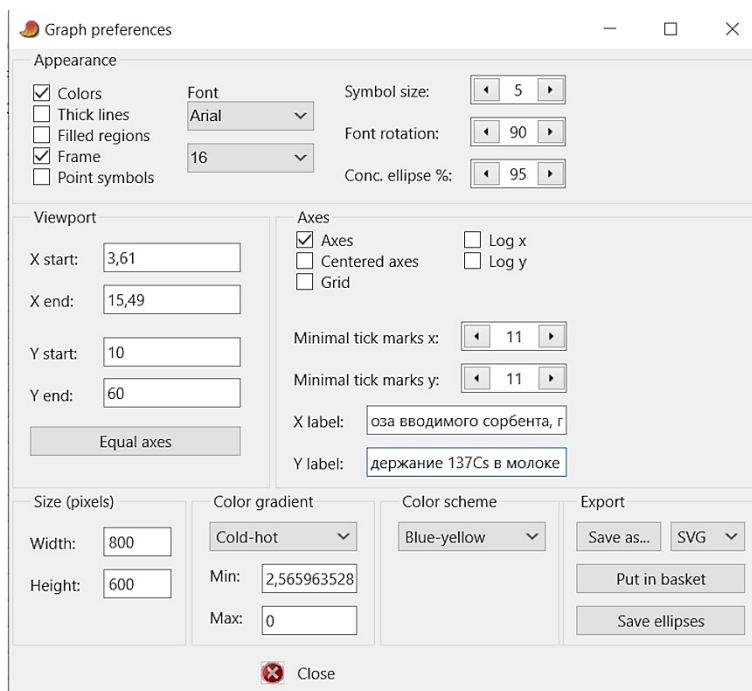


Рисунок 6 – Окно настроек графика криволинейной регрессии

Сравните полученный график с рисунком 7.

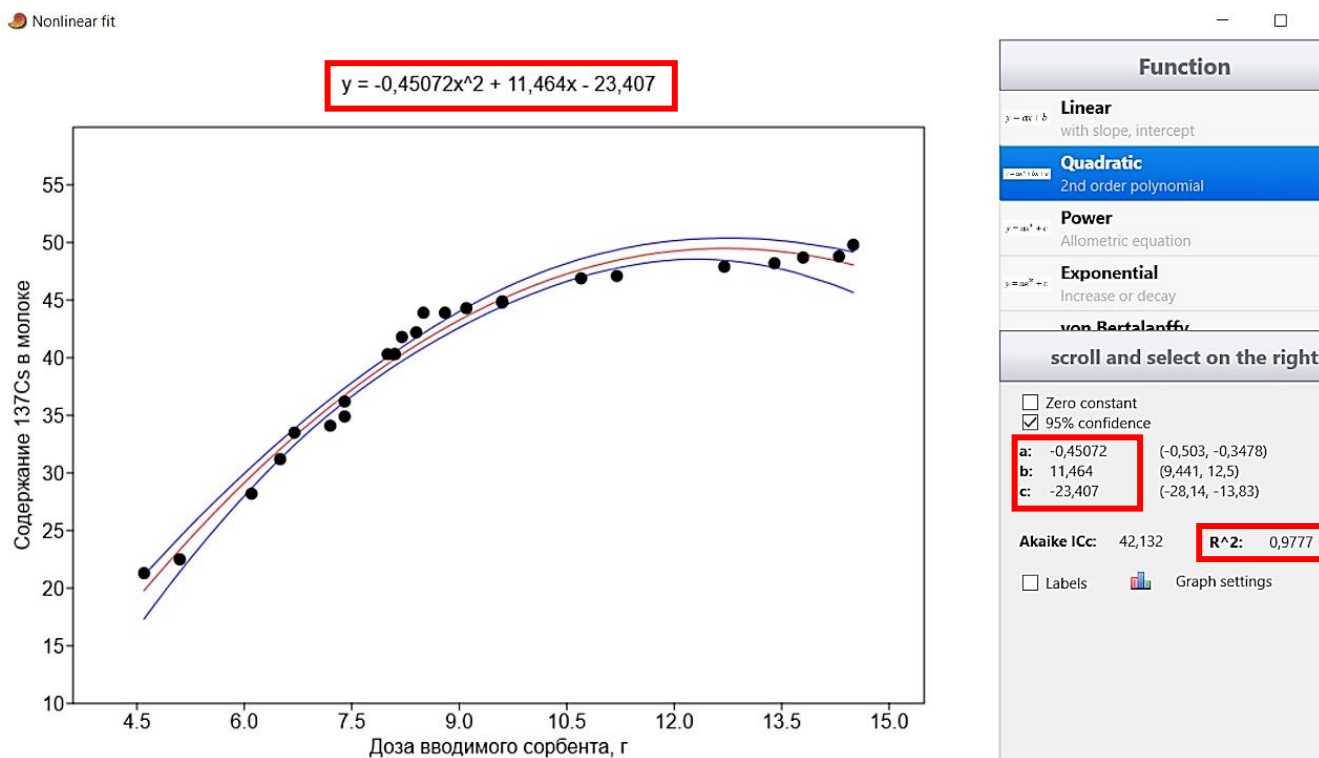


Рисунок 7 – Итоговый график криволинейной регрессии

Опишем полученный график:

- коэффициент детерминации (R^2) составил 0,9777, т.е. ~98% всех данных включены в полученную модель;

- полученные коэффициенты квадратного уравнения имеют значения: $a = -0,45072$; $b = 11,464$; $c = -23,407$;

- полученное квадратное уравнение регрессии $y = ax^2 + bx + c$ имеет вид $y = -0,45072x^2 + 11,464x - 23,407$.

Выполните задания для самоконтроля.

Задания для самоконтроля

Выполните расчеты криволинейной корреляции и регрессии. Объясните полученный результат.

Варианты заданий													
1		2		3		4		5		6		7	
X	Y	X	Y	X	Y	X	Y	X	Y	X	Y	X	Y
14,5	49,8	33,3	0,03	9,06	2,13	0,9	66,9	0,97	0,4	1,5	0,7	0,87	0,15
5,1	22,5	7,1	0,2	8,95	2,12	0,01	0,01	1,21	2,79	1,47	0,69	0,98	0,11
6,1	28,2	11,2	0,13	5,58	1,61	-0,8	-46,4	0,83	0,09	1,58	0,73	1,32	0,04
6,5	31,2	14,3	0,08	11,26	2,38	1,2	165,1	1,16	1,83	1,57	0,73	1,22	0,05
6,7	33,5	17,2	0,06	9,14	2,14	1,5	314,8	0,95	0,31	1,45	0,68	0,81	0,18
8,2	41,8	22,2	0,05	2,87	0,83	1,1	141,1	1,53	22,9	1,45	0,69	1,47	0,02
7,4	34,9	34,7	0,03	12,01	2,42	-0,1	-0,1	1,09	1,13	1,02	0,51	0,96	0,11
10,7	46,9	11,2	0,1	6,86	1,81	0,4	5,8	0,48	0	1,18	0,58	0,77	0,2
14,3	48,8	3,7	0,37	7	1,89	0,8	60,5	1,1	1,15	0,92	0,46	0,93	0,12
8,1	40,3	33,8	0,03	2,86	0,99	-0,1	-0,3	1,15	1,73	0,96	0,48	0,88	0,14
8,1	40,3	10,9	0,12	7,59	1,97	1,2	175	1,04	0,73	1,78	0,79	0,96	0,11
4,6	21,3	13,1	0,08	10,47	2,25	2,6	1845,6	1,23	3,15	0,55	0,2	0,73	0,23
8,4	42,2	18,1	0,07	6,78	1,77	1,6	421,1	0,75	0,04	1,21	0,6	1,56	0,02
9,6	44,9	20,1	0,06	2,57	0,69	1,2	192,9	1,02	0,62	0,86	0,43	1,25	0,05
8,8	43,9	23,4	0,05	7,59	1,96	0,7	42,4	1,11	1,3	0,61	0,26	1,27	0,04
9,1	44,3	30,1	0,04	9,59	2,17	-0,01	-0,01	1,14	1,63	1,09	0,54	0,68	0,26
9,6	44,8	27	0,04	11,08	2,35	0,02	0,02	1,35	7,22	0,86	0,43	1,12	0,07
7,4	36,2	25,4	0,04	12,33	2,43	0,2	1,1	1,11	1,31	1,12	0,56	1,53	0,02
8,5	43,9	25,6	0,04	11,76	2,42	1,6	443,8	1,78	88,7	0,53	0,19	0,72	0,23
11,2	47,1	30,1	0,04	9,49	2,2	0,4	8,4	0,79	0,06	1,16	0,58	1,13	0,07
12,7	47,9	27,7	0,04	12,2	2,48	-0,3	-2,8	0,84	0,1	1,05	0,52	1,12	0,07
13,4	48,2	26	0,05	9,62	2,22	-0,5	-13,1	0,83	0,1	0,57	0,22	0,94	0,12
13,8	48,7	23,2	0,05	12,53	2,48	0,8	50	1,21	2,76	0,9	0,45	1,32	0,04
7,2	34,1	19	0,06	3,51	1,06	0,3	4,3	1,2	2,61	0,75	0,36	0,95	0,12
8	40,3	32,1	0,03	4,76	1,46	0,6	19,8	1,31	5,75	0,86	0,43	1,2	0,06
		16,8	0,07	10,41	2,3	0,2	1,1	1,67	49,21	0,77	0,37	0,72	0,23
		12,1	0,09	8,71	2,09	0,6	23,2	1,18	2,29	1,61	0,74	0,69	0,25
		9,3	0,17	8,7	2,12	1,4	255,9	0,65	0,01	0,68	0,31	0,96	0,11
		6,7	0,26	9,68	2,17	0,8	44,8	0,89	0,17	1,01	0,51	0,9	0,14
		21,5	0,05	10,04	2,24	1,5	326,4	0,68	0,01	1,36	0,65	1,67	0,01