Решение задачи аппроксимации.

- Нелинейная регрессия.
- Нахождение коэффициентов нелинейной аппроксимирующей зависимости путём сведения её к линейной.
- Выбор лучшей аппроксимирующей зависимости.

Содержание

- Основные виды нелинейных зависимостей, сводящихся к линейным;
- Пример 1:
  - вычисление параметров каждой из трёх теоретических зависимостей;
  - вычисление суммы квадратов отклонений для каждой зависимости;
  - отображение на графиках экспериментальных точек и теоретических зависимостей;
- Определение лучшей из предложенных теоретических зависимостей, которые описывают набор экспериментальных данных.

Нелинейная регрессия. Нахождение коэффициентов нелинейной аппроксимирующей зависимости путём сведения её к линейной

Пусть известно, что полученные экспериментальные данные $\{x_i\}$, $\{y_i\}$, i=1,2, …, n описываются нелинейной зависимостью общего вида:

$y = f(x, a, b)$

Задача состоит в нахождении параметров этой зависимости, т. е. в вычислении коэффициентов a и b.

Часто нелинейная зависимость путём элементарных математических преобразований может быть сведена к линейной.

В этом случае для вычисления коэффициентов в Excel можно будет воспользоваться функциями НАКЛОН(…) и ОТРЕЗОК (…), рассмотренными ранее.

Основные виды нелинейных зависимостей, сводящихся к линейным

1) $y = a \cdot b^x$ - показательная функция;

2) $y = a \cdot e^{b \cdot x}$ – экспоненциальная зависимость;

3) $y = \dfrac{1}{a \cdot x + b}$ - дробно-линейная функция;

4) $y = a \cdot ln(x) + b$ - логарифмическая функция;

5) $y = a \cdot x^b$ - степенная функция;

6) $y = a + \dfrac{b}{x}$ - гиперболическая функция;

7) $y = \dfrac{x}{a \cdot x + b}$ - дробно-рациональная функция.

На следующих примерах рассмотрим некоторые приёмы сведения нелинейной зависимости к линейной.

Пример 1.

Известно, что приведённые в таблице экспериментальные данные $\{x_i\}$, $\{y_i\}$, i=1,2, …, n

x	2	3,5	5	6	7
y	4,53	3,12	2,7	1,88	1,55

могут быть описаны с помощью следующих теоретических зависимостей общего вида:

$$y1(x) = a \cdot e^{b \cdot x}$$

$$y2(x) = m + k \cdot ln(x)$$

$$y3(x) = \frac{c}{x + d}$$

Пример 1.

1 Определить, какая из предложенных теоретических зависимостей наилучшим образом описывает набор экспериментальных данных $\{x_i\}$, $\{y_i\}$, для чего:

а) вычислить параметры каждой теоретической зависимости;

б) вычислить сумму квадратов отклонений для каждой зависимости;

в) отобразить на графиках (отдельно для каждой зависимости) экспериментальные точки и теоретические зависимости;

2 Предсказать значение Y при X = 11. Показать соответствующие точки на графиках.

Пример 1. Решение

Для первой зависимости:

$$y1(x) = a \cdot e^{b \cdot x}$$

Чтобы от произведения перейти к сумме и избавиться от возведения числа *e* в степень – прологарифмируем обе части выражения. Получим:

$$ln(y) = ln(a) + b \cdot x$$

Затем, выполним замену переменных:

$z = ln(y),$ $\qquad$ $c = ln(a),$

сводим зависимость к линейной:

$$z = c + x \cdot b.$$

Далее в Excel с помощью функций НАКЛОН и ОТРЕЗОК найдём коэффициенты *c* и *b*.

Затем вычислим коэффициент a : $\quad a = exp(c).$

Пример 1. Решение

<u>Решение в Excel:</u>

1 В ячейки A2:B7 введём исходные данные.

2 В ячейку C3 введём формулу =LN(B3) и скопируем её в ячейки C4:C7.

3 Для вычисления коэффициента a в ячейку B11 введём формулу =EXP(ОТРЕЗОК(C3:C7;A3:A7)).

4 Для вычисления коэффициента b в ячейку B12 введём формулу =НАКЛОН(C3:C7;A3:A7).

	A	B	C	D	E
1	$y=a*e^{b*x}$				
2	x	y	Ln(y)		
3	2	4,53	1,510722		
4	3,5	3,12	1,137833		
5	5	2,7	0,993252		
6	6	1,88	0,631272		
7	7	1,55	0,438255		
8					
9					
10					
11	a=	6,884			
12	b=	-0,210			
13					

=EXP(ОТРЕЗОК(C3:C7;A3:A7))

=НАКЛОН(C3:C7;A3:A7)

Пример 1. Решение

<u>Решение в Excel:</u>

5 Для вычисления квадратов отклонений заданной зависимости от экспериментальных данных в ячейку D3 введём формулу =($B$11*EXP($B$12*A3)-B3)^2
и скопируем её в ячейки D4:D7.

6 В ячейке D8 вычислим сумму квадратов отклонений: =СУММ(D3:D8).

D3 fx =($B$11*EXP($B$12*A3)-B3)^2

	A	B	C	D	E	F
1	$y=a*e^{b*x}$					
2	x	y	Ln(y)	Квадраты отклонений		
3	2	4,53	1,510722	0,000		
4	3,5	3,12	1,137833	0,033		
5	5	2,7	0,993252	0,085		
6	6	1,88	0,631272	0,005		
7	7	1,55	0,438255	0,001		
8			Сумма	0,124	=СУММ(D3:D7)	
9						
10			=EXP(ОТРЕЗОК(C3:C7;A3:A7))			
11	a=	6,884				
12	b=	-0,210	=НАКЛОН(C3:C7;A3:A7)			
13						

Пример 1. Решение

<u>Решение в Excel:</u>

7 Для построения теоретической кривой, используя найденные коэффициенты, в ячейку E3 введём формулу =$B$11*EXP($B$12*A3) и скопируем её в ячейки E4:E7.

8 Для предсказания значения Y при X=11 в ячейку A9 введём 11, а в ячейку E11 скопируем полученную формулу.

E3 fx =$B$11*EXP($B$12*A3)

	A	B	C	D	E
1	$y=a*e^{b*x}$				
2	x	y	Ln(y)	Квадраты отклонений	Теоретическая кривая 1
3	2	4,53	1,510722	0,000	4,523
4	3,5	3,12	1,137833	0,033	3,301
5	5	2,7	0,993252	0,085	2,409
6	6	1,88	0,631272	0,005	1,953
7	7	1,55	0,438255	0,001	1,583
8			Сумма	0,124	
9	11				0,683
10					
11	a=	6,884			
12	b=	-0,210			
13					

=EXP(ОТРЕЗОК(C3:C7;A3:A7))

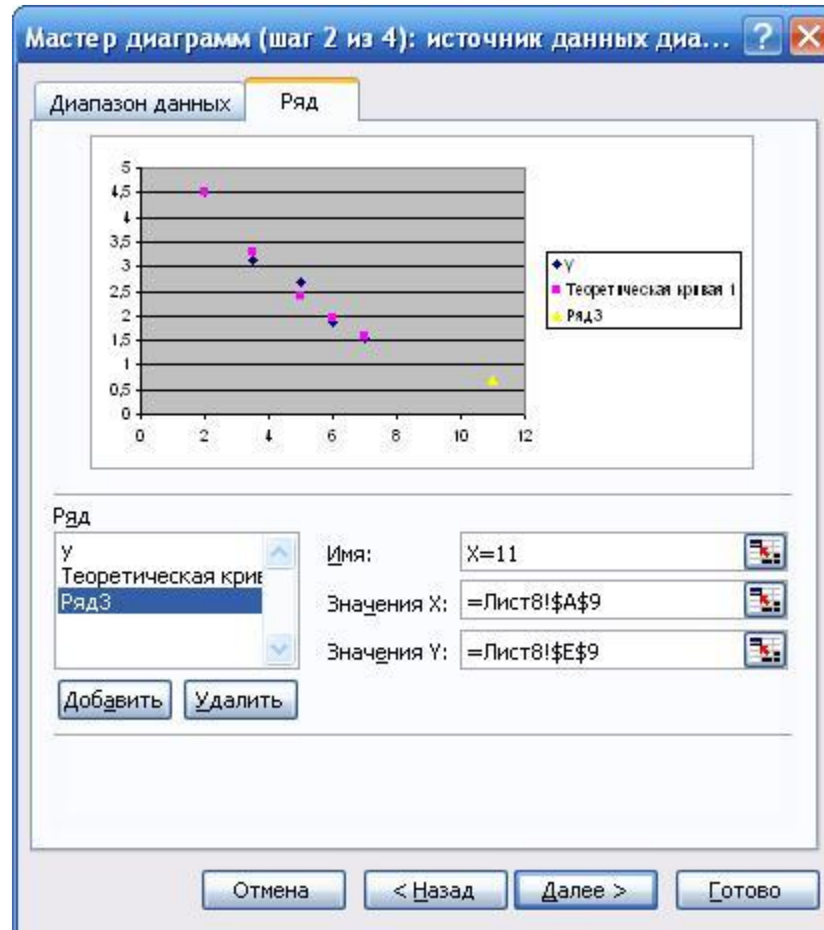
=НАКЛОН(C3:C7;A3:A7)

Пример 1. Решение

Решение в Excel:

9 Выделим диапазоны A2:B7 и E2:E7. С помощью **Мастера диаграмм** построим точечный график.

10 Для добавления на график предсказанного значения Y при X=11 на вкладке **Ряд** щёлкнем по кнопке **Добавить** и заполним соответствующие поля.



11 Щёлкнем по кнопке **Готово**.

12 На полученном графике с помощью форматирования представим теоретическую кривую в виде гладкой линии без маркеров.

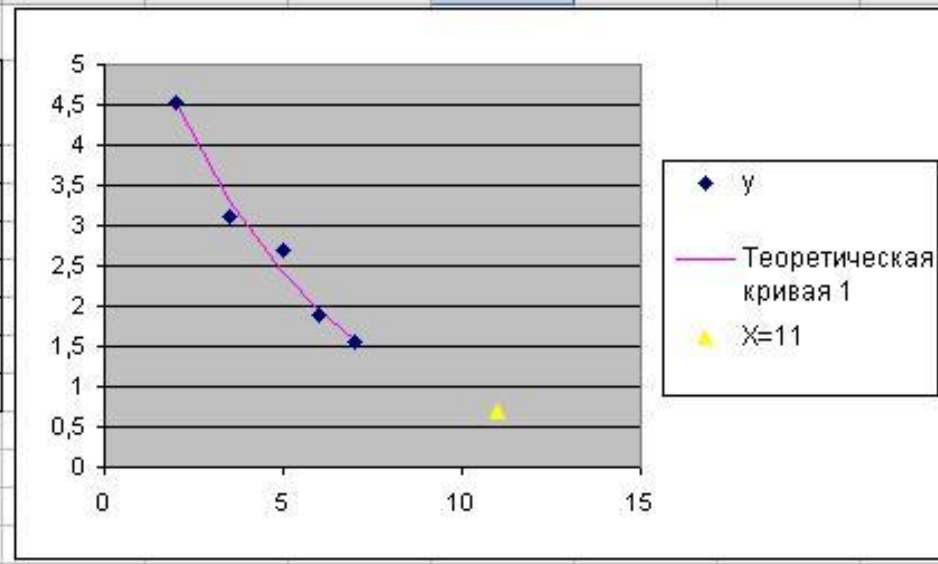
Пример 1. Решение

Результат решения для первой зависимости в Excel:

	A	B	C	D	E
1	$y=a*e^{b*x}$				
2	x	y	Ln(y)	Квадраты отклонений	Теоретическая кривая 1
3	2	4,53	1,510722	0,000	4,523
4	3,5	3,12	1,137833	0,033	3,301
5	5	2,7	0,993252	0,085	2,409
6	6	1,88	0,631272	0,005	1,953
7	7	1,55	0,438255	0,001	1,583
8			Сумма	0,124	
9	11				0,683
10					
11	a=	6,884			
12	b=	-0,210			
13					

=EXP(ОТРЕЗОК(C3:C7;A3:A7))

=НАКЛОН(C3:C7;A3:A7)



Пример 1. Решение

Для второй зависимости:

$$y2(x) = m + k \cdot ln(x)$$

Чтобы свести данную зависимость к линейной выполним замену переменных:

$$z = ln(x).$$

Получим линейную зависимость:

$$y = m + k \cdot z.$$

Далее в Excel с помощью функций НАКЛОН и ОТРЕЗОК найдём коэффициенты *m* и *k*.

Пример 1. Решение

Решение в Excel:

1 В ячейки A16:B21 введём (скопируем) исходные данные.

2 В ячейку C17 введём формулу =LN(A17) и скопируем её в ячейки C18:C21.

3 Для вычисления коэффициента m в ячейку B25 введём формулу =ОТРЕЗОК(B17:B21;C17:C21).

4 Для вычисления коэффициента k в ячейку B26 введём формулу =НАКЛОН(B17:B21;C17:C21).

15	**y=m+k*ln(x)**			
16	x	y	Ln(x)	
17	2	4,53	0,6931	
18	3,5	3,12	1,2528	
19	5	2,7	1,6094	
20	6	1,88	1,7918	
21	7	1,55	1,9459	
22				
23				
24				
25	m=	6,143		=ОТРЕЗОК(B17:B21;C17:C21)
26	k=	-2,322		=НАКЛОН(B17:B21;C17:C21)
27				

Пример 1. Решение

<u>Решение в Excel:</u>

5 Для вычисления квадратов отклонений заданной зависимости от экспериментальных данных в ячейку D17 введём формулу =($B$25+$B$26*LN(A17)-B17)^2
и скопируем её в ячейки D18:D21.

6 В ячейке D22 вычислим сумму квадратов отклонений: =СУММ(D17:D21).

15	y=m+k*ln(x)			
16	x	y	Ln(x)	Квадраты отклонений
17	2	4,53	0,6931	0,000
18	3,5	3,12	1,2528	0,013
19	5	2,7	1,6094	0,087
20	6	1,88	1,7918	0,010
21	7	1,55	1,9459	0,006
22			Сумма	0,116
23				
24				
25	m=	6,143		
26	k=	-2,322		
27				

=($B$25+$B$26*LN(A17)-B17)^2

=СУММ(D17:D21)

=ОТРЕЗОК(B17:B21;C17:C21)

=НАКЛОН(B17:B21;C17:C21)

Пример 1. Решение

Решение в Excel:

7 Для построения теоретической кривой, используя найденные коэффициенты, в ячейку E17 введём формулу =$B$25+$B$26*LN(A17) и скопируем её в ячейки E18:E21.

8 Для предсказания значения Y при X=11 в ячейку A23 введём 11, а в ячейку E23 скопируем полученную формулу.

E17 fx =$B$25+$B$26*LN(A17)

	A	B	C	D	E
15	y=m+k*ln(x)				
16	x	y	Ln(x)	Квадраты отклонений	Теоретическая кривая 2
17	2	4,53	0,6931	0,000	4,533
18	3,5	3,12	1,2528	0,013	3,234
19	5	2,7	1,6094	0,087	2,406
20	6	1,88	1,7918	0,010	1,982
21	7	1,55	1,9459	0,006	1,624
22			Сумма	0,116	
23	11				0,575
24					
25	m=	6,143			
26	k=	-2,322			
27					

=ОТРЕЗОК(B17:B21;C17:C21)

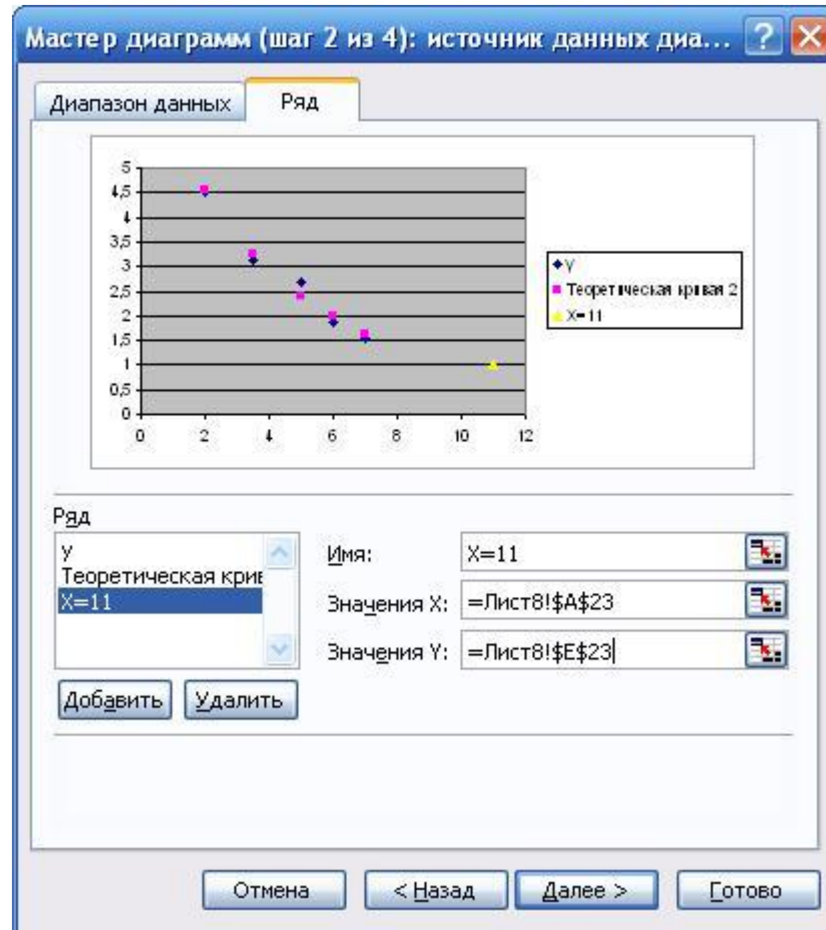
=НАКЛОН(B17:B21;C17:C21)

Пример 1. Решение

<u>Решение в Excel:</u>

9 Выделим диапазоны A16:B21 и E16:E21. С помощью **Мастера диаграмм** построим точечный график.

10 Для добавления на график предсказанного значения Y при X=11 на вкладке **Ряд** щёлкнем по кнопке **Добавить** и заполним соответствующие поля.



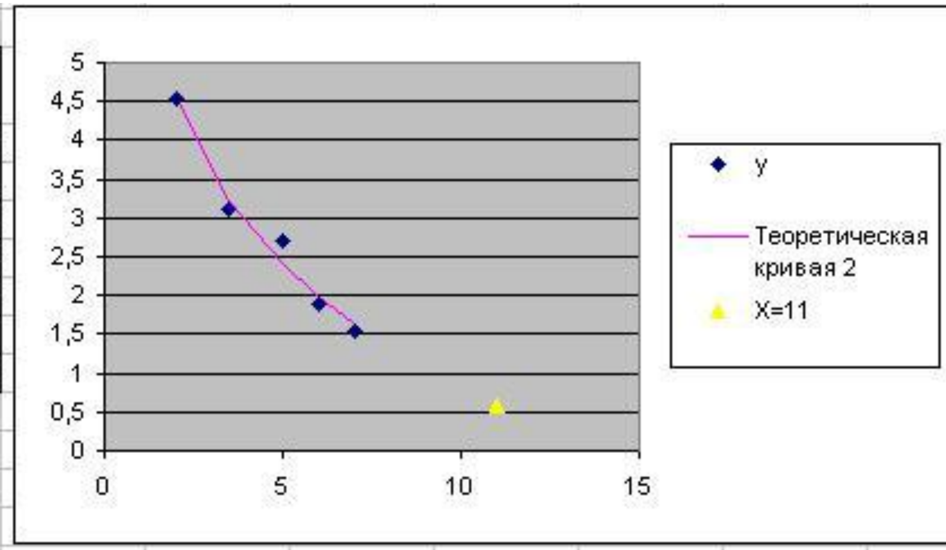
11 Щёлкнем по кнопке **Готово**.

12 На полученном графике с помощью форматирования представим теоретическую кривую в виде гладкой линии без маркеров.

Пример 1. Решение

Результат решения для второй зависимости в Excel:

15	y=m+k*ln(x)				
16	x	y	Ln(x)	Квадраты отклонений	Теоретическая кривая 2
17	2	4,53	0,6931	0,000	4,533
18	3,5	3,12	1,2528	0,013	3,234
19	5	2,7	1,6094	0,087	2,406
20	6	1,88	1,7918	0,010	1,982
21	7	1,55	1,9459	0,006	1,624
22			Сумма	0,116	
23	11				0,575
24					
25	m=	6,143	=ОТРЕЗОК(B17:B21;C17:C21)		
26	k=	-2,322	=НАКЛОН(B17:B21;C17:C21)		
27					



Пример 1. Решение

<u>*Для третьей зависимости:*</u>

$$y3(x) = \frac{c}{x+d}$$

Чтобы свести данную зависимость к линейной перевернём обе части исходной зависимости:

$$\frac{1}{y} = \frac{x+d}{c} \quad \Rightarrow \quad \frac{1}{y} = \frac{x}{c} + \frac{d}{c}$$

и выполним замену переменных:

$$z = \frac{1}{y}, \qquad a = \frac{1}{c}, \qquad b = \frac{d}{c}.$$

В результате получим линейную зависимость:

$$z = a \cdot x + b.$$

Далее в Excel с помощью функций НАКЛОН и ОТРЕЗОК найдём коэффициенты *a* и *b, и затем* вычислим *c* и *d:*

$$c = \frac{1}{a}, \qquad d = b \cdot c\ .$$

Пример 1. Решение

Решение в Excel:

1 В ячейки A30:B35 введём (скопируем) исходные данные.

2 В ячейку C31 введём формулу =1/B31 и скопируем её в ячейки C32:C35.

3 Для вычисления коэффициента c в ячейку B39 введём формулу =1/НАКЛОН(C31:C35;A31:A35).

4 Для вычисления коэффициента d в ячейку B26 введём формулу =ОТРЕЗОК(C31:C35;A31:A35)*B39.

	A	B	C	D	E
29	y=c/(x+d)				
30	x	y	1/y		
31	2	4,53	0,2208		
32	3,5	3,12	0,3205		
33	5	2,7	0,3704		
34	6	1,88	0,5319		
35	7	1,55	0,6452		
36					
37					
38					
39	c=	12,100			
40	d=	0,355			
41					

=1/B31

=1/НАКЛОН(C31:C35;A31:A35)

=ОТРЕЗОК(C31:C35;A31:A35)*B3

Пример 1. Решение

<u>Решение в Excel:</u>

5 Для вычисления квадратов отклонений заданной зависимости от экспериментальных данных в ячейку D31 введём формулу =($B$39/(A31+$B$40)-B31)^2
и скопируем её в ячейки D32:D35.

6 В ячейке D36 вычислим сумму квадратов отклонений: =СУММ(D31:D35).

	A	B	C	D	E	F	G
29	**y=c/(x+d)**						
30	x	y	1/y	Квадраты отклонений			
31	2	4,53	0,2208	0,371			
32	3,5	3,12	0,3205	0,000			
33	5	2,7	0,3704	0,194			
34	6	1,88	0,5319	0,001			
35	7	1,55	0,6452	0,009			
36			Сумма	0,575			
37							
38							
39	c=	12,100					
40	d=	0,355					
41							

=($B$39/(A31+$B$40)-B31)^2

=СУММ(D31:D35)

=1/НАКЛОН(C31:C35;A31:A35)

=ОТРЕЗОК(C31:C35;A31:A35)*B39

Пример 1. Решение

Решение в Excel:

7 Для построения теоретической кривой, используя найденные коэффициенты, в ячейку E31 введём формулу =$B$39/(A31+$B$40) и скопируем её в ячейки E32:E35.

8 Для предсказания значения Y при X=11 в ячейку А37 введём 11, а в ячейку E37 скопируем полученную формулу.

E31 fx =$B$39/(A31+$B$40)

	A	B	C	D	E
29	y=c/(x+d)				
30	x	y	1/y	Квадраты отклонений	Теоретическая кривая 3
31	2	4,53	0,2208	0,371	5,139
32	3,5	3,12	0,3205	0,000	3,139
33	5	2,7	0,3704	0,194	2,260
34	6	1,88	0,5319	0,001	1,904
35	7	1,55	0,6452	0,009	1,645
36			Сумма	0,575	
37	11				1,066
38					
39	c=	12,100			
40	d=	0,355			
41					

=1/НАКЛОН(C31:C35;A31:A35)

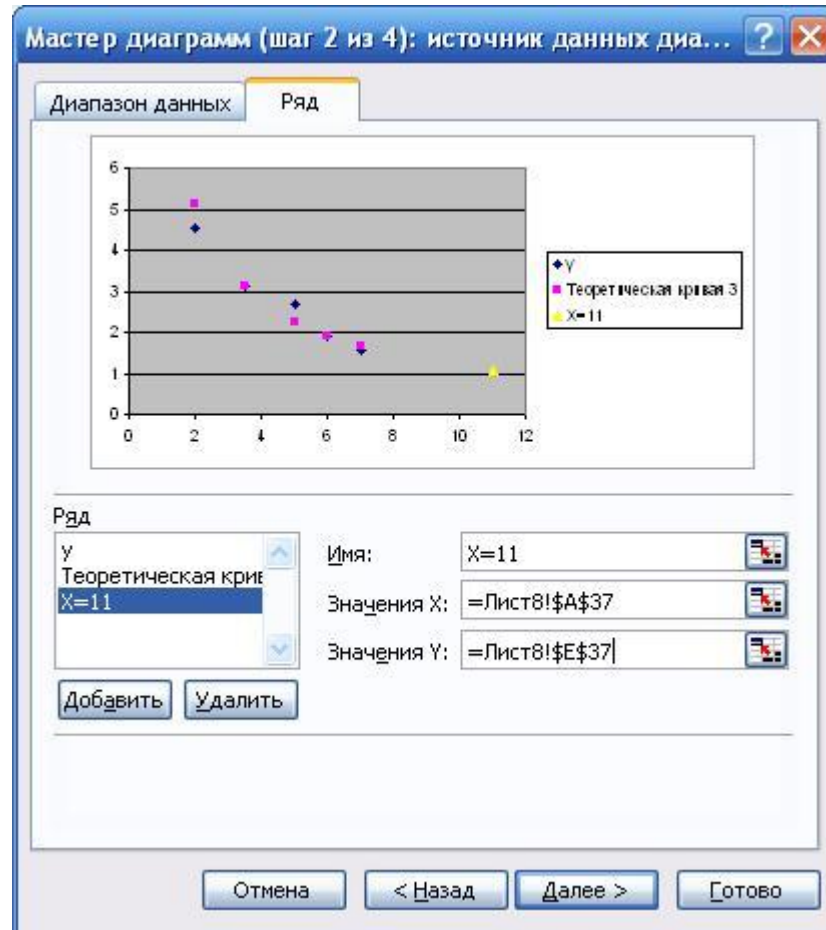
=ОТРЕЗОК(C31:C35;A31:A35)*B39

Пример 1. Решение

<u>Решение в Excel:</u>

9 Выделим диапазоны A30:B35 и E30:E35. С помощью **Мастера диаграмм** построим точечный график.

10 Для добавления на график предсказанного значения Y при X=11 на вкладке **Ряд** щёлкнем по кнопке **Добавить** и заполним соответствующие поля.



11 Щёлкнем по кнопке **Готово**.

12 На полученном графике с помощью форматирования представим теоретическую кривую в виде гладкой линии без маркеров.

Пример 1. Решение

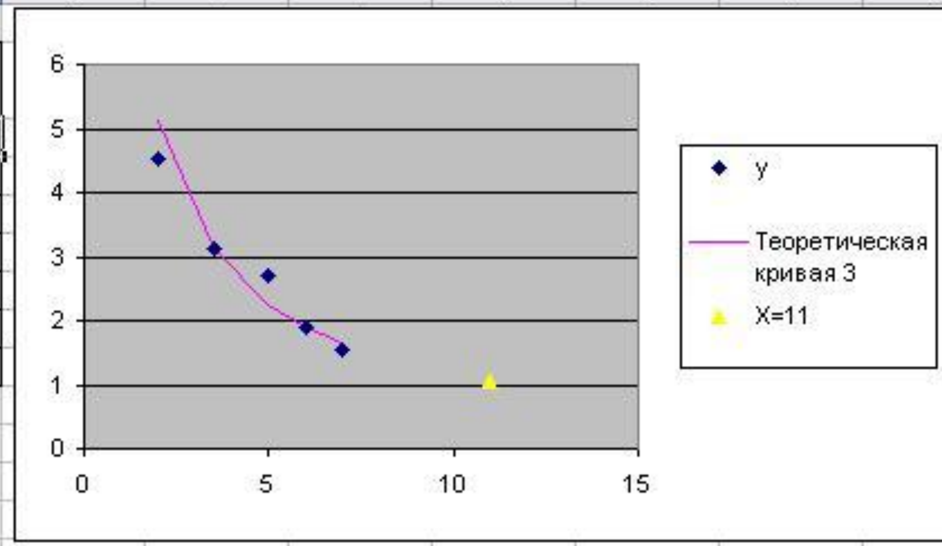
Результат решения для третьей зависимости в Excel:

E31 | f_x =$B$39/(A31+$B$40)

	A	B	C	D	E
29	y=c/(x+d)				
30	x	y	1/y	Квадраты отклонений	Теоретическая кривая 3
31	2	4,53	0,2208	0,371	5,139
32	3,5	3,12	0,3205	0,000	3,139
33	5	2,7	0,3704	0,194	2,260
34	6	1,88	0,5319	0,001	1,904
35	7	1,55	0,6452	0,009	1,645
36			Сумма	0,575	
37	11				1,066
38					
39	c=	12,100			
40	d=	0,355			
41					

=1/НАКЛОН(C31:C35;A31:A35)

=ОТРЕЗОК(C31:C35;A31:A35)*B39



Выбор лучшей аппроксимирующей зависимости

Рассмотрим результаты, полученные в ходе решения задачи.

<u>Суммы квадратов отклонений</u>:

- для первой зависимости 0,124;
- для второй зависимости 0,116;
- для третьей зависимости 0,575.

Лучшей аппроксимирующей зависимостью является та, сумма квадратов отклонений которой от экспериментальных данных является **наименьшей**.

Следовательно, в нашем примере, лучшей является вторая зависимость

$y2(x) = m + k \cdot ln(x)$

Спасибо за внимание!