Математическая и статистическая обработка данных в ЭТ

1. Решение трансцендентных уравнений
2. Решение систем линейных уравнений
3. Метод Монте-Карло
4. Регрессионный анализ

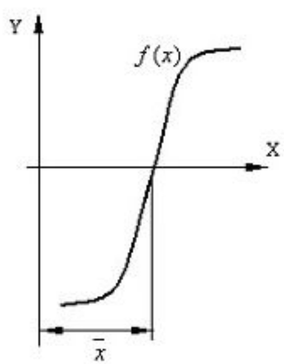
1. Решение трансцендентных уравнений

- Если законы функционирования модели нелинейны, а моделируемые процесс или система обладают одной степенью свободы (т.е. имеют одну независимую переменную), то такая модель, как правило, описывается одним нелинейным уравнением.

Дано нелинейное уравнение:

$$f(x) = 0$$

Необходимо решить это уравнение, т. е. найти его корень $\overline{x}$.



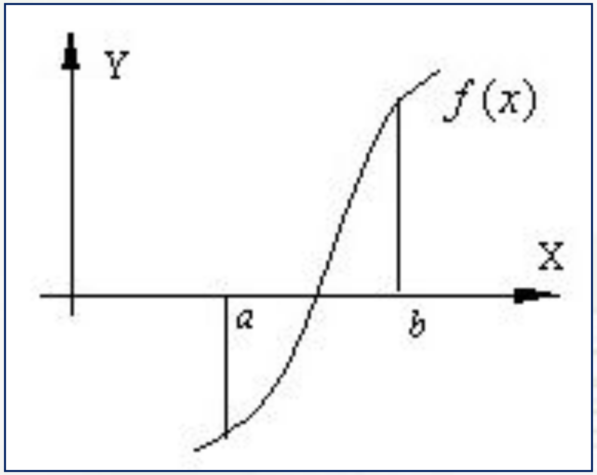
Определение

- Если функция f(x) включает в себя тригонометрические или экспоненциальные функции от некоторого аргумента x, то уравнение f(x)=0 называется трансцендентным уравнением.
  - например уравнения:
    - $\cos x = x$
    - $\log x = x - 5$
    - $2^x = \log x + x^5 + 40$
- Большинство трансцендентных уравнений не может быть решено точно. Такие уравнения обычно имеют бесконечное множество решений.

В общем случае решение уравнения находится в следующей последовательности:

- Этапы
  1. отделение (локализация) корня;
  2. приближённое вычисление корня до заданной точности.
- Задачу отыскания корней уравнения можно считать решенной, если найти корни уравнения с заданной степенью точности. Для этого используются приближенные (численные) методы решения.
  – точность нахождения корня (Сервис → Параметры → Вкладка Вычисления).

- Большинство употребляющихся приближенных методов решения уравнений являются, по существу, способами уточнения корней. Для их применения необходимо знание интервала изоляции [a,b], в котором лежит уточняемый корень уравнения



Решение уравнения: уточнение корней, т.е. сужение интервала [a,b] до величины равной заданной степени точности ε.

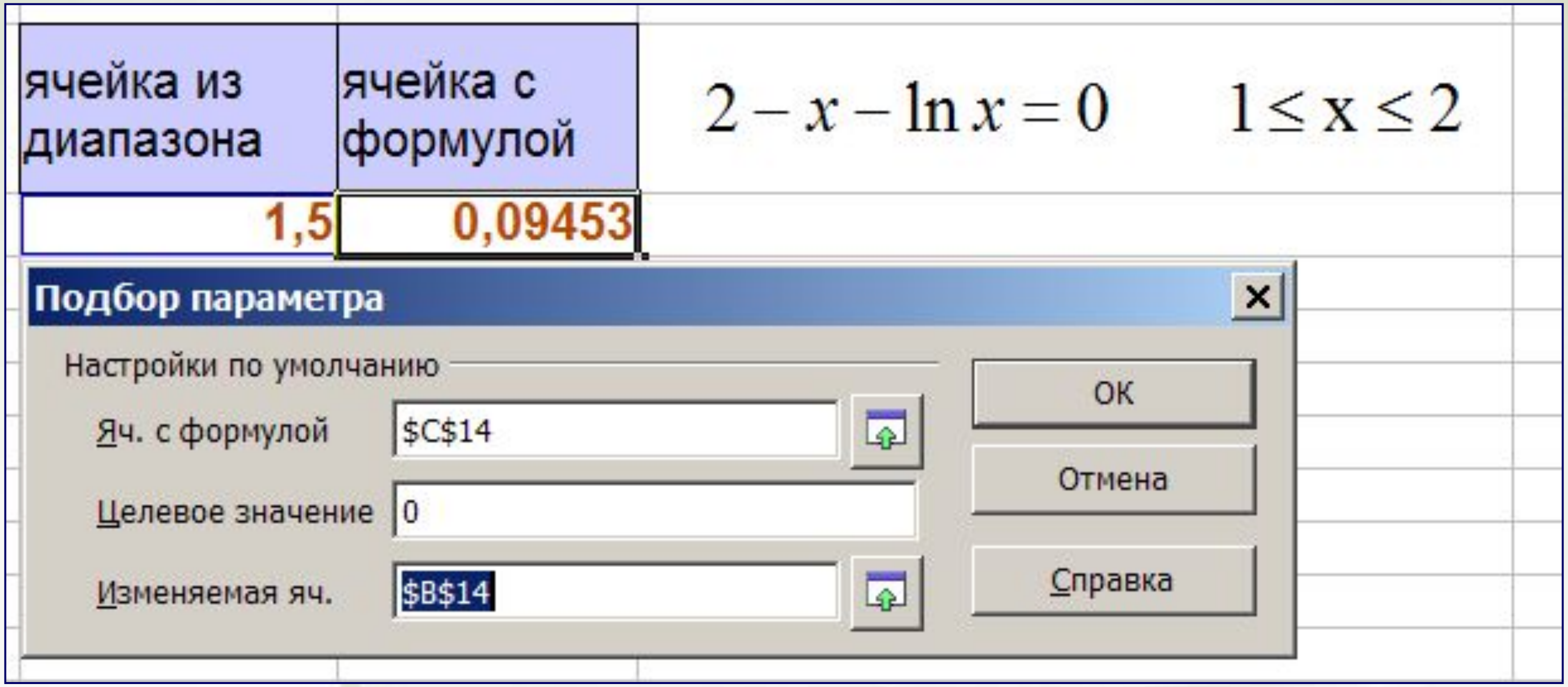
Для трансцендентных уравнений пригодны следующие методы уточнения приближенных значений корней:

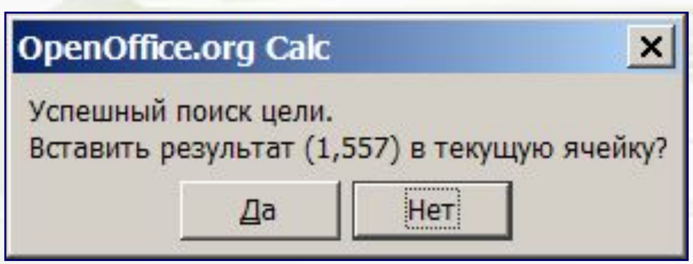
- Метод подбора параметра;
- метод простых итераций;
- графический метод;
- метод половинного деления (метод дихотомии);
- метод Ньютона (метод касательных);
- метод хорд.

1. Подбор параметра

- При **подборе параметра** OOo Calc изменяет значение в одной конкретной ячейке до тех пор, пока формула, зависимая от этой ячейки, не возвращает нужный результат.
  - **СЕРВИС\ПОДБОР ПАРАМЕТРА**.
    - **Ячейка с формулой** – уравнение.
    - **Целевое значение –** результат уравнения.
    - **Изменяемая ячейка** – ссылка на ячейку, значение которой нужно подобрать (из диапазона).
    - **ОК**.

Пример:


ячейка из диапазона
ячейка с формулой
$2 - x - \ln x = 0 \quad 1 \le x \le 2$
1,5
0,09453
Подбор параметра
Настройки по умолчанию
Яч. с формулой
C14
Целевое значение
0
Изменяемая яч.
B14
ОК
Отмена
Справка


OpenOffice.org Calc
Успешный поиск цели.
Вставить результат (1,557) в текущую ячейку?
Да
Нет

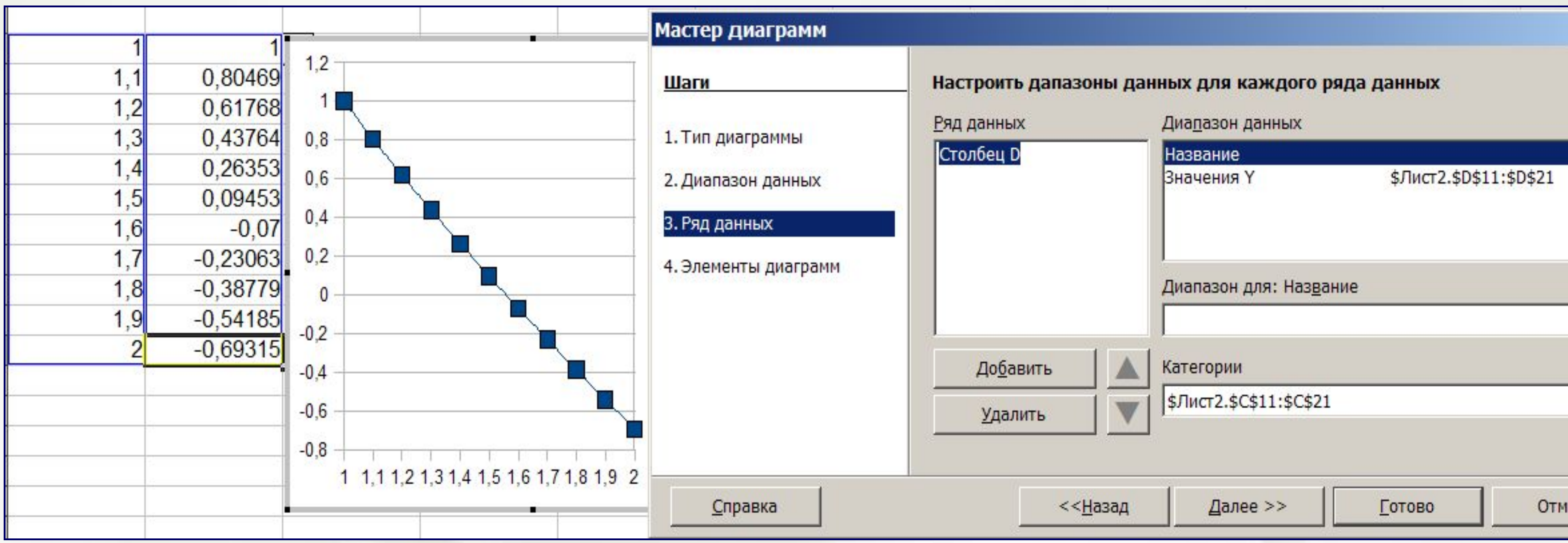

ячейка из диапазона
ячейка с формулой
1,557
0,00024

2. Графический метод отделения корней (наиболее наглядный)

- Для простой функции
- строится график функции y=f(x), и определяются абсциссы точек пересечения этого графика с осью *OX*, которые и являются корнями уравнения f(x)=0
- Для построения графика необходимо построить **таблицу значений**, аргумент которой изменяется с фиксированным шагом.
- Шаг выбирают небольшим, и используя **Мастер диаграмм** строится график.

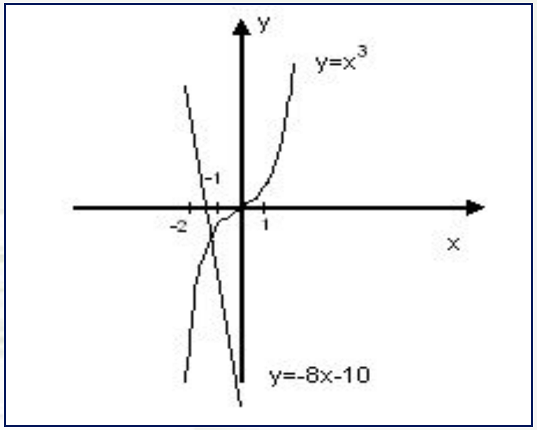
Пример

- Строим **таблицу значений**
- Шаг выбирают небольшим (0,1), строится график.
- Пересечение с осью *OX* – есть решение уравнения



Для сложной функции

- Если $f(x)$ – сложная, то ее надо представить в виде $\mathbf{f(x)=\phi^1(x) - \phi^2(x)}$ так, чтобы легко строились графики функций $\mathbf{y=\phi^1(x)}$ и $\mathbf{y=\phi^2(x)}$
- Абсциссы точек пересечения графиков и будут корнями уравнения
  - **Пример:** $x^3+8x+10=0$
  - Представим левую часть уравнения в виде $f(x)=\phi^1(x) - \phi^2(x)$
  - Получим: Построим графики функций $y=x^3$ и $y= -8-10x$.



Корень уравнения на отрезке [–2;1]

Метод итераций
(метод последовательных приближений)

- Указанный интервал (отрезок) делится на части. Процесс деления отрезка для нахождения корней уравнения продолжается до указанной точности **ε**
- Среди всех интервалов, выбирается тот, при котором значение "*у*" меняет знак с **"+" на "–"** (пересечение с осью *OX*).
- Ближайшее к 0 число – есть решение уравнения

Матрицы

- Значительная часть математических моделей различных объектов и процессов записывается в простой и компактной матричной форме.
- При решений систем уравнений используются матрицы и арифметические действиями над ними.
- **Матрицей** размера $m \times n$ называется прямоугольная таблица чисел, содержащая m – строк и n – столбцов.

Мастер функций → категория Массив

- **Массив** – набор ячеек или значений, которые обрабатываются как одна группа.
- **Формула массива** действует для нескольких ячеек и является не копией, а общей формулой для всех ячеек матрицы.
- Формула массива вставляется как формула матрицы и обозначается фигурными скобками {}.

Операции над матрицами

- **TRANSPOSE () –** транспонирование (столбцы исходной матрицы заменяются строками с соответствующими номерами)
- **MDETERM () –** вычисление определителя (число, вычисляемое на основе значений элементов массива).

Умножение матриц

- Произведение матриц определено, если число столбцов первой матрицы равно числу строк второй.

Перемножение матриц осуществляется по правилу:

$$C = A * B = \begin{pmatrix} 1\,\text{стр} * 1\,\text{стб} & 1\,\text{стр} * 2\,\text{стб} & \ldots. & 1\,\text{стр} * p\,\text{стб} \\ 2\,\text{стр} * 1\,\text{стб} & 2\,\text{стр} * 2\,\text{стб} & \ldots. & 2\,\text{стр} * p\,\text{стб} \\ \ldots\ldots\ldots\ldots & \ldots\ldots\ldots\ldots & \ldots.. & \ldots\ldots\ldots\ldots \\ m\,\text{стр} * 1\,\text{стб} & m\,\text{стр} * 2\,\text{стб} & \ldots. & m\,\text{стр} * p\,\text{стб} \end{pmatrix}.$$

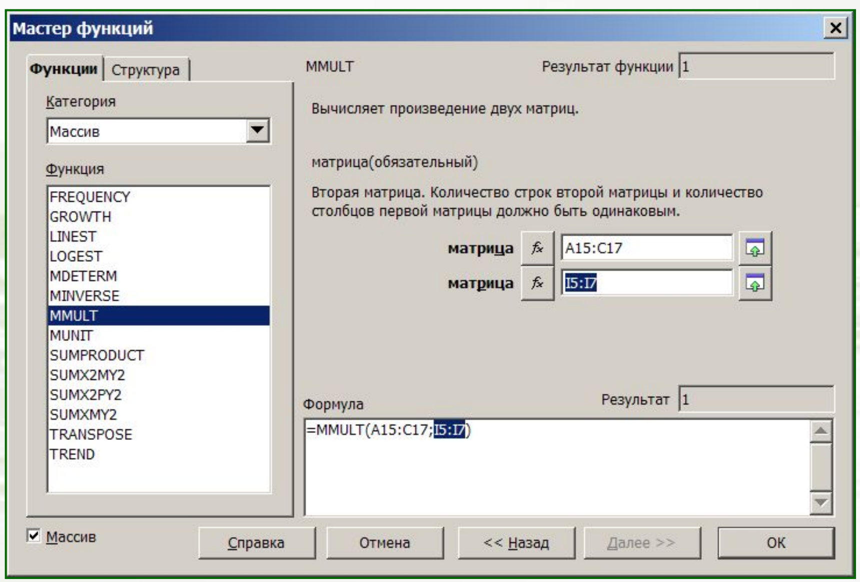
Пусть, например,

$$C = A * B = \begin{pmatrix} 1 & 3 & 4 & 2 \\ 3 & 2 & 0 & -1 \\ 0 & 1 & -1 & 2 \end{pmatrix} * \begin{pmatrix} 1 & 3 \\ 2 & 2 \\ 10 & 0 \\ 12 & -1 \end{pmatrix} =$$

$$= \begin{pmatrix} 1*1+3*2+4*10+2*12 & 1*3+3*2+4*0+2*(-1) \\ 3*1+2*2+0*10-1*12 & 3*3+2*2+0*0-1*(-1) \\ 0*1+1*2-1*10+2*12 & 0*3+1*2-1*0+2*(-1) \end{pmatrix} = \begin{pmatrix} 71 & 7 \\ -5 & 14 \\ 16 & -1 \end{pmatrix}.$$

Умножение матриц в ЭТ

- **Последовательность действий:**
  - Ввод 2 матриц;
  - <u>Выделение блока результатов</u>;
  - Ввести **= MMULT**(массив1; массив2);
  - ОК
  - Результаты появляются в выделенном блоке



Нахождение обратной матрицы MINVERSE ()

- Необходимым и достаточным условиям существования обратной матрицы является невырожденность исходной матрицы
  - Матрица называется **невырожденной** или **регулярной**, если ее определитель отличен от 0 ($|A| \neq 0$);
  - в противном случае (при $|A|=0$) матрица называется **вырожденной** или **сингулярной**
- Последовательность действий аналогична умножению матриц

Дополнительные операции:

- **INDEX ()** – извлечение элемента по номеру строки и столбца
- **ROWS ()** – определение числа строк
- **COLUMNS ()** – определение числа столбцов
- **Сложение (вычитание)** матриц можно производить только одного размера. Все элементы складываются или вычитаются поэлементно.

$a_{ij} b_{ij} (i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, n)$

Система n линейных уравнений с n неизвестными.

- Многие задачи в технике, экономике и других областях сводятся к решению системы линейных уравнений.
- Пусть дана система n уравнений с n неизвестными, где $a_{ij}\, b_{ij}$ (i=1,2,…,n; j=1,2,…,n) – произвольные числа, называемые, соответственно, **коэффициентами** при переменных и **сводными членами уравнений**.

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

Решение системы

- Ввести матрицу коэффициентов (под значением *x*),
- ввести матрицу свободных членов.

- Вычислить определитель **MDETERM ()**.
- Получить обратную матрицу (выделить диапазон ячеек, столько же, сколько в матрице свободных членов. Функция **MINVERSE ()**.
- Решить систему. Выделить 3 ячейки, т.к. неизвестных значения 3 – для x1,x2,x3. Обратная матрица × матрицу свободных членов. Функция **MMULT ().**

Проверка

Дана система уравнений

$$\begin{cases} 2x_1 - x_2 + x_3 = 3 \\ x_1 + 3x_2 - 2x_3 = 1 \\ x_2 + 2x_3 = 8 \end{cases}$$

матрица коэффициентов

2	-1	1
1	3	-2
0	1	2

матрица свободных членов

3
1
8

19 матрица - регулярная, т.к. определитель не равен 0

Обратная матрица

0,421	0,158	-0,05
-0,11	0,211	0,263
0,053	-0,11	0,368

обратная матрица для матрицы коэффициентов

Решение СЛАУ

1	обратную матрицу
2	перемножаем на матрицу
3	свободных членов

1. Подстановка найденных неизвестных в уравнение.
2. Функция **MMULT ()**. Умножение матрицы коэффициентов на полученную матрицу неизвестных (решение), предварительно выделив 3 ячейки. Получаем матрицу свободных членов.

Решить задачу:

- Ресторан специализируется на выпуске трех видов фирменных блюд: **B1, B2, B3,** при этом используются ингредиенты трех типов **S1, S2, S3**. Нормы расхода каждого из них на одно блюдо и объем расхода ингредиентов на 1 день заданы таблицей:
- Найти ежедневный объем выпуска.

Ингредиент	Нормы расхода ингредиентов на одно блюдо (у.е.)			Расход ингредиентов на 1 день (у.е.)
	B1	B2	B3	
S1	5	3	4	2700
S2	2	1	1	900
S3	3	2	2	1600

Приближенное вычисление определенных интегралов

- С помощью нахождения первообразных можно вычислить интегралы для довольно незначительного класса функций, поэтому возникает необходимость в приближенных методах вычисления интегралов.

- определенный интеграл $\boxed{\int_a^b f(x)\,dx}$

где f(x) непрерывная на [a, b] функция.

Простые способы приближенного вычисления

- формула прямоугольников,
- формула трапеций,
- формула Симпсона или параболическое интегрирование,
- метод Монте-Карло.

Метод Монте-Карло

- **Метод статистических испытаний**, численный метод решения задач при помощи моделирования случайных процессов и событий.
- Название метод получил от **г. Монте-Карло** в Монако. Этот метод требует применения случайных чисел.

Пример:

$$\int_{0}^{0,5} \frac{\arcsin^5 x\,dx}{\sqrt{1-x^2}}$$

- Для вычисления интеграла: используется функция **RAND** () – возвращает дробное случайное число от 0 до 1.

Для получения случайного числа между a и b, используется формула:

=RAND() * (верхний – нижний) + нижний
=RAND() * (b – a) + a

После вычисления случайных чисел вводится значения подынтегральной функции для этих чисел.

Далее:

Среднее значение подынтегральной функции:

AVERAGE (1000 значений подынтегральной функции**)**

Значение интеграла считается по формуле:

= Average * (b – a),

- где a - нижний предел интегрирования,
- b - верхний предел

$$\int_{0}^{0,5} \frac{\arcsin^5 x \, dx}{\sqrt{1 - x^2}}$$

Как зафиксировать значение?

- Скопировать значение интеграла (сл. число)
- Установить курсор в другой ячейке
- Зайти в Верхнее меню:
  - Правка → Вставить как
  - Убрать флажок Вставить все – Установить флажок Числа.

Использование:

- Расчеты по методу Монте-Карло **сравнительно просты** и не требуют **большой оперативной памяти**.
- **Используется** для построения и изучения моделей (живых и неживых систем, инженерных конструкций, разнообразных процессов – экономических, химических, социальных) и т.д.

Примеры вывода случайных чисел:

- Для вывода случайного числа, используя функцию: =**RAND ()** → 0,12345 (от 0 до 1)

Интервал (60;200)

=rand()*(60-200)+200 или

=rand()*(200-60)+60

Результат: 65,2356

165,123

Функция Round()

Округляет число до указанного количества десятичных разрядов

<u>Синтаксис</u>

Round(число; число разрядов)

Число – округляемое число

Число разрядов – количество десятичных разрядов, до которого нужно округлить число

Пример: round (14,2356; 0) **результат** 14

round ($\underbrace{\text{rand()*(15-10)+10}}_{\textbf{число}}$; 1) 12,1

разряд

Функция RANDBETWEEN ( )

- Возвращает случайное число между двумя заданными числами. При каждом действие на рабочем листе возвращается новое случайное число.
- Если функция недоступна или возвращает ошибку #ИМЯ?, нужно установить надстройку анализа.

Решение задач аппроксимации средствами ЭТ

- Регрессионный анализ

- На практике при ***моделировании*** различных ***процессов*** (экономических, технических, социальных) используются различные способы вычисления приближенных значений функций по известным значениям в некоторых фиксированных точках.

Применение

- **Такого рода задачи приближения функций часто возникают:**
  - при оценке развития процесса;
  - изучении взаимосвязи переменных значений, полученных в результате эксперимента;
  - прогнозировании некоторых показателей …
- Для решения задач данного класса применяются математические методы (метод корреляционно-регрессионного анализа).

- Если для моделирования процесса, заданного таблицей, построить **функцию**, приближенно описывающую данный процесс на основе метода наименьших квадратов, она будет называться **аппроксимирующей функцией (регрессией)**, а сама **задача** построения этих функций - **задачей аппроксимации.**

- **Важным направлением** в изучении закономерностей динамики инженерно-экономических процессов является исследование тенденции развития (тренда).

В основе составления тренда лежит использование метода регрессионного анализа, который позволяет подобрать аналитическую функцию, максимально точно описывающую изменение уровня динамики во времени.

Регрессионный анализ

- Форма **статистического анализа**, используемого в основном для прогнозов.
- Регрессионный анализ позволяет оценить степень связи между переменными, предлагая механизм вычисления предлагаемого значения переменной из нескольких уже известных значений.

В Calc для построения регрессий имеются две возможности:

1. Добавление выбранных регрессий (**линий тренда**) в диаграмму, построенную на основе таблицы данных (наличие построенной диаграммы);

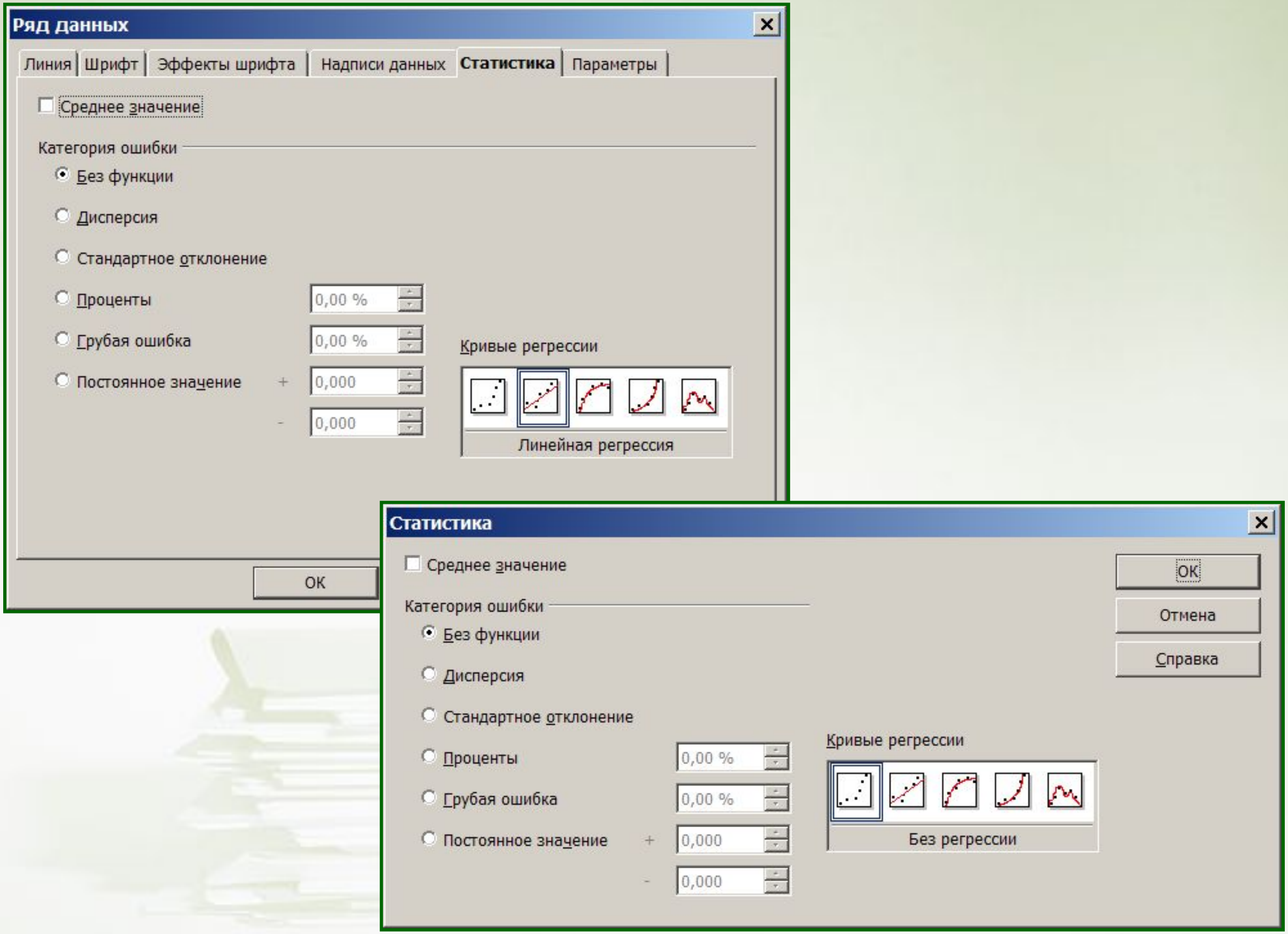
2. Использование встроенных статистических функций Calc, позволяющих получать регрессии (линии тренда) непосредственно на основе таблицы исходных данных.

ТРЕНД

- **Тренд** (кривая регрессии) – это функция заданного вида , с помощью которой можно аппроксимировать построенный по данным таблицы график.
- Тренд служит для выявления тенденций развития процесса, представленного в виде диаграммы, и обеспечивает прогноз на заданный период .

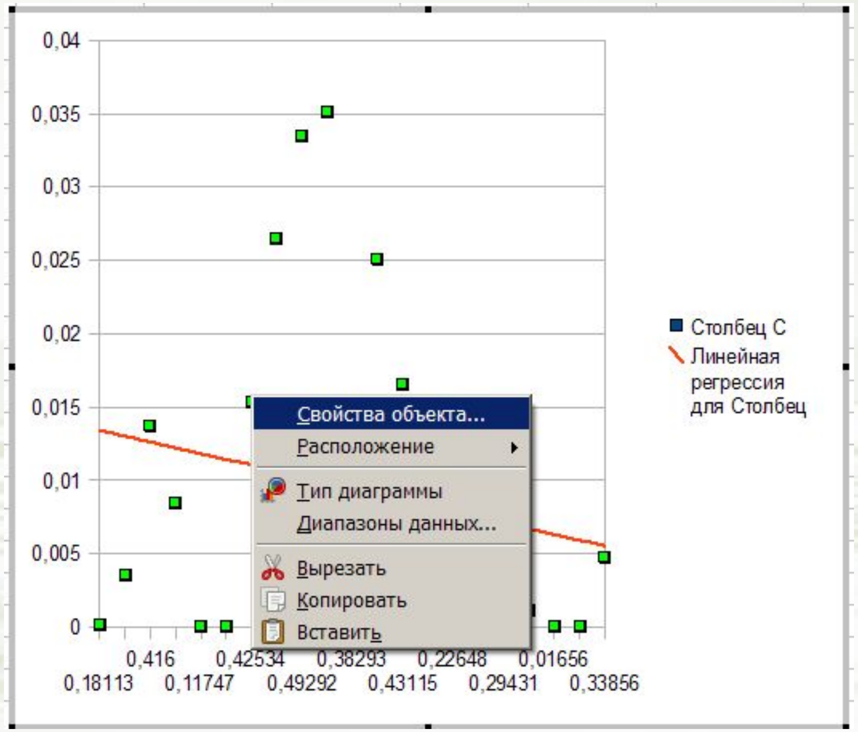
Добавление линии тренда

- Необходимо построить **<u>точечный</u>** график для экспериментальных данных.
- Выделить построенную диаграмму и в контекстном меню выбрать Свойства объекта → вкладка Статистика → Кривые регрессии (**ЛИНИЯ ТРЕНДА)** или вызвать меню Вставка → Статистика.


Ряд данных
Линия
Шрифт
Эффекты шрифта
Надписи данных
Статистика
Параметры
Среднее значение
Категория ошибки
Без функции
Дисперсия
Стандартное отклонение
Проценты
Грубая ошибка
Постоянное значение
0,00 %
0,00 %
0,000
0,000
Кривые регрессии
Линейная регрессия
OK
Статистика
Среднее значение
Категория ошибки
Без функции
Дисперсия
Стандартное отклонение
Проценты
Грубая ошибка
Постоянное значение
0,00 %
0,00 %
0,000
0,000
Кривые регрессии
Без регрессии
OK
Отмена
Справка

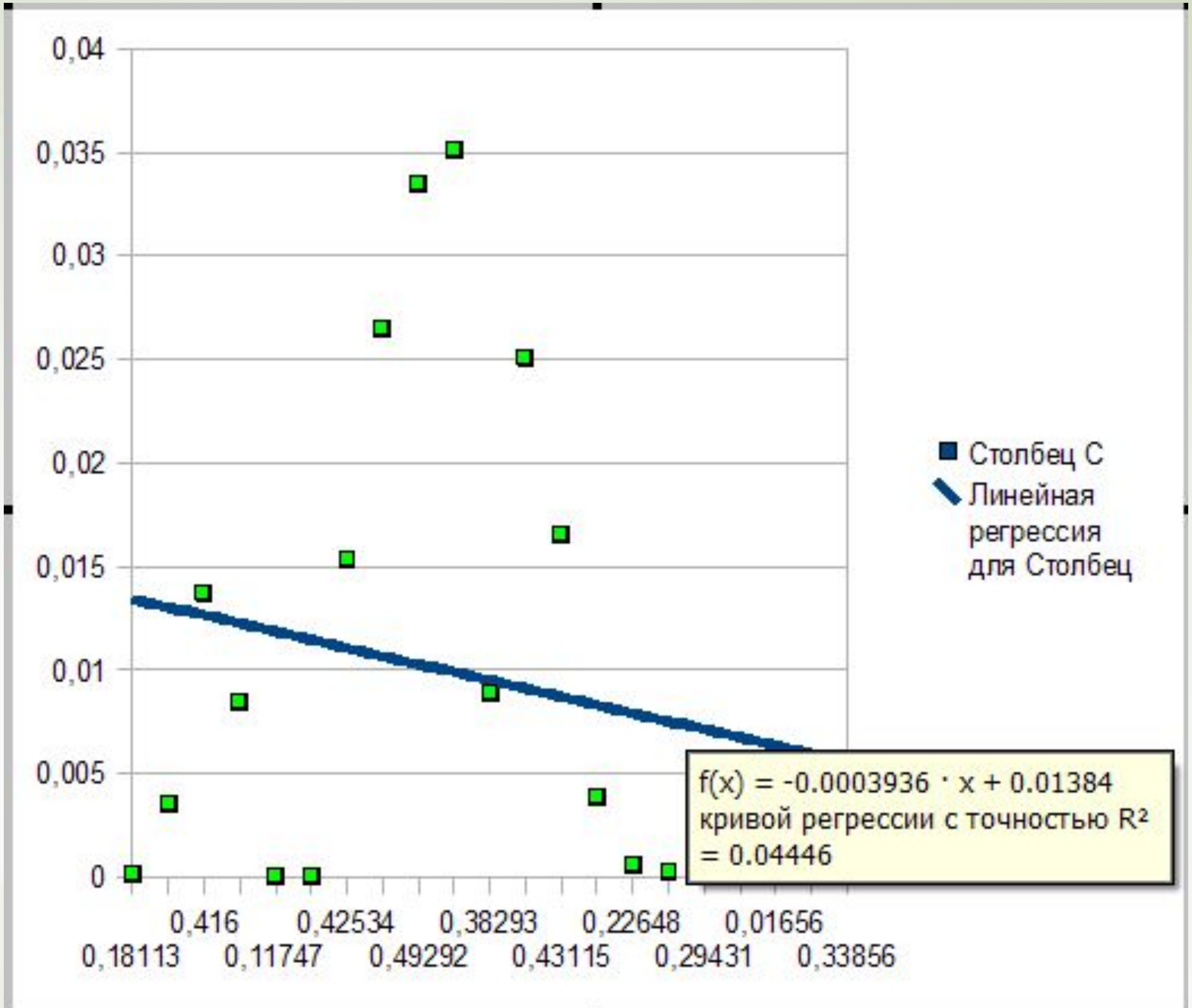
1. Стандартные типы тренда

Существует 4 различных видов линий тренда, которые могут быть добавлены на диаграмму. Способ следует выбирать в зависимости от типа данных.



Точность аппроксимации

- Близость значений линии тренда к фактическим данным – **R^2 (коэффициент корреляции, аппроксимации).**
- Если значение **R^2** = или близко к 1 линия тренда наиболее соответствует действительности. При построении линии тренда значение **R^2** рассчитывается автоматически.
- Полученный результат можно вывести на диаграмме (Справка → Что это такое?).



0,04
0,035
0,03
0,025
0,02
0,015
0,01
0,005
0
0,416
0,42534
0,38293
0,22648
0,01656
0,18113
0,11747
0,49292
0,43115
0,29431
0,33856
Столбец C
Линейная регрессия для Столбец
f(x) = -0.0003936 · x + 0.01384
кривой регрессии с точностью R²
= 0.04446

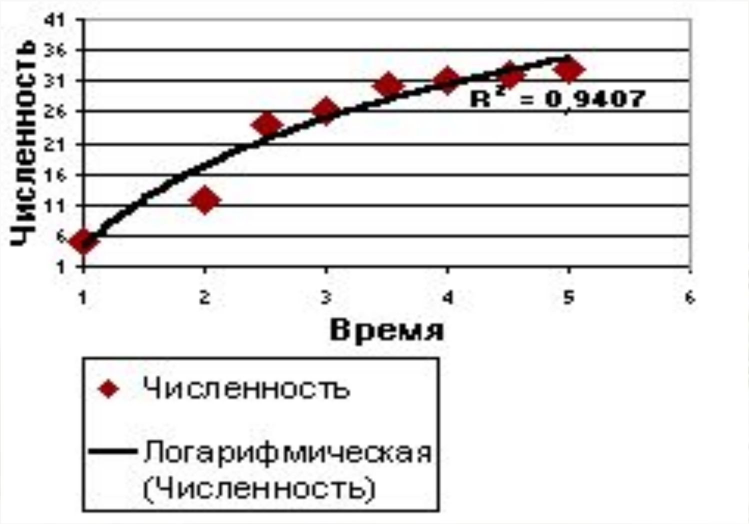
Линейная аппроксимация

- Это прямая линия, наилучшим образом описывающая набор данных. Применяется в самых простых случаях, когда точки данных расположены близко к прямой.
  - *линейная аппроксимация хороша для величины, которая увеличивается или убывает с постоянной скоростью*

$$y = mx + b$$

Логарифмическая аппроксимация

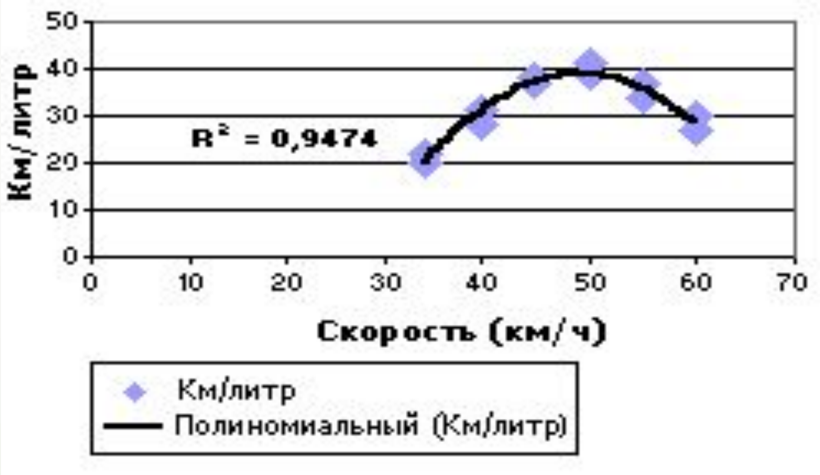
- Описывает величину, которая вначале быстро растет или убывает, а затем постепенно стабилизируется.
  – использует как "–", так и "+" величины



$$y = c\ln(x) + b$$

Полиномиальная аппроксимация

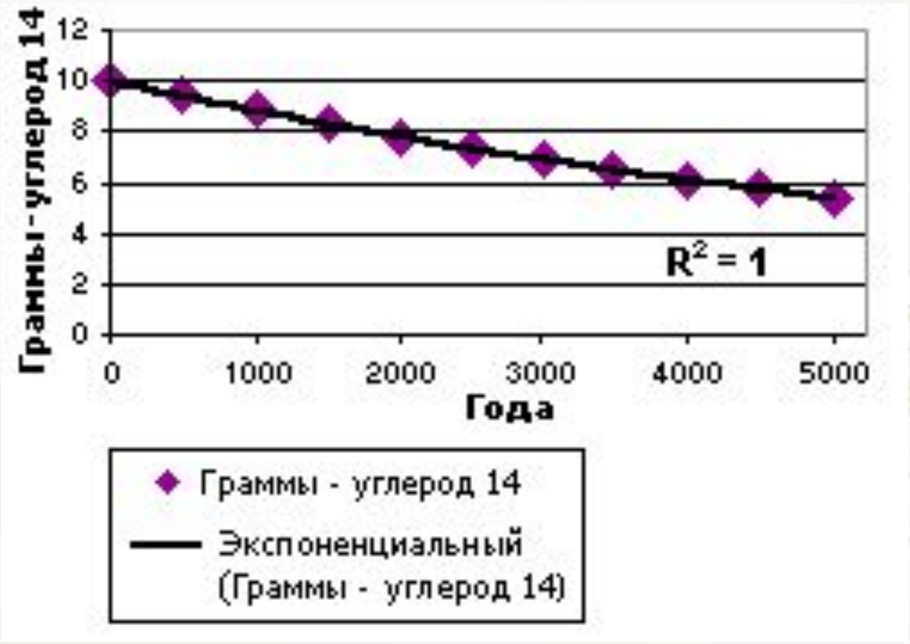
- Описывает величины, попеременно возрастающие и убывающие. Полезна для анализа большого набора данных о нестабильной величине.
  - *Степень полинома определяется количеством экстремумов (максимумов и минимумов) кривой.*



$$y = c_0 + c_1 x + c_2 x^2 + c_3 x^3 \ldots$$

Экспоненциальная аппроксимация

- Полезна в том случае, если скорость изменения данных непрерывно возрастает.
  - *При 0 и «–» этот вид приближения неприменим.*



$$y = ce^{bx}$$

2. Встроенные функции для построения регрессий

- Используется для построения линий тренда вне области диаграммы.
- Эти функции позволяют строить лишь линейные или экспоненциальные регрессии.

Функции для построения линейной регрессии

- **Linest** – рассчитывает статистику для ряда с применением метода наименьших квадратов, чтобы вычислить прямую линию, которая наилучшим образом аппроксимирует имеющиеся данные.
- Данная функция вычисляет параметры линейной регрессии в виде массива

Функция Linest

Уравнение для прямой линии имеет следующий вид: **y = mx + b**

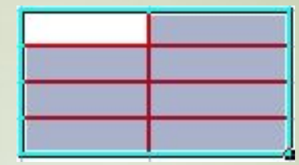
y – функция независимого значения x

x – независимая переменная

m – тангенс угла наклона линейной регрессии к оси абсцисс

b – координата точки пересечения линейной регрессии с осью ординат

Чтобы получить регрессионную статистику при помощи функции Linest



1. *Выделить 8 ячеек в 2 ряда*
2. *Применить функцию Linest*
3. *Ctrl+Shift+Enter*

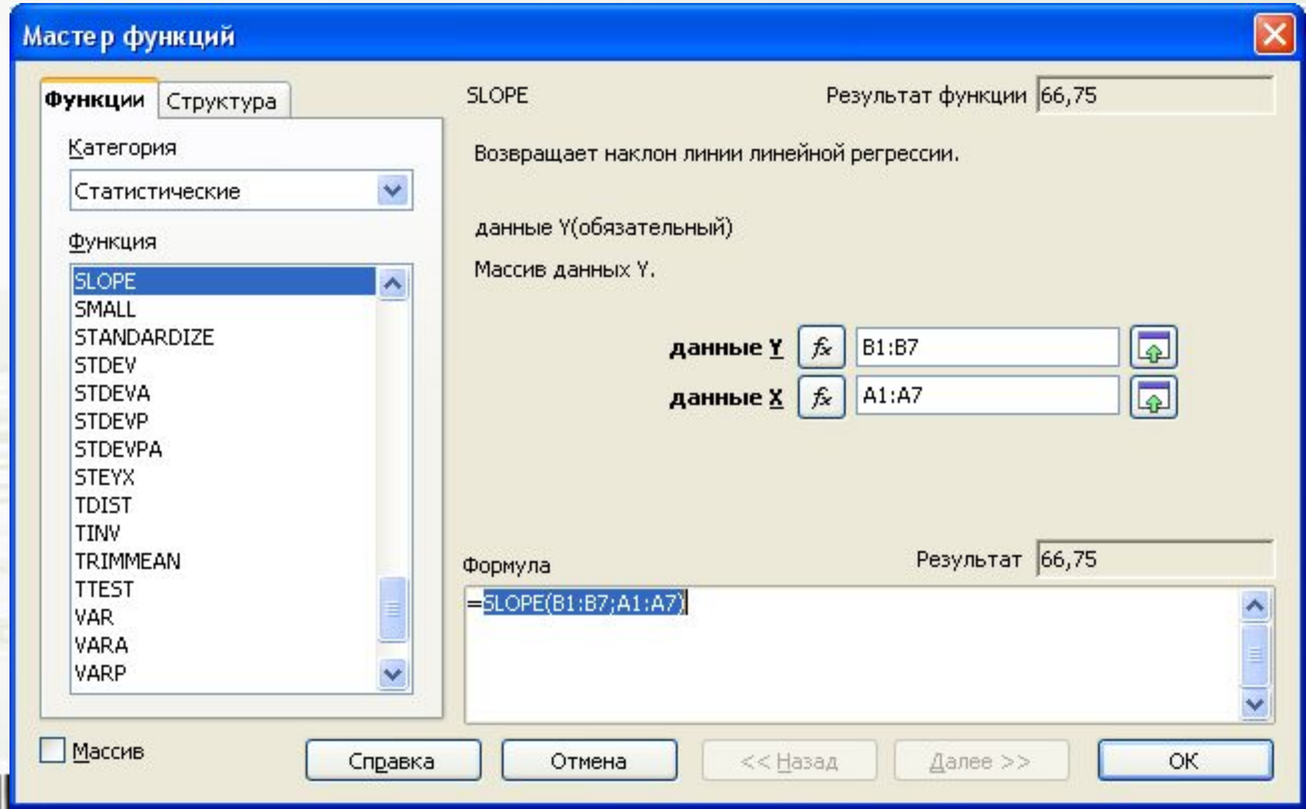
   *В **1 ячейке** получим значение **m**,*

   ***2 ячейка** – **b**,*

   ***5 ячейка** – коэффициент корреляции (**R^2**).*

4. *Подставить в уравнение данные, растянуть на диапазон ячеек.*

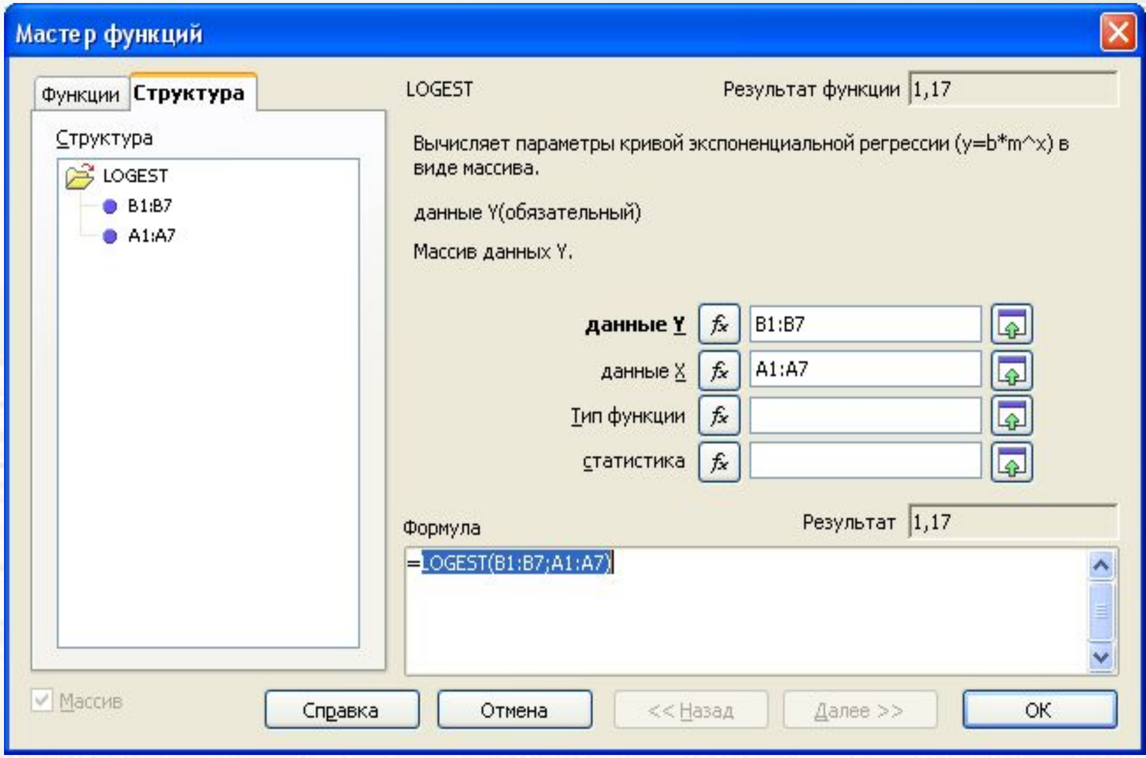
Кроме функции **Linest** для получения параметров линейной регрессии в программе Calc можно использовать функции **Slope** (возвращает наклон линии линейной регрессии) **Intersept** (вычисляет отрезок, отсекаемый линией линейной регрессии).



Функция Logest

- Рассчитывает прогнозируемый

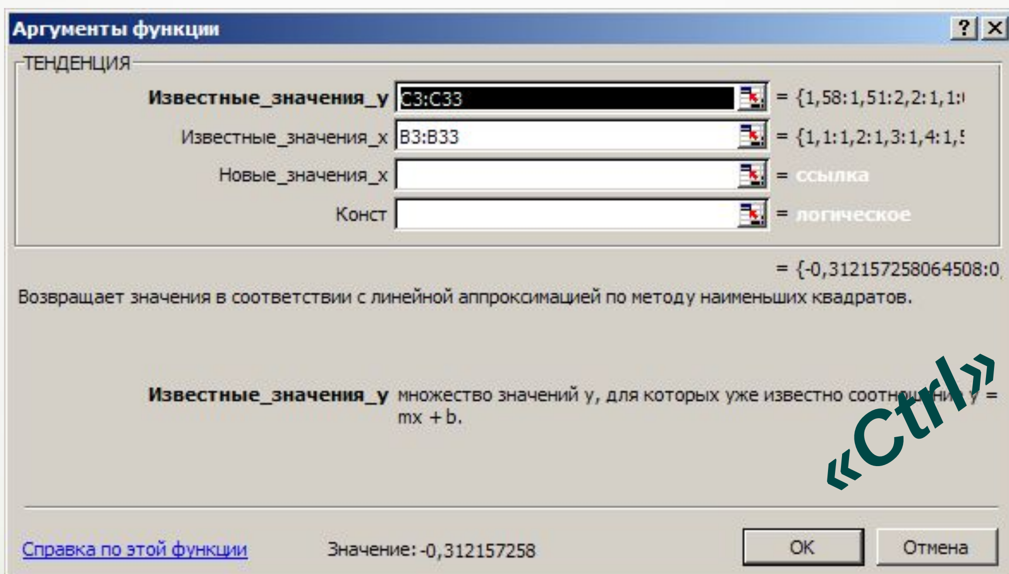
экспоненциальный рост на основании имеющихся данных. Приемы построения регрессий с помощью функций Linest, Slope, Intersept, Logest практически совпадают.



Функция Trend

Возвращает значения в соответствии с линейным трендом.

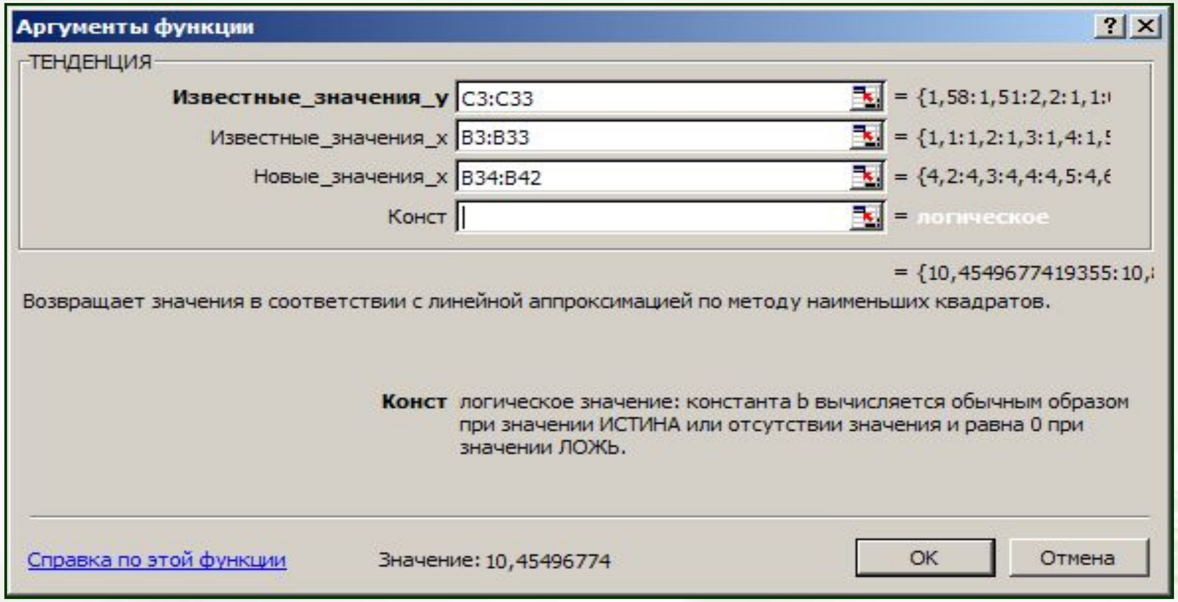
- Выделить диапазон ячеек, где будут располагаться значения. Вызывать функцию **Trend.**



«Ctrl» + «Shift» + «Enter».

Получение прогноза при помощи функции TREND

- Выделить диапазон ячеек, для значений, прогнозируемых функцией Trend.
- ***Вызвать функцию Trend.***



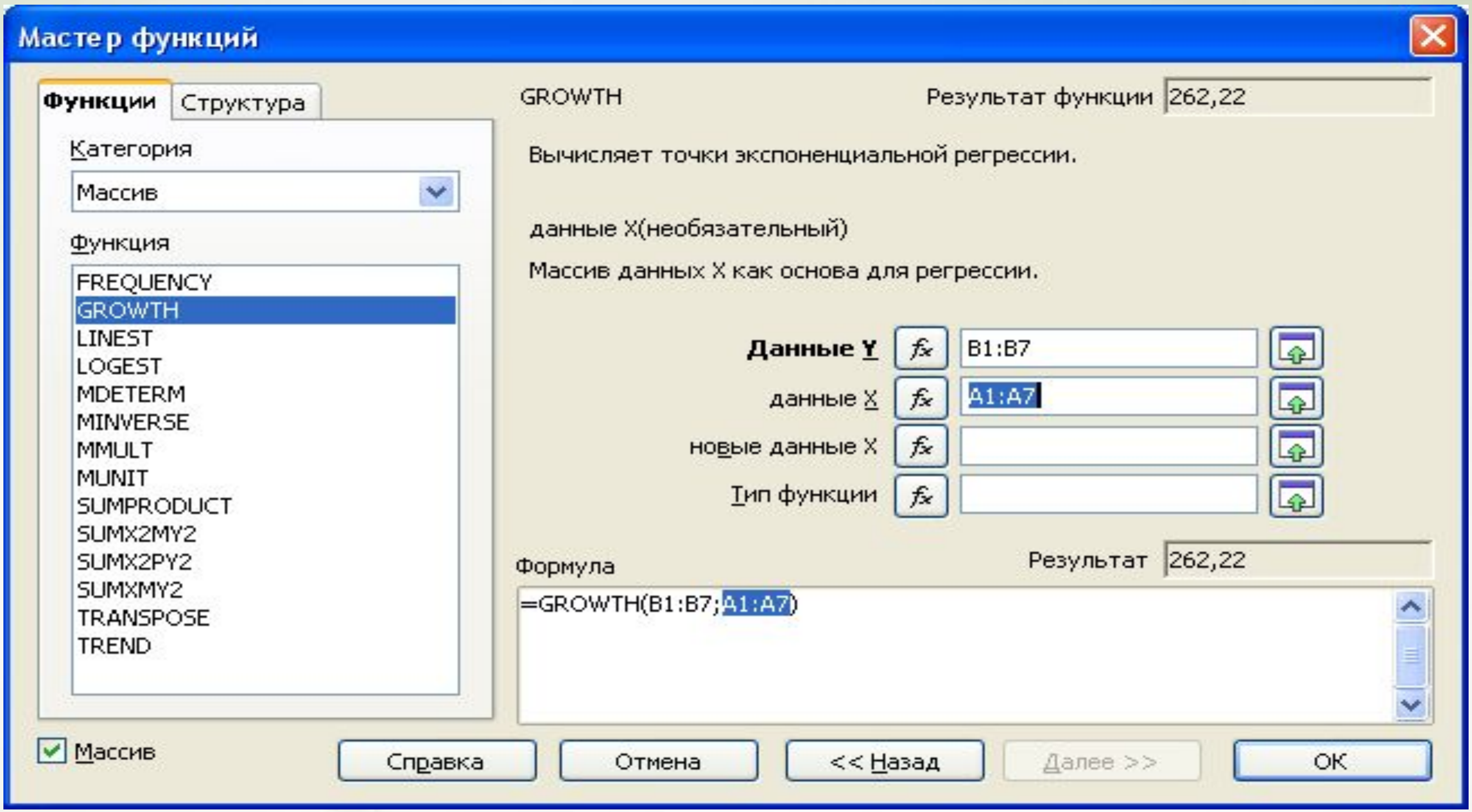
Использовать комбинацию клавиш **<u>«Ctrl» + «Shift» + «Enter».</u>**

Функции для построения экспоненциальной (нелинейной) регрессии

- Для данных, содержащих нулевые или отрицательные значения, этот вид приближения неприменим.
- **Growth – возвращает параметры экспоненциального тренда.**

Уравнение, описывающее кривую экспоненциальной регрессии имеет вид:

$$y = b*m^x$$


Мастер функций
Функции
Структура
Категория
Массив
Функция
FREQUENCY
GROWTH
LINEST
LOGEST
MDETERM
MINVERSE
MMULT
MUNIT
SUMPRODUCT
SUMX2MY2
SUMX2PY2
SUMXMY2
TRANSPOSE
TREND
GROWTH
Результат функции
262,22
Вычисляет точки экспоненциальной регрессии.
данные X(необязательный)
Массив данных X как основа для регрессии.
Данные Y
B1:B7
данные X
A1:A7
новые данные X
Тип функции
Формула
Результат
262,22
=GROWTH(B1:B7;A1:A7)
Массив
Справка
Отмена
<< Назад
Далее >>
OK

Следует отметить, что приемы построения регрессий с помощью функций TREND и GROWTH практически совпадают. То же самое можно сказать и о паре функций LINEST и LOGEST.

Отметим, что, в отличие от функций TREND и GROWTH, ни одна из перечисленных выше функций (SLOP, INTERCEPT, LINEST, LOGEST) не является регрессией. Эти функции играют лишь вспомогательную роль, определяя необходимые параметры регрессии.

Для линейной и экспоненциальной регрессий, построенных с помощью функций SLOP, INTERCEPT, LINEST, LOGEST, внешний вид их уравнений всегда известен, в отличие от линейной и экспоненциальной регрессий, соответствующих функциям TREND и GROWTH.

Ограничения: при расчете кривой регрессии учитываются только пары данных со следующими значениями:

- логарифмическая регрессия: учитываются только положительные значения x,
- экспоненциальная регрессия: учитываются только положительные значения y,
- потенциальная регрессия: учитываются только положительные значения x и положительные значения y.

Достоинствами инструмента встроенных функций:

- достаточно простой однотипный процесс формирования рядов данных исследуемой характеристики для всех встроенных статистических функций, задающих линии тренда;
- стандартная методика построения линий тренда на основе сформированных рядов данных;
- возможность прогнозирования поведения исследуемого процесса на необходимое количество шагов вперед или назад.

К недостаткам следует отнести:

- отсутствие встроенных функций для создания других (кроме линейного и экспоненциального) типов линий тренда. Это обстоятельство часто не позволяет подобрать достаточно точную модель исследуемого процесса, а также получить близкие к реальности прогнозы.
- при использовании функций TREND и GROWTH не известны уравнения линий тренда.

Достоинства встроенных функций

- *Простой процесс получения данных;*
- *Возможность прогнозирования результатов;*

Недостатки

- *нет встроенных функций для создания других (кроме линейного и экспоненциального) типов линий тренда.*
- *не известны уравнения линий тренда*