МЕТОДЫ ВЫЧИСЛЕНИЙ

ТЕМА 3. ОПТИМИЗАЦИЯ

Оптимизация

Оптимизация – это поиск оптимального (наилучшего) варианта в заданных условиях.

Оптимальное решение – такое, при котором некоторая заданная функция (*целевая функция*) достигает минимума или максимума.

Постановка задачи:

- **целевая функция**

$f(x) \to \min$ (расходы, потери, ошибки)

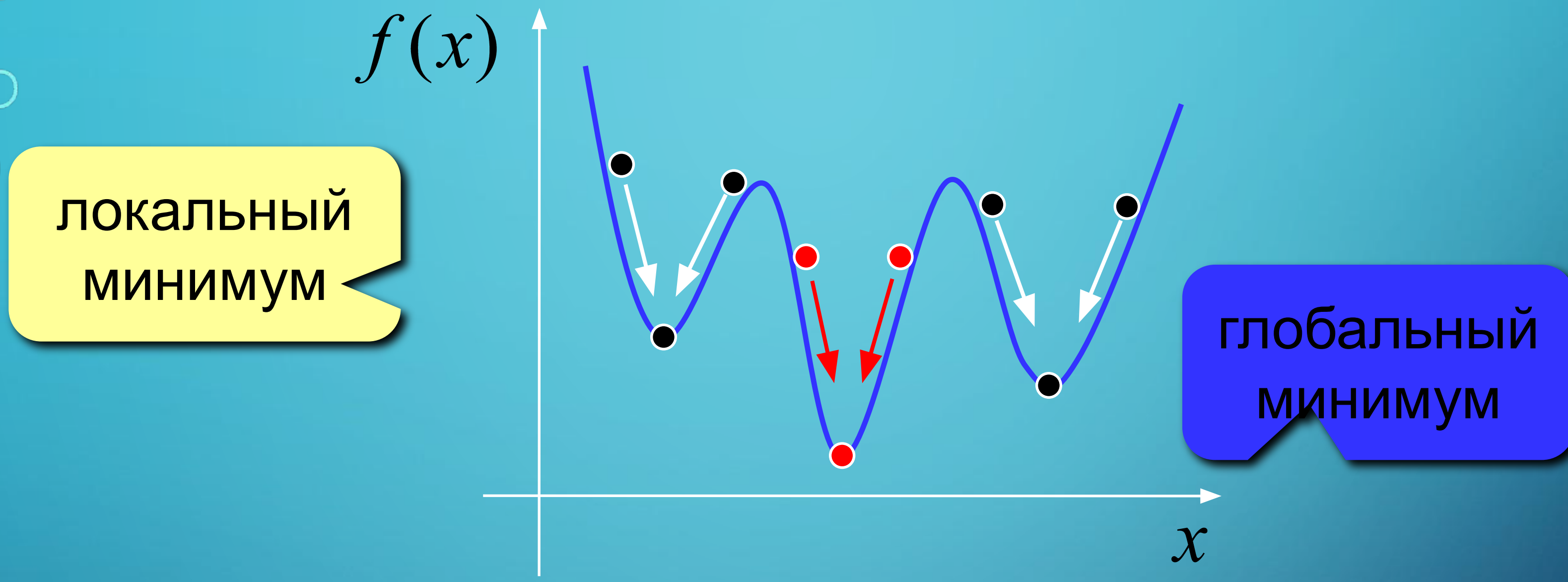
$f(x) \to \max$ (доходы, приобретения)

- **ограничения**, которые делают задачу осмысленной

Задача без ограничений: построить дом при минимальных затратах.

Решение: не строить дом вообще.

Оптимизация

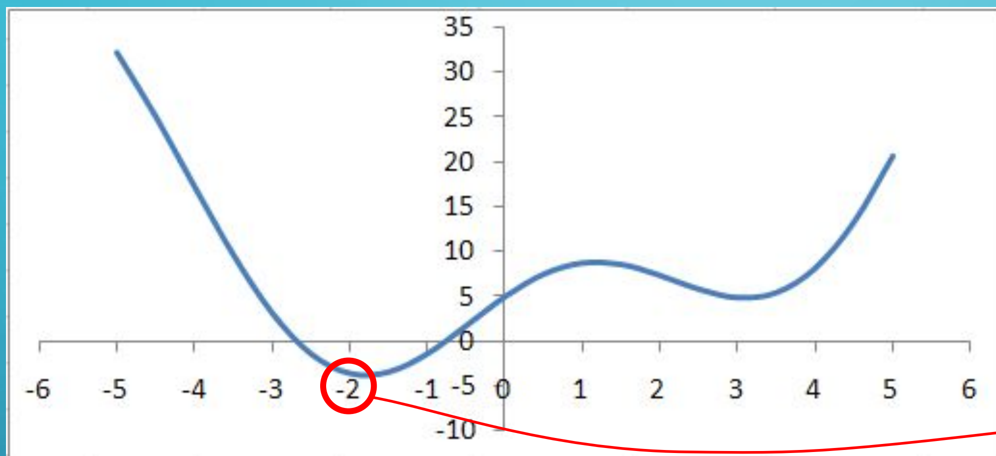


- обычно нужно найти **глобальный минимум**
- большинство численных методов находят только **локальный минимум**
- минимум, который найдет *Excel,* зависит от выбора начального приближения («шарик на горке скатится в ближайшую ямку»)

Поиск минимума функции

$$y = x^2 + 6\sin x + 5\cos x$$

1. Строим график функции (диаграмма «Точечная»)



? Зачем нужен график?

начальное приближение

$$x_0 = -2$$

2. Подготовка данных

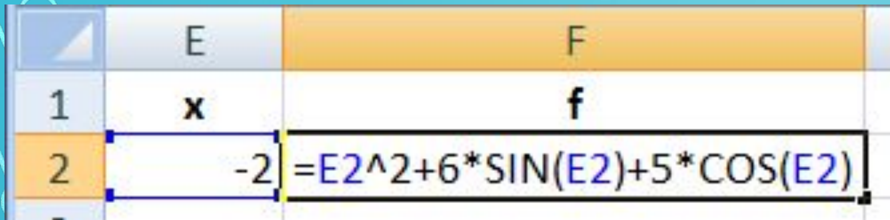
начальное приближение

	E	F
1	x	f
2	-2	=E2^2+6*SIN(E2)+5*COS(E2)

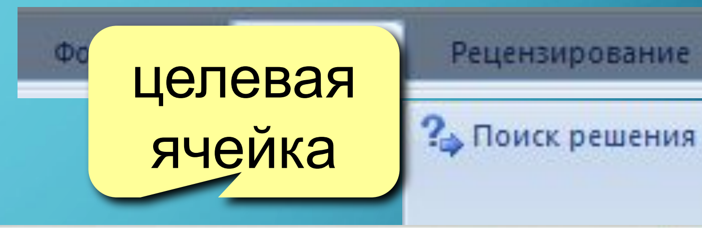
целевая ячейка

! Изменение E2 должно влиять на F2!

Поиск минимума функции



3. Надстройка «Поиск решения»



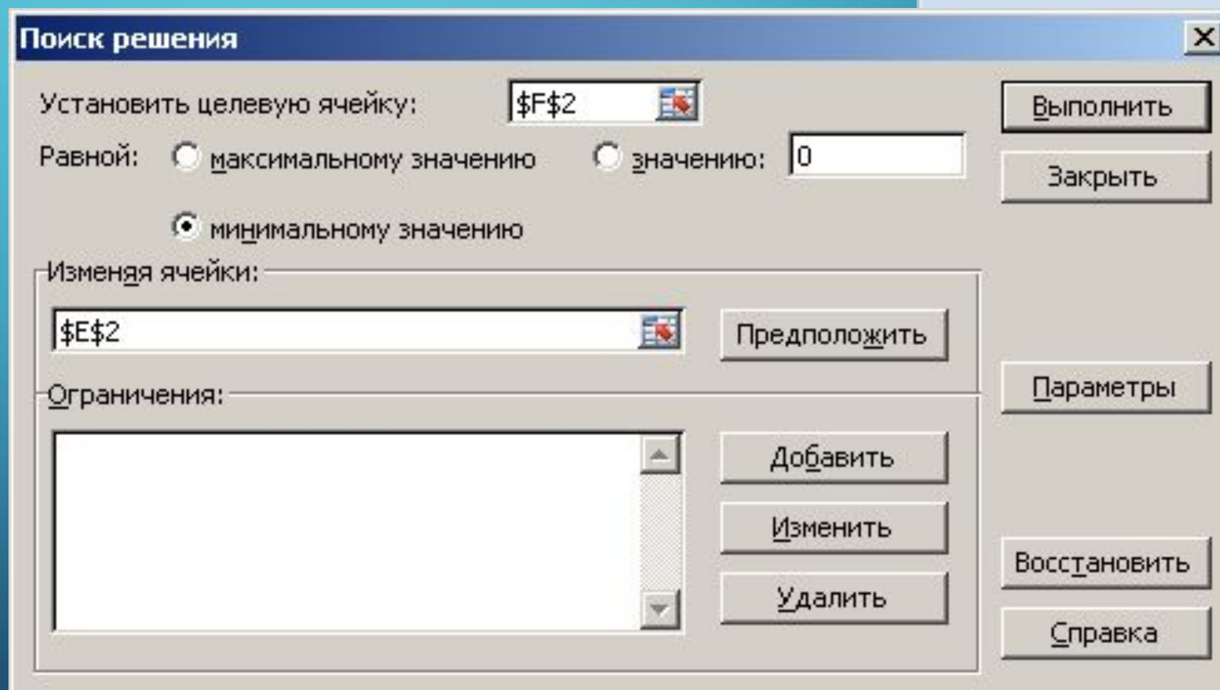
целевая ячейка

изменяемые ячейки:

E2

D2:D6

D2:D6; C5:C8



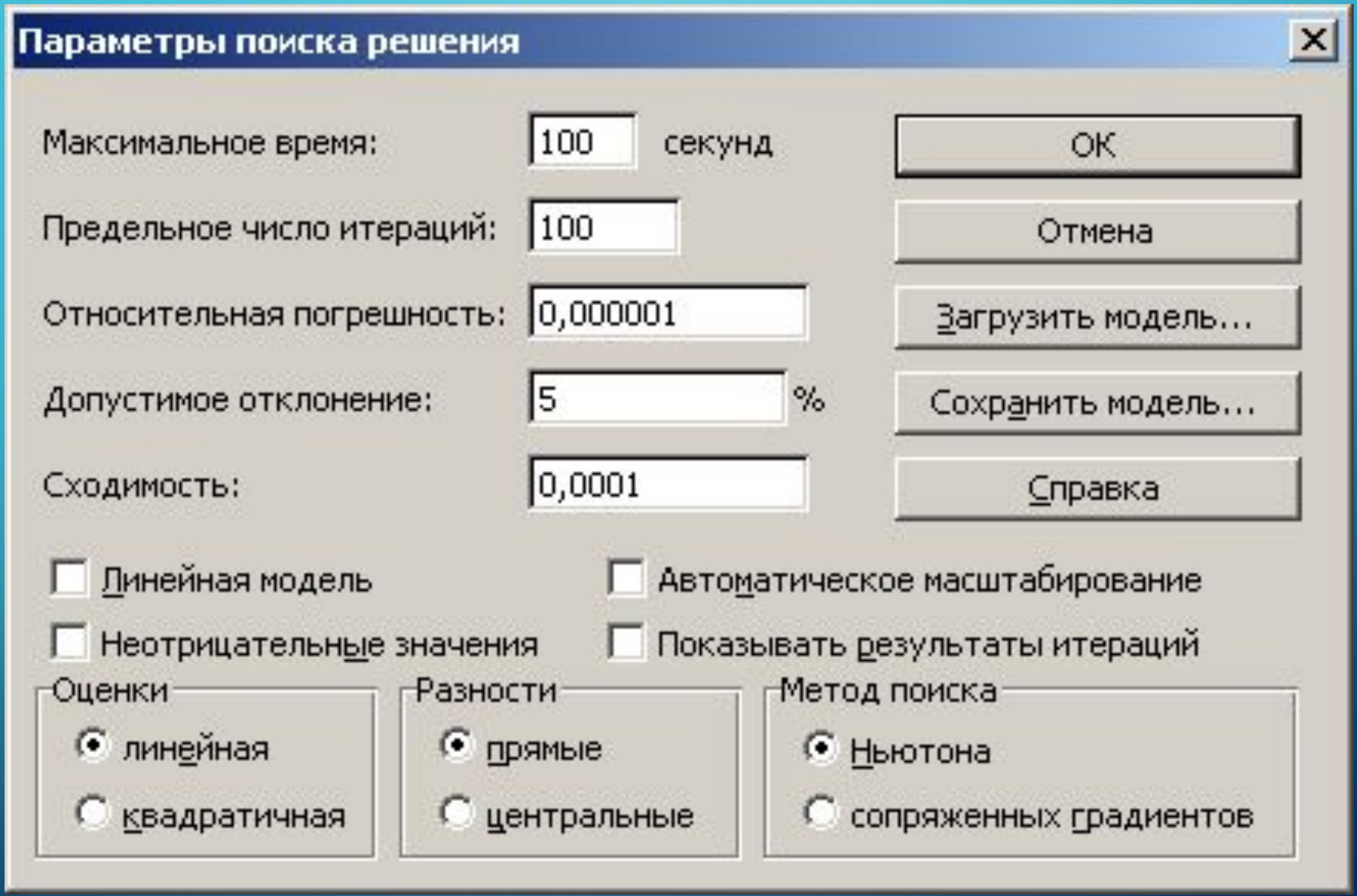
ограничения

A1 <= 20

B2:B8 >= 5

A1 = целое

Параметры оптимизации



Оптимизация

Подбор параметра – это оптимизация?

Надстройка «Поиск решения» позволяет:

- искать минимум и максимум функции
- использовать несколько изменяемых ячеек и диапазонов
- вводить ограничения (<=, >=, целое, двоичное)

Как влияет ограничение «А1-целое» на сложность решения задачи?

МЕТОДЫ ВЫЧИСЛЕНИЙ

ТЕМА 4. ВОССТАНОВЛЕНИЕ ЗАВИСИМОСТЕЙ

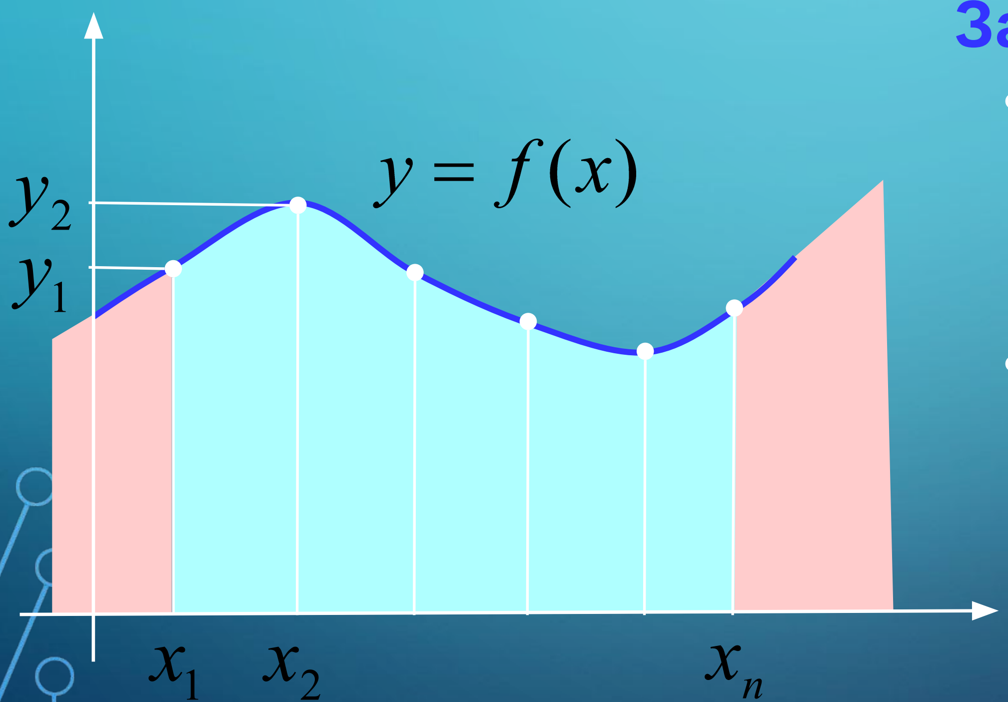
Восстановление зависимостей

Пары значений (аргумент-функция):

какую?

$$(x_1, y_1), (x_2, y_2), \; ..., \; (x_n, y_n)$$

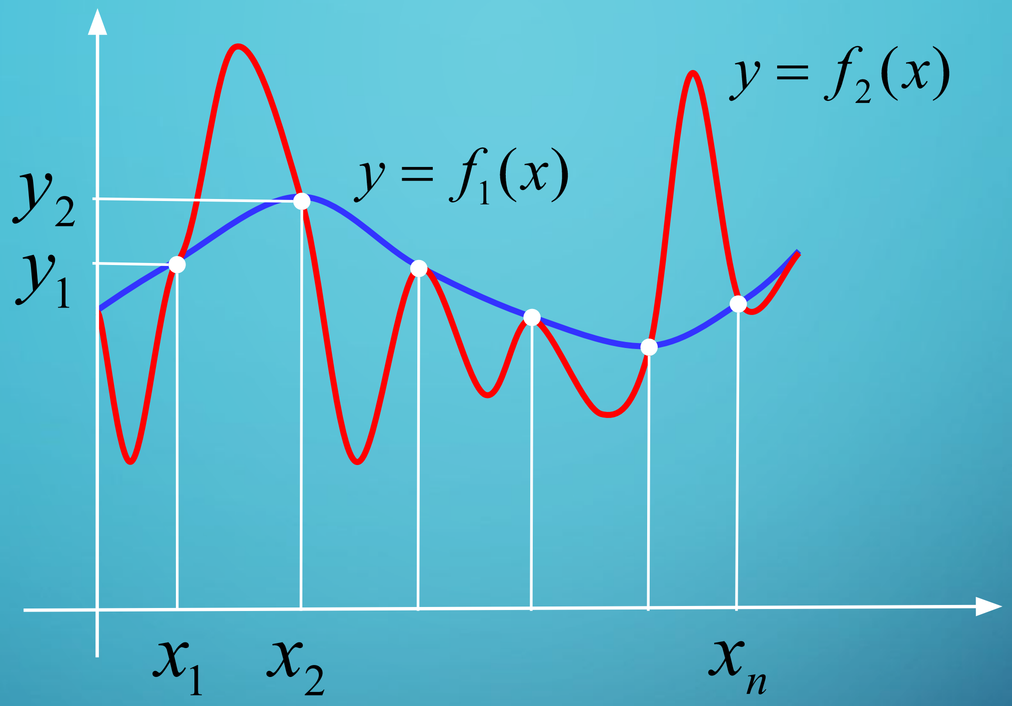
задают некоторую неизвестную функцию $y = f(x)$



Зачем:

- найти y в промежу-точных точках *(интерполяция)*
- найти y вне диапазона измерений *(экстраполяция, прогнозирование)*

Какое решение нам нужно?



Через заданный набор точек проходит бесконечно много разных кривых!

Вывод: задача **некорректна**, поскольку решение неединственно.

Восстановление зависимостей

Корректная задача: найти функцию **заданного вида**, которая лучше всего соответствует данным.

Примеры:

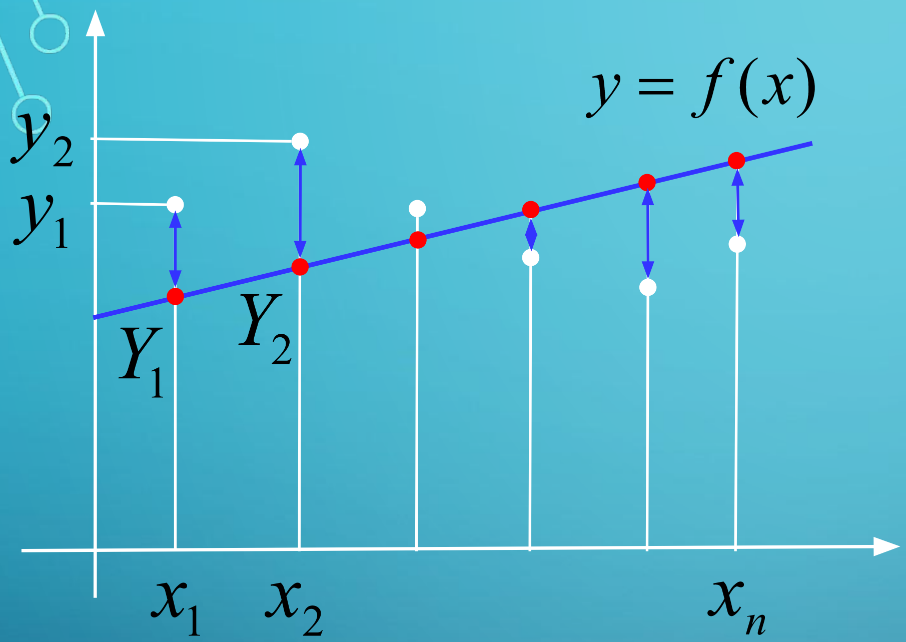
- линейная $y = a \cdot x + b$
- полиномиальная $y = a_3 x^3 + a_2 x^2 + a_1 x + a_0$
- степенная $y = a \cdot x^b$
- экспоненциальная $y = a \cdot e^{bx}$
- логарифмическая $y = a \cdot \ln x + b$

! **График функции не обязательно проходит через заданные точки!**

? **Как выбрать функцию?**

Что значит «лучше всего соответствует»?

Метод наименьших квадратов (МНК):



(x_i, y_i) заданные пары значений

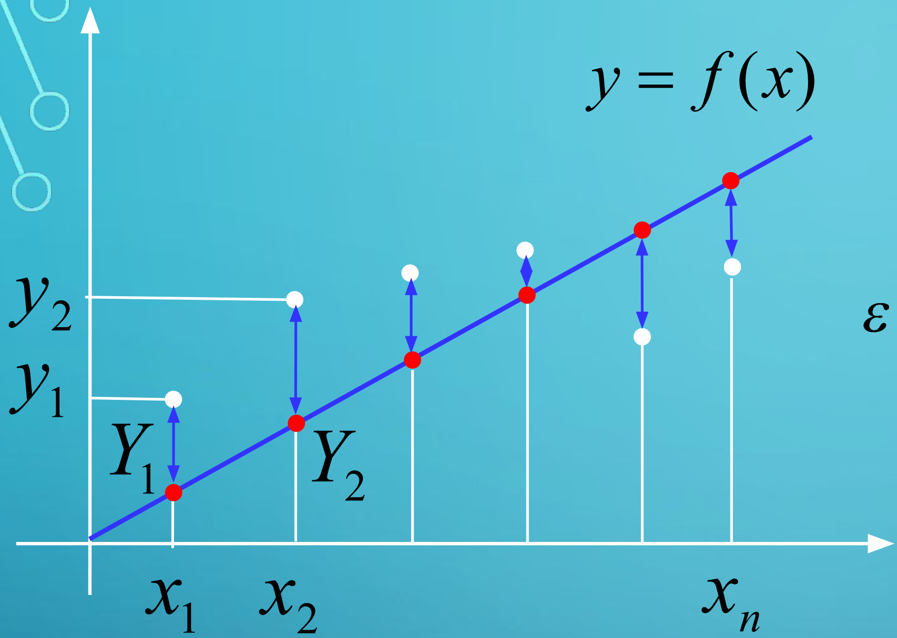
$$Y_i = f(x_i)$$

$$\varepsilon = \sum_{i=1}^{n} (y_i - Y_i)^2 \to \min$$

? Зачем возведение в квадрат?

1) чтобы складывать положительные значения
2) решение сводится к системе линейных уравнений (просто решать!)

МНК для линейной функции



неизвестно!

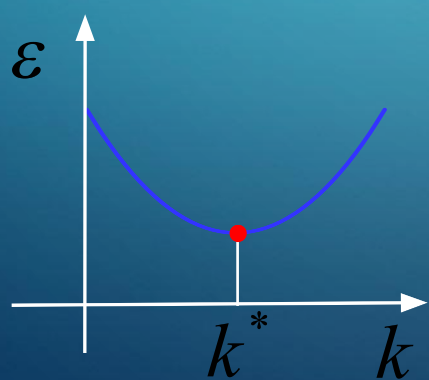
$$Y_i = k \cdot x_i$$

$$\varepsilon(k) = \sum_{i=1}^{n}(y_i - Y_i)^2 = \sum_{i=1}^{n}(y_i - kx_i)^2$$

$$= k^2 \cdot \boxed{\sum_{i=1}^{n} x_i^2} - k \cdot \boxed{2\sum_{i=1}^{n} x_i \cdot y_i} + \boxed{\sum_{i=1}^{n} y_i^2}$$

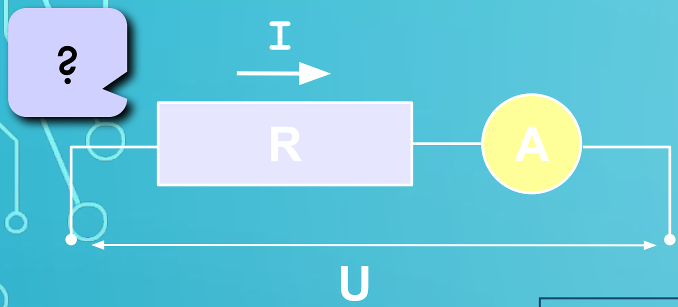
a $-b$ c

$$\varepsilon(k) = ak^2 + bk + c \to \min$$



$$k^* = -\frac{b}{2a} = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2}$$

Сопротивление проводника



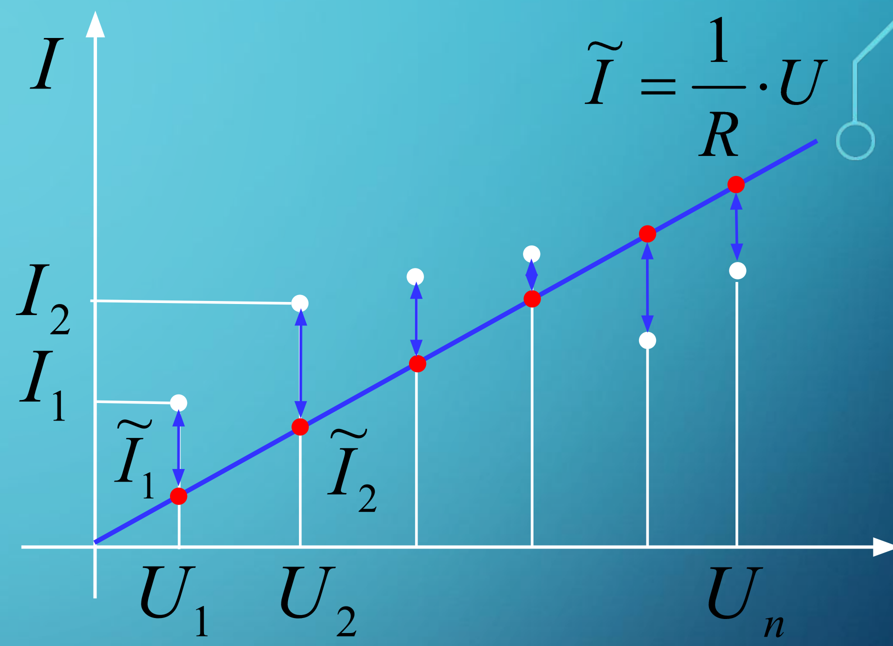
?

Закон Ома

$$I = \frac{1}{R} \cdot U$$

Точки на линии:

$$\tilde{I}_k = \frac{1}{R^*} \cdot U_k$$



$$\varepsilon(R) = \sum_{k=1}^{n} (I_k - \tilde{I}_k)^2 = \sum_{k=1}^{n} \left(I_k - \frac{1}{R} \cdot U_k \right)^2$$

$$= \frac{1}{R^2} \cdot \sum_{k=1}^{n} U_k^2 - \frac{1}{R} \cdot 2 \sum_{k=1}^{n} I_k \cdot U_k + \sum_{k=1}^{n} I_k^2$$

a $-b$

$$\frac{1}{R^*} = -\frac{b}{2a} = \frac{\sum_{k=1}^{n} I_k U_k}{\sum_{k=1}^{n} U_k^2}$$

Обработка результатов эксперимента

Задача. В файле `mnk.txt` записаны в столбик 10 пар чисел (напряжение, ток), полученные в результате эксперимента с одним резистором. Найти (приближенно) его сопротивление по методу наименьших квадратов.

$$\frac{1}{R^*} = \frac{\sum_{k=1}^{n} I_k U_k}{\sum_{k=1}^{n} U_k^2} \quad \Rightarrow \quad \boxed{R^* = \frac{\sum_{k=1}^{n} U_k^2}{\sum_{k=1}^{n} I_k U_k}}$$

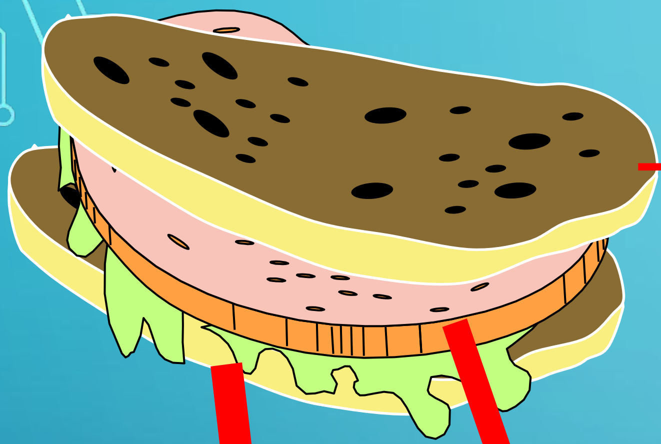
Этапы решения:

1. Прочитать данные из файла в массивы `U` и `I`.
2. Вычислить $\sum_{k=1}^{n} I_k U_k$ и $\sum_{k=1}^{n} U_k^2$.
3. Вычислить R^*.

Работа с файлами: принцип с

Переменная типа «текстовый файл»:

```
var f: text;
```



I этап. открыть файл :

- связать переменную **f** с файлом

```
Assign(f, 'mnk.txt');
```

- открыть файл (сделать его активным, приготовить к работе)

```
Reset(f); {для чтения}
```

```
Rewrite(f); {для записи}
```

II этап: работа с файлом

```
Read ( f, n );  { ввести значение n }
```

```
Write ( f, n );  { записать значение n }
Writeln ( f, n );{с переходом на нов.строку }
```

III этап: закрыть файл

```
Close(f);
```

Обработка результатов эксперимента

Чтение данных:

```
var f: text;
...
begin
  Assign(f, 'mnk.txt');
  Reset(f);
  for k:=1 to 10 do begin
    Read(f, U[k], I[k]);
    Writeln(U[k]:0:3, ' ', I[k]:0:3);
  end;
  Close(f);
end.
```

```
U, I: array[1..10] of real;
   k: integer;
```

? **Какие переменные и массивы надо объявить?**

Обработка результатов эксперимента

Вычисления:

```
var UU: real;
...
UU := 0;
for k:=1 to 10 do begin
  UU := UU + U[k]*U[k];
end;
```

$$\sum_{k=1}^{n} U_k^2$$

Что вычисляем?

Как найти $\sum_{k=1}^{n} I_k U_k$?

Как найти R^*?

$$R^* = \frac{\sum_{k=1}^{n} U_k^2}{\sum_{k=1}^{n} I_k U_k}$$

Задания

«4»: Используя метод наименьших квадратов, найти приближенное значение сопротивления по данным файла `mnk.txt`.

«5»: Сделать то же самое, предполагая, что в файле неизвестное количество пар значений, но не более 100. Цикл ввода должен выглядеть так:

пока не достигнут конец файла (`eof` = *end of file*)

```
while not eof(f) do begin
  { читаем U[k] и I[k] }
  { тут еще что-то надо сделать }
end;
```

Коэффициент достоверности (*Excel*)

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - Y_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}$$

(x_i, y_i) заданные пары значений

$Y_i = f(x_i)$

$\bar{y}$ – среднее значение y_i

Крайние случаи:

•если график проходит через точки:

$$R^2 = 1$$

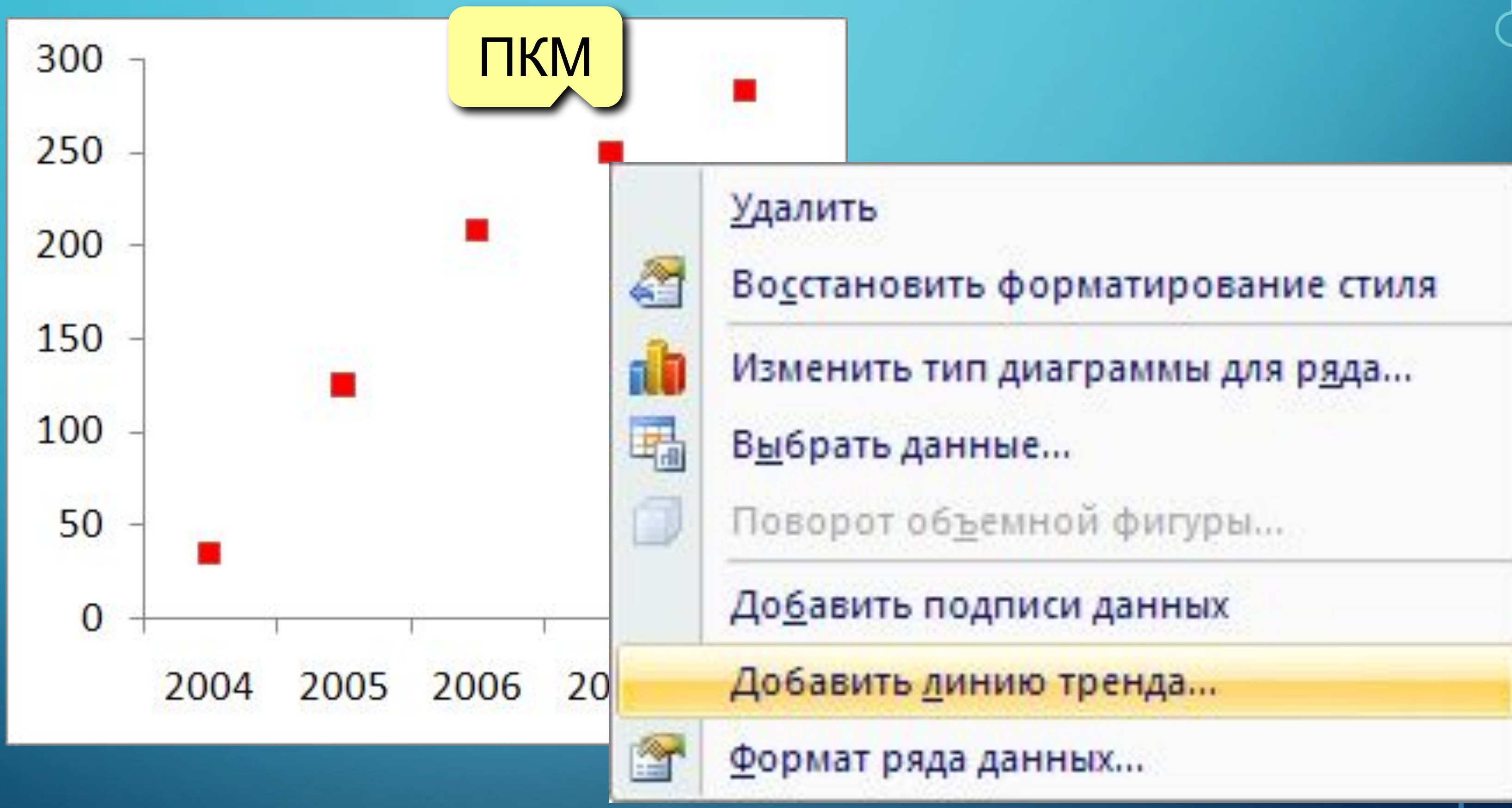
•если считаем, что y не меняется и $Y_i = \bar{y}$:

$$R^2 = 0$$

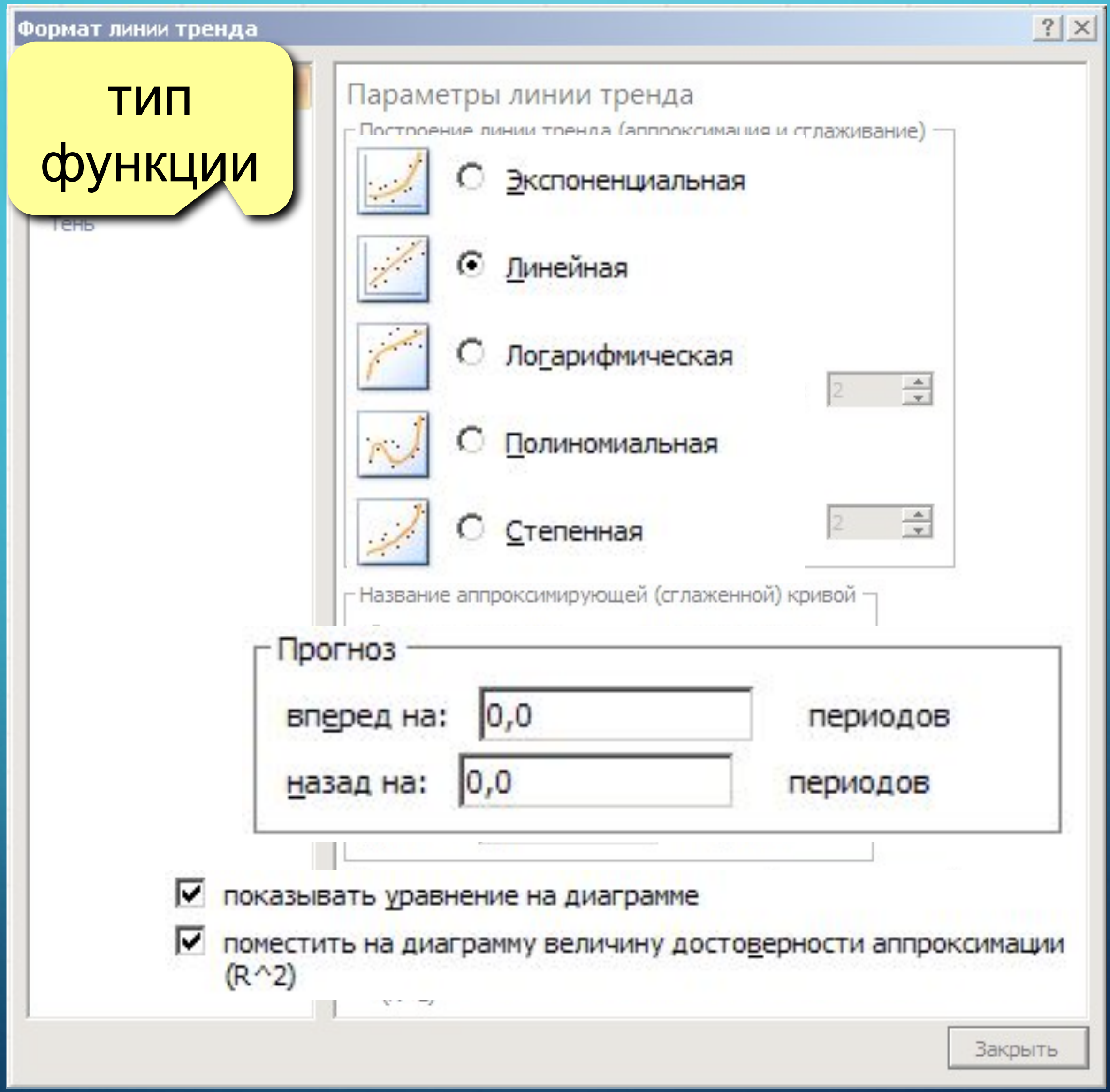
! Фактически – метод наименьших квадратов!

Восстановление зависимостей

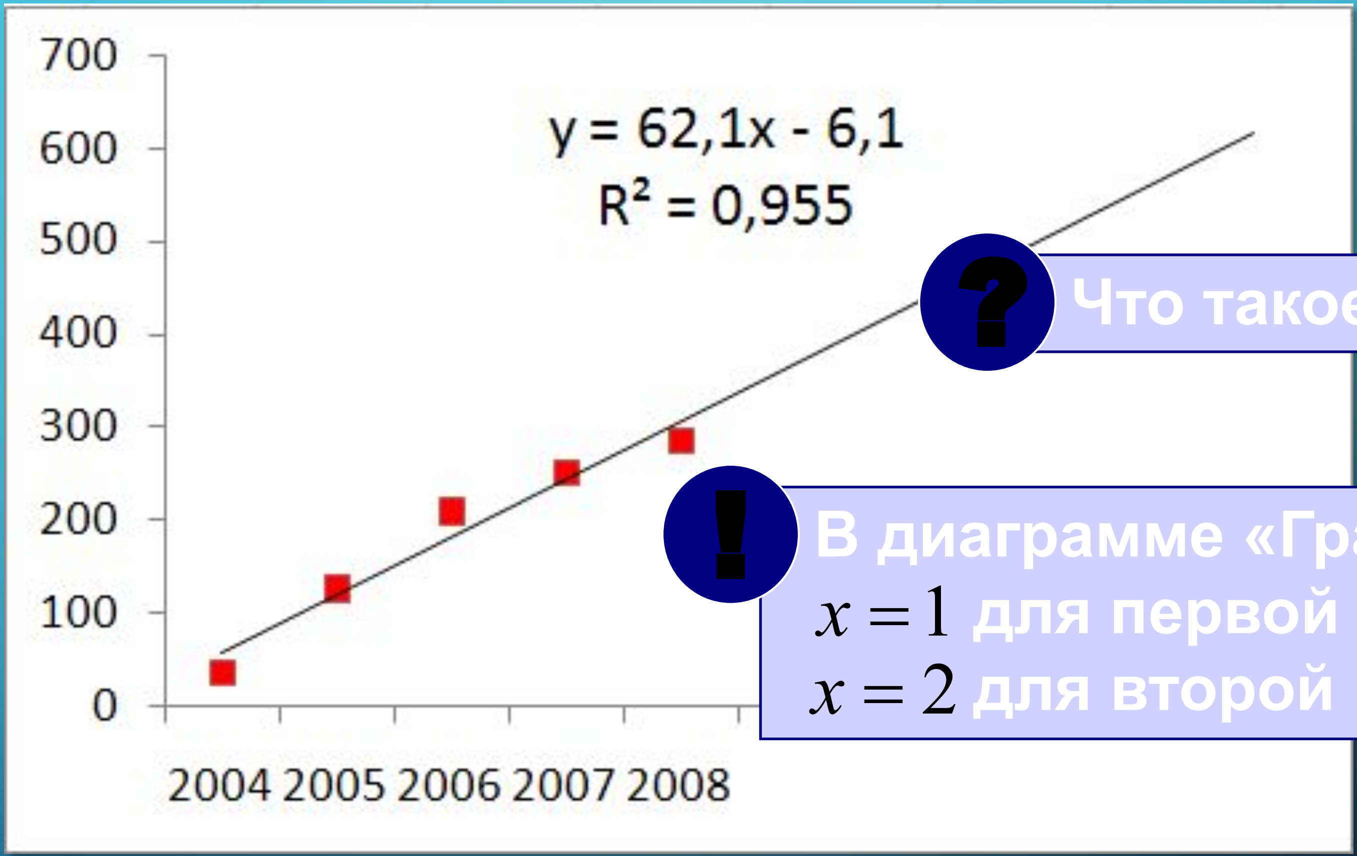
Диаграмма «График»:



Восстановление зависимостей



Восстановление зависимостей



? Что такое x ?

! В диаграмме «График» $x = 1$ для первой точки, $x = 2$ для второй и т.д.

? Насколько хорошо выбрана функция?

Восстановление зависимостей

Сложные случаи (нестандартная функция):

$$f(x) = a \cdot \sin kx + b$$

? Что делать?

Алгоритм:

1) выделить ячейки для хранения a, k, b
2) построить ряд $Y_i = f(x_i)$ для тех же x_i
3) построить на одной диаграмме ряды y_i и Y_i
4) попытаться подобрать a, k, b так, чтобы два графика были близки
5) вычислить R^2 в отдельной ячейке

   функции: СУММКВРАЗН – сумма квадратов разностей рядов
   ДИСПР – дисперсия

6) Поиск решения: $R^2 \to \min$

! Это задача оптимизации!

МЕТОДЫ ВЫЧИСЛЕНИЙ

ТЕМА 5. СТАТИСТИКА

Ряд данных и его свойства

Ряд данных – это упорядоченный набор значений

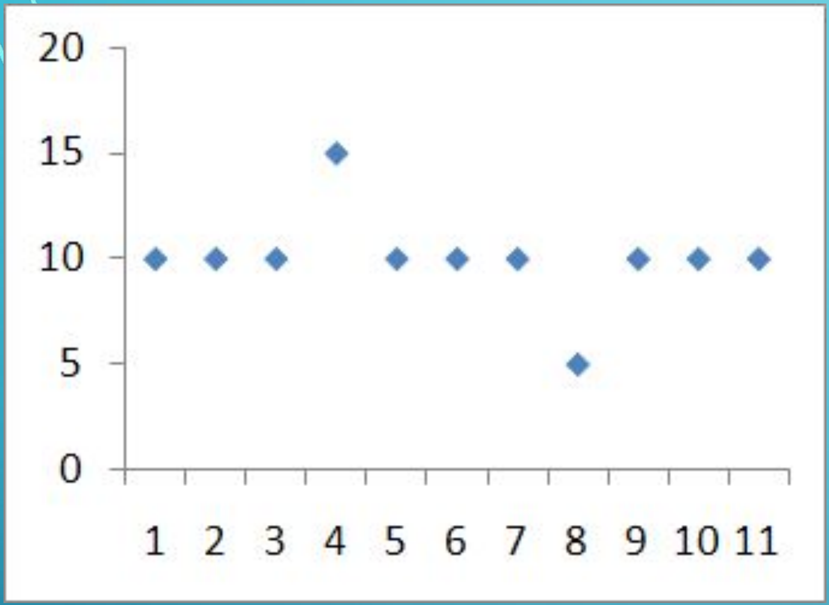
$$x_1,\ x_2,\ \ldots,\ x_n$$

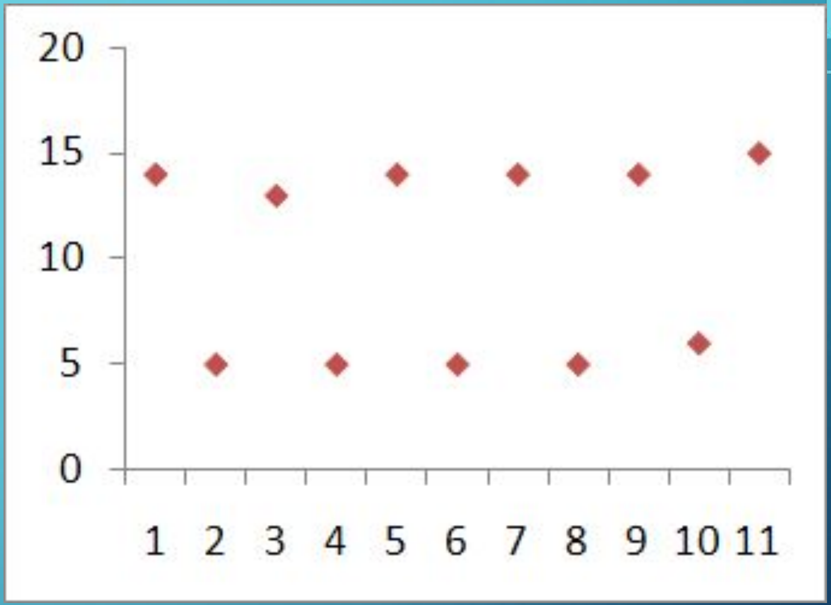
Основные свойства (ряд **A1:A20**):

- количество элементов **=СЧЕТ(A1:A20)**
- количество элементов, удовлетворяющих некоторому условию:
  = СЧЕТЕСЛИ(A1:A20;"<5")
- минимальное значение **=МИН(A1:A20)**
- максимальное значение **=МАКС(A1:A20)**
- сумма элементов **=СУММ(A1:A20)**
- среднее значение **=СРЗНАЧ(A1:A20)**

Дисперсия

Для этих рядов одинаковы МИН, МАКС, СРЗНАЧ





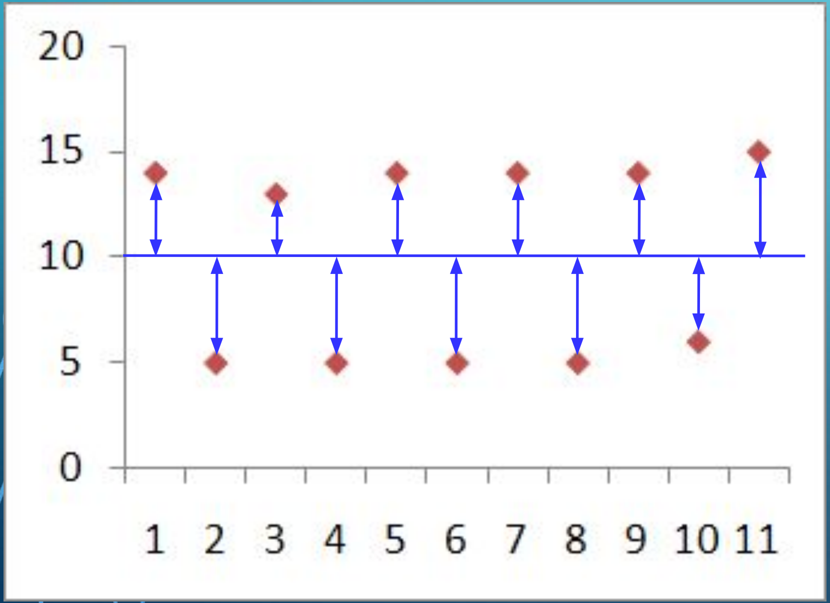
В чем различие?

Дисперсия («разброс») – это величина, которая характеризует разброс данных относительно среднего значения.

Дисперсия

$$D_x = \frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}$$

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$ среднее арифметическое



$(x_1 - \bar{x})^2$ квадрат отклонения x_1 от среднего

D_x *средний* квадрат отклонения от среднего значения

Дисперсия и СКВО

Стандартная функция

=ДИСПР(A1:A20)

Функции – Другие – Статистические

Что неудобно:

если x измеряется в метрах,

то D_x – в $м^2$

В каких единицах измеряется?

СКВО = среднеквадратическое отклонение

$$\sigma_x = \sqrt{D_x}$$

=СТАНДОТКЛОНП(A1:A20)

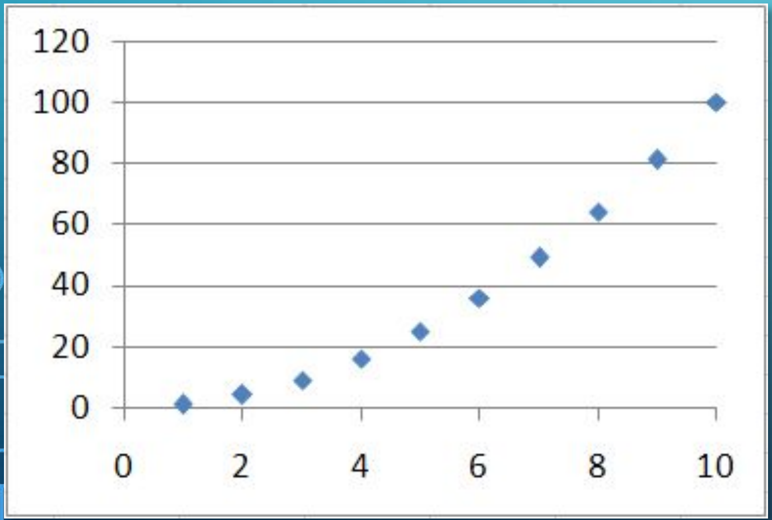
Взаимосвязь рядов данных

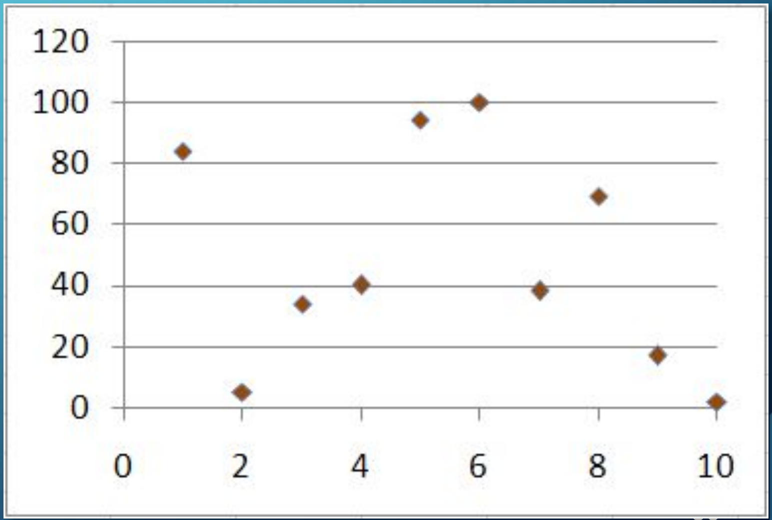
Два ряда одинаковой длины:

$$x_1,\ x_2,\ \ldots,\ x_n \qquad\qquad y_1,\ y_2,\ \ldots,\ y_n$$

Вопросы:

- есть ли связь между этими рядами (соответствуют ли пары (x_i, y_i) какой-нибудь зависимости $y = f(x)$)
- насколько сильна эта связь?





Взаимосвязь рядов данных

Ковариация:

$$K_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n}$$

? Если x и y – один и тот же ряд? $K_{xx} = D_x$

в среднем!

Как понимать это число?

- если $K_{xy} > 0$ увеличение x приводит к увеличению y
- если $K_{xy} < 0$ увеличение x приводит к уменьшению y
- если $K_{xy} \approx 0$ связь обнаружить не удалось

Что плохо?

- единицы измерения: если x в метрах, y в литрах, то K_{xy} – в м · л
- K_{xy} зависит от абсолютных значений x и y, поэтому ничего не говорит о том, насколько сильна связь

Взаимосвязь рядов данных

Коэффициент корреляции:

$$\rho_{xy} = \frac{K_{xy}}{\sigma_x \cdot \sigma_y}$$

$\sigma_x,\ \sigma_y$ – СКВО рядов x и y

? Какова размерность? безразмерный!

$$\boxed{-1 \le \rho_{xy} \le 1}$$

Как понимать это число?

- если $\rho_{xy} > 0$: увеличение x приводит к увеличению y
- если $\rho_{xy} < 0$: увеличение x приводит к уменьшению y
- если $\rho_{xy} \approx 0$: связь обнаружить не удалось

```
=КОРРЕЛ(A1:A20;B1:B20)
```

Взаимосвязь рядов данных

Как понимать коэффициент корреляции?

$0 < |\rho_{xy}| \le 0{,}2$: очень слабая корреляция

$0{,}2 < |\rho_{xy}| \le 0{,}5$: слабая

$0{,}5 < |\rho_{xy}| \le 0{,}7$: средняя

$0{,}7 < |\rho_{xy}| \le 0{,}9$: сильная

$0{,}9 < |\rho_{xy}| \le 1$: очень сильная

$\rho_{xy} = 1$: линейная зависимость $y = ax + b, \quad a > 0$

$\rho_{xy} = -1$: линейная зависимость $y = ax + b, \quad a < 0$

? **Если $\rho_{xy} \approx 0$, то связи нет?**

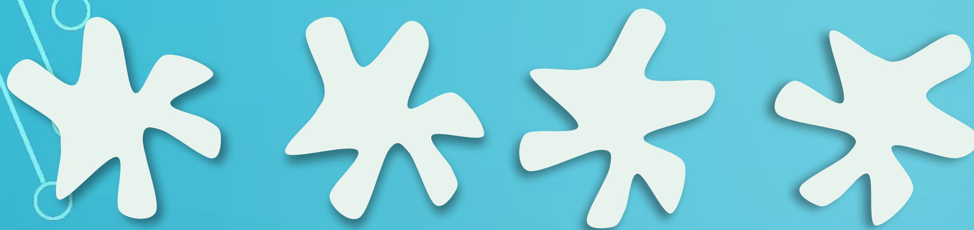
! **Метод для определения линейной зависимости!**

МЕТОДЫ ВЫЧИСЛЕНИЙ

ТЕМА 6. МОДЕЛИРОВАНИЕ

(по мотивам учебника А.Г. Гейна и др., Информатика и ИКТ, 10 класс, М.: Просвещение, 2008)

Модель деления

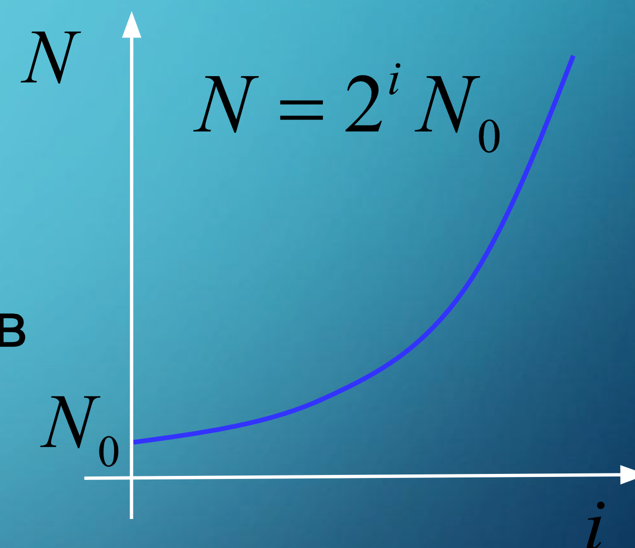


N_0 – начальная численность

$N_1 = 2N_0$ – после 1 цикла деления

$N_2 = 2N_1 = 4N_0$ – после 2-х циклов

$N_i = 2N_{i-1} = 2^i N_0$



Особенности модели:

1) не учитывается смертность
2) не учитывается влияние внешней среды
3) не учитывается влияние других видов

Модель неограниченного роста (Т. Мальтус)

$$N_i = N_{i-1} + K_p \cdot N_{i-1} - K_c \cdot N_{i-1}$$

K_p – коэффициент рождаемости

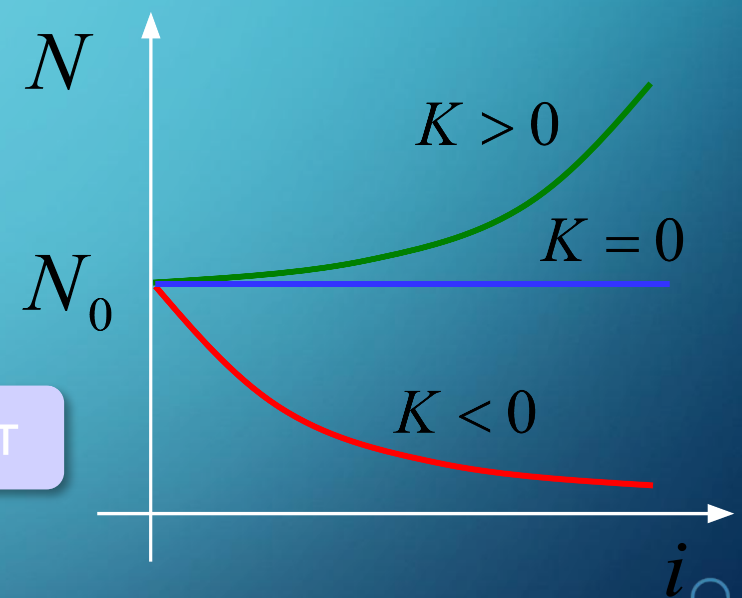
K_c – коэффициент смертности

Коэффициент прироста $K = K_p - K_c$

$$\boxed{N_i = (1+K) \cdot N_{i-1}}$$

$$N_i = N_{i-1} + K \cdot N_{i-1}$$

прирост



Особенности модели:

1) не учитывается влияние численности N и внешней среды на K
2) не учитывается влияние других видов на K

Модель ограниченного роста (П. Ферхюльст)

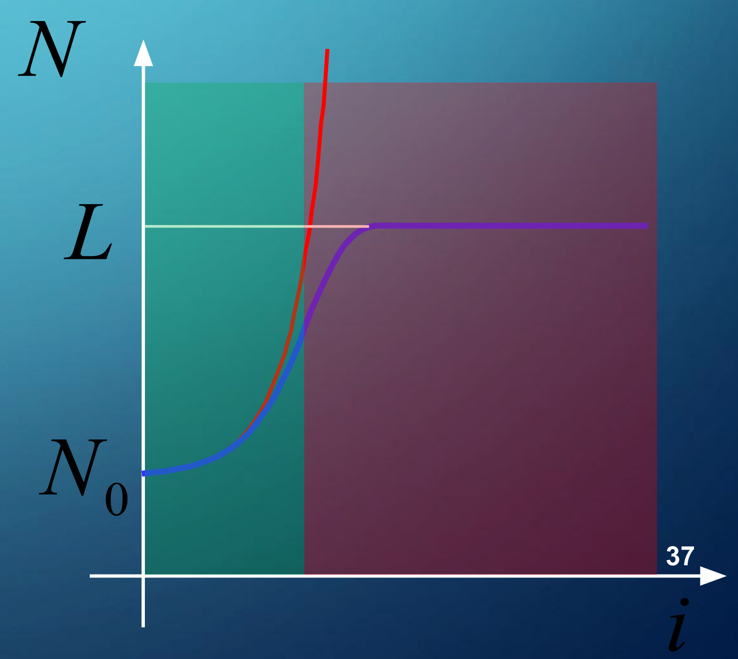
L – предельная численность животных

$$N_i = (1 + K_L) \cdot N_{i-1}$$

Идеи:

1) коэффициент прироста K_L зависит от численности N
2) при $N=0$ должно быть $K_L=K$ (начальное значение)
3) при $N=L$ должно быть $K_L=0$ (достигнут предел)

$$N_i = \left(1 + K\frac{L - N_{i-1}}{L}\right) \cdot N_{i-1}$$



! Модель адекватна, если ошибка < 10%!

Модель с отловом

Примеры: рыбоводческое хозяйство, разведение пушных зверей и т.п.

$$N_i = \left(1 + K\frac{L - N_{i-1}}{L}\right)\cdot N_{i-1} \boxed{- R}$$

отлов

? Какая будет численность?

$N_i = N_{i-1}$, прирост = отлову

$$N = N + K\frac{L - N}{L}N - R \quad \Rightarrow \frac{K}{L}\cdot N^2 - K\cdot N + R = 0$$

? Сколько можно отловить?

Модель эпидемии гриппа

L – всего жителей

N_i – больных в i-ый день

Z_i – заболевших в i-ый день

V_i – выздоровевших

W_i – всего выздоровевших за i дней

Основное уравнение:

$$N_i = N_{i-1} + Z_i - V_i$$

Ограниченный рост:

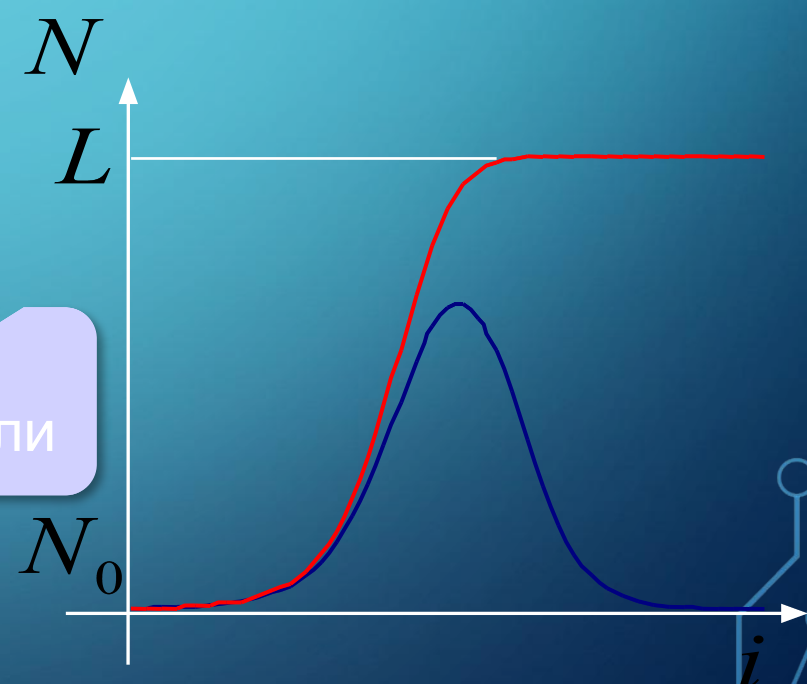
$$Z_i = K\frac{L - N_{i-1} - W_{i-1}}{L}\cdot N_{i-1}$$

болели и выздоровели

Выздоровление (через 7 дней):

$$V_i = Z_{i-7}$$

$$W_i = W_{i-1} + V_i$$



Влияние других видов

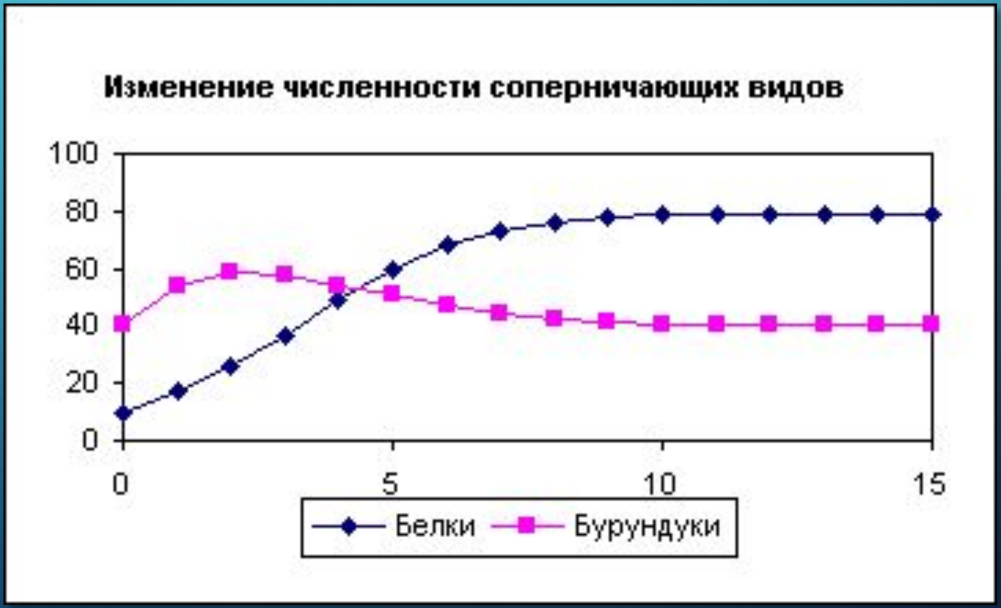
N_i – численность белок, M_i – численность бурундуков

$$N_i = N_{i-1}(2 - K_1 \cdot N_{i-1} - K_2 \cdot M_{i-1})$$
$$M_i = M_{i-1}(2 - K_3 \cdot M_{i-1} - K_4 \cdot N_{i-1})$$

? Откуда видно влияние?

K_2, K_4 – взаимное влияние

если $K_2 > K_1$ или $K_4 > K_3$ – враждующие виды



Моделирование двух популяций

N_0 M_0

	A	B	C	D	E	F
1	*i*	*N*	*M*		*K1*	0,05
2	0	20	20		*K2*	0,01
3	1	=B2*(2-F1*B2-F2*C2)			*K3*	0,05
4	2				*K4*	0,01
5	3					

$$N_i = N_{i-1}(2 - K_1 \cdot N_{i-1} - K_2 \cdot M_{i-1})$$

? Как скопировать формулы «вниз»?

Модель системы «хищник-жертва»

Модель – не-система:

караси

щуки

$$N_i = \left(1 + K\frac{L - N_{i-1}}{L}\right)\cdot N_{i-1}$$

$$Z_i = (1 - d)\cdot Z_{i-1}$$

вымирают без еды

Модель – система:

1) число встреч пропорционально $N_i \cdot Z_i$
2) «эффект» пропорционален числу встреч

$$N_i = \left(1 + K\frac{L - N_{i-1}}{L}\right)\cdot N_{i-1} - b_1 \cdot N_{i-1} \cdot Z_{i-1}$$

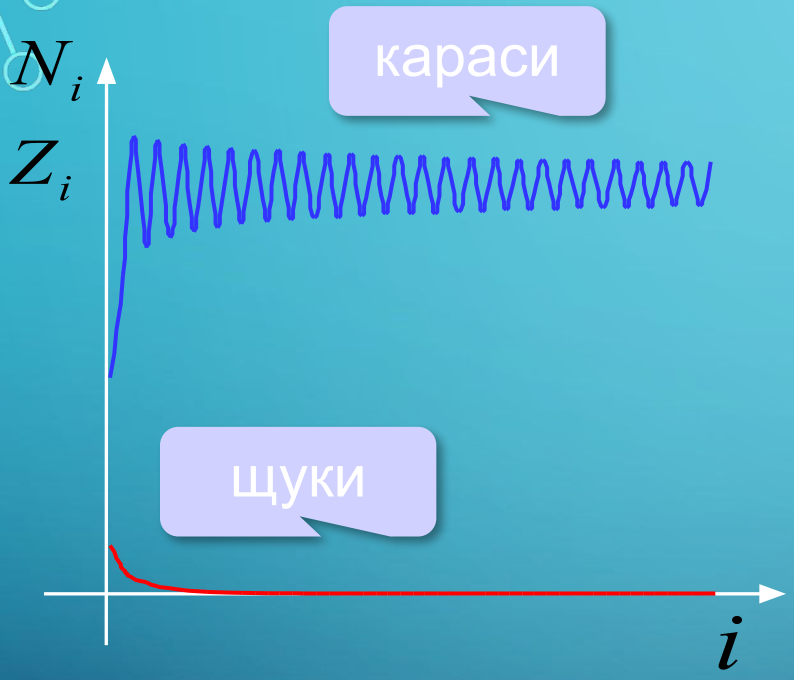
$$Z_i = (1 - d)\cdot Z_{i-1} + b_2 \cdot N_{i-1} \cdot Z_{i-1}$$

численность уменьшается

численность увеличивается

Модель системы «хищник-жертва»

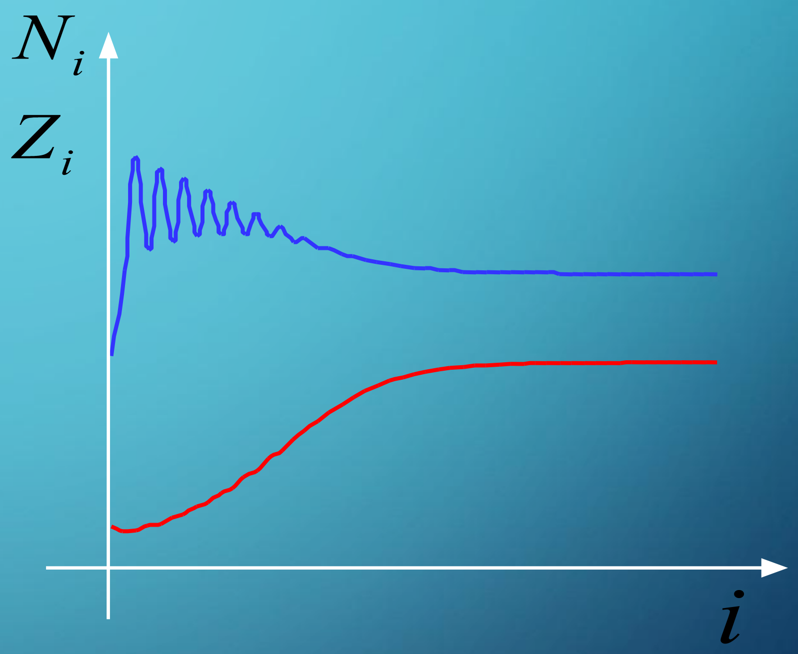
Хищники вымирают:



$$d = 0{,}8$$

$$b_1 = b_2 = 0{,}5$$

Равновесие:

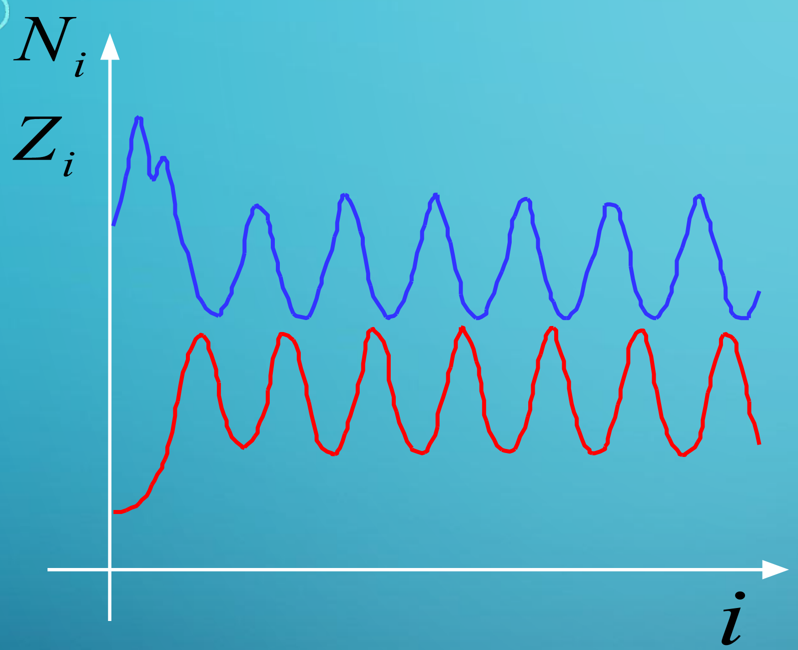


$$d = 0{,}8$$

$$b_1 = 0{,}5; \quad b_2 = 1$$

Модель системы «хищник-жертва»

Колебания:



$$d = 0{,}8$$

$$b_1 = 0{,}5; \; b_2 = 2$$

Случайные процессы

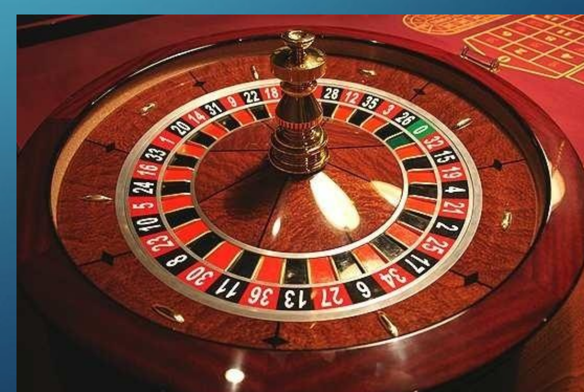
Случайно…

1)встретить друга на улице
2)разбить тарелку
3)найти 10 рублей
4)выиграть в лотерею

Случайный выбор:

1)жеребьевка на соревнованиях
2)выигравшие номера в лотерее

Как получить случайность?



Случайные числа на компьютере

Электронный генератор

- нужно специальное устройство
- нельзя воспроизвести результаты

Псевдослучайные числа – обладают свойствами случайных чисел, но каждое следующее число вычисляется по заданной формуле.

Метод середины квадрата (Дж. фон Нейман)

564321

в квадрате

318**458191**041

209**938992**481

малый период (последовательность повторяется через 10^6 чисел)

Случайные числа на компьютере

Линейный конгруэнтный метод

остаток от деления

$$x_n = (a \cdot x_{n-1} + c) \mod m$$

a, c, m - целые числа

$$x_n = (16807 \cdot x_{n-1} + 12345) \mod 1073741823$$

простое число

2^{30}-1

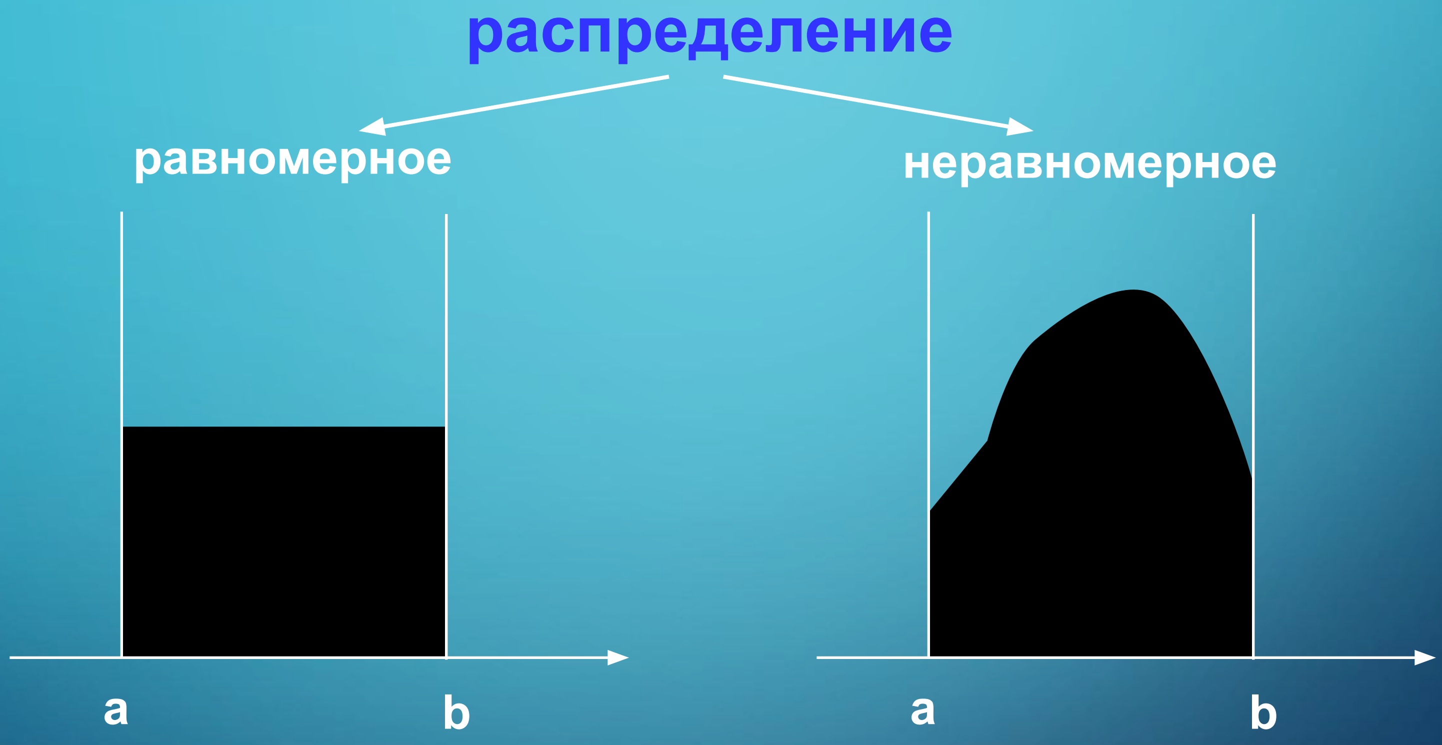
? Какой период?

период m

«Вихрь Мерсенна»: период 2^{19937}-1

Распределение случайных чисел

Модель: снежинки падают на отрезок [a,b]

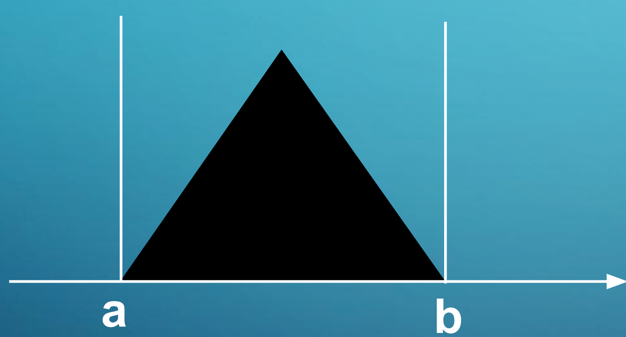


Сколько может быть разных распределений?

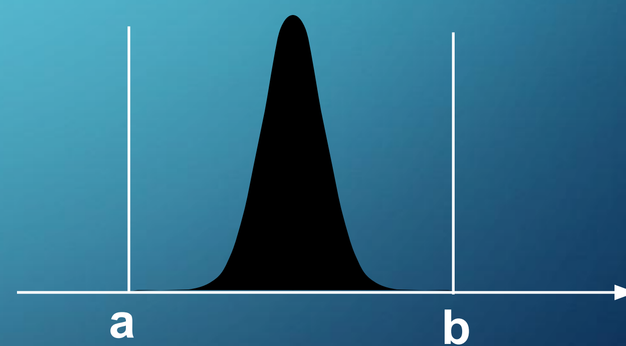
Распределение случайных чисел

Особенности:

- распределение – это характеристика **всей последовательности**, а не одного числа
- **равномерное** распределение одно, компьютерные датчики (псевдо)случайных чисел дают равномерное распределение
- неравномерных – много
- любое неравномерное можно получить с помощью равномерного



$$x = \frac{x_1 + x_2}{2}$$

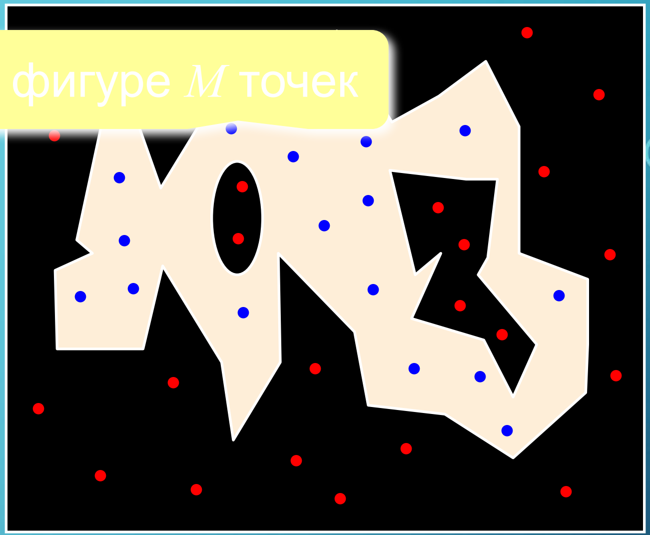


$$x = \frac{x_1 + x_2 + \dots + x_{12}}{12}$$

$x_1, x_2, \dots$ равномерное распределение

Вычисление площади (метод Монте-Карло)

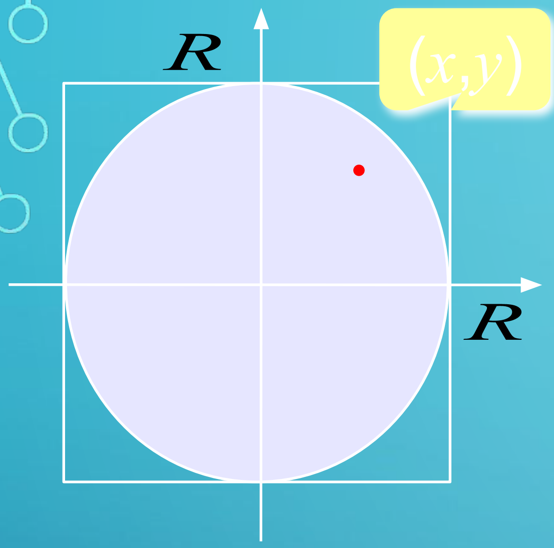
1. Вписываем сложную фигуру **в другую фигуру**, для которой легко вычислить площадь (прямоугольник, круг, …).
2. **Равномерно** N точек со случайными координатами внутри прямоугольника.
3. Подсчитываем количество точек, **попавших на фигуру**: M.
4. Вычисляем **площадь**: $\frac{S}{S_0} \approx \frac{M}{N} \Rightarrow S \approx S_0 \cdot \frac{M}{N}$



!

1. Метод приближенный.
2. Распределение должно быть равномерным.
3. Чем больше точек, тем точнее.
4. Точность ограничена датчиком случайных чисел.

Вычисление площади



Случайные координаты:

```
x := R*random;
y := R*random;
```

Когда точка внутри круга?

$$x^2 + y^2 \le R^2$$

Программа:

```
for i:=1 to N do begin
  { найти случайные координаты }
  if x*x + y*y <= R*R then M := M+1;
end;
S := 4*R*R*M / N;
```

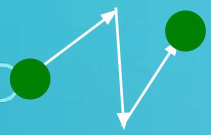
? Как найти число π?

Задания

«4»: Вычислите площади кругов с радиусами
`R = 1, 2, 3, 4, 5.`
Используя электронные таблицы, найдите приближенную формулу для вычисления площади круга.

«5»: Вычислите объем шаров с радиусами
`R = 1, 2, 3, 4, 5.`
Используя электронные таблицы, найдите приближенную формулу для вычисления объема шара.

Броуновское движение



Случайное направление (в рад):

```
alpha := 2*pi*random;
```

Случайный шаг:

```
h := hMax*random;
```

Программа:

```
for i:=1 to N do begin
  { найти случайное направление и шаг }
  x := x + h*cos(alpha);
  y := y + h*sin(alpha);
end;
```

Графика (АЛГО)

Начальное положение частицы:

```
x:= 200; y:= 250;
MoveTo(round(x), round(y));
```

Задать цвет линии:

```
Pen(1, 0, 255, 0);
```

толщина линии

R(red) 0..255

G(green) 0..255

B(blue) 0..255

Движение частицы:

```
for i:=1 to N do begin
  { определить новые координаты }
  LineTo(round(x), round(y));
end;
```

Задания

«4»: Постройте траектории движения двух частиц в течение 200 шагов. Частицы должны двигаться одновременно.

«5»: Постройте траектории движения 10 частиц в течение 200 шагов. **Частицы должны двигаться одновременно. Используйте массивы для хранения координат частиц.**

Системы массового обслуживания

Примеры:

1) звонки на телефонной станции
2) вызовы «скорой помощи»
3) обслуживание клиентов в банке

сколько линий?

сколько бригад?

сколько операторов?

Особенности:

1) клиенты (запросы на обслуживание) поступают постоянно, но через случайные интервалы времени
2) время обслуживание каждого клиента – случайная величина

! Нужно знать характеристики (распределения) «случайностей»!

Клиенты в банке



Вход клиентов:

1)за 1 минуту – до I_{max} человек

2)равномерное распределение

Обслуживание:

1)от T_{min} до T_{max} минут

2)равномерное распределение

?

Сколько нужно касс, чтобы клиенты стояли в очереди не более M минут?

Клиенты в банке

Число клиентов в помещении банка:

было | пришли | ушли

```
N := N + in - out;
```

! *Допущение*: клиенты распределены по кассам равномерно!

Количество касс: K

Средняя длина очереди: $\dfrac{N}{K}$

Q – длина очереди

Время ожидания: $M \le Q \cdot T_{\max}$

Допустимая длина очереди: $\dfrac{N}{K} \le Q_{\max} = \dfrac{M}{T_{\max}}$

Клиенты в банке

Пришли за очередную минуту:

округление

```
in := round(inMax*random);
```

Случайное время обслуживания:

```
T := Tmin + (Tmax - Tmin)*random;
```

! Каждый оператор за эту минуту обслужит $\frac{1}{T}$ клиентов!

Обслужены за очередную минуту и выходят:

```
out := round(K / T);
```

Клиенты в банке (программа)

период моделирования L минут

```
count := 0; { счетчик «плохих» минут }
for i:=1 to L do begin
  in := { случайное число входящих }
  out := { случайное число обслуженных }
  N := N + in - out;
  if N/K > Qmax then
    count := count + 1;
end;
writeln(count/L:10:2);
```

? Что выводится?

Клиенты в банке (исходные данные)

```
inMax := 10; { max число входящих за 1 мин }
Tmin := 1; { min время обслуживания }
Tmax := 5; { max время обслуживания }
L := 1000; { период моделирования в минутах }
M := 10; { допустимое время ожидания }
```

Задача: найти минимальное *K*, при котором время ожидания в 90% случаев не больше *M* минут.