

# Модели статистического прогнозирования (11класс)

0011 0010 1010 1101 0001 0100 1011

1 2  
4 5

## Цели урока:

- **Выяснить особенности статистических моделей,**
- **вспомнить обобщённую формулу квадратичной функции,**
- **ознакомиться с методом наименьших квадратов,**
- **изучить несколько новых функций Excel(СУММ, поиск решений).**

# Статистика- наука о сборе, измерении и анализе массовых количественных данных

- медицинская статистика
- экономическая статистика
- социальная статистика ...

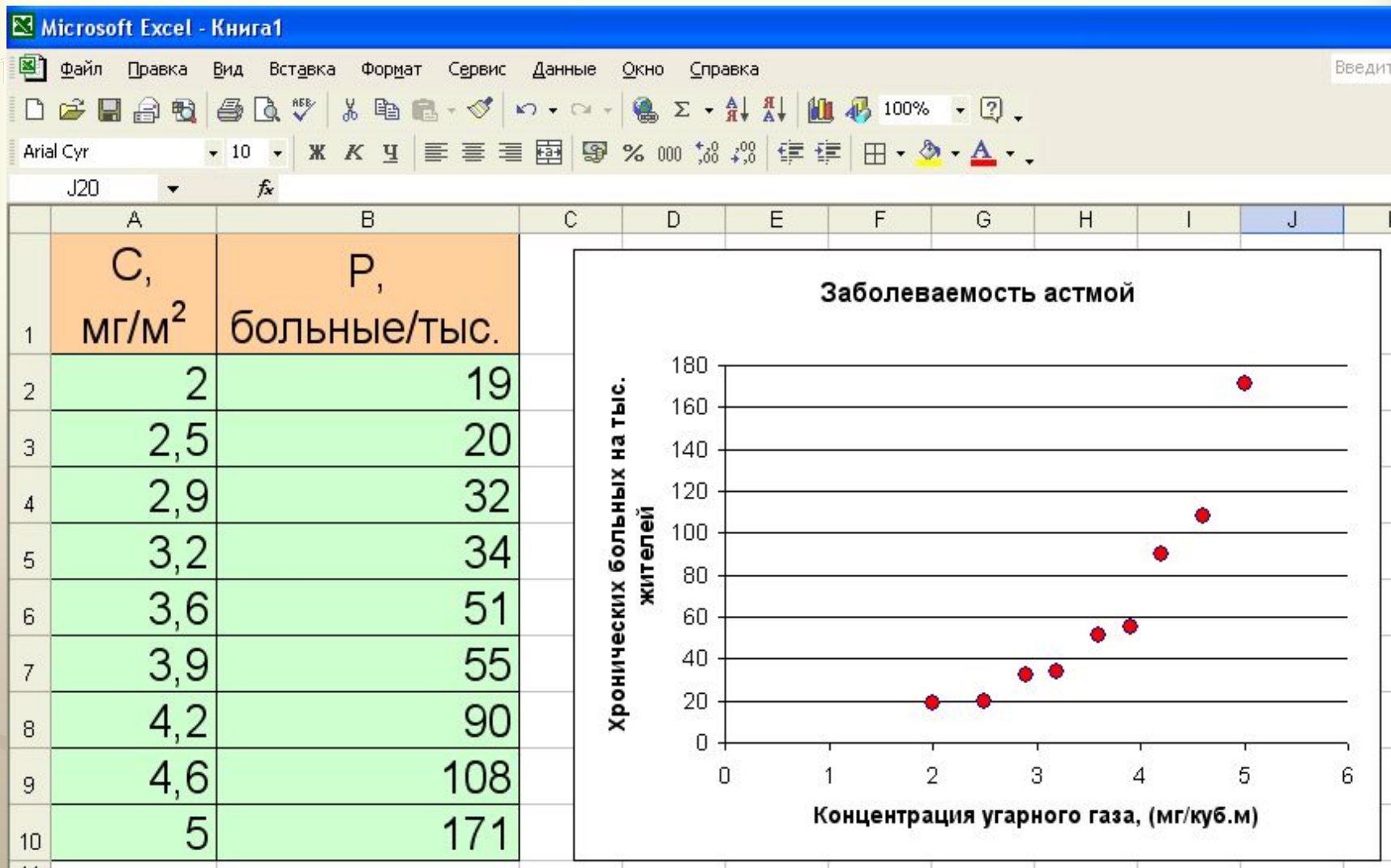
математический аппарат -  
**математическая статистика**

Зависимости устанавливаются экспериментальным путем:

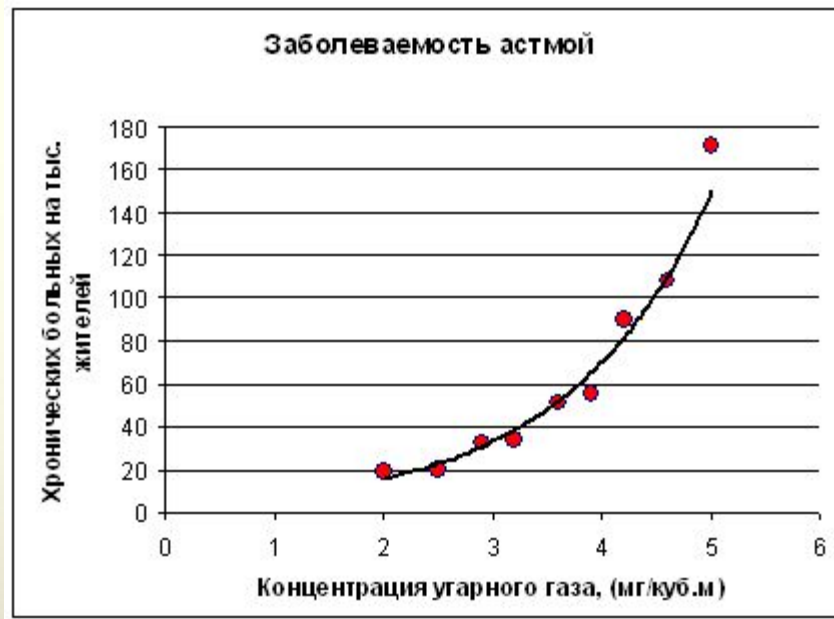
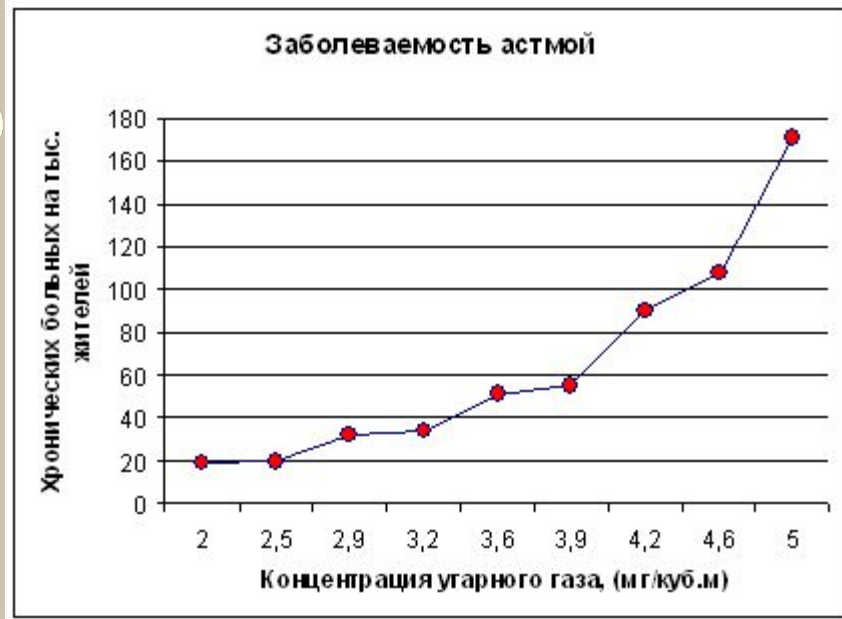
- сбор данных
- анализ
- обобщение



# Сведения о средней концентрации угарного газа в атмосфере С и о заболеваемости астмой (число хронических больных на 1000 жителей Р.



## 2 варианта построения графической зависимости по экспериментальным данным



Основные требования к искомой функции:

- она должна быть **достаточно простой** для использования ее в дальнейших вычислениях;
- график функции должен проходить **вблизи экспериментальных точек** так, чтобы отклонения этих точек от графика были минимальны и равномерны.

Полученная таким образом функция называется в статистике **регрессионной моделью**.

# Два этапа получения регрессионной модели

## 1) подбор вида функции:

$y = ax + b$  - линейная функция;

$y = ax^2 + bx + c$  - квадратичная функция (полиномиальная);

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

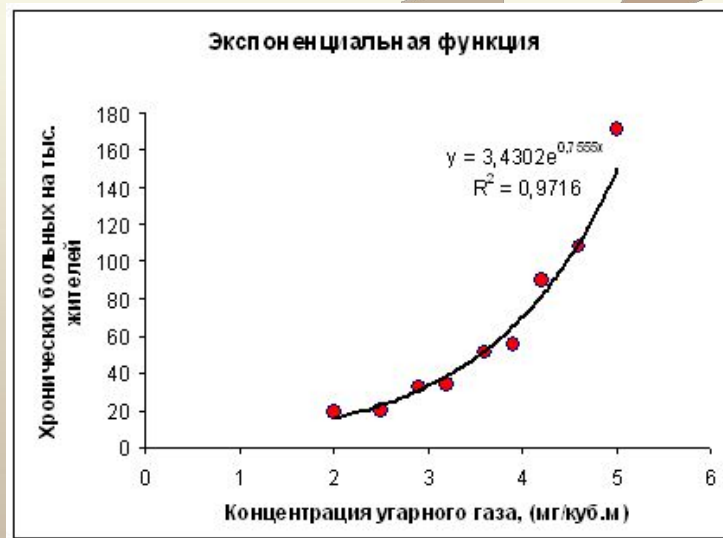
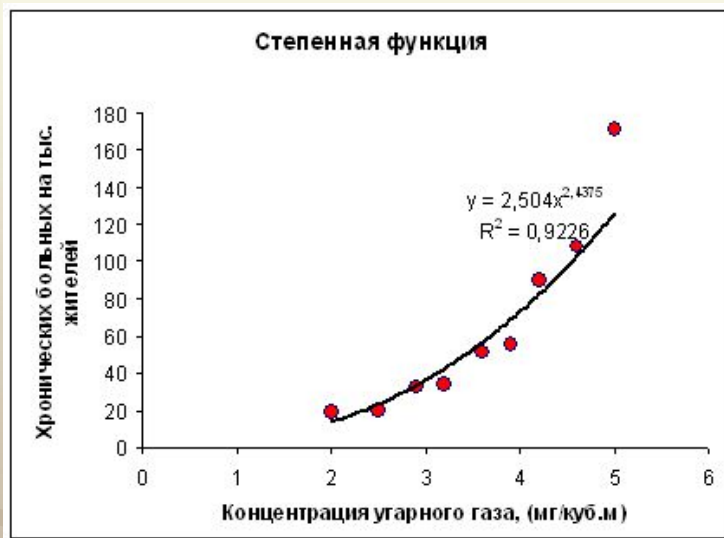
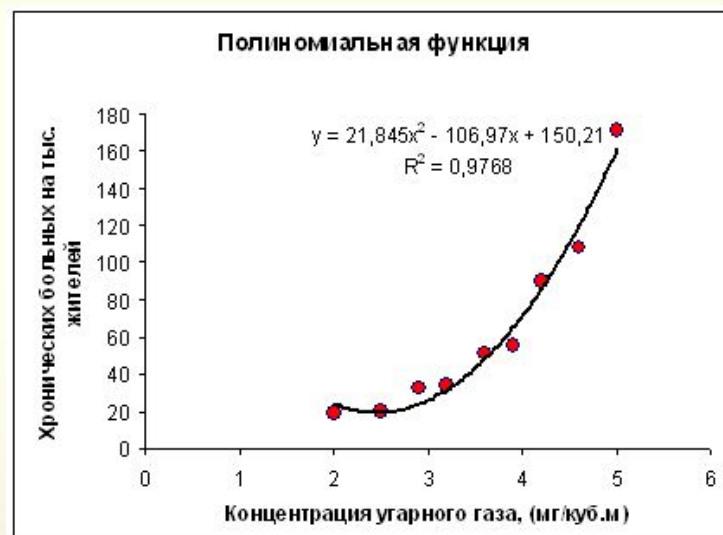
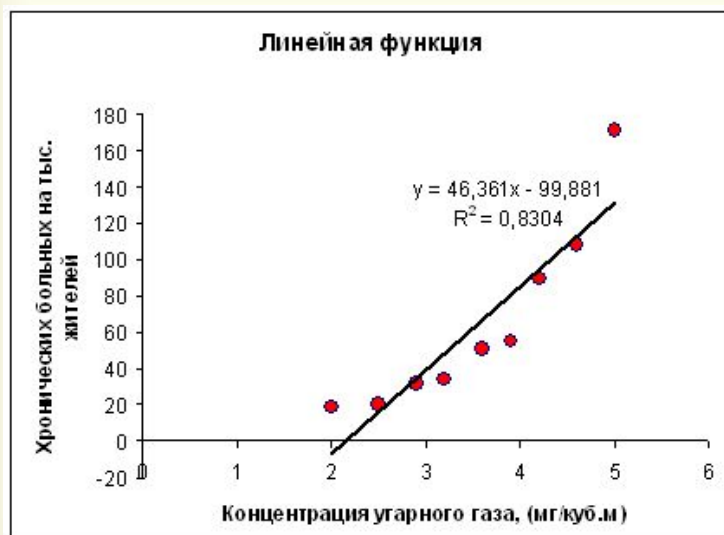
$y = ae^{bx}$  - экспоненциальная функция;

$y = ax^b$  - степенная функция.

## 2) вычисление параметров функции:

метод наименьших квадратов - *сумма квадратов отклонений у-координат всех экспериментальных точек от у-координат графика функции должна быть минимальной.*

# Графики функций, построенные по МНК, - тренды



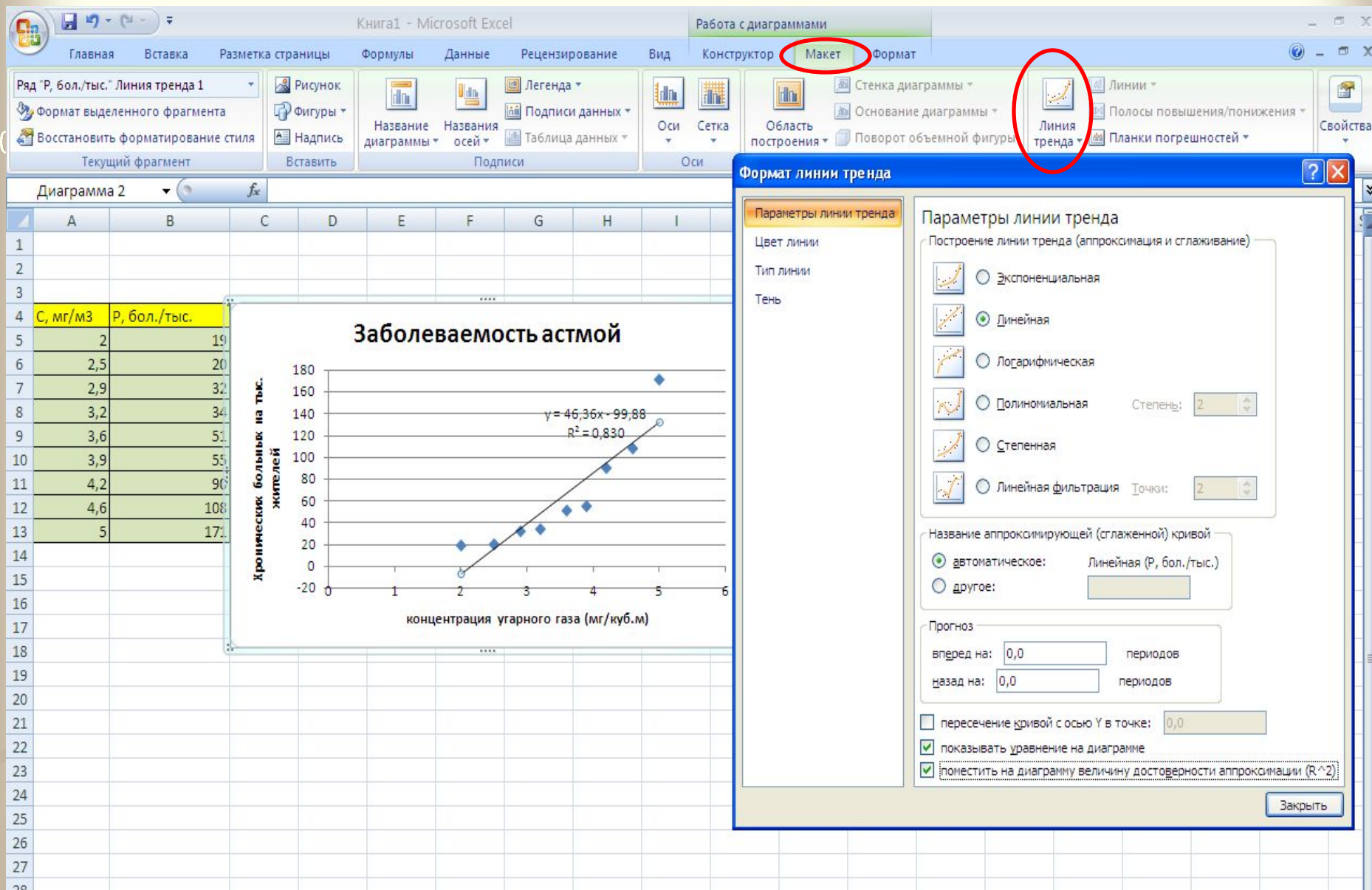
$R^2$  – коэффициент детерминированности (от 0 до 1)

# Алгоритм построения регрессионной модели по МНК с помощью MS Excel (линейный тренд)

- Ввести табличные данные зависимости заболеваемости **Р** от концентрации угарного газа **С**.
- Построить точечную диаграмму. (В качестве подписи к оси ОХ выбрать название тренда - «Линейный», остальные надписи и легенду можно игнорировать).
- Щелкнуть мышью по полю диаграммы; выполнить команду **Диаграмма – Добавить линию тренда**;
- В открывшемся окне на вкладке **Тип** выбрать **Линейный** тренд;
- Перейти на вкладку **Параметры** и установить галочки на флажках **показывать уравнения на диаграмме** и **поместить на диаграмме величину достоверности аппроксимации  $R^2$**
- щелкнуть ОК.



# Построение регрессионной модели по МНК с помощью MS Excel 2007 (линейный тренд)



# Практическая работа 3.16

## Получение регрессионных моделей в MS Excel

0011 0010 1010 1101 0001 0100 1011

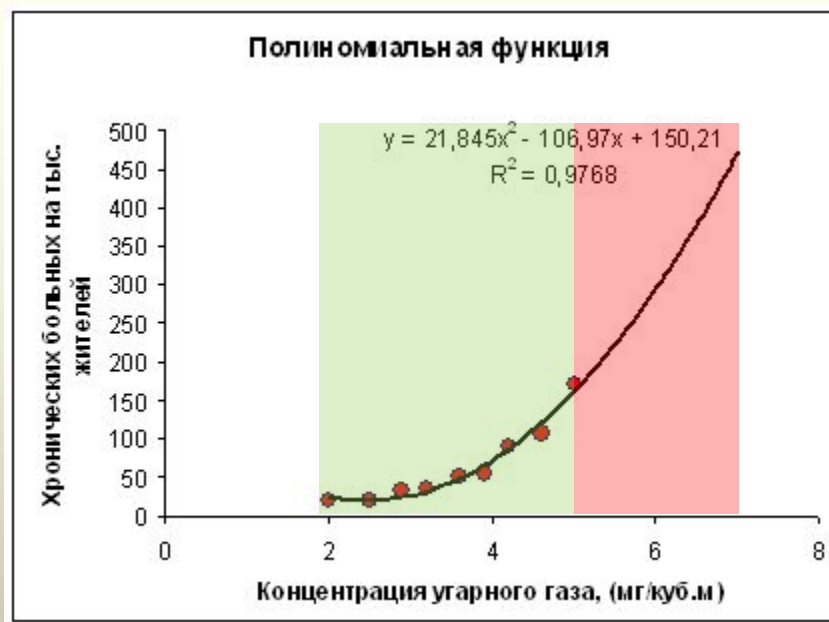
**Цель работы:** освоение способов построения по экспериментальным данным регрессионной модели и графического тренда средствами табличного процессора MS Excel.

Семакин И.Г. Практикум. Информатика и ИКТ 10-11 кл., стр.105



# Прогнозирование по регрессивной модели:

- **Восстановление значений** – прогноз в пределах экспериментальных значений независимой переменной.



- **Экстраполяция** – прогнозирование за пределами экспериментальных данных

## Ограничения при экстраполяции !

Применимость регрессионной модели ограничена,  
т.к. экстраполяция строится на гипотезе, что  
за пределами экспериментальной области  
закономерность зависимости сохраняется.

На практике – разным областям данных могут  
лучше соответствовать разные модели.

**Вывод:** применять экстраполяцию можно  
только в областях данных, близких к  
экспериментальной

## Используемая литература

- И.Г.Семакин и др. Информатика 10-11. Практикум, М.: Бином. Лаборатория знаний, 2012
- И.Г.Семакин и др. Информатика 11. Базовый уровень, М.: Бином. Лаборатория знаний, 2013

