What is mathematical modeling?

Konstantin Avilov
researcher at INM RAS
IMMC Expert Panel member

Mathematical modeling and school

- In Russian schools, mathematical modeling is not taught as a separate subject. Mathematics, physics, chemistry, programming and so on are studied.
- In fact, all natural sciences are one, and mathematics and computer science are just instruments created at the request of physics and other “practical” sciences.
- Mathematical modeling is essentially the method of all quantitative research in all sciences.
- From the school point of view, mathematical modeling is the rejoining of all natural sciences into a single “science”.

Mathematical modeling and natural sciences

- Any research looking for a connection between measurable (quantitative) parameters of something generates a mathematical model.
- Mathematical model is always a certain simplification, a certain conceptualization that disregards some properties of reality.
- The study of an *adequate/valid* mathematical model allows one to determine some of the properties of the simulated object.
- The adequacy/validity of mathematical models is an extremely difficult question.

Building mathematical models

Strategies for building mathematical models:

- **"From theory"**: in areas where almost everything is already known and there are well-confirmed theories, mathematical models are constructed as a special case of the theory. Usually there is a "construction set" of primitives that allows one to "assemble" a mathematical model.

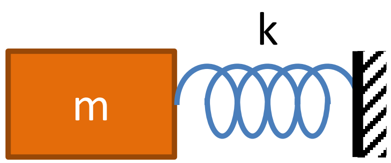
    Examples: classical mechanics and other classical physics

- **"From data":** in areas where there are no general theories, or it is not clear what is happening in the experiment, mathematical models are built as a tool for finding dependencies in the data.

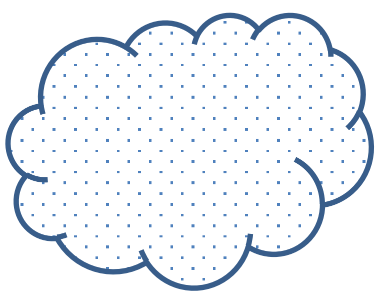
    Examples: biology / medicine, cutting edge of all natural sciences

Types of “simple” mathematical models

1) Algebraic models:



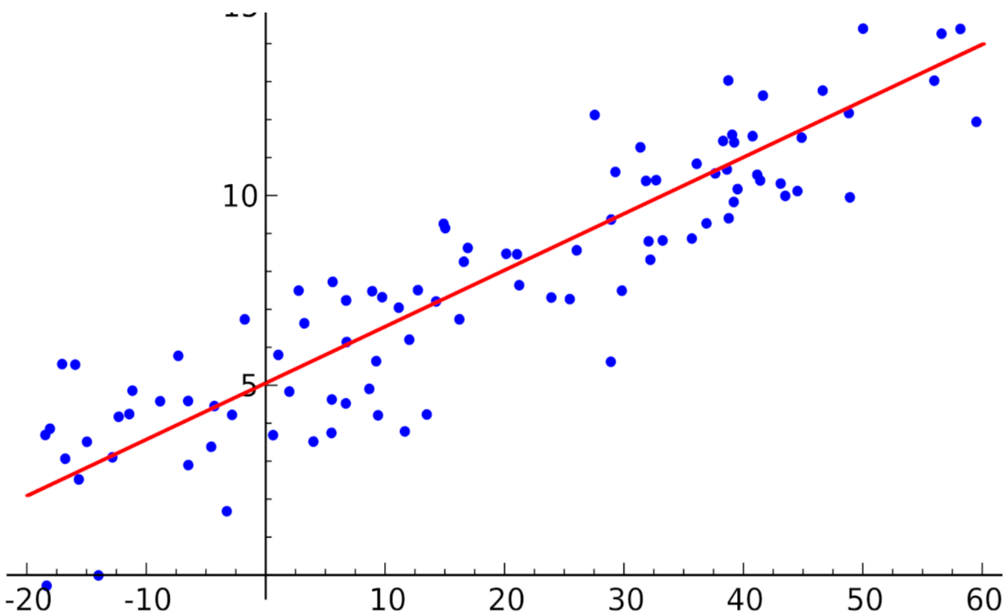
$ma = F = kx$



$PV = vRT$

Types of "simple" mathematical models

2) Regressions



Data: (x_i, y_i), $i=1\ldots N$

Regression:

$$y=f(x,\Theta)+\varepsilon$$

Θ - parameters, ε - error

Linear regr.: $\boldsymbol{f=a+bx}$,
a и b – parameters (Θ)

Quadratic regr.: $\boldsymbol{f=a+bx+cx^2}$,
a, b и c – parameters (Θ)

There exist standard methods of calculating the optimal parameters Θ from data (x_i, y_i) ("regression fitting methods", "least squares method", etc.)

Types of “simple” mathematical models

3) Evolutionary models: when something changes with time

- Such models are often described by differential equations
- Discrete-algebraic version

$$\boldsymbol{X(t+1)=f(X(t), \Theta, \ldots)}, \quad t \text{ – time}$$

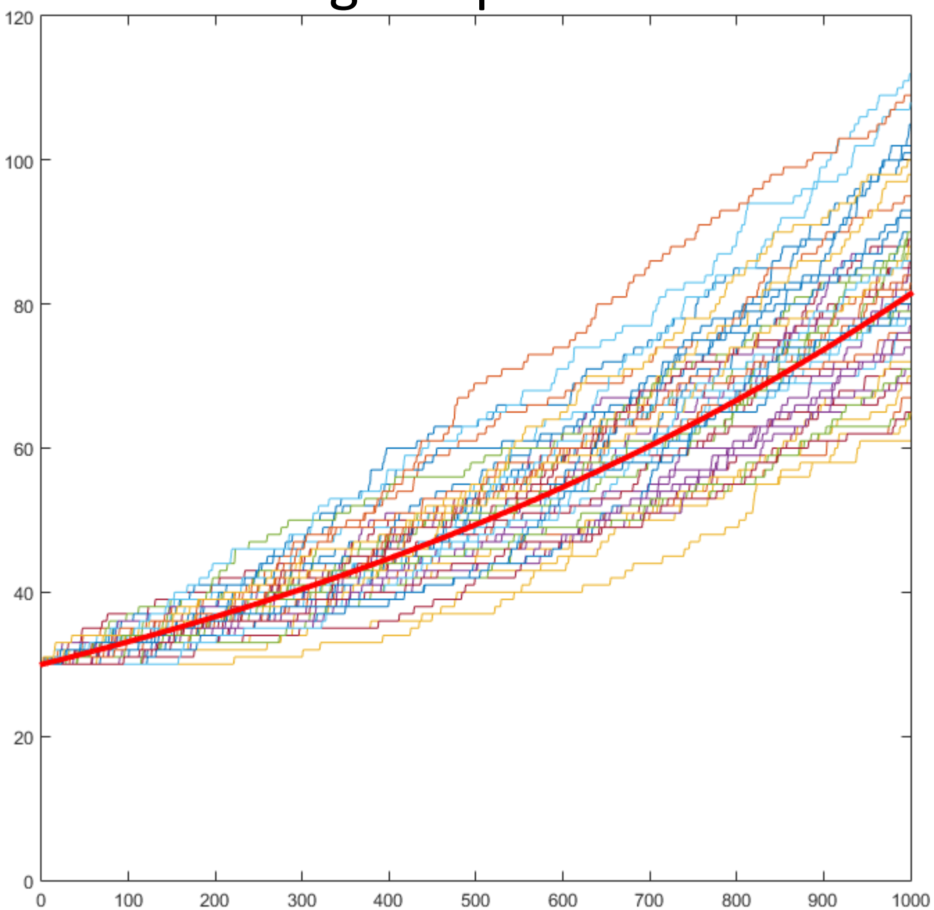
Example (Malthus model - growth or extinction of a population with a specific rate a):

$$\boldsymbol{X(t+1)=(1+a)X(t)}$$

Types of “simple” mathematical models

4) Stochastic models (models with randomness)

- Purely probabilistic models (based on probability theory) - calculating the probabilities of events based on distributions



- Stochastic versions of evolutionary models: instead of the rates of the processes, their probabilities are set at each step

On the chart: bold line - deterministic Malthus,

the rest of the lines are “stochastic Malthus” with the same growth rate *a.*

Types of “simple” mathematical models

5) Agent-based (simulational) models

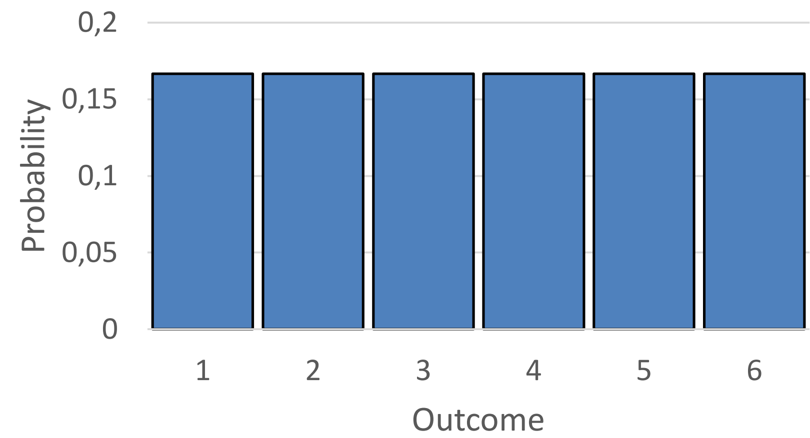
- A “computer game” is built in which some entities (agents) interact and create some kind of complex collective behavior.
- Agent models allow one to describe complex patterns of interaction between agents.
- Examples (in NetLogo):
  - formation of flocks of birds based on simple flight rules for each bird
  - spread of infections in structured populations

Random variables-1

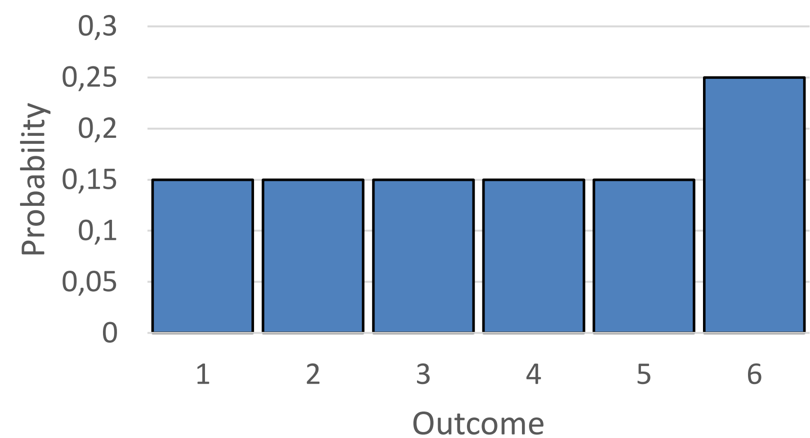
- **Random variable** is the basic concept of probability theory (which is not taught in school). At the “common sense” level, this is a kind of process, by repeating which you can get different values (for example, throwing a dice or a coin).
- If the process is “unchanged”, then with a very large number of independent repetitions, the frequencies of different **outcomes** (“values”) stabilize and converge to what is called the **distribution of the random variable**. It characterizes the RV: “with what probability do we get each of the possible outcomes in a single experiment?"
- The more real measurements are collected, the closer the **sample frequencies** measured from them are to the true distribution.

Random variables-2 (distributions)

- “Fair” 6-sided dice:
  outcomes 1,2,3,4,5,6 with equal probabilities 1/6=16.(6)%



- “Unfair” 6-sided dice :
  outcomes 1,…,5 with probab-es 15%
  outcome 6 with probability 25%



Random variables-3 (mean)

- **The expected value of a random variable (mean)** is the value that will be obtained if a lot of independent values of a random variable are obtained and averaged.
- Calculating the mean: sum up the products of the outcome values by their probabilities:
  - x_i – *i*-th possible outcome
  - p_i – the probability of the *i*-th outcome
  - *Mx* – the mean of *x*

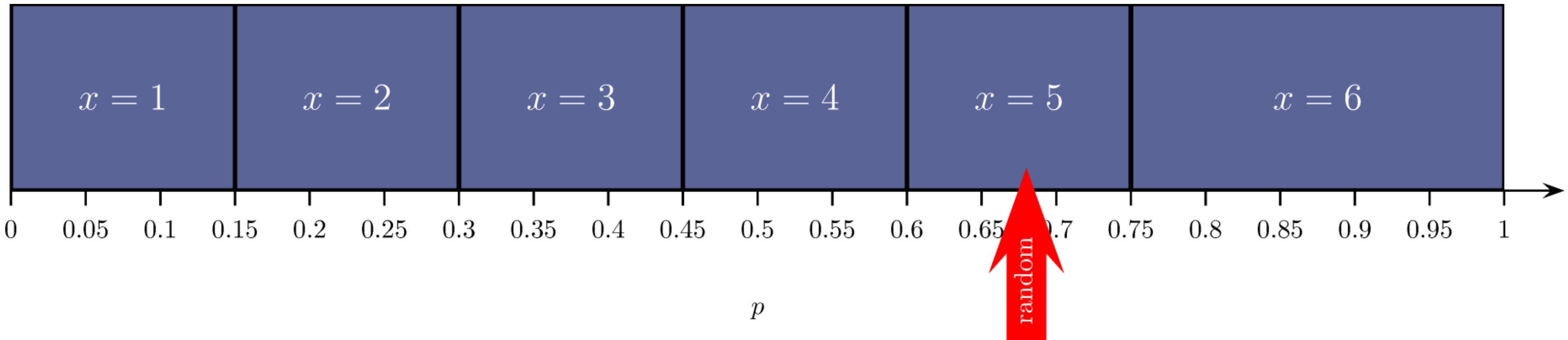
$$Mx = \sum_i x_i p_i$$

“Fair” dice: *Mx*=(1*1/6 + 2*1/6 + … + 6*1/6) = 3.5

“Unfair” dice : *Mx*=(1*0,15 + … + 5*0,15 + 6*0,25) = 3.75

Random variables-4 (numerical)

- It is possible to calculate the parameters of random variables numerically (using the Monte Carlo method): generate many realizations of random variables according to a given distribution, and then simply calculate the parameters (e.g., the mean) from the synthetic data obtained.
- In programming languages, there is usually a function (rand, random, rnd, etc.) that generates a pseudo-random number distributed evenly from 0 to 1 or from 0 to MAX_INT.
- To generate a discrete RV with a known distribution, you just need to place the value of *rnd* onto the cumulative distribution:



How to build a mathematical model

Creating a mathematical model is similar to solving a problem in physics:

1. Understand “what is happening in general”, what “variables” describe the situation/object in hand (i.e., create a “conceptual model”).
2. Determine the laws/rules that govern the relationships between the “variables”. It can be both previously known laws, or some approximate rules with unknown parameters.
3. If there are unknown parameters, estimate them from the available data (“fit the model to the data” or evaluate the parameters directly from the data).
4. Study the properties of the obtained mathematical model (descriptions of relationships between the “variables”), get useful conclusions about the properties of the object of modeling.

What should be in the MMC paper

- Many important requirements for the MMC paper are listed in the Contest Rules. Read them carefully.
- The main aim of the article is a clear description of the work done and the results obtained.
- The article should not contain unnecessary “padding”, this confuses the reader and does not allow them to see your real achievements.
- One of the most important criteria for evaluating a paper is the validity of the research decisions made in it: you should explain why you did it this way, and not otherwise (except for obvious decisions, of course). Even if you know a “weakness” of a “modeling move”, you should write about it openly.
- A clear statement of the results.
- Russian schools do not teach “scientific writing”. But one always has to start somewhere and try to write a coherent text.

Thanks for your attention!