

Что такое математическое моделирование?

Константин Авилов
научный сотрудник ИВМ РАН
член Expert Panel IMMC

(Работа выполнена при поддержке гранта РФФИ №19-29-14217)

What is mathematical modeling?

Konstantin Avilov
researcher at INM RAS
IMMC Expert Panel member

(The work was funded by RFBR, grant number 19-29-14217)

Математическое моделирование и школа

- В российской школе, как считается, не изучают математическое моделирование. Но изучают математику, физику, химию, программирование и т.д.
- На самом деле все естественные науки едины, а математика и информатика – это инструменты, созданные по запросам физики и других «практических» наук.
- Математическое моделирование – это, по сути, метод всех количественных исследований в любых науках.
- С точки зрения школы, математическое моделирование – это обратное соединение всех естественных наук в единую «science».

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”.

Математическое моделирование и естественные науки

- Любое исследование, ищущее связь между измеримыми (количественными) параметрами чего-либо, порождает математическую модель.
- Математическая модель – это всегда некоторое упрощение, некоторая концептуализация, отсекающая какие-то свойства реальности.
- Изучение *адекватной* математической модели позволяет узнать некоторые свойства объекта моделирования.
- Адекватность мат.моделей – вопрос крайне сложный.

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.

Построение математических моделей

Стратегии построения математических моделей:

- **«От теории»:** в областях, где уже почти всё известно и имеются хорошо подтвержденные теории, мат.модели строятся как частный случай теории. Обычно есть «конструктор» из примитивов, позволяющий «собирать» мат.модель.

Примеры: классическая механика и пр. классическая физика

- **«От данных»:** в областях, где общих теорий нет, или непонятно, что происходит в эксперименте, мат.модели строятся как инструмент поиска зависимостей в данных.

Примеры: биология/медицина, «передний край» любых естественных наук

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.

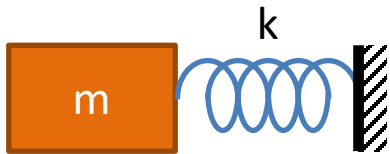
Examples: classical mechanics and other classical physics

- **“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

Типы «простых» математических моделей

1) Алгебраические модели:



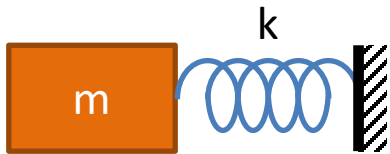
$$ma = F = kx$$



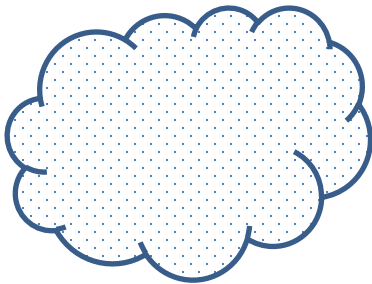
$$PV = \nu RT$$

Types of “simple” mathematical models

1) Algebraic models:



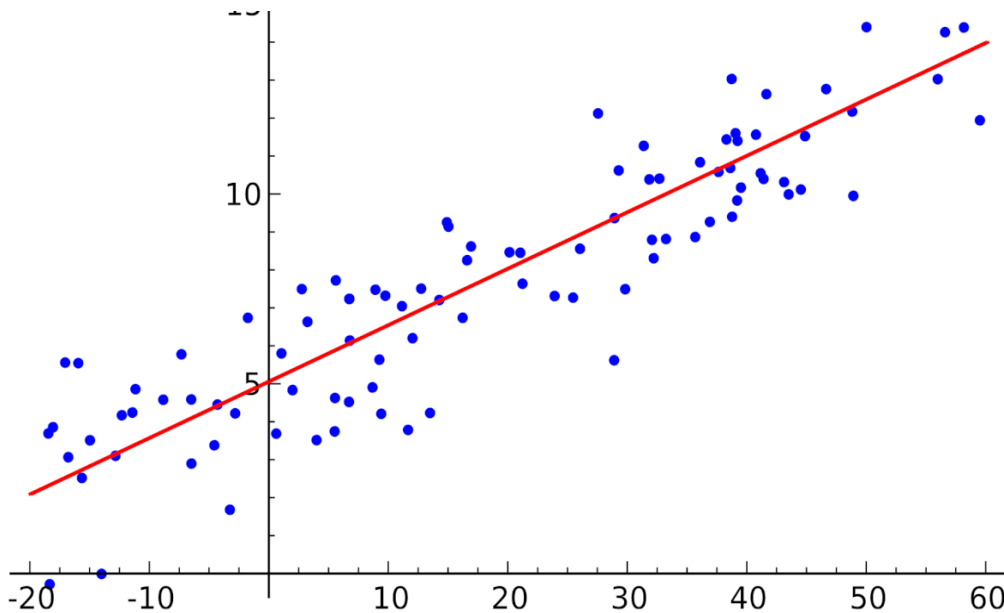
$$ma = F = kx$$



$$PV = \nu RT$$

Типы «простых» математических моделей

2) Регрессии



Данные: $(x_i, y_i), i=1...N$

Регрессия:

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

θ - параметры, ε - ошибка

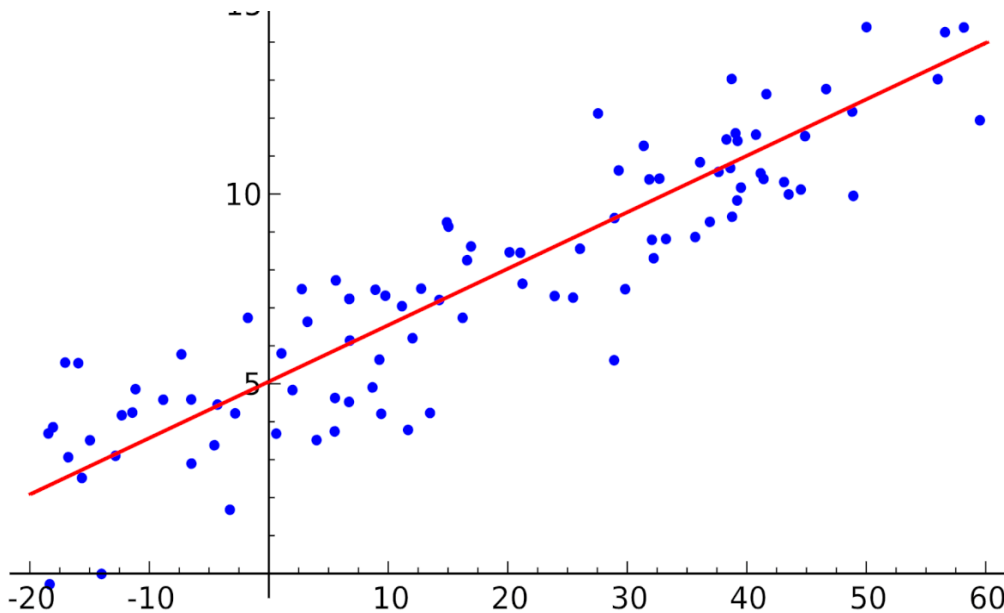
Линейная рег.: $f=a+bx$,
 a и b – параметры (θ)

Квадратичная рег.: $f=a+bx+cx^2$,
 a, b и c – параметры (θ)

Существуют стандартные методы вычисления наилучших параметров θ по данным (x_i, y_i) («методы подгонки регрессии», «метод наименьших квадратов», и т.п.)

Types of “simple” mathematical models

2) Regressions



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

Regression:

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

θ - parameters, ε - error

Linear regr.: **$f=a+bx$** ,

a и b – parameters (θ)

Quadratic regr.: **$f=a+bx+cx^2$** ,

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.)

Типы «простых» математических моделей

3) Эволюционные модели: когда что-то изменяется во времени

- Часто такие модели описываются дифференциальными уравнениями
- Дискретно-алгебраический вариант

$$X(t+1)=f(X(t), \theta, \dots), \quad t - \text{время}$$

Например (модель Мальтуса – рост или вымирание популяции с удельной скоростью a):

$$X(t+1)=(1+a)X(t)$$

Types of “simple” mathematical models

- 3) Evolutionary models: when something changes with time
- Such models are often described by differential equations
 - Discrete-algebraic version

$$X(t+1)=f(X(t), \theta, \dots), \quad t - \text{time}$$

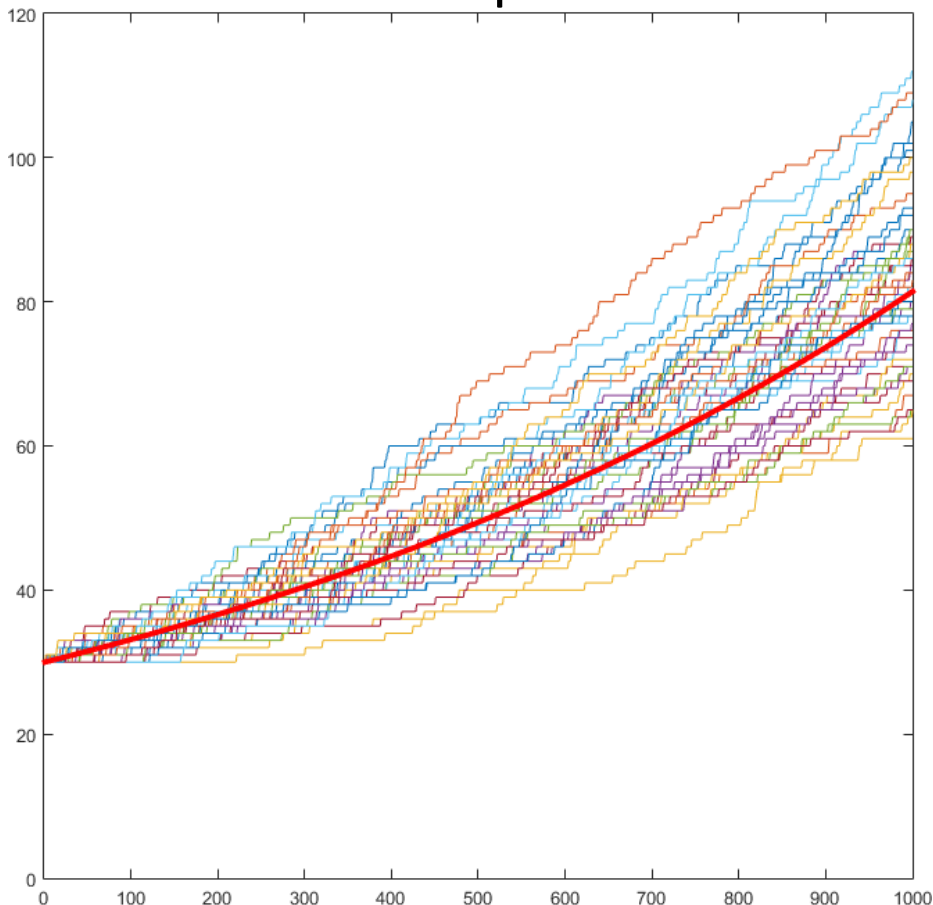
Example (the Malthus model: the growth or extinction of population with a specific rate a):

$$X(t+1)=(1+a)X(t)$$

Типы «простых» математических моделей

4) Стохастические модели (модели с элементом случайности)

- Чисто вероятностные модели (на основе теории вероятностей) – вычисление вероятностей событий на основе распределений



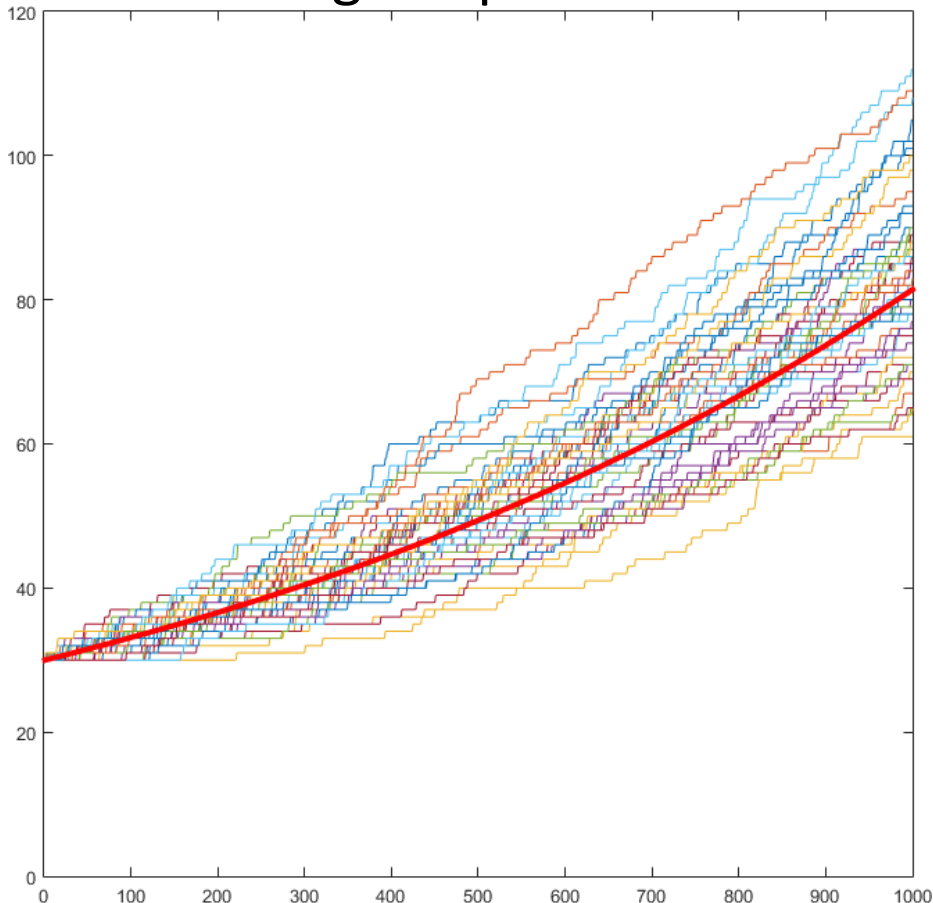
- Стохастический вариант эволюционных моделей: вместо скоростей процессов задаются их вероятности на каждом шаге

Жирная линия на графике – детерминированный «Мальтус», остальные линии – «стохастический Мальтус» с той же скоростью роста a .

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 processes, their probabilities at each step are set

The bold line on the chart shows the “deterministic Malthus,” other lines represent the “stochastic Malthus” with the same growth rate α .

Типы «простых» математических моделей

5) Агентные (симуляционные) модели

- Создается «компьютерная игра», в которой некоторые сущности (агенты) взаимодействуют и создают какое-то сложное общее поведение.
- Агентные модели позволяют отобразить сложные схемы взаимодействия между агентами.
- Примеры (NetLogo):
 - формирование стай птиц на основе простых правил полёта каждой птицы
 - распространение инфекций в структурированных популяциях

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

Случайные величины-1

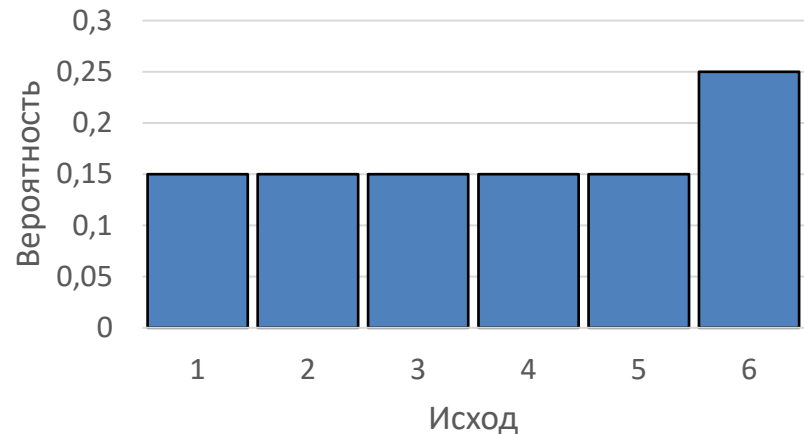
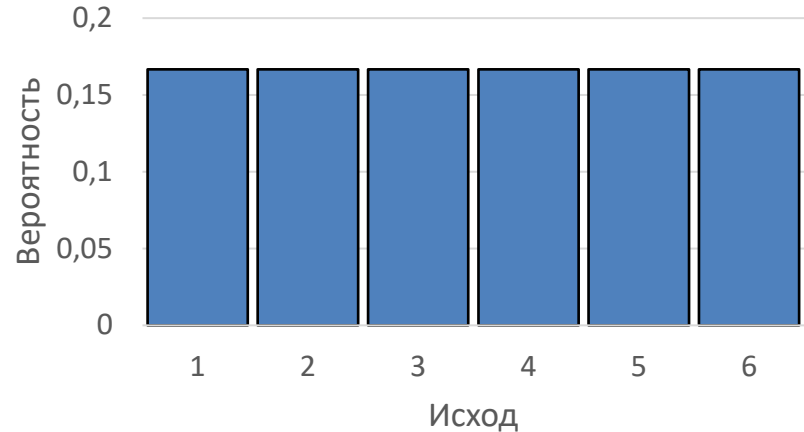
- **Случайная величина** – основное понятие теории вероятностей (которую не рассказывают в школе). На «бытовом» уровне, это некий процесс, повторяя который можно получать разные значения (напр., броски игрального кубика или монеты)
- Если процесс «неизменен», то при очень большом количестве независимых повторений частоты разных **исходов** («значений») стабилизируются и сходятся к тому, что называется **распределением случайной величины** и характеризует собой её: «с какой вероятностью мы получим каждый из возможных исходов в единичном эксперименте».
- Чем больше собрать реальных измерений, тем ближе измеренные по ним **выборочные частоты** исходов к истинному распределению.

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.

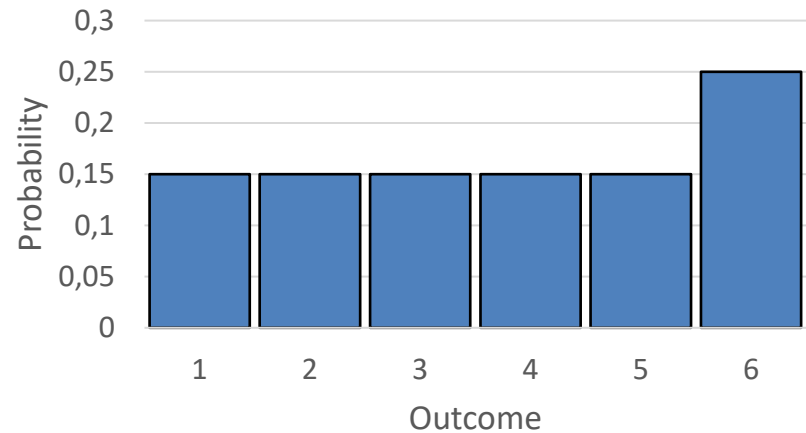
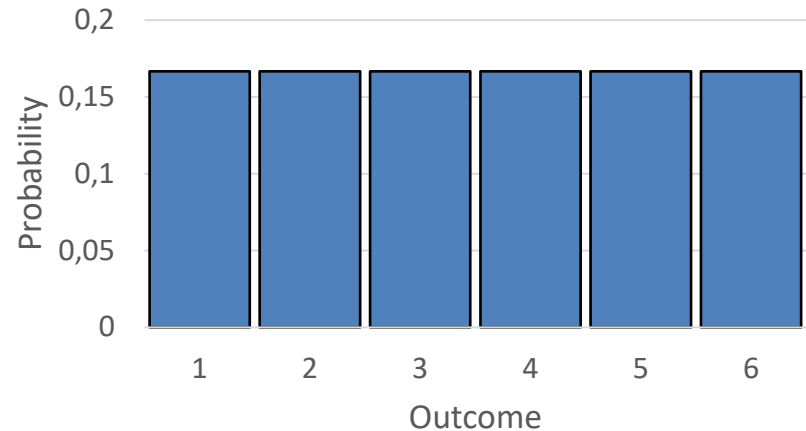
Случайные величины-2 (распределения)

- «Честный» шестигранный кубик:
исходы 1,2,3,4,5,6 с равными
вероятностями $1/6=16,6\%$
- «Нечестный» шестигранный кубик:
исходы 1,...,5 с вероятностями 15%
исход 6 с вероятностью 25%



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%



Случайные величины-3 (среднее)

- **Среднее значение случайной величины (математическое ожидание)** – то значение, которое получится, если получить очень много независимых значений случайной величины и усреднить.
- Подсчёт мат. ожидания: суммирование произведений значений исходов на их вероятности:
 - x_i – i -ый возможный исход
 - p_i – вероятность i -го исхода
 - Mx – мат. ожидание сл. величины x

$$Mx = \sum_i x_i p_i$$

Для «честного» кубика: $Mx = (1 \cdot 1/6 + 2 \cdot 1/6 + \dots + 6 \cdot 1/6) = 3,5$

Для «нечестного» кубика: $Mx = (1 \cdot 0,15 + \dots + 5 \cdot 0,15 + 6 \cdot 0,25) = 3,75$

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

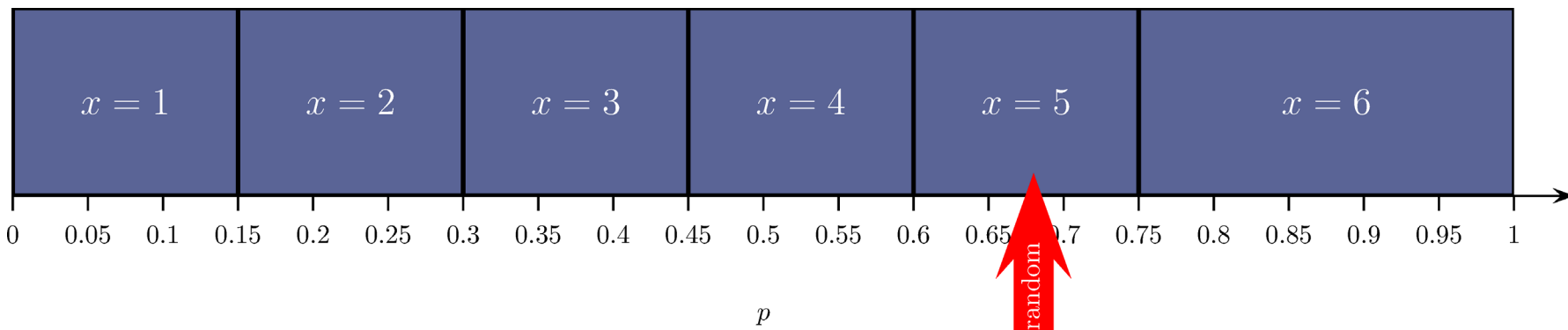
$$Mx = \sum_i x_i p_i$$

“Fair” dice: $Mx = (1 \cdot 1/6 + 2 \cdot 1/6 + \dots + 6 \cdot 1/6) = 3.5$

“Unfair” dice : $Mx = (1 \cdot 0,15 + \dots + 5 \cdot 0,15 + 6 \cdot 0,25) = 3.75$

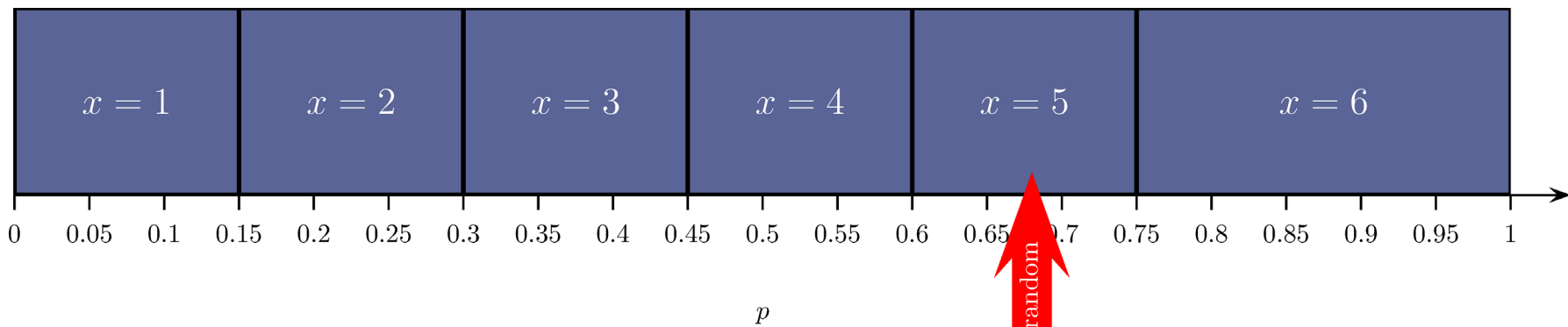
Случайные величины-4 (численные)

- Вычислять параметры случайных величин можно и численно (методом Монте-Карло): генерировать много реализаций случайных величин по заданному распределению, а затем «в лоб» считать какие-то параметры (напр., среднее) по так полученным синтетическим данным.
- В языках программирования обычно есть функция (rand, random, rnd и т.д.), генерирующая псевдослучайное число, равномерно распределённое от 0 до 1 или от 0 до MAX_INT.
- Для генерации дискретной сл. величины с известным распределением надо просто отсчитать значение rnd по кумулятивному распределению:



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:



Как создавать мат. модель

Создание мат. модели похоже на решение задачи по физике:

1. Надо понять, «что вообще происходит», какими «переменными» описывается ситуация/объект (т.е. создать «концептуальную модель»).
2. Определить законы/правила взаимосвязи между «переменными». Это могут быть и ранее известные законы, и какие-то приблизительные правила с неизвестными параметрами.
3. Если есть неизвестные параметры, оценить их из имеющихся данных («настройка модели на данные» или прямая оценка параметров из данных).
4. Изучить свойства полученного математической модели (описания связей между «переменными»), получить полезные выводы о свойствах объекта изучения.

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.

Что должно быть в Мамонт-статье

- Многие важные требования к Мамонт-статье описаны в Правилах конкурса. Прочтите их внимательно.
- Главная задача статьи – внятное описание проделанной работы и полученных результатов.
- В статье не должно быть излишней «воды», это запутывает читателя и не дает ему увидеть ваши реальные достижения.
- Один из главнейших критериев оценки работы – это обоснованность принятых в ней исследовательских решений: следует объяснять, почему вы сделали так, а не иначе (кроме очевидных решений, разумеется). Даже если вы знаете «слабость» какого-то решения, об этом надо писать прямо.
- Внятная формулировка результатов.
- В российских школах не обучают «scientific writing». Но всегда надо когда-то начинать и пытаться написать связный текст.

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.

Спасибо за внимание!