

Что такое математическое моделирование?

Константин Авилов

научный сотрудник ИВМ РАН

член Expert Panel IMMS

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».

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.

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.

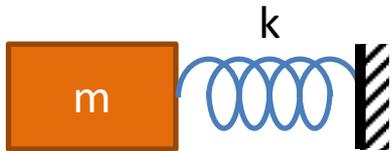
Примеры: классическая механика и пр. классическая физика

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

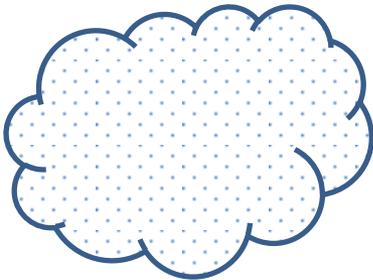
Примеры: биология/медицина, «передний край» любых естественных наук

Типы «простых» математических моделей

1) Алгебраические модели:



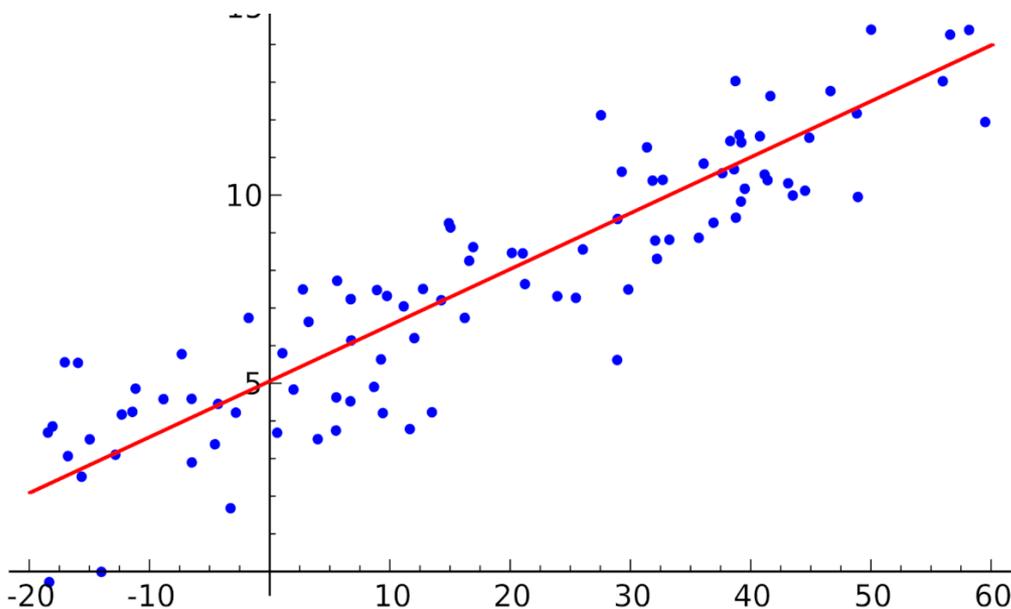
$$ma = F = kx$$



$$PV = \nu RT$$

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.)

Типы «простых» математических моделей

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

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

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

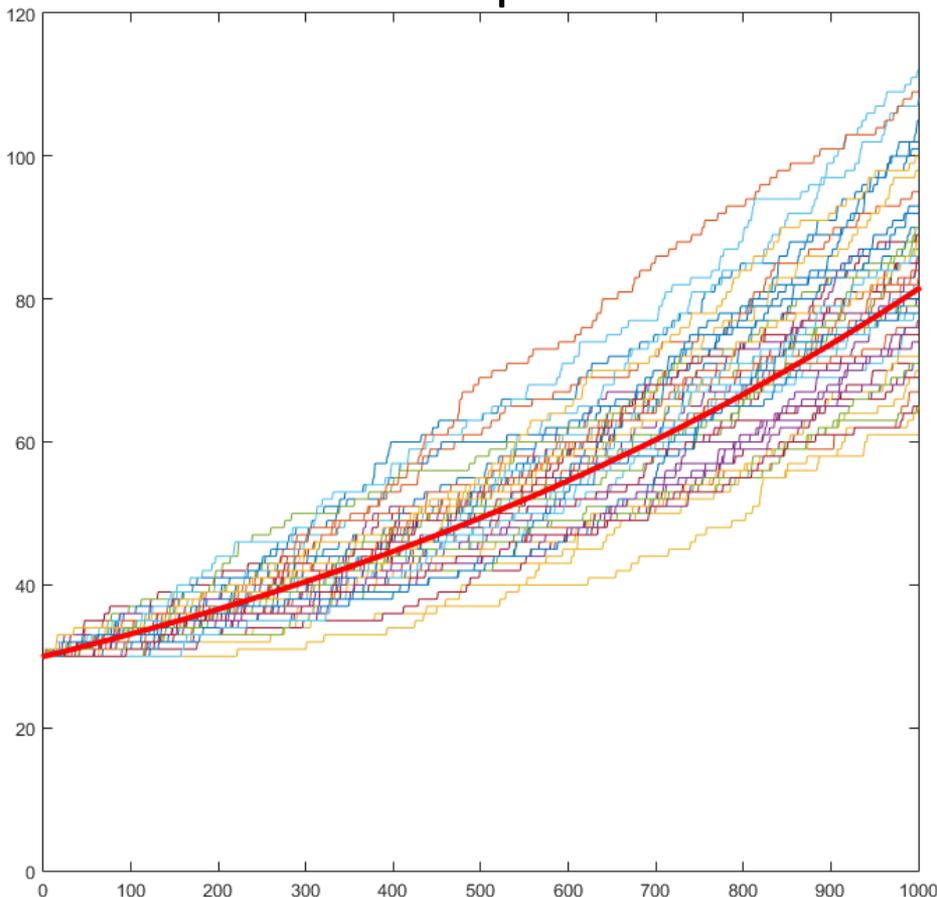
Например (модель Мальтуса – рост или вымирание популяции с удельной скоростью a):

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

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

На графике: жирная линия – детерминированный «Мальтус», остальные линии – «стох.Мальтус» с той же скоростью роста a .

Типы «простых» математических моделей

5) Агентные (симуляционные) модели

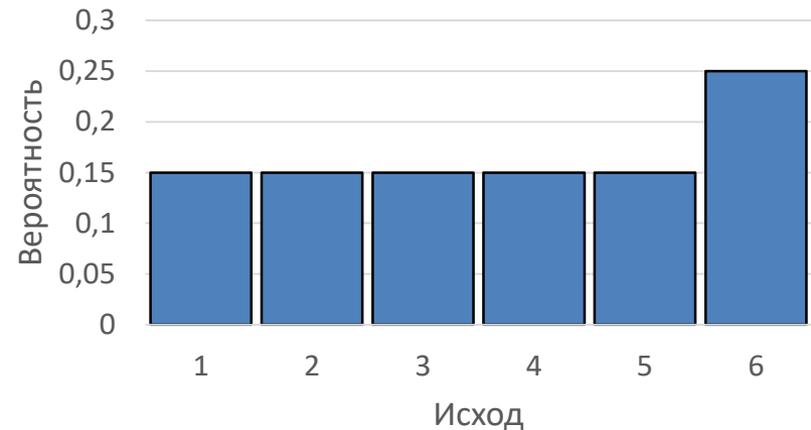
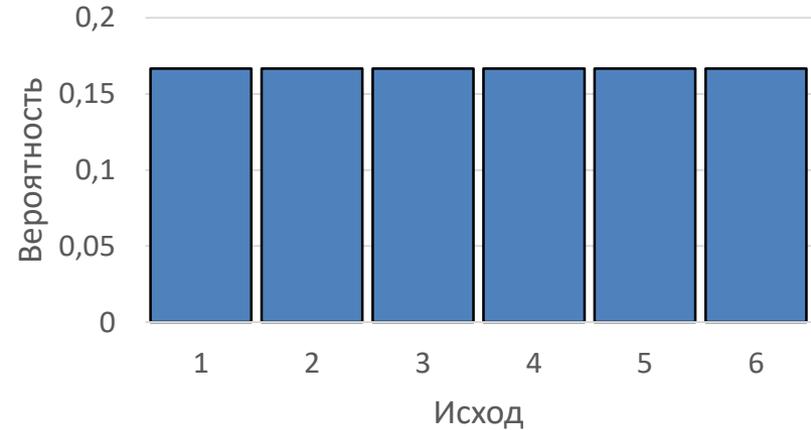
- Создается «компьютерная игра», в которой некоторые сущности (агенты) взаимодействуют и создают какое-то сложное общее поведение.
- Агентные модели позволяют отобразить сложные схемы взаимодействия между агентами.
- Примеры (NetLogo):
 - формирование стай птиц на основе простых правил полёта каждой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

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.

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%



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

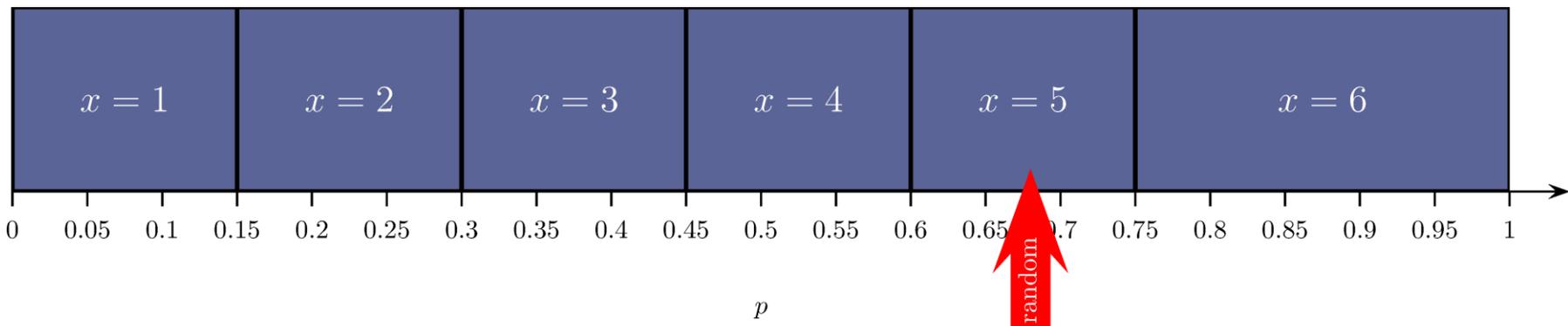
$$Mx = \sum_i x_i p_i$$

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

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

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:



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.

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.

Спасибо за внимание!