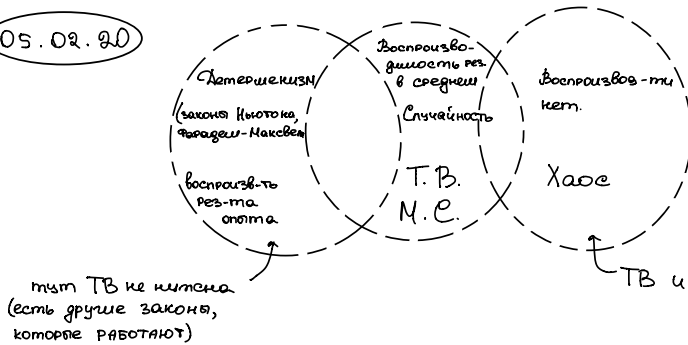


# Шевшиков Георгий Леонидович.

05.02.20



ТВ - теория вероятности  
МС - математическая статистика

Laplace, Gauss, Чебышев, Марков, Колмогоров, Карл Пирсон, ...

## § Формы представления данных и пути их анализа

1. Выборка вещ. чисел  $x_1, \dots, x_n$   $x_i \in \mathbb{R}$
2.  $X_1, X_2, \dots, X_n$  - числа / векторы (присоединяем или не присоединяем)  
 $X_i \in \mathbb{R}^m$   $X_i = (x_{i1}, \dots, x_{im})^T$
3. Континуальные реализации во времени  $x(t), t \in [0, T]$
4. Факторы нечисловой природы: анкетные данные, ..., семантики.

МС имеет дело с 1-3 т.е. это числа, функции.

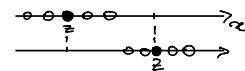
Что будем делать с данными?

## План статистических кар-к

### 1. Кар-ки средних значений

какие примеры выборки:  $x_1, \dots, x_n$   $x_i \in \mathbb{R}$

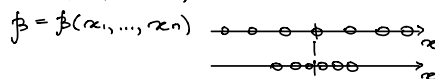
какое значение на выборке  $z \in \mathbb{R}$   $z = z(x_1, \dots, x_n)$



### 2. Кар-ки рассеивания данных

$x_1, \dots, x_n$ ,  $x_i \in \mathbb{R}$

(среднее число одно, а рассеивание разное)



### 3. Корреляционные кар-ки. (взаимосвязи)

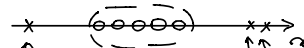
$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$



### 4. Кар-ки экстремальных значений

$x_1, \dots, x_n$   $x_i \in \mathbb{R}$

(надо знать методы их выявления)



подозрительные числа (выбросы) которые тоже нужно изучать

5. кар-ки распределений (важна природа величин)  
\* выборка  $x_1, \dots, x_n$   $x_i \sim F(x), f(x)$   
хотим получить оценки, плотности  $\hat{F}(x), \hat{f}(x) = ?$

6. кар-ки спектров данных  
преобраз-я Фурье, ...  
Имеется выбор преобр-я

### Цели анализа данных

1. Компактное представление данных
2. Оценка параметров моделей массовых явлений
3. Прогнозирование

### § Кар-ки средних одномерных данных