Министерство образования Республики Беларусь

Учреждение образования

«Брестский государственный технический университет»

Кафедра ИИТ

Лабораторная работа №7

за 7 семестр

По дисциплине: «КМиАД»

Выполнил:

Студент 4 курса

Группы ПО-4(2)

Тупик Д. Л.

Проверил:

Чичурин А. В.

2022

**Статистический анализ данных**

Wolfram Language объединяет многие аспекты статистического анализа данных, от получения и изучения данных до построения высококачественных моделей и вывода последствий.

Wolfram Language предоставляет несколько способов получения данных, начиная со встроенных кураторских источников данных, импорта из различных форматов файлов или подключения к базам данных. Базовая обработка данных, включая вычисление статистических величин, сглаживание, тестирование и визуализацию, дает первый уровень анализа.

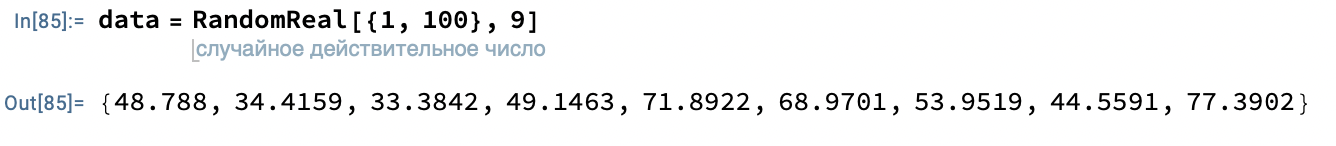
Добавляя в микс модели, такие как модели распределения или регрессии, можно ответить на более широкий круг вопросов анализа или даже предоставить возможности прогнозирования.

1. **Элементарная описательная статистика**

Описательные статистические функции Wolfram Language работают как на явных данных, так и на символьных представлениях статистических распределений.

При работе с явными данными функции регулярно обрабатывают огромные наборы данных, которые могут содержать не только числа, но и символьные элементы, представляющие, например, параметризованные или неизвестные данные.

Наиболее часто статистические программы используются для генерации случайных чисел и массивов данных с ними. Для отдельных псевдослучайных реальных чисел предусмотрена функция RandomReal[]. При каждом обращении к этой функции генерируется случайное (точнее, псевдослучайное) число в интервале от 0 до 1 с равномерным распределением. Функция RandomReal[xmin, xmax] генерирует случайное число в заданном интервале изменения переменной x.



**Mean**

При наличии списка с n элементами  *среднее* [значение Mean](https://reference.wolfram.com/language/ref/Mean.html)[*list*] определяется как Рисунок 16.



**Mediana**

*Медиана* [Median](https://reference.wolfram.com/language/ref/Median.html)[*list*] эффективно дает значение в середине отсортированной версии *списка*. Его часто считают более надежной мерой центра распределения, чем среднее, поскольку оно меньше зависит от отдаленных значений.

[Variance](https://reference.wolfram.com/language/ref/Variance.html)  
[Variance](https://reference.wolfram.com/language/ref/Variance.html)[*list*] определяется какРисунок 18 для реальных данных.

**[StandardDeviation](https://reference.wolfram.com/language/ref/StandardDeviation.html)**

[StandardDeviation](https://reference.wolfram.com/language/ref/StandardDeviation.html)[*list*] определяется как. Рисунок 20



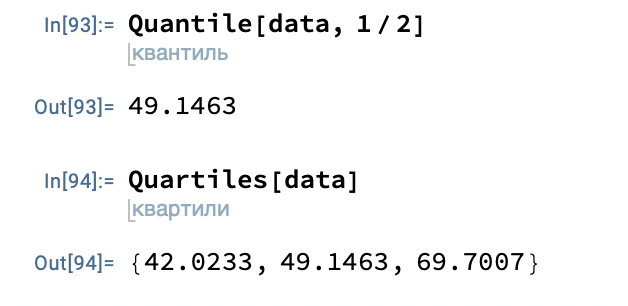
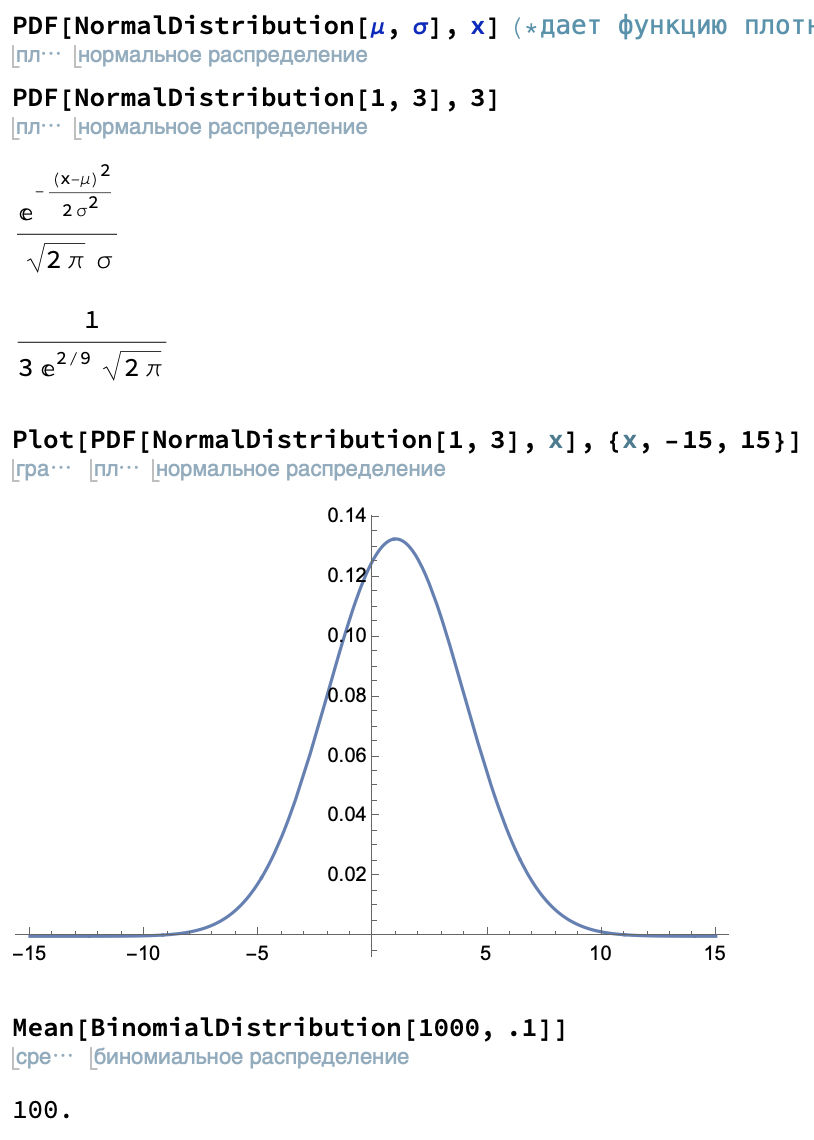
Если элементы в списке считаются выбранными случайным образом в соответствии с некоторым распределением вероятностей, то среднее дает оценку того, где находится центр распределения, в то время как стандартное отклонение дает оценку того, насколько широка дисперсия в распределении.

**Quantile**

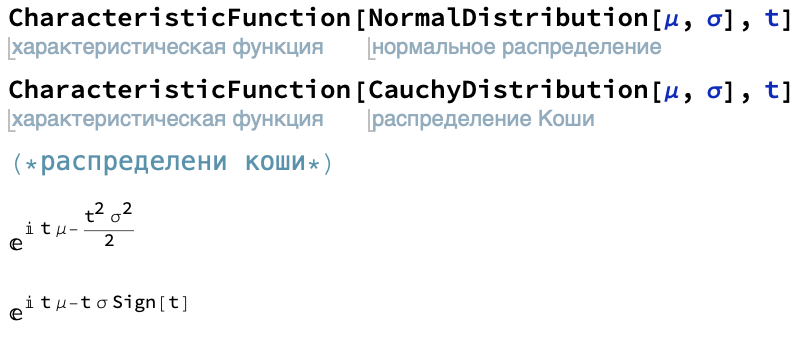
Тем https://reference.wolfram.com/language/tutorial/Files/NumericalOperationsOnData.en/7.png Й *квантиль* Quantile [*список]*,*q*]эффективно дает значение пути https://reference.wolfram.com/language/tutorial/Files/NumericalOperationsOnData.en/8.pngчерез отсортированную версию *списка*.

**2. Работа со статистическими распределениями**

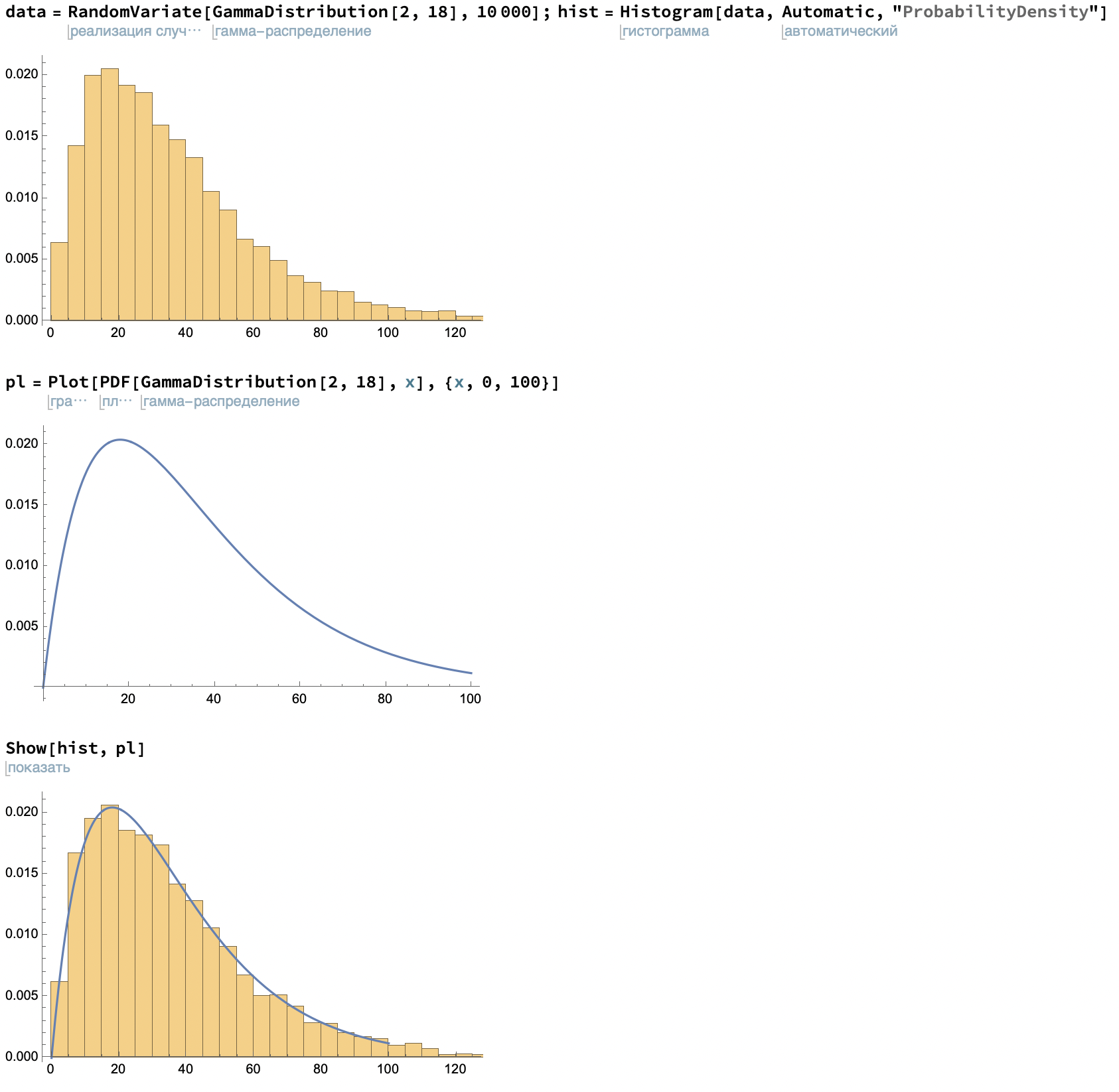
**PDF**

PDF[dist,x]вычисляет массовую функцию при x, если x является числовым значением, и в противном случае оставляет функцию в символьной форме, когда это возможно.

**CharacteristicFunction**

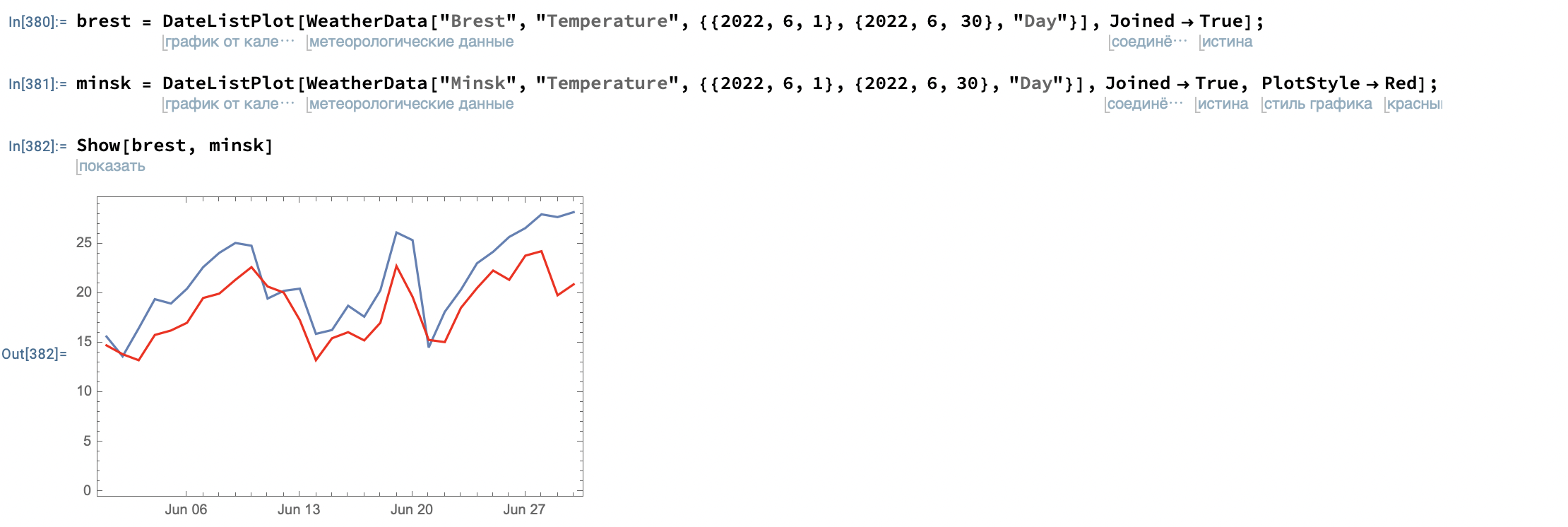
CharacteristicFunction дает характеристическую функцию для распределения в зависимости от переменной t.

**Histogram&Plot**

Массивы данных часто представляются гистограммами по столбцам или строками. Если массив одномерный — его представляют столбиковыми гистограммами. При этом данные по горизонтали разбиваются на N участков по числу столбцов (или строк) будущей гистограммы. В каждом из частичных участков подсчитывается сумма данных — число, и оно определяет высоту (или длину) столбца. Таким образом, проводится первичная статистическая обработка данных.

**Задача 1.**

График температур в городе Бресте и Минске за июнь 2022 года



**Задача 2.**

Методом Монте-Карло определить вероятность прибытия на занятия вовремя, если:

1. выход ровно в 8
2. начало занятий в 8:40
3. первую половину пути надо проехать на автобусе (20 минут ± 10 минут)
4. вторую половину пути надо пройти на автобусе (10 минут ± 5 минут)

