

Лабораторная работа №1

МОДЕЛИРОВАНИЕ НЕПРЕРЫВНЫХ И ДИСКРЕТНЫХ
СЛУЧАЙНЫХ ВЕЛИЧИН

Выполнил работу: Студент группы ИА-231

Зырянов Иван

2024

Введение

Вариант 4

Непрерывные случайные величины

4	$\frac{1}{x \cdot \ln 10}$	$(0; 10]$
---	----------------------------	-----------

Дискретные случайные величины

4	$P\{X = k_i\} = p_i; p = \{0.05, 0.1, 0.2, 0.3, 0.2, 0.1, 0.05\} \quad k = 4, 5, \dots, 10$
---	---

Цель лабораторной работы – получить знание о генераторах случайных величин и их практической реализации. Работа также закрепляет практические навыки применения теории вероятности и основ статистического анализа.

Генераторы случайных чисел используются для получения последовательностей чисел, которые можно считать случайными. В большинстве языков программирования (включая MATLAB) реализованы генераторы псевдослучайных чисел

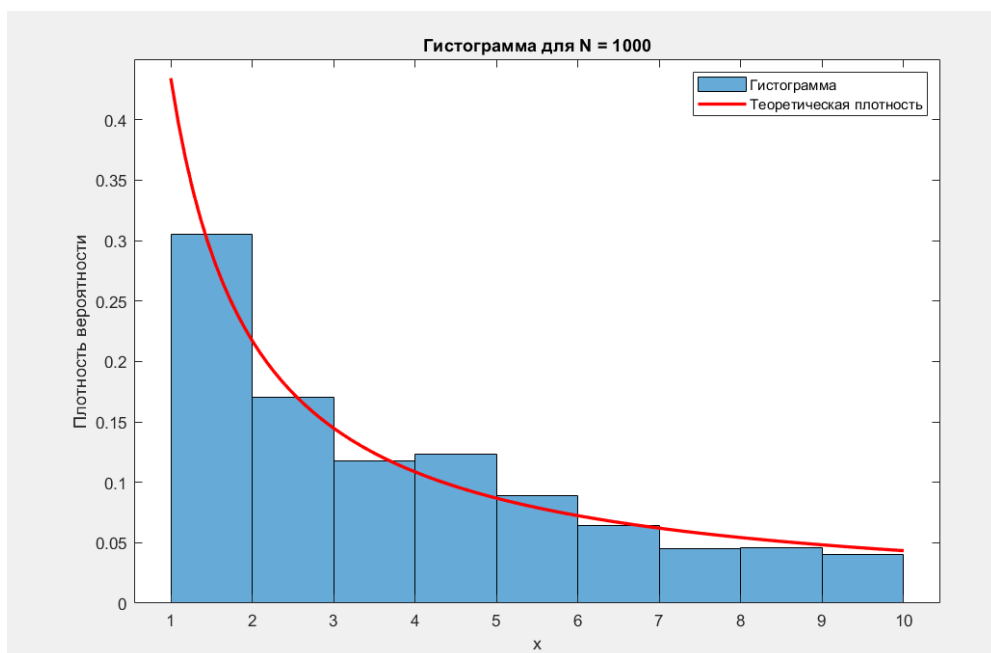
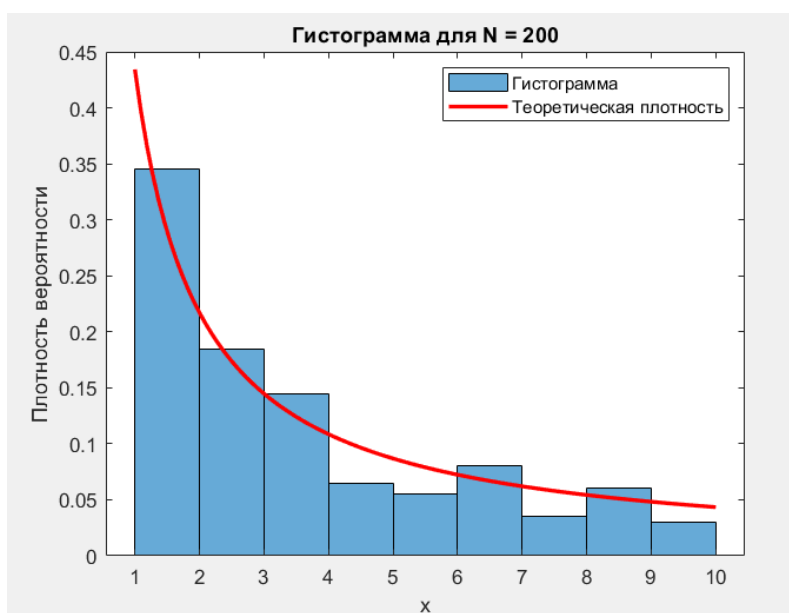
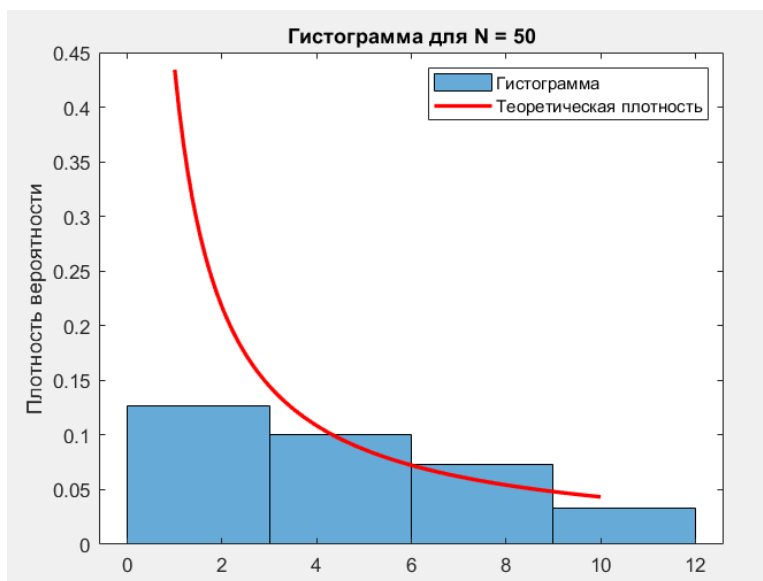
Результаты Работы по шагам

1-

функция распределения: $F(x) = \log_{10}(x)$

Следовательно, обратная функция: $x = 10^y$





Точечные оценки для N = 50

{'N'}	{'Оценка среднего'}	{'Оценка дисперсии'}	{'Оценка СКО'}
{{[50]}}	{{[3.2778]}}	{{[4.7558]}}	{{[2.1808]}}

Интервальные оценки для N = 50

{'N'}	{'alpha'}	{'Интервал для среднего'}	{'Интервал для дисперсии'}
{{[50]}}	{{[0.1000]}}	{'[2.7607, 3.7948]'}	{'[3.5128, 6.8681]'}
{{[50]}}	{{[0.0500]}}	{'[2.6580, 3.8975]'}	{'[3.3185, 7.3851]'}
{{[50]}}	{{[0.0100]}}	{'[2.4512, 4.1043]'}	{'[2.9788, 8.5520]'}

Точечные оценки для N = 200

{'N'}	{'Оценка среднего'}	{'Оценка дисперсии'}	{'Оценка СКО'}
{{[200]}}	{{[4.1754]}}	{{[7.2346]}}	{{[2.6897]}}

Интервальные оценки для N = 200

{'N'}	{'alpha'}	{'Интервал для среднего'}	{'Интервал для дисперсии'}
{{[200]}}	{{[0.1000]}}	{'[3.8611, 4.4897]'}	{'[6.1813, 8.6023]'}
{{[200]}}	{{[0.0500]}}	{'[3.8004, 4.5505]'}	{'[5.9997, 8.8965]'}
{{[200]}}	{{[0.0100]}}	{'[3.6808, 4.6701]'}	{'[5.6650, 9.5110]'}

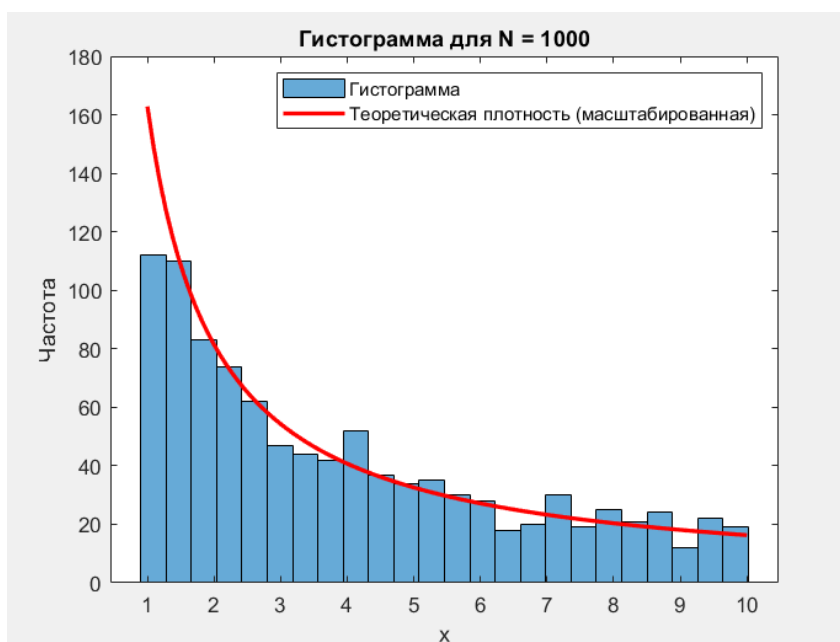
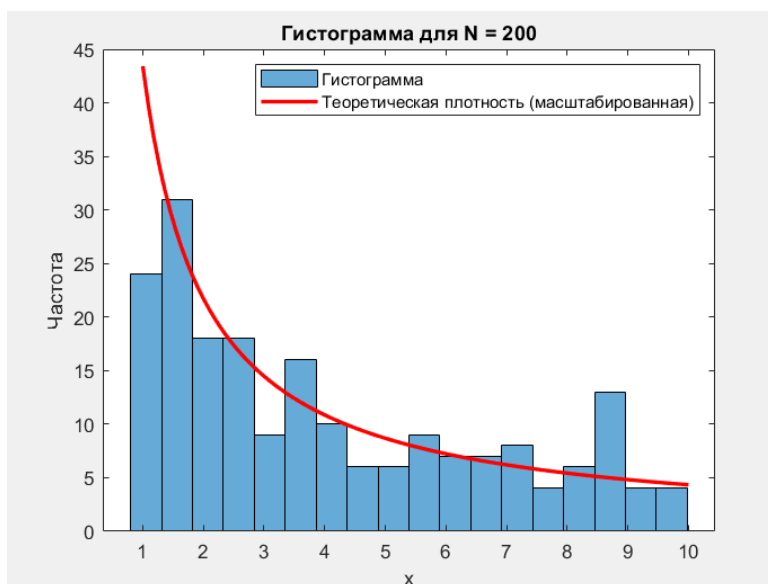
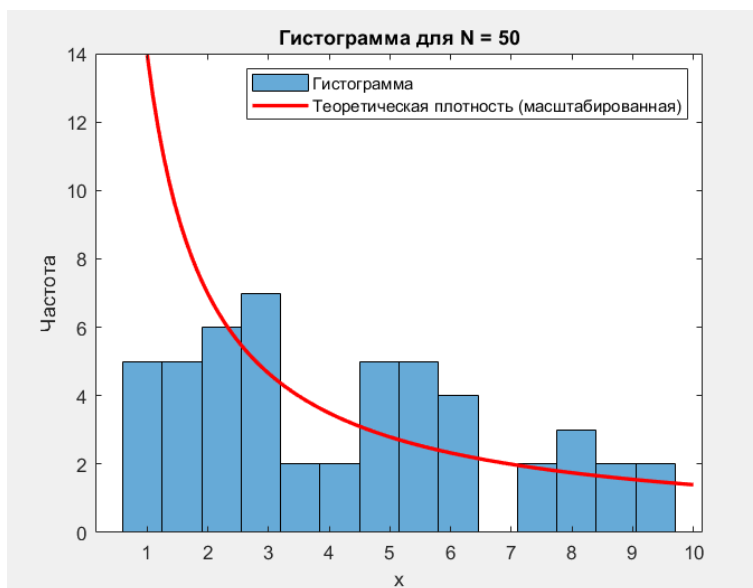
Точечные оценки для N = 1000

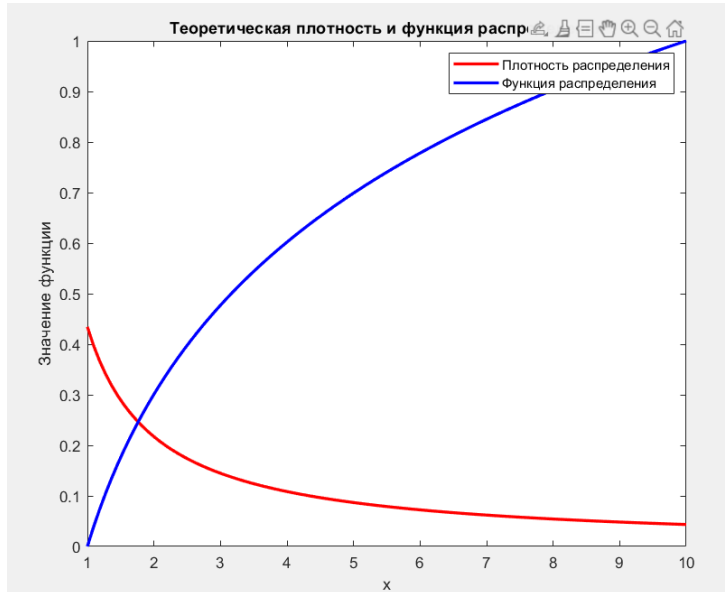
{'N'}	{'Оценка среднего'}	{'Оценка дисперсии'}	{'Оценка СКО'}
{{[1000]}}	{{[3.9661]}}	{{[6.1784]}}	{{[2.4856]}}

Интервальные оценки для N = 1000

{'N'}	{'alpha'}	{'Интервал для среднего'}	{'Интервал для дисперсии'}
{{[1000]}}	{{[0.1000]}}	{'[3.8367, 4.0955]'}	{'[5.7488, 6.6609]'}
{{[1000]}}	{{[0.0500]}}	{'[3.8118, 4.1203]'}	{'[5.6704, 6.7581]'}
{{[1000]}}	{{[0.0100]}}	{'[3.7632, 4.1689]'}	{'[5.5213, 6.9537]'}

6,7-





1. Визуальный анализ гистограмм:

- **Форма:** Гистограммы для всех трёх выборок ($N=50$, $N=200$, $N=1000$) визуально напоминают форму графика теоретической плотности распределения. Это говорит о том, что сгенерированные выборки действительно похожи на выборки из заданного распределения.
- **Сглаживание:** С увеличением объёма выборки (от 50 до 1000) гистограмма становится более сглаженной и лучше приближается к графику теоретической плотности, что согласуется с законом больших чисел – чем больше объём выборки, тем точнее она отражает свойства генеральной совокупности.

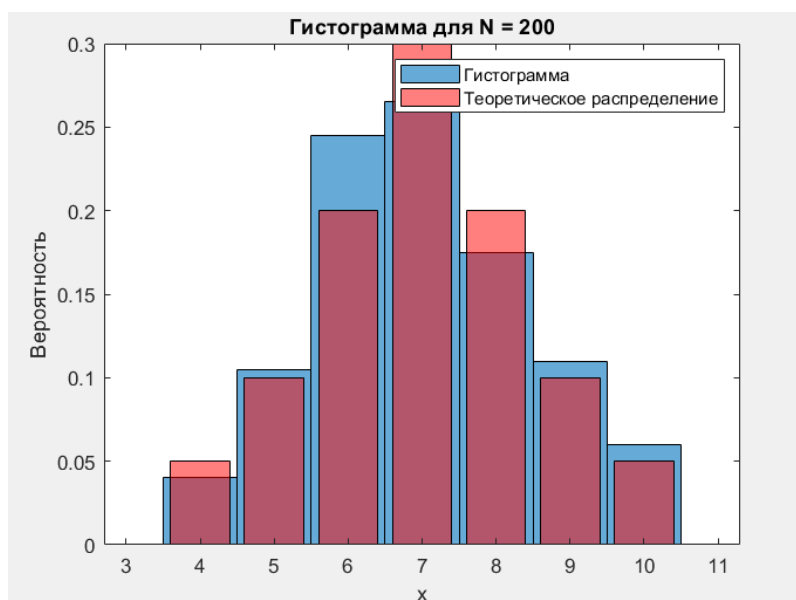
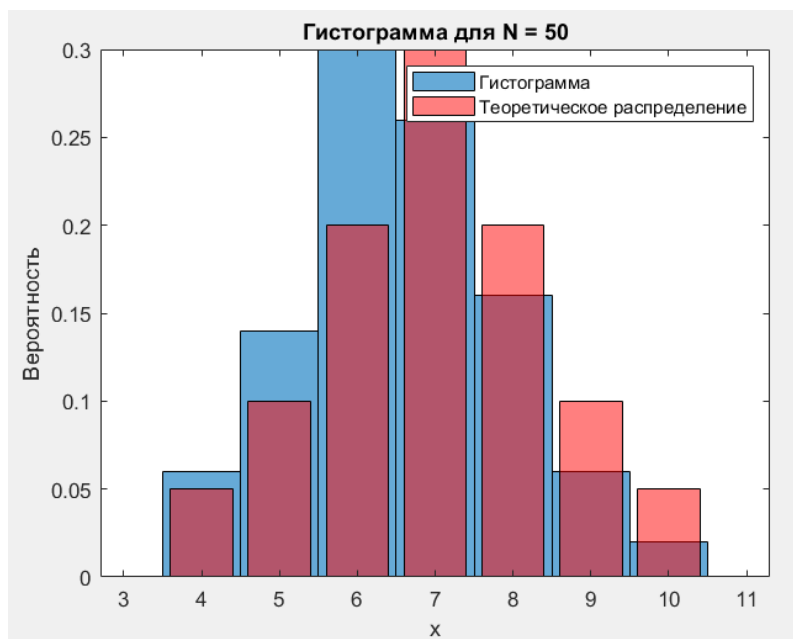
2. Анализ точечных и интервальных оценок:

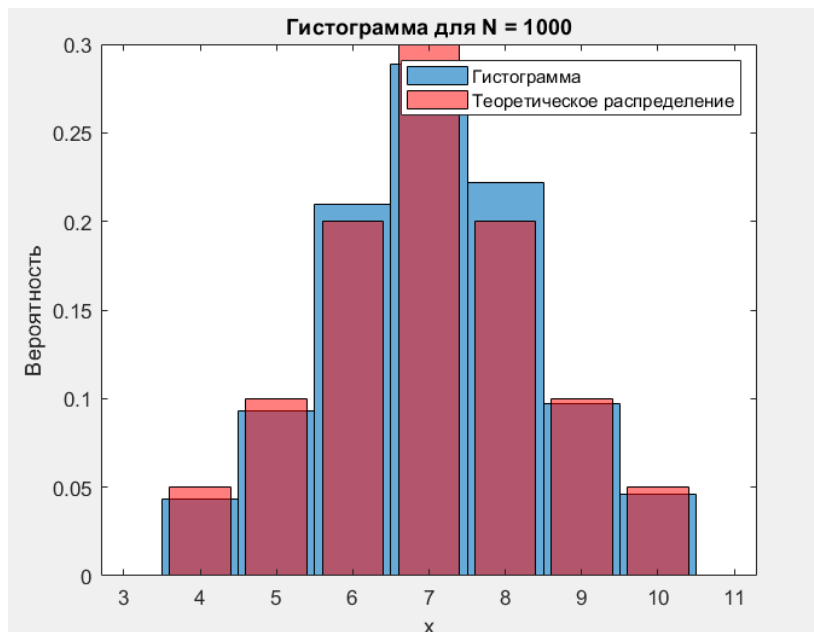
- **Сходимость оценок:** С увеличением объёма выборки точечные оценки среднего, дисперсии и СКО приближаются к значениям, которые можно было бы ожидать, глядя на теоретическую плотность распределения.
- **Сужение интервальных оценок:** С увеличением объёма выборки доверительные интервалы для среднего и дисперсии сужаются. Это говорит о том, что с увеличением объёма выборки мы получаем более точные оценки параметров распределения.
- **Уровень значимости:** Доверительные интервалы, построенные для разных уровней значимости, показывают, что с уменьшением уровня значимости (от 0.1 до 0.01) интервалы становятся шире, что логично.

3. Сравнение с теоретическими функциями:

- **Соответствие функции распределения:** График эмпирической функции распределения (который можно получить из гистограммы) будет тем ближе к графику теоретической функции распределения, чем больше объём выборки.

10- Повторяем шаги с 1 по 9, но уже для дискретно распределенной случайной





Точечные оценки для N = 50

{'N'}	{'Оценка среднего'}	{'Оценка дисперсии'}	{'Оценка СКО'}
{[50]}	{[6.5800]}	{[1.8404]}	{[1.3566]}

Интервальные оценки для N = 50

{'N'}	{'alpha'}	{'Интервал для среднего'}	{'Интервал для дисперсии'}
{[50]}	{[0.1000]}	{'[6.2583, 6.9017]'}	{'[1.3594, 2.6578]'}
{[50]}	{[0.0500]}	{'[6.1945, 6.9655]'}	{'[1.2842, 2.8579]'}
{[50]}	{[0.0100]}	{'[6.0658, 7.0942]'}	{'[1.1527, 3.3094]'}

Точечные оценки для N = 200

{'N'}	{'Оценка среднего'}	{'Оценка дисперсии'}	{'Оценка СКО'}
{[200]}	{[7]}	{[2.1910]}	{[1.4802]}

Интервальные оценки для N = 200

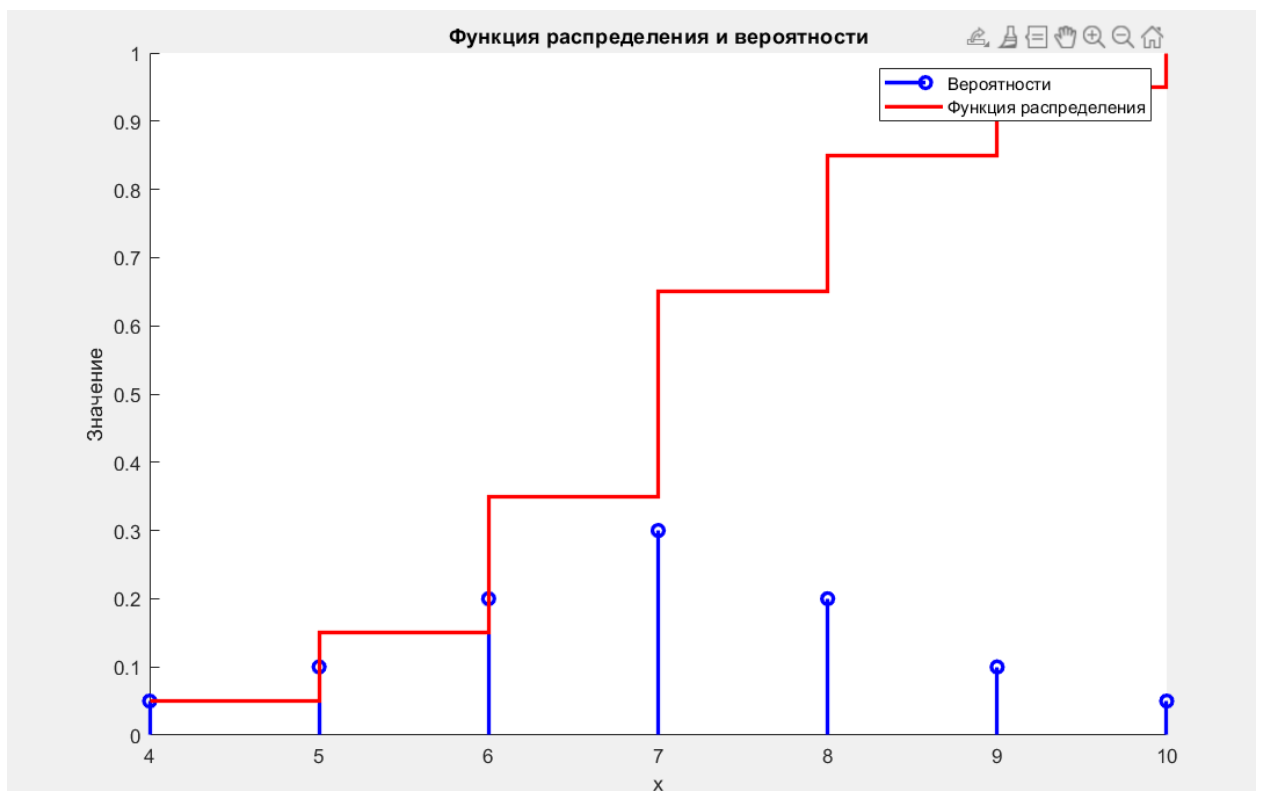
{'N'}	{'alpha'}	{'Интервал для среднего'}	{'Интервал для дисперсии'}
{[200]}	{[0.1000]}	{'[6.8270, 7.1730]'}	{'[1.8720, 2.6051]'}
{[200]}	{[0.0500]}	{'[6.7936, 7.2064]'}	{'[1.8170, 2.6942]'}
{[200]}	{[0.0100]}	{'[6.7278, 7.2722]'}	{'[1.7156, 2.8804]'}

Точечные оценки для $N = 1000$

{'N'}	{'Оценка среднего'}	{'Оценка дисперсии'}	{'Оценка СКО'}
{[1000]}	{[7.0290]}	{[1.9942]}	{[1.4121]}

Интервальные оценки для $N = 1000$

{'N'}	{'alpha'}	{'Интервал для среднего'}	{'Интервал для дисперсии'}
{[1000]}	{[0.1000]}	{'[6.9555, 7.1025]'}	{'[1.8555, 2.1499]'}
{[1000]}	{[0.0500]}	{'[6.9414, 7.1166]'}	{'[1.8302, 2.1813]'}
{[1000]}	{[0.0100]}	{'[6.9138, 7.1442]'}	{'[1.7821, 2.2444]'}



Выводы

Визуальный анализ гистограмм и графика распределения:

- Гистограммы для выборок разных размеров ($N = 50, 200, 1000$) демонстрируют схожую с теоретическим распределением форму. Чем больше размер выборки, тем ближе гистограмма к теоретическим вероятностям, представленным столбцами.
- Как и в случае с непрерывной случайной величиной, с увеличением размера выборки гистограмма становится более сглаженной и точнее отражает теоретическое распределение.
- График функции распределения имеет ярко выраженный ступенчатый характер, что типично для дискретных распределений. Каждая ступенька соответствует вероятности конкретного значения.

2. Анализ точечных и интервальных оценок:

- Сходимость к теоретическим значениям: Точечные оценки среднего и дисперсии, рассчитанные для выборок, с ростом N приближаются к значениям, которые можно было бы ожидать, исходя из теоретических вероятностей.
- Уменьшение ширины доверительных интервалов: Подобно непрерывному случаю, с увеличением размера выборки доверительные интервалы для среднего и дисперсии сужаются, что указывает на повышение точности оценок.
- Влияние уровня значимости: С уменьшением уровня значимости (от 0.1 до 0.01) доверительные интервалы становятся шире. Это связано с тем, что при более высоком уровне доверия (меньшем уровне значимости) требуется более широкий интервал, чтобы гарантировать попадание истинного значения параметра в интервал с заданной вероятностью.

Общие выводы:

- Сгенерированные выборки из дискретного распределения хорошо согласуются с заданными теоретическими вероятностями.
- Точечные и интервальные оценки параметров (среднего и дисперсии) становятся более точными с увеличением размера выборки.
- Выбор уровня значимости влияет на ширину доверительных интервалов: чем меньше уровень значимости, тем шире интервал.

11-Выводы:

В ходе лабораторной работы были исследованы свойства дискретного распределения вероятностей и подтверждены на практике основные положения математической статистики.

1. Моделирование дискретного распределения:

- Были сгенерированы выборки различного объёма (50, 200, 1000) из заданного дискретного распределения с вероятностями p и значениями k .
- С помощью гистограмм наглядно продемонстрировано, как с увеличением объёма выборки эмпирическое распределение всё точнее приближается к теоретическому.

2. Оценивание параметров:

- Для каждой выборки были вычислены точечные оценки среднего значения, дисперсии и среднеквадратического отклонения.
- Было показано, что с ростом объёма выборки точечные оценки приближаются к значениям, ожидаемым для данного теоретического распределения.

3. Интервальное оценивание:

- Построены доверительные интервалы для среднего значения и дисперсии при различных уровнях значимости (0.1, 0.05, 0.01).
- Продemonстрировано, что с увеличением объёма выборки доверительные интервалы сужаются, что говорит о повышении точности оценок.
- Также видно, что с уменьшением уровня значимости (т.е. с повышением требований к надёжности) доверительные интервалы расширяются.

4. Визуализация распределения:

- Построены графики функции распределения и вероятностей для заданного дискретного распределения.
- Наглядно показано, как связаны между собой вероятности отдельных значений и значения функции распределения.

Заключение:

Лабораторная работа позволила закрепить теоретические знания о дискретных случайных величинах, получить практические навыки моделирования случайных выборок, построения гистограмм и оценивания параметров распределения.

Доп задания-

--- Теоретические моменты ---

Непрерывное распределение:

Теоретическое мат. ожидание: 3.9087

Теоретическая дисперсия: 6.2200

Дискретное распределение:

Теоретическое мат. ожидание: 7.0000

Теоретическая дисперсия: 2.1000

--- Асимметрия и эксцесс (для выборок) ---

N = 50:

Непрерывное распределение:

Асимметрия: 0.9144

Эксцесс: 2.4605

Дискретное распределение:

Асимметрия: 0.0347

Эксцесс: 2.4215

N = 200:

Непрерывное распределение:

Асимметрия: 0.8488

Эксцесс: 2.4655

Дискретное распределение:

Асимметрия: 0.1627

Эксцесс: 2.6290

N = 1000:

Непрерывное распределение:

Асимметрия: 0.7328

Эксцесс: 2.3483

Дискретное распределение:

Асимметрия: 0.0750

Эксцесс: 2.6648

```

--- Критерий Колмогорова-Смирнова ---
N = 50 (непрерывное распределение):
  Уровень значимости | h | p-value
      0.10           | 0 | 0.2146
      0.05           | 0 | 0.2146
      0.01           | 0 | 0.2146
N = 50 (дискретное распределение):
  Уровень значимости | h | p-value
      0.10           | 1 | 0.0003
      0.05           | 1 | 0.0003
      0.01           | 1 | 0.0003

N = 200 (непрерывное распределение):
  Уровень значимости | h | p-value
      0.10           | 0 | 0.2091
      0.05           | 0 | 0.2091
      0.01           | 0 | 0.2091
N = 200 (дискретное распределение):
  Уровень значимости | h | p-value
      0.10           | 1 | 0.0000
      0.05           | 1 | 0.0000
      0.01           | 1 | 0.0000

N = 1000 (непрерывное распределение):
  Уровень значимости | h | p-value
      0.10           | 0 | 0.6042
      0.05           | 0 | 0.6042
      0.01           | 0 | 0.6042
N = 1000 (дискретное распределение):
  Уровень значимости | h | p-value
      0.10           | 1 | 0.0000
      0.05           | 1 | 0.0000
      0.01           | 1 | 0.0000

```

Ссылка на код программы:

[Sibsutis/3st Year/5th semester/TMO/1lab at main · kibatora/Sibsutis \(github.com\)](https://github.com/kibatora/Sibsutis)