Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет ИТМО»

Факультет программной инженерии и компьютерной техники

Учебно исследовательская работа №1

По Моделированию

Вариант №119

Выполнил:

Ступин Тимур Русланович

Группа № P3308

Поток № 1.3

Преподаватель:

Авксентьева Елена Юрьевна

Санкт-Петербург 2025

Содержание

Цель работы

Применение методов обработки и статистического анализа результатов измерений на примере заданной числовой последовательности путем оценки числовых моментов и выявления свойств последовательности на основе корреляционного анализа, а также аппроксимация закона распределения заданной последовательности по двум числовым моментам случайной величины.

Порядок выполнения работы

Этап 1

Оценки математического ожидания, дисперсии, среднеквадратичного отклонения, коэффициента вариации заданной числовой последовательности и доверительные интервалы для оценки математического ожидания с доверительными вероятностями 0.9; 0.95 и 0.99, сведенные в таблицу.

При расчете использовались следующие формулы:

- Математическое ожидание

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

- Дисперсия

$$\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2$$

- Среднеквадратическое отклонение

$$\sigma = \sqrt{\sigma^2}$$

- Коэффициент вариации

$$\nu = \frac{\sigma}{\overline{x}}$$

- Полуинтервал доверительного интервала (доверительная вероятность p)

$$\varepsilon_p = t_p \cdot \frac{\sigma}{\sqrt{n}}$$

- Относительное отклонение от эталонного значения (в процентах)

$$\delta = \mid \frac{X_n - X_{300}}{X_{300}} \mid \cdot 100\%$$

Форма 1

Характеристи-ка		**Количество случайных величин**					
3-8		10	20	50	100	200	300
Мат.ож.	Знач.	19.7870	13.9422	22.4223	23.0236	22.5223	22.8924
	%	13.5654	39.0969	2.0539	0.5727	1.6168	
Дов. инт. (0.9)	Знач.	0.8794	0.5455	0.5965	0.4867	0.3386	0.2830
	%	210.7519	92.7807	110.7812	71.9964	19.6587	
Дов. инт. (0.95)	Знач.	0.4388	0.2722	0.2976	0.2429	0.1690	0.1412
	%	210.7519	92.7807	110.7812	71.9964	19.6587	
Дов. инт. (0.99)	Знач.	0.0877	0.0544	0.0595	0.0485	0.0338	0.0282
	%	210.7519	92.7807	110.7812	71.9964	19.6587	
Дисперсия	Знач.	489.7042	376.9324	1126.5235	1500.1853	1452.1956	1521.3447
	%	67.8111	75.2237	25.9521	1.3908	4.5453	
С.к.о.	Знач.	22.1293	19.4147	33.5637	38.7322	38.1077	39.0044
	%	43.2647	50.2242	13.9489	0.6979	2.2991	
К-т вариации	Знач.	1.1184	1.3925	1.4969	1.6823	1.6920	1.7038
	%	34.3605	18.2705	12.1445	1.2634	0.6935	

% - относительные отклонения рассчитанных значений от значений, полученных для выборки из трехсот величин

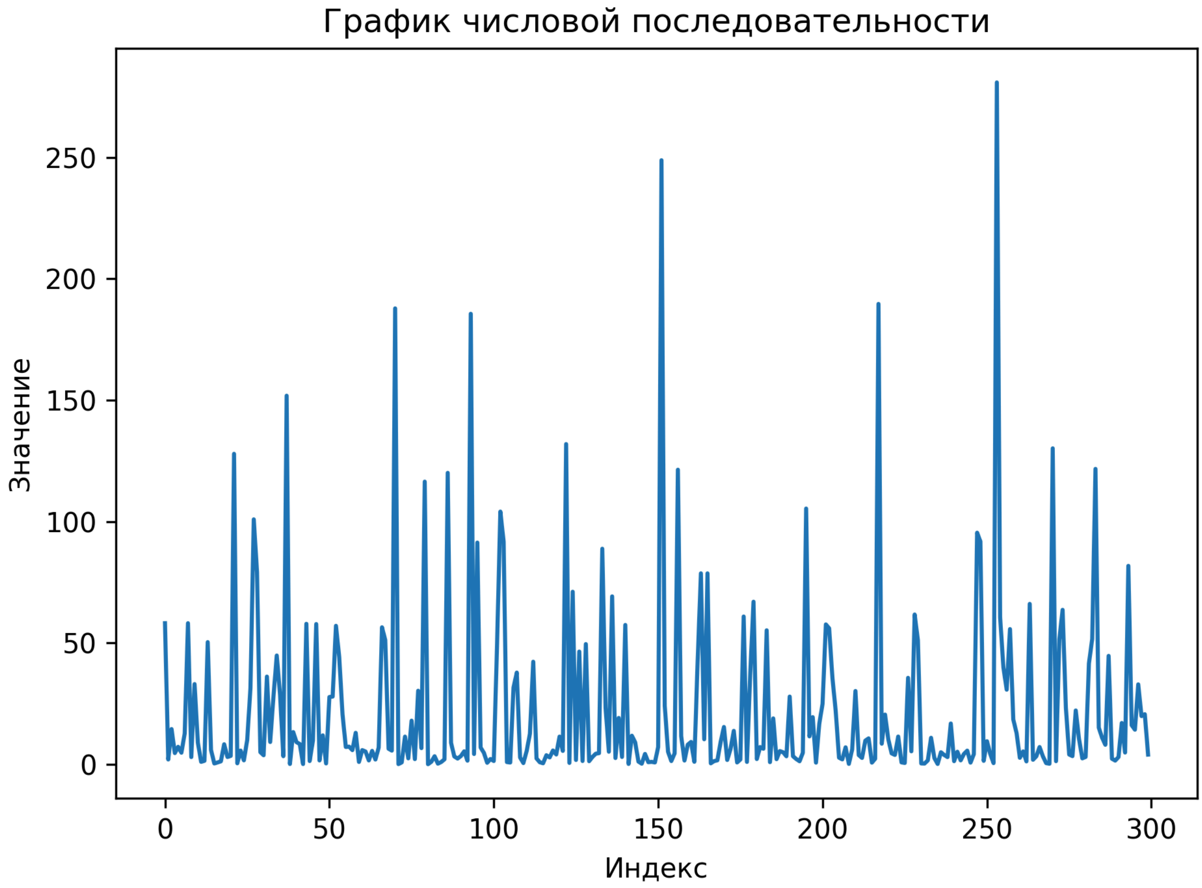
Вывод

Математическое ожидание, дисперсия и среднеквадратичное отклонение стабилизируются с ростом объёма выборки, что свидетельствует о сходимости оценок к истинным значениям. Коэффициент вариации сохраняет относительно устойчивое значение при $n \geq 100$, что указывает на стабильность относительного разброса данных и позволяет надёжно использовать его для выбора аппроксимирующего закона распределения.

Этап 2

Значений заданной числовой последовательности с результатами анализа характера числовой последовательности (возрастающая, убывающая, периодичная и т.п.);

График 1



Вывод

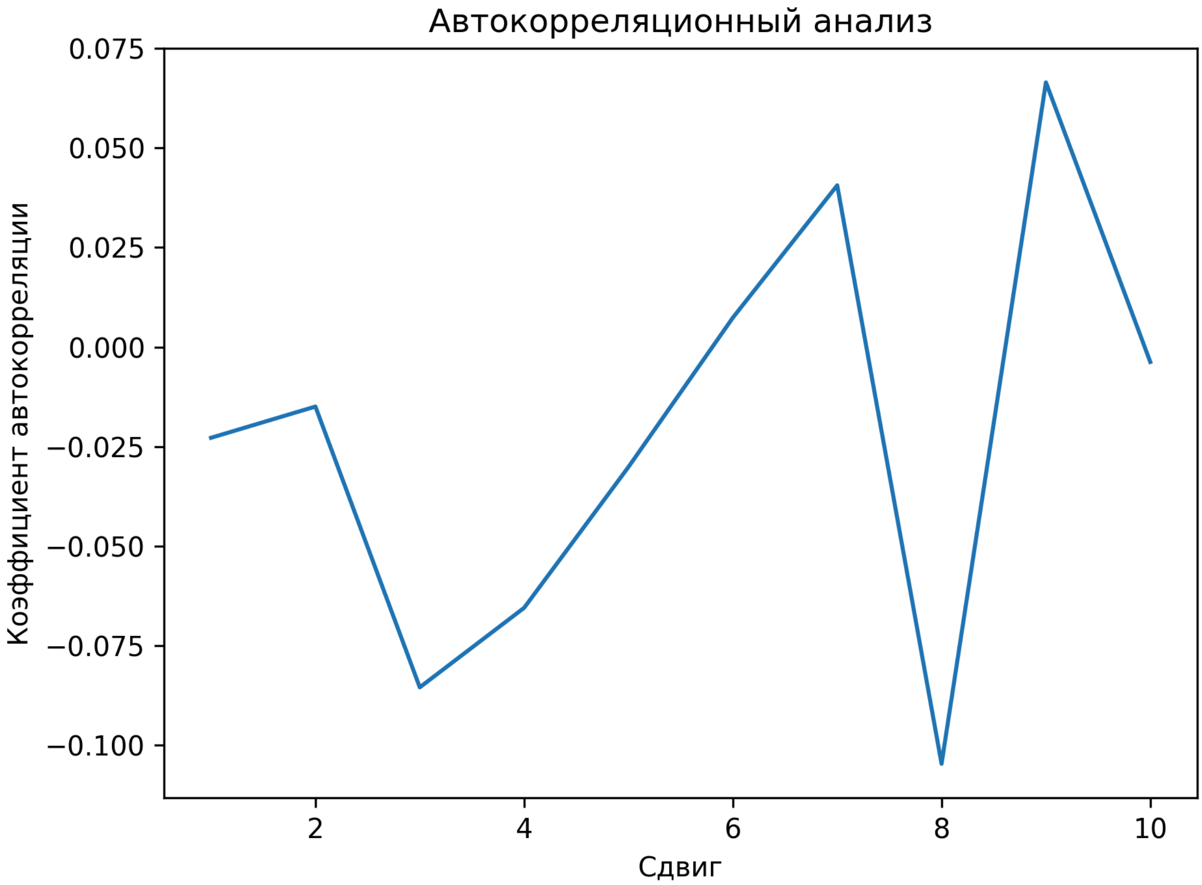
Анализируя график, можно сделать вывод, что исходная последовательность не является периодической, возрастающей или убывающей.

Этап 3

Результаты автокорреляционного анализа (значения коэффициентов автокорреляции со сдвигом 1, 2, 3, ...), представленные как в числовом (форма 3), так и графическом виде, с обоснованным выводом о характере заданной числовой последовательности (можно ли ее считать случайной).

Форма 3

Сдвиг ЧП	1	2	3	4	5	6	7	8	9	10
К-т АК	−0.0228	−0.0150	−0.0855	−0.0655	−0.0301	0.0074	0.0406	−0.1047	0.0665	−0.0038



Вывод

Коэффициенты автокорреляции при сдвигах от 1 до 10 находятся в пределах ± 0.1, что указывает на отсутствие значимой корреляции между соседними элементами последовательности. Наблюдаемые колебания носят случайный характер, следовательно, заданную числовую последовательность можно считать случайной и независимой.

Этап 4

Гистограмма распределения частот для заданной числовой последовательности (график 2).

График 2

Вывод

Анализ гистограммы распределения частот показывает, что заданная числовая последовательность имеет сильно скошенное вправо распределение: большинство значений сконцентрировано в левой части графика (в интервале $0.01–15.62$, где частота превышает 200), а затем наблюдается быстрое снижение частоты по мере увеличения значений. Такая форма распределения не соответствует нормальному или равномерному закону, но хорошо описывается распределениями с экспоненциальным хвостом, в частности, гиперэкспоненциальным. Это согласуется с ранее полученными статистическими характеристиками — высоким коэффициентом вариации ($\nu = 1.7038 > 1$) и отсутствием автокорреляции, что указывает на случайный характер данных. Таким образом, гистограмма визуально подтверждает выбор типа аппроксимирующего закона на следующем этапе.

Этап 5

Параметры, рассчитанные по двум начальным моментам и определяющие вид аппроксимирующего закона распределения заданной случайной последовательности (равномерный; экспоненциальный; нормированный Эрланга; гипоэкспоненциальный;).

Для данной по варианту выборки коэффициент вариации больше единицы, следовательно, для аппроксимации последовательности будем использовать гиперэкспоненциальный закон.

Значения математического ожидания ($\overline{x} = 22.8924$) и коэффициента вариации $\nu = 1.7038$) были определены ранее.

$$q \leq \frac{2}{1+\nu^2} \approx 0.5111$$

Выберем $q = 0.5$, тогда:

$$\overline{x}_1 = \left[1 + \sqrt{\frac{1-q}{2q}(\nu^2 - 1)}\right]\overline{x} = 45.1483$$

$$\overline{x}_2 = \left[1 + \sqrt{\frac{q}{2(1-q)}(\nu^2 - 1)}\right]\overline{x} = 0.5622$$

где $\overline{x}$ – математическое ожидание

Соответственно получаем следующий аппроксимирующий закон распределения:

$$F(r_1, r_2 = \overline{x}_1 \cdot -\ln(1 - r_2) \quad \text{при} \quad r_1 < q$$

$$F(r_1, r_2 = \overline{x}_2 \cdot -\ln(1 - r_2) \quad \text{при} \quad r_1 \geq q$$

Вывод

Исходя из прошлого этапа и вычислений в данном этапе, можем сказать, что аппроксимирующий закон распределения данной числовой последовательности: **гиперэкспоненциальный**.

Этап 6

Описание алгоритма (программы) формирования аппроксимирующего закона распределения и расчета значений всех числовых характеристик с иллюстрацией (при защите отчета) его работоспособности. генерации случайной последовательности, соответствующей подобранному гиперэкспоненциальному распределению.

Я реализовал программный модуль на языке Python. Алгоритм основан на использовании двух независимых равномерных случайных чисел r_1 и r_2 , сгенерированных функцией `np.random.uniform()`.

Код программы, реализующей данный алгоритм:

```
def generate_hyperexp(q, t1, t2, n_samples=300):
    np.random.seed(42)
```

```python
    generated_numbers = []

    for i in range(n_samples):
        r1 = np.random.uniform(0, 1)
        r2 = np.random.uniform(0, 1)

        if r1 < q:
            x = t1 * -np.log(1 - r2)
        else:
            x = t2 * -np.log(1 - r2)

        generated_numbers.append(x)

    return generated_numbers
```

Вывод

Мне удалось сформировать числовую последовательность по аппроксимирующему закону и вычисленным параметрам на языке Python.

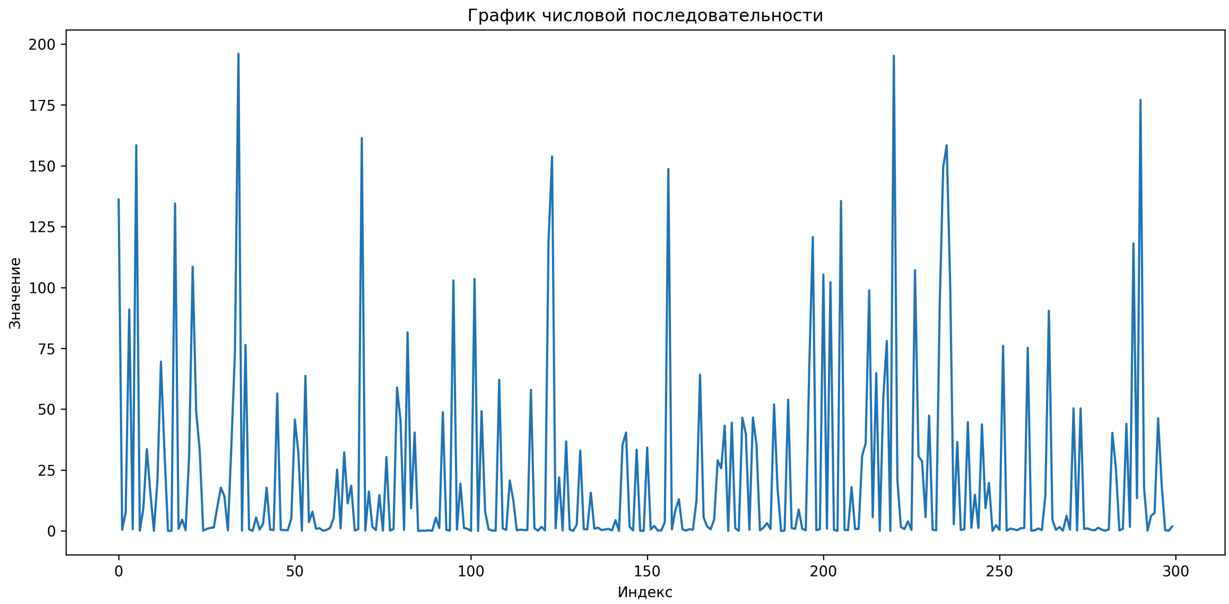
Этап 7

Выводы по результатам сравнения сгенерированной в соответствии с полученным аппроксимирующим законом распределения последовательности случайных величин и заданной числовой последовательности.

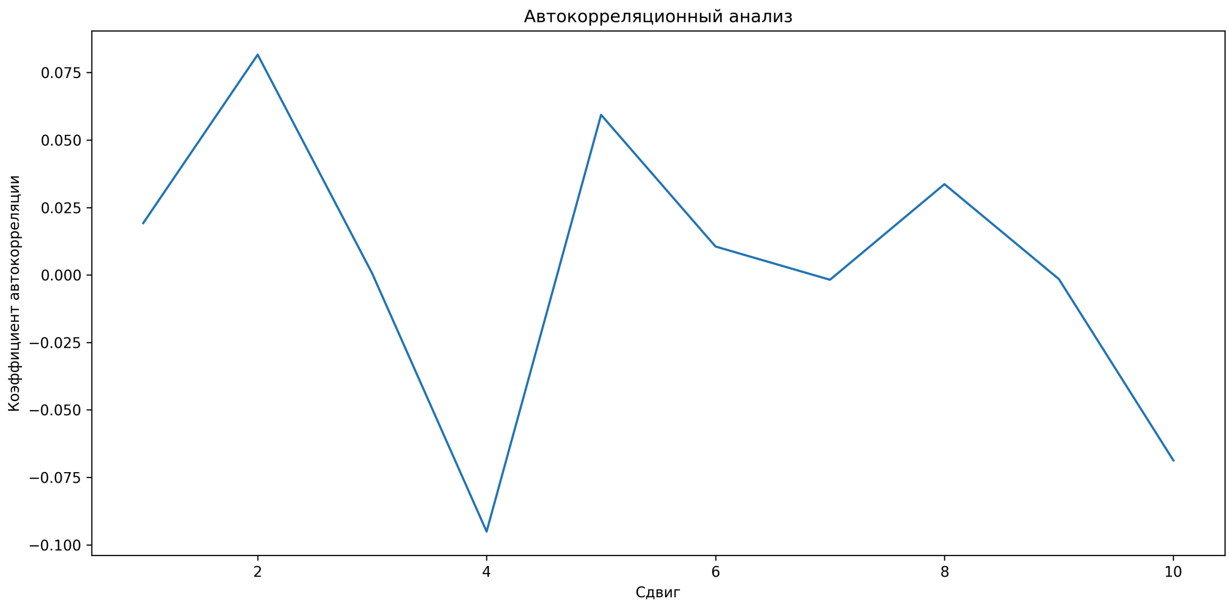
Форма 2

Закон распределения: гиперэкспоненциальный закон							
Характеристика		**Количество случайных величин**					
		10	20	50	100	200	300
Мат.ож.	Знач.	45.2900	35.8167	29.1321	23.5226	20.6412	22.9573
	%	97.2792	56.0141	26.8967	2.4622	10.0890	
Дов. инт. (0.9)	Знач.	2.4030	1.4728	0.8295	0.5020	0.3211	0.2833
	%	748.1682	419.8563	192.7829	77.2013	13.3241	
Дов. инт. (0.95)	Знач.	1.1991	0.7350	0.4139	0.2505	0.1602	0.1414
	%	748.1682	419.8563	192.7829	77.2013	13.3241	
Дов. инт. (0.99)	Знач.	0.2397	0.1469	0.0827	0.0501	0.0320	0.0283
	%	748.1682	419.8563	192.7829	77.2013	13.3241	
Дисперсия	Знач.	3656.8748	2747.5320	2178.7506	1596.1734	1305.6310	1524.9912
	%	139.7964	80.1671	42.8697	4.6677	14.3844	
С.к.о.	Знач.	60.4721	52.4169	46.6771	39.9521	36.1335	39.0511
	%	54.8536	34.2263	19.5281	2.3072	7.4713	
К-т вариации	Знач.	1.3352	1.4635	1.6023	1.6985	1.7506	1.7010
	%	21.5053	13.9653	5.8068	0.1512	2.9114	

Математическое ожидание отличается от математического ожидания исходной выборки на величину, не превосходящую доверительные интервалы. Это говорит о том, что аппроксимация выполнена качественно.



При сравнении полученных гистограмм видно, что полученная нами последовательность похожа на исходную. Это говорит о том, что параметры аппроксимации были рассчитаны достаточно точно.



Коэффициент автокорреляции интервалов от 1 до 10 приближены к нулю, следовательно, можно сказать, что выборка случайна.

Коэффициент корреляции между двумя числовыми последовательностями:

$$r = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \cdot \sum_{i=1}^{n}(y_i - \overline{y})^2}} = 0.0331$$

Для сгенерированной и полученной последовательности мы рассчитали корреляционную зависимость. Как видно из значения r, корреляции между исходной и сгенерированной случайной последовательностями практически нет.

Вывод

Сравнение гистограммы распределения частот исходной числовой последовательности и плотности распределения гиперэкспоненциального закона показало, что действительно исходная ЧП соотносится с гиперэкспоненциальным аппроксимирующим законом. Сравнение числовых характеристик исходной и сгенерированной ЧП показало явное сходство характеристик.

Вывод

В ходе выполнения лабораторной работы я исследовал данную числовую последовательность, определив основные её параметры, такие как математическое ожидание, дисперсия коэффициент вариации. Далее я проанализировал построенный график числовой последовательности, по которому не выявил возрастания, убывания или периодичности последовательности. Из этого я сделал вывод что исследуемая последовательность является случайной случайной, что подтвердилось результатами автокорреляционного анализа. Затем я вычислил параметры аппроксимирующего закона и по ним сгенерировали новую последовательность. Коэффициенты вариации сгенерированной последовательности также близится к 1.7, что говорит о корректности генерации гиперэкспоненциального закона. Коэффициент автокорреляции сгенерированной последовательности варьируется около нуля, что говорит о её случайности. Также, рассчитав коэффициент корреляции между исходной и сгенерированной выборкой я обнаружил что его значение очень мало, что говорит о достаточно точной генерации и оценке параметров исходной числовой последовательности. Математическое ожидание и дисперсия сгенерированной последовательности отличаются от исходной, но отличие не выходит за пределы доверительных интервалов.