

Федеральное государственное автономное образовательное учреждение высшего
образования

«Национальный исследовательский университет ИТМО»

Факультет программной инженерии и компьютерной техники

Направление подготовки 09.03.04 «Программная инженерия» –

Системное и прикладное программное обеспечение

Отчёт

По учебно-исследовательской работе №1

По моделированию

Вариант: 16

Выполнили:

студенты 3 курса

Барсуков Максим Андреевич

Горляков Даниил Петрович

Группа: P3315

Принял:

Алиев Тауфик Измайлович

Отчёт принят «__»____2024 г.

Оценка: _____

Задание

Цель работы

Изучение 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.

Содержание отчета

1. оценки *математического ожидания, дисперсии, среднеквадратического 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 (форма 1);
2. график (график 1) значений заданной числовой последовательности с результатами анализа характера числовой последовательности (возрастающая, убывающая, периодичная и т.п.);
3. окрестности автокорреляционного анализа (значения коэффициентов авторреляции 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);
4. гистограмма распределения частот для заданной числовой последовательности (график 2);
5. параметры, рассчитанные по двум начальным моментам и определяющие *вид аппроксимирующего закона распределения* заданной случайной последовательности (равномерный; экспоненциальный; нормированный Эрланга; гипотенциальный; гиперэкспоненциальный);
6. *описание алгоритма (программы) формирования* аппроксимирующего закона распределения и расчета значений всех числовых характеристик с иллюстрацией (при защите отчета) его работоспособности;
7. выводы по результатам сравнения сгенерированной в соответствии с полученным аппроксимирующим законом распределения последовательности случайных величин и заданной числовой последовательности, а именно:
 1. сравнения *плотности распределения* аппроксимирующего закона с *гистограммой распределения* частот для исходной числовой последовательности (график 3);

2. расчета числовых характеристик *сгенерированной* в соответствии с аппроксимирующим законом распределения случайной последовательности: математического ожидания, дисперсии, среднеквадратического отклонения, коэффициента вариации (представленные в таблице по форме 2) и коэффициентов автокорреляции при разных значениях сдвигов (в таблице по форме 3), а также сравнения (в %) полученных значений со значениями, рассчитанными для *заданной* числовой последовательности;
 3. проведения *корреляционного анализа* сгенерированной в соответствии с аппроксимирующим законом распределения последовательности случайных величин и заданной числовой последовательности на основе *коэффициента корреляции*.
8. ***по каждому из перечисленных выше пунктов отчета должны быть сформулированы результативные выводы и заключения.***

Ход работы

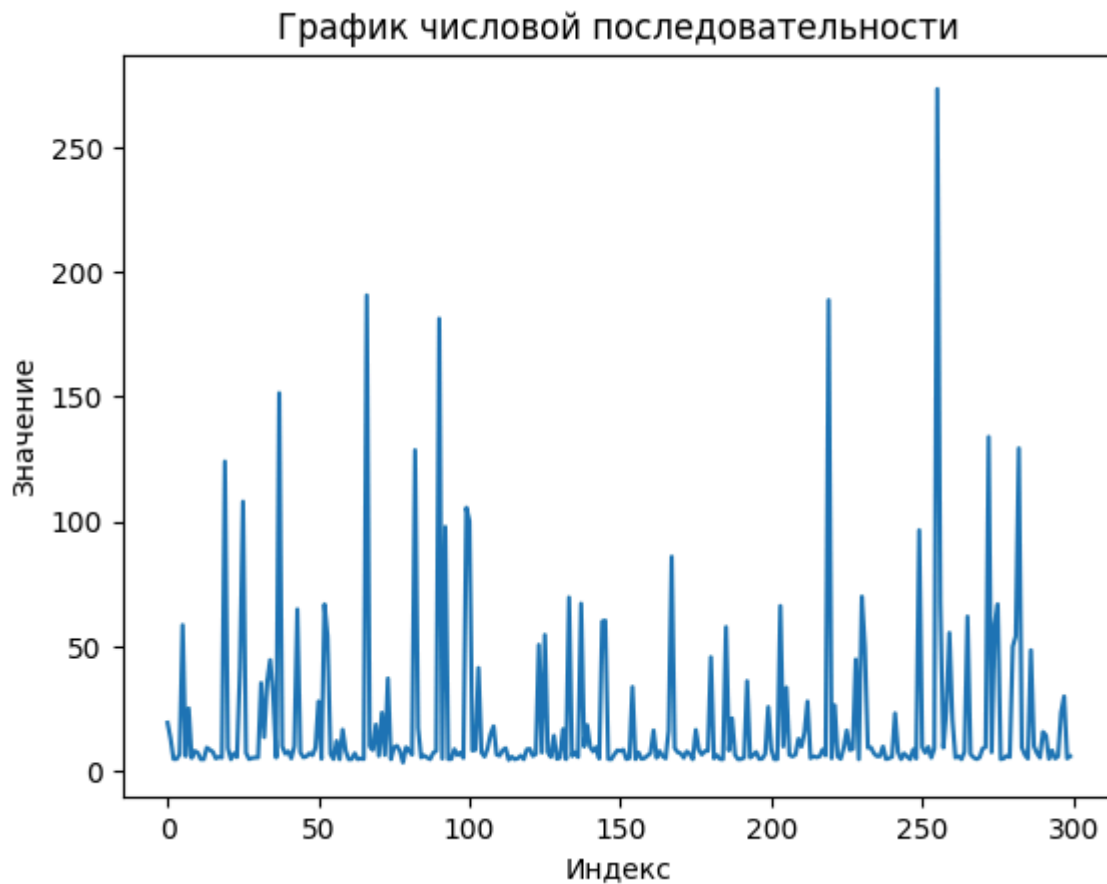
Этап 1. Форма №1. Оценки математического ожидания, дисперсии, среднеквадратического отклонения, коэффициента вариации заданной числовой последовательности и доверительные интервалы для оценки математического ожидания с доверительными вероятностями 0,9; 0,95 и 0,99, сведенные в таблицу.

Характеристика		Количество случайных величин					
		10	20	50	100	200	300
Мат. ож.	Знач,	15,03840	16,57860	19,88642	21,97183	18,15293	19,62616
	%	76,62	84,47	101,33	111,95	92,49	
Дов. инт. (0,9)	Знач,	±8,2608	±10,0877	±7,1891	±6,1162	±3,4519	±3,1558
	%	±261,766	±319,66	±227,81	±193,88	±109,38	
Дов. инт. (0,95)	Знач,	±9,8434	±12,0203	±8,5663	±7,2879	±4,1132	±3,7603
	%	±261,77	±319,66	±227,81	±193,81	±109,38	
Дов. инт. (0,99)	Знач,	±12,9364	±15,7973	±11,258	±9,5779	±5,4056	±4,9419
	%	±261,77	±319,66	±227,81	±193,81	±109,38	
Дисперсия	Знач,	252,226	752,249	955,127	1382,6	880,81	1104,282
	%	22,84	68,12	86,49	125,21	79,769	
С. к. о.	Знач,	15,8816	27,4272	30,9051	37,1836	29,6785	33,2307
	%	47,79	82,54	93,00	111,90	89,31	
К-т вариации	Знач,	1,056072	1,654371	1,554082	1,692331	1,634913	1,693186
	%	62,37	97,71	91,78	99,95	96,56	

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

Вывод из 1 этапа: Доверительные интервалы уменьшаются с ростом выборки, что свидетельствует о повышении точности оценок математического ожидания. Дисперсия и среднеквадратическое отклонение возрастают с увеличением выборки, что говорит о большей вариативности в данных, но при больших объемах выборки наблюдается стабилизация. Коэффициент вариации показывает умеренные изменения, что указывает на относительную стабильность отношения стандартного отклонения к среднему при увеличении выборки.

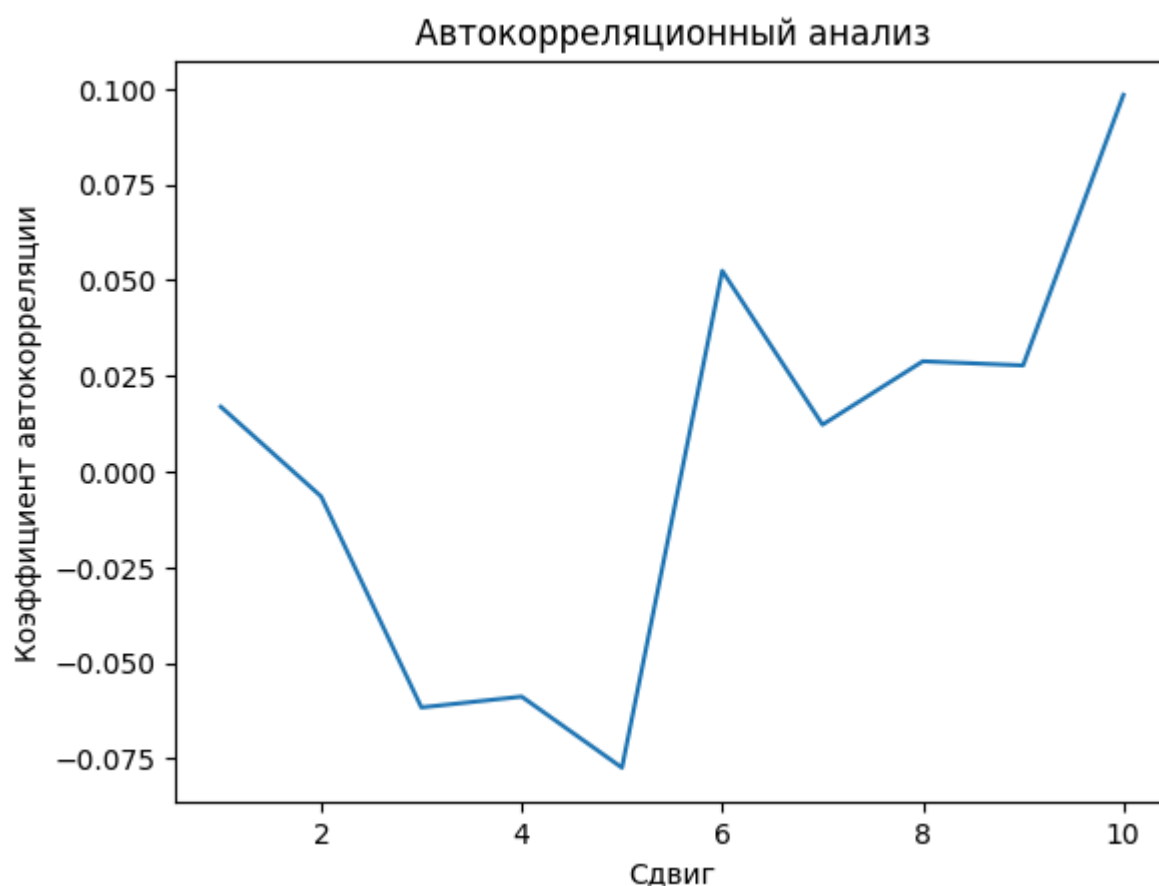
Этап 2. График №1. Значений заданной числовой последовательности с результатами анализа характера числовой последовательности.



Вывод из 2 этапа: Изучив график, можно сделать вывод, что исходная последовательность не является периодической, возрастающей или убывающей.

Этап 3. Форма 3. Результаты автокорреляционного анализа (значения коэффициентов автокорреляции со сдвигом 1, 2, 3, ...), представленные как в числовом (форма 3), так и графическом виде.

Сдвиг ЧП	1	2	3	4	5	6	7	8	9	10
К-т АК	0,0169	-0,0061	-0,0601	-0,0569	-0,0753	0,0500	0,0112	0,0310	0,0293	0,0965



Вывод из 3 этапа: Последовательность можно считать случайной так как данные коэффициенты указывают на то, что между числами не было выявлено зависимости, нет тенденции и периодичности.

Этап 4. График 2. Гистограмма распределения частот для заданной числовой последовательности (график 2).

Интервалы																		
№	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Левая граница	3,29	18,291	33,292	48,292	63,293	78,294	93,295	108,295	123,296	138,297	153,298	168,299	183,299	198,301	213,301	228,302	243,302	258,303
Правая граница	18,291	33,292	48,292	63,293	78,294	93,295	108,295	123,296	138,297	153,298	168,299	183,299	198,301	213,301	228,302	243,302	258,303	273,304
Частота	236	15	11	14	8	1	5	0	4	1	0	1	2	0	0	0	0	0



Вывод из 4 этапа: Исходя из гистограммы мы можем видеть, что большая часть значений располагается в промежутке от 3,29 до 18,291, еще часть располагается до 138,297 и наименьшая часть значений располагается в диапазоне больше. Из коэффициента вариации (близок к 1,7), можно предположить, что закон распределения ЧП – гиперэкспоненциальный.

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

Для данной по варианту выборки коэффициент вариации больше единицы, Следовательно, для аппроксимации последовательности будем использовать гиперэкспоненциальный закон. Значения математического ожидания ($M(X) = 19,62616$) и коэффициента вариации $v = 1,693186$) были определены ранее.

$$q \leq \frac{2}{1+v^2} \approx 0,517$$

Примем $q = 0,285$:

$$t_1 = [1 + \sqrt{\frac{1-q}{2q}} (v^2 - 1)]t = 51,7828t$$

$$t_2 = [1 - \sqrt{\frac{q}{2(1-q)}} (v^2 - 1)]t = 8,4450t,$$

где t – математическое ожидание

Соответственно, получаем следующий аппроксимирующий закон распределения:

$$F(r1, r2) = t1 * - \ln(1 - r2) \text{ при } r1 < q$$

$$F(r1, r2) = t2 * - \ln(1 - r2) \text{ при } r1 \geq q$$

Вывод из 5 этапа: Исходя из прошлого этапа и вычислений в данном этапе, можем сказать, что аппроксимирующий закон распределения данной ЧП: гиперэкспоненциальный.

6 этап. *Описание алгоритма (программы) формирования аппроксимирующего закона распределения и расчета значений всех числовых характеристик с иллюстрацией (при защите отчета) его работоспособности.*

Описание:

Для генерации случайной последовательности используем Excel. В листе Генератор создадим 300 пар случайных чисел с помощью =СЛМАССИВ(300; 2).

Используем эти данные для генерации гиперэкспоненциального распределения:

$$=ЕСЛИ(Генератор!A1 < F2;H1;H2) * -LN(1 - Генератор!B1)$$

Получим числовую последовательность.

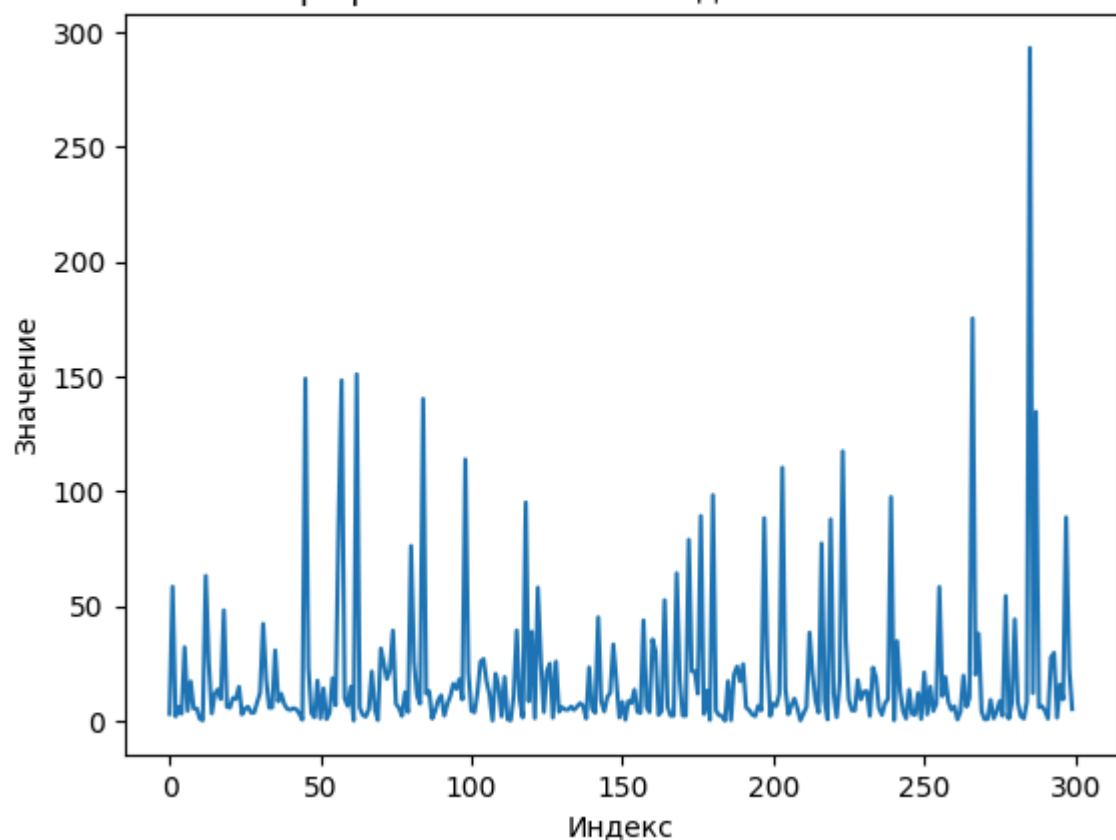
Вывод из 6 этапа: Нам удалось сформировать ЧП по аппроксимирующему закону в Excel и описать алгоритм формирования ЧП.

7 этап. График 3. Форма 2. Выводы по результатам сравнения сгенерированной в соответствии с полученным аппроксимирующим законом распределения последовательности случайных величин и заданной числовой последовательности.

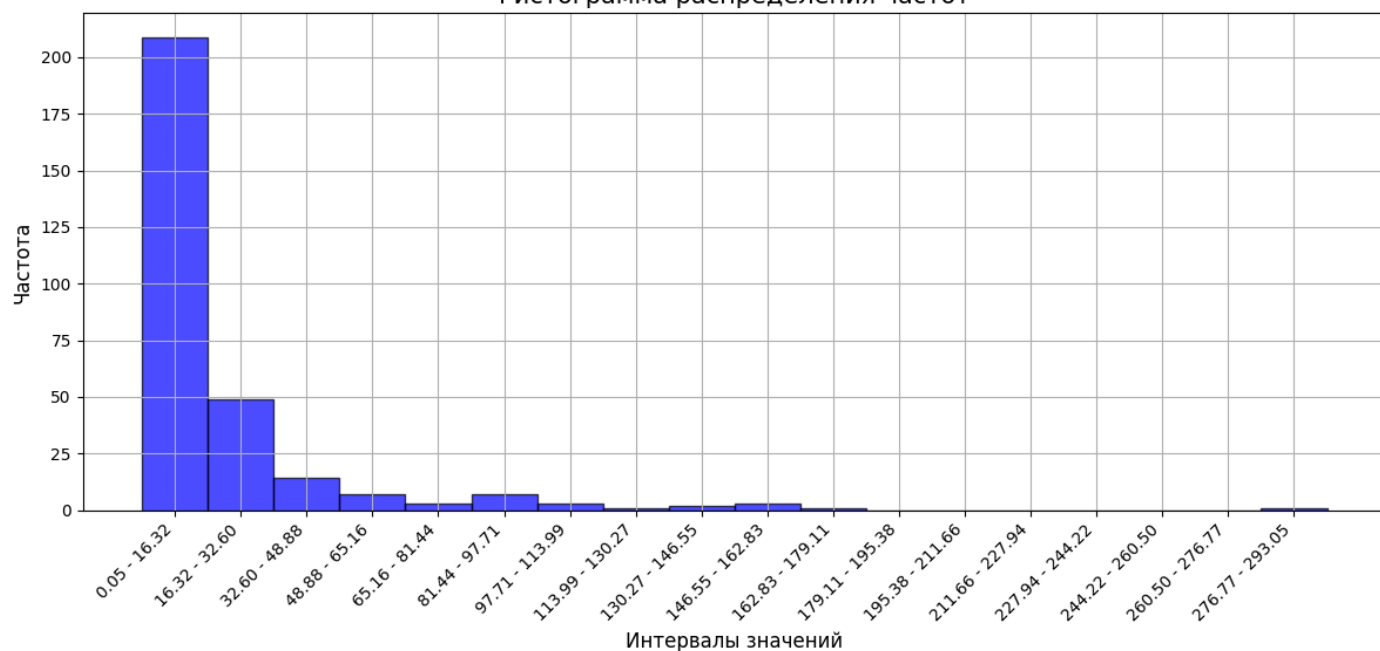
<u>Закон распределения: гиперэкспоненциальный закон</u>							
Характеристика		Количество случайных величин					
		10	20	50	100	200	300
Мат. ож.	Знач,	13,654073	15,910140	14,964120	19,414181	17,988566	19,26444
	%	70,88	82,59	77,68	100,78	93,38	
Дов. инт. (0,9)	Знач,	±9,0061	±6,9756	±5,5409	±5,2282	±3,16	±3,082
	%	±292,22	±226,33	±179,78	±169,64	±102,53	
Дов. инт. (0,95)	Знач,	±10,7315	±8,3119	±6,6024	±6,2298	±3,7654	±3,6724
	%	±292,22	±226,33	±179,78	±169,64	±102,53	
Дов. инт. (0,99)	Знач,	±14,1035	±10,9237	±8,6771	±8,1873	±4,9485	±4,8263
	%	±292,22	±226,34	±179,79	±169,64	±102,53	
Дисперсия	Знач,	299,7927	359,6961	567,3919	1010,2911	738,1539	1053,219
	%	28,46	34,15	53,87	95,92	70,09	
С. к. о.	Знач,	17,3145	18,9657	23,8200	31,7851	27,1690	32,4533
	%	53,35	58,44	73,4	97,94	83,72	
К-т вариации	Знач,	1,268085	1,192048	1,591807	1,637209	1,510348	1,684625
	%	75,27	70,76	94,49	97,19	89,65	

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

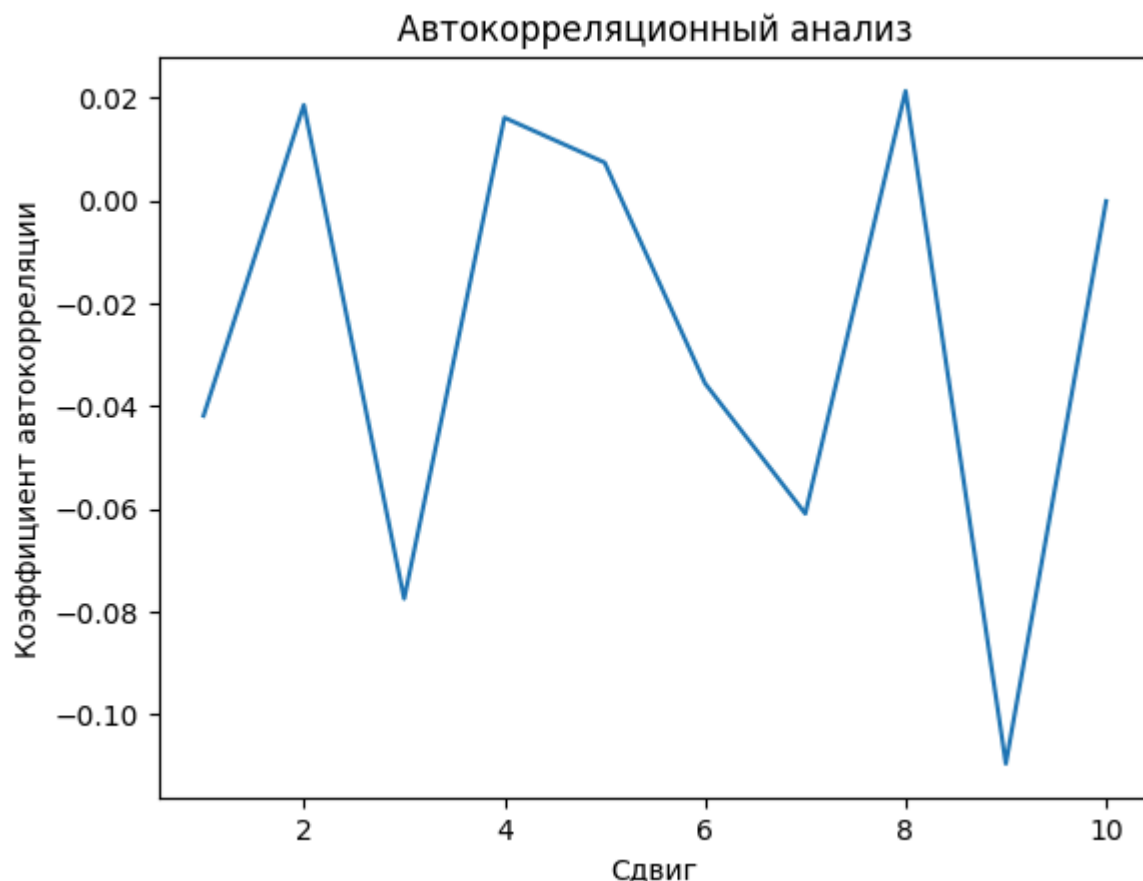
График числовой последовательности



Гистограмма распределения частот



При сравнении полученных гистограмм видно, что полученная нами последовательность похожа на исходную. Тем самым, мы доказали, что выбранная нами аппроксимация подходит.



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

Коэффициент корреляции между двумя числовыми последовательностями:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \times \sum_{i=1}^n (y_i - \bar{y})^2}} = 0,007997$$

Для сгенерированной и полученной последовательности мы рассчитали корреляционную зависимость. Как мы видим из $0.09 \geq |r|$, корреляции между исходной и сгенерированной случайной последовательностями нет.

Вывод из 7 этапа: 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.

Выводы

В рамках лабораторной работы была дана числовая последовательность, для которой мы определили математическое ожидание, дисперсию и другие параметры. Далее мы проанализировали построенную гистограмму, по которой не было выявлено возрастания, убывания или периодичности последовательности. Исследуемую последовательность можно назвать случайной исходя из автокорреляционного анализа. Затем мы вычислили параметры аппроксимирующего закона и по ним сгенерировали новую последовательность. Коэффициент вариации первой и второй последовательности приближен к 1.7. Коэффициент автокорреляции первой и второй последовательности варьируется около нуля, исходя из этого можно сделать вывод о том, что выборка случайна. Математическое ожидание и дисперсия отличаются, но отличие не выходит за пределы доверительных интервалов.