

**Федеральное государственное автономное
образовательное учреждение высшего образования
«Санкт-Петербургский национальный исследовательский университет
информационных технологий, механики и оптики»**



ITMO UNIVERSITY

Факультет Программной Инженерии и Компьютерной Техники

Учебно-исследовательская работа №1

по дисциплине

«МОДЕЛИРОВАНИЕ»:

Вариант 73

Выполнил:

Ястребов-Амирханов Алекси

ГРУППА: Р3332

ПРЕПОДАВАТЕЛЬ: Алиев Тауфик Измайлович

Санкт-Петербург,
2025

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.

Задание

В процессе исследований необходимо выполнить обработку заданной числовой последовательности (ЧП) для случаев, когда путем измерений получено 10, 20, 50, 100, 200 и 300 значений случайной величины, а именно:

- рассчитать значения следующих числовых моментов заданной числовой последовательности:
 - математическое ожидание;
 - дисперсию;
 - среднеквадратическое отклонение;
 - коэффициент вариации;
 - доверительные интервалы для оценки математического ожидания с доверительными вероятностями 0,9; 0,95 и 0,99;
 - относительные отклонения (в процентах) полученных значений от наилучших значений, полагая, что наилучшими (эталонными) являются значения, рассчитанные для наиболее представительной выборки из трехсот случайных величин;
- построить график значений для заданной числовой последовательности и определить ее характер, а именно: является эта последовательность возрастающей/убывающей, периодичной (при наличии периодичности оценить по графику длину периода);
- выполнить автокорреляционный анализ и определить, можно ли заданную числовую последовательность считать случайной;
- построить гистограмму распределения частот для заданной числовой последовательности;
- выполнить аппроксимацию закона распределения заданной случайной последовательности по двум начальным моментам, используя, в зависимости от значения коэффициента вариации, одно из следующих распределений:
 - ✓ равномерный;
 - ✓ экспоненциальный;
 - ✓ нормированный Эрланга k -го порядка или гипотенциальный с заданным коэффициентом вариации;
 - ✓ гиперэкспоненциальный с заданным коэффициентом вариации;
- реализовать генератор случайных величин в соответствии с полученным аппроксимирующим законом распределения (в EXEL или программно) и проиллюстрировать на защите его работу;

- сгенерировать последовательность случайных величин с использованием реализованного генератора и рассчитать значения числовых моментов по аналогии с заданной числовой последовательностью;
- выполнить автокорреляционный анализ сгенерированной последовательности случайных величин;
- выполнить сравнительный анализ сгенерированной последовательности случайных величин с заданной последовательностью, построив соответствующие зависимости на графике значений и гистограмме распределения частот;
- оценить корреляционную зависимость сгенерированной и заданной последовательностей случайных величин.

Результаты проводимых исследований представить в виде таблиц и графиков. На основе полученных промежуточных и конечных результатов следует сделать обоснованные выводы об исследуемой числовой последовательности, предложить закон распределения для ее описания и оценить качество аппроксимации этим законом.

Ход работы

Этап 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.

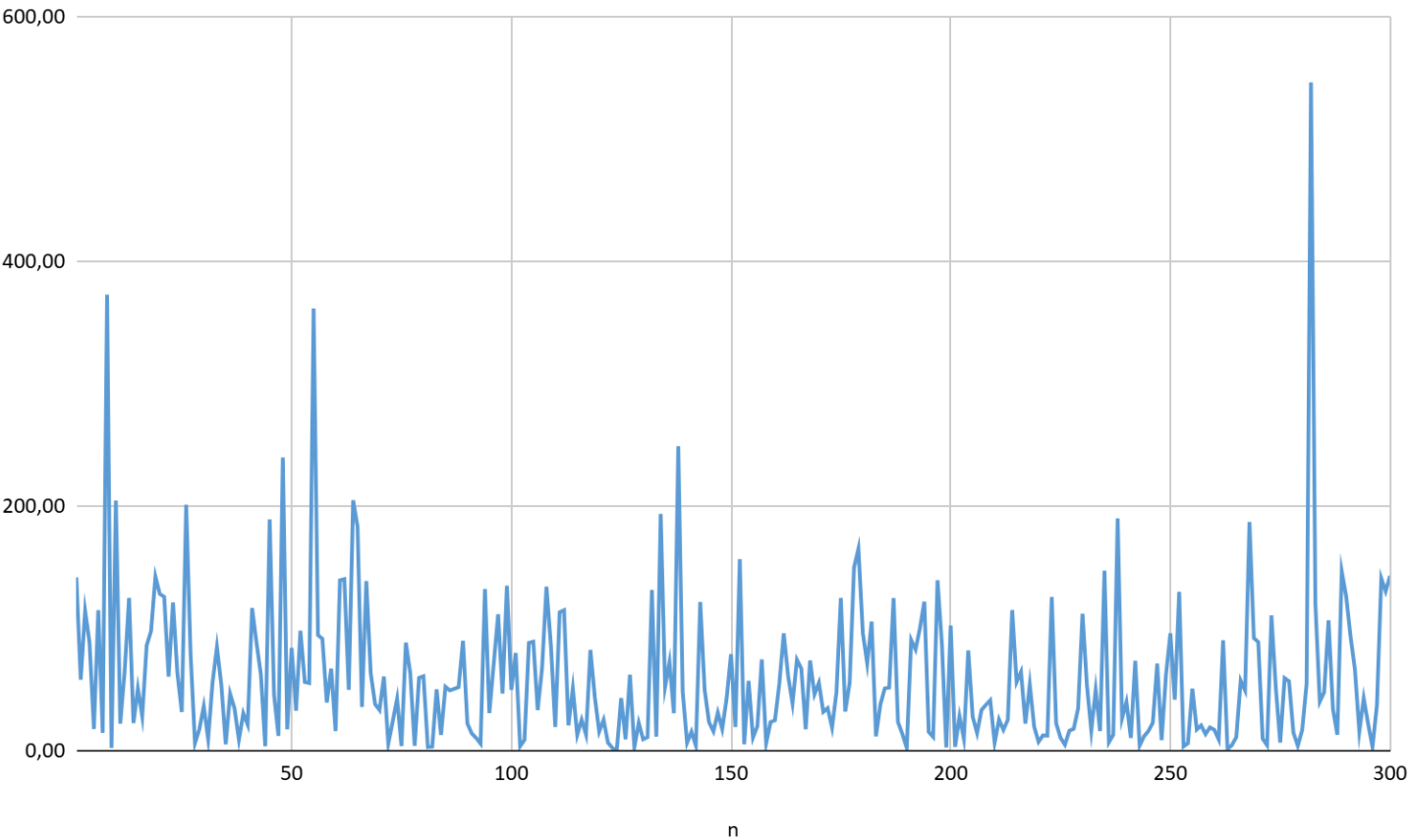
Характеристика		Количество случайных величин					
		10	20	50	100	200	300
Мат.ож.	Знач.	112,695	94,623	76,763	71,695	63,123	59,232
	%	90,262	59,751	29,598	21,041	6,570	
Дов. инт. (0,9)	Знач.	57,839	31,188	16,742	11,137	6,889	5,902
	%	879,956	428,413	183,658	88,687	16,722	
Дов. инт. (0,95)	Знач.	68,998	37,205	19,972	13,285	8,218	7,041
	%	879,956	428,413	183,658	88,687	16,722	
Дов. инт. (0,99)	Знач.	90,683	48,898	26,249	17,461	10,801	9,254
	%	879,956	428,413	183,658	88,687	16,722	
Дисперсия	Знач.	12392,610	7206,544	5191,718	4594,436	3516,316	3871,427
	%	220,104	86,147	34,103	18,676	-9,173	
с.к.о.	Знач.	111,322	84,891	72,054	67,782	59,299	62,221
	%	78,915	36,436	15,803	8,938	-4,697	
К.В.	Знач.	0,988	0,897	0,939	0,945	0,939	1,050
	%	-5,964	-14,595	-10,644	-9,999	-10,572	

Вывод:

Математическое ожидание убывает при увеличении объема выборки, что может говорить о превосходстве больших величин в начале выборки. В общем математическое ожидание показывает убывающую тенденцию. Процент отклонения значений доверительного интервала уменьшается пропорционально увеличению количества значений в выборке.

Этап 2. Значений заданной ЧП с результатами анализа характера числовой последовательности.

Заданная числовая последовательность

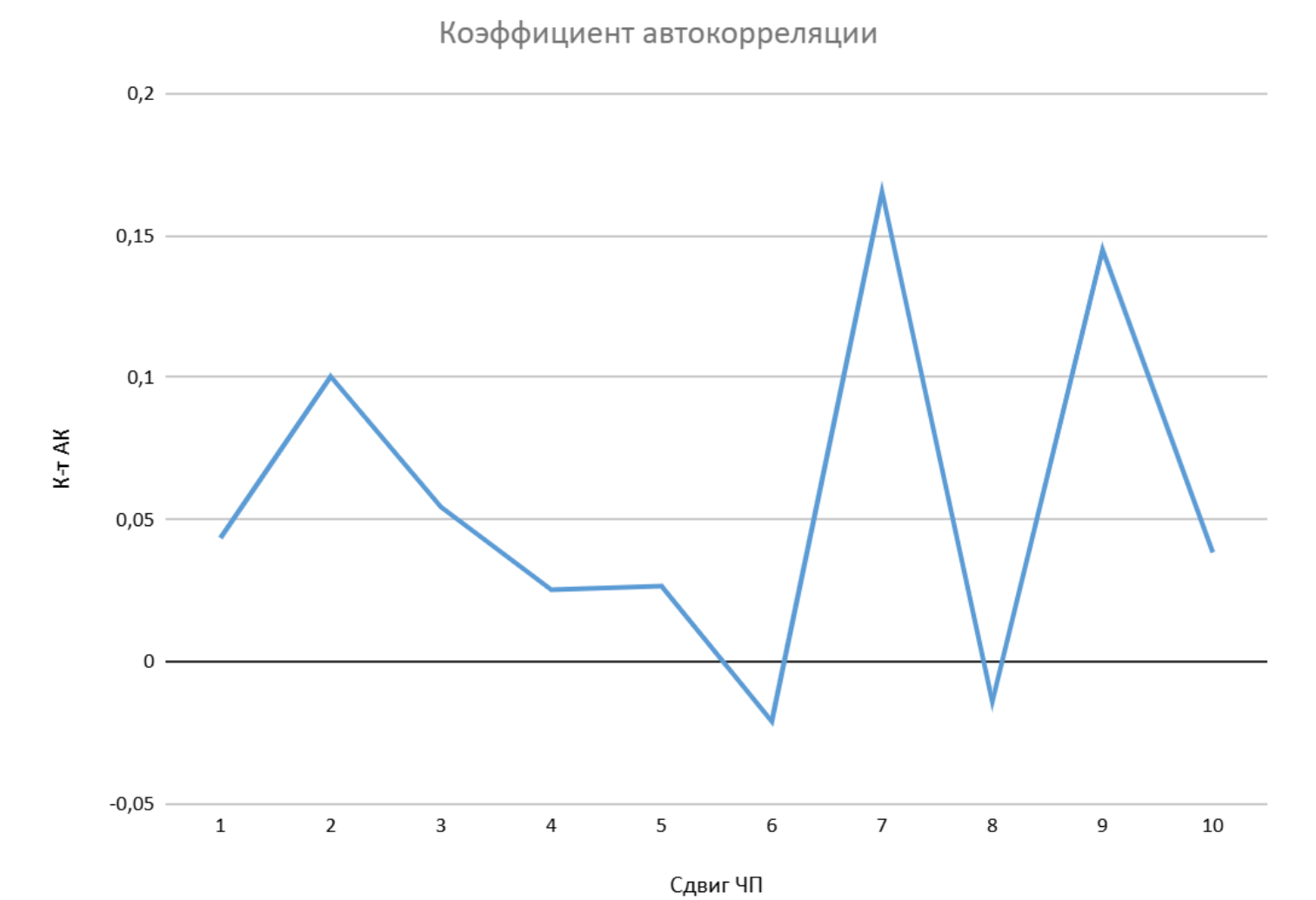


Вывод:
Исходная последовательность не является возрастающей, убывающей или периодичной.

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

Коэффициенты автокорреляции

Сдвиг ЧП	1	2	3	4	5	6	7	8	9	10
К-т АК для задан. ЧП	0,0435	0,1003	0,0544	0,0252	0,0265	-0,0210	0,1652	-0,0141	0,1449	0,0384



Вывод:
Коэффициенты указывают, что зависимости между значениями последовательности нет. Последовательность можно считать случайной.

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



Вывод:
По гистограмме частот можно определить, что большая половина значений находится в диапазоне от 0 до 50, еще одна большая часть находится от 50 до 300, наименьшее количество значений располагается в диапазоне от 400 до 450, и от 550 до 600.

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

Часть 1

Проводим вычисления:

Первый начальный момент (математическое ожидание)

$$M[X] = \alpha_1 = 59.232$$

Второй начальный момент

$$\alpha_2 = 7366.907$$

Дисперсия

$$D[X] = \alpha_2 - (\alpha_1)^2 = 3858.528$$

Стандартное отклонение

$$\sigma = \sqrt{D[x]} = 62.117$$

Коэффициент вариации

$$v = \frac{\sigma}{M[x]} = 1.049$$

Поскольку $v > 1$, для аппроксимации закона распределения используем гиперэкспоненциальное распределение с заданным коэффициентом вариации.

Часть 2

Аппроксимация гиперэкспоненциальным распределением

Вычисляем параметр p (вероятность первой фазы):

$$p = \frac{1 - \sqrt{1 - \frac{2}{1+v^2}}}{2} = 0.391$$

Параметры распределения

$$\alpha_1 = \frac{2}{M[X]} = 0.013$$

$$\alpha_2 = \frac{2(1-p)}{M[X]} = 0.021$$

Часть 3

Закон распределения аппроксимируется гиперэкспоненциальным:

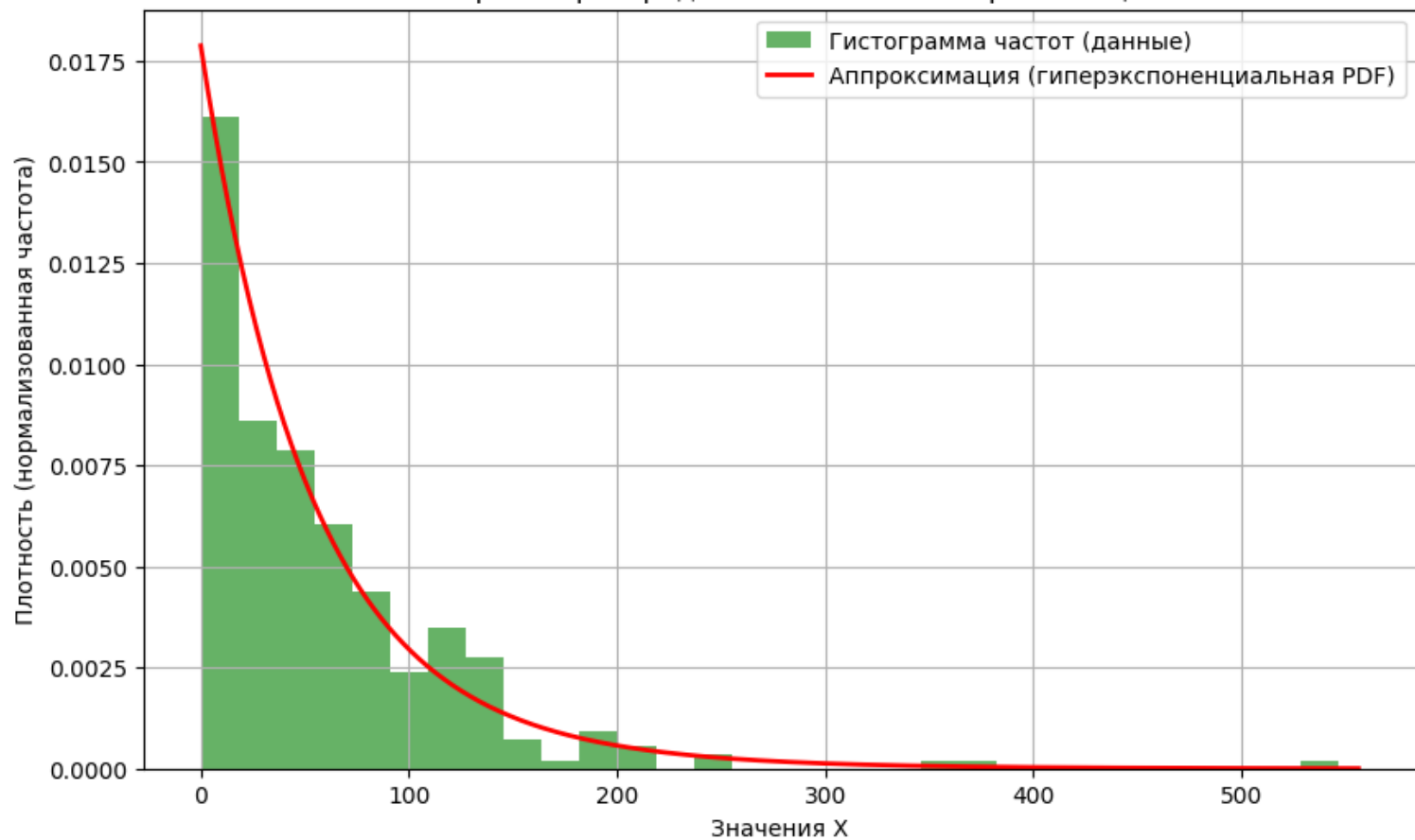
Плотность распределения:

$$f(x) = 0.391 \cdot 0.013 \cdot e^{-0.013x} + 0.609 \cdot 0.021e^{-0.021x}$$

Функция распределения:

$$F_{(x)} = 1 - 0.391 \cdot e^{-0.013x} - 0.609 \cdot e^{-0.021x}$$

Гистограмма распределения частот с аппроксимацией



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

Генератор выполнен в среде Excel

мы аппроксимировали распределение двухфазным гиперэкспоненциальным законом с параметрами:

$p = 0.391$ (вероятность выбора первой фазы),

$\lambda_1 = 0.013$ (интенсивность первой экспоненциальной фазы),

$\lambda_2 = 0.021$ (интенсивность второй фазы).

Генерация будет выполняться в одной ячейке по формуле

`=IF(RAND() < 0.391; -LN(RAND())/0.013; -LN(RAND())/0.021)`

Здесь три вызова RAND(): один для выбора фазы, один для V. Они независимы, так что это работает корректно.

Чтобы убедиться, что распределение соответствует аппроксимации:

- 1) Вычислить среднее: В отдельной ячейке используем `=AVERAGE(C2:C301)` (для 300 значений). Получилось 58,66313275, что близко к 59.232 (для изначальных данных)
- 2) Вычислить дисперсию: В отдельной ячейке используем `=VAR.S(C2:C301)`. Получилось 3808,750603 что близко к 3871,427 (для изначальных данных)
- 3) Далее коэффициент вариации: В отдельной ячейке используем `=STDEV.S(C2:C301) / AVERAGE(C2:C301)` (STDEV.S — это выборочное стандартное отклонение (σ , аналог $\sqrt{\text{VAR.S}}$). Получили 1,052024893, что так же близко к 1,049 (для изначальных данных)

Вывод:

Алгоритм генерации величин по гиперэкспоненциальному закону был реализован. Проверка и совпадение Среднего, Дисперсии и коэффициент вариации показывают, что алгоритм был реализован правильно.

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

1. расчет числовых характеристик сгенерированной в соответствии с аппроксимирующим законом распределения случайной последовательности: математического ожидания, дисперсии, среднеквадратического отклонения, коэффициента вариации и коэффициентов автокорреляции при разных значениях сдвигов, а также сравнения (в %) полученных значений со значениями, рассчитанными для заданной числовой последовательности

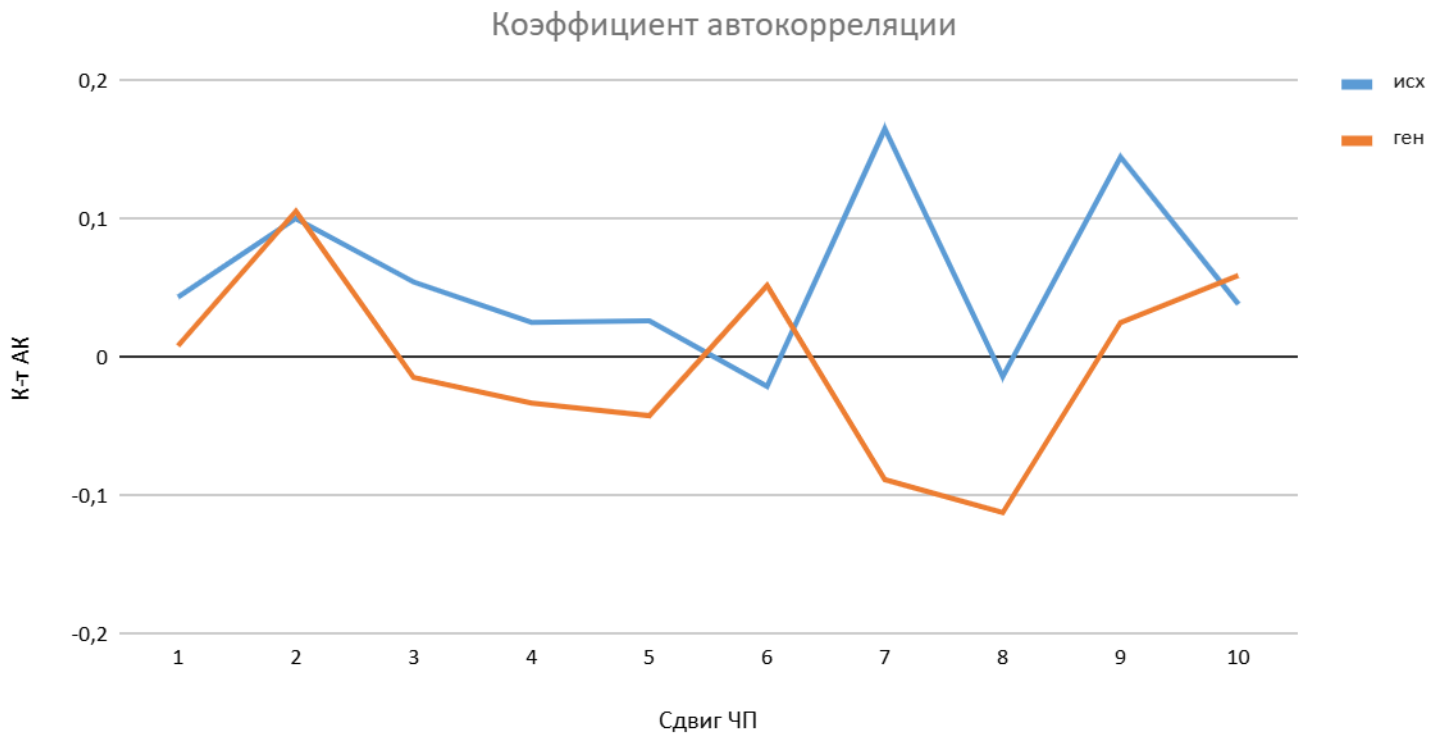
Характеристика		Количество случайных величин					
		10	20	50	100	200	300
Мат.ож.	Знач	62,273	62,052	64,460	56,350	55,000	56,952
	%	44,742	34,422	16,027	21,403	12,869	3,848
Дов. инт. (0,9)	Знач	35,659	23,254	14,897	10,002	6,675	5,669
	%	38,347	25,440	11,023	10,192	3,103	3,948
Дов. инт. (0,95)	Знач	42,540	27,740	17,771	11,931	7,963	6,763
	%	38,347	25,440	11,023	10,192	3,103	3,948
Дов. инт. (0,99)	Знач	55,909	36,459	23,356	15,681	10,466	8,888
	%	38,347	25,440	11,023	10,192	3,103	3,948
Дисперсия	Знач	4710,581	4006,293	4110,278	3705,639	3301,488	3571,753
	%	61,989	44,408	20,830	19,345	6,109	7,741
с.к.о.	Знач	68,634	63,295	64,111	60,874	57,459	59,764
	%	38,347	25,440	11,023	10,192	3,103	3,948
к.в.	Знач	1,102	1,020	0,995	1,080	1,045	1,049
	%	11,573	13,697	5,959	14,264	11,209	0,104

Математическое ожидание отличается от математического ожидания исходной выборки на величину, не превышающую доверительные интервалы, что говорит о достоверной аппроксимации

2. проведение корреляционного анализа сгенерированной в соответствии с аппроксимирующим законом распределения последовательности случайных величин и заданной числовой последовательности на основе коэффициента корреляции

Сдвиг ЧП	1	2	3	4	5	6	7	8	9	10
К-т АК для задан. ЧП	0,0435	0,1003	0,0544	0,0252	0,0265	-0,0210	0,1652	-0,0141	0,1449	0,0384
К-т АК для сгенерир. ЧП	0,0083	0,1056	-0,0144	-0,0332	-0,0422	0,0519	-0,0883	-0,1122	0,0252	0,0592
%	80,9	5,3	126,6	231,3	259,2	346,8	153,5	698,9	82,6	54,3

Коэффициенты автокорреляции приближен к 0 и сравнение с заданной ЧП не показывает сходства, что говорит о случайности выборки



Значения близки к 0, что доказывает случайность последовательности.

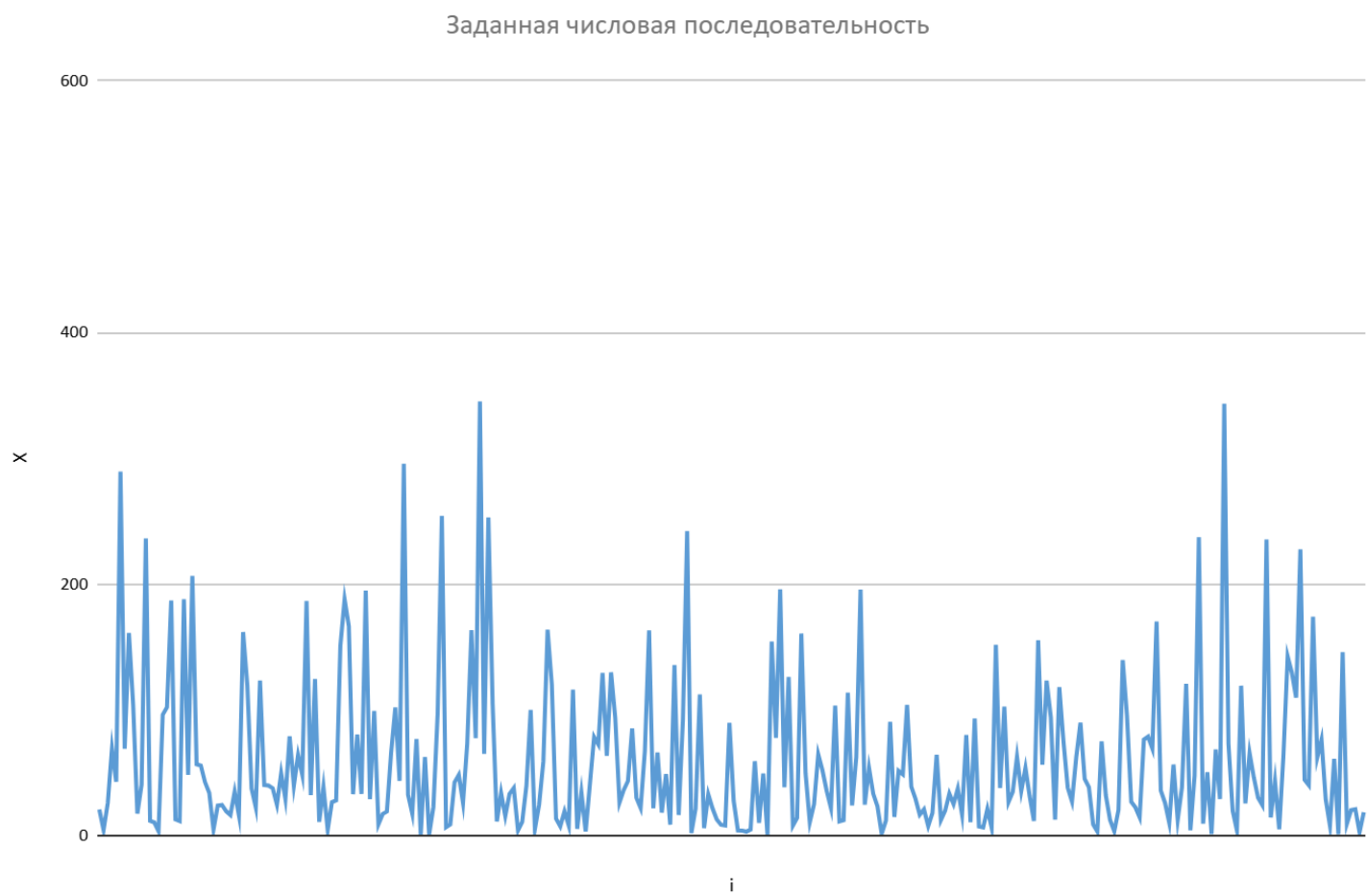
Коэффициент корреляции между сгенерированной и исходной ЧП также близок к 0

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} = 0.0209$$

Это доказывает независимость сгенерированной и исходной последовательностей

Вывод: Для сгенерированной и полученной последовательности был рассчитан коэффициент корреляции, значение которого очень мало. Следовательно линейная корреляция крайне слаба, поскольку при аппроксимации основной упор делался на совпадение плотностей распределения, а не на линейную зависимость.

Значения сгенерированной ЧП



Вывод: Сгенерированная ЧП внешне не похожа на исходную ЧП. Так как алгоритм моделирует распределение, а не зависимость от временной составляющей.

Выводы

1. В лабораторной работе были вычислены математическое ожидание, дисперсия, среднеквадратическое отклонение, коэффициент вариации и доверительные интервалы.
2. Провели автокорреляционный анализ и доказали случайность числовой последовательности.
3. По первым двум моментам и коэффициенту вариации сделали предположение о виде аппроксимирующего закона распределения – гиперэкспоненциального распределение.
4. На основании рассчитанных характеристик был смоделирован генератор случайных величин.
5. В целом, исследования по аппроксимации и сравнению исходной и сгенерированной последовательностей подтвердили высокую степень соответствия как по визуальным, так и по числовым характеристикам.