**Факультет Программной Инженерии и Компьютерной техники**

**Моделирование**

**Лабораторная работа №1**

**«Обработка результатов измерений: статистический анализ числовой последовательности»**

**Вариант 46**

**Выполнили:**

**Тюрин И.Н.**

**Сосновцев Г.А.**

**Группа P34102**

[**Порядок выполнения работы 2**](#_b25gz3102e4t)

[**Ход работы 4**](#_rm6ycxnb2xyy)

[Этап 1: расчёт значений числовых моментов исходной числовой последовательности 4](#_rglvaht92sbc)

[Этап 2: построение графика значений для заданной числовой последовательности и определение ее характера 4](#_nf22szd2wn8y)

[Этап 3: автокорреляционный анализ 5](#_fckah2ykze33)

[Этап 4: построение гистограммы распределения частот для заданной числовой последовательности 6](#_h1y4cl6ci50o)

[Этап 5: выполнение аппроксимации закона распределения заданной случайной последовательности по двум начальным моментам, используя распределение Эрланга k-го порядка 7](#_ajxcyo6k5dtw)

[Этап 6: реализация генератора случайных величин в соответствии с полученным аппроксимирующим законом распределения 8](#_yw33ndum9wq)

[Этап 7: генерация последовательности случайных величин в соответствии с полученным законом распределения и расчёт значения числовых моментов по аналогии с заданной числовой последовательностью 8](#_9y106eyp6khm)

[Этап 8: автокорреляционный анализ сгенерированной последовательности случайных величин 9](#_7xghnotd2rfk)

[Этап 9: сравнительный анализ сгенерированной последовательности случайных величин с заданной последовательностью, построив соответствующие зависимости на графике значений и гистограмме распределения частот 10](#_jwmcx72bshsb)

[Этап 10: оценка корреляционной зависимости сгенерированной и заданной последовательностей случайных величин 12](#_7o2xmz4dtoes)

[**Вывод 12**](#_ic1ckxnz1du1)

# **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.

# **Порядок выполнения работы**

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 гипоэкспоненциальный 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**

## Исследование статистических характеристик заданной числовой последовательности

Изобразим заданную ЧП на графике. По графику (см. рисунок №1) можно сделать вывод, что последовательность не возрастающая/убывающая и не периодичная. На графике можно видеть пики, что говорит о наличии выбрасов в последовательности.

Рисунок №1. График значений исходной числовой последовательности

## 

Провели анализ числовых моментов, результаты которых занесли в таблицу №1, и выявили факты требующие детального рассмотрения:

1. относительная погрешность матожидани для 200 измерений после 100 измерений увеличивается по абсолютному значению;
2. дисперсия (и СКО) увеличивается для 100 измерений после 50 измерений.

Таблица №1. Числовые моменты исходной последовательности

| Характеристика |  | Количество случайных величин | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 10 | 20 | 50 | 100 | 200 | 300 |
| Мат. ож. | Знач. | 259,680 | 205,271 | 154,357 | 168,370 | 162,965 | 168,502 |
| % | 54,111 | 21,821 | -8,395 | -0,079 | -3,286 |
| Дов. инт. (0,9) | Знач. | ±71,883 | ±47,712 | ±28,066 | ±21,132 | ±13,878 | ±11,061 |
| % | 42,660 | 28,315 | 16,656 | 12,541 | 8,236 |
| Дов. инт. (0,95) | Знач. | ±85,752 | ±56,918 | ±33,481 | ±25,209 | ±16,556 | ±13,195 |
| % | 50,891 | 33,779 | 19,870 | 14,960 | 9,825 |
| Дов. инт. (0,99) | Знач. | ±112,702 | ±74,806 | ±44,003 | ±33,131 | ±21,760 | ±17,342 |
| % | 66,885 | 44,395 | 26,114 | 19,662 | 12,913 |
| Дисперсия | Знач. | 19141,472 | 16866,061 | 14589,874 | 16541,952 | 14270,406 | 13596,141 |
| % | 40,786 | 24,050 | 7,309 | 21,667 | 4,959 |
| СКО | Знач. | 138,353 | 129,869 | 120,789 | 128,616 | 119,459 | 116,602 |
| % | 18,653 | 11,378 | 3,590 | 10,303 | 2,450 |
| К-т вариации | Знач. | 0,533 | 0,633 | 0,783 | 0,764 | 0,733 | 0,692 |
| % | -23,008 | -8,572 | 13,083 | 10,389 | 5,931 |

% — относительная погрешность (относительно всей выборки), рассчитывается как **.**

Пункт 1 и 2 можно объяснить тем, что мы имеем дело со случайной величиной и в следующих после 50 и 100 первых чисел есть несколько значительных выбрасов, что можно видеть на рис. 1., они сильно оттягивают мат. ожидание и влияют на дисперсию с СКО.

В целом значения характеристик стремятся к характерисикам для полной выборки, относительная погрешность снижается с увеличением числа элементов, что подтверждает закон больших чисел.

По гистограмме (см. рисунок №2) видно, что 68% значений находятся в промежутке от 0 до 200. Дальше частота убывает. Для подтверждения наблюдения можно посчитать квартили.

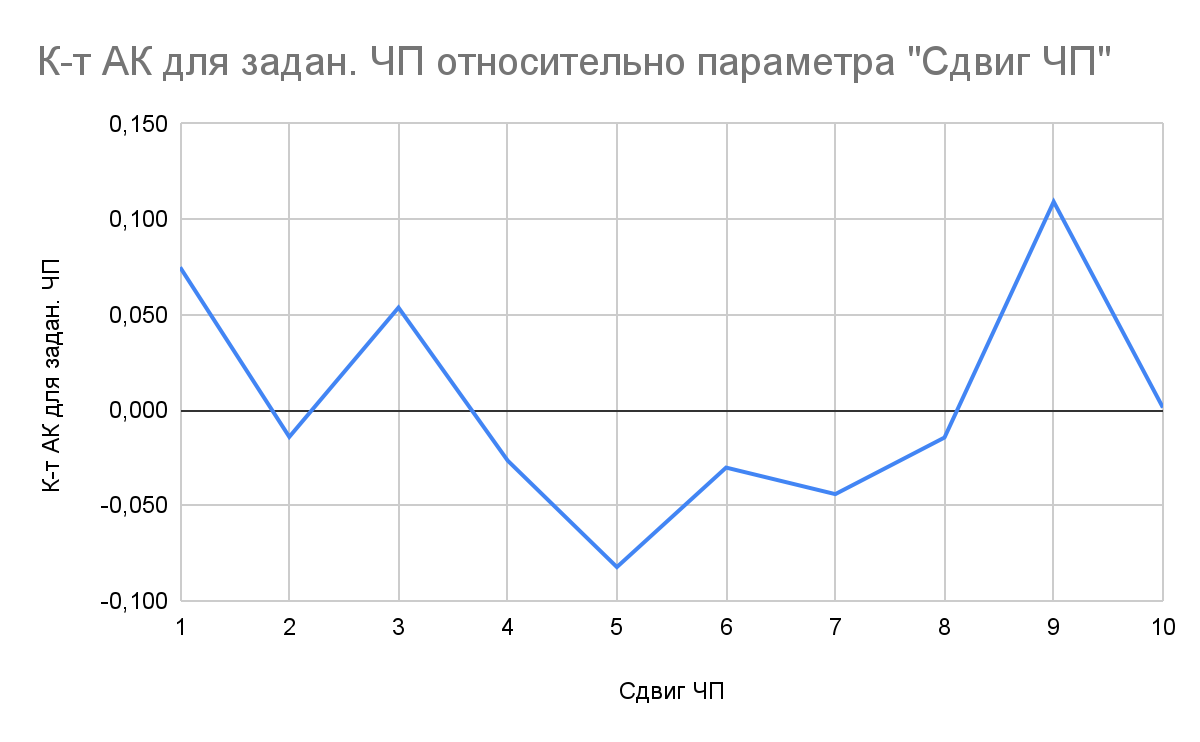
Минимальное значение ЧП – 7,93; максимальное – 650,38; Q1 – 83,39; Q2 – 142,00; Q3 – 220, 91 подтверждают наблюдение.

Рисунок №2. Гистограмма распределения частот для исходной последовательности



Таблица №2. Коэффициенты автокорреляции данной ЧП

| Сдвиг ЧП | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| К-т АК для задан. ЧП | 0,075 | -0,014 | 0,054 | -0,026 | -0,082 | -0,030 | -0,044 | -0,014 | 0,109 | 0,001 |

Рисунок №3. График коэффициентов автокорреляции

Выполнив автокорреляционный анализ с разными значениями смещения (от 1 до 10), можно сказать, что внутри последовательности между значениями измерений нет зависимости, тенденции или периодичности. Заметим, что есть несколько значения больше нуля (см. рис. №3), остальные меньше, но все они близки к нулю, что позволяет нам считать эту числовую последовательность случайной.

## **Выполнение аппроксимации закона распределения заданной случайной последовательности по двум начальным моментам**

Получено математическое ожидание и коэффициентом вариации :

| v | 0,692 |
| --- | --- |
| t | 168,502 |

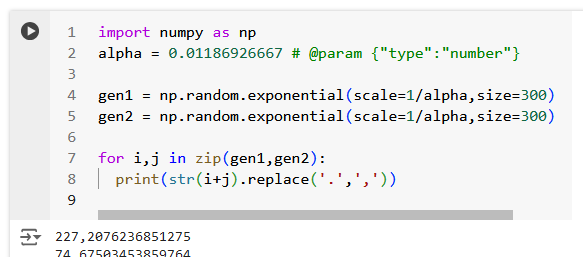
Коэффициент вариации лежит в интервале , поэтому рассмотрим распределение Эрланга k-го порядка для аппроксимации нашего распределения. Значения соответствующих параметров найдем по формулам:

; ; ; ; .

Коэффициен очень близко, поэтому при округлении было использовано округление вниз, хотя и предписывается округлить вверх до ближайшего целого.

| k | 2 |
| --- | --- |
| M[τ] | 84,25120333 |
| M\_E\_k | 252,75361 |
| v\_E\_k | 0,7071067812 |
| α | 0,01186926667 |

Генератор случайных чисел для распределения Эрланга 2-го порядка строится на основе двух генераторов экспоненциального распределения с вычесленным коэффициентом , а результирующее распределение получается как их сумма.

Рисунок №4. Фрагмент кода на ЯП Python для генерации экспоненциальных распределений в составе распределения Эрланга 2-го порядка.

Для реализации генератора согласно аппроксимации с помощью распределения Эрланга k-го порядка использовали сумму двух полученных экспоненциальных распределения, сгенерированных с помощью библиотеки numpy для Python (см. рис. №4).

Альтернативно можно было бы использовать табличный процессор, сгенерировав массив случайных чисел (вероятностей в интервале (0;1)) с помощью =RANDARRAY(301,1) и использовав их для получения экспонциального распределения с помощью формулы обратной функции распределения:.

## 

## 

## **Анализ статистических характеристик сгенерированной числовой последовательности и сравнение с исходной**

Рисунок №4. График сравнения значений сгенерированной числовой последовательности

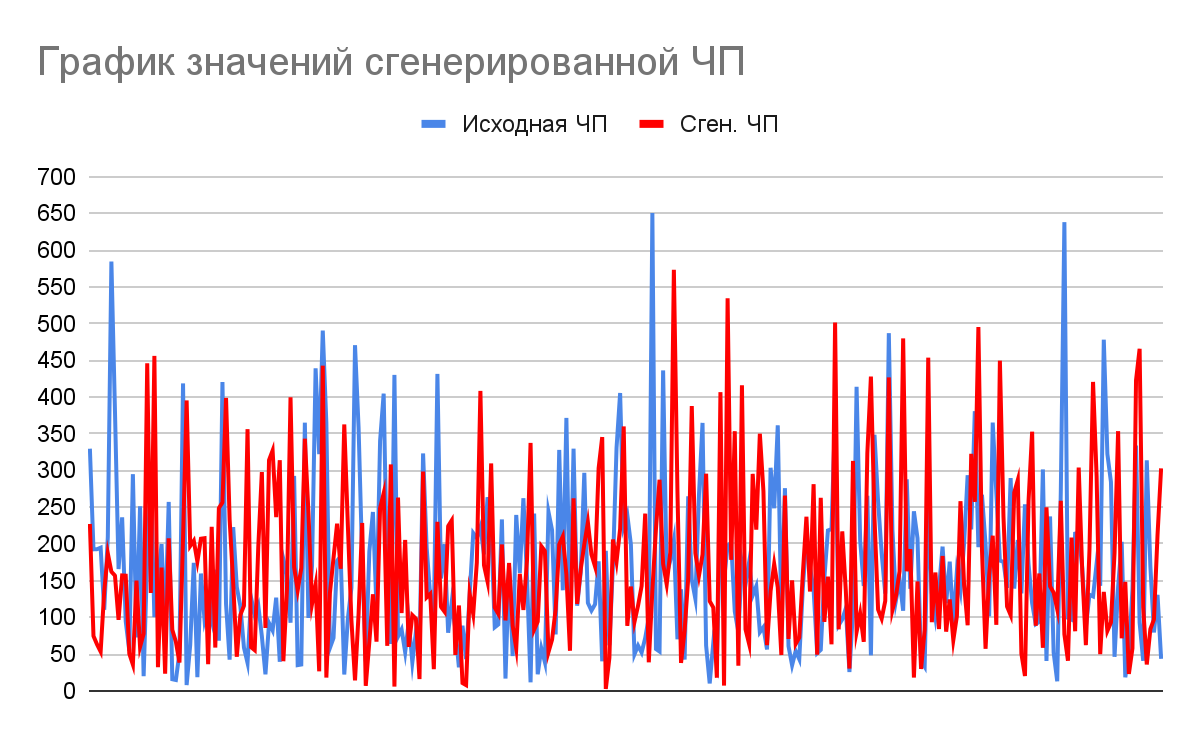


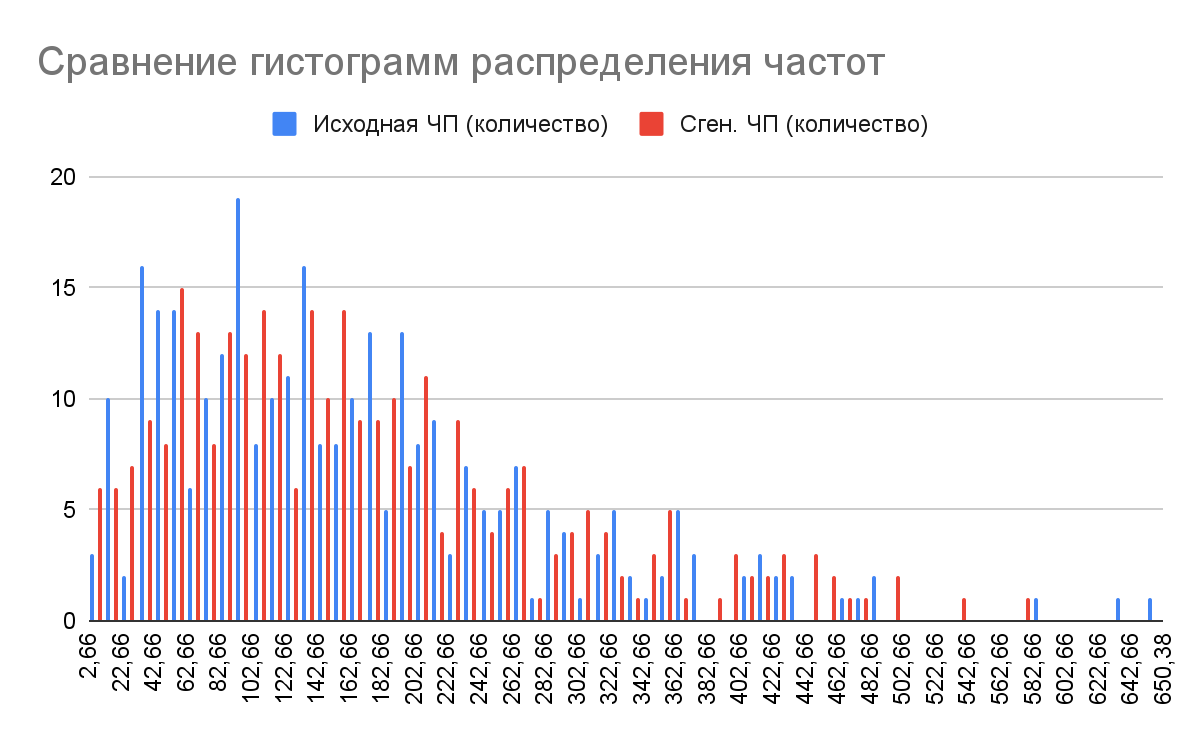
Рисунок №5. Гистограмма распределения частот для исходной и сгенерированной последовательностей

Таблица №3. Характеристики сгенерированной случайной ЧП

| Характеристика |  | Количество случайных величин | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 10 | 20 | 50 | 100 | 200 | 300 |
| Мат. ож. | Знач. | 131,429 | 145,626 | 160,302 | 164,506 | 168,330 | 171,633 |
| % | -49,388 | -29,056 | 3,851 | -2,295 | 3,292 | 1,858 |
| Дов. инт. (0,9) | Знач. | ±29,863 | ±43,339 | ±25,655 | ±18,183 | ±12,940 | ±10,987 |
| % | -58,455 | -9,166 | -8,590 | -13,955 | -6,765 | -0,670 |
| Дов. инт. (0,95) | Знач. | ±35,625 | ±51,701 | ±31,548 | ±21,656 | ±15,416 | ±13,174 |
| % | -58,455 | -9,166 | -5,774 | -14,094 | -6,887 | -0,158 |
| Дов. инт. (0,99) | Знач. | ±41,733 | ±68,349 | ±41,427 | ±28,379 | ±20,270 | ±17,375 |
| % | -62,971 | -8,632 | -5,855 | -14,345 | -6,843 | 0,190 |
| Дисперсия | Знач. | 3303,714 | 13915,834 | 12191,031 | 12247,116 | 12404,948 | 13414,465 |
| % | -82,741 | -17,492 | -16,442 | -25,963 | -13,072 | -1,336 |
| С.к.о. | Знач. | 57,478 | 117,965 | 110,413 | 110,667 | 111,378 | 115,821 |
| % | -58,455 | -9,166 | -8,590 | -13,955 | -6,765 | -0,670 |
| К-т вариации | Знач. | 0,437 | 0,810 | 0,689 | 0,673 | 0,662 | 0,675 |
| % | -17,916 | 28,037 | -11,980 | -11,934 | -9,736 | -2,482 |

% — относительная погрешность (относительно заданной ЧП), рассчитывается как .

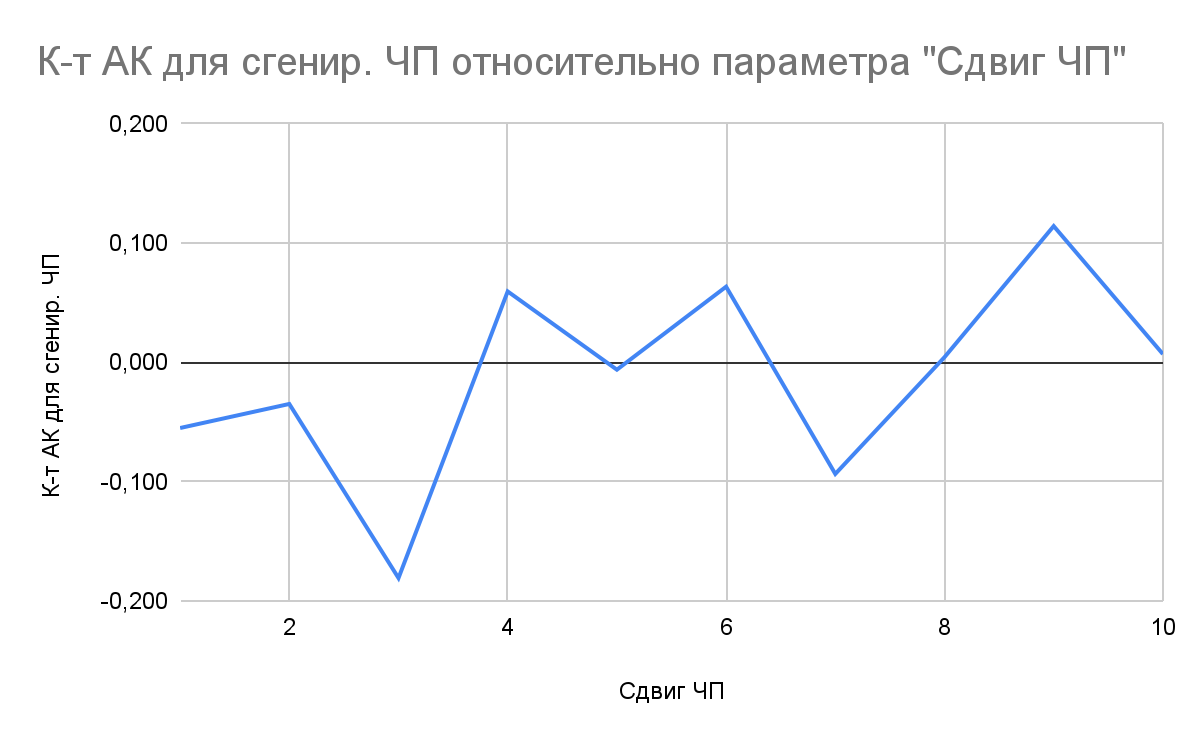
В таблице №3 представлены характеристики сгенерированного распределения в сравнении с исходным. Мат. ожидание сгенерированной последовательности отличается от мат. ожидания исходной последовательности на величину, не превосходящую доверительные интервалы. Так же присутствуют “аномалии” связанные с выбросами, которые в принципе даже похожи на онные в исходной ЧП.

Автокорелляционный анализ сгенерированной последовательности показал (см. таблицу №4), что она очень схожа с данной ЧП. Коэффициенты приближены к нулю и довольно далеки от единицы (см. рис. №6). То есть, автокорреляционный анализ не показывает между числами лаговой/сдвиговой зависимости, тенденции или периодичности.

Таблица №4. Коэффициенты автокорреляции

| Сдвиг ЧП | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| К-т АК для задан. ЧП | 0,075 | -0,014 | 0,054 | -0,026 | -0,082 | -0,030 | -0,044 | -0,014 | 0,109 | 0,001 |
| К-т АК для сгенир. ЧП | -0,055 | -0,035 | -0,181 | 0,059 | -0,006 | 0,063 | -0,094 | 0,004 | 0,114 | 0,007 |
| % | -235,963 | -59,721 | -129,705 | -144,540 | 1239,135 | -147,509 | -52,914 | -431,223 | -4,454 | -81,110 |
| К-т корреляции | 0,069 |  |  |  |  |  |  |  |  |  |

% — отношение полученных значений: × 100%

Рисунок №6. График коэффициентов автокорреляции

Коэффициент корреляции исходной и сгенерированной последовательностей равен 0,069. Между последовательностями корреляция очень низкая, они не имеют зависимости при схожих моментах, что говорит о хорошем качестве построенной модели распределения.

Сгенерированная последовательность очень близка к исходной, судя по графику сгенерированных значений (см. рисунок №4) и гистограмме частот (см. рисунок №5). Синяя линия и синие столбики — сгенерированная последовательность, красные — исходная последовательность.

Рисунок №7. График плотности сгенерированного распределения



Видим, что плотность сгенерированной ЧП (см. рисунок №7) похожа на гистограмму распределенных частот для заданной ЧП (см. график на рисунке №5), значит распределение Эрланга было посчитано верно. Из сравнения двух гистограмм видим совпадения в поведении ЧП.

# Вывод

В ходе выполнения работы мы изучили методы статистического анализа результатов измерений на примере заданной исходной числовой последовательности через оценку числовых моментов и выявление свойств этой последовательности.

Провери анализ заданной числовой последовательности. По первым двум числовым моментам определили, что заданная последовательност, соответствует распределению Эрланга 2-го порядка. Учитывая это, мы синтезировали модель и сгенерировали числовую последовательность.

Провели анализ сгенерированной ЧП и на основе первых двух моментов и автокорреляции определили, что она хорошо моделирует распределению исходной числовой последовательности. Выраженной корреляции между исходной и сгенерированной последовательностями не было обнаружено, что опять же говорит о высоком качестве построенной модели.