

Университет ИТМО

Факультет ПИиКТ

## Компьютерные сети

Лабораторная работа №11

«Обработка результатов измерений: статистический анализ  
числовой последовательности»

Работу выполнили

Юнусов Роман

Лалин Алексей

Группа

P34102

Санкт-Петербург

2024

## 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 гипоекспоненциальный 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..

## Выполнение работы

### Анализ исходной последовательности

Таблица 1 Характеристики исходной числовой последовательности

Характеристика		количество случайных величин					
		10	20	50	100	200	300
Мат. Ож.	знач	18,28949	20,03665	27,03092	28,11718	25,74401	25,90467
	%	29,39696	22,65237	4,347683	8,540949	0,620196	
Дов. Инт(0.9)	знач	±11,376	±11,377	±8,086	±6,799	±4,535	3,792911
	%	199,9291	199,9456	113,1877	79,25709	19,55787	
Дов. Инт(0.95)	знач	±13,571	±13,572	±9,646	±8,111	±5,410	4,524715
	%	199,9291	199,9456	113,1877	79,25709	19,55787	
Дов. Инт(0.99)	знач	±17,836	±17,837	±12,678	±10,660	±7,110	5,946768
	%	199,9291	199,9456	113,1877	79,25709	19,55787	
Дисперсия	знач	479,4103	958,9265	1211,057	1712,47	1523,55	1598,79
	%	70,01419	40,02175	24,25167	7,110353	4,706098	
с.к.о	знач	21,89544	30,96654	34,80024	41,382	39,03268	39,98488
	%	45,2407	22,55437	12,96649	3,494132	2,381404	
К-т вариации	знач	1,19716	1,545495	1,287423	1,471769	1,516185	1,543539
	%	22,44059	0,126691	16,59277	4,64969	1,772199	

Собственно найдем для получившегося ряда характеристики.

Формулы

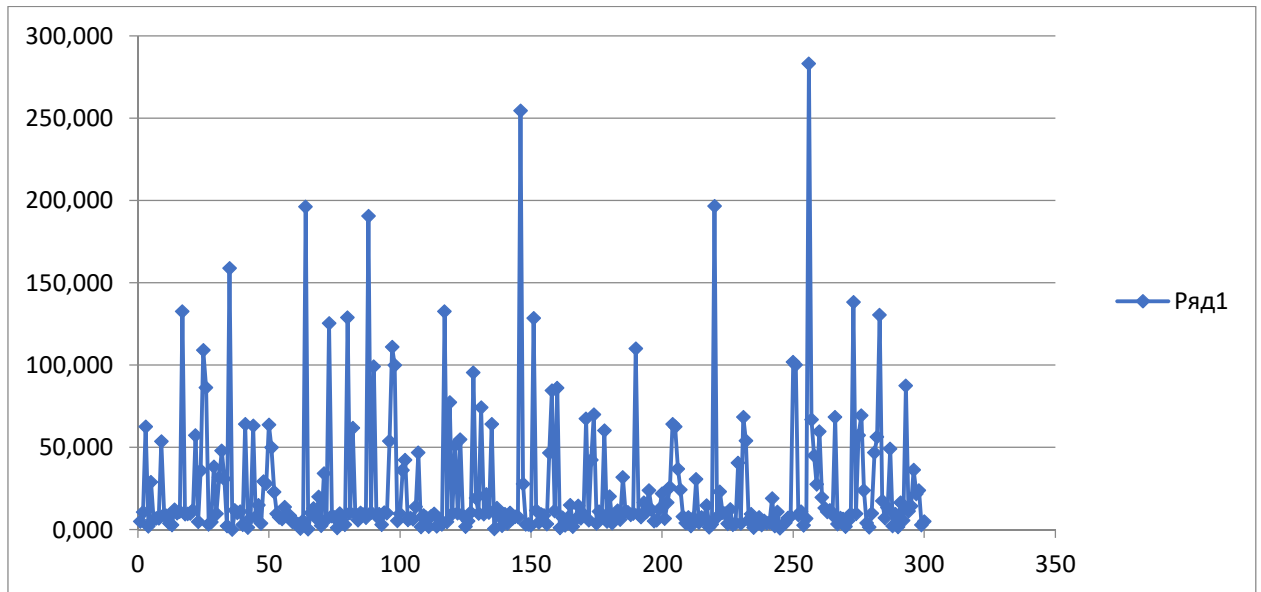
Мат ожидание  $\bar{X} = \frac{\sum x}{n}$

Дисперсия считается несмещенной  $D = \frac{\sum (x_i - \bar{X})^2}{n-1}$ ,  $\sigma$ (ско) – корень из дисперсии

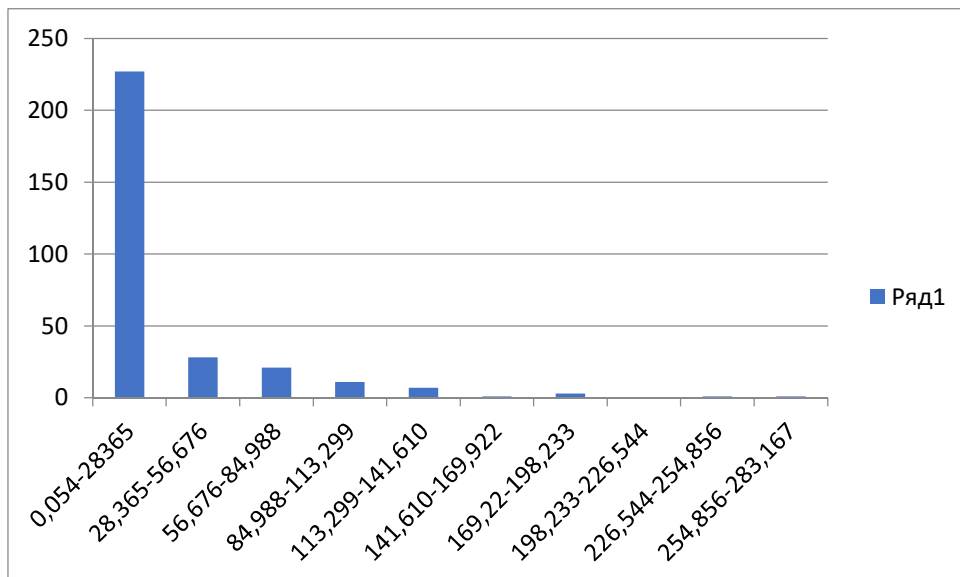
Интервалы считаются как  $= k * \frac{\sigma}{n}$

Коэффициент вариации отвечает за разброс относительно среднего и считается как  $= \frac{\sigma}{\bar{X}}$

## 2 График значений исходной последовательности



## 3 гистограмма исходной последовательности



Как видим подавляющая часть значений не превышает 29.

Сделаем автокорреляционный анализ последовательности с самой собой.

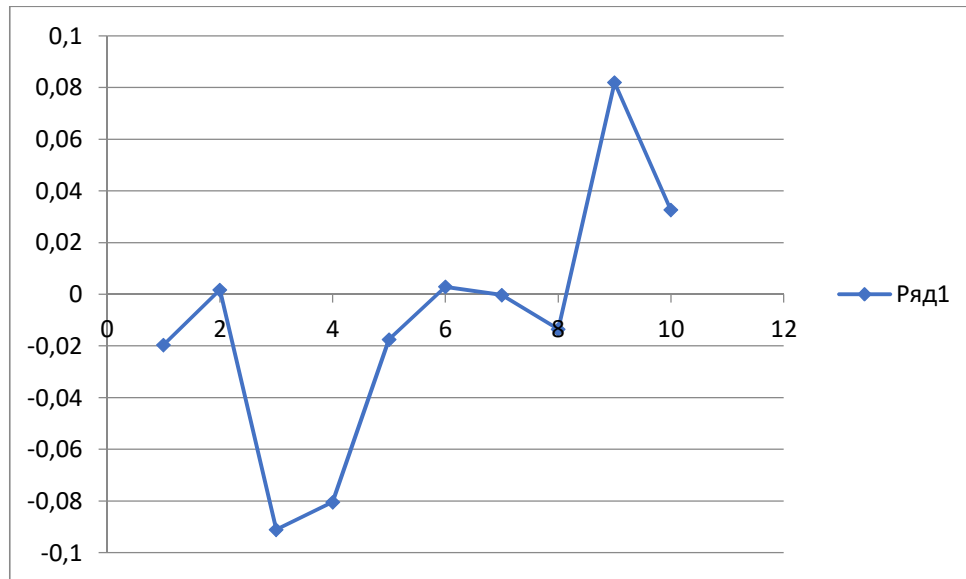
Для этого нужно посчитать значения функции  $\frac{cov}{\sigma^2} = \frac{\sum (x_i - \bar{X}) * (x_{i+k} - \bar{X})}{n * \sigma^2}$ .

Собственно мы находим ковариацию нашего ряда со сдвинутой версией самого себя на 1,2,3...k значений.

Таблица 4 коэффициенты автокорреляции

сдвиг чп	1	2	3	4	5	6	7	8	9	10
	-0,0196	0,001714	-0,09104	-0,08043	-0,01747	0,002887	-0,00027	-0,01353	0,081924	0,032586

5 график коэффициентов автокорреляции



Поскольку значения близки к 0, то корреляции между ними нет, а значит наша величина и правда случайная.

### Генерация случайного ряда

Поскольку коэффициент вариации  $> 1$ , то стоит попытаться вывести нашу случайную функцию через гиперэкспоненту.

$$\left. \begin{aligned} F(x) &= qp(1 - e^{-\alpha_1 x}) + (1 - q)(1 - e^{-\alpha_2 x}); \\ f(x) &= q\alpha_1 e^{-\alpha_1 x} + (1 - q)\alpha_2 e^{-\alpha_2 x}. \end{aligned} \right\}$$

По условию наше  $q$  должно быть меньше чем  $q \leq \frac{2}{1+\sigma^2} = 0.78$ , мы решили взять  $q=0.3$

Дальше из этого  $q$ , мат ожидания  $t$  и СКО исходной функции определяем математические ожидания  $t_1$  и  $t_2$  экспонент по формулам.

$$t_1 = \left[ 1 + \sqrt{\frac{1-q}{2q}} (v^2 - 1) \right] t; \quad t_2 = \left[ 1 - \sqrt{\frac{q}{2(1-q)}} (v^2 - 1) \right] t.$$

Из них выходит  $t_1 = 58.8$ ,  $t_2 = 11.8$ . Но нам прежде всего интересны именно коэффициент  $a$  в экспоненциальной функции, который обратен мат ожиданию функции, значит  $a_1 = 0.017$  и  $a_2 = 0.0847$ .

Теперь поговорим о генерации, будем делать случайное число  $q$ , если  $q$  меньше выбранного 0.3, то будем генерировать по второй экспоненте (с  $t_2$ ), иначе по первой. Очевидно что теперь в зависимости от  $q$  функция распределения ведет себя как экспонента, то давайте решим уравнение.

$$F(x) = U = 1 - e^{-a \cdot x}$$

Очевидно, что если мы возьмем случайное  $U$  в пределах от 0 до 1, то получим случайное  $X$  в соответствии с данной функцией распределения.

Тогда сделаем два ряда из 300 случайных чисел от 0 до 1, первый служит для определения функции экспоненты(q), второй служит для определения случайного U.

$$x = -\frac{1}{a} * \ln(1 - U) = -t * \ln(1 - U)$$

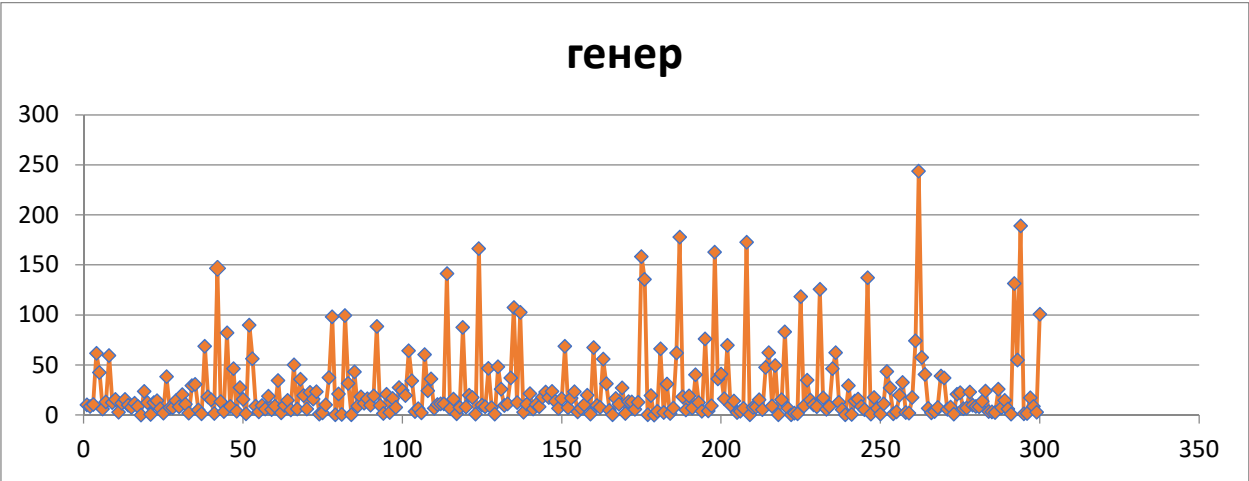
Получили случайный ряд.

Анализ случайного ряда

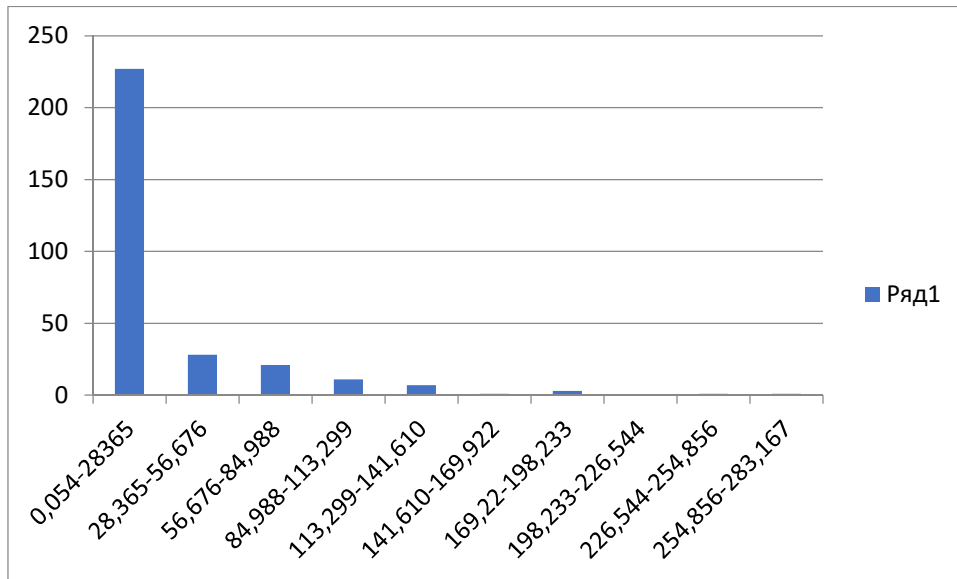
Характеристика		количество случайных величин					
		10	20	50	100	200	300
Мат. Ож.	знач	18,28949	20,03665	27,03092	28,11718	25,74401	25,73057
	%	29,39696	22,65237	4,347683	8,540949	0,620196	
Дов. Инт(0.9)	знач	±11,376	±11,377	±8,086	±6,799	±4,535	3,497886
	%	199,9291	199,9456	113,1877	79,25709	19,55787	
Дов. Инт(0.95)	знач	±13,571	±13,572	±9,646	±8,111	±5,410	4,172767
	%	199,9291	199,9456	113,1877	79,25709	19,55787	
Дов. Инт(0.99)	знач	±17,836	±17,837	±12,678	±10,660	±7,110	5,484208
	%	199,9291	199,9456	113,1877	79,25709	19,55787	
Дисперсия	знач	479,4103	958,9265	1211,057	1712,47	1523,55	1359,745
	%	70,01419	40,02175	24,25167	7,110353	4,706098	
с.к.о	знач	21,89544	30,96654	34,80024	41,382	36,87471	36,87471
	%	45,2407	22,55437	12,96649	3,494132	7,778348	
К-т вариации	знач	1,19716	1,545495	1,287423	1,471769	1,516185	1,433109
	%	22,44059	0,126691	16,59277	4,64969	1,772199	

Как видим, мат ожидание нового ряда и старого ряда различаются не больше чем дов интервал для 0.99, а значит генерация выглядит правильной.

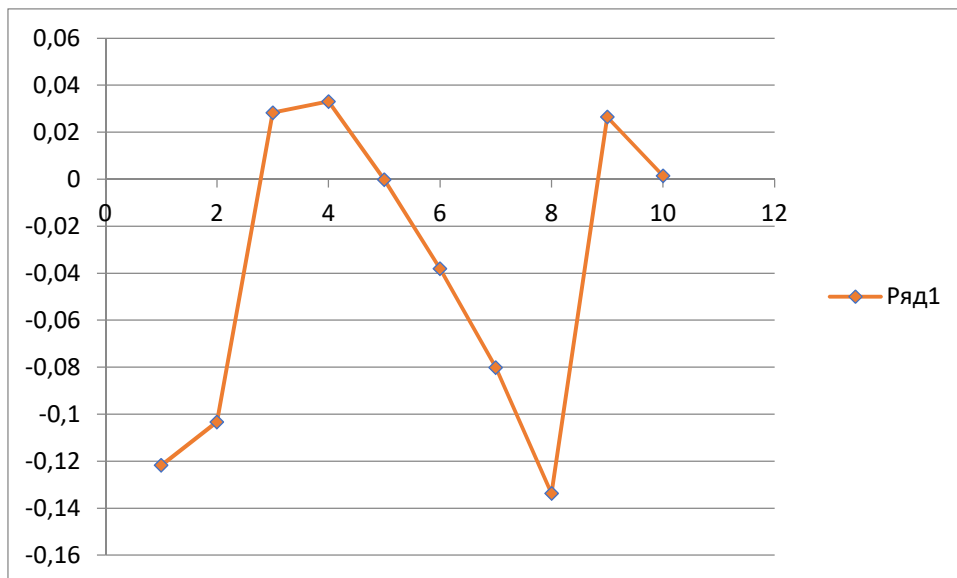
6 график значений сгенерированной последовательности



#### 7 гистограмма сгенерированной последовательности



#### 8 автокорреляционный анализ

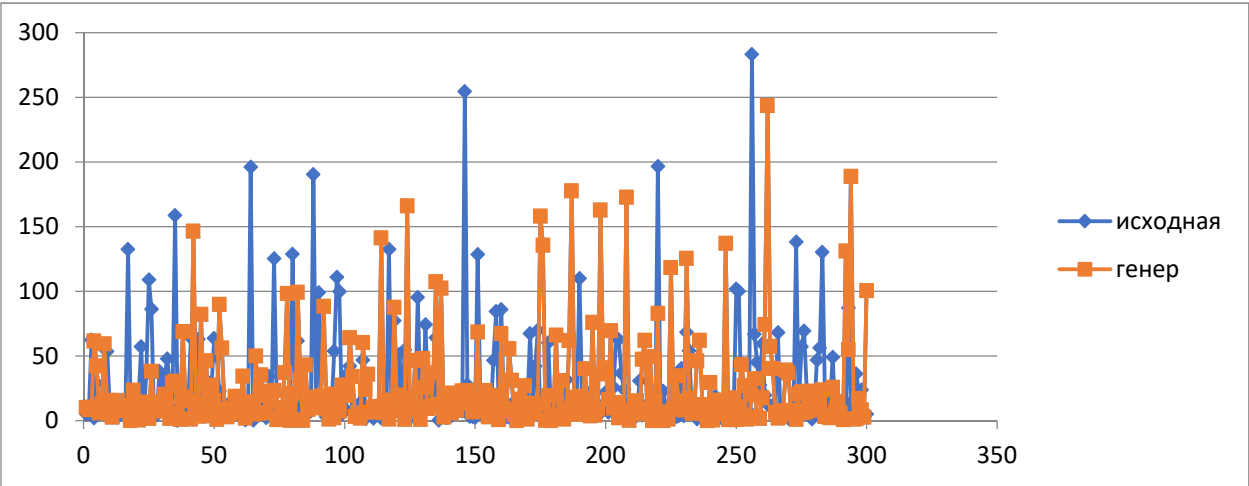


Автокорреляционный анализ возможно немного великоват на некоторых сдвигах, но всё ещё достаточно хаотичен и близок к 0.

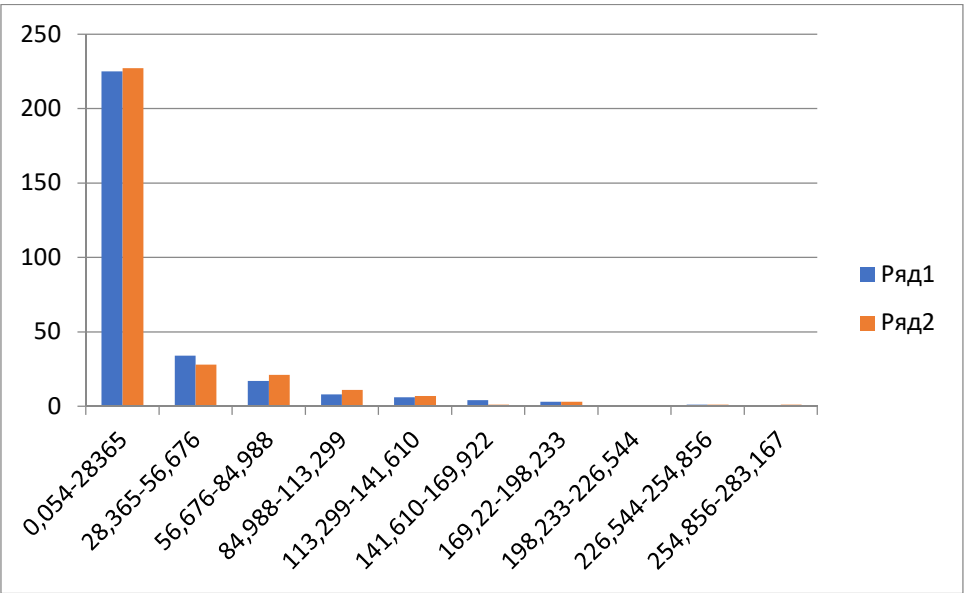


Сравнение

9 график общий

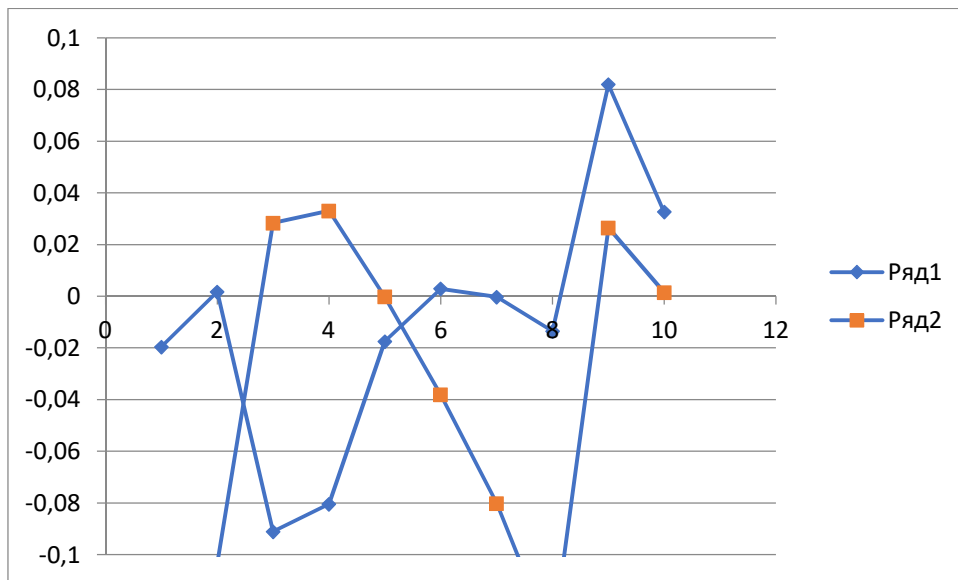


10гистограмма общая



11 автокорреляционный анализ общий

сдвиг чп	1	2	3	4	5	6	7	8	9	10
	-0,0196	0,001714	-0,09104	-0,08043	-0,01747	0,002887	-0,00027	-0,01353	0,081924	0,032586
	-0,12177	-0,10329	0,028318	0,033058	-0,00021	-0,03814	-0,08016	-0,13369	0,026473	0,001467



Как мы можем видеть по гистограмме, последовательности достаточно похожи, что даёт лишний повод утверждать то, что генерация выполнена корректно.

Осталось определить корреляционную зависимость рядов, для этого отсортируем данные в них и проведем корреляционный анализ, формулу которого я уже приводил в автокорреляции.

$$\frac{cov(x, y)}{\sigma_x * \sigma_y}$$

Для сгенерированной последовательности получилось, что ответ равен 0.99, примерно равно 1, а значит, наши последовательности с огромной вероятностью связаны.

## Вывод

В рамках работы были выполнены все цели задания. Мы научились вычислять разнообразные характеристики случайного ряда. После этого мы смогли сгенерировать новую последовательность похожую на исходную и проверить валидность результата генерации.