Министерство науки и высшего образования Российской Федерации

Федеральное государственное бюджетное образовательное учреждение высшего образования

ТОМСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ СИСТЕМ УПРАВЛЕНИЯ И РАДИОЭЛЕКТРОНИКИ (ТУСУР)

Кафедра экономической математики, информатики и статистики (ЭМИС)

СТАТИСТИЧЕСКИЙ АНАЛИЗ ДАННЫХ

Отчет по практической работе по дисциплине «Информационно-аналитические системы управления»

Студент гр. 590-1

\_\_\_\_\_\_\_\_\_\_\_\_/Г.К. Петров

«\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_ 2023 г.

Ассистент кафедры ЭМИС

\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_/ К.В. Никитин

оценка подпись

«\_\_\_» \_\_\_\_\_\_\_\_\_\_\_\_\_ 2023 г.

Томск 2023

Практическая работа №2

«Статистический анализ данных»

**Цель работы:** закрепить навыки получения основных статистических показателей при анализе выборок данных, условно аналогичных эмпирическим, и закрепить базовые навыки статистического анализа на примере пакета Scilab.

**Задание.**

1.1 В Scilab определить квадратичную матрицу M размером 10 и заполнить ее случайными числами по какому-либо из законов распределения (например, по геометрическому, параметр «успеха» p = 0.2);

1.2 Визуализировать полученные данные при помощи построения 3D-графика типа «поверхность». Проанализировать точки экстремума. Найти наибольший и наименьший элемент матрицы в виде значения и индексов элемента (в случае, если равнозначных минимумов/максимумов несколько – указать первый встречающийся);

1.3 Вычислить выборочное среднее числовых данных (среднее всех элементов, среднее по столбцам, среднее по строкам);

1.4 Определить аналогичную по размеру матрицу весов W в виде чередующихся нулей и единиц: Используя полученную матрицу весов, найти средневзвешенное выборочных средних для матрицы M;

1.5 Найти медиану, дисперсию и стандартное отклонение и средневзвешенное стандартное отклонение (в качестве весов использовать матрицу W из пункта 3) для данных из матрицы M;

1.6 Определить абсолютные (пренебрегая интервалами) частоты появления значений в матрице.;

1.7 Используя формулу Стерджесса, сгруппировать данные из матрицы по интервалам и определить частоты появления значений: n = 1 + 3.332 \* lg(N). Количество интервалов округлить в меньшую сторону;

1.8 Упорядочить элементы матрицы по возрастанию;

2 Получить набор случайных чисел, распределённых по Пуассону со средним равным 10. Проверить гипотезу о соответствии экспериментального распределения распределению Пуассону, показать результаты графически при помощи гистограммы. Использовать критерий Пирсона. В ходе решения вычислить среднее и дисперсию;

3 При проведении 25 измерений толщины плоской детали после штамповки было получено значение средней толщины 91.25 мм со средним квадратичным отклонением 0.09 мм. Определить доверительный интервал для истинной толщины детали с 95-типроцентной вероятностью.

**Результат выполнения задания.**

Пример выполнения заданий 1.1-1.8 представлен на рисунках 1-8.

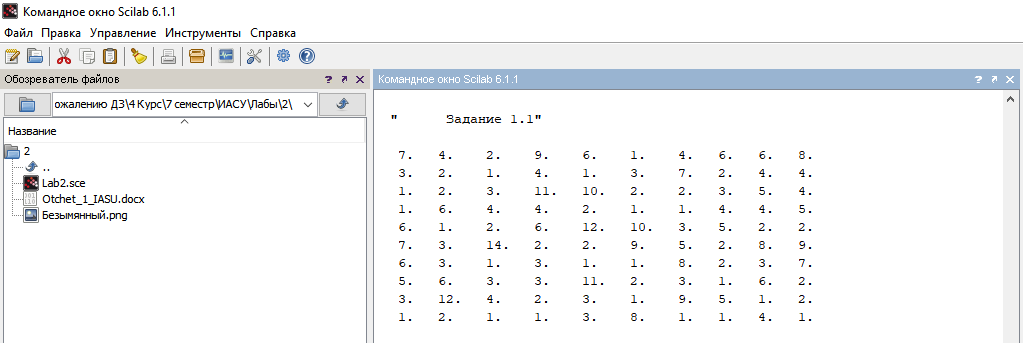


Рисунок 1 – Результат выполнения задания 1.1.

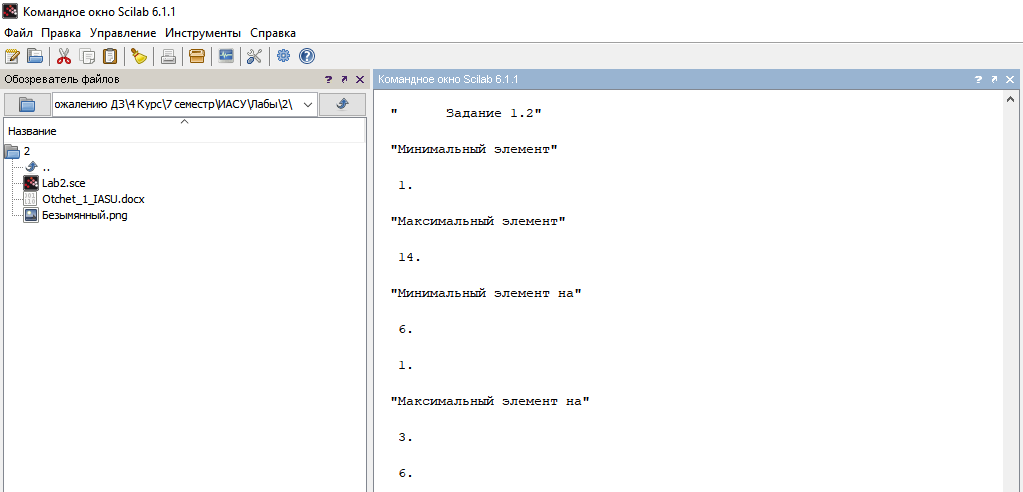


Рисунок 2 – Результат выполнения задания 1.2.

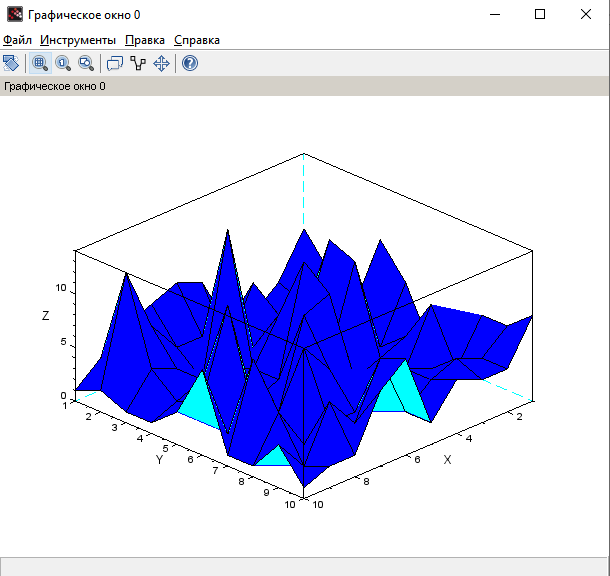


Рисунок 3 – График для задания 1.2.

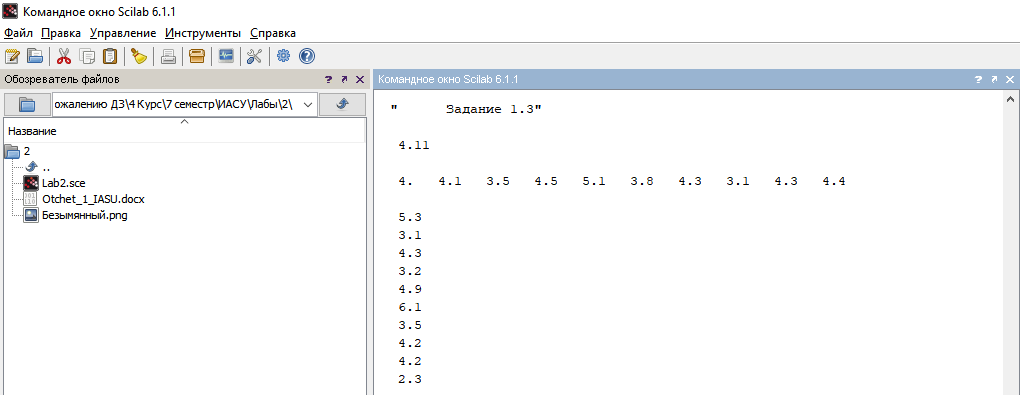


Рисунок 4 – Результат выполнения задания 1.3.

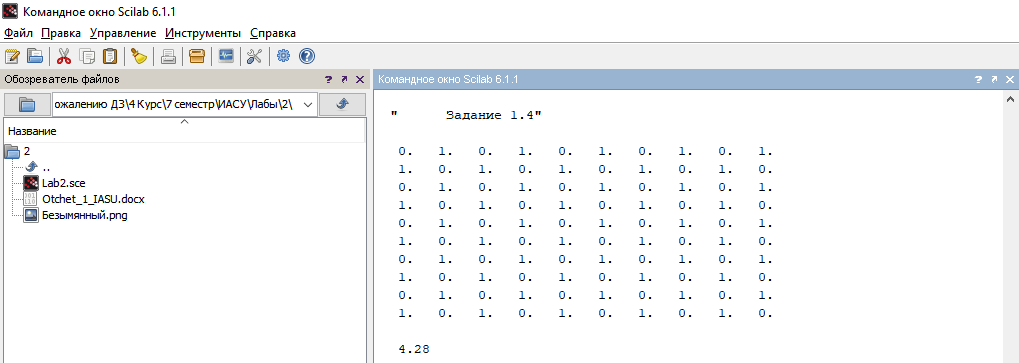


Рисунок 5 – Результат выполнения задания 1.4.

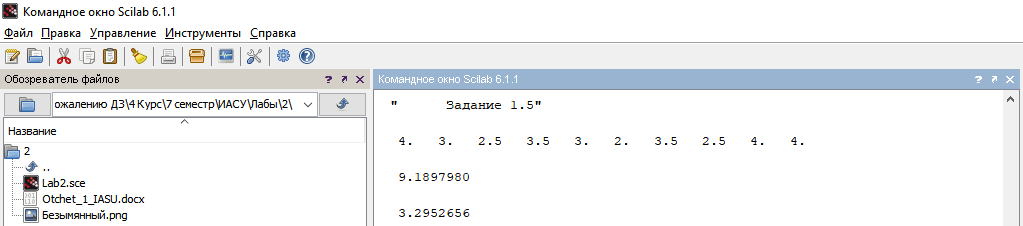


Рисунок 6 – Результат выполнения задания 1.5.

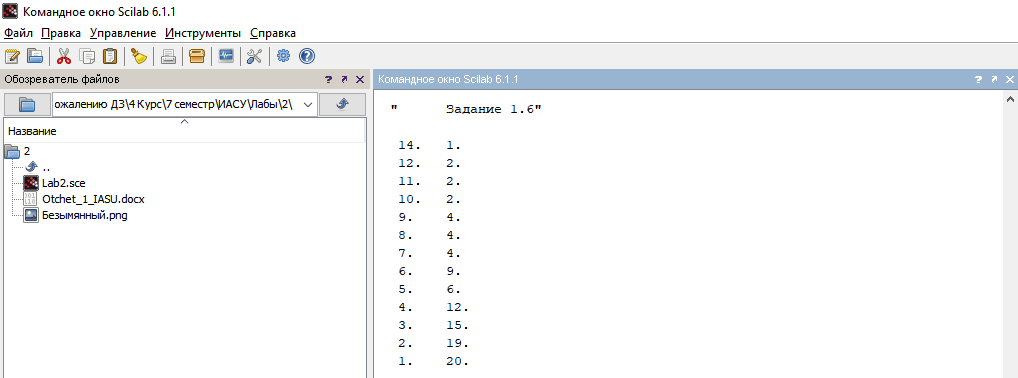


Рисунок 7 – Результат выполнения задания 1.6.

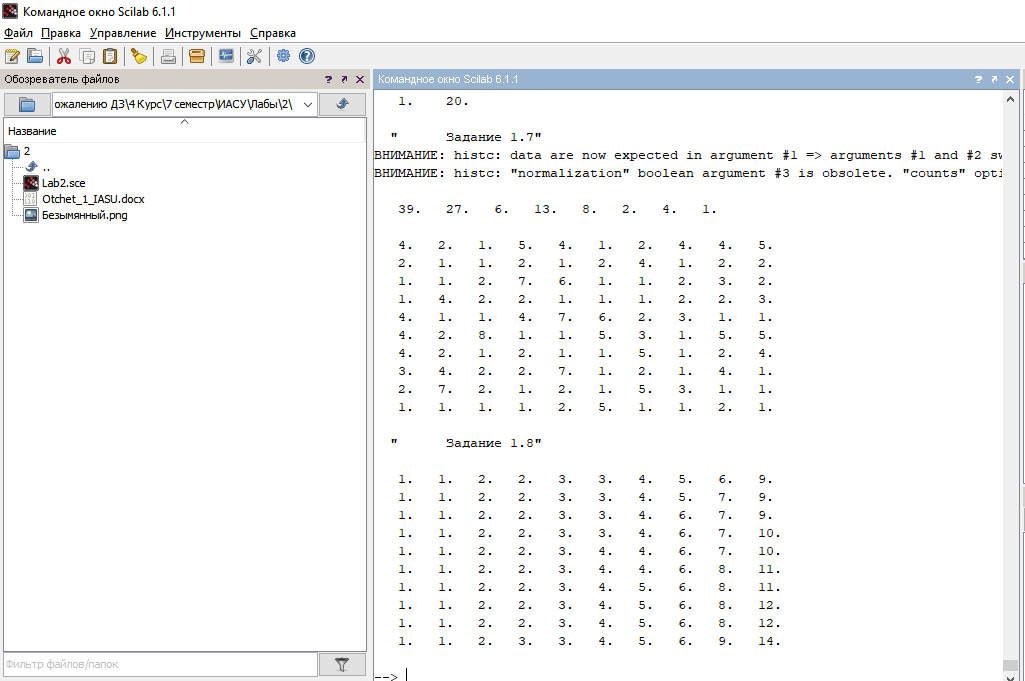


Рисунок 8 – Результат выполнения заданий 1.7 и 1.8.

Пример выполнения задания 2 представлен на рисунке 9.

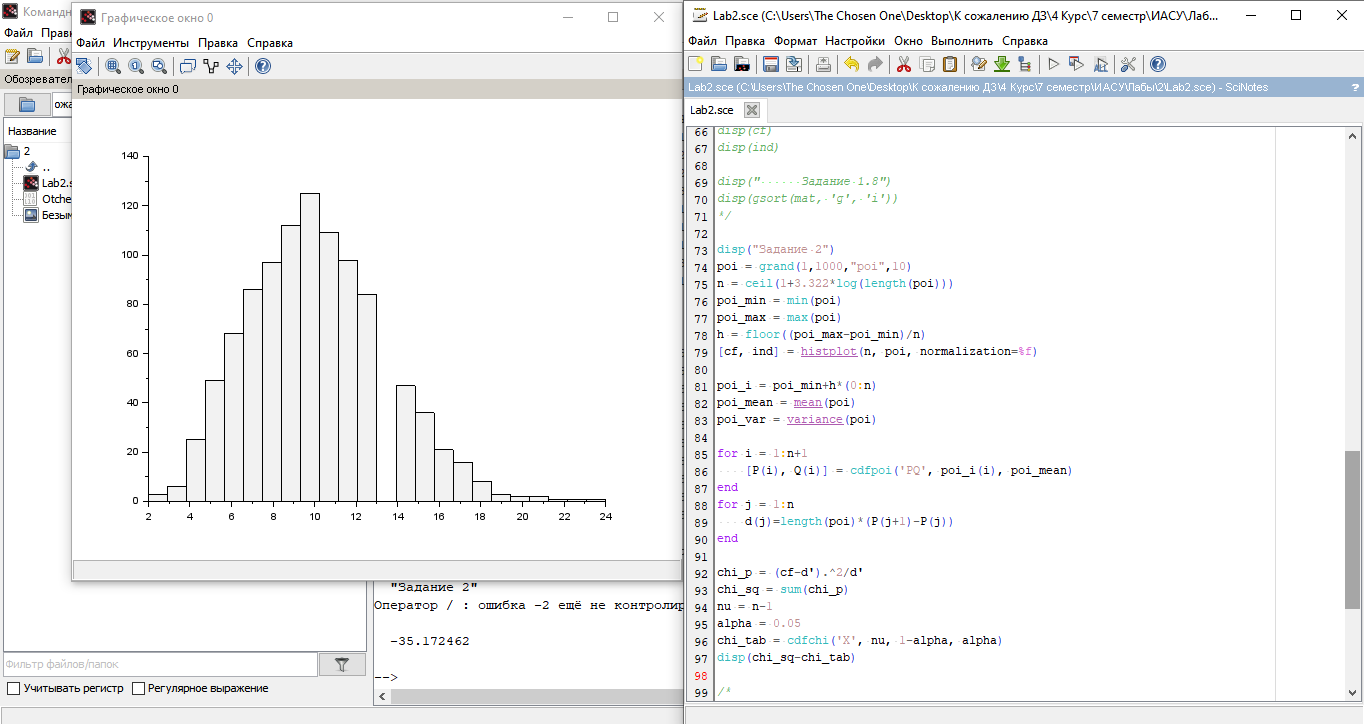


Рисунок 9 – Результат выполнения задания 2.

Пример выполнения задания 3 представлен на рисунке 10.

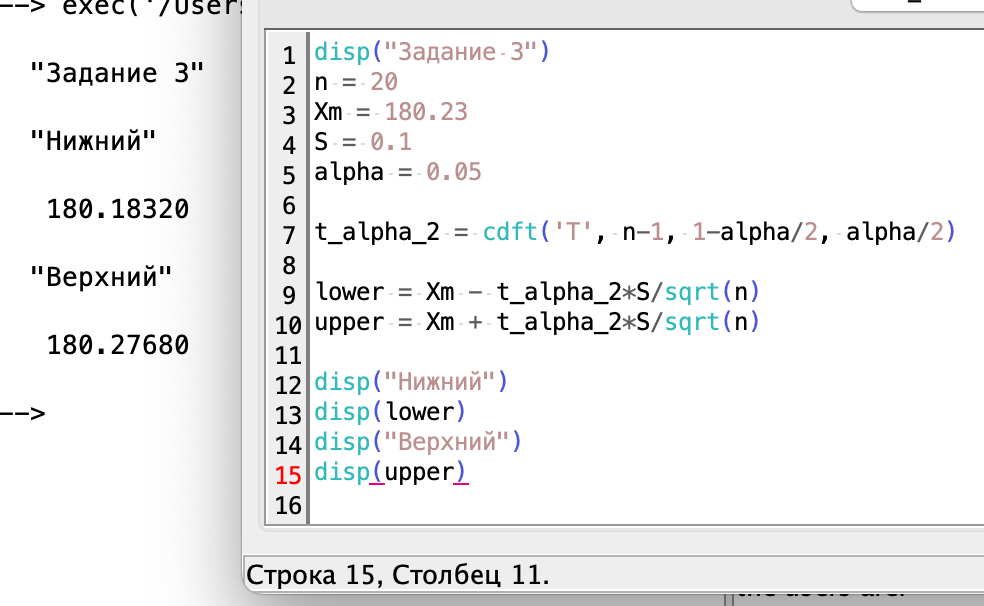


Рисунок 10 – Результат выполнения задания 3 Вариант №7 .

Код, использованный для выполнения работы, представлен в Приложении А.

**Вывод:** в процессе выполнения работы были закреплены навыки получения основных статистических показателей при анализе выборок данных, условно аналогичных эмпирическим, а также базовые навыки статистического анализа на примере пакета Scilab.

**Приложение А**

(обязательное)

Код для выполнения работы

disp(" Задание 1.1")

mat = grand(10, 10, "geom", 0.2)

disp(mat)

disp(" Задание 1.2")

plot3d(mat)

disp("Минимальный элемент")

disp(min(mat))

disp("Максимальный элемент")

disp(max(mat))

*//Я не разобрался в find() и, видимо, другая половина интернета тоже не разобралась*

min\_x=-1

min\_y=-1

max\_x=-1

max\_y=-1

for i=1:10

for j = 1:10

if (mat(i,j)==min(mat) && min\_x==-1)

min\_y=i

min\_x=j

end

if (mat(i,j)==max(mat) && max\_x==-1)

max\_y=i

max\_x=j

end

end

end

disp("Минимальный элемент на")

disp(min\_x)

disp(min\_y)

disp("Максимальный элемент на")

disp(max\_x)

disp(max\_y)

disp(" Задание 1.3")

disp(mean(mat))

disp(mean(mat, 'r'))

disp(mean(mat, 'c'))

disp(" Задание 1.4")

mat2=[]

for i=1:10

for j=1:10

mat2(i,j)=modulo(i+j,2)

end

end

disp(mat2)

disp(meanf(mat, mat2))

disp(" Задание 1.5")

disp(median(mat, 'r'))

disp(variance(mat))

disp(stdevf(mat, mat2))

disp(" Задание 1.6")

disp(tabul(mat))

disp(" Задание 1.7")

n = floor(1+3.322\*log(10))

[cf, ind] = histc(n, mat, normalization=%f)

disp(cf)

disp(ind)

disp(" Задание 1.8")

disp(gsort(mat, 'g', 'i'))

disp("Задание 2")

poi = grand(1,1000,"poi",10)

n = ceil(1+3.322\*log(length(poi)))

poi\_min = min(poi)

poi\_max = max(poi)

h = floor((poi\_max-poi\_min)/n)

[cf, ind] = histplot(n, poi, normalization=%f)

poi\_i = poi\_min+h\*(0:n)

poi\_mean = mean(poi)

poi\_var = variance(poi)

for i = 1:n+1

[P(i), Q(i)] = cdfpoi('PQ', poi\_i(i), poi\_mean)

end

for j = 1:n

d(j)=length(poi)\*(P(j+1)-P(j))

end

chi\_p = (cf-d').^2/d'

chi\_sq = sum(chi\_p)

nu = n-1

alpha = 0.05

chi\_tab = cdfchi('X', nu, 1-alpha, alpha)

disp(chi\_sq-chi\_tab)

disp("Задание 3")

n = 25

Xm = 91.25

S = 0.09

alpha = 0.05

t\_alpha\_2 = cdft('T', n-1, 1-alpha/2, alpha/2)

lower = Xm - t\_alpha\_2\*S/sqrt(n)

upper = Xm + t\_alpha\_2\*S/sqrt(n)

disp("Нижний")

disp(lower)

disp("Верхний")

disp(upper)