ГУАП

КАФЕДРА № 43

ОТЧЕТ   
ЗАЩИЩЕН С ОЦЕНКОЙ

ПРЕПОДАВАТЕЛЬ

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| к.т.н., доцент |  |  |  | В. В. Мышко |
| должность, уч. степень, звание |  | подпись, дата |  | инициалы, фамилия |

|  |
| --- |
| ОТЧЕТ О ЛАБОРАТОРНОЙ РАБОТЕ |
| Выравнивание статистических распределений и проверка гипотез о законах распределения случайных величин |
| по курсу: Обработка экспериментальных данных |
|  |
|  |

РАБОТУ ВЫПОЛНИЛ

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| СТУДЕНТ гр. № | 4131з |  |  |  | М. Д. Быстров |
|  |  |  | подпись, дата |  | инициалы, фамилия |

Санкт-Петербург 2025

**Цель работы**

По заданному интервальному статистическому ряду построить статистическое распределение экспериментальных данных в виде гистограммы, произвести ее выравнивание теоретической плотностью нормального распределения и проверить гипотезу о соответствии статистического и теоретического распределений.

**Задание на лабораторную работу**

1. Найти статистические вероятности попадания значений случайной величины в интервалы по заданному числу попаданий *ml* (табл. 2.1).

2. Построить гистограмму распределения экспериментальных данных.

3. Найти теоретическую плотность нормального распределения в соответствии с методом моментов, полученную кривую нанести на гистограмму распределения.

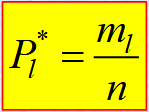
4. Проверить гипотезу о соответствии статистического и теоретического распределений (т. е. гипотезу о нормальном распределении случайной величины) методом К. Пирсона при уровне значимости:

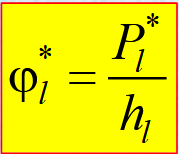
α = 0,025 – для четных вариантов; α = 0,05 – для нечетных вариантов.

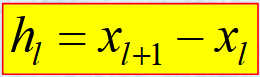
**Таблица 2.1**

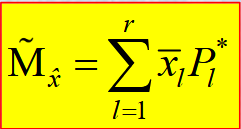
|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| № варианта | | Интервальный статистический ряд | | | | | | |
| 4 | *Jl* | 0; 0,1 | 0,1; 0,2 | 0,2; 0,3 | 0,3; 0,4 | 0,4; 0,5 | 0,5; 0,6 | 0,6; 0,7 |
| *ml* | 3 | 16 | 22 | 27 | 15 | 11 | 6 |

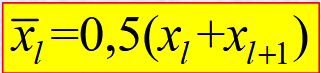
**Математические формулы**

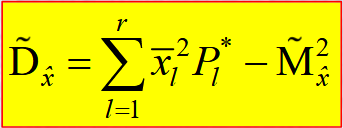
Частота (статистическая вероятность) попадания значений случайной величины в l-й диапазон; n - объем выборки:  (1)

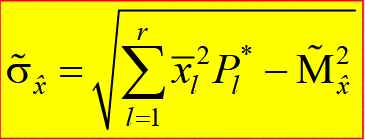
Высота ступенек гистограммы:  (2)

Ширина l-го диапазона:  (3)

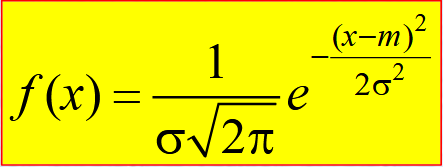
Математическое ожидание интервального стат. ряда:  (4)

Среднее значение диапазона l:  (5)

Дисперсия ИСР:  (6)

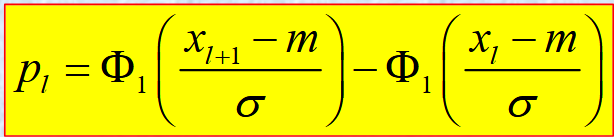
СКО ИСР:  (7)

Плотность распределения вероятностей при нормальном законе распределения:

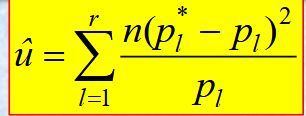
 (8)

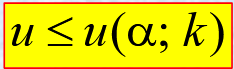
Число степеней свободы:  (9)

Теоретические вероятности нормального распределения:

 (10)

Показатель согласованности гипотезы нормального распределения:

 (11)

Условие подтверждение гипотезы о н.р.:  (12)

**Результат выполнения работы**

1. **Найти статистические вероятности попадания значений случайной величины в интервалы по заданному числу попаданий *ml* (табл. 2.1).**

Для вычисления вероятности попадания используется формула (1).

def P(M):

    """Получить вероятность попадания величины в диапазон"""

    return [m / sum(M) for m in M]

# интервалы

J = [0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7]

# количество попаданий в каждый интервал

M = [3, 16, 22, 27, 15, 11, 6]

# найти статистические вероятности попадания

# значений случайной величины в интервалы

p = P(M)

print(f"Интервалы: {J}")

print(f"Кол-во попаданий: {M}")

print(f"Вероятности: {p}")

Результат:

Интервалы: [0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7]

Кол-во попаданий: [3, 16, 22, 27, 15, 11, 6]

Вероятности: [0.03, 0.16, 0.22, 0.27, 0.15, 0.11, 0.06]

1. **Построить гистограмму распределения экспериментальных данных.**

Для расчета высоты ступенек гистограммы используются формулы (2),(3).

def draw\_hist(J, P\_):

    """Построить график в plt"""

    vals = []

    x\_bar\_positions = []

    # расчет высот столбцов на основе вероятностей

    for i, p in enumerate(P\_):

        vals += [p / (J[i+1] - J[i])]

        x\_bar\_positions += [J[i] + (J[i+1] - J[i]) / 2]

    plt.bar(

        x\_bar\_positions,

        vals,

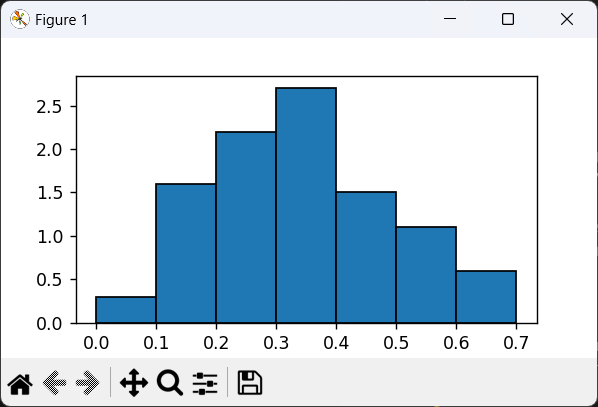
        width=J[1]-J[0], # одинаковая ширина столбцов

        edgecolor = "black")

draw\_hist(J, p)

plt.show()

Результат:



*Рисунок 1 Гистограмма интервального статистического ряда*

1. **Найти теоретическую плотность нормального распределения в соответствии с методом моментов, полученную кривую нанести на гистограмму распределения.**

Использованы формулы (4), (5), (6), (7), (8)

def M(J, P\_):

    """Вычислить мат. ожидание

    Интервального статистического ряда"""

    m = 0

    for i, p in enumerate(P\_):

        m += 0.5 \* (J[i] + J[i+1]) \* p

    return m

def D(J, P\_):

    """Вычислить дисперсию ИСР"""

    m = M(J, P\_)

    d = 0

    for i, p in enumerate(P\_):

        d += ((0.5 \* (J[i] + J[i+1])) \*\* 2) \* p

    d -= m \*\* 2

    return d

def SKO(J, P\_):

    """Вычислить СКО ИСР"""

    return math.sqrt(D(J, P\_))

m = M(J,P\_)

d = D(J, P\_)

sko = SKO(J, P\_)

print(f"Мат. ожидание: {m}")

print(f"Дисперсия: {d}")

print(f"СКО: {sko}")

def f\_approx(x, m, sko):

    """Функция аппроксимирующей кривой

    плотности распределения вероятностей"""

    f = 1 / (sko \* math.sqrt(2 \* math.pi))

    power = - ((x - m) \*\* 2)/(2 \* (sko \*\* 2))

    f = f \* (math.e \*\* power)

    return f

def draw\_normal(m, sko, start, end):

    """Нарисовать выравнивающий график плотности распределения"""

    X = [x for x in np.arange(start, end, 0.01)]

    Y = [f\_approx(x, m, sko) for x in X]

    plt.plot(X, Y, color="red")

draw\_normal(m, sko, -0.1, 0.8)

plt.show()

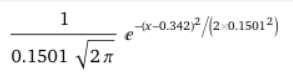
Результат:

Мат. ожидание: 0.34199999999999997

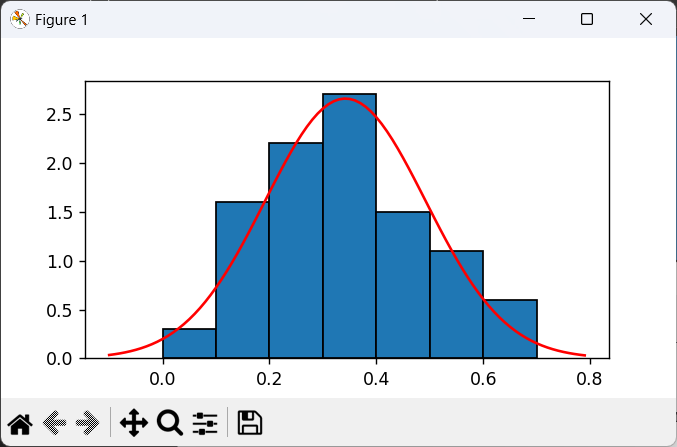
Дисперсия: 0.022536

СКО: 0.15011995203836165

Функция теоретической кривой распределения принимает вид:



Дополненный график изображен на рисунке 2.



*Рисунок 2 Кривая теоретической плотности нормального распределения*

Видно, что теоретическая кривая распределения сохраняет существенные особенности статистического распределения.

1. **Проверить гипотезу о соответствии статистического и теоретического распределений (т. е. гипотезу о нормальном распределении случайной величины) методом К. Пирсона при уровне значимости:**

**α = 0,025 – для четных вариантов; α = 0,05 – для нечетных вариантов.**

Для варианта 4: уровень значимости 0,025.

Поскольку интервал изменения случайной величины разбит на 7 диапазонов (r = 8), число степеней свободы в соответствии с (9) составляет:

По таблице приложения 7 определяется критическая граница:

В расчетах используются формулы (10), (11).

Значения функции Ф1 взяты из приложения 3.

Код расчета:

import pandas as pd

np.set\_printoptions(suppress=True)

def prepare\_f1():

    """Подготовить функцию F1 из приложения"""

    df = pd.read\_excel("./Таблицыприложения.xlsx", sheet\_name="Приложение3")

    values = {}

    for v in zip(df.iloc[:, 1], df.iloc[:, 2]):

        values[v[0]]=v[1]

    def f(x):

        """Функция Ф1"""

        x=round(x, 2)

        if x not in values:

            x=round(x, 1)

        return values[x] if x in values else None

    return f

f1 = prepare\_f1()

def calc\_p(J, m, sko):

    """Вычислить вероятности попадания

    случайной величины в диапазоны

    по нормальному закону"""

    p\_list = []

    for i in range(len(J) - 1):

        p = f1((J[i+1] - m)/sko) - f1((J[i] - m)/sko)

        p\_list += [p]

    return p\_list

p\_n = calc\_p(J, m, sko)

# вывести вероятности из статистического ряда

# и из нормального распределения

print(f"p\*l: {P\_}")

print(f"pl: {np.round(p\_n, 4)}")

# разность между вероятностями

diff = np.array(P\_, np.float64) - np.array(p\_n, np.float64)

print(f"p\*l-pl: {diff}")

# квадрат разностей

powered = np.pow(diff, 2)

print(f"(p\*l-pl)^2: {np.round(powered, 4)}")

# умножение на кол-во измерений

# и деление на теоретическую вероятность

result = (powered / np.array(p\_n, np.float64)) \* N

print(f"n(p\*l-pl)^2/p1: {np.round(result, 4)}")

# сумма

u = sum(result)

print(f"u = {u}")

Результат:

p\*l: [0.03, 0.16, 0.22, 0.27, 0.15, 0.11, 0.06]

pl: [0.043 0.1174 0.2186 0.262 0.2014 0.1042 0.0345]

p\*l-pl: [-0.013 0.0426 0.0014 0.008 -0.0514 0.0058 0.0255]

(p\*l-pl)^2: [0.0002 0.0018 0. 0.0001 0.0026 0. 0.0007]

n(p\*l-pl)^2/pl: [0.393 1.5458 0.0009 0.0244 1.3118 0.0323 1.8848]

u = 5.193003610962049

Вычисленное значение показателя согласованности:

Согласно условию (12), гипотеза о нормальном распределении случайной величины принимается.

**Анализ полученных результатов и выводы**

В ходе выполнения второй лабораторной работы проведено выравнивание статистического распределения и проверка гипотезы о нормальном распределении случайной величины.

Для интервального статистического ряда построена гистограмма распределения, рассчитаны числовые характеристики.

Произведено выравнивание распределения с использованием аппроксимирующей кривой плотности распределения, рассчитанной с использованием числовых характеристик ряда. Подтверждено сходство распределения с нормальным законом.

Проведена проверка гипотезы о нормальном распределении случайной величины методом К.Пирсона с уровнем значимости 0,025. Гипотеза о нормальном распределении подтверждена.

**Список литературы**

1. Статистические методы обработки экспериментальных данных [Текст] : Учеб. пособие / В. И. Сеньченков. - СПб. : ГУАП, 2006 (СПб.). - 243 с. - Библиогр.: с. 227 (12 назв.). - 100 экз. - **ISBN** 5-8088-0213-X
2. Теория вероятностей и математическая статистика : учеб. пособие / Е. А. Трофимова, Н. В. Кисляк, Д. В. Гилёв ; [под общ. ред. Е. А. Трофимовой] ; М-во образования и науки Рос. Федерации, Урал. федер. ун-т. – Екатеринбург : Изд-во Урал. ун-та, 2018. – 160 с. ISBN 978-5-7996-2317-3