**Федеральное государственное автономное образовательное учреждение высшего образования «Национальный исследовательский университет ИТМО»**

**Факультет программной инженерии и компьютерной техники**

**Практическая работа № 6**

**Вариант 2**

Выполнил:

Брагин Роман Андреевич

Группа P3216

Проверил:

Селина Елена Георгиевна

г. Санкт-Петербург

2024 г

*Задание:*

Необходимо определить следующие статистические характеристики: вариационный ряд, экстремальные значения и размах, оценки математического ожидания и среднеквадратического отклонения, эмпирическую функцию распределения и её график, гистограмму и полигон приведенных частот группированной выборки.

*Выборка чисел:*



*Листинг программы:*

*import math*

*import matplotlib.pyplot as plt*

*# выборка*

*data = [0.83, -0.48, -1.35, 0.31, 0.59, 1.35, -0.30, -0.24, 0.51, 0.26,*

*0.73, 0.00, 1.59, 0.17, -0.45, 1.60, -0.18, -1.73, 0.03, 1.70]*

*# алгоритм сортировки*

*def bubble\_sort(arr):*

*n = len(arr)*

*for i in range(n):*

*for j in range(0, n - i - 1):*

*if arr[j] > arr[j + 1]:*

*arr[j], arr[j + 1] = arr[j + 1], arr[j]*

*return arr*

*# Сортировка выборки*

*sorted\_data = bubble\_sort(data.copy())*

*print("Вариационный ряд:", sorted\_data)*

*# максимальный и минимальный элементы*

*max\_value = sorted\_data[-1]*

*min\_value = sorted\_data[0]*

*# 1. Вариационный ряд и статистический ряд*

*list\_of\_frequency = []*

*statistical\_list = []*

*for i in range(len(sorted\_data)):*

*if i == 0 or sorted\_data[i-1] != sorted\_data[i]:*

*statistical\_list.append(sorted\_data[i])*

*list\_of\_frequency.append(1)*

*else:*

*list\_of\_frequency[-1] += 1*

*# 2. Выводим статистический ряд*

*print(f"Статистический ряд: \n Элементы: {statistical\_list}")*

*print(f" Частоты: {list\_of\_frequency}")*

*# 3. Экстремальные значения и размах*

*range\_val = max\_value - min\_value*

*print(f"Минимум: {min\_value}, Максимум: {max\_value}, Размах: {range\_val}")*

*# 4. Математическое ожидание*

*sum\_val = 0*

*for i in range(len(statistical\_list)):*

*sum\_val += statistical\_list[i] \* list\_of\_frequency[i]*

*mean\_val = sum\_val / len(data)*

*print(f"Математическое ожидание: {mean\_val}")*

*# 5. Выборочная дисперсия*

*sum\_squared\_diff = 0*

*for i in range(len(statistical\_list)):*

*sum\_squared\_diff += ((statistical\_list[i] - mean\_val) \*\* 2) \* list\_of\_frequency[i]*

*sample\_variance = sum\_squared\_diff / len(data)*

*print(f"Выборочная дисперсия: {sample\_variance}")*

*# 6. Среднеквадратическое отклонение*

*std\_dev = math.sqrt(sample\_variance)*

*print(f"Выборочное среднеквадратическое отклонение: {std\_dev}")*

*# 7. Исправленная дисперсия*

*corrected\_variance = sample\_variance \* (len(data) / (len(data) - 1))*

*print(f"Исправленная дисперсия: {corrected\_variance}")*

*# 8. Исправленное среднеквадратическое отклонение*

*corrected\_std\_dev = math.sqrt(corrected\_variance)*

*print(f"Исправленное среднеквадратическое отклонение: {corrected\_std\_dev}")*

*# 9. Интервальный ряд (по правилу Стерджеса)*

*bins = int(1 + math.log2(len(sorted\_data))) # Количество интервалов*

*bin\_width = (max\_value - min\_value) / bins*

*# Создание интервалов и подсчёт частот*

*bin\_counts = [0] \* bins*

*interval\_start = min\_value - bin\_width / 2 # Начало интервала от x\_min - половина ширины интервала*

*for value in sorted\_data:*

*index = int((value - interval\_start) / bin\_width)*

*if index >= bins: # Обработка максимального значения*

*index = bins - 1*

*bin\_counts[index] += 1*

*# Формирование интервалов для вывода*

*intervals = []*

*for i in range(bins):*

*left = interval\_start + i \* bin\_width*

*right = left + bin\_width*

*intervals.append(f"[{left:.2f}, {right:.2f})")*

*# Построение гистограммы*

*bin\_centers = [interval\_start + (i + 0.5) \* bin\_width for i in range(bins)]*

*plt.bar(bin\_centers, [count / len(data) for count in bin\_counts], width=bin\_width \* 0.9, label='Гистограмма')*

*plt.xlabel('Значение')*

*plt.ylabel('Частота')*

*plt.title('Гистограмма распределения')*

*plt.legend()*

*plt.show()*

*# Построение полигона частот*

*plt.plot(bin\_centers, [count / len(data) for count in bin\_counts], marker='o', linestyle='-', color='r', label='Полигон частот')*

*plt.xlabel('Значение')*

*plt.ylabel('Частота')*

*plt.title('Полигон частот')*

*plt.legend()*

*plt.grid()*

*plt.show()*

*# 10. Эмпирическая функция распределения*

*cdf = []*

*for i in range(len(sorted\_data)):*

*cdf.append((i + 1) / len(sorted\_data))*

*print(f"\nЭмпирическая функция распределения:")*

*print(f"При x <= {sorted\_data[0]}: 0")*

*for i in range(1, len(sorted\_data)):*

*print(f"При {sorted\_data[i-1]} < x <= {sorted\_data[i]} :", cdf[i])*

*print(f"При x > {sorted\_data[-1]}: 1.0")*

*print(f"\nКоличество интервалов: {bins}")*

*print(f"Длина интервалов: {bin\_width}")*

*# Вывод интервального ряда с частотами и частотностью (в процентах)*

*print("\nИнтервальный ряд:")*

*total\_data\_points = len(data) # Общее количество точек данных*

*for i in range(bins):*

*frequency = bin\_counts[i]*

*percentage = (frequency / total\_data\_points) # Расчёт частотности*

*print(f"Интервал {intervals[i]}: частота = {frequency}, частотность = {percentage:.2f}")*

*# Построение ЭФР*

*plt.step(sorted\_data, cdf, where='post', label='Эмпирическая функция распределения')*

*plt.xlabel('Значение')*

*plt.ylabel('Эмпирическая вероятность')*

*plt.title('Эмпирическая функция распределения (ЭФР)')*

*plt.grid()*

*plt.legend()*

*plt.show()*

*Ответ:*

*Вариационный ряд: [-1.73, -1.35, -0.48, -0.45, -0.3, -0.24, -0.18, 0.0, 0.03, 0.17, 0.26, 0.31, 0.51, 0.59, 0.73, 0.83, 1.35, 1.59, 1.6, 1.7]*

*Статистический ряд:*

*Элементы: [-1.73, -1.35, -0.48, -0.45, -0.3, -0.24, -0.18, 0.0, 0.03, 0.17, 0.26, 0.31, 0.51, 0.59, 0.73, 0.83, 1.35, 1.59, 1.6, 1.7]*

*Частоты: [1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1]*

*Минимум: -1.73, Максимум: 1.7, Размах: 3.4299999999999997*

*Математическое ожидание: 0.24700000000000005*

*Выборочная дисперсия: 0.801611*

*Выборочное среднеквадратическое отклонение: 0.8953273144498608*

*Исправленная дисперсия: 0.8438010526315789*

*Исправленное среднеквадратическое отклонение: 0.9185864426560948*

*Эмпирическая функция распределения:*

*При x <= -1.73: 0*

*При -1.73 < x <= -1.35 : 0.1*

*При -1.35 < x <= -0.48 : 0.15*

*При -0.48 < x <= -0.45 : 0.2*

*При -0.45 < x <= -0.3 : 0.25*

*При -0.3 < x <= -0.24 : 0.3*

*При -0.24 < x <= -0.18 : 0.35*

*При -0.18 < x <= 0.0 : 0.4*

*При 0.0 < x <= 0.03 : 0.45*

*При 0.03 < x <= 0.17 : 0.5*

*При 0.17 < x <= 0.26 : 0.55*

*При 0.26 < x <= 0.31 : 0.6*

*При 0.31 < x <= 0.51 : 0.65*

*При 0.51 < x <= 0.59 : 0.7*

*При 0.59 < x <= 0.73 : 0.75*

*При 0.73 < x <= 0.83 : 0.8*

*При 0.83 < x <= 1.35 : 0.85*

*При 1.35 < x <= 1.59 : 0.9*

*При 1.59 < x <= 1.6 : 0.95*

*При 1.6 < x <= 1.7 : 1.0*

*При x > 1.7: 1.0*

*Количество интервалов: 6*

*Длина интервалов: 0.6445032587522389*

*Интервальный ряд:*

*Интервал [-2.05, -1.41): частота = 1, частотность = 0.05*

*Интервал [-1.41, -0.76): частота = 1, частотность = 0.05*

*Интервал [-0.76, -0.12): частота = 5, частотность = 0.25*

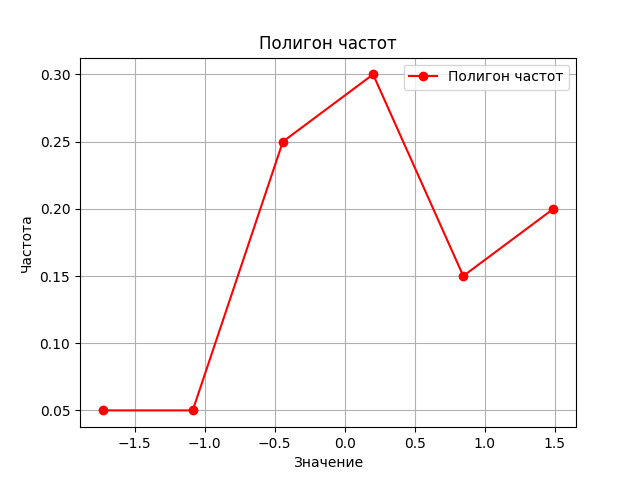
*Интервал [-0.12, 0.53): частота = 6, частотность = 0.30*

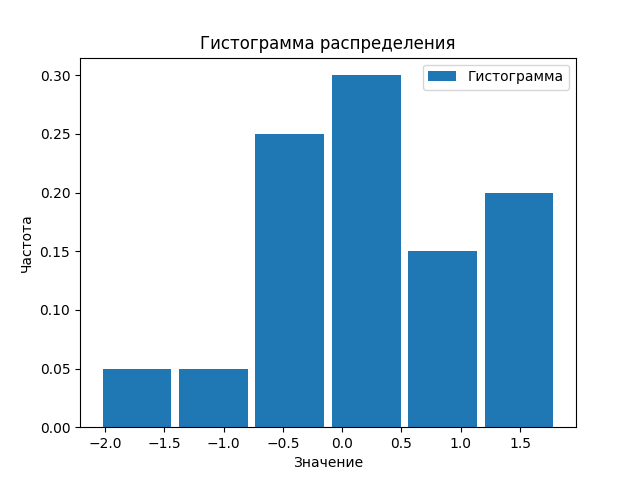
*Интервал [0.53, 1.17): частота = 3, частотность = 0.15*

*Интервал [1.17, 1.81): частота = 4, частотность = 0.20*

**Картинки:**

**

**

**

**Решение 2 и 3 задач:**