Университет ИТМО

Теория вероятностей

Практическая работа №5

Статистическое исследование выборки

Вариант 3

ФИО студента: Готовко Алексей Владимирович
Направление подготовки: 09.03.04 (СППО)
Учебная группа: P32101
ФИО преподавателя: Селина Е.Г.

Санкт-Петербург

2023г.

1 Задание 1

1.1 Исходные данные

Выборка:

-0.03	0.73	-0.59	-1.59	0.38	1.49	0.14	-0.62	-1.59	1.45	-0.38	-1.49	-0.15	0.63	0.06	-1.59	0.61	0.62	-0.05	1.56

2 Исходный код

```python
from sympy import symbols, Piecewise, plot
from math import log10
from matplotlib import pyplot as plt


def plot_show(intervals):
    x = symbols('x')
    f = Piecewise(
        *[(i[2], (i[0] < x) & (x <= i[1])) for i in intervals]
    )
    plot(f)


def hist_show(x_hist, y, h):
    plt.bar(x_hist, y, width=h)
    plt.xlabel("Value")
    plt.ylabel("Density frequency")
    plt.show()


def polygon_show(x_polygon, y):
    plt.plot(x_polygon, y)
    plt.show()


def print_results(data, series, data_extreme_values,
                  data_range, series_expected_value,
                  series_standard_deviation):
    print(f'\nData:\n  |\t{", ".join([str(i) for i in data])}\n')

    print("    " + '_' * 15 + ' ' + "\n  | {:^5s} | {:^5s} |"
          .format("X", "P"))
    print('  |' + '_' * 15 + '|')
    for k in series:
        print("  | {:5.2f} | {:5.2f} |"
              .format(k, series[k]))
    print("  |" + '_' * 15 + "|\n")

    print("Extreme values:\n  |\tmin: {}\n  |\tmax: {}\n"
          .format(data_extreme_values[0], data_extreme_values[1]))

    print("Data range:\n  |\t{:5.2f}\n"
          .format(data_range))

    print("Expected value:\n  |\t{:5.2f}\n"
          .format(series_expected_value))

    print("Standard deviation:\n  |\t{:5.2f}\n"
          .format(series_standard_deviation))

```

```python
52  def print_function(intervals):
53      print("F(x) =\n    ___\n   |")
54      for i in intervals:
55          print("   |\t{:5.2f}\tIF\t{:5.2f} < x <= {:4.2f}".format(i[2], i[0], i[1]))
56      print("   |___\n")


59  def main():
60      data = [-0.03, 0.73, -0.59, -1.59, 0.38,
61              1.49, 0.14, -0.62, -1.59, 1.45,
62              -0.38, -1.49, -0.15, 0.63, 0.06,
63              -1.59, 0.61, 0.62, -0.05, 1.56]
64      size = len(data)

66      data.sort()

68      data_set = sorted(set(data))

70      data_map = {}
71      series = {}
72      series_expected_value = 0
73      series_squared_expected_value = 0
74      for n in data_set:
75          data_map[n] = data.count(n)
76          series[n] = data_map[n] / size

78          series_expected_value += n * series[n]
79          series_squared_expected_value += n ** 2 * series[n]

81      data_extreme_values = [data[0], data[-1]]
82      data_range = data_extreme_values[1] - data_extreme_values[0]

84      series_variance = series_squared_expected_value - series_expected_value ** 2

86      series_standard_deviation = series_variance ** 0.5

88      intervals = [[float("-inf"), data_set[0], 0]]
89      for i in range(len(data_set) - 1):
90          p = series[data_set[i]] + intervals[i][2]

92          intervals.append([data_set[i], data_set[i + 1], p])

94      intervals.append([data_set[-1], float("+inf"), 1])

96      print_results(data, series, data_extreme_values,
97                    data_range, series_expected_value,
98                    series_standard_deviation)
99      print_function(intervals)
100     plot_show(intervals)

102     # ----------------------------------------------------------------------------------

104     k = int(1 + 3.322 * log10(size))
105     h = data_range / k
106     offset = data_extreme_values[0]
107     hist_intervals = []
108     for i in range(k):
109         hist_intervals.append([])

111     int_idx = 1
112     ceil = offset + int_idx * h
```

```python
        for i in data:
            if i >= ceil:
                int_idx += 1
                ceil += h
            hist_intervals[int_idx - 1].append(i)

    hist_data = []
    for i, e in enumerate(hist_intervals):
        hist_data.append(
            {
                "interval": e,
                "density_frequency": len(e) / size / h,
                "floor": offset + i * h,
                "pivot": offset + i * h + h / 2
            }
        )

    x_hist = [i["floor"] for i in hist_data]
    x_polygon = [i["pivot"] for i in hist_data]
    y = [i["density_frequency"] for i in hist_data]

    hist_show(x_hist, y, h)
    polygon_show(x_polygon, y)


if __name__ == "__main__":
    main()
```

3   Результат работы программы

```
Data:
|   -1.59, -1.59, -1.59, -1.49, -0.62, -0.59, -0.38, -0.15, -0.05, -0.03,
|    0.06,  0.14,  0.38,  0.61,  0.62,  0.63,  0.73,  1.45,  1.49,  1.56

    _____
  |   X   |   P   |
  |_____|
  | -1.59 |  0.15 |
  | -1.49 |  0.05 |
  | -0.62 |  0.05 |
  | -0.59 |  0.05 |
  | -0.38 |  0.05 |
  | -0.15 |  0.05 |
  | -0.05 |  0.05 |
  | -0.03 |  0.05 |
  |  0.06 |  0.05 |
  |  0.14 |  0.05 |
  |  0.38 |  0.05 |
  |  0.61 |  0.05 |
  |  0.62 |  0.05 |
  |  0.63 |  0.05 |
  |  0.73 |  0.05 |
  |  1.45 |  0.05 |
  |  1.49 |  0.05 |
  |  1.56 |  0.05 |
  |_____|
```

```
 29  Extreme values:
 30  |          min: -1.59
 31  |          max:  1.56
 32
 33  Data range:
 34  |          3.15
 35
 36  Expected value:
 37  |          -0.02
 38
 39  Standard deviation:
 40  |          0.98
 41
 42  F(x) =
 43     ___
 44  |
 45  |          0.00       IF       -inf < x <= -1.59
 46  |          0.15       IF      -1.59 < x <= -1.49
 47  |          0.20       IF      -1.49 < x <= -0.62
 48  |          0.25       IF      -0.62 < x <= -0.59
 49  |          0.30       IF      -0.59 < x <= -0.38
 50  |          0.35       IF      -0.38 < x <= -0.15
 51  |          0.40       IF      -0.15 < x <= -0.05
 52  |          0.45       IF      -0.05 < x <= -0.03
 53  |          0.50       IF      -0.03 < x <= 0.06
 54  |          0.55       IF       0.06 < x <= 0.14
 55  |          0.60       IF       0.14 < x <= 0.38
 56  |          0.65       IF       0.38 < x <= 0.61
 57  |          0.70       IF       0.61 < x <= 0.62
 58  |          0.75       IF       0.62 < x <= 0.63
 59  |          0.80       IF       0.63 < x <= 0.73
 60  |          0.85       IF       0.73 < x <= 1.45
 61  |          0.90       IF       1.45 < x <= 1.49
 62  |          0.95       IF       1.49 < x <= 1.56
 63  |          1.00       IF       1.56 < x <=  inf
 64  |___
 65
```