

```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
import plotly.express as px
from IPython.display import display, Math, Latex
from utils import Answer
```

Лабораторная работа №1

Задание 2

{Фамилия Имя}, {Номер группы}, Вариант {Номер варианта}, ({Дата})

Задание

По данным выборки ($n = 150$) требуется:

1. Составить вариационный ряд и список вариантов.
2. Составить ряд распределения частот вариационного ряда и построить полигон частот.
3. Составить ряд распределения относительных частот вариационного ряда и построить полигон относительных частот.
4. Составить эмпирическую функцию распределения и построить график эмпирической функции распределения.
5. Найти основные числовые характеристики вариационного ряда:
 - Выборочное среднее - \bar{x}_e
 - Выборочную дисперсию - S^2
 - Стандартное отклонение (Среднеквадратическое отклонение) - S
 - Коэффициент вариации - CV
6. Пояснить смысл полученных результатов

Данные

```
LB_NUM = 1
LB_SURNAME = 'Фамилия'
LB_NAME = 'Имя'
LB_VARIANT = 59
LB_EXERCISE_NUM = 2
print(f'Фамилия, Имя: {LB_SURNAME}, {LB_NAME}\nНомер лабораторной работы:
```


3, 3, 3, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4, 4,
4, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 6, 6, 6, 6, 6, 6, 6,
6]

Варианты, x_i

```
unique = sorted(data.unique())  
print(unique)
```

```
[0, 1, 2, 3, 4, 5, 6]
```

Пункт 2

Составить ряд распределения частот вариационного ряда и построить полигон частот.

Частота, n_i

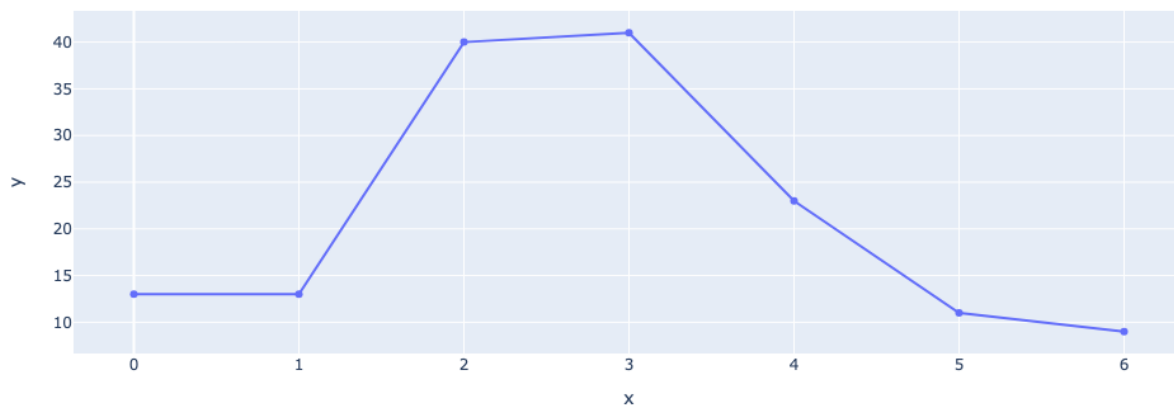
```
freq = pd.Series(data_sort).value_counts().sort_index()
freq = pd.DataFrame(freq, index=unique)
freq.columns=["Частота"]
freq = freq.T
freq
```

	0	1	2	3	4	5	6
Частота	13	13	40	41	23	11	9

Полигон частот вариационного ряда

```
fig = px.line(x=unique, y=freq.iloc[0], markers=True, title='Полигон  
частот вариационного ряда')  
fig.show()
```

Полигон частот вариационного ряда



```
def plot_freq_polygon(x, y, title):  
    sns.set_style("darkgrid")  
    fig, ax = plt.subplots(figsize=(10, 5))  
    ax.set_title(title)  
    ax.plot(x, y, marker='o')  
    ax.set_xlabel("x")  
    ax.set_ylabel("n")  
    plt.plot()
```

```
plot_freq_polygon(x=unique, y=freq.iloc[0], title='Полигон частот  
вариационного ряда')
```



Пункт 3

Составить ряд распределения относительных частот вариационного ряда и построить полигон относительных частот.

Относительная частота, w_i

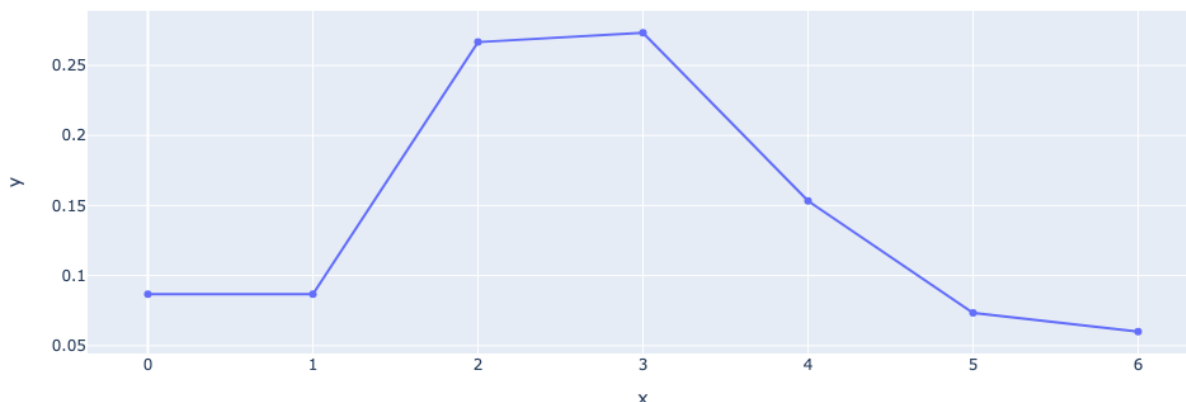
```
rel_freq = pd.Series(data_sort).value_counts().sort_index() /  
len(data_sort)  
rel_freq = pd.DataFrame(rel_freq, index=unique)  
rel_freq.columns=["Относительная частота"]  
rel_freq = rel_freq.T  
rel_freq
```

	0	1	2	3	4	5	6
Относительная частота	0.086667	0.086667	0.266667	0.273333	0.153333	0.073333	0.06

Полигон относительных частот вариационного ряда

```
fig = px.line(x=unique, y=rel_freq.iloc[0], markers=True, title='Полигон  
относительных частот вариационного ряда')  
fig.show()
```

Полигон относительных частот вариационного ряда



```
def plot_rel_freq_polygon(x, y, title):  
    sns.set_style("darkgrid")  
    fig, ax = plt.subplots(figsize=(10, 5))  
    ax.set_title(title)  
    ax.plot(x, y, marker='o')  
    ax.set_xlabel("x")  
    ax.set_ylabel("w")  
    plt.plot()
```

```
plot_rel_freq_polygon(x=unique, y=rel_freq.iloc[0], title='Полигон частот  
вариационного ряда')
```



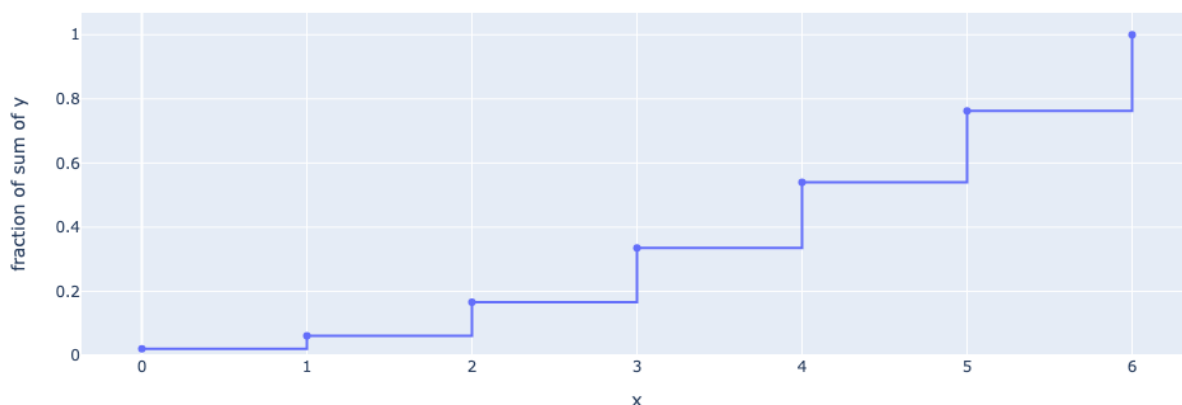
Пункт 4

Составить эмпирическую функцию распределения и построить график эмпирической функции распределения.

Эмпирическая функция распределения F^*

```
emp_func = rel_freq.iloc[0].cumsum()
emp_func.name = "F*"
emp_func = pd.DataFrame(emp_func, index=unique).T
emp_func
```

График эмпирической функции распределения



	0	1	2	3	4	5	6
F*	0.086667	0.173333	0.44	0.713333	0.866667	0.94	1.0

```
def emp_func_latex(emp_func):
    latex_text = f'$$\nF^*(x) = \n\begin{{cases}}\n'
    for i in range(len(emp_func.values.tolist()[0])):
        if i == 0:
            latex_text += f'0, && x \le {unique[0]:.2f} \\\\'
        elif i == len(emp_func.values.tolist()[0]) - 1:
            latex_text += f'{emp_func.values.tolist()[0][i - 1]:.2f}, && \\unique[i - 1]:.2f} < x \le {unique[i]:.2f} \\\\'
            latex_text += f'1.0, && x > {unique[i]:.2f} \\\\'
        else:
            latex_text += f'{emp_func.values.tolist()[0][i - 1]:.2f}, && \\unique[i - 1]:.2f} < x \le {unique[i]:.2f} \\\\'
            latex_text += "\\n"
    latex_text += '\\end{{cases}}\n$$'
    return Math(latex_text)

display(emp_func_latex(emp_func))
```

$$F^*(x) = \begin{cases} 0, & x \leq 0.00 \\ 0.09, & 0.00 < x \leq 1.00 \\ 0.17, & 1.00 < x \leq 2.00 \\ 0.44, & 2.00 < x \leq 3.00 \\ 0.71, & 3.00 < x \leq 4.00 \\ 0.87, & 4.00 < x \leq 5.00 \\ 0.94, & 5.00 < x \leq 6.00 \\ 1.0, & x > 6.00 \end{cases}$$

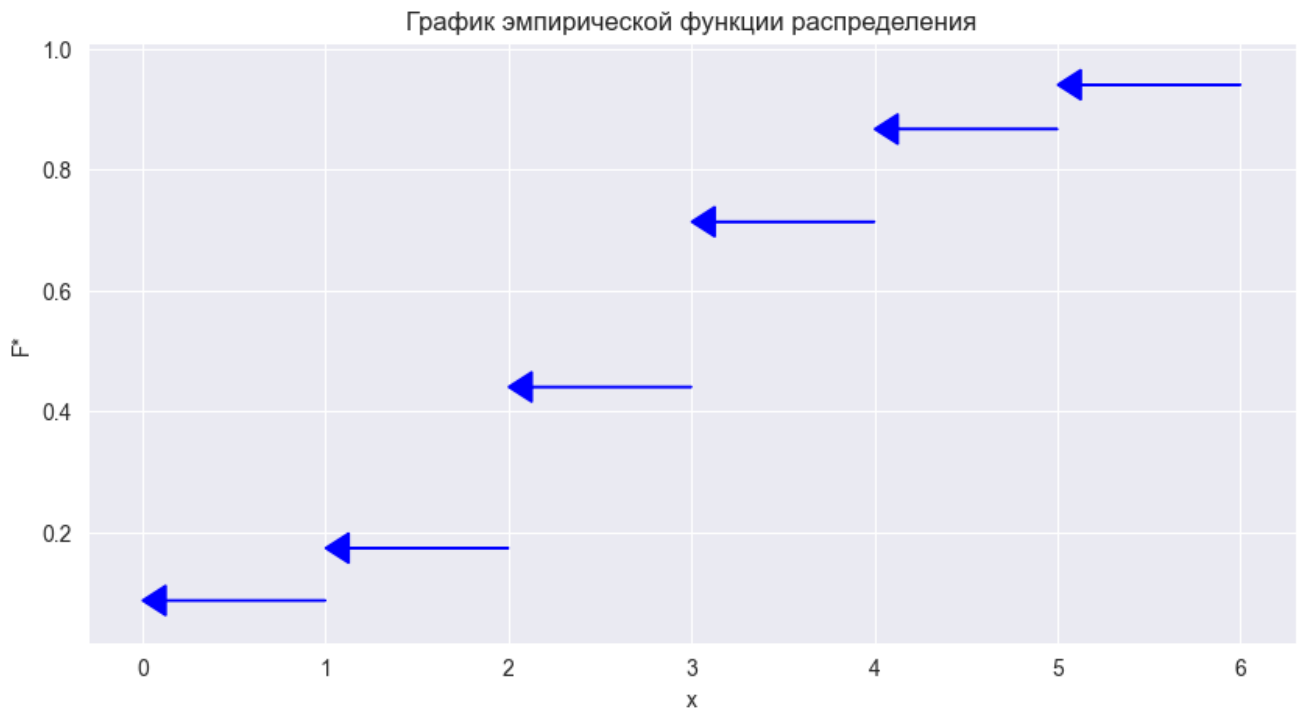
График эмпирической функции распределения

```
fig = px.ecdf(x=unique, y=emp_func.values.tolist()[0], markers=True,
title='График эмпирической функции распределения')
fig.show()
```

```
def plot_cdf_func(x, y, title):
    fig, ax = plt.subplots(figsize=(10, 5))
    min_y, max_y = min(y), max(y)
    arrow_length = ((max(x) - min(x)) / len(x)) * .15
    for idx in range(len(y) - 1):
        dx = x[idx] - x[idx + 1]
        ax.arrow(x=x[idx + 1], y=y[idx], dx=dx, dy=0,
            color="blue", head_width=.05, head_length=arrow_length,
            length_includes_head=True)
    ax.set_title(title)
```

```
ax.set_ylabel("F*")
ax.set_xlabel("x")
plt.plot()
```

```
plot_cdf_func(unique, emp_func.iloc[0].tolist(), 'График эмпирической функции распределения')
```



Пункт 5

Найти основные числовые характеристики вариационного ряда

Выборочное среднее, \bar{x}

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

```
mean = np.mean(data)
print(f'Выборочное среднее: {mean:.2f}')
```

Выборочное среднее: 2.78

Выборочная дисперсия, s_x^2

$$s_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$


```
var = np.var(data, ddof=1)
print(f'Выборочная дисперсия: {var:.2f}')
```

Выборочная дисперсия: 2.35

Стандартное отклонение (Среднеквадратическое отклонение), s_x

$$s_x = \sqrt{s^2}$$

```
std = np.std(data, ddof=1)
print(f'Среднеквадратическое отклонение: {std:.2f}')
```

Среднеквадратическое отклонение: 1.53

Коэффициент вариации, CV

$$CV = \frac{s}{\bar{x}} \cdot 100$$

```
cv = (std / mean) * 100
cv
```

55.11054713812624

Пункт 6

Пояснить смысл полученных результатов.

$$\bar{x} = 2.77$$

$$s_x^2 = 2.28$$

$$s = 1.51$$

$$CV = 54.49$$