

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ “ЛЬВІВСЬКА ПОЛІТЕХНІКА”
Інститут прикладної математики та фундаментальних наук

Кафедра прикладної математики

Звіт

Про виконання лабораторної роботи № 2
з курсу “Математична Статистика”

Виконав:
Студент групи ПМ-33
Маркевич Леонід
Перевірив:
Янішевський В.С.

Львів-2023

Лабораторна робота №2

Варіант 12

Завдання:

- 1) знайти вибірки для заданих випадкових величин X і Y;
- 2) побудувати кореляційне поле в координатах (X, Y);
- 3) знайти вибіркові оцінки характеристик розподілів;
- 4) перевірити гіпотезу про значущість вибіркового коефіцієнта кореляції;
- 5) навести висновки до роботи;

$$x_i = \left(\sum_{j=1}^{12} r_j - 6 \right) \sigma_X + a_X, \quad y_i = \left(\sum_{j=1}^{12} r_j - 6 \right) \sigma_Y + a_Y \quad (1)$$

де $a_x = 12 - 10 = 2$, $\sigma_x = 3 + 12/10 = 4.2$;

$a_y = 12 + 5 = 17$, $\sigma_y = 5 + 12/5 = 7.4$;

Хід роботи:

Реалізація виконана на мові програмування Python з використанням бібліотек pandas, numpy, math, matplotlib, scipy

- 1) знайти вибірки для заданих випадкових величин X і Y;

Реалізація формули програмно:

```
V = 12

n = 350
ax = V - 10
ay = V + 5
sigma_x = 3 + V / 10
sigma_y = 5 + V / 5
sum_limit = 12

def generation_x(self):
    r = sum(random.uniform(0, 1) for _ in range(self.sum_limit))
    return (r - 6) * self.sigma_x + self.ax
```

```

def generation_y(self):
    r = sum(random.uniform(0, 1) for _ in range(self.sum_limit))
    return (r - 6) * self.sigma_y + self.ay

def generate(self):
    x = [self.generation_x() for _ in range(self.n)]
    y = [self.generation_y() for _ in range(self.n)]
    return {'x': x, 'y': y}

```

Приклад роботи:

[6.474;9.922]	[-5.632;13.316]
[4.015;18.402]	[8.422;18.83]
[3.905;26.574]	[5.632;20.803]
[8.841;26.955]	[1.7;15.257]
[1.421;20.825]	[5.305;17.151]
[1.763;15.677]	[10.365;7.337]
[3.785;27.276]	[-4.265;19.497]
[0.083;16.361]	[0.121;12.531]
[-3.435;22.722]	[3.862;18.612]
[-8.431;13.057]	[1.817;17.335]
[2.403;1.23]	[7.941;21.937]
[10.754;28.154]	[-1.06;15.92]
[3.119;28.473]	[4.163;12.199]
[0.972;25.966]	[-4.165;8.029]
[-1.268;13.324]	[-1.13;22.938]
[7.685;21.847]	[-0.776;25.189]
[0.674;14.336]	[0.57;6.664]
[10.191;22.643]	[6.478;12.657]
[5.421;12.931]	[7.598;4.595]
[-0.273;17.476]	[4.336;11.625]
[-1.811;14.511]	[0.646;19.447]
[3.717;15.861]	[1.958;6.068]
[-5.464;10.995]	[3.0;11.353]
[4.908;16.035]	[8.573;16.737]
[10.533;11.266]	[9.071;14.747]
[2.331;18.371]	[4.677;6.716]
[6.315;23.553]	[2.458;23.767]
[-0.67;4.534]	[5.155;3.684]
[-2.282;26.069]	[-5.675;10.782]
[1.028;13.708]	[5.075;30.625]
[-3.954;15.1]	[0.352;5.153]
[-1.139;21.224]	[-1.782;22.032]
	[3.133;24.326]

2) Побудувати кореляційне поле в координатах (X ,Y);

Програмна реалізація:

```
x_coords = sample['x']
y_coords = sample['y']

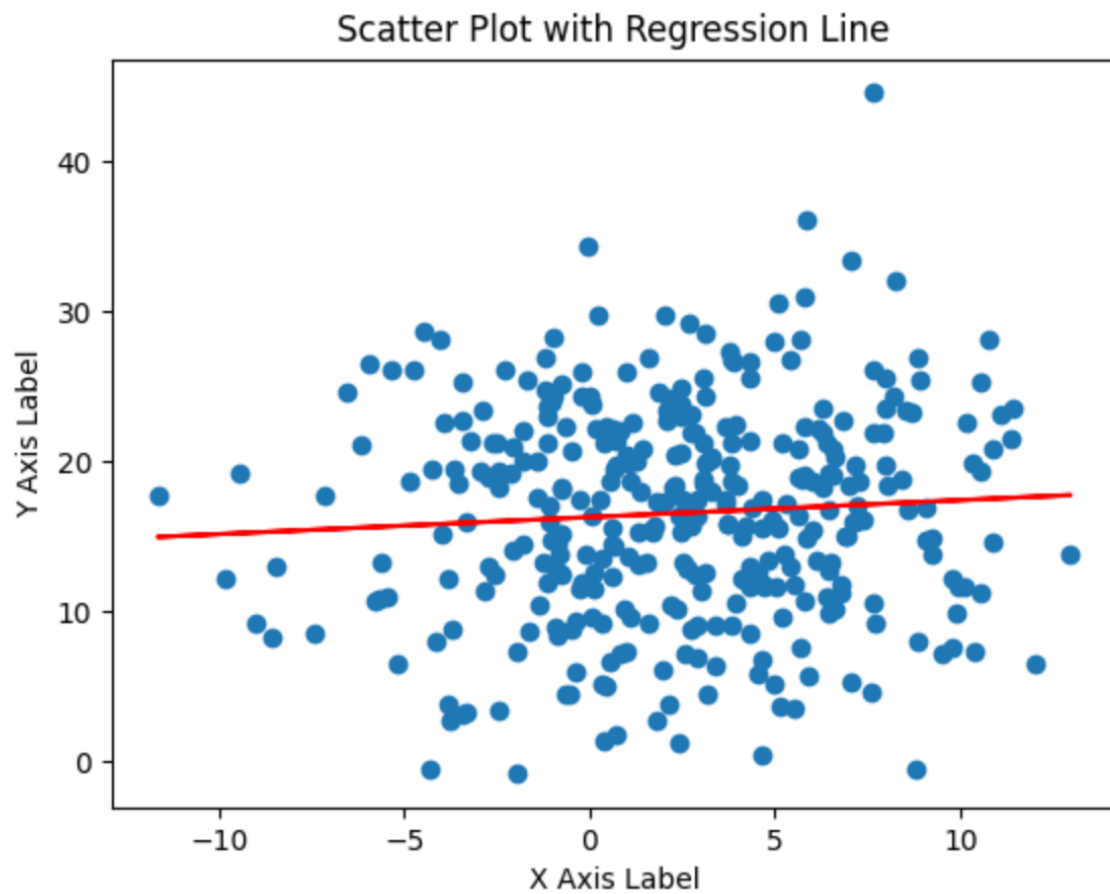
# емпірична лінія(лінія регресії)
slope, intercept = np.polyfit(x_coords, y_coords, 1)
regression_line_y = np.array(x_coords) * slope + intercept
```

```
plt.scatter(x_coords, y_coords)
plt.plot(x_coords, regression_line_y, color='red')

plt.xlabel('X Axis Label')
plt.ylabel('Y Axis Label')
plt.title('Scatter Plot with Regression Line')

plt.show()
```

Приклад роботи:



3) знайти вибіркові оцінки характеристик розподілів;

Програмна реалізація:

```
def вибіркве_середнє(дані):  
    return просумувати(дані) / довжина(дані)  
  
def sample_variance(я, дані):  
    вибіркве_середнє = я.вибіркве_середнє(дані)  
    сума = 0  
    for число in дані:  
        сума += pow((число - вибіркве_середнє), 2)  
    return сума / довжина(дані)  
  
def sample_root_mean_square_deviation(self, data):  
    return sqrt(self.sample_variance(data))  
  
def sample_correlation_coefficient(self, data_x, data_y):
```

```

root_deviation_x = self.sample_root_mean_square_deviation(data_x)
root_deviation_y = self.sample_root_mean_square_deviation(data_y)
Xv = self.вибіркове_середнє(data_x)
Yv = self.вибіркове_середнє(data_y)

```

```

suma = 0
for i in range(len(data_y)):
    suma += (data_x[i] - Xv) * (data_y[i] - Yv)
return suma / (len(data_x) * root_deviation_x * root_deviation_y)

```

Приклад роботи:

```

[16.828, 10.212]
sample mean x: 2.400817981086285
sample mean y: 16.551109789389475
sample variance x: 19.375810015412103
sample variance y: 50.95667441566298
sample root mean square deviation x: 4.40179622602093
sample root mean square deviation y: 7.138394386391311
sample correlation coefficient x and y: 0.0698936333780307

```

4) Перевірити гіпотезу про значущість вибіркового коефіцієнта кореляції;

Програмна реалізація:

```

r = numbers_analyzer.sample_correlation_coefficient(data_y=sample['y'], data_x=sample['x'])
T = (r * math.sqrt(350 - 2)) / math.sqrt(1 - r ** 2)
tkp = t.ppf(1 - 0.05 / 2, 350 - 2)

if np.abs(T) < tkp:
    print("Нема підстав відкидати нульову гіпотезу")
else:
    print("Нульову гіпотезу відкидаємо")

```

Приклад роботи:

Для даного набору даних результатом є такий рядок

```
Нема підстав відкидати нульову гіпотезу
```

Висновки:

В ході виконання лабораторної роботи опрацював набір даних згенерований на основі двовимірної випадкової величини, опанував способи визначення точкових оцінок для таких даних та визначення ступеня їхньої залежності.

sample mean x: 1.6848384266894163

sample mean y: 16.878150792963318

sample variance x: 17.500725656849436

sample variance y: 57.28528096135828

sample root mean square deviation x: 4.183386864353982

sample root mean square deviation y: 7.568704047679383

sample correlation coefficient x and y: 0.009884464342460807.

На основі критерію

$$T = \frac{r_{XY} \sqrt{n-2}}{\sqrt{1-r_{XY}^2}}$$

і по таблиці критичних точок розподілу Стьюдента, за заданим рівнем

