

Analiza Danych Eksperymentalnych

Sprawozdanie - Laboratorium 4

Statystyka opisowa

Michał KORDASZ 241289

Łukasz ŚMIERZCHAŁA 222276

6 grudnia 2021

Termin zajęć: Środa 11:15 TN

Data laboratorium: 1 grudnia 2021

Prowadzący: Dr inż. Krzysztof Halawa

Spis treści

1	Dane	3
2	Histogramy	3
2.1	Implementacja w Pythonie	4
3	Statystyka Opisowa	4
3.1	Średnia arytmetyczna	5
3.2	Momenty centralne	6
3.3	Odchylenie standardowe	6
3.4	Współczynnik asymetrii	6
3.5	Współczynnik koncentracji	7
3.6	Współczynnik zmienności	7
4	Wnioski	7

Tabela 1: Dane

środkie klas x_i	Liczność klas n_i			
	szereg I	szereg II	szereg III	szereg IV
10	0	6	0	6
20	18	6	6	12
30	36	30	60	30
40	42	66	36	36
50	36	30	30	60
60	18	6	12	6
70	0	6	6	0

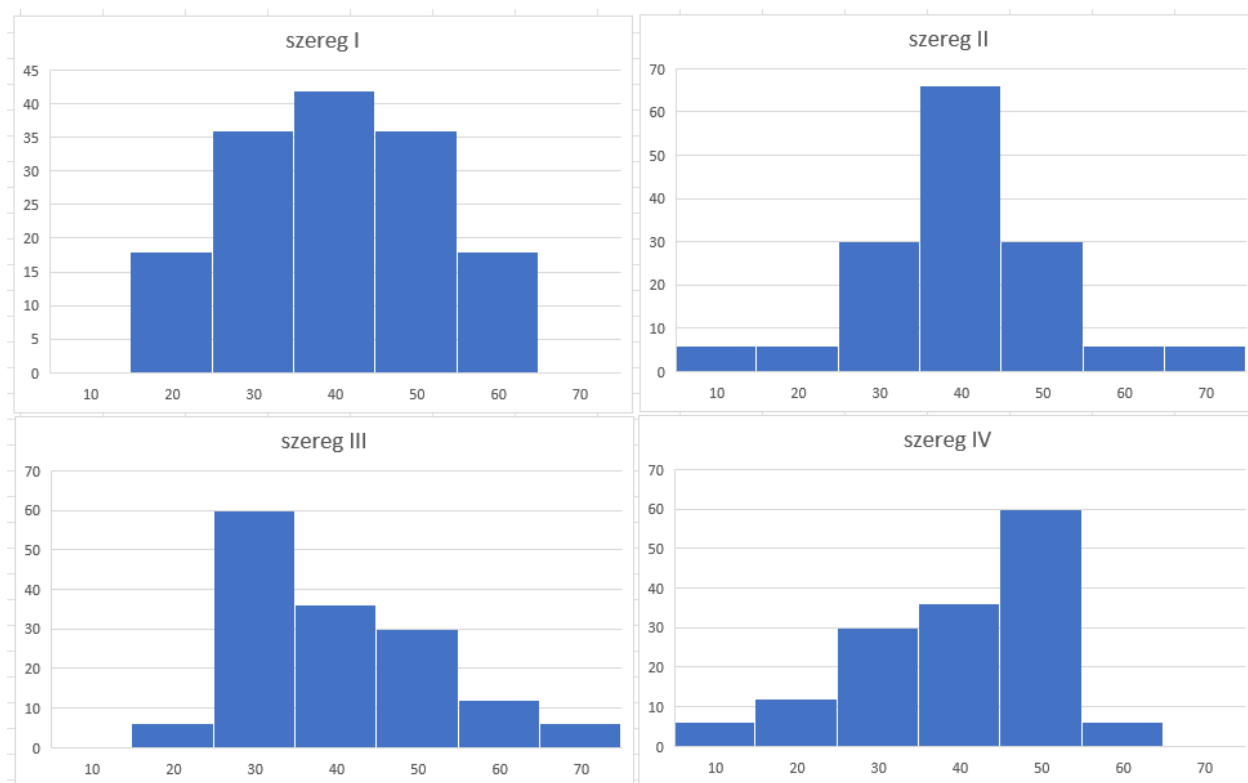
1 Dane

Do przeprowadzenia ćwiczenia wykorzystane zostały 4 szeregi rozdzielcze danych. Dane przedstawione zostały w tabeli 1.

2 Histogramy

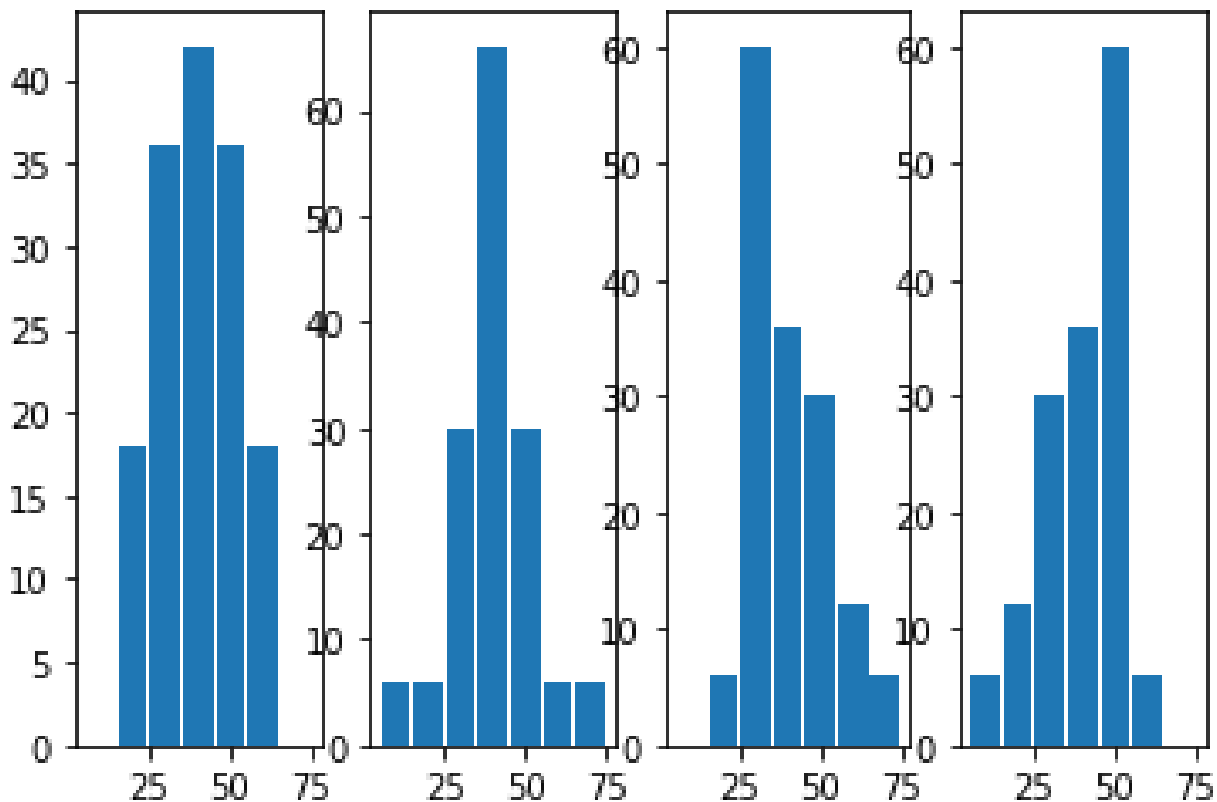
Histogramy szeregów zostały przedstawione na rys. 2.

W Excelu histogramy zostały utworzone za pomocą operacji **Insert -> Histogram**.



Rysunek 1: Histogramy szeregów rozdzielczych

2.1 Implementacja w Pythonie



Rysunek 2: Histogramy szeregów rozdzielczych w Pythonie

```
1 plt.subplot(1,4,1).bar(data[:,0], data[:,1], width=9)
2 plt.subplot(1,4,2).bar(data[:,0], data[:,2], width=9)
3 plt.subplot(1,4,3).bar(data[:,0], data[:,3], width=9)
4 plt.subplot(1,4,4).bar(data[:,0], data[:,4], width=9)
```

3 Statystyka Opisowa

W ramach laboratoriów obliczono następujące wartości charakteryzujące szeregi rozdzielcze:

- Średnia arytmetyczna
- 1, 2, 3, 4 Moment centralny
- Odchylenie standardowe
- Współczynnik asymetrii
- Współczynnik koncentracji
- Współczynnik zmienności

Wyniki zostały przedstawione na rys. 3.

	A	B	C	D	E	F
1		xi	ni			
2			I	II	III	IV
3		10	0	6	0	6
4		20	18	6	6	12
5		30	36	30	60	30
6		40	42	66	36	36
7		50	36	30	30	60
8		60	18	6	12	6
9		70	0	6	6	0
10						
11	średnia arytmetyczna		40	40	40	40
12						
13	momenty centralne	1	0	0	0	0
14		2	144	144	144	144
15		3	0	0	1200	-1200
16		4	43200	81600	57600	57600
17						
18	odchylenie standardowe		12	12	12	12
19						
20	współczynnik asymetrii		0	0	0.694	-0.694
21						
22	współczynnik koncentracji		2.083	3.935	2.778	2.778
23						
24	współczynnik zmienności		30%	30%	30%	30%

Rysunek 3: Statystyka Opisowa

3.1 Średnia arytmetyczna

Średnia arytmetyczna szeregu rozdzielczego dana jest wzorem:

$$\bar{x} = \frac{\sum x_i n_i}{\sum n_i}$$

W excelu średnia arytmetyczna została obliczona przy pomocy następującej formuły:

```
=SUM(B3:B9*C3:C9)/SUM(C3:C9)
```

Implementacja w Pythonie:

```
1 means = np.zeros(4)
2 for i in range(len(means)):
```

```

3   for j in range(len(data)):
4       means[i] += data[j,0]*data[j,i+1]
5 for i in range(len(means)):
6     means[i] = means[i]/150

```

3.2 Momenty centralne

Moment centralny r-rzędu szeregu rozdzielczego dany jest wzorem:

$$M_r = \frac{1}{n} \sum_{i=1}^k (x_i - \bar{x})^r n_i$$

W excelu momenty centralne zostały obliczone przy pomocy następującej formuły:

```

1 =SUM(POWER($B$3:$B$9-C$11, $B13)*C$3:C$9)/SUM(C$3:C$9)

```

Implementacja w Pythonie:

```

1 momentums = np.zeros((4,4))
2   for i in range(len(means)):
3       for j in range(len(data)):
4           #print(f'({data[j,0]} - {means[i]}) * {data[j, i+1]}')
5           momentums[i, 0] += (pow((data[j,0] - means[i]), 1) * data[j, i+1])
6           momentums[i, 1] += (pow((data[j,0] - means[i]), 2) * data[j, i+1])
7           momentums[i, 2] += (pow((data[j,0] - means[i]), 3) * data[j, i+1])
8           momentums[i, 3] += (pow((data[j,0] - means[i]), 4) * data[j, i+1])
9
10  for i in range(len(momentums)):
11      momentums[i] = momentums[i]/150

```

3.3 Odchylenie standardowe

Odchylenie standardowe jest pierwiastkiem z wariancji (a wariancją jest drugi moment centralny), może więc być reprezentowane przy pomocy wzoru:

$$\sigma = \sqrt{C_2}$$

W excelu odchylenie standardowe został obliczony przy pomocy następującej formuły:

```

1 =SQRT(C14)

```

Implementacja w Pythonie:

```

1 dev = np.zeros(4)
2 for i in range(len(dev)):
3     dev[i] = math.sqrt(momentums[i,1])

```

3.4 Współczynnik asymetrii

Współczynnik asymetrii dany jest wzorem:

$$g = \frac{\sqrt{C_3}}{\sigma^3}$$

W excelu współczynnik asymetrii został obliczony przy pomocy następującej formuły:

```

1 =C15/POWER(C18, 3)

```

Implementacja w Pythonie:

```

1 skew = np.zeros(4)
2 for i in range(len(skew)):
3     skew[i] = momentums[i, 2] / dev[i]**3

```

3.5 Współczynnik koncentracji

Współczynnik koncentracji dany jest wzorem:

$$K = \frac{\sqrt{C_4}}{\sigma^4}$$

W excelu współczynnik koncentracji został obliczony przy pomocy następującej formuły:

```
1 =C16/POWER(C18, 4)
```

Implementacja w Pythonie:

```
1 coef_cont = np.zeros(4)
2 for i in range(len(coef_cont)):
3     coef_cont[i] = momentums[i, 3] / dev[i]**4
```

3.6 Współczynnik zmienności

Współczynnik zmienności dany jest wzorem:

$$V = \frac{\sigma}{\bar{x}} 100\%$$

W excelu współczynnik zmienności został obliczony przy pomocy następującej formuły:

```
1 =C18/C11
```

Implementacja w Pythonie:

```
1 change_coef = np.zeros(4)
2 for i in range(len(change_coef)):
3     change_coef[i] = (dev[i] / means[i])*100
```

4 Wnioski

Output z Pythona:

```
1 Dane:
2 [[10  0  6  0  6]
3  [20 18  6  6 12]
4  [30 36 30 60 30]
5  [40 42 66 36 36]
6  [50 36 30 30 60]
7  [60 18  6 12  6]
8  [70  0  6  6  0]]
9 Momenty:
10 [[ 0.  144.  0. 43200.]
11  [ 0.  144.  0. 81600.]
12  [ 0.  144. 1200. 57600.]
13  [ 0.  144. -1200. 57600.]]
14 Odchylenie standardowe:
15 [12. 12. 12. 12.]
16 Współczynnik asymetrii:
17 [ 0.  0.  0.69444444 -0.69444444]
18 Współczynnik koncentracji skupienia:
19 [2.08333333 3.93518519 2.77777778 2.77777778]
20 Współczynnik zmiany [%]:
21 [30. 30. 30. 30.]
```

Wszystkie otrzymane wyniki w Excelu są tożsame z kodem w Pythonie.

Pierwszy moment centralny jest zawsze równy 0.

Drugi moment centralny dla szeregów o takiej samej średniej przyjmuje takie same wartości, niezależnie od rozłożenia elementów.

Trzeci moment centralny dla szeregów symetrycznych przyjmuje wartość 0, natomiast dla szeregów asymetrycznych, przyjmuje wartości dodatnie, gdy asymetria jest prawostronna, oraz ujemne, gdy asymetria jest lewostronna. Wartość momentu jest tym większa, im większa jest asymetria.

Czwarty moment centralny przyjmuje tym mniejsze wartości, im bardziej skoncentrowane ('zbite') są elementy szeregu.

Współczynnik asymetrii oblicza się w oparciu o trzeci moment centralny i określa czy / jak bardzo dany szereg jest asymetryczny.

Współczynnik koncentracji oblicza się w oparciu o czwarty moment centralny i określa jak bardzo elementy szeregu są skoncentrowane wokół wartości średniej.

Współczynnik zmienności jest miarą zróżnicowania rozkładu cechy.