

Politechnika Gdańska



Mariusz Kaszubowski
Katedra Statystyki
Wydział Zarządzania i Ekonomii
Politechnika Gdańska

Grupowanie i prezentacja danych

Ćwiczenia nr 2
Podstawy Statystyki

Podstawowe pojęcia

- Populacja statystyczna
- Jednostka statystyczna
- Cecha statystyczna (jakościowa-dychotomiczna, jakościowa - wielodzielna, ilościowa - skokowa, ilościowa - ciągła)
- Grupowanie typologiczne i wariacyjne
- Szereg statystyczny (szczegółowy, rozdzielczy, strukturalny, punktowy, przedziałowy, czasowy, rozdzielczy kumulacyjny)
- Wykresy statystyczne (punktowe, obrazkowe, powierzchniowe, liniowe, mapowe, złożone)

Tworzenie szeregów rozdzielczych

- Ustalenie liczby klas (k)
- Ustalenie rozpiętości przedziałów (h)
- Ustalenie górnych i dolnych granic przedziałów
- Przyporządkowanie danych do przedziałów klasowych
- Zliczenie jednostek w poszczególnych klasach

Przykład

- Stopę bezrobocia rejestrowanego w % w Polsce wg województw na koniec roku 1996 przedstawia następujący szereg (Źródło: Rocznik Statystyczny 1997, GUS, tabl. 21(214), s. 141):

4,1; 11,9; 11,7; 9,6; 16,7; 13,1; 19,0; 12,1; 23,4; 10,6; 17,0; 18,2;
14,9; 8,4; 15,2; 17,3; 24,7; 6,1; 14,9; 16,9; 12,0; 11,7; 14,3; 16,2;
12,8; 23,6; 12,9; 17,0; 17,4; 16,8; 16,6; 6,2; 14,2; 17,2; 14,6; 10,7;
12,7; 10,7; 25,7; 24,6; 13,1; 13,9; 12,4; 18,6; 21,7; 21,5; 9,8; 12,4;
15,3.

1. Określ badaną zbiorowość, jednostkę oraz cechę statystyczną.
2. Przeprowadź grupowanie statystyczne.
3. Zaprezentuj graficznie otrzymany szereg.

Rozwiązanie nr 1

Przedział klasowy $x_{0i} - x_{1i}$	Pomiary z przedziału	Liczba jednostek n_i	Środek przedziału \dot{x}_i	Średnia z przedziału \bar{x}_i	$ \dot{x}_i - \bar{x}_i $
$(2,3; 5,9)$					
$(5,9; 9,5)$					
$(9,5; 13,1)$					
$(13,1; 16,7)$					
$(16,7; 20,3)$					
$(20,3; 23,9)$					
$(23,9; 27,5)$					
Suma	X	49	X	X	

Rozwiązanie 1 c.d.

Przedział klasowy $x_{0i} - x_{1i}$	Pomiary z przedziału	Liczba jednostek n_i	Środek przedziału \dot{x}_i	Średnia z przedziału \bar{x}_i	$ \dot{x}_i - \bar{x}_i $
$(2,3; 5,9)$	4,1	1	4,1	4,1	0
$(5,9; 9,5)$	6,1; 6,2; 8,4	3	7,7	6,9	0,8
$(9,5; 13,1)$	9,6; 9,8; 10,6; 10,7; 10,7; 11,7; 11,7; 11,9; 12,0; 12,1; 12,4; 12,4; 12,7; 12,8; 12,9; 13,1; 13,1	17	11,3	11,8	0,5
$(13,1; 16,7)$	13,9; 14,2; 14,3; 14,6; 14,9; 14,9; 15,2; 15,3; 16,2; 16,6; 16,7	11	14,9	15,2	0,3
$(16,7; 20,3)$	16,8; 16,9; 17,0; 17,0; 17,2; 17,3; 17,4; 18,2; 18,6; 19,0	10	18,5	17,5	1
$(20,3; 23,9)$	21,5; 21,7; 23,4; 23,6;	4	22,1	22,6	0,5
$(23,9; 27,5)$	24,6; 24,7; 25,7	3	25,7	25	0,7
Suma	X	49	X	X	3,8

Warunki poprawnego grupowania

$$R = x_{\max} - x_{\min}$$

$$k \leq \sqrt{n} \quad \text{lub} \quad k \leq 1 + 3,322 \cdot \log n$$

$$h = \frac{R}{k}$$

$$x_{01} = x_{\min} - 0,5 \cdot h$$

$$x_{11} = x_{01} + h$$

⋮

$$x_{0i} = x_{01} + (i-1) \cdot h$$

$$x_{1i} = x_{0i} + h$$

$$i = 2, 3, 4, \dots, k$$

liczba obserwacji (n)	liczba zalecanych klas
40 – 60	6 – 8
60 – 100	7 – 10
100 – 200	9 – 12
200 – 500	11 – 17
500 – 1500	18 – 25
ponad 1500	25

$$n_i \geq 5$$

$$\sum_{i=1}^k |\dot{x}_i - \bar{x}_i| \leq 0,5 \cdot h$$

Rozwiązanie nr 2

$$k = \sqrt{49} = 7$$

$$k \approx 1 + 3,322 \log 49 \approx 6,615$$

$$k \leq 6,615 \Rightarrow k = 6$$

$$R = 25,7 - 4,1 = 21,6$$

$$h = \frac{21,6}{6} = 3,6 \approx 4$$

$$x_{01} = 4,1 - 0,5 \cdot 4 = 2,1 \approx 2$$

Przedział klasowy	Liczba jednostek	% jednostek	% skumulowany
Suma	49	100%	X

Rozwiązanie (Statistica 8.0)

Klasa	Tabela liczności: Tabela liczności dla kategorii użytkownika. (stopa_bezrobocia_w_woj)			
	Liczba	Skumulow. Liczba	Procent	Skumulow. Procent
Włącz v1<=5,9	1	1	2,04082	2,0408
Włącz v1<=9,5	3	4	6,12245	8,1633
Włącz v1<=13,1	17	21	34,69388	42,8571
Włącz v1<=16,7	11	32	22,44898	65,3061
Włącz v1<=20,3	10	42	20,40816	85,7143
Włącz v1<=23,9	4	46	8,16327	93,8776
Włącz v1<=27,5	3	49	6,12245	100,0000
Nie określone	0	49	0,00000	100,0000

Od Do		Tabela liczności: Zmn1 (stopa_bezrobocia_w_woj_na_31.12.1996)			
		Liczba	Skumulow. Liczba	Procent	Skumulow. Procent
2,000000<=x<6,000000		1	1	2,04082	2,0408
6,000000<=x<10,00000		5	6	10,20408	12,2449
10,000000<=x<14,00000		16	22	32,65306	44,8980
14,000000<=x<18,00000		17	39	34,69388	79,5918
18,000000<=x<22,00000		5	44	10,20408	89,7959
22,000000<=x<26,00000		5	49	10,20408	100,0000
Braki		0	49	0,00000	100,0000