

## POJĘCIA WSTĘPNE

**STATYSTYKA** - nauka traktująca o metodach ilościowych badania prawidłowości zjawisk (procesów) masowych.

**BADANIE STATYSTYCZNE** - ogół prac mających na celu poznanie struktury określonej zbiorowości statystycznej.

**ZBIOROWOŚĆ (POPULACJA) STATYSTYCZNA** - zbiór dowolnych elementów (osób, przedmiotów, faktów) podobnych pod względem określonych cech (ale nie identycznych) poddanych badaniu statystycznemu.

**JEDNOSTKA STATYSTYCZNA** - składowe (elementy) zbiorowości (obiekty badania), które podlegają bezpośrednio obserwacji lub pomiarowi.

**$n$**  - oznaczenie liczby jednostek statystycznych w populacji

**ZBIOROWOŚĆ (POPULACJA) GENERALNA** - wszystkie elementy będące przedmiotem badania, co do których chcemy formułować wnioski ogólne.

**ZBIOROWOŚĆ PRÓBNA (PRÓBA)** - podzbiór populacji generalnej; wyniki badań próby są uogólniane na zbiorowość generalną. Próba musi być reprezentatywna.

Reprezentatywność zależy od: sposobu wyboru jednostek (celowy, losowy) oraz liczebności próby.

**$n > 30$**  - duża próba

**$n \leq 30$**  - mała próba

# RODZAJE BADANIA STATYSTYCZNEGO

1. całkowite (wyczerpujące)

2. częściowe (reprezentacyjne, ankietowe, itp.)

**CECHA STATYSTYCZNA** - podlegająca badaniu właściwość jednostki statystycznej.

Cechę oznaczamy dużą literą (np. X, Y, Z, ...).

Wartość cechy dla konkretnej jednostki (np. jednostki o numerze  $i$ ) oznaczamy małą literą z indeksem dolnym (np.  $x_i, y_i, z_i, \dots$ ).

Cechy statystyczne dzielimy na:

- cechy mieralne (ilościowe , kwantytatywne). Cechy te dzielimy dalej na: ciągłe i skokowe (dyskretne),
- cechy niemieralne (jakościowe , kwalitatywne)

Inny podział cech to podział na:

- STYMULANTY
- DESTYMULANTY
- NOMINANTY

**poczytać w domu:**

**rozdz. 1 i 2 z książki [2] oraz rozdz. 2 z książki [3]**

# GRUPOWANIE materiału statystycznego

**SZEREGI STATYSTYCZNE** - odpowiednio usystematyzowany i uporządkowany surowy materiał statystyczny.

Szeregi statystyczne dzielimy na szeregi:

- szczegółowe
- rozdzielcze (punktowe, przedziałowe)
- czasowe (momentów, okresów)

## PRZYKŁAD 1 (szereg szczegółowy i szereg rozdzielczy)

Przedmiotem badania jest wadliwość produkcji na III zmianie w firmie DINO. Liczba wyprodukowanych wyrobów wynosi 50 ( $n=50$ ). Cecha badana ( $X$ ) oznacza liczbę usterek w wyrobie.

Surowy materiał statystyczny to ciąg 50 liczb oznaczający liczbę braków stwierdzonych w kolejnym wyrobie. Ma on postać:

3, 0, 1, 0, 0, 0, 0, 1, 0, 0, 4, 2, 0, 0, 1, 1, 0, 3, 0, 2,  
0, 1, 0, 0, 1, 0, 2, 0, 0, 0, 3, 0, 0, 4, 0, 2, 0, 0, 1, 0,  
0, 1, 3, 0, 2, 0, 0, 0, 0, 2

SZEREG SZCZEGÓŁOWY otrzymamy sortując te liczby rosnąco (najczęściej) lub malejąco (rzadziej). Liczby braków posortowane rosnąco dają następujący ciąg  $\{x_i\}$  50 liczb:

0,  
0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2,  
2, 2, 2, 2, 3, 3, 3, 3, 4, 4

Liczby te spełniają warunek definicji szeregu szczegółowego:

$$x_1 \leq x_2 \leq x_3 \leq \dots \leq x_{50}$$

## SZEREG ROZDZIELCZY PUNKTOWY

<i>numer klasy</i>	<i>liczba braków</i>	<i>liczba wyrobów (liczebność)</i>
<i>i</i>	<i>x<sub>i</sub></i>	<i>n<sub>i</sub></i>
<b>1</b>	<b>0</b>	
<b>2</b>	<b>1</b>	
<b>3</b>	<b>2</b>	<b>6</b>
<b>4</b>	<b>3</b>	<b>4</b>
<b>5</b>	<b>4</b>	<b>2</b>
<b>razem</b>	<b>×</b>	

<i>numer klasy</i>	<i>liczba braków</i>	<i>liczebność skumulowana</i>
<i>i</i>	<i>x<sub>i</sub></i>	<i>n<sub>i sk</sub></i>
<b>1</b>	<b>0</b>	
<b>2</b>	<b>1</b>	
<b>3</b>	<b>2</b>	
<b>4</b>	<b>3</b>	
<b>5</b>	<b>4</b>	
<b>razem</b>	<b>×</b>	<b>×</b>

## WSKAŹNIK STRUKTURY ( $w_i$ )

**Wskaźnik struktury (inaczej częstość) nazywany jest też: liczebnością względną, frakcją, odsetkiem. Wylicza się go następująco:**

$$w_i = \frac{n_i}{n} \quad \text{dla} \quad i = 1, 2, \dots, k$$

<i>numer klasy</i>	<i>liczba braków</i>	<i>liczba wyrobów (liczebność)</i>	<i>wskaźnik struktury</i>
<i>i</i>	<i>x<sub>i</sub></i>	<i>n<sub>i</sub></i>	<i>w<sub>i</sub></i>
<b>1</b>	<b>0</b>	<b>30</b>	
<b>2</b>	<b>1</b>	<b>8</b>	
<b>3</b>	<b>2</b>	<b>6</b>	
<b>4</b>	<b>3</b>	<b>4</b>	<b>0,08</b>
<b>5</b>	<b>4</b>	<b>2</b>	<b>0,04</b>
<b>razem</b>	<b>×</b>	<b>50</b>	

**Kolumna liczb {  $w_i$  } nazywana jest rozkładem empirycznym (liczby usterek).**

## **SKUMULOWANY WSKAŹNIK STRUKTURY ( $w_{i\ sk}$ )**

**Skumulowany wskaźnik struktury (inaczej: częstość skumulowana).  
Wylicza się go następująco:**

$$w_{i\ sk} = \frac{n_{i\ sk}}{n} \quad \text{dla} \quad i = 1, 2, \dots, k$$

<i>numer klasy</i>	<i>liczba braków</i>	<i>liczebność skumulowana</i>	<i>skumulowany wskaźnik struktury</i>
<i>i</i>	<i>x<sub>i</sub></i>	<i>n<sub>i sk</sub></i>	<i>w<sub>i sk</sub></i>
<b>1</b>	<b>0</b>	<b>30</b>	
<b>2</b>	<b>1</b>	<b>38</b>	
<b>3</b>	<b>2</b>	<b>44</b>	
<b>4</b>	<b>3</b>	<b>48</b>	
<b>5</b>	<b>4</b>	<b>50</b>	
<b>razem</b>	<b>×</b>	<b>×</b>	<b>×</b>

**Kolumna liczb  $\{ w_{i\ sk} \}$  nazywana jest  
dystribuantą empiryczną (liczby usterek).**

## **ZALECENIA przy grupowaniu w szereg rozdzielczy przedziałowy**

$k$  - liczba klas

$$k \approx \sqrt{n}$$

$h$  - rozpiętość przedziału

$$h \approx \frac{x_{\max} - x_{\min}}{k}$$

### **PRZYKŁAD 2 (szereg rozdzielczy przedziałowy)**

**Przedmiotem badania jest czas dojazdu do pracy w dwóch firmach: DINO i ZAUR.**

**Surowy materiał empiryczny znajduje się na kserówce.**

**Czas dojazdu pracowników firmy DINO [w minutach]**

	1	2	3	4	5	6	7	8	9	10
00	17	28	22	20	53	36	37	21	19	21
10	22	17	32	33	22	31	9	30	19	24
20	16	22	20	34	19	15	21	17	17	27
30	36	51	23	24	16	38	21	13	42	19
40	17	25	19	16	19	29	31	24	39	28
50	65	18	19	20	48	23	60	21	24	37
60	19	22	30	22	21	19	27	31	18	43
70	24	23	58	30	47	19	32	20	18	20
80	20	35	17	5	24	27	31	51	32	39
90	60	58	39	19	24	21	29	14	18	16

**Czas dojazdu pracowników firmy ZAUR [w minutach]**

	1	2	3	4	5	6	7	8	9	10
00	42	49	36	46	21	47	38	15	64	51
10	52	41	25	36	24	33	39	36	47	53
20	49	8	43	54	44	7	51	55	43	8
30	40	47	40	25	49	19	55	56	29	40
40	28	19	33	46	29	35	47	34	41	40
50	32	53	53	54	59	18	34	52	49	17
60	38	36	49	43	49	44	38	17	54	30
70	51	41	50	21	19	49	49	22	44	54
80	25	60	39	34	37	54	17	52	11	12
90	32	31	17	11	32	43	62	39	22	40
100	49	31	46	50	50	33	47	12	64	53
110	55	43	28	55	63	49	28	38	51	46
120	48	40	55	5	38	37	50	49	46	51
130	33	53	47	26	65	46	26	47	28	50
140	24	65	45	46	40	42	38	21	39	52
150	42	49	19	46	49	51	39	31	38	48
160	50	52	47	33	37	24	54	47	44	53
170	31	41	43	49	53	32	48	41	53	35
180	41	28	34	50	37	46	41	49	54	50
190	39	48	28	10	53	63	47	55	45	50

*Pogrupuj dane w szeregi rozdzielcze następującej postaci :*

DINO	
czas dojazdu	liczba pracowników
5 - 15	
15 - 25	
25 - 35	
35 - 45	
45 - 55	
55 - 65	
razem	

ZAUR	
czas dojazdu	liczba pracowników
5 - 15	
15 - 25	
25 - 35	
35 - 45	
45 - 55	
55 - 65	
razem	



## WSKAŹNIK PODOBIEŃSTWA STRUKTUR

Wskaźnik podobieństwa struktur ( $w_p$ ) jest najprostszą miarą statystyczną pozwalającą ocenić podobieństwo kształtowania się badanej cechy w dwóch różnych zbiorowościach.

Wyliczamy go następująco:

$$w_p = \sum_{i=1}^k \min\{w_{1i}, w_{2i}\}$$

<i>numer klasy</i>	<i>czas dojazdu</i>	<i>częstość DINO</i>	<i>częstość ZAUR</i>	<i>obliczenia wskaźnika</i>
<i>i</i>	$x_{0i} - x_{1i}$	$w_{1i}$	$w_{2i}$	$w_p$
<b>1</b>	<b>5 – 15</b>	<b>0,05</b>	<b>0,05</b>	
<b>2</b>	<b>15 – 25</b>	<b>0,55</b>	<b>0,10</b>	
<b>3</b>	<b>25 – 35</b>	<b>0,20</b>	<b>0,15</b>	
<b>4</b>	<b>35 – 45</b>	<b>0,10</b>	<b>0,25</b>	
<b>5</b>	<b>45 – 55</b>	<b>0,05</b>	<b>0,40</b>	<b>0,05</b>
<b>6</b>	<b>55 – 65</b>	<b>0,05</b>	<b>0,05</b>	<b>0,05</b>
<b>razem</b>	<b>×</b>	<b>1,00</b>	<b>1,00</b>	

## **PREZENTACJA GRAFICZNA SZEREGÓW STATYSTYCZNYCH**

**HISTOGRAM - wykres słupkowy**

**DIAGRAM - wykres liniowy**

**Oba typy wykresów mogą być sporządzane w wariantach dla:**

- **liczebności**
- **liczebności skumulowanej**
- **częstości**
- **częstości skumulowanej**

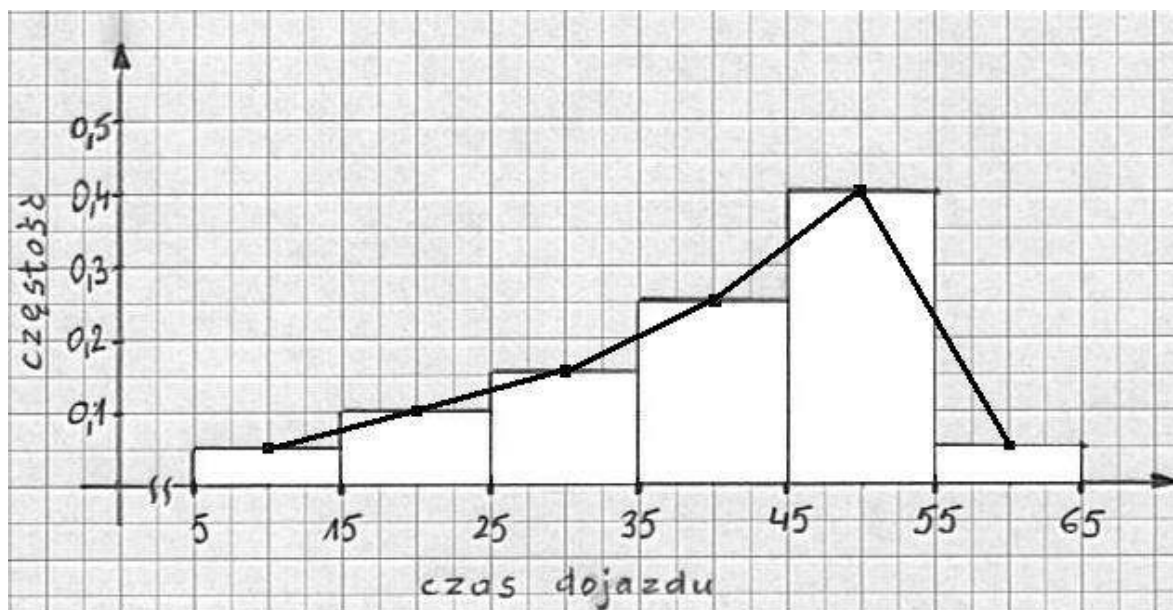
**Przykłady histogramów i diagramów dla ZAUR (s. 15).**

**Dla wzrokowego porównania rozkładu badanej cechy w dwóch (lub więcej) zbiorowościach używamy wyłącznie wykresów częstościowych (s.16).**

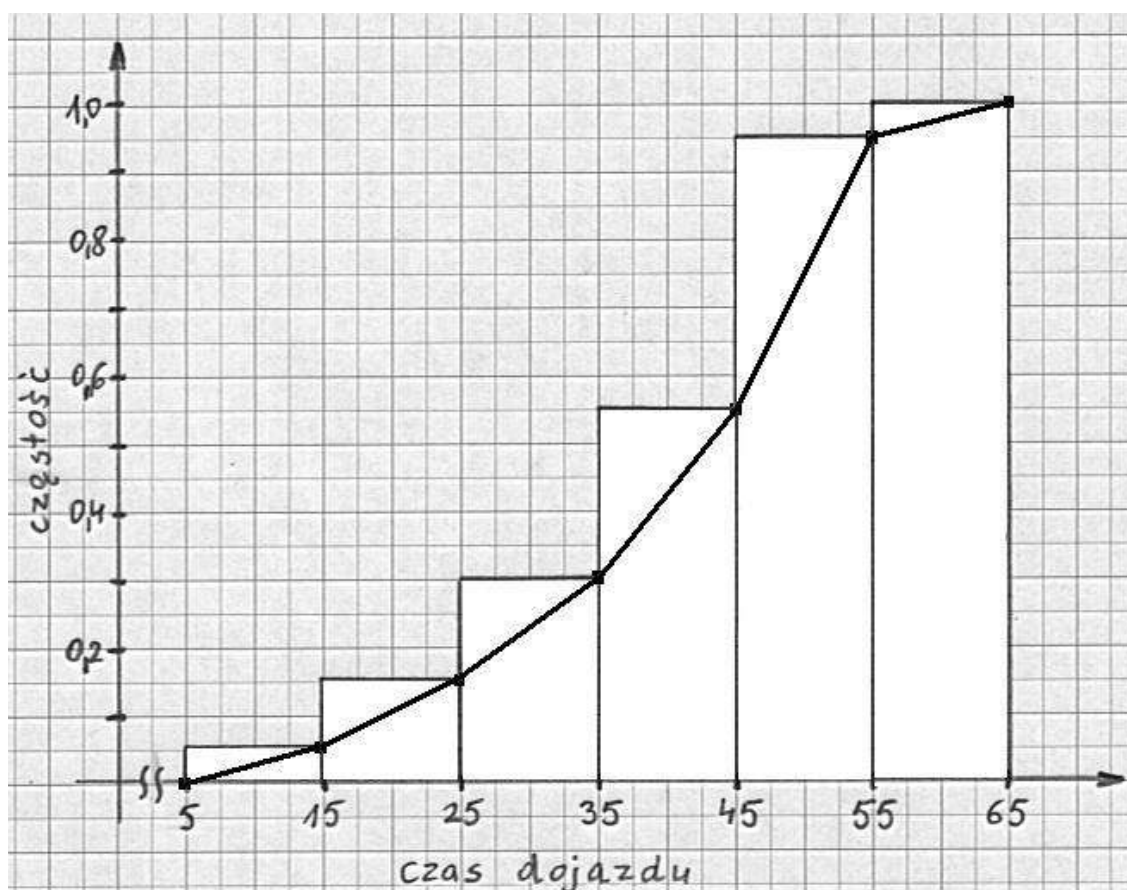
**Dla DINO wykonać je samodzielnie w domu.**

**O innych typach wykresów poczytać samodzielnie we wskazanych wcześniej rozdziałach.**

## Histogram i diagram częstości dla czasu dojazdu pracowników firmy ZAUR



## Histogram i diagram częstości skumulowanej dla czasu dojazdu pracowników firmy ZAUR



## Diagramy częstości dla czasu dojazdu pracowników firm DINO i ZAUR

