

**Zagadnienia:** Szereg rozdzielczy, histogram, miary położenia i zmienności dla szeregu szczegółowego i rozdzielczego, wykres pudełkowy.

**Zad.1** Dane pochodzą z bazy danych gestów wykonywanych przez różne osoby. Pomiary parametrów gestykulacji odbywały się z wykorzystaniem czujnika przyspieszenia umieszczonego na palcu. Dane są przedstawione w szeregu szczegółowym, zawierają informację w milisekundach na temat czasu potrzebnego kolejnym osobom do wykonania gestu – „zatoczenie ręką koła”. Każda z osób wykonała 36 powtórzeń tego gestu.

Wykorzystując wzór na liczbę klas szeregu rozdzielczego:  $k = \sqrt{n}$ , gdzie  $k$  – liczba klas,  $n$  – liczebność próby, zbuduj z tych danych szereg rozdzielczy.

$i$	Dolna granica przedziału $x_{1i}$	Środek przedziału $\hat{x}$	Górna granica przedziału $x_{2i}$	Liczebność przedziału $n_i$	Liczebność skumulowana $n(x_{2i})$	Częstość $w_i$	Częstość skumulowana $F(x_{2i})$
1.							
2.							
...							
...							

**Zad.2** Wykorzystując szereg rozdzielczy, narysuj histogram. Na osi Y zaznacz częstości odpowiadające poszczególnym przedziałom. Do wykresu dodaj krzywą przedstawiającą częstość skumulowaną.

**Zad.3** Dysponując wzorami na średnią, wariancję, odchylenie standardowe dla szeregu szczegółowego, wyprowadź wzory na podstawie których można byłoby te parametry wyliczyć dla szeregu rozdzielczego. Dla szeregu rozdzielczego przyjmij takie oznaczenia jak w tabeli w zadaniu 3.

Wzory dla szeregu szczegółowego:

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

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

Odchylenie standardowe: 
$$s = \sqrt{s^2}$$

Gdzie:  $n$  – liczebność próby;  $x_i$  – kolejne wartości cechy z szeregu szczegółowego;

**Zad.4** Kwantyl to wartość cechy w uporządkowanym szeregu statystycznym, poniżej której znajduje się  $n_{kw}$  jednostek statystycznych, natomiast wartości pozostałych  $n - n_{kw}$  jednostek są nie mniejsze od tej wartości. Wśród kwantyli można wyróżnić:

- Kwartyle* – dzielą zbiorowość na 4 równe części;
- Decyle* – dzielą zbiorowość na 10 równych części;
- Percentyle* – dzielą zbiorowość na 100 równych części.

W praktyce najczęściej wykorzystuje się kwartyle. Oznacza się je jako:

- $Q_1$  - pierwszy kwartył, dzieli zbiorowość w stosunku 25% : 75%
- $Q_2 = Me$  – drugi kwartył, czyli mediana, dzieli zbiorowość w stosunku 50% : 50%
- $Q_3$  – trzeci kwartył, dzieli zbiorowość w stosunku 75% : 25%

Poprzez *IQR* oznacza się rozstęp międzykwartyłowy:

$$IQR = Q_3 - Q_1$$

Medianę dla szeregu rozdzielczego wyznacza się wg wzoru:

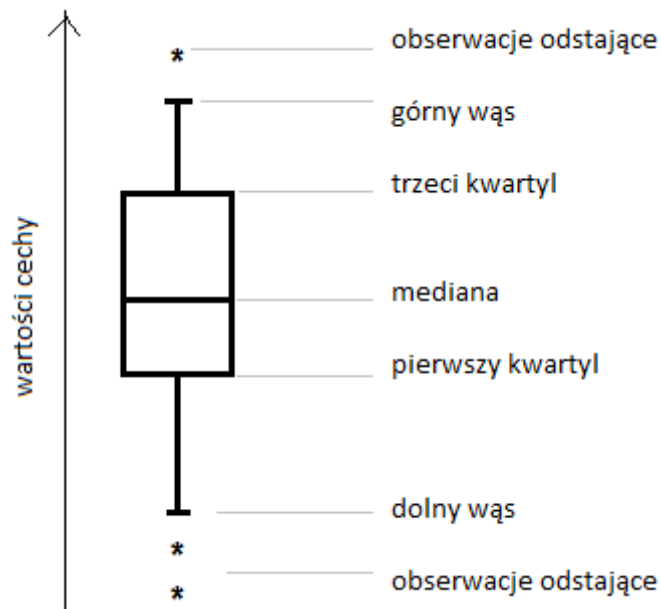
$$Me = x_{1Me} + \left(\frac{1}{2}n - \sum_{i=1}^m n_i\right) \cdot \frac{c}{n_{Me}}$$

Gdzie:  $x_{1Me}$  – dolna granica przedziału zawierającego medianę;  $m$  – liczba przedziałów poprzedzających przedział z medianą;  $c$  – długość przedziału, w którym znajduje się mediana;  $n_{Me}$  – liczebność przedziału, w którym znajduje się mediana;

**W jaki sposób dla szeregu rozdzielczego można wyliczyć kwartyle  $Q_1$  i  $Q_3$ ?**

**Zad.5** Wylicz medianę, kwartyle oraz rozstęp międzykwartyłowy dla danych z szeregu rozdzielczego. Następnie narysuj wykres pudełkowy.

Wykres pudełkowy:



Długość każdego z wąsów jest równa  $1,5IQR$ , chyba, że:

- wartość maksymalna jest mniejsza niż  $Q_3 + 1,5IQR$ ;
- wartość minimalna jest większa niż  $Q_1 - 1,5IQR$ .

W takim przypadku długość wąsa jest zdeterminowana przez odpowiednio wartość maksymalną lub minimalną. Obserwacje znajdujące się poza 3 rozstępami  $IQR$  to obserwacje odstające.

**Zad.6** Średnie i odchylenia standardowe dla czasu wykonywania gestu (w ms) dla kolejnych osób zestawiono w tabeli:

	Osoba 1	Osoba 2	Osoba 3	Osoba 4	Osoba 5	Osoba 6
średnia	200	98	143	290	439	206
odchylenie	22	7	23	18	41	19

Narysuj wykresy średnia-wąsy umożliwiające porównanie tych parametrów dla wszystkich osób. Przyjmij, że zarówno dolny jak i górny wąs ma długość odchylenia standardowego.