Statystyki opisowe

Agnieszka Murkowska, klasa II, IV Liceum Ogólnokształcące im. Komisji Edukacji Narodowej w Elblągu, pod kierunkiem mgr Joanny Gronowskiej

Przybliżenia

Błąd bezwzględny – różnica między dokładną wartością a wartością przybliżoną. Obliczamy go odejmując od większej wartości – wartość mniejszą.

Istnieją dwa rodzaje błędów bezwzględnych:

- gdy przybliżenie jest mniejsze od dokładnej wartości (np. 18,95 ≈ 18), mówimy, że jest to *przybliżenie z niedomiarem*, w takiej sytuacji obliczamy to odejmując przybliżenie od dokładnej wartości ($|18{,}95 - 18| = 0{,}95$)
- gdy przybliżenie jest większe od dokładnej wartości (np. 18,95 ≈ 19), mówimy, że jest to *przybliżenie z nadmiarem*, w takiej sytuacji obliczamy to odejmując dokładną wartość od przybliżenia ($|18 - 18{,}95| = 0{,}95$)

Przykład 1.

Jeśli cena jakiejś usługi wynosi np. 18,95 zł, klienci przeważnie zapamiętują ją w przybliżeniu około „18zł", „około 19zł", „około 20zł".
Wszystkie z tych przybliżeń różnią się o pewną kwotę od dokładnej ceny.
Ta kwota to właśnie błąd przybliżenia.

$18{,}95zł \ \approx \ 18zł\ z\ błędem\ 0{,}95zł$

$18{,}95zł \ \approx \ 19zł\ z\ błędem\ 0{,}05zł$

$18{,}95zł \ \approx \ 20zł\ z\ błędem\ 1{,}05zł$

$18{,}95 \ \approx \ 18{,}90zł\ z\ błędem\ 0{,}05zł$

Błąd względny – obliczając stosunek do dokładnej wartości, możemy lepiej ocenić błąd przybliżenia. Stosunek ten nazywamy błędem względnym i zwykle wyrażamy go w procentach.

$$Błąd\ względny\ = \frac{błąd\ bezwzględny\ przybliżenia}{dokładna\ wartość}$$

Przykład 2.

Obliczamy, jaki jest błąd względny tego przybliżenia na przykładzie masy słonia i psa

	Słoń	Pies
masa	5282,4kg	8kg
masa w przybliżeniu	5282kg (przybliżenie z niedomiarem)	8,4 kg (przybliżenie z nadmiarem)
Błąd bezwzględny przybliżenia	0,4kg	0,4 kg
Błąd względny przybliżenia	$\frac{0{,}4}{5282{,}4} \approx 0{,}0076\%$	$\frac{0{,}4}{8} = 5\%$

Średnia arytmetyczna

Średnia arytmetyczna liczb $x_1 + x_2 + \ldots + x_n$ to liczba x obliczana według wzoru:

$$\bar{x} = \frac{x_1 + x_2 + .. + x_n}{n}$$

Przykład 3.

Średnia liczba guzików u dziewcząt to średnia liczba dziesięciu liczb: 0, 1, 1, 2, 2, 2, 4, 4, 5, 6

Wynosi ona:

$$\bar{x} = \frac{0 + 1 * 2 + 2 * 3 + 4 * 2 + 5 + 6}{10} = \frac{27}{10} = 2{,}7$$

Średnia liczba guzików u chłopców to średnia liczba dziesięciu liczb: 1,1, 1, 1, 2, 2, 2, 4, 4, 4.

$$\bar{x} = \frac{1 \cdot 4 + 3 \cdot 2 + 4 \cdot 3}{10} = \frac{22}{10} = 2{,}2$$

Średnia liczba guzików dziewcząt i chłopców wynosi:

$$\bar{x} = \frac{27 + 22}{10 + 10} = \frac{49}{20} = 2{,}45$$

Mediana

Niech $a_1, a_2, \ldots, a_n$ oznacza ciąg liczb, w którym każda następna liczba jest nie mniejsza od poprzedniej, czyli dla dowolnych dwóch kolejnych liczb a_n, a_{n+1} spełniony jest warunek $a_n \leq a_{n+1}$ (chodzi o to żeby liczby były ułożone w kolejności rosnącej).

Gdy *n* (liczba wyrazów w ciągu) jest liczbą nieparzystą, to medianą liczb $a_1, a_2, \ldots, a_n$ nazywamy środkowy wyraz w tym ciągu.

$$\cancel{1}, \cancel{2}, \cancel{3}, \cancel{4}, 5, \cancel{6}, \cancel{7}, \cancel{8}, \cancel{9}$$

W powyższym zestawieniu liczb mediana jest równa 5

Gdy *n* jest liczbą parzystą, to medianą liczb $a_1, a_2, \ldots, a_n$ nazywamy średnią arytmetyczną dwóch środkowych wyrazów tego ciągu.

$$\cancel{1}, \cancel{2}, \cancel{3}, \cancel{4}, 5, 6, \cancel{7}, \cancel{8}, \cancel{9}, \cancel{10}$$

i jest równa średniej arytmetycznej liczb 5 i 6, czyli $\frac{5+6}{2}$=5,5

Dominanta

Dominanta zestawu danych to wartość, która w danym zestawie liczb występuje najczęściej. Jeśli w zestawie kilka wartości występuje z tą samą (najwyższą) częścią, to każda z tych wartości jest dominantą.

Jeśli wszystkie wartości w zestawie występują z tą samą częstotliwością, to przyjmujemy, że zestaw nie ma dominanty.

Przykład 4.

5, 2, 8, 2, 9, 2, 4, 6

dominanta wynosi 2

1, 2, 5, 6, 8, 1, 2, 6

dominanty są równe 1, 2, 6

Średnia ważona

Gdy obliczamy średnią arytmetyczną liczb, wszystkie te liczby traktujemy jednakowo (żadna z nich nie jest wyróżniona).

Czasami jednak do niektórych danych przywiązujemy inną wagę i chcemy to uwzględnić przy obliczaniu wartości przeciętnej.

Wzór:

$$\bar{x}_{\omega} = \frac{a_1x_1 + a_2x_2 + \ldots + a_nx_n}{a_1 + a_2 + a_n}$$

Przykład 5.

Oblicz średnią ważoną liczb 3, 8, 8, 8, 32, 34, 35, jeśli liczby parzyste – mają wagę 0,4, a nieparzyste – mają wagę 0,6.

$$\frac{3 \cdot 0{,}6 + 8 \cdot 0{,}4 + 8 \cdot 0{,}4 + 8 \cdot 0{,}4 + 32 \cdot 0{,}4 + 34 \cdot 0{,}4 + 35 \cdot 0{,}6}{0{,}6 + 0{,}4 + 0{,}4 + 0{,}4 + 0{,}4 + 0{,}4 + 0{,}6} = \frac{58{,}8}{3{,}2} = 18{,}38$$

Odchylenie standardowe

Odchyleniem standardowym zestawu danych liczbowych $x_1, x_2, \ldots, x_n$ nazywamy liczbę *S*, którą obliczamy ze wzoru:

$$S = \sqrt{\frac{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2}{n}}$$

Przykład 6.

Oblicz średnią arytmetyczną i odchylenie standardowe temperatury w Morzu Bałtyckim obserwowane przez 12 miesięcy.

I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
2,7	1,7	1,9	3,7	7,5	12,8	15,8	15,9	14,2	11,5	8,9	4,9

$$\bar{x} = \frac{2{,}7 + 1{,}7 + 1{,}9 + 3{,}7 + 7{,}5 + 12{,}8 + 15{,}8 + 15{,}9 + 14{,}2 + 11{,}5 + 8{,}9 + 4{,}9}{12} = \frac{101{,}5}{12} = 8{,}46$$

Dla wygody obliczymy przede wszystkim S^2 (nazywane wariancją), a następnie przez policzenie pierwiastka otrzymamy odchylenie standardowe *S*:

$$S^2 = \frac{(2{,}7 - 8{,}46)^2 + (1{,}7 - 8{,}46)^2 + (1{,}9 - 8{,}46)^2 + (3{,}7 - 8{,}46)^2 + (7{,}5 - 8{,}46)^2 + (12{,}8 - 8{,}46)^2}{12} +$$

$$+ \frac{(15{,}8 - 8{,}46)^2 + (15{,}9 - 8{,}46)^2 + (14{,}2 - 8{,}46)^2 + (11{,}5 - 8{,}46)^2 + (8{,}9 - 8{,}46)^2 + (4{,}9 - 8{,}46)^2}{12} =$$

$$= \frac{328,08}{12} = 27,34$$

Stąd otrzymujemy $S = \sqrt{27,34} = 5,23$.