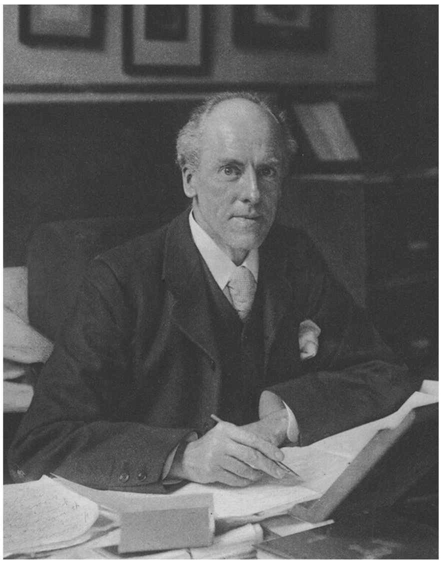
Odchylenie standardowe

Źródło: dostępny w internecie: www.pixabay.com.


Karl Pearsone
Źródło: dostępny w internecie: commons.wikimedia.org, domena publiczna.

Znamy już wariancję – jedną z miar rozproszenia. Nie jest to jednak najlepszy środek do wnioskowania, bowiem podnoszenie do kwadratu odchyleń liczb od średniej powoduje, że rozrzut określany jest w kwadratowych jednostkach pomiaru.

Z tego powodu do analizy rozrzutu wartości jakiejś wielkości (np. inflacji, kursu akcji) wokół średniej, wykorzystuje się odchylenie standardowe. Pojęcie to zostało wprowadzone stosunkowo niedawno, bo w 1894 r. Wprowadził je angielski matematyk, prekursor statystyki Karl Pearson.

Warto wiedzieć, że w 1911 r. Pearson utworzył w Londynie pierwszy na świecie uniwersytecki wydział statystyki.

Twoje cele

- Obliczysz odchylenie standardowe danych przedstawionych w różny sposób.
- Przeanalizujesz i zinterpretujesz odchylenie standardowe danego zestawu danych statystycznych.

Przeczytaj

Odchylenie standardowe – co to takiego?

Odchylenie standardowe jest najczęściej stosowaną miarą rozproszenia. Jest miarą określającą przeciętne zróżnicowanie poszczególnych wartości cechy statystycznej od poziomu średniej arytmetycznej. Odchylenie standardowe to pierwiastek kwadratowy ze średniej arytmetycznej kwadratów odchyleń poszczególnych wartości cechy od wartości średniej arytmetycznej.

Definicja: Odchylenie standardowe

Odchyleniem standardowym zestawu danych statystycznych $x_1,\ x_2,\ \ldots,\ x_n$ od średniej arytmetycznej $\bar{x}$ nazywamy liczbę σ równą pierwiastkowi kwadratowemu z wariancji.

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

Podstawowe własności

Odchylenie standardowe:

- to miara mianowana – ma miano takie, jak badana cecha statystyczna,
- jest liczone na podstawie wszystkich obserwacji,
- bazuje na średniej arytmetycznej, a więc nie może być wyznaczone w szeregach, w których nie można wyznaczyć średniej,
- określa miarę rozrzutu jednej zbiorowości pod względem jednej cechy,
- im ma wyższą wartość, tym bardziej zróżnicowana jest badana zbiorowość statystyczna.

Wyznaczanie odchylenia standardowego w szeregach szczegółowych

Przykład 1

Zbadano liczbę czekoladek w pudełkach z napisem „zawartość 500 g”. Otrzymano wyniki: 20, 21, 23, 19, 17. Obliczymy odchylenie standardowe tych danych.

Rozwiązanie:

Porządkujemy dane według rosnących wartości.

x_1	x_2	x_3	x_4	x_5
17	19	20	21	23

Określamy liczbę danych: $n = 5$.

Obliczamy średnią arytmetyczną.

$$\bar{x} = \frac{17+19+20+21+23}{5} = 20$$

Obliczamy odchylenia od średniej dla każdej z danych.

$|17 - 20| = 3$

$|19 - 20| = 1$

$|20 - 20| = 0$

$|21 - 20| = 1$

$|23 - 20| = 3$

Obliczamy odchylenie standardowe, podstawiając do wzoru wyznaczone odchylenia od średniej.

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

$$\sigma = \sqrt{\frac{3^2+1^2+0^2+1^2+3^2}{5}}$$

$$\sigma = \sqrt{\frac{20}{5}} = 2$$

Odpowiedź:

Odchylenie standardowe jest równe 2, co oznacza, że liczba czekoladek w pudełkach różni się przeciętnie od średniej o 2 czekoladki.

Przykład 2

Zbadano cenę pączków w kilku sklepach. Otrzymano następujące wyniki: 2 zł, $2,5$ zł, 3 zł, $1,5$ zł, 4 zł, 2 zł. Obliczymy odchylenie standardowe ceny pączków od średniej.

Rozwiązanie:

Porządkujemy dane: $1,5$ zł, 2 zł, 2 zł, $2,5$ zł, 3 zł, 4 zł.

Obliczamy średnią arytmetyczną cen.

$$\bar{x} = \frac{1{,}50+2+2+2{,}50+3+4}{6}$$

$$\bar{x} = \frac{15}{6}$$

$$\bar{x} = 2,5 \text{ zł}$$

Obliczamy odchylenie od średniej dla każdej z danych.

$$|1,50 - 2,50| = 1$$

$$|2 - 2,50| = 0,50$$

$$|2,50 - 2,50| = 0$$

$$|3 - 2,50| = 0,50$$

$$|4 - 2,50| = 1,50$$

Obliczamy odchylenie standardowe.

$$\sigma = \sqrt{\frac{1^2+(0{,}50)^2+(0{,}50)^2+0^2+(0{,}5)^2+(1{,}5)^2}{6}}$$

$$\sigma = \sqrt{\frac{4}{6}} = \frac{2}{\sqrt{6}} = \frac{\sqrt{6}}{3}$$

$$\sigma \approx 0{,}82 \ zł$$

Odpowiedź:

Odchylenie standardowe jest równe w przybliżeniu $0{,}82$ *zł*, co oznacza, że ceny różnią się przeciętnie o $0{,}82$ *zł* od średniej ceny.

Wyznaczanie odchylenia standardowego w szeregach rozdzielczych punktowych

Przykład 3

W tabeli zapisano dane na temat wieku uczniów. Obliczymy odchylenie standardowe wieku uczniów od średniej.

Wiek uczniów (w latach)	15	16	17	18
Liczba uczniów	2	1	4	3

Rozwiązanie:

Określamy liczbę uczniów.

$2 + 1 + 4 + 3 = 10$

Obliczamy średnią arytmetyczną wieku.

$$\bar{x} = \frac{2 \cdot 15 + 1 \cdot 16 + 4 \cdot 17 + 3 \cdot 18}{10}$$

$$\bar{x} = \frac{168}{10} = 16,8$$

Obliczamy odchylenie od średniej dla każdej z wartości danych.

$$|15 - 16,8| = 1,8$$

$$|16 - 16,8| = 0,8$$

$$|17 - 16,8| = 0,2$$

$$|18 - 16,8| = 1,2$$

Obliczamy odchylenie standardowe.

$$\sigma = \sqrt{\frac{2 \cdot (1,8)^2 + 1 \cdot (0,8)^2 + 4 \cdot (0,2)^2 + 3 \cdot (1,2)^2}{10}}$$

$$\sigma = \sqrt{\frac{11,6}{10}} = \sqrt{1,16} \approx 1$$

Odpowiedź:

Odchylenie standardowe od średniej wieku uczniów jest równe w przybliżeniu 1. Oznacza to, że wiek uczniów różni się od średniej wieku mniej więcej o rok.

Odchylenie standardowe można wykorzystać do porównywania parametrów statystycznych danych liczbowych dotyczących tych samych cech w kilku zbiorowościach statystycznych.

Przykład 4

W tabeli przedstawiono dane dotyczące wzrostu (z dokładnością do 5 cm) dwóch grup uczniów. Obliczymy średnią arytmetyczną i odchylenie standardowe wzrostu w każdej z grup. Ocenimy, w której grupie zróżnicowanie wzrostu jest mniejsze.

Wzrost (w cm)	160	165	170	175	180
Grupa 1 Liczba uczniów	1	2	1	2	4
Grupa 2 Liczba uczniów	1	2	4	2	1

Obliczamy średnią arytmetyczną wzrostu.

Grupa 1

$$\bar{x} = \frac{1 \cdot 160 + 2 \cdot 165 + 1 \cdot 170 + 2 \cdot 175 + 4 \cdot 180}{10}$$

$$\bar{x} = \frac{1730}{10}$$

$$\bar{x} = 173 \text{ cm}$$

Grupa 2

$$\bar{x} = \frac{1 \cdot 160 + 2 \cdot 165 + 4 \cdot 170 + 2 \cdot 175 + 1 \cdot 180}{10}$$

$$\bar{x} = \frac{1700}{10}$$

$$\bar{x} = 170 \text{ cm}$$

Obliczamy odchylenie standardowe.

Grupa 1

$$\sigma = \sqrt{\frac{1 \cdot (160-173)^2 + 2 \cdot (165-173)^2 + 1 \cdot (170-173)^2 + 2 \cdot (175-173)^2 + 4 \cdot (180-173)^2}{10}}$$

$$\sigma = \sqrt{\frac{510}{10}} \approx 7,14$$

$$\sigma \approx 7,14 \text{ cm}$$

Grupa 2

$$\sigma = \sqrt{\frac{1 \cdot (160-170)^2 + 2 \cdot (165-170)^2 + 4 \cdot (170-170)^2 + 2 \cdot (175-170)^2 + 1 \cdot (180-170)^2}{10}}$$

$$\sigma = \sqrt{\frac{300}{10}} \approx 5,48$$

$$\sigma \approx 5,48 \text{ cm}$$

Odpowiedź:

Odchylenie standardowe w grupie drugiej jest znacznie mniejsze niż w pierwszej – zróżnicowanie wzrostu w grupie drugiej jest mniejsze niż w grupie pierwszej.

Słownik

odchylenie standardowe

odchyleniem standardowym zestawu danych statystycznych od średniej arytmetycznej nazywamy liczbę równą pierwiastkowi kwadratowemu z wariancji

Animacja

Polecenie 1

Przeanalizuj przykłady wyznaczania odchylenia standardowego. Zastanów się, w jaki sposób można zinterpretować wyniki obliczeń.

Trwa wczytywanie danych...

Film dostępny pod adresem https://zpe.gov.pl/a/Dr0HHdz4a

Film nawiązujący do treści materiału dotyczącego zagadnienia odchylenia standardowego.

Polecenie 2

Korzystające ze wzoru zapisanego w animacji, oblicz odchylenie standardowe zestawu danych: 0, 2, 6, 8. Wynik podaj z dokładnością do $0,1$.

Sprawdź się

Pokaż ćwiczenia:

Ćwiczenie 1

Ćwiczenie 2

Ćwiczenie 3

Ćwiczenie 4

Ćwiczenie 5

Ćwiczenie 6

Przeanalizowano liczbę uczniów nieobecnych w dwóch klasach w marcu. Wyniki przedstawiono w tabelce.

Liczba dni nieobecności	1	2	4	5
Liczba uczniów klasa III A	8	5	3	4
Liczba uczniów Klasa III B	2	3	9	6

Ćwiczenie 7

Oblicz odchylenie standardowe danych zapisanych w tabelce. Wynik zaokrąglij do całości.

Wartość x_i cechy	1	2	4	9
Liczebność n_i	6	1	1	2

Ćwiczenie 8

Anka otrzymała za rozwiązania trzech zadań po 4 punkty, za rozwiązanie dwóch zadań po 5 punktów, za rozwiązania czterech zadań po 3 punkty i za jedno zadanie 6 punktów. Oblicz wariancję i odchylenie standardowe dla tych danych.

Dla nauczyciela

Autor: Justyna Cybulska

Przedmiot: Matematyka

Temat: Odchylenie standardowe

Grupa docelowa: III etap edukacyjny, liceum, technikum, zakres rozszerzony

Podstawa programowa:

XII. Rachunek prawdopodobieństwa i statystyka. Zakres podstawowy.

Uczeń:

3) oblicza średnią arytmetyczną i średnią ważoną, znajduje medianę i dominantę;
4) oblicza odchylenie standardowe zestawu danych (także w przypadku danych odpowiednio pogrupowanych), interpretuje ten parametr dla danych empirycznych.

Kształtowane kompetencje kluczowe:

- kompetencje w zakresie rozumienia i tworzenia informacji
- kompetencje matematyczne oraz kompetencje w zakresie nauk przyrodniczych, technologii i inżynierii
- kompetencje cyfrowe
- kompetencje osobiste, społeczne i w zakresie umiejętności uczenia się

Cele operacyjne:

Uczeń:

- oblicza odchylenie standardowe danych przedstawionych w postaci szeregów szczegółowych i w postaci szeregów rozdzielczych punktowych
- analizuje i interpretuje odchylenie standardowe danego zestawu danych statystycznych
- dostrzega podobieństwa oraz analogie, formułuje wnioski i na ich podstawie porównuje dane dwóch zbiorowości statystycznych
- dobiera i tworzy modele matematyczne przy rozwiązywaniu problemów teoretycznych i praktycznych

Strategie nauczania:

- konstruktywizm

Metody i techniki nauczania:

- burza mózgów
- tworzenie przez analogię
- drzewo statystyczne

Formy pracy:

- praca indywidualna
- praca w grupach
- praca całego zespołu klasowego

Środki dydaktyczne:

- komputery z dostępem do Internetu w takiej liczbie, żeby każdy uczeń miał do dyspozycji komputer

Przebieg lekcji

Faza wstępna:

Uczniowie metodą burzy mózgów przypominają poznane pojęcia statystyczne, ze szczególnym zwróceniem uwagi na miary rozproszenia.

Wspólnie zastanawiają się czy wariancja jest wystarczającą miarą, pozwalająca na określenie stopnia rozproszenia danych wokół średniej.

Nauczyciel podaje temat i cele zajęć oraz kryteria sukcesów.

Faza realizacyjna:

Uczniowie w grupach zapoznają się z określeniem odchylenia standardowego i metodą – tworzenie przez analogię – opracowują algorytm obliczania odchylenia standardowego dla danych zapisanych w postaci szeregu szczegółowego. A następnie porównują z odpowiednimi treściami zapisanymi w sekcji „Przeczytaj”.

W podobny sposób postępują przy opracowywaniu algorytmu obliczeń dla danych zapisanych w postaci szeregu rozdzielczego.

Uczniowie zapoznają się z animacją i tworzą drzewko statystyczne, na którym graficznie obrazują uzyskane informacje i powiązania między miarami rozproszenia.

Faza podsumowująca:

Uczniowie indywidualnie wykonują proponowane ćwiczenia interaktywne.

Wybrani uczniowie podsumowują zajęcia – omawiają pracę swoich grup, wskazują na problemy i uzyskane umiejętności.

Nauczyciel wskazuje mocne i słabe strony pracy uczniów, ocenia pracę grup.

Praca domowa:

Zadaniem uczniów jest zebranie danych na temat wzrostu członków swojej rodziny i na podstawie zebranych danych, opisanie wyników za pomocą wszystkich poznanych miar rozproszenia.

Materiały pomocnicze:

Miary rozproszenia

Wskazówki metodyczne:

Animację i materiały z sekcji „Przeczytaj" uczniowie mogą przeanalizować w domu i wtedy zajęcia mogą być przeprowadzone metodą odwróconej klasy.