

Podstawy analizy danych

Ćwiczenie 2: Mierniki statystyczne

Program ćwiczeń obejmuje następująca zadania:

1. Na produkcję przedsiębiorstwa wytwarzającego odkuwki składa się produkcja pięciu jego oddziałów. Każdy z tych oddziałów wyprodukował taką samą liczbę odkuwek, jednak koszt jednostkowy w każdym z tych oddziałów jest inny i odpowiednio wynosi 17, 18, 17.5, 18.5, 17.5 tys. zł. Jaki jest średni koszt jednostkowy odkuwki w tym przedsiębiorstwie?
2. Stan zatrudnienia w pewnym przedsiębiorstwie na dzień 31 grudnia kształtował się następująco:

rok	liczba osób
2009	1270
2010	1321
2011	1400
2012	1526
2013	1587
2014	1666

Wyznaczyć wskaźniki wzrostu zatrudnienia w ciągu roku dla lat 2010–2014, a następnie wyznaczyć średni wskaźnik wzrostu zatrudnienia dla tego okresu.

3. Chromonikielina Glowray jest stopem oporowym zawierającym 65% niklu o gęstości $\mu_1 = 8.90 \text{ g/cm}^3$, 20% żelaza o gęstości $\mu_2 = 7.87 \text{ g/cm}^3$ i 15% chromu o gęstości $\mu_3 = 7.20 \text{ g/cm}^3$. Wyrazić gęstość μ chromonikieliny jako pewną średnią gęstości μ_1, μ_2, μ_3 , a następnie obliczyć wartość μ .
4. Trzy oddziały pewnej firmy produkują te same urządzenia. Wartość miesięcznej produkcji w każdym z tych oddziałów jest taka sama, jednak cena jednostkowa urządzenia wyprodukowanego przez każdy z tych oddziałów jest inna i odpowiednio wynosi 2565 zł, 2250 zł i 2406 zł. Jaka jest średnia cena jednego urządzenia wyprodukowanego przez tę firmę?
5. Pokazać, że dla średniej arytmetycznej zachodzi wzór

$$\sum_{i=1}^n (x_i - \bar{x}) = 0$$

6. Strukturę rodzin wg liczby członków w pewnej miejscowości charakteryzuje rozkład

liczba członków	2	3	4	5	6	7	8
odsetek rodzin	15	30	20	15	10	5	5

Za pomocą miar przeciętnych scharakteryzować ten rozkład. Wyniki zinterpretować.

7. W pewnym zakładzie zbadano pracowników bezpośrednio produkcyjnych pod względem stażu pracy. Okazało się, że 25% tych pracowników pracowało poniżej 6 lat, połowa od 6 do 12 lat, natomiast wśród pozostałych najdłuższy staż wyniósł 18 lat. Średni staż pracy pracowników inżynieryjno-technicznych w tym zakładzie wynosił 12 lat. Jaki był średni staż pracy ogółu pracowników tego zakładu, jeśli wiadomo, że grupa pracowników bezpośrednio produkcyjnych była 2.5-krotnie liczniejsza od inżynieryjno-technicznych?

8. W szeregach rozdzielczych przedziałowych bezpośrednio można określić tylko przedział, w którym znajduje się modalna (jest to przedział o największej liczebności). Konkretną wartość liczbową należącą do tego przedziału i będącą dobrym przybliżeniem modalnej wyznacza się często wg wzoru

$$Mo \approx x_m + \frac{n_m - n_{m-1}}{(n_m - n_{m-1}) + (n_m - n_{m+1})} h_m,$$

gdzie: x_m – dolna granica przedziału, w którym znajduje się modalna, n_m – liczebność przedziału modalnej, n_{m-1} – liczebność przedziału poprzedzającego przedział modalnej, n_{m+1} – liczebność przedziału następującego po przedziale modalnej, h_m – długość przedziału modalnej. Uzasadnić poprawność tego wzoru.

9. W spółdzielni mieszkaniowej dominują mieszkania o powierzchni 45 – 50 m². Na podstawie badań statystycznych ustalono, że najwięcej mieszkań (30%) miało powierzchnię 48 m², a 20% mieszkań miało powierzchnię od 40 do 45 m². Jaki odsetek mieszkań ma powierzchnię od 50 do 55 m²?

10. Udowodnić, że wariancję

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

można równoważnie wyznaczyć ze wzoru

$$s^2 = \left(\frac{1}{n} \sum_{i=1}^n x_i^2 \right) - \bar{x}^2.$$

Wyznaczyć liczby operacji potrzebnych do wykorzystania obydwu wzorów.

11. Lekkoatleta *A* uzyskał w skoku w dal następujące wyniki na zawodach w całym sezonie (w m): 6.82, 6.96, 7.23, 7.05, 7.80, 7.75. Lekkoatleta *B*, startujący w tych samych zawodach, uzyskał takie wyniki, że ich średnia arytmetyczna wyniosła 7.5 m, a suma ich kwadratów 450.2592 m². Który z tych lekkoatletów osiągał regularniejsze wyniki?

12. Dzielne zużycie energii elektrycznej (w kWh) w pewnym bloku mieszkalnym kształtowało się następująco:

zużycie	2-4	4-6	6-8	8-10	10-12	12-14
liczba rodzin	6	10	30	40	10	4

Obliczyć średnią arytmetyczną, medianę, modalną, współczynnik zmienności oraz wyznaczyć obszar zmienności.