

Raport

Porównanie wynagrodzeń kobiet i mężczyzn w
Polsce w maju 2024 r.

1 Metryki statystyczne

Poniżej znajdują się wzory i definicje metryk, które zastosujemy oraz wyjaśnienie, co one oznaczają w kontekście analizowanych danych.

1.1 Miary położenia

1.1.1 Średnia arytmetyczna

Średnią arytmetyczną wyrażamy wzorem

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

1.1.2 Średnia harmoniczna

Średnią harmoniczną wyrażamy wzorem

$$\bar{x}_h = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}.$$

1.1.3 Średnia geometryczna

Średnią geometryczną wyrażamy wzorem

$$\bar{x}_g = \sqrt[n]{\prod_{i=1}^n x_i}.$$

1.1.4 Średnia ucinana

Średnią ucinaną wyrażamy wzorem

$$\bar{x}_u = \frac{1}{n - 2k} \sum_{i=k+1}^{n-k} x_i.$$

1.1.5 Średnia winsorowska

Średnią winsorowską wyrażamy wzorem

$$\bar{x}_w = \frac{1}{n} \left[(k+1)x_{k+1} + \sum_{i=k+2}^{n-k-1} x_i + (k+1)x_{n-k} \right].$$

1.1.6 Mediana

Medianę z uporządkowanej rosnąco próby x_1, \dots, x_n wyrażamy wzorem

$$\text{Med}(x) = \begin{cases} x_{\frac{n+1}{2}}, & \text{gdy } n \text{ jest liczbą nieparzystą} \\ \frac{1}{2}(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}), & \text{gdy } n \text{ jest liczbą parzystą.} \end{cases}$$

1.1.7 Kwartyle

Kwartyle dzielą zbiór danych na cztery grupy:

- drugi kwartył ($Q2$) to mediana,
- pierwszy kwartył ($Q1$) to mediana grupy obserwacji mniejszych od $Q2$,
- trzeci kwartył ($Q3$) to mediana grupy obserwacji większych od $Q2$.

1.1.8 Dominanta

Dominanta to wartość, która występuje najczęściej w badanej próbie.

1.2 Miary rozproszone

1.2.1 Rozstęp z próby

Rozstęp z uporządkowanej rosnąco próby x_1, \dots, x_n wyrażamy wzorem

$$R = x_n - x_1.$$

1.2.2 Rozstęp międzykwartyłowy

Rozstęp międzykwartyłowy wyrażamy wzorem

$$IQR = Q3 - Q1.$$

1.2.3 Wariancja z próby

Wariancję z próby wyrażamy wzorem

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

1.2.4 Odchylenie standardowe z próby

Odchylenie standardowe z próby wyrażamy wzorem

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

1.2.5 Odchylenie przeciętne od wartości średniej

Odchylenie przeciętne od wartości średniej wyrażamy wzorem

$$d_1 = \frac{1}{n} \sum_{i=1}^n |x_i - \bar{x}|.$$

1.2.6 Współczynnik zmienności

Współczynnik zmienności wyrażamy wzorem

$$\nu = \frac{s}{\bar{x}} \cdot 100\%.$$

1.3 Miary asymetrii i spłaszczenia

1.3.1 Współczynnik skośności

Współczynnik skośności wyrażamy wzorem

$$\alpha = \frac{n}{(n-1)(n-2)} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s} \right)^3.$$

1.3.2 Kurtoza

Kurtozę wyrażamy wzorem

$$K = \frac{n-1}{(n-2)(n-3)} ((n+1)K_1 - 3(n-1)) + 3,$$

gdzie

$$K_1 = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^2}.$$