

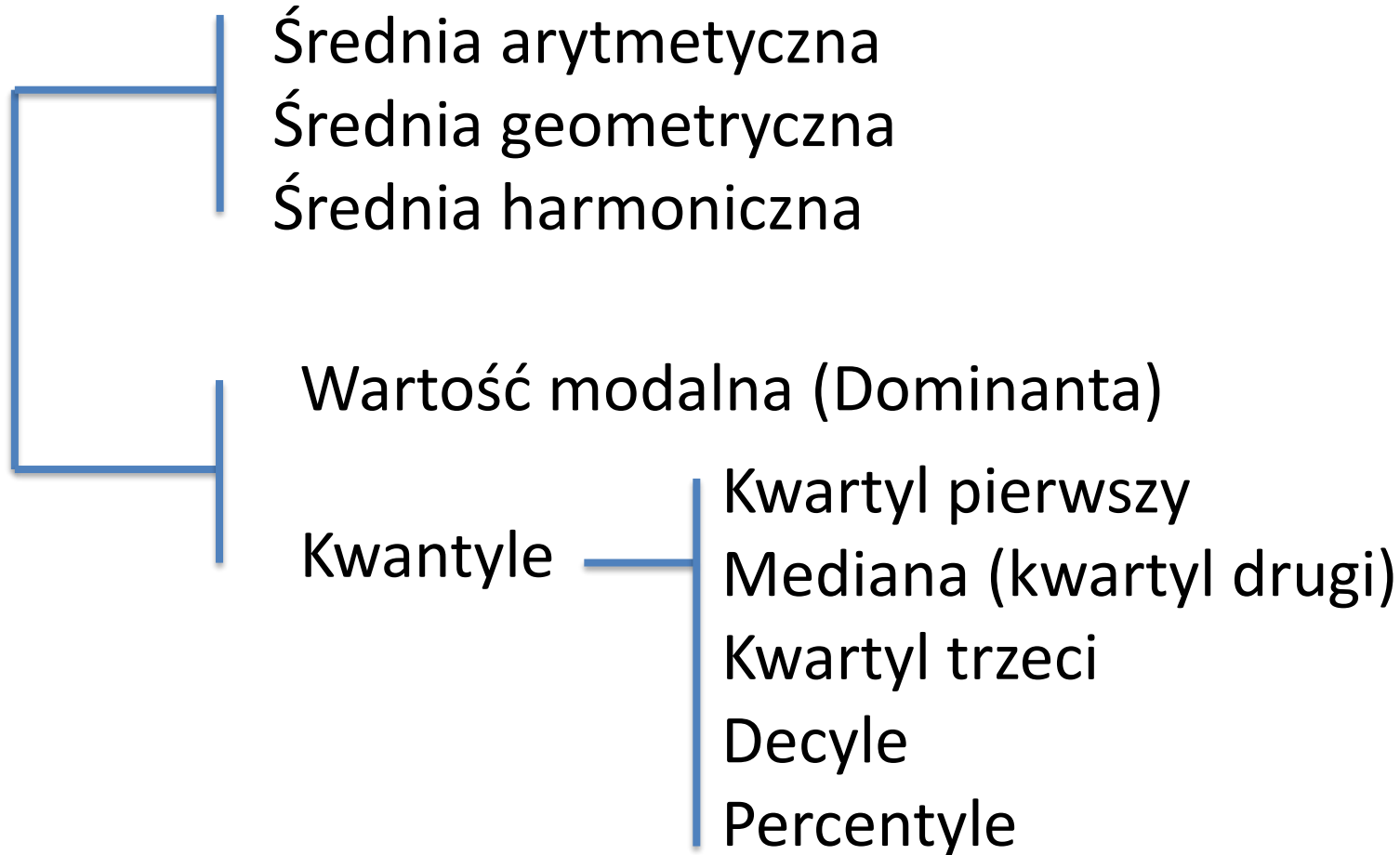
Statystyka opisowa

XIII Lubelski Konkurs Statystyczno-Demograficzny Sigma Kwadrat

Statystyka opisowa:

- Określenie przeciętnego poziomu i rozmieszczenia wartości zmiennej
- Określenie granic obszaru zmienności wartości zmiennej
- Określenie skupienia i spłaszczenia oraz stopnia odejścia od idealnej symetrii

Miary położenia:



Średnia arytmetyczna

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

gdzie x_1, x_2, \dots, x_n ,
ciąg wartości
zmiennnej

Średnia geometryczna

$$\overline{x_g} = \sqrt[n]{x_1 * x_2 * \dots * x_n}$$

Średnia harmoniczna

$$\overline{x_h} = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$$

Przykłady

Średnia arytmetyczna

Średni wzrost 4 ludzi w cm:

$$\bar{x} = \frac{150+155+160+180}{4} = 161,25$$

Przykłady

Średnia geometryczna – gdy zjawiska ujmowane są dynamicznie

$$\text{Dynamika zjawiska} = \frac{\text{wielkość zjawiska w okresie } t+1}{\text{wielkość zjawiska w okresie } t} * 100$$

$$\text{Dynamika produkcji} = \frac{2\,000\,000}{1\,500\,000} * 100 = 133,3$$

$$\text{Względna zmiana liczby ludności} = \frac{10\,000}{10\,500} * 100 = 95,2$$

Średnia harmoniczna

Stosowana do obliczania średnich z wartości względnych

Przykład

Przyrost naturalny na 1000 ludności w miastach:

A – 2, B – 3, C - 4

Średnia harmoniczna dla tych 3 miast:

$$X_{\text{śr}} = \frac{3}{\frac{1}{2} + \frac{1}{3} + \frac{1}{4}} = 2,77$$

Wartość modalna (dominanta)

Wartość występująca w zbiorze najczęściej

Wzrost koszykarzy	200	201	202	203	204	205	206	207	208
Liczba koszykarzy	2	3	5	8	12	11	4	4	1

W przedziałach:

Wynik testu	8-12	13-17	18-22	23-27	28-32	33-37	38-42
Liczba studentów	4	29	39	81	35	9	4

Kwantyle

Wartości cechy badanej zbiorowości przedstawionej w postaci szeregu statystycznego, które dzielą zbiorowość na określone części pod względem liczby jednostek. Części te pozostają względem siebie w określonych proporcjach.

Najczęściej stosowane kwantyle to:

Kwartyle (podział na 4 części)

Decyle (podział na 10 części)

Percentyle (podział na 100 części)

Kwartyle

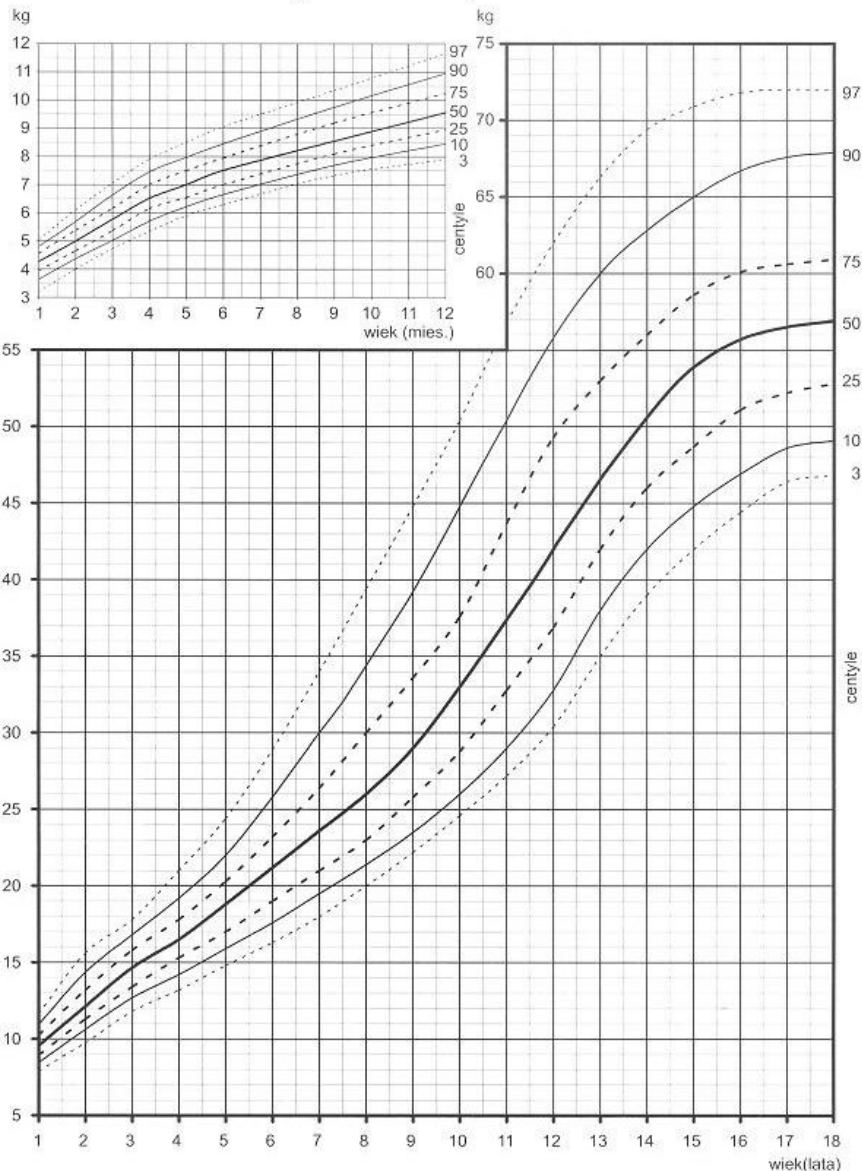
Kwartyl pierwszy (kwartyl dolny) jest to wartość, poniżej której znajduje się 25 proc. obserwacji w danej zbiorowości

Kwartyl drugi (mediana), wartość która dzieli daną zbiorowość na dwie równe części

Kwartyl trzeci (kwartyl górny) jest to wartość, poniżej której znajduje się 75 proc. obserwacji w danej zbiorowości

Percentyle

Siatka centylowa masy ciała dziewcząt



Miary zmienności

Wariancja – w zbiorze wyników obserwacji jest to przeciętne kwadratowe odchylenie poszczególnych wyników od ich średniej

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

Odchylenie standardowe – pierwiastek kwadratowy z wariancji

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

Miary asymetrii i koncentracji

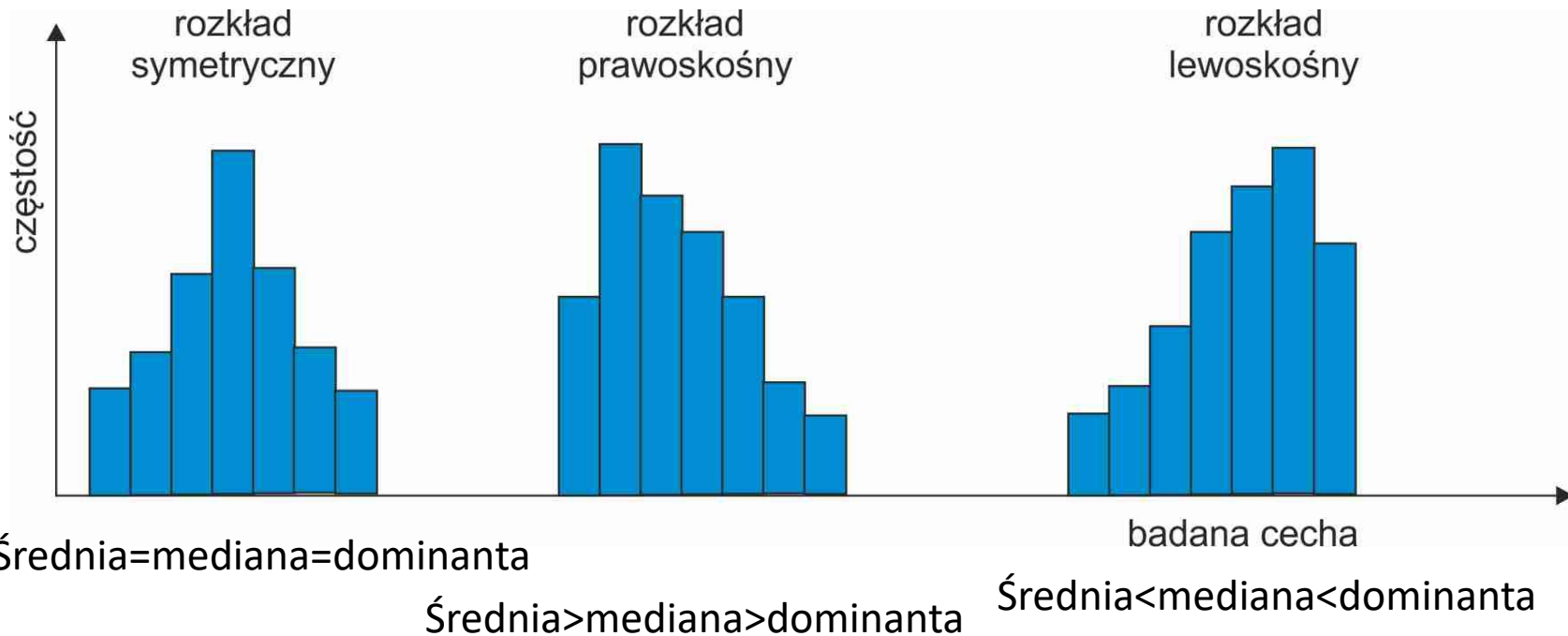
Asymetria (Skośność)

Współczynnik skośności

=0

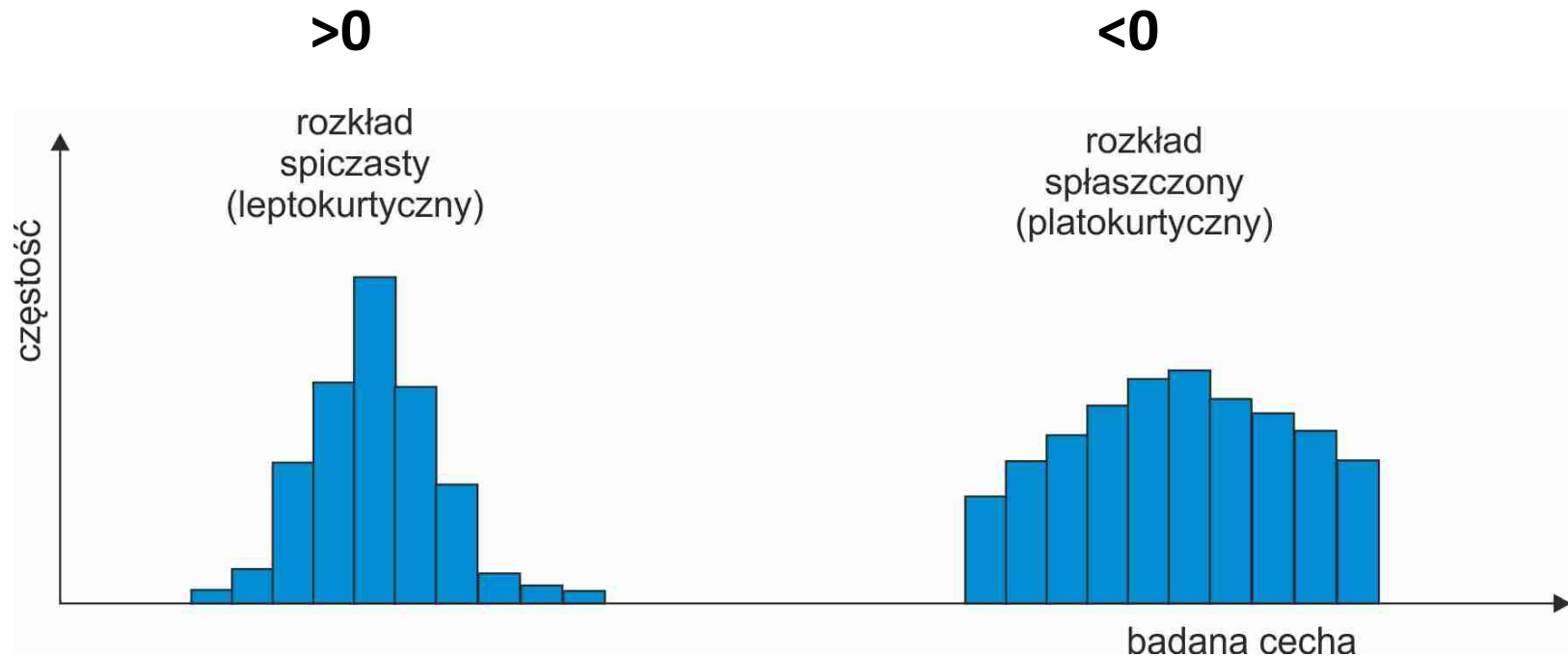
>0

<0



Miary asymetrii i koncentracji

Koncentracja (kurtoza)





Lubelski Ośrodek Badań Regionalnych

Urząd Statystyczny w Lublinie

