Analiza struktury - badanie asymetrii i koncentracji

Miary asymetrii

Zdarza się, że badanie średniego poziomu cechy i rozproszenia jej wartości nie wskazuje na istnienie różnic między porównywanymi zbiorowościami.

Okazuje się bowiem, że istotne są przeciętny poziom i wewnętrzne zróżnicowanie cechy, ale także to czy przeważająca liczba badanych jednostek ma wartości cechy powyżej czy też poniżej przeciętnego poziomu cechy.

Miary asymetrii

Miary asymetrii (skośności) określają kierunek rozkładu cech zmiennych w **zbiorowości** (rozkład może być symetryczny lub asymetryczny lewostronnie lub prawostronnie) oraz **stopień odchylenia rozkładu cechy zmiennej** od rozkładu symetrycznego.

Asymetria oznacza deformację rozkładu cechy zmiennej w związku z wydłużeniem ramienia krzywej liczebności w prawo lub w lewo w stosunku do dominanty.

Im asymetria rozkładu jest większa, tym mniejsza jest wartość poznawcza średniej arytmetycznej oraz pozostałych miar klasycznych, i odwrotnie.

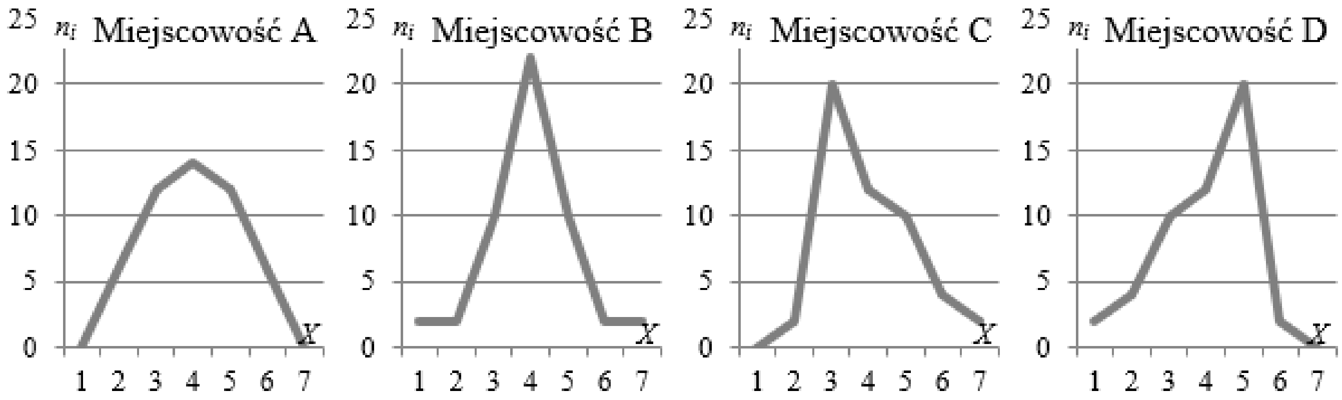
Przykład. Znaczenie konieczności przeprowadzania kompleksowej analizy zjawisk pokazuje przykład liczby zatrudnionych w małych sklepach spożywczych w czterech miejscowościach (dane umowne za W. Krysicki i inni, s. 28):

Liczba zatrudnionych x_i	Liczba sklepów n_i			
	Miejscowość A	Miejscowość B	Miejscowość C	Miejscowość D
1	0	2	0	2
2	6	2	2	4
3	12	10	20	10
4	14	22	12	12
5	12	10	10	20
6	6	2	4	2
7	0	2	2	0
-	50	50	50	50

Z każdej z czterech miejscowości wybrano po 50 sklepów. Dane ułożono w postaci szeregów jednowariantowych.

Wyznaczając wartości średniej arytmetycznej, odchylenia standardowego, współczynnika asymetrii oraz współczynnika koncentracji uzyskano rezultaty:

Miejscowość A:	$\bar{x} = 4$;	$S(x) = 1{,}2$;	$A = 0$;	$W_K = 2{,}08$.
Miejscowość B:	$\bar{x} = 4$;	$S(x) = 1{,}2$;	$A = 0$;	$W_K = 3{,}94$.
Miejscowość C:	$\bar{x} = 4$;	$S(x) = 1{,}2$;	$A = 0{,}69$;	$W_K = 2{,}78$.
Miejscowość D:	$\bar{x} = 4$;	$S(x) = 1{,}2$;	$A = -0{,}69$;	$W_K = 2{,}78$.



Rysunek. Rozkłady liczby zatrudnionych w sklepach spożywczych

Mimo różnych rozkładów liczebności liczby zatrudnionych w sklepach w czterech badanych miejscowościach w każdej stwierdzono taką samą wartość średnią – 4 oraz takie samo zróżnicowanie mierzone odchyleniem standardowym – 1,2. Natomiast rozkłady różnią się asymetrią: pierwszy i drugi są symetryczne, ale trzeci jest prawostronny, a czwarty co do siły taki sam jak trzeci, ale lewostronny. Rozkłady różnią się także koncentracją, która najmniejsza jest w szeregu pierwszym, a największa w szeregu drugim, z kolei szeregi trzeci i czwarty charakteryzują się taką samą koncentracją. Przykład ten pokazuje, że ograniczenie analiz do najbardziej popularnych średniej i zróżnicowania może być niewystarczające, a ważna w ocenie rozkładu jest także asymetria i koncentracja.

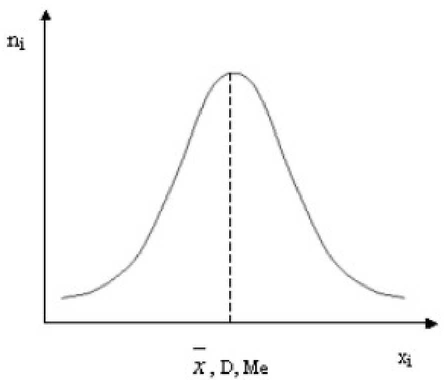
Miary asymetrii

Szereg symetryczny to szereg, w którym liczebności (albo częstości) są rozłożone proporcjonalnie po obu stronach wartości średniej.

Szereg niesymetryczny nie spełnia warunku symetrii. Szereg niesymetryczny będzie miał przewagę niskich albo wysokich wyników.

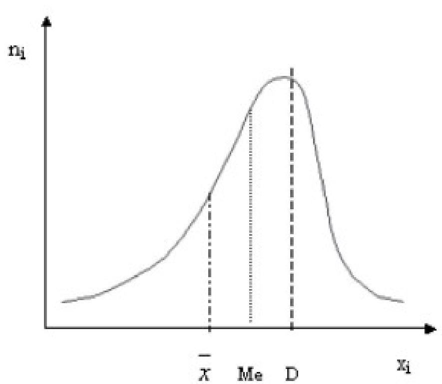
Miary asymetrii

Punktem wyjścia analizy kierunku i siły asymetrii jest wzajemne położenie względem siebie miar średnich – **średniej arytmetycznej,**

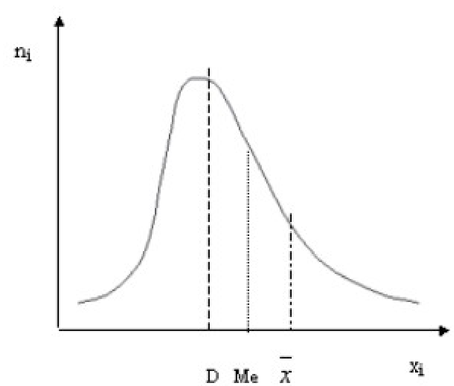


Rozkład symetryczny: $\bar{x} = Me = D$

$\bar{x}$ =Me=D dla rozkładu symetrycznego



Rozkład asymetryczny lewostronnie: $\bar{x} < Me < D$



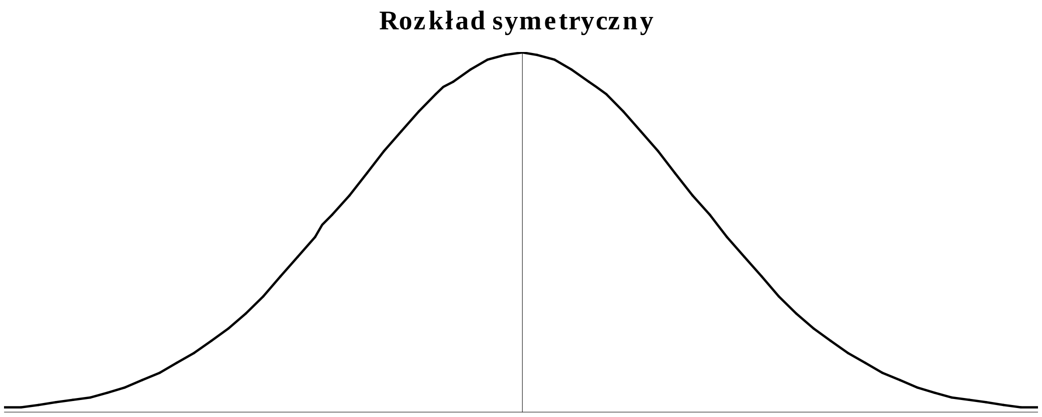
Rozkład asymetryczny prawostronnie: $\bar{x} > Me > D$

$\bar{x}$ >Me>D dla rozkładu o asymetrii prawostronnej

$\bar{x}$ <Me<D dla rozkładu o asymetrii lewostronnej

ANALIZA ASYMETRII

Excel → Skośność



$$\bar{x} = Me = D$$

Rozkład normalny, rozkład Gaussa (w literaturze francuskiej zwany rozkładem Laplace'a-Gaussa) – jeden z najważniejszych rozkładów prawdopodobieństwa, odgrywający ważną rolę w statystyce. Wykres funkcji prawdopodobieństwa tego rozkładu jest krzywą w kształcie dzwonu.

Przyczyną jego znaczenia jest częstość występowania w naturze. Jeśli jakaś wielkość jest sumą lub średnią bardzo wielu drobnych losowych czynników, to niezależnie od rozkładu każdego z tych czynników jej rozkład będzie zbliżony do normalnego (centralne twierdzenie graniczne) – dlatego można go bardzo często zaobserwować w danych.

Rachunek momentów

Podstawowy wskaźnik asymetrii oparty jest na **trzecim momencie centralnym**. Ogólnie moment centralny definiuje się jako:

$$M_l = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^l \text{ , gdzie } l \in N .$$

Pierwszy moment centralny (l=1) jest równy zero:

$$M_1 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}) = 0 .$$

Dzieje się tak ponieważ suma wartości odchyleń obserwacji od średniej arytmetycznej jest równa zero.

Drugi moment centralny (l=2) jest wariancją:

$$M_2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 = S^2(x) .$$

W związku z tym, że kwadraty odchyleń $x_i - \bar{x}$ są dodatnie, drugi moment centralny pozwala na ocenę rozproszenia wyników wokół poziomu przeciętnego.

Natomiast z punktu widzenia asymetrii interesujący jest trzeci moment centralny:

$$M_3 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3 .$$

W szeregach symetrycznych suma trzecich potęg dodatnich i ujemnych odchyleń $x_i - \bar{x}$ jest równa 0, ponieważ, jeżeli każdej wartości $x_i < \bar{x}$ odpowiada wartość $x_j > \bar{x}$ spełniająca warunek :

$$-(x_i - \bar{x}) = \bar{x} - x_j \text{ i } n_i = n_j ,$$

to trzecie potęgi tych odchyleń znoszą się wzajemnie i M_3=0.

Miary asymetrii

Natomiast w przypadku występowania asymetrii suma trzecich potęg odchyleń dodatnich będzie inna (większa lub mniejsza) od sumy odchyleń ujemnych. Im bardziej wartości x_i odchylać będą się od wartości średniej (na plus lub minus) tym wyższa będzie wartość trzeciej potęgi tego odchylenia. W związku z tym, że trzecia potęga nie zmienia znaku różnicy $x_i - \bar{x}$, to suma takich odchyleń wskazywać będzie kierunek asymetrii. Dlatego:

- $M_3 = 0$ oznacza rozkład symetryczny;
- $M_3 > 0$ rozkład o asymetrii dodatniej (rozkład prawostronny);
- $M_3 < 0$ rozkład o asymetrii ujemnej (rozkład lewostronny).

Trzeci moment centralny pozwala jedynie na określenie kierunku asymetrii, aby dodatkowo interpretować siłę asymetrii wyznacza się **współczynnik asymetrii (skośności)**, który jest ilorazem trzeciego momentu centralnego i trzeciej potęgi odchylenia standardowego:

Miary asymetrii

- **Współczynnik asymetrii** (skośności):

$$A = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3}{S^3(x)}$$

wartości: <-1 ; 1 >

Dla danych pogrupowanych w szereg jednowariantowy lub przedziałowy przyjmuje postać:

$$A = \frac{\frac{1}{n}\sum_{i=1}^{k}(x_i - \bar{x})^3 n_i}{S^3(x)}$$

A = 0 oznacza rozkład symetryczny

A > 0 rozkład o asymetrii dodatniej (rozkład prawostronny)

A < 0 rozkład o asymetrii ujemnej (rozkład lewostronny)

Miary asymetrii

Odnosząc różnice te do odchylenia standardowego, średniej i dominanty oraz odchylenia ćwiartkowego uzyskuje się **współczynnik asymetrii, zwany współczynnikiem skośności Pearsona (**<-1; 1>)**:**

$$A_{\bar{x}} = \frac{\bar{x} - D}{S(x)} \qquad A_{Me} = \frac{Me - D}{Q}$$

Jeżeli:

rozkład jest symetryczny, to $A_{\bar{x}} = 0$,

rozkład jest asymetryczny prawostronnie, to $A_{\bar{x}} > 0$,

rozkład jest asymetryczny lewostronnie, to $A_{\bar{x}} < 0$.

Miary asymetrii

Pozycyjny współczynnik asymetrii, przedstawiony za pomocą miar pozycyjnych (uzupełnienie miar klasycznych):

$$A_Q = \frac{(Q_3 - Me) - (Me - Q_1)}{(Q_3 - Me) + (Me - Q_1)}$$

Większa rozpiętość III ćwiartki $Q_3 – Me > Me – Q_1$ (wówczas $A_Q > 0$) oznacza asymetrię prawostronną.

Większa rozpiętość II ćwiartki $Q_3 – Me < Me – Q_1$ (wówczas $A_Q < 0$) oznacza asymetrię lewostronną.

ZADANIE
W firmie Radio-Taxi przeprowadzono badanie taksówek ze względu na tygodniową liczbę przejechanych kilometrów.
Stosując metodę momentów należy przeprowadzić analizę struktury.

Przebieg w tys. km	Odsetek taksówek	x_i'	$x_i' n_i$
0.8 - 1.2	2	1	2
1.2 - 1.6	**15**	**1,4**	**21**
1.6 - 2.0	41	1,8	73,8
2.0 - 2.4	**33**	**2,2**	**72,6**
2.4 - 2.8	6	2,6	15,6
2.8 - 3.2	**3**	**3**	**9**
suma	N =100		194

Miara przeciętna

$$\bar{x} = \frac{\sum_{i=1}^{k} x_i' n_i}{N} = \frac{194}{100} = 1,94$$

ZADANIE

W firmie Radio-Taxi przeprowadzono badanie taksówek ze względu na tygodniową liczbę przejechanych kilometrów.
Stosując metodę momentów należy przeprowadzić analizę struktury.

Przebieg w tys. km	Odsetek taksówek	x'_i	$x'_i n_i$	$(x'_i - \bar{x})^2 n_i$
0.8 - 1.2	2	1	2	1,77
1.2 - 1.6	**15**	**1,4**	**21**	**4,37**
1.6 - 2.0	41	1,8	73,8	0,8
2.0 - 2.4	**33**	**2,2**	**72,6**	**2,23**
2.4 - 2.8	6	2,6	15,6	2,61
2.8 - 3.2	**3**	**3**	**9**	**3,37**
suma	N =100		194	15,16

Miary zróżnicowania

$$\sigma^2 = S^2(x) = \frac{\sum (x'_i - \bar{x})^2 n_i}{N} = \frac{15,16}{100} = 0,1516$$

$$S_{(x)} = \sqrt{0,1516} = 0,4$$

$$V_x = \frac{S_x}{\bar{x}} \cdot 100\% = \frac{0,4}{1,94} \cdot 100\% = 20\%$$

$$1,54 < x_{typ} < 2,34$$

ZADANIE

W firmie Radio-Taxi przeprowadzono badanie taksówek ze względu na tygodniową liczbę przejechanych kilometrów.

Stosując metodę momentów należy przeprowadzić analizę struktury.

Przebieg w tys. km	Odsetek taksówek	x'_i	$x'_i n_i$	$(x'_i - \bar{x})^2 n_i$	$(x'_i - \bar{x})^3 n_i$
0.8 - 1.2	2	1	2	1,77	-1,66
1.2 - 1.6	**15**	**1,4**	**21**	**4,37**	**-2,36**
1.6 - 2.0	41	1,8	73,8	0,8	-0,11
2.0 - 2.4	**33**	**2,2**	**72,6**	**2,23**	**0,58**
2.4 - 2.8	6	2,6	15,6	2,61	1,73
2.8 - 3.2	**3**	**3**	**9**	**3,37**	**3,57**
suma	N =100		194	15,16	1,742

Miary asymetrii

$$\mu_3 = \frac{\sum (x'_i - \bar{x})^3 n_i}{N} = \frac{1,742}{100} = 0,01742 \qquad \alpha_3 = \frac{\mu_3}{s^3(x)} = \frac{0,01742}{0,4^3} = 0,3$$

Asymetria prawostronna

Przykład Dane dotyczą zatrudnienia w sześciu przedsiębiorstwach.

Tabela Obliczenia pomocnicze przy obliczaniu współczynnika asymetrii w szeregu wyliczającym

x_i	$x_i-\bar{x}$	$(x_i-\bar{x})^2$	$(x_i-\bar{x})^3$
8	-11,33	128,3689	-1454,42
9	-10,33	106,7089	-1102,30
19	-0,33	0,1089	-0,04
21	1,67	2,7889	4,66
25	5,67	32,1489	182,28
34	14,67	215,2089	3157,11
116	-	485,3334	787,30

Na podstawie danych zamieszczonych w tabeli

$$\bar{x}=\frac{116}{6}=19{,}33\,;\quad S(x)=\sqrt{\frac{1}{6}\cdot 485{,}3334}=8{,}99\,;$$

oraz:

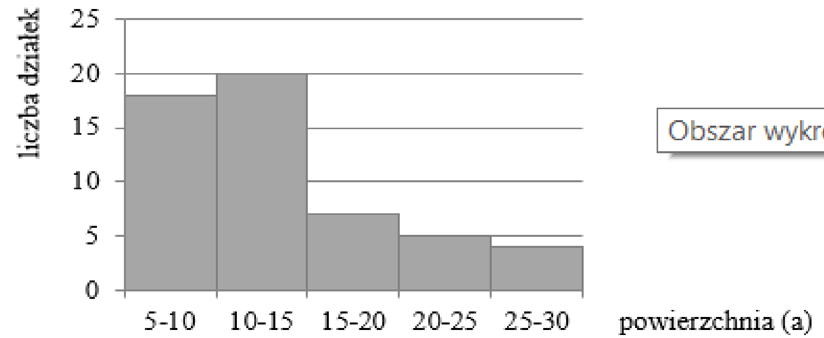
$$A=\frac{\frac{1}{6}\cdot 787{,}30}{8{,}99^3}=0{,}18\,.$$

Uzyskany wynik oznacza, że rozkład liczby zatrudnionych w badanych sześciu przedsiębiorstwach charakteryzuje się słabą asymetrią prawostronną, czyli wyższe wartości (powyżej średniej) są bardziej rozciągnięte (bardziej zróżnicowane) niż wartości niższe, które są bardziej skupione.

Nieobciążony współczynnik asymetrii przyjmuje tutaj wartość $\hat{A}$ =0,25, wniosek jest zatem identyczny jak powyższy.

Przykład Dane dotyczą wielkości działek gruntu (w arach) przeznaczonych pod budownictwo jednorodzinne (dane przykładowe).
Obliczenia pomocnicze przy obliczaniu współczynnika asymetrii w szeregu przedziałowym

	x_i	n_i	$n_{i(sk)}$	$x_i \cdot n_i$	$x_i - \overline{x}$	$(x_i - \overline{x})^2 \cdot n_i$	$(x_i - \overline{x})^3 \cdot n_i$
5-10	7,5	18	18	135	-6,02	652,3272	-3927,01
10-15	12,5	20	38	250	-1,02	20,8080	-21,22
15-20	17,5	7	45	122,5	3,98	110,8828	441,31
20-25	22,5	5	50	112,5	8,98	403,2020	3620,75
25-30	27,5	4	54	110	13,98	781,7616	10929,03
	-	54	-	730	-	1968,9816	11042,86



Rysunek. Rozkład powierzchni działek budowlanych

Na podstawie danych zamieszczonych w tabeli

$$\overline{x} = \frac{730}{54} = 13{,}52\ ; \quad S(x) = \sqrt{\frac{1}{54} \cdot 1968{,}9816} = 6{,}04\ ;$$

oraz:

$$A = \frac{\frac{1}{54} \cdot 11042{,}86}{6{,}04^3} = 0{,}93.$$

Uzyskany wynik oznacza bardzo silną asymetrię prawostronną rozkładu powierzchni działek budowlanych, co oznacza, że wyższe wartości są bardziej rozciągnięte niż wartości niższe, powierzchnia większości działek skupia się na dolnych wartościach.

Wyznaczając dodatkowe miary:

$$D = 10 + \frac{20-18}{(20-18)+(20-7)} \cdot 5 = 10{,}67\ ,$$

$$Q_1 = 5 + \frac{13{,}5-0}{18} \cdot 5 = 8{,}75\ ,$$

$$Me = 10 + \frac{27-18}{20} \cdot 5 = 12{,}25\ ,$$

$$Q_3 = 15 + \frac{40{,}5-38}{7} \cdot 5 = 16{,}79\ ,$$

$$Q = \frac{16{,}79-8{,}75}{2} = 4{,}02\ ,$$

można wyznaczyć uzupełniające współczynniki asymetrii:

$$A_{\bar{x}} = \frac{13{,}52-10{,}67}{6{,}04} = 0{,}47\ , \quad A_{Me} = \frac{12{,}25-10{,}67}{4{,}02} = 0{,}39\ ,$$

$$A_Q = \frac{(16{,}79-12{,}25)-(12{,}25-8{,}75)}{(16{,}79-12{,}25)+(12{,}25-8{,}75)} = \frac{4{,}54-3{,}5}{4{,}54+3{,}5} = 0{,}13\ .$$

Wszystkie uzupełniające współczynniki asymetrii wskazują na asymetrię prawostronną, ich siła jest od słabej do przeciętnej. Uzyskane wartości dość znacznie różnią się od wartości współczynnika A. Jednak inne jest tutaj podejście do problemu asymetrii, stąd też nie można oczekiwać takich samych rezultatów.

Z wartości współczynnika $A_{\bar{x}}$ wnioskować można, że dominująca liczba działek budowlanych ma powierzchnię wyraźnie mniejszą niż wartość średnia. Podobnie z wartości współczynnika A_{Me} wnioskować można, że dominująca liczba działek budowlanych ma powierzchnię wyraźnie mniejszą niż wartość środkowa. Z kolei z wartości współczynnika A_Q wnioskować można, że rozpiętość III ćwiartki jest nieznacznie szersza od rozpiętości II ćwiartki, czyli działki z III ćwiartki są nieznacznie bardziej zróżnicowane ze względu na powierzchnię niż działki z II ćwiartki.

Miary koncentracji

Koncentracja wokół poziomu przeciętnego jest jednym z elementów opisu zbiorowości statystycznej. Jednostki statystyczne mogą wykazywać silne bądź słabe skupienie. Z reguły, im większe zróżnicowanie wartości, mierzone np. odchyleniem standardowym, tym mniejsza jest koncentracja wokół poziomu przeciętnego i odwrotnie.

Jednak mogą zdarzyć się sytuacje, że przy takim samym zróżnicowaniu zbiorowości będą różnić się koncentracją, która dość silne zależy od obszaru zmienności badanej cechy.

Miary koncentracji

Podstawowy wzór **współczynnika koncentracji**:

$$W_k = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{S^4(x)} \quad \text{lub} \quad W_k = \frac{\frac{1}{n}\sum_{i=1}^{k}(\dot{x}_i - \bar{x})^4 n_i}{S^4(x)}$$

Jeżeli:

W_k=3, koncentracja normalna

W_k>3, koncentracja silniejsza niż normalna, rozkład liczebności (częstości) takiej cechy nazywa się leptokurtycznym;

W_k<3, koncentracja słabsza niż normalna, rozkład liczebności (częstości) takiej cechy nazywa się platykurtycznym

Miary koncentracji

Odejmując od współczynnika koncentracji wartość 3 otrzymuje się miernik zwany **kurtozą** (ekscesem).

<u>Interpretacja:</u> Kurtuoza jest miarą skupienia poszczególnych obserwacji wokół średniej.

Kurtoza przedstawiona wzorem:

$$K = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^4}{S^4(x)} - 3$$

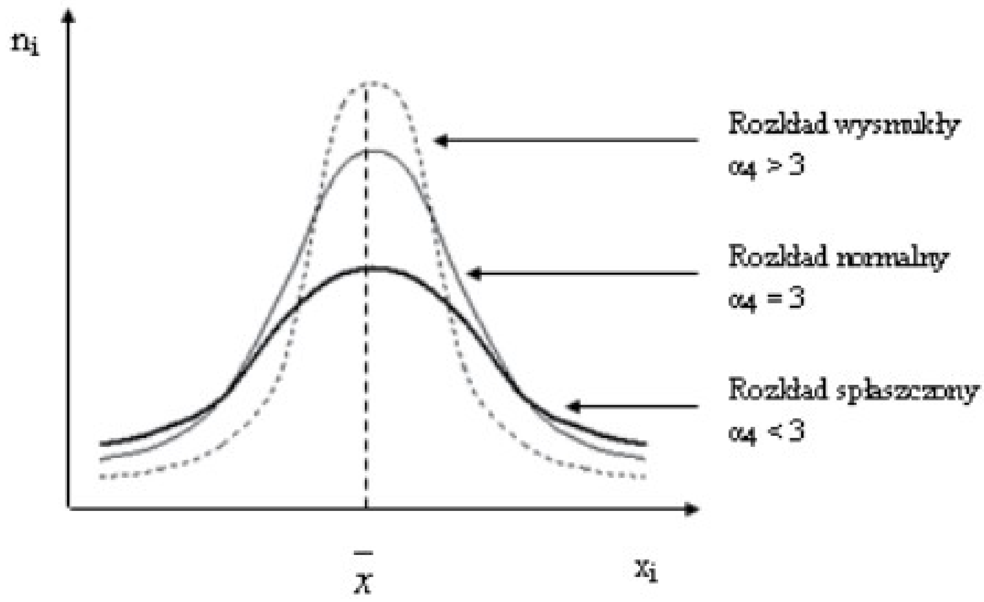
<u>Jeżeli:</u>

K=0 koncentracja normalna

K>0 koncentracja silniejszą niż normalną,

K<0 koncentrację słabszą niż normalną.

Rozkład cech o różnym skupieniu



Im krzywa liczebności (częstości) bardziej wysmukła tym koncentracja wartości jednostek statystycznych wokół poziomu średniego silniejsza.

Przykład Dane dotyczą zatrudnienia w sześciu przedsiębiorstwach.

Obliczenia pomocnicze przy obliczaniu współczynnika koncentracji w szeregu wyliczającym

x_i	$x_i - \bar{x}$	$(x_i - \bar{x})^2$	$(x_i - \bar{x})^4$
8	-11,33	128,3689	16 478,57
9	-10,33	106,7089	11 386,79
19	-0,33	0,1089	0,01
21	1,67	2,7889	7,78
25	5,67	32,1489	1 033,55
34	14,67	215,2089	46 314,87
116	-	485,3334	75 221,58

Na podstawie danych zamieszczonych w tabeli

$$\bar{x} = \frac{116}{6} = 19{,}33\,; \quad S(x) = \sqrt{\frac{1}{6} \cdot 485{,}3334} = 8{,}99\,;$$

$$W_k = \frac{\frac{1}{6} \cdot 75\,221{,}58}{8{,}99^4} = 1{,}92\,; \quad (K{=}1{,}92{-}3{=}{-}1{,}08)$$

Uzyskany wynik oznacza, że rozkład zatrudnienia w badanych sześciu przedsiębiorstwach jest słabo skoncentrowany wokół poziomu przeciętnego.

Nieobciążony wskaźnik kurtozy przyjmuje tutaj wartość $\hat{K}$ =-0,66, wniosek jest zatem analogiczny do powyższego.

Przykład Dane dotyczą wielkości działek gruntu (w arach) przeznaczonych pod budownictwo jednorodzinne (dane przykładowe).

Obliczenia pomocnicze przy obliczaniu współczynnika koncentracji w szeregu przedziałowym

	x_i	n_i	$n_{i(sk)}$	$x_i \cdot n_i$	$x_i - \bar{x}$	$(x_i - \bar{x})^2 \cdot n_i$	$(x_i - \bar{x})^4 \cdot n_i$
5-10	7,5	18	18	135	-6,02	652,3272	23 640,6
10-15	12,5	20	38	250	-1,02	20,8080	21,6
15-20	17,5	7	45	122,5	3,98	110,8828	1 756,4
20-25	22,5	5	50	112,5	8,98	403,2020	32 514,4
25-30	27,5	4	54	110	13,98	781,7616	152 787,8
	-	54	-	730	-	1968,9816	210 720,8

Na podstawie danych zamieszczonych w tabeli

$$\bar{x} = \frac{730}{54} = 13{,}52\ ; \quad S(x) = \sqrt{\frac{1}{54} \cdot 1968{,}9816} = 6{,}04\ ;$$

oraz:
$$W_k = \frac{\frac{1}{54} \cdot 210\,720{,}8}{6{,}04^4} = 2{,}93\,.$$

Uzyskany wynik oznacza, że rozkład wielkości działek gruntu wokół poziomu przeciętnego charakteryzuje się koncentracją zbliżoną do normalnej.

Z uwagi na silną asymetrię wynik ten nie odpowiada w pełni koncentracji normalnej.

Z rysunku nie jest łatwo odczytać poziom koncentracji, gdyż skale na osi odciętych i rzędnych przedstawiają inne wielkości. Na osi odciętych przedstawiamy warianty zmiennej, a na osi rzędnych ich liczebności. Nie jest to klasyczny układ współrzędnych. Trudność ta nie występuje w przypadku asymetrii, gdzie porównuje się rozkłada liczebności wokół średniej.

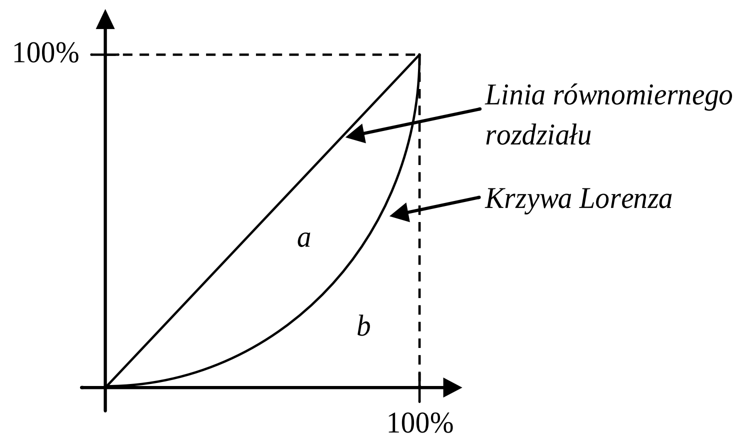
Nierównomierność podziału zjawiska w zbiorowości

Koncentracja rozumiana jako nierównomierny podział zjawiska w zbiorowości oznacza nierównomierne rozłożenie wartości cechy zmiennej pomiędzy poszczególne jednostki tej zbiorowości.

Jeśli wszystkie jednostki zbiorowości dysponują taką samą wartością cechy zmiennej, to koncentracja nie występuje, natomiast jeśli jedna jednostka zbiorowości dysponuje całą sumą wartości cechy zmiennej to występuje wówczas koncentracja zupełna.

Inny rodzaj koncentracji

Krzywa Lorenza:



Rys. Krzywa Lorenza

Na osi odciętych (x) odkłada się skumulowaną częstość jednostek statystycznych, a na osi rzędnych (Y) udział w łącznym funduszu cechy.

Na podstawie stosunku pola *a* do pola trójkąta zawierającego krzywą Lorenza (połowa kwadratu o boku 100) wyznacza się współczynnik koncentracji Giniego.

Wielobok koncentracji Lorenza

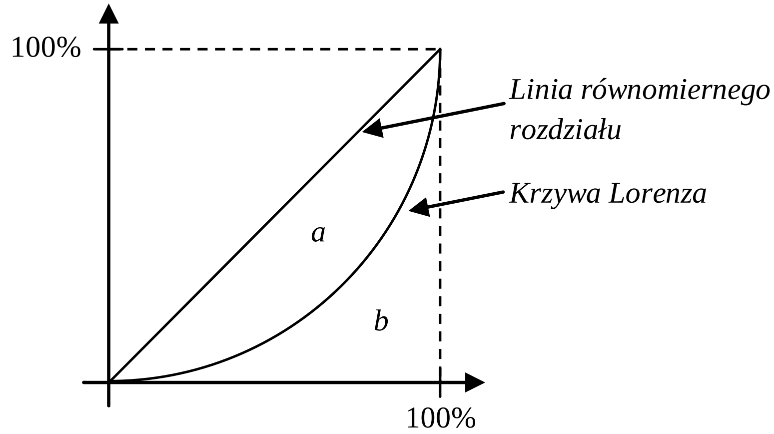
Wielobok koncentracji Lorenza pozwala ocenić koncentrację zjawiska w sposób graficzny, za pomocą krzywej Lorenza.

Koncentracja zjawiska jest tym większa, im większa jest powierzchnia zawarta między krzywą Lorenza, a linią równomiernego rozdziału (krzywa Lorenza jest bardziej wypukła w stosunku do linii równomiernego rozdziału), i odwrotnie.

Jeżeli krzywa Lorenza pokrywa się z linią równomiernego rozdziału to koncentracja nie występuje.

Inny rodzaj koncentracji

- Krzywa Lorenza:



Rys. Krzywa Lorenza

Na podstawie stosunku pola *a* do pola trójkąta zawierającego krzywą Lorenza wyznacza się **współczynnik koncentracji Giniego**:

$$K_L \square \frac{a}{500}$$

$$K_L \square \frac{a}{5000} \square \frac{5000 \square b}{5000}$$

gdzie:

a - jest polem pomiędzy linią równomiernego rozdziału a krzywą Lorenza.

b – *j*est powierzchnią pola pod krzywą Lorenza

Miary koncentracji

Współczynnik określony w ten sposób jest współczynnikiem unormowanym i przyjmuje wartości z przedziału **<0; 1>.**

K_L=0 – oznacza brak koncentracji (rozdział równomierny), krzywa Lorenza pokrywa się z linią równomiernego rozdziału, wtedy pole *a=0.*

K_L=1 - oznacza koncentrację zupełną (cały fundusz cech przypada na jedną jednostkę statystyczną), gdy pole *b*=0 (*a*=5000), a krzywa Lorenza staje się łamaną przebiegającą po podstawie i prawym boku trójkąta