PODSTAWOWE MIARY OPISU STRUKTURY ZBIOROWOŚCI STATYSTYCZNEJ

Agnieszka Rossa

Szkic wykładu

Analiza struktury zbiorowości

Analiza struktury polega na przetworzeniu szeregów strukturalnych w syntetyczne miary opisowe takie, jak:

I. **Wskaźniki natężenia i struktury** – wskaźnik natężenia wyraża kształtowanie się wielkości jednego zjawiska na tle innego, logicznie z nim związanego; wskaźniki struktury reprezentują z kolei liczebności względne występowania określonych wartości w badanej zbiorowości.
II. **Miary średnie** (tendencji centralnej) – opisują przeciętne położenie wartości liczbowych danej cechy statystycznej.
III. **Miary zróżnicowania** (dyspersji, rozrzutu, zmienności, rozproszenia) – opisują stopień rozproszenia wartości badanej cechy wokół średniej.
IV. **Miary asymetrii** (skośności) – informują, czy większa część jednostek zbiorowości ma wartości cechy większe czy też mniejsze od wartości centralnej.

Wskaźniki natężenia

- **Wskaźnik natężenia** jest ilorazem liczebności jednej zbiorowości do liczebności innej zbiorowości, logicznie z nią związanej.

Wskaźniki natężenia

- **Wskaźnik natężenia** jest ilorazem liczebności jednej zbiorowości do liczebności innej zbiorowości, logicznie z nią związanej.
- Typowym wskaźnikiem natężenia jest wskaźnik gęstości zaludnienia, obliczany jako iloraz liczby mieszkańców do powierzchni danego obszaru (na przykład, w Łodzi gęstość zaludnienia w 2008 roku była na poziomie ok. 2548 osób na kilometr kwadratowy).

Wskaźniki struktury

- Załóżmy, że wartości badanej cechy w *n*-elementowej zbiorowości zostały pogrupowane w szereg rozdzielczy (punktowy lub z przedziałami klasowymi).

Wskaźniki struktury

- Załóżmy, że wartości badanej cechy w n-elementowej zbiorowości zostały pogrupowane w szereg rozdzielczy (punktowy lub z przedziałami klasowymi).

- Niech n_i, $i = 1, 2, \ldots, k$ będą liczebnościami empirycznymi poszczególnych klas szeregu.

Wskaźniki struktury

- Załóżmy, że wartości badanej cechy w n-elementowej zbiorowości zostały pogrupowane w szereg rozdzielczy (punktowy lub z przedziałami klasowymi).
- Niech $n_i,\ i = 1, 2, \ldots, k$ będą liczebnościami empirycznymi poszczególnych klas szeregu.
- **Wskaźnikiem struktury** w_i nazywamy liczebność względną i-tej klasy zdefiniowaną jako iloraz

$$w_i = \frac{n_i}{n}, \text{ przy czym } \sum_{i=1}^{k} w_i = 1.$$

Wskaźnik podobieństwa struktur

- Wskaźniki struktury można wykorzystać do oceny podobieństwa struktur zbiorowości ze względu na wybraną cechę.

Wskaźnik podobieństwa struktur

- Wskaźniki struktury można wykorzystać do oceny podobieństwa struktur zbiorowości ze względu na wybraną cechę.
- Załóżmy, że wartości cechy w dwóch zbiorowościach pogrupowano w szeregi rozdzielcze o jednakowych klasach. Niech w_{1i} oraz w_{2i} oznaczają wskaźniki struktury dla i-tej klasy w obu szeregach.

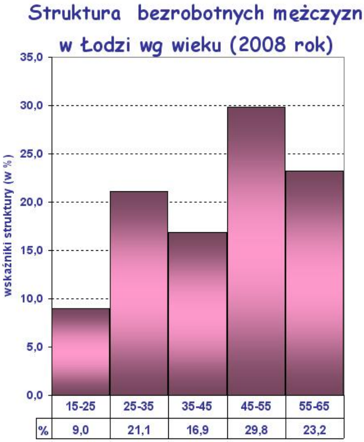
Wskaźnik podobieństwa struktur

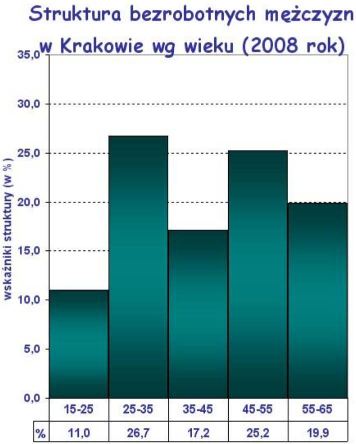
- Wskaźniki struktury można wykorzystać do oceny podobieństwa struktur zbiorowości ze względu na wybraną cechę.
- Załóżmy, że wartości cechy w dwóch zbiorowościach pogrupowano w szeregi rozdzielcze o jednakowych klasach. Niech w_{1i} oraz w_{2i} oznaczają wskaźniki struktury dla i-tej klasy w obu szeregach.
- Wówczas **wskaźnik w_p podobieństwa struktur** dla obu zbiorowości obliczamy ze wzoru

$$w_p = \sum_{i=1}^{k} \min(w_{1i}, w_{2i}).$$

Wskaźnik podobieństwa struktur

W tym przykładzie jest wysoki: $w_p = 9 + 21,1 + 16,9 + 25,2 + 19,9 = 92,1\%$





Miary średnie
Podział

Miary średnie dzielimy na:

Miary średnie
Podział

Miary średnie dzielimy na:

1. **Miary średnie klasyczne.** Są to miary obliczane dla cechy ilościowej na podstawie jej wartości odnotowanych dla wszystkich jednostek zbiorowości. Do miar średnich zaliczamy:

Miary średnie
Podział

Miary średnie dzielimy na:

1. **Miary średnie klasyczne.** Są to miary obliczane dla cechy ilościowej na podstawie jej wartości odnotowanych dla wszystkich jednostek zbiorowości. Do miar średnich zaliczamy:

   a. średnia arytmetyczna (oznaczana jako $\bar{x}$),

Miary średnie
Podział

Miary średnie dzielimy na:

1. **Miary średnie klasyczne.** Są to miary obliczane dla cechy ilościowej na podstawie jej wartości odnotowanych dla wszystkich jednostek zbiorowości. Do miar średnich zaliczamy:

   a. średnia arytmetyczna (oznaczana jako $\bar{x}$),

   b. średnia harmoniczna ($\bar{x}_h$) stosowana głównie w odniesieniu do cech stosunkowych (np. wydajność, prędkość itp.),

Miary średnie
Podział

Miary średnie dzielimy na:

1. **Miary średnie klasyczne.** Są to miary obliczane dla cechy ilościowej na podstawie jej wartości odnotowanych dla wszystkich jednostek zbiorowości. Do miar średnich zaliczamy:

   a. średnia arytmetyczna (oznaczana jako $\bar{x}$),

   b. średnia harmoniczna ($\bar{x}_h$) stosowana głównie w odniesieniu do cech stosunkowych (np. wydajność, prędkość itp.),

   c. średnia geometryczna (G) stosowana np. w odniesieniu do wskaźników dynamiki (będzie przedstawiona przy okazji analizy dynamiki zjawisk).

Miary średnie
Podział – c.d.

2. **Miary średnie pozycyjne.** Nazwa tych miar pochodzi stąd, iż są obliczane na podstawie wartości cechy tylko tych jednostek, które zajmują określoną pozycję w uporządkowanym szeregu lub które mogą być uznane za szczególnie charakterystyczne dla danej zbiorowości.

   Do tej grupy miar zaliczamy:

Miary średnie
Podział – c.d.

2. **Miary średnie pozycyjne.** Nazwa tych miar pochodzi stąd, iż są obliczane na podstawie wartości cechy tylko tych jednostek, które zajmują określoną pozycję w uporządkowanym szeregu lub które mogą być uznane za szczególnie charakterystyczne dla danej zbiorowości.

   Do tej grupy miar zaliczamy:

   a. wartość modalną inaczej nazywana dominantą (*Do*),

Miary średnie
Podział – c.d.

2. **Miary średnie pozycyjne.** Nazwa tych miar pochodzi stąd, iż są obliczane na podstawie wartości cechy tylko tych jednostek, które zajmują określoną pozycję w uporządkowanym szeregu lub które mogą być uznane za szczególnie charakterystyczne dla danej zbiorowości.

   Do tej grupy miar zaliczamy:

   a. wartość modalną inaczej nazywana dominantą (Do),

   b. kwartyle, w tym: kwartyl pierwszy (Q_1), kwartyl drugi (Q_2), kwartyl trzeci (Q_3); szczególne znaczenie ma kwartyl drugi zwany także medianą lub wartością środkową i oznaczany symbolem Me.

Klasyczne miary średnie
Średnia arytmetyczna

- **Średnia arytmetyczna** jest sumą wszystkich wartości badanej cechy, podzieloną przez ich liczbę.

Klasyczne miary średnie
Średnia arytmetyczna

- **Średnia arytmetyczna** jest sumą wszystkich wartości badanej cechy, podzieloną przez ich liczbę.
- Przykładem może być średnia ocen w indeksie (każdy student z pewnością ją obliczał).

Klasyczne miary średnie
Średnia arytmetyczna

- **Średnia arytmetyczna** jest sumą wszystkich wartości badanej cechy, podzieloną przez ich liczbę.
- Przykładem może być średnia ocen w indeksie (każdy student z pewnością ją obliczał).
- Dla zapisu formalnego wzoru na średnią arytmetyczną przyjmijmy następujące oznaczenia. Niech $x_1, x_2, \ldots, x_n$ oznaczają kolejne wartości badanej cechy (np. kolejne oceny w indeksie). Wówczas średnią arytmetyczną zapiszemy wzorem:

$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$ lub krócej: $\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$

Klasyczne miary średnie
Średnia arytmetyczna

- **Średnia arytmetyczna** jest sumą wszystkich wartości badanej cechy, podzieloną przez ich liczbę.
- Przykładem może być średnia ocen w indeksie (każdy student z pewnością ją obliczał).
- Dla zapisu formalnego wzoru na średnią arytmetyczną przyjmijmy następujące oznaczenia. Niech $x_1, x_2, \ldots, x_n$ oznaczają kolejne wartości badanej cechy (np. kolejne oceny w indeksie). Wówczas średnią arytmetyczną zapiszemy wzorem:
  $\bar{x} = \frac{x_1+x_2+\ldots+x_n}{n}$ lub krócej: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
- Przedstawiona średnią, zwana jest średnią arytmetyczną prostą, ponieważ oblicza się ją na podstawie szeregów szczegółowych prostych.

Klasyczne miary średnie
Średnia arytmetyczna – c.d.

- W przypadku szeregów rozdzielczych korzystamy z **formuł ważonych**, w których rolę wag pełnią liczebności n_i:

$$\bar{x} = \begin{cases} \frac{1}{n}\sum_{i=1}^{k} x_i n_i, & \text{dla szeregu punktowego,} \\ \frac{1}{n}\sum_{i=1}^{k} \dot{x}_i n_i, & \text{dla szeregu z przedziałami klasowymi,} \end{cases}$$

gdzie $\dot{x}_i$ dla $i = 1, 2, \ldots, k$ oznaczają środki przedziałów klasowych, natomiast k jest liczbą wierszy szeregu rozdzielczego.

Klasyczne miary średnie
Średnia arytmetyczna – ważona średnia ze średnich

Formułę **średniej ważonej** stosujemy także w przypadku obliczania średniej ze średnich.

Przykład 1.

- Załóżmy, że mamy trzy zbiory danych A, B, C:

$$A = \{4, 6, 5\}, \quad B = \{7, 9\}, \quad C = \{5, 4, 5, 3, 3\}$$

Klasyczne miary średnie
Średnia arytmetyczna – ważona średnia ze średnich

Formułę **średniej ważonej** stosujemy także w przypadku obliczania średniej ze średnich.

Przykład 1.

- Założmy, że mamy trzy zbiory danych A, B, C:

$$A = \{4, 6, 5\}, \quad B = \{7, 9\}, \quad C = \{5, 4, 5, 3, 3\}$$

- Można sprawdzić, że średnie arytmetyczne wyznaczone z danych ze zbiorów A, B, C równe są odpowiednio: $5, \ 8, \ 4$.

Klasyczne miary średnie
Średnia arytmetyczna – ważona średnia ze średnich

Formułę **średniej ważonej** stosujemy także w przypadku obliczania średniej ze średnich.

Przykład 1.

- Załóżmy, że mamy trzy zbiory danych A, B, C:

$$A = \{4, 6, 5\}, \quad B = \{7, 9\}, \quad C = \{5, 4, 5, 3, 3\}$$

- Można sprawdzić, że średnie arytmetyczne wyznaczone z danych ze zbiorów A, B, C równe są odpowiednio: $5,\ 8,\ 4$.
- **Pytanie:** Ile wynosi średnia arytmetyczna dla danych z połączonych zbiorów?

Klasyczne miary średnie
Średnia arytmetyczna – ważona średnia ze średnich

Rozwiązanie:

- Pierwszy sposób polega na połączeniu danych ze zbiorów A, B, C i wyznaczeniu z nich średniej arytmetycznej, czyli $\frac{4+6+5+7+9+5+4+5+3+3}{10} = 5,1$.

Klasyczne miary średnie
Średnia arytmetyczna – ważona średnia ze średnich

Rozwiązanie:

- Pierwszy sposób polega na połączeniu danych ze zbiorów A, B, C i wyznaczeniu z nich średniej arytmetycznej, czyli $\frac{4+6+5+7+9+5+4+5+3+3}{10} = 5,1$.
- Drugi sposób polega na wykorzystaniu średnich cząstkowych obliczonych dla zbiorów A, B, C. Błędem byłoby jednak obliczenie zwykłej średniej ze średnich, tj. $\frac{5+8+4}{3}$. Otrzymany wynik (ok. $5,7$) nie zgadza się z uzyskanym wyżej.

Klasyczne miary średnie
Średnia arytmetyczna – ważona średnia ze średnich

Rozwiązanie:

- Pierwszy sposób polega na połączeniu danych ze zbiorów A, B, C i wyznaczeniu z nich średniej arytmetycznej, czyli $\frac{4+6+5+7+9+5+4+5+3+3}{10} = 5,1$.
- Drugi sposób polega na wykorzystaniu średnich cząstkowych obliczonych dla zbiorów A, B, C. Błędem byłoby jednak obliczenie zwykłej średniej ze średnich, tj. $\frac{5+8+4}{3}$. Otrzymany wynik (ok. $5,7$) nie zgadza się z uzyskanym wyżej.
- Poprawne rozwiązanie wymaga zastosowania formuły średniej ważonej, w której wagami są liczebności zbiorów:

$$\frac{5\cdot 3+8\cdot 2+4\cdot 5}{10}=\frac{15+16+20}{10}=5,1.$$

Klasyczne miary średnie – Własności średniej arytmetycznej

1. Spełnia relację $x_{min} < \bar{x} < x_{max}$, gdzie x_{min}, x_{max} oznaczają wartość minimalną i maksymalną w zbiorze danych.

Klasyczne miary średnie – Własności średniej arytmetycznej

1. Spełnia relację $x_{min} < \bar{x} < x_{max}$, gdzie x_{min}, x_{max} oznaczają wartość minimalną i maksymalną w zbiorze danych.
2. Zachodzą następujące równości (wynikają z definicji):

Klasyczne miary średnie – Własności średniej arytmetycznej

1. Spełnia relację $x_{min} < \bar{x} < x_{max}$, gdzie x_{min}, x_{max} oznaczają wartość minimalną i maksymalną w zbiorze danych.
2. Zachodzą następujące równości (wynikają z definicji):
   - dla szeregu szczegółowego

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

Klasyczne miary średnie – Własności średniej arytmetycznej

1. Spełnia relację $x_{min} < \bar{x} < x_{max}$, gdzie x_{min}, x_{max} oznaczają wartość minimalną i maksymalną w zbiorze danych.
2. Zachodzą następujące równości (wynikają z definicji):
   - dla szeregu szczegółowego

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

   - dla szeregu rozdzielczego punktowego

$$\sum_{i=1}^{n} x_i n_i = \bar{x} \sum_{i=1}^{k} n_i,$$

Klasyczne miary średnie – Własności średniej arytmetycznej

1. Spełnia relację $x_{min} < \bar{x} < x_{max}$, gdzie x_{min}, x_{max} oznaczają wartość minimalną i maksymalną w zbiorze danych.
2. Zachodzą następujące równości (wynikają z definicji):
   - dla szeregu szczegółowego

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

   - dla szeregu rozdzielczego punktowego

$$\sum_{i=1}^{n} x_i n_i = \bar{x} \sum_{i=1}^{k} n_i,$$

   - dla szeregu rozdzielczego z przedziałami klasowymi

$$\sum_{i=1}^{n} \dot{x}_i n_i = \bar{x} \sum_{i=1}^{k} n_i.$$

Klasyczne miary średnie – Własności średniej arytmetycznej – c.d.

3. Suma ochyleń poszczególnych wartości obserwowanej cechy od jej średniej arytmetycznej jest równa 0, czyli w przypadku:

Klasyczne miary średnie – Własności średniej arytmetycznej – c.d.

3. Suma ochyleń poszczególnych wartości obserwowanej cechy od jej średniej arytmetycznej jest równa 0, czyli w przypadku:
   - szeregu szczegółowego

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

Klasyczne miary średnie – Własności średniej arytmetycznej – c.d.

3. Suma ochyleń poszczególnych wartości obserwowanej cechy od jej średniej arytmetycznej jest równa 0, czyli w przypadku:
   - szeregu szczegółowego

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

   - szeregu rozdzielczego punktowego

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

Klasyczne miary średnie – Własności średniej arytmetycznej – c.d.

3. Suma ochyleń poszczególnych wartości obserwowanej cechy od jej średniej arytmetycznej jest równa 0, czyli w przypadku:
   - szeregu szczegółowego

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

   - szeregu rozdzielczego punktowego

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

   - szeregu rozdzielczego z przedziałami klasowymi

$$\sum_{i=1}^{k}(\dot{x}_i - \bar{x})n_i = 0.$$

Klasyczne miary średnie – Własności średniej arytmetycznej – c.d.

4. Suma kwadratów ochyleń poszczególnych wartości cechy od jej średniej arytmetycznej jest minimalna, czyli dla dowolnej stałej *a* spełnione są nierówności:

Klasyczne miary średnie – Własności średniej arytmetycznej – c.d.

4. Suma kwadratów ochyleń poszczególnych wartości cechy od jej średniej arytmetycznej jest minimalna, czyli dla dowolnej stałej *a* spełnione są nierówności:
   - w przypadku szeregu szczegółowego

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 \leq \sum_{i=1}^{n}(x_i - a)^2,$$

Klasyczne miary średnie – Własności średniej arytmetycznej – c.d.

4. Suma kwadratów ochyleń poszczególnych wartości cechy od jej średniej arytmetycznej jest minimalna, czyli dla dowolnej stałej *a* spełnione są nierówności:
   - w przypadku szeregu szczegółowego

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 \leq \sum_{i=1}^{n}(x_i - a)^2,$$

   - w przypadku szeregu rozdzielczego punktowego

$$\sum_{i=1}^{k}(x_i - \bar{x})^2 n_i \leq \sum_{i=1}^{k}(x_i - a)^2 n_i,$$

Klasyczne miary średnie – Własności średniej arytmetycznej – c.d.

4. Suma kwadratów ochyleń poszczególnych wartości cechy od jej średniej arytmetycznej jest minimalna, czyli dla dowolnej stałej *a* spełnione są nierówności:
   - w przypadku szeregu szczegółowego

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 \leq \sum_{i=1}^{n}(x_i - a)^2,$$

   - w przypadku szeregu rozdzielczego punktowego

$$\sum_{i=1}^{k}(x_i - \bar{x})^2 n_i \leq \sum_{i=1}^{k}(x_i - a)^2 n_i,$$

   - w przypadku szeregu rozdzielczego z przedziałami kl.

$$\sum_{i=1}^{k}(\dot{x}_i - \bar{x})^2 n_i \leq \sum_{i=1}^{k}(\dot{x}_i - a)^2 n_i.$$

Klasyczne miary średnie
Własności średniej arytmetycznej – c.d.

5. Średniej arytmetycznej nie można obliczyć dla szeregu rozdzielczego z otwartymi przedziałami klasowymi.

   Jeśli otwarte przedziały klasowe mają niewielkie liczebności (do 5% ogólnej liczebności), to przed obliczeniem średniej arytmetycznej można je domknąć.

Klasyczne miary średnie
Własności średniej arytmetycznej – c.d.

5. Średniej arytmetycznej nie można obliczyć dla szeregu rozdzielczego z otwartymi przedziałami klasowymi.

   Jeśli otwarte przedziały klasowe mają niewielkie liczebności (do 5% ogólnej liczebności), to przed obliczeniem średniej arytmetycznej można je domknąć.

6. Średnia arytmetyczna jest **”wrażliwa” na nietypowe wartości** cechy (tj. znacznie różniące się od pozostałych wartości w zbiorze); wielkości odstające mogą zniekształcić (zawyżyć lub zaniżyć) wartość średniej arytmetycznej.

   Istnieją jednak pewne sposoby radzenia sobie z taką sytuacją. Jeśli mamy podstawy przypuszczać, że wartość odstająca pojawiła się przypadkowo, wówczas przed obliczeniem średniej usuwamy tę wartość ze zbioru danych. Drugim sposobem jest przekształcenie wszystkich danych np. za pomocą funkcji logarytmicznej, dzięki czemu wartości w zbiorze po transformacji będą do siebie bardziej zbliżone.

Klasyczne miary średnie
Średnia harmoniczna

- **Średnia harmoniczna** jest odwrotnością średniej arytmetycznej z odwrotności zaobserwowanych wartości $x_1, x_2, \ldots, x_n$ cechy ilościowej w badanej zbiorowości.

Klasyczne miary średnie
Średnia harmoniczna

- **Średnia harmoniczna** jest odwrotnością średniej arytmetycznej z odwrotności zaobserwowanych wartości $x_1, x_2, \ldots, x_n$ cechy ilościowej w badanej zbiorowości.
- Formalnie zapisujemy ją wzorem:

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

Przedstawiona formuła odnosi się do szeregów prostych.

Klasyczne miary średnie
Średnia harmoniczna

- **Średnia harmoniczna** jest odwrotnością średniej arytmetycznej z odwrotności zaobserwowanych wartości $x_1, x_2, \ldots, x_n$ cechy ilościowej w badanej zbiorowości.
- Formalnie zapisujemy ją wzorem:

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

Przedstawiona formuła odnosi się do szeregów prostych.

- W przypadku szeregów rozdzielczych korzystamy z formuł ważonych:

$$\bar{x}_h = \begin{cases} \frac{n}{\sum_{i=1}^{k} \frac{1}{x_i} n_i}, & \text{dla szeregu punktowego,} \\ \frac{n}{\sum_{i=1}^{k} \frac{1}{\dot{x}_i} n_i}, & \text{dla szeregu z przedziałami klasowymi.} \end{cases}$$

Klasyczne miary średnie
Średnia harmoniczna – przykłady

Średnie harmoniczne stosuje się do obliczania poziomu średniego dla cechy o charakterze stosunkowym, takich jak: wydajność, prędkość, siła nabywcza pieniądza itp.

Przykład 2.

Klasyczne miary średnie
Średnia harmoniczna – przykłady

Średnie harmoniczne stosuje się do obliczania poziomu średniego dla cechy o charakterze stosunkowym, takich jak: wydajność, prędkość, siła nabywcza pieniądza itp.

Przykład 2.

- Długość linii kolejowej łączącej miasta A i B jest równa 100 km. Pociąg pospieszny jedzie z miasta A do miasta B z prędkością 100 km/h, a pociąg osobowy – z prędkością 50 km/h.

Klasyczne miary średnie
Średnia harmoniczna – przykłady

Średnie harmoniczne stosuje się do obliczania poziomu średniego dla cechy o charakterze stosunkowym, takich jak: wydajność, prędkość, siła nabywcza pieniądza itp.

Przykład 2.

- Długość linii kolejowej łączącej miasta A i B jest równa 100 km. Pociąg pospieszny jedzie z miasta A do miasta B z prędkością 100 km/h, a pociąg osobowy – z prędkością 50 km/h.
- **Pytanie:** Jaka jest średnia prędkość obu pociągów na tej trasie?

Klasyczne miary średnie
Średnia harmoniczna – przykłady c.d.

- Nasuwa się pozornie oczywista odpowiedź, że średnia prędkość obu pociągów jest równa $\frac{100+50}{2} = 75$ km/h.

Klasyczne miary średnie
Średnia harmoniczna – przykłady c.d.

- Nasuwa się pozornie oczywista odpowiedź, że średnia prędkość obu pociągów jest równa $\frac{100+50}{2} = 75$ km/h.
- Przyglądając się jednak bliżej, zauważymy, że obydwa pociągi pokonują łącznie trasę 200 km w czasie 3 godzin, a zatem (poprawna) średnia prędkość wynosi $\frac{200}{3} \approx 66,7$ km/h.

Klasyczne miary średnie
Średnia harmoniczna – przykłady c.d.

- Nasuwa się pozornie oczywista odpowiedź, że średnia prędkość obu pociągów jest równa $\frac{100+50}{2} = 75$ km/h.
- Przyglądając się jednak bliżej, zauważymy, że obydwa pociągi pokonują łącznie trasę 200 km w czasie 3 godzin, a zatem (poprawna) średnia prędkość wynosi $\frac{200}{3} \approx 66,7$ km/h.
- Ten sam wynik uzyskamy, obliczając średnią harmoniczną z obu prędkości. Mamy bowiem

$$\bar{x}_h = \frac{2}{\frac{1}{100} + \frac{1}{50}} = \frac{2}{\frac{1}{100} + \frac{2}{100}} = \frac{2 \cdot 100}{3} \approx 66,7 \text{ km/h}.$$

Klasyczne miary średnie
Średnia harmoniczna – przykłady c.d.

Przykład 3.

- W pewnym banku przy okienkach kasowych zatrudnionych jest 10 pracowników.

Klasyczne miary średnie
Średnia harmoniczna – przykłady c.d.

Przykład 3.

- W pewnym banku przy okienkach kasowych zatrudnionych jest 10 pracowników.
- Zmierzono czas obsługi klientów w ciągu wybranego 8-godzinnego dnia pracy.

Klasyczne miary średnie
Średnia harmoniczna – przykłady c.d.

Przykład 3.

- W pewnym banku przy okienkach kasowych zatrudnionych jest 10 pracowników.
- Zmierzono czas obsługi klientów w ciągu wybranego 8-godzinnego dnia pracy.
- Pięciu pracowników potrzebowało na realizację transakcji zleconych przez pojedynczego klienta po 20 min, trzech pracowników – 15 min, a dwóch pracowników – 10 min.

Klasyczne miary średnie
Średnia harmoniczna – przykłady c.d.

Przykład 3.

- W pewnym banku przy okienkach kasowych zatrudnionych jest 10 pracowników.
- Zmierzono czas obsługi klientów w ciągu wybranego 8-godzinnego dnia pracy.
- Pięciu pracowników potrzebowało na realizację transakcji zleconych przez pojedynczego klienta po 20 min, trzech pracowników – 15 min, a dwóch pracowników – 10 min.
- **Pytanie:** Jaki jest średni czas obsługi klientów banku w badanym dniu pracy?

Średnia harmoniczna – przykłady c.d.

- Ze wzoru na ważoną średnią arytmetyczną otrzymamy w tym przypadku błędny wynik $\frac{20 \cdot 5 + 15 \cdot 3 + 10 \cdot 2}{10} = 16,5$ min.

Średnia harmoniczna – przykłady c.d.

- Ze wzoru na ważoną średnią arytmetyczną otrzymamy w tym przypadku błędny wynik $\frac{20 \cdot 5 + 15 \cdot 3 + 10 \cdot 2}{10} = 16,5$ min.
- Zauważymy, że pracownicy potrzebujący 20, 15 lub 10 min na wykonanie operacji zleconych przez klienta, w ciągu 8-godzinnego dnia pracy zrealizują zlecenia odpowiednio $24 \cdot 5$, $32 \cdot 3$ i $48 \cdot 2$ klientów, obsługując łącznie 312 osób, przy czym czas przepracowany w tym dniu przez wszystkich pracowników wyniesie $8 \cdot 60 \cdot 10 = 4800$ min. Średni czas obsługi klienta przy okienku w danym dniu jest więc równy $\frac{4800}{312} \approx 15,38$ min.

Średnia harmoniczna – przykłady c.d.

- Ze wzoru na ważoną średnią arytmetyczną otrzymamy w tym przypadku błędny wynik $\frac{20 \cdot 5 + 15 \cdot 3 + 10 \cdot 2}{10} = 16,5$ min.
- Zauważymy, że pracownicy potrzebujący 20, 15 lub 10 min na wykonanie operacji zleconych przez klienta, w ciągu 8-godzinnego dnia pracy zrealizują zlecenia odpowiednio $24 \cdot 5$, $32 \cdot 3$ i $48 \cdot 2$ klientów, obsługując łącznie 312 osób, przy czym czas przepracowany w tym dniu przez wszystkich pracowników wyniesie $8 \cdot 60 \cdot 10 = 4800$ min. Średni czas obsługi klienta przy okienku w danym dniu jest więc równy $\frac{4800}{312} \approx 15,38$ min.
- Taki sam wynik otrzymamy ze wzoru na średnią harmoniczną ważoną:

$$\frac{10}{\frac{1}{20} \cdot 5 + \frac{1}{15} \cdot 3 + \frac{1}{10} \cdot 2} = \frac{10}{0,25 + 0,2 + 0,2} \approx 15,38 \text{ min.}$$

Pozycyjne miary średnie
Dominanta (wartość modalna)

- **Dominantą** *Do* nazywamy tę wartość cechy, która w badanej zbiorowości występuje najczęściej.

Pozycyjne miary średnie
Dominanta (wartość modalna)

- **Dominantą** *Do* nazywamy tę wartość cechy, która w badanej zbiorowości występuje najczęściej.
- W szeregach szczegółowych lub rozdzielczych punktowych dominantę można wskazać, odnajdując wartość najliczniej reprezentowaną przez jednostki zbiorowości.

Pozycyjne miary średnie
Dominanta (wartość modalna)

- **Dominantą** *Do* nazywamy tę wartość cechy, która w badanej zbiorowości występuje najczęściej.
- W szeregach szczegółowych lub rozdzielczych punktowych dominantę można wskazać, odnajdując wartość najliczniej reprezentowaną przez jednostki zbiorowości.
- W szeregach rozdzielczych z przedziałami klasowymi (a więc w przypadku cechy ilościowej) można określić jedynie przedział, w którym dominanta występuje. Jest to przedział o największej liczebności, oczywiście pod warunkiem, że przedzial ten i przedziały bezpośrednio sąsiadujące mają taką samą rozpiętość. Przedział taki nazywamy **przedziałem dominanty**.

Pozycyjne miary średnie
Dominanta (wartość modalna) – c.d.

- Przybliżoną wartość dominanty w szeregu rozdzielczym z przedziałami klasowymi wyznaczamy za pomocą następującego wzoru interpolacyjnego

$$Do = x_s + h_s \frac{n_s - n_{s-1}}{n_s - n_{s-1} + n_s - n_{s+1}},$$

gdzie:

x_s – początek przedziału dominanty,

h_s – rozpiętość przedziału dominanty,

n_s, n_{s-1}, n_{s+1} – liczebności odpowiednio przedziału dominanty, przedziału poprzedniego i następnego.

Pozycyjne miary średnie – Własności dominanty

1. Charakteryzuje jednostki o typowym poziomie cechy, a więc nie wszystkie jednostki.

Pozycyjne miary średnie – Własności dominanty

1. Charakteryzuje jednostki o typowym poziomie cechy, a więc nie wszystkie jednostki.
2. W przeciwieństwie do miar klasycznych dominantę możemy określić także dla cechy jakościowej.

W przypadku cechy ilościowej wymienić można dodatkowo takie własności dominanty, jak:

Pozycyjne miary średnie – Własności dominanty

1. Charakteryzuje jednostki o typowym poziomie cechy, a więc nie wszystkie jednostki.
2. W przeciwieństwie do miar klasycznych dominantę możemy określić także dla cechy jakościowej.

W przypadku cechy ilościowej wymienić można dodatkowo takie własności dominanty, jak:

3. W szeregu symetrycznym dominanta równa jest średniej arytmetycznej.

Pozycyjne miary średnie – Własności dominanty

1. Charakteryzuje jednostki o typowym poziomie cechy, a więc nie wszystkie jednostki.
2. W przeciwieństwie do miar klasycznych dominantę możemy określić także dla cechy jakościowej.

W przypadku cechy ilościowej wymienić można dodatkowo takie własności dominanty, jak:

3. W szeregu symetrycznym dominanta równa jest średniej arytmetycznej.
4. Dominantę można wyznaczać ze wzoru przybliżonego, gdy przedział dominanty oraz przedziały poprzedni i następny mają jednakową rozpiętość.

Pozycyjne miary średnie – Własności dominanty

1. Charakteryzuje jednostki o typowym poziomie cechy, a więc nie wszystkie jednostki.
2. W przeciwieństwie do miar klasycznych dominantę możemy określić także dla cechy jakościowej.

W przypadku cechy ilościowej wymienić można dodatkowo takie własności dominanty, jak:

3. W szeregu symetrycznym dominanta równa jest średniej arytmetycznej.
4. Dominantę można wyznaczać ze wzoru przybliżonego, gdy przedział dominanty oraz przedziały poprzedni i następny mają jednakową rozpiętość.
5. Dominantę można wyznaczać w szeregach rozdzielczych z otwartymi przedziałami klasowymi (o ile przedziały te nie sąsiadują bezpośrednio z przedziałem dominanty).

Pozycyjne miary średnie
Dominanta (wartość modalna)

Przykład 4.

- Załóżmy, że badamy kolor oczu osób zamieszkujących kraje skandynawskie i afrykańskie.

Pozycyjne miary średnie
Dominanta (wartość modalna)

Przykład 4.

- Załóżmy, że badamy kolor oczu osób zamieszkujących kraje skandynawskie i afrykańskie.
- Wówczas prawdopodobnie okaże się, że typowym (tj. dominującym) kolorem oczu wśród mieszkańców Skandynawii jest kolor niebieski, a wśród mieszkańców Afryki – brązowy.

Pozycyjne miary średnie – Kwartyle

- **Kwartyl pierwszy** Q_1 dzieli uporządkowaną niemalejąco zbiorowość na dwie części w ten sposób, że 25% jednostek ma wartości cechy nie większe niż kwartyl pierwszy Q_1, a 75% jednostek ma wartości cechy nie mniejsze niż Q_1.

Pozycyjne miary średnie – Kwartyle

- **Kwartyl pierwszy** Q_1 dzieli uporządkowaną niemalejąco zbiorowość na dwie części w ten sposób, że 25% jednostek ma wartości cechy nie większe niż kwartyl pierwszy Q_1, a 75% jednostek ma wartości cechy nie mniejsze niż Q_1.
- **Kwartyl drugi** Q_2 (nazywany także medianą i oznaczany symbolem Me) dzieli uporządkowaną niemalejąco zbiorowość na dwie części w ten sposób, że 50% jednostek ma wartości cechy nie większe niż mediana Me, a pozostałe 50% jednostek ma wartości cechy nie mniejsze niż Me.

Pozycyjne miary średnie – Kwartyle

- **Kwartyl pierwszy** Q_1 dzieli uporządkowaną niemalejąco zbiorowość na dwie części w ten sposób, że 25% jednostek ma wartości cechy nie większe niż kwartyl pierwszy Q_1, a 75% jednostek ma wartości cechy nie mniejsze niż Q_1.
- **Kwartyl drugi** Q_2 (nazywany także medianą i oznaczany symbolem Me) dzieli uporządkowaną niemalejąco zbiorowość na dwie części w ten sposób, że 50% jednostek ma wartości cechy nie większe niż mediana Me, a pozostałe 50% jednostek ma wartości cechy nie mniejsze niż Me.
- **Kwartyl trzeci** Q_3 dzieli uporządkowaną niemalejąco zbiorowość na dwie części w ten sposób, że 75% jednostek ma wartości cechy nie większe niż kwartyl trzeci Q_3, a 25% jednostek ma wartości cechy nie mniejsze niż Q_3.

Pozycyjne miary średnie
Kwartyle

- W szeregach szczegółowych medianę obliczamy ze wzoru

$$Me = \begin{cases} \frac{x_{n/2}+x_{(n+2)/2}}{2}, & \text{gdy } n \text{ jest parzyste,} \\ x_{(n+1)/2}, & \text{gdy } n \text{ jest nieparzyste,} \end{cases}$$

gdzie $x_{(n+1)/2}$, $x_{n/2}$, $x_{(n+2)/2}$ oznaczają wartości cechy dla tych jednostek zbiorowości, które w uporządkowanym (niemalejąco lub nierosnąco) szeregu znajdują się na miejscach o numerach odpowiednio

$$\frac{n+1}{2}, \quad \frac{n}{2}, \quad \frac{n+2}{2}.$$

Pozycyjne miary średnie
Kwartyle w szeregu szczegółowym i rozdzielczym punktowym

- Wyznaczanie kwartyla pierwszego i trzeciego z szeregu szczegółowego lub rozdzielczego punktowego rozpoczynamy od znalezienia mediany, która dzieli uporządkowaną zbiorowość na połowy.

Pozycyjne miary średnie
Kwartyle w szeregu szczegółowym i rozdzielczym punktowym

- Wyznaczanie kwartyla pierwszego i trzeciego z szeregu szczegółowego lub rozdzielczego punktowego rozpoczynamy od znalezienia mediany, która dzieli uporządkowaną zbiorowość na połowy.
- Dla pierwszej cześci (tj. dla połowy obserwacji nie większych od mediany) ponownie wyznaczamy medianę. Wyznaczona wartość będzie odpowiadała kwartylowi pierwszemu Q_1. Z kolei mediana wyznaczona dla drugiej części – kwartylowi trzeciemu Q_3.

Pozycyjne miary średnie
Kwartyle w szeregu szczegółowym i rozdzielczym punktowym

- Wyznaczanie kwartyla pierwszego i trzeciego z szeregu szczegółowego lub rozdzielczego punktowego rozpoczynamy od znalezienia mediany, która dzieli uporządkowaną zbiorowość na połowy.
- Dla pierwszej cześci (tj. dla połowy obserwacji nie większych od mediany) ponownie wyznaczamy medianę. Wyznaczona wartość będzie odpowiadała kwartylowi pierwszemu Q_1. Z kolei mediana wyznaczona dla drugiej części – kwartylowi trzeciemu Q_3.
- W przypadku szeregu rozdzielczego punktowego odnalezienie obserwacji reprezentujących kwartyle ułatwia kumulacja liczebności, która polega na sumowaniu kolejnych liczności n_i w szeregu.

Pozycyjne miary średnie
Kwartyle w szeregu rozdzielczym z przedziałami klasowymi

- Obliczanie kwartyli w szeregach rozdzielczych z przedziałami klasowymi opiera się na wzorach przybliżonych.

Pozycyjne miary średnie
Kwartyle w szeregu rozdzielczym z przedziałami klasowymi

- Obliczanie kwartyli w szeregach rozdzielczych z przedziałami klasowymi opiera się na wzorach przybliżonych.
- W pierwszym kroku odnajdujemy przedziały, do których należą jednostki o numerach $\frac{1}{4}n$, $\frac{1}{2}n$, $\frac{3}{4}n$. Przedziały te nazywamy odpowiednio przedziałem kwartyla pierwszego, przedziałem mediany i przedziałem kwartyla trzeciego.

Pozycyjne miary średnie
Kwartyle w szeregu rozdzielczym z przedziałami klasowymi

- Obliczanie kwartyli w szeregach rozdzielczych z przedziałami klasowymi opiera się na wzorach przybliżonych.
- W pierwszym kroku odnajdujemy przedziały, do których należą jednostki o numerach $\frac{1}{4}n$, $\frac{1}{2}n$, $\frac{3}{4}n$. Przedziały te nazywamy odpowiednio przedziałem kwartyla pierwszego, przedziałem mediany i przedziałem kwartyla trzeciego.
- W następnym kroku obliczamy poszczególne kwartyle.

Pozycyjne miary średnie
Kwartyle w szeregu rozdzielczym z przedziałami klasowymi – c.d.

- Kwartyl pierwszy wyznaczamy z następującego wzoru przybliżonego

$$Q_1 = x_s + \frac{h_s}{n_s}\left(\frac{1}{4}n - \sum_{i=1}^{s-1} n_i\right),$$

gdzie:

x_s – początek przedziału kwartyla pierwszego,

h_s – rozpiętość przedziału kwartyla pierwszego,

n_s – liczebność przedziału kwartyla pierwszego,

$\sum_{i=1}^{s-1} n_i$ – liczebność skumulowana od przedziału pierwszego do przedziału poprzedzającego przedział kwartyla pierwszego.

Pozycyjne miary średnie
Kwartyle w szeregu rozdzielczym z przedziałami klasowymi – c.d.

- W analogiczny sposób wyznaczamy Me i Q_3

$$Me = x_s + \frac{h_s}{n_s}\left(\frac{1}{2}n - \sum_{i=1}^{s-1} n_i\right),$$

$$Q_3 = x_s + \frac{h_s}{n_s}\left(\frac{3}{4}n - \sum_{i=1}^{s-1} n_i\right),$$

przy czym x_s, h_s, n_s w tych wzorach odnoszą się do odpowiednio przedziału mediany lub przedziału kwartyla trzeciego.

Miary zmienności

Dlaczego konieczne jest obliczanie miar zmienności dla zbioru danych?

Przykład 5.

- Wyobraźmy sobie, że mamy przeprowadzić zajęcia z matematyki w pewnej klasie uczniów liczących 10 osób.

Miary zmienności

Dlaczego konieczne jest obliczanie miar zmienności dla zbioru danych?

Przykład 5.

- Wyobraźmy sobie, że mamy przeprowadzić zajęcia z matematyki w pewnej klasie uczniów liczących 10 osób.
- Od dyrektora szkoły wiemy, że średni iloraz inteligencji w tej klasie jest równy 100.

Miary zmienności

Dlaczego konieczne jest obliczanie miar zmienności dla zbioru danych?

Przykład 5.

- Wyobraźmy sobie, że mamy przeprowadzić zajęcia z matematyki w pewnej klasie uczniów liczących 10 osób.
- Od dyrektora szkoły wiemy, że średni iloraz inteligencji w tej klasie jest równy 100.
- Mało obeznany metodologicznie nauczyciel może uznać tę informację za sygnał, że wystarczy przygotować zadania dla ”przeciętniaków”.

Miary zmienności

Dlaczego konieczne jest obliczanie miar zmienności dla zbioru danych?

Przykład 5.

- Wyobraźmy sobie, że mamy przeprowadzić zajęcia z matematyki w pewnej klasie uczniów liczących 10 osób.
- Od dyrektora szkoły wiemy, że średni iloraz inteligencji w tej klasie jest równy 100.
- Mało obeznany metodologicznie nauczyciel może uznać tę informację za sygnał, że wystarczy przygotować zadania dla "przeciętniaków".
- Ale czy takie rozumowanie jest poprawne?

Miary zmienności

Dlaczego konieczne jest obliczanie miar zmienności dla zbioru danych?

Miary zmienności

Dlaczego konieczne jest obliczanie miar zmienności dla zbioru danych?

- Wiemy, że średnia może powstać z różnych danych. Jeśli w klasie wszyscy uczniowie mają IQ równe 100, to średnia też wyniesie 100. Powiemy wówczas, że zbiorowość uczniów jest **jednorodna**.

Miary zmienności

Dlaczego konieczne jest obliczanie miar zmienności dla zbioru danych?

- Wiemy, że średnia może powstać z różnych danych. Jeśli w klasie wszyscy uczniowie mają IQ równe 100, to średnia też wyniesie 100. Powiemy wówczas, że zbiorowość uczniów jest **jednorodna**.
- Ale średni iloraz inteligencji równy 100 można otrzymać także wtedy, gdy w klasie (przykładowo) połowa uczniów ma iloraz 120, a druga połowa 80. Z prostych rachunków wynika, że średnia jest tu też równa 100: $\frac{120 \cdot 5 + 80 \cdot 5}{10} = 100$. Tym razem jednak zbiorowość uczniów jest **heterogeniczna**.

Miary zmienności

Dlaczego konieczne jest obliczanie miar zmienności dla zbioru danych?

- Wiemy, że średnia może powstać z różnych danych. Jeśli w klasie wszyscy uczniowie mają IQ równe 100, to średnia też wyniesie 100. Powiemy wówczas, że zbiorowość uczniów jest **jednorodna**.
- Ale średni iloraz inteligencji równy 100 można otrzymać także wtedy, gdy w klasie (przykładowo) połowa uczniów ma iloraz 120, a druga połowa 80. Z prostych rachunków wynika, że średnia jest tu też równa 100: $\frac{120 \cdot 5 + 80 \cdot 5}{10} = 100$. Tym razem jednak zbiorowość uczniów jest **heterogeniczna**.
- Nauczyciel powinien zerknąć na indywidualne wartości IQ, aby ocenić, czy w klasie są sami "przeciętniacy", czy też zarówno "geniusze", jak i "słabeusze".

Miary zmienności
Podział

- Miary, które pozwalają ocenić stopień heterogeniczności danej zbiorowości (czyli stopień zróżnicowania) nazywamy **miarami zmienności** lub zamiennie – miarami zróżnicowania, dyspersji, rozproszenia.

Miary zmienności
Podział

- Miary, które pozwalają ocenić stopień heterogeniczności danej zbiorowości (czyli stopień zróżnicowania) nazywamy **miarami zmienności** lub zamiennie – miarami zróżnicowania, dyspersji, rozproszenia.
- Miary zmienności dzielimy na bezwzględne i względne.

Miary zmienności
Podział

- Miary, które pozwalają ocenić stopień heterogeniczności danej zbiorowości (czyli stopień zróżnicowania) nazywamy **miarami zmienności** lub zamiennie – miarami zróżnicowania, dyspersji, rozproszenia.
- Miary zmienności dzielimy na bezwzględne i względne.
- Do **miar bezwzględnych** zaliczamy:

Miary zmienności
Podział

- Miary, które pozwalają ocenić stopień heterogeniczności danej zbiorowości (czyli stopień zróżnicowania) nazywamy **miarami zmienności** lub zamiennie – miarami zróżnicowania, dyspersji, rozproszenia.
- Miary zmienności dzielimy na bezwzględne i względne.
- Do **miar bezwzględnych** zaliczamy:
  1. **Klasyczne miary zróżnicowania**, w tym:

Miary zmienności
Podział

- Miary, które pozwalają ocenić stopień heterogeniczności danej zbiorowości (czyli stopień zróżnicowania) nazywamy **miarami zmienności** lub zamiennie – miarami zróżnicowania, dyspersji, rozproszenia.
- Miary zmienności dzielimy na bezwzględne i względne.
- Do **miar bezwzględnych** zaliczamy:
  1. **Klasyczne miary zróżnicowania**, w tym:
     a. odchylenie przeciętne d_x,

Miary zmienności
Podział

- Miary, które pozwalają ocenić stopień heterogeniczności danej zbiorowości (czyli stopień zróżnicowania) nazywamy **miarami zmienności** lub zamiennie – miarami zróżnicowania, dyspersji, rozproszenia.
- Miary zmienności dzielimy na bezwzględne i względne.
- Do **miar bezwzględnych** zaliczamy:
  1. **Klasyczne miary zróżnicowania**, w tym:
     a. odchylenie przeciętne d_x,
     b. wariancję s_x^2,

Miary zmienności
Podział

- Miary, które pozwalają ocenić stopień heterogeniczności danej zbiorowości (czyli stopień zróżnicowania) nazywamy **miarami zmienności** lub zamiennie – miarami zróżnicowania, dyspersji, rozproszenia.
- Miary zmienności dzielimy na bezwzględne i względne.
- Do **miar bezwzględnych** zaliczamy:
  1. **Klasyczne miary zróżnicowania**, w tym:
     a. odchylenie przeciętne d_x,
     b. wariancję s_x^2,
     c. odchylenie standardowe s_x.

Miary zmienności
Podział

- Miary, które pozwalają ocenić stopień heterogeniczności danej zbiorowości (czyli stopień zróżnicowania) nazywamy **miarami zmienności** lub zamiennie – miarami zróżnicowania, dyspersji, rozproszenia.
- Miary zmienności dzielimy na bezwzględne i względne.
- Do **miar bezwzględnych** zaliczamy:
  1. **Klasyczne miary zróżnicowania**, w tym:
     a. odchylenie przeciętne d_x,
     b. wariancję s_x^2,
     c. odchylenie standardowe s_x.
  2. **Pozycyjne miary zróżnicowania**, w tym:

Miary zmienności
Podział

- Miary, które pozwalają ocenić stopień heterogeniczności danej zbiorowości (czyli stopień zróżnicowania) nazywamy **miarami zmienności** lub zamiennie – miarami zróżnicowania, dyspersji, rozproszenia.
- Miary zmienności dzielimy na bezwzględne i względne.
- Do **miar bezwzględnych** zaliczamy:
  1. **Klasyczne miary zróżnicowania**, w tym:
     a. odchylenie przeciętne d_x,
     b. wariancję s_x^2,
     c. odchylenie standardowe s_x.
  2. **Pozycyjne miary zróżnicowania**, w tym:
     a. rozstęp R_x,

Miary zmienności
Podział

- Miary, które pozwalają ocenić stopień heterogeniczności danej zbiorowości (czyli stopień zróżnicowania) nazywamy **miarami zmienności** lub zamiennie – miarami zróżnicowania, dyspersji, rozproszenia.
- Miary zmienności dzielimy na bezwzględne i względne.
- Do **miar bezwzględnych** zaliczamy:
  1. **Klasyczne miary zróżnicowania**, w tym:
     a. odchylenie przeciętne d_x,
     b. wariancję s_x^2,
     c. odchylenie standardowe s_x.
  2. **Pozycyjne miary zróżnicowania**, w tym:
     a. rozstęp R_x,
     b. odchylenie ćwiartkowe Q_x.

Miary zmienności
Podział – c.d.

Miary zmienności
Podział – c.d.

Do **miar względnych** zaliczamy:

Miary zmienności
Podział – c.d.

Do **miar względnych** zaliczamy:

1. **Klasyczne współczynniki zmienności**, w tym:

Miary zmienności
Podział – c.d.

Do **miar względnych** zaliczamy:

1. **Klasyczne współczynniki zmienności**, w tym:
   a. współczynnik zmienności oparty na odchyleniu przeciętnym V_{d_x},

Miary zmienności
Podział – c.d.

Do **miar względnych** zaliczamy:

1. **Klasyczne współczynniki zmienności**, w tym:
   a. współczynnik zmienności oparty na odchyleniu przeciętnym V_{d_x},
   b. współczynnik zmienności oparty na odchyleniu standardowym V_{s_x}.

Miary zmienności
Podział – c.d.

Do **miar względnych** zaliczamy:

1. **Klasyczne współczynniki zmienności**, w tym:
   a. współczynnik zmienności oparty na odchyleniu przeciętnym V_{d_x},
   b. współczynnik zmienności oparty na odchyleniu standardowym V_{s_x}.
2. **Pozycyjny współczynnik zmienności**:

Miary zmienności
Podział – c.d.

Do **miar względnych** zaliczamy:

1. **Klasyczne współczynniki zmienności**, w tym:
   a. współczynnik zmienności oparty na odchyleniu przeciętnym V_{d_x},
   b. współczynnik zmienności oparty na odchyleniu standardowym V_{s_x}.
2. **Pozycyjny współczynnik zmienności**:
   a. współczynnik zmienności oparty na odchyleniu ćwiartkowym Q_x.

Klasyczne miary zmienności
Jak obliczyć odchylenie przeciętne?

Przykład 6.

- W odniesieniu do przykładu 5 (dotyczącego IQ) załóżmy dalej, że ilorazy inteligencji w 10-osobowej grupie uczniów kształtował się następująco:

$$85, 85, 95, 95, 95, 100, 105, 110, 115, 115.$$

Klasyczne miary zmienności
Jak obliczyć odchylenie przeciętne?

Przykład 6.

- W odniesieniu do przykładu 5 (dotyczącego IQ) załóżmy dalej, że ilorazy inteligencji w 10-osobowej grupie uczniów kształtował się następująco:

$$85, 85, 95, 95, 95, 100, 105, 110, 115, 115.$$

- Średnia wartość IQ w tej grupie wynosi 100, ale ma tu miejsce spore zróżnicowanie pomiędzy uczniami.

Klasyczne miary zmienności
Jak obliczyć odchylenie przeciętne?

Przykład 6.

- W odniesieniu do przykładu 5 (dotyczącego IQ) załóżmy dalej, że ilorazy inteligencji w 10-osobowej grupie uczniów kształtował się następująco:

$$85, 85, 95, 95, 95, 100, 105, 110, 115, 115.$$

- Średnia wartość IQ w tej grupie wynosi 100, ale ma tu miejsce spore zróżnicowanie pomiędzy uczniami.
- Oznaczmy poszczególne wyniki symbolami

$$x_1, x_2, \ldots, x_{10},$$

natomiast średnią z tych wyników symbolem $\bar{x}$.

Klasyczne miary zmienności
Jak obliczyć odchylenie przeciętne?

Mamy więc następujące wartości i ich odchylenia od średniej

$$\begin{aligned}
x_1 &= 85, & x_1 - \bar{x} &= -15,\\
x_2 &= 85, & x_2 - \bar{x} &= -15,\\
x_3 &= 95, & x_3 - \bar{x} &= -5,\\
x_4 &= 95, & x_4 - \bar{x} &= -5,\\
x_5 &= 95, & x_5 - \bar{x} &= -5,\\
x_6 &= 100, & x_6 - \bar{x} &= 0,\\
x_7 &= 105, & x_7 - \bar{x} &= 5,\\
x_8 &= 110, & x_8 - \bar{x} &= 10,\\
x_9 &= 115, & x_9 - \bar{x} &= 15,\\
x_{10} &= 115, & x_{10} - \bar{x} &= 15.
\end{aligned}$$

Klasyczne miary zmienności
Jak obliczyć odchylenie przeciętne?

Mamy więc następujące wartości i ich odchylenia od średniej

$$\begin{aligned}
x_1 &= 85, & x_1 - \bar{x} &= -15,\\
x_2 &= 85, & x_2 - \bar{x} &= -15,\\
x_3 &= 95, & x_3 - \bar{x} &= -5,\\
x_4 &= 95, & x_4 - \bar{x} &= -5,\\
x_5 &= 95, & x_5 - \bar{x} &= -5,\\
x_6 &= 100, & x_6 - \bar{x} &= 0,\\
x_7 &= 105, & x_7 - \bar{x} &= 5,\\
x_8 &= 110, & x_8 - \bar{x} &= 10,\\
x_9 &= 115, & x_9 - \bar{x} &= 15,\\
x_{10} &= 115, & x_{10} - \bar{x} &= 15.
\end{aligned}$$

Ale suma wszystkich odchyleń jest równa **0**!

Klasyczne miary zmienności
Jak obliczyć odchylenie przeciętne?

Mamy więc następujące wartości i ich odchylenia od średniej

$$
\begin{aligned}
x_1 &= 85, & x_1 - \bar{x} &= -15,\\
x_2 &= 85, & x_2 - \bar{x} &= -15,\\
x_3 &= 95, & x_3 - \bar{x} &= -5,\\
x_4 &= 95, & x_4 - \bar{x} &= -5,\\
x_5 &= 95, & x_5 - \bar{x} &= -5,\\
x_6 &= 100, & x_6 - \bar{x} &= 0,\\
x_7 &= 105, & x_7 - \bar{x} &= 5,\\
x_8 &= 110, & x_8 - \bar{x} &= 10,\\
x_9 &= 115, & x_9 - \bar{x} &= 15,\\
x_{10} &= 115, & x_{10} - \bar{x} &= 15.
\end{aligned}
$$

Ale suma wszystkich odchyleń jest równa **0**!
Możemy jednak obliczyć sumę odchyleń bezwzględnych, która w tym przykładzie wynosi **90**, a następnie podzielić przez ich liczbę (tj. przez 10). W ten sposób otrzymamy odchylenie przeciętne d_x równe **9**.

Klasyczne miary zmienności
Odchylenie przeciętne w szeregu prostym

A zatem, jeśli dysponujemy zbiorem danych (o liczebności *n*) zestawionych w szereg szczegółowy, prosty, to **odchylenie przeciętne** obliczamy ze wzoru:

$$d_x = \frac{|x_1 - \bar{x}| + |x_2 - \bar{x}| + |x_3 - \bar{x}| + \ldots + |x_n - \bar{x}|}{n}$$

Klasyczne miary zmienności
Odchylenie przeciętne w szeregu prostym

A zatem, jeśli dysponujemy zbiorem danych (o liczebności *n*) zestawionych w szereg szczegółowy, prosty, to **odchylenie przeciętne** obliczamy ze wzoru:

$$d_x = \frac{|x_1 - \bar{x}| + |x_2 - \bar{x}| + |x_3 - \bar{x}| + \ldots + |x_n - \bar{x}|}{n}$$

Formułę tę zapisujemy w skrócie wzorem:

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

Klasyczne miary zmienności
Odchylenie przeciętne w szeregu prostym

A zatem, jeśli dysponujemy zbiorem danych (o liczebności *n*) zestawionych w szereg szczegółowy, prosty, to **odchylenie przeciętne** obliczamy ze wzoru:

$$d_x = \frac{|x_1 - \bar{x}| + |x_2 - \bar{x}| + |x_3 - \bar{x}| + \ldots + |x_n - \bar{x}|}{n}$$

Formułę tę zapisujemy w skrócie wzorem:

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

Interpretacja: Jest to średnia arytmetyczna z bezwzględnych odchyleń wartości cechy od jej średniej arytmetycznej.

Klasyczne miary zmienności

Odchylenie przeciętne w szeregu rozdzielczym punktowym – c.d. przykładu 6

Załóżmy, że dane z przykładu 6 pogrupowane zostały w szereg rozdzielczy punktowy:

wartość IQ x_i	liczby uczniów n_i
85	2
95	3
100	1
105	1
110	1
115	2

Klasyczne miary zmienności
Odchylenie przeciętne w szeregu rozdzielczym punktowym – c.d. przykładu 6

Załóżmy, że dane z przykładu 6 pogrupowane zostały w szereg rozdzielczy punktowy:

wartość IQ x_i	liczby uczniów n_i
85	2
95	3
100	1
105	1
110	1
115	2

Zauważymy, że są to te same dane, ale inaczej przedstawione. Odchylenie przeciętne dla tego szeregu powinno pozostać więc bez zmian.

Klasyczne miary zmienności

Odchylenie przeciętne w szeregu rozdzielczym punktowym – c.d. przykładu 6

Aby je obliczyć, wygodnie jest przeprowadzić obliczenia pośrednie w dodatkowych kolumnach tablicy.

wartość IQ x_i	liczby uczniów n_i	odchylenia bezwzględne $\lvert x_i - \bar{x} \rvert$	ważone odchylenia $\lvert x_i - \bar{x} \rvert \cdot n_i$
85	2	15	30
95	3	5	15
100	1	0	0
105	1	5	5
110	1	10	10
115	2	15	30
Razem	10	$\times$	90

Klasyczne miary zmienności

Odchylenie przeciętne w szeregu rozdzielczym punktowym – c.d. przykładu 6

Aby je obliczyć, wygodnie jest przeprowadzić obliczenia pośrednie w dodatkowych kolumnach tablicy.

wartość IQ x_i	liczby uczniów n_i	odchylenia bezwzględne $\lvert x_i - \bar{x} \rvert$	ważone odchylenia $\lvert x_i - \bar{x} \rvert \cdot n_i$
85	2	15	30
95	3	5	15
100	1	0	0
105	1	5	5
110	1	10	10
115	2	15	30
Razem	10	$\times$	90

Mamy: $d_x = \frac{90}{10} = 9$. Ogólny wzór: $d_x = \frac{1}{n} \sum_{i=1}^{k} |x_i - \bar{x}| \cdot n_i$

Klasyczne miary zmienności

Odchylenie przeciętne w szeregu rozdzielczym z przedziałami klasowymi – c.d. przykładu 6

Pogrupujmy dane z poprzedniego szeregu punktowego w szereg rozdzielczy z przedziałami klasowymi o rozpiętości 10.

przedziały wartości IQ x_i	**liczby uczniów** n_i
(85, 95]	5
(95,105]	2
(105,115]	3

Klasyczne miary zmienności

Odchylenie przeciętne w szeregu rozdzielczym z przedziałami klasowymi – c.d. przykładu 6

Pogrupujmy dane z poprzedniego szeregu punktowego w szereg rozdzielczy z przedziałami klasowymi o rozpiętości 10.

przedziały wartości IQ x_i	liczby uczniów n_i
(85, 95]	5
(95,105]	2
(105,115]	3

W tym przypadku obliczona wartość odchylenie przeciętnego będzie tylko **przybliżeniem** rzeczywistej wartości, ponieważ nie mamy pełnej informacji o poziomie IQ dla wszystkich uczniów.

Klasyczne miary zmienności

Odchylenie przeciętne w szeregu rozdzielczym z przedziałami klasowymi – c.d. przykładu 6

Pogrupujmy dane z poprzedniego szeregu punktowego w szereg rozdzielczy z przedziałami klasowymi o rozpiętości 10.

przedziały wartości IQ x_i	**liczby uczniów** n_i
(85, 95]	5
(95,105]	2
(105,115]	3

W tym przypadku obliczona wartość odchylenie przeciętnego będzie tylko **przybliżeniem** rzeczywistej wartości, ponieważ nie mamy pełnej informacji o poziomie IQ dla wszystkich uczniów. Aby znaleźć d_x musimy przyjąć dla każdego przedziału klasowego jakąś uśrednioną wartość IQ. Wartościami tymi niech będą środki poszczególnych przedziałów.

Klasyczne miary zmienności

Odchylenie przeciętne w szeregu rozdzielczym z przedziałami klasowymi – c.d.

przedziały IQ x_i	liczby uczniów n_i	środki przedz. $\dot{x}_i$	$\dot{x}_i \cdot n_i$	odchylenia bezwzględne $\lvert\dot{x}_i - \bar{x}\rvert$	ważone odchylenia $\lvert\dot{x}_i - \bar{x}\rvert \cdot n_i$
85–95	5	90	450	8	40
95–105	2	100	200	2	4
105–115	3	110	330	12	36
Razem	10	$\times$	980	$\times$	80

Klasyczne miary zmienności

Odchylenie przeciętne w szeregu rozdzielczym z przedziałami klasowymi – c.d.

przedziały IQ x_i	liczby uczniów n_i	środki przedz. $\dot{x}_i$	$\dot{x}_i \cdot n_i$	odchylenia bezwzględne $\lvert\dot{x}_i - \bar{x}\rvert$	ważone odchylenia $\lvert\dot{x}_i - \bar{x}\rvert \cdot n_i$
85–95	5	90	450	8	40
95–105	2	100	200	2	4
105–115	3	110	330	12	36
Razem	10	$\times$	980	$\times$	80

Średnia arytmetyczna IQ obliczona na podstawie tego szeregu wynosi $\frac{980}{10} = 98$, a zatem jest tylko przybliżeniem faktycznej średniej. Dalej w obliczeniach przyjęto to przybliżenie.

Klasyczne miary zmienności

Odchylenie przeciętne w szeregu rozdzielczym z przedziałami klasowymi – c.d.

przedziały IQ x_i	liczby uczniów n_i	środki przedz. $\dot{x}_i$	$\dot{x}_i \cdot n_i$	odchylenia bezwzględne $\lvert\dot{x}_i - \bar{x}\rvert$	ważone odchylenia $\lvert\dot{x}_i - \bar{x}\rvert \cdot n_i$
85–95	5	90	450	8	40
95–105	2	100	200	2	4
105–115	3	110	330	12	36
Razem	10	×	980	×	80

Średnia arytmetyczna IQ obliczona na podstawie tego szeregu wynosi $\frac{980}{10} = 98$, a zatem jest tylko przybliżeniem faktycznej średniej. Dalej w obliczeniach przyjęto to przybliżenie. Uzyskujemy przybliżoną wartość odchylenia przeciętnego $\frac{80}{10} = 8$. Ogólnie wzór: $d_x = \frac{1}{n}\sum_{i=1}^{k} |\dot{x}_i - \bar{x}| \cdot n_i$

Klasyczne miary zmienności
Odchylenie przeciętne – wzory

Podsumowanie: Skorzystaliśmy z trzech formuł na d_x.

Klasyczne miary zmienności
Odchylenie przeciętne – wzory

Podsumowanie: Skorzystaliśmy z trzech formuł na d_x.

- W szeregu szczegółowym:

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

Klasyczne miary zmienności
Odchylenie przeciętne – wzory

Podsumowanie: Skorzystaliśmy z trzech formuł na d_x.

- W szeregu szczegółowym:

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

- W szeregu rozdzielczym punktowym:

$$d_x = \frac{1}{n}\sum_{i=1}^{k}|x_i - \bar{x}| \cdot n_i$$

Klasyczne miary zmienności
Odchylenie przeciętne – wzory

Podsumowanie: Skorzystaliśmy z trzech formuł na d_x.

- W szeregu szczegółowym:

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

- W szeregu rozdzielczym punktowym:

$$d_x = \frac{1}{n}\sum_{i=1}^{k}|x_i - \bar{x}| \cdot n_i$$

- W szeregu rozdzielczym z przedziałami klasowymi:

$$d_x = \frac{1}{n}\sum_{i=1}^{k}|\dot{x}_i - \bar{x}| \cdot n_i$$

Klasyczne miary zmienności – Wariancja

W analogiczny sposób konstruujemy wzory na inną klasyczną miarę zmienności, zwaną **wariancją**.

Klasyczne miary zmienności – Wariancja

W analogiczny sposób konstruujemy wzory na inną klasyczną miarę zmienności, zwaną **wariancją**.

- W szeregu szczegółowym:

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

Klasyczne miary zmienności – Wariancja

W analogiczny sposób konstruujemy wzory na inną klasyczną miarę zmienności, zwaną **wariancją**.

- W szeregu szczegółowym:

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

- W szeregu rozdzielczym punktowym:

$$s_x^2 = \frac{1}{n}\sum_{i=1}^{k}(x_i - \bar{x})^2 \cdot n_i$$

Klasyczne miary zmienności – Wariancja

W analogiczny sposób konstruujemy wzory na inną klasyczną miarę zmienności, zwaną **wariancją**.

- W szeregu szczegółowym:

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

- W szeregu rozdzielczym punktowym:

$$s_x^2 = \frac{1}{n}\sum_{i=1}^{k}(x_i - \bar{x})^2 \cdot n_i$$

- W szeregu rozdzielczym z przedziałami klasowymi:

$$s_x^2 = \frac{1}{n}\sum_{i=1}^{k}(\dot{x}_i - \bar{x})^2 \cdot n_i$$

Klasyczne miary zmienności – Odchylenie standardowe

Pierwiastek kw. z wariancji – **odchylenie standardowe.**

Klasyczne miary zmienności – Odchylenie standardowe

Pierwiastek kw. z wariancji – **odchylenie standardowe.**

- W szeregu szczegółowym:

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

Klasyczne miary zmienności – Odchylenie standardowe

Pierwiastek kw. z wariancji – **odchylenie standardowe.**

- W szeregu szczegółowym:

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

- W szeregu rozdzielczym punktowym:

$$s_x = \sqrt{\frac{1}{n}\sum_{i=1}^{k}(x_i - \bar{x})^2 \cdot n_i}$$

Klasyczne miary zmienności – Odchylenie standardowe

Pierwiastek kw. z wariancji – **odchylenie standardowe.**

- W szeregu szczegółowym:

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

- W szeregu rozdzielczym punktowym:

$$s_x = \sqrt{\frac{1}{n}\sum_{i=1}^{k}(x_i - \bar{x})^2 \cdot n_i}$$

- W szeregu rozdzielczym z przedziałami klasowymi:

$$s_x = \sqrt{\frac{1}{n}\sum_{i=1}^{k}(\dot{x}_i - \bar{x})^2 \cdot n_i}$$

Klasyczne miary zmienności

Wariancja i odch. stand. w szeregu szczegółowym – c.d. przykładu 6

W przypadku szeregu szczegółowego z przykładu 6 mamy następujące wartości IQ i kwadraty ich odchyleń od średniej:

$$
\begin{aligned}
x_1 &= 85, & (x_1-\bar{x})^2 &= 225,\\
x_2 &= 85, & (x_2-\bar{x})^2 &= 225,\\
x_3 &= 95, & (x_3-\bar{x})^2 &= 25,\\
x_4 &= 95, & (x_4-\bar{x})^2 &= 25,\\
x_5 &= 95, & (x_5-\bar{x})^2 &= 25,\\
x_6 &= 100, & (x_6-\bar{x})^2 &= 0,\\
x_7 &= 105, & (x_7-\bar{x})^2 &= 25,\\
x_8 &= 110, & (x_8-\bar{x})^2 &= 100,\\
x_9 &= 115, & (x_9-\bar{x})^2 &= 225,\\
x_{10} &= 115, & (x_{10}-\bar{x})^2 &= 225.
\end{aligned}
$$

Stąd otrzymujemy

$$s_x^2 = \frac{1}{10}\sum_{i=1}^{10}(x_i-\bar{x})^2 = \frac{1100}{10} = 110, \qquad s_x = \sqrt{110} \approx 10,5.$$

Klasyczne miary zmienności

Wariancja i odch. stand. w szeregu rozdzielczym punktowym – c.d. przykładu 6

wartość IQ x_i	liczby uczniów n_i	kwadraty odchyleń $(x_i - \bar{x})^2$	ważone kwadraty odchyleń $(x_i - \bar{x})^2 \cdot n_i$
85	2	225	450
95	3	25	75
100	1	0	0
105	1	25	25
110	1	100	100
115	2	225	450
Razem	10	×	1100

Klasyczne miary zmienności

Wariancja i odch. stand. w szeregu rozdzielczym punktowym – c.d. przykładu 6

wartość IQ x_i	**liczby uczniów** n_i	**kwadraty odchyleń** $(x_i - \bar{x})^2$	**ważone kwadraty odchyleń** $(x_i - \bar{x})^2 \cdot n_i$
85	2	225	450
95	3	25	75
100	1	0	0
105	1	25	25
110	1	100	100
115	2	225	450
Razem	10	$\times$	1100

Wariancja równa jest więc $s_x^2 = \frac{1100}{10} = 110$, a odchylenie standardowe $s_x = \sqrt{110} \approx 10,5$.

Klasyczne miary zmienności

Wariancja i odch. stand. w szeregu rozdzielczym z przedziałami klasowymi – c.d. przykładu 6

przedziały IQ x_i	liczby uczniów n_i	środki przedz. $\dot{x}_i$	kwadraty odchyleń $(\dot{x}_i - \bar{x})^2$	ważone kwadraty odchyleń $(\dot{x}_i - \bar{x})^2 \cdot n_i$
85– 95	5	90	64	320
95–105	2	100	4	8
105–115	3	110	144	432
Razem	10	×	×	760

Klasyczne miary zmienności

Wariancja i odch. stand. w szeregu rozdzielczym z przedziałami klasowymi – c.d. przykładu 6

przedziały IQ x_i	liczby uczniów n_i	środki przedz. $\dot{x}_i$	kwadraty odchyleń $(\dot{x}_i - \bar{x})^2$	ważone kwadraty odchyleń $(\dot{x}_i - \bar{x})^2 \cdot n_i$
85– 95	5	90	64	320
95–105	2	100	4	8
105–115	3	110	144	432
Razem	10	×	×	760

W obliczeniach przyjęto średnie IQ równe 98 (zamiast dokładnej wartości 100).

Klasyczne miary zmienności

Wariancja i odch. stand. w szeregu rozdzielczym z przedziałami klasowymi – c.d. przykładu 6

przedziały IQ x_i	liczby uczniów n_i	środki przedz. $\dot{x}_i$	kwadraty odchyleń $(\dot{x}_i - \bar{x})^2$	ważone kwadraty odchyleń $(\dot{x}_i - \bar{x})^2 \cdot n_i$
85– 95	5	90	64	320
95–105	2	100	4	8
105–115	3	110	144	432
Razem	10	$\times$	$\times$	760

W obliczeniach przyjęto średnie IQ równe 98 (zamiast dokładnej wartości 100).

Wariancja jest tu wyznaczona w przybliżeniu i wynosi $\frac{760}{10} = 76$. Podobnie, przybliżeniem odchylenia standardowego jest liczba $\sqrt{76} \approx 8,7$.

Pozycyjne miary zmienności

Kiedy obliczamy pozycyjne miary zmienności?

Pozycyjne miary zmienności

Kiedy obliczamy pozycyjne miary zmienności?

- Wariancja i odchylenie standardowe są miarami zróżnicowania, najczęściej stosowanymi w praktyce. Jednak nie zawsze istnieje możliwość ich obliczenia.

Pozycyjne miary zmienności
Kiedy obliczamy pozycyjne miary zmienności?

- Wariancja i odchylenie standardowe są miarami zróżnicowania, najczęściej stosowanymi w praktyce. Jednak nie zawsze istnieje możliwość ich obliczenia.
- Na przykład, gdy szereg rozdzielczy ma **otwarte przedziały klasowe** i nie jest możliwe znalezienie średniej arytmetycznej, wówczas nie jest możliwe także wyznaczenie klasycznych miar zmienności.

Pozycyjne miary zmienności

Kiedy obliczamy pozycyjne miary zmienności?

- Wariancja i odchylenie standardowe są miarami zróżnicowania, najczęściej stosowanymi w praktyce. Jednak nie zawsze istnieje możliwość ich obliczenia.
- Na przykład, gdy szereg rozdzielczy ma **otwarte przedziały klasowe** i nie jest możliwe znalezienie średniej arytmetycznej, wówczas nie jest możliwe także wyznaczenie klasycznych miar zmienności.
- Innym przykładem jest występowanie w zbiorze danych **obserwacji nietypowych (odstających)**. Wówczas często nie jest wskazane obliczanie średniej arytmetycznej, a tym samym także klasycznych miar zmienności.

Pozycyjne miary zmienności
Kiedy obliczamy pozycyjne miary zmienności?

- Wariancja i odchylenie standardowe są miarami zróżnicowania, najczęściej stosowanymi w praktyce. Jednak nie zawsze istnieje możliwość ich obliczenia.
- Na przykład, gdy szereg rozdzielczy ma **otwarte przedziały klasowe** i nie jest możliwe znalezienie średniej arytmetycznej, wówczas nie jest możliwe także wyznaczenie klasycznych miar zmienności.
- Innym przykładem jest występowanie w zbiorze danych **obserwacji nietypowych (odstających)**. Wówczas często nie jest wskazane obliczanie średniej arytmetycznej, a tym samym także klasycznych miar zmienności.
- W takich sytuacjach zamiast miar klasycznych obliczamy **miary pozycyjne**.

Pozycyjne miary zmienności
Rozstęp

- **Rozstęp** definiujemy jako różnicę między wartością największą i najmniejszą badanej cechy w zbiorowości, czyli

$$R_x = x_{max} - x_{min},$$

gdzie
$x_{max} = \max\{x_1, x_2, \ldots, x_n\}$,
$x_{min} = \min\{x_1, x_2, \ldots, x_n\}$.

Pozycyjne miary zmienności
Rozstęp

- **Rozstęp** definiujemy jako różnicę między wartością największą i najmniejszą badanej cechy w zbiorowości, czyli
$$R_x = x_{max} - x_{min},$$
gdzie
$x_{max} = \max\{x_1, x_2, \ldots, x_n\}$,
$x_{min} = \min\{x_1, x_2, \ldots, x_n\}$.

- Na podstawie danych z przykładu 6 mamy: $x_{min} = 85$, $x_{max} = 115$, zatem rozstęp wynosi $R_x = 30$.

Pozycyjne miary zmienności
Odchylenie ćwiartkowe

- **Ochylenie ćwiatkowe** Q_x jest miarą zróżnicowania opartą na kwartylach. Definiujemy je jako połowa różnicy pomiędzy trzecim a pierwszym kwartylem

$$Q_x = \frac{Q_3 - Q_1}{2}.$$

Pozycyjne miary zmienności
Odchylenie ćwiartkowe

- **Ochylenie ćwiatkowe** Q_x jest miarą zróżnicowania opartą na kwartylach. Definiujemy je jako połowa różnicy pomiędzy trzecim a pierwszym kwartylem

$$Q_x = \frac{Q_3 - Q_1}{2}.$$

- Odchylenie ćwiartkowe interpretujemy jako połowę rozpiętości przedziału, w którym znajduje się 50% jednostek skupionych najbliżej mediany.

Pozycyjne miary zmienności
Odchylenie ćwiartkowe

- **Ochylenie ćwiatkowe** Q_x jest miarą zróżnicowania opartą na kwartylach. Definiujemy je jako połowa różnicy pomiędzy trzecim a pierwszym kwartylem

$$Q_x = \frac{Q_3 - Q_1}{2}.$$

- Odchylenie ćwiartkowe interpretujemy jako połowę rozpiętości przedziału, w którym znajduje się 50% jednostek skupionych najbliżej mediany.
- Na wartość odchylenia ćwiartkowego nie mają wpływu wartości mniejsze od kwartyla pierwszego i wartości większe od kwartyla trzeciego.

Pozycyjne miary zmienności
Odchylenie ćwiartkowe

- **Ochylenie ćwiatkowe** Q_x jest miarą zróżnicowania opartą na kwartylach. Definiujemy je jako połowa różnicy pomiędzy trzecim a pierwszym kwartylem

$$Q_x = \frac{Q_3 - Q_1}{2}.$$

- Odchylenie ćwiartkowe interpretujemy jako połowę rozpiętości przedziału, w którym znajduje się 50% jednostek skupionych najbliżej mediany.
- Na wartość odchylenia ćwiartkowego nie mają wpływu wartości mniejsze od kwartyla pierwszego i wartości większe od kwartyla trzeciego.
- Miara ta zatem, w przeciwieństwie do klasycznych miar zróżnicowania, nie jest wrażliwa na wartości skrajne.

Pozycyjne miary zmienności
Odchylenie ćwiartkowe

- W przykładzie 6 mieliśmy następujące dane:

$$85, 85, 95, 95, 95, 100, 105, 110, 115, 115.$$

Pozycyjne miary zmienności
Odchylenie ćwiartkowe

- W przykładzie 6 mieliśmy następujące dane:

$$85, 85, 95, 95, 95, 100, 105, 110, 115, 115.$$

- Kwartyl pierwszy i trzeci są tu równe odpowiednio: $Q_1 = 95$, $Q_3 = 110$, a zatem odchylenie ćwiartkowe wynosi

$$Q_x = \frac{110 - 95}{2} = 7,5$$

Pozycyjne miary zmienności
Odchylenie ćwiartkowe

- W przykładzie 6 mieliśmy następujące dane:

$$85, 85, 95, 95, 95, 100, 105, 110, 115, 115.$$

- Kwartyl pierwszy i trzeci są tu równe odpowiednio: $Q_1 = 95$, $Q_3 = 110$, a zatem odchylenie ćwiartkowe wynosi

$$Q_x = \frac{110 - 95}{2} = 7,5$$

- Między miarami zróżnicowania obliczonymi dla tego samego szeregu zachodzą nierówności

$$Q_x < d_x < s_x.$$

Istotnie, w przykładzie $Q_x = 7,5$, $d_x = 9$, $s_x = 10,5$.

Względne miary zróżnicowania
Współczynniki zmienności – czyli rzecz o porównywaniu zmienności

Przykład 7.

- Załóżmy, że dwóch skoczków narciarskich wykonało na pewnych zawodach po 4 skoki (np. 2 podczas treningu i 2 podczas konkursu).

Względne miary zróżnicowania
Współczynniki zmienności – czyli rzecz o porównywaniu zmienności

Przykład 7.

- Załóżmy, że dwóch skoczków narciarskich wykonało na pewnych zawodach po 4 skoki (np. 2 podczas treningu i 2 podczas konkursu).
- Średnia długość skoków skoczka A wyniosła 130 m, a skoczka B – 110 m (a zatem skoczek A był lepszy!)

Względne miary zróżnicowania
Współczynniki zmienności – czyli rzecz o porównywaniu zmienności

Przykład 7.

- Załóżmy, że dwóch skoczków narciarskich wykonało na pewnych zawodach po 4 skoki (np. 2 podczas treningu i 2 podczas konkursu).
- Średnia długość skoków skoczka A wyniosła 130 m, a skoczka B – 110 m (a zatem skoczek A był lepszy!)
- Odchylenia standardowe długości skoków dla obu zawodników były jednakowe i wynosiły 2,5 m.

Względne miary zróżnicowania
Współczynniki zmienności – czyli rzecz o porównywaniu zmienności

Przykład 7.

- Załóżmy, że dwóch skoczków narciarskich wykonało na pewnych zawodach po 4 skoki (np. 2 podczas treningu i 2 podczas konkursu).
- Średnia długość skoków skoczka A wyniosła 130 m, a skoczka B – 110 m (a zatem skoczek A był lepszy!)
- Odchylenia standardowe długości skoków dla obu zawodników były jednakowe i wynosiły 2,5 m.
- **Pytanie:** Czy można powiedzieć, że pod względem regularności skoków obydwaj zawodnicy byli podobni?

Względne miary zróżnicowania
Współczynniki zmienności – czyli rzecz o porównywaniu zróżnicowania

- Wbrew pozorom odpowiedź, nie jest twierdząca. Pod względem regularności skoków zawodnik A także okazuje się być trochę lepszym!

Względne miary zróżnicowania
Współczynniki zmienności – czyli rzecz o porównywaniu zróżnicowania

- Wbrew pozorom odpowiedź, nie jest twierdząca. Pod względem regularności skoków zawodnik A także okazuje się być trochę lepszym!
- Odchylenie standardowe równe 2,5 m bowiem "więcej znaczy" w przypadku, gdy średnia odległość skoku wynosi 110 m, niż w przypadku średniej równej 130 m.

Względne miary zróżnicowania
Współczynniki zmienności – czyli rzecz o porównywaniu zróżnicowania

- Wbrew pozorom odpowiedź, nie jest twierdząca. Pod względem regularności skoków zawodnik A także okazuje się być trochę lepszym!
- Odchylenie standardowe równe 2,5 m bowiem ”więcej znaczy” w przypadku, gdy średnia odległość skoku wynosi 110 m, niż w przypadku średniej równej 130 m.
- Aby to liczbowo wykazać, wystarczy obliczyć iloraz odchylenia standardowego do średniej.

Względne miary zróżnicowania
Współczynniki zmienności – czyli rzecz o porównywaniu zróżnicowania

- Wbrew pozorom odpowiedź, nie jest twierdząca. Pod względem regularności skoków zawodnik A także okazuje się być trochę lepszym!
- Odchylenie standardowe równe 2,5 m bowiem ”więcej znaczy” w przypadku, gdy średnia odległość skoku wynosi 110 m, niż w przypadku średniej równej 130 m.
- Aby to liczbowo wykazać, wystarczy obliczyć iloraz odchylenia standardowego do średniej.
- Dla zawodnika A wspomniany iloraz wynosi $\frac{2{,}5}{130} \approx 0,019$, a dla zawodnika B $\frac{2{,}5}{110} \approx 0,023$, a więc dla zawodnika B zmienność skoków stanowi większy ułamek średniej.

Względne miary zróżnicowania

Współczynniki zmienności – czyli rzecz o porównywaniu zróżnicowania

- Wbrew pozorom odpowiedź, nie jest twierdząca. Pod względem regularności skoków zawodnik A także okazuje się być trochę lepszym!
- Odchylenie standardowe równe 2,5 m bowiem "więcej znaczy" w przypadku, gdy średnia odległość skoku wynosi 110 m, niż w przypadku średniej równej 130 m.
- Aby to liczbowo wykazać, wystarczy obliczyć iloraz odchylenia standardowego do średniej.
- Dla zawodnika A wspomniany iloraz wynosi $\frac{2{,}5}{130} \approx 0,019$, a dla zawodnika B $\frac{2{,}5}{110} \approx 0,023$, a więc dla zawodnika B zmienność skoków stanowi większy ułamek średniej.
- Obliczone ilorazy można wyrazić w procentach odpowiednio: $1,9\%$ i $2,3\%$.

Względne miary zmienności – Współczynniki zmienności

- Względne miary zróżnicowania (inaczej **współczynniki zmienności**), definiujemy jako ilorazy bezwzględnych miar zróżnicowania do odpowiednich miar średnich.

Względne miary zmienności – Współczynniki zmienności

- Względne miary zróżnicowania (inaczej **współczynniki zmienności**), definiujemy jako ilorazy bezwzględnych miar zróżnicowania do odpowiednich miar średnich.
- Współczynniki te stosujemy przy porównaniach.

Względne miary zmienności – Współczynniki zmienności

- Względne miary zróżnicowania (inaczej **współczynniki zmienności**), definiujemy jako ilorazy bezwzględnych miar zróżnicowania do odpowiednich miar średnich.
- Współczynniki te stosujemy przy porównaniach.
  - współczynik zmienności V_{d_x} oparty na odchyleniu przeciętnym

$$V_{d_x} = \frac{d_x}{\bar{x}} \cdot 100\%,$$

Względne miary zmienności – Współczynniki zmienności

- Względne miary zróżnicowania (inaczej **współczynniki zmienności**), definiujemy jako ilorazy bezwzględnych miar zróżnicowania do odpowiednich miar średnich.
- Współczynniki te stosujemy przy porównaniach.
  - współczynnik zmienności V_{d_x} oparty na odchyleniu przeciętnym

$$V_{d_x} = \frac{d_x}{\bar{x}} \cdot 100\%,$$

  - współczynnik zmienności V_{s_x} oparty na odchyleniu standardowym

$$V_{s_x} = \frac{s_x}{\bar{x}} \cdot 100\%,$$

Względne miary zmienności – Współczynniki zmienności

- Względne miary zróżnicowania (inaczej **współczynniki zmienności**), definiujemy jako ilorazy bezwzględnych miar zróżnicowania do odpowiednich miar średnich.
- Współczynniki te stosujemy przy porównaniach.
  - współczynnik zmienności V_{d_x} oparty na odchyleniu przeciętnym

$$V_{d_x} = \frac{d_x}{\bar{x}} \cdot 100\%,$$

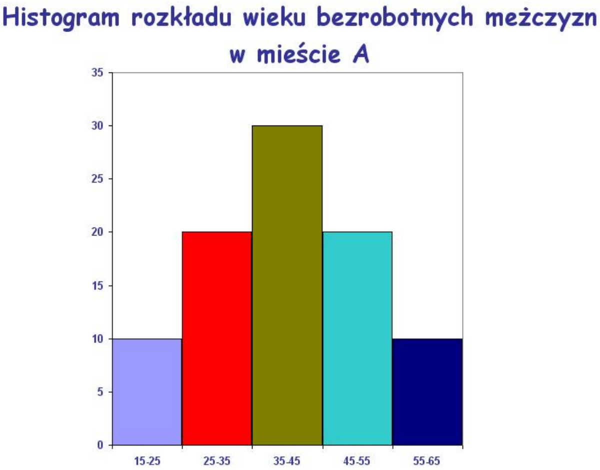
  - współczynnik zmienności V_{s_x} oparty na odchyleniu standardowym

$$V_{s_x} = \frac{s_x}{\bar{x}} \cdot 100\%,$$

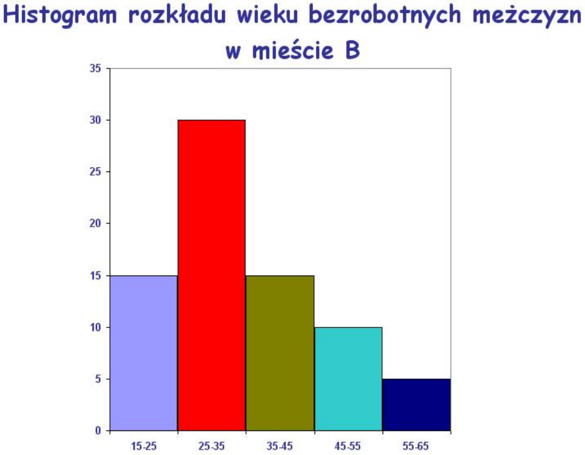
  - współczynnik zmienności V_{Q_x} oparty na odch. ćwiartkowym

$$V_{Q_x} = \frac{Q_x}{Me} \cdot 100\%.$$

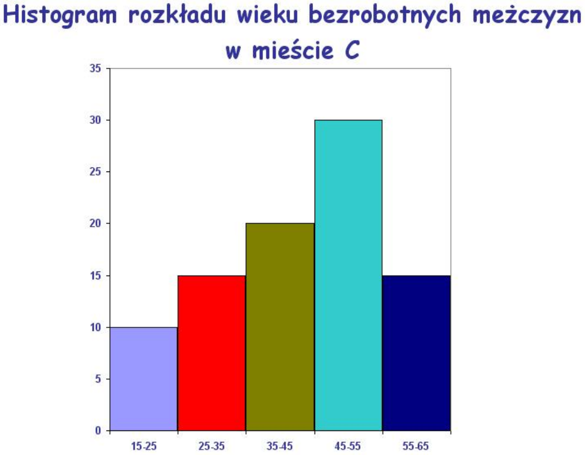
Analiza asymetrii – Przykład histogramu szeregu symetrycznego



Analiza asymetrii – Przykład histogramu szeregu asymetrycznego prawostronnie



Analiza asymetrii – Przykład histogramu szeregu asymetrycznego lewostronnie



Miary asymetrii

- W szeregach symetrycznych średnia arytmetyczna równa jest medianie i dominancie, czyli

$$\bar{x} = Me = Do,$$

natomiast różnica między kwartylem trzecim a medianą równa jest różnicy między medianą a kwartylem pierwszym, czyli

$$Q_3 - Me = Me - Q_1.$$

Miary asymetrii

- W szeregach symetrycznych średnia arytmetyczna równa jest medianie i dominancie, czyli

$$\bar{x} = Me = Do,$$

natomiast różnica między kwartylem trzecim a medianą równa jest różnicy między medianą a kwartylem pierwszym, czyli

$$Q_3 - Me = Me - Q_1 .$$

- Przy asymetrii lewostronnej zachodzą nierówności

$$\bar{x} \leq Me \leq Do, \qquad (Q_3 - Me) - (Me - Q_1) < 0,$$

natomiast przy asymetrii prawostronnej mają miejsce nierówności odwrotne.

Miary asymetrii – Wskaźnik skośności

- Własności te wykorzystuje się przy konstrukcji wybranych mierników asymetrii.

Miary asymetrii – Wskaźnik skośności

- Własności te wykorzystuje się przy konstrukcji wybranych mierników asymetrii.
- **Wskaźnik skośności** definiujemy wzorem

$$M_s = \bar{x} - Do.$$

Miary asymetrii – Wskaźnik skośności

- Własności te wykorzystuje się przy konstrukcji wybranych mierników asymetrii.
- **Wskaźnik skośności** definiujemy wzorem

$$M_s = \bar{x} - Do.$$

- Znak tego wskaźnika informuje o kierunku asymetrii: **znak ujemny** oznacza asymetrię lewostronną (nazywaną także asymetrią ujemną), natomiast **znak dodatni** oznacza asymetrię prawostronną (dodatnią). W przypadku szeregów symetrycznych mamy $M_s = 0$.

Miary asymetrii – Wskaźnik skośności

- Własności te wykorzystuje się przy konstrukcji wybranych mierników asymetrii.
- **Wskaźnik skośności** definiujemy wzorem

$$M_s = \bar{x} - Do.$$

- Znak tego wskaźnika informuje o kierunku asymetrii: **znak ujemny** oznacza asymetrię lewostronną (nazywaną także asymetrią ujemną), natomiast **znak dodatni** oznacza asymetrię prawostronną (dodatnią). W przypadku szeregów symetrycznych mamy $M_s = 0$.
- Wskaźnik M_s jest **miarą mianowaną**, o jego wartości decyduje nie tylko stopień skośności szeregu, ale również ogólny poziom cechy w danej zbiorowości. Z tego powodu częściej obliczany jest (względny) współczynnik skośności.

Miary asymetrii – Współczynnik skośności

- **Współczynnik skośności** obliczamy ze wzoru:

$$W_s = \frac{M_s}{s_x} = \frac{\bar{x} - Do}{s_x} \quad \text{lub} \quad W_s = \frac{M_s}{d_x} = \frac{\bar{x} - Do}{d_x}.$$

Miary asymetrii – Współczynnik skośności

- **Współczynnik skośności** obliczamy ze wzoru:

$$W_s = \frac{M_s}{s_x} = \frac{\bar{x} - Do}{s_x} \quad \text{lub} \quad W_s = \frac{M_s}{d_x} = \frac{\bar{x} - Do}{d_x}.$$

- Współczynnik W_s przyjmuje na ogół wartości z przedziału $[-1, 1]$ (w przypadku skrajnej asymetrii może się zdarzyć, że jego wartość wykroczy poza podany przedział).

Miary asymetrii – Współczynnik skośności

- **Współczynnik skośności** obliczamy ze wzoru:

$$W_s = \frac{M_s}{s_x} = \frac{\bar{x} - Do}{s_x} \quad \text{lub} \quad W_s = \frac{M_s}{d_x} = \frac{\bar{x} - Do}{d_x}.$$

- Współczynnik W_s przyjmuje na ogół wartości z przedziału $[-1, 1]$ (w przypadku skrajnej asymetrii może się zdarzyć, że jego wartość wykroczy poza podany przedział).
- Znak współczynnika W_s informuje o **kierunku asymetrii**, a wartość bezwzględna – **o sile asymetrii**.

Miary asymetrii – Klasyczny współczynnik asymetrii

- **Klasyczny współczynnik asymetrii** jest często stosowaną miarą asymetrii wyrażoną wzorem:

$$A_s = \frac{\mu_3}{s_x^3},$$

gdzie s_x oznacza odchylenie standardowe, natomiast μ_3 jest tzw. momentem centralnym trzeciego rzędu, który definiujemy następująco:

$$\mu_3 = \begin{cases} \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3, & \text{dla szeregu szczegółowego,} \\ \frac{1}{n}\sum_{i=1}^{k}(x_i - \bar{x})^3 n_i, & \text{dla szeregu r. punktowego,} \\ \frac{1}{n}\sum_{i=1}^{k}(\dot{x}_i - \bar{x})^3 n_i, & \text{dla szeregu r. z przedziałami.} \end{cases}$$

Miary asymetrii – Klasyczny współczynnik asymetrii

- **Klasyczny współczynnik asymetrii** jest często stosowaną miarą asymetrii wyrażoną wzorem:

$$A_s = \frac{\mu_3}{s_x^3},$$

gdzie s_x oznacza odchylenie standardowe, natomiast μ_3 jest tzw. momentem centralnym trzeciego rzędu, który definiujemy następująco:

$$\mu_3 = \begin{cases} \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^3, & \text{dla szeregu szczegółowego,} \\ \frac{1}{n}\sum_{i=1}^{k}(x_i - \bar{x})^3 n_i, & \text{dla szeregu r. punktowego,} \\ \frac{1}{n}\sum_{i=1}^{k}(\dot{x}_i - \bar{x})^3 n_i, & \text{dla szeregu r. z przedziałami.} \end{cases}$$

- Współczynnik A_s przyjmuje na ogół wartości z przedziału $[-2, 2]$ (w przypadku skrajnej asymetrii jego wartość może wykroczyć poza ten przedział).

Miary asymetrii – Pozycyjny współczynnik asymetrii

- **Pozycyjny współczynnik asymetrii** określa siłę i kierunek asymetrii dla tych jednostek zbiorowości, które znajdują się między pierwszym i trzecim kwartylem, a więc w zawężonym obszarze zmienności cechy.

Miary asymetrii – Pozycyjny współczynnik asymetrii

- **Pozycyjny współczynnik asymetrii** określa siłę i kierunek asymetrii dla tych jednostek zbiorowości, które znajdują się między pierwszym i trzecim kwartylem, a więc w zawężonym obszarze zmienności cechy.
- Definiujemy go wzorem

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

$$= \frac{Q_3 - 2Me + Q_1}{Q_3 - Q_1}.$$

Miary asymetrii – Pozycyjny współczynnik asymetrii

- **Pozycyjny współczynnik asymetrii** określa siłę i kierunek asymetrii dla tych jednostek zbiorowości, które znajdują się między pierwszym i trzecim kwartylem, a więc w zawężonym obszarze zmienności cechy.
- Definiujemy go wzorem

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

$$= \frac{Q_3 - 2Me + Q_1}{Q_3 - Q_1}.$$

- Współczynnik A_Q przyjmuje wartość z przedziału $[-1, 1]$. Podobnie, jak mierniki W_s i A_s, jego znak informuje o kierunku, a wartość bezwzględna – o sile asymetrii.