

Miary tendencji centralnej (położenia).

Miara rozkładu jest to liczbową charakterystyką rozkładu cechy, dostarczająca informacji na temat właściwości tego rozkładu. Taką charakterystykę określa się w zależności od przedmiotu badania statystycznego jako:

- parametr - jeśli badane są dane z pełnej populacji
- statystyka - jeśli przedmiotem badania są dane z próby losowej

Miarą tendencji centralnej rozkładu nazywamy miarę rozkładu, która określa położenie wartości centralnych rozkładu (wartości typowych, średnich). Miary tendencji centralnej zaliczamy do miar położenia.

Miarą położenia rozkładu nazywamy miarę rozkładu, która określa relację między dwoma identycznymi rozkładami, ale przesuniętymi względem osi odciętych układu współrzędnych.

Przejdźmy tymczasem do omówienia najważniejszych miar tendencji centralnej. Są to:

Mediana

Medianą lub wartością środkową, którą oznaczamy m_e - próbki x_1, \dots, x_n nazywamy środkową liczbę w uporządkowanej niemalejąco próbce

$$x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$$

gdy n jest liczbą nieparzystą, albo średnią arytmetyczną dwóch środkowych liczb, gdy n jest liczbą parzystą, tzn.

$m_e = x_{(n+1)/2}$ gdy n nieparzyste, lub

$m_e = (x_{(n/2)} + x_{(n/2+1)})/2$ gdy n parzyste

Moda

Wartością modalną (modą, dominantą) m_0 próbki x_1, \dots, x_n o powtarzających się wartościach nazywamy najczęściej powtarzającą się wartość (o ile istnieje!), nie będącą x_{\min} ani też x_{\max} . Jeśli wartość taka nie istnieje lub istnieje ich wiele to mówimy, że nie ma mody.

Średnia arytmetyczna

Średnią arytmetyczną liczb x_1, \dots, x_n nazywamy liczbę

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

Jeżeli w próbce wynik pomiaru x_i wystąpił n_i razy, $i=1, \dots, k$ gdzie $\sum_{i=1}^k n_i = n$, to średnią arytmetyczną oblicza się według równoważnego wzoru:

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

Średnia ta bywa również nazywana *średnią arytmetyczną ważoną*. Liczności n_i pełnią tu rolę tzw. wag.

Miary rozproszenia

Miarą rozproszenia (zmienności, dyspersji) nazywamy miarę rozkładu, która opisuje relację pomiędzy rozkładami różniącymi się rozproszeniem wartości cechy wokół wartości centralnych.

Rozstęp

Rozstęp - najprostsza miara rozproszenia. Jest to miara charakteryzująca empiryczny obszar zmienności badanej cechy. Wyraża się wzorem:

$$R = x_{\max} - x_{\min}$$

Rozstęp jest miarą, która nie uwzględnia wszystkich wartości. Ponadto jest bardzo podatna na wartości odstające.

Wariancja

Wariancja - wariancją zmiennej X nazywamy średnią arytmetyczną kwadratów odchyleń poszczególnych wartości zmiennej od średniej arytmetycznej całej zbiorowości:

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

Wartość oczekiwana

1. Dla rozkładu dyskretnego $\{(x_i, p_i), i=1, \dots\}$ mamy $EX = \sum_k x_k p_k$ (o ile szereg jest zbieżny). W przeciwnym wypadku EX nie istnieje!
2. W przypadku absolutnie ciągłym i gęstości f mamy $EX = \int_{-\infty}^{\infty} xf(x)dx$ (o ile całka jest zbieżna).

Wartość oczekiwana to funkcja liniowa, ponieważ $EX(aX + b) = aEX + b$. Ponadto, jeśli X to zmienna losowa, g - funkcja mierzalna to $g(X)$ - zmienna losowa i $EG(X) = \sum_i g(x_i) p_i$ (dla rozkładu dyskretnego) lub $EG(X) = \int_{-\infty}^{\infty} g(u) f(u) du$

Wariancja (zmiennej losowej)

$$VarX = E(X - EX)^2 \quad (\text{o ile istnieją obie wartości oczekiwane}).$$

W przypadku dyskretnym mamy:

$$\begin{aligned} VarX &= \sum_k (x_k - EX)^2 p_k = \sum_k (x_k^2 - 2 \cdot x_k \cdot EX + (EX)^2) p_k \\ &= \sum_k x_k^2 p_k - 2 \cdot EX \sum_k x_k p_k + (EX)^2 \sum_k p_k = EX^2 - 2 \cdot EX \cdot EX + (EX)^2 = EX^2 - (EX)^2 \end{aligned}$$

Miedzy tymi dwiema liniami powinien stać znak równości, ale OOffice nie pozwala ;). Analogicznie w przypadku absolutnie ciągłym:

$$\begin{aligned} VarX &= \int_{-\infty}^{\infty} (x - EX)^2 f(x) dx = \int_{-\infty}^{\infty} x^2 f(x) dx - 2EX \int_{-\infty}^{\infty} xf(x) dx + (EX)^2 \int_{-\infty}^{\infty} f(x) dx \\ &= \int_{-\infty}^{\infty} x^2 f(x) dx - (EX)^2 = EX^2 - (EX)^2 \end{aligned}$$

Wariancja nie jest funkcją liniową, gdyż

$$Var(aX+b)=E((aX+b)-E(aX+b))^2=E(aX+b-aEX-b)^2=a^2 E(X-EX)^2=a^2 VarX$$