Gdy miara staje się celem, przestaje być dobrą miarą

Wariancja stanowi różnicę między momentem zwykłym drugiego rzędu $\langle x^2 \rangle$, a kwadratem momentu zwykłego pierwszego rzędu. Wariancja jest równa zero wtedy i tylko wtedy gdy rozkład zmiennej losowej X jest **zdegenerowany**, tzn. $x_i = \langle x \rangle$.

$$(\Delta x)^2 = \sum_{i=1}^{n} [x_i - \langle x \rangle]^2 \, p_i = \langle x^2 \rangle - \langle x \rangle^2 \geq 0 \quad \rightarrow \quad \langle x^2 \rangle \geq \langle x \rangle^2$$

Kwadraty **odchyleń** $\Delta x = x_i - \langle x \rangle$ występujące w definicji wariancji powiększają wpływ obserwacji znacznie odbiegających od średniej arytmetycznej: $|x_i| \gg \langle x \rangle$ na wartość sumy

$$\sum_{i=1}^{n} [x_i - \langle x \rangle]^2 \, p_i$$

Jednocześnie kwadraty odchyleń Δx występujące w definicji wariancji pomniejszają wpływ obserwacji zbliżonych do średniej arytmetycznej: $x_i \approx \langle x \rangle$ na wartość sumy

$$\sum_{i=1}^{n} [x_i - \langle x \rangle]^2 \, p_i$$

Przykład: jeżeli odchylenie Δx wynosi 5, to jego udział (waga) w definicji wariancji wynosi 25. Jeżeli odchylenie Δx wynosi $1/10$, to jego udział w definicji wariancji wynosi zaledwie $1/100$. Niekiedy pożądane jest aby wszystkie obserwacje x_i miały udział w definicji wartości wskaźnika rozproszenia proporcjonalny do wielkości odchyleń Δx. Tak jest np. w definicji **odchylenia przeciętnego** d.

$$d = \sum_{i=1}^{n} |x_i - \langle x \rangle| \, p_i = \sum_{i=1}^{n} |\Delta x_i| \, p_i$$

Fakty dotyczące wariancji $(\Delta x)^2$

1) jest miarą rozproszenia wartości cechy x_i wokół średniej arytmetycznej $\langle x \rangle$
2) im większe rozproszenie wartości cechy tym większą wartość przyjmuje wariancja
3) jest równa zero gdy każdy pomiar występujący w analizowanej zbiorowości statystycznej jest identyczny: $x_i = \langle x \rangle$
4) jest wyrażona w jednostkach miary wartości cechy x_i podniesionych do kwadratu – jest wielkością mianowaną
5) opiera się na kwadratach odchyleń $\Delta x_i = x_i - \langle x \rangle$
6) jest ściśle związana z pojęciami takimi jak nieoznaczoność (chemia kwantowa) i fluktuacje (termodynamika statystyczna)
7) nie jest miarą rozproszenia odporną na wartości odstające x_O (ang. *outliers*) występujące w analizowanej zbiorowości statystycznej

$$x_O = |x_i| \gg \langle x \rangle$$

Zadanie 1. Wykonano serię pomiarów pewnej bezwymiarowej wielkości fizycznej. Otrzymano wyniki:

0, 2, 4, 6, 8, 10

Proszę obliczyć ile wynosi wariancja.

$$(\Delta x)^2 = \langle x^2 \rangle - \langle x \rangle^2 = \frac{220}{6} - 5^2 = \frac{220}{6} - 25 = \frac{220}{6} - \frac{150}{6} = \frac{70}{6} = 11\frac{4}{6}$$

Zadanie 2. Wykonano serię pomiarów pewnej bezwymiarowej wielkości fizycznej. Otrzymano wyniki:

2, 4, 5, 5, 6, 8

Proszę obliczyć ile wynosi wariancja.

$$(\Delta x)^2 = \langle x^2 \rangle - \langle x \rangle^2 = \frac{170}{6} - 5^2 = \frac{170}{6} - 25 = \frac{170}{6} - \frac{150}{6} = \frac{20}{6} = 3\frac{2}{6}$$

Zadanie 3. Wykonano serię pomiarów pewnej bezwymiarowej wielkości fizycznej. Otrzymano wyniki:

5, 5, 5, 5, 5, 5

Proszę obliczyć ile wynosi wariancja.

$$(\Delta x)^2 = \langle x^2 \rangle - \langle x \rangle^2 = 25 - 5^2 = 0$$

Odchylenie standardowe Δx – klasyczna miara zmienności, obok średniej arytmetycznej $\langle x \rangle$ najczęściej stosowane pojęcie statystyczne. Odchylenie standardowego zostało wprowadzone do statystyki w 1894 roku przez **Karla Pearsona**. Odchylenie standardowe to pierwiastek kwadratowy z wariancji.

$$\Delta x = \sqrt{(\Delta x)^2} = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}$$

Karl Pearson (1857-1936) – brytyjski matematyk i statystyk, uznawany za jednego z twórców nowoczesnej statystyki. Był profesorem University College London. Jego prace wywarły duży wpływ na statystykę, biometrię i analizę danych.

Przewaga odchylenia standardowego nad wariancją polega na tym, że odchylenie standardowe jest wyrażone w tych samych jednostkach co badana wielkość, podczas gdy wariancja jest wyrażona w jednostkach podniesionych do kwadratu. Odchylenie standardowe określa, jak bardzo poszczególne wyniki pomiarów różnią się od średniej arytmetycznej, a tym samym informuje nas o jakości danych (precyzji pomiarowej). Odchylenie standardowe stosuje się w naukach chemicznych m.in. do oceny powtarzalności miareczkowań, pomiarów absorbancji, pH, przewodnictwa, pomiarów potencjometrycznych itd. Odchylenie standardowe stanowi również punkt odniesienia przy identyfikacji wartości odstających x_O w analizowanej zbiorowości statystycznej.