Wariancja $(\Delta x)^2$ - różnica między momentem zwykłym drugiego rzędu $\langle x^2 \rangle$ a kwadratem momentu zwykłego pierwszego rzędu $\langle x \rangle$.

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

Wariancja $(\Delta x)^2$ przyjmuje zawsze wartości większe lub równe zero (nieujemne). Jeżeli wszystkie wartości x_i zmiennej mają identyczną wartość liczbową, tzn. $x_i = \langle x \rangle$, wtedy $(\Delta x)^2 = 0$. Zauważmy, że wprowadzenie kwadratów odchyleń $[x_i - \langle x \rangle]^2$ powiększa wpływ obserwacji znacznie odbiegających od wartości średniej $\langle x \rangle$ na wartość sumy

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

Z tego samego powodu, pomniejszany jest wpływ na sumę obserwacji bliskich wartości średniej $x_i \approx \langle x \rangle$. Jeżeli odległość wartości cechy x_i w próbie od średniej $\langle x \rangle$ wynosi 2, to jej udział (waga) w sumie definiującej wariancję wynosi 4; jeśli odległość wartości cechy x_i w próbie od średniej $\langle x \rangle$ wynosi $1/10$, to jej udział (waga) w sumie definiującej wariancję wynosi $1/100$. Niekiedy pożądane jest by wszystkie obserwacje miały udział w ogólnej wartości wskaźnika rozproszenia proporcjonalny do wielkości **odchyłek** tych obserwacji od wartości średniej: $x_i - \langle x \rangle$. Wówczas stosuje się wskaźnik rozproszenia zwany **odchyleniem przeciętnym** d (patrz dalej).

Kilka faktów dotyczących wariancji $(\Delta x)^2$

a) jest **miarą rozproszenia** wartości x_i wokół wartości średniej $\langle x \rangle$,
b) im większe rozproszenie wartości x_i wokół wartości średniej $\langle x \rangle$, tym większą wartość liczbową przyjmuje wariancja $(\Delta x)^2$,
c) wariancja przyjmuje wartość $(\Delta x)^2 = 0$ wyłącznie wtedy, gdy w analizowanej próbie nie występuje rozproszenie wyników x_i wokół wartości średniej $\langle x \rangle$ - wszystkie wyniki przyjmują identyczną wartość liczbową,
d) wariancja $(\Delta x)^2$ jest **wielkością mianowaną** - jest wyrażona w jednostkach miary wartości x_i podniesionych do kwadratu,

e) opiera się na kwadratach **odchyłek** $\Delta x = x_i - \langle x \rangle$,
f) w naukach ścisłych wariancja $(\Delta x)^2$ nazywana jest również **dyspersją**, **niezonaczonością**, **nieokreślonością** (chemia kwantowa) lub **fluktuacją** (termodynamika statystyczna),
g) nie jest odporna na wartości odstające (ang. *outliers*) występujące w próbie.

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

0, 2, 4, 6, 8, 10

Proszę obliczyć ile wynosi wariancja.

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

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

2, 4, 5, 5, 6, 8

Proszę obliczyć ile wynosi wariancja.

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

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 - stanowi klasyczną miarę zmienności, obok średniej arytmetycznej $\langle x \rangle$ jest najczęściej stosowanym pojęciem statystycznym. Pojęcie odchylenia standarowego zostało wprowadzone do statystyki opisowej w 1884 roku przez **Karla Pearsona**. Odchylenie standardowe jest pierwiastkiem kwadratowym z wariancji $(\Delta x)^2$. Odchylenie standardowe (podobnie jak wariancja) jest liczbą nieujemną: $\Delta x \geq 0$. Im większa wartość liczbowa odchylenia standardowego, tym większe zróżnicowanie badanej zbiorowości statystycznej.

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

Odchylenie standardowe Δx jest równe zero jedynie wtedy gdy wszystkie wyniki w próbie są identyczne. W przypadku gdy wariancja jest wyznaczana na podstawie **próby** (losowej) pobranej z **populacji generalnej** będziemy ją oznaczać symbolem $(\Delta s)^2$. W definicji $(\Delta s)^2$ suma iloczynów $[x_i - \langle x \rangle]^2 n_i$ jest dzielona nie przez całkowitą liczbę obserwacji N, ale przez różnicę $N - 1$ (dlaczego?).

$$(\Delta x)^2 = \sum_{i=1}^{n} [x_i - \langle x \rangle]^2 \, p_i = \frac{1}{N} \sum_{i=1}^{n} [x_i - \langle x \rangle]^2 \, n_i$$

$$(\Delta s)^2 = \frac{1}{N-1} \sum_{i=1}^{n} [x_i - \langle x \rangle]^2 \, n_i = \frac{1}{N-1} \cdot \frac{N}{N} \sum_{i=1}^{n} [x_i - \langle x \rangle]^2 \, n_i$$

$$\downarrow$$

$$(\Delta s)^2 = \frac{N}{N-1} \cdot \frac{1}{N} \sum_{i=1}^{n} [x_i - \langle x \rangle]^2 \, n_i = \frac{N}{N-1} (\Delta x)^2 \quad \rightarrow \quad (\Delta s)^2 \geq (\Delta x)^2$$

$$\downarrow$$

$$\Delta s = \sqrt{\frac{N}{N-1}} \Delta x$$

Uzasadnienie tego faktu jest następujące: przyjmijmy, że w badanej populacji generalnej występuje **wartość najmniejsza** x_{min} i **wartość największa** x_{max} oraz stosunkowo niewiele wartości, które są do nich zbliżone. Większość z wartości x_i jakie przyjmuje zmienna X są porównywalne z wartością średnią $\langle x \rangle$, tzn. $x_i \approx \langle x \rangle$. Gdy z populacji generalnej pobieramy próbę to prawdopodobieństwo, że wybierzemy wartości skrajne jest bardzo małe – w związku z tym wariancja szacowana (**estymowana**) dla populacji generalnej na podstawie próby losowej będzie zaniżona! Różnica $N - 1$, która pojawia się w definicji $(\Delta s)^2$ ma na celu korektę tego faktu. Dla $N \to \infty$ zachodzi jednak

$$\frac{N}{N-1} \approx 1 \quad \rightarrow \quad (\Delta s)^2 \approx (\Delta x)^2$$

Rozpatrzymy rozkład zmiennej X, w którym każda z wartości x_i występuje tylko raz: $n_i = 1$.

$$p_i = \frac{n_i}{N} = \frac{1}{N}$$

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

Zauważmy, że wariancję $(\Delta x)^2$ możemy interpretować jako **średnią arytmetyczną kwadratów odchyleń** poszczególnych wartości x_i zmiennej X od wartości oczekiwanej $\langle x \rangle$. Wyznaczmy odchylenie standardowe Δx.

$$\Delta x = \sqrt{\frac{1}{N} \sum_{i=1}^{N} [x_i - \langle x \rangle]^2} = \sqrt{\frac{1}{N}} \cdot \sqrt{\sum_{i=1}^{N} [x_i - \langle x \rangle]^2} = \frac{1}{\sqrt{N}} \sqrt{\sum_{i=1}^{N} [x_i - \langle x \rangle]^2}$$

Odchylenie standardowe Δx stanowi funkcję N zmiennych: $\Delta x = f(x_1, x_2, x_3, \dots, x_N)$. Ile wynosi pochodna cząstkowa odchylenia standardowego Δx względem i-tej obserwacji? Uwaga: Δx jest funkcją potrójnie złożoną (superponowaną)!

$$\frac{\partial}{\partial x_i}\Delta x = \frac{\partial}{\partial x_i}\frac{1}{\sqrt{N}}\sqrt{\sum_{i=1}^{N}[x_i - \langle x\rangle]^2} = \frac{1}{\sqrt{N}}\frac{\partial}{\partial x_i}\left(\sum_{i=1}^{N}[x_i - \langle x\rangle]^2\right)^{\frac{1}{2}}$$

$$\frac{\partial}{\partial x_i}\Delta x = \frac{1}{\sqrt{N}}\cdot\frac{1}{2}\left(\sum_{i=1}^{N}[x_i - \langle x\rangle]^2\right)^{-\frac{1}{2}}\cdot\ \frac{\partial}{\partial x_i}\sum_{i=1}^{N}[x_i - \langle x\rangle]^2$$

Usuwamy niewymierność z mianownika (rozszerzamy ułamek $1/\sqrt{N}$).

$$\frac{\partial}{\partial x_i}\Delta x = \frac{1}{\sqrt{N}}\cdot\frac{\sqrt{N}}{\sqrt{N}}\cdot\frac{1}{2}\left(\sum_{i=1}^{N}[x_i - \langle x\rangle]^2\right)^{-\frac{1}{2}}\cdot\ \frac{\partial}{\partial x_i}\sum_{i=1}^{N}[x_i - \langle x\rangle]^2$$

$$\frac{\partial}{\partial x_i}\Delta x = \frac{\sqrt{N}}{2N}\cdot\left(\sum_{i=1}^{N}[x_i - \langle x\rangle]^2\right)^{-\frac{1}{2}}\cdot\ \frac{\partial}{\partial x_i}\sum_{i=1}^{N}[x_i - \langle x\rangle]^2$$

$$\frac{\partial}{\partial x_i}\Delta x = \frac{1}{\frac{2N}{\sqrt{N}}}\cdot\left(\sum_{i=1}^{N}[x_i - \langle x\rangle]^2\right)^{-\frac{1}{2}}\cdot\ \frac{\partial}{\partial x_i}\sum_{i=1}^{N}[x_i - \langle x\rangle]^2$$

$$\frac{\partial}{\partial x_i}\Delta x = \frac{1}{2N\Delta x}\cdot\frac{\partial}{\partial x_i}\sum_{i=1}^{N}[x_i - \langle x\rangle]^2 = \frac{1}{2N\Delta x}\cdot\frac{\partial}{\partial x_i}[x_i - \langle x\rangle]^2$$

$$\frac{\partial}{\partial x_i}\Delta x = \frac{1}{2N\Delta x}\cdot 2[x_i - \langle x\rangle]\cdot 1 = \frac{x_i - \langle x\rangle}{N\Delta x}$$

Pochodna cząstkowa odchylenia standardowego Δx względem obserwacji x_i wynosi

$$\frac{\partial\Delta x}{\partial x_i} = \frac{x_i - \langle x\rangle}{N\Delta x} = \frac{x_i - \langle x\rangle}{N}\frac{1}{\Delta x}$$

Pochodna cząstkowa odchylenia standardowego Δx względem x_i jest więc proporcjonalna do odwrotności odchylenia standardowego! Zapiszmy warunek konieczny dla istnienia ekstremum funkcji wielu zmiennych

$$\frac{\partial \Delta x}{\partial x_i} = \frac{x_i - \langle x \rangle}{N \Delta x} = 0 \quad \rightarrow \quad x_i - \langle x \rangle = 0 \quad \rightarrow \quad x_i = \langle x \rangle$$

Wpływ pojedynczej obserwacji x_i na wartość odchylenia standardowego Δx maleje gdy liczba obserwacji N rośnie. W granicznym przypadku, gdy $N \to \infty$ wpływ pojedynczej obserwacji x_i na wartość odchylenia standardowego jest znikomy (zaniedbywalnie mały).

$$\lim_{N \to \infty} \frac{x_i - \langle x \rangle}{N \Delta x} = 0$$

Wpływ pojedynczej obserwacji x_i na wartość odchylenia standardowego Δx jest tym większy im większa jest wartość odchylenia $x_i - \langle x \rangle$: w mianowniku występuje iloczyn dwóch stałych.