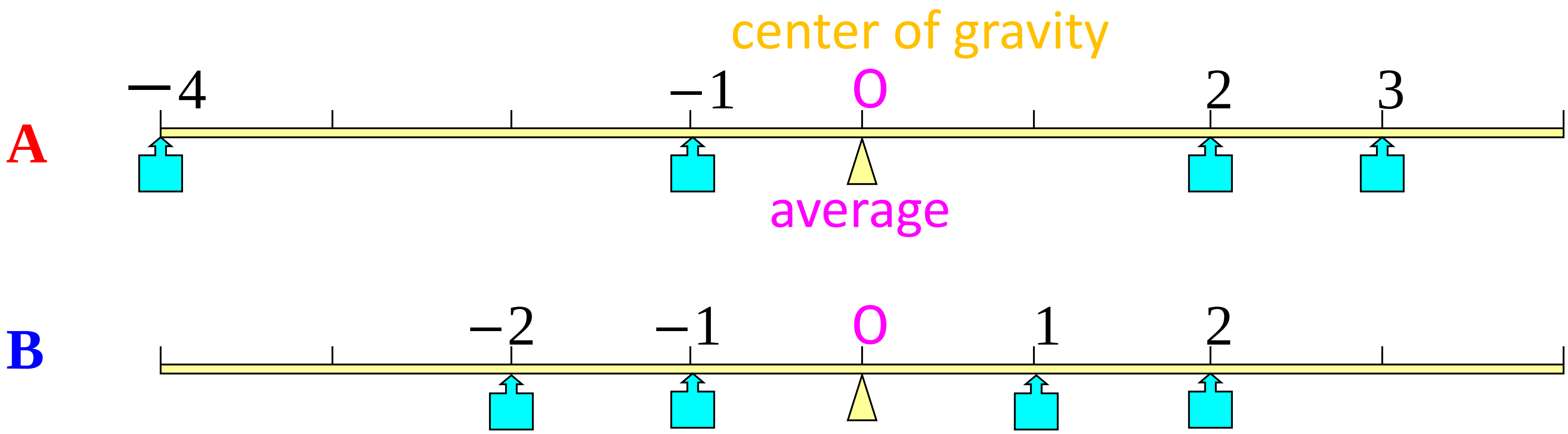
How to quantify "**Dispersion**"



Explain using a balance with a suspended weight.

Calculate the distance from the "center of gravity"

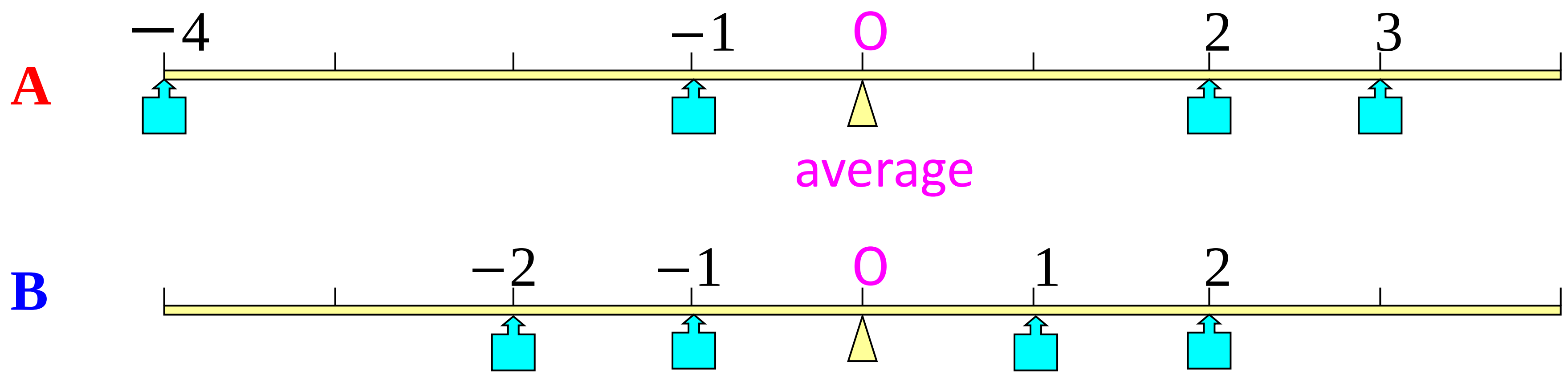
A：−4，−1，2，3　　　B：−2，−1，1，2

Calculate the sum and average of the distances.

A：$\{(-4)+(-1)+2+3\}/4=0$

B：$\{(-2)+(-1)+1+2\}/4=0$

The average value cannot be used as an indicator of “Dispersion”.



Calculate the sum and average of the squares of the distances.

A: $\{(-4-0)^2+(-1-0)^2+(2-0)^2+(3-0)^2\}/4$

$=\{(-4)^2+(-1)^2+2^2+3^2\}/4=7.5$

B: $\{(-2-0)^2+(-1-0)^2+(1-0)^2+(2-0)^2\}/4$

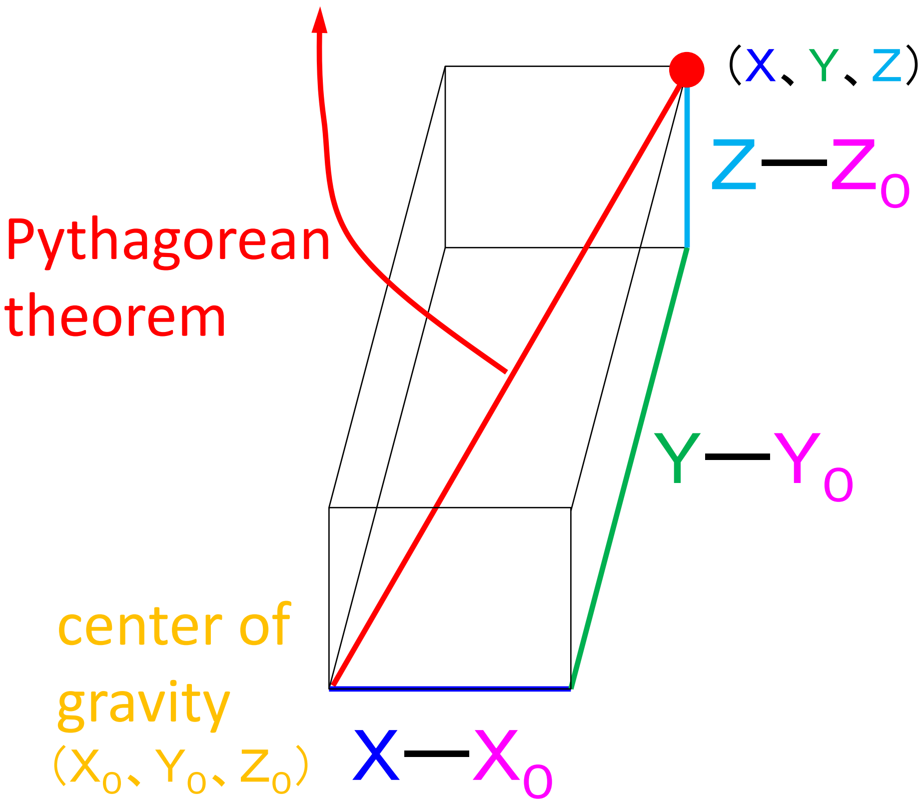
$=\{(-2)^2+(-1)^2+1^2+2^2\}/4=2.5$

A with a large "Dispersion" has a larger number than B.

Which one of Euclidean distance and Manhattan distance represents dispersion?

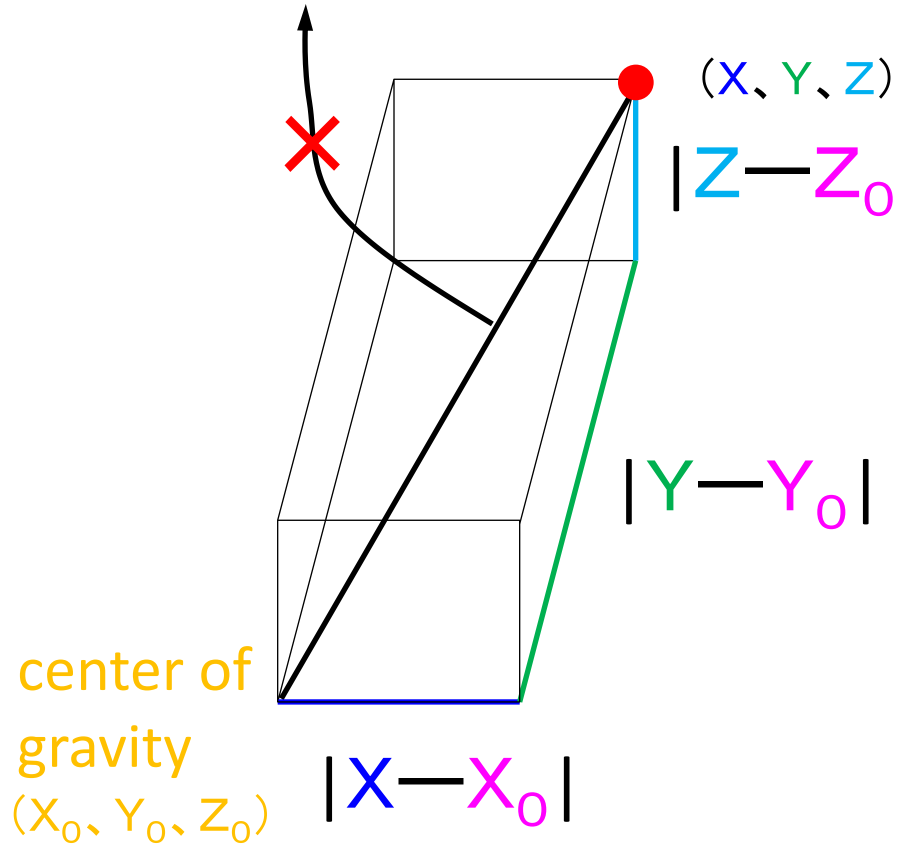
Euclidean distance

$$\sqrt{(X-X_0)^2+(Y-Y_0)^2+(Z-Z_0)^2}$$

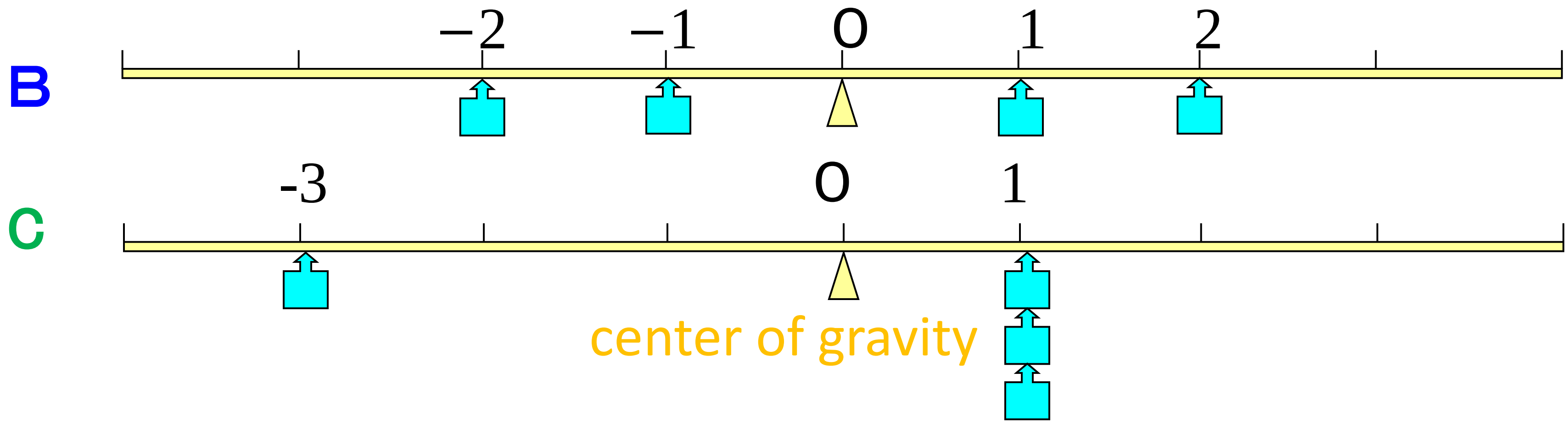


Manhattan distance

$$|X-X_0|+|Y-Y_0|+|Z-Z_0|$$



Which of B and C has larger dispersion?



Calculate the sum and the average of the distance from the center of gravity to the weight as the absolute value

A: $(|-4|+|-1|+|2|+|3|)/4=2.5$

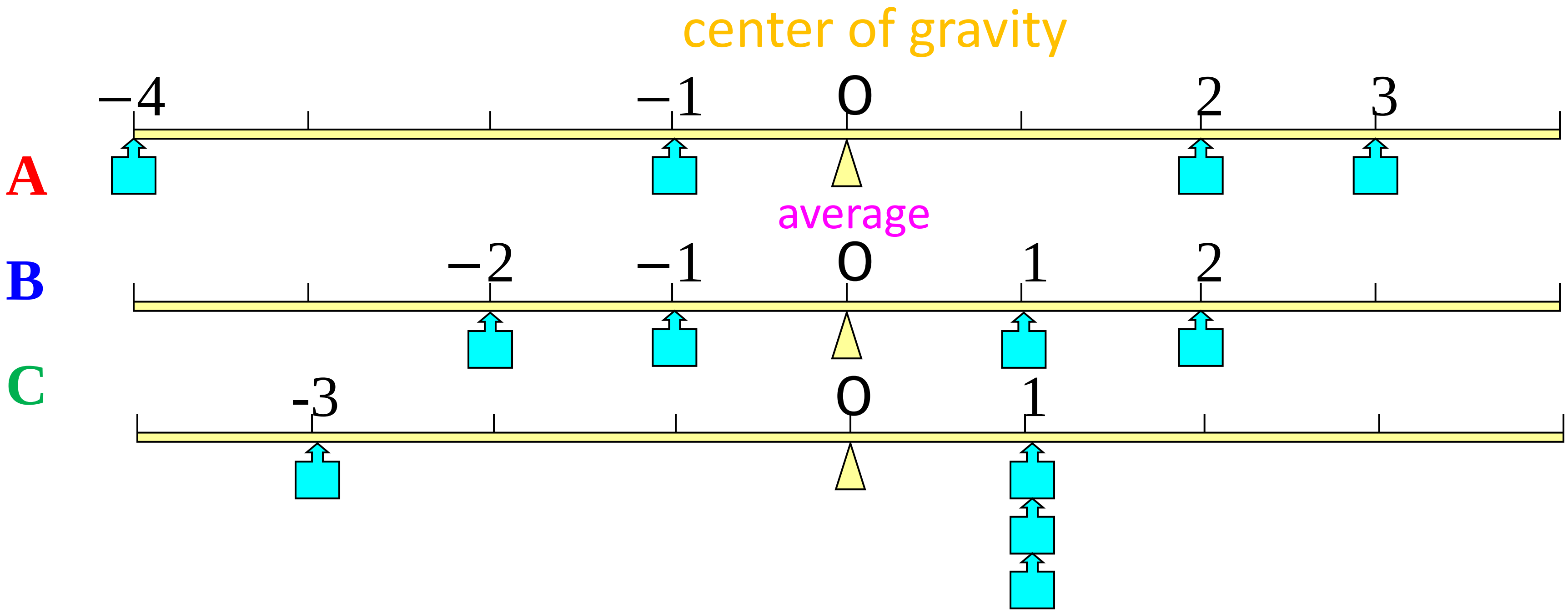
B: $(|-2|+|-1|+|1|+|2|)/4=1.5$

C: $(|-3|+|1|+|1|+|1|)/4=1.5$

B and C: same

Calculate the sum and average of the squares of the distances.

C: $\{(-3)^2+1^2+1^2+1^2\}/4=3.0$



Degree of dispersion→Variance $V = \sum(x - \bar{x})^2/n$

A: $\{(-4-0)^2+(-1-0)^2+(2-0)^2+(3-0)^2\}/4=7.5$

B: $\{(-2-0)^2+(-1-0)^2+(1-0)^2+(2-0)^2\}/4=2.5$

C: $\{(-3-0)^2+(1-0)^2+(1-0)^2+(1-0)^2\}/4=3.0$

Standard deviation $\sigma = \sqrt{V} = \sqrt{\sum(x - \bar{x})^2/n}$

A: $\sigma_A=\sqrt{7.5}=2.73$ ←dispersion is largest

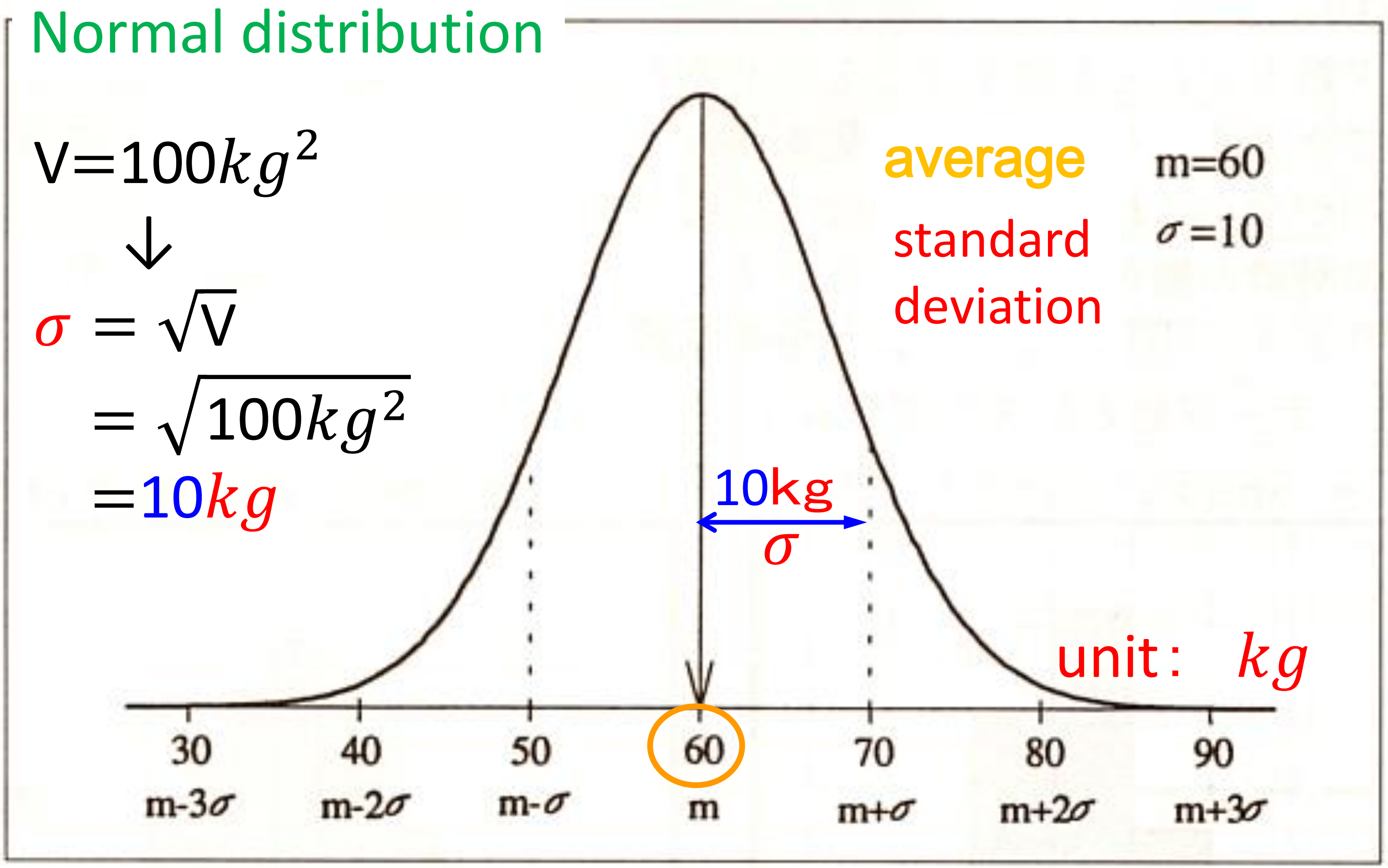
B: $\sigma_B=\sqrt{2.5}=1.58$

C: $\sigma_C=\sqrt{3.0}=1.73$

$\bar{x}$: average

n : number of weights

This graph shows a normal distribution with an average body weight of 60 kg and a standard deviation of σ: 10kg.



Why do we need to calculate the square root of V?

→ To fit the unit dimensions