

# • 2강 평균, 분산, 표준편차

- 평균: 데이터를 나타내는 대표값  $\rightarrow$  공학에서는 Mean 이라는 의미로  $m$ 을 씀

$$m = \frac{a+b+c}{N}$$

- 분산 & 표준편차

• 분산: 데이터가 얼마나 퍼져 있는지 나타내는 값

$$V = \frac{(a-m)^2 + (b-m)^2 + (c-m)^2}{N} \quad \leftarrow (\text{데이터} - \text{평균})^2 \quad \leftarrow \text{제곱}$$

값 - 평균  $\Rightarrow$  편차: 얼마나 퍼져 있나?

분산: 편차의 제곱의 평균

제곱을 하는 이유?

음수인 이 되는 걸 막기 위해

그러나 제곱을 하면

값이 너무 커져서 감이 안 와.

$\rightarrow$  그래서  $\sqrt{\quad}$ 를 씀  $\Rightarrow$  표준편차

• 표준편차 (std  $\rightarrow$  standard deviation): 데이터와 비슷한 크기로 분산을 변환

$$\sigma = \sqrt{V} \quad 3 \pm 1 \quad \leftarrow \text{표준편차}$$

문제) 1, 2, 3의 분산과 표준편차

$$V = \frac{(1-2)^2 + (2-2)^2 + (3-2)^2}{3} = \frac{2}{3}$$

$$\sigma = \sqrt{\frac{2}{3}} = \frac{\sqrt{6}}{3}$$

대략 알아두면 좋음)

$$\begin{bmatrix} m - \sigma \\ m + \sigma \end{bmatrix}$$

데이터의  
68% 차지

