



기초통계

□ 평균: N개의 자료에 대해 산술평균을 구하는 방법

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

□ 자주 사용하는 기호

i 번째 관찰값 : x_i

모집단의 개수 : N

모집단 평균 : μ

표본의 개수 : n

표본평균 : \bar{x}

□ 중앙값: N개의 자료에 대해 정확히 가운데 있는 값을 의미

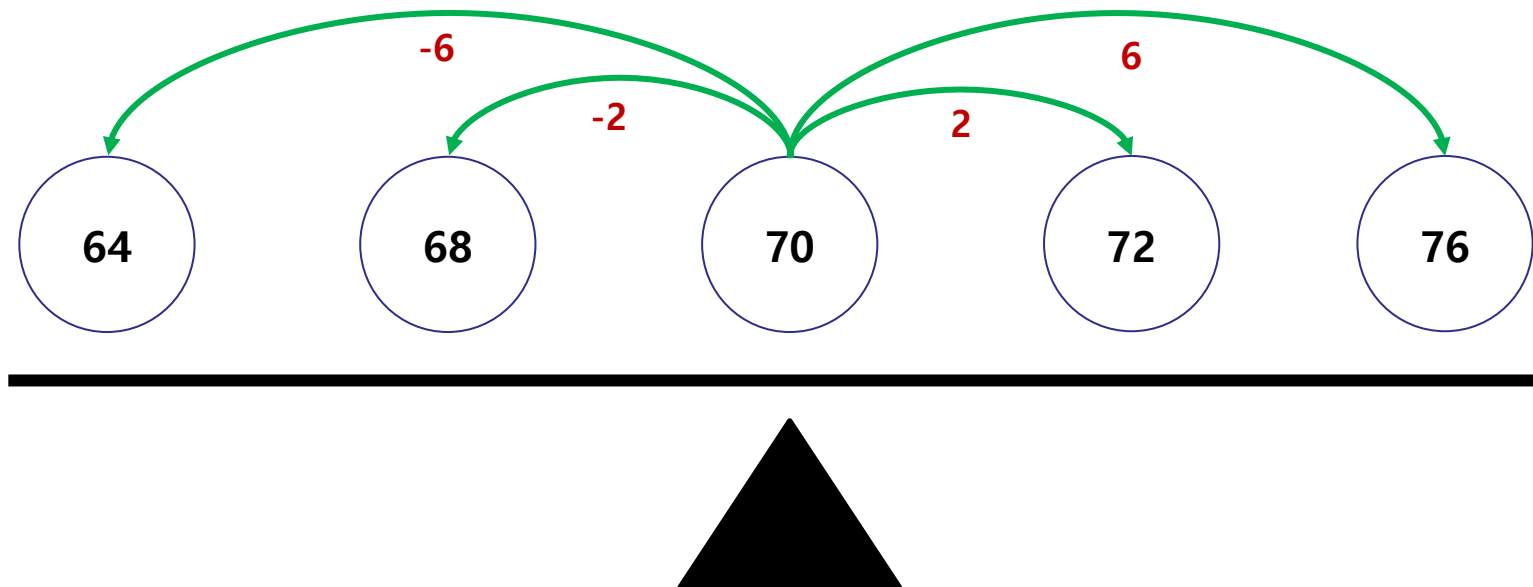
✓ 전체 자료의 개수(n)가 홀수일 때 $\rightarrow \frac{n+1}{2}$

✓ 전체 자료의 개수(n)가 짝수일 때 $\rightarrow \frac{n}{2}, \left(\frac{n}{2} + 1\right)$ 번째 값들의 평균

□ 편차

- ✓ 개별 자료와 평균 간의 개별 거리 측정

“각 자료들이 평균에 대해서 평균적으로 얼마나 떨어져 있을까?”



□ 편차

- ✓ 개별 자료와 평균 간의 개별 거리 측정

$$Deviation = X_i - \bar{X}$$

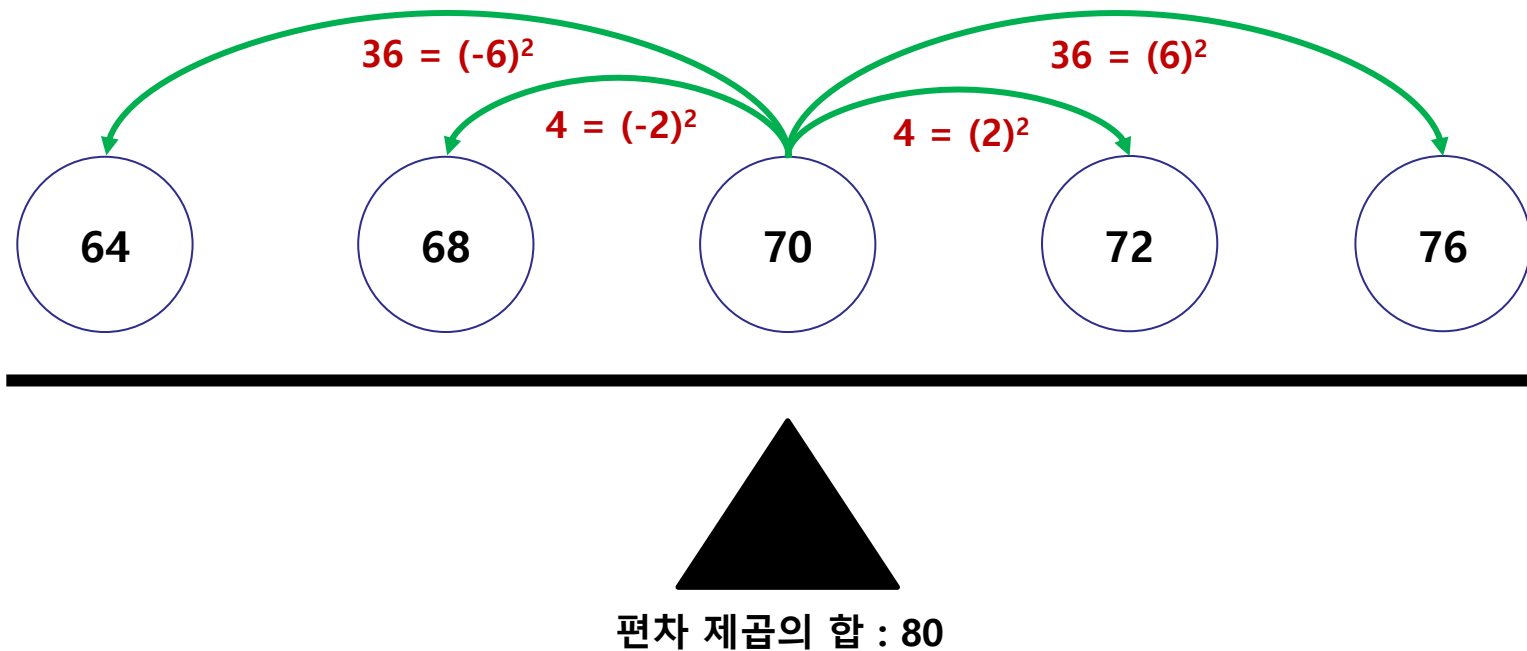
□ 주요 용어

- ✓ ***Deviation*** : 편차
- ✓ X_i : 개별 관찰값
- ✓ \bar{X} : 평균

□ 편차의 제곱

- ✓ 편차들을 양수로 만들기 위해 제곱을 사용함

거리 편차를 제곱하면 면적이 된다



□ 분산 (Variance)

- ✓ 편차 제곱의 평균

$$\text{Variance} = \frac{\sum (x_i - \bar{x})^2}{n}$$

□ 주요 용어

- ✓ *Variance* : 분산 (= 편차 제곱의 평균)
- ✓ X_i : 개별 관찰값
- ✓ \bar{X} : 평균

□ 표준편차 (Variance)

- ✓ 편차 제곱의 평균

$$\text{Standard Deviation} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n}}$$

□ 주요 용어

- ✓ **Standard Deviation** : 표준편차 (= 분산의 제곱근)
- ✓ x_i : 개별 관찰값
- ✓ \bar{x} : 평균

□ 모집단과 표본의 분산과 표준편차

구 분	모 집 단	표 본
분 산	$\sigma^2 = \frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N}$	$s^2 = \frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N - 1}$
표준편차	$\sigma = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N}}$	$s = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N - 1}}$