2.2 표준편차와 자유도

1) 표준편차와 자유도의 직관적 의미
2) 실생활에서의 표준편차와 자유도

1. 표준편차와 자유도의 직관적 의미

제곱근-평균-제곱 (Root Mean Square)

계산은 표현의 역순(제곱 후 평균, 최종적으로 제곱근)

(1) 제곱 (S) : 모든 수를 제곱하여 부호를 없앤다.

(2) 평균 (M) : 제곱된 값들의 평균을 구한다.

(3) 제곱근 (R) : 제곱-평균된 값에 제곱근을 취한다.

$$RMS = \sqrt{\text{숫자들의제곱의평균}}$$

예: 0, 1, -3, 3, -1의 RMS

$$\sqrt{\frac{0^2 + 1^2 + (-3)^2 + 3^2 + (-1)^2}{5}} = \sqrt{4} = 2$$

1. 표준편차와 자유도의 직관적 의미

표준편차의 계산

표준편차(SD)는 “평균으로부터의 편차들”의 RMS와 “대략” 비슷

표본 분산 및 표본 표준편차는

$$S^2 = \frac{(X_1 - \bar{X})^2 + (X_2 - \bar{X})^2 + \cdots + (X_n - \bar{X})^2}{n-1}$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2, \quad S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

예: 20, 10, 15, 15 에 대한 표준편차

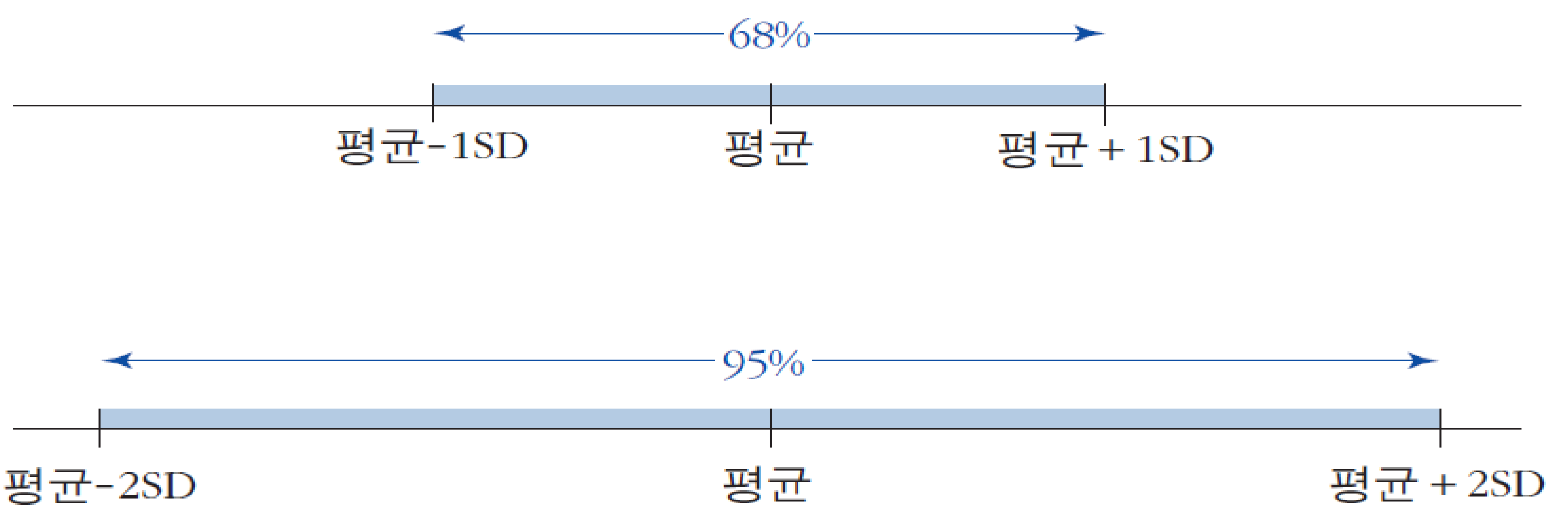
- 평균은 (20+10+15+15)/4 = 15이고, 평균으로부터의 편차들은 5, -5, 0, 0 이므로, 표준편차는

$$\sqrt{\frac{5^2 + (-5)^2 + 0^2 + 0^2}{4-1}} = \sqrt{16.7} \approx 4.1$$

1. 표준편차와 자유도의 직관적 의미

표준편차의 의미

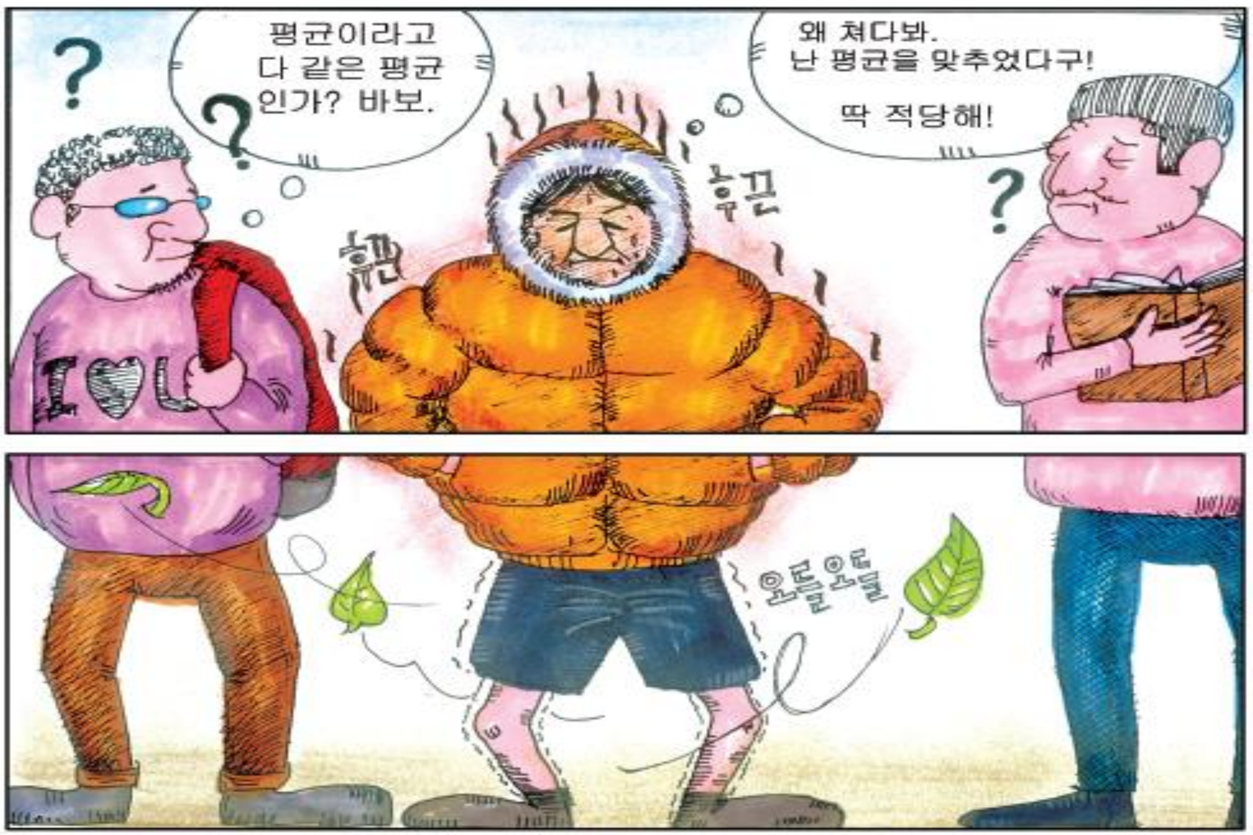
- 표준편차는 관측치들이 평균으로부터 얼마나 떨어져 있는지 알려줌
- 68-95법칙
  - 관측치들의 약 68% 정도가 평균으로부터 1 표준편차 이내로 떨어져 있다.
  - 관측치들의 약 95% 정도가 평균으로부터 2 표준편차 이내로 떨어져 있다.



류근관. (2013). 통계학, 제 3 판. 서울: 법문사. P. 61

1. 표준편차와 자유도의 직관적 의미

평균뿐만 아니라 표준편차도 중요



평균만으론 곤란!

류근관. (2013). 통계학, 제 3 판. 서울: 법문사. P. 52

1. 표준편차와 자유도의 직관적 의미

자유도의 정의

- 자유도는 합쳐진 값들 중에서 실질적으로 독립인 값들의 개수
- 표준편차 계산하는 경우의 자유도는 ”자료의 개수 - 1”
- 표준편차 계산의 대상이 되는 편차들의 합은 0이 됨. 편차들의 합이 0이 된다는 하나의 제약조건이 자유도를 1만큼 감소시킨 것임

1. 표준편차와 자유도의 직관적 의미

자유도의 정의

예: 극단적으로 자료의 개수가 하나인 경우

- 편차는 단 하나뿐이고 그 값은 0임
- 0에 대해 제곱의 평균, 즉 MS(mean square)를 구할 때 자유도 고려치 않으면 0/1=0이고 자유도를 고려하면 0/(1-1)=0/0으로 부정형(indefinite form)이 됨
- 단 하나의 자료만을 가지고는 퍼진 정도를 알 수 없음. 이 때 퍼진 정도는 0이 아니라 '알 수 없다(부정형)'가 정답임. 즉, 자유도를 고려해야 함

1. 표준편차와 자유도의 직관적 의미

정의

측정오차(measurement error)

- 관측치와 실제 값의 차이
- 측정오차가 존재하면, (관측치)=(실제 값)+(측정오차)
- 측정오차의 대략적인 크기는 관측치들의 표준편차(SD)를 통해 알 수 있음
- 표준편차(SD)의 크기는 한 번의 관측에서 측정오차가 어느 정도 될지 알려 줌

편의(bias)

- 방향성을 갖는 하나의 체계적인 오차
- 측정오차와 함께 편의가 있으면, (관측치)=(실제 값)+(편의)+(측정오차)

이탈값(outlier)

- 극단적인 관측치