

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

원래 분산의 정의(=편차의 제곱의 평균)에 따르면 우변에서 $n-1$ 이 아닌 n 으로 나눠줘야 합니다만, **통계적 추정에서 사용하는 표본분산** s^2 은 위와같이 구합니다. 그 이유는, 모집단으로부터 추출된 표본들의 실제 분산(n 으로 나뉘었을 때의 분산)은 모분산($=\sigma^2$)보다 항상 작거나 같게 나오는데 이를 달리 표현하면, 실제 분산의 정의를 통해 구해진 표본들의 분산 값이 모분산으로 쓰기에 다소 저평가되어있다는 말입니다. 따라서 그 값을 모분산 값과 유사하게 맞춰주기 위해 (n 으로 나누는 대신) 그보다 더 작은 $n-1$ 로 나누어서 값을 키워줍니다. 이런 연유로 표본분산을 표기할 때에는 실제 변량들의 분산과 구분하기 위해 s^2 으로 나타냅니다. (여기서 s 는 표준편차(standard deviation)의 약자입니다.) 이에 대한 내용은 Bessel이라는 사람이 제안했으며, 통계학에서는 **Bessel's correction**이라고 부릅니다.

분량이 작아질수록, 그 값은 더 커진다!

* 모평균 추릴 때, 표본분산을 모분산 대신 사용한다!!