
자료의 수치적 표현

1. 척도

A. 위치척도 (중심, location measure)

- i. 평균 (Mean, Average)
- ii. 중앙값 (Median)
- iii. 최빈값 (Mode)
- iv. 사분위수(Quartile) & 백분위수 (Percentile)

B. 산포척도 (dispersion measure)

- i. 범위 (Range)
- ii. 사분위수 범위 (Interquartile Range)
- iii. 분산 & 표준편차 (Variation & Standard Deviation)
- iv. 변동계수 (Coefficient of Variation)

C. 모양척도 (Shape measure)

- i. 왜도 (Skewness)
- ii. 첨도 (Kurtosis)

2. 위치척도 - 평균 (Mean)

- A. Mean (통계학) : 모평균, 표본 평균, ...
- B. Average (수학) : 산술, 기하, 조화, 이차, 가중 ...

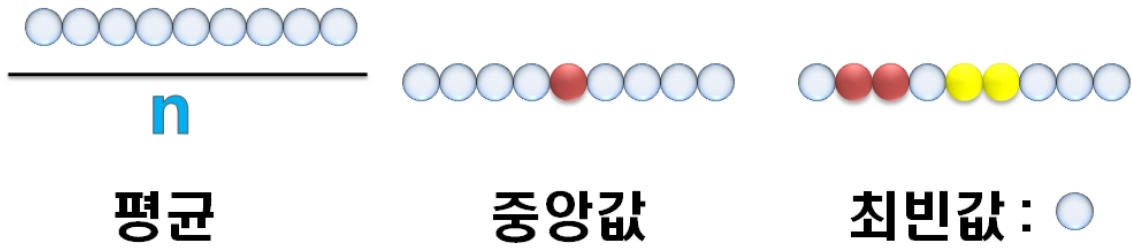
$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

3. 위치척도 - 중앙값 (Median)

- A. 자료를 순서대로 나열
- B. 가장 가운데 있는 값
- C. 자료의 수가 홀수인 경우
- D. 자료의 수가 짝수인 경우

4. 위치척도 - 최빈값 (Mode)

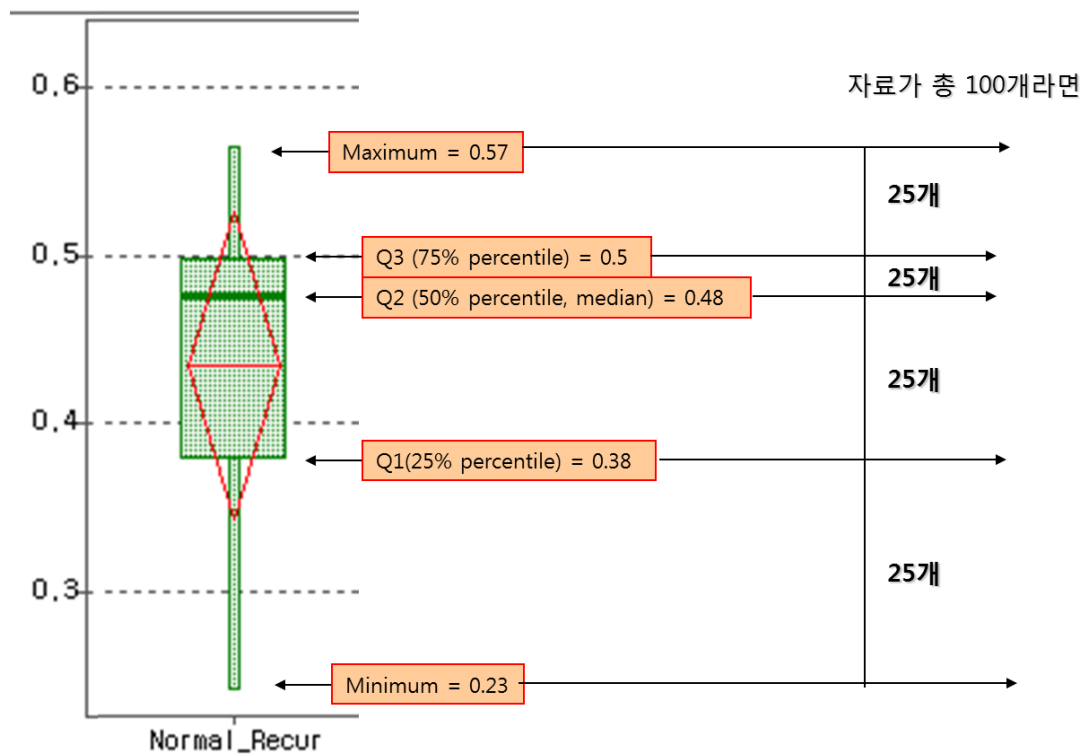
A. 가장 빈도(frequency)가 높은 값



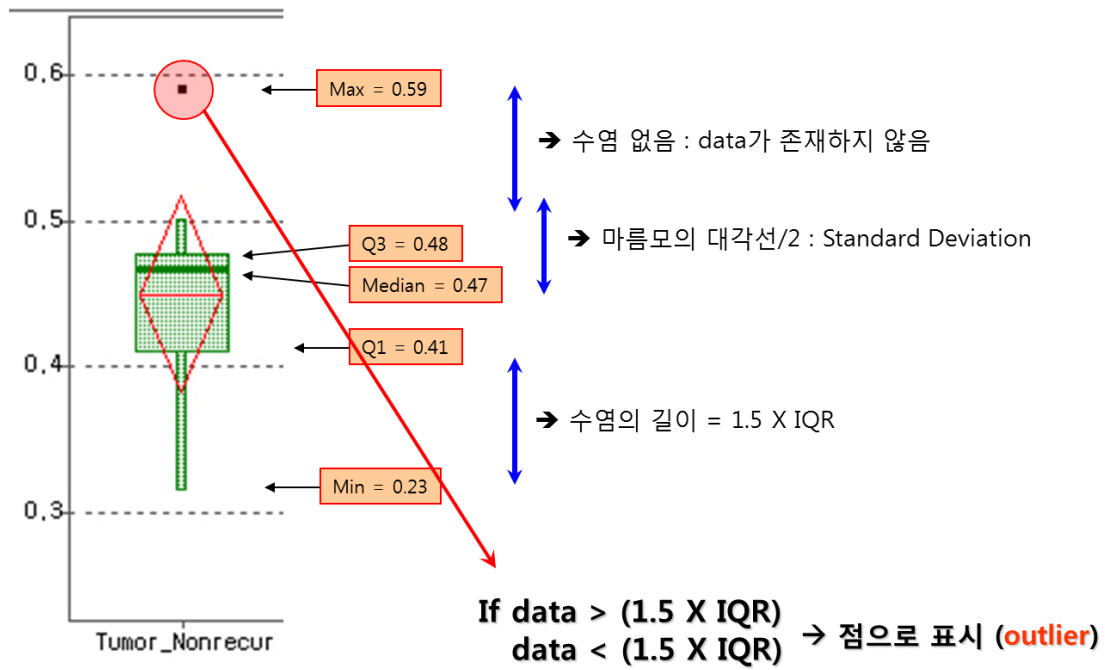
5. 위치척도 - 사분위수, 백분위수

A. 사분위수 (Quartile)

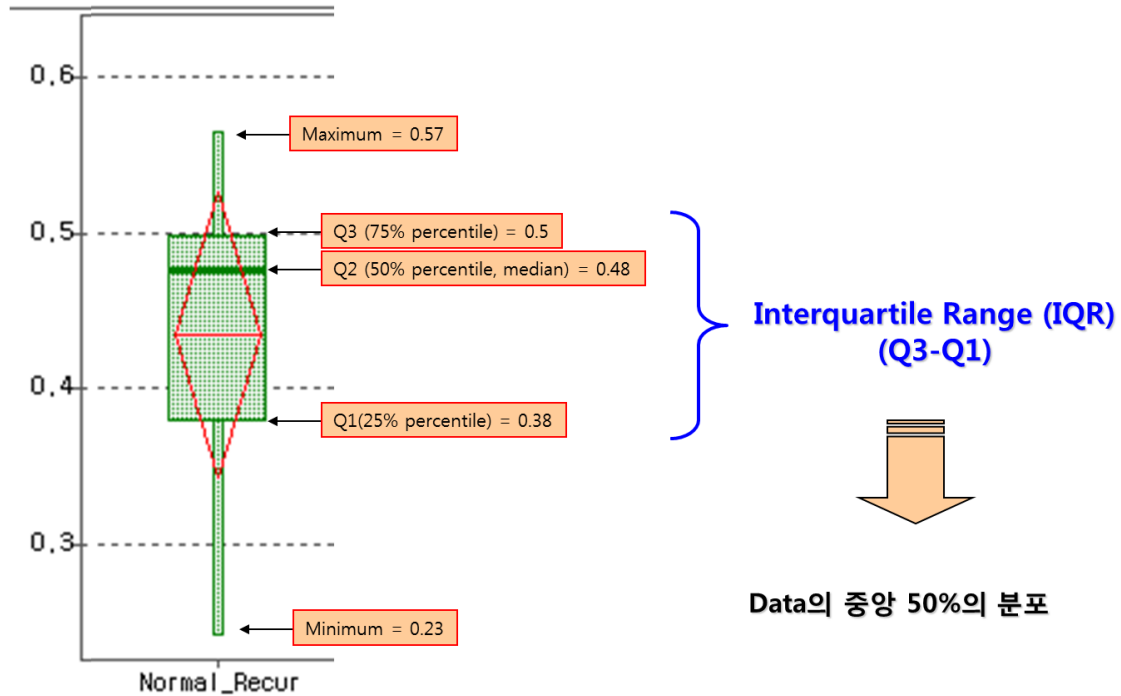
B. 백분위수 (Percentile)



6. 산포척도 - 범위 (Range)



7. 산포척도 - IQR, 분산, 표준편차



분산

$$Var(X) = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$$

자유도 (degree of freedom)

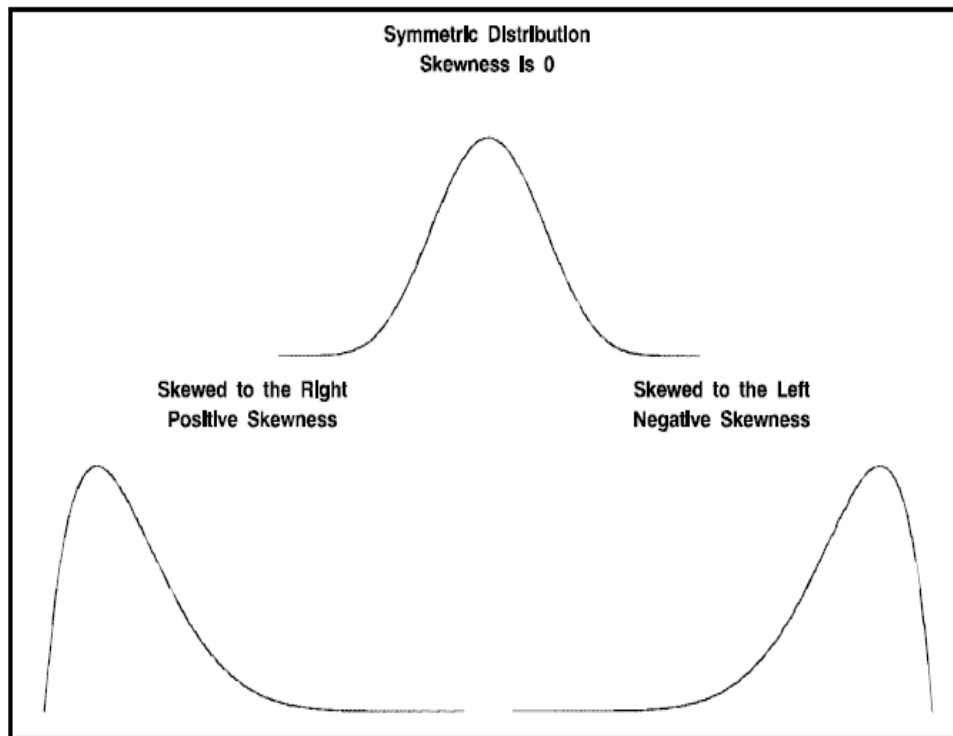
8. 산포척도 - 변동계수 (Coefficient of Variation)

A. 두 집단의 분포 비교

B. 단위에 대한 보정

$$CV = \frac{S}{\bar{X}}$$

9. 분포의 모양 척도 - 왜도 (Skewness)



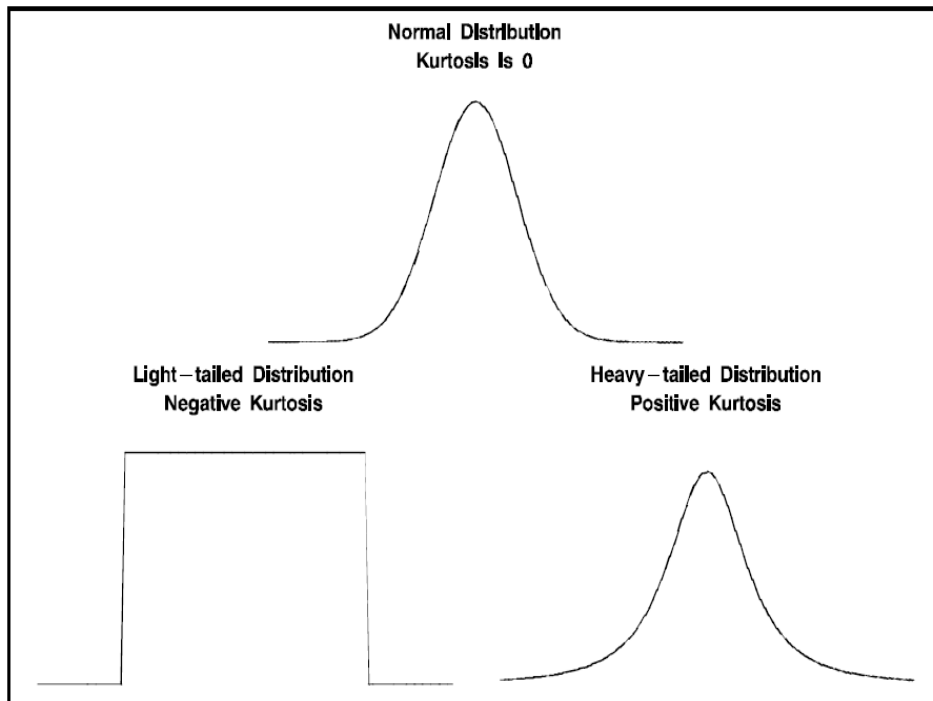
$$S_k = \sum_{i=1}^n \frac{[(X_i - \bar{X})/s]^3}{n-1}$$

$S_k=0$, 좌우대칭

$S_k<0$, 좌측으로 긴 꼬리

$S_k>0$, 우측으로 긴 꼬리

10. 분포의 모양 척도 - 첨도 (Kurtosis)



$$K = \sum_{i=1}^n \frac{[(X_i - \bar{X})/s]^4}{n-1} - 3$$

$K=0$, 표준정규분포와 뽕족한 정도가 같다.

$K<0$, 표준정규분포보다 납작하다.

$K>0$, 표준정규분포보다 뽕족하다.

11. 분포의 비교

- A. 평균 = 중앙값
- B. 평균 < 중앙값
- C. 평균 > 중앙값

