

상관분석

숙명여자대학교 경영학부 오중산

상관관계

- 상관관계(correlation)란?

- ◆ 상관관계는 두 변수 간의 선형적(linear)인 관계를 의미함
- ◆ 상관관계에서는 두 변수의 변화 ‘방향’과 ‘강도’, 두 가지 측면을 고려해야 함
 - 두 변수가 같은 방향으로 변화하는지, 반대 방향으로 움직이는지?
 - 두 변수가 변화하는 강도(크기)가 얼마나 되는지?
- ◆ 상관계수(correlation coefficient)
 - 보통 표본상관계수(추정치)로 모집단 상관계수를 추정하는데, 통계적 유의성 확인이 중요함
 - 추정된 상관계수의 부호(변화 방향)와 크기(변화 강도)를 통해 상관관계를 파악할 수 있음

상관관계

- 피어슨 상관계수(Pearson's coefficient of correlation)

- ◆ 전제조건: 두 변수가 정량적 변수여야 하며, 정규성 조건을 만족해야 함

- ◆ 피어슨 상관계수의 특징

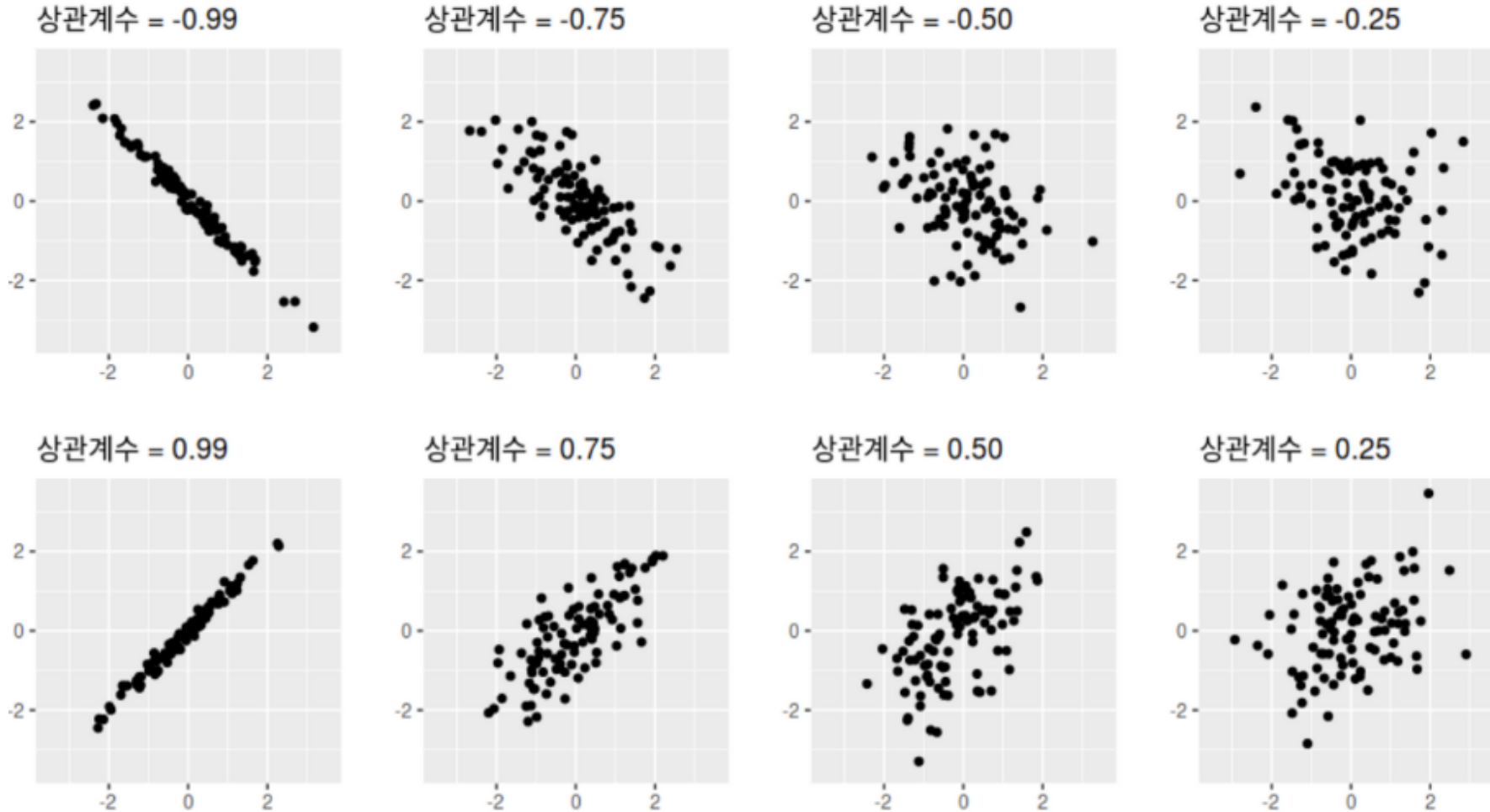
- -1에서 1사이 값으로 음수면 두 변수가 반대방향으로, 양수면 같은 방향으로 변화함
- 0에 가까울수록 상관관계가 약하고, ± 1 에 가까울수록 상관관계가 강함

$$\text{(모집단상관계수)} \quad \rho_{XY} = \frac{\text{Cov}(X,Y)}{\sigma_X \cdot \sigma_Y} = \frac{E[(X-E[X]) \cdot (Y-E[Y])]}{\sigma_X \cdot \sigma_Y} = \frac{E[XY] - E[X]E[Y]}{\sigma_X \cdot \sigma_Y}$$

$$\text{(표본상관계수)} \quad r_{XY} = \frac{\sum (X - \bar{X}) \cdot (Y - \bar{Y})}{\sqrt{\sum (X - \bar{X})^2 \cdot \sum (Y - \bar{Y})^2}} = \frac{1}{n-1} \cdot \sum \frac{(X - \bar{X})}{s_X} \cdot \frac{(Y - \bar{Y})}{s_Y} \quad t(n-2) = r_{XY} \cdot \sqrt{\frac{n-2}{1-r_{XY}^2}}$$

상관관계

- 피어슨 상관계수(Pearson's coefficient of correlation)



상관관계

- 상관관계와 인과관계(causality)

- ◆ 인과관계 → 상관관계

- 상관관계는 필요조건이고, 인과관계는 충분조건
- 인과관계는 상관관계의 특수한 형태

- ◆ X(원인)와 Y(결과) 간의 인과관계 성립 요건

- 시간: X가 Y보다 선행해야 함
- 전제: X가 없으면 Y가 존재할 수 없음
- 설명력: Y에 대해 X가 논리적으로 잘 설명할 수 있어야 함
- 지속성: 시간이 흘러도 Y에 대한 X의 설명력이 유지되어야 함

