KMO와 Bartlett 검사

# 1. KMO (Kaiser-Meyer-Olkin) 검사

KMO 검사는 변수들 간의 상관관계가 요인분석에 적합한지를 평가하는 지표입니다. 이는 변수들 간의 상관관계의 정도와 공통성(communality)을 비교하여 계산됩니다.

## KMO 값의 의미

KMO 값은 0과 1 사이의 값을 가지며, 값이 클수록 요인분석에 적합하다는 것을 의미합니다.

- 0.90 이상: 매우 적합  
- 0.80 ~ 0.89: 적합  
- 0.70 ~ 0.79: 보통  
- 0.60 ~ 0.69: 약간 부적합  
- 0.50 ~ 0.59: 부적합  
- 0.50 미만: 매우 부적합, 요인분석을 수행하지 않는 것이 좋습니다.

## KMO 검사의 계산 방법

KMO 공식:

KMO = ΣΣ r\_ij^2 / (ΣΣ r\_ij^2 + ΣΣ p\_ij^2)

여기서, r\_ij는 상관계수, p\_ij는 편상관계수입니다.

## KMO 검사의 목적

KMO 검사는 요인분석을 수행하기 전에 데이터의 적합성을 판단하는 데 중요한 역할을 합니다. 상관관계가 높은 변수들 간에 공통 요인이 존재할 가능성이 크기 때문에, KMO 값이 높을수록 요인분석의 결과가 더 신뢰할 수 있게 됩니다.

# 2. Bartlett의 구형성 검정 (Bartlett's Test of Sphericity)

Bartlett의 구형성 검정은 변수들 간의 상관행렬이 단위행렬(Identity Matrix)인지 여부를 검정합니다. 단위행렬이란, 대각선 요소가 1이고 나머지 요소가 모두 0인 행렬을 의미합니다.

## Bartlett 검정의 목적

Bartlett 검정은 요인분석을 수행하기 전에 변수들 간의 상관관계가 통계적으로 유의미한지를 평가하는 데 사용됩니다. 만약 상관행렬이 단위행렬이라면, 변수들 간의 상관관계가 없으므로 요인분석이 의미가 없습니다.

## Bartlett 검정의 계산 방법

기본 가정: 상관행렬이 단위행렬과 다르다 (즉, 상관관계가 있다).

검정 결과: p-value가 0.05보다 작으면, 상관행렬이 단위행렬이 아니므로 요인분석이 적합하다는 결론을 내릴 수 있습니다.

검정 통계량:

χ^2 = -[n - 1 - (2p + 5)/6] log |R|

여기서, n은 표본 수, p는 변수의 수, R은 상관행렬입니다.

## Bartlett 검정의 해석

- p-value < 0.05: 상관행렬이 단위행렬과 유의미하게 다르므로, 요인분석이 적합합니다.  
- p-value ≥ 0.05: 상관행렬이 단위행렬과 크게 다르지 않으므로, 요인분석이 부적합할 수 있습니다.

# 요약

KMO 검사는 데이터의 요인분석 적합성을 평가하는 지표로, 값이 높을수록 요인분석이 적합하다는 것을 의미합니다. Bartlett의 구형성 검정은 변수들 간의 상관행렬이 단위행렬인지 여부를 검정하여, 요인분석이 의미 있는지 판단합니다. 요인분석을 수행하기 전에 KMO 검사와 Bartlett 검정을 통해 데이터가 요인분석에 적합한지 확인하는 것이 중요합니다. KMO 값이 높고 Bartlett 검정에서 p-value가 낮을수록 요인분석 결과의 신뢰성이 높아집니다.