고전적 최소자승법(OLS)과 최대우도법(ML) 비교

# 1. 고전적 최소자승법(OLS: Ordinary Least Squares)

## 이론

OLS는 회귀분석에서 종속변수와 독립변수 간의 관계를 설명하는 모형의 파라미터를 추정하는 방법입니다. 이 방법은 예측된 종속변수와 실제 종속변수 사이의 차이, 즉 잔차(residuals)의 제곱합을 최소화하는 파라미터를 찾는 것을 목표로 합니다.

## 수식

OLS의 기본 수식은 다음과 같습니다.  
- 회귀모형: Y\_i = β\_0 + β\_1 X\_i + ε\_i  
 - Y\_i : 종속변수  
 - X\_i : 독립변수  
 - β\_0 , β\_1 : 회귀계수 (추정하고자 하는 값)  
 - ε\_i : 오차항  
- OLS의 목표: 잔차의 제곱합을 최소화  
 - Minimize ∑(ε\_i^2) = ∑(Y\_i - Ŷ\_i)^2 = ∑(Y\_i - β\_0 - β\_1 X\_i)^2

## 실제 사례

예를 들어, 주택 가격(종속변수)과 방의 수(독립변수) 간의 관계를 분석할 때 OLS를 사용할 수 있습니다.  
1. 데이터를 수집하고, 주택 가격과 방의 수 데이터를 확보합니다.  
2. OLS를 통해 주택 가격을 방의 수로 설명하는 회귀계수를 추정합니다.  
 - 예를 들어, 방의 수가 1개 증가할 때 주택 가격이 평균적으로 50만 원 증가하는 것으로 추정될 수 있습니다.  
3. 결과적으로 방의 수가 3개인 주택의 예측 가격은 OLS 회귀식에 따라 계산할 수 있습니다.

# 2. 최대우도법(ML: Maximum Likelihood)

## 이론

최대우도법은 주어진 데이터가 발생할 확률을 최대화하는 파라미터를 찾는 방법입니다. 이 방법은 특히 복잡한 모형이나 비정규 분포 데이터에서도 유용하게 사용됩니다. ML은 데이터가 특정 분포를 따른다고 가정하고, 그 분포에서 관측된 데이터를 설명하는 파라미터를 추정합니다.

## 수식

ML의 기본 원리는 다음과 같습니다.  
- L(θ | X) = P(X | θ)  
 - L(θ | X): 주어진 데이터 X에서 파라미터 θ의 우도(likelihood)  
 - P(X | θ): 주어진 파라미터 θ에서 데이터 X가 발생할 확률  
- 최대우도추정법의 목표: 우도함수 L(θ | X)를 최대화하는 θ를 찾음  
 - θ̂ = arg max\_θ L(θ | X)

## 실제 사례

예를 들어, 정규 분포를 가정한 데이터에서 평균과 분산을 추정하는 경우 최대우도법을 사용할 수 있습니다.  
1. 데이터가 정규 분포를 따른다고 가정합니다.  
2. 관측된 데이터에서 정규 분포의 평균과 분산을 추정하기 위해 최대우도법을 적용합니다.  
 - μ̂와 σ̂^2를 추정합니다.  
3. 이 추정치를 사용해 새로운 데이터가 발생할 확률을 계산하거나, 모델을 구축할 수 있습니다.

# 3. OLS와 ML의 비교

- OLS는 주로 선형 회귀 모형에서 사용되며, 잔차의 제곱합을 최소화하는 방향으로 회귀계수를 추정합니다.  
 - 사용 예: 선형 회귀, 다중 회귀  
- ML은 보다 일반적인 추정 방법으로, 비선형 회귀, 로지스틱 회귀 등 다양한 통계 모형에 적용될 수 있습니다.  
 - 사용 예: 로지스틱 회귀, 프로빗 모형, 포아송 회귀  
  
실제 분석 상황에서 OLS와 ML은 데이터의 특성, 모형의 복잡성, 가정된 분포에 따라 선택됩니다.