**가우스-마르코프 정리 (Gauss-Markov Theorem)**

# 1. 이론 개요

가우스-마르코프 정리는 선형 회귀 분석에서 최소자승법(OLS)으로 추정된 회귀 계수들이 어떤 조건 하에서 가장 효율적인 추정량이 되는지를 설명하는 중요한 통계 이론입니다.

- 핵심 가정: 가우스-마르코프 정리가 성립하기 위해서는 다음의 기본 가정들이 필요합니다.

1. **모형의 선형성**: 종속 변수 Y와 독립 변수 X 사이의 관계가 선형적이어야 합니다.

2. **오차항의 기댓값이** 0: 모든 관측치에 대해 오차항의 기댓값이 0이어야 합니다. 즉, E(ε\_i) = 0.

3. **오차항의 등분산성**: 오차항의 분산이 일정해야 합니다. 즉, Var(ε\_i) = σ^2 (모든 i에 대해 동일).

4. **오차항 간의 비상관성**: 서로 다른 관측치의 오차항들 간에 상관관계가 없어야 합니다. 즉, Cov(ε\_i, ε\_j) = 0 for i ≠ j.

5. 독립 변수들 간의 완전한 선형 독립성: 독립 변수들이 서로 독립적이어야 하며, 다중공선성이 없어야 합니다.

- **결론**: 위의 가정들이 충족될 경우, 최소자승법(OLS)으로 계산된 회귀 계수들은 **최량선형불편추정치(BLUE: Best Linear Unbiased Estimator)**임이 보장됩니다.

- **최량**: **최소 분산**을 가짐.

- **선형**: **독립 변수의 선형 결합**으로 표현됨.

- **불편**: **추정량의 기댓값이 참값과 일치함**.

# 2. 수식적 설명

- 선형 회귀 모형: Y = Xβ + ε

- 여기서,

- Y는 n × 1 벡터로, 종속 변수 (관측치)

- X는 n × k 행렬로, 독립 변수 (설명 변수)

- β는 k × 1 벡터로, 회귀 계수

- ε는 n × 1 벡터로, 오차항

- 최소자승법 (OLS) 추정량: OLS 추정량 β̂는 다음과 같이 계산됩니다.

β̂ = (X'X)^(-1)X'Y

여기서 X'는 X의 전치 행렬을 의미합니다.

- 추정량의 분산: OLS 추정량 β̂의 분산-공분산 행렬은 다음과 같이 주어집니다.

Var(β̂) = σ^2 (X'X)^(-1)

- 가우스-마르코프 정리: 위의 기본 가정들이 모두 충족될 경우, β̂는 최량선형불편추정치(BLUE)가 됩니다.

즉, OLS 추정량이 최적의 선형 추정량이며, 분산이 최소화됩니다.

# 3. 실제 사례

- 사례 1: 경제학에서 소득 예측

한 나라의 개인 소득을 예측하기 위해 교육 수준, 직업 유형, 거주 지역 등의 변수들을 사용해 선형 회귀 모형을 설정할 수 있습니다.

이때, OLS 방법을 사용하여 회귀 계수를 추정하면, 가우스-마르코프 정리의 가정이 충족된다면 이 추정량은 BLUE가 됩니다.

예를 들어, 특정 지역에서 교육 수준이 소득에 미치는 영향을 분석할 때, OLS 추정량은 가장 효율적인 추정치가 됩니다.

- 사례 2: 주식 시장 분석

주식 수익률을 예측하기 위해 경제적 지표들(예: GDP 성장률, 금리, 인플레이션 등)을 독립 변수로 사용하여 회귀 분석을 수행할 수 있습니다.

이 경우, 주식 수익률이 경제 지표와의 선형 관계를 따른다는 가정 하에 OLS 방법을 사용하여 회귀 계수를 추정합니다.

가우스-마르코프 정리가 적용될 수 있다면, 이 추정치는 가장 신뢰할 수 있는 예측치가 됩니다.

- 사례 3: 의료 데이터 분석

환자의 건강 상태(예: 혈압)를 나이, 체중, 운동 습관 등으로 설명하는 회귀 모형을 설정할 수 있습니다.

이 모형에서 OLS를 사용하여 각 변수들이 혈압에 미치는 영향을 추정할 수 있으며, 가정이 충족된다면 이 추정량은 최량선형불편추정치(BLUE)가 됩니다.

예를 들어, 나이가 혈압에 미치는 영향이 얼마나 되는지 정확하게 평가할 수 있습니다.