SPSS 선형 회귀분석의 방법론

# 1. 입력(Enter)

## 이론적 배경

수식: Y = β₀ + β₁X₁ + β₂X₂ + ... + βₙXₙ + ε  
입력 방법에서는 모든 독립변수 X₁, X₂, ... , Xₙ를 한꺼번에 포함하여 모델을 구성합니다. 이 방법은 주어진 변수들이 모두 중요하다고 가정하고, 전체 회귀모델을 한 번에 평가합니다.

## 구체적인 방법론

절차:  
1. 모든 독립변수를 모델에 포함시킵니다.  
2. 회귀계수(β)와 유의성(p-값)을 계산합니다.  
3. 전체 모델의 적합성을 평가합니다 (예: R², F-통계량).  
4. 각 변수의 영향력을 개별적으로 평가하고 해석합니다.

# 2. 단계선택(Stepwise)

## 이론적 배경

수식: 단계선택 방법에서는 여러 회귀모델을 평가하며, 각 모델에서 변수의 추가와 제거를 반복합니다. 이 과정에서 AIC (Akaike Information Criterion)나 BIC (Bayesian Information Criterion) 같은 정보기준을 사용해 모델 적합도를 평가합니다.

## 구체적인 방법론

절차:  
1. 모델에 아무것도 포함하지 않은 상태에서 시작합니다.  
2. 모든 독립변수를 평가하여 가장 유의미한 변수를 추가합니다.  
3. 새 변수를 추가한 후 모델의 적합도를 평가하고, 필요 시 유의하지 않은 변수를 제거합니다.  
4. 이 과정을 유의미한 변화가 없을 때까지 반복합니다.  
5. 최종적으로 선택된 변수들로 회귀모델을 확정합니다.

# 3. 제거(Remove)

## 이론적 배경

수식: 제거 방법은 특정 변수 X\_j를 회귀모델에서 제거함으로써 전체 모델의 적합도를 다시 평가합니다. 이를 통해 제거된 변수가 모델의 설명력에 미치는 영향을 분석합니다.  
수식적으로는, 제거된 변수 X\_j를 포함하지 않은 상태에서 회귀모델을 다시 구성하게 됩니다: Y = β₀ + Σ (i ≠ j) βᵢXᵢ + ε

## 구체적인 방법론

절차:  
1. 초기 모델에 포함된 변수를 기준으로 시작합니다.  
2. 연구자가 지정한 변수를 모델에서 제거합니다.  
3. 모델의 적합도를 평가하고, 변수 제거가 모델에 미치는 영향을 분석합니다.  
4. 필요시 추가로 변수를 제거하여 최종 모델을 확정합니다.

# 4. 후진(Backward Elimination)

## 이론적 배경

수식: 후진 제거 방법은 모든 변수를 포함한 모델로 시작한 후, 가장 유의하지 않은 변수부터 하나씩 제거해 나갑니다. 모델의 적합도는 잔차 제곱합(Residual Sum of Squares, RSS)이나 AIC, BIC 등으로 평가합니다.  
RSS = Σ (i=1 to n) (yᵢ - ŷᵢ)²  
AIC = 2k - 2ln(L)  
BIC = ln(n)k - 2ln(L)  
여기서 k는 모델의 변수 수, L은 최대우도함수, n은 샘플 크기입니다.

## 구체적인 방법론

절차:  
1. 모든 독립변수를 포함한 초기 모델에서 시작합니다.  
2. 가장 유의하지 않은 변수(높은 p-값을 가진 변수)를 모델에서 제거합니다.  
3. 모델의 적합도를 다시 평가합니다.  
4. 변수를 계속 제거하여, 유의미한 변수만 남을 때까지 이 과정을 반복합니다.  
5. 최종적으로 남은 변수들로 회귀모델을 확정합니다.

# 5. 전진(Forward Selection)

## 이론적 배경

수식: 전진 선택 방법은 빈 모델에서 시작하여 가장 유의미한 변수를 하나씩 추가하는 방식입니다. 새로운 변수를 추가할 때마다 모델의 적합도를 평가하여, 선택된 변수가 모델에 실질적인 기여를 하는지 판단합니다.  
- 초기 모델: Y = β₀ + ε  
- 첫 번째 추가된 변수: Y = β₀ + β₁X₁ + ε  
- 계속해서 추가: Y = β₀ + β₁X₁ + β₂X₂ + ... + ε

## 구체적인 방법론

절차:  
1. 아무런 변수가 포함되지 않은 상태에서 시작합니다.  
2. 모든 독립변수를 평가하여, 가장 유의미한 변수(p-값이 가장 낮은 변수)를 모델에 추가합니다.  
3. 모델의 적합도를 평가하고, 필요시 유의미한 변수를 추가합니다.  
4. 변수를 더 이상 추가해도 모델의 적합도가 개선되지 않을 때까지 이 과정을 반복합니다.  
5. 최종적으로 선택된 변수들로 회귀모델을 확정합니다.