SPSS에서의 선형 회귀분석 통계량 및 진단 지표 설명

# 회귀계수 관련 통계량

## 1. 회귀계수의 추정값(Estimate)

회귀모델에서 독립변수가 종속변수에 미치는 영향을 수치로 표현한 값입니다.  
예를 들어, 회귀계수 값이 0.5라면, 독립변수가 1 단위 증가할 때 종속변수가 평균적으로 0.5 단위 증가함을 의미합니다.

## 2. 신뢰구간(Confidence Interval)

회귀계수의 참값이 포함될 것으로 기대되는 구간입니다.  
보통 95% 신뢰구간을 사용하며, 이 구간이 넓을수록 추정값의 불확실성이 크다는 것을 의미합니다.

## 3. 공분산 행렬(Covariance Matrix)

회귀계수 간의 공분산을 나타내는 행렬로, 각 회귀계수 간의 관계를 보여줍니다.  
이 행렬을 통해 회귀계수 간의 상관성을 파악할 수 있습니다.

# 잔차 관련 통계량

## 1. PRESS (Predicted Residual Error Sum of Squares)

회귀분석에서 각 관측치에 대해 해당 관측치를 제외하고 추정된 모델로 예측한 값과 실제 값의 차이의 제곱합입니다.  
이 값이 작을수록 모델의 예측력이 좋음을 의미합니다.

## 2. Durbin-Watson 통계량

잔차의 자기상관성을 검정하는 통계량입니다.  
2에 가까운 값은 자기상관이 없음을, 2에서 멀어질수록 자기상관이 있음을 의미합니다.

## 3. 케이스별 진단(Casewise Diagnostics)

각 관측치에 대해 잔차가 일정 기준 이상으로 클 때, 이 관측치가 이상치(outlier)일 가능성을 진단하는 데 사용됩니다.

# 그 외 주요 통계량

## 1. 모형적합(Overall Model Fit)

회귀모델 전체의 적합성을 평가하는 통계량으로, 대표적으로 F-통계량과 p-값이 사용됩니다.  
F-통계량이 크고 p-값이 0.05 이하이면 모델이 통계적으로 유의미함을 나타냅니다.

## 2. R제곱 변화량(Change in R-squared)

모델에 새로운 독립변수를 추가했을 때 설명력이 얼마나 증가했는지를 보여줍니다.  
증가량이 클수록 추가된 변수가 모델에 중요한 기여를 했음을 나타냅니다.

## 3. 기술통계(Descriptive Statistics)

분석에 사용된 데이터의 평균, 표준편차, 최소값, 최대값 등을 나타내는 기본적인 통계량입니다.

## 4. 부분상관 및 편상관계수(Partial and Part Correlations)

부분상관계수는 특정 변수를 통제한 후 다른 두 변수 간의 상관성을 나타내고,  
편상관계수는 회귀분석에서 다른 변수의 영향을 배제한 상태에서의 두 변수 간의 순수한 상관성을 나타냅니다.

## 5. 공선성 진단(Collinearity Diagnostics)

독립변수 간의 상관성이 높은 경우(다중공선성)를 진단하는 지표입니다.  
VIF(Variance Inflation Factor)와 Condition Index 등이 사용되며, VIF 값이 10을 넘으면 다중공선성이 있다고 판단할 수 있습니다.

## 6. 선택기준(Selection Criteria)

모델 선택 시 사용하는 기준으로, AIC(Akaike Information Criterion)와 BIC(Bayesian Information Criterion) 등이 있습니다.  
값이 작을수록 모델이 데이터에 더 잘 적합된다고 평가합니다.