1. 회귀분석

회귀분석: 자료를 통해 종속변수와 독립변수간의 함수관계를 통계적으로 찾아보는 분석방법

독립변수(설명변수): 다른 변수에 영향을 주는 변수

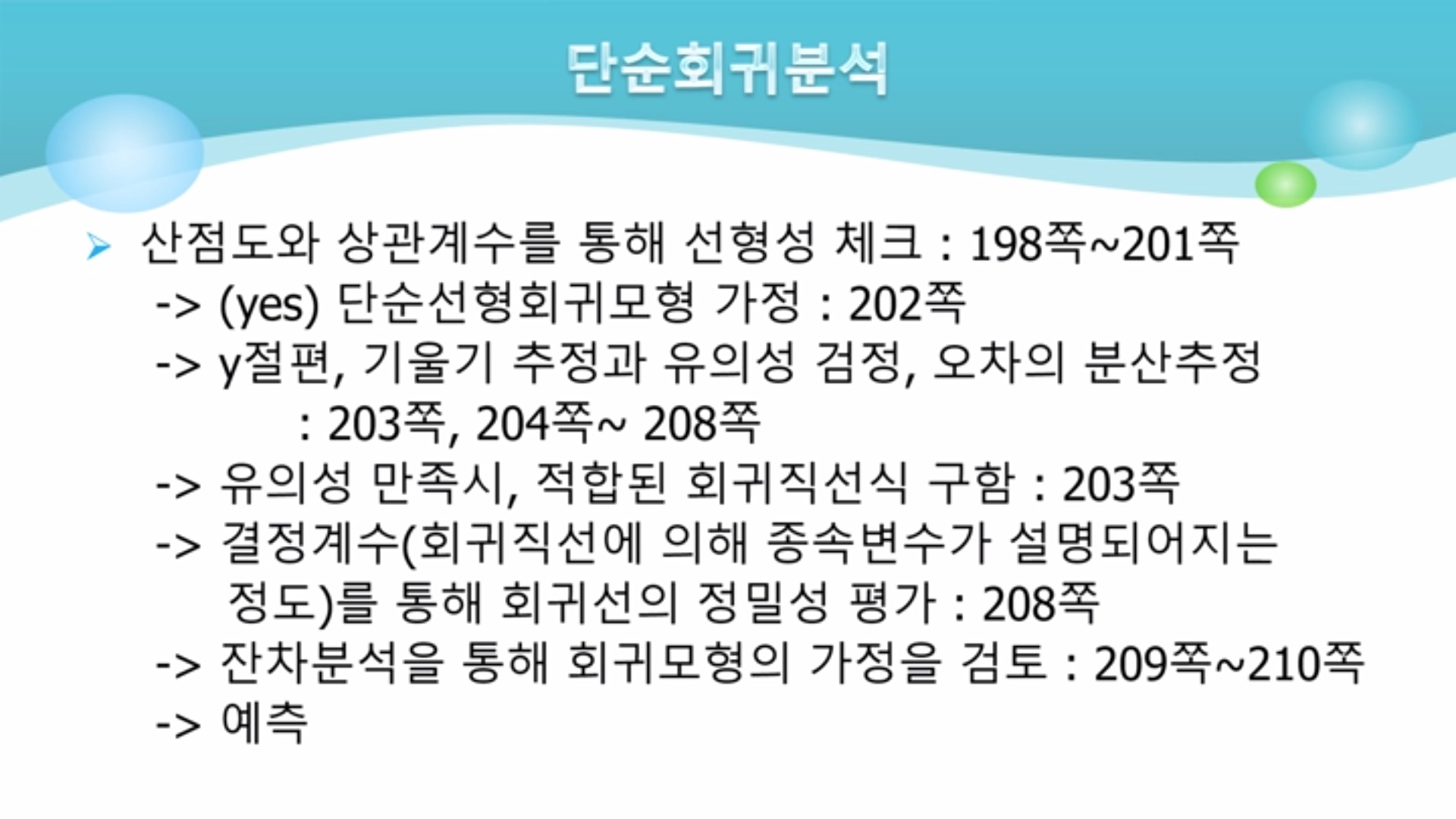
종속변수(반응변수): 다른 변수의 영향을 받는 변수

ㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡ

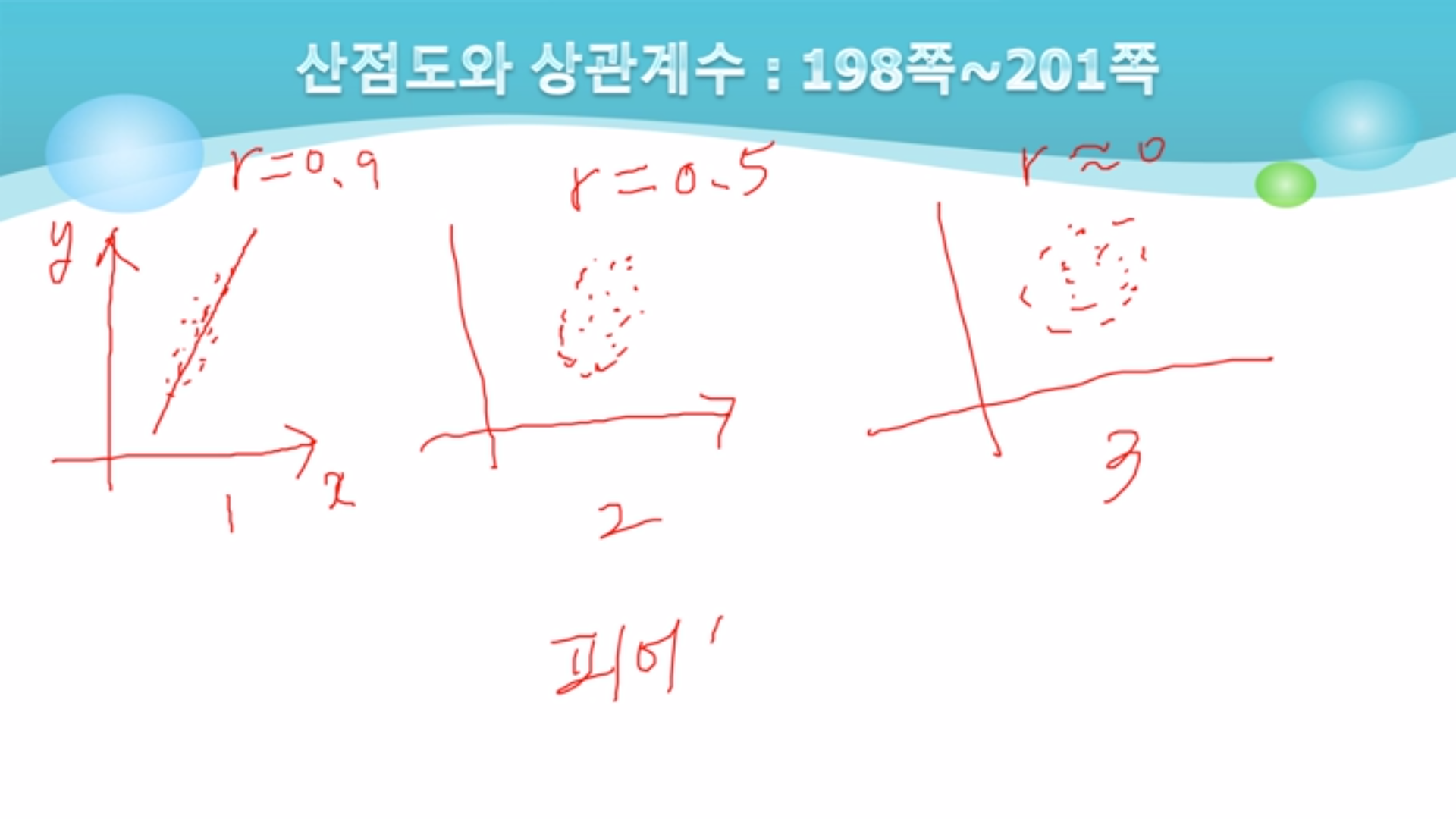
회귀분석 과정: 그림을 통해 두 변수들 사이의 연관성 파악 (산점도 이용) -> 변수들 사이의 연관성을 계량화된 척도로 파악 (상관계수) -> 인과관계를 지니는 경우 함수적인 관련성을 규명하기 위해 통계적 모형을 가정하고, 측정된 변수들의 자료를 이용하여 모형에 대한 추론을 하고, 정밀성을 평가하고 가정을 검토 (회귀분석)

ㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡㅡ

단순회귀분석: 1개의 독립변수가 1개의 종속변수에 영향을 주는 선형모형을 사용한 회귀분석



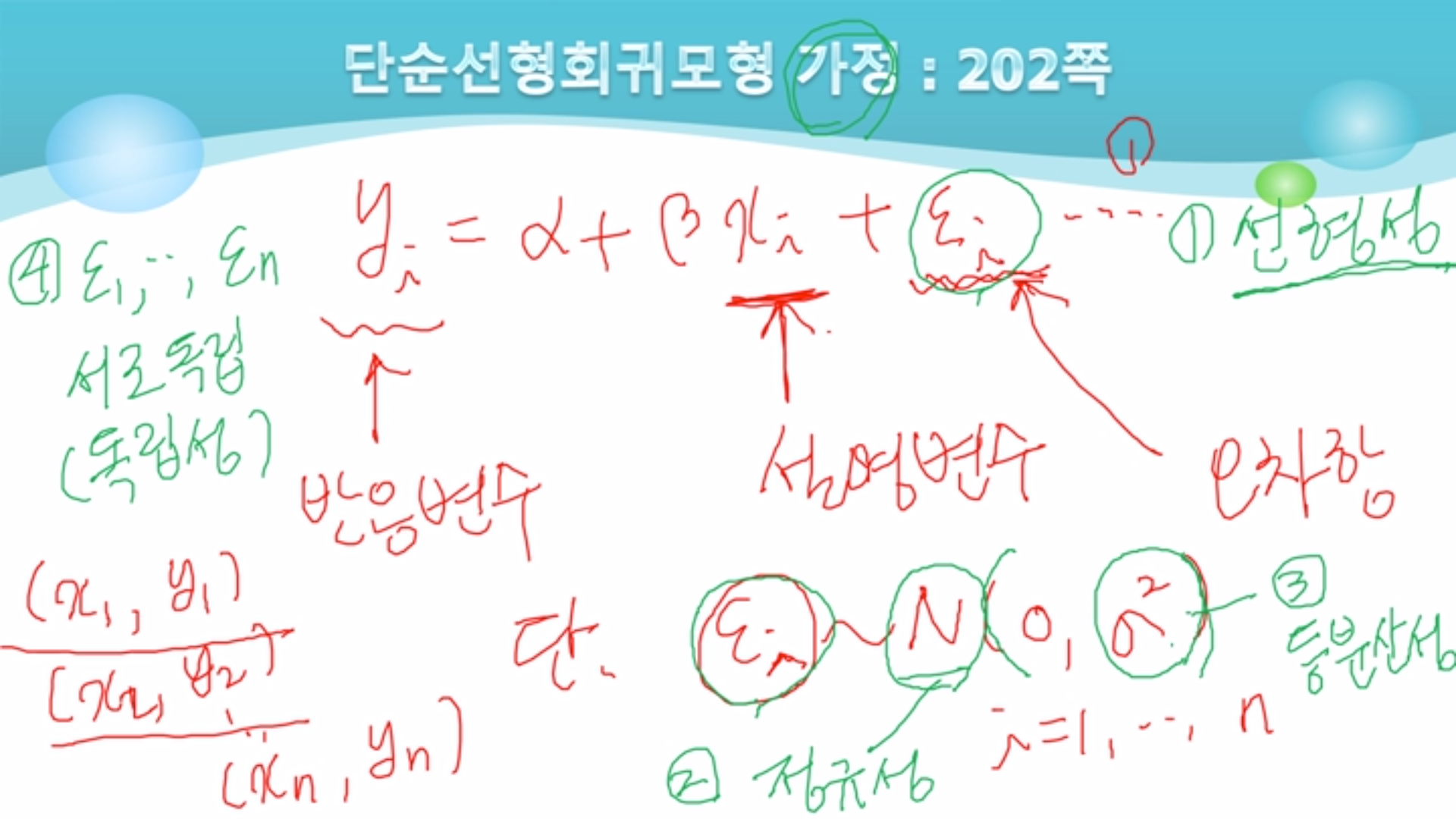
2. 피어슨 상관계수 (직선성, 선형성을 알아보는 상관계수)



- r이 1에 가까울수록 점들이 직선에 가까움

- 기울기가 음수이면 피어슨 상관계수도 음수 ( -1<=r<=1 )

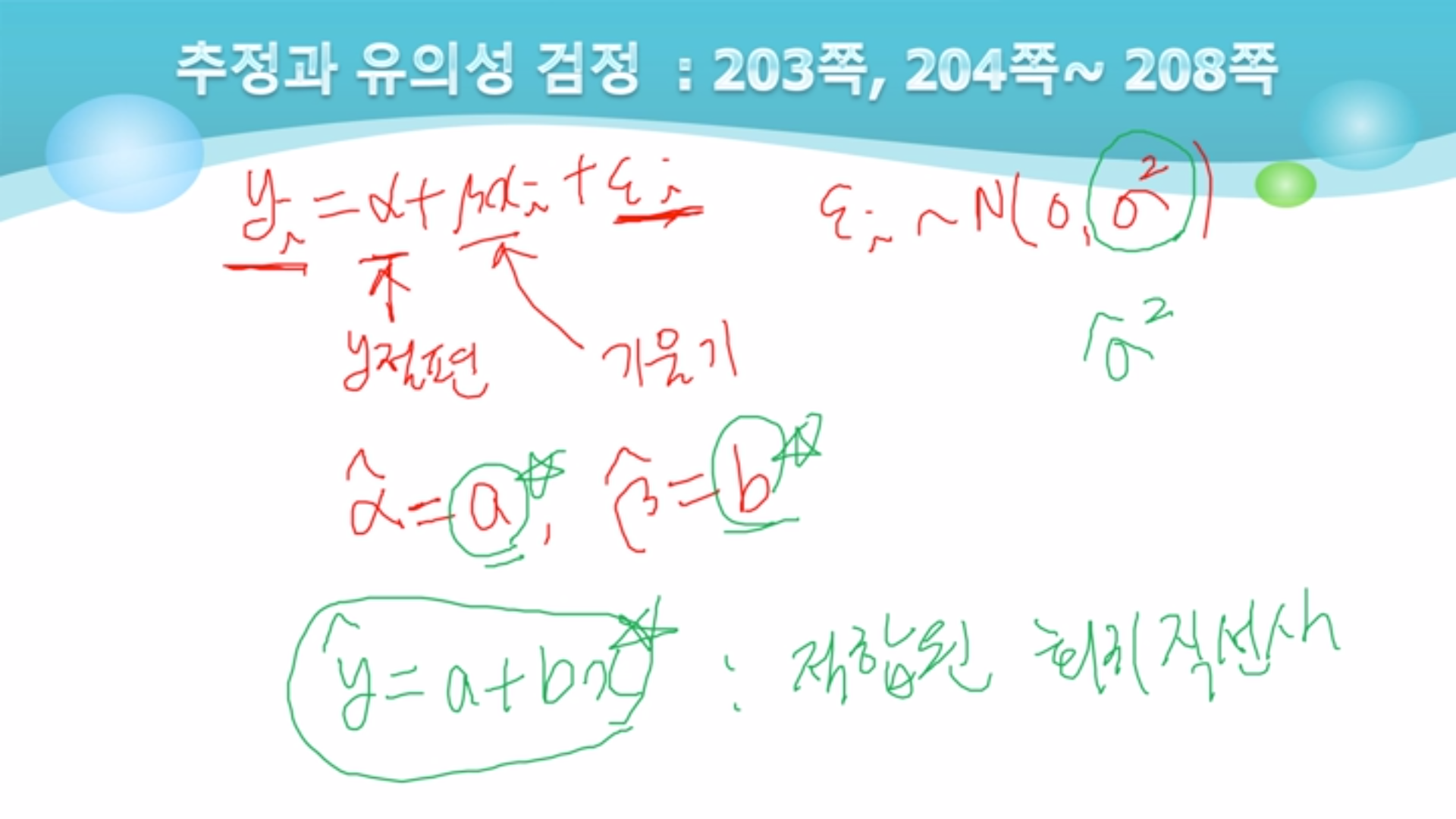
3. 단순선형회귀모형 가정



- ex) Xi가 아버지의 키라고 할 때 뒤의 오차항 때문에 아들들의 키가 다르게 나옴

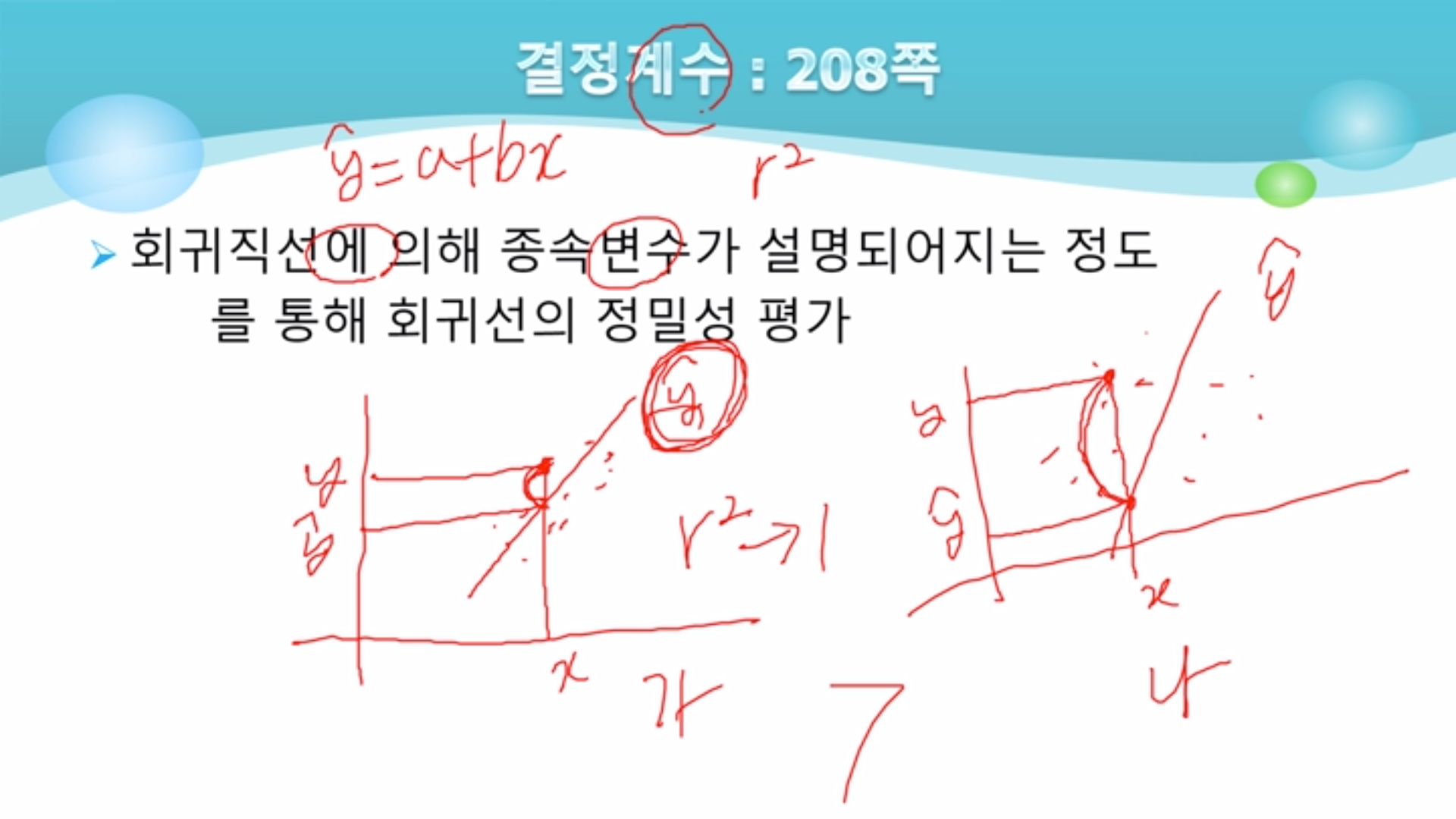
- 4가지 가정이 들어감: 직선 모형을 사용했기 때문에 선형성, 정규분포 가정한 정규성, n개의 오차항들의 동일한 분산사용한 등분산성, n개의 오차항들이 서로 독립이라는 독립성

4. 추정과 유의성 검정



- 적합된 회귀직선식, a, b 값이 유의한 값인지 검증을 하게 됨

5. 결정계수



- 원래의 y값과 추정된 회귀직선식의 y값간의 차이 크기를 바탕으로 더 정밀한 회귀직선식을 찾을 수 있음