오차(Error)와 잔차(Residual), 잔차(Residual) VS 오차(Error)

모집단에서 참값에서 관찰된 값을 뺀 만큼 즉, 실제와 관찰치(값) 사이의 차이를 오차라고 한다. <mark>전차</mark>는 참값이 아닌 <u>관찰</u> <u>치(값)에서 예측값과의 차이</u>를 의미한다.

I.오차 회생(=취(k)위의 ૠ.

- 1. 오차(error)의 정의
- 모집단에서 실제값이 회귀선과 비교해 볼 때 나타나는 차이(정확치와 관측치의 차이)
- 1. 잔차의 정의
- 표본에서 나온 관측값이 회귀선과 비교해볼 때 나타나는 차이
- <u>회귀모형에서 오차항은 측정할 수 없으므로 잔차를 오차항의 관찰값으로 해석하여 오차</u>항에 대한 가정들의 성립 여부를 조사함

단순 선형회귀분석

■ 분산 모델 이용

$$\sum_{i=1}^{n} (x_i - \bar{x})^2 \longrightarrow \sum_{i=1}^{n} (y_i - b_0 + b_1 x_i)^2$$

- 편차(deviation)
 - 자료가 평균을 중심으로 어떻게 분포 → 분산 및 표준편차
 - 편차(개체값-평균): (x_i x̄)
- 잔차(residuals)
 - 회귀분석에서 사용 → 모델의 적합도(회귀직선)
 - $\bullet \ e_i = (y_i \widehat{y}_i)$
- 오차(error)
 - 데이터마이닝 성능평가에서 사용 → 모형의 성능 (실제값 예측)
 - $(y_i \hat{y}_i)$

24

 \mathcal{E}'

义。(明朝): 24

P : 1