# 4주_오차 및 관측값처리

가천대학교 토목환경공학과
박 홍 기

# 유효숫자(Significant Figures)

- 유효숫자(Significant figures)는 수의 정확도에 영향을 주는 숫자이다. 보통 다음의 경우를 제외하고 모든 숫자는 유효숫자이다.
  - 0.00012의 1 앞에 있는 0들처럼 자릿수를 표시하기 위한 0
  - 유효숫자가 아닌 자리의 숫자와 연산하여 영향받은 자리의 숫자
  - 측정 기구의 한계로 정확하지 않은 자리의 숫자

- 101.12는 5개의 유효숫자 (1, 0, 1, 1, 2)
- 0.00012는 2개의 유효숫자 (1, 2)
- 120.00은 5개의 유효숫자 (1, 2, 0, 0, 0)
- 15000의 유효숫자가 네 개라면, 반드시 $1.500 \times 10^4$으로 표기해야 한다.

# 오차의 종류

- 착오(mistake) 또는 과대오차(blunder, gross error)
  - 측량자의 과실, 부주의, 미숙 등이 원인, 오차제거 가능
- 정오차(constant error)
  - 누차로서 오차의 원인과 결과가 명확한 불변의 오차 - 자연오차, 기계오차, 개인오차
  - 계통적 오차(systematic error): 주로 기계오차
- 우연오차(random error; 부정오차)
  - 착오, 정오차, 계통오차를 소거한 후 남는 것.
  - 원인 모름, 통계적 처리

# 오차 발생원인에 의한 분류

**① 기계오차 (Instrumental Error)**

- 장비오차라고도 하며, 관측기계가 불완전하여 발생하는 오차
- 오차를 최소화하기 위해서는 관측하기 전에 기계에 대한 점검과 조정(검교정)을 해야 하는데, 기계를 교정을 할 수 없는 경우에는 보정량을 계산하여 관측값을 보정해야 한다.

**② 자연오차 (Natural Error)**

- 환경오차라고도 하며, 관측을 수행할 때의 여러 가지 자연 환경 조건의 변화에 의해 발생하는 오차를 말한다. 바람, 습도, 중력, 온도 등이 대표적인 원인이 된다.
- 오차를 최소화하기 위해서는 관측과 동시에 온도, 습도 등 외부 환경 조건을 기록하여 조건에 따른 오차 보정량을 이용하여 관측값을 보정한다.

**③ 개인오차 (Personal Error)**

- 관측자의 숙련도나 관측 습관 등에 의해 발생하는 오차
- 예를 들면, 눈금을 읽을 때 최소눈금 사이를 읽는 방법이 관측자마다 다르기 때문에, 또는, 측량기계 수평맞추기에서 관측자마다 숙련도가 다르기 때문에 발생하는 오차가 개인오차라고 할 수 있다.

# 정확도와 정밀도

- 정밀도(precision):측정값들의 상대적인 편차가 적은 것
  - 측정 과정과 밀접한 관계
  - 측정 방법과 장비에 크게 영향을 받음
  - 우연오차와 매우 밀접
- 정확도(Accuracy):측정값이 참 값에 얼마나 일치하는가를 표시하는 척도
  - 측정결과와 관련
  - 정오차, 착오 제거 관련

# 오차의 일반 법칙

- 확률곡선이 가지는 “오차의 일반 법칙”

① +오차와 –오차는 같은 빈도로 발생
② 매우 큰 오차는 거의 발생하지 않음
③ 작은 오차는 큰 오차보다 더 자주 발생

# 분산 (Variance)

① 자료집합 내에서의 개별 관측의 정밀도를 나타내는 척도로서, 자료의 흩어진 정도를 숫자로 표현

② 잔차의 단순합과 잔차의 평균은, 오차의 일반적인 특성에 따라 0이 된다. 따라서, 개별관측의 흩어진 정도를 표현하기 위해서는 잔차의 제곱의 합을 관측수로 나눈 값(잔차제곱의 평균)을 사용하며, 이것을 분산이라고 한다.

③ 모집단의 흩어진 정도를 표현한 모집단 분산과 표본집단의 흩어진 정도를 표현한 표본집단 분산이 있음

④ 모집단 분산 (모분산) :

$$\sigma^2 = \frac{\sum (y_i - \mu)^2}{n} = \frac{\sum \varepsilon^2}{n}$$

⑤ 표본집단 분산 (표본분산) :

$$S^2 = \frac{\sum (y_i - \bar{y})^2}{n-1} = \frac{\sum \upsilon^2}{n-1}$$

# 표준편차 (Standard Deviation)

① 분산의 경우 단위가 관측값의 제곱으로 나타나기 때문에, 관측값의 정밀도를 관측값의 단위와 일치하는 값으로 나타내기 위하여 분산에 제곱근을 취하여 이것을 "표준편차"라고 하였다.

② 모집단 분산의 제곱근을 특별히 표준오차(Standard Error, σ )라고 부르고, 표본집단 분산의 제곱근을 표준편차(Standard Deviation, *S* )라고 구분하여 부르기도 한다.

③ 표준오차 :

$$\pm\sigma = \pm\sqrt{\sigma^2} = \pm\sqrt{\frac{\Sigma(y_i-\mu)^2}{n}} = \pm\sqrt{\frac{\Sigma\varepsilon^2}{n}}$$

④ 표준편차 :

$$\pm S = \pm\sqrt{S^2} = \pm\sqrt{\frac{\Sigma(y_i-\bar{y})^2}{n-1}} = \pm\sqrt{\frac{\Sigma\upsilon^2}{n-1}}$$

⑤ 공간위치결정에서는 관측에서 모집단을 사용하는 경우는 거의 없으므로, 일반적으로 일컫는 '분산'과 '편차'는 '표본분산'과 '표준편차'를 말한다.

⑥ RMSE(Root Mean Square Error): 영어 해석 그대로, 잔차 제곱의 평균에 제곱근을 취한 것으로서, 표준편차의 정의와 동일하다. 그러나, 위의 분산과 표준편차는 미지수 1개에 대한 반복관측에 의한 개별관측의 정밀도인데 반하여, RMSE는 미지수 2개 이상이 포함된 관측의 정밀도를 나타낼 때 사용하는 용어이다.

# 최확값

- 최확값
  - 경중률이 일정할 때

$$L_0 = \frac{L_1 + L_2 + \cdots + L_n}{n}$$

  - 경중률을 고려할 때

$$L_0 = \frac{P_1 L_1 + P_2 L_2 + \cdots + P_n L_n}{P_1 + P_2 + \cdots + P_n}$$

여기서, $L_1, L_2, L_3, \cdots L_n$ : 관측값

$P_1, P_2, P_3, \cdots P_n$ : 경중률

# 평균제곱근오차

- 최확값에 대한 평균제곱근오차

  - 경중률이 일정할 때 - 경중률을 고려할 때

$$M_0 = \pm\sqrt{\frac{[vv]}{n(n-1)}} \qquad M_0 = \pm\sqrt{\frac{[pvv]}{[p](n-1)}}$$

- 1관측값에 대한 평균제곱근오차

  - 경중률이 일정할 때 - 경중률을 고려할 때

$$M_0 = \pm\sqrt{\frac{[vv]}{(n-1)}} \qquad M_0 = \pm\sqrt{\frac{[pvv]}{(n-1)}}$$

여기서, $M_0$ : 평균제곱근 오차

$n$ : 관측회수

$v$ : 잔차(관측값-최확값)

# 확률오차와 정도

- 확률오차 : 분포함수의 면적의 50%

$$v_0 = \pm 0.6745(M_0)$$

- 정도 :
  - 정도는 관측이 수행된 후 취득된 관측값이 그 신뢰도가 얼마나 되는지 평가해 보는 수단
  - 최확값과 표준 오차와의 비로 표현되며, 측량학의 경우, 분자가 1로 표현됨
  - Coefficient of Variation (COV) 정도 = $\dfrac{v_0}{L_0}$ , 또는 $\dfrac{M_0}{L_0}$
  - 상대오차(relative error): 측정값의 크기와 오차의 비
  - 거리측량의 정밀도 표시 시 사용
    - 10m 관측값에 +1 mm 잔차는 상대오차가 1/10,000 임

# 경중률

- 경중률(weight): 무게, 중량
  - 한 측정값과 이와 연관된 다른 측정값에 대한 상대적인 신뢰성을 표현하는 척도
    - 한번 관측한 결과보다 100번 관측한 것이 신뢰
    - 10m 거리를 측정한 것이 10km를 측정한 것보다 신뢰
    - 중량은 분산의 크기에 반비례(분산이 크면 오차가 크므로)

$$\text{weight} \propto N \quad (\text{관측회수}) \qquad P_1 : P_2 : P_3 = N_1 : N_2 : N_3$$

$$\text{weight} \propto \frac{1}{S} \; (\text{노선거리}) \qquad P_1 : P_2 : P_3 = \frac{1}{S_1} : \frac{1}{S_2} : \frac{1}{S_3}$$

$$\text{weight} \propto \frac{1}{m^2} \; (m: \text{평균제곱근 오차}) \qquad P_1 : P_2 : P_3 = \frac{1}{m_1^2} : \frac{1}{m_2^2} : \frac{1}{m_3^2}$$

# 최확값과 평균제곱근오차

- 최확값의 결정

  경중률이 같을 때

  $$L_0 = \frac{L_1 + L_2 + \cdots + L_n}{n}$$

  경중률이 다를 때

  $$L_0 = \frac{P_1L_1 + P_2L_2 + \cdots + P_nL_n}{P_1 + P_2 + \cdots + P_n}$$

- 평균제곱근오차의 결정(RMSE): 분포함수의 0.6826 즉 68%
  - 최확값에 대한 평균제곱근오차

    경중률이 같을 때

    $$M_0 = \pm\sqrt{\frac{[vv]}{n(n-1)}}$$

    경중률이 다를 때

    $$M_0 = \pm\sqrt{\frac{[Pvv]}{[P](n-1)}}$$

  - 1관측값에 대한 평균제곱근오차

    경중률이 같을 때

    $$M_0 = \pm\sqrt{\frac{[vv]}{n-1}}$$

    경중률이 다를 때

    $$M_0 = \pm\sqrt{\frac{[Pvv]}{n-1}}$$

# 오차 전파(우연오차 또는 부정오차)

- 관측값들을 사용하여 어떤 결과값을 얻게 될 때, 결과값의 평균제곱오차는 각 평균제곱근 오차의 제곱근이 된다.

- 오차전파식

$$M^2 = \left(\frac{\partial f}{\partial x}\right)^2 m_x{}^2 + \left(\frac{\partial f}{\partial y}\right)^2 m_y{}^2 + \cdots$$

$$\text{또는} \quad M = \pm\sqrt{\left(\frac{\partial f}{\partial x}\right)^2 m_x{}^2 + \left(\frac{\partial f}{\partial y}\right)^2 m_y{}^2 + \cdots}$$

# 측량학에서 필요한 기초수학

- 각도 : 60진법, 라디안, 호도법
- 삼각함수
- 삼각형 : sine 법칙, 피타고라스 정리
- 미분
- 적분
- 오차전파
- 면적 공식, 체적 공식

# The End