

단위가 엄밀히 정의되면 다음으로 측정값을 결정하여야 한다. 이 과정을 측정이라 한다. 그러나 어떤 측정에서든지 정확히 측정한다는 것은 불가능하다. 그러므로 측정값의 착오가 발생하고 측정값과 참값의 차이를 오차로 정의할 수 있다. 한편 참값은 정확히 알 수 없는 양이므로 오차도 정확히는 알 수 없고 단지 추측할 수 있는 수치일 뿐이다. 일반적으로 오차는 부당오차, 계통오차, 우연오차, 확률오차 등으로 구분할 수 있다. 부당오차는 계기조작 상 분명한 실수를 범하여서 측정값이 신빙성이 없는 경우에 생기는 오차이다. 길이를 재는데 한쪽 원점을 맞추지 않았다든지, 저항측정에서 원점을 확인하지 않는 경우 등이다. 이때에는 그 원인이 명백하므로 얻은 데이터를 무시하는 것이 보통이다. 이들을 포함시킨다면 다른 측정값들의 신뢰도를 떨어뜨리기 때문이다.

계통오차는 측정 계기의 불비한 점에 기인되는 오차로서 그 크기와 부호를 추정할 수 있고, 부정할 수 있는 오차이다. 전압을 여러 번 측정하여 평균값을 얻었는데 확인 결과 사용한 전압계의 눈금이 원점에서 벗어나 있다든지, 어떤 자로 길이를 재었는데 온도에 따른 길이의 변화를 고려하지 않는 경우 등이다. 이때에는 계통오차를 추정할 수 있으므로 별도로 추정하여 결과적으로 얻은 측정값에 직접 반영할 수 있다.

우연오차는 반복 측정할 때마다 상이한 결과를 얻게 되는 측정값들의 변동에 기인한다. 우연오차를 줄이는 문제가 실험결과를 향상시키는 제일 중요한 요소이다. 측정값의 정밀도는 이 우연오차를 어떻게 처리하고 분석하는가에 달려 있기 때문이다. 우연오차를 줄이는 방법은 주로 더 정밀한 측정계기를 사용한다든지 또는 이와 더불어 여러 번 반복 측정하는 것이다.

확률오차는 측정값을 얻을 때 추정되는 오차의 크기를 나타낸다. 예를 들면, 어떤 측정값이 $x = \bar{x} \pm \sigma$ 로 나왔다면 이는 x 의 오차가 σ 라는 의미가 아니다. 올바른 뜻은 결과가 틀리더라도 $x \pm \sigma$ 이상 벗어날 확률은 작다는 것을 의미한다. 따라서 σ 는 측정값이 주어진 값 $\bar{x}$ 에서 어느 정도 벗어날 수 있는지 확률적 척도를 제시해 준다. 정상분포를 이루는 오차에서는 평균값의 표준편차를 σ_m 으로 나타내면 참값이 $\bar{x} + \sigma_m$ 와 $\bar{x} - \sigma_m$ 사이에 있을 확률이 약 68%임을 의미한다. 그러므로 $\sigma = 0.67\sigma_m$ 을 택하여 $\bar{x} + \sigma$ 와 $\bar{x} - \sigma$ 사이에 있을 확률이 50%가 되게 하는 σ 를 확률오차라고 부른다.

1. 측정값의 유효숫자

측정값은 숫자로 표시하여야 하지만 그 의미는 수학에서의 표기법과 다르다. 모든 측정값은 근삿값이므로 무의미한 자릿수들을 나열할 필요가 없다. 그래서 효력이 있는 숫자, 즉 유효숫자만을 표시하여야 하는데 일반적으로 다음과 같이 정해진다.

- (1) 0이 아닌 맨 왼쪽의 숫자가 최상 유효숫자이다.
- (2) 소수점이 없을 경우에는 0이 아닌 맨 오른쪽의 숫자가 최하 유효숫자이다.
- (3) 소수점이 있을 경우에는 맨 오른쪽의 숫자가 0이더라도 이 수가 최하 유효숫자이다.
- (4) 최상 유효숫자간의 모든 숫자가 유효숫자이다.

보기로 아래에 열거한 숫자들은 오직 고딕체로 나타낸 부분만이 유효숫자이다.

914.38	23000	0.0094
4.000	700.	3.140×10^4

측정값들을 가감승제할 때에는 불필요한 계산시간의 낭비를 줄이기 위하여 다음의 요령으로 결과를 얻는다.

• 덧셈과 뺄셈 : 보기로 4.5와 0.3352의 합을 계산하면

$$\begin{array}{r}
 4.5 \\
 +) 0.3352 \\
 \hline
 4.8352
 \end{array}
 \rightarrow
 \begin{array}{r}
 4.5 \\
 +) 0.3352 \\
 \hline
 4.83
 \end{array}
 =
 \begin{array}{r}
 4.5 \\
 +) 0.33 \\
 \hline
 4.8
 \end{array}$$

반올림

숫자 4.5는 소수점 이하 두 자리에서는 유효숫자가 없으므로 0.3352도 그곳에서 자른 다음 두 숫자를 더하여 반올림하면 된다.

- 곱셈과 나눗셈 : 보기로 4.5와 0.3352의 곱을 계산하면

$$4.5 \times 0.3352 = 1.5080 \Rightarrow 4.5 \times 0.335 = 1.5075 \xrightarrow{\text{반올림}} 1.5$$

숫자 4.5는 유효숫자가 두 자리이므로 0.3352는 0.335에서 세 자리만 택하여 곱한 후 결과는 두 자리가 되도록 반올림하면 된다.

2. 표준오차

측정값들은 우연오차 때문에 매번 측정할 때마다 다른 값을 얻게 되고 어떤 분포를 이룬다. 이러한 측정값들의 분포특성을 기술하기 위하여 이들을 대표할 수 있는 수치와 분포된 정도를 나타내는 척도가 필요하다.

측정자료를 대표할 수 있는 수치로서는 최빈값, 중앙값 및 평균값 등을 사용한다. 최빈값은 측정자료들을 나열했을 때 빈도가 가장 많은 측정값이고, 중앙값은 이보다 작은 자료와 많은 자료가 똑같은 측정값의 분포에서 중앙에 위치한 측정값이다.

평균값은 측정값의 산술평균이다. 측정값들이 $x_1, x_2, x_3, \dots, x_N$ 이라 하여 모두 N 개의 자료를 얻었으면 평균값은 다음과 같이 계산한다.

$$\text{평균값} \quad \bar{x} = \frac{1}{N} \sum_{i=1}^N x_i$$

경우에 따라서는 구간을 설정하여 측정하기 때문에 $x_1, x_2, x_3, \dots, x_N$ 에 대한 빈도가 각각 $f_1, f_2, f_3, \dots, f_N$ 인 형태로 자료를 얻을 수 있다. 이 때문에 빈도를 가중치로 택하여 아래와 같이 평균값을 계산한다.

$$\text{평균값} \quad \bar{x} = \frac{\sum_{i=1}^N f_i x_i}{\sum_{i=1}^N f_i}$$

측정자료들이 분포된 정도를 나타내기 위하여 편차를 $d_i = x_i - \bar{x}$ 를 사용하는 것이 편리하다. 이들을 평균하면 0이 되므로 그 절대값의 평균을 평균편차라 한다. 그러나 통계적으로는 다음과 같이 정의된 표준편차 σ 가 더 중요한 의미를 갖는다.

$$\sigma^2 = \frac{1}{N-1} \sum_{i=0}^N (x_i - \bar{x})^2$$

σ^2 자체는 분산(variance)이라고 불리운다.

한 물리량을 1회에 N 번씩 여러 차례 되풀이하여 측정하면 매회 얻어지는 측정값에 대한 평균값과 표준편차는 일반적으로 달라진다. 평균값 $\bar{x}$ 의 표준오차를 σ_m 이라 하면 σ, σ_m, N 사이에는 다음과 같은 관계가 성립한다.

$$\sigma_m = \frac{\sigma}{\sqrt{N}} = \left\{ \frac{1}{N(N-1)} \sum_{i=1}^N (x_i - \bar{x})^2 \right\}^{\frac{1}{2}}$$

측정자료들로부터 보고할 측정값은 다음과 같다.

$$x = \bar{x} \pm \sigma_m$$

오차를 표기하는 방법은 흔히 절대오차, 상대오차 및 퍼센트 오차의 세 가지가 있다.

절대오차 σ

상대오차 $\frac{\sigma}{|\bar{x}|}$

퍼센트오차 $\frac{\sigma}{|\bar{x}|} \times 100$

상대오차는 측정값의 정밀도를 나타내므로 편리하고, 퍼센트오차는 이와 관련하여 많이 사용된다.

3. 오차의 전파

직육면체를 가로, 세로 및 높이를 각각 열 번씩 측정하여 그 부피를 구하는 실험을 생각해 보자. 이들 데이터를 이용하면 1,000개의 부피에 관한 데이터를 얻을 수 있고 이로부터 평균값과 표준편차를 구할 수 있다. 그러나 보다 합리적인 방법은 가로, 세로,