

## 0.1 Teoria błędów

Wskutek niedoskonałości przyrządów jak również niedoskonałości naszych zmysłów - wszystkie pomiary są dokonywane z określonym stopniem dokładności. Często nie otrzymujemy prawidłowych wartości mierzonej wielkości lecz wartości do nich zbliżone.

**Wszystkie wartości wielkości fizycznych są obarczone pewnym błędem**

**Pomiarem** nazywamy czynności związane z ustaleniem wartości liczbowej miary danej wielkości fizycznej. Istotą pomiaru fizycznego jest porównanie wielkości mierzonej z ustalonym wzorcem jednostką.

### Narzędzia pomiarowe

- wzorce;
- przyrządy pomiarowe.

### Sposób pomiaru

- wielkości proste - pomiar bezpośredni;
- wielkości złożone - pomiar pośredni.

W pracowni spotykamy się z następującymi po sobie procesami:

1. Pomiar
  - a ustawienie przyrządu;
  - b obserwacja zjawiska;
  - c odczyt mierzonej wielkości.
2. obliczenia - krytyczna analiza prawidłowości i stopnia ich pewności.

#### 0.1.1 Błędy popełniane podczas pomiarów

1. błędy przypadkowe - są to błędy niepowtarzające się. Mogą one przyjmować wartość dodatnią lub ujemną. Spowodowane są przez różne niekontrolowane przez eksperymentatora czynniki, działające w chwili pomiaru (np. zmiany napięcia w sieci elektrycznej, do której podłączone jest urządzenie pomiarowe, ograniczona dokładność obserwacji eksperymentatora).

Statystyka umożliwia nie tylko oszacowanie wartości błędu przypadkowego, ale również określenie jego prawdopodobieństwa.
2. błędy grube - są to duże błędy spowodowane nieuwagą lub niestarannością eksperymentatora. Wynik zawierający błąd gruby bardzo różni się od pozostałych wyników. Na tym też opierają się niektóre kryteria ich eliminacji.
3. błędy systematyczne - są to błędy powtarzające się, w większości tego samego znaku. Powodują je czynniki działające w jednakowy sposób w czasie wielokrotnego powtarzania tego samego pomiaru. Przyczyną tych błędów może być: niedokładność przyrządów, niedokładność metod pomiarowych oraz wzorców stosowanych do ostatecznych obliczeń.

W praktyce analitycznej występują zarówno błędy systematyczne, jak i przypadkowe, natomiast błędy grube można łatwo eliminować. Najważniejszym zadaniem przy szacowaniu błędu wyniku końcowego jest znalezienie błędu dominującego (przypadkowego, systematycznego) i zminimalizowanie jego wpływu przez dobór odpowiednich warunków pomiarowych.

## 0.2 Miary popełnianych błędów

Błąd pomiaru stanowi różnicę pomiędzy zmierzoną wartością  $x_i$  a wartością rzeczywistą (prawdziwą)  $x_0$

$$\delta x = x_i - x_0$$

Równanie to jest nierozwiązywalne, zawiera ono bowiem dwie niewiadome tj. wartość rzeczywistą  $x_0$  i błąd pomiaru  $\delta x$ . W celu rozwiązania oszacowuje się bądź wartość rzeczywistą ( $x_0 \approx \mu \approx \bar{x}$ ) bądź  $\delta x$ . Na przykład za  $\delta x$  można przyjąć czułość przyrządu pomiarowego. Stosuje się także:

1. błąd bezwzględny wielkości mierzonej  $|x - x_0| = \delta$ ;
2. błąd względny  $\delta/x_0$ ;
3. błąd procentowy  $(\delta/x_0) \cdot 100\%$ .

## 0.3 Rzetelność

(miarodajność) oznacza stopień zgodności z wartością prawdziwą (rzeczywistą).

Dla materiału biologicznego nie znamy i nie możemy znać wartości prawdziwej (rzeczywistej) zakładamy tutaj, że są to średnie wartości  $\bar{x}$  pomiarów uzyskane przez różne laboratoria przy zastosowaniu tej samej metody.

W analityce chemicznej do wyznaczenia wartości rzeczywistej służą substancje wzorcowe, w których oznaczoną na podstawie wzoru chemicznego zawartość poszukiwanego składnika przyjmuje się jako wartość prawdziwą. Podstawą oceny rzetelności może być odchylenie procentowe.

**Rzetelność pojedynczej próby** określa błąd  $\delta x$  tj.  $\delta x = x_i - x_0$ . Na wartość tej różnicy składa się zespół błędów (błąd przypadkowy -  $\delta x_i$ , błąd systematyczny -  $\delta_{wyw}$ , błąd grubo -  $\delta_g$ ). Zatem

$$x_i = x_0 + \delta x_i + \delta_{sys} + \delta_g.$$

Błąd przypadkowy pojedynczego wyniku można oszacować korzystając z wiedzy o metodzie przedziałów ufności dla parametru położenia. Załóżmy, że obserwacje mają rozkład normalny, wtedy

$$\delta x_i = \pm z_{\alpha/2} S_n.$$

Jest to tak naprawdę połowa przedziału ufności dla parametru położenia obliczonego na podstawie pojedynczej obserwacji. ( $z_{\alpha/2}$  to odpowiednia wartość kwantyla z rozkładu normalnego. **Ważne** wzoru tego używamy, gdy do oszacowania dysponujemy dużą próbą ( $n \geq 30$ ). W przypadku małych prób użyjemy oczywiście rozkładu  $t$  - *Studenta*.

**Rzetelność wyniku końcowego** analizy to różnica pomiędzy wartością  $\bar{x}$  i wartością prawdziwą  $\mu$ .

$$\delta \bar{x} = \bar{x} - \mu.$$

Błąd przypadkowy w tym przypadku to połowa przedziału ufności dla parametru położenia:

$$\delta \bar{x} = \pm z_{\alpha/2} \frac{S_n}{\sqrt{n}}$$

**Metoda odzysku** Metoda odzysku polega na oznaczeniu nieznanego stężenia w badanej próbce oraz równolegle w tej samej próbce z dodatkiem określonej ilości wzorca (wzbogaconej). Uzyskana różnica pomiędzy stężeniami obu próbek: wzbogaconej i niewzbogaconej jest miarą odzysku rzetelności.

$$\% \text{odzysku} = \frac{a - b}{c} 100$$

gdzie  $a$  - stężenie próbki wzbogaconej,  $b$  - stężenie próbki niewzbogaconej  $c$  - ilość dodanego wzorca.

## 0.4 Dokładność

(precyzja, rozrzut) to zgodność wyników analizy powtarzanej wielokrotnie na tym samym materiale w tych samych warunkach (odczynniki, aparat mierzący)

## 0.5 Powtarzalność

to uzyskiwanie tych samych wartości na tym samym materiale w różnym czasie, przez różnych analityków, różnymi odczynnikami.

### 0.5.1 Błędy pomiarowe

**Błąd grubzy** - kryterium eliminacji

**Zadanie 1** *Zmierzono ekspresję genu BRCA1 u 10 pacjentek. Wyniki to*

$$X = 4, 15, 9, 16, 6, 5, 16, 4, 11, 8, 35$$

*Pytania:*

- *Czy któraś z obserwacji jest obarczona błędem grubym?*
- *Ile obserwacji jest obarczonych błędem?*

### 0.5.2 Test Grubbsa

Do testowania hipotezy

$$H_0: \text{brak obserwacji odstających}$$

przy dwustronnej alternatywie możemy wykorzystać następującą statystykę testową:

$$T(X) = \frac{|X_i - \bar{X}|}{S_X}$$

Wartość krytyczną wyznacza się ze wzoru:

$$c_\alpha = \frac{N-1}{\sqrt{N}} \sqrt{\frac{t_{\alpha/(2N), N-2}^2}{N-2 + t_{\alpha/(2N), N-2}^2}}$$

gdzie  $t_{\alpha/(2N), N-2}$  to kwantyl rzędu  $1 - \alpha/(2N)$  rozkładu  $t$ -Studenta o  $N - 2$  stopniach swobody. Dla jednostronnej alternatywy wykorzystuje się kwantyl rzędu  $t_{\alpha/N, N-2}$ .

#### Zastosowanie testu Grubbsa

1. Obserwacje

$$4, 15, 9, 16, 6, 5, 16, 4, 11, 8, 35$$

2. liczymy średnią i odchylenie standardowe

$$\bar{X} = 11.72, S_X = 8.99$$

3. obliczamy kolejne wartości  $|X_i - \bar{X}|/S_X$

$$1.17, 0.49, 0.41, 0.64, 0.86, 1.02, 0.64, 1.17, 0.11, 0.56, 3.53$$



Wygodny sposób postępowania w powyżej wymienionych przypadkach przedstawił Grubbs. Korzystamy w tym teście z odpowiednich wartości tabelarycznych  $S_n^2/S^2$ ,  $S_1^2/S^2$ ,  $S_{n-1,n}^2/S^2$  i  $S_{1,2}^2/S^2$ .

1. do sprawdzenia największej wartości z próby o liczności  $n$  posługujemy się wzorem

$$S_n^2/S^2 = \frac{\sum_{i=1}^{n-1} (x_i - \bar{x}_n)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

gdzie  $x_1 \leq x_2 \leq \dots \leq x_n$ , oraz

$$\bar{x}_n = \frac{1}{n-1} \sum_{i=1}^{n-1} x_i, \quad \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

2. dla sprawdzenia wartości najniższej stosujemy

$$S_1^2/S^2 = \frac{\sum_{i=2}^n (x_i - \bar{x}_1)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

gdzie

$$\bar{x}_1 = \frac{1}{n-1} \sum_{i=1}^{n-1} x_i$$

3. dla sprawdzenia dwóch największych wartości stosujemy

$$S_{n-1,n}^2/S^2 = \frac{\sum_{i=1}^{n-2} (x_i - \bar{x}_{n-1,n})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

gdzie  $\bar{x}_{n-1,n} = \frac{1}{n-2} \sum_{i=1}^{n-2} x_i$

4. dla sprawdzenia dwóch najmniejszych postępujemy podobnie jak w punkcie 3.

**Zadanie 4** Sprawdzić czy wśród otrzymanych wyników: 13, 17, 42, 43, 46, 47, 49, 49, 54, 56, 67 wartości 13, 17 należą do zbioru.

**Zadanie 5** Sprawdzić czy wśród otrzymanych wyników: 13, 17, 42, 43, 46, 47, 49, 49, 54, 56, 67, 85, 90 należy odrzucić wartości 13 i 17 oraz 85 i 90.