

K

VIỆN NĂNG LƯỢNG NGUYÊN TỬ VIỆT NAM  
VIỆN CÔNG NGHỆ XẠ HIỂM

---

## KHOA HỌC VÀ CÔNG NGHỆ VỀ NHIÊN VẬT LIỆU HẠT NHÂN

Huỳnh Văn Trung (Chủ biên), Thái Bá Cầu  
Đỗ Ngọc Liên, Lê Bá Thuận, Đỗ Quý Sơn, Cao Hùng Thái,  
Nguyễn Bá Tiến, Cao Đình Thanh, Nguyễn Lanh

## XỬ LÝ THỐNG KÊ CÁC SỐ LIỆU TRONG HÓA PHÂN TÍCH

HUỲNH VĂN TRUNG

HÀ NỘI - 2005

6100-4

18/9/06

## Chương I

### CÁC THUẬT NGỮ VÀ ĐỊNH NGHĨA

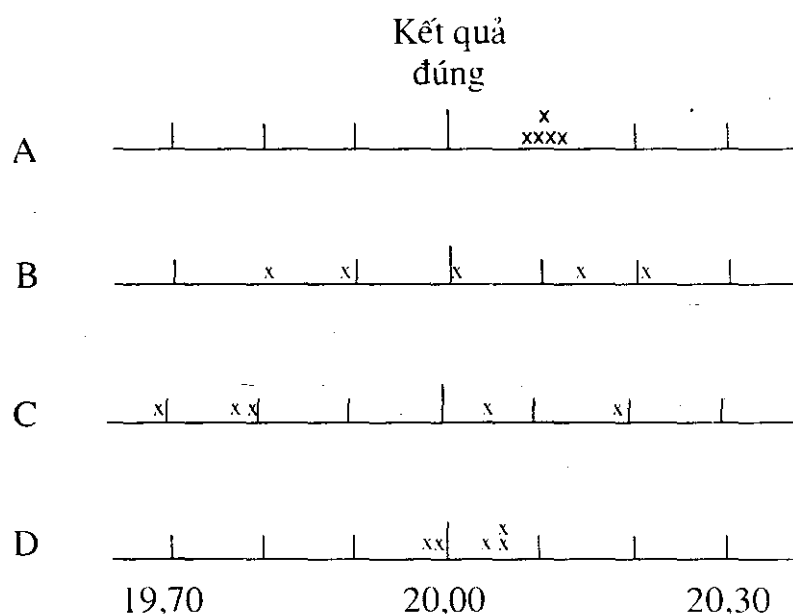
**Mê du răng (Mesurand )** Theo định nghĩa của ISO (Cơ quan Tiêu chuẩn hoá quốc tế) là một đại lượng cụ thể nào đó cần đo hoặc cần phân tích. trong danh pháp Hoá học của IUPAC còn gọi đó là chất cần phân tích (Analyte) chúng tôi mạnh dạn Việt hoá thuật ngữ này theo cách đọc của tiếng Pháp Mê du răng (Mésurand).

**Phép đo:** Là tập hợp một số thao tác nhằm mục đích thu được giá trị của một đại lượng nhất định. Bản chất của phép đo chính là việc so sánh đại lượng cần đo với một đại lượng cùng loại được chọn làm đơn vị. Ví dụ đo độ dài theo đơn vị mét; đo khối lượng theo đơn vị Kg, đo nồng độ theo đơn vị phần trăm (%) hay phần triệu (ppm), đo dung tích theo mililit ...Ta chia phép đo thành 2 loại: phép đo trực tiếp và phép đo gián tiếp. Phép đo trực tiếp là phép đo trong đó giá trị của đại lượng đo được đọc ngay trên phương tiện đo, còn phép đo gián tiếp là phép đo trong đó giá trị của đại lượng đo được tính toán qua mối liên hệ đã biết giữa nó và các đại lượng liên quan có giá trị xác định được bằng phép đo trực tiếp. Độ chính xác của phép đo gián tiếp phụ thuộc vào độ chính xác phép đo trực tiếp có liên quan.

**Độ đúng (Accuracy) và độ chính xác (Precision ) của phép đo:** Đây là hai khái niệm rất quan trọng trong lý thuyết sai số và cũng thường hay bị lẫn lộn với nhau. Cần phải làm rõ và phân biệt hai khái niệm này. Để dễ hình dung ta mô tả hai đặc trưng này bằng ví dụ cụ thể biểu diễn trên Hình 1. Đây là kết quả chuẩn độ thể tích của 4 phân tích viên A, B, C, D.. Mỗi phân tích viên cho 5 kết quả chuẩn độ. Giá trị chuẩn đã biết là 20.00 ml. Các giá trị chuẩn độ được biểu diễn trên thang tuyến tính một chiều (Hình 1). Từ Hình 1 ta có các nhận xét sau:

- Các giá trị trung bình của 2 phân tích viên B và D rất gần với giá trị thật (20,00 ml). Ta nói chúng có **độ đúng tốt**.
- Các giá trị trung bình của hai phân tích viên A và C hoặc cao hơn hoặc thấp hơn giá trị thật. Ta nói chúng có **độ đúng kém**.
- Năm giá trị chuẩn độ của phân tích viên A và 5 giá trị của phân tích viên D rất chụm với nhau. Ta nói chúng có **độ lặp lại hay độ chính xác tốt**.
- Năm giá trị của phân tích viên B và 5 giá trị của C khá phân tán. Ta nói chúng có **độ lặp lại hay độ chính xác kém**.

**Độ đúng** nói chung là một đặc trưng rất quan trọng nếu như không muốn nói là quan trọng nhất của phép phân tích định lượng. Đây là một đại lượng biểu thị độ phù hợp giữa các kết quả đo và giá trị thực (hoặc được chấp nhận là giá trị thực) của Mê du răng. Độ đúng miêu tả tính đúng của kết quả thực nghiệm. Nói chính xác thì chỉ có loại phép đo liên quan đến các đối tượng đếm là hoàn toàn đúng. Tất cả các phép đo khác đều mắc sai số và chỉ cho ta một sự gần đúng.



Hình 1. Các kết quả chuẩn độ thể tích của 4 phân tích viên A, B, C, D.

Độ đúng là một khái niệm tương đối, bởi vì một phương pháp được coi là đúng hay không đúng phụ thuộc rất nhiều vào yêu cầu của nhà khoa học và khó khăn của vấn đề phân tích. Ví dụ, phương pháp phân tích cho ta kết quả nằm trong giới hạn  $\pm 10\%$  hoặc một phần tỷ (ppb) của lượng đúng thủy ngân trong một mô tế bào của cá chứa 10 ppb kim loại thường được coi là có độ đúng hợp lý. Ngược lại, một phương pháp cho kết quả nằm trong giới hạn  $\pm 10\%$  lượng đúng của thủy ngân trong một mẫu vàng chứa 20% kim loại được coi là phương pháp không đúng và không chấp nhận được.

Độ đúng thường được biểu thị theo sai số tương đối hoặc sai số tuyệt đối. Sai số tuyệt đối  $E_a$  của giá trị trung bình  $\bar{x}$  của một dãy nhỏ các phép phân tích lặp lại được cho bởi hệ thức:

$$E_a = \bar{x} - x_i \quad (1)$$

$x_i$  là giá trị được chấp nhận của đại lượng đo. Thường người ta hay biểu diễn độ đúng theo sai số tương đối.

$$\text{Sai số tương đối} = \frac{\bar{x} - x_i}{x_i} \cdot 100\% \quad (2)$$

Thông thường sai số tương đối được biểu diễn theo % như ở trên, trong các trường hợp khác thương số được nhân với 1.000 để tính sai số theo phần nghìn (ppt). Cần chú ý rằng cả sai số tương đối lẫn sai số tuyệt đối đều có dấu, dấu dương cho biết kết quả đó lớn hơn giá trị thực, còn dấu âm thì ngược lại.

Chúng ta sẽ còn đề cập đến hai loại sai số: sai số ngẫu nhiên hoặc sai số bất định và sai số hệ thống hoặc sai số xác định. Sai số của một dãy phép đo lặp lại là tổng của hai loại sai số đó:

$$E_a = E_r + E_s \quad (3)$$

Ở đây  $E_r$  là sai số ngẫu nhiên của phép đo và  $E_s$  là sai số hệ thống.

**Độ chính xác** là thuật ngữ chỉ sự trùng hợp lặp lại của các kết quả đo thu được trong một điều kiện như nhau. Như vậy độ chính xác chỉ phụ thuộc vào sự phân bố của các sai số ngẫu nhiên và không liên quan đến giá trị thực của Mê du răng. Độ chính xác được biểu diễn và tính toán qua độ lệch chuẩn của các giá trị đo đồng thời.

**Độ lệch chuẩn:** Độ lệch chuẩn  $\sigma$  hay S (Standard deviation) được định nghĩa bởi biểu thức :

$$\sigma = \sqrt{\frac{\sum_{x=1}^n (x_i - \bar{x})^2}{n - 1}} \quad (4)$$

Độ lệch chuẩn tương đối (Relative Standards Deviation – RSD được định nghĩa bởi biểu thức:

$$RSD = \frac{\sigma}{\bar{x}} \cdot 100 \quad (5)$$

Độ lệch chuẩn càng lớn thì độ chính xác càng kém . Các giá trị đo được dùng trong tính toán cần được tiến hành độc lập, khách quan trong cùng điều kiện như nhau. Nhưng vấn đề này ta sẽ đi sâu vào các chương sau liên quan đến bản chất thống kê các sai số đo ngẫu nhiên.

**Giá trị thật (True Value)** là giá trị (thường là hàm lượng) thực có của Mê su rang, Giá trị này có được là do ta tự pha chế (Mẫu chuẩn giả) hoặc là

xác định được bởi các phương pháp phân tích hoàn hảo nhất của nhiều phòng thí nghiệm tin cậy (Sẽ nói rõ hơn trong chương về Mẫu chuẩn).

**Sai số (Error):** Sai số (của phép đo) là giá trị chênh lệch giữa giá trị đo được và giá trị thật của Mê du rằng. Người ta thường biểu diễn chúng dưới dạng sai số tuyệt đối và giá trị tương đối.

**Sai số ngẫu nhiên (Random Error):** Là kết quả phép đo trừ đi giá trị trung bình thu được từ một tập hợp các phép đo đồng thời trong cùng điều kiện như nhau.

**Sai số hệ thống (Systematic Error)** là giá trị chênh lệch giữa kết quả thật và giá trị trung bình thu được từ một tập hợp các kết quả xác định song song trong cùng điều kiện như nhau.

Như vậy cả sai số ngẫu nhiên và sai số hệ thống đều đóng góp vào phần sai số chung của phép đo. Về mặt số học chúng liên hệ với nhau như sau:

**Sai số chung = Sai số hệ thống + sai số ngẫu nhiên .**

**Độ lệch chuẩn tổng cộng (Pooled Standard Deviation):** Đây là độ lệch thu được do sự cộng hợp của các độ lệch riêng phần của các tập đo trong cùng điều kiện như nhau. Độ lệch chuẩn tổng cộng được biểu diễn bằng công thức sau:

$$S_{\text{pool}} = \sqrt{\frac{\sum_{i=1}^{n_1} (x_i - \bar{x}_1)^2 + \sum_{i=1}^{n_2} (x_i - \bar{x}_2)^2 + \dots + \sum_{i=1}^{n_k} (x_i - \bar{x}_k)^2}{\sum_{i=1}^k n_i = k}} \quad (6)$$

Ở đây  $n_1, n_2, n_3 \dots$  là số các kết quả xác định trong mỗi tập k và  $x_1, x_2, x_3, \dots, x_n$  là giá trị trung bình tương ứng của mỗi tập k.

**Hệ số biến sai (Coefficient of Variation):** Được định nghĩa qua độ lệch chuẩn và giá trị trung bình qua biểu thức sau :

$$CV = \frac{s}{x} \cdot 100\% \quad (7)$$

**Trung vị (Median)** Trung vị là một giá trị trong tập đo mà so với giá trị này thì số các giá trị khác lớn hơn hoặc nhỏ hơn là như nhau .

- Nếu số tập hợp là lẻ thì lấy điểm trung tâm là trung vị.

Ví dụ ta có tập hợp 5 giá trị sau :

2,9

2,6

2,4

2,3

2,2

Giá trị trung bình là

$$\bar{x} = 2,5$$

Trung vị được chọn là số giữa : 2,4

Nếu số tập hợp là chẵn thì trung vị là giá trị trung bình của hai số giữa .

Ví dụ ta có tập hợp 4 giá trị :

0,1000

0,0902

0,0886

0,0884

$$\bar{x} = 0,0918$$

$$\text{Trung vị: } (0,0902 + 0,0886) / 2 = 0,0894$$

Trong trường hợp lý tưởng thì giá trị trung bình và trung vị trùng nhau, tuy nhiên trong thực tế ít khi gặp trường hợp như vậy đặc biệt khi số tập hợp đo là không lớn .

**Độ lặp lại (Repeatability):** Độ lặp lại là đại lượng chỉ ra mức độ phân tán các kết quả đo, nó đặc trưng cho độ gần về giá trị tuyệt đối của hai hay nhiều hơn các phép đo thu được trong cùng điều kiện như nhau. Độ lặp lại được biểu diễn dưới hai dạng là độ lặp lại tuyệt đối và độ lặp lại tương đối.

**Độ lặp lại tuyệt đối:** Đó là hiệu số giữa giá trị  $x_i$  đo được nào đó trong tập đo và giá trị trung bình cộng ( $x_i - \bar{x}$ )

Ví dụ khi phân tích hàm lượng clorua người ta nhận được 3 kết quả sau :

Số xác định      Hàm lượng ( %)      lệch khỏi giá trị trung bình ( $x_i - \bar{x}$ )

$x_1$                       24, 39                      0, 077

$x_2$                       24, 19                      0, 123

$x_3$                       24, 36                      0, 047

$$\Sigma = 72, 94$$

$$\Sigma = 0, 247$$

$$\bar{x} = 72,94/3 = 24,31$$

Trung bình của độ lệch trung bình :       $0, 247 / 3 = 0, 08$

Độ lệch giữa giá trị cao nhất  $x_{Max}$  và giá trị thấp nhất  $x_{Min}$

$$24,39 - 24,19 = 0,20.$$

Ta cũng có thể biểu thị độ lặp lại bằng độ lặp lại khỏi trung vị. Tuy nhiên cách này ít khi được dùng.

Độ lặp lại có thể biểu diễn qua cách tính thống kê qua độ lệch chuẩn và độ phân tán(sẽ đi sâu ở chương sau)

**Độ lặp lại tương đối:** Đó là tỷ số giữa độ lệch khỏi giá trị trung bình chia cho giá trị trung bình:



$$(x_i - \bar{x})/\bar{x}$$

Trong ví dụ trên ở phép xác định  $x_i$  là:  $\frac{x_i - \bar{x}}{x_2} = 0.32$

### **Giới hạn tin cậy ( Confidence limit ) và Khoảng tin cậy (Confidence interval )**

Giá trị trung bình thực  $\mu$  là hằng định và giá trị thật của nó luôn luôn không biết được . Tuy nhiên nhờ thống kê có thể tìm được các giới hạn của vùng quanh các giá trị thực nghiệm trung bình  $\bar{x}$  , trong đó cân chờ đợi với mức độ đã cho của xác suất tìm thấy giá trị trung bình thực. Các giới hạn tìm được như vậy gọi là giới hạn tin cậy. Khoảng được giới hạn bởi các cận trên và cận dưới được gọi là khoảng tin cậy. Cần lưu ý rằng giá trị của khoảng tin cậy phụ thuộc vào xác suất của khả năng giá trị trung bình rơi vào đó. Xác suất càng nhỏ thì khoảng tin cậy càng nhỏ và ngược lại:

$$x \pm \frac{ts}{\sqrt{n}}$$

Khoảng giá trị từ  $-\frac{ts}{\sqrt{n}}$  đến  $+\frac{ts}{\sqrt{n}}$  được gọi là khoảng tin cậy

## Chương II

### BẢN CHẤT, NGUỒN GỐC CÁC LOẠI SAI SỐ VÀ PHÂN LOẠI

#### 2.1. Các loại sai số và phân loại

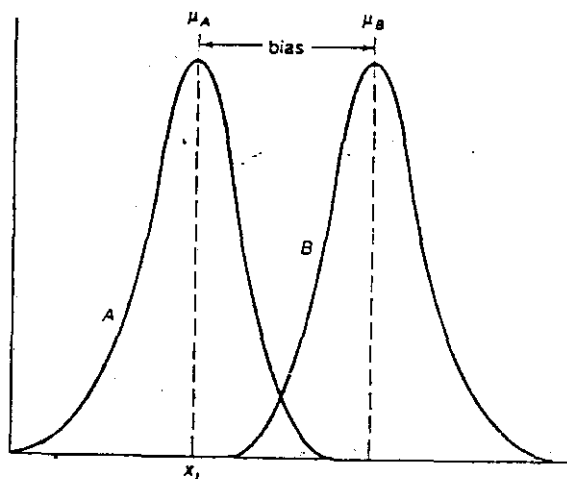
Bất kỳ một kết quả đo nào cũng gặp phải sai số. Sai số này không đơn thuần do một sai số đơn lẻ nào gây ra mà là tổng hợp của nhiều loại sai số có nguồn gốc khác nhau. Nghiên cứu sâu vào bản chất sai số người ta có thể chia thành hai loại sai số tùy theo nguồn gốc xuất hiện của chúng. Đó là **sai số hệ thống** (còn gọi là **sai số xác định**) và **sai số ngẫu nhiên** (còn gọi là **sai số bất định**). Trong thực tế còn một loại sai số thứ ba thường gặp là **sai số thô**, loại này trong nhiều trường hợp tăng lên do tính không cẩn thận, vụng về, lười biếng hoặc sự rủi ro của người làm thí nghiệm. Các nguồn gốc điển hình của loại sai số này bao gồm sự dịch chuyển số trong số liệu ghi, đánh đổ mẫu, dùng thang đo sai, đưa đột nhiên chất bản vào và làm đảo dấu trên máy đọc. Một sai số thô trong một tập các phép đo lặp lại xuất hiện như là một số lạc tức là số liệu khác hẳn với các số khác trong tập. Hiển nhiên loại sai số này không có gì phải đi sâu nghiên cứu. Các sai số hệ thống là các sai số mà về nguyên tắc có thể tìm ra nguyên nhân bản chất và đại lượng của sai số này có thể tính toán và xác định được. Các sai số hệ thống có giá trị xác định, có nguyên nhân nhất định và có cùng dấu và cùng độ lớn đối với các phép đo lặp lại được thực hiện theo y hệt nhau. Các sai số hệ thống gây ra sự sai lệch trong kỹ thuật đo. Độ sai lệch được minh họa bởi 2 đường cong trên Hình 2. trên đó có biểu diễn sự phân bố tần số các kết quả đo lặp lại trong phân tích các mẫu đồng nhất bằng hai phương pháp có sai số ngẫu nhiên cùng độ lớn.

Phương pháp A không có sai lệch, nên giá trị trung bình giới hạn được coi là giá trị thực  $x_t$ .

Phương pháp B có độ sai lệch:

$$\text{Độ sai lệch} = \mu_B - x_t = \mu_B - \mu_A$$

Chú ý rằng độ sai lệch ảnh hưởng tới toàn bộ số liệu trong dãy đo và có dấu. Các sai số hệ thống có 3 nguồn gốc: sai số công cụ, sai số cá nhân và sai số phương pháp.



Hình 2. Minh hoạ độ sai lệch, Độ sai lệch =  $\mu_B - x_i$

Còn sai số ngẫu nhiên thì phức tạp hơn nhiều, khó biết được nguồn gốc xuất hiện của chúng, còn các đại lượng của chúng thì dao động ngẫu nhiên và không thể đo và loại trừ dễ dàng được. Phải tìm bản chất và nguồn gốc của nó trong các hiện tượng và quá trình ngẫu nhiên. Các hiện tượng và quá trình này sẽ có quy luật khi số phép đo càng lớn và tuân theo các quy luật của thống kê xác suất. Và đây cũng chính là phần quan trọng và lý thú nhất của cuốn sách này.

## 2.2. Nguồn gốc của sai số hệ thống

Nguồn gốc của sai số hệ thống có rất nhiều, nhưng tựu trung nó bao gồm từ các nguyên nhân sau :

- Sai số của kỹ thuật đo hay phương pháp phân tích (nồng độ dung dịch, thuốc thử, độ tinh khiết của hoá chất)
- Các sai số có nguồn gốc từ máy móc thiết bị đo ( hệ thống Đề tectơ, sự khuếch đại tín hiệu, sự trôi và dịch chuyển phổ, hiệu ứng nhiệt độ, ảnh hưởng pH, sai số chuẩn máy ...

- Kỹ năng và sai sót của người phân tích.

Ta thử đi sâu vào phân tích các nguyên nhân trên:

a. Các sai số mang tính chủ quan của phân tích viên .

Đây là sai số xuất hiện do sự thiếu hiểu biết và kỹ năng tay nghề của phân tích viên hoặc sự thiếu cẩn trọng trong quá trình thực nghiệm. không tuân thủ nghiêm túc quy trình phân tích, làm vung vãi và mất mát thậm chí đưa thêm các chất lạ gây ảnh hưởng đến mẫu, hoặc bỏ qua các hiệu chuẩn về thể tích, đọc sai các vạch chia trên thang đo, lọc rửa kết tủa không cẩn thận...Ở đây cũng phải lưu ý đến tính khách quan của phân tích viên nhiều khi cứ định kiến cố ghi kết quả lệch về phía chiều mong muốn và dự đoán. Đây là chưa kể những khuyết tật bẩm sinh của người phân tích (như bị bệnh mù màu..) không phân biệt được các màu sắc khi chuyển màu ở điểm tương đương trong quá trình chuẩn độ.

b. Các sai số do phương pháp.

Nhiều khi trong phương pháp và quy trình phân tích có một công đoạn nào đó (ví như kết tủa, chiết ...) không chuyển hoá hoàn toàn một trăm phần trăm như lý thuyết, hoặc độ chuẩn của thuốc thử không chính xác hoặc bị biến đổi theo thời gian, hay quá trình phân huỷ mẫu không hoàn toàn, chất cần phân tích không tan hết vào dung dịch hoặc bị hấp phụ vào kết tủa.

c. Sai số do dụng cụ.

Các sai số này gây ra do sự không hoàn thiện của dụng cụ hoặc thiết bị đo. Chẳng hạn thể tích của bình đo (pipet, bu-ret) không chính xác hoặc không được chuẩn hoá cẩn thận tương ứng theo nhiệt độ, hiệu suất đo của đề tec tơ kém, xuất hiện các hiện tượng chuyển dịch và trôi phổ ...

### **2.3. Phát hiện, xác định và loại trừ các sai số hệ thống .**

Về nguyên tắc sai số hệ thống có thể phát hiện đánh giá đại lượng và loại trừ. Người ta thường điều chỉnh các sai số dụng cụ khi chuẩn hoá máy đo. Nên định kỳ chuẩn hoá máy đo vì các thông số của máy dễ bị thay đổi theo nhiệt độ, độ ẩm, điện áp ...

Nhiều sai số chủ quan do trình độ và tay nghề có thể khắc phục được bằng các lớp huấn luyện và thi tay nghề (proficiency test). Tuy nhiên việc chọn đúng phương pháp phân tích cho một đối tượng có hàm lượng nhất định là rất quan trọng tránh được sai số hệ thống lớn. Sau đây là một số cách được khuyến cáo áp dụng để phát hiện và loại trừ sai số hệ thống:

- *Phân tích các mẫu chuẩn đã biết chính xác hàm lượng* (CRM-Certified Reference Materials). Đây thường là các mẫu chuẩn quốc tế hay quốc gia ( như IAEA-International Atomic Energy Agency, NBS - National Bureau of Standards, ASTM-American Standards for Testing of Materials, NIST-National Institute for Standards and Technology)
- *Phân tích mẫu chuẩn giả hay mẫu thử cấp*: Các mẫu chuẩn này cần được chuẩn bị và pha chế rất cẩn thận sao cho nồng độ của các cấu tử được biết với độ tin cậy và chính xác rất cao.
- *Phân tích bằng các phương pháp độc lập*: Các phương pháp độc lập cần nhất là phải có độ tin cậy và độ chính xác đảm bảo, không nên giống về bản chất hay kiểu phân tích của phương pháp đã dùng
- *Phân tích mẫu trắng*: Các sai số cố định dùng trong các phép đo có thể được loại trừ nếu tiến hành phân tích mẫu trắng với đầy đủ các bước như khi làm với mẫu thật. Các kết quả thu được từ mẫu trắng dùng làm giá trị hiệu chuẩn cho phép phân tích thực.

Như trên đã nói, ta có thể tìm được nguyên nhân và nguồn gốc sai số hệ thống và từ đó tìm cách loại trừ được nó, còn sai số ngẫu nhiên thì rất phức tạp, khó tìm được các nguyên nhân gây ra nó mà chỉ có thể cố gắng giảm nó đến mức tối thiểu bằng cách phân tích cẩn thận nhiều lần rồi xử lý thống kê toán học các dữ kiện thu được. Do vậy phần sai số ngẫu nhiên được tách thành một chương riêng.

### Chương III

## SAI SỐ NGẪU NHIÊN

Như ta đã biết để đánh giá các kết quả đo người ta phải dựa vào hai tiêu chuẩn quan trọng nhất là độ đúng và độ lặp lại (hay còn gọi là độ chính xác). Độ lặp lại phụ thuộc nhiều vào các sai số ngẫu nhiên của toàn bộ quá trình phân tích. Sai số ngẫu nhiên càng lớn thì độ tán mạn của các kết quả phân tích càng cao hay nói khác đi độ chụm kém và tất nhiên là độ chính xác thấp. Trong thực tế ta hay gặp sai số ngẫu nhiên nhỏ hơn là sai số ngẫu nhiên lớn. Nếu không chú ý ta dễ dàng bỏ qua chúng (loại sai số bé). Trong trường hợp này ta có cảm giác là thu được các kết quả như nhau nhưng thực tế là khác nhau. Làm sao tăng được độ chính xác tức giảm sai số ngẫu nhiên. Muốn vậy trước hết phải nắm vững bản chất của hiện tượng ngẫu nhiên và sai số ngẫu nhiên. Tính ngẫu nhiên mang trong mình nó tính quy luật, tính tất nhiên khi số các phép đo là nhiều vô hạn. Đó là quy luật ngẫu nhiên. Để hiểu bản chất của vấn đề này ta bắt đầu đi vào nghiên cứu các đại lượng ngẫu nhiên và các loại phân bố.

### 1. Các đại lượng ngẫu nhiên và sai số ngẫu nhiên

Ta nhắc tóm tắt lại các thuật ngữ, định nghĩa cũng như các công thức liên quan đến các đại lượng ngẫu nhiên và diễn đạt chúng theo ngôn ngữ thống kê toán học:

**Độ đúng** Độ đúng thường được biểu thị theo sai số tương đối hoặc sai số tuyệt đối. Sai số tuyệt đối  $E_d$  của giá trị trung bình  $\bar{x}$  của một dãy nhỏ các phép phân tích lặp lại được cho bởi hệ thức:

$$E_d = \bar{x} - x_i \quad (8)$$

### Độ chính xác

Độ chính xác được biểu diễn và tính toán qua độ lệch chuẩn  $\sigma$  (hay  $s$ ) của các giá trị đo đồng thời :

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n}} \quad (9)$$

Đối với một số không lớn các phép đo lặp lại thì phương trình (9) trên không thể dùng trực tiếp vì rằng giá trị trung bình của một số vô hạn phép đo được ký hiệu là  $\mu$  chính là giá trị thực khi không có sai số hệ thống – giá trị này không bao giờ biết được. Thay cho giá trị này ta phải dùng giá trị trung bình của tập nhỏ các phép đo  $\bar{x}$ . Trong nhiều trường hợp  $\bar{x}$  khác với  $\mu$ . Ở đây thay cho phương trình độ lệch chuẩn  $\sigma$  trong phương trình (9) ta dùng  $S$ .

$$S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} \quad (10)$$

**Độ lặp lại:** Độ lặp lại là đại lượng chỉ ra mức độ phân tán các kết quả đo, nó đặc trưng cho độ gần về giá trị tuyệt đối của hai hay nhiều hơn các phép đo thu được trong cùng điều kiện như nhau. Trong sai số ngẫu nhiên có thể coi hai khái niệm độ lặp lại và độ chính xác là một.

**Khoảng tin cậy:** Để tính khoảng tin cậy, ta phải quay về với phân bố chuẩn  $t$ . Có thể tính khoảng tin cậy qua giá trị trung bình  $\bar{x}$  theo công thức sau:

$$\bar{x} \pm \frac{t(P.F)s}{\sqrt{n}} = \bar{x} \pm \Delta x \quad (11)$$

Đồng thời cần phải chọn trước xác suất  $P$ . Ta có thể lấy các giá trị cần thiết của  $t(P, f)$  ở bảng 3.

## 2. Xử lý thống kê các sai số ngẫu nhiên

Trong một số mục dưới đây sẽ xem xét các kỹ năng thống kê để phân tích các dữ liệu tuân theo phân bố ngẫu nhiên.

### 2.1. Tập toàn thể và tập chọn lọc (mẫu)

Khi xử lý thống kê các số liệu, người ta thường giả thiết rằng một nhóm các kết quả thực nghiệm lặp lại thu được trong phòng thí nghiệm là một phần nhỏ của một số vô hạn các kết quả về nguyên tắc có thể thu được trong khoảng thời gian không hạn định với lượng mẫu vô hạn.

Các nhà thống kê đã gọi nhóm số liệu là một tập chọn lọc (mẫu) và coi nó như là một bộ phận của một tập toàn thể vô hạn các số liệu mà về nguyên tắc có tồn tại. Các định luật thống kê chỉ áp dụng hạn chế cho các tập toàn thể. Khi áp dụng các định luật này cho một tập chọn lọc các số liệu thí nghiệm cần phải coi tập chọn lọc là thực sự đại biểu cho tập toàn thể. Bởi vì giả thiết này không có gì bảo đảm nên các trình bày về sai số ngẫu nhiên không tránh khỏi mang tính chất bất định và phải được trình bày theo ngôn ngữ xác suất.

**Trung bình tập:** Trung bình tập ( $\mu$ ) hoặc trung bình giới hạn của một dãy số liệu lặp lại được định nghĩa theo phương trình:

$$\mu = \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n x_i}{n} \quad (12)$$

Ở đây,  $x_i$  là giá trị của lần đo thứ  $i$ . Từ phương trình trên ta thấy giá trị trung bình của dãy phép đo càng gần với giá trị trung bình tập khi số phép đo



n càng gần với vô hạn. Điều quan trọng cần bổ sung là, khi không có sự sai lệch thì  $\mu$  là giá trị thực đối với đại lượng cần đo.

### Độ lệch chuẩn tập ( $\sigma$ ) và phương sai tập ( $\sigma^2$ )

Độ lệch chuẩn tập và phương sai tập, về mặt thống kê là thước đo quan trọng về độ chính xác của tập toàn thể các số liệu. Như vậy:

$$\sigma = \sqrt{\lim_{n \rightarrow \infty} \frac{\sum_{i=1}^n (x_i - \mu)^2}{n}} \quad (13)$$

Ở đây,  $x_i$  vẫn là giá trị ở lần đo thứ  $i$ .

Để ý rằng độ lệch chuẩn tập là căn bậc hai của các độ lệch riêng lẻ với giá trị trung bình của toàn tập. Các nhà thống kê thường hay biểu diễn độ chính xác của các số liệu theo phương sai, chẳng qua chỉ là bình phương của độ lệch chuẩn ( $\sigma^2$ ), bởi vì các phương sai là tổ hợp cộng tích. Điều đó có nghĩa là, nếu trong hệ có nguồn độc lập các sai số ngẫu nhiên thì phương sai tổng  $\sigma_t^2$  được tính theo hệ thức:

$$\sigma_t^2 = \sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \dots + \sigma_n^2 \quad (14)$$

Ở đây,  $\sigma_1^2, \sigma_2^2, \sigma_3^2, \dots, \sigma_n^2$  là các phương sai riêng lẻ.

Các nhà hoá học thường hay mô tả độ chính xác của phép đo theo độ lệch chuẩn hơn là phương sai bởi vì độ lệch chuẩn được biểu diễn cùng đơn vị của chính phép đo.

### Giá trị trung bình mẫu ( $\bar{x}$ )

Giá trị trung bình mẫu là giá trị trung bình của một dãy hữu hạn phép đo. Bởi vì trong trường hợp này  $n$  là một số hữu hạn, nên  $\bar{x}$  thường khác ít

nhiều với giá trị trung bình tập  $\mu$  và do đó khác với giá trị thực của đại lượng đo. Trong trường hợp này việc dùng các ký hiệu khác nhau đủ để nhấn mạnh sự phân biệt quan trọng này.

### **Độ lệch chuẩn mẫu (s) và phương sai mẫu ( $s^2$ )**

Độ lệch chuẩn (s) của một mẫu số liệu có kích thước hạn chế được tính theo phương trình:

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad (15)$$

Để ý rằng, độ lệch chuẩn mẫu khác với độ lệch chuẩn tập đã được định nghĩa theo phương trình 13 ở ba khía cạnh: Thứ nhất,  $\sigma$  được thay bằng s để nhấn mạnh sự khác nhau giữa hai khái niệm này. Thứ hai, giá trị trung bình thực  $\mu$  được thay bằng giá trị trung bình mẫu  $\bar{x}$ . Cuối cùng, n-1 được định nghĩa như là số bậc tự do xuất hiện ở mẫu số thay cho n.

### **Độ lệch chuẩn tương đối (RSD) và hệ số biến sai (CV)**

Độ lệch chuẩn tương đối thường cho nhiều thông tin hơn độ lệch chuẩn tuyệt đối. Độ lệch chuẩn tương đối của một mẫu số liệu được tính theo:

$$RSD = \frac{s \cdot 10^4}{\bar{x}} \quad (16)$$

Khi  $z = 2$  thì độ lệch chuẩn tương đối được tính theo phần trăm (như phương trình 5), khi  $z = 3$  thì độ lệch được tính theo phần nghìn. Trong trường hợp  $z = 2$ , độ lệch chuẩn tương đối cũng chính là hệ số biến sai (phương trình 7):

$$CV = \frac{s}{\bar{x}} \cdot 100\% \quad (17)$$

Với tập toàn thể các số liệu thì  $\sigma$  và  $\mu$  được dùng thay cho  $s$  và  $\bar{x}$  trong các phương trình (16) và (17).

### Cách tính độ lệch chuẩn mẫu

Khi dùng máy tính tay không có chức năng tính độ lệch chuẩn để tính  $s$  có thể sử dụng tiện lợi phương trình sau đây, tương đương với (15)

$$s = \sqrt{\frac{\sum x_i^2 - (\sum x_i)^2 / n}{n-1}} \quad (18)$$

**Ví dụ:** Cho các số liệu lặp lại sau đây về nồng độ  $\text{SO}_2$  trong khí quyển gần một nhà máy giấy: 1,96, 1,91, 1,88 và 1,94 phần triệu (ppm).

Tính: a) Giá trị trung bình ; b) Độ lệch chuẩn tuyệt đối và c) Hệ số biến sai của số liệu?

	ppm $\text{SO}_2$	
	$x_i$	$x_i^2$
$x_1$	1,96	3,8416
$x_2$	1,91	3,6481
$x_3$	1,88	3,5344
$x_4$	1,94	3,7636
	$\Sigma x_i = 7,69$	$\Sigma x_i^2 = 14,7877$

a)  $\bar{x} = 7,69/4 = 1,9225 = 1,92 \text{ ppm } \text{SO}_2$

b) Áp dụng phương trình 17 ta có :

$$S = \sqrt{\frac{14.7877 - (7.69)^2}{4-1}} = \sqrt{\frac{14.7877 - (7.69)^2}{3}} = 0.035 \text{ ppm SO}_2$$

$$\text{c) } CV = \frac{0.035 \text{ ppm}}{1.92 \text{ ppm}} \cdot 100\% = 1.8 \%$$

Chú ý rằng, sự khác nhau giữa  $\sum x_i^2$  và  $(\sum x_i)^2/N$  trong ví dụ 1.2 rất nhỏ, vì vậy nếu làm tròn vội vã thì sẽ phạm sai số nghiêm trọng khi tính giá trị của  $s$ . Do nguồn gốc này của sai số mà phương trình 18 không bao giờ được dùng để tính độ lệch chuẩn đối với các số chứa từ 5 chữ số trở lên, trong trường hợp này nên dùng phương trình 14. Điều quan trọng cũng cần phải lưu ý là các máy tính tay và các máy tính nhỏ có chức năng độ lệch chuẩn thường sử dụng một dạng của phương trình 18. Vì vậy, có thể thấy là khi dùng các loại máy tính đó để tính  $s$  cho các số liệu có từ 5 số có nghĩa trở lên sẽ mắc sai số lớn.

## 2.2. Áp dụng các phương pháp thống kê để kiểm tra các giả thiết

Nhiều nghiên cứu khoa học cần thiết kiểm tra các giả thiết. Người ta tạo ra các mô hình giả thuyết, trên cơ sở các mô hình này người ta xây dựng các thực nghiệm để kiểm tra tính đúng đắn của mô hình đặt ra. Nếu các cứ liệu thực nghiệm không xác nhận mô hình giả thiết thì người ta bỏ đi và lại tìm kiếm mô hình khác. Ngược lại nếu tìm thấy có sự phù hợp giữa các kết quả thực nghiệm với giả thiết thì mô hình giả định có thể làm cơ sở cho các thực nghiệm tiếp tục. Nếu giả thiết được xác nhận bởi một số lượng đáng kể số liệu thực nghiệm thì người ta công nhận giả định này trở thành lý thuyết. Tuy

nhiên các số liệu thực nghiệm ít khi phù hợp sít với các số liệu của mô hình lý thuyết, do vậy các nhà khoa học phải đi sâu giải quyết vấn đề: sự không phù hợp này có phải là do giả thuyết sai lầm hay hay là do các sai số ngẫu nhiên không tránh khỏi của phép đo. Do vậy hiển nhiên là cần có sự kiểm tra thống kê. Sau đó người ta đánh giá xác suất xuất hiện của sự sai khác quan sát được trong kết quả các sai số ngẫu nhiên theo các định luật thống kê. Thông thường thì sự sai khác bằng hay lớn hơn trong xác suất tin cậy 95% thì giả thuyết không được chấp nhận và sự sai khác được xem là đáng kể. Ta sẽ trở lại vấn đề này trong phần nói đến các loại phân bố.

Các nhà phân tích thường gặp các phép kiểm tra như: So sánh hai giá trị trung bình của hai tập ( $\bar{x}_1$  và  $\bar{x}_2$ ), giá trị trung bình của phép phân tích  $\bar{x}$  và đại lượng  $\mu$  được chấp nhận là thực, các độ lệch chuẩn  $\sigma_1$  và  $\sigma_2$  của hai dãy phép đo và cả độ lệch chuẩn  $S_1$  và  $S_2$  của một tập nhỏ và độ lệch chuẩn  $\sigma$  của tập lớn.

### **3. Các loại phân bố**

#### **3.1. Phân bố tần số thực nghiệm .**

Khác với sai số hệ thống, không thể loại trừ các sai số ngẫu nhiên từ các phép đo mà chỉ có thể giảm thiểu nó từ nhiều phép đo song song trong cùng một điều kiện như nhau. Trước hết cần phải thống kê và hệ thống hoá các số liệu thực nghiệm và phân tích xem sự phân bố các giá trị thực nghiệm tuân theo những quy luật thống kê nào. Bước đầu ta trình bày chúng dưới dạng bảng biểu, sau đó sẽ biểu diễn chúng dưới dạng đồ thị. Ở đây phương pháp đồ thị tỏ ra rất quan trọng và bổ ích. Việc biểu diễn các kết quả đo theo tần số xuất hiện của nó rất quan trọng. Chúng được trình bày dưới dạng các điểm riêng biệt trên trục số đã chia tuyến tính (một chiều) và sau đó đánh giá về mật độ xuất hiện. Sự phân bố này được gọi là phân bố một chiều. Trực quan hơn ta có thể biểu diễn theo giản đồ bậc thang theo đó giá trị của một đại lượng đo  $x$  nào đó được phân bố theo cấp tùy theo tần xuất xuất hiện  $h$  của nó.

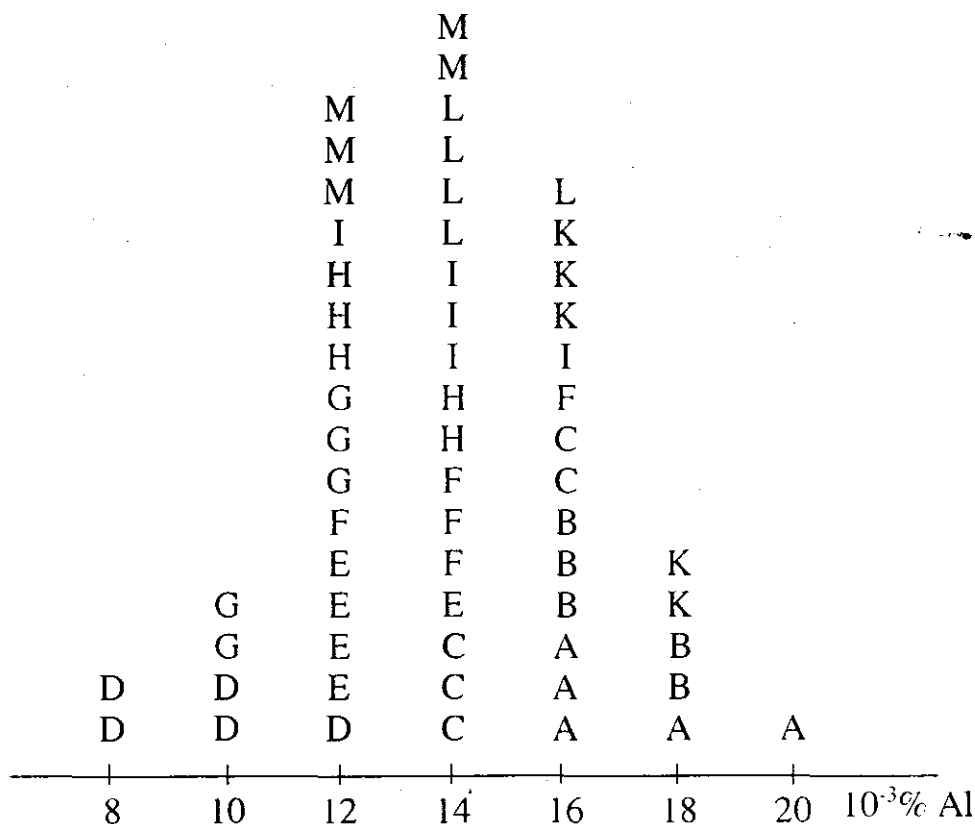
Cột cao nhất của giản đồ bậc thang này chỉ ra giá trị đo thường được xuất hiện nhiều nhất và tương ứng với điểm này ở trục tung là giá trị mật độ. Để xây dựng giản đồ bậc thang ta tập hợp các giá trị riêng rẽ thành  $k$  cấp có bề rộng  $d$ . Số các cấp  $k$  phải bằng khoảng căn bậc hai của tổng số các giá trị đo được, nhưng số này phải không được nhỏ hơn 5 và lớn hơn 20. Nếu  $k$  quá bé sẽ làm phai mờ đi những chi tiết đặc trưng của phân bố tần số, còn nếu  $k$  quá lớn (chia quá chi tiết) có thể xoá nhoà đi hình ảnh tổng quát do các độ lệch ngẫu nhiên bé. Chẳng hạn số phép đo là 100 thì nên chia số cấp là 10 ( $k=10$ ). Khi chọn các giới hạn của cấp cần chú ý để cho giới hạn trên của cấp dưới nhỏ hơn giới hạn dưới của cấp trên liền kề.

Khi các kết quả đo một mẫu được sắp xếp theo cách trên trong các điều kiện thí nghiệm hoàn hảo thường ta thu được phân bố đối xứng với một cực đại tần số. Những phân bố tần số không đối xứng có các cực đại lệch về bên trái hoặc bên phải chứng tỏ rằng có những nhược điểm của điều kiện thí nghiệm hoặc là sự phân chia cấp không thích hợp.

Ta hãy xem ví dụ về các kết quả xác định hàm lượng nhôm (% Al) trong thép ở 12 phòng thí nghiệm, mỗi phòng thí nghiệm cho 5 kết quả thu được trong các ngày khác nhau như sau:

Phòng thí nghiệm	A	0,016	0,015	0,017	0,016	0,019
-	B	0,017	0,016	0,016	0,016	0,018
-	C	0,015	0,014	0,014	0,014	0,015
-	D	0,011	0,007	0,008	0,010	0,009
-	E	0,011	0,011	0,013	0,012	0,012
-	F	0,012	0,014	0,013	0,013	0,015
-	G	0,011	0,009	0,012	0,010	0,012
-	H	0,011	0,011	0,012	0,014	0,013
-	I	0,012	0,014	0,015	0,013	0,011
-	K	0,015	0,018	0,016	0,017	0,016
-	L	0,015	0,014	0,013	0,014	0,014
-	M	0,012	0,014	0,012	0,013	0,012

Như vậy ta có  $n = 60$  giá trị. Giá trị thấp nhất là của phòng thí nghiệm D với  $x_{D2} = 0,007\%$  nhôm, giá trị cao nhất là của phòng thí nghiệm A với  $x_{A5} = 0,019\%$  nhôm. Sau khi tập hợp các số liệu thành 7 cấp ( $k=7$ ) với độ rộng của cảo  $d = 0,002\%$  Al, ta có  $k \approx \sqrt{n}$ . Cấp thấp nhất bao gồm các giá trị 0,007 và 0,008%, cấp thứ hai-0,09 và 0,010 % ... Với sự phân chia các cấp như vậy ta có phân bố tần số như trên hình 3.



Hình 3. Phân bố tần số khi xác định đồng thời nhôm trong thép tại 12 Phòng thí nghiệm

Mặc dù xuất xứ các giá trị khác nhau, phân bố tần số vẫn có cực đại rõ rệt. Điều đáng ngạc nhiên nữa là các giá trị lặp lại trong các phòng thí nghiệm riêng biệt lại nằm kế nhau, trong khi đó khi so sánh các dữ kiện trong các phòng thí nghiệm riêng biệt lại có các số liệu khác biệt đáng kể.

Từ dạng phân bố tần số ta có được một hình ảnh tổng quát về sự xuất hiện sai số ngẫu nhiên. Đó là khi sai số ngẫu nhiên lớn ta có phân bố rộng, khi sai số ngẫu nhiên bé thì đường phân bố trở nên hẹp và nhọn. Rõ ràng ở đây

không có bất kỳ một thông tin nào về sai số hệ thống, bởi vì sai số này không làm thay đổi dạng phân bố. Sai số hệ thống không đồng nhất có thể xác định bằng các phương pháp khác nhau

*Bảng 1- Các phép đo lặp lại độ hấp thụ*

Số TT đo	Độ hấp thụ A	Số TT đo	Độ hấp thụ A	Số TT đo	Độ hấp thụ A
1	0,488	18	0,475	35	0,476
2	0,480	19	0,480	36	0,490
3	0,486	20	0,494*	37	0,488
4	0,473	21	0,492	38	0,471
5	0,475	22	0,484	39	0,486
6	0,482	23	0,481	40	0,478
7	0,486	24	0,487	41	0,486
8	0,482	25	0,478	42	0,482
9	0,481	26	0,483	43	0,477
10	0,490	27	0,482	44	0,477
11	0,480	28	0,491	45	0,486
12	0,489	29	0,481	46	0,478
13	0,478	30	0,469**	47	0,483
14	0,471	31	0,485	48	0,480
15	0,482	32	0,477	49	0,483
16	0,483	33	0,476	50	0,479
17	0,488	34	0,483	-	-

Giá trị trung bình độ hấp thụ = 0,482

Độ lệch chuẩn = 0,0056

Số liệu ghi theo thứ tự thu được.

\* Giá trị cực đại,

\*\* Giá trị cực tiểu



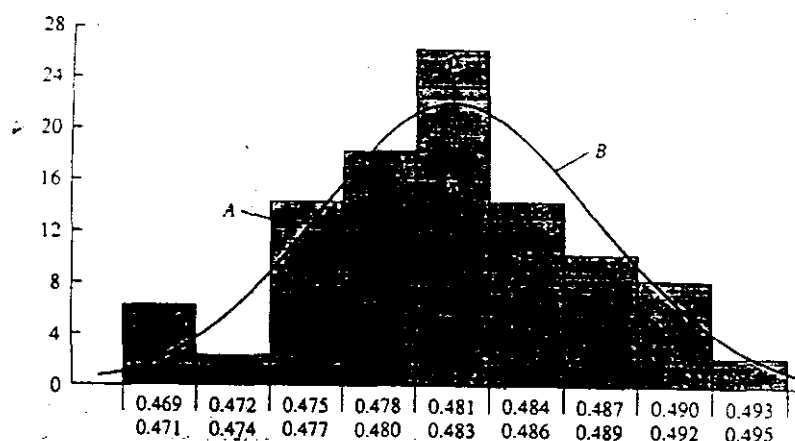
**Ví dụ:** Trong một thí nghiệm chung, nhiều phòng thí nghiệm có cùng sai số hệ thống (về cả độ lớn cũng như dấu). Do vậy ta thấy xuất hiện phân bố tần số với hai hoặc một số cực đại. Cực đại thứ hai có thể là vai của cực đại chính, nó hình như tạo thành phân bố giả khi sự chuyển dịch hệ thống không quá lớn. Ta lấy một thí dụ khác:

Khi đo mật độ quang của dung dịch phức màu đỏ tạo thành giữa Fe (III) và một lượng Thioxianat đủ dư, ta có các kết quả đo 50 lần như trên bảng 1

*Bảng 2. Sự phân bố tần số các số liệu từ Bảng 1*

Khoảng độ hấp thụ A	Số giá trị trong mỗi khoảng y	Tần số tương đối y/n*
0,469 đến 0,471	3	0,06
0,472 - 0,474	1	0,02
0,475 - 0,477	7	0,14
0,478 - 0,480	9	0,18
0,481 - 0,483	13	0,26
0,484 - 0,486	7	0,14
0,487 - 0,489	5	0,10
0,490 - 0,492	4	0,08
0,493 - 0,495	1	0,02
n* = Tổng số phép đo.		

Điều hợp lý là nếu coi số phép phân tích lớn hơn rất nhiều so với số liệu đã ghi trong Bảng 2. và nếu kích thước các tế bào nhỏ hơn rất nhiều thì trong trường hợp tối ưu ta sẽ được một đường cong đều như đã vẽ trên Hình 4.1.B. Đường cong thuộc loại này được gọi là đường cong Gauss hoặc đường cong chuẩn sai số. Bằng thực nghiệm người ta tìm thấy rằng các kết quả của các phép phân tích hoá học lặp lại thường được phân bố gần đúng với dạng Gauss hoặc dạng phân bố chuẩn khác mà ta sẽ đề cập đến trong phần sau.



Hình 4.1.A. Giản đồ chỉ ra sự phân bố của 50 kết quả đo trong bảng 1.

B. Đường cong Gauss đối với các số liệu có cùng giá trị trung bình và độ lệch chuẩn như các số liệu trong A

Như đã nói ở trên các nhà thống kê thường dùng  $\mu$  để biểu diễn giá trị trung bình của một tập vô hạn số liệu (xem hình 2) và  $\bar{x}$  là giá trị trung bình của một dãy nhỏ số liệu lặp lại. Sai số ngẫu nhiên E, đối với giá trị trung bình của dãy nhỏ là:

$$E_r = \bar{x} - \mu \quad (19)$$

Người ta tìm thấy rằng đối với giá trị trung bình của dãy hữu hạn đạt nhanh tới giá trị thực khi số phép đo N lớn hơn 20 hoặc 30. Như vậy, qua ví dụ sau đây, ta có thể xác định sai số ngẫu nhiên của từng số liệu riêng lẻ cũng như giá trị trung bình của một dãy số liệu.

#### Ví dụ:

Tính sai số ngẫu nhiên đối với: a) Số liệu thứ hai của Bảng 1 và b) Giá trị trung bình đối với 3 số liệu đầu tiên trong Bảng.

Giá trị trung bình đối với toàn bộ dãy là 0,482 và do số phép đo là 50 nên có thể coi sai số ngẫu nhiên bằng 0. Như vậy giá trị  $\mu$  được chấp nhận 0,482.

a) Ở đây, sai số ngẫu nhiên đối với phép đo duy nhất  $x_2$  là:

$$E_r = x_2 - \mu = 0,480 - 0,482 = -0,002.$$

b) Giá trị trung bình  $\bar{x}$  đối với 3 giá trị đầu tiên trong bảng là:

$$\bar{x} = \frac{0,488 + 0,480 + 0,486}{3} = 0,485$$

Thay thế vào phương trình 18 ta được:

$$E_r = \bar{x} - \mu = 0,485 - 0,482 = +0,003.$$

### 3.2. Phân bố lý thuyết

Khi hệ thống hoá các kết quả đo và biểu diễn chúng bằng đồ thị ta thu được một hình ảnh chung về sự phân bố các sai số ngẫu nhiên mang tính quy luật, điều này cho phép ta giả thiết về sự tồn tại của một quy luật toán học nào đó chi phối sự phân bố các kết quả của sai số ngẫu nhiên. Đó chính là quy luật về phân bố lý thuyết các đại lượng đo ngẫu nhiên.

#### 3.2.1. Phân bố chuẩn Gauss

Ta giả thiết số phép đo  $n$  là vô cùng lớn ( $n \rightarrow \infty$ ). Khi đó độ rộng của lớp chia là rất bé ( $d \rightarrow 0$ ) và sự phân bố tần số của một giá trị có thể mô tả bằng hàm xác suất sau:

$$y = P(x)dx = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dx \quad (20)$$

Ở đây  $P(x)dx$  là xác suất của sai số ngẫu nhiên trong khoảng  $x$  đến  $x+dx$  ;

$\sigma$ - Độ lệch chuẩn, là tham số đặc trưng cho một dãy đã cho của một số lớn phép đo.

$\mu$ - Giá trị trung bình tổng

Trong phương trình trên  $x$  là giá trị của một phép đo riêng biệt,  $\mu$  là giá trị trung bình đại số của một tập hợp vô tận phép đo. Đại lượng  $(x-\mu)$  là độ lệch khỏi giá trị trung bình,

$y$  là tần số xuất hiện của mỗi giá trị  $(x-\mu)$ .

Hàm phân bố này được gọi là phân bố chính quy hay phân bố chuẩn Gauss. Phân bố này do nhà toán học Carl Friedrich Gauss (1777-1855) tìm ra khi nghiên cứu phân tích các kết quả đo đạc trắc địa.

Phương trình (20) mô tả mật độ xác suất của phân bố chuẩn, trong đó hai tham số  $\mu$  và  $\sigma$  có ý nghĩa rất quan trọng. Vị trí và dạng của đường cong phân bố được xác định bởi hai tham số này. Cực đại đường cong nằm ở điểm  $x = \mu$ . Các Điểm uốn ở  $x = \mu - \sigma$  và  $\mu + \sigma$  ( Hình 4.2. )

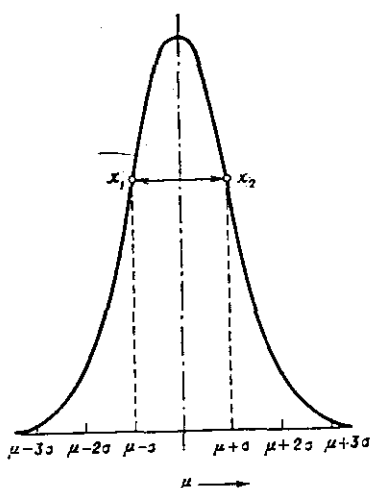
Độ lệch chuẩn  $\sigma$  thể hiện độ phân tán của phép đo hay nói khác đi là độ chính xác của phép đo. Độ lệch chuẩn càng bé độ chính xác càng cao và đường cong càng nhọn, độ lệch chuẩn càng lớn độ chính xác kém và đường cong tù.

Sự phân bố tần số được biểu thị bởi đường cong Gauss có các đặc điểm sau:

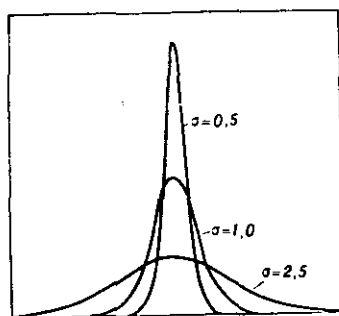
- 1- Kết quả thường gặp nhất là giá trị trung bình  $\mu$  của dãy số liệu.
- 2- Các kết quả phân bố một cách đối xứng xung quanh giá trị trung bình.
- 3- Các giá trị sai lệch nhỏ với giá trị trung bình trung tâm có tần số cao hơn các giá trị sai lệch lớn.
- 4- Khi không có sai số hệ thống thì giá trị trung bình của một dãy lớn số liệu gần với giá trị thực.

Đặc điểm thứ 4 cho thấy về nguyên tắc ta có thể làm giảm sai số ngẫu nhiên của phép phân tích xuống giá trị tối thiểu gần với 0. Đáng tiếc là trong thực tế rất hiếm khi có thể thực hiện được mục đích này, bởi vì yêu cầu này

đòi hỏi phải thực hiện ít nhất 20 phép phân tích lặp lại. Thông thường nhà khoa học chỉ có thời gian thực hiện 2 hoặc 3 phép đo lặp lại, do đó có thể thấy được là sẽ có sai số ngẫu nhiên đáng kể đối với giá trị trung bình của một số nhỏ như vậy các phép đo lặp lại.



Hình 4.2. Đường cong phân bố chuẩn



Hình 4.3. Dạng đường cong phân bố chuẩn với các độ lệch chuẩn khác nhau

Đường cong đạt giá trị  $y=0$  ở  $x=\pm \infty$ . Nhưng trong thực tế ta chỉ xét trong khoảng  $x = \mu \pm 3\sigma$

Hình 4.3. chỉ ra rằng 3 đường cong Gauss có diện tích bằng nhau với các giá trị trung bình  $\mu$  như nhau nhưng với các giá trị khác nhau của sai số bình phương trung bình  $\sigma$ .

Khi  $\sigma$  giảm thì các đường cong trở nên càng hẹp, càng nhọn. Các giá trị nằm trong khoảng  $\pm\sigma$  chiếm 68%, trong khoảng  $\pm 2\sigma$  chiếm 95% và trong khoảng  $\pm 3\sigma$  chiếm 99,7%

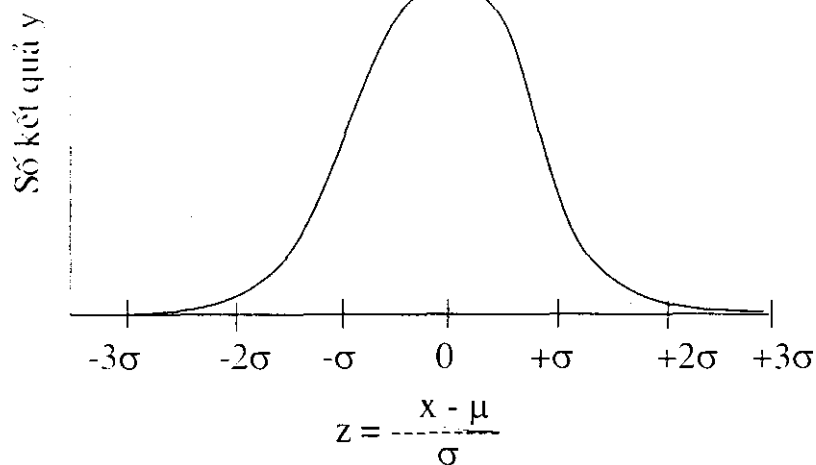
Để ứng dụng người ta chuẩn hoá phương trình (20) bằng cách đặt  $z = (x - \mu) / \sigma$

khi đó :

$\sigma_u = 1$  và hàm chuẩn hoá có dạng:

$$y = P(x) = \frac{1}{\sqrt{2\pi}} \cdot \exp. \frac{-x^2}{2} \quad (21)$$

Và đường cong theo biến số  $z$  này có dạng như trên hình 5. (Giá trị của  $z$  đặt trên trục hoành).



Hình 5. Đường cong phân bố theo biến số  $z$

Để có được sự tin cậy của các dữ kiện thực nghiệm thì tập hợp các kết quả thực nghiệm cần kiểm tra xem sự phân bố có tuân theo phân bố chuẩn hay không.

### 3.2.2. Phân bố Student T

Trong thực nghiệm nhiều khi các phép đo song song không đủ lớn, để tính độ lệch chuẩn  $S$  người ta đưa ra hàm phân bố Student T. Hàm này có dạng:

$$P(t, f) = K(f) \left( 1 + \frac{t^2}{f} \right)^{-\frac{f+1}{2}} \quad (22)$$

Với  $t = \frac{(\bar{x} - \mu)}{S_x}$  trong đó  $S_x = \frac{Sx}{\sqrt{n}} \rightarrow t = \frac{\bar{x} - \mu}{S} \cdot \sqrt{n}$

$K(f)$  là hằng số tùy thuộc vào bậc tự do  $f = n - 1$ ;  $t$  là hàm phân bố Student;  $S_x$  là độ lệch khỏi giá trị trung bình.  $P(t, f)$  là giá trị xác suất xuất hiện đại lượng  $t$  ứng với bậc tự do nhất định  $f$ . Hàm này có dạng đối xứng  $P(-t, f) = P(t, f)$ . Khi  $f \rightarrow \infty$  thì  $P(t, f) = 1$  ứng với giá trị

$$t = \pm \frac{\bar{x} - \mu}{S_x}$$

Giá trị  $t$  luôn luôn ứng với một xác suất và bậc tự do nhất định :

$$P(-t, t) = 2 \int_t^{\infty} P(t, f) dt = \alpha \quad (23)$$

Số thí nghiệm càng ít, xác suất  $P$  càng lớn thì giá trị  $t$  càng lớn. Trong thực tế người ta lập bảng tính sẵn  $t(\alpha, f)$  (xem bảng 3) và phụ lục. Tiến hành tra bảng với  $\alpha$  và  $f$  ta có  $t(\alpha, f)$ .

Từ biểu thức 22 ta có thể tính được  $\mu$  :

$$\mu = \bar{x} \pm \frac{t.s}{\sqrt{n}}$$

Hoặc  $\mu = \bar{x} \pm \varepsilon$  ,trong đó  $\varepsilon = \pm \frac{t.s}{\sqrt{n}}$

Như vậy giá trị thật  $\mu$  nằm trong khoảng  $\bar{x} - \varepsilon < \mu < \bar{x} + \varepsilon$  được coi là sai số ngẫu nhiên (tuyệt đối) và sai số tương đối (phần trăm) là :

$$\varepsilon = \pm \frac{t.s.100}{\bar{x}\sqrt{n}}$$

Giá trị từ  $\mu - \varepsilon$  đến  $\mu + \varepsilon$  được gọi là khoảng tin cậy (tương ứng với một xác suất nhất định), với xác suất lớn thì khoảng tin cậy rộng. xác suất bé thì khoảng tin cậy hẹp .

*Bảng 3. Giá trị t ứng với độ tin cậy  $\alpha$  và số bậc tự do  $f = n-1$*

f	P=0,90	P=0,95	P=0,99
1	6,31	12,7	63,7
2	2,92	4,3	9,92
3	2,35	3,18	5,84
4	2,13	2,78	4,60
5	2,01	2,57	4,03
6	1,94	2,45	3,71
7	1,89	2,36	3,50
8	1,86	2,31	3,36
9	1,83	2,26	3,25
10	1,81	2,23	3,17
15	1,75	2,13	2,95
20	1,73	2,06	2,79

**3.2.3. Phân bố chuẩn F (Fisher)**Phân bố này được dùng để so sánh độ chính xác của hai tập dữ liệu của hai phân tích viên trong cùng một phòng thí



thực nghiệm hoặc bởi hai phương pháp phân tích khác nhau trên cùng một mẫu phân tích hoặc là các kết quả từ hai phòng thí nghiệm khác nhau. Về mặt thống kê F được định nghĩa như là tỷ số của hai phương sai, hoặc đơn giản hơn là tỷ lệ của hai độ phân tán của hai tập dữ liệu :

$$F = \frac{S_1^2}{S_2^2} \quad (F > 1) \quad (24)$$

Trong đó  $S_1$  là phương sai lớn hơn và  $S_2$  là phương sai bé hơn. Do đó F luôn luôn lớn hơn 1 (xem phụ lục 4) Giá trị F có trong bảng thống kê đối với nhiều mức độ hiển nhiên khác nhau. Nó phụ thuộc vào số bậc tự do. Phụ lục 4 được thiết lập sao cho  $F > 1$ , bởi vậy độ phân tán S được so sánh trong điều kiện  $S_1 > S_2$ .

Trong bảng 4 sự lựa chọn được đưa ra các giá trị F tương ứng với xác suất tin cậy 95%. Tất cả các giá trị này ( với xác suất 95% ) phải có độ lớn sao cho tỷ số độ phân tán không vượt quá giá trị nêu trong bảng. Có một số bảng đầy đủ hơn với các xác suất tin cậy và số bậc tự do khác nhau .

Bảng 4. Các giá trị F tương ứng với xác suất 95% (  $P = 0,95$  ) với các bậc tự do  $k_1 = n_1 - 1$  và  $k_2 = n_2 - 2$

Giá trị f ở $\nu_1$						
$\nu_2$	2	3	4	5	6	$\infty$
2	19.00	19.16	19.25	19.30	19.33	19.50
3	9.55	9.28	9.12	9.01	8.94	8.53
4	6.94	6.59	6.39	6.26	6.16	5.63
5	5.79	5.41	5.19	5.05	4.95	4.36
6	5.14	4.76	4.53	4.39	4.28	3.67
$\infty$	3.00	2.60	2.37	2.29	2.10	1.00

**Ví dụ :** Giả thiết rằng ta có 2 dãy quan sát thực nghiệm với 6 kết quả xác định song song của dãy 1 và 4 kết quả của dãy 2. Dãy 1 có độ lệch chuẩn  $S_1 = 0,05$  và dãy thứ hai với  $S_2 = 0,02$ . Cần phải kiểm tra liệu  $S_1$  có cao hơn  $S_2$  một cách hiển nhiên hay không. Trong trường hợp đã cho số bậc tự do của dãy 1 :  $f = n - 1 = 5$  và dãy 2 có  $f = n - 1 = 3$  và từ bảng 4 ta tìm thấy  $F = 9,01$ . Tỷ số độ phân tán tìm thấy từ thực nghiệm là:

$$S_1 / S_2 = 0,0025 / 0,004 = 6,25 .$$

Về mặt lý thuyết chỉ có ( 1- 0,95 %) tức 5% khả năng  $F$  có thể lớn hơn 9.01. Từ đó cho kết luận giả thiết không là có cơ sở và không có sự khác biệt về mặt thống kê.

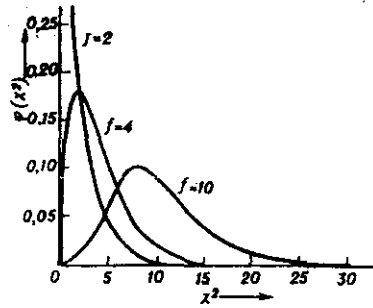
### 3.2.4. Phân bố $\chi^2$

Nếu tập hợp các đại lượng ngẫu nhiên  $x_1, x_2, \dots, x_n$  có phân bố chuẩn thì có thể thu được đại lượng ngẫu nhiên với số bậc tự do  $f = n - 1$  ta có hàm phân bố  $\chi^2$  :

$$\chi^2 = \sum_{i=1}^n \left( \frac{x_i - \bar{x}}{\sigma} \right)^2 = (n - 1) \frac{s^2}{\sigma^2} \quad (25)$$

Trong đó  $S^2$  là phương sai của tập chọn lọc và  $\sigma^2$  là phương sai của tập toàn thể.

Hàm phân bố  $\chi^2$  nằm trong góc phần tư thứ nhất trong miền từ  $\chi^2 = 0$  đến  $\chi^2 = \infty$ . Dạng của nó phụ thuộc rất nhiều vào số bậc tự do  $f$  (Hình 6). Đối với số bậc tự do bé thì đường cong là bất đối xứng, nhưng với sự tăng của  $f$ , thì sự bất đối xứng giảm và khi số bậc tự do lớn ta có đường cong Gauss với  $\mu > 0$ . Lấy tích phân hàm phân bố trong giới hạn từ 0 đến  $\chi_p^2$  ( $\chi_p^2 < \infty$ ) cho ta phần  $\bar{P}$  tổng diện tích dưới đường cong mà diện tích này ứng với xác suất để cho giá trị  $\chi^2 = \sum x_i^2$  thu được từ  $f$  quan sát độc lập trong khoảng ( 0, 1 ...  $\chi_p^2$ ). Phụ lục 3 là các giá trị cận tích phân của phân bố  $\chi^2$  theo các giá trị  $f$  ở các xác suất khác nhau.



Hình 6. Phân bố  $\chi^2$  đối với  $f=2, 4$  và  $10$

### 3.2.5. Phân bố Poát xông (Poisson )

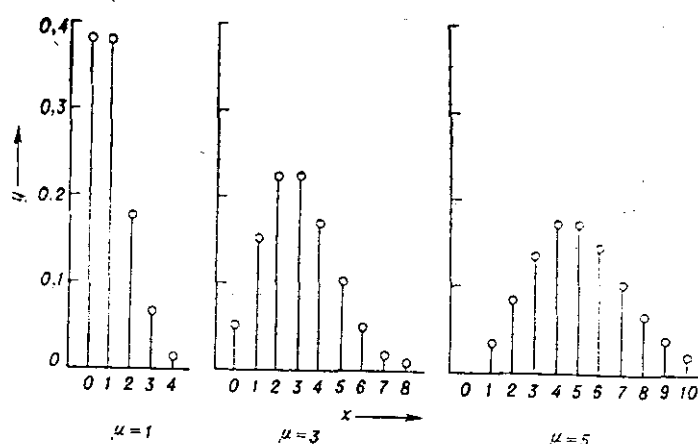
Có nhiều phép phân tích cho các kết quả đo dưới dạng các đại lượng nguyên, đặc biệt là các phương pháp hạt nhân hay hoá phóng xạ (phương pháp Huỳnh quang tia X, các phương pháp phổ  $\alpha, \beta, \gamma$ , phổ Ronghen). Các phương pháp này thường có kết quả đo là số đếm hay số xung trên một đơn vị thời gian. Do tính chất hiếm hoi của các sự kiện này trong khoảng thời gian quan sát thành phần của mẫu thay đổi không đáng kể. Nếu lặp lại nhiều lần cùng một thí nghiệm thì mối quan hệ giữa giá trị đo và tần xuất xuất hiện có thể biểu diễn như sau :

$$y = \frac{\mu^x e^{-\mu}}{x!} \quad (26)$$

Đây gọi là phân bố Poat xông. Loại phân bố này được đặc trưng bằng chỉ một tham số đó là giá trị trung bình  $\mu$ . Giữa giá trị trung bình  $\mu$  và sai số bình phương trung bình  $\sigma$  có mối quan hệ như sau:

$$\sigma = \sqrt{\mu} \quad (27)$$

Khác với phân bố chuẩn, phân bố Poat xong là rất rời rạc. Với các giá trị nhỏ nó rất bất đối xứng ( Hình 7)



Hình 7. Phân bố Poat xong với các giá trị khác nhau của  $\mu$

Sự bất đối xứng giảm xuống rất nhanh khi tăng  $\mu$  và dạng đường phân bố tiến tới phân bố chuẩn có giá trị trung bình  $\mu$  và sai số bình phương trung bình  $\sigma = \sqrt{\mu}$ . Trong thực tế khi  $x > 15$  có thể coi là xấp xỉ với phân bố chuẩn. Khi đó có 68,3% tất cả các giá trị nằm trong giới hạn  $\mu - \sqrt{\mu}$  và  $\mu + \sqrt{\mu}$

### 3.2.6. Quan hệ giữa các loại phân bố

Các loại phân bố vừa trình bày trên thoạt nhìn có vẻ không liên quan gì với nhau cả. Nhưng thực tế chúng có quan hệ với nhau trong những điều kiện nhất định. Ở đây ta không đi sâu vào toán học thống kê, nhưng chỉ biết rằng chẳng hạn phân bố Poat xong có thể tiến tới phân bố chuẩn nếu thoả mãn điều kiện  $x > 15$ . Hay phân bố Student cũng có thể chuyển thành phân bố chuẩn khi  $f \rightarrow \infty$ . Cũng có thể có những quan hệ tương tự giữa các loại phân bố khác. Ta biết rằng các phân bố đặc biệt ( t , f ,  $\chi^2$  ) chỉ là trường hợp riêng biệt của phân bố chuẩn với bài toán được cho nhất định và khi số bậc tự do có giới hạn .

## Chương IV

# ỨNG DỤNG ĐỂ TÍNH VÀ BIỂU DIỄN CÁC LOẠI SAI SỐ

### 1. Cách tính các sai số

Trước khi đi vào tính toán các ví dụ cụ thể (từ các số liệu thực nghiệm đã được công bố) ta nhắc lại khái niệm cơ bản nhất là sai số đo và sự cộng hợp các sai số.

**Sai số đo :** Như đã biết nguyên nhân dẫn đến các sai số đo có nhiều và độ lớn của giá trị này cũng rất thay đổi. Điều này dẫn tới tính bất định (uncertainty) của các kết quả phân tích. Tuy nhiên sai số đo có thể làm giảm thiểu, thậm chí trong một số trường hợp có thể loại trừ bằng nhiều phép thực nghiệm và kiểm tra. Sai số này có thể được đánh giá, tính toán dùng phương pháp thống kê và phép chemometry. Các sai số thô bạo có thể mắc phải do thiết bị hay phòng thí nghiệm kém. Việc hoàn chỉnh thiết bị đo thực nghiệm cũng như nâng cao kỹ năng phân tích cũng loại trừ được loại sai số này.

Tuy nhiên khi đọc các thông số thực nghiệm như các chỉ số trên buret, nhiệt độ trên nhiệt kế, trọng lượng mẫu trên cân phân tích hay các tín hiệu điện thế, cường độ dòng ... luôn luôn xảy ra sự biến động cố hữu của các thông số đo nếu việc đọc được lặp lại nhiều lần trong cùng điều kiện như nhau. Hơn nữa sai số có thể không phát hiện được nếu không biết giá trị thật (hay giá trị chuẩn) của mẫu phân tích.

### 2. Sự cộng hợp các sai số ngẫu nhiên và định luật lan truyền sai số

Trong phần nói về sai số hệ thống đã nêu có 3 nguồn gốc dẫn đến sai số liên quan đến kỹ năng người làm thực nghiệm, thiết bị dụng cụ đo, phương pháp và quy trình phân tích. Trong sai số ngẫu nhiên cũng bao gồm nhiều sai số bộ phận hay còn gọi là sai số riêng phần. Các sai số riêng phần quá nhỏ có thể bỏ qua không cộng hợp vào sai số toàn phần. Như đã biết, tiêu chí cơ bản

nhất để đánh giá sai số ngẫu nhiên các kết quả thực nghiệm là đo độ lệch tuyệt đối hoặc độ lệch tương đối. Để cho sai số toàn phần nhỏ, khi đo thực nghiệm cần phải tìm các điều kiện tối ưu. Có thể xác lập các điều kiện này dựa trên định luật lan truyền sai số. Nội dung của định luật này như sau: Nếu giữa n biến số không tương quan có quan hệ hàm số có dạng:

$$Y = \varphi(x_1, x_2, \dots, x_n)$$

thì sai số chung  $\sigma_y$  có thể biểu diễn gần đúng dưới dạng:

$$\sigma_y^2 = \left( \frac{\partial \varphi}{\partial x_1} \right)^2 \sigma_{x_1}^2 + \left( \frac{\partial \varphi}{\partial x_2} \right)^2 \sigma_{x_2}^2 + \dots + \left( \frac{\partial \varphi}{\partial x_n} \right)^2 \sigma_{x_n}^2 \quad (28)$$

Sai số tổng  $\sigma_y$  từ phương trình 28 lớn hơn sai số riêng phần. Sai số tổng này tăng lên khi cộng hợp các sai số riêng phần. Vì vậy thật vô nghĩa khi cho kết quả sai số riêng phần có rất nhiều số thập phân sau dấu phẩy trong khi đó bản thân các sai số riêng phần chỉ bao gồm rất ít các số thập phân có nghĩa. Người ta thống kê thấy rằng trong hoá phân tích thường gặp các dạng có liên quan đến 4 phương trình cơ bản. Nếu chỉ xét 2 biến số  $x_1$  và  $x_2$  ta có các quan hệ cơ bản dưới dạng các phương trình sau:

Công thức tính

Sai số tổng

$$y = x_1 + x_2$$

$$\sigma_y^2 = \sigma_{x_1}^2 + \sigma_{x_2}^2 \quad (28a)$$

$$y = x_1 - x_2$$

$$y = x_1 / x_2$$

$$\left( \frac{\sigma_y}{y} \right)^2 = \left( \frac{\sigma_{x_1}}{x_1} \right)^2 + \left( \frac{\sigma_{x_2}}{x_2} \right)^2 \quad 28b$$

$$y = x_1 \cdot x_2$$

Khi phát sinh sai số, người ta luôn cộng phương sai tương ứng bằng cách là đối với tổng hay hiệu thì cộng các phương sai của sai số tuyệt đối, còn đối với tích và thương lại cộng phương sai của các sai số tương đối.

Một phương pháp phân tích công cụ điển hình thường bao gồm một vài phép đo đặc thực nghiệm. Mỗi phép đo này chịu một sự bất định ngẫu nhiên của kết quả cuối cùng. Nhằm mục đích tìm hiểu xem các bất định ngẫu nhiên này ảnh hưởng như thế nào tới kết quả của phép phân tích, ta giả thiết rằng kết quả  $x$  phụ thuộc vào các biến thực nghiệm  $p, q, r, \dots$ , mỗi biến này dao động theo một cách ngẫu nhiên và độc lập. Điều này có nghĩa là  $x$  là hàm của  $p, q, r, \dots$  và ta có thể viết:

$$x = f(p, q, r, \dots) \quad (29)$$

Độ bất định  $dx_i$  (độ lệch với giá trị trung bình) ở phép đo thứ  $i$  đối với  $x$  sẽ phụ thuộc vào kích thước và dấu của các độ bất định ứng với  $dp_i, dq_i, dr_i, \dots$  và ta có thể viết:

$$dx_i = f(dp_i, dq_i, dr_i, \dots)$$

Độ biến đổi  $dx$  như là hàm của độ bất định về  $p, q, r$  có thể đánh giá bằng cách lấy tổng vi phân phương trình 28.

Như vậy:

$$dx = \left( \frac{\partial x}{\partial p} \right)_{q,r} dp + \left( \frac{\partial x}{\partial q} \right)_{p,r} dq + \left( \frac{\partial x}{\partial r} \right)_{p,q} dr + \dots \quad (30)$$

Để thiết lập mối liên hệ giữa độ lệch chuẩn của  $x$  với các độ lệch chuẩn của  $p, q, r, \dots$  cần phải lấy bình phương phương trình ở trên. Để làm như vậy, ta sẽ bỏ các kí tự liên quan đến các đạo hàm riêng phần. Như vậy:

$$(dx)^2 = \left[ \left( \frac{\partial x}{\partial p} \right) dp + \left( \frac{\partial x}{\partial q} \right) dq + \left( \frac{\partial x}{\partial r} \right) dr + \dots \right]^2 \quad (31)$$

Phương trình này sau đó phải được lấy tổng giữa các giới hạn từ  $i = 1$  đến  $i = n$  ở đây  $n$  lại là tổng số phép đo lặp lại. Khi lấy bình phương phương trình 31 có nổi lên hai số hạng ở bên phải của phương trình:

- 1) Số hạng bình phương.
- 2) Số hạng giao hoán.

Các số hạng bình phương có dạng:

$$\left( \frac{\partial x}{\partial p} \right)^2 dp^2 + \left( \frac{\partial x}{\partial q} \right)^2 dq^2 + \left( \frac{\partial x}{\partial n} \right)^2 dn^2 + \dots$$

Các số hạng bình phương luôn luôn dương và do đó không bao giờ triệt tiêu. Ngược lại, các số hạng giao hoán có thể có dấu dương hoặc âm. Ví dụ :

$$\left( \frac{\partial x}{\partial p} \right) \left( \frac{\partial x}{\partial q} \right) dp dq + \left( \frac{\partial x}{\partial p} \right) \left( \frac{\partial x}{\partial n} \right) dp dn + \dots$$

Nếu  $dp$ ,  $dq$ , và  $dr$  biểu diễn các sai số độc lập và ngẫu nhiên thì một số hạng giao hoán sẽ là âm hoặc dương. Như vậy, tổng của tất các số hạng này sẽ gần với không, đặc biệt khi  $n$  lớn.

Do khuynh hướng triệt tiêu của các số hạng giao hoán mà tổng của phương trình 31 từ  $i = 1$  đến  $i = n$  có thể được coi như là của các số hạng bình thường. Khi ấy tổng này có dạng:

$$\sum (dx)^2 = \left( \frac{\partial x}{\partial p} \right)^2 \sum (dp)^2 + \left( \frac{\partial x}{\partial q} \right)^2 \sum (dq)^2 + \left( \frac{\partial x}{\partial r} \right)^2 \sum (dr)^2 + \dots \quad (32)$$

Chia tổng này cho  $n$ , ta được:



$$\frac{\sum(dx_i)^2}{n} = \left(\frac{\partial x}{\partial p}\right)^2 \frac{\sum(dp_i)^2}{n} + \left(\frac{\partial x}{\partial q}\right)^2 \frac{\sum(dq_i)^2}{n} + \left(\frac{\partial x}{\partial r}\right)^2 \frac{\sum(dr_i)^2}{n} + \dots \quad (33)$$

Mặt khác, từ phương trình 32 ta thấy:

$$\frac{\sum(dr_i)^2}{n} = \frac{\sum(x_i - \mu)^2}{n} = \sigma_x^2$$

Ở đây,  $\sigma_x^2$  là phương sai của x. Tương tự:

$$\frac{\sum(dp_i)^2}{n} = \sigma_p^2$$

Như vậy, phương trình 33 có thể được viết theo các phương sai của các đại lượng, nghĩa là:

$$\sigma_x^2 = \left(\frac{\partial x}{\partial p}\right)^2 \sigma_p^2 + \left(\frac{\partial x}{\partial q}\right)^2 \sigma_q^2 + \left(\frac{\partial x}{\partial r}\right)^2 \sigma_r^2 + \dots \quad (34)$$

Trong hoá học phân tích ta thường gặp hiệu các giá trị đo. Từ phương trình 34 ta thấy ở đây phải sử dụng sai số tuyệt đối, sai số này chỉ tăng lên một ít và không phụ thuộc vào giá trị của hiệu số. Thường người ta thay hiệu thu được vào công thức chứa tích hay thương của các đại lượng. Trong trường hợp này người phân tích phải quan tâm đến sai số tương đối của hiệu  $\sigma_{x_1-x_2} / x_1-x_2$ . Nó càng lớn khi hiệu  $x_1-x_2$  càng nhỏ và đại lượng này trở nên rất lớn khi  $x_1$  và  $x_2$  gần bằng nhau. Vì vậy cần phải cố tránh lấy hiệu số của các giá trị đo khi các giá trị này xấp xỉ nhau về trị số.

Chẳng hạn khi xác định clo ta thu được các kết quả sau :

Chén nung với AgCl       $x = 8,3453 \text{ g}$

Chén nung không       $x_0 = 8,0875 \text{ g}$

$$\text{AgCl} \quad y=x-x_0=0,2578 \text{ g}$$

Có thể xác định sai số cân theo cách bình phương của phương pháp phân tích khối lượng. Đối với cả hai phép cân sai số này là như nhau và là  $\sigma_y=0,0002 \text{ g}$ . Từ phương trình 28a ta thu được sai số tuyệt đối của hiệu  $\sigma_y = \sigma_x \sqrt{2} \approx 0,0003 \text{ g}$ , con số này không lớn hơn bao nhiêu so với sai số của một phép cân.

Sai số tương đối của mỗi phép cân riêng bằng  $\sigma_y/x \approx 0,00003 \sim 0,003 \%$ . Ngược lại hiệu số  $y=x-x_0$  có sai số tương đối cao hơn khá rõ rệt:  $\sigma_y/y \approx 0,001$  (0,1%). Mặc dù độ chính xác của phép cân rất cao nhưng khi xác định qua hiệu số lại mắc phải sai số tương đối khá lớn 0,1%.

Khi nhân một số có sai số với một số không có sai số ( thừa số chuyển đổi tỷ lệ ) chỉ có sai số tuyệt đối tăng lên còn sai số tương đối vẫn không đổi. Đối với phép chia cũng vậy.

Trong phép xác định Niken bằng phương pháp khối lượng người ta cân 0,3124 gam Nikendiaxetylglyoxim ứng với 0,2250.  $0,3124=0,0703 \text{ g}$  Niken. Khi sai số cân  $\sigma_y=0,0003 \text{ g}$  thì sai số tuyệt đối của phép xác định qua Diaxetylglyoxim là  $0,0003 \text{ g}$  tương ứng với  $0,2250 \cdot 0,00003=0,00007 \text{ g}$  Niken, còn sai số tương đối trong cả hai trường hợp là  $0,001=0,1\%$ .

Việc xem xét các sai số dạng này trong các phương pháp phân tích hoá học chỉ cho một khái niệm về độ chính xác của phép đo. Các sai số đo thường rất nhỏ so với những sai số gây ra bởi những vi phạm trong quá trình các phản ứng hoá học. Mặc dù vậy việc bàn luận sai số như thế không nên bỏ qua vì nó cho phép người làm phân tích xác lập các điều kiện đo tối ưu. Ngoài ra trong các phép đo đặc biệt có thể thu được khái niệm về sai số tối thiểu có thể có.

Ta đi sâu xem xét đối với phương pháp phân tích bằng đo phóng xạ và đo quang:

*Phương pháp phân tích phóng xạ :*

Các phương pháp phân tích phóng xạ dựa trên việc đo số đếm các xung phóng xạ phát ra từ mẫu phân tích. Bản chất của hiện tượng phân rã phóng xạ mang tính ngẫu nhiên, do vậy ta dễ dàng áp dụng các quy luật của phân bố

ngẫu nhiên để xử lý và tính toán sai số. Người ta lặp lại nhiều lần phép đo số xung phóng xạ phát ra và từ các giá trị bằng số riêng rẽ ta tìm được giá trị trung bình  $\bar{x}$ . Để có thể so sánh các giá trị này với nhau thường người ta quy về đơn vị thời gian (trong một phút hoặc giây). Nếu trong thời gian T phút ta đo được tất cả là  $\bar{x}$  xung thì tần số xung sẽ là:

$$\nu = \frac{\bar{x}}{t} = \text{xung / phút}$$

Để xác định tần số xung có thể hoặc là thu số xung trong một đơn vị thời gian đã cho hoặc là xác định số xung rồi đo thời gian cần thiết. Phương pháp đầu gọi là phương pháp chọn thời gian không đổi, phương pháp sau gọi là chọn số xung không đổi. Dãy các đại lượng rời rạc thường hay tuân theo phân bố Poisson. Nhờ quan hệ  $\sigma \sim \sqrt{\bar{x}}$  ta có thể dễ dàng đánh giá được sai số tương đối của các sự kiện đếm được.

Ta tìm thấy:

$$\frac{\sigma}{\bar{x}} = \frac{\sqrt{\bar{x}}}{\bar{x}} = \frac{1}{\sqrt{\bar{x}}} \quad (35)$$

Chẳng hạn hãy tìm số xung tối thiểu cần đo được đối với một chế phẩm phóng xạ để có độ chính xác 1% (tương đối).

Giá trị trung bình  $\bar{x}$  thu được từ dãy các phép đo theo phương trình 35 bằng:

$$\frac{\sigma}{\bar{x}} = 0.01 = \frac{1}{\sqrt{\bar{x}}}$$

Từ đó thấy  $\bar{x} = 10.000$  thì cần phải có ít nhất 10000 để đạt được độ chính xác yêu cầu 1% tương đối.

Người ta thu được tần số xung  $\gamma$  cần tìm từ hiệu tần số của chế phẩm cộng với nền và tần số của riêng nền, do đó ta có :

$$\gamma = \gamma_p + \gamma_m \quad (36)$$

Ở đây  $\gamma_p$ - tần số xung của chế phẩm nền

$\gamma_m$  – tần số xung của nền

Nếu cả hai đại lượng  $\gamma_p$  và  $\gamma_m$  đều được xác định trong thời gian T đã cho (chọn thời gian không đổi ) thì

$$\gamma = \frac{\bar{x}p}{T} + \frac{\bar{x}m}{T} = \frac{\bar{x}}{T} \quad (37)$$

Bởi vì hai giá trị được lấy trong cùng một thời gian đếm T, nên số xung có thể dùng trực tiếp như một đại lượng cân đo. Phương trình chuyển thành :

$$\bar{x} = \bar{x}p + \bar{x}m \quad (38)$$

ở đây  $\bar{x}_p$  số xung của chế phẩm với nền

$\bar{x}_m$  số xung của nền

Từ phương trình 38 theo định luật lan truyền sai số và coi  $\sigma \approx \sqrt{\bar{x}}$  ta rút ra :

$$\sigma_x = \sqrt{\sigma_p^2 + \sigma_m^2} = \sqrt{\bar{x}p + \bar{x}_m} \quad (39)$$

Có thể cho tần số xung  $\gamma_m$  là một phân phân số của đại lượng  $\gamma_p$  :

$$\gamma_m = \xi \gamma_p \quad (\xi < 1) \quad (40)$$

Nếu đã chọn phương pháp thời gian không đổi thì:

$$\bar{x}_m = \xi \bar{x}_p$$

Thay đẳng thức này vào phương trình 39 và ta thu được sai số tương đối của phép xác định số xung theo phương pháp thời gian không đổi là:

$$\frac{\sigma_{\bar{x}}}{\bar{x}} = \frac{\sqrt{(1+\xi)}}{\sqrt{x_p(1-\xi)}} \quad (41)$$

Trong phương pháp số xung không đổi ta đã biết  $\bar{x}_p = \bar{x}_m = \bar{x}$ . Để thu được số xung đó người ta đo thời gian cần thiết  $T_p$  và  $T_m$  bằng cách dùng phương trình 35 ta tìm được tần số xung. Phương trình 36 dẫn tới:

$$v = v_p - v_m = \frac{\bar{x}_p}{T_p} - \frac{\bar{x}_m}{T_m} = \bar{x} \left( \frac{1}{T_p} - \frac{1}{T_m} \right) \quad (42)$$

Trong phương trình 42 do  $\frac{1}{T_p} \neq \frac{1}{T_m}$ . Vì vậy tính sai số theo phương trình 36

Biết rằng  $\sigma = \sqrt{\bar{x}}$  ta có:

$$\sigma_v = \sqrt{\frac{\sigma_p^2}{T_p^2} + \frac{\sigma_m^2}{T_m^2}} = \sqrt{\frac{\bar{x}_p}{T_p^2} + \frac{\bar{x}_m}{T_m^2}} \quad (43)$$

Ta lại biểu diễn tần số xung  $\gamma_m$  như phân số của  $\gamma_p$  tương tự như phương trình 40. Nếu phương pháp đã chọn là phương pháp số xung không đổi thì ta có đẳng thức:

$$\frac{1}{T_m} = \xi \frac{1}{T_p}$$

Thay đẳng thức đó vào phương trình 43 khi  $\bar{x}_i = \bar{x}_m$  ta tìm được sai số tương đối của tần số xung trong phương pháp xung không đổi:

$$\frac{\sigma_i}{\bar{x}} = \frac{\sqrt{1+\xi}}{\sqrt{\bar{x}(1-\xi)}} \quad (44)$$

Để so sánh các sai số tương đối của phương pháp thời gian không đổi và phương pháp số xung không đổi người ta tìm tương quan giữa chúng bằng cách dùng phương trình 41 và 44. Nếu ký hiệu sai số của phương pháp thời gian không đổi là  $T$ , còn sai số của phương pháp số xung không đổi là  $n$  thì khi  $\sigma_i = (\bar{x} - 1)$  ta có:

$$\frac{\left( \frac{\sigma_i}{\bar{x}} \right)}{\left( \frac{\sigma_i}{\bar{x}} \right)_n} = \sqrt{\frac{1 + \xi}{1 - \xi}} \quad (45)$$

Chỉ số này được xác định chỉ bởi tử số của số xung thu được của chế phẩm nền và của riêng nền. Sai số của phương pháp thời gian không đổi hầu như luôn luôn lớn hơn sai số của phương pháp số xung không đổi, nếu như trong cả hai phương pháp người ta tính một số xung gần như nhau khi đo độ phóng xạ của nền cũng như độ phóng xạ của chế phẩm + nền.

Khi biểu diễn phương trình 45 bằng đồ thị ta thấy rằng sai số ngẫu nhiên tăng lên.

Trong phương pháp thời gian không đổi phụ thuộc vào tỷ số của các tần số số xung đo được  $\gamma_p$  và  $\gamma_n$ . Trong trường hợp bất lợi nhất ( $\xi \sim 0,4$ ) khi  $\gamma_p = \gamma_n = \bar{x}$ , phương pháp thời gian không đổi cho sai số ngẫu nhiên (theo (45) tương đối) lớn hơn phương pháp số xung không đổi.

Để phân tích định lượng, người ta giả thiết có mối quan hệ giữa số xung  $\gamma$  và nồng độ  $C$  của mẫu. Khi sử dụng mẫu chuẩn đã biết nồng độ, ta tìm được hàm lượng của mẫu cần phân tích là:

$$C = \frac{C_E \nu_A}{\nu_E} \quad (46)$$

Ở đây  $\nu_A, \nu_E$  tần số xung của mẫu phân tích và mẫu chuẩn theo phương trình 36;  $C_E$  là hàm lượng của mẫu chuẩn.

Nếu khi đếm xung ta dùng phương pháp thời gian không đổi, ta có :

$$C = \frac{C_E \bar{x}_A}{\bar{x}_E} \quad (47)$$

Theo định luật lan truyền sai số:

$$\frac{\sigma_C}{C} = \sqrt{\left(\frac{\sigma_{C_E}}{C_E}\right)^2 + \left(\frac{\sigma_{x_A}}{\bar{x}_A}\right)^2 + \left(\frac{\sigma_{x_E}}{\bar{x}_E}\right)^2}$$

Giả thiết rằng  $\bar{x}_A \approx \bar{x}_E = \bar{x}$  ta có:

$$\frac{\sigma_C}{C} = \sqrt{\left(\frac{\sigma^2}{C_E}\right)^2 + \frac{1}{\bar{x}_E} + \frac{1}{\bar{x}_A}} \quad (48)$$

Sai số chung khi xác định nồng độ gồm sai số khi ước lượng nồng độ chuẩn và sai số khi tính  $x_A$  và  $x_E$ . Có thể giảm sai số chung bằng cách chuẩn hoá chính xác và tăng các giá trị của đại lượng cần đếm. Tuy nhiên cách này là không hoàn toàn có lợi nếu không đồng thời giảm sai số khi xác định nồng độ chuẩn  $C_E$ .

Ví dụ khác: Khi nghiên cứu hiện tượng cộng kết người ta tiến hành kết tụ bari sunfat khi có sắt phóng xạ. Ở đây người ta dùng dung dịch có chứa 22.9mg  $^{59}\text{FeCl}_2 \cdot 6\text{H}_2\text{O}$  ( $\approx 4.73 \text{ mg } ^{59}\text{Fe}^{3+}$ ), lượng muối sắt này được cân chính xác và phân tích và hoà tan trong 50 ml. Tần số xung thu được từ 5 phút là 122004 xung /phút. Nếu nồng độ của dung dịch này,  $v_E = 122004$  xung /phút. Sau khi kết tụ người ta tiến hành đo tần số xung  $v$  trong thể tích đó:  $v_A = 103410$  xung / phút. Nếu với phương trình 47 nồng độ sắt còn lại sẽ là:

$$C = \frac{4.73.103410}{122004} = 4.01 \text{ mg } ^{59}\text{Fe} / \text{ml}$$

Khi tính sai số theo công thức 48 cần chú ý là đại lượng  $C_E$  thu được từ các phép xác định khối lượng và thể tích. Vì vậy:

$$\begin{aligned} \frac{\sigma}{C} &= \sqrt{\left(\frac{0.03}{22.9}\right)^2 + \left(\frac{0.04}{50}\right)^2 + \frac{1}{122004} + \frac{1}{103410}} \\ &= \sqrt{1.71.10^{-6} + 6.4.10^{-7} + 8.2.10^{-6} + 9.8.10^{-6}} \\ &= 0.0045 = 0.45\% \text{ (tương đối)} \end{aligned}$$

Bằng cách chọn các điều kiện thích hợp, ta thu được các sai số tương đối trong mỗi phép đo có giá trị gần như nhau. Sai số chung nằm trong giới hạn của sai số ngẫu nhiên thường có trong các phương pháp cân chuẩn loại.

Các phương pháp phân tích phóng xạ hoặc các phương pháp liên quan đến các phân rã hạt nhân (phân tích hạt nhân) được ứng dụng rất rộng rãi và có thể xác định được đến hàm lượng rất nhỏ, tuy nhiên nó đòi hỏi sự biện luận rất chi tiết và chặt chẽ về sai số đo để tìm ra các điều kiện đo tối ưu nhất.



### Phương pháp phân tích đo quang ( trắc quang)

Người ta hay dùng phương pháp trắc quang để phân tích gián tiếp hỗn hợp các chất. Chẳng hạn ta phải xác định hỗn hợp nhiều cấu tử các chất hữu cơ trong vùng tử ngoại: Ta đo mật độ quang của tổng n cấu tử ở n bước sóng  $\lambda$  khác nhau. Nếu biết hệ số hấp thụ phân tử  $\varepsilon_{\lambda}$  đối với từng cấu tử tinh khiết cũng như chiều dày lớp dung dịch hấp thụ ánh sáng thì theo định luật Lambert Beer ta dễ dàng lập hệ n phương trình sau:

$$\frac{E(\lambda_1)}{l} = \varepsilon_1(\lambda_1)C_1 + \varepsilon_2(\lambda_1)C_2 + \dots \varepsilon_n(\lambda_1)C_n$$

$$\frac{E(\lambda_2)}{l} = \varepsilon_1(\lambda_2)C_1 + \varepsilon_2(\lambda_2)C_2 + \dots \varepsilon_n(\lambda_2)C_n$$

$$\frac{E(\lambda_n)}{l} = \varepsilon_1(\lambda_n)C_1 + \varepsilon_2(\lambda_n)C_2 + \dots \varepsilon_n(\lambda_n)C_n$$

Người ta xác định nồng độ cần tìm  $C_i$  bằng cách giải hệ phương trình này. Sai số khi xác định từng cấu tử riêng biệt trước hết phụ thuộc vào các sai số đo thông thường. Khi giải hệ phương trình thành phần định lượng của hỗn hợp có ý nghĩa quyết định Trường hợp lý tưởng nhất là cả n cấu tử đều có hàm lượng tương đương nhau. Hàm lượng của cấu tử càng nhỏ thì sai số đo càng lớn. Ngoài ra với cùng một bước sóng thì hệ số hấp thụ phân tử các cấu tử phải khác nhau. Phải tiến hành xác định tất cả các đại lượng cần đo (mật độ quang và hệ số hấp thụ phân tử) hết sức cẩn thận vì sai số hệ thống của một trong các đại lượng này sẽ trở thành sai số hệ thống của tất cả các nồng độ được tính ra.

Nói chung phải nói rằng trong các phương pháp phân tích gián tiếp như phương pháp đo quang, đo phóng xạ cũng như các phương pháp phân tích dụng cụ khác, sai số đối với từng cấu tử riêng biệt luôn luôn lớn hơn trong phép xác định trực tiếp. Vì vậy điều đặc biệt quan trọng phải đạt được khi phân

tích là chọn các điều kiện đo tối ưu. Bảng 6 sau đây tóm tắt về sự lan truyền sai số trong các phép tính số học.

Bảng 6. Sự lan truyền sai số

Phép tính số học	Ví dụ	Độ lệch chuẩn
Cộng và trừ	$x=p+q-r$	$s_x = \sqrt{s_p^2 + s_q^2 + s_r^2}$ (1)
Nhân và chia	$x=p.q/r$	$\frac{s_x}{x} = \sqrt{\left(\frac{s_p}{p}\right)^2 + \left(\frac{s_q}{q}\right)^2 + \left(\frac{s_r}{r}\right)^2}$ (2)
Hàm	$x=p^y$	$\frac{s_x}{x} = y \cdot \frac{s_p}{p}$ (3)
Logarit	$x=\log P$	$s_x = 0.434 \frac{s_p}{p}$ (4)
Đối logarit	$x=\text{antilog} P$	$\frac{s_x}{x} = 2.303 s_p$ (5)

\*p, q và r là các biến thực nghiệm có các độ lệch chuẩn tương ứng  $s_p$ ,  $s_q$  và  $s_r$ ; y là hằng số

### 3. Các ví dụ cụ thể

#### Ví dụ 1.

Tính sai số ngẫu nhiên đối với: a) Số liệu thứ hai của Bảng 1 và b) Giá trị trung bình đối với 3 số liệu đầu tiên trong Bảng.

Giá trị trung bình đối với toàn bộ dãy là 0,482 và do số phép đo là 50 nên có thể coi sai số ngẫu nhiên bằng 0. Như vậy giá trị  $\mu$  được chấp nhận 0,482.

a)- Ở đây, sai số ngẫu nhiên đối với phép đo duy nhất  $x_2$  là:

$$E_r = x_2 - \mu = 0,480 - 0,482 = -0,002.$$

b)- Giá trị trung bình  $\bar{x}$  đối với 3 giá trị đầu tiên trong bảng là:

$$\bar{x} = \frac{0,488 + 0,480 + 0,486}{3} = 0,485$$

Thay thế vào phương trình a1.4 ta được:

$$E_r = \bar{x} - \mu = 0,485 - 0,482 = +0,003.$$

**Ví dụ 2-** Cho các số liệu lặp lại sau đây về nồng độ  $\text{SO}_2$  trong khí quyển gần một nhà máy giấy: 1,96, 1,91, 1,88 và 1,94 phần triệu (ppm). Tính: a)- Giá trị trung bình ; b)- Độ lệch chuẩn tuyệt đối và c)- Hệ số biến sai của số liệu?

ppm $\text{SO}_2$		
	$x_i$	$x_i^2$
$x_1$	1,96	3,8416
$x_2$	1,91	3,6481
$x_3$	1,88	3,5344
$x_4$	1,94	3,7636
	$\Sigma x_i = 7,69$	$\Sigma x_i^2 = 14,7877$

a)-  $\bar{x} = 7,69/4 = 1,9225 = 1,92 \text{ ppm } \text{SO}_2$

b)- Áp dụng phương trình 18 về độ lệch chuẩn mẫu :

$$s = \sqrt{\frac{\sum x_i^2 - (\sum x_i)^2 / n}{n-1}}$$

ta có :

$$s = \sqrt{\frac{14.7877 - (7.69)^2 / 4}{4 - 1}}$$

$$= \sqrt{\frac{14.7877 - 14.784025}{3}}$$

$$= \sqrt{\frac{0.003675}{3}} = 0.035 \text{ ppmSO}_2$$

c)-Và hệ số biến sai CV:

$$CV = \frac{0,035 \text{ ppm}}{1.92 \text{ ppm}} \cdot 100\% = 1,8 \%$$

Chú ý rằng, sự khác nhau giữa  $\sum x_i^2$  và  $(\sum x_i)^2/N$  trong ví dụ 2 rất nhỏ, vì vậy nếu làm tròn vội vã thì sẽ phạm sai số nghiêm trọng khi tính giá trị của s. Do nguồn gốc này của sai số mà phương trình 18 không bao giờ được dùng để tính độ lệch chuẩn đối với các số chứa từ 5 chữ số trở lên, trong trường hợp này nên dùng phương trình 14. Điều quan trọng cũng cần phải lưu ý là các máy tính tay và các máy tính nhỏ có chức năng độ lệch chuẩn thường sử dụng một dạng của phương trình 17. Vì vậy, có thể thấy là khi dùng các loại máy tính đó để tính s cho các số liệu có từ 5 số có nghĩa trở lên sẽ mắc sai số lớn.

**Ví dụ 3-** Thuỷ ngân trong các mẫu của bảy loại cá lấy từ sông Mississippi đã được xác định bằng phương pháp hấp thụ bức xạ phát ra từ hơi thuỷ ngân nguyên tố. Tính kết quả đánh giá gộp độ lệch chuẩn của phương pháp dựa trên ba cột số liệu đầu tiên trong Bảng sau:

Số loại mẫu	Số mẫu đo	Hàm lượng Hg [ppm]	Giá trị trung bình [ppm]	Tổng bình phương độ lệch với giá trị trung bình
1	3	1,80; 1,58; 1,64	1,673	0,0259
2	4	0,96; 0,98; 1,02; 1,10	1,015	0,0115
3	2	0,13; 3,35	3,240	0,0242
4	6	2,06; 1,93; 2,12; 1,89; 1,95	2,018	0,0611
5	4	0,57; 0,58; 0,64; 0,49	0,570	0,0114
6	5	2,35; 2,44; 2,70; 2,48; 2,44	2,482	0,0685
7	4	1,11; 1,15; 1,22; 1,04	1,130	0,0170

Tổng số phép đo: 28

Tổng của tổng các bình phương: 0.2196

Các giá trị trong các cột 4 và 5 đối với loại mẫu 1 được mô tả như sau:

$x_i$	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
1,80	0,127	0,0161
1,58	0,093	0,0087
1,64	0,033	0,0011

-----  
5,02                      Tổng bình phương    = 0,0259

$$\bar{x} = \frac{5,02}{3} = 1,673$$

Các số liệu khác trong các cột 4 và 5 của Bảng cũng thu được theo cách tương tự. Như vậy:

$$s_{\text{gop}} = \sqrt{\frac{0.2196}{28-7}} = 0.102 \text{ ppmHg}$$

Để ý rằng, trong ví dụ trên mỗi một trong số 7 loại mẫu mất một bậc tự do. Tuy vậy, số bậc tự do còn lại vẫn lớn hơn 20 nên giá trị  $s$  tính được được coi là rất gần với  $\sigma$ , vì ta có thể giả thiết rằng:  $s \rightarrow \sigma$ .

#### Ví dụ 4.

Một nhà hoá học thu được số liệu sau đây đối với hàm lượng rượu của một mẫu máu, hàm lượng % của etanol: 0,084; 0,089; và 0,079. Tính giới hạn tin cậy 95% đối với giá trị trung bình, giả thiết rằng:

- a)- Không biết gì về độ chính xác của phương pháp,
- b)- Các thí nghiệm sơ bộ đã cho biết :  $s \rightarrow \sigma = 0.006\%$  etanol

$$\begin{aligned} \text{a)-} \quad \sum x_i &= 0,084 + 0,089 + 0,079 = 0,252 \\ \sum x_i^2 &= 0,0070566 + 0,007921 + 0,006241 = 0,021218 \end{aligned}$$

Ở đây:  $\bar{x} = 0,252/3 = 0,084$ . Bảng 3 cho thấy :  $t = \pm 4.03$  đối với hai bậc tự do và độ tin cậy 95%. Như vậy:

$$95\%CL = \bar{x} \pm \frac{ts}{\sqrt{n}} = 0.084 \pm \frac{4.3 \cdot 0.0050}{\sqrt{3}} = 0.084 \pm 0.012$$

b)- Vì đã biết giá trị  $\sigma$ :

$$95\%CL = \bar{x} \pm \frac{z\sigma}{\sqrt{n}} = 0.084 \pm \frac{1.96 \cdot 0.006}{\sqrt{3}} = 0.084 \pm 0.007$$

Qua ví dụ 4 ta để ý rằng việc biết chắc chắn  $\sigma$  gần như làm giảm khoảng tin cậy xuống còn một nửa.

#### 4. Kiểm tra tính sai lệch

Trong một phương pháp phân tích, độ sai lệch thường được phát hiện bằng cách phân tích một hay nhiều chất so sánh chuẩn có thành phần đã biết. Về mặt xác suất giá trị trung bình thực nghiệm  $\bar{x}$  của phép phân tích đó sẽ khác với giá trị thực  $\mu$  của mẫu chuẩn. Trong trường hợp này, cần xem xét sự sai khác nói trên là do sai số ngẫu nhiên khi phân tích chất chuẩn hay là do sai số hệ thống của phương pháp gây nên.

Cách tổng quát để xử lý vấn đề này về mặt thống kê là so sánh hiệu  $\bar{x} - \mu$  nhận được bằng thực nghiệm với hiệu có thể có ứng với một mức xác suất đã cho. nếu phương pháp phân tích không có sự lệch. Nếu giá trị thực nghiệm  $\bar{x} - \mu$  lớn hơn hiệu tính toán được thì phương pháp có sai lệch. Mặt khác, nếu giá trị thực nghiệm bằng hoặc nhỏ hơn giá trị tính thì không thể minh chứng có sai lệch.

Để kiểm tra tính sai lệch cần dùng đại lượng thống kê như đã trình bày ở trên. Ở đây, phương trình 10 được sắp xếp lại như sau:

$$\bar{x} - \mu = \pm \frac{ts}{\sqrt{n}} \quad (49)$$

Trong đó:  $n$  là số phép đo lặp lại được dùng khi kiểm tra (Nếu đánh giá tốt  $\sigma$  thì trong phương trình trên  $t$  được thay bằng  $Z$  và  $s$  bằng  $\sigma$ ). Nếu giá trị thực nghiệm của  $\bar{x} - \mu$  lớn hơn giá trị  $\bar{x} - \mu$  tính được từ phương trình 34 thì có thể khẳng định là phương pháp có sai lệch. Mặt khác, nếu giá trị tính được từ phương trình 35 lớn hơn thì không thể minh chứng là có sai lệch phương pháp.

### Ví dụ 5.

Một phương pháp mới để định lượng nhanh lưu huỳnh trong dầu hoả được kiểm tra trên một mẫu có thành phần đã biết từ phương pháp chuẩn bị mẫu chứa 0.123% S. Kết quả phân tích như sau:

$$\%S = 0,112; 0,118; 0,115; \text{ và } 0,119.$$

Các số liệu này chứng tỏ rằng phương pháp phân tích có sai lệch không?

$$\sum x = 0,112 + 0,118 + 0,115 + 0,119 = 0,464$$

$$\bar{x} = 0,464 / 4 = 0,116\%S$$

$$\bar{x} - \mu = 0,116 - 0,123 = 0,007\%S$$

$$\sum x_i^2 = 0,012544 + 0,013924 + 0,013225 + 0,014161 = 0,053854$$

$$s = \sqrt{\frac{0,053854 - (0,464)^2 / 4}{4 - 1}} = \sqrt{\frac{0,000030}{3}} = 0,0032$$

Từ Bảng 3 ta tìm thấy rằng ở mức tin cậy 95%, t có giá trị 3,18 ứng với 3 bậc tự do. Như vậy:

$$\frac{ts}{\sqrt{4}} = \frac{3,18 \cdot 0,0032}{\sqrt{4}} = \pm 0,0051$$

Giá trị trung bình thu được từ thực nghiệm dao động trong phạm vi  $\pm 0,0051$  hoặc lớn hơn không lặp lại quá 5 lần trong số 100 phép đo. Như vậy, ta có thể kết luận là  $\bar{x} - \mu = -0,007$  là một sai khác đáng kể và có sai lệch phương pháp. Nói chung ta mắc phải sai số hơi lớn hơn 5 phần trăm. Nếu ta tiến hành tính toán theo cách tương tự bằng cách dùng giá trị t ở mức tin cậy 99% thì có giá trị bằng 0,0093. Như vậy, nếu chúng ta muốn sai số không quá



1 phần trăm thì  $\frac{ts}{\sqrt{n}}$  có thể kết luận rằng sự không có sai lệch đã được minh chứng. Chú ý rằng, kết luận này khác với cách nói là không có tồn tại sai lệch.

### Ví dụ 6

Giả thiết rằng, từ thực nghiệm ta biết được phương pháp đã được sử dụng trong ví dụ 5 có độ lệch chuẩn toàn tập là 0,0032% S. Điều này có nghĩa là  $s \rightarrow \sigma = 0,0032$  với mức tin cậy 99%. Có thể khẳng định rằng phương pháp có sai lệch không?

Ở đây, ta viết:

$$\bar{x} - \mu = \pm \frac{z\sigma}{\sqrt{4}} = \pm \frac{2,58 \cdot 0,0032}{\sqrt{4}} = \pm 0,00413$$

Kết quả sai khác thực nghiệm – 0,007 là lớn hơn đáng kể so với giá trị tính được ở trên. Vì vậy, sự sai lệch là hoàn toàn khẳng định.

### Ví dụ 7.

Số đĩa n trong cột sắc ký có thể tính được theo phương trình 24-27:

$$n = 16 \left( \frac{t_R}{W} \right)^2$$

Ở đây,  $t_R$  là thời gian lưu và W là độ rộng của các pick sắc ký theo cùng đơn vị như là  $t_R$ .

Hexachlorobenzen làm hiện pick sắc ký lỏng hiệu quả cao với thời gian lưu 13,36 phút. Độ rộng của pick ở đáy là 2,18 phút. Độ chuẩn s ở hai lần đo là tương ứng bằng 0,043 và 0,061 phút. Hãy tính:

a)- Số đĩa trong cột

b)- Độ lệch chuẩn đối với kết quả tính được.

$$n = 16 \left( \frac{13.36 \text{ min}}{2.18 \text{ min}} \right)^2 = 601 \text{ đĩa}$$

Thay  $\sigma$  bằng  $s$  trong phương trình 33 ta có :

$$s_u^2 = \left( \frac{\partial u}{\partial t} \right)_{t_n}^2 s_{t_n}^2 + \left( \frac{\partial u}{\partial n} \right)_{t_n}^2 s_n^2$$

Lấy đạo hàm riêng phần phương trình gốc :

$$\left( \frac{\partial u}{\partial t} \right)_{t_n} = \frac{32t_R}{w^2} \quad \text{và} \quad \left( \frac{\partial u}{\partial n} \right)_{t_R} = \frac{-32t_R^2}{w^3}$$

Rút ra :

$$\begin{aligned} s_u^2 &= \left( \frac{32t_R}{w^2} \right)^2 s_{t_n}^2 + \left( \frac{-32t_R^2}{w^3} \right)^2 s_n^2 \\ &= \left( \frac{32 \cdot 13.36 \text{ min}}{(2.18 \text{ min})^2} \right)^2 (0.061 \text{ min})^2 + \left( \frac{-32(13.36 \text{ min})^2}{(2.18 \text{ min})^3} \right)^2 (0.043 \text{ min})^2 = 5921 \end{aligned}$$

$$s_u = \sqrt{5921} = 24.3 = 24$$

như vậy:  $n = 6.0(\pm 0.2) \times 10^2$  đĩa

## 5. Làm tròn kết quả từ các phép tính số học

Phương trình 34 tỏ ra tiện lợi khi cần làm tròn các kết quả tính toán số học. Ví dụ, xét kết quả trường hợp  $x$  được tính theo biểu thức:

$$x = p + q - r$$

Ở đây,  $p$ ,  $q$ ,  $r$  là các đại lượng thực nghiệm có các độ lệch chuẩn tương ứng  $s_p$ ,  $s_q$  và  $s_r$ . áp dụng phương trình 34 (Dùng độ lệch chuẩn mẫu thay cho độ lệch chuẩn toàn tập), ta được:

$$s_x^2 = \left( \frac{\partial x}{\partial p} \right)_{p,q,r}^2 s_p^2 + \left( \frac{\partial x}{\partial q} \right)_{p,r}^2 s_q^2 + \left( \frac{\partial x}{\partial r} \right)_{p,q}^2 s_r^2$$

nhưng vì:

$$s_x^2 = \left( \frac{\partial x}{\partial p} \right)_{q,r} + \left( \frac{\partial x}{\partial q} \right)_{p,r} = 1 \quad \text{và} \quad \left( \frac{\partial x}{\partial r} \right)_{p,q} = -1$$

bởi vậy, sự biến thiên của x được cho bởi phương trình:

$$s_x^2 = (1)^2 s_p^2 + (1)^2 s_q^2 + (-1)^2 s_r^2$$

và độ lệch chuẩn của bất quả là:

$$s_x = \sqrt{s_p^2 + s_q^2 + s_r^2}$$

Như vậy, độ lệch chuẩn tuyệt đối của tổng hoặc hiệu bằng căn bậc hai tổng bình phương độ lệch chuẩn tuyệt đối của các số lập nên tổng hoặc hiệu.

Tiến hành theo cách tương tự ta được các biểu thức có ghi trong cột 3, Bảng 6 đối với các phép tính số học khác. Chú ý rằng, trong một số phép tính, các phương sai tương đối như  $(s_x/x)^2$  và  $(s_p/p)^2$  được tổ hợp chứ không phải các độ lệch chuẩn tuyệt đối.

### Ví dụ 1.

Tính độ lệch chuẩn của kết quả sau:

$$\frac{[14.3(\pm 0.2) - 11.6(\pm 0.2)] 0.050(\pm 0.001)}{[820(\pm 10) + 1030(\pm 5)] 42.3(\pm 0.4)} = 1.725(\pm ?) \cdot 10^{-6}$$

Ở đây, các số trong dấu ngoặc là độ lệch chuẩn tuyệt đối. Trước hết ta phải tính độ lệch chuẩn của tổng và hiệu. Độ lệch chuẩn  $s_p$  đối với hiệu ở tử số được cho bởi:

$$s_p = \sqrt{(\pm 0.2)^2 + (\pm 0.2)^2} = 0.283$$

Đối với tổng ở mẫu số, độ lệch chuẩn  $s_q$  được cho bởi:

$$s_p = \sqrt{(\pm 10)^2 + (\pm 5)^2} = \pm 11.2$$

Bây giờ ta có thể viết lại phương trình như sau:

$$\frac{2.7(\pm 0.283)0.050(\pm 0.001)}{1850(\pm 11.2)42.3(\pm 0.4)} = 1.725(\pm ?)10^{-6}$$

Phương trình này bây giờ chỉ chứa các tích và thương và phương trình (2) của Bảng 6 được áp dụng:

$$\frac{s_x}{x} = \sqrt{\left(\frac{\pm 0.283}{2.7}\right)^2 + \left(\frac{\pm 0.001}{0.050}\right)^2 + \left(\frac{\pm 11.2}{1850}\right)^2 + \left(\frac{\pm 0.4}{42.3}\right)^2} = \pm 0.107$$

Để thu được độ lệch chuẩn tuyệt đối ta viết:

$$s_n = \pm 0.107x = \pm 0.107(1.725 \cdot 10^6) = \pm 0.18510^6$$

và đáp số được làm tròn  $1.7(\pm 0.2)10^{-6}$

## Ví dụ 2.

Tính độ lệch tuyệt đối của các kết quả của các phép tính sau đây. Độ lệch chuẩn tuyệt đối với mẫu đại lượng có cho trong ngoặc:

a)-  $x = \log[2.00(\pm 0.02)10^{-4}] = -3.6990 \pm ?$

b)-  $x = \text{anti log}[1200(\pm 0.00^3)] = 15.849 \pm ?$

c)-  $x = \text{anti log}[45.4(\pm 0.3)] = 2.5119 \cdot 10^{15} \pm ?$

từ phương trình(4) trong bảng 6 ta thấy

$$s_x = \pm 0,434 \cdot \frac{0,02 \cdot 10^{-4}}{2,00 \cdot 10^{-4}} = \pm 0,004$$

như vậy:  $\log[2,00(\pm 0,02) \cdot 10^4] = -3,699(\pm 0,004)$

sử dụng phương trình (4) trong bảng(6) ta có :

$$\frac{s_x}{x} = 2,303(\pm 0,003) = \pm 0,0069$$

$$s_x = \pm 0,0069x = \pm 0,0069 \cdot 18.849 = 0,109$$

Vì vậy  $\text{antilog}[1,200(\pm 0,003)] = 15,8 \pm 0,1$

$$\frac{c/s_x}{x} = 2,303(\pm 0,3) = \pm 0,691; s_x = \pm 0,691 \cdot 2.511 \cdot 10^{45} = \pm 1.7 \cdot 10^{45}$$

Do đó:  $= \text{anti log}[45.54(\pm 0,3)] = 2.5(\pm 1.7) 10^{45}$

Ví dụ 2 cho thấy sai số tuyệt đối lớn liên quan với anti logarit của một số có vài chữ số ngoài điểm thập phân. Tính bất định lớn tăng lên do ở chỗ các số nằm bên trái phần thập phân (phần đặc trưng) chỉ dùng để định vị điểm thập phân. Sai số lớn về anti logarit là do độ bất định tương đối lớn trong phần định trị của các số (Chẳng hạn:  $0,4 \pm 0,3$ ).

### **Phương pháp bình phương tối thiểu**

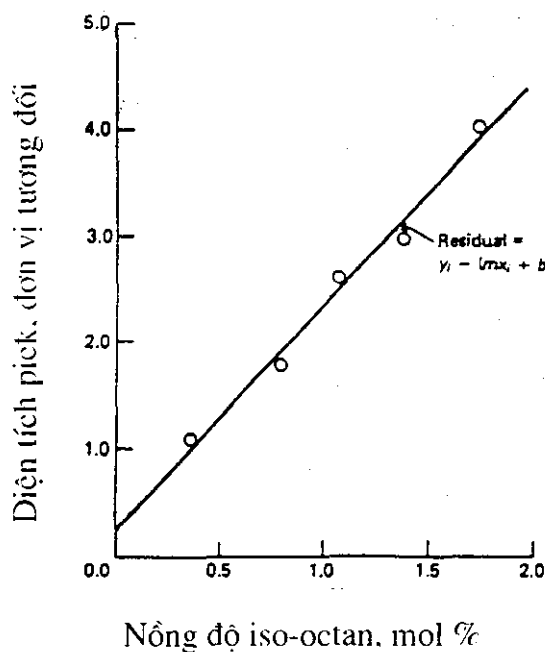
Nhiều phương pháp phân tích dựa trên đường chuẩn, ở đây đại lượng đo y được đặt như là hàm của nồng độ đã biết x của một dãy các mẫu chuẩn. Hình 6 là một đường chuẩn điển hình, được xây dựng từ phép xác định sắc ký iso-octan trong các mẫu hydrocarbon. Tung độ (biến độc lập) là diện tích dưới peak sắc ký đối với iso-octan và hoành độ (biến độc lập) là % mol iso-octan. Như là trường hợp điển hình (và mong muốn), đồ thị gần với đường thẳng. Tuy vậy, cần để ý rằng do có sai số ngẫu nhiên trong quá trình đo không phải tất cả mọi số liệu đều nằm đúng trên đường thẳng. Do đó ta phải thử dựng một đường thẳng "tốt nhất" đi qua các điểm. Cách khái quát để dựng đường thẳng như vậy được gọi là phương pháp bình phương tối thiểu. Khi áp dụng phương

pháp bình phương tối thiểu ta giả thiết rằng có một liên hệ tuyến tính giữa diện tích của pick (y) và nồng độ phân tích (x) được mô tả bằng phương trình:

$$y = mx + b.$$

Ở đây, m là độ dốc của đường thẳng và b là đoạn cắt. Chúng ta cũng giả thiết rằng mọi sự sai lệch của các điểm riêng lẻ ra khỏi đường thẳng đều là do sai số khi đo diện tích và không có sai số khi đo x, nghĩa là nồng độ của các dung dịch chuẩn đã được biết chính xác.

Độ sai lệch theo chiều thẳng đứng được gọi là phần dư (residual). Đường thẳng tạo được bằng phương pháp bình phương tối thiểu là đường thẳng làm giảm tổng bình phương của các phần dư tại các điểm đến mức thấp nhất.



Hình 8. Đường chuẩn xác định iso-octan trong các hỗn hợp hydro-cacbon

Để cho tiện, chúng tôi quy ước ba đại lượng:  $s_{xx}$ ,  $s_{yy}$ , và  $s_{xy}$  như sau:

$$s_{xx} = \sum (x_i - \bar{x})^2 = \sum x_i^2 - \frac{(\sum x_i)^2}{n} \quad (50)$$

$$s_{yy} = \sum (y_i - \bar{y})^2 = \sum y_i^2 - \frac{(\sum y_i)^2}{n} \quad (51)$$

$$s_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - \frac{\sum x_i \sum y_i}{n} \quad (52)$$

Ở đây,  $x_i$  và  $y_i$  là các toạ độ tại các điểm số liệu riêng lẻ,  $n$  là số cặp số liệu được dùng để xây dựng đường chuẩn, và  $\bar{x}$  và  $\bar{y}$  là các giá trị trung bình đối với các biến, hoặc:

$$\bar{x} = \sum x_i / n \quad \text{và} \quad \bar{y} = \sum y_i / n$$

Để ý rằng,  $s_{xx}$  và  $s_{yy}$  là tổng bình phương độ lệch với giá trị trung bình của các giá trị riêng lẻ  $x$  và  $y$ . Các biểu thức tương đương ở phần cuối bên phải của các phương trình 50, 51 và 52 được sử dụng tiện lợi khi dùng máy tính tay.

Có 6 đại lượng hay dùng được tính từ  $s_{xx}$ ,  $s_{yy}$  và  $s_{xy}$ :

$$1- \text{Độ dốc của đoạn thẳng } m: m = \frac{s_{xy}}{s_{xx}} \quad (53)$$

$$2- \text{Đoạn cắt } b: b = \bar{y} - m\bar{x} \quad (54)$$

3-Độ lệch chuẩn  $S_y$  cùng các số dư được tính theo

$$s_y = \sqrt{\frac{s_{yy} - m^2 s_{xx}}{n - 2}} \quad (55)$$

4- Độ lệch chuẩn của độ dốc  $S_m$ :

$$s_m = s_y / \sqrt{s_{xx}} \quad (56)$$

5- Độ lệch chuẩn  $s_b$  của đoạn cắt:

$$s_h = s_v = \sqrt{\frac{\sum x_i^2}{n \sum x_i^2 - (\sum x_i)^2}} = \sqrt{\frac{1}{n - (\sum x_i)^2 / \sum x_i^2}} \quad (57)$$

6- Độ lệch chuẩn  $S_c$  của kết quả phân tích thu được theo đường chuẩn:

$$s_c = \frac{s_v}{m} \sqrt{\frac{1}{2} + \frac{1}{n} + \frac{\bar{x}_s - \bar{y}}{m^2 s_{xx}}} \quad (58)$$

Phương trình 58 cho phép tính độ lệch chuẩn của giá trị trung bình  $\bar{y}_c$  của một tập gồm L phép đo lặp lại khi sử dụng đường chuẩn chứa n điểm, cần nhắc lại là  $\bar{y}$  đối với n giá trị chuẩn hoá.

*Bảng 7. Số liệu chuẩn hoá đối với phương pháp sắc kí xác định iso-octan trong hỗn hợp hydro-cácbon*

% mol của iso-octan [ $x_i$ ]	Diện tích pick [ $y_i$ ]	$x_i^2$	$y_i^2$	$x_i y_i$
0,352	1,09	0,12390	1,1881	0,38368
0,083	1,78	0,64481	3,1684	1,42964
1,08	2,60	1,16640	6,7600	2,80800
1,38	3,09	1,90440	9,1809	4,18140
1,75	4,01	3,06250	16,0801	7,01750
Cộng: 5,365	Cộng: 12,51	Cộng: 6,90201	Cộng: 36,3775	Cộng: 15,81992

### **Ví dụ 1.**

Tiến hành phân tích bình phương tối thiểu các số liệu thực nghiệm cho trong hai cột đầu tiên của Bảng 7 và vẽ trên Hình 6.



Các cột 3, 4, và 5 của Bảng chứa các giá trị tính toán đối với  $x_i^2, y_i^2$  và  $x_i y_i$  tổng của chúng được ghi ở dưới mỗi cột. Chú ý rằng, số các chữ số trong các giá trị tính toán là số cực đại ghi được trên máy tính và ta sẽ không làm tròn khi phép tính chưa kết thúc.

Bây giờ chú ý nếu thay thế vào trong các phương trình (50), (51) và (52) ta sẽ được:

$$s_{xx} = \sum x_i^2 - (\sum x_i)^2 / n = 6,90201 - (5,305)^2 / 5 = 1,14537$$

$$s_{yy} = \sum y_i^2 - (\sum y_i)^2 / n = 36,3775 - (12,51)^2 / 5 = 5,07748$$

$$s_{xy} = \sum x_i y_i - \sum x_i \sum y_i / n = 15,81992 - 5,365 \cdot 12,51 / 5 = 2,39669$$

Thay thế các đại lượng vào các phương trình (53), (54) ta có:

$$m = 2,39669 / 1,14537 = 2,095 = 2,09$$

$$b = 12,51 / 5 - 2,095 \cdot 5,365 / 5 = 0,2567 = 0,26$$

Như vậy phương trình bình phương tối thiểu sẽ là:

$$y = 2,09x + 0,26$$

Thay vào phương trình 55 ta sẽ được độ lệch chuẩn đối với số dư:

$$s_1 = \sqrt{\frac{s_{yy} - m^2 s_{xx}}{n - 2}} = \sqrt{\frac{5,07748 - (2,0925)^2 \cdot 1,14537}{5 - 2}} = 0,14$$

Và việc thay thế vào phương trình (56) cho ta độ lệch chuẩn của độ dốc:

$$s_m = \frac{s_1}{\sqrt{s_{xx}}} = 0,14 \sqrt{1,14537} = 0,13$$

Độ lệch chuẩn của đoạn cắt thu được từ phương trình a1.36. Như vậy:

$$s_b = 0,14 \sqrt{\frac{1}{5 - (5,365)^2 / 6,90201}} = 0,16$$

## Ví dụ 2.

Đường cong chuẩn rút ra từ ví dụ 1. được dùng để xác định sắc kí iso-octan trong hỗn hợp hydro-cácbon. Diện tích pick bằng 2.65. Tính phần trăm mol của iso-octan và độ lệch chuẩn của kết quả nếu diện tích là:

- a)- Kết quả của một phép đo duy nhất,
- b)- Giá trị trung bình của 4 phép đo.

Trong trường hợp khác:

$$x = \frac{7 - 0.26}{0.209} = \frac{2.65 - 0.26}{2.09} = 1.14 \text{ mol}\%$$

- a)- Thay thế vào phương trình (58), ta được:

$$s_c = \frac{0.14}{3.09} \sqrt{\frac{1}{1} + \frac{1}{5} + \frac{(2.65 - 12.51/5)^2}{(2.09)^2 \cdot 1.145}} = 0.074 \text{ mol}\%$$

- b)- Đối với giá trị trung bình của 4 phép đo:

$$s_c = \frac{0.14}{3.09} \sqrt{\frac{1}{1} + \frac{1}{5} + \frac{(2.65 - 12.51/5)^2}{(2.09)^2 \cdot 1.145}} = 0.046 \text{ mol}\%$$

## Chương V

# MẪU CHUẨN VÀ QUÁ TRÌNH HIỆU CHUẨN

### 5.1. Mẫu chuẩn (Certified Reference Materials )

#### 5.1.1. Định nghĩa Mẫu chuẩn

Độ chính xác của các kết quả phân tích có thể đạt được dựa trên việc sử dụng hệ thống đo phân tích hoàn hảo và kỹ năng thuần thục của phân tích viên cũng như tăng số phép đo lặp lại lên nhiều lần (để giảm thiểu các sai số thô bạo và ngẫu nhiên). Trong khi đó độ đúng của các kết quả phân tích chỉ có thể thu được khi loại trừ được sai số hệ thống do các chất có trong chất nền (matrix) và các yếu tố ảnh hưởng khác của chất nền gây ra. Mẫu chuẩn được dùng để loại trừ sai số này. Vậy Mẫu chuẩn là gì ? Theo định nghĩa của Cơ quan đo lường hợp pháp quốc tế (OIML ) *Mẫu chuẩn là một vật liệu hay một chất mà một hay nhiều tính chất của chúng đủ đồng nhất và được tạo nên để chuẩn hoá một dụng cụ, đánh giá một phương pháp đo hoặc để xác định giá trị đối với các vật liệu.* Các giá trị chuẩn trong mẫu chuẩn được tạo nên bởi một quy trình kỹ thuật có giá trị và hoàn hảo nhất. Giá trị chứng chỉ này phải được phê chuẩn bởi cơ quan pháp quy có thẩm quyền. Mẫu chuẩn được sản xuất ra trong các dạng khác nhau và cho các mục tiêu sử dụng khác nhau.

#### 5.1.2. Phân loại

Căn cứ vào độ chính xác có thể phân loại mẫu chuẩn thành *chuẩn đầu* và *chuẩn thứ cấp* :

- *Chuẩn đầu* (Primary Standards): Chuẩn được chỉ định hay được thừa nhận có chất lượng cao nhất và hoàn hảo nhất về mặt đo lường. Giá trị chứng chỉ của nó được công nhận không dựa vào các chuẩn khác của cùng đại lượng.

- *Chuẩn thứ cấp* (Secondary Standards): Là chuẩn mà giá trị của nó được ấn định bằng cách so sánh với chuẩn đầu của cùng đại lượng .

- *Chuẩn bậc I, II ,...* Là chuẩn mà giá trị của nó được ấn định bằng cách so sánh với chuẩn thứ cấp hoặc chuẩn có độ chính xác cao hơn có mục đích khác nhau và chúng có thể chứa một hay nhiều thành phần chuẩn

Căn cứ vào chức năng ,mục đích sử dụng có thể phân loại chuẩn thành 4 loại sau :

- *Chuẩn quốc tế*: Là chuẩn được một hiệp định quốc tế công nhận để làm cơ sở ấn định giá trị cho các chuẩn khác của đại lượng có liên quan trên phạm vi quốc tế .

- *Chuẩn quốc gia*: Là chuẩn được phê duyệt bởi một cơ quan có thẩm quyền của quốc gia có thể lấy nó làm cơ sở ấn định giá trị chuẩn cho các chuẩn khác có liên quan trong nước.

- *Chuẩn chính* : Là chuẩn có chất lượng cao nhất về mặt đo lường có thể ở một địa phương hoặc một tổ chức xác định mà các phép đo ở đó đều được dẫn xuất này.

- *Chuẩn công tác* (working Standards): Là chuẩn được dùng thường xuyên để hiệu chuẩn hoặc kiểm tra vật đo, phương tiện đo hoặc mẫu chuẩn .

Các chuẩn có một loạt những ứng dụng, như :

- Xác định giá trị của những phương pháp phân tích mới;
- Tiêu chuẩn hoá/chuẩn hoá các vật liệu chuẩn khác;
- Khẳng định giá trị của các phương pháp tiêu chuẩn;
- Trợ giúp các sơ đồ QC & QA.

Ở đây chúng tôi thấy cần làm rõ thêm về khái niệm các *chất chuẩn* (Chemical Standards ) hay *Hoá chất chuẩn*. Đây là một hoá chất tinh khiết dạng lỏng hay rắn đã biết chính xác nồng độ hay thành phần được dùng để chuẩn hoá một một thuốc thử hay hoá chất khác hoặc một dụng cụ đo. Chúng phải có thành phần rõ ràng và có độ sạch và độ ổn định cao. Phần lớn chúng đều có sẵn trên thị trường dưới nhãn hiệu AnalaR. Những chất chuẩn đầu về nguyên tắc được dùng trong chuẩn độ để xác lập độ chuẩn một hoá chất hay

thuốc thử (titrant) (Ví dụ như để tạo nên nồng độ chính xác của nó) phải được công nhận ở cấp quốc tế và phải đáp ứng được những yêu cầu sau:

- Dễ mua và dễ bảo quản ở trạng thái tinh khiết cao và có thành phần hoá học rõ ràng;
- Không hút ẩm và bền vững trong không khí để có trọng lượng chuẩn;
- Có tạp chất không vượt quá 0,02% trọng lượng;
- Dễ tan trong nước hay trong các dung môi khác;
- Phản ứng nhanh với chất phân tích trong dung dịch;
- Khác với các nguyên tố tinh khiết khác, chúng phải có phân tử lượng tương đối cao để giảm thiểu sai số cân trọng lượng.

Các chất chuẩn đầu thường được dùng trực tiếp trong các phương pháp chuẩn độ hay để chuẩn hoá các dung dịch chuẩn thứ cấp hay chuẩn công tác (secondary or working standards). Một số chuẩn cấp một dùng cho phân tích bằng phương pháp chuẩn độ được trình bày trong Bảng 8.

*Bảng 8. Một số chất chuẩn dùng trong phân tích chuẩn độ*

Phương pháp chuẩn độ	Chất chuẩn
Acid-base	Natri carbonat, $\text{Na}_2\text{CO}_3$ Natri tetrabonat, $\text{Na}_2\text{B}_4\text{O}_7 \cdot 10\text{H}_2\text{O}$ Kali Biphthalate $\text{KH}(\text{C}_8\text{H}_4\text{O}_4)$ Benzoic acid, $\text{C}_6\text{H}_5\text{COOH}$
ôxy hoá khử	Bicromat Kali, $\text{K}_2\text{Cr}_2\text{O}_7$ Kali iodate, $\text{KIO}_3$ Natri oxalate, $\text{Na}_2\text{C}_2\text{O}_4$

Kết tủa	Nitrat bạc, AgNO Natri clorua
Complexon	Kẽm , Zn Manhê , Mg EDTA

### 5.1.3. Một số đặc trưng của Mẫu chuẩn và các phương pháp thể hiện

a/ Mẫu chuẩn có các đặc trưng cơ bản:

- Các giá trị chứng chỉ: Đó là giá trị thật hoặc được công nhận là giá trị thật được phê duyệt bởi một cơ quan có thẩm quyền và được ghi trong văn bản chứng chỉ
- Độ tin cậy của các giá trị chứng chỉ: Nó được biểu diễn bằng khoảng giá trị tương ứng với một xác suất nhất định (trong thời gian MC có hiệu lực)
- Độ đồng đều.
- Độ ổn định theo thời gian: Được biểu diễn bằng khoảng thời gian trong đó giá trị thực của đại lượng vật lý của MC được giữ nguyên trong giới hạn đã định trong điều kiện bảo quản quy định

Do vậy, việc chế tạo mẫu chuẩn là một quá trình rất công phu và phải tuân thủ nghiêm ngặt các quy trình và quy định của ISO Guide30, ISO Guide 31 và quy định của tổ chức đo lường hợp pháp quốc tế OIML .

- Các thông số ảnh hưởng: Các thông số ảnh hưởng của MC được biểu diễn hoặc là bằng các giá trị số học của các yếu tố được xác định trong khi chứng chỉ hoặc bằng danh sách các yếu tố số học đặc trưng cho thành phần hoặc tính chất của MC và sự ảnh hưởng đến các kết quả đo cũng như độ không tin cậy của đại lượng được tái lập qua việc dùng MC dưới điều kiện quy định.

- Các hàm số ảnh hưởng: Các hàm số này được biểu diễn dưới dạng toán học hoặc đồ thị miêu tả sự phụ thuộc của các đặc trưng đo lường của MC vào các yếu tố ảnh hưởng ( độ ẩm ,nhiệt độ ,áp suất ,v.v... )

b/Nguyên tắc xác định các giá trị chứng chỉ của MC

-Các giá trị chứng chỉ của MC có thể được biểu diễn bằng các đại lượng có thứ nguyên hoặc không có thứ nguyên định lượng cho các tính chất của MC ( độ dẫn nhiệt ,hằng số điện môi , v.v....)hay thành phần của chúng ( thành phần hoá học , thành phần cấu trúc v.v....) .

-Các giá trị chứng chỉ của MC được xác định trong quá trình phê chuẩn . Việc xác định bao gồm công việc đo đạc ,tính toán và thao tác khác nhờ đó các giá trị chứng chỉ được thiết lập.

-Giá trị chứng chỉ có thể được xác định hoặc là cho từng MC riêng biệt, chứng chỉ riêng lẻ hoặc cho mẫu đại diện của lô MC với điều kiện là độ đồng đều của MC phải thoả mãn các yêu cầu ổn định ( chứng chỉ toàn lô ) . Trường hợp này các giá trị chứng chỉ được ấn định cho toàn bộ lô mẫu .

- Các phương pháp và dụng cụ đo để xác định giá trị chứng chỉ của MC cần được lựa chọn trên cơ sở tính đến các yếu tố như : mục tiêu và phạm vi ứng dụng MC, độ không tin cậy tối đa cho phép của các giá trị chứng chỉ. Số lượng MC cần dùng theo từng đợt và trọng lượng mỗi MC sẽ được dùng, các khía cạnh kinh tế v.v....

c/Để xác định các giá trị chứng chỉ của MC có thể sử dụng các phương tiện sau:

-Dùng MC sơ cấp , các dụng cụ đo . các phương pháp so sánh tương đối và tuyệt đối .

-Dùng MC đo lường cao cấp kết hợp với phương pháp so sánh .

- Các kết quả phân tích liên phòng thí nghiệm (sẽ đề cập kỹ hơn trong phần sau )

#### **5.1.4. Quy trình chế tạo mẫu chuẩn**

Việc sản xuất chế tạo mẫu chuẩn được tiến hành theo các bước sau :

- a. Lựa chọn vật liệu
- b. Tiến hành phân tích và đánh giá các thông số cơ bản của vật liệu chọn làm mẫu chuẩn như thành phần hoá học, các tính chất cơ lý.
- c. Chuẩn hoá thiết bị và phương tiện, phương pháp đo.
- d. Nghiền trộn đồng đều và đánh giá độ đồng đều theo các tiêu chuẩn thống kê.
- e. Đánh giá độ bền của vật liệu
- f. Tổ chức phân tích liên phòng thí nghiệm
- g. Xử lý thống kê các kết quả phân tích liên phòng thí nghiệm để thu các kết quả có độ tin cậy cao chất.
- h. Đề xuất để được cơ quan có thẩm quyền phê duyệt các giá trị chứng chỉ (Certified Values) .
- k/ Đóng gói và bảo quản theo quy định .

Trong các bước trên của quy trình sản xuất mẫu chuẩn ta cần đi sâu vào quá trình nghiền trộn đồng đều và đánh giá độ đồng đều cũng như xử lý thống kê các kết quả phân tích so sánh là những khâu quan trọng nhất .

*Nghiền trộn đồng đều* : Sau khi chọn được vật liệu làm mẫu chuẩn, nếu là quặng phải dùng máy đập nhỏ tới kích thước cần thiết rồi đem nghiền. Nghiền xong cần loại bỏ phần hạt có độ biến đổi lớn nhất về hàm lượng của cấu tử được nghiền cứu, chỉ lấy phần hạt có thành phần ít thay đổi và thích hợp với giải hàm lượng mong muốn. Đem trộn lẫn và nghiền tiếp cho tới khi đạt độ đồng đều nhất định.

Tiếp theo là *đánh giá độ đồng đều* : Độ đồng đều hay độ đồng nhất của mẫu được đánh giá qua hàm lượng của chất chính hoặc dùng đồng vị đánh dấu. Mẫu chỉ được coi là đồng đều khi độ lệch cho phép của các phép xác định của các phần lấy ngẫu nhiên trong mẫu nằm trong khoảng sai số thống kê của phương pháp phân tích được dùng, nói theo ngôn ngữ thống kê là phương sai của việc đo  $n$  mẫu  $\sigma_n$  cũng xấp xỉ với phương sai của  $n$  lần đo song



song của một mẫu. Ta có thể đánh giá độ đồng đều theo các tiêu chuẩn Student T, tiêu chuẩn Fisher F hoặc tiêu chuẩn Entropy S :

Độ đồng đều  $\delta$  theo tiêu chuẩn T được tính theo công thức:

$$\delta = \frac{2.576\sigma_n \sqrt{\frac{n+1}{n}}}{\bar{c}}$$

Ở đây  $\sigma_n$  là độ lệch chuẩn với số lần đo là n.  $\bar{c}$  là hàm lượng trung bình của n mẫu đo. Giá trị 2,576 được lấy để đảm bảo độ tin cậy là 95%.

Độ đồng đều theo tiêu chuẩn Entropy S được đánh giá như sau:

Xuất phát từ quan niệm rằng vật chất càng phân tán thì càng đồng đều, do vậy có thể dùng đại lượng S để làm thước đo cho độ đồng đều .

Ta biết rằng Entropy tăng khi hệ chuyển từ trạng thái trật tự sang hỗn loạn . Sự biến thiên S được biểu diễn theo hệ thức :

$$\Delta S_{1-2} = a \lg \frac{\Psi_1}{\Psi_2}$$

Trong đó  $\psi_1$  và  $\psi_2$  là xác suất tương ứng đối với trạng thái 1 và 2. Tuy nhiên việc đánh giá độ đồng đều qua biến thiên S chỉ thuận lợi khi phân tích được hàm lượng nhiều cấu tử trong hệ phân tán. Do vậy thay vì tính giá trị Entropy người ta đánh giá xác suất của giá trị  $\chi^2$ -đại lượng liên quan đến tính phân bố chuẩn và đặc trưng cho mức độ đồng đều của hệ :

$$\chi^2 = n \sum_{i=1}^n \frac{(w_i - p_i^2)}{p_i}$$

trong đó :

$W_i$  : tần số tương đối của khoảng hàm lượng thứ i

n: Số mẫu chọn ngẫu nhiên để phân tích

$P_i$ : Xác suất nhận được theo luật phân bố chuẩn

Có thể phân loại độ đồng đều theo xác suất P (f,..) của giá trị  $\chi^2$

Nếu xác suất P nằm trong khoảng 0.0 đến 0.1 ta có độ đồng đều rất kém. Xác suất trong khoảng 0.1 đến 0.3 ta có độ đồng đều kém, từ 0.3 đến 0.5 độ đồng đều tạm được, từ 0.5 đến 0.7 là độ đồng đều tốt và trên 0.7 tương ứng với độ đồng đều rất tốt.

Mẫu sau khi đạt được tiêu chuẩn đồng đều được gửi đi phân tích ở nhiều phòng thí nghiệm tin cậy. Các phòng thí nghiệm tham gia phân tích so sánh phải báo cáo kết quả của 3 lần xác định song song. Tất cả các giá trị trung bình  $\bar{x}_i$  của các phòng thí nghiệm được xử lý như một bộ phận các dữ kiện được đưa vào bảng (có kèm theo mã số phòng thí nghiệm).

#### *Xử lý thống kê các tập dữ liệu*

Các bước chính của quá trình đánh giá thống kê các số liệu thực nghiệm thu được như sau:

1. Giá trị trung bình của từng phòng thí nghiệm được tính toán trên cơ sở các kết quả báo cáo.
2. Tất cả các giá trị trung bình của từng PTN được xử lý như một tập dữ liệu và được sắp xếp theo thứ tự tăng dần.
3. Loại các giá trị bất thường.

Bộ các dữ liệu được xử lý trên máy vi tính để loại các giá trị bất thường. Theo kinh điển thường dùng chuẩn Q hay chuẩn Dison để loại các giá trị bất thường. Giá trị Q được tính theo biểu thức 59:

$$Q = \frac{|x - x_{n-1}|}{x_{\max} - x_{\min}} \quad (59)$$

trong đó:

$x_n$  là giá trị nghi ngờ

$x_{n+1}$  là giá trị lân cận  $x_n$

$x_{\max}, x_{\min}$  là giá trị lớn nhất và nhỏ nhất.

So sánh giá trị Q tính được: Q thực nghiệm ( $Q_m$ ) với giá trị Q lý thuyết ( $Q_l$ ) tương ứng với độ tin cậy nào đó và số kết quả thực nghiệm. Nếu  $Q_m > Q_l$  cần phải loại bỏ giá trị  $x_n$  và ngược lại. Phương pháp này thường áp dụng cho tập hợp dữ liệu n nhỏ hơn 10.

Ở đây giá trị  $h$  của điểm dữ liệu được dùng là cơ sở để xem xét dữ liệu có bị loại hay không. Giá trị  $h$  được tính theo biểu thức 60:

$$h = (x_i - x_{n-1})(S_{n-1})^{-1} \left( \frac{n}{n-1} \right)^{-1/2} \quad (60)$$

trong đó:

- $x_j$  là giá trị của điểm cần xem xét (bị loại hay không bị loại);
- $x_{n-1}$  là giá trị trung bình số học của tập hợp các giá trị trừ giá trị  $x_j$ ;
- $S_{n-1}$  là độ lệch chuẩn của tập hợp các giá trị trừ giá trị  $x_j$ ;
- $n$  là tổng các điểm dữ liệu trong tập hợp bộ số liệu.

Nếu giá trị  $h$  tìm được lớn hơn 3,162, giá trị  $x_j$  bị loại với độ không tin cậy là 0,05, và quá trình được tiếp tục với điểm dữ liệu có khoảng cách gần nhất.

Nếu  $h$  nhỏ hơn 3,162, điểm xem xét được chấp nhận thuộc tập dữ liệu và quá trình xem xét tiếp tục với điểm dữ liệu tiếp theo. Nếu  $h$  của điểm dữ liệu tiếp theo cũng nhỏ hơn 3,162 thì cả hai điểm đều được chấp nhận. Ngược lại  $h > 3,162$  thì cả hai điểm đều bị loại.

4. Giá trị trung vị (overall median) được tính theo phương pháp thông thường.
5. Để so sánh đánh giá, các giá trị trung bình số học, giới hạn tin cậy và giá trị "Mode" cũng được tính toán. Giá trị Mode là giá trị có tần suất lớn nhất trong tập dữ liệu. Ở đây giá trị Mode tính được là giá trị tương ứng với giá trị hàm mật độ địa phương LD ( Local Density ) lớn nhất.
6. Sự phân bố các dữ liệu được biểu diễn như một hàm LD của các dữ liệu thu được. Giá trị LD của dữ liệu được tính như sau:

Tất cả các giá trị được chấp nhận (trung bình PTN)  $\bar{x}_i$  được xử lý như là một tập dữ liệu và được sắp xếp theo thứ tự tăng dần. Khoảng cách giữa điểm  $\bar{x}_k$  và các điểm lân cận  $\bar{x}_{k-1}$  và  $\bar{x}_{k+1}$  có thể sử dụng như là một thước đo mật độ này. Tuy nhiên,  $\bar{x}_i$  là một biến ngẫu nhiên của một tập hợp hữu hạn dữ liệu nên có thể hình thành một "đám" ngẫu nhiên các điểm với mật độ địa phương rất cao

nhưng không phải là kết quả của sự đóng góp của toàn bộ tập dữ liệu. Sự không bình thường này được loại trừ bằng cách chính xác hoá các số liệu ban đầu. Giá trị đúng  $\bar{x}_k$  của mỗi một điểm thực nghiệm ban đầu  $\bar{x}_k$  là trung bình số học của hai điểm lân cận và giá trị chính nó.

$$\bar{x}_k = \frac{\bar{x}_{k-1} + \bar{x}_k + \bar{x}_{k+1}}{3}$$

Giá trị điểm đầu và giá trị điểm cuối không thay đổi.

$$\bar{x}_1 = \bar{x}_1 \text{ và } \bar{x}_n = \bar{x}_n$$

Khoảng cách trung bình  $a_k$  từ điểm  $\bar{x}_k$  đến các điểm lân cận  $\bar{x}_{k-1}$  và  $\bar{x}_{k+1}$  được tính như sau:

$$a_k = \frac{\bar{x}_{k-1} + \bar{x}_{k+1}}{2}$$

$$k = 2, \dots, (n-1)$$

$$a_1 = \bar{x}_2 - \bar{x}_1 \text{ và } a_n = \bar{x}_n - \bar{x}_{n-2}$$

Giá trị  $a_i$  thu được tỷ lệ nghịch với mật độ số liệu và đại lượng nghịch đảo của nó:

$$\hat{\phi}(x_i) = \frac{1}{a_i}$$

khi đó tỷ lệ thuận với mật độ địa phương của dữ liệu tại các giá trị  $x_i$ , tức là số điểm trên một đơn vị nồng độ trong vùng lân cận ngay sát điểm được biết. Bằng cách dùng qui trình “là phẳng” ta có thể loại trừ hoặc giảm thiểu các cực đại ngẫu nhiên có biên độ nhỏ. Cách tính như sau:

$$\bar{\phi}(x_i) = \frac{\bar{\phi}_{k-1} + \bar{\phi}_k + \bar{\phi}_{k+1}}{3}$$

$$\bar{\phi}_1 = \bar{\phi}_1 \text{ và } \bar{\phi}_n = \bar{\phi}_n$$

Đặt giá trị  $\bar{\phi}(x_i)$  cực đại là 100, ta thu được các giá trị hàm  $\phi(x_i)$  có dạng đồng nhất với các giá trị ban đầu và nó cho phép so sánh trực tiếp sự phân bố của các tập dữ liệu khác nhau không phụ thuộc vào số dữ liệu có trong tập dữ liệu

xem xét. Như vậy bằng cách biểu diễn qua hàm mật độ địa phương LD ta thu được đường cong biểu thị sự phụ thuộc của mật độ xác suất vào các kết quả phân tích của các phòng thí nghiệm độc lập để từ đây đề xuất các giá trị có mật độ xác suất cao nhất làm giá trị chứng chỉ cho Mẫu chuẩn .

## 5.2. Quá trình hiệu chuẩn

Trừ những phương pháp phân tích tuyệt đối liên quan đến các phản ứng hoá học có hệ số tỷ lượng (Stoichiometry) đã biết (Chẳng hạn xác định bằng phương pháp trọng lượng hay phương pháp chuẩn độ), một sự chuẩn hoá hay một quy trình tiêu chuẩn hoá được đặt ra để xác định mối liên quan giữa cường độ của một thông số hoá lý đo được của chất phân tích và một số lượng hay nồng độ của chất phân tích tương ứng với cường độ đó.

Kỹ thuật và phương pháp chuẩn hoá thường đòi hỏi là một dụng cụ, và một đầu dò (detector ) để đo các tín hiệu điện tử. Sự tác động của thành phần nền lên tín hiệu đầu dò- điều có thể ngăn cản được và được hiểu là hiệu ứng nền- là rất quan trọng. Hiệu ứng này ảnh hưởng rất nhiều đến các kết quả đo. Thành phần nền của mẫu chuẩn càng giống với chất phân tích bao nhiêu càng tốt bấy nhiêu.

Có nhiều phương pháp hiệu chuẩn hoặc chuẩn hoá. Việc chọn phương pháp thích hợp nhất phụ thuộc vào những đặc tính của kỹ thuật phân tích được sử dụng, vào đặc tính tự nhiên của mẫu và vào hàm lượng. Cụ thể bao gồm:

- Ngoại chuẩn (External standardization)- Một seri gồm ít nhất chất chuẩn đã biết nồng độ của chất cần phân tích và của các thành phần nền (nếu cần) hoặc là được chuẩn bị từ các hoá chất trong phòng thí nghiệm có độ sạch được đảm bảo (AnalaR hay mức tương đương) hoặc các fixanal thương mại đã biết chính xác nồng độ có thể sử dụng ngay. Các giá trị của tín hiệu đo thu được trên đầu dò đối với chất chuẩn trong điều kiện ổn định và thêm vào đó là mẫu trắng (blank) được chuẩn bị giống hệt như đối với mẫu phân tích ,trong mẫu trắng đó có tất cả thành phần như của mẫu phân tích trừ chất cần phân tích . Các số liệu được đưa lên đồ thị và được coi là đồ thị chuẩn hoặc được

dùng để tính hệ số chuyển đổi kết quả đo được đối với chất phân tích trong mẫu thành khối lượng hay nồng độ tương đương.

**Phương pháp thêm (Standard addition).** Đây là phương pháp chuẩn hoá nhằm tránh ảnh hưởng của chất nền bằng cách đo mẫu phân tích song song với việc đo mẫu thật có thêm một lượng đã biết của chất chuẩn.

**Phương pháp nội chuẩn (Internal standardization).** Ở đây chất nội chuẩn được thêm vào mẫu phân tích và mẫu chuẩn sau đó đo tỷ lệ tín hiệu của hai đại lượng trên và lập đường chuẩn theo nồng độ chất phân tích.

Các dụng cụ và thiết bị dùng cho công việc phân tích phải được bảo dưỡng thật chuẩn và chuẩn hoá theo các giá trị chuẩn để bảo đảm rằng các số đo được là chính xác và thực tế. Các thao tác phải được kiểm tra đều đặn, kết quả ghi chép phải được giữ gìn cẩn thận, như vậy mọi sai lệch đều được phát hiện và sửa chữa nhanh chóng. Các thiết bị kiểm tra như microcomputer và microprocessor thường có bộ phận kiểm tra thao tác, được lắp đặt bên trong máy và thường tự động làm việc mỗi khi thiết bị được đóng điện.

Sau đây là một vài ví dụ về chuẩn hoá thiết bị và dụng cụ:

- Chuẩn hoá bằng tay cân điện tử bằng các quả cân chuẩn;
- Chuẩn hoá các dụng cụ thuỷ tinh bằng cách cân khối lượng nước sạch;
- Chuẩn hoá các vạch hấp phụ hay bước sóng của máy quang phổ bằng các đặc tính hấp phụ hay phát xạ chuẩn;
- Chuẩn hoá các thang nhiệt độ, các thang điện thế hoặc các thang vạch khác bằng thiết bị đo chuẩn.

Việc hiệu chuẩn hoặc chuẩn hoá các thông số liên quan đến phân tích hoá học và kiểm tra vật liệu phải được đặc biệt lưu ý, bởi vì nếu sao nhãng hay bỏ qua những nguyên tắc cơ bản trong phương pháp luận áp dụng trong phân tích sẽ phạm phải những sai số lớn. Dưới đây sẽ trình bày một số những chỉ

dẫn cơ bản cho những người làm phân tích, kể cả những người làm việc trong phòng thí nghiệm và những người đánh giá kết quả.

Mỗi một phép đo, đặc biệt mỗi một phép phân tích định lượng đều phải sử dụng các chất chuẩn hoặc mẫu chuẩn để kiểm tra và hiệu chuẩn. Đây là điều kiện cần thiết bảo đảm độ chính xác của các kết quả.

Chất lượng phép đo chuẩn hoá phụ thuộc vào:

- Tính bất định nội tại của mẫu chuẩn (một bộ khối lượng chuẩn hoá, dung dịch chuẩn độ, hỗn hợp khí, thành phần của MC, v.v...)

- Độ tương thích (hay độ phù hợp với mục tiêu) của mẫu chuẩn trong điều kiện thực tế sử dụng, kể cả phương pháp phân tích được sử dụng và mẫu phân tích).

Tính bất định tổng của chuẩn hoá bao hàm cả hai yếu tố trên và nó phải được tối ưu hoá mà không được xem nhẹ bất kỳ một yếu tố nào.

Người phân tích phải biết so sánh tính bất định của chuẩn hoá với những bất định phân tích tổng (Thường vẫn được nhất trí giữa người phân tích và khách hàng). Sự so sánh này sẽ cho ta những chỉ dẫn cần thiết để chọn lựa giữa những trình tự chuẩn hoá khác nhau đã có và xa hơn nữa, là để cải tiến phương pháp và các quy trình.

Trong các phép thử dựa trên các phép đo định lượng, nguyên tắc của tính liên kết chuẩn (traceability) của các chất chuẩn và của các dụng cụ đo của một phòng thí nghiệm có uy tín theo tiêu chuẩn cơ bản quốc gia thông qua hệ thống tiêu chuẩn hoá quốc gia, vẫn thường được áp dụng.

### 5.3. Lựa chọn mẫu chuẩn

Tiếp cận ban đầu là xem và so sánh các đặc trưng của mẫu chuẩn cần tìm trong danh mục các mẫu chuẩn (CRMs) có trên thị trường. Người dùng cần tham khảo:

- Các Catalogue của các nhà sản xuất,
- Ngân hàng dữ liệu về mẫu chuẩn COMAR (COMAR data bank),

- Các ấn phẩm hay các bài giới thiệu, kiểm tra những chọn lựa tốt nhất về CRM trong lĩnh vực riêng nếu có trên thị trường.

Các phòng thí nghiệm cần phải nắm chắc rằng các CRM đã qua các thử nghiệm bước đầu. Đó là :

- Các nguyên tố cần phân tích đã được chứng chỉ ( certified )
- Quy trình chứng chỉ phải thể hiện ở mức tương đối độ tin cậy của phép đo (xem ISO Guide REMCO 35) và chúng phải được thể hiện đầy đủ trong tài liệu (ISO Guide 31). Mọi tính liên kết chuẩn ( traceability ) mới chỉ được xác định bằng nguyên tắc mà chưa có đánh giá về độ bất định thì sẽ không có giá trị pháp lý .

Về chất nền (matrix) của mẫu chuẩn , phòng thí nghiệm phải cân nhắc một thực tế là cả về mặt kinh tế lẫn kỹ thuật là không thể có sự đồng nhất hoàn toàn về chất nền giữa MC và mẫu phân tích. Một sự đồng nhất hợp lý phải được chấp nhận. Nếu không, toàn bộ quy trình phân tích phải được xem xét lại.

Việc sử dụng một MC có sẵn trên thị trường thường đảm bảo cho:

- Có được độ đúng tốt nhất
- Có được tỷ lệ: thực hành/chi phí tốt nhất.

Một phòng thí nghiệm khi đã quyết định không sử dụng MC tương ứng có sẵn thì phải giải thích lý do cho quyết định của mình theo trình tự của nó.

Các quy định chung khi sử dụng MC được nêu trong văn bản kỹ thuật số 18 của Tổ chức Đo lường hợp pháp quốc tế ( N<sup>o</sup> 18-OIML ), ở đây cần lưu ý một số điểm sau :

- Sự hiệu chỉnh một phần trọn vẹn của phân tích và chi phí chi cho nó phải là một tỷ lệ hợp lý và đầy đủ cho phân tích . Điều đó phải được lên kế hoạch và phải được cung cấp, đặc biệt khi phải mua MC hoặc hoàn thiện MC nội bộ. Nếu dự tính thiếu các khoản chi phí trên đây thì sẽ không đảm bảo quy trình chuẩn hoá một cách đầy đủ được .

Việc chuẩn hoá các quá trình phân tích hoá học phải đáp ứng được một số yêu cầu cần thiết, như đã nêu ở trên .Việc đáp ứng những yêu cầu trên có



thể ở những dạng khác nhau trong những lĩnh vực khác nhau. Những giải trình trên đây chưa phải là những điều kiện đầy đủ về chất lượng trong việc chuẩn hoá. Mỗi người sử dụng cần phải:

- Nêu lên những điều kiện cụ thể riêng của mình,
- Đưa ra những ngoại lệ so với những quy định chung.

Trong mọi trường hợp, người sử dụng phải xác định và phân tích những nhu cầu của mình trong các phương diện khác nhau, đồng thời rút ra và thực hiện những đáp ứng cụ thể cho mỗi trường hợp.

## CÂU HỎI VÀ BÀI TẬP

1. Người ta cân 5 lần một vật cân với các kết quả như sau: 9,990; 9,986; 9,973 và 9,980 gam. Tính độ lệch chuẩn  $S$  của các kết quả trên.
2. Hàm lượng đồng trong một hợp kim được xác định bằng phương pháp so màu với 3 kết quả song song như sau: 0,084; 0,089; và 0,079%  
Tính khoảng tin cậy của giá trị trung bình với xác suất 95%
3. Các kết quả xác định hàm lượng sắt tổng số trong một mẫu đất là như sau: 0,112; 0,118; 0,115; và 0,119%. Hàm lượng sắt được coi là thực (được chứng chỉ) trong mẫu trên là 0,123% .Hãy xét xem phép xác định có mắc phải sai số hệ thống hay không.
4. Một mẫu A có hàm lượng nước 1,31% được phân tích 3 lần với các kết quả là 1,28; 1,26; và 1,29% . Một mẫu B khác có hàm lượng nước 8,67% cũng được phân tích 3 lần với các kết quả là 8,48; 8,53; và 8,50% . Hãy so sánh :
  - a. Độ lệch tuyệt đối và tương đối của giá trị trung bình của hai dãy phân tích trên.
  - b. Sai số tuyệt đối và tương đối của giá trị trung bình của B
5. Tính khoảng tin cậy của giá trị trung bình đối với kết quả 3 lần đo một mẫu phân tích là 7,74; 6,98; và 7,27% với xác suất 90%
6. Kết quả xác định hàm lượng Can xi trong máu bằng 3 thí nghiệm là 3,15; 3,25; và 3,26. Tính khoảng tin cậy đối với giá trị trung bình tương ứng xác suất 95%

7. Khi phân tích một mẫu chuẩn có hàm lượng axeton 10,3% người ta nhận được 3 kết quả sau : 10,2; 9,9 và 10,3% axeton .Một mẫu chuẩn khác chứa 0,40% axeton cho các kết quả 0,38; 0,34; và 0,35% axeton. Hãy so sánh :

- Độ lặp lại của hai phép phân tích, về đại lượng các độ lệch tương đối và tuyệt đối so với các giá trị trung bình
- Các sai số của phép phân tích về giá trị tuyệt đối và tương đối

8. Xét các tập con số liệu sau đây:

A	B	C	D
61,45	3,27	12,06	2,7
61,53	3,26	12,14	2,4
61,32	3,24		2,6
	3,28		2,9
	3,23		

Tính:

- Giá trị trung bình mỗi tập con số liệu và xác định các bậc tự do liên quan đến diện tích  $\bar{x}$ ;
- Độ lệch chuẩn tuyệt đối của mỗi tập con và hãy xác định các bậc tự do liên quan với việc tính  $s$ ;
- Sai số chuẩn của giá trị trung bình của mỗi tập con;
- Hệ số biến sai đối với mỗi điểm số liệu riêng lẻ so với giá trị trung bình.

9. Hãy tính S dựa theo các số liệu tổ hợp của phép xác định phổ trắc quang của axit nitriloaxetic (NTA ) trong các mẫu nước sông:

Số mẫu	NTA, $\mu\text{g/ml}$
1	13 ; 17 ; 14,9
2	38 ; 37 ; 38
3	25 ; 29 ; 23 ; 29 ; 26

10. Giá trị nhận được đối với mỗi tập số liệu trong Bảng của bài tập 13 lần lượt là:

A: 61,71, B: 3,28, C: 12,23 và D: 2,75. Tính:

a) Sai số tuyệt đối đối với giá trị trung bình của mỗi tập con:

b) Sai số tương đối tính theo phần trăm đối với mỗi giá trị trung bình.

11. Khi xác định Fe trong đuôi thải bằng phương pháp hấp thụ nguyên tử người ta tìm thấy độ lệch chuẩn  $S \rightarrow \sigma = 2,4 \mu\text{g/ml}$  theo các số liệu tổ hợp của 30 phép phân tích ( trong mỗi phép phân tích có 3 phép song song ). Hãy tính khoảng tin cậy đối với kết quả  $18,5 \mu\text{g/ml}$  Fe với xác suất tin cậy 80 và 95 % bằng cách sử dụng:

a. Phép xác định riêng

b. Giá trị trung bình từ hai phép xác định

c. Giá trị trung bình từ bốn phép xác định

12. Phương pháp riêng để phân tích đồng cho ta kết quả bị thấp đi 0,5 mg. Hãy tính sai số tương đối theo % do người này gây ra nếu khối lượng đồng trong mẫu lần lượt là:

a) 25 mg; b) 100 mg; c) 250 mg; d) 500 mg.

13. Người ta tìm thấy rằng phương pháp xác định Br trong các hợp chất hữu cơ có sai số hằng định 0,20 mg Br . Hãy tính sai số tương đối (%) của kết quả phân tích mẫu chứa gần 10% BR nếu đã lấy các khối lượng sau :  
a/ 10mg ; b/ 50 mg ; c/ 100mg ; d/ 500mg d/ 1000mg

14. Phương pháp mô tả trong một bài tập được sử dụng để phân tích một loại quặng chứa cỡ 4,8% đồng. Tính khối lượng mẫu tối thiểu cần phải lấy nếu sai số tương đối do việc mất 0,5 mg phải nhỏ hơn:

a) 0,1% ; b) 0,5% ; c) 0,8% và d) 1,2%.

15. Một nhà hoá học đã nhận được các kết quả sau của ba phép xác định song song hàm lượng lindan trong mẫu thuốc trừ sâu: 7,47; 6,98; 7,27. Hãy tính khoảng tin cậy đối với giá trị trung bình từ ba số liệu với xác suất tin cậy 90% giả thiết rằng:
- các số liệu của 3 phép đo là thông tin duy nhất để đánh giá độ lặp lại của phương pháp
  - Trên cơ sở các thực nghiệm sơ bộ ,nhà hoá học giả thiết rằng  $S \rightarrow \sigma = 0,28\%$  lindan
16. Một công cụ có độ lệch chuẩn 0,1%. Hãy tính số phép đo lặp lại cần thiết phải có để sai số chuẩn của giá trị trung bình là 0,01%?
17. Khi xác định lưu huỳnh trong mẫu dầu hoả nhiễm bẩn người ta thu được các kết quả sau : 0,742; 0,693; và 0,755 %. Hãy tính khoảng tin cậy đối với giá trị trung bình của phép phân tích này với xác suất tin cậy 95%
18. Một phương pháp đo đã biết cho giá trị trung bình 0,500 và độ lệch chuẩn là  $1,84 \times 10^{-3}$ . Biết rằng thống kê Gauss áp dụng được. Hãy tính số phép đo lặp lại cần thiết nếu sai số chuẩn của giá trị trung bình không vượt quá 0,100%?
19. Hàm lượng Ca trong huyết thanh của một bệnh nhân được xác định với 3 kết quả như sau : 3,15 ; 3,25 ; 3,26 mili đlg /l .Hãy xác định khoảng tin cậy đối với giá trị trung bình của các kết quả trên với xác suất tin cậy 95% nếu giả thiết rằng:
- Thông tin sơ bộ về độ lặp lại của phép phân tích không có
  - $S \rightarrow \sigma = 0,05$  mili đlg/l

20. Trong một phương pháp riêng để xác định crom trong các mẫu địa chất thường xuyên có độ hao hụt độ tan vào cỡ 1.8 mg. Người ta dùng phương pháp này để phân tích một mẫu chứa khoảng 18% Cr. Hãy dự đoán sai số tương đối (theo phần nghìn) của kết quả thu được, do nguồn sai số của hệ thống này, nếu mẫu lấy để phân tích có khối lượng 0.400 mg?
21. Để kiểm tra phép chuẩn độ xác định Ca trong đá vôi người ta tiến hành phân tích mẫu chuẩn đá vôi chứa 30,15% CaO. Kết quả trung bình của 4 phép phân tích bằng 30,26% CaO với độ lệch chuẩn 0,085% CaO. Bằng cách tổ hợp các số liệu của một số phép phân tích người ta tìm ra rằng  $S \rightarrow \sigma = 0.094\% \text{ CaO}$ . Hãy cho biết:
- Các số liệu có chỉ ra được sự có mặt của sai số hệ thống ở xác suất tin cậy 95% hay không
  - Hãy lập lại các phép tính nếu giả thiết rằng giá trị tổ hợp  $\sigma$  không có
22. Các số liệu sau đây do việc nghiên cứu liên tục ion can xi trong huyết tương của một số đối tượng:

Đối tượng	Hàm lượng TB của Ca	Số phép đo	Độ lệch của các kết quả riêng lẻ với giá trị trung bình
1	3.16	5	0,14; 0,09; 0,06; 0,00; 0,11
2	4.08	4	0,07; 0,12; 0,10; 0,01
3	3.75	5	0,13; 0,05; 0,08; 0,14; 0,07
4	3.49	3	0,10; 0,13; 0,07
5	3.32	6	0,07; 0,10; 0,11; 0,03; 0,04; 0,05

- Tính s đối với mỗi tập hợp giá trị?
- Gộp các số liệu lại và tính s của phương pháp phân tích?

23. Phương pháp xác định hàm lượng riêng của chì trong các mẫu khí quyển dựa trên đại lượng đo được của không khí đi qua tấm lọc và tiến hành phân tích các vòng cắt từ tấm lọc. Tính các giá trị riêng lẻ cũng như giá trị gộp của s đối với các số liệu sau đây:

Mẫu	$\mu\text{g Pb/m}^3$ không khí
1	1,5; 1,2; 1,3
2	2,0; 2,3; 2,3; 2,2
3	1,8; 1,7; 1,4; 1,6
4	1,6; 1,3; 1,2; 1,5; 1,6

24. ánh giá độ lệch chuẩn tuyệt đối và hệ số biến sai đối với các kết quả của các phép tính sau đây. Cần làm tròn kết quả sao cho nó chỉ chứa một số các số có nghĩa. Các số trong dấu ngoặc là các độ lệch chuẩn tuyệt đối:

- a)-  $x = \lg 878(\pm 4) = 2,94349$   
 b)-  $x = \lg 0,4957(\pm 0,0004) = -0,30478$   
 c)-  $p = \text{antilogarit } 3,64(\pm 0,01) = 4,36516$   
 d)-  $p = \text{antilogarit } -7,191(\pm 0,002) = 6,44169 \cdot 10^{-8}$

25. Sự khác nhau ở xác suất tin cậy 95% giữa các giá trị trung bình thực nghiệm  $x_1$  và  $x_2$  dẫn ra trong các trường hợp dưới đây có nghĩa không? Độ lệch tổ hợp được ký hiệu là S, độ lệch chuẩn được ký hiệu là  $\sigma$ .

	$x_1$	$x_2$	$n_1$	$n_2$	Độ lệch chuẩn
a/	14,1	14,5	2	3	$S = 0,2$
b/	14,1	14,5	6	8	$S = 0,2$
c/	14,1	14,5	2	3	$S = 0,2$

26.Đánh giá độ lệch chuẩn tuyệt đối và hệ số biến sai đối với các kết quả của các phép tính sau đây. Hãy làm tròn mỗi kết quả sao cho chúng chỉ gồm các số có nghĩa. Các số trong dấu ngoặc là các độ lệch chuẩn tuyệt đối.

a)-  $y = 6.75(\pm 0.03) + 0.843(\pm 0.001) - 7.0121(\pm 0.001) = 0.572$

b)-  $y = 67.1(\pm 0.3).1.03(\pm 0.02).10^{-17} = 6.9113.10^{-16}$

c)-  $y = 243(\pm 1) \cdot \frac{760(\pm 2)}{1.006(\pm 0.006)} = 183578.5$

d)  $y = \frac{143(\pm 6) - 64 \pm 3}{1249(\pm 1) + 77(\pm 8)} = 5.9578.10^{-2}$

e)-  $y = \frac{1.97(\pm 0.01)}{243(\pm 3)} = 8.106996.10^{-3}$

27.Kết quả xác định nguyên tử lượng của cacbon là 12.0112; 12.0210; 12.0102; 12.0118; 12.0111; 12.0106; 12.0113; 12.0101; 12.00970; 12.0095; 12.0080

a.Tính giá trị trung bình

b.Tính giá trị tuyệt đối của độ lệch chuẩn

c.Biểu diễn nguyên tử lượng của cacbon bằng cận tin cậy với xác suất 95%

d.Kết quả của thí nghiệm thêm là 12.0138 .Nên bỏ hay giữ lại kết quả này ?

28.Biết rằng độ lệch chuẩn đối với một phương pháp phân tích vàng trong nước biển là 0,025 ppb. Hãy tính giới hạn tin cậy 99% đối với phép phân tích bằng phương pháp đã nêu, dựa trên:

a)- Một phép đo duy nhất;

b)- Ba phép đo;

c)- Năm phép đo?

29.Một phương pháp đã được xây dựng để xác định các dẫn xuất clo của các hydro cacbon trong các mẫu khí quyển có độ lệch chuẩn là 0,030 ppm:

a)-Tính giới hạn tin cậy 95% đối với giá trị trung bình của 4 phép đo thu được bằng phương pháp trên?



b)- Có bao nhiêu phép đo phải thực hiện nếu giới hạn tin cậy 95% là  $\pm 0,017$ ?

30. Hãy đánh giá độ lệch chuẩn tuyệt đối của các kết quả được dẫn ra dưới đây .Hãy làm tròn kết quả cho đến con số thích hợp.các con số có nghĩa.

Trong ngoặc có chỉ ra các độ lệch chuẩn tuyệt đối .

a/  $y = 6.75 ( \pm 0,03 ) + 0.834( \pm 0,001 ) - 7.021 ( \pm 0,001 ) = 0.572$

b/  $y = 19.97 ( \pm 0,04 ) + 0.0030( \pm 0,0001 ) + 1.29 ( \pm 0,008 ) = 21.263$

c/  $y = 67.1 ( \pm 0,3 ) . 1.03 ( \pm 0,02 ) . 10^{-17} = 6,9113.10^{-16}$

31. Độ lệch chuẩn trong phương pháp phân tích cacbon monoxit trong khí thải ô tô dựa trên các thí nghiệm đã được tiến hành rộng rãi là 0,80 ppm:

a)- Đánh giá giới hạn tin cậy 90% đối với 3 phép phân tích?

b)- Cần có bao nhiêu phép đo để đạt tới giới hạn tin cậy 90% là 0.50ppm?

32. Hàm lượng phần trăm theo lý lịch của Niken trong một mẫu thép chuẩn đặc biệt của NIST là 1,12%. Kết quả phân tích bằng một phương pháp phổ trắc quang mới cho ta hàm lượng phần trăm như sau: 1,10; 1,08; 1,09; 1,10; 1,12; 1,09 với mức tin cậy 95% có thể kết luận là phương pháp có sai lệch không?

33. Có bao nhiêu con số có nghĩa trong các số sau :

a/0.0607

b/9966

c/0.0003644

d/735702

e/9004

g/31814

34. Để giải quyết vấn đề : Lấy hay loại bỏ kết quả ra khỏi các tập dẫn ra dưới đây .hãy dùng chuẩn Q:

a/ 41,37 ; 41,61 ; 41,84 ; 41,70

b/ 7,300 ; 7,284 ; 7,388 ; 7,292

35. Tính hàm lượng phần trăm của Cu trong quặng từ những dữ kiện sau đây với con số có nghĩa cần thiết. 1.000g quặng được chuẩn độ bằng 4.000ml Natri thiosunfat 0.1000 N với buret 50ml. Cu = 63,54

36. Một phương pháp chuẩn độ để xác định can xi trong limestone đã được kiểm tra bằng cách phân tích một mẫu limestone NIST chứa 30,15% CaO. Giá trị trung bình của bốn lần phân tích là 30,26% CaO với độ lệch chuẩn là 0,085%. Bằng cách gộp số liệu từ một số lần phân tích người ta xác định được  $s \rightarrow \sigma = 0,094\%$  CaO.

— Các số liệu có cho thấy có sai số hệ thống ở mức tin cậy 95% không?

— Các số liệu có cho thấy có sai số hệ thống ở mức tin cậy 95% không nếu không cho giá trị gộp về  $\sigma$ ?

37. Để kiểm tra chất lượng công việc của một phòng thí nghiệm thương mại, người ta tiến hành phân tích hai lần axit benzoic tinh chế (68,8% C; 4,953% H). Giả thiết rằng độ lệch chuẩn tương đối của phương pháp là  $s \rightarrow \sigma = 4$  ppt đối với cac bon và 6 ppt đối với Hydro. Các giá trị trung bình của kết quả thu được là 68,5% C và 4,882 H. Với mức tin cậy 95% có thể coi có sai số hệ thống trong các phép phân tích trên hay không?

38. Đường kính của một quả cầu đo được là 2,15 cm và độ lệch chuẩn giá trị trung bình là 0,02 cm. Hãy tính giá trị tốt nhất của thể tích quả cầu và tìm độ lệch chuẩn thể tích đó?

39. Tích số tan  $T_{AgX}$  của muối bạc AgX bằng  $4,0 (\pm 0,4) \cdot 10^{-5}$ . Tính sai số khi tính độ tan của AgX trong nước?

40. Có thể đọc trên một pH kế với độ lệch chuẩn  $\pm 0,01$  đơn vị pH trong phạm vi pH từ 2 – 12. Tính độ lệch chuẩn của  $[H_3O^+]$  tại mỗi điểm cuối của khoảng đo đó?

41. Một dung dịch được pha chế bằng cách cân x g chất X hoà tan trong bình định mức 100 ml. Cân có độ chính xác 0,2 mg và được coi như là độ lệch chuẩn và bình đó có thể đổ đầy với độ chính xác 0,15 ml và cũng được coi như là độ lệch chuẩn. Tính độ lệch chuẩn của nồng độ theo g/ml?

42. Đánh giá độ lệch chuẩn tuyệt đối của kết quả thu được từ các phép tính sau (Các số trong dấu ngoặc là các độ lệch chuẩn tuyệt đối đối với các số đã cho). Hãy biểu diễn các kết quả với các số có nghĩa thích hợp:

a)  $x = \lg 878 (\pm 4) = 2.94349$

b)  $x = \lg 0,4957 (\pm 0,0004) = -0,30478$

c)  $p = \text{antilogarit } 3,64 (\pm 0,01) = 4365,16$

d)  $p^* = \text{antilogarit } -7,191 (\pm 0,002) = 6,44169 \times 10^{-8}$

43. Nồng độ ion sulphat trong nước tự nhiên có thể được xác định bằng cách đo độ đục sinh ra khi cho lượng đủ  $BaCl_2$  vào một lượng xác định của

10. c. Đục kế là dụng cụ phân tích được chuẩn hoá với một dãy các dung dịch  $\text{Na}_2\text{SO}_4$  chuẩn. Các số liệu sau đây thu được khi chuẩn hoá:

$\text{SO}_4^{2-}$ (l. C <sub>s</sub> )	Chỉ số trên đục kế, R
0,00	0,06
5,00	1,48
10,00	2,28
15,00	3,98
20,00	4,61

- Giả thiết có sự phụ thuộc tuyến tính giữa chỉ số đọc trên máy và nồng độ ion  $\text{SO}_4^{2-}$ , vẽ các số liệu trên lên đồ thị và kẻ đường thẳng đi qua các điểm (điểm gốc cũng như một)?
- Viết phương trình bình phương tối thiểu đối với sự phụ thuộc giữa các  $\text{SO}_4^{2-}$  và R.
- Viết ra các đường thẳng suy ra từ sự phụ thuộc ở (b) với đường thẳng đi qua gốc (nếu có)?
- Tính độ lệch chuẩn đối với độ dốc và đoạn cắt đối với đường bình phương tối thiểu.
- Giả thiết nồng độ ion sulphat trong mẫu có số đọc trên đục kế là 3,67? Tính độ lệch chuẩn tuyệt đối của kết quả và hệ số biến phân.
- Viết công thức phép tính trong c) giả thiết rằng 3,67 là giá trị trung bình của 6 lần đo trên đục kế?

11. Các số liệu sau đây thu được khi chuẩn hoá điện cực can xi để xác định nồng độ  $\text{Ca}^{2+}$ . Liên hệ tuyến tính giữa thế E và pCa như sau:

pCa	E, mV
5,00	+53,8

4,00	-27,7
3,00	+2,7
2,00	+31,9
1,00	+65,1

a)- Đưa các số liệu lên đồ thị và kẻ đường thẳng đi qua các điểm (Ước lượng bằng mắt)?

b)- Lập biểu thức bình phương tối thiểu đối với đường thẳng độc nhất qua các điểm. Đưa lên đồ thị?

=====00000=====

## ĐÁP ÁN

1. $x_i$	$(x_i - \bar{x})$	$(x_i - \bar{x})^2$
9,990	$7,6 \cdot 10^{-3}$	$57,8 \cdot 10^{-6}$
9,986	$3,6 \cdot 10^{-3}$	$13,0 \cdot 10^{-6}$
9,973	$9,4 \cdot 10^{-3}$	$88,4 \cdot 10^{-6}$
9,983	$0,6 \cdot 10^{-3}$	
9,980	$2,4 \cdot 10^{-3}$	

$$\bar{x} = 9,9824 \quad \sum (x_i - \bar{x})^2 = 165,4 \cdot 10^{-6}$$

$$s = \sqrt{\frac{165,4 \times 10^{-6}}{4}} = 6,4 \times 10^{-3}$$

2.  $\bar{x} = 0,084$

$$S = 0,0050$$

Tra bảng Student ở bậc tự do là 2 ta có  $t = 4,30$

Vậy khoảng tin cậy với xác suất 95% là :

$$0,084 \pm \frac{t \cdot s}{\sqrt{n}} = 0,084 \pm 0,012$$

3.  $\bar{x} = 0,1160$

$$(x - \mu) = 0,1160 - 0,123 = -0,07$$

$$S = 0,0032$$

tra bảng ở bậc tự do 3 ta có  $t = 3,18$  với xác suất 95%

Vì  $-0,07 < -0,0059$  nên phép xác định không mắc sai số hệ thống với xác suất 95%

4. Mẫu A : a/ 0,01% , 0,9% ; b/ -0,03% và -2,5%

Mẫu B : a/ 0,036% ; 0,42% ; b/ -0,04% và 0,48%

5. Đáp số:  $7,24\% \pm 0,42\%$  ;  $S = 0,25$  (  $P = 90\%$  )

6.  $S = 0,061$  (  $P = 95\%$  ) ;  $3,22 \pm 0,15$

11.	$P = 86\%$	$P = 95\%$
a/	$18,5 \pm 3,1$	$18,5 \pm 4,7$
b/	$18,5 \pm 3,2$	$18,5 \pm 3,3$
c/	$18,5 \pm 1,5$	$18,5 \pm 2,4$

13. a/  $-20\%$  ; b/  $-4,0\%$  ; c/  $-2,0\%$  ; d/  $-0,4\%$  ; đ/  $-0,2\%$

15. a/  $S = 0,25$  (  $P = 90\%$  ) ;  $7,24 \pm 0,42$  :

b/  $7,24 \pm 0,26$  (  $P = 90\%$  )

17.  $S = 0,031$  (  $P = 95\%$  ) ,  $0,72 \pm 0,077$

19. a/  $S = 0,061$  (  $P = 95\%$  ) ,  $3,22 \pm 0,015$

b/  $3,22 \pm 0,06$  (  $P = 95\%$  )

21. a/  $\bar{x} - \mu = 0,11\%$  CaO ;  $Z\sigma\sqrt{n} = 0,092\%$  sai số hệ thống được bỏ qua.

b/  $\bar{x} - \mu = 0,11\%$  CaO ;  $t\sigma\sqrt{n} = 0,14\%$  sai số hệ thống không hiện rõ

25. a/ Không phát hiện sự khác nhau có nghĩa .

b/ Sự khác nhau có nghĩa

c/ Sự khác nhau có nghĩa

30.  $S_y$   $y$

a/ 0,03	0,57 ( $\pm 0,03$ )
b/ 0,09	21,26 ( $\pm 0,09$ )
c/ 0,1 .10 <sup>-16</sup>	6,9 ( $\pm 0,1$ ) .10 <sup>-16</sup>
d/ 0,01 .10 <sup>-5</sup>	1,84 ( $\pm 0,01$ ) .10 <sup>-5</sup>
d/ 0,5	6,0 ( $\pm 0,5$ )
e/ 0,1 .10 <sup>-3</sup>	8,1 ( $\pm 0,1$ ) .10 <sup>-3</sup>
g/ 55	8,8 ( $\pm 0,6$ ) .10 <sup>-2</sup>
h/ 0,06 .10 <sup>-6</sup>	4,69 ( $\pm 0,06$ ) .10 <sup>-6</sup>

33. Đáp số: a/ 3 ; b/ 4 ; c/ 4 ; d/ 6 ; e/ 4 ; g/ 5

35. Đáp số: 2,54 %



**Phụ lục 1**

Bảng hàm Gauss  $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$

x	0	1	2	3	4	5	6	7	8	9
0.0	0.3989	3989	3989	3986	3986	3984	3982	3980	3977	3973
0.1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0.2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0.3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0.4	3683	3668	9653	3637	3621	3605	3589	3572	3555	3538
0.5	3521	5303	3485	3467	3448	3929	3410	3391	3372	3352
0.6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0.7	3123	3101	3079	3056	3034	3011	2983	2966	2943	2920
0.8	2879	2874	2850	2827	2803	2780	2756	2732	2709	2685
0.9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1.0	0.2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1.1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1.2	1942	1919	1895	1872	1849	1826	1804	1781	1758	1736
1.3	1714	1691	1669	1647	1626	1640	1582	1561	1539	1518
1.4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1.5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1.6	1109	1092	1074	1057	1040	1023	1006	0889	0973	0957
1.7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1.8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1.9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2.0	0.0840	0579	0519	0508	0498	0468	0479	0466	0459	0049

2.1	0440	0431	0422	0413	0404	0396	0388	0379	0371	0363
2.2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2.3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2.4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2.5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2.6	0136	0132	1029	0126	0122	0119	0116	0113	0110	0107
2.7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2.8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2.9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3.0	0.0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3.1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3.2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3.3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3.4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3.5	0009	0008	0008	0008	0008	0007	0007	0007	0006	0006
3.6	0006	0006	0006	0006	0006	0005	0005	0005	0004	0004
3.7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3.8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3.9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001
x	0	1	2	3	4	5	6	7	8	9

**Phụ lục 2.** Cận tích phân của phân bố t phụ thuộc vào xác suất P (hai phía) và  
(một phía ) bậc tự do f

f	P=0.50	0.75	0.90	0.95	0.98	0.99
1	1.00	2.41	6.31	12.7	31.82	63.7
2	0.816	1.60	2.92	4.30	6.97	9.92
3	0.765	1.42	2.35	3.18	4.54	5.81
4	0.740	1.31	2.13	2.78	3.75	4.60
5	0.727	1.30	2.01	2.57	3.37	4.03
6	0.718	1.27	1.94	2.45	3.14	3.71
7	0.711	1.25	1.89	2.36	3.00	3.50
8	0.706	1.24	1.86	2.31	2.90	3.36
9	0.703	1.23	1.83	2.26	2.82	3.25
10	0.700	1.22	1.81	2.23	2.76	3.17
11	0.697	1.21	1.80	2.20	2.72	3.11
12	0.695	1.21	1.78	2.08	2.68	3.05
13	0.691	1.20	1.77	2.16	2.65	3.01
14	0.402	1.20	1.76	2.14	2.62	2.98
15	0.691	1.20	1.75	2.13	2.60	2.95
16	0.690	1.19	1.75	2.12	2.58	2.92
17	0.689	1.19	1.71	2.11	2.57	2.90
18	0.688	1.19	1.73	2.10	2.55	2.88

19	0.688	1.19	1.73	2.09	2.54	2.86
20	0.687	1.18	1.73	2.09	2.53	2.85
21	0.680	1.18	1.72	2.08	2.52	2.83
22	0.686	1.18	1.72	2.07	2.51	2.82
23	0.685	1.18	1.71	2.07	2.50	2.81
24	0.685	1.18	1.71	2.06	2.49	2.80
25	0.684	1.18	0.71	2.06	2.49	2.79
26	0.684	1.18	1.71	2.06	2.48	2.78
27	0.684	1.18	1.71	2.05	2.47	2.77
28	0.683	1.17	1.70	2.05	2.47	2.76
29	0.683	1.17	1.70	2.05	2.46	2.76
30	0.683	1.17	1.70	2.04	2.46	2.75
40	0.681	1.17	1.68	2.02	2.42	2.70
60	0.679	1.16	1.67	2.00	2.39	2.66
120	0.677	1.16	1.66	1.98	2.36	2.62
$\infty$	0.674	1.15	1.64	1.96	2.33	2.58
f	$\bar{P} = 0.75$	0.875	0.95	0.975	0.99	0.995

**Phụ lục 3.** Cận tích phân của phân bố  $\chi^2$  phụ thuộc vào xác suất  $\overline{P}$  và bậc tự do f

f	$\overline{P} = 0.01$	0.5	0.10	0.50	0.90	0.95	0.99
1	0.0 <sup>3</sup> 157	0.0 <sup>3</sup> 393	0.0158	0.455	2.71	3.84	6.64
2	0.0201	0.103	0.211	1.39	1.61	5.99	9.21
3	0.115	0.352	0.518	2.37	6.25	7.81	11.3
4	0.287	0.711	1.06	3.30	7.78	9.49	13.3
5	0.554	1.15	1.60	4.35	9.21	11.1	15.1
6	0.872	1.64	2.20	5.35	10.6	12.6	16.8
7	1.24	2.17	2.83	6.35	12.0	14.1	18.5
8	1.65	2.73	3.49	7.34	13.4	15.5	20.1
9	2.09	3.33	4.17	8.34	14.7	16.9	21.7
10	2.56	3.94	4.87	9.34	16.0	18.3	23.2
11	3.55	4.57	5.58	10.3	17.3	19.7	24.7
12	3.57	5.23	6.30	11.3	18.5	21.0	26.2
13	4.11	5.89	7.01	12.3	19.8	22.4	27.7
14	4.66	6.57	7.79	13.3	21.1	23.7	29.1
15	5.23	7.26	8.55	14.3	22.3	25.0	30.6
16	5.81	7.96	9.31	15.3	23.5	26.3	32.0
17	6.41	8.67	10.1	15.3	24.8	27.6	33.4
18	7.01	9.39	10.9	17.3	26.0	28.9	34.8
19	7.63	10.1	11.7	18.3	27.2	30.4	36.2
20	8.26	10.1	12.4	19.3	28.4	34.4	37.6

21	8.90	11.6	13.2	20.3	29.6	32.7	38.9
22	9.54	12.3	14.0	21.3	30.8	33.9	40.3
23	10.2	13.1	14.8	22.3	32.0	35.2	41.6
24	10.9	13.8	15.7	23.3	33.2	36.4	43.0
25	11.5	14.6	16.5	24.3	34.4	37.7	44.3

**Phụ lục 4:** Giới hạn tích phân của phân bố F trong sự phụ thuộc vào  $f_1$ ,  $f_2$ , đối với  $P=0.95$ .

$f_2$	$f_1=1$	2	3	4	5	6	7
1	161	200	246	225	230	231	239
2	18.51	19.00	19.16	19.25	19.30	19.33	19.37
3	10.13	9.55	9.28	9.12	9.01	8.04	8.84
4	7.71	6.94	6.59	6.39	6.26	6.16	6.04
5	6.61	5.79	5.41	5.49	5.05	4.95	4.82
6	5.99	5.14	4.76	4.53	4.34	4.28	4.45
7	5.59	4.74	4.35	4.12	3.97	3.87	3.73
8	5.32	4.46	4.07	3.84	3.69	3.58	3.44
9	5.12	4.26	3.86	3.63	3.48	3.37	3.23
10	4.96	4.10	3.71	3.48	3.33	3.22	3.07
11	4.84	3.98	3.59	3.36	3.20	3.09	2.95
12	4.75	3.88	3.49	3.26	3.11	3.00	2.85
13	4.67	3.80	3.41	3.18	3.02	2.92	2.77
14	4.60	3.74	3.34	3.11	2.96	2.85	2.70
15	4.51	3.68	3.29	3.06	2.90	2.79	2.64
16	4.49	3.63	3.24	3.01	2.85	2.74	2.59
17	4.45	3.59	3.20	2.96	2.81	2.70	2.55
18	4.41	3.55	3.16	2.93	2.77	2.66	2.51
19	4.38	3.52	3.13	2.90	2.74	2.63	2.48
20	4.35	3.49	3.10	2.87	2.71	2.60	2.45
21	4.32	3.47	3.07	2.84	2.68	2.57	2.42
22	4.30	3.44	3.05	2.82	2.66	2.55	2.40
23	4.28	3.42	3.03	2.80	2.64	2.53	2.38
24	4.26	3.40	3.01	2.78	2.62	2.51	2.36
25	4.24	3.38	2.99	2.76	2.60	2.49	2.34
26	4.22	2.37	2.98	2.74	2.59	2.47	2.32

27	4.21	3.35	2.96	2.73	2.57	2.46	2.30
28	4.20	2.34	2.95	2.71	2.56	2.44	2.29
29	4.18	3.33	2.93	2.70	2.54	2.43	2.28
30	4.17	3.32	2.92	2.69	2.53	2.42	2.27
40	4.08	3.23	2.84	2.61	2.45	2.34	2.18
60	4.00	3.15	2.76	2.52	2.37	2.25	2.10
120	3.92	3.07	2.68	2.45	2.29	2.17	2.02
$\infty$	3.84	2.99	2.60	2.37	2.21	2.09	1.94
$f_2$	$f_1=1$	2	3	4	5	6	7



## TÀI LIỆU THAM KHẢO

- 1.K.Doerffel : Statistik in der analytischen Chemie. Leipzig 1966  
Bản tiếng Nga ,NXB Mir 1968
- 2.D.Kealey and P.J.Haines :Analytical Chemistry  
Bios;2002
- 3.Douglas A. Skoog and James J. Leary: Principle of Instrumental Analysis  
Sauders College Publishing ; 1992

## MỤC LỤC

Trang	
<b>Chương 1 Các thuật ngữ và định nghĩa</b>	<b>1</b>
Mê du răng, Phép đo , Độ đúng , Độ chính xác	1
Độ lệch chuẩn	5
Giá trị thật	5
Sai số , Sai số ngẫu nhiên , sai số hệ thống	5
Hệ số biến sai và trung vị	7
Độ lặp lại	8
Giới hạn tin cậy và khoảng tin cậy	9
<b>Chương 2 Bản chất nguồn gốc các loại sai số và phân loại</b>	<b>10</b>
2.1. Các loại sai số và phân loại	10
2.1. Nguồn gốc của sai số hệ thống	11
2.3. Phát hiện , xác định và loại trừ các sai số hệ thống	12
<b>Chương 3 Sai số ngẫu nhiên</b>	<b>14</b>
1. Các đại lượng ngẫu nhiên và sai số ngẫu nhiên	14
2. Xử lý thống kê các sai số ngẫu nhiên	16
2.1. Tập toàn thể và tập chọn lọc	16
2.2. Áp dụng các phương pháp thống kê để kiểm tra các giả thiết	20
3. Các loại phân bố	21
3.1. Phân bố tần số thực nghiệm	21
3.2. Phân bố lý thuyết	27
3.2.1. Phân bố chuẩn Gauss	27
3.2.2. Phân bố Student T	31
3.2.3. Phân bố chuẩn F	32
3.2.4. Phân bố $\chi^2$	34
3.2.5. Phân bố Poisson	35
3.2.6. Quan hệ giữa các loại phân bố	36

<b>Chương 4 Ứng dụng để tính và biểu diễn các loại sai số</b>	<b>37</b>
1.Cách tính các sai số	37
2.Sự cộng hợp các sai số ngẫu nhiên và định luật lan truyền sai số	37
3.Các ví dụ cụ thể	50
4.Kiểm tra tính sai lệch	55
5. Làm tròn các kết quả từ các phép tính số học	59
6.Phương pháp bình phương tối thiểu	61
<b>Chương 5 Mẫu chuẩn và quá trình hiệu chuẩn</b>	<b>66</b>
5.1. Mẫu chuẩn	66
5.1.1. Định nghĩa Mẫu chuẩn	66
5.1.2. Phân loại	71
5.1.3. Một số đặc trưng của Mẫu chuẩn và phương pháp thể hiện	70
5.1.4.Quy trình chế tạo Mẫu chuẩn	77
5.2. Quá trình hiệu chuẩn	87
5.3. Lựa chọn Mẫu chuẩn	80
<b>Câu hỏi và bài tập</b>	<b>82</b>
<b>Phụ lục 1</b>	<b>97</b>
<b>Phụ lục 2</b>	<b>99</b>
<b>Phụ lục 3</b>	<b>101</b>
<b>Phụ lục 4</b>	<b>103</b>
<b>Tài liệu tham khảo</b>	<b>105</b>