



**UNIVERSIDAD DE LOS
ANDES**
Facultad de Ciencias
Departamento de
Química



*Introducción al tratamiento de datos experimentales en
Química.*

Dr. Ricardo Hernández

Prof. Asistente - Facultad de Ciencias - ULA

Mérida, 11 de Marzo de 2004

*Núcleo "Pedro Rincón Gutiérrez", La Hechicera, Edificio "A" Ciencias, Cuarto Piso
OPT la Hechicera, No 53*

Mérida, Mérida, ZP 5101 – Venezuela

E-mail: rmhr@ula.ve

Tel.: 0274-2401392; Fax: 0274-2401286

El peso de un átomo de carbono es $1,99513 \times 10^{-23}$ g. Esta cantidad contiene seis dígitos; entendiendo por dígito a cada uno de los diez números básicos del sistema decimal (0-9). Una *cifra significativa* es un dígito que denota un valor discreto en el lugar de la cantidad en la que se encuentra. El dígito 0 (cero) es una cifra significativa excepto cuando está ubicado delante de los restantes dígitos básicos del sistema decimal (1-9) o puede ser incorporado en una potencia de diez. De esta forma, el peso del átomo de carbono señalado anteriormente contiene seis cifras significativas, aún si no se usa la notación científica. De igual forma, el valor del peso de un mol de átomos de carbono señalado arriba, contiene seis cifras significativas. Se debe tener presente que se presentan dudas cuando varios dígitos con valor de cero preceden al indicador decimal; por ejemplo, el Número de Avogadro, ¿Contiene tres o veintitrés cifras significativas? Siendo que esta cantidad es una constante aceptada como $6,02 \times 10^{23}$, el número de cifras significativas es 3. En otros casos, el número de cifras significativas estará dado por la incertidumbre o apreciación del instrumento de medida.

El número de cifras significativas puede ser definido como el número de dígitos necesarios para expresar la incertidumbre asociada a una medida. De esta forma, las cantidades o valores observados deben ser recogidos hasta la primera cifra significativa con incertidumbre; es decir, hasta donde la apreciación o límite de detección del instrumento indique. Así, en la mayoría de los análisis los pesos se determinan hasta la décima de miligramo, por ejemplo, 1,0234 g; esto significa que el peso medido es menor que 1,0235 y mayor que 1,0233. Si este peso se expresa como 1,023 g, significa que se ha determinado el peso con una incertidumbre en el dígito correspondiente al miligramo (0,001 g) y que el peso medido se ubica entre 1,022 g y 1,024 g.

Cálculos y cifras significativas

Existen una serie de reglas que deben ser tomadas en cuenta al momento de manipular las cifras significativas cuando se efectúan cálculos:

- i. Se deben recoger los valores hasta la primera cifra significativa con incertidumbre.
- ii. Se deben redondear cantidades hasta el número correcto de cifras significativas.
- iii. La suma o diferencia de cualesquiera cantidades no puede ser más precisa que la cantidad que tenga la mayor incertidumbre: $122,2 + 12,345 + 1,01 = 135,5$
- iv. El porcentaje de precisión del resultado de una operación de multiplicación o división no puede contener más cifras significativas que el porcentaje de precisión de la variable con la menor cantidad de cifras significativas, y el menor valor absoluto sin tomar en cuenta el indicador decimal (variable clave). Así, por ejemplo, en el cálculo siguiente:

$$\frac{25,63(-0,6481)3,301}{1,1579} = -47,35491899$$

La variable clave es 25,63; pues es la que posee la menor cantidad de cifras significativas y el menor valor absoluto; por lo tanto, el resultado debe expresarse como -47,35

Sin tomar en cuenta el indicador decimal, 2563 es menor que 3301; de aquí se tiene que la incertidumbre porcentual para la primera cantidad será $(1/2563)100 = 0,039$; es decir 0,04. Por otro lado, la incertidumbre porcentual de la segunda será $(1/3301)100 = 0,03$. La incertidumbre que debe tomarse es la mayor, 0,04

c. Redondeo

- Si el dígito que sigue a la última cifra significativa es mayor que 5, ésta debe ser redondeada a el cantidad próximo mayor. Si en el cálculo anterior el resultado fuese -47,35691899, deberíamos redondear a -47,36
- Si el dígito es menor que 5, entonces la última cifra significativa queda inalterada y se eliminan todos los dígitos que la siguen. Ver el resultado del redondeo en el cálculo anterior.
- Si el dígito que sigue a la última cifra significativa es 5, la última cifra significativa es redondeada al cantidad par más cercano:

$$-47,35591899 \Rightarrow -47,36$$

$$-47,36591899 \Rightarrow -47,36$$

$$-47,37591899 \Rightarrow -47,38$$

2. **Nociones de Estadística**

Población

Conjunto de todos los individuos, objetos en estudio. Como regularmente resulta imposible, o nada práctico, recoger los datos de una población, especialmente si ésta es muy numerosa, se recurre a la toma de una pequeña porción representativa denominada *la muestra*.

Tipos de datos, arreglos y frecuencias

Los datos pueden ser de dos tipos: *datos en bruto o sin procesar* y *datos organizados o procesados*. Un *arreglo* es un conjunto de datos organizados; la diferencia entre el mayor y el menor valor (límites superior e inferior) se denomina *rango o intervalo de datos*. Cuando se organizan grandes cantidades de datos, generalmente aparecen conjuntos de datos similares a los que se denomina *clases o categorías*. El número de datos pertenecientes a una determinada clase define la *frecuencia de la clase*. Al conjunto organizado de frecuencias se le denomina *distribución de frecuencias* y generalmente se representa mediante un grafico de barras.

Promedios

Un promedio, es un valor que es típico o representativo de un conjunto de datos. Los promedios tienden a ubicarse de manera central dentro de un intervalo de datos. Se pueden definir varios tipos de promedios, de los cuales, los más comunes son: *la media aritmética, la mediana y la moda*. Cada uno tiene ventajas y desventajas dependiendo de los datos y del propósito de los cálculos.

a) La media aritmética

La media aritmética, o media, de un conjunto de N datos, $x_1, x_2, x_3, x_4, \dots, x_N$, se denota por \bar{x} y se define como:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_N}{N} = \frac{1}{N} \sum_{j=1}^N x_j$$

Si los mismos datos ocurren con cierta frecuencia f , entonces la media aritmética será:

$$\bar{x} = \frac{f_1x_1 + f_2x_2 + f_3x_3 \dots + f_Nx_N}{f_1 + f_2 + f_3 \dots + f_N} = \frac{\sum_{j=1}^N x_j}{\sum f}$$

b) La media aritmética ponderada

En ocasiones se asocia a los datos con ciertos factores ponderados (o pesos), w , que definen la importancia de los mismos o su significación. En estos casos la media adquiere un valor que difiere de la media aritmética y se define como:

$$\bar{x} = \frac{w_1x_1 + w_2x_2 + w_3x_3 \dots + w_Nx_N}{\sum w}$$

c) La mediana y la moda

La mediana de un arreglo de datos (conjunto de datos acomodados) es el valor de en medio –si el número de datos es impar- o bien la media aritmética de los dos valores centrales –si el número de datos es par. Así, por ejemplo, el arreglo de números 3, 4, 4, 5, 6, 8, 8, 8, 10, tiene una mediana igual a 6; mientras, el arreglo 5, 5, 7, 9, 11, 12, 15, 18, tiene una mediana igual a $(9+11)/2 = 10$.

La moda de un conjunto de números es el valor que ocurre con mayor frecuencia; es decir el valor más común. La moda puede no existir en ciertos conjuntos, y si existe puede no ser única. Un conjunto que tiene una sola moda se denomina *unimodal*.

Dispersión de datos y error de una medida

El grado con el cual los diversos datos tienden a esparcirse respecto de un valor promedio se denomina dispersión, variación o precisión. Existen varias formas de calcular la dispersión de los datos, siendo las más comunes: el *rango*, la *desviación estándar* y la *varianza*.

- a) El **rango** es la diferencia entre el límite superior y el límite inferior de un arreglo de datos.
- b) La **desviación estándar** se define de la siguiente manera:

$$s = \sqrt{\frac{\sum_{j=1}^N (x_j - \bar{x})^2}{N}}$$

Generalmente se reemplaza N por $(N-1)$ cuando el número de datos es pequeño –como suele ocurrir van de 3 a 5 – ya que el valor resultante es una mejor estimación de la desviación estándar de los datos.

- c) La **varianza** se define como el cuadrado de la desviación estándar.

Es imposible realizar un análisis químico sin que los resultados estén totalmente libres de errores. Sin embargo, es la función del químico minimizarlos y estimar su magnitud con una exactitud aceptable. Cada medición esta sujeta a muchas imprecisiones (errores) que cuando se combinan producen una diversidad de resultados alejados, en mayor o menor grado del valor verdadero. Sin embargo es posible determinar de imprecisión definiendo un intervalo de valores dentro del cual se encuentra centrado el valor verdadero. En los análisis químicos es común encontrar al menos dos tipos generales de errores: (a) *errores aleatorios o indeterminados* y (b) *los errores sistemáticos o determinados*. Los errores aleatorios surgen como consecuencia de variaciones no controladas de variables experimentales; por ejemplo variaciones del voltaje de la red de energía eléctrica o las vibraciones causadas por vehículos al paso. Estos errores no tienen una magnitud o signo definido o constante. Por el contrario, los errores sistemáticos tienen un valor y signo definidos, y por tanto una causa común. Este error suele ser de mayor magnitud que el error aleatorio y ocasiona la mayor desviación respecto al valor real. En las técnicas en las que la apreciación del investigador es necesaria, este error es, generalmente, el de mayor importancia; por ejemplo, al estimar el color de una solución, o la posición de una aguja entre dos divisiones de la escala de un instrumento, o la posición del menisco de un líquido en un matras aforado.

El término *precisión* se emplea para describir la reproducibilidad o concordancia de las medidas. Un conjunto de medidas con alto grado de precisión implica un conjunto de valores poco dispersos y con valores de error aleatorio bajos. El término *exactitud* indica que tan cercano esta una medición de su valor verdadero o aceptado. La exactitud se expresa en términos de *error absoluto* y *error relativo*. El error absoluto es la diferencia entre el valor hallado experimentalmente y el valor verdadero o aceptado como verdadero.

$$E = x_i - x_t$$

Donde x_t es el valor verdadero y x_i es el valor encontrado. El error relativo suele ser una medida más empleada y más útil, generalmente expresada en términos de porcentaje:

$$E_r = \frac{x_i - x_t}{x_t}$$

Bibliografía

- Murray R. Spiegel y David P. Lindstrom, “Estadística Schaum”, 1ra Edición, McGraw-Hill, México 2000
- D. A. Skoog, D. M. West, F. J. Holler y S. R. Crouch, “Química Analítica”, 7ma Edición, McGraw-Hill, México, 2001