UNIVERSIDAD DE PUERTO RICO EN HUMACAO
DEPARTAMENTO DE QUÍMICA
(www.uprh.edu / ~quimgen)

QUIM 3003 - 3004

MEDIDAS: TRATAMIENTO DE LOS DATOS EXPERIMENTALES

I. Introducción

La mayor parte de los experimentos que se realizarán durante las siguientes sesiones de laboratorio serán cuantitativos, esto es, experimentos en que se encuentre el valor numérico de alguna propiedad de un sistema. Toda medida realizada debe de tener sus unidades. Debido a que toda medida está sujeta a errores bien sean determinados o indeterminados para obtener una medida con buena precisión y/o exactitud se debe repetir la determinación el mayor número de veces posible (no menos de tres veces). Las medidas se expresarán al número correcto de cifras significativas y se utilizará la Prueba de Rechazo de Valores Dudosos que mejor aplique *(Ver Parte VI*) para eliminar el o los valores que resulten inaceptables. Luego se continua con el análisis estadístico de las determinaciones restantes hasta obtener el mejor valor a reportar o tendencia central de los rebultados.

II. Cifras significativas de una medida

Son aquellos dígitos que conocemos con certeza más el primer dígito incierto (el último dígito que posee un número es el primero que generalmente se considera incierto), por ejemplo, en el caso de la pesada de 25.0100 g hay 6 cifras significativas, pues el primer dígito incierto es el de las décimas de mg.

Otros ejemplos: 0.134

0.0134 todos poseen 3 cifras significativas

1.34

134 , escrito como 1.34×10^2

La regla general para contar las cifras significativas es la de “leer el número de izquierda a derecha, comenzando a contar los dígitos con aquel que, en ese orden, es el primero que no es cero”. La posición del punto decimal se ignora porque está determinada por las unidades que se utilicen y no por la exactitud del instrumento.

A. Reglas prácticas para determinar el número de cifras significativas:

- **ceros significativos**:

Los ceros que forman parte de un número pueden contar o no para la determinación de las cifras significativas, de acuerdo con el lugar que ocupen en número y a la manera cómo se haya escrito dicho número.

Ejemplos:

10400	expresión dudosa: no se sabe si los ceros están señalando la exactitud de la medida o señalando el punto decimal. Para indicar un número específico de cifras significas hay que escribirlo en notación científica.
1.04×10^4	tres cifras significativas indica que el instrumento tiene una exactitud que solamente alcanza $\pm$ 100 unidades. La notación científica denota con exactitud las cifras significativas
1.040×10^4	cuatro cifras significativas: indica que el instrumento tiene una exactitud de $\pm$ 10 unidades

- **Reglas:**

a) Los ceros intercalados entre los números, siempre cuentan como ceros significativos, ej.: 2041 tiene cuatro cifras significativas

b) Los ceros a la izquierda del punto decimal cuentan como cifras significativas si están precedidos de otros números, pero resulta dudoso el contarlas. Es siempre preferible expresarlo en notación científica teniendo en cuenta la exactitud del instrumento.

Ejemplos:	73,000:	dudoso, posiblemente cinco cifras significativas
	7.30×10^2	tres cifras significativas
	0.730	tres cifras significativas

c) Los ceros a la derecha del punto decimal sólo cuentan si están precedidos de otros números.

Ejemplos:	0.008	una sola cifra significativa
	0.0080	dos cifras significativas
	0.08010	cuatro cifras significativas

B. Incertidumbre de una medida

1) *Incertidumbre absoluta :* se asume que es de más o menos una unidad en la última cifra significativa, sino se especifica un valor.

Ejemplos:	0.0240	tiene una incertidumbre de ± 0.0001
	2.40	tiene una incertidumbre de ± 0.01
	2.4×10^3	tiene una incertidumbre de $\pm 0.1 \times 10^3 = \pm 100$
	2400	es dudoso, si los ceros indican además del punto decimal, la exactitud del aparato, la incertidumbre absoluta es de ±1

2) *Incertidumbre relativa :* se encuentra dividiendo la incertidumbre absoluta entre la medida.

Incertidumbre absoluta: 0.0001

Iincertidumbre relativa: $\dfrac{0.0001}{0.0240} = \dfrac{1}{240} = 0.0046$

C. Operaciones considerando cifras significativas y sus incertidumbres

1) Suma y Resta

La incertidumbre absoluta del resultado se determina sumando las incertidumbres absolutas de las cantidades envueltas.

Ejemplo 1:

$$\begin{array}{rll} 21.3532 & \pm\ 0.01 & g \\ 1.613 & \pm\ 0.001 & g \\ 152.1 & \pm\ 0.1 & g \\ \hline 175.0662 & \pm\ 0.1011\,g & = 175.1 \pm 0.1\,g \end{array}$$

$$(\textit{incertidumbre absoluta} = 0.1\,g)$$

Ejemplo 2:

$$\begin{array}{rl} 52.34732 & \pm 0.00001\ g \\ -\ 50.212 & \pm 0.001\ \ \ \ g \\ \hline 2.13532 & \pm 0.00101\ g\ = 2.135\ \pm 0.001\ g \end{array}$$

$$(\textit{incertidumbre absoluta permisible de } 0.001\ g)$$

En ambos casos ha sido necesario ajustar el resultado a la incertidumbre de la medida menos exacta. Esta operación se llama "redondeo" del resultado. Una regla a menudo usada para redondear es aumentar el último dígito que se conserva en 1, si el primer dígito que se desprecia es mayor que cinco y conservar el último dígito que se deja si el primero que se desecha es menor que cinco. Si el primer dígito que se desecha resulta ser cinco, se sigue la regla de conservar el primer dígito que se deja si es par y aumentarlo en uno si es impar.

2) Multiplicación y división

- La incertidumbre relativa del resultado es igual a la suma de las incertidumbres relativas de los factores.

Ejercicio: Escriba el resultado de la siguiente multiplicación con su incertidumbre absoluta correspondiente

$$24.2\ mL\ x\ 0.98271 \frac{g}{mL} = ?$$

Solución:

- Primero : calcular la incertidumbre relativa de cada medida:

Incertidumbre relativa de 24.2 es: $\frac{0.1}{24.2} = \frac{1}{242}$

Incertidumbre relativa de 0.98271 es: $\frac{0.00001}{0.98271} = \frac{1}{98271}$

- Segundo: comparar las incertidumbres relativas:

Incertidumbre relativa 23.78 es : $\frac{0.01}{23.78} = \frac{1}{2378}$ $\left(\textit{menor que } \frac{1}{242} \right)$

Incertidumbre relativa de 23.8 es: $\frac{0.1}{23.8} = \frac{1}{238} \quad \left(mayor\ que\ \frac{1}{242} \right)$

- Tercero: calcular la incertidumbre del resultado:

$$\frac{1}{242} + \frac{1}{98271} = \frac{1}{242} = 0.0041322$$

- Cuarto: Calcular la incertidumbre absoluta(redondear a la primera cifra significativa)

$$23.78582 \times 0.0041322 = 0.0982885 = 0.1$$

- Quinto: Escribir el **RESULTADO: 23.8 ± 0.1g**

- Un método que a menudo da el mismo resultado que la regla anterior es redondear el resultado al número de cifras significativas del factor de menos cifras.

  Ejemplo: 24.2 x 0.98271 = 23.781582 = 23.8 (tres cifras significativas, igual que en el factor 24.2)

III. Tendencia central de un grupo de medidas

Una vez que un grupo de resultados ha sido calculado haciendo uso adecuado de las cifras significativas y se han aplicado las pruebas de rechazo de valores dudosos es necesario determinar el mejor valor para poder compararlo con el "verdadero valor" (μ) y así tener una idea de la exactitud de los resultados.

La selección del mejor valor se hace calculando la tendencia central del grupo de medidas. Esta tendencia central se mide generalmente encontrando el promedio aritmético ($\bar{x}$)

A. Promedio aritmético $(\overline{X})$

Se determina con la expresión:

$$\overline{X} = \frac{X_1 + X_2 + \ldots + X_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

donde el símbolo $\sum_{i=1}^{n} x_i$ se lee como " la suma de todos los valores obtenidos".

B. Mediana (M)

Se colocan las determinaciones aceptadas (luego de aplicar la prueba de rechazo de valores dudosos) en orden ascendente.

1) Para n impar: el valor central será la mediana

Ejm.:

X_i (g)
3.18
3.19
3.20

$M = 3.19\ g$

2) Para n par: el promedio de los valores centrales será la mediana

X_i (g)
3.16
3.18
3.19
3.20

$$M = \frac{3.19\ g + 3.18\ g}{2} = 3.185\ g$$

$$M = 3.18\ g$$

Se le aplicó la regla del redondeo

IV. Exactitud, Precisión e Intervalo

Siempre que en el laboratorio se determina un valor, la medida debe hacerse varias veces (generalmente tres (3) veces). El analista necesita determinar el "mejor valor" para reportarlo. El mejor valor se puede determinar calculando el promedio aritmético de los resultados individuales, pero en ocasiones algunos de dichos valores son muy diferentes de los demás y entonces surge la duda de si incluirlo o no en el promedio. Además, el analista debe conocer el límite de error a que estén sujetas sus medidas, si quiere saber hasta qué punto son confiables. El objeto de este bosquejo es describir los métodos estadísticos que con más frecuencia se utilizan para ambos propósitos.

A. Exactitud

Es lo cercano que está una determinación (X_i) del verdadero valor (μ). La exactitud se expresa generalmente en término del "error" de dichas determinaciones midiéndose el error por la expresión (X_i - μ) (léase como: "la diferencia entre la medida y el verdadero valor.

Errores

- *Errores absolutos (Ea) y Errores relativos (Er)*

a) Error absoluto de una determinación: viene dado por la diferencia entre el valor de la determinación y el valor real o aceptado de la medida. Las unidades serán las mismas que las unidades envueltas en la determinación (g, ml, %, etc.). Por ejemplo, un objeto cuyo verdadero valor del peso es 0.1000 g y que al pesarlo la pesada resulta de 0.1001 g posee un error absoluto de 0.0001 g en dicha determinación.

$$E_a = X_i - \mu$$

b) Error relativo de una determinación X_1 (E_r): es la relación que existe entre el error absoluto y el verdadero valor. Viene dado por la expresión:

$$E_r = \frac{X_i - \mu}{\mu} = \frac{E_a}{\mu}$$

Como el error relativo es una relación entre magnitudes del mismo tipo de unidades, las unidades del numerador se cancelan con las del denominador y por lo tanto, la expresión carece de unidades. Los conceptos relativos se expresan en porciento (multiplicar por 100) o en partes por millón: ppm (multiplicar por 10^6). El error puede determinarse comparando el promedio de las determinaciones ($\bar{x}$) con el valor correcto (μ).

$$E_a = \overline{X} - \mu \qquad E_r = \frac{\overline{X} - \mu}{\mu} x\, 10^2 \qquad E_r = \frac{\overline{X} - \mu}{\mu} x\, 10^6$$

- *Errores determinados e indeterminados:*

a) Errores determinados o sistemáticos son aquellos causados por un método analítico defectuoso, por mal funcionamiento de alguno de los instrumentos utilizados o por mala manipulación del analista. Pueden ser eliminados encontrando la causa que los produce.

b) Errores indeterminados o al azar son errores que son imposibles de eliminar debido a que en toda medida hay siempre un grado de incertidumbre imposible de eliminar, por ejemplo: el más cuidadoso analista es incapaz de leer una bureta de 50 ml con una exactitud mayor de 0.01 ó 0.02 ml, o es incapaz de pesar en una balanza analítica estándar con una exactitud mayor de 0.01 ó 0.02 mg. Estos errores son debido al grado máximo de exactitud con que fue diseñado el instrumento utilizado y son la causa de que un grupo sucesivo de determinaciones no reproduzcan el mismo valor del fenómeno, no importa lo preciso del método y de la técnica del analista

B. Precisión

Es lo cercano que están unas medidas de las otras dentro de un grupo de determinaciones. La precisión se mide en términos de la desviación de las medidas límites del promedio aritmético del grupo. La precisión lo que mide realmente es la habilidad del analista a la hora de reproducir un resultado experimental. Aunque generalmente la buena precisión coincide con la buena exactitud, es totalmente posible conseguir una buena exactitud y una mala precisión o viceversa.

Formas de Medir Precisión

La precisión se expresa en términos de desviación. Hay varios tipos de desviación:

1) Desviación promedio Absoluta $(\overline{d}_a)$ o Desviación promedio relativa $(\overline{d}_r)$:

Desviación de una medida: es la diferencia entre la medida y el valor promedio

$$d_i = (X_i - \overline{X})$$

Desviación promedio de un grupo de medidas: es el promedio de las desviaciones absolutas

$$\overline{d}_a = \sum d_i \qquad ; \qquad \overline{d}_a = \frac{1}{n}\sum_{i=1}^{n} \left|X_i - \overline{X}\right|$$

$$\% \, d_r = \frac{\bar{d}_a}{\bar{X}} \, X \, 10^2 \quad ; \quad ppm \, d_r = \frac{\bar{d}_a}{\bar{X}} \, X \, 10^6$$

2) Desviación estándar Absoluta (s_a) o Desviación estándar relativa (s_r)

Es una medida más representativa y más útil de la desviación. Se define como la relación existente entre la raíz cuadrada de la suma de los cuadrados de las desviaciones y la raíz cuadrada del número de determinaciones menos una.

- Desviación Estándar Absoluta (s_a):

$$s_a = \left[\frac{1}{n-1} \sum_{i=1}^{n} \left(X - \bar{X} \right)^2 \right]^{1/2}$$

$$s_a = \sqrt{\sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{n-1}}$$

Cuando (n - 1) es muy grande, para facilitar los cálculos, se puede emplear la fórmula:

$$\sum_{i-1}^{n} (X_i - \bar{X})^2 - \sum X_1^2 - (\bar{X})^2$$

La desviación estándar es más sensible al efecto de las determinaciones que distan mucho entre sí de lo que es la determinación de la desviación promedio y esa es la razón por la cual se prefiere la desviación estándar.

Tanto la desviación estándar como la desviación promedio pueden expresarse como desviaciones relativas utilizando las siguientes expresiones:

- Desviación Estándar Relativa (s_r) expresada :

  - $en \quad \% \, s = \frac{s}{\bar{x}} \, x \, 100$

  - $en \quad ppm \, s = \frac{s}{\bar{x}} \, x \, 10^6$

C. Intervalo

Se entiende por intervalo la diferencia entre las medidas extremas de un grupo de determinaciones, o sea, (X_n - X_1) suponiendo que las determinaciones están ordenadas en orden creciente de su magnitud.

El intervalo ha sido recomendado repetidamente en la literatura como una mejor medida de la precisión que la desviación estándar en el caso de que el número de determinaciones oscila entre 3 y 10, o sea, cuando el número de determinaciones es igual o menor que 10, pues obviamente, a grupos de dos determinaciones no se le pueden aplicar métodos estadísticos para eliminar valores dudosos.

VI. Pruebas de Rechazo de valores dudosos

Muchas veces ocurre que en un grupo de determinaciones, aparecen algunos que divergen bastante de los demás y surge la duda de si se deben incluir en el cálculo del mejor valor. A continuación describimos algunas técnicas que se usan para saber cuándo un valor se debe eliminar. Es importante indicar que aunque las pruebas de rechazo de valores dudosos se explican al final de este módulo *es lo primero que se hace al realizar un análisis estadístico*.

A. La Prueba de las 3S

Cuando el número de determinaciones es mayor de 10 y hay alguna determinación dudosa, dicha determinación puede eliminarse del cálculo de la desviación estándar, si se cumple que:

$$(X_{dudoso} - \overline{X}) > 3s$$

B. Prueba utilizando la desviación promedio

Si el grupo de medidas incluyera una determinación dudosa, se puede calcular la desviación promedio sin incluir dicha medida, si se cumple que (X_{dudoso} - X) > 4 d. Si la anterior especificación no se cumpliera, lo que se recomienda es hacer nuevas determinaciones para minimizar el peso de la determinación dudosa. La regla antes descrita solamente es aplicable cuando el grupo de medidas es de cuatro o más determinaciones.

C. La Prueba "Q"

Cuando el número de determinaciones es igual o menor de diez, también es muy frecuente el uso de la Prueba "*Q*" cuando se quiere determinar si un dato se puede o no eliminar. Dicha prueba se conduce a un nivel de un 90% de probabilidad o confiabilidad, o dicho de otro modo, si un valor dudoso se rechaza según la prueba Q, hay un 90% de que tal valor fuera realmente diferente del resto de las determinaciones.

Método:

- Divida la diferencia entre la determinación dudosa y la determinación más cercana a ella entre el valor del intervalo (incluyendo el valor dudoso), obteniéndose así el cociente Q.

- Si el valor de Q obtenido es mayor que el cociente $Q_{0.90}$ que aparece en la tabla correspondiente, el valor dudoso puede ser desechado. Los valores dudosos son X_1 o X_n. Se le hace la prueba primero al valor donde la diferencia entre éste y el número más cercano en valor sea mayor.

- Si es posible rechazar alguno de los límites, se recalcula el intervalo sin incluir el valor eliminado y se vuelven a probar los nuevos límites, hasta llegar a valores confiables de ambos.

Caso especial: grupo de tres determinaciones

Aquí sólo se puede rechazar uno de los límites, por lo que se puede calcular $(X_2 - X_1)$ ó $(X_3 - X_2)$ y dividir dicha diferencia entre el valor del intervalo $(X_3 - X_1)$. La Q así calculada se compara con $Q_{0.090}$ para n = 3, sólo podrán rechazarse valores muy divergentes. Por ejemplo, supóngase que se corre un análisis por triplicado y los tres valores obtenidos son 6.00%, 5.00% y 5.07%). La intuición dice que el valor de 6.00% es dudoso, sin embargo, al aplicar la Prueba Q no puede rechazarse tal valor. En casos como éste, lo recomendable es hacer dos o tres determinaciones más y aplicar de nuevo la prueba Q. Si el valor dudoso aún no se puede rechazar, es entonces mejor reportar la mediana como el mejor valor.

TABLA DE VALORES DE $Q_{0.90}$		Ecuaciones
n	$Q_{0.090}$	
3	0.941	
4	0.765	*Valor menor:* $Q = \left\lvert \dfrac{X_2 - X_1}{X_n - X_1} \right\rvert$
5	0.642	
6	0.560	
7	0.507	*Valor mayor:* $Q = \left\lvert \dfrac{X_n - X_{n-1}}{X_n - X_1} \right\rvert$
8	0.47	
9	0.44	
10	0.41	