

Universidad de los Andes.
Departamento de Química
Prácticas de Laboratorio de Química

Tratamiento Estadístico de Datos

1. Objetivos:

1. Definir los conceptos de exactitud y precisión
2. Clasificar los errores en las distintas mediciones y dar alternativas para llevarlos al mínimo.
3. Definir y aplicar el concepto de cifras significativas para obtener resultados confiables en una medición
4. Aplicar algunos de los principios de la estadística a valores experimentales obtenidos en el laboratorio para determinar el valor mas probable para las mediciones.

2. Introducción:

Si el resultado de una medición física tiene que ser de algún valor para quien la desarrolla, es esencial que tal persona tenga una idea acerca del grado de confianza que puede otorgársele a la mencionada determinación. Desafortunadamente, toda medición de una cantidad física se encuentra sujeta a un grado de incertidumbre y el experimentador solamente puede pretender alcanzar una aproximación aceptable al verdadero valor de la cantidad. En consecuencia, una labor esencial, aunque no muy sencilla, es la evaluación de la incertidumbre que afecta a toda medición.

3. Definiciones:

Cada una de las mediciones realizadas está sujeta a incertidumbre. Estas incertidumbres son responsables de las variaciones que se observan entre los resultados de varias mediciones idénticas de una misma cantidad. En la práctica se realizan varias mediciones, partiendo del razonamiento de que la confianza con que puede sostener sus resultados se ve aumentada si se demuestra su reproducibilidad. Para definir y estimar el mejor valor de la medición se dispone de dos cantidades:

- a. El valor medio o promedio: el valor obtenido dividiendo la suma de un grupo de mediciones por el número de resultados individuales que constituyen dicho grupo.
- b. La mediana: es aquel valor ordenado alrededor del cual se distribuyen los otros simétricamente, siendo la mitad de ellos numéricamente mayores y la

otra mitad numéricamente menores. Si el grupo consta de un número impar de mediciones, la elección de la mediana se puede hacer directamente; pero si el grupo consta de un número par de mediciones, se toma como mediana el valor medio del par de resultados centrales.

- c. La moda: es aquel valor que se repite con mayor frecuencia en un conjunto de datos. Es posible que exista más de una moda (conjunto multimodal).

4. Precisión:

La precisión de un conjunto de datos es la medida de la magnitud del error indeterminado (error inevitable asociado al proceso de medición); también se puede entender como una medida de la dispersión de los datos, es decir, qué tan juntos o separados están los datos entre sí.

Una medida de la fiabilidad de los datos, conocida como **prueba Q**, puede emplearse para decidir si un resultado que es notablemente diferente de los otros deberá rechazarse o no, para calcular la media. Para efectuar esta prueba las n mediciones se acomodan en orden creciente y se marcan como $X_1, X_2, \dots, X_{n-1}, X_n$. El valor “dudoso” que se evaluará para ser descartado mediante la prueba Q puede ser el primero (X_1) o el último (X_n)

$$Q_{calc} = \frac{|\text{Valor extremo} - \text{Valor siguiente}|}{\text{Rango}} = \frac{(x_n - x_{n-1}) \text{ ó } (x_2 - x_1)}{x_n - x_1}$$

Para que un valor se elimine, Q calculado debe ser igual o mayor al valor de Q reportado en la tabla (el 90% significa la confianza con la que se puede hacer esa determinación):

N	Q (90 %)	N	Q (90 %)
3	0.94	9	0.44
4	0.76	10	0.41
5	0.64	11	0.39
6	0.56	12	0.38
7	0.51	15	0.34
8	0.47	20	0.30

Hecha la prueba Q, se procede a calcular la media, y a partir de ahí, el parámetro estadístico que con más frecuencia se emplea para medir la precisión es la desviación estándar.

$$s = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{X})^2}{N - 1}}$$

Esta *estimación puntual* de la precisión puede usarse para crear un estimador más robusto y completo, un **intervalo de confianza**, el cual, cuanto más estrecho sea, indicará una mayor precisión. Este intervalo se ayuda del estadístico **t de Student** y se calcula de la siguiente manera:

$$I.C.(95\% \text{ confianza}) = \bar{x} \pm t \cdot \frac{s}{\sqrt{n}}$$

Donde \bar{x} es la media o promedio, n es el número de datos tomados, s es la desviación estándar y t es un valor que se revisa en la siguiente tabla, según el número de datos tomados (n):

n	t	n	t
2	12.706	8	2.365
3	4.303	9	2.306
4	3.182	10	2.262
5	2.776	11	2.228
6	2.571	12	2.131
7	2.447	13	2.086

5. Exactitud:

Parámetro que indica que tan cercanos están los resultados obtenidos respecto al valor real. Es importante determinar ese “valor verdadero” a partir de una fuente confiable, tomada por un técnico o por el profesor, por ejemplo.

Ojo: el promedio calculado a partir de los datos de una práctica no es el valor verdadero. De hecho, ese promedio es el que se compara con el valor verdadero proporcionado por el profesor o técnico, y de ahí se concluye sobre la exactitud obtenida en la práctica.

6. Error e Incertidumbre:

Términos que tienden a ser confundidos a la hora de evaluar la precisión y la exactitud de un experimento, pero que en realidad están bien definidos y discriminados.

a) Error Absoluto y Error Relativo: La exactitud de un resultado puede expresarse en términos de error relativo y error absoluto.

El error absoluto de un resultado X_i es la magnitud o valor absoluto de la diferencia entre el valor medido y el valor definido como verdadero (es decir, igual a $|X_i - X|$) y se expresa en las unidades de los dos números, es decir, en gramos, mililitros, segundos, etc.

Por ejemplo, la medición de cierta propiedad ha dado los siguientes resultados: 24,39, 24,20 y 24,28. Si tomamos el valor medio de la serie: 24,29, pero el valor verdadero es de 25,00 entonces tenemos un error absoluto de $|24,29 - 25,00| = 0,71$ unidades.

Por su parte, el error relativo es un porcentaje, sin unidades, y se calcula de la siguiente manera para valores esperados diferentes a 0:

$$\% \text{ Error relativo} = \frac{|X_i - X_{\text{real}}|}{X_{\text{real}}} \times 100 \%$$

b) Incertidumbre absoluta y relativa: Mientras el error es un indicador de exactitud, la incertidumbre es una medida de precisión. La incertidumbre absoluta expresa el margen de imprecisión asociado a una medida. Para el caso de una serie de datos, se define simplemente como la magnitud de la diferencia entre un dato particular y el promedio calculado.

$$\text{Incert. Absoluta del dato } i - \text{simo} = |x_i - \bar{x}|$$

Por su parte, la incertidumbre relativa es un porcentaje, pero se divide la incertidumbre absoluta entre el valor medido

$$\% \text{ Incertidumbre relativa} = \frac{\text{Incert. absoluta}}{\text{Valor medido}} \times 100 \% = \frac{|x_i - \bar{x}|}{x_i} \times 100 \%$$

Volviendo al ejemplo de las tres mediciones usado para determinar el error:

Muestra	Medición	Incertidumbre absoluta	Incertidumbre relativa
1	24,39	0,10	0,41%
2	24,20	0,09	0,37%

3	24,28	0,01	0,04%
Valor medio =	24,29		

La incertidumbre da una idea de la dispersión de los datos de manera individual; para hablar de una incertidumbre del conjunto de datos como un todo, se prefiere la desviación estándar u otra serie de ecuaciones basadas en la desviación estándar. Incluso, para hablar de una incertidumbre relativa del grupo de datos, se puede usar el **coeficiente de variación**:

$$C_v = \frac{\text{Desv. Estándar}}{\text{Promedio}} \times 100\% = \frac{s}{\bar{x}} \cdot 100\%$$

Este coeficiente ofrece una estimación rápida de la reproducibilidad (y por lo tanto, la buena aplicación) del experimento: generalmente se dice que un Cv de máximo un 5% representa una buena reproducibilidad. No obstante, en la vasta mayoría de los casos, es preferible analizar la precisión a partir del intervalo de confianza ya descrito anteriormente.

7. Cifras Significativas.

El número de cifras significativas es el número de dígitos que se necesitan para expresar científicamente un valor sin que se pierda exactitud. En la práctica se expresan todos los datos que se conocen con certeza, más un dígito final incierto que por lo general tiene un valor de +/- 1.

Por ejemplo los siguientes números todos tienen 3 cifras significativas: 0.104; 1.04; 104 y $1.04 \cdot 10^{-4}$; el 1 y el 0 intermedios se conocen con certeza, el 4 es incierto, pero significativo. Obsérvese que el exponencial NO tiene efecto sobre el número de cifras significativas.

Manejo de los ceros: los ceros que aparecen entre otros dígitos como en 10.04, son siempre significativos; los ceros iniciales en un decimal como en 0.000998 nunca son significativos, (solo sirven para señalar el punto decimal); los ceros al final del número pueden o no ser significativos, así el número 500 puede tener una, dos o tres cifras significativas, para evitar confusiones suele colocarse un punto que indica el número de cifras significativas, o mejor aún, expresar el número en notación exponencial o científica por ejemplo el número 500., con el punto, claramente tiene tres cifras significativas. También se puede expresar como $5,00 \times 10^2$.

Redondeo: para expresar el resultado de un cálculo, se aproxima hacia el número mayor si el último dígito es mayor a 5 y se aproxima hacia el número menor si el último dígito es menor que 5. Por ejemplo, si el resultado es 4.56; y debe expresarse solo con dos cifras significativas, se debe escribir 4,6; asimismo, si el resultado es 2.23; y debe expresarse con dos cifras significativas, se debe escribir 2.2. si el último dígito del resultado es el 5, deberá redondearse al número par más próximo (2,95 se aproxima a 3).

Suma y Resta: Es necesario asegurarse que el número de lugares decimales del resultado, tenga el mismo número de lugares decimales del dato que menos lugares decimales tenga. ($0.10 \text{ g} + 0.028 \text{ g} = 0.13 \text{ g}$)

Multiplicación y División: es necesario asegurarse que el número de cifras significativas del resultado sea igual al número de cifras significativas del dato que menos cifras significativas posea. ($8.62 \text{ g} / 2.0 \text{ cm}^3 = 4.3 \text{ g/cm}^3$)

Nombre: _____ Código: _____

Nombre: _____ Código: _____

1. -Las mediciones repetidas de la estatura de una persona dieron los siguientes datos:

122,38 cm; 122,31 cm; 122,35 cm; 121,26 cm; 122,32 cm; 122,30 cm; 122,32 cm; 122,35 cm; 122,40 cm y 122,32 cm;

-Los siguientes resultados se obtuvieron al medir la temperatura de una muestra de etanol:

65,4 °C; 66,6 °C; 63,3°C; 66,0 °C; 66,1 °C; 65,4 °C; 65,8 °C; 66,3 °C; 65,9 °C; 66,6 °C; 65,4 °C;

Para los anteriores grupos de datos determine:

- Media
 - Moda
 - Mediana
 - Incertidumbre relativa
 - Incertidumbre absoluta
 - Desviación estándar
 - Intervalo de confianza al 95%
 - ¿Es válido excluir el valor mayor más alejado? Desarrolle una prueba Q.
2. Convierta los siguientes números a notación científica o normal –según el caso- e indique el número de cifras significativas:
- 0,000778
 - $3,45 \times 10^{-5}$
 - 6.778×10^6
 - 0.005005
 - 3.5×10^{-3}
3. Desarrolle las siguientes operaciones teniendo en cuenta el manejo de las cifras significativas:
- $(3.556 + 1.42) - 2.55$
 - $(56.87 + 2.556) \times 0,34$
 - $(789.65 - 12.3) / 3,5$
 - $(3.56 \times 12.6) + (2.45 \times 6.8)$
 - $(23.6 / 555.78) + (3.78 \times 2.5)$
4. Se midieron los siguientes niveles de absorbancia (“qué tan oscuro se ve una solución”) en función de la concentración de permanganato de potasio:

Concentración	Absorbancia
---------------	-------------

KMnO ₄ (M)	
0	0,001
0,005	0,010
0,010	0,019
0,015	0,031
0,020	0,038
0,025	0,047

Elabore un bosquejo de la gráfica, realice la regresión lineal y determine la pendiente y el intercepto del conjunto de datos.