

Análisis de Errores

1.- Medida y concepto de error

1.1.- Valor aceptado y error absoluto de una medida

Todo proceso de medida es una aproximación del verdadero valor de lo que se pretende medir (es la estimación que hace una o varias personas valiéndose de diferentes métodos y/o aparatos). Así, si pretendemos medir una distancia, el resultado de la medida va a depender de varios factores: la sensibilidad del dispositivo empleado, el método empleado para llevar a cabo la medida, el propio observador y de un conjunto de factores difíciles de controlar que pueden originar alteraciones en el proceso de la medida. Todo ello hace que si realizamos varias medidas éstas no coincidan necesariamente, y entonces cada vez obtengamos un valor.

De todas formas, a la hora de expresar el resultado de una medida, tendremos que especificar un único valor numérico. Este valor se conoce habitualmente como *valor aceptado* de la medida. Debe entenderse que el valor aceptado de la medida no coincide con el verdadero valor de la misma, que nunca podemos conocer con exactitud. Además, se debe indicar también la incertidumbre que la acompaña. Esta incertidumbre (que se conoce como *error absoluto*) nos proporciona una indicación sobre cuán cerca está el resultado obtenido del verdadero valor. Así pues, el resultado R de una medida contendrá tres elementos: el valor aceptado de la medida, V_A , el error absoluto, ε_A , y las correspondientes unidades, y se expresa normalmente de la forma:

$$R = (V_A \pm \varepsilon_A) \text{ unidades.} \quad (1)$$

Por ejemplo, supongamos que el resultado de la medida de una longitud L es

$$L = (36.5 \pm 0.7) \text{ cm.} \quad (2)$$

Con esta forma de escribir L indicamos que no conocemos el valor verdadero de L , pero aceptamos como resultado un cierto valor $V_A = 36.5$ cm. Indicamos también que estamos seguros de que el valor verdadero de L se encuentra en algún punto del intervalo $V_A - \varepsilon_A = 35.8$ cm, $V_A + \varepsilon_A = 37.2$ cm (véase la Fig. 1).

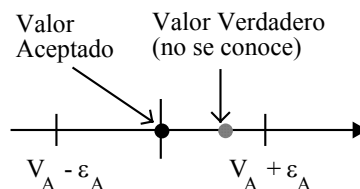


Figura 1. Significado de valor aceptado y error absoluto.

En la sección 2.2 se describe la forma de asignar valores aceptados y errores absolutos a las medidas obtenidas en el laboratorio.

1.2.- Error relativo

Otro índice de error muy útil es el *error relativo*, que se define de la forma

$$\varepsilon_r = \frac{\varepsilon_A}{V_A} \cdot 100, \quad (3)$$

que se expresa en %. Este índice se interpreta como un criterio de calidad de la medida, independientemente de la magnitud de que se trate.

En general, los errores absolutos de medidas diferentes no son comparables entre sí. Por ejemplo, imaginemos que medimos una longitud L y un tiempo t y que los resultados son

$$L = (5.2 \pm 0.4) \text{ m},$$

$$t = (4563 \pm 6) \text{ s}. \quad (4)$$

Como se trata de medidas de magnitudes diferentes, no sabemos cuál de las dos es más precisa. Para poder compararlas, usamos el error relativo. Aplicando la ecuación (3), se obtiene:

$$\varepsilon_r(L) = \frac{0.4}{5.2} \cdot 100 = 7.7\%,$$

$$\varepsilon_r(t) = \frac{6}{4563} \cdot 100 = 0.13\%, \quad (5)$$

de lo que deducimos que, en términos relativos, la medida de t es más precisa que la de L .

1.3.- Fuentes de error

Los errores de las medidas se clasifican, de acuerdo con los motivos que los originan, en sistemáticos (ε_{sis}) y accidentales (ε_{acc}).

Los primeros son originados por defectos del método o del aparato de medida y siempre actúan en el mismo sentido. Son evitables y realmente se trata de equivocaciones, ya que es una equivocación utilizar un método inadecuado o un instrumento defectuoso.

Por ejemplo, imaginemos que se pretende medir un tiempo cuyo valor verdadero (V_V) es de 28.4 s. Imaginemos que el cronómetro que se utiliza adelanta. Si, guiados por nuestra confianza en el cronómetro, lo utilizamos para efectuar tres medidas, obtendríamos probablemente valores semejantes a los siguientes

$$1^{\text{a}} \text{ medida} \rightarrow t_1 = 29.1 \text{ s}$$

$$2^{\text{a}} \text{ medida} \rightarrow t_2 = 29.5 \text{ s}$$

$$3^{\text{a}} \text{ medida} \rightarrow t_3 = 29.4 \text{ s}$$

Representando estos resultados en un eje horizontal, obtendríamos el resultado que se muestra en la Fig. 2.

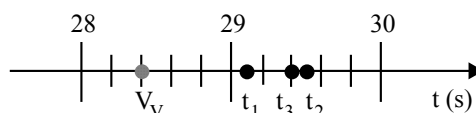


Figura 2. Ejemplo de error sistemático

En esta figura apreciamos que las tres medidas se desvían en la misma dirección y que la información que obtenemos está falseada, ya que los valores obtenidos no nos dan una indicación de por dónde se encuentra el valor verdadero.

Por el contrario, los errores accidentales no se pueden evitar ni controlar y proceden de una multitud de causas sobre las que no se puede actuar. Son fortuitos y por lo tanto se producen al azar. Esto no debe desanimarnos, ya que, como no actúan en el mismo sentido, se "compensan", y dan lugar a un resultado que proporciona una información valiosa de por dónde se encuentra el valor verdadero de la magnitud.

A modo de ejemplo, supongamos ahora que pretendemos medir la aceleración de la gravedad, g , a partir del tiempo que invierte un móvil en recorrer una altura h , en caída libre, utilizando la ecuación

$$h = \frac{1}{2} g t^2, \quad (6)$$

donde t es el tiempo de vuelo de la caída libre. Para ello, la experiencia consistirá en medir el tiempo t utilizando un cronómetro. Obviamente, la precisión de nuestras medidas dependerá de nuestra habilidad en hacer coincidir los tiempos de salida y llegada del móvil con los de puesta en marcha y parada del cronómetro, respectivamente. Parece lógico pensar que en unos casos nuestra acción de poner en marcha el cronómetro se adelantará a la de la

salida del móvil, mientras que en otros casos se retrasará. Algo similar sucederá con la acción de detener el cronómetro y la llegada del móvil. Así, si imaginamos que el verdadero valor del tiempo de vuelo es 4.45 s, parece razonable pensar que en unos casos excederemos este valor, mientras que en otros nos quedaremos cortos. Una muestra de tres medidas realizadas podría ser:

$$1^{\text{a}} \text{ medida} \rightarrow t_1 = 4.66 \text{ s}$$

$$2^{\text{a}} \text{ medida} \rightarrow t_2 = 4.54 \text{ s}$$

$$3^{\text{a}} \text{ medida} \rightarrow t_3 = 4.31 \text{ s}$$

Haciendo uso de una representación semejante a la utilizada en el anterior ejemplo, una situación como la esquematizada en la Fig. 3.

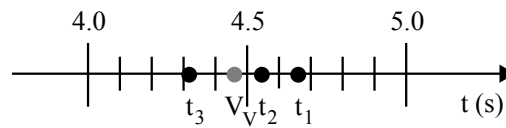


Figura 3. Ejemplo de error accidental

Ahora apreciamos que, como habíamos indicado anteriormente, los errores accidentales actúan en diferentes sentidos, de forma que el intervalo de medidas, a diferencia de lo que sucedía con el ejemplo anterior, contiene al verdadero valor.

Por todo ello, los errores más peligrosos son los sistemáticos, ya que en el caso de errores accidentales es de esperar con igual probabilidad que unas veces el resultado de la medida sea superior al verdadero valor y que otras veces sea inferior a él, por lo que algún tipo de valor medio será una buena aproximación del verdadero valor de la medida. Puede demostrarse que la dependencia funcional del error accidental con el número n de medidas efectuadas sigue una ley del tipo.

$$\text{Error accidental} \rightarrow \varepsilon_{\text{acc}} \propto \frac{1}{\sqrt{n}} \quad (7)$$

Esto significa que aumentando el número de medidas podemos reducir la incertidumbre tanto como queramos (aunque este hecho supone invertir mucho más tiempo en el proceso de medida).

De lo expuesto anteriormente, y excluidos los errores sistemáticos, puede parecer a primera vista que el error o incertidumbre puede hacerse tan pequeño como se quiera, y aumentar así la exactitud de las medidas realizadas, sin más que aumentar el número de medidas realizadas. Sin embargo, cada aparato de medida tiene una sensibilidad determinada (mínimo valor apreciable de la magnitud a determinar). Por ello, la reducción de los errores anteriores no nos eliminará la imprecisión de la sensibilidad del aparato, conocida habitualmente como *error de escala*, ε_{esc} .

En general se suele asumir que el error global que afecte a una medida está conformado por las tres fuentes de error citados y, por lo tanto, el error total en el caso más desfavorable se considerará la suma de los tres errores. Sin embargo, en general convendrá siempre realizar la experiencia habiendo eliminado los errores sistemáticos por lo que entonces en el error total sólo contribuirán el error de escala y el error accidental.

1.4.- Regla de redondeo de números. Cifras significativas

Dado el significado de cota de imprecisión que tiene el error absoluto, éste no se debe expresar nunca con más de una cifra significativa distinta de cero, salvo cuando la primera de ellas sea un 1 (10, 11, ..., 19), en cuyo caso se añadirá la siguiente cifra significativa. Entenderemos como cifras significativas de un número a todas las cifras 1, 2, 3, ..., 9, que entran en este número así como el cero, si éste se encuentra en medio del número o a la derecha del mismo. Además, en todos los casos incrementaremos la última cifra significativa no truncada en una unidad si la primera cifra eliminada es mayor o igual que 5.

Además, el valor aceptado de la magnitud debe tener sólo las cifras necesarias para que su última cifra significativa ocupe la misma posición decimal que la última del error absoluto.

Ejemplo: Expresa correctamente los siguientes datos experimentales:

- | | |
|------------------------|-------------------------|
| a) 1.234 ± 0.157 | d) 0.01683 ± 0.0058 |
| b) 3.4182 ± 0.0234 | e) 6.3 ± 0.097 |
| c) 6288 ± 1551 | f) 428.3514 ± 0.008 |

Teniendo en cuenta las reglas anteriores, el error del caso a) tendrá que expresarse con dos cifras significativas ya que la primera de ellas es un 1. Así, el error tendrá que escribirse como ± 0.16 , puesto que la segunda cifra significativa ha sido redondeada por exceso al ser la tercera mayor que 5. Además, el valor aceptado de la magnitud solamente puede contener dos cifras decimales atendiendo a la posición decimal que ocupa la última cifra significativa del error con lo cual habremos de escribir 1.23 ± 0.16 .

El error del caso b) tiene cuatro cifras decimales, de las que sólo tres son significativas. Como la primera es un 2, tenemos que dejar sólo una. Puesto que la siguiente es un tres, que es menor que cinco, el error queda simplemente ± 0.02 . Después ajustamos el valor aceptado, dejándolo con el mismo número de cifras decimales que el error. Tendremos que "cortar" a partir del uno, pero como la cifra siguiente es un 8 (que es mayor que 5), tenemos que redondear el 1 a 2. Por esta razón, la escritura correcta del caso b) es 3.42 ± 0.02 .

Ejercicio: Siguiendo un proceso análogo, obténgase el resultado para el resto de casos del ejemplo anterior.

2.- Cálculo de errores

En este apartado veremos cómo asignar el valor aceptado y el error a las medidas experimentales. Distinguiremos dos tipos de medidas: *directas* e *indirectas*. Las medidas *directas* son el resultado de la comparación de una cierta magnitud que se desea medir con otra estándar (unidad) de la misma naturaleza, empleando un instrumento de medida más o menos complejo. Mientras que las medidas *indirectas* son las obtenidas mediante el uso de ecuaciones que involucran valores de medidas directas.

2.1.- Error de escala de una medida directa

Partiremos de los supuestos de los errores sistemáticos han sido eliminados y de que los errores accidentales son despreciables. De este modo, la única fuente de error a considerar será la precisión del aparato empleado para realizar la medida. En cualquier caso, tomaremos como valor aceptado de la medida la lectura del aparato.

Cualquier aparato o dispositivo de medida tiene una precisión limitada. Por ejemplo, con una regla graduada común no se pueden apreciar décimas de milímetro, ya que la menor división corresponde a milímetros. Así pues, al efectuar una medida tenemos, como mínimo una incertidumbre que corresponde a la sensibilidad del aparato, que denominamos *error de escala* (ε_{esc}).

En los dispositivos más simples, tales como la mencionada regla graduada o un termómetro de columna de mercurio, es muy fácil asignar el error de escala: corresponde a la menor división de la escala.

Ejemplo: Imaginemos que estamos realizando una experiencia en el laboratorio en el que necesitamos conocer la temperatura de un fluido en un instante determinado. En el laboratorio, sólo se dispone de un termómetro de mercurio de $0.5\text{ }^{\circ}\text{C}$ de sensibilidad, y a la hora de realizar la medida nos encontramos con la lectura que aparece en la Fig. 4.

En la figura vemos que el error de escala es $0.5\text{ }^{\circ}\text{C}$, y la lectura es $17.5\text{ }^{\circ}\text{C}$. Luego la medida la expresamos como:

$$T = (17.5 \pm 0.5)\text{ }^{\circ}\text{C}.$$

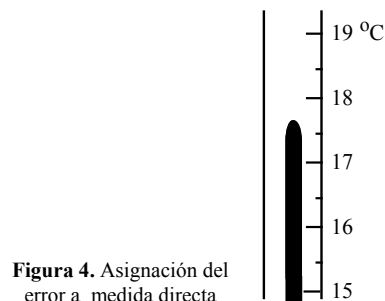


Figura 4. Asignación del error a medida directa

Ahora bien, en dispositivos más complejos, tales como instrumentos electrónicos, pueden intervenir diversos factores en el error de una medida. En tales casos, el fabricante suministra una serie de especificaciones

para determinar este error de cada medida, en función de cómo ésta haya sido realizada, y es necesario atender a estas instrucciones para determinar ese error.

Por ejemplo, en el caso de polímetros eléctricos, los errores básicos a tener en cuenta se deben a la propia imprecisión del aparato (que consiste en un cierto % de la lectura) y al redondeo que efectúa el aparato, debido a que la pantalla presenta un número de dígitos limitado. Generalmente, el fabricante dará una tabla en la que especifica el error que se comete en función del rango o fondo de escala escogido (máximo valor posible de la lectura) y de la resolución (mínimo valor absoluto, distinto de 0, que puede marcar el aparato con el rango escogido). Este error vendrá dado, típicamente, como suma de un % de la lectura y una cierta función de los dígitos de resolución. En la Tabla 1 se muestran las especificaciones del fabricante para varios polímetros Hewlett-Packard.

Ejemplo: Al medir una corriente continua (*dc current*) I con un polímetro HP E2378A, usando un fondo de escala de 300 mA, se obtiene una lectura de 165.2 en la pantalla. Encuentra el valor de la medida I con su error. Emplea para ello las especificaciones que figuran en la Tabla 1.

Function	Range	Resolution	Accuracy (%rdg + number of digits)
dc voltage	300 mV	0.100 mV	0.3 % + 2
	3 V	0.001 V	0.3 % + 2
	30 V	0.01 V	0.4 % + 1
	300 V	0.1 V	0.4 % + 1
	1000 V	1 V	0.4 % + 1
ac voltage	3 V	0.001 V	1.0 % + 3
	30 V	0.01 V	1.0 % + 3
	300 V	0.1 V	1.0 % + 3
	750 V	1 V	1.0 % + 3
dc current	300 μ A	0.1 μ A	1.0 % + 2
	3 mA	0.001 mA	1.0 % + 2
	30 mA	0.01 mA	1.0 % + 2
	300 mA	0.1 mA	1.5 % + 2
	10 A	0.01 A	1.5 % + 2
ac current	300 μ A	0.1 μ A	2.0 % + 5
	3 mA	0.001 mA	2.0 % + 5
	30 mA	0.01 mA	2.0 % + 5
	300 mA	0.1 mA	2.0 % + 5
	10 A	0.01 A	2.0 % + 5
Resistance	300 Ω	100 m Ω	0.7 % + 2
	3 k Ω	1 Ω	0.7 % + 1
	30 k Ω	10 Ω	0.7 % + 1
	300 k Ω	100 Ω	0.7 % + 1
	3 M Ω	1 k Ω	0.7 % + 1
	300 M Ω	10 k Ω	2.0 % + 1

Tabla 1: Especificaciones técnicas del polímetro HP E2378A.

En la columna que da la precisión (*Accuracy*) vemos que con el fondo de escala (*Range*) de 300 mA el error consiste en un 1.5% de la lectura más 2 dígitos.

El 1.5% de 165.2 es 2.48. Por otra parte, con el fondo de escala de 300 mA la resolución es de 0.1 mA, ya que la salida de pantalla es de 4 dígitos. Esto viene también indicado en la tabla (columna *Resolution*). Por tanto, un dígito de resolución corresponde a 0.1 mA, y 2 dígitos son 0.2 mA, de modo que el error total es:

$$\varepsilon = \left(\frac{165.2 \cdot 1.5}{100} + 0.2 \right) \text{ mA} = (2.48 + 0.2) \text{ mA} = 2.68 \text{ mA} = 3 \text{ mA},$$

en donde se ha efectuado el redondeo. Así pues, la medida debe darse como

$$I = (165 \pm 3) \text{ mA}.$$

Ejercicio: Al medir una resistencia (*Resistance*) R con un polímetro Hewlett-Packard E2378A, empleando un fondo de escala de 30 k Ω , se obtiene una lectura de 26.45 en la pantalla. Encuentra el valor de la medida de R .

2.2.- Error de una magnitud medida directamente al realizar un número n de medidas

En esta ocasión, partiremos de los supuestos de que hemos sido capaces de eliminar los errores sistemáticos y de que trabajaremos con un instrumento de medida de una gran sensibilidad. De este modo, los únicos errores que afectarán a nuestras medidas experimentales serán los errores accidentales.

Supongamos que hacemos un número n de medidas de una misma cantidad obteniendo los valores $x_1, x_2, x_3, \dots, x_n$, que difieren entre ellos como consecuencia de la única influencia de los errores aleatorios. Obviamente se nos plantean dos cuestiones: a) ¿qué valor tomamos como medida de esa cantidad?, y b) ¿cuál debe ser el índice de error que se debe asociar a esa cantidad?. Respecto a la primera pregunta, parece bastante razonable tomar el *valor medio* de todos los datos como *mejor valor* (m_v). Esto es debido a que si los errores son al azar tan probable es que ocurran por exceso como por defecto y al hacer la media quedarán compensados (si se dispone de un número suficiente de medidas). El valor medio \bar{x} , o media, está definido de la siguiente manera:

$$x_m = \bar{x} = \frac{\sum_{i=1}^n x_i}{n} . \quad (8)$$

Queda ahora por responder a la segunda pregunta, ¿qué error se tomará para la magnitud?. En principio parece bastante apropiado tomar como índice de error un número que de algún modo mida la dispersión de los datos en torno al valor medio. Obviamente la desviación con respecto al valor medio se puede conocer a partir de las diferencias entre cada dato individual y la media, $(x_i - \bar{x})$. Pero estas diferencias son positivas para unas medidas y negativas para otras, de manera que el valor medio de estas dispersiones, la desviación media \bar{d} , es nulo, es decir,

$$\bar{d} = \frac{\sum_{i=1}^n (x_i - \bar{x})}{n} = 0 .$$

Por ello, para estimar la desviación media evitando que se compensen unas desviaciones con otras, conviene tomar valores absolutos en la forma siguiente

$$\bar{d} = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n} . \quad (9)$$

De la expresión anterior se observa que la desviación media y, por lo tanto, el error van a depender a su vez del valor medio. Esto nos permite elegir un criterio más objetivo que el utilizado anteriormente a la hora de determinar el mejor valor. Este criterio podría ser escoger como *mejor valor*, m_v , aquél que minimice la desviación media de las medidas en torno a él. En términos matemáticos, el mejor valor ha de minimizar la cantidad u

$$u = \frac{\sum_{i=1}^n |x_i - m_v|}{n} . \quad (10)$$

De esta forma, el mejor valor sería aquél que distara menos de todos los datos. Matemáticamente, este problema se resuelve determinando el valor m_v que anula la primera derivada respecto de m_v de la función anterior. Sin embargo, existen dificultades matemáticas a la hora de determinar la derivada de una función afectada de un valor absoluto. Una forma de evitar este problema es trabajar con una función diferente que continúe dando información sobre la disparidad de las medidas, pero derivable en el intervalo exigido. Por ello se suele trabajar con una función, u , que proporciona el valor medio de los cuadrados de las desviaciones, es decir,

$$u = \frac{\sum_{i=1}^n (x_i - m_v)^2}{n} . \quad (11)$$

Con esta nueva función desaparecen los inconvenientes matemáticos anteriores. Si aplicamos, ahora que podemos, el criterio de elegir un mejor valor m_v que minimice la función u (a este criterio se le denomina *criterio de Gauss*), obtenemos

$$\left. \frac{du}{dm_v} \right|_{m_v} = \frac{1}{n} \sum_{i=1}^n 2(x_i - m_v) = - \left[\frac{2}{n} \sum_{i=1}^n x_i - 2m_v \right] = 0, \quad (12)$$

y despejando m_v se tiene

$$m_v = \frac{\sum_{i=1}^n x_i}{n} = \bar{x}, \quad (13)$$

obteniendo así que el valor que hace mínima la suma de las dispersiones al cuadrado es, precisamente, la media aritmética de las medidas obtenidas. Por ello, de aquí en adelante admitiremos que el *mejor valor* m_v de un conjunto de n medidas x_i es el valor medio \bar{x} , también denotado o x_m . Además, una adecuada medida de la dispersión de los datos vendrá dada por la siguiente expresión

$$\sigma^2(x) = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}, \quad (14)$$

que recibe el nombre de **varianza** o **desviación cuadrática media**. Algo más parecido a la desviación media es la raíz cuadrada positiva de la varianza

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

que recibe el nombre de **desviación típica**, o **desviación estándar** (en la bibliografía se suele expresar como SD, *standard deviation*). Cuando n es pequeño, la desviación estándar suele denotarse como $\sigma_{n-1}(x)$, o simplemente $s(x)$, definida en la forma

$$\sigma_{n-1}(x) = s(x) = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}, \quad (16)$$

fundamentalmente por dos razones: a) Porque, aunque el número de medidas es n , el número de grados de libertad es $n-1$ ya que las n medidas están relacionadas por la ecuación (13), que constituye una ecuación de ligadura y que por tanto reduce en uno el número de grados de libertad iniciales. b) Porque esta nueva definición asigna una desviación estándar indeterminada (error indeterminado) para el caso de una sola medida, mientras que la anterior asignaba una desviación estándar nula, y lo primero parece más apropiado.

En general, el parámetro $\sigma(x)$ o, en su caso, $s(x)$, representa la dispersión de la muestra x_i , es decir, una especie de desviación media de los valores x_i respecto del valor medio \bar{x} . De este modo, proporciona la probabilidad de encontrar una medida dentro del intervalo definido por $\bar{x} \pm \sigma(x)$. Pero si queremos tomar el valor medio \bar{x} como resultado de la medida, nos interesa más definir un parámetro que nos exprese la probabilidad de encontrar el valor verdadero, V_v , dentro del intervalo de error. Para ello, supongamos que realizamos para una misma magnitud, sujeta tan solo a errores aleatorios, varios grupos de observaciones de una misma medida. Como se puede suponer, los valores medios obtenidos para cada grupo, $\bar{x}^1, \bar{x}^2, \dots, \bar{x}^m$, no coincidirán, pero ellos, al igual que los valores de las medidas realizados, se situarán también alrededor del V_v , aunque es posible demostrar que dispuestos alrededor de éste de forma que su dispersión es más pequeña que la que corresponde a las medidas, tal y como se muestra en la Fig. 5.

A la dispersión que afecta a los valores medios la representaremos por $\sigma(\bar{x})$ o σ_μ , y es posible demostrar que se encuentra relacionada con la dispersión de los valores de la muestra mediante la expresión

$$\sigma^2(\bar{x}) = \sigma_\mu^2 = \frac{\sigma^2(x)}{n},$$

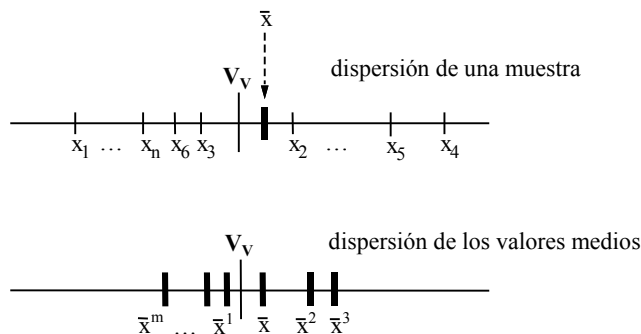


Figura 5. Estimación de la dispersión del valor medio

es decir,

$$\sigma(\bar{x}) = \sigma_{\mu} = \frac{\sigma(x)}{\sqrt{n}} \tag{17}$$

Como podemos observar, la dispersión del valor medio se puede reducir tanto como se quiera simplemente aumentando el número de medidas realizadas en la muestra puesto que σ_{μ} es inversamente proporcional a $n^{1/2}$. En definitiva, la estimación del error ϵ de la medida estará basado, en general, en la desviación estándar de la media. Un criterio habitual es tomar como error el doble del valor de la desviación estándar de la media, es decir:

$$\epsilon(\bar{x}) = 2 \sigma(\bar{x}) = \frac{2 \sigma(x)}{\sqrt{n}} \tag{18}$$

Para entender que grado de incertidumbre en la determinación de la medida proporciona este cálculo del error puede hacerse la siguiente discusión. Siempre que la medida de una magnitud, x , esté sujeta a un comportamiento aleatorio, se observa que, en el caso en que se realice un número muy elevado de medidas, los valores obtenidos se distribuyen de acuerdo con una función de distribución denominada *distribución normal* y que viene representada en la Fig. 6a.

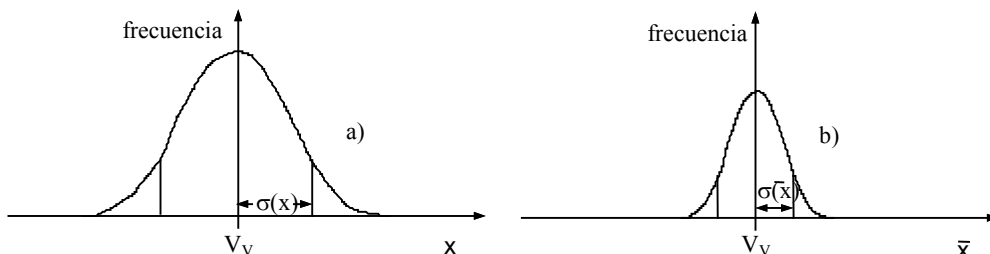


Figura 6. Representación gráfica de la distribución normal: a) para un conjunto de datos de un experimento, b) para los valores medios resultado de varios experimentos.

La función de distribución normal proporciona la probabilidad de que al realizar una medida de la magnitud se obtenga un determinado valor. En este sentido, en el supuesto teórico de que se pudiese efectuar un número ilimitado de medidas de la magnitud, el valor de la distribución normal en un determinado valor de la variable x_i proporciona la probabilidad de obtención de dicho valor. Sin embargo, en una experiencia real el número de medidas que pueden realizarse es limitado y por lo tanto obtendríamos un conjunto discreto de valores que se ubicarían a lo largo de la curva de la Fig. 6a. Si fuesen los valores medios de cada muestra los que se representarían frente a la frecuencia obtendríamos una distribución semejante a la que se representa en la Fig. 6b que resulta ser más estrecha que la distribución representada en la Fig. 6a. Esto significa que la dispersión de los valores medios es menor que la de las medidas (en un factor inversamente proporcional a $n^{1/2}$ tal y como ha sido establecido anteriormente). La función de distribución Normal, o de Gauss, es perfectamente conocida y se encuentra tipificada y tabulada. Así podemos conocer el número de medidas del valor medio cuyos valores quedan comprendidos dentro de un intervalo específico. Puede demostrarse que existe un 32.7% de probabilidad de que al realizar una medida de \bar{x} ésta se aparte del V_v una cantidad superior a $\sigma(\bar{x})$; sin embargo la

probabilidad de que el resultado obtenido difiera del V_V en una cantidad superior a $2\sigma(\bar{x})$ es sólo del 4.6%. La probabilidad de que la medida difiera en más de $3\sigma(\bar{x})$ del V_V es muy baja, tan solo el 0.3%. De este modo, el criterio para la medida del error que hemos establecido con la ecuación (17) implica que la probabilidad de que el valor verdadero de la medida, V_V , se encuentre dentro del intervalo $\bar{x} \pm \sigma(\bar{x})$ es del 95%.

Ejemplo : Se ha llevado a cabo la medida de la resistencia de una bobina, obteniéndose los siguientes valores. Determinad el valor estimado de la resistencia de la bobina, así como las dispersiones $\sigma_{n-1}(x)$ y $\sigma(\bar{x})$ del conjunto de medidas. Expresad el valor de la medida de R con su error.

resistencia (Ω)	resistencia (Ω)
5.615	5.620
5.622	5.633
5.624	5.628
5.618	5.624
5.613	----

Como hemos indicado anteriormente, tomamos como valor estimado el valor medio \bar{R} , esto es:

$$\bar{R} = \frac{\sum_{i=1}^n R_i}{n} = \frac{(5.615 + 5.622 + 5.624 + 5.618 + 5.613 + 5.620 + 5.633 + 5.628 + 5.624)}{9} = 5.622 \Omega$$

$$s(R) = \sqrt{\frac{\sum_{i=1}^n (R_i - \bar{R})^2}{n-1}} = \frac{\sqrt{(5.615 - 5.622)^2 + (5.622 - 5.622)^2 + (5.624 - 5.622)^2 + (5.618 - 5.622)^2 + (5.620 - 5.622)^2 + (5.633 - 5.622)^2 + (5.628 - 5.622)^2 + (5.624 - 5.622)^2}}{8} = 0.006 \Omega$$

$$\sigma(\bar{R}) = \frac{s(R)}{\sqrt{n}} = \frac{0.006}{\sqrt{9}} = 0.002 \Omega$$

$$\varepsilon(\bar{R}) = 2 \sigma(\bar{R}) = 0.004 \Omega$$

En definitiva, el valor estimado de R con su error es:

$$R = \bar{R} \pm \varepsilon(\bar{R}) = 5.622 \pm 0.004 \Omega$$

2.3.- Organigrama del proceso de medida.

Con el fin de ser operativos en el laboratorio a la hora de realizar una medida directa, y para que sirva de recetario, proponemos el siguiente organigrama del proceso de medida directa:

- i) En primer lugar corregiremos y eliminaremos cualquier error sistemático que pueda afectar a la experiencia que realizamos y al proceso de medida. De esta forma tan solo tendremos que considerar los posibles errores accidentales y la sensibilidad del instrumento utilizado en la medida.
- ii) Determinaremos la sensibilidad del instrumento que utilizamos para medir, a la que denominaremos *error de escala* ε_{esc} .
- iii) Efectuaremos *inicialmente 3 medidas*. En el caso de que todas ellas coincidan se tomará como valor de la magnitud el valor obtenido y como error el de escala del instrumento utilizado. En caso contrario, continuaremos con el resto de los pasos del organigrama.
- iv) Determinaremos su valor medio \bar{x} , la dispersión de las medidas $s(x)$ y la dispersión de su valor medio $s(\bar{x})$.
- v) Si alguna de las medidas obtenidas es tal que $|x_i - \bar{x}| > 2s(x)$, la rechazaremos, entendiendo por lo que hemos expuesto hasta ahora que su elevada dispersión no se debe a causas accidentales; por ello,

procederemos a realizar otra medida que la sustituya, calculando para la nueva muestra sus correspondientes valor medio \bar{x} , dispersión de las medidas $s(x)$ y dispersión de la media $s(\bar{x})$.

- vi)** Compararemos el doble de la dispersión del valor medio $s(\bar{x})$, con el error de escala del aparato ϵ_{esc} , y procederemos de la siguiente forma:
- a) Si $2s(\bar{x}) \gg \epsilon_{esc}$ efectuaremos 3 medidas más y volveremos al paso iv).
- b) Si $2s(\bar{x}) \approx \epsilon_{esc}$ no realizaremos más medidas, y tomaremos como error el mayor de los dos índices comparados: $s(\bar{x})$ o ϵ_{esc} .
- vii)** Tomaremos como valor de la medida el valor medio \bar{x} , y como índice de error el mayor de entre $2s(\bar{x})$ y el ϵ_{esc} expresando el resultado final en la forma: $\bar{x} \pm 2s(\bar{x})$, o bien: $\bar{x} \pm \epsilon_{sc}$ según el caso, indicando en cualquiera de ellos las unidades que hemos utilizado.

Ejemplo: Medida del periodo de un péndulo con un cronómetro que aprecia 5 centésimas de segundo

Aplicando el organigrama anterior, debemos llevar a término los siguientes pasos:

i) Eliminamos errores sistemáticos posibles:

- a) Comprobamos, que el cronómetro que vamos a utilizar no atrasa ni adelanta. Para ello comparamos su funcionamiento con otro "reloj" cuyo funcionamiento sea correcto.
- b) Para ganar precisión en lugar de medir el tiempo de una oscilación, medimos el de 50 oscilaciones de forma que si introducimos un error en el proceso de medida, éste tendrá menos peso
- c) Sincronizamos (manual o electrónicamente) los momentos de puesta en marcha y parada del cronómetro con los de inicio de la primera oscilación, y la finalización de la oscilación 50 del péndulo respectivamente.
- d) Evitamos corrientes de aire u otras causas que puedan perturbar de forma notoria la trayectoria del péndulo.

ii) Tal como se ha expresado anteriormente el error de escala del aparato coincide con la sensibilidad del cronómetro utilizado, esto es:

$$\epsilon_{esc} = 0.05 \text{ s.}$$

iii) Realizamos tres medidas para 50 oscilaciones en los tres casos. Supongamos que obtenemos:

$$x_1 = 96.44 \text{ s,}$$

$$x_2 = 96.12 \text{ s,}$$

$$x_3 = 96.37 \text{ s.}$$

iv) Calculamos los parámetros \bar{x} , $s(x)$ y $s(\bar{x})$, y obtenemos

$$\bar{x} = \frac{(96.44 + 96.12 + 96.37)}{3} = 96.31 \text{ s,}$$

$$s(x) = \sqrt{\frac{(96.44 - 96.31)^2 + (96.12 - 96.31)^2 + (96.37 - 96.31)^2}{2}} = 0.1682 \text{ s,}$$

$$s(\bar{x}) = \frac{s(x)}{\sqrt{3}} = 0.0971 \text{ s.}$$

v) En este caso ninguna medida difiere del valor medio en una cantidad superior a $2s(x) = 2 \cdot 0.1682 = 0.3364 \text{ s}$, luego no es necesario desestimar ninguna de las medidas que hemos realizado.

vi) Comparamos $2s(\bar{x})$ con el error de escala del aparato y vemos que $2s(\bar{x}) = 0.1942 \text{ s}$, y que $\epsilon_{esc} = 0.05 \text{ s}$, por lo tanto $2s(\bar{x}) \gg \epsilon_{esc}$. Ello significa que debemos aumentar el número de medidas, hacer 3 medidas más, y retornar a iv). Supongamos que al hacer las tres medidas, la muestra completa queda de la siguiente manera

$$x_1 = 96.44 \text{ s,}$$

$$x_4 = 96.31 \text{ s,}$$

$$x_2 = 96.12 \text{ s,}$$

$$x_5 = 96.30 \text{ s,}$$

$$x_3 = 96.37 \text{ s,}$$

$$x_6 = 92.50 \text{ s.}$$

Al calcular ahora \bar{x} , $s(x)$ y $s(\bar{x})$, utilizando las ecuaciones (13), (15) y (16) respectivamente, obtenemos

$$\bar{x} = 95.6733 \text{ s,}$$

$$s(x) = 1.5583 \text{ s,}$$

$$s(\bar{x}) = 0.6362 \text{ s.}$$

Volvemos a aplicar v) y vemos que la medida x_6 difiere del valor medio calculado en más de $2s(x)$ (efectivamente $95.6733 - 92.5000 = 3.1733 > 2 \cdot 1.5583 = 3.1166$). Por tanto desestimamos la medida x_6 y realizamos una nueva medida. Supongamos que después de realizar la medida la muestra queda de la siguiente manera:

$$x_1 = 96.44 \text{ s,}$$

$$x_4 = 96.31 \text{ s,}$$

$$x_2 = 96.12 \text{ s,}$$

$$x_5 = 96.30 \text{ s,}$$

$$x_3 = 96.37 \text{ s,}$$

$$x_6 = 96.31 \text{ s.}$$

De acuerdo con el organigrama volvemos al punto iv), y calculamos de nuevo \bar{x} , $s(x)$ y $s(\bar{x})$, obteniendo:

$$\bar{x} = 96.3083 \text{ s,}$$

$$s(x) = 0.1065 \text{ s,}$$

$$s(\bar{x}) = 0.0435.$$

Ahora vemos que no existe ningún valor que se desvíe del valor medio en una cantidad superior a $2s(x)$, lo que significa que no hay que sustituir ningún valor de la muestra, de acuerdo con el criterio v). A continuación volvemos (según el punto vi) a comparar los valores del error de escala y de $2s(\bar{x})$ y obtenemos:

$$2s(\bar{x}) = 0.087 \text{ s,}$$

$$\epsilon_{\text{esc}} = 0.05 \text{ s,}$$

con lo que $2s(\bar{x})$ y ϵ_{esc} , son del mismo orden, lo que significa que no hay que realizar más medidas, y dado que $0.087 \text{ s} > 0.05 \text{ s}$ tomaremos como índice de error $2s(\bar{x})$.

vii) El valor final de la medida (el tiempo que invierte el péndulo en realizar 50 oscilaciones) se expresa como $\Delta t = (96.31 \pm 0.09) \text{ s}$. De esta forma el periodo es (ver errores de medidas indirectas en el capítulo 3):

$$T = \frac{\Delta t}{50} = (1.9262 \pm 0.0018) \text{ s.}$$

2.4.- Error de una medida indirecta

Es usual que una magnitud física no se determine directamente con una medida, sino a través de una fórmula, gráfica, tabla, es decir, a partir de otras magnitudes (que a su vez habrán sido medidas directamente o no). Se habla entonces de una *medida indirecta*. Por ejemplo, podemos usar la longitud y el período de las pequeñas oscilaciones de un péndulo para determinar la aceleración de la gravedad g , que en tal caso se habrá medido indirectamente. Obviamente, la precisión con que se ha determinado g dependerá de algún modo de la de la longitud y el período.

Ejemplo: Para entender mejor este efecto de "propagación" de los errores, consideremos un ejemplo geométrico. Normalmente, no medimos directamente el área S de un rectángulo, sino que determinamos la longitud de cada lado a y b , y usamos la expresión

$$S = a \cdot b.$$

Como la medida de S no es directa, hay que asignarle algún error a partir de los errores ϵ_a y ϵ_b . Es fácil darse cuenta de que el valor verdadero de S estará, con seguridad, comprendido entre los dos valores siguientes:

$$(a + \epsilon_a)(b + \epsilon_b) = a b + (b \epsilon_a + a \epsilon_b);$$

$$(a - \epsilon_a)(b - \epsilon_b) = a b - (b \epsilon_a + a \epsilon_b),$$

en donde se ha despreciado el término pequeño $\epsilon_a \epsilon_b$. Así pues, es natural escribir el resultado como

$$S = a b \pm \epsilon_S; \quad \epsilon_S = b \epsilon_a + a \epsilon_b.$$

Los dos métodos más utilizados para calcular el error de una medida indirecta son los conocidos como método de los logaritmos y método de las derivadas parciales. Veamos por separado cada uno de estos métodos.

El método de los logaritmos

El resultado del ejemplo anterior nos recuerda la regla de diferenciación del producto. De hecho, siempre que la fórmula para determinar la magnitud R sea de la forma

$$R = aX^bY^c, \quad (19)$$

siendo $a, b, c...$ constantes y $X, Y...$ las magnitudes de partida con error conocido, puede emplearse un algoritmo rápido de determinación del error, que consiste en tomar logaritmos neperianos en la ecuación inicial

$$\ln R = \ln a + b \ln X + c \ln Y, \quad (20)$$

diferenciando a continuación,

$$\frac{dR}{R} = b \frac{dX}{X} + c \frac{dY}{Y}, \quad (21)$$

y considerando que la "regla de propagación de errores" coincide con la "regla de propagación de diferenciales", con lo que finalmente se obtiene

$$\frac{\varepsilon_R}{R} = \left| b \frac{\varepsilon_X}{X} \right| + \left| c \frac{\varepsilon_Y}{Y} \right|, \quad (22)$$

es decir, una expresión para el error relativo de R en términos de los errores relativos de las magnitudes de partida. En la ecuación (21) se ha escrito el módulo de cada sumando para tener en cuenta el hecho de la propagación de errores siempre es aditiva.

Ejercicio: Obténgase el resultado del ejemplo anterior empleando el método de los logaritmos.

Ejemplo: El periodo de las pequeñas oscilaciones de un péndulo simple de longitud $l = (22.1 \pm 0.1)$ cm es de $T = (0.943 \pm 0.006)$ s. ¿Qué valor de la aceleración de la gravedad g puede aceptarse a partir de estos datos?

Considerando el péndulo como un oscilador armónico simple, podemos emplear la conocida relación

$$g = 4\pi^2 \frac{l}{T^2},$$

de donde resulta

$$g = 4\pi^2 \frac{0.221 \text{ m}}{(0.943 \text{ s})^2} = 9.811 \text{ m/s}^2.$$

Tomando logaritmos, diferenciando y sustituyendo diferenciales por errores obtenemos

$$\begin{aligned} \frac{\varepsilon_g}{g} &= \frac{\varepsilon_l}{l} + 2 \frac{\varepsilon_T}{T} = \frac{0.1}{22.1} + 2 \frac{0.006}{0.943} = \\ &= 4.5 \cdot 10^{-3} + 1.27 \cdot 10^{-2} = 0.017, \end{aligned}$$

Sabemos así que la constante g se ha determinado, aproximadamente, con un 2% de error. El error absoluto de g es por tanto

$$\varepsilon_g = 0.017 (9.81 \text{ m/s}^2) = 0.17 \text{ m/s}^2,$$

con lo que el valor que se acepta para g resulta ser

$$g = (9.81 \pm 0.17) \text{ m/s}^2$$

Como acabamos de ver con este ejemplo, el método de los logaritmos es muy sencillo de emplear. Por esta razón, lo usaremos siempre que sea posible. Sin embargo, como no todas las ecuaciones son de la forma (19), existen situaciones en las que no es posible hacerlo (por ejemplo, en el caso muy común de que aparezcan sumas). En estos casos, se puede emplear el método de las derivadas parciales, que pasamos a describir a continuación.

Método de las derivadas parciales

Supongamos que una magnitud R depende de otras magnitudes X, Y, \dots , de las que se conocen sus valores aceptados y sus errores, a través de una función arbitraria de la forma

$$R = R(X, Y, \dots) \quad (23)$$

Diferenciando la expresión anterior, obtendremos

$$dR = \frac{dR}{dX} dX + \frac{dR}{dY} dY + \dots \quad (24)$$

De un modo similar a como se procedió en el método de los logaritmos, se supone ahora que la "regla de propagación de errores" coincide con la "regla de propagación de diferenciales", es decir, se sustituye los diferenciales en la ecuación (24) por los respectivos errores absolutos. Sin embargo, después de hacer esto debe tomarse el "módulo" de ε_R y considerar que el módulo de la suma es igual a la suma de los módulos. Así, la expresión final para el error indirecto (que admite una deducción rigurosa) es

$$\varepsilon_R = \left| \frac{dR}{dX} \varepsilon_X \right| + \left| \frac{dR}{dY} \varepsilon_Y \right| + \dots \quad (25)$$

Ejemplo: Determinese el valor que debe aceptarse para g , determinada con los datos del Ejemplo anterior, si se emplea el procedimiento de las derivadas parciales.

Aplicando la expresión general (24) a la ecuación para g , se obtiene

$$\begin{aligned} \varepsilon_g &= \left| \frac{dg}{dt} \varepsilon_t \right| + \left| \frac{dg}{dT} \varepsilon_T \right| + \dots = \\ &= \left| \frac{4\pi^2}{T^2} \varepsilon_t \right| + \left| \frac{8\pi^2}{T^3} \varepsilon_T \right| = \left| \frac{4\pi^2 l}{0.943^2} 0.001 \right| + \left| \frac{8\pi^2 \cdot 0.221}{0.943^3} 0.006 \right| = \\ &= 4.44 \cdot 10^{-2} + 1.25 \cdot 10^{-1} = 0.17 \end{aligned}$$

Por tanto, el valor aceptado para g es ahora

$$g = (9.81 \pm 0.17) \text{ m/s}^2$$

El método de las derivadas parciales es completamente general, y por ello puede también utilizarse en los casos en que es posible emplear el método de los logaritmos. Insistiremos de todas formas en que, como hemos visto en el ejemplo anterior, el método de las derivadas parciales conlleva un proceso de cálculo más laborioso que el de los logaritmos. Por ello, y debido a lo cómodo que resulta su empleo, utilizaremos el método de los logaritmos en los casos en los que pueda llevarse a cabo.

3. Representaciones gráficas y tablas

3.1.- Normas para la representación gráfica de datos experimentales

El método gráfico es muy conveniente e intuitivo para representar los datos experimentales. Al estudiar una dependencia funcional cualquiera, es conveniente escribir los resultados en forma de tabla, donde a cada valor de un parámetro x corresponde un valor determinado de otro parámetro y . Normalmente, conviene también construir la gráfica correspondiente a esta tabla.

Para concretar, imaginemos que los valores que han resultado de un experimento en el cual se ha determinado la diferencia de potencial, V , entre los extremos de un condensador, y , en función del tiempo, x , cuando se coloca el condensador en un circuito de corriente continua que contiene una resistencia de valor conocido, son

$y[V (V)]$	34.7	14.4	8.5	4.2	2.5	1.6
$x[t (s)]$	0	50	100	150	200	250

Si se toma el sistema de coordenadas rectangulares y se trazan por el eje de ordenadas los valores del voltaje, y por el eje de abscisas los valores del tiempo, se obtiene un sistema de puntos en el plano (ver Fig. 7).

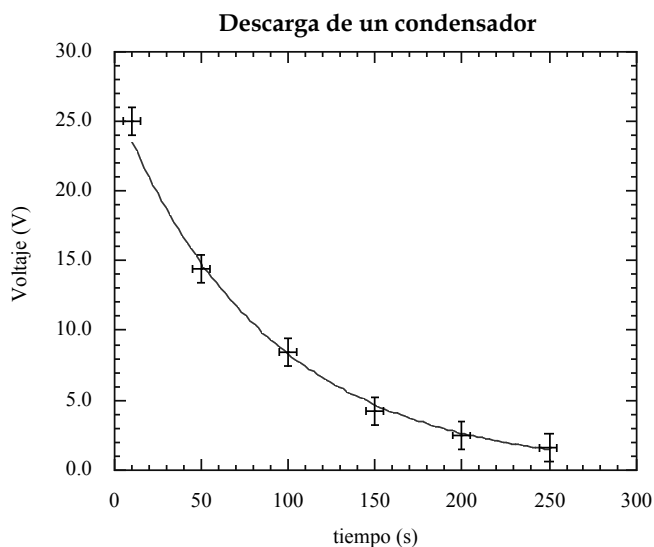


Figura 7. Ejemplo de representación gráfica de resultados experimentales.

Las gráficas deben elaborarse siguiendo los siguientes criterios:

Cada gráfica que elaboremos contará, como en el ejemplo anterior, con un breve título que se refiera a la experiencia física a la que corresponden los datos. Alternativamente, puede ir acompañada de un pie de figura numerado, en el que se aparezca éste u otro tipo de información (descripción de diferentes símbolos utilizados, condiciones de medida, ...). Por su parte, el título de cada uno de los ejes contendrá la información sobre la magnitud representada y las unidades en la que se expresa.

La escala de los ejes será la adecuada para que los puntos ocupen el máximo espacio en la gráfica. Las divisiones del eje, regularmente espaciadas, indicarán el valor correspondiente, pero no así las posibles subdivisiones. La escala no tiene por qué comenzar en las coordenadas (0,0), ya que en muchos casos conviene trasladar el origen a un punto arbitrario para utilizar totalmente el área de la gráfica. En el caso de existir diferencias de varios órdenes de magnitud en los valores de los datos representados, conviene adoptar una escala logarítmica en vez de lineal.

Como norma general, en la gráfica han de ser claramente apreciables los puntos experimentales y para ello se indica directamente en la gráfica el error de las coordenadas del punto, trazando por él una línea vertical y otra horizontal, de modo que la distancia del punto hasta los extremos de estas líneas es igual al error en la correspondiente coordenada (ver Fig. 7).

3.2.- Interpolación en una gráfica

Es frecuente que se necesite obtener valores de algunas magnitudes físicas a partir de gráficas. Así, por ejemplo, considérese el caso anterior en que tenemos una gráfica que representa los valores de la diferencia de potencial entre los extremos de un condensador en función del tiempo para un circuito de corriente continua RC.

Nuestro objetivo será determinar el valor del voltaje, V , para un determinado valor del tiempo no coincidente con ninguno de los tabulados para la construcción de la gráfica. Para ello se localizará sobre el eje de abscisas dicho valor temporal, t , y se representará sobre él con su correspondiente intervalo de error $[t-\varepsilon_t, t+\varepsilon_t]$. A continuación se trazarán sobre el valor de t y los extremos del intervalo de error sendos segmentos

paralelos al eje de ordenadas hasta determinar su punto de intersección con la gráfica que representa V en función del tiempo, tal y como se muestra en la Fig. 8. A partir de dicho punto de intersección se trazarán ahora segmentos paralelos al eje de abscisas y se determinará su punto corte con el eje de ordenadas. El valor de V se obtiene como el valor leído sobre el eje de ordenadas para el punto de corte que parte del valor de t mientras que el error de V se obtendrá como la semidiferencia del intervalo de error que aparecerá sobre el eje considerado, es decir, como $\varepsilon_V = (V_M - V_m)/2$, ver Fig. 8.

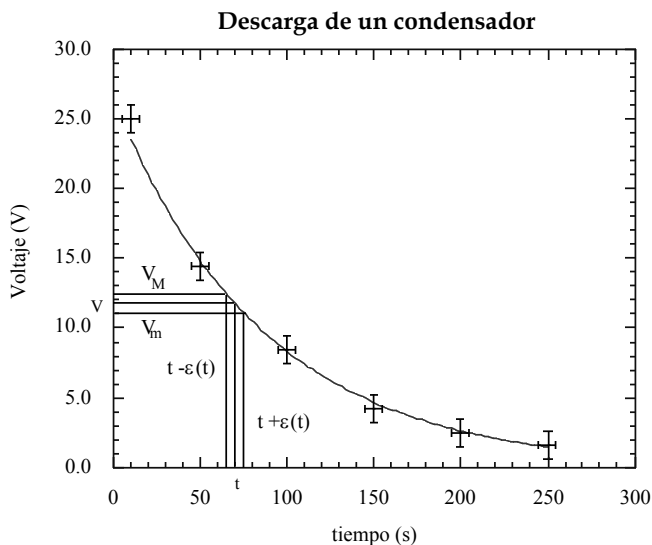


Figura 8. Interpolación en una gráfica y determinación del intervalo de error.

3.3.- Interpolación lineal en una tabla.

Es frecuente que se necesiten obtener valores de algunas magnitudes físicas a partir de tablas numéricas. Nosotros vamos a considerar el caso de tablas de simple entrada que son aquellas en que la variable dependiente, y , es sólo función de una variable independiente, x , es decir, $y = f(x)$. Así, por ejemplo, considérese el caso en que aparecen tabulados los valores de la viscosidad del agua, η , medida en centipoises para diferentes temperaturas, T , medidas en grados centígrados, tal y como se muestra en la siguiente tabla

$x[T(^{\circ}\text{C})]$	0	5	10	15	20
$y[\eta(\text{cp})]$	1.7865	1.5138	1.3037	1.1369	1.0019

Nuestro objetivo será determinar el valor de y (que para el ejemplo propuesto es la viscosidad del agua) para un determinado valor de x , en nuestro caso la temperatura. En el caso en que el valor de x coincida con alguno de los valores de la tabla se tomará como valor de y el valor tabulado y se asignará un error de ± 1 en la última cifra decimal que aparezca tabulada. Así, en el ejemplo propuesto si nos preguntan cuál es el valor de la viscosidad del agua para una temperatura de 10°C , la respuesta sería que ésta es $\eta = (1.3037 \pm 0.0001)$ cp.

Estamos ahora interesados en encontrar el valor de y para un valor de x no incluido exactamente en la tabla. Para ello, se comienza por encontrar aquellos valores de x entre los que se encuentra nuestro valor no tabulado. De este modo, la tabla presentará entonces la forma

x_1	x_2
y_1	y_2

Considerando que para el intervalo de x_1 a x_2 la expresión $y=f(x)$ puede asimilarse a una recta, podremos escribir

$$y = y_1 + \frac{y_2 - y_1}{x_2 - x_1}(x - x_1), \quad (26)$$

que permitirá determinar y en función de x o viceversa. Aplicando, por ejemplo, el método de las derivadas parciales el error de y , $\varepsilon(y)$, vendrá dado por la expresión

$$\varepsilon_y = \left| \frac{y_2 - y_1}{x_2 - x_1} \right| \varepsilon_x. \quad (27)$$

Así, si en el ejemplo propuesto estamos interesados en conocer el valor de la viscosidad del agua a una temperatura de (12 ± 1) °C, aplicando las expresiones anteriores resulta para él un valor

$$\eta = 1.3037 + \frac{1.1369 - 1.3037}{15 - 10}(12 - 10) = 1.23698 \text{ cp}, \quad (28)$$

con un error de

$$\varepsilon_\eta = \left| \frac{1.1369 - 1.3037}{15 - 10} \right| 1 = 0.033 \text{ cp}, \quad (29)$$

con lo cual, expresado correctamente, resultaría $\eta = (1.24 \pm 0.03) \text{ cp}$.

4.- Ajuste lineal por mínimos cuadrados

4.1.-Ajuste de una recta por mínimos cuadrados

En un apartado anterior nos hemos ocupado de la manera de obtener el mejor valor de una magnitud a partir de una o varias medidas o conjuntos de medidas. Un problema diferente es determinar la relación funcional entre dos magnitudes x e y como resultado de un experimento. Supongamos que, por razones teóricas bien fundadas, sabemos que entre x e y existe la relación lineal

$$y = ax + b, \quad (30)$$

y estamos interesados en determinar los mejores valores de los parámetros a y b compatibles con un conjunto de n medidas y_i correspondientes a los n valores, también experimentales, x_i . Para concretar, imaginemos que los valores que han resultado de un experimento en el cual se ha determinado la diferencia de potencial entre los extremos de una resistencia (y_i), cuando la recorren corrientes de diferente intensidad (x_i) son los siguientes

y_i (V)	1.5	2.5	4.0	3.6	5.9	6.1
x_i (A)	1.0	2.0	3.0	4.0	5.0	6.0

Como la relación existente entre los datos experimentales es lineal (ley de Ohm). La forma más simple de resolver nuestro problema es utilizar un método gráfico. Representaremos para ello los resultados experimentales tal y como muestra la Fig. 9.

El método más inmediato para evaluar los valores de a y b consiste en dibujar "a ojo" la recta que "parece acercarse lo máximo posible" a los puntos experimentales. Decidida la recta la evaluación de su pendiente a y su ordenada en el origen b es inmediata ($a = \text{tg } \alpha$ y b es el punto de corte de la recta con el eje de las ordenadas). Así para el caso particular que nos ocupa obtendríamos

$$y = 0.9x + 0.8,$$

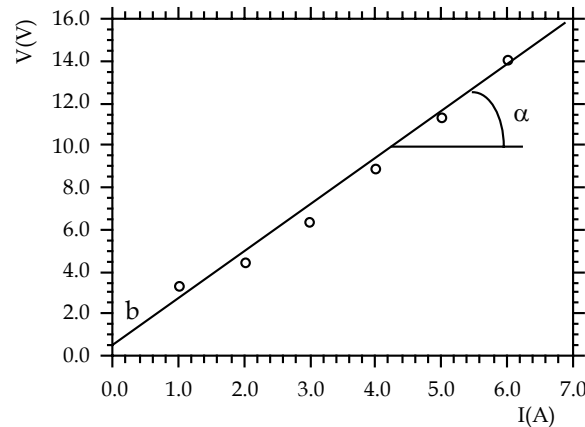


Figura 9. Representación gráfica del voltaje en función de la intensidad

Sin embargo, este método es muy subjetivo, pues dos experimentadores distintos probablemente dibujarían rectas diferentes para ajustar un mismo conjunto de datos. Por este motivo, conviene utilizar un método objetivo para obtener los parámetros de la relación buscada.

Para ello, supondremos conocidos los valores de a y b con lo cual es posible calcular, para cada uno de los datos experimentales, las desviaciones de los valores experimentales y_i respecto de los calculados a partir de la ecuación (9) para $x = x_i$, a los que denotaremos como y'_i . Esta diferencia es

$$\Delta_i = y_i - y'_i, \quad (31)$$

Un buen criterio para la determinación de los valores de a y b consiste en minimizar la suma de los cuadrados de estas diferencias, lo que garantiza que la suma de las desviaciones de los puntos experimentales respecto a los ajustados es mínima. Es decir, hemos de exigir que a y b minimicen la cantidad

$$R = \sum_{i=1}^n \Delta_i^2 = \sum_{i=1}^n (y_i - y'_i)^2 = \sum_{i=1}^n (y_i - ax_i - b)^2. \quad (32)$$

Para ello basta con exigir que las derivadas parciales de la anterior función con respecto a ambos parámetros sean iguales a cero; es decir,

$$\frac{\partial R}{\partial a} = 0, \quad \frac{\partial R}{\partial b} = 0. \quad (33)$$

Las dos condiciones expresadas en la ecuación anterior se traducen en un sistema lineal de dos ecuaciones con dos incógnitas, los parámetros a y b , que resulta ser

$$\begin{aligned} a \sum x^2 + b \sum x &= \sum xy, \\ a \sum x + bn &= \sum y, \end{aligned} \quad (34)$$

siendo n el número de datos. Estas ecuaciones se conocen como ecuaciones normales para la determinación de a y b . En estas ecuaciones hemos omitido los subíndices i para simplificar la escritura, pero hay que entender que

$$\sum x = \sum_{i=1}^n x_i \quad (35)$$

e igual para el resto de los sumatorios. Las soluciones de las ecuaciones normales son

$$a = \frac{n \sum xy - \sum x \sum y}{n \sum x^2 - (\sum x)^2}, \quad b = \frac{\sum y - a \sum x}{n}. \quad (36)$$

Volviendo al ejemplo propuesto anteriormente, tenemos que

$$n = 6, \quad \sum x = 21 \text{ A}, \quad \sum x^2 = 91 \text{ A}^2, \quad \sum y = 23.6 \text{ V}, \quad \sum xy = 99 \text{ AV},$$

con lo cual $a = 0.94 \text{ V/A}$ y $b = 0.65 \text{ V}$, por lo que la relación lineal

$$y = 0.94x + 0.65$$

es el mejor ajuste de una recta a los datos experimentales según el método de los mínimos cuadrados. Recuerdese que por el método gráfico y a ojo habíamos obtenido $a = 0.9 \text{ A/V}$ y $b = 0.8 \text{ V}$ valores muy aproximados a los que hemos obtenido de este modo objetivo.

4.2.- Determinación de los errores en la pendiente y la ordenada en el origen

Los valores que se obtienen por el método de los mínimos cuadrados son, evidentemente, inexactos porque se han determinado a partir de n pares de datos los cuales son imprecisos. Para el cálculo del error asociado a ambos parámetros podría, en principio, utilizarse un método subjetivo consistente en la determinación de éstos a partir de las pendientes y ordenadas en el origen de una pareja extrema de rectas que estimen "a ojo" los valores máximos y mínimos de a y b compatibles con los datos experimentales de que se dispone. De este modo tomaríamos como errores de a y b la semidiferencia de las pendientes y ordenadas en el origen de estas rectas.

Al igual que en el apartado anterior es conveniente, no obstante, disponer de un criterio objetivo que permita estimar siempre del mismo modo el error asociado a ambos parámetros de la recta (el anterior método además sobrestima los valores del error en un factor \sqrt{n}). Un razonamiento estadístico riguroso que queda fuera del alcance de este manual proporciona para estos errores las siguientes expresiones

$$\varepsilon_a = \left(\frac{s^2(y)}{\sum (x_i - \bar{x})^2} \right)^{1/2}, \quad \varepsilon_b = \left(\frac{s^2(y)}{n} \right), \quad (37)$$

en donde \bar{x} no es más que el valor medio de los datos x_i y en donde

$$s^2(y) = \frac{\sum (y_i - ax_i - b)^2}{n-2}, \quad (38)$$

está relacionado con las desviaciones al cuadrado de los valores experimentales y_i respecto a los calculados a partir de la recta ajustada, $ax_i + b$. Así, para el ejemplo que nos ocupa ocurre que

$$s^2(y) = 0.322 \text{ V}^2,$$

con lo cual

$$\varepsilon_a = 0.14 \text{ V/A}, \quad \varepsilon_b = 0.23 \text{ V},$$

son los valores estimados para los errores de ambos parámetros. Por lo tanto, la expresión correcta de los parámetros a y b , con sus respectivos errores, sería $a = (0.94 \pm 0.14) \text{ V/A}$ y $b = (0.65 \pm 0.23) \text{ V}$.

En estadística se define el coeficiente de correlación lineal r para cualquier conjunto de parejas de valores de la siguiente manera

$$r = \frac{n \sum xy - \sum x \sum y}{\left(n \sum x^2 - (\sum x)^2 \right) \left(n \sum y^2 - (\sum y)^2 \right)} \quad (39)$$

Puede demostrarse que los valores de r están comprendidos entre $+1$ y -1 . Los valores positivos corresponden a rectas crecientes y los negativos a rectas decrecientes. Además, el valor del coeficiente de correlación nos indica

el grado de alineación de nuestros puntos experimentales. El valor $r=1$ (-1 en el caso de rectas decrecientes) corresponde a un conjunto de puntos perfectamente alineados. Para puntos menos alineados el valor de r es menor. De este modo, también es posible estimar el error de los parámetros de la recta en términos de este coeficiente de correlación. Según este método, los errores de a y b quedan

$$\varepsilon_a = a \sqrt{\frac{1}{r^2} - 1}, \quad \varepsilon_b = \varepsilon_a \sqrt{\frac{\sum x^2}{N}}, \quad (40)$$

donde r viene dado por la ecuación (38).

4.3.- Ajuste de otro tipo de funciones mediante linearización

El caso de una relación lineal que hemos tomado como ejemplo no es tan especial como podría pensarse porque muchas relaciones funcionales de interés pueden linearizarse con algún truco. He aquí algunos ejemplos

$$y = c e^{dx} \quad \rightarrow \quad \ln y = \ln c + dx$$

$$y = c d^x \quad \rightarrow \quad \ln y = \ln c + x \ln d$$

$$y = \frac{1}{c + dx} \quad \rightarrow \quad \frac{1}{y} = c + dx.$$

Así, por ejemplo, para convertir una relación de tipo exponencial, tal como la expresada en la primera de las anteriores ecuaciones, en una relación lineal basta con tomar logaritmos neperianos en ambos lados de la igualdad y, posteriormente, ajustar linealmente por mínimos cuadrados la variable $y' = \ln(y/c)$ frente a la variable independiente x .

4.4.- Ajuste de datos mediante ordenador

El ajuste de conjuntos de datos es un procedimiento engorroso y totalmente repetitivo. Por ello, es habitual usar un programa de ordenador diseñado a este efecto, en el cual generalmente se insertarán los datos x e y , tras lo cual el programa nos suministrará todos los resultados de interés (valores de a y b con sus respectivos errores). En el mercado existen muchos programas de tratamiento de datos que permiten llevar a cabo de forma automática ajustes por mínimos cuadrados. Nosotros emplearemos los programas denominados 'SigmaPlot' o 'Kaleidagraph'.