

DETERMINACIÓN DE ERRORES Y TRATAMIENTO DE DATOS

I. Introducción.

II. Error en una medida: determinación y expresión de errores.

- 1. Clasificación de los errores.**
- 2. Exactitud, precisión y sensibilidad.**
- 3. Error absoluto y error relativo.**
- 4. Expresión del error.**
- 5. Determinación de errores en medidas directas.**
- 6. Determinación de errores en medidas indirectas.**

III. Tratamiento de datos.

- 1. El método de los mínimos cuadrados.**
- 2. Construcción de gráficas.**
- 3. Interpolación en tablas.**

I. INTRODUCCIÓN

Las medidas experimentales están afectadas de cierta imprecisión en sus valores debido a las imperfecciones del aparato de medida o a las limitaciones de nuestros sentidos en el caso de que sean ellos los que deben registrar la información. El valor de las magnitudes físicas se obtiene experimentalmente efectuando una medida; ésta puede ser directa sobre la magnitud en cuestión o indirecta, es decir, obtenida por medio de los valores medidos de otras magnitudes ligadas con la magnitud problema mediante una fórmula física. Así pues, resulta imposible llegar a conocer el valor exacto de ninguna magnitud, ya que los medios experimentales de comparación con el patrón correspondiente en las medidas directas viene siempre afectado de imprecisiones inevitables. El problema es establecer los límites dentro de los cuales se encuentra dicho valor.

El principal objetivo de estos apuntes es presentar al estudiante algunos conceptos básicos de la denominada Teoría de Errores; con ello, se pretende que el alumno se desenvuelva con agilidad en las diversas prácticas, permitiéndole reconocer los factores que influyen en el error, así como el cálculo del mismo. Además, se ofrecen algunas nociones sobre tratamiento de datos que incluye el ajuste de rectas mediante el método de mínimos cuadrados.

II. ERROR EN UNA MEDIDA Y DETERMINACIÓN Y EXPRESIÓN DE ERRORES.

1. Clasificación de los errores

El error se define como la diferencia entre el valor verdadero y el obtenido experimentalmente. Los errores no siguen una ley determinada y su origen está en múltiples causas. Atendiendo a las causas que los producen, los errores se pueden clasificar en dos grandes grupos: errores sistemáticos y errores accidentales.

Los errores sistemáticos son aquellos que permanecen constantes a lo largo de todo el proceso de medida y, por tanto, afectan a todas las mediciones de un modo definido y es el mismo para todas ellas; se pueden subclasificar en errores instrumentales, personales o por la elección del método. Los errores instrumentales son los debidos al aparato de medida; por ejemplo, un error de calibrado generaría este tipo de imprecisión. Los errores personales se deben a las limitaciones propias del experimentador; así, una persona con algún problema visual

puede cometer errores sistemáticos en la toma de ciertos datos. Finalmente, el error en la elección del método se presenta cuando se lleva a cabo la determinación de una medida mediante un método que no es idóneo para tal fin; por ejemplo, la medida del tiempo de caída de un objeto por mera inspección visual.

Los errores accidentales son aquellos que se producen en las variaciones que pueden darse entre observaciones sucesivas realizadas por un mismo operador. Estas variaciones no son reproducibles de una medición a otra y su valor es diferente para cada medida. Las causas de estos errores son incontrolables para el observador. Los errores accidentales son en su mayoría de magnitud muy pequeña y para un gran número de mediciones se obtienen tantas desviaciones positivas como negativas. Aunque con los errores accidentales no se pueden hacer correcciones para obtener valores más concordantes con el real, si se emplean métodos estadísticos se puede llegar a algunas conclusiones relativas al valor más probable en un conjunto de mediciones.

2. Exactitud, precisión y sensibilidad.

La exactitud es el grado de concordancia entre el valor verdadero y el experimental. Un aparato es exacto si las medidas realizadas con él son todas muy próximas al valor "verdadero" de la magnitud medida.

La precisión es el grado de concordancia entre una medida y otras de la misma magnitud realizadas en condiciones sensiblemente iguales. Un aparato es preciso cuando la diferencia entre diferentes medidas de una misma magnitud sean muy pequeñas.

La sensibilidad de un aparato es el valor mínimo de la magnitud que es capaz de medir. Así, si la sensibilidad de una balanza es de 5 mg significa que para masas inferiores a la citada la balanza no presenta ninguna desviación. Normalmente, se admite que la sensibilidad de un aparato viene indicada por el valor de la división más pequeña de la escala de medida.

La exactitud implica normalmente precisión, pero la afirmación inversa no es cierta, ya que pueden existir aparatos muy precisos que posean poca exactitud debido a los errores sistemáticos tales como error de cero, etc. En general, se puede decir que es más fácil conocer la precisión de un aparato que su exactitud.

3. Error absoluto y error relativo

El error absoluto en una medida x de determinada magnitud es la diferencia entre dicho valor y el valor verdadero de la medida; se notará por Δx y, por tanto, su expresión es:

$$\Delta x = x - x_0$$

donde x_0 representa el valor verdadero de la medida. El error absoluto cuantifica la desviación en términos absolutos respecto al valor verdadero. No obstante, en ocasiones es más interesante resaltar la importancia relativa de esa desviación. Por ello, se define el error relativo como el cociente entre el error absoluto y el valor verdadero; notándolo por ε su expresión es:

$$\varepsilon = \frac{|\Delta x|}{x_0}$$

y suele expresarse porcentualmente sin más que multiplicar por 100.

4. Expresión del error

En Física, presentar una medida experimental significa dar el valor de dicha cantidad y expresar cual es su error; no tiene sentido establecer un determinado valor si no se acota debidamente el mismo. Así, la expresión correcta de una medida debe ser:

$$x \pm |\Delta x|$$

Dado el significado de cota de imprecisión que tiene el error absoluto, éste siempre se expresa con una única cifra significativa, es decir, con el primer dígito comenzando por la izquierda distinto de cero; este número ser redondeado por exceso en una unidad si la segunda cifra significativa es 5 o mayor de 5. Este convenio de expresión del error encuentra dos excepciones: que la primera cifra significativa sea un 1 o que siendo la primera un 2, la segunda no llega 5; en estos casos, el error vendrá dado por las dos primeras cifras significativas, procediéndose al redondeo de la segunda en el mismo sentido que ya se ha explicado.

Hay que resaltar que el valor de una magnitud debe tener el mismo orden decimal que el error absoluto. Esto es razonable dado que no tendría sentido encontrar el valor de una magnitud con un grado de precisión superior al del error de la medida. Así, no podemos medir décimas de milímetro con una regla cuya sensibilidad es del milímetro. Finalmente, se acepta como criterio que si el valor de una medida es leído de una tabla u otro lugar, sin indicación de su error, se tomará como error una unidad del orden de la última cifra con que se expresa; por ejemplo, si en una tabla aparece que el valor de una medida es de 0.056 sin ninguna indicación de error, se conviene en que el mismo es de ± 0.001 . En la siguiente tabla se dan distintos ejemplos.

Valores incorrectos	Valores correctos
3.418 ± 0.123	3.42 ± 0.12
6.3 ± 0.09	6.30 ± 0.09
46288 ± 1551	46300 ± 1600
428.351 ± 0.27	428.4 ± 0.3
0.01683 ± 0.0058	0.017 ± 0.006

5. Determinación de errores en medidas directas.

Como ya se ha explicado, cuando se realice la medida de cualquier magnitud hay que indicar el error asociado a la misma. Dado que no conocemos el valor verdadero de la magnitud que deseamos medir, se siguen ciertos procedimientos para hacer una estimación del mismo y de su cota de error.

Con el fin de alcanzar cierta validez estadística en los resultados de las medidas es muy conveniente repetir varias veces su determinación; por convenio, se ha establecido en 3 este número mínimo. No obstante, es posible que en alguna ocasión no tenga sentido llevar a cabo estas repeticiones, en cuyo caso se considera que el error absoluto coincide con el valor de la sensibilidad del aparato utilizado para realizar la medida. En el caso habitual, cuando son 3 las medidas tomadas, pueden presentarse poco o muy dispersas y en función de esta dispersión ser conveniente aumentar o no el número de determinaciones del valor de la magnitud. Para decidir el número de determinaciones del valor de una magnitud física que se desea medir se sigue el siguiente procedimiento: se realizan las 3 mediciones x_i de la magnitud en cuestión y se calcula su valor medio:

$$\bar{x}_3 = \frac{\sum_{i=1}^3 x_i}{3}$$

A continuación se determina su dispersión D, esto es, la diferencia entre los valores extremos de las medidas:

$$D = x_{\text{máximo}} - x_{\text{mínimo}}$$

Finalmente, se obtiene el tanto por ciento de dispersión, T, que viene dado por:

$$T = 100 \cdot \frac{D}{\bar{x}_3}$$

Con estos parámetros se pasa al siguiente cuadro que establece la casuística que puede darse; S representa la sensibilidad del aparato de medida, D₆ es la dispersión para seis medidas y N el número de medidas necesarias en cada caso. Así, por ejemplo, si se ha obtenido que la dispersión es mayor que la sensibilidad y el tanto por ciento de dispersión está comprendido entre el 2% y el 8%, son necesarias 6 medidas; el valor verdadero queda establecido en la media aritmética de las 6 medidas y su error corresponde al máximo de entre la dispersión de las seis medidas dividido por 4 o la sensibilidad.

D	T	N	x ₀	Δx
D < S		3	$\bar{x}_N = \frac{\sum_{i=1}^N x_i}{N}$	S
D > S	T ≤ 2%	3		S
	2% ≤ T ≤ 8%	6		máx {D ₆ /4, S}
	8% ≤ T ≤ 15%	15		$\Delta x = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x}_N)^2}{N \cdot (N - 1)}}$
	15% ≤ T	> 50		

Si se han realizado 15 o más medidas, en realidad se está buscando que el conjunto de las mismas sea una distribución gaussiana o normal, en cuyo caso, el error que se considera corresponde con el error cuadrático medio (ECM) o desviación standard; el significado de este parámetro puede encontrarse en cualquier volumen de estadística básica aunque podemos sintetizarlo de forma cuantitativa como sigue. En el intervalo:

$$\bar{x} - ECM < x < \bar{x} + ECM$$

se encuentra el 68,3% de las medidas realizadas en una gran serie de las mismas. De igual forma, es posible demostrar que en el intervalo:

$$\bar{x} - 2 \cdot ECM < x < \bar{x} + 2 \cdot ECM$$

se encuentra el 95,4% de las medidas realizadas. Por último, en el intervalo:

$$\bar{x} - 3 \cdot ECM < x < \bar{x} + 3 \cdot ECM$$

se encuentra el 99,7% de las medidas realizadas en una gran serie de las mismas.

6. Determinación de errores en medidas indirectas.

Como ya se ha indicado, la medida indirecta de una magnitud se alcanza por aplicación de una fórmula a un conjunto de medidas directas, (variables independientes o datos), que las relacionan con la magnitud problema. Mediante dicha fórmula se obtiene también el error de la medida. Debe tenerse muy presente que si en la expresión matemática que relaciona las magnitudes aparecen números irracionales (tales como π o e) se deben elegir con un número de cifras significativas que no afecten a la magnitud del error absoluto de la magnitud que queremos determinar. En cualquier caso, esta elección determinará el valor del error asignado a dicha constante; en muchas ocasiones, sobre todo cuando se trabaja con calculadora u ordenador, lo más conveniente es tomar todos los decimales que aparecen para el número en cuestión: de esta manera, su error es muy pequeño y puede despreciarse frente a los del resto de las magnitudes que intervengan.

El procedimiento para determinar el error de la medida hecha de manera indirecta es el siguiente. Supongamos que la magnitud F es función de otras magnitudes físicas, estando relacionadas con ellas por la expresión genérica:

$$F = f(x_1, x_2, \dots, x_N)$$

Supongamos, además, que se han realizado medidas de las variables, x_i , y se han determinado su valor y su error. Se obtiene la diferencial total de F en función de las diferenciales de las variables x_i :

$$dF = \frac{\partial F}{\partial x_1} dx_1 + \frac{\partial F}{\partial x_2} dx_2 + \dots + \frac{\partial F}{\partial x_N} dx_N = \sum_{i=1}^N \frac{\partial F}{\partial x_i} dx_i$$

A continuación se asimilan las diferentes diferenciales a los errores absolutos y además consideramos que en el cálculo del error de F debemos ponernos en el caso más desfavorable, es decir, el error mayor posible; así, se tomarán los valores absolutos de las derivadas parciales con el fin de tener una suma de términos positivos. Por tanto, el error en F viene dado por:

$$\Delta F = \sum_{i=1}^N \left| \frac{\partial F}{\partial x_i} \right| \Delta x_i$$

En el caso en el que la función considerada sea de la forma:

$$F = \prod_{i=1}^N x_i^{\alpha_i}$$

con α_i constantes positivas o negativas, se presenta una notable simplificación si se procede a tomar logaritmos neperianos antes de llevar a cabo el análisis anterior. En efecto, si se lleva a cabo esta operación se tiene que:

$$\ln F = \ln \left(\prod_{i=1}^N x_i^{\alpha_i} \right) = \sum_{i=1}^N \alpha_i \ln x_i$$

de donde, obteniendo su diferencial:

$$d(\ln F) = \sum_{i=1}^N \alpha_i d(\ln x_i)$$

Ahora, teniendo en cuenta la diferencial del logaritmo neperiano, se concluye que:

$$\frac{dF}{F} = \sum_{i=1}^N \alpha_i \frac{dx_i}{x_i}$$

Finalmente, asimilando de nuevo los diferenciales totales a los errores absolutos se obtiene:

$$\frac{\Delta F}{F} = \sum_{i=1}^N \left| \alpha_i \frac{\Delta x_i}{x_i} \right|$$

que en función del error relativo quedaría:

$$\Delta F = F \sum_{i=1}^N |\alpha_i \cdot \varepsilon_i|$$

El siguiente ejemplo sirve para ver de forma práctica las actuaciones descritas hasta aquí. Supongamos que se quiere determinar el volumen de un cilindro; para ello, puesto que este parámetro viene dado por:

$$V = \pi r^2 h$$

se procederá a calcular el radio r y la altura h del cuerpo. Supongamos que tales valores son $r = 5.00 \pm 0.05$ y $h = 100.0 \pm 0.5$. Entonces, el volumen vale $V = 7853.9816\dots$. Para expresar correctamente este resultado hay que determinar cuanto vale su error; así, se calcula el valor de la diferencial de V :

$$dV = 2\pi r h dr + \pi r^2 dh$$

y se sustituyen los diferenciales por errores:

$$\Delta V = |2\pi r h| \Delta r + |\pi r^2| \Delta h$$

de donde obtenemos que $\Delta V = 196.34954\dots$. Por tanto, el resultado de la medición del volumen es $V = 7900 \pm 200$.

III. TRATAMIENTO DE LOS DATOS

En muchas ocasiones, los resultados obtenidos de los datos se interpretan mejor con ayuda de una representación gráfica. Además, este procedimiento muestra una tendencia que permite estimar los valores en otros puntos diferentes a los experimentales o demuestra una determinada relación matemática entre las variables representadas. Por ello, conviene exponer el método de ajuste de datos más frecuente, el de mínimos cuadrados, y resumir las características que debe tener una buena gráfica. Para acabar, es muy posible que a menudo encontremos datos en unas tablas y no se encuentre en ellas el valor exacto en el que estamos

interesados; para tal fin, se exponen las reglas de interpolación básica en tablas de simple o doble entrada.

1. El Método de Mínimos Cuadrados.

Con frecuencia, se plantea el problema de encontrar una expresión matemática $y = f(x)$ de la ley física que rige el comportamiento de un determinado fenómeno a partir de una serie de N medidas (x_i, y_i) de las magnitudes x e y que lo caracterizan. En este apartado se estudiará el caso de que la representación gráfica del fenómeno estudiado proporcione una distribución de los puntos experimentales en forma prácticamente lineal; esto debe interpretarse como la dependencia lineal de las dos variables físicas y , por ello, es necesario determinar la ecuación de la recta que será la expresión de la ley física que rige el fenómeno estudiado. El método más frecuente para llevar a cabo este ajuste se denomina de mínimos cuadrados.

Se pretende, por tanto, encontrar una recta $y = ax + b$ de forma que se ajuste lo mejor posible a los datos experimentales. Ahora bien, esta bondad de ajuste puede establecerse de varias maneras. El método de mínimos cuadrados toma como mejor ajuste aquel que hace mínima la siguiente cantidad:

$$\sum_{i=1}^N (y_i - y_{\text{teórico}})^2 = \sum_{i=1}^N (y_i - ax_i - b)^2$$

Observe que los parámetros que determinan la recta son su pendiente a y su ordenada en el origen b . Por tanto, estamos frente a un problema de extremos que depende de las variables a y b . Es condición necesaria entonces que:

$$\frac{\partial \sum_{i=1}^N (y_i - ax_i - b)^2}{\partial a} = 0$$
$$\frac{\partial \sum_{i=1}^N (y_i - ax_i - b)^2}{\partial b} = 0$$

Se tienen pues dos ecuaciones con dos incógnitas, a y b ; resolviendo se obtiene que:

$$\sum_{i=1}^N 2(y_i - ax_i - b)(-x_i) = 0$$

$$\sum_{i=1}^N 2(y_i - ax_i - b)(-1) = 0$$

de forma que despejando a y b adecuadamente se llega a:

$$a = \frac{\sum_{i=1}^N x_i \sum_{i=1}^N y_i - N \sum_{i=1}^N x_i y_i}{\left(\sum_{i=1}^N x_i\right)^2 - N \sum_{i=1}^N x_i^2}$$

$$b = \frac{\sum_{i=1}^N x_i \sum_{i=1}^N x_i y_i - \sum_{i=1}^N y_i \sum_{i=1}^N x_i^2}{\left(\sum_{i=1}^N x_i\right)^2 - N \sum_{i=1}^N x_i^2}$$

De estas expresiones es posible encontrar el valor del error en cada parámetro:

$$\Delta a = \sqrt{\frac{\sum_{i=1}^N (y_i - ax_i - b)^2}{(N-2) \sum_{i=1}^N (x_i - \bar{x})^2}}$$

$$\Delta b = \sqrt{\left(\frac{\sum_{i=1}^N (y_i - ax_i - b)^2}{(N-2)}\right) \cdot \left(\frac{1}{N} + \frac{\bar{x}^2}{\sum_{i=1}^N (x_i - \bar{x})^2}\right)}$$

Además de los valores de pendiente y ordenada en el origen sería interesante obtener algún factor que cuantificara la bondad del ajuste; esto permitiría comparar los resultados de diferentes ajustes; este factor se denomina coeficiente de correlación lineal "r". La expresión de r es:

$$r = \frac{N \sum_{i=1}^N x_i y_i - \sum_{i=1}^N x_i \sum_{i=1}^N y_i}{\sqrt{\left(N \sum_{i=1}^N x_i^2 - \left(\sum_{i=1}^N x_i \right)^2 \right) \left(N \sum_{i=1}^N y_i^2 - \left(\sum_{i=1}^N y_i \right)^2 \right)}}$$

Puede probarse que el valor de r está acotado en valor absoluto entre 0 y 1, siendo tanto mejor el ajuste cuanto más cercano a la unidad sea r.

Como ejemplo de aplicación del método de mínimos cuadrados supongamos un experimento en el que se mide el alargamiento de un muelle debido a la acción de una pesa; se trata de comprobar la ley de Hooke. Los datos que se han obtenido son:

Medida número "i"	"x" l_i	"y" m_i
1	42.0	2
2	48.4	4
3	51.3	6
4	56.3	8
5	58.6	10

Esta tabla puede ser tratada para encontrar las sumatorias oportunas como sigue:

Medida número "i"	"x" l_i	"y" M_i	l_i^2	$m_i l_i$
1	42.0	2	1764	84
2	48.4	4	2342.56	194
3	51.3	6	2631.69	308
4	56.3	8	3169.69	450
5	58.6	10	3433.96	586
N = 5	$\sum l_i = 256.6$	$\sum m_i = 30$	$\sum m_i^2 = 13341.9$	$\sum m_i l_i = 1622$

de donde ahora, fácilmente, obtenemos que:

$$a = \frac{\sum_{i=1}^N m_i \sum_{i=1}^N l_i - N \sum_{i=1}^N m_i l_i}{\left(\sum_{i=1}^N m_i \right)^2 - N \sum_{i=1}^N m_i^2} = 0.47463$$

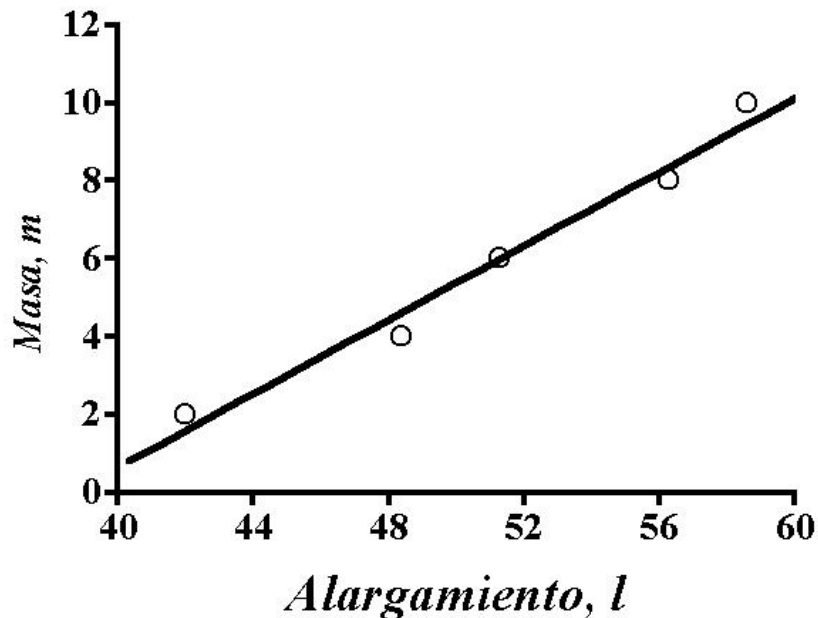
$$b = \frac{\sum_{i=1}^N m_i \sum_{i=1}^N m_i l_i - \sum_{i=1}^N l_i \sum_{i=1}^N m_i^2}{\left(\sum_{i=1}^N m_i \right)^2 - N \sum_{i=1}^N m_i^2} = -18.3579$$

Además se tiene que $\Delta a = \pm 0.04$ y $\Delta b = \pm 2.2$. Así pues, concluimos:

$$a = 0.47 \pm 0.04$$

$$b = -18.4 \pm 2.1$$

Finalmente, con los valores de a y b puede realizarse una representación gráfica; en efecto, basta tomar dos puntos cualesquiera que pertenezcan a la recta para poder dibujarla. El aspecto de la gráfica sería, por tanto, como el de la figura.



2. Construcción de gráficas

La representación gráfica de los fenómenos físicos que estudiemos deben ajustarse a las siguientes normas de uso general que clarifican y estandarizan los resultados. Se pueden enumerar como sigue:

a) Las gráficas se harán en papel milimetrado con los ejes bien trazados y en cuyos extremos se indique la magnitud representada en ellos y la unidad en que ha sido medida. El título de la gráfica será claro y vendrá indicado en la parte superior.

b) La variable independiente del fenómeno debe ir representada en abscisas y la dependiente en ordenadas.

c) Las escalas, sobre ambos ejes, han de permitir una lectura rápida y sencilla. Para ello se elegirán las escalas con intervalos de medida adecuados.

d) Las escalas deben abarcar todo el intervalo de medidas realizadas y sólo el citado intervalo.

e) Sobre los ejes solo se indican los valores correspondientes a las divisiones enteras de la escala de forma que queden uniformemente espaciadas. En general, no se señalan los valores correspondientes a las medidas realizadas.

f) Los valores medidos se representan sobre el papel milimetrado por el punto correspondiente a sus dos coordenadas y rodeado por el denominado rectángulo de error. Este tiene por base la longitud comprendida entre $x_i - \Delta x$ y $x_i + \Delta x$ y por altura se extiende desde $y_i - \Delta y$ hasta $y_i + \Delta y$, siendo x_i e y_i las coordenadas del punto experimental. En el caso de que x o y sean despreciables en comparación con la escala utilizada el rectángulo de error queda reducido a un simple segmento vertical u horizontal, según el caso.

g) Las líneas que aparezcan en las gráficas representan la tendencia de los puntos experimentales y se obtienen por medio del método de ajuste correspondiente; por ello, han de ser líneas finas continuas pero nunca quebradas y determinadas por los valores experimentales.

3. Interpolación en tablas

La interpolación en tablas permite encontrar valores de las variables dependientes para valores concretos de la variable independiente que no están explícitamente en la tabla.

X	Y
x_1	y_1
x	y
x_2	y_2

Así, en una tabla de simple entrada para las magnitudes X e Y, se localizarán los valores entre los que se quiere interpolar, x_1 y x_2 , a los que corresponden y_1 e y_2 . Para un valor dado de x comprendido entre x_1 y x_2 , el valor de y correspondiente viene dado por:

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

y el error correspondiente es:

$$\Delta y = \left| \frac{y_2 - y_1}{x_2 - x_1} \right| \cdot \Delta x$$

En las tablas de doble entrada para cada pareja de valores x, y se proporciona el valor correspondiente a una tercera variable z relacionada con las dos anteriores. Si la tabla responde al tipo:

	y_1	y	y_2
x_1	z_{11}		z_{12}
x		z	
x_2	z_{21}		z_{22}

donde z es el valor buscado y comprendido entre x_1 y x_2 e y_1 e y_2 , entonces:

$$z = \frac{z_{21} - z_{11}}{x_2 - x_1} (x - x_1) + \frac{z_{12} - z_{11}}{y_2 - y_1} (y - y_1) + z_{11}$$

y el correspondiente error viene dado por:

$$\Delta z = \left| \frac{z_{21} - z_{11}}{x_2 - x_1} \right| \Delta x + \left| \frac{z_{12} - z_{11}}{y_2 - y_1} \right| \Delta y$$