ERROR IMAGINARIO

Carlos Antonio Rivera Orozco, María del Rosario Rodríguez Báez
Bufete de ingenieros industriales, S.C.
Francisco Pimentel 4104 B; colonia Las granjas; C.P. 31100; Chihuahua, Chih. México
614- 419 3700; carlos@bii.mx; rosario@bii.mx

Resumen: El error aleatorio es contra intuitivo porque su representación lineal muestra de forma simultánea los signos positivo y negativo para la misma referencia. La literatura típica en los campos de la estadística y de la metrología ofrece pocos apoyos gráficos para comprender este fenómeno. Este trabajo expone el proceso para calcular el error aleatorio por medio de la desviación estándar, analiza las características de los errores en el plano cartesiano y concluye que el error aleatorio es imaginario. Como resultado se proponen las definiciones de error real y error imaginario.

1. INTRODUCCIÓN

El error aleatorio tiene su expresión más conocida en la desviación estándar (o típica)[1]; las propiedades matemáticas del concepto se describen con mayor o menor detalle en literatura de los campos de la estadística y la metrología.

Es común encontrar definiciones en prosa con la expresión algebraica que le corresponde, en algunas ocasiones se acompañan con un gráfico lineal.

En raras ocasiones se explica la razón para transformar los cálculos lineales de la desviación estándar como una suma de cuadrados y en menos ocasiones se ofrecen gráficos que soporten su comprensión.

Por lo anterior, el apartado dos explica los conceptos básicos por cuatro medios: prosa, álgebra, números y gráficos[2]; el énfasis está orientado para vincular conceptos de metrología con estadística.

Durante la explicación se analiza el error aleatorio en el plano cartesiano y presenta los argumentos que permiten concluir que los cuadrantes II y IV son simultáneos, que los valores que se registran en esas áreas son imaginarios y que la desviación estándar pertenece a estos cuadrantes.

El apartado tres propone dos definiciones de error como resultado de este trabajo.

[1] Si bien los conceptos desviación estándar y desviación típica se refieren a lo mismo, en este trabajo se utiliza el primero.
[2] Esta idea se toma del prefacio del libro "Cálculo. Trascendentes tempranas", por James Stewart [1].

Este trabajo tiene su origen en dos preguntas que se plantearon en «Incertidumbre a cuadros»: "¿qué ocurre gráficamente con las áreas que obtienen signo negativo? ¿Cómo interactúan los cuadrantes 2º, 3º y 4º del plano cartesiano?" [2] Igual que en el trabajo anterior, la intensión es ofrecer un apoyo didáctico para el análisis del error aleatorio.

2. CONCEPTOS BÁSICOS

El error aleatorio es un concepto que se construye con los conceptos y cálculos que se describen en los subapartados que siguen.

2.1. Error de medida

En el campo de las matemáticas el error es la "diferencia entre el valor medido o calculado y el real" [3] en la metrología es la "diferencia entre el valor medido de una magnitud y un valor de referencia" [4]; ambas definiciones son coherentes.

Considerando lo anterior, el error (e) siempre hace referencia a una operación aritmética, la resta, donde un valor de referencia (vr) se sustrae de un valor observado (vo), su expresión algebraica es

$$e = vo - vr \tag{1}$$

donde el residuo puede adquirir uno de tres valores:

$$\text{Si } vo < vr \rightarrow e < 0 \tag{1.1}$$
$$\text{Si } vo = vr \rightarrow e = 0 \tag{1.2}$$
$$\text{Si } vo > vr \rightarrow e > 0 \tag{1.3}$$

Para el caso (1.1) el residuo tiene valor negativo, $-e$, esto significa que el valor observado está por debajo del valor de referencia.

Para el caso (1.2) el residuo tiene valor cero, $e = 0$, esto significa que el valor observado es idéntico al valor de referencia, no hay error.

Para el caso (1.3) el residuo tiene valor positivo, $+e$, esto significa que el valor observado está por arriba del valor de referencia.

De los casos anteriores se concluye que los signos del error de medida son mutuamente excluyentes: es positivo o es negativo.

El error de medida se interpreta con dos elementos que lo forman: primero el signo que informa sobre la ubicación del valor observado con respecto al valor de referencia, segundo el valor numérico que cuantifica la diferencia entre los valores.

Como ejemplo, considere un patrón de trabajo con valor de $10\ mm$, el instrumento de medida presenta $9.8\ mm$, el error de medida e se calcula como sigue

$$\begin{gathered} e = 9.8\ mm - 10\ mm \\ e = -\ 0.2\ mm \end{gathered} \tag{2}$$

La ***figura 1*** muestra su representación gráfica

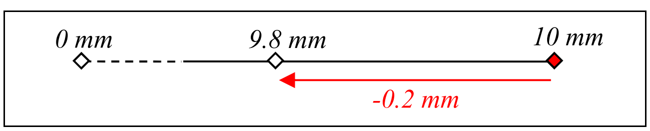

Fig. 1 *Sentido y cuantificación del error de medida.*

2.2. Media aritmética

En el campo de la estadística, la media aritmética, en lo sucesivo la media, es el valor esperado de un conjunto de datos, su cálculo representa una medida de tendencia central. Este concepto no se define para el campo de la metrología. La ecuación (3) presenta su forma de cálculo.

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

Es importante observar que, a diferencia del concepto del error, el cálculo de la media no requiere de valores externos al proceso, la media se puede calcular sin necesidad de una referencia.

Como ejemplo, considere que se realizan, con el mismo instrumento, en la misma pieza, cuatro mediciones de longitud en la misma posición cuyos resultados son: $5.33\ mm$; $5.34\ mm$; $5.36\ mm$ y $5.37\ mm$; al sustituir los valores en la ecuación (3) se produce el resultado que se muestra a continuación

$$\begin{gathered} \bar{x} = ((5.33 + 5.34 + 5.36 + 5.37)\ mm\ )/\ 4 \\ \bar{x} = 5.35\ mm \end{gathered} \tag{4}$$

con la representación gráfica de la ***figura 2***

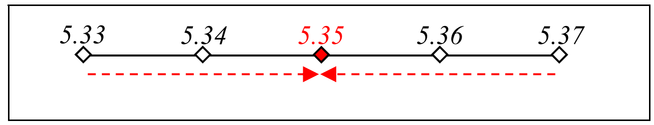

Fig. 2 *La media en el conjunto de 4 datos.*

Como se observa en el ejemplo, la media es el valor en el centro del conjunto, se puede decir que $5.35\ mm$ es un valor adecuado para representar ese conjunto de datos.

2.3. Error aleatorio de medida

En el campo de la metrología el error aleatorio es el “componente del error de medida que, en condiciones repetidas, varía de manera impredecible”, la nota 1 del VIM para este concepto señala “el valor de referencia para un error aleatorio es la media que se obtendría de un número infinito de mediciones repetidas del mismo mensurando” [5]. El concepto estadístico más cercano a lo anterior es la desviación estándar.

2.4. Desviación estándar

En el campo de la estadística la desviación estándar se define como la raíz cuadrada de la varianza que a su vez representa el promedio del cuadrado de la distancia de cada punto con respecto de la media.

En el campo de la metrología el concepto se define con mayor precisión según su uso, por ejemplo, la GUM[3] la define como la “desviación estándar experimental” [6], la norma ISO 5725-1 presenta las definiciones “desviación estándar de repetibilidad” y desviación estándar de reproducibilidad” [7], según las condiciones experimentales en las que se obtienen los datos.

La desviación estándar se calcula de dos formas, según las características del conjunto de datos: al calcular el total de los datos posibles se obtiene una desviación estándar poblacional; al calcular un

[3] Guide to the expression of uncertainty of measurements.

subconjunto del total de los datos posibles se obtiene una desviación estándar muestral.

La ecuación (5) presenta el cálculo de la desviación estándar poblacional, la ecuación (6) presenta el cálculo de la desviación estándar muestral.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(X_i-\mu)^2}{N}} \qquad (5)$$

$$s = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n-1}} \qquad (6)$$

Las diferencias de forma entre las ecuaciones se observa en los símbolos utilizados. La convención en el campo de la estadística es representar la media poblacional con la letra mi, μ del alfabeto griego [8]; la desviación estándar poblacional con la letra sigma, σ del alfabeto griego y la desviación estándar muestral con la letra s. Las letras X y N de la ecuación (5) representan la totalidad de los datos posibles a diferencia de la ecuación (6) que corresponde a datos muestra.

La diferencia de cálculo entre las ecuaciones se observa en el denominador del segundo miembro. Cuando se calcula la desviación estándar muestral la división es entre n-1, un número más pequeño que el que el total de datos n, con el fin de corregir el sesgo que se produce al utilizar sólo una porción de los datos posibles. Esto se conoce como la corrección de Bessel [9]. La relación entre las desviaciones estándar se muestra a continuación

$$s > \sigma \qquad (7)$$

La cantidad de datos posibles de una medición es infinito, luego entonces, sin importar el número de datos que se generen, siempre se trabaja con una muestra de los valores de medida. Lo anterior explica de forma razonable porque la convención en metrología es utilizar la desviación estándar muestral s.

Al analizar la ecuación (6) se observa la interacción del error y la media. El primer concepto se representa con la ecuación (1) de tal forma que al sustituir vo por x_i y vr por $\bar{x}$ se obtiene la ecuación (8)

$$e_i = x_i - \bar{x} \qquad (8)$$

al sustituir (8) en (6) se obtiene la ecuación (9)

$$s = \sqrt{\frac{\sum_{i=1}^{n} e_i^2}{n-1}} \qquad (9)$$

El concepto de la media se observa en la ecuación (9), dentro de la raíz cuadrada, donde se aprecia la división de la suma de n errores que se divide entre n-1 errores. Esta expresión se conoce como el promedio del cuadrado de los errores.

Como ejemplo, se calcula la desviación estándar para los valores de la ***figura 2***. La ***tabla 1*** presenta los resultados de los errores por punto de medida.

i	x	$\bar{x}$	e	e^2
1	5.33	5.35	-0.02	0.000 4
2	5.34	5.35	-0.01	0.000 1
3	5.36	5.35	0.01	0.000 1
4	5.37	5.35	0.02	0.000 4

Tabla 1. *Errores por punto de medida.*

Según se observa, para calcular la desviación estándar en un grupo de cuatro datos se obtienen cuatro errores de medida, en el entendido que el valor de referencia es la media.

Considerando que el objetivo es generar un valor que represente al conjunto de errores, lo razonable es calcular la media de los errores, pero esto genera un problema porque la adición de los errores en la columna e es cero. Este fenómeno es natural porque al estar la referencia en el centro de los datos, al sumar los errores, estos se compensan. La ***figura 3*** muestra su representación gráfica

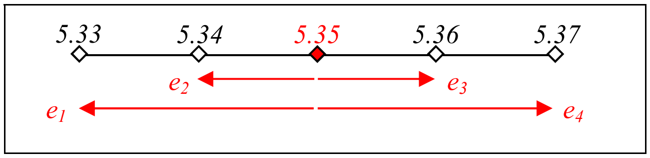

Fig. 3 *Errores en el cálculo de la desviación estándar en un conjunto de cuatro datos.*

De lo anterior se concluye que la suma de los errores con respecto a la media, expresados como segmentos de recta, es cero. Algebraicamente se tiene la identidad (10)

$$\sum_{i=1}^{n} e_i \equiv 0 \qquad (10)$$

Para evitar la compensación lineal de los errores, estos se transforman a una expresión cuadrática, la ecuación (10) cambia a la identidad (11) como la suma es en términos de áreas, lo razonable es trasladar la representación de los errores a un gráfico de dos dimensiones.

$$\sum_{i=1}^{n} e_i^2 \geq 0 \tag{11}$$

Las ***figuras 4a y 4b*** presentan los errores de la ***tabla 1*** en el plano cartesiano.

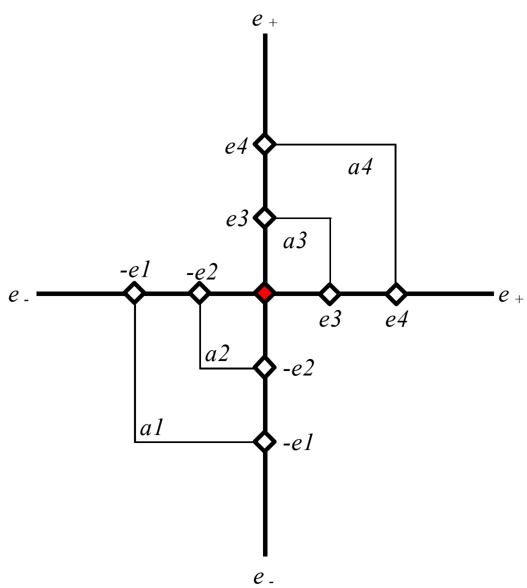

Fig. 4a *Plano cartesiano con el origen en la media.*

La ***figura 4a*** muestra el origen trasladado a la media de $5.35\ mm$, que es el valor de referencia representado por el rombo rojo. Los rombos blancos representan los errores calculados en la ***tabla 1***, proyectados de forma simultánea en las abscisas y las ordenadas. Esta es la representación gráfica de los errores elevados al cuadrado.

El cuadrante I presenta los errores positivos, el cuadrante III presenta los errores negativos, el área que corresponde al error cuadrado se representa con la letra a. Las áreas en los cuadrantes I o III aportan información sobre su ubicación con respecto a la media, por ejemplo, el área $a2$ ubicada en el cuadrante III corresponde al error negativo observado en la medición de $5.34\ mm$. La ***figura 4b*** presenta los errores de cada cuadrante.

Los cuadrantes II y IV corresponden a errores imaginarios porque cualquier valor registrado en ese espacio implica que el error tiene dos signos diferentes de forma simultánea y como ya se indicó en el subapartado 2.1, esto no es coherente en la expresión lineal. Por ejemplo, el área ax en el cuadrante II significa que el valor observado x es mayor y menor que la media al mismo tiempo, el registro tiene una representación simultánea en el cuadrante IV que permite registrar un valor menor y mayor con respecto a la media.

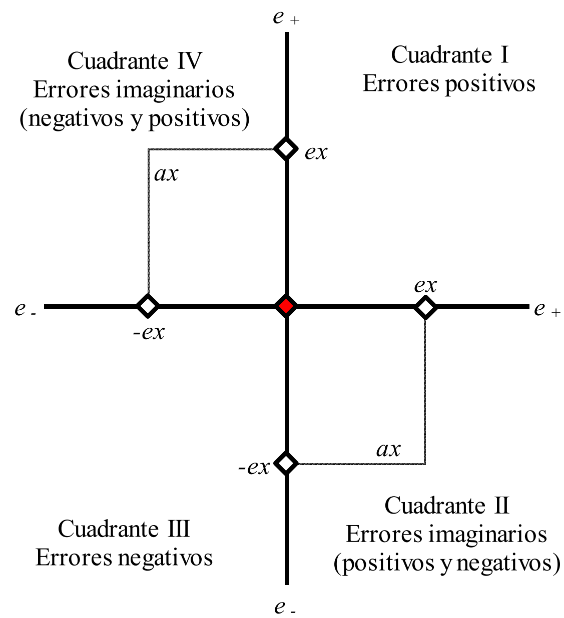

Fig. 4b *Errores en el plano cartesiano.*

El gráfico de los errores en el plano cartesiano con el origen en la referencia permite concluir (1) que los errores positivos se expresan como áreas en el cuadrante I, (2) que los errores negativos se expresan como áreas en el cuadrante III, (3) que los cuadrantes II y IV presentan errores imaginarios ya que no es posible observar en una medición un error cuyo sentido sea positivo y negativo al mismo tiempo.

Regresando al cálculo de la desviación estándar, se tiene entonces que los errores expresados como áreas, presentan la propiedad señalada en la ecuación (11), la suma de áreas corresponde al numerador en la ecuación (9) de tal forma que el cociente dentro de la raíz cuadrada es un área promedio que representa a los errores con respecto a la media.

El área promedio en la ***figura 4a*** es $0.000\ 25\ mm^2$, esto es cuando $n = 4$, el valor corresponde a σ^2. El área promedio no sesgada es $0.000\ 33\ mm^2$, esto es considerando $n\text{-}1 = 3$, el valor corresponde a s^2.

El párrafo anterior presenta los resultados de las varianzas poblacional y muestral. Para expresar el resultado de la desviación estándar es necesario transformar los resultados a segmentos de líneas, calculando la raíz cuadrada de la varianza.

La transformación de un valor de dos dimensiones a un valor de una dimensión produce dos resultados, el primero es un valor x con signo positivo, el segundo es el mismo valor x con signo negativo; es decir que al transformar el valor cuadrático a lineal se obtiene un resultado asociado a los signos positivo y negativo de forma simultánea.

Lo anterior es posible por dos razones, la primera, la desviación estándar produce su propia referencia que está en el centro del conjunto de datos, por lo tanto el promedio lineal de los errores tiene ambos sentidos; la segunda, la desviación estándar es un estimador de la dispersión del conjunto de datos, por lo tanto no está obligado a pertenecer al conjunto de los errores reales.

Una vez considerado lo anterior, la desviación estándar muestral de los datos en la ***figura 3*** es $0.0182\ mm$

La ***figura 5a*** presenta las áreas imaginarias de la desviación estándar en los cuadrantes II y IV del plano cartesiano.

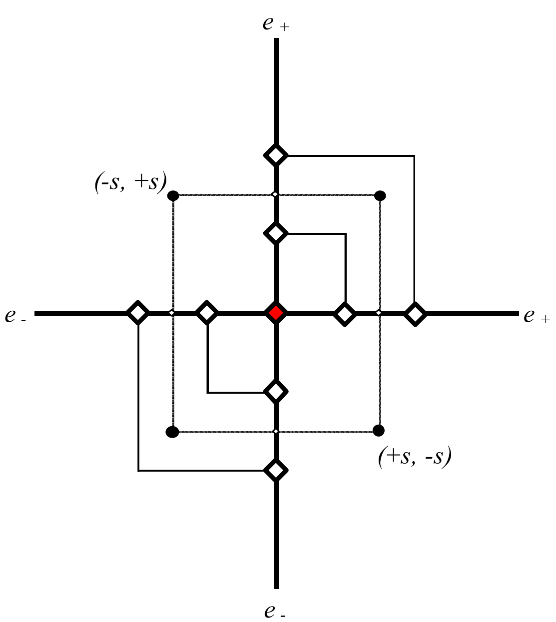

Fig. 5a *Proyección de la desviación estándar.*

Las coordenadas $(+s,\ -s)$ y $(-s,\ +s)$ presentan el orden de los signos para el valor s. Las áreas imaginarias que se proyectan en los cuadrantes I y III muestran el área promedio en los campos de los errores reales.

Considerando lo expuesto, se puede concluir (1) que el área promedio de los errores es un área imaginaria, (2) que al considerar los cuadrantes II y IV como simultáneos, el área imaginaria total es dos veces el producto del área imaginaria; (3) que las áreas imaginarias conservan propiedades aditivas, (4) que las áreas imaginarias pueden interactuar con las áreas reales al proyectarse sus valores en los cuadrantes I y III.

La ***figura 5b*** presenta los segmentos de recta que le corresponden a la desviación estándar en un gráfico lineal donde se puede observar que a partir de la referencia se presentan el sentido y la cuantificación de dos errores de medida de forma simultánea.

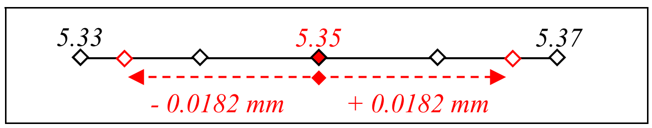

Fig. 5b *Desviación estándar en el gráfico lineal.*

La representación de la ***figura 5b*** muestra los valores ordenados en la forma $(-s,\ +s)$, en sentido estricto, esta es la representación lineal que le corresponde al cuadrante IV de la ***figura 5a.***

3. RESULTADOS

El error es un tema de interés especial en la metrología, esto se observa en los recursos dedicados para definirlo, clasificarlo y ejemplificarlo en el VIM y la GUM [4]. Lo anterior se justifica por los niveles de abstracción requeridos para tratar los conceptos relacionados con el tema (como valor verdadero, valor realista e incertidumbre) y los conceptos derivados (como errores sistemático y aleatorio).

Considerando lo anterior, es razonable pensar que el aparato de conceptos está maduro y es funcional,

[4] Como ejemplos se pueden revisar el sexto párrafo del apartado 0.1 y las seis definiciones que empiezan con el vocablo "error" en el VIM (ver referencia [4]), por otra parte el subapartado 3.2, el primer párrafo del anexo D, el subanexo D.4 y el anexo E en la GUM (ver referencia [6]).

sin embargo aún existe espacio para exponer las propiedades de los errores que se describen en el apartado dos de este trabajo. Por lo anterior se proponen dos definiciones nuevas.

3.1. error real

Error cuya expresión cuadrática en un plano cartesiano, con el origen trasladado al valor de referencia, se observa en el cuadrante I o en el cuadrante III.

Nota 1. En un gráfico lineal, el error real tiene sólo un valor y sentido con respecto a su referencia.

Nota 2. Son errores reales: error de medida, error sistemático, sesgo y veracidad.

3.2. error imaginario

Error cuya expresión cuadrática en un plano cartesiano, con el origen trasladado al valor de referencia, se observa de forma simultánea en los cuadrantes II y IV.

Nota 1. En un gráfico lineal, el error imaginario tiene dos valores y dos sentidos con respecto a su referencia.

Nota 2. Son errores imaginarios: error aleatorio, repetibilidad, reproducibilidad, precisión, precisión intermedia e incertidumbre de medida.

4. DISCUSIÓN

No se deben confundir los conceptos error imaginario y número imaginario. El primero se define en función del espacio que ocupa en un plano cartesiano; el segundo cuenta con una definición formal en el campo de las matemáticas.

En la práctica se observa la separación en el tratamiento de datos por tipo de errores, por ejemplo, la ley de la propagación de las incertidumbres está basada en las propiedades aditivas (de las áreas) de los errores imaginarios; por otra parte, los ajustes de los instrumentos de medida se basan en el estudio de los errores reales.

La GUM “da por sentado que todas las componentes de incertidumbre son de la misma naturaleza y son tratadas idénticamente” [10], bajo este punto de vista se justifica que los efectos sistemáticos y aleatorios sean tratados de la misma forma (ver anexo E.3 de la referencia [6]).

Si, para el caso anterior, se sustituyen los conceptos de efectos sistemático y aleatorio por errores real e imaginario, el punto de vista de la GUM ganaría en claridad al evidenciar la forma en que interactúan los errores en el plano cartesiano, según se muestra en la ***figura 5a***.

Por último, el análisis de los errores reales e imaginarios por el método gráfico que se propone aporta elementos didácticos para comprender el concepto del error aleatorio.

5. CONCLUSIONES

Este trabajo ofrece una explicación coherente sobre los errores real e imaginario; también ofrece herramientas gráficas con potencial didáctico para la diseminación de los conceptos relacionados con el error aleatorio; por último se ofrecen dos definiciones formales que se pueden incluir como parte del vocabulario de metrología.

REFERENCIAS

[1] J. Stewart “Cálculo. Trascendentes tempranas” Cuarta Edición: Colombia, Thomson Learning, p. ix, 2002.
[2] C. Rivera y R. Rodríguez, “Incertidumbre a cuadros”, en Simposio Metrología 2012, Centro Nacional de Metrología, p. 553, 2012.
[3] Real Academia Española, “error”, http://lema.rae.es/drae/?val=error, 2014-04-10.
[4] CENAM Traducción al español del VIM-3ª. Marzo 2009, “Vocabulario Internacional de Metrología – Conceptos fundamentales y generales, y términos asociados (VIM)”, Centro Nacional de Metrología, p. 30, 2009.
[5] ibídem., p. 31.
[6] Norma Mexicana NMX-CH-140-IMNC-2002, “Guía para la expresión de incertidumbre en las mediciones”, Instituto Mexicano de Normalización y Certificación, p. 13.
[7] Norma Mexicana NMX-CH-5725-1-IMNC-2006, “Exactitud (veracidad y precisión) de resultados y métodos de medición – Parte 1: Principios Generales y definiciones”, Instituto Mexicano de Normalización y Certificación, pp. 4-5.
[8] Real Academia Española, “mi”, http://lema.rae.es/drae/, 2014-04-13.
[9] J. b. Kennedy y A.M. Neville, “Estadística para ciencias e ingeniería” Segunda edición: México, Harla, pp. 35-36, 1982.
[10] NMX-CH-140-IMNC-2002, Op.cit., pp. 67.