Error estadístico (σ_e) o Desviación estándar de los promedios

Hasta ahora vimos dos estimadores para una distribución: el promedio ($\bar{x}$) y la desviación estándar (S o σ). Ambas son propiedades de la distribución específica de la que estamos hablando.

Si la distribución es gaussiana, podemos usar esos dos estimadores para dar una primera imagen del valor de la magnitud física asociada a la distribución y a su incerteza. En ese caso, sabemos que alrededor del 68% de los datos se encuentran en el intervalo $\bar{x} \pm S$, por lo que podemos decir: "si realizo una nueva medición de la magnitud física *x*, dentro de los parámetros y condiciones que definen a la distribución, tengo una probabilidad de ~ 68% de obtener un valor dentro del intervalo $\bar{x} - S$ y $\bar{x} + S$ (que no se aleja más que S respecto del promedio)." Ese es el significado entonces de la expresión $\bar{x} \pm S$.

Si tomamos más de una muestra de datos cuál será la dispersión de los promedios de dichas muestras?. Para esto vamos a trabajar con la varianza de la magnitud física (o variable aleatoria VA) x, *V*(x), que es simplemente S^2, y con sus propiedades.

Supongamos que tenemos una muestra "A" con N mediciones de una magnitud física (x), el promedio será

$$\bar{x} = \sum_{i=1}^{N} \frac{x_i}{N} \qquad (1)$$

Y la varianza de x:

$$V(x) = \sum_{i=1}^{N} \frac{(x_i - \bar{x})^2}{N} \qquad (2)$$

Algunas propiedades de la varianza:

1) Si la magnitud física es constante, su varianza es cero (la variancia de una cte. es cero). Lo cual resulta trivial según su definición (Eq. 2).

2) Si tenemos varias VA medidas independientemente, la varianza de la suma de las VA es la suma de las varianzas.

3) Si conocemos la varianza de x, la varianza del producto de x por una constante k cumple que:

$$V(k.x) = k^2\, V(x) \qquad (3)$$

(demostrar)

Supongamos entonces que tenemos nuestra variable aleatoria *x*, y queremos saber qué tan bien está definido el promedio de la muestra tomada con N mediciones de una magnitud física (x). Para ello, tomemos una cantidad n de muestras tomadas del mismo experimento, bajos las mismas condiciones (n experimentos), cada uno de los cuales tendrá un valor medio ($\bar{x}$).

El promedio de la muestra con n datos de los promedios será

$$< \bar{x} >= \sum_{i=1}^{n} \frac{\bar{x}_i}{n} \qquad (4)$$

Así como antes nuestro objetivo fue definir un intervalo de confianza para cualquier dato medido de la experiencia con N mediciones de la magnitud física (x), ahora queremos hacer lo mismo con la experiencia que toma los promedios de la anterior.

Para esto, podemos buscar cuánto vale la varianza del promedio que está definido en Eq. (1); y utilizando las propiedades de la varianza, podemos llegar a que:

$$V(\bar{x}) = V\left(\sum_{i=1}^{N} \frac{x_i}{N}\right) = V\left(\frac{1}{N}\sum_{i=1}^{N} x_i\right) \underset{\text{Propiedad 3)}}{=} \frac{1}{N^2} V\left(\sum_{i=1}^{N} x_i\right) \underset{\text{Propiedad 2)}}{=} \frac{1}{N^2}\left(\sum_{i=1}^{N} V(x_i)\right) = \frac{1}{N^2}\left(NV(x_i)\right) \underset{\substack{\text{Dado que } x_i \text{ es una medición de la VA } x, \\ \text{la varianza de } x_i \text{, es la varianza de } x}}{=} \frac{V(x)}{N}$$

Entonces,

$$V(\bar{x}) = \frac{V(x)}{N} \qquad (5)$$

Lo cual quiere decir que la varianza del promedio de una series de mediciones, descripta por la distribución de la VA x, es inversamente proporcional al número de mediciones que forman cada serie (N).

Si volvemos a hablar en términos de la desviación estándar, aplicando raíz cuadrada a $V(\bar{x})$ de Eq. 5, se tiene:

$$S_{<\bar{x}>} = \frac{S}{\sqrt{N}} = \sigma_e \qquad (6)$$

La ecuación (6) nos indica que si analizamos el promedio a partir de realizar una serie de muestras, existe un 68% de probabilidades que el promedio de una muestra de N datos, se encuentre en el intervalo $< \bar{x} > \pm \sigma_e$. Donde $< \bar{x} >$ es el promedio de los promedios de varias muestras con N datos.

En la práctica, podemos usar cualquiera de los promedios obtenidos de una de esas series y su distribución será tal que el intervalo de confianza puede expresarse como la desviación estándar de los promedios σ_e.