

# Varianza y desviación estándar

Sanabria Ugarte, Anahí.<sup>1</sup>

<sup>1</sup>*Departamento de Informática y Sistemas \**

(Dated: 26 de abril de 2023)

Las mediciones tomadas de forma directa, es decir, mediante un instrumento, deben ser expresadas de la forma correcta y en la notación internacional para el entendimiento de personas que no manejan el mismo idioma del autor o la misma simbología nativa que puede existir. Es por ello que saber nociones básicas de la estadísticas, nos permiten tener una escritura propia de ellas. A su vez, esto ayuda a leer otras mediciones, generalmente comerciales, y poder entender que es lo que se está por comprar.

**Palabras clave:** : medición, estadística, expresión, medida, dataframe.

## I. INTRODUCCIÓN

Al reportar mediciones de laboratorio en experimentos y, sobretodo, si son medidas directas, se encuentra una notación específica que ayuda a leer la medición de mejor forma. Generalmente, las mediciones viene del siguiente modo:

$$3,66 \pm 0,01[cm]; 0,29 \%$$

De esta forma, a simple vista y con pocos conocimientos, no se llega a entender más que las primeras cifras, que son lo que mide el objeto al que se consulta. Sin embargo, para hallar esos números tan raros, se han necesitado procesos matemáticos y estadísticos por detrás.

Cuando se habla de mediciones con instrumentos que poseen una cierta calibración, se habla, como se ha recalado en el artículo de mediciones de laboratorio [1], de incertidumbre o errores en la medición. En el trabajo anteriormente mencionado, se hace énfasis al estudio superficial de lo que puede llegar a ser un error. Sin embargo, este error no es el único dato necesario para expresar mediciones con su notación respectiva. Para ello, la estadística y la comparación con lo que es la media de un conjunto de datos, serán las mejores aliadas. Los conceptos que se verán en este artículo son considerados conceptos básicos para entender la estadística y el origen de muchas notaciones especiales de mediciones.

## II. OBJETIVO Y PLANTEAMIENTO DEL PROBLEMA

En este experimento, se centrarán los esfuerzos en dejar claro conceptos como varianza y desviación estándar, además de indagar un poco en el concepto de mediciones directas. Se recuerda que los conceptos de media se tocaron en el primer artículo con un experimento similar [1].

Cuando se tiene una medición expresada como anteriormente se mencionó, se debe empezar a reconocer

términos. Generalmente la notación es la anteriormente mencionada en la sección I. En el primer término se tendrá la media de una serie de mediciones o la medición única tomada con el instrumento. El segundo término, luego del símbolo más menos, es la desviación estándar, que delimita los intervalos en el que puede estar nuestro error respecto al primer término. El tercer término hace referencia a la unidad en la que se expresa la medición, siempre se debe escribir entre corchetes. Por último, el tercer término indica el porcentaje de error que hubo respecto a la media de la medición.

Cabe resaltar, el cálculo del porcentaje se hace mediante el error absoluto y el valor verdadero o media, términos que se profundizaron en el trabajo de mediciones de laboratorio [1].

Primeramente, se define:

**Varianza:** es el estudio de la dispersión de los datos respecto a su media [2].

**Desviación estándar:** es el intervalo cuantificado donde los datos dispersos están más cerca de la media. Los que se encuentren por fuera del intervalo que marca, son datos alejados de la media [2].

Una vez definidos, servirán para ver la utilidad de los cálculos respectivos. Adicionalmente podemos definir de forma muy sutil:

**Medidas directas:** son medidas tomadas por ecuaciones algebraicas e incapaces para los instrumentos de medición.

Entre las más comunes podemos encontrar el área y el volumen.

## III. MATERIALES Y MÉTODO

Para comenzar el experimento, se necesitará las herramientas de medición y las figuras o cuerpos a los que se realizará la medición. En el caso de este laboratorio, tomaremos una gran base de datos, la cual nos permitirá sacar resultados más precisos y ver el comportamiento. Se cuenta con una base de datos para 4 figuras y 10 tipos por cada una. Cada tipo tendrá mil datos a analizar.

Si se replica este experimento, no es necesario tener tantos datos, con unos cuantos podrían ser suficientes,

---

\* 202300536@umss.edu.bo

sin embargo tener muchos facilitan a tener más precisión en los cálculos de errores y desviaciones.

Adicionalmente, para los cálculos matemáticos se empleó el lenguaje de programación R en RStudio, ahí se evaluaron y generaron las tablas correspondientes.

La metodología que se usó fue empezar a reducir lo más que se pueda cada dataframe. Sacando medias y errores sin necesidad de sacar el de cada uno de los datos. A su vez, se usaron fórmulas específicas de cada propósito:

$$d_i = x_i - \hat{x} \quad (1)$$

donde:

$x_i$ : representa la medición realizada.

$\hat{x}$ : representa la media de las mediciones.

$d$ : discrepancia.

$$s^2 = \frac{1}{n} \sum_{i=1}^n d_i^2 \quad (2)$$

donde:

$s^2$ : representa la varianza.

$n$ : número de mediciones.

$d_i^2$ : representa la discrepancia de cada elemento al cuadrado.

$$\sigma = \sqrt{s^2} \quad (3)$$

donde:

$\sigma$ : representa la desviación estándar.

$s^2$ : representa la varianza.

$$densidad = \frac{masa}{volumen} \quad (4)$$

Se recalca que se profundizó en [1] acerca de los errores y la media, por lo tanto no se les hará mención en este trabajo.

#### IV. DESARROLLO EXPERIMENTAL

Para comenzar con el experimento, se requiere leer las bases de datos propuestas con las mediciones de cada figura. Esto se lo exporta en R como base de datos y se comienza a revisar las filas y columnas con las que cuenta. Para este caso tenemos figuras como: cilindro hueco, cilindro, esfera y moneda. Cada figura cuenta con tantas mediciones se haga como bases de datos. Cada base de

datos cuenta con 10 columnas para diez figuras del mismo tipo. Por lo tanto, tendremos que sacar las medias de cada base y cada columna.

Posterior, se usará una función creada llamada FunMed, la cual con el método apply se aplicará a cada columna y devolverá otro data frame con las medias respectivas.

Con el uso de las medias de cada medición, se procederá a calcular las mediciones indirectas. Para el caso de cada figura, se desea saber la densidad y el volumen respectivo. Para ello se crea la función **FunDen** y para la función **FunVol** se especificará una diferente de acuerdo a la figura que se tenga, es decir a base las ecuaciones (1), (2), (3) y usando las medias de cada dato. Así obtendremos tablas con muchos menos datos que si se sacara la densidad y volumen de cada una de las mediciones, que serían al rededor de cuatrocientos datos cada uno.

A cada media de las mediciones se aplicará una función llamada **estadística**, con ella se debe calcular las varianzas (2), desviación estándar y además encontrar los errores respecto a sus medias, como, el valor del error absoluto y el error porcentual. Con ello se podrá obtener los términos de la expresión, mencionada en la sección [II]. Para el primer término(media), segundo término(desviación estándar) (3) y tercer término (error porcentual)  $E\%$ .

Se procederá a graficar algunas mediciones con la función plot de R y armar las tablas correspondientes.

#### V. RESULTADOS

Cuadro I. Cálculos de mediciones indirectas de volumen y densidad hechos para el cilindro hueco.

Nro	Volúmen	Masas	Densidad
1	45.04021	45.56902	1.0117408
2	74.26053	55.23445	0.7437929
3	120.89146	58.20629	0.4814756
4	170.38827	81.17701	0.4764237
5	3192.40553	89.53687	0.4653550
6	200.27907	105.73528	0.5279397
7	231.47604	108.85233	0.470253
8	375.37922	128.44803	0.3421820
9	600.10677	128.95991	0.2148950
10	564.75175	139.75296	0.2474591

Cuadro II. Aplicación de las fórmulas estadísticas en los cilindros huecos.

	CH1	CH2	CH3	CH4	CH5	CH6	CH7	CH8	CH9	CH10
Media	3.66	4.25	5.20	5.87	7.34	8.15	8.60	8.73	9.04	9.90
DesvEst	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
ErrorPorc	0.29	0.25	0.19	0.18	0.16	0.13	0.13	0.14	0.12	0.11

## VI. DISCUSIÓN

Como se puede apreciar en la tabla 1 [V], las mediciones directas nos sirven para sacar las mediciones indirectas. En el caso del volumen y la densidad, se usaron los datos de los cilindros huecos, hablando del radio externo, interno y su alto.

Igualmente en el cuadro 2 [V] se puede ver que estas mismas mediciones, directas e indirectas, nos fueron de utilidad para hallar sus estadísticas. Con estas se puede armar una serie de cosas interesantes. En esta tabla se muestran la media, desviación estándar y error porcentual de las alturas de los 10 cilindros huecos evaluados con los datos de la tabla 1 [V]. Por lo tanto ya se puede hacer la interpretación de la escritura de las mediciones de la siguiente forma: primer dato a ser puesto será la media de la medición. Tomando como referencia el primer cilindro hueco, se tendrá que el primer término será 3.66. Posterior, se ubicará en la desviación estándar, la cual es 0.01. Las unidades de medida de cada figura están en [cm], por lo tanto, esta será la unidad de medida que se usará. Finalmente el error porcentual será el tercer y último término de nuestra expresión. Es así que nos queda de la siguiente forma:

$$3,66 \pm 0,01[cm]; 0,29 \%$$

Con esto, se puede decir que se ha hecho una expresión correctamente escrita y además que habla por sí misma sin necesidad de hacer mucha explicación sobre ella.

Continuando, se ve el comportamiento de algunas mediciones en relación de otras. Para ello se recurre a la primera figura [??]. Para una esfera se puede deducir lo siguiente: a mayor volumen, menor densidad. Esto también se puede comprobar al revisar la fórmula (4), ya que

el volumen inversamente proporcional a la densidad, por lo que el gráfico está correcto.

Además en la figura [??] se puede observar otro comportamiento en mediciones de una moneda. Evidenciando, por gráfico, que, a medida que el diametro de la moneda es mayor, la densidad de la moneda descende. Esto se debe a que la masa y el volumen son dependientes del diametro de la moneda y la densidad depende de ambos. Por lo tanto se concluye que la figura hace sentido.

## VII. CONCLUSIONES

Como principales conclusiones, se obtiene: Primeramente, el conocimiento de herramientas y ecuaciones básicas de estadística son útiles para conocer e interpretar muchos símbolos y resultados expresados con una notación especial. A su vez estos ayudan a la hora de realizar uno mismo un trabajo de medición y poder presentarlos correctamente y en forma de que todos puedan comprender lo que se quiere decir y expresar con ellos.

Como siguiente conclusión, se puede decir que en las ecuaciones de una medición respecto a otra, se puede deducir su comportamiento, hablando de si son inversas, proporcionales o dependientes.

Por último, se resalta que el experimento y el trabajo en general puede ser replicado por cualquier persona con el sumo cuidado de las mediciones y asegurándose que las fórmulas o ecuaciones utilizadas den los resultados sin mucho redondeo y de la forma más precisa posible. Se recomienda tener un gran volumen de datos para trabajarlas. Si las mediciones son tomadas a mano con un vernier o tornillo micrométrico, entonces entre 15 a 20 podrán ser un número ideal de mediciones.

## VIII. BIBLIOGRAFÍA

- 
- [1] A. Sanabria Ugarte, Mediciones y errores, Laboratorio de física general. , 1 (2023).  
 [2] P. Cervantes-Hernández, Media, varianza y desviación estándar, Ciencia y Mar **12**, 29 (2008).  
 [2] [1]