REVISTA ARGENTINA
DE TERAPIA INTENSIVA
2019 - 36 N° 1

¿Qué es y para qué sirve el Intervalo de Confianza?

Ladislao P. Díaz Ballve, Fernando G. Ríos, Marcela Mariano

Gabinete de Apoyo para la Producción de Información Hospitalaria (GAPIH), Hospital Nacional "Prof. Alejandro Posadas", El Palomar, Buenos Aires
Cátedra de Metodología de la Investigación Científica, Universidad Nacional de la Matanza, San Justo, Buenos Aires

Correspondencia:
Lic. Ladislao Díaz Ballve
ladislaodiaz@gmail.com

Los autores no declaran conflictos de intereses.

Para introducirnos en el concepto de intervalo de confianza, se debe saber que la pregunta de investigación implica, de manera implícita o explícita, a la población de estudio o población diana. Cuando se formula una pregunta de investigación, se debe identificar a quiénes o qué conjunto de individuos u objetos cumplen el(los) criterio(s) para responder la pregunta de investigación.

Para ilustrar la utilidad de la estimación por intervalos, supongamos que, ya cansado de especular acerca del peso de los pacientes que ingresan en la Terapia Intensiva, usted quiere conocer cuál es el peso promedio de los habitantes del área metropolitana de Buenos Aires. Entonces, no tiene más que salir a la calle con su balanza y pesar a toda la población, es decir, a varios millones de personas. Esta tarea es difícil de llevar a cabo o directamente imposible; sin embargo, respuestas a preguntas como esta se encuentran en una sencilla búsqueda bibliográfica en una base de datos de revistas de salud.

Entonces, medir a toda una población no es sencillo y digamos que, en realidad, es imposible, pero de todas maneras, se realizan investigaciones que informan los valores normales de diferentes poblaciones. ¿Cómo puede ser posible? La respuesta usted ya la conoce, las investigaciones generalmente se llevan a cabo con muestras y, a partir de dichas muestras, se realizan inferencias de la población de la cual proviene la muestra, y se señala la variabilidad que podría encontrarse en otras muestras (obtenidas de la misma población).

Continuando con nuestro ejemplo y conociendo que es posible utilizar una muestra, procedemos a seleccionar 200 habitantes, los pesamos, obtenemos un peso promedio de 72 kg ($\bar{x}$) con ± 5,6 kg de desviación estándar (DE). Sabemos ahora el valor promedio de nuestra pequeña muestra de 200 habitantes, pero nuestro objetivo, en realidad, es conocer el peso promedio de la población. ¿Existe solución a este problema? Una forma es pesando a los millones de habitantes que conforman la población, una vía poco viable. Habría otra manera de conocer el dato poblacional, a través de determinar el grado de incertidumbre, calculando un rango de valores entre los que se encuentra el verdadero valor poblacional. Este rango de valores se conoce como intervalo de confianza (IC).[1] Podemos definir a los IC como los rangos de valores entre los que se encuentra el verdadero promedio de población.

La construcción de estos IC se fundamenta en distintas teorías, una de ellas y en la que nos basaremos es la distribución de muestreo que es la distribución de probabilidad de una muestra de una población.[2]

Imaginemos que, de una determinada población, tomamos todas las muestras posibles de tamaño n y calculamos una estadística (por ejemplo, media) de todas las muestras. Si realizo una distribución de probabilidad de este estadístico (la media), obtendré una distribución de muestreo (no es lo mismo que la distribución poblacional), este tipo de distribución tiene ciertas características entre ellas la aleatorización, el tamaño, el conocimiento de la población. Al tomar muestras aleatorias y repetidas de tamaño n de una población y calculamos el promedio de cada una de estas muestras, no importa la forma de la distribución original de la población, la distribución de promedios seguirá una distribución normal. En la Figura 1, se presenta una muestra de distribución Uniforme (No normal) donde al aumentar el número de muestras adquiere una distribución normal o Gaussiana (Figura 2).

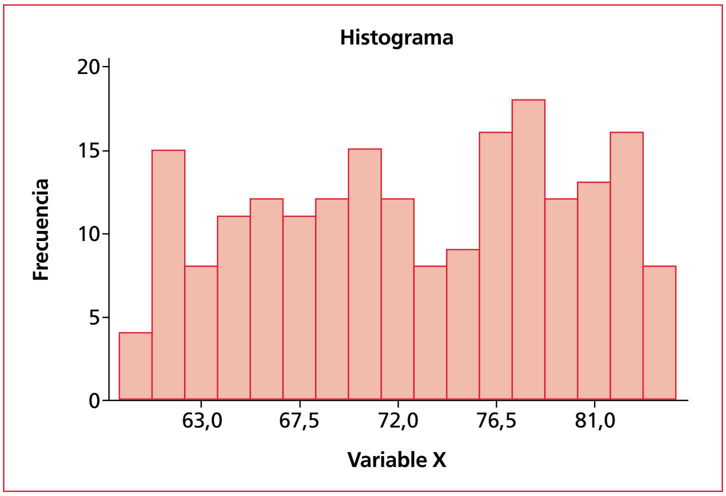

Figura 1. Distribución uniforme una muestra de la variable X.

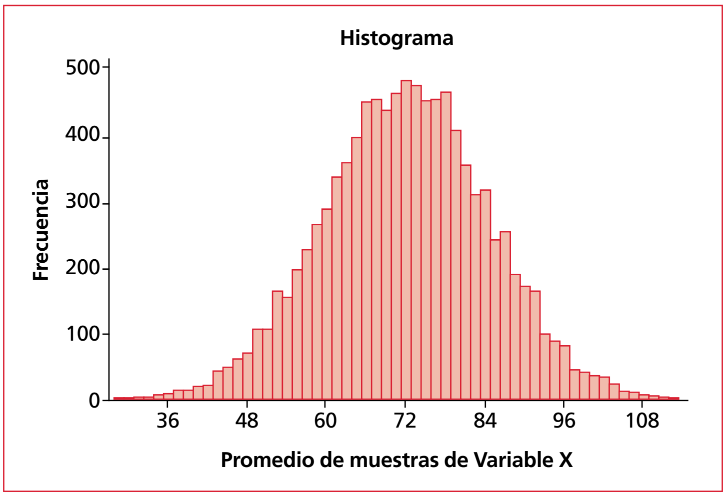

Figura 2. Al aumentar la cantidad de muestras provenientes de la misma población que la de la Figura 1, se observa una distribución Normal.

Este teorema se denomina teorema central del límite y es la base de toda la estadística inferencial, de este teorema se determinan las siguientes propiedades de la distribución de muestreo: la media de medias es igual a la media poblacional, la desviación estándar será igual a la desviación estándar dividida la raíz cuadrada de n ($\frac{s}{\sqrt{n}}$) llamado error estándar (EE) que es la variabilidad entre las medias de las muestras.[2,3]

Con esta información, regresemos al ejemplo inicial, contamos ahora con un recurso para conocer el verdadero valor promedio del peso de la población de los pacientes que ingresan en Terapia Intensiva sin medir a todos los habitantes. ¿Cómo es esto? Pues si tomamos varias muestras al azar se comportarán como una distribución Normal, cuantas más mejor. La media de esta distribución de muestras nos aproxima mucho a la media poblacional (μ) y su EE.

¿Y ahora cómo seguimos? Habiendo puesto al descubierto ciertos secretos ocultos de la estadística, decimos que el objetivo de la estimación por IC es obtener un intervalo, donde encuentra el verdadero valor del parámetro con una determinada probabilidad. Dicha probabilidad se denomina nivel de confianza, sabemos ahora que podemos basarnos en características de la distribución Normal donde queda establecido que el 95% del área central bajo la curva normal estandarizada se encuentra entre el valor $z = \pm 1{,}96$. Este valor de Z será usado para calcular el IC sabiendo que corresponde a un 95% de probabilidad de que contenga al parámetro poblacional.[4]

Por lo tanto, a través de ciertos artilugios algebraicos, la construcción del IC se realiza sumando y restándole al parámetro encontrado el valor de confianza estipulado (z) por el EE ($\bar{x} \pm z^*EE$) (Figura 3).

Cuanto más estricto sea el valor de confianza que defina el investigador, por ejemplo 90% (z=1,65) o 95% (z=1,96) o si desea una confianza aún mayor 99% (z=2,575) más ancho serán los intervalos. En la variabilidad del IC, también influye el tamaño de la muestra, cuanto menor es el n de la muestra será más amplio el intervalo, esto es para dar cobertura a la incertidumbre derivada de las muestras pequeñas, por eso un tamaño muestral grande produce IC más precisos. Otro factor es la variabilidad del objeto de estudio que influye, de manera importante, en la construcción del intervalo (Figura 4).[5]

¿Pero cuál sería la utilidad del IC? Para conocer el valor poblacional inaccesible al investigador, por ejemplo, la media poblacional ya no necesito "medir" a *todos los sujetos* de una población, puedo inferir el rango de valores entre los que se encuentra el verdadero valor poblacional. Y esto lo podemos lograr a partir de una sola muestra tomada al azar. El margen de error o de equivocarme en la estimación es del 5%

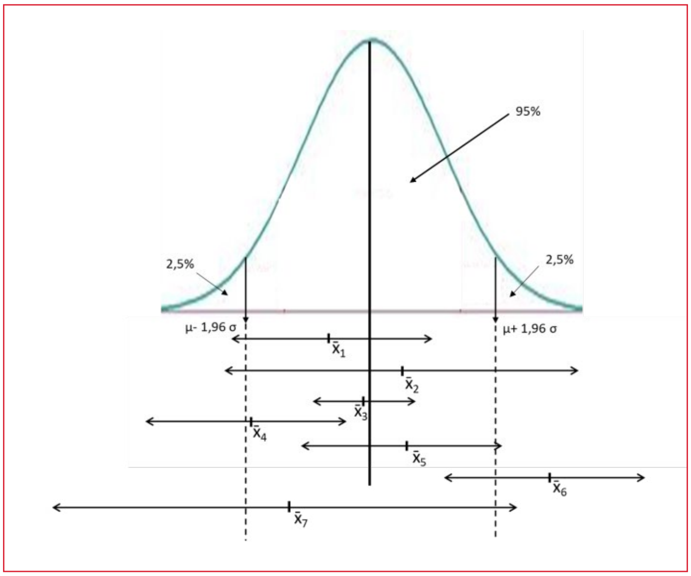

Figura 3. Interpretación del nivel de confianza según el intervalo de la media de una distribución Normal.

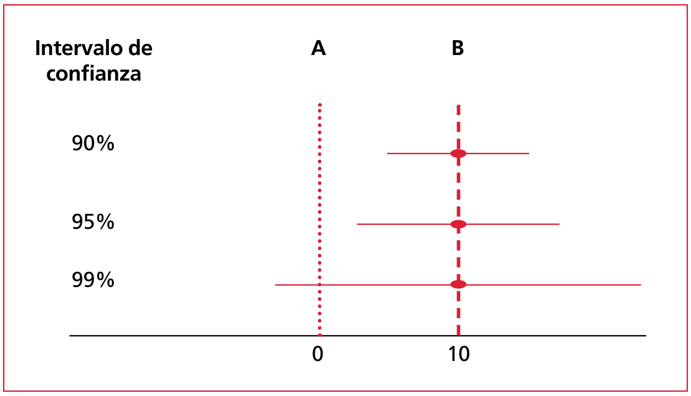

Figura 4. **A.** La hipótesis nula (hipótesis de no diferencias), **B.** La estimación puntual de la diferencia encontrada. Las barras muestran la amplitud de los intervalos de confianza del 90%, 95% y 99%, y el valor de P inferido del hecho de que el intervalo de confianza respectivo incluye o no a la hipótesis nula.

si calculo un IC del 95%, aunque se puede ser más preciso y disminuir el grado de error al 1% calculando un IC del 99%.[2]

Se debe aclarar que es posible calcular un IC para cualquier estadístico (DE, Varianza, Mediana, *odds ratio* o cualquier estadístico que el investigador necesite utilizar). Es importante remarcar que el cálculo del IC no varía sustancialmente entre los distintos estadísticos (IC = Estadístico ± z *EE), lo que sí varía es la manera en que se calcula el EE, para el caso de la media, no es un procedimiento complejo. Pero para otros estadísticos suelen ser fórmulas más complicadas. En la actualidad, casi todos los programas informáticos que se utilizan para cálculos estadísticos pueden estimar el EE de cualquier estadístico.

A modo de conclusión, podemos decir que resulta útil presentar los resultados con IC, mediante ellos se representa la variabilidad encontrada en los estadísticos utilizados. De esta manera, le proporcionamos al lector información de mayor calidad y se complementan los resultados de una prueba estadística única donde solo se informa el valor p (hoy cuestionado en muchos aspectos). Sin embargo, utilizar la estimación por intervalos como único determinante, al igual que el valor p como método para la toma de decisiones, son enfoques incorrectos y ambos recursos al igual que otros deben ser interpretados con cautela y alejados de posiciones rígidas y definitivas.

Bibliografía

1. Fethney JBA. Statistical and clinical significance, and how to use confidence intervals to help interpret both. Aust Crit Care 2010; 23(2): 93-97.
2. Dawson B, Trapp RG. *Basic & clinical biostatistics*, 4th ed. New York: McGraw-Hill, Medical Pub. Division; 2004:438.
3. Daniels W. *Bioestadística*, 3ra ed. México D.F.: LIMUSA; 2006:735.
4. Ranstam J. Why the P-value culture is bad and confidence intervals a better alternative. *Osteoarthr Cartil* 2012;20: 805-808.
5. Mc Cormack J, Vandermeer B, Allan GM. How confidence intervals become confusion intervals. *BMC Med Res Methodol* 2013;13(1):134.