



Revista Colombiana de Psiquiatría

ISSN: 0034-7450

revista@psiquiatria.org.co

Asociación Colombiana de Psiquiatría
Colombia

Castañeda, Javier A.; Fabián Gil, Jacky
Una mirada a los intervalos de confianza en investigación
Revista Colombiana de Psiquiatría, vol. XXXIII, núm. 2, 2004, pp. 193-201
Asociación Colombiana de Psiquiatría
Bogotá, D.C., Colombia

Disponible en: <http://www.redalyc.org/articulo.oa?id=80633208>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica
Red de Revistas Científicas de América Latina, el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

Metodología

de investigación y lectura crítica de estudios

Una mirada a los intervalos de confianza en investigación

Javier A. Castañeda¹, Jacky Fabián Gil²

Resumen

Frecuentemente los resultados de una investigación provienen del análisis de una muestra, y es necesario generalizar esta información sobre el total de elementos que conforman la población de estudio. Diversas metodologías para este fin están disponibles a través de procedimientos estadísticos, como las pruebas de hipótesis y los intervalos de confianza. Sin embargo, especialmente en los últimos treinta años, las publicaciones científicas médicas han visto la creciente necesidad de acompañar los resultados publicados con medidas que permitan determinar su grado de precisión e incertidumbre, cuando son generalizados sobre la población. De esta forma, la estimación de los intervalos de confianza se convierte en la metodología estadística más ampliamente usada y la que responde a esta creciente necesidad.

El presente artículo ilustra la filosofía que se maneja en la construcción de los intervalos de confianza, presenta y discute los elementos que los componen y explica la forma correcta de interpretación que se debe hacer de ellos.

Palabras clave: estimación, nivel de confianza, error muestral, precisión.

Title: A Look at Confidence Intervals in Research

Abstract

Research findings frequently rise from analysis of a sample and it is necessary to infer the conclusions for the population where the sample elements come from. Statistical methods such as confidence intervals and hypothesis testing allow inferences to be made from a sample to a specific population. Notably over the last 30 years, medical scientific publications have increasingly needed to report research results using measures that specify the degree of precision and uncertainty when conclusions are generalized to a population. Thus, confidence intervals have become the most common methodology used to report scientific results. This paper shows the philosophy behind the construction of confidence intervals as well as the presentation and discussion of their elements beside an explanation of the correct way to interpret them.

¹ Estadístico, profesor instructor de la Unidad de Epidemiología Clínica y Bioestadística de la Pontificia Universidad Javeriana. Bogotá, Colombia.

² Bioestadístico, profesor instructor de la Unidad de Epidemiología Clínica y Bioestadística de la Pontificia Universidad Javeriana. Bogotá, Colombia.

Key words: estimation, confidence level, sampling error, precision.

Introducción

Un grupo de investigadores en psiquiatría interesado en conocer la prevalencia del trastorno depresivo en adultos mayores de sesenta años residentes en Bogotá seleccionó una muestra de noventa personas mayores de sesenta años de la ciudad. Luego de analizar la información recolectada durante la evaluación de los sujetos estudiados, se encontró que nueve personas de la muestra sufren trastorno depresivo. Los investigadores planean comunicar los resultados de la investigación, y ante ello surgen las tres preguntas siguientes: ¿la prevalencia del 10% encontrada en la muestra significa que el 10% de la población de adultos mayores en Bogotá sufre del trastorno depresivo? Puesto que la selección de los sujetos se realizó de forma aleatoria, ¿habría obtenido el mismo valor de prevalencia otra muestra seleccionada del mismo tamaño?, y ¿qué efecto tiene el tamaño de la muestra del estudio para la generalización de los resultados? Dada la relevancia que los investigadores suponen para los resultados de la investigación, desean responder estas preguntas antes de continuar con la comunicación de sus hallazgos.

Las tres preguntas anteriores se refieren a un aspecto común: la

generalización de los resultados, es decir, cómo derivar conclusiones para los individuos de la población de estudio a partir de los resultados del análisis de las unidades contenidas en la muestra. Por fortuna todos estos interrogantes tienen respuesta dentro de la teoría estadística, específicamente a través de la inferencia estadística.

La inferencia estadística es el procedimiento por medio del cual se obtienen conclusiones de la población a partir de una muestra o subconjunto de unidades seleccionadas de ésta. La inferencia comprende dos áreas de trabajo denominadas estimación y pruebas de hipótesis (1). La estimación implica derivar, a partir de la información de la muestra, un valor o un rango de valores (llamado intervalo de confianza), que permitan concluir sobre una medida de interés en la población (2). De esta forma, las preguntas de los investigadores se refieren a la estimación de un parámetro en la población: la proporción de adultos mayores con trastorno depresivo residentes en Bogotá.

Desde mediados de los años ochenta, casos como el presentado en el anterior ejemplo han motivado un gradual incremento en el uso y reporte de medidas y metodología estadísticas en los artículos publicados en revistas médicas (3). Por ello, el presente artículo tiene los objetivos de ilustrar la filosofía que se

maneja en la construcción de los intervalos de confianza (una de las medidas estadísticas más ampliamente usadas para reportar resultados), presentar y discutir los elementos que los componen y precisar la interpretación correcta que se debe hacer de ellos.

Determinación de valores de interés en la población

Cuando el objetivo es determinar un valor desconocido en la población, se cuenta con diferentes alternativas, las cuales dependen de los recursos de que se disponga en la investigación. Estos recursos pueden ser de tipo económico, logístico o de tiempo. Cuando se cuenta con los recursos suficientes, la mejor opción para determinar un valor de interés en la población es realizar un censo, es decir, analizar todos los elementos que componen la población de estudio; de esta forma es posible calcular de forma exacta el parámetro (valor de interés) en la población.

En otro caso, cuando no se cuenta con los suficientes recursos o existen conflictos éticos que impiden la aplicación de una intervención a toda una población (por ejemplo, al probar nuevos tratamientos en seres vivos), se debe seleccionar un subconjunto de individuos de la población sobre los cuales se llevará a cabo el desarrollo de la investigación. Estos individuos

seleccionados de la población corresponden a lo que en el campo de la estadística se denomina muestra (4).

En las investigaciones en las que se hace uso de una muestra, la idea es usar la información que se obtiene sobre ella para dar conclusiones acerca de la población de estudio (inferencia). De esta forma, los resultados derivados de la muestra corresponden a estimadores de los valores en la población; por ello, es necesario garantizar que la muestra sea representativa de la población de interés y que las estimaciones derivadas gocen de precisión. Ésa es la razón por la cual la selección de una muestra requiere, entre otros aspectos, la planeación de un método para escoger los elementos de la población y el cálculo del tamaño de la muestra o número de unidades que la conformarán. Las estimaciones derivadas de la muestra pueden ser de dos tipos: estimaciones puntuales o estimaciones por intervalos de confianza.

Estimación puntual

Como su nombre lo indica, en la estimación puntual el valor del parámetro poblacional es calculado por medio de un único valor derivado de la muestra (5). Para el caso de nuestro ejemplo, los investigadores —basados en los resultados de la muestra— comunican que el 10%

de la población de adultos mayores de Bogotá presenta trastorno depresivo, puesto que este valor es producto del cálculo de la prevalencia sobre la muestra recolectada. Este tipo de cálculo, a pesar de ser muy precisa al reportar un único valor como estimador, presenta fuertes limitaciones desde el punto de vista estadístico, dado que es poco confiable, porque el resultado obtenido depende de la muestra seleccionada. De esta forma, si el grupo de investigadores tomara una muestra diferente, los resultados para prevalencia del trastorno depresivo probablemente serían diferentes.

La estimación puntual no permite especificar las variaciones de la estimación sobre otras posibles muestras, por lo cual no es viable derivar una medida que permita determinar con qué grado de certidumbre el valor obtenido en la muestra refleja (infiere) el verdadero valor en la población. Para corregir esta deficiencia se crearon los intervalos de confianza (6).

Estimación por intervalos
de confianza

Un intervalo de confianza es un conjunto de valores, con un límite inferior y uno superior, que con una determinada confianza se espera que contenga el valor del parámetro de interés en una población específica (7). Además, los intervalos de confianza controlan el efecto de

obtener resultados diferentes debidos al azar en la selección de muestras distintas (error muestral). Cabe anotar que los intervalos de confianza no pueden controlar errores no muestrales, como los sesgos en el diseño o en la conducción del estudio (3).

La estimación por intervalos se basa en la idea del efecto del error muestral sobre los resultados de la investigación, pues, particularmente, el hecho de seleccionar diferentes muestras puede arrojar diferentes resultados, pero alrededor del verdadero valor del parámetro poblacional (8). De forma general, un intervalo de confianza se compone de una estimación puntual, de una medida de variabilidad llamada error estándar y de un coeficiente de confiabilidad (5). El cálculo de los límites del intervalo de confianza consiste en restar y sumar a la estimación puntual el resultado del producto entre el coeficiente de confianza (CC) y el error estándar (ee) (fórmula 1).

Fórmula 1. (Estimador puntual)
± [(Coeficiente de confianza) x
(error estándar)]

Para garantizar la validez de la estimación por intervalos de confianza la muestra debe ser seleccionada a través de un método de muestreo probabilístico (2), es decir, el método de muestreo debe ser objetivo, a fin de garantizar que todos los elemen-

tos de la población tengan alguna oportunidad de conformar la muestra. A la vez, la muestra debe reflejar el comportamiento o distribución de las variables de interés de la población de estudio (9).

El cálculo de intervalos de confianza es posible para diversos estimadores, como la media o la proporción muestral, la diferencia de medias o proporciones muestrales y riesgos relativos o riesgos relativos indirectos muestrales, entre otros. Para un gran número de parámetros, por ejemplo, la media o la proporción de una población, la estimación puntual corresponde al valor central del intervalo de confianza; sin embargo, en algunos otros casos, como la estimación del riesgo relativo, la estimación puntual no coincide con el valor medio del intervalo de confianza.

El intervalo de confianza representa un rango de valores posibles para el parámetro estimado en la población. Su amplitud está determinada de manera conjunta por el error estándar del estimador y el nivel de confianza deseado. De esta forma, el intervalo de confianza será más amplio si la magnitud del error estándar es grande o el nivel de confianza deseado es cercano al 100%.

El error estándar es la medida de variabilidad del estimador usada en la construcción del intervalo de confianza, y corresponde a una

función de la desviación estándar de la variable de interés y del tamaño de la muestra. Así, si el interés de un estudio fuera estimar la edad media de las personas que padecen esquizofrenia, el estimador que se debe emplear corresponde a la media muestral y su error estándar es σ/\sqrt{n} , donde σ es la desviación estándar poblacional para la variable edad y n es el tamaño de la muestra.

Para observar el efecto de la desviación estándar sobre la amplitud del intervalo de confianza, considérese que en el ejemplo anterior las personas de la población de estudio tienen edades muy similares. De esta forma, es de esperar que las edades de los individuos seleccionados en la muestra sean muy homogéneas y la desviación estándar muestral tendrá una magnitud pequeña, que se traducirá en una menor amplitud del intervalo de confianza. En el caso contrario, si las edades de los individuos fueran muy heterogéneas, se tendría un aumento en la amplitud del intervalo de confianza.

El efecto del tamaño de la muestra sobre la amplitud del intervalo de confianza se puede notar fácilmente, puesto que en el cálculo del error estándar éste se halla contenido en el denominador, de donde se establece una relación inversa. Así, al disminuir el tamaño de la muestra, aumenta la magnitud del error estándar y, por ende, aumen-

ta la amplitud del intervalo de confianza y se reduce la precisión en la estimación.

El coeficiente de confiabilidad corresponde a un valor percentil en la correspondiente función de distribución muestral del estimador. Es decir, dependiendo del parámetro que se desee estimar (medias, proporciones, etc.) y del tamaño de la muestra, se establece la distribución y el valor correspondiente al nivel de confianza deseado para el cálculo del intervalo de confianza. Por ejemplo, si se toma una muestra de 250 pacientes y se quiere calcular un intervalo del 90% de confianza para estimar el promedio de edad de las personas que padecen esquizofrenia, el valor del coeficiente

de confianza (CC) se determina a partir de la distribución normal estándar, el cual corresponde al valor 1,64.

A pesar de que el nivel de confianza más frecuentemente usado en el reporte de resultados es 95%, el investigador igualmente puede emplear otros valores que expresen mayores o menores niveles de confianza, como 99% o 90%, respectivamente. En la Tabla 1 pueden observarse los valores del coeficiente de confianza para los niveles de confianza usados con mayor frecuencia para algunos parámetros. Cabe notar que a mayor nivel de confianza, mayor el valor del coeficiente de confianza, lo cual implica mayor amplitud del intervalo calculado.

Tabla 1. Valores del coeficiente de confianza para estimar la media (con tamaño de muestra grande) y la proporción de una población

Nivel de confianza	9 0 %	9 5 %	9 9 %
Coeficiente de confianza	1,64	1,96	2,57

A continuación se presentan dos ejemplos del uso de intervalos de confianza y la forma correcta de reportar los resultados.

Ejemplo 1: supóngase que a partir de una muestra aleatoria de 144 pacientes se quiere calcular un intervalo del 95% de confianza para la edad media de las personas que padecen esquizofrenia, cuando se

conoce que la desviación estándar de la edad en la población es 6,2 años y el promedio de edad calculado en la muestra es de 17 años. A partir de la información conocida y puesto que el tamaño de la muestra es suficientemente grande, se puede aceptar el supuesto de normalidad y determinar el error estándar como $0/\sqrt{n} = 6,2/12 = 0,52$. De la misma forma, para un nivel de confianza

del 95% el coeficiente de confianza es de 1,96 y el cálculo del intervalo de confianza se realiza reemplazando los términos en la Fórmula 1:

$$17 \pm z(1,96 * 0,52) = 17 \pm z1,01 = [16; 18]$$

La forma de reportar el resultado del intervalo de confianza es la siguiente: a un nivel de confianza del 95%, la edad media de la población de personas que padecen esquizofrenia está entre 16 y 18 años.

Para ver el efecto que tiene el tamaño de la muestra en la amplitud (precisión) del intervalo, supongamos un $n = 30$. De esta forma el intervalo de confianza es [14,7; 19,2], el cual tiene más del doble de la amplitud del intervalo con $n = 144$.

Ejemplo 2: continuando con el caso de los investigadores interesados en conocer la prevalencia del trastorno depresivo en adultos mayores de sesenta años residentes en Bogotá, se sabe que tomaron una muestra aleatoria de noventa individuos y que la estimación puntual de la prevalencia fue 10%. Con esta información podemos construir un intervalo de confianza.

Primero, se debe especificar el nivel de confianza del intervalo, digamos 99%. Basados en el nivel de confianza y en el tamaño de la muestra, se determina el coeficiente de confianza (CC = 2,57). El siguiente

elemento que se va a calcular es el error estándar para la estimación de una proporción (5), ($ee = \hat{p}(1-\hat{p})/n = 0,1(1-0,1)/90 = 0,031$). Reemplazando en la Fórmula 1:

$$0,10 \pm z(2,57 * 0,03) = 0,10 \pm z0,08 = [0,02; 0,18]$$

La forma de reportar el intervalo de confianza es la siguiente: a un nivel de confianza del 99%, la prevalencia del trastorno depresivo en adultos mayores de sesenta años residentes en Bogotá está entre 2% y 18%.

Para ver el efecto que tiene el tamaño de la muestra en la amplitud (precisión) del intervalo para la prevalencia, supongamos un $n = 450$. De esta forma, el intervalo de confianza es [6,3%; 13,6%], el cual es aproximadamente de la mitad de la amplitud del intervalo con $n = 90$.

¿Qué quiere decir que un intervalo es del 95% de confianza? La interpretación es la siguiente: si se toman cien muestras de un mismo tamaño y se calculan sus respectivos intervalos de confianza del 95%, es de esperar que 95 de los intervalos contengan el verdadero valor del parámetro y los restantes cinco intervalos no lo contengan (Figura 1). En la práctica, sólo se puede tomar una muestra y construir un intervalo de confianza. De esta forma, si

el intervalo de confianza es del 95%, la oportunidad de que el intervalo sea desafortunado, es decir, de que no contenga el verdadero valor del parámetro de la población, es del 5%.

Figura 1. Conjunto de 100 intervalos de confianza del 95% para la prevalencia del trastorno depresivo en adultos mayores de sesenta años residentes en Bogotá, construidos de muestras de tamaño 450. Prevalencia (P) en la población del 11%.



La confianza frente al error. Estos dos conceptos siempre desempeñan un papel importante en el momento de calcular intervalos de confianza. El siguiente ejemplo ilustra su naturaleza: una persona quiere fijar una hora de encuentro en una cafetería con otra persona y establece las siguientes dos opciones: (a) encontrarse entre la 1:00 y la 1:10 de la tarde en la cafetería,

con una confianza del 50% de que llega a la cita y (b) encontrarse entre las 9:00 de la mañana y las 4:00 de la tarde en la cafetería, con una confianza del 100% de que llega a la cita en ese período. Fijémonos en que el margen de error en la primera opción es de solo diez minutos, pero la confianza de que llegue es muy baja, lo cual en términos prácticos no resulta conveniente o útil. De igual forma, la segunda opción muestra que a pesar de manejarse un nivel alto de confianza, esto no es útil cuando el nivel de error es demasiado grande. Lo anterior también sucede en la construcción de los intervalos de confianza; por lo tanto, se deben construir intervalos de confianza lo suficientemente precisos y confiables para que sean útiles de acuerdo con los objetivos del estudio.

Comentarios

Se aconseja que el investigador establezca los niveles de error y de confianza más convenientes para alcanzar resultados útiles y prácticos, de acuerdo con los objetivos de su estudio. De igual forma, si se planea usar un intervalo de confianza, es recomendable hacer un cálculo del tamaño de la muestra para niveles de confianza y de precisión establecidos; así es posible hacer una evaluación preliminar de la factibilidad del estudio, en términos de los recursos económicos, logísticos y de tiempo.

Bibliografía

1. Casella G, Berger R. Statistical inference. 2nd ed. Pacific Grove (CA): Duxbury; 2002.
2. Pagano M, Gauvreau K. Principles of biostatistics. 2nd ed. Pacific Grove (CA): Duxbury Thomson Learning; 2000.
3. Altman D, Machin D, Bryant T, Gardner M. Statistics with confidence. 2nd ed. Bristol: BMJ books. 2000.
4. Cochran WG. Sampling techniques. 3rd ed. Nueva York: John Wiley and Sons; 1977.
5. Daniel W. Bioestadística base para el análisis de las ciencias de la salud. México: Limusa; 2002.
6. Dawson B, Trapo RG. Basic and clinical biostatistics. 3rd ed. Nueva York: McGraw-Hill; 2001.
7. Hogg R, Craig A. Introduction to mathematical statistics. New York: Macmillan Publishing; 1978.
8. Sackett DL, Richardson WS, Rosenberg W, Haynes RB. Evidence based medicine: how to practice and teach EBM. Londres: Churchill Livingstone; 1997.
9. Ospina D. Introducción al muestreo. Bogotá: Unibiblos; 2001.

Correspondencia: Javier A. Castañeda

Hospital Universitario San Ignacio

Pontificia Universidad Javeriana

Unidad de Epidemiología Clínica y Bioestadística

Cra. 7 No. 40-62, piso 2

Correo electrónico: javier.castaneda@javeriana.edu.co