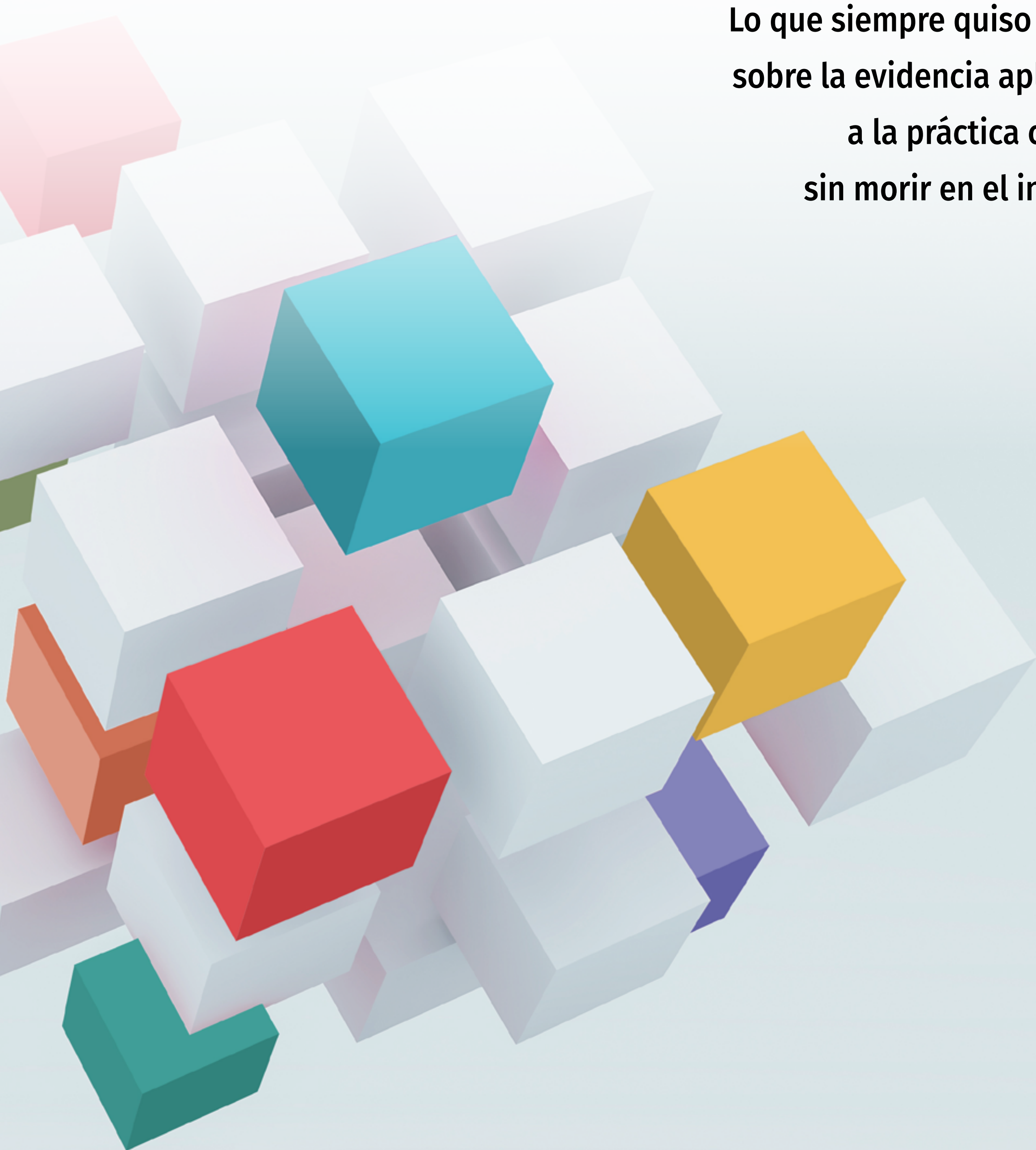
Medicina
Basada en la Evidencia

Lo que siempre quiso saber
sobre la evidencia aplicada
a la práctica clínica
sin morir en el intento



Grupo Pediatría Basada en Evidencias AEPap

aepCPBE Comité de Pediatría Basada En la Evidencia

Cómo citar este documento:
Comité de Pediatría Basada en la Evidencia de la AEP; Grupo de Trabajo de Pediatría Basada en la Evidencia de la AEPap (coords.). Medicina Basada en la Evidencia. Lo que siempre quiso saber la evidencia aplicada a la práctica clínica sin morir en el intento. Madrid: Lúa Ediciones; 2024.

Han participado como coordinadores del libro:

Comité de Pediatría Basada en la Evidencia de la AEP:
Paz González Rodríguez, María Aparicio Rodrigo, Pilar Aizpurua Galdeano, M.ª Jesús Esparza Olcina, Eduardo Ortega Páez

Grupo de Trabajo de Pediatría Basada en la Evidencia de la AEPap:
M.ª Salomé Albi Rodríguez, Javier González de Dios, Rafael Martín Masot, Carlos Ochoa Sangrador, Manolo Molina Arias

Realización y edición:
Lúa Ediciones 3.0, S.L.
www.luaediciones.com

ISBN: 978-84-128758-1-2

6.4.

Inferencia estadística: estimación puntual y por intervalos

Eduardo Cuestas Montañés
Servicio de Pediatría y Neonatología. Hospital Privado Universitario de Córdoba. Argentina

Manuel Molina Arias
Servicio de Gastroenterología Pediátrica. Hospital Infantil Universitario La Paz. Madrid. España

Carlos Ochoa Sangrador
Servicio de Pediatría. Complejo Asistencial de Zamora. Zamora. España

"La estimación puntual es como tratar de clavar un clavo en movimiento: nunca estás seguro de si realmente lo has acertado"

Bradley Efrom

OBJETIVOS:

- Distinguir entre estimación puntual y por intervalos de confianza
- Comprender el concepto de error estándar de una distribución de muestras
- Saber interpretar un intervalo de confianza
- Aprender a calcular un intervalo de confianza para una media
- Aprender a calcular un intervalo de confianza para una proporción

Una población representa la totalidad de personas que nos interesa estudiar. En general, estudiar a toda una población es laborioso y costoso en términos de tiempo y recursos y, en algunos casos, imposible, porque la población puede ser en principio abstracta (por ejemplo, cuando estudiamos pacientes que podrían desarrollar una complicación en el futuro). Por lo tanto, habitualmente recolectamos datos de una muestra o subconjunto de personas representativas de esa población, es decir, que tienen características similares a las de la población en estudio, y los utilizamos para obtener conclusiones, o sea, hacer inferencias sobre esa población.

Los datos obtenidos de este modo no nos informan un valor exacto e invariable, ya que se obtienen solo de una parte de toda la población. Además, si tomáramos varias muestras de la misma población obtendríamos valores diferentes en cada una de ellas.

La inferencia estadística pretende obtener un valor que se acerque lo más posible al supuesto valor real, que nunca conoceremos. Esta aproximación puede tener dos enfoques diferentes: la estimación puntual y la estimación por intervalos de confianza.

PARÁMETROS, ESTADÍSTICOS Y ESTIMADORES

Como ya se ha dicho, pocas veces podemos acceder al valor exacto de un parámetro de toda la población, como una media o una proporción, por lo que habitualmente estimamos el valor del parámetro utilizando los datos recogidos de una muestra.

Esta estimación de un estadístico muestral es una estimación puntual del parámetro, es decir, toma un único valor, que será solo una de las múltiples estimaciones puntuales teóricas, obtenidas del análisis de otras posibles muestras seleccionadas a partir de la población.

Las estimaciones puntuales varían de una muestra a otra. La estimación por intervalos de confianza nos permite tener en cuenta esta imprecisión y establecer unos límites entre los que se encontrará el parámetro poblacional que queremos estimar

Por ello, la forma más prudente de facilitar cualquier estimación es presentarla como un rango de valores entre los que puede estar el parámetro poblacional a estimar. Este rango constituye lo que denominamos intervalo de confianza. Habitualmente estimamos intervalos con un 95% de confianza, que interpretaremos, de forma sencilla, como que el parámetro poblacional tiene una probabilidad del 95% de encontrarse dentro de esos límites.

INTERVALO DE CONFIANZA Y ERROR ESTÁNDAR

Para estimar los intervalos de confianza partiremos de las estimaciones puntuales de nuestra muestra, habitualmente, una frecuencia relativa (proporción) para una variable nominal dicotómica, o una media (media muestral) para una variable continua, que son nuestras mejores aproximaciones al parámetro poblacional que hay que estimar: una proporción (p) o una media (μ). Pero como han sido estimadas en muestras, por prudencia, solo podemos decir que los parámetros a estimar tendrán valores cercanos a los que hemos obtenido en nuestra muestra.

Pero, ¿cómo calculamos el intervalo de error alrededor de nuestras medidas muestrales? La aproximación más intuitiva es intentar saber si las proporciones o medias muestrales siguen algún tipo conocido de distribución de probabilidad.

Hagamos un ejercicio teórico a partir de los datos de una amplia muestra de casos de partos (12 000), de la que vamos a seleccionar muestras de creciente tamaño muestral. Nuestro objetivo es estimar la proporción de partos distócicos que hay en esa población. No olvidemos que, incluso con muestras de gran tamaño muestral, las muestras solo son aproximaciones a la población. Asimismo, recordemos que, en el mundo real, este ejercicio teórico no es posible, ya que en nuestros estudios solo vamos a contar con una única muestra (que podría ser la que aquí vamos a usar como población).

Para calcular el error estándar de una proporción debemos conocer su valor en la muestra, su complementario y el tamaño muestral

Hemos seleccionado 120 muestras aleatorias de tamaño n = 10 (cada muestra 10 partos). En cada una de las 120 muestras estimamos la frecuencia relativa de parto distócico, con lo que obtendremos 120 proporciones, que utilizaremos como si fueran observaciones individuales, que adoptarán valores entre 0 y 1, para confeccionar un histograma de frecuencias (**Figura 1**). Como la proporción de partos distócicos en la muestra que hemos empleado como población es 0,15 (15%), la mayoría de los valores van a estar entre 0,10 y 0,20; es importante advertir de nuevo que esa información que sabemos aquí es siempre desconocida. Junto al histograma podemos ver la media y desviación típica de esas observaciones, que han pasado a ser consideradas como variables continuas. El histograma izquierdo nos dice que la media es 0,15 y la desviación típica 0,115.

Ahora vamos a seleccionar 120 muestras aleatorias de tamaño n = 20 (cada muestra 20 partos). Haciendo el mismo procedimiento obtenemos un nuevo histograma de frecuencias (el de la posición central), con estimaciones de media 0,15 y desviación típica 0,081.

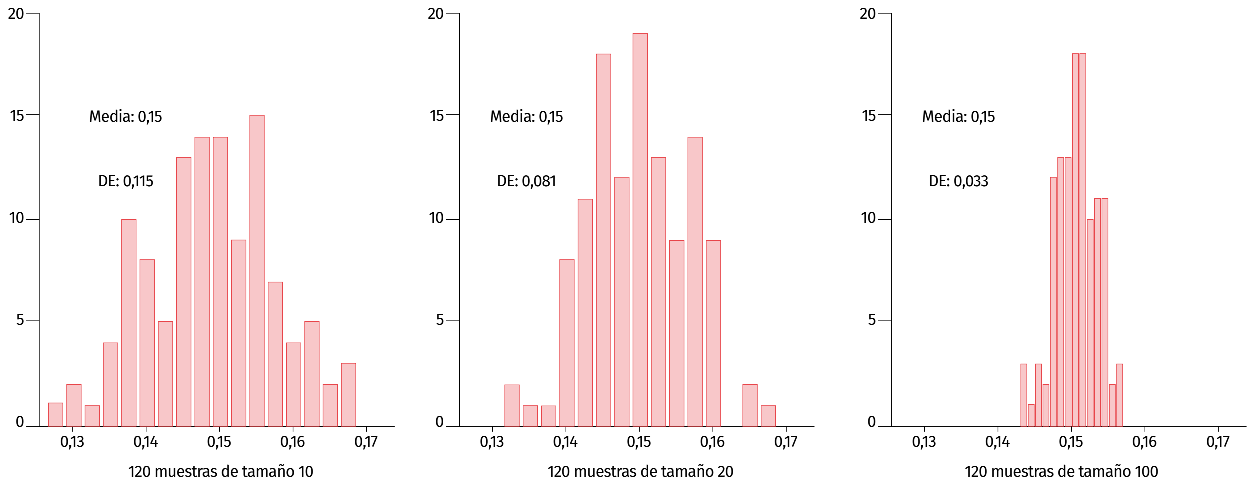
Finalmente, vamos a seleccionar 120 muestras aleatorias de tamaño n = 100 (cada muestra 100 partos). Haciendo el mismo procedimiento obtenemos un nuevo histograma de frecuencias (el de la derecha de la figura), con estimaciones de media 0,15 y desviación típica 0,033.

En los histogramas (**Figura 1**) podemos ver que, aunque el valor teórico poblacional era 0,15 (en el conjunto de los 12 000 partos), las proporciones de partos distócicos en las muestras individuales son muy variables, aunque la media de todas las proporciones coincide.

Podemos ver también cómo a mayor tamaño muestral los valores obtenidos son menos dispersos, lo que se traduce en un histograma más estrecho, con una apariencia cada vez más acampanada, y en una desviación típica progresivamente menor (0,115; 0,081; 0,033). Curiosamente, se puede deducir que el factor que pondera dicha dispersión es el tamaño muestral siguiendo una relación fija: raíz cuadrada del cociente de la proporción (0,15) por su complementario (0,85) y dividido por el tamaño de las muestras (10 o 20 o 100):

$$\sqrt{\frac{0,15\times0,85}{10}} = 0,112 \qquad \sqrt{\frac{0,15\times0,85}{20}} = 0,080 \qquad \sqrt{\frac{0,15\times0,85}{100}} = 0,035$$

Figura 1. Distribuciones muestrales con muestras de diferente tamaño. DE: desviación estándar



Esta relación corresponde a lo que conocemos como error estándar o error típico, que se comporta como el factor de dispersión (desviación típica) de la distribución de proporciones muestrales que, al igual que otros estimadores, sigue una distribución normal (según el teorema del límite central para muestras n ≥30). El error estándar se puede calcular a partir de la proporción que hemos encontrado en nuestra muestra, su complementario (1-p) y el tamaño muestral (la aproximación a la normal es válida si el producto [n × p × 1-p] es mayor que 5).

$$Error\ estándar_{proporción} = \sqrt{\frac{p(1-p)}{n}}$$

Como ya dijimos, nuestro objetivo era estimar un parámetro poblacional, pero no sabíamos cómo cuantificar nuestra incertidumbre para dar los valores entre los que es verosímil que se encuentre. Ahora ya tenemos los elementos que necesitamos para estimar el parámetro poblacional a partir de las medidas descriptivas de nuestra muestra. Como solo vamos a tener una muestra, asumiremos la proporción observada como el punto central de la estimación (estimación puntual). Utilizando esa proporción y el tamaño muestral, calcularemos el error estándar, y utilizando las propiedades de la distribución normal, podemos cuantificar entre qué rango de valores se encuentra.

Una vez conocida la estimación puntual de una proporción y su error estándar, es posible utilizar las características de la distribución normal para calcular el intervalo de confianza de la proporción

Imaginemos que hemos realizado un estudio con una muestra de 100 partos y que hemos encontrado una proporción de 0,17 (17 de los 100, 17%, tuvieron parto distócico). Lo más probable es que hubiéramos encontrado un 15%, pero por azar hemos encontrado otro resultado cercano. Asumimos ese 0,17 como media y calculamos el error estándar con la fórmula:

$$Error\ estándar_{proporción} = \sqrt{\frac{0{,}17(1-0{,}17)}{100}}$$

Sabemos que en una distribución de probabilidad normal el rango entre 1,96 (Z1-α/2) veces por debajo y por arriba de la media se encontraban el 95% de los valores. Usemos esta propiedad para estimar el rango de valores entre los que tengo una confianza del 95% de que se encuentre la proporción poblacional. Este cálculo corresponde a:

$$p \pm Z_{1-\propto/2} \times error\ estándar \gg 0{,}17 \pm 1{,}96 \times 0{,}0375$$

Intervalo de confianza al 95% (proporción poblacional p): 0,0965 a 0,2435 (9,65 a 24,35%).

Habitualmente, este rango se conoce como intervalo de confianza del 95% (IC 95). Se puede interpretar diciendo que tenemos un 95% de confianza de que la proporción poblacional se encuentre entre 9,65 y 24,35% (con un error menor del 5%). Aunque esta es la interpretación más sencilla, la interpretación real corresponde a que, si hiciéramos 100 estudios con un tamaño muestral similar al nuestro, el verdadero parámetro poblacional estaría incluido en 95 de los 100 intervalos de confianza estimados.

Un IC 95% es un rango de valores que tiene una probabilidad del 95% de contener el valor real de la población, en el sentido de que, si se extrajera un número infinito de muestras para estimar el valor de interés, el 95% de sus IC del 95% contendrían el verdadero valor de la población

PROCEDIMIENTO DE CÁLCULO DEL INTERVALO DE CONFIANZA

Al igual que hemos deducido la fórmula del error estándar para la estimación de proporciones, el lector interesado puede consultar en textos especializados la deducción de las fórmulas de error estándar de otros parámetros. Lo importante es entender que ese error estándar es el factor de dispersión del conjunto teórico de estimaciones de dicho parámetro en sucesivas muestras. En el capítulo 6.6. correspondiente al cálculo del tamaño muestral se muestran las fórmulas para el cálculo del error estándar de los parámetros más utilizados en Epidemiología.

La desviación estándar describe la dispersión de los valores de la muestra alrededor de su media. El error estándar es un parámetro de la distribución de las distintas muestras que refleja la dispersión de las distintas estimaciones muestrales y su margen de error

El procedimiento siempre es el mismo: 1) calculamos la medida descriptiva o la medida de frecuencia, riesgo o impacto de nuestra muestra; 2) estimamos a partir de dichos datos el error estándar; y 3) usamos dicho error estándar y el valor Z correspondiente (para un intervalo de confianza del 95%: 1,96) para estimar el intervalo de confianza.

Cuando manejemos variables continuas, definidas por sus medias y desviaciones típicas, no debemos confundir la desviación típica de nuestra muestra, que describe nuestros datos, con el error estándar, que nos permite estimar el rango de error en la estimación de la media poblacional.

Aunque el cálculo para proporciones, medias, diferencias de proporciones o diferencias de medias resulta sencillo, no recomendamos hacer dichos cálculos de forma manual. Para ello tenemos calculadoras epidemiológicas que realizan los cálculos fácilmente o, si tenemos los datos tabulados, podemos recurrir a paquetes estadísticos.

Veamos un ejemplo práctico:

Calculemos el intervalo de confianza para una media, utilizando para ello el programa EPIDAT 4.2, *software* gratuito que puede ser descargado libremente (https://www.sergas.es/Saude-publica/EPIDAT-4-2?idioma=es) y que no requiere instalación (el fichero descargado se puede ejecutar). En el **Anexo 1** de este capítulo pueden consultar los pasos para resolver este ejercicio.

Supongamos que tenemos una muestra de 10 recién nacidos y que conocemos sus pesos al nacimiento (3388, 2703, 2735, 2980, 3865, 2967, 3014, 3807, 3504 y 3975). Una vez calculada su media (3293,8 g) y su desviación estándar (477,4 g), podríamos calcular manualmente el error estándar para obtener los límites del intervalo de confianza.

Lo más sencillo es utilizar el programa informático, que nos muestra para el intervalo de confianza del 95% un límite inferior de 2952,29 g y un límite superior de 3635,31 g.

Veamos otro ejemplo:

Veamos ahora el cálculo del intervalo de confianza de una proporción, utilizando el programa EPIDAT. En el **Anexo 2** de este capítulo puede consultar los pasos para resolver este ejercicio.

Supongamos que hemos estudiado una muestra de 100 niños escolares y hemos detectado una proporción de obesidad del 15%. Estamos interesados en obtener una estimación de la prevalencia de obesidad en la población con un nivel de confianza del 95%.

Si introducimos los datos, el programa nos informa de que el intervalo de confianza del 95% se encuentra entre los límites 8,64% y 23,53%.

Veamos un tercer ejercicio:

Utilizaremos nuevamente EPIDAT para calcular el intervalo de confianza de una diferencia de proporciones. En el **Anexo 3** de este capítulo puede consultar los pasos para resolver este ejercicio.

Supongamos que comparamos la muestra de escolares del ejercicio anterior con otra de 120 escolares de otra ciudad y obtenemos una proporción de obesos del 20% (24 escolares). Estamos interesados en estimar la diferencia de la prevalencia de obesidad entre las dos poblaciones de las que provienen las muestras.

Si introducimos los datos, el programa nos informa que el intervalo de confianza del 95% para la diferencia de proporciones es de -0,15 a 0,05 o, si lo preferimos, de -15% a 5%. Más adelante comentaremos el significado de este intervalo.

INTERPRETACIÓN DE LOS INTERVALOS DE CONFIANZA

Al momento de interpretar un intervalo de confianza nos interesan varias cuestiones.

Primero, debemos observar la amplitud del intervalo. Un intervalo amplio indica que la estimación es imprecisa; uno estrecho indica una estimación precisa. La amplitud del intervalo de confianza depende del tamaño del error estándar que, a su vez, depende del tamaño de la muestra y, cuando se trata de una variable continua, de la variabilidad de los datos. Por lo tanto, los estudios pequeños con datos variables dan intervalos de confianza más amplios que los estudios más grandes con datos menos variables.

El intervalo de confianza informa sobre la precisión de la estimación del tamaño del efecto y es útil para realizar algunos contrastes de hipótesis

Segundo, debemos reconocer qué implicaciones clínicas pueden derivarse del intervalo. Los límites superiores e inferiores permiten evaluar si los resultados son clínicamente importantes. Por ejemplo, si 27 de 64 preescolares presentan enfermedad periodontal, la prevalencia es de 42,2% (IC 95: 30 a 55%). Se trata de un intervalo de confianza bastante amplio, lo que indica poca precisión. Sin embargo, el límite inferior de este intervalo de confianza indica que un porcentaje sustancial de estos niños (al menos el 30%) tiene enfermedad periodontal, por lo que se justificaría realizar estudios para obtener una estimación más precisa y valorar la necesidad de una intervención sanitaria.

Y, por último, debemos comprobar si el intervalo incluye algún valor de interés especial. Podemos comprobar si un valor hipotético del parámetro poblacional se encuentra dentro del intervalo de confianza; por ejemplo, tras una intervención nutricional queremos saber la prevalencia de obesidad en adolescentes de nuestro medio y compararla con la que se ha descrito en nuestro país en publicaciones previas del 30%. Seleccionamos una muestra de 400 adolescentes que han participado en la intervención encontrando una prevalencia del 25%, con un intervalo de confianza del 95% entre 20,8 a 29,5%. Como en el intervalo de confianza no está incluido el valor de referencia, podemos decir que nuestros adolescentes tienen menor obesidad que los de la población nacional, con una confianza de al menos el 95%.

Esta propiedad puede utilizarse también en el contraste de hipótesis, como se trata en el capítulo 6.5. Como vimos en el ejemplo anterior sobre las diferencias de proporciones de obesidad en dos muestras, el intervalo de confianza de la diferencia de proporciones incluía el valor cero, por lo que ya sabemos, sin conocer el valor de p, que la diferencia entre las dos muestras no es estadísticamente significativa.

Anexo 1
1. Lanzar EPIDAT
2. Seleccionar las opciones de menú Módulos → Inferencia sobre parámetros → Una población → Media
3. Introducir los datos en la ventana emergente: ■ Media: 3293,8 ■ Desviación estándar: 477,4 ■ Tamaño de muestra: 10 ■ Fijar el nivel de confianza: por defecto, 95%
4. Seleccionar la casilla “Intervalo de confianza”. Podemos deseleccionar las casillas correspondientes al contraste de hipótesis (no se necesitan para este ejercicio)
5. Pulsar el botón “Calcular”

Anexo 2
1. Lanzar EPIDAT
2. Seleccionar las opciones de menú Módulos → Inferencia sobre parámetros → Una población → Proporción
3. Introducir los datos en la ventana emergente: ■ Número de casos: 15 ■ Tamaño de muestra: 100 ■ Fijar el nivel de confianza: por defecto, 95%
4. Seleccionar la casilla “Intervalo de confianza”. Podemos deseleccionar las casillas correspondientes al contraste de hipótesis (no se necesitan para este ejercicio)
5. Pulsar el botón “Calcular”

Anexo 3
1. Lanzar EPIDAT
2. Seleccionar las opciones de menú Módulos → Inferencia sobre parámetros → Dos poblaciones → Proporciones independientes
3. Introducir los datos en la ventana emergente: ■ Muestra 1: número de casos 15, tamaño de muestra 100 ■ Muestra 2: número de casos 24, tamaño de muestra 120 ■ Fijar el nivel de confianza: por defecto, 95%
4. Seleccionar la casilla “Intervalo de confianza”. Podemos deseleccionar las casillas correspondientes al contraste de hipótesis (no se necesitan para este ejercicio)
5. Pulsar el botón “Calcular”

Ver respuestas

PREGUNTAS DE AUTOEVALUACIÓN

1. **El intervalo de confianza para una proporción puede aproximarse mediante una distribución normal si se cumple la regla empírica de que el producto de *n* (tamaño muestral) por *p* (proporción) y por el complementario de *p* (1-p) es mayor de:**
   a) 1
   b) 4
   c) 3
   d) 5
2. **Si aumenta el nivel de confianza, el intervalo de confianza:**
   a) Aumenta.
   b) Disminuye.
   c) Permanece igual.
   d) Depende del tamaño muestral.
3. **De una muestra de 3500 recién nacidos la media de la talla fue de 49,9 cm y su desviación típica de 1,9 cm. Se calculó un IC 95% de 49,8 cm a 50,0 cm. Podemos inferir que:**
   a) Si repetimos el estudio con un número elevado de muestras diferentes, el valor medio de la talla se encontrará entre 49,8cm y 50 cm en el 95% de las muestras.
   b) Nos informa sobre la variabilidad que presentan los valores medios de la talla en las diferentes muestras que podríamos extraer de la misma población.
   c) Nos indica el grado de precisión con el que se estimó el valor poblacional.
   d) Si hubiésemos utilizado un nivel más alto de confianza podríamos haber obtenido una estimación más precisa.

BIBLIOGRAFÍA

Beyene J, Moineddin R. Methods for confidence interval estimation of a ratio parameter with application to location quotients. BMC Med Res Methodol. 2005;5:32.

Daly LE. Confidence Limits Made Easy: Interval Estimation Using a Substitution Method. American Journal of Epidemiology. 1998;147:783-90.

De Maris A, Selman SH. Converting Data into Evidence. New York: Springer New York; 2013.

Douglas G. Altman, David Machin, TN. Bryant, Michael J. Gardner-Statistics with Confidence-BMJ Books; 2000.

Everitt BS. Medical Statistics from A to Z: A Guide for Clinicians and Medical Students. 3.ª ed. Cambridge University Press; 2021.

Michael Harris, Gordon Taylor - Medical Statistics Made Easy. Scion Publishing; 2020.

Ochoa Sangrador C, Ortega Páez E, Molina Arias M. Inferencia estadística: estimación por intervalos. Evid Pediatr. 2019;15:40.

Ochoa Sangrador C, Molina Arias M, Ortega Páez E. Inferencia estadística: probabilidad, variables aleatorias y distribuciones de probabilidad. Evid Pediatr. 2019;15:27.

Ochoa Sangrador C, Molina Arias M. Estadística descriptiva. Evid Pediatr. 2018;14:43.

Smith DJ, Samuel S. Basic Statistical Techniques for Medical and Other Professionals: A Course in Statistics to Assist in Interpreting Numerical Data. 1.ª ed. New York: Productivity Press; 2021.