El nivel de confianza del intervalo de confianza para la proporción

Antonio Francisco Roldán López de Hierro
Carmen Batanero

SUMA+ núm. 93
pp. 29-37
Artículo recibido en *Suma* en septiembre de 2018 y aceptado en septiembre de 2019

A menudo observamos que el alumnado calcula el intervalo de confianza para la proporción aplicando una y otra vez la misma fórmula, lo que le lleva a una interpretación errónea del resultado final. En el presente artículo desarrollamos una propuesta para determinar el intervalo de confianza para la proporción basada en la aproximación de Wilson. Los cálculos asociados están al alcance del alumnado de los cursos finales de la ESO y sirven para que el alumnado aprenda a interpretar correctamente tanto el significado del intervalo obtenido como el nivel de confianza.

Palabras clave: Intervalo de confianza, Nivel de confianza, Proporción, Interpretación.

The confidence level of the proportion confidence interval // Students often work out automatically the confidence interval for proportions by applying the same formula over and over again, which leads to a wrong interpretation of the final result. In this article we develop a proposal to determine the confidence interval for proportions based on Wilson's approach. Computations associated to this approach are within the reach of the Secondary School upper grades students and they help them learn to correctly interpret both the meaning of the obtained interval and the involved confidence level.

Keywords: Confidence interval, Confidence coefficient, Proportion, Interpretation.

Uno de los problemas más interesantes (y con un mayor número de aplicaciones) al que nos podemos enfrentar es el de estimar la proporción de individuos de una población que poseen una cierta propiedad. Veamos algunos ejemplos en diversos ámbitos del conocimiento, donde a veces se utiliza el término *porcentaje*, el cual es más comprensible para el alumnado:

— En el ámbito de la política, ¿cuál es el porcentaje de electores que finalmente votará en las próximas elecciones? Y, más en concreto, ¿qué porcentaje de votantes votará finalmente a un partido político determinado?

— En el campo de la medicina, preguntas como ¿qué porcentaje de la población padece diabetes?, o ¿qué porcentaje de la población pasará la gripe el próximo invierno? Conocer estos porcentajes permite a las autoridades sanitarias prever qué actuaciones deben llevarse a cabo de manera que el servicio a la población sea el más adecuado posible a la vez que no se desperdicien recursos.

- En el terreno comercial, es muy importante conocer el porcentaje de visitantes de una página web que finalmente termina comprando el producto que se ofrece, y cuáles son los contenidos de la página web que llevan a un mayor porcentaje de compradores.
- En un contexto biológico, podría ser interesante conocer el porcentaje de descendientes de una cierta población que sobrevivirá al primer año de vida. Por ejemplo, si tenemos una piscifactoría, es razonable que estemos informados del número de peces que cabe esperar cuando se reproduzca una determinada generación.

Los anteriores son solo unos pocos ejemplos, accesibles al alumnado, que tratan de justificar la importancia de estimar una cierta proporción que, en general, debido al enorme tamaño que suele poseer una población, es desconocida (por ejemplo, algunas aplicaciones en el ámbito de la medicina pueden encontrarse en Erdogan y Gülhan, 2016). Como es lógico, su mayor o menor conocimiento determina nuestras vidas ya que es una de las herramientas fundamentales a la hora de tomar decisiones políticas o empresariales.

Dentro del currículo de las Matemáticas en el sistema educativo español, el tema de la Inferencia Estadística está únicamente incluido en la asignatura de Matemáticas Aplicadas a las Ciencias Sociales II de 2.º de Bachillerato de Ciencias Sociales. No obstante, puesto que existe una corriente de alumnado de la rama científica y tecnológica que se interesa por estos contenidos, especialmente por la posibilidad de presentarse y obtener una buena calificación en el correspondiente examen de las Pruebas de Acceso a la Universidad, los centros educativos ofrecen al alumnado de 2.º de Bachillerato una asignatura optativa de Estadística y Probabilidad.

El escaso tiempo disponible lleva a que tanto el proceso de deducción del intervalo de confianza para la proporción como su aplicación en un contexto práctico sean procedimientos absolutamente mecánicos, sin tener en cuenta métodos alternativos. Puesto que este método se basa en la aproximación normal, se introduce un error que hace que el coeficiente de confianza real del intervalo construido por el procedimiento usual sea menor que el deseado.

Esta cuestión no se clarifica, ni siquiera en los actuales planes de estudio de las asignaturas universitarias de la mayoría de grados, donde se imparten estos contenidos: se sigue enseñando la fórmula para el cálculo del intervalo de confianza en la metodología de Neyman-Pearson para la proporción, aunque existen diferentes aproximaciones tanto en el método frecuencial como en otras escuelas de estadística. Tampoco se suele discutir el significado del intervalo de confianza, el cual se suele interpretar incorrectamente.

En este contexto, en el presente artículo nos planteamos varios objetivos:

- Reflexionar sobre el método de construcción más frecuente del intervalo de confianza para la proporción, describiendo sus virtudes y sus debilidades, mostrando ejemplos sencillos para analizar el error introducido por la aproximación normal.
- Proponer un nuevo método para determinar un intervalo de confianza para la proporción de mayor fiabilidad. Aunque la técnica de deducción posiblemente se deba dejar para el alumnado universitario, se podría trabajar en bachillerato si se proporciona una hoja de cálculo con el correspondiente algoritmo programado.
- Sugerir dedicar menor tiempo al cálculo y mayor a la interpretación del proceso y de los resultados para desterrar entre el alumnado y el profesorado errores comunes de interpretación del intervalo de confianza.

Dificultades en la comprensión del intervalo de confianza

Nuestra experiencia docente nos ha llevado a observar que son muchos los alumnos que, tras desarrollar en papel el proceso mecánico de obtener el intervalo de confianza para una proporción desconocida p al 95% de confianza, afirman que la probabilidad de que el verdadero valor de p pertenezca a dicho intervalo es del 95%. Esto es un error de interpretación descrito en la investigación didáctica, por ejemplo, en Olivo (2008), quien sugiere que esta es una interpretación bayesiana (sería aceptable en esta metodología, pero no en la frecuencial).

La interpretación correcta es que la confianza está depositada en el método, no en el intervalo, de manera que confiamos en que en el 95% de las ocasiones, al desarrollar el procedimiento con muestras aleatorias simples e independientes, y cuando se repite muchas veces el proceso, el verdadero valor de la proporción desconocida p esté contenido en el intervalo determinado. Por tanto los límites del intervalo son variables y el intervalo concreto que se construye en cada caso puede, o no, contener al verdadero valor de la proporción poblacional.

Otras dificultades que aparecen a menudo provienen de la incomprensión de las relaciones entre el ancho del intervalo, el tamaño de la muestra y la proporción muestral, e incluso de la confusión entre la precisión y la confianza, pues se supone que un intervalo más estrecho implica siempre mayor confianza (Behar, 2001).

Por otro lado, al trabajar con el intervalo de confianza y, en general, con la inferencia, debemos tener en cuenta tres proporciones que los estudiantes confunden entre sí (Schuyten, 2001):

- La proporción en la población, que es un parámetro desconocido.
- La proporción muestral en la muestra que hemos tomado para calcular el intervalo y es constante.
- La proporción muestral, considerada como variable aleatoria en las infinitas muestras posibles que pueden formarse de la misma población, tamaño y condiciones.

El tener que trabajar con esto a tres niveles simultáneamente causa una gran dificultad a los estudiantes.

Análisis del método de construcción del intervalo de confianza para la proporción

Además de los posibles errores de interpretación, debe mejorarse el método de construcción. El procedimiento que describimos a continuación es el habitual para el cálculo del intervalo de confianza (véase, por ejemplo, Canavos, 1988). Supongamos que estamos interesados en conocer la proporción de individuos de una población que cumplen una cierta condición. La proporción debe tomar un valor concreto dentro del intervalo real, pero dicho valor es desconocido.

Para estimarla, tomamos una muestra aleatoria simple de tamaño n dentro de la población, y observamos la variable aleatoria X que mide el número de individuos de la muestra que verifica dicha propiedad. Como es conocido, la variable aleatoria X sigue una distribución binomial de parámetros n y p, lo que denotamos por $X \mapsto B(n,p)$. Una sencilla aplicación del teorema central del límite, permite aproximar, cuando n es suficientemente grande, la distribución discreta binomial $B(\text{n},p)$ mediante la distribución (continua) normal de la misma media y desviación típica. En este caso, podemos plantear la posibilidad de aproximar la variable X mediante una nueva variable $\widetilde{X}$ con distribución normal $N(np,npq)$, donde utilizamos la notación $q=1-p$ para la proporción de individuos de la población que no cumple la propiedad estudiada. Obsérvese que el segundo argumento de la distribución normal es la varianza. De esta forma, hemos aproximado

$$X \simeq \widetilde{X} \mapsto N(np,npq)\,.$$

Al dividir estas variables entre n, damos paso al estudio de variables aleatorias que miden la proporción de individuos de la muestra que cumple la cualidad estudiada. En tal caso,

$$p = \frac{X}{n} \simeq \frac{\widetilde{X}}{n} \mapsto N\left(p, \frac{pq}{n}\right).$$

Tipificando, llegamos a la conclusión de que, aproximadamente, cuando el tamaño de la muestra es suficientemente grande,

$$\frac{P-p}{\sqrt{\frac{pq}{n}}} \mapsto N(0,1)\,. \qquad [1]$$

A partir de esta distribución en el muestreo se suele deducir el intervalo de confianza para la proporción de la siguiente forma. Fijemos un nivel de confianza $1-\alpha$, donde α representa el nivel de significación. Usualmente suele tomarse un 95% de confianza, aunque también es razonable tomar otros valores entre 0,9 y 0,99. Se trata de encontrar el valor Z tal que varíe en el intervalo central del $(1-\alpha)\%$ de la distribución normal $N(0,1)$, es decir:

$$P\left(-z_{1-\alpha/2} < Z < z_{1-\alpha/2}\right) = 1-\alpha\,.$$

La expresión anterior es válida para cualquier variable normal estándar Z. En consecuencia, si como variable Z elegimos la aproximación descrita en [1], obtenemos la igualdad:

$$P\left(-z_{1-\alpha/2} < \frac{P-p}{\sqrt{\frac{p\,(1-p)}{n}}} < z_{1-\alpha/2}\right) = 1-\alpha\,.$$

Puesto que p es desconocido, se sustituye su valor en el denominador por la proporción muestral, que es su estimador puntual (de máxima verosimilitud). De esta forma, se llega a:

$$P\left(-z_{1-\alpha/2} < \frac{P-p}{\sqrt{\frac{P\,(1-P)}{n}}} < z_{1-\alpha/2}\right) = 1-\alpha\,. \qquad [2]$$

Despejando el parámetro p de una forma estándar deducimos la expresión usual que se utiliza para determinar el intervalo de confianza, al nivel de confianza $1-\alpha$, para la proporción poblacional p que, en forma abreviada, es:

$$IC(p) = \left[P \pm z_{1-\alpha/2}\sqrt{\frac{P\,(1-P)}{n}}\right]. \qquad [3]$$

La principal bondad de esta expresión es que es muy sencilla de utilizar en la práctica y, además, produce buenos resultados. Su principal debilidad es que ha resultado obtenida a través de un proceso en el que, sin aparente justificación, se sustituye el verdadero valor de p por su estimación P. Además, este reemplazo solo se hace en el denominador de la expresión considerada, con la única intención de simplificar los cálculos posteriores. En cierta forma, se transforma un problema no lineal en un problema lineal. Por otro lado, la expresión resulta de una aproximación de una variable discreta (binomial) mediante una variable continua (normal) como consecuencia del teorema central del límite. De esta forma, debemos ser conscientes de que, con todo lo satisfactoria que pueda ser, no se trata más que de una aproximación.

¿Garantiza el método el nivel de confianza elegido?

Para analizar el error de aproximación del método descrito utilizamos un ejercicio que fue propuesto en el examen titular de las pruebas de acceso y admisión a la universidad de la convocatoria de septiembre de 2017 en Andalucía (un análisis más detallado de ejercicios similares puede consultarse López-Martín, Batanero, Díaz-Batanero y Gea, 2016).

> EJERCICIO 1. En una muestra, elegida al azar, de 100 estudiantes de una universidad, se ha observado que 25 desayunan en la cafetería del campus.
>
> a) (1.25 puntos) Determine, con un nivel de confianza del 95 %, un intervalo de confianza para estimar la proporción de estudiantes de esa universidad que desayunan en la cafetería. […]

Una posible resolución de este ejercicio es la siguiente. Dado que la muestra está formada por $n=100$ estudiantes (consideramos que este es un tamaño suficiente para poder desarrollar la aproximación normal por binomial) y resultó que 25 de ellos desayunan en la cafetería, por lo que la proporción muestral de estudiantes que desayunan en la cafetería es $P=25/100=0{,}25$. Dado que el nivel de confianza es $1-\alpha=0{,}95=95\%$, resulta que $\alpha=0{,}05$ y $1-\alpha/2=0{,}975$. De esta forma, el nivel crítico al 95% de confianza es $z_{0{,}975}\simeq 1{,}96$. Así, el intervalo de confianza al 95% de confianza para la proporción de alumnado de esa universidad que desayuna en la cafetería es:

$$IC(p)=\left[P\pm z_{1-\alpha/2}\sqrt{\frac{P\,(1-P)}{n}}\right]=$$
$$=\left[0,25\pm 1,96\sqrt{\frac{0,25\,(1-0,25)}{n}}\right]\simeq[0,25\pm 0,085]=$$
$$=\left[0,25-0,085;\,0,25+0,085\right]=\left[0,165;\,0,035\right].$$

Por consiguiente, el intervalo de confianza para la proporción p de alumnado de esa universidad que desayuna en la cafetería, asociado a la muestra considerada y al 95% de confianza, es [0,165;0,335].

Una vez construido el intervalo nos preguntamos, ¿es verdad que se conserva el valor del coeficiente de confianza, es decir, podemos confiar en que el 95% de los intervalos que así se construyen contienen al verdadero valor desconocido? Hemos desarrollado una simulación computacional para estudiar este extremo siguiendo los pasos que detallamos a continuación.

1. Fijamos el nivel de confianza $1-\alpha$ al que desarrollaremos el experimento (por ejemplo, al 95%) y el tamaño n de las muestras que consideraremos (por ejemplo, $n=30$).
2. Generamos un número al azar en el intervalo $(0,1)$ para simular una posible proporción p.
3. Generamos, con ayuda del ordenador, un valor aleatorio de una distribución binomial $B(\mathrm{n},p)$, que sirve como indicador del número de éxitos que se podrían obtener en una muestra aleatoria de tamaño n.
4. Dividimos dicho valor entre n obteniendo un posible valor muestral para la proporción de individuos de la muestra que poseen la característica estudiada (conoceríamos el valor muestral de la variable P).
5. Aplicamos la fórmula [3] y obtenemos el intervalo de confianza para p asociado a la posible muestra generada.
6. Determinamos si el verdadero valor de p pertenece al intervalo de confianza calculado.
7. Repetimos los pasos 2 a 6 un número muy grande de veces (por ejemplo, 10 000 o 1 000 000), de manera que las muestras consideradas sean independientes, y contamos cuántas veces el verdadero valor del parámetro pertenecía al intervalo calculado.

Resumimos en la tabla 1 los resultados que obtuvimos al repetir el experimento, tantas veces como se indica en ella, con muestras de tamaño 30 al 95% de confianza.

Método usual (basado en la aproximación normal)		
Tamaño muestral	$n=30$	
Número de iteraciones	10 000	1 000 000
Número de iteraciones en que p pertenece al intervalo de confianza	8 719	875 354
Proporción de éxitos de los intervalos de confianza calculados	87,19 %	87,54 %

Tabla 1. Porcentaje de intervalos de confianza para la proporción que contienen la proporción poblacional en la simulación de la fórmula tradicional

Los resultados obtenidos en la tabla 1 muestran, experimentalmente, que la fórmula del intervalo de confianza que utilizamos en clase no llega a conseguir el resultado esperado de un 95% de intervalos que contienen al verdadero valor de la proporción. De hecho, está muy lejos de dicho valor. Realmente, habría que considerar tamaños muestrales alrededor de 1 000 para que se obtengan resultados cercanos al valor del nivel de confianza, lo que hace, en muchos casos, imposible la toma de la muestra (en la práctica, suele ser muy costoso, tanto en dinero como en tiempo, conseguir 1 000 datos de la variable considerada).

Un nuevo método para el cálculo del intervalo de confianza para la proporción

En esta sección presentamos una nueva metodología para determinar un intervalo de confianza para la proporción p introducida por Wilson (1927), cuya idea esencial consiste en despejar el parámetro p de la expresión [2] utilizando una ecuación de segundo grado en vez de realizar una aproximación de primer grado. Es por ello que, aunque las cuentas no sean tan triviales como en el caso anterior, sí están al alcance del alumnado del bachillerato (e incluso de los últimos cursos de la ESO). Despejamos el parámetro p de la siguiente ecuación (sabiendo que los elementos que intervienen son positivos) utilizando la notación z para el valor $z_{1-\alpha/2}$ o el valor $-z_{1-\alpha/2}$:

$$\frac{P-p}{\sqrt{\frac{p(1-p)}{n}}}=z \quad \Rightarrow \quad (P-p)^2=z^2\frac{p(1-p)}{n} \quad \Leftrightarrow$$
$$\Leftrightarrow \quad P^2-2pP+p^2=\frac{z^2}{n}p-\frac{z^2}{n}p^2 \quad \Leftrightarrow$$
$$\Leftrightarrow \quad \left(n+z^2\right)p^2-\left(2nP+z^2\right)p+nP^2=0\,.$$

Dado que el discriminante de esta ecuación es estrictamente positivo, la ecuación de segundo grado posee dos soluciones reales, que son las siguientes:

$$p=\frac{\left(2nP+z^2\right)\pm\sqrt{z^2\left(z^2+4nP(1-P)\right)}}{2\left(n+z^2\right)}=$$
$$=\frac{2nP+z^2\pm z\sqrt{z^2+4nP(1-P)}}{2\left(n+z^2\right)}\,.$$

Estas expresiones determinan, precisamente, los extremos del intervalo de confianza propuesto por el método de Wilson, el cual, por comodidad, se escribe como sigue:

$$IC_W(p)=$$
$$\left[\frac{2nP+z^2_{1-\alpha/2}\pm z^2_{1-\alpha/2}\sqrt{z^2_{1-\alpha/2}+4nP(1-P)}}{2\left(n+z^2_{1-\alpha/2}\right)}\right]. \quad [4]$$

ALGUNAS PROPIEDADES

No es difícil demostrar que los dos extremos del intervalo de confianza propuesto por Wilson para p están contenidos dentro del intervalo $[0,1]$ (lo que, por cierto, puede no ocurrir con la aproximación usual). Escrito en la forma [4] podemos observar que se trata de un intervalo centrado en el punto:

$$\frac{2nP+z^2_{1-\alpha/2}}{2\left(n+z^2_{1-\alpha/2}\right)}=\frac{\omega_1}{\omega_1+\omega_2}P+\frac{\omega_1}{\omega_1+\omega_2}\cdot\frac{1}{2}$$

(donde $\omega_1=n$ y $\omega_2=z^2_{1-\alpha/2}$), que puede interpretarse como una media ponderada entre la proporción muestral P y el valor ½. De esta forma, este método *traslada* hacia el centro del intervalo $[0,1]$ la estimación puntual que puede hacerse de la proporción poblacional, que ya no sería la proporción muestral sino un valor intermedio entre esta y el centro del intervalo.

Por otro lado, el radio del intervalo de confianza, es decir, la cantidad que se suma y se resta al centro del intervalo, si bien posee una expresión más complicada, hace intervenir los mismos factores que afectaban al radio del intervalo de confianza calculado por el método usual.

¿Conserva el método de Wilson para el cálculo del intervalo de confianza para la proporción el coeficiente de confianza?

Como hemos puesto de manifiesto, la principal diferencia entre el método usual y el método de Wilson es que el primero realiza una aproximación normal, y produce un intervalo muy sencillo de interpretar y de calcular, mientras que el segundo resuelve directamente la ecuación de segundo grado, lo que conduce a un intervalo un poco más complicado. Esta dificultad puede superarse si se proporciona al estudiante un programa de cálculo para su determinación pues, cuando se emplean métodos informáticos (como ocurre ya hoy en día), no es un gasto computacional significativamente superior la determinación de un intervalo de confianza por el método de Wilson.

Como se ha comentado anteriormente, la interpretación del nivel de confianza es el porcentaje de intervalos que, teóricamente, consiguen contener al verdadero valor del parámetro a estudiar (cuando se toman muchas muestras aleatorias independientes). Para desarrollar una comparativa a este respecto entre las metodologías usual y de Wilson, detallamos en la tabla 2 los resultados que se habrían obtenido exactamente con las mismas muestras que dieron lugar a la tabla 1.

Método de Wilson		
Tamaño muestral	$n=30$	
Número de iteraciones	10 000	1 000 000
Número de iteraciones en que *p* pertenece al intervalo de confianza	9 488	952 323
Proporción de éxitos de los intervalos de confianza calculados	94,88 %	95,23 %

Tabla 2. Porcentaje de intervalos de confianza para la proporción que contienen la proporción poblacional en la simulación de la fórmula de Wilson

Los resultados expuestos en la tabla 2 ponen de manifiesto que el método de Wilson alcanza resultados muy cercanos al nivel de confianza incluso con muestras de pequeño tamaño. De hecho, este porcentaje no varía significativamente independientemente de los parámetros que intervienen en el experimento. Es por ello que podemos afirmar que el método de Wilson es más fiable que el método usual o, al menos, proporciona resultados más cercanos al nivel de confianza que se fija desde el principio.

Implicaciones para la enseñanza

El principal objetivo del presente artículo ha sido el de abrir las mentes de los usuarios y profesores de Inferencia Estadística en la etapa secundaria y universitaria. A menudo, tendemos a pensar que el único método que nos han explicado en clase es el único procedimiento admisible, y que cuando existen otros métodos, estos no están al alcance de nuestros conocimientos matemáticos. Sin embargo, cuando el alumnado conoce la existencia de varios métodos para un mismo propósito, es capaz de reflexionar sobre la conveniencia de cada uno de ellos, identificando sus bondades y sus inconvenientes.

En el presente artículo hemos demostrado que técnicas estadísticas y algebraicas muy sencillas permiten construir una nueva metodología para el cálculo de intervalos de confianza que proporcionan una mejor solución, pues conservan el nivel de confianza que se fija antes de su cálculo. Aunque sea deseable utilizar métodos sencillos de cálculo, la búsqueda de simplicidad debe tener en cuenta también la solución obtenida; en nuestra exposición hemos visto la ventaja del método de Wilson respecto a la conservación del coeficiente de confianza. Además, la tecnología nos permite salvar el escollo del cálculo.

Por otro lado, deseamos contribuir a la labor del profesorado en la formación del razonamiento estadístico (Batanero, 2011) (y, en general, matemático), y desterrar ideas preconcebidas acerca de interpretaciones erróneas del significado del intervalo de confianza. Como hemos comentado, un error común al interpretar el intervalo de confianza construido en el ejercicio 1 es interpretar este intervalo diciendo que «la probabilidad de que la proporción p de estudiantes que desayunan en la universidad esté entre el 16,5 % y el 33,5 % es del 95 %». Esta interpretación no es correcta. En realidad, solo hay dos casos: toda vez que el intervalo de confianza se ha calculado (asociado a la muestra aleatoria considerada), o bien el verdadero valor p (que es desconocido) pertenece al intervalo (en tal caso, la probabilidad anterior sería 1) o bien no pertenece (en cuyo caso la probabilidad anterior sería 0). No hay más posibilidades. La confianza está depositada en el método de construcción del intervalo de confianza: cuando se calculan muchos intervalos de confianza para un mismo parámetro (en este caso, p) asociados a muestras aleatorias independientes, confiamos en que el 95 % de ellos contengan al verdadero valor del parámetro (si bien no sabemos cuáles de ellos lo contienen y cuáles no).

Nuestra propuesta hace también hincapié en la conveniencia de dedicar una parte del tiempo de la enseñanza a actividades interpretativas que permitan superar estos errores. Una ayuda para ello es la simulación con *applets* como el disponible en el servidor de Rossman y Chance (ver en la bibliografía) donde se visualizan las muestras, la distribución muestral y los intervalos de confianza para la proporción o la media, y se puede variar el coeficiente de confianza, el tamaño de muestra y el parámetro de la población (ver figura 1).

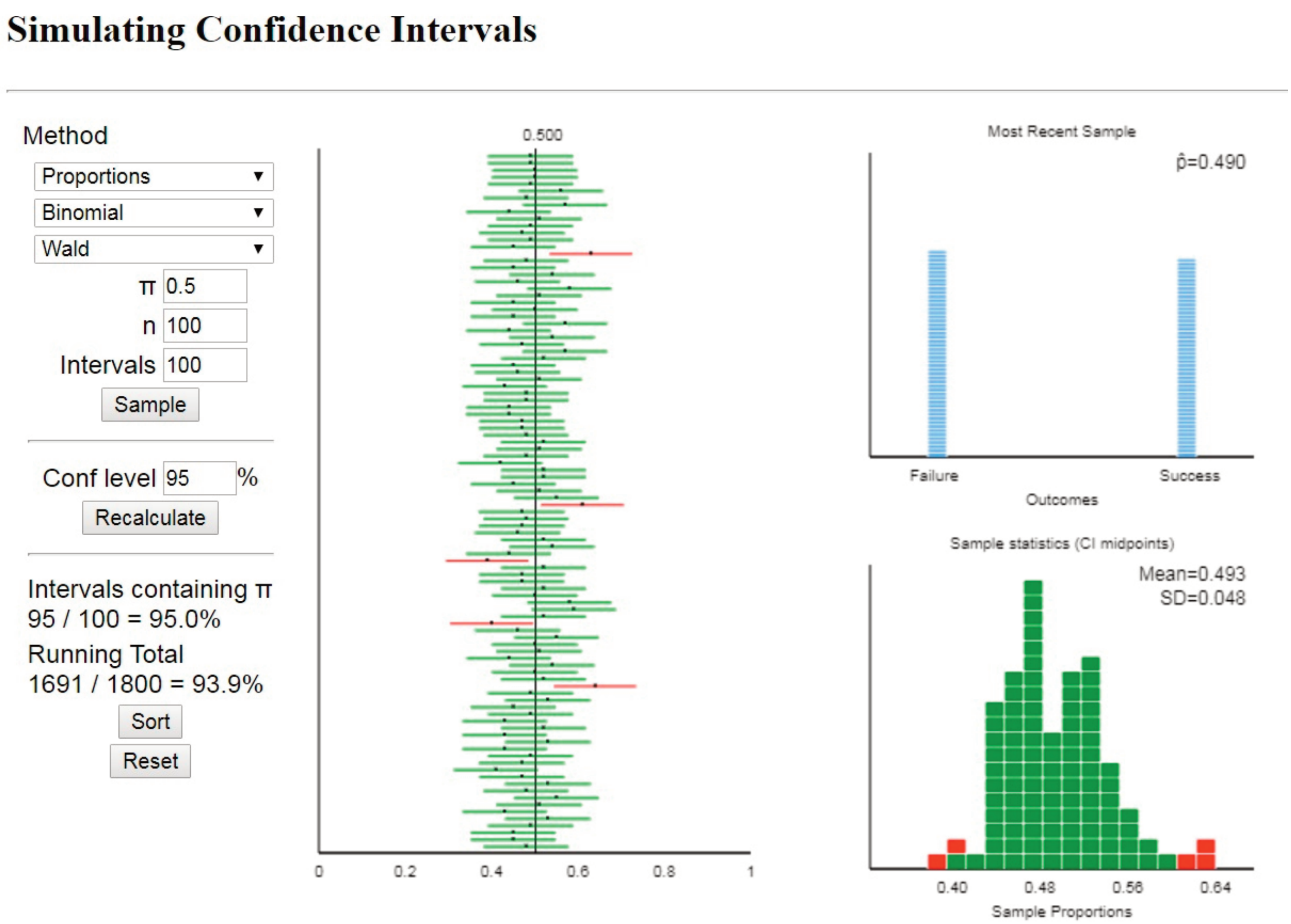

Figura 1. Simulación de intervalos de confianza

Igualmente se debe poner atención a otras dificultades al estudiar este tema (véase, a este respecto, Olivo, Batanero y Díaz, 2008), por ejemplo, que el alumnado no está nada familiarizado con la palabra *proporción*. Si el alumnado no se siente capaz de comprender este término, manifiesta una conducta de rechazo a cualquiera de los contextos de enseñanza-aprendizaje que traten de desarrollarse de cara a la asimilación de los contenidos relacionados. Sin necesidad de entrar en una visión de la proporción como la razón entre dos cantidades de una misma magnitud, el término análogo que mejor se adapta casi siempre a las capacidades del alumnado es el de *porcentaje*. Utilizando este sencillo cambio en la nomenclatura, conseguimos una mejor predisposición del alumnado para afrontar tareas relacionadas con la proporción: en cierta forma, el término porcentaje nos ayuda a comprender que existe un *todo* y que, dentro de ese todo, nos centramos en una *parte*, que evaluamos numéricamente a través de números expresados en porcentajes. La utilización de la palabra porcentaje frente al término proporción hace accesible al alumnado el planteamiento de toda una variedad de problemas que ponen de manifiesto la importancia del problema planteado inicialmente.

Esperamos que estas sugerencias interesen al profesorado y contribuyan a la mejora del uso y la enseñanza de la Inferencia Estadística.

Agradecimientos: Proyecto EDU2016-74848-P (AEI, FEDER) y Grupos FQM126 y FQM245 (Junta de Andalucía).

Referencias bibliográficas

Batanero, C. (2011), «Del análisis de datos a la inferencia: Reflexiones sobre la formación del razonamiento estadístico», *XIII Conferência Interamericana de Educação Matemática*, Belo Horizonte, Brasil.

Behar, R. (2001). *Aportaciones para la mejora del proceso de enseñanza-aprendizaje de la estadística*, Tesis Doctoral, Universidad Politécnica de Cataluña.

Canavos, G. C. (1988), *Probabilidad y Estadística*, McGraw-Hill/Interamericana de México, México.

Erdogan, S., y O. T. Gülhan (2016), «Alternative confidence interval methods used in the diagnostic accuracy studies», *Computational and Mathematical Methods in Medicine*, vol. 2016, article ID 7141050, 1-7.

López-Martín, M. M., C. Batanero, C. Díaz-Batanero y M. M. Gea (2016), «La inferencia estadística en las pruebas de acceso a la universidad en Andalucía», *Revista Paranaense de Educaçao Matemática*, n.º 5(8), 33-59.

Olivo, E. (2008), *Significados del intervalo de confianza en la enseñanza de la ingeniería en México*, Tesis Doctoral. Universidad de Granada.

Olivo, E., C. Batanero y C. Díaz (2008), «Dificultades de comprensión del intervalo de confianza en estudiantes universitarios», *Educación Matemática*, n.º 20(3), 55-82.

Rossman, A., y B. Chance, Rossman/Chance Applet Collection (Simulating Confidence Intervals), Consultada el 5 de septiembre de 2018 en <http://www.rossmanchance.com/applets/ConfSim.html>.

Schuyten, G. (1991), «Statistical thinking in psychology and education», en D.Vere-Jones (Ed.) *Proceeding of the Third International Conference on Teaching Statistics,* Otago, Nueva Zelanda: International Statistical Institute, 486-490.

Wilson, E. B. (1927), «Probable inference, the law of succession, and statistical inference», *Journal of the American Statistical Association*, n.º 22, 209-212.

Antonio Francisco Roldán López de Hierro
Universidad de Granada
<aroldan@ugr.es>

Carmen Batanero
Universidad de Granada
<batanero@ugr.es>