DESARROLLO DEL RAZONAMIENTO DE ESTUDIANTES DE BACHILLERATO SOBRE INTERVALOS ALEATORIOS

Development of the reasoning of high school students on random intervals

Martínez-Pérez, S.[a] y Sánchez, E.[a]

[a]Centro de Investigación y Estudios Avanzados del IPN

Resumen

Se presenta una actividad de aprendizaje que requiere que los estudiantes, con apoyo de software, construyan intervalos aleatorios y determinen si un valor dado se encuentra en ellos; consideramos que esto puede ser usados para introducir el estudio de los intervalos de confianza ya que se ha documentado que en el tratamiento de los intervalos de confianza existen una serie de falsas concepciones (Castro-Sotos, Vanhoof, Noortgate y Onghena, 2007). En los resultados se muestran los razonamientos de los estudiantes después de haber trabajado en la actividad.

Palabras clave: *intervalos aleatorios, nivel de confianza, variable aleatoria, probabilidad, frecuencia relativa.*

Abstract

A learning activity is presented that requires students, with software support, to construct random intervals and determine if a given value is in them; we consider that this can be used to introduce the study of confidence intervals since it has been documented that in the treatment of confidence intervals there are a series of false conceptions (Castro-Sotos, Vanhoof, Noortgate and Onghena, 2007). The results show the students' reasoning after having worked with the activity.

Keywords: *random intervals, confidence level, random variable, probability, relative frequency.*

PROBLEMÁTICA

La importancia de la inferencia estadística radica en el hecho de que es la herramienta principal de la estadística, la cual se refiere, de acuerdo con Pratt, Johnson-Wilder, Ainley y Manson (2008, p. 2) a la "identificación de patrones en forma de tendencias o parámetros estadísticos en la población" utiliza las herramientas proporcionadas por la matemática, para hacer afirmaciones sobre poblaciones a partir del análisis de una muestra, con el fin de elaborar predicciones y tomar decisiones en situaciones de incertidumbre, muchas veces presente en la vida cotidiana; la estimación de parámetros y el contraste de hipótesis son dos temas importantes en la inferencia estadística (Ben-Zvi, 2006).

Respecto al contraste de hipótesis y la estimación de parámetros, Cumming (2012) realizó una comparación entre los resultados que arrojan los dos procedimientos a un problema dado, llegó a la conclusión de que el contraste de hipótesis suele ser limitado mientras que los IC proporcionan un panorama amplio ya que dan un rango de posibles resultados.

Entender un intervalo de confianza requiere la coordinación de conceptos como población, muestra, distribución, variación, representatividad, probabilidad, azar, muestreo, conceptos que presentan de manera individual dificultades de comprensión. Esto genera falsas concepciones que han sido documentadas en algunos trabajos (Fidler, 2006; Fidler et al., 2004; Castro-Sotos et al., 2007 y Kalinovski, 2010), entre las más frecuentes Garfield, delMas, y Chance (2004) señalan: hay un 95% de probabilidad de que el intervalo de confianza incluya la media muestral, hay un 95% de probabilidad de que la media poblacional este entre los dos valores (límite superior y límite

Martínez-Pérez, S. y Sánchez, E. (2018). Desarrollo del razonamiento de estudiantes de bachillerato sobre intervalos aleatorios. En L. J. Rodríguez-Muñiz, L. Muñiz-Rodríguez, A. Aguilar-González, P. Alonso, F. J. García García y A. Bruno (Eds.), *Investigación en Educación Matemática XXII* (pp. 330-337). Gijón: SEIEM.

inferior), 95% de todos los datos están incluidos en el intervalo, un intervalo ancho significa menos confianza, un intervalo de confianza más angosto es siempre mejor (a pesar del nivel de confianza).

En Hogg y Craig (1978) se presentan a los intervalos aleatorios IA, como una forma de introducir a los IC. Los IA se construyen considerando una variable aleatoria y una distribución de probabilidad. Aunque el texto está dirigido a estudiantes de nivel superior, en el presente trabajo se considera que, con ayuda de la tecnología, en este caso el software Fathom, se puede elaborar una actividad de aprendizaje, que introduzca una red de nociones de los IA's en estudiantes de bachillerato.

ANTECEDENTES

Aunque los intervalos de confianza se han enseñado durante años en cursos introductorios de estadística, muy poca investigación ha sido realizada sobre la comprensión de ellos de los estudiantes (Belia et al., 2005; Castro-Sotos et al., 2007). Algunas de ellas se desarrollan con estudiantes de nivel universitario, por ejemplo: Olivo (2008) a través de un test busca conocer si los estudiantes construyen e interpretan de manera correcta a los IC.

En niveles menores se encuentran las investigaciones de Pfannkuch, Wild y Parsonage (2012) quienes realizaron un trabajo con estudiantes de 14 años utilizando el método Bootstrap para generar actividades que permitieran tener un acercamiento a los intervalos de confianza a través visualizaciones dinámicas desarrolladas con el software VIT, de tal forma que los estudiantes generaran una noción intuitiva sobre lo que es un intervalo de confianza y que además hicieran inferencias. Propusieron un marco teórico para el desarrollo de ideas sobre intervalo de confianza.

Por otro lado, Prodromu (2013) realizó un trabajo de investigación con estudiantes de secundaria a quienes aplicó una actividad en la cual debían calcular una proporción a través de la toma repetida de muestras de diferentes tamaños, con ayuda de un software. Para cada muestra debían estimar la proporción requerida y luego establecer un intervalo alrededor de dicha proporción. En sus conclusiones, la autora menciona que se debe tocar el tema del nivel de confianza ya que es importante.

En el presente trabajo se retoma el hecho de que se debe abordar el nivel de confianza ya que es el que influye en las predicciones o inferencias que se elaboran sobre el intervalo de confianza.

METODOLOGÍA

Experimentos de diseño

Para realizar el presente trabajo se utilizaron los experimentos de diseño (Cobb y Gravemeijer, 2008) que se desarrolla en tres fases las cuales se describen a continuación.

En la primera se desarrolló la actividad, nosotros usamos los principios de diseño para apoyar el desarrollo del razonamiento estadístico de estudiantes propuestos por Cobb y McClain (2004) para elaborarla. Dichos principios deben considerar cinco aspectos: 1) El enfoque de las ideas estadísticas centrales, en este trabajo serán población, muestra, intervalos aleatorios, frecuencia relativa, probabilidad; 2) la actividad de instrucción; 3) estructura de la actividad de clase, se debe comenzar señalando aspectos importantes, variables a considerar y la manera en que se mediaran, el tema que se va tratar, desarrollo de las actividades y por último discusión hecha de los estudiantes acerca de los resultados que obtuvieron; 4) herramientas informáticas utilizadas por los estudiantes, se elabora una simulación en el software Fathom; 5) discurso en el aula, se refiere a el lenguaje que se debe utilizar, el cual deberá contemplar las posibles sentencias que se espera que el estudiante haga sobre las ideas centrales, por ejemplo: la muestra representa a la población, la frecuencia relativa tiende a la probabilidad.

En la segunda fase se llevó a cabo la experimentación, la toma de datos fue a través de hojas de trabajo donde los estudiantes escribieron sus respuestas, para el análisis de ellas buscamos palabras o ideas que fueran comunes y las colocamos en códigos, este proceso de agrupación de respuestas se plantea en la Teoría Fundamentada propuesta por Glaser y Strauss (1967).

En la tercera fase se hizo un análisis retrospectivo, mismo que se llevó a cabo al final de cada aplicación para rediseñar la actividad y volver a aplicar.

Actividad

Se planteó la siguiente situación a los estudiantes:

De una urna con 10 bolas numeradas del 0 al 9 se sacan dos bolas, considera el ***intervalo*** *formado por los valores enteros que se encuentran entre el mínimo de los valores sacados y el máximo de ellos (considerando los extremos), ¿cuál es la probabilidad de que dicho intervalo contenga al número ocho?*

*Contener al ocho significa que el ocho se encuentra entre los valores mínimo y máximo o que es uno de ellos.

Se elaboraron preguntas orientadas a observar tres aspectos: las nociones sobre frecuencia relativa, probabilidad y ley de los grandes números; la variabilidad y la generalización.

RESULTADOS Y DISCUSIÓN

A continuación, se describen el desarrollo y los resultados de cada aplicación

Primera versión, primera aplicación (33 estudiantes de 17-18 años, cursaban Estadística y Probabilidad I)

Se presentó la situación a los estudiantes, se discutió sobre la forma en que se resolvería, se mencionó que se podría hacer una simulación física, se les pidió que escribieran los pasos en una tabla de dos columnas (plan de simulación); en la primera columna tenían que escribir los pasos de una simulación física y en la otra columna los pasos para hacer la simulación en Fathom, finalmente tenían que responder las preguntas basándose en los resultados observados en la simulación.

La actividad incluye 8 incisos divididos en dos partes, la primera parte tiene la intención de conocer si los estudiantes observan la variabilidad y la ley de los grandes números, incisos 1 al 5; en los incisos 1 y 3 se les solicita a los estudiantes que con ayuda de la simulación obtuvieran 10 y 100 intervalos cinco veces, respectivamente y que escribieran sus frecuencias relativas, en los incisos 2 y 4 se les pidió que con base en las frecuencias relativas obtenidas estimaran la probabilidad de que el intervalo contenga al número 2. La intención de aumentar el número de intervalos era que observaran que entre más intervalos la frecuencia relativa presentaba menos variación, mismo que se les cuestionó en el inciso 5.

Después de revisar las respuestas del inciso 4 se encontraron tres códigos: moda, media y máximo/mínimo. En el primero se colocaron todas las respuestas que mencionan que la probabilidad de que un intervalo contuviera al dos era el valor que más se repite; en el segundo, las respuestas de los alumnos que calcularon la media y en el último, las respuestas donde los estudiantes indicaron que la probabilidad es el valor máximo o mínimo de todos los que obtuvieron. Los ejemplos de las respuestas se muestran en la Tabla 1.

Tabla 1. Ejemplos de respuestas del inciso 4

Código	Respuesta
Moda	0.6 porque es el que se repite
Media	5 porque es el promedio
Máximo/mínimo	0.3 porque es el más chico

Para el inciso 5 se determinaron tres códigos: obviedad, desapercibido y pelotas; en el primero se colocaron las respuestas en las que se manifestaba que la diferencia consistía en la cantidad de intervalos, en el segundo código se colocaron las respuestas en las que se escribió que no había diferencia y en el tres las respuestas en las que se escribió que si hay cambio pero que se debe a que el número de bolas en la urna es la que aumenta. Los ejemplos de las respuestas se presentan en la Tabla 2.

Tabla 2. Ejemplos de respuestas del inciso 5

Código	Respuesta
Obviedad	Si, la causa es el aumento de intervalos
Desapercibido	No, no hay mucha diferencia, de hecho, son muy similares, si acaso varían por los decimales.
Pelotas	Si, porque tiene más pelotas

Los incisos 6, 7 y 8 tienen como objetivo observar si los estudiantes utilizan los resultados anteriores para generalizar; en el inciso 6 se le cuestiona al estudiante que si tuviera 1000 intervalos ¿cuántos contendrían al dos? En el inciso 7 se plantea que la probabilidad de que un intervalo contenga al 3 es del 43% y se le cuestiona al estudiante sobre el significado de esta aseveración. Finalmente, en el inciso 8 se pregunta que si de 100 intervalos 23 contienen al siete ¿qué significa?

Para el inciso 6 se determinaron 3 códigos: similitud, mitad y software. En similitud se colocaron las respuestas en las que el estudiante considera que los resultados serían similares a los incisos 2 y 4; en mitad los estudiantes respondieron que sería la mitad o 500 y en el código software los estudiantes escribieron el valor que la simulación les mostró.

En el inciso 7 se encontraron cuatro códigos: baja, repetición, valor y preciso. En el código baja se colocaron las respuestas que mencionan que la probabilidad es poca o baja; en el código repetición los estudiantes respondieron usando el mismo enunciado del inciso; en el código valor están las respuestas de los estudiantes que consideran que el 43% corresponde a la probabilidad de obtener un tres; en el código preciso se pusieron las respuestas que indican que de 100 intervalos 43 (430 de 1000) contendrían al tres.

Finalmente, para el inciso 8 se encontraron cuatro códigos: baja, repetición, valor y preciso. En el código baja se colocaron las respuestas que mencionan que la probabilidad es poca o baja; en el código repetición los estudiantes respondieron usando el mismo enunciado del inciso; en el código valor están las respuestas de los estudiantes que consideran que la probabilidad de obtener un siete en 23%; en el código preciso se pusieron las respuestas que indican que hay una probabilidad del 23% de que un intervalo contenga al siete.

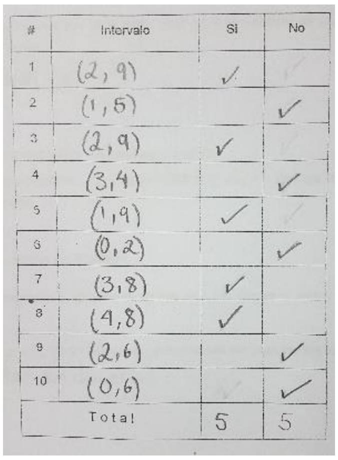
De esta aplicación concluimos que los estudiantes para entender mejor la situación debían hacer una simulación física ya que perdieron de vista que la cantidad de bolas numeradas era la misma y lo que iba aumentando era el número de intervalos obtenidos, para ellos el aumento consistía en meter más bolas a la urna; también se consideró que un polígono de frecuencias podría mostrar al estudiante la diferencia entre pocos y muchos intervalos además de que los estudiantes podrían determinar hacia que valor tienden las frecuencias relativas. También se decidió cambiar el dos por el ocho porque en algunas respuestas se observó que los estudiantes escribieron que la probabilidad de que un intervalo contenga al dos (esta probabilidad es 0.511…) es la mitad y cuando obtenían resultados se podría pensar que llevar a cabo la toda la actividad no tenía sentido porque al final confirmaban que estaban en lo correcto.

Por último, los estudiantes no estaban familiarizados con el software Fathom, lo que implicó que para la elaboración de la simulación tuvimos que hacer un paréntesis para explicar las herramientas y las funciones, cuando estuvieron haciendo la simulación la mayoría de los estudiantes presentaron dificultades tanto para el llenado del plan de simulación como para el uso de las herramientas, lo que provocó que fueran perdiendo interés y como consecuencia no pusieran atención en el desarrollo de la actividad; por lo anterior se determinó que para la siguiente versión ya no se pediría el llenado del plan de simulación.

Segunda versión, segunda aplicación (20 estudiantes de 17-18 años, cursaban Estadística y Probabilidad II)

Se plantea la misma situación al estudiante, solo que ahora se trabajará con el número ocho, la actividad se dividió en dos partes, en la primera se hace una simulación física y en la segunda se debe hacer la simulación en Fathom, en ambas se deben contestar preguntas.

Para la simulación física (incisos 1 al 5), se proporcionó a cada estudiante una bolsa que contenía las 10 bolas numeradas del 0 al 9, extrajeron 2 bolas sin reemplazo, esto último fue cuestionado, la mayoría de los estudiantes argumentó, que si se hacía el reemplazo podría salir la primera bola en la segunda extracción y entonces ya no se formaría un intervalo. La pareja de números extraída la anotaban en una tabla (Figura 1), tachaban en la casilla de SI, si el ocho estaba contenido en el intervalo o tachaban en la casilla NO en caso contrario.



#	Intervalo	SI	No
1	(2, 9)	✓	
2	(1, 5)		✓
3	(2, 9)	✓	
4	(3, 4)		✓
5	(1, 9)	✓	
6	(0, 2)		✓
7	(3, 8)	✓	
8	(4, 8)	✓	
9	(2, 6)		✓
10	(0, 6)		✓
	Total	5	5

Figura 1. Ejemplo de la tabla llenada por los estudiantes

Se solicitó a cada estudiante obtener 10 intervalos (inciso 1), después en un pizarrón se anotaron todos los resultados y se construyó una tabla de frecuencias (incisos 2 y 3), se generó la discusión, los alumnos observaron y determinaron que la probabilidad era 3 o 4 (inciso 4).

Elaboraron la simulación, pero de acuerdo con la experiencia anterior sólo hicieron una parte, la que consistía en insertar las bolas, tomar dos y determinar si se captura o no al ocho, la otra parte que era la de elaborar el polígono de frecuencias ya se entregó hecha.

Se pidió en la hoja de trabajo que obtuvieran 10 intervalos, luego 100 y luego 1000 (inciso 6), se les pidió que describieran lo que sucedía en la gráfica (inciso 8), con las respuestas de este inciso el código generado fue *gráfica* ya que se pudo respondieron que con pocos intervalos los puntos en la gráfica están muy dispersos y conforme va aumentando el número de intervalos los puntos de la gráfica (Figura 2) se "alinean" a una recta que tiene cierto valor (entre 0.3 y 0.4) y al mismo tiempo contestan que la probabilidad está entre 0.3 y 0.4.

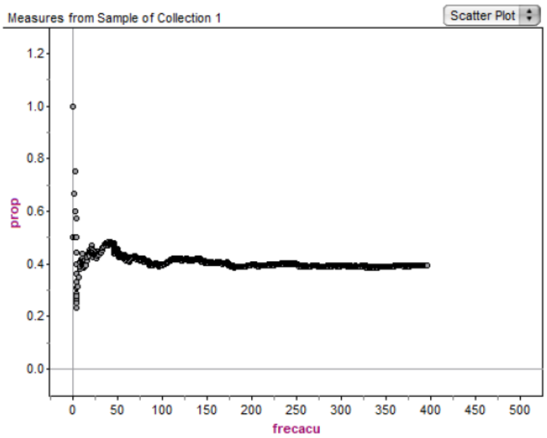


Figura 2. Gráfica elaborada en la simulación

En el inciso 9 (inciso seis de la primera versión), se pregunta que si tuvieran 1000 intervalos ¿cuántos contendrían al ocho?, se les hizo la aclaración que no usaran software; los códigos resultantes son *valor* y *aproximación*, en el primero en los estudiantes responden con el valor que observan en la gráfica (0.3 o 0.4) 300 o 400; en el código *aproximación* las respuestas contienen afirmaciones como: aproximadamente 300 o 400 o daban un rango de valores.

Los incisos 10 y 11, incisos 7 y 8 de la primera versión, los códigos generados son los mismos que en la primera versión.

De esta aplicación se concluye que la simulación física ayuda a entender mejor la situación, sin embargo, siguen las dificultades para elaborar la simulación en Fathom, aun cuando los pasos son casi similares que en la simulación física. Siguen presentando pocas nociones de probabilidad, frecuencia relativa y ley de los grandes números. La gráfica fue muy útil para que los estudiantes observaran que para pocos intervalos no se puede dar una estimación precisa ya que las frecuencias están muy dispersas y que para un número de intervalos grande las frecuencias se estabilizan alrededor de un valor además que ayudó al estudiante a hacer una aproximación sobre cuántos intervalos contendrían al ocho.

Tercera versión, tercera aplicación (16 estudiantes entre 16-17 años, no han cursado Estadística y Probabilidad)

La situación es la misma que en la segunda aplicación, se realiza la simulación física, la modificación fue que ahora los datos que se anotan en el pizarrón son frecuencias relativas, una vez que están todas, se pregunta por la probabilidad y mencionan que es el valor que más se repite.

Ahora se les propone que se considere que fueron 160 intervalos los que se obtuvieron y que de esos 49 contienen al 8 y se comparó con el valor que más se repite (son similares).

La simulación ya se les dio hecha pues siguen presentando problemas en el momento de elaborarla, se les explicó en qué consistía haciendo la comparación con la simulación física. Las preguntas son las mismas que en la versión anterior excepto que no deben hacer una tabla de frecuencias.

La intención en esta aplicación era observar que tanto influía en los estudiantes el hecho de que no hicieran la simulación pues en las aplicaciones anteriores se confundían y distraían mucho. El hecho de que la simulación se les proporcionara ya terminada evitó que se dispersaran cuando tratan de hacerla ya que las dificultades que se les van presentando genera que pierdan concentración.

Los códigos resultantes en esta aplicación fueron iguales a los de la aplicación anterior, lo que nos lleva a concluir que los razonamientos son similares aún cuando los estudiantes de las primeras dos aplicaciones se encontraban cursando la materia de Estadística y Probabilidad.

Conclusiones generales

De las dos últimas aplicaciones concluimos que los estudiantes logran entender que dadas muchas repeticiones habrá poca variabilidad, mismo que fue documentado por Inzunza (2017), quien en sus resultados encuentra que es posible que estudiantes con pocos antecedentes matemáticos y de probabilidad pueden realizar conexiones correctas entre el enfoque clásico y frecuencial de la probabilidad a partir de la visualización del comportamiento de los datos esto fue posible con el polígono de frecuencias generado por la simulación esto se pudo constatar con las aproximaciones que dieron en sus respuestas.

Otra dificultad observada es que los estudiantes tienen dificultades para expresar lo que observan ya que de acuerdo con Sánchez y Valdez (2017) no cuentan con lenguaje probabilista el cual es fundamental que el alumno lo adquiera.

Finalmente hemos determinado que con la actividad se ha trabajado la parte probabilística sin embargo falta considerar que tanto usa el estudiante los resultados en la siguiente situación: si se metieran en una urna todos los posibles intervalos que se pueden formar con los números del 0 al 9, ¿qué tan confiando estarías de obtener uno que contenga al ocho?

Referencias

Belia, S., Fidler, F., Williams, J. y Cumming, G. (2005). Researchers misunderstand confidence intervals and standard error bars. *Psychological Methods*, *10*(4), 389-96.

Ben-Zvi, D. (2006). Scaffolding students informal inference and argumentation. En A. Rossman y B. Chance (Eds.), *Working cooperatively in statistics education: Proceedings of the Seventh International Conference on Teaching Statistics*, Salvador, Brazil. Recuperado de http://www.stat.auckland.ac.nz/~iase/publications/17/2D1_BENZ.pdf

Castro-Sotos, A. E., Vanhoof, S., Noortgate, W. y Onghena, P. (2007). Students' misconceptions of statistical inference: A review of the empirical evidence from research on statistics education. *Educational Research Review*, *2*, 98-113.

Cobb, P. y Gravemeijer, K. (2008). Experimenting to support and understand learning processes. En A. E. Kelly, R. A. Lesh y J. Y. Baek (Eds.), *Handbook of design research methods in education. Innovations in Science, Technology, Engineering and Mathematics Learning and Teaching* (pp. 68-95). Mahwah, NJ: Lawrence Erlbaum Associates.

Cobb, P. y McClain, K. (2004). Principles of Instructional Design for Supporting the Development of Students Statistical Reasoning. En D. Ben-Zvi y J. Garfield (Eds.), *The Challenge of Developing Statistical Literacy, Reasoning, and Thinking* (pp. 375-396). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Cumming, G. (2012). Inference by eye: Pictures of confidence intervals and thinking about levels of confidence. *Teaching Statistics*, *29*(3), 89-93.

Fidler, F., Thomason, N., Cumming, G., Finch, S. y Leeman, J. (2004). Editors can lead researchers to confidence intervals, but can't make them think: Statistical reform lessons from medicine. *Psychological Science*, *15*, 119-126.

Fidler, F. (2006). Should Psychology abandon p values and teach CIs instead? Evidence based reforms in statistic education. En A. Rossman y B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics.* Voorburg, The Netherlands: International Statistical Institute. Institute. Recuperado de http://www.stat.auckland.ac.nz/*iase/publications.php (online).

Garfield, J. B., del Mas, R. C. y Chance, B. L. (2004). Tools for teaching and assessing statistical inference. Recuperado de http://www.gen.umn.edu/research/stat_tools.

Glaser, B. y Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New York: Aldine Publishing Company.

Hogg, R. y Craig, A. (1978). *Introduction to Mathematical Statistics* (pp. 200-234). New York, USA: Macmillan Publishing Co., Inc.

Huck, S. W. (2009). *Statistical Misconceptions*. New York: Taylor and Francis Group.

Inzunza Casares, S. (2017). Conexiones entre las aproximaciones clásicas y frecuencial de la probabilidad en un ambiente de modelación computacional. *Avances de Investigación en Educación Matemática, 11*, 69-86.

Kalinowski, P. (2010). Identifying misconceptions about confidence intervals. En C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics*, Ljubljana, Slovenia. Voorburg, The Netherlands: International Statistical Institute.

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

Pfannkuch, M., Wild, C. y Parsonage, R. (2012). A conceptual pathway to confidence intervals. *ZDM Mathematics Education,* 44, 899-911.

Pfannkuch, M., Forbes, S., Harrawary, J., Budgett, S. y Wild, C. (2013). "Bootstrapping" students' understanding of statical inference. *Teaching y Learning, Research Iniciative*, 1-18.

Pratt, D., Johnston-Wilder, P., Ainley, J. y Mason, J. (2008). Local and global thinking in statistical inference. *Statistics Education Research Journal, 7*(2), 107-129.

Prodromu, T. (2013). Informal Inferential Reasoning: Interval Estimates of Parameters. *International Journal of Statistics and Probability, 2*(2).

Sánchez, E. y Valdez, J. C. (2017). Las ideas fundamentales de probabilidad en el razonamiento de estudiantes de bachillerato. *Avances de Investigación en Educación Matemática, 11*, 127-143.