

XIV EVENTO INTERNACIONAL “MATECOMPU2012”
“LA ENSEÑANZA DE LA MATEMÁTICA, LA ESTADÍSTICA Y LA
COMPUTACIÓN”

SIGNIFICAR LOS INTERVALOS DE CONFIANZA:
UN RETO PARA LA ENSEÑANZA¹

Felipe Fernández Hernández ffernandez@pedagogica.edu.co²

Luísa Andrade Escobar landrade@pedagogica.edu.co

Ingrith Álvarez Alfonso ialvarez@pedagogica.edu.co

INTRODUCCIÓN

La línea de Educación Estadística de la Universidad Pedagógica Nacional, en uno de sus proyectos de investigación titulado *Experimentos de enseñanza y razonamiento estadístico en situaciones de probabilidad e inferencia estadística*, luego de un proceso de estudio acerca de algunas de las conexiones existentes entre la probabilidad y estadística, centra su atención en el significado que se le da al nivel de confianza, antes y después de que se calcule e interprete el intervalo. Bajo esta inquietud, realiza una revisión documental, la cual permite reconocer errores respecto a la interpretación de intervalos de confianza, e inicia así, el diseño de una secuencia de tareas para desarrollar con docentes de matemáticas en formación, con el fin de conceptualizar lo que significa un nivel de confianza, pero con la intención de evitar generar los errores reportados por otras investigaciones.

¹ La información aquí reportada hace parte de resultados parciales del proyecto de investigación Experimentos de enseñanza y razonamiento estadístico en situaciones de probabilidad e inferencia estadística, (DMA-328-12) desarrollado por la Línea de Educación Estadística y financiado por Departamento de Matemáticas y el Centro de Investigaciones de la Universidad Pedagógica Nacional.

² Los autores son docentes de la Universidad Pedagógica Nacional, Colombia, integrantes del grupo de investigación Didáctica de las Matemáticas en la Línea de Educación Estadística.

PROBLEMA DE INVESTIGACIÓN

El trabajo en el proyecto se plantea inicialmente con la idea de establecer conexiones más explícitas entre conceptos de probabilidad y estadística, que permitan un aprendizaje de los contenidos a través de la comprensión de su significado, de su razón de ser, de su surgimiento en la historia y de su utilidad, de forma que se promueva el razonamiento estadístico. En la búsqueda de hacer más transparentes las conexiones entre estadística y probabilidad, se exploran ideas como la de equiparar los conjuntos que intervienen en la relacional funcional asociada a una variable aleatoria, es decir, el espacio muestral y los reales, con la población de estudio y la población de datos del análisis de datos en la estadística, e incluso llevar dicha equiparación al territorio inferencial al considerar la noción de distribución muestral.

Aunque la reflexión realizada no llegó a establecer claramente las ventajas que éstos vínculos podrían representar en la comprensión y desarrollo del razonamiento estadístico de los estudiantes, en la indagación llevada a cabo sí se pone de relieve la importancia de conceptos de la estadística como el de muestra, muestreo y distribución muestral, para inferir acerca de parámetros de poblaciones, en donde la interpretación frecuencial de la probabilidad cobra especial importancia para interpretar los niveles de confianza asociados a los llamados intervalos de confianza. Se nota así mismo, la escasa atención que tales conceptos reciben en los currículos escolares y universitarios, según se aprecia en la falta de estándares al respecto en los Estándares Básicos de Competencias en Matemáticas del Ministerio de Educación Nacional (2006)³. Además, sobresalen las dificultades que experimentan los estudiantes en cuanto a las distribuciones de muestreo, reportadas por la investigación realizada por Fernández, Andrade y Sarmiento (2011)⁴, donde se afirma que mientras muchos estudiantes son capaces de llevar a cabo los cálculos que se requieren en un proceso de inferencia, con frecuencia son incapaces de

³ Ministerio de Educación Nacional (2006). *Estándares Básicos de Competencias en Matemáticas*. Bogotá, Colombia.

⁴ Fernández, F. Andrade, L. & Sarmiento, B. (2010). Variación y diseño de experimentos de enseñanza para la educación estadística (reporte de investigación). Bogotá: Universidad Pedagógica Nacional.

explicar el proceso delineado o de interpretar apropiadamente los resultados de los cálculos.

MARCO TEÓRICO

La conceptualización teórica se concreta en dos sentidos. Por un lado, los resultados del proyecto (Fernández, Andrade y Sarmiento, 2011) sugieren posibilidades de formular una propuesta en la que se explicita la preocupación por explorar conexiones entre conceptos de probabilidad y estadística, para ampliar la comprensión de las conceptualizaciones del grupo de investigación y enriquecer el avance en el desarrollo de propuestas que promuevan razonamiento estadístico en los estudiantes; y por otro lado, los resultados de discusiones sostenidas acerca de cómo concretar conexiones entre estadística y probabilidad, conduce a valorar la relevancia didáctica y práctica de establecerlas y a precisar qué conexiones concretas considerar y la metodología a seguir. El resultado de dicha valoración, lleva a adoptar como metodología el ‘análisis didáctico’ sugerido por Gómez (2002)⁵, y se elige el tema ‘intervalos de confianza’. Esta decisión conduce al abordaje de problemáticas de la enseñanza y aprendizaje como las mencionadas en trabajos de Behar (2007)⁶, Behar y Yañez (2010)⁷, Olivo y Batanero (2007)⁸ y Kalinowski (2010)⁹.

Respecto a la elección de trabajar con el tema de intervalos de confianza, cabe señalar que obedece a la intención de abordar situaciones de muestreo, en las que el foco de atención aborde, entre otros asuntos: la identificación de poblaciones y

⁵Gómez; P. (2002). Análisis didáctico y diseño curricular. *Revista EMA*, 7 (3), 251-292.

⁶Behar, R. (2007). ¿Estamos buscando el ahogado aguas arriba? El caso de la estimación con intervalos de confianza. Primer Encuentro Nacional de Educación Estadística (ENAES), Bogotá.

⁷Yañez, G. & Behar, R. (2010). The confidence intervals: a difficult matter, even for experts. In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*.

⁸Olivo, R. & Batanero, C. (2007). Un estudio exploratorio de dificultades de comprensión de intervalo de confianza. *Revista Unión*, 12, 37-51.

⁹Kalinowski, P. (2010). Identifying misconceptions about confidence intervals In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*.

muestras de datos y de estudio; la variación que se origina en el muestreo entre muestras de un mismo tamaño y entre muestras de diferente tamaño; y el estimar medidas de agrupación como la media poblacional, con base en medias muestrales. En particular, la consideración de asuntos conceptuales relativos a la variación en contextos de estimación por intervalos de confianza, se vio como una opción razonable para dar continuidad a trabajos como los reportados en Fernández, Andrade y Sarmiento (2010) y Arboleda y Urbina (2012)¹⁰. Además, y como asunto de interés especial para el grupo de investigación, se empezó a trabajar en torno a la conceptualización del nivel de confianza de un intervalo.

En relación con este tema, Behar (2007) señala las deficiencias en el conocimiento sobre estimación mediante intervalos de confianza y la falta de comprensión al respecto por parte de los profesores e incluso de los expertos estadísticos, que son motivo de las dificultades de los estudiantes y sugiere que seguir pensando en mejorar el aprendizaje de los estudiantes solo a través de intervenciones e innovaciones en el aula, bajo el supuesto de que los profesores tienen los conceptos claramente entendidos, no parece ser suficiente.

Para Behar “el concepto de estimación por medio de intervalos de confianza es de los más importantes y útiles en estadística”. Un intervalo de confianza para estimar una media poblacional, es un intervalo que estima este parámetro desconocido con base en una muestra aleatoria. Es un conjunto plausible de valores para el parámetro de interés, generado por unos datos que se obtuvieron en un muestreo aleatorio. El intervalo de confianza para la media poblacional (μ), se construye a partir de un estadístico muestral (la media muestral \bar{x}) al que se le suma y resta la medida del margen de error, dependiendo del nivel de confianza elegido. Behar plantea que el significado de confianza es una idea importante que con frecuencia no es clara; al respecto observa que la construcción del intervalo de confianza no se basa en la variable original X , sino en la media de una muestra aleatoria, y por consiguiente la confianza no es una característica de la población de donde se tomó la muestra y no se refiere a las unidades de la población. El nivel de confianza indica la probabilidad de que el método para construir intervalos de confianza, genere al menos uno que contenga la media poblacional μ , es decir, se refiere al método y no

¹⁰Arboleda, C. & Urbina, G. (2012). Razonamiento estadístico de estudiantes universitarios alrededor de situaciones de muestreo apoyadas en tareas de simulación (tesis). Bogotá: Universidad Pedagógica Nacional.

al parámetro; si se repite el muestreo muchas veces y por lo tanto la construcción del intervalo, a la larga, el porcentaje de los intervalos generados que contienen la media poblacional μ , está indicado por el nivel de confianza.

Así, Behar señala que “con el enfoque frecuentista, el contexto en el cual se explica el significado de ‘confianza’, es el de la potencial repetición de las estimaciones, de esta manera “la confianza” no la tiene el intervalo concreto que se obtuvo mediante un procedimiento aleatorio como resultado de la realización de una muestra aleatoria, sino que la confianza está asociada al procedimiento aleatorio que genera el intervalo”. Es decir, la confianza está asociada al procedimiento de construcción o generación de intervalos; esta confianza existe antes de conocer el intervalo de confianza que se va a establecer a partir de los datos de la muestra aleatoria que se tomó. Así que si se tiene esta información a priori, se puede suponer que el 95% –si ese es el nivel de confianza que se determinó– de los intervalos que se establezcan con ese procedimiento dado, va a contener la media poblacional μ . Añade que la verdad es que “jamás se sabrá” si el intervalo establecido contiene la media poblacional μ o no; “no obstante, por las credenciales del procedimiento, si toca decidir, yo puedo actuar como si el intervalo particular hubiera atrapado la media verdadera, con el riesgo asociado al procedimiento generador”. Entonces, suponer que en un intervalo específico está la media poblacional μ , significa que los valores del intervalo son valores plausibles para el parámetro μ .

De la estructura matemática mediante la cual se construye el intervalo de confianza se puede deducir, que entre mayor confianza se establezca para generar los intervalos, más amplios serán éstos. Si se aumenta el tamaño de la muestra para el mismo nivel de confianza, se obtienen intervalos más cortos. Si se aumenta el nivel de confianza, sin aumentar el tamaño de muestra, no se mejora la estimación, sino simplemente se tiene otra expresión de la calidad de la estimación, pues “el aumento de la confianza se hace en detrimento del margen de error”. Lo deseable sería aumentar el nivel de confianza con menor margen de error, o sea con intervalos más cortos, lo cual solo se logra cuando se aumenta el tamaño de la muestra.

Por otra parte, Behar identifica diversas concepciones relacionadas con la comprensión de conceptos asociados con la estimación por medio de intervalos de confianza, las cuales identifica a través de preguntas formuladas a profesores y expertos. Los resultados de este estudio ponen en evidencia “la necesidad de

orientar los esfuerzos de investigación en la comprensión de los conceptos estadísticos en el colectivo de profesores, pues su correcta comprensión es condición sin la cual, los demás esfuerzos de mejoramiento del aprendizaje en nuestros estudiantes tendrá tímidos resultados. No tiene sentido esforzarse en lograr que los estudiantes comprendan lo que muchos de nosotros los profesores no comprendemos, pero no somos conscientes de ello”.

Algunas de estas concepciones correctas y erradas son: un $(1-\alpha) \times 100\%$ de los datos están incluidos en el intervalo de confianza (falsa); hay un $(1-\alpha) \times 100\%$ de probabilidad de que el intervalo de confianza incluya la media muestral (falsa); si se generan 200 intervalos con el 95% de confianza bajo el mismo proceso, aproximadamente 10 de aquellos intervalos no contendrán la media de la población (verdadera); la probabilidad de que el intervalo de confianza construido incluya la media poblacional es 95% (falsa); si se conserva el tamaño de la muestra fijo, el intervalo de confianza se amplía cuando se aumenta el nivel de confianza (verdadera); si se fija el nivel de confianza, el intervalo de confianza se hace más corto cuando se aumenta el tamaño de la muestra (verdadera).

METODOLOGÍA

Al realizar las actividades, tales como el estudio cuidadoso del contenido matemático y el contenido didáctico a tratar con los estudiantes, se inicia el ciclo dinámico de planeación, diseño, implementación y análisis de una secuencia de tareas. Así vemos que el trabajo del grupo en torno al diseño y la elaboración de las tareas de instrucción, se ha acercado en esta investigación al procedimiento para organizar la enseñanza de las matemáticas propuesto por Gómez (2002) y denominado ‘análisis didáctico’, no solo por el hecho de que el trabajo de diseño y planeación se ha llevado a cabo previamente y no de manera simultánea a la implementación de las tareas con los estudiantes, sino también porque las acciones realizadas por el grupo encajan en los procesos que componen un ciclo de análisis didáctico.

En primer lugar, luego de la consulta bibliográfica se fijó como objetivo de aprendizaje para los estudiantes, la comprensión sobre intervalo de confianza. La indagación realizada sobre el contenido matemático, específicamente para

establecer las definiciones e interpretaciones de intervalo y nivel de confianza, y determinar los conceptos ligados a estas nociones y sus relaciones, se ajusta a un análisis de contenido. En cuanto al análisis cognitivo, se han estudiado las dificultades y concepciones erradas de los estudiantes alrededor de la interpretación del nivel de confianza de los intervalos, con base en la literatura revisada y en la experiencia docente de los integrantes del grupo de investigación. También, a partir de la información recogida en pruebas piloto de la secuencia de tareas, se han evidenciado otras dificultades relativas a conceptos relacionados con el objeto de estudio, tales como la diferenciación entre: población y muestra de datos e individuos, parámetros y estadísticos, y la selección de muestras aleatorias, las cuales se han examinado y abordado en el análisis de instrucción, identificando y diseñando tareas como parte de un taller dirigido a que el estudiante mediante la exploración amplíe su conceptualización de intervalo y nivel de confianza de manera adecuada, atendiendo de manera paralela algunos de los conceptos en los que se evidencia dificultades y que son necesarios para generar un ámbito propicio para la conceptualización de intervalos de confianza. Por último, el análisis e interpretación de los desarrollos de los estudiantes en las pruebas piloto que llevaron a considerar cambios en el taller, podría verse como parte del análisis de actuación respecto al docente.

El siguiente, es un esquema representativo de los aspectos que caracterizan el ciclo de análisis didáctico (Gómez, 2002).

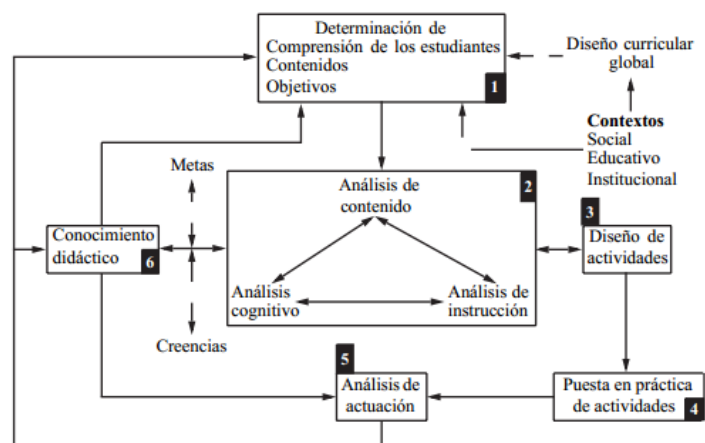


Ilustración 1. Ciclo de análisis didáctico (Gómez, 2002)

La información que hasta el momento se ha recolectado consiste en los desarrollos escritos por parte de algunos estudiantes, a partir de pruebas piloto que se han

implementando con estudiantes de cuarto, quinto y séptimo semestre, que han tomado alguno de los cursos de estadística, probabilidad o inferencia estadística que ofrece la Licenciatura en Matemáticas de la Universidad Pedagógica Nacional, con quienes se puso a prueba la primera y segunda parte de las versiones iniciales del taller. El posterior análisis realizado en torno a esta información, ha generado la posibilidad de evaluar dichas secciones del y mejorar su diseño en cuanto a las preguntas y tareas que se le proponen a los estudiantes, y a las orientaciones didácticas que se le siguieren al docente que implementará el taller.

Más adelante se espera implementar la totalidad de la secuencia de tareas con un grupo de estudiantes de la Especialización en Educación Matemática de la misma Universidad.

PROPUESTA DE TAREAS

Para conceptualizar el significado intervalos de confianza, y como parte de los resultados de la investigación en curso, se ha ido diseñando y ajustando una secuencia de tareas¹¹, constituida en tres partes, con sus respectivos propósitos y guía de gestión de clase.

Primera parte

Propósito: Se busca que el estudiante diferencie entre los individuos observados y la característica medida, mejore la comprensión en lo que se refiere a la muestra y la población, tanto de estudio como de datos.

Gestión de clase: La primera parte se debe trabajar de manera individual. Al estudiante se le entrega la guía de preguntas y el profesor estará atento a resolver inquietudes de los estudiantes y a observar la comprensión de ellos respecto a las preguntas planteadas.

Para la socialización, al terminar el desarrollo de las nueve tareas, el docente debe procurar resaltar que si se quiere describir una característica con una variable de agrupación, lo que se toma para el cálculo son muestras de datos y no muestras de los sujetos mismos. También es pertinente enfatizar en la diferencia entre población de estudio y de datos.

¹¹ Se presenta la última versión ajustada a partir de resultados de pruebas piloto. El taller sigue en construcción y ajuste a partir de la aplicación de este con diversos grupos de estudiantes.

SITUACIÓN

Mundialmente se ha reconocido que aquellas personas que tienen un coeficiente intelectual (C.I.) igual o superior a 125 puntos, son superdotados. En el colegio “Los Pilos” en donde hay 1000 alumnos en los grados de sexto a once, se quiere realizar una prueba para determinar el puntaje del coeficiente intelectual representativo de estos cursos. Los costos de aplicar esta prueba son muy altos y el colegio no dispone de suficientes recursos para aplicarla a todos.

De acuerdo con el enunciado de la situación planteada:

1. ¿Cuál es la característica que interesa observar en este estudio?
2. ¿A quiénes se podría estudiar para observar dicha característica?
3. Proponga un ejemplo de la información que se podría recolectar, identificando la característica que se observa y a quién se le observa.
4. Compare su ejemplo con los siguientes y explique cuál de ellos ilustra mejor la información que se podría recoger :

Ejemplo 1.

Alumno	a1	a2	a3	aj	...
Puntaje de C.I.	112	126	120	98	...

Ejemplo 2.

Alumno	Juan	Luis	Pepe	Lucia	...
--------	------	------	------	-----	-----	-------	-----

Ejemplo 3.

Puntaje de C.I.	112	126	120	98	...
-----------------	-----	-----	-----	-----	-----	----	-----

5. Complete el ejemplo escogido teniendo en cuenta su respuesta al ítem 2.
6. A partir del ejemplo que usted completó liste algunos de:
 - a. Los alumnos seleccionados.
 - b. Las medidas de la característica observada.

7. De la información anterior deben emerger dos colecciones de elementos, una se conoce como *muestra de estudio* y la otra como *muestra de datos*. Identifique cuál es cada una de ellas.
8. Con base en lo analizado en el punto anterior, describa la población de estudio y la población de datos.
9. Describa la diferencia entre la población de datos y la muestra de datos.

Segunda parte

Propósito: Se busca que el estudiante seleccione la muestra y el tamaño adecuado de la misma para realizar el estudio, y diferencie entre parámetro y estadístico.

Gestión de clase: Organizar el trabajo por grupos de 2 o 3 estudiantes. Inicialmente cada grupo selecciona la muestra de estudio de tamaño 20 (se explicará cómo seleccionar una muestra a partir de la selección de números aleatorios del 0 al 9 con calculadora o bolsa) y luego solicitan al profesor los datos correspondientes a los números aleatorios obtenidos. El docente debe contar con un archivo en donde tiene registrado el CI de los 1000 estudiantes de la población. Se debe entregar inicialmente la sección A y después de que hayan contestado, entregar las preguntas de la sección B.

Sección A

10. ¿Qué tamaño de muestra de estudio sería conveniente seleccionar para determinar el puntaje del coeficiente intelectual de los grados de sexto a once y por qué?
11. Seleccionen una muestra aleatoria y registre la muestra de datos.
12. Teniendo en cuenta esta muestra ¿qué podría decir acerca del puntaje del coeficiente intelectual de los cursos? Explique.

Sección B

13. ¿La media que se obtiene de la muestra, en su grupo, será igual a la media obtenida en los otros grupos? Expliquen.
14. Comparen la media obtenida a partir de su muestra con las medias obtenidas por otros grupos. Expliquen si mantienen o no su anterior afirmación.

15. Comenten la validez de la siguiente afirmación.

El hecho de que las medias determinadas puedan ser diferentes, implica que la media de la población no sea única.

16. La respuesta dada al ítem anterior, ¿seguirá siendo igual para otras medidas de resumen estadístico, tales como la mediana, la desviación, etc.? Expliquen.

17. ¿Es posible afirmar que la media del puntaje del coeficiente intelectual de la población, es la que calcularon para su muestra? Expliquen.

Socialización orientada por el docente: Se comparten las respuestas dadas para procurar hacer evidente la diferencia entre parámetro y estadístico o estimador. Si algún grupo afirma que la media de la población no es única se propondrá pensar en cuántas muestras se pueden tomar del tamaño de la población para que se dé cuenta que sólo hay una muestra posible y por lo tanto solo hay una opción para calcular la media.

Tercera parte

Propósito: Confrontar al estudiante con la conceptualización de lo que significa nivel de confianza.

Gestión de clase: Se recordará a los estudiantes el algoritmo para determinar los límites de un intervalo de confianza. Se debe tener la tabla de la distribución normal y calculadora o computador.

18. Si para estimar la media del puntaje del coeficiente intelectual de la población se utilizan intervalos con un nivel del 90% de confianza, describa su interpretación acerca de lo que significa el nivel de confianza antes de calcular el intervalo.

19. Para la siguiente muestra de coeficientes intelectuales de veinte estudiantes, construya un intervalo para la media de la población con un nivel de confianza del 90%.

149 129 119 130 97 128 129 107 98 122
136 113 115 117 118 142 137 120 134 140

20. Describa su interpretación acerca de lo que significa el nivel de confianza respecto al intervalo resultante.
21. Para la siguiente muestra responda las dos preguntas anteriores
- 101 107 130 101 115 104 91 91 121 109
- 104 125 113 98 110 119 102 92 111 120
22. Explicar cómo se pueden hacer compatibles las anteriores interpretaciones sobre los dos intervalos de confianza construidos.
23. Pedro, estudiante de estadística, quien construyó los dos intervalos encuentra contradictorias las dos interpretaciones, ¿puede usted explicar por qué?

Socialización orientada por el docente: La socialización se debe ir dando a medida que el docente lo considere pertinente. Se recomienda tener en cuenta:

Para la pregunta #18, se espera que el estudiante diga algo correcto como:

- ✓ El intervalo contendrá el valor del parámetro con 90% de probabilidad (interpretación a priori).
- ✓ Pero es mejor si dice que en 90 intervalos de 100 contendrá el valor.

Para la pregunta #20, se espera que el estudiante diga algo como:

- ✓ el parámetro estará entre los valores obtenidos con la muestra, con un 90% de probabilidad,
- ✓ el estimador estará en el intervalo con un 90% de probabilidad,
- ✓ el 90 % de los datos de la población estará en ese intervalo.
- ✓ el promedio muestral es el valor del parámetro con un 90% de probabilidad.
- ✓ No se espera que dé una interpretación frecuencial.

CONCLUSIONES

Los avances y logros de la investigación se dan desde lo conceptual y lo metodológico. Desde lo conceptual, atendiendo a lo que dice Behar (2007) acerca de la formación de docentes, los mismos investigadores avanzan en la reconceptualización respecto al significado del nivel e intervalo de confianza, logrando superar errores tales como, que después de haber calculado un intervalo de confianza, se afirme que “el 95% de confianza del intervalo, significa que en dicho

intervalo estará la media poblacional con un 95% de probabilidad”. De igual forma al enfrentarnos a preguntas tales como: ¿para qué sirve el cálculo de un solo intervalo si la interpretación del nivel de confianza es frecuencial? y ¿qué se podría concluir del intervalo calculado?, se ha podido conceptualizar que la probabilidad de 90% que el intervalo contenga la verdadera media, es correcta antes de calcular el intervalo, es decir, dicha afirmación pierde sentido después de que el intervalo se ha calculado; afirmación que se reporta como errada en varios de los documentos consultados.

En cuanto a lo metodológico se logra ver la pertinencia del ‘análisis didáctico’ como orientador del proceso en curso, ya que como fruto del mismo se presenta la secuencia de tareas, a la cual ha contribuido la reflexión frente a los resultados obtenidos en las pruebas piloto como parte del análisis de didáctico que se ha hecho al objeto de estudio, y el análisis de actuación del docente al lograr asociar otros objetos matemáticos necesarios para la significación de los intervalos de confianza, en pro de asumir el reto de su enseñanza.