

Muestra Aleatoria e Intervalo de Confianza. Su conceptualización en el ámbito de los profesionales de la salud

María Florencia Walz

(Universidad Autónoma de Entre Ríos. Argentina)

Fecha de recepción: 03 de mayo de 2015

Fecha de aceptación: 23 de septiembre de 2015

Resumen Diferentes investigaciones informan sobre el significado conceptual que, de Intervalo de Confianza y Muestra Aleatoria, tienen los estudiantes de Estadística de carreras universitarias con orientación no matemática. Sin embargo, poco se sabe acerca de la conceptualización que de tales conceptos subsiste en los graduados. En este estudio se trabaja con trece profesionales de la salud, con la intención de conocer cuál es la concepción general que ellos conservan de éstos objetos y observar la influencia que tiene en su significación, una explicación coloquial bajo una dialéctica inferencial sin formalismos disciplinares teóricos.

Palabras clave Conceptualización, muestra aleatoria, intervalos de confianza, profesionales de la salud

Title **Random Sample and Confidence Interval. Its conceptualization in the field of health professionals**

Abstract Different studies report information about what conceptual meaning of Confidence Interval and Random Sample is grasped by college students studying Statistics in non-mathematically orientated studies. However, little is known about the conceptualization which from such concepts remains in graduates. In this study it worked with thirteen health professionals in order to know which general meaning from these concepts they still keep and assess the influence which colloquial explanation has, under inferencial dialectics, without theoretical formalisms, in such signification.

Keywords Conceptualization, random sample, confidence intervals, health professionals.

1. Introducción

En numerosos artículos científicos relativos a las ciencias de la salud es común observar que los autores detallan el tipo de muestra con la que se trabajó en el estudio (que, generalmente, se dice ser “aleatoria”) y cuál fue o cómo se calculó su tamaño. Tales detalles, en este tipo de publicaciones, son casi obligatorios de cumplimentar; puesto que de esta manera se da comienzo a los aspectos metodológicos del trabajo y permiten dilucidar el alcance inferencial del mismo.

Sin embargo, el concepto de Muestra Aleatoria no siempre es bien empleado y bajo su mote se incluyen muestras restringidas a las unidades experimentales accesibles en el momento de la



investigación, o a muestras obtenidas bajo un diseño de asignación no aleatoria a los grupos experimentales, entre otros casos. En relación a esto, Dawson y Trapp (2005) alegan que en algunos estudios clínicos los investigadores suelen utilizar a todos los pacientes que tienen a la mano y que satisfacen los criterios de entrada (o de inclusión).

Particularmente, el significado de aleatoriedad es el que no resulta del todo claro. Y, de acuerdo con Bennett (2000), sin la noción correcta de aleatoriedad -esencia de la estadística inferencial- los resultados de los estudios pueden llevar a conclusiones sin sentido, con sesgos o no muy fiables.

Por su parte, el cálculo del Tamaño Muestral se interpreta, usualmente, como un resultado determinístico, resultante de una expresión que se supone predeterminada. Así, considerando una de las situaciones más simple, como sería calcular el tamaño muestral para estimar una media poblacional, teniendo como únicas pretensiones, fijar la confianza de trabajo y el error máximo admisible en la estimación, la expresión del cálculo, presentada en muchos libros de estadística aplicada (Walpole y col. 1999, pág. 246; Miller y col. 1992, pág. 212), sirve a los objetivos del investigador aunque no se comprenda la fundamentación teórica que la sustenta; por la existencia, tal vez, de obstáculos epistemológicos que, entre otros, podrían ser los conceptos de Estadísticos y Distribuciones Muestrales (Manzano y Braña, 2003).

Otro aspecto preocupante, que se observa con frecuencia en artículos científicos de ciencias de la salud y biológicas, es que se exponen en tablas los resultados de las medidas descriptivas calculadas con los datos muestrales (media y desvío) y, conjuntamente en la misma tabla, se detallan los Intervalos de Confianza; incluso, para las características que forma parte de los criterios de inclusión de la investigación y que no son una variable aleatoria de interés inferencial. Ignorándose el significado conceptual de la estimación de un parámetro mediante esta técnica.

Como es bien sabido, un Intervalo de Confianza brinda una herramienta inferencial de gran valor, tanto para estimar algún parámetro como para decidir si éste podría tomar un valor supuesto (función análoga a Pruebas de Hipótesis). Así, por ejemplo, en la comparación de dos medias poblacionales; la interpretación de los valores obtenidos en un Intervalo de Confianza construido para estimar el parámetro Diferencia de Medias, lleva a decidir si es altamente probable o no que las medias poblacionales sean iguales. Metodología, muy empleada en los análisis estadísticos realizados por los profesionales de las ciencias de la salud.

Tales cuestiones han motivado la realización de este estudio con la intención de conocer: a) qué conceptualización perdura en los profesionales de las ciencias de la salud sobre los conceptos Muestra Aleatoria e Intervalo de Confianza y b) analizar la influencia de una intervención dialéctica coloquial inferencial, sin formalismos ni derivaciones teóricas, en la significación de los mismos; considerando que, en estos ámbitos (los de la salud), las prioridades disciplinares inducen a una instrucción más pragmática y el entrenamiento en un pensamiento estocástico y matemático es limitado.

2. Estado del arte

El objeto Muestra Aleatoria es un concepto que se torna conflictivo a la hora de seleccionar una de ella de una población. Tanto por mala interpretación de su definición o por dificultades relacionadas con el diseño científico disciplinar de la investigación.

Montgomery define Muestra Aleatoria de la siguiente manera:

“...si una población contiene N elementos, y una muestra de n de ellos será seleccionada, entonces el procedimiento empleado se denomina muestreo aleatorio si cada una de las $N!/[(N-n)!n!]$ posibles muestras tiene la misma probabilidad de ser elegida”. (Montgomery, 1994, pág.18)

Esa probabilidad (de la que habla el autor) es desconocida en casi la totalidad de las investigaciones; por lo que, bajo el pensamiento sin entrenamiento en la lógica inferencial, tal aspecto se vuelve descartable y una Muestra Aleatoria lo es, sólo si respeta ciertas creencias populares acerca de la aleatoriedad y representatividad; que, en términos generales, el primero, es considerado como un sinónimo de lo que “toque en suerte” y el segundo, como sinónimo de grande. Desconociéndose que la aplicación de un Método de Muestreo Aleatorio podría garantizar la equiprobabilidad en las unidades experimentales a seleccionar y, por ende, obtener una Muestra Aleatoria.

Muchos autores deliberaron y deliberan acerca del término “aleatorio” o “azar”; término que al parecer es bastante sencillo de entender, pero que, en su aplicación práctica genera confusión. Al respecto Lorenz reflexiona:

“La verdad es que tengo la sospecha de que la idea popular de que algo se produce al azar es la idea de algo que no presenta una pauta discernible, en lugar de ser algo que comporta la propiedad menos fácilmente detectable de la dependencia sensible”. (Lorenz, 2000, pág. 19)

Enunciado que encierra una complejidad extrema para quienes no están acostumbrados al empleo del pensamiento probabilístico.

Bennett (2000) comenta que el concepto que de azar se tiene, imposibilita pensar que la voluntad, el conocimiento o la inteligencia puedan influenciar en el resultado del destino. Este mismo autor dejó entrever un posible causal de la mala interpretación del vocablo aleatoriedad, advirtiendo que al descuidar la teoría de probabilidad nos alejamos de la comprensión del concepto “aleatoriedad” y en consecuencia es difícil dilucidar lo inherente a la estadística inferencial. Así mismo, el empleo del cálculo probabilístico sin independencia de lo matemático propio, lleva a interpretar la probabilidad como un valor inmutable, puesto que éste es, en sí mismo, determinístico.

Relativo al concepto Intervalo de Confianza, numerosas investigaciones aportan sus valoraciones sobre las dificultades más frecuentes para su comprensión o sobre el significado que tiene en los estudiantes universitarios de diferentes orientaciones. Otras tantas, evalúan distintas metodologías didácticas para mejorar su entendimiento. Por ejemplo, Behar (2001) señala que no se comprende la definición de Intervalo de Confianza, al detectar la creencia de que el rango de valores obtenido corresponde al del estadístico usado como estimador. Este autor dice, también, que existen conflictos para asociar el nivel de confianza con la probabilidad que hay de que el intervalo incluya al verdadero valor del parámetro.

Olivo, Batanero y Díaz (2008), por su parte, identificaron que es frecuente interpretar la definición de Intervalo de Confianza en forma bayesiana, suponiéndose que los extremos del intervalo son fijos y no aleatorios.

Autores como Cumming, William y Fidler (2004) aducen que el Intervalo de Confianza para la media poblacional suele interpretarse como una probabilidad alta de replicación; es decir, una nueva muestra producirá un Intervalo de Confianza parecido al original. Sin embargo, y atendiendo a los resultados de sus investigaciones, mencionan que la inferencia con la estimación por Intervalo de Confianza se comprende más sencillamente que la relacionada con el Contraste de Hipótesis.



En otro sentido, Schenker y Gentleman (2001) han detectado que, en la interpretación de dos Intervalos de Confianza para estimar dos medias de poblaciones independientes, existe la creencia de que sólo son significativamente diferentes cuando se tocan extremo con extremo o no están solapados.

En cuanto a las metodologías didácticas recomendadas para la enseñanza del tema, en general, implican reforzar ciertos conceptos previos, tales como: Población y Muestra, Estadístico y Parámetro (Walz, 2011). Por otra parte, si se considera que la Probabilidad es lo primero que surge en la historia de la Estadística; bajo un paradigma epistemológico, éste tema, en acción conjunta con la educación en el razonamiento probabilístico, debería encabezar toda enseñanza formal de la disciplina (Walz, 2015). Estimulándose, así, desde el comienzo, la concepción y aceptación de la inevitable incertidumbre que encierra la ocurrencia de los distintos sucesos en los fenómenos estocásticos. Lo que contribuirá, por ende, al desarrollo del pensamiento que concibe la posibilidad de anticipar un acontecer futuro. Transmitiendo, de esta manera, el *sentido estadístico* que requiere la instrucción actual de la Estadística aplicada (Batanero y col., 2013).

Por su parte Garfield, delMas y Chance (1999) sugieren estimular el aprendizaje de Intervalo de Confianza y de Muestreo Aleatorio, mediante simulación y Terán (2006) discute cómo la interacción con la computadora mejora los conocimientos relativos a éstos.

Relativo a las dificultades entorno al cálculo del Tamaño Muestral, no son tantos los estudios específicos que refirieran sobre las conceptualizaciones que de este concepto se hacen en los distintos perfiles del estudiantado universitario y, mucho menos, en los profesionales de cualquier orientación. Pero, considerando que sus expresiones se derivan del planteo teórico empleado en la deducción de un Intervalo de Confianza, es posible que los mismos obstáculos existentes para la comprensión de éste último, sean los que se interpongan en la comprensión de aquel, tanto en términos de la confianza, como del error deseado; simplificándolo a una fórmula o receta al momento del diseño de la investigación.

3. Metodología de trabajo

El estudio fue diagramado en tres partes, con el objeto de alcanzar los objetivos planteados en el mismo. Se trabajó con una muestra restringida, sujeta a que los participantes cumplieran los siguientes criterios de inclusión: a) ser profesional de la salud, residente en la ciudad de Paraná o de Santa Fe (capitales de provincias argentinas), b) ser uno de los autores del artículo científico (seleccionado por el investigador de este trabajo) en el que se especificara que se trabajó con una Muestra Aleatoria y que se tuviera como objetivo la estimación de algún parámetro, c) haber realizado algún curso formal de Estadística y d) aceptara participar, ateniéndose a las pautas y condiciones diagramadas para el estudio.

Con tales criterios, el número de potenciales participantes se redujo considerablemente; por lo que la selección no pudo ser aleatoria. Principalmente, los criterios a), b) y d) determinaron la restricción, debido a la imposibilidad de contactar a los profesionales autores y que, además, accedieran a participar. Por tales razones la Muestra Restringida quedó conformada por trece profesionales: cuatro médicos, seis bioquímicos y tres nutricionistas licenciados. Ocho de la ciudad de Santa Fe y el resto de la ciudad de Paraná.

El escaso Tamaño Muestral y la no aleatoriedad en la selección de la muestra deberá ser tenido en cuenta, en caso de que este artículo sea empleado por otros autores para realizar comparaciones con sus estudios o inferir acerca de cuestiones poblacionales semejantes. Tales aspectos podrían estar

sesgando las conclusiones; debido a las semejanzas pedagógicas y didácticas, tanto en la formación de grado como en la de estadística de los participantes (por ser egresados de universidades de la región).

De los artículos seleccionados de cada participante, cuatro tenían dos muestras para comparar sus medias, los demás tenían solo una muestra con la que se pretendía estimar una media poblacional.

Primera parte: Indagatoria sobre los artículos científicos

En esta primera instancia, y con el objeto de analizar la conceptualización de Muestra Aleatoria e Intervalo de Confianza, se le pidió a cada participante qué explicara algunas cuestiones expresadas en su artículo científico, relativas a los objetos estadísticos enunciados.

La primera pregunta fue: *¿Por qué la muestra recolectada es una Muestra Aleatoria?*, incluso para los artículos que trabajaban con dos muestras “aleatorias” (con el objeto de comparar dos medias poblacionales entre un grupo control y el otro con alguna patología, síntoma o hábito).

Luego, las preguntas estuvieron dirigidas a evaluar el significado dado a los valores obtenidos en los intervalos presentados. En cinco de los nueve artículos con una sola muestra, se presentaba solamente un Intervalo de Confianza construido bajo los lineamientos de una estimación de media poblacional para muestras chicas; en los otros cuatro se detallaban, además, los intervalos correspondientes a la media \pm el desvío estándar de cada característica evaluada. Los cuatro artículos que trabajaban con dos grupos, presentaban un Intervalo de Confianza para la diferencia de medias (para muestras independientes).

A todos los participantes se les preguntó (señalándosele en su artículo, el intervalo de interés para el investigador de este trabajo) *¿Qué significa este rango de valores aquí presentado? Y ¿Qué podría inferir con él?* A los que presentaron los intervalos para la media \pm el desvío se les consultó *¿Qué información brinda este rango de valores?*

Durante la consulta, el entrevistador solo tomó nota de las respuestas no intervino ni con acotaciones de opinión ni con aclaraciones pertinentes a errores.

Segunda parte: Cuestionario

En esta oportunidad, realizada inmediatamente después de la indagatoria anterior, se le solicitó a cada voluntario, que respondiera un cuestionario escrito (bajo presencia del entrevistador) con la recomendación de que solo apelaran a la memoria y/o conocimientos, es decir, sin darles la posibilidad de refrescarlos previamente con lectura bibliográfica.

El instrumento contenía siete preguntas abiertas, destinadas a evidenciar con sus propias palabras, el significado conceptual de los conceptos objetos de estudio. El ítem 8 es un ejercicio para indicar si la proposición enunciada es Verdadera o Falsa. Los restantes ítems (del 9 al 13) son de múltiples opciones de respuestas; todos consignados para evaluar los conocimientos teóricos formales subsistentes. La estructura final del mismo se detalla a continuación:

1. ¿Qué entiende por población y qué por muestra?
2. ¿Cuándo una muestra es una Muestra Aleatoria?
3. ¿Cómo debe ser una muestra si lo que se quiere es inferir resultados a una población?
Explique lo que enuncie.
4. ¿Qué tamaño debe tener una muestra para hacer inferencias?



5. ¿Cómo calcula un Tamaño Muestral?
6. Cuando calcula un promedio muestral (con los valores obtenidos en una muestra aleatoria), qué puede inferir con éste valor.
7. ¿Qué información brinda el rango de valores de un Intervalo de Confianza del 95% para una media poblacional?
8. Responder con Verdadero o Falso las proposiciones enunciadas para el caso en el que usa la expresión: $n=(Z_{\alpha/2} \sigma^2)/e^2=(Z_{0,05} \sigma^2)/5^2$ en el cálculo del Tamaño Muestral, cuando lo que se quiere estimar es un promedio poblacional. Justificar porqué asevera su falsedad.
 - a) Con ese Tamaño Muestral podemos decir, con una confianza del 90%, que el error que se cometerá en la estimación de la media diferirá como mucho en 5 unidades del verdadero promedio poblacional. (V)
 - b) Con ese Tamaño Muestral podemos decir que el error que se cometerá en la estimación será del 5% con una confianza del 95%. (F)
 - c) Al pretender menor error en la estimación, el n resultante es más grande. (V)

Hasta aquí, las preguntas fueron redactadas por el autor de esta investigación. Las mismas fueron revisadas por un experto en cuanto a la factibilidad de interpretación. La comprensión semántica y disciplinar de todo el cuestionario, fue puesta a prueba con ocho alumnos que finalizaban el cursado de la asignatura Probabilidad y Estadística del Profesorado en Química de la Facultad de Ciencia y Tecnología de la Universidad Autónoma de Entre Ríos. Con este cuestionario se intenta evaluar diferenciación entre Población/Muestra, concepto de Muestra Aleatoria, inferencias por estimación y relaciones incorrectas entre ancho del intervalo y el coeficiente de confianza, entre el Tamaño Muestral y la precisión y entre el Tamaño Muestral y la confianza.

Los ítems consecutivos del 9 al 13, detallados a continuación, corresponden a los ítems 1, 2, 3 y 4 extraídos, textualmente, del Anexo 1 de Olivo, Batanero y Díaz (2008) y al ítem 10 del mismo anexo, modificado según cursivas, de la siguiente manera “La distribución muestral utilizada en la construcción de un Intervalo de Confianza para *estimar la media de una población supuestamente normal en muestras pequeñas es: ...*” Lo correcto indicado por los alumnos, debía ser, además, justificado.

9. El Intervalo de Confianza de 50% para la media de una población μ es:
 - a. El rango dentro del cual cae 50% de los valores de la media de la muestra.
 - b. Un intervalo más ancho que el Intervalo de Confianza de 95%
 - c. *Un intervalo de valores calculado a partir de los datos de la muestra. En 50% de las muestras de una población, el intervalo calculado contiene a la media de la población.*
 - d. Dos veces más ancho que el Intervalo de Confianza de 100%.
10. Comparado con los Intervalos de Confianza calculados en muestras de tamaño $n=4$, el ancho de los Intervalos de Confianza de la media de una población normal calculado en muestras de tamaño $n=50$:
 - a. Variará más que los anchos de los intervalos para muestras de tamaño $n=4$.
 - b. *Variará, pero no tanto como los anchos de los intervalos para muestras de tamaño $n=4$.*
 - c. Tomarán valores parecidos.
11. Si, manteniendo todos los demás datos fijos, el nivel de confianza se reduce (por ejemplo, de 90% a 80%):
 - a. El Intervalo de Confianza no cambia.
 - b. El Intervalo de Confianza será más ancho.
 - c. *El Intervalo de Confianza será más angosto.*
 - d. El cambio en el Intervalo de Confianza no es predecible.
12. En un Intervalo de Confianza de 95% para la media:

- a. Si se toman muchas muestras y con cada una se construye el intervalo, la media muestral caerá dentro del Intervalo de Confianza 95% de las veces.
 - b. La probabilidad de que caiga en un Intervalo de Confianza calculado de una muestra específica es 0.95.
 - c. Si se toman muchas muestras de igual tamaño, 95% de los Intervalos de Confianza calculado contendrían a μ .
13. La distribución muestral utilizada en la construcción de Un Intervalo de Confianza para estimar la media de una población supuestamente normal en muestras pequeñas es:
- a. Distribución *t* de Student.
 - b. Distribución Chi-cuadrada.
 - c. Distribución normal.
 - d. Distribución F.

Nota: la opción correcta en cada ítem del 9 al 13, se indicó mediante escritura *itálica*.

Tercera parte: Coloquio y revisión teórica

Finalizado y entregado el cuestionario anterior, a cada participante se le explicó qué es y cómo se obtiene una Muestra Aleatoria y su importancia en la inferencia estadística, cuales son las diferencias entre Estadísticos Muestrales y Parámetros Poblacionales y el concepto global y funcional de un Intervalo de Confianza para estimar una media poblacional. La metodología empleada fue coloquial en todo su sentido, haciendo uso permanente de una dialéctica inferencial. No se escribió ninguna expresión matemática o estadística relacionada con distribuciones ni se derivó ninguna expresión de Intervalo de Confianza.

Finalizadas las explicaciones, sin mediar anotaciones sobre las mismas, se le pidió a cada participante que respondiera nuevamente las preguntas abiertas 2, 3, 6 y 7 del cuestionario anterior; considerando que las mismas dan la posibilidad de que expongan lo entendido sobre la explicación dada, haciendo uso de sus palabras y no mediante opciones de interpretaciones teóricas formales y cerradas.

4. Resultados y discusión

De la primera parte: Indagatoria sobre los artículos científicos

El argumento unánime para justificar que la muestra era aleatoria, fue que los pacientes fueron seleccionados al azar. Sin embargo, ninguno empleó algún método de Muestreo Aleatorio. Es decir: cada unidad experimental ingresó a la investigación en la medida que el sujeto era reconocido por el profesional como potencial participante, por satisfacer los criterios de inclusión de su investigación. Completado el Tamaño Muestral o el período disponible para recabar información, se cerraba la muestra. Esto reafirma lo expresado por Dawson y Trapp (2005) y da indicios del desconocimiento de la definición real de Muestra Aleatoria y de Muestreo Aleatorio. Las respuestas fueron concretas y claras; ninguno demostró ambigüedad, ni tenía, al menos, la duda de si su muestra era o no, realmente, una Muestra Aleatoria.

Esta selección de “lo que toque al azar por destino” podría resultar ser, en sí misma, una aproximación a un Muestreo Aleatorio; pero no siempre es el caso. Es más, podría decirse que, en estudios clínicos, casi nunca lo es (por la gran variabilidad biológica, social, cultural, ... de los seres humanos). En general, al no prever un diseño con Muestreo Aleatorio en una investigación, las



inferencias se desvían hacia otra población. Al decir de Colton (1974) se estarían involucrando a dos poblaciones no bien diferenciadas: una es sobre la cual se quiere inferir o generalizar los resultados y la otra es la población muestreada.

Cierto es que no siempre se dispone ni de tiempo ni de casuística para responder a los lineamientos y requerimientos de un Muestreo Aleatorio; en tales situaciones, informar las condiciones en las que se toma la muestra y esclarecer los sesgos que en ella podrían estar operando, fortalecería la confiabilidad de los resultados. Sin embargo, describirlo requiere que se reconozca el escenario y, de hecho, entender significativamente los conceptos de Muestra Aleatoria y otras Muestras; indispensable para la inferencia que se hará.

Todas las explicaciones recibidas por los que tenían una muestra y acompañaban las medidas resúmenes de las variables estudiadas con un Intervalo de Confianza, coinciden con lo observado por Behar (2001), en cuanto a que se le adjudica ese rango a la movilidad del estadístico. Una de las respuestas que más reflejaba ese sentido decía: “La media calculada con datos muestrales toma valores dentro de ese intervalo porque depende de los datos que conformen la muestra”. Quienes presentaron el intervalo correspondiente a la media \pm el desvío estándar muestral, por su parte, desconocían la no equivalencia conceptual de éste en comparación con Intervalo de Confianza. Por ende, desconocían también su aporte referido al porcentaje de los datos observados contenidos en ese rango de valores (según teorema de Chebyshev) para el caso de las variables con distribución normal, como afirmaban tener. Los cuatro que presentaron el Intervalo de Confianza para la diferencia de medias, no supieron interpretar el concepto de *Diferencia* como parámetro y *Diferencia* como estadístico. Sin poder, tampoco, explicar claramente por qué se puede inferir que las medias poblacionales son o no iguales cuando el intervalo incluye o no al cero, respectivamente.

De la segunda parte: Cuestionario

Pregunta 1. Todas las respuestas fueron correctas, de acuerdo a las definiciones formales que de ellas se hacen en muchos de los libros de Bioestadística, por ejemplo, lo presentado por Dawson y Trapp (2005, pp. 371-372) o por Milton (2001, pág. 1).

Pregunta 2. Todas las respuestas a esta pregunta tuvieron en común el vocablo “azar”. En palabras más o menos, expresaban: “...cuando se eligen al azar las unidades que la conforman”. Ninguno advirtió acerca de la necesidad de equiprobabilidad de las unidades para ser elegidas, ni de los métodos de Muestreo Aleatorio; que, implícitamente, son los que garantizan esa equidad en la selección. Resultados que dan cuenta del desconocimiento de su definición elemental.

Pregunta 3. Todas las respuestas expresaban que debía ser *Aleatoria* y *Representativa* de la población. Tras un análisis cualitativo de las explicaciones dadas respecto a lo concerniente a *Representativa*, su significado puede resumirse en sinónimo de grande; dando cuenta, también, de la necesidad de introducir en la muestra, unidades experimentales de los distintos estratos que pudieran existir en la población a estudiar (espontánea e independientemente, cuatro de los entrevistados, ejemplificaron como debería hacerse una encuesta en boca de urna para predecir resultados de elecciones); acotando, que de ese modo, se minimizaba el “error en los resultados”. O sea, intentaron reflejar la representatividad a través de un Muestreo Estratificado.

Pregunta 4. Todas las respuestas aducían que el tamaño debe ser tal que garantice la representación de la población. Ninguno dejó entrever, la relación del Tamaño Muestral con las necesidades o deseos del investigador en cuanto a la confianza y el error máximo aceptado en la estimación del parámetro en estudio. Coincidentemente con lo manifestado en la entrevista previa.

Pregunta 5. Las respuestas a esta pregunta fueron variadas, por ejemplo: usando un tamaño parecido al que tienen otros investigadores en estudios semejantes o que lo calculan según la “fórmula que correspondiera” o que lo calculaban con un programa estadístico, entre otras. Ante tales respuestas queda claro la visión determinista que, en general, se tiene del cálculo del Tamaño Muestral.

Pregunta 6. Las respuestas obtenidas dan indicios de que el promedio muestral calculado es considerado como un descriptor, con atribuciones solamente para la muestra, descartándosele incumbencias o funciones en la inferencia. Ninguno de los entrevistados dijo que la media muestral fuera un estimador puntual de un parámetro. Resultados que concuerdan con la apreciación de Fidler y Cumming (citado por Olivo, Batanero y Díaz, 2008) en cuanto a que se ignora la naturaleza inferencial que tiene. Vale aclarar que uno de los participantes perfiló una noción inferencial al decir (en tiempo verbal conjuntivo): “La media muestral es el promedio que tendría la variable en la población”. Sin embargo, pareciera ser que no considera a la media muestral como variable aleatoria, unificando el valor del estadístico y del parámetro en uno solo.

Pregunta 7. Dos de los participantes interpretaron la información brindada por un Intervalo de Confianza como el rango de valores que podrían tomar futuras observaciones, los restantes lo asociaron al rango de valores que podrían tomar las medias de distintas muestras; esto último vislumbra que reconocen la variabilidad muestral de los estadísticos pero refuerza lo dicho anteriormente, en cuanto al desconocimiento de su significado estimador. Por otra parte, no todos se condijeron con lo que contestaron en la primera parte (Indagatoria sobre los artículos científicos). Es decir, bajo argumentación propia, vale la representación que le han dado al objeto en una aplicación real y, bajo presión de sapiencia de significados teóricos (segunda parte: Cuestionario), vale otra.

Ejercicio 8 de Verdadero y Falso, destinado a evaluar el conocimiento relativo a precisión y confianza en el cálculo del Tamaño Muestral para realizar estimaciones, pone de manifiesto que el uso de la estadística por usuarios no estadísticos se basa, considerablemente, en el empleo de reglas mecanizadas e interpretaciones estereotipadas (instauradas, evidentemente, como obstáculos epistemológicos) que desvirtúan el significado real del objeto estadístico. A las proposiciones a y b no las respondieron bien ninguno de los entrevistados. Al ítem c, solo lo contestaron bien, tres personas, pero remitiéndose al objeto matemático: al deducir que si disminuían el valor del denominador en la “fórmula” el resultado del cociente daba un número más grande. Claramente, una conceptualización determinista.

Los ítems 9, 10, 11, 12 y 13, fueron empleados para conocer qué significado teórico perdura de estos objetos estadísticos estudiados, al menos, dos años antes.

Ítem 9. A la opción correcta (la c) todos los participantes la respondieron mal, adjudicándole lo correcto a la opción “a”. Este ejercicio, según sus autores: Cruise, Dudley y Thayer (citados por Olivo, Batanero y Díaz, 2008) intenta detectar los siguientes sesgos: el intervalo se refiere a la media muestral (distractor a), relación incorrecta entre ancho del intervalo y coeficiente de confianza (distractores b y d). Lo obtenido sigue marcando la tendencia observada en las respuestas anteriores de relacionar el intervalo con los posibles valores del estadístico.

El ítem 10 formulado por Gardfield, delMas y Chance (citados por Olivo, Batanero y Díaz, 2008) evalúa la comprensión del efecto del Tamaño Muestral sobre la precisión (ancho de intervalos) cuando se mantiene constante el coeficiente de confianza. Cinco lo respondieron bien (opción b), los restantes no contestaron por no interpretar lo expresado en cada ítem. Esto da indicios de la no existencia de un aprendizaje significativo de los contenidos teóricos, que permita comprender el origen de la expresión de cálculo como así también el sentido de la estimación de parámetros a través de Un Intervalo de Confianza.



Los ítems 11 y 12, formulados por Cruise, Dudley y Thayer (citado por Olivo, Batanero y Díaz, 2008) se refieren al efecto del Coeficiente de Confianza sobre el ancho del intervalo y su variación en diferentes muestras. Solo dos respondieron bien las opciones correctas de ambos ítems (c y c); pero uno solo argumentó, para el ítem 11, que cuando más confianza se pide se sacrifica precisión. El resto no contestó nada, aduciendo no saber a que se refería lo enunciado. Evidentemente, la estructura conceptual que le da el sentido a la Estimación por Intervalo de Confianza es muy frágil.

El ítem 13 fue elaborado por Olivo, Batanero y Díaz (2008). El mismo fue modificado por el autor del presente trabajo, por necesidad de adaptación a la situación de su estudio; pero no interfiere en lo que se evalúa, que son los conocimientos relativos a la distribución muestral de un estadístico. Siete participantes marcaron la opción correcta “a” pero sin justificar; los otros no contestaron. Detectándose, así, incompreensión del sustento teórico que lo avala y de quién deriva la construcción del Intervalo de Confianza. En tal sentido, habría falencias conceptuales en cuanto a estimador puntual como variable aleatoria y la distribución muestral a la que se ajusta.

De la tercera parte: Coloquio

Tras la explicación teórica coloquial, el 100% respondió correctamente las preguntas 2 y 3 (sobre Muestra Aleatoria y la necesidad de contar con una de ésta para hacer inferencias estadísticas).

Las preguntas abiertas 6 y 7 sobre Estimación Puntual y por Intervalo de Confianza relativa al promedio, bajo una lupa de análisis cualitativo, en general, estuvieron bien respondidas. El 100% contestó correctamente que el verdadero promedio poblacional puede ser estimado puntualmente con el promedio muestral y que el Intervalo de Confianza indica los posibles valores que “podría” tomar aquel. Sin embargo, solo cinco de los trece, expresaron explícitamente el grado de certeza implícito en el “podría”, estableciendo que en la confianza de trabajo estaba inmiscuida la probabilidad de la veracidad enunciada.

La mayoría de las respuestas a la pregunta 7 expresaban lo siguiente (transcribo textualmente una de ellas): “Se estima que el promedio poblacional podría tomar cualquiera de los valores del rango del intervalo”. La palabra *podría*, la empleó el 90% de los entrevistados, dando indicios de que contemplan la incertidumbre en la estimación. Con lo obtenido se evidenció un cambio conceptual; al perderse la idea de que el Intervalo de Confianza es un rango de valores concerniente al estadístico, vislumbrándose un constructo mas adecuado del objeto estadístico.

5. Conclusión

El significado inicial que de Intervalo de Confianza tienen los participantes de este estudio no difiere de lo observado por alguno de los autores mencionados en este trabajo; especialmente, el hecho de considerar el Intervalo de Confianza como el rango de valores en el que se movería el estadístico muestral. Mientras que sobre Muestra Aleatoria puede decirse que, prácticamente, desconocen tanto su definición formal como su significado para la inferencia estadística.

Respecto a la expresión de cálculo del Tamaño Muestral, podría aseverarse que en su conceptualización hay un alto contenido de valoración determinista. Evidenciándose la desunión entre éste y el de estimación por Intervalo de Confianza.

Dado que la muestra observada en este estudio, no es aleatoria, no se realizan inferencias hacia la población de interés. Sin embargo, la valoración cuantitativa y cualitativa de las respuestas a las

preguntas abiertas (de la Tercera Parte), evidencian cierta tendencia de mejora, que permiten reflexionar que una intervención coloquial impacta positivamente en la comprensión de los conceptos evaluados. Es decir, la dialéctica inferencial empleada en la explicación coloquial (sin formalismos teóricos) de los conceptos Muestra Aleatoria e Intervalo de Confianza, aparentemente, tiene un rol importante en el mejoramiento o restitución de su conceptualización correcta. La que, sin desmedro del rigor científico por el retaceo teórico, antepone el discernimiento perceptivo al conocimiento formal, repercutiendo favorablemente en el aprendizaje significativo del objeto de estudio.

Reflexión final: La enseñanza de la *Estadística aplicada* como ciencia inserta en la currícula de casi todas las carreras universitarias comienza, poco más de dos décadas atrás, bajo la tutela exclusiva de los matemáticos, que eran quienes la “estudiaban y entendían”, empleándose esencialmente la didáctica de la Matemática al enseñarla. Hoy en día, sabemos que ambas disciplinas distan entre sí; puesto que la *Estadística aplicada* requiere ser generosamente sazonada con la didáctica de la ciencia en la que se aplica, confrontándola permanentemente con la naturaleza estocástica de la realidad que se investiga. La que no está ni en los números ni en las teorías sino en su propio tejido; el cual hay que intentar dilucidar bajo la lupa de un razonamiento probabilístico. En palabras de Wolfowitz:

“De lo que debemos protegernos es del desarrollo de una teoría que, por una parte, tiene poca o ninguna relación con los problemas reales de la Estadística, y que, por otra parte, cuando se ve como Matemática pura, no es lo suficientemente interesante, por si misma, ni para sobrevivir”. (Wolfowitz, 1969, pág. 42).

Agradecimientos

A todos los profesionales de la salud que accedieron a participar en este estudio.

Nota

Los artículos científicos analizados en este estudio, autoría de los participantes, no son referenciados bibliográficamente con el objeto de preservar su identidad.

Bibliografía

- Batanero, C., Arteaga, P., Serrano, L. y Ruiz, B. (2013). Prospective primary school teacher's Perception of Randomness. E. J. Chernoff y B. Sriraman (eds). *Probabilistics thinking: presenting plural perspectives. Advances in Mathematics Education*. Series Springer, 345-366. Recuperado el 20 de marzo de 2014 de <http://link.springer.com/book/10.1007/978-94-007-7155-0>
- Batanero, C., Tauber, L. y Sánchez, M. V. (2001). Significado y comprensión de la distribución normal en un curso introductorio de análisis de datos. *Cuadrante*, 10(1), 59-92. Recuperado el 20 de noviembre de 2014 de <http://www.ugr.es/~batanero/ARTICULOS/Cuadrante.pdf>
- Behar, R. (2001). *Aportaciones para la mejora del proceso de enseñanza-aprendizaje de la estadística*. Tesis Doctoral de la Universidad Politécnica de Cataluña.
- Bennett, D. (2000). *Aleatoriedad*. Alianza Editorial: Madrid.
- Colton, T. (1974). *Statistics in Medicine*. Little, Brown and Company: Boston.
- Cumming, G., Williams, J. y Fidler, F. (2004). Replication and researchers' understanding of confidence intervals and standard error bars. *Understanding Statistics*, 3, 299-311.
- Dawson, B. y Trapp, R.G. (2005). *Bioestadística Médica*. El Manual Moderno: México.



- Garfield, J.B., delMas, R.C. y Chance, B.L. (1999). The role of assessment in research on teaching and learning statistics. *Annual Meeting: American Educational Research Association*. Montreal, Canadá. Recuperado el 20 de octubre de 2014 de http://www.scielo.org.mx/scielo.php?script=sci_nlinks&ref=1000387&pid=S1665-5826200800030000200011&lng=es
- Lorenz, E.N. (2000). *La esencia del caos*. Debate Pensamiento, S.A: Madrid.
- Manzano, V. y Braña, T. (2003). Análisis de datos y técnicas de muestreo. J.P. Lévy y J. Varela (eds). *Análisis multivariable para las ciencias sociales*, 91-143.
- Miller, I. R., Freund, J.E. y Johnson, R. (1992). *Probabilidad y estadística para ingenieros*. Prentice-Hall Hispanoamericana S.A: Méjico.
- Milton, S. (2001). *Estadística para Biología y Ciencias de la Salud*. Mc Graw Hill/Interamericana de España S.A.: Madrid.
- Montgomery D. (1994). *Diseño y análisis de experimentos*. Grupo editorial iberoamericana: Méjico.
- Olivo, E., Batanero, C. y Díaz, C. (2008). Dificultades de comprensión del intervalo de confianza en estudiantes universitarios. *Educación Matemática*, 20(3), 5-32. Recuperado el 26 de febrero de 2014 de <http://www.revista-educacion-matematica.com/volumen-20/numero-3/5-32>.
- Schenker, N. y Gentleman, J.F. (2001). On judging the significance of differences by examining the overlap between confidence intervals. *The American Statistician*, 55, 182-186.
- Terán, T. (2006). Elements of meaning and its role in the interaction with a computational program. A. Rossman y B. Chance (eds.), *Proceedings of the Seventh International Conference on Teaching Statistics*. IASE. Salvador de Bahía, Brasil. Recuperado el 20 de febrero de 2014 de http://www.scielo.org.mx/scielo.php?script=sci_nlinks&ref=1000417&pid=S1665-5826200800030000200026&lng=es
- Walpole, R.E., Myers, R.H. y Myers, S.L. (1999). *Probabilidad y estadística para ingenieros*. Prentice-Hall Hispanoamericana, S.A.: Méjico.
- Walz, M.F. (2011). *Hacia el aprendizaje significativo de los test de hipótesis en las Ciencias experimentales*. Tesis de Maestría. Universidad Nacional del Litoral, Argentina. Recuperado el 30 de marzo de 2015 de <http://www.bibliotecavirtual.unl.edu.ar:8180/tesis/bitstream/1/279/1/tesis>
- Walz, M.F. (2015). Programa de Estadística aplicada a la Biología: una propuesta. *Números. Revista de Didáctica de la Matemática*, 88, 17-29. Recuperado el 03 de marzo de 2015 de http://www.sinewton.org/numeros/numeros/88/Articulos_02.pdf
- Wolfowitz, J. (1969). Reflections on the future of mathematical statistics. R.C. Bose et al. (eds.), *Essays in Probability and Statistics*, 21-45.

María Florencia Walz. Facultad de Ciencia y Tecnología (FCyT). Universidad Autónoma de Entre Ríos. Paraná, Entre Ríos. Argentina. Nacida en Paraná, el 3 de mayo de 1968. Bioquímica, Especialista en Estadística, Master en Didáctica de las Ciencias. Profesor de las asignaturas: Bioestadística del Profesorado en Biología y de Estadística de Bioquímica, Licenciatura en Nutrición y Licenciatura en Terapia Ocupacional de la Facultad de Bioquímica y Cs. Biológicas de la Universidad Nacional del Litoral (Santa Fe, Argentina). Directora del Gabinete de Asesoramiento Estadístico para tesis e investigaciones de la FCyT. Profesor responsable permanente de Estadística de la carrera de posgrado Maestría en Geomática aplicada a la gestión de riesgos ambientales de la FCyT. Publicaciones sobre didáctica de la Estadística: *Programa de estadística aplicada a la Biología: una propuesta* (Rev. Número, 2015), *El pensamiento estocástico: entorno indispensable en la enseñanza de la Estadística Inferencial* (Rev. Aula Universitaria, 2015), *Obstáculos epistemológicos en el aprendizaje significativo de los test estadísticos en las ciencias experimentales* (Rev. Bilateral Brasil Argentina, 2012), *Hacia un aprendizaje significativo de los test de hipótesis* (Revista Anual de la Facultad de Bioquímica y Ciencias Biológicas, 2011), *Adecuación de la asignatura Estadística al perfil del alumno de la Licenciatura en Nutrición* (Rev. Aula Universitaria, 2010).
Email: florencia.walz@gmail.com