Construcción de significados sobre distribuciones muestrales y conceptos previos a la inferencia en un ambiente de simulación computacional

Santiago Inzunsa Cázares

Resumen: En el presente artículo, presentamos resultados de una investigación sobre los significados que un grupo de 11 estudiantes universitarios construyeron sobre distribuciones muestrales y otros conceptos previos a la inferencia estadística, en un ambiente de simulación computacional como el que proporciona Fathom, un software que ha sido desarrollado para la enseñanza de la estadística y la probabilidad. El estudio se enfocó en identificar los diferentes elementos de significado puestos en juego por los estudiantes y su evolución como consecuencia del ambiente de simulación. Los principales conceptos abordados en el estudio fueron la variabilidad muestral, el efecto del tamaño de muestra en el comportamiento de las distribuciones muestrales y en las probabilidades de resultados muestrales. Se consideraron distribuciones muestrales de medias y proporciones.

Palabras clave: distribuciones muestrales, inferencia estadística, simulación computacional, educación estadística.

Building meanings for sample distributions and inference concepts in a computer environment

Abstract: In the present paper we report results of a research about the meanings that a group of 11 university students constructed about sampling distributions and other concepts previous to the statistical inference in a computer simulation environment like the one provided by Fathom, a software that has been developed for teaching statistics and probability. The study focused in identifying the different elements of meaning constructed by the students and their evolution as a result of the computer simulation environment. The main concepts considered in the study were the sampling variability, the effect of sample size on the behavior of the sampling distributions and the probabilities of samples results. Sampling distributions of means and proportions were considered in the study.

Keywords: sampling distributions, statistical inference, computer simulation, statistics education.

Fecha de recepción: 24 de abril de 2008.

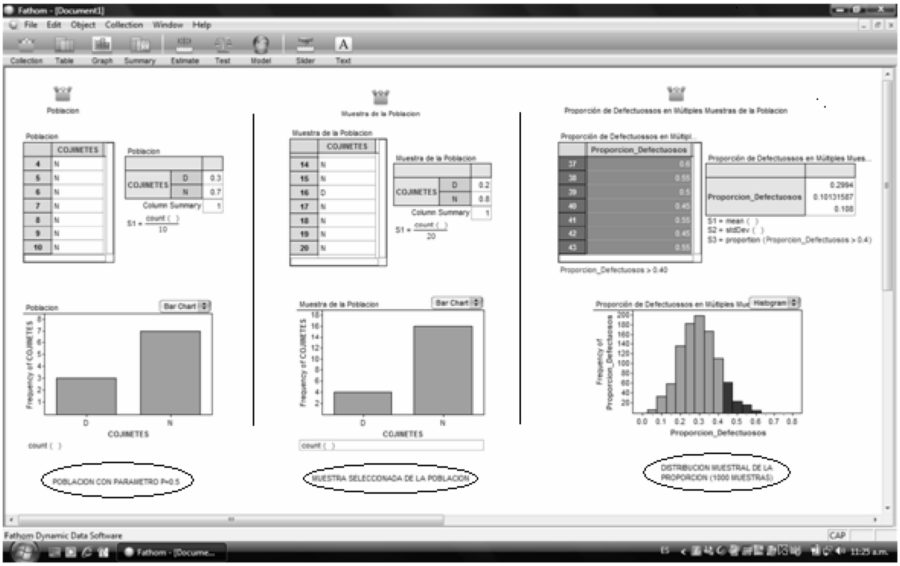
PLANTEAMIENTO DEL PROBLEMA

Una de las áreas con mayor aplicación de la estadística práctica es la inferencia estadística. Su importancia radica en que, mediante la aplicación de sus métodos, es posible obtener conclusiones significativas acerca de toda una población con base en la información que proporcionan los datos de una sola muestra o un experimento. Un concepto clave en el que se fundamentan los métodos de inferencia estadística son las distribuciones muestrales, ya que representan el valor que puede tomar un estadístico (por ejemplo, la media o la proporción) en cada una de las muestras aleatorias de un tamaño dado que se pueden extraer de una misma población.

La experiencia de muchos profesores y diversos resultados de investigación en educación estadística señalan que las distribuciones muestrales constituyen un concepto notoriamente difícil para muchos estudiantes. Por ejemplo, Chance, DelMas y Garfield (2004) consideran que la dificultad de las distribuciones muestrales radica en que requieren integración y combinación de muchas ideas de estadística, tales como distribución, muestra, población, variabilidad y muestreo. Lipson (2002) asocia la dificultad a la idea de muestra, al proceso dinámico de muestreo y a la diversidad de representaciones matemáticas y simbólicas del concepto. Por su parte, Moore (1992) asocia las dificultades de comprensión de la inferencia estadística y las distribuciones muestrales a la enseñanza tradicional que ha prevalecido hasta ahora, donde se utiliza un enfoque basado en probabilidad, la cual es reconocida como una de las áreas más complejas de las matemáticas.

Ante las dificultades que entraña el enfoque formal basado en variables aleatorias y distribuciones teóricas de probabilidad para los estudiantes con poca formación matemática, muchos educadores estadísticos (Shaughnessy, 1992; Gordon y Gordon, 1992; Thisted y Velleman, 1992; Mills, 2002; Sánchez e Inzunsa, 2006) han sugerido un acercamiento empírico basado en simulación computacional. En este enfoque, la distribución muestral de un estadístico se obtiene seleccionando repetidamente muestras de un tamaño dado de una misma población, calculando en cada una de ellas el valor del estadístico y acumulándolo para formar la distribución. Este enfoque está conceptualmente más relacionado con el proceso real de inferencia y requiere pocos antecedentes matemáticos y probabilísticos por parte de los estudiantes (Lipson, 2002; y Moore, 1990).

En el proceso anterior, el software con el que se desarrolla el proceso de simulación desempeña un papel muy importante. En el presente estudio hemos utilizado Fathom (Finzer *et al.*, 2002), un software que ha sido diseñado con propósitos específicos de enseñanza, el cual le permite al propio usuario construir la distribución muestral y relacionar durante el proceso de simulación los conceptos involucrados a través de diversas representaciones (gráficas, tablas y fórmulas) ligadas entre sí. Su naturaleza dinámica permite que el muestreo sea observado conforme se va desarrollando y que el usuario pueda observar cómo cambia una de las representaciones cuando un dato o parámetro es modificado en otra representación. El proceso completo de simulación de una distribución muestral en Fathom se puede observar en la siguiente imagen:



En el proceso se pueden identificar tres etapas o niveles importantes: modelación de la población, construcción de la distribución muestral (selección de muestras y cálculo de estadísticos) y seccionamiento de la distribución muestral para calcular probabilidades de valores muestrales. En cada etapa se involucran diferentes conceptos. En la primer etapa, se involucra la población y se calculan medidas que la describen (parámetros); en la segunda etapa, se visualiza la distribución de una muestra y sus medidas de centro y dispersión, mientras que, en la tercera etapa, se puede visualizar la distribución muestral seccionada con los estadísticos obtenidos en cada una de las muestras seleccionadas y las medidas de centro y dispersión.

Consideramos que las características del software y la manera como se utiliza en las actividades de enseñanza pueden tener una influencia importante en la construcción de significados por parte de los estudiantes. Sin embargo, esta línea requiere aún de mucha investigación que ayude a aclarar los puntos más problemáticos del aprendizaje de las distribuciones muestrales y su relación con la inferencia estadística. Nuestra investigación está orientada en este sentido.

Específicamente, en el presente trabajo nos hemos planteado las siguientes preguntas: ¿cuáles son los elementos de significado que los estudiantes atribuyen a las distribuciones muestrales y conceptos previos a la inferencia estadística como resultado de una enseñanza basada en un ambiente de simulación computacional? y ¿cómo evolucionan dichos significados cuando trabajan en este tipo de ambientes?

ANTECEDENTES

Actualmente han empezado a surgir algunos resultados de investigación donde se ha utilizado simulación computacional en el estudio de las distribuciones muestrales y la inferencia estadística. Dichos resultados empiezan a dar luz sobre la complejidad del concepto en este tipo de ambientes y las ventajas que éstos aportan respecto al enfoque tradicional basado en probabilidad. Por ejemplo, DelMas, Garfield y Chance (1999) realizaron una investigación con estudiantes universitarios sobre el papel que desempeña la simulación computacional en la comprensión de las distribuciones muestrales y su relación con el teorema del límite central. Los resultados indican que, en algunas versiones de la investigación, se obtuvieron resultados positivos; sin embargo, un número significativo de estudiantes no logró comprender las implicaciones básicas del teorema del límite central, a pesar de que las actividades estaban dirigidas a ello. Los autores concluyen que una presentación sencilla y clara no conduce necesariamente a una sólida comprensión de las distribuciones muestrales.

Por su parte, Lipson (2000) llevó a cabo una investigación con estudiantes de posgrado acerca del papel que representa un ambiente de simulación computacional en la formación de esquemas sobre las distribuciones muestrales y su impacto en la comprensión de la inferencia estadística. Los resultados señalan que un alto porcentaje de estudiantes tuvieron claras muchas de las proposiciones del esquema adecuado de las distribuciones muestrales, como es el caso de la relación entre poblaciones y muestras, el parámetro expresa las características de una población, las distribuciones muestrales son descritas por estadísticos y un estadístico es una variable. Sin embargo, la mayoría de ellos falló en varias proposiciones que se consideran fundamentales, como la relación entre el tamaño de muestra y la variabilidad de una distribución muestral, la distribución muestral está centrada en el parámetro y la distribución muestral de un estadístico puede ser modelada por una distribución conocida de probabilidad.

También interesados en la conexión entre distribuciones muestrales y el proceso de inferencia en un ambiente de simulación computacional, Saldanha y Thompson (2002) realizaron una investigación con estudiantes de bachillerato. Los investigadores identifican que un esquema adecuado en el proceso de muestreo para construir una distribución muestral e interpretar los resultados de la simulación involucra tres niveles: selección de las muestras de la población y registro del estadístico de interés, repetición del muestreo un gran número de veces y acumulación de los valores del estadístico para formar una colección, y sección de la colección para determinar qué proporción de estadísticos se encuentra entre dos valores dados. Los resultados obtenidos señalan que la mayoría de los estudiantes tuvo dificultad para desarrollar dicho esquema.

Por su parte, Meletiou-Mavrotheris (2004) desarrolló un experimento de enseñanza con el propósito de explorar el potencial de Fathom para introducir a los estudiantes en la inferencia estadística. Los resultados indican que el uso del software condujo a los estudiantes a la construcción de un modelo mental bastante coherente de las distribuciones muestrales y otros conceptos relacionados con la inferencia estadística. Ella sugiere que un enfoque informal de la inferencia estadística, empleando ambientes con las características de Fathom, puede ser más efectivo que ambientes convencionales que promueven simulaciones del tipo "caja negra", donde los estudiantes simplemente observan cómo la computadora construye la distribución muestral sin permitirles conexiones directas entre lo formal y lo informal.

MARCO TEÓRICO

El marco teórico que hemos utilizado consta de dos componentes. La primera tiene que ver con una concepción acerca del significado y la comprensión de los objetos matemáticos y estadísticos propuesta por Godino y Batanero (1994; 1998) y Godino (2002); la segunda consiste en un enfoque mediante el cual las computadoras se consideran como herramientas cognitivas con potencial para generar cambios en la actividad mental de los estudiantes cuando exploran conceptos y resuelven problemas matemáticos (Dörfler, 1993; Pea, 1987).

La teoría de los significados propuesta por Godino y Batanero (1994, 1998) y Godino (2002) se caracteriza por la adopción de dos enfoques en la definición de sus elementos teóricos:

1. Un *enfoque antropológico* en el estudio de la cognición matemática, en el cual se parte de que las personas trabajan e interactúan dentro de diferentes grupos o instituciones donde las matemáticas son objeto de estudio y, por lo tanto, el significado no es único y está mediado por las características del contexto donde se abordan los objetos matemáticos.
2. Un *enfoque pragmático* en el estudio de los objetos matemáticos, que implica que el significado de un objeto matemático está en función del sistema de prácticas o uso que se hace de él para resolver situaciones-problema.

El enfoque antropológico conduce a tener en cuenta dos dimensiones interdependientes del significado de los objetos matemáticos: *significado institucional* y *significado personal*. El significado institucional es el significado que un objeto matemático tiene para un determinado grupo de profesionales interesados en las situaciones-problema de donde éste emerge; mientras que el significado personal es el significado que una persona (por ejemplo un estudiante) le atribuye al objeto matemático.

En nuestro caso, el objeto de interés ha consistido en las distribuciones muestrales en un curso introductorio de estadística de nivel universitario. A partir de una muestra representativa de libros de estadística, hemos identificado el sistema de prácticas prototípicas que utiliza la institución de los educadores estadísticos en este nivel para resolver problemas donde surge este concepto; es decir, el significado institucional. Dicho significado nos ha servido para diseñar la secuencia de actividades que hemos presentado a los estudiantes; además, es importante para contrastar los significados personales que han construido durante las actividades con simulación.

Por su parte, el enfoque pragmático permite definir el significado de un objeto matemático, como el sistema de prácticas (institucionales o personales) que se utilizan para resolver un campo específico de problemas en un determinado contexto. En dicho sistema de prácticas, intervienen diversos elementos que se ponen en juego de manera interrelacionada en la actividad matemática involucrada en la resolución de situaciones-problema. Para dar cuenta de dicho sistema de prácticas y describir la actividad matemática realizada, en Godino (2002) se propone una tipología de cinco elementos, de acuerdo con las funciones que desempeñan en el trabajo matemático. Esta tipología constituye los *elementos de significado* que permiten establecer tanto el significado personal como institucional de los objetos matemáticos. Dichos elementos así como sus funciones específicas son:

1. *Problemas y situaciones (elementos fenomenológicos).* Se refiere al campo de problemas o situaciones de las que surge el concepto en estudio.
2. *Representaciones y lenguaje (elementos representacionales).* Es cualquier representación verbal o escrita que se utiliza para referirse o representar los conceptos y propiedades que intervienen en un problema.
3. *Acciones y procedimientos (elementos procedimentales).* Son los procedimientos o estrategias que se utilizan para resolver los problemas o situaciones presentadas.
4. *Conceptos y propiedades (elementos conceptuales).* Son los conceptos, propiedades y sus relaciones con otros conceptos que se ponen en juego cuando se resuelve un problema.
5. *Validaciones y argumentaciones (elementos argumentativos).* Los argumentos o validaciones que se utilizan para obtener un resultado, comunicarlo a otros o convencerse a uno mismo de su validez.

La comprensión de un objeto matemático desde esta perspectiva consiste en un proceso continuo y progresivo donde los estudiantes adquieren y relacionan los distintos elementos que componen el significado del concepto.

La otra componente que hemos utilizado tiene que ver con la manera como puede usarse la computadora en el proceso de aprendizaje de las matemáticas y, particularmente, en el de la estadística. Pea (1987) sostiene que la inteligencia no es una cualidad de la mente sola, sino un producto de la relación entre estructuras mentales y herramientas del intelecto proporcionados por la cultura, como es el caso de las computadoras. Para Pea (1987, p. 91), "una herramienta cognitiva es cualquier medio que ayuda a trascender las limitaciones de la mente en el pensamiento, el aprendizaje y las actividades de resolución de problemas". En particular, en el caso de las computadoras constituyen una extraordinaria y potente herramienta cognitiva para aprender a pensar matemáticamente; con ellas se pueden operar, no sólo números, sino también símbolos; permiten, además, almacenar y manipular símbolos dinámicamente y permiten interacciones con los usuarios en tiempo real.

Apoyado en estas ideas, Dörfler (1993) propone un marco conceptual sobre el uso de la computadora en la educación matemática e identifica algunas maneras en las que una herramienta computacional puede generar cambios en la actividad mental de los estudiantes:

1. Cambio de las actividades a un nivel cognitivo más alto (meta-nivel).
2. Cambio de objetos con los que se realizan las actividades.
3. Enfoca las actividades en transformación y análisis de representaciones.
4. Apoya la cognición situada y resolución de problemas.

METODOLOGÍA

El tipo de investigación que más se adapta para responder las preguntas planteadas es una investigación de tipo cualitativo, en cuanto que lo que interesa es interpretar, describir y comprender los significados que estudiantes universitarios atribuyen a las distribuciones muestrales en un ambiente de simulación computacional.

PARTICIPANTES

El estudio se llevó a cabo con 11 estudiantes voluntarios (19-20 años) que tomaban un curso de Inferencia Estadística en la carrera de Ingeniería en Informática del Instituto Politécnico Nacional (México), los cuales, al inicio del estudio, ya habían estudiado el tema de distribuciones muestrales desde una perspectiva tradicional, utilizando fórmulas y tablas de probabilidad. Las actividades se desarrollaron en un aula donde cada estudiante disponía de una computadora y el investigador no era su profesor de la clase.

INSTRUMENTOS

El estudio consistió en nueve actividades, las cuales se desarrollaron durante 20 sesiones de hora y media (que incluyen sesiones de familiarización con el software, aplicación de cuestionarios y entrevistas). Entre las distribuciones poblacionales consideradas en las actividades se encontraban distribuciones binomiales, uniformes discretas, normales e irregulares discretas. Todos los problemas correspondían a distribuciones muestrales de medias y proporciones, dos estadísticos muy importantes que cubren una amplia gama de aplicaciones en los libros de estadística y a los que con frecuencia se reduce la enseñanza de las distribuciones muestrales en los primeros cursos universitarios.

Antes de iniciar las actividades con la computadora, los estudiantes respondieron un examen de distribuciones muestrales y un cuestionario para conocer sus significados acerca de los conceptos investigados. Durante el estudio recibieron hojas de trabajo para cada una de las actividades. Cada actividad fue guardada en un archivo de computadora, lo cual fue complementado con videograbación de varias sesiones y un diario de observación que se llevó durante el transcurso del estudio. Al final de las actividades con la computadora, se les suministró de nuevo un cuestionario para explorar sus significados sobre los conceptos involucrados y conocer el efecto de las actividades con el software. Algunos estudiantes fueron entrevistados en torno a los mismos conceptos con el propósito de profundizar en los significados construidos.

Procedimientos

Como fundamento para el estudio de los significados, diseñamos un proceso de aprendizaje en el que se planteó que fuera el mismo estudiante con su acción sobre los problemas propuestos quien construyera significados sobre los conceptos involucrados. De ahí que el investigador únicamente sirviera de guía y proporcionara ayuda sólo cuando era estrictamente necesario para continuar con las actividades. Cada problema involucraba diferente distribución poblacional y la manera de abordarlos consistió en construir distribuciones para muestras de diferente tamaño, con un incremento gradual en el número de muestras seleccionadas para posibilitar que los estudiantes descubrieran propiedades y construyeran conceptos como la variabilidad muestral, las implicaciones del teorema del límite central en el comportamiento de las distribuciones y el cálculo de probabilidades de valores muestrales.

En el análisis de los datos, hemos tenido en cuenta primeramente las cinco componentes de significado de un objeto matemático que se establecen en Godino (2002), así como las diferentes etapas que se han identificado en el proceso de simulación con Fathom, en las cuales tienen lugar diferentes elementos de significado que nos interesa resaltar. Dichas etapas son:

a) Modelación de la población.
b) Construcción de la distribución muestral.
c) Seccionamiento de la distribución muestral para calcular probabilidades de valores muestrales.

RESULTADOS Y ANÁLISIS

Situaciones-problema (elementos fenomenológicos)

Las situaciones planteadas en el estudio pertenecen a un campo de problemas de las distribuciones muestrales en el cual se busca calcular probabilidades de algunos valores muestrales mediante un proceso deductivo, donde se supone conocida la distribución de la población y sus parámetros. Este tipo de problemas es muy común en los libros antes del estudio de la inferencia estadística. Un ejemplo representativo del tipo de situaciones que se abordaron es el siguiente:

En una fábrica de cojinetes para automóvil se ha desajustado una máquina y 30% de su producción está saliendo defectuosa.

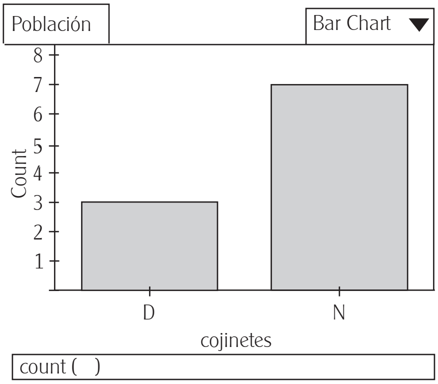
a) Si se toma una muestra de tamaño 80, ¿cuántos cojinetes defectuosos esperarías en la muestra?
b) ¿Cuál es la probabilidad de que en una muestra de tamaño 80 existan 30 o más defectuosos?
c) ¿Cuál es la probabilidad de que aparezcan 20 o menos defectuosos en la muestra de tamaño 80?

En los textos de estadística, estos problemas se resuelven por lo general mediante la estandarización de la distribución real y con el apoyo de tablas de probabilidad. En nuestro caso, los problemas se resolvieron mediante la técnica de simulación con Fathom, en la cual se requieren acciones diferentes por parte de los estudiantes, como la formulación de un modelo poblacional, la extracción de muestras, y el cálculo y acumulación de los resultados muestrales. No se requiere estandarización y las probabilidades se calculan de manera frecuencial a partir de la colección de resultados que comprenden la distribución muestral empírica.

Figura 1 Población con variable discreta

Población

	cojinetes
1	D
2	D
3	D
4	N
5	N
6	N
7	N
8	N
9	N
10	N



Representaciones y lenguaje (elementos representacionales)

El tipo de representaciones (gráficas, numéricas y simbólicas) que proporciona el software, así como la manera en que se relacionan entre sí, fueron elementos clave que tuvieron implicaciones en los demás componentes de significado (acciones, conceptos, propiedades y argumentaciones) que los estudiantes construyeron a lo largo de las actividades, desde la formulación del modelo hasta el cálculo de probabilidades.

Por ejemplo, un factor importante para tener éxito en la formulación del modelo está ligado a la selección de la representación adecuada (véanse las figuras 1 y 2). Si la población es discreta, los elementos hipotéticos se escriben en una tabla de casos que reproduzca el valor del parámetro; si la población es continua, los elementos se generan mediante una fórmula. Sin embargo, estos criterios no fueron comprendidos por muchos estudiantes hasta las actividades finales del estudio.

Un ejemplo de ello lo tomamos de la primera actividad introductoria (población con distribución binomial, $p = 0.5$), en la cual se pedía simular muestras de 10 lanzamientos de una moneda y calcular las proporciones de águilas en cada muestra (véase la figura 3). En los primeros dos modelos, se observa que los estudiantes no identifican los elementos de la población, en cambio, se introducen posibles proporciones de águilas que podrían suceder en el muestreo. Estos modelos estuvieron influidos por la simulación física que realizaron los estudiantes con una moneda antes de las actividades, en la cual se pedía realizar 10 conjuntos

Figura 2 Población con variable continua

Población

	refrescos
=	randomNormal (200, 5)
794	195.086
795	204.824
796	198.312
797	200.198
798	196.317
799	197.383
800	207.547
801	198.605

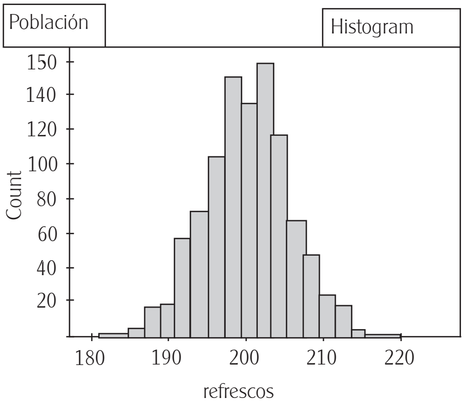


Figura 3 Modelos poblacionales construidos en la primera actividad

Collection

	Moneda
1	0.5
2	0.5
3	0.5
4	0.5
5	0.5
6	0.5
7	0.5
8	0.5
9	0.5
10	0.5

Collection

	Monedas
1	0.4
2	0.5
3	0.3
4	0.5
5	0.6
6	0.7
7	0.5
8	0.6
9	0.4
10	0.5

Volado

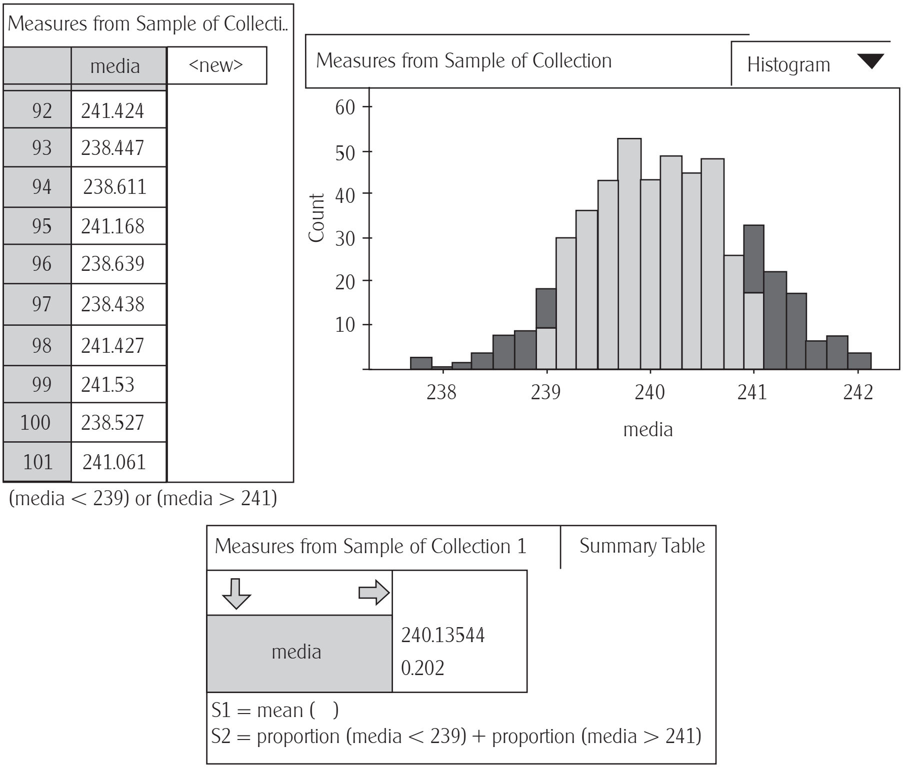
	Aguilas
1	águila
2	sol

Moneda

	Tabla
1	s
2	s
3	s
4	s
5	s
6	A
7	A
8	A
9	A
10	A

de 10 lanzamientos. Por su parte, los dos últimos modelos son correctos, ya que ambos reproducen el valor del parámetro, aunque con diferente cantidad de elementos. En posteriores actividades, se observó que muchos estudiantes seguían teniendo problemas para identificar los criterios correctos para la formulación del modelo, no obstante que en algunos casos se trataba de poblaciones similares pero con diferentes parámetros.

Figura 4 Conexión entre diversas representaciones de una distribución muestral



Una vez construido el modelo de la población, fue práctica frecuente calcular sus medidas descriptivas (media y desviación estándar) en una tabla resumen, así como su representación gráfica. Estas representaciones sirvieron de referencia para que los estudiantes establecieran relaciones con las medidas descriptivas de las muestras que seleccionaban, lo cual les permitió observar la variabilidad muestral alrededor del parámetro.

La representación gráfica de la distribución muestral más utilizada fue el histograma, el cual se forma a partir de la tabla con estadísticos acumulados. Adicionalmente, se puede utilizar una fórmula para filtrar muestras que cumplen con ciertos valores y reflejarlos en la representación gráfica (véase la figura 4). De esta manera, se conectan las tres representaciones y se visualizan en el sombreado de una región de la distribución. A pesar de la importancia de esta conexión entre representaciones para darle sentido al cálculo de probabilidades, sólo un par de estudiantes logró establecerlas al final del estudio, cuando ya se habían apropiado de la herramienta computacional.

En suma, durante todo el proceso de simulación, los estudiantes pudieron establecer relaciones entre los diversos conceptos involucrados, como son población, muestra, variabilidad muestral y el efecto de tamaño de muestra, el comportamiento de la distribución muestral y las probabilidades de resultados muestrales, mediante las diversas representaciones, como histogramas, tablas de casos, tablas con medidas descriptivas, fórmulas para calcular estadísticos y proporciones de resultados muestrales y realizando ajustes de una distribución teórica con una distribución empírica.

Esta conexión entre representaciones fue importante para que muchos estudiantes construyeran significados correctos de diversos conceptos involucrados. Un ejemplo de ello es el caso de Omar, quien en el examen diagnóstico mostró dificultades en la comprensión de la variabilidad muestral. En una entrevista final respondió a los cuestionamientos del investigador en el contexto de la siguiente actividad: en una fábrica de cojinetes para automóvil se ha desajustado una máquina y 30% de su producción está saliendo defectuosa.

INVESTIGADOR: Toma una muestra de tamaño 10, repite el proceso y observa lo que sucede con la proporción de defectuosos.

OMAR: No se aleja mucho de 3.

INVESTIGADOR: ¿Cómo esperarías que variara esa proporción?

OMAR: Que no fuera mucho mayor que 3 ni mucho menor que 3.

INVESTIGADOR: Proporciona límites entre los cuales podrían estar los resultados.

OMAR: Pues casi siempre va a salir entre 2 y 4, más o menos.

INVESTIGADOR: Ahora observa cómo es la proporción de defectuosos cada vez que tomas una muestra y compárala con la proporción de defectuosos en la población. ¿Hay alguna relación entre ambas proporciones?

OMAR: Pues de cierto modo, porque la proporción poblacional es la proporción real y la muestral no sale exactamente igual, porque no estamos tomando toda la población, simplemente es una muestra. Pues, obviamente la de la muestra va a ser diferente a la de la población.

INVESTIGADOR: Pero ¿qué tan diferente?

OMAR: No mucho, pero entre mayor sea la muestra más se va acercar a la real.

En las respuestas de Omar se puede observar que ha construido una idea correcta de la variabilidad de la proporción alrededor del parámetro poblacional y de la representatividad de una muestra. Esto sin duda es un reflejo de las situaciones que trabajó en el ambiente de simulación y estadística dinámica.

Acciones y procedimientos (elementos procedimentales)

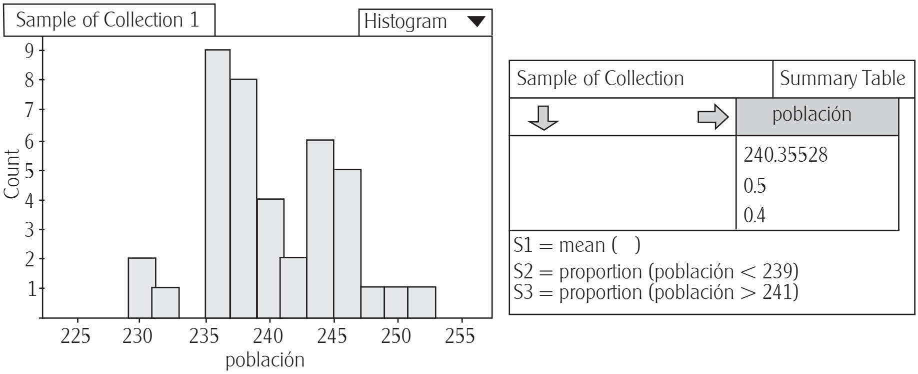
En el diseño de las actividades y las características del software, desde luego mediaron las acciones y estrategias de los estudiantes. En la parte correspondiente al modelo poblacional, sus estrategias de formulación no siempre estuvieron bien dirigidas, ya que muchos estudiantes no tenían claro que era necesario identificar el tipo de variable y el valor del parámetro para una correcta formulación; como ejemplo de ello podemos mencionar que en una de las actividades que trataba de una población binomial con 30% de cojinetes defectuosos, la cual se describió en párrafos anteriores, muchos estudiantes consideraron equiprobables ambos resultados, mientras que en otra actividad con variable continua intentaron modelar la población como si fuera discreta.

En la etapa de la selección de muestras se observaron acciones por parte de algunos estudiantes que dificultaban la construcción de la distribución muestral. Por ejemplo, cuando se pedía seleccionar 1 000 muestras de tamaño 10, seleccionaban 1 muestra de tamaño 1 000. Esta confusión también ha sido identificada por Saldanha y Thompson (2002) y se relaciona con la dificultad de pasar de una muestra individual a una distribución de muestras. Este paso es considerado muy importante para adquirir un buen razonamiento e interpretación de los resultados de distribuciones muestrales.

Se observó otra confusión importante relacionada con la idea de muestra individual en la siguiente actividad: una máquina de refrescos está diseñada para que la cantidad de bebida que sirve promedie 240 ml con una desviación estándar de 5 ml. La máquina se verifica periódicamente, para ello, se toma una muestra de 40 bebidas y se calcula el contenido promedio (media). Si la máquina recibe mantenimiento cuando el contenido promedio de la muestra es menor que 239 ml o mayor que 241 ml, ¿qué proporción de veces recibirá mantenimiento la máquina?

Una pareja de estudiantes calcularon la probabilidad de un resultado muestral utilizando una muestra particular en lugar de utilizar la distribución que contenía las medias de 1 000 muestras (véase la figura 5). Es decir, confundieron los resultados de una muestra particular con la distribución muestral, por la simi-

Figura 5 Cálculo de probabilidades utilizando los resultados de una muestra



litud que en este caso existía entre ambas, lo cual implica que calcularon la proporción de datos de la muestra que se encontraban entre ciertos límites, en lugar de calcular la proporción de muestras.

En este mismo sentido, en otra actividad que fue videograbada, se observó que otra pareja de estudiantes utilizaron la última muestra de una distribución de 1 000 muestras seleccionadas de una población binomial con un 30% de cojinetes defectuosos, en lugar de utilizar la distribución muestral. Cuando el investigador los cuestiona sobre lo anterior, señalan que el enunciado "calcular la probabilidad de que en una muestra" los hizo proceder de esa manera. Un extracto de la conversación con los estudiantes se muestra a continuación:

Libnia: No hay ninguna probabilidad de que existan más de 30 defectuosos en una muestra de tamaño 80, pues son 24 en total. [Está observando el resultado de la última muestra.]

Edgar: Falta considerar las 1 000 muestras de tamaño 80. [Acertadamente opina Edgar.]

Libnia: Me confundo, aquí se me pide una muestra de tamaño 80. Ésta es una muestra de tamaño 80.

Edgar: Exacto.

Libnia: Si yo trabajo sobre las 1 000 muestras de tamaño 80, no estoy contestando lo que se me pide, porque aquí dice que trabaje sobre la muestra de tamaño 80 y la muestra de tamaño 80 es ésta.

Conceptos y propiedades (elementos conceptuales)

Los principales elementos conceptuales que nos propusimos investigar y que se tuvieron en cuenta en el diseño de las actividades son: variabilidad muestral, las implicaciones del teorema del límite central en las distribuciones muestrales y el cálculo de probabilidades de resultados muestrales. Para el estudio de las implicaciones del teorema del límite central, se hizo énfasis en la construcción de distribuciones para diferentes tamaños de muestra sobre una misma gráfica, con la idea de facilitar la comparación entre ellas.

En las primeras actividades observamos que no era sencillo que los estudiantes tuvieran una comprensión inmediata de sus implicaciones del teorema con sólo variar la población y el tamaño de muestra. Si bien algunos estudiantes señalaban algunos rasgos de comportamiento, otros centraban su atención en situaciones triviales, como el grosor de las barras de los histogramas o en los decimales de las medidas descriptivas. Hubo necesidad entonces de rediseñar preguntas en las siguientes actividades para enfocar más la atención en las propiedades de forma, centro y distribución de las distribuciones muestrales. Veamos a continuación las respuestas en las actividades 4 y 5 (distribuciones con forma irregular) de dos estudiantes que no habían apreciado correctamente las propiedades en las actividades anteriores.

> Realmente el centro no es muy variable, está entre 17.2 y 17.4 sin importar el tamaño de la muestra, pero entre mayor sea el tamaño de la muestra, menor será la dispersión [...] Entre más se incremente el tamaño de muestra la distribución toma la forma de distribución normal [Viridiana].
>
> No importa el tamaño de la muestra, la media es casi igual, sólo es diferente por decimales, en cambio la desviación sí afecta, entre mayor sea el número, más pequeña es la desviación [...] El centro, tanto en las muestras como en la población, es igual con diferencia en decimales, pero la dispersión sí varía e influye el tamaño de muestra [...] Cuando la muestra se incrementa, hay un mayor ajuste en la curva de la distribución normal [Ana Lilia].
>
> Si redondeamos el centro de la población, es igual al centro de cada una de las distribuciones, mientras que la dispersión tiende a disminuir cada vez que aumenta el tamaño de la muestra [...] Las barras se adelgazan mientras el tamaño de la muestra aumenta [Jorge].

En la deducción de las propiedades, la mayoría de los estudiantes utilizaron más las representaciones numéricas (tabla resumen con medidas descriptivas) que las representaciones gráficas. Es decir, centraban su atención en los valores de la media y la desviación estándar de la población y en la media y desviación estándar (error estándar) de las distribuciones muestrales y las relacionaban con el tamaño de muestra.

Al final de las actividades, entrevistamos a Mónica para conocer el significado que había desarrollado sobre las implicaciones del límite central. En la entrevista, ella dispuso de una representación gráfica (véanse las figuras 6 y 7) que fue clave tanto para entender las propiedades de las distribuciones muestrales, como para interpretar el efecto del tamaño de muestra en las probabilidades.

INVESTIGADOR: ¿Tú crees que haya alguna relación entre la población y la distribución muestral?

MÓNICA: Bueno, aquí en la población son los valores del dado y aquí son las medias de los valores de los dados.

INVESTIGADOR: ¿Entonces, no tienen por qué parecerse?

MÓNICA: Yo digo que no. En la población son valores originales y en la distribución muestral son valores ya procesados.

INVESTIGADOR: Si en lugar de esta forma uniforme de la población, se tuviera una población como columpio, ¿crees que eso repercutiría en la forma de la distribución muestral?

MÓNICA: No, por lo regular la forma de la distribución muestral siempre nos da normal.

INVESTIGADOR: ¿Independientemente de la forma de la población?

MÓNICA: Sí.

INVESTIGADOR: ¿Qué pasa con la media de la distribución muestral comparada con la media de la población?

MÓNICA: La media por lo general siempre se acerca a la media poblacional.

INVESTIGADOR: ¿Qué pasa con la desviación estándar?

MÓNICA: La desviación estándar es la que sí varía, se hace más pequeña, conforme aumenta el tamaño de muestra.

Para calcular e interpretar las probabilidades, Mónica vinculó una tabla con los valores de las medias con el histograma y utilizó la opción de filtro para sombrear los resultados de interés (véase la figura 7).

Figura 6 Distribuciones muestrales para $n = 5$ y $n = 15$

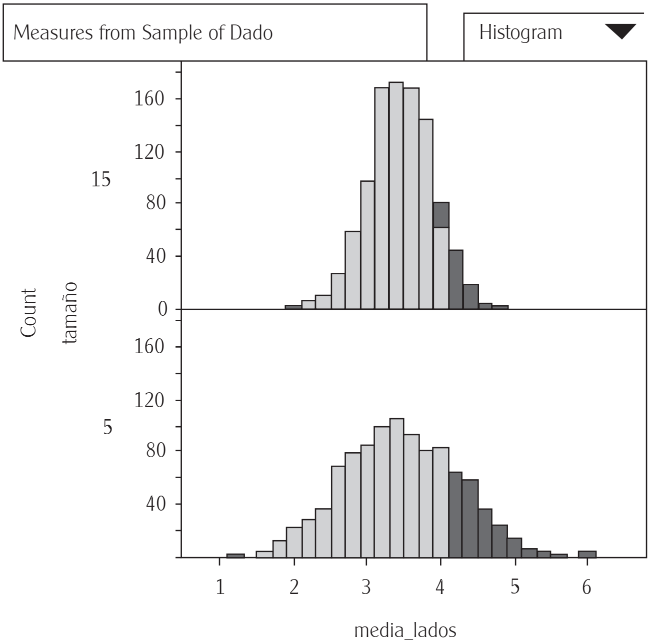
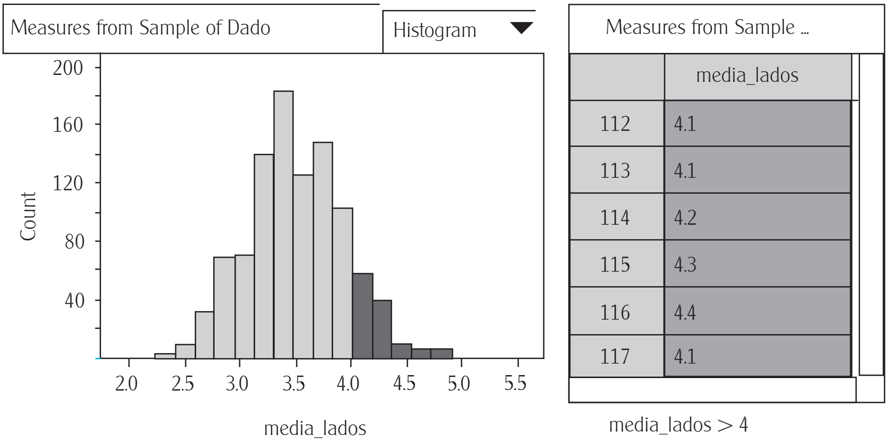


Figura 7 Distribución muestral para $n = 30$



INVESTIGADOR: ¿En cual de las dos distribuciones es más probable que aparezca una media de 4 o más?
MÓNICA: En la de tamaño 5.
INVESTIGADOR: ¿Por qué?
MÓNICA: Porque tenemos muchos más valores hacia la derecha de 4.
INVESTIGADOR: Si se tratara de hacer apuestas, ¿a qué valores le apostarías?
MÓNICA: Entre 3 y 4.
INVESTIGADOR: ¿Y si quisieras aumentar tu probabilidad de ganar?
MÓNICA: Entonces elegiría entre el 2 y el 5.

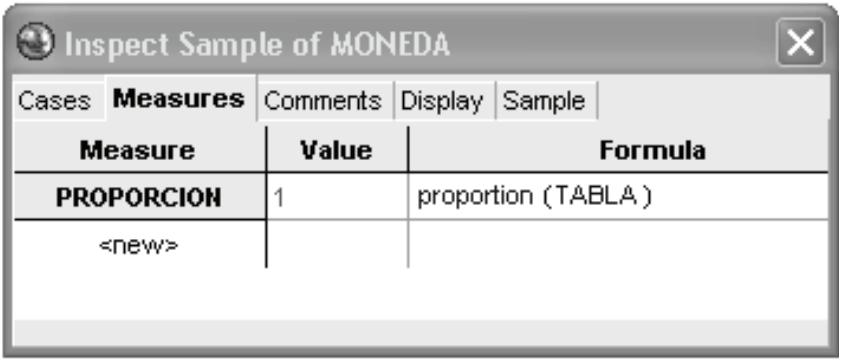
Como vemos, en el ambiente computacional, Mónica pudo elaborar nociones fundamentales concernientes a la teoría de las distribuciones muestrales, a saber, la variabilidad muestral, la dependencia de ésta en relación con el tamaño de muestra, la convergencia de la distribución muestral hacia una distribución normal y el efecto del tamaño de muestra en la variabilidad de las distribuciones y en la probabilidad de resultados muestrales.

ARGUMENTACIONES (ELEMENTOS ARGUMENTATIVOS)

La conexión entre representaciones gráficas, numéricas y simbólicas fue un elemento importante en la producción de argumentos, como recurso de control y validación del proceso de solución. Por ejemplo, los estudiantes pudieron establecer relaciones sobre el efecto del tamaño de muestra en el comportamiento de las distribuciones muestrales mediante la comparación de medidas descriptivas y visualizando la forma de las gráficas. De igual manera, pudieron explorar la variabilidad muestral tomando muestras una a una y observando cómo cambiaban sus medidas descriptivas. Por su parte, las representaciones simbólicas proporcionadas en el inspector y el editor de fórmulas fueron utilizadas como recurso de advertencia cuando los estudiantes cometían errores de sintaxis. Por ejemplo, en la actividad 1, Ana Lilia cometió un error al introducir la fórmula de la proporción (véase la figura 8) y se percató cuando tomó 1 000 muestras y todas le dieron 1 como resultado.

Otro recurso de validación que se utilizó en las últimas actividades fue la superposición de la distribución teórica con la distribución empírica. En dicho proceso, la mayoría de los estudiantes tuvieron dificultades para identificar los parámetros de la distribución teórica e introducirlos correctamente en el editor

Figura 8 Inspector de fórmulas como recurso de validación



Inspect Sample of MONEDA

Cases | Measures | Comments | Display | Sample

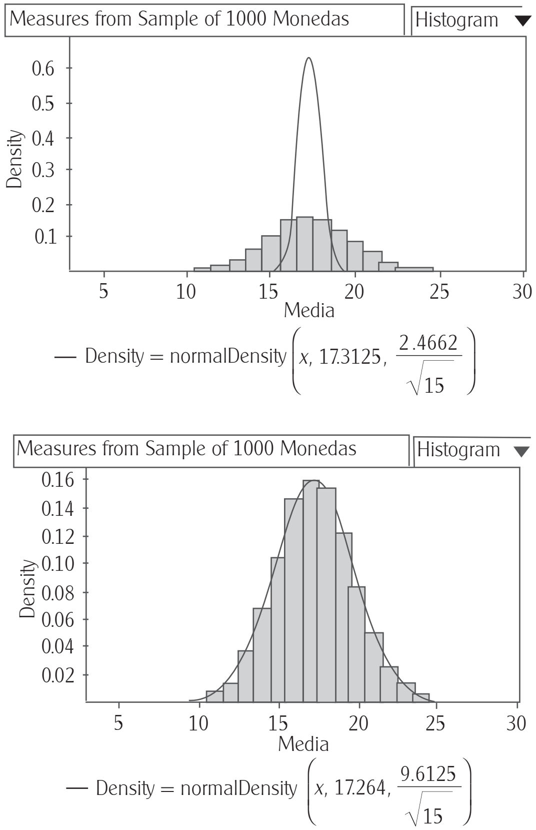
Measure	Value	Formula
PROPORCION	1	proportion (TABLA)
<new>		

Measures from Sample of MONEDA 10		
	PROPORCIÓN 1	<new>
992	1	
993	1	
994	1	
995	1	
996	1	
997	1	
998	1	
999	1	
1000	1	

de fórmulas. Por ejemplo, en el caso de la distribución muestral de medias, se tiene que: $\mu_{\bar{x}} = \mu$ y $\sigma_{\bar{x}} = \dfrac{\sigma}{\sqrt{n}}$.

Un ejemplo de ello lo tomamos de la actividad de Gerardo en la actividad 4 (véase la figura 9), quien introduce la desviación estándar de la distribución muestral en lugar de la desviación estándar poblacional.

Figura 9 Distribuciones teórica y empírica obtenidas por Gerardo



Análisis comparativo entre significados previos y posteriores a las actividades de simulación

Los resultados del examen previo a las actividades de simulación nos proporcionan una idea de los significados construidos por los estudiantes en la enseñanza tradicional que recibieron del tema. El análisis refleja que sus acciones principales se centraron en sustituir los datos de los problemas en la fórmula de estandarización para determinar la probabilidad de un resultado muestral mediante tablas. La mayoría de los estudiantes no utilizaron de manera sistemática las representaciones gráficas y el sombreado de regiones de la distribución, no obstante su importancia para transitar del ámbito de los datos reales a los datos estandarizados, lo cual los condujo a cometer diversos errores en el cálculo de las probabilidades. El principal elemento conceptual que se puso en juego fue el cálculo de probabilidades de resultados muestrales.

Como consecuencia de lo anterior, en el cuestionario diagnóstico –el cual se enfocó a aspectos conceptuales más que procedimentales–, se observó que los significados personales de muchos estudiantes eran superficiales y en muchos casos, erróneos. Por ejemplo, para explorar la relación entre el tamaño de la muestra y la variabilidad de las distribuciones muestrales, un ítem del cuestionario disponía de cinco distribuciones para tamaños de muestra 5, 10, 15, 20 y 25, respectivamente, las cuales fueron extraídas de una población con distribución normal. Se requería que asignaran el tamaño correspondiente a cada una de ellas e identificaran el valor de la media de la población. Sólo dos estudiantes asignaron correctamente el tamaño de muestra y cuatro identificaron el valor de la población, aunque con argumentos poco fundamentados que nos dejan ver un conocimiento superficial de los conceptos involucrados.

En resumen, los resultados del examen y el cuestionario previos muestran que, en muchos casos, el significado personal de los estudiantes no contempla importantes elementos conceptuales necesarios para una comprensión adecuada de las distribuciones muestrales, como es el caso de la apreciación de la variabilidad muestral y las implicaciones del teorema del límite central. Su lenguaje (representaciones verbales) ha sido insuficiente en algunos casos para argumentar correctamente muchos de los resultados y las representaciones gráficas no se utilizaron de manera sistemática. Su significado está ligado principalmente al uso de fórmulas y tablas de probabilidad.

Por su parte, en el cuestionario posterior se propusieron diversos ítems que nos permitieran analizar el significado construido por los estudiantes en las actividades de simulación y poder así establecer comparaciones con los significados previos. Un concentrado de las respuestas a ambos cuestionarios se muestra en el cuadro 1. Un ejemplo representativo de este cuestionario es el siguiente:

En las figuras 10 y 11 se presentan la distribución de una población y dos distribuciones muestrales de medias para muestras de tamaño 2 y 30. Coloca el tamaño de muestra en la distribución que corresponda.

El ítem fue respondido correctamente por siete estudiantes, de los cuales cinco dan argumentaciones que involucran a la variabilidad de manera correcta.

Figura 10 Distribución de la población

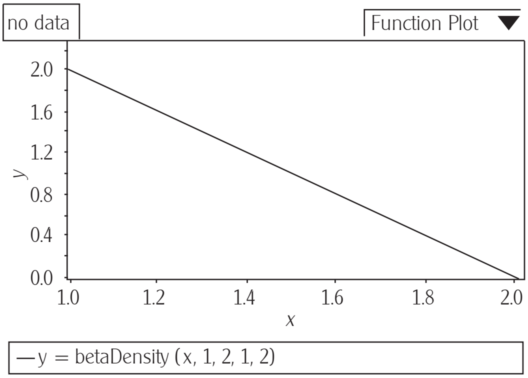
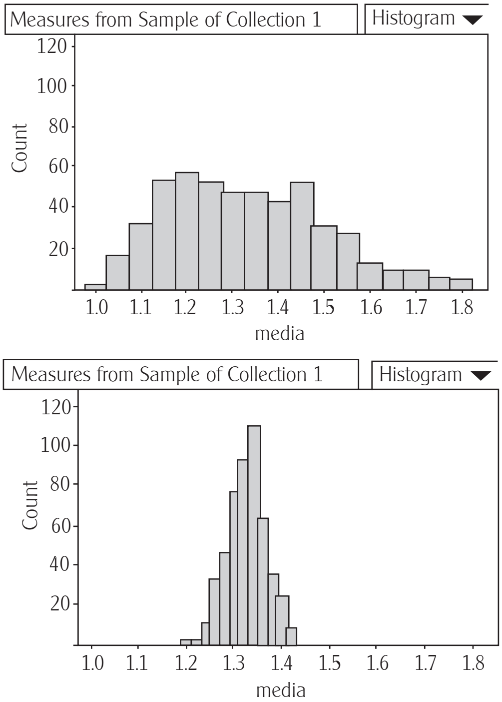


Figura 11 Distribuciones muestrales



Cuadro 1 Comparación de resultados en los cuestionarios previo y posterior

Conceptos y propiedades investigados	Previo	Posterior
Variabilidad de un conjunto de datos en una gráfica	3	5
Variabilidad muestral	5	5
Efecto del tamaño de muestra en la variabilidad de una distribución muestral	3	7
La media de la distribución muestral está centrada en la media de la población	4	9
La diferencia entre la media muestral y la media poblacional puede deberse a la variabilidad muestral	3	9
Efecto del tamaño de muestra en la probabilidad de un resultado muestral	1	8
Efecto del tamaño de muestra en la forma de la distribución muestral	2	5

Ninguno hizo referencia a la forma, que era otra propiedad a la que se podía hacer alusión, pues la muestra más grande genera una distribución más normal que la muestra pequeña. Por su parte, los cuatro estudiantes que contestaron incorrectamente argumentan que en la primera distribución hay más datos, y otros señalan la similitud con la población conforme mayor es la muestra. En diversos ítems estos estudiantes han mostrado esa idea incorrecta de que, a mayor número de datos, hay más variabilidad, además, han leído de manera incorrecta las gráficas, señalando que hay más datos en una que en otra.

Otro de los ítems con el que nos propusimos evaluar el efecto del tamaño de muestra en la probabilidad fue el siguiente: En una población de estudiantes, el peso medio es de 60 kg y una desviación estándar de 5 kg. Si se seleccionan dos muestras aleatorias de dicha población, una de tamaño 10 y otra de tamaño 100, ¿en cuál consideras que es más probable tener una media muestral menor que 55 kg?

a) En la muestra de tamaño 10.
b) En la muestra de tamaño 100.
c) En ambos muestras se tiene la misma probabilidad.

Ocho estudiantes contestaron correctamente y en sus argumentos se observa que han construido un significado correcto del efecto del tamaño de muestra en la probabilidad, la cual la relaciona con la variabilidad.

GERARDO: Porque hay más datos esparcidos y, en una de 100, por lo general los resultados se centrarían más en la media de 60.
ANA LILIA: En la muestra de 10, por tener mayor dispersión que la muestra de tamaño 100.
MÓNICA: Puesto que tiene más variabilidad en los datos, la media muestral se aleja más de la poblacional.
OMAR: Porque la dispersión es mayor, por lo tanto, los posibles resultados se pueden alejar más de la media.

DISCUSIÓN

El análisis de los diferentes elementos de significado (representaciones, acciones, conceptos, propiedades y argumentaciones) utilizados por los estudiantes en el ambiente computacional señala que muchos estudiantes construyeron correctamente diversos conceptos y nociones contemplados en el significado institucional de las distribuciones muestrales. En la construcción de dichos significados, fueron importantes tanto las características de herramienta cognitiva del software, como la manera en la cual se implementó mediante el diseño de las actividades de simulación.

Por ejemplo, la posibilidad de incrementar poco a poco la cantidad de muestras seleccionadas de una población y construir sobre una misma gráfica distribuciones muestrales para diferentes tamaños de muestra permitió a los estudiantes establecer relaciones entre diversos conceptos, como el tamaño de muestra, la forma, el centro, la variabilidad muestral de las distribuciones muestrales y la probabilidad de resultados muestrales a través de representaciones gráficas, numéricas y simbólicas entrelazadas entre sí.

Sus acciones se orientaron a la construcción por ellos mismos de las distribuciones muestrales, estableciendo el modelo de la población, extrayendo muestras, calculando y acumulando estadísticos para, finalmente, calcular probabilidades de resultados muestrales. Este proceso se pudo observar a través de la transformación y análisis de representaciones que frecuentemente realizaron los estudiantes. Superponer una distribución teórica a una distribución empírica, sombrear áreas correspondientes a la probabilidad calculada, comunicarse con la computadora mediante la introducción de una fórmula y recibir retroalimentación de los posibles errores son ejemplos de recursos de validación y generación de argumentos que muestran el nivel cognitivo en el que se involucraron los estudiantes, el cual no está presente en las actividades que realizan en un ambiente tradicional, donde la actividad se resume al empleo y sustitución de fórmulas y al manejo de tablas de probabilidad, principalmente.

Los objetos sobre los que se realizaron las acciones fueron cualitativamente diferentes. Por ejemplo, en un ambiente tradicional, la población se define por lo general mediante un enunciado que señala el tipo de distribución y sus parámetros, sin ninguna referencia concreta. En cambio, en el ambiente de simulación, los estudiantes se vieron obligados a modelar la población y visualizarla por medio de representaciones menos abstractas, como es una tabla con los elementos poblacionales hipotéticos, complementándola en algunas actividades con una gráfica y sus medidas descriptivas. De esta manera, el objeto "población" se volvió visible y más concreto para los estudiantes.

Además, acciones que son importantes en el enfoque tradicional, como la estandarización de la distribución muestral para poder utilizar las tablas de probabilidad, y que son fuente de numerosos errores de cálculo, no fueron necesarias en el ambiente de simulación. Por el contrario, acciones que no se requieren en el enfoque tradicional, como es el caso de la formulación de un modelo de población, fueron un elemento importante en el enfoque de simulación.

Sin embargo, el proceso de simulación no estuvo exento de dificultades y los estudiantes exhibieron algunas confusiones en algunas etapas del proceso. Las principales dificultades estuvieron relacionadas principalmente con el uso de las representaciones simbólicas del software, como es el caso de la formulación del modelo poblacional, la definición de la expresión para calcular el estadístico en las muestras y la definición de las expresiones para calcular las probabilidades de resultados muestrales. En cuanto a las confusiones, destacan la confusión entre distribución de una muestra con una distribución muestral (dado el parecido de ambas) y la tendencia a ver a las muestras de manera aislada, en lugar de verlas como distribuciones.

Por todo ello, consideramos que el concepto de distribuciones muestrales resultó complejo para los estudiantes, ya que, a pesar del diseño de las actividades y las características educativas del software, en el cuestionario posterior y las entrevistas con algunos estudiantes, se observa que algunos estudiantes tuvieron cierta evolución en los significados, aunque de manera diferenciada, llegando a obtener muy buenos resultados en algunos ítems, mientras que en otros, hubo un incremento apenas notable en las respuestas correctas y sus argumentaciones; no fue el caso de otros, que continuaron sin comprender importantes conceptos como la variabilidad muestral y el efecto del tamaño de muestra sobre ella, así como en la forma de las distribuciones muestrales.

IMPLICACIONES PARA LA ENSEÑANZA DE LAS DISTRIBUCIONES MUESTRALES

Entre las ventajas que observamos al resolver problemas mediante simulación podemos mencionar las siguientes:

1. Los estudiantes fueron partícipes de todo el proceso, desde la formulación de la población hasta el cálculo de probabilidades, utilizando para ello diversas representaciones (aspecto representacional al que se refiere Biehler, 1991) que estuvieron a su alcance en la mayoría de los casos. Esto permitió que los estudiantes se enfocaran en el proceso y también en el resultado, a diferencia del ambiente tradicional, donde usualmente el énfasis se centra en los resultados.
2. La manera en la que los estudiantes pudieron calcular e interpretar las probabilidades de resultados muestrales como proporciones de casos de interés en un total de casos observados y que ellos mismos generaron mediante la toma de muestras, ayudados sobre todo por las características del software, el cual les permitió calcular los resultados mediante una expresión definida por ellos mismos y, en algunos casos, filtrarlos y hasta sombrearlos.
3. El ambiente de simulación volvió triviales, y en algunos casos hizo innecesarios (aspecto computacional que señala Biehler, 1991), procesos que son laboriosos y complicados en un ambiente de lápiz y papel, como la estandarización de la distribución muestral y el uso de tablas de probabilidad, los cuales son fuente de diversos errores al resolver los problemas.
4. Observamos, además, que los estudiantes le encuentran sentido a la resolución de problemas de distribuciones muestrales mediante simulación, ya que han construido por sí mismos las distribuciones, generando las poblaciones, tomando muestras, definiendo estadísticos y calculando sus probabilidades como proporciones de casos que se han presentado.

Por todo ello, los resultados sugieren que es importante que, antes de que los estudiantes se inicien en el estudio de la inferencia estadística, tengan experiencias de enseñanza donde exploren el concepto de distribuciones muestrales y sus propiedades, el teorema del límite central y el efecto del tamaño de muestra en la probabilidad de un determinado resultado muestral; un ambiente de estadística dinámica, apoyado en representaciones múltiples, parece ser apropiado para ello.

Consideramos que los significados que los estudiantes atribuyen a las distribuciones muestrales después de la experiencia de enseñanza con simulación les pueden ayudar a comprender más fácilmente procedimientos inferenciales. Por ejemplo, en los procesos de estimación, es importante el hecho de entender cuáles son los valores más probables de ciertos resultados muestrales y alrededor de qué valor se encuentran. Así como entender que, cuando la variabilidad muestral rebasa ciertos límites, puede deberse a causas no atribuibles al azar, lo que es importante en procesos de prueba de hipótesis.

REFERENCIAS BIBLIOGRÁFICAS

Biehler, R. (1991), "Computers in probability education", en R. Kapadia y M. Borovcnik (eds.), *Chance Encounters: Probability in Education*, Dordrecht, Kluwer Academic Publishers, pp. 169-211.

Chance, B., R. DelMas y J. Garfield (2004), "Reasoning about sampling distributions", en D. Ben-Zvi y J. Garfield (eds.), *The Challenge of Developing Statistical Literacy, Reasoning and Thinking*, Dordrecht, Kluwer Academic Publishers, pp. 295-323.

DelMas, R., J. Garfield y B. Chance (1999), "Exploring the role of computer simulations in developing understanding of sampling distributions", artículo presentado en la Reunión Anual de AERA (American Educational Research Asociation).

Dörfler, W. (1993), "Computer use and views of the mind", en Ch. Keitel y K. Ruthven (eds.), *Learning from Computers: Mathematics Education and Technology*, Springer Verlag, NATO ASI Series, núm 121, pp. 159-189.

Finzer, W., T. Erickson y J. Binker (2002), *Fathom Dynamic Statistics Software*, Key Curriculum Press Technologies.

Godino, J. (2002), "Un enfoque ontológico y semiótico de la cognición matemática", *Recherches en Didactique des Mathématiques*, vol. 22, pp. 237-284.

Godino, J. y C. Batanero (1994), "Significado institucional y personal de los objetos matemáticos", *Recherches en Didactique des Mathématiques*, vol. 14, núm. 3, pp. 325-355.

——— (1998), "Clarifying the meaning of mathematical objets as a priority area of research in Mathematics Education", en A. Sierpinska y J. Kilpatrick (eds.), *Mathematics Education as a Research Domain: A Search for Identity*, Dordrecht, Kluwer Academic Publishers, pp. 177-195.

Gordon, F. y S. Gordon (1992), "Sampling + simulation = statistical understanding computer graphics simulations of sampling distributions", en F. Gordon y S. Gordon (eds.), *Statistics for the Twenty-First Century*, MAA Notes, núm. 26, The Mathematical Association of America, pp. 207-216.

Lipson, K. (2000), *The Role of the Sampling Distribution in Developing Understanding of Statistical Inference*, tesis de doctorado, University of Technology of Swinburne, Australia.

——— (2002), "The role of computer based technology in developing understanding of the concept of sampling distribution", en B. Phillips (ed.), *Proceedings of the Sixth International Conference on Teaching Statistics*, Ciudad del Cabo, Sudáfrica, Voorburg International Statistics Institute.

Meletiou-Mavrotheris, M. (2004), "Technological tools in the introductory statistics classroom: Effects on student understanding of Inferential Statistics", *International Journal of Computers for Mathematical Learning*, núm. 8, Dordrecht, Kluwer Academic Publishers, pp. 265-297.

Mills, J. D. (2002), "Using computer simulation methods to teach statistics: A review of the literature", *Journal of Statistics Education*, vol. 10, núm. 1, recuperado en enero de 2008, http://www.amstat.org/publications/jse/v10n1/mills.html.

Moore, D. S. (1990), "Uncertainty", en L. Steen (ed.), *On the Shoulders of Giants: New Approaches to Numeracy*, Estados Unidos, National Academy Press, pp. 95-137.

——— (1992), "What is statistics?", en D. Hoaglin y D. Moore (eds.), *Perspectives on Contemporary Statistics*, MAA Notes, núm. 21, Mathematical Association of America, pp. 1-17.

Pea, R. (1987), "Cognitive technologies for mathematics education", en A. Schoenfeld (ed.), *Cognitive Science and Mathematics Education*, Lawrence Erlbaum Associates Publishers, pp. 89-122.

Saldanha, L. y P. Thompson (2002), "Student's scheme-based conceptions of sampling and its relationship to statistical inference", en D. Meuborn, P. Sztajn, D. White, H. Wiegel y K. Nooney (eds.), *Proceedings of the Twenty-*

Fourth Annual Meeting of North American Chapter of the International Group for the Psychology of Mathematics Education, Athens, Georgia, pp. 1305-1316.

Sánchez, E. y S. Inzunsa (2006), "Meaning's construction about sampling distributions in a dynamic statistics environment", en A. Roosman y B. Chance (eds.), *Proceedings of the 7th International Conference on Teaching Statistics,* IASE-ISI, Salvador, Brasil, pp. 1-6.

Shaughnessy, M. (1992), "Research in probability and statistics: Reflections and directions", en D. A. Grouws (ed.), *Handbook of Research on Mathematics Teaching and Learning,* Nueva York, Macmillan, pp. 465-494.

Thisted, R. A. y P. F. Velleman (1992), "Computers and modern statistics", en D. Hoaglin y D. Moore (eds.), *Perspectives on Contemporary Statistics,* MAA Notes, núm. 21, Mathematical Association of America, pp. 19-39.

DATOS DEL AUTOR

Santiago Inzunsa Cázares
Universidad Autónoma de Sinaloa, México
sinzunza@uas.uasnet.mx