Educación matemática
ISSN: 0187-8298
ISSN: 2448-8089
Editorial Santillana

Silvestre Castro, Eleazar; Sánchez Sánchez, Ernesto Alonso; Inzunza Cazares, Santiago

El razonamiento de estudiantes de bachillerato sobre
el muestreo repetido y la distribución muestral empírica

Educación matemática, vol. 34, núm. 1, 2022, pp. 100-130
Editorial Santillana

DOI: https://doi.org/doi.org/10.24844/EM3401.04

 https://doi.org/10.24844/EM3401.04

El razonamiento de estudiantes de bachillerato sobre el muestreo repetido y la distribución muestral empírica

The reasoning of high school students about repeated sampling and the empirical sampling distribution

Eleazar Silvestre Castro,[1] Ernesto Alonso Sánchez Sánchez,[2]
Santiago Inzunza Cazares[3]

Resumen: Presentamos los resultados de la implementación de una trayectoria de aprendizaje cuyo propósito fue explorar el razonamiento que desarrollan estudiantes de bachillerato sobre el muestreo y distribuciones muestrales empíricas (DME). Se diseñaron tres tareas que trataban de la estimación de una proporción, las cuales fueron resueltas por los estudiantes simulando el muestreo en forma física y con el apoyo de un software. Las respuestas a las tareas fueron analizadas y codificadas con técnicas de teoría fundamentada. Los resultados indican que, en un primer momento, la adherencia a una concepción aditiva del muestreo empuja a muchos estudiantes a "reconstruir" la población a través de la unificación de muestras, pero una vez que construyen y emplean una distribución muestral empírica, su razonamiento comienza a transitar hacia una concepción de tipo multiplicativa, mostrando indicios de asimilar algunas propiedades básicas de la distribución muestral. Con base en los patrones de

---

Fecha de recepción: 19 de mayo de 2020. **Fecha de aceptación:** 9 de noviembre de 2021.

[1] Departamento de Matemáticas, Universidad de Sonora, Hermosillo, Sonora, eleazar.silvestre@gmail.com, orcid.org/0000-0002-9472-483X.

[2] Departamento de Matemática Educativa, CINVESTAV-IPN, Ciudad de México, esanchez@cinvestav.mx, orcid.org/0000-0002-8995-7962.

[3] Facultad de Informática, Universidad Autónoma de Sinaloa, Culiacán Sinaloa, sinzunza@uas.edu.mx, orcid.org/0000-0003-4014-6031.

respuesta se proponen tres niveles de razonamiento: ingenuo, transicional y multiplicativo-distribucional.

Palabras clave: *Niveles de razonamiento, concepción aditiva y multiplicativa del muestreo repetido, distribución muestral empírica, simulación computacional, bachillerato.*

Abstract: We present the results of the implementation of a learning trajectory whose purpose was to explore the development of high school students' reasoning about sampling and empirical sampling distributions (ESD). We designed three tasks related to the estimation of a proportion, which were solved by the students by physically simulating the sampling process and with the support of a software. The answers to the tasks were analyzed and coded with grounded theory techniques. Results indicate that, at first, the adherence to an additive conception of sampling pushes many students to "reconstruct" the population through the unification of samples, but once they build and use an empirical sampling distribution, their reasoning begins to move towards a multiplicative conception, showing signs of assimilating some basic properties of the sampling distribution. Based on the response patterns, three levels of reasoning are proposed: naive, transitional and distributional-multiplicative.

Keywords*: Levels of reasoning, additive and multiplicative conception of repeated sampling, empirical sampling distribution, computer simulation, high school.*

1. INTRODUCCIÓN

La actividad central de la inferencia estadística es hacer afirmaciones probables acerca de una población, a partir de la información que ofrece una muestra. Aunque esta idea es aparentemente simple, esconde un sofisticado aparato conceptual e importantes aspectos sutiles. Chance *et al*. (2004) señalan que mientras los estudiantes universitarios suelen aprender a hacer los cálculos en tareas de inferencia estadística, generalmente no logran entender los procesos subyacentes que explican las razones de esos cálculos y cuya comprensión les proporcionaría las claves para interpretar los resultados de manera fructífera.

Dichos autores sugieren que tales dificultades se explican en gran medida por una ausente o precaria comprensión, por parte de los estudiantes, del concepto de *distribución muestral*. En efecto, este concepto es central en cualquier procedimiento de inferencia estadística y cuando los estudiantes lo llegan a entender se aprovisionan de una poderosa herramienta. Pero para hacerse de esta herramienta se requiere recorrer un camino sinuoso ya que el concepto teórico de distribución muestral es muy complejo y multifacético, pues está ligado a otros conceptos como población, muestra, variabilidad y distribución; la adquisición de estos y sus relaciones también requiere de grandes esfuerzos (Begué *et al*., 2018). Para contribuir en la comprensión de entender cómo los estudiantes integran estos conceptos para darle sentido al muestreo y su relación con la inferencia estadística, en el presente estudio se realizó un experimento de enseñanza vía el diseño y aplicación de una trayectoria hipotética de aprendizaje. El objetivo de esta es introducir a los estudiantes al concepto de distribución muestral a partir de algunas actividades. En este contexto se formulan las siguientes preguntas de investigación:

- ¿Qué *niveles de razonamiento* emergen cuando estudiantes de bachillerato responden a preguntas que involucran una DME?
- ¿Qué *rasgos* de las *concepciones aditiva* o *multiplicativa* del muestreo repetido simulado presentan las respuestas de los estudiantes?

La segunda pregunta se relaciona con las concepciones de muestreo repetido de los estudiantes; para esto se caracterizan las concepciones aditiva y multiplicativa del muestro repetido, haciéndonos eco de las concepciones de muestra formuladas por Saldanha y Thompson (2002).

2. ANTECEDENTES

2.1 Concepciones de la noción de muestra

Watson y Moritz (2000) examinaron las respuestas de poco más de 3,000 estudiantes de Tasmania de los grados 3, 6 y 9 a la pregunta “Si te dieran una ‘muestra’, ¿qué tendrías?”. En el análisis de las respuestas utilizaron el modelo SOLO (Biggs y Collis, 1982) para identificar que, en los niveles básicos, el concepto de muestra más sofisticado implica las nociones *parte-todo* y *representatividad*, por ejemplo: "una muestra es una parte representativa de un todo"; Watson y Moritz clasifican a esta como una concepción relacional de muestra. El rasgo central que lleva a considerarla relacional es la *representatividad*. No obstante, en estudios con adultos, Kahneman *et al*. (1982) identificaron la heurística de representatividad; esta, en el caso de muestras, se traduce en el sesgo de que "cualquier muestra es representativa", que resulta una idea errónea. Así, mientras que en el nivel básico la representatividad es un avance, con el desarrollo se convierte en un sesgo.

En la revisión de Shaughnessy (1992) sobre las varias falsas concepciones estadísticas populares entre la gente, menciona la creencia de que toda muestra es representativa. Rubin *et al*. (1990) destacan que la dificultad en el concepto de muestra deriva de la aparente oposición de las nociones de representatividad y variabilidad muestrales; una muestra no es totalmente representativa, pero tampoco está gobernada por una variabilidad incontrolada; conciliar ambas ideas es un desafío central para el aprendizaje del muestreo. Saldanha y Thompson (2002) identifican las concepciones aditiva y multiplicativa de muestra de estudiantes de bachillerato en el contexto de un experimento de enseñanza. La concepción aditiva de muestra es, simplemente, un subconjunto de la población; mientras que la multiplicativa implicaría una idea de (casi) proporcionalidad con relación a la variable en cuestión, es decir, una cierta representatividad.

Por su parte, Ainley *et al*. (2015) llaman la atención sobre la importancia de distinguir dos clases de situaciones de muestreo en función de la noción de población que se considera, a saber: "A) Situaciones en las que la población es material, finita y contable" y, "B) Situaciones en las que la población es una formulación matemática". La concepción aditiva de muestra (Saldanha y Thompson, 2002; ver el marco conceptual) se relaciona con poblaciones de las situaciones de tipo A, pues solo en estas poblaciones se puede pensar en un subconjunto. En efecto, cuando la población es una distribución no es inmediato dilucidar a una muestra como subconjunto. El caso más sencillo de distribución es la binomial, en la que una muestra es una secuencia de "éxitos" (E) y "fracasos" (F), por ejemplo, la muestra EEFEF no es un subconjunto de la población.

2.2 Muestreo repetido simulado y DME

Hesterberg (1998) escribe que los métodos de simulación "ofrecen una forma de enseñar estadística y probabilidad mediante el uso de la computadora para obtener experiencia directa y comprensión intuitiva a través de gráficos" (p. 1). Con esta idea, en los años 90 varios trabajos comenzaron a proponer el uso del muestreo repetido simulado para la enseñanza de estimación de parámetros (Taffe y Garnham, 1996), de las pruebas de hipótesis (Ricketts y Berry, 1994) o de los intervalos de confianza (Johnson, 2001). A partir de entonces, se han multiplicado las propuestas de tal manera que sería imposible mencionarlas en este espacio (ver por ejemplo Watson y Chance, 2012). Una de las aportaciones más importantes de las técnicas de muestreo repetido simulado, es la posibilidad de generar distribuciones muestrales *empíricas*, recurso clave para investigaciones y propuestas didácticas que analizan y desarrollan el razonamiento estadístico en temas de inferencia desde un *enfoque informal*. Al respecto, Ben-Zvi *et al.* (2015) comentan que "la tecnología permite que los estudiantes se involucren directamente en la 'construcción' de la distribución muestral, enfocándose en el proceso involucrado, en lugar de que se les presente solo el resultado final" (p. 298). En este sentido, Garfield y Ben-Zvi (2008) dedican un capítulo de su libro al tema de aprender a razonar con muestras y distribuciones muestrales.

Por un lado, algunas investigaciones empíricas evidencian un impacto incipiente en el aprendizaje sobre distribuciones. A pesar del desarrollo de herramientas visuales y flexibles, Chance *et al.*, (2004, 2007) sugieren que hacer demostraciones de simulaciones no conduce necesariamente a una mejor comprensión y razonamiento. Por otro lado, hay resultados que exhiben el potencial de utilizar estos recursos; por ejemplo, Burril (2019) documenta cómo crear imágenes mentales de las ideas fundamentales de la estadística con apoyo de recursos tecnológicos en estudiantes para profesor, entre dichas ideas fundamentales se incluye a la distribución muestral. Jacob y Doerr (2013) señalan que, con la instrucción y guía necesarias, los estudiantes pueden llegar a establecer relaciones entre la distribución muestral y el parámetro con el que se generó. Inzunza e Islas (2019) y van Dijke-Droogers *et al.* (2019) reportan que un enfoque informal basado en simulación del muestreo ayuda a los estudiantes a desarrollar un razonamiento normativamente correcto sobre las distribuciones muestrales.

Un componente de la investigación es la creación de marcos que describan o modelen el desarrollo y el crecimiento del pensamiento de los estudiantes sobre la variación, distribución y distribución muestral (Langrall *et al.*, 2017). En particular, Chance *et al.* (2004) proponen un marco de desarrollo del razonamiento de estudiantes universitarios sobre la noción de distribución muestral, mientras que Noll y Shaughnessy (2012) propusieron un marco enfocado al desarrollo de la noción de variabilidad muestral en estudiantes de secundaria y bachillerato. Consideramos que hace falta un marco que describa rasgos del razonamiento sobre distribución muestral de estudiantes de bachillerato; en este trabajo se avanza en este sentido.

3. MARCO CONCEPTUAL

Estamos interesados en explorar y entender cómo es que los estudiantes comienzan a formar el concepto de distribución muestral. En ciertos acercamientos se pretende que ellos se apropien del concepto partiendo de la definición y deduciendo sus propiedades (ver por ejemplo Chance *et al.*, 2004). En otro enfoque, como el que aquí se adopta, se parte de tareas de muestreo y se utilizan recursos auxiliares para que los estudiantes asocien las propiedades al concepto. Este es el caso de una DME generada mediante muestreo repetido simulado en una computadora; se espera que los estudiantes deduzcan de esta, propiedades similares a las del concepto teórico. Consideramos que la DME es un modelo más concreto del concepto teórico de distribución muestral. Como el concepto de distribución muestral es muy complejo, es necesario elegir el tipo de caracterización y las propiedades que serían viables de aprender por los estudiantes de bachillerato y que prefiguren debidamente el concepto. A continuación, avanzaremos en dicha caracterización.

3.1 La DME de proporciones para estudiantes de bachillerato

El concepto de DME se refiere a una distribución finita de valores de un estadístico de muchas muestras de un mismo tamaño de una población (Noll y Shaughnessy, 2012). Una característica importante de una DME para la enseñanza es que se puede obtener mediante una serie de acciones de muestreo repetido que el estudiante realiza a partir de pocas instrucciones; en una primera etapa se puede hacer de forma rudimentaria con manipulativos y a lápiz y papel, y en una etapa posterior de manera más elaborada con ayuda de algún software (van Dijke-Droogers *et al.*, 2019); en nuestro caso utilizamos el software Fathom (Finzer, 2001).

Una DME es una versión concreta y simplificada de una distribución muestral teórica; mientras esta se obtiene mediante complejos procedimientos matemáticos, aquella se basa en un procesamiento computacional. En este punto conviene mencionar un aspecto sutil, pero importante, y es que una distribución muestral para proporciones en general y una DME para proporciones en particular, dependen solo de tres condiciones dadas: 1) El muestreo es aleatorio y con remplazo, 2) Un tamaño de muestra determinado y, 3) La proporción en la población del atributo en estudio.

Cuando las muestras son de un tamaño (n) mayor o igual a 10 y el número mínimo de muestras representadas cercano a 100, las propiedades más importantes de una DME son: 1) el centro de la DME es una aproximación a la proporción poblacional, 2) la forma de la DME tiende a ser normal y 3) la desviación estándar de una DME disminuye conforme aumenta el valor de n. Una distribución muestral teórica permite calcular la probabilidad de que el estadístico de una muestra tome un valor determinado o esté en un rango de valores (en el caso discreto) o se encuentre en un intervalo determinado (en el caso continuo); la DME asociada permite estimar dichas probabilidades con base en las frecuencias de los eventos correspondientes.

3.2 Concepciones aditiva y multiplicativa del muestreo

En el contexto del muestreo repetido, Saldanha y Thompson (2002) propusieron la "imagen aditiva" de una muestra como el "ver una muestra simplemente como un subconjunto de una población y ver múltiples muestras como múltiples subconjuntos" (p. 265). Bajo esta perspectiva, las muestras tienen un carácter estático porque son vistas como subconjuntos aislados de la población de donde fueron extraídas; se enfatiza únicamente la relación parte-todo de la muestra con la población y se evalúa lo inusual de un valor muestral específico a través de su cercanía con el parámetro poblacional. En cambio, los autores definen una "imagen multiplicativa" de una muestra como:

> ... una mini versión cuasi-proporcional de la población muestreada, donde la imagen de 'cuasi-proporcionalidad' emerge al anticipar una variedad limitada de resultados si se repitiera el proceso de muestreo...llamamos a esta imagen de muestra una concepción multiplicativa de muestra (MCS) porque su constitución implica operaciones conceptuales de razonamiento multiplicativo. (Saldanha y Thompson, 2002, p. 266)

Según este planteamiento, cuando un estudiante ve a una muestra como casi-representativa de la población es porque tiene conciencia de la variabilidad muestral; de esta manera podría llegar a evaluar lo inusual de un valor muestral específico contrastándolo ante una colección de valores muestrales que se encuentran *distribuidos* alrededor del parámetro poblacional. Encontramos estas definiciones potencialmente útiles para caracterizar y analizar el razonamiento de los estudiantes respecto a sus concepciones de muestra y muestreo, pero también notamos que cada una refiere a rasgos fundamentales de distinta naturaleza entre sí: el concepto de "subconjunto" en el caso de la imagen aditiva, y el concepto de "cuasi–proporcionalidad" para la imagen multiplicativa.

Un problema con las caracterizaciones de la concepción aditiva y multiplicativa de Saldanha y Thompson es que no se relacionan con las concepciones de *población* que puedan tener los estudiantes, lo que resulta desconcertante debido a que población y muestra están íntimamente relacionadas. En general, no se le ha dado mucha importancia al concepto de *población* en las investigaciones de educación estadística sobre muestra, muestreo y distribuciones muestrales. Por ejemplo, Watson y Moritz (2000, p. 112) se preguntan "¿cómo es que los estudiantes de diferentes edades y sexos definen muestras y qué propósito ven ellos en el muestreo?" Pero no formulan una pregunta análoga para la noción de población. Es evidente, sin embargo, que hay dos concepciones de este concepto; la primera que suele ser adoptada por novatos y la segunda por expertos:

> *Población 1:* Un conjunto de individuos u objetos que están bajo estudio (Novato). Bajo este esquema el muestreo se realiza sin reemplazo y, por tanto, cada muestra es mutuamente excluyente con cualquier otra obtenida mediante el mismo proceso.
> *Población 2:* Un conjunto de valores de una variable observada en una población (población 1) (Expertos). Bajo este esquema el muestreo se realiza con reemplazo y, por tanto, cada muestra es estocásticamente independiente de cualquier otra.

El paso de la primera a la segunda definición implica un salto de abstracción importante ya que en la primera se suele pensar en la naturaleza de las personas o cosas, mientras que la segunda es de naturaleza puramente aritmética y abstracta.

4. MÉTODO

4.1 Sujetos de estudio y escenario de investigación

La metodología de estudio es de carácter cualitativo. El experimento de enseñanza se llevó a cabo en condiciones escolarizadas, se realizó con un grupo de estudiantes de bachillerato (17 a 19 años) del Colegio de Ciencias y Humanidades de la Universidad Nacional Autónoma de México (UNAM) plantel Azcapotzalco, ubicado en la Ciudad de México. La realización del experimento ocurrió mientras los estudiantes iniciaban su curso de Estadística y Probabilidad II. Consistió de la aplicación de tres tareas, A, B y C, que se desarrollaron en tres sesiones de clase consecutivas con duración aproximada de 90 minutos cada una.

En la primera tarea participaron 34 estudiantes, pero debido a sus compromisos extraescolares, no previstos y ajenos a nuestro control, en las otras dos solo se presentaron 12. Los antecedentes de contenido estadístico de los estudiantes consistían de medidas descriptivas (medidas de tendencia central y dispersión) y probabilidad (enfoques de probabilidad y reglas para el cálculo de probabilidades en eventos simples y compuestos) que adquirieron en el curso anterior de Estadística y Probabilidad I.

La tarea A se trabajó en el aula de clases regular, con los estudiantes sin acceso a computadoras y organizados en 10 equipos de entre 3 y 4 integrantes cada uno; las tareas B y C se llevaron a cabo en un centro de cómputo de la institución escolar, con los estudiantes trabajando de manera individual. Al momento de la aplicación de las tareas, la participación del profesor-investigador (el primer autor del presente artículo) se limitó a responder dudas de los estudiantes para que resolvieran las actividades sin proveerles las respuestas normativas.

4.2 Descripción de la trayectoria de actividades

Las tareas que resolvieron los estudiantes fueron enmarcadas en una trayectoria hipotética de aprendizaje; la tabla 1 resume el conjunto de tareas, objetivos e hipótesis de aprendizaje correspondientes:

Tabla 1. Componentes de la trayectoria de actividades sobre muestreo y la DME

Nombre	Descripción	Objetivo de aprendizaje	Hipótesis de aprendizaje
Tarea A, primer momento. Estimación de la proporción p con manipulativos	Obtención de muestras de una urna para estimar la proporción p de un atributo.	Que los estudiantes relacionen el muestreo repetido con la estimación de la proporción de la población.	H1. Los estudiantes entienden que se puede estimar la proporción de la población a partir de obtener muestras repetidas y observar la proporción en cada muestra. H2. Perciben algunos rasgos característicos del muestreo repetido, en particular, la variabilidad muestral. H3. Se basan en teorías personales para proponer un procedimiento de estimación de la proporción con base en las observaciones de las muestras obtenidas y proponen un valor para dicha proporción.
Tarea A, momento 2. Estimación de la proporción p a partir de una DME	Reorganizar los valores del estadístico en una distribución de frecuencias (DME) para estimar p con base en esta colección.	Que los estudiantes adquieran una noción de DME y la representen con ayuda de lápiz y papel y con Fathom. Que utilicen la DME para estimar la proporción de una población.	H4. Los estudiantes aprenden a organizar los valores del estadístico en una DME y que el software Fathom hace esta tarea. H5. Perciben rasgos de una DME (su moda, que no es uniforme y tiende a ser simétrica). H6. Reorganizan el procedimiento para estimar la proporción que se elaboró en la tarea anterior, para hacerlo con base en la DME.

Tarea B. Estimación de probabilidades a partir de la DME	Describir los cambios que ocurren en la DME cuando se simulan más valores del estadístico. Estimar la probabilidad de un valor del estadístico utilizando la DME.	Que los estudiantes consoliden su noción de DME y la utilicen para estimar probabilidades.	H7. Los estudiantes perciben la analogía entre el muestreo repetido que realizaron físicamente en las tareas previas y el muestreo simulado que se realiza con el software. H8. Observan que el centro de la DME se corresponde con la proporción en la población y que su distribución es aproximadamente simétrica. H9. Utilizan las frecuencias (relativas) en la DME para estimar las probabilidades de que el estadístico de una muestra tenga un valor determinado.
Tarea C. Identificación de una DME	Identificar la representación correspondiente a una DME con parámetros n=10 y p=.38	Que los estudiantes reconozcan una DME a partir de una situación específica.	H10. Los estudiantes determinan el centro de una DME y lo asocian con la proporción de la población de la cual se obtuvo la DME. H11. Reconocen cuando una distribución es aproximadamente simétrica y en forma de campana y asocian estos rasgos a una DME. H12. Tomando como criterio los anteriores puntos reconocen la DME que corresponde a una situación dada.

La tarea A contempló dos momentos; en el primero se atiende la recomendación de van Dijke *et al.* (2019) de iniciar con los estudiantes explorando el muestreo repetido vía la manipulación de materiales físicos, pues argumentan que de esta manera puede darse sentido a conceptos difíciles e intangibles como el muestreo y la variabilidad muestral. La población que exploraron consistía en 1,000 cuadros blancos y 1,000 cuadros negros de papel de aproximadamente 1 cm^2 cada uno, mezclados en una urna opaca. Los estudiantes desconocían la proporción de cuadros de cada color y lo que podían observar a simple vista no era suficiente para hacer alguna hipótesis sobre dicha proporción. Se les pidió estimar el porcentaje de cuadros negros de la urna. Para esto se sugirió que obtuvieran muestras de tamaño 10, pero se les dejó en libertad de que extrajeran el número de muestras y el tipo de muestreo (con o sin reemplazo) que consideraran pertinente.

En el segundo momento de la tarea, utilizando la misma urna, se instruyó a los estudiantes que obtuvieran 74 muestras de tamaño 10 para que contaran el número de cuadritos negros en ella, registrando dicho número; luego se les pidió que introdujeran los 74 resultados obtenidos en el software Fathom representando así una DME. Finalmente, tenían que estimar el valor del parámetro tomando como base la información de la distribución representada por el software (ver figura 1).

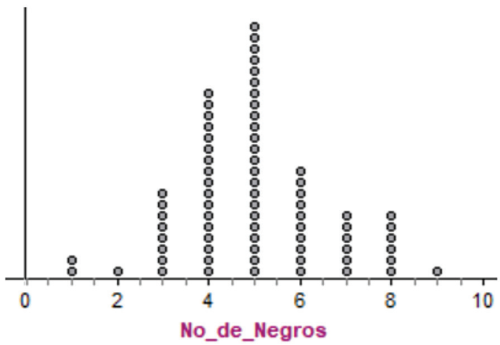
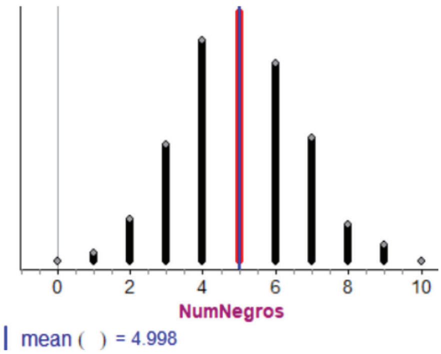


Figura 1. DME de un muestreo repetido 74 veces (n=10)

La tarea B consistió en explorar la DME construida en la tarea A, pero ahora seleccionando las muestras en forma repetida mediante simulaciones computarizadas para generar una DME de 2 mil valores del estadístico; para ello se reveló a los estudiantes el valor de la proporción poblacional al inicio de la tarea, pues para realizar la simulación se requiere partir de un parámetro conocido. En la exploración se solicitó a los estudiantes expresar los cambios y patrones observados en la DME a medida que aumentaba el número de muestras, para finalmente estimar la probabilidad de obtener 5 cuadros negros en la muestra de 10. La habilidad para estimar una probabilidad o proporción a través de la

DME es fundamental para desarrollar conceptos y técnicas del dominio estocástico (Saldanha y Thompson, 2002; Jacob y Doerr, 2013).



DME de 2,000 valores

NumNegros	0	3
	1	18
	2	85
	3	231
	4	435
	5	489
	6	388
	7	243
	8	73
	9	33
	10	2
	Column Summary	2000

Figura 2. DME con 2 mil valores de la variable "número de cuadritos negros en la muestra de n=10

La tarea C consistió de un ítem inspirado en los trabajos de Chance *et al.* (2004); primero se les describió el siguiente experimento a los estudiantes:

> De una urna que contiene 2 mil frijoles, de los cuales 1,240 son pintos y 760 negros, se extraen 3 mil muestras aleatorias, cada una de ellas de tamaño 10, tomadas con reemplazo. Para cada muestra se calcula la proporción de frijoles negros que se presentó en cada muestra. El experimento se simula en el software Fathom, el cual proporciona la representación de una DME.

De tal manera que eligieran una de las siguientes opciones:

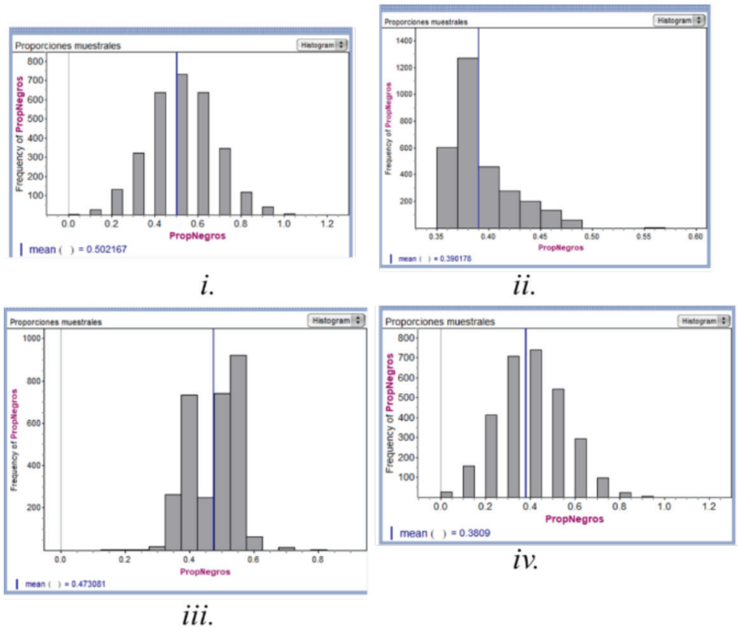


Figura 3. Opciones de gráficos para una DME con n=10 y p=.38 (opción correcta: iv)

La resolución de la tarea no requiere del uso de simulaciones en Fathom sino del conocimiento de las propiedades básicas de la DME (que su centro se aproxima a la proporción poblacional y que es simétrica). La tarea continúa enmarcada en la exploración de un modelo de urna, donde la población a la que se hace referencia es ahora un conjunto grande de frijoles de dos tipos (pintos y negros). Al igual que en el estudio de Chance y colaboradores, con esta tarea pretendemos observar y analizar qué conocimientos y razonamientos desarrollaron los estudiantes con relación al concepto de DME.

4.3 Proceso de implementación

El primer momento de la tarea A se realizó con los 10 equipos disponiendo de una urna que contenía la población de cuadros mencionada anteriormente. Tras describir el contenido de la urna sin declarar el valor de la proporción poblacional, el investigador tomó una muestra de 10 cuadros de la urna sin reposición, los regresó a la urna e hizo la pregunta abierta a la clase: "¿es este resultado, x% muestral de cuadros negros, suficiente evidencia para asegurar que la urna posee X% de cuadros negros?" Hemos observado en pilotajes de la tarea que casi la totalidad de los estudiantes rechazan sin titubeo la propuesta de tomar el porcentaje muestral como la proporción poblacional argumentando que la muestra es muy pequeña, y/o que si se repitiera el experimento se podría obtener un valor drásticamente distinto al primero. Por tanto, enseguida se solicitó a los estudiantes recabaran tantas muestras como desearan para estimar la proporción poblacional, pero respetando el tamaño de muestra en 10 elementos. Se abrió entonces un espacio de aproximadamente 45 minutos para que los estudiantes realizaran el proceso de muestreo y análisis de datos.

Tras finalizar la exploración del muestreo de manera física, se registraron las propuestas para la proporción poblacional en el pizarrón, y la atención se focalizó en la variabilidad que presentaron tales valores. El profesor preguntó nuevamente a la clase: "¿qué pasaría si hacemos este experimento muchas más veces?, ¿podría ayudarnos eso a encontrar una mejor estimación de p?" Bajo esta consigna es que inicia el segundo momento de la tarea, en el que se propone a los estudiantes tomar más muestras, pero esta vez utilizando un muestreo con reemplazo y registrando los valores del estadístico directamente en un gráfico de frecuencias (figura 1). De esta manera la tarea A finalizó con los estudiantes generando una nueva estimación de p a partir de la DME constituida por 74 valores del estadístico.

La tarea B se llevó a cabo en la segunda sesión, teniendo lugar en el centro de cómputo e inició con la intención de explorar qué hubiese pasado con la DME de la tarea A si se hubiese seleccionado un mayor número de muestras. Tras tomarse un tiempo de aproximadamente 35 minutos para ingresar los comandos necesarios de la programación en Fathom, los estudiantes registraron los cambios y patrones en la DME a medida que aumentaba su convergencia al modelo probabilístico binomial.

Arribado al punto en el que la DME se constituyó de 2,000 valores del estadístico, se solicitó a los estudiantes estimar la probabilidad de obtener el valor del estadístico igual a 5 cuadros negros en la muestra de 10. Para resolver la situación, los estudiantes podían volver a simular casi instantáneamente la DME a través de un comando simple que generaba una nueva colección de valores del estadístico del mismo tamaño; con este recurso los estudiantes podían apreciar la variabilidad entre distribuciones muestrales y brindar una mejor estimación de la probabilidad solicitada.

La tercera y última sesión se llevó a cabo en el mismo centro de cómputo y se trabajó en la tarea C. Previo a resolver la tarea, la sesión inició con una discusión breve en la que se resumieron los hallazgos anteriores por parte de los estudiantes al explorar la DME, utilizando las simulaciones computarizadas. La discusión se centró en la persistencia del acumulamiento en los valores del 4 a 6 cuadros negros por muestra, y que la forma tipo acampanada de la distribución parecía estabilizarse a medida que aumentaba el número de valores simulados. Una observación importante en esta exploración fue que los estudiantes reconocieron al promedio de la DME como una aproximación a p cuando se divide por *n*, y como un valor aparentemente constante mientras la distribución converge al modelo binomial. Acto seguido, los estudiantes respondieron la tarea C trabajando de manera individual.

4.4 Instrumentos de recolección de datos y procedimiento de análisis

Los instrumentos utilizados en la investigación son congruentes con su naturaleza cualitativa. Estos fueron: a) una bitácora de campo y un conjunto de memorándums personales del profesor-investigador sobre el desarrollo general de la investigación; b) registro de entrevistas breves, no estructuradas y generadas "en vivo" con algunos estudiantes mientras resolvían los cuestionamientos; c) las respuestas brindadas por los estudiantes ante las tareas de la secuencia: la tarea A se realizó con hojas de trabajo físicas, en las dos restantes se utilizaron hojas de trabajo electrónicas; y (4), registro de material audiovisual del desarrollo de la secuencia a lo largo de las tres sesiones. Para este reporte tomamos las respuestas de los estudiantes como nuestra fuente primaria de datos; en particular, consideramos solo aquellas que en nuestra visión guardan una vinculación estrecha con las concepciones aditiva o multiplicativa del muestreo, en tanto que hemos descartado aquellas que obedecen a juicios idiosincráticos o heurísticos que difícilmente pueden ser relacionadas a tales concepciones o el de la DME.

Las respuestas fueron codificadas utilizando principios de codificación abierta de la *Teoría Fundamentada* (Birks y Mills, 2011), de acuerdo con nuestro cuerpo de datos y sin pretender cubrir todas las etapas que la teoría prescribe. En particular, analizamos los datos en dos niveles de codificación. El primer nivel se hizo analizando las similitudes y diferencias en las estrategias y argumentos de los estudiantes ante las distintas tareas, tomando como ejes analíticos los conceptos de población, muestra, muestreo y DME. Los autores de este trabajo compararon y refinaron los códigos asignados a cada respuesta de manera reiterativa (método de comparación constante) para obtener un primer conjunto de categorías, que denominamos *patrones de razonamiento*, los cuales representan manifestaciones de modos específicos de actuación y razonamiento intrínsecos a nuestros estudiantes y la micro cultura de la clase.

El segundo nivel de codificación se hizo vinculando los patrones de razonamiento con los rasgos normativos de los conceptos mencionados para cada tarea de la trayectoria de aprendizaje. De este proceso se obtuvo un nuevo conjunto de categorías que condensan a los códigos previos y con ellas se generaron tres niveles de razonamiento: *ingenuo (aditivo)*, *transicional* y *multiplicativo-distribucional*. Tales niveles constituyen un modelo empírico sobre el desarrollo gradual acerca de la complejidad estructural y de abstracción que alcanza el razonamiento de nuestros estudiantes sobre el concepto de distribución muestral.

Como punto final del análisis, contrastamos las hipótesis de aprendizaje de la trayectoria con los resultados expresados en los niveles de razonamiento de nuestro modelo empírico; la guía para este procedimiento fue valorar en qué medida se cumplen las hipótesis de aprendizaje correspondientes. Es importante mencionar que este contraste no tiene como objetivo principal valorar a detalle la efectividad de la trayectoria para el aprendizaje, sino contribuir a la identificación de aspectos no previstos en el experimento que podrían ser considerados en una nueva experimentación cuyo objetivo sea profundizar en la saturación y refinamiento de nuestras categorías (muestreo teórico).

5. RESULTADOS

Primero describimos ocho patrones de razonamiento de los estudiantes, obtenidos a partir del análisis de las respuestas a los problemas planteados e indicando en cada caso la frecuencia con la que se presentó. Posteriormente describimos los niveles de razonamiento en los que se organizan y jerarquizan las respuestas de nuestros estudiantes.

5.1 Tarea A, primer momento. Estimación de la proporción con manipulativos

Sin saber el contenido específico de la urna, todos los equipos consideraron que el muestreo debía efectuarse sin reemplazo, pero la mayoría (ocho equipos) lo hizo extendiendo el muestreo sin reemplazo hacia la muestra completa (i.e., primero seleccionaron una muestra elemento por elemento y sin sustitución, y después tomaron otra muestra del mismo tamaño sin haber retornado los elementos previamente extraídos a la urna; los dos equipos restantes tomaron los elementos sin reemplazo y los retornaron a la urna, antes de sacar una nueva muestra). Independientemente de las diferencias en los métodos de muestreo, a todos los equipos se les pidió hacer una estimación de la proporción de cuadros negros en la población. A continuación, se muestran los patrones de razonamiento que emergieron del análisis de sus respuestas.

Patrón 1. Proporción en la unión de muestras (4/10 equipos). Sacan N muestras de tamaño *n* sin remplazo y, en cada una cuentan el número de éxitos, obteniendo las frecuencias absolutas: n_1, n_2, ..., n_N. El estimador es la proporción de la suma de éxitos con relación al total de muestras. Por ejemplo, el equipo E6 respondió:

> Sacamos 20 muestras de 10 unidades. De las 200 unidades obtenidas, 86 son negras, por lo tanto, suponemos que el 43% de la población son negras.

Patrón 2. Promedio de éxitos o proporciones (3/10 equipos). Se sacan N muestras de tamaño *n* sin remplazo y para cada muestra se calcula la proporción p_i de la i-ésima muestra, luego se estima la proporción con el promedio de la p_i's. Por ejemplo, el equipo E1, argumenta:

> Sacamos [20] muestras de 10 en 10 y sumamos los porcentajes de cada una y obtuvimos el promedio al dividirlo entre 20.

El patrón 1 se propone diluir las muestras en una sola muestra más grande y de esta obtener la proporción solicitada. Como piensan en un muestreo sin remplazo, tiene sentido para ellos la acumulación de las muestras en una sola. Una vez que tienen la muestra grande consideran que esta es representativa (Rubin *et al.*, 1990). El procedimiento es aditivo pues considera muestras ajenas que se reúnen en una sola.

En contraste, quienes siguen el patrón 2 obtienen el porcentaje de cada muestra, no las diluyen en una sola, sino ya como promedios. Aún mantienen una idea aditiva de muestreo, pues siguen un procedimiento sin remplazo. No obstante, y aunque este procedimiento es aritméticamente equivalente al anterior, calculan el estadístico para cada muestra; este es un rasgo (aún muy básico) de la noción de distribución de un estadístico de muestreo, ya que, teniendo la lista de los porcentajes del atributo por muestra, antes de promediarlos, se podrían agrupar para obtener sus frecuencias y representarlas gráficamente.

Una vez que los estudiantes experimentan un proceso de muestro repetido con dispositivos manipulables (el recipiente con cuadritos negros y blancos) aprenden a construir una distribución empírica de éxitos o proporciones. Esto les permite entender un muestreo repetido virtual. En la siguiente sesión los estudiantes trabajaron con una DME utilizando Fathom.

5.2 Tarea A, momento 2. Estimación de p a partir de una DME

Patrón 3. Recuperación de una muestra grande (1/10 equipos). Respuestas clasificadas bajo este patrón evidencian este razonamiento: si x es el número de cuadritos negros en una muestra de n=10 (x = 0,1, ...,10) y f_X la frecuencia de muestras con x cuadritos negros, entonces la proporción se estima de la siguiente manera: $\hat{p} = \frac{\sum_{x=0}^{10} x f_x}{740}$. Para el caso de la gráfica de la figura 1, el procedimiento se desarrolla en la tabla 1, y el resultado sería $\hat{p} = \frac{370}{740} \cong 0.48$.

Tabla 2. Tabla de frecuencias de los resultados mostrados en la Figura 1

x	0	1	2	3	4	5	6	7	8	9	10	Total
f_x	0	2	1	8	17	23	10	6	6	1	0	74
xf_x	0	2	2	24	68	115	60	42	48	9	0	370

Al respecto, el equipo E8 comentó:

> Contabilizamos el número de cuadritos negros, se sacó el promedio y se realizó una regla de 3 y llegamos a la conclusión del porcentaje aproximado es 47% de cuadritos negros.

La regla de tres que refiere el equipo es: $\frac{n}{N} = \frac{Y}{740}$, así $\hat{p} = \frac{n}{N}$ donde N es el tamaño de la población, n el número de cuadritos negros en la población y Y el número de cuadritos negros contados en la DME, en el caso del ejemplo de la Tabla 1, $Y = 370$. En la simulación que hizo el equipo obtuvo el valor $Y = 348$.

Patrón 4. Conjunto modal o Moda de la DME (9/10 equipos). Toman como referencia un conjunto modal o directamente la moda. En ambos casos identifican los valores de la variable con su porcentaje respecto al tamaño de la muestra (10 en el caso del problema actual). En el primer caso estiman la proporción de la población mediante el promedio de los porcentajes de los valores modales y en el segundo toman el porcentaje de la moda respecto al tamaño de la muestra. A partir de esta calculan el estadístico dividiendo por el tamaño de la muestra, o equivalente, considerando el porcentaje de la moda respecto al tamaño de la muestra. Es decir, para ellos el valor 5 representa el 50%. Por ejemplo, el equipo E10 escribió:

> Consideramos que el % más plausible está entre los valores del 40% y 60%, ya que presentan una mayor frecuencia la cantidad de cuadritos negros de 4, 5 y 6. De igual forma podría ser 50, tomando en cuenta el promedio de 4, 5 y 6.

Los que siguen el procedimiento del patrón 3, aunque muestran que pueden extraer información de la distribución muestral, su estrategia para la estimación es recuperar una muestra grande a partir de las muestras de tamaño 10; se mueven implícitamente pensando que las muestras representadas en la DME son ajenas, lo cual ya no es el caso porque el muestreo es con remplazo. Por esta razón, consideramos que detrás de este procedimiento hay una concepción aditiva del muestreo (Saldanha y Thompson, 2002). En el patrón 4 se aplica la idea intuitiva de que las muestras más frecuentes son las más representativas, de tal manera que la concepción de muestreo deja de ser puramente aditiva (las muestras no se unifican) y se aprovecha la DME para identificar los porcentajes más frecuentes en las muestras.

Noll y Shaughnessy (2012) encontraron respuestas a las tareas de predicción en que los estudiantes se centran en la moda o grupos modales; juzgan que se apoyan en "centros débiles" y, respecto a su marco, la ubican en un nivel transicional. Una cuestión relevante de este patrón es que se basa en una propiedad de la distribución muestral (grupo modal o moda). Este razonamiento evoca la tensión entre representatividad y variabilidad señalada por Rubin *et al.* (1990), pues implica que se comienza a conciliar representatividad (grupo modal) con variabilidad en la forma siguiente: la proporción de las muestras del grupo modal varía poco de la proporción de la población, mientras que las que no pertenecen al grupo modal tienen una proporción más alejada.

5.3 Tarea B. Estimación de probabilidades a partir de la DME

Patrón 5. Tanteo cualitativo (4/12 estudiantes). Para estimar la probabilidad de que $x = 5$, proponen una probabilidad mediante una comparación cualitativa, observando dónde hay más o menos puntos en la distribución muestral, pero no consideran las frecuencias precisas. Por ejemplo, el estudiante A9 escribió:

> De acuerdo con la gráfica [Figura 2], la mayor probabilidad de que salgan valores oscila entre los números 4, 5 y 6 por lo que se observa que el 5 tenga mayor oportunidad de obtener el resultado de la muestra 2001 [una nueva muestra] y pienso que sería un 50% de probabilidad al ser el más repetido el 5.

Patrón 6. Uso de las frecuencias (6/12 estudiantes). Se estima la probabilidad de un evento calculando su frecuencia relativa; se apoyan en la tabla que despliega el software. Por ejemplo, el estudiante A7 responde así:

> Tiene una probabilidad de 24.65% ya que, volviendo a la regla de tres, tengo en mi tabla de muestreo que el número cinco salió 493 ocasiones y la cantidad de muestras (2,000 veces) es mi 100%, por lo tanto tengo ese porcentaje como las repeticiones del número cinco y lo probable de que salga en esa ocasión.

Como vemos en la respuesta del estudiante A9, clasificada en el patrón 5, sólo se menciona a un conjunto modal y se infiere que esos valores son los más probables, no obstante, no se entiende cómo concluye que la probabilidad pedida es 50%. Consideramos que es probable que tales estudiantes hagan un tanteo cualitativo, aunque también puede ser que repitan un razonamiento que recuerda al patrón 4, en el que el 50%, más bien, era una estimación del porcentaje de la población. Pero, en general, las respuestas que clasificamos en este patrón se caracterizan por no utilizar las frecuencias que exhibe la distribución sino por hacer estimaciones subjetivas e idiosincráticas; hay que recordar que ellos sabían cómo recuperar las frecuencias con en el software, pero no utilizaron este conocimiento.

En cambio, los estudiantes cuyas respuestas se clasifican en el patrón 6, leen correctamente la distribución y estiman la probabilidad con base en las frecuencias. Este logro es importante pues refleja que, comienzan a entender la función de la DME, como un instrumento para evaluar la probabilidad de que el estadístico tome determinados valores; con esto, van captando el sentido normativo de una distribución muestral. Esto nos sugiere definir dos concepciones asociadas a una DME, a saber: *concepción descriptiva* (no probabilística) y *concepción frecuencial-probabilística* de la DME.

5.4 Tarea C. Identificación de una DME

Patrón 7. Enfoque en la asimetría de la DME (5/12 estudiantes). Buscan la distribución que represente la asimetría de la población, por lo que eligen la opción ii o iv. Por ejemplo, el estudiante A7 respondió:

> La cantidad de frijoles negros es menos a la de frijoles pintos, por lo que las proporciones que se pueden obtener deben de ser bajas y la gráfica dos representa proporciones bajas, por lo mismo de que, como ya dije, los frijoles pintos les ganan en cantidad, por una gran diferencia a los frijoles negros.

Patrón 8. Semejanza con experiencia anterior (1/12 estudiantes). Elige la opción iv, cuya distribución tiene forma acampanada, pues observó en experiencias previas esta propiedad. A1 fue el único estudiante que argumentó así:

> Conforme a los cálculos que hemos hecho en el anterior experimento, lo más próximo a los resultados que se obtuvieron en este, están en la gráfica iv, deduzco que al ser el mismo procedimiento en el momento de la experimentación los resultados saldrán similares al experimento 1.

Los estudiantes cuya respuesta corresponde al patrón 7 creen que el hecho de que la probabilidad sea diferente de 0.50 implica que la distribución debe ser asimétrica. Consideramos que este patrón obedece a una concepción ingenua del muestreo pues ven en la distribución una representación de la población. Esto es similar a otras observaciones reportadas en la literatura, por ejemplo, Chance *et al.* (2004) consideran dentro de la lista de las falsas concepciones comunes sobre las distribuciones muestrales, la siguiente: "Creer que la distribución muestral debe ser similar a la población (para tamaños de muestra *n*>1)" (p. 302). Por el contrario, el estudiante que sigue el patrón 8 tiene en cuenta la forma acampanada y simétrica de la distribución, aunque solo sea como una generalización de sus observaciones sobre la forma de las distribuciones muestrales previamente simuladas. Esto coloca a este tipo de respuestas a nivel transicional.

6. DISCUSIÓN Y CONCLUSIONES

6.1 Modelo de desarrollo del razonamiento sobre DME de estudiantes de bachillerato

La experiencia llevada a cabo con los estudiantes y el análisis de sus respuestas a las tareas, nos permite destacar algunos contrastes entre respuestas ingenuas y aquellas que tienen algún o algunos rasgos pertinentes, aunque incipientes, con relación al conocimiento normativo que deben ir adquiriendo sobre el tema. En particular, en la tabla 2, se organizan los rasgos acerca de los conceptos de población, muestra, muestreo y DME que diferencian las respuestas más ingenuas de aquellas que contienen rasgos pertinentes al conocimiento normativo. Los conceptos de la cuarta columna son los prescritos por la norma y los que se espera que a la larga logren los estudiantes, pero tales conceptos no fueron encontrados en la población que examinamos. En lo que sigue describimos los componentes de este modelo.

Tabla 2. Modelo de desarrollo del razonamiento sobre la DME

Conceptos	**Niveles de razonamiento**		
	Ingenuo (Aditivo)	**Transicional**	**Multiplicativa y Distribucional**
Población	Material	Simulada	Matemática
Muestra	Real	Simulada	Real/Abstracta
Muestreo	Sin remplazo	Con remplazo	Con remplazo
DME	Descriptiva	Frecuencial/ Probabilística	Probabilística/ Estándar de variación

Población. La diferencia señalada por Ainley *et al.* (2015) entre población material y población matemática (i.e. Una distribución) es fundamental para entender que la actividad de simulación implica desprenderse (al menos al principio) del concepto de población material, para entender que la población es una distribución. Varios estudiantes, razonablemente, pretenden trasladar razonamientos derivados a los procedimientos que hicieron con los cuadritos blancos y negros a los procedimientos simulados. No obstante, comenzar a trabajar con la simulación computacional puede ser útil para transitar a la idea de población matemática, siempre que haya consciencia de que es un paso delicado y no se trate como si fuera automático.

Muestras. Una clara diferenciación entre una muestra real obtenida de la población material, y las muestras repetidas mediante un proceso de simulación es necesaria para entender cómo funciona el proceso de inferencia. Una muestra real implica una incursión en el campo, aunque, en nuestro caso, la incursión al campo consista simplemente en ir al recipiente y sacar una muestra de cuadritos. En cambio, el proceso de simulación implica una distribución (población matemática) que no ofrece información sobre la población real, sino información general sobre la variabilidad muestral. En los patrones 1 y 3 se piensa en muestras reales, que unidas van a dar una muestra grande que se considera representativa.

Muestreo. Una intuición común es que entre más grande sea una muestra reflejará mejor las características de la población material, es decir, será más representativa (Watson y Moritz, 2000). Con esta idea, se puede concebir que un muestreo repetido (de muestras reales) permite obtener una gran muestra representativa; pero para construir la muestra grande a partir de muestras pequeñas, estas tienen que ser ajenas, de ahí que parezca natural que el muestreo adecuado sea sin remplazo; esta manera de concebir el muestro es consistente con la idea de muestra aditiva de Saldanha y Thompson (2002). En cambio, detrás del muestreo con remplazo hay ideas matemáticas más profundas.

El modelo matemático de una muestra (obtenida mediante un procedimiento con remplazo) es el de un punto en un espacio n-dimensional, idea que fue propuesta por el estadístico E. Pearson a J. Neyman (previamente había sido utilizada por R. Fisher) y es significativo que a Neyman le haya costado un poco asimilarla y aceptarla (Lehman, 2011). Pero sin ir tan lejos, en cualquier software es relativamente fácil generar muestras de una población matemática (distribución) si el muestreo es con remplazo, mientras que un muestreo sin remplazo se complica mucho. La inclinación de los estudiantes hacia el muestreo sin remplazo se explica porque consideran poblaciones materiales y la dificultad de darle sentido al muestreo con remplazo en este contexto.

DME. Para que los estudiantes le den un sentido (tendiendo al normativo) al concepto de DME deben transitar de concepciones concretas de las nociones de población, muestra y muestreo, apoyadas por el sentido común, a concepciones matemáticas (más abstractas) sustentadas por el concepto de distribución. La concepción de DME más ingenua, pero con sentido, que se revela en las respuestas de los estudiantes es la de una organización de las muestras que ofrece información sobre la población; en el caso presente, ven en la DME información que permite contar cuantos cuadritos negros se obtuvieron en el conjunto de muestras generadas (patrón 3).

En cambio, otros ven en la DME información sobre las proporciones más frecuentes (patrón 5), aunque sin asociarlo con la estimación de probabilidades de eventos relacionados con valores precisos del estadístico. Afortunadamente, se pudo observar que la mitad de los equipos ven las frecuencias de los valores del estadístico en la DME y las utilizan para estimar probabilidades de obtener un valor del estadístico en una muestra o la probabilidad de un evento formado por valores del estadístico. Las concepciones de la DME descritas las hemos llamado, respectivamente, concepción descriptiva y concepción frecuencial/probabilística.

6.2 Comparación con otros modelos de desarrollo

Aunque el modelo de Chance *et al.* (2004) y el aquí presentado tienen en común que el objetivo es el desarrollo de la noción de DME, no son del todo comparables porque los sujetos que ellos estudiaron eran estudiantes universitarios y en el presente estudio eran de bachillerato. En parte por esta razón, ellos se centran en tres propiedades de las distribuciones muestrales derivadas del *Teorema Central de Límite* y nosotros nos ocupamos de las nociones previas más básicas de población, muestra y muestreo.

La definición de los niveles en el modelo de Chance *et al.* (2004) se describen en términos de conocimiento y comprensión de dichas propiedades y de las relaciones entre ellas. En cambio, nuestro modelo es más abstracto y menos descriptivo, pues solo contrastamos concepciones de esas nociones que marcan un avance: Población material vs. Matemática, Muestras reales vs. Abstractas y Muestreo sin remplazo vs. Con remplazo. Quizá estas diferencias no son tan cruciales en el nivel universitario, pero parecen importantes para que los estudiantes transiten a un concepto abstracto de distribución muestral.

Nuestra propuesta tampoco es del todo comparable con el marco de Noll y Shaughnessy (2012) pero las semejanzas y diferencias pueden ser ilustrativas. Coinciden en que se examinaron estudiantes de nivel bachillerato (aunque ellos incluyeron también de secundaria), que el tema es el razonamiento alrededor de muestreo repetido y que en ambos hubo episodios de instrucción. También coinciden en que la población en las tareas es material, pero una diferencia es que en nuestro estudio nos dirigimos principalmente a preguntas asociadas a la DME generadas por un proceso de simulación; así, transitamos rápidamente de una población material a una matemática.

Noll y Shaughnessy (2012) centran la atención en los tipos de razonamiento sobre la variabilidad; en nuestro caso, al enfrentar a los estudiantes a preguntas dirigidas hacia la DME, no emergen sus concepciones espontáneas sobre variabilidad muestral, pues esta ya está representada en la DME. Es decir, en nuestras tareas los estudiantes no tienen que reflexionar sobre cómo varía el estadístico, sino interpretar la variabilidad a partir de la DME. Así, los términos aditivo y transicional tienen contenido muy distinto en cada modelo, pero podrían ser complementarios.

6.3 Contraste de resultados con el diseño de la Trayectoria de Aprendizaje

En el primer momento de la Tarea A, la concepción aditiva de nuestros estudiantes acerca de las relaciones entre los conceptos de población, muestra y muestreo, produce que las hipótesis de aprendizaje H_1 y H_2 no se cumplan; la mayoría aplica un método de muestreo que no se corresponde con el método del muestreo repetido, sino con la idea de generar una muestra representativa vía la conformación de una muestra grande, con ello, evitan lidiar con la variabilidad muestral presente en los valores del estadístico y considerarla para elaborar la estimación de la proporción desconocida. En cambio, en el segundo momento de la tarea, todas las hipótesis de aprendizaje se validan en mayor medida dado que, si bien pocos estudiantes mantienen la estrategia de recuperar una muestra grande a partir de la DME, la mayoría es receptiva a utilizar la moda o grupo modal de la DME para elaborar la estimación de la proporción. Este cambio notorio en las respuestas de los estudiantes sugiere que la estrategia de animarlos a explorar el muestreo por cuenta propia y luego instruirlos para la construcción de una DME, puede resultar benéfica como estrategia didáctica orientada a confrontarlos con la noción de variabilidad muestral y así transitar hacia una concepción multiplicativa sobre el muestreo repetido.

Respecto a la tarea B, estimación de una probabilidad a partir de la DME, los resultados sugieren que la hipótesis de aprendizaje que se cumple parcialmente es el uso de las frecuencias relativas de los valores del estadístico, que se extraen de la DME, para brindar la estimación de la probabilidad solicitada (H_9). Este hecho es razonable de presentarse dada la dificultad observada en estudiantes de bachillerato para hacer la asociación de manera intuitiva o espontánea (Jacob y Doerr, 2013; van Dijke-Droogers *et al.*, 2019). En este sentido, en vista de que no se plantearon alternativas o cuestionamientos adicionales para que los estudiantes fueran inducidos a realizar dicha asociación, el rediseño de la trayectoria debiera contemplarlos para que la transición hacia una concepción probabilista de la DME sea facilitada.

Respecto a la tarea C, identificación de una DME, destacamos que la hipótesis de aprendizaje que se cumple parcialmente es la H_{11} puesto que los estudiantes asocian un valor de p diferente de .5 a una distribución que no necesariamente es simétrica. Aunque varios llegan a seleccionar la representación correcta de la DME, sus respuestas sugieren que se guían con una sola propiedad del concepto (el promedio de la DME es una aproximación a la proporción o que la DME suele tener forma acampanada) y no con ambas a la vez. Este hecho es razonable de observar si consideramos que nuestros estudiantes apenas inician el estudio del concepto de DME. En consecuencia, sería necesario que continúen con la exploración del concepto, en particular sobre el rol de los parámetros *n* y p, en aras de que transiten hacia una concepción más consolidada y abstracta del mismo.

CONCLUSIONES

La primera pregunta de investigación se responde con el modelo de desarrollo del razonamiento sobre la DME (tabla 2). Este se construyó destacando las nociones de población, muestra, muestreo y DME, el cual puede ser útil como punto de partida para caracterizar una trayectoria de aprendizaje y, en general, para el diseño de la enseñanza del tema. Sin embargo, tiene la limitación de no ofrecer descripciones del desarrollo de otras nociones también importantes como aleatoriedad, variabilidad, distribución y probabilidad que también son importantes en la formación del concepto de distribución muestral. Otros marcos como el de Noll y Shaughnessy (2012) pueden complementar el modelo.

Respecto a nuestra segunda pregunta de investigación, destacamos que la concepción aditiva y multiplicativa de la presente investigación estuvo influenciada por el trabajo de Saldanha y Thompson (2002), pero en su desarrollo abandonamos las nociones de muestra aditiva y multiplicativa que ellos sugirieron para sustituirlas por las nociones de muestreo aditivo y multiplicativo. Mientras que para Saldanha y Thompson (2002) una muestra aditiva es "un subconjunto de una población" (p. 257), sin relación con el muestreo repetido, para nosotros el rasgo aditivo de una muestra ocurre precisamente en el muestro repetido sin remplazo; la importancia de esta noción de muestreo sin y con remplazo en el razonamiento de los estudiantes ha pasado desapercibida hasta ahora.

Nosotros observamos que las muestras sin remplazo, tanto a nivel de elementos como a nivel de muestras, tienen sentido para los estudiantes porque al reunirlas obtienen una muestra más grande que se aproxima más a la población. Para Saldanha y Thompson (2002) una muestra multiplicativa es "una versión cuasi proporcional a pequeña escala de la población"; nosotros argumentamos que el carácter multiplicativo de una muestra resulta de un muestreo con remplazo pues este se imbrica con operaciones conjuntistas que permiten representar una muestra como un punto en un espacio n-dimensional; en consecuencia, el cálculo de las muestras posibles se obtiene multiplicando la cardinalidad de la población (en caso de ser finita) tantas veces como el tamaño de la muestra. No hemos profundizado en esta característica del muestreo porque nuestros estudiantes estuvieron lejos aún de utilizar propiedades multiplicativas en sus razonamientos.

Como reflexiones finales, consideramos que los conceptos y el modelo presentado en este estudio tienen el potencial de aclarar y ayudar a explicar algunas de las dificultades que enfrentan los estudiantes en sus razonamientos sobre las distribuciones muestrales; en este sentido podría contribuir a orientar investigaciones futuras en el tema de DME. Una limitación de esta investigación fue partir de una concepción ingenua del razonamiento de los estudiantes acerca del paso de las tareas con manipulativos, al de las tareas con simulación computacional, sin distinguir diferencias que finalmente se revelaron en el curso de la investigación. Esta limitación, que se nos revela ahora, se manifiesta en defectos importantes en los instrumentos de observación, no obstante, nos anima a continuar con otras investigaciones alrededor del mismo tema corrigiendo los errores de diseño detectados.

Agradecimientos

Se agradece el apoyo otorgado por el Consejo Nacional de Ciencia y Tecnología (CONACYT) para estudios doctorales. Número de becario 333940.

REFERENCIAS

Ainley, J. Gould, R., y Pratt, D. (2015). Learning to reason from samples: Commentary from the perspectives of task design and the emergence of “big data”. *Educational Studies in Mathematics*, *88*(3), 405–412. https://doi.org/10.1007/s10649-015-9592-4

Begué, N., Batanero, C., y Gea, M. (2018). Comprensión del valor esperado y variabilidad de la proporción muestral por estudiantes de educación secundaria obligatoria. *Revista Enseñanza de las Ciencias*, *36*(2), 63–79.

Ben-Zvi, D., Bakker, A., y Makar, K. (2015). Learning to reason from samples. *Educational Studies in Mathematics*, *88*(3), 291–303. https://doi.org/10.1007/s10649-015-9593-3

Biggs, J. B., y Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy*. Academic Press.

Birks, M., y Mills, J. (2011). *Grounded theory: A practical guide*. Sage

Burril, G. (2019). Building concept image of fundamental ideas in statistics: The role of technology. En G. Burril y D. Ben-Zvi (Eds.), *Topics and trends in current statistics education research* (pp. 123–149). Springer. https://doi.org/10.1007/978-3-030-03472-6_6

Chance, B., Ben-Zvi, D., Garfield, J., y Medina, E. (2007). The role of technology in improving student learning of statistics. *Technology Innovations in Statistics Education*, *1*(1). https://doi.org/10.5070/T511000026

Chance, B., delMas, R. C., y Garfield, J. (2004). Reasoning about sampling distributions. En D. Ben-Zvi y J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 295–323). Kluwer Academic Publishers.

Finzer, W. (2001). *Fathom Dynamic Statistics* (v. 2.1). Key Curriculum Press.

Garfield, J. B., y Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching practice*. Springer. https://doi.org/10.1007/978-1-4020-8383-9

Hesterberg, T. C. (1998), Simulation and bootstrapping for teaching statistics. *American Statistical Association. Proceedings of the Section on Statistical Education* (pp. 44–52). American Statistical Association. Recuperado de https://drive.google.com/file/d/1Q-QF2dIsaQ18y8e4JaM8-zsU-JLyeGQLk/

Inzunza, S., e Islas, E. (2019). Análisis de una trayectoria de aprendizaje para desarrollar razonamiento sobre muestras, variabilidad y distribuciones muestrales. *Educación Matemática*, *31*(3), 203–230. https://doi.org/10.24844/em3103.08

Jacob, B. L., y Doerr, H. M. (2013). Students' informal inferential reasoning when working with the sampling distribution. En B. Ubuz, C. Haser, y M. A. Mariotti (Eds.), *Proceedings of the 8th Congress of the European Society for Research in Mathematics Education* (pp. 829–839). CERME.

Johnson, R. W. (2001). An introduction to the bootstrap. *Teaching Statistics*, *23*(2), 49–54. https://doi.org/10.1111/1467-9639.00050

Kahneman, D., Slovic, P., y Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press. https://doi.org/10.1017/CBO9780511809477

Langrall, C. W., Makar, K., Nilsson, P., y Shaughnessy, J. M. (2017). Teaching and learning probability and statistics: An integrated perspective. En J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 490–525). National Council of Teachers of Mathematics.

Lehman, E. L. (2011). *Fisher, Neyman, and the creation of classical statistics*. Springer. https://doi.org/10.1007/978-1-4419-9500-1

Noll, J., y Shaughnessy, J. M. (2012). Aspects of students' reasoning about variation in empirical sampling distributions. *Journal for Research in Mathematics Education*, *43*(5), 509–556. https://doi.org/10.5951/jresematheduc.43.5.0509

Ricketts, C., y Berry, J. (1994). Teaching statistics through resampling. *Teaching Statistics*, *16*(2), 41–44.

Rubin, A., Bruce, B., y Tenney, Y. (1990). Learning about Sampling: Trouble at the core of Statistics. En D. Vere-Jones (Ed.), *Proceedings of 3rd International Conference of Teaching Statistics* (pp. 314–319). ICOTS.

Saldanha, L., y Thompson, P. (2002). Conceptions of sample and their relationship to statistical inference. *Educational Studies in Mathematics*, *51*(3), 257–270. https://doi.org/10.1023/A:1023692604014

Shaughnessy, J. M. (1992). Research in probability and statistics: Reflections and directions. En D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 465–494). Macmillan.

Taffe, J., y Garnham, N. (1996). Resampling, the bootstrap and Minitab. *Teaching Statistics*, *18*(1), 24–25. http://doi.org/10.1111/j.1467-9639.1996.tb00891.x

van Dijke-Droogers, M., Drijvers, P., y Bakker, A. (2019). Repeated sampling with a black box to make informal statistical inference accessible. *Mathematical Thinking and Learning*, *21*(2), 1–23. https://doi.org/10.1080/10986065.2019.1617025

Watson J. M., y Chance, B. (2012). Building intuitions about statistical inference based on resampling. *Australian Senior Mathematics Journal*, *26*(1), 6–18.

Watson, J. M., y Moritz J. B. (2000). Developing concepts of sampling. *Journal for Research in Mathematics Education*, *31*(1), 44–70. https://doi.org/10.2307/749819

Eleazar Silvestre Castro

Dirección: Departamento de Matemáticas, Universidad de Sonora,
Boulevard Luis Encinas y Rosales s/n Col. Centro,
C.P. 83000, Hermosillo, Sonora, México