

Análisis de una tarea dirigida a introducir el muestreo y distribuciones muestrales en el bachillerato

Analysis of a task aimed to introduce sampling and sampling distributions in high school

Eleazar Silvestre Castro¹, Ernesto Sánchez¹ y Carmen Batanero²

¹CINVESTAV-IPN, ²Universidad de Granada

Resumen

Se describe la tarea inicial de una Trayectoria Hipotética de Aprendizaje sobre muestreo y distribución muestral dirigida a estudiantes mexicanos de bachillerato (17-19 años). Se utilizan elementos teóricos del Enfoque Ontosemiótico para identificar los objetos matemáticos implícitos en la tarea, así como para analizar algunas respuestas de los estudiantes.

Palabras clave: Muestreo, distribuciones muestrales, trayectoria hipotética de aprendizaje, enfoque ontosemiótico, bachillerato.

Abstract

We describe the initial task of a Hypothetical Learning Trajectory about sampling and sampling distributions directed to Mexican high school students (17-19 years-old). Some theoretical tools from the Onto-Semiotic Approach are used to identify the mathematical objects embedded in the task, and to analyse some students' responses.

Keywords: Sampling, sampling distributions, hypothetical learning trajectory, onto-semiotic approach, high school.

1. Introducción

En las últimas décadas se han realizado diversas investigaciones que documentan errores y limitaciones relacionadas con la inferencia (descritas, por ejemplo, en Harradine, Batanero y Rossman, 2011). En particular, se describen creencias e intuiciones erróneas sobre muestreo e inferencia como lo son la heurística de la representatividad, ley de los números pequeños, entre otras. Algunos investigadores (Chance, delMas y Garfield, 2004; Liu y Thompson, 2007) también destacan que una comprensión deficiente o limitada del muestreo y distribución muestral (DM) dificulta el desarrollo de un razonamiento inferencial.

La enseñanza en México no ayuda a superar estos conflictos, pues suele enfocarse en cálculos y procedimientos; además el currículum matemático pre-universitario incorpora pocos contenidos estadísticos y bajo un enfoque tradicional. Para incidir en esta problemática, algunos autores han sugerido versiones informales de estos métodos (Zieffler, Garfield, delMas y Reading; 2008), con el fin de que el estudiante se familiarice con la lógica subyacente en la inferencia, basadas en software especializado. Este es el enfoque tomado en este trabajo, que es parte de una investigación en desarrollo sobre el razonamiento que exhiben los estudiantes mexicanos de bachillerato en situaciones que requieren trabajar con la DM empírica utilizando Fathom (que permite construir simulaciones dinámicas). En particular, se analiza la tarea inicial de

Silvestre, E., Sánchez, E. y Batanero, C. (2017). Análisis de una tarea dirigida a introducir el muestreo y distribuciones muestrales en el bachillerato. En J. M. Contreras, P. Arteaga, G. R. Cañadas, M. M. Gea, B. Giacomone y M. M. López-Martín (Eds.), *Actas del Segundo Congreso Internacional Virtual sobre el Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos*. Disponible en, enfoqueontosemiotico.ugr.es/civeos.html

una *Trayectoria Hipotética de Aprendizaje (THA)* utilizando el *Enfoque Ontosemiótico del Conocimiento y la Instrucción Matemáticos (EOS)* para el análisis de los objetos matemáticos implicados en la misma, así como para describir y analizar brevemente las respuestas obtenidas por parte de los estudiantes.

2. Marco teórico

Una THA es una herramienta constructivista que permite asistir tanto procesos de investigación como de enseñanza y aprendizaje de conceptos matemáticos (Simon, 1995; Simon y Tzur, 2004). Para un conjunto de objetos matemáticos que el estudiante debe aprender, una THA es un vehículo de planeación cognitivo-didáctica que describe de forma anticipada su razonamiento en ciertas tareas matemáticas. Se distinguen tres elementos relacionados: a) objetivos de aprendizaje; b) un conjunto de tareas o actividades didácticas; y c) hipótesis sobre el proceso de aprendizaje.

La generación de una THA requiere de la comprensión del conocimiento actual de los estudiantes involucrados; las tareas son un aspecto clave, debido a su papel en el desarrollo de conocimiento matemático; debido a la naturaleza incierta del proceso, el profesor/investigador a menudo se involucra en la revisión de los respectivos componentes (ciclo de enseñanza de las matemáticas). A pesar del posible “éxito” de una THA con un grupo de estudiantes, no existe garantía de que ésta presente el mismo impacto con otros debido a cambios en factores socioculturales e institucionales. Así, una THA no se asume como el único camino hacia la construcción de conocimiento, sino como una apuesta por una vía prometedora de aprendizaje significativo.

El EOS (Godino, 2003; Godino y Batanero, 1994; Godino, Batanero y Font, 2007) es un marco teórico dirigido al análisis y valoración de procesos de enseñanza y aprendizaje de contenidos matemáticos, donde se consideran elementos fundamentales las nociones de práctica, objeto y significado. Una práctica (matemática) es toda aquella expresión (lingüística o no) de un sujeto al enfrentarse a una determinada tarea matemática; un sistema de prácticas asociadas a cierto tipo/campo de problemas constituye el significado de objeto (matemático). Cuando se habla de un sistema de prácticas socialmente aceptadas o compartidas dentro de una comunidad/institución, se habla de significado institucional; cuando el sistema de prácticas es referenciado a un sujeto (prácticas personales), se habla de significado personal. Se consideran seis tipos de objetos primarios interrelacionados entre sí, los cuales intervienen en las prácticas y emergen de las mismas: conceptos, proposiciones, propiedades, argumentos, procesos y lenguaje/representaciones. Todos los objetos primarios que el estudiante posee en cierto momento y que pone en acción al enfrentar una situación problema son considerados objetos intervinientes; aquellos que resultan de un nuevo sistema de prácticas se denominan objetos emergentes.

Existen varios tipos de significado institucional: el sistema de prácticas referentes a las de los expertos (y que determina lo que es un objeto matemático para las instituciones) es denominado significado institucional de referencia; aquellas que el profesor selecciona y delimita para proponer a sus estudiantes es denominado significado institucional pretendido; el que efectivamente tiene lugar en el aula, por ejemplo a través de lecciones o actividades didácticas, se clasifica como significado institucional implementado; y, finalmente, las prácticas implicadas en la solución de las tareas o cuestiones que el profesor selecciona para realizar evaluaciones es el significado institucional evaluado.

3. Método

Contexto y muestra. Los participantes del estudio fueron 42 estudiantes del último año de un bachillerato mexicano; la THA se aplicó al inicio de su curso de Estadística y Probabilidad II (EyP II). Previamente completaron y aprobaron un curso (tradicional) de EyP I cuyo contenido fue: representaciones gráficas, medidas de tendencia central y de variabilidad, enfoque clásico y frecuencial de la probabilidad, cálculo de probabilidad para eventos simples y compuestos. Los datos se recogen en hojas de trabajo electrónicas en donde los estudiantes respondieron los distintos cuestionamientos; observaciones de dos profesores e investigadores (en algunas sesiones); y algunas entrevistas breves durante y después de las actividades.

Resumen de la THA implementada. Los objetivos globales de la THA son: a) promover una concepción estocástica sobre el proceso de muestreo y estimación, basado en la exploración de la DM empírica de proporciones; b) explorar y describir el razonamiento de los estudiantes en las actividades con base en sus estrategias y argumentos. La THA incluye ocho actividades basadas en Fathom y organizadas en tres fases: 1. Introducción a la DM (estimación de probabilidades de obtener ciertos valores muestrales; generación de un método para clasificarlos como probables o no probables de obtener; obtención del promedio de la DM y análisis de su relación con el parámetro poblacional); 2. Análisis de la DM al variar el tamaño de muestra (forma, variabilidad muestral, estimación de probabilidades y valores probables y no probables); 3. Estimación de un porcentaje poblacional desconocido con base en su DM empírica (cambiar el tamaño de muestra y realizar estimaciones sobre el porcentaje poblacional). Una vez aplicada la THA los estudiantes se enfrentaron a una situación problema que implica la realización informal de una prueba de hipótesis, con el objetivo de analizar el grado en que utilizan su nuevo conocimiento (Silvestre y Sánchez, 2016).

4. Análisis de la tarea inicial

Una consideración clave es la falta de interacción de los estudiantes con situaciones de muestreo, estimación e inferencia en su formación previa; lo que sugiere presentarán una alta dependencia de pre-concepciones e intuiciones, a menudo erróneas, al enfrentarse a situaciones que involucran estos conceptos (por ejemplo, concebir una muestra “representativa” como una versión en miniatura de la población). Por ello una posible vía para comenzar el proceso de instrucción es dirigir la THA a re-orientar dichas intuiciones hacia un razonamiento más apropiado. En adición, para la selección y diseño de la actividad que se analiza en este trabajo se han tomado algunos resultados de Saldanha y Thompson (2002, 2007) sobre el razonamiento de estudiantes de bachillerato sobre muestreo e inferencia y algunas heurísticas/sesgos al trabajar con situaciones de incertidumbre (Kahneman, Slovic y Tversky, 1982).

Componentes de las actividades. La tarea analizada engloba dos subtarefas relacionadas. La primera es una situación de estimación de un porcentaje poblacional a través de muestras pequeñas (población física, no involucra la tecnología); en la segunda se analiza la información que puede ser extraída de una DM empírica para esta estimación (uso de Fathom para simular múltiples muestras). Los objetivos de aprendizaje de la primera subtarea son probar de forma empírica las siguientes propiedades: 1.1) la variabilidad muestral hace problemática cualquier afirmación determinista sobre la composición de la población; 1.2) la imposibilidad de predecir con certeza los resultados “favorables” (que coinciden o son muy cercanos al porcentaje poblacional)

genera la necesidad de cuantificar su probabilidad; 1.3) una DM (empírica) es una colección ordenada de valores muestrales que puede proveer dicha información. Los objetivos de aprendizaje de la segunda subtarea son comprobar que: 2.1) dado un número grande de ensayos, la frecuencia relativa de aparición de cierto (s) valor (es) muestral brinda una estimación de su probabilidad; 2.2) los porcentajes muestrales más probables son los de mayor frecuencia en el total de ensayos (muestras); 2.3) la DM puede ser resumida a través del promedio (y en este caso particular la moda).

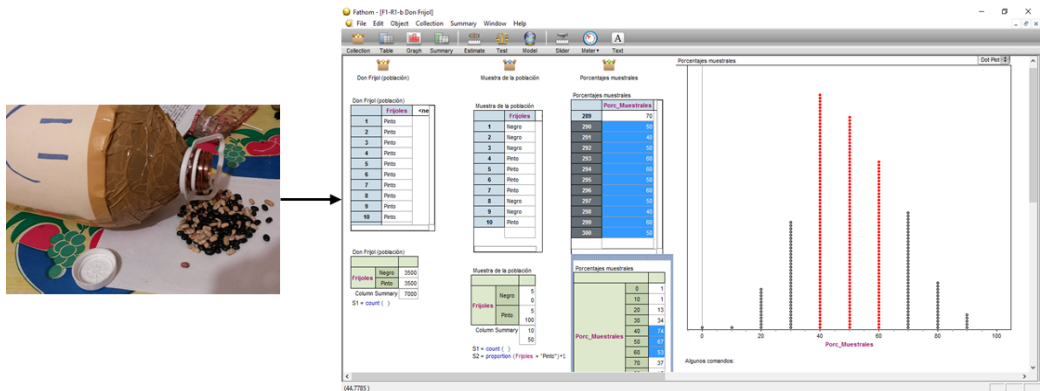


Figura 1. Don Frijol (izquierda); y DM empírica de porcentajes muestrales (derecha).

La primer sub-tarea comienza con la presentación de Don Frijol a los estudiantes: un contenedor que contiene 7,000 frijoles donde el 50% son pintos y el restante 50% son negros (Figura 1). D. Frijol representa una población física dicotómica cuya composición es desconocida para los estudiantes; al comienzo el interés es la estimación del porcentaje poblacional de frijoles pintos con base en una muestra pequeña. Una vez que un estudiante ha extraído una muestra aleatoria de 10 frijoles, la pregunta inicial dirigida a la clase es: “¿es este resultado, x % muestral, suficiente evidencia para asegurar que Don Frijol posee X% de frijoles pintos?” Se inicia una discusión abierta, orientada por el profesor, que comienza por la respuesta *a priori* (casi instantánea y propia de un razonamiento intuitivo) de los estudiantes, rechazando la proposición. Usan mayoritariamente dos argumentos: que la muestra es muy pequeña para confiar en ella y que puede presentarse variación en el porcentaje muestral si se repite el experimento. En el estudio exploratorio se observó que son pocos los estudiantes que pueden justificar apropiadamente cómo y porqué el tamaño de muestra influye. Se continúa la discusión utilizando preguntas que buscan propiciar que el estudiante exhiba sus creencias:

1. Recolecta tantas muestras, de igual tamaño (10), como creas conveniente y registra su porcentaje muestral, ¿de qué manera se puede predecir un nuevo porcentaje muestral? (Identificación de la variabilidad muestral).
2. ¿Qué sugieren los porcentajes recolectados sobre el porcentaje de frijoles pintos en Don Frijol? (Desarrollo de un procedimiento para estimar el porcentaje poblacional con pocas muestras de tamaño pequeño)
3. Con base al porcentaje poblacional revelado a los estudiantes ¿se obtuvieron resultados favorables (iguales o cercanos al porcentaje poblacional) o no?
4. ¿Consideras igual de probables los resultados favorables y no favorables?
5. ¿Cómo podría ser cuantificada esta probabilidad? Por ejemplo, la probabilidad de obtener porcentajes muestrales entre 40%-60%.

6. Imagina que se generan 100 muestras aleatorias de tamaño 10; el porcentaje muestral 20% ha aparecido en 4 ocasiones y el 50% en 36 ocasiones, ¿cuál de estos dos porcentajes muestrales tiene mayor probabilidad de aparecer en la muestra 101? (Estimación de la probabilidad de obtención con un número grande de muestras).

----- Inicio de la segunda subtarea -----

7. Con base en la simulación (simulación de un número grande de muestras y sus respectivos porcentajes muestrales de una población de parámetro 50% mediante Fathom), ¿qué probabilidad asignarías al evento “obtener un porcentaje de 70% en una nueva muestra”?
8. Clasifica todos los posibles porcentajes muestrales como probables o no probables.
9. Da un valor numérico que resuma todos los porcentajes muestrales obtenidos en la simulación.

La primer sub-tarea se inicia según proponen Saldanha y Thompson (2007) pero con un énfasis en los porcentajes muestrales; también la expresión “resultado típico” es reemplazada por “resultado favorable”. Se espera que las preguntas 4 y 5 mostrarán la incapacidad de los estudiantes para cuantificar la probabilidad (Saldanha y Thompson, 2002), por lo que la pregunta 6 provee un medio para intuitivamente realizar esta estimación. Este conjunto de prácticas deriva en la emergencia de una DM empírica (binomial con $n=10$ y $P=.5$). Los ítems de la segunda sub-tarea están dirigidos hacia la exploración y extracción de información a partir de la DM usando Fathom. Dado que la realización de estas tareas implica un cierto ir y venir de respuestas correctas y parcialmente correctas se considera más apropiado hablar de las *hipótesis del proceso de razonamiento* en sustitución de las “hipótesis de aprendizaje”, las cuales son:

1. Con base en la simulación física y la variabilidad muestral, los estudiantes identificarán la imposibilidad de predecir con total certeza un nuevo porcentaje muestral.
2. Los estudiantes no brindarán mucha atención a los altos niveles de variabilidad (colección y tamaño de muestra pequeños) al proponer un porcentaje poblacional.
3. Una vez revelado el porcentaje poblacional, se presentará un conflicto entre la expectativa y no haber obtenido resultados favorables.
4. Los estudiantes reconocerán la necesidad de cuantificar la probabilidad de obtener ciertos porcentajes muestrales, pero difícilmente propondrán un método para esto.
5. El análisis de la frecuencia relativa de aparición permitirá estimar la probabilidad de obtener determinados porcentajes muestrales.
6. Los estudiantes propondrán generar un número grande de muestras para analizar las frecuencias relativas de aparición y así estimar probabilidades.

----- Segunda sesión/sub-tarea -----

7. A partir de la simulación de la DM empírica, los estudiantes estimarán la probabilidad de obtener un cierto porcentaje muestral a partir de la frecuencia relativa de aparición.
8. Clasificarán los porcentajes muestrales con mayor frecuencia como más probables

y los de menor frecuencia como poco probables.

9. Utilizarán el promedio (probablemente calculado de forma incorrecta) y la moda para resumir la DM.

Tabla 1. Tipos de objetos implícitos en las tareas

<i>Tipo</i>	<i>Objeto</i>	<i>Subtarea</i>	
		<i>1</i>	<i>2</i>
<i>Situación problema</i>	Estimar un porcentaje poblacional a través de muestras pequeñas con base en una exploración física.	x	-
	Explorar y caracterizar una DM empírica.	-	x
<i>Conceptos</i>	Población, muestra, tamaño de muestra, parámetro y estimador, variable aleatoria, variable estadística.	x	x
	Experimento aleatorio, resultado, suceso seguro e imposible, casos favorables y posibles. Probabilidad (subjctiva/enfoque frecuencial/estimación).	x	x
	Frecuencias absolutas, porcentajes.	x	x
	Variabilidad muestral.	x	x
	Distribución muestral (binomial/empírica/discreta).	x	x
	Convergencia a la DM teórica.	-	x
	Estimación puntual.	x	-
	Promedio de la DM.	-	x
Ley de los grandes números.	-	x	
<i>Propiedades/argumentos</i>	Con una sola muestra, la mejor estimación del porcentaje de la población es el porcentaje de la muestra.	x	-
	Debido a la variabilidad muestral, la estimación del parámetro basada en una sola muestra es aproximada.	x	-
	A partir de la DM, la mejor estimación del porcentaje poblacional sería la media.	-	x
	A partir de la DM, la mejor estimación del porcentaje poblacional sería la moda si la muestra es pequeña.	-	x
	La frecuencia relativa es una estimación de la probabilidad.	x	x
	En una DM, a mayor frecuencia, mayor probabilidad de obtener cierto valor muestral (y viceversa).	x	x
	La DM empírica se aproxima a la DM teórica.	-	x
	El promedio de la DM se aproxima al parámetro.	-	x
	La frecuencia relativa converge a la probabilidad cuando se aumenta el número de ensayos (muestras).	-	x
La DM se estabiliza cuando aumenta el número de ensayos.	-	x	
<i>Procedimientos</i>	Identifica los elementos en la simulación física.	x	x
	Identifica los elementos en la simulación con Fathom.	-	x
	Genera y registra un conjunto de muestras independientes con sus respectivos porcentajes muestrales.	x	-
	Identifica los porcentajes muestrales y sus frecuencias.	-	x
	Propone un método no ambiguo para clasificar los porcentajes muestrales más y menos probables.	-	x
	Calcula el promedio/moda de la DM.	-	x
	Identifica la variabilidad de los porcentajes muestrales.	x	x
	A partir de la simulación física, propone un método para estimar el porcentaje poblacional.	x	-
	Identifica la variabilidad de la DM en distintas simulaciones de la misma.	-	x
	Estimar probabilidades a través de las frecuencias relativas.	x	x
<i>Lenguaje/</i>	Población, muestra, parámetro y estimador; DM: representaciones	x	x

<i>Representaciones</i>	físicas y numéricas; a partir de la simulación con Fathom, representación tabular y gráfica e icónica. Argumentos y proposiciones: lenguaje verbal, numérico y simbólico.	x	x
-------------------------	--	---	---

Análisis semiótico de las tareas. La Tabla 1 muestra los principales objetos implícitos en la realización de las tareas. Algunos (población, muestra, variable aleatoria, promedio probabilidad) son conocimientos previos (intervenientes) y otros emergen a raíz de la exploración y desarrollo de la situación (estimación de probabilidades, ley de los grandes números y convergencia de la DM empírica a la teórica). Este análisis permite identificar los tipos de objetos y la forma en que intervienen en la THA para complementar el análisis de la tarea implementada o modificarla. Por ejemplo, el lenguaje y las proposiciones/argumentos se introducen de forma que no es necesario dominar las expresiones y fundamentos matemáticos que las sustentan; por lo tanto, el sistema de prácticas que aquí se genera son consideradas de naturaleza informal.

6. Resumen de algunas respuestas a la tarea

Las respuestas a la pregunta 7 se muestran en la Tabla 2: El primer grupo (seis parejas), se caracteriza por haber realizado algún cálculo para estimar la probabilidad de obtener un 70% de frijoles pintos; utilizan las frecuencias relativas y una o más simulaciones de la DM para brindar una estimación. Por ejemplo, la pareja 16 (P16) respondió:

Estimamos que la muestra 301 tiene la probabilidad del 10% al 13% de que salga el 70% de frijoles pintos debido a que al tomar diversas muestras y observar el valor del porcentaje 70, su rango es de 30 a 40 repeticiones que equivale al 10% - 13%.

Tabla 2. Respuestas (procedimientos y argumentos) al ítem 7

	<i>Códigos</i>	<i>Frec.</i>
<i>Realiza cálculos (6 parejas)</i>	Calcula correctamente una estimación de la probabilidad: intervalo o valor particular basado en distintas simulaciones o una sola simulación.	6
<i>No realiza cálculos (16 parejas)</i>	Asigna una probabilidad de 30% o 40% y señala uno de los siguientes atributos: el valor muestral (70%) es menos probable debido a su baja frecuencia; considera que los valores de 40%, 50% y 60% son los de mayor frecuencia: especifica la frecuencia del valor muestral.	6
	No asigna valor numérico y señala uno de los atributos mencionados.	3
	Asigna una probabilidad de 35% o 0% debido ya que considera que los porcentajes muestrales tienden a estos valores.	2
	Otros: Asume la frecuencia máxima del porcentaje muestral como su probabilidad; asigna un 10% de probabilidad debido a que es una de diez posibilidades (de la variable aleatoria); asigna 35% como la probabilidad.	5

Otro grupo (16 parejas) no realiza cálculos; seis parejas de este grupo asignaron un valor de 30% o 40% para la probabilidad de obtener un 70% de frijoles pintos. Además, nueve parejas señalaron algún atributo relacionado a altas o bajas frecuencias. Por ejemplo, P9 mencionó “40% porque en la simulación el valor de 70% no fue un valor constante o que apareciera continuamente en los valores de las muestras anteriores”. De las siete parejas restantes dos mencionaron una supuesta tendencia de los porcentajes muestrales hacia el 0% o 35%, y otras dos asignaron un 10% debido a que el 70% es uno de diez posibles valores de la variable aleatoria.

Las respuestas correspondientes a la pregunta 8 (porcentajes muestrales más probables

de obtener) se muestran en la Tabla 3. 15 parejas clasificaron los porcentajes muestrales como probables o no por sus frecuencias absolutas de aparición, mencionando en su mayoría los valores del 40% al 60%. Sin embargo, 14 no declararon qué criterio debe cumplirse para que estas frecuencias sean consideradas altas. Así, P2 indicó: “40%, 50% y 60% ya que estos porcentajes muestrales, la frecuencia con la que salen es mayor a 55”; mientras que P22 contestó: “El 40%, 50% y 60% tienen mayor probabilidad de aparecer por que en la gráfica de Fathom aparecía con más frecuencia”.

Tabla 3. Procedimientos y argumentos en el ítem 8

	<i>Códigos</i>	<i>Frec.</i>
<i>Argumenta con base en frecuencias (14 parejas)</i>	Porcentajes muestrales de 40%-60% porque presentan las mayores frecuencias (especifica un valor para la frecuencia).	1
	Porcentajes muestrales de 40%-60% porque presentan las mayores frecuencias (señala algunos valores o el intervalo).	13
<i>Se basa en el criterio de cercanía (6 parejas)</i>	Porcentajes muestrales de 40%-60% porque presentan las mayores frecuencias y porque el porcentaje poblacional es igual a 50%.	1
	Porcentajes muestrales de 40%-60% porque son los más cercanos al porcentaje poblacional (señala algunos valores o el intervalo).	4
	Porcentajes muestrales de 40%-60% porque el porcentaje poblacional es de 50%.	1
<i>Otros (2 parejas)</i>	Valores muestrales de 40%-60%, argumento redundante o incoherente.	2

El siguiente grupo mayoritario de respuestas, constituido por seis parejas, utiliza o se basa en el argumento de que, dado que se conoce el porcentaje poblacional (50%), se asume que los porcentajes muestrales más probables de obtener son los que están más cercanos al parámetro poblacional. Este criterio es denominado como criterio de la cercanía y es ejemplificado por P15, que respondió:

40%, 50% y 60% por la cantidad de frijoles pintos y frijoles negros es la misma (50-50) pero no es probable que en todas las muestras salga el 50% de frijoles pintos, así que los valores más cercanos al 50% es 40% y 60%.

Este tipo de argumento es considerado como un sesgo o heurística pues deja de lado el utilizar la información (datos) que se muestra en el simulador; es decir, una limitación derivada de este criterio es que sólo basta conocer el parámetro poblacional para estimar qué valores muestrales se espera obtener en una nueva muestra aleatoria (sin importar el tamaño de la muestra o la desviación estándar de la población).

Por cuestiones de espacio no se incluyen las respuestas correspondientes a los porcentajes muestrales menos probables; sin embargo, las respuestas obtenidas son similares para el caso de los porcentajes más probables. En el caso de los porcentajes muestrales menos probables se mencionan en su mayoría valores de 20% o menores y de 80% o mayores (especial énfasis en el 10% y 90%); sólo una pareja incluyó los valores de 30% y 70% en su respuesta.

Las respuestas de la pregunta 9 se encuentran en la Tabla 4. Diez parejas decidieron utilizar el promedio para resumir/representar los porcentajes muestrales pero sólo tres de éstas lo calcularon de forma correcta. Tres parejas tomaron como datos las frecuencias de algunos porcentajes muestrales (de mayor frecuencia) para aplicar un procedimiento similar al del promedio; tres más plantearon utilizar el promedio pero no lo calcularon. Como ejemplo de estas respuestas, P11 indicó:

Realizamos un promedio de los porcentajes que se repetían con mayor frecuencia y llegamos a la conclusión de que por lo menos el valor porcentual sea 68% de los 300 porcentajes. Los cálculos que se realizaron fue sacar un porcentaje de los porcentajes que se repetían con mayor frecuencia: 40 con 60 veces, 50 con 82 veces y 60 con 62 veces. Cálculo: $60 + 82 + 62 = 204 / 300 = 0.68 * 100 = 68\%$.

Tabla 4. Procedimientos y argumentos en el ítem 9.

	<i>Código</i>	<i>Frec.</i>
<i>Utiliza (o propone) el promedio de la DM (10 parejas)</i>	Cálculo correcto.	3
	Cálculo incorrecto, aplica el método a frecuencias porcentajes muestrales con mayor frecuencia.	2
	Cálculo incorrecto, aplica una "regla de tres" con base en un solo porcentaje muestral de alta frecuencia.	1
	Cálculo incorrecto, no proporciona el método. Propone utilizar el promedio, no lo calcula.	1 3
<i>Propone utilizar la moda de la DM (10 parejas)</i>	Utiliza la moda como representante (50%) debido a su mayor frecuencia.	6
	Utiliza la moda sesgada como representante (porcentajes mayores a 50% y que no forman parte de la variable aleatoria, considera con base en sus simulaciones que ahí se presentan mayores frecuencias).	3
	Utiliza la moda sesgada como representante (menor al 50% debido a que hay una alta frecuencia de muestras con esos porcentajes).	1
<i>Otros (2 parejas)</i>	Porcentajes de 55%/60%, no brinda argumentos o justifica de forma imprecisa.	2

Diez parejas propusieron utilizar la moda de la DM pero cuatro realizaron un cálculo incorrecto; estas parejas propusieron una moda "sesgada" al proponer un porcentaje muestral diferente de 50% y que no están incluidos en la variable aleatoria. Por ejemplo, P21 respondió "De los 300 porcentajes simulados nosotros estimamos que el valor que más se repite es de 54% pues ninguna muestra es igual a la anterior y la mayoría de las veces sale mayor a 50%".

7. Conclusiones y discusión

Puede considerarse que las respuestas descritas muestran el razonamiento de los estudiantes sobre nociones de muestreo, estimación y probabilidad antes de la enseñanza. A pesar de que los estudiantes no habían sido instruidos previamente en este contenido, puede notarse que el significado identificado en su razonamiento no es del todo incorrecto (por ejemplo, el cálculo de estimaciones para la probabilidad, uso del promedio y moda para representar los porcentajes muestrales). Dada la complejidad involucrada en la enseñanza de la Estadística y Probabilidad, analizar el razonamiento de los estudiantes constituye un posible primer acercamiento para el profesor e investigador que desea incidir y reorientar estas prácticas dentro de un proceso de instrucción o formación. En general, el uso de la THA nos ha permitido diseñar y plantear situaciones problema que promueven acciones ostensibles de los estudiantes, lo que permite analizar su significado personal a través de ciertas prácticas. El uso del EOS nos ayuda a clarificar qué objetos, tanto pretendidos e implementados, se encuentran implícitos en la THA y en el razonamiento que los estudiantes exhiben. Posibles aplicaciones de este tipo de análisis son el robustecimiento de la comprensión del funcionamiento y estructura de la actividad misma, así como sobre las implicaciones en el desarrollo del significado generado en los estudiantes.

Reconocimiento: Trabajo realizado parcialmente en el marco del Proyecto EDU2016-74848-P (MEC)

Referencias

- 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). Amsterdam: Kluwer Academic Publishers.
- Godino, J. D. (2003). Teoría de las funciones semióticas. *Departamento de Didáctica de la Matemática. Universidad de Granada*. [en línea]. Recuperable en <http://www.ugr.es/local/jgodino>.
- Godino, J. D. y Batanero, C. (1994). Significado personal e institucional de los objetos matemáticos. *Recherches en Didactiques des Mathématiques*, 14(3), 325-355.
- Godino, J.D., Batanero, C. y Font, V. (2007). The onto-semiotic approach to research in mathematics education *ZDM, The International Journal on Mathematics Education*, 39(1-2) 127-135.
- Harradine, A., Batanero, C. y Rossman, A. (2011). Students and teachers' knowledge of sampling and inference. En C. Batanero, G. Burrill y C. Reading (Eds.), *Teaching statistics in school mathematics. Challenges for teaching and teacher education* (pp. 235-246). New York: Springer.
- Kahneman, D., Slovic, P. y Tversky, A. (1982). *Judgment under uncertainty: Heuristics and biases*. New York: Cambridge University Press.
- Liu, Y. y Thompson, P. (2007) Teachers' Understandings of Probability, *Cognition and Instruction*, 25(2-3), 113-160.
- Saldanha. L., y Thompson, P. (2002) Conceptions of sample and their relationship to statistical inference. *Educational Studies in Mathematics*, 51, 257-270.
- Saldanha. L., y Thompson, P. (2007) Exploring connections between sampling distributions and statistical inference: an analysis of students' engagement and thinking in the context of instruction involving repeated sampling. *International Electronic Journal of Mathematics Education*, 3, 270-297.
- Silvestre, E. y Sanchez, E. (2016). Patrones en el desarrollo del razonamiento inferencial informal: introducción a las pruebas de significancia en el bachillerato. En *Memorias del XX Simposio de la Sociedad Española de Investigación en Educación Matemática (SEIEM XX)*, Málaga, España, 2016.
- Simon, M. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education*, 26, 114–145.
- Simon, M., y Tzur R. (2004). Explicating the Role of Mathematical Tasks in Conceptual Learning: An Elaboration of the Hypothetical Learning Trajectory. *Mathematical Thinking and Learning*, 6(2), 91-104.
- Zieffler, A., Garfield, J., delMas, R. y Reading, C. (2008). A framework to support research on informal inferential reasoning. *Statistics Education Research Journal*, 7(2), 40–58.