Experiencia de enseñanza para fomentar el razonamiento inferencial informal de estudiantes de secundaria

Juan Carlos Huerta Orozco[1] y Ana Luisa Gómez Blancarte[2]

RESUMEN

El objetivo de este artículo es mostrar cómo estudiantes de secundaria pueden desarrollar elementos de un razonamiento inferencial informal en una experiencia de enseñanza con el apoyo de la tecnología. Se describe el diseño e implementación de una secuencia didáctica, desarrollada bajo la heurística de crecer la muestra, a fin de que los estudiantes construyeran argumentos para responder a la pregunta: *¿Quiénes tienen más riesgo de morir por COVID-19 en México, hombres o mujeres?* Los resultados ilustran los modelos de argumentación de los estudiantes y ofrecen un entendimiento de los elementos del razonamiento inferencial informal que surgieron. Se concluye el potencial de la experiencia de enseñanza para fomentar en los estudiantes la articulación de argumentos basados en generalizaciones, evidencias que integran un conocimiento estadístico y del contexto, uso de expresiones de incertidumbre y la aparición de principios generales como consecuencia de la heurística de crecer la muestra.

PALABRAS CLAVE

Razonamiento inferencial informal, Inferencia estadística, Educación secundaria, Heurística de crecer la muestra, Software Fathom.

---

[1] juan.huerta.oro@chih.nuevaescuela.mx
Escuela Secundaria Técnica #19
https://orcid.org/0009-0004-7317-1510

[2] algomezb@ipn.mx
Instituto Politécnico Nacional
https://orcid.org/0000-0002-8837-8643

Huerta Orozco, J. C., & Gómez Blancarte, A. L. (2023). Experiencia de enseñanza para fomentar el razonamiento inferencial informal de estudiantes de secundaria. En A. Castañeda, (Ed.), *Aportes y recursos para la innovación en la educación matemática* (pp. 123–146). SOMIDEM Editorial.
https://doi.org/10.24844/SOMIDEM/S1/2023/01-05

ANTECEDENTES

La inferencia estadística provee de herramientas tanto para generar argumentos válidos como para generalizar un aspecto de una población a partir del análisis de una muestra tomada de esta (Zieffler et al., 2008). La investigación en educación estadística ha mostrado que el estudio de la inferencia estadística es difícil para los alumnos, pues demanda conocimientos de diferentes conceptos estadísticos y de una lógica probabilística (Makar & Rubin, 2018). Por ello, enseñar un enfoque informal de inferencia estadística ha sido sugerido como una estrategia para apoyar a los alumnos de educación básica a construir bases de un entendimiento de ideas formales de inferencia estadística (Tobías-Lara & Gómez-Blancarte, 2019). Dicho enfoque informal implica enseñar conceptos de inferencia desde una "versión simplificada" para que los estudiantes puedan comprenderlos (Makar, 2013), así como centrarse en fomentar un razonamiento inferencial informal (RII).

Promover un RII desde edades tempranas favorece en los estudiantes la posibilidad de debatir empleando términos estadísticos (centro, variación y distribución) a un nivel apropiado para su edad, y de usar los datos como evidencia para generar conclusiones con cierto nivel de incertidumbre (Makar, 2013). De acuerdo con Ben-Zvi et al. (2012), "prestar atención a las articulaciones emergentes de incertidumbre de los alumnos al emitir juicios sobre patrones y tendencias en los datos puede brindar la oportunidad de desarrollar comprensiones más sofisticadas de la inferencia estadística" (p. 913).

Las tareas que pueden contribuir al desarrollo del RII involucran: (1) estimar o dibujar la gráfica de una población a partir de los datos de una muestra, (2) comparar dos muestras de datos para inferir si hay una diferencia real entre sus poblaciones y (3) juzgar cuál de dos modelos o argumentos es más probable que sea verdad (Zieffler et al., 2008). Estos tipos de tareas buscan que el alumno presente inferencias hacia una población a partir de los datos de una muestra y priorizan la presentación de argumentos para soportar las inferencias expuestas. Se pueden desarrollar dentro de actividades que plantean preguntas de investigación sobre temas diversos a fin de que los estudiantes las respondan mediante un ciclo de investigación (e.g., Makar, 2013; Pfannkuch, 2011). Es necesario tomar en cuenta factores importantes que impactan en el desarrollo del RII, como son el contexto de los datos –conocimiento del contexto del problema–, el contexto de la experiencia de aprendizaje –los antecedentes del estudiante sobre la tarea y el ambiente físico y social en que se trabaja– (Pfannkuch, 2011), y la estructura de la pregunta de investigación (Makar, 2013).

Además de las tareas y actividades, la tecnología es otro factor importante para fomentar un RII. Los programas de cómputo Tinkerplots y Fathom son dos herramientas tecnológicas destacadas en la educación estadística

que apoyan ese razonamiento. Según Henriques y Oliveira (2016), la construcción de gráficos en Tinkerplots coadyuvó a que los alumnos de educación secundaria "generalizaran a partir de las características observadas usando los datos como evidencia en lugar de basar sus inferencias en su experiencia personal" (p. 76). Por su parte, Hernández Trevethan (2018) señala que el uso combinado de ambos programas (Tinkerplots y Fathom) apoyó el entendimiento y aplicación, por parte de estudiantes de bachillerato, de temas de inferencia estadística (e.g., pruebas de ANOVA e independencia) desde un enfoque informal. No obstante, cabe aclarar que el uso de la tecnología por sí misma no garantiza un entendimiento y razonamiento de conceptos abstractos de estadística (Kula & Koçer, 2020). En este sentido, es importante considerar que los recursos tecnológicos en la enseñanza y aprendizaje de la inferencia estadística deben usarse para promover un entendimiento de los conceptos estadísticos y no del uso mismo del software.

PLANTEAMIENTO DEL PROBLEMA

Dada la dificultad para entender nociones de inferencia estadística por parte de estudiantes de niveles superiores, diversos investigadores han atendido la sugerencia de promover un RII en estudiantes de educación preuniversitaria (e.g., Ben-Zvi et al., 2012; Henriques & Oliveira, 2016; Makar, 2013; Pfannkuch, 2011; Zieffler et al., 2008). En México, este tipo de investigaciones también son necesarias, pues los estudiantes de Educación Superior (ES), incluso los de Educación Media Superior (EMS), presentan dificultades en temas de inferencia estadística como lo son las pruebas de hipótesis (e.g., Silvestre Castro & Urrea Bernal, 2021). Además, el enfoque de enseñanza de la estadística que suelen fomentar tanto los programas de estudio como los profesores que la imparten, con pocas excepciones, no favorece un razonamiento inferencial estadístico.

El estudio de la inferencia estadística tiene lugar en la mayoría de los cursos de estadística que se ofertan en la ES y en pocas asignaturas de Probabilidad y Estadística de la EMS. De acuerdo con Gómez Blancarte et al. (2022), de ocho programas de estudio de dicha asignatura que se usan en instituciones públicas representativas de la EMS del país, solo dos de ellos fomentan algunos elementos del razonamiento inferencial, y uno de los dos programas oferta la asignatura de manera optativa. En consecuencia, es muy posible que la mayoría de los estudiantes de EMS tengan dificultades en temas de inferencia estadística cuando ingresan a la ES.

Los estudios de Cruz López y Ojeda Ramírez (2021) y Pinto Sosa (2020) dan cuenta de que los programas de estudio de la ES para la enseñanza de la estadística suelen enfatizar los procedimientos estadísticos y cálculos matemáticos (e.g., realizar gráficos y cálculos de medidas descriptivas). Cruz López y Ojeda Ramírez analizaron 410 programas de cursos de esta-

dística de diversas instituciones mexicanas de ES. Los autores reportan una omisión considerable de aspectos importantes para resolver problemas mediante la investigación estadística (e.g., formular preguntas de investigación, plantear preguntas acerca de las generalizaciones hechas a una muestra, discutir resultados basados en datos) y una carencia en la promoción de habilidades necesarias para transmitir resultados y construir argumentos basados en datos. Tanto el proceso de investigación estadística como las habilidades de argumentación son claves para desarrollar un razonamiento inferencial.

Por su parte, Rivera Pérez y Gómez Blancarte (2022), en su estudio, señalan que los profesores de ES parecen enfocar su enseñanza en la aplicación de métodos estadísticos para resolver problemas, pero se centran más en el análisis de datos sin tomar en cuenta el planteamiento de preguntas de investigación ni la interpretación de los resultados estadísticos para emitir conclusiones. En este sentido, lo que hacen los profesores participantes en dicho estudio coincide con las propuestas curriculares de los programas analizados por Cruz López y Ojeda Ramírez (2021). Además, el estudio de Sepúlveda Vega y Sánchez Sánchez (2022) muestra que algunos profesores de EMS desconocen el enfoque informal de inferencia estadística para resolver problemas de significación estadística y para interpretar el razonamiento que subyace en la solución, por parte de sus estudiantes, a problemas similares.

Desarrollar un RII es uno de los enfoques que se sugieren para apoyar el entendimiento formal de nociones de inferencia estadística. Así pues, son necesarios estudios que ofrezcan al profesor estrategias para promover el RII de sus estudiantes. Derivado de una investigación más amplia (ver Huerta Orozco, 2021), el presente estudio reporta una experiencia de enseñanza, basada en la heurística de "crecer la muestra" (Ben-Zvi et al., 2018) y en el uso de la tecnología, para fomentar el RII en estudiantes desde nivel secundaria. El objetivo es analizar los argumentos de los estudiantes cuando resuelven una situación inferencial, a fin de responder a la pregunta: ¿Qué elementos del RII desarrollan estudiantes de secundaria cuando resuelven una situación inferencial con apoyo de la tecnología?

MARCO TEÓRICO

De acuerdo con Makar y Rubin (2018), el RII se caracteriza por: (1) presentar una generalización, predicción o estimaciones de parámetros y conclusiones, que van más allá de describir los datos, (2) usar lenguaje probabilístico al desarrollar la generalización, que incluye una referencia informal de la incertidumbre de las conclusiones obtenidas, (3) utilizar los datos como evidencia de las generalizaciones, (4) centrarse en la cualidad de conjunto de una distribución, y (5) integrar un conocimiento del contexto.

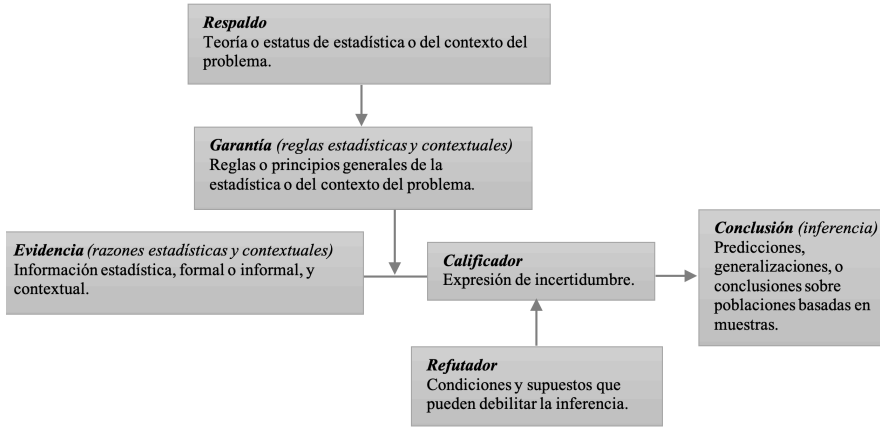
Por su parte, Zieffler et al. (2008) consideran que el desarrollo de un RII integra argumentos fundados en evidencia para hacer juicios, afirmaciones o predicciones de una población según sus muestras.

Por un lado, el RII demanda un conocimiento informal de temas estadísticos aprendidos en la escuela y de experiencias fuera de ella, así como de un razonamiento limitado a la experiencia de los estudiantes (Zieffler et al., 2008). Por otro, requiere de un conocimiento del contexto del problema y de los datos que serán analizados para realizar la inferencia, lo que Pfannkuch (2011) denomina "contexto de los datos".

Para analizar el razonamiento inferencial estadístico (RIE), Gómez-Blancarte y Tobías-Lara (2018; 2023), con base en el modelo de argumentación de Toulmin et al. (1984), proponen el modelo de argumentación del RIE (ver Figura 1). El tipo de argumentos que el RIE demanda se clasifican en "argumentos por generalización", los cuales implican afirmaciones basadas en la examinación de características de una "muestra suficientemente amplia y representativa" de una población (Toulmin et al., 1984, p. 219).

Figura 1

Modelo de argumentación del RIE



Nota: Tomado de Gómez Blancarte y Tobías Lara (2023).

El modelo de argumentación del RIE se explica en términos de los seis elementos de un argumento: *conclusión*, *evidencia*, *garantía*, *respaldo*, *calificador* y *refutador*. A continuación, basados en la propuesta de Gómez-Blancarte y Tobías-Lara (2023), se describen cada uno de estos elementos.

Conclusión. Enunciado que se presenta para su aceptación en general. Este elemento es semejante a la primera característica del RII: presentar o

emitir generalizaciones, predicciones, estimaciones de parámetros, conclusiones o juicios, que van más allá de describir los datos (Makar & Rubin, 2018).

Evidencia. Declaraciones que especifican determinados hechos que actúan como evidencia para apoyar la conclusión. En el RIE, la evidencia se basa en un análisis estadístico de los datos o en otra información del contexto. En este sentido corresponde con la tercera característica del RII: utilizar los datos como evidencia de las generalizaciones (Makar & Rubin, 2018).

Tanto el conocimiento estadístico como el del contexto juegan un papel primordial en la generación de razones sobre las cuales se apoya una inferencia. Por un lado, el conocimiento estadístico da lugar a las razones estadísticas, las cuales se desprenden del análisis estadístico de los datos; las razones contextuales provienen del contexto del problema o situación de la cual se derivan los datos que son analizados para realizar la inferencia. El conocimiento del contexto juega un papel fundamental porque sirve para interpretar y darle sentido a los datos.

Garantía. Son reglas o principios generales necesarios para demostrar que el paso de la evidencia a la afirmación es legítimo; pueden ser principios y estatutos legales, reglas empíricas, formulas o generalizaciones comunes basadas en intuiciones o experiencias. En el RIE, las garantías pueden ser normas o reglas procedentes de la estadística (reglas estadísticas) o del contexto del problema en el que se produce la inferencia (reglas contextuales). Dado que la garantía respalda el paso de la evidencia a la conclusión, su función es validar la conexión entre los datos y la conclusión.

Respaldo. Al igual que la garantía apoya a la evidencia, el respaldo apoya a la garantía. Se trata de un cuerpo de conocimiento aceptado en determinada ciencia, o de estatutos generales que actúan como argumentos autorizados donde descansan las garantías. La propuesta para el RIE es que ese cuerpo de conocimiento o estatuto general pueden provenir de la disciplina estadística o del contexto del problema. Además, el respaldo apoya a conclusiones que pueden ser conflictivas porque existen diferentes caminos de razonamiento para llegar a ellas, no por la verdad o falsedad de la evidencia.

Calificador y ***Refutador*****.** Estos elementos determinan la fuerza de un argumento. El calificador se refiere al "grado y el tipo de certeza que acompaña a las diferentes conclusiones" (Toulmin et al., 1984, p. 82), se expresa con adverbios modales como: "ciertamente", "presumiblemente", "muy probablemente, "muy posiblemente" (Toulmin et al., 1984, p. 86). En inferencia estadística, las conclusiones son por naturaleza inciertas, pues están basadas en un examen de la muestra de la población hacia la cual se infiere. En este sentido, este elemento se corresponde con la segunda característica

del RII: el uso del lenguaje probabilístico al desarrollar la generalización, que incluye una referencia informal de la incertidumbre de las conclusiones obtenidas (Makar & Rubin, 2018). Por su parte, el refutador actúa como las condiciones o circunstancias extraordinarias o excepcionales que podrían debilitar los argumentos de apoyo a la conclusión. En otras palabras, la conclusión está garantizada sólo en ausencia de esas condiciones.

METODOLOGÍA

El estudio se enmarcó en la metodología de *experimento de enseñanza*, la cual se compone de episodios de enseñanza cuyo propósito es explorar el aprendizaje de conceptos matemáticos y el razonamiento de los estudiantes (Steffe & Thomson, 2000). Cada episodio cuenta con un agente de enseñanza (el docente), uno o más alumnos participantes, un testigo (e.g., el asesor de tesis) y un método de grabación de las sesiones.

Secuencia didáctica

El experimento de enseñanza consistió en el desarrollo de una secuencia didáctica de seis actividades (Tabla 1) que giran en torno a responder las preguntas de investigación estadística: *¿Quiénes tienen más riesgo de morir por COVID-19 en México, hombres o mujeres? ¿A qué crees que se deba ese riesgo?* La primera pregunta cuestiona sobre una diferencia en la proporción de decesos por COVID-19 entre las poblaciones de hombres y mujeres. La segunda cuestiona las razones que pudieron provocar la posible diferencia de decesos. Estas preguntas buscan que el alumno se motive a encontrar una respuesta a través del análisis de datos, como lo sugieren Makar (2013) y Pfannkuch (2011).

Los datos que los estudiantes requerían para realizar el análisis estadístico y así poder responder a las preguntas de investigación, se descargaron de las bases de datos de la página de la Secretaría de Salud (disponibles en el sitio web: https://datos.covid-19.conacyt.mx/). Dicha base se preparó a fin de que los estudiantes exploraran únicamente los decesos confirmados (85,252) en el periodo del 28 de febrero al 15 de octubre de 2020. Se mantuvieron las variables de sexo, semana epidemiológica del fallecimiento, edad de la persona y factores de riesgo de cada individuo fallecido. Posteriormente, para estudiar si existía una diferencia entre las proporciones de decesos de ambos géneros, se aplicó un muestreo aleatorio para generar muestras de 25, 50, 100 y 250 casos, respectivamente. Además, para realizar un análisis del factor de riesgo responsable de la diferencia en las proporciones de decesos entre ambos géneros, se realizó otro muestreo aleatorio estratificado a la base de datos de decesos para conseguir una muestra de 500 casos, 250 hombres y 250 mujeres.

Tabla 1

Actividades de la secuencia didáctica

Actividades	Propósito
Actividad 1	
Tarea 1. Conocimiento acerca de COVID-19: a. ¿Qué es el COVID-19? b. ¿Cuáles son los diagnósticos para saber si se tiene COVID-19? c. ¿En qué consiste una prueba SARS-COV-2 (en adelante Prueba)? d. ¿Cuáles son los factores de riesgo para que una persona muera al contraer COVID-19?	Generar un conocimiento del contexto del problema.
Tarea 2. Información sobre la cantidad de decesos por COVID-19, en el periodo del 28 de febrero al 15 de octubre de 2020, y sobre la diferencia en la mortalidad por COVID-19 entre hombres y mujeres que reportan algunos estudios de investigación. Se plantean las preguntas de investigación: *¿Quiénes tienen más riesgo de morir por COVID-19 en México, los hombres o las mujeres? ¿A qué crees que se deba ese riesgo?*	Presentar las preguntas de investigación para motivar la realización de la investigación estadística y conocer las respuestas de los alumnos, previo a la exploración de datos.
Tarea 3. Inicio de la heurística "crecer la muestra". Se entrega a los estudiantes la base de 25 datos (n = 25) para que en Fathom construyan una gráfica de puntos y obtengan el porcentaje de decesos de hombres y mujeres.	Analizar las gráficas y porcentajes a fin de responder quienes murieron más, hombres o mujeres, y justificar la respuesta.
Tarea 4. Dibujar, a lápiz y papel, la gráfica de puntos que, según el estudiante, debería formarse con n = 50.	Explicar la gráfica de puntos dibujada y compararla, durante la Actividad 2, con la construida en Fathom.
Actividad 2	
Tarea 1. Se entrega a los estudiantes la base de 50 datos (n = 50) para que en Fathom construyan una gráfica de puntos y obtengan el porcentaje de decesos de hombres y mujeres.	Analizar las gráficas y porcentajes a fin de responder quienes murieron más, hombres o mujeres, y justificar la respuesta.
Tarea 2. Comparación de la gráfica de predicción (para n = 50), realizada en la Actividad 1, con la gráfica construida en la tarea anterior.	Promover la observación de similitudes y diferencias entre la gráfica de predicción (dibujada por el estudiante) y la gráfica con los datos reales (construida en Fathom).
Tarea 3. Dibujar, a lápiz y papel, la gráfica de puntos que, según el estudiante, debería formarse con n = 100.	Explicar la gráfica de puntos dibujada y compararla, durante la Actividad 3, con la construida en Fathom.

Actividad 3	
Tarea 1. Se entrega a los estudiantes la base de 100 datos (n = 100) para que en Fathom construyan una gráfica de puntos y obtengan el porcentaje de decesos de hombres y mujeres.	Analizar las gráficas y porcentajes a fin de responder quienes murieron más, hombres o mujeres, y justificar la respuesta.
Tarea 2. Comparación de la gráfica de predicción (para n = 100), realizada en la Actividad 2, con la gráfica construida en la tarea anterior.	Promover la observación de similitudes y diferencias entre la gráfica de predicción (dibujada por el estudiante) y la gráfica con los datos reales (construida en Fathom).
Tarea 3. Dibujar, a lápiz y papel, la gráfica de puntos que, según el estudiante, debería formarse con n = 250.	Explicar la gráfica de puntos dibujada y compararla, durante la Actividad 4, con la construida en Fathom.
Actividad 4	
Tarea 1. Se entrega a los estudiantes la base de 250 datos (n = 250) para que en Fathom construyan una gráfica de puntos y obtengan el porcentaje de decesos de hombres y mujeres.	Analizar las gráficas y porcentajes a fin de responder quienes murieron más, hombres o mujeres, y justificar la respuesta.
Tarea 2. Comparación de la gráfica de predicción (para n = 250), realizada en la Actividad 3, con la gráfica construida en la tarea anterior.	Promover la observación de similitudes y diferencias entre la gráfica de predicción (dibujada por el estudiante) y la gráfica con los datos reales (construida en Fathom).
Tarea 3. Fin de la heurística "crecer la muestra". Dibujar la gráfica que debería formarse con la totalidad de la población (N = 85,252), según lo observado en las gráficas previas.	Estimar y explicar la distribución de decesos en la población.
Actividad 5	
Tarea 1. Se entrega a los estudiantes una base de 500 datos (n = 500), 250 hombres y 250 mujeres fallecidas, de factores de riesgo (enfermedad cardiovascular, pulmonar, tabaquismo, obesidad e inmunosupresión) y edad. a. Construcción de gráficas por edad y sexo. b. Construcción de gráficas por cada factor de riesgo y sexo. c. Obtención del porcentaje de decesos de hombres y mujeres para cada factor de riesgo.	Analizar las gráficas y porcentajes para considerar si alguna de las enfermedades o la edad pueden ser un factor de riesgo que determine una posible diferencia entre las poblaciones de donde se desprenden dichas muestras.

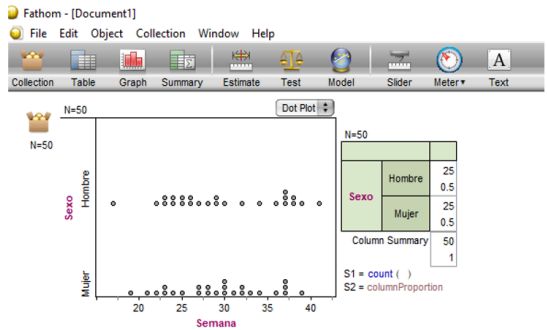
Actividad 6

Tarea 1. Cierre de la secuencia didáctica. De acuerdo con las actividades previas, los estudiantes respondieron las siguientes preguntas: a. A qué conclusión llegas sobre quiénes tienen más riesgo de morir por COVID-19 en México, ¿los hombres o las mujeres? b. ¿Cómo llegaste a esa conclusión? ¿Cuáles son tus evidencias? c. ¿Es confiable tu conclusión? ¿Por qué? d. ¿Cuáles son las condiciones para considerar que tu conclusión se cumple?	Explicitar los elementos del modelo de argumentación, de acuerdo con las actividades previamente realizadas.

Las primeras cuatro actividades de la secuencia didáctica estuvieron basadas en la heurística de "crecer la muestra" (Ben-Zvi et al., 2018). Dicha heurística consiste en presentar una gráfica de una muestra de datos tomada de la población en estudio para, posteriormente, realizar una inferencia respecto a cómo sería la gráfica de una muestra de mayor tamaño. Después, los alumnos realizan una comparación entre su inferencia y una gráfica de las muestras con mayor tamaño. Esta idea permite un ciclo de inferencias en las que se presenta una gráfica, se realiza una predicción y se hace una comparación que se puede repetir múltiples ocasiones con muestras cada vez mayores. Durante la aplicación de la Actividad 1, los alumnos construyeron en el programa Fathom (https://fathom.concord.org) una primera gráfica de puntos (Figura 2) para una muestra de tamaño 25. Los puntos de la gráfica representan los casos de decesos por género y por semana epidemiológica. Además, se determinó la proporción de decesos de cada género según la muestra graficada. Posteriormente, se solicitó al alumno dibujar en papel la forma que tomaría la gráfica al agregar mayor cantidad de datos (crecer la muestra). El ciclo de inferencias continuó al entregar una muestra con mayor cantidad de datos para ser analizados a través del programa Fathom, tal y como se describió en la Tabla 1.

Figura 2

Ejemplo de la gráfica de puntos y la proporción de decesos construida en Fathom



La actividad 5 gira en torno al estudio de una muestra de n = 500, con 250 hombres y 250 mujeres. Los alumnos, con apoyo del mismo programa, construyeron las gráficas de comparación por sexo para explorar el factor de riesgo, enfermedades (Figura 3) y edad (Figura 4) que podía generar la diferencia entre las proporciones de decesos de ambos géneros.

Figura 3

Ejemplo de algunas de las gráficas de enfermedades construidas en Fathom

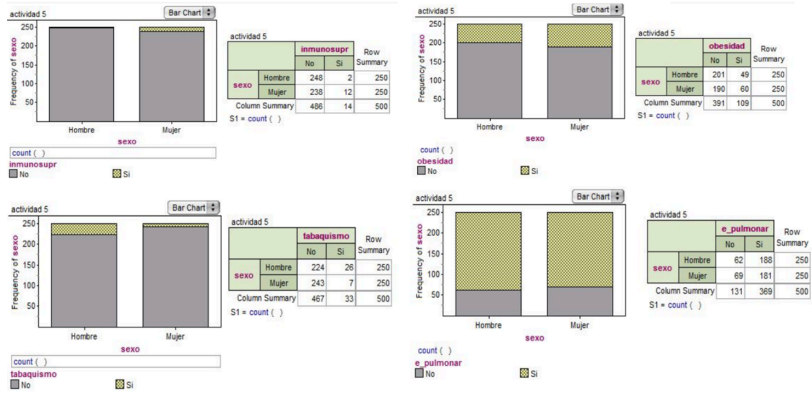
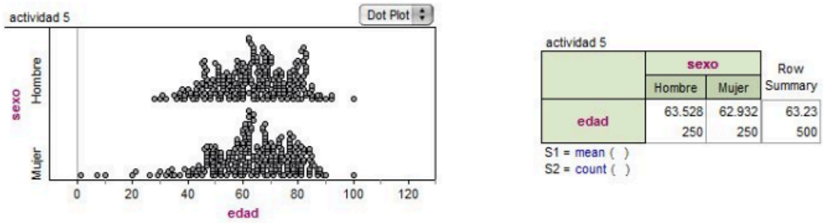


Figura 4

Ejemplo de las gráficas de edad construidas en Fathom



Finalmente, la actividad 6 consistió en un debate guiado para fomentar la aparición de los elementos del modelo de argumentación del RIE, al proveer un espacio de discusión y puesta en común de lo observado a lo largo de las actividades realizadas.

Implementación de la secuencia didáctica

El experimento de enseñanza se desarrolló en un contexto educativo en el que participaron estudiantes de tercer grado de educación secundaria de la Escuela Secundaria Técnica #19, ubicada en el municipio de Ignacio Zaragoza, una zona semi rural del estado de Chihuahua. El docente (Autor 1) invitó a sus alumnos a que, de manera voluntaria, participaran en el experimento de enseñanza. Además, informó a los padres para contar con su aprobación y apoyo durante la aplicación de la secuencia didáctica. Debido

al contexto de la zona donde se ubica el plantel y a la situación de la pandemia que obligó a tener clases en línea, solo participaron cuatro alumnos que contaban con equipo de cómputo y acceso a internet. Se usaron seudónimos para identificar a los participantes: Arturo, Estrella, Sol y Lucero. Según el promedio de calificaciones, uno de ellos se consideró de aprovechamiento alto, dos de aprovechamiento medio y uno de aprovechamiento bajo.

Se llevaron a cabo cinco sesiones de enseñanza a través de la plataforma *Meet*, las cuales duraron entre 45 y 70 minutos cada una. Antes de iniciar el desarrollo de las actividades, el docente solicitó a los alumnos realizar la instalación (versión de prueba) del programa Fathom en cada uno de sus equipos de cómputo. Al inicio de las primeras cuatro actividades, el docente enseñó a los alumnos a usar las herramientas básicas necesarias de Fathom para que pudieran llevar a cabo las tareas de las actividades de manera autónoma. Los alumnos lograron comprender cómo utilizar el programa, por lo que las dudas respecto a su uso fueron disminuyendo conforme se progresó en las tareas de la secuencia didáctica. Así, durante las primeras cuatro actividades, el docente realizó 1) una breve discusión de aprendizajes previos sobre el tema o lo que se observó en las sesiones anteriores, 2) una explicación del uso del programa Fathom. Posteriormente, los alumnos 3) realizaron las tareas de las actividades y, al finalizar, 4) expusieron y justificaron sus respuestas a cada una de las tareas de la actividad.

Para el desarrollo de la actividad 5, los alumnos ya contaban con experiencia manejando el programa. Sin embargo, debido a la necesidad de crear gráficas diferentes a las realizadas en las actividades previas, el docente explicó el proceso para generar una gráfica de comparación y encontrar los valores de las proporciones de decesos por cada uno de los factores de riesgo. Posteriormente, los alumnos construyeron de manera independiente sus propias gráficas. En caso de duda, el docente guiaba verbalmente las acciones a realizar en el programa para que los estudiantes prosiguieran con el manejo autónomo del programa. Para finalizar la aplicación de la secuencia, durante la actividad 6, únicamente se discutieron las respuestas a las preguntas correspondientes.

Recolección y análisis de datos

Los datos recolectados provienen de la transcripción de los videos de las sesiones de clase, las respuestas escritas a cada actividad y de mensajes de *WhatsApp* que los alumnos escribían al profesor. Con estos datos se realizó un análisis de contenido para distinguir las respuestas de los alumnos a las preguntas de investigación en cada una de las actividades que conformaron la secuencia didáctica. Dicho análisis se refiere a desarrollar una categorización de los datos para proveer de estructura al texto y llevar a cabo un

proceso sistemático de interpretación (Kuckartz, 2019). En este caso, se realizó una categorización de las respuestas de acuerdo con los elementos del modelo de argumentación del RIE, se utilizaron los códigos C, E, G, B, Q y R. La C corresponde a la conclusión a la pregunta de investigación (*¿Quiénes tienen más riesgo de morir por COVID-19 en México, hombres o mujeres?*); la E hace referencia a las razones estadísticas, razones o explicaciones contextuales que respaldaban a la conclusión y que estaban motivadas por la justificación que dieron a la respuesta de la pregunta de investigación. Se utilizó Q para identificar las expresiones de incertidumbre o certeza expuestas por los alumnos. G, B y R representaron a la garantía, el respaldo y el refutador, respectivamente. Dado que el análisis se hizo por cada una de las actividades, se utilizaron subíndices numéricos en cada código para enumerar los elementos presentados a lo largo de la secuencia didáctica. Por ejemplo, C_1 representa la conclusión presentada en la Actividad 1, E_4 representa a la evidencia número 4, pero no necesariamente hace referencia a la Actividad 4, pues el alumno podía proporcionar más de una evidencia para una sola conclusión. Finalmente, se construyó el modelo de argumentación de cada alumno por actividad.

RESULTADOS

Por razones de espacio, en esta sección se expone el RII de Arturo. Se eligió a este participante porque, a diferencia de sus compañeras, proporcionó un mayor número de evidencias para apoyar sus conclusiones. Esto nos permitió ejemplificar con mayor detalle la integración de razones estadísticas y razones contextuales, así como exponer el tipo de explicaciones contextuales usadas para la interpretación de los datos. De hecho, su compañera Lucero también proporcionó evidencias que detallan dicha integración, pero, por razones de conectividad, no desarrolló la Actividad 1. Por su parte, el RII de Sol y Estrella se caracterizó por proporcionar evidencias cuyos argumentos enfatizaban más un tipo de razones: razones estadísticas (Sol) o razones y explicaciones contextuales (Estrella). Es decir, aunque integraban razones estadísticas y contextuales, el soporte principal de su conclusión estaba en uno de esos dos tipos.

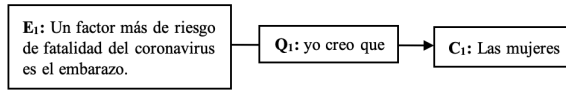
Modelo de argumentación de Arturo previo a la exploración de datos

Durante la Tarea 2 de la Actividad 1, previo a la exploración de los datos, los alumnos respondieron, de acuerdo con su conocimiento del contexto, a las preguntas: *¿Quiénes tienen más riesgo de morir por COVID-19 en México, los hombres o las mujeres? ¿A qué crees que se deba ese riesgo?* Las respuestas de Arturo mostraron tres elementos del modelo de argumentación (Figura 5). Su conclusión (C1) fue que las mujeres tenían mayor riesgo, utilizó la expresión "yo creo que" (Q1) para denotar cierta incertidumbre en su con-

clusión. Su principal evidencia (E1) fue una razón contextual, la cual se apoyó en el hecho de que las mujeres tienen un factor de riesgo que no tienen los hombres: el embarazo. En este sentido, para Arturo la población de mujeres tenía más riesgo de fallecer por COVID-19.

Figura 5

Modelo de argumentación de Arturo, previo a la exploración de dato

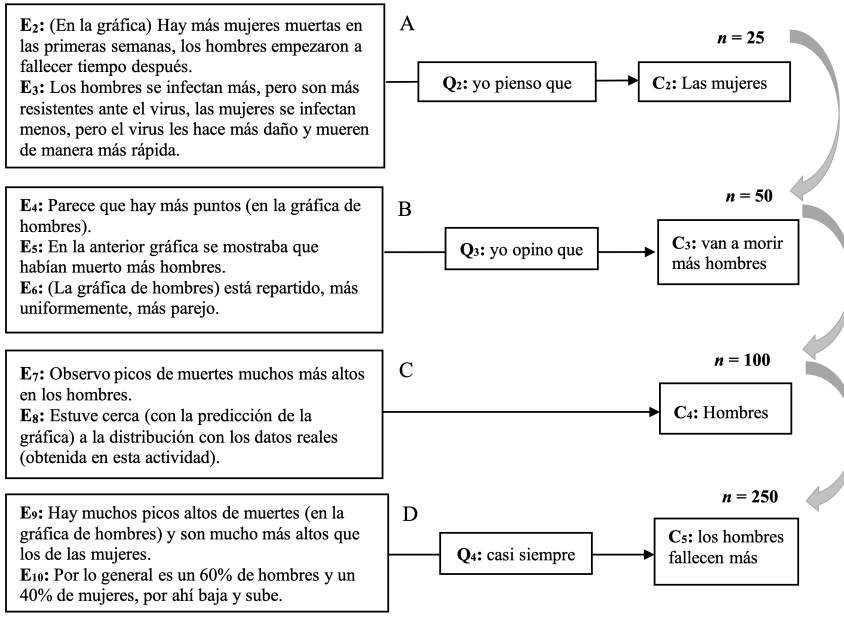


Modelos de argumentación de Arturo durante la "heurística de crecer la muestra"

A partir de la Tarea 3 de la Actividad 1 y hasta la Tarea 3 de la Actividad 4, los alumnos hicieron uso del programa Fathom para analizar los datos. En los argumentos producidos después de la exploración de los datos de la muestra n = 25 (Figura 6A), Arturo mantuvo su argumento de que las mujeres tenían mayor riesgo de fallecer por COVID-19 que los hombres (C_2). Conforme fue explorando más datos, su argumento cambió y concluyó que eran los hombres quienes corrían el mayor riesgo (Figuras 6B, 6C, 6D).

Figura 6

Modelos de argumentación de Arturo, durante la "heurística de crecer la muestra"



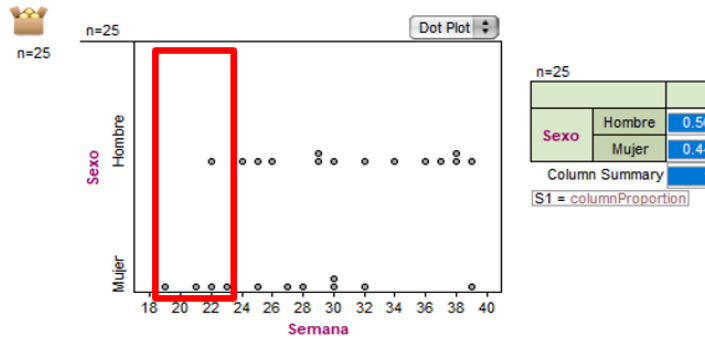
En los cuatro argumentos, derivados de la heurística de crecer la muestra, Arturo utilizó en tres de ellos calificadores modales como "yo pienso que"

(Q_2), "yo opino que" (Q_3) y "casi siempre" (Q_4). El Q_2 manifiesta un nivel de incertidumbre que puede ser atribuido a que él mantuvo su pensamiento de que "el virus les hace más daño" (E_3) a las mujeres, posiblemente por el factor embarazo que ya había expresado en su E_1. En cambio, como se mostrará más adelante, el calificador Q_4 denota un poco más de confianza en su conclusión debido a las observaciones que hizo durante la exploración de los datos.

En la gráfica de la muestra de 25 casos (Figura 7), Arturo comparó los decesos de las primeras semanas y observó que "hay más mujeres muertas en las primeras semanas, los hombres empezaron a fallecer tiempo después" (E_2). En la muestra analizada, la gráfica de mujeres tuvo un par de fallecimientos antes de la aparición del primer fallecimiento de los hombres (Figura 7, barra roja); estos datos fueron información suficiente para que Arturo pensará que las mujeres fallecían más.

Figura 7

Gráfica construida por Arturo para n =25



La E_2 es una razón estadística porque está basada en la exploración de los datos presentados en la gráfica. No obstante, Arturo solo utilizó como evidencia una parte del conjunto de los datos, posiblemente para sostener su pensamiento previo sobre que son las mujeres las de mayor riesgo. Además de la E_2, Arturo explicó que "los hombres se infectan más pero son más resistentes ante el virus" (E_3). Arturo parecía reconocer que la muestra mostraba más fallecimientos de hombres, tal vez observó las proporciones de fallecimientos (0.56 para hombres y 0.44 para mujeres) que también realizó en el programa Fathom (Figura 7, cuadro derecho), pero su idea de que las mujeres tenían mayor riesgo la explicó diciendo que a ellas "el virus les hace más daño y mueren de manera más rápida" (E_3). En este sentido, la explicación contextual (E_3) que utilizó Arturo muestra la manera en que él interpretó los datos a fin de darle sentido a su conclusión (C_2).

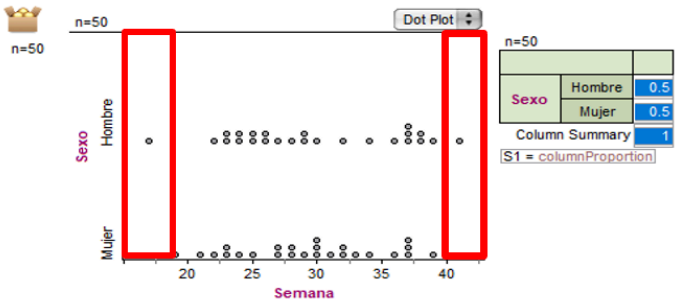
Después del análisis de la muestra de 50 casos, Arturo cambió su argumento de que las mujeres tenían el mayor riesgo y concluyó que eran los

hombres (C_3). La expresión "van a morir más hombres" (Figura 6B) es una predicción que denota una generalización que va más allá de la muestra analizada. Además, su conclusión se apoyó únicamente en razones estadísticas (E_4, E_5, E_6). La primera razón (E_4) se basó en la observación de la gráfica construida (Figura 8), pues vio "más puntos" en la gráfica de hombres. Dado que son la misma cantidad de puntos, posiblemente él vio los puntos extremos de la gráfica (Figura 8, rectángulos rojos) y se enfocó en el rango de los datos de la de hombres, lo que lo hizo percibir "más puntos".

La E_5 ("En la anterior gráfica se mostraba que habían muerto más hombres") es una razón estadística basada en la actividad previa. El alumno tomó en cuenta que en la gráfica realizada con la muestra de $n = 25$ había mayor proporción de hombres fallecidos. Es decir, esta razón se basa en un resultado previo que Arturo utiliza como un hecho para apoyar una conclusión posterior. En otras palabras, la misma gráfica que el alumno usó para apoyar su C_2 ("Las mujeres"), la usó ahora para apoyar su C_3 ("van a morir más hombres"), la diferencia radica en el tipo de interpretación de los datos que hizo Arturo sobre la misma gráfica. Lo anterior ejemplifica cómo las evidencias estadísticas (una mejor interpretación de las gráficas) empezaron a influir en su razonamiento.

Figura 8

Gráfica construida por Arturo para n = 50



A pesar de que la proporción de fallecidos en la muestra de 50 datos fue igual en ambos grupos, Arturo concluyó que eran los hombres según el comportamiento que los datos estaban mostrando desde la gráfica de 25 datos. Además, en la E_6, Arturo comenzó a tomar en cuenta la forma de la distribución de los datos, utilizó las palabras "repartido", "uniformemente" y "parejo" para referirse de manera informal a la distribución de los puntos observados. De hecho, su C_4 (Figura 6C) se apoyó en una razón estadística (E_7) que da cuenta de otra observación en la distribución de los datos: se enfocó en la cantidad de puntos apilados en la gráfica de la muestra de hombres, que llamó "picos de muertes"; realizó una comparación entre ellos y observó "picos de muertes mucho más altos en los hombres". Es decir,

observó la frecuencia de fallecimientos por semana y las comparó en ambas gráficas (en la Figura 9 se utilizaron un par de rectángulos rojos, con las mismas dimensiones, para mostrar la diferencia entre los puntos apilados).

En la E_8 ("Estuve cerca a la distribución con los datos reales") Arturo utilizó el término "distribución" para referirse a la forma en que dibujó la gráfica de predicción solicitada en la Tarea 3 de la Actividad 2. El alumno consideró que su predicción de la gráfica de datos (Figura 10) coincidió parcialmente con el comportamiento de la gráfica realizada en Fathom (Figura 9). En otras palabras, Arturo ya predecía que serían los hombres quienes presentarían un mayor número de muertes en la siguiente muestra. En este sentido, el análisis de las gráficas estaba modificando y sosteniendo el razonamiento de Arturo.

Figura 9

Gráfica construida por Arturo para n = 100

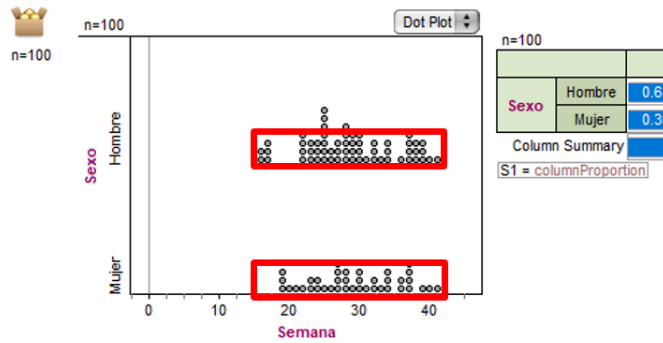
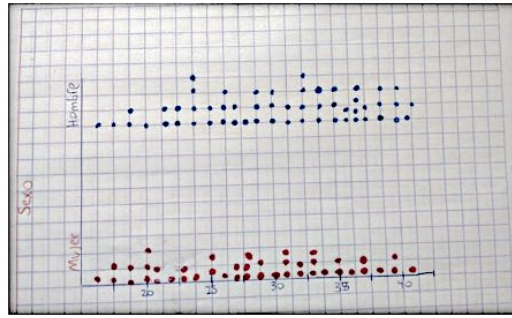


Figura 10

Gráfica de predicción de una muestra con n = 100 realizada por Arturo



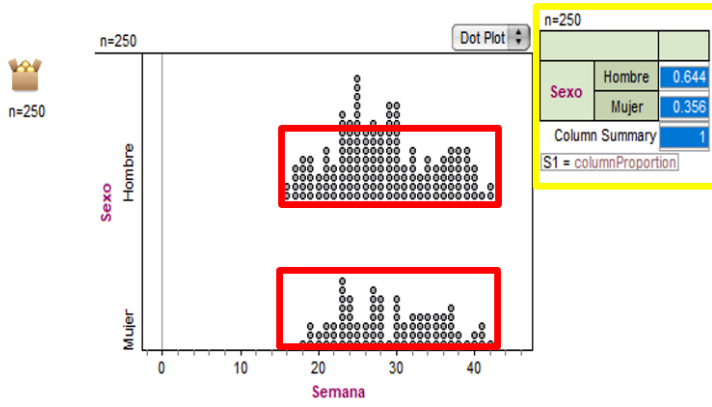
Cabe aclarar que en su C_4 Arturo no manifestó una frase que denotara incertidumbre sobre su conclusión (Figura 6C). En cambio, en su argumento correspondiente al análisis de la muestra de 250 datos (Figura 6D), Arturo utilizó una frase de incertidumbre (Q_4: "casi siempre"), diferente a las previas (Q_2: "yo pienso que", Q_3: "yo opino que"), la cual denota un poco más de

confianza en su conclusión, debido a los resultados estadísticos que había observado en los análisis previos. Con la muestra de 250 casos, Arturo se apoyó en dos razones estadísticas (E_9 y E_{10}) para concluir que los "hombres fallecen más" (C_5). La E_9 es una interpretación de las características observadas en la gráfica (Figura 11), en la cual Arturo comparó la altura de los puntos, E_9: "Hay muchos picos altos de muertes y son mucho más altos que los de las mujeres", similar a lo observado en la E_7.

La E_{10} ("Por lo general es un 60% de hombres y un 40% de mujeres, por ahí baja y sube") está basada en los valores de las proporciones observadas (Figura 11, rectángulo amarillo). Sin embargo, es importante notar que Arturo no dio los valores exactos que arrojó el programa; en su lugar, señaló que esos valores estaban alrededor ("por ahí baja y sube") de 60% de muertes de hombres y 40% para el caso de las mujeres. El hecho de observar que la proporción esperada de muertes para cada población se encuentra alrededor de cierto valor es una característica importante del RII, pues muestra una inferencia de un valor poblacional que fue determinado por la exploración de los datos de las muestras graficadas. Además, toma en cuenta la variabilidad de los datos, lo cual es un aspecto fundamental del razonamiento estadístico.

Figura 11

Gráfica construida por Arturo para n = 250



Modelo de argumentación de Arturo sobre los factores de riesgo

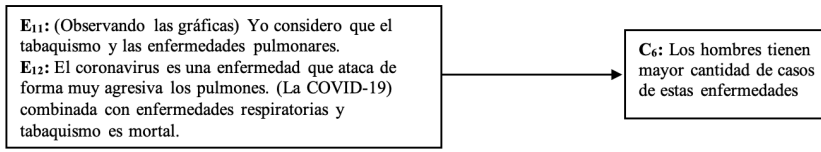
Posterior a realizar el análisis de las gráficas sobre los factores de riesgo, Arturo concluyó que "los hombres tienen mayor cantidad de casos de estas enfermedades" (C_6, Figura 12), lo que significa que su razonamiento fue más descriptivo, centrado en la muestra explorada, que inferencial.

Las evidencias usadas por Arturo estuvieron formadas por una razón estadística E_{11} y una explicación contextual E_{12}. Sobre E_{11}, surgió directamente de la observación de las gráficas exploradas. El tabaquismo y las

enfermedades pulmonares fueron los únicos dos factores en los que se observó una mayor proporción en los hombres (Figura 13, rectángulo rojo y rectángulo azul) en comparación con las mujeres. En el resto de los factores, fueron las mujeres quienes presentaron mayor proporción.

Figura 12

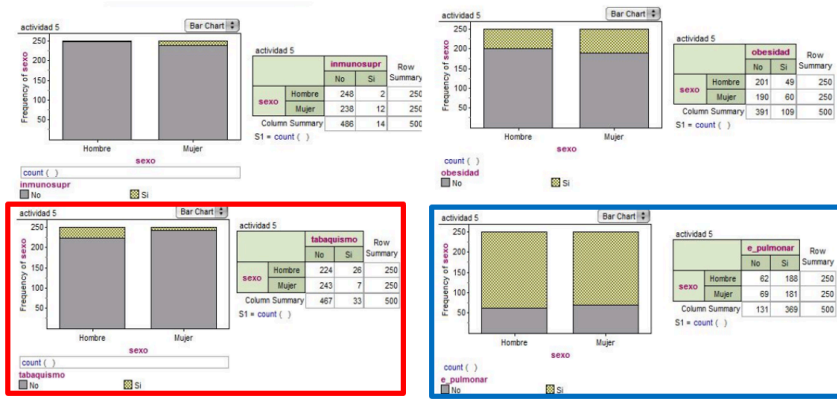
Modelo de argumentación de Arturo sobre los factores de riesgo



Arturo presentó una hipótesis para dar sentido al hecho de que los hombres tenían una mayor proporción en factores de riesgo como tabaquismo y enfermedades pulmonares. Esta hipótesis es una explicación contextual (E_{12}) de la E_{11}, ya que explica que la COVID-19 "ataca de forma muy agresiva a los pulmones" y que "combinada con enfermedades respiratorias y tabaquismo es mortal" (E_{12}). Estas razones contextuales que ofreció Arturo pueden derivarse del conocimiento contextual sobre el virus, expuesto a través de los medios de comunicación.

Figura 13

Gráficas de comparación por factores de riesgo para n = 500



Modelo de argumentación final de Arturo

Durante la Actividad 6, Arturo concluyó que "los hombres son más afectados" (C_7, Figura 14), es decir, los hombres son la población con mayor prevalencia a fallecer por COVID-19. Arturo mostró un nivel de incertidumbre sobre su C_7 al considerar que su conclusión no es confiable, Q_5: "Mi respuesta no es tan confiable, apenas me estoy metiendo a este terreno". En este sentido, el lenguaje de incertidumbre parece denotar una falta de confianza en los

conocimientos de Arturo respecto a los procedimientos estadísticos implicados en la tarea.

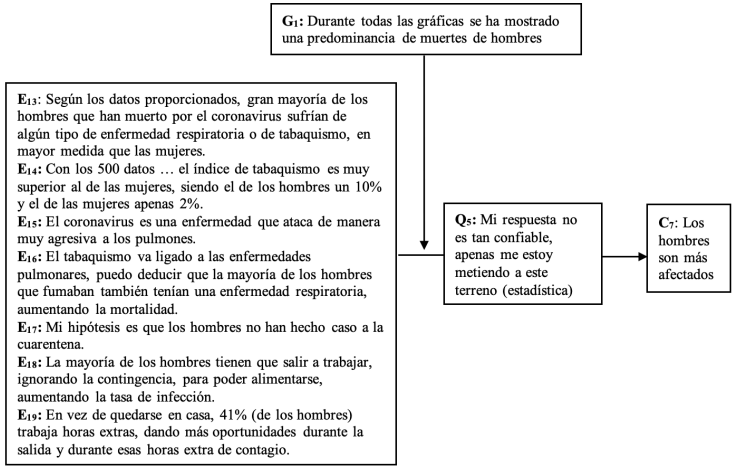
Arturo utilizó diferentes razones (E_{13} – E_{19}), tanto estadísticas como contextuales, para sostener su conclusión. Además, logró expresar, por primera vez, una garantía (G_1) en la cual apoyó el paso de las evidencias a la conclusión.

La E_{13} es una razón estadística porque hace referencia al análisis de la muestra de datos ("Según los datos proporcionados …") sobre los factores de riesgo causantes de la muerte de hombres ("…gran mayoría de los hombres que han muerto por el coronavirus sufrían de algún tipo de enfermedad respiratoria o de tabaquismo, en mayor medida que las mujeres").

La evidencia E_{14} es otra razón estadística que complementa a la E_{13}, pues hace referencia al tamaño de la muestra analizada (n = 500) y a los valores de las proporciones de las personas fallecidas (10% vs. 2%) que padecían tabaquismo, advirtiendo un mayor valor para el caso de los hombres.

Figura 14

Modelo de argumentación final de Arturo



En la E_{15} se puede apreciar una razón contextual, la cual usó Arturo como una explicación contextual para asociar el tabaquismo con la enfermedad del COVID-19: "El coronavirus es una enfermedad que ataca de manera muy agresiva a los pulmones" (E_{15}). En otras palabras, Arturo usa esta razón para evidenciar que los hombres mueren más por el tabaquismo, factor que afecta en los pulmones al igual que lo hace el virus del COVID-19. De hecho, en su E_{16} conecta el tabaquismo con las enfermedades respiratorias, pues supone que los hombres con tabaquismo tenían una enfermedad respiratoria que los hace tener un mayor riesgo de morir. Luego supone que los hombres no hicieron caso a la cuarentena obligatoria (E_{17}) para reducir

el número de contagios de COVID-19. Al parecer, Arturo contaba con información sobre el porcentaje de hombres que salían a trabajar horas extras (E_{19}). Esta información se convirtió en una razón más para apoyar su hipótesis explicativa (E_{18}) sobre la mayor exposición de contagio por parte de los hombres.

En el argumento de Arturo se puede observar el uso de la garantía G_1: "Durante todas las gráficas se ha mostrado una predominancia de muertes de hombres". Es importante recalcar que el alumno utilizó los resultados estadísticos de las gráficas realizadas como un hecho general de sus evidencias particulares (E_{13} y E_{14}) para dar lugar a su conclusión.

CONCLUSIONES

El análisis del RII de Arturo ofrece una visión de la posibilidad de fomentar en el aula ese razonamiento cuando se desarrolla una situación didáctica de inferencia estadística con el uso de una herramienta computacional para el análisis de los datos. El desarrollo de esta secuencia motivó, en mayor o menor medida, el surgimiento de elementos del RII, según el modelo de argumentación del RIE, como se muestra en las principales conclusiones del estudio:

- La heurística de "crecer la muestra" coadyuvó a que las conclusiones de algunos de los alumnos manifestaran una generalización más allá de la muestra. Por ejemplo, la C_3 de Arturo ("van a morir más hombres") indica un pensamiento más inferencial que determinístico. En este sentido, se coincide con Ben-Zvi et al. (2012) sobre que dicha heurística contribuye a que los alumnos formulen predicciones, mejorando en aspectos de inferencia informal.
- El programa Fathom apoyó la visualización de los datos, lo que permitió que los alumnos se centraran más en el análisis de estos que en la construcción de sus representaciones. En este sentido, la tecnología fue útil para promover el surgimiento de razones estadísticas y fomentar un razonamiento estadístico; en particular, un razonamiento estadístico sobre distribuciones, ya que la interpretación de las gráficas dio lugar para que los alumnos visualizaran y mencionaran, de manera informal (e.g., "uniformemente", "picos"), características de la distribución de los datos.
- La información del contexto que utilizaron los alumnos comenzó a tomar en cuenta datos de la situación estudiada, lo que permitió una integración más equilibrada de razones estadísticas y razones contextuales. El fenómeno de sustentar las conclusiones en creencias (e.g., E_1: "Un factor más de riesgo de fatalidad del coronavirus es el embarazo") se redujo conforme se avanzó en la secuencia didáctica, pues se comenzaron a presentar mayor cantidad de razones estadísticas.

- Aunque el uso de calificadores modales no fue constante, se identificaron dos tipos: aquellos que expresaban incertidumbre y aquellos que expresaban más certidumbre sobre las conclusiones. Según Henriques y Oliveira (2016), los alumnos tienden a expresar la incertidumbre de manera cualitativa, lo cual es algo esperado cuando se desarrolla un RII, ya que se aceptan frases apropiadas al nivel del estudiante. Las frases que utilizaron los estudiantes (e.g., "yo pienso", "yo opino", "casi siempre", "mi respuesta no es tan confiable") son reconocidas como indicadores de incertidumbre en el sentido de que le otorgan a la conclusión un estatus de duda, es decir, que no son conclusiones incuestionables, sino, como señalan Ben-Zvi et al. (2012), son "tentativas, temporales y relativas a la información disponible" (p. 924). No obstante, es importante enfatizar el uso del lenguaje de incertidumbre como un indicador inherente de los procedimientos estadísticos, y no de la confianza que se tenga sobre ellos.
- Los alumnos identificaron un patrón en las actividades realizadas, lo que ayudó a formular sus garantías. Al trabajar durante varias sesiones con la misma pregunta de investigación, pero con diferentes tamaños de muestra, los alumnos iban observando patrones en los datos. Luego, observaban que la evidencia actual era un caso particular de esos patrones, garantizando con ello su conclusión. En otras palabras, los patrones se convirtieron en principios que autorizaban la conexión entre la evidencia y la conclusión.

En general, la situación didáctica permitió el desarrollo de un RII en los estudiantes porque se apoyó en: 1) tareas que motivaron la presentación de inferencias hacia una población a partir de los datos de las muestras de esa población (Zieffler et al., 2008), 2) preguntas de investigación que despertaron el interés en los estudiantes y motivaron la exploración de los datos haciendo uso de la tecnología (Makar, 2013), y 3) uso de un contexto familiar para los estudiantes. Como señala Pfannkuch (2011), el contexto expande el RII de los estudiantes porque les permite generar una conexión entre la información abstracta de los datos (e.g., las proporciones de fallecidos, las distribuciones de los datos en las gráficas) y la generación de una historia plausible sobre ellos. Además, el modelo de argumentación del RIE fue un referente clave para guiar los procesos de argumentación de los propios estudiantes. En este sentido, el estudio aporta ideas importantes y necesarias para fomentar un RII en los estudiantes de educación secundaria, y da cuenta de la posibilidad de llevar al aula estrategias didácticas que apoyen dicho razonamiento a fin de potenciar el aprendizaje tanto de conceptos estadísticos como de un uso de la estadística para juzgar situaciones reales basados en argumentos válidos.

REFERENCIAS

Ben-Zvi, D., Aridor, K., Makar, K., & Bakker, A. (2012). Students' emergent articulations of uncertainty while making informal statistical inferences. *ZDM - International Journal on Mathematics Education*, *44*(7), 913–925. https://doi.org/10.1007/s11858-012-0420-3

Ben-Zvi, D., Gravemeijer, K., & Ainley, J. (2018). Design of statistics learning environments. En D. Ben-Zvi, K. Makar & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 473–502). Springer. https://doi.org/10.1007/978-3-319-66195-7_16

Cruz López, C., & Ramírez Ojeda, M. M. (2021). Un diagnóstico de los cursos de servicio en estadística en México. *Statistics Education Research Journal, 20*(2). https://bit.ly/3EyQcH0

Gómez Blancarte, A. L., Chávez Aguilar, R. D., & Miranda Viramontes, I. (2022). Enfoques de la enseñanza de la estadística en los programas de estudio de educación superior. *IE Revista de investigación educativa de la REDIECH*, *13*. https://doi.org/10.33010/ie_rie_rediech.v13i0.1394

Gómez-Blancarte, A., & Tobías-Lara, M. G. (2018). Using the Toulmin model of argumentation to validate students' inferential reasoning. En M. A. Sorto, A. White, & L. Guyot (Eds.), *Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS-10)*. International Statistical Institute. https://bit.ly/3PyeOpK

Gómez-Blancarte, A., & Tobías-Lara, M. G. (2023). The integration of undergraduate students' informal and formal inferential reasoning. *Educational Studies in Mathematics, 113*(2), 251–269. https://doi.org/10.1007/s10649-022-10205-w

Henriques, A., & Oliveira, H. (2016). Students' expressions of uncertainty in making informal inference when engaged in a statistical investigation using Tinkerplots. *Statistics Education Research Journal, 15*(2), 62–80. https://bit.ly/3PasOnZ

Hernández Trevethan, H. (2018). High school, Tinkerplots, Fathom… and independence and ANOVA. How come? En M. A. Sorto, A. White, & L. Guyot (Eds.), *Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS-10)*. International Statistical Institute. https://bit.ly/3r9NPYc

Huerta Orozco, J. C. (2021). *Secuencia didáctica para fomentar el desarrollo del razonamiento inferencial informal en alumnos de secundaria* [Tesis de maestría, Centro de Investigación en Ciencia Aplicada y Tecnología Avanzada, Unidad Legaria (CICATA-Legaria), del Instituto Politécnico Nacional]. Repositorio de CICATA-Legaria. https://bit.ly/3Rg1dV7

Kuckartz, U. (2019). Qualitative text analysis: A systematic approach. En G. Kaiser & N. Presmeg (Eds.), *Compendium for early career researchers in mathematics education*, (pp. 181–197). Springer. https://doi.org/10.1007/978-3-030-15636-7

Kula, F., & Koçer, R. G. (2020). Why is it difficult to understand statistical inference? Reflections on the opposing directions of construction and application of inference framework. *Teaching Mathematics and its Applications: An International Journal of the IMA, 39*(4), 248–265. https://doi.org/10.1093/teamat/hrz014

Makar, K. (2013). Predict! Teaching statistics using informal statistical inference. *Australian Mathematics Teacher, 69*(4), 34–41. https://bit.ly/44TSV8Q

Makar, K., & Rubin, A. (2018). Learning about statistical inference. En D. Ben-Zvi, K. Makar & J. Garfield (Eds.), *International handbook of research in statistics education* (pp. 261–294). Springer. https://doi.org/10.1007/978-3-319-66195-7_8

Pfannkuch, M. (2011). The role of context in developing informal statistical inferential reasoning: A classroom study. *Mathematical Thinking and Learning*, *13*(1–2), 27–46. https://doi.org/10.1080/10986065.2011.538302

Pinto Sosa, J. E. (2020). El cambio en el currículo en Educación Estadística en el ámbito universitario: dificultades y retos. *Matemáticas, Educación y Sociedad, 3*(2), 56–74. https://bit.ly/3Ptd29b

Rivera Pérez, J. G., & Gómez Blancarte, A. L. (2022). Características de la enseñanza de la estadística en disciplinas de STEM en la Educación Superior. *Innovación Educativa, 22*(88), 9–31. https://bit.ly/3Rflj1W

Sepúlveda Vega, F., & Sánchez Sánchez, E. A. (2022). Conocimiento de profesores de bachillerato sobre un acercamiento informal a las pruebas de significación estadística. En A. Saucedo & D. Díaz-Levicoy (Eds.), *Formación del profesorado para enseñar estadística: retos y oportunidades* (pp. 457–481). Centro de investigación de educación matemática y estadística de la Universidad Católica del Maule. https://bit.ly/3sQgyBJ

Silvestre Castro, E., & Urrea Bernal, M. A. (2021). Análisis del razonamiento de estudiantes de bachillerato frente a una tarea introductoria al contraste de hipótesis. *SahuarUS Revista electrónica de Matemáticas, 5*(1), 16–34. https://bit.ly/44KUIg3

Steffe, L. P., & Thompson, P. W. (2000). Teaching experiment methodology: Underlying principles and essential elements. En A. E. Kelly & R. A. Lesh (Eds.), *Handbook of Research Design in Mathematics and Science Education* (pp. 267–307). Erlbaum.

Tobías-Lara, M. G., & Gómez-Blancarte, A. L. (2019). Assessment of informal and formal inferential reasoning: A critical research review. *Statistics Education Research Journal*, *18*(1), 8–25. https://bit.ly/44PP50a

Toulmin, S., Rieke, R., & Janik, A. (1984). *An introduction to reasoning* (2a ed.). Macmillan Publishing Company.

Zieffler, A., Garfield, J., delMas, R., & Reading, C. (2008). A framework to support research on informal inferential reasoning. *Statistics Education Research Journal*, *7*(2), 40–58. https://bit.ly/488xRhu