

UNIVERSIDAD DE SONORA

Facultad Interdisciplinaria de Ciencias Exactas y Naturales

Departamento de Matemáticas

Doctorado en Ciencias con Especialidad en Matemática Educativa



**Avance de Tesis para
presentación de examen
predoctoral**



Título del trabajo:

Una trayectoria hipotética de aprendizaje para la enseñanza de la prueba de hipótesis desde un enfoque inferencial estadístico informal para estudiantes de nivel superior

Presentado por:

M.C. Jovan Israel Segundo Rosas

Dirección de tesis:

Dr. Eleazar Silvestre Castro

Noviembre de 2025

1. Introducción

A lo largo de las prácticas de los docentes se ha constatado que las pruebas de hipótesis (PH) es un tema de Inferencia Estadística que los estudiantes encuentran muy difícil de entender. Este tema, además de ser fundamental en cualquier área profesional, suele resultar complejo, en gran medida porque para comprenderlo no basta con dominar procedimientos técnicos, sino que es necesario haber construido previamente una serie de ideas clave.

Entre esas ideas se encuentra, en primer lugar, el reconocimiento de que los fenómenos que estudiamos involucran aleatoriedad, lo cual implica aceptar que nuestras inferencias nunca serán seguras y que siempre existirá la incertidumbre. A esto se suma la necesidad de estar familiarizado con los procesos de muestreo y con los principios de la teoría de la probabilidad, que constituyen la base sobre la cual se construye el razonamiento inferencial formal.

Diversos estudios (mencionados más adelante) han señalado que, en los cursos de estadística, es común que los estudiantes enfrenten las PH desde una perspectiva meramente procedimental, limitándose a memorizar fórmulas y algoritmos para aplicarlos mecánicamente. Sin embargo, con frecuencia esto no se acompaña de una comprensión real de las nociones subyacentes, lo que provoca que los alumnos no sean capaces de interpretar de forma adecuada los resultados obtenidos, ni qué papel juegan la incertidumbre y la variabilidad en dicho proceso.

En este contexto, el interés de este trabajo es abonar al desarrollo de material didáctico para la enseñanza de las PH, con el propósito de facilitar su comprensión y hacer más accesible un tema que los estudiantes suelen percibir como abstracto y distante. La idea central de esta intervención didáctica es facilitar el acceso a las nociones que conforman las PH mediante una aproximación informal, que permita a los estudiantes transitar progresivamente hacia el formalismo matemático propio de estas pruebas.

Para ello, se asume como foco teórico el enfoque de la Inferencia Estadística Informal (IEI), el cual ofrece un marco pertinente para introducir a los estudiantes en la toma de decisiones bajo incertidumbre a partir de datos, antes de entrar en la formalidad matemática. Para este trabajo, dicho enfoque se visualiza como una estrategia que puede acercar a los alumnos a construir significados a partir de la experimentación, la simulación y la argumentación, lo cual abre el camino hacia una comprensión más sólida de las PH y de la inferencia estadística en general.

En concordancia con lo anterior, este trabajo se orienta al diseño de una trayectoria hipotética de aprendizaje (THA) que articule progresivamente las ideas necesarias para comprender las PH. Dicha propuesta busca responder a áreas de oportunidad identificadas en investigaciones previas sobre la enseñanza de las PH y se plantea como un proyecto de

intervención didáctica, cuyo propósito es ofrecer alternativas pedagógicas que favorezcan un aprendizaje más significativo en los estudiantes de nivel superior, específicamente en las áreas de las Ciencias Económicas y Administrativas.

2. Antecedentes

2.1 Inferencia Estadística y la Prueba de Hipótesis

En el contexto de la Estadística, uno de sus fines, es la Inferencia. Una inferencia es una conclusión de que los patrones en los datos están presentes en un contexto más amplio (Ramsey y Schafer, 2002; citado por Rossman, 2008). Refiriéndonos a la Inferencia Estadística, ésta es una inferencia justificada por un modelo probabilístico que vincula a los datos a un contexto más amplio (Ramsey y Schafer, 2002; citado por Rossman, 2008). Desde este punto de vista, podemos identificar que se trata de un proceso inductivo, es decir, a partir de una muestra recolectada, se emiten juicios o predicciones acerca de cómo podría ser la población de los datos con respecto a una cierta variable aleatoria.

No obstante, los métodos propuestos en la literatura sólo son alternativas que nos ayudan a esbozar cómo pudiera ser tal población con una argumentación teóricamente sólida pero que en realidad pudiera diferir de dicha teoría, por lo que estos sólo nos aportan un acercamiento a dicha inferencia con un cierto grado de incertidumbre; dada la naturaleza de la inferencia, hasta el momento no existe alguna estrategia que nos lleve a predecir la población de una manera certera.

Lo anterior suele ser un problema al momento de introducir al estudiante al estudio de la Estadística Inferencial, y más aún, en la toma de decisión que se procura en una PH, como una de las técnicas utilizadas en dicha área de estudio, donde no hay certeza si la hipótesis nula es cierta o no. Las PH son un proceso inductivo en el cual tratamos de encontrar suficiente evidencia como para rechazar o no una hipótesis en particular; dichas hipótesis suelen ser estimaciones, posibles o no, de un parámetro (Konold y Pollatsek, 2002).

Cuando obtenemos una muestra de datos podemos especular sobre cómo podría ser alguno de los parámetros de la población de donde salió dicha muestra, sin embargo, ¿qué tan certeros podemos ser? En ocasiones podemos proponer una hipótesis de cómo pudiera ser algún parámetro de la población, aunque con la incertidumbre de si ésta será verdadera o falsa, lo que sí sabemos es que, en definitiva, una hipótesis siempre será verdadera o falsa, nunca será verdadera algunas veces y falsa en otras.

Para resolver una PH, actualmente se puede recurrir a tres métodos, un par, propuestos desde la inferencia frecuentista, el enfoque de Ronald Fisher y el método propuesto por Neyman y Pearson; y un tercero, pensado desde un enfoque probabilístico llamado Inferencia Bayesiana (Batanero, 2000; Batanero, Díaz, 2015), aunque en este trabajo sólo se enfocará en los primeros dos mencionados debido a que éste sigue un sentido probabilístico que difiere totalmente a la filosofía frecuentista.

El enfoque de Fisher, propuesto desde la inferencia frecuentista, en su libro “The design of experiments” publicado en 1935, introdujo las pruebas de significación donde se permite rechazar una hipótesis, si el valor-p es muy pequeño (García, 2017). Por otro lado, Neyman y Pearson, se interesaron en las pruebas de hipótesis como proceso de decisión que permite elegir entre una hipótesis nula (H_0) u otra alternativa (H_1) (Rivadula, 1991). Este enfoque tiene mayor sentido en pruebas que se efectúan de forma repetitiva bajo las mismas condiciones. Algunas de sus diferencias pueden apreciarse en la Tabla 1.

	Enfoque de Fisher	Enfoque de Neyman y Pearson
Hipótesis en contraste	Sólo la hipótesis nula (H_0)	Hipótesis nula (H_0) contra Hipótesis alternativa (H_1)
Objetivo	Buscar evidencia suficiente para rechazar H_0 . Si el valor-p es muy pequeño se rechaza H_0 .	Tomar una decisión entre: - Rechazar H_0 y aceptar H_1 . - Aceptar H_0 y rechazar H_1 . Esto también depende del nivel de significación (α).
Implicaciones de la solución a un problema	No está claro qué tan pequeño debe ser el valor-p para poder rechazar H_0 . Es subjetivo.	- Si rechazamos H_0 y aceptamos H_1 , siendo H_0 cierta, se cometería un error tipo 1, aunque puede ser controlable mediante un α pequeño. - Si aceptamos H_0 y rechazamos H_1 , siendo H_1 cierta, se cometería un error tipo 2, error que puede ser muy difícil controlar.
Otra característica	La prueba se realiza una vez.	La prueba se puede realizar más de una vez.

Tabla 1: Características de las filosofías de la PH desde el enfoque frecuentista.

Estas diferencias evidencian la necesidad de repensar la forma en que el tema se aborda en la enseñanza universitaria, sin embargo, en los programas de nivel superior de la Universidad de Sonora, las PH, que suelen hacer acto de presencia en áreas como las Ciencias Económicas y Administrativas (Programa de Estadística Inferencial, Licenciatura en Finanzas, UNISON, 2022), se propone introducir las PH como un híbrido entre los dos

enfoques frecuentistas (Batanero, 2000), esto sin ninguna aclaración inicial en particular, por lo que se considera que la transposición didáctica sugerida no comparte los orígenes epistemológicos de éstos, lo que, en algún momento, puede causar un obstáculo epistemológico que impida, por ejemplo, una correcta interpretación del resultado de una PH.

En esta área, las PH se consideran de mucha importancia. En el campo de las finanzas y los negocios se desea que los estudiantes puedan “determinar propiedades de una población estadística, a partir de una parte de ésta y obtener conclusiones útiles para hacer deducciones sobre una totalidad, basándose en la información numérica de la muestra” (UNISON, 2022), y una de las estrategias que se propone incluir es la PH. Las PH tienen muchas aplicaciones, por ejemplo, si se desea probar una suposición de cómo se comporta un mercado, o ¿qué tan cierta pueda ser alguna situación financiera para poder tomar la decisión de si se invierte o no en un mercado en específico? (UNISON, 2022)

Así mismo, en otro programa, el objetivo que se persigue es el que menciona “el alumno aplicará los fundamentos y las técnicas de la inferencia estadística en el estudio de situaciones de incertidumbre que enfrentan la contabilidad y los negocios, adoptando modelos pertinentes para la variabilidad en muestras, utilizando intervalos de confianza y formulando contrastes de hipótesis adecuados para la toma de decisiones a partir de evidencia empírica acerca de las variables de interés presentes en muestras.” (UNISON, 2018).

Ciertamente, el algoritmo de cómo resolver una PH ya es conocido y divulgado en la literatura, sin embargo, ¿los estudiantes entienden las ideas que hay detrás de una PH? o ¿pueden interpretar el resultado obtenido? Diversos investigadores han reportado dificultades que los estudiantes demuestran a través de los errores cometidos al momento de resolver una PH, que también se cometen al interpretar el resultado obtenido. Dichas dificultades se describen más adelante.

Desde un punto de vista epistemológico, según Batanero y Díaz (2015) y Rossman (2008), al resolver una PH, el resultado obtenido puede no ser lo suficientemente contundente como para no equivocarnos, quien resuelve una PH debe estar consciente de que puede equivocarse al rechazar o no una hipótesis y, de ser posible, con qué probabilidad se podría estar equivocado. Aunque resolvamos una PH por los métodos existentes, ello no nos asegura que podamos tener la certeza de la veracidad de una hipótesis.

2.2 Dificultades y retos en la enseñanza y aprendizaje de las pruebas de hipótesis

2.2.1 Algunas dificultades previstas desde la epistemología

Ambos enfoques, previamente mencionados, provenientes de la inferencia frecuentista que, si bien son similares, parten de pensamientos distintos y persiguen diferentes objetivos. En el currículo escolar se suele encontrar una hibridación entre éstos, sin embargo, por lo mencionado anteriormente, considero que, en principio, ésta pueda ocasionar dificultades en el proceso de enseñanza-aprendizaje, así como lo señalan Batanero (2000) y Batanero y Díaz (2015).

En la literatura se han documentado, ampliamente, distintos tipos de dificultades y errores cometidos por los estudiantes, desde los conceptuales hasta los procedimentales, en la enseñanza tradicional de las PH. A continuación, se mencionan algunos ejemplos.

En Batanero y Díaz (2015), se reportan algunos errores que los estudiantes comenten con frecuencia, estos derivados de interpretaciones incorrectas de las pruebas de hipótesis. Por ejemplo, el concepto de nivel de significancia α , siendo la probabilidad de cometer el error tipo 1 al momento de rechazar H_0 suponiendo sea cierta, es decir, $\alpha = P(\text{Rechazar } H_0 \mid H_0 \text{ es cierta})$; el error más común cometido por los estudiantes es el de intercambiar los dos términos de la probabilidad condicionada, es decir, $\alpha = P(H_0 \text{ es cierta} \mid \text{se ha rechazado } H_0)$.

En Falk (1986) muchos de sus estudiantes pensaban que α era la probabilidad de equivocarse al rechazar H_0 . Este tipo de error se repitió en muchos más estudiantes (Vallecillos, 1994). También se presentan confusiones entre las hipótesis nula y alternativa, algunos estudiantes piensan que la hipótesis nula es la que queremos demostrar (no la que queremos rechazar) (Batanero y Díaz, 2015).

Vallecillos (1994) reporta que existen dificultades para identificar si la hipótesis de la prueba se refiere al parámetro poblacional o al estadístico muestral, siendo que el parámetro es constante a diferencia del estadístico que puede variar según la muestra recabada. Vallecillos (1999) también describe ciertas creencias de los estudiantes que se considera hay que tomar en cuenta. Dichas creencias van desde confundir los dos enfoques simplemente pensando que la PH siempre consiste en decidirse por alguna de las dos hipótesis planteadas; pensar que una PH sólo se realiza una vez, sería una concepción errónea si se piensa en el enfoque de Neyman y Pearson, entre otras.

Vera, Díaz y Batanero (2011) reportan errores que cometen los estudiantes al confundir si la prueba será unilateral o bilateral. Además, los estudiantes, en lugar de plantear un problema de PH de media en una única población, plantean pruebas de comparación entre dos medias en poblaciones diferentes.

En Case y Jacobbe (2018) se menciona que los estudiantes interpretan incorrectamente el sentido del valor-p, para ellos, esto es como un valor probabilístico que

indica qué tan probable es que la hipótesis nula sea falsa o verdadera, como si fuese verdadera unas veces y falsa en otras.

En Inzunza y Jiménez, (2013), enfatizan que “las principales concepciones erróneas asociadas al aprendizaje y uso de las PH, se les atribuye a dos aspectos fundamentales”, y uno de ellos es la hibridación entre los enfoques de Fisher y Neyman-Pearson que aún se siguen promoviendo en la bibliografía sugerida en los programas de cursos.

Castro Sotos et al. (2009) y Krishnan e Idris (2014) reportan que, “entre los errores más frecuentes, los estudiantes creyeron que una PH es una prueba matemática de la hipótesis nula, o que es una prueba probabilística por contradicción.”, es decir, como tratándose de tener que demostrar la veracidad de la hipótesis nula. Inclusive, es estadísticamente incorrecto tomar la decisión de aceptar la hipótesis nula dado que no podemos tener la certeza de que así sea cualquiera de las hipótesis planteadas.

En Retamal, Martínez, Sáez y Lugo-Armenta (2023), en su estudio realizado con estudiantes de ingeniería, se reportan dificultades en la formulación de las hipótesis en contraste. Es común que los estudiantes cometan errores al momento de establecer cuáles son las hipótesis de la prueba, mencionan que existe una falta de comprensión de la lógica en cuestión. También identifican errores en las interpretaciones de los resultados de una PH, escritura inadecuada de las conclusiones y la mala interpretación de la información del problema.

Pero también, se han reportado errores cometidos por los profesores al momento de la puesta en escena de su intervención didáctica. Birnbaum (1982) pidió a sus estudiantes que expresaran el significado de “nivel de significación”, a lo que, la mayoría de ellos contestó “Un nivel de significancia del 5% significa que, en promedio, 5 de cada 100 veces que rechazamos la hipótesis nula estaremos equivocados”. Resultados como este, fueron encontrados por Krauss y Wassner (2002) en profesores de universidad encargados de impartir dichos contenidos. Cobb (2007) menciona que, en la enseñanza tradicional, hacer cambios drásticos en el proceso de enseñanza como comparar entre dos poblaciones en vez de realizar inferencia sobre una, representa un salto muy grande en la didáctica asociada, como para que el estudiante entienda las ideas que den continuidad.

2.2.2 Pruebas de hipótesis: Algunas complejidades en el razonamiento inductivo y retos de su enseñanza

Otras dificultades han sido reportadas en la literatura respecto al razonamiento estadístico involucrado. En general, los conceptos y el razonamiento que requiere la

inferencia estadística son difíciles para la mayoría de los estudiantes, incluso para profesores e investigadores que la aplican en su vida profesional (Vallecillos y Batanero, 1997; Lipson, 2000; Liu y Thompson, 2005; Canal y Behar, 2010; Dolor y Noll, 2015; Nilsson, 2023). En las pruebas de hipótesis son muchos los conceptos que intervienen como: muestreo, variabilidad muestral, población, distribución muestral, distribución de medias muestrales, nivel de significancia, los tipos de errores, media muestral, media poblacional, distribución de probabilidad, etc. Realizar una prueba de hipótesis involucra entender cada uno de estos conceptos, distinguir entre población y muestra, reconocer que el tamaño de la muestra puede afectar en la toma de decisión realizada, además de aplicarlos bajo la lógica de un razonamiento inductivo.

Así mismo, Konold y Pollatsek (2002) e Inzunza e Islas (2019), reportan dificultades presentadas en la comprensión, razonamiento y pensamiento sobre diversos conceptos de la Estadística Inferencial. Para la mayoría de los estudiantes, la Estadística Inferencial es un área sumamente compleja, quizás dicha concepción sea producto de lo que señala Cobb (2007):

“cuya enseñanza tradicional ha estado basada en una serie de simplificaciones y enfocada en el desarrollo de habilidades matemáticas para uso de fórmulas y procedimientos, las cuales, muchas veces resultan insuficientes para que los estudiantes razonen y piensen estadísticamente”.

destacando que su formación previa es deficiente.

Batanero (2006), Castro, Vanhoof, Van den Noortgate y Onghena (2007), Batanero y Díaz (2015) y Van Dijke-Droogers, Drijvers y Bakker (2019), mencionan dificultades para aplicar métodos de inferencia estadística, para interpretar los resultados de éstos, que pueden estar muy relacionados con los distintos enfoques existentes (para la prueba de hipótesis), específicamente hablando de los enfoques de Fisher y Neyman y Pearson. En la enseñanza tradicional, se observa una falta de especificidad al respecto de cuál de estas filosofías es la empleada en el aula de clase para resolver e interpretar una prueba de hipótesis, inclusive libros de texto como “Estadística para Administración y Economía” de Anderson, Sweeney y Williams (2008) o también “Estadística Aplicada a los Negocios y la Economía” de Webster (2000) se observa una especie de hibridación entre ambas y se trabaja a través de una serie de pasos y procedimientos, planteando un par de hipótesis (nula y alternativa) con el objetivo de elegir una de las dos (Yates, Moore y McCabe, 1998, citados por Liu y Thompson, 2009).

La enseñanza tradicional, que se ha centrado en la formalidad matemática de la estadística, ha propiciado este tipo de dificultades en el entendimiento de la prueba de hipótesis. Batanero y Díaz (2015) describen cómo se suele resolver una prueba de hipótesis en el aula de clase, desde la formalidad, sin embargo, también se dejan entrever implicaciones

que tiene esta dinámica; si el estudiante no hace uso de su conocimiento estadístico previo, o si no ha comprendido los conceptos involucrados como distribución de medias muestrales, distribución de probabilidad o nivel de significación, sólo observará la prueba de hipótesis con el propósito de memorizar la estrategia utilizada.

En contraposición, en la literatura revisada se observan intentos iniciales sobre bosquejos de propuestas de lo que se podría hacer en el aula de clase al momento de la enseñanza de las pruebas de hipótesis para el nivel superior, tanto Rossman (2008), como Batanero y Díaz (2015), señalan que un buen camino para acercarse de manera empírica o informal es a través de la simulación de muestras por computadora. Generar más muestras mediante simulación puede permitir la observación de lo típico y lo atípico de una población (Cobb, 2007).

2.2.3 Los profesores y las PH.

Aunado a lo anterior, y cómo ya se empezaba a encontrar en la literatura, existen dificultades en los mismos profesores acerca de cómo emplear, e interpretar el resultado de una PH; algunos docentes poseen concepciones erróneas que después se transfieren a los estudiantes. Liu y Thompson (2009) reportaron errores en la forma de pensar de algunos profesores que imparten los contenidos propios de la estadística inferencial, específicamente en la PH; ellos mencionan que hay profesores que se resisten a rechazar una hipótesis nula, aunque los datos muestrales hayan resultado atípicos suponiendo la hipótesis nula sea cierta, es decir, los docentes dicen que, no pueden rechazar la hipótesis nula, porque quizás sólo tuvimos la suerte de recolectar una muestra atípica.

Diversas investigaciones han evidenciado que tanto los profesores en activo como los profesores en formación presentan limitaciones conceptuales en torno a la lógica inferencial que sustenta las pruebas de hipótesis. Estas dificultades se manifiestan en la comprensión parcial de conceptos como la hipótesis nula, el nivel de significancia, el valor-p y el vínculo entre los datos muestrales, la variabilidad y la decisión estadística. Por ejemplo, Liu y Thompson (2007) reportaron que muchos docentes tienden a interpretar la PH desde un enfoque determinista, sin reconocer el razonamiento lógico subyacente ni la naturaleza estocástica del proceso.

Una dificultad recurrente es la escasa comprensión de la variabilidad muestral y de las distribuciones muestrales. Los profesores suelen tener problemas para concebir el muestreo repetido como un proceso generador de distribuciones y, en consecuencia, para interpretar la probabilidad asociada a la ocurrencia de resultados extremos (Matuszewski, 2018). Esta falta de comprensión limita su habilidad para explicar cómo la inferencia estadística se basa en la comparación entre la evidencia empírica y lo que se esperaría bajo la hipótesis nula.

Asimismo, se ha documentado que muchos docentes presentan una dependencia procedimental en la enseñanza de las PH, centrandó su atención en la aplicación mecánica de fórmulas y en la toma de decisiones basadas en el valor crítico o el valor-p, sin promover una comprensión profunda del sentido inferencial de estos procedimientos (Liu y Thompson, 2007; Matuszewski, 2018). Dicho enfoque puede conducir a una visión reduccionista de la inferencia estadística, en la que el razonamiento probabilístico y la interpretación contextual de los resultados se ven relegados a un segundo plano.

Las investigaciones más recientes apuntan a que el uso de simulaciones y actividades prácticas contribuye significativamente a superar estas dificultades. Matuszewski (2018) demostró que, cuando los docentes participan en experiencias de aprendizaje que incorporan simulaciones computacionales, logran visualizar mejor la variabilidad muestral y comprender de manera más intuitiva el significado del valor-p como una medida de rareza bajo la hipótesis nula. Del mismo modo, estudios actuales sostienen que el enfoque de inferencia estadística informal, basado en la generación empírica de distribuciones muestrales mediante simulación, favorece el desarrollo de un pensamiento inferencial más coherente y conectado con la noción de azar (Garfield y Ben-Zvi, 2008; Pratt et al., 2020).

En el caso de los profesores en formación, Yeh y Santagata (2015) encontraron que, mediante programas de desarrollo profesional que incluyen análisis de datos, formulación de hipótesis y argumentación basada en evidencia, los futuros docentes mejoran su comprensión sobre cómo se construye y se evalúa una hipótesis estadística. Sin embargo, incluso con estas mejoras, persisten malentendidos respecto al valor-p frecuentemente interpretado como la probabilidad de que la hipótesis nula sea verdadera, así como dificultades para articular los resultados de una prueba con el contexto del problema o con el razonamiento probabilístico subyacente (Chance & Rossman, 2008; Liu & Thompson, 2007).

En síntesis, esto también es causante de los errores que cometen los estudiantes y de las dificultades que los mismos presentan al no comprender la lógica de las PH, que a su vez impide la correcta interpretación de los resultados obtenidos al resolver una PH.

2.3 Inferencia Estadística Informal (IED): alternativa para la enseñanza de la inferencia

En años recientes, se han producido diversos trabajos de investigación con los que se ha buscado la manera de hacer más accesible a los estudiantes los contenidos de la Estadística Inferencial. Dentro de éstas, algunas se han pronunciado por realizar el acercamiento a estos conceptos mediante la Inferencia Informal (Zieffler, Garfield, delMas y Reading, 2008) y otros, pensando específicamente en las PH, o en las ideas que la conforman, mediante la

Inferencia Estadística Informal (Batanero y Díaz, 2015; Van Dijke-Droogers, Drijvers y Bakker, 2019; Nilsson, 2023; Silvestre, Armenta e Inzunza, 2024).

Para entender qué es la Inferencia Estadística Informal (IEI), si bien, algunos autores han tratado de dar una definición a este concepto, me parece que no están del todo completas. Pfannkuch (2006) definió el término Inferencia Estadística Informal como la extracción de conclusiones a partir de datos que se basa principalmente en observar, comparar y razonar a partir de distribuciones de datos (citado en Estrella, Méndez-Reina y Vidal-Szabó, 2023). Por su parte, Makar y Rubin (2009), identifican tres principios clave: generalización más allá de los datos, datos como evidencia de estas generalizaciones y razonamiento probabilístico sobre la generalización; aunque utilizando un enfoque informal también se puede utilizar la experiencia adquirida para hacer tales inferencias (citados en Van Dijke-Droogers, Drijvers y Bakker, 2019)

En Segundo (2018), a partir de los trabajos de investigación de Zieffler, et al. (2008) y Rossman (2008), se define al Razonamiento Inferencia Estadístico Informal como el proceso y el resultado de extender la información y características encontradas en una muestra de datos a un contexto más amplio, acompañada por una valoración o medición del grado de incertidumbre de esta generalización apoyada en un modelo probabilístico, función de distribución o distribución muestral.

Por lo anterior y considerando lo expuesto por los autores previamente mencionados, puedo decir que la Inferencia Estadística Informal es aquel juicio emitido, a partir de una muestra de datos a un contexto más amplio, acompañado de una valoración o medición del grado de incertidumbre apoyada por un modelo probabilístico, función de distribución o distribución muestral. La Inferencia Estadística Informal (IEI), aunque parte de lo informal o empírico, incluso puede ir acompañada de una valoración en términos de la creencia y experiencia del estudiante, así como lo encontrado en los trabajos de Van Dijke-Droogers, Drijvers y Bakker, (2019) y Nilsson (2023).

La Inferencia Estadística Informal ha ido tomando fuerza en los últimos años dado que ha mostrado tener resultados interesantes al momento de implementar dicho enfoque en el proceso de enseñanza-aprendizaje de los estudiantes. En particular, algunas investigaciones dan cuenta de que apoyarse en este enfoque permite una mayor accesibilidad al concepto de PH sin tener que mencionar directamente los conceptos básicos de la inferencia; esto con el fin de formalizarlos en un momento futuro, ya que el concepto de PH sea mejor comprendido por los estudiantes. En la siguiente sección describiremos con detalle algunas de estas propuestas.

La IEI puede aportar una base cognitiva sólida para comprender las PH. Antes de introducir fórmulas, los estudiantes podrían ya entender que:

- los datos varían;
- algunos resultados son raros bajo ciertos supuestos;
- la rareza constituye evidencia contra H_0 .

3 Análisis de aportaciones clave

3.1 Propuestas de enseñanza desde un acercamiento informal.

En años recientes se han publicado propuestas de enseñanza de conceptos de la Estadística Inferencial, desde el enfoque de la Inferencia Estadística Informal (IEI), y en particular con respecto a las PH. Este enfoque surge a partir de la identificación de errores y dificultades de los estudiantes en torno al aprendizaje de la Estadística Inferencial. Como ya se mencionó en la sección anterior, tradicionalmente se recurre a la formalidad matemática desde el inicio del curso, lo que deviene en la mecanización a través del uso de fórmulas y algoritmos para la realización de cálculos, lo que a su vez produce que el estudiante no logre argumentar sus resultados de la forma correcta.

En Stohl, Angotti y Tarr (2010) se realizó una investigación con estudiantes de 11 y 12 años en la que se observa cómo los estudiantes razonan con respecto a los resultados de experimentos de lanzamiento de dados simulados por computadora para apoyar o refutar la suposición de que los resultados de un dado son equiprobables (ver Figura 1). Los autores examinaron las acciones de los estudiantes con el software y sus interacciones sociales para inferir sus expectativas y si creían que sus datos empíricos podían usarse para rechazar la hipótesis de resultados equiprobables.

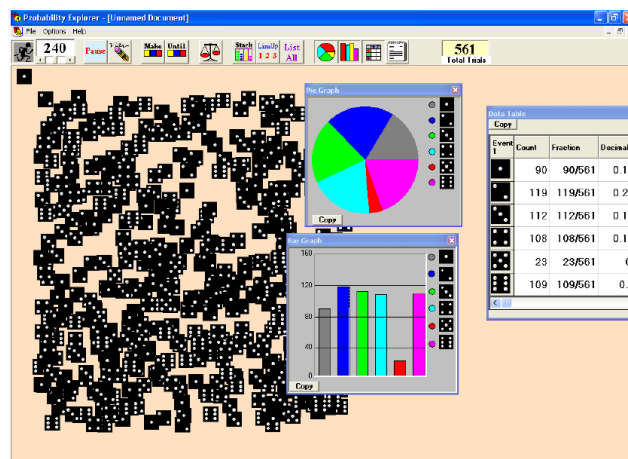


Figura 1

En éste utilizan los resultados experimentales como una distribución muestral empírica (aunque simulada) que después es vinculada con la distribución teórica para argumentar la decisión tomada. Sin embargo, en ese nivel educativo los grados de razonamiento, así como el conocimiento y la experiencia adquirida hasta el momento, suelen presentarse más débiles en comparación con los estudiantes de nivel superior, a quienes está dirigido este proyecto de intervención. Sin embargo, este tipo de propuesta no se apega a la epistemología del concepto PH, por lo que pudiera ocasionar otro tipo de dificultades.

En el trabajo de García y Sánchez (2014) se aplicó un cuestionario a estudiantes de bachillerato de entre 16 y 17 años, para identificar su grado de razonamiento inferencial informal con respecto a las PH de proporciones. Se les planteó a los estudiantes un par de problemas en los que una cierta industria ofreció un producto llamado “Gender Choice”, el cual ayuda a favorecer la probabilidad de que una pareja conciba una niña en el próximo embarazo. Se les mostró a los estudiantes un par de casos, el primero donde el 52% de los nacimientos fueron niñas y el segundo donde 90 de 100 parejas tuvieron una niña.

Dichos estudiantes se encontraban iniciando su tercer semestre del bachillerato y no estaban cursando alguna asignatura en la que se tratara la Estadística Inferencial, por lo que su conocimiento al respecto era escaso. Análogamente, los autores comentan que, al tratar de analizar los datos mediante las componentes del marco de Zieffler et al. (2008), encontraron que éstos se cumplían de manera incipiente, por lo que optaron por utilizar otras categorías. Esto provoca considerar la posibilidad de encontrar una mayor oportunidad de acercar al estudiante a las PH en el nivel superior y después de haber cursado alguna asignatura previa, o en general, haber adquirido una mayor madurez en su capacidad de razonamiento.

3.2 Propuestas que integran una THA.

Algunos autores proponen que a través de una THA es posible construir un diseño instruccional que pueda movilizar las pautas educativas propuestas en trabajos previos de investigación, con una cierta secuencia, en la que también podemos especular a priori hipótesis de aprendizaje, es decir, los razonamientos que se espera que el estudiante demuestre y así tener una evidencia de su grado de aprendizaje. En el apartado de los elementos teóricos se hablará de esto con más detenimiento.

Por lo anterior, otros trabajos se han ido desarrollando, siguiendo el enfoque informal, y en particular a través del diseño de una THA, que le abonan a la construcción de diseños instruccionales que puedan llevar al estudiante por una dirección más favorable hacia el aprendizaje de los conceptos de la Estadística Inferencial y en particular de las PH.

La Trayectoria Hipotética de Aprendizaje (THA) siendo introducida por Simon (1995) y Simon y Tzur (2004), es un modelo teórico para el diseño instruccional en educación matemática (Simon, 2020). Consiste en tres componentes, un objetivo de aprendizaje, un conjunto de tareas e hipótesis de aprendizaje.

Estrella, Méndez-Reina y Vidal-Szabó (2023) propusieron una THA conformada de 4 pasos, donde a partir de un juego llamado Frog Race (carrera de ranas) estudiantes de 8 años registraban en una hoja de trabajo lo sucedido al lanzar un par de monedas. Los 4 pasos se establecieron de la siguiente manera:

- Paso 1: Hacer predicciones que contrasten con los datos obtenidos.
- Paso 2: Visualizar y reconocer la variación en la muestra.
- Paso 3: tomar conciencia del comportamiento regular de las muestras asignando un nivel de confianza a cada evento.
- Paso 4: Generar afirmaciones más allá de los datos disponibles como evidencia utilizando expresiones de incertidumbre.

En la hoja se observaban tres ranas, una rana anaranjada, otra rosa y una tercera de color azul (ver Figura 2). Al lanzar las dos monedas se identifica si ambas caras de la moneda son iguales o diferentes, y si son iguales qué lado ocurrió. Si en el lanzamiento en ambas monedas caía cruz, la rana anaranjada avanzaba una posición; si en ambas monedas caía cara, entonces la rana rosa avanzaba una posición y si el lado mostrado en ambas monedas era distinto entre sí, la rana azul se movía una posición. Los estudiantes llevaron a cabo el juego (experimento) y se terminó cuando alguna de las ranas ganaba, es decir, llegaba a la meta; posteriormente se les preguntó a los estudiantes ¿qué rana crees que gane en una segunda carrera?, para esto, dado que el estudiante ya tendría una primera experiencia de lo que podría suceder, en varios casos, hicieron uso de ese conocimiento para argumentar su predicción sobre el resultado de una segunda carrera.

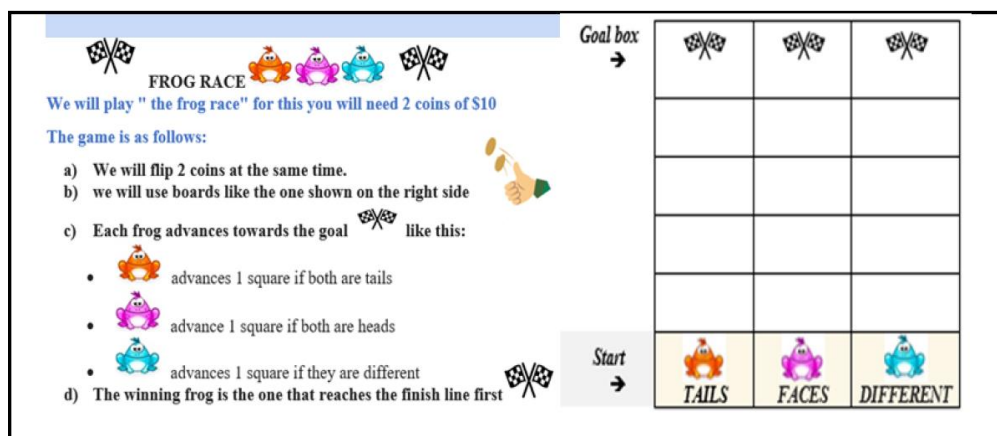


Figura 2

Esta THA comienza a acercar a los estudiantes a algunas de las ideas que conforman la PH, como el planteamiento de una hipótesis, así como después sembrar la duda con respecto a ésta de acuerdo con las muestras recabadas en cada juego. Sin embargo, no se llega a utilizar algún modelo de probabilidad, ni siquiera se construye una distribución muestral empírica (DME) que pueda ser utilizada como modelo para la toma de decisión. En otras palabras, no se llega a la resolución de una PH.

Por su parte, Nilsson (2023), también para estudiantes de nivel básico (11 y 12 años), considerando que es un nivel en el que se puede comenzar a plantear las ideas de la Inferencia Estadística, aunque informalmente, propone una THA que consiste en 3 lecciones:

- Lección 1: Promover el razonamiento del espacio muestral modificando dicho espacio muestral en Color Run.
- Lección 2: Comparar muestras del mismo generador aleatorio.
- Lección 3: Comparar muestras de diferentes generadores aleatorios.

Cada lección se conforma por cada una de las tareas propuestas en Zieffler, et al. (2008). Esta trayectoria sigue el juego llamado “Color Run”, en el que se presenta una botella llena de canicas y un tablero de juego, la idea es agitar la botella, sacar una canica y registrar en el tablero el color de la canica extraída. El juego finaliza cuando uno de los colores llega hasta 7 repeticiones (ver Figura 3).

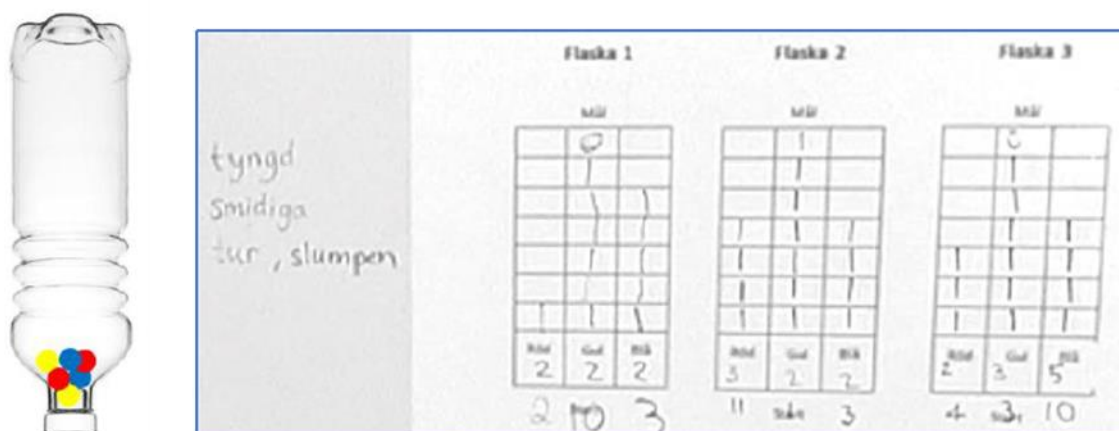


Figura 3

En la lección 1, la botella es transparente, de tal forma que las canicas sean visibles, y que el estudiante pueda observar completamente lo que sucede en el proceso del muestreo, cómo es la muestra extraída con respecto a las canicas que se veían en la botella al inicio. En la lección 2 se pidió a los estudiantes que compararan entre muestras que fueron extraídas del mismo generador aleatorio (misma botella). Para la lección 3 se prepararon 3 botellas distintas, esta vez no se podían ver las canicas que había dentro de las botellas; de éstas, una

botella contenía una cantidad distinta de canicas y los estudiantes debían adivinar cuál era la botella diferente. Al realizar las actividades, se observó que los estudiantes emitieron sus predicciones con base en los resultados de los muestreos, incluso, emitieron una valoración del grado de certeza que ellos asignaban a la probabilidad de haber tenido la botella distinta.

Así como en la anterior, esta THA tampoco tiene por objetivo culminar con la resolución de una PH, sólo se queda en algunas ideas que conforman su lógica. Del mismo modo, sólo se realizan muestreos físicos, lo que limita la construcción de una DME competente para usar como modelo de referencia.

van Dijke-Droogers, Drijvers y Bakker (2019) proponen una THA de 3 pasos, basada en diversas pautas educativas, previamente reportadas en la literatura, para construir una secuencia didáctica, destinada a estudiantes de 15 años.

Con ésta se introduce a los estudiantes a los conceptos clave de la Inferencia Estadística Informal. Los estudiantes investigaron el contenido de una caja negra llena de esferas recopilando, intercambiando y comparando resultados de muestras físicas y posteriormente simuladas con diferentes tamaños y diferente número de repeticiones. Después preguntan a los estudiantes ¿qué pasaría si se repite el experimento? Para esto, los estudiantes comparten sus expectativas acerca de cuántas esferas amarillas habrá en una muestra de 40 esferas, siendo que la caja negra contiene 750 esferas amarillas y 250 esferas no amarillas y discutir los posibles resultados en la muestra. Sin embargo, en el tercer paso de la THA, proponen el uso de una herramienta computacional para simular el proceso de muestreo tratando de observar si dicha herramienta puede influir en los estudiantes para mejorar las predicciones efectuadas. Al pasar a las simulaciones por computadora se observó que los estudiantes eran capaces de simular las distribuciones muestrales, cambiando los tamaños de muestra y variando el número de repeticiones, de tal forma que utilizaron dichas distribuciones como modelos para interpretar la variación y la incertidumbre existente.

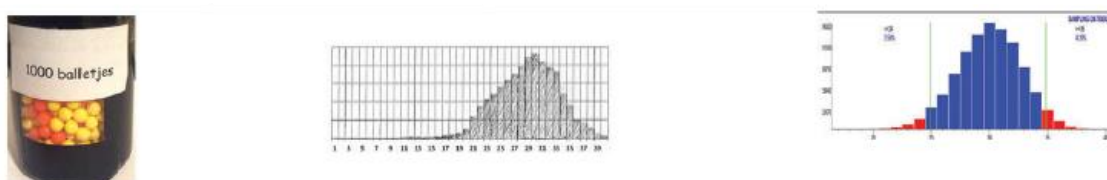


Figura 4: Caja negra con bolitas, distribución muestral empírica y distribución muestral simulada.

Fuente: van Dijke-Droogers, et al. (2019).

Si bien, en esta THA, los estudiantes construyeron la distribución muestral, tanto física como simulada, sólo llegan a interpretar la variación y la incertidumbre existente con

cierto grado de argumentación, mas no se realiza el ejercicio principal de la PH. En ésta ya se incluye el uso de una herramienta tecnológica digital, el simulador VU Stat.

Los autores también identificaron varios retos que aún siguen vigentes. Ellos reportan que “los estudiantes no saben cómo interpretar la variación de los datos, ya que notaron que no estaban completamente seguros de sus estimaciones debido a la variación de los resultados”. Por otro lado, “los estudiantes experimentaron y notaron que un tamaño de muestra más grande (generalmente) conduce a una menor variación en la estimación de la proporción de la población y, por lo tanto, a una mejor inferencia” (van Dijke-Droogers et al. 2019); sin duda el tamaño de muestra es un concepto clave para la inferencia, por lo que, lograr que los estudiantes tomen conciencia de ello, es de suma importancia.

También se ha identificado que no todos los estudiantes hacen uso de su conocimiento previo, por lo que recurren a factores más intuitivos para realizar la inferencia (van Dijke-Droogers et al. 2019).

Hasta este momento, las THA analizadas abordan, desde un enfoque informal, nociones fundamentales de la inferencia estadística, pero sin tener como objetivo la PH. Por otro lado, dos de éstas no incorporan el uso de tecnología digital, lo que limita la posibilidad de construir una distribución muestral. Esta carencia resulta particularmente relevante, pues la construcción de distribuciones muestrales constituye un paso necesario para plantear y resolver una PH. Cabe señalar que sólo una de las trayectorias revisadas recurre a simulaciones digitales de muestreo repetido, aunque su aplicación se queda en el tratamiento de conceptos elementales. Recordemos también que estas propuestas se trabajaron con estudiantes de niveles básico y preuniversitario.

A continuación, se presentarán algunos trabajos que sí se centran en la enseñanza de las PH a partir de la construcción de distribuciones muestrales, aunque empíricas simuladas, con apoyo de herramientas tecnológicas digitales.

En García (2017), siendo su tesis doctoral, a través del diseño de una THA intenta promover el desarrollo de razonamiento inferencia informal (RII) sobre las PH de proporciones de acuerdo a la filosofía de las pruebas de significación de Fisher, a través de la simulación de distribuciones muestrales utilizando Fathom como herramienta de software. Se implementó con 36 estudiantes de bachillerato organizados en 18 parejas. Se realizó en 4 sesiones de dos horas, donde los estudiantes resolvieron 4 problemas de PH de proporciones.

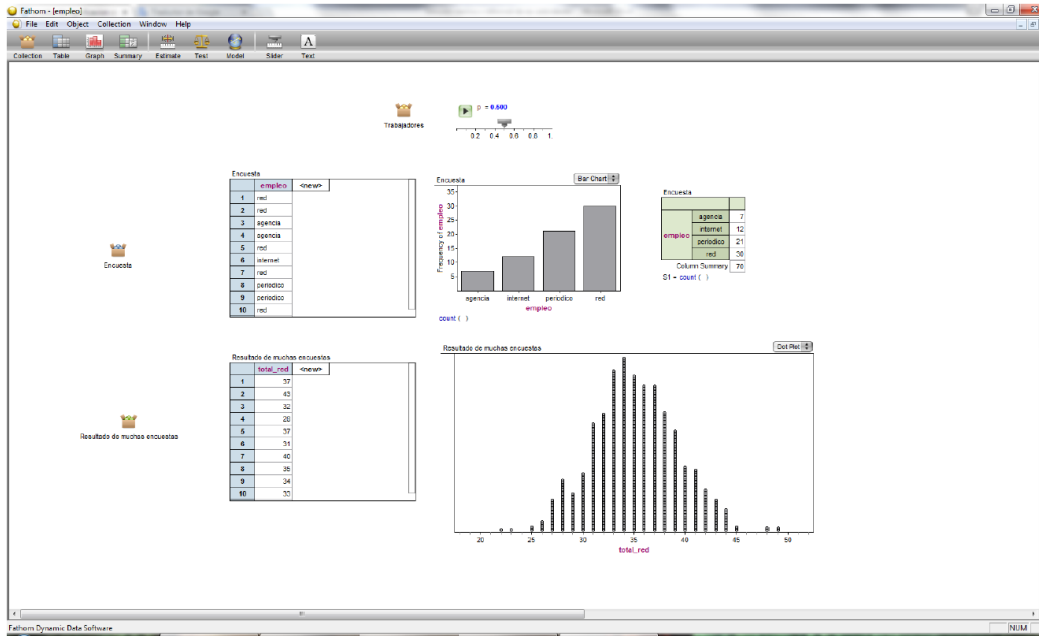


Figura 5: Distribución muestral simulada en Fathom.

Fuente: García (2017).

Sin embargo, esta intervención se realizó con estudiantes de bachillerato que, además, no estaban cursando alguna asignatura relacionada con la Probabilidad y la Estadística. Entre las observaciones que hace el autor, se menciona que, al inicio, hubo confusiones en el uso de la simulación, esto da indicios de que quizás es necesario iniciar por medio de la simulación física antes de utilizar la computacional.

Silvestre, Armenta e Inzunza (2024) realizaron un estudio, con estudiantes de nivel superior, de áreas económicas y sociales. Reportaron un primer ciclo de diseño de una THA, con el enfoque informal a la inferencia y haciendo uso de la filosofía de Ronald Fisher para la enseñanza de las PH. Implementaron el software Fathom como herramienta para la simulación de muestreo repetido para la creación de una DME, esto fue el punto de partida de la intervención.

Se formaron equipos de tres integrantes. La THA fue construida vía pautas educativas propuestas por autores como Zieffler et al. (2008), Rossman (2008) y Batanero y Díaz (2015). Ésta fue conformada por tres tareas, como se puede apreciar en la Tabla 2.

Tarea	Hipótesis general	Referencias
1	Partiendo de una situación inferencial de interés, en la que se trabaja con datos categóricos, los estudiantes, inicialmente, utilizarán su conocimiento contextual para evaluar la hipótesis nula de “no diferencia / no efecto / modelo 50-50”. En un segundo momento, analizando la misma situación, los estudiantes utilizarán la DME que modeliza la hipótesis nula de no diferencia como	<ul style="list-style-type: none"> - Partir de y utilizar conocimientos previos de estudiantes (ZIEFFLER <i>et al.</i>, 2008) - El enfoque de Fisher y el trabajo con datos

	vehículo para clasificar el resultado experimental como atípico, por lo cual la rechazarán. En un tercer momento, los estudiantes trasladarán estos razonamientos para analizar una situación inferencial distinta, en la que realizarán una prueba de hipótesis bajo el modelo 50-50 por cuenta propia.	categoricos son vía de acceso favorable a la prueba de hipótesis (ROSSMAN, 2008; BATANERO & DÍAZ, 2015)
2	Los estudiantes realizarán cambios en el resultado experimental y repetirán la prueba, para identificar que éste afecta al cálculo del p-valor. Realizarán cambios en el tamaño de muestra y repetirán la prueba, para identificar que se altera la variabilidad de la distribución de probabilidad que modeliza la hipótesis nula. Identificarán que el resultado de la prueba puede alterarse al modificar alguno de estos elementos.	- Un acercamiento informal facilita identificar roles de elementos clave involucrados en la PH (ROSSMAN, 2008; BATANERO & DÍAZ, 2015)
3	Los estudiantes utilizarán los razonamientos desarrollados en las tareas previas para realizar una prueba de hipótesis en la que el modelo de la hipótesis nula difiera del modelo 50-50 o de no diferencia.	- Los estudiantes deben ser capaces, eventualmente, de investigar afirmaciones más allá del modelo 50-50 (ROSSMAN, 2008)

Tabla 2: Vista general de hipótesis y su relación con referencias educativas, que enmarcan la THA.

Fuente: Silvestre, et al. (2024).

En dicha propuesta se estructuraron tres tareas, en las que se introduce al estudiante, a los conceptos clave de las PH desde el enfoque informal. Primero, al estudiante se le presenta un problema, bajo un modelo de 50-50 (espacio muestral equiprobable entre dos posibles sucesos), donde el estudiante deberá razonar mediante su conocimiento previo si se puede rechazar o no una hipótesis dada. Posteriormente se le muestra, a través de la simulación por computadora (mediante Fathom), una DME, que le pueda servir de argumento para apoyar su inferencia (ver Figura 6), lo anterior guiado por el docente. Al final se le pide al estudiante resolver una PH similar, pero en un contexto diferente y por cuenta propia.

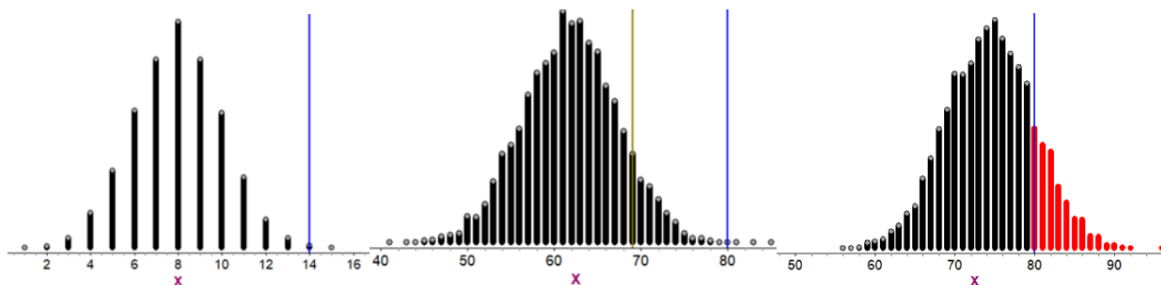


Figura 6: Distribuciones utilizadas para resolver tareas de la THA.

Fuente: Silvestre et al. (2024)

Si bien, este trabajo se centra en la enseñanza de las PH, utilizando el muestreo repetido mediante la simulación computacional, que además se implementó con estudiantes universitarios, éste aplica problemas de situaciones de modelo 50-50, es decir, con espacios muestrales teóricamente equiprobables. En general, las PH también se realizan en situaciones de espacios muestrales no equiprobables, por lo que aquí se identifica un área de oportunidad

y donde este trabajo pretende incidir. Cabe mencionar que tanto en García (2017) como en Silvestre, et al. (2024) se inicia la propuesta directamente con el uso del software para realizar el muestreo repetido mediante la simulación computacional. Para este trabajo, se considera como etapa preliminar el uso de simulaciones físicas, semejante a lo visto en los primeros trabajos revisados referentes a la construcción de una THA.

Por otra parte, se vislumbran dificultades muy serias en el momento de pasar a un modelo nulo diferente al de no diferencia, en este caso, provocó que los estudiantes retrocedieran a estrategias fallidas, como juzgar la hipótesis de manera determinista. Si se desea implementar una THA con situaciones como esta, o simplemente donde el espacio muestral fuese no equiprobable habría que realizar reajustes a la misma, quizás desde el contexto del problema.

Si nos centramos en las propuestas que involucran el diseño de una THA, algunas tienen como objetivo el aprendizaje de las PH, aunque otras no. Según Batanero y Díaz (2015), desde la matemática formal, la decisión de rechazar o no una hipótesis debe estar argumentada con base en la DMT; sin embargo, mediante un proceso de muestreo repetido (Batanero y Díaz, 2015; Garfield et al., 2015; Manor y Ben-Zvi, 2015; Watson y Chance, 2012; citados en van Dijke-Droogers et al., 2019) con apoyo de simuladores digitales, es posible construir una DME, aunque simulada (DMES), dado que puede estar muy cercana a la teórica, por lo que puede ser usada, para realizar la inferencia, dicho uso es lo que se considera como enfoque informal. Este tipo de constructo se considera clave para resolver una PH desde el enfoque informal.

En algunos trabajos se observa que la DME se construye mediante simulación por software, es decir, se elabora una DMES y otros construyen una DME mediante muestreo repetido práctico (o físico) a lo que Chance y Rossman (2006) llaman simulación práctica. Ciertamente, existe una gran diferencia entre éstas, en términos de la robustez o refinamiento de la DME construida; dicho de otro modo, una DME construida desde simulación física no podría tener suficientes datos recabados por su naturaleza, sin embargo, se prevé que iniciar con este tipo de simulación, para después pasar a la digital podría evitar el efecto de “caja negra”.

De estos trabajos revisados, en particular en van Dijke-Droogers et al. (2019), se implementaron ambos tipos de simulación. Aunque en este no se realizan PH, se observó que los estudiantes utilizaron la distribución de muestreo simulada como un modelo para interpretar la variación y la incertidumbre, además, pudieron simular fácilmente las distribuciones y usarlas para determinar los resultados más comunes. Por otro lado, los estudiantes tuvieron dificultades en la comprensión conceptual de una de las herramientas digitales utilizadas.

En el trabajo de García (2017), sí se centró en el aprendizaje las PH y se utilizó la simulación digital, mediante el software Fathom para realizar el proceso de muestreo repetido

y así construir la DMES. Sin embargo, se reporta que hubo confusión sobre la función de la simulación y la hipótesis nula. Muchos estudiantes no comprendieron que el software (Fathom) simulaba muestras de una población hipotética (basada en la hipótesis nula) y no de la población verdadera. Los estudiantes tendían a pensar que la simulación representaba la población real del problema. Al no entender que la simulación parte de un supuesto, los estudiantes a menudo no sabían qué valor poner en el parámetro del simulador; en varias ocasiones, utilizaron el valor del estadístico de la muestra como parámetro para la simulación, en lugar del valor de la hipótesis nula que se quería poner a prueba.

A su vez, aunque el objetivo era que los estudiantes midieran informalmente el valor- p , este fue un concepto difícil de asimilar; en lugar de contar la frecuencia de la muestra y los valores más extremos, los estudiantes contaban únicamente la frecuencia del valor exacto del estadístico de la muestra. De igual modo, muchos estudiantes creían que, si aplicaban el método correctamente, su conclusión era irrefutable y no consideraban la posibilidad de error.

En Silvestre, Armenta e Inzunza (2024), siendo un trabajo que se llevó a cabo con estudiantes de nivel superior, también se centró en el aprendizaje de las PH. Sin embargo, hubo dificultades que aún persistieron en los estudiantes al argumentar sus inferencias, por ejemplo, el uso de la DME para justificar el rechazo de la hipótesis nula fue valorado como incipiente; a pesar de que los estudiantes tenían disponible la DME, se les dificultó desprenderse de la perspectiva real, por lo que usar la distribución para justificar el rechazo fue un gran reto para la mayoría.

3.3 Aportaciones para la enseñanza de las PH desde el enfoque informal

Rossmann (2008) propone introducir la Inferencia Estadística de manera informal, especialmente para las PH, siguiendo la filosofía de Fisher. Sugiere un proceso de seis pasos:

1. Razonamiento intuitivo: partir de creencias iniciales, evaluar la improbabilidad de los datos si estas fueran ciertas y rechazarlas si la probabilidad es muy baja.
2. Simulación y aleatorización: usar pruebas de aleatorización simulada para relacionar el diseño del estudio con la inferencia y comprender el valor- p , siguiendo el enfoque de las “3R” (aleatorizar, repetir, rechazar).
3. Conexión diseño-conclusiones: enseñar que el alcance de las conclusiones depende del método de recolección de datos.
4. Ejemplos concretos: usar situaciones familiares y variables categóricas para facilitar el razonamiento lógico-probabilístico.

5. Conceptos clave: destacar el papel de la proporción muestral y el tamaño de muestra en la fuerza de la evidencia.
6. Enfoque de Fisher primero: priorizar la idea de fuerza de evidencia contra el modelo nulo antes que el enfoque de Neyman-Pearson centrado en decisiones y errores.

Para el sexto paso es de considerar que el proceso de razonamiento de la inferencia inductiva fisheriana se relaciona con un argumento de modus tollens en lógica, pero con un aspecto probabilístico añadido. Reconsiderando el ejemplo de los “datos 7/11” en estos términos, se razonó de la siguiente manera:

1. Una vez definida la hipótesis nula, suponerla cierta y preguntarse ¿sería muy sorprendente observar una muestra como la obtenida o más extrema?
2. Realizar el proceso de muestreo repetido a través de la simulación digital y observar lo que pudiera pasar de ser cierta la hipótesis nula.
3. Con suficiente evidencia, determinar si se rechaza o no la hipótesis nula.

Asimismo, en Batanero y Díaz (2015), para el nivel medio superior, se propuso una manera informal de acercarse a las PH, través de la simulación por computadora, generando muestras aleatorias de un mismo tamaño para observar lo que sucede con las medias muestrales, es decir, partir de lo empírico, obteniendo muchas muestras, las que quizás el estudiante pueda utilizar como argumento para inferir informalmente, cómo puede ser la población que originó las muestras observadas; pensando en estudiantes de nivel superior, inclusive proponer una valoración de qué tan seguros puedan estar ante la predicción realizada; en este caso, la simulación propuesta se observa en la figura 1.

Desde la filosofía de Fisher, para poder realizar una PH es necesario utilizar la distribución muestral teórica (DMT) y el cálculo del valor-p para poder tomar la decisión de rechazar o no una hipótesis determinada. Sin embargo, Batanero y Díaz (2015) proponen una manera de acercarse a dicha inferencia desde el enfoque informal. Para esto, es necesario llevar a cabo un proceso de muestreo repetido mediante un simulador digital; ellas argumentan que al realizar la simulación de suficientes experimentos se puede aproximar a la misma decisión sin realizar los cálculos formales.

Distributions of Sample Means

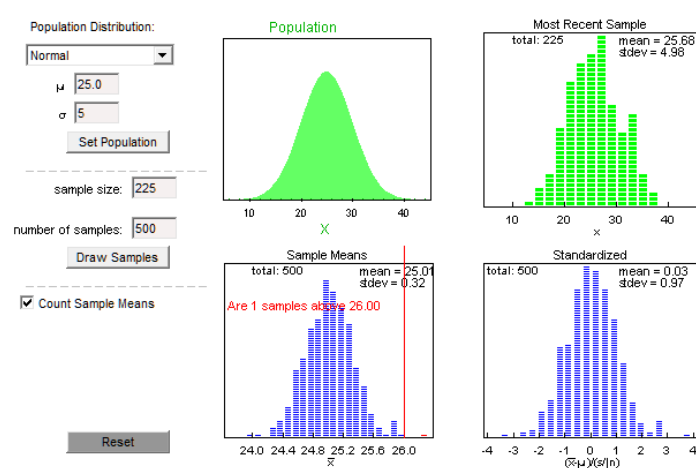


Figura 1: Simulación de la distribución muestral (500 muestras)

Fuente: Batanero y Díaz (2015)

Cabe aclarar que la intención de esto no es evitar los cálculos, sino en un momento introductorio, apoyar al estudiante en la comprensión de la lógica del concepto, y a su vez tratar de evitar los errores o las dificultades que surgen, sobre todo en la interpretación del resultado, para después pasar a la formalidad.

Aunque, tanto Rossman (2008) como Batanero y Díaz (2015), proponen el uso de la simulación digital como herramienta de apoyo para introducir al estudiante en el aprendizaje de las PH desde el enfoque informal, no muestran evidencia de experimentación, por lo que aún queda inconcluso cómo pudiera ser un diseño didáctico a partir de sus propuestas teóricas.

Case y Jacobbe (2018) realizaron una investigación al respecto de este tipo de ejercicio, para identificar dificultades que podrían enfrentar los estudiantes al momento de aprender inferencia estadística. Los autores explican que la PH a menudo se malinterpreta y que los métodos basados en la simulación tienen el potencial de hacer visible el pensamiento de los estudiantes sobre la inferencia.

Case y Jacobbe se centraron en dos desafíos principales: la lucha por coordinar el esquema multinivel y la dificultad para coordinar las dos perspectivas (mundo real y mundo hipotético) (ver Figura 2). Mencionan que, si bien las simulaciones ofrecen ventajas pedagógicas, los hallazgos demostraron que persisten dificultades, como la confusión entre el tamaño de la muestra y el número de ensayos simulados, y la confusión entre simulación y replicación.

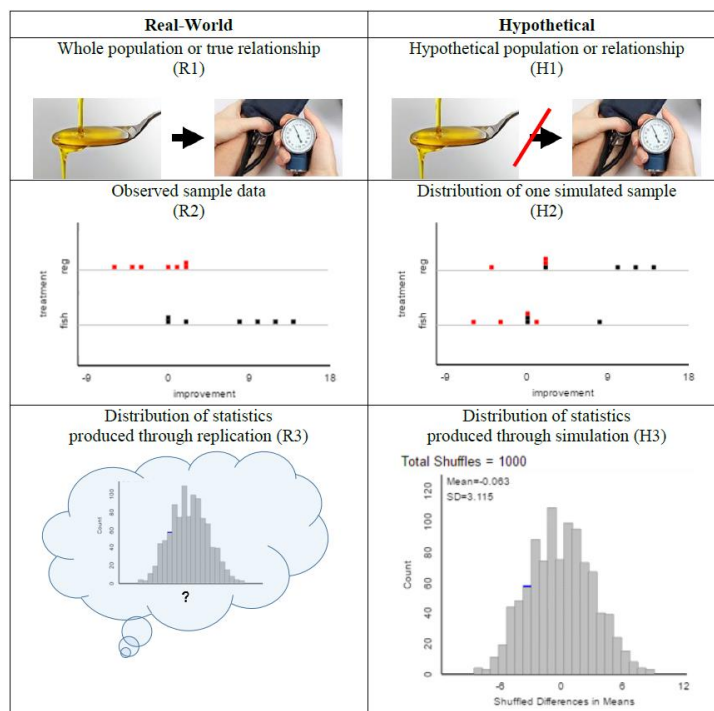


Figura 2: Mundo real y mundo hipotético.

Fuente: Case y Jacobbe (2018)

En términos generales, los autores afirman que el uso de métodos basados en simulación no es una solución perfecta que elimina todas las dificultades asociadas con el aprendizaje de la inferencia. El estudio también declara como inestable la capacidad de los estudiantes para coordinar ambos niveles y perspectivas, sin embargo, también ofrecen una evaluación optimista ya que los estudiantes a menudo resolvieron los problemas con herramientas como applets o con preguntas de seguimiento.

Tanto el trabajo de Case y Jacobbe (2018) como el de Silvestre et al. (2024) son particularmente relevantes para esta revisión. Por un lado, se realiza una propuesta teórica para la enseñanza de las PH desde el enfoque informal, tratando de reducir los errores y dificultades que surgen a partir del uso de simuladores y por otro, una aportación en cuanto al proceso del diseño didáctico para estudiantes del nivel superior.

3.4 Tecnología digital en la enseñanza desde el enfoque informal

En las últimas décadas, la incorporación de herramientas tecnológicas en la enseñanza de la estadística ha desempeñado un papel fundamental en el desarrollo del razonamiento del pensamiento y razonamiento estadísticos (Wild y Pfannkuch, 1999; Batanero, 2000; Garfield

y Ben-Zvi, 2009; Biehler, Ben-Zvi, Bakker y Makar, 2013). Algunos de estos recursos se han utilizado para el desarrollo de propuestas de enseñanza de la PH desde el enfoque informal.

Bajo un enfoque informal, se busca que los estudiantes construyan significados acerca de la variabilidad, el muestreo y la inferencia a partir de experiencias empíricas y exploratorias, en lugar de partir de procedimientos formales o teoremas probabilísticos. Para ello, el uso de simuladores se ha consolidado como un recurso didáctico clave, pues permite representar de manera dinámica procesos de muestreo repetido, visualizar la dispersión de las estadísticas muestrales y construir distribuciones muestrales que pueden ser utilizadas como modelos de inferencia empírica (Batanero y Díaz, 2015; García, 2017; van Dijke-Droogers et al., 2019; Silvestre et al., 2024).

Diversos estudios han evidenciado que este tipo de entornos promueve una comprensión más profunda de la aleatoriedad y la variabilidad (e.g. Inzunza, 2013). Rossman (2008) sostiene que la simulación computacional ofrece a los estudiantes la posibilidad de experimentar con el azar, observar patrones emergentes y reflexionar sobre la incertidumbre asociada a los resultados muestrales. De igual forma, Batanero y Díaz (2015) argumentan que la simulación constituye un puente conceptual entre el razonamiento informal y la inferencia formal, al permitir que el estudiante relacione la variabilidad observada en las muestras con la estabilidad relativa de las estadísticas.

Entre las herramientas tecnológicas empleadas en la literatura revisada que abordan este enfoque se encuentran Fathom, TinkerPlots, VUStat y los applets desarrollados por Rossman y Chance.

El software Fathom ha sido ampliamente utilizado por su capacidad para generar y representar visualmente distribuciones muestrales mediante simulaciones controladas, lo que facilita la exploración de conceptos como la media muestral, el error estándar o la variabilidad inter-muestral. Investigaciones como las de García (2017) y Silvestre et al. (2024) reportan el uso de Fathom en actividades que promueven la comprensión del proceso de muestreo repetido y la construcción empírica de modelos inferenciales.

Por su parte, TinkerPlots se ha empleado especialmente en niveles educativos básicos y medios, debido a su interfaz intuitiva y sus posibilidades de manipulación directa de datos y simulaciones (Garfield & Ben-Zvi, 2008; Pfannkuch et al., 2018). Este entorno permite a los estudiantes visualizar cómo la variabilidad surge naturalmente de los procesos aleatorios y cómo, mediante la repetición de muestras, se configura una distribución muestral representativa.

Asimismo, VUStat ha sido utilizado como una herramienta de transición hacia contextos más formales, al incorporar funcionalidades de análisis estadístico combinadas con

simulación, promoviendo la comprensión de la inferencia mediante actividades de estimación y PH en contextos auténticos. Finalmente, los applets de la colección de Rossman y Chance constituyen recursos accesibles y versátiles para la enseñanza de la estadística, especialmente en la ilustración de conceptos inferenciales básicos a través de simulaciones rápidas y visualmente comprensibles, integradas en cursos universitarios introductorios (Chance y Rossman, 2008).

En conjunto, estas herramientas digitales comparten el propósito de favorecer la comprensión de la inferencia estadística a través de la experiencia con la variabilidad y la simulación, permitiendo que los estudiantes participen activamente en la construcción de modelos empíricos. Su uso dentro del enfoque inferencial informal ha demostrado ser un medio eficaz para conectar la observación de fenómenos aleatorios con la formulación de conclusiones inferenciales, contribuyendo así al desarrollo de una comprensión más significativa y fundamentada del razonamiento estadístico.

El uso de la tecnología ha adquirido un papel estratégico en la enseñanza de la inferencia estadística y, particularmente, en la didáctica de las pruebas de hipótesis. Diversas investigaciones en educación estadística han señalado que las herramientas tecnológicas no solo facilitan el cálculo, sino que transforman de manera sustantiva los procesos de razonamiento que los estudiantes desarrollan al enfrentarse con situaciones de incertidumbre. Como afirman Garfield y Ben-Zvi (2008), “la tecnología no es únicamente un medio para realizar procedimientos, sino un elemento que posibilita nuevas formas de pensar estadísticamente”. Esta idea se vuelve especialmente relevante en el estudio de las pruebas de hipótesis, un contenido que históricamente ha sido enseñado desde una perspectiva formalista y procedimental.

En este campo, el uso de simuladores y entornos digitales interactivos ha permitido que los estudiantes exploren la variabilidad muestral de manera dinámica y comprensiva. Konold y Miller (2015) sostienen que la simulación computacional “acerca al estudiante a la lógica inferencial antes de introducir los procedimientos algebraicos”, lo que facilita una comprensión más profunda del razonamiento subyacente al valor-p, la hipótesis nula y la lógica de la aleatorización. Desde el enfoque de la inferencia estadística informal, esta aproximación resulta particularmente potente, ya que permite construir significado a partir de experimentos virtuales de muestreo repetido, apoyando preguntas del tipo “¿qué tan raro sería observar este resultado solo por azar?”.

En el contexto particular de las pruebas de hipótesis, las tecnologías digitales permiten visualizar distribuciones teóricas y empíricas, manipular parámetros, cambiar tamaños de muestra y contrastar escenarios alternativos en tiempo real. Wild y Pfannkuch (1999) destacan que las representaciones dinámicas favorecen la construcción de un “razonamiento transnumeral”, es decir, la habilidad de interpretar, transformar y conectar

diferentes formas de representación estadística. Esta capacidad se vuelve especialmente relevante cuando el estudiantado transita desde aproximaciones intuitivas, basadas en simulación, hacia métodos más formales sustentados en la distribución normal, la distribución t o modelos de proporciones, lo cual exige comprender la relación entre datos reales, modelos y supuestos.

Asimismo, el uso de applets, plataformas estadísticas y lenguajes como R o Python —en versiones diseñadas para la educación— ofrece entornos que fomentan la experimentación y la formulación de conjeturas. Garfield y Ben-Zvi (2008) subrayan que estas herramientas permiten que los estudiantes “manipulen datos, visualicen procesos y evalúen la variabilidad de resultados bajo diferentes condiciones”, lo que potencia la autonomía cognitiva y el desarrollo de una comprensión más profunda de los conceptos inferenciales.

En lo relativo a la enseñanza de las PH con un enfoque formal, la tecnología también desempeña un papel crucial como mediadora entre la teoría y la práctica. Por ejemplo, el uso de simuladores facilita la transición entre la variabilidad observada en muestras obtenidas de forma virtual y las propiedades matemáticas de las distribuciones teóricas, permitiendo construir conexiones explícitas entre la aleatorización empírica y los resultados que Ronald Fisher formalizó en su propuesta de la prueba exacta y la noción de significación estadística. En palabras de Fisher (1956), “la lógica de la significación depende de comparar lo observado con lo que podría esperar el azar bajo un supuesto específico”; esta comparación, que en el método clásico se fundamenta en distribuciones teóricas, puede ser introducida de forma más accesible a través de simulaciones que permitan visualizar dicha lógica sin recurrir inicialmente al formalismo matemático.

No obstante, el empleo de estas herramientas también plantea retos y limitaciones que deben considerarse en el diseño didáctico. En primer lugar, el uso de entornos visuales no garantiza por sí mismo la comprensión conceptual, ya que los estudiantes pueden centrarse en la manipulación de los simuladores sin establecer conexiones entre las representaciones y los significados estadísticos subyacentes (Batanero & Díaz, 2015; Pratt et al., 2020).

Case y Jacobbe (2018) y Silvestre et al. (2024) apoyan esta idea y pueden generar nuevos retos y dificultades. Por ejemplo, causar confusión entre el tamaño de la muestra y el número de pruebas, es decir, razonan que un gran número de ensayos simulados hace que el método sea más fiable que el enfoque teórico cuando el tamaño de la muestra original es pequeño. También pueden causar confusión en las unidades de observación ya que, a menudo, los estudiantes tienen dificultades para distinguir la distribución de la muestra de la distribución muestral simulada.

Uno de los desafíos que surgen es la falta de coordinación de perspectivas (mundo real contra hipotético). Los estudiantes luchan por coordinar la perspectiva del mundo real, donde se recogieron los datos de la muestra, y la perspectiva hipotética, donde se asume que la hipótesis nula es verdadera. Asimismo, fallo al reconocer el rol de la hipótesis nula, dicho de otro modo, los estudiantes pueden fallar en reconocer el papel de la hipótesis nula en el proceso de simulación o el propósito de la misma. Hasta creer que el profesor manipula el simulador a conveniencia, alterando incluso el estadístico muestral.

En segundo lugar, la rápida evolución del software educativo plantea el desafío de mantener la vigencia y compatibilidad de las herramientas, así como de integrar su uso en contextos curriculares diversos (por ejemplo, Fathom). Los creadores de Fathom han decidido discontinuar el proyecto y concentrarse en la construcción de otra herramienta de software que pueda cumplir con las demandas y necesidades actuales como aprovechar los servicios web, lo que ha dado como resultado el producto CODAP. Aún con estas limitaciones, las investigaciones coinciden en que las tecnologías interactivas, cuando se emplean con una orientación didáctica adecuada, constituyen un componente esencial para la enseñanza y aprendizaje de la inferencia estadística desde una perspectiva informal.

3.5 Reflexiones y aspectos clave de las propuestas revisadas.

Si bien, las THA revisadas que tienen por objetivo la enseñanza de la PH, siguen pautas educativas de investigaciones recientes, algunas no siguen de manera completa los fundamentos teóricos, o bien, no logran cubrir todos los aspectos que subyacen las PH. Por ejemplo, los trabajos que procuran apegarse a lo sugerido en Batanero y Díaz (2015), si bien concuerdan con la lógica planteada, no se hace uso de la misma herramienta tecnológica, lo que podría estar provocando rupturas en el proceso de aprendizaje.

En estos trabajos se utilizan contextos diversos, desde situaciones sociales, médicas, situaciones cotidianas, o ejemplos de experimentos realizados previamente que se pueden abordar para la enseñanza de conceptos de Inferencia Estadística, sin embargo, esto también juega un papel muy importante puesto que se ha mencionado que, al menos de inicio, se utilicen contextos familiares a los estudiantes dado que su conocimiento previo puede abonar a mejorar el entendimiento de las PH (Makar y Rubin, 2009; van Dijke-Droogers et al., 2017).

Por ejemplo, aunque en Silvestre, Armenta e Inzunza (2024) se trabajó con estudiantes de nivel superior y estudiantes de áreas económicas y administrativas, los contextos utilizados no eran a fin a dicha área, esto también podría provocar desinterés por parte de los estudiantes.

Por otro lado, si bien, varios trabajos se enfocaron en fundamentos de la inferencia estadística, y tocando elementos de la lógica de las PH, estos se trabajaron con estudiantes de niveles previos al universitario, lo que da para considerar que el estudiante de nivel universitario podría tener mayor conocimiento contextual que se pueda utilizar a favor para desarrollar su razonamiento estadístico informal para introducir la lógica de las PH.

A su vez, se identifica que tales propuestas no han sido implementadas en momentos posteriores a haber atendido el tema de Distribuciones Muestrales, algo que para este proyecto se considera clave. Otro aspecto clave y relacionado a este es haber tenido experiencia previa con el muestreo repetido y la distribución muestral empírica (DME) como antecedente a la PH desde lo informal.

Como se ha estado mencionando, se han reportado avances significativos en el diseño de THA orientadas a la didáctica de las PH. Sin embargo, aún persisten aspectos esenciales que no han sido suficientemente explorados y que resultan fundamentales para fortalecer la comprensión inferencial de los estudiantes. Entre estos aspectos, destaca la necesidad de proponer actividades que permitan examinar de manera explícita cómo varía el comportamiento de los estadísticos de prueba bajo diferentes tamaños de muestra, incorporando análisis contrafactuales del tipo “qué habría ocurrido si el tamaño muestral hubiera sido mayor o menor”. Este tipo de razonamiento es particularmente relevante tanto en el caso de proporciones, como —y sobre todo— en el estudio de la media, donde la distribución normal y sus aproximaciones juegan un papel central en la construcción de la distribución muestral y en la interpretación del valor-p.

Asimismo, se vuelve indispensable integrar en la trayectoria actividades que favorezcan una reflexión informada sobre el nivel de significancia, de modo que el estudiante comprenda su naturaleza subjetiva y las implicaciones que conlleva su elección en el proceso de toma de decisiones estadísticas. Esta comprensión situada permitiría que los estudiantes no solo apliquen mecánicamente un umbral convencional, sino que desarrollen criterios más conscientes y críticos sobre su selección en función de un problema concreto.

En conjunto, estas consideraciones apuntan a la necesidad de diseñar una nueva trayectoria hipotética de aprendizaje que articule de manera coherente el análisis del tamaño de muestra, el trabajo explícito con proporciones y, de manera prioritaria, con la media bajo el marco de la distribución normal, así como una reflexión fundamentada sobre el nivel de significancia. Una trayectoria con estas características contribuiría a ampliar las posibilidades didácticas en el estudio de las pruebas de hipótesis desde un enfoque de inferencia estadística informal, permitiendo que los estudiantes desarrollen una comprensión más robusta, flexible y conceptualmente significativa.

4 Problemática y Objetivos del Proyecto

4.1 Problemática

A lo largo de mi práctica docente en cursos de estadística, he constatado que la enseñanza de las PH representa uno de los desafíos más complejos dentro de la didáctica de la inferencia estadística. Con el paso del tiempo, me he convencido de que llevar a cabo una transposición didáctica eficaz de este contenido exige no solo un profundo dominio conceptual, sino también estrategias pedagógicas que permitan a los estudiantes aproximarse a la construcción de conocimiento a partir de sus propios razonamientos. Sin embargo, esta aspiración se ve frecuentemente limitada por la naturaleza abstracta y multifacética del tema (Batanero, 2001; Garfield y Ben-Zvi, 2008).

En particular, he observado que cuando los estudiantes no comprenden de manera sólida los fundamentos de la inferencia estadística (fundamentos que, por sí mismos, son numerosos y conceptualmente exigentes) se vuelve prácticamente imposible que logren interiorizar de forma significativa las herramientas propias de la Estadística Inferencial, entre ellas la PH. La comprensión de ideas como variabilidad muestral, comportamiento de las estadísticas bajo muestreo repetido, distribución de probabilidad o nivel de significancia constituye un requisito indispensable para avanzar hacia un entendimiento genuino del proceso de contraste estadístico (delMas, Garfield y Chance, 2007; Zieffler et al., 2008). Cuando estos elementos no están consolidados, la PH termina reduciéndose a un conjunto de pasos mecanizados que los estudiantes ejecutan sin interpretar adecuadamente sus implicaciones, fenómeno ampliamente documentado en la literatura (Chance, delMas, y Garfield, 2004; Konold y Kazak, 2008).

A ello se suma la persistencia de un pensamiento predominantemente determinista en los estudiantes, el cual parece derivar de su experiencia previa con las matemáticas escolares, generalmente orientadas hacia procedimientos exactos y resultados únicos. Este marco cognitivo contrasta de manera notable con la noción de aleatoriedad y variabilidad inherente a la estadística, y dificulta la transición hacia formas de razonamiento más probabilistas (Wilensky, 1997; Pratt y Ainley, 2008). Como consecuencia, los estudiantes suelen recurrir a la aplicación literal de fórmulas o algoritmos sin problematizar los supuestos, condiciones o significados que subyacen a cada decisión inferencial, situación descrita por Lovett (2001) como un aprendizaje basado en “procedimientos sin comprensión”.

En conjunto, estas observaciones ponen de manifiesto la necesidad de repensar la didáctica de la PH desde enfoques que permitan trabajar de manera explícita la variabilidad, promover la comprensión gradual de los conceptos fundamentales de la inferencia y favorecer la adopción de un pensamiento no determinista. Diversos autores han señalado que aproximaciones fundamentadas en la inferencia informal, el uso de simulación y la

experiencia con muestreo repetido pueden favorecer estos procesos (Pfannkuch, 2005; Rossman y Chance, 2014; Makar y Rubin, 2009). Ésta es una vía que se menciona como fuertemente provechosa para que los estudiantes desarrollen interpretaciones más profundas y argumentadas de los procedimientos inferenciales, dejando atrás el uso mecánico de recetas y fortaleciendo su capacidad para otorgar sentido a los resultados estadísticos.

En el ámbito de la educación estadística a nivel superior, la enseñanza y el aprendizaje de los contenidos relacionados con la inferencia estadística continúan representando un desafío tanto para docentes como para estudiantes. En particular, el aprendizaje de las PH suele abordarse desde un enfoque predominantemente formal y procedimental, centrado en la aplicación de fórmulas y en la interpretación mecánica de resultados. Esta tendencia limita la comprensión conceptual de los procesos inferenciales y reduce las oportunidades de desarrollar un razonamiento estadístico adecuado.

Como ejemplo, en los programas de las asignaturas de Estadística de la Universidad de Sonora, como el Programa de Estadística en los Negocios de la Licenciatura en Contaduría Pública (UNISON, 2018) y el Programa de Estadística Inferencial de la Licenciatura en Finanzas (UNISON, 2022), o del Tecnológico Nacional de México, como el Programa de Estadística para Administración 2 (TecNM, 2016), se sugiere la implementación de situaciones problema contextualizadas, el uso de hojas de trabajo, la discusión de ideas y la incorporación de herramientas tecnológicas como software estadístico; sin embargo, en la bibliografía y materiales sugeridos no se especifica el uso de dichas herramientas con un propósito didáctico ni se ofrecen orientaciones sobre cómo integrarlas pedagógicamente. Entonces, la enseñanza de la PH queda relegada al conocimiento y pericia del docente hacia el diseño didáctico.

Dado que el programa brinda escasos lineamientos para la enseñanza de las PH y otros conceptos de la inferencia, resulta pertinente revisar la bibliografía sugerida en éste, así como otros textos que puedan aportar ideas para el diseño instruccional. Este análisis resulta inclusive necesario dado que la enseñanza de las PH en los cursos de estadística ha sido objeto de análisis crítico debido a la persistencia de enfoques tradicionales que privilegian la mecanización de procedimientos sobre la comprensión conceptual de acuerdo con lo que se muestra en dicha bibliografía. Dentro de la bibliografía recomendada del curso (UNISON, 2018; UNISON, 2022) se mencionan libros como los siguientes:

- a) Estadística para administración y economía (Anderson, Sweeney y Williams, 2008).
- b) Estadística aplicada a los negocios (Webster, 2000).
- c) Fundamentos de estadística para administración (Wonnacott y Wonnacott, 1989).

En dichos textos se incluye el tema de las PH y se sugieren problemas y soluciones para estos, como una guía para que el profesorado los retome en sus prácticas docentes. La propuesta de resolución que se sugiere en éstos sigue la formalidad matemática y mediante el uso de fórmulas o recetas bien definidas los que parece provocar la eliminación de la incertidumbre que suele causar la variabilidad. A continuación, se presentan algunos ejemplos de problemas propuestos (ver figuras 3, 4 y 5).

Aplicaciones

Autoexamen

15. Las declaraciones de impuestos presentadas antes del 31 de marzo obtienen un reembolso que en promedio es de \$1056. Considere la población de los declarantes de "última hora" que presentan su declaración los últimos cinco días del periodo para este trámite (normalmente del 10 al 15 de abril).
 - a. Un investigador sugiere que la razón por la que estos declarantes esperan hasta los últimos días se debe a que en promedio obtienen un reembolso menor que los que declaran antes del 31 de marzo. Dé las hipótesis apropiadas de manera que el rechazo de H_0 favorezca la sugerencia de este investigador.
 - b. En una muestra de 400 personas que presentaron su declaración entre el 10 y el 15 de abril, la media de los reembolsos fue \$910. Por experiencia se sabe que es posible considerar que la desviación estándar poblacional es $\sigma = \$1600$. ¿Cuál es el valor- p ?
 - c. Con $\alpha = 0.05$, ¿cuál es su conclusión?
 - d. Repita la prueba de hipótesis anterior usando el método del valor crítico.

Figura 3: Un problema de prueba de hipótesis para la media solicitado la realización de cálculos.

Fuente: Anderson, Sweeney y Williams (2008, p. 358).

10. Como gerente de compras para una gran empresa de seguros usted debe decidir si actualizar o no los computadores de la oficina. A usted se le ha dicho que el costo promedio de los computadores es de US\$2,100. Una muestra de 64 minoristas revela un precio promedio de US\$2,251, con una desviación estándar de US\$812. ¿A un nivel de significancia del 5% parece que su información es correcta?

Figura 4: Un problema de prueba de hipótesis.

Fuente: Webster (2000).

9-9 Los registros informan que en una muestra aleatoria de 100 horas una máquina produjo un promedio horario de 678 artículos con una desviación estándar de 25. Después de instalar un dispositivo de control, la máquina produjo un promedio de 674 artículos por hora, con una desviación estándar de 5, en una muestra aleatoria de 500 horas. Al señalar la disminución de 4 artículos por hora en la media muestral, el gerente sostuvo que el dispositivo de control reducía la producción. Por el contrario, el gremio afirmó que esa disminución era "mera fluctuación estadística".
Para resumir objetivamente la evidencia acerca del cambio verdadero o no de la producción, calcule el valor de probabilidad.

Figura 5: Un problema de prueba de hipótesis.

Fuente: Wonnacott y Wonnacott (1989).

Como se puede observar, ningún problema solicita o sugiere el uso de software como recurso de apoyo tecnológico, sino que deben resolverse usando la teoría formal de la probabilidad (teoría de variables aleatorias). Por ejemplo:

La propuesta de la fuente a) para resolver PH consiste en 5 pasos (ver tabla 3):

Paso	Descripción	
1	Establecer las hipótesis nula y alternativa (H_0 y H_1).	
2	Especificar el nivel de significancia.	
3	Recabar los datos muestrales y calcular el valor del estadístico de prueba (valor de Z) mediante la fórmula de estandarización $\left(Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \right).$	
Método de valor-p	4	Ubicar el valor de Z en la tabla de la distribución normal estándar y encontrar el valor-p.
	5	Rechazar H_0 si el valor-p $\leq \alpha$.
Método de valor crítico	4	Utilizar la tabla de la distribución normal estándar (al revés) para buscar el valor de α como área bajo la curva y encontrar el valor de Z correspondiente y así determinar el valor crítico de las zonas de rechazo y de aceptación.
	5	Identificar la zona en la que se encuentra el estadístico de prueba para determinar si se rechaza H_0 .

Tabla 3: Pasos para resolver pruebas de hipótesis según Anderson et al. (2008, p. 354)

Para resolver el problema de la figura 3, según Anderson et al. (2008), se inicia mediante el establecimiento de las hipótesis nula y alternativa (H_0 y H_1) (siendo lo que se pide en el primer inciso):

$$H_0: \mu \geq \$1,056$$

$$H_1: \mu < \$1,056$$

El segundo inciso solicita el cálculo del valor-p, en vez de especificar el nivel de significancia, particularidad que no coincide con los pasos sugeridos y que se repite con otros problemas propuestos en la misma fuente. De cualquier modo, para calcular el valor-p se debe calcular el estadístico de prueba como en el paso 3:

$$Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{910 - 1056}{\frac{1600}{\sqrt{400}}} = -1.825$$

Por lo que

$$\text{valor} - p = 0.034$$

Después del cálculo del valor-p seguimos con el tercer inciso que solicita concluir con respecto a H_0 con un nivel de significancia de 0.05. Aquí se esperaría que el estudiante sólo concluya si se rechaza o no H_0 , que en este caso sí sería rechazada, pero podría no añadir la argumentación correspondiente, lo que devendría en una interpretación con un tono muy determinista. El cuarto inciso sólo solicita volver a realizar la prueba, pero con el método del valor crítico.

Las fuentes b) y c) siguen los mismos pasos como propuesta para resolver las PH.

También coinciden en la característica de promover lo procedimental antes que lo conceptual. No obstante, se identifica que los libros de texto utilizados en la formación estadística no han incorporado los avances propuestos en la investigación con respecto a dicho tema, lo que puede limitar el desarrollo del RIEI en los estudiantes. Entre los tipos de problemas que se identifican en éstos, muchos suelen ser de carácter procedimental, además de no promover el uso de la tecnología computacional ni hacer uso de su RIEI, es decir, aquellos que sólo solicitan al estudiante el uso de fórmulas o recetas mecanizadas para resolverlos.

Por otro lado, tampoco se ven indicios del uso del RIEI de los estudiantes. Estos se enfocan mayormente en solicitar la realización de cálculos y como en rara ocasión se solicita argumentar las conclusiones emitidas (si es que se solicita concluir) pocos estudiantes propondrían una interpretación del resultado obtenido acompañado de su argumentación correspondiente en términos del contexto del problema.

Asimismo, se ha identificado la presencia de un enfoque epistemológico híbrido en la enseñanza de la PH, combinando de manera inconsistente los enfoques de Fisher y Neyman-Pearson. Por ejemplo, en la figura 4, a pesar de que sólo se realizará la prueba una vez (enfoque de Fisher), se establece un nivel de significancia (enfoque de Neyman y Pearson). Esta fusión, puede generar confusión en los estudiantes y dificultar su aplicación correcta en contextos reales.

Después de evaluar las condiciones de dicha bibliografía, en la siguiente tabla (ver tabla 4) identificamos los siguientes porcentajes que corresponden a la aparición de problemas con las características antes mencionadas:

Característica	Evaluación		
	Anderson et al. (2008) (45 problemas)	Webster (2000) (62 problemas)	Wonnacott y Wonnacott (1989) (5 problemas)
Carácter procedimental (solicitar cálculos como acción principal)	82.2 %	95.2%	80%
Enfoque híbrido	71.1%	96.8%	60%

Tabla 4: Evaluación de la identificación de algunas características en las fuentes sugeridas en los programas de estudios.

Uno de los principales problemas identificados es la predominancia de problemas procedimentales sobre problemas conceptuales. Los primeros conducen a la aplicación mecánica de algoritmos sin una comprensión profunda, mientras que los segundos fomentan la conexión de conocimientos y la adaptabilidad a nuevas situaciones. La carencia de problemas conceptuales de PH puede restringir la capacidad de los estudiantes para desarrollar su razonamiento estadístico en el tema.

Otra limitación relevante es la escasa o nula integración del RIEI, siendo que ha demostrado ser una herramienta metacognitiva potencialmente útil para facilitar la transición hacia enfoques formales de inferencia. A pesar de su potencial, las fuentes bibliográficas analizadas no promueven su uso. De hecho, es poco común encontrar textos que promuevan dicho acercamiento a la PH, por ejemplo, en Watkins, Scheaffer y Cobb (2008) aparece un apartado especial para Pruebas de Significancia Informal (ver figura 6).

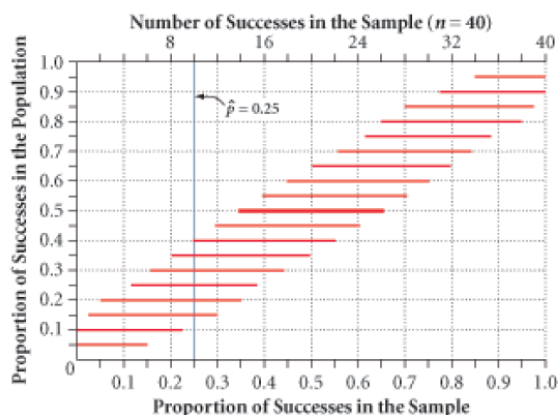
Example: Miguel and Kevin's Spins

Miguel and Kevin spin pennies and get 10 heads out of 40 spins for a sample proportion, \hat{p} , of 0.25. Is this a statistically significant result?

Solution

As shown in Display 8.8, if 0.5 is the true proportion of heads when a penny is spun, it would be a rare event to get only 10 heads in a sample of 40 spins. Miguel and Kevin have a statistically significant result. This leads them to conclude that 0.5 probably is not the true proportion of heads when a penny is spun.

Another way of seeing this is to note that $p = 0.5$ isn't included in the 95% confidence interval for the true proportion of heads when a penny is spun. Look at the vertical line at $\hat{p} = 0.25$ in Display 8.9. It doesn't intersect the horizontal line that represents the population with $\hat{p} = 0.5$. That is, the sample proportion $\hat{p} = 0.25$ is far enough from the standard of 0.5 to reject 0.5 as a plausible value for the proportion of times heads appear when a penny is spun.



Display 8.9 A sample proportion of $\hat{p} = 0.25$ isn't reasonably likely when $p = 0.5$.

Figura 6: Prueba de significancia informal para la proporción.

Fuente: Watkins et al. (2008).

Además, el papel de la tecnología digital en la enseñanza de la PH sigue siendo relegado. A pesar de que en los programas sí se sugiere, los textos sugeridos en éstos no apoyan tal sugerencia, es decir, existe una contradicción. La simulación computacional, que permite generar distribuciones muestrales y fortalecer la comprensión de su lógica, no es promovida en los textos revisados. Integrar estas herramientas podría mejorar significativamente el aprendizaje y la aplicación de la estadística en contextos reales.

En síntesis, los libros de texto recomendados abordan las PH mediante una hibridación entre los enfoques de Fisher y de Neyman-Pearson, solicitando al estudiante establecer las hipótesis nula y alternativa y comparar el nivel de significancia con el valor p o los valores críticos correspondientes. De acuerdo con Batanero y Díaz (2015), este enfoque

contribuye a malas interpretaciones frecuentes, tales como asumir que el valor-p representa la probabilidad de que H_0 sea cierta, o que un resultado no significativo implica la inexistencia de un efecto.

De esta manera, la transposición didáctica sugerida en los programas y materiales tradicionales no refleja los fundamentos epistemológicos de las PH. Aunque diversas investigaciones en educación estadística (Zieffler et al., 2008; Garfield y Ben-Zvi, 2008; Pfannkuch y Budgett, 2014; Pfannkuch et al., 2014; Batanero y Díaz, 2015; Silvestre et al., 2024, entre otros) han mostrado los beneficios del enfoque de la IEI como herramienta cognitiva para introducir conceptos inferenciales, no se identifican indicios de su inclusión explícita en los contenidos curriculares o en los materiales institucionales existentes.

Ante ello, la perspectiva de la IEI ha emergido como una alternativa pedagógica valiosa, al promover el razonamiento estadístico a partir de la exploración de datos, la simulación y la interpretación de la variabilidad muestral. No obstante, a pesar del creciente reconocimiento de este enfoque en la literatura especializada, se observa una escasez de materiales didácticos y diseños de intervención que orienten su implementación en la enseñanza de las PH. La mayoría de los recursos disponibles se concentran en aspectos introductorios de la inferencia estadística, como la comprensión de la variabilidad y de los procesos de muestreo, sin llegar a proponer actividades integrales que conduzcan a la formulación, contrastación e interpretación de hipótesis desde una perspectiva informal. Esta carencia de propuestas concretas conlleva a que los docentes continúen utilizando materiales tradicionales, poco coherentes con los principios del enfoque informal.

En este contexto, resulta necesario diseñar y validar propuestas didácticas innovadoras que permitan introducir las PH desde la IEI, mediante estrategias de simulación a través de herramientas de software, que favorezcan la comprensión progresiva de la lógica involucrada en el concepto.

Una vía de interés que se ha encontrado en la literatura para el diseño de este tipo de propuestas es el uso de las THA. Éstas se han empleado como medio para planificar secuencias didácticas que promuevan el razonamiento estadístico desde el enfoque IEI (García, 2017; van Dijke-Droogers et al., 2019; Nilsson, 2023; Silvestre et al., 2024), Aunque pocas de estas trayectorias se han centrado en el aprendizaje de las PH.

Esta situación abre la posibilidad de diseñar una THA que integre los principios de la IEI hacia el aprendizaje de las PH, articulando tanto actividades exploratorias como herramientas tecnológicas de simulación, que guíen a los estudiantes en la construcción progresiva de la inferencia estadística. Sin embargo, dado que aún hay poco material de referencia en cuanto al diseño didáctico, además de que dichas propuestas no logran atender todos los aspectos de la PH, de aquí surge la siguiente cuestión:

¿Cómo diseñar una propuesta didáctica para el aprendizaje de las PH, de acuerdo con el enfoque de la IEI y mediante una THA, de modo que mejore la interpretación de los resultados por parte de los estudiantes?

Otra área de oportunidad identificada en la literatura revisada radica en la necesidad de profundizar en la comprensión del proceso general de diseño de las trayectorias hipotéticas de aprendizaje (THA) y, de manera complementaria, en cómo los estudiantes se desenvuelven dentro de una determinada secuencia didáctica. Si bien las THA han demostrado ser un recurso valioso para anticipar y orientar el aprendizaje, persiste una limitada exploración sistemática sobre los fundamentos teóricos y metodológicos que guían su construcción y validación.

Al respecto, diversos estudios han incorporado marcos conceptuales particulares para sustentar el diseño de sus propuestas. Por ejemplo, García (2017) recurrió a la taxonomía SOLO para caracterizar los niveles de complejidad cognitiva en el aprendizaje y utilizarlos como base para la elaboración de su trayectoria. Por su parte, Case y Jacobbe (2018) estructuraron su diseño a partir de las perspectivas real e hipotética, diferenciando entre cómo se espera que evolucione el pensamiento estudiantil y cómo se manifiesta en la práctica. De manera distinta, van Dijke, Droogers y Bakker (2019) fundamentaron su propuesta en la teoría de la objetivación, centrada en los procesos socioculturales mediante los cuales los estudiantes construyen significados matemáticos.

La diversidad de enfoques evidencia, no solo la riqueza conceptual que sustenta el diseño de THA, sino también la importancia de analizar por qué una trayectoria específica podría funcionar o no en contextos determinados. Esta reflexión resulta crucial, pues permite identificar las condiciones que favorecen u obstaculizan el aprendizaje, así como realizar los reajustes necesarios para optimizar el diseño. En consecuencia, el interés por comprender estos mecanismos constituye un componente indispensable para fortalecer la propuesta desarrollada en este trabajo y avanzar hacia procesos de enseñanza más informados, coherentes y sensibles a la complejidad del aprendizaje de la PH.

Así pues, el interés no sólo radica en la faceta del diseño didáctico hacia la enseñanza de las PH mediante una THA, sino también se desprenden aspectos relacionados con la investigación en cuanto al proceso de aprendizaje que intenta promover dicho diseño y por qué funciona o no. Es decir, también se cuestiona:

¿Cómo evolucionan los argumentos a las interpretaciones y recursos estadísticos de los estudiantes conforme avanzan en la resolución de la THA hacia la enseñanza de las PH desde el enfoque informal?

En relación con lo anterior, se identifica la necesidad de elegir los Experimentos de Diseño como metodología para el presente proyecto de intervención, que abarca ambas componentes, tanto desarrollo de investigación como diseño didáctico en formato cíclico y, por tanto, permite conceptualizar este trabajo como un proyecto centrado en un problema de diseño.

4.2 Objetivo general y específicos

Después de la revisión literaria realizada, se observa que existe interés en la construcción de THA, enmarcadas en experimentos de diseño, para el diseño y elaboración de secuencias de actividades didácticas enfocadas en el aprendizaje de la Estadística Inferencial y desde el enfoque IEI; pero, a pesar de que hay mucha teoría al respecto de dicho enfoque, aún existen muy pocas propuestas, y en específico para las PH, para estudiantes de nivel superior de las áreas económicas y administrativas.

Por lo que, para este trabajo, el objetivo general se establece como sigue:

Diseñar una secuencia de actividades didácticas, a través de la construcción de una THA y mediante un experimento de diseño, para la enseñanza de las PH desde el enfoque IEI, para estudiantes de nivel superior en el área de ciencias económicas y administrativas.

Este es un proyecto de intervención didáctica, sin embargo, al llevar a cabo un trabajo como este, siempre surgen elementos que pueden formar parte de la teoría referente a cómo elaborar un diseño instruccional y por qué funciona, por lo que, adicionalmente, también se considera la creación de un modelo de argumentación de respuestas con base en el Modelo de Argumentación de Toulmin para el razonamiento inferencial estadístico (RIE) (Gómez-Blancarte y Tobías-Lara, 2018), con el propósito de aprovechar y sistematizar los posibles aportes teóricos emergentes que puedan surgir durante la implementación del proyecto.

Asimismo, para lograr dicho objetivo, se proponen los siguientes objetivos específicos:

- Definir los componentes de una THA (las tareas, los objetivos y las hipótesis de aprendizaje) que aborden las PH desde un enfoque informal, para estudiantes universitarios en el área de las ciencias económicas y administrativas.
- Implementar la THA como un primer ciclo de diseño.
- Analizar respuestas y producciones de estudiantes, aciertos, errores y modos de argumentación siguiendo el Modelo de Argumentación de Toulmin.

- Utilizar el modelo de argumentación identificado en los estudiantes para rediseñar la THA, concretando un primer ciclo de diseño.
- Implementar el rediseño de la THA en un segundo ciclo y robustecer el modelo de argumentación identificado en los estudiantes.
- Utilizar el modelo de argumentación identificado en los estudiantes en el segundo ciclo para modificar y robustecer los componentes de la THA, generando una versión final de la misma.

5 Marco conceptual

Garfield y Ben-Zvi (2008), señalan que "la inferencia estadística no es simplemente un conjunto de procedimientos, sino un marco de razonamiento que vincula los datos muestrales con afirmaciones sobre poblaciones", característica que las PH heredan como una de las herramientas del campo disciplinar. Esta afirmación permite comprender que su naturaleza es fundamentalmente epistemológica, por lo que no sólo consiste en aplicar fórmulas, sino en construir conocimiento a partir de datos sujetos a variabilidad.

Para este proyecto se constituye un marco conceptual basado en la epistemología de la PH, en particular de la propuesta de Fisher, el uso de la IEI de acuerdo a lo sugerido por Batanero y Díaz (2015) y Rossman (2008) y las THA como vehículo para el diseño didáctico. Por otro lado, como herramienta para analizar las argumentaciones de los estudiantes con respecto a las PH realizadas, se utilizará el Modelo de Argumentación de Toulmin adaptado para el Razonamiento Inferencial Estadístico (RIE) por Gómez-Blancarte y Tobías-Lara (2018).

5.1 Prueba de significancia: la filosofía de Fisher.

Para Fisher, el objetivo de una prueba de significación es evaluar si los datos observados podían considerarse plausibles bajo un modelo asumido, representado por la hipótesis nula. Fisher (1935) afirmaba que "la prueba de significación evalúa la consistencia de los datos con la hipótesis nula, y un resultado improbable constituye evidencia para dudar de ella". Esta visión sitúa la PH como un proceso epistemológico de evaluación de evidencia, no como una decisión dicotómica absoluta.

La filosofía fisheriana rechaza la idea de aceptar H_0 . Bajo su marco conceptual, el propósito es medir la fuerza de la evidencia contra H_0 , lo cual se operacionaliza mediante el valor-p. Según Fisher, el valor-p debe interpretarse como una medida continua de evidencia, no como un umbral rígido. En sus palabras, el valor-p indica la "significación del desajuste entre el modelo y los datos".

La epistemología fisheriana resalta tres elementos clave que deben incorporarse en la enseñanza:

- El carácter evidencial del valor-p, entendido como indicador de rareza bajo un modelo.
- La naturaleza no dicotómica de la inferencia, pues Fisher no concebía el rechazo o aceptación absoluta.
- La importancia de la plausibilidad del modelo, evaluada mediante comparación observado–esperado.

Integrar explícitamente esta filosofía en la enseñanza de las PH, especialmente con experiencias basadas en simulación, eventualmente, podría evitar confusiones epistemológicas y fortalecer la comprensión conceptual del proceso inferencial (Rossman, 2008; Batanero y Díaz, 2015; García, 2017; Silvestre et al., 2024).

La comprensión adecuada de las PH requiere una conceptualización profunda de la variabilidad muestral, de la naturaleza probabilística de los datos y del significado de la extremidad. Sin estos elementos, el valor-p carece de sentido, y las decisiones estadísticas se perciben como criterios arbitrarios.

Esta filosofía se considera conveniente para introducir la enseñanza de las PH, pues no nos preocupamos de la hipótesis alternativa o el error tipo II (Batanero y Díaz, 2015).

5.2 Inferencia estadística informal en la prueba de hipótesis

5.2.1 Razonamiento inferencial estadístico informal

El enfoque de la IEI constituye una corriente didáctica relativamente reciente que propone introducir a los estudiantes en los principios de la inferencia antes de abordar formalmente tópicos como las PH. De acuerdo con autores como Zieffler et al. (2008), Rossman (2008) y Segundo (2018), este enfoque se basa en el desarrollo del razonamiento inferencial por parte de los estudiantes, quienes interpretan datos y extraen conclusiones sobre una población a partir de una muestra o valoran si un resultado es inusual, sin recurrir a procedimientos formales como el cálculo del valor-p, el uso de una distribución muestral teórica (DMT) como modelo o el nivel de significancia.

Por ejemplo, ante la situación: “*Un estudiante lanza una moneda 10 veces y obtiene 9 sellos*”, puede preguntarse si la moneda está cargada. Si decide realizar más ensayos para observar si el resultado inicial se repite, estaría utilizando su razonamiento inferencial informal (RII) para realizar una inferencia informal, sustentada en la evidencia empírica y no en el uso de fórmulas estadísticas. Y si además, se argumenta la inferencia mediante la distribución muestral empírica (DME), es decir, incluyendo en la argumentación un proto valor-p correspondiendo a la cantidad de muestras recabadas como la obtenida en el estudio original o más extrema, consideramos que estaría haciendo uso de un razonamiento más profundo, es decir, al que llamamos razonamiento inferencial estadístico informal (RIEI).

Si bien una limitación del enfoque informal radica en la cantidad de ensayos que pueden efectuarse físicamente, hoy en día existen herramientas de software (como Fathom, hojas de cálculo y applets interactivos, entre ellos *One Variable with Sampling* de la colección de Rossman y Chance) que permiten simular múltiples repeticiones del proceso de muestreo

y facilitan la toma de decisiones con base en la distribución muestral empírica simulada (DMES).

5.2.2 Aproximación informal a la prueba de hipótesis

La IEI constituye un enfoque didáctico y cognitivo que busca facilitar la comprensión profunda de la inferencia sin recurrir inicialmente a los formalismos matemáticos. Como afirman Batanero y Díaz (2015), "el razonamiento informal constituye una etapa necesaria en el desarrollo cognitivo del pensamiento inferencial", ya que precede a la formalización y permite construir significado antes de aplicar procedimientos.

Este enfoque se basa en la idea de que los estudiantes pueden elaborar conclusiones razonables sobre una población mediante la interpretación de datos, la observación de la variabilidad y la evaluación de patrones. De acuerdo con Garfield y Ben-Zvi (2008), "la inferencia informal se apoya en comparaciones, en la identificación de tendencias y en el uso de evidencia basada en datos sin necesidad de modelos matemáticos precisos".

Además, reconoce que la inferencia es un proceso de razonamiento que emerge de la interacción con los datos. La IEI invita a los estudiantes a construir modelos mentales sobre cómo se comportan los datos bajo diferentes supuestos, a identificar valores típicos y extremos y a justificar sus conclusiones.

Para adaptar estas sugerencias a un diseño didáctico se rescatan los siguientes elementos (ver tabla 5):

Elemento	Descripción
La hipótesis nula	Rossmann (2008) propone iniciar mediante el establecimiento de una hipótesis nula de no diferencia (50-50), para la que se buscaría evidencia con la intención de poderla rechazar o no.
Construcción de la DMES	Mediante el uso de la simulación para llevar a cabo el proceso de muestreo repetido, se ha de promover la construcción de una distribución muestral empírica simulada.
Obtener un proto valor-p	Identificar en la DMES la proporción de muestras simuladas con media igual o más extrema que la observada en la muestra del estudio. Le llamamos proto valor-p dado que no ha sido calculado mediante la formalidad matemática.
Toma de decisión	Se incitará al estudiante a tomar una decisión con respecto a H_0 mediante el proto valor-p; si este es muy pequeño se podría rechazar H_0 .

Tabla 5: Elementos seleccionados para el diseño de las actividades didácticas de la THA.

La simulación se ha convertido en una herramienta clave para la enseñanza de la inferencia y, por consiguiente, de las PH. Rossman (2008) señala que la simulación permite aproximar la DMES a la DMT mediante datos generados por computadora, proporcionando así una comprensión intuitiva del proceso inferencial. A su vez, podría transformar conceptos abstractos en experiencias concretas. Además, permite observar la variabilidad sin necesidad de fórmulas y facilita construir significado alrededor del valor-p.

Uno de los aportes más relevantes de Rossman es la reinterpretación del valor-p como una frecuencia relativa. Según Rossman (2008), "los estudiantes comprenden mejor el valor-p cuando lo observan emerger de una distribución simulada que cuando se les presenta como un número obtenido mediante una fórmula".

5.3 Trayectoria hipotética de aprendizaje

5.3.1 Definición y sus elementos

Las THA fueron introducidas por Simon (1995) y Simon y Tzur (2004) como parte de un enfoque constructivista para la enseñanza y el aprendizaje de las matemáticas. Esta idea surge como herramienta para planificar, anticipar y analizar el aprendizaje de los estudiantes en contextos de enseñanza. Es decir, sirve como apoyo para diseñar actividades didácticas desde su planeación hasta el análisis de resultados, esto a través de diversas etapas, en las que se describen hipótesis de aprendizaje que el investigador, a priori, considera que serán logradas en la implementación, y que después serán revisadas para finalmente realizar un análisis retrospectivo.

La THA "consiste en una meta de aprendizaje para los estudiantes, los procesos de aprendizaje subyacentes y los medios (tareas, actividades) utilizados para apoyar estos procesos y lograr las metas de aprendizaje (hipótesis sobre el proceso de aprendizaje)" (Nilsson, 2023, p. 4). La trayectoria es "hipotética" porque está basada en teorías del aprendizaje y en la experiencia del docente, pero no garantiza resultados fijos; debe ajustarse dinámicamente en función de la retroalimentación obtenida después del proceso de implementación en el aula.

De la revisión de literatura previa se han identificado algunos trabajos que intentan promover el uso de la IEI para realizar inferencia como proceso introductorio, en el que se elaboran THA (van Dijke-Droogers et al., 2019; Silvestre et al., 2024, entre otros). Si bien, estos trabajos han sido esfuerzos que promueven la inferencia desde el enfoque informal, no todos se centran en las PH, no todos están dirigidos a estudiantes universitarios y no todos utilizan la simulación computacional.

5.3.2 Implicaciones para la construcción de la THA

Diseñar una THA para la enseñanza de la PH mediante el enfoque IEI requiere integrar actividades que permitan observar variabilidad, generar DMES y comparar resultados. Las tareas deben fomentar la exploración, discusión y argumentación. Konold (2007) indica que "los datos deben ser el punto de partida para el razonamiento estadístico". Es decir, se sugiere que, a partir de una muestra con datos reales y en un contexto del área económica y administrativa, dichas tareas correspondan a solicitar a los estudiantes que realicen:

- Muestreo repetido mediante simulación.
- Construcción de una DMES bajo el supuesto H_0 como verdadero.
- Identificación de valores iguales o más extremos como el observado en la muestra.
- Argumentación basada en la evidencia.

En el capítulo 7 se describirán con mayor detalle las tareas y características específicas de la THA que se pretende elaborar.

5.4 Modelo de argumentación de Toulmin

Se utilizará el Modelo de Argumentación de Toulmin con el propósito de analizar las argumentaciones que los estudiantes expresarán en cada PH realizada. El modelo fue adaptado para el RIE por Gómez-Blancarte y Tobías-Lara (2018).

Éste es un modelo ampliamente utilizado para analizar y construir argumentos de manera clara y estructurada. Fue desarrollado por el filósofo británico Stephen Toulmin en su obra "The Uses of Argument", publicada en 1958. Gómez-Blancarte y Tobías-Lara (2018) adaptaron este modelo al contexto del RIE, con el objetivo de analizar cómo los estudiantes universitarios construyen argumentos al realizar PH. Esta adaptación busca integrar tanto las razones estadísticas como las contextuales que los estudiantes utilizan para sustentar y validar sus generalizaciones a partir de los datos muestrales

El modelo identifica seis componentes clave en la estructura de un argumento: afirmación, datos, garantía, respaldo, calificador modal y contrargumento. Para este proyecto, se muestran los componentes de dicho modelo y a qué se refiere cada uno, adaptado al RIEI (ver Tabla 6):

Componente	Relación con la IEI
Afirmación o Conclusión	Generalización más allá de los datos. (ejemplo: “la moneda está cargada”).
Datos	Información o evidencia que apoya a la afirmación. (ejemplo: “en 30 lanzamientos, se obtuvieron 23 sellos”).
Garantía	La conexión lógica entre los datos y la afirmación. (ejemplo: “si la moneda fuera justa, sería raro obtener 23 sellos”).
Respaldo	Apoyo adicional para la garantía. (ejemplo: “las simulaciones muestran que obtener 23 sellos o más, con una moneda justa, es muy poco probable”).
Calificador Modal	Indica el nivel de certeza de la afirmación. (ejemplo: “probablemente”, “es muy probable”, “es posible que”).
Contraargumento	Considera condiciones en las que la afirmación no se sostendría. (ejemplo: “a menos que se haya cometido un error en los registros de los lanzamientos”).

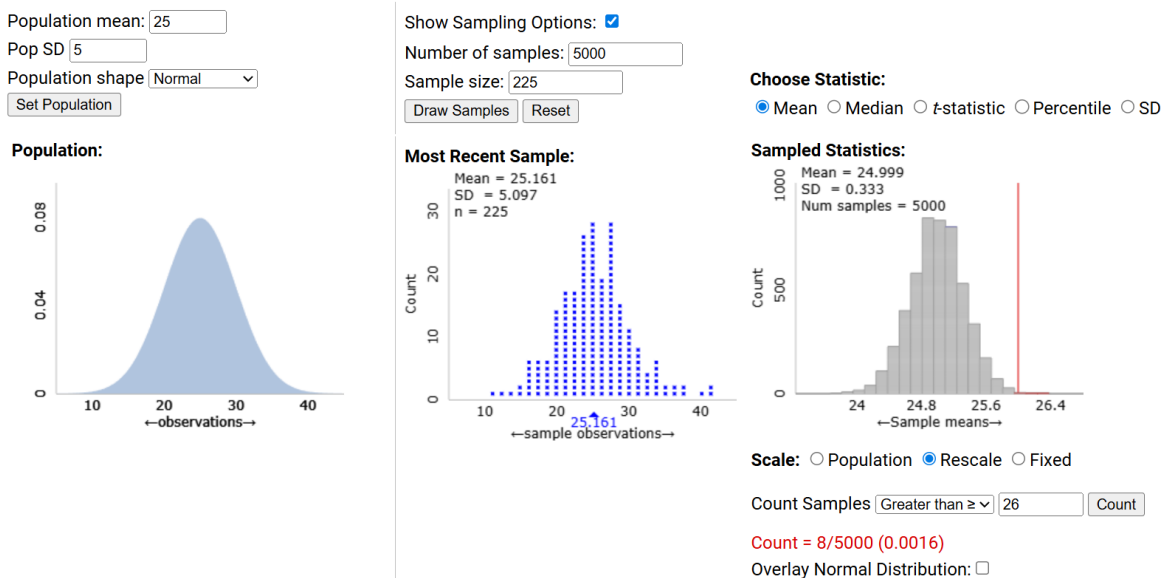
Tabla 6: Componentes del Modelo de Argumentación de Toulmin

En el caso de haber realizado una PH desde el enfoque IEI, el modelo se implementaría como se muestra a continuación partiendo de la siguiente situación problema tomada de Batanero y Díaz (2015):

Un director sanitario sostiene que el Índice de Masa Corporal (IMC) medio de los adolescentes de su distrito no supera el nivel 25 (sobrepeso). Para contrastar su afirmación toma una muestra aleatoria de 225 adolescentes que da como resultado un IMC medio de 26. Sabiendo que el IMC sigue una distribución Normal con desviación típica 5 discuta, mediante un contraste de hipótesis con $H_0: \mu \leq 25$, si la afirmación del director sanitario es correcta, con un nivel de significación del 5%.

Al resolver la PH mediante el enfoque IEI, los estudiantes realizarían la simulación que podría verse así:

One Variable with Sampling



Una posible interpretación del resultado por parte de los estudiantes podría ser la siguiente: “Después de simular 5000 muestras, se rechaza la hipótesis nula, es decir, la media poblacional del IMC supera el nivel 25, dado que sólo en 8 de 5000 muestras sucedió una media igual a la de la muestra del estudio o más extremo, por lo que es muy probable que el IMC medio sea mayor que 25. Las simulaciones muestran que obtener muestras con media de 26 o más, con dicha hipótesis nula, es muy poco probable. O quizás la muestra obtenida en el estudio es atípica y por ello resultó con una media muy alejada”.

Componente	Relación con la IEI
Afirmación o Conclusión	Generalización más allá de los datos. (ejemplo: “la media poblacional del IMC supera el nivel 25”).
Datos	Información o evidencia que apoya a la afirmación. (ejemplo: “en 8 de 5000 muestras sucedió una media igual a la de la muestra del estudio o más extremo”).
Garantía	La conexión lógica entre los datos y la afirmación. (En este caso no se expresó una garantía)
Respaldo	Apoyo adicional para la garantía. (ejemplo: “Las simulaciones muestran que obtener muestras con media de 26 o más, con dicha hipótesis nula, es muy poco probable”).
Calificador Modal	Indica el nivel de certeza de la afirmación. (ejemplo: “probablemente”, “es muy probable”, “es posible que”, en este caso “es muy probable”).
Contraargumento	Considera condiciones en las que la afirmación no se sostendría.

	(ejemplo: “O quizás la muestra obtenida en el estudio es atípica y por ello resultó con una media muy alejada”).
--	--

Tabla 7: Ejemplo de uso del modelo de argumentación para la PH.

6 Aspectos metodológicos

6.1 Características iniciales

El proyecto de intervención se implementará con un grupo de estudiantes de alguno de los programas educativos del área económica y administrativa de la Universidad de Sonora, bajo las condiciones de haber tenido cierto grado de experiencia realizando simulaciones de muestreo repetido y haber trabajado un poco con distribuciones de muestreo.

Se requerirán 3 sesiones de 1 hora cada una y el autor de este documento fungirá como el docente en el momento de la implementación.

6.2 Metodología para el desarrollo de la intervención: experimentos de diseño

Para el desarrollo de este trabajo se ha considerado el uso de los experimentos de diseño como enfoque metodológico del proyecto. Comúnmente, los experimentos de diseño tienen el propósito de desarrollar, tanto teorías sobre cierto proceso de aprendizaje, como características para la elaboración de un diseño didáctico. Por lo que, siendo este trabajo un proyecto de intervención didáctica, también se puede hacer uso de esta metodología dado que es altamente intervencionista, además de permitir un diseño iterativo basado en el análisis retrospectivo de conjeturas.

Según Cobb (2007), citado por Cortina (2014), los experimentos de diseño son una metodología de la educación matemática cuyo fin principal es el desarrollo de recursos educativos, que contribuyan al mejoramiento de la enseñanza y el aprendizaje de las matemáticas en el ámbito escolar.

Para este proyecto, el experimento de diseño constará de 5 etapas:

1. *Análisis preliminar o exploratorio*. Primeramente, se identificará la problemática, el contenido matemático, es decir, las PH y el contexto de enseñanza. (Revisión de la literatura, identificación de dificultades de aprendizaje y el establecimiento de metas de aprendizaje).
2. *Diseño de la intervención*. Se creará una secuencia didáctica o material educativo mediante una THA con base en los elementos teóricos previamente mencionados. Aquí se diseñarán actividades y se formularán hipótesis de aprendizaje y/o de diseño, adicionalmente se puede considerar el rol del docente y del entorno).
3. *Implementación y experimentación (fase cíclica)*. La intención es aplicar el diseño de la THA en un entorno real (aula) y observar cómo funciona. En esta etapa se recolectarán datos cualitativos y/o cuantitativos (registros escritos y grabaciones

en audio), se observará la interacción estudiante-actividad-docente y se ajustará el diseño en función de lo que se observa. Esta etapa podrá repetirse tantas veces se considere necesario para mejorar el diseño según se detecten oportunidades para esto.

4. *Análisis retrospectivo*. Aquí se analizarán los datos recogidos a través de las componentes del modelo de argumentación de Toulmin adaptado al RIE y se contrastará con las hipótesis de diseño. Se evaluarán los logros de aprendizaje, se identificarán patrones en la construcción del conocimiento, y adicionalmente, permitirá el refinamiento de teorías de aprendizaje.
5. *Revisión del diseño y teoría*. En esta etapa el propósito es ajustar la intervención y/o los elementos teóricos con base en la experiencia de cada implementación.

La finalidad del experimento de diseño es la construcción de una propuesta didáctica mejorada y, además, la obtención de un conocimiento más profundo del proceso de aprendizaje. Cabe señalar que, según Cortina (2013), un experimento de diseño implica la puesta a prueba de una THA.

Por otro lado, como ya se mencionó, el diseño instruccional se apoyará con el applet “Inferencia de una proporción” de la colección de applets del sitio web de Rossman y Chance. Otros trabajos se han apoyado en el software Fathom (Inzunza, 2013; Silvestre et al., 2024; Sepúlveda y Sánchez, 2024), o VUStat (van Dijke-Droogers et al., 2019), sin embargo, éste quedó discontinuado por parte de los desarrolladores. Además, fue sugerido por Chance y Rossman (2006) y Batanero y Díaz (2015). A su vez, es de fácil acceso desde cualquier dispositivo con acceso a Internet.

7 La propuesta y sus características

7.1 Características generales

La propuesta consiste en la construcción de una THA para introducir la enseñanza de las PH a estudiantes de nivel superior en las áreas de ciencias económicas y administrativas, desde el enfoque IEI. Para dicha construcción se revisaron algunos diseños propuestos en la literatura y las trayectorias elaboradas.

La más cercana al tipo de propuesta que se presenta en este trabajo, es la construida en Silvestre et al. (2024). Sin embargo, como se mencionó previamente, hay elementos no considerados y que han sido sugeridos en la investigación, uno de ellos es el contexto. Los contextos de las situaciones problema no corresponden al área de estudio, lo que podría ser un factor clave en el desarrollo de su RIEI.

Un ejemplo se muestra a continuación:

En una publicación de “Datos abiertos del Gobierno del Estado de Sonora” (ver: <https://datos.sonora.gob.mx/dataset/Beneficiarios%20de%20becas%20o%20est%C3%ADmulos%20educativos/resource/6846a873-bec6-4e86-9ce7-cd6225d557ee>) se reportó que en una muestra de 12,792 estudiantes de licenciatura becarios entre 2020 y 2022, el importe mensual medio fue de \$12,081 con una desviación estándar muestral de \$12,792.

Las PH que se solicitarán realizar a los estudiantes estarán centradas en pruebas para la media bajo el supuesto de normalidad. Si bien varios trabajos han propuesto la implementación de PH para proporciones o para la distribución binomial, lo cual constituye una posible vía introductoria, pocos se han centrado en la media siendo una medida muy utilizada. De ahí surge el interés de este proyecto de intervención por adoptar dicha característica como eje principal.

7.2 Características consideradas para el diseño de la THA

Para este trabajo se considera la necesidad de establecer cuatro tareas que se desglosan a continuación:

Tarea 1:

Esta tiene por objetivo dar entrada a los estudiantes en el tema de las PH para la media poblacional mediante el enfoque IEI. La tarea consiste en resolver una PH usando el RIEI, en un contexto del área económica y administrativa, con datos reales para inferir con respecto a la media poblacional. A su vez, ésta se dividirá en tres momentos.

En el primer momento se les presentará a los estudiantes una situación problema con datos reales y se les pedirá resolver una PH con respecto a la media poblacional, con cierta desviación estándar poblacional conocida. Ellos deberán hacer uso de la simulación para realizar el proceso de muestreo repetido para poder tomar una decisión, interpretar su conclusión con base en el contexto y acompañarla por una argumentación tratando de que ésta muestre suficiente información como para cubrir todos los componentes del modelo de argumentación de Toulmin. Para esta resolución, el profesor intervendrá como guía.

En el segundo momento se les pedirá a los estudiantes realizar una PH por cuenta propia a partir de una nueva situación problema y se especificará un cierto nivel de significancia. En el tercer momento, se retomará la situación 2 para pedir a los estudiantes realizar la PH pero bajo el supuesto de tener una media muestral distinta y con el mismo nivel de significancia.

Tarea 2:

El objetivo de la tarea 2 es que el estudiante reconozca la dispersión de la distribución como una característica importante a considerar para la PH. La tarea consiste en resolver una PH usando el RIEI y suponiendo una desviación estándar distinta. Esta tarea se dividirá en dos momentos.

El primer momento consistirá en partir de la situación 1 y solicitar a los estudiantes la realización de una PH bajo el supuesto de tener una desviación estándar más pequeña. Para el segundo momento se partirá de la situación 2 pero bajo el supuesto de tener una desviación estándar más grande.

Tarea 3:

El objetivo de esta tarea es que el estudiante identifique como importante el tamaño de muestra y que esto sea considerado en la interpretación y argumentación al realizar una PH. La tarea consistirá en resolver una PH usando el RIEI, suponiendo un tamaño de muestra mayor. Para ello, dicha tarea se dividirá en dos momentos.

El primer momento retomará la situación 1 para realizar una PH pero con un tamaño de muestra más grande. El profesor intervendrá como guía para este momento de la tarea. Para el segundo momento, se le pedirá al estudiante resolver una PH, retomando la situación 2, bajo el supuesto de tener un tamaño de muestra mayor, por cuenta propia.

Tarea 4:

El objetivo de la tarea 4 es que el estudiante considere importante el nivel de significancia como otro aspecto esencial de la PH, por lo que la tarea 4 solicitará partir de la situación 2 para resolver una PH usando el RIEI, pero con distinto nivel de significancia, por cuenta propia.

8 Cronograma

En la siguiente tabla se ilustran las actividades que se planean realizar para el siguiente semestre (2026-1):

Actividad	Descripción
Primer bosquejo de diseño de la THA con sus respectivas actividades. Entre enero y febrero de 2026.	Se llevará a cabo la construcción de la THA considerando los componentes, tareas, objetivos y otras características descritas. También se diseñarán las actividades de cada tarea de la trayectoria.
Implementación de un primer ciclo de diseño. Entre abril y mayo de 2026.	Se llevará a cabo la implementación de un primer ciclo de diseño de la THA en un grupo de estudiantes de la Universidad de Sonora que se encuentren estudiando algún programa educativo del área económica y administrativa.
Análisis del primer ciclo de diseño. Entre mayo y julio de 2026	Se llevará a cabo el análisis de las argumentaciones de los estudiantes mediante el modelo de argumentación de Toulmin adaptado para el RIE.

9 Referencias

Anderson, D., Sweeney, D. J. y Williams, T. (2008). *Estadística para administración y economía, 10a. Ed* (10a ed.). Cengage Learning Editores S.A. de C.V.

Batanero, C. (2000). Controversias sobre el papel de los contrastes de hipótesis en la investigación experimental. *Mathematical thinking and learning*, 2(1–2), 75–97. https://doi.org/10.1207/s15327833mtl0202_4

Batanero, C. (2001). *Didáctica de la estadística*. Grupo Editorial Universitario.

Batanero, C. (2006). The challenges of teaching statistical inference. *Jornadas de Classificação e Analise de Dados*, Universidad de Lusiada Lisboa: CLAD. CD ROM

Batanero, C. y Díaz, C. (2015). Aproximación informal al contraste de hipótesis. *II Jornadas Virtuales de Didáctica de la Estadística, la Probabilidad y la Combinatoria*. Granada: SEIEM.

Biehler, R., Ben-Zvi, D., Bakker, A., y Makar, K. (2013). Technology for enhancing statistical reasoning at the school level. En *Third International Handbook of Mathematics Education* (pp. 643–689). Springer New York.

Birnbaum, I. (1982). Interpreting statistical significance. *Teaching Statistics*, 4, 24–27.

Canal, G. y Behar, R. (2010). The confidence intervals: a difficult matter, even for experts. En C. Reading (Ed.), *Proceedings of the Eighth International Conference on Teaching Statistics*. Voorburg, The Netherlands: International Statistical Institute. En línea: https://iase-web.org/documents/papers/icots8/ICOTS8_C143_CANAL.pdf?1402524973

Case, C., y Jacobbe, T. (2018). A framework to characterize student difficulties in learning inference from a simulation-based approach. *Statistics education research journal*, 17(2), 9–29. <https://doi.org/10.52041/serj.v17i2.156>

Castro Sotos, A. E., Vanhoof, S., Van den Noortgate, W., y Onghena, P. (2007). Students' misconceptions of statistical inference: A review of the empirical evidence from research on statistics education. *Educational Research Review*, 2(2), 98–113. <https://doi.org/10.1016/j.edurev.2007.04.001>

Castro Sotos, A. E., Vanhoof, S., Van den Noortgate, W., y Onghena, P. (2009). How Confident are Students in their Misconceptions about Hypothesis Tests? *Journal of Statistics*

Education: An International Journal on the Teaching and Learning of Statistics, 17(2).
<https://doi.org/10.1080/10691898.2009.11889514>

Chance, B., delMas, R. y Garfield, J. (2004). Reasoning about sampling distributions. *The Statistics Teacher Network*, 64, 1–6.

Cobb, G. W. (2007). The introductory statistics course: A Ptolemaic curriculum? *Technology Innovations in Statistics Education*, 1(1), p. 1-15. <https://doi.org/10.5070/t511000028>

delMas, R., Garfield, J. y Chance, B. (2007). Assessing the effects of a simulation activity on students' conceptual understanding of sampling distributions. *Statistics Education Research Journal*, 6(2), 28–58.

Dolor, J., y Noll, J. (2015). Using guided reinvention to develop teachers' understanding of hypothesis testing concepts. *Statistics education research journal*, 14(1), 60–89.
<https://doi.org/10.52041/serj.v14i1.269>

Estrella, S., Méndez-Reina, M., & Vidal-Szabó, P. (2023). Exploring informal statistical inference in early statistics: A learning trajectory for third-grade students. *Statistics education research journal*, 22(2), 10. <https://doi.org/10.52041/serj.v22i2.426>

Falk, R. (1986). Misconceptions of statistical significance. *Journal of Structural Learning*, 9(1), 83–96.

Falk, R. y Konold, C. (1997). Making sense of randomness: Implicit encoding as a basis for judgment. *Psychological Review*, 104(2), 301–318. <https://doi.org/10.1037/0033-295X.104.2.301>

García, V. y Sánchez, E. (2014). Razonamiento inferencial informal: el caso de la prueba de significación con estudiantes de bachillerato. En M. T. González, M. Codes, D. Arnau y T. Ortega (Eds.), *Investigación en Educación Matemática XVIII* (pp. 345-354). Salamanca: SEIEM

García, V. (2017). Diseño de una trayectoria hipotética de aprendizaje para la introducción y desarrollo del razonamiento sobre el contraste de hipótesis en nivel medio superior [Tesis de Doctorado, CINVESTAV]. Repositorio Institucional – CINVESTAV.

Garfield, J. y Ben-Zvi, D. (2008). *Developing Students' Statistical Reasoning: Connecting Research and Teaching Practice*. Springer.

Inzunza Cázares, S. (2013). Un Acercamiento Informal a La Inferencia Estadística Mediante Un Ambiente Computacional Con Estudiantes De Bachillerato. *Revista AMIUTEM*, 1(1),

61–76. Recuperado a partir de <https://revista.amiutem.edu.mx/index.php/relecamiutem/article/view/7>

Inzunza, S., y Jiménez, J. (2013). Caracterización del razonamiento estadístico de estudiantes universitarios acerca de las pruebas de hipótesis. *Revista latinoamericana de investigación en matemática educativa*, 16(2), 179–211.

<https://doi.org/10.12802/relime.13.1622>

Inzunza, S., e Islas, E. (2019). Diseño y evaluación de una trayectoria hipotética de aprendizaje para intervalos de confianza basada en simulación y datos reales. *Bolema Boletim de Educação Matemática*, 33(63), 1–26. <https://doi.org/10.1590/1980-4415v33n63a01>

Krauss, S. y Wassner, C. (2002). How significance tests should be presented to avoid the typical misinterpretations. En B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching Statistics*. Cape Town, South África: International Association for Statistics Education. En línea: https://iase-web.org/documents/papers/icots6/3m4_krau.pdf?1402524961

Krishnan, S. e Idris, N. (2014). Students' misconceptions about hypothesis test. *Journal of Research in Mathematics Education*, 3(3), 276–293. <https://doi.org/10.4471/redimat.2014.54>

Konold, C. y Pollatsek, A. (2002). Data analysis as the search for signals in noisy processes. *Journal for research in mathematics education*, 33(4), 259. <https://doi.org/10.2307/749741>.

Konold, C. y Kazak, S. (2008). Reconnecting data and chance. *Technology Innovations in Statistics Education*, 2(1).

Lehmann, E. (1999). *Elements of Large Sample Theory*. Springer, New York. <https://doi.org/10.1007/b98855>

Lipson, K. (2000). *The role of the sampling distribution in developing understanding of statistical inference*. Edu.au. Sitio web:

<https://researchbank.swinburne.edu.au/file/fc076d2d-b4a5-4c00-bfd7-84f2c3d0eb9e/1/Kay%20Lipson%20Thesis.pdf>

Liu, Y., y Thompson, P. W. (2005). Teachers' understanding of hypothesis testing. In *Proceedings of the Twenty-seventh Annual Meeting of the International Group for the Psychology of Mathematics Education*.

Liu, Y., y Thompson, P. W. (2009). Mathematics teachers' understandings of proto-hypothesis testing. *Pedagogies An International Journal*, 4(2), 126–138.

<https://doi.org/10.1080/15544800902741564>

- Lovett, M. (2001). A collaborative convergence on studying reasoning processes: A case study in statistics. En *Cognitive Science Approaches to Testing* (pp. 135–175).
- Makar, K. y Rubin, A. (2009). A framework for thinking about informal statistical inference. *Statistics Education Research Journal*, 8(1), 82-105. Recuperado de [http://www.stat.auckland.ac.nz/~iase/serj/SERJ8\(1\)_Makar_Rubin.pdf](http://www.stat.auckland.ac.nz/~iase/serj/SERJ8(1)_Makar_Rubin.pdf)
- Nilsson, P. (2023). Hypothetical learning trajectory on informal hypothesis testing in a probability context. *Statistics Education Research Journal*, 22(2), 1–14. <https://doi.org/10.52041/serj.v22i2.425>
- Pfannkuch, M. (2005). Probability and statistical reasoning. En J. Kilpatrick et al. (Eds.), *International Handbook of Mathematics Education* (pp. 949–987). Springer.
- Pfannkuch, M. (2006). Informal Inferential Reasoning. En A. Rossman y B. Chance (Ed.), *Proceedings of the Seventh International Conference on Teaching Statistics*. Salvador, Bahia, Brazil. En línea: https://www.ime.usp.br/~abe/ICOTS7/Proceedings/PDFs/InvitedPapers/6A2_PFAN.pdf
- Pratt, D. y Ainley, J. (2008). Making sense of uncertainty in the mathematics classroom. *ZDM*, 40, 611–620.
- Ramsey, F., y Schafer, D. (2002). *The statistical sleuth: A course in methods of data analysis* (2nd ed.). Belmont, CA: Duxbury Press. ISBN: 0534386709
- Retamal, M., Martínez, H., Sáez, R. y Lugo-Armenta, J. (2023). Exploring challenges and strategies in teaching hypothesis testing to engineering students from the perspective of educators. *Eurasia journal of mathematics science and technology education*, 19(12), em2371. <https://doi.org/10.29333/ejmste/13865>
- Rivadulla, A. (1991). Probabilidad e inferencia científica. *Barcelona: Anthropos*.
- Rossman, A. (2008). Reasoning about informal statistical inference: One statistician’s view. *Statistics education research journal*, 7(2), 5–19. <https://doi.org/10.52041/serj.v7i2.467>
- Rossman, A. y Chance, B. (2014). Using simulation-based inference for learning introductory statistics. *Wiley Interdisciplinary Reviews: Computational Statistics*, 6(4), 211–221.
- Segundo, J. (2018). *Un acercamiento al razonamiento inferencial estadístico informal en carreras del área económica y administrativa* [Tesis de Maestría, Universidad de Sonora]. Repositorio Institucional – Universidad de Sonora.

Silvestre, E., Armenta, M. e Inzunza, S. (2024). Diseño y Evaluación de una Trayectoria Hipotética de Aprendizaje orientada a introducir la Prueba de Hipótesis desde un Acercamiento Informal. *Bolema: Boletim de Educação Matemática*.

Simon, M. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education*, 26(2), 114–145.
<https://doi.org/10.5951/jresematheduc.26.2.0114>

Simon, M. (2020). Hypothetical Learning Trajectories in Mathematics Education. En S. Lerman (Ed.), *Encyclopedia of Mathematics Education*, 354-357.

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. <https://doi.org/10.4324/9780203063279>

Stohl, H., Angotti, R. y Tarr, J. (2010). Making comparisons between observed data and expected outcomes: Students' informal hypothesis testing with probability simulation tools. *Statistics Education Research Journal*, 9(1), 68–96. <https://doi.org/10.52041/serj.v9i1.388>

TecNM (2016). Programa de estudio de curso de Estadística para Administración 2 para la Licenciatura en Administración del Tecnológico Nacional de México. Recuperado de: <https://ith.mx/documentos/reticulas/admin/SEMESTRE%203/LAD-1017%20Estad%20C3%ADstica%20para%20la%20administraci%20C3%B3n%20II.pdf>

UNISON (2018). Programa de estudio de curso Estadística en los Negocios en la Facultad Interdisciplinaria de Ciencias Económicas y Administrativas de la Universidad de Sonora. Recuperado de https://www.mat.uson.mx/sitio/documentos/dcea/estadistica_en_los_negocios.pdf

UNISON (2022). Programa de estudio de curso Estadística Inferencial en la Facultad Interdisciplinaria de Ciencias Económicas y Administrativas de la Universidad de Sonora. Recuperado de https://www.mat.uson.mx/web/wp-content/uploads/Lic_Finanzas_estadistica_inferencial.pdf

Vallecillos, A. (1994). Estudio teórico-experimental de errores y concepciones sobre el contraste estadístico de hipótesis en estudiantes universitarios [Tesis de Doctorado, Universidad de Granada].

Vallecillos, A. (1999). Some empirical evidence on learning difficulties about testing hypotheses. *Proceedings of the 52 session of the International Statistical Institute* (Vol.2, pp. 201–204). Helsinki: International Statistical Institute.

Vallecillos, A., y Batanero, C. (1997). Conceptos activados en el contraste de hipótesis estadísticas y su comprensión por estudiantes universitarios. *Recherches En Didactique Des Mathématiques*, 17(1), 29–48. <https://revue-rdm.com/1997/conceptos-activados-en-el/>

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*, 22(2), 116–138. <https://doi.org/10.1080/10986065.2019.1617025>

Vera, O., Díaz, C. y Batanero, C. (2011). Dificultades en la formulación de hipótesis estadísticas por estudiantes de Psicología. *Unión*, 27, 41-61.

Webster, A. (2000). *Estadística aplicada a los negocios y la economía* (3.a ed.). Irwin McGraw-Hill.

Wild, C. y Pfannkuch, M. (1999). Statistical thinking in empirical enquiry. *Revue Internationale de Statistique [International Statistical Review]*, 67(3), 223–248. <https://doi.org/10.1111/j.1751-5823.1999.tb00442.x>

Wilensky, U. (1997). What is normal anyway? Therapy for epistemological anxiety. *Educational Studies in Mathematics*, 33, 171–202.

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. <https://doi.org/10.52041/serj.v7i2.469>