UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
ESCUELA NACIONAL COLEGIO DE CIENCIAS Y HUMANIDADES

PROGRAMAS DE ESTUDIO 2024

ÁREA DE MATEMÁTICAS

ESTADÍSTICA Y PROBABILIDAD I-II

Primera edición: julio de 2024.

ÍNDICE

ESTADÍSTICA Y PROBABILIDAD I

ESTADÍSTICA Y PROBABILIDAD II

PRESENTACIÓN DE LA MATERIA

Ubicación de la materia en el marco del mapa curricular

La asignatura Estadística y Probabilidad se enmarca en el Área de Matemáticas y se ofrece como materia optativa en los semestres quinto y sexto. Se conforma de las asignaturas Estadística y Probabilidad I y II, en las que se brinda una visión de conjunto de los principales conceptos y procedimientos de la estadística y se promueven en el alumnado habilidades para llevar a cabo investigaciones estadísticas en el campo profesional de su preferencia y para el análisis e interpretación de la información en la vida cotidiana.

En el desarrollo de la materia se aplican y recuperan aprendizajes de los cursos previos del Área de Matemáticas, entre ellos, algunos relacionados con conceptos aritméticos, variación proporcional, funciones y geometría analítica. Al mismo tiempo, la necesidad de usar softwares especializados y aplicaciones conduce a aprovechar la formación del alumnado en Taller de Cómputo.

La materia de Estadística y Probabilidad contribuye al desarrollo de otras áreas de estudio en el nivel medio superior y superior ya que, como herramienta para la investigación y la resolución de problemas, se puede utilizar de diversas formas en todas las áreas del conocimiento contempladas en el Colegio: Matemáticas, Ciencias Experimentales, Histórico Social y Talleres de Lenguaje y Comunicación.

Por ejemplo, en Biología se utilizan conceptos de probabilidad para estudiar la genética de poblaciones animales o vegetales; en Química y Física se requiere tomar mediciones, diseñar experimentos, sistematizar observaciones y hacer inferencias; en Geografía, Economía y Antropología se manejan datos cuantitativos y cualitativos como materia prima de trabajo; en Administración se emplea la información estadística para tomar decisiones en escenarios de incertidumbre; en Letras se aplican técnicas estadísticas para validar la autoría de textos literarios mediante el análisis de patrones; en Derecho, las investigaciones penales deben considerar la incertidumbre asociada a las evidencias y los testimonios; entre otras muchas aplicaciones.

Enfoque disciplinario y didáctico

La Estadística es una disciplina que se ocupa de resolver problemas que implican la recolección, el análisis y la interpretación de datos. Su estudio se inscribe en los objetivos del Área de Matemáticas por ser un instrumento poderoso para que el alumnado conozca y descubra su entorno físico y social en un proceso que admite titubeos, conjeturas y aproximaciones y, al mismo tiempo, se apoya en el rigor de los conceptos probabilísticos para la conformación de su estructura teórica.

Los datos reales no se comportan de acuerdo con ningún modelo matemático exacto, como los que han aprendido previamente quienes ingresan a los cursos de Estadística y Probabilidad. Esta forma de variar es lo que se denomina variabilidad de los datos.

Las fuentes de la variabilidad son muy diversas, entre ellas se encuentra la heterogeneidad natural de los individuos que componen una población, las discrepancias que se producen al medir repetidamente un mismo objeto o sujeto con fines de precisión, las fluctuaciones en los resultados de elecciones o asignaciones aleatorias en diferentes experimentos y procedimientos, y muchas otras.

La variabilidad es una característica inherente a los fenómenos naturales (no hay dos seres vivos idénticos) y a los procesos sociales (no hay dos colectivos que se relacionen exactamente de la misma forma). Para analizar estos fenómenos y procesos y ser capaces de tomar decisiones frente a ellos, es fundamental comprender el concepto de variabilidad, asumir sus implicaciones y cuantificarla. De ahí que la Estadística se pueda describir como el estudio de la variabilidad.

La Estadística se basa en la probabilidad para construir un marco teórico que permite estudiar una característica o variable de interés en una población, o bien el análisis de las posibles relaciones entre dos o más características observadas en cada elemento de la población. Los estudios estadísticos emplean una metodología basada en la recolección de información parcial a través de la selección de una muestra aleatoria. El análisis inicial de esta información posibilita realizar inferencias informales para después avanzar a la construcción de modelos probabilísticos que sustentan teóricamente las inferencias estadísticas formales.

El contenido disciplinario de la materia consta de tres bloques: el bloque de análisis de la información, el de probabilidad y el de inferencias estadísticas. Cada uno de ellos abarca dos unidades.

La primera unidad de Estadística y Probabilidad I incluye una serie de técnicas para organizar y visualizar los datos de una variable estadística de manera que se puedan identificar fácil y rápidamente sus aspectos más relevantes. Más allá de construir tablas y gráficas o calcular medidas, las actividades de esta unidad deben conducir a que el alumnado formule diversas observaciones acerca de la información que proporcionan los datos y responda las preguntas que se le planteen para facilitar el análisis de esa información.

Se debe propiciar que el alumnado desarrolle una visión global de los datos que le permita analizar su comportamiento y comparar diferentes colecciones de datos. Para ello, es necesario que pase de enfocarse en valores individuales (como el mínimo, el máximo o el más frecuente), a considerar la distribución de frecuencias como un todo, con su centro y su medida de dispersión.

En la segunda unidad de Estadística y Probabilidad I se extiende el análisis de la información a la posible existencia de una relación entre dos características que se expresan en un conjunto de datos bivariados, ya sea que correspondan a dos variables cualitativas o a dos variables cuantitativas.

Para que los datos sean útiles y relevantes, es necesario situarlos en contextos específicos y vinculados a alguna toma de decisión o resolución de un problema. Los datos no pueden analizarse de manera significativa si no se tiene en cuenta el contexto en el que se generaron y los métodos que se utilizaron para recolectarlos.

El profesorado debe fomentar el intercambio de ideas entre el alumnado sobre cómo generalizar a las poblaciones los hallazgos obtenidos en muestras. Esto se logra promoviendo que el alumnado formule inferencias informales, que puedan explicar en qué evidencias se apoyan y qué factores pueden introducir sesgos o limitaciones en sus conclusiones.

Asimismo, será necesaria la utilización de aplicaciones tecnológicas que permitan, por ejemplo, visualizar en qué consiste el criterio de mínimos cuadrados o bien que faciliten la construcción de tablas de contingencia, de porcentajes por renglón o por columna, de diversos tipos de gráficas y el cálculo de las medidas, para destinar más tiempo al análisis de la información.

El bloque de probabilidad abarca la Unidad 3 de la primera asignatura y la Unidad 1 de Estadística y Probabilidad II, y consiste en los conceptos y resultados probabilísticos que serán necesarios para introducir las inferencias estadísticas formales. En la parte final de Estadística y Probabilidad I se trabajará la asignación de una probabilidad a cada evento a través del enfoque clásico, así como su aproximación mediante el enfoque frecuencial y el subjetivo.

El uso de algún software o aplicación para calcular las frecuencias relativas asociadas a un resultado en un número cada vez más grande de repeticiones de un experimento aleatorio, permitirá que el alumnado visualice la tendencia de esas cantidades a largo plazo, comprenda la regularidad estadística y le dé un sentido adecuado al valor de una probabilidad.

La deducción de las reglas básicas para el cálculo de probabilidades y de la regla para la probabilidad condicional deberán conducir a la identificación de eventos mutuamente excluyentes y eventos independientes. La independencia es una hipótesis esencial en el desarrollo de los temas posteriores.

Al inicio de la segunda asignatura, se introducirá el concepto clave de variable aleatoria, que permite enfocarse en una sola característica numérica de los resultados de un experimento aleatorio. Este concepto es esencial para pasar del análisis de probabilidades de eventos dispersos a la construcción de funciones que representen las probabilidades de todos los eventos de un mismo tipo, conocidas como distribuciones de probabilidad o funciones de densidad, con su centro y su medida de variabilidad.

Para estudiar las variables aleatorias continuas, se usarán ejemplos donde el conteo se reemplaza por el cálculo de una medida geométrica (longitud o área) al aplicar el enfoque clásico; de esta manera el alumnado podrá descubrir que la probabilidad de que este tipo de variables tomen un valor particular es cero, ya que la longitud y el área de un punto es cero, y solo puede ser positiva la probabilidad de que tomen valores en intervalos.

Este bloque termina con el estudio de dos importantes modelos probabilísticos que serán de gran utilidad en la última parte del curso: las distribuciones binomial y normal.

El bloque de inferencias estadísticas abarca las dos últimas unidades de la segunda asignatura e inicia con el análisis de la media y la proporción de una muestra aleatoria, y su identificación como estimadores de los parámetros media y proporción. Una de las ventajas de los métodos inferenciales es que toman en cuenta la probabilidad de cometer un error al aplicarlos. Por ello, para emplear los estadísticos como estimadores es necesario conocer qué relación guarda su distribución de probabilidad o función de densidad con los parámetros de interés. Lo anterior da lugar al estudio de las distribuciones muestrales.

Para facilitar la comprensión y el uso de las distribuciones muestrales, se requiere nuevamente recurrir a softwares y aplicaciones que permitan visualizar una aproximación frecuencial de esas distribuciones. De esta forma, el alumnado podrá descubrir y comprobar que, independientemente de la forma de la distribución poblacional, la distribución de la media y de la proporción muestral tiende a una distribución normal conforme aumenta el tamaño de la muestra. Mediante estas actividades de indagación, el alumnado aprecia la enorme potencia y el alcance del Teorema del Límite Central cuyas hipótesis se reducen a la construcción de muestras aleatorias simples.

Se espera que, en la parte final del curso, el profesorado haya podido transmitir al alumnado la importancia de todo lo aprendido previamente para el estudio de las inferencias estadísticas. Un tratamiento mecánico de estos métodos contribuye a la confusión, desafortunadamente muy extendida, de que con ellos se *demuestran* afirmaciones acerca del valor del parámetro.

En la última unidad se pretende que el alumnado aprenda a estimar el valor de un parámetro o a contrastar dos hipótesis sobre dicho valor, basándose en la probabilidad de obtener los resultados observados al realizar una única muestra aleatoria. Aunque es posible que el valor observado en la muestra esté lejos del valor real del parámetro, el Teorema del Límite Central garantiza que es muy poco probable que eso ocurra para una muestra de tamaño adecuado y elegida correctamente.

Hay que resaltar en este enfoque disciplinario, la importancia de que se contemple la diversidad de aplicaciones de la estadística en el desarrollo científico y humanista en diferentes campos, por ejemplo, en estudios sobre el cuidado de especies en peligro de extinción, la protección del medio ambiente, la protección de cultivos con especies nativas; el respeto y la tolerancia a la diversidad, tanto en la naturaleza como en las relaciones humanas. Por otro lado, es conveniente incluir ejemplos de procesos en los que se requiere un control de la variabilidad, por ejemplo, en el campo de la producción, el de la industria farmacéutica o en la de alimentos, y en estudios sobre la calidad de la educación.

En el enfoque didáctico se reconocen tres niveles en el proceso de enseñanza y aprendizaje de la estadística, que no son ajenos entre sí.

- **La cultura estadística**
  Se refiere a la capacidad de comprender y utilizar el lenguaje, los símbolos y las técnicas estadísticas básicas, así como de interpretar las representaciones de los datos. Implica tener un buen conocimiento de la información que proporcionan los datos en contextos específicos y su aplicación para tomar decisiones o posiciones.
- **El razonamiento estadístico**
  Es la habilidad de argumentar los conceptos e ideas estadísticas y darle sentido a la información estadística; implica reconocer las conexiones entre los conceptos y combinar ideas acerca de los datos y la probabilidad. Requiere comprender y ser capaz de explicar procesos estadísticos y, por tanto, ser capaz de dar interpretaciones adecuadas de sus resultados.
- **El pensamiento estadístico**
  Implica la comprensión del por qué y el cómo se realizan las investigaciones estadísticas, así como la utilización de los modelos adecuados para simular los fenómenos aleatorios, reconociendo cómo, cuándo y por qué se pueden utilizar los instrumentos deductivos existentes. Incluye lo que hace un profesional de la estadística para cuestionar e investigar los problemas y los datos involucrados en un contexto específico.

En un curso introductorio a la estadística, como el que se imparte en el Colegio, se trata de abarcar fundamentalmente los niveles de cultura y razonamiento estadísticos. Dada la intersección entre los niveles, algunas actividades pueden ayudar también a desarrollar el pensamiento estadístico. Sin embargo, este último nivel requiere de una mayor profundización que no se puede alcanzar en el bachillerato.

Uno de los principales objetivos del Modelo Educativo del CCH es buscar el desarrollo del alumnado independiente, que tenga una visión crítica y que pueda analizar temas de su interés, defender sus puntos de vista con argumentos sólidos y participar en la toma de decisiones. Para ello, la educación estadística es un instrumento indispensable siempre y cuando el profesorado tenga claro el objetivo de fomentar estas habilidades al diseñar e impartir los cursos.

Se requiere diseñar experiencias de aprendizaje que fomenten la colaboración, el debate de los conceptos estadísticos, la capacidad de formular conjeturas, razonar y enfrentarse a retos metodológicos mediante una estrategia de resolución de problemas. Se trata de crear en el aula un ambiente de aprendizaje que brinde al alumnado muchas oportunidades para pensar, razonar y argumentar, así como para discutir y reflexionar con sus pares.

Las principales cuestiones que conviene tener en cuenta para generar este tipo de ambiente son: identificar las ideas estadísticas fundamentales y diseñar actividades orientadas a ellas; plantear tareas interesantes para el alumnado basadas en datos reales o realistas; usar herramientas tecnológicas acompañadas de preguntas que fomenten la reflexión del alumnado; crear una cultura en el aula que los anime a argumentar sus opiniones con criterios estadísticamente válidos y respetando ciertas normas para el desarrollo de las discusiones; fomentar un compromiso del alumnado con su propio aprendizaje y el de sus pares, y diseñar métodos de evaluación apropiados. Considerar todos estos aspectos en la planeación de las clases, modifica completamente la dinámica de la interacción dentro y fuera del aula.

En lugar de iniciar cada tema o subtema definiendo conceptos y procedimientos, se deben diseñar actividades y generar discusiones que faciliten al alumnado acercarse a ellos, contrastarlos con sus ideas previas y comprender su necesidad y utilidad para luego formalizar esas ideas en las definiciones.

Utilizar herramientas digitales para facilitar el aprendizaje de conceptos teóricos y teoremas importantes, permite que el alumnado visualice y explore los conceptos, así como descubrir los resultados por sí mismos, mediante secuencias didácticas que fomentan el desarrollo de su razonamiento estadístico. De esta manera se evita la exposición pasiva de los aspectos teóricos en el pizarrón.

Sin desarrollar el razonamiento probabilístico, los procedimientos para realizar inferencias estadísticas formales se reducen a la aplicación mecánica de fórmulas, impidiendo al alumnado alcanzar una comprensión real y una interpretación adecuada de sus resultados, por lo que otro objetivo importante es desarrollar el razonamiento probabilístico.

El aprendizaje de la probabilidad enfrenta abundantes desafíos y en su tratamiento suelen surgir intuiciones contrarias a las establecidas en la teoría, paradojas y concepciones erróneas. El razonamiento probabilístico es la forma de razonar acerca de juicios, afirmaciones y justificaciones para la toma de decisiones en presencia de la incertidumbre. Este tipo de razonamiento está relacionado con el razonamiento estadístico y el desarrollo de uno contribuye al del otro.

El tratamiento de problemas de probabilidad debe incluir la simulación, tanto física como usando software especializado. La simulación física ayuda a entender cómo se modelan las características esenciales del experimento real, mientras que la simulación con software permite realizar una gran cantidad de repeticiones para estudiar el comportamiento a largo plazo y mostrar la regularidad estadística de los experimentos aleatorios.

En lo que respecta a la evaluación, se debe asegurar que sea integral y continua, empezando por evaluaciones diagnósticas que verifiquen los prerrequisitos y muestren los conocimientos previos antes de abordar los diferentes contenidos. La continuidad se logrará mediante evaluaciones formativas que permitan valorar los procesos requeridos para los aprendizajes que se van logrando durante el desarrollo de los temas.

Para evaluar el aprendizaje global de cada unidad o tema, se aplicará una evaluación sumativa que permita verificar si el alumnado ha logrado una comprensión adecuada de los conceptos y procedimientos estadísticos, así como su combinación e interacción en la solución de problemas relacionados con la toma de decisiones.

El profesorado elegirá el tipo de evaluación que realizará (mediante proyectos, basado en problemas, estudios de caso u otros) y los instrumentos de evaluación adecuados a su forma particular de impartir los cursos. Lo importante es que sea integral y continua, y que brinde retroalimentación al alumnado y a él o la propia docente, identificando cuestiones que permitan un continuo ajuste a su planeación para mejorar el aprendizaje del alumnado a su cargo. Al diseñar las evaluaciones, es necesario tomar en cuenta que no se trata de evaluar la capacidad de memorizar las fórmulas o los pasos de un procedimiento, sino de fomentar el razonamiento estadístico frente a diversas situaciones.

Una forma de enriquecer la evaluación es incorporar la autoevaluación y la coevaluación como estrategias complementarias, de modo que el proceso evaluativo no dependa únicamente del criterio del o la docente.

Finalmente, para que este enfoque didáctico se materialice, se requiere de la iniciativa y creatividad del profesorado. Las cartas descriptivas de las dos asignaturas de esta materia ofrecen sugerencias sintéticas de estrategias que pretenden inspirar al profesorado a enriquecer su trabajo docente con un desarrollo más amplio y diverso que el que se muestra en esos ejemplos.

Concreción en la materia de los principios del Modelo Educativo del Colegio: *aprender a aprender, aprender a hacer, aprender a ser*

Los principios del Modelo Educativo del Colegio se concretan en la materia Estadística y Probabilidad de la siguiente forma:

Aprender a aprender

La disponibilidad casi inmediata de una gran cantidad de información en todos los campos del conocimiento y en todos los terrenos de la vida social y política, produce también una gran cantidad de información errónea o falsa que puede conducir a la manipulación mediante la generación de ideas y conceptos que no se apegan a la realidad o no corresponden a las teorías científicas y humanistas. A esto hay que agregar que la extensión en la aplicación de la inteligencia artificial (IA), junto al uso generalizado de dispositivos móviles, puede conducir a la idea de que la acción de aprender consiste solamente en buscar en la red.

El principio de *aprender a aprender* se concreta, entre otros aspectos, en formar un alumnado que desarrolle una fuerte autonomía para hacerse de conocimientos propios de manera continua a lo largo de su vida. Esta autonomía requiere que sepan investigar la información necesaria, sea a través de libros y revistas, de experimentos o levantamiento de encuestas, o sea a través de internet. La tendencia apunta a que la primera forma de acercarse a la información es la red.

Para que esa información contribuya al aprendizaje en cualquier área del conocimiento es necesario el alumnado que formamos sea capaz de discriminar la calidad y la veracidad de la información, que tenga una actitud crítica frente a su origen y a la forma en que se recolectó, y que sea capaz de analizar esa información. La estadística es una gran herramienta para el desarrollo de estas habilidades.

En la medida en que los cursos de la materia conviertan las aulas en espacios para aprender a argumentar las opiniones con criterios estadísticamente válidos, y que se vinculen los conceptos y procedimientos a la toma de decisiones, se brindará una formación que contribuya a que el alumnado forme su propio criterio en diversos temas y, al mismo tiempo, pueda modificar un punto de vista cuando encuentre evidencias que apunten en otra dirección. Esta actitud es fundamental para desarrollar un aprendizaje constante.

Aprender a hacer

Todo conocimiento adquiere sentido cuando se convierte en acción práctica. La estadística es una herramienta indispensable para resolver problemas que requieren datos y en ese sentido contribuye a la aplicación del conocimiento.

Un estudio estadístico empieza con la identificación de un problema en cualquier área del conocimiento y la formulación de las preguntas que se deben responder para arribar a una solución, y termina con la interpretación de la información obtenida mediante el análisis de los datos para dar respuesta a esas preguntas. Esto permite tomar decisiones frente a fenómenos naturales o procesos sociales caracterizados por la variabilidad. Lo mismo puede tratarse de estudios relacionados con el cuidado ambiental y la sustentabilidad, que estudios relacionados con la brecha salarial entre hombres y mujeres u otros problemas sociales vinculados a la perspectiva de género, entre otros muchos tipos de estudios.

En el desarrollo de los estudios estadísticos, se requiere la aplicación de herramientas computacionales para un buen manejo de grandes cantidades de datos. Los cursos de Estadística y Probabilidad incluyen en su desarrollo el acercamiento del alumnado a diversas herramientas de este tipo que contribuyen a una formación que contempla el eje transversal de uso de las TIC.

Por eso, el alumnado que haya adquirido educación estadística durante su estancia en el CCH contarán con una buena preparación para hacer realidad el principio de *aprender a hacer.*

Aprender a ser

En la formación del alumnado en el Colegio, no solo se integran saberes sino también valores. Los cursos de Estadística y Probabilidad buscan contribuir a formar un alumnado que:

- Tengan un pensamien:o crítico al analizar cualquier tema de su interés.
- Adquieran responsabilidad social a través del estudio de problemas que afectan a diversos sectores de la sociedad o de la naturaleza.
- Logren la socialización de su trabajo y la discusión de sus ideas con confianza, seguridad y respeto.
- Adquieran seguridad en sí mismos al defender sus puntos de vista con argumentos sólidos y al participar en la toma de decisiones.
- Se desarrollen con autonomía para la construcción de su conocimiento y para aprender de sus propias experiencias de vida.

Contribución de la materia al Perfil del Egresado

La formación que brinda la materia Estadística y Probabilidad contribuye a que el alumnado adquiera los siguientes conocimientos, habilidades, actitudes y valores:

- Desarrolla una cultura y un razonamiento estadístico que le permiten resolver problemas y tomar decisiones sobre diversos eventos relacionados con fenómenos naturales y procesos sociales, a partir de estudios estadísticos.
- Adquiere recursos adecuados para construir argumentaciones, interpretaciones y valoraciones sólidas y coherentes fundamentadas en la información que brindan los datos.
- Contribuye a la validez científica de estudios en cualquier área del conocimiento mediante la aplicación de inferencias estadísticas y la correcta interpretación de sus resultados.
- Utiliza herramientas como dispositivos tecnológicos y softwares computacionales en los estudios estadísticos para el manejo eficiente de grandes cantidades de datos.
- Comprende textos científicos o de divulgación.
- Asume una visión crítica frente a las fuentes de la información, la forma en que es recolectada y los métodos de análisis empleados en su tratamiento, para ser capaz de discriminar la información pertinente de aquella que distorsiona o manipula los datos sobre el objeto de estudio.
- Fortalece su seguridad y su autoestima mediante una construcción del conocimiento cada vez más autónoma y sólida.
- Adquiere valores como la tolerancia y el respeto a la diversidad.

- Asume un compromiso con acciones que fortalezcan la equidad de género, el respeto a la naturaleza y la defensa de la sustentabilidad.
- Asume un compromiso social basado en la socialización del trabajo y el debate respetuoso de ideas para adquirir una buena formación ciudadana.

Propósitos generales de la materia

El alumnado será capaz de:

- Contar con una cultura y un razonamiento estadísticos, que le permitirán la evaluación de la validez y relevancia de la información estadística, así como la interpretación de estudios estadísticos y la elaboración de inferencias formales, para la toma de decisiones y la definición de posiciones.
- Diseñar y aplicar los cuatro pasos de una investigación estadística: formulación de preguntas, recolección de datos, análisis de datos e interpretación de los resultados, para la construcción de conocimientos en diversos campos del saber considerando la variabilidad de los datos y el papel del azar.
- Emplear herramientas tecnológicas para ordenar y presentar datos, calcular medidas y construir gráficas, analizar tendencias y correlaciones entre dos variables estadísticas, realizar simulaciones, calcular probabilidades y para descubrir y comprender conceptos teóricos y conclusiones de grandes teoremas, con el fin de lograr la comprensión y aplicación de las principales técnicas estadísticas.

Panorama general de las unidades

	Estadística y probabilidad I	Estadística y probabilidad II
Unidad 1	28 hrs.	24 hrs.
	Análisis de información estadística para una variable	**Modelos de probabilidad y sus aplicaciones**
Unidad 2	10 hrs.	18 hrs.
	Datos bivariados	**Distribuciones muestrales**
Unidad 3	26 hrs.	22 hrs.
	Azar y probabilidad	**Inferencia estadística**
Total	64 hrs.	64 hrs.

Estadística y probabilidad I

PRESENTACIÓN DE LA ASIGNATURA DE ESTADÍSTICA Y PROBABILIDAD I

La asignatura de Estadística y Probabilidad I es el primer acercamiento del alumnado a una visión integral de los métodos para analizar la información que brinda una colección de datos, es decir, un conjunto de valores que toman ciertas variables estadísticas de interés en un estudio o en la resolución de un problema.

A diferencia de las variables involucradas en una relación funcional, las variables estadísticas representan características de los elementos de una población, por ejemplo, la preferencia por un candidato, el tiempo que tarda en hacer efecto un medicamento, la estatura de los perros de cierta raza o el largo promedio de las hojas de algún tipo de plantas. Su estudio requiere métodos diferentes a los que se usan en las ramas de las matemáticas que ha estudiado previamente el alumnado. La introducción a estos métodos inicia por identificar preguntas cuya solución requiere datos y reconocer la omnipresencia de la variabilidad, así como la necesidad de tomar decisiones en condiciones en las que usualmente no hay respuestas seguras o garantizadas.

En la primera unidad de la asignatura, con una duración de 28 horas, se abordan técnicas de estadística descriptiva para el análisis de la información que brindan los datos de una sola variable estadística. Además de construir e interpretar tablas y gráficas, y calcular medidas en contextos vinculados a alguna toma de decisiones, se busca que el alumnado sea capaz de usar distribuciones de frecuencias como herramientas de análisis global, con su tendencia y variabilidad, para investigar y comparar colecciones de datos.

La segunda unidad continúa el desarrollo de métodos de estadística descriptiva, en este caso centrándose en la relación entre dos variables estadísticas, ya sean numéricas o categóricas, dentro del contexto de una investigación o problema. Este tipo de análisis tiene muchas aplicaciones, pero también puede conducir a conclusiones equivocadas si se aplica de manera mecánica. El desarrollo del razonamiento covariacional se facilita mediante el uso de la computadora para analizar ejemplos en contextos diversos, para la construcción y análisis de tablas de contingencia y gráficas adecuadas, para determinar los modelos lineales que describan la tendencia de un diagrama de dispersión y evaluar la fuerza de la relación entre las variables aparejadas. Esta unidad tendrá una duración de 10 horas.

Las dos primeras unidades brindarán al alumnado una buena formación para analizar la información con una visión crítica y argumentar sus posiciones y conclusiones con base en datos. Para introducir la tercera unidad, el alumnado debe entender que los métodos inferenciales formales, que estudiará en la siguiente asignatura, se basan en ciertos conceptos y procedimientos probabilísticos por ser técnicas para deducir información sobre una población con base exclusivamente en lo que se observa en una muestra aleatoria simple. Esta introducción es necesaria porque los métodos probabilísticos que se van a abordar en la tercera unidad tienen una lógica distinta a lo visto hasta ese momento en la asignatura.

La última unidad, con una duración de 26 horas, brinda un acercamiento al estudio del azar y la probabilidad mediante el análisis de diversas experiencias aleatorias que permitan la comprensión de qué es la probabilidad de un evento y cómo se mide en ciertos casos mediante el enfoque clásico, o bien se aproxima mediante los enfoques frecuencial y subjetivo. Adicionalmente, se deducirán y aplicarán algunas reglas básicas para el cálculo de probabilidades reconociendo conceptos como los eventos mutuamente excluyentes, la probabilidad condicional y la independencia entre eventos.

El propósito general del programa de Estadística y Probabilidad I es proveer al alumnado de una cultura estadística que le permita organizar, comprender e interpretar información basada en datos, e iniciar el desarrollo de su razonamiento estadístico para evaluar la validez de la información y para diseñar e interpretar estudios estadísticos que permitan la construcción de conocimientos en diversos campos del saber.

Lograr este propósito requiere emplear herramientas tecnológicas para ordenar y presentar los datos, para calcular medidas y construir gráficas, analizar tendencias y correlaciones entre variables aparejadas, realizar simulaciones, así como para calcular probabilidades o estimarlas de manera frecuencial. El uso de la tecnología debe incorporarse en secuencias didácticas diseñadas para favorecer el descubrimiento y la indagación que conduzcan a la comprensión de cuestiones teóricas y a un análisis crítico de la información. Los propósitos específicos de la asignatura son los siguientes. El alumnado:

1. Analizará la información que brinda una colección de datos acerca de una característica de interés, a partir de su distribución, tendencia y variabilidad, y usando diversas representaciones, en el contexto de una toma de decisión o la resolución de un problema, para desarrollar su cultura y razonamiento estadísticos.
2. Analizará la relación entre dos variables estadísticas tanto cualitativas como cuantitativas y realizará predicciones válidas, a partir de la modelación de la tendencia de esa relación y midiendo su grado de intensidad, con la finalidad de interpretar y evaluar críticamente la información estadística en dos variables aparejadas.
3. Iniciará el estudio formal de la probabilidad mediante el tratamiento de experiencias aleatorias, aplicando los tres enfoques para la asignación de probabilidades a eventos y deduciendo las reglas básicas para el cálculo de probabilidades con la finalidad de desarrollar su razonamiento probabilístico.

UNIDAD 1. ANÁLISIS DE INFORMACIÓN ESTADÍSTICA PARA UNA VARIABLE

Presentación de la unidad

El contenido de esta unidad contempla conceptos y procedimientos de la estadística descriptiva para el análisis de la información que puede deducirse de una colección de datos de una variable estadística. Se trata de brindar al alumnado una formación fundamentalmente a nivel de cultura estadística que le permita comprender y utilizar la información en la toma de decisiones y en la resolución de problemas.

La unidad incluye cuatro temas. Para iniciar, el primero, que aborda un conjunto de nociones básicas preliminares, se propone diseñar actividades que permitan que el alumnado identifique preguntas cuya respuesta requiere datos, que accedan a diversas colecciones de datos de su entorno, que comprendan el papel del contexto en los estudios estadísticos y asimilen el concepto de variabilidad.

Es conveniente recolectar datos reales obtenidos por levantamiento, experimentación o acudiendo a bases de datos confiables, vinculando su análisis a alguna toma de decisión. Se puede iniciar planteando el problema o la pregunta a responder e invitar al alumnado a formular conjeturas que luego se corroborarán o descartarán mediante el análisis de los datos. La discusión grupal también puede aprovecharse para invitar al estudiantado a realizar inferencias informales.

Por ejemplo, para diseñar una campaña contra el tabaquismo en el plantel, se desea recabar información acerca de los hábitos de tabaquismo del alumnado. ¿Qué información se requiere?, ¿cómo se puede obtener?, ¿creen ustedes que haya diferencias por género en el consumo de tabaco? Si contamos con la información de tres o cuatro grupos de quinto semestre, ¿podríamos extender los resultados a lo que ocurre en todo el plantel?, ¿en qué basan esa conclusión?

Un levantamiento de datos útiles se puede realizar mediante una encuesta aplicada en los grupos de la materia con ayuda de alguna herramienta tecnológica, con preguntas como: ¿cuál es la carrera que deseas estudiar?, ¿cuántas materias adeudas?, ¿cuántos hermanos tienes?, ¿aproximadamente cuántos minutos te lleva trasladarte de tu casa a la escuela?, ¿cuál es tu deporte favorito?, ¿qué red social utilizas con mayor frecuencia?, ¿fumas, dejaste de fumar o nuca has fumado?, ¿en qué colonia vives?, y otras similares. La información recabada puede emplearse a lo largo de toda la unidad en varios temas y subtemas: por ejemplo, para organizar los datos en tablas, identificar las variables estadísticas, construir gráficas, formular observaciones sobre la información que brindan los datos y aplicar medidas adecuadas para representarlos.

El segundo tema es la representación tabular de los datos. La construcción de tablas, tanto por valores (de frecuencias simples) como por intervalos (de datos agrupados), puede abordarse en el aula con pocos datos y apoyarse en softwares especializados para su construcción cuando se analizan grandes colecciones de datos reales o realistas. Es fundamental que ninguna actividad termine con la mera construcción de una tabla de frecuencias, sino con la discusión sobre la información que brindan los datos ordenados, invitando al alumnado a formular sus observaciones.

El tercer tema se refiere a la representación gráfica. Se debe buscar que la explicación de cómo se construye cada tipo de gráfica no represente gran dificultad técnica, y apoyarse en herramientas computacionales o aplicaciones accesibles desde dispositivos móviles, para destinar la mayor parte del tiempo al análisis de la información representada. El uso de gráficas para comparar distintas colecciones de datos se puede acompañar de invitaciones al alumnado a que describa el comportamiento global de los datos.

Para finalizar el análisis exploratorio de datos, en la última parte el alumnado empleará medidas estadísticas que indiquen los valores representativos de una colección de datos y medidas que muestran qué tan dispersos se encuentran. El alumnado no sólo podrá calcular medidas estadísticas, sino que comprenderá la información que brinda cada una de ellas.

Es muy importante promover la observación de la tendencia y la variabilidad de cada distribución de frecuencias y usar esa distribución como herramienta para comparar comportamientos diversos en distintas colecciones. Asimismo, es importante diseñar actividades en las que se integren todos los elementos que se abordan en la unidad para el análisis de la información.

La presentación de la regla empírica resaltará su utilidad al brindar porcentajes aproximados de datos en cada rango, pero sin dejar de mencionar que solo es aplicable cuando la distribución de frecuencias es unimodal y aproximadamente simétrica.

Esta unidad permite abordar datos relacionados con los ejes transversales de los programas del CCH, por ejemplo, se pueden analizar datos sobre la evolución de la contaminación de los mares o los efectos del uso de combustibles fósiles en México y promover discusiones acerca de la sustentabilidad. En páginas especializadas se encuentran cuestionarios acerca de diversos tipos de violencia que pueden haber vivido desde la infancia tanto los hombres como las mujeres. Recoger y analizar este tipo de datos brinda un marco para reflexionar en torno a la normalización de la violencia de género y los estereotipos. En cuanto al eje de la educación ciudadana se pueden analizar datos relacionados con el acoso escolar o bullying u otras prácticas que se deben prevenir en las relaciones entre el alumnado. El conocimiento y aplicación de las tecnologías de la información y la comunicación se aborda en prácticamente todos los temas de la unidad. En las estrategias sugeridas se proponen temas sobre perspectiva de género como la muerte materna y el ciberacoso. También se sugieren temas vinculados a los ejes transversales en la actividad para la construcción de gráficas. La formación ciudadana se aborda al invitar al alumnado a identificar, cuestionar y valorar la discriminación y la desigualdad mediante el análisis de datos.

Con el tratamiento de estos temas y con actividades que permitan vincular los ejercicios trabajados con la toma de decisiones, será posible avanzar también en el desarrollo del razonamiento estadístico del alumnado en la medida en que se logre que incorporen las principales ideas estadísticas a su forma de enfrentar problemas que requieren datos en su vida cotidiana.

Carta descriptiva

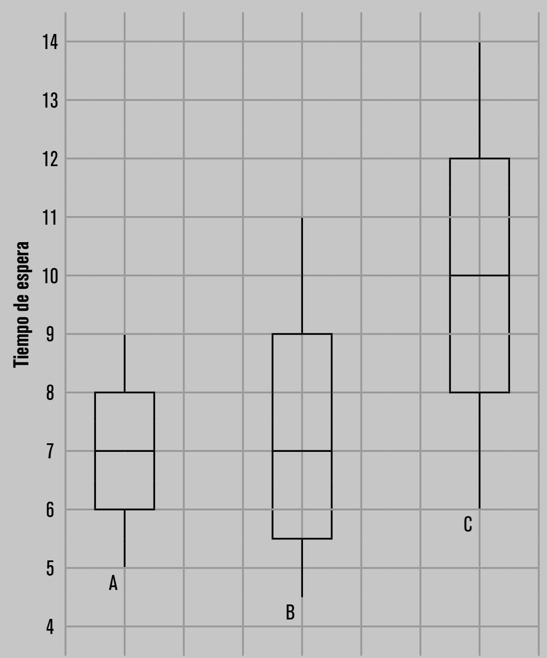
Propósito	Tiempo
Al finalizar la unidad, el alumnado: Analizará la información que brinda una colección de datos acerca de una característica de interés, a partir de su distribución, tendencia y variabilidad, y usando diversas representaciones, en el contexto de una toma de decisión o la resolución de un problema, para desarrollar su cultura y razonamiento estadístico	28 hrs.

Aprendizajes	Temática	Estrategias sugeridas
El alumnado: 1. Discute que la estadística estudia la variabilidad de las características de la población, considerando la homogeneidad y la heterogeneidad en los valores observados. 2. Explica las nociones de variable estadística, población y muestra. 3. Reconoce que los datos se obtienen por levantamiento o por experimentación. 4. 4. Examina la importancia de la recopilación y representación de datos en la investigación estadística. 5. Reconoce la importancia del muestreo. 6. Aplica algún procedimiento de selección aleatoria que le permita comparar una característica de una población conocida con la misma característica en muestras aleatorias. 7. Concluye que el azar es una de las causas de la variabilidad en los datos estadísticos.	**Nociones básicas** • Variable, población y muestra. • Variabilidad. • Investigación estadística. • Recolección de datos. • Muestreo, azar y probabilidad.	Se busca que el alumnado realice actividades que le permitan convivir con la variabilidad de los datos y con conceptos como población y muestras aleatorias, por ejemplo: El profesorado selecciona un artículo de la *Gaceta UNAM* que muestre los resultados de una investigación relevante para la sociedad, como el artículo “La muerte materna en México”, de la edición 541, del 18 de septiembre de 2023. Plantear preguntas del siguiente tipo: ¿cuál es la problemática?, ¿cuál es la población de estudio?, ¿de dónde se obtuvieron los datos?, ¿qué variables se han estudiado y cómo se han medido?, ¿en todos los países pasará igual?, ¿estos datos han cambiado a lo largo del tiempo?, ¿será necesario estudiar dicho cambio?, ¿en qué casos sí y en qué casos no? o ¿para qué puede ser necesario? Tratar de fomentar el interés por la investigación estadística y que el alumnado reconozca los conceptos clave que se emplean en el artículo y comprenda el significado de los resultados. Se construye una lista numerada en el aula donde el alumnado indique su género y la materia que le representa mayor dificultad en el semestre que está cursando. Se trabaja en equipos pequeños con esa información contestando preguntas como:

Aprendizajes	Temática	Estrategias sugeridas
		¿cuántas materias diferentes se eligieron en el grupo?, ¿cuáles son las dos materias que aparecen con más frecuencia?, ¿existen diferencias significativas por género?, ¿Serán diferentes las materias de mayor dificultad en otro grupo u otro plantel?, ¿qué factores pueden influir en que sean iguales o en que sean diferentes? Posteriormente, se usa cualquier generador de números aleatorios en el celular para que cada equipo seleccione una muestra aleatoria de 8 elementos y conteste preguntas como: ¿cuál es la materia que más se repite en la muestra?, ¿coincide con alguna de las dos materias que más se repiten en todo el grupo?, ¿en la muestra aparece alguna de esas dos materias? Se solicita a cada equipo que seleccione otra muestra aleatoria de 8 elementos y compare los resultados con la muestra anterior y con los de la población. Discutir las respuestas en plenaria y reunir la información de cuántas de las muestras seleccionadas en todo el grupo sí incluyeron las materias que más se repiten en la población y cuántas no.
8. Distingue los diferentes tipos de variables estadísticas. 9. Construye tablas de distribución de frecuencias, incorporando el uso de la computadora, para describir el comportamiento de una variable. 10. Formula conclusiones e inferencias informales con base en datos reales o realistas organizados en tablas de frecuencias, argumentando su validez y señalando posibles sesgos o limitaciones.	Representación tabular • Tipos de variables. • Tablas de distribución de frecuencias.	Para que el alumnado viva la necesidad de organizar la información y de identificar sus características más importantes, se pueden usar actividades como las siguientes: - Mediante un cuestionario en línea, se recogen datos sobre el alumnado que incluyan distintos tipos de variables (edad, género, estatura, peso, número de hermanos, tiempo de traslado a la escuela, hora de salida, hábitos de tabaquismo, promedio, carrera que desea estudiar y otros). Se promueve una discusión sobre los datos de alguna de las variables, tal cual fueron recabados, para motivar la reflexión, por ejemplo, sobre el promedio se puede preguntar: ¿qué tipo de variable es?, ¿entre qué cantidades está su valor?, ¿en qué rango están los promedios más frecuentes? ¿en qué rango están los menos frecuentes?, ¿qué porcentaje del alumnado tendrá un promedio mayor a 9? Posteriormente, construir una tabla de frecuencias con los datos de la encuesta y revisar las mismas preguntas para concluir que la organización de la información facilita su análisis. Usar la información reunida en los grupos para construir tablas de frecuencias con datos de distintos tipos de variables usando herramientas digitales. Cada tabla se acompañará de observaciones sobre cuestiones como el rango de valores, el valor más frecuente, el rango en el que se concentran más las frecuencias y otras.

Aprendizajes	Temática	Estrategias sugeridas
		Promover que el alumnado describa lo mejor posible la tendencia que muestran las distribuciones de frecuencias y comparen esa tendencia en distintas colecciones de datos, por ejemplo, entre hombres y mujeres o entre grupos diferentes. - Plantear una situación como la siguiente: El CCH pondrá en marcha un programa de intercambio estudiantil y será asignado de acuerdo con los promedios del alumnado. Si se quisiera que, al menos, un 5% del alumnado pueda beneficiarse del programa, ¿qué decisión se puede tomar sobre el promedio mínimo con la información obtenida?, ¿se puede asumir que en todo el colegio los promedios tendrán un comportamiento similar al observado en algunos grupos de la materia? Orientar la discusión para identificar los posibles sesgos en las inferencias informales.
11. Construye gráficas, incorporando el uso de herramientas tecnológicas, para interpretar la información que brindan los datos en el contexto de una investigación o problema.	Representación gráfica • Gráfica de barras. • Gráfica circular. • Gráfica de puntos. • Histograma de frecuencias. • Polígono de frecuencias. • Ojivas.	Se busca que el alumnado identifique diversas variables estadísticas, las grafique e interprete adecuadamente la información con ayuda de la tecnología, por ejemplo, con actividades como las siguientes. - Dividir al estudiantado en equipos de cuatro o cinco personas y pedirles que elijan un tema de interés como: participación en deportes, prácticas de acoso escolar en el CCH, cuidado de áreas verdes en el plantel, interés por la ciencia y tecnología, preferencias en las próximas elecciones, prácticas de acoso sexual y violencia de género, u otros. Cada grupo elabora una encuesta con cinco preguntas cerradas sobre su tema, que pueden tener respuestas numéricas o categóricas. Las encuestas se pasan entre el alumnado buscando recabar la mayor cantidad de respuestas posible. Cada equipo recoge las respuestas y las tabula en una hoja de cálculo. Eligen el tipo de gráfico más adecuado para representarlas y analizan las gráficas obtenidas. Presentan su gráfico al grupo, explicando el tema, las preguntas, la razón para elegir ese gráfico y el análisis de la información. - Solicitar al grupo, organizado en parejas que comparen varias gráficas del mismo tipo, por ejemplo, ojivas sobre fallecimientos por Covid-19 durante 2020-2022 en tres países diferentes, gráficas de puntos sobre la duración de focos de tres marcas distintas, gráficas de barras de calificaciones en Matemáticas IV de tres grupos diferentes, u otras. Promover un análisis de la información lo más completo posible y escribir una conclusión en cada caso.

Aprendizajes	Temática	Estrategias sugeridas
12. Calcula medidas de tendencia central, de dispersión y de posición con apoyo de herramientas tecnológicas para describir el comportamiento de una variable. 13. Interpreta los valores de las medidas de tendencia central y escoge la más adecuada para representar los datos en el contexto de una investigación o problema. 14. Interpreta los valores del rango, la desviación estándar y los cuartiles en el contexto de una investigación o problema. 15. Compara la variabilidad entre colecciones de datos con medias muy distintas por medio de sus coeficientes de variación.	Medidas estadísticas • Tendencia Central: Media, mediana y moda. • Dispersión: Rango, varianza y desviación estándar, coeficiente de variación. • Posición: Cuartiles y diagrama de caja, deciles y percentiles. • Regla empírica.	Para que el alumnado trabaje en el significado de cada una de las medidas estadísticas, se proponen actividades como: - Discutir qué es el ciberacoso y preguntar al grupo sus conjeturas sobre la gravedad del problema. Para analizar qué tan grave es el ciberacoso por edad y sexo de las personas afectadas, revisar los resultados de la encuesta de Inegi, presentados en el Módulo sobre Ciberacoso (MOCIBA) de 2022 o de años posteriores. (https://www.inegi.org.mx/programas/mociba/2022/#tabulados). Trabajar con los datos de las frecuencias absolutas de mujeres y hombres acosados y no acosados en la entidad donde se encuentra el plantel, que se presentan en las bases de datos "Población de hombres (o mujeres) de 12 años y más que utilizó internet o celular durante los últimos tres meses, por entidad federativa y grupos de edad según condición de haber vivido ciberacoso". Solicitar primero que se trabaje con la población de mujeres y realizar el histograma de frecuencias, calcular las medidas de tendencia central, dispersión y posición.
16. Formula inferencias informales con base en medidas estadísticas de datos reales o realistas, argumentando su validez y señalando posibles sesgos o limitaciones. 17. Discute la Regla empírica en distribuciones aproximadamente simétricas y unimodales en términos de relación entre tendencia y dispersión. 18. Analiza el comportamiento de diversas colecciones de datos a partir de su distribución, tendencia y dispersión para compararlas y tomar decisiones en el contexto de una investigación o problema.		Posteriormente, revisar los datos para la población de hombres y analizar las diferencias entre ambas distribuciones. Plantear preguntas para la reflexión como: ¿cuál es la edad promedio de hombres y de mujeres que sufrieron ciberacoso?, ¿en qué edad se presenta mayor ciberacoso en cada sexo?, ¿en qué edad es menos frecuente sufrir ciberacoso para hombres y mujeres? ¿cuál es el rango de edad del 50% más joven de las mujeres y hombres que sufrieron ciberacoso?, ¿cuál es el rango de edad del 50% mayor? y ¿qué edad tiene el 25% de la población más joven que sufre ciberacoso en cada sexo? Introducir algunas preguntas que involucren el uso del diagrama de caja y su interpretación. Propiciar la participación del alumnado al compartir el análisis del problema y conducir la discusión para llegar a alguna conclusión o toma de posición. - Plantear la siguiente situación: El alumnado del CCH mide el tiempo de espera en tres lugares de venta de comida en su plantel (A, B, C) para evitar retardos al entrar a su siguiente clase. Los datos se muestran en los siguientes diagramas de caja. Analice los datos y formule algunas preguntas, por ejemplo, ¿qué diferencias se observan entre los tiempos que brindan cada lugar?,

Aprendizajes	Temática	Estrategias sugeridas
		¿cuál lugar muestra el menor rango de espera?, ¿cuál lugar ofrece un rango mayor de tiempo de espera?, Si se considera que el alumnado sólo puede esperar 7 minutos, ¿cuál lugar ofrece mejor servicio?, ¿qué riesgos hay en el lugar B en comparación con el lugar A? 

Evaluación

Evaluación diagnóstica	Evaluación formativa	Evaluación sumativa
La evaluación diagnóstica consiste en explorar los conocimientos previos del alumnado. Para esta unidad se requiere el manejo de operaciones aritméticas básicas y jerarquía de operaciones, así como notación y lenguaje algebraico. Explorar también las nociones que tiene el alumnado sobre la estadística, qué estudia y para qué sirve, así como las gráficas y medidas que conoce.	La evaluación formativa permite evaluar los conocimientos que se van adquiriendo a lo largo del desarrollo de los temas y ajustar la planeación didáctica a sus resultados. En esta unidad se sugiere evaluar la habilidad del alumnado para usar tablas, gráficas y medidas como instrumentos para el análisis de la información que brinda una colección de datos.	La evaluación sumativa permite determinar el grado en el cual el alumnado logró los aprendizajes. En esta unidad se requiere observar si el alumnado es capaz de utilizar las diversas herramientas estadísticas en el análisis de una situación problemática, un caso o un proyecto. En particular, si logra ver la forma, tendencia y variabilidad de una distribución de frecuencias como instrumento para el análisis y la comparación.
Ejemplos	**Ejemplos**	**Ejemplos**
Se puede realizar a través de: • Una actividad introductoria que puede ser lúdica y que requiera realizar operaciones aritméticas e identificar algunos conceptos. • Un cuestionario oral o escrito, con preguntas sobre lenguaje algebraico e interpretación de promedios y porcentajes. • Una lluvia de ideas en torno a problemas que requieran datos y la utilidad de la estadística. • Una discusión guiada acerca de una situación que conduzca a una investigación estadística, analizando qué preguntas se deben responder, qué información se requiere y cómo recolectarla. • Elaboración de un organizador gráfico sobre las herramientas de estadística descriptiva que conocen	Se pueden aplicar métodos de evaluación como: • Basada en problemas. Para cada tipo de herramientas estadísticas, se plantea uno o más problemas en los que se use esa herramienta para arribar a una solución. Ejemplos de esto son las estrategias sugeridas. • Mediante estudios de caso. Se analiza toda la información de un caso y se van usando las herramientas estadísticas para formular conclusiones. Conviene acompañarlos con rúbricas o listas de cotejo. Por ejemplo, analizar la política gubernamental frente al ciberacoso y sus efectos o consecuencias. • Con base en una investigación. Cada equipo acuerda el tema de investigación al inicio de la unidad, define las preguntas a responder, la información que se requiere y la forma de recabarla. Se va revisando la aplicación por parte de cada equipo, de las diversas herramientas que se van abordando en clase.	Se puede basar en productos como: • Los problemas atacados durante el desarrollo de la unidad colectados en un portafolio de evidencias o la resolución de problemas nuevos en los que se deban usar las distintas herramientas. • La conclusión fundamentada de un estudio de caso desarrollado a lo largo de la unidad o la realización de un nuevo estudio de caso usando todo lo que se aprendió. • La conclusión fundamentada de una investigación desarrollada a lo largo de la unidad. En las opciones anteriores, será necesario precisar si se realizarán presentaciones o trabajos escritos y de qué tipo. • Un examen o cuestionario individual. Conviene tener cuidado de que los reactivos no se centren en los procedimientos aritméticos o técnicos sino en el análisis estadístico de diversas situaciones. Cualquiera de las opciones se puede acompañar de coevaluaciones y autoevaluaciones.

UNIDAD 2. DATOS BIVARIADOS

Presentación de la unidad

En esta unidad se introducirá al alumnado por primera vez al razonamiento sobre la asociación entre dos variables estadísticas, también conocido como razonamiento covariacional, donde se explorará cómo se relacionan estas dos variables a medida que experimentan cambios.

Este razonamiento se debe enfocar en la observación de dos características en cada elemento de la población o muestra. Las características observadas pueden corresponder a dos variables estadísticas cualitativas, dos cuantitativas, o una combinación de ambos tipos. Esta unidad aborda los dos primeros escenarios, es decir, el análisis de dos variables cualitativas y dos cuantitativas. En el estudio de datos bivariados, el punto central es indagar si existe una relación entre las dos características observadas.

Es importante señalar que, aunque el estudio formal del razonamiento covariacional solo puede desarrollarse mediante el uso de inferencias estadísticas, en esta unidad se ofrece un primer acercamiento del alumnado a este tipo de razonamiento desde la perspectiva del análisis de la información que brinda una colección de datos bivariados sobre la relación entre dos variables estadísticas aparejadas, mediante la creación e interpretación de tablas, gráficos y medidas dentro del contexto de una investigación o problema, apoyándose con el uso de la computadora. En el análisis se incorporan inferencias informales sobre la extensión de la relación observada en los datos de una muestra bivariada a toda una población.

La unidad consta de tres temas. En el primero de ellos se busca familiarizar al alumnado con los datos bivariados de características aparejadas de distintos tipos. El segundo tema consiste en el estudio de la relación entre dos variables estadísticas categóricas mediante la construcción de tablas de contingencia, el cálculo de frecuencias relativas y el análisis de los porcentajes por renglón y por columna en esas tablas. Este análisis de la información se apoyará en la construcción de gráficas adecuadas y culminará con la formulación de inferencias informales, en contextos que resulten de interés para el alumnado.

El tercer tema es la relación entre dos variables numéricas. Se busca que el alumnado desarrolle la habilidad de reconocer cuándo una relación entre dos variables de este tipo es razonablemente lineal. Lo anterior se aborda mediante la construcción e interpretación del diagrama de dispersión, así como el cálculo e interpretación del coeficiente de correlación lineal de Pearson, que actúa como medida de la fuerza y dirección de una relación lineal. Posteriormente, se les involucra en la construcción e interpretación de un modelo matemático lineal que refleje la tendencia en la relación, utilizando el criterio de mínimos cuadrados para construir la recta de mejor ajuste.

El enfoque didáctico de la materia conduce a centrar el trabajo del alumnado en la formulación de observaciones, primero como conjeturas acerca de la relación entre dos variables y posteriormente como conclusiones del análisis de la información. La deducción de la ecuación de la recta de mínimos cuadrados tiene un fundamento teórico que trasciende al nivel del alumnado. Por ello, es importante acudir al uso de softwares especializados o simulaciones interactivas (por ejemplo, https://phet.colorado.edu/sims/html/least-squares-regression/latest/least-squares-regression_all.html?locale=es) para que el alumnado visualice en qué consiste ese criterio y pueda explorar cómo cambia la recta al mover algunos puntos del diagrama de dispersión. Entender de dónde vienen las fórmulas o por qué funcionan, le da sentido a su uso e interpretación, sin perder de vista que la parte técnica no debe ser el centro del trabajo del alumnado.

En cuanto a los ejes transversales, en las estrategias sugeridas se incluyen ligas a varias encuestas nacionales y un informe país, que brindan información que podemos usar en las aulas al desarrollar los temas de esta unidad. Por ejemplo, en la presentación de resultados de la Encuesta Nacional sobre Discriminación se presentan varias gráficas de barras con divisiones porcentuales sobre percepción del respeto a diversos grupos, reconocimiento de derechos y aceptación de medidas para la igualdad, y sobre prejuicios, estigmas sociales y estereotipos, entre otros temas. Cada una de estas colecciones de barras se puede usar en una actividad donde el alumnado interprete la información que brindan esas gráficas y redacte un pequeño texto con su reflexión sobre lo que esa información significa en la formación ciudadana de los mexicanos.

Carta descriptiva

Propósito	Tiempo
Al finalizar la unidad, el alumnado: Analizará la relación entre dos variables estadísticas tanto cualitativas como cuantitativas y realizará predicciones válidas, a partir de la modelación de la tendencia de esa relación y midiendo su grado de intensidad, con la finalidad de interpretar y evaluar críticamente la información estadística en dos variables aparejadas.	10 hrs.

Aprendizajes	Temática	Estrategias sugeridas
El alumnado: 1. Reconoce que los datos bivariados se obtienen al observar dos características en cada elemento de la población. 2. Distingue que, entre dos variables estadísticas representadas en datos bivariados, puede existir alguna relación.	Correlación entre dos variables estadísticas • Ejemplos de covariación de variables de diversos tipos.	Para que el alumnado se familiarice con los datos bivariados en contextos cercanos se pueden desarrollar actividades como la siguiente: - Usar la encuesta realizada en los grupos de la materia, y solicitar al alumnado organizado en equipos pequeños que seleccione dos variables cualitativas contempladas en la encuesta y escriba cinco ejemplos de datos bivariados.

Aprendizajes	Temática	Estrategias sugeridas
		Preguntar si consideran que esas variables están relacionadas y argumentar por qué. Hacer lo mismo para dos variables cuantitativas y para una variable de cada tipo. Después solicitar que elijan dos variables cualitativas que crean que sí están relacionadas, lo mismo para dos cuantitativas y para una de cada tipo. Siempre solicitar que argumenten en qué basan la creencia de que están relacionadas. Cerrar la actividad con una discusión grupal en la que se fomentará el intercambio de argumentos entre los equipos favoreciendo que en los casos polémicos se contrasten puntos de vista a favor y en contra. Al terminar esta discusión grupal el profesor o profesora puede explicar el contenido de esta unidad.
3. Construye tablas de contingencia, barras múltiples o barras con subdivisiones para presentar e interpretar la información correspondiente a dos variables estadísticas cualitativas aparejadas. 4. Examina la información vertida en una tabla de contingencia, en términos de la relación entre dos variables estadísticas cualitativas, dentro del contexto de una investigación o problema. 5. Formula inferencias informales acerca del comportamiento de dos variables estadísticas cualitativas en el contexto de una investigación o problema.	Relación entre dos variables estadísticas cualitativas. • Tablas de contingencia y su interpretación. • Gráficos y su interpretación.	- Con base en un ejemplo que puede ser ficticio, analizar en el aula la interpretación correcta de las frecuencias conjuntas y las frecuencias marginales de una tabla de contingencia, cuidando que en la descripción de las conjuntas se use la conjunción "y". - Tomar dos variables cualitativas de la encuesta que pueden estar relacionadas, por ejemplo, género y hábitos de tabaquismo, género y carrera elegida, o cualquier otra pareja que el profesor o profesora considere pertinente. Iniciar la discusión con una pregunta detonadora como ¿Existe alguna relación entre estas variables estadísticas? Anotar las conjeturas iniciales y construir tablas de contingencia con ayuda de un software especializado y usando los datos de la encuesta (por ejemplo, puede usarse CODAP, EXCEL, Fathom o GeoGebra). Construir también las tablas de frecuencias relativas en porcentajes y de porcentajes por renglón y por columna. Para cada una de estas tablas, solicitar que el alumnado, organizado en equipos pequeños, formule observaciones. Es muy importante que la formulación considere cuál es el total del que se calculan los porcentajes en cada caso, por ejemplo, que se sepa diferenciar entre el porcentaje de encuestados que son hombres y fuman, el porcentaje de las personas que fuman del total de hombres y el porcentaje de hombres en el total de quienes fuman. Construir también gráficos de distintos tipos. Una vez que se llegue a una conclusión sobre la existencia o no de una relación entre las variables, y se contraste con las conjeturas iniciales, cerrar la actividad preguntando en discusión grupal si la conclusión puede extenderse a todo el alumnado del plantel, y solicitar que argumenten en qué basan esa conclusión.

Aprendizajes	Temática	Estrategias sugeridas
		- Para usar datos de bases confiables, el profesorado puede seleccionar diversas parejas de variables cualitativas provenientes de encuestas nacionales como las siguientes: Encuesta Nacional de Salud y Nutrición Informe país sobre la calidad de la ciudadanía en México Encuesta Nacional de Hábitos de Reciclaje de Plásticos Encuesta Nacional sobre discriminación - Con la información que se elija, diseñar una actividad de interpretación de información.
6. Construye diagramas de dispersión para describir el comportamiento de dos variables estadísticas cuantitativas aparejadas. 7. Examina la información vertida en un diagrama de dispersión en el contexto de una investigación o problema. 8. Formula inferencias informales acerca del comportamiento de dos variables estadísticas cuantitativas aparejadas en el contexto de una investigación o problema. 9. Interpreta el valor del coeficiente de correlación lineal al ajustar los puntos del diagrama de dispersión desde una relación aproximadamente lineal hasta datos no relacionados, utilizando una aplicación interactiva o software especializado. 10. Descubre que la recta de mínimos cuadrados es la que mejor modela la correlación entre dos variables estadísticas cuantitativas, cuando ésta se presenta de manera aproximadamente lineal. 11. Analiza los parámetros de la recta de mejor ajuste interpretándolos dentro del contexto de una investigación o problema.	Relación entre dos variables estadísticas cuantitativas. • Diagrama de dispersión. • Coeficiente de correlación (Pearson). • Regresión lineal.	Con el fin de familiarizar al alumnado con los diagramas de dispersión, se pueden aplicar actividades como las siguientes. • Con ayuda de un software especializado, construir diagramas de dispersión usando parejas de variables de la encuesta en los grupos de la materia, por ejemplo, tiempo de traslado de la casa a la escuela y la hora de salida, estatura y peso, o el largo del brazo y la longitud del palmo de la mano. Antes de construir los diagramas de dispersión, solicitar que se formulen conjeturas sobre el grado de linealidad de la relación, entre nula, débil, moderada y fuerte. Una vez construido, contrastar las conjeturas con lo que se observa. Buscar también ejemplos de relaciones con pendiente negativa como tiempo de uso de dispositivos electrónicos antes de dormir y horas de sueño, tiempo dedicado a actividades de ocio y tiempo de estudio. Para introducir el coeficiente de correlación como medida de la intensidad de una relación lineal, se sugieren actividades como las siguientes: • Usando un software, presentar al alumnado una serie de diagramas de dispersión que representen diferentes niveles de correlación, por ejemplo, un diagrama con puntos muy dispersos, otro con puntos muy cercanos a una línea recta y varios diagramas intermedios, algunos con tendencia lineal de pendiente positiva y otros de pendiente negativa. Solicitar que en parejas ordenen los diagramas desde lo que consideren la relación lineal más débil hasta la que consideren más fuerte, en dos colecciones dependiendo de la pendiente. En discusión grupal llegar a un acuerdo acerca de la clasificación. • Usar el software para calcular el coeficiente de correlación de los diagramas de dispersión analizados, y formular preguntas como:

Aprendizajes	Temática	Estrategias sugeridas
12. Aproxima, considerando las limitaciones del dominio y el valor del coeficiente de correlación, el valor de una variable regresora de un valor de la variable de respuesta, por medio de la recta de mejor ajuste.		¿qué valores puede tomar el coeficiente de correlación?, ¿a qué número se acercan los coeficientes de diagramas de dispersión que están muy cercanos a una recta de pendiente positiva?, ¿a qué número se acercan los que corresponden a diagramas muy cercanos a una recta de pendiente negativa? ¿a qué número se acercan mientras más lejana está la nube de puntos de una recta? Concluir la actividad cambiando un punto en un diagrama de dispersión con coeficiente de correlación cercano a 1 o a -1, para que esté cada vez más lejos de la nube de puntos. Pedir al alumnado que describan lo que ocurre con el valor del coeficiente de correlación. Para que el alumnado visualice la construcción de la recta de mejor ajuste por mínimos cuadrados, se pueden desarrollar actividades como las siguientes: • En un diagrama de dispersión en hojas de papel, solicitar al alumnado organizado en parejas que trace la recta que a su juicio es la que mejor se ajusta al diagrama, definiendo de antemano cuál será el criterio que se usará. Socializar las distintas opiniones acerca de cómo trazar la recta. Conducir la discusión hacia la necesidad de que el criterio contemple todos los puntos del diagrama, no solo algunos de ellos. • Usar un software especializado o simulación interactiva para colocar una recta móvil sobre un diagrama de dispersión y visualizar el cuadrado de las distancias verticales a todos los puntos (por ejemplo, en el sitio https://phet.colorado.edu/sims/html/least-squares-regression/latest/least-squares-regression_all.html?locale=es). Solicitar al alumnado que mueva la recta hasta que la suma de esos cuadrados sea lo más chica posible. Posteriormente, con ayuda del software o simulación, trazar la recta de mínimos cuadrados para que la comparen con la que construyeron manualmente. • Se propone la división del alumnado en equipos de cuatro personas. Cada equipo se dedicará a recopilar diversas medidas, por ejemplo, los datos referentes al peso corporal y la presión sistólica, o un par de medidas antropométricas (estatura, ancho de brazos o piernas, circunferencia del cráneo, apertura de manos, etcétera) u otra pareja de variables cuantitativas asociadas.

Aprendizajes	Temática	Estrategias sugeridas
		Trazar el diagrama de dispersión, calcular el coeficiente de correlación y trazar la gráfica de la recta de mejor ajuste. Analizar las tres herramientas y escribir una conclusión acerca de la linealidad de la relación entre las variables. Formular inferencias informales acerca de la posible extensión de la conclusión a todo el alumnado del colegio.

Evaluación

Evaluación diagnóstica	Evaluación formativa	Evaluación sumativa
Para diagnosticar el punto de partida en el desarrollo de esta unidad, considerar los siguientes conocimientos: operaciones con números racionales y cálculo de porcentajes, jerarquía de operaciones, conceptos de correlación y función, ecuación de una recta en la forma pendiente-ordenada al origen, interpretación de los parámetros y gráficas de rectas, tipos de variables estadísticas y gráficas de barras.	Se espera que a lo largo de esta unidad el alumnado aprenda a reconocer datos bivariados, identificar las dos variables asociadas a esos datos, formular conjeturas sobre la relación entre las variables y utilizar herramientas de análisis, como tablas de contingencia, gráficos, diagramas de dispersión, coeficiente de correlación lineal, entre otros, para determinar la validez de sus conjeturas.	Al concluir la unidad, será necesario evaluar los alcances del alumnado en la identificación de datos bivariados, análisis de la información que brindan las tablas de contingencia y gráficos en torno a una posible relación entre dos variables cualitativas; análisis de la información a partir de un diagrama de dispersión, interpretación del coeficiente de correlación lineal y la construcción de la recta de mejor ajuste para dos variables cuantitativas.
Ejemplos	**Ejemplos**	**Ejemplos**
Se puede realizar a través de: • Un cuestionario con ejercicios de operaciones aritméticas. • Lluvia de ideas acerca de conceptos como correlación y función. • En discusión grupal preguntar qué hacer para graficar la recta que corresponde a una ecuación de la forma $y = mx + b$ y la determinación de la ecuación que corresponde a una gráfica, así como la interpretación de los parámetros en situaciones contextualizadas.	Se pueden aplicar métodos de evaluación como: • Basada en problemas. Ejemplos de esto son las estrategias sugeridas además de actividades en las que el alumnado realice mediciones antropométricas en clase para verificar conjeturas sobre posibles relaciones. También se pueden pedir trabajos en casa, individuales o en equipos, para contestar preguntas con la información de bases de datos.	Se puede basar en productos como: • Los problemas atacados durante el desarrollo de la unidad colectados en un portafolio de evidencias o la resolución de uno o más problemas nuevos en los que se deban usar las distintas herramientas. • La conclusión fundamentada de un estudio de caso desarrollado a lo largo de la unidad o la realización de un nuevo estudio de caso usando todo lo que se aprendió.

Evaluación diagnóstica	Evaluación formativa	Evaluación sumativa
• En el desarrollo de la unidad, motivar la recuperación de conceptos y procedimientos abordados en la unidad anterior y corregir el lenguaje.	• Mediante estudios de caso. Por ejemplo, estudios sobre el seguimiento de las prácticas de discriminación y su relación con la condición étnica o la situación de migración de los afectados. Conviene acompañarlos con rúbricas o listas de cotejo. • Con base en una investigación. Se pueden implementar proyectos pequeños, por ejemplo, realizar una encuesta de 4 o 5 preguntas entre sus compañeros de las distintas clases acerca de la posible relación entre dos variables cualitativas y dos cuantitativas como género y hábitos de lectura o promedio y horas de sueño u otras. Conviene acompañarlos con rúbricas o listas de cotejo.	• La conclusión fundamentada de una investigación desarrollada a lo largo de la unidad. En las opciones anteriores, será necesario precisar si se realizarán presentaciones o trabajos escritos y de qué tipo. • Un examen o cuestionario individual. Conviene tener cuidado de que los reactivos no se centren en los procedimientos aritméticos o técnicos sino en el análisis estadístico de diversas situaciones. Cualquiera de las opciones se puede acompañar de coevaluaciones y autoevaluaciones.

UNIDAD 3. AZAR Y PROBABILIDAD

Presentación de la unidad

En la presente unidad el alumnado iniciará el estudio de los conceptos de la Teoría de Probabilidad que le serán útiles para comprender las técnicas de inferencia estadística que abordará el siguiente semestre. Incluye cuatro temas.

Se propone introducir el primer tema, sobre el objeto de estudio de la probabilidad, presentando situaciones en las que se vincule la probabilidad con la frecuencia de ocurrencia de algún resultado de interés cuando se hace un gran número de repeticiones, para después abordar la caracterización de los experimentos aleatorios y experimentos deterministas, la determinación del espacio muestral de un experimento aleatorio y la identificación de los eventos con subconjuntos del espacio muestral.

En el segundo tema se abordará la asignación del valor de la probabilidad a un evento mediante tres enfoques: el clásico, el frecuencial y el subjetivo. Es importante que el alumnado comprenda que a cada evento solo se le puede asignar una probabilidad, que no es otra cosa que una medida de la facilidad con la que ocurre al realizar el experimento. Cuando todos los elementos del espacio muestral tienen la misma probabilidad de ocurrir, ese valor se puede obtener mediante el conteo usando el enfoque clásico. Se propone incluir algún ejemplo en el que sea natural considerar dos formas de describir el espacio muestral y analizar con el alumnado cuál descripción conduce a un espacio muestral equiprobable antes de aplicar el enfoque clásico.

La probabilidad de un evento dice poco acerca de lo que puede ocurrir en la siguiente realización del experimento aleatorio; su valor adquiere significado cuando se repite el experimento un gran número de veces. El enfoque frecuencial evidencia este significado de la probabilidad de un evento y proporciona una aproximación a su valor numérico mediante una frecuencia relativa obtenida tras una gran cantidad de repeticiones, sin importar si los resultados son equiprobables o no. Para que el alumnado descubra que las frecuencias relativas de cualquier evento tienden a estabilizarse alrededor de un número al ir aumentando la cantidad de repeticiones del experimento, hecho al que se conoce como regularidad estadística, conviene usar simulaciones computacionales de eventos sencillos en los que se conozca la probabilidad por la aplicación del enfoque clásico. Las simulaciones digitales se pueden usar también para aproximar probabilidades de eventos complejos en los que al alumnado no le sea sencillo usar el enfoque clásico.

Por último, el enfoque subjetivo tiene aplicaciones en dos contextos: cuando no se puede aplicar ninguno de los enfoques anteriores, en cuyo caso debe basarse en la opinión de expertos, y cuando se formula una conjetura inicial sobre una probabilidad para luego ir ajustando ese valor con base en la recolección de nueva información. Una de las aplicaciones cotidianas de este enfoque se presenta en los informes meteorológicos. La información obtenida mediante satélites acerca de los fenómenos climáticos se acompaña por la experiencia de los expertos para brindar probabilidades de ocurrencia de lluvias, vientos o el arribo de un huracán en el futuro inmediato.

Conviene iniciar el tercer tema, sobre el cálculo de probabilidades, escribiendo los elementos del espacio muestral de cada experimento e identificando los resultados favorables para diversos eventos. Posteriormente, podrán calcularse determinando solo la cantidad de elementos tanto del espacio muestral como de los eventos. Esto requiere el uso de algunas técnicas de conteo.

El fundamento de las reglas para el cálculo de probabilidades (complemento, unión y regla de la suma) se puede mostrar usando el enfoque clásico y apoyándose en diagramas de Venn o en tablas de contingencia, evitando simplemente asumirlas.

En el último tema es importante deducir la regla para el cálculo de probabilidades condicionales mostrando que se trata de una reducción del espacio muestral como efecto del conocimiento de alguna información parcial.

Cuando la ocurrencia o no ocurrencia de un evento no altera la probabilidad de que ocurra otro, los eventos son independientes. Esta caracterización se expresa en términos de probabilidad condicional mediante la igualdad $P(A|B) = P(A)$, o bien $P(B|A) = P(B)$, de la que se deduce la regla del producto. Este concepto se puede abordar mediante la presentación de ejemplos de eventos independientes y eventos dependientes. Por ejemplo, en el lanzamiento de dos monedas comunes se pueden analizar eventos como A: se obtiene águila en el primer lanzamiento, B: se obtienen dos caras iguales y C: se obtienen dos águilas. Comparando probabilidades condicionales o mediante la regla del producto el alumnado podrá verificar que A y B son independientes, pero no lo son A y C ni B y C. Es importante considerar que diversos estudios en educación estadística evidencian que con frecuencia se conoce la regla del producto, pero no se comprende la independencia.

En cuanto a los ejes transversales, en las estrategias sugeridas se presenta un ejemplo que puede usarse para motivar una discusión grupal acerca del valor que se le da al nacimiento de una niña y al nacimiento de un niño en una política de control de natalidad. Aunque es posible encontrar otros ejemplos en los que se contemplen ejes como el de sustentabilidad y el de formación ciudadana, en esta unidad resalta especialmente el eje transversal de *Conocimiento y aplicación de las Tecnologías de la Información y la Comunicación,* que se verá reflejado en cada uno de los temas.

Carta descriptiva

Propósito	Tiempo
Al finalizar la unidad, el alumnado: Iniciará el estudio formal de la probabilidad mediante el tratamiento de experiencias aleatorias, aplicando los tres enfoques para la asignación de probabilidades a eventos y deduciendo las reglas básicas para el cálculo de probabilidades con la finalidad de desarrollar su razonamiento probabilístico.	26 hrs.

Aprendizajes	Temática	Estrategias sugeridas
El alumnado: 1. Reconoce la relación entre la probabilidad y la cantidad de ocurrencias de un resultado en un gran número de repeticiones. 2. Comprende las características de experimentos aleatorios y experimentos deterministas. 3. Identifica a la probabilidad como la medida de la posibilidad de ocurrencia de un resultado en un experimento aleatorio. 4. Define los conceptos de espacio muestral y evento.	La probabilidad y su objeto de estudio 1. Probabilidad y cantidad de ocurrencias de un resultado en un gran número de repeticiones. • Experimentos aleatorios y deterministas. • Probabilidad como medida de la ocurrencia. • Espacio muestral y eventos.	Se busca que el alumnado analice alguna situación en la que se relacione la probabilidad con la cantidad de ocurrencias en un gran número de repeticiones, por ejemplo: • Una política de control de natalidad consiste en que cada familia puede tener un solo hijo hombre y todas las hijas que nazcan antes del primer hombre. Recoger conjeturas acerca de las posibles repercusiones de esta política en la cantidad de hijos e hijas de cada familia, y en la cantidad de hombres y mujeres en la población. Asumiendo que es igualmente probable que nazca una niña o un niño, cada nacimiento se puede simular mediante el lanzamiento de una moneda, considerando, por ejemplo, que águila representa niña y sol niño. Organizar al grupo en parejas y solicitar que cada pareja simule lo que ocurre en 50 familias, lanzando la moneda hasta que salga el primer sol. Se registrará la cantidad de descendientes y la de niñas en cada familia y el promedio de esas cantidades en las 50 familias. El análisis se puede completar haciendo la simulación en computadora o reuniendo las cantidades obtenidas por todas las parejas para calcular un promedio de varios cientos de familias. Invitar al alumnado a contrastar sus conjeturas iniciales con los resultados obtenidos y señalar la relación entre la probabilidad del nacimiento de cada sexo y el número promedio de hijos e hijas por familia. Para discutir la diferencia entre experimentos aleatorios y deterministas se pueden realizar actividades como: • Se pide a algún miembro del estudiantado que salga del salón, se coloca una bola negra dentro de la “caja 1” y una bola blanca dentro de la “caja 2”. Regresa quien salió y se le pide que extraiga la bola de la “caja 1” e indique su color.

Aprendizajes	Temática	Estrategias sugeridas
		Preguntar al alumnado si se trata de un experimento aleatorio o determinista, pidiendo que fundamenten su respuesta. Una vez que en la discusión quede claro que cada vez que ese experimento se repite en las mismas condiciones produce el mismo resultado, preguntar todas las formas posibles para convertirlo en aleatorio.
5. Comprende el enfoque clásico para el cálculo de probabilidades cuando los resultados son igualmente probables y lo aplica en experimentos sencillos. 6. Observa la tendencia a estabilizarse de las frecuencias relativas de un evento al realizar un gran número de repeticiones mediante simulaciones físicas y por computadora. 7. Aproxima la probabilidad de ocurrencia de algún resultado de un experimento aleatorio a partir de observaciones experimentales o de simulaciones físicas y por computadora. 8. Identifica condiciones en las que puede ser necesario acudir a una aproximación subjetiva de la probabilidad.	Enfoques de la probabilidad • Enfoque clásico de la probabilidad. • Enfoque frecuencial de la probabilidad. • Enfoque subjetivo de la probabilidad	Plantear alguna situación en la que resulte natural medir la posibilidad de ocurrencia de un resultado con la fracción que indica qué parte del total de resultados son favorables a un evento (enfoque clásico), por ejemplo: • En un juego, se tiene que elegir una bola al azar de una urna y se gana si sale blanca. Hay tres urnas con bolas idénticas salvo por el color. Una tiene 3 bolas negras y 4 blancas, otra tiene una bola negra y 2 blancas y la otra tiene 4 bolas negras y 5 blancas. Preguntar cuál escogerían para hacer la selección y por qué. Hacer notar que la fracción indicada pierde su sentido si hay algún resultado más probable que otro e introducir la formulación del enfoque clásico para espacios muestrales equiprobables. Para contribuir a la identificación de espacios muestrales equiprobables y no equiprobables, se puede plantear una situación como la siguiente: • Al lanzar dos monedas comunes idénticas, ¿en cuál de los siguientes espacios muestrales todos los resultados tienen la misma probabilidad de ocurrir? Ω_1 = {dos águilas, dos soles, un águila y un sol}, Ω_2 = {águila-águila, águila-sol, sol-águila, sol-sol}. Escuchar los argumentos del alumnado por cada opción. Organizar al grupo en parejas para que cada una repita el lanzamiento 50 veces y registre cuántas veces obtuvo cada resultado. El análisis se puede completar haciendo la simulación en computadora o reuniendo en el pizarrón las cantidades obtenidas por todas las parejas. Concluir que “un águila y un sol” es más probable que “dos águilas” y que “dos soles” por lo que Ω_1 no es equiprobable. Se busca fundamentar el enfoque frecuencial analizando cómo van variando las frecuencias relativas de un resultado del que se conoce la probabilidad en un experimento aleatorio sencillo, con apoyo computacional.

<table>
<tr><th>Aprendizajes</th><th>Temática</th><th>Estrategias sugeridas</th></tr>
<tr><td></td><td></td><td>

- En algún software especializado o simulación interactiva, obtener la trayectoria de las frecuencias relativas de un resultado al ir aumentando la cantidad de repeticiones de un experimento aleatorio (por ejemplo, www.stapplet.com/largenum.html). Repetir varias veces la simulación de al menos 500 realizaciones para observar diversas trayectorias. Invitar al alumnado a describir lo mejor posible el comportamiento que se observa. Resaltar que las frecuencias relativas tienden a estabilizarse alrededor de la probabilidad conocida, pero no de forma regular ni siempre igual.

</td></tr>
<tr><td>

9. Construye espacios muestrales, eventos simples, y eventos compuestos aplicando operaciones entre conjuntos: unión, intersección y complemento.
10. Calcula probabilidades de eventos simples y compuestos mediante conteo directo y construyendo tablas de contingencia para representar las relaciones entre dos eventos.
11. Explica el concepto de mutua exclusividad entre eventos y aplica la regla de la suma.

</td><td>

Cálculo de la probabilidad de eventos simples y compuestos

- Cálculo de probabilidades mediante conteo directo.
- Deducción de las reglas para la probabilidad del complemento y de la unión de eventos arbitrarios.
- Eventos mutuamente excluyentes y regla de la suma.

</td><td>

Para deducir las reglas básicas de la probabilidad, conviene usar un experimento aleatorio cuyos resultados se puedan recoger en una tabla de contingencia. Por ejemplo:

- En un grupo de 300 estudiantes de bachillerato se recogió la siguiente información sobre hábitos de tabaquismo:

	Nunca ha fumado	Dejó de fumar	Fuma	Total
Hombre	85	12	43	140
Mujer	103	17	40	160
Total	188	29	83	300

Se selecciona un estudiante de ese grupo al azar. Identificar eventos como: H: es hombre, M: es mujer, N: nunca ha fumado, D: dejó de fumar, F: fuma.

Calcular las probabilidades de los eventos anteriores y construir nuevos eventos como F^c, $M \cap N$ y $H \cap D$. Invitar al alumnado a buscar relaciones con las probabilidades de los eventos simples.

Las operaciones se pueden describir en términos de ocurrencia de eventos así: A^c: no ocurre A, $A \cap B$: ocurre uno de los eventos o ambos, y $A \cap B$: ocurren ambos eventos simultáneamente.

Se busca que el alumnado analice alguna situación en la cual se pueda construir el espacio muestral pero que dicho trabajo no resulte tan simple, por ejemplo:

- Tres alumnas y tres alumnos se organizan en equipo para preparar una exposición en su clase de filosofía, tres deben preparar los materiales y los otros tres deberán exponer.

</td></tr>
</table>

Aprendizajes	Temática	Estrategias sugeridas
		Deciden realizar la repartición al azar. Definir los siguientes eventos A: expondrán tres mujeres, B: al menos una mujer, C: expondrá al menos un hombre, y D: expondrán tres hombres. Pedir al alumnado que construya el espacio muestral en discusión grupal hasta incluir los 20 elementos que lo componen. Obtener las probabilidades siguientes: P(A), P(B), P(C), P(D), $P(A \cap D)$, $P(B \cap C)$, $P(A \cup D)$ y $P(A^c)$. Se pretende además que el alumnado desarrolle el concepto de mutua exclusividad y deduzca la regla de la suma. Se puede preguntarles: ¿qué significa que la probabilidad de la intersección de dos eventos resulte cero?, ¿cómo se relaciona la probabilidad de dos eventos mutuamente excluyentes con la probabilidad de su unión?
12. Construye la expresión para el cálculo de la probabilidad condicional entre dos eventos, a partir de una tabla de contingencia o usando diagramas de Venn. 13. Calcula probabilidades condicionales utilizando la expresión correspondiente. 14. Calcula probabilidades conjuntas. 15. Reconoce el concepto de independencia y deduce la regla del producto.	Probabilidad condicional y de eventos independientes. • Reducción del espacio muestral cuando se conoce información parcial acerca de los resultados. • Regla para el cálculo de probabilidades condicionales y conjuntas. • Eventos independientes.	Para presentar la probabilidad condicional se puede usar una situación en la que el conocimiento de alguna información parcial sobre el resultado induce una reducción del espacio muestral. Un ejemplo típico es: • Dos amigos A y B se encuentran después de muchos años. A indica que tiene dos descendientes sin especificar su sexo. Días después, B llega de visita a casa de A y le abre una niña que dice: “mi papá lo está esperando”. Preguntar al alumnado cuál es la probabilidad de que el otro hijo de A sea hombre y en qué se fundamenta esa respuesta. Usualmente, responden 50% porque puede ser hombre o mujer. Analizar la situación paso a paso. En un principio, sin ninguna información, los resultados posibles para el sexo de los dos descendientes son {hh, hm, mh, mm}. Asumiendo que es igualmente probable que nazca un hombre (h) a una mujer (m), la probabilidad de que los descendientes sean un hombre y una mujer es 2/4 = ½. Cuando ya se sabe que una es mujer, ya no puede ocurrir hh y los resultados posibles se reducen a {hm, mh, mm} de donde la probabilidad que se busca es 2/3. Identificar los eventos A: hay un hombre y una mujer, B: al menos una es mujer, identificar la última probabilidad obtenida con $P(A\|B)$ y presentar la probabilidad condicional.

Aprendizajes	Temática	Estrategias sugeridas
		Para introducir la independencia entre dos eventos se propone analizar en clase dos situaciones como las siguientes: • Una urna contiene 2 bolas negras y 3 blancas, se extraerán dos bolas al azar y sin reemplazo. Definir los eventos: A: la primera bola extraída es negra, y B: la segunda bola extraída es blanca. • Dos urnas contienen cada una 2 bolas negras y 3 blancas, se extraerá una bola de cada urna. Definir los eventos: A: la bola extraída de la primera urna es negra, y B: la bola extraída de la segunda urna es blanca. • Discutir las diferencias entre ambas situaciones y analizar si la ocurrencia de A en cada situación afecta a la probabilidad de B y las razones de esto. Observar que en la primera situación $P(B\|A)$ no es igual a la $P(B)$ mientras que en la segunda sí se cumple la igualdad. Identificar independencia con situaciones en las que $P(B\|A) = P(B)$ o bien $P(A\|B) = P(A)$. ¿Cómo se calcula $P(A \cap B)$ en cada una de estas situaciones?

Evaluación

Evaluación diagnóstica	Evaluación formativa	Evaluación sumativa
Los prerrequisitos para esta unidad se reducen a las operaciones con números racionales y su significado. Conviene explorar los conocimientos previos del alumnado sobre la probabilidad de un evento y el cálculo de probabilidades.	Durante el desarrollo de la unidad será necesario retroalimentar oportunamente al alumnado para contribuir a que vaya desarrollando una intuición probabilística, mediante la evaluación de cada tema y subtema con actividades continuas. El contenido de cada tema requiere una buena comprensión de los temas y subtemas anteriores.	Al concluir la unidad, se debe evaluar la capacidad del alumnado para resolver problemas de probabilidad e identificar eventos mutuamente excluyentes y eventos independientes.

Evaluación diagnóstica	Evaluación formativa	Evaluación sumativa
Ejemplos	**Ejemplos**	**Ejemplos**
Se puede realizar a través de: • Una lluvia de ideas sobre conceptos como aleatoriedad, probabilidad y frecuencia. • Una discusión guiada acerca de qué indica o qué mide la probabilidad de un resultado.	Se pueden aplicar métodos de evaluación como: • Basada en problemas. Ejemplos de esto son las estrategias sugeridas además de otras actividades que promuevan la comprensión y el uso de los enfoques para asignar probabilidades, así como las reglas para el cálculo de probabilidades.	Se puede basar en productos como: • Los problemas atacados durante el desarrollo de la unidad colectados en un portafolio de evidencias o la resolución de uno o más problemas nuevos en los que se calculen probabilidades. • La conclusión fundamentada de un estudio de caso desarrollado a lo largo de la unidad o la realización de uno nuevo
• Una actividad focal introductoria, por ejemplo, si se lanza 10 veces una moneda común, ¿pueden decir aproximadamente cuántas veces saldrá águila? Después de recoger respuestas pedirle a cada estudiante que haga los lanzamientos y verifique su predicción. Finalizar recogiendo el número de águilas obtenidas en todo el grupo y motivar una conclusión acerca de que existe una relación entre la probabilidad y la frecuencia en un gran número de repeticiones, pero no en pocas.	Conviene recoger conjeturas del alumnado antes de resolver cada problema y contrastarlas con los resultados. • Mediante estudios de caso. Por ejemplo, estudiar la mejor estrategia para un juego en el que se combina una decisión personal con la presencia del azar, como el juego de las tres puertas: se coloca un buen regalo detrás de una de las puertas; el jugador elige una puerta al azar, el conductor abre otra puerta que no tiene el regalo y el jugador debe decidir si mantener su elección inicial o cambiar a la otra puerta aún cerrada. • Con base en una investigación. Se pueden implementar proyectos pequeños, por ejemplo, investigar cómo se usa el enfoque frecuencial para aproximar probabilidades en una tabla de mortalidad y su utilidad.	• La conclusión fundamentada de una investigación desarrollada a lo largo de la unidad. • En cada caso, será necesario precisar si se realizarán presentaciones o trabajos escritos y de qué tipo. • Un examen o cuestionario individual. Conviene tener cuidado de que los reactivos abarquen todo el contenido de la unidad. • Cualquiera de las opciones se puede acompañar de coevaluaciones y autoevaluaciones.

Estadística y probabilidad II

PRESENTACIÓN DE LA ASIGNATURA DE ESTADÍSTICA Y PROBABILIDAD II

Estadística y Probabilidad II es una asignatura introductoria a los métodos para realizar inferencias estadísticas, que son una valiosa herramienta para la toma de decisiones en condiciones de incertidumbre con un amplio campo de aplicaciones en diversas áreas del saber. Conocer estos métodos será de gran utilidad para el alumnado en sus estudios superiores, además de ser importantes para la formación ciudadana y la vida cotidiana.

Se estructura en tres unidades. La primera unidad, que se desarrollará en 24 horas, cierra el bloque de los conocimientos de probabilidad que se requieren para abordar las técnicas inferenciales. En esta unidad, se mostrará que las variables aleatorias permiten elegir una sola característica numérica de los resultados de un experimento aleatorio. Una vez definida una variable aleatoria, es posible construir una función de probabilidad de todos los eventos relacionados con esa característica, a la que le llamamos distribución de probabilidad. El alumnado será capaz de construir distribuciones de probabilidad de distintas variables aleatorias discretas, comprenderá el significado de la esperanza y desviación estándar de una variable aleatoria y conocerá la distribución binomial que es un modelo que aparece con mucha frecuencia y resulta muy útil para lo que se estudiará después.

La dificultad teórica más relevante de esta unidad es el tratamiento de las variables aleatorias continuas, y el paso de las distribuciones de probabilidad discretas a las funciones de densidad. Una forma de abordarlo es construyendo histogramas de frecuencias relativas con una gran cantidad de datos de alguna variable continua, reales o ficticios, y mostrar que el área proporcional de cada barra es igual a la frecuencia relativa que se indica en la barra, es decir, la parte del área total que abarca cada barra es igual a la frecuencia relativa que corresponde a su altura. Mostrar que una forma de aproximar la probabilidad de que la variable tome valores en algún intervalo es calculando el área sobre ese intervalo. De esta manera, al ir aumentando el número de repeticiones y disminuyendo la longitud de cada intervalo, se verá como las tapas superiores de las barras van tendiendo a formar una curva continua y la probabilidad de que la variable tome valores en un intervalo corresponde al área debajo de esa curva y sobre el intervalo.

En la parte final de la unidad se presenta la distribución normal. Se puede introducir como resultado de aproximaciones frecuenciales de distribuciones binomiales, usando un software especializado o simulación interactiva, en las que se toma un valor de n cada vez más grande. El alumnado comprobará que independientemente de qué tan asimétrica sea la primera distribución binomial, el proceso descrito siempre tiende a una campana de Gauss. Conviene analizar las principales características de esa distribución, es decir, que el punto donde se levanta el eje de simetría es la media y que la probabilidad de obtener valores cercanos a la media es grande. Esto se desprende de que el área debajo de la curva es mayor en intervalos localizados en el centro de la campana.

El propósito de la segunda unidad es analizar los dos estadísticos que se usarán como estimadores en la parte final del curso: la media y la proporción muestrales. Tiene asignado un tiempo didáctico de 18 horas. El paso inicial es mostrar que se trata de variables aleatorias ya que, en el desarrollo anterior, el alumnado ha tratado a la media como un número. No es difícil descubrir que en cada muestra aleatoria estos estadísticos toman distintos valores y antes de realizar la muestra no es posible determinar el valor que tomarán. De esto se deduce que tienen una distribución de probabilidad, una esperanza y una desviación estándar. Se plantea el objetivo de analizar qué relación hay entre la esperanza y desviación estándar de la media y la proporción muestrales y los mismos parámetros poblacionales, para lo cual, en esta unidad se asumirá que se conocen esos parámetros.

El resultado teórico fundamental de esta unidad es el Teorema del Límite Central (TLC) que establece que, al aumentar el tamaño de las muestras, la distribución de la media muestral siempre tiende a una distribución normal con una esperanza igual a la media poblacional. Este resultado se cumple para cualquier distribución poblacional, sin importar si es discreta o continua, ni si es muy asimétrica o irregular. El resultado no requiere ninguna hipótesis sobre la población más allá de tener parámetros μ y σ finitos. Con relación a las muestras, el Teorema requiere que sean muestras aleatorias simples, es decir, colecciones de variables aleatorias independientes e idénticamente distribuidas.

Para evitar la exposición pasiva del texto del Teorema, se propone la elaboración de secuencias didácticas que empleen apoyo computacional o de aplicaciones en dispositivos móviles, para que el alumnado pueda visualizar el acercamiento a la forma de campana en las distribuciones de los estimadores aproximadas de manera frecuencial. Las preguntas que se le plantee al alumnado para ayudarlo a descubrir la potencia del TLC deben incluir la observación de que la probabilidad de que la media muestral tome valores cercanos a la media poblacional es grande para muestras de un tamaño adecuado.

La tercera unidad tiene como propósito realizar inferencias formales sobre los valores de los parámetros para fundamentar la toma de decisiones en técnicas estadísticas estructuradas. Se exploran procedimientos de estimación por intervalos y contrastes de hipótesis sobre los valores de la media y la proporción de la población, en un lapso de 22 horas de trabajo académico.

Los métodos para realizar inferencias estadísticas requieren prácticamente todo lo que el alumnado ha estudiado en esta materia. Combinan procedimientos probabilísticos con el manejo de datos y las distribuciones muestrales. Un intervalo de confianza se construye alrededor del valor de la media o la proporción que se obtiene en una sola muestra aleatoria realizada, considerando una probabilidad de acertar, es decir, una probabilidad de que el parámetro caiga en el intervalo. El contraste de hipótesis estriba en el análisis de qué tan inusual es el valor del estimador que se observa al realizar una sola muestra aleatoria, si la hipótesis nula fuera cierta, aceptando una probabilidad de cometer cierto tipo de error.

Ambos procedimientos están atravesados por la presencia del azar. No se debe perder de vista que, aunque es alta la probabilidad de que el valor del estimador esté cerca del parámetro, es poco probable pero no imposible obtener valores alejados de ese parámetro. La interpretación adecuada de los resultados, en contextos específicos, es uno de los aspectos centrales de la propuesta de este programa.

El propósito general del programa de Estadística y Probabilidad II es fomentar el desarrollo de la cultura y el razonamiento estadísticos del alumnado, fortaleciendo sus habilidades para interpretar y comunicar resultados de investigaciones formales de manera efectiva, iniciando además el desarrollo de su pensamiento estadístico a través de una comprensión profunda y duradera de las inferencias estadísticas básicas y su aplicación en diversos contextos. Los propósitos específicos de la asignatura son los siguientes.

El alumnado:

1. Se apropiará del concepto de variable aleatoria y analizará las distribuciones de probabilidad con sus medidas de tendencia y variabilidad, abordando en particular los modelos Binomial y Normal para resolver problemas.
2. Comprenderá la relación de la probabilidad con la estadística mediante el análisis de las distribuciones de la media y la proporción muestrales a través del Teorema del Límite Central para desarrollar su razonamiento estadístico.
3. Realizará inferencias formales sobre los valores de los parámetros, a partir del análisis de los estimadores en una muestra aleatoria, para respaldar la toma de decisiones en una investigación estadística, iniciando el desarrollo de su pensamiento estadístico.

UNIDAD 1. MODELOS DE PROBABILIDAD Y SUS APLICACIONES

Presentación de la unidad

En esta unidad el alumnado abordará el concepto de variable aleatoria, iniciando con el caso discreto, y deducirá su distribución de probabilidad, lo que le permitirá pasar del cálculo de probabilidades de eventos dispersos a la construcción de funciones de probabilidad.

Una variable aleatoria es una función del espacio muestral de un experimento aleatorio al conjunto de números reales, que cumple ciertas condiciones matemáticas. En lugar de centrar la enseñanza de este importante concepto en su definición matemática, se trata de mostrar su utilidad para seleccionar una característica numérica de los resultados del experimento y visualizar las probabilidades de todos los eventos de un mismo tipo.

Se propone abordar el concepto de variable aleatoria discreta junto con el de distribución de probabilidad. La investigación en educación estadística muestra que resulta muy frecuente el error de asignar la misma probabilidad a cada uno de los posibles valores de una variable aleatoria (sesgo de equiprobabilidad). Para evitarlo, se sugiere que los primeros ejemplos que se aborden incluyan todos los elementos del espacio muestral que corresponden a cada valor de la variable aleatoria. Por ejemplo, si la variable es la suma de los dos números que se obtienen al lanzar dos dados comunes, el alumnado debe identificar que hay solo una pareja que tiene suma 2 y hay 6 parejas que tienen suma 7, por lo que la probabilidad de que la variable aleatoria tome el valor 2 es 1/36 mientras que la probabilidad de que tome el valor 7 es 6/36.

Con los conocimientos de probabilidad adquiridos por el alumnado en la Unidad 3 de Estadística y Probabilidad I, es posible construir diversas variables aleatorias discretas, en las que se debe calcular el valor esperado o esperanza matemática y la desviación estándar, resaltando el significado de estos parámetros: el primero, como aproximación del promedio de los valores que toma una variable aleatoria cuando se repite muchas veces el experimento, o bien como promedio ponderado por la probabilidad de cada valor. El segundo como aproximación del promedio a largo plazo de las desviaciones respecto a la media.

Una distribución de probabilidad discreta que modela un tipo de experimentos que aparece con mucha frecuencia es la distribución binomial, que indica la probabilidad de cada número posible de éxitos que se puede obtener al realizar n pruebas independientes de un experimento aleatorio, todas ellas con la misma probabilidad de éxito. Para facilitar la identificación de las probabilidades involucradas en el modelo binomial, se sugiere empezar por presentar los ensayos Bernoulli, en los que solo se realiza una vez un experimento aleatorio cuyos resultados posibles se agrupan en éxito y fracaso. La construcción de la distribución binomial se puede ilustrar en un ejemplo sencillo, analizando paso a paso lo que debe ocurrir para que el número de éxitos sea cero, uno, dos o tres. En cada caso, se usará la independencia para multiplicar probabilidades y se contarán las formas de obtener cada cantidad posible de éxitos, identificando que se trata del cálculo de las combinaciones que indican de cuántas formas se pueden seleccionar los lugares de los éxitos en las n repeticiones del experimento.

Se presentarán, sin deducirlas, las fórmulas para el cálculo de la esperanza y la desviación estándar de una binomial, invitando al alumnado a interpretar sus valores en una diversidad de problemas en distintos campos de conocimiento. Utilizando herramientas digitales se podrán aproximar frecuencialmente diversas distribuciones binomiales para que el alumnado identifique la forma que toma esa distribución cuando la probabilidad de éxito es cercana a cero, a 1 o a ½.

Después se introducirán las variables aleatorias continuas que pueden tomar cualquier valor dentro de un intervalo. Estas variables surgen cuando la característica numérica observada es resultado de una medición. Se sugiere introducirlas abordando algún problema de medida en el que se usen longitudes o áreas. Por ejemplo, considerar el experimento aleatorio del lanzamiento al azar de un dardo sobre un blanco circular y trabajar con la variable aleatoria X que indica la distancia del punto de contacto al centro del blanco. Así, los puntos de contacto que satisfacen una desigualdad del tipo $X<b$ corresponden a un círculo de radio b y los que satisfacen desigualdades del tipo $a<X<b$ están en un anillo circular. Usando el enfoque clásico para calcular probabilidades, no es difícil pasar del conteo de casos favorables y del total de casos posibles, al área de la región favorable y el área total. Un ejemplo de este tipo permite observar que la región que corresponde a cualquier igualdad $X=c$ tiene área cero porque forma una curva dentro del blanco circular (la circunferencia de radio c).

Para introducir la función de densidad o distribución de probabilidad continua, se puede trabajar con histogramas de barras muy delgadas. Si bien en la construcción del histograma se usan las frecuencias relativas como alturas de las barras, se puede mostrar que la frecuencia relativa entre dos valores de la variable es igual al área proporcional de las barras entre esos dos puntos y es muy similar al área debajo de una curva suave construida sobre esas barras. Esa curva será la función de densidad. La intención de una construcción de este tipo es que se identifique la probabilidad de que una variable aleatoria continua tome valores en un intervalo con un área debajo de la curva, no con las imágenes de la función. Hay dos propiedades matemáticas que caracterizan a estas funciones: sus imágenes son mayores o iguales a cero, y el área total debajo de la curva es 1.

Para finalizar la unidad, se aborda la distribución normal, cuyo estudio es muy importante por sus aplicaciones prácticas y por su trascendencia en resultados teóricos como el Teorema del Límite Central. La función de densidad

Normal tiene forma de campana de Gauss, una campana simétrica respecto a un eje vertical levantado sobre el valor que será la esperanza matemática o media de la distribución. Esta forma garantiza que las probabilidades más grandes estén en el centro de la gráfica, ya que las áreas son mayores alrededor del eje de simetría, lo que hace muy probable que la variable aleatoria tome valores cercanos a la media y muy poco probable que tome valores alejados de ella, hacia la derecha o la izquierda.

La distribución normal es un modelo adecuado para el tratamiento de diversas variables que surgen en el estudio de características morfológicas de los individuos de poblaciones numerosas, como la estatura, la talla de los pies o el diámetro de la cabeza; características fisiológicas en grandes poblaciones, como el efecto de un medicamento; características sociológicas como el consumo de cierto producto; en el estudio de errores cometidos al medir ciertas magnitudes, sobre todo cuando se trata de magnitudes muy pequeñas o muy grandes, y en muchas otras situaciones.

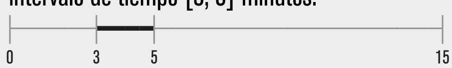
Es importante que el alumnado tome en cuenta que las funciones de densidad normales pueden tener distintos ejes de simetría y distintas alturas. Como el área debajo de la curva debe ser 1, si alcanzan mayor altura deben ser más delgadas y si tienen menor altura deben ser más anchas. La desviación estándar de la distribución normal es la distancia del eje de simetría al punto de inflexión de la campana. En el bachillerato no es posible calcular la media y la desviación estándar de distribuciones continuas, pero es importante que el alumnado reconozca esos parámetros en la representación gráfica y su interpretación en las diversas aplicaciones. Esto facilita la identificación de la distribución normal estándar y la comparación de sus áreas con las de cualquier otra normal.

En cuanto a los ejes transversales, en esta unidad hay que resaltar que la comprensión de las distribuciones Binomial y Normal prepara al alumnado para abordar estudios interdisciplinarios que incluyan a la estadística, la probabilidad y casi cualquier otra área de conocimiento, debido a la gran diversidad de aplicaciones de esos modelos.

Carta descriptiva

Propósito	Tiempo
Al finalizar la unidad, el alumnado: Se apropiará del concepto de variable aleatoria y analizará las distribuciones de probabilidad con sus medidas de tendencia y variabilidad, abordando en particular los modelos Binomial y Normal para resolver problemas.	24 hrs.

Aprendizajes	Temática	Estrategias sugeridas
El alumnado: 1. Reconoce una variable aleatoria discreta y construye su distribución de probabilidad. 2. Construye distribuciones aproximadas frecuencialmente, utilizando simuladores digitales. 3. Interpreta el valor de la esperanza matemática y la desviación estándar en el contexto de una investigación o problema.	Variables aleatorias discretas y sus distribuciones de probabilidad • Ejemplos de variables aleatorias discretas y sus distribuciones de probabilidad. • Esperanza matemática y desviación estándar de una variable aleatoria discreta.	Se sugiere introducir esta unidad planteando alguna situación en la que se requiera asignar números a los resultados de un experimento aleatorio. Por ejemplo: - Una persona paga \$15 para entrar a un juego en el que lanza 4 monedas comunes, una de \$1, una de \$5 y dos de \$10. Gana todas las monedas que caigan en sol. ¿Conviene participar en este juego? ¿Qué se tendría que analizar para fundamentar si conviene o no? Listar los resultados que se pueden obtener en los lanzamientos y las ganancias netas correspondientes. Deducir la probabilidad de cada posible ganancia neta. Puede resolverse también usando un simulador para aproximar frecuencialmente la distribución de probabilidad de la ganancia neta. Para darle un sentido práctico al valor esperado de una variable aleatoria, se puede usar un ejemplo como el que sigue: - Una variable aleatoria indica el mayor de los números obtenidos al lanzar dos dados comunes. La pregunta que nos proponemos contestar es: ¿cuál es el promedio de los valores que toma la variable aleatoria? Se puede hacer una simulación en cualquier software o trabajar en el aula solicitando que cada estudiante repita el experimento 15 o 20 veces y registre el máximo en cada lanzamiento doble. Se reúne la información de todo el grupo en el pizarrón. Supongamos que, en 600 lanzamientos de 2 dados, el máximo fue 1 en 19 ocasiones, 2 en 49 de las repeticiones, 3 en 90 ocasiones, 4 en 103, 5 en 160 y 6 en 179 de las repeticiones. El promedio de los resultados obtenidos es: $\frac{1(19)+2(49)+3(90)+4(103)+5(160)+6(179)}{600} =$ $1\left(\frac{19}{600}\right)+2\left(\frac{49}{600}\right)+3\left(\frac{90}{600}\right)+4\left(\frac{103}{600}\right)+5\left(\frac{160}{600}\right)+6\left(\frac{179}{600}\right)=4.455$

Aprendizajes	Temática	Estrategias sugeridas
		Las frecuencias relativas aproximan el valor de las probabilidades correspondientes, por lo que, si queremos tener una aproximación de este promedio sin repetir un gran número de veces el experimento, podemos calcular la esperanza dada por: $E(X)=1P[X=1]+2P[X=2]+3P[X=3]+4P[X=4]+5P[X=5]+6P[X=6]$ $=1\left(\frac{1}{36}\right)+2\left(\frac{3}{36}\right)+3\left(\frac{5}{36}\right)+4\left(\frac{7}{36}\right)+5\left(\frac{9}{36}\right)+6\left(\frac{11}{36}\right)=4.47$
4. Identifica las características de un proceso binomial. 5. Construye el modelo para la distribución binomial y visualiza su forma para distintos valores de n y p, apoyándose en simuladores digitales. 6. Aplica el modelo binomial, su valor esperado y su desviación estándar en problemas contextualizados interpretando los resultados.	Distribución binomial • Experimentos Bernoulli y Binomial. • Función de distribución binomial, esperanza matemática y desviación estándar. • Aplicaciones.	Se busca que la fórmula de la distribución binomial no solo se presente, sino que sea deducida en un ejemplo sencillo como el siguiente: • Analizar la cantidad de águilas en cuatro lanzamientos de una moneda común. Tras verificar que cumple las condiciones de un experimento Binomial, analizar cuántos resultados hay en los que ocurren 0, 1, 2, 3 o 4 águilas. Calcular la probabilidad de cada uno de esos resultados usando la regla del producto para las probabilidades de éxito y fracaso. Acudir a las combinaciones para contar las formas en que se pueden acomodar las diversas cantidades de águilas en los cuatro lanzamientos. Cuando haya quedado claro el ejemplo, preguntar al alumnado cómo sería la fórmula general para n repeticiones y x éxitos.
7. Comprende que, en el caso de variables aleatorias continuas, la probabilidad debe calcularse para valores dentro de un intervalo. 8. Identifica que la probabilidad de que una variable continua tome valores en un intervalo es el área debajo de su función de densidad y sobre el intervalo.	Variables aleatorias continuas y sus funciones de densidad • Modelos de probabilidad continuos. • La probabilidad como área bajo la curva.	Se sugiere abordar algún problema en donde se requiera usar mediciones geométricas como longitud o área. Por ejemplo: • Un delincuente tiene una conversación telefónica de 15 minutos y habla de información comprometedora en el intervalo de tiempo [3, 5] minutos.  La policía, que le sigue la pista desde hace tiempo, logra interceptar la llamada en un instante aleatorio x y escucha toda la conversación a partir de ese momento. ¿En qué intervalo debe caer x para que la policía escuche toda la información comprometedora? ¿Y para que escuche parte de esa información? ¿Cómo se podría calcular la probabilidad de esos eventos? Discutir primero la necesidad de cambiar el conteo por la longitud para aplicar el enfoque clásico en este caso. Calcular las probabilidades usando el cociente de longitud favorable entre longitud total. Preguntar por la probabilidad de que la policía se conecte exactamente en el momento en que empieza la información comprometedora y llegar a la conclusión de que esa probabilidad es cero porque la longitud favorable es la de un punto.

Aprendizajes	Temática	Estrategias sugeridas
9. Esboza la curva de densidad para una variable aleatoria aproximadamente *normal*, a partir de la suavización de un polígono de frecuencias con ayuda de un simulador digital. 10. Identifica a la media como el punto donde se levanta el eje de simetría y descubre que la probabilidad de obtener valores cercanos a esa media es grande. 11. Calcula probabilidades en una distribución normal aplicando la Regla Empírica, dentro de problemas contextualizados interpretando los resultados. 12. Calcula probabilidades en una distribución normal estándar. 13. Aplica el proceso de estandarización para resolver problemas con cualquier distribución normal. 14. Contrasta el área bajo la gráfica para una situación de comportamiento normal con su correspondiente área en la gráfica del modelo estandarizado.	Distribución normal • Características de una distribución normal y su utilidad. • Identificación geométrica de la esperanza y la desviación estándar. • Normal estándar y el cálculo de probabilidades. • Otras distribuciones normales y su relación con la normal estándar. • Problemas de aplicaciones.	Se busca mostrar a la distribución normal como resultado de un proceso de refinamiento de una distribución Binomial cuando se va aumentando el número de repeticiones independientes del experimento Bernoulli. • Usar un software como GeoGebra para visualizar el histograma y el polígono de frecuencias de una distribución Binomial con $p = 0.05$ o bien $p = 0.95$, y $n= 5$. Ver cómo van cambiando las gráficas cuando n va en aumento usando valores como $n = 20, 50, 100, 125, 150$, acercándose cada vez más a una campana de Gauss con eje de simetría en $E(X) = np$. También se pueden usar simulaciones en las que el valor de n se mueve mediante un deslizador, por ejemplo, https://www.geogebra.org/m/gv5jfcy4 Para acercar al alumnado a las características especiales de esta importante distribución, mostrar que se cumple lo siguiente en cualquier gráfica normal usando simuladores digitales: a) El punto donde se levanta el eje de simetría es la media o esperanza. b) La distancia del eje de simetría al punto donde cambia la concavidad de la campana (hacia cualquiera de los lados) es la desviación estándar. c) El área total debajo de la curva es 1 y del eje de simetría hacia uno de los lados es 0.5. d) La forma de campana garantiza que las probabilidades más grandes se obtienen en intervalos simétricos respecto a la media y que son igualmente pequeñas las probabilidades de intervalos que se alejan de la media hacia la derecha que hacia la izquierda. Para mostrar la utilidad práctica de la distribución normal, se puede: • Emplear la regla empírica para calcular probabilidades en intervalos con centro en la media y longitudes laterales de 1, 2 o 3 veces la desviación estándar, en contextos como la estatura, la talla de los pies u otras medidas morfológicas en poblaciones grandes, el tiempo en que hace efecto un medicamento o la magnitud de errores cometidos en mediciones, entre otros.

Aprendizajes	Temática	Estrategias sugeridas
		Para mostrar la relación entre la Normal estándar y otras distribuciones normales, se propone: • Identificar la distribución normal estándar en términos de su eje de simetría y su desviación estándar y calcular probabilidades usando aplicaciones disponibles en celulares. En tanto se sigan usando tablas en los exámenes extraordinarios, será necesario que se explique la forma de operarlas. Explicar en qué consiste la estandarización de una variable con otra distribución normal y visualizar en un software como GeoGebra o un simulador digital que el área que corresponde a $P[a \leq X < b]$ es igual al área que corresponde a $P[\frac{a-\mu}{\sigma} \leq Z < \frac{b-\mu}{\sigma}]$ en varios casos particulares antes de presentar la formulación general.

Evaluación

Evaluación diagnóstica	Evaluación formativa	Evaluación sumativa
Para esta unidad, el alumnado requiere un buen conocimiento del concepto de probabilidad y la aplicación de los enfoques y las reglas básicas de la probabilidad para resolver problemas. También se requiere un buen conocimiento de las medidas de tendencia central y de dispersión en una colección de datos.	Durante el desarrollo de la unidad será necesario retroalimentar oportunamente al alumnado para contribuir a que vaya adquiriendo un buen manejo de las variables aleatorias discretas y aplique sus conocimientos de probabilidad para determinar su distribución de probabilidad, que deduzca y use el modelo binomial, que comprendan las características particulares de una variable aleatoria continua y manejen adecuadamente el modelo normal.	Al término de esta unidad se evaluarán los conocimientos y habilidades para que el alumnado distinga variables aleatorias discretas y continuas, calcule e interprete los parámetros media y desviación estándar, reconozca los modelos binomial y normal y los pueda usar para resolver problemas.
Ejemplos	**Ejemplos**	**Ejemplos**
Se puede realizar a través de: • Sesiones de resolución de problemas en equipos con discusión grupal para comparar resultados. • Lluvia de ideas para recordar los conceptos de media y desviación estándar de una colección de datos.	Se pueden aplicar métodos de evaluación como: • Basada en problemas. Ejemplos de esto son las estrategias sugeridas y otros problemas diseñados por el o la docente sobre cada uno de los temas y subtemas.	Se puede basar en productos como: • Los problemas atacados durante el desarrollo de la unidad colectados en un portafolio de evidencias o la resolución de uno o más problemas nuevos en los que se requiera determinar distribuciones de probabilidad.
• Un cuestionario con reactivos sobre el significado de la probabilidad de un evento y las distintas formas de calcularla.	• Esto se puede acompañar con rúbricas de autoevaluación, discusiones guiadas sobre las diferencias de las variables aleatorias discretas y las continuas, y la resolución y revisión de tareas para indagar sobre la identificación de los dos modelos que se estudian.	• Bitácora digital de las prácticas realizadas con software y simuladores estadísticos que muestren las habilidades adquiridas por el alumnado. • Un examen o cuestionario individual con reactivos que abarquen todo el contenido de la unidad.

UNIDAD 2. DISTRIBUCIONES MUESTRALES

Presentación de la unidad

Después de revisar los temas de la unidad 1, el alumnado debe estar preparado para seguir con el estudio de variables aleatorias a otro nivel. En la Unidad 2 se desarrollará el aprendizaje sobre cómo las muestras de datos pueden revelar información sobre las características más importantes de las poblaciones de las que fueron extraídas. El concepto sobresaliente en esta unidad es el de distribución muestral.

Una muestra aleatoria se representa de diferentes formas dependiendo del momento en que se observe. Si ya se determinó que la muestra será de n elementos, pero aún no se realiza la extracción, la muestra se representa por una colección de n variables aleatorias $(X_1, X_2, X_3, \ldots, X_n)$, donde X_1 indica el valor que se obtendrá en el primer elemento que se elija, X_2 el que se obtendrá en el segundo, y así sucesivamente. Cuando ya se realizó la extracción y se observó la característica de interés en cada elemento seleccionado, la muestra aleatoria pasa a ser una nada de números $(x_1, x_2, x_3, \ldots, x_n)$ porque ya se conoce el valor que tomó cada una de las variables, esto es, $x_1=x_1$, $x_2=x_2$, y así sucesivamente.

Se le llama estadístico o estadística a cualquier función de las n variables que forman una muestra aleatoria. Un estadístico es una nueva variable aleatoria cuyos valores dependen de los valores que tomen los elementos de la muestra. Se conoce como distribución muestral a la distribución de probabilidad de un estadístico.

Para cada parámetro poblacional hay una colección de estadísticos que se pueden usar como estimadores. La determinación de cuál estadístico es mejor estimador, es un tema extenso que no abordaremos aquí, pero cabe señalar que está comprobado que un muy buen estimador de la media poblacional es la media muestral y un muy buen estimador de la proporción poblacional es la proporción muestral.

Al tratar el tema con el alumnado, más que abordar cuestiones teóricas, es conveniente apoyarse en la variabilidad muestral para comprobar que la media y la proporción muestrales son variables aleatorias. Por ejemplo, se puede plantear el problema de seleccionar una muestra aleatoria de 300 estudiantes del CCH y calcular la edad promedio de esa muestra y la proporción de mujeres que contiene. Antes de realizar la muestra no es posible determinar cuál será el valor promedio ni la proporción de mujeres que contendrá. Si, por ejemplo, una vez realizada la muestra se obtiene una edad promedio de 16.8 años y una proporción de mujeres de 56%, ¿se puede esperar que en otra muestra del mismo tamaño se obtengan los mismos resultados? La respuesta es no, pero es bastante probable que en una nueva muestra se obtengan resultados similares. En una nueva realización se puede obtener una edad promedio de 16.6 años y una proporción de 59% de mujeres. ¿Es posible afirmar que estos valores son iguales en la población? Tampoco es posible, pero lo que el alumnado va a aprender en esta unidad es que es muy probable que los valores que se obtengan en una muestra aleatoria estén cerca de los parámetros o valores poblacionales, siempre que haya sido bien seleccionada y tenga un tamaño adecuado.

Una vez que se haya reconocido que la media y la proporción muestrales son variables aleatorias que pueden tomar distintos valores en cada muestra aleatoria que se realice, el problema que se le planteará al alumnado es averiguar cómo es su distribución de probabilidad, su media y su desviación estándar. En particular, nos interesa saber cómo se relacionan la media y la desviación estándar muestrales con los mismos parámetros poblacionales. Es importante recalcar que en esta unidad vamos a asumir que se conocen los parámetros poblacionales para investigar esta relación.

Conviene llevar a cabo una actividad en la que sea posible enumerar todas las muestras aleatorias simples y calcular en ellas una media o una proporción. Para ello, será necesario utilizar una distribución poblacional hipotética muy sencilla y que las muestras sean de 2 elementos. El alumnado podrá observar cómo cambia el valor del estadístico de una muestra a otra y descubrir que puede haber dos o más muestras que conduzcan a los mismos valores. Es importante recalcar que la distribución de la media y de la proporción muestrales no se parece a la distribución poblacional.

Después de observar que este tipo de análisis se complica mucho si la distribución de probabilidad poblacional incluye 10 o más valores posibles y las muestras son de 5 o más elementos, se evidenciará la necesidad de aproximar frecuencialmente las distribuciones muestrales con apoyo tecnológico.

Uno de los objetivos más destacados de esta unidad es desarrollar en el alumnado una comprensión precisa del Teorema del Límite Central (TLC), destacando su potencia y utilidad para aproximar las distribuciones muestrales cuando el tamaño de la muestra es suficientemente grande. El alumnado debe visualizar, a través de simulaciones digitales, que independientemente de cómo sea la distribución poblacional, las distribuciones muestrales de la media y la proporción tienden a parecerse a una campana de Gauss al ir aumentando el tamaño de la muestra. Adicionalmente, el error estándar va disminuyendo, lo que hace que los valores de los estadísticos se tornen cada vez más concentrados alrededor de la media. Como la media de la distribución muestral es precisamente el valor del parámetro en los casos que estudiamos aquí, esto garantiza que la probabilidad de que los estimadores tomen valores cercanos al parámetro es cada vez más grande mientras mayor sea el tamaño de la muestra.

Después de ver lo anterior con ayuda de un software especializado o simulación interactiva, la formulación del TLC tendrá mucho sentido para el alumnado y se facilitará su aplicación. La hipótesis del teorema que es importante resaltar es que las muestras aleatorias deben cumplir que cada elemento de la población tenga la misma probabilidad de ser seleccionado en cualquier paso del procedimiento. Esto es a lo que se le llama una muestra aleatoria simple. La hipótesis formal consiste en tratar con colecciones de variables aleatorias independientes e idénticamente distribuidas. Son idénticamente distribuidas por referirse a la misma característica numérica definida en la misma población, y son independientes por la condición descrita anteriormente. Para el alumnado de bachillerato, la formulación de este importante teorema puede ser la siguiente:

Teorema del Límite Central. Considérese una muestra aleatoria simple $X_1, X_2, X_3, \ldots, X_n$ tomada de una población con media μ y desviación estándar σ. Al ir aumentando el tamaño de la muestra n, la media muestral $\overline{X} = \frac{X_1 + X_2 + X_3 + \ldots + X_n}{n}$ tiende a distribuirse como una Normal con media μ y desviación estándar $\frac{\sigma}{\sqrt{n}}$, es decir, la variable $\frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$ tiende a distribuirse como una Normal estándar.

La conclusión del TLC se extiende fácilmente a la proporción muestral ya que esta es el promedio de n variables con distribución Bernoulli, es decir, la proporción es el promedio de las variables que toman solo los valores 1 o 0 dependiendo de si el elemento de la muestra tiene o no la característica que se estudia.

Finalmente, se emplearán las distribuciones muestrales para calcular probabilidades sobre la media y la proporción muestral cuando se conocen los parámetros.

Cabe resaltar que en esta unidad resulta esencial recurrir al apoyo tecnológico, pues no es posible calcular o simular manualmente cientos de muestras de 30, 50 o más elementos. Esto evidencia el poder de esas herramientas para facilitar la comprensión del alumnado de conceptos teóricos abstractos, sea mediante aplicaciones que únicamente requieren internet o mediante softwares que se pueden trabajar en salas de cómputo. Pero es indispensable que su uso esté integrado en secuencias didácticas cuidadosamente diseñadas para conducir la indagación y discusión del alumnado hacia las ideas fundamentales.

El eje transversal que más se desarrolla en esta unidad es el de conocimiento y aplicación de las tecnologías de la información y de la comunicación. Esto se debe a que no es posible desarrollar los conceptos teóricos que conducen a los principales resultados que se requieren en los temas de la unidad con la formación matemática que tiene el alumnado de nivel bachillerato, por lo que nuestra mejor herramienta consiste en que visualicen esos resultados con apoyo tecnológico.

Carta descriptiva

Propósito	Tiempo
Al finalizar la unidad, el alumnado: Comprenderá la relación de la probabilidad con la estadística mediante el análisis de las distribuciones de la media y la proporción muestrales a través del Teorema del Límite Central, para desarrollar su razonamiento estadístico.	18 hrs.

Aprendizajes	Temática	Estrategias sugeridas
El alumnado: 1. Identifica a los estadísticos como indicadores del posible valor puntual de sus correspondientes parámetros. 2. Construye muestras aleatorias simples a partir de una población con distribución conocida y calcula los valores de un estimador en cada muestra. 3. Concluye que los estimadores son variables aleatorias. 4. Elabora la distribución de probabilidad de un estimador en un ejemplo sencillo. 5. Fórmula conjeturas acerca del comportamiento de una variable en la población, a partir de los datos de una muestra, en el contexto de una investigación o un problema.	Población y muestra • Variabilidad muestral. • Parámetros y estadísticos • Muestreo aleatorio simple. • Cálculo de una distribución muestral.	Con el objetivo de reconocer la variabilidad de las muestras y la posible relación entre los valores de los estimadores y los parámetros, se sugiere utilizar un problema de este tipo: • En un cuestionario digital (Microsoft Forms, Google Forms u otro), realizar una pequeña encuesta entre el alumnado en la que se pregunte el género (hombre, mujer, no binario) y cuánto dinero trae en los bolsillos. Los resultados de la encuesta serán los datos poblacionales. Solicitar al alumnado que se calcule la media poblacional de dinero (μ_X) y la proporción poblacional de mujeres (p). Organizar al alumnado en equipos de 5 personas de manera aleatoria, y ya agrupados, cada uno debe decir cuánto dinero trae en su bolsillo, para que se calcule el promedio por persona del dinero en cada equipo. Recoger la información de cada equipo en el pizarrón. Organizar nuevos equipos aleatorios de 5 integrantes para que hagan lo mismo y recoger los resultados. Organizar una discusión grupal con preguntas como: ¿Los promedios obtenidos en las muestras son parecidos entre sí?, ¿en cuántas muestras estos promedios estuvieron muy alejados del resto? ¿Cuántos de los promedios de las muestras se parecen al promedio poblacional? Si se forman equipos por tercera vez al azar, ¿los promedios serán similares? Usando la lista numerada de los datos poblacionales y con el apoyo de algún generador digital de números aleatorios, indique a los equipos que tomen un par de muestras aleatorias de 25 elementos de la población, en las que calculen la media muestral de dinero ($\overline{X}$) y la proporción muestral de mujeres ($\widehat{p}$). Solicitar que comparen esos resultados con la media y proporción poblacionales. Reúna los resultados de todas las muestras en el pizarrón. Finalice la actividad con una discusión grupal en la que se analice la cercanía de los valores de las medias y proporciones muestrales entre sí y con los parámetros. Pregunte si se observó algún cambio al aumentar el tamaño de las muestras.

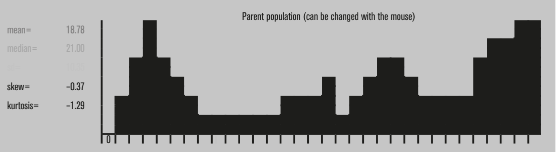
<table>
<tr><th>Aprendizajes</th><th>Temática</th><th>Estrategias sugeridas</th></tr>
<tr><td></td><td></td><td>
Se propone determinar una distribución muestral haciendo cálculos directos en una situación como la siguiente:
• Suponer que, en una población de estudiantes, la variable aleatoria X = edad de un alumnado elegido al azar, puede tomar los valores 15, 16, 17 y 18 años con las siguientes probabilidades:
<table>
<tr><td>x</td><td>15</td><td>16</td><td>17</td><td>18</td></tr>
<tr><td>$P[X = x]$</td><td>0.2</td><td>0.3</td><td>0.2</td><td>0.3</td></tr>
</table>
Con base en la información de la tabla calcular la media (μ_X) y su varianza (σ^2_X). Se van a comparar esos parámetros con los valores de las medias muestrales en muestras aleatorias simples de tamaño 2.

Solicitar que enlisten todas las posibles parejas de edades que se pueden obtener y calculen la media de cada una ($\bar{X}$). Se sugiere preguntar ¿en cuántas muestras de dos elementos la media de la muestra coincide con la media de la población? ¿Cuáles son todos los valores que puede tomar la media de una muestra de dos elementos? Guiar una discusión utilizando las respuestas anteriores para que el alumnado construya la distribución de probabilidad de las medias muestrales:
<table>
<tr><td>$\bar{x}$</td><td>15</td><td>15.5</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>$P[X = \bar{x}]$</td><td>0.04</td><td>0.12</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
</table>
Solicitar que calculen la esperanza de la media muestral ($\mu_{\bar{X}}$) y la varianza de la media muestral ($\sigma^2_{\bar{X}}$).
Realizar algunas preguntas en discusión grupal como: ¿Qué relación hay entre μ_X y $\mu_{\bar{X}}$?, ¿Entre qué se debe dividir σ^2_X para obtener $\sigma^2_{\bar{X}}$?
</td></tr>
</table>

<table>
<tr><th>Aprendizajes</th><th>Temática</th><th>Estrategias sugeridas</th></tr>
<tr>
<td>
6. Comprueba que los estimadores media y proporción muestral se distribuyen de manera aproximadamente normal, al trabajar con muestras grandes con la ayuda de la simulación por computadora.
7. Construye las distribuciones muestrales para la media y la proporción, bajo las condiciones del Teorema del Límite Central.
8. Emplea las distribuciones muestrales para calcular probabilidades sobre la media y la proporción muestral cuando se conocen los parámetros.
</td>
<td>
Las variables aleatorias media y proporción muestrales

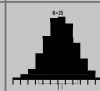
• Teorema del Límite Central.
• Distribución muestral de medias.
• Distribución muestral de proporciones.
</td>
<td>
Para que el alumnado visualice las conclusiones del Teorema del Límite Central y se dé cuenta de que no importa como sea la distribución poblacional, se puede aplicar una sugerencia como la siguiente.

• Solicitar al alumnado que utilice algún simulador digital con la ayuda de su celular e internet, en el que pueda elegir la forma de la distribución poblacional, su media y su desviación estándar, así como el tamaño de la muestra y el número de muestras que se simularán. Por ejemplo, se puede usar https://onlinestatbook.com/stat_sim/sampling_dist/

Este simulador ofrece varias distribuciones para ser usadas como distribución poblacional y también ofrece la posibilidad de que el usuario construya la forma de esa distribución moviendo barras de un histograma de manera táctil. Esta es una opción:


Solicitar al alumnado que observe los valores de los parámetros poblacionales que muestra el simulador:
μ_X (mean) y σ_X (sd).
Para la cantidad de muestras que se simulan, se ofrecen tres opciones: 5, 10000 y 100000, y para el tamaño de las muestras se ofrecen seis posibilidades: 2, 5, 10, 16, 20 y 25. Una vez que se realizan las simulaciones, se presenta la distribución de la media muestral en azul:


</td>
</tr>
</table>

<table>
<tr><th>Aprendizajes</th><th>Temática</th><th>Estrategias sugeridas</th></tr>
<tr><td></td><td></td><td>
Solicitar al alumnado que cambie el tamaño de las muestras y analice las variaciones que se producen. También puede cambiar la distribución poblacional. Conviene que completen una tabla como la siguiente:
<table>
<tr><td colspan="5">Media poblacional μ_X =</td></tr>
<tr><td colspan="5">Desviación estándar poblacional σ_X =</td></tr>
<tr><td>Tamaño de la muestra</td><td>n = 2</td><td>n = 5</td><td>n = 10</td><td>n = 25</td></tr>
<tr><td>Esperanza de las medias muestrales</td><td>$\mu_{\bar{X}}$ =</td><td>$\mu_{\bar{X}}$ =</td><td>$\mu_{\bar{X}}$ =</td><td>$\mu_{\bar{X}}$ =</td></tr>
<tr><td>Desviación estándar de las medias muestrales (error estándar)</td><td>$\sigma_{\bar{X}}$ =</td><td>$\sigma_{\bar{X}}$ =</td><td>$\sigma_{\bar{X}}$ =</td><td>$\sigma_{\bar{X}}$ =</td></tr>
<tr><td>Gráfica</td><td></td><td></td><td></td><td></td></tr>
</table>
Cerrar la actividad solicitando que contesten preguntas como:
¿Se parecen los histogramas de la distribución poblacional y la distribución muestral?
El histograma de la distribución muestral, ¿se parece a alguna distribución que conozcas?, ¿cuál?
Al cambiar el tamaño de las muestras de 2 a 25, ¿qué sucede con el ancho y la altura de las gráficas?
Compara la media poblacional con la esperanza de las medias muestrales para diferentes tamaños de muestra.
¿Qué relación hay entre μ_X y $\mu_{\bar{X}}$?
¿Cómo es la dispersión en la distribución de las medias muestrales cuando el tamaño de la muestra va aumentando?
</td></tr>
</table>

Evaluación

Evaluación diagnóstica	Evaluación formativa	Evaluación sumativa
En esta unidad se recomienda realizar una evaluación diagnóstica que contemple los siguientes puntos: • Definiciones básicas de Estadística, como población, muestra y muestreo. • Temas de Probabilidad, como azar, aleatorio, espacio muestral y cálculo de probabilidad de diferentes clases de eventos. • Variables Aleatorias y sus Distribuciones de Probabilidad, en particular, la Distribución Normal.	Los aspectos centrales que se requiere tomar en cuenta para la evaluación formativa de esta unidad incluyen la comprensión de: • Que los estimadores son variables aleatorias. • La relación entre los parámetros y la media y el error estándar de los estimadores. • La potencia y el alcance del TLC.	En esta evaluación se requiere comprobar qué nivel de dominio adquirió el alumnado sobre las distribuciones muestrales de la media y la proporción, su comprensión del TLC y las condiciones que se requieren para su aplicación y la interpretación adecuada del hecho de que media de cada una de estas distribuciones muestrales es el parámetro que se desea estimar.
Ejemplos	**Ejemplos**	**Ejemplos**
Se puede realizar a través de: • Una discusión guiada acerca de conceptos como población, muestra, muestreo y variabilidad muestral. • Una actividad de simulación física, usando cualquier generador de números aleatorios en dispositivos móviles, para extraer muestras de una población conocida. • Planteamiento de un problema para que se identifique el objetivo de una investigación, la población, el procedimiento de muestreo y el tamaño de la muestra.	Se pueden aplicar métodos de evaluación como: • Basada en problemas. Ejemplos de esto son las estrategias sugeridas y otros problemas diseñados por el o la docente sobre la utilidad del TLC en diferentes situaciones, sobre las diferencias y similitudes entre las distribuciones muestrales de la media y la proporción, y sobre el cálculo de probabilidades acerca de la distancia entre la media muestral y la poblacional, entre otros temas. Esto se puede acompañar con rúbricas de autoevaluación. • Con base en una investigación. Se pueden implementar proyectos pequeños, por ejemplo, investigar qué pasa con la distribución de la media o proporción muestrales al ir aumentando el tamaño de muestra mediante aproximaciones frecuenciales en un software o simulación interactiva. Escribir una buena descripción del comportamiento observado.	Se puede basar en productos como: • Los problemas atacados durante el desarrollo de la unidad colectados en un portafolio de evidencias o la resolución de uno o más problemas nuevos en los que se requiera calcular probabilidades sobre la media o la proporción muestral. • Textos con las conclusiones de las pequeñas investigaciones realizadas. • Planteamiento de un caso o de un problema en el cual apliquen las herramientas adquiridas. • Análisis crítico de un artículo de investigación para verificar los resultados aplicando distribuciones muestrales. • Cuestionario con preguntas que requieran interpretación y toma de decisiones basada en los resultados obtenidos.

UNIDAD 3. INFERENCIA ESTADÍSTICA

Presentación de la unidad

La presente unidad es la culminación del trabajo realizado durante todas las unidades anteriores de ambas asignaturas. La inferencia estadística es uno de los temas más enseñados en la universidad, por tratarse de una herramienta fundamental para tomar decisiones cuando se desconocen los parámetros de una población y solo se cuenta con la información que brinda una muestra aleatoria.

En esta unidad, se trata de resolver dos problemas fundamentales: la estimación de parámetros poblacionales desconocidos a partir de los estadísticos elegidos, y analizar una hipótesis sobre un parámetro cuyo valor se ignora.

La estimación puntual consiste en proporcionar un solo valor posible del parámetro desconocido, estimado por el valor que toma un estadístico al realizar una muestra aleatoria. Otra posibilidad consiste en no brindar un valor particular sino un intervalo en el que puede caer el parámetro desconocido. Se parte de los estimadores media y proporción muestrales para generar dos límites que son aleatorios antes de realizar la muestra. Se le llama nivel de confianza a la probabilidad de que el parámetro caiga en ese intervalo aleatorio. Para el cálculo de los límites se utiliza el error de estimación, y se calcula considerando el error estándar del estadístico utilizado como estimador y el valor Z para el nivel de confianza que elija el investigador.

Un error común es definir el nivel de confianza como la probabilidad de que el valor del parámetro esté contenido en el intervalo numérico obtenido una vez realizada la muestra. Esta interpretación ignora el hecho de que una vez calculado el intervalo, este contiene al parámetro o no lo contiene, es decir, la probabilidad ya tiene un valor particular que puede ser 1 (evento seguro) o bien 0 (evento imposible).

El nivel de confianza indica la proporción de intervalos que contendrán el valor del parámetro, para distintas realizaciones de una muestra con un tamaño dado, pero no indica si el intervalo calculado con base en una realización lo contiene o no. Se recomienda el uso de simulaciones interactivas para que el alumnado entienda la interpretación correcta del nivel de confianza. (Por ejemplo, seeing-theory.brown.edu, rossmanchance.com, stapplet.com, digitalfirst.bfwpub.com, istats.shinyapps.io).

Cuando se interpreta un resultado obtenido en una muestra, dentro del contexto de una investigación o problema, surge la pregunta de si pudo haber ocurrido por casualidad. Por ejemplo, si se quiere analizar la hipótesis de que la edad promedio del alumnado del CCH es mayor o igual a 17 años y en una muestra aleatoria de 300 estudiantes se obtiene una edad promedio de 16.8 años, es muy factible que la diferencia entre la hipótesis y el resultado experimental se deba a la variabilidad muestral. Pero si en la muestra se obtiene una edad promedio de 16.2 años, no parece factible que la diferencia se deba al azar. La pregunta que surge es: ¿desde qué edad promedio conviene rechazar la hipótesis?

El contraste de hipótesis es un procedimiento estadístico para evaluar si el azar es una explicación plausible del resultado obtenido en una muestra aleatoria. Su finalidad es decidir entre dos hipótesis, es decir, inclinarse por rechazar o no una determinada hipótesis o conjetura sobre el valor de un parámetro desconocido, a partir de lo observado en una sola muestra aleatoria. La hipótesis que afirma que la variación del resultado experimental con respecto a lo esperado se debe únicamente al azar, se llama hipótesis nula.

La falsa idea de que al finalizar hemos logrado probar la verdad o falsedad de una hipótesis es común tanto entre el alumnado como entre profesionales. Para ayudar a prevenir esta interpretación errónea, en esta unidad se debe revisar la lógica del contraste de hipótesis, en la cual se contemplan dos posibles decisiones respecto a la hipótesis nula (H_0): rechazar esta hipótesis, asumiendo que es falsa y aceptando la alternativa, o abstenerse de esa acción.

Partiendo de la información contenida en la muestra aleatoria se calcula el valor del estadístico de prueba del contraste, dejando a consideración del profesorado comparar el p-valor con el nivel de significación o, equivalentemente, comparar el valor del estadístico calculado con el valor crítico que limita una región de rechazo.

La decisión de rechazar la hipótesis nula se toma cuando lo observado en la muestra se considera suficientemente atípico si la hipótesis nula fuera cierta. El nivel de significancia o de significación, α, es la máxima probabilidad que el investigador acepta de cometer el error de rechazar la hipótesis nula cuando esta es cierta, es decir, $\alpha=P$ (Rechazar $H_0|H_0$ es cierta). Su interpretación errónea más común es confundirla con la probabilidad de que la hipótesis nula sea cierta, cuando se ha tomado la decisión de rechazarla, esto es, intercambiar el papel del condicionante y el condicionado en la probabilidad anterior.

Así pues, en esta unidad se emplean las medidas estadísticas y las técnicas de manejo de la información para analizar la información que brindan las muestras; se usan las herramientas probabilísticas y las distribuciones muestrales en la construcción de los extremos de un intervalo y en el procedimiento para rechazar o no una hipótesis nula, y se hace uso de la formación que ha ido desarrollando en el alumnado la constante formulación de inferencias informales, para una mejor comprensión de cómo interpretar los resultados.

En las inferencias estadísticas formales, la decisión se basa en la información obtenida en una sola muestra aleatoria, y, por ello, es muy importante tener en cuenta que si bien el TLC permite afirmar que hay una probabilidad alta de que la media y proporción muestrales estén cercanas a los parámetros, resulta poco probable pero no imposible que se obtengan valores alejados de los parámetros al realizar una muestra. Se debe hacer consciencia en el alumnado de que al llevar a cabo los procedimientos no hay forma de que sepamos si estamos en esa situación o no.

Las técnicas que se presentan en esta unidad se pueden aplicar en una gran diversidad de situaciones que abarcan, en particular temas relacionados con los ejes transversales sustentabilidad, formación ciudadana y perspectiva de género. Por ejemplo, en la estrategia sugerida sobre la recolección de tapas de plástico se pretende guiar una discusión que atienda los ejes de formación ciudadana (pues la recolección es para apoyar a niños con cáncer) y de sustentabilidad (pues se apoya al correcto manejo de residuos). Su aplicación en estudios interdisciplinarios requiere, por lo general, de algunas modificaciones que eviten la necesidad de conocer la desviación estándar poblacional, lo que conduce a una función de densidad ligeramente diferente a la normal: la distribución t de *Student*.

Carta descriptiva

Propósito	Tiempo
Al finalizar la unidad, el alumnado: Realizará inferencias formales sobre los valores de los parámetros, a partir del análisis de los estimadores en una muestra aleatoria, para respaldar la toma de decisiones en una investigación estadística, iniciando el desarrollo de su pensamiento estadístico.	22 hrs.

Aprendizajes	Temática	Estrategias sugeridas
El alumnado: 1. Distingue los tipos de inferencias estadísticas que se abordarán y reconoce cuándo se requiere estimar un parámetro y cuándo verificar una hipótesis respecto a un parámetro desconocido.	Tipos de inferencias estadísticas • Estimación. • Contraste de hipótesis.	Se busca generar discusiones que muestren la necesidad de tener información sobre el valor de un parámetro desconocido, o verificar una conjetura, solo con la información de una muestra aleatoria. Por ejemplo, plantear situaciones como las siguientes: • Se pretende juntar tapas de plástico para apoyar una Asociación Civil que ayuda a niños con cáncer, pero no se cuenta en el plantel con un espacio para almacenarlas. Por lo tanto, se deberán colectar el mismo día que serán recogidas. La Asociación necesita conocer un aproximado de la cantidad de tapitas que le serán entregadas para organizar la recolección. Discutir con el alumnado: ¿cómo podemos realizar esta predicción?, ¿cuál es la población de interés?, ¿tenemos acceso a ella previo al día de recolección?, ¿qué nos interesa conocer de esa población? Partiendo de una prueba piloto en un grupo, ¿qué podríamos medir en ella para predecir la cantidad de tapas que en promedio llevará cada estudiante del plantel el día de la recolección?

Aprendizajes	Temática	Estrategias sugeridas
		• ¿Cuántos han visto la película Avengers Infinity War? En esa película se afirma que Thanos consigue eliminar a la mitad de la población del Universo. Al final de la película se pueden identificar personajes que se esfumaron y otros que no lo hicieron. Con esa información, ¿es posible corroborar la afirmación de Thanos? • Discutir con el alumnado: si se desea verificar la proporción de eliminados, ¿cuál sería la población de interés?, ¿tenemos acceso a ella?, ¿cómo se puede usar la información con la que se cuenta para verificar si es verdad lo que afirma Thanos?
2. Identifica el concepto de estimación por intervalos en el contexto de una investigación o un problema. 3. Construye las expresiones para el cálculo de intervalos de confianza para la media y la proporción, bajo las condiciones del Teorema del Límite Central. 4. Calcula el intervalo de confianza para la media y la proporción e interpreta el resultado obtenido en el contexto de una investigación o problema.	Intervalos de confianza para la media y la proporción de la población • Estimación puntual y por intervalos. • Significado del nivel de confianza. • Determinación de un intervalo de confianza para la media y la proporción de la población. • Relaciones entre nivel de confianza, tamaño de muestra y longitud del intervalo. • Aplicación e interpretación de resultados.	Para analizar los elementos que componen un intervalo de confianza, se puede dar continuidad al ejemplo de la recolección de tapas de plástico de la siguiente manera. • Se solicita que cada estudiante colecte tapas de plástico y las lleve al salón un día preestablecido. Reunir el conjunto de tapas y calcular el promedio por estudiante. Si no es posible realizar la experimentación, considerar que se tiene ese dato. Discutir sí lo observado en el grupo se puede generalizar a la recolección en todo el plantel. ¿Cómo se puede utilizar esta información del grupo para estimar la cantidad de tapas que en promedio llevará cada estudiante participante del plantel? Guiar la discusión para concluir que conviene calcular un intervalo y mostrar el procedimiento para calcularlo. Se busca que el alumnado comprenda el significado del nivel de confianza, con actividades como: • Utilizar un simulador interactivo o un software especializado y plantear la situación en la que se conoce el parámetro, se extrae un gran número de muestras y con cada muestra se determina un intervalo. Esto permite verificar en la práctica qué proporción de intervalos sí contienen el parámetro. Cambiar el nivel de confianza y verificar cómo cambia esta proporción. • Discutir las relaciones entre el nivel de confianza, el tamaño de muestra y la longitud del intervalo, para lo cual se recomienda el uso de herramientas tecnológicas.

Aprendizajes	Temática	Estrategias sugeridas
		Conviene concluir la discusión del tema anterior construyendo un intervalo de confianza en los contextos iniciales, por ejemplo: • En la película *Avengers Infinity War*, un espectador observó que de 40 personajes que participaron en la película, 18 se esfumaron. Considerar ese resultado como una muestra aleatoria para calcular un intervalo para estimar la proporción de la población del Universo que se esfumó con un nivel de confianza del 95%.
5. Comprende el concepto de contraste de hipótesis, para una media y para una proporción, bajo las condiciones del Teorema del Límite Central. 6. Construye contrastes de hipótesis para la media y para la proporción, preferentemente con el uso de herramientas tecnológicas, en el contexto de una investigación o un problema. 7. Concluye si se rechaza o no una hipótesis estadística para la media o para la proporción, en el contexto de una investigación o un problema.	Contraste de hipótesis para la media y la proporción de la población. • Tipos de parejas de Hipótesis. • Interpretación del nivel de significancia. • Estadístico de Prueba. • Desarrollo del contraste: región de rechazo o valor p. • Aplicación e interpretación de resultados.	Para presentar los elementos de un contraste de hipótesis y el procedimiento a realizar, se sugiere emplear los contextos trabajados con anterioridad. Por ejemplo: • En el ejemplo de recolección de tapas de plástico, agregar la siguiente consideración: la Asociación solo puede hacer la recolección a domicilio cuando la cantidad de tapas es de al menos 10 mil 000 tapas. Considerar la información obtenida en el grupo como una muestra aleatoria para responder: ¿se alcanzará el promedio por estudiante necesario para lograr la recolección de la Asociación en el plantel? Para responder, considerar que participan 500 estudiante. ¿Y si fueran 1000 estudiante? Para realizar una aplicación, se puede usar el otro contexto. • Thanos afirma que consigue eliminar a la mitad de la población del Universo. Un espectador de Avengers Infinity War considera que el desastre fue mayor pues de 64 personajes que participaron 35 se esfumaron. Para un nivel de significancia del 5%, ¿se acepta lo dicho por Thanos? El profesorado decidirá trabajar contraste de hipótesis a través de los valores críticos propios del enfoque clásico, o a partir del p–valor, ya que esta última vía es la recurrente en los paquetes computacionales.

Evaluación

Evaluación diagnóstica	Evaluación formativa	Evaluación sumativa
Los requerimientos básicos para esta unidad incluyen una correcta comprensión de los conceptos: parámetro, estadístico, distribución Normal, distribución muestral, así como el TLC. También conviene revisar las ideas previas acerca de las inferencias estadísticas formales e informales.	Para recoger evidencias de la comprensión de cada tema y subtema, se sugiere proponer situaciones en contextos que reflejen los ejes transversales, es decir, temas relacionados con la formación ciudadana, la perspectiva de género, la sustentabilidad, el uso de las TIC y la interdisciplinariedad. Incluir también actividades sobre temas de interés del alumnado.	Es importante observar si el alumnado es capaz de aplicar los procedimientos inferenciales abordados, si los puede explicar y usar para fundamentar conclusiones de manera adecuada, y si es capaz de conectar los conceptos estadísticos para elegir el procedimiento que conviene aplicar en un problema real de análisis de datos.
Ejemplos	**Ejemplos**	**Ejemplos**
Se puede realizar a través de: • Una lluvia de ideas sobre la diferencia entre parámetro y estadístico. • Una discusión guiada sobre las características de una distribución Normal y las dos distribuciones muestrales estudiadas. • Un organizador gráfico sobre las inferencias estadísticas.	Se pueden aplicar métodos de evaluación como: • Basada en problemas. Plantear preguntas que conduzcan a la necesidad de la inferencia y presentar los elementos de cada procedimiento en contextos de situaciones problemáticas. • Estudio de caso. Por ejemplo, buscar en las estadísticas del Colegio una característica que se pueda medir en el alumnado, calcular el estadístico con los datos del grupo y a partir de este realizar una estimación por intervalos o plantear una hipótesis a verificar.	Se puede basar en productos como: • Los problemas atacados durante el desarrollo de la unidad colectados en un portafolio de evidencias o la resolución de uno o más problemas nuevos en los que se requiera aplicar lo aprendido en esta unidad. • Un examen o cuestionario individual con reactivos que abarquen todo el contenido de la unidad.

Referencias

Para el alumnado

Básicas

Mendenhall, W., Beaver, R. J., & Beaver, B. M. (2018). *Probabilidad y Estadística para el Bachillerato.* CENGAGE Learning.
https://issuu.com/cengagelatam/docs/9786075267326

Kelmansky, D. (2009) *Estadística para todos, estrategias de pensamiento y herramientas para la solución de problemas*. Artes Gráficas Rioplatense. https://acrobat.adobe.com/id/urn:aaid:sc:VA6C2:d6618d41-8dbb-433d-a73f-fed3c4281cf7

Triola, M. F. (2018). *Estadística* (12 ed.). Pearson. http://librodigital.sangregorio.edu.ec/librosusgp/B0038.pdf

Complementarias

Aguilar, A., Altamira, J., & García O. (2010). *Introducción a la inferencia estadística.* Pearson. http://biblioteca.usfa.edu.bo/cgi-bin/koha/opac-retrieve-file.pl/?id=1a7263c447fb8b7b6ddf02133d39d6ab

Castillo, J., Gómez J. (1998). *Estadística inferencial básica*. Iberoamérica

Coladarci, T., Cobb, C. (2013). *Fundamentals of Statistical Reasoning in Educatio*n (4 edición). Library of Congress Cataloging-in-Publication Data. https://acrobat.adobe.com/id/urn:aaid:sc:VA6C2:0fe33dcd-4d12-4a40-8db5-5220ad26c1df

Para el profesorado

Básicas

Coladarci, T., Cobb, C. (2013). *Fundamentals of Statistical Reasoning in Educatio*n (4 edición). Library of Congress Cataloging-in-Publication Data. https://acrobat.adobe.com/id/urn:aaid:sc:VA6C2:0fe33dcd-4d12-4a40-8db5-5220ad26c1df

Gutiérrez, E. y Panteleeva, V. O. (2014). *Probabilidad y Estadística, Aplicaciones a la ingeniería y las ciencias.* Patria. https://gc.scalahed.com/recursos/files/r161r/w25418w/sem8_probabilidadyestadistica.pdf

Mendenhall, W., Beaver, R. J., & Beaver, B. M. (2023). *Introducción a la probabilidad y estadística*. Cengage Learning.
https://www.fcfm.buap.mx/jzacarias/cursos/estad2/libros/book5e2.pdf

Complementarias

Álvarez, I. y Romero, V. A. (2019). *Enseñanza y aprendizaje de la estadística y la probabilidad: Propuesta de intervención para el aula*. Universidad Pedagógica Nacional. Xpress Estudios Gráficos Digital. https://acrobat.adobe.com/id/urn:aaid:sc:VA6C2:11ee8264-095b-4c6d-8dae-2dc5bd87c669

Castillo, J., Gómez J. (1998). *Estadística inferencial básica*. Iberoamérica

Devore J. L. (2012) *Probabilidad y Estadística para ingeniería y ciencias* (8ª ed). CENGAGE Learning. https://bibliotecaia.ism.edu.ec/Repo-book/p/ProbabilidadEstadistica.pdf

Infante, S., Zárate, G. (2012). *Métodos Estadísticos. Un enfoque interdisciplinario* (3ª ed.). Colegio de Postgraduados.

Núñez del Prado, A. (2006) *Estadística básica para planificación* (17ª ed). Siglo XXI.

Ross, S. M. (2014). *Introducción a la Estadístic*a (2ª edición). Editorial Reverté.

Referencias de contenido transversal

Álvarez, M., Castillo, J. (2019) *Panorama estadístico de la violencia contra niñas, niños y adolescentes en México*. (1ª ed) UNICEF. https://www.unicef.org/mexico/media/1731/file/UNICEF_PanoramaEstadistico.pdf

Educaverde Eco-reto (n.d.). ECOCE. https://educa-verde.ecoce.mx/

López, P. (2019) *Políticas públicas para la inclusión financiera de las mujeres para la movilidad social en México*. (1ª ed) Centro de Estudios Espinosa Yglesias A. C. https://ceey.org.mx/wp-content/uploads/2020/08/Inclusio%CC%81n-Financiera-final.pdf

Encuesta Nacional de Salud y Nutrición. https://ensanut.insp.mx

Informe país sobre la calidad de la ciudadanía en México. https://portalanterior.ine.mx/archivos2/portal/DECEYEC/EducacionCivica/informePais/

Encuesta Nacional de Hábitos de Reciclaje de Plásticos. https://www.ecoce.mx/files/boletines/40_informe.pdf

Encuesta Nacional sobre Discriminación. https://www.inegi.org.mx/programas/enadis/2022/

Páginas de internet y simuladores

Estadística y Probabilidad I

Applet for simulation and theory-based analysis of one binary variable. (n.d.). RossmanChance. https://www.rossmanchance.com/applets/2021/oneprop/OneProp.htm

CalculatorSoup. (2023). *Descriptive statistics calculator*. https://www.calculatorsoup.com/calculators/statistics/descriptivestatistics.php

Descriptive statistics Applet. (n.d.). RossmanChance. https://www.rossmanchance.com/applets/2021/descstats/Dotplot.htm

Guess the correlation Applet. (n.d.). RossmanChance. https://www.rossmanchance.com/applets/2021/guesscorrelation/GuessCorrelation.html

Monty hall game. (n.d.). RossmanChance. https://www.rossmanchance.com/applets/2021/montyhall/Monty.html

Squares Regression-least 1.1.32. (n.d.). PhET: Free online physics, chemistry, biology, earth science and math simulations. https://phet.colorado.edu/sims/html/least-squares-regression/latest/least-squares-regression_all.html?locale=es

T-TesT, CHI-square, ANOVA, regression, correlation... (n.d.). Online Statistics Calculator: Hypothesis testing, t-test, chi-square, regression, correlation, analysis of variance, cluster analysis. https://datatab.net/statistics-calculator/descriptive-statistics

Two quantitative variables. (n.d.). RossmanChance. https://www.rossmanchance.com/applets/2021/regshuffle/regshuffle.htm

Estadística y Probabilidad II

Amar, R. *(n.d.). Simulating confidence intervals. Stapplets*. https://www.stapplet.com/confint.html

CLT for means. (n.d.). https://openintro.shinyapps.io/CLT_mean/

Confidence intervals of the mean. (n.d.). Department of Zoology at UBC. https://www.zoology.ubc.ca/~whitlock/Kingfisher/CIMean.htm

Confidence intervals. (n.d.). Digital First subtypes on staging server (digitalfirst.bfwpub.com). https://digitalfirst.bfwpub.com/stats_applet/stats_applet_4_ci.html

Confidence intervals. (n.d.). Digital First subtypes on staging server (digitalfirst.bfwpub.com). https://digitalfirst.bfwpub.com/stats_applet/stats_applet_4_ci.html Coverage. (n.d.). https://istats.shinyapps.io/ExploreCoverage/ Kunin, D. (n.d.). Seeing Theory. https://seeing-theory.brown.edu/

Matt Bognar. *(n.d.). Mathematical Sciencess--College of Liberal Arts & Sciences, The University of Iowa*. https://homepage.divms.uiowa.edu/~mbognar/

Normal probability calculator. (n.d.). RossmanChance. https://www.rossmanchance.com/applets/2021/normcalc/NormCalc.html

Plinko Probability. (n.d.). PhET: Free online physics, chemistry, biology, earth science and math simulations. https://phet.colorado.edu/sims/html/plinko-probability/latest/plinko-probability_all.html?locale=es

Simulating confidence intervals. (n.d.). RossmanChance. https://www.rossmanchance.com/applets/2021/confsim/ConfSim.html

DIRECTORES DE PLANTELES

PROGRAMAS
DE ESTUDIO
2024

Escuela Nacional Colegio de Ciencias y Humanidades.
Los Programas de Estudio del Área de Matemáticas
se terminaron de imprimir en el mes de julio de 2024.

COLEGIO DE CIENCIAS Y HUMANIDADES ®

PROGRAMAS DE ESTUDIO 2024