

PENSAMIENTO MATEMÁTICO I

BACHILLERATO SEMIESCOLARIZADO

UAS



Faustino Vizcarra Parra
Rolando Alberto Forneiro Rodríguez
Victoria Bárbara Arencibia Sosa
Armando Flórez Arco

Contenido

Dedicatoria y agradecimientos	2
Presentación	4
Tabla de categorías, subcategorías, aprendizajes de trayectoria y metas de aprendizaje de Pensamiento Matemático I	6
Progresión de aprendizaje 1. El cálculo de probabilidades y la hipótesis de equiprobabilidad	8
Progresión de aprendizaje 2. La probabilidad y las técnicas de conteo	20
Progresión de aprendizaje 3. La probabilidad condicionada	28
Progresión de aprendizaje 4. Recolección de datos estadísticos	36
Progresión de aprendizaje 5. Representación de la información	46
Progresión de aprendizaje 6. La relación entre variables categóricas	62
Progresión de aprendizaje 7. La relación entre variables cuantitativas	73
Progresión de aprendizaje 8. Los valores atípicos y las variables de confusión en afirmaciones estadísticas y gráficas	84
Progresión de aprendizaje 9. Estudio de la población a partir de una muestra	95
Progresión de aprendizaje 10. Los estudios observacionales y el diseño de experimentos	105
Progresión de aprendizaje 11. Las medidas estadísticas en el estudio de un fenómeno	118
Progresión de aprendizaje 12. La distribución normal	131
Autoevaluación para el aprendizaje	144
Coevaluación para el aprendizaje	145
Bibliografía consultada	146
Referencia a las fuentes de consulta de tablas	146
Referencia a las fuentes de consulta de figuras	148
Referencia a las fuentes de consulta de los códigos QR	153
Bibliografía de consulta para el estudiante y el profesor	154

Dedicatoria y agradecimientos

A nuestros queridos docentes.

Con profundo agradecimiento y admiración dedicamos este libro a aquellas y aquellos profesores excepcionales que, con su pasión por la enseñanza y su compromiso inquebrantable, han guiado la creación de estas páginas. Su dedicación en el desarrollo del pensamiento probabilístico y estadístico ha iluminado el camino a otros docentes del Nivel Medio Superior (NMS) de la Universidad Autónoma de Sinaloa (UAS).

Sus enseñanzas han sido como faros de sabiduría, iluminando mentes, inspirando la curiosidad y fomentando el gusto por el aprendizaje. En este sentido, este libro es un testimonio de ese arduo trabajo y devoción, y esperamos que sea bien recibido para formar las nuevas generaciones de estudiantes. Así, su uso en el proceso de enseñanza y aprendizaje nos invita a seguir aprendiendo y creciendo como un equipo de docentes, que ven en la innovación la importancia de incorporar la inteligencia artificial (IA) como un aliado.

Tomemos en cuenta que la integración de la inteligencia artificial en el proceso de enseñanza y aprendizaje emerge como un catalizador fundamental para el desarrollo del pensamiento probabilístico y estadístico. Al aprovechar sus capacidades, la IA puede facilitar el acceso al conocimiento, personalizar el aprendizaje, optimizar los métodos pedagógicos y evaluar los resultados. Además, el estudiante la puede utilizar como un tutor en su proceso de aprendizaje. En definitiva, se reconoce el impacto de la IA como un aliado del docente en el proceso educativo.

En agradecimiento por sembrar las semillas para el desarrollo del pensamiento matemático en las y los estudiantes del NMS de la UAS, les extendemos nuestro más sincero agradecimiento. Que este libro sea un tributo a su legado en la formación de mentes brillantes y pensadores en el quehacer de las matemáticas.

Colaboradores:

Paloma Sandoval Gámez	UAP C. U. Mochis
José Humberto Romero Fitch	
Christian Marcel López Nieblas	
Ramiro Amezcua Reyes	
Heriberto Carlos Ayala Cruz	
Isaías López Romero	
Oscar Mauricio Heredia Ruiz	
Horacio Gabriel López Ramírez	
María del Pilar Madrid Solís	
Clarissa López Aboyte	UAP Casa Blanca
Jorge Aldivar Contreras Espinoza	
Lorena Leal Montoya	
María Del Rosario Llanes Molinero	UAP Central Diurna
Eva Angelina Martínez Campaña	

Nadiezhdá Román Valenzuela	
Luis Felipe Flores Tirado	UAP Dr. Salvador Allende
Adán Meza Sánchez	UAP El Rosario
Jorge Ramos Martínez	UAP Escuinapa
Juana María Armenta Trasviña	
Silvia Bojórquez Soto	UAP Guasave Diurna
Reyna Jesús Trasviña López	
Sandra Araceli Arreola Mora	UAP Heraclio Bernal
César Pilar Quintero Campos	
Martín Luna Belmar	UAP La Cruz
Abril Liseth Fierro Romero	
Jesús Antonio García Duarte	
Yadira Esmeralda Gutiérrez Esquivel	
Adriana Gutiérrez García	UAP Los Mochis
Edith Ivett Ocampo Manjarrez	
Paola Elifelet Reyes Álvarez	
Erick Eduardo Romero Gómez	
Zayto Baltazar Peñuelas Borboa	
Izeth Sarai Rivera Diaz	UAP Los Mochis Extensión Macapule
Sumiko Martínez Campaña	UAP Rafael Buelna Tenorio
Asia Cecilia Carrasco Valenzuela	
Iliana Tirado Olivas	UAP Rubén Jaramillo
Eva Edith Verdugo Serrano	
Alma Rosario Gámez Vázquez	UAP Ruiz Cortines
Daniela Castro Miranda	
Daniel Martínez Campaña	UAP Semiescolarizada
Jorge Radney Montgomery Leyva	UAP Valle del Carrizo
Anarelli Corona Cárdenas	
Fernando Tomás Gil Camacho	UAP Victoria del Pueblo
Yoanna Marisol Mercado Lizarde	
María Esther Franco González	UAP Vladimir Ilich Lenin
Carmen Leonor Castro Millán	
Fernando Eleazar Acosta Cruz	DGEP

Presentación

Es un placer presentarles este libro de Pensamiento Matemático I para el bachillerato semiescolarizado de la UAS (adaptado del libro Pensamiento Matemático I para el bachillerato escolarizado de la UAS), que ha sido cuidadosamente diseñado para acompañar a las y los estudiantes de bachillerato semiescolarizado en su fascinante travesía por el mundo de esa maravillosa forma matemática de pensar denominada pensamiento matemático, que proporciona una base sólida y estimulante para el aprendizaje.

Así, este libro es para utilizarse en la Unidad de Aprendizaje Curricular (UAC) Pensamiento Matemático I del Recurso Sociocognitivo Pensamiento Matemático, correspondiente al primer semestre del componente fundamental y extendido del plan de estudios (UAS, 2024) del Currículo del bachillerato de la Universidad Autónoma de Sinaloa 2024 que, de acuerdo con el Marco Curricular Común de la Educación Media Superior (MCCEMS) establecido por la Secretaría de Educación Pública (SEP, 2023a), enfatiza el desarrollo del pensamiento matemático.

El pensamiento matemático, según el MCCEMS, se define como: un Recurso Sociocognitivo que involucra diversas actividades cognitivas que van desde la ejecución de operaciones y el desarrollo de procedimientos y algoritmos hasta abarcar procesos mentales abstractos, incluida la intuición, que se dan cuando el sujeto participa del quehacer matemático al resolver problemas, usar o crear modelos, elaborar tanto conjeturas como argumentos y organizar, sustentar y comunicar sus ideas. (SEP, 2023c, p. 17)

La secuencia de este libro está basada en progresiones de aprendizaje, cada una diseñada para desarrollar sobre la anterior un pensamiento matemático; para el caso particular de esta UAC, un pensamiento probabilístico y estadístico.

En el sentido anterior, las progresiones de aprendizaje (PA) de la UAC Pensamiento Matemático I desarrollan el pensamiento probabilístico y estadístico para el logro de las metas de aprendizaje en la siguiente secuencia:

- PA 1. El cálculo de probabilidades y la hipótesis de equiprobabilidad.
- PA 2. La probabilidad y las técnicas de conteo.
- PA 3. La probabilidad condicionada.
- PA 4. Recolección de datos estadísticos.
- PA 5. Representación de la información.
- PA 6. La relación entre variables categóricas.
- PA 7. La relación entre variables cuantitativas.
- PA 8. Los valores atípicos y las variables de confusión en afirmaciones estadísticas y gráficas.
- PA 9. Estudio de una población a partir de una muestra.

- PA 10. Los estudios observacionales y el diseño de experimentos.
- PA 11. Las medidas estadísticas en el estudio de un fenómeno.
- PA 12. La distribución normal.

Bajo esta lógica del proceso de desarrollo del pensamiento matemático, las progresiones de aprendizaje están estructuradas y secuenciadas, en el sentido de que cada una es más compleja que la anterior, de acuerdo al nivel de pensamiento matemático que demande cada progresión. Cada una de ellas, se inicia con una evaluación diagnóstica; luego, le sigue una situación contextualizada, ejemplos formativos y una evaluación formativa diseñadas atendiendo a las subcategorías de las categorías del pensamiento matemático, mismas que orientan hacia el logro de las metas de aprendizaje; al final cuenta con instrumentos para la autoevaluación y coevaluación.

Además, en cada PA se consideran tres momentos claves de la evaluación: diagnóstica, formativa (mientras se aprende) y final; haciendo énfasis en la evaluación formativa en función de la retroalimentación, para que, durante el proceso de realizar las actividades de aprendizaje, las y los docentes puedan determinar el nivel de logro por los estudiantes, en particular, de las metas de aprendizaje que contribuyen a los aprendizajes de trayectoria. Es decir, se utiliza la evaluación formativa como herramienta para comprender su progreso y ajustar, en consecuencia, las estrategias activas.

También, durante el proceso de aprendizaje, en cada PA se lleva a cabo la autoevaluación (A), coevaluación (C) y heteroevaluación (H); para ello, se implementa como técnica principal de evaluación la observación, utilizando guías específicas para tal fin. Los resultados se reflejarán en la tabla que aparece al inicio de cada progresión en correspondencia con el desempeño. Por otra parte, se sugiere usar los códigos QR (generados en parzibyte: <https://parzibyte.me/apps/generador-qr/>), así como la Inteligencia Artificial y las aplicaciones de celular como aliados en este proceso de aprendizaje.

Finalmente, el desarrollar un pensamiento matemático no solo les abrirá las puertas en el aula, sino que también los acompañará a lo largo de sus vidas, dotándoles de la capacidad de enfrentar cualquier desafío con ingenio y perspicacia.

¡Adentrémonos juntos en el fascinante universo del pensamiento matemático!

Tabla de categorías, subcategorías, aprendizajes de trayectoria y metas de aprendizaje de Pensamiento Matemático I

Pensamiento Matemático I			
Categorías			
C1 Procedural	C2 Procesos de intuición y razonamiento	C3 Solución de problemas y modelación	C4 Interacción y lenguaje matemático
Subcategorías			
S1 Elementos aritmético-algebraicos	S1 Capacidad para observar y conjeturar	S1 Uso de modelos	S1 Registro escrito, simbólico, algebraico e iconográfico
S2 Elementos geométricos	S2 Pensamiento intuitivo	S2 Construcción de modelos	S2 Negociación de significados
S4 Manejo de datos e incertidumbre	S3 Pensamiento formal	S3 Estrategias heurísticas y ejecución de procedimientos no rutinarios	S3 Ambiente matemático de comunicación
Aprendizajes de Trayectoria			
Valora la aplicación de procedimientos automáticos y algorítmicos, así como la interpretación de sus resultados para anticipar, encontrar y validar soluciones a problemas matemáticos, de áreas del conocimiento y de su vida personal.	Adopta procesos de razonamiento matemático tanto intuitivos como formales tales como observar, intuir, conjeturar y argumentar, para relacionar información y obtener conclusiones de problemas (matemáticos, de las ciencias naturales, experimentales y tecnología, sociales, humanidades y de la vida cotidiana).	Modela y propone soluciones a problemas tanto teóricos como de su entorno, empleando lenguaje y técnicas matemáticas.	Explica el planteamiento de posibles soluciones a problemas y la descripción de situaciones en el contexto que les dio origen empleando lenguaje matemático y lo comunica a sus pares para analizar su pertinencia.
Metas de Aprendizaje			
M1-C1 Ejecuta cálculos y algoritmos para resolver problemas matemáticos, de las ciencias y de su entorno.	M1-C2 Observa y obtiene información de una situación o fenómeno para establecer estrategias o formas de visualización que ayuden a entenderlo.	M1-C3 Selecciona un modelo matemático por la pertinencia de sus variables y relaciones para explicar una situación, fenómeno o resolver un problema tanto teórico como de su contexto.	M1-C4 Describe situaciones o fenómenos empleando rigurosamente el lenguaje matemático y el lenguaje natural.

<p>M2-C1 Analiza los resultados obtenidos al aplicar procedimientos algorítmicos propios del pensamiento matemático en la resolución de problemáticas teóricas y de su contexto.</p>	<p>M2-C2 Desarrolla la percepción y la intuición para generar conjeturas ante situaciones que requieran explicación o interpretación.</p>	<p>M2-C3 Construye un modelo matemático, identificando las variables de interés, con la finalidad de explicar una situación o fenómeno y/o resolver un problema tanto teórico como de su entorno.</p>	<p>M2-C4 Socializa con sus pares sus conjeturas, descubrimientos o procesos en la solución de un problema tanto teórico como de su entorno.</p>
<p>M3-C1 Comprueba los procedimientos usados en la resolución de problemas utilizando diversos métodos, empleando recursos tecnológicos o la interacción con sus pares.</p>	<p>M3-C2 Compara hechos, opiniones o afirmaciones para organizarlos en formas lógicas útiles en la solución de problemas y explicación de situaciones y fenómenos.</p>	<p>M3-C3 Aplica procedimientos, técnicas y lenguaje matemático para la solución de problemas propios del pensamiento matemático, de áreas de conocimiento, recursos sociocognitivos, recursos socioemocionales y de su entorno.</p>	
	<p>M4-C2 Argumenta a favor o en contra de afirmaciones acerca de situaciones, fenómenos o problemas propios de la matemática, de las ciencias o de su contexto.</p>	<p>M4-C3 Construye y plantea posibles soluciones a problemas de áreas de conocimiento, recursos sociocognitivos, recursos socioemocionales y de su entorno, empleando técnicas y lenguaje matemático.</p>	

El cálculo de probabilidades y la hipótesis de equiprobabilidad

Progresión de aprendizaje 1

Reconoce la incertidumbre como consecuencia de la variabilidad e identifica la equiprobabilidad como una hipótesis que, si se puede asumir, facilita el estudio de la probabilidad y observa que, al incrementar el número de repeticiones de una simulación, la frecuencia del evento estudiado tiende a aproximarse a su probabilidad teórica.

Metas de aprendizaje	En proceso de logro	Bueno	Sobresaliente
M1-C1 Ejecuta cálculos y algoritmos para resolver problemas matemáticos, de las ciencias y de su entorno.	A		
	C		
	H		
M1-C3 Selecciona un modelo matemático por la pertinencia de sus variables y relaciones para explicar una situación, fenómeno o resolver un problema tanto teórico como de su contexto.	A		
	C		
	H		
M1-C4 Describe situaciones o fenómenos empleando rigurosamente el lenguaje matemático y el lenguaje natural.	A		
	C		
	H		

Evaluación diagnóstica 1.1

1. En la vida cotidiana se presentan fenómenos deterministas o fenómenos aleatorios. Los deterministas son aquellos que, realizados en las mismas circunstancias, siempre se obtiene el mismo resultado. Por ejemplo: si lanzas una pelota hacia arriba, esta subirá durante un periodo de tiempo y luego caerá al suelo. Respecto a los aleatorios, estos dan lugar a varios resultados, sin saber con certeza cuál de ellos va a resultar antes de que ocurra, por ejemplo: si lanzas un dado al aire, no sabes qué número saldrá. Analiza cuáles de los siguientes sucesos consideras son deterministas (D) o aleatorios (A).
 - a) Congelar una botella de agua en el refrigerador: _____
 - b) Sacar una bola par de una bolsa con 10 bolas numeradas del 1 al 10: __
 - c) Personas que van al cine en un día: _____
 - d) Voltar una ficha de dominó y que esta sea la ficha “doble 5”: _____

2. ¿Qué entiendes como evento equiprobable?
 - a) Evento o proceso cuyo resultado no se puede predecir con certeza.
 - b) Son los experimentos cuyos resultados pueden predecirse.
 - c) Son los resultados que tienen la misma posibilidad de ocurrir.

No todos los eventos que se nos presentan tienen un carácter aleatorio o determinista. El contexto en el que te encuentras impone situaciones que pueden presentar variabilidad y, consecuentemente, cierta incertidumbre en las que, de acuerdo con tus gustos, necesidades, informaciones, consideraciones o criterios, habrás de tomar una decisión.

Una forma de ayudarnos a tomar decisiones, de manera más objetiva, es considerar e interpretar la variabilidad que puede tener un *fenómeno* o la incertidumbre sobre su manifestación, lo cual ayudaría a decidir sobre el mismo, ya sea de manera intuitiva o analizando información.

Así pues, bajo ciertas condiciones, para tomar decisiones se requiere de un pensamiento sobre lo probable o *pensamiento probabilístico*; la *incertidumbre* es la falta de conocimiento o información precisa sobre las diferentes probabilidades de ocurrencia de un suceso, mientras que la *variabilidad* se refiere a la heterogeneidad o a la variación intrínseca de una cantidad o un fenómeno que propician resultados con diferentes valores de probabilidad.

La *variabilidad* en un fenómeno puede producir *incertidumbre*. Cuando un fenómeno es variable, sus características o resultados pueden cambiar o fluctuar en diferentes situaciones o momentos. Esta variabilidad puede generar incertidumbre porque no se puede predecir con certeza cómo se comportará el fenómeno en cada caso específico.

En ese sentido, *las probabilidades*, como rama de las matemáticas, resultan de gran utilidad cuando se tiene una base de información que nos permita hacer inferencias útiles para la toma de decisiones; una recopilación de datos estadísticos nos proporcionará elementos objetivos, que pueden ser utilizados, para conocer la probabilidad de ocurrencia de un evento y así tener un mejor criterio al identificar y evaluar las alternativas de solución.

El siguiente tema, por ejemplo, precisa la recolección de datos para su posterior análisis: **“La participación de la ciudadanía en el cuidado y promoción de la salud: el dengue”**

La salud de las personas y de las poblaciones no depende exclusivamente de las intervenciones del sistema sanitario, sino que también es susceptible a los determinantes biológicos, ambientales y sociales. Por ello, resulta acertado informarnos, acudiendo a fuentes de consultas y explorando los datos para hacer las interpretaciones necesarias al aplicar procedimientos adecuados y determinar las probabilidades a partir de un *enfoque frecuencial*.

La incertidumbre en la variabilidad respecto al dengue se relaciona con la imprevisibilidad de la propagación y el brote de la enfermedad ocasionada por factores como el clima y el agua estancada que propician la reproducción del mosquito. Esto

dificulta la predicción precisa de los momentos y lugares donde se producirán los brotes. Analizar los datos de una tabla sobre los casos confirmados de dengue puede servir para identificar patrones, tendencias y áreas de mayor incidencia de la enfermedad, lo que puede ayudar en la toma de decisiones para implementar medidas de prevención y control.

La Tabla 1.1 muestra datos estadísticos sobre los casos confirmados de dengue por cada municipio del estado de Sinaloa, comparando los años 2022 y 2023. Con base en ellos, es posible ordenar la información para reconocer datos relevantes y encontrar algún patrón en la información, para lo cual es pertinente e imprescindible introducir la frecuencia absoluta (f_a) y la frecuencia relativa (f_r):

- **Frecuencia absoluta (f_a).** Es el número de veces que se presenta cada modalidad.
- **Frecuencia relativa (f_r).** Es el valor que representa la proporción del número de veces que se repite el dato en la colección.

Se obtiene dividiendo la frecuencia absoluta entre el total de datos, $f_r = \frac{f_a}{N}$.

Jurisdicción		Casos confirmados de dengue en plataforma vectores, comparativo 2022-2023 acumulados hasta semana 39								
		Municipio	Dengue no grave		Dengue con signos de alarma		Dengue grave		Total	
			2022	2023	2022	2023	2022	2023	2022	2023
1-1	Ahome	11	2	25	7	0	1	36	10	
	Choix	0	0	0	0	5	0	5	0	
	El Fuerte	12	1	13	2	2	0	27	3	
1-2	Guasave	7	0	5	2	0	0	12	2	
	Sinaloa	1	0	0	0	0	0	1	0	
1-3	Angostura	1	0	2	1	0	0	3	1	
	Mocorito	0	0	3	0	0	0	3	0	
	S. Alvarado	7	1	19	3	0	0	26	4	
1-4	Badiraguato	10	0	9	0	2	0	21	0	
	Culiacán	34	5	45	4	2	1	81	10	
	Navolato	7	0	8	0	0	0	15	0	
	Cósala	0	0	0	0	0	0	0	0	
1-5	Elota	0	0	0	1	0	0	0	1	
	San Ignacio	1	0	0	0	0	0	1	0	
	Concordia	0	0	0	1	0	0	0	1	
	Mazatlán	3	5	6	4	1	0	10	9	
1-6	Escuinapa	11	45	3	14	1	3	15	62	
	Rosario	0	0	1	0	0	0	1	0	
Total		105	59	139	39	13	5	257	103	

Tabla 1.1. Casos confirmados de dengue correspondiente a la semana epidemiológica No. 39. de los años 2022 y 2023.

Fuente: Secretaría de Salud Sinaloa, 2017.

Así, con base en la información de la Tabla 2.1, se puede representar de manera conveniente los datos de las variables donde se describe el número de veces que se presenta cada evento (f_a) y la parte que representa respecto al total de la población (f_r) en 2023.

Ello permite diferenciar no sólo los lugares donde se presentan la mayor cantidad de casos de dengue, sino también la proporción de los mismos respecto a la población.

Municipio	Casos totales 2023 f_a	$f_r = \frac{f_a}{N}$
Ahome	51	0.0001
Choix	0	0
El fuerte	4	0.000035
Guasave	11	0.000033
Sinaloa	0	0
Angostura	1	0.000019
Mocorito	0	0
S. Alvarado	4	0.000044
Badiraguato	0	0
Culiacán	22	0.000021
Navolato	0	0
Cósala	0	0
Elota	3	0.000065
San Ignacio	2	0.000083
Concordia	2	0.000061
Mazatlán	33	0.000063
Escuinapa	104	0.00163
Rosario	0	0

Ejemplo formativo 1.1

Veamos otro ejemplo en el que es importante considerar la frecuencia relativa. En el beisbol, el promedio ofensivo de bateo de un bateador se calcula dividiendo la cantidad de hits que ha bateado entre las veces que ha ido al bate en forma acumulativa. Si en una temporada de juegos, ha bateado 91 hits en 275 turnos al bate, su promedio ofensivo es $91/275 \approx 0.33$. Ello significa que la probabilidad de dar un hit, ese bateador cada vez que va al bate es de 0.33 o bien que da un hit una de cada tres veces que batea.

Si en un juego el bateador ha fallado dos veces consecutivas, ¿eso significa que la próxima vez que esté al bate dará un hit? No, eso no significa que la próxima vez que esté al bate dará un hit. El promedio de bateo es una medida de la probabilidad de dar un hit, no una predicción de lo que va a pasar en cada turno al bate. La probabilidad de dar un hit es la misma en cada turno al bate, independientemente de lo que haya pasado antes. El promedio de bateo solo refleja el desempeño histórico del bateador.

Un jugador de beisbol puede participar en cinco, seis o más temporadas de juegos y, por tanto, tener un promedio ofensivo de bateo acumulado que ofrece una información más fidedigna sobre la consistencia de dar hit cuando batea.

Yasmany Tomás, cubano, es un jugador activo de beisbol en la Liga Mexicana del Pacífico; en Los Cañeros de Los Mochis y fue elegido el Jugador más Valioso de la Temporada 22-23 en dicha liga¹.

En dicha temporada tuvo $0.328 \approx 0.33$ como promedio ofensivo de bateo, lo que significa que, como promedio, da un hit cada tres veces que batea. En la tabla siguiente aparece su promedio ofensivo durante cuatro temporadas anteriores:

Temporada	J	AB	R	H	2B	3B	HR	RBI	BB	HBP	K	SB	CS	AVG	OBP	SLG	OPS	WAR
2015	118	406	40	111	19	3	9	48	17	2	110	5	2	.273	.305	.401	.707	-1.3
2016	140	530	72	144	30	1	31	83	31	1	136	2	4	.272	.313	.508	.820	-0.4
2017	47	166	19	40	11	1	8	32	13	0	50	0	0	.241	.294	.464	.758	-0.5
2019	4	6	0	0	0	0	0	0	0	0	3	0	0	.000	.000	.000	.000	-0.2
Carrera	309	1108	131	295	60	5	48	163	61	3	299	7	6	.266	.306	.459	.765	--
Promedios de temporada	77	277	33	74	15	1	12	41	15	1	75	2	2	.266	.306	.459	.765	-

Tabla 1.2. Promedio ofensivo durante cuatro temporadas de Yasmany Tomás.

Fuente: https://www.espn.com.mx/beisbol/mlb/jugador/estadisticas/_id/33662/yasmany-tomas

En la penúltima fila se aprecia, que en 1,108 veces al bate conectó 295 hits para un promedio acumulado de $0.266 > 0.25$, lo que significa que conectaba de hits, aproximadamente algo más de una vez cada cuatro veces.

Analizando su desempeño en la temporada 22-23 respecto a su frecuencia de bateo histórica, en las cuatro temporadas analizadas, se puede inferir que ha mejorado su técnica de bateo, lo cual resulta de utilidad para la dirección del equipo y decidir sobre mantenerlo contratado por su desempeño.

Consideremos ahora la siguiente situación: en El Rosario, Sinaloa, una de las principales fuentes de empleo de las familias está en las diferentes empacadoras, cuyo producto principal es el mango, generalmente como un producto no procesado, mismo que es seleccionado y empacado, atendiendo a las diferentes normas exigidas por los mercados internacionales; uno de los principales requisitos, entre otros, es la presentación adecuada del producto, el cual se empaca en cajas que deben presentar ciertas características para que puedan cubrir las exigencias de los diferentes importadores del mundo.

En esas empresas trabajan personas encargadas de realizar proyecciones por medio de modelos probabilísticos, que permiten apreciar cuánto mango de buena calidad llegará aproximadamente cada mes. En lo que sigue vas a estudiar una forma en la que pueden estimar esa proyección.

1.1 Eventos equiprobables y no equiprobables

La equiprobabilidad es un concepto que se utiliza en ciertos eventos aleatorios, en los que se conoce el número de resultados posibles y cada uno de los n casos posibles tiene la misma posibilidad de ocurrir; por ejemplo, al lanzar una moneda, se sabe que podemos obtener dos resultados posibles, $n = 2$, y se tiene $1/2 = 0.5$, o bien 50%, de

¹ <https://www.septimaentrada.com/liga-mexicana-del-pacifico/lmp-yasmany-tomas-jugador-valioso-2022-23>

probabilidad de obtener cada uno de los dos resultados; las dos situaciones son igualmente probables de ocurrir.

Otro ejemplo sobre equiprobabilidad sería el lanzamiento de un dado, suponiendo que no está cargado, pesa y mide lo mismo en todas sus caras; la diferencia con el caso anterior es que $n = 6$; la probabilidad de que ocurra cada evento es igual a $1/6 = 0.167$, o bien 16.7%, pero como en el caso anterior, cada uno de los eventos posibles tienen la misma probabilidad de ocurrir.

Si dos eventos tienen la misma posibilidad de ocurrir se dice que son *equiprobables* y $P(A) = P(B)$; es decir, la probabilidad de que ocurra A es igual a la probabilidad de que ocurra B . En cambio, los eventos "*no equiprobables*" son aquellos donde la probabilidad de que ocurra el evento A , es diferente a que ocurra la del evento B , es decir $P(A) \neq P(B)$.

Así, lanzar un dado que no esté cargado, dos veces y que caiga cada vez el número 3, es un evento equiprobable y lanzar un dado que esté cargado, dos veces y que caiga cada vez el número 3, es un evento no equiprobable.

Hay fenómenos que se pueden modelar matemáticamente, por una relación causa-efecto, por ejemplo, determinar la distancia que recorre un vehículo a velocidad constante si se conoce el tiempo en que se va a mover; otros no, como saber qué día de la semana puede llover más. Si bien no es posible una descripción causa-efecto, eso no significa que no se pueda encontrar una forma de referir matemáticamente este tipo de fenómenos (llamados, como ya sabes, aleatorios o producto del azar); para ello se estudia la frecuencia con que ocurren o pueden ocurrir y se emplea una descripción estadística, con un concepto matemático central: el de *probabilidad*.

1.2 Probabilidad teórica y probabilidad frecuencial

Para explicar la probabilidad teórica partiremos de un ejemplo específico: lanzar un dado que no se encuentre cargado, y el evento deseado es que se obtenga el número 3. El número total de eventos posibles N es igual a 6, ya que $N = \{1, 2, 3, 4, 5, 6\}$ y el número de eventos favorables es $n = 1$, porque el evento deseado A es sólo el número 3.

Por tanto, la probabilidad de que salga el número 3 es igual a:

$$P(3) = \frac{1}{6} = 0.167,$$

y este resultado es lo que conocemos como **probabilidad teórica**:

$$P(A) = \frac{\text{número de eventos favorables}}{\text{número de casos posibles}} = \frac{n}{N}$$

También la probabilidad suele expresarse en tanto por ciento, para lo cual el valor obtenido al calcular la probabilidad se multiplica por 100, es decir, $P(A) = \frac{n}{N} \cdot 100$, por lo que en el ejemplo anterior $P(3) = 16.7\%$.

Ejemplo formativo 1.2

Un juego de barajas trae 52 cartas y estas se dividen en 4 grupos de 13 cartas cada uno: picas (♠), corazones (♥), diamantes (♦) y tréboles (♣). Las cartas de picas y tréboles son negras, mientras que las cartas de corazones y diamantes son rojas. Las cartas de cada grupo son: as, rey, reina, sota, 10, 9, 8, 7, 6, 5, 4, 3 y 2.

La extracción de una carta en un juego de barajas se basa en la probabilidad y se calcula dividiendo el número de cartas favorables por el número total de cartas en la baraja. Por tanto, la probabilidad de extraer:

- Un 10 de color rojo es $\frac{2}{52} = \frac{1}{26} = 0.038$ o 3,8 % (hay 2 dieces rojos en la baraja: uno de corazones y uno de diamantes)
- Una carta de color negro es $\frac{26}{52} = \frac{1}{2} = 0.5$ o 50 % (hay 26 cartas negras en total: picas y tréboles)
- Una carta de corazones es $\frac{13}{52} = \frac{1}{4} = 0.25$ o 25 % (hay 13 cartas de corazones)
- Un rey o una reina $\frac{8}{52} = \frac{2}{13} = 0.15$ o 15 % (hay 4 reyes y 4 reinas en total, sumando 8 cartas favorables)

Ahora, si el mismo experimento de lanzar un dado y querer obtener un tres lo llevas a cabo repetidas veces, podrás apreciar que la posibilidad de que se obtenga tres, no necesariamente coincide con la probabilidad teórica; pero, si el evento se repite un número considerable de veces, los valores cambiarán y se apreciará una tendencia: mientras más veces se repita el experimento, la posibilidad de obtener un tres se acercará a la probabilidad teórica. A este tipo de probabilidad se le denomina **probabilidad frecuencial**.

Para calcular esta probabilidad, es necesario realizar el experimento un número elevado de veces, ya que cuanto más se repite el experimento, los resultados serán más cercanos a la probabilidad teórica.

A diferencia de la probabilidad teórica, la probabilidad frecuencial de que ocurra un suceso A , se calcula al dividir el número de casos favorables (s) entre el número de veces que se repite el experimento (N):

$$P(A) = \frac{\text{número de casos favorables del experimento}}{\text{número total de intentos}} = \frac{s}{N}$$

Ahora podemos aplicar estos conceptos en el siguiente ejemplo.

Ejemplo formativo 1.3

En la siguiente tabla se muestra el experimento donde se realizaron 10 lanzamientos de un dado y los resultados, que incluyen la frecuencia relativa y la probabilidad frecuencial.

Número del dado	Probabilidad teórica $P(A) = \frac{1}{6}$	Frecuencia absoluta (f_a)	Frecuencia relativa (f_r)	Probabilidad como porcentaje (%)
1	0.167	1	0.100	10%
2	0.167	3	0,300	30%
3	0.167	1	0.100	10%
4	0.167	1	0.100	10%
5	0.167	2	0.200	20%
6	0.167	2	0,200	20%
Total	1	10	1	100%

Podrás observar que la probabilidad frecuencial de cada uno de los eventos posibles no se corresponde con la probabilidad teórica. A continuación, se repitió el experimento una mayor cantidad de veces, 100 veces, con los resultados siguientes:

Número del dado	Probabilidad teórica $P(A) = \frac{1}{6}$	Frecuencia absoluta (f_a)	Frecuencia relativa (f_r)	Probabilidad como porcentaje (%)
1	0.167	18	0.18	18%
2	0.167	18	0.8	18%
3	0.167	16	0.16	16%
4	0.167	19	0.19	19%
5	0.167	20	0.20	20%
6	0.167	9	0.09	9%
Total	1	100	1	100%

Según los resultados, se aprecia que hay cierta similitud entre la probabilidad teórica y la frecuencia relativa obtenida y en comparación con el lanzamiento del dado 10 veces, en 100 lanzamientos se obtienen valores más cercanos a la probabilidad teórica en la mayoría de los casos.

Ahora, la siguiente tabla muestra los resultados de completar hasta 1000 lanzamientos, calculando los datos de frecuencia absoluta, frecuencia relativa y probabilidad frecuencial.

Número del dado	Probabilidad teórica $P(A) = \frac{1}{6}$	Frecuencia absoluta (f_a)	Frecuencia relativa (f_r)	Probabilidad como porcentaje (%)
1	0.167	167	0.167	16.7%
2	0.167	173	0.173	17.3%
3	0.167	169	0.169	16.9%
4	0.167	176	0.176	17.6%
5	0.167	192	0.192	19.2%
6	0.167	123	0.123	12.3%
Total	1	1000	1	100%

Al analizar el comportamiento de los resultados, comparando la probabilidad teórica con la frecuencia relativa obtenida con 10, 100 y 1000 anotaciones de lanzamientos del dado, se concluye que en la medida que se ha aumentado el número de lanzamientos, las anotaciones se acercan más a la probabilidad teórica.

Evaluación formativa 1.1

1. Lanzamiento de una moneda para determinar la probabilidad teórica y frecuencial.
 - a) Calcula la probabilidad teórica de la ocurrencia de sol y de águila en el lanzamiento de una moneda.
 - b) Lanza una moneda un total de 20 veces y marca el resultado con una "X" en el siguiente formato (A significa águila; S significa sol):

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
A																				
S																				

- c) Responde lo siguiente:
 - i. Número de casos favorables para águila. _____
 - ii. Número de casos favorables para sol. _____
 - iii. Probabilidad frecuencial para águila. _____
- d) Ahora reúne los resultados junto con los de otros nueve compañeros (total 100 lanzamientos, $N = 100$) y vuelve a contestar lo siguiente.
 - i. Número de casos favorables para águila. _____
 - ii. Número de casos favorables para sol. _____
 - iii. Probabilidad frecuencial para águila. _____
- e) ¿A qué conclusión llegas? Compártela con tus compañeros.

2. En los empaques de El Rosario, llegan los camiones con jabas llenas de mangos recolectadas directamente de los diferentes huertos, en donde cada camión llega con aproximadamente 600 jabas, este mango se separa en mango de exportación y en mango de rezaga que no es apto para la exportación.

a) Los resultados de las 10 últimas semanas se muestran en la siguiente tabla. Con ellos, calcula la frecuencia relativa y la probabilidad en porcentaje.

Semana	Jabas recibidas	Jabas aptas	Jabas aptas		
			Porcentaje (%)	Frecuencia relativa (f_r)	Probabilidad (%)
1	5500	5055	91.91		
2	5320	4831	90.81		
3	5100	4529	88.80		
4	5120	4572	89.30		
5	5360	4824	90.00		
6	5600	5096	91.00		
7	5650	4972	88.58		
8	5500	5005	90.00		
9	5470	4923	90.00		
10	5510	4986	90.49		

b) Con dichos resultados los encargados del empaque quieren conocer la probabilidad de que en la semana en curso los resultados estén como mínimo en el 90%. Utilizando la probabilidad frecuencial calcula qué probabilidad hay de que los resultados cumplan esa expectativa.

c) En la semana actual cada camión llega con 600 jabas y se obtienen los siguientes datos:

- El lunes llega un camión con 90 jabas de rezaga.
- El martes dos camiones, uno con 18 y otro con 54 jabas de rezaga.
- El miércoles llega un camión con 66 jabas de rezaga.
- El jueves tres camiones con 60, 96 y 30 jabas de rezaga respectivamente.
- El viernes dos camiones, uno con 48 y el otro con 60 jabas de rezaga.

Confecciona una tabla en la que se relacione por cada día de la semana los camiones que llegaron a la empacadora, la cantidad de jabas de mango recibidas y de ellas, cuántas resultan de mangos aptos para exportar. Con dicha información, calcula el porcentaje respecto a cada camión, la frecuencia relativa y la probabilidad en por ciento.

Día	Núm. de camiones	Jabas recibidas	Jabas aptas	Jabas aptas		
				Porcentaje (%)	Frecuencia relativa (f_r)	Probabilidad (%)
Lunes						
Martes						
Miércoles						
Jueves						
Viernes						
Total						

d) ¿Cuál es el total de jabas de mangos recibidas y de ellas, el total de jabas aptas?

e) ¿Cuál es la probabilidad de jabas aptas? Expresa el resultado en porcentaje.

f) ¿Se cumplen las expectativas de los encargados del empaque? _____

Fundamenta tu respuesta: _____

Autoevaluación y coevaluación 1.1

Nombre: _____ Plantel: _____ Grupo: _____ Turno: _____

Autoevaluación para el aprendizaje

Selecciona en la columna la opción que mejor refleje tu nivel de desempeño en el proceso para el aprendizaje de la progresión de aprendizaje 1. Responde con honestidad a la evaluación de cada uno de los criterios que se enlistan.

Desempeño	En proceso de logro	Bueno	Sobresaliente
Propicié un clima de comunicación favorable para el aprendizaje con mis compañeros.			
Participé activamente con ideas para la toma razonada de decisiones.			
Contribuí colaborativamente en la retroalimentación de dudas de mis compañeros.			
Apliqué el incremento de la frecuencia de ocurrencia de un evento para comprobar su aproximación a la probabilidad teórica de que ocurra (M1-C1).			
Entendí la diferencia entre eventos equiprobables y no equiprobables (M1-C3).			
Interpreté adecuadamente, en una situación dada, la diferencia entre los casos posibles de ocurrencia de un suceso y el total de casos, para aplicar el concepto de probabilidad (M1-C4).			

Coevaluación para el aprendizaje

Solicita a un compañero del equipo que marque en la columna la opción que mejor describa tu desempeño durante el trabajo colectivo, concluida la progresión de aprendizaje 1, y que responda con honestidad la evaluación de cada uno de los criterios que se enlistan a continuación.

Desempeño	En proceso de logro	Bueno	Sobresaliente
Propició un clima de comunicación favorable para el aprendizaje con mis compañeros.			
Participó activamente con ideas para la toma razonada de decisiones.			
Contribuyó colaborativamente en la retroalimentación de dudas de mis compañeros.			
Aplicó el incremento de la frecuencia de ocurrencia de un evento para comprobar su aproximación a la probabilidad teórica de que ocurra (M1-C1).			
Supo apreciar la diferencia entre eventos equiprobables y no equiprobables (M1-C3).			
Interpretó adecuadamente, en una situación dada, la diferencia entre los casos posibles de ocurrencia de un suceso y el total de casos, para aplicar el concepto de probabilidad (M1-C4).			

Nombre y firma de quien coevalúa

La probabilidad y las técnicas de conteo

Progresión de aprendizaje 2

Elige una técnica de conteo (ordenaciones con repetición, ordenaciones, permutaciones, combinaciones) para calcular el número total de casos posibles y casos favorables para eventos simples con la finalidad de hallar su probabilidad y con ello generar una mayor conciencia en la toma de decisiones.

Las técnicas de conteo se introducen para entender la probabilidad de eventos aleatorios en los que la expresión explícita de su espacio muestral es poco factible.

Metas de aprendizaje		En proceso de logro	Bueno	Sobresaliente
M2-C1 Analiza los resultados obtenidos al aplicar procedimientos algorítmicos propios del Pensamiento Matemático en la resolución de problemáticas teóricas y de su contexto.	A			
	C			
	H			
M3-C1 Comprueba los procedimientos usados en la realización de problemas utilizando diversos métodos, empleando recursos tecnológicos o la interacción con sus pares.	A			
	C			
	H			
M3-C3 Aplica procedimientos, técnicas y lenguaje matemático para la solución de problemas propios del Pensamiento Matemático, de Áreas de Conocimiento, Recursos Sociocognitivos, Recursos Socioemocionales y de su entorno.	A			
	C			
	H			

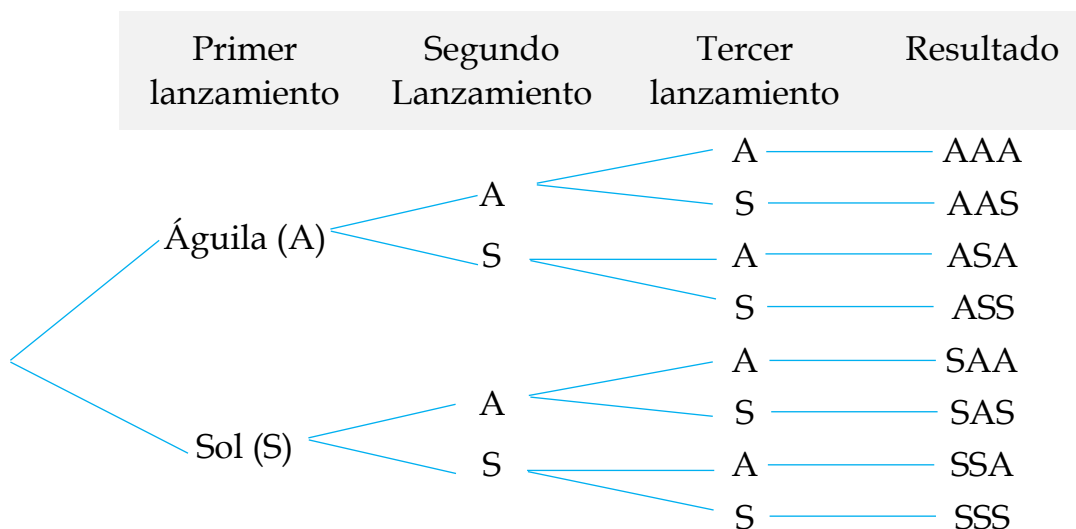
Evaluación diagnóstica 2.1. Resuelve los siguientes ejercicios:

- Si lanzas una moneda al aire tres veces, ¿qué probabilidad hay de obtener sol, águila, sol en ese orden?
- En una clase de Pensamiento Matemático I hay 10 estudiantes, de ellos, seis son mujeres y cuatro son hombres. El profesor quiere seleccionar un comité de tres miembros para organizar una actividad, pero no importa el orden en que sean seleccionados. ¿De cuántas formas diferentes se puede formar este comité?

En el inciso a) de la evaluación diagnóstica puede resolverse usando un diagrama de árbol como se muestra a continuación.

Si lanzas una moneda al aire tres veces, ¿qué probabilidad hay de obtener sol, águila, sol en ese orden?

Una forma de calcularla es dibujar un diagrama de árbol:



Luego, hay que identificar el número de casos posibles y de casos favorables. El total de casos posibles es ocho y hay una sola variante en que se obtiene SAS, luego la probabilidad de que se obtenga ese resultado es $P(SAS) = 1/8 = 0.125$.

Ahora, si lanzas una moneda al aire 10 veces, ¿cuántos resultados posibles hay?, ¿qué probabilidad hay de obtener sol, águila, sol, sol, águila, sol, sol, águila, sol, sol en ese orden?

El uso del diagrama de árbol se vuelve inoperante de dibujar por la cantidad grande de resultados posibles, en consecuencia, mediante esta técnica no es posible calcular la probabilidad:

$$P(SASSSASSSS) = \frac{\text{número de resultados favorables}}{\text{número total de resultados posibles}}$$

Entonces, ¿cómo la calcularías?

Más aún, si retomamos el inciso b) de la evaluación diagnóstica, con dos preguntas adicionales:

En una clase de Pensamiento matemático I hay 10 estudiantes, de ellos, seis son mujeres y cuatro son hombres. El profesor quiere seleccionar un presidente, un vicepresidente y un secretario del grupo.

- ¿De cuántas formas diferentes se pueden asignar estos tres cargos?
- Si Juan, María y Sofía son estudiantes de dicha clase, ¿cuál es la probabilidad de que Sofía sea la presidenta, Juan el vicepresidente y María la secretaria?
- ¿Cuál es la probabilidad de que la presidencia la ocupe un hombre, la vicepresidencia una mujer y el puesto de secretaria una mujer?

Con el conocimiento que tienes acerca del cálculo de probabilidades, ¿puedes resolverlo?

Técnicas de conteo

Una técnica de conteo se define como: un conjunto de métodos utilizados en combinatoria para determinar el número de maneras en que se pueden organizar, seleccionar o distribuir elementos dentro de un conjunto bajo ciertas condiciones. Estas técnicas, son importantes para el cálculo de probabilidades en los casos en que la técnica del diagrama de árbol, se vuelve inoperante para conocer el número total de casos posibles y el número de resultados favorables.

Las técnicas de conteo se usan en el cálculo de probabilidades, cuando la técnica del diagrama de árbol se vuelve ineficiente para conocer el número de resultados posibles, como consecuencia de un número grande de resultados. A continuación, veremos la técnica de permutaciones y de combinaciones, así como su uso en el cálculo de probabilidades.

Permutaciones

Las permutaciones de n elementos tomados de r en r , denotado como ${}_n P_r$, es el número de formas posibles de ordenar, r elementos seleccionados de un conjunto de n elementos distintos. En este tipo de permutaciones, el orden de los elementos seleccionados importa.

La fórmula para calcular el número de permutaciones de n elementos tomados de r en r :

$${}_n P_r = \frac{n!}{(n-r)!}, \text{ donde } n \geq r.$$

Dónde: $n!$ es el factorial de n y $(n-r)!$ es el factorial de $(n-r)$.

Ejemplo formativo 2.1

En una clase de Pensamiento matemático I hay 10 estudiantes, de ellos, seis son mujeres y cuatro son hombres. El profesor quiere seleccionar un presidente, un vicepresidente y un secretario del grupo.

- ¿De cuántas formas diferentes se pueden asignar estos tres cargos?
- Si Juan, María y Sofía son estudiantes de dicha clase, ¿cuál es la probabilidad de que Sofía sea la presidenta, Juan el vicepresidente y María la secretaria?
- ¿Cuál es la probabilidad de que la presidencia la ocupe un hombre, la vicepresidencia una mujer y el puesto de secretaria una mujer?

Resolución:

- ¿De cuántas formas diferentes se pueden asignar estos tres cargos?

El número total de estudiantes es $n = 10$ y el número de cargos posibles es $r = 3$.

Utilizando la fórmula de permutaciones de n en r :

$${}_{10}P_3 = \frac{10!}{(10-3)!} = \frac{10!}{7!} = \frac{10 \cdot 9 \cdot 8 \cdot 7!}{7!} = 10 \cdot 9 \cdot 8 = 720.$$

Por lo tanto, hay 720 formas diferentes de asignar los tres cargos entre los 10 estudiantes.

b) Si Juan, María y Sofía son estudiantes de dicha clase, ¿cuál es la probabilidad de que Sofía sea la presidenta, Juan el vicepresidente y María la secretaria?

Sea el evento $A = \{\text{Sofía la presidenta, Juan el vicepresidente y María la secretaria}\}$.

El número de evento favorables es uno y el número total de resultados posibles es 720.

$$P(A) = \frac{1}{720} \approx 0.0014$$

Por lo tanto, la probabilidad de que Sofía sea la presidenta, Juan el vicepresidente y María la secretaria es aproximadamente 0.14 %.

c) ¿Cuál es la probabilidad de que la presidencia la ocupe un hombre, la vicepresidencia una mujer y el puesto de secretaria una mujer?

Sea el evento $B = \{\text{la presidencia hombre, la vicepresidencia mujer y la secretaria mujer}\}$.

Para calcular esta probabilidad, consideramos los siguientes pasos:

1) Asignar la presidencia a un hombre:

Hay 4 hombres en la clase.

Número de formas de elegir al presidente:

$${}_4P_1 = \frac{4!}{(4-1)!} = \frac{4!}{3!} = \frac{4 \cdot 3!}{3!} = 4$$

2) Asignar la vicepresidencia a una mujer:

Después de asignar la presidencia a un hombre, quedan 9 estudiantes, de estos, 6 son mujeres.

Número de formas de elegir a la vicepresidenta:

$${}_6P_1 = \frac{6!}{(6-1)!} = \frac{6!}{5!} = \frac{6 \cdot 5!}{5!} = 6$$

3) Asignar el puesto de secretaria a una mujer:

Después de asignar los dos primeros cargos, quedan 8 estudiantes, de estos, 5 son mujeres.

Número de formas de elegir a la secretaria:

$${}_5P_1 = \frac{5!}{(5-1)!} = \frac{5!}{4!} = \frac{5 \cdot 4!}{4!} = 5$$

El número de formas favorables de asignar los cargos según las condiciones es:

$${}_4P_1 \cdot {}_6P_1 \cdot {}_5P_1 = 4 \cdot 6 \cdot 5 = 120$$

Y el número total de formas de asignar los tres cargos, calculado en el inciso a), es 720.

Así, la probabilidad de que la presidencia la ocupe un hombre, la vicepresidencia una mujer y el puesto de secretaria una mujer es:

$$P(B) = \frac{\text{número de resultados favorables}}{\text{número total de resultados posibles}} = \frac{120}{720} = \frac{1}{3} \approx 0.1667.$$

Por lo tanto, la probabilidad es aproximadamente del 16.67 %.

Combinaciones

Las combinaciones, es una técnica de conteo utilizada para determinar el número de formas en que se pueden seleccionar r elementos, de un conjunto de n elementos sin tener en cuenta el orden. A diferencia de las permutaciones, en las combinaciones el orden de los elementos seleccionados no importa.

La combinación de n elementos tomados de r en r , se denota como ${}_n C_r$. La fórmula para calcularla es:

$${}_n C_r = \frac{n!}{(n-r)! \cdot r!}, \text{ donde } n > r.$$

Donde: $n!$ es el factorial de n , $r!$ es el factorial de r y $(n-r)!$ es el factorial de $(n-r)$.

Ejemplo formativo 2.2

En una clase de Pensamiento Matemático I hay 10 estudiantes, de ellos, seis son mujeres y cuatro son hombres. El profesor quiere seleccionar un comité de tres miembros para organizar una actividad, pero no importa el orden en que sean seleccionados.

- ¿De cuántas formas diferentes se puede formar este comité?
- Si Juan, María y Sofía son estudiantes de dicha clase, ¿cuál es la probabilidad de que sean seleccionados para el comité?
- ¿Cuál es la probabilidad de que el comité esté compuesto por un hombre y dos mujeres?

Resolución:

- ¿De cuántas formas diferentes se puede formar este comité?

Para encontrar el número de formas diferentes de seleccionar un comité de tres miembros de un total de 10 estudiantes, sin importar el orden, usaremos combinaciones.

El número total de estudiantes es $n = 10$ y el número de lugares posibles es $r = 3$.

$${}_{10}C_3 = \frac{10!}{(10-3)! \cdot 3!} = \frac{10!}{7! \cdot 3!} = \frac{10 \cdot 9 \cdot 8 \cdot 7!}{7! \cdot 3!} = \frac{10 \cdot 9 \cdot 8 \cdot 7!}{3 \cdot 2 \cdot 1} = 120.$$

Por lo tanto, hay 120 formas diferentes de formar un comité de tres miembros.

b) Si Juan, María y Sofía son estudiantes de dicha clase, ¿cuál es la probabilidad de que sean seleccionados para el comité?

En las combinaciones, las siguientes ternas son la misma, debido a que no importa el orden.

Sofía, Juan y María
Juan, María y María

Sofía, María y Juan
María, Sofía y Juan

Juan, Sofía y María
María, Juan y Sofía

Sea el evento $A = \{\text{elegir tres integrantes: Sofía, Juan y María}\}$.

El número de evento favorables es uno y el número total de resultados posibles es 120.

$$P(A) = \frac{1}{120} \approx 0.0083$$

Por lo tanto, la probabilidad de que Sofía, Juan y María formen el comité es aproximadamente 0.83 %.

c) ¿Cuál es la probabilidad de que el comité esté compuesto por un hombre y dos mujeres?

Sea el evento $B = \{\text{un hombre y dos mujeres}\}$.

Para encontrar esta probabilidad, consideramos los siguientes pasos:

1) Seleccionar uno de cuatro hombres:

Número de formas de seleccionar un hombre de cuatro:

$${}_4C_1 = \frac{4!}{(4-1)! \cdot 1!} = \frac{4!}{3!} = \frac{4 \cdot 3!}{3!} = 4$$

2) Seleccionar dos de seis mujeres:

Número de formas de seleccionar un hombre de dos mujeres:

$${}_6C_2 = \frac{6!}{(6-2)! \cdot 2!} = \frac{6!}{4! \cdot 2!} = \frac{6 \cdot 5 \cdot 4!}{4! \cdot 2!} = \frac{6 \cdot 5}{2} = \frac{30}{2} = 15$$

El número de formas favorables de seleccionar un hombre y dos mujeres es:

$${}_4C_1 \cdot {}_6C_2 = 4 \cdot 15 = 60$$

El número total de formas de formar un comité de 3 miembros es 120.

Así, la probabilidad de que el comité esté compuesto por un hombre y dos mujeres es:

$$P(B) = \frac{\text{número de resultados favorables}}{\text{número total de resultados posibles}} = \frac{60}{120} = \frac{1}{2} \approx 0.5.$$

Por lo tanto, la probabilidad es aproximadamente del 50 %.

De manera general, para el cálculo de probabilidades usando permutaciones se pueden obtener mediante la fórmula:

$$P(A) = \frac{\text{número de resultados favorables}}{\text{número total de resultados posibles}} = \frac{{}_s P_t \cdot {}_{n-s} P_{r-t}}{{}_n P_r}, s < n, r < t.$$

También, de forma general, para el cálculo de probabilidades usando combinaciones se pueden obtener mediante la fórmula:

$$P(A) = \frac{\text{número de resultados favorables}}{\text{número total de resultados posibles}} = \frac{{}_s C_t \cdot {}_{n-s} C_{r-t}}{{}_n C_r}, s < n, r < t.$$

Diferencia entre permutaciones y combinaciones	
Permutaciones: son los arreglos de los elementos de un conjunto en los que sí importa la posición que ocupa cada uno de los elementos que integran el arreglo.	Combinaciones: son los arreglos de los elementos de un conjunto en los que no importa la posición que ocupa cada uno de los elementos que integran el arreglo.
${}_n P_r \neq {}_n C_r$	
${}_n P_r = \frac{n!}{(n-r)!}$	${}_n C_r = \frac{n!}{(n-r)! r!}$

Evaluación formativa 2.1

1. En una competencia de atletismo participan 12 corredores, de los cuales 7 son hombres y 5 son mujeres. Se deben otorgar medallas de oro, plata y bronce.
 - a) ¿De cuántas formas diferentes se pueden otorgar estas tres medallas?
 - b) Si Juan, María y Sofía son corredores en la competencia, ¿cuál es la probabilidad de que Sofía gane la medalla de oro, Juan la de plata y María la de bronce?
 - c) ¿Cuál es la probabilidad de que la medalla de oro la gane una mujer, la de plata un hombre y la de bronce una mujer?
2. En un club de ajedrez hay 16 miembros, de los cuales 9 son adultos y 7 son jóvenes. El presidente del club necesita seleccionar un comité de cinco miembros para organizar un torneo, sin importar el orden en que sean seleccionados.
 - a) ¿De cuántas formas diferentes se puede formar este comité?
 - b) Si Pedro, Ana y Luis son miembros del club, ¿cuál es la probabilidad de que los tres sean seleccionados para el comité?
 - c) ¿Cuál es la probabilidad de que el comité esté compuesto por tres adultos y dos jóvenes?
3. Un lote de 100 focos contiene 4 defectuosos y 96 en buen estado. Se seleccionan al azar cinco focos. ¿Cuál es la probabilidad de obtener 3 focos buenos y 2 defectuosos?

Autoevaluación y coevaluación 2.1

Nombre: _____ Plantel: _____ Grupo: _____ Turno: _____

Autoevaluación para el aprendizaje

Selecciona en la columna la opción que mejor refleje tu nivel de desempeño en el proceso para el aprendizaje de la progresión de aprendizaje 2. Responde con honestidad a la evaluación de cada uno de los criterios que se enlistan a continuación.

Desempeño	En proceso de logro	Bueno	Sobresaliente
Propicié un clima de comunicación favorable para el aprendizaje con mis compañeros.			
Participé activamente con ideas para la toma razonada de decisiones.			
Contribuí colaborativamente en la retroalimentación de dudas de mis compañeros.			
En el análisis de una situación que lo requería diferencié cuándo utilizar las permutaciones y cuándo las combinaciones (M2-C1).			
Empleé aplicaciones informáticas para el cálculo de permutaciones y combinaciones (M3-C1).			
Utilicé la técnica de conteo apropiada para determinar la cantidad de casos posibles en un evento (M3-C3).			

Coevaluación para el aprendizaje

Solicita a un compañero del equipo que marque en la columna la opción que mejor describa tu desempeño durante el trabajo colectivo, concluida la progresión de aprendizaje 2, y que responda con honestidad la evaluación de cada uno de los criterios que se enlistan a continuación.

Desempeño	En proceso de logro	Bueno	Sobresaliente
Propició un clima de comunicación favorable para el aprendizaje con mis compañeros.			
Participó activamente con ideas para la toma razonada de decisiones.			
Contribuyó colaborativamente en la retroalimentación de dudas de mis compañeros.			
En el análisis de una situación que lo requería diferenció cuándo utilizar las permutaciones y cuándo las combinaciones (M2-C1).			
Empleó aplicaciones informáticas para el cálculo de permutaciones y combinaciones (M3-C1).			
Utilizó la técnica de conteo apropiada para determinar la cantidad de casos posibles en un evento (M3-C3).			

Nombre y firma de quien coevalúa.

La probabilidad condicionada

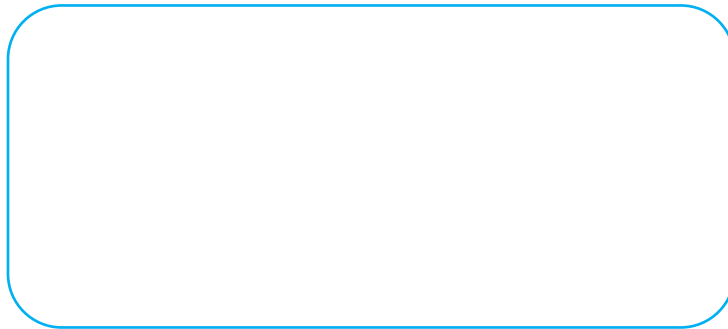
Progresión de aprendizaje 3

Observa cómo la probabilidad de un evento puede actualizarse cuando se obtiene más información al respecto y considera eventos excluyentes e independientes para emplearlos en la determinación de probabilidades condicionales.

Meta de aprendizaje		En proceso de logro	Bueno	Sobresaliente
M2-C2 Desarrolla la percepción y la intuición para generar conjeturas ante situaciones que requieren explicación o interpretación.	A			
	C			
	H			
M4-C2 Argumenta a favor o en contra de afirmaciones acerca de situaciones, fenómenos o problemas propios de la matemática, de las ciencias o de su contexto.	A			
	C			
	H			

Evaluación diagnóstica 3.1

1. En una escuela técnica donde se ofertan las carreras de inglés y computación, hay 20 alumnos inscritos en inglés y 43 inscritos en computación. Dentro del total de alumnos, hay 12 alumnos inscritos en ambos. Construye un diagrama de Venn para ilustrar la situación:



2. Si se elige un alumno al azar:
 - a) ¿Cuál será la probabilidad de que estudie inglés? _____.
 - b) ¿Cuál será la probabilidad de que estudie computación? _____.
 - c) ¿Cuál será la probabilidad de que estudie ambas carreras? _____.
 - d) ¿Cuál será la probabilidad de que estudie inglés, si se sabe que estudia computación? _____.

¿Podrías dar respuesta al inciso d? A continuación, aprenderás cómo.

Probabilidad condicional

Un diagrama de Venn, es una representación gráfica de conjuntos que utiliza círculos superpuestos para mostrar las relaciones entre ellos. Por ejemplo, en la Figura 3.1, donde U representa el conjunto universo y A , B y C son conjuntos cualesquiera. Cualquier conjunto, deberá quedar incluido dentro del área del rectángulo.

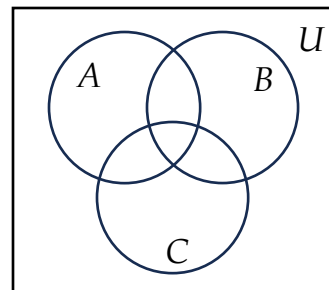


Figura 3.1. Diagrama de Venn.

Fuente: Elaboración propia (Word, 2024).

Ejemplo formativo 3.1

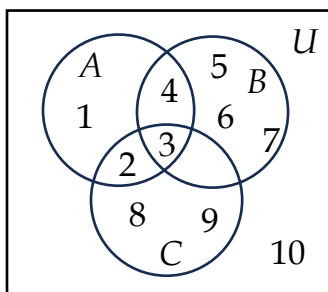


Figura 3.2. Diagrama de Venn.

Fuente: Elaboración propia (Word, 2024).

Por ejemplo, si $U = \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$

$$A = \{1, 2, 3, 4\} \quad B = \{3, 4, 5, 6, 7\} \quad C = \{2, 3, 8, 9\}$$

Construye un diagrama de Ven con los conjuntos dados.

El diagrama de Venn queda como la Figura 3.2.

Ya conoces qué es y sabes cómo se llena un diagrama de Venn; a continuación, se usa en el cálculo de probabilidades.

Ejemplo formativo 3.2

En una escuela primaria en el grupo de segundo grado hay 38 alumnos; se sabe que 12 juegan futbol, 14 practican taekwondo, seis alumnos están inscritos en futbol y a la vez en taekwondo. Si se selecciona un alumno al azar, ¿cuál es la probabilidad de que practique futbol si se sabe que practica taekwondo?

Ante esta nueva situación vamos a dar respuesta a la pregunta y para ello, organicemos la información en un diagrama de Venn, de acuerdo con la Figura 3.3.

Sean:

A : el suceso "Practica futbol"

B : el suceso "Practica taekwondo"

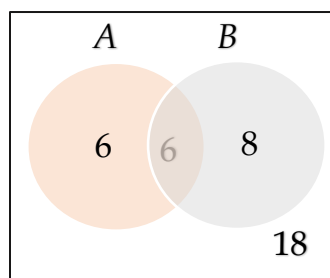


Figura 3.3. Diagrama de Venn.

Fuente: Elaboración propia (Word, 2024).

$$P(A) = \frac{12}{38} = \frac{6}{19}$$

$$P(B) = \frac{14}{38} = \frac{7}{19}$$

$$P(A \cap B) = \frac{6}{38} = \frac{3}{19}$$

El problema pide calcular la probabilidad de que un alumno practique futbol, si ya se sabe que practica taekwondo, por lo que nos concentramos en los que practican taekwondo (14) y de ellos los que practican futbol (6); lo representamos mediante la Figura 3.4.

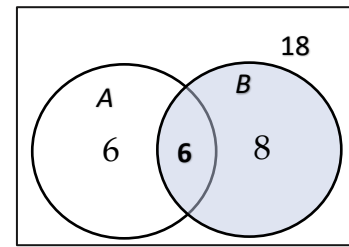


Figura 3.4. Diagrama de Venn.
Fuente: Elaboración propia (Word, 2024).

$P(\text{practica futbol} / \text{practica taekwondo})$ o bien $P(A/B)$ y considera lo siguiente:

Selecciona un alumno que se sabe practica taekwondo (información adicional), de forma que solo se consideran los alumnos que lo practican, excluyendo a aquellos que no.

$$\text{Probabilidad de que practique taekwondo: } P(B) = \frac{7}{19}$$

$$\text{Probabilidad de que practique futbol y taekwondo: } P(A \cap B) = \frac{3}{19}$$

$$P(A/B) = \frac{P(\text{practica futbol y taekwondo})}{P(\text{taekwondo})} = \frac{\frac{3}{19}}{\frac{7}{19}} = \frac{3}{7}$$

Por lo tanto, si se selecciona un alumno al azar, la probabilidad de que practique futbol si se sabe que practica taekwondo es $3/7$.

La probabilidad condicional o probabilidad condicionada, es la posibilidad de que ocurra un evento, al que denominamos A , como consecuencia de que ha tenido lugar otro evento, al que denominamos B .

Para designar la probabilidad de un evento B , condicionada a otro evento A , escribimos:

$P(B/A) \rightarrow$ La probabilidad de B , dado que ya ocurrió A .

$$P(B/A) = \frac{P(A \cap B)}{P(A)}, \quad \text{si } P(A) > 0$$

Es decir, la probabilidad de que suceda B , dado que ha acontecido A , es igual a la probabilidad de que ocurra A y B , al mismo tiempo, entre la probabilidad de A .

De forma análoga, la probabilidad de que suceda A , dado que ha acontecido B , es igual a la probabilidad de que ocurra A y B , al mismo tiempo, entre la probabilidad de B , se representa como sigue.

$P(A/B) \rightarrow$ La probabilidad de A dado que ya ocurrió B.

$$P(A/B) = \frac{P(A \cap B)}{P(B)}, \quad \text{si } P(B) > 0$$



QR 3.1. Probabilidad condicional: introducción.
Fuente: Parzibyte 2024.

¿Cómo identificas cuándo utilizar la fórmula de probabilidad condicional?

Ejemplo formativo 3.3

En primer grado los alumnos deben realizar por lo menos una actividad cocurricular fuera del horario de clases, y se sabe que de 300 alumnos: 40 están inscritos en teatro, 60 están inscritos en porristas, 20 están inscritos en porristas y en teatro a la vez; el resto realiza otra actividad cocurricular. ¿Cuál es la probabilidad de que un alumno esté inscrito en porristas, dado que está inscrito en teatro?

Sean:

A: el suceso “está inscrito en teatro”. B: el suceso “está inscrito en porristas”.

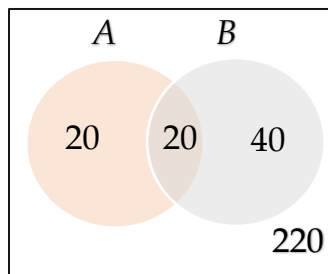


Figura 3.5. Diagrama de Venn.
Fuente: Elaboración propia (Word, 2024).

$$P(B/A) = \frac{P(B \cap A)}{P(A)}$$

$$P(A) = \frac{40}{300} = \frac{2}{15}$$

$$P(B \cap A) = \frac{20}{300} = \frac{1}{15}$$

$$P(B/A) = \frac{\frac{1}{15}}{\frac{2}{15}} = \frac{15}{30} = \frac{1}{2}$$

Ahora, ¿cuál es la probabilidad de que un alumno esté inscrito en teatro, dado que está inscrito en porristas?

$$P(A/B) = \frac{P(B \cap A)}{P(B)} = \frac{\frac{1}{15}}{\frac{1}{5}} = \frac{1}{3}$$

Ejemplo formativo 3.4

En un estudio realizado por alumnos de la Facultad de Medicina, de la Universidad Autónoma de Sinaloa a personas de la tercera edad, se encontró que el 33% padece diabetes, el 41% padece hipertensión y 17% padece ambas enfermedades, mientras que el resto de la población no padece ninguna de las dos enfermedades. ¿Cuál es la probabilidad de que una persona padezca diabetes, dado que es hipertenso?

Sean D : padecer diabetes H : ser hipertenso

$$P(D \cap H) = 0.17 \text{ y } P(H) = 0.41$$

$$P(D/H) = \frac{P(D \cap H)}{P(H)} = \frac{0.17}{0.41} = 0.4146$$

La probabilidad de que una persona de la tercera edad padezca diabetes, dado que es hipertenso, es de 41.46 %.

Probabilidad de eventos mutuamente excluyentes

Dos o más eventos son mutuamente excluyentes, si al ocurrir uno de ellos, excluye la posibilidad de que ocurra el otro.

Ejemplo formativo 3.6

Si se lanza una moneda.

Sean A : que al caer muestre águila.

B : que al caer muestre sello.



Figura 3.6. Lanzamiento de una moneda.

Fuente: Fotografía. Rolando Alberto Forneiro Rodríguez, 2024.

Si sucede el evento A , no puede suceder el evento B de acuerdo con la Figura 3.7.

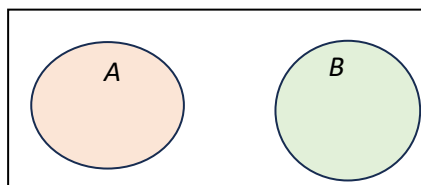


Figura 3.7. Diagrama de Venn.

Fuente: Elaboración propia (Word, 2024).

Así, al lanzar una moneda, solo puede ocurrir uno de los dos eventos a la vez, o cae águila o cae sello.

Probabilidad de eventos independientes

Dos eventos A y B son independientes si y solo si, la probabilidad del evento B no está influida por el suceso del evento A o viceversa.

Ejemplo formativo 3.7

¿Cuál es la probabilidad de que, al lanzar dos dados, en los dos se obtenga número impar?

Evento A : al lanzar el dado 1 caiga número impar (1, 3, 5), $P(A) = \frac{3}{6}$.

Evento B : al lanzar el dado 2 caiga número impar (1, 3, 5), $P(B) = \frac{3}{6}$.

$$P(A \cap B) = P(A) \cdot P(B)$$

$$P(A \cap B) = \frac{3}{6} \cdot \frac{3}{6} = \frac{9}{36} = \frac{1}{4}$$

	1	2	3	4	5	6
1	1,1	1,2	1,3	1,4	1,5	1,6
2	2,1	2,2	2,3	2,4	2,5	2,6
3	3,1	3,2	3,3	3,4	3,5	3,6
4	4,1	4,2	4,3	4,4	4,5	4,6
5	5,1	5,2	5,3	5,4	5,5	5,6
6	6,1	6,2	6,3	6,4	6,5	6,6



Figura 3.8. Lanzamiento de dos dados.
Fuente: www.freepik.com, 2024.

Evaluación formativa 3.1

1. En un estudio de 100 bebedores de café, se determinó que 70 utilizan azúcar, 60 usan crema y 50 de ellos, prefieren azúcar y crema. ¿Cuál es la probabilidad de que los bebedores de café no utilizan ni azúcar, ni crema en el café?
2. Una caja contiene 7 contactos defectuosos y 3 aceptables; otra caja contiene 4 defectuosos y 6 aceptables. Se traslada un contacto de la primera caja a la segunda; a continuación, se extrae un contacto de la segunda caja. ¿Cuál es la probabilidad de que este último sea aceptable?
3. Dos personas por separado, eligen al azar un número del 0 al 9, ¿cuál es la probabilidad de que las dos personas no piensen en el mismo número?
4. En una ciudad, el 30 % de la población tiene el cabello castaño, el 35 % tiene ojos castaños y el 20 % tiene cabello y ojos castaños. Se elige una persona al azar:
 - a) Si tiene el cabello castaño, ¿cuál es la probabilidad de que tenga también ojos castaños?
 - b) Si tiene ojos castaños, ¿cuál es la probabilidad de que no tenga cabello castaño?
5. En una clase de inglés hay 100 alumnos, de los cuales 55 son hombres, 40 usan reloj, y 20 son hombres y usan reloj. Se selecciona al azar un alumno de dicho curso:
 - a) Si sabemos que el alumno seleccionado es hombre, ¿cuál es la probabilidad de que use reloj?

- b) Si sabemos que el alumno seleccionado no usa reloj, ¿qué probabilidad hay de que sea hombre?
6. Un inversionista está considerando dos empresas (A y B) para invertir. La probabilidad de que la empresa A sea rentable este año es del 70 %, mientras que para la empresa B es del 50 %. La probabilidad de que ambas sean rentables es del 35 %.
- a) Determina si las rentabilidades de las empresas A y B son eventos independientes.
- b) Luego, el inversionista recibe un informe que indica que el sector económico de la empresa A está creciendo más de lo esperado este año. ¿Cómo debería actualizar el inversionista, la probabilidad de rentabilidad de la empresa A basándose en esta nueva información?
7. Una nueva enfermedad ha comenzado a propagarse en una comunidad cerrada. Inicialmente, la probabilidad de que una persona dentro de esta comunidad se infecte es del 30 %. Una vacuna se introduce y se administra al 40 % de la población. Quienes están vacunados tienen solo un 10 % de probabilidad de infectarse. Un residente de la comunidad da positivo a la enfermedad.
- a) Calcula la probabilidad de que esta persona haya sido vacunada.
- b) ¿Cómo cambia esta probabilidad, si te informan que solo el 20 % de las pruebas positivas corresponden a personas vacunadas?
8. Escribe un ejemplo de eventos que sean mutuamente excluyentes.
- _____
- _____
9. Calcula la probabilidad de que al lanzar dos monedas el resultado sea sol en ambas. _____

Autoevaluación y coevaluación 3.1

Nombre: _____ Plantel: _____ Grupo: _____ Turno: _____

Autoevaluación para el aprendizaje

Selecciona en la columna la opción que mejor refleje tu nivel de desempeño en el proceso para el aprendizaje de la progresión de aprendizaje 3. Responde con honestidad, a la evaluación de cada uno de los criterios que se enlistan a continuación.

Desempeño	En proceso de logro	Bueno	Sobresaliente
Propicié un clima de comunicación favorable para el aprendizaje con mis compañeros.			
Participé activamente con ideas para la toma razonada de decisiones.			

Contribuí colaborativamente en la retroalimentación de dudas de mis compañeros.			
Generé conjeturas ante situaciones sobre el cálculo de probabilidades que requieren explicación o interpretación. (M2-C2).			
Identifiqué cuándo utilizar la probabilidad condicional en la resolución de problemas contextualizados (M4-C2).			

Coevaluación para el aprendizaje

Solicita a un compañero del equipo que marque en la columna, la opción que mejor describa tu desempeño durante el trabajo colectivo, concluida la progresión de aprendizaje 3, y que responda con honestidad la evaluación de cada uno de los criterios que se enlistan a continuación.

Desempeño	En proceso de logro	Bueno	Sobresaliente
Propició un clima de comunicación favorable para el aprendizaje con mis compañeros.			
Participó activamente con ideas para la toma razonada de decisiones.			
Contribuyó colaborativamente en la retroalimentación de dudas de mis compañeros.			
Generó conjeturas ante situaciones sobre el cálculo de probabilidades que requieren explicación o interpretación. (M2-C2).			
Identificó cuándo utilizar la probabilidad condicional en la resolución de problemas contextualizados (M4-C2).			

Nombre y firma de quien coevalúa.

Recolección de datos estadísticos

Progresión de aprendizaje 4

Selecciona una problemática o situación de interés, con la finalidad de recolectar información y datos de fuentes confiables e identifica las variables relevantes para su estudio.

Metas de aprendizaje	En proceso de logro	Bueno	Sobresaliente
M1-C1 Ejecuta cálculos y algoritmos para resolver problemas matemáticos, de las ciencias y de su entorno.	A		
	C		
	H		
M1-C2 Observa y obtiene información de una situación o fenómeno para establecer estrategias o formas de visualización que ayuden a entenderlo.	A		
	C		
	H		
M2-C4 Socializa con sus pares sus conjeturas, descubrimientos o procesos en la solución de un problema tanto teórico como de su entorno.	A		
	C		
	H		

Evaluación diagnóstica 4.1

Investigando la relación entre vivir solo y la depresión.

Un grupo de estudiantes de primer año de bachillerato visitó a sus abuelos para indagar sobre la depresión. Los resultados revelaron que aquellos que viven solos son más propensos a deprimirse, que aquellos que viven con otras personas.

1. ¿Qué variables sondearon los estudiantes, variables cuantitativas o cualitativas?
2. ¿Cómo las medirías?
3. ¿Cómo comprobarías si hay una relación entre vivir solo y la depresión?

Una fuente confiable de información es el INEGI (Instituto Nacional de Estadística, Geografía e Información). Este es un organismo que recopila información mediante la aplicación de censos y encuestas y usa la estadística para producir indicadores que proporcionen un panorama completo de diversos temas de México como: demografía y sociedad; economía y sectores productivos; geografía y medio ambiente; así como gobierno, seguridad y justicia.

Por ejemplo, en el tema de demografía y sociedad, la Encuesta Nacional de Ingresos y Gastos en los hogares (ENIGH) 2022, reveló que el gasto corriente monetario

promedio trimestral de las y los sinaloenses es de aproximadamente \$43, 000 pesos de los cuales, el rubro de alimentos, bebidas y tabaco representó el mayor gasto rondando los \$14,000 pesos, en contraste con el rubro de cuidados de la salud con aproximadamente \$1,500 pesos. Por lo que las y los sinaloenses gastamos más en alimentos, bebidas y tabaco que en el cuidado de la salud.

Rubros de gasto	Gasto promedio en pesos
Alimentos, bebidas y tabaco.	\$14,000
Transporte y comunicaciones.	\$10,000
Vivienda y servicios.	\$4,000
Cuidados personales.	\$3,500
Educación y esparcimiento.	\$3,500
Limpieza y cuidados de la casa.	\$3,000
Vestido y calzado.	\$2,000
Transferencias de gastos.	\$1,500
Cuidados de la salud.	\$1,500
Total (gasto corriente monetario trimestral)	\$43,000

Tabla 4.1. Gasto corriente monetario promedio trimestral del año 2022 por hogar en Sinaloa, según grandes rubros del gasto en pesos.
Fuente: Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) 2022. Nueva Serie. Disponible en: <https://www.inegi.org>.

[mx/programas/enigh/nc/2022/](https://www.inegi.org.mx/programas/enigh/nc/2022/)En el rubro de alimentos, bebidas y tabaco de las 20 principales variables de gasto del 2022 la Tabla 4.2, nos muestra que las y los sinaloenses gastamos más en cuidados personales (\$2,663.00) y en electricidad y combustibles (\$2,640.00).

Rubros específicos del gasto	Gasto promedio en pesos	Rubros específicos del gasto	Gasto promedio en pesos
Cuidados personales	\$2,663	Vestido	\$1,374
Electricidad y combustibles	\$2,640	Transporte público	\$1,344
Educación	\$2,450	Leche y derivados	\$1,187
Alimentos fuera del hogar	\$2,411	Verduras	\$1,114
Carnes	\$2,268	Huevo	\$749
Comunicaciones	\$2,191	Alquileres brutos	\$700
Cereales	\$2,102	Esparcimiento	\$688
Cuidados de la casa	\$1,821	Otros gastos diversos	\$676
Adquisición de vehículos	\$1,639	Calzado y su reparación	\$630
Bebidas	\$1,475	Enseres domésticos	\$596

Tabla 4.2. Gasto corriente monetario promedio trimestral en los veinte principales rubros de gasto del año 2022 por hogar en Sinaloa, según grandes rubros del gasto en pesos.

Fuente: Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) 2022. Nueva Serie. Disponible en: <https://www.inegi.org.mx/programas/enigh/nc/2022/>

Hay varias técnicas para recopilar los datos estadísticos, su elección depende del tipo de investigación. Por ejemplo, para la recopilación de los datos de la Tabla 4.1 y la Tabla 4.2 sobre el gasto corriente promedio trimestral de las y los sinaloenses, la técnica

implementada fue la encuesta y el instrumento aplicado fue el cuestionario sobre gastos del hogar.

Ahora, hablando de **probabilidad**, esta se refiere a la medida de la certeza o posibilidad de que ocurra un evento y se utiliza para predecir resultados en situaciones inciertas o aleatorias. Un ejemplo clásico es el experimento de lanzar un dado, en el que la probabilidad de obtener un número específico es $1/6$, ya que hay 6 caras equiprobables

Por su parte, la **estadística** se enfoca en recopilar, organizar, analizar e interpretar datos para obtener conclusiones significativas. Utiliza técnicas estadísticas para sacar conclusiones sobre una población, basándose en muestras para recopilar datos mediante encuestas aplicadas en estudios científicos, análisis de mercados y muchos otros campos. La estadística contribuye a entender la variabilidad de los datos, a tomar decisiones informadas y hacer inferencias sobre la población en general, a partir de los datos obtenidos de la muestra.

Una **encuesta** constituye una técnica de investigación que el investigador implementa para recopilar información a través de un cuestionario. Los resultados se pueden presentar en diversos formatos como trípticos, gráficos, tablas o informes escritos. El cuestionario se aplica para recolectar datos de diferentes tipos de variables estadísticas: **cualitativas**, que pueden ser nominales u ordinales y **cuantitativas**, que se dividen en discretas y continuas.

Las **variables cualitativas** también son conocidas como variables categóricas, que representan características o cualidades no numéricas (es decir, no pueden medirse en términos numéricos). Se representan mediante etiquetas expresadas en palabras o símbolos. Estas variables se dividen en dos tipos: las nominales y las ordinales.

Las **variables cualitativas nominales** representan distintas categorías sin un orden inherente. Por ejemplo, el color de los ojos (negro, café, ámbar, avellana, verde, azul), el género (hombre, mujer, no binario), el tipo de sangre (A+, A-, B+, B-, AB+, AB-, O+, O-). No existe una jerarquía entre las categorías.

Las **variables cualitativas ordinales** son atributos o cualidades, que representan categorías que tienen un orden o jerarquía intrínseca. Por ejemplo, el nivel educativo (primaria, secundaria, bachillerato y universidad), la escala de satisfacción de una serie de televisión (muy mala, mala, regular, buena, muy buena y excelente) o el nivel de acuerdo o desacuerdo que tienes con el reglamento escolar (totalmente en desacuerdo, en desacuerdo, neutral, de acuerdo y totalmente de acuerdo).

Ejemplo formativo 4.1

Identifica si las siguientes variables cualitativas son nominales u ordinales según corresponda.

a) Raza de gato	<u>Nominal</u>	b) Color de ropa preferido	<u>Nominal</u>
c) Estado civil	<u>Nominal</u>	d) Número de celular	<u>Nominal</u>
e) Preferencia musical	<u>Nominal</u>	f) Nivel socioeconómico	<u>Ordinal</u>
g) Código postal	<u>Nominal</u>	h) Grado del dolor de cabeza	<u>Ordinal</u>
i) Tipo de transporte	<u>Nominal</u>	j) Nivel de atención en un restaurante	<u>Ordinal</u>
k) Marca de automóvil	<u>Nominal</u>	l) Clasificación de tallas	<u>Ordinal</u>

Las **variables cuantitativas** se expresan mediante cantidades numéricas con las que puedes realizar operaciones matemáticas tales como la suma, la resta, la multiplicación y la división para calcular la media, la mediana, la moda y otras medidas estadísticas. Estas variables se dividen en dos tipos: las variables cuantitativas discretas y las variables cuantitativas continuas.

Las **variables cuantitativas discretas** asumen un número contable de valores. Por ejemplo: el número de hijos, la cantidad de automóviles en un estacionamiento o el número de estudiantes en una clase. Estas variables toman valores enteros y se cuentan en unidades individuales.

Las **variables cuantitativas continuas** pueden tomar cualquier valor en un intervalo. Ejemplos comunes son la altura, el peso, la temperatura o la cantidad de dinero. Estas variables pueden tener infinitos valores posibles dentro de un intervalo y se representan generalmente mediante los números reales (\mathbb{R}).

Ejemplo formativo 4.2

Identifica si las siguientes variables cuantitativas son discretas (D) o continuas (C) según corresponda.

a) Pares de zapato por persona	<u>D</u>	b) Velocidad de un coche	<u>C</u>
c) Ancho de una casa	<u>C</u>	d) Distancia recorrida al día	<u>C</u>
e) Costo de un desayuno	<u>C</u>	f) Habitantes de una casa	<u>D</u>
g) Cantidad de mascotas en el hogar	<u>D</u>	h) Sueldo de un profesor	<u>C</u>
i) Duración de tu canción preferida	<u>C</u>	j) Horas de sueño	<u>C</u>
k) Cantidad de sal que consumes	<u>C</u>	l) Automóviles vendidos	<u>D</u>

Al emprender una investigación, el definir la variable y su tipo, es un factor esencial para el manejo de datos. En estadística, tener claro qué estás estudiando es vital y la variable es el elemento que te proporciona esa información clave. Por lo que debes

desarrollar habilidades para reconocer en una investigación, si una variable es cuantitativa (discreta o continua) o cualitativa (ordinal o nominal).

Ejercicio formativo 4.3

En las siguientes situaciones identifica la variable y determina su tipo: si es cuantitativa discreta, cuantitativa continua, cualitativa ordinal o cualitativa nominal.

1. Se realizará un estudio sobre el uso de energías renovables en hogares. Con esta finalidad, se llevará a cabo un censo de paneles solares instalados en viviendas y con ello evaluar la adopción y eficacia de estas fuentes alternativas.
La variable es censo de paneles solares. Es una variable de tipo cuantitativa discreta.
2. Un estudio detalla el incremento de casos de estrés en la población laboral. Los datos recabados muestran un aumento del 35% en comparación con el año anterior.
La variable es incremento de estrés laboral. Es una variable de tipo cuantitativa discreta.
3. Un almacén de productos electrónicos está llevando a cabo un inventario, clasificando los dispositivos por marcas.
La variable es marca del dispositivo. Es una variable de tipo cualitativa nominal.
4. El campo de golf local está elaborando una tabla con los puntajes de un torneo de 18 hoyos, indicando las puntuaciones de cada golfista en cada hoyo del recorrido.
La variable es puntaje en el torneo del golf. Es una variable de tipo cuantitativa discreta.
5. Un fisioterapeuta está monitoreando la movilidad de un paciente postoperatorio para otorgarle el alta cuando recupere un cierto nivel de movimiento.
La variable es nivel de movilidad de un paciente postoperatorio. Es una variable de tipo cualitativa ordinal.

Al estudiar una problemática o situación de interés, puedes hacerlo mediante un censo, como el INEGI al llevar a cabo el censo poblacional, debido a que necesita información detallada y exhaustiva de toda la población; sin embargo, debido al costo que esto implica, puedes recurrir a la selección de una muestra de la población. La **población**, se refiere al conjunto completo de elementos que se busca investigar o analizar. Por otro lado, la **muestra** es una parte cuidadosamente seleccionada de la población, que representa sus características principales. Como lo verás más adelante, la muestra la puedes seleccionar por medio de un muestreo probabilístico o no probabilístico.

Evaluación formativa 4.1

1. Aplica la siguiente encuesta a un mínimo de 30 personas que lleven el gasto de su hogar. Puedes crear un formulario de Google y cada uno lo aplica a quien lleve el gasto en su casa (asegurar un mínimo de 30 encuestados).

Encuesta de gastos del hogar

Hola, gracias por participar en nuestra encuesta sobre los artículos de limpieza, para cuidado personal y servicios educativos. Tus respuestas son valiosas y nos ayudarán a entender mejor las necesidades de las personas en relación con estos productos.

- I. Edad: _____
- II. ¿Cuál es tu género?
 - Masculino
 - Femenino
 - No binario
 - Prefiero no decirlo
- III. ¿Cuántas personas viven en tu hogar? _____

Artículos de limpieza

- IV. ¿Cuáles de los siguientes artículos de limpieza utilizas regularmente en tu hogar?
 - Detergente para ropa
 - Limpiador multiusos
 - Desinfectante
 - Limpiavidrios
 - Aromatizantes
 - Escobas
 - Trapeadores
 - Aspiradora
 - Productos para el cuidado de pisos (mopas, ceras, etc.)
 - Otro (especificar) _____
- V. ¿Cómo decides qué productos de limpieza comprar?
 - Precio
 - Recomendación de amigos o familiares
 - Publicidad
 - Ingredientes o componentes específicos
 - Experiencia previa con el producto
 - Otro (especificar) _____
- VI. ¿Qué aspecto consideras más importante al elegir productos de limpieza?
 - Efectividad
 - Seguridad para el medio ambiente

- Seguridad para la salud
- Precio
- Otras consideraciones (especificar)_____

VII. ¿Cuánto gastas al mes en artículos para limpieza del hogar? _____

Artículos para el cuidado personal

VIII. ¿Qué productos de cuidado personal usan comúnmente los hombres en tu hogar?
(Puedes seleccionar más de una opción)

- Afeitadoras
- Gel de afeitar
- Champú
- Gel o espuma para el cabello
- Desodorante
- Cuidado facial (cremas, lociones, etc.)
- Otros (especificar)_____

IX. ¿Qué productos de cuidado personal usan comúnmente las mujeres en tu hogar?
(Puedes seleccionar más de una opción)

- Champú y acondicionador
- Jabón o gel de ducha
- Cremas hidratantes
- Maquillaje
- Productos de higiene íntima
- Desodorante
- Afeitadoras
- Gel de afeitar
- Otros (especificar)_____

X. ¿Qué productos de cuidado personal usan comúnmente los niños en tu hogar?
(Puedes seleccionar más de una opción)

- Champú para niños
- Jabón infantil
- Crema para pañales (si hay bebés)
- Cepillos de dientes para niños
- Pasta dental para niños
- Otros (especificar)_____

XI. ¿Cuánto gastas sal mes en artículos para el cuidado personal de los que viven en su hogar? _____

Servicios educativos

XII. ¿Cuál es el máximo nivel educativo actual de la mamá?

- Educación preescolar
- Primaria
- Secundaria
- Bachillerato
- Licenciatura
- Maestría
- Doctorado

XIII. ¿Cuál es el máximo nivel educativo actual del papá?

- Educación preescolar
- Primaria
- Secundaria
- Bachillerato
- Licenciatura
- Maestría
- Doctorado

XIV. ¿Qué recursos educativos utilizan en el hogar para apoyar la educación de los miembros en edad escolar?

(Puedes seleccionar más de una opción)

- Libros y materiales didácticos
- Acceso a internet y tecnología
- Aplicaciones educativas
- Clases particulares
- Otros (especificar)_____

XV. ¿Cuáles crees que son los principales desafíos en el acceso a servicios educativos para los miembros de tu hogar?

(Puedes seleccionar más de una opción)

- Falta de recursos tecnológicos
- Dificultades de acceso a internet
- Necesidades educativas especiales
- Otros (especificar)_____

XVI. ¿Cómo calificas la calidad de los servicios educativos que reciben los miembros del hogar en edad escolar?

- Pésimo
- Malo
- Regular
- Bueno

- Excelente

XVII. ¿Cuánto gastas al mes en artículos escolares de los miembros en edad escolar?

XVIII. ¿Hay algún comentario adicional que te gustaría compartir sobre los servicios educativos en tu hogar?

2. Identifica la variable en cada pregunta y clasifícala según su tipo (si es cuantitativa discreta, cuantitativa continua, cualitativa ordinal o cualitativa nominal).

Número de pregunta	Variable	Tipo de variable
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		
16		
17		
18		

3. Resume los datos de cada pregunta en forma de gráfica o tabla según sea más conveniente.

Autoevaluación y coevaluación 10.1

Nombre: _____ Plantel: _____ Grupo: _____ Turno: _____

Autoevaluación para el aprendizaje

Selecciona en la columna la opción que mejor refleje tu nivel de desempeño en el proceso para el aprendizaje de la progresión de aprendizaje 4. Responde con honestidad a la evaluación de cada uno de los criterios que se enlistan a continuación.

Desempeño	En proceso de logro	Bueno	Sobresaliente
Propicié un clima de comunicación favorable para el aprendizaje con mis compañeros.			
Participé activamente con ideas para la toma razonada de decisiones.			
Contribuí colaborativamente en la retroalimentación de dudas de mis compañeros.			
Calculé la media, mediana o moda según el tipo de variable (M1-C1).			
Diferencíé entre los tipos de variables cualitativas y cuantitativas (M1-C2).			
Compartí mis propuestas sobre el diseño de cuestionarios para recolectar información de una situación de interés (M2-C4).			

Coevaluación para el aprendizaje

Solicita a un compañero del equipo que marque en la columna la opción que mejor describa tu desempeño durante el trabajo colectivo, concluida la progresión de aprendizaje 4, y que responda con honestidad la evaluación de cada uno de los criterios que se enlistan a continuación.

Desempeño	En proceso de logro	Bueno	Sobresaliente
Propició un clima de comunicación favorable para el aprendizaje con mis compañeros.			
Participó activamente con ideas para la toma razonada de decisiones.			
Contribuyó colaborativamente en la retroalimentación de dudas de mis compañeros.			
Calculó la media, mediana o moda según el tipo de variable (M1-C1).			
Diferenció entre los tipos de variables cualitativas y cuantitativas (M1-C2).			
Compartió sus propuestas sobre el diseño de cuestionarios para recolectar información de una situación de interés (M2-C4).			

Nombre y firma de quien coevalúa

Representación de la información

Progresión de aprendizaje 5

Analiza datos categóricos y cuantitativos de alguna problemática o situación de interés para el estudiantado, a través de algunas de sus representaciones gráficas más sencillas como las gráficas de barras (variables cualitativas) o gráficos de puntos e histogramas (variables cuantitativas).

Metas de aprendizaje	En proceso de logro	Bueno	Sobresaliente
M1-C1 Ejecuta cálculos y algoritmos para resolver problemas matemáticos, de las ciencias y de su entorno.	A		
	C		
	H		
M2-C1 Analiza los resultados obtenidos al aplicar procedimientos algorítmicos propios del Pensamiento Matemático, en la resolución de problemáticas teóricas y de su contexto.	A		
	C		
	H		
M2-C2 Desarrolla la percepción y la intuición para generar conjeturas, ante situaciones que requieren explicación o interpretación.	A		
	C		
	H		

Evaluación Diagnóstica 5.1

En los deportes generalmente se recaban, organizan, analizan e interpretan datos relativos al rendimiento de los jugadores o equipo con el objetivo de evaluar el desempeño durante la temporada. A través de la cuantificación y el ordenamiento de los datos, se intenta explicar los resultados, por lo que resulta una herramienta de suma utilidad para tomar decisiones basadas en la evidencia y proporciona instrumentos para la mejora en el entrenamiento y en la competencia.

En la Liga Mexicana del Pacífico (LMP) se obtuvieron los resultados que aparecen en la Tabla 5.1, durante la temporada 2022-2023.

Equipo	Juegos jugados	Juegos ganados	Juegos perdidos
Algodoneros de Guasave	68	36	32
Tomateros de Culiacán	68	26	42
Venados de Mazatlán	68	33	35
Mayos de Navojoa	68	31	37
Yaquis de Obregón	68	38	30
Naranjeros de Hermosillo	68	43	25
Cañeros de Los Mochis	68	41	27
Charros de Jalisco	68	29	39
Sultanes de Monterrey	68	32	36
Águilas de Mexicali	68	31	37

Tabla 5.1. Frecuencia de juegos de beisbol jugados en la LMP.

Fuente:

https://es.wikipedia.org/wiki/Liga_Mexicana_del_Pac%C3%ADfico_2022-23

1. Calcula el promedio (media) de los juegos ganados de los equipos.

2. ¿Qué valores se repiten en los juegos ganados perdidos? _____
3. ¿Qué tipo de análisis representan los datos anteriores? ¿Cualitativo o cuantitativo? _____. Argumenta tu respuesta: _____

4. Observa la Figura 5.1 y analiza la información que se solicita:

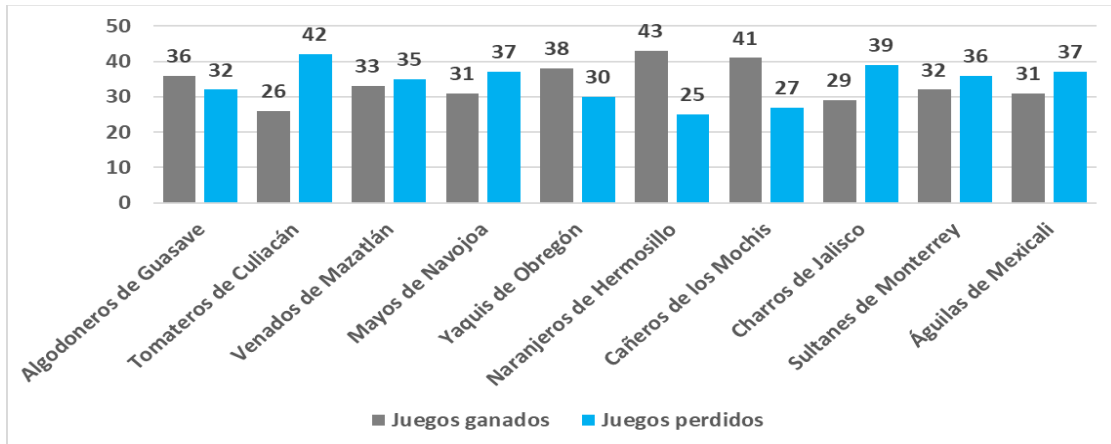


Figura 5.1. Resultados de juegos ganados y juegos perdidos de la LMP 2022-2023.

Fuente: Elaboración propia, con base en: https://es.wikipedia.org/wiki/Liga_Mexicana_del_Pacífico_2022-23 (Excel, 2024).

- a) ¿Qué información te muestra la gráfica presentada, y qué tipo de gráfico es?

- b) ¿Qué deduces con base en los datos presentados de la gráfica anterior?

5. A la vista de los datos presentados de la gráfica de la Figura 5.2, ¿qué deduces?

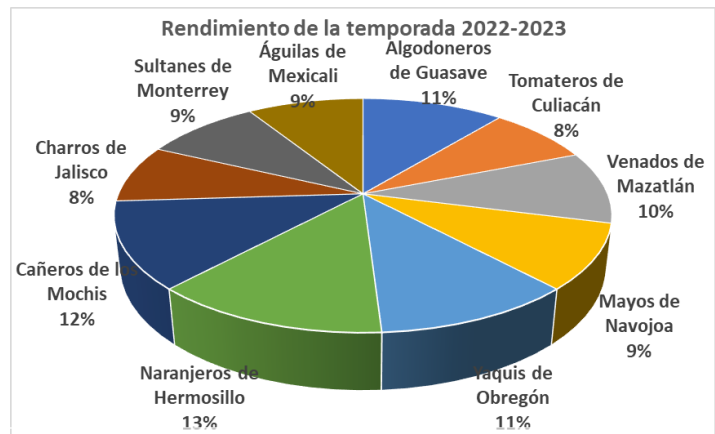


Figura 5.2. Rendimiento de la temporada 2022-2023 en la LMP.

Fuente: Elaboración propia, con base en:

https://es.wikipedia.org/wiki/Liga_Mexicana_del_Pacífico_2022-23 (Excel, 2024).

6. ¿Qué tipo de gráfico es?

De acuerdo con el análisis realizado en la evaluación diagnóstica, ¿en qué piensas que te pueden ayudar las gráficas en tu vida diaria?

Las personas, a lo largo de su vida, toman decisiones importantes relacionadas con las actividades que desempeñan día a día, por ello el pensamiento estadístico ha cobrado cada vez más importancia y el uso de gráficas para analizar y visualizar información es cada vez más frecuente, ya que te facilitan la comunicación y la visualización de los datos y son utilizadas como una herramienta en el proceso de análisis.

Existen diversos tipos de gráficas estadísticas que se pueden crear como se muestra en la Figura 5.3, dependiendo de cómo tengas organizada tu información y tipo de variable de estudio; ello te permitirá analizar, observar y verificar algunas tendencias que presenten los datos.

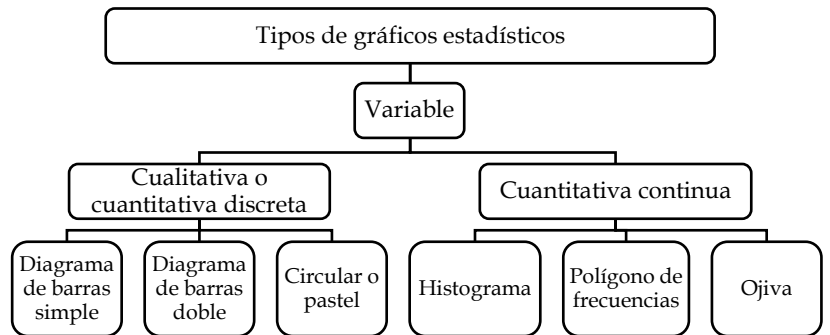


Figura 5.3. Tipos de gráficos.
Fuente: Elaboración propia (Word, 2024).

Gráfica de barras simple, barra doble y circulares

Las gráficas de barras, se usan para organizar y presentar datos diversos y pueden ser horizontales o verticales. Se utilizan principalmente, para representar variables cualitativas como se muestra en las Figuras 5.4 y 5.5.



Figura 5.4. Gráfica de barra simple.
Fuente: Elaboración propia (Excel, 2024).

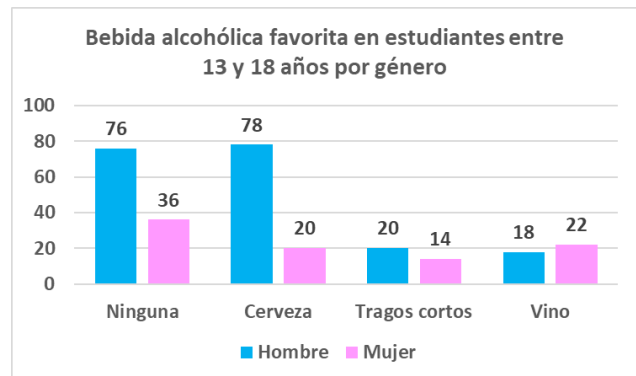


Figura 5.5. Gráfica de barra doble.
Fuente: Elaboración propia (Excel, 2024).

Por ejemplo, la Tabla 5.2 se obtuvo de una investigación para conocer las 10 cuentas de Instagram con más seguidores en el 2023, y en la Figura 5.6, se muestra mediante una gráfica de barras.

Cuenta	Seguidores (millones) 2023 ²	Seguidores (millones) 2022 ¹
Instagram	658	569
Cristiano Ronaldo	607	500
Lionel Messi	488	376
Selena Gomez	430	357
Kylie Jenner	400	372
Dwayne Johnson	391	348
Ariana Grande	380	340
Kim Kardashian	364	333
Beyoncé	317	257
Khloé Kardashian	312	243

Tabla 5.2. Cuentas de Instagram con más seguidores.
Fuente: Elaboración propia, con base en: <https://www.ayudacuentas.com/instagram/cuentas-con-mas-seguidores/> y <https://www.merca20.com/los-mas-seguidos-instagram-2022/> (Word, 2023).

Las gráficas de barra doble, son similares a las gráficas de barras simples, pero en estas, se grafican los datos de dos variables de estudio por cada valor de la categoría, utilizando barras adyacentes para comparar las cantidades entre las dos categorías. Por ejemplo, la Tabla 5.2 muestra las cuentas de Instagram con más seguidores tanto en el 2022 como en el 2023, a partir de la cual puedes realizar una gráfica de barra doble para representar los datos como se muestra en la Figura 5.7.

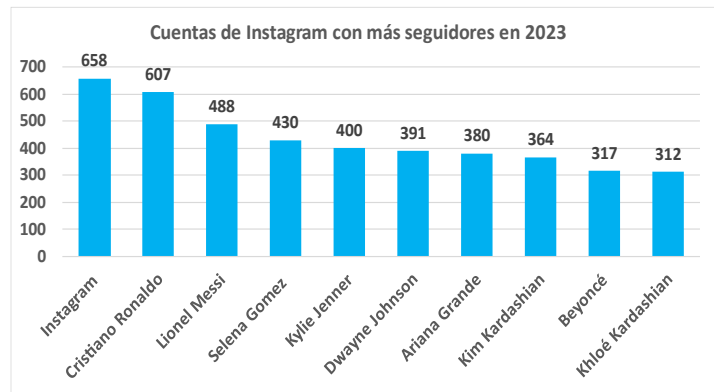


Figura 5.6. Cuentas de Instagram con más seguidores en 2023.
Fuente: Elaboración propia, con base en AyudaCuentas.com (Excel, 2024).

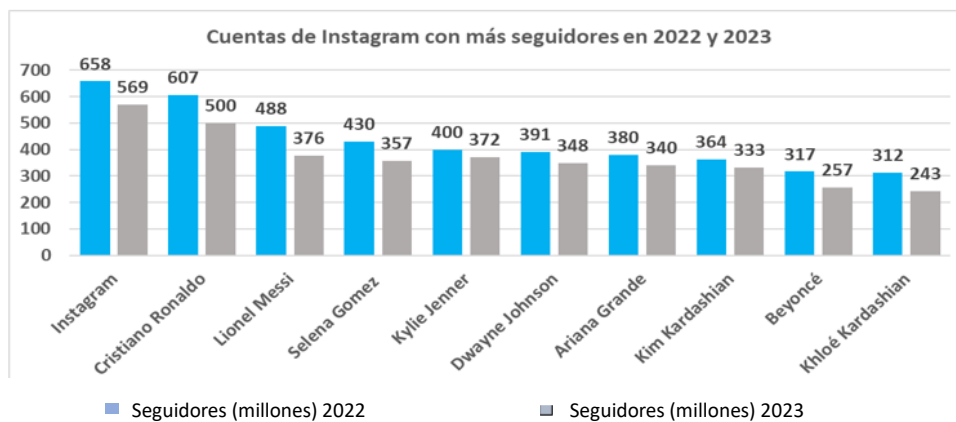


Figura 5.7. Cuentas de Instagram con más seguidores en 2022 y 2023.
Fuentes: Elaboración propia, con base en: <https://www.ayudacuentas.com/instagram/cuentas-con-mas-seguidores/> y <https://www.merca20.com/los-mas-seguidos-instagram-2022/>, (Excel, 2024).

El gráfico de doble barra permite observar, por ejemplo, que la comparación de las cuentas de Instagram en 2022 y 2023, muestra una ligera diferencia entre Kylie Jenner y Selena Gómez, estando Kylie Jenner en la cuarta posición en 2023, a diferencia del 2022, año en el que ocupó la cuarta posición Selena Gómez, por un margen de 30 millones. ¿Podrías hacer otra comparación extraída de esta tabla?

Ejemplo formativo 5.1

1. Con los datos obtenidos de una investigación sobre las carreras con mayor demanda en la Universidad Nacional Autónoma de México (UNAM) en los periodos 2022-2023 y 2023-2024 que aparecen en la Tabla 5.3, se ha elaborado un gráfico de doble barra que aparece a la derecha de la tabla.

Carreras con más demanda en la UNAM	2023-2024 ¹	2022-2023 ²
Medicina	23,153	31,213
Derecho	13,993	23,313
Psicología	13,029	21,538
Administración	9,058	17,266
Arquitectura	7,729	10,650
Cirujano dentista	7,550	9,303
Contaduría	7,446	13,390
Enfermería	6,980	7,883
Pedagogía	6,427	9,394
Medicina veterinaria y zootecnia	5,889	8,006



Tabla 5.3. Carreras con más demandada en la unam.

Fuente: Elaboración propia con base en: <https://blog.unitips.mx/carreras-mas-demandadas-de-la-unam> y <https://oferta.unam.mx> (Word, 2023).

2. De la observación de la gráfica, se aprecia rápidamente que en el ciclo 2023-2024 hubo menos demanda en todas las carreras que en el ciclo 2022-2023 y que en ambos ciclos, las carreras con mayor demanda fueron: Medicina, Derecho y Psicología y la de menos demanda, fue Medicina Veterinaria y Zootecnia.

¿Qué es un gráfico circular? Un gráfico circular es un diagrama que muestra los datos en sectores fáciles de entender. Cada sector, representa una categoría de datos y el tamaño, es proporcional a la cantidad que representa cada valor que toma la variable (para ello se te sugiere usar porcentajes).

Para elaborar un gráfico circular puedes usar una hoja de cálculo o bien a lápiz y papel, como se indica en el siguiente enlace:

<https://es.wikihow.com/hacer-un-gr%C3%A1fico-circular>

Ejemplo formativo 5.2

Con los datos obtenidos de la investigación, sobre las carreras con mayor demanda en la UNAM en el período 2023-2024 que aparecen en la Tabla 5.3, también se puede elaborar un gráfico circular con Microsoft Excel o a lápiz y papel. Para ello:

1. Se calcula el total de la demanda de carreras y el porcentaje que representó la cantidad de cada carrera.
2. En una hoja de cálculo de Excel, se escribe la etiqueta de cada dato en la columna de la izquierda, después se anotan los datos porcentuales en la fila del costado.
3. En insertar se hace clic en la opción “Gráficos” y luego en “Gráfico circular” para elegir el estilo que tenga el gráfico.

Se obtiene un gráfico como el que aparece en la Figura 5.8.



Figura 5.8. Carreras más demandadas UNAM 2023-2024.

Fuente: Elaboración propia, con base en: <https://blog.unitips.mx/carreras-mas-demandadas-de-la-unam> (Excel, 2024).

Gráfico de puntos e histograma

Un gráfico de puntos es útil para datos cuantitativos. Este se representa sobre una línea horizontal, organizando del menor de los datos hasta el mayor, y apilando los puntos hacia arriba uno del otro en los valores repetidos.

Supón que un comentarista de deportes, analizó el desempeño de la Liga MX Varonil, temporada de clausura 22-23, en el que participaron 18 equipos y cada uno jugó 34 partidos, en ambas clausuras. En correspondencia, elaboró un gráfico de puntos, como se muestra en la gráfica de la Figura 7.9 con el número de goles a favor (*GF*) anotados por cada equipo, utilizando los siguientes datos:

$$GF = \{46, 60, 53, 55, 62, 40, 50, 41, 48, 34, 51, 37, 48, 48, 28, 27, 37, 39\}$$

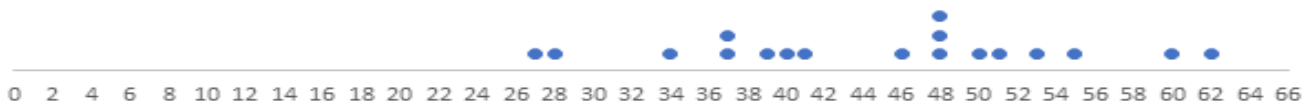


Figura 5.9. Goles a favor en la Liga MX temporada 2022-2023.

Fuente: Elaboración propia, con base en: <https://www.mediotiempo.com/futbol/liga-mx/tabla-general/385-2022> (Excel, 2024).

Observa en la gráfica anterior, que los datos presentan mucha dispersión y hay valores sin frecuencia, por lo que es más apropiado como gráfico un **histograma** y para crearlo, debes realizar los siguientes **pasos**:

Paso 1. Determina el Rango (R) = $Valor_{máximo} - Valor_{mínimo} = 62 - 27 = 35$

Paso 2. Determina el número de intervalos (que se denota con la letra k). Cálculalo con la \sqrt{n} , en donde n es el número de la muestra. $\sqrt{18} = 4.24$, se redondea a 4.

Paso 3. Determina la amplitud de los intervalos que se denota con la letra i , con la fórmula siguiente: $i = \frac{R}{k}$; por lo que $i = \frac{35}{4} = 8.75$, redondea a 9 unidades (Recuerda que siempre se debe cumplir la siguiente condición, $k \cdot i > R$). Así, para el primer intervalo, el límite inferior (dato mínimo igual a 27) le sumas 9 y dará el límite superior que es $36 = 27 + 9$; es decir, el primer intervalo es $[27, 36)$. De manera análoga construye los demás intervalos: $[36, 45)$, $[45, 54)$ y $[54, 63)$.

Paso 4. Organiza los datos por intervalos. Cuenta los datos que hay en cada intervalo y con ellos, obtén la frecuencia absoluta. El corchete al inicio del primer intervalo $[27 - 36)$ indica que ahí contarás los datos que son mayores o iguales a 27 (límite inferior) y menores que 36 (límite superior), como en la Tabla 5.4.

	Intervalos	Conteo	Frecuencia absoluta
A partir del 27 hasta antes del 36 →	$[27, 36)$	III	3
A partir del 36 hasta antes del 45 →	$[36, 45)$	HHH	5
A partir del 45 hasta antes del 54 →	$[45, 54)$	HHH II	7
A partir del 54 hasta antes del 63 →	$[54, 63)$	III	3

Tabla 5.4. Distribución de frecuencias agrupadas de goles a favor en la Liga MX temporada 2022-2023.
Fuente: Elaboración propia (Word, 2023).

Paso 5. Construye el histograma. En el eje X (abscisas), cada una de las barras tiene la amplitud del intervalo y en el eje Y (ordenadas), la frecuencia absoluta. Obsérvalo en la Figura 5.10.

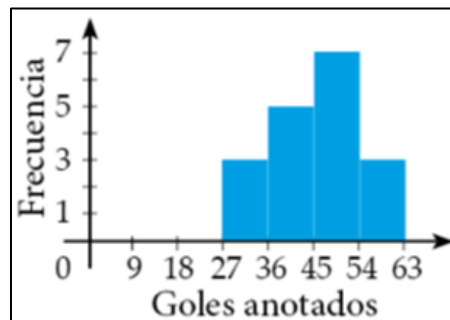


Figura 5.10. Histograma de goles a favor en la Liga MX temporada 2022-2023.
Fuente: Elaboración propia (Excel, 2024).

Observar el histograma nos ayuda a analizar mejor las variables cuantitativas en datos agrupados. Se observa el centro, la extensión y la forma del conjunto de datos, así como los valores atípicos o extremos, para verificar si hay posibles errores o problemas en los datos.

Polígono de frecuencias

Cálculo de la marca de clase	Intervalos	Marca de clase	Frecuencia absoluta
$\frac{27 + 36}{2} = 31.5 \rightarrow$	[27, 36)	31.5	3
$\frac{36 + 45}{2} = 40.5 \rightarrow$	[36, 45)	40.5	5
$\frac{45 + 54}{2} = 49.5 \rightarrow$	[45, 54)	49.5	7
$\frac{54 + 63}{2} = 58.5 \rightarrow$	[54, 63)	58.5	3

Un polígono de frecuencias se construye a partir de un histograma. Para ello, debes unir los puntos medios de cada intervalo en el histograma (marca de clase, que es el promedio de cada intervalo).

Tabla 5.5. Distribución de frecuencias agrupadas con marcas de clase.
Fuente: Elaboración propia (Word, 2023).

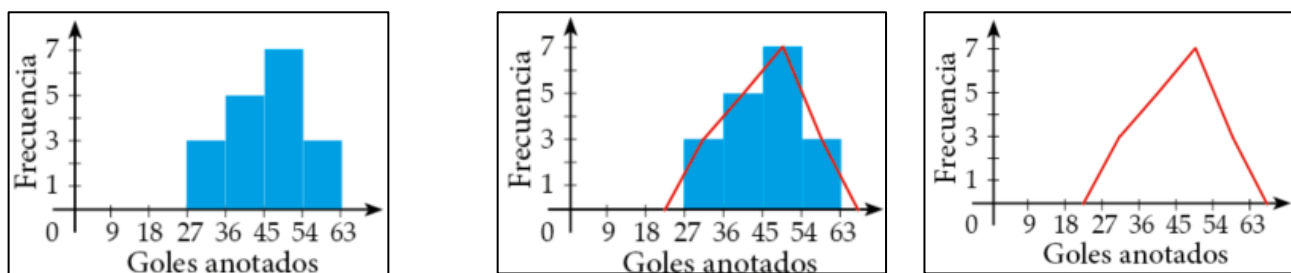


Figura 5.11. Proceso de construcción de un polígono de frecuencias a partir del histograma.
Fuente: Elaboración propia (Excel, 2024)

Ejemplo formativo 5.3

Gráfico de puntos, histograma y polígono de frecuencias.

Los datos siguientes, muestran el puntaje obtenido en 20 reactivos, por 40 alumnos de un grupo, en el último examen bimestral de la asignatura Pensamiento Matemático I.

10, 15, 11, 08, 12, 10, 13, 10, 12, 10, 12, 17, 10, 12, 11, 14, 15, 20, 10, 12, 10, 20, 14, 13, 05, 16, 05, 05, 14, 18, 07, 05, 12, 11, 02, 04, 14, 18, 16, 17.

Primeramente, se expresan los datos en orden ascendente:

02 04 05 05 05 05 07 08 10 10 10 10 10 10 10 11 11 11 12 12
12 12 12 12 13 13 14 14 14 14 15 15 16 16 17 17 18 18 20 20

1. Vamos a crear un gráfico de puntos, un histograma y un polígono de frecuencias, para ello te puedes apoyar en alguno de los siguientes dos softwares, por ejemplo: Canva® y GeoGebra.



Código QR 5.1. Cómo crear gráficos en Canva.
Fuente: Parzibyte, 2024.



Código QR 5.2. Cómo crear un gráfico de puntos en GeoGebra.
Fuente: Parzibyte, 2024.

a) Se obtiene, para el gráfico de puntos:

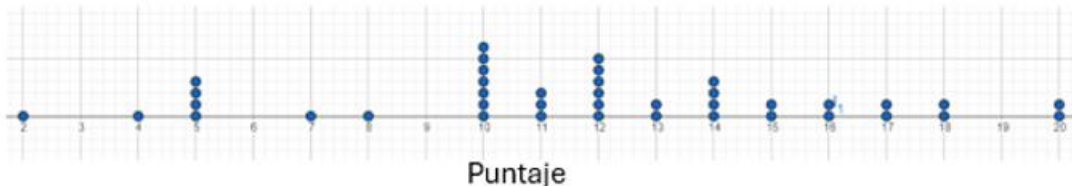


Figura 5.12a. Puntaje obtenido
Fuente: Elaboración propia, (Excel, 2024).

b) Para confeccionar el histograma:

$$\text{Rango: } (R) = \text{Valor}_{\text{máximo}} - \text{Valor}_{\text{mínimo}} = 20 - 2 = 18$$

Número de intervalos, si $n = 40$, $k = \sqrt{n} = \sqrt{40} \approx 6.32$, se redondea a 6

$$\text{Amplitud de los intervalos, } i = \frac{R}{k} = \frac{18}{6} = 3$$

El producto $i \cdot k = 18$, **no es** mayor estricto que el valor del rango, también 18, por lo que se añade un intervalo más y se organizan los datos en una tabla según los intervalos siguientes determinados:

[2,5), [5,8), [8,11), [11,14), [14,17), [17,20), [20,23)

Intervalos	Conteo	Frecuencia absoluta
[2,5)	II	2
[5,8)	II	5
[8,11)	III III	8
[11,14)	III III I	11
[14,17)	III III	8
[17,20)	III	4
[20,23)	II	2

Así, se obtiene el siguiente histograma:

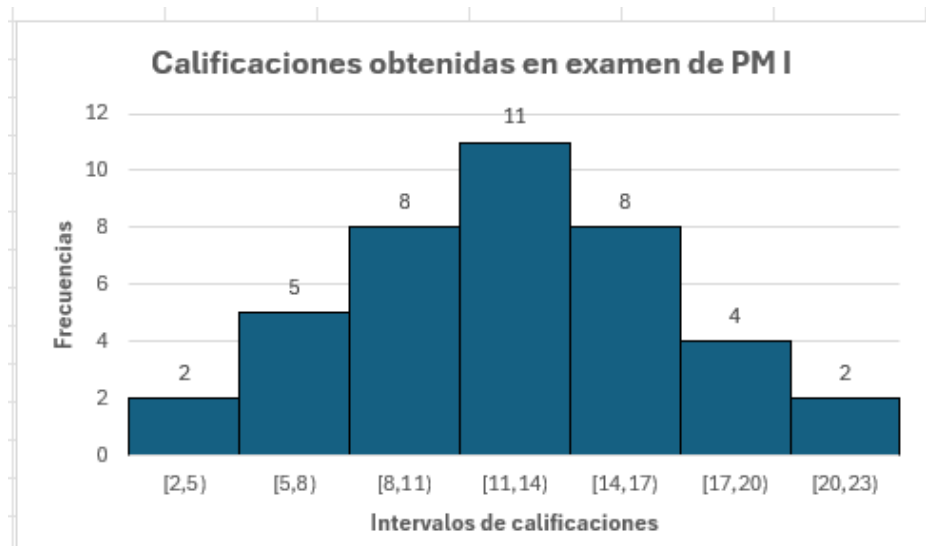


Figura 5.12b. Puntaje obtenido.
Fuente: Elaboración propia, (Excel, 2024).

c) El polígono de frecuencias se obtiene uniendo los puntos medios del borde superior de cada barra del histograma:

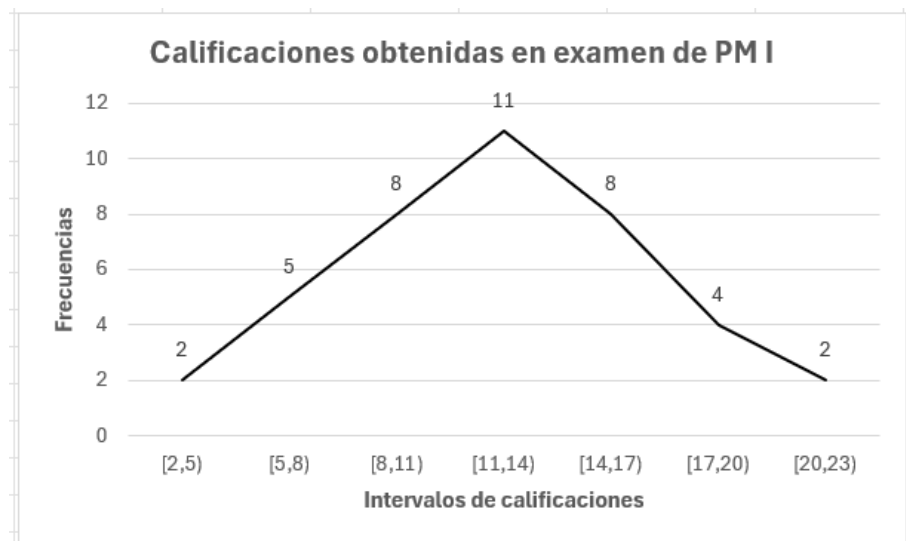


Figura 5.12c. Puntaje obtenido.
Fuente: Elaboración propia, (Excel, 2024).

- De cada uno de estos gráficos, puedes obtener información sobre las calificaciones registradas:
 - En el gráfico de puntos se observa que el rango es 18, hay cierta concentración de las calificaciones entre los valores 10 y 12, por lo que se aprecia una mayor distribución de calificaciones hacia el lado derecho del valor 12.
 - En el histograma, se observa una cierta uniformidad de la distribución hacia ambos lados respecto a un valor central, con determinada simetría.

- En el polígono de frecuencias, se observa esa distribución hacia ambos lados del polígono, y por tanto no hay sesgo respecto a ninguno de los dos lados de la distribución de las calificaciones.
3. Sobre las frecuencias de las calificaciones, también a partir de los gráficos se pueden apreciar determinadas regularidades
- Se concentran 16 de las 40 calificaciones entre 10 y 12. A la izquierda de este grupo hay 8 calificaciones y a la derecha hay 16.
 - Las notas 10, 12, y 14 aparecen con la mayor frecuencia. A esos valores, están cercanos los valores 11 y 13, lo que representa cierta agrupación alrededor de estas notas.
 - Por otra parte, no hay muchos valores extremadamente bajos o altos, lo que indica cierta distribución normal alrededor de la media en los resultados.

Al analizar los datos en las representaciones gráficas, determinadas formas en su comportamiento, pueden indicar una regularidad estadística sobre dichos datos, misma que se utiliza para describirlos o analizar en qué medida los datos analizados se asemejan o difieren entre sí.

Tendencias observables

Las tablas y gráficas estadísticas, permiten el análisis y facilitan la comprensión de la información. A través de estas herramientas visuales, es posible identificar patrones y comportamientos específicos conocidos como tendencias observables. En este contexto, se distinguen tres grupos fundamentales de estas: tendencia estadística, tendencia central y tendencia de dispersión de los datos.

Tendencia estadística. Es la dirección, o el patrón que presentan los datos al graficarlos, ya sea que provengan de una o más variables. Esta tendencia se clasifica como ascendente, descendente o estacionaria. Algunos ejemplos son:

- Si un puesto de esquite registra un crecimiento constante en sus ventas en los meses de enero a mayo de 2023, eso representa una tendencia estadística ascendente.
- Si de las 18:00 a las 23:00 horas del 8 de enero de 2024, los datos climáticos muestran una disminución gradual de la temperatura en Culiacán, eso indica una tendencia estadística descendente.

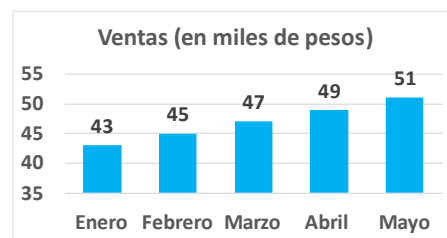


Figura 5.13. Ventas.

Fuente: Elaboración propia (Excel, 2024).

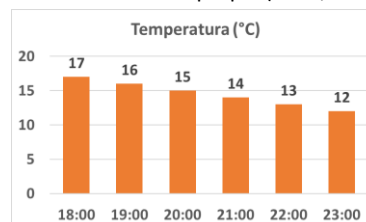


Figura 5.14. Temperatura en °C.

Fuente: Elaboración propia (Excel, 2024).

- Si el número de pares de zapatos, que tienen los estudiantes de un grupo de primer grado de la Facultad de Medicina muestran un incremento o decremento similar, eso revela una tendencia estadística estacionaria.



Figura 5.15. Pares de zapatos.

Fuente: Elaboración propia (Excel, 2024).

Tendencia central. Se refiere al valor de referencia que representa a la mayoría de los datos de un conjunto. Se emplean tres medidas principales para identificarla: la media aritmética, la mediana y la moda. La **mediana**, te indica el punto central de los datos y como se concentran éstos a su alrededor. La **mediana**,



Figura 5.16. Televisores por hogar.

Fuente: Elaboración propia (Excel, 2024).

te muestra el punto central de los datos sólo si éstos están ordenados. La **moda**, te indica el dato o los datos con mayor frecuencia del conjunto de datos.

La gráfica de la izquierda, muestra el número de televisores por hogar en un conjunto habitacional privado de 15 casas. En él, se observa que la mediana es tres televisores y la moda es dos televisores por hogar.

Tendencia de dispersión. Es el valor de referencia que representa la distribución de los datos, alrededor de sus medidas de tendencia central. Una cualidad de esta dispersión, es que un valor grande indica que los datos están más alejados o dispersos de la medida de tendencia central, mientras un valor pequeño (cercano al cero), indica que los datos tienden a estar más concentrados alrededor de la medida de tendencia central. Una medida de tendencia de dispersión muy usada es la desviación estándar, la cual indica el grado de agrupamiento de los datos alrededor de la media aritmética.

Imagina que estás analizando el tiempo destinado al día para ver videos de TikTok, por estudiantes del grupo A (Figura 5.17a) y del grupo C (Figura 5.17b). Observa que en el histograma del grupo A, los datos están más dispersos que en el histograma del grupo C, por lo que el grupo A tiene una mayor desviación estándar que el grupo C.



Figura 5.17a. Tiempo destinado al día para ver videos de TikTok por estudiantes del grupo A.
Fuente: Elaboración propia (Excel, 2024).

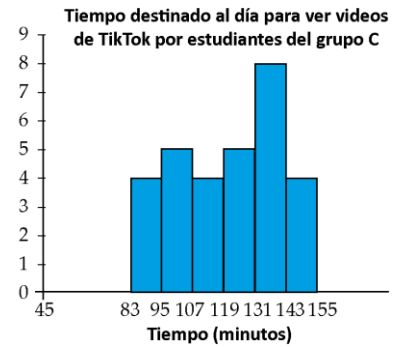


Figura 5.17b. Tiempo destinado al día para ver videos de TikTok por estudiantes del grupo C.
Fuente: Elaboración propia (Excel, 2024).

Ejemplo formativo 5.4

Analicemos si se aprecia alguna tendencia observable en las siguientes gráficas:

a) En la Figura 5.18 aparece la distribución del peso de 20 estudiantes.

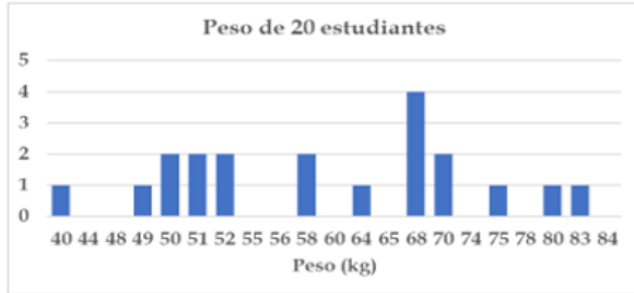


Figura 5.18. Peso de 20 estudiantes.
Fuente: Elaboración propia (Excel, 2024).

Se aprecia una tendencia de dispersión, con valores alejados de las medidas de tendencia central. (Media: 61.6, Mediana: 61, Moda: 68)

b) En la Figura 5.19 se muestra el tiempo invertido en plataformas y dispositivos digitales



Figura 5.19. Tiempo invertido en plataformas y dispositivos digitales.
Fuente: Elaboración propia (Excel, 2024).

Tendencia estadística descendente. La Moda muestra una tendencia descendente ya que la mayoría de los datos se refieren al tiempo invertido en internet y en las demás formas va disminuyendo.

c) En la figura 5.20 aparecen los resultados de juegos ganados y juegos perdidos de la Liga Mexicana del Pacífico 2022-2023



Figura 5.20. Resultados de JG y JP de la LMO 2022-2023.
 Fuente: Elaboración propia (Excel, 2024).

Tendencia estadística estacionaria.

En general se aprecia una tendencia estacionaria, tanto en los juegos ganados como perdidos.

Evaluación formativa 5.1

1. Pregunta a 20 compañeros de la clase, su estatura en centímetros y elabora: el gráfico de puntos, histograma y polígono de frecuencias.

Datos obtenidos:

2. Argumenta el comportamiento de dichas gráficas:

3. De la gráfica con las calificaciones de los estudiantes del grupo en el Ejemplo formativo 5.3, analiza si se destaca:

- a) Una tendencia creciente, decreciente o estacionaria: _____
- b) Un valor que más se repite o moda en la cantidad de datos: _____
- c) El punto medio o mediana de los datos: _____
- d) Un valor que se aleje sustancialmente de los datos: _____

4. Calcula el promedio o media aritmética de los datos obtenidos: _____

5. ¿Se aprecia si los datos se concentran, o dispersan respecto a alguno de los valores de tendencia central?

Justifica tu respuesta: _____

6. Analiza si se aprecia alguna tendencia observable en la gráfica de la Figura 5.21:

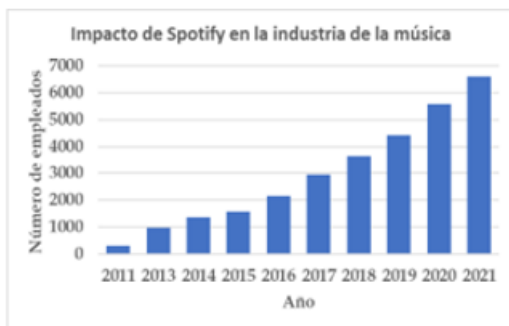


Figura 5.21. Impacto de Spotify en la industria de la música.

Fuente: Elaboración propia (Excel, 2024).

Autoevaluación y coevaluación 5.1

Nombre: _____ Plantel: _____ Grupo: _____ Turno: _____

Autoevaluación para el aprendizaje

Selecciona en la columna la opción que mejor refleje tu nivel de desempeño en el proceso para el aprendizaje de la progresión de aprendizaje 5. Responde con honestidad a la evaluación de cada uno de los criterios que se enlistan a continuación.

Desempeño	En proceso de logro	Bueno	Sobresaliente
Propicié un clima de comunicación favorable para el aprendizaje con mis compañeros.			
Participé activamente con ideas para la toma razonada de decisiones.			
Contribuí colaborativamente en la retroalimentación de dudas de mis compañeros.			
Desarrollé en el orden previsto los pasos necesarios para la construcción de diferentes tipos de gráficos (M1-C1).			
Seleccioné, en correspondencia con el tipo de datos obtenidos de un evento, el tipo de gráfico apropiado para representarlos (M2-C1).			
Identifiqué tendencias y comportamientos específicos de los datos correspondientes a determinados eventos a través de las tablas y gráficas estadísticas (M2-C2).			

Coevaluación para el aprendizaje

Solicita a un compañero del equipo, que marque en la columna la opción que mejor describa tu desempeño durante el trabajo colectivo, concluida la progresión de aprendizaje 5, y que responda con honestidad la evaluación de cada uno de los criterios que se enlistan a continuación.

Desempeño	En proceso de logro	Bueno	Sobresaliente
Propició un clima de comunicación favorable para el aprendizaje con mis compañeros.			
Participó activamente con ideas para la toma razonada de decisiones.			
Contribuyó colaborativamente en la retroalimentación de dudas de mis compañeros.			
Desarrolló en el orden previsto los pasos necesarios para la construcción de diferentes tipos de gráficos (M1-C1).			
Seleccionó, en correspondencia con el tipo de datos obtenidos de un evento, el tipo de gráfico apropiado para representarlos (M2-C1).			
Identificó tendencias y comportamientos específicos de los datos correspondientes a determinados eventos, a través de las tablas y gráficas estadísticas (M2-C2).			

Nombre y firma de quien coevalúa

La relación entre variables categóricas

Progresión de aprendizaje 6

Analiza cómo se relacionan entre sí, dos o más variables categóricas a través del estudio de alguna problemática o fenómeno de interés para el estudiantado, con la finalidad de identificar si dichas variables son independientes.

Metas de aprendizaje	En proceso de logro	Bueno	Sobresaliente
M1-C2 Observa y obtiene información de una situación o fenómeno para establecer estrategias o formas de visualización que ayuden a entenderlo.	A		
	C		
	H		
M3-C2 Compara hechos, opiniones o afirmaciones para organizarlos en formas lógicas útiles en la solución de problemas y explicación de situaciones y fenómenos.	A		
	C		
	H		
M4-C2 Argumenta a favor o en contra de afirmaciones acerca de situaciones, fenómenos o problemas propios de la matemática, de las ciencias o de su contexto.	A		
	C		
	H		

Ya has aprendido a representar gráficamente de diferentes formas, la información o datos que se obtienen sobre un conjunto de personas, objetos o sobre la ocurrencia de un evento.

Para el estudio de una o más variables, sobre un determinado suceso, es importante diferenciar la naturaleza de aquellas variables que son **cuantitativas**, es decir las que se pueden medir en una escala numérica, de las que tienen un carácter **cualitativo**. Se le llama **variable categórica**, a la que se utiliza en el campo de la estadística, para designar una característica o propiedad no-numérica o cualitativa a algún individuo, condición o suceso. Se utilizan básicamente, para clasificar o calificar los elementos de un conjunto considerado.

Son variables categóricas las siguientes:

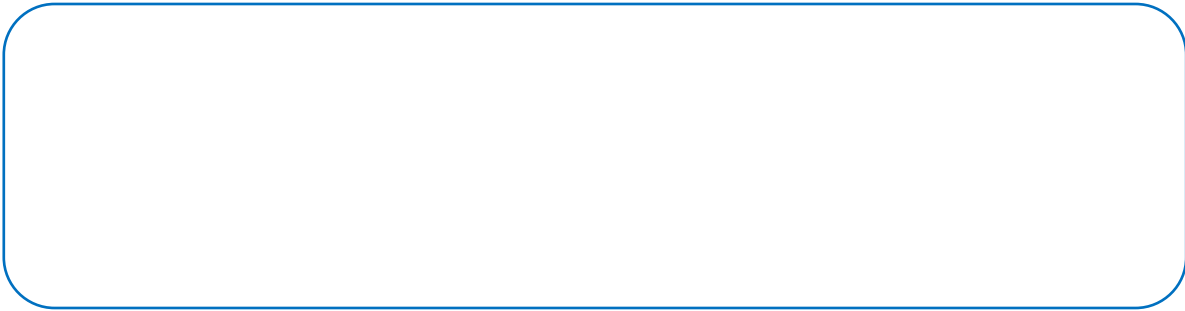
- Color, cuyos valores pueden ser: azul, verde, rojo, blanco, entre otros.
- Sexo, con los valores: masculino y femenino.
- Grupo sanguíneo, con rango de valores: A , B , AB , O .
- Nivel de educación: primaria, secundaria, medio superior, superior.

Por ejemplo, si en un hospital hay 170 enfermos hospitalizados, el grupo sanguíneo, que es una variable categórica, se registra dentro de la historia clínica. Si bien es una variable no numérica, es posible saber cuántos, del total de enfermos, están en cada rango de valores. Supongamos que se obtiene lo siguiente:

Grupo:	A	B	AB	O
Cantidad:	45	50	35	40

Evaluación diagnóstica 6.1

Con los datos anteriores, utilizando Excel, confecciona un gráfico de barras para representarlos.



1. ¿Se aprecia algún tipo de tendencia en dichos datos? En ese caso, ¿cuál?

2. ¿Qué medida de tendencia central es más apropiada para caracterizar estos datos?

Para conocer la tendencia o valor más probable de una variable categórica, se utiliza como medida de tendencia central la *moda*, pues significa el valor que más se repite en un conjunto de datos.

Ejemplo formativo 6.1

Vamos a identificar en cada una de las siguientes expresiones, las variables categóricas y las cuantitativas.

1. En una investigación epidemiológica, se analiza la influencia que tiene la **vacunación, contra cierta enfermedad en las personas que contraen esa enfermedad.**
Vacunación contra cierta enfermedad: categórica
Personas que contraen esa enfermedad: cuantitativa
2. En una investigación sobre la anemia, se estudia la relación que hay entre **el consumo de determinados alimentos y la concentración de hemoglobina.**
Consumo de determinados alimentos: categórica
Concentración de hemoglobina: cuantitativa
3. En una investigación realizada por oculistas, se busca determinar la influencia que tiene **el tiempo que utiliza el celular en la fatiga ocular.**
Tiempo que se utiliza el celular: cuantitativa
Fatiga ocular: categórica
4. En una investigación realizada por médicos, se analiza la relación que hay entre **el tipo de grasas y la presión arterial.**
Tipos de grasas que consume: categórica
Nivel de la presión arterial: cuantitativa

5. En un estudio realizado por médicos, se analiza cómo **la actividad física mejora la circulación de la sangre.**
Actividad física: categórica
Circulación de la sangre: categórica
6. En una investigación realizada por gastroenterólogos, se busca determinar cómo **la cantidad de café, puede modificar los niveles de reflujo del ácido estomacal.**
Cantidad de café: cuantitativa
Nivel de reflujo del ácido estomacal: categórica
7. En una investigación realizada por estudiantes de ciencias de la educación, se busca establecer cómo **el método de estudio elegido, influye en el rendimiento académico.**
Método de estudio: categórica
Rendimiento académico: cuantitativa

El análisis del comportamiento de una variable, al estudiar un individuo o suceso, nos permite valorar qué repercusión puede tener sobre él mismo. Sin embargo, también es posible considerar, el estudio simultáneo de dos variables y su efecto sobre un mismo suceso o situación, preguntarnos además, si estudiamos a la vez esas dos variables lo siguiente: ¿habrá una relación entre ambas?, ¿influirán los valores de una en los valores de la otra?

Las **variables bidimensionales** son aquellas que, desde el punto de vista estadístico, miden dos o más características diferentes de un mismo individuo, suceso o situación y posibilitan estudiar la relación entre ellas.

Como ejemplo, podemos mencionar la problemática relacionada con la pandemia del COVID-19, en la que las autoridades del sector salud, luego de realizar estudios estadísticos, determinaron que la aplicación de vacunas a la población, reduce de manera muy significativa los riesgos de contagio.

Tomando como referencia el ejemplo anterior, ¿cuál es el suceso o situación que se valora?, ¿cuáles serían las variables mencionadas?, ¿consideras que existe relación de dependencia entre ellas?

Pero, no siempre se puede presuponer que una serie de variables sobre un determinado suceso tienen una relación, puesto que pueden aparentar que sí, y sin embargo, no tener ninguna; por lo que es acertado, estudiar bajo qué situaciones hay una relación entre dos variables y en cuáles no se manifiesta influencia entre ellas.

La importancia de su estudio, radica en que la relación entre dos variables juega un papel muy relevante, tanto en el campo social y económico, como para la salud y el

examen de fenómenos naturales, pues su estudio y análisis adecuados favorecen la toma de decisiones asertivas.

En el análisis de variables bidimensionales, se dice que dos variables son dependientes, si el conocer el valor de una de ellas, afecta la distribución de la otra variable Y son independientes, si el conocer el valor de una de ellas, no afecta la distribución de la otra variable.

Una de las formas más apropiadas, para el registro de las variables bidimensionales, son las tablas de doble entrada o de contingencia, las cuales se construyen a partir de una matriz de filas y columnas en las que se registran los datos; en las columnas se expresan las categorías de información y en las filas se agregan los datos.

Ejemplo formativo 6.2

La siguiente, es una tabla de doble entrada o de contingencia, que muestra el número de jugadores de béisbol de la Liga Mexicana del Pacífico de México, que hacen estiramiento antes de cada juego, y cuánto tuvieron lesiones durante la temporada regular.

	Lesión durante el año pasado	No lesión durante el año pasado	Total
Hace estiramientos	45	255	300
No hace estiramientos	189	211	400
Total	234	466	700

- Las variables que se contrastan en la situación de este ejemplo son: los estiramientos y las lesiones.
- Ambas variables son categóricas, dado que:
 - La variable estiramientos, tiene como categorías: sí y no.
 - La variable lesión durante el año pasado, tiene como categorías: sí y no.
- La siguiente tabla muestra los datos con los porcentajes correspondientes.

	Lesión durante el año pasado	%	No lesión durante el año pasado	%	Total
Hace estiramientos	45	$\frac{45}{300} = 0.15$ 15 %	255	$\frac{255}{300} = 0.85$ 85 %	300
No hace estiramientos	189	$\frac{189}{400} = 0.4725$ 47.25%	211	$\frac{211}{400} = 0.5275$ 52.75 %	400
Total	234	62.25 %	466	137.75 %	700

Si se analizan los datos de la tabla en forma porcentual es posible inferir algunas conclusiones no tan fáciles de determinar si estos estuvieran solo listados. El número de **lesiones** cuando **se hace estiramiento** es equivalente al 15%, mientras que la proporción de **no lesiones** cuando **se hace estiramiento** es igual al 85%. A su vez, la proporción de **lesiones** cuando **se hace estiramiento** con respecto a cuando **no se hace estiramiento** también es diferente. Todo parece indicar que hay menos lesiones si se hace estiramiento y más lesiones si no se hace. Como resultado, se puede inferir que si un deportista hace estiramientos tiene menos probabilidad de sufrir lesiones.

Cuando se comparan dos variables categóricas con valores binarios, es decir, solo dos valores posibles para la variable, el análisis de las proporciones puede ser útil para determinar si existe una relación entre ambas variables. En otras palabras, si las proporciones de dos variables son similares, entonces no hay una relación significativa entre ellas y pueden ser estadísticamente independientes. Por otro lado, si las proporciones son distintas, entonces hay una relación significativa entre las dos variables y pueden ser estadísticamente dependientes.

Sin embargo, en general, el análisis de las proporciones no es suficiente para determinar si existe una relación entre las variables categóricas. Para determinar con más precisión si hay relación entre las variables se utiliza una prueba estadística denominada Chi-cuadrado.

La **prueba Chi-cuadrado** es un procedimiento estadístico utilizado para analizar la relación entre dos variables categóricas. Para calcular la prueba Chi-cuadrado, se utilizan los datos de una tabla de contingencia, que muestra la frecuencia de los valores de cada variable. A través de esta prueba se busca determinar si una diferencia, entre

los datos obtenidos y los datos esperados, se debe al azar o si existe una relación entre las variables consideradas.

Ejemplo formativo 6.3

Supón que necesitas conocer si la preferencia por el consumo de sodas o refrescos está relacionada con el género de los jóvenes y, para ello en la Preparatoria Guasave Diurna realizas una encuesta a 100 estudiantes, con el resultado que muestra la Tabla 6.1.

Frecuencia observada (f_o)	Preferencia Baja	Preferencia Alta	Total
Mujeres	21	26	47
Hombres	15	38	53
Total	36	64	100

Tabla 6.1. Preferencia del consumo de refrescos por género.

Fuente: Elaboración propia (Word, 2023).

Como investigador quieres especificar la respuesta con un nivel de significancia de $\alpha = 0.05$. El **nivel de significancia** es un valor que se utiliza en estadística para determinar si un resultado es estadísticamente significativo. Un nivel de significancia del 0.05% indica que hay un 5% de probabilidad de que los resultados observados sean debidos al azar y no a una causa real.

Se parte del criterio de que la preferencia, por las sodas y el género de las personas son independientes o que no existe relación entre ellas y se le llama hipótesis 0 o hipótesis nula. En contraposición, se llama hipótesis 1 o alternativa a la afirmación de que sí hay relación entre ambas:

H_0 : El género y la preferencia por las sodas son independientes.

H_1 : El gusto por las sodas está relacionado con el género de la persona.

La prueba Chi-cuadrado, (se denota por χ^2), sirve para determinar si se acepta la hipótesis nula H_0 o si se rechaza y, en consecuencia, se acepta la hipótesis alternativa H_1 .

Una vez confeccionada la tabla de contingencia, para aplicar la prueba Chi-cuadrado se procede de la siguiente manera:

- **Calcula la frecuencia esperada (f_e)** de cada celda, que es el número de observaciones que se esperaría en cada celda si no hubiera relación entre las variables, como se expresa en H_0 . Para ello, se multiplica la frecuencia total de la fila correspondiente a esa celda por la frecuencia total de la columna correspondiente a la misma celda y se divide por el tamaño total de la muestra que aparece en la Tabla 6.2, y se obtiene:

Frecuencia esperada	Preferencia Baja	Preferencia Alta	Total
Hombres	$f_e = \frac{36 \cdot 47}{100} = 16.92$	$f_e = \frac{64 \cdot 47}{100} = 30.08$	47
Mujeres	$f_e = \frac{36 \cdot 53}{100} = 19.08$	$f_e = \frac{64 \cdot 53}{100} = 33.92$	53
Total	36	64	100

Tabla 6.2. Frecuencia esperada del consumo de refrescos por género.

Fuente: Elaboración propia (Word, 2023).

- **Calcula la desviación entre la frecuencia observada y la frecuencia esperada** de cada celda. Para calcular la desviación, se resta la frecuencia esperada de la frecuencia observada y se eleva al cuadrado; luego, se divide por la frecuencia esperada. A continuación, suma todas las desviaciones obtenidas. El resultado obtenido es la estadística de prueba Chi-cuadrado.

$$\chi^2 = \frac{(21 - 16.92)^2}{16.92} + \frac{(26 - 30.08)^2}{30.08} + \frac{(15 - 19.08)^2}{19.08} + \frac{(38 - 33.92)^2}{33.92}$$

$$= 0.9838 + 0.5534 + 0.8725 + 0.4908 = 2.9005$$

- **Determina el número de grados de libertad de la prueba.** El número de grados de libertad se calcula restando uno del número de filas y uno del número de columnas de la tabla de contingencia; por tanto, en este ejemplo, los datos registrados están en $f = 2$ filas y $c = 2$ columnas, por lo tanto, el número de grados de libertad es: $k = (2 - 1)(2 - 1) = 1$.
- **Consulta la tabla de distribución Chi-cuadrado para determinar el valor crítico de la prueba** (una sección de esta aparece en la Tabla 6.3). El valor crítico depende del nivel de significancia y el número de grados de libertad. En este caso para un nivel de significancia de $\alpha = 0.05$ y un grado de libertad se obtiene como valor crítico 3.841, como se muestra en la Tabla 6.3.

	0.995	0.99	0.975	0.95	0.9	0.5	0.2	0.1	0.05	0.025	0.02	0.01
1	0.0000397	0.000157	0.000982	0.00393	0.0158	0.455	1.642	2.706	3.841	5.024	5.412	6.635
2	0.0100	0.020	0.051	0.103	0.211	1.386	3.219	4.605	5.991	7.378	7.824	9.210
3	0.072	0.115	0.216	0.352	0.584	2.366	4.642	6.251	7.815	9.348	9.837	11.345
4	0.207	0.297	0.484	0.711	1.064	3.357	5.989	7.779	9.488	11.143	11.668	13.277
5	0.412	0.554	0.831	1.145	1.610	4.351	7.289	9.236	11.070	12.833	13.388	15.086
6	0.676	0.872	1.237	1.635	2.204	5.348	8.558	10.645	12.592	14.449	15.033	16.812
7	0.989	1.239	1.690	2.167	2.833	6.346	9.803	12.017	14.067	16.013	16.622	18.475
8	1.344	1.646	2.180	2.733	3.490	7.344	11.030	13.362	15.507	17.535	18.168	20.090
9	1.735	2.088	2.700	3.325	4.168	8.343	12.242	14.684	16.919	19.023	19.679	21.666
10	2.156	2.558	3.247	3.940	4.865	9.342	13.442	15.987	18.307	20.483	21.161	23.209

Tabla 6.3. Tabla de distribución de la Chi-cuadrado.

Fuente: Disfruta las Matemáticas, <https://www.disfrutalasmaticas.com/datos/tabla-chi-cuadrado.html>

- **Compara el valor de la estadística de prueba Chi-cuadrado obtenida con el valor crítico obtenido.** Si el valor de la estadística de prueba es mayor que el valor crítico, se rechaza la hipótesis nula de que no hay relación entre las variables. Si el valor de la estadística de prueba es menor o igual que el valor crítico, se acepta la hipótesis nula.

En este ejemplo, $\chi^2 = 2.9005 < 3.841$, por lo que se acepta la hipótesis nula que expresa que el género y la preferencia por las sodas son independientes, es decir que la preferencia por las sodas o refrescos no está vinculada con el género de la persona, con un nivel de significación de 5%.

Siguiendo estos pasos es posible, para otras situaciones, calcular la prueba Chi-cuadrado y determinar si existe o no una relación entre dos o más variables que sean categóricas; pero los cálculos pueden simplificarse mucho haciendo uso de herramientas digitales. Para ello escanea el código QR 6.1.

Al abrir la calculadora digital para la prueba Chi-cuadrado, aparece una tabla en la que puedes rellenar los valores obtenidos de las variables y obtienes de inmediato el valor de χ^2 , como se muestra en la Figura 6.1 para el ejemplo anterior.

Para realizar una prueba de independencia de Chi-cuadrado, simplemente completa las celdas a continuación para obtener una tabla de contingencia de hasta 5 filas y 5 columnas. Si su tabla es más pequeña que 5×5 , simplemente deje las otras celdas en blanco.

	Grupo 1	Grupo 2
Categoría 1	<input type="text" value="21"/>	<input type="text" value="26"/>
Categoría 2	<input type="text" value="15"/>	<input type="text" value="38"/>
Categoría 3	<input type="text"/>	<input type="text"/>
Categoría 4	<input type="text"/>	<input type="text"/>
Categoría 5	<input type="text"/>	<input type="text"/>

Estadístico de prueba χ^2 : 2.900317



QR 6.1. Calculadora de prueba de independencia de Chi-cuadrado.
Fuente: Parzibyte, 2024.

Figura 6.1. Tabla de contingencia en la calculadora digital de *statologos*, para realizar la prueba Chi-cuadrado.
Fuente: <https://statologos.com/calculadora-prueba-de-independencia-chi-cuadrado/>

Observa, que el valor dado por la calculadora digital es: $\chi^2 = 2.9005$, el cual coincide con el calculado anteriormente. También, puedes usar aplicaciones de celular como Mathematics, GeoGebra, entre otras.

Ejemplo formativo 6.4

Vamos a aplicar la prueba de independencia Chi-cuadrado, con un nivel de significancia 0.01, a lo analizado en el ejemplo formativo 6.2, para determinar si existe relación entre las variables: jugadores de beisbol que hacen estiramiento o no antes de cada juego y la posibilidad de tener o no lesiones.

1. Partimos de la tabla de contingencia 2x2.

Frecuencia observada (f_o)	Lesión durante el año pasado	No lesión durante el año pasado	Total
Hace estiramientos	45	255	300
No hace estiramientos	189	211	400
Total	234	466	700

2. Se formulan las hipótesis H_0 y H_1 .

H_0 : los estiramientos y las lesiones, son independientes.

H_1 : Los estiramientos, están relacionados con las lesiones de los jugadores de beisbol.

3. Se aplica la calculadora prueba de independencia Chi-cuadrado y se obtiene:

$$\chi^2 \approx 80.1313.$$

4. Se determina el grado de libertad:

$$(\text{Número de filas} - 1)(\text{Número de columnas} - 1) = (2 - 1)(2 - 1) = (1)(1) = 1$$

5. Se obtiene, en la tabla de distribución Chi-cuadrado el valor crítico de la prueba para ese grado de libertad y nivel de significancia:

Para un valor de significancia $\alpha = 0.01$ y un grado de libertad, el valor crítico es igual a 6.635.

6. Finalmente se comparan los valores obtenidos por la calculadora y por la tabla:

$$\chi^2 \approx 80.1313 > 6.635$$

7. Dado que el valor del estadístico de prueba, es mayor que el valor crítico, $\chi^2 \approx 80.1313 > 6.635$, se rechaza la hipótesis H_0 . Es decir, los estiramientos están relacionados con las lesiones de los jugadores de beisbol.

Si quieres saber más sobre la prueba Chi-cuadrado escanea el código QR 6.2.



Evaluación [formativa 6.1](#)

La siguiente tabla muestra la relación que existe entre el consumo de comida chatarra y la obesidad en alumnos del primer grado de la preparatoria Guasave Diurna.

	Presenta obesidad	No presenta obesidad	Total
Consume comida chatarra	300	290	590
No consume comida chatarra	290	320	610
Total	590	610	1200

1. ¿Qué variables se contrastan? _____

2. ¿A qué tipo de variables hace referencia? ¿cuantitativas o categóricas? Argumenta tu respuesta. _____

3. Elabora una hipótesis nula H_0 y la alternativa H_1 , después aplica la prueba de independencia Chi-cuadrado con un nivel de significancia de 0.01, utilizando alguna herramienta digital.
 H_0 : _____
 H_1 : _____
4. ¿A qué conclusión llegas? _____

Autoevaluación y coevaluación 6.1

Nombre: _____ Plantel: _____ Grupo: _____ Turno: _____

Autoevaluación para el aprendizaje

Selecciona en la columna la opción que mejor refleje tu nivel de desempeño en el proceso para el aprendizaje de la progresión de aprendizaje 6. Responde con honestidad a la evaluación de cada uno de los criterios que se enlistan a continuación.

Desempeño	En proceso de logro	Bueno	Sobresaliente
Propicié un clima de comunicación favorable para el estudio de la relación entre dos variables categóricas.			
Participé activamente con ideas para identificar si dos variables categóricas son independientes.			
Contribuí colaborativamente en la retroalimentación de dudas sobre variables categóricas.			
Analicé dos variables categóricas para verificar su independencia (M3-C2).			
Usé las proporciones para verificar si existe relación entre dos variables categóricas (M4-C2).			
Justifiqué mediante la prueba Chi-cuadrado si existe relación entre dos variables categóricas (M4-C2).			

Coevaluación para el aprendizaje

Solicita a un compañero del equipo que marque en la columna la opción que mejor describa tu desempeño durante el trabajo colectivo, concluida la progresión de aprendizaje 6, y que responda con honestidad la evaluación de cada uno de los criterios que se enlistan a continuación.

Desempeño	En proceso de logro	Bueno	Sobresaliente
Propició un clima de comunicación favorable para el estudio de la relación entre dos variables categóricas.			
Participó activamente con ideas para identificar si dos variables categóricas son independientes.			
Contribuyó colaborativamente en la retroalimentación de dudas sobre la relación entre dos variables categóricas.			
Analizó dos variables categóricas para verificar su independencia (M3-C2).			
Usó las proporciones para verificar si existe relación entre dos variables categóricas (M4-C2).			
Justificó mediante la prueba Chi-cuadrado si existe relación entre dos variables categóricas (M4-C2).			

Nombre y firma de quien coevalúa

La relación entre variables cuantitativas

Progresión de aprendizaje 7

Analiza dos o más variables cuantitativas a través del estudio de alguna problemática o fenómenos de interés para el estudiantado, con la finalidad de identificar si existe correlación entre dichas variables.

Metas de aprendizaje		En proceso de logro	Bueno	Sobresaliente
M1-C2 Observa y obtiene información de una situación o fenómeno para establecer estrategias o formas de visualización que ayuden a entenderlo.	A			
	C			
	H			
M3-C2 Compara hechos, opiniones o afirmaciones para organizarlos en formas lógicas útiles en la solución de problemas y explicación de situaciones y fenómenos.	A			
	C			
	H			
M4-C2 Argumenta a favor o en contra de afirmaciones acerca de situaciones, fenómenos o problemas propios de la matemática, de las ciencias o de su contexto.	A			
	C			
	H			

Evaluación diagnóstica 7.1

- ¿Qué son las variables cuantitativas?
 - Variables que se miden en categorías.
 - Variables que se miden en términos numéricos.
 - Variables que se describen en términos cualitativos.
 - Variables que no se pueden medir.
- En un estudio sobre el rendimiento académico, ¿qué tipo de variable es la calificación promedio de los estudiantes?
 - Cualitativa nominal.
 - Cualitativa ordinal.
 - Cuantitativa discreta.
 - Cuantitativa continua.
- ¿Qué es un gráfico de dispersión?
 - Un gráfico que muestra la distribución de una variable.
 - Un gráfico que muestra la relación entre dos variables mediante puntos en un plano cartesiano.
 - Un gráfico que muestra la frecuencia de una variable.
 - Un gráfico que muestra la mediana y los cuartiles de una variable.

En la progresión anterior, analizaste la relación entre dos variables categóricas, así como, la dependencia e independencia de las variables. En esta progresión vas a relacionar dos o más variables cuantitativas a través de problemáticas de interés mediante el coeficiente de correlación de Pearson para lo cual aprenderás a construir el gráfico de dispersión, que será de apoyo para interpretar dicho coeficiente.

Los organizadores de la Liga Mexicana del Pacífico quieren realizar un estudio sobre el impacto que ha tenido el llamado rey de los deportes (beisbol). De las siguientes opciones, ¿cuáles son variables cuantitativas?

- a) Las marcas de los patrocinadores.
- b) El número de espectadores por juego.
- c) Ganancias de la venta de suvenires.
- d) Los colores de las camisetas más populares.

¿Las variables que elegiste apoyan a los organizadores en el estudio sobre el impacto del beisbol en nuestra región?

Para estudiar la relación entre dos variables cuantitativas debes saber que, dicha relación es **la medida de asociación entre dos variables que se miden en una escala cuantitativa o numérica**. Para ello, recuerda que las variables cuantitativas son aquellas que se pueden medir en una escala numérica como el peso, la altura, el precio o el tiempo. A continuación, se describen algunos tipos de relaciones entre variables cuantitativas.

Correlación lineal

La gráfica de una función lineal es una recta que tiene por ecuación $y = mx + b$, donde $m \neq 0$ y (x, y) son las coordenadas de cualquier punto de la recta; m es la pendiente de la recta y b es la ordenada en el origen.

La pendiente de una recta se determina mediante la fórmula:

$$m = \frac{y_2 - y_1}{x_2 - x_1}, \text{ donde } x_1 \neq x_2.$$

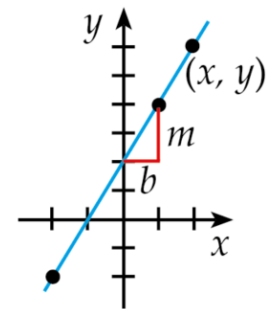


Figura 7.1. Gráfica de la ecuación $y = mx + b$.
Fuente: Elaboración propia (Desmos, 2024).

La **relación funcional** es una relación en la que los puntos de un gráfico se distribuyen de forma lineal como se muestra en la Figura 7.2. En este caso la relación entre las dos variables se puede modelar con una línea recta, lo cual hace referencia a una relación funcional cuya **correlación es perfecta**, resultado de modelos determinísticos.

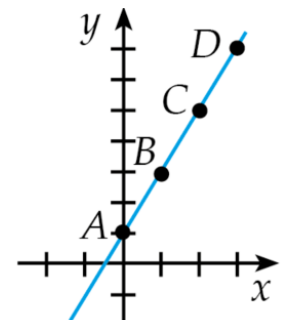


Figura 7.2. Gráfica de la ecuación $y = mx + b$.
Fuente: Elaboración propia (Desmos, 2024).

La gráfica de la función lineal $f(x) = mx + b$ representa una relación funcional ideal, es decir, utilizar ecuaciones matemáticas para pronosticar una cantidad exacta en términos de otra, algo posible para las ciencias exactas; sin embargo, para otras áreas como las ciencias sociales y humanas se recurre a la llamada **relación estadística**, que

viene siendo una medida de asociación entre dos o más variables. La relación estadística se puede determinar de forma visual mediante el *gráfico de dispersión* y de forma numérica, mediante el *coeficiente de correlación de Pearson*.

El **gráfico de dispersión**, es una herramienta gráfica que permite visualizar la relación entre dos variables. En él, cada punto representa un par de valores de las dos variables. Por ejemplo, si estamos analizando la relación entre el peso y la altura de un grupo de personas, cada punto del gráfico representaría el peso y la altura de una persona específica.

Un gráfico de dispersión se caracteriza por tener *dos ejes*, horizontal y vertical; *puntos*, donde cada punto representa un par de valores de las dos variables; y *forma*, la cual puede indicar la naturaleza de la relación entre las dos variables. Al graficar los datos de la Tabla 7.1, obtienes un gráfico de dispersión (ver Figura 7.3), de la relación de la calificación obtenida por doce estudiantes en la asignatura de Matemáticas y Mecánica.

Alumno	Calificación en Matemáticas	Calificación en Mecánica
1	3	4
2	4	5
3	5	4
4	6	5
5	6	7
6	9	8
7	7	6
8	8	5
9	10	9
10	10	10
11	8	8
12	10	8

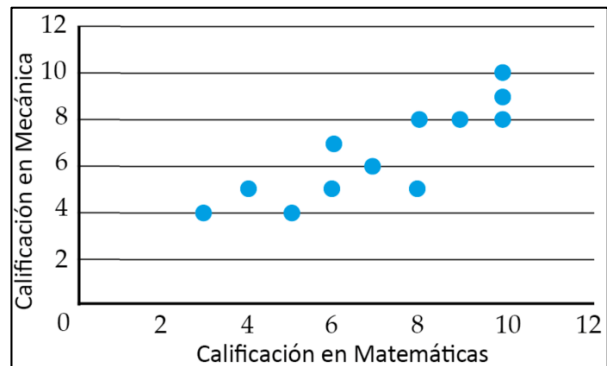


Figura 7.3. Gráfico de dispersión entre la calificación obtenida en Matemáticas y Mecánica.

Fuente: Elaboración propia (Excel, 2024).

Tabla 7.1. Datos de la calificación obtenida en Matemáticas y Mecánica.

Fuente: Elaboración propia (Word, 2023).

La relación entre variables cuantitativas es una herramienta importante para la investigación, el control de calidad y el *marketing*. En investigación, la relación entre variables cuantitativas se utiliza para explorar las asociaciones entre diferentes variables. En control de calidad, la relación entre variables cuantitativas se utiliza para monitorear la variación de dos variables en un proceso de producción. En el *marketing* se utiliza, para analizar las asociaciones entre diferentes variables en un estudio de mercado. En este sentido, podemos decir que la correlación examina la relación entre dos variables; sin embargo, no significa necesariamente que una variable sea la causa

de la otra. Por eso, solemos decir que la **correlación no implica causalidad**; posiblemente, tener buenas calificaciones en Matemáticas no significa que también las tendrás en Mecánica, pues, aunque hay cierta similitud entre ellas, no necesariamente una materia es causa de la otra.

Como observaste, el gráfico de dispersión, es especialmente útil cuando se desea examinar si existe una correlación o patrón entre dos conjuntos de datos numéricos. Puede ayudarte a identificar patrones, tendencias y correlaciones que son fundamentales en el análisis de datos y la toma de decisiones.

Para saber si existe correlación, debes observar la dirección de la relación entre las dos variables. Si los puntos en el gráfico tienden a agruparse en una dirección específica, esto indica una correlación. Si los puntos se alejan de manera uniforme, esto indica una falta de correlación. Se sugiere trazar una línea recta siguiendo el patrón de los puntos, partiendo del punto coordenado de las medias de las variables (\bar{x}, \bar{y}) , conocida como **recta de regresión**. Observa que en la Figura 7.4, la línea de tendencia es ascendente, por lo que las variables están correlacionadas de forma directa. Si es descendente como en la Figura 7.5, están correlacionadas de forma inversa. Y si no existe patrón alguno que marque una tendencia como en la Figura 7.6, entonces su correlación es nula o inexistente.

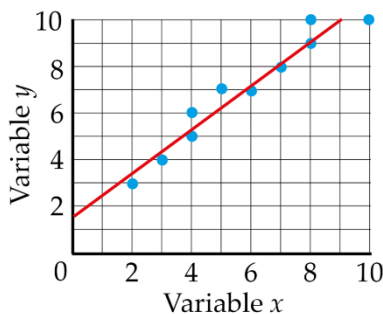


Figura 7.4 Correlación directa o positiva.
Fuente: Elaboración propia (Excel, 2024).

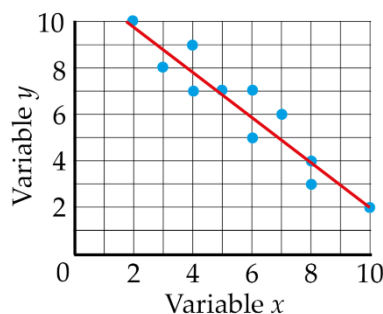


Figura 7.5 Correlación inversa o negativa.
Fuente: Elaboración propia (Excel, 2024).

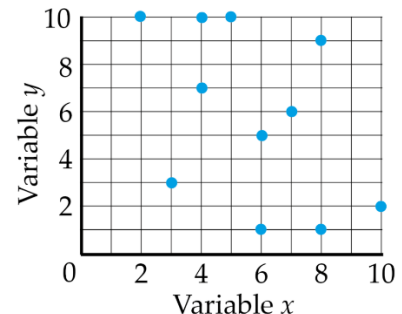


Figura 7.6 Correlación nula.
Fuente: Elaboración propia (Excel, 2024).

De la misma forma, la fuerza de la correlación se puede evaluar observando cuán cerca están los puntos del gráfico a una línea de tendencia. Si los puntos están cerca de la línea, la correlación es **fuerte** (ver Figura 7.7); si están parcialmente dispersos, la correlación es **débil** (ver Figura 7.8); si los puntos están totalmente dispersos, no hay fuerza de correlación alguna (ver Figura 7.9).

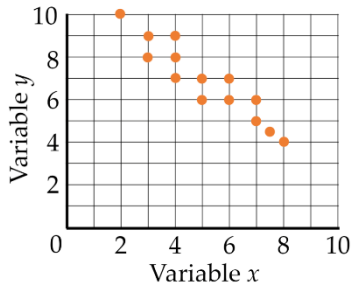


Figura 7.7 Fuerza de correlación fuerte.
Fuente: Fuente: Elaboración propia (Excel, 2024).

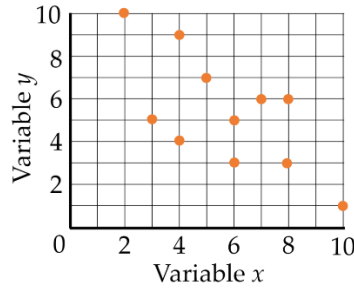


Figura 7.8 Fuerza de correlación débil.
Fuente: Elaboración propia (Excel, 2024).

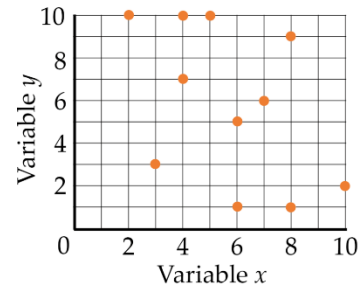


Figura 7.9 Fuerza de correlación nula.
Fuente: Elaboración propia (Excel, 2024).

Otro factor que puede influir en la interpretación de la correlación entre variables, incluye la detección de valores atípicos o puntos que se desvíen significativamente del patrón general en el gráfico. Teniendo en cuenta dicha información, se pueden establecer conclusiones basadas en la existencia de correlación, su naturaleza y fuerza de tal forma que sea útil para tomar decisiones o realizar predicciones.



QR 7.1. Video sobre la correlación entre dos variables.
Fuente: Parzibyte, 2024.

Como observaste en las figuras anteriores, el gráfico de dispersión facilita de forma intuitiva analizar la correlación entre dos variables. Otra forma es mediante el coeficiente de correlación, el cual es una medida numérica que indica la fuerza y la dirección de la relación entre dos variables que varía de -1 a 1 . Un coeficiente de correlación de 1 indica una **relación lineal positiva perfecta**, lo que haría referencia a una función lineal creciente; un coeficiente de correlación de -1 indica una **relación lineal negativa perfecta** o función lineal decreciente; y un coeficiente de correlación de 0 indica que **no hay relación** entre las dos variables.

Para medir la correlación de variables cuantitativas, puedes utilizar el **coeficiente de correlación de Pearson**, que es una medida de la fuerza y la dirección de la relación lineal entre dos variables cuantitativas y se puede calcular mediante la siguiente fórmula:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

Donde:

- r es el coeficiente de correlación de Pearson.
- x_i es i -ésimo valor de la variable x .
- y_i es i -ésimo valor de la variable y .
- \bar{x} es la media de la variable x .
- \bar{y} es la media de la variable y .
- \sum es la suma de los valores.



QR 7.2. Hoja de cálculo de Google Docs.
Fuente: Parzibyte, 2024.

El cálculo del coeficiente de correlación de Pearson, lo puedes hacer desde diferentes aplicaciones de celular como Desmos y GeoGebra; también, mediante la aplicación hojas de cálculo de Google, para analizar el nivel de correlación entre dos variables cuantitativas x y y sobre una población.

Para calcular el coeficiente de correlación de Pearson en la aplicación hojas de cálculo de Google, utiliza la función **COEF.DE.CORREL**, la cual emplea dentro de su algoritmo programado, una expresión propia de dicho coeficiente.

Procedimiento:

1. Captura la tabla de datos como se muestra en la imagen.
2. Da clic en insertar función.
3. Elige Estadísticas y busca COEF.DE.CORREL.
4. Selecciona el rango de datos de la variable y , escribe coma y selecciona el rango de datos de x .
5. Da clic en Enter.

	A	B	C
1	No. Examen	Horas de estudio	Calificación
2	1	20	6.5
3	2	16	6
4	3	34	8.5
5	4	23	7
6	5	27	9
7	6	32	9.5
8	7	18	7.5
9	8	22	8
10			
11	Coeficiente de correlación:		
12			

fx Ingresa texto o una fórmula

Figura 7.10. Datos de horas de estudio y de calificación.
Fuente: Google Docs, 2024, Disponible en: <https://docs.google.com/spreadsheets/>.

	A	B	C
1	No. Examen	Horas de estudio	Calificación
2	1	20	6.5
3	2	16	6
4	3	34	8.5
5	4	23	7
6	5	27	9
7	6	32	9.5
8	7	18	7.5
9	8	22	8
10			
11	Coeficiente de correlación:		0.827922179
12			

Figura 7.11. Cálculo del coeficiente de correlación.
Fuente: Google Docs, 2024, Disponible en: <https://docs.google.com/spreadsheets/>.

	A	B	C
1	No. Examen	Horas de estudio	Calificación
2	1	20	6.5
3	2	16	6
4	3	34	8.5
5	4	23	7
6	5	27	9
7	6	32	9.5
8	7	18	7.5
9	8	22	8
10			
11	Coeficiente de correlación:		=COEF.DE.CORREL(B2:B9,C2:C9)
12			

fx =COEF.DE.CORREL(B2:B9,C2:C9)

Figura 7.12. Función para calcular el coeficiente de correlación.
Fuente: Google Docs, 2024, Disponible en: <https://docs.google.com/spreadsheets/>.

Evaluación formativa 7.1

1. Analiza y determina si existe correlación entre las variables de los siguientes ejercicios. Argumenta basándote en la lectura anterior.

a) Altura de alumnos x y sus puntajes de exámenes y .

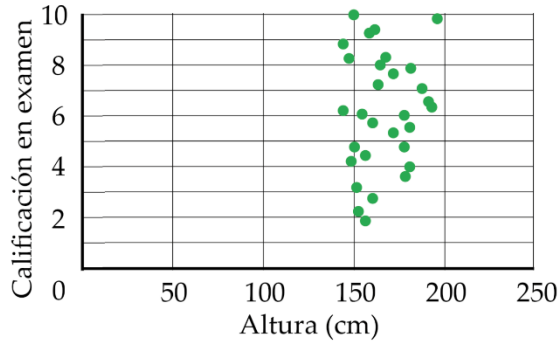


Figura 7.13. Gráfico de dispersión.

Fuente: Elaboración propia (Excel, 2024).

¿Existe correlación? _____

Argumenta:

b) Semanas trabajadas x y la paga a un ingeniero y .

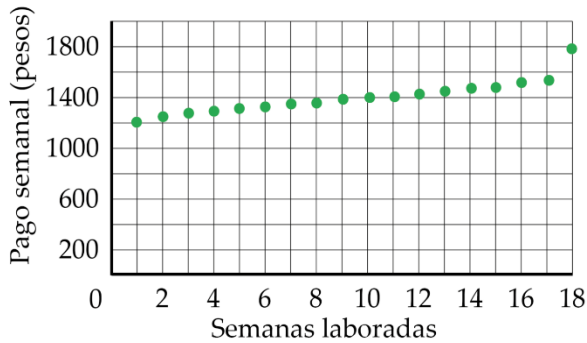


Figura 7.14. Gráfico de dispersión.

Fuente: Elaboración propia (Excel, 2024).

¿Existe correlación? _____

Argumenta:

c) Talla del calzado x y películas vistas en un año por los estudiantes y .

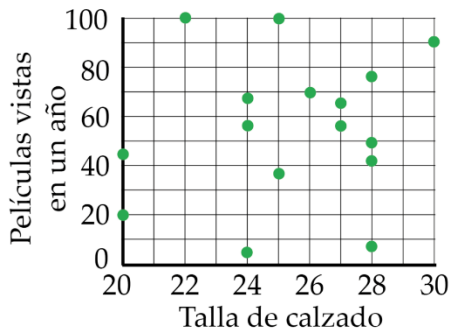


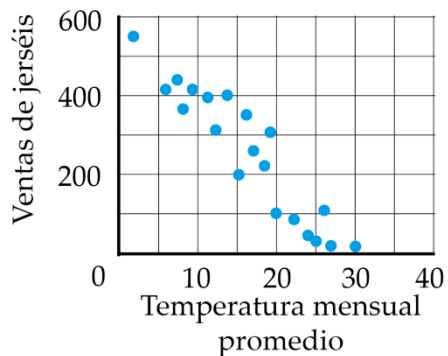
Figura 7.15. Gráfico de dispersión.

Fuente: Elaboración propia (Excel, 2024).

¿Existe correlación? _____

Argumenta:

d) La temperatura x y venta de jerséis y .



¿Existe correlación? _____

Argumenta:

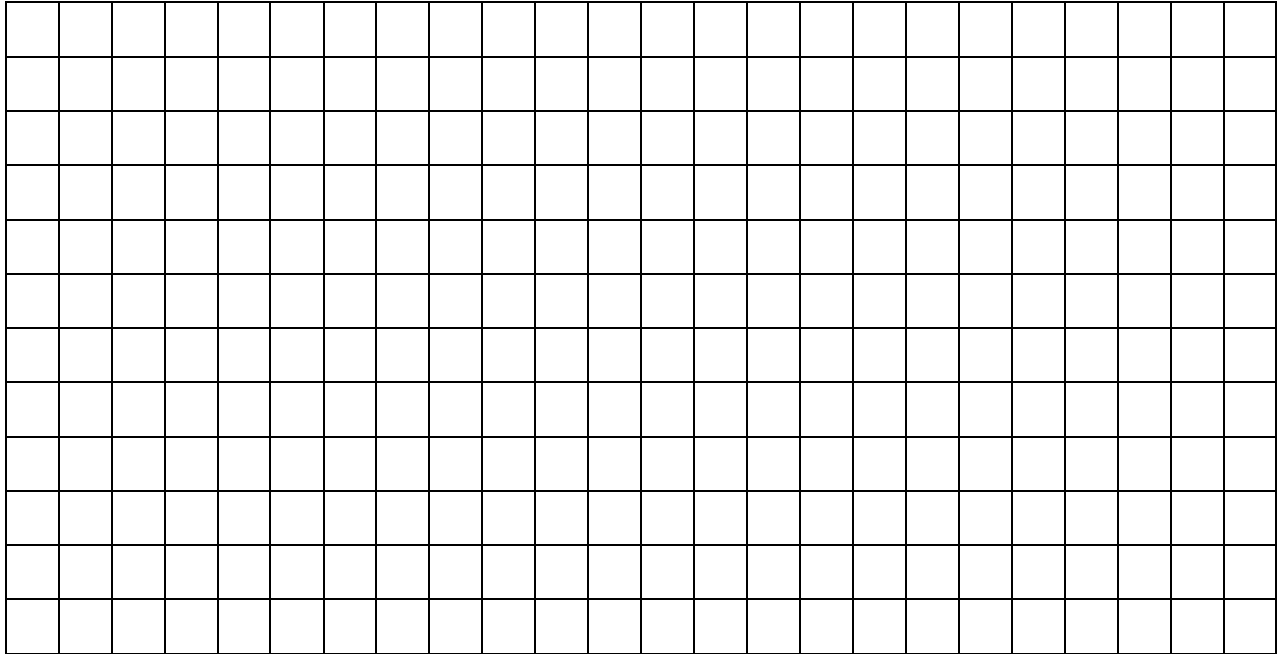
Figura 7.16. Gráfico de dispersión.
Fuente: Elaboración propia (Excel, 2024).

2. Organízate y solicita la estatura en metros y el peso en kilogramos a 30 estudiantes. Usa la variable x para representar la estatura en metros y la variable y para el peso en kg.

a) Registra los datos de cada alumno las siguientes tablas:

Alumn o	Estatura x	Peso y	Alumn o	Estatura x	Peso y
1			16		
2			17		
3			18		
4			19		
5			20		
6			21		
7			22		
8			23		
9			24		
10			25		
11			26		
12			27		
13			28		
14			29		
15			30		

b) Usa los datos de la tabla para construir el gráfico de dispersión. Observa que cada punto (x_1, y_1) en el gráfico, corresponde a los datos de un alumno.



c) Argumenta con base a la naturaleza y fuerza de la correlación, ¿qué tipo de relación hay entre las variables estatura y peso? _____

3. Usa los datos del ejercicio dos y realiza lo siguiente:

- a) Abre la aplicación **Hojas de Cálculo** en el dispositivo móvil.
- b) Captura en la tabla los datos que registraste en el ejercicio 2.
- c) Utiliza la función **COEF.DE.CORREL**, luego, selecciona los datos de las variables estatura (x) y peso (y) y calcula el coeficiente de correlación.
- d) Comenta acerca los resultados obtenidos.
- e) El coeficiente de correlación de Pearson es igual a: _____
- f) Toma captura de pantalla e inserta la a continuación.
- g) Interpreta el coeficiente de correlación de las variables analizadas.

Autoevaluación y coevaluación 7.1

Nombre: _____ Plantel: _____ Grupo: _____ Turno: _____

Autoevaluación para el aprendizaje

Selecciona en la columna, la opción que mejor refleje tu nivel de desempeño en el proceso para el aprendizaje de la progresión de aprendizaje 7. Responde con honestidad la evaluación de cada uno de los criterios que se enlistan a continuación.

Desempeño	En proceso de logro	Bueno	Sobresaliente
Propicié un clima de comunicación favorable para el estudio de la relación entre dos variables cuantitativas.			
Participé activamente con ideas para identificar si dos variables cuantitativas están correlacionadas.			
Contribuí colaborativamente en la retroalimentación de dudas sobre la relación entre dos variables cuantitativas.			
Analice dos variables cualitativas para verificar si están correlacionadas (M3-C2).			
Usé el grafico de dispersión para estudiar la correlación entre dos variables cuantitativas (M4-C2).			
Justifiqué mediante el coeficiente de correlación de Pearson si dos variables cuantitativas están correlacionadas (M4-C2).			

Coevaluación para el aprendizaje

Solicita a un compañero del equipo que marque en la columna, la opción que mejor describa tu desempeño durante el trabajo colectivo, concluida la progresión de aprendizaje 7, y que responda con honestidad la evaluación de cada uno de los criterios que se enlistan a continuación.

Desempeño	En proceso de logro	Bueno	Sobresaliente
Propició un clima de comunicación favorable para el estudio de la relación entre dos variables cuantitativas.			
Participó activamente con ideas para identificar si dos variables cuantitativas están correlacionadas.			
Contribuyó colaborativamente en la retroalimentación de dudas sobre la relación entre dos variables cuantitativas.			
Analizó dos variables cualitativas para verificar si están correlacionadas (M3-C2).			

Usó el grafico de dispersión para estudiar la correlación entre dos variables cuantitativas (M4-C2).			
Justificó mediante el coeficiente de correlación de Pearson si dos variables cuantitativas están correlacionadas (M4-C2).			

Nombre y firma de quien coevalúa

Los valores atípicos y las variables de confusión en afirmaciones estadísticas y gráficas

Progresión de aprendizaje 8

Cuestiona afirmaciones estadísticas y gráficas, considerando valores atípicos (en el caso de variables cuantitativas) y la posibilidad de que existan factores o variables de confusión.

Metas de aprendizaje	En proceso de logro	Bueno	Sobresaliente
M1-C2 Observa y obtiene información de una situación o fenómeno para establecer estrategias o formas de visualización que ayuden a entenderlo.	A		
	C		
	H		
M2-C4 Socializa con sus pares sus conjeturas, descubrimientos o procesos en la solución de un problema tanto teórico como de su entorno.	A		
	C		
	H		

Evaluación diagnóstica 8.1

1. ¿La media es la medida de tendencia central más representativa de calcular en un conjunto de datos ampliamente disperso?
 a) Verdadero b) Falso
2. ¿Cuál es el método utilizado para calcular la medida de tendencia central que consiste en sumar todos los valores y luego dividirlos por el número total de valores?
 a) La media b) La mediana c) La moda d) El rango
3. En un conjunto de datos con valores: 3, 6, 8, 10, 12, ¿cuál es la mediana?
 a) 6 b) 8 c) 10 d) 9
4. ¿Cuál es la media del siguiente conjunto de datos: 5, 7, 0, 15, 12, 10?
 a) 6 b) 8.2 c) 9.8 d) 10.5
5. ¿Cuál de las siguientes afirmaciones es verdadera sobre la mediana de un conjunto de datos?
 a) Siempre es igual a la moda
 b) Es el valor que ocurre con mayor frecuencia
 c) Divide el conjunto de datos ordenado en dos partes iguales
 d) Es igual a la suma de todos los valores dividida por el número total de valores

La **Paradoja de Simpson**. Un hospital está evaluando la efectividad de dos tratamientos para una enfermedad. Se divide a los pacientes en dos grupos de edad: menores de 50 años y mayores de 50 años. Los resultados muestran que el tratamiento A es más efectivo que el tratamiento B en ambos grupos de edad por separado. Sin

embargo, al combinar los datos de todos los pacientes, resulta que el tratamiento B es más efectivo en general.

¿Cómo es posible que el tratamiento A sea mejor en cada grupo de edad, pero el tratamiento B sea mejor en general? ¿Qué factores podrían explicar esta aparente contradicción?

Los **Datos Atípicos**. Una empresa de fabricación de electrodomésticos, está analizando los tiempos de vida útil de sus lavadoras. Se recopilan datos de 100 lavadoras, y se encuentra que la mayoría tiene una vida útil entre 8 y 12 años. Sin embargo, hay dos lavadoras que duraron solo 2 años y una que duró 20 años.

¿Cómo deberían tratarse estos datos atípicos al analizar la vida útil de las lavadoras?, ¿deberían incluirse en el análisis o excluirse? y ¿qué impacto podrían tener estos datos atípicos en las conclusiones y decisiones de la empresa?

Las variables de confusión y la paradoja de Simpson

En el campo de la estadística y la investigación, es importante comprender la relación entre las variables y cómo estas pueden influir en los resultados de un estudio. Una de las situaciones más intrigantes y contraintuitivas que puede surgir debido a la interacción de las variables es la paradoja de Simpson. Esta paradoja se produce, cuando la relación entre dos variables se invierte o desaparece al introducir una tercera variable, conocida como variable de confusión.

Las variables de confusión, son aquellas que están asociadas tanto con la variable independiente como con la variable dependiente en un estudio, pero no son el objeto principal de la investigación. Estas variables pueden enmascarar o distorsionar la verdadera relación entre las variables de interés, llevando a conclusiones erróneas si no se tienen en cuenta adecuadamente.

Un ejemplo clásico de la paradoja de Simpson, se presenta en el ámbito educativo. Una universidad está evaluando el rendimiento académico de sus estudiantes en dos facultades diferentes: la Facultad de Ciencias y la Facultad de Humanidades. La universidad quiere determinar, qué facultad tiene un mejor rendimiento académico general.

Al analizar los datos de todos los estudiantes, se observa que la Facultad de Ciencias tiene un promedio de calificaciones de 85%, mientras que la Facultad de Humanidades tiene un promedio de calificaciones de 80%. Basándose en estos datos generales, parecería que los estudiantes de la Facultad de Ciencias tienen un mejor rendimiento académico.

Sin embargo, al desglosar los datos por género, se descubre un patrón interesante. En la Facultad de Ciencias, las mujeres tienen un promedio de calificaciones de 90%, mientras que los hombres tienen un promedio de calificaciones de 80%. Por otro lado,

en la Facultad de Humanidades, los hombres tienen un promedio de calificaciones de 75%, mientras que las mujeres tienen un promedio de calificaciones de 85%.

En este caso, el género actúa como una variable de confusión. Al considerarlo, se observa que las mujeres tienen un mejor rendimiento académico en la Facultad de Humanidades en comparación con la Facultad de Ciencias, mientras que los hombres tienen un mejor rendimiento en la Facultad de Ciencias en comparación con la Facultad de Humanidades.

Si no se tiene en cuenta la variable de confusión del género, se podría llegar a la conclusión errónea de que la Facultad de Ciencias tiene un mejor rendimiento académico en general. Sin embargo, al considerar el género, se revela un patrón más complejo y se evita una interpretación simplista de los datos. Este ejemplo ilustra cómo la paradoja de Simpson puede surgir en un contexto educativo, donde una variable de confusión como el género, puede influir en la relación entre la facultad y el rendimiento académico de los estudiantes.

La presencia de variables de confusión puede tener implicaciones significativas en la interpretación de los resultados de una investigación. Si no se identifican y controlan adecuadamente, estas variables pueden llevar a asociaciones erróneas o enmascarar relaciones importantes. Por lo que, es importante que seas consciente de la posible existencia de variables de confusión y utilices técnicas estadísticas apropiadas, para abordar su efecto. Por último, comprender y tener en cuenta estas variables es necesario para evitar interpretaciones erróneas de los resultados y así, obtener conclusiones válidas en una investigación.

Ejemplo formativo 8.1

En una empresa se está investigando la relación entre el tipo de dieta (vegetariana o no vegetariana) y el nivel de colesterol alto en los empleados.

		Número de Participantes con Nivel de colesterol		Total de participantes
		Bajo	Alto	
Tipo de dieta	Vegetariana	1300	200	1500
	No vegetariana	1100	400	1500

Tabla 8.1. Relación entre el tipo de dieta y el nivel de colesterol.

Fuente: Elaboración propia (Word, 2024).

Análisis de la Tabla 8.1. Relación entre la dieta y el colesterol alto:

- Dieta vegetariana: 350 participantes con colesterol alto de 1500 (23.33%).
- Dieta no vegetariana: 250 participantes con colesterol alto de 1500 (16.67%).

Al analizar los datos, parece que la dieta vegetariana está asociada con un mayor riesgo de colesterol alto en comparación con la dieta no vegetariana.

Ahora, considerando el factor edad. Los datos se dividen en dos grupos de edad: menores de 40 años y mayores de 40 años. Los datos recopilados son los siguientes:

Grupo de edad	Tipo de dieta	Número de Participantes con Nivel de colesterol		Total de participantes
		Bajo	Alto	
Menores de 40	Vegetariana	450	50	500
	No vegetariana	850	150	1000
Mayores de 40	Vegetariana	700	300	1000
	No vegetariana	400	100	500

Tabla 8.2. Relación entre el tipo de dieta y el nivel de colesterol considerando el factor edad.

Fuente: Elaboración propia (Word, 2024).

Análisis de la Tabla 8.2. Relación entre la dieta y el colesterol alto, separados por grupo de edad:

1. Menores de 40 años:

- Dieta vegetariana: 50 participantes con colesterol alto de un total de 500 participantes (10%).
- Dieta no vegetariana: 150 participantes con colesterol alto de un total de 1000 participantes (15%).

2. Mayores de 40 años:

- Dieta vegetariana: 300 participantes con colesterol alto de un total de 1000 participantes (30%).
- Dieta no vegetariana: 100 participantes con colesterol alto de un total de 500 participantes (20%).

Al separar los datos por grupo de edad, se observa que, tanto en el grupo de menores de 40 años como en el grupo de mayores de 40 años, la dieta vegetariana está asociada con un menor riesgo de colesterol alto en comparación con la dieta no vegetariana.

En este caso, la edad actúa como una variable de confusión que influye en la relación entre el tipo de dieta y el riesgo de colesterol alto. La distribución desigual de participantes en los grupos de edad y los tipos de dieta, crea la ilusión de una relación

cuando no se considera la edad, mientras que al estratificar por edad se revela la verdadera relación subyacente.

Así, mediante la Tabla 8.2 se demuestra cómo la presencia de una variable de confusión, puede ocasionar una paradoja de Simpson y la importancia de considerar y controlar adecuadamente las variables de confusión en el análisis de datos.

En el ejemplo del hospital, donde evalúan la efectividad de dos tratamientos para una enfermedad. En el que se divide a los pacientes en dos grupos de edad: menores de 50 años y mayores de 50 años. Y los resultados muestran que el tratamiento A es más efectivo que el tratamiento B en ambos grupos de edad por separado. Sin embargo, al combinar los datos de todos los pacientes, resulta que el tratamiento B es más efectivo en general.

¿Cómo es posible que el tratamiento A sea mejor en cada grupo de edad, pero el tratamiento B sea mejor en general?

Esta situación es un ejemplo de la paradoja de Simpson, donde los resultados de subgrupos individuales (en este caso, grupos de edad), muestran una tendencia opuesta a los resultados generales cuando se combinan los datos. Es posible que el tratamiento A sea más efectivo en cada grupo de edad por separado, pero el tratamiento B sea más efectivo en general, debido a la distribución desigual de los pacientes en los grupos de edad y la diferente proporción de pacientes que reciben cada tratamiento en cada grupo.

¿Qué factores podrían explicar esta aparente contradicción?

Pueden existir variables de confusión como la edad, el sexo o el estilo de vida que influyan en la efectividad de los tratamientos, lo que ilustra una variable de confusión y a la Paradoja de Simpson.

Los datos atípicos

Es común encontrarse con observaciones que se desvían significativamente del patrón general de los datos estadísticos. Estas observaciones, conocidas como datos atípicos o *outliers*, pueden tener un impacto significativo en los resultados y las conclusiones derivadas de un conjunto de datos. Comprender qué son los datos atípicos, cómo identificarlos y cómo manejarlos adecuadamente es necesario para garantizar la integridad y la validez de cualquier conclusión.

Los datos atípicos, se definen como observaciones que difieren sustancialmente de la mayoría de los datos en un conjunto. Estas observaciones pueden ser mucho más grandes o más pequeñas que el resto de los valores. Estos pueden surgir por diversas razones, como errores de medición, errores de registro, eventos inusuales o simplemente por la presencia de observaciones genuinamente extremas.

La presencia de datos atípicos puede tener varios efectos en el análisis de datos. Por un lado, pueden distorsionar las medidas de tendencia central, como la media, y las medidas de dispersión, como la desviación estándar. Esto puede llevar a una representación inexacta de la distribución de los datos y a conclusiones erróneas.

Por otro lado, los datos atípicos también pueden proporcionar información valiosa sobre fenómenos inusuales o eventos raros que merecen una investigación más profunda. En algunos casos, pueden representar descubrimientos interesantes o indicar la presencia de subgrupos distintos dentro de los datos.

Cuando se trabaja con un conjunto de datos que contiene datos atípicos, es importante abordarlos de manera adecuada. El primer paso, es identificar los datos atípicos mediante técnicas visuales, como diagramas de caja o gráficos de dispersión. Una vez identificados, es necesario investigar el origen de estos datos atípicos y determinar si son errores u observaciones legítimas.

Si se determina que los datos atípicos son errores, como valores mal registrados o mediciones incorrectas, lo más apropiado es corregirlos o eliminarlos del conjunto de datos. Sin embargo, si los datos atípicos representan observaciones válidas, se deben considerar diferentes enfoques para manejarlos. Una opción es utilizar medidas robustas, como la mediana en lugar de la media, que son menos sensibles a la presencia de datos atípicos. Otra alternativa es realizar análisis, con y sin los datos atípicos para evaluar su impacto en los resultados.

Ejemplo formativo 8.2

Una empresa de comercio electrónico, está analizando los ingresos diarios generados por las ventas en su plataforma durante el último mes. Los datos recopilados son los siguientes:

Día	Ingreso	Día	Ingreso	Día	Ingreso
1	\$1,500	11	\$1,650	21	\$1,900
2	\$1,800	12	\$1,800	22	\$1,300
3	\$1,200	13	\$1,100	23	\$1,550
4	\$1,600	14	\$1,700	24	\$1,700
5	\$1,400	15	\$10,000	25	\$1,450
6	\$1,700	16	\$1,600	26	\$1,650
7	\$1,900	17	\$1,750	27	\$1,800
8	\$1,300	18	\$1,200	28	\$1,200
9	\$1,550	19	\$1,500	29	\$1,600
10	\$1,450	20	\$1,400	30	\$1,500

Tabla 8.3. Ingresos diarios generados por las ventas mediante plataforma durante el último mes.

Fuente: Elaboración propia (Word, 2024).

En este conjunto de datos de la Tabla 8.3, se puede observar que el día 15 presenta un valor atípico de \$10,000 pesos, que es significativamente mayor que los ingresos

diarios típicos. Este dato atípico podría deberse a un evento especial de ventas o una promoción extraordinaria que tuvo lugar ese día. En un gráfico de puntos como el de la Figura 8.1 se observa el valor atípico de un ingreso de \$10,000 pesos.

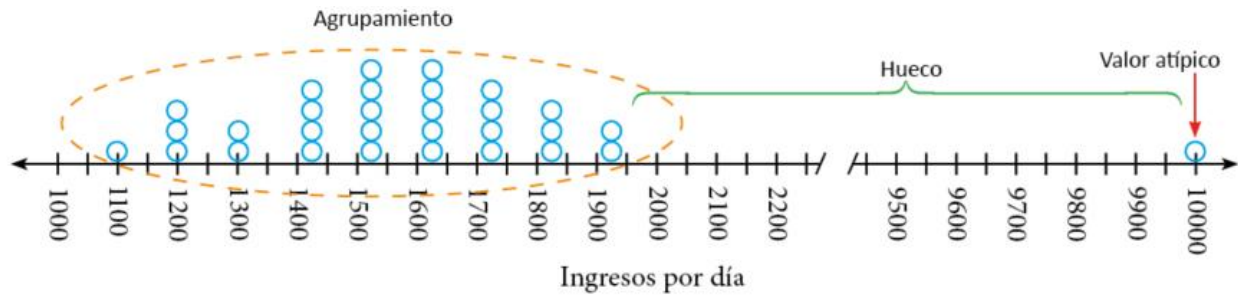


Figura 8.1. Gráfico de puntos con un valor atípico.
Fuente: Elaboración propia (Word, 2024).

Otra forma de observar un dato atípico es mediante el gráfico de cajas.



Figura 8.2. Gráfico de cajas con un valor atípico.
Fuente: Elaboración propia (Excel, 2024).

Recomendación sobre qué medida de tendencia calcular:

En presencia de datos atípicos como en este caso, se recomienda utilizar la mediana como medida de tendencia central en lugar de la media. La mediana es menos sensible a los valores extremos y proporciona una representación más robusta de la tendencia central de los datos.

La mediana se calcula ordenando los datos de menor a mayor y seleccionando el valor central. En este caso, con 30 observaciones, la mediana sería el promedio de los valores en las posiciones 15 y 16 después de ordenar los datos.

\$1 100, \$1 200, \$1 200, \$1 200, \$1 300, \$1 300, \$1 400, \$1 400, \$1 450, \$1 450, \$1 500, \$1 500, \$1 500, \$1 550, \$1 550, \$1 600, \$1 600, \$1 600, \$1 650, \$1 650, \$1 700, \$1 700, \$1 700, \$1 750, \$1 800, \$1800, \$1 800, \$1 900, \$1 900, \$10 000.

$$Me = \frac{1550 + 1600}{2} = 1575$$

Con la mediana, la venta media mensual es de \$1,575 pesos y con la media, $\bar{x} = \$1,825$, la venta promedio mensual es de \$1,825 pesos. Observa que el valor atípico tiende a inflar el valor de la venta promedio mensual.

Al utilizar la mediana, se obtiene una medida de tendencia central que es más resistente a la influencia del dato atípico del día 15. Esto proporciona una visión más precisa de los ingresos diarios típicos de la empresa, sin que el valor extremo distorsione la representación de la tendencia central.

Retomado la situación de la empresa de fabricación de electrodomésticos que está analizando los tiempos de vida útil de sus lavadoras. Se recopilan datos de 100 lavadoras, y se encuentra que la mayoría tiene una vida útil entre 8 y 12 años. Sin embargo, hay dos lavadoras que duraron solo 2 años y una que duró 20 años.

- ¿Cómo deberían tratarse estos datos atípicos al analizar la vida útil de las lavadoras?

Los datos atípicos deben ser identificados y analizados cuidadosamente para comprender por qué difieren tanto de la mayoría de los datos. Es importante investigar si estos casos son resultado de errores en la recolección de datos, errores de medición, o si representan verdaderas anomalías en el comportamiento de las lavadoras.

- ¿Deberían incluirse en el análisis o excluirse?

La inclusión o exclusión de los datos atípicos dependerá del enfoque del análisis y de los objetivos de la empresa. Si el objetivo es obtener una visión precisa de la vida útil típica de las lavadoras en condiciones normales, puede ser razonable excluir los datos atípicos. Sin embargo, si la empresa está interesada en comprender y abordar problemas potenciales de calidad o durabilidad, sería más apropiado incluir estos casos en el análisis.

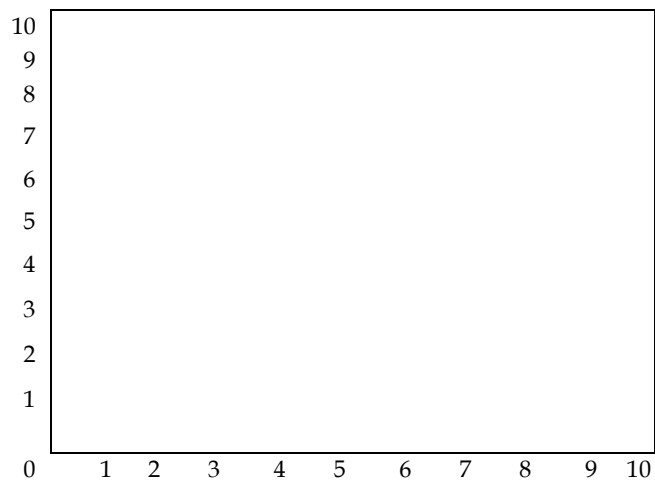
- ¿Qué impacto podrían tener estos datos atípicos en las conclusiones y decisiones de la empresa?

Los datos atípicos pueden tener un impacto significativo en las conclusiones y decisiones de la empresa. Si se excluyen sin una comprensión adecuada de su causa subyacente, podrían sesgar las conclusiones hacia una visión demasiado optimista o pesimista de la vida útil de las lavadoras. Por otro lado, si se incluyen y se analizan adecuadamente, podrían proporcionar información sobre posibles problemas de calidad o durabilidad que la empresa necesita abordar. En cualquier caso, es importante considerar el contexto específico y los objetivos del análisis al interpretar el impacto de los datos atípicos.

Actividad evaluativa 8.1

1. Un estudio examinó la relación entre el consumo de café y el riesgo de enfermedad cardíaca en hombres y mujeres. Los resultados mostraron que:
 - En los hombres: los bebedores de café tenían un menor riesgo de enfermedad cardíaca que los no bebedores.
 - En las mujeres: las bebedoras de café tenían un mayor riesgo de enfermedad cardíaca que las no bebedoras.
 - Al combinar los datos de hombres y mujeres, no se encontró relación entre el consumo de café y el riesgo de enfermedad cardíaca.
 - a) ¿Qué relación se observa entre el consumo de café y el riesgo de enfermedad cardíaca cuando se consideran hombres y mujeres por separado?
 - b) ¿Qué sucede con esta relación cuando se combinan los datos de ambos sexos?
 - c) ¿Cómo explicarías estos resultados aparentemente contradictorios?
 - d) ¿Qué precauciones deberían tomar los investigadores al interpretar estos datos?
2. Considera el siguiente escenario: En un hospital, se compararon las tasas de éxito de dos tratamientos, A y B, para una enfermedad. Los resultados fueron:
 - Pacientes menores de 40 años: Tratamiento A (80% de éxito), Tratamiento B (90% de éxito)
 - Pacientes de 40 años o más: Tratamiento A (70% de éxito), Tratamiento B (60% de éxito)
 - En general, el Tratamiento A tuvo una tasa de éxito del 75%, mientras que el Tratamiento B tuvo un 70% de éxito.
 - a) Elabora tablas o gráficas que te ayuden a visualizar la información.
 - b) ¿Qué tratamiento parece ser más efectivo cuando se consideran los grupos de edad por separado?
 - c) ¿Qué tratamiento parece ser más efectivo cuando se combinan los datos de ambos grupos de edad?
 - d) ¿Estamos ante un caso de la paradoja de Simpson? Explica tu respuesta.
 - e) ¿Qué implicaciones tiene este escenario en la toma de decisiones médicas?
3. a) Realiza un gráfico de puntos con los datos de la siguiente tabla. En el eje horizontal coloca el número de veces y la frecuencia en la parte vertical.

Número de veces	Frecuencia
0	1
1	5
2	10
3	8
4	7
5	9
9	1



- b) ¿Cuál consideras que sea el hueco? ¿Por qué?
- c) ¿Cuál consideras que es el agrupamiento? ¿Por qué?
- d) ¿Cuál es el valor atípico? ¿Por qué?
- e) Calcula la media y la mediana de los datos; luego, indaga en la inteligencia artificial, ¿cómo afecta un valor atípico a la media y a la mediana de los datos?

4. Observa las siguientes gráficas y sus medidas de tendencia central.

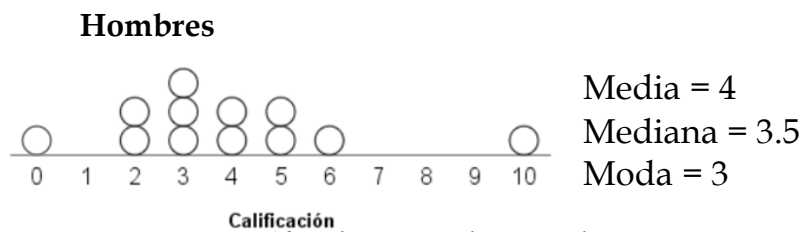
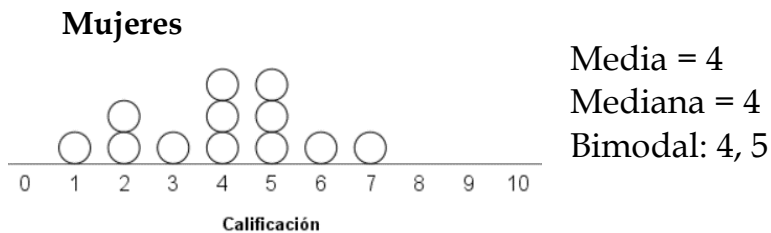


Figura 10.3. Calificación en matemáticas por género.
Fuente: Elaboración propia (Excel, 2024).

- a) ¿Consideras que las mujeres y los hombres tienen el mismo promedio de calificación?
- b) ¿Qué medida de tendencia central consideras que es más representativa de la calificación promedio de los hombres? Justifica tu respuesta.

Autoevaluación y coevaluación 8.1

Nombre: _____ Plantel: _____ Grupo: _____ Turno: _____

Autoevaluación para el aprendizaje

Selecciona en la columna la opción que mejor refleje tu nivel de desempeño en el proceso para el aprendizaje de la progresión de aprendizaje 8. Responde con honestidad a la evaluación de cada uno de los criterios que se enlistan a continuación.

Desempeño	En proceso de logro	Bueno	Sobresaliente
Propicié un clima de comunicación favorable para el estudio de datos atípicos y variables de confusión.			
Participé activamente con ideas para identificar datos atípicos y variables de confusión.			
Contribuí colaborativamente en la retroalimentación de dudas sobre datos atípicos y variables de confusión.			
Hice gráficos de puntos o diagramas de cajas para identificar datos atípicos (M1-C2).			
Comuniqué las posibles causas de la paradoja de Simpson que ayuden a identificar variables de confusión (M2-C4).			
Hice gráficos y tablas para evitar el efecto de la paradoja de Simpson (M1-C2).			

Coevaluación para el aprendizaje

Solicita a un compañero del equipo que marque en la columna la opción que mejor describa tu desempeño durante el trabajo colectivo, concluida la progresión de aprendizaje 8, y que responda con honestidad la evaluación de cada uno de los criterios que se enlistan a continuación.

Desempeño	En proceso de logro	Bueno	Sobresaliente
Propició un clima de comunicación favorable para el estudio de datos atípicos y variables de confusión.			
Participó activamente con ideas para identificar datos atípicos y variables de confusión.			
Contribuyó colaborativamente en la retroalimentación de dudas sobre datos atípicos y variables de confusión.			
Hizo gráficos de puntos o diagramas de cajas para identificar datos atípicos (M1-C2).			
Comunicó las posibles causas de la paradoja de Simpson que ayuden a identificar variables de confusión (M2-C4).			
Hizo gráficos y tablas para evitar el efecto de la paradoja de Simpson (M1-C2).			

Nombre y firma de quien coevalúa

Estudio de la población a partir de una muestra

Progresión de aprendizaje 9

Identifica, ante la imposibilidad de estudiar la totalidad de una población, la opción de extraer información de ésta a través del empleo de técnicas de muestreo, en particular, valora la importancia de la aleatoriedad al momento de tomar una muestra.

Metas de aprendizaje	En proceso de logro	Bueno	Sobresaliente
M1-C2. Observa y obtiene información de una situación o fenómeno para establecer estrategias o formas de visualización que ayuden a entenderlo.	A		
	C		
	H		
M2-C3 Construye un modelo matemático, identificando las variables de interés, con la finalidad de explicar una situación o fenómeno y/o resolver un problema tanto teórico como de su entorno.	A		
	C		
	H		
M4-C3 Construye y plantea posibles soluciones a problemas de Áreas de Conocimiento, Recursos Sociocognitivos, Recursos Socioemocionales y de su entorno, empleando técnicas y lenguaje matemático.	A		
	C		
	H		

Evaluación diagnóstica 9.1

1. ¿Cuál de las siguientes afirmaciones describe mejor el muestreo probabilístico?
 - a) Todos los elementos de la población tienen la misma posibilidad de ser seleccionados.
 - b) Los elementos se seleccionan de acuerdo con el juicio del investigador.
 - c) Se eligen los elementos más fáciles de acceder.
 - d) La muestra se elige en función de cuotas predefinidas.
2. ¿Cuál es una ventaja principal del muestreo probabilístico sobre el no probabilístico?
 - a) Es más rápido y económico.
 - b) Permite calcular estimaciones precisas y cuantificar la incertidumbre.
 - c) Requiere menos planificación.
 - d) Se basa en el juicio subjetivo del investigador.
3. ¿Cuál es la característica principal del muestreo aleatorio simple?
 - a) Todos los elementos tienen una probabilidad igual de ser seleccionados.
 - b) Los elementos se eligen basándose en la accesibilidad.
 - c) Se eligen elementos específicos para cumplir con cuotas.
 - d) La selección se basa en el juicio del investigador.
4. Si un investigador selecciona a los participantes basándose en su propio criterio para confirmar su hipótesis, ¿qué tipo de muestreo está utilizando?
 - a) Muestreo aleatorio simple
 - b) Muestreo por juicio
 - c) Muestreo sistemático

d) Muestreo por conglomerados

En la mayoría de las investigaciones, pocas veces, es posible estudiar a toda la población, por lo que se selecciona una muestra y desde luego, se pretende que esta muestra sea representativa de la población. Las muestras se utilizan para ahorrar tiempo y recursos. Las técnicas de muestreo son un conjunto de métodos estadísticos que permiten seleccionar una muestra representativa de una población.

La **muestra** es un subgrupo de la población de interés sobre el cual se recolectarán datos. Digamos que es un subconjunto de elementos que pertenecen a ese conjunto definido en sus características al que llamamos **población** y deberá ser representativo de dicha población.



Figura 9.1. Población y muestra.
Fuente: Fotografía, Asia Cecilia Carrasco Valenzuela e Iliana Tirado Olivas (Android, 2024).

Técnicas de muestreo

Existen dos tipos de técnicas de muestreo: el *muestreo no probabilístico* y el *muestreo probabilístico*.

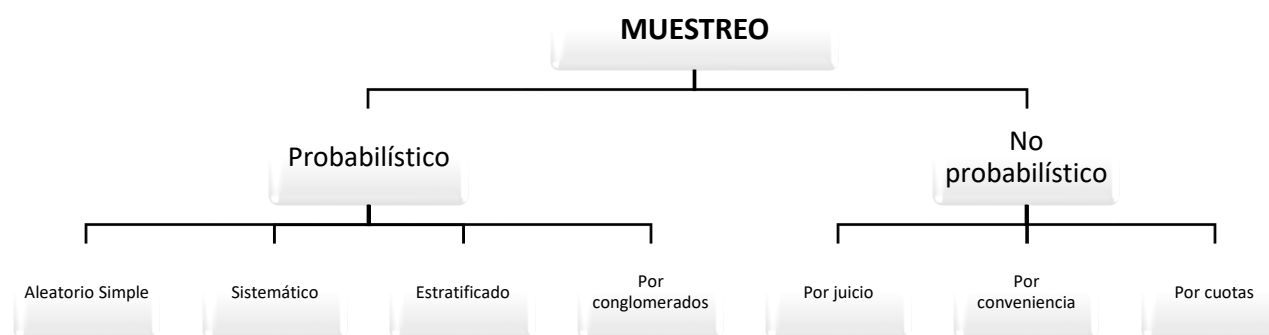


Figura 9.2. Tipos de muestreo.
Fuente: Elaboración propia (Word, 2024).

En las **muestras probabilísticas** todos los elementos de la población tienen la misma posibilidad de ser seleccionados y se obtienen definiendo las características de la población y el tamaño de la muestra. Este tipo de muestreo es considerado el más riguroso y confiable, ya que permite calcular estimaciones precisas y cuantificar la incertidumbre asociada con ellas.

¿Cómo se obtiene el número premiado en un sorteo de lotería? En las loterías tradicionales, a partir de las esferas con un dígito que se extraen (después de revolverlas mecánicamente en una tómbola) hasta formar el número, de manera que todos los números tienen la misma probabilidad de ser elegidos.

En las **muestras no probabilísticas** la elección de los elementos **no depende de la probabilidad**, sino, mediante un juicio subjetivo del investigador. Este tipo de muestreo puede ser útil en situaciones donde no es posible realizar un muestreo probabilístico debido a limitaciones de dinero y tiempo.

Ejemplo formativo 9.1

Determina si cada uno de los siguientes métodos para seleccionar a los participantes de una encuesta, corresponde a un muestreo probabilístico (P) o no probabilístico (NP).

Una empresa encuestadora requiere conocer la opinión de los ciudadanos sobre quién le gustaría que fuera el próximo candidato a gobernador. Para ello, consultó el directorio telefónico y decidió seleccionar a algunos de ellos para preguntarles si desean participar en dicha encuesta telefónica.

1. (NP) La empresa selecciona a los primeros 500 ciudadanos del directorio telefónico.
2. (P) Elige al azar a 500 clientes del directorio telefónico.
3. (NP) La empresa selecciona a los 500 ciudadanos que más han gastado en telefonía en los últimos años.
4. (P) La empresa selecciona a 500 ciudadanos al zar, de los cuales, 50 corresponden a cada uno de los códigos postales de la ciudad.

Muestreo probabilístico

El **muestreo probabilístico** es esencial en los diseños de investigación donde se pretende hacer estimaciones de variables en la población, pues su objetivo es obtener una muestra representativa que te permita hacer inferencias sobre la población a partir de los resultados obtenidos.

Para obtener una muestra probabilística son necesarios dos procedimientos:

- 1) Calcular un tamaño de **muestra** (n) que sea representativo de la población.
- 2) Seleccionar los elementos muestrales de manera que al inicio todos tengan la misma posibilidad de ser elegidos.

El cálculo de un tamaño de **muestra** (n) puede hacerse mediante la aplicación estadística “EST calcula tu muestra” en tu celular. También, usando las fórmulas clásicas que se han desarrollado, pero es más tardado y el resultado es el mismo o muy similar al que proporciona la aplicación.



Figura 9.3. Aplicación para calcular tamaño de una muestra.
Fuente: EST calcula tu muestra (Android, 2024).

Luego, debes aplicar un método de muestreo probabilístico como el aleatorio simple, estratificado, sistemático o por conglomerados, para seleccionar los elementos muestrales de manera que al inicio todos tengan la misma posibilidad de ser elegidos.

Los tipos de muestreo probabilísticos son:

Aleatorio simple. Para poder realizar este tipo de muestreo, todos los individuos de la población deben estar numerados en un listado, a partir de un listado de números aleatorios. Si no dispones del listado de individuos, no se podrá utilizar esta técnica de muestreo.



Figura 9.4. Muestreo aleatorio simple.

Fuente: Fotografía. Asia Cecilia Carrasco Valenzuela e Iliana Tirado Olivas (Android, 2024).

Ejemplo: para seleccionar una muestra de 5 estudiantes de una población de 50, asigna un número a cada individuo y selecciona 5 números al azar. Puedes utilizar la aplicación "Generar números aleatorios" en tu celular.

Estratificado. En este tipo de muestreo se divide a la población en subgrupos o estratos que tienen alguna característica común. La selección de sujetos dentro de cada estrato se hace aleatoriamente. La estratificación se suele hacer en función de diferentes variables o características: género, edad, situación laboral, etc. La elección de los elementos en cada estrato se realiza mediante algún método de muestreo aleatorio simple.

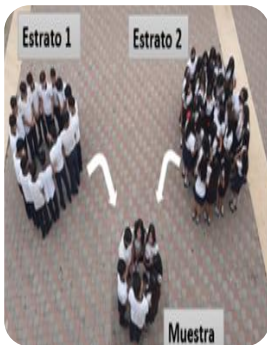


Figura 9.5. Muestreo estratificado.

Fuente: Fotografía. Asia Cecilia Carrasco Valenzuela e Iliana Tirado Olivas (Android, 2024).

Ejemplo: en una población de estudiantes de preparatoria, considera los estratos, hombres y mujeres, y selecciona una muestra de cada uno.

Sistemático. Es similar al aleatorio simple: los sujetos deben estar identificados. Éstos no se eligen a partir de un listado de números aleatorios, sino que, se hace sistemáticamente eligiendo a uno de cada cierto número de sujetos. En el muestreo sistemático, primero se selecciona un elemento aleatoriamente, y el resto de los elementos de la muestra se seleccionan utilizando un intervalo fijo, que indica la separación entre los elementos sucesivos de la muestra.



Figura 9.6. Muestreo sistemático.

Fuente: Fotografía. Asia Cecilia Carrasco Valenzuela e Iliana Tirado Olivas (Android, 2024).

Ejemplo: para seleccionar una muestra de 4 estudiantes de una población de 22, elige una persona al azar y posteriormente una de cada 5 en un listado con todos los estudiantes.

Conglomerados. En este tipo de muestreo, cada conglomerado se pretende que sea una parte de toda la población, es decir, grupos heterogéneos. Posteriormente, se seleccionan algunos conglomerados al azar que conforman la muestra y se analiza cada individuo perteneciente a los conglomerados seleccionados. En este tipo de muestreo hay una selección en varias etapas, todas con procedimientos probabilísticos. Ejemplo: para seleccionar una muestra de estudiantes de una escuela, es posible dividirla en grados, después, selecciona algunos de ellos al azar.

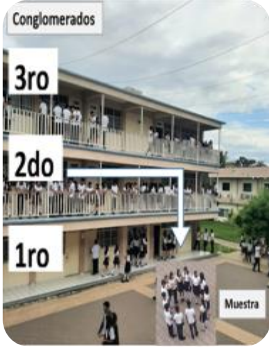


Figura 9.7. Muestreo por conglomerados.

Fuente: Fotografía. Asia Cecilia Carrasco Valenzuela e Iliana Tirado Olivas (Android, 2024).

Actividad formativa 9.3

Determina el tipo de muestreo aplicado para cada situación.

1. Se realizó un juego en el que se elige a la ganadora del juego y a partir de ahí, se empezó a numerar cada cinco lugares, estas fueron las muestras finales. **Muestreo aleatorio sistemático.**
2. Asignar un número a cada alumno de la lista y luego utilizar un generador de números aleatorios, para seleccionar 20 estudiantes. **Muestreo aleatorio simple.**
3. Se quiere estimar el número de butacas en mal estado de la preparatoria, se sabe que hay 30 aulas, cada una con aproximadamente 50 butacas, por lo que, se eligen al azar 5 aulas y se procede a revisar cada una de las butacas de dichas aulas. **Muestreo aleatorio por conglomerados.**
4. Se dividió al grupo de primer semestre en subconjuntos según edades, de 15 y 16 años, y luego, se seleccionó una muestra de cada uno utilizando un muestreo aleatorio simple. **Muestreo aleatorio estratificado.**

Muestreo no probabilístico

En el muestreo no probabilístico, la elección de los elementos no depende de la probabilidad, sino de causas relacionadas con las características de la investigación o de quien selecciona la muestra. Aclaramos que el procedimiento no es mecánico, ni se usan fórmulas de probabilidad, tan solo, depende del proceso de toma de decisiones de una investigación o de un grupo de investigadores, y desde luego, las muestras seleccionadas obedecen a otros criterios de investigación.

Este tipo de muestreo es importante en la investigación porque permite seleccionar una muestra de una población de manera rápida y económica. Aunque esta técnica no garantiza que la muestra sea representativa de la población, es útil en situaciones en las que es difícil o costoso obtener una muestra probabilística.

El muestreo no probabilístico es ampliamente utilizado en estudios exploratorios y en investigaciones cualitativas, así como, en estudios de mercado y en la investigación social y política.

Los tipos de muestreo no probabilísticos más utilizados son:

Muestreo	En qué consiste	Ejemplo
Por conveniencia	Consiste en elegir la muestra, de acuerdo con ciertas características que le sean convenientes al investigador, así como, a los encuestados en función de su disponibilidad y facilidad de acceso.	El investigador pregunta a las personas si están dispuestas a participar en el estudio.
Por juicio	El investigador emplea su propio juicio para elegir los individuos que debe incluir en la muestra. El investigador elige de manera deliberada una muestra para confirmar su opinión.	Si quieres estimar cuánto gasta una persona que va de compras a un centro comercial extraes una muestra entre los compradores que parecen haber gastado cierta cantidad.
Por cuotas	Es similar al muestreo aleatorio estratificado, pero la diferencia radica en que la muestra se elige de forma subjetiva.	El investigador elige los elementos que van a pertenecer a la muestra tomando en cuenta sus propios términos.

Actividad formativa 9.4

Determina qué tipo de muestreo no probabilístico se está utilizando.

Por conveniencia

Por juicio

Por cuotas

1. Deseas conocer la opinión sobre una crema aromática corporal. Para la selección optas por abordar a mujeres en el supermercado, y les preguntas si están dispuestas a participar en el estudio. **Por conveniencia**
2. Si queremos estimar cuánto gastan las personas en el cine, una opción es encuestar a las familias que asisten al cine ese día para extraer una muestra. **Por juicio**
3. Se está estudiando el machismo en el bachillerato, se decidió dividir la población total entre hombres y mujeres, luego, elige un número determinado de participantes de cada grupo. **Por cuotas**

La encuesta

La encuesta es una herramienta fundamental en la investigación. Permite recopilar datos sobre actitudes, percepciones o características de los participantes a través de preguntas cerradas; así como, información sobre las experiencias, opiniones o características de los sujetos observados, mediante preguntas abiertas. Para diseñar una encuesta es importante que definas claramente, el objetivo de la investigación, elabores preguntas claras y concisas, elijas la técnica de muestreo y selecciones el tipo de encuesta más adecuado (como encuestas presenciales, en línea o telefónicas).

La **definición del objetivo** es fundamental para comprender qué aspectos se buscan indagar dentro de la población. Este es un punto clave para el diseño preciso de las preguntas en la encuesta y para la elección acertada de la técnica de muestreo. Un objetivo claro, proporciona la dirección necesaria para recabar datos relevantes y significativos, asegurando que la encuesta sea efectiva y capaz de ofrecer conclusiones sólidas.

Es imprescindible que las **preguntas** en la encuesta sean claras, concisas y pertinentes con respecto al objetivo planteado. Se sugiere la inclusión, tanto preguntas abiertas como cerradas para obtener una visión integral de la información requerida. Asimismo, es esencial evitar sesgos y confusiones en las preguntas, ya que podrían comprometer la validez de los resultados obtenidos.

La selección de la **técnica de muestreo** apropiada está ligada al propósito de la encuesta y las particularidades de la población estudiada. En busca de representatividad se sugiere el empleo de técnicas de muestreo probabilístico, como el muestreo aleatorio simple o el estratificado, que contribuyen a asegurar una muestra representativa y confiable de la población.

Una vez definida la técnica de muestreo, se procede a la selección de participantes y la **aplicación de la encuesta**. La obtención de resultados confiables depende en gran medida de la rigurosidad aplicada en la ejecución precisa de la técnica de muestreo seleccionada.

Por último, los resultados son sometidos a un **análisis** exhaustivo utilizando diversas técnicas con el objetivo de extraer información valiosa sobre la población estudiada. Buscando extraer información significativa y relevante que pueda ser útil para la toma de decisiones o la formulación de recomendaciones.

A manera de ejemplo, si el objetivo es evaluar los hábitos alimenticios de los estudiantes de bachillerato, para identificar los tipos de alimentos más consumidos durante el horario escolar, sus preferencias alimenticias, el acceso a opciones saludables en la escuela y la percepción sobre la influencia de la alimentación en su rendimiento académico. Para ello, puedes aplicar la siguiente encuesta.

Encuesta sobre los alimentos que consumen los estudiantes de bachillerato.

Datos demográficos. Edad: _____ Género: _____ Grado escolar: _____

Circula, en cada caso, la letra con la opción que consideres más adecuada.

Hábitos alimenticios:

1. ¿Con qué frecuencia consumes alimentos durante el horario escolar?
a) Nunca b) Ocasionalmente c) A menudo d) Siempre
2. ¿Sueles desayunar antes de ir a la escuela?
a) Sí, siempre a) A veces c) No, nunca
3. ¿Prefieres llevar comida de casa o comprar en la cafetería de la escuela?
a) Llevo comida de casa siempre
b) A veces llevo comida de casa y a veces compro en la cafetería
c) Siempre compro en la cafetería
d) Nunca compro en la cafetería

Tipo de alimentos consumidos:

4. ¿Qué tipo de alimentos consumes con mayor frecuencia durante el horario escolar?
(Selecciona todas las que correspondan)
a) Frutas
b) Verduras
c) Bocadillos (galletas, panes, etc.)
d) Sándwiches o bocadillos preparados en casa
e) Comida rápida (hamburguesas, pizza, tortas, etc.)
f) Comida caliente (arroz, pasta, consomé, cazuela, etc.)
g) Otros (especifica) _____
5. ¿Tienes acceso a opciones de alimentos saludables en la escuela?
a) Sí, hay opciones saludables disponibles
b) No, las opciones saludables son limitadas
c) No estoy seguro

Preferencias y opiniones:

6. ¿Qué mejorarías o cambiarías en las opciones de alimentos ofrecidas en la escuela?
a) Mayor variedad de opciones saludables
b) Mejor calidad de los alimentos
c) Precios más accesibles
d) Otros (especifica) _____
7. ¿Crees que tu alimentación durante el horario escolar afecta tu rendimiento académico?
a) Sí
b) No
c) No estoy seguro
8. ¿Qué cambiarías o mejorarías sobre el tipo de alimentos que venden en la escuela?

Evaluación formativa 9.1

Implementación de una encuesta.

Tema: Los adolescentes utilizan sus teléfonos móviles para realizar diversas actividades. En equipo investiga qué actividades realizan los estudiantes con sus celulares.

1. **Diseño de preguntas.** Elabora al menos cinco preguntas cerradas y una pregunta abierta que te permitan obtener información detallada sobre el uso de teléfonos móviles entre los estudiantes. Asegúrate de que las preguntas cerradas incluyan opciones de respuestas predefinidas.
2. **Muestreo y aplicación.** Realiza un muestreo probabilístico para seleccionar una muestra representativa de estudiantes. Aplica la encuesta a al menos 20 estudiantes seleccionados de manera aleatoria.
3. **Análisis de resultados.** Analiza los resultados obtenidos de la encuesta y elabora un informe escrito. Incluye gráficos o tablas que ilustren los resultados obtenidos y cita las fuentes de información consultadas.
4. Contesta las siguientes preguntas:
 - a) ¿Cuál es la población? _____

 - b) ¿Cuál es la muestra en este estudio? _____

 - c) ¿Qué tipo de muestreo probabilístico utilizaste? _____

Autoevaluación y coevaluación 9.1

Nombre: _____ Plantel: _____ Grupo: _____ Turno: _____

Autoevaluación para el aprendizaje

Selecciona en la columna la opción que mejor refleje tu nivel de desempeño en el proceso para el aprendizaje de la progresión de aprendizaje 9. Responde con honestidad a la evaluación de cada uno de los criterios que se enlistan a continuación.

Desempeño	En proceso de logro	Bueno	Sobresaliente
Propicié un clima de comunicación favorable para el estudio del muestreo probabilístico y no probabilístico.			
Participé activamente con ideas para identificar el tipo de muestro sen el tipo de estudio a realizar.			
Contribuí colaborativamente en la retroalimentación de dudas sobre los tipos de muestreos y sus características.			
Elegí el tipo de muestreo según el tipo de estudio que se realiza: diseño de experimentos o estudio observacionales confusión (M1-C2).			

Determiné los parámetros que describen a una población en el cálculo de una muestra (M2-C3).			
--	--	--	--

Coevaluación para el aprendizaje

Solicita a un compañero del equipo que marque en la columna la opción que mejor describa tu desempeño durante el trabajo colectivo, concluida la progresión de aprendizaje 9, y que responda con honestidad la evaluación de cada uno de los criterios que se enlistan a continuación.

Desempeño	En proceso de logro	Bueno	Sobresaliente
Propició un clima de comunicación favorable para el estudio del muestreo probabilístico y no probabilístico.			
Participó activamente con ideas para identificar el tipo de muestro sen el tipo de estudio a realizar.			
Contribuyó colaborativamente en la retroalimentación de dudas sobre los tipos de muestreos y sus características.			
Eligió el tipo de muestreo según el tipo de estudio que se realiza: diseño de experimentos o estudio observacionales (M1-C2).			
Determinó los parámetros que describen a una población en el cálculo de una muestra (M2-C3).			

Nombre y firma de quien coevalúa

Los estudios observacionales y el diseño de experimentos

Progresión de aprendizaje 10

Valora las ventajas y limitaciones de los estudios observacionales y los compara con el diseño de experimentos, a través de la revisión de algunos ejemplos tomados de diversas fuentes.

Meta de aprendizaje	En proceso de logro	Bueno	Sobresaliente
M1-C4 Describe situaciones o fenómenos empleando rigurosamente el lenguaje matemático y el lenguaje natural.	A		
	C		
	H		

Actividad formativa 10.1

Identifica, a tu juicio, si cada uno de los siguientes ejemplos corresponden a un estudio de carácter observacional o experimental.

1. Un investigador observa y registra el comportamiento de las personas en una plaza pública durante diferentes momentos del día, sin interferir en sus acciones. Observa cuántas personas están usando sus teléfonos móviles, qué tipo de actividades realizan (navegar por internet, enviar mensajes, realizar llamadas, etc.), la duración de cada actividad y otros detalles que consideres importantes.

El investigador podría utilizar un cuaderno para tomar notas, o incluso una cámara para registrar visualmente la actividad sin identificar a las personas. Anota la hora del día, la ubicación de la persona en la plaza, el tipo de dispositivo utilizado, etc.

Al revisar los datos recopilados, el investigador puede encontrar patrones sobre cuándo y cómo las personas utilizan sus dispositivos móviles en espacios públicos. Puede identificar tendencias, como un aumento en el uso de dispositivos móviles durante ciertas horas del día o preferencias de actividades en función de la edad o el género.

Es un estudio _____.

2. Se seleccionan dos grupos de participantes al azar. El Grupo A se designa como el grupo control y el Grupo B como el grupo experimental. Ambos grupos son similares en edad, sexo y nivel de condición física.

Durante cuatro semanas, el Grupo B realizó una rutina de ejercicio cardiovascular (correr, caminar rápido, etc.) durante 30 minutos al día, cinco días a la semana. El Grupo A no efectuó ningún cambio en su rutina diaria.

Al inicio del estudio, se midió la frecuencia cardíaca en reposo de ambos grupos; luego, se tomaron de nuevo las medidas de la frecuencia cardíaca; finalmente, después de cuatro semanas se toma la frecuencia cardíaca a ambos grupos y se compararán.

Se compararon los datos de la frecuencia cardíaca en reposo entre el Grupo A (grupo de control), y el Grupo B (grupo experimental) después de las cuatro semanas. Si la frecuencia cardíaca en reposo del Grupo B disminuye significativamente, en comparación con el Grupo A, se podría concluir que el ejercicio tuvo un efecto en la reducción de la frecuencia cardíaca en reposo.

Es un estudio _____.

Imagina que eres un investigador interesado en estudiar la relación entre el consumo de café y el rendimiento académico de los estudiantes universitarios.

Decides realizar un estudio observacional para lo cual seleccionas una muestra de 500 estudiantes universitarios y les pides que completen una encuesta sobre sus hábitos de consumo de café y su promedio de calificaciones. Recopilas los datos y analizas la relación entre el consumo de café y el promedio de calificaciones. Observas que los estudiantes que consumen más café tienden a tener un promedio de calificaciones más alto en comparación con aquellos que consumen menos o no consumen café.

Sin embargo, no puedes concluir que el consumo de café causa un mejor rendimiento académico, ya que puede haber otros factores que influyen en esta relación. Por ejemplo, los estudiantes que consumen más café pueden ser más dedicados o tener mejores hábitos de estudio, lo que a su vez conduce a un promedio de calificaciones más alto. El estudio observacional solo te permite identificar una asociación, pero no una relación causal.

Así que, para establecer una relación causal, puedes optar por diseñar un experimento. Seleccionas aleatoriamente a 100 estudiantes y los divides en dos grupos: un grupo de control y un grupo experimental. Durante un mes, el grupo experimental consume una cantidad específica de café todos los días, mientras que el grupo control no consume café. Todos los demás factores, como los hábitos de estudio y las horas de sueño, se mantienen constantes en ambos grupos.

Al final del mes, comparas el rendimiento académico de ambos grupos. Si observas una diferencia significativa en él, entre el grupo experimental y el grupo control, puedes concluir que el consumo de café tiene un efecto causal en el rendimiento académico. Este diseño experimental te permite controlar otros factores y aislar el efecto del consumo de café en el rendimiento académico.

En conclusión, los estudios observacionales son útiles para identificar asociaciones y generar hipótesis, pero no pueden establecer relaciones causales. Los experimentos, por otro lado, permiten manipular variables y controlar factores externos para determinar relaciones causales. Ambos tipos de estudios son valiosos en la investigación científica, pero es importante comprender sus diferencias y limitaciones al interpretar los resultados.

Los estudios observacionales y experimentales tienen enfoques diferentes, así como ventajas y desventajas; en este sentido, hay dos preguntas a las que debes dar respuesta al término de la progresión de aprendizaje:

- ¿Cuáles son las diferencias entre los estudios observacionales y experimentales?
- ¿Cuáles son las ventajas y desventajas de cada tipo de estudio?

Estudios observacionales

El uso masivo del celular es una tendencia que ha cambiado la forma de comunicarse y de interactuar entre las personas. Este fenómeno ha llamado la atención de investigadores, debido a que no sólo proporciona información valiosa sobre la interacción humana con la tecnología, sino que también, sobre las preferencias de los usuarios, con base en las consultas hechas en Internet. Esto tiene implicaciones significativas en la mercadotecnia móvil; por otra parte, su uso excesivo presenta la disminución de las horas de sueño o miedo a salir sin el celular, pues su dependencia ha llevado a algunas personas a ser incapaces de controlar o disminuir su uso. Estas problemáticas y otras más, las puedes investigar mediante la implementación de un estudio observacional.

Un **estudio observacional** implica la observación de fenómenos o comportamientos sin intervenir directamente en ellos. Esto es, el investigador observa individuos, fenómenos o poblaciones, sin intervenir ni modificar las condiciones naturales (tal como se desarrollan en su entorno habitual). Con el objetivo de recopilar datos sobre su comportamiento (costumbres o reacciones de un sujeto o grupo) mediante métodos como la observación directa, encuestas, análisis de registros médicos o sociales, entre otros. Todos ellos se analizan para identificar patrones, tendencias o relaciones entre variables (no se garantiza el descubrimiento de relaciones causales) para comprender mejor cómo interactúan diferentes factores en un entorno dado.

Por ejemplo, el comité de padres de familia puede solicitar a la dirección de un centro escolar un estudio sobre el consumo de comida chatarra por estudiantes en las cafeterías de la escuela, para comprender patrones alimenticios y preferencias dentro del entorno escolar. Esto implica la observación y recopilación de datos sobre los patrones de consumo de alimentos del estudiantado sin la intervención del investigador en sus elecciones alimenticias. Este estudio proporcionaría una comprensión detallada de cómo los estudiantes eligen y consumen comida chatarra, en un entorno escolar; sin embargo, no puedes establecer relaciones causales entre el consumo de productos chatarra y otros factores, ya que no hay manipulación directa de variables, ni control sobre las condiciones.

Ejemplo formativo 10.1

Justifica por qué cada uno de los siguientes estudios es observacional.

1. Un estudio que examina la relación entre el consumo de bebidas azucaradas y la obesidad en una población en un momento dado, recopilando datos sobre el consumo de bebidas y el índice de masa corporal de los participantes.

Justificación: este es un estudio observacional porque se recopilan datos en un solo punto en el tiempo sin intervención ni manipulación de las variables. Los investigadores observan y analizan las relaciones entre las variables tal como existen en la población estudiada.

2. Un estudio que sigue a un grupo de niños desde el nacimiento hasta la adolescencia para investigar cómo la exposición temprana a pesticidas afecta su desarrollo cognitivo.

Justificación: este es un estudio observacional porque los investigadores no manipulan la exposición a pesticidas. En cambio, registran y analizan cómo las diferentes exposiciones a pesticidas en los niños afectan su desarrollo cognitivo a lo largo del tiempo.

Existen diferentes tipos de estudios observacionales, cada uno con enfoques específicos para recolectar datos e información sin intervenir en las variables observadas. Los más comunes son:

Estudios de cohortes: se da seguimiento a un conjunto de individuos a lo largo de un período determinado con el propósito de investigar un evento particular. Una cohorte consiste en un grupo de personas que comparten una característica común, como la edad, el sexo o la ocupación, por ejemplo, dar seguimiento a un grupo de fumadores y no fumadores, a lo largo de 10 años, para observar el desarrollo de enfermedades pulmonares.

Estudios de casos y controles: comparan individuos con una condición particular (casos) con aquellos que no la tienen (controles) para identificar factores de riesgo, por ejemplo, si tienes una empresa en la que la eficiencia del uso de la tecnología permite obtener mayores ingresos, te interesaría saber si las diferencias en el uso de la tecnología están asociadas con un mejor desempeño laboral. En este estudio, los casos son los empleados con alto rendimiento laboral y los controles, los empleados con rendimiento laboral promedio o bajo.

Estudios transversales: recopilan datos de una población en un momento específico en el tiempo (sin continuidad en el tiempo). Se observan múltiples variables, al mismo tiempo, para establecer asociaciones; por ejemplo, si quieres analizar los hábitos de estudio que tus compañeros de grupo implementaron durante el primer año de estudio en el bachillerato puedes encuestarlos al final del ciclo escolar preguntando por las horas dedicadas al trabajo académico, uso de recursos educativos, etc.

Ejemplo formativo 10.2

Justifica por qué cada uno de los siguientes estudios es de cohortes, de casos y controles o transversal.

1. Un estudio que sigue a un grupo de personas con una dieta alta en grasas durante 10 años para observar la incidencia de enfermedades cardiovasculares.

Justificación: Este es un estudio observacional de cohorte porque los investigadores no manipulan la dieta de los participantes, ni asignan la dieta alta en grasas. Simplemente seleccionan a personas que ya siguen una dieta alta en grasas y las siguen a lo largo de 10 años para observar la incidencia de enfermedades cardiovasculares.

2. Un estudio que compara a mujeres con cáncer de mama (casos) con mujeres sin cáncer de mama (controles) para investigar su historial de uso de anticonceptivos orales.

Justificación: Este es un estudio observacional de casos y controles porque los investigadores no manipulan el uso de anticonceptivos orales. En lugar de eso, seleccionan a mujeres con y sin cáncer de mama y comparan retrospectivamente su exposición al uso de anticonceptivos orales para identificar posibles factores de riesgo asociados con la enfermedad.

3. Un estudio que evalúa la prevalencia de diabetes tipo 2 en una población urbana y su relación con el nivel de actividad física, midiendo ambos aspectos en un solo momento.

Justificación: Este es un estudio observacional transversal, porque se recopilan datos sobre la diabetes y la actividad física en un solo punto en el tiempo. Los investigadores no manipulan ninguna variable, sino que observan y analizan la relación entre la actividad física y la prevalencia de diabetes tal como existen en la población estudiada en ese momento.

Si bien, los estudios observacionales ofrecen información valiosa sobre las relaciones entre variables; sin embargo, su limitación radica en la dificultad para establecer relaciones causales directas, dado que no involucran la manipulación controlada de variables independientes. No obstante, estos estudios pueden proporcionar indicios que orienten hacia la realización de un estudio experimental posterior, permitiendo poner a prueba la relación identificada en el estudio observacional.

Estudios experimentales

En contraste con un estudio observacional, un **estudio experimental** implica la manipulación de variables para la observación y registro de los efectos causales, por

ejemplo, un docente para innovar su proceso de enseñanza y aprendizaje, diseña un experimento para investigar cómo diferentes métodos de enseñanza afectan el rendimiento académico del estudiantado. Para tal fin, asigna aleatoriamente a los estudiantes a diferentes grupos (grupo control y grupo experimental); luego, en cada uno implementa un método de enseñanza diferente para medir su efecto en el rendimiento académico, y así, saber cuál método es mejor.

En ejemplo anterior, el profesor investiga si hay una relación causal entre métodos de enseñanza y rendimiento académico, tomando en cuenta que puede haber otras variables (como el género, edad, nivel económico, estrato social, etc.), que no se controlaron y podrían influir en los resultados; a estas variables las conoces como variables de confusión.

Ejemplo formativo 10.3

Justifica por qué cada uno de los siguientes estudios es experimental.

1. Un estudio que investiga la eficacia de un nuevo medicamento para la hipertensión. Los participantes se asignan aleatoriamente a dos grupos: uno recibe el nuevo medicamento y el otro recibe un placebo. Los investigadores miden la presión arterial de ambos grupos durante seis meses.

Justificación: este es un estudio experimental porque los investigadores controlan la intervención (administración del nuevo medicamento) y asignan aleatoriamente a los participantes a diferentes grupos. La aleatorización ayuda a minimizar sesgos y permite una comparación directa entre los efectos del medicamento y el placebo.

2. Ejemplo: Un estudio que evalúa el impacto de un programa educativo intensivo en el rendimiento académico de los estudiantes de bachillerato. Se seleccionan escuelas al azar para implementar el programa, mientras que el resto continúan con el currículo estándar. Los resultados académicos se miden al final del año escolar.

Justificación: este es un estudio experimental porque los investigadores implementan una intervención específica (programa educativo intensivo) y controlan qué escuelas participan en el programa a través de una selección aleatoria. Esto permite evaluar el impacto del programa en comparación con el currículo estándar.

3. En un estudio que analiza el impacto de la música en la productividad laboral. Los empleados de una oficina se dividen aleatoriamente en dos grupos: uno trabaja con música de fondo y el otro sin música. La productividad se mide a través del número de tareas completadas en una semana.

Justificación: Este es un estudio experimental porque los investigadores manipulan la variable independiente (presencia de música) y asignan aleatoriamente a los

participantes a diferentes condiciones. Esto permite evaluar el efecto de la música en la variable dependiente (productividad laboral).

¿Cómo diseñar un experimento estadístico?

Como has observado, un estudio experimental implica que el investigador manipule conscientemente, una o varias variables independientes para observar el impacto que esta manipulación tiene sobre una variable dependiente. A fin de asegurar la validez y confiabilidad de los resultados, es importante seguir una serie de pasos en el diseño de un experimento estadístico. Estos pasos incluyen:

- Definir la **pregunta de investigación** y establecer una **hipótesis** sobre la variable dependiente a estudiar con relación a la independiente.
- Seleccionar la **población** y seleccionar mediante un método de muestreo adecuado una **muestra** representativa de la población.
- Identificar la variable dependiente y las independientes asegurándose de que sean cuantificables y observables.
- **Asignar aleatoriamente** los elementos de la muestra a dos grupos: el grupo de control, sin manipulación de variables independientes, y el grupo experimental, que sí experimenta con dichas variables.
- **Realizar la experimentación** siguiendo procedimientos estandarizados que garanticen resultados coherentes en repeticiones posteriores del experimento.
- **Analizar los datos** sometiendo los resultados a pruebas y análisis estadísticos para aceptar o rechazar la hipótesis.
- **Generar conclusiones** a partir del análisis de los datos para obtener y discutir las conclusiones sobre el experimento y la hipótesis.

En un estudio experimental:

- **Las variables independientes** son aquellas que el investigador manipula o controla deliberadamente. Se espera que estas variables causen un efecto en la variable dependiente.
- **La variable dependiente** es la variable que se observa y mide para determinar el efecto de la variable independiente o las variables manipuladas.
- **El grupo experimental** recibe la intervención o tratamiento experimental, es decir, la manipulación de la variable independiente.
- **El grupo de control** es el grupo de comparación que no recibe la intervención o tratamiento experimental. Sirve como punto de referencia para comparar los efectos observados en el grupo experimental.
- **La asignación aleatoria**, es un procedimiento estadístico para asignar a los participantes al azar a los grupos experimental y de control, y así, controlar mejor las variables de confusión y asegurar que las diferencias observadas sean el resultado de la intervención y no de otras causas.

Ejemplo formativo 10.4

1. En cada experimento plantea la pregunta de investigación y la hipótesis; identifica la variable independiente y la variable dependiente; describe el método.

Experimento 1. Estudio sobre el efecto del ejercicio en la reducción del estrés:

- **Pregunta de investigación:** ¿cuál es el efecto de un programa de ejercicio aeróbico supervisado durante ocho semanas en la reducción de los niveles de estrés percibido en adultos sanos?
- **Hipótesis:** se espera que un programa de ejercicio aeróbico durante ocho semanas reduzca los niveles de estrés en adultos sanos.
- **Variable independiente:** programa de ejercicio (grupo A: ejercicio aeróbico tres veces por semana durante ocho semanas, grupo B: sin ejercicio).
- **Variable dependiente:** niveles de estrés medidos mediante cuestionarios validados (por ejemplo, escala de estrés percibido).
- **Método:** se reclutan 50 adultos sanos y se asignan aleatoriamente a dos grupos. El grupo A sigue un programa de ejercicio aeróbico supervisado durante ocho semanas, mientras que el grupo B no realiza ejercicio adicional. Antes y después del período de 8 semanas, se administran cuestionarios para medir los niveles de estrés percibido en ambos grupos.

Experimento 2. Investigación sobre el efecto de la música en el aprendizaje de nuevos idiomas:

- **Pregunta de investigación:** ¿tiene la música clásica un impacto positivo en la retención de vocabulario en estudiantes de idiomas?
- **Hipótesis:** se espera que estudiar un vocabulario escuchando música clásica mejore su retención.
- **Variable independiente:** tipo de música (grupo A: música clásica, grupo B: sin música).
- **Variable dependiente:** retención de vocabulario medida mediante pruebas de conocimiento de palabras en el nuevo idioma.
- **Método:** se recluta a un grupo de 40 estudiantes de idiomas y se dividen aleatoriamente en dos grupos. Durante un mes, el grupo A estudia vocabulario mientras escucha música clásica, mientras que el grupo B estudia en silencio. A fin de mes se administran pruebas de conocimiento de palabras en el nuevo idioma a ambos grupos.

2. En cada experimento identifica las variables indicadas.

Experimento 1. Efecto de las baterías sobre el brillo de una bombilla.

- **Pregunta de investigación:** ¿cómo afecta la cantidad de baterías al brillo de la bombilla?
- **Variable independiente:** número de baterías.
- **Variable dependiente:** brillo de la bombilla.

- **Control de variables:** grosor de los cables, tamaño de las baterías, tipo de bombillas, temperatura del cable.

Experimento 2. Aplicación de un nuevo medicamento a enfermos de diabetes tipo

- **Pregunta de investigación:** ¿Cómo afecta la dosis de un nuevo medicamento en el control de los niveles de glucosa en sangre en pacientes con diabetes tipo 2?
- **Variable independiente:** Dosis del nuevo medicamento administrado.
- **Variable dependiente:** Niveles de glucosa en sangre de los pacientes.
- **Control de variables:** edad de los pacientes, género, tiempo de estar enfermo de diabetes tipo 2, peso y altura de los pacientes (índice de masa corporal), dieta y hábitos alimenticios y nivel de actividad física.

Los estudios experimentales son importantes para establecer relaciones de causa y efecto entre variables, ya que permiten controlar la influencia de factores externos y manipular variables para observar sus efectos directos. Son comunes en campos como en las humanidades, las ciencias sociales, las ciencias naturales y las ciencias exactas.

Evaluación formativa 10.1

1. Características, ventajas y desventajas de un estudio observacional.

Usa la inteligencia artificial para identificar las características principales, ventajas y desventajas de un estudio observacional.

Características principales de un estudio observacional	
Ventajas	Desventajas

2. Características, ventajas y desventajas de un estudio experimental.

Usa la inteligencia artificial para identificar las características principales, ventajas y desventajas de un estudio experimental.

Características principales de un estudio experimental	
Ventajas	Desventajas

3. Diferencia un estudio observacional de un estudio de diseño de experimentos justificando tu respuesta.

Estudio 1. Se hizo un estudio en el que se tomó una muestra aleatoria de adultos y se les preguntó sobre sus hábitos al levantarse por la mañana. Los datos mostraron que la gente que tomaba una caminata después de levantarse los volvería más activos y los agilizaría en las actividades del día.

¿Qué tipo de estudio es?

a) Observacional

b) Experimental

Justificación: _____

Estudio 2. Se tomó un grupo de agricultores de maíz y se dividió aleatoriamente en dos. A uno se le pidió agregar por una semana fertilizante orgánico a base de lixiviado de lombriz; mientras que al otro, se le pidió agregar fertilizante orgánico compuesto por lixiviado de guano de murciélago. Luego los investigadores compararon el aspecto (características físicas) de cada grupo de agricultores.

¿Qué tipo de estudio es?

a) Observacional

b) Experimental

Justificación: _____

Estudio 3: Se asignaron aleatoriamente voluntarios a uno de los dos grupos: a un grupo se le indicó usar redes sociales como de costumbre (*Facebook* y *TikTok*), al otro grupo se les bloqueó el acceso a *Facebook* y *TikTok*. Los investigadores analizaron cuál grupo tendía a ser más feliz.

¿Qué tipo de estudio es?

a) Observacional

b) Experimental

Justificación: _____
_____.

Estudio 4. Se seleccionó una muestra aleatoria de estudiantes y se examinaron sus hábitos académicos. Cada estudiante se clasificó como usuario esporádico, moderado o frecuente en la realización de actividades escolares como tareas y trabajos en clase. Los investigadores observaron cuál grupo tendía a estresarse menos.

¿Qué tipo de estudio es?

a) Observacional

b) Experimental

Justificación: _____
_____.

4. Identifica el tipo de estudio observacional.

a) Encuesta de satisfacción laboral.

Objetivo: determinar el nivel de satisfacción laboral en una empresa.

Método: se realiza un estudio donde se distribuye una encuesta a todos los empleados de la empresa en un momento específico para evaluar su satisfacción en relación con el ambiente laboral, la remuneración, las oportunidades de crecimiento, entre otros aspectos.

Este es un estudio observacional de tipo _____.

b) Enfermedades cardíacas y dieta.

Casos: personas con enfermedades cardíacas diagnosticadas.

Controles: personas sin enfermedades cardíacas (grupo de comparación).

Método: se desarrolla un estudio para comparar los patrones dietéticos entre ambos grupos buscando asociaciones entre ciertos tipos de alimentación y la incidencia de enfermedades cardíacas.

Este es un estudio observacional de tipo _____.

c) Seguimiento de desarrollo académico.

Objetivo: observar el impacto del estilo de enseñanza en el rendimiento académico.

Método: se elige un grupo de niños en edad preescolar y se les da seguimiento durante 11 años. Se registran datos sobre el tipo de educación recibida, el ambiente familiar, la participación en actividades educativas extracurriculares, y se evalúa su rendimiento académico en diferentes etapas.

Este es un estudio observacional de tipo _____.

5. Analiza con cuidado el siguiente experimento y completa lo faltante en los demás experimentos presentados.

Experimento. Efecto de fertilizantes en cultivos.

Pregunta de investigación: _____

Variable independiente: _____

Variable dependiente: _____

Control de variables: _____

6. Diseña un experimento donde puedas identificar las variables experimentales.

Pregunta de investigación: _____

Variable independiente: _____

Variable dependiente: _____

Control de variables: _____

Método: _____

Autoevaluación y coevaluación 10.1

Nombre: _____ Plantel: _____ Grupo: _____ Turno: _____

Autoevaluación para el aprendizaje

Selecciona en la columna, la opción que mejor refleje tu nivel de desempeño en el proceso para el aprendizaje de la progresión de aprendizaje 10. Responde con honestidad a la evaluación de cada uno de los criterios que se enlistan a continuación.

Desempeño	En proceso de logro	Bueno	Sobresaliente
Propicié un clima de comunicación favorable para el diseño de experimentos y estudios observacionales.			
Participé activamente con ideas para identificar ventajas y desventajas de estudios observacionales y diseño de experimentos.			
Contribuí colaborativamente en la retroalimentación de dudas sobre estudios observacionales y diseño de experimentos.			
Comuniqué diferencias y similitudes entre el diseño de experimentos y un estudio observacional (M1-C4).			

Coevaluación para el aprendizaje

Solicita a un compañero del equipo que marque en la columna la opción que mejor describa tu desempeño durante el trabajo colectivo, concluida la progresión de aprendizaje 10, y que responda con honestidad la evaluación de cada uno de los criterios que se enlistan a continuación.

Desempeño	En proceso de logro	Bueno	Sobresaliente
Propició un clima de comunicación favorable para el diseño de experimentos y estudios observacionales.			
Participó activamente con ideas para identificar ventajas y desventajas de estudios observacionales y diseño de experimentos.			
Contribuyó colaborativamente en la retroalimentación de dudas sobre estudios observacionales y diseño de experimentos.			
Comunicó diferencias y similitudes entre el diseño de experimentos y un estudio observacional (M1-C4).			

Nombre y firma de quien coevalúa

Las medidas estadísticas en el estudio de un fenómeno

Progresión de aprendizaje 11

Describe un fenómeno, problemática o situación de interés para el estudiantado utilizando las medidas de tendencia central (media, mediana y moda) y de dispersión (desviación estándar, varianza, rango intercuartil, etc.) adecuadas al contexto y valora qué tipo de conclusiones puede extraer a partir de dicha información.

Metas de aprendizaje		En proceso de logro	Bueno	Sobresaliente
M4-C2 Argumenta a favor o en contra de afirmaciones acerca de situaciones, fenómenos o problemas propios de la matemática, de las ciencias o de su contexto.	A			
	C			
	H			
M3-C3 Aplica procedimientos, técnicas y lenguaje matemático para la solución de problemas propios del Pensamiento Matemático, de Áreas de Conocimiento, Recursos Sociocognitivos, Recursos Socioemocionales y de su entorno.	A			
	C			
	H			

Actividad formativa 11.1

Selecciona la opción que consideres correcta:

1. ¿Cuál de las siguientes opciones es una medida de tendencia central?

- a) Rango b) Desviación estándar c) Media Aritmética

2. ¿Qué mide la mediana de un conjunto de datos?

- a) El valor más común b) El valor central c) La suma de todos los valores

3. ¿Cuál es la definición de moda en un conjunto de datos?

- a) El promedio de todos los valores.
b) El valor central cuando los datos están ordenados.
c) El valor más frecuente.

4. ¿Por qué es importante conocer las medidas de tendencia central y de dispersión en situaciones de la vida real? _____

La situación económica de las familias en la actualidad plantea un desafío significativo en la educación. Los padres desean que sus hijos estudien a pesar de los ingresos limitados que tienen, pues el costo de enviar a sus hijos a la preparatoria puede ser un obstáculo importante. Esto puede llevar a decisiones financieras difíciles, como sacrificar necesidades básicas o buscar ayuda adicional, y se deben considerar costos adicionales como transporte y actividades extracurriculares.

La inversión en educación es esencial para el futuro de los hijos, sin embargo, las dificultades económicas pueden crear una carga financiera significativa. Se necesita planificación cuidadosa y, a veces, recursos adicionales para superar estos desafíos económicos. Las medidas estadísticas pueden servir para comprender la magnitud de

estos desafíos y desarrollar estrategias para apoyar a las familias en su búsqueda de una educación de calidad.

En el contexto de la investigación y el análisis de datos, las medidas de tendencia central [Media Aritmética (\bar{X}), Mediana (Me), Moda (Mo)], las de posición (Cuartiles, Deciles, Percentiles) y las de dispersión [Rango, Varianza (s^2), Desviación Estándar (s)], son herramientas que nos sirven para comprender problemas del mundo real. Estas medidas, ampliamente utilizadas en diversas disciplinas, como economía, salud, educación y ciencias, permiten resumir y analizar datos para tomar decisiones informadas y abordar problemas de la vida cotidiana y complejos.

Medidas de tendencia central

Las **medidas de tendencia central** son estadísticas descriptivas que se utilizan para representar el centro de la distribución de una población o una muestra. Son útiles para resumir datos, facilitando con ello la interpretación de estos. Además, dependiendo de su naturaleza y contexto, una medida de tendencia central puede ser más apropiada que otra para representar un conjunto de datos.

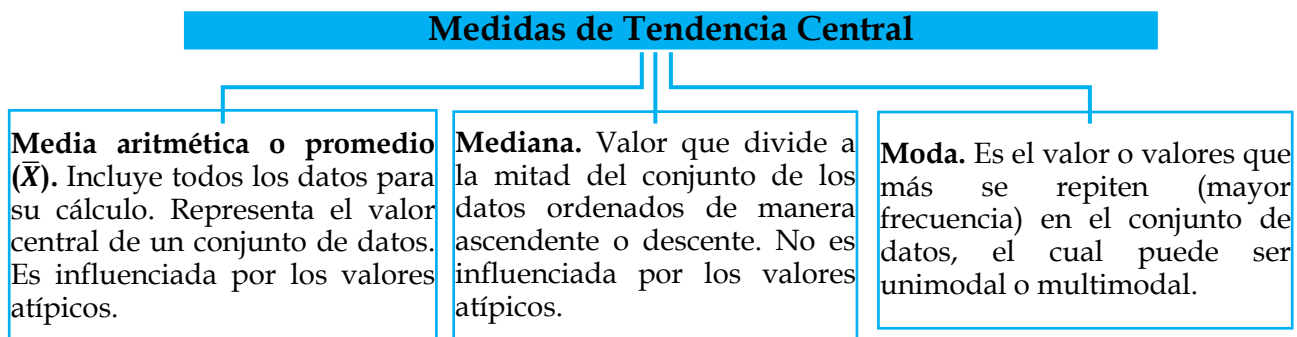


Figura 11.1 Medidas de tendencia central.
Fuente: Elaboración propia (Word, 2024).

Media Aritmética (\bar{X}). Se calcula sumando todas las observaciones obtenidas y dividiendo dicha suma entre el número total de elementos involucrados.

Mediana (Me). Una vez ordenados los datos de forma ascendente, si el conjunto de datos tiene un número impar de valores, la mediana es simplemente el valor central y si el conjunto de datos tiene un número par de valores, la mediana es el promedio de los dos valores centrales.

Moda (Mo). Es especialmente útil cuando se trata de datos categóricos o discretos, indica la opción más popular o frecuente.

Tipo de datos	Media Aritmética	Mediana	Moda
Datos Simples	$\bar{X} = \frac{\sum x}{n}$	Identifica la posición de la mediana con $\frac{n+1}{2}$ y después, localiza la mediana.	Es el dato que más se repite de un conjunto de datos.
Datos ordenados simples	$\bar{X} = \frac{\sum f \cdot x}{\sum f}$	Identifica la posición de la mediana con $\frac{\sum f + 1}{2}$ y después localiza la mediana.	Es el dato que más se repite de un conjunto de datos.

Tabla 11.1 Fórmulas para calcular las medidas de tendencia central.

Fuente: Elaboración propia (Word, 2024).

Ejemplo formativo 11.1

Los gastos relacionados con la educación de los hijos pueden incluir inscripción, libros de texto, uniformes, transporte, almuerzos escolares y otros gastos relacionados con la vida estudiantil. Por ejemplo, el costo semanal de la asistencia de los hijos a la preparatoria para un grupo de padres es: 100, 150, 90, 110, 200, 90, 110, 110, 140, 170.

1. Calcula la **media aritmética** (\bar{X}) para conocer el promedio del costo semanal:

$$\bar{X} = \frac{100 + 150 + 90 + 110 + 200 + 90 + 110 + 110 + 140 + 170}{10} = \frac{1270}{10} = 127$$

La media indica que, en promedio, para este grupo de padres, el costo semanal de enviar a su hijo a la preparatoria es de \$127.00.

2. Calcula la **mediana** (Me):

Primero ordena los datos, después determina la posición de la mediana, y por último, identifica la mediana.

Datos ordenados de menor a mayor: 90, 90, 100, 110, 110, 110, 140, 150, 170, 200.

$$\text{Posición de la mediana (PM)} = \frac{n+1}{2} = \frac{10+1}{2} = 5.5$$

El número de datos, $n = 10$ es par, por lo que, la posición de la mediana se localiza entre las cantidades 110 y 110.

$$Me = \frac{110 + 110}{2}$$

La mediana sugiere que la mitad de las familias gastan \$110.00 por semana.

3. Calcula la **moda** (Mo):

Recuerda que es el valor que aparece con mayor frecuencia en el conjunto de datos. En los datos: 90, 90, 100, 110, 110, 110, 140, 150, 170, 200, el valor 110 es el que más se repite.

$$Mo = 110$$

La moda indica que el costo semanal más frecuente es de \$110.00.

La gráfica de puntos con las medidas de tendencia central del costo semanal es:

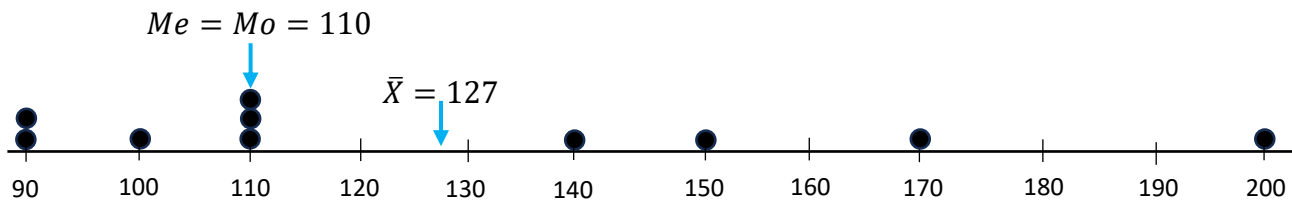


Figura 11.2. Gráfico de puntos.

Fuente: Elaboración propia (Word, 2024).

¿Qué medida es mejor representante de cada conjunto de datos? _____

¿Por qué? _____

Ejemplo formativo 11.2

Determina las medidas de tendencia central para los siguientes datos organizados en una tabla de frecuencias.

Tabla de frecuencias de datos simples			Medidas de tendencia central																										
<table border="1"> <thead> <tr> <th>Costo semanal (x)</th> <th>Frecuencia (f)</th> <th>$f \cdot x$</th> </tr> </thead> <tbody> <tr> <td>90</td> <td>2</td> <td>180</td> </tr> <tr> <td>100</td> <td>1</td> <td>100</td> </tr> <tr> <td>110</td> <td>3</td> <td>330</td> </tr> <tr> <td>140</td> <td>1</td> <td>140</td> </tr> <tr> <td>150</td> <td>1</td> <td>150</td> </tr> <tr> <td>170</td> <td>1</td> <td>170</td> </tr> <tr> <td>200</td> <td>1</td> <td>200</td> </tr> <tr> <td>Totales</td> <td>10</td> <td>1270</td> </tr> </tbody> </table>	Costo semanal (x)	Frecuencia (f)	$f \cdot x$	90	2	180	100	1	100	110	3	330	140	1	140	150	1	150	170	1	170	200	1	200	Totales	10	1270	<p>Media aritmética</p> $\bar{X} = \frac{\sum f \cdot x}{\sum f} = \frac{1270}{10} = 127$ <p>La media es 127, lo que indica que en promedio, los padres gastan \$127.00 a la semana en la educación de sus hijos.</p> <p>Mediana</p> $PM: \frac{\sum f + 1}{2} = \frac{10 + 1}{2} = 5.5$ $Me = \frac{110 + 110}{2} = 110$ <p>La mediana es 110, lo que indica que la mitad de los padres gasta menos de \$110.00 y la otra mitad gasta más de \$110.00 a la semana en la educación de sus hijos.</p> <p>Moda</p> $Mo = 110$ <p>La moda es 110, lo que indica que el gasto más común entre los padres es de \$110.00 a la semana.</p>	
Costo semanal (x)	Frecuencia (f)	$f \cdot x$																											
90	2	180																											
100	1	100																											
110	3	330																											
140	1	140																											
150	1	150																											
170	1	170																											
200	1	200																											
Totales	10	1270																											

Medidas de posición

Las **medidas de posición** son estadísticas que se utilizan para describir la posición relativa de un valor dentro de un conjunto de datos. Estas medidas ayudan a comprender dónde se encuentra un valor en relación con el resto de los datos. Los cinco números de resumen son: mínimo, primer cuartil, segundo cuartil (mediana), el tercer

cuartil y el máximo. Los **cuartiles** Q_1, Q_2, Q_3 , son valores que dividen a un conjunto de datos ordenados en cuatro partes iguales, y el cuartil Q_2 coincide con la mediana del conjunto de datos.

Medidas de posición	
<p>Cuartiles. <i>mínimo</i> (mín.), $Q_1, Q_2=Me,$ $Q_3,$ <i>máximo</i> (máx.)</p>	<p>Localiza la posición del cuartil</p> $P_k = \begin{cases} \frac{kn}{4} + 0.5, & \text{si } n \text{ es par} \\ \frac{k(n+1)}{4}, & \text{si } n \text{ es impar} \end{cases}$ <p>k es el cuartil que se busca. n es el número de datos.</p> <p>Calcular los cuartiles implica dividir un conjunto de datos ordenado en cuatro partes iguales, donde cada parte representa el 25 % de los datos. Los cuartiles se utilizan para describir la distribución de los datos. Los tres cuartiles principales son:</p> <ul style="list-style-type: none"> • Primer Cuartil (Q_1): El valor que separa el 25 % inferior de los datos. • Segundo Cuartil (Q_2): También conocido como la mediana, el valor que separa el 50 % inferior de los datos. • Tercer Cuartil (Q_3): El valor que separa el 75 % inferior de los datos. <p>Para datos entablas de frecuencias simple, busca la posición del cuartil en la frecuencia absoluta acumulada y se identifica el dato que corresponde al cuartil.</p>

Tabla 11.2 Fórmulas para calcular las medidas de posición.

Fuente: Elaboración propia (Word, 2024).

Ejemplo formativo 11.3

Continuando con los datos simples del Ejemplo formativo 11.1, correspondientes al costo semanal de enviar a los hijos a la preparatoria: 100, 150, 90, 110, 200, 90, 110, 110, 140, 170. Calcula las medidas de posición datos simples o cinco números de resumen.

Medidas de posición.

Datos ordenados: 90, 90, 100, 110, 110, 110, 140, 150, 170, 200

$$\text{mín.} = 90 \text{ y } \text{máx.} = 200$$

Dado que el número de datos $n = 10$, es par, la posición de cada cuartil es:

$$P = \frac{kn}{4} + 0.5$$

Posición de Q_1	Posición de Q_2	Posición de Q_3
$P_1 = \frac{1 \cdot 10}{4} + 0.5 = 3$	$P_2 = \frac{2 \cdot 10}{4} + 0.5 = 5.5$	$P_3 = \frac{3 \cdot 10}{4} + 0.5 = 8$

$$Q_2 = \frac{110 + 110}{2} = 110$$

$90, 90, 100, 110, 110, 110, 140, 150, 170, 200$
 $\uparrow \qquad \qquad \downarrow \qquad \qquad \uparrow$
 $Q_1 = 100 \qquad \qquad \qquad Q_3 = 100$

El *mín.* = 90, $Q_1 = 100$, $Q_2 = Me = 110$, $Q_3 = 150$ y *máx.* = 200

Primer Cuartil ($Q_1 = 100$): El 25 % de los padres gasta \$100.00 o menos en la educación semanal de sus hijos.

Mediana ($Q_2 = 110$): El 50 % de los padres gasta \$110.00 o menos en la educación semanal de sus hijos.

Tercer Cuartil ($Q_3 = 150$): El 75 % de los padres gasta \$150.00 o menos en la educación semanal de sus hijos.

La gráfica de cajas y de puntos con las medidas de posición.

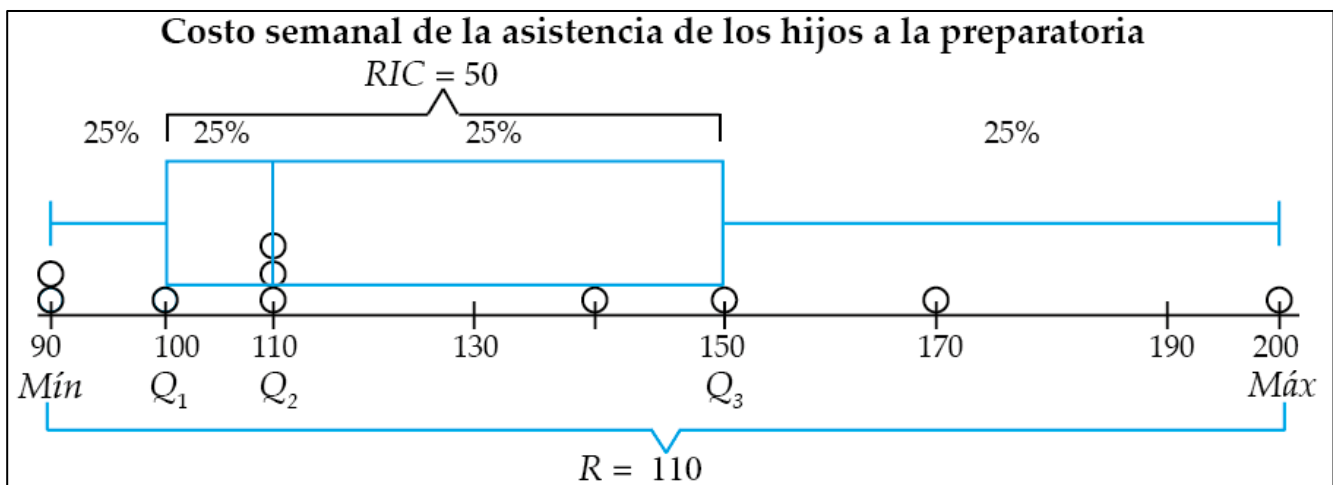


Figura 11.3 Costo semanal de la asistencia de los hijos a la preparatoria.
Fuente: Elaboración propia (Word, 2024).

Ejemplo formativo 11.4

Continuando con los datos organizados en una tabla de frecuencias simple, del Ejemplo formativo 11.2, correspondientes al costo semanal de enviar a los hijos a la preparatoria. Cálculo de las medidas de posición.

Tabla de frecuencias de datos simples			Medidas de posición																										
<table border="1"> <thead> <tr> <th>Costo semanal (x)</th> <th>Frecuencia (f)</th> <th>Frecuencia acumulada</th> </tr> </thead> <tbody> <tr><td>90</td><td>2</td><td>2</td></tr> <tr><td>100</td><td>1</td><td>3</td></tr> <tr><td>110</td><td>3</td><td>6</td></tr> <tr><td>140</td><td>1</td><td>7</td></tr> <tr><td>150</td><td>1</td><td>8</td></tr> <tr><td>170</td><td>1</td><td>9</td></tr> <tr><td>200</td><td>1</td><td>10</td></tr> <tr><td>Totales</td><td>10</td><td></td></tr> </tbody> </table>	Costo semanal (x)	Frecuencia (f)	Frecuencia acumulada	90	2	2	100	1	3	110	3	6	140	1	7	150	1	8	170	1	9	200	1	10	Totales	10		<p>Cálculo de los cuartiles</p> <p>Posición del cuartil: $P_k = \frac{kn}{4} + 0.5$</p> $P_1 = \frac{(1)(10)}{4} + 0.5 = 3 \quad Q_1 = 100$ $P_2 = \frac{(2)(10)}{4} + 0.5 = 5.5 \quad Q_2 = 110$ $P_3 = \frac{(3)(10)}{4} + 0.5 = 8 \quad Q_3 = 150$	
Costo semanal (x)	Frecuencia (f)	Frecuencia acumulada																											
90	2	2																											
100	1	3																											
110	3	6																											
140	1	7																											
150	1	8																											
170	1	9																											
200	1	10																											
Totales	10																												
<p>El <i>mín.</i> = 90, $Q_1 = 100$, $Q_2 = Me = 110$, $Q_3 = 150$ y <i>máx.</i> = 200</p> <p>Primer Cuartil ($Q_1 = 100$): El 25 % de los padres gasta \$100.00 o menos en la educación semanal de sus hijos.</p> <p>Mediana ($Q_2 = 110$): El 50 % de los padres gasta \$110.00 o menos en la educación semanal de sus hijos.</p> <p>Tercer Cuartil ($Q_3 = 150$): El 75 % de los padres gasta \$150.00 o menos en la educación semanal de sus hijos.</p>																													

Medidas de dispersión

Las **medidas de dispersión** son esenciales para comprender la variabilidad en los datos, evaluar la consistencia de los resultados y tomar decisiones informadas en una amplia gama de disciplinas, desde la estadística hasta la investigación científica; ayudan a estudiar la dispersión de un conjunto de datos, mediante la cuantificación de la dispersión. Por ejemplo:

- El **rango** de un conjunto de datos mide la diferencia entre el valor máximo y el valor mínimo del conjunto; es una medida de dispersión que proporciona una idea de cuán dispersos están los datos, es decir, la extensión total en la cual se encuentran distribuidos los valores.
- El **rango intercuartil (RIC)**, también conocido como rango intercuartílico o recorrido intercuartílico. Este mide la dispersión de los datos en un conjunto, al

considerar el rango entre el primer cuartil (Q_1) y el tercer cuartil (Q_3). Específicamente, mide la amplitud del intervalo que contiene el 50 % central de los datos, eliminando el efecto de los valores atípicos o extremos.

- La **varianza** mide la dispersión de los datos respecto a su media.
- La **desviación estándar** mide cuánto se desvían los valores con respecto a la media. Una desviación estándar alta, indica mayor dispersión de los datos, mientras que una desviación estándar baja indica menor dispersión.

A continuación, se describen las fórmulas:

Medidas de dispersión			
Rango	Se calcula restando el valor mínimo del valor máximo en un conjunto de datos. Representa la diferencia entre el puntaje más alto y el puntaje más bajo, y da una idea de la extensión total de los datos. $\text{Rango} = \text{máx.} - \text{mín.}$		
Rango intercuartil	$RIC = Q_3 - Q_1$		
Varianza	<table border="0" style="width: 100%;"> <tr> <td style="width: 50%; vertical-align: top;"> De una población $s^2 = \frac{\sum(x - \bar{X})^2}{N}$ De una muestra $s^2 = \frac{\sum(x - \bar{X})^2}{n - 1}$ Método abreviado $s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n - 1}$ </td> <td style="width: 50%; vertical-align: top;"> De una muestra $s^2 = \frac{\sum f(x_i - \bar{X})^2}{n - 1}$ Método abreviado $s^2 = \frac{\sum fx^2 - \frac{(\sum fx)^2}{n}}{n - 1}$ </td> </tr> </table>	De una población $s^2 = \frac{\sum(x - \bar{X})^2}{N}$ De una muestra $s^2 = \frac{\sum(x - \bar{X})^2}{n - 1}$ Método abreviado $s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n - 1}$	De una muestra $s^2 = \frac{\sum f(x_i - \bar{X})^2}{n - 1}$ Método abreviado $s^2 = \frac{\sum fx^2 - \frac{(\sum fx)^2}{n}}{n - 1}$
De una población $s^2 = \frac{\sum(x - \bar{X})^2}{N}$ De una muestra $s^2 = \frac{\sum(x - \bar{X})^2}{n - 1}$ Método abreviado $s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n - 1}$	De una muestra $s^2 = \frac{\sum f(x_i - \bar{X})^2}{n - 1}$ Método abreviado $s^2 = \frac{\sum fx^2 - \frac{(\sum fx)^2}{n}}{n - 1}$		
Desviación Estándar	<table border="0" style="width: 100%;"> <tr> <td style="width: 50%; text-align: center;"> $s = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n - 1}}$ </td> <td style="width: 50%; text-align: center;"> $s = \sqrt{\frac{\sum fx^2 - \frac{(\sum fx)^2}{n}}{n - 1}}$ </td> </tr> </table>	$s = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n - 1}}$	$s = \sqrt{\frac{\sum fx^2 - \frac{(\sum fx)^2}{n}}{n - 1}}$
$s = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n - 1}}$	$s = \sqrt{\frac{\sum fx^2 - \frac{(\sum fx)^2}{n}}{n - 1}}$		

Tabla 11.3 Fórmulas para calcular las medidas de dispersión.

Fuente: Elaboración propia (Word, 2024).

Ejemplo formativo 11.4

Cálculo de las medidas de dispersión para los datos simples del Ejemplo formativo 11.1

90, 90, 100, 110, 110, 110, 140, 150, 170, 200

1. **Rango.** Para calcular el rango, primero identificamos el valor mínimo y el valor máximo en el conjunto de datos. Luego, restamos el valor mínimo del valor máximo.

$$\text{Rango} = 200 - 80 = 110$$

El rango es 110, lo que significa que hay una diferencia de \$110.00 entre el gasto semanal más alto y el gasto semanal más bajo en la educación de los hijos. Es decir, mientras

algunos padres pueden estar gastando relativamente poco (como \$90 por semana), otros están gastando mucho más (hasta \$200 por semana).

2. **Rango intercuartil.** Se calcula como la diferencia entre Q_3 y Q_1 .

$$\text{El rango intercuartil, } RIC = Q_3 - Q_1 = 150 - 100 = 50$$

El rango intercuartil de 50 significa que la mitad central de los gastos semanales en educación de los hijos de los padres está distribuida dentro de un intervalo de \$50.00.

Ahora vamos a realizar el cálculo de las medidas de dispersión para datos simples:

3. **Varianza de los datos (para una muestra).** se calcula como el promedio de las diferencias al cuadrado entre cada valor y la media.

90, 90, 100, 110, 110, 110, 140, 150, 170, 200.

$$s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n - 1}$$

$$s^2 = \frac{173500 - \frac{(1270)^2}{10}}{9}$$

$$s^2 = \frac{12210}{9}$$

$$s^2 \approx 1356.67$$

x	x^2
90	8100
90	8100
100	10000
110	12100
110	12100
110	12100
140	19600
150	22500
170	28900
200	40000
Total	1270
	173500

4. **Desviación estándar.** Es la raíz cuadrada de la varianza.

$$s \approx \sqrt{1356.67} \approx 36.83$$

La desviación del costo semanal de la asistencia de los hijos a la preparatoria con respecto a la media es de \$36.83 en promedio.

Ejemplo formativo 11.5

Continuando con los datos organizados en una tabla de frecuencias simple, del Ejemplo formativo 11.2, correspondientes al costo semanal de enviar a los de enviar a los hijos a la preparatoria. Cálculo de las medidas de dispersión.

Medidas de dispersión a partir de una tabla de frecuencias de datos simples

Costo semanal (x)	Frecuencia (f)	$f \cdot x$	x^2	$f \cdot x^2$
90	2	180	8100	16200
100	1	100	10000	10000
110	3	330	12100	36300
140	1	140	19600	19600
150	1	150	22500	22500
170	1	170	28900	28900
200	1	200	40000	40000
Totales	10	1270		173500

Rango: $200 - 90 = 110$

Rango intercuartílico: $RIC = 150 - 100 = 50$

Cálculo de la varianza

$$\begin{aligned} s^2 &= \frac{\sum fx^2 - \frac{(\sum fx)^2}{n}}{n - 1} \\ &= \frac{173500 - \frac{(1270)^2}{10}}{10 - 1} \\ &= \frac{173500 - 161290}{9} \\ &\approx 1356.67 \end{aligned}$$

Cálculo de la desviación estándar

$$s \approx \sqrt{1356.67} \approx 36.83$$

La desviación del costo semanal de la asistencia de los hijos a la preparatoria con respecto a la media es de \$36.83 en promedio.

Evaluación formativa 11.1

1. Recaba las calificaciones de la Progresión 9 de Pensamiento Matemático I de al menos 15 compañeros.

Datos:

- a) Determina las medidas de tendencia central e interprétalas con base al contexto de los datos.

- b) Determina las medidas de posición e interprétalas con base al contexto de los datos.

$mín.$ = _____, Q_1 = _____, Q_2 = _____, Q_3 = _____, $máx.$ = _____

- c) Calcula la desviación estándar e interprétala con base al contexto de los datos.

2. Según lo aprendido, ¿en qué otros aspectos de la vida cotidiana se pueden utilizar las medidas de tendencia central y de dispersión?

3. Proporciona una situación de la vida cotidiana en donde puedas implementar lo aprendido.

4. En una pequeña tienda de ropa, Juan el dueño, se enfrentaba a un desafío. Quiere fijar los precios de sus camisetas, de manera que sean atractivas tanto para sus clientes como para su rentabilidad. Decidió analizar el precio de las camisetas en su localidad para tomar una decisión informada. Estos son precios que Juan recopiló de diferentes negocios y que puede considerar para tomar decisiones sobre la fijación de precios en su tienda de ropa:

\$150.00, \$200.00, \$180.00, \$250.00, \$220.00, \$190.00, \$170.00, \$210.00, \$160.00, \$190.00

- a) Calcula el precio promedio de las camisetas.

- b) Encuentra la moda de los precios y explica su importancia para la fijación de precios.
- c) ¿Cómo interpretarías la mediana en el contexto de la tienda de ropa?
- d) Calcula el percentil 25 (Q_1) y discute su relevancia para la estrategia de precios.
- e) ¿Qué información proporciona el percentil 75 (Q_3)? Y, ¿cómo podría influir en las decisiones de fijación de precios?
- f) ¿Por qué podría ser útil considerar diferentes estrategias de precios para distintos percentiles?
- g) Calcula el rango de los precios de las camisetas y explora cómo podría ser utilizado para establecer precios mínimos y máximos.
- h) Analiza la varianza y la desviación estándar. ¿Cómo podrían estas medidas influir en la estabilidad de los precios en la tienda?
- i) ¿Qué estrategias de fijación de precios podrías recomendar basándote en las medidas de tendencia central?
- j) Considera la dispersión de precios. ¿Cómo podrías sugerir a Juan que ajuste sus precios para mantener una gama atractiva pero consistente?

Autoevaluación y coevaluación 11.1

Autoevaluación y coevaluación

Nombre: _____ Plantel: _____ Grupo: _____ Turno: _____

Autoevaluación para el aprendizaje

Selecciona en la columna la opción que mejor refleje tu nivel de desempeño en el proceso para el aprendizaje de la progresión de aprendizaje 11. Responde con honestidad a la evaluación de cada uno de los criterios que se enlistan a continuación.

Aspectos	En proceso de logro	Bueno	Sobresaliente
Propicié un clima de comunicación favorable para el estudio de Las medidas estadísticas en el estudio de un fenómeno.			
Interpreté el resultado de las medidas estadísticas de acuerdo con el contexto de los datos.			

Coevaluación:

Solicita a un compañero marque en la columna la opción que mejor describa tu desempeño durante la progresión de aprendizaje 11, pide que responda con honestidad la evaluación de cada uno de los criterios que se enlistan a continuación.

Desde tu punto de vista, el desempeño de tu compañero es:

Aspectos	En proceso de logro	Bueno	Sobresaliente
Propició un clima de comunicación favorable para el estudio de Las medidas estadísticas en el estudio de un fenómeno.			
Participó activamente con ideas en el aprendizaje de las medidas estadísticas (tendencia central, de posición y de dispersión).			
Contribuyó colaborativamente en la retroalimentación de dudas sobre el cálculo de las medidas estadísticas.			
Su desempeño al realizar las actividades fue.			
Interpretó el resultado de las medidas estadísticas de acuerdo con el contexto de los datos.			

Nombre y firma de quien coevalúa

La distribución normal

Progresión de aprendizaje 12

Explica un evento aleatorio cuyo comportamiento puede describirse a través del estudio de la distribución normal y calcula la probabilidad de que dicho evento suceda.

Metas de aprendizaje		En proceso de logro	Bueno	Sobresaliente
M3-C1 Comprueba los procedimientos usados en la resolución de problemas utilizando diversos métodos, empleando recursos tecnológicos o la interacción con sus pares.	A			
	C			
	H			
M4-C2 Argumenta a favor o en contra de afirmaciones acerca de situaciones, fenómenos o problemas propios de la matemática, de las ciencias o de su contexto.	A			
	C			
	H			
M3-C3 Aplica procedimientos, técnicas y lenguaje matemático para la solución de problemas propios del Pensamiento Matemático, de Áreas de Conocimiento, Recursos Sociocognitivos, Recursos Socioemocionales y de su entorno.	A			
	C			
	H			



Figura 12.1. Distribución normal de la estatura de estudiantes.
Fuente: Fotografía. César Pilar Quintero.

Evaluación diagnóstica 12.1

Lee el ejercicio y elige la respuesta correcta.

- Si un dado es lanzado 50 veces y el número 6 sale 15 veces, ¿cuál es la probabilidad de que salga un 6 en el próximo lanzamiento?
 - Mayor que en los lanzamientos anteriores, debido a la cantidad de veces que ya ha salido.
 - Menor que en los lanzamientos anteriores, debido a la cantidad de veces que ya ha salido.
 - Exactamente $1/6$, ya que cada lanzamiento es independiente.
 - No hay suficiente información para determinar la probabilidad.

2. Un equipo de baloncesto tiene jugadores con los siguientes puntos por partido: 10, 15, 20, 25, 30. Si el jugador que anota 30 puntos por partido se lesiona y es reemplazado por un nuevo jugador que anota 15 puntos por partido, ¿qué efecto tiene esto en la media de puntos por partido del equipo?
- A) La media disminuye.
 B) La media aumenta.
 C) La media se mantiene igual.
 D) La media se duplica.
3. En una distribución normal, ¿qué representa el valor z en el cálculo de probabilidades?
- A) Un valor atípico en el conjunto de datos.
 B) La media de la distribución.
 C) Un valor estandarizado que indica cuántas desviaciones estándar un punto de datos está de la media.
 D) El valor mínimo en el conjunto de datos.

4. Observa la gráfica de la derecha:
 Con base en la región sombreada, ¿qué probabilidad se está calculando?
- A) $P(Z > -3)$
 B) $P(Z \leq 0)$
 C) $P(Z = 0)$
 D) $P(Z < 3)$

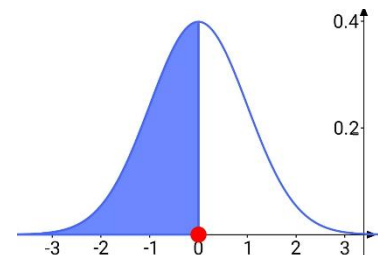


Figura 12.2. Gráfica de la distribución normal estándar.
 Fuente: Probability Distributions (Android, 2024).

5. Observa la gráfica de la derecha:
 Con base en la región sombreada, ¿qué probabilidad se está calculando?
- A) $P(Z > -2)$
 B) $P(Z \leq -2)$
 C) $P(Z \leq 1)$
 D) $P(Z \leq 1) - P(Z \leq -2)$

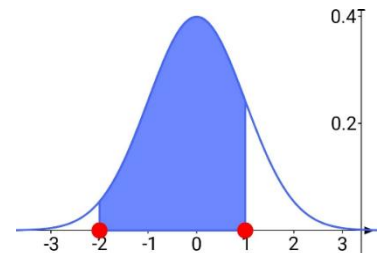


Figura 12.3. Gráfica de la distribución normal estándar.
 Fuente: Probability Distributions (Android, 2024).

6. Observa la gráfica de la derecha:
 Con base en la región sombreada, ¿qué probabilidad se está calculando?
- A) $P(Z \geq -1)$
 E) $P(Z \leq -1)$
 F) $P(-1 \leq Z \leq 3)$
 G) $P(Z \geq -1) - P(Z \geq 3)$

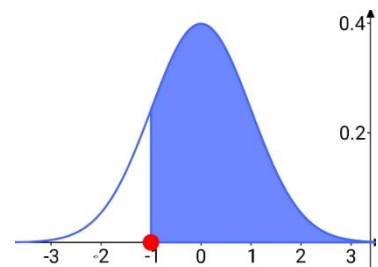
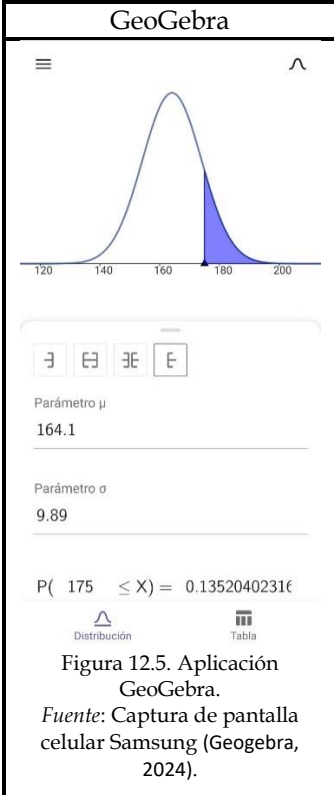
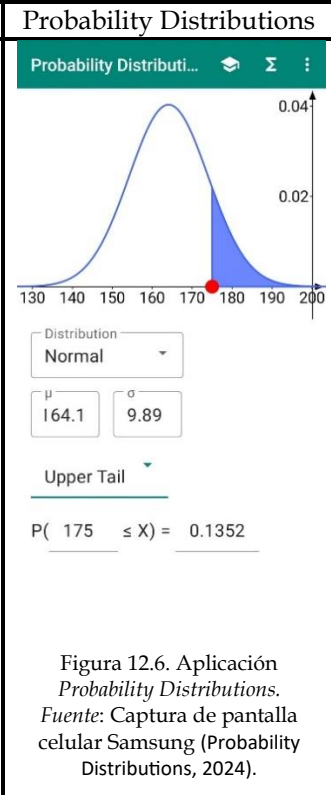
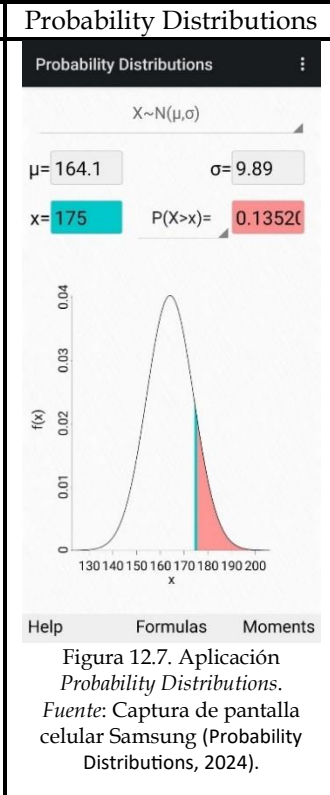



Figura 12.4. Gráfica de la distribución normal estándar.
 Fuente: Probability Distributions (Android, 2024).

Los y las estudiantes de una escuela del nivel bachillerato realizaron medidas sobre la estatura de 170 estudiantes, dando como resultado la **estatura media** de 164.1 cm y la **desviación estándar** de 9.89 cm. Si las estaturas se distribuyen normalmente, ¿cómo realizarías los siguientes incisos?:

- Calcular la probabilidad de que un estudiante elegido al azar mida más de 175 cm.
- ¿Cuántos estudiantes esperas que midan más de 175 cm?

Para resolver la situación anterior, puedes estandarizar la variable estatura de un estudiante y usar los valores de una tabla de la distribución normal z (ver Tabla 12.1); así como, usar una de las siguientes aplicaciones en el celular.

 <p>Figura 12.5. Aplicación GeoGebra. Fuente: Captura de pantalla celular Samsung (Geogebra, 2024).</p>	 <p>Figura 12.6. Aplicación Probability Distributions. Fuente: Captura de pantalla celular Samsung (Probability Distributions, 2024).</p>	 <p>Figura 12.7. Aplicación Probability Distributions. Fuente: Captura de pantalla celular Samsung (Probability Distributions, 2024).</p>	 <p>Figura 12.8. Aplicación Distribución normal. Fuente: Captura de pantalla celular Samsung (Probability Distributions, 2024).</p>
---	---	--	---

La distribución normal

La distribución de probabilidad normal se usa en la vida cotidiana para explicar eventos aleatorios, como en estadística en el análisis de datos en finanzas, en la modelación del comportamiento de variables aleatorias, en ingeniería para el diseño de sistemas robustos, en psicología para el estudio del comportamiento humano, en medicina para evaluar y diagnosticar enfermedades basadas en valores de referencia, entre otros.

Los parámetros de la distribución normal son la media μ y la desviación estándar σ . En este caso se dice que la variable aleatoria X tiene una distribución normal $N(\mu, \sigma)$ y se denota como $X \sim N(\mu, \sigma)$.

Es una distribución de probabilidad continua que se utiliza para modelar variables aleatorias en las cuales, los valores se agrupan alrededor de la media y se dispersan simétricamente. Como se muestra en la Figura 12.9, esta tiene una forma de campana y está completamente definida por su media μ y desviación estándar σ .

En el caso de que $\mu = 0$ y $\sigma = 1$, así como se ve en la Figura 12.10, se usa la variable Z para representar la variable aleatoria con una **distribución normal estándar**, es decir, $Z \sim N(0,1)$; la cual se usa para el cálculo de probabilidades determinando el área bajo la curva en un rango específico de valores.

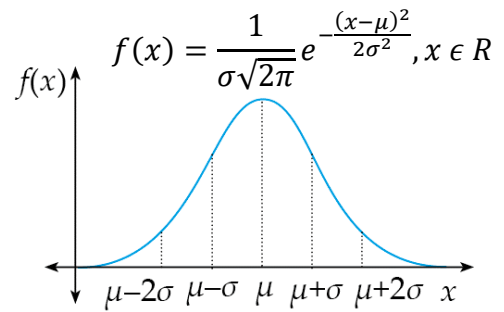


Figura 12.9. Curva de la distribución normal $N(\mu, \sigma)$.

Fuente: Elaboración propia (Excel, 2024).

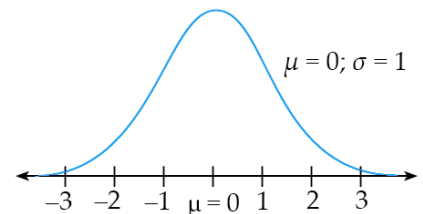


Figura 12.10. Curva de la distribución normal $N(0, 1)$.

Fuente: Elaboración propia (Excel, 2024).

También se utiliza como referencia para comparar y analizar otras distribuciones normales dadas en las Figuras 12.11 y 12.12.

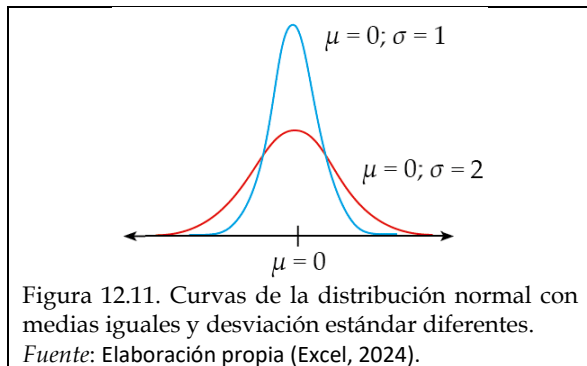


Figura 12.11. Curvas de la distribución normal con medias iguales y desviación estándar diferentes.
Fuente: Elaboración propia (Excel, 2024).

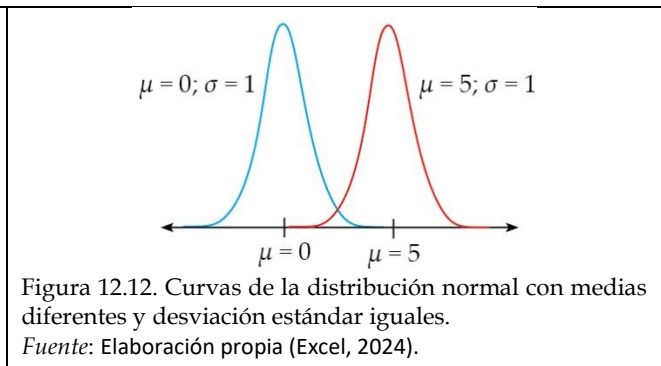


Figura 12.12. Curvas de la distribución normal con medias diferentes y desviación estándar iguales.
Fuente: Elaboración propia (Excel, 2024).

Cálculo de probabilidades utilizando la curva normal estándar

La tabla del área bajo la curva de la distribución normal estándar se utiliza para calcular probabilidades de la variable Z asociada a la distribución normal estándar, $Z \sim N(0, 1)$. Mediante ella puedes determinar la probabilidad de que el valor z de una variable aleatoria Z , caiga dentro de un rango específico de valores, por ejemplo, $P(Z \leq z)$, $P(Z \geq z)$ o $P(-z \leq Z \leq z)$, como se muestra en la Figura 12.13.

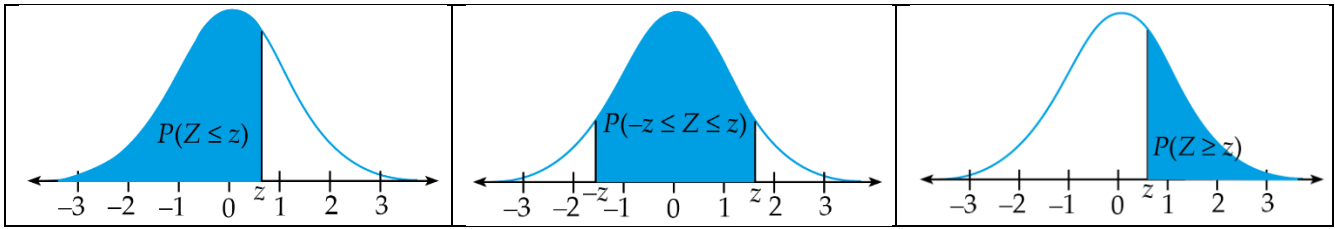


Figura 12.13. El área bajo la curva de la distribución normal estándar $N(0, 1)$ y su relación con la probabilidad.

Fuente: Elaboración propia (GeoGebra, 2024).

Este cálculo se basa en las siguientes consideraciones:

1. Es importante entender que, en el cálculo de áreas bajo la curva de la distribución normal estándar no se trabaja con valores x , sino con valores z ; el acceso a los valores de Tabla 12.1 es con valores z .
2. Debido a que la tabla estándar utiliza valores z debemos transformar los valores x en valores z . Esto se llama estandarizar los valores x . La estandarización se hace aplicando la fórmula: $x = \mu + z\sigma$.

$$\text{Despejando } z: z = \frac{x - \mu}{\sigma}$$

3. El área bajo toda la curva normal es igual a 1.
4. En la Tabla 12.1 se enumeran todas las posibles áreas en los intervalos que empiezan en la media (localizada en $z = 0.0000$) y terminan en un valor específico de z .
5. Las probabilidades (áreas) de otros intervalos, se calculan usando los elementos de la tabla y aplicando operaciones de suma y resta, según las propiedades de la curva normal.

La regla empírica en la distribución normal $N(0,1)$, como se muestra en la Figura 12.14, establece que aproximadamente el 68% de los datos se encuentran dentro de una desviación estándar alrededor de la media, el 95% se encuentra dentro de dos desviaciones estándar y el 99.7% se encuentra dentro de tres desviaciones estándar.

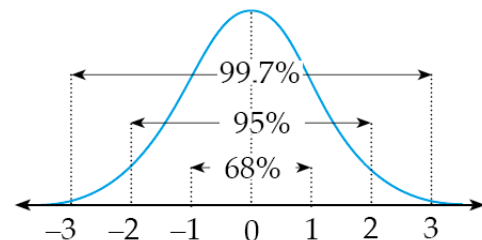


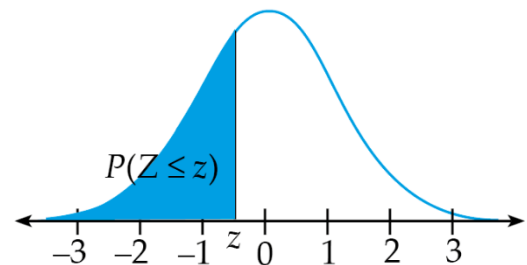
Figura 12.14. Regla empírica en la distribución normal $N(0,1)$.

Fuente: Elaboración propia (Excel, 2024).

A continuación, se presenta el valor de la tabla para z a cuatro decimales, que es el área bajo la curva de la distribución normal estándar a la izquierda de z .

Figura 12.15. Área bajo la curva de la normal estándar a la izquierda de z .

Fuente: Elaboración propia (GeoGebra, 2024).



Z	0.0000	0.0100	0.0200	0.0300	0.0400	0.0500	0.0600	0.0700	0.0800	0.0900
-3	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133

Z	0.0000	0.0100	0.0200	0.0300	0.0400	0.0500	0.0600	0.0700	0.0800	0.0900
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Tabla 12.1. Tabla de probabilidad $Z \sim N(0, 1)$. Representa el área acumulada bajo la curva de la función de densidad normal estándar.
Fuente: Elaboración propia (Excel, 2023).

Aplicaciones de la distribución normal

Ejemplo formativo 12.1

Los y las estudiantes de una escuela de nivel bachillerato, realizaron medidas sobre la estatura 170 estudiantes, dando como resultado la **estatura media** de $\mu = 164.1$ cm y la **desviación estándar** de $\sigma = 9.89$ cm. Si las estaturas se distribuyen normalmente, como podemos observar en la Figura 12.16, realiza lo siguiente:

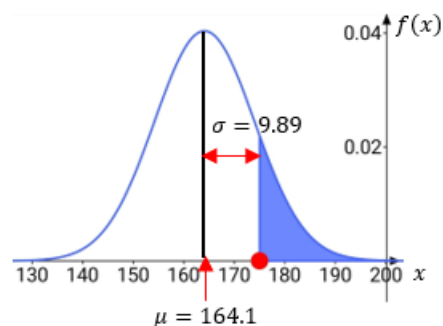


Figura 12.16. Distribución normal con media 164.1 cm y desviación estándar 9.89 cm.
Fuente: Elaboración propia (GeoGebra, 2024).

de

- Calcula la probabilidad de que un estudiante elegido al azar mida más de 175 cm.
- ¿Cuántos estudiantes esperas que midan más de 175 cm?

Resolución:

a) Para calcular la probabilidad tienes que encontrar el área sombreada bajo la curva usando la Tabla 12.1.

Por lo anterior, estandariza el valor de $x = 175$.

$$z = \frac{x - \mu}{\sigma} = \frac{175 - 164.1}{9.89} = \frac{10.9}{9.89} = 1.1$$

Consulta la Tabla 12.1 y determina el área bajo la curva para $z = 1.102$. Para este caso, la probabilidad de que un estudiante elegido al azar mida más de 175 cm es:

$$P(Z > 1.1) = 1 - P(Z \leq 1.1) = 1 - 0.8643 = 0.1357 \text{ o } 13.57\%$$

De acuerdo con la Tabla 12.1, la probabilidad de que un estudiante elegido al azar mida más de 175 centímetros es: **0.1357 o 13.57 %**.

Lo anterior, también se puede calcular mediante la aplicación *Probability Distributions*, usando la variable aleatoria $X \sim N(164.1, 9.89)$ en vez de $Z \sim N(0, 1)$.

$$P(X > 175) = 0.1352$$

De acuerdo con la aplicación *Probability* la probabilidad de que un estudiante elegido al azar mida más de 175 centímetros es: **0.1352 o 13.52 %**.

Figura 12.18. Probabilidad de que un estudiante elegido al azar mida más de 175 cm.

Fuente: *Probability Distributions* (Android, 2024).

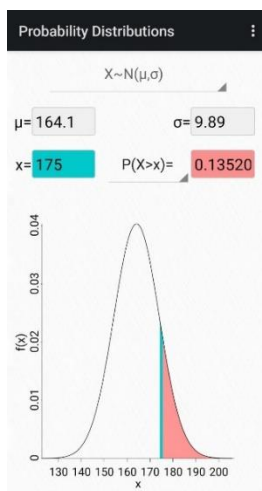


Figura 12.17. Altura de una estudiante.

Fuente: Fotografía. César Pilar Quintero (Android, 2024).

b) Para saber cuántos se espera que midan más de 175 cm, hay que calcular el 13.57% de 170 estudiantes.

$$(170)(0.1357) \approx 23$$

Se espera que 23 estudiantes de los 170 midan más de 175 cm.

Ejemplo formativo 12.2

Con referencia al Ejemplo formativo 12.1 obtén la probabilidad de que la estatura del estudiante elegido se encuentre:

1. Entre 160 cm y 170 cm; $P(160 < X < 170)$
2. Sea menor que 155 cm; $P(X < 155)$
3. Sea mayor que 180 cm; $P(X > 180)$

Resolución:

1. Entre 160 cm y 170 cm; $P(160 < X < 170) = 0.3854$ o 38.54%

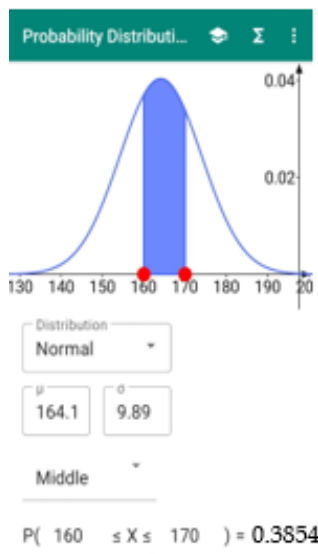


Figura 12.20. Cálculo de probabilidad de la forma $P(a \leq X \leq b)$.

Fuente: *Probability Distributions*, (Android, 2024).

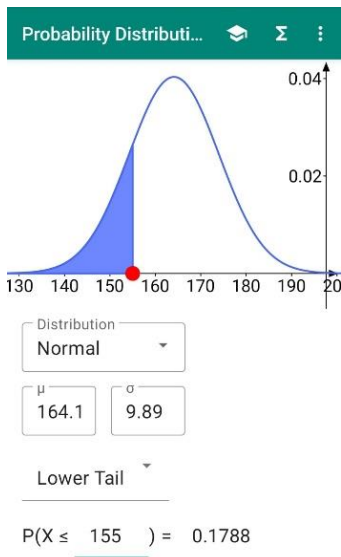


Figura. 12.19. Altura de un estudiante.

Fuente: Fotografía. César Pilar Quintero (Android, 2024).

La probabilidad de que la estatura del estudiante elegido se encuentre entre 160 cm y 170 cm es del 38.54%.

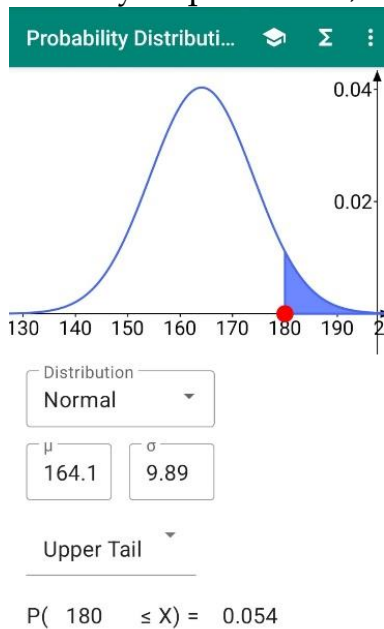
2. Sea menor que 155 cm; $P(X < 155) = 0.1788$ o 17.88%.



La probabilidad de que la estatura del estudiante elegido sea menor que 155 cm es del 17.88%.

Figura 12.21. Cálculo de probabilidad de la forma $P(X \leq b)$.
Fuente: *Probability Distributions* (Android, 2024).

3. Sea mayor que 180 cm; $P(X > 180) = 0.054$ o 5.4%



La probabilidad de que la estatura del estudiante elegido sea mayor que 180 cm es del 5.4%.

Figura 12.22. Cálculo de probabilidad de la forma $P(X \geq b)$.
Fuente: *Probability Distributions* (Android, 2024).

Ejemplo formativo 12.3

Responde la siguiente situación.

Un estudiante del bachillerato entra a la escuela a las 7:00 a.m. y hace un promedio de 20.5 minutos desde que sale de su casa hasta que llega a su escuela, con una desviación estándar de 2.9 minutos. Supón que la distribución de los tiempos de viaje es normal. Si siempre sale de su casa a las 6:40 a.m., ¿qué porcentaje de las veces llegará tarde?

Resolución:

X: Minutos que tarda en llegar a la escuela.

$$P(X > 20) = 0.5684 \text{ o } 56.84 \%$$

Por lo tanto, el estudiante llegará tarde a la escuela el 56.84% de las veces.

Ejemplo formativo 12.5

Realiza una encuesta (40 personas) entre familiares y/o vecinos. Suponiendo que los pesos se distribuyen normalmente obtener la probabilidad de que la persona elegida se encuentre:

1. Por arriba de la media.
2. Por abajo de la media.
3. Una desviación estándar por abajo de la media y una desviación estándar por arriba de la media.

Resolución:

Encuesta de pesos (40 personas):

68, 72, 61, 75, 63, 70, 58, 65, 73, 69, 62, 71, 67, 64, 60, 74, 66, 59, 76, 68,
65, 70, 62, 72, 61, 63, 75, 67, 71, 64, 60, 73, 69, 58, 66, 74, 65, 62, 70, 68

Calculando la media y desviación estándar:

$$\text{Media} = 2676/40 = 66.9 \text{ kg}$$

$$\text{Desviación estándar} = 5.148 \text{ kg}$$

Suponiendo que los pesos se distribuyen normalmente obtener la probabilidad de que la persona elegida se encuentre:

1. Probabilidad de estar por encima de la media (66.9 kg):

$$P(X > 66.9) = 0.5 = 50\%$$

2. Probabilidad de estar por debajo de la media (66.9 kg):

$$P(X < 66.9) = 0.5 = 50\%.$$

3. Probabilidad de estar entre 61.725 kg (1 desviación estándar por debajo de la media) y 72.048 kg (1 desviación estándar por encima de la media):



Figura. 12.23. Peso de una estudiante.
Fuente: Fotografía. César Pilar Quintero
(Android, 2024).

$$P(61.725 < X < 72.048) = 0.6827 = 68.4\%.$$

Evaluación formativa 12.1

1. Indaga si existe alguna diferencia en el resultado al calcular $P(Z \leq 2)$ y $P(Z < 2)$. Justifica tu respuesta.
2. En un hospital, el tiempo medio que tardan los pacientes en ser atendidos es de 40 minutos, con una desviación estándar de ocho minutos.
 - a) Calcula la probabilidad de que un paciente espere menos de media hora.
 - b) Calcula la probabilidad de que un paciente espere más de un cuarto de hora.
 - c) Calcula la probabilidad de que un paciente espere entre 30 y 60 minutos.
 - d) Explica las posibles implicaciones para la percepción del servicio del hospital.
 - e) Comprueba los resultados usando aplicaciones de celular e inteligencias artificiales.
 - f) Compara tus resultados con los de la clase y escribe tus conclusiones al respecto.
3. Considerando una distribución normal con media 100 y desviación estándar 15.
 - a) Explica por qué la probabilidad de obtener un valor mayor que 130 es menor que obtener un valor menor que 70.
 - b) Comprueba los resultados usando aplicaciones de celular e inteligencias artificiales.
 - c) Compara tus resultados con los de la clase y escribe tus conclusiones al respecto.
4. Para llevar a cabo un análisis estadístico del consumo diario de tortillas por familia, primero determina el número promedio de tortillas consumidas, así como la desviación estándar de dicha cantidad. Considera que la variable X , que representa el número de tortillas consumidas por una familia en un día, sigue una distribución normal. Basándote en esta distribución, calcula la probabilidad de que el consumo de tortillas por una familia en un día sea:
 - a) Al menos 30 tortillas.
 - b) Entre 25 y 43 tortillas.
 - c) Menos de 20 tortillas.

Autoevaluación y coevaluación 12.1

Nombre: _____ Plantel: _____ Grupo: _____
Turno: _____ Fecha: _____

Autoevaluación para el aprendizaje

Selecciona en la columna la opción que mejor refleje tu nivel de desempeño en el proceso para el aprendizaje de la progresión de aprendizaje 12. Responde con honestidad a la evaluación de cada uno de los criterios que se enlistan a continuación.

Desempeño	En proceso de logro	Bueno	Sobresaliente
Propicié un clima de comunicación favorable para el estudio de elementos y propiedades de la distribución normal.			
Participé activamente con ideas para estandarizar valores z.			
Contribuí colaborativamente en la retroalimentación de dudas sobre el cálculo de probabilidades.			
Verifiqué los cálculos de las probabilidades utilizando aplicaciones de celular o IA (M3-C1).			
Interpreté el valor de probabilidad con base al contexto de los datos (M4-C2).			
Apliqué el cálculo de probabilidades a eventos aleatorios con distribución normal (M3-C3).			

Coevaluación para el aprendizaje

Solicita a un compañero del equipo que marque en la columna la opción que mejor describa tu desempeño durante el trabajo colectivo, concluida la progresión de aprendizaje 12, y que responda con honestidad la evaluación de cada uno de los criterios que se enlistan a continuación.

Desempeño	En proceso de logro	Bueno	Sobresaliente
Propició un clima de comunicación favorable para el estudio de elementos y propiedades de la distribución normal.			
Participó activamente con ideas para estandarizar valores z.			
Contribuyó colaborativamente en la retroalimentación de dudas sobre el cálculo de probabilidades.			
Verificó los cálculos de las probabilidades utilizando aplicaciones de celular o IA (M3-C1).			
Interpretó el valor de probabilidad con base al contexto de los datos (M4-C2).			
Aplicó el cálculo de probabilidades a eventos aleatorios con distribución normal (M3-C3).			

Nombre y firma de quien coevalúa

Autoevaluación para el aprendizaje

Nombre: _____ Plantel: _____ Grupo: _____ Turno: _____

La presente autoevaluación debe servir para que valores tu labor durante el semestre. Selecciona en la columna la opción que mejor refleje tu nivel de desempeño en el proceso de aprendizaje de la Unidad de Aprendizaje Curricular Pensamiento Matemático I, en cumplimiento de las metas de aprendizaje. Responde con honestidad a la evaluación de cada uno de los criterios que se enlistan a continuación.

Mis desempeños	Puedo mejorar	Alcancé	Sobresaliente
Efectúo cálculo de probabilidades y de medidas de tendencia central (M1-C1).			
Valoro la selección adecuada del tipo de gráfico apropiado para representar datos y cuándo se utilizan las permutaciones o las combinaciones en el análisis de un evento (M2-C1).			
Empleo aplicaciones informáticas para el cálculo de permutaciones y combinaciones, así como para otros estadísticos de prueba de hipótesis (M3-C1).			
Reconozco que la organización de datos y su visualización gráfica contribuyen a tomar decisiones en relación con la incertidumbre y variabilidad, en la determinación entre los tipos de variables cualitativas y cuantitativas, así como para apreciar el efecto de datos atípicos (M1-C2).			
Identifico tendencias y comportamientos específicos de los datos correspondientes a determinados fenómenos o eventos a través de su representación en tablas y en gráficas estadísticas (M2-C2).			
Analizo la relación de independencia que puede haber entre dos variables categóricas, así como si existe correlación entre variables cualitativas (M3-C2).			
Utilizo diferentes propiedades y pruebas estadísticas para valorar relaciones entre variables o interpretar medidas de tendencia relacionadas con variadas situaciones o eventos. (M4-C2)			
Aprecio la diferencia entre eventos equiprobables y no equiprobables (M1-C3).			
Aplico las técnicas de conteo apropiadas, las medidas de tendencia y el cálculo de probabilidades para el análisis y solución de problemas relacionados con eventos diversos. (M3-C3)			
Realizo el cálculo de probabilidades usando la distribución normal (M4-C3).			
Explico la diferencia entre los casos posibles de ocurrencia de un suceso y el total de casos, así como las diferencias y similitudes entre el diseño de experimentos y un estudio observacional (M1-C4).			
Comparto mis propuestas sobre el diseño de cuestionarios para recolectar información de una situación de interés (M2-C4).			

Coevaluación para el aprendizaje

Solicita a un compañero de equipo que valore tu labor durante el semestre y que seleccione en la columna la opción que mejor refleje tu nivel de desempeño en el proceso de aprendizaje de la Unidad de Aprendizaje Curricular Pensamiento Matemático I, en cumplimiento de las metas de aprendizaje. Responde con honestidad a la evaluación de cada uno de los criterios que se enlistan a continuación.

Desempeños de mi compañero	Puede mejorar	Alcanzó	Sobresaliente
Efectúa cálculo de probabilidades y de medidas de tendencia central (M1-C1).			
Valora la selección adecuada del tipo de gráfico apropiado para representar datos y cuándo ha sido necesario utilizar las permutaciones o las combinaciones en el análisis de un evento (M2-C1).			
Emplea aplicaciones informáticas para el cálculo de permutaciones y combinaciones, así como para otros estadísticos de prueba de hipótesis (M3-C1).			
Reconoce que la organización de datos y su visualización gráfica contribuyen a tomar decisiones en relación con la incertidumbre y variabilidad, en la determinación entre los tipos de variables cualitativas y cuantitativas, así como para apreciar el efecto de datos atípicos (M1-C2).			
Identifica tendencias y comportamientos específicos de los datos correspondientes a determinados fenómenos o eventos a través de su representación en tablas y en gráficas estadísticas (M2-C2).			
Analiza la relación de independencia que puede haber entre dos variables categóricas, así como si existe correlación entre variables cualitativas (M3-C2).			
Utiliza diferentes propiedades y pruebas estadísticas para valorar relaciones entre variables o interpretar medidas de tendencia relacionadas con variadas situaciones o eventos. (M4-C2)			
Aprecia la diferencia entre eventos equiprobables y no equiprobables (M1-C3).			
Aplica las técnicas de conteo apropiadas, las medidas de tendencia y el cálculo de probabilidades para el análisis y solución de problemas relacionados con eventos diversos. (M3-C3)			
Realiza el cálculo de probabilidades usando la distribución normal (M4-C3).			
Explica la diferencia entre los casos posibles de ocurrencia de un suceso y el total de casos, así como las diferencias y similitudes entre el diseño de experimentos y un estudio observacional (M1-C4).			
Comparte sus propuestas sobre el diseño de cuestionarios para recolectar información de una situación de interés (M2-C4).			

Nombre y firma de quien coevalúa.

Bibliografía consultada

- Juárez, J. D. et al. (2023). Estadística. Exploración de datos. Once Ríos.
- Juárez, J. D. et al. (2023). Probabilidad. Once Ríos.
- SEP (2023a). Acuerdo número 09/08/23 por el que se establece y regula el Marco Curricular Común de la Educación Media Superior. Diario Oficial de la Federación.
- SEP (2023b). Orientaciones Pedagógicas del recurso sociocognitivo pensamiento matemático. SEMS.
- SEP (2023c). Progresiones de aprendizaje del recurso sociocognitivo pensamiento matemático. SEMS.
- SEP (2023d). Programa de estudios del Recurso Sociocognitivo Pensamiento Matemático I. Secretaría de Educación Pública.
- Triola, M. F. (2004). Estadística (9a ed.). Pearson Educación.

Referencia a las fuentes de consulta de tablas

Índice de tablas

Tabla 1.1. Casos confirmados de dengue correspondiente a la semana epidemiológica No. 39. de los años 2022 y 2023.

Fuente: Secretaría de Salud Sinaloa, 2017. <http://saludsinaloa.gob.mx/wp-content/uploads/2017/epidemiologia/tercer-trimestre-2023/Boletin%20Semanal%20Convencional%20Sinaloa%202023-SEM%2039.pdf>

Tabla 1.2. Promedio ofensivo durante cuatro temporadas de Yasmany Tomás.

Fuente:

https://www.espn.com.mx/beisbol/mlb/jugador/estadisticas/_/id/33662/yasmany-tomas

Tabla 4.1. Gasto corriente monetario promedio trimestral del año 2022 por hogar en Sinaloa, según grandes rubros del gasto en pesos.

Fuente: Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) 2022. Nueva Serie. Disponible en: <https://www.inegi.org>.

Tabla 4.2. Gasto corriente monetario promedio trimestral en los veinte principales rubros de gasto del año 2022 por hogar en Sinaloa, según grandes rubros del gasto en pesos.

Fuente: Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) 2022. Nueva Serie. Disponible en: <https://www.inegi.org.mx/programas/enigh/nc/2022/>

Tabla 5.1. Frecuencia de juegos de beisbol jugados en la LMP.

Fuente: https://es.wikipedia.org/wiki/Liga_Mexicana_del_Pac%C3%ADfico_2022-23

Tabla 5.2. Cuentas de Instagram con más seguidores.

Fuente: Elaboración propia, con base en: <https://www>.

ayudacuentas.com/instagram/cuentas-con-mas-seguidores/
y <https://www.merca20.com/los-mas-seguidos-instagram-2022/> (Word, 2023).

Tabla 5.3. Carreras con más demandada en la unam.

Fuente: Elaboración propia con base en: <https://blog.unitips.mx/carreras-mas-demandadas-de-la-unam> y <https://oferta.unam.mx> (Word, 2023).

Tabla 5.4. Distribución de frecuencias agrupadas de goles a favor en la Liga MX temporada 2022-2023.

Fuente: Elaboración propia (Word, 2023).

Tabla 5.5. Distribución de frecuencias agrupadas con marcas de clase.

Fuente: Elaboración propia (Word, 2023).

Tabla 6.1. Preferencia del consumo de refrescos por género.

Fuente: Elaboración propia (Word, 2023).

Tabla 6.2. Frecuencia esperada del consumo de refrescos por género.

Fuente: Elaboración propia (Word, 2023).

Tabla 6.3. Tabla de distribución de la Chi-cuadrado.

Fuente: Disfruta las Matemáticas,
<https://www.disfrutalasmaticas.com/datos/tabla-chi-cuadrado.html>

Tabla 7.1. Datos de la calificación obtenida en Matemáticas y Mecánica.

Fuente: Elaboración propia (Word, 2023).

Tabla 8.1. Relación entre el tipo de dieta y el nivel de colesterol.

Fuente: Elaboración propia (Word, 2024).

Tabla 8.2. Relación entre el tipo de dieta y el nivel de colesterol considerando el factor edad.

Fuente: Elaboración propia (Word, 2024).

Tabla 8.3. Ingresos diarios generados por las ventas mediante plataforma durante el último mes.

Fuente: Elaboración propia (Word, 2024).

Tabla 11.1 Fórmulas para calcular las medidas de tendencia central.

Fuente: Elaboración propia (Word, 2024).

Tabla 11.2 Fórmulas para calcular las medidas de posición y de dispersión.

Fuente: Elaboración propia (Word, 2024).

Tabla 11.3 Fórmulas para calcular las medidas de dispersión.

Fuente: Elaboración propia (Word, 2024).

Tabla 12.1. Tabla de probabilidad $Z \sim N(0, 1)$. Representa el área acumulada bajo la curva de la función de densidad normal estándar.

Fuente: Elaboración propia (Excel, 2023).

Referencia a las fuentes de consulta de figuras

Índice de figuras

Figura 3.1. Diagrama de Venn.

Fuente: Elaboración propia (Word, 2024).

Figura 3.2. Diagrama de Venn.

Fuente: Elaboración propia (Word, 2024).

Figura 3.3. Diagrama de Venn.

Fuente: Elaboración propia (Word, 2024).

Figura 3.4. Diagrama de Venn.

Fuente: Elaboración propia (Word, 2024).

Figura 3.5. Diagrama de Venn.

Fuente: Elaboración propia (Word, 2024).

Figura 3.6. Lanzamiento de una moneda.

Fuente: Fotografía. Rolando Alberto Forneiro Rodríguez, 2024.

Figura 3.7. Diagrama de Venn.

Fuente: Elaboración propia (Word, 2024).

Figura 3.8. Lanzamiento de dos dados.

Fuente: www.freepik.com, 2024.

Figura 5.1. Resultados de juegos ganados y juegos perdidos de la LMP 2022-2023.

Fuente: Elaboración propia, con base en:

https://es.wikipedia.org/wiki/Liga_Mexicana_del_Pacífico_2022-23 (Excel, 2024).

Figura 5.2. Rendimiento de la temporada 2022-2023 en la LMP.

Fuente: Elaboración propia, con base en:

https://es.wikipedia.org/wiki/Liga_Mexicana_del_Pacífico_2022-23 (Excel, 2024).

Figura 5.3. Tipos de gráficos.

Fuente: Elaboración propia (Word, 2024).

Figura 5.4. Gráfica de barra simple.

Fuente: Elaboración propia (Excel, 2024).

Figura 5.5. Gráfica de barra doble.

Fuente: Elaboración propia (Excel, 2024).

Figura 5.6. Cuentas de Instagram con más seguidores en 2023.

Fuente: Elaboración propia, con base en: <https://www.ayudacuentas.com/instagram/cuentas-con-mas-seguidores> (Excel, 2024).

Figura 5.7. Cuentas de Instagram con más seguidores en 2022 y 2023.

Fuentes: Elaboración propia, con base en: <https://www.ayudacuentas.com/instagram/cuentas-con-mas-seguidores/> y <https://www.merca20.com/los-mas-seguidos-instagram-2022/>, (Excel, 2024).

Figura 5.8. Carreras más demandadas UNAM 2023-2024.

Fuente: Elaboración propia, con base en: <https://blog.unitips.mx/carreras-mas-demandadas-de-la-unam> (Excel, 2024).

Figura 5.9. Goles a favor en la Liga MX temporada 2022-2023.

Fuente: Elaboración propia, con base en: <https://www.mediotiempo.com/futbol/liga-mx/tabla-general/385-2022> (Excel, 2024).

Figura 5.10. Histograma de goles a favor en la Liga MX temporada 2022-2023.

Fuente: Elaboración propia (Excel, 2024).

Figura 5.11. Proceso de construcción de un polígono de frecuencias a partir del histograma.

Fuente: Elaboración propia (Excel, 2024).

Figura 5.12a. Puntaje obtenido

Fuente: Elaboración propia, (Excel, 2024).

Figura 5.12b. Puntaje obtenido.

Fuente: Elaboración propia, (Excel, 2024).

Figura 5.12c. Puntaje obtenido.

Fuente: Elaboración propia, (Excel, 2024).

Figura 5.13. Ventas.

Fuente: Elaboración propia (Excel, 2024).

Figura 5.14. Temperatura en °C.

Fuente: Elaboración propia (Excel, 2024).

Figura 5.15. Pares de zapatos.

Fuente: Elaboración propia (Excel, 2024).

Figura 5.16. Televisores por hogar.

Fuente: Elaboración propia (Excel, 2024).

Figura 5.17a. Tiempo destinado al día para ver videos de TikTok por estudiantes del grupo A.

Fuente: Elaboración propia (Excel, 2024).

Figura 5.17b. Tiempo destinado al día para ver videos de TikTok por estudiantes del grupo C.

Fuente: Elaboración propia (Excel, 2024).

Figura 5.18. Peso de 20 estudiantes.

Fuente: Elaboración propia (Excel, 2024).

Figura 5.19. Tiempo invertido en plataformas y dispositivos digitales.

Fuente: Elaboración propia (Excel, 2024).

Figura 5.20. Resultados de JG y JP de la LMO 2022-2023.

Fuente: Elaboración propia (Excel, 2024).

Figura 5.21. Impacto de Spotify en la industria de la música.

Fuente: Elaboración propia (Excel, 2024).

Figura 6.1. Tabla de contingencia en la calculadora digital de *statologos*, para realizar la prueba Chi-cuadrado.

Fuente: <https://statologos.com/calculadora-prueba-de-independencia-chi-cuadrado/>

Figura 7.1. Gráfica de la ecuación $y = mx + b$.

Fuente: Elaboración propia (Desmos, 2024).

Figura 7.2. Gráfica de la ecuación $y = mx + b$.

Fuente: Elaboración propia (Desmos, 2024).

Figura 7.3. Gráfico de dispersión entre la calificación obtenida en Matemáticas y Mecánica.

Fuente: Elaboración propia (Excel, 2024).

Figura 7.4 Correlación directa o positiva.

Fuente: Elaboración propia (Excel, 2024).

Figura 7.5 Correlación inversa o negativa.

Fuente: Elaboración propia (Excel, 2024).

Figura 7.7 Fuerza de correlación fuerte.

Fuente: *Fuente:* Elaboración propia (Excel, 2024).

Figura 7.8 Fuerza de correlación débil.

Fuente: Elaboración propia (Excel, 2024).

Figura 7.9 Fuerza de correlación nula.

Fuente: Elaboración propia (Excel, 2024).

Figura 7.10. Datos de horas de estudio y de calificación.

Fuente: Google Docs, 2024, Disponible en: <https://docs.google.com/spreadsheets/>.

Figura 7.11. Cálculo del coeficiente de correlación.

Fuente: Google Docs, 2024, Disponible en: <https://docs.google.com/spreadsheets/>.

Figura 7.12. Función para calcular el coeficiente de correlación.
Fuente: Google Docs, 2024, Disponible en: <https://docs.google.com/spreadsheets/>.

Figura 7.13. Gráfico de dispersión.
Fuente: Elaboración propia (Excel, 2024).

Figura 7.14. Gráfico de dispersión.
Fuente: Elaboración propia (Excel, 2024).

Figura 7.15. Gráfico de dispersión.
Fuente: Elaboración propia (Excel, 2024).

Figura 7.16. Gráfico de dispersión.
Fuente: Elaboración propia (Excel, 2024).

Figura 8.1. Gráfico de puntos con un valor atípico.
Fuente: Elaboración propia (Word, 2024).

Figura 8.2. Gráfico de cajas con un valor atípico.
Fuente: Elaboración propia (Excel, 2024).

Figura 10.3. Calificación en matemáticas por género.
Fuente: Elaboración propia (Excel, 2024).

Figura 9.1. Población y muestra.
Fuente: Fotografía, Asia Cecilia Carrasco Valenzuela e Iliana Tirado Olivas (Android, 2024).

Figura 9.2. Tipos de muestreo.
Fuente: Elaboración propia (Word, 2024).

Figura 9.3. Aplicación para calcular tamaño de una muestra.
Fuente: EST calcula tu muestra (Android, 2024).

Figura 9.4. Muestreo aleatorio simple.
Fuente: Fotografía. Asia Cecilia Carrasco Valenzuela e Iliana Tirado Olivas (Android, 2024)..

Figura 9.5. Muestreo estratificado.
Fuente: Fotografía. Asia Cecilia Carrasco Valenzuela e Iliana Tirado Olivas (Android, 2024)..

Figura 9.6. Muestreo sistemático.
Fuente: Fotografía. Asia Cecilia Carrasco Valenzuela e Iliana Tirado Olivas (Android, 2024)..

Figura 9.7. Muestreo por conglomerados.
Fuente: Fotografía. Asia Cecilia Carrasco Valenzuela e Iliana Tirado Olivas (Android, 2024)..

Figura 11.1 Medidas de tendencia central.

Fuente: Elaboración propia (Word, 2024).

Figura 11.2. Gráfico de puntos.

Fuente: Elaboración propia (Word, 2024).

Figura 11.3 Costo semanal de la asistencia de los hijos a la preparatoria.

Fuente: Elaboración propia (Word, 2024).

Figura 12.1. Distribución normal de la estatura de estudiantes.

Fuente: Fotografía. César Pilar Quintero (Android, 2024).

Figura 12.2. Gráfica de la distribución normal estándar.

Fuente: *Probability Distributions* (Android, 2024).

Figura 12.3. Gráfica de la distribución normal estándar.

Fuente: *Probability Distributions* (Android, 2024).

Figura 12.4. Gráfica de la distribución normal estándar.

Fuente: *Probability Distributions* (Android, 2024).

Figura 12.5. Aplicación GeoGebra.

Fuente: Captura de pantalla celular Samsung (Geogebra, 2024).

Figura 12.6. Aplicación *Probability Distributions*.

Fuente: Captura de pantalla celular Samsung (*Probability Distributions*, 2024).

Figura 12.7. Aplicación *Probability Distributions*.

Fuente: Captura de pantalla celular Samsung (*Probability Distributions*, 2024).

Figura 12.8. Aplicación Distribución normal.

Fuente: Captura de pantalla celular Samsung (*Probability Distributions*, 2024).

Figura 12.9. Curva de la distribución normal $N(\mu, \sigma)$.

Fuente: Elaboración propia (Excel, 2024).

Figura 12.10. Curva de la distribución normal $N(0, 1)$.

Fuente: Elaboración propia (Excel, 2024).

Figura 12.11. Curvas de la distribución normal con medias iguales y desviación estándar diferentes.

Fuente: Elaboración propia (Excel, 2024).

Figura 12.12. Curvas de la distribución normal con medias diferentes y desviación estándar iguales.

Fuente: Elaboración propia (Excel, 2024).

Figura 12.13. El área bajo la curva de la distribución normal estándar $N(0, 1)$ y su relación con la probabilidad.

Fuente: Elaboración propia (GeoGebra, 2024).

Figura 12.14. Regla empírica en la distribución normal $N(0, 1)$.

Fuente: Elaboración propia (Excel, 2024).

Figura 12.15. Área bajo la curva de la normal estándar a la izquierda de z .

Fuente: Elaboración propia (GeoGebra, 2024).

Figura 12.16. Distribución normal con media 164.1 cm y desviación estándar 9.89 cm.

Fuente: Elaboración propia (GeoGebra, 2024).

Figura. 12.17. Altura de una estudiante.

Fuente: Fotografía. César Pilar Quintero (Android, 2024).

Figura 12.18. Probabilidad de que un estudiante elegido al azar mida más de 175 cm.

Fuente: *Probability Distributions* (Android, 2024).

Figura. 12.19. Altura de un estudiante.

Fuente: Fotografía. César Pilar Quintero (Android, 2024).

Figura 12.20. Cálculo de probabilidad de la forma $P(a \leq X \leq b)$.

Fuente: *Probability Distributions* (Android, 2024).

Figura 12.21. Cálculo de probabilidad de la forma $P(X \leq b)$.

Fuente: *Probability Distributions* (Android, 2024).

Figura 12.22. Cálculo de probabilidad de la forma $P(X \geq b)$.

Fuente: *Probability Distributions* (Android, 2024).

Figura. 12.23. Peso de una estudiante.

Fuente: Fotografía. César Pilar Quintero (Android, 2024).

Referencia a las fuentes de consulta de los códigos QR

Índice de códigos QR

QR 3.1. Probabilidad condicional: introducción. Disponible en:

<https://youtu.be/pko0dqQidnI>

Fuente: Parzibyte 2024. Disponible en: <https://parzibyte.me/apps/generador-qr/>

Código QR 5.1. Cómo crear gráficos en Canva. Disponible en:

<https://youtu.be/qfImncZdoOU>

Fuente: Parzibyte, 2024. Disponible en: <https://parzibyte.me/apps/generador-qr/>

Código QR 5.2. Cómo crear un gráfico de puntos en GeoGebra. Disponible en:

<https://youtu.be/g4xA5skm1ZA>

Fuente: Parzibyte, 2024. Disponible en: <https://parzibyte.me/apps/generador-qr/>

QR 6.1. Calculadora de prueba de independencia de Chi-cuadrado. Disponible en:

<https://statologos.com/calculadora-prueba-de-independencia-chi-cuadrado/>

Fuente: Parzibyte, 2024. Disponible en: <https://parzibyte.me/apps/generador-qr/>

Código QR 6.2. Prueba Chi-cuadrado. Disponible en:

<https://www.disfrutalasmaticas.com/datos/prueba-chi-cuadrado.html>

Fuente: Parzibyte, 2024. Disponible en: <https://parzibyte.me/apps/generador-qr/>

QR 7.1. Video sobre la correlación entre dos variables. Disponible en:

<https://youtu.be/qhT0Db4pmqk>

Fuente: Parzibyte, 2024. Disponible en: <https://parzibyte.me/apps/generador-qr/>

QR 7.2. Hoja de cálculo de Google Docs.

Fuente: Parzibyte, 2024. Disponible en: Disponible en:

<https://parzibyte.me/apps/generador-qr/>

Bibliografía de consulta para el estudiante y el profesor

Escalante, L. (2023). *Pensamiento matemático I*. Book Mart.

Colavita, E, (2023). *Pensamiento matemático I*. Macmillan Education.

Resumen del libro "Pensamiento Matemático I: Bachillerato Semiescolarizado de la UAS"

Este libro, desarrollado por Faustino Vizcarra Parra, Rolando Alberto Forneiro Rodríguez, Victoria Bárbara Arencibia Sosa y Armando Flórez Arco, es un material didáctico esencial para los estudiantes de bachillerato semiescolarizado. Diseñado para la Unidad de Aprendizaje Curricular (UAC) Pensamiento Matemático I, se alinea con el plan de estudios de la Universidad Autónoma de Sinaloa (UAS) 2024, conforme al Marco Curricular Común de la Educación Media Superior (MCCEMS).

El texto se estructura en doce progresiones de aprendizaje que cubren desde los fundamentos de la probabilidad y técnicas de conteo hasta la recolección de datos estadísticos y la interpretación de la distribución normal. Cada progresión incluye evaluaciones diagnósticas, situaciones contextualizadas, ejemplos formativos y evaluaciones finales, fomentando un aprendizaje integral y reflexivo.

Los autores dedican este libro a los docentes que inspiran curiosidad y amor por las matemáticas, y reconocen el papel de la inteligencia artificial como un aliado en el proceso educativo. Con una combinación de teoría, práctica y herramientas tecnológicas, este libro busca desarrollar el pensamiento probabilístico y estadístico, preparando a los estudiantes para enfrentar desafíos con ingenio y perseverancia.